St. Petersburg State University

6th ST. PETERSBURG WORKSHOP ON SIMULATION

Proceedings

Volume II

St. Petersburg, June 28 – July 4

Edited by S.M.Ermakov, V.B. Melas & A.N.Pepelyshev

St. Petersburg, 2009

VVM com. Ltd.

PROCEEDINGS OF THE 6th ST. PETERSBURG WORKSHOP ON SIMULATION. St. Petersburg, June 28 – July 4, 2009. Volume II. Ed. by S.M.Ermakov, V.B. Melas and A.N.Pepelyshev — St. Petersburg. VVM com. Ltd., 2009. — 568 pp.

ISBN

The second volume contains ninety three papers. These papers are extended abstracts of reports to be presented in the last two days of the 6th St.Petersburg Workshop.

© Authors 2009

ISBN

Contents

Optimal design for discrimination 581

Tommasi C. *Robust optimum designs to a misspecified model* **583**

Bogacka B., Atkinson A.C., Patan M. *Designs for Discriminating Between Models by Testing the Equality of Parameters* **589**

Amo-Salas M., López-Fidalgo J.,López-Ríos V. *Optimal designs for discriminating between pharmacokinetic models with correlated observations* **595**

Statistical inference from complex data 603

Brunner E. *Repeated Measures under Non-Sphericity* **605**

Darkhovsky B. *Non-asymptotic estimation of functionals in noise* **611**

Ryabko B. *Using Universal Source Coding for Statistical Analysis of Time Series* **617**

Ghosh S., Dey D. *Linear and Nonlinear Approximations of the Ratio of the Standard Normal Density and Distribution Functions for the Estimation of the Skew Normal Shape Parameter* **623**

Causeur D., Friguet C. *Conditional FDR estimation based on a factor analytic approach of multiple testing* **628**

Garel B., Kodia B. *Performance of a new dependence measure between stable non-Gaussian random variables and comparison with other measures: a simulation study* **634**

Guo L., Guan R. *Study on Linear Regression Analysis of Multivariate Compositional Data* **640**

Lifetime data analysis **647**

Huber C. *Simulations for frailty models with general censoring and truncation* **649**

Lee M-L.T., Whitmore G.A. *Proportional Hazards and Threshold Regression: Their Theoretical and Practical Connections* **650**

Caroni C. *Trend tests for recurrent events with incomplete and misclassified information* **651**

Hu J. *Analysis of Event History Data With Covariate Missing Not At Random* **656**

Nikulin M., Saaidia N. *Inverse Gaussian family and its applications in reliability. Study by simlulation* **657**

Recent advances in change-point analysis **665**

Gut A., Steinebach J.G. *Sequential change-point analysis for renewal counting processes* **667**

Hušková M., Marušiaková M. *M-procedures for detection of changes for dependent observations* **673**

Steland A. *A Note on Data-Adaptive Bandwidth Selection for Sequential Kernel Smoothers* **679**

Reliability and survival analysis **685**

Simino J., Hollander M., McGee D. *Calibration of proportional hazards models* **687**

Young L.J., Gotway C.A., Kearney G., DuClos C. *Models for Assessing Uncertainty for Local Regression* **693**

Olkin I. *Inequalities: Some Probabilistic, Some Matrix, Some Moment, and Some Classical* **699**

Stichastic simulation of rare events and stiff systems **701**

Cancela H., El Khadiri M., Rubino G., Tuffin B. *A Recursive Variance Reduction Technique with Bounded relative Error for Communication Network Reliability Estimation* **703**

Krull C., Horton G. *Proxel-Based Simulation: Theory and Applications* **709**

Sandmann W. *Multistep Methods for Markovian Event Systems* **715**

Lagnoux A. *An adaptive branching splitting algorithm under cost constraint for rare event analysis* **721**

Broniatowsky M. *Rare events simulation and conditioned random walks: the moderate deviation case* **726**

Goodness-of-fit and related methods 729

Huskova M., Meintanis S.G. *Tests for symmetric error distribution in regression models* **731**

Henze N., Nikitin Y., Ebner B. *L_p-type goodness-of-fit tests and their asymptotic comparison* **737**

Meintanis S.G., Taufer E. *Inference Procedures for Stable–Paretian Stochastic Volatility Models* **743**

Allison J.S., Swanepoel J.W.H. *A new method of evaluating the performance of bootstrap-based tests* **749**

Střelec L., Stehlík M. *On Robust Testing for Normality* **755**

Volkova K. *Tests of the exponentiality based on properties of order statistics* **761**

Martynov G. *Goodness-of-fit tests for the Weibull and Pareto distributions* **765**

Performance analysis of biological and queuing models 771

Lopez-Herrero M.J. *Analysis of the time to infection and disease extinction time in SIS epidemic models* **773**

Artalejo J. *Number of removals in the stochastic SIR model* **779**

Delgado R. *A sufficient condition for stability of fluid limit models* **785**

Ridder A. *Importance sampling simulation of the fork-join queue* **791**

Atencia I., Pechinkin A.V. *A discrete-time queueing system with total renewal discipline* **797**

Nobel R. *A discrete-time retrial queueing model with multiple servers* **803**

Economou A. *The maximum number of infectives in SIS epidemic models: Computational techniques and quasi-stationary distributions* **804**

Section reports 811

Queuing and other discrete systems 813

Klimenok V., Khramova V., Dudin A., Eom H., Kim C. *The $BMAP/PH/N$ queue operating in random environment* **815**

Klimenok V.I, Kim C.S., Taramin O.S. *Steady-state analysis of a dual tandem queue with arbitrary inter-arrival time distribution* **821**

De Clercq S., Steyaert B., Bruneel H. *Analysis of a Multi-Class Discrete-time Queueing System under the Slot-Bound Priority rule* **827**

Eom H.E., Kim C.S., Melikov A., Fattakhova N. *Approximate Method for QoS Analysis of Multi-Threshold Queuing Model of Multi-Service Wireless Networks* **833**

Boukouvalas A., Cornford C., Singer A. *Managing Uncertainty in Complex Stochastic Models: Design and Emulation of a Rabies Model* **839**

Koskinen J. *Using latent variables to account for heterogeneity in exponential family random graph models* **845**

Lukyanenko A., Gurtov A., Morozov E. *An adaptive backoff protocol with Markovian contention window control* **851**

Mosyagina E.N. *On estimations of periodically non-stationary stochastic automaton behavior under fuzzy condition* **857**

Khokhulina V.A., Tchirkov M.K. *On Decomposition of Fuzzy Automata Models* **863**

Shevchenko A.S., Tchirkov M.K. *On simulation of stochastic languages by nondeterministic automata* **869**

Krivulin N. *Evaluation of the Lyapunov exponent for generalized linear second-order exponential systems* **875**

Monte-Carlo methods and stochastic modeling 883

L'Ecuyer P., Tuffin B. *Limiting Distributions for Randomly-Shifted Lattice Rules* **885**

Sushkevich T., Strelkov S., Maksakova S. *Kinetic approach and method of influence function to modelling of polarized radiation transfer* **891**

Walker J.H., Iskedjian M. *A Semi-Markov Model for Patients Following a Clinically Isolated Syndrome Event Prior to Progression to Clinically Definite Multiple Sclerosis* **897**

Paramonov Y., Andersons J., Kleinhofs M. *Modelling of tensile strength of fiber and composite using MinMaxDM distribution family* **903**

Miretskiy D., Scheinhardt W., Mandjes M. *An efficient Multilevel Splitting scheme* **909**

Alfares H.K. *Simulation-based multi-craft workforce scheduling* **915**

Ermakov S., Timofeev K. *On two new Monte-Carlo methods for solving nonlinear equations* **922**

Ermakov S.M., Rukavishnikova A.I. *Application of the Monte-Carlo and Quasi Monte-Carlo methods to solving systems of linear equations* **929**

Harlamov B.P. *Stochastic Modeling in Chromatography* **935**

Nekrutkin V., Rumyantzev N. *Artificial Monte Carlo interactions for linear problems* **941**

Kolnogorov A.V., Shelonina T.N. *Monte-Carlo Simulations for the Multi-Armed Bandit Problem* **947**

Alexeyeva N. *The generalized geometric distribution and the associated Galton-Watson model with a medical statistical application* **953**

Heussen N., Hilgers R.-D., Ackermann D. *Choice of the reference set in randomization tests in the presence of missing values - a simulation study* **960**

Burnaeva E. *On Discrepancy in Connection with the Residual of the Cubature Formulae* **969**

Probabilistic models **973**

Rozovsky L. *A remark to large deviation probabilities for sums of i.i.d. random variables in the domain of attraction of a stable law* **975**

Petrov V.V., Korchevsky V.M. *On the strong law of large numbers for sequences of dependent random variables* **977**

Nevzorov V., Saghatelyan V. *On one new model of records* **981**

Gouet R., López F.J., Sanz, G. *Limit theorems for the counting process of near-records* **985**

Orsingher E., Cammarota V. *Random motions in hyperbolic spaces* **991**

Applied stochastic procedures **993**

Grané A., Veiga H. *Wavelet-based detection of outliers in volatility models* **995**

Dyudenko I., Morozov E., Pagano M., Sandmann W. *Comparative Study of Effective Bandwidth Estimators: Batch Means and Regenerative Cycles* **1003**

Jørgensen B., Petersen H.C. *Efficient Estimation for Incomplete Multivariate Data* **1010**

Andronov A. *Maximal likelihood estimates for modified gravitation model by aggregated data* **1016**

Ermakov S.M., Nechaeva M.L. *On estimation of the special form nonlinear regression parameters* **1022**

Korobeynikov A. *On the Consistency of ML-estimates for the Special Model of Survival Curves with Incomplete Data* **1027**

Bochkina N., Lewin A. *Fuzzy Fisher test for contingency tables with application to genomics data* **1033**

Ermakov M. *Bootstrap from moderate deviation viewpoint* **1034**

Experimental design **1039**

Malyutov M., Sadaka H. *On Capacity of Screening Experiments under Linear Programming Analysis* **1041**

Roth K. *Adaptive Designs for Dose Escalation Studies* **1048**

Atkinson A.C. *Adaptive Covariate Adjusted Designs for Clinical Trials that Seek Balance* **1054**

Melas V.B. *On the functional approach to studying optimal Bayesian design* **1060**

Mielke T. *Sparse Sampling D-Optimal Designs in Quadratic Regression With Random Effects* **1066**

Liang Y., Carriere KC. *Multiple-Objective response-adaptive repeated measurement designs for clinical trials with dichotomous outcomes* **1072**

Sedunov E.V., Sedunova A.N. *Unbiased procedures in time-dependent regression experiments* **1079**

Dette H., Melas V.B., Shpilev P. *Optimal designs for estimating the linear combination of the coefficients in trigonometric regression models* **1085**

Pepelyshev A. *Improvement of random LHD for high dimensions* **1091**

Harman R., Štulajter F. *Optimality of Equidistant Sampling Designs for a Nonstationary Ornstein-Uhlenbeck Process* **1097**

Schiffl K., Hilgers R.D. *A-Optimal Designs for Two-Color Microarray Experiments for interesting contrasts* **1103**

Optimization techniques **1109**

Tikhomirov A. *On some properties of the simulated annealing algorithm* **1111**

Kreimer J., Ianovsky E. *Optimization of Real-Time Systems* **1117**

Laurini F., Grossi L., Gozzi G. *Robust Lagrange multiplier test with forward search simulation envelopes* **1124**

Granichin O., Gurevich L., Vakhitov A. *Parameter Estimation and Tracking For the Rosenbrock Function Using Simultaneous Perturbation Stochastic Approximation Algorithm* **1130**

Friguet C., Causeur D.*Estimation of the proportion of true null hypotheses among dependent tests* **1137**

Index **1142**

Session

Optimal design for discrimination

organized by Jesus López Fidalgo (Spain)

6th St.Petersburg Workshop on Simulation (2009) 583-587

Robust optimum designs to a misspecified model

Chiara Tommasi[1]

Abstract

Usually, in the Theory of Optimal Experimental Design the model is assumed to be known at the design stage. In practice, however, more competing models may be plausible for the same data. In this paper, instead of finding a design which is "good" for both model discrimination and parameter estimation we find an optimum design which is "good" for estimating the unknown parameters whether or not the assumed model is correct.

1. Introduction

Optimum designs are derived under the assumption that the statistical model is known at the design stage. In practice, however, more rival models may provide similar data fit. The optimum design problem of estimating some aspects of the model has been considered by many authors while less attention has been paid to the problem of model discrimination. Some references about this subject are [9], [4], [5], [11] and [11], among others. Only few authors deal with the dual problem of model discrimination and parameter estimation, see for instance [5] , [19], [16], [15], [1], [13] and the references therein.

In this paper instead of finding a design useful for both model discrimination and parameter estimation we propose another approach. We compute a design which is optimum for parameter estimation but takes into account a possible specification error in the model. In other words, we find an optimum design which is "robust" to a misspecified model. Robust designs to a variety of specification model errors have been studied by [7], [1], among others. Usually, the experimental conditions are chosen in order to maximize some criterion function of the inverse of the asymptotic covariance matrix of the maximum likelihood estimator (MLE). When the selected model is correct then the asymptotic covariance matrix of the MLE is the inverse of Fisher's information matrix. In order to have a precise estimate of the parameters, for instance the D-criterion should be applied to the Fisher information matrix. However, if there is a specification error in the model, i.e. the chosen model is wrong, then the likelihood function is not correct. Under some regularity conditions, the MLE is still consistent but with asymptotic covariance matrix given by the information sandwich variance matrix (see for instance [8] or [18]). In this case the D-criterion should be applied to the inverse of the information sandwich variance matrix. Since the chosen model may be correct

[1]University of Milano, E-mail: chiara.tommasi@unimi.it

or not, we propose a compound criterion given by a weighted geometric mean of D-efficiencies based on Fisher's information matrix and on the inverse of the information sandwich variance matrix, respectively. Maximizing this compound criterion we get an optimum design which is "good" whether or not the selected model is correct. In this sense this optimum design is robust to a misspecified model and it is called DR-optimum design. We suggest to use a DR-optimum design whenever a sandwich variance estimator (SVE) is used for estimating the asymptotic covariance matrix of the MLE, since a SVE estimates the information sandwich variance matrix which simplifies to the inverse of Fisher's information matrix in absence of misspecification, see for instance [17] and [9].

The formal definition of information sandwich variance matrix is given in Section 2. In Section 3 we describe the new optimality criterion for getting a DR-optimum design. Finally, in Section 4 an explanatory example is widely developed.

2. Information sandwhich variance matrix

Let $g(y;x,\theta)$ and $f(y;x,\theta)$ be two statistical models for the response variable Y, where $x \in \mathcal{X}$ is an experimental condition and $\theta \in \Theta \subseteq \mathbb{R}^m$ is an unknown vector of parameters. Notice that the interpretation of the parameters must be the same in both models.

From now on, ξ is an approximate design, i.e. it is a discrete probability distribution with a finite number of support points, $x_1, \ldots, x_k$. Thus, ξ_x denotes the design which puts the whole mass at point x. Furthermore, hereafter $f(y;x,\theta)$ is the model used for computing the likelihood function, $\hat{\theta}$ is the MLE of θ under this model and θ_0 is the unknown "true" value of θ under the "true" model.

If $g(y;x,\theta_0)$ is the "true" probability density function (pdf) of Y and θ_g minimizes the Kullback-Leibler discrepancy between $g(y;x,\theta_0)$ and $f(y;x,\theta)$ with respect to θ for any $x \in \mathcal{X}$, then under usual regularity conditions, θ_g is the solution of the following system of equations

$$\frac{\partial}{\partial\theta}\int \log\frac{g(y;x,\theta_0)}{f(y;x,\theta)}\, g(y;x,\theta_0)\, dy = -\int \frac{\partial}{\partial\theta}\, \log[f(y;x,\theta)]\, g(y;x,\theta_0)\, dy = 0, \quad x \in \mathcal{X}. \tag{1}$$

The assumption that there exists a common θ_g which minimizes the Kullback-Leibler discrepancy between $g(y;x,\theta_0)$ and $f(y;x,\theta)$, for any $x \in \mathcal{X}$, is fundamental for proving that

$$\sqrt{n}(\hat{\theta} - \theta_g) \to N(0, M_g(\xi;\theta_g)^{-1} K(\xi;\theta_g) M_g(\xi;\theta_g)^{-1}), \tag{2}$$

(for the proof, see for instance [18]) where

$$K(\xi;\theta) = \int_{\mathcal{X}} \int \frac{\partial \log[f(y;x,\theta)]}{\partial\theta}\, \frac{\partial \log[f(y;x,\theta)]}{\partial\theta^T}\, g(y;x,\theta_0)\, dy\, d\xi(x)$$

and

$$M_g(\xi;\theta) = \int_{\mathcal{X}} \int -\frac{\partial \log[f(y;x,\theta)]}{\partial\theta\partial\theta^T}\, g(y;x,\theta_0)\, dy\, d\xi(x).$$

The matrix $IS(\xi;\theta_g) = M_g(\xi;\theta_g)^{-1}K(\xi;\theta_g)M_g(\xi;\theta_g)^{-1}$ is called information sandwich variance matrix and, except for the constant of proportionality n, is the asymptotic covariance matrix of the MLE when the wrong model is used.

On the other hand, if the "true" pdf is $f(y;x_i,\theta_0)$, i.e. if the used model is correctly specified, then from (2) we get the standard result $\sqrt{n}(\hat{\theta}-\theta_0) \to N(0, M(\xi;\theta_0)^{-1})$, where $M(\xi;\theta_0)$ is the Fisher information matrix (except for the constant of proportionality n).

Thus, from equation (2) we have that whenever $\theta_g = \theta_0$, the MLE $\hat{\theta}$ is consistent for θ_0, whether or not the used model is correctly specified.

3. A robust optimum design to a wrong model

Let $\tilde{\theta}_0$ be a nominal value for θ_0. If the used model for computing the likelihood function is correctly specified, then $\Phi_D[M(\xi;\tilde{\theta}_0)] = \log|M(\xi;\tilde{\theta}_0)|$ has to be maximized in order to have a precise estimate of θ_0. Otherwise, if $g(y;x,\theta_0)$ is the "true" pdf of Y, then $\Phi_D[IS(\xi;\tilde{\theta}_0)^{-1}] = -\log|IS(\xi;\tilde{\theta}_0)|$ should be maximized. The corresponding optimum design, i.e. $\xi^*_{IS} = \arg\max_\xi \Phi_D[IS(\xi;\tilde{\theta}_0)^{-1}]$, can be found through standard algorithms and its optimality can be checked using the result here below.

Theorem 1. *A design ξ maximizes $\Phi_D[IS(\xi;\tilde{\theta}_0)^{-1}]$ if and only if*

$$\psi_{IS}(x,\xi) = 2\,\mathrm{tr}[M_g^{-1}(\xi;\tilde{\theta}_0)M_g(\xi_x;\tilde{\theta}_0)] - \mathrm{tr}[K^{-1}(\xi;\tilde{\theta}_0)K(\xi_x;\tilde{\theta}_0)] - m \le 0, \quad x \in \mathcal{X}. \tag{3}$$

Let $\mathrm{Eff}_D(\xi) = (|M(\xi;\tilde{\theta}_0)|/|M(\xi^*_D;\tilde{\theta}_0)|)^{1/m}$ and $\mathrm{Eff}_{IS}(\xi) = (|IS(\xi^*_{IS};\tilde{\theta}_0)|/|IS(\xi;\tilde{\theta}_0)|)^{1/m}$ be the efficiencies of a design ξ with respect to $\xi^*_D = \arg\max_\xi \Phi_D[M(\xi;\tilde{\theta}_0)]$ and ξ^*_{IS}, respectively. When there is not complete confidence that the chosen model for drawing inferences is correct, then the following geometric mean of efficiencies may be maximized,

$$\mathrm{Eff}_D(\xi)^\alpha\, \mathrm{Eff}_{IS}(\xi)^{(1-\alpha)}, \tag{4}$$

where $0 \le \alpha \le 1$ may be chosen to balance between the two possibilities about the truth of model $f(y;x,\theta)$. To maximize (4) or its logarithm is equivalent to maximize the following criterion function

$$\Phi_{DR}(\xi;\tilde{\theta}_0) = \frac{1}{m}\left\{\alpha\,\Phi_D[M(\xi;\tilde{\theta}_0)] + (1-\alpha)\,\Phi_D[IS(\xi;\tilde{\theta}_0)^{-1}]\right\}, \tag{5}$$

which is called DR-optimality criterion. The corresponding optimum design, i.e. $\xi^*_{DR} = \arg\max_\xi \Phi_{DR}(\xi;\tilde{\theta}_0)$, is called DR-optimum design. Design criterion (5) is concave thus the following equivalence theorem may be stated.

Theorem 2. *A design ξ^*_{DR} is DR-optimum if and only if it fulfils the following inequality*

$$\psi_{DR}(x,\xi^*_{DR}) = \alpha\,\psi_D(x,\xi) + (1-\alpha)\,\psi_{IS}(x,\xi) \le 0, \quad x \in \mathcal{X},$$

where $\psi_D(x,\xi) = \mathrm{tr}[M^{-1}(\xi;\tilde{\theta}_0)M(\xi_x;\tilde{\theta}_0)] - m$.

A SVE is a consistent estimator for $IS(\xi;\tilde{\theta}_0)$, which simplifies to Fisher's information matrix in absence of model misspecification. Since $\Phi_{DR}(\xi;\tilde{\theta}_0)$ combines $\Phi_D[M(\xi;\tilde{\theta}_0)]$ with $\Phi_D[IS(\xi;\tilde{\theta}_0)^{-1}]$, ξ^*_{DR} is a "good" design when the model is correctly specified and also when there is model misspecification (at least for some values of α). For this reason ξ^*_{DR} may be considered a robust design to a misspecified model and it should be computed whenever a SVE is used for estimating the asymptotic covariance matrix of the MLE.

4. An application

Let $f(y;\mu)$ be the family of exponential pdf's with mean μ and $g(y;\sigma^2)$ be the family of standard log-normal pdf's with mean $\mu = e^{\sigma^2/2}$ and variance $e^{\sigma^2}(e^{\sigma^2}-1)$. Let us assume that $\mu = \theta_1 + \theta_2 x$ where $\theta = (\theta_1,\theta_2)^T$ is a vector of unknown parameters and $x \in \mathcal{X} = [0,1]$ is an experimental condition. Under model $g(y;\sigma^2)$, $\mu = e^{\sigma^2/2} > 1$ thus the parametric space of θ must be $\Theta = \{(\theta_1,\theta_2) : (\theta_1 > 1, \theta_2 > 0), (\theta_1+\theta_2 > 1, \theta_2 < 0)\}$. The interest is in the estimation of θ_0, i.e. the unknown "true" value of θ under the true pdf. Model $f(y;x,\theta)$ is used for drawing inferences.

If $f(y;x,\theta_0)$ is actually the "true" model then the D-optimum design for estimating θ_0 must be computed by maximizing $\Phi_D[M(\xi;\tilde{\theta}_0)]$, where $\tilde{\theta}_0 = (\tilde{\theta}_{01},\tilde{\theta}_{02})^T$ is a nominal value for θ_0. It is well known that a design ξ is D-optimum if and only if $\psi_D(x,\xi) \le 0$, for any $x \in \mathcal{X}$. Let ξ^*_D be the equally weighted design with support points 0 and 1. We have that

$$\psi_D(x,\xi^*_D) = \frac{4\,\tilde{\theta}_{01}\,x\,(x-1)\,(\tilde{\theta}_{01}+\tilde{\theta}_{02})}{(\tilde{\theta}_{01}+\tilde{\theta}_{02}\,x)^2} \le 0, \quad x \in [0,1],$$

thus ξ^*_D, which does not depend on $\tilde{\theta}_0$, is the D-optimum design.

If the "true" pdf of Y is $g(y;x,\theta_0)$, then from (1) there exists a common value $\theta_g = (\theta_{1g},\theta_{2g})^T$ which minimizes the Kullback-Leibler discrepancy between $f(y;x,\theta)$ and $g(y;x,\theta_0)$ for any $x \in [0,1]$ and it is $\theta_g = \theta_0$. Thus, the MLE of θ is still consistent for θ_0 and the optimum design should be chosen by maximizing $\Phi_D[IS(\xi;\tilde{\theta}_0)^{-1}]$. From equation (3), $\psi_{IS}(x,\xi)$ evaluated at ξ^*_D is

$$\psi_{IS}(x,\xi^*_D) = \frac{4\,x\,(x-1)\,(a+b\,x+c\,x^2)}{(\tilde{\theta}_{01}^2-1)(\tilde{\theta}_{01}+\tilde{\theta}_{02})^2(\tilde{\theta}_{01}+\tilde{\theta}_{02}x)^2[(\tilde{\theta}_{01}+\tilde{\theta}_{02})^2-1]} \tag{6}$$

where a, b and c are bivariate polynomials of degree 8 in $\tilde{\theta}_{01}$ and $\tilde{\theta}_{02}$. The denominator in the right-hand side of equation (6) is always positive and

$$p(x) = a + b\,x + c\,x^2 \ge 0, \quad x \in [0,1],$$

in the following cases:

1. if $c > 0$ and $\Delta = b^2 - 4ac \le 0$;
2. if $c > 0$, $\Delta > 0$ and $x_{min} < x_{max} \le 0$, where x_{min} and x_{max} are the solutions of equation $p(x) = 0$;

3. if $c > 0$, $\Delta > 0$ and $x_{max} > x_{min} \geq 1$;

4. if $c < 0$, $\Delta > 0$, $x_{min} \leq 0$ and $x_{max} \geq 1$.

If one of the previous conditions is fulfilled, then

$$\xi_D^* = \arg\max_{\xi} \log |M(\xi; \tilde{\theta}_0)| = \arg\max_{\xi} \log |IS(\xi; \tilde{\theta}_0)^{-1}|,$$

i.e. ξ_D^* is optimum even when the underling true model is the log normal. For instance, if $\tilde{\theta}_0 = (3.2, -0.5)^T$ then $a = 2909.42$, $b = 1534.27$, $c = -102.075$, $\Delta > 0$, $x_{min} = -1.70$ and $x_{max} = 16.73$. Thus, condition 4 is satisfied and ξ_D^* is an optimum design as shown by Figure 1, which displays function $\psi_{IS}(x, \xi_D^*)$.

Otherwise, when $\tilde{\theta}_0$ is such that none of the previous conditions is satisfied then, ξ_D^* is not robust to the wrong model and $\xi_{DR}^* = \arg\max_{\xi} \Phi_{DR}(\xi; \tilde{\theta}_0)$ should be computed. This happens for instance if $\tilde{\theta}_0 = (2, 3)^T$, since $a = 32400$, $b = -23287.5$, $c = -19237.5$, $\Delta > 0$, $x_{min} = -2.04$ and $x_{max} = 0.83$. Indeed, $\xi_{IS}^* = \arg\max_{\xi} \Phi_D[IS(\xi; \tilde{\theta}_0)^{-1}]$ is the equally weighted design with support points 0 and 0.6355 (Figure 2) while ξ_D^* is equally supported at 0 and 1. In this case,

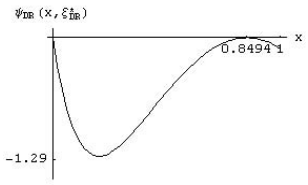


Figure 1: Function $\psi_{IS}(x, \xi_D^*)$.

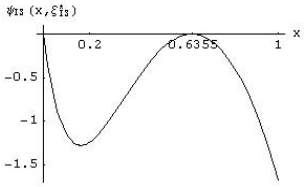


Figure 2: Function $\psi_{IS}(x, \xi_{IS}^*)$.

a DR-optimum design should be computed, which depends on the choice of the weight α. Figure 3 displays $\text{Eff}_D(\xi_{DR}^*)$ and $\text{Eff}_{IS}(\xi_{DR}^*)$ for different values of α. The intersection point gives the value of α for which there is equal efficiency. This value, denoted by α^*, should be chosen whenever there is complete uncertainty about the truth of the model and the common efficiency is quite large.

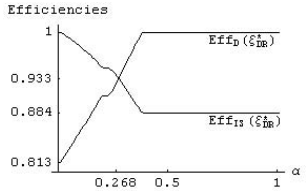


Figure 3: Efficiencies of ξ_{DR}^* with respect to ξ_D^* and ξ_{IS}^*, respectively.

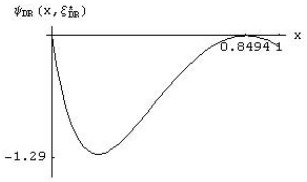


Figure 4: Function $\psi_{DR}(x, \xi_{DR}^*)$.

In this example, for $\alpha^* = 0.268$, ξ_{DR}^* is the design which puts equal weights at points 0 and 0.8494 (as shown in Figure 4) and provides a common efficiency of about 93%.

References

[1] Adewale, A., Wiens, D.P., 2009 Robust Designs for Misspecified Logistic Models , *J. Statist. Plann. Inference*, 139, 3-15.

[2] Atkinson, A. C., 2008 DT-optimum designs for model discrimination and parameter estimation, *J. Statist. Plann. Inference*, 1, 56-64.

[3] Atkinson, A.C., Fedorov, V.V., 1975a The design of experiments for discriminating between two rival models, *Biometrika*, 62(1), 57-70.

[4] Atkinson, A.C., Fedorov, V.V., 1975b Optimal design: Experiments for discriminating between several models, *Biometrika*, 62(2), 289-303.

[5] Dette, H. 1993 On a mixture of the D- and D_1-optimality criterion in polynomial regression. *J. Statist. Plann. Inference*, 35, 233–249.

[6] Dette, H., Titoff, S., 2009 Optimal discrimination designs, *Ann. Statist.*, to appear.

[7] Wiens, D.P., 1998 Minimax robust designs and weights for approximately specified regression models, *J. Am. Stat. Assoc.*, 93, 1440-1450.

[8] Huber, P.J., 1967 The behaviour of maximum likelihood estimates under nonstandard conditions. *Proceedings of the Fifth Berkley Symposium on Mathematical Statistics and Probability*, 1, 221-233.

[9] Kauermann, G., Carrol, J. 2001 A note on the efficiency of sandwich covariance matrix estimation, *J. Am. Stat. Assoc.*, 96, 1387-1396.

[10] López-Fidalgo, J., Tommasi, C., Trandafir, P.C., 2008 Discrimination between some extensions of the Michaelis-Menten model. *J. Statist. Plann. Inference*, 138, 3797-3804.

[11] López-Fidalgo, J., Tommasi, C., Trandafir, P.C., 2007 An optimal experimental design criterion for discriminating between non-Normal models, *J. R. Statist. Soc. B*, 69, 231-242.

[12] Tsai, M., Zen, M., 2004 Criterion-robust optimal designs for model discrimination and parameter estimation: multivariate polynomial regression case, *Statistica Sinica*, 14, 591-601.

[13] Tommasi, C., 2008 Optimal designs for both model discrimination and parameter estimation, preprint, http://services.bepress.com/unimi/statistics/art34.

[14] Waterhouse, T.H., Woods, D.C., Eccleston, J.A., Lewis, S.M., 2008 Design selection criteria for discrimination/estimation for nested models and a binomial response, *J. Statist. Plann. Inference*, 138, 132-144.

[15] White, H., 1980 A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity, *Econometrica*, 48 (4), 817-834.

[16] White, H., 1982 Maximum likelihood estimation of misspecified models, *Econometrica*, 50 (1), 1-25.

[17] Wiens, D.P., 2009. Robust discrimination designs, *J. R. Statist. Soc. B*, to appear.

6th St.Petersburg Workshop on Simulation (2009) 589-593

Designs for Discriminating Between Models by Testing the Equality of Parameters

Barbara Bogacka[1], Anthony C. Atkinson[2], Maciej Patan[3]

Abstract

A general model for enzyme kinetics with inhibition, the "mixed" inhibition model, simplifies to the non-competitive inhibition model when two of the parameters are equal. We find D_S-optimum designs for testing the equality of the parameters in this non-linear model and in a linear model. Connections with T-optimum designs for model discrimination are considered.

1. Introduction

We consider the problem of the optimum design of experiments to determine whether two parameters have the same value. The motivation for this work arose from experiments in enzyme kinetics, where a meaningful simpler model is obtained when the values of two parameters coincide. The kinetic models are nonlinear. In order to motivate our approach we start our discussion in § with a simple linear example. The kinetic example is considered in §. The designs that we find are D_S-optimum. In § we ponder the relationship, for nonlinear models, between our designs and T-optimum designs for model discrimination.

Throughout we work with the customary second-order assumptions of additive independent errors of constant variance. We are then able to use the standard theory of optimum design for regression models as described in several books including [7], [3] and [1]. We focus on continuous designs expressed as a measure ξ over a design region $\mathcal{X}$.

2. Two Variable Regression

We start with simple linear regression on two variables, that is with the model

$$y_{ij} = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \epsilon_{ij} \quad i = 1, \ldots, n \quad \text{and} \quad j = 1, \ldots, n_i, \tag{1}$$

where the ϵ_{ij} have zero mean, are independent and have constant variance so that efficient estimation is by least squares. The experimental region $\mathcal{X}$ is the square

[1]Queen Mary, University of London, E-mail: b.bogacka@qmul.ac.uk
[2]London School of Economics, E-mail: a.c.atkinson@lse.ac.uk
[3]University of Zielona Góra, E-mail: m.patan@issi.uz.zgora.pl

Table 1: Two Variable Regression: D-optimum design for all three parameters and D_S-optimum design for parameter equality

Criterion	i	1	2	3	4	Efficiency %
	$\begin{bmatrix} x_{1i} \\ x_{2i} \end{bmatrix}$	$\begin{bmatrix} -1 \\ -1 \end{bmatrix}$	$\begin{bmatrix} 1 \\ -1 \end{bmatrix}$	$\begin{bmatrix} -1 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$	
D	w_i	1/4	1/4	1/4	1/4	50.
D_S	w_i	0	1/2	1/2	0	100.

$[-1, 1] \times [-1, 1]$. The D-optimum design for all three parameters is the 2^2 factorial shown in Table 1 putting weight 1/4 at the four corners of the design region.

To find designs for testing whether $\beta_1 = \beta_2$ we reparameterise (1) by writing $\beta_1 = \beta + \delta$ and $\beta_2 = \beta - \delta$. The deterministic part of the model then is

$$\mathrm{E}\,(y) = \beta_0 + \beta(x_1 + x_2) + \delta(x_1 - x_2). \tag{2}$$

Testing whether $\delta = 0$ is identical to testing whether $\beta_1 = \beta_2$. The design yielding the most powerful test will be that which minimises the variance of the estimator $\hat{\delta}$. This is the D_S-optimum design for δ. It is straightforward to show that this design, given in Table 1 puts weight 1/2 on the two corners of the design region where $x_1 \neq x_2$. As with all designs found in this paper, we checked its optimality by using the General Equivalence Theorem relating the criteria of D- and G-optimality [4]. Here we require the extension to D_S-optimality [1, p. 139]. Indeed the maximum of the derivative function $d_s(x, \xi)$ over $\mathcal{X}$ is one.

In these designs information about the equality of β_1 and β_2 comes only from comparison of the responses where $x_1 \neq x_2$, the points numbered 2 and 3 in the table. In the D-optimum design these points receive half of the observations and the efficiency of this design for testing parameter equality is only 50%. Of course, because the D_S-optimum design only has two support points it is not possible to estimate all three parameters in the model.

In more complicated examples we require a formal definition of efficiency. Let $\mathrm{var}\,\hat{\delta}(\xi)$ be the variance of $\hat{\delta}$ when some design ξ is used with $\mathrm{var}\,\hat{\delta}(\xi^*)$ the variance for the D_S-optimum design ξ^*. Clearly, $\mathrm{var}\,\hat{\delta}(\xi^*) \leq \mathrm{var}\,\hat{\delta}(\xi)$. Then the efficiency of the design ξ is defined as

$$\mathrm{Eff}\,(\xi) = \mathrm{var}\,\hat{\delta}(\xi^*)/\mathrm{var}\,\hat{\delta}(\xi). \tag{3}$$

As we see in Table 1, a design with an efficiency of 50% requires twice the number of trials to give the same variance for $\hat{\delta}$, and so a test with the same power, as does the D_S-optimum design. This definition of efficiency is an extension to D_S-optimality of the customary definition of D-efficiency in terms of a power of the ratio of determinants of information matrices [1, p. 151].

3. Enzyme Kinetics

The models arise in the assessment of drug metabolism. Endogenous enzymes typically metabolize the drug of interest. In studies of inhibition, interest is rather in the effect of the drug in preventing the metabolism of another drug. The importance is in the study of unwanted adverse drug reactions.

The models relate the velocity of reaction v to concentrations of substrate $[S]$ and of inhibitor $[I]$. Different types of binding lead to a variety of models for the reaction. Here we consider two possible models.

Linear Mixed Inhibition. In this four-parameter model the deterministic velocity equation is:

$$v = \frac{V[S]}{K_m\left(1+\dfrac{[I]}{K_c}\right)+[S]\left(1+\dfrac{[I]}{K_u}\right)}, \tag{4}$$

with the parameters V, K_m, K_u and K_c to be determined experimentally. In designing experiments it is assumed that the errors in the measurements of v follow the second-order assumptions.

Non-competitive Inhibition. When $K_u = K_c$ the model has a specific interpretation and becomes

$$v = \frac{V[S]}{(K_m+[S])\left(1+\dfrac{[I]}{K_c}\right)}. \tag{5}$$

To obtain efficient designs for testing the equality of K_c and K_u we rewrite the model (4) in a form analogous to (2). If we let $\theta_1 = 1/K_c$ and $\theta_2 = 1/K_u$, (4) becomes

$$v = \frac{V[S]}{K_m\left(1+\theta_1[I]\right)+[S]\left(1+\theta_2[I]\right)}. \tag{6}$$

We now make a reparameterization similar to that of § and write $\theta_1 = \theta + \delta$ and $\theta_2 = \theta - \delta$, when (6) becomes

$$v = \frac{V[S]}{(K_m+[S])\left(1+\theta[I]\right)+\delta[I]\left(K_m-[S]\right)}, \tag{7}$$

which reduces to (5) when $\delta = 0$. Efficient designs for testing this reduction will minimize the variance of the estimator of δ. An experimental design involves the choice of concentrations $x_i = ([S]_i, [I]_i)^{\mathrm{T}}$ at which measurements are to be taken.

Let the vector of four parameters in (7) be written as $\psi = (V, K_m, \theta, \delta)^{\mathrm{T}}$. The information matrix is a function of the vector of partial derivatives

$$f_i(x_i, \psi^0) = \left.\frac{\partial v(x_i,\psi)}{\partial \psi}\right|_{\psi^0} \tag{8}$$

of the response function with respect to the parameters ψ, often called the parameter sensitivities. See Chapter 17 of [1].

Since (7) is nonlinear in three of the parameters, optimum designs will depend on the values of all parameters except V. In this paper we find locally optimum designs that depend on the value of ψ^0, a prior point estimate of the parameters.

In an unpublished technical report we present analytical expressions for D-optimum designs for the nonlinear model (4) or, equivalently, for the reparameterized form (7). When the design region is a rectangle $\mathcal{X} = \big[[S]_{\min}, [S]\max\big] \times \big[[I]_{\min}, [I]\max\big]$ the optimum design has the form

$$\xi^* = \begin{Bmatrix} ([S]_{\max}, [I]_{\min})^T & (s_2, [I]_{\min})^T & ([S]_{\max}, i_3)^T & (s_4, i_4)^T \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{Bmatrix}, \qquad (9)$$

so that four settings of experimental values (s_2, s_4, i_3 and i_4) have to be calculated. We found the locally D-optimum design for the initial parameter values $(V^0, K_m^0, K_c^0, K_u^0) = (1513, 6.59, 1.35, 1.35)$ when $\mathcal{X}$ is such that $[S] \in [0, 100]$ and $[I] \in [0, 100]$. The resulting D-optimum design is in Table 2. Although our interest was not in testing the equality of K_c and K_u, the chosen parameter values support this hypothesis and we use them here in designing our experiments.

In Table 2 we compare three designs for testing whether δ in (7) is zero. Since we take K_c^0 and K_u^0 equal, we are finding designs under this hypothesis. The first design is the D-optimum design given in the report. For the second design we keep the support points of the D-optimum design but perform a three-dimensional numerical search to find the D_S-optimum weights. The results, given in Table 2 are surprising. To five decimal places (we haven't explored further) the weights are the fractions 1/9, 2/9, 2/9 and 4/9. It seems that another analytical result may be obtained here. However, this design is not the D_S-optimum design, as application of the Equivalence Theorem shows. A coarse search over 1681 grid points in $\mathcal{X}$ gave a maximum value of 1.1498 for the maximum of $d_s(x, \xi)$ as opposed to one for the optimum design. The closeness of this value to one shows that the design is not markedly suboptimum in its properties.

The analytical results do not extend to D_S-optimality. To find an optimum design with four support points when $\mathcal{X}$ is two dimensional requires an 11-dimensional numerical search. Instead we finesse the problem, using the equivalence theorem to check our results.

The D-optimum design of Table 2, found when $\delta = 0$, is of the form

$$\xi^* = \begin{Bmatrix} ([S]_{\max}, [I]_{\min})^T & (s_2, [I]_{\min})^T & ([S]_{\max}, i_3)^T & (s_2, i_3)^T \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{Bmatrix}, \qquad (10)$$

a special case of (9) with only two unknowns. To find the D_S-optimum design we performed a five dimensional search over the weights and the values of s_2 and i_3. The resulting design is in Table 2. To check the optimality of this design we performed a search over a dense grid of 1,002,001 points in $\mathcal{X}$. Because we are using a numerical method, we started our search from a variety of different initial designs. On our third search the maximum value of the derivative function $d_s(x, \xi)$ was 1.0001. The D_S-optimum design for parameter equality certainly has the structure of design points given by (10).

The design weights for the D_S-optimum design in the last row of Table 2 are suggestive of further numerical simplifications. The numerical results suggest that

Table 2: Enzyme Kinetics: D-optimum design and D_S-optimum designs for parameter equality

Criterion	i	1	2	3	4	Eff.%
D	$\begin{bmatrix}[S_i]\\[I_i]\end{bmatrix}$	$\begin{bmatrix}100.\\0.\end{bmatrix}$	$\begin{bmatrix}5.8226\\0.\end{bmatrix}$	$\begin{bmatrix}100.\\1.35\end{bmatrix}$	$\begin{bmatrix}5.8226\\1.35\end{bmatrix}$	
	w_i	1/4	1/4	1/4	1/4	72.12
D_s weights	$\begin{bmatrix}[S_i]\\[I_i]\end{bmatrix}$	$\begin{bmatrix}100.\\0.\end{bmatrix}$	$\begin{bmatrix}5.8226\\0.\end{bmatrix}$	$\begin{bmatrix}100.\\1.35\end{bmatrix}$	$\begin{bmatrix}5.8226\\1.35\end{bmatrix}$	
	w_i	1/9	2/9	2/9	4/9	89.04
D_S	$\begin{bmatrix}[S_i]\\[I_i]\end{bmatrix}$	$\begin{bmatrix}100.\\0.\end{bmatrix}$	$\begin{bmatrix}4.1877\\0.\end{bmatrix}$	$\begin{bmatrix}100.\\1.9093\end{bmatrix}$	$\begin{bmatrix}4.1877\\1.9093\end{bmatrix}$	
	w_i	0.0858	0.2071	0.2071	0.5000	100.

point 4 has a weight one half, that points 2 and 3 share the same weight and that the residue is on point 1. The design is itself quite close in support points and weights to the second design in the table, that for the D-optimum points with weights found to maximize the criterion of D_S-optimality. This design has an efficiency of 89.04% in line with the value of 1.1498 found for $d_s(x, \xi)$. The D-optimum design has a lower efficiency of 72.12%. An important practical point is whether a compound design can be found that has both good D-efficiency for all four parameters and good D_S-efficiency for testing the equality of parameter values. Chapter 21 of [1] describes methods for finding efficient multi-criterion compound designs.

4. T-optimality

A seemingly different approach to optimum experimental designs for discriminating between models is T-optimality introduced by [2]. To discriminate between the four-parameter model (4) and the three-parameter model (5) experiments should be performed where the models are "furthest apart". Where in $\mathcal{X}$ this occurs depends on the values of the parameters in the larger model. In the construction of optimum designs the parameter values for the smaller model are estimated from the expected response of the larger model, and depend on the design as well as on the parameter values.

The procedure applies to both linear and nonlinear models, which need not be nested. For linear models such as (1) and (2) with $\delta = 0$, that differ by a single

parameter, the T-optimum design does not depend on the parameters of the larger model and is identical to the D_S-optimum design for δ. However, the relationship between the two approaches for nonlinear models is more complex.

We found D_S-optimum designs for the nonlinear model for enzyme kinetics with δ in (7) equal to zero. It is clear from the construction of T-optimum designs outlined above that the two models have to differ, that is that $\delta \neq 0$. For very small δ, T-optimum designs are obtained which are very close to the D_S-optimum designs for the same δ, but the designs become increasingly different as δ increases. Discussion and an example, for an extension of the Michaelis-Menten model, are given by [5]. The construction also suggests that T-optimum designs may be much less stable than D_S-optimum designs when observational error is relatively large compared to the effects to be estimated.

Designs for model discrimination, with an emphasis on T-optimality, are in Chapter 20 of [1] and §21.8 describes compound DT-optimum designs for simultaneous parameter estimation and model discrimination. [8] extend T-optimality to designs in which the factors can be time traces of, for example, temperature in a chemical reaction. [6] extend T-optimality to non-normal models.

References

[1] A. C. Atkinson, A. N. Donev, and R. D. Tobias. *Optimum Experimental Designs, with SAS*. Oxford University Press, Oxford, 2007.

[2] A. C. Atkinson and V. V. Fedorov. The design of experiments for discriminating between two rival models. *Biometrika*, 62:57–70, 1975.

[3] V. V. Fedorov and P. Hackl. *Model-Oriented Design of Experiments*. Lecture Notes in Statistics 125. Springer Verlag, New York, 1997.

[4] J. Kiefer and J. Wolfowitz. Optimum designs in regression problems. *Annals of Mathematical Statistics*, 30:271–294, 1959.

[5] J. López-Fidalgo, C. Tommasi, and C. Trandafir. Optimal designs for discriminating between some extensions of the Michaelis-Menten model. *Journal of Statistical Planning and Inference*, 138:3797–3804, 2008. doi: 10.1016/j.jspi.2008.01.014.

[6] J. López-Fidalgo, C. Trandafir, and C. Tommasi. An optimal experimental design criterion for discriminating between non-normal models. *Journal of the Royal Statistical Society, Series B*, 69:231–242, 2007.

[7] F. Pukelsheim. *Optimal Design of Experiments*. Wiley, New York, 1993.

[8] D. Uciński and B. Bogacka. T-optimum designs for discrimination between two multiresponse dynamic models. *Journal of the Royal Statistical Society, Series B*, 67:3–18, 2005.

6th St.Petersburg Workshop on Simulation (2009) 595-599

Optimal Designs for Discriminating Between Pharmacokinetic Models with Correlated Observations[1]

Mariano Amo-Salas[2], Jesús López-Fidalgo[3],Víctor López-Ríos[4]

Abstract

Two nested pharmacokinetic models are considered in this work. The observations are taken on the same subject so the samples are correlated. The covariance function assumed is an exponential covariance function. Optimal exact designs are computed for each model with different criteria. Moreover, compound designs to estimate the parameters and nonlinear functions of the parameters are computed. An iterative algorithm based on T-optimality and an algorithm from Brimkulov, Krug and Savanov [1] is adapted in order to compute T-optimal designs with correlated observations. Finally, compound designs to discriminate between the models and estimate the nonlinear functions are considered. A test power study is provided as well.

1. Introduction

Pharmacokinetics is the study of various biological processes affecting a drug: dissolution, absorption, distribution, metabolism, and elimination. A pharmacokinetic model is used to describe the concentration of such substances into the organism over time. Pharmacokinetic data is collected for each subject over time, so the first issue is to define the optimal sampling times in order to estimate several characteristics in the compartmental model of interest.

Optimal designs for discrimination between models have usually very poor properties for estimation of the parameters in the chosen model [2]. Here we use Optimal Design Theory to provide designs of known properties with a specified balance between estimation and discrimination. In this work, we are interested in a couple of compartmental models, one with four components with reversible rates of transfer and the other with three components (see [3], [4], [5] and [6] for similar works). The aim of this paper is to find optimal experimental designs for a nonlinear model arising in a particular compartment, C (Figure **??**), in order to discriminate between both models and estimate in an optimal way several

[1]This work was supported by grant MTM2007 67211 C03-01 and PAI07-0019.

[2]University of Castilla-La Mancha, E-mail: Mariano.Amo@uclm.es

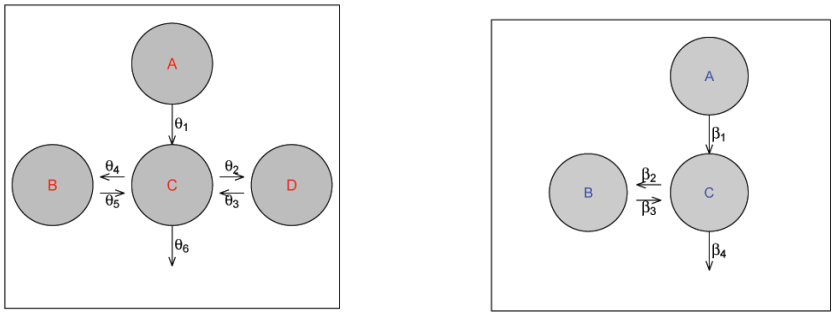
[3]University of Castilla-La Mancha, E-mail: Jesus.LopezFidalgo@uclm.es

[4]National University of Colombia, E-mail: victorignaciolopez@gmail.com

nonlinear functions of the parameters (rates of transfer) used in pharmacokinetics. The optimization is being performed in the correlated case with respect to various criteria, which depend on the Fisher information matrix.

2. A couple of Compartmental Models

We are interested in a couple of compartmental models, see Figure **??**. Both models describe, for example, the oral drug administration. The difference between them is in one compartment.



(a) (b)

Figure 1: Two compartmental models, (a) Model I: Four-compartment model, with reversible rates between the last three compartments, with rates of transfer $\theta_i, i = 1, \ldots, 6$, (b) Model II: Three-compartment model, with reversible rates between the last two compartments.

The objective is to obtain optimal sampling times in order to estimate the rates of transfer. Measures are taken in the central compartment, C.

The models considered is this work are:

$$\eta_C^I(t;\Theta_1) = \theta_1 A_0 \left[h_1(\Theta_1)\, e^{-\theta_1 t} + h_2(\Theta_1)\, e^{\lambda_2 t} + h_3(\Theta_1)\, e^{\lambda_3 t} + h_4(\Theta_1)\, e^{\lambda_4 t}\right], \qquad (1)$$

where $\Theta_1^T = (\theta_1, \theta_2, \theta_3, \theta_4, \theta_5, \theta_6)$, h_i are functions of θ_i and λ_i and λ_i, $i = 2, 3, 4$ are solutions of the cubic polynomial:

$$\lambda^3 + \lambda^2(\theta_2 + \theta_3 + \theta_4 + \theta_5 + \theta_6) + \lambda\left[(\theta_2 + \theta_3 + \theta_6)\,\theta_5 + (\theta_4 + \theta_6)\,\theta_3\right] + \theta_3\theta_5\theta_6 = 0.$$

$$\eta_C^{II}(t;\Theta_2) = \frac{A_0\beta_1}{\lambda_2 - \lambda_3}\left[r_1(\Theta_2)\, e^{-\beta_1 t} + r_2(\Theta_2)\, e^{\lambda_2 t} + r_3(\Theta_2)\, e^{\lambda_3 t}\right], \qquad (2)$$

where $\Theta_2^T = (\beta_1, \beta_2, \beta_3, \beta_4)$, r_i are functions of β_i and λ_i and λ_i, $i = 2, 3$ are solutions of the quadratic polynomial:

$$\lambda^2 + \lambda(\beta_2 + \beta_3 + \beta_4) + \beta_3\beta_4 = 0.$$

In both models, we consider the following characteristics of interest, which are nonlinear functions of Θ_i $(i = 1, 2)$:

- Area under the curve: either $F_1(\Theta_1^0) = \int_0^\infty \eta_C^I(t; \Theta_1^0)dt$ for the model I or $G_1(\Theta_2^0) = \int_0^\infty \eta_C^{II}(t; \Theta_2^0)dt$ for the model II.
- Time to maximum concentration: either $F_2(\Theta_1^0) = t^I_{\max} = \arg\max_t \eta_C^I(t; \Theta_1^0)$ for the model I or $G_2(\Theta_2^0) = t^{II}_{\max} = \arg\max_t \eta_C^{II}(t; \Theta_2^0)$ for the model II.
- Maximum concentration: either $F_3(\Theta_1^0) = \eta_C^I(t^I_{\max}; \Theta_1^0)$ for the model I or $G_3(\Theta_2^0) = \eta_C^{II}(t^{II}_{\max}; \Theta_2^0)$ for the model II.
- First time to attains the 50% of the maximum concentration: either $F_4(\Theta_1^0) = T^I_{0.50}$ for the model I or $G_4(\Theta_2^0) = T^{II}_{0.50}$ for the model II,

where Θ_1^0 and Θ_2^0 are local values for Θ_1 and Θ_2 respectively.

Let a general nonlinear regression model be

$$Y(t) = \eta(t; \Theta) + \epsilon(t), \;\; t \in \chi,$$

where the random variables $\epsilon(t)$ are assumed normally distributed with zero mean and $Cov(\epsilon(t), \epsilon(t')) = c(t, t', \rho)$, where $c(.,.)$ is a known function. Therefore $\eta(t; \Theta)$ may be one of the two partially known functions $\eta_C^I(t; \Theta_1)$ or $\eta_C^{II}(t; \Theta_2)$, where $\Theta_1 \in \Omega_1 \subseteq \mathbb{R}^6$, and $\Theta_2 \in \Omega_2 \subseteq \mathbb{R}^4$ are unknown parameter vectors. We assume that

$$\text{Cov}\left(\epsilon(t), \epsilon(t')\right) = \sigma^2 exp\left(-\rho\left|t - t'\right|\right), \tag{3}$$

the so-called exponential covariance function. The covariance matrix will be denoted by Σ.

The models are given by a sum of four and three exponential terms respectively, so these models are composed by nonlinear functions and nominal values are needed. These nominal values use here come from the results given in a scientific paper [7]. This paper provides the values of the transfer rates for the reduced model and these values are adapted for the four-compartment model. These nominal values considered in this example are:

$$A_0 = 68.96, \Theta_1^T = [1.10, 0.03, 0.06, 0.15, 0.10, 0.40], \Theta_2^T = [1.099, 0.17, 0.09, 0.40].$$

3. T-optimal designs

Both models describe the same process; the second model (η_C^{II}) considers three compartments whereas the first model (η_C^I) takes four compartments. In this point we are interested in the discrimination between both models. For that, T-optimality criterion is used. In this criterion one of the two rival models is assumed as the true model, η_C^I, therefore the parameters of this model, Θ_1, are assumed known before the experiment is realized. The criterion function is defined as follows:

$$\Phi(\xi_n) = \min_{\Theta_2} [\eta_C^I(\xi_n; \Theta_1) - \eta_C^{II}(\xi_n; \Theta_2)]^T \Sigma^{-1} [\eta_C^I(\xi_n; \Theta_1) - \eta_C^{II}(\xi_n; \Theta_2)].$$

This is a generalization of the T-optimality criterion function for the uncorrelated case. Optimal designs are computed by maximizing this criterion. These designs are exact designs due to the correlation between observations.

In order to compute exact optimal designs the numerical algorithm presented by Brimkulov, Krug and Savanov [1] is adapted to T-optimality in this section. This algorithm consists on finding D-optimal sampling points by estimating parameters in linear models for expectations of random fields. A general scheme of the algorithm for D-optimality with correlated observations is detailed [8]. It is an exchange-type algorithm that starts from an arbitrary initial n-point design. In the case of exact optimal designs the number of trials or points of the design are fixed by the practitioner and none of the points is repeated. At each iteration one support point is deleted from the current design and a new point is included in its place in maximizing the value of the criterion function. The algorithm is detailed below.

Algorithm

Step 1. Select an initial design $\xi_n^{(0)} = \{t_1^{(0)}, \ldots, t_n^{(0)}\}$ such that $t_i^{(0)} \neq t_j^{(0)}$ for $i, j \in I = \{1, 2, \ldots, n\}$ and $i \neq j$.

Step 2. Set $l = 0$ and compute:
$\tilde{\Theta}_2^{(l)} = \arg\min_{\Theta_2} [\eta_C^I(\xi_n^{(l)}; \Theta_1) - \eta_C^{II}(\xi_n^{(l)}; \Theta_2)]^T {\Sigma^{(l)}}^{-1} [\eta_C^I(\xi_n^{(l)}; \Theta_1) - \eta_C^{II}(\xi_n^{(l)}; \Theta_2)]$,
$\triangle(\xi_n^{(l)}) = [\eta_C^I(\xi_n^{(l)}; \Theta_1) - \eta_C^{II}(\xi_n^{(l)}; \tilde{\Theta}_2^{(l)})]^T {\Sigma^{(l)}}^{-1} [\eta_C^I(\xi_n^{(l)}; \Theta_1) - \eta_C^{II}(\xi_n^{(l)}; \tilde{\Theta}_2^{(l)})]$.

Step 3. Determine

$$(i^*, t^*) = \arg \max_{(i,t) \in IxT} \triangle(\xi_{n,t_i \rightleftarrows t}^{(l)}),$$

where $\xi_{n,t_i \rightleftarrows t}^{(l)}$ means that the support point t_i in design $\xi_n^{(l)}$ is changed by t and $\chi = [0, t_\infty]$.

Step 4. If $\triangle(\xi_{n,t_{i^*} \rightleftarrows t^*}^{(l)}) - \triangle(\xi_{n,t}^{(l)}) \leq \delta$, where δ is the given tolerance, then STOP. Otherwise, $\xi_n^{(l+1)} = \xi_{n,t_{i^*} \rightleftarrows t^*}^{(l)}$ and set $l \leftarrow l + 1$ and go to Step 2.

In previous works the uncorrelated case for these models has been studied. The T-optimal design for this case has five points, so we assume an initial design with five support points.

Table 1 shows T-optimal designs for the correlated case with five support points where η_C^I is assumed as the true model.

And now we compute the efficiencies to obtain the robustness of these designs.

Table 1 shows the 5-point exact designs to discriminate between the models with an exponential correlation matrix. The value of ρ expresses the degree of correlation between observations. When the value of ρ is higher than 10, designs are similar due to the correlation between observations is small. On the other hand, Table 1 shows that designs are quite different when correlation is higher.

Table 2 shows the efficiencies of the designs for different values of ρ and ρ_0, where ρ is the true value of the parameter and ρ_0 is the nominal value. These outcomes show that the designs are not robust. The worst case, is for a nominal

value of 0.001 and a real value of 10, then the efficiency is 7.68%. Therefore it is important to know the correlation between observations in order to select an adequate design.

Table 1: T-optimal designs for different values of ρ

ρ	optimal design					optimal value
0.001	1.61	2.51	4.97	11.02	26.09	5.75
0.01	1.17	2.18	5.03	12.54	36.63	7.03
0.1	0.53	1.71	4.93	13.72	49.81	8.14
1	0.37	1.79	5.75	16.27	52.56	8.38
10	0.34	1.94	5.86	16.35	52.64	8.39

Table 2: Efficiencies of T-optimal designs for different values of ρ

ρ/ρ_0	0.001	0.01	0.1	1	10
0.001	100.00	60.73	22.44	18.22	18.38
0.01	60.25	100.00	65.61	53.89	53.83
0.1	23.13	67.21	100.00	92.06	91.26
1	8.85	38.31	83.91	100.00	99.82
10	7.68	35.00	79.13	99.73	100.00

Table 1 shows the 5-point exact designs to discriminate between the models with an exponential correlation matrix. The value of ρ expresses the degree of correlation between observations. When the value of ρ is higher than 10, designs are similar due to the correlation between observations is small. On the other hand, Table 1 shows that designs are quite different when correlation is higher. Table 2 shows the efficiencies of the designs for different values of ρ and ρ_0, where ρ is the true value of the parameter and ρ_0 is the nominal value. These outcomes show that the designs are not robust. The worst case, is for a nominal value of 0.001 and a real value of 10, then the efficiency is 7.68%. Therefore it is important to know the correlation between observations in order to select an adequate design.

4. L-optimality for discriminating and estimating

In this section we will compute optimal designs to discriminate between two rival models and estimate nonlinear functions by means of the L-optimality criterion (a design ξ_n is an L-optimal design if maximizes $[Tr(K^T I^{-1}(\xi_n, \Theta)K)]^{-1})$). For that, matrix K was created in such a way the entries were the ratios between the gradients of the functions and their averages.

Now L-optimality will be used to discriminate between models. These models are nested and we are interested in estimating the parameters θ_2 and θ_3 of model η_C^I because when their parameters are zero model η_C^{II} is obtained. Therefore K will be completed to estimate these parameters too. Thus, K will have two more columns with the ratios between the gradients of the parameters and their averages. With this matrix we compute the L-optimal designs for model η_C^I.

Table 3 shows the values of L-optimal designs in order to estimate the nonlinear functions and the parameters θ_2 and θ_3 simultaneously. The efficiencies of these

designs are computed in order to study the robustness of the designs with respect to the parameter of the correlation function, ρ, Table 4 shows the efficiencies of L-optimal designs. For values of ρ higher than 1 the designs obtained are similar and therefore efficiency is high. But designs are different when the ρ-value is low and efficiencies are less than 50% in some cases.

Table 3: L-optimal designs for different values of ρ for estimation and discrimination

ρ	optimal design						optimal value
0.01	0.4969	1.646	4.562	11.12	27.04	63.62	$3.136e-05$
0.1	0.5129	1.672	4.671	11.79	30.28	78.57	$8.040e-06$
1	0.3661	1.749	5.345	13.70	33.22	80.90	$5.973e-06$
10	0.3386	1.896	5.473	13.78	33.30	80.98	$5.899e-06$
100	0.3386	1.896	5.473	13.78	33.30	80.98	$5.899e-06$
1000	0.3386	1.896	5.473	13.78	33.30	80.98	$5.899e-06$

Table 4: Efficiencies of L-optimal designs for different values of ρ for estimation and discrimination

ρ/ρ_0	0.01	0.1	1	10	100	1000
0.01	100.00	70.54	54.23	53.26	53.26	53.26
0.1	76.34	100.00	90.93	89.95	89.95	89.95
1	46.91	83.04	100.00	99.78	99.78	99.78
10	42.35	77.85	99.67	100.00	100.00	100.00
100	42.35	77.85	99.67	100.00	100.00	100.00
1000	42.35	77.85	99.67	100.00	100.00	100.00

References

[1] Brimkulov U.N., Krug G.K. and Savanov V.L. (1986) *Design of Experiments in Investigating Random Fields and Processes.* Nauka, Moscow.

[2] Atkinson A.C. (2008) *DT-optimum designs for the Model Discrimination and Parameter Estimation.* Journal of Statistical Planning and Inference, 138: 56–64.

[3] Atkinson A.C., Chaloner K., Herzberg A.M., and Juritz J. (1993) *Optimum experimental designs for properties of a compartmental model.* Biometrics, 49 (2):325–337.

[4] Allen D.M. (1983) *Parameter estimation for nonlinear models with emphasis on compartmental models.* Biometrics, 39:629–637.

[5] Stroud J.R., Müller P., and Rosner G.L. (2001) *Optimal sampling times in population pharmacokinetic studies.* Appl. Statist., 50 Part 3:345–359.

[6] Waterhouse T.H. (2005) *Optimal Experimental Design for Nonlinear and Generalised Linear Models.* PhD thesis, University of Queensland, Australia.

[7] Davis J.L., Papich M.G., Morton A.J., Gayle J., Blikslarger A.T., and Campbell N.B. (2007) *Pharmacokinetics of etodolac in the horse following oral and intravenous administration.* J. vet. Pharmacol. Therap., 30 :43–48.

[8] Ucinski D. and Atkinson A.C. (2004) *Experimental design for time-dependent models with correlated observations.* Studies in Nonlinear Dynamics & Econometrics, 8(2), Article 13.

Session

Statistical inference from complex data

organized by Subig Ghosh (USA)

6th St.Petersburg Workshop on Simulation (2009) 605-609

Repeated Measures under Non-Sphericity

Edgar Brunner[1]

1. Introduction

We consider one group of n independent subjects observed under d different conditions. The repeated measures $Y_{k1}, \ldots, Y_{kd}$ observed on subject k, are assumed to follow a multivariate normal distribution

$$\mathbf{Y}_k = (Y_{k1}, \ldots, Y_{kd})' \ \sim \ N(\boldsymbol{\mu}, \mathbf{S}), \ k = 1, \ldots, n, \tag{1}$$

where $\boldsymbol{\mu} = E(\mathbf{Y}_k)$ denotes the expectation and $\mathbf{S} = Cov(\mathbf{Y}_k)$ denotes an unknown covariance matrix. We do not assume a particular structure of $\mathbf{S}$ which is commonly referred to as an *unstructured covariance matrix*. In classical analysis of variance (ANOVA) models, the observations Y_{ks} are decomposed as

$$\begin{aligned} Y_{ks} \quad = \quad & \mu_s + B_k + \epsilon_{ks}, \quad s = 1, \ldots, d; \ k = 1, \ldots, n \\ & B_k \sim N(0, \sigma_B^2), \quad \text{independent} \\ & \epsilon_{ks} \sim N(0, \sigma^2), \quad \text{independent} \\ & B_k \text{ and } \epsilon_{ks} \quad \text{independent.} \end{aligned}$$

Since $E(B_k) = E(\epsilon_{ks}) = 0$ and since B_k and ϵ_{ks} are assumed to be independent, one obtains

$$Var(Y_{ks}) = Var(B_k + \epsilon_{ks}) \ = \ \sigma_B^2 + \sigma^2$$

$$Cov(Y_{ks}, Y_{ks'}) = E\left[(B_k + \epsilon_{ks})(B_k + \epsilon_{ks'})\right] \ = \ E(B_k^2) = \sigma_B^2.$$

This generates the covariance matrix

$$\mathbf{S} = \begin{pmatrix} \sigma_B^2 + \sigma^2 & \cdots & \sigma_B^2 \\ \vdots & \ddots & \vdots \\ \sigma_B^2 & \cdots & \sigma_B^2 + \sigma^2 \end{pmatrix} \ = \ \sigma^2 \mathbf{I}_d + \sigma_B^2 \mathbf{J}_d,$$

where $\mathbf{I}_d$ denotes the d-dimensional unit matrix and $\mathbf{J}_d = \mathbf{1}_d \mathbf{1}_d'$ the $d \times d$ matrix of ones. The structure of this covariance matrix is referred to as *compound symmetric* which is a special case of sphericity where it is assumed that the variances of any differences Y_{ks} and $Y_{ks'}$ are identical.

[1]University of Göttingen, Humboldt Allee 32, 37073 Göttingen, Germany. E-mail: brunner@ams.med.uni-goettingen.de

In his seminal paper, Box (1954) suggested to approximate the distribution of the classical ANOVA statistic in this general model by a scaled F-distribution with degrees of freedom $f_1 = (d-1)\epsilon$ and $f_2 = (d-1)(n-1)\epsilon$. The quantity

$$\epsilon = \frac{[\mathrm{tr}(\mathbf{S})]^2}{(d-1)\mathrm{tr}(\mathbf{S}^2)} \tag{2}$$

is called Box's ϵ. If $\mathbf{S}$ is known then this approximation is quite accurate in the upper tail $F_q(f_1, f_2)$ for $q \geq 0.8$ of the distribution. The estimation of Box's ϵ has been considered in many papers in the literature, see e.g., Geisser and Greenhouse (1958), Greenhouse and Geisser (1959), Huynh and Feldt (1970), and Ahmad, Werner and Brunner (2008) among many others. For further discussion and applications see Keselman (1998) and Keselman, Algina, Wilcox, and Kowalchuk (2000).

Geisser and Greenhouse warned against replacing $\mathbf{S}$ in (2) with the the sample covariance matrix $\widehat{\mathbf{S}}_n = \frac{1}{n}\sum_{k=1}^{n}(\mathbf{Y}_k - \bar{\mathbf{Y}}.)(\mathbf{Y}_k - \bar{\mathbf{Y}}.)'$ to obtain an estimator of ϵ because of a potential bias. They suggested a conservative F-test using the degrees of freedom $f_1 = 1$ and $f_2 = (n-1)$. This procedure, however, may be extremely conservative and thus may have very low power.

It is our intention to present quite accurate estimators of the degrees of freedom f_1 and f_2. We will apply the estimators obtained by Ahamd, Werner and Brunner (2008) in a slightly different context. At the same time, we will admit a factorial structure of the repeated measures in (1). As the procedures suggested do not require $d \leq n$, it turns out that the suggested approximation is also valid for very large values of d, i.e. for high dimensional data with an unstructured covariance matrix $\mathbf{S}$. As the procedures are not sensitive to the assumption of the normal distribution, it is briefly discussed that this assumption may be relaxed.

The extension of these ideas to a design with two groups of independent subjects is not straightforward and requires separate considerations which will be briefly indicated in Section 3.

2. The Simple Repeated Measures Model

We start with the simple repeated measures model in (1) and assume that the vectors $\mathbf{Y}_k \sim N(\boldsymbol{\mu}, \mathbf{S})$ are i.i.d. for $k = 1, \ldots, n$, with an unstructured and unknown covariance matrix $\mathbf{S}$, where $d > n$ is also admitted. We want to test the hypothesis $H_0 : \mathbf{H}\boldsymbol{\mu} = \mathbf{0}$ and note that a factorial structure of the repeated measures can be included by imposing a corresponding structure on the index s. In what follows, we will need the symmetric formulation of a hypothesis by using the projection matrix $\mathbf{T} = \mathbf{H}'(\mathbf{H}\mathbf{H}')^{-}\mathbf{H}$, where $(\cdot)^{-}$ denotes some g-inverse. Note that $\mathbf{T}\boldsymbol{\mu} = \mathbf{0} \Longleftrightarrow \mathbf{H}\boldsymbol{\mu} = \mathbf{0}$.

To simplify the notation, let $\mathbf{Z}_k = \mathbf{T}\mathbf{Y}_k$. Then under H_0, $E_{H_0}(\mathbf{Z}_k) = \mathbf{0}$, and $Cov(\mathbf{Z}_k) = \mathbf{T}\mathbf{S}\mathbf{T} = \boldsymbol{\Sigma}$, $k = 1, \ldots, n$.

We note that $\mathrm{tr}(\mathbf{T}\mathbf{S}\mathbf{T}) = \mathrm{tr}(\mathbf{T}\mathbf{S}) = \mathrm{tr}(\boldsymbol{\Sigma})$ since $\mathbf{T}$ is a projection matrix. Let $\bar{\mathbf{Y}}. = \frac{1}{n}\sum_{k=1}^{n}\bar{\mathbf{Y}}_k$ and $\bar{\mathbf{Z}}. = \mathbf{T}\bar{\mathbf{Y}}. = \frac{1}{n}\sum_{k=1}^{n}\mathbf{T}\mathbf{Y}_k$. Then, for testing H_0, we consider the quadratic form $Q_n = (\mathbf{T}\bar{\mathbf{Y}}.)'(\mathbf{T}\bar{\mathbf{Y}}.) = \bar{\mathbf{Z}}.'\,\bar{\mathbf{Z}}.$. As $Cov(\bar{\mathbf{Z}}.) = \frac{1}{n}\boldsymbol{\Sigma}$, one

obtains approximately under H_0

$$F_n = \frac{n \cdot \bar{\mathbf{Z}}_{\cdot}' \, \bar{\mathbf{Z}}_{\cdot}}{\mathrm{tr}(\widehat{\mathbf{\Sigma}}_n)} \stackrel{\cdot}{\sim} F(f, (n-1)f), \quad \text{where} \quad f = \frac{[\mathrm{tr}\mathbf{\Sigma}]^2}{(\mathrm{tr}\mathbf{\Sigma}^2)}. \tag{3}$$

Here,

$$\widehat{\mathbf{\Sigma}}_n = \frac{1}{n-1} \sum_{k=1}^{n} (\mathbf{Z}_k - \bar{\mathbf{Z}}_{\cdot})(\mathbf{Z}_k - \bar{\mathbf{Z}}_{\cdot})' \tag{4}$$

denotes the sample covariance matrix of $\mathbf{\Sigma}$. Then, the unknown quantities $[\mathrm{tr}(\mathbf{\Sigma})]^2$ and $\mathrm{tr}(\mathbf{\Sigma}^2)$ are estimated directly without using $\widehat{\mathbf{\Sigma}}_n$ (Ahmad, Werner, and Brunner, 2008). Let $A_k = \mathbf{Z}_k' \mathbf{Z}_k, \ k = 1, \ldots, n$, and $A_{kl} = \mathbf{Z}_k' \mathbf{Z}_l, \ k \neq l = 1, \ldots, n$, and let further

$$B_1 \ = \ \frac{1}{n(n-1)} \sum_{k \neq l}^{n} A_k A_l \ \text{ and } \ B_2 \ = \ \frac{1}{n(n-1)} \sum_{k \neq l}^{n} A_{kl}^2. \tag{5}$$

Then, under $H_0 : \mathbf{T}\boldsymbol{\mu} = \mathbf{0}$, it follows that

$$E_{H_0}(B_1) \ = \ [\mathrm{tr}(\mathbf{\Sigma})]^2 \quad \text{and} \quad E_{H_0}(B_2) \ = \ \mathrm{tr}(\mathbf{\Sigma}^2)$$
$$Var\left(\frac{B_1}{[\mathrm{tr}(\mathbf{\Sigma})]^2}\right) \leq \frac{8}{n-1} \quad \text{and} \quad Var\left(\frac{B_2}{\mathrm{tr}(\mathbf{\Sigma}^2)}\right) \ \leq \ \frac{8}{n-1}.$$

Thus, one obtains an asymptotically unbiased estimator $\hat{f} = B_1/B_2$ of the degrees of freedom f and finally, $F_n \ \stackrel{\cdot}{\sim} \ F(\hat{f}, (n-1)\hat{f})$. We note that the bias of $\hat{f}$ is smaller than $8/(n-1)$ and therefore uniformly bounded with respect to the dimension d. Simulation studies show that between $0.1 \leq \alpha \leq 0.01$ the pre-assigned level α is maintained quite accurately for $2 \leq d \leq 1000$ even for small $n \geq 10$.

3. The Two-Groups Repeated Measures Design

The model in Section 2 can be generalized to two groups of independent subjects

$$\mathbf{Y}_{ik} = (Y_{ik1}, \ldots, Y_{ikd})' \ \sim \ N(\boldsymbol{\mu}_i, \mathbf{S}_i), \quad i = 1, 2; \ k = 1, \ldots, n_i, \tag{6}$$

where $\mathbf{S}_i = Cov(\mathbf{Y}_{ik})$ and $\boldsymbol{\mu}_i = (\mu_{i1}, \ldots, \mu_{id})' = E(\mathbf{Y}_{ik}), \ \ i = 1, 2$. Further let $N = n_1 + n_2$ denote the total number of all subjects. We do not assume that the two covariance matrices are equal (multivariate Behrens-Fisher problem) and we admit that $d > n_i, \ \ i = 1, 2$ (high-dimensional case).

We want to test the hypothesis $H_0 : \mathbf{H}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2) = \mathbf{0}$. Let $\mathbf{T} = \mathbf{H}'(\mathbf{H}\mathbf{H}')^{-}\mathbf{H}$ similar as in Section 2. Then, $\mathbf{T}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2) = \mathbf{0} \Longleftrightarrow \mathbf{H}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2) = \mathbf{0}$. Further we admit a factorial structure of the repeated measures by imposing an appropriate structure on the index s.

Let $\mathbf{Z}_{ik} = \mathbf{T}\mathbf{Y}_{ik}$ then $E(\mathbf{Z}_{ik}) = \mathbf{T}\boldsymbol{\mu}_i$ and $Cov(\mathbf{Z}_{ik}) = \mathbf{T}\mathbf{S}_i\mathbf{T} = \mathbf{\Sigma}_i, \ i = 1, 2$. For testing H_0 we use the quadratic form $Q_N = (\bar{\mathbf{Z}}_{1\cdot} - \bar{\mathbf{Z}}_{2\cdot})'(\bar{\mathbf{Z}}_{1\cdot} - \bar{\mathbf{Z}}_{2\cdot})$ where

$\bar{\mathbf{Z}}_{i\cdot} = \frac{1}{n_i}\sum_{k=1}^{n_i} \mathbf{T}\mathbf{Y}_{ik}$. Then it follows that $E_{H_0}(\bar{\mathbf{Z}}_{1\cdot} - \bar{\mathbf{Z}}_{2\cdot}) = \mathbf{T}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2) = \mathbf{0}$ and that

$$Cov(\bar{\mathbf{Z}}_{1\cdot} - \bar{\mathbf{Z}}_{2\cdot}) = \mathbf{S}_N \;=\; \frac{1}{n_1}\boldsymbol{\Sigma}_1 + \frac{1}{n_2}\boldsymbol{\Sigma}_2. \tag{7}$$

The unknown covariance matrices $\mathbf{S}_N$, $\boldsymbol{\Sigma}_1$, and $\boldsymbol{\Sigma}_2$ are estimated by the empirical covariance matrices

$$\widehat{\mathbf{S}}_N = \frac{1}{n_1}\widehat{\boldsymbol{\Sigma}}_1 + \frac{1}{n_2}\widehat{\boldsymbol{\Sigma}}_2, \quad \text{and} \tag{8}$$

$$\widehat{\boldsymbol{\Sigma}}_i = \frac{1}{n_i - 1}\sum_{k=1}^{n_i}(\mathbf{Z}_{ik} - \bar{\mathbf{Z}}_{i\cdot})(\mathbf{Z}_{ik} - \bar{\mathbf{Z}}_{i\cdot})'. \tag{9}$$

Finally, one obtains under H_0 that

$$F_N = \frac{(\bar{\mathbf{Z}}_{1\cdot} - \bar{\mathbf{Z}}_{2\cdot})'(\bar{\mathbf{Z}}_{1\cdot} - \bar{\mathbf{Z}}_{2\cdot})}{\mathrm{tr}(\widehat{\mathbf{S}}_N)} \;\dot{\sim}\; F(f, f_0) \quad \text{where} \tag{10}$$

$$f \;=\; \frac{\left[\mathrm{tr}(\mathbf{S}_N)\right]^2}{\mathrm{tr}(\mathbf{S}_N^2)} \quad \text{and} \quad f_0 \;= \frac{\left[\mathrm{tr}(\mathbf{S}_N)\right]^2}{\sum_{i=1}^{2}\mathrm{tr}(\boldsymbol{\Sigma}_i^2)/[n_i^2(n_i-1)]}. \tag{11}$$

To estimate f and f_0 let

$$\mathrm{tr}(\widehat{\mathbf{S}}_N) = \frac{1}{n_1}\mathrm{tr}(\widehat{\boldsymbol{\Sigma}}_1) + \frac{1}{n_2}\mathrm{tr}(\widehat{\boldsymbol{\Sigma}}_2),$$

where $\widehat{\boldsymbol{\Sigma}}_i$ is given in (9). Using (7) it follows that

$$\begin{aligned}
\left[\mathrm{tr}(\mathbf{S}_N)\right]^2 &= \frac{1}{n_1^2}\left[\mathrm{tr}(\widehat{\boldsymbol{\Sigma}}_1)\right]^2 + \frac{1}{n_2^2}\left[\mathrm{tr}(\widehat{\boldsymbol{\Sigma}}_2)\right]^2 + \frac{2}{n_1 n_2}\mathrm{tr}(\widehat{\boldsymbol{\Sigma}}_1)\cdot\mathrm{tr}(\widehat{\boldsymbol{\Sigma}}_2), \\
\mathrm{tr}(\mathbf{S}_N^2) &= \frac{1}{n_1^2}\mathrm{tr}(\widehat{\boldsymbol{\Sigma}}_1^2) + \frac{1}{n_2^2}\mathrm{tr}(\widehat{\boldsymbol{\Sigma}}_2^2) + \frac{2}{n_1 n_2}\mathrm{tr}(\widehat{\boldsymbol{\Sigma}}_1\widehat{\boldsymbol{\Sigma}}_2).
\end{aligned}$$

To estimate the unknown quantities $\left[\mathrm{tr}(\boldsymbol{\Sigma}_i)\right]^2, \mathrm{tr}(\boldsymbol{\Sigma}_i^2), i = 1, 2,\ \mathrm{tr}(\boldsymbol{\Sigma}_1)\cdot\mathrm{tr}(\boldsymbol{\Sigma}_2)$, and $\mathrm{tr}(\boldsymbol{\Sigma}_1\boldsymbol{\Sigma}_2)$, note that

$$\begin{aligned}
E(\mathbf{Z}_{ik} - \mathbf{Z}_{i\ell}) &= \mathbf{T}(\boldsymbol{\mu}_i - \boldsymbol{\mu}_i) \;=\; \mathbf{0}, \\
Cov(\mathbf{Z}_{ik} - \mathbf{Z}_{i\ell}) &= 2\cdot\boldsymbol{\Sigma}_{\mathbf{i}}
\end{aligned} \tag{12}$$

and let

$$\begin{aligned}
A_{kl}^{(i)} &= (\mathbf{Z}_{ik} - \mathbf{Z}_{il})'(\mathbf{Z}_{ik} - \mathbf{Z}_{il}), \quad k \neq l,\ i = 1, 2, \\
A_{klrs}^{(i)} &= (\mathbf{Z}_{ik} - \mathbf{Z}_{il})'(\mathbf{Z}_{ir} - \mathbf{Z}_{is}), \quad k \neq l \neq r \neq s,\ i = 1, 2, \\
A_{klrs}^{(1,2)} &= (\mathbf{Z}_{1k} - \mathbf{Z}_{1l})'(\mathbf{Z}_{2r} - \mathbf{Z}_{2s}), \quad k \neq l \neq r \neq s.
\end{aligned}$$

Also note that $E(A_{kl}^{(i)}) = \mathrm{tr}\left[Cov(\mathbf{Z}_{ik} - \mathbf{Z}_{il})\right] + (\boldsymbol{\mu}_i - \boldsymbol{\mu}_i)'\mathbf{T}(\boldsymbol{\mu}_i - \boldsymbol{\mu}_i) \;=\; 2\mathrm{tr}(\boldsymbol{\Sigma}_i)$. Thus, it follows that

$$\begin{aligned}
E\left(A_{kl}^{(i)} A_{rs}^{(i)}\right) &= \left[\mathrm{tr}(2\boldsymbol{\Sigma}_i)\right]^2 \;=\; 4 \cdot \left[\mathrm{tr}(\boldsymbol{\Sigma}_i)\right]^2,\; i = 1, 2, \\
E\left[(A_{klrs}^{(i)})^2\right] &= 4 \cdot \mathrm{tr}(\boldsymbol{\Sigma}_i^2),\; i = 1, 2, \\
E\left[A_{kl}^{(1)} A_{rs}^{(2)}\right] &= 4 \cdot \mathrm{tr}(\boldsymbol{\Sigma}_1) \cdot \mathrm{tr}(\boldsymbol{\Sigma}_2), \\
E\left[(A_{klrs}^{(1,2)})^2\right] &= 4 \cdot \mathrm{tr}(\boldsymbol{\Sigma}_1 \boldsymbol{\Sigma}_2)
\end{aligned}$$

and one obtains

$$B_1^{(i)} = \frac{1}{n_i(n_i-1)(n_i-2)(n_i-3)} \sum_{k \neq l \neq r \neq s}^{n_i} A_{kl}^{(i)} A_{rs}^{(i)} \quad \text{for } \left[\mathrm{tr}(\boldsymbol{\Sigma}_i)\right]^2,\; i = 1, 2, \tag{13}$$

$$B_2^{(i)} = \frac{1}{n_i(n_i-1)(n_i-2)(n_i-3)} \sum_{k \neq l \neq r \neq s}^{n_i} (A_{klrs}^{(i)})^2 \quad \text{for } \mathrm{tr}(\boldsymbol{\Sigma}_i^2),\; i = 1, 2, \tag{14}$$

$$C_1 = \frac{1}{n_1(n_1-1)n_2(n_2-1)} \sum_{k \neq l}^{n_1} \sum_{r \neq s}^{n_2} A_{kl}^{(1)} A_{rs}^{(2)} \quad \text{for } \mathrm{tr}(\boldsymbol{\Sigma}_1) \cdot \mathrm{tr}(\boldsymbol{\Sigma}_2), \tag{15}$$

$$C_2 = \frac{1}{n_1(n_1-1)n_2(n_2-1)} \sum_{k \neq l}^{n_1} \sum_{r \neq s}^{n_2} \left(A_{klrs}^{(1,2)}\right)^2 \quad \text{for } \mathrm{tr}(\boldsymbol{\Sigma}_1 \boldsymbol{\Sigma}_2). \tag{16}$$

One obtains an unbiased estimator of $\left[\mathrm{tr}(\mathbf{S}_N)\right]^2$ from

$$\widehat{[\mathrm{tr}(\mathbf{S}_N)]^2} = \frac{1}{n_1^2} B_1^{(1)} + \frac{1}{n_2^2} B_1^{(2)} + \frac{2}{n_1 n_2} C_1$$

and of $\mathrm{tr}(\mathbf{S}_N^2)$ from

$$\widehat{\mathrm{tr}(\mathbf{S}_N^2)} = \frac{1}{n_1^2} B_2^{(1)} + \frac{1}{n_2^2} B_2^{(2)} + \frac{2}{n_1 n_2} C_2$$

and finally estimators of f and f_0 in (11) from

$$\hat{f} = \frac{\sum_{i=1}^2 B_1^{(i)}/n_i^2 + 2\; C_1/(n_1 n_2)}{\sum_{i=1}^2 B_2^{(i)}/n_i^2 + 2C_2/(n_1 n_2)} \tag{17}$$

$$\hat{f}_0 = \frac{\sum_{i=1}^2 B_1^{(i)}/n_i^2 + 2C_1/(n_1 n_2)}{\sum_{i=1}^2 B_2^{(i)}/[n_i^2(n_i-1)]}. \tag{18}$$

Simulation studies show that the approximation obtained by this procedure is quite accurate for $n_1, n_2 \geq 10$ for $2 \leq d \leq 1000$. It may be noted that for computational purposes or performing simulations all the above given estimators can easily be obtained by using certain matrix techniques reducing computing time and memory space needed considerably.

References

[1] Ahmad, M. R., Werner, C., and Brunner, E. (2008). Analysis of High Dimensional Repeated Measures Designs: The One Sample Case. *Computational Statistics and Data Analysis* **53**, 416-427.

[2] Box, G. E. P. (1954). Some Theorems on Quadratic Forms Applied in the Study of Analysis of Variance Problems, II. Effects of Inequality of Variance and of Correlation Between Errors in the Two-Way Classification. *Annals of Mathematical Statistics*, **25**, 484-498.

[3] Geisser, S. and Greenhouse, S. W. (1958). An Extension of Box's Result on the Use of the F Distribution in Multivariate Analysis. *Annals of Mathematical Statistics*, **29**, 885-891.

[4] Greenhouse, S.W. and Geisser, S. (1959). On Methods in the Analysis of Profile Data. *Psychometrika*, **24** (2), 95-112.

[5] Huynh, H. and Feldt, L. S. (1970). Conditions Under Which Mean Square Ratios in Repeated Measurement Designs Have Exact F-Distributions. *Journal of the American Statistical Association*, **65**, 1582-1589.

[6] Keselman, H.J. (1998). Testing treatments effects in repeated measures designs: An update for psychological researchers. *Psychophysiology*, **35**, 470-478.

[7] Keselman, H.J., Algina, J., Wilcox, R.R. and Kowalchuk, R.K. (2000). Testing repeated measures hypotheses when covariance matrices are heterogeneous: Revisiting the robustness of the Welch-James test again. *Educational and Psychological Measurements*, **60**, 925-938.

6th St.Petersburg Workshop on Simulation (2009) 611-615

Non-asymptotic estimation of functionals in noise

Boris Darkhovsky[1]

1. Introduction

An estimation of functions and/or functionals by means of noisy data is a traditional problem of mathematical statistics. In particular, the problem of regression functions estimation is of this type. The minimax approach to the problem arises naturally in the case of the regression function belonging to some known class (in mathematical statistics, they call it non-parametric estimation of a regression function). The problem as well as its generalization in the form of the risk minimization theory is thoroughly studied in the fundamental book [7]. The recent book [6] contains a lot of interesting results in the field of non-parametric estimation of functions.

The problems of risk minimization and non-parametric functions estimation are usually considered under the assumption that the noisy parts of observations (Gaussian, as a rule) tend to zero in an appropriate sense. Under such assumptions, it is possible to obtain estimates of the minimax risk and to find the *asymptotically optimal* decision in many cases. The work [10] stands out in this landscape because it gives the best linear non-asymptotic minimax estimate of the value at zero for a polynomial regression function of a given degree and for its derivative, too.

The problem of estimating functionals using incomplete data is also considered in the frame of the deterministic approach. We mean the so called recovery problem (see, for example, [13, 12, 11] and the references therein). One of the most important result in the theory of recovery problems is Smolyak theorem: the minimax estimate of a linear functional in many cases is a linear function of the data. This result does not hold for the stochastic recovery problem.

The approach to non-parametric functions estimation in the context of the stochastic learning theory is actively developed in the recent publications (see, for example, [2, 5, 9]). Based on general results from functions theory, these works show that in many interesting cases there exists an approximation of the regression function such that the probability of its deviation from the true value tends to zero exponentially as the size of the data tends to infinity.

In this paper, we consider the problem of functionals or a collection of functionals estimation using incomplete and noisy data. A non-parametric estimation of a regression function is a particular case of this problem. We propose a new formalization of the problem which, in the case of finite many observations, allows

[1]Institute for Systems Analysis RAS, Moscow, Russia. E-mail: darbor@isa.ru

us to get *non-asymptotically optimal* estimates without losing the substance of the initial problem. The approach was applied to the estimation of a linear functional in [3] and further developed in [4]. The idea of the approach is to introduce a natural probabilistic measure on a finite dimensional set of all possible values of data. The estimation problem is considered as a game of "*Nature*" (N) vs. Statistician (S) such that N chooses a point from the set of all possible values according to the probabilistic measure, whereas S wants to minimize her/his losses. In this paper this idea is extended to a collection of functionals parametrized by an index set.

The paper is organized as follows. In section 2 the stochastic recovery problem in the standard setting is given. In section 3 an informative statement, proposed formalization and basic results are given. In section 4 some examples are considered.

2. Standard setting of the stochastic recovery problem

Let X be a linear normed space, X' its dual, $W \subset X$ a convex central symmetric set, $x' \in X'$, and let $\langle x', x\rangle$ denote the value of a linear functional at an element x. Let Y be a linear normed space, $A : X \to Y$ a linear operator, and ξ is a Y-valued random element. Consider the following problem

$$R(\phi) \equiv \sup_{x\in\mathfrak{M}} \mathbf{E}\big|\langle x', x\rangle - \phi(Ax+\xi)\big| \to \inf_{\phi\in M}$$

Here $\mathbf{E}$ is a symbol of mathematical expectation and the infimum is over some set M of functionals on Y.

Let Φ be a set of all uniformly continuous functionals over Y, Y^* be the topological dual of Y and

$$R \stackrel{\text{def}}{=} \inf_{\phi\in\Phi} R(\phi), \quad R^* \stackrel{\text{def}}{=} \inf_{y^*\in Y^*} R(y^*)$$

Let $\tilde{y}^*$ be such that

$$\sup_{x\in\mathfrak{M}} \big|\langle x', x\rangle - \langle \tilde{y}^*, Ax\rangle\big| = \inf_{y^*\in Y^*} \sup_{x\in\mathfrak{M}} \big|\langle x', x\rangle - \langle y^*, Ax\rangle\big|$$

Theorem 1. *[3] The following estimate holds*

$$R^* - \parallel \tilde{y}^* \parallel \mathbf{E} \parallel \xi \parallel \le R \le R^*$$

3. New setting of the stochastic recovery problem

3.1. Stochastic recovery problem for a single functional

Let X, Y be linear normed spaces, $A : X \to Y$, $\varphi : X \to \mathbb{R}$ be given operators, $\mathfrak{M} \subset X$ be given set. Let $\Theta \stackrel{\text{def}}{=} \mathrm{Im}A(\mathfrak{M}) = \{y \in Y : \exists\, x \in \mathfrak{M}\,\text{such that}\, A(x) = y\}$.

The following problem we call a *special recovery problem*(SRP):

$$\forall\theta\in\Theta\quad \Phi(\theta,z)=\sup_{x\in\mathfrak{M},\ A(x)=\theta}|\varphi(x)-z|\longrightarrow\inf_{z\in\mathbb{R}}. \tag{1}$$

Suppose that the operator A in (1) has a *finite-dimensional* image, that is $\Theta\subset\mathbb{R}^k$. Let the collection of finite-dimensional vectors $Y=\{y_i\}$, $y_i=\theta+\xi_i$, $i=1,\ldots,n$ represents (for the simplest observation scheme) the available information after n observations (as it's usual for statistical problems we suppose that it's possible to do several measurements of the unknown vector $\theta\in\Theta$). Suppose that the collection of random vectors $(\xi_1,\ldots,\xi_n)$ has the density $p(\cdot)$ with respect to Lebesgue measure. Then, one can introduce on a finite dimensional space $\mathbb{R}^k$ a probability measure $\mathbf{P}_Y$ with the density

$$q(\theta\big|Y)=\left\{\begin{array}{ll}\dfrac{p(y_1-\theta,\ldots,y_n-\theta)}{\int\limits_{\Theta}p(y_1-\theta,\ldots,y_n-\theta)d\theta}, & \text{if } \theta\in\Theta,\\ 0, & \text{if } \theta\notin\Theta.\end{array}\right) \tag{2}$$

The density $q(\theta\big|Y)$ could be considered as a certain "*degree of confidence*" of a statistician to every point of the image $\mathrm{Im}A(\mathfrak{M})$ upon obtaining the information Y containing a random error.

Another possible interpretation is as follows. Given a bounded set Θ and a random vector $\theta\in\Theta$ with prior uniform distribution over Θ, the density $q(\theta\big|Y)$ is nothing but posterior density θ under condition that observation Y was obtained.

For an unbounded Θ one cannot formally speak of a uniform distribution, whereas an expression for $q(\theta\big|Y)$ still retains the meaning of posterior density. The assumption of uniformity of the vector θ prior distribution reflects the fact that *no probability measure is defined* over $\mathfrak{M}$, and so the appearance of various points of the image $\mathrm{Im}A\,(\mathfrak{M})$ is, so to say, "equally possible".

Consider on the real line for each vector $\theta\in\Theta$ the set $K(\theta)=\{v\in\mathbb{R}:\varphi(x)=v,\ A(x)=\theta,\ x(\cdot)\in\mathfrak{M}\}$ and suppose that it is bounded. This is the collection of all possible values of the functional $\varphi(\cdot)$ under the condition that the information image of the set $\mathfrak{M}$ coincides with the point θ under the map A.

Consider the following game with the "$Nature$": the "$Nature$" chooses a vector $\theta\in\Theta$ with a probability determined by the density $q(\theta|Y)$ and the statistician has to choose some number u as an estimate of the value of the functional $\varphi(x)$.

As under $\theta\in\Theta$ the true values of the functional being estimated are contained in the set $K(\theta)$, then a natural way to estimate the loss from the decision is to use some function $h(\cdot)$ of the distance between the chosen number and the points of the set $K(\theta)$, that is, $\sup_{t\in K(\theta)}h(|u-t|)$. We assume that this value is a measurable and bounded function of θ for each u (such assumption is usually true in practice). Since a probability measure has been introduced on the set Θ, $\sup_{t\in K(\theta)}h(|u-t|)$ is a random variable and a natural statement of the problem is as follows:

$$\mathbf{E}\left(\sup_{t\in K(\theta)}h(|\ u-t\ |)\Big|Y\right)\longrightarrow\inf_{u\in\mathbb{R}}, \tag{3}$$

where $\mathbf{E}(\cdot|Y) \stackrel{\text{def}}{=} \mathbf{E}_Y(\cdot)$ denote the mean over the density $q(\cdot \,|\, Y)$.

Put

$$\alpha(\theta) \stackrel{\text{def}}{=} \sup_{x\in\mathfrak{M},\, A(x)=\theta} \varphi(x), \quad \beta(\theta) \stackrel{\text{def}}{=} \inf_{x\in\mathfrak{M},\, A(x)=\theta} \varphi(x)$$
$$m(\theta) = 1/2\,(\alpha(\theta) + \beta(\theta))\,,\ r(\theta) = 1/2\,(\alpha(\theta) - \beta(\theta))$$
$$H_Y(v) = \int\limits_{\{\theta\in\Theta: m(\theta)\geq v\}} r(\theta) q(\theta|Y) d\theta,\ M_Y = \mathbf{E}_Y m(\theta),\ R_Y = \mathbf{E}_Y r(\theta)$$

and consider the equation

$$2H_Y(v) - v = R_Y - M_Y \tag{4}$$

Theorem 2. *Let $\alpha(\theta)$ and $\beta(\theta)$ are bounded and measurable functions and for any $Y \in \mathbb{R}^{kn}$ the functions $m(\theta)$, $r(\theta)$ are square integrable w.r.t. measure $q(\theta|Y)d\theta$. Let the observations are: $y_i = \theta + \xi_i$, $i = 1, \ldots, n$, $\theta \in \Theta = \mathrm{Im}A\,(\mathfrak{M})\,$, $Y = \{y_i\}$, where vectors ξ_i are random and there exists a density w. r. t. Lebesgue measure for the collection $\{\xi_i\}_{i=1}^n$. Then:*

i) the optimal(non-asymptotic) solution $u^*(Y)$ of problem () under the loss function $h(t) = t$ is the "posterior" median of the random variable $m(\theta) = \frac{1}{2}\,(\alpha(\theta) + \beta(\theta))$.

ii)the optimal(non-asymptotic) solution $u^*(Y)$ of problem () under the loss function $h(t) = t^2$ is the (unique) root of the equation (4).

3.2. Stochastic recovery problem for a collection of functionals

It is possible to consider a collection of functionals, namely $\Phi = \{\varphi_t(x)\}_{t\in T}, x \in \mathfrak{M}$, where T be given collection of indices. Then the results can be generalized for this situation under natural assumptions.

Problem 0.

$$\mathbf{E}_Y W_0\,(u(\cdot), \theta) \longrightarrow \inf_{u(t)\in U},$$

where U is the set of all measurable functions on T.

Problem 1.

$$\mathbf{E}_Y W_1\,(u(\cdot), \theta) \longrightarrow \inf_{u(t)\in L_1(\mu)},$$

where $L_1(\mu)$ is the set of all μ-integrable functions on T.

Problem 2.

$$\sqrt{\mathbf{E}_Y W_2\,(u(\cdot), \theta)} \longrightarrow \inf_{u(t)\in L_2(\mu)},$$

where $L_2(\mu)$ is the set of all μ-square-integrable functions on T.

Problem 0

Theorem 3. *Let assumptions (A.1), (A.2) hold. Then the following inequalities hold for the value $\mathbb{L}$ of the Problem 0:*

$$\sup_{t\in T} \mathbf{E}_Y\,(r(t,\theta) + |m(t,\theta) - \mathbb{M}(t,Y)|) \leq \mathbb{L} \leq \mathbf{E}_Y \sup_{t\in T}\,(r(t,\theta) + |m(t,\theta) - \mathbb{M}(t,Y)|)\,,$$

where $\mathbb{M}(t,Y)$ *is the median of the random variable* $m(t,\theta)$ *over the distribution* $\mathbf{P}_Y$.

In particular, if $T=\{t_0\}$ *then the optimal solution* $u_0^*(t_0,Y)$ *of Problem 0 is the median of process* $m(\cdot,\theta)$, $\quad u_0^*(t_0,Y)=\mathbb{M}(t_0,Y)$.

Problem 1

Theorem 4. *Let assumptions (A.1)–(A.4) hold. The optimal solution* $u_1^*(t,Y)$ *of problem 1 is equal to* $u_0^*(t,Y)$, *i.e. to the median of process* $m(\cdot,\theta)$, $u_1^*(t,Y)=\mathbb{M}(t,Y)$.

Problem 2

Theorem 5. *Let assumptions (A.1)-(A.4) hold. The optimal solution* $u_2^*(t,Y)$ *of problem 2 is given by the unique root of equation (**??**).*

3.3. Asymptotic of optimal solutions

In previous subsections we found *non-asymptotic* solutions of problems 0,1,2. We can also consider these problems in deterministic statement, i.e. for the case $\xi_i\equiv 0$, $Y=y_1\stackrel{\text{def}}{=}y=\theta$. Obviously, in deterministic case the optimal solution for each problem is $m(t,\theta)$. It is interesting to investigate an asymptotic of the stochastic solutions as the number of observations n tends to infinity. Here we formulate the corresponding theorem.

Assume that $\{\xi_i\}$ are independent and identically distributed random vectors with the d.f. $f(\cdot)$. We assume that the density satisfies the standard regularity conditions (see [1]) in some neighborhood of Θ.

Let $\mathbf{P}_{\theta^*}(\mathbf{E}_{\theta^*})$ be a probability measure (mean) on a sample space corresponding to the true (unknown) value of the parameter $\theta^*\in\Theta$. Let us introduce the following function (the Kullback–Leibler information)

$$\mathbb{K}(\theta^*,\theta)=\int\left(\log\frac{f(y-\theta^*)}{f(y-\theta)}\right)f(y-\theta^*)\,dy.$$

Theorem 6. *Suppose that the following conditions hold:*

1) *the set* $\Theta=\mathrm{Im}A(\mathfrak{M})$ *is a compact;*

2) *the random vectors* $\{\xi_i\}$ *are independent, identically distributed with d.f.* $f(\cdot)$ *(w.r.t. Lebesgue measure) that satisfies the standard regularity conditions;*

3) *the Kullback–Leibler information* $\mathbb{K}(\theta^*,\theta)$ *is twice continuously differentiable with respect to* θ *for every* $\theta^*,\theta\in\Theta$*;*

4) *the optimal solution of the deterministic problems* $m(t,\theta)$ *satisfies for every* $t\in T$ *Lipschitz condition w.r.t.* θ, $\theta\in\Theta$.

Then

(a) *for every interior point* $\theta^*\in\Theta$ *any finite collection* $(u_i^*(t_1,y),\dots,u_i^*(t_s,y))$ *of the optimal solutions* $u_i^*(t,y), i=0,1,2$ *converges* $\mathbf{P}_{\theta^*}$*-a.s. to* $(m(t_1,\theta^*),\dots,m(t_s,\theta^*))$ *as* $n\to\infty$*;*

(c) *the following estimate holds:*

$$\sup_{\theta^*\in\Theta}\mathbf{E}_{\theta^*}\left|u_i^*(t,y)-m(t,\theta^*)\right|\le O\big(\sqrt{n^{-1}\log n}\big),\ i=0,1,2.$$

4. Examples

In this section we consider the estimation (in different cases) of the following functionals: 1)a function value at given point, 2) a function derivative at given point, 3) an integral of a function over given segment 4)a maximum of a function on a given segment.

The optimal deterministic solutions will be given. According to the previous sections, to get the optimal stochastic solution one has to construct the "posterior" distribution and to calculate (for example) the median of the deterministic solution w.r.t. the "posterior" distribution.

References

[1] Borovkov, A.A., Mathematical Statistics (Gordon and Breach, Amsterdam, 1990).

[2] Cucker, F., Smale, S., 2001. On mathematical foundations of learning. Bulletin of AMS, 39, 1-49.

[3] Darkhovsky, B.S., 1998. On a stochastic recovery problem. Theory Probab. Appl., 43, 282–288.

[4] Darkhovsky, B.S., 2004. A new approach to the stochastic recovery problem. Theory Probab. Appl., 49, 51–64.

[5] DeVore, R., Kerkyacharian, G., Picard, D, Temlyakov, V., 2006. Approximation methods for supervised learning. Found. Comput. Math., 6, no. 1, 3-38.

[6] Efromovich, S., 1999. Nonparametric Curve Estimation, Methods, Theory, and Applications. Springer-Verlag, New York.

[7] Ibragimov,I.A., Khasminskii, R.Z., 1981. Statistical Estimation: Asymptotic Theory. Springer-Verlag, New York.

[8] Ioffe, A. D., and Tikhomirov, V. M., *Theory of Extremal Problems*, New York: North-Holland, 1979.

[9] Konyagin, S.V., Lifshits, E.D., 2008. On adaptive estimators in statistical learning theory. Steklov Trudi,

[10] Legostaeva, I.L., Shiryaev, A.N., 1971. Minimax weights in a trend detection problem of a random process. Theory Probab. Appl., 16, 344–349.

[11] Magaril-Il'yaev, G.G., Tikhomirov, V.M., 2003. Convex Analysis: Theory and Applications, AMS, Providence, RI.

[12] Micchelli, C. A., Rivlin, T. J., 1977. A survey of optimal recovery. In: Optimal Estimation in Approximation Theory, Micchelli, C. A., Rivlin, T.J. eds., Plenum, New York, 1–54.

[13] Traub, J., Wozniakowski,H., 1980. A General Theory of Optimal Algorithms. Academic Press, New York.

6th St.Petersburg Workshop on Simulation (2009) 617-621

Using Universal Source Coding for Statistical Analysis of Time Series[1]

Boris Ryabko[2]

Abstract

We show how universal codes can be used for solving some of the most important statistical problems for time series. By definition, a universal code (or a universal lossless data compressor) can compress any sequence generated by a stationary and ergodic source asymptotically to the Shannon entropy, which, in turn, is the best achievable ratio for lossless data compressors.

We consider finite-alphabet and real-valued time series and the following problems: estimation of the limiting probabilities for finite-alphabet time series and estimation of the density for real-valued time series, the on-line prediction, for both types of the time series and the following problems of hypothesis testing: goodness-of-fit testing, or identity testing, and testing of serial independence. It is important to note that all problems are considered in the framework of classical mathematical statistics and, on the other hand, everyday methods of data compression (or archivers) can be used as a tool for the estimation and testing. It turns out, that quite often the suggested methods and tests are more powerful than other known ones when they are applied in practice.

1. Introduction

In this report we describe and develop a new approach to estimation, prediction and hypothesis testing for time series, which was suggested recently [6, 7, 9]. This approach is based on ideas of universal coding (or universal data compression) and has shown a high efficiency. In particular, this approach was applied to the problem of randomness testing [9]. This problem is quite important for practice [4, 5]; for example, the National Institute of Standards and Technology of USA (NIST)has suggested "A statistical test suite for random and pseudorandom number generators for cryptographic applications" [5], which consists of 16 tests. It

[1]This work was supported by the Russian Foundation of Basic Research; grant no. 06-07-89025.

[2]Siberian State University of Telecommunications and Informatics and Institute of Computational Technology of Siberian Branch of Russian Academy of Science , E-mail: boris@ryabko.net

has turned out that tests which are based on universal codes are more powerful than the tests suggested by NIST [9].

We consider finite-alphabet and real-valued time series and the nonparametric estimation of the limiting probabilities for finite-alphabet time series and the estimation of the density for real-valued time series, the on-line prediction for both types of the time series and the nonparametric goodness-of-fit testing and the testing of serial independence.

We would like to emphasize that everyday methods of data compression (or archivers) can be directly used as a tool for estimation, prediction and hypothesis testing. It is important to note that the modern archivers (like *zip*, *arj, rar,* etc.) are based on deep theoretical results of the source coding theory and have shown their high efficiency in practice because archivers can find many kinds of latent regularities and use them for compression.

Proofs of all theorems can be found here: http://arxiv.org/abs/0809.1226

2. Finite Alphabet Processes

We will consider stationary ergodic processes generating letters from a finite alphabet A.

First we briefly describe lossless codes, or methods of (lossless) data compression. A data compression method (or code) φ is defined as a set of mappings φ_n such that $\varphi_n : A^n \to \{0,1\}^*$, $n = 1, 2, \dots$ and for each pair of different words $x, y \in A^n$ $\varphi_n(x) \neq \varphi_n(y)$. It is also required that each sequence $\varphi_n(u_1)\varphi_n(u_2)...\varphi_n(u_r)$, $r \geq 1$, of encoded words from the set $A^n, n \geq 1$, could be uniquely decoded into $u_1u_2...u_r$. Such codes are called uniquely decodable. For example, let $A = \{a, b\}$, the code $\psi_1(a) = 0, \psi_1(b) = 00$, obviously, is not uniquely decodable. In what follows we call uniquely decodable codes just "codes".

There exist so-called universal codes. For their description we recall that sequences $x_1 \dots x_t$, generated by a source P, can be "compressed" to the length $-\log P(x_1...x_t)$ bits and, on the other hand, for any source P there is no code ψ for which the average codeword length ($\Sigma_{u \in A^t} P(u)|\psi(u)|$) is less than $-\Sigma_{u \in A^t} P(u) \log P(u)$. Universal $-\log P(x_1...x_t)$ asymptotically for any stationary and ergodic source P in average and with probability 1. The formal definition is as follows: a code U is universal if for any stationary and ergodic source P the following equalities are valid:

$$\lim_{t \to \infty} |U(x_1 \dots x_t)|/t = h_\infty(P)$$

with probability 1, and $\lim_{t \to \infty} E(|U(x_1 \dots x_t)|)/t = h_\infty(P)$, where $E(f)$ is the expected value of f, $h_\infty(P)$ is the Shannon entropy of P; see for definition, for example, [2, 3]. So, informally speaking, a universal code estimates the probability characteristics of a source and uses them for efficient "compression".

The following theorem shows how universal codes can be applied for probability estimation.

Theorem 1. *Let U be a universal code and*

$$\mu_U(u) = 2^{-|U(u)|}/\Sigma_{v\in A^{|u|}}\, 2^{-|U(v)|}. \tag{1}$$

Then, for any stationary and ergodic source P the following equalities are valid:

$$\lim_{t\to\infty}\frac{1}{t}(-\log P(x_1\cdots x_t) - (-\log \mu_U(x_1\cdots x_t))) \;=\; 0$$

with probability 1 and $\lim_{t\to\infty}\frac{1}{t}\sum_{u\in A^t} P(u)\log(P(u)/\mu_U(u)) \;=\; 0.$

An informal outline of the proof ia as follows: $\frac{1}{t}(-\log P(x_1\cdots x_t)$ and $\frac{1}{t}(-\log \mu_U(x_1\cdots x_t))$ go to the Shannon entropy $h_\infty(P)$, that is why the difference is 0.

As we mentioned above, any universal code U can be applied for prediction. Namely, the measure μ_U (1) can be used for prediction as the following conditional probability: $\mu_U(x_{t+1}|x_1...x_t) = \mu_U(x_1...x_tx_{t+1})/\mu_U(x_1...x_t)$. The following theorem shows that such a predictor is quite reasonable. Moreover, it gives a possibility to apply practically used data compressors for prediction of real

data (like EUR/USD rate) and obtain quite precise estimation [8] .

Theorem 2. *Let U be a universal code and P be any stationary and ergodic process. Then*

$$i)\ \lim_{t\to\infty}\frac{1}{t}\ E\{\log\frac{P(x_1)}{\mu_U(x_1)} + \log\frac{P(x_2|x_1)}{\mu_U(x_2|x_1)} + \ldots + \log\frac{P(x_t|x_1...x_{t-1})}{\mu_U(x_t|x_1...x_{t-1})}\} \;=\; 0,$$

$$ii)\ \lim_{t\to\infty} E(\frac{1}{t}\ \sum_{i=0}^{t-1}(P(x_{i+1}|x_1...x_i) - \mu_U(x_{i+1}|x_1...x_i))^2) \;=\; 0\,,$$

and iii) $\lim_{t\to\infty} E(\frac{1}{t}\ \sum_{i=0}^{t-1}|P(x_{i+1}|x_1...x_i) - \mu_U(x_{i+1}|x_1...x_i)|) \;=\; 0\,.$

An informal outline of the proof is as follows:

$$\frac{1}{t}\ \{E(\log\frac{P(x_1)}{\mu_U(x_1)}) + E(\log\frac{P(x_2|x_1)}{\mu_U(x_2|x_1)}) + \ldots + E(\log\frac{P(x_t|x_1...x_{t-1})}{\mu_U(x_t|x_1...x_{t-1})})\}$$

is equal to $\frac{1}{t}E(\log\frac{P(x_1...x_t)}{\mu_U(x_1...x_t)})$. Taking into account Theorem 1, we obtain the first statement of the theorem.

3. Hypothesis Testing

Let the hypothesis H_0^{id} be that the source has a particular distribution π and the alternative hypothesis H_1^{id} that the sequence is generated by a stationary and ergodic source which differs from the source under H_0^{id}. We consider the problem of testing H_0^{id} against H_1^{id}. Let the required level of significance (or the Type I error) be α, $\alpha \in (0,1)$. We describe a statistical test which can be constructed based on any code φ.

The main idea of the suggested test is quite natural: compress a sample sequence $x_1...x_t$ by a code φ. If the length of the codeword ($|\varphi(x_1...x_t)|$) is significantly less than the value $-\log\pi(x_1...x_t)$, then H_0^{id} should be rejected. The key observation is that the probability of all rejected sequences is quite small for any φ, that is why the Type I error can be made small. The precise description of the test is as follows: *The hypothesis H_0^{id} is accepted if*

$$-\log\pi(x_1...x_t) - |\varphi(x_1...x_t)| \leq -\log\alpha. \tag{2}$$

Otherwise, H_0^{id} is rejected. We denote this test $T_\varphi^{id}(A,\alpha)$.

Theorem 3. *i) For each distribution $\pi, \alpha \in (0,1)$ and a code φ, the Type I error of the described test $T_\varphi^{id}(A,\alpha)$ is not larger than α and ii) if, in addition, π is a finite-order stationary and ergodic process over A^∞ (i.e. $\pi \in M^*(A)$) and φ is a universal code, then the Type II error of the test $T_\varphi^{id}(A,\alpha)$ goes to 0, when t tends to infinity.*

Let us recall that the null hypothesis H_0^{SI} is that the source is Markovian of order not larger than m, ($m \geq 0$), and the alternative hypothesis H_1^{SI} is that the sequence is generated by a stationary and ergodic source which differs from the source under H_0^{SI}. In particular, if $m = 0$, this is the problem of testing for independence of time series.

Let there be given a sample $x_1...x_t$ generated by an (unknown) source π. The main hypothesis H_0^{SI} is that the source π is Markovian whose order is not greater than m, ($m \geq 0$), and the alternative hypothesis H_1^{SI} is that the sequence is generated by a stationary and ergodic source which differs from the source under H_0^{SI}. The described test is as follows.

Let φ be any code. By definition, the hypothesis H_0^{SI} is accepted if

$$(t-m)\, h_m^*(x_1...x_t) - |\varphi(x_1...x_t)| \leq \log(1/\alpha)\,, \tag{3}$$

where $\alpha \in (0,1)$. Otherwise, H_0^{SI} is rejected. We denote this test by $T_\varphi^{SI}(A,\alpha)$.

Theorem 4. *i) For any code φ the Type I error of the test $T_\varphi^{SI}(A,\alpha)$ is less than or equal to $\alpha, \alpha \in (0,1)$ and, ii) if, in addition, φ is a universal code, then the Type II error of the test $T_\varphi^{SI}(A,\alpha)$ goes to 0, when t tends to infinity.*

4. Real-Valued Time Series

Here we address the problem of nonparametric estimation of the density for time series. Let X_t be a time series and the probability distribution of X_t is unknown, but it is known that the time series is stationary and ergodic. In this part we will use the generalization of the Shannon-MacMillan-Breiman theorem to the processes with densities, which was established by Barron [1]. First we describe considered processes with some properties needed for the generalized Shannon-MacMillan-Breiman theorem to hold. In what follows, we restrict our attention to processes that take bounded real values. However, the main results may be

extended to processes taking values in a compact subset of a separable metric space.

Let B denote the Borel subsets of R, and B^k denote the Borel subsets of R^k, where R is the set of real numbers. Let R^∞ be the set of all infinite sequences $x = x_1, x_2 \ldots$ with $x_i \in \mathsf{R}$, and let B^∞ denote the usual product sigma field on R^∞, generated by the finite dimensional cylinder sets $\{A_1, \ldots A_k, \mathsf{R}, \mathsf{R}, \ldots\}$, where $A_i \in B, i = 1, \ldots, k$. Each stochastic process $X_1, X_2, \ldots, X_i \in \mathsf{R}$, is defined by a probability distribution on $(\mathsf{R}^\infty, B^\infty)$. Suppose that the joint distribution P_n for $(X_1, X_2, \ldots, X_n)$ has a probability density function $p(x_1 x_2 \ldots x_n)$ with respect to a sigma-finite measure M_n. Assume that the sequence of dominating measures M_n is Markov of order $m \geq 0$ with a stationary transition measure. A familiar case for M_n is Lebesgue measure. Let $p(x_{n+1}|x_1 \ldots x_n)$ denote the conditional density given by the ratio $p(x_1 \ldots x_{n+1})\ /p(x_1 \ldots x_n)$ for $n > 1$. It is known that for stationary and ergodic processes there exists a so- called relative entropy rate $\tilde{h}$ defined by $\tilde{h} = \lim_{n\to\infty} -E(\log p(x_{n+1}|x_1 \ldots x_n))$, where E denotes expectation with respect to P. We will use the following generalization of the Shannon–MacMillan–Breiman theorem:

Claim 1 ([1]). *If $\{X_n\}$ is a $P-$stationary ergodic process with density $p(x_1 \ldots x_n) = dP_n/dM_n$ and $\tilde{h}_n < \infty$ for some $n \geq m$, the sequence of relative entropy densities $-(1/n)\log p(x_1 \ldots x_n)$ convergence almost surely to the relative entropy rate, i.e., $\lim_{n\to\infty}(-1/n)\log p(x_1 \ldots x_n) = \tilde{h}$ with probability 1 (according to P).*

Now we return to the estimation problems. Let $\{\Pi_n\}, n \geq 1$, be an increasing sequence of finite partitions of R that asymptotically generates the Borel sigma-field B and let $x^{[k]}$ denote the element of Π_k that contains the point x. (Informally, $x^{[k]}$ is obtained by quantizing x to k bits of precision.) For integers s and n we define the following approximation of the density $p^s(x_1 \ldots x_n) = P(x_1^{[s]} \ldots x_n^{[s]})/M_n(x_1^{[s]} \ldots x_n^{[s]})$. We also consider $\tilde{h}_s = \lim_{n\to\infty} -E(\log p^s(x_{n+1}|\ x_1 \ldots x_n))$. Applying Claim 1 to the density $p^s(x_1 \ldots x_t)$, we obtain that a.s. $\lim_{t\to\infty} -\frac{1}{t}\log p^s(x_1 \ldots x_t) = \tilde{h}_s$. Let U be a universal code, which is defined for any finite alphabet. In order to describe a density estimate we will use the probability distribution $\{\omega = \omega_1, \omega_2, ...\}$ on integers $\{1, 2, ...\}$ by $\omega_1 = 1 - 1/\log 3$, ... , $\omega_i = 1/\log(i+1) - 1/\log(i+2)$, ... (In what follows we will use this distribution, but results described below are obviously true for any distribution with nonzero probabilities.) Now we can define the density estimate r_U as follows:

$$r_U(x_1 \ldots x_t) = \sum_{i=0}^{\infty} \omega_i\ \mu_U(x_1^{[i]} \ldots x_t^{[i]})/M_t(x_1^{[i]} \ldots x_t^{[i]})\ ,$$

where the measure μ_U is defined by (1). (It is assumed here that the code $U(x_1^{[i]} \ldots x_t^{[i]})$ is defined for the alphabet, which contains $|\Pi_i|$ letters.)

It turns out that, in a certain sense, the density $r_U(x_1 \ldots x_t)$ estimates the unknown density $p(x_1 \ldots x_t)$.

Theorem 5. *Let X_t be a stationary ergodic process with densities $p(x_1 \ldots x_t) = dP_t/dM_t$ such that $\lim_{s\to\infty} \tilde{h}_s = \tilde{h} < \infty$. Then $\lim_{t\to\infty} \frac{1}{t}\log \frac{p(x_1 \ldots x_t)}{r_U(x_1 \ldots x_t)} = 0$ with probability 1 and $\lim_{t\to\infty} \frac{1}{t}\ E(\log \frac{p(x_1 \ldots x_t)}{r_U(x_1 \ldots x_t)})\ =\ 0$.*

The following theorem are devoted to the conditional probability $r_U(x|x_1...x_m) = r_U(x_1...x_m x)/r_U(x_1...x_m)$. We will see that the conditional density $r_U(x|x_1...x_m)$ is a reasonable estimation of the unknown density $p(x|x_1...x_m)$.

Theorem 6. *Let $B_1, B_2, ...$ be a sequence of measurable sets. Then the following equalities are true:*

$$i)\ \lim_{t\to\infty} E(\frac{1}{t}\sum_{m=0}^{t-1}(P(x_{m+1}\in B_{m+1}|x_1...x_m) - R_U(x_{m+1}\in B_{m+1}|x_1...x_m))^2) = 0\,,$$

$$ii)\ E(\frac{1}{t}\sum_{m=0}^{t-1}|P(x_{m+1}\in B_{m+1}|x_1...x_m) - R_U(x_{m+1}\in B_{m+1}|x_1...x_m))| = 0\,,$$

where $R_U(x_{m+1}\in B_{m+1}|x_1...x_m) = \int_{B_{m+1}} r_U(x|x_1...x_m)dM_{1/m}$

References

[1] Barron A.R. (1985) The strong ergodic theorem for dencities: generalized Shannon-McMillan-Breiman theorem, *The annals of Probability*, **13** (4), 1292–1303.

[2] Gallager R. G. (1968) *Information Theory and Reliable Communication.* John Wiley & Sons, New York.

[3] Krichevsky R. (1993) *Universal Compression and Retrival.* Kluver Academic Publishers.

[4] L'Ecuyer P., Simard R. J. (2007) TestU01: A C library for empirical testing of random number generators. *ACM Transactions on Mathematical Software (TOMS),* **33**, (4).

[5] Rukhin A. and others. (2001) *A statistical test suite for random and pseudorandom number generators for cryptographic applications.* (NIST Special Publication 800-22 (with revision dated May,15,2001)). http://csrc.nist.gov/rng/SP800-22b.pdf

[6] Ryabko B. Ya. (1988) Prediction of random sequences and universal coding. *Problems of Inform. Transmission*, **24**(2), 87-96.

[7] Ryabko B., Astola J.,Gammerman A. (2006) Application of Kolmogorov complexity and universal codes to identity testing and nonparametric testing of serial independence for time series, *Theoretical Computer Science,* **359**, 440-448.

[8] Ryabko B., Monarev V. (2005) Experimental Investigation of Forecasting Methods Based on Data Compression Algorithms. *Problems of Information Transmission,* **41**(1), 65-69.

[9] Ryabko B., Monarev V. (2005) Using Information Theory Approach to Randomness Testing, *Journal of Statistical Planning and Inference,* **133**(1), 95–110.

6th St.Petersburg Workshop on Simulation (2009) 623-627

Linear and Nonlinear Approximations of the Ratio of the Standard Normal Density and Distribution Functions for the Estimation of the Skew Normal Shape Parameter

Subir Ghosh[1] and Debarshi Dey

Abstract

We introduce a linear approximation and a nonlinear approximation of the ratio of the standard normal density and distribution functions in presence of an unknown constant representing the shape parameter of the skew normal distribution. The purpose of these approximations is to estimate the skew normal shape parameter. We present a new estimation method of the shape parameter based on these approximations. The simulation results demonstrate that the approximations strongly resemble their true values in the regions of interest and the estimated biases of the shape parameter are small.

Key words and phrases: Likelihood, Linear approximation, Nonlinear approximation, Skew normal, Standard normal.

1. Introduction

We consider the ratio R of the standard normal $N(0,1)$ density function ϕ and the distribution function Φ as

$$R_\lambda(z) = \frac{\phi(\lambda z)}{\Phi(\lambda z)}, \tag{1}$$

where λ is a fixed unknown constant. The numerical value of $R_\lambda(0)$ is 0.7978846. Figure 1 is showing the graphs of $R_\lambda(z)$ against z for $\lambda = \pm 2, \pm 1$, and ± 0.5. The graphs in Figure 1 intersect at $z = 0$.

For a given λ, we want to approximate $R_\lambda(z)$ by a linear function and a nonlinear function for $-3 \leq z \leq 3$. The approximating functions are

$$A_\lambda(z) = 0.7978846 + \lambda\alpha z, \tag{2}$$

$$B_\lambda(z) = 0.7978846 + \lambda\beta(e^{-\delta z} - 1), \tag{3}$$

[1]Department of Statistics, University of California, Riverside, CA 92521,USA. E-mail: subir.ghosh@ucr.edu

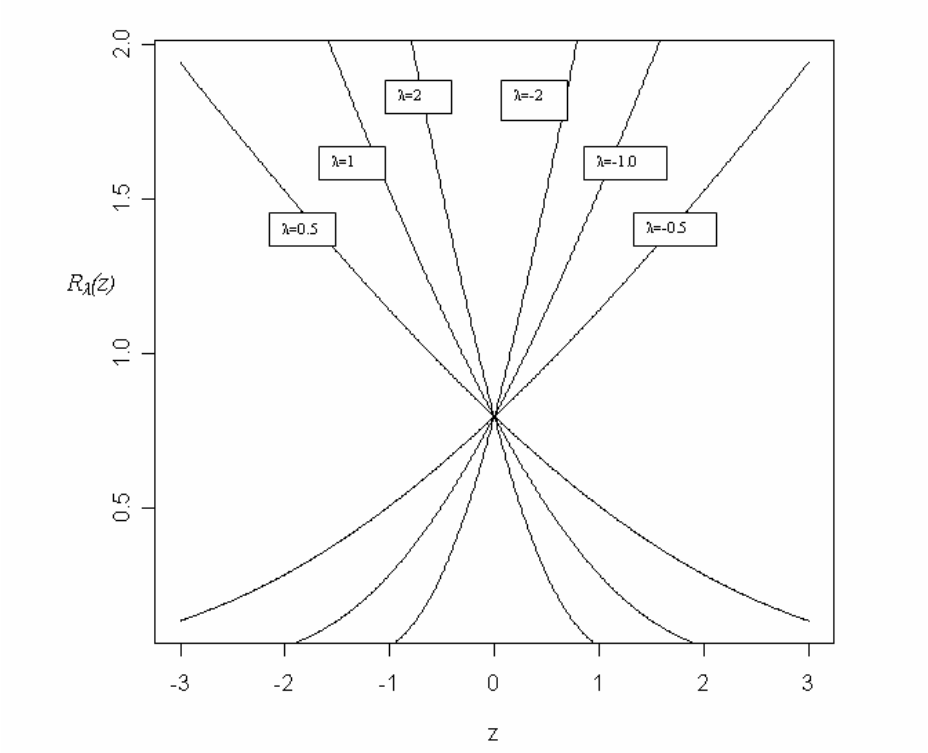


Figure 1: Plots of $R_\lambda(z)$ against z for $\lambda = \pm 2,\ \pm 1$, and ± 0.5.

where α and β are unknown constants and

$$\delta = \begin{cases} +1, & \text{if } \lambda \geq 0; \\ -1, & \text{if } \lambda \leq 0. \end{cases}$$

Figures 2 and 3 are counterparts of Figure 1 for $A_\lambda(z)$ and $B_\lambda(z)$, respectively for the given values of λ and $\lambda\alpha$ in Figure 2 and the given values of λ and $\lambda\beta$ in Figure 3. The graphs are now labeled by the values of λ as well as $\lambda\alpha$ in Figure 2 and $\lambda\beta$ in Figure 3. Again all the graphs in Figures 2 and 3 intersect at $z = 0$. The estimation methods of $\lambda, \lambda\alpha$, and $\lambda\beta$ are given in Section 3.

We want to estimate the unknown parameters λ in (1), $\lambda\alpha$ in (2), and $\lambda\beta$ in (3). and to evaluate the goodness of two approximating functions $A_\lambda(z)$ and $B_\lambda(z)$ for $R_\lambda(z)$ using the simulation method.

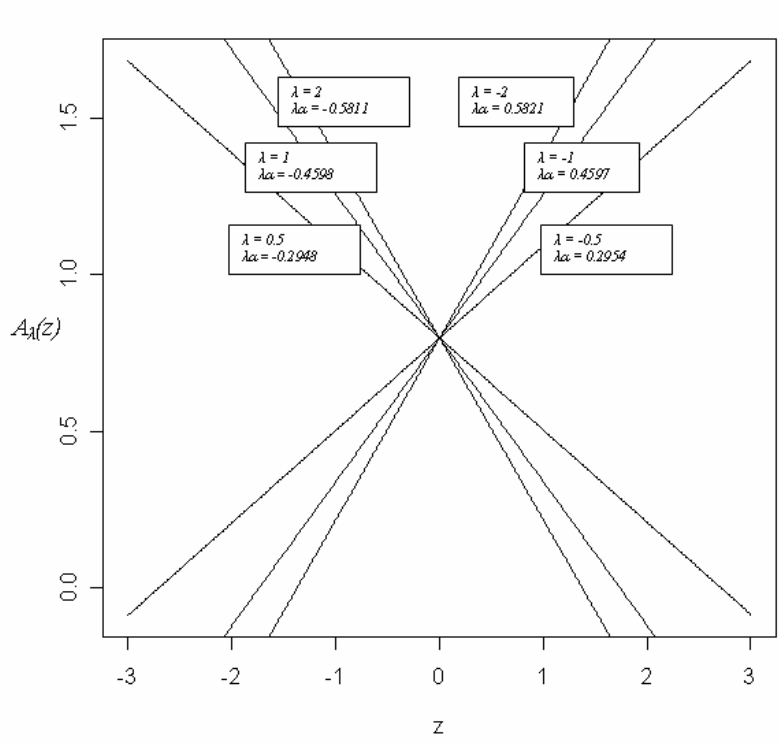


Figure 2: Plots of $A_\lambda(z)$ against z for $\lambda = \pm 2,\ \pm 1$, and ± 0.5.

2. Motivation

The skew normal distribution was introduced in Azzalini (1985, 1986). A random variable Z is said to have a skew normal distribution with the shape parameter λ if its density function at $Z = z$ is

$$f(z;\lambda) = 2\phi(z)\Phi(\lambda z). \tag{4}$$

We denote the distribution with the density in (4) as SN(0,1,λ). When $\lambda = 0$, $\Phi(\lambda z) = \Phi(0) = \frac{1}{2}$, $f(z;0) = \phi(z)$ and the random variable Z becomes a standard normal variable with the distribution N(0,1) . We now consider a random variable Y for a given value x of a random variable X satisfying

$$Y = \gamma_0 + \gamma_1 x + \sigma Z, \tag{5}$$

where γ_0, γ_1, and $\sigma(> 0)$ are unknown fixed constants. The random variable Y is distributed skew normal with density

$$f(y;\lambda,\gamma_0,\gamma_1,\sigma) = \frac{2}{\sigma}\phi\left(\frac{y-\gamma_0-\gamma_1 x}{\sigma}\right)\Phi\left(\lambda\frac{y-\gamma_0-\gamma_1 x}{\sigma}\right). \tag{6}$$

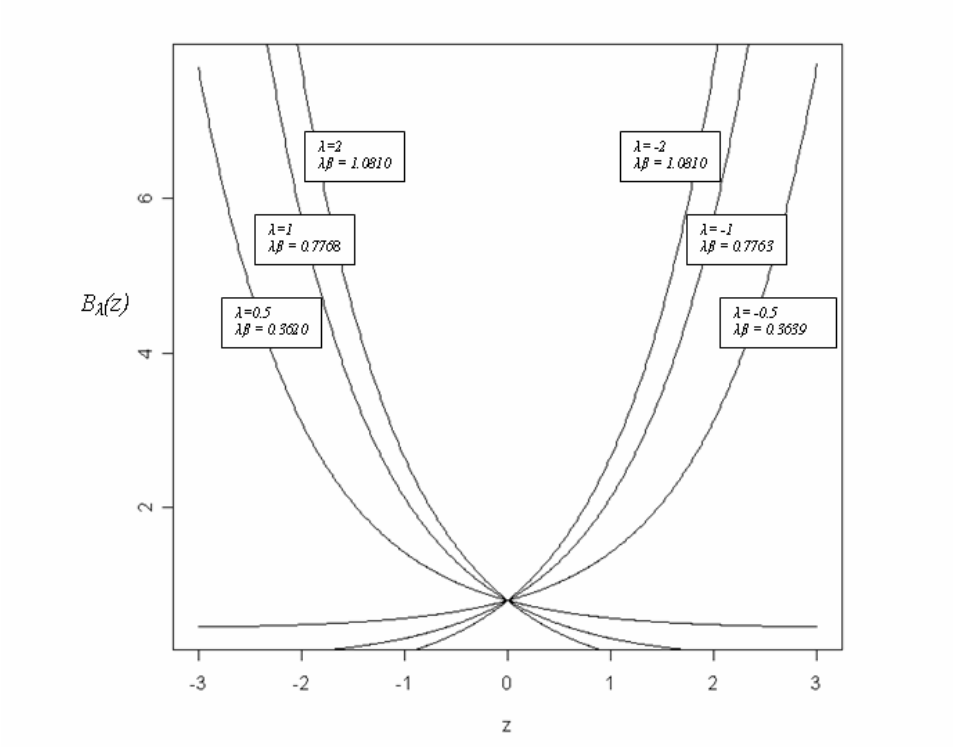


Figure 3: Plots of $B_\lambda(z)$ against z for $\lambda = \pm 2,\ \pm 1$, and $\pm .5$.

We denote the distribution with the density in (6) as $\mathrm{SN}(\gamma_0 + \gamma_1 x, \sigma, \lambda)$. We now consider n independent observations (y_i, x_i) from the skew normal distribution with density in (6). We denote $z_i = \frac{y_i - \gamma_0 - \gamma_1 x_i}{\sigma}, i = 1, ..., n$. The Maximum Likelihood Estimating Equations can be expressed as

$$\sum_{i=1}^{n} z_i = \lambda \sum_{i=1}^{n} R_\lambda(z_i), \tag{7}$$

$$\sum_{i=1}^{n} x_i z_i = \lambda \sum_{i=1}^{n} x_i R_\lambda(z_i), \tag{8}$$

$$\sum_{i=1}^{n} z_i R_\lambda(z_i) = 0, \tag{9}$$

$$\sum_{i=1}^{n} z_i^2 = n. \tag{10}$$

We use the equations (7)-(10) to estimate the parameters λ in $R_\lambda(z)$, $\lambda\alpha$ in $A_\lambda(z)$, and $\lambda\beta$ in $B_\lambda(z)$ based on the values of $z_i, i = 1, ..., n$.

The complexity for maximum likelihood estimation of the shape parameter

λ of the skew normal distribution discussed in Sartori (2006) where the author proposed a modified maximum likelihood estimation. We argue that the source of complexity is from the presence of a complex function $R_\lambda(z)$. The proposed method of estimation of λ in this paper deals with the complexity by approximating $R_\lambda(z)$ by $A_\lambda(z)$ in (1) or $B_\lambda(z)$ in (2). Our simulations demonstrate that the approximations (2) and (3) perform satisfactorily in the regions of interest and the biases in the estimates of λ using the approximations (2) and (3) are in fact sufficiently small.

References

[1] Azzalini, A. (1985). *A class of distributions which includes the normal ones. Scand. J. Statist.* 12, 171-178.

[2] Azzalini, A. (1986). *Further results on a class of distributions which includes the normal ones. Statistica* 46, 199-208.

[3] Dalla Valle, A. (2004). *The skew-normal distribution. In: Skew-Elliptical Distributions and their Applications*, Ed. Genton, M.G., 3-24. Chapman & Hall/CRC: Boca Raton, Florida.

[4] Mendenhall, W., Beaver, R.J., and Beaver, B.M. (2009). *Introduction to Probability & Statistics.* 13th edition. *Cengage/Duxbury: Florence, Kentucky*

[5] Rao, C.R. (1973). *Linear Statistical Inference and its applications.* 2nd edition. *Wiley: New York*

[6] Sartori, N. (2006). *Bias prevention of maximum likelihood scalar skew normal and skew t distributions. J. Statist. Planning and Inference.* 136, 4259-4275.

6th St.Petersburg Workshop on Simulation (2009) 628-632

Conditional FDR estimation based on a factor analytic approach of multiple testing

David Causeur[1], Chloé Friguet[2]

Abstract

Since the seminal paper by [1] introducing the False Discovery Rate (FDR) as an overall type-I error rate in multiple testing, the properties of the initial procedure have been studied under different dependence assumptions. Many recent papers dealing with this subject have highlighted the instability generated by a high amount of correlation among test statistics and major improvements of Benjamini and Hochberg's adjustment for multiplicity have been proposed to ensure a trustworthy error control. The present paper focuses on the estimation of the FDR for dependent data. Under assumption of a factor analysis model for the correlation among the test statistics, a FDR estimate is proposed, taking advantage of the latent structure to remove a conditional bias due to dependence and reduce the variance of estimation.

1. Introduction

Multiple testing procedures have for long been mainly designed for simultaneous inference on linear contrasts in Analysis of Variance settings. Exactly as single-hypothesis testing and confidence interval estimation can be viewed as two faces of the same coin, multiple testing is generally considered in this context as equivalent to estimation of simultaneous confidence intervals for the linear contrasts. Consequently, the probability of an erroneous rejection, which is the usual type-I error rate in single-hypothesis testing, has naturally been extended to the probability of at least one erroneous rejection, also called the Family-Wise Error Rate (FWER), as an overall type-I error rate for multiple tests. Generally, in Analysis of Variance settings, only a limited number of hypotheses are tested simultaneously, most often the pairwise comparisons of mean levels in different groups, which usually makes it efficient to control the FWER by Šidák or Bonferroni-type adjustment (see [10] for a comprehensive review).

In the last two decades, the development of high-throughput technologies such as remote sensing, infrared spectroscopy or genome-wide scans by microarray techniques has pointed out the limits of FWER-controlling multiple testing procedures.

[1]European University of Brittany, E-mail: david.causeur@agrocampus-ouest.fr
[2]European University of Brittany, E-mail: chloe.friguet@agrocampus-ouest.fr

The huge number of tests involved in these highly dimensional contexts indeed suggests an alternative type-I error rate, defined for the set of rejected hypotheses rather than for the whole set of tests. The False Discovery Rate (FDR), introduced by [1] as the expected proportion of errors among the rejections, has turned out to show desirable properties, especially for large datasets. FDR-controlling multiple testing procedures derived from the Benjamini and Hochberg (BH) method lead indeed to less conservative decision rules than Šidák or Bonferroni-type adjustments.

The BH procedure has initially been shown to control the FDR under assumption of independence between the test statistics (see [1]). However, independence is unrealistic for most highly dimensional multiple testing issues. Many extensions of the BH procedure have therefore been motivated by a better error control under various dependence assumptions (see [2] for a review and [6] for a recent contribution under a general dependence structure). As shown for instance by [8] and [4], a high amount of dependence among test statistics results in an increased variability of the FDR estimation and, consequently, a loss of power. Another important, yet surprising, impact of dependence is presented by [3] as a conditional bias in FDR estimation, which can lead to strongly misleading strategies. To remove this bias, [3] suggests a FDR estimate, accounting for dependence by means of a summary statistics for the dispersion in the distribution of correlations among test statistics. In the present paper, dependence among the variables is explicitly accounted for by a factor analysis model of the correlation structure, in which a limited number of latent factors is supposed to concentrate the shared variability between test statistics. Analogously with [3], a conditional FDR estimate, given the factors, is deduced and shown to provide a faithful prediction of the proportion of errors among the rejections.

Section 2 gives motivating arguments for a factor analysis modeling of the correlation in large scale significance tests. In section 3, the conditional FDR estimate is introduced, showing large improvements with respect to the usual unconditional approach.

2. Dependence in multiple testing

Let $Y = (Y^{(1)}, Y^{(2)}, \ldots, Y^{(m)})'$ be a random $m-$vector which conditional distribution with respect to some explanatory variables $x = (x^{(1)}, \ldots, x^{(p)})'$, $p \geq 1$, is normal with expectation $\mathbb{E}_x(Y) = (\beta_0^{(k)} + x'\beta^{(k)})_{k=1\ldots m}$ and variance Σ, assumed to be constant with respect to x and positive definite.

For variables with indices k in a subset $\mathcal{M}_0$ of $\{1, 2, \ldots, m\}$ of size m_0, a particular linear contrast of interest $\lambda'\beta^{(k)}$ is zero, whereas for k in $\mathcal{M}_1 = \mathcal{M}_0^c$, $\lambda'\beta^{(k)} \neq 0$. Our aim is to find out the response variables for which $\mathrm{H}_0^{(k)} : \lambda'\beta^{(k)} = 0$ is not true, or in other words, to test simultaneously the m null hypotheses. Hereafter, the least-squares estimator of $\beta^{(k)}$, calculated on a sample of n independent items, is denoted $\hat{\beta}^{(k)}$ and $\lambda'\hat{\beta}^{(k)} - \lambda'\beta^{(k)}$ is of course normally distributed with expectation 0 and variance $\sigma_k^2 n^{-1}\lambda' S_{xx}^{-1}\lambda$, where σ_k is the conditional standard deviation of $Y^{(k)}$ given x and S_{xx} is the sample variance-covariance matrix of

the explanatory variables. A usual testing procedure is based on the individual Student's test statistics $T^{(k)} = \sqrt{n}\lambda'\hat{\beta}^{(k)}/\ (s_k\sqrt{\lambda' S_{xx}^{-1}\lambda})$, where s_k^2 is the residual mean square error of the linear model relating $Y^{(k)}$ to x.

A factor analysis model for dependence

For $k = 1, \ldots, m$, let $\varepsilon^{(x,k)}$ denote the $n-$vector of residual errors in the kth regression model introduced above. [6] show that, provided there is no Borel measurable function g such that $\varepsilon^{(x,k)} = g(\varepsilon^{(x,1)}, \ldots, \varepsilon^{(x,k-1)}, \varepsilon^{(x,k+1)}, \ldots, \varepsilon^{(x,m)})$ almost surely, there exists a $m \times q$ matrix B, with $q \leq n$, a $n \times q$ matrix Z and a $n \times m$ matrix ε such that, for $k = 1, \ldots, m$,

$$\varepsilon^{(x,k)} = Z\mathbf{b}_k + \varepsilon^{(k)}, \tag{1}$$

where b_k is the kth row of B and $\varepsilon^{(k)}$ is the kth column of ε. Moreover, the columns of ε are uncorrelated and $\mathrm{Var}(\varepsilon^{(k)}) = \psi_k^2 I_n$.

The random effects regression modeling of $\varepsilon^{(x)}$ given in expression (1) is equivalent to a factor model in which B is the matrix of loadings, ψ_k^2 is the kth specific variance and Z is the matrix of unobservable latent factors. As in the exploratory factor model (see [7]), Z is assumed to be normally distributed with expectation 0 and variance I_q. Therefore, the conditional variance Σ is decomposed into the following specific and common parts:

$$\Sigma = \Psi + BB', \tag{2}$$

where Ψ is the $m \times m$ diagonal matrix of specific variances.

[4] show that a large proportion of common variance in the factor analysis decomposition has a negative impact on the stability of FDR-controlling multiple testing procedures and suggest a factor-analytic method to reduce correlation between test statistics.

EM factor analysis

Principal Factoring is probably the most famous estimation method in the factor analysis model (see [7]). It consists in an iterated Principal Component Analysis, which requires, at each step of the algorithm, the Singular Value Decomposition of an updated correlation matrix. However, in high-dimensional situations, this is computationally time and memory consuming. Since factor analysis is a particular latent variable model, an EM algorithm (see [9]) can be implemented to achieve the maximum likelihood solution and avoid SVD of large matrices. The algorithm we propose slightly modifies the initial EM algorithm to apply to the modeling of a conditional covariance matrix, given x. In the following, $\hat{\varepsilon}^{(x)}$ is the $n \times m$ residual matrix for the fixed effects of models (1). After a primary estimation $\widehat{\mathcal{M}}_0$ of $\mathcal{M}_0$ by the set of variables for which the p-value of the usual Student's test for H_0 exceeds a given significance level α, the kth column of $\hat{\varepsilon}^{(x)}$ is derived by fitting either an unrestricted linear model relating $Y^{(k)}$ to x if $k \notin \widehat{\mathcal{M}}_0$ or the same model under restriction $H_0^{(k)}$ if $k \in \widehat{\mathcal{M}}_0$. The iterative algorithm is now described through its E and M steps:

- *E step* - $\hat{Z}$ is first computed: for $i = 1, \ldots, n$, $\hat{Z}_i = G\hat{B}'\hat{\Psi}^{-1}\hat{\varepsilon}_i^{(x)}$ and $S_i^{(z)} = G + \hat{Z}_i\hat{Z}_i'$, where $G = (I_q + \hat{B}'\hat{\Psi}^{-1}\hat{B})^{-1}$ and $\hat{Z}_i$ denotes the ith row of $\hat{Z}$.
- *M step* - The uniquenesses and factor loadings are derived:

$$\hat{B} = \big[\sum_{i=1}^{n} \hat{\varepsilon}_i^{(x)}\hat{Z}_i'\big]\big[\sum_{i=1}^{n} S_i^{(z)}\big]^{-1}$$

$$\hat{\Psi} = \mathrm{diag}\big[S - \hat{B}\frac{1}{n}\sum_{i=1}^{n}\hat{Z}_i\hat{\varepsilon}_i^{(x)'}\big],$$

  where diag is here the matrix operator that sets all the off-diagonal elements to zero and $S = n^{-1}\sum_{i=1}^{n}\hat{\varepsilon}_i^{(x)}\hat{\varepsilon}_i^{(x)'}$ stands for the usual sample estimate of Σ.

It is especially important in the present multiple testing context to avoid underestimation of the uniquenesses ψ_k^2 because it would inflate artificially the test statistics. Let us focus on the estimators of the uniquenesses resulting from the above EM algorithm, viewed as residual mean square errors:

$$\hat{\psi}_k^2 \quad = \quad \frac{1}{n}\hat{\varepsilon}^{(k,x)'}P_z\hat{\varepsilon}^{(k,x)},$$

where $\hat{\varepsilon}^{(k,x)}$ is the kth column of $\hat{\varepsilon}^{(x)}$ and $P_z = I_n - \hat{Z}\big(\sum_{i=1}^{n} S_i^{(z)}\big)^{-1}\hat{Z}'$. By analogy with other nonlinear smoothing methods, we propose to account for the parametric complexity of the factor analysis model by replacing the denominator n in the above expression by the trace of P_z.

3. FDR estimation

Basically, multiple testing procedures can be viewed as the sequence of a single-hypothesis testing method applied to each test and the choice of a threshold t for the p-values, under which the null hypothesis is rejected. For each t, let V_t denote the number of erroneous rejections and R_t the number of rejections. The thresholding procedure aims at controlling an overall type-I error rate at a given level α and, for high-throughput data, it is now quite commonly accepted that a reasonable choice of type-I error is the actual False Discovery Proportion $\mathrm{FDP}_t = V_t/R_t$, namely the proportion of rejected hypotheses which are erroneously rejected. The expected FDP_t, also called the False Discovery Rate and denoted FDR_t, is defined by [1] as $\mathrm{FDR}_t = \mathbb{E}(\mathrm{FDP}_t | R_t > 0)$.

For a given type-I level α, the following method is also proposed by [1] to choose a threshold t_α with $\mathrm{FDR}_{t_\alpha} \leq \alpha$: $t_\alpha = \max_t\Big\{t \in [0,1],\ \widehat{\mathrm{FDR}}_t \leq \alpha\Big\}$, where $\widehat{\mathrm{FDR}}_t = m_0 t/R_t$ is an FDR estimate if m_0 is assumed to be known. Substituting m_0 by an accurate estimation results in a more precise control of the FDR (see for instance [5] for a review of estimation procedures).

Figure 1 illustrates the effects of correlation on the estimation of the FDR. The simulation study involves 1000 $n \times m$ datasets, with $n = 60$ and $m = 400$, which

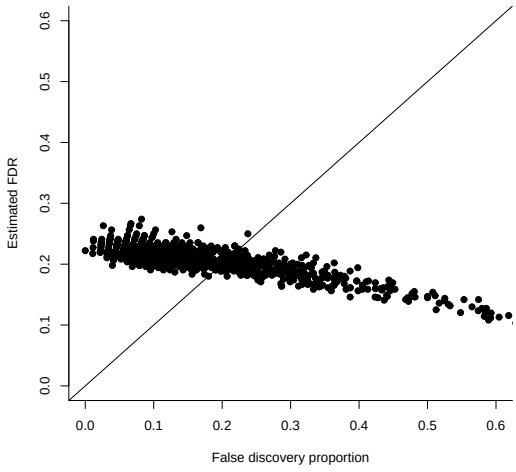


Figure 1: Estimated FDR versus the observed false discovery proportion FDP_t with $t = 0.05$ for 1000 simulated datasets. The correlation structure in the simulation study is defined by a Factor Analysis model with 5 factors and a large proportion $\pi = \mathrm{trace}(BB')/\mathrm{trace}(\Sigma)$ of common variance ($\pi = 0.67$).

rows are distributed according to a normal distribution with expectation 0 and variance $\Sigma = \Psi + BB'$, where $\mathrm{rank}(B) = q = 5$ and $\pi = \mathrm{trace}(BB')/\mathrm{trace}(\Sigma) = 0.67$. For each dataset, the means of the variables in two groups of 30 items are compared by a t-test and $\widehat{\mathrm{FDR}}_t$ is calculated with $t = 0.05$. As also mentioned by [3], Figure 1 shows that $\widehat{\mathrm{FDR}}_t$ and the observed False Discovery Proportion are negatively correlated, which can lead to strongly misleading conclusions about the error control.

In the sequel, we propose an alternative FDR estimate based on the distribution of V_t conditionally on the factors. It is shown in [4] that, under assumption (2), V_t is distributed, conditionally on the factors Z, as the sum of independent Bernoulli variables with expectations $t_z^{(k)}$, k in $\mathcal{M}_0$, given by:

$$t_z^{(k)} \quad = \quad 1 - \Phi^{(1)}\Big(-\frac{b_k'}{\psi_k}\tau(\mathbf{Z}) + \frac{\sigma_k}{\psi_k}[-u_t; u_t]\Big),$$

where $\Phi^{(1)}(A)$ is the probability that a random variable normally distributed with mean 0 and standard deviation 1 belongs to $A \subseteq \mathbb{R}$, u_t is the $(1 - t/2)-$quantile of the standard normal distribution, $\tau(\mathbf{Z}) = \sqrt{n}\hat{\beta}_z'\lambda/\sqrt{\lambda'\mathbf{S}_{xx}^{-1}\lambda}$ and $\hat{\beta}_z$ denotes the least-squares estimator of the $p \times q$ matrix of the slope coefficients in the multivariate regression of $\mathbf{Z}$ onto the explanatory variables.

Hence, the conditional expectation of V_t, given the factors, is $\mathbb{E}(V_t \mid \mathbf{Z}) = \sum_{k\in\mathcal{M}_0} t_z^{(k)}$ and the following conditional FDR estimate is deduced: $\mathrm{Fdr}_t(\mathbf{Z}) =$

$\mathbb{E}(V_t \mid \mathbf{Z})/R_t = \sum_{k \in \mathcal{M}_0} t_z^{(k)}/R_t$. The presentation will focus on the distribution of this FDR estimate and the beneficial impact of conditioning on the factors to improve error control.

References

[1] Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing, *Journal of the Royal Statistical Society B* **57**, 289–300.

[2] Dudoit, S. Shaffer, J.P. and Boldrick J.C. (2003). Multiple hypothesis testing in microarray experiments, *Statistical Science* **18** (1), 71–103.

[3] Efron, B. (2007). Correlation and large-scale simultaneous testing, *J. Amer. Statist. Assoc.* **102**, 93–103.

[4] Friguet, C., Kloareg, M. and Causeur, D. (2009). A factor model approach to multiple testing under dependence, *Submitted*.

[5] Langaas, M., Lindqvist, B.H. and Ferkingstad, E. (2005). Estimating the proportion of true null hypotheses, with application to DNA microarray data, *Journal of the Royal Statistical Society. Series B* **67**, 216. 555–572.

[6] Leek, J.T. and Storey, J.D. (2008). A general framework for multiple testing dependence, *Proceedings of the National Academy of Sciences, USA*, **105** 18718–18723.

[7] Mardia, K. V., Kent, J. T., and Bibby, J. M. (1979). Multivariate Analysis. *Probability and Mathematical Statistics*. Academic Press, London.

[8] Owen, A.B. (2005). Variance of the number of false discoveries, *J. R. Stat. Soc. Ser. B Stat. Methodol.* **67** (3), 411–426.

[9] Rubin, D.B. and Thayer, D.T. (1982). EM algorithms for ML factor analysis, *Psychometrika* **47**, 1, 69–76

[10] Shaffer, J.P. (1995). Multiple Hypothesis Testing, *Annual Review of Psychology* **46** 561–584.

6th St.Petersburg Workshop on Simulation (2009) 634-638

Performance of a new dependence measure between stable non-Gaussian random variables and comparison with other measures: a simulation study[1]

Bernard Garel[2], Bernedy Kodia[3]

Abstract

The signed symmetric covariation coefficient is a new dependence measure between symmetric α-stable random variables. This coefficient satisfies most properties of the classical Pearson coefficient. In the case of sub-Gaussian random vectors, it is shown that this coefficient coincides with the generalized association parameter (gap) proposed by Paulauskas. In the case of linear combinations of stable random variables the exact values of these measures of dependance is given. Then we propose two estimators of this coefficient, based respectively on fractional lower-order moments and on screened ratio. A comparison based on simulation results is made in the case of sub-Gaussian random vectors and in the case of linear combinations of independent symmetric stable random variables.

1. Introduction

Many types of physical phenomena and financial data exhibit a very high variability and stable distributions are often used for their modeling. Since the seminal work of Mandelbrot (1963) who suggested the stable laws as possible models for the distributions of income and speculative prices, the interest for these laws greatly increased and now they are widely used in telecommunications and many other fields such as physics, biology, genetic and geology (see Uchaikin and Zolotarev, 1999).

Stable non-Gaussian random vectors do not possess moments of second order. Therefore, the correlation matrix which allows to understand the association between the coordinates of a random vector, does not exist. Then we need other measures of dependence.

[1]This work was supported by grant MiPy Region 06001715 and 07005628.

[2]National Polytechnic Institute, E-mail: garel@n7.fr

[3]Toulouse University, E-mail: bernedy.kodia@enseeiht.fr

Kanter and Steiger (1974) showed that, under some conditions, the conditional expectation of a stable variable given another one is linear. Then Paulauskas (1976) proposed the generalized association parameter (gap). After that Miller (1978) proposed a new dependence measure called covariation. The constant of linearity of conditional expectation has been expressed by means of this measure and then called the covariation coefficient. Garel *et al.* (2004) introduced the symmetric covariation. Then Garel and Kodia (2009) introduced the signed symmetric covariation coefficient. In the case of sub-Gaussian random vectors, this new coefficient coincides with the gap.

This paper is organized as follow: Section 2 is devoted to a reminder of basic definitions and some properties of stable random vectors, the dependence measures and the above mentioned coefficients. We give their first properties. Other properties of these coefficients are discussed in the context of sub-Gaussian random vectors in Section 3. We also give the exact expressions of the signed symmetric covariation coefficient and the gap in the case of linear transformations of independent stable random variables. Then, in Section 4, we compare two estimates of the signed symmetric covariation coefficient from a simulation study.

2. Stable random vectors and dependence measures

Here we adopt Samorodnitsky and Taqqu (1994)'s conventions and denote by $S_\alpha(\gamma,\beta,\delta)$ the law of a stable random variable, with $0<\alpha\leq 2, \gamma\geq 0, -1\leq\beta\leq 1$ and δ a real parameter. If the distribution is symmetric, it is denoted by $S\alpha S(\gamma)$ or shortly $S\alpha S$, which means that $\beta=\delta=0$. Let $0<\alpha<2$. The random vector $X=(X_1,X_2)$ is $S\alpha S$ iff there exists a unique finite symmetric measure, called spectral measure, $\mathbf{\Gamma}$ on the unit circle $S_2=\{\mathbf{s}\in\mathbb{R}^2 : \|\mathbf{s}\|=1\}$ such that for all $\boldsymbol{\theta}\in\mathbb{R}^2$:

$$\phi_{\mathbf{X}}(\boldsymbol{\theta}) = E\exp\{i\langle\boldsymbol{\theta},\mathbf{X}\rangle\} = \exp\left\{-\int_{S_2}|\langle\boldsymbol{\theta},\mathbf{s}\rangle|^\alpha\mathbf{\Gamma}(d\mathbf{s})\right\} \tag{1}$$

where $\langle .,.\rangle$ denotes the inner product. In the sequel, unless differently specified , we assume $\alpha>1$ and consider symmetric stable random variables or vectors.

Definition 1. *The covariation of X_1 on X_2, introduced by Miller (1978), is the real number*

$$[X_1,X_2]_\alpha = \int_{S_2} s_1 s_2^{<\alpha-1>}\mathbf{\Gamma}(d\mathbf{s}),$$

where $a^{<p>}=|a|^p\text{sign}(a)$ is called signed power.

Let γ_{X_1} be the scale parameter of the $S\alpha S$ random variable X_1. Then we have

$$[X_1,X_1]_\alpha = \int_{S_2}|s_1|^\alpha\mathbf{\Gamma}(d\mathbf{s}) = \gamma_{X_1}^\alpha,$$

We denote $\|X_1\|_\alpha = ([X_1,X_1]_\alpha)^{1/\alpha}=\gamma_{X_1}$. Then $\|\cdot\|_\alpha$ defines a norm, called covariation norm.

Definition 2. *Let (X_1, X_2) be a bivariate $S\alpha S$ random vector with $\alpha > 1$. The signed symmetric covariation coefficient between X_1 and X_2, see Garel and Kodia (2009), is the quantity:*

$$\operatorname{scov}(X_1, X_2) = \kappa_{(X_1,X_2)} \left| \frac{[X_1, X_2]_\alpha [X_2, X_1]_\alpha}{\|X_1\|_\alpha^\alpha \|X_2\|_\alpha^\alpha} \right|^{\frac{1}{2}}, \tag{2}$$

where

$$\kappa_{(X_1,X_2)} = \begin{cases} \operatorname{sign}([X_1, X_2]_\alpha) & \text{if} \quad \left| \frac{[X_1,X_2]_\alpha}{\|X_2\|_\alpha^\alpha} \right| \geq \left| \frac{[X_2,X_1]_\alpha}{\|X_1\|_\alpha^\alpha} \right|, \\ \operatorname{sign}([X_2, X_1]_\alpha) & \text{if} \quad \left| \frac{[X_1,X_2]_\alpha}{\|X_2\|_\alpha^\alpha} \right| < \left| \frac{[X_2,X_1]_\alpha}{\|X_1\|_\alpha^\alpha} \right|. \end{cases} \tag{3}$$

Then $\kappa_{(X_1,X_2)}$ is the sign of the coefficient of the covariation which has the greatest absolute value.

The signed symmetric covariation coefficient has the following properties: $-1 \leq \operatorname{scov}(X_1, X_2) \leq 1$ and if X_1, X_2 are independent, then $\operatorname{scov}(X_1, X_2) = 0$; for all $a \neq 0$, $|\operatorname{scov}(X, aX)| = 1$; if a and b are two non-zero reals, then $\operatorname{scov}(aX_1, bX_2) = \pm \operatorname{scov}_\alpha(X_1, X_2)$; for $\alpha = 2$, $\operatorname{scov}(X_1, X_2)$ coincides with the usual correlation coefficient.

Another measure of dependence: the generalized association parameter (gap) was introduced by Paulauskas (1976). Let (U_1, U_2) be a random vector on S_2 with probability distribution $\widetilde{\mathbf{\Gamma}} = \mathbf{\Gamma}/\mathbf{\Gamma}(S_2)$. Because of the symmetry of $\mathbf{\Gamma}$, one has $EU_1 = EU_2 = 0$. Then the gap is defined as:

$$\widetilde{\rho} = \frac{EU_1U_2}{(EU_1^2 EU_2^2)^{1/2}}.$$

It is a measure of dependence for (X_1, X_2). For a bivariate stable vector with characteristic function ((1)) the gap $\widetilde{\rho}$ has the following properties valid for all $0 < \alpha \leq 2$: $-1 \leq \widetilde{\rho} \leq 1$ and if a distribution corresponds to a random vector with independent coordinates, then $\widetilde{\rho} = 0$; $|\widetilde{\rho}| = 1$ if and only if the distribution is concentred on a line; for $\alpha = 2$, $\widetilde{\rho}$ coincides with the correlation coefficient of the Gaussian random vector; $\widetilde{\rho}$ is independent of α and depends only on the spectral measure $\mathbf{\Gamma}$; if the characteristic function of $\mathbf{X} = (X_1, X_2)$ is given by

$$\phi_{\mathbf{X}}(\boldsymbol{t}) = \exp\left\{ -C(\gamma_1^2 t_1^2 + 2r\gamma_1\gamma_2 t_1 t_2 + \gamma_2^2 t_2^2)^{\alpha/2} \right\},$$

where C is an appropriate constant, then r is the gap.

3. Sub-Gaussian case and linear combination case

We recall the definition of a sub-Gaussian random vector.

Definition 3. *Let $0 < \alpha < 2$, let G_1, G_2 be zero mean jointly normal random variables and let A be a positive random variable such that $A \sim S_{\alpha/2}\big(\big(\cos\frac{\pi\alpha}{4}\big)^{2/\alpha}, 1, 0\big)$, independent of (G_1, G_2), then $\mathbf{X} = A^{1/2}\mathbf{G} = (A^{1/2}G_1, A^{1/2}G_2)$ is a sub-Gaussian random vector with underlying Gaussian vector $\mathbf{G} = (G_1, G_2)$.*

Proposition 1. *Let $1 < \alpha < 2$ and $\mathbf{X}$ a sub-Gaussian random vector. Then the signed symmetric covariation coefficient between the components of $\mathbf{X}$ is equal to the gap.*

Let $1 < \alpha \leq 2$ and let X_1, X_2 be independent random variables such that $X_k \sim S_\alpha(\gamma_k, 0, 0), k = 1, 2$. Let $\mathbf{A} = \{a_{jk}\}$, $1 \leq j \leq k \leq 2$, be a real matrix. The random vector $\mathbf{Y} = (Y_1, Y_2) = \mathbf{AX}$, whose components are linear combinations of the X_k:

$$Y_j = \sum_{k=1}^{2} a_{jk} X_k, \quad j = 1, 2, \tag{4}$$

is symmetric α-stable. Then we have:

Proposition 2. *Let $1 < \alpha \leq 2$ and let $\mathbf{Y} = \mathbf{AX}$ an α-stable random vector, then the signed symmetric covariation coefficient is given by:*

$$scov(Y_1, Y_2) = \kappa_{(Y_1,Y_2)} \frac{|a_{11}a_{21}|^\alpha + |a_{12}a_{22}|^\alpha + a_{11}a_{22}(a_{12}a_{21})^{\langle\alpha-1\rangle} + a_{12}a_{21}(a_{11}a_{22})^{\langle\alpha-1\rangle}}{|a_{11}a_{21}|^\alpha + |a_{12}a_{22}|^\alpha + |a_{12}a_{21}|^\alpha + |a_{11}a_{22}|^\alpha}^{\frac{1}{2}}{}^{1/2},$$

and the gap is given by:

$$\widetilde{\rho} = \frac{a_{11}a_{21}(a_{11}^2 + a_{21}^2)^{\frac{\alpha}{2}-1} + a_{12}a_{22}(a_{12}^2 + a_{22}^2)^{\frac{\alpha}{2}-1}}{\left[\begin{array}{cc} a_{11}^2(a_{11}^2 + a_{21}^2)^{\frac{\alpha}{2}-1} + a_{12}^2(a_{12}^2 + a_{22}^2)^{\frac{\alpha}{2}-1} & a_{21}^2(a_{11}^2 + a_{21}^2)^{\frac{\alpha}{2}-1} + a_{22}^2(a_{12}^2 + a_{22}^2)^{\frac{\alpha}{2}-1} \end{array}\right]^{\frac{1}{2}}}.$$

4. Estimation of the signed symmetric covariation coefficient

First we begin by a classical result.

Lemma 1. *Let (X_1, X_2) be $S\alpha S$ with $\alpha > 1$. Then for all $1 \leq p < \alpha$,*

$$\frac{[X_1, X_2]_\alpha}{\|X_2\|_\alpha^\alpha} = \frac{EX_1X_2^{<p-1>}}{E|X_2|^p}.$$

Using this lemma we propose the following estimator for our coefficient:

$$\widehat{\mathrm{scov}}(X_1, X_2) = \widehat{\kappa}_{(X_1,X_2)} \frac{\left|\left(\sum_{i=1}^{n} X_{1i}\,\mathrm{sign}(X_{2i})\right)\left(\sum_{i=1}^{n} X_{2i}\,\mathrm{sign}(X_{1i})\right)\right|^{1/2}}{\left[\left(\sum_{i=1}^{n} |X_{1i}|\right)\left(\sum_{i=1}^{n} |X_{2i}|\right)\right]^{1/2}} \tag{5}$$

where

$$\widehat{\kappa}_{(X_1,X_2)} = \begin{cases} [c]ccc\text{sign}\Big(\sum_{i=1}^{n} X_{1i}\ \text{sign}(X_{2i})\Big) \text{ if} \\ \left|\frac{\sum_{i=1}^{n} X_{1i}\ \text{sign}(X_{2i})}{\sum_{i=1}^{n}|X_{1i}|}\right| \geq \left|\frac{\sum_{i=1}^{n} X_{2i}\ \text{sign}(X_{1i})}{\sum_{i=1}^{n}|X_{2i}|}\right|, \\ \text{sign}\Big(\sum_{i=1}^{n} X_{2i}\ \text{sign}(X_{1i})\Big) \text{ if} \\ \left|\frac{\sum_{i=1}^{n} X_{1i}\ \text{sign}(X_{2i})}{\sum_{i=1}^{n}|X_{1i}|}\right| < \left|\frac{\sum_{i=1}^{n} X_{2i}\ \text{sign}(X_{1i})}{\sum_{i=1}^{n}|X_{2i}|}\right| \end{cases} \tag{6}$$

and $(X_{11}, X_{21}), ..., (X_{1n}, X_{2n})$ are iid copies of (X_1, X_2). By the classical law of large number this estimator is strongly consistent. It depends neither on the unknown value of α nor on the spectral measure of X.

A second estimator of our coefficient is based on screened ratio: $\widehat{\text{scov}}^{SR}(X_1, X_2) =$

$$\widehat{\kappa}_{(X_1,X_2)} \frac{\left|\Big(\sum_{i=1}^{n} X_{1i}X_{2i}^{-1}\mathbb{I}_{]c_1,c_2[}(|X_{2i}|)\Big)\Big(\sum_{i=1}^{n} X_{2i}X_{1i}^{-1}\mathbb{I}_{]c_1,c_2[}(|Y_i|)\Big)\right|^{1/2}}{\Big[\Big(\sum_{i=1}^{n}\mathbb{I}_{]c_1,c_2[}(|X_{1i}|)\Big)\Big(\sum_{i=1}^{n}\mathbb{I}_{]c_1,c_2[}(|X_{2i}|)\Big)\Big]^{1/2}}, \tag{7}$$

where c_1 et c_2 are constants to specify. It results from Kanter and Steiger (1974) that ((7)) is also strongly consistent towards the signed symmetric coefficient of covariation.

Performances of the two estimators can be evaluated from Table (1) in the sub-Gaussian case and from Table (2) in the case of linear combinations ((4)) of symmetric independent stable random variables. The scale parameters of X_1 and X_2 are respectively denoted by γ_1 and γ_2. The size of the simulated samples is n. The number of replications is 100. In Formula (7) we took $c_1 = 1$ and $c_2 = \infty$. In each table the displayed value is the mean over the replications. The positive values showed below are the mean absolute deviations to the mean displayed upper.

References

[1] Garel B., Kodia B. (2009) Signed symmetric covariation coefficient for alpha-stable dependence modeling. *C. R. Acad. Sci. Paris.,* Ser. **I 347**, 315–320.

[2] Garel B., d'Estampes L. and Tjostheim D. (2004) Revealing some unexpected dependence properties of linear combinations of stable random variables using symmetric covariation. *Communications in Statistics: Theory and Methods*, **33 (4)**, 768–786.

[3] Kanter M., Steiger W. L. (1974) Regression and autoregression with infinite variance. *Advanced Applied Probabilities,* **6**, 768–783.

[4] Mandelbrot B. (1963) The variation of speculative prices. *J. Business,* **36**, 394–419.

Table 1: Estimation results in the sub-Gaussian case.

Data	$\alpha = 1.5$, n=100, $\gamma_1 = 10$ and $\gamma_2 = 150$								
scov	-1.00	-0.80	-0.60	-0.40	-0.20	0.00	0.10	0.30	0.50
$\widehat{\text{scov}}$	-1.00	-0.79	-0.57	-0.40	-0.22	0.02	0.07	0.26	0.48
	0.00	0.07	0.13	0.14	0.16	0.16	0.18	0.18	0.13
$\widehat{\text{scov}}^{SR}$	-1.00	-0.81	-0.60	-0.41	-0.26	0.02	0.10	0.36	0.53
	0.00	0.25	0.25	0.32	0.31	0.41	0.37	0.39	0.30
Data	$\alpha = 1.5$, n=500, $\gamma_1 = 10$ and $\gamma_2 = 150$								
scov	-1.00	-0.80	-0.60	-0.40	-0.20	0.00	0.10	0.30	0.50
$\widehat{\text{scov}}$	-1.00	-0.79	-0.61	-0.39	-0.19	0.02	0.10	0.28	0.48
	0.00	0.06	0.06	0.11	0.08	0.10	0.10	0.09	0.07
$\widehat{\text{scov}}^{SR}$	-1.00	-0.78	-0.60	-0.40	-0.25	-0.00	0.14	0.34	0.49
	0.00	0.14	0.18	0.15	0.14	0.17	0.17	0.13	0.14

Table 2: Estimation results for linear combinations of stable variables.

Data	$\alpha = 1.5$, n=100, $\gamma = 10$								
a_{11}	100	30	12	2	2	10	-13	-11	-12
a_{12}	16	45	-17	2	2	0	9	13	21
a_{21}	-50	-12	2	-39	-18	0	2	2	2
a_{22}	-8	-2	12	10	10	10	5	4	5
scov	-1.00	-0.80	-0.60	-0.40	-0.20	0.00	0.11	0.30	0.50
$\widehat{\text{scov}}$	-1.00	-0.79	-0.61	-0.42	-0.23	-0.01	0.12	0.29	0.49
	0.00	0.06	0.13	0.15	0.15	0.16	0.18	0.15	0.14
$\widehat{\text{scov}}^{SR}$	-1.00	-0.77	-0.60	-0.41	-0.23	-0.02	0.07	0.27	0.51
	0.00	0.13	0.11	0.15	0.11	0.24	0.18	0.12	0.09

[5] Miller G. (1978) Properties of certain symmetric stable distributions. *Journal of Multivariate Analysis*, **8** (3) 346–360.

[6] Paulauskas V. J. (1976) Some remarks on Multivariate Stable Distributions. *Journal of Multivariate Analysis,* **6**, 356–368.

[7] Samorodnitsky G., Taqqu M. S. (1994) Stable non-Gaussian random processes. *Stochastic Modeling. Chapman & Hall, New York-London.*

[8] Uchaikin V.V., Zolotarev V.M. (1999) Chance and Stability. Stable Distributions and their Applications. *de Gruyter, Berlin-New York*

6th St.Petersburg Workshop on Simulation (2009) 640-644

Study on Linear Regression Analysis of Multivariate Compositional Data

Guo Lijuan[1], Guan Rong[2]

Abstract

Compositional data analysis is used to reflect the proportion structure of research objects.In this paper constrained regression is applied to compositional data regression and a method of multivariate regression analysis, based on the inner product in space $R^{n\times m}$, is put forward, which aims at the situation when dependent and independent variables are all correlated compositional data. This method transforms the unit-sum constraint within components into constrained regression, and eliminates the adverse effects from multicollinearity.At the same time, by keeping all the information from original variables, the method ensures the regression model can be explained by compositional data with different meanings.
Key words: Compositional data, Inner product based on space $R^{n\times m}$, Multivariate compositional data linear regression, Linear constrained regression

1. Introduction

Compositional data is widely applied in data analysis in the fields of social sciences, economic and management. It can be used to reflect the proportional structure of studying objects.

Denote m-dimensioned compositional data as$X = (x_1, x_2, \cdot, x_m)'$whose m components satisfy the unit-sum constraint as follows:

$$\sum_{j=1}^{m} x_j = 1, x_j \geq 0;$$

The concept of compositional data can be traced back to the work of Ferrers in the 19th century[1].Not until Aitchison had published the statistical analysis of compositional data in 1986, there was no systematical theory monograph of Compositional Data[2]. This book not only deeply studied theoretical methods about compositional,but also proposed logratio transformation for compositional data analysis. Zhang Yaoting (2000) discussed the regression modeling method of one dependent compositional data to another independent compositional data by

[1]Beihang University,E-mail: glj_buaa@yahoo.com.cn
[2]Beihang University, E-mail: funny_2000@163.com

using asymmetric logratio transformation[3]. Based on the symmetrical logratio transformation[4] and Partial Least Squares (PLS) regression, Huiwen Wang et al. (2003) put forward a simple linear regression modeling method of compositional data. But this model can only interpret the relativity between components of dependent variable and independent variable. By combining symmetrical logratio transformation with PLS path modeling analysis[5], Huiwen Wang et al. (2006) proposed a multivariate linear regression modeling method, aiming at the situation when dependent and independent variables are all correlated compositional data. In this model,extracting latent variables from original?compositional data may cause information loss which reduce the accuracy of the model.

Based on such research background, constrained regression, based on the inner product of space, is applied to compositional data regression of a single compositional dependent variable and one or more independent variables. In the proposed model, all compositional variables are treated as a whole and no latent variables need to be extracted. And it is easy to interpret the meanings and functions of different compositional variables in the regression model. Simulation study validates the effectiveness of this method.

2. Linear regression analysis of multivariate compositional data

Compositional Data is a special distribution data in the space of$R^{n\times m}$.Based on inner product in$R^{n\times m}$,the multivariate linear regression model of generalized distribution data shall be introduced first.

2.1. Multivariate linear regression model based on the space of $R^{n\times m}$

In space of $R^{n\times m}$,definition of inner product is given by

$$\langle X,Y\rangle = tr(X'Y) = \sum_{j=1}^{m}\sum_{i=1}^{n} x_{ij}y_{ij} \tag{1}$$

Assume that the dependent variable Y and the independent variables X_1,X_2,$\cdots$,X_p are distribution data in $R^{n\times m}$, and denote

$$Y = (y_1, y_2, \cdots, y_m)', X_k = (x_1^{(k)}, x_2^{(k)}, x_m^{(k)})', k = 1, 2, \cdots p$$

For n sample points, it follows

$$Y = \begin{bmatrix} y_{11} & \cdots & y_{1m} \\ \vdots & \ddots & \vdots \\ y_{n1} & \cdots & y_{nm} \end{bmatrix} \in R^{n\times m}, X_k = \begin{bmatrix} x_{11}^k & \cdots & x_{1m}^k \\ \vdots & \ddots & \vdots \\ x_{n1}^k & \cdots & x_{nm}^k \end{bmatrix} \in R^{n\times m}, k = 1, 2, \cdots p \tag{2}$$

Therefore, the multivariate linear regression model of distribution data can be defined as

$$Y = \beta_0 E + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_p X_p + \varepsilon, \tag{3}$$

where, $E = \begin{pmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{pmatrix} \in R^{n\times m}, \varepsilon \in R^{n\times m}$stands for random error, and β_0, β_1,$\cdots$,$\beta_p$are model parameters to be estimated. Denote estimated values as $\hat{\beta}_0, \hat{\beta}_1, \cdots, \hat{\beta}_p$,and the regression model is expressed as

$$\hat{Y} = \hat{\beta}_0 E + \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2 + \ldots + \hat{\beta}_p X_p \tag{4}$$

According to the formula(1),sum of square errors $SSE$of the multivariate linear regression model can be given by

$$\begin{aligned} SSE &= \| Y - \hat{\beta}_0 E - \sum_{k=1}^{p} \hat{\beta}_k X_k \|^2 \\ &= \sum_{j=1}^{m}\sum_{i=1}^{n}(y_{ij} - \hat{\beta}_0 - \sum_{k=1}^{p} \hat{\beta}_k x_{ij}^{(k)})^2 \end{aligned} \tag{5}$$

According to Ordinary Least Squares (OLS),SSE shall be minimized. In order to seek the least-squares solution of $\beta_0, \beta_1, \cdots, \beta_p$,calculate the partial derivatives of the formula(5).

Thus, the normal equations are obtained as follows:

$$\begin{aligned} &mn\hat{\beta}_0 + \sum_{k=1}^{p} \hat{\beta}_k \sum_{j=1}^{m}\sum_{i=1}^{n} x_{ij}^{(k)} = \sum_{j=1}^{m}\sum_{i=1}^{n} y_{ij} \\ &\hat{\beta}_0 \sum_{j=1}^{m}\sum_{i=1}^{n} x_{ij}^{(l)} + \sum_{k=1}^{p} \hat{\beta}_k \sum_{j=1}^{m}\sum_{i=1}^{n} x_{ij}^{(l)} x_{ij}^{(k)} = \sum_{j=1}^{m}\sum_{i=1}^{n} x_{ij}^{(l)} y_{ij}, l = 1, 2, \cdots p \end{aligned} \tag{6}$$

By formula (1) and (2), calculation formula of $\hat{\beta}_1, \hat{\beta}_2, \cdots \hat{\beta}_p$ is

$$\begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \vdots \\ \hat{\beta}_p \end{bmatrix} = \begin{bmatrix} \langle E, E\rangle & \langle E, X_1\rangle & \cdots & \langle E, X_p\rangle \\ \langle X_1, E\rangle & \langle X_1, X_1\rangle & \cdots & \langle X_1, X_P\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle X_p, E\rangle & \langle X_P, X_1\rangle & \cdots & \langle X_p, X_P\rangle \end{bmatrix}^{-1} \begin{bmatrix} \langle E, Y\rangle \\ \langle X_1, Y\rangle \\ \vdots \\ \langle X_p, Y\rangle \end{bmatrix} \tag{7}$$

Similar to ordinary multivariate linear regression model, denote

$$E = (E \quad X_1 \quad X_2 \quad \cdots \quad X_m), \tag{8}$$

and the least-squares solutions of distribution data regression can be defined as

$$\hat{\beta} = (X'X)^{-1}X'Y. \tag{9}$$

2.2. Linear regression model of multivariate compositional data

The method proposed above gives a good solution to regression analysis on distribution data, and effectively overcomes multicollinearity within components of compositional data, which is a specific distribution data characterized by unit-sum constraint. In this method, the unit-sum constraint will be converted into constraints over regression coefficients.

2.2.1. Conversion of unit-sum constraint

Let dependent variable Y and independent variables $X_1, X_2, \cdots, X_p$ be compositional Data in $R^{n\times m}$. According to the unit-sun constraint, it follows

$$\sum_{j=1}^{m} y_j = 1, y_j \geq 0; \sum_{j=1}^{m} x_j^{(k)} = 1, x_j^{(k)} \geq 0, k = 1, 2, \cdots p \tag{10}$$

Thus, multivariate compositional data regression model $\hat{Y} = \hat{\beta}_0 E + \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2 + \ldots + \hat{\beta}_p X_p$ can be expressed as

$$\begin{array}{c} y_1 = \beta_0 + \beta_1 x_1^{(1)} + \cdots + \beta_p x_1^{(p)} \\ y_1 = \beta_0 + \beta_1 x_2^{(1)} + \cdots + \beta_p x_2^{(p)} \\ \vdots \\ y_1 = \beta_0 + \beta_1 x_m^{(1)} + \cdots + \beta_p x_m^{(p)} \end{array} \tag{11}$$

Integrate both sides of equations in formula (11) and by unit-sum constraint it follows

$$m\beta_0 + \beta_1 + \cdots + \beta_p = 1 \tag{12}$$

Therefore, a conclusion can be drawn out that the multivariate compositional data regression model is a regression model with linear constraints, which will be provided in the following paragraphs.

2.2.2. Regression model with linear constraints

In general, regression model $y = X\beta + \varepsilon$ has no constraints on β. However, when applied to practical problems, regression models will become constrained. The constraints include equality constraints, inequality constraints, linear constraints, nonlinear constraints, random constraints, etc.Therefore, study on regression models with constraints is also of great significance.

The following discussion covers regression with linear constraints.

Given model

$$\begin{cases} y = X\beta + \varepsilon \\ H\beta = c \\ E(\varepsilon) = 0, Var(\varepsilon) = \sigma^2 I_n \end{cases} \tag{13}$$

where X is a matrix of $n \times (p+1)$ with rank of $P+1$, and H is a matrix of $q \times (p+1)$ with rank of q, and c is a q-dimensioned vector. According to OLS, the estimators of β can be obtained by solving the minimum solution of

$$Q = (y - X\beta)'(y - X\beta) + \lambda'(H\beta - c) \tag{14}$$

where λ the Lagrange multiplier, is a undetermined q-dimensioned vector and $\hat{\beta}_H$ is the estimators of β.

then

$$\begin{aligned} \hat{\beta}_H &= (X'X)^{-1}X'y - \tfrac{1}{2}H(X'X)^{-1}H'\hat{\lambda}_H \\ &= \hat{\beta} - \tfrac{1}{2}(X'X)^{-1}H\hat{\lambda}_H \end{aligned} \tag{15}$$

$$c = H\hat{\beta}_H = H\hat{\beta} - \frac{1}{2}H(X'X)^{-1}H'\hat{\lambda}_H \tag{16}$$

where, $\hat{\beta} = (X'X)^{-1}X'y$ is the least-squares solution of unconstrained regression.

Considering that $X'X$ is a positively definite matrix and H has full-row rank, it is easy to prove that $H(X'X)^{-1}H'$ is also a positively definite matrix.

According to formula (15) and (16), parameter estimation of linear regression model of multivariate compositional data can be expressed as

$$\hat{\beta}_H = \hat{\beta} + (X'X)^{-1}H'[H(X'X)^{-1}H']^{-1}(c - H\hat{\beta}), \tag{17}$$

where, $H = (m, 1, \cdots, 1), c = 1$.

3. Simulation study

To validate the method of multivariate compositional data regression analysis proposed above, simulation data of 3-dimensioned compositional data sets and with 14 sample points each will be used. Random generation goes as follows:

Generate two random data sets obeying normal distribution $N(100, 50)$, each with 3 variables and 14 sample points,denoted as $N_1 = (n_{11}, n_{12}, n_{13})', N_2 = (n_{21}, n_{22}, n_{23})'$.Let $x_{1i} = \frac{n_{1i}}{\sum_{i=1}^{3} n_{1i}}, x_{2i} = \frac{n_{2i}}{\sum_{i=1}^{3} n_{2i}}, i = 1, 2, 3$,and compositional data can be obtained, denoted as $X_1 = (x_{11}, x_{12}, x_{13})', X_2 = (x_{21}, x_{22}, x_{23})'$. By the linear constraint in regression model of multivariate compositional data, it follows $Y = 0.2 + 0.2X_1 + 0.2X_2 + \varepsilon$,where ε is a 3-dimensioned random error, obeying normal distribution $N(0, 0.0001)$.Table 2 shows comparisons between the

Compared with the designed model, results of parameter estimation in the regression model are satisfactory, which verifies the effectiveness of this method.

4. Conclusion

The paper puts forward a new analysis method of multivariate compositional data regression, which is an effective solution to regression model analysis when both the dependent variable and independent variables are all the compositional data.

Table 1: Independent variables of compositional data by random generation

Sample Point	X_1			X_2		
1	0.017699	0.319322	0.662979	0.087616	0.523921	0.388463
2	0.018698	0.355263	0.626039	0.076007	0.486784	0.437209
3	0.019403	0.412438	0.568159	0.068636	0.487816	0.443548
4	0.0066	0.472659	0.520742	0.062036	0.480342	0.457622
5	0.007285	0.36582	0.626895	0.068974	0.461094	0.469932
6	0.00702	0.348053	0.644928	0.058385	0.441017	0.500599
7	0.005629	0.351628	0.642742	0.051655	0.422806	0.5255
8	0.001546	0.34547	0.652981	0.046876	0.408024	0.5451
9	0.002197	0.295358	0.702445	0.043037	0.391213	0.565751
10	0.002609	0.275365	0.722026	0.040231	0.386409	0.573361
11	0.004474	0.234452	0.761074	0.036296	0.380638	0.583066
12	0.002842	0.206379	0.790778	0.03271	0.362167	0.605124
13	0.002719	0.222381	0.7749	0.030519	0.347535	0.621945
14	0.002884	0.223853	0.773263	0.026098	0.358139	0.615763

designed model and the regression model.

Table 2: Comparisons between designed model and regression model

Regression Coefficient	β_0	β_1	β_2
Designed Model	0.2	0.2	0.2
Regression Model	0.2	0.2001	0.1998

By treating each component as a whole, the method overcomes adverse effects from the multicollinearity among components, which therefore makes OLS effective in regression. At the same time, the method makes it easy to understand compositional data with different meanings in the model. Through converting the unit-sum constraint within components of compositional data into parameter constraint in the regression model, the method transforms multivariate compositional data regression into a constrained regression problem based on the space. The method can not only analyze the regression model of multivariate compositional data effectively, but also provided a fresh perspective for exploring multi-analysis of compositional data.

References

[1] Ferrers N M. An Elementary Treatise on Trilinear Coordinates[M]. London: Macmillan, 1866.

[2] Aitchison J. The statistical analysis of compositional data[M]. London: Chapman and Hall, 1986.

[3] Zhang Yaoting. The Statistical Analysis Generality of Compositional Data[M]. Beijing: Science Press, 2000.

[4] Wang Huiwen, Huang Wei. Linear regression model of compositional data[J]. System Engineering, 2003, 21(2): 102 -106.

[5] Wang Heiwen, Zhang Zhihui, Tenenhaus M. Multiple Linear Regression Modeling Method Based on the Compositional Data[J]. Journal of Management Sciences In China, 2006, 9(4): 27-32.

[6] Fang Kaitai, Quan Hui, Chen Qingyun. Practical Regression Analysis[M]. Beijing: Science Press, 1988.

Session

Lifetime data analysis

organized by Mei-Ling Ting Lee
(USA)

6th St.Petersburg Workshop on Simulation (2009) 649-653

Simulations for frailty models with general censoring and truncation

Catherine Huber[1]

The need to apply frailty models to analyze survival data arises when the assumption of a homogeneous population seems questionable. In order to model unobserved heterogeneity in the population one introduces a random effect into the model, called frailty, defined to act multiplicatively on the hazard rate $h(t|z)$ of an individual with covariate vector z. A frailty model therefore arises naturally from a Cox model, $h(t|z) = \exp(< \beta, z >)$ with unobserved covariates which materialize the frailty parameter η: $h(t|z) = \eta \exp(< \beta, z >)$. A frailty parameter η is also introduced to model dependence between survival times if the standard assumption of independence seems unrealistic. The frailty we consider here is meant to take into account a possible heterogeneity among the population. It is not a shared frailty as the individuals are assumed to be independent. We consider such a model when the data are both interval censored and truncated. Using several types of frailty distributions, gamma, log-normal and inverse Gaussian, we derive a process to estimate the coefficients β of the covariate z under simulated data.

[1]Universit? Paris Descartes, 45 rue des Saints-P?res, 75 006 Paris, France/ E-mail: catherine.huber@univ-paris5.fr

6th St.Petersburg Workshop on Simulation (2009) 650-654

Proportional Hazards and Threshold Regression: Their Theoretical and Practical Connections

Mei-Ling Ting Lee[1] and G. A. Whitmore[2]

Abstract

Proportional hazards (PH) regression is an established methodology for analyzing survival and time-to-event data. The proportional hazards assumption of PH regression, however, is not always appropriate and, thus, statistical researchers have explored many alternatives. Threshold regression (TR) is one of these alternative methodologies. The connection between PH regression and TR has been examined in previous published work but the investigations have been limited in scope. In this article, we study the connections between these two regression methodologies in greater depth and show that, in fact, PH regression is, for most situations, a special case of TR. We show two methods of construction by which TR models can yield PH functions for survival times, one based on altering the TR time scale and the other based on varying the TR boundary. We discuss how to estimate the TR time scale and boundary, with or without the PH assumption. Finally, we demonstrate the potential benefits of setting PH regression in the first-hitting-time context of TR regression. Simulation results will be presented.

[1]University of Maryland, College Park, MD. USA. E-mail:MLTLEE@UMD.EDU
[2]McGill University, Montreal, Quebec, Canada

6th St.Petersburg Workshop on Simulation (2009) 651-655

Trend tests for recurrent events with incomplete and misclassified information

Chrys Caroni[1]

Abstract

Undetected withdrawal of units from a study implies longer than expected times between the last recorded event and the end of the study. A test is given for this, using approximations to the sum of Beta random variables. The effect on trend tests is examined. The loss of power involved by dropping the time after the last event from the analysis is examined. Tests for use when only one event per unit can be recorded are also considered.

1. Time limited and failure limited data

The common theme of the topics considered in this paper is possible unreliability of follow up data in a study of times to events, particularly recurrent events such as failure times in a repairable system. We start with some observations on the data generating process.

The times of occurrence of a sequence of events in a non-homogeneous Poisson process with intensity $\lambda(t)$ can be generated as follows. The time t_1 from zero to the first event is generated by observing that the survival function $S(t) \sim U[0,1]$. But $S(t) = exp\left(-\Lambda(t)\right)$ where

$$\Lambda(t) = \int_0^t \lambda(\tau) d\tau$$

is the cumulative intensity. Consequently t_1 can be simulated as

$$t_1 = \Lambda^{-1}\left\{-\ln u_1\right\}$$

where u_1 is a random deviate from $U[0,1]$. Next, t_2 can be generated by observing that the same argument applies to the conditional survival function $S(t_2|t_1) = \exp\left\{\Lambda(t_2) - \Lambda(t_1)\right\}$, so that $\Lambda(t_2) - \Lambda(t_1) = -\ln u_2$ where u_2 is another random deviate from $U[0,1]$. Hence $t_2 = \Lambda^{-1}\left\{\Lambda(t_1) - \ln u_2\right\}$ and so on for $t_3, t_4, \ldots$. In the particular case of the power law intensity $\lambda(t) = \alpha\beta t^{\beta-1}$ this gives

$$t_i = \left\{t_{i-1}^{\beta} - \alpha^{-1} \ln u_i\right\}^{1/\beta} \quad \text{with} \quad t_0 = 0.$$

[1]National Technical University of Athens, Greece, E-mail: ccar@math.ntua.gr

Time limited data where observation is limited to the interval $(0, T]$ can be obtained by generating event times until $t_{n+1} > T$ while $t_n \leq T$, giving n events within the interval. In generating *failure limited* data, careful attention should be paid to what is meant by this term. The strict definition of failure limited data is that the number of events n is fixed in advance, so that the total observation time $T = t_n$ is random. (For time limited data, T is fixed and n is random.) However, the term is often applied to data that consist of the reported event times and therefore end in a failure, but without pre-determination of the number of failures that would be recorded. This is another form of time limited data, because the failures are those that occurred in the time available for the study. Thus n continues to be random. Data ending in a failure in this way have different properties from strictly failure limited data; for example, inter-event times in a homogeneous Poisson process are negatively correlated instead of being independent.

The data generation process has important consequences under some circumstances. For example, simulation shows that Kvaloy and Lindqvist's tests for trend based on total time on test in multiple systems [2] do not have the claimed size unless "failure limited" data have been generated under the strict definition.

If an observation is made only when an event happens, how should we treat the time remaining until the end of the study period after the last event recorded for a subject or unit? Let the time on study of unit i be T_i, with n_i events recorded at ordered times $t_{i1} < t_{i2} \cdots < t_{in_i}$. The remaining time $r_i = T_i - t_{in_i}$ appears to be a right-censored observation of the time until event $n_i + 1$. However, if it is possible that the unit has left the study at some time following the last event (e.g. if machines are being taken elsewhere for repair) then event $n_i + 1$ could actually have happened, although it has not happened where it could be observed.

2. Testing for excessively large remaining times

Assume a homogeneous Poisson process (HPP) with intensity λ_i for unit i. The problem is to test whether the remaining times r_i after the last event in each unit are "too large" compared to the observed inter-failure times.

For a single unit, the theory is well known (e.g. [3]). The following analysis extends the theory to a set of units. The n_i event times t_{ij} of unit i in the interval $(0, T_i)$ are uniformly distributed in an HPP and, by symmetry, the remaining time r_i after the last event has the same distribution as the time t_{i1} from the beginning of observation until the first event. Rescaling to $s_i = r_i/T_i$, this distribution conditional on n_i is Beta $(1, n_i)$. An obvious test statistic combining information from all units is $S = \sum_{i=1}^{k} s_i$, which is distributed as the sum of k independent Beta statistics. Unfortunately, the sum of Beta random variables does not have a simple distribution. However, because the moments can be expressed very simply, it is possible to approximate the distribution by equating moments. Following [1], the distribution of S can be approximated by taking

$$S = kZ, \quad Z \sim \text{Beta}(e, f)$$

where

$$f = E(1-E)^2/V - (1-E)$$
$$e = f(1-E)^{-1}$$

with

$$kE = \sum 1/(1+n_i)$$

the sum of the expected values of the S_i and k^2V equal to the sum of the variances of the S_i.

Simulation results obtained under the null hypothesis of the HPP (Table 1) indicate that the true size of the test is rather well approximated by this method.

Table 1: Simulated exceedance probabilities under the null hypothesis of nominal percentage points obtained from the Beta approximation to the distribution of the statistic S. $E(N)$ is the expected number of failures per unit. 10,000 simulations per combination

Units	$E(N)$	1%	5%	10%
5	10	0.94	4.89	9.88
10	5	0.82	4.64	9.29
5	20	1.09	5.30	10.13
10	10	0.79	4.87	9.89
20	5	0.87	4.49	9.50

Table 2: Simulated powers using nominal 1% and 5% percentage points obtained from the Beta approximation to the distribution of the statistic S. p is the probability of withdrawal from the study after any event

Units	$E(N)$	p	1%	5%
20	5	0.1	41.22	67.14
20	5	0.05	14.74	34.62
10	5	0.05	7.81	23.21
10	10	0.05	59.51	79.85
5	10	0.05	32.48	57.11

To examine the powers, it is necessary to specify an alternative data generation mechanism. The results in Table were obtained by the following process:

1. Generate the first event time t_1
2. Generate a uniform random variable u. If $u < p$ where p is a chosen small probability, then cease generating for this unit and go to the next unit
3. Otherwise, generate event time t_2. If $t_2 > T$, the chosen period of study, then go to the next unit and start again at step 1
4. Repeat step 2.

This simulates the early withdrawal of the unit from the study. Note that the withdrawal of a unit is unknown to the investigator, who believes that all the remaining time until the end of the study represents a right-censored observation for that unit.

3. Allowing for unreliability in the time remaining

If the values of the r_i's tend to be too large, this may have an important effect on statistical analyses. Table 3 shows that tests for trend can indeed be very sensitive to errors introduced into the data in the way described above.

Table 3: Simulated exceedance probabilities (%) of 5% significance level tests for trend when units may have withdrawn from the study. MH, combined Military Handbook test; AD Anderson-Darling test [2]. p is the probability of withdrawal from the study after any event

Units	$E(N)$	p	MH	AD
10	5	0.1	13.4	17.1
10	5	0.05	7.89	9.34
10	10	0.01	6.05	7.09
10	20	0.01	13.8	17.6

To allow for unreliability of the remaining time intervals r_i, the data can be treated as ending with failures at the times t_{in_i} instead of as time limited at times T_i. However, if it was not in fact necessary to remove the right-censored intervals, power will have been lost by not using the information in these intervals. The question is, how much?

Table 4: Relative powers (%) of combined Military Handbook test and Anderson-Darling test for trend, with 5 or 10 units and a total expected number of 60 failures. Data generated under power-law model with $\beta = 1.5$. Test statistics calculated treating data as time limited (T) and terminating in failure (F). 10,000 simulations

Units x events	Combined Military Handbook		Anderson-Darling	
	T	F	T	F
4×5, 1×40	85.9	82.0	97.0	96.5
5×12	85.7	82.0	77.7	73.6
10×6	85.8	78.2	77.6	71.3
8×4, 2×14	80.6	79.7	89.8	86.9

Table 4 presents a selection of simulation results for time limited data generated from non-homogeneous Poisson processes with power law intensity $\beta\lambda^{\beta-1}$ with $\beta = 1.5$, for 5 or 10 units with a total expected number of failures equal to 60.

The combined Military Handbook test and the Anderson-Darling test statistic of [2] were calculated both for the time limited original data and the data obtained by treating each unit's data as terminating in failure at its last event time.

It can be seen that the loss of power in these circumstances is small, compared to the risk of distortion of tests by including unreliable remaining time intervals.

4. Units with no events

Because the Beta statistic depends on time remaining after the last event, it cannot include units without events. To include such units, an obvious way to proceed is to construct a randomization test, assuming equal intensities. If n events in total have been observed in the k units, then simulate the allocation of n balls to k urns with unequal probabilites proportional to T_i and thus simulate the distribution of total time on test of the units with no events (empty urns).

The brief discussion so far has been in terms of recurrent events (repairable systems). However, similar points apply to studies that examine the time to a single event (death, first breakdown, and so on) for each unit. Undetected withdrawal in this case implies that the time occupied by the apparently right-censored observations (units where no event occured by the end of the study), is larger than expected compared to the observed lifetimes.

References

[1] Johannesson, B., Giri, N. (1995). On approximations involving the Beta distribution. *Commun. Statist. Simul. Comp.*, **24**, 489-503.

[2] Kvaloy, J.T., Lindqvist, B.H. (1998). TTT-based test for trend in repairable systems data. *Reliab. Eng. Syst. Saf.*, **60**, 13-28.

[3] Solow, A.R. (1993). Inferring extinction from sighting data. *Ecology*, **74**, 962-964.

6th St.Petersburg Workshop on Simulation (2009) 656-660

Analysis of Event History Data With Covariate Missing Not At Random

Joan Hu[1]

Abstract

Motivated by a study for disease control, we consider estimation under the Cox proportional hazards model based on a set of right-censored survival times with missing covariates, where the missing mechanism is not independent of the missing data conditional on the observed data. We present a likelihood based estimation procedure with the current data supplemented with some readily available information. The medical study that motivated this research is used throughout the talk for illustration.

[1]Simon Fraser University, Canada E-mail:joanh@stat.sfu.ca

6th St.Petersburg Workshop on Simulation (2009) 657-661

Inverse Gaussian family and its applications in reliability. Study by simlulation

M. Nikulin[1], N. Saaidia[2]

1. Introduction

The inverse Gaussian distribution (IGD) is named so because it satisfies the relationship with normal distribution. The IGD was introduced first by Schrodinger in 1915 [Seshadri 1993], and its found to have many applications in a various fields such biology, economics, cardiology, demography, linguistic, etc. See for examples [Seshadri 1993], [Seshadri 1999], [Chhikara and Folks, 1989], [Voinov and Nikulin 1993],[Lawless 2003] for more details. The IGD is very competitor of Weibull, generalized Weibull and de lognormal distribution.
We study the possibilities (Outlooks) of applications of the IGD in reliability and survival analysis, and we study by simulation the proprieties of dynamic regression models based on the family of IGD.

2. Properties and Characterization

Let as consider $\mathbb{X} = (X_1, X_2, \cdots, X_n)^T$ be n independent and identically distributed (iid) random variables, Wa say that X_i fellow the inverse gaussian distribution (and we note $X_i \sim IG(\mu, \lambda)$) if the density function is defined by:

$$f(x, \mu, \lambda) = \left(\frac{\lambda}{2\pi x^3}\right)^{\frac{1}{2}} exp\{-\frac{\lambda(x-\mu)^2}{2\mu^2 x}\}, \quad x \geq 0, \quad \mu > 0, \quad \lambda > 0. \qquad (1)$$

where μ is the mean and λ is the shape parameter.

This function is unimodal. We can show that

$$E(X_i) = \mu \ , Var(X_i) = \frac{\mu^3}{\lambda}.$$

[1]IMB, Université Victor Segalen Bordeaux 2, France. E-mail: ierrr01@ie.technion.ac.il

[2]IMB, Université Victor Segalen Bordeaux 2, France & and Université Badji Mokhtar, Annaba, Algérie

3. Parameters estimation

The likelihood function $L(\mathbb{X}, \theta)$, with $\theta = (\mu, \lambda)^T$ of a sample $\mathbb{X} = (X_1, X_2, \cdots, X_n)^T$ is given by the following formula:

$$L(\mathbb{X}, \theta) = \prod_{i=1}^{n} f(X_i, \theta) = \lambda^{\frac{n}{2}} (2\pi)^{\frac{n}{2}} \prod_{i=1}^{n} X_i^{-\frac{3}{2}} exp\left(\sum_{i=1}^{n} \frac{-\lambda(X_i - \mu)^2}{2\mu^2 X_i} \right). \quad (2)$$

Let as $V = \sum_{i=1}^{n} (X_i^{-1} - \overline{X}^{-1})$, then the MLE's estimators of μ and λ are successively:

$$\begin{cases} \hat{\mu} = \overline{X} \\ \hat{\lambda} = \dfrac{n}{\sum_{i=1}^{n} (X_i^{-1} - \overline{X}^{-1})} = \dfrac{n}{V} \end{cases} \quad (3)$$

The MVUE's estimators of μ and λ are successively:

$$\begin{cases} \hat{\mu} = \overline{X} \\ \hat{\lambda} = \dfrac{n-1}{V} \end{cases} \quad (4)$$

Notice that for the family $IG(\mu, \lambda)$, the statistics $\overline{X}$ and V are independent, and the bivariate statistic $T = \left(\overline{X}, V\right)^T$ is minimal sufficient and complete, see for example [Voinov and Nikulin 1993], [Seshadri 1993]

4. Chi-squared goodness-of-fit test for the IGD

Consider the problem of testing the hypothesis H_0 that the distribution of the sample $\mathbb{X} = (X_1, X_2, \cdots, X_n)^T$ of n (idd) random variables belongs to the family $\{IG(\mu, \lambda)\}$, which

$$H_0 : P(X_i \leq x) = F(x, \theta), \quad (5)$$

where $F(x, \theta)$ is the distribution function of (1).
Dividing the real line into r intervals : $I_1, I_2, \cdots, I_r$ mutually disjoint, by the points:

$$0 = a_0 < a_1 < \cdots < a_{r-1} < a_r = +\infty,$$

and grouping the sample $\mathbb{X} = (X_1, X_2, \cdots, X_n)^T$ over the intervals $I_1, I_2, \cdots, I_r$. It results the vector of frequencies $\nu = (\nu_1, \nu_2, \cdots, \nu_r)^T$.

Suppose that: $p_i = \int_{a_{i-1}}^{a_i} f(x, \theta) dx,\ i = 1, 2, \cdots, r,$
and we denoted by $X_n(\theta)$ the vector with components $\dfrac{\nu_i - np_i}{\sqrt{np_i}}, i = 1, 2, \cdots, r.$
The Fisher information of the sample $\mathbb{X}$ is:

$$I_n(\theta) = ni(\theta) = \begin{pmatrix} \frac{n\lambda}{\mu^3} & 0 \\ 0 & \frac{n}{2\lambda} \end{pmatrix}. \quad (6)$$

We consider the RRN statistic Y_n^2 proposed by [Nikulin 1973] et [Rao and Robson 1974](see, also [Drost 1988], [van der vaart 1998]):

$$Y_n^2(\hat{\theta}_n) = X_n^2(\hat{\theta}_n) + \frac{1}{n} l^T(\hat{\theta}_n)(i(\hat{\theta}_n) - J(\hat{\theta}_n))^{-1} l(\hat{\theta}_n) \tag{7}$$

with
$J(\theta) = B(\theta)^T B(\theta)$ is the Fisher information of the vector of frequencies ν, where:

$$B(\theta) = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ \vdots & \vdots \\ b_{r1} & b_{r2} \end{pmatrix}; \;\; b_{i1}(\theta) = \frac{1}{\sqrt{p}_i} \frac{\partial p_i}{\partial \mu} \;\; , \;\; b_{i2}(\theta) = \frac{1}{\sqrt{p}_i} \frac{\partial p_i}{\partial \lambda}, \;\; i = 1, 2, \cdots, r.$$

$$l(\theta) = (l_1(\theta), l_2(\theta))^T; \;\; l_1(\theta) = \sum_{i=1}^{r} \frac{\nu_i}{p_i} \frac{\partial p_i}{\partial \mu}(\theta), \;\; l_2(\theta) = \sum_{i=1}^{r} \frac{\nu_i}{p_i} \frac{\partial p_i}{\partial \lambda}(\theta).$$

Under H_0, the statistic Y_n^2 possesses the chi-squared distribution χ^2_{r-1} of $r-1$ degrees of freedom in the limit [Greenwood and Nikulin 1996].

Case 1. $\theta = (\mu, \lambda)^T$ is known.

In this case, the statistic

$$Y_n^2 = X_n(\theta)^T X_n(\theta) = X_n^2(\theta) = \sum_{i=1}^{r} \frac{(\nu_i - np_i)^2}{np_i}, \tag{8}$$

possesses the chi-squared distribution χ^2_{r-1} of $r-1$ degrees of freedom in the limit.

For a given level α, if we choose the critical value:

$$C_\alpha = \chi^2_{r-1,1-\alpha},$$

the hypothesis H_0 is accepted if $Y_n^2 \leq C_\alpha$, otherwise H_0 is rejected.

Case 2. One of the parameters is unknown.

Then we have

$$Y_n^2 = \begin{cases} X_n^2(\hat{\lambda}_n) + \frac{1}{n} \dfrac{\left(\sum_{i=1}^{r} \frac{\nu_i}{p_i} \frac{\partial p_i}{\partial \lambda}(\hat{\lambda}_n) \right)^2}{\frac{1}{2\hat{\lambda}_n} - \sum_{i=1}^{r} \frac{1}{p_i} \left(\frac{\partial p_i}{\partial \lambda}(\hat{\lambda}_n) \right)^2} \;\; , \\ \qquad \text{if } \; \mu \; \text{ is known and } \lambda \; \text{ is unknown} \\ X_n^2(\hat{\mu}_n) + \frac{1}{n} \dfrac{\left(\sum_{i=1}^{r} \frac{\nu_i}{p_i} \frac{\partial p_i}{\partial \mu}(\hat{\mu}_n)) \right)^2}{\frac{\lambda}{\hat{\mu}^3} - \sum_{i=1}^{r} \frac{1}{p_i} (\frac{\partial p_i}{\partial \mu}(\hat{\mu}_n))^2} \;\; , \\ \qquad \text{if } \; \mu \; \text{ is unknown and } \lambda \; \text{ is known} \end{cases} \tag{9}$$

For a given level α, if we choose the critical value:

$$C_\alpha = \chi^2_{r-1,1-\alpha},$$

the hypothesis H_0 is accepted if $Y_n^2 \leq C_\alpha$, otherwise H_0 is rejected.

Case 3. θ is unknown.

The MLE's of $\widehat{\theta} = (\widehat{\mu}, \widehat{\lambda})^T$ is given by (3). The statistic (7) can be written as:

$$Y_n^2(\hat{\theta}_n) = X_n^2(\hat{\theta}_n) + \frac{1}{n \mid M \mid} \left[\left(\frac{1}{2\hat{\lambda}_n^2} - \sum_{i=1}^{r} \frac{1}{p_i} \omega_{i2}^2(\hat{\theta}_n) \right) \left(\sum_{i=1}^{r} \frac{\nu_i}{p_i} \omega_{i1}(\hat{\theta}_n) \right)^2 \right] +$$

$$2 \left[\left(\sum_{i=1}^{r} \frac{1}{p_i} \omega_{i1}(\hat{\theta}_n) \omega_{i2}(\hat{\theta}_n) \right) \left(\sum_{i=1}^{r} \frac{\nu_i}{p_i} \omega_{i1}(\hat{\theta}_n) \right) \left(\sum_{i=1}^{r} \frac{\nu_i}{p_i} \omega_{i2}(\hat{\theta}_n) \right) \right] +$$

$$\left[\left(\frac{\widehat{\lambda}_n}{\widehat{\mu}_n^3} - \sum_{i=1}^{r} \frac{1}{p_i} \omega_{i1}^2(\hat{\theta}_n) \right) \left(\sum_{i=1}^{r} \frac{\nu_i}{p_i} \omega_{i2}(\hat{\theta}_n) \right)^2 \right], \tag{10}$$

where

$$M = i(\hat{\theta}_n) - J(\hat{\theta}_n) = \begin{pmatrix} \frac{\hat{\lambda}_n}{\hat{\mu}_n^3} - \sum_{i=1}^{r} b_{i1}^2 & -\sum_{i=1}^{r} b_{i1} b_{i2} \\ -\sum_{i=1}^{r} b_{i1} b_{i2} & \frac{1}{2\hat{\lambda}_n^2} - \sum_{i=1}^{r} b_{i2}^2 \end{pmatrix}, \tag{11}$$

$$\omega_{i1}(\hat{\theta}_n) = \frac{\partial p_i}{\partial \mu}(\hat{\theta}_n) \quad and \quad \omega_{i1}(\hat{\theta}_n) = \frac{\partial p_i}{\partial \lambda}(\hat{\theta}_n),$$

and $\mid M \mid$ is the determinant of the matrix M.

For a given level α, if we choose the critical value:

$$C_\alpha = \chi^2_{r-1,1-\alpha},$$

the hypothesis H_0 is accepted if $Y_n^2 \leq C_\alpha$, otherwise H_0 is rejected.

5. Chi-squared goodness-of-fit test for the IGD for randomly censored data

In Reliability and survival analysis, we often encounter incomplete observations, and in this situation the usual methods are no longer valid. In the case of random censorship, we use the Pearson-type chi-squared statistics $\hat{Q}_n$ proposed by [Habib and Thomas (1996)] which compare the Kaplan-Meier estimate $\hat{S}_n(x)$ to the parametric $S(x, \hat{\theta}_n)$.

Let as $(\mathbb{X}, \Delta) = (X_1, \delta_1), (X_2, \delta_2) \cdots, (X_n, \delta_n)$ be n independent and identically distributed random variable, which $X_i = min(T_i, C_i)$, T_i is the failure time, C_i is the censure and δ_i is the indicator function defined as:

$$\delta_i = \begin{cases} 1 & , \ if \quad T_i \leq C_i \\ 0 & , \ otherwise \end{cases} .$$

Consider the problem of testing the hypothesis H_0 that the distribution of the sample $(\mathbb{X}, \Delta)$ belongs to the family $\{IG(\mu, \lambda)\}$ which:

$$H_0 : P(X_i > x) = S(x, \theta),$$

where $S(x, \theta) = 1 - F(x, \theta)$ is the survival function (or reliability function) of IGD.

Habib et Thomas (1996) have shown that $\sqrt{n}\left(\hat{S}_n(x) - S(x, \hat{\theta}_n)\right)$ converges to the Gaussian process under the hypothesis H_0 .

We Divide the real line into r intervals: $I_1, I_2, \cdots, I_r$ mutually disjoint by the points:

$$0 = a_0 < a_1 < \cdots < a_{r-1} < a_r = +\infty.$$

Let as the victor

$$\hat{Z}_n = \sqrt{n}\left(\hat{S}_n - S_{\hat{\theta}_n}\right),$$

where $\hat{S}_n = (\hat{S}_n(a_1), \hat{S}_n(a_2), \cdots, \hat{S}_n(a_{r-1}))^T$ and $S_{\hat{\theta}_n} = (S(a_1, \hat{\theta}_1), S(a_2, \hat{\theta}_2), \cdots, S(a_{r-1}, \hat{\theta}_n))^T$.

The statistic

$$\hat{Q}_n = \hat{Z}_n^T \hat{\Sigma}^{-1} \hat{Z}_n,$$

where $\hat{\Sigma}$ is the estimator for the covariance matrix Σ (see [Habib and Thomas 1996]) possesses the chi-squared distribution χ^2_{r-1} of $r-1$ degrees of freedom in the limit.

For a given level α, if we choose the critical value:

$$C_\alpha = \chi^2_{r-1,1-\alpha},$$

the hypothesis H_0 is accepted if $\hat{Q}_n \leq C_\alpha$, otherwise H_0 is rejected.

We Note that for uncensored data, the statistic $\hat{Q}_n$ reduces to that proposed by Rao-Robson-Nikulin Y_n^2. We can also consider the case of doubly censored data (see [Ionescu and Lmmnios 1999]).

6. Dynamic regression models based on the family of IGD and their applications in reliability and survival analysis

In this section, we consider two survival flexible regression models $\{S_{x(\cdot)}, x(\cdot) \in E\}$ on a set E of all admissible stresses,where $S_{x(\cdot)}$ is the *survival* function of $T = T_{x(\cdot)}$ given $x(\cdot) = (x_1(\cdot), ..., x_m(\cdot))^T$.

We remind that the *survival function* and the *hazard rate function* given $x(\cdot)$ are :

$$S_{x(\cdot)}(t) = \mathbf{P}\left(T_{x(\cdot)} \geq t \mid x(s); 0 \leq s \leq t\right), \quad \lambda_{x(\cdot)}(t) = -\frac{S'_{x(\cdot)}(t)}{S_{x(\cdot)}(t)}.$$

We have the *simple frailty model* if the hazard rate of an individual is influenced by a *non-observable* positive random variable Z, called the *frailty variable*, in the manner:

$$\lambda_{x(\cdot)}(t|Z=z) = z\lambda_0(t)r\{x(t)\}, \quad x(\cdot) \in E,$$

where $\lambda_0(\cdot)$ is the hazard rate of the baseline distribution. If we suppose that the density of *frailty variable* Z belongs to the family of *inverse Gaussian distributions* $IG(\mu, \lambda)$, then we obtain the so called *inverse Gaussian frailty model on* E.

Now we consider the famous *accelerated failure time* (AFT) model on E, (see Bagdonavičius and Nikulin (2002)). The AFT model is the most used in *accelerated trials*.

We say that the family $\{S_{x(\cdot)}, x(\cdot) \in E\}$ of survival function on E *forms the* AFT *model on* E if there exists a positive function $r : E \to \mathbf{R}^1$ and a baseline survival function S_0 such that the elements of the family $\{S_{x(\cdot)}, x(\cdot) \in E\}$ verifies the next relation:

$$S_{x(\cdot)}(t) = S_0\left(\int_0^t r(x(u))du\right), x(\cdot) \in E,$$

where the so-called baseline function S_0 does not depend on $x(\cdot)$.

Very often is supposed that baseline distribution in the AFT model and the law of the frailty variable in the frailty model are from the family lognormal, Gamma, Weibull, and Generalized Weibull failure-time distributions. In these cases the Inverse Gaussian family of distributions is an interesting competing family for these 4 families. We study the properties of the Frailty and the AFT models when they are based on the Inverse Gaussian family of distributions. At the end we consider the degradation model based on the family of IGD.

References

[1] Bagdonavičius, V., Nikukin, M., (2002) Accelereted Life Models: Modeling and Statistical Analysis. Chapman and Hall.

[2] Chhikara, R.S., Folks, J.L., (1989). The Inverse Gaussian Distribution. Marcel Dekker, New York..

[3] Drost F. (1988), Asymptotics for Generalized Chi-squared Goodness-of-fit Tests, amsterdam: Centre for Mathematics and Computer Sciences,CWI Tracs, 48.

[4] Greenwood, P. S. and Nikulin, M. (1996), A guide to Chi-squared Testing, John Wiley and Sons, New York.

[5] Habib, M.G., Thomas, D.R. (1996). Chi-squared Goodness-of-Fit Tests For Randomly Censored Data. Annals of Statistics, 14, 759-765.

[6] Huber, C., Limnios, N., Mesbah, M., Nikulin, M.,(Eds) (2008) Mathematical Methods in Survival analysis, Reliability and Quality of Life, Willy-ISTE

[7] Ionescu D.C, Lmmnios, N. (Eds) (1999) Statistical and Probabilistic Models in Reliability, Birkhauser Boston.

[8] Lawless, J.F. (2003) Statistical Models and Methods for Lifetime Data, 2nd ed.,New York: John Wiley.

[9] Seshadri, V., (1993), The Inverse Gaussian Distribution: A Case Study in Exponential Families, Clarendon Press:Oxford,

[10] Seshadri, V., (1999), The inverse Gaussian Distribution: statistical theory and applications. Springer,New York

[11] Nikulin M.S.(1973),On a Chi-square Test For Continuous Distributions,Thery of Probability and Applications, 18 , p.638-639.

[12] Nikulin M.S. (1973), Chi-square Test For Continuous Distributions with Shift and Scaie Parameters," Teor. Veroyatn.Primen., 18, No. 3, 559-568

[13] van der vaart, A.W., (1998) Asymptotic Statistics. Cambridge Series in Statistics and probabilistic Mathematics. Cambridge: Cambridge University Press.

[14] Voinov, M., Nikulin,M. (1993) Unbiased Estimators and Their Applications, V.1 Univariate case, Dordrecht: Kluwer Academic Publishers.

[15] Rao, K.C. and Robson, D.S. (1974) A chi-square statistic for goodness-of-fit tests within the exponential family. Commun. Statist. 3 1139-1153

Session

Recent advances in change-point analysis

organized by Ansgar Steland (Germany)

6th St.Petersburg Workshop on Simulation (2009) 667-671

Sequential change-point analysis for renewal counting processes

Allan Gut[1], Josef G. Steinebach[2]

Abstract

The standard approach in change-point theory is to base the statistical analysis on a sample of fixed size. Alternatively, one observes some random phenomenon sequentially and takes action as soon as one observes some statistically significant deviation from the "normal" behaviour. In [2], we introduced some (truncated) sequential testing procedures for detecting a change-point in a sequence of renewal counting data. In the present note, we first review some of these results and and discuss some recent work (cf. [3]), in which we look in more detail into the behaviour of the relevant stopping times under alternatives, in particular the time it takes from the actual change-point until the change is detected.

1. Introduction

In [2], we suggested some truncated sequential monitoring procedures for detecting a structural break ("change-point") in a series of counting data, e.g., the number of claims of an insurance portfolio, which are sequentially observed at equidistant time-points up to a "truncation point" (say) n, i.e., we have a "closed-end" procedure. Some limiting extreme value asymptotics (as $n \to \infty$) could be derived in [2] under the null hypothesis of "no change", thus allowing for a choice of the critical boundaries in the monitoring schemes such that the false alarm rate (asymptotically) attains a prescribed level α. Moreover, some limiting properties under the alternative could also be proved showing that the statistical procedures have asymptotic power 1. The present note reviews some of the results from [2] and also discusses some recent work (cf. [3]), in which we look in more detail into the behaviour of the relevant stopping times, in particular the time it takes from the (unknown) change-point until one detects that a change actually has occurred, in other words, asymptotics for stopping times under alternatives are proved.

As in [2], we observe counting data $N(0), N(1), \ldots, N(n)$ at time-points $t = 0, 1, \ldots, n$, where $\{N(t)\}_{0 \le t \le n}$ is a renewal counting process with drift coefficient θ and variance parameter η^2 up to some (unknown) change-point k_n^*, after which it

[1]Uppsala University, E-mail: allan.gut@math.uu.se
[2]University of Cologne, E-mail: jost@math.uni-koeln.de

changes to an independent second renewal counting process with drift coefficient θ^* and variance parameter ${\eta^*}^2$. More precisely, we assume that $\{N(t)\}_{0\le t\le n}$ has the following structure:

$$N(t) = \begin{cases} N_0(t), & \text{for} \quad 0 \le t \le k_n^*, \\ N_0(k_n^*) + N_1(t - k_n^*), & \text{for} \quad k_n^* < t \le n, \end{cases} \tag{1}$$

with

$$N_0(t) = \min\{k : \sum_{i=1}^{k} X_i > t\}, \quad N_1(t) = \min\{k : \sum_{i=1}^{k} X_i^* > t\}, \quad t \ge 0,$$

and independent sequences $\{X_i\}_{i=1,2,\dots}$ and $\{X_i^*\}_{i=1,2,\dots}$ of i.i.d. r.v.'s satisfying

$$EX_1 > 0,\ EX_1^* > 0,\ EX_1 \ne EX_1^*;\ \mathrm{Var}\,(X_1) > 0,\ \mathrm{Var}\,(X_1^*) > 0. \tag{2}$$

On setting $\theta = 1/\mu$, $\eta^2 = \sigma^2/\mu^3$, $\theta^* = 1/\mu^*$, ${\eta^*}^2 = {\sigma^*}^2/{\mu^*}^3$, where $\mu = EX_1$, $\mu^* = EX_1^*$, $\sigma^2 = \mathrm{Var}\,(X_1)$, ${\sigma^*}^2 = \mathrm{Var}\,(X_1^*)$, we want to test, for example, the null hypothesis

$$H_0 : k_n^* = n \quad (\text{"no change"})$$

versus the "two-sided" alternative

$$H_1 : 1 \le k_n^* < n,\ \theta^* \ne \theta \quad (\text{"change in the drift at } k_n^*\text{"}),$$

taking sequentially into account the observed counting data $N(0), N(1), \dots, N(n)$.

Our asymptotic results below are based on the following strong invariance principle (cf. [2], Proposition 3.1), which shows that, under an r-th moment condition (with some $r > 2$), the counting process $\{N(t)\}_{0\le t\le n}$ above can almost surely (a.s.) be approximated (with a rate $o(n^{1/r})$) by a Gaussian process $\{V(t)\}_{0\le t\le n}$ possessing a similar structure, i.e., also having a drift coefficient θ and variance parameter η^2 up to the change-point k_n^*, and changing thereafter to an independent second Gaussian process with drift coefficient θ^* and variance parameter ${\eta^*}^2$:

Proposition 1. *Assume that* $E|X_1|^r < \infty$ *and* $E|X_1^*|^r < \infty$ *for some* $r > 2$. *Then*

$$\sup_{0\le t\le n} |N(t) - V(t)| \overset{a.s.}{=} o(n^{1/r}), \tag{3}$$

with

$$V(t) = \begin{cases} t\theta + \eta W_0(t), & \textit{for} \quad 0 \le t \le k_n^*, \\ V(k_n^*) + (t - k_n^*)\theta^* + \eta^* W_1(t - k_n^*), & \textit{for} \quad k_n^* < t \le n, \end{cases} \tag{4}$$

where θ, η^2; $\theta^*, {\eta^*}^2$ *are as in* (2), *and where* $\{W_0(t),\ t \ge 0\}$, $\{W_1(t),\ t \ge 0\}$ *are two independent (standard) Wiener processes.*

Remark 1. *Although our analysis is based on the strong approximation above, a weak invariance principle, which, in turn, is available for a much wider class of stochastic processes, would have been sufficient (cf. [4], Section 1).*

In Section we review, for the reader's convenience, some results under the null hypothesis from [2], before we discuss some recent work on the asymptotic normality of stopping times under the alternative in Section . In Section , we add some concluding remarks showing that, based on the sequential monitoring, an asymptotic confidence interval for the change-point k_n^* can also be obtained.

Since our results are based on the strong invariance principle of Proposition 1, we tacitly assume throughout in the following that the conditions required for the application of (3) *and* (4) *are fulfilled.*

2. Null asymptotics

From the sequential observations $N(0), N(1), \ldots, N(n)$, we compute the variables

$$Y_k = Y_{k,n} = \frac{N(k) - N(k-h_n) - h_n\theta}{\eta\sqrt{h_n}}, \quad k = h_n, \ldots, n,$$

$$Z_k = Z_{k,n} = \frac{N(k) - k\theta}{\eta\sqrt{k}}, \quad k = k_n, \ldots, n,$$

and the stopping times

$$\tau_n^{(1)} = \min\{h_n \le k \le n : |Y_k| > c_n^{(1)}\},$$

$$\tau_n^{(2)} = \min\{k_n \le k \le n : |Z_k| > c_n^{(2)}\}$$

($\min\emptyset := +\infty$), where $c_n^{(1)}, c_n^{(2)}$ are suitable critical values and h_n, k_n are the lengths of the respective "training periods".

Remark 2. *For the sake of simplicity, we assume that the "in-control" parameters θ, η are known, but they can also be replaced by "suitable" sequential estimates (see [2], Section 5).*

The critical values $c_n^{(1)}, c_n^{(2)}$ are chosen such that the false alarm rates (asymptotically) attain a prescribed level α, i.e.,

$$P_{H_0}(\tau_n^{(1)} < \infty) = P_{H_0}\big(\max_{h_n \le k \le n} |Y_k| > c_n^{(1)}\big) \approx \alpha, \qquad \text{and}$$

$$P_{H_0}(\tau_n^{(2)} < \infty) = P_{H_0}\big(\max_{k_n \le k \le n} |Z_k| > c_n^{(2)}\big) \approx \alpha,$$

which can be achieved via the following extreme value asymptotics from [2], Section 4:

Theorem 1. *If $h_n \ll n$, but $\sqrt{h_n} \gg n^{1/r}$, then, under H_0, with normalizations $a_n^{(1)} = \sqrt{2\log(n/h_n)}$ and $b_n^{(1)} = 2\log(n/h_n) + \frac{1}{2}\log\log(n/h_n) - \frac{1}{2}\log\pi$,*

$$a_n^{(1)} \max_{h_n \le k \le n} |Y_k| - b_n^{(1)} \xrightarrow{d} E \quad as \quad n \to \infty,$$

where $P(E \le x) = \exp(-2e^{-x})$, $x \in \mathbb{R}$, that is, the critical value $c_n^{(1)}$ can (asymptotically) be chosen as

$$c_n^{(1)} = \frac{E_{1-\alpha} + b_n^{(1)}}{a_n^{(1)}} \quad \big(\sim \sqrt{2\log(n/h_n)} \,\big),$$

where $E_{1-\alpha}$ denotes the $(1-\alpha)$-quantile of the (two-sided) Gumbel distribution.

Theorem 2. *If $k_n \ll n$, but $\sqrt{k_n} \gg n^{1/r}$, then, under H_0, with normalizations $a_n^{(2)} = \sqrt{2\log\log(n/k_n)}$ and $b_n^{(2)} = 2\log\log(n/k_n) + \frac{1}{2}\log\log\log(n/k_n) - \frac{1}{2}\log(4\pi)$,*

$$a_n^{(2)} \max_{k_n \le k \le n} |Z_k| - b_n^{(2)} \xrightarrow{d} E \quad as \quad n \to \infty,$$

that is, the critical value $c_n^{(2)}$ can (asymptotically) be chosen as

$$c_n^{(2)} = \frac{E_{1-\alpha} + b_n^{(2)}}{a_n^{(2)}} \quad \big(\sim \sqrt{2\log\log(n/k_n)} \,\big),$$

with E and $E_{1-\alpha}$ as in Theorem 1.

Remark 3. *The asymptotics of Theorems 1 and 2 retain, if the "in-control" parameters θ, η are replaced by suitable sequential estimates (cf. [2], Section 5).*

Now, the question is how quickly a possible change-point k_n^* can be detected by the monitoring procedure, that is, what can be said about the behaviour of the stopping times $\tau_n^{(1)}, \tau_n^{(2)}$ or the detection delays $\tau_n^{(1)} - k_n^*, \tau_n^{(2)} - k_n^*$ under the alternative H_1 ? In the next section it will turn out that the limiting distributional behaviour of the stopping times, suitably normalized, is an asymptotically normal one.

3. Asymptotics under the alternative

Similar to [1], we consider an "early change" scenario here, that is, we assume that the change-point k_n^* does not occur too late compared to the length of the training period in the following technical sense:

$$k_n^* = \mathcal{O}\big(h_n \log^\gamma(n/h_n)\big) \quad \text{as} \quad n \to \infty \quad (\text{for some } \gamma > 0). \tag{5}$$

Then we have the following asymptotic normality (see [3], Section 5):

Theorem 3. *Assume that* (5) *holds. If* $\{h_n\}$ *is as in Theorem 1, then, under* H_1,

$$\frac{\tau_n^{(1)} - k_n^*}{\frac{\eta}{|\theta^* - \theta|}\sqrt{h_n}} - c_n^{(1)} \ \xrightarrow{d}\ N(0,1) \quad as \quad n \to \infty.$$

Remark 4. *It is obvious from the proof that*

$$P_{H_1}(\tau_n^{(1)} \geq k_n^*) \ \to \ 1 \quad as \quad n \to \infty.$$

For the second stopping time we similarly assume that

$$k_n^* = \mathcal{O}\big(k_n \log^\gamma(n/k_n)\big) \quad \text{as} \quad n \to \infty \quad (\text{for some } \gamma > 0). \tag{6}$$

Theorem 4. *Assume that* (6) *holds. If* $\{k_n\}$ *is as in Theorem 2, then, under* H_1,

$$\frac{\tau_n^{(2)} - k_n^*}{\frac{\eta}{|\theta^* - \theta|}\sqrt{k_n^*}} - c_n^{(2)} \ \xrightarrow{d}\ N(0,1) \quad as \quad n \to \infty.$$

Remark 5. *Here it is also obvious from the proof that*

$$P_{H_1}(\tau_n^{(2)} \geq k_n^*) \ \to \ 1 \quad as \quad n \to \infty.$$

In [2], Section 5, we also considered

$$\hat{Y}_k = \hat{Y}_{k,n} = \frac{N(k) - N(k - h_n) - h_n\hat{\theta}_k}{\hat{\eta}_k\sqrt{h_n}}, \quad k = \hat{h}_n, \ldots, n,$$

and, for testing H_0 against the two-sided alternative H_1,

$$\hat{\tau}_n^{(1)} = \min\{\hat{h}_n \leq k \leq n : |\hat{Y}_k| > \hat{c}_n^{(1)}\}$$

($\min\emptyset := +\infty$), where $\hat{c}_n^{(1)}$ again is a critical value and $\hat{\theta}_k$ and $\hat{\eta}_k$ above are "suitable" sequential estimates. The "null asymptotics" from Theorem 1 (in the case of θ, η known) retain, so that $\hat{c}_n^{(1)}$ can also be chosen from an extreme value asymptotic, that is, we have

Theorem 5. *If* $h_n \ll \hat{h}_n \ll n$, *but* $\sqrt{h_n} \gg n^{1/r}$, *then, under* H_0, *with the same normalizing sequences* $\{a_n^{(1)}\}$ *and* $\{b_n^{(1)}\}$ *as in Theorem 1,*

$$a_n^{(1)} \max_{\hat{h}_n \leq k \leq n} |\hat{Y}_k| - b_n^{(1)} \ \xrightarrow{d}\ E \quad as \quad n \to \infty,$$

i.e., the critical value $\hat{c}_n^{(1)}$ *can (asymptotically) be chosen as*

$$\hat{c}_n^{(1)} = c_n^{(1)} = \frac{E_{1-\alpha} + b_n^{(1)}}{a_n^{(1)}} \quad \big(\sim \sqrt{2\log(n/h_n)}\ \big),$$

with E *and* $E_{1-\alpha}$ *as in Theorem 1.*

Moreover, with $\hat{\theta} = \hat{\theta}_{\hat{\tau}_n^{(1)}}, \hat{\eta} = \hat{\eta}_{\hat{\tau}_n^{(1)}}$, we have (see [3], Section 6):

Theorem 6. *If* $\{h_n\}$ *and* $\{\hat{h}_n\}$ *are as in Theorem 5, then, under* H_1,

$$\frac{\hat{\tau}_n^{(1)} - k_n^*}{\frac{\hat{\eta}}{|\theta^* - \hat{\theta}|}\sqrt{h_n}} - c_n^{(1)} \ \xrightarrow{d}\ N(0,1) \quad as \quad n \to \infty.$$

4. Some concluding remarks

It is obvious from the proof of Theorem 6 that, if there is an estimate $\hat{\theta}^*$ of the unknown parameter θ^* satisfying

$$\hat{\theta}^* - \theta^* \;=\; o_P\big(1/\sqrt{\log(n/h_n)}\,\big) \quad \text{as} \quad n \to \infty, \tag{7}$$

then the asymptotic normality retains, i.e., we have

Theorem 7. *Assume that* (7) *holds. If* $\{h_n\}$ *and* $\{\hat{h}_n\}$ *are as in Theorem 5, then, under* H_1,

$$\frac{\hat{\tau}_n^{(1)} - k_n^*}{\frac{\hat{\eta}}{|\hat{\theta}^* - \hat{\theta}|}\sqrt{h_n}} - c_n^{(1)} \;\xrightarrow{\;d\;}\; N(0,1) \quad \textit{as} \quad n \to \infty.$$

As an immediate consequence we obtain the following asymptotic confidence interval for the change-point k_n^* (see [3], Section 7):

Corollary 1. *Under the assumptions of Theorem 7, as* $n \to \infty$,

$$P_{H_1}\Big(\hat{\tau}_n^{(1)} - \big(c_n^{(1)} + z_{1-\alpha}\big)\frac{\hat{\eta}}{|\hat{\theta}^* - \hat{\theta}|}\sqrt{h_n} \;\le\; k_n^* \;\le\; \hat{\tau}_n^{(1)}\Big) \;\to\; 1 - \alpha,$$

where $z_{1-\alpha}$ *denotes the* $(1-\alpha)$*-quantile of an* $N(0,1)$*-random variable.*

Remark 6. *A natural estimate* $\hat{\theta}^*$ *of* θ^* *can be obtained by*

$$\hat{\theta}^* = \frac{N(\hat{\tau}_n^{(1)} + h_n^*) - N(\hat{\tau}_n^{(1)})}{h_n^*},$$

where $h_n^* \to \infty$ *(as* $n \to \infty$*) at an "appropriate" rate (cf. [3], Section 7).*

References

[1] Aue A., Horváth L., Kokoszka P., Steinebach J. (2008) Monitoring shifts in mean: Asymptotic normality of stopping times. *Test*, **17**, 515–530.

[2] Gut A., Steinebach J. (2002) Truncated sequential change-point detection based on renewal counting processes. *Scand. J. Statist.*, **29**, 693–719.

[3] Gut A., Steinebach J. (2008) Truncated sequential change-point detection based on renewal counting processes II. *J. Statist. Plann. Infer.*, 16 pp. (online available: DOI 10.1016/j.jspi.2008.08.021).

[4] Horváth L., Steinebach, J. (2000) Testing for changes in the mean or variance of a stochastic process under weak invariance. *J. Statist. Plann. Infer.*, **91**, 365–376.

6th St.Petersburg Workshop on Simulation (2009) 673-677

M-procedures for detection of changes for dependent observations[1]

Marie Hušková[2], Miriam Marušiaková[3]

Abstract

The paper concerns M- type procedures for detection of changes in location models for dependent observations. CUSUM type test statistics are studied when the distribution of the error terms fulfill α-mixing conditions. Theoretical results are accompanied by a simulation study. As special procedures we get L_1 type tests. The results can be extended to more general models.

1. Introduction

We assume that the observations $Y_{1n}, \ldots, Y_{nn}$ obtained at time points $t_1 < \ldots < t_n$ follow the model:

$$Y_{in} = \mu_0 + \delta I\{i > k_n^*\} + e_i, \quad i = 1 \ldots, n, \tag{1}$$

where $k_n^* (\leq n)$, μ_0 and $\delta \neq 0$ are unknown parameters. Function $I\{A\}$ denotes the indicator of the set A. Finally, $e_1, \ldots, e_n$ are random errors fulfilling regularity conditions specified below.
We consider the testing problem no change in location versus there is a change:

$$H_0 : k_n^* = n \quad versus \quad H_1 : k_n^* < n, \tag{2}$$

where k_n^* is unknown.
This problem was considered in a large number papers and books, e.g. Andrews (1993), Csörgő and Horváth (1997), Bai and Perron (1999).
We consider test procedures based on the partial sums of M-residuals:

$$S_k(\psi) = \sum_{i=1}^{k} \widehat{e}_i(\psi), \quad k = 1, \ldots, n, \tag{3}$$

where ψ is a score function,

$$\widehat{e}_i(\psi) = \psi(Y_i - \widehat{\mu}_n(\psi)), \tag{4}$$

[1]This work was supported by grant MSM 0021620839 and GAČR 201/09/0755
[2]Charles University in Prague, E-mail: huskova@karlin.mff.cuni.cz
[3]Charles University in Prague, E-mail: maruskay@gmail.com

$\widehat{\mu}_n(\psi)$ is the M-estimator of μ_0 with score function ψ, i.e., it is defined as a solution of minimization problem $argmin_t \sum_{i=1}^n \rho(Y_i - t)$, where ρ is a convex loss function such that $\rho' = \psi$. Sometimes the estimator is defined as a solution of the equation

$$\sum_{i=1}^{n} \psi(Y_i - t) = 0.$$

The above introduced partial sum test statistics $S_k(\psi)$, $k = 1, \ldots, n$, can be called also *score statistics*.
We assume the following:
(A.1) $\{e_i\}_i$ is a strictly stationary α- mixing sequence with $\{\alpha(k)\}$ and with distribution function F symmetric around 0 and such that for $\delta > 0$ and $\Delta > 0$

$$\sum_{k=0}^{\infty} (k+1)^{\delta/2} \alpha(k)^{\Delta/(2+\delta+\Delta)} \leq C \tag{5}$$

for some positive constant C depending on δ and Δ only.
The score function ψ, the distribution function F and the function $\lambda(t) = -\int \psi(e-t)dF(e)$, $t \in R^1$, satisfy
(A.2) ψ is non-decreasing, antisymmetric, the derivative $\lambda'(.)$ of the function $\lambda(.)$ exists and is Lipschitz in a neighborhood of zero, $\lambda(0) = 0$ and $\lambda'(0) > 0$,
(A.3) $\int |\psi(x)|^{2+\delta+\Delta} dF(x) < \infty$ and

$$\int |\psi(x+t_2) - \psi(x+t_1)|^{2+\delta+\Delta} dF(x) \leq C_1 |t_2 - t_1|^a, \quad |t_j| \leq C_2(\delta, \Delta),\ j = 1, 2$$

for some constants $1 \leq a \leq 2 + \delta + \Delta$, $\delta > 0$, $\Delta > 0$ from ass. (A.1) and C_1 and C_2 are positive constants depending on δ, Δ.
(A.4) Let

$$0 < \sigma^2(\psi) = E\psi^2(e_1) + 2\sum_{i=1}^{\infty} E\psi(e_1)\psi(e_{1+i}) < \infty. \tag{6}$$

• It is known that under very mild conditions linear processes satisfy (A.1), see, e.g. Withers (1981) and Doukhan (1994). The advantage of α-mixing is that if $\{e_i\}_i$ is α-mixing so is $\{q(e_i)\}_i$ with the same coefficients for any measurable function q that suits for our situation.
• The assumptions (A.2)-(A.3) are standard assumptions imposed on the score function ψ and the error distribution F.
• Typical considered ψ's: (a) For $\psi(x) = x$, $x \in R^1$ the procedures reduce to classical L_2 ones that were treated under large spectrum of dependencies, e.g., Csörgő and Horváth (1997), Antoch et al (1997), Perron (2006), Kirch (2006). Assumptions (A.2)-(A.3) reduce to the moment restrictions, no symmetry is needed, $a = 2$.
(b) $\psi(x) = sign\, x$, $x \in R^1$, the procedures reduce to L_1 procedures and assumptions (A.2)-(A.3) are satisfied if the error distribution F is symmetric and has a continuous density f in a neighborhood of 0 and $f(0) > 0$. In this case $a = 1$ for any $\delta > 0$ and $\Delta > 0$.

(c) For the Huber ψ function defined as $\psi(x) = xI\{|x| \le K\} + K\, sign\,(x) I\{|x| \ge K\}$ for some $K > 0$, the assumptions (A.2)-(A.3) are satisfied for F symmetric if the continuous density f exists in a neighborhood of $\pm K$ and $f(K) > 0$. In this case $a = 2$ for any $\delta > 0$ and $\Delta > 0$.
(d) In case the distribution F is known and smooth enough, ψ is a function related to F, usually called score function.
We present the results for the test statistics:

$$T_n(\psi) = \max_{1 \le k < n} \frac{|S_k(\psi)|}{\sqrt{n}\widehat{\sigma}_n(\psi)} \tag{7}$$

$$T_n(\psi, \eta) = \max_{\eta n \le k < n(1-\eta)} \sqrt{\frac{n}{k(n-k)}} \frac{|S_k(\psi)|}{\widehat{\sigma}_n(\psi)} \tag{8}$$

where $\eta \in (0, 1/2)$ and $\widehat{\sigma}_n(\psi)$ is a proper standardization.
These test statistics can be also called *M-type test* due to the closeness to the M-estimators or *score statistics* due to the relation to a likelihood ratio type test. For $\psi(x) = x$, $x \in R$ both test statistics reduce to the most often test statistics used for the testing problem (2). It is related to the likelihood test statistic for the considered testing problem when the error distribution are i.i.d. normal. They are reasonably sensitive w.r.t to large spectrum of the error distribution as soon as the error terms are i.i.d. with some moment restrictions. If either of these assumptions is violated the behavior of the tests can be quite poor. As soon as the error terms are dependent there is a problem to find a reasonable standardization $\widehat{\sigma}_n(\psi)$ with $\psi(x) = x$. This was discussed, e.g., in Hušková et al (1997). Another problem with $\psi(x) = x$, $x \in R$ arises if there is an outlier. Then the test statistics can be considerably influenced by a single observation and erroneously reject the null hypothesis. In case of heavy distribution F standardization $\widehat{\sigma}_n(\psi)$ becomes quite large and the null hypothesis is not rejected in case of alternative. This is connected with so called robustness. Concerning M-type procedures with i.i.d. errors were explored, e.g., Koul et al (2003), Hušková and Picek (2004, 2005).
In the next section we formulate assertions on limit behavior of the above introduced test statistics.

2. Main results

We formulate here the assertions on the limit behavior of the introduced test statistics and also introduce and study a suitable class of the estimators $\widehat{\sigma}_n(\psi)$.

Theorem 1. *Let $Y_{1n}, \dots, Y_{nn}$ follow model (1) with $\delta = 0$. Let assumptions (A.1) – (A.4) be satisfied and let $\widehat{\sigma}_n^2(\psi)$ be a consistent estimator of $\sigma^2(\psi)$, i.e., as $n \to \infty$,*

$$\widehat{\sigma}_n^2(\psi) \to^P \sigma^2(\psi). \tag{9}$$

Then under H_0, as $n \to \infty$,

$$\max_{1 \le k < n} \frac{|S_k(\psi)|}{\sqrt{n}\widehat{\sigma}_n(\psi)} \to^d \max_{0 < t < 1} |B(t)|$$

and

$$\max_{\eta n\le k<n(1-\eta)} \sqrt{\frac{n}{k(n-k)}}\frac{|S_k(\psi)|}{\widehat{\sigma}_n(\psi)} \to^d \max_{\eta<t<1-\eta} \frac{|B(t)|}{\sqrt{t(1-t)}}$$

where $\{B(t);\, t\in(0,1)\}$ *a Brownian motion and* $0<\eta<1/2$.

The proof proceeds along the line of M type procedures with i.i.d. error, i.e., through asymptotic linearities it is shown that the limit behavior of the process $\{S_{\lfloor nt\rfloor}(\psi)/\sqrt{n};\, t\in[0,1]\}$ has the same limit distribution as $\{\big(\sum_{i=1}^{\lfloor nt\rfloor}\psi(e_i) - t\sum_{j=1}^{n}\psi(e_j)\big)/\sqrt{n}; t\in[0,1]\}$. The proof of asymptotic linearity is based on max type inequalities for dependent errors e_i. The proof is then finished applying the Functional central limit theorem.
There is a question of a suitable estimator of $\sigma^2(\psi)$ satisfying (9). We propose the following Bartlett type estimator:

$$\widehat{\sigma}_n^2(\psi) = \widehat{R}(0,\psi) + 2\sum_{k=1}^{\Lambda_n} w(k/\Lambda_n)\widehat{R}(k,\psi), \tag{10}$$

$$\widehat{R}(k,\psi) = \frac{1}{n}\sum_{i=1}^{n-k}\widehat{e}_i(\psi)\widehat{e}_{i+k}(\psi), \tag{11}$$

where

$$w(t) = (1-t)I\{0\le t\le 1\}, \quad t\in R^1.$$

Additionally, to assumptions (A.1) – (A.4) we assume:
(A.5) For some $q>4$

$$E|\psi(e_i)|^q<\infty, \quad \sum_{j=1}^{\infty}\alpha^{1-4/q}(j)<\infty.$$

Theorem 2. *Let* $Y_{1n},\dots,Y_{nn}$ *follow model (1) with* $\delta=0$. *Let assumptions (A.1) – (A.5) be satisfied and let, as* $n\to\infty$

$$\Lambda_n\to\infty, \qquad \Lambda_n n^{-\min(1/3,a/(2(2+\Delta+\delta)))}\to 0 \tag{12}$$

where a,Δ,δ *are from assumptions (A.1) and (A.3). Then under* H_0 *(9) holds true.*

Alternatively, we can consider a flat top kernel:

$$w(t) = 1, \quad |t|\le 1/2$$
$$= 2(1-|t|), \quad 1/2<|t|<1$$
$$= 0, \quad |t|\ge 1.$$

Theorem 2 remains true even for this kernel. More information on the suitability of the respective estimator can be found, e.g., Politis (2003) and Hušková and Kirch (2008). Both estimators are consistent under H_0 or local alternatives but

this is not generally true under fixed alternatives. Typically, the estimators of $\sigma(\psi)$ become much larger, however the resulting tests remain consistent see Theorem 4 below. Both type of estimators of $\sigma(\psi)$ can be modified to take into account possible change in order to improve the power. Confer also Hušková and Kirch (2008). At last we present results on the limit behavior under local as well as fixed alternatives. The results are quite close to those for independent observations.

Theorem 3. *Let $Y_{1n}, \ldots, Y_{nn}$ follow model (1) with $\delta = \delta_n = O(n^{-1/2})$ and $k_n^* = \lfloor n\gamma \rfloor$ for some $\gamma \in (0,1)$. Let assumptions (A.1) – (A.4) and (12) be satisfied and let $\widehat{\sigma}_n^2(\psi)$ be defined in (10). Then $T_n(\psi)$ and $T_n(\psi, \eta)$ have the same limit distribution as $\max_{0<t<1} |B(t) - \delta_n \sqrt{n} \lambda'(0) h_\gamma(t)/\sigma(\psi)|$ and*

$$\max_{\eta<t<1-\eta} \{|B(t) - \delta_n \sqrt{n} \lambda'(0) h_\gamma(t)/\sigma(\psi)|/\sqrt{t(1-t)}\},$$

respectively, where $\{B(t); t \in (0,1)\}$ a Brownian bridge and $0 < \eta < 1/2$ and $h_\gamma(t) = \min(t, \gamma)(1 - \max(t, \gamma))$, $t \in (0,1)$.

Theorem 4. *Let $Y_{1n}, \ldots, Y_{nn}$ follow model (1) with $\delta \neq 0$ with $k_n^* = \lfloor n\gamma \rfloor$ for some $\gamma \in (0,1)$. Let assumptions (A.1) – (A.5) and (12) be satisfied and let $\widehat{\sigma}_n^2(\psi)$ be defined in (10). Moreover, let function $\lambda(t)$ has the derivative in a neighborhood of the points δ^0 and $\delta^0 - \delta$ for δ^0 being the solution of the equation $\gamma\lambda(\delta^0) + (1-\gamma)\lambda(\delta^0 - \delta) = 0$ and δ from the model (1). Moreover, assume $\lambda'(\delta^0) > 0$ and $\lambda'(\delta^0 - \delta) > 0$ and $\int (\psi^2(x+\delta) + \psi^2(x-\delta)) dF(x) < \infty$. Then, as $n \to \infty$*

$$T_n(\psi) \to^P \infty, \quad T_n(\psi, \eta) \to^P \infty.$$

Remark 1. *All above assertions are known for some times for either $\psi(x) = x$, $x \in R^1$ (L_2 procedures) or general ψ but i.i.d. error terms. The assertions can be easily extended to other test statistics based on the partial sums $S_k(\psi)$, $k = 1, \ldots, n$. Also the change point $\widehat{k}_n^*$ defined as*

$$|S_{\widehat{k}_n^*}|(\psi) = \max_{1 \le k \le n} |S_k(\psi)|$$

is a reasonable estimator of the change point k_n^* and has expected asymptotic properties.

Remark 2. *Theorem 1 provides an approximation for critical values for test procedures based on either test statistics. Alternatively, we can also apply a suitable version of resampling methods. Particularly, the block bootstrap studied in Kirch (2006) can be adjusted to our situation.*

Remark 3. *Theorem 1–4 imply that the considered test statistics provide consistent tests.*

References

[1] Andrews, D.W.K. (1993) Tests for parameter instability and structural change with unknown change point. *Econometrica* 61: 821-856.

[2] Antoch J., Hušková M. and Z. Prášková (1998) Effect of dependency on statistics for determination of change. *Statist. Planning and Inference* 60: 291–310.

[3] Csörgő, M., and Horváth, L. (1997) Limit Theorems in ChangePoint Analysis. Wiley, Chichester.

[4] Doukhan P. (1994) *Mixing: properties and examples.* Lecture Notes in Statistics 85, Springer, New York.

[5] de Jong, R. M. and Davidson J.(2000) The functional central limit theorem and weak convergence to stochastic integrals. I. Weakly dependent processes. *Econometric Theory* 16, no. 5, 621–642.

[6] Hušková, M. and Picek J. (2004) Some remarks on permutation type tests in linear models in Regression Models. *Discussiones Mathematicae, Probability and Statistics* 24: 151–181.

[7] Hušková M. and Picek J. (2005) Bootstrap in Detection of Changes in Linear Regression. *Sankhya : The Indian Journal of Statistics Special Issue on Quantile Regression and Related Methods*, Volume 67, Part 2, pp 1-27.

[8] Hušková M. and Kirch C. (2008) A note on studentized confidence intervals for the change-point, submitted.

[9] Kirch C. (2006) *Resampling Methods for the Change Analysis of Dependent Data.* PhD thesis, University of Cologne, Cologne. http://kups.ub.uni-koeln.de/ volltexte/2006/1795/.

[10] Kirch, C. (2007) Block permutation principles for the change analysis of dependent data. *J. Statist. Plann. Inference* 137: 2453-2474.

[11] Koul H., L. Qian and D. Surgailis (2003) Asymptotics of M-estimators in two phase linear regression models. *J. Stochastic Processes and Applications* 103/1: 123–154.

[12] Politis, D. N. (2003) Adaptive bandwidth choice. *J. Nonparametr. Stat.* 15: 517-533.

[13] Withers, C.S. (1981) Conditions for linear processes to be strong-mixing. *Z. Wahrsch. verw. Geb.* 57: 477-480.

6th St.Petersburg Workshop on Simulation (2009) 679-683

A Note on Data-Adaptive Bandwidth Selection for Sequential Kernel Smoothers

Ansgar Steland[1]

Abstract

Sequential kernel smoothers form a class of procedures covering various known methods for the problem of detecting a change in the mean as special cases. In applications, one often aims at estimation, prediction and detection of changes. We propose to use sequential kernel smoothers and study a sequential cross-validation algorithm to choose the bandwidth parameter assuming that observations arrive sequentially at equidistant time instants. An uniform weak law of large number and a consistency result for the cross-validated bandwidth is discussed.

1. Introduction

Let us assume that observations $Y_n = Y_{Tn}$, $1 \leq n \leq T$, T the maximum sample size, arrive sequentially and satisfy the model equation

$$Y_n = m(n/T) + \epsilon_n, \qquad n = 1, 2, \ldots, T,\ T \geq 1,$$

for some bounded and piecewise continuous function $m : [0, \infty) \to \mathbb{R}$. The errors $\{\epsilon_n : n \in \mathbb{N}\}$ form a sequence of i.i.d. random variables such that

$$E(\epsilon_n) = 0, \qquad E(\epsilon_1^4) < \infty. \tag{1}$$

Consequently, $m(t)$, $t \in [0, 1]$, models the process mean during the relevant time frame $[0, T]$. In practice, an analysis has often to solve three problems. (i) Estimation of the current process mean. (ii) One-step prediction of the process mean. (iii) Signaling when there is evidence that the process mean differs from an assumed (null) model. Usually, different statistics are used for these problems. To ease interpretation and applicability, we will base the detector on the same statistic used for estimation and prediction. Our reasoning is that a method which fits the data well and has convincing prediction properties should also possess reasonable detection properties for a large class of alternatives models.

We confine ourselves to closed end procedures where monitoring stops at a (usually large) time horizon T. The proposed kernel smoother is controlled by

[1]RWTH Aachen University, E-mail: steland@stochastik.rwth-aachen.de

a bandwidth parameter which controls the degree of smoothing. As well known, its selection is crucial, particularly for estimation and prediction accuracy. We propose to select the bandwidth sequentially by minimizing a sequential version of the cross-validation criterion. The topic has been quite extensively studied in the literature assuming the classic regression estimation framework where the data gets dense as the sample size increases. A comprehensive monograph of the general methodology is [6]. For references to the literature on estimation of regression functions that are smooth except some discontinuity (change-) points see the recent work of [1].

Before proceeding, let us discuss our assumptions on m. Often the information about the problem of interest is not sufficient to setup a (semi-) parametric model for the process mean m and the distribution of the error terms, which would allow us to use methods based on, e.g., likelihood ratios. In this paper it is only assumed that

$$m \in \mathrm{Lip}, \quad m(t) > 0, \qquad t > 0, \qquad \text{and } \|m\|_\infty < \infty\ , \tag{2}$$

where Lip denotes the class of Lipschitz continuous functions. Under these general conditions, one should use detectors which avoid (semi-) parametric specifications about the shape of m, and nonparametric smoothers $\widehat{m}_n$ which estimate some monotone functional of the process mean and are sensitive with respect to changes of the mean. For these reasons, we confine our study to detectors of the form

$$S_T = \inf\{\lfloor s_0 T \rfloor \le t \le T : \widehat{m}_t > c\}.$$

Here c is a threshold (control limit), $s_0 \in (0,1)$ determines through $\lfloor T s_0 \rfloor$ the start of monitoring, $\lfloor x \rfloor$ denoting the integer part of x, and $\{\widehat{m}_n : n \in \mathbb{N}\}$ is a sequence of $\sigma(Y_1, \ldots, Y_n)$-measureable statistics. Specifically, we study the following sequential kernel smoother

$$\widetilde{m}_n = \widetilde{m}_{n,h} = \frac{1}{h} \sum_{i=1}^{n} K([i-n]/h) Y_i, \qquad n = 1, 2, \ldots$$

and the associated normed version

$$\widehat{m}_n = \widehat{m}_{n,h} = \widetilde{m}_h / \frac{1}{h} \sum_{j=1}^{n} K([i-n]/h),$$

respectively, which are related to the classic Nadaraya-Watson estimator. It is worth noting that various classic control chart statistics are obtained as special cases. Denoting the target value by μ_0, the CUSUM chart is based on $C_n = \sum_{i=1}^{n} [X_i - (\mu_0 + K)]$ where $\{X_n\}$ denotes the observed process and K is the reference value. This chart corresponds to the choice $K(z) = 1$ if $Y_n = X_n - (\mu_0 + K)$ for all n. The EWMA recursion, $\widehat{m}_n = \lambda Y_n + (1-\lambda)\widehat{m}_{n-1}$ with starting value $\widehat{m}_0 = Y_0$, $\lambda \in (0,1)$ a smoothing parameter, corresponds to the kernel $K(z) = e^{|z|}$ and the bandwidth $h = 1/\log(1-\lambda)$.

Our assumptions on the smoothing kernel are as follows.

$$K \in \mathrm{Lip}(\mathbb{R}; [0,\infty)),\ \|K\|_\infty < \infty,\ \mathrm{supp}(K) \subset [-1,1],\ \text{and } K > 0 \text{ on } (0,1). \tag{3}$$

For the bandwidth $h > 0$ we assume that

$$\lim_{T\to\infty} T/h = \xi \tag{4}$$

for some constant $\xi \in (0, \infty)$, which guarantees that in our design the number of observations on which $\widehat{m}_T$ depends converges to ∞, as $T \to \infty$. In practice, one can select ξ and put $h = T/\xi$.

In [3, 4, 5] procedures based on the sequential smoother $\widehat{m}_n$ are studied, which allow us to detect changes in the mean of a stationary or random walk series of observations. The asymptotic theory was studied as well. Specifically, in [4] it is shown that under the assumptions of the present paper the process $\{\sqrt{T}\widehat{m}_{\lfloor Ts\rfloor,h} : s \in [0,1]\}$ satisfies a functional central limit theorem when $m = 0$, i.e.,

$$\sqrt{T}\widehat{m}_{\lfloor Ts\rfloor,h} \Rightarrow \mathbb{M}(s),$$

for some centered Gaussian process $\{\mathbb{M}(s) : s \in [0,1]\}$ which depends on ξ.This result can be used to construct detection procedures with pre-specified statistical properties. E.g., when choosing the control limit such that the type I error rate satisfies $P(S_T \le T) = \alpha$ when $m = 0$ for some given significance level $\alpha \in (0,1)$, the control limit also depends on ξ. The question arises, how one can or should select the bandwidth $h \sim T$ and the parameter ξ, respectively.

In this paper we propose to select the bandwidth $h > 0$ such that the Y_t are well approximated by sequential predictions $\widehat{m}_t$ which are calculated from past data $Y_1, \ldots, Y_{t-1}$. For that purpose we propose a sequential version of the cross-validation criterion based on sequential leave-one-out estimates.

2. Sequential Cross-Validation

The idea of cross-validation is to choose parameters such that the corresponding estimates provide a good fit on average. To achieve this goal, one may consider the average squared distance between observations, Y_i, and predictions as an approximation of the integrated squared distance. To avoid over-fitting and interpolation, the prediction of Y_i is determined using the reduced sample where Y_i is omitted. Since, additionally, we aim at selecting the bandwidth h to obtain a good fit when using the sequential estimate, we consider

$$\widehat{m}_{h,-i} = N_{T,-i}^{-1}\frac{1}{h}\sum_{j=1}^{i-1} K([j-i]/h)Y_j, \qquad i = 2, 3, \ldots$$

with the constant $N_{T,-i} = h^{-1}\sum_{j=1}^{i-1} K([j-i]/h)$. Notice that $\widehat{m}_{h,-i}$ can be regarded as a sequential leave-one-out estimate. The corresponding *sequential leave-one-out cross-validation criterion* is defined as

$$CV_s(h) = \frac{1}{T}\sum_{i=2}^{\lfloor Ts\rfloor}(Y_i - \widehat{m}_{h,-i})^2, \qquad h > 0.$$

The cross-validation bandwidth at time s is now obtained by minimizing $CV_s(h)$ for fixed s. Notice that $CV_s(h)$ is a sequential unweighted version of the criterion studied by [2] in the classic regression function estimation framework. We do not consider a weighted CV sum, since we have in mind that the selected bandwidth is used to obtain a good fit for past and current observations. However, similar results as presented here can be obtained for the weighted criterion $T^{-1}\sum_{i=1}^{n} K([i-n]/h)(Y_i - \widehat{m}_{h,-i})^2$ as well. Notice that due to

$$CV_s(h) = \frac{1}{T}\sum_{i=1}^{\lfloor Ts\rfloor} Y_i^2 - \frac{2}{T}\sum_{i=2}^{\lfloor Ts\rfloor} Y_i \widehat{m}_{h,-i} + \frac{1}{T}\sum_{i=2}^{\lfloor Ts\rfloor} \widehat{m}_{h,-i}^2,$$

minimizing $CV_s(h)$ is equivalent to minimizing

$$C_{T,s}(h) = -\frac{2}{T}\sum_{i=2}^{\lfloor Ts\rfloor} Y_i \widehat{m}_{h,-i} + \frac{1}{T}\sum_{i=2}^{\lfloor Ts\rfloor} \widehat{m}_{h,-i}^2.$$

Thus, we will study $C_{T,s}(h)$ in the sequel. Cross-validation is expensive in terms of computational costs and minimizing $C_{T,s}$ for all s is not feasible in many case. Therefore and to simplify exposition, let us fix a finite number of time points

$$0 < s_1 < \cdots < s_N,$$

$N \in \mathbb{N}$. Later we shall relax this assumption and allow that N is an increasing function of T. At time s_i the cross-validation criterion is minimized to select the bandwidth, $h_i^* = h_i^*(Y_1, \ldots, Y_{s_i})$, and that bandwidth is used during the time interval $[s_i, s_{i+1})$, $i = 1, \ldots, N$.

3. Asymptotic Results

The question arises which function is estimated by $C_{T,s}(h)$. Our first result identifies the limit and shows convergence in mean.

Theorem 1. *We have*

$$E(C_{T,s}(h)) \to C_\xi(s) = -2\frac{\int_0^s \int_0^r \xi K(\xi(u-r))m(\xi u)\,du dr}{\int_0^s \xi K(\xi(r-s))\,dr} + \frac{\int_0^s \xi^2 \int_0^r \int_0^r K(\xi(u-r))K(\xi(v-r))m(u)m(v)\,du\,dv\,dr}{\int_0^s \xi K(\xi(r-s))\,dr}, \tag{5}$$

as $T \to \infty$, uniformly in $s \in [s_0, 1]$.

Before proceeding, let us consider an example where the function $C_\xi(s)$ possesses a well-separated minimum.

Example 1. *Suppose K is given by $K(z) = (1-|z|)1_{[0,1]}(z)$ for $z \in \mathbb{R}$. Further, let us consider the nonlinear function $m(t) = x(x-0.2)(x-0.4)$. Clearly, $C_\xi(s)$ is a polynomial of order 4 with coefficients which depend on s. Figure 1 depicts $C_\xi(s)$ for some values of ξ. The locations of the (real) roots of $\frac{\partial}{\partial \xi} C_\xi(s)$ depend on $s \in [0,1]$ and are shown in Figure 1 as well.*

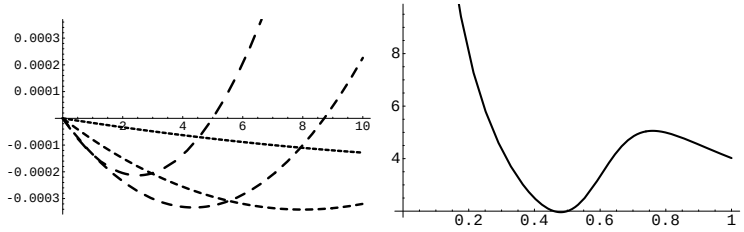


Figure 1: *Left panel:* The function $C_s(\xi)$, $\xi \in (0, 20]$, for $s \in \{0.1, 0.2, 0.3, 0.4\}$. *Right panel:* The optimal values for ξ as a function of $s \in (0, 1]$.

We will now study the uniform mean squared convergence of the random function $C_{T,s}(h)$. Define $\mathcal{S}_N = \{s_i : 1 \leq i \leq N\}$.

Theorem 2. *We have*

$$E \sup_{s \in \mathcal{S}_N} |C_{T,s}(h) - E(C_{T,s}(h))|^2 = O(T^{-1}),$$

as $T \to \infty$.

Current research focuses on the following generalization which allows that the number of time points where cross validation is conducted is a function of the maximum sample size T.

Conjecture 3. *Assume* $N = N_T$ *and*

$$0 < s_0 < s_{N1} < \cdots < s_{NN} \leq 1, \qquad N \geq 1, \tag{6}$$

and put $\mathcal{S}_N = \{s_{Ni} : 1 \leq i \leq N\}$. *Given the assumptions of Theorem 2 there exists some* $\gamma > 0$ *with* $\frac{N_T}{T^\gamma} = o(1)$, *such that*

$$E \sup_{s \in \mathcal{S}_N} |C_{T,s}(h) - E(C_{T,s}(h))|^2 = o(1)$$

Combining the above statements, we obtain

Theorem 4. *Suppose that* (1) *and* (6)

$$E \sup_{s \in \mathcal{S}_N} |C_{T,s}(h) - C_s(\xi))|^2 \to 0,$$

as $T \to \infty$.

We shall now extend the above results to study weak consistency of the cross-validation bandwidth under fairly general and weak assumptions. Having in mind the fact that $h \sim T$, let us simplify the setting by assuming that

$$h = h(\xi) = T/\xi, \qquad \xi \in [1, \Xi],$$

for some fixed $\Xi \in (1, \infty)$. This means, h and ξ are now equivalent parameters for each T. We also restrict the optimization to a compact interval, which is not restrictive for applications. Now $\widehat{m}_{h,-i}$ can be written as

$$\widehat{m}_{h,-i} = \frac{1}{(i-1)h} \sum_{j=1}^{i-1} K(\xi(j-i)/T) Y_j.$$

With some abuse of notation, let us also write

$$C_{T,s}(\xi) = C_{T,s}(T/\xi).$$

Theorem 5. *For any* $s \in [s_0, 1]$

$$\sup_{\xi \in [1,\Xi]} |C_{T,s}(\xi) - EC_{T,s}(\xi)| = o_P(1), \tag{7}$$

and

$$\sup_{\xi \in [1,\Xi]} |C_{T,s}(\xi) - C_{\xi}(s)| = o_P(1), \tag{8}$$

as $T \to \infty$.

We are now in a position to formulate the following conjecture on the asymptotic behavior of the the cross-validated sequential bandwidth selector.

Conjecture 6. *Suppose* $C_\xi(s)$ *possesses a well-separated minimum* $\xi^* \in [1, \Xi]$, *i.e.,*

$$\sup_{\xi \in [1,\Xi]: |\xi - \xi^*| \geq \varepsilon} C_\xi(s) > C_{\xi^*}(s).$$

for every $\varepsilon > 0$. *Then*

$$\operatorname{argmin}_{\xi \in [1,\Xi]} C_{T,s}(\xi) \xrightarrow{P} \xi^*.$$

References

[1] Gijbels I., Goderniaux A.C. (2004) Bandwidth Selection for Changepoint Estimation in Nonparametric Regression. *Technometrics*, **46**, 1, 76–86. Oliver and Boyd, Edinburgh.

[2] Härdle W., Marron J.S. (1985) Optimal bandwidth selection in nonparametric regression function estimation. *Ann. Statist.*, **13**, 1465–1481.

[3] Schmid W., Steland A. (2000) Sequential control of non-stationary processes by nonparametric kernel control charts. *AstA Adv. Stat. Anal.*, **84**, 315–336.

[4] Steland A. (2004) Sequential control of time series by functionals of kernel-weighted empirical processes under local alternatives. *Metrika*, **60**, 229–249.

[5] Steland A. (2005) Random walks with drift - A sequential approach. *J. Time Ser. Anal.*, **26** (6), 917-942.

[6] Wand M.P., Jones M.C. (1995) *Kernel Smoothing*. Chapman & Hall, Boca Raton.

Session

Reliability and survival analysis

organized by Ingram Olkin
(USA)

6th St.Petersburg Workshop on Simulation (2009) 687-691

Calibration of proportional hazards models

Jeannette Simino[1], Myles Hollander, Dan McGee

Abstract

Given a prognostic survival model based on one population, the question arises as to whether this model may be used to accurately predict disease in a different population. When the underlying rate of disease differs in the new population, the model must be calibrated. Following work by van Houwelingen (2000), we examine whether a model based on the Framingham Heart Study can be applied to diverse studies from around the world.

1. Introduction

The American Heart Association urges all CHD-free adults aged 40 and older to have their global CHD risk computed every 5 years [1]. Clinicians base treatment decisions on this assessment of underlying global CHD risk, thus the prediction equations used should be accurate and feasible [1]. Framingham Study based risk functions are commonly recommended for use in the United States [1] because of the Framingham study methodology, long term follow up, and inclusion of females [2]. The validity of Framingham-based equations in other populations is questionable because the study consisted of white middle-class individuals [3]. Framingham risk functions may not accurately estimate absolute CHD risk in populations with low or high CHD risk based on major risk factors [2].

There are no generally accepted algorithms for assessing the validity of an established survival model in a new population. Van Houwelingen developed a method called validation by calibration which allows a clinician to assess the validity of a well-accepted published survival model on his/her own patient population and adjust the published model to fit that population [4]. Van Houwelingen embeds the published model into a new model containing only 3 parameters to test if the published model is strictly valid in the new population. As van Houwelingen points out, this helps combat the overfitting that occurs when models with many covariates are fit on small datasets. His method also enables researchers to conduct statistical tests to determine if the shape and scale of the hazard functions, as well as the hazard ratios, are properly specified. Each component can be adjusted if necessary.

[1]All authors are affiliated with Florida State University. Correspondence can be addressed to `simino@stat.fsu.edu` (Jeannette Simino), `holland@stat.fsu.edu` (Myles Hollander), and `dan@stat.fsu.edu` (Dan McGee).

We use van Houwelingen's validation by calibration method to judge the validity of the Framingham Cox models in cohorts from the Diverse Populations Collaboration (DPC), a collection of studies encompassing many ethnic, geographic, and socioeconomic groups [5]. Validation by calibration can be useful since one female and three male DPC cohorts have 23 CHD deaths or less. We will also perform simulations to gauge the power of the validation by calibration method to reject an invalid Weibull proportional hazards model under various shape and scale misspecifications.

2. The Framingham Model

We first fit sex-specific models of CHD death to participants examined during the lipoprotein phenotyping project that roughly corresponded to the eleventh exam (1971) of the original Framingham study and the first exam of the Framingham Offspring study. The models were fit on cardiovascular disease-free individuals with complete predictor information who were censored at 15 years. Details of the model building process and final model parameter estimates can be found elsewhere [5]. The male and female models both include age, HDL, diabetes, and current smoking status. Additionally, the female model includes cholesterol and DBP whereas the male model includes the inverse of cholesterol squared and SBP. To employ the validation by calibration method we need the baseline cumulative hazard function estimated at each observed time in the external data sets. We use nonlinear least squares to fit the simpler mathematical function presented by van Houwelingen and Thorogood to the cumulative hazard of each Framingham sex-specific model [5].

3. Discrimination and Hazard Ratio Test

The estimated concordance indices of the Framingham Cox models in the DPC cohorts ranged from 0.80 to 0.88 except in the ARIC Black (0.67), ARIC Nonblack (0.72), and CHS White (0.66) male cohorts. Harrell et al. have suggested that if a model adequately discriminates then it can be re-calibrated [6]. We pursue the calibration that assumes that the relative weight of each predictor is correctly specified in the Framingham model and hence does not alter discrimination [4]. Although we proceed to calibrate the male ARIC and CHS White cohorts, a revision of predictor relative weights or creation of a new model might be necessary.

Before undertaking the validation by calibration procedure, we conduct a preliminary test of the Framingham sex-specific hazard ratio validity. To each individual in the external cohort we calculate the prognostic index, $\mathbf{x}_i\beta$, using the coefficient vector from the Framingham sex-specific Cox model. We fit a cohort-specific Cox proportional hazards model with the prognostic index as the single covariate [4]. We test that the coefficient of the prognostic index is unity at the 0.05 level. The Framingham hazard ratio is rejected in the female NHANES II White cohort, the male CORDIS cohort, and both male and female LRC cohorts.

4. Van Houwelingen's Validation by Calibration

Let Z, a continuous random variable representing the time to failure, follow a semi-parametric Cox proportional hazards model with baseline cumulative hazard function $\Lambda_0(z)$ and p-length coefficient vector β. Let $\mathbf{x}$ represent a p-length row vector of risk factors determined prior to study commencement. Let the model satisfy the condition that $\Lambda(z|\mathbf{x}) = e^{\mathbf{x}\beta}\Lambda_0(z)$. The transformation $Z^* = \Lambda_0(Z)$ represents a random variable following an accelerated failure time model with a baseline exponential distribution of mean 1. Van Houwelingen suggests fitting the calibration model

$$ln(Z^*) = \psi + \omega(\mathbf{x}\beta) + \rho\epsilon \tag{1}$$

treating $\mathbf{x}\beta$ as a single fixed covariate and ϵ as Gumbel. By transforming from Z to Z^* the proportional hazards model has been reformatted into a Weibull accelerated failure time model. The parameters ψ and ω can take any real numbered values but ρ must be positive. If the underlying model of the failure time is the proportional hazards model corresponding to the cumulative hazard Z^* and the coefficient vector β, then $\psi = 0$, $\omega = -1$, and $\rho = 1$. He notes that ρ relates to the shape of the baseline hazard, ψ controls the overall level of failure, and ω controls the effect of the published linear predictor. Specifically, the shape of the baseline hazard has been properly specified if $\rho = 1$. The quantity $e^{-\frac{\psi}{\rho}}$ is a scaling factor in the recalibrated hazard, thus if both $\rho = 1$ and $\psi = 0$ the baseline hazard is considered valid. The published model correctly specifies the hazard ratios if $\frac{\omega}{\rho} = -1$.

5. Calibrating the Framingham Model to DPC Cohorts

A description of the DPC cohorts and analysis inclusion/exclusion criteria is described elsewhere [5]. We calculate the transformed outcome $z_i^* = \Lambda_0(t_i)$ using the values of the observed time of the individual in the DPC cohort, t_i, and the smoothed Framingham baseline cumulative hazard. We use Stata to fit a Weibull accelerated failure time model with z_i^* as the outcome, the prognostic index $\mathbf{x}_i\beta$ as the single covariate, and CHD death as the indicator of failure. If the Framingham proportional hazards model is correctly specified for the new external sample, $\psi = 0$, $\omega = -1$, and $\rho = 1$. Stata parameterizes the model slightly different thus we actually test $H_0 : \psi = 0$, $\omega = -1$, $\ln(\frac{1}{\rho}) = 0$. Table 1 contains the validation by calibration results for the sex-specific Framingham Cox models applied to each cohort of the DPC.

The global hypothesis test of the Framingham sex-specific model is rejected in half of the cohorts for both genders. The males and female cohorts within a study behave similarly with regard to the global test. In the male CORDIS cohort, the failure to reject the global hypothesis test contradicts the rejection of the Framingham hazard ratio in Section . The global validation by calibration hypothesis test may have insufficient power to detect the invalidity.

The adjustments should meet some statistical criteria since any calibrated model might ultimately be applied to the whole parent population. Applying the 3-

Table 1: Validation by Calibration of the Framingham Sex-specific Cox Models onto the Cohorts from the DPC. The p-value corresponds to the test of the joint hypothesis that $\psi = 0, \omega = -1$, and $\rho = 1$.

Model	Cohort	ψ	$se(\psi)$	ω	$se(\omega)$	ρ	$se(\rho)$	p-value
Female	CHS White	-1.58	1.88	-0.92	0.31	0.74	0.19	0.14
	Glostrup	-1.53	0.80	-0.75	0.12	0.86	0.11	0.13
	LRC	-1.79	0.58	-0.74	0.09	0.90	0.10	0.00
	NHANES II White	0.30	0.68	-0.93	0.11	1.24	0.11	0.00
Male	ARIC Black	-3.78	0.78	-0.46	0.19	0.66	0.12	0.00
	ARIC Nonblack	-4.42	0.42	-0.42	0.11	0.43	0.05	0.00
	CHS White	-0.12	1.76	-0.91	0.37	0.95	0.20	0.38
	CORDIS	1.23	0.87	-1.34	0.20	1.07	0.12	0.25
	Glostrup	1.03	0.69	-1.19	0.14	1.11	0.10	0.42
	LRC	1.59	0.62	-1.45	0.15	1.23	0.10	0.00
	NHANES II Black	4.82	3.00	-1.95	0.68	1.65	0.40	0.23
	NHANES II White	1.61	0.58	-1.32	0.13	1.25	0.09	0.02

parameter adjusted model to the whole population could be just as egregious as directly applying the Framingham, especially in the DPC cohorts with small number of events. We perform a backward elimination of the calibration parameters at the 0.05 level. The three separate hypotheses are $H_1 : \psi = 0$, $H_2 : \omega = -1$, and $H_3 : \rho = 1$. We set the parameter to its hypothesized value if it is deemed insignificant with the greatest p-value. If the Wald and likelihood ratio tests yield different eliminations, we will give preference to the latter.

Except the males of CORDIS, all cohorts that failed to reject the global hypothesis test backward eliminated all three calibration parameters. We fail to reject conditional Wald tests of $\psi = 0$ given $\omega = -1$ and $\rho = 1$ at the 0.05 level in these cohorts as well. The parameters relating to the overall death level and impact of the predictors remained in the calibration model of the male CORDIS cohort. This result is consistent with the rejection of the Framingham hazard ratio in the male CORDIS group. Among cohorts that reject the global validation by calibration hypothesis test, the females keep two calibration parameters whereas the males keep all three.

We fit alternate models to cohorts using a significance of 0.10 for the backward elimination and the conditional Wald test. Bar graphs are generated to see how the backward eliminated and alternate models fared relative to the Framingham model. We divide each cohort into quintiles based on the Framingham Cox model predicted CHD death probabilities. For each quintile we graph the observed probability of CHD death within 10 years (or 5 years if data dictates) computed using Kaplan-Meier, as well as the mean predicted probabilities using the Framingham model and any backward elimination or alternate models. Figure 1 depicts the poor calibration both the Framingham and backward eliminated models achieve in the ARIC Black cohort following the poor discrimination indicated by the concordance index. The figure also shows the improvement in calibration in the male LRC cohort which began with good discrimination.

The parameter estimates and standard errors for the chosen backward eliminated and alternate models are listed in Table 2. We could not reject the Framingham model for the omitted cohorts. The low number of events in the male NHANES

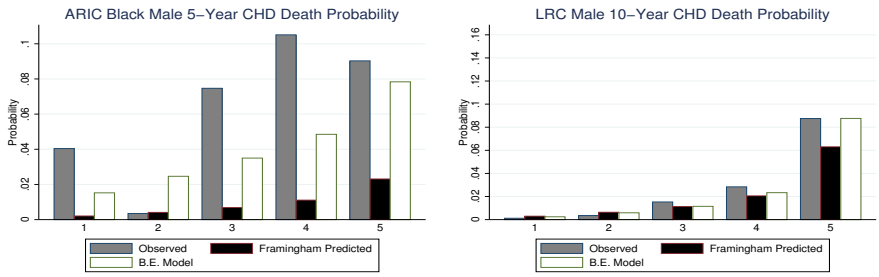


Figure 1: Observed and Predicted CHD Death Probabilities by Quintile of Framingham Model Risk

II Black cohort and CHS White cohorts of both sexes causes large standard errors of parameter estimates which complicates the assessment of calibration parameter inclusion. Although we exclude the ARIC cohorts from further discussion, we include them in the table for completeness, noting that they require a revision of risk factor weights or the fitting of new cohort-specific models. Wald tests of the hazard ratios of the backward eliminated models are consistent with the results of the cohort-specific tests. The cohorts that reject the hazard ratio test have an asterisk next to their name.

Table 2: Backward Eliminated and Alternate Models for Cohorts of the DPC

Model	Cohort	ψ	$se(\psi)$	ω	$se(\omega)$	ρ	$se(\rho)$
Female	LRC with $\rho = 1$ *	-1.23	0.24	-0.82	0.06		
	NHANES II White with $\psi = 0$ *			-0.88	0.04	1.20	0.05
Male	ARIC Black	-3.78	0.78	-0.46	0.19	0.66	0.12
	ARIC Nonblack	-4.42	0.42	-0.42	0.11	0.43	0.05
	Alt. CHS White with $\omega = -1$, $\rho = 1$	0.36	0.21				
	CORDIS with $\rho = 1$ *	0.77	0.37	-1.25	0.13		
	LRC *	1.59	0.62	-1.45	0.15	1.23	0.10
	NHANES II White	1.61	0.58	-1.32	0.13	1.25	0.09

Our baseline survivor function represents a non-smoking, non-diabetic 45-year old individual with a cholesterol of 160 mg/dL, HDL of 60 mg/dL, and either SBP of 120 mmHg (males) or DBP of 80 mmHg (females). According to the chosen backward eliminated or alternate models, the Framingham model overestimates the survival probability in the LRC and NHANES II female cohorts for all times between 0 and 15 years but underestimates the survival probability for times between 10 to 15 years in the male CHS White, CORDIS, LRC, and NHANES II White cohorts. The hazard ratio for a q unit increase of any predictor yields a value closer to unity under the backward eliminated female LRC and NHANES II White models than under the Framingham model. Under the backward eliminated male CORDIS, LRC, and NHANES II White models the hazard ratios are farther away from unity than in Framingham, although the difference is not statistically significant for the NHANES II White cohort.

6. Power Simulations

We apply sex-specific Framingham-fit Weibull proportional hazards model of CHD death to simulated data following a Weibull model of different scale and/or shape from Framingham but with the same hazard ratio. Each simulated data set includes 2500 observations with 4 continuous and two dichotomous predictors. We perform 10000 repetitions to calculate the proportion of data sets from the simulated Weibull distribution that reject the Framingham model via the global test. The power to reject the null hypothesis depends on the complicated interplay between the distance of the estimated parameter values from their hypothesized values, the covariance matrix of the parameter estimates, and potential skewness/non-normality of the parameter estimates. Thus we advocate using van Houwelingen's method more as an exploratory tool for adjustments rather than relying on it as a strict test of validity.

References

[1] Sheridan S., Pignone M., Mulrow C. (2003) Framingham-based tools to calculate the global risk of coronary heart disease: a systematic review of tools for clinicians. *Journal of General Internal Medicine*, **18**, 1039–1052.

[2] Haq I., Ramsay L., Yeo W., et al. (1999) Is the Framingham risk function valid for northern European populations? A comparison of methods for estimating absolute coronary risk in high risk men. *Heart*, **81**, 40–46.

[3] D'Agostino R., Grundy S., Sullivan L., et al. (2001) Validation of the Framingham coronary heart disease prediction scores: results of a multiple ethnic groups investigation. *Journal of the American Medical Association*, **286**, 180–187.

[4] van Houwelingen H. (2000) Validation, calibration, revision and combination of prognostic survival models. *Statistics in Medicine*, **19**, 3401–3415.

[5] Simino J. (2008) *Discrimination and Calibration of Prognostic Survival Models (dissertation)*. Florida State University, Tallahassee, FL.

[6] Harrell Jr. F., Lee K., Mark D. (1996) Tutorial in biostatistics. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Statistics in Medicine*, **15**, 361–387.

6th St.Petersburg Workshop on Simulation (2009) 693-697

Models for Assessing Uncertainty for Local Regression[1]

Linda J. Young[2], Carol A. Gotway[3], Greg Kearney[4], Chris DuClos[5]

Abstract

Many studies and programs rely on existing data from multiple sources (e.g., surveillance systems, health registries, governmental agencies) as the foundation for analysis and inference. Numerous statistical issues are associated with combining such disparate data. Florida's efforts to move toward implementation of The Centers for Disease Control and Prevention (CDC)s Environmental Public Health Tracking (EPHT) Program (EPHT) aptly illustrate these issues, which are typical of almost any study designed to measure the association between environmental hazards and health outcomes. In this paper, we consider the inferential issues that arise when a potential explanatory variable is measured on one set of spatial units, but then must be predicted on a different set of spatial units. We compare methods for assessing uncertainty and the potential bias that arises from using predicted variables in spatial regression models. Our focus is on relatively simple methods and concepts that can be transferred to the states departments of health, the organizations responsible for implementing EPHT.

1. Introduction

The Centers for Disease Control and Preventions Environmental Public Health Tracking (EPHT) program is to track exposure and health effects that may be related to environmental hazards. EPHT is relying on health and environmental exposures and hazards information collected by other programs; little or no primary data are being collected. Now in the implementation stage, states are beginning to develop websites that provide and map these existing data sets. As standards

[1]This publication was supported by the Florida Department of Health, Division of Environmental Health and Grant/Cooperative Agreement Number 5 U38 EH000177-02 from the Centers for Disease Control and Prevention (CDC). The findings and conclusions in this report are those of the authors and do not necessarily represent the views of the CDC.

[2]University of Florida, E-mail: LJYoung@ufl.edu

[3]Centers for Disease Control and Prevention, E-mail: cdg7@cdc.gov

[4]Florida Department of Health, E-mail: Greg_Kearney@doh.state.fl.us

[5]Florida Department of Health, E-mail: Chris_DuClos@doh.state.fl.us

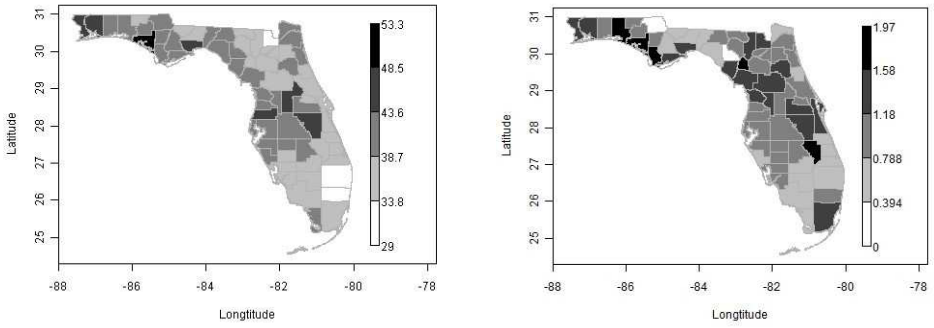
for a unified national reporting system are being developed, the focus has turned to statistical inference based on combining data from disparate sources, each source associated with different sampling units and hence different spatial support. To link data on a common spatial scale for subsequent analysis, some variables must be predicted on that scale using data collected on a different scale, e.g., data collected from ozone monitors, which have point support, must be used to predict ozone values at the county level (see [6],[2], [11] for similar efforts). If a potential explanatory variable must be predicted on a different spatial unit, the resulting predicted values are generally smoother than the true ones and, in some cases, this can lead to bias in the estimated effects [5], [3]. In addition, smoothing can further impact proper uncertainty assessment. The bias from smoothing has been studied in the context of moving from point support to point support and from point support support to areal support followed by global regression [10]. In the study considered here, the smoothing process takes us from point support to areal support, and the subsequent analysis allows for spatial variation in the association between public health and ozone. Further, uncertainty assessment and bias reduction in studies with spatially misaligned data is explored. Although our focus is on data collected in support of the EPHT program, the methodology, concepts, and key ideas we present pertain to most studies linking health with environmental factors.

2. The Study and Supporting Data

Two of EPHT's core measures are ozone and myocardial infarction (MI). For concreteness and to illustrate a few of the many statistical challenges encountered in this basic study, we will consider the data from August, 2005, to model the relationship between ozone (environmental exposure) and MI (health outcome) at the county level. When assessing the potential relationship between ozone and MI, care must be taken to account for potential confounders. The ozone, MI, and sociodemographic data used in this effort have been gathered from four different sources.

Ozone measurements, recorded from a network of air monitors placed throughout the state, were obtained from Floridas Department of Environmental Protection (FDEP). During the study period of August 2005, ozone data were available from 56 monitors. The maximum of the daily maximum 8-hour average ozone values during a month is used as the monthly data value for a particular monitor and is referred to as the monthly maximum ozone.

FDOH has a data-sharing agreement (DSA) with Florida's Agency for Health Care Administration (AHCA), providing us with access to two data sources: confidential hospitalization records and emergency room records. The hospitalization records contain all admissions to Florida's public and private hospitals where either the primary or secondary cause of admission was MI (ICD-10 codes 410.0–414.0) [8]. If a person presents to the emergency room and is subsequently admitted to the hospital, that person's record is included in the hospitalization records, but not the emergency room records. Non-Florida residents were excluded from the analysis. Consistent with our DSA, AHCA provided both the zip code and county



(a) (b)

Figure 1: The (a) predicted maximum ozone from block kriging and (b) the MI SER for Each County from August 2005.

of residence for each patients record. Selected patient demographic information is also recorded, including sex, age, and race/ethnicity.

Selected sociodemographic data were obtained from two sources: the U.S. Census Bureau and CDCs Behavioral Risk Factor Surveillance System (BRFSS). Census estimates (e.g., age, race/ethnicity, sex, education) at the state and county level are available on an annual basis. BRFSS is a state-based system of health surveys that collect annual information on health risk behaviors (e.g., the percentage of smokers in each county), preventive health practices, and health-care access, primarily related to chronic disease and injury [7].

3. Data Linkage and Analysis

Before analysis, the data must be linked on the same geographical scale. The county level is used here. For the ozone data, block kriging was used to obtain a predicted maximum ozone value for each county that accounts for the change of support from the points at which monitors are located (Figure 2(a)). Evidence of autocorrelation in the smoothing effect of the support-adjusted predictions may be observed here. The MI data were indirectly standardized by age (aged =45, 45–55, 55–65, and >65 years), race/ethnicity (black, white, or other), and sex (female or male). The information regarding age, race/ethnicity, and sex for the MI cases were obtained from the hospital records. Florida's population was used as the comparison standard to calculate the number of expected MI cases. This gave an MI standardized event ratio (MI SER), defined as the ratio of the number of observed MI cases to that expected among the Florida population, for each county and each month [9]. These are displayed in Figure 2(b). Socio-demographic data from the BRFSS and the the U.S. Census were available at the county level.

After the data were linked, the focus became analyzing the association between MI SER and ozone. A linear mixed model was used to relate MI SER to ozone,

adjusting for covariates, and allowing the intercept and coefficient of ozone to vary with county. So that the assumptions of the models were more nearly met, the natural logarithm of MI SER was used as the response variable. The expected MI SER for a county was used as a weight for that county to account for non-constant variance. A number of covariates were considered. The significant covariates were percent of adults aged 25 and over who have completed high school but do not have any college education, percentage of adults aged 25 and over who smoke, and social economic status (SES) as measured by an indicator variable of whether or not the county median income was above or below the median income of the state. Random effects for county and the interaction between county and maximum ozone allowed for an intercept and slope for each county. A spherical covariance model was used to account for spatial correlation among the counties'intercepts and slopes associated with ozone. A spherical covariance structure was also used to account for correlation among the residuals that is expected from the Berksen error arising from using the predicted monthly maximum ozone for county i from block kriging instead of the true value [3]. The relative MI SER is the exponentiation of the sum of the estimated coefficients associated with ozone and the interaction between ozone and county. The delta method was used to obtain the standard error of relative SER. The relative SERs for each county from this "Krige and Regress, Accounting for Covariance"approach are displayed in Figure 3(a).

The "Krige and Regress, Accounting for Covariance"method provides an unbiased estimate of relative SER and the appropriate standard error if the model is the true one. To explore the possibility of error in the exposure model, we use a leave-one-out cross-validation regression calibration approach based on a simple measurement error model in which the observed maximum monthly ozone at each monitor is regressed against the covariates and the predicted ozone from leave-one-out-cross-validation for the county [1]. The calibration equation obtained from the monitors was applied to the predicated county maximum ozone levels. The calibrated county ozone values were then used in the linear mixed models described above. This is called the "Krige, Calibrate, and Regress"method, and the relative SERs are displayed in Figure 3(b).

Although these results account for the effect of prediction error in the regression, the uncertainty may still be too small as kriging predictors were used in both previous models. Thus, $K = 1000$ conditional simulations of ozone were conducted. Support-adjusted predictions of average ozone values for each county were obtained, giving a distribution of predicted ozone values for each county that defines the uncertainty distribution for predicted ozone. The linear mixed model in the "Krige and Regress, Accounting for Covariance"approach was then fit for each simulation. The results from the 1000 simulations were combined to obtain relative SERs [4] (see Figure 3(c)). This is described as the "Conditional Simulation"method.

The predicted monthly maximum ozone values from conditional simulation have the same variability as the true monthly maximum ozone values. Failure of the two to be equal should be considered classical measurement error (not Berksen error as in the earlier models, see [3]). Thus, the parameter estimates obtained from the linear mixed model are biased. Although [3] recognized the measure-

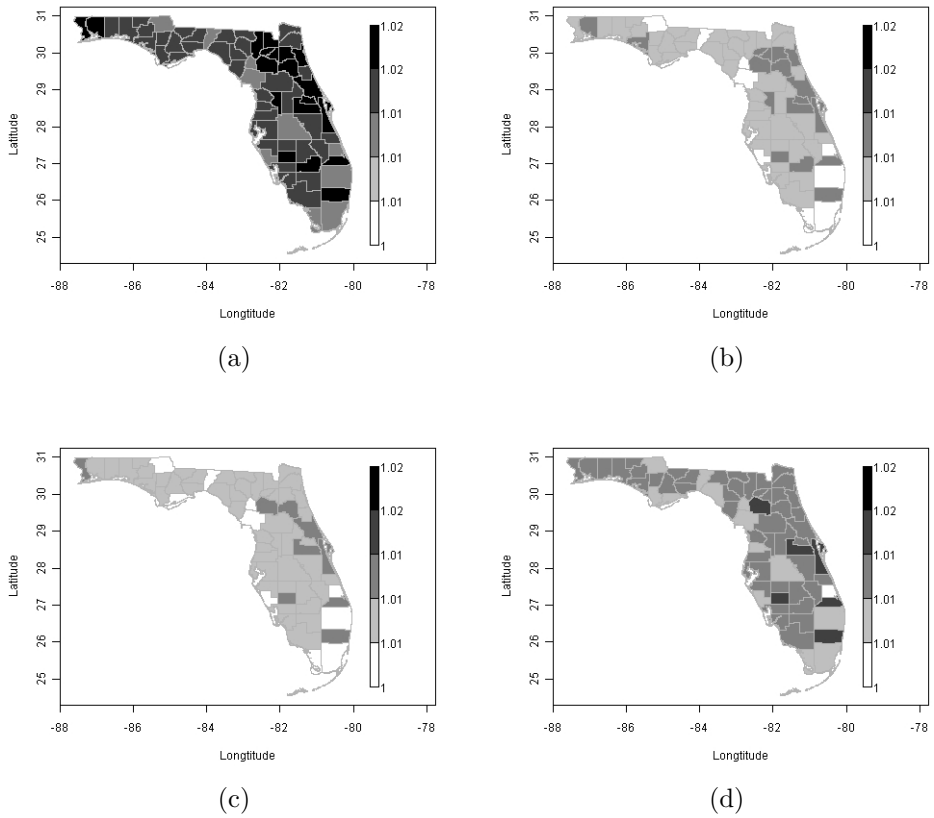


Figure 2: Relative SER from (a) the "Krige and Regress, Accounting for Covariance", (b)"Krige, Calibrate, and Regress", (c) "Conditional Simulation", and (d) the "Conditional Simulation and Calibrate"methods.

ment error associated with simulation, they made no effort to correct for that error. To do that, we use the leave-one-out cross-validation regression calibration approach based on the simple measurement error model (see [1]. For each of the 1000 potential environmental exposure realizations from conditional simulation, the predicted ozone exposure values were calibrated. Using the calibrated ozone values in the modeling process instead of the simulated values provides an adjustment for the bias that is anticipated from using the simulated values and error in the model predicting environmental exposure. The resulting predicted relative SERs from this "Conditional Simulation and Calibrate"approach are displayed in Figure 3(d).

References

[1] Carroll R.J., Ruppert D., Stefanski L.A., and Crainiceanu, C.M. 2006. *Measurement Error in Nonlindar Models, 2nd ed.* Chapman & Hall: Boca Raton, Fl.

[2] Gelfand A.E., Zhu L., and Carlin B.P. (2001) On the change of support problem for spatio-temporal data. *Biostatistics* 2: 3145.

[3] Gryparis A, Paciorek CJ, Zeka A, Schwartz J and Coull B. 2008. Measurement error caused by spatial misalignment in environmental epidemiology. *Biostatistics* doi:10.1093/biostatistics/kxn033.

[4] Little R.J.A. and Rubin D.B. (2002) *Statitical Analysis with Missing Data, 2nd ed.* Wiley: New York.

[5] Madsen L., Ruppert D., and Altman N.S. (2008) Regression with Spatially Misaligned Data. *Environmetrics* 19, 453-467

[6] Mugglin A.S., Carlin B.P., and Gelfand A.E. (2000) Fully model-based approaches for spatially misaligned data. *Journal of the American Statistical Association* 95: 877887.

[7] U.S. Centers for Disease Control and Prevention (CDC). (2006) *Health risks in the United States: Behavioral Risk Factor Surveillance System 2006.* US Department of Health and Human Services, CDC: Atlanta, GA.

[8] World Health Organization. (2005) *International Classification of Diseases and Related Health Problems (ICD-10)*, 2nd Ed. WHO.

[9] Woodward M. 2004. *Epidemiology: Study Design and Data Analysis*, 2nd Ed. Chapman & Hall/CRC: Boca Raton, Florida.

[10] Young L.J., Gotway C.A., Yang J., Kearney G., and Duclos C. (2009) Assessing uncertainty in support-adjusted spatial misalignment problems. *Communications in Statistics, Theory and Methods. In press.*

[11] Zhu L., Carlin B.P. and Gelfand A.E. (2008) Hierarchical regression with misaligned spatial data: relating ambient ozone and pediatric asthma ER visits in Atlanta. *Environmetrics* 14:537-557.

6th St.Petersburg Workshop on Simulation (2009) 699-703

Inequalities: Some Probabilistic, Some Matrix, Some Moment, and Some Classical

Ingram Olkin[1]

Abstract

There is a long history of probabilistic and statistical inequalities, of eigenvalue and matrix inequalities, and also moment inequalities. That there is a connection between these inequalities has not always been visible. Indeed, that probabilistic inequalities can yield matrix inequalities is somewhat elusive. Probabilistic inequalities often have the advantage of providing intuitive proofs. Furthermore, many probabilistic inequalities achieve equality for two-point distributions, in which case sharpness is readily exhibited. We here review the connection between probabilistic inequalities and moment and matrix inequalities. Some of the well-known probabilistic inequalities discussed are the Chebychev, Hajek-Renyi, Lyapunov inequalities, and more.

[1]Department of Statistics, Sequoia Hall, 390 Serra Mall, Stanford University, E-mail: iolkin@stat.stanford.edu

Session

Stochastic simulation of rare events and stiff systems

organized by Werner Sandmann (Germany)

6th St.Petersburg Workshop on Simulation (2009) 703-705

A Recursive Variance Reduction Technique with Bounded relative Error for Communication Network Reliability Estimation

Hector Cancela[1], Mohamed El Khadiri[2], Gerardo Rubino[3], Bruno Tuffin[4]

Abstract

Static network reliability estimation is an NP-hard problem and Monte Carlo simulation is therefore a relevant tool to provide an estimation. On the other hand, crude Monte Carlo is inefficient when dealing with rare events. This paper reviews a previously proposed Recursive Variance Reduction (RVR) algorithm and shows that it is not asymptotically efficient as reliabilities of individual links increase. We propose variations that are shown to verify the so-called Bounded relative Error (BRE) property.

1. Model

Reliability estimation is an important problem in networking, to know the probability that a group of nodes can communicate. We consider the case of *static* models, widely used in engineering, where time does not play any specific role. More specifically, a communication network is represented by an undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{K})$ where $\mathcal{V}$ is the set of nodes, $\mathcal{E} = \{1, \ldots, m\}$ is the set of links connecting nodes, and $\mathcal{K}$ is a subset of the node-set, called the terminal-set. Nodes are assumed perfect, i.e., they do not fail, while links can fail, link $e \in \mathcal{E}$ failing with probability $q_e = 1 - r_e$. All failure events of individual links are assumed independent.

The random state of the network is given by vector

$$X = (X_1, \ldots, X_m)$$

where X_e for $1 \leq e \leq m$ is a Bernoulli random variable giving 1 if link e is working, and 0 if it is failed.

We aim at investigating the probability r that that all nodes in the terminal-set $\mathcal{K}$ are connected. We define a structure function Φ of $\{0,1\}^m$ in $\{0,1\}$ such

[1]Universidad de la República, E-mail: cancela@fing.edu.uy
[2]Saint-Nazaire Institute of Technology, E-mail: Mohamed.El-Khadiri@univ-nantes.fr
[3]INRIA Rennes, E-mail: rubino@irisa.fr
[4]INRIA Rennes, E-mail: btuffin@irisa.fr

that $\Phi(x) = 1$ if all nodes in $\mathcal{K}$ are connected when the state is $x = (x_1, \ldots, x_m)$, and $\Phi(x) = 0$ otherwise. We then have $\Phi(X)$ Bernoulli random variable such that $\mathbb{E}[\Phi(X)] = r = r(\mathcal{G})$.

The number of states is 2^m, and increases exponentially with the number m of links. Actually the computation of r is known to be NP-hard [2] in general, and Monte Carlo simulation [1] is therefore a relevant solution method for this problem when $\mathcal{E}$ is of moderate to large size [5].

Furthermore, we are often interested in situations where links, and as a consequence the overall network, are highly reliable. In order to model this high reliability, and to study the robustness of the considered method with respect to it, we introduce a so-called rarity parameter ϵ such that $q_e = a_e \epsilon^{b_e}$ with $a_e, b_e > 0$ constants independent of ϵ. As $\epsilon \to 0$, $q_e \to 0$ and $r \to 1$ if we assume that $\Phi(1, \ldots, 1) = 1$ (otherwise, we are in the trivial case where $r = 0$). We willn rather look at $Y = Y(\mathcal{G}) = 1 - \Phi(X)$, and $\mathbb{E}[Y] = q = q(\mathcal{G}) = 1 - r << 1$, the network *unreliability*, to emphasize the quantity that is difficult to estimate, i.e., how far we are from 1.

2. Monte Carlo simulation

Crude Monte Carlo method consists in using n independent copies $X^{(i)} = (X_1^{(i)}, \ldots, X_m^{(i)})$ for $1 \leq i \leq n$ of X, that is copies of the graph $\mathcal{G}$, and computing $Y^{(i)} = 1 - \Phi(X^{(i)})$. The crude estimator of q is then

$$\hat{Y}_n = \frac{1}{n} \sum_{i=1}^{n} Y^{(i)}$$

and its variance is $\mathrm{Var}[\hat{Y}_n] = rq/n = q(1-q)/n$, easily estimated by $\hat{Y}_n(1 - \hat{Y}_n)/(n-1)$. The Central Limit Theorem then provides a confidence interval at confidence level α

$$\left(\hat{Y}_n - c_\alpha \sqrt{\frac{\hat{Y}_n(1-\hat{Y}_n)}{n-1}},\ \hat{Y}_n + c_\alpha \sqrt{\frac{\hat{Y}_n(1-\hat{Y}_n)}{n-1}} \right)$$

where $c_\alpha = \phi^{-1}((1+\alpha)/2)$ and ϕ is the cumulative distribution function of the standard Normal law.

There are unfortunately difficulties when applying the crude Monte Carlo estimator when q is small. First, it is unlikely that one of the $Y^{(i)}$ is equal to one, i.e., that the rare event is reached, unless the sample size n is large (it requires in average $n = 1/q$ to get the event once). Also, even if n is large enough, the *relative error*, that is the relative half-size of the confidence interval, verifies

$$c_\alpha \frac{\sqrt{\mathrm{Var}[\hat{Y}_n]}}{\mathbb{E}[Y]} =_\alpha \frac{\sqrt{q(1-q)}}{q\sqrt{n-1}} \to \infty$$

as $q \to 0$ (i.e., as $\epsilon \to 0$ for our model). It means that again that, in order to get a given relative precision, we need to increase the length of the simulation.

One of the main stream in rare event simulation is to find out estimators $\hat{Y}'_n$ of $\mathbb{E}[Y]$ enjoying the so-called *Bounded relative Error* (BRE) property, i.e., for which the relative error $c_\alpha \sqrt{\mathrm{Var}[\hat{Y}'_n]}/\mathbb{E}[Y]$ remains bounded as $\mathbb{E}[Y] \to 0$, or equivalently [8],

$$\text{we have BRE if } \frac{\mathbb{E}[Y^2]}{(\mathbb{E}[Y])^2} \text{remains bounded as } \mathbb{E}[Y] \to 0.$$

Many variance reduction algorithms have been applied to cope with the rare event issue for the static models [5]. Here, we are going to focus on the *Recursive Variance Reduction* (RVR) algorithm, one of the most successful methods.

3. Recursive Variance Reduction (RVR) algorithm and implementation

RVR algorithm proposed in [3] considers a $\mathcal{K}$-cutset, i.e., a set $\mathcal{C}$ of links whose failure ensures the system failure. If $q_{\mathcal{C}}$ is the probability that all links in $\mathcal{C}$ are failed, RVR replaces each $Y^{(i)}$ by $Y^{(i)}_{RVR}$, for n independent copies of Y_{RVR} with

$$Y_{RVR} = q_{\mathcal{C}} + (1 - q_{\mathcal{C}}) \sum_{j=1}^{|\mathcal{C}|} \mathbf{1}_{B'_j} Y_{RVR}(\mathcal{G}_j), \tag{1}$$

where

- $\mathcal{G}_j$ is the graph $\mathcal{G}$ with the $j-1$ first links of $\mathcal{C}$ failed, but the j-th working.
- $(B'_j)_{j=1..|\mathcal{C}|}$ with $B'_j = [B_j|A]$ is a sequence of disjoint events, where B_j the event "the $j-1$ first links of $\mathcal{C}$ are down, but the j-th is up" and A is "at least one link on $\mathcal{C}$ is up". The (conditional) probability of B'_j is $p_j = \mathbb{P}[B_j]/(1 - q_{\mathcal{C}}) = (\prod_{k=1}^{j-1} q_k) r_j/(1 - q_{\mathcal{C}})$.

It is shown in [3] that $\mathbb{E}[Y_{RVR}] = q(\mathcal{G}) = q$ and it can be checked that

$$\begin{aligned} \mathbb{E}[(Y_{RVR})^2] &= q_{\mathcal{C}}^2 + 2q_{\mathcal{C}}(1 - q_{\mathcal{C}}) \left(\sum_{j=1}^{|\mathcal{C}|} \frac{\mathbb{P}[B_j]}{1 - q_{\mathcal{C}}} \mathbb{E}[Y_{RVR}(\mathcal{G}_j)] \right) \\ &\quad + (1 - q_{\mathcal{C}})^2 \left(\sum_{j=1}^{|\mathcal{C}|} \frac{\mathbb{P}[B_j]}{1 - q_{\mathcal{C}}} \mathbb{E}[(Y_{RVR}(\mathcal{G}_j))^2] \right). \end{aligned} \tag{2}$$

In other words, the principle is to select a $\mathcal{K}$-cutset and to determine the first working link J on the cutset, forcing the event to happen (there are all failed with probability $q_{\mathcal{C}}$). Then Y_{RVR}, is replaced by $q_{\mathcal{C}} + (1 - q_{\mathcal{C}}) Y_{RVR}(\mathcal{G}_J)$ with $\mathcal{G}_J$ the resulting graph.

A new $\mathcal{K}$-cutset method can be found for $\mathcal{G}_J$, and the algorithm is applied recursively up to a graph with terminal nodes not connected (returning 1), or connected (returning 0).

A more efficient implementation of RVR algorithm is in [4] and consists, instead of recomputing the cutsets at each independent iterations, in distributing the n copies of Y_{RVR} among the $|\mathcal{C}|$ different $q_{\mathcal{C}} + (1 - q_{\mathcal{C}})Y_{RVR}(\mathcal{G}_j)$ according to a multinomial distribution with respective probability p_j. This significantly reduces the computational burden.

4. Asymptotic analysis and robust version

The RVR has been shown to be an efficient technique, but we prove now that it does not verify BRE property as $\epsilon \to 0$. Consider the simple graph $\mathcal{G}$ of Figure 1, where we want to compute the probability q that the two nodes s and t are not connected. Each link e unreliability is assumed to be $q_e = \epsilon$.

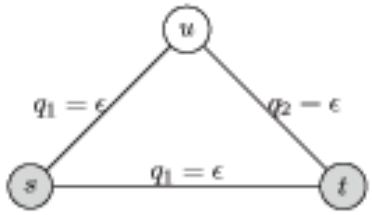


Figure 1: Simple graph topology for which RVR does not verify BRE.

Considering the cut with the two links starting from node s and ordering them as first the one from s to u, $q_{\mathcal{C}} = \epsilon^2$, and from (2),

$$\begin{aligned}\mathbb{E}[Y_{RVR}^2] &= \epsilon^4 + 2\epsilon^2\left((1-\epsilon)\mathbb{E}[Y_{RVR}(\mathcal{G}_1)] + \epsilon(1-\epsilon)\mathbb{E}[Y_{RVR}(\mathcal{G}_2)]\right) \\ &\quad +(1-\epsilon^2)\left((1-\epsilon)\mathbb{E}[(Y_{RVR}(\mathcal{G}_1))^2] + \epsilon(1-\epsilon)\mathbb{E}[(Y_{RVR}(\mathcal{G}_2))^2]\right).\end{aligned}$$

For the graph $\mathcal{G}_2$, the single link of interest is the one from s to t, and $Y_{RVR}(\mathcal{G}_2) = \epsilon$. Thus $\mathbb{E}[Y_{RVR}(\mathcal{G}_2)] = \epsilon$ and $\mathbb{E}[(Y_{RVR}(\mathcal{G}_2))^2] = \epsilon^2$. For the graph $\mathcal{G}_1$, as the link from s to u is working, this link can be contracted by merging s and u. The graph becomes a digraph with two links from s to t, each failing with probability ϵ. $\mathbb{E}[Y_{RVR}(\mathcal{G}_1)] = \epsilon^2$, but RVR is again used with the (single) cut containing the two links, and again leading to $\mathbb{E}[(Y_{RVR}(\mathcal{G}_1))^2] = \epsilon^4$. Finally, $\mathbb{E}[Y_{RVR}^2] = \epsilon^3 + 5\epsilon^4 - 6\epsilon^5 + \epsilon^7 = \Theta(\epsilon^3)$, and $\mathbb{E}[Y_{RVR}^2]/(\mathbb{E}[Y_{RVR}])^2 = \Theta(\epsilon^{-1}) \to \infty$ as $\epsilon \to 0$ (since $\mathbb{E}[Y_{RVR}] = q(\mathcal{G}) = 2\epsilon^2 - \epsilon^3$). As a consequence, we have the following property:

Proposition 1. *RVR algorithm does not verify Bounded Relative Error property.*

The problem actually comes from the sampling of the uniform random variable to select the first link of the cutset that does work.

Consider now the case where we combine RVR with Importance Sampling (IS) [1, 8]. Typically, we want to change the probability of selecting B'_j in (1) from $\mathbb{P}[B_j]/(1 - q_{\mathcal{C}})$ to a new one. Assume that we rather assign a uniform probability to the B'_j $(1 \le j \le |\mathcal{C}|)$, i.e., use new measure $\tilde{\mathbb{P}}[B'_j] = 1/|\mathcal{C}|$. This method, that

we call *Balanced RVR* (BRVR), requires to add the likelihood ratio $\mathbb{P}[B'_j]/\tilde{\mathbb{P}}[B'_j]$ in the estimator to keep it unbiased. This new estimator is, using $\tilde{\mathbb{P}}$ instead of $\mathbb{P}$,

$$Y_{BRVR} = q_{\mathcal{C}} + (1-q_{\mathcal{C}}) \sum_{j=1}^{|\mathcal{C}|} \mathbf{1}_{B'_j} \frac{\mathbb{P}[B'_j]}{\tilde{\mathbb{P}}[B'_j]} Y_{BRVR}(\mathcal{G}_j) \quad (3)$$

$$= q_{\mathcal{C}} + |\mathcal{C}| \sum_{j=1}^{|\mathcal{C}|} \mathbf{1}_{B'_j} \mathbb{P}[B_j] Y_{BRVR}(\mathcal{G}_j). \quad (4)$$

We then have the desired property:

Proposition 2. *BRVR algorithm verifies Bounded Relative Error property.*

As a sketch of proof, first notice that

$$\mathbb{E}[(Y_{BRVR})^2] = q_{\mathcal{C}}^2 + 2q_{\mathcal{C}}|\mathcal{C}| \left(\sum_{j=1}^{|\mathcal{C}|} \mathbb{P}[B_j]\mathbb{E}[Y_{BRVR}(\mathcal{G}_j)] \right) + |\mathcal{C}|^2 \left(\sum_{j=1}^{|\mathcal{C}|} (\mathbb{P}[B_j])^2 \mathbb{E}[(Y_{BRVR}(\mathcal{G}_j))^2] \right). \quad (5)$$

The proof then proceeds by induction. Let c, d, f_j and d_j be constants respectively such that $q_{\mathcal{C}} = \Theta(\epsilon^c)$, $\mathbb{E}[Y_{BRVR}] = \Theta(\epsilon^d)$, $\mathbb{P}[B_j] = \Theta(\epsilon^{f_j})$ and $\mathbb{E}[Y_{BRVR}(\mathcal{G}_j)] = \Theta(\epsilon^{d_j})$. If $\forall j$, $\mathbb{E}[(Y_{BRVR}(\mathcal{G}_j))^2] = \Theta((\mathbb{E}[Y_{BRVR}(\mathcal{G}_j)])^2)$, we have from (4) $d = \min_j(c, f_j + d_j)$. From (5), $\mathbb{E}[(Y_{BRVR})^2] = \Theta(\epsilon^{\min_j(2c, c+f_j+d_j, 2f_j+2d_j)})$ thus

$$\mathbb{E}[(Y_{BRVR})^2] = \Theta(\epsilon^{2d}) = \Theta(\mathbb{E}[Y_{BRVR}^2]).$$

For the simplest topologies, (the ones working as soon as a single link is working), the property is also verified, which completes the proof.

The multinomial version can be applied to BRVR as well.

5. Zero-variance version of RVR algoritm

Instead of sampling the first working component in $\mathcal{C}$ by a uniform distribution in BRVR, we can look at an importance sampling strategy which would lead to variance zero. Zero-variance approximation is indeed a promising area in Monte Carlo simulation [7]. The idea is, somewhat similarly to [6], to select B'_j with probability $\tilde{\mathbb{P}}[B_j]/(1-q_{\mathcal{C}})$ in the estimator (3), with

$$\tilde{\mathbb{P}}[B_j] = \frac{\mathbb{P}[B_j] q(\mathcal{G}_j)}{\sum_{k=1}^{|\mathcal{C}|} \mathbb{P}[B_k] q(\mathcal{G}_k)} \quad (6)$$

It can then be proved that the resulting estimator Y_{ZRVR} has variance $\mathrm{Var}[Y_{ZRVR}] = 0$, the subscrit $ZRVR$ being for *Zero-Variance RVR*. This can be proved by induction playing with variances, or just by showing that Y_{ZRVR} always gives q, using

the fact that

$$q = q_{\mathcal{C}} + \sum_{k=1}^{|\mathcal{C}|} \mathbb{P}[B_k] q(\mathcal{G}_k).$$

Unfortunately, implementing this estimator requires the knowledge of the $q(\mathcal{G}_i)$, but if we knew them, there would be no need for simulation. Instead, we suggest to use as an approximation of $q(\mathcal{G}_i)$ the probability $\tilde{q}(\mathcal{G}_i)$ of the cutset (generally a mincut) that used to compute $q(\mathcal{G}_i)$ in the RVR algorithm, and to plug it into (6).

Note that approximating the probability $q(\mathcal{G}_j)$ by the probability of a cutset is at no or minor cost additional cost, especially when implementing the multinomial version, since, computing cutsets has to be performed for RVR and its variants.

All missing proofs, as well as numerical illustrations, will be provided in the extended version of the paper.

References

[1] S. Asmussen and P. W. Glynn. *Stochastic Simulation*. Springer-Verlag, New York, 2007.

[2] Michael O. Ball. Computational complexity of network reliability analysis: An overview. *IEEE Transactions on Reliability*, 35(3):230–239, Aug. 1986.

[3] H. Cancela and M. El Khadiri. A recursive variance-reduction algorithm for estimating communication-network reliability. *IEEE Transactions on Reliability*, 44(4):595–602, 1995.

[4] H. Cancela and M. El Khadiri. On the rvr simulation algorithm for network reliability evaluation. *IEEE Transactions on Reliability*, 52(2):207–212, 2003.

[5] H. Cancela, M. El Khadiri, and G. Rubino. In G. Rubino and B. Tuffin, editors, *Rare Event Simulation using Monte Carlo Methods*, chapter Rare events analysis by Monte Carlo techniques in static models. John Wiley & Sons, 2009. To appear.

[6] H. Cancela, P. L'Ecuyer, M. Lee, G. Rubino, and B. Tuffin. Combining path-based methods, rqmc and zero-variance approximations in static model analysis. In *Seventh International Conference on Monte Carlo and quasi-Monte Carlo Methods*, July 2008.

[7] P. L'Ecuyer and B. Tuffin. Approximate zero-variance simulation. In *Proceedings of the 2008 Winter Simulation Conference*, pages 170–181. IEEE Press, 2008.

[8] G. Rubino and B. Tuffin, editors. *Rare Event Simulation using Monte Carlo Methods*. John Wiley & Sons, 2009. To appear.

6th St.Petersburg Workshop on Simulation (2009) 709-713

Proxel-Based Simulation: Theory and Applications

Claudia Krull[1], Graham Horton[2]

Abstract

Discrete stochastic models are widely used to describe current engineering and logistics problems. The stochastic simulation of such models can get very expensive if the models are stiff or rare system events are of interest. Proxels are a state space-based simulation technique that does not have these drawbacks. They implicitly use a discrete-time Markov chain to deterministically discover all possible system states at discrete points in time. Some applications have shown that Proxels are especially suitable for the analysis of small stiff models, and can outperform stochastic simulation techniques in that area.

1. Introduction

Discrete stochastic models can be used to describe some current problems in the industry. Their analysis is often performed using discrete event-based simulation (DES). Unfortunately, DES can get very expensive. When stiff models and rare events are involved, many replications are required to gain statistically meaningful results. The performance of DES is dependent on the degree of stiffness of the model or rareness of the event of interest. Existing methods for rare event simulation try to relieve that by modifying either the model or the problem specification. However, these methods can be very complex and are usually problem dependent in their application. Proxel-based simulation is a recently developed state space-based simulation approach, which is based on discrete-time Markov chains (DTMC). It is a deterministic algorithm and does not suffer a significant performance decrease when rare events are involved. Proxels are especially suitable for the simulation of small stiff models, discovering all possible system developments in one run and assigning them probabilities. In contrast to partial or ordinary differential equations, Proxels are more intuitive to use and not inherently limited to specific model classes. Using a generic implementation, Proxels can in principle be applied to any discrete stochastic model, instead of stochastic simulation techniques. The paper describes the basic idea of Proxels, two successful applications and some current extensions.

[1]Otto-von-Guericke-University Magdeburg, E-mail: claudia@sim-md.de
[2]Otto-von-Guericke-University Magdeburg, E-mail: graham@sim-md.de

2. State of the Art

2.1. Stochastic Simulation of Rare Event Models

The stochastic simulation of models involving rare events can become unfeasibly expensive. Many replications are needed to discover the rare events and even more to obtain statistically significant results for them. In general, the cost of a DES is dependent on the number of state changes performed per simulation run and on the number of simulation runs necessary. This implies that the cost increases with increasing degree of rareness of the event. Rare event simulation methods try to relieve this problem. Importance sampling modifies the model definition by changing transitions specifications to make the event of interest more frequent. Importance splitting defines intermediate thresholds that have to be crossed before reaching the rare event. Development paths are split at these thresholds in order to increase the number of times a rare event is encountered within the simulation runs. Both methods require a subsequent rescaling of the results to make them applicable to the original model. However, both methods are mathematically complex and usually require problem knowledge to be applied properly. Their performance still suffers somewhat when the degree of rareness of the event increases.

2.2. Discrete-Time Markov Chains

Discrete-time Markov chains (DTMC) are a well researched area of mathematical modeling (see [1] for a thorough introduction). They can represent the state space of a model including the state transitions defined by one step transition probabilities. If one can build a DTMC representing a models states and behavior, then the solution of that DTMC is comparably easy using existing algorithms. However, the state transitions in a DTMC are memoryless, they can only directly represent discretized exponential or geometric distributions. Continuous non-Markovian distributions, such as Normal, Weibull or Lognormal, cannot be represented directly in a Markov chain. A direct DTMC representation of a real system involving time dependent behavior is often not detailed enough to draw conclusions about the systems dynamics. The easy solution of a Markov chain comes at the expense of loosing details of the time dependent system behavior. This limits the applicability of DTMC solutions to problems were the exact dynamic system behavior is not of much importance.

3. Proxel Background and Theory

One approach applying the advantages of Markov chains to the analysis of non-Markovian models are supplementary variables, which extend a system state by logging the age of that state. [2] This leads to partial differential equations (PDE) as system description that can then be solved. Extending this idea, it is possible to make any process memoryless by logging the ages of all currently activated or relevant transitions. Doing this at discrete points in time enables to use an algorithmic approach, rather than setting up and solving PDEs analytically. Using

supplementary variables to make all non-Markovian processes of a model memoryless is the basic idea of the Proxel-based simulation method [3, 6].

A Proxel as defined in Equation (1) is a point S in the extended state space of the model – the discrete system state dS extended by the age of the relevant transitions $\vec{\tau}$ for a specific point in the simulation time t – , with the probability of that state p. Proxels are only generated at discrete points in time, which are multiples of the simulation time step. The probability to perform any active state change within one of these time steps can the be determined by the so-called instantaneous rate function (IRF), which is defined as in Equation (2).

$$P = (S, p) = ((dS, \vec{\tau}, t), p) \tag{1}$$

$$\mu(\tau) = \frac{f(\tau)}{1 - F(\tau)} \tag{2}$$

The following is a sketch of the Proxel-based simulation approach based on these ideas. `start` represents the initial system state, and `dt` the discrete simulation time step.

```
1  Create initial Proxel with initial system state
   at start of simulation time $((start,0,0),1)$
2  For each activated transition r of each Proxel at time t
3    Create Proxel for t+dt with the probability
     of state change r within dt, reset transition age r
4    Create Proxel for t+dt for the case of no state change
     with leftover probability, increase transition ages by dt
5    Store newly created Proxels in data structure
6  Repeat 2-5 until end of simulation time
```

The algorithm implicitly builds a DTMC of the reachable model state space. By extending the discrete system states by the transition activation times, all processes are made memoryless. This enables to determine a transient solution of the discrete stochastic model algorithmically.

The performance of the exact implementation largely depends on the data structure chosen for Proxel storage. In general, the Proxel approach is much more flexible than the original supplementary variables, because it does not require to setup and solve differential equations. In contrast to DES, the cost of the method is not influenced by the stiffness of the model. The algorithm deterministically discovers all possible states at discrete points in time. The smaller the simulation time step is, the more accurate these probabilities are. On the other hand, the simulation time step also determines the cost of the simulation. This enables a trade-off between accuracy and computation cost. Some further features and problems of Proxel-based simulation will be discussed in Section .

4. Two Example Applications of Proxels

This section describes two example applications of Proxel-based simulation to small stiff models. These nicely demonstrate the properties of Proxel-based simulation and exemplify the area where Proxels can outperform existing simulation methods.

4.1. Analysis of a Vehicle Warranty Model

The problem described here comes from an industry project carried out for the DaimlerChrysler AG (DC). [7] The task was to determine costs of different warranty strategies for the following scenario: The expiration of a car warranty is based on a maximum mileage and a maximum time (e.g. 10000 miles vs. 1 year). Failures within the warranty period incur costs for the manufacturer. The failure has a much smaller rate than the warranty expiration, however, the occurrence of a failure generates considerable cost. Therefore we are dealing with a stiff model.

The DES employed by DC needed a runtime of 20 to 30 hours to compute a cost estimate to the accuracy of one cent for one parameter set (years, mileage, cost per failure), using given failure and time to mileage distributions. The special-purpose Proxel-based algorithm developed for this case used mileage as the basic time unit and one mile as discrete simulation time step. The failure probability within the warranty period multiplied by the cost of a failure directly yielded the desired warranty cost. This approach was already quite fast, needing only few minutes to obtain a comparable result. In a second attempt, rough estimates obtained for larger simulation time steps were used to extrapolate a more accurate solution. This was possible because of a linear convergence of the solution parameter with decreasing simulation time step. This decreased the computation time to mere seconds for one parameter set. This application enabled DC to gain faster and more precise predictions for the warranty costs in only a fraction of the original time, eventually enabling a faster decision between warranty strategies.

4.2. Proxel-Based Queuing Simulation

Queuing analysis is an old subject in modeling and simulation. [1] It has recently become of interest again, since many problems in electronic communication can be described using queuing models. The goal of classical queuing analysis is to find an analytical expressions for the performance measures of a class of queuing systems. However, this is not always possible, depending on the system specification. As an alternative, DES can be employed, even though it is a lot less accurate and more expensive. Proxels can be a good alternative to DES, when no analytical solution is available. They are especially suitable for queuing simulation, because the discrete state space of a queuing model is usually small and the number of processes is limited to arrival and service of customers. Furthermore, queuing models can be very stiff, or rare system states are of interest, such as the overflow of a buffer in a switch and the resulting packet loss.

An example queuing system of a small call center is of type M/G/c/K. The exact problem specification is $Exp(1.25)/N(1; 0.2)/2/17$ with a Markovian arrival process, a normally distributed service process, two call-center agents as servers and a holding queue capacity of 15, system capacity of 17. The rare event of interest is the filling up of the queue and the resulting possibility to loose incoming calls. The Proxel solution needed only seconds to produce a meaningful result for the queue overflow probability. In contrast, a discrete event-based simulation of the system needed 15 minutes of computation time. The event of the queue filling up did not happen often enough, resulting in an inappropriate confidence interval

for the measure. The problem was by far to stiff for standard DES. See [5] for more details and examples.

For queuing simulation in general Proxels can be used to obtain exact results for analytically not tractable systems. They can also provide answers for systems that cannot be tackled using DES. Furthermore, Proxels can help obtain rough estimates for not yet formally analyzed problems.

5. Special Issues

This section discusses special issues and problems of the Proxel-based simulation method, as well as extensions that were already performed to reduce these problems. One major drawback of Proxel simulation and state space-based methods in general is the drastic increase in the number of system states due to the extension with supplementary variables. This so-called state space explosion limits the applicability of the methods to models with a small discrete state space. The effects of this state space explosion can be dampened somewhat by intelligent storage and retrieval strategies for Proxels. Two more fundamental strategies to tackle that problem have been implemented so far and will be described here briefly.

The first problem leading to state space explosion is that every continuous distribution is split into as many separate time steps as the support of the distribution needs, also covering very smooth parts with too many sampling points. Each one of those sampling points leads to a different age value and consequently to a separate Proxel that needs to be stored and processed. One solution to this is the combination with discrete phase-type distributions (DPH). [4] These can represent smooth distribution functions with much less sampling points, leading to a drastic reduction in the size of the expanded state space. This increases the size of the models that can be feasibly analyzed using Proxels. The combination of Proxels and DPH is possible because both are ways to represent a non-Markovian distribution with a segment of a DTMC.

The second problem leading to state space explosion is related to stiff models, because using the original algorithm, the fastest model transition determines the size of the time step that is used to discretize all distributions. If the model is stiff, this time step needs to be very small, and is inefficient for much slower transitions. The use of so-called variable time steps can help relieve that problem. [8] Here, every transition can be performed using a time step of optimal size. This strategy can reduce the computation cost for stiff models significantly, again enabling the analysis of larger models using Proxels.

The cost of a Proxel-based simulation algorithm increases with an increasing discrete state space and increasing number of concurrently activated transitions. It also increases with decreasing simulation time step size, leading to the above mentioned state space explosion, but on the other hand it also enables a trade-off between accuracy and cost of a computation. An extrapolation of the simulation results that were computed using larger time steps can be used to obtain more accurate results while reducing computation time. Summing up, current extensions and special purpose implementations of Proxels make the simulation method applicable to a significant group of real world problems.

6. Conclusion

Proxel-based simulation is a state space-based method well suitable for the analysis of small stiff models or models containing rare events. Two specific applications were described, demonstrating that. In contrast to DES and current methods for Rare event simulation, Proxels can deterministically discover all possible system states in one run and assign probabilities to them. The state space explosion inherent to this class of approaches limits the applicability to small models. However, it can be dampened somewhat through the use of discrete phase-type distributions or variable time steps. Proxels can be used to obtain accurate results in a limited computation time for some problems, where DES cannot be feasibly applied.

References

[1] Bolch G., Greiner S., de Meer H., Trivedi K. S. (1998) *Queuing Networks and Markov Chains*. John Wiley & Sons, New York.

[2] German R., Lindemann C. (1994) Analysis of stochastic petri nets by the method of supplementary variables. In *Proceedings of Performance Evaluation*, 20:317-335.

[3] Horton G. (2002) A new paradigm for the numerical simulation of stochastic petri nets with general firing times. In *Proceedings of the European Simulation Symposium 2002*, pp.129-136. SCS European Publishing House.

[4] Krull C. (2008) *Discrete-Time Markov Chains: Advanced Applications in Simulation*. PhD thesis, Otto-von-Guericke-University Magdeburg.

[5] Krull C., Horton G. (2007) Application of proxels to queuing simulation. In *Simulation and Visualization 2007*, Magdeburg, Germany, pp.299-310.

[6] Lazarova-Molnar S. (2005) *The Proxel-Based Method: Formalisation, Analysis and Applications*. PhD thesis, Otto-von-Guericke-University Magdeburg.

[7] Lazarova-Molnar S., Horton G. (2004) Proxel-based simulation of a warranty model. In *European Simulation Multiconference 2004*, pp.221-224. SCS European Publishing House.

[8] Wickborn F., Horton G. (2005) Feasible state space simulation: Variable time steps for the proxel method. In *Proceedings of the 2nd Balkan conference in informatics*, Ohrid, Macedonia, pp.446-453.

6th St.Petersburg Workshop on Simulation (2009) 715-719

Multistep Methods for Markovian Event Systems

Werner Sandmann[1]

Abstract

We consider multistep methods for accelerated trajectory generation in the simulation of Markovian event systems, which is particularly useful in cases where the length of trajectories is large, e.g. when regenerative cycles tend to be long, when we are interested in transient measures over a finite but large time horizon, or when multiple time scales render the system stiff.

1. Markovian Event Systems

Markovian models are widespread for modeling stochastic phenomena in a variety of domains. Typically, the models are given in a high-level description such as queueing networks, Petri nets, stochastic automata networks, or sets of coupled chemical reactions, amongst many others. In principle, they can be mapped to the stochastic process level in that they are uniquely defined by an initial probability distribution and a generator matrix. But in practice models tend to be very large. The size of the state space typically increases exponentially with the number of system components or, in other words, the model dimensionality. This effect is known as state space explosion and often causes models to be numerically intractable. One major advantage of simulation is that the state space need not be explicitly enumerated. Thus, a model description that reflects the event system character of the model is well suited, in particular for simulation purposes. In almost all relevant cases the structure of the underlying Markov chain is not arbitrary but state transitions correspond to certain events where similar events essentially have the same effect. Hence, they can be taken as specific discrete event systems [4], which provides a structured model description on an intermediate level of abstraction. For Markovian models the events need not be scheduled and the setting of Markovian event systems is also useful for numerical solution [6].

In order to describe a Markovian event system we have to define its state space and to specify all relevant events that may trigger state transitions. It is necessary to define under which conditions a certain event may occur, how it affects the system state and at which rate it occurs. Diverse formal specifications of Markovian event systems can be found in the literature. Here, we adopt the transition class formalism of [10]. Without loss of generality we assume that the state space is $\mathcal{S} \subseteq \mathbb{N}^d$. All events that trigger state transitions are classified according to their effects which yields transition classes. Formally, a transition

[1]University of Bamberg, Germany, E-mail: werner.sandmann@uni-bamberg.de

class is a triplet $\mathcal{C} = (\mathcal{U}, u, \alpha)$ where $\mathcal{U} \subseteq \mathbb{N}^d$ is the source state space containing all states in which the event or the corresponding state transition, respectively, is possible, $u : \mathcal{U} \to \mathbb{N}^d$ is the destination state function giving the new state $u(x) \in \mathbb{N}^d$ according to the state transition when the event occurs in state $x \in \mathcal{U}$, and $\alpha : \mathcal{U} \to \mathbb{R}$ is the transition rate function giving the rate $\alpha(x) \in \mathbb{R}$ at which the event or transition occurs in state $x \in \mathcal{U}$. Any Markovian model can be uniquely described by a set of such transition classes together with an initial distribution.

As a queueing network example consider a d-node tandem network with exponentially distributed service times where arrivals occur only at the first node according to a Poisson process with arrival rate λ. The service rates are denoted by $\mu_1, \ldots, \mu_d$ and the buffer capacities by $\nu_1, \ldots, \nu_d$. Hence, the different types of transitions are arrivals at node 1, moves from node i to node $i+1,\ 0 < i < d$ and departures from node d. Therefore, $d+1$ transition classes are sufficient:

$\mathcal{C}_1 = (\mathcal{U}_1, u_1, \alpha_1)$, where

- $\mathcal{U}_1 = \{(x_1, \ldots, x_d) \in \mathbb{N}^d : x_1 < \nu_1\}$,
- $u_1 : \mathbb{N}^d \to \mathbb{N}^d, x \mapsto u_1(x) = (x_1 + 1, x_2, x_3, \ldots, x_d)$,
- $\alpha_1 : \mathbb{N}^d \to \mathbb{R}, x \mapsto \alpha_1(x) = \lambda$;

$\mathcal{C}_i = (\mathcal{U}_i, u_i, \alpha_i), i = 2, \ldots, d$, where

- $\mathcal{U}_i = \{(x_1, \ldots, x_d) \in \mathbb{N}^d : x_{i-1} > 0, x_i < \nu_i\}$,
- $u_i : \mathbb{N}^d \to \mathbb{N}^d, x \mapsto u_i(x) = (x_1, \ldots, x_{i-2}, x_{i-1}-1, x_i+1, x_{i+1}, \ldots, x_d)$,
- $\alpha_i : \mathbb{N}^d \to \mathbb{R}, x \mapsto \alpha_i(x) = \mu_{i-1}$;

$\mathcal{C}_{d+1} = (\mathcal{U}_{d+1}, u_{d+1}, \alpha_{d+1})$, where

- $\mathcal{U}_{d+1} = \{(x_1, \ldots, x_d) \in \mathbb{N}^d : x_d > 0\}$,
- $u_{d+1} : \mathbb{N}^d \to \mathbb{N}^d, x \mapsto u_d(x) = (x_1, \ldots, x_{d-1}, x_d - 1)$,
- $\alpha_{d+1} : \mathbb{N}^d \to \mathbb{R}, x \mapsto \alpha_d(x) = \mu_d$;

It becomes clear that state-dependent rates can be easily incorporated just by corresponding transition rate functions. Also the state space may be infinite, which is then implicitly given by dropping the restrictions on the source state spaces. Phase-type distributed interarrival and service times can be modeled by properly defined transition classes for any change from one to the next phase.

As a chemical reaction set consider the enzyme-catalyzed substrate conversion

$$E + S \underset{c_2}{\overset{c_1}{\rightleftharpoons}} ES \overset{c_3}{\rightharpoonup} E + P \tag{1}$$

where c_1, c_2, c_3 denote associated reaction rate constants such that the corresponding state-dependent reaction rate computes as c_i times the number of possible combinations of the required reactants. States of corresponding Markovian models are similarly defined as states of a queueing network, namely by the number of molecules of each species. If we successively number the species E, S, ES, P, a state $x = (x_1, x_2, x_3, x_4)$ expresses that there are x_1 E-molecules, x_2 S-molecules, x_3 ES-molecules, and x_4 P-molecules. Then the transition classes corresponding to the stoichiometric equation (1) are the following.

$\mathcal{C}_1 = (\mathcal{U}_1, u_1, \alpha_1)$, where

- $\mathcal{U}_1 = \{(x_1, \ldots, x_4) \in \mathbb{N}^4 : x_1, x_2 > 0\}$,
- $u_1 : \mathbb{N}^4 \to \mathbb{N}^4, x \mapsto u_1(x) = (x_1 - 1, x_2 - 1, x_3 + 1, x_4)$,
- $\alpha_1 : \mathbb{N}^4 \to \mathbb{R}, x \mapsto \alpha_1(x) = c_1 x_1 x_2$;

$\mathcal{C}_2 = (\mathcal{U}_2, u_2, \alpha_2)$, where

- $\mathcal{U}_2 = \{(x_1, \ldots, x_4) \in \mathbb{N}^4 : x_3 > 0\}$,
- $u_2 : \mathbb{N}^4 \to \mathbb{N}^4, x \mapsto u_2(x) = (x_1 + 1, x_2 + 1, x_3 - 1, x_4)$,
- $\alpha_2 : \mathbb{N}^4 \to \mathbb{R}, x \mapsto \alpha_2(x) = c_2 x_3$;

$\mathcal{C}_3 = (\mathcal{U}_3, u_3, \alpha_3)$, where

- $\mathcal{U}_3 = \{(x_1, \ldots, x_4) \in \mathbb{N}^4 : x_3 > 0\}$,
- $u_3 : \mathbb{N}^4 \to \mathbb{N}^4, x \mapsto u_3(x) = (x_1 + 1, x_2, x_3 - 1, x_4 + 1)$,
- $\alpha_3 : \mathbb{N}^4 \to \mathbb{R}, x \mapsto \alpha_3(x) = c_3 x_3$.

Obviously, compared to a desription via generator matrices the transition class formalism for Markovian event systems provides a huge gain in storage requirements and is also well suited for immediate implementation. An important point regarding computer implementations is that the state space and the generator matrix of the underlying Markov chain is implicitly coded by logical predicates and simple functions that are both easy to implement.

2. Multistep Methods

Simulation of Markovian models is straightforward. It essentially consists of generating trajectories according to the Markov chain dynamics. In discrete-time the next state is chosen according to the transition probabilities. In continuous-time the same is done according to the jump probabilities after generating the exponentially distributed state holding time. If the interest is in steady-state distributions, generation of holding times can be skipped even in the continuous-time case by simulating the uniformized discrete-time Markov chain instead but for transient measures trajectories of the CTMC are generated. However, if the time horizon is large or the system is stiff corresponding to multiple time scales this becomes exceedingly slow such that accelerated trajectory generation is desirable.

The basic idea of multistep methods is to accelerate the trajectory generation via advancing the simulation by appropriately chosen time steps rather than simulating each single event explicitly. Multistep simulation methods for stochastic models have been proposed in several contexts, including computer and communication networks [1, 7, 8] that need not be Markovian. Here, we cast multistep simulation approaches for Markovian networks in the setting of Markovian event systems, which is inspired by approaches in chemical physics [2, 5, 9]. where state spaces are potentially infinite, the reaction system typically evolves on multiple time scales.

Let C be the number of transition classes. For $i = 1, \ldots, C$ define $v_i = u_i(x) - x$ and denote by K_i the random variable describing the number of times that an event/transition according to $\mathcal{C}_i$ occurs in the time interval $[t, t + \tau)$. Then

$$X(t + \tau) = X(t) + \sum_{i=1}^{C} v_i K_i. \tag{2}$$

Accordingly, a general algorithmic framework for approximate trajectory generation where the simulation is advanced by pre-defined time steps instead of simulating every single event is as follows:

Init $t := t_0$, $x := x_0$ and t_{end};

while $t < t_{end}$

1. Compute all $\alpha_i(x)$ and $\alpha(x) := \alpha_1(x) + \cdots + \alpha_C(x)$;
2. Choose a step size τ according to some appropriate rule;
3. Compute suitable estimates $\hat{k}_1, \ldots, \hat{k}_C$ for $K_1, \ldots, K_C$;
4. Set $t := t + \tau$ and update the system state x according to (2).

If $X(t) = x$ and all transition rate functions are constant in $[t, t + \tau)$, the random variable K_i is Poisson distributed with mean $\tau\alpha_i(x)$, that is for $k \in \mathbb{N}_0$:

$$P(K_i = k) = \frac{(\tau\alpha_i(x))^k \exp(-\tau\alpha_i(x))}{k!}, \qquad i = 1, \ldots, C. \tag{3}$$

Note that even state-independent transition rate functions are not necessarily constant over time but some can vanish if the corresponding transition is not any longer possible in a state that has been reached in the meantime. Handling all transition rate functions as if they were indeed constant gives an appropriate rule for Step 2 of the above algorithm, which then yields an approximate scheme for trajectory generation. The quality of the approximation relies on the validity of the assumption of approximately constant transition rate functions, which must be formally specified and can then be used to control the approximation error.

Direct multistepping computes $\hat{k}_1, \ldots, \hat{k}_M$ as realizations of the corresponding Poisson random variables. Obviously, (2) then becomes an explicit deterministic expression for $X(t + \tau)$ as a function of x and obeys similarities to the explicit (forward) Euler method for solving systems of ordinary differential equations (ODEs). If the state components x_i are large and the Poisson random variates are approximated by their means, (2) becomes the explicit Euler formula for the deterministic event rate equations. Therefore, direct multistepping is also referred to as explicit tau-leaping in the context of chemically reacting systems [5, 12].

However, explicit ODEs solvers are known to become instable for stiff ODE systems and this effect turns over to direct multistepping for stiff Markovian systems. For stiff ODEs, implicit methods such as the implicit (backward) Euler method are often better suited, which motivates similar approaches to multistepping for Markovian models. Unfortunately, a completely implicit multistep simulation method would require to generate random variates according to the Poisson distribution with means $\tau\alpha_i(X(t+\tau)), i = 1, \ldots, C$, which depend on the unknown

random state $X(t+\tau)$. Instead, a partially implicit version is considered [9, 12]. Rewriting the random variables K_i as $K_i - \tau\alpha_i(X(t)) + \tau\alpha_i(X(t))$ and evaluating all transition rate functions in the last term at $X(t+\tau)$ instead of $X(t)$ yields

$$X(t+\tau) = X(t) + \sum_{i=1}^{C} v_i \Big(K_i - \tau\alpha_i(X(t)) + \tau\alpha_i(X(t+\tau)) \Big). \tag{4}$$

Then, in a first step, all K_i are approximated by computing realizations of Poisson random variables as with direct multistepping. Once these realizations, now denoted by $k_1, \ldots, k_C$, have been generated and given $X(t) = x$, (4) becomes an implicit deterministic equation that can be solved by, e.g., Newton iteration. Typically, the resulting estimate $\hat{x}(t+\tau)$ for $X(t+\tau)$ is not integer-valued. Therefore, in practice, the estimates to be used for the updating in Step 4 of the above algorithm are obtained by rounding the corresponding term in (4) to the nearest integer. That is

$$\hat{k}_i = \text{round}(k_i - \tau\alpha_i(x) + \tau\alpha_i(\hat{x}(t+\tau))). \tag{5}$$

As an alternative to (4), motivated by the properties of the trapezoidal rule for solving systems of deterministic ODEs, [3] proposed to substitute (4) by

$$X(t+\tau) = X(t) + \sum_{i=1}^{C} v_i \Big(K_i - \frac{\tau}{2}\alpha_i(X(t)) + \frac{\tau}{2}\alpha_i(X(t+\tau)) \Big), \tag{6}$$

which sometimes yields higher accuracy. However, it depends on the specific problem at hand whether (4) or (6) should be preferred.

3. Summary of Simulation Experiences

The concepts as presented in the previous section have been mostly applied to chemically reacting systems in recent years. In this context it has been often empirically demonstrated that the simulation of *some* stiff systems can be significantly accelerated. The integrated framework of Markovian event systems and the transition class formalism render possible to apply similar multistep methods to the simulation of Markovian models in a broader class of application domains. We have studied the simulation of Markovian queueing networks where phase-type distributions are allowed for both interarrival and service time distributions. Multiple time scales and associated stiff systems in the queueing network context results from, e.g., servers that may fail and can be repaired, where failure rates are by orders of magnitude smaller than repair rates. It turns out that multistep methods are particularly suitable for accelerated simulation of queueing networks. Even if the systems are not stiff, they are usually enormously complex and direct multistepping significantly accelerates simulation at the expense of only a small loss in accuracy. In particular, for state-independent interarrival and service rates the loss of accuracy is most often negligible. Moreover, combining direct/explicit and implicit multistepping is likely to yield further improvements [2, 11, 12]. Hence, multistep methods for Markovian event systems are very promising to provide efficient simulation methods within a broad range of application domains.

References

[1] S. Bohacek, J. P. Hespanha, J. Lee, and K. Obraczka. A hybrid systems modeling framework for fast and accurate simulation of data communication networks. In *Proceedings of ACM SIGMETRICS'03*, (2003). 58–69.

[2] Y. Cao, D. T. Gillespie, and L. R. Petzold. The adaptive explicit-implicit tau-leaping method with automatic tau selection. *J. Chemical Physics*, 126(22):224101–224101–9, (2007).

[3] Y. Cao and L. R. Petzold. Trapezoidal tau-leaping formula for the stochastic simulation of biochemical systems. In *Proceedings of the 1st Conference on Foundations of Systems Biology in Engineering*, (2005), 149–152.

[4] C. G. Cassandras and S. Lafortune. *Introduction to Discrete Event Systems.* Springer, 2nd edition, (2008).

[5] D. T. Gillespie. Approximate accelerated stochastic simulation of chemically reacting systems. *J. Chemical Physics*,(2001), 115:1716–1732.

[6] W. K. Grassmann. Finding transient solutions in Markovian event systems through randomization. In W. J. Stewart, editor, *Numerical Solution of Markov Chains*, (1991), ch. 18, 357–372. Marcel Dekker, Inc.

[7] Y. Guo, W. Gong, and D. Towsley. Time-stepped hybrid simulation (TSHS) for large scale networks. In *Proceedings of IEEE INFOCOM 2000*, (2000), vol. 2, 441–450.

[8] A. Kochut and A. U. Shankar. Timestep stochastic simulation of computer networks using diffusion approximation. In *Proceedings of MASCOTS 2006*, (2006), 247–254.

[9] M. Rathinam, L. R. Petzold, Y. Cao, and D. T. Gillespie. Stiffness in stochastic chemically reacting systems: The implicit tau-leaping method. *J. Chemical Physics*, (2003), 119:12784–12794.

[10] W. Sandmann. Structured description of Markovian network models and its potentials for efficient rare event simulation. In *Proceedings of HetNets'04*, (2004), P39/1–10.

[11] W. Sandmann. *Exposition and streamlined formulation of adaptive explicit-implicit tau-leaping.* Technical report, University of Bamberg (2009).

[12] W. Sandmann. Rare event simulation methodologies in systems biology. In G. Rubino and B. Tuffin, editors, *Rare Event Simulation using Monte Carlo Methods*, (2009), ch. 11. John Wiley & Sons.

6th St.Petersburg Workshop on Simulation (2009) 721-725

An adaptive branching splitting algorithm under cost constraint for rare event analysis

Agnès Lagnoux[1]

Abstract

This paper deals with the splitting method first introduced in rare event analysis. In this technique, the sample paths are split into R multiple copies at various stages to speed up simulation. Given the cost, the optimization of the algorithm suggests to take all the transition probabilities equal; nevertheless, in practice, these quantities are unknown. In this paper, we present an algorithm in two steps that copes with that problem.
Keywords : splitting method, simulation, cost function, Laplace transform, Galton-Watson, branching processes, iterated functions, rare event

The study of rare events is an important area in the analysis and prediction of major risks as earthquakes, floods, air collision risks, etc. Studying the major risks can be taken up by two main approaches which are the statistical analysis of collected data and the modelling of the processes leading to the accident. The statistical analysis of extreme values needs a long observation time since the very low probability of the events considered. The modelling approach consists first in formalizing the system considered and then in using mathematical ([1] and [15]) or simulation tools to obtain some estimates.

Analytical and numerical approaches are useful, but may require many simplifying assumptions. On the other hand, Monte Carlo simulation is a practical alternative when the analysis calls for fewer simplifying assumptions. Nevertheless, obtaining accurate estimates of rare event probabilities, say about 10^{-9} to 10^{-12}, using traditional techniques require a huge amount of computing time.

Many techniques for reducing the number of trials in Monte Carlo simulation have been proposed, like importance sampling (see e.g. [4] and [8]) or trajectory splitting. In this technique, we suppose there exists some well identifiable intermediate system states that are visited much more often than the target states themselves and behave as gateway states to reach the rare event. Thus we consider a decreasing sequence of events B_i leading to the rare event L:

$$L = L_{M+1} \subset L_M \subset \cdots \subset L_1 \,.$$

Then $\mathbb{P}(L)$ is given by

$$\mathbb{P}(L) = \mathbb{P}(L|L_M)\mathbb{P}(L_M|L_{M-1}) \cdots \mathbb{P}(L_2|L_1)\mathbb{P}(L_1), \tag{1}$$

[1]Université Toulouse 2. E-mail: lagnoux@univ-tlse2.fr. URL: http://www.math.univ-toulouse.fr/ lagnoux/

where on the right hand side, each conditioning event is "not rare". For the applications we have in mind, these conditional probabilities are in general not available explicitly. Instead we know how to make evolve the particles from level L_i to the next level L_{i+1} (e.g. Markovian behavior).

The principle of the algorithm is at first to run simultaneously several particles starting from the level L_i; after a while, some of them have evolved "badly", the other have evolved "well" i.e. have succeeded in reaching the threshold L_{i+1}. Then "bad" particles are moved to the position of the "good" ones and so on until L is reached. In such a way, the more promising particles are favoured; unfortunately that algorithm is hard to analyse directly because of the interaction introduced between particles and may be difficult to apply. Examples of this class of algorithms can be found in [2] with the "go with the winners" scheme, in [9, 6] in the context of the approximate counting and in [7, 5, 3] in a more general setting.

Nevertheless, all these algorithms lie on a common base, simpler to analyse and called branching splitting model. In this technique, we make a $\{0, 1\}$ Bernoulli trial to check whether or not the set event L_1 has occured. In that case, we split this trial in R_1 Bernoulli subtrials, and for each of them we check again whether or not the event L_2 has occured. This procedure is repeated at each level, until L is reached. If an event level is not reached, neither is L, then we stop the current retrial. Using N independent replications of this procedure, we have then considered $NR_1 \cdots R_M$ trials, taking into account for example, that if we have failed to reach a level L_i at the i-th step, the $R_i \cdots R_M$ possible retrials have failed. Clearly the particles reproduce and evolve independently.

An unbiased estimator of $\mathbb{P}(L)$ is given by the quantity

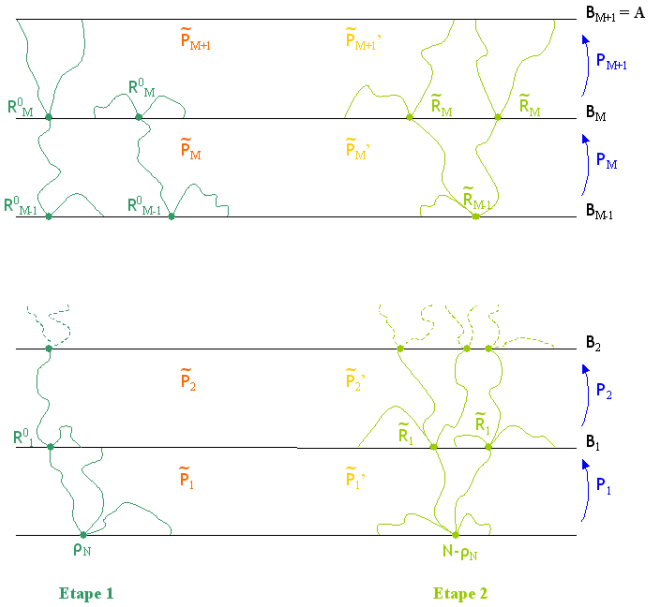
$$\widehat{P} = \frac{N_L}{N \prod_{i=1}^{M} R_i}, \tag{2}$$

where N_L is the total number of trajectories having reached the set L. Considering that this algorithm is represented by N independent Galton-Watson branching processes $(Z_n)_n$, as done in [10], the variance of $\widehat{P}$ can be then derived and depends on the probability transitions and on the mean numbers (m_i) of particles successes at each level. Lead by the heuristic presented in [16, 17], an optimal algorithm is derived by minimizing the variance of the estimator for a given budget (computational cost), defined as the expected number of trials generated during the simulation, where each trial is weighted by a cost function.

The optimization of the algorithm [10] suggests to take all the transition probabilities equal to a constant P_0 and the numbers of splitting equal to the inverse of this constant. We then deduce the number of thresholds M and finally N is given by the cost. This result is not surprising since it means that the branching processes are critical Galton-Watson processes. In other words, optimal values are chosen in such a way to balance the loss of variance from too little splitting and the exponential growth in computational effort from too much splitting.

Some pratical problems arises while applying the optimal algorithm in concrete models issued from reality. First, the optimal splitting number can be non integer. In [11], the author proposes three algorithms to face this problem. Then, for the applications we have in mind, the thresholds B_i are fixed but the conditional

probabilities are unknown as said before (instead we know how to make evolve the particles from level L_i to the next level L_{i+1}). Moreover, we assume here that they lie in some compact $[a,b] \subset]0,1[$. This hypothesis is essential: otherwise, nothing can be done algorithmically. In practice, it is generally implicit but nothing is said about it. We propose here an algorithm in two phases based on the branching



splitting model: the first one is a learning phase in which we sample ρ_N particles. The algorithm proceeds as in the classical branching splitting method with splitting numbers $(R_i^0)_{i=1\cdots M}$ chosen arbitrarly at the beginning. In the second phase, we run $N-\rho_N$ particles that we make evolve as in the first phase but with splitting numbers estimators of the optimal splitting numbers $(R_i)_{i=1\cdots M}$; the estimators being obtained during the first learning phase and following the optimal rule given in [10]. Since the complexity of the formulas, we will simply lead an asymptotic study when the cost C goes to infinity. Assuming the transition probabilities lie in a compact implies that the cost by particle is bounded below and above which allows us to lead the survey when N goes to infinity. A precise analysis shows that we shall dedicate asymptotically $\mu_s C^{2/3}$ particles to the learning phase and $C/C_{opt} - \mu_s C^{2/3}$ to the second phase, where C_{opt} is an explicit constant and μ_s is derived by the optimization of the algorithm; i.e. assuming that the number of particles generated during the learning phase behaves like $\mu_\alpha(C)C^{1-\alpha}$, we shall take $\alpha = 1/3$. Moreover, we note that N is linear in C, and so dedicating $\mu_s C^{2/3}$ particles to the first phase amounts to dedicate it $\lambda_s N^{2/3}$ particles, for some λ_s depending on μ_s.

Finally, we compare the results given by this algorithm and those obtained following the techniques detailled in [14] and [3] in the context of a modified Ornstein–Ulhenbeck diffusion.

All these results can be found in [13].

References

[1] Aldous, David. Probability Approximations via the Poisson Clumping Heuristic. of *Applied Mathematical Sciences* v. 77. (1989).

[2] Aldous, Davis and Vazirani, Umesh V.,≪Go with the winners≫ algorithms, IEEE Symposium on Fundations of Computer Science, 7 (1994) 492-501.

[3] F. Cérou and A. Guyader, Adaptive multilevel splitting for rare event analysis. *Rapport de recherche de l'INRIA - Rennes , Equipe : ASPI*, 2005.

[4] P.T. de Boer, Analysis and efficient simulation of queueing models of telecommunication systems, PhD thesis, University of Twente, The Nederlands, 2000.

[5] P. Del Moral, Feynman-Kac formulae, Genealogical and interacting particle systems with applications, Probability and its Applications (New York), Springer-Verlag, New York, 2004.

[6] P. Diaconis and S. Holmes, Three examples of Monte-Carlo Markov chains: at the interface between statistical computing, computer science, and statistical mechanics, Discrete probability and algorithms (Minneapolis, MN, 1993), IMA Vol. Math. Appl., 72 (1995) 43-56.

[7] A. Doucet and N. de Freitas and N. Gordon, An introduction to sequential Monte Carlo methods, Sequential Monte Carlo methods in practice, Stat. Eng. Inf. Sci., Springer, (2001) 3-14.

[8] P. Heidelberger, Fast Simulation of Rare Events in Queueing and Reliability Models, ACM Transactions on Modeling and Simulation, 5(1) (1995) 43-85.

[9] M. Jerrum and A. Sinclair, The Markov chain Monte Carlo method: an approach to approximate counting and integration, Approximation Algorithms for NP-hard Problems, (1997) 482-520.

[10] A. Lagnoux, Rare event simulation, Probability in the Engineering and Informational Sciences, 20 (2006) 45-66.

[11] A. Lagnoux-Renaudie, Effective branching splitting method under cost constraint, Stochastic Processes and their Applications, 118 (2008) 1820-1851.

[12] A. Lagnoux-Renaudie, A Two-Step branching splitting model under cost constraint for rare event analysis, Submitted to Applied Probability Journals.

[13] A. Lagnoux-Renaudie, A Two-Step branching splitting model under cost constraint for rare event analysis, Submitted to Applied Probability Journals.

[14] F. LeGland and N. Oudjane, A Sequential Particle Algorithm that Keeps the Particle System Alive, Stochastic Hybrid Systems : Theory and Safety Critical Applications, Lecture Notes in Control and Information Sciences 337, Springer-Verlag, Berlin (2006).

[15] J.S. Sadowsky, On Monte Carlo estimation of large deviations probabilities, Ann. Appl. Probab., 6 (1996) 2 399-422.

[16] M. Villén-Altamirano and J. Villén-Altamirano, RESTART: a Method for Accelerating Rare Event Simulations, 13th International Teletraffic Congress, Copenhagen, (1991) 71-76.

[17] Villén-Altamirano, Manuel and Villén-Altamirano, José, Restart: An Efficient and General Method for Fast Simulation of Rare Event, Tech. Rept. No 7, Departamento de Mathematica Aplicada, E.U. Informática, Universidad Polytéchnica de Madrid, 1997.

6th St.Petersburg Workshop on Simulation (2009) 726-731

Rare events simulation and conditioned random walks: the moderate deviation case

M. Broniatowsky[1]

This talk focuses on Importance Sampling for moderate deviations of the sample mean of i.i.d. real valued summands under the Cramer condition. Applications to M estimators are presented, as well as numerical results.

The r.v's $\mathbf{X}_i's$ are i.i.d. , are centered with variance 1, with common density $p_{\mathbf{X}}$ on $\mathbb{R}$, and

$$\mathbf{Z} := \frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_i =: \frac{1}{n}\mathbf{S}_1^n$$

is the empirical mean of the $\mathbf{X}_i's$. The set A is the interval (a_n, ∞) where a_n tends slowly to $E(\mathbf{X}_1)$ from above. Denote

$$P_n := P\left(\frac{1}{n}\mathbf{S}_1^n \in A\right).$$

Many asymptotic results provide sharp estimates for $P(\mathbf{Z} \in A)$ but it is a known fact that asymptotic expansions are not always good tools when dealing with numerical approximations for fixed (even large) n. For example, citing Ermakov (2004, Thory of Probability and Applications, p 624), the Berry-Esseen approximation for the evaluation of risks of order 10^{-2} in testing is pertinent for sample sizes of order 5000-10000; also the accuracy of available moderate deviation probabilities as developped by Inglot, Kallenberg and Ledwina (Annals probability 1992) has not been investigated. This motivates our interest in numerical techniques in this field.

The basic estimate of P_n is defined as follows: generate L i.i.d. samples $X_1^n(l)$ with underlying density p_X and define

$$P^{(n)}(\mathcal{E}_n) := \frac{1}{L}\sum_{l=1}^{L}\mathbf{1}_{\mathcal{E}_n}\left(X_1^n(l)\right)$$

where

$$\mathcal{E}_n := \{(x_1, ..., x_n) \in \mathbb{R}^n : s_1^n/n > a_n\}.$$

Here $s_1^n := x_1 + ... + x_n$. The statistics $P^{(n)}(\mathcal{E}_n)$ estimates the *moderate deviation* probability of the sample mean of the $\mathbf{X}_i's$. Also denoting g a sampling density of

[1]Université Paris 6, France, Michel.Broniatowski@upmc.fr, joint work with Yaakov Ritov, Hebrew University, Jerusalem, Israel

the vector Y_1^n the associated IS estimate is

$$P_g^{(n)}(\mathcal{E}) := \frac{1}{L}\sum_{l=1}^{L}\frac{p_X\left(Y_1^n(l)\right)}{g\left(Y_1^n(l)\right)}\mathbf{1}_{\mathcal{E}}\left(Y_1^n(l)\right). \tag{1}$$

when g In the range of moderate deviations the two major contributions to IS schemes for the estimation of P_n are Fuh and Hu (Biometrika 2004) and Ermakov (Statistics and Decision, 2007). The paper by Fuh and Hu does not consider events of moderate deviations as intended here; it focuses on IS schemes for the estimation of $P(Z \in A)$ where Z is a given multinormal random vector and A is a fixed set in $\mathbb{R}^d$. The authors consider efficiency with respect to the variance of the estimate and state that for the case of interest the efficient sampling scheme is deduced from the distribution of Z by a shift in the mean inside the set A. The papers by Ermakov instead handle similar problems as we do. Ermakov's 2007 considers a sampling scheme where g is the density of i.i.d. components. He proves that this scheme is efficient in the sense that the computational burden necessary to obtain a relative precision of the estimate with respect to P_n does not grow exponentially as a function of n. He considers statistics of greater generality than the sample mean, such as M and L estimators; in the range of moderate deviations the asymptotic behavior of those objects is captured however through their linear part which is the empirical mean of their influence function, which puts the basic situation back at the center of the scene. We discuss efficiency and present some results in connection with Ermakov's pertaining to M and L estimators.

The numerator in the expression (1) is the product of the $p_{\mathbf{X}_1}(Y_i)'s$ while the denominator need not be a density of i.i.d. copies evaluated on the $Y_i's$. Indeed the optimal choice for g is the density of $\mathbf{X}_1^n$ conditioned upon $\mathcal{E}_n$, say $p_{\mathbf{X}_1^n/\mathcal{E}_n}$.

Since the optimal solution is known to be $p_{\mathbf{X}_1^n/\mathcal{E}_n}$, the best its approximation, the best the sampling scheme. Classical sampling schemes consist in simulation of independent copies of r.v.'s $Y_i(l)$,$1 \leq i \leq n$, and efficiency is defined in terms of variance of the estimate inside this class of sampling, which, by nature, is suboptimal with respect to sampling under good approximations of $p_{\mathbf{X}_1^k/\mathcal{E}_n}$ for long runs, i.e. for large $k = k_n$.The present paper explores the choice of good sampling schemes from this standpoint. Looking at (1) it is also remarked that a sharp approximation of $p_{\mathbf{X}_1^n/\mathcal{E}_n}$ need not be completed on the entire space $\mathbb{R}^n$ but only on typical paths under the sampling scheme Obviously mimicking the optimal scheme results in a net gain on the number L of replications of the runs which are necessary to obtain a given accuracy of the estimate with respect to P_n . However the criterion which we consider is different from the variance, and results as an evaluation of the MSE of our estimate on specific subsets of the runs generated by the sampling scheme, which we call typical subsets, namely having probability going to 1 under the sampling scheme as n increases. On such sets, the MSE is proved to be of very small order with respect to the variance of the classical estimate, whose MSE cannot be diminuished on any such typical subsets. We believe that this definition makes sense and prove it also numerically. It will be shown that the relative gain in terms of simulation runs necessary to perform an

$\alpha\%$ relative error on P_n drops by a factor $\sqrt{n-k}/\sqrt{n}$ with respect to the classical IS scheme. Here $k = k_n$ satisfies $\lim_{n\to\infty} k/n = 1$.

Our proposal therefore hinges on the local approximation of the conditional distribution of longs runs $\mathbf{X}_1^k$ from $\mathbf{X}_1^n$. This cannot be achieved through the classical theory of moderate deviations; at the contrary the ad hoc procedure developped in the range of large deviations by Diaconis and Freedman (Theoretical Probability 1988) for the local approximation of the conditional distribution of $\mathbf{X}_1^k$ given the value of $\mathbf{S}_1^n$ is the starting point of the present approach. We will also present some numerical results in the range of large deviations

Session

Goodness-of-fit and related methods

organized by Simos Meintanis (Greece)

6th St.Petersburg Workshop on Simulation (2009) 731-735

Tests for symmetric error distribution in regression models[1]

Marie Hušková[2], Simos G. Meintanis[3]

Abstract

In linear and nonparametric regression models, the problem of testing for symmetry of the distribution of errors is considered. We propose a test statistic which utilizes the empirical characteristic function of the corresponding residuals. Here asymptotic properties of the test statistic are stated and discussed. The talk will contain also more detailed asymptotic results, discussions as well as a simulation study compares bootstrap versions of the proposed test to other more standard procedures.

1. Introduction

Assume $(Y, \boldsymbol{X})$ are observations from the general model

$$Y = m(\boldsymbol{X}) + \sigma(\boldsymbol{X})\varepsilon, \tag{1}$$

where $m(\cdot)$ and $\sigma^2(\cdot)$ denote the regression and variance functions, respectively, and the error ε is assumed to have an unspecified distribution function (DF), F_ε –with some properties specified later. The corresponding characteristic function (CF) and its imaginary part are denoted by $\varphi_\varepsilon(t)$ and $S_\varepsilon(t)$, respectively.

On the basis of independent observations $\{Y_j, \boldsymbol{X}_j\},\ j = 1, 2, ..., n$, with $\boldsymbol{X}_j = (X_{j1}, X_{j2}, ..., X_{jp})^{\mathrm{T}}$, we wish to test the null hypothesis of symmetry

$$H_0 : F_\varepsilon(u) = 1 - F_\varepsilon(-u),\ u \in \mathbb{R} \tag{2}$$

for the error distribution. In view of a well known characterization of symmetry around the origin, the null hypothesis may equivalently be stated as $S_\varepsilon(t) = 0,\ t \in \mathbb{R}$.

Two cases of the general regression model (1.1) will be considered: The linear regression case with homoscedasticity with general p and nonrandom design, and the nonparametric heteroskedastic regression model with unspecified heteroscedasticity, $p = 1$ and random design.

[1]This work was supported by grant MSM 0021620839, and GAČR 201/09/0755.

[2]Charles University in Prague, E-mail: huskova@karlin.mff.cuni.cz

[3]National and Kapodistrian University, E-mail: simonmei@econ.uoa.gr

In both cases the test statistic incorporates the standardized residuals,

$$e_j = (Y_j - \widehat{m}_n(\boldsymbol{X}_j))/\widehat{\sigma}_n(\boldsymbol{X}_j),\ j = 1, 2, ..., n. \tag{3}$$

In the linear regression case $\widehat{m}_n(\boldsymbol{X}_j) = \boldsymbol{X}_j^{\mathrm{T}}\widehat{\beta}_n$, with $\widehat{\beta}_n$ being an appropriate estimator of β, and $\widehat{\sigma}_n(\boldsymbol{X}_j) = \widehat{\sigma}_n$, while in nonparametric regression $\widehat{m}_n(\cdot)$ and $\widehat{\sigma}_n^2(\cdot)$ are kernel estimators of $m(\cdot)$ and $\sigma^2(\cdot)$, respectively.
Following earlier works on testing symmetry with i.i.d. data, see for instance Feuerverger and Mureika (1977), Koutrouvelis (1985), Ghosh and Ruymgaart (1992), Neuhaus and Zhu (1998), and Henze et al. (2003), we propose to use the corresponding empirical CF

$$\varphi_n(t) = \frac{1}{n}\sum_{j=1}^{n} e^{ite_j},$$

in the test statistic

$$T_{n,w} = n\int_{-\infty}^{\infty} S_n^2(t)w(t)dt, \tag{4}$$

where $S_n(t)$ denotes the imaginary part of the empirical CF, and $w(t)$ is an appropriate weight function. Rejection of the null hypothesis is for large values of $T_{n,w}$.
As for motivation, it should be noted that many authors have stressed the significance of symmetric errors in linear and nonparametric regression. Bickel (1982) for instance argues that in case of symmetry around the origin the slope parameters in linear regression can be adaptively estimated so that the resulting estimators share the efficiency of the maximum likelihood estimators computed by forming a likelihood with the actual (but unknown) error distribution correctly specified; refer also to Newey (1988), Kappenman (1988), Fan and Gencay (1995), Neumeyer et al. (2005), Hettmansperger et al. (2002), Dette et al. (2002), and Neumeyer and Dette (2007).
Section 2 contains basic theoretical properties of the proposed test procedures. Various comments and remarks are in Section 3.

2. Assumptions and basic asymptotic results

The basic theoretical results will be carried out in the convenient setting of the Hilbert space of measurable functions $f : \mathbb{R} \to \mathbb{R}$, endowed with the norm $\|f\| = \left(\int_{-\infty}^{\infty} f^2(t)w(t)\,dt\right)^{1/2}$. The notation $\to^D$ means convergence in distribution of random elements and random variables.

(a) Linear Regression case

Here we assume that $Y_1, \ldots, Y_n$ are independent observations following the linear model

$$Y_j = \mathbf{x}_j^{\mathrm{T}}\beta + \varepsilon_j,\ j = 1, 2, ..., n, \tag{5}$$

where $\mathbf{x}_j = (1, x_{j2}, ..., x_{jp})^{\mathrm{T}} \in \mathbb{R}^p,\ j = 1, 2, ..., n$, are known regressors, $\beta \in \mathbb{R}^p$ denote unspecified regression parameters, errors $\varepsilon_j,\ j = 1, 2, ..., n$, are assumed to be independent copies on a random variable having distribution function $F_\varepsilon(\cdot)$. We wish to test the null hypothesis (2).
In this case the test procedures utilizes the residuals,

$$\widehat{e}_j = Y_j - \mathbf{x}_j^{\mathrm{T}} \widehat{\boldsymbol{\beta}}_n,\ j = 1, 2, ..., n, \tag{6}$$

where $\widehat{\boldsymbol{\beta}}_n$ is an estimator of $\boldsymbol{\beta}$.

We assume that the characteristic function of error terms $\varphi_\varepsilon(.)$, estimators $\widehat{\beta}_n$ of regression parameters, and the regressors fulfil:

(A.1) $\varepsilon_1, \ldots, \varepsilon_n$ are i.i.d. random variables with the symmetric distribution function $F_\varepsilon(.)$ and the characteristics function $\varphi_\varepsilon(.)$.

(A.2) $\widehat{\boldsymbol{\beta}}_n = \widehat{\boldsymbol{\beta}}_n\left(\{\mathbf{x}_j, Y_j\}; j = 1, 2, ..., n\right)$ is a regression invariant estimator of $\boldsymbol{\beta}$, i.e., $\widehat{\beta}_n\left(\{\mathbf{x}_j, Y_j + \mathbf{x}_j^{\mathrm{T}}\boldsymbol{v}\}; j = 1, 2, ..., n\right) = \hat{\boldsymbol{\beta}}_n\left(\{\boldsymbol{x}_j, Y_j\}; j = 1, 2, ..., n\right) + \boldsymbol{v}$, for each $\mathbf{v} \in \mathbb{R}^p$. Moreover, it is assumed that as $n \longrightarrow \infty$,

$$\Big(\sum_{j=1}^{n} \boldsymbol{x}_j \boldsymbol{x}_j^T\Big)^{1/2} (\widehat{\boldsymbol{\beta}}_n - \boldsymbol{\beta}) = O_P(1)$$

$$\frac{1}{\sqrt{n}} \sum_{j=1}^{n} \mathbf{x}_j^{\mathrm{T}} (\widehat{\beta}_n - \beta_0) = \frac{d_\beta}{\sqrt{n}} \sum_{j=1}^{n} \psi_\beta(\varepsilon_j) + o_{\mathrm{P}}(1), \tag{7}$$

where ψ_β is a measurable antisymmetric function with $\mathrm{E}\psi_\beta(e_1) = 0$ and $\mathrm{E}\psi_\beta^2(e_1) < \infty$.

(A.3) $\lim_{n \longrightarrow \infty} \max_{1 \le v \le n} \boldsymbol{x}_v^T \Big(\sum_{j=}^{n} \boldsymbol{x}_j \boldsymbol{x}_j^T\Big)^{-1} \boldsymbol{x}_v = 0.$

(A.4) The weight function $w(.)$ is nonnegative, symmetric and $\int_{-\infty}^{\infty} t^2 w(t) dt < \infty$.

Next, the asymptotic null distribution of the test statistic is stated.

Theorem 0.1. *Let $Y_1, \ldots, Y_n$ follow the model (1) and let the assumptions (A.1) – (A.4) , be satisfied. Then*

$$T_{n,w} = ||W_n||^2 + o_{\mathrm{P}}(1), \quad \text{as} \quad n \to \infty, \tag{8}$$

where

$$W_n(t) = \frac{1}{\sqrt{n}} \sum_{j=1}^{n} W_{nj}(t) \tag{9}$$

with

$$W_{nj}(t) := W_n(t, \varepsilon_{j,n}) = \sin(t\varepsilon_{j,n}) - tC_\varepsilon(t) d_\beta \psi_\beta(\varepsilon_{j,n})). \tag{10}$$

for $j = 1, \ldots, n$, $t \in \mathbb{R}$ with $C_\varepsilon(.)$ denoting the real part of the characteristic function $\varphi_\varepsilon(.)$ Moreover, as $n \to \infty$, there is a zero–mean Gaussian process $\mathcal{W} = \{\mathcal{W}(t);\ t \in \mathbb{R}\}$ such that, as $n \longrightarrow \infty$,

$$W_n \xrightarrow{\mathcal{D}} \mathcal{W}, \quad T_{n,w} \xrightarrow{\mathcal{D}} \|\mathcal{W}\|^2. \tag{11}$$

(b) Nonparametric Regression case

For simplicity we assume a single regressor ($p = 1$), with values denoted by $X_j,\ j = 1, 2, ..., n$. Suppose that $(Y_1, X_1), \ldots, (Y_n, X_n)$, are i.i.d. random vectors such that

$$Y_j = m(X_j) + \sigma(X_j)\varepsilon_j, \quad j = 1, \ldots, n, \tag{12}$$

where $\varepsilon_1, \ldots, \varepsilon_n$, $X_1, \ldots, X_n$, $m(\cdot)$ and $\sigma(\cdot)$ satisfy:

(B.1) $\varepsilon_1, \ldots, \varepsilon_n$, are i.i.d. random variables with symmetric distribution F_ε, zero mean, unit variance and $\mathrm{E}(\varepsilon_j^4) < \infty$.

(B.2) $X_1, \ldots, X_n$, are i.i.d. on $[0, 1]$ with common positive continuous density f_X.

(B.3) $(\varepsilon_1, \ldots, \varepsilon_n)$ and $(X_1, \ldots, X_n)$, are independent.

(B.4) Let m be a function on $[0, 1]$ with Lipschitz first derivative.

(B.5) $\sigma(x)$, $x \in [0, 1]$ is positive on $[0, 1]$ with Lipschitz first derivative.

(B.6) The weight function w is nonnegative and symmetric, and

$$0 < \int_{-\infty}^{\infty} t^4 w(t)dt < \infty.$$

Our procedure depends on the estimators of $m(\cdot)$ and $\sigma(\cdot)$. The used residuals defined in (3) are based on the following estimators of the density function $f_X(\cdot)$ of X_j's, regression function $m(\cdot)$ and variance function $\sigma^2(\cdot)$:

$$\widehat{f}_X(x) = \frac{1}{nh_n} \sum_{j=1}^{n} K((x - X_j)/h_n), \quad x \in [0, 1], \tag{13}$$

$$\widehat{m}_n(x) = \frac{1}{nh_n \widehat{f}_X(x)} \sum_{j=1}^{n} K((x - X_j)/h_n) Y_j, \quad x \in [0, 1], \tag{14}$$

$$\widehat{\sigma}_n^2(x) = \frac{1}{nh_n \widehat{f}_X(x)} \sum_{j=1}^{n} K((x - X_j)/h_n)(Y_j - \widehat{m}_n(x))^2, \quad x \in [0, 1]. \tag{15}$$

It is assumed that the kernel $K(\cdot)$ and the bandwidth $h = h_n$ involved in the estimation of $m(\cdot)$ and $\sigma(\cdot)$ satisfy

(B.7) Let K be a symmetric twice continuously differentiable density on $[-1, 1]$ with $K(-1) = K(1) = 0$.

(B.8) Let $\{h_n\}$ be a sequence of the bandwidth such that $\lim_{n\longrightarrow\infty} nh_n^2 = \infty$ and $\lim_{n\longrightarrow\infty} nh_n^{3+\delta} = 0$ for some $\delta > 0$.

The limit behavior of the test statistic $T_{n,w}$ under the null hypothesis H_0 of symmetry in the present setup is quite similar the case of linear regression setup.

Theorem 0.2. *Let $Y_1, \ldots, Y_n$ follow the model (12) and let the assumptions (B.1) – (B.8) be satisfied. Then the assertion (8) remains true if $W_{nj}(t)$, $j = 1, \ldots, n$ are replaced by*

$$W^o_{nj}(t) := W_n(t, \varepsilon_j) = \sin(t\varepsilon_j) - \varepsilon_j t\phi_\varepsilon(t)). \tag{16}$$

for $j = 1, \ldots, n$, $t \in R$. Moreover, the assertions(11) for a zero–mean Gaussian process $\mathcal{W}^0 = \{\mathcal{W}^0(t);\ t \in \mathbb{R}\}$ with the covariance structure as the process $\{n^{-1/2} \sum_{j=1}^n W^o_{nj}(t);\ t \in \mathbb{R}\}$.

3. Comments and remarks

- Assumptions (A.1)–(A.4) and (B.1)–(B.8) are quite standard; for discussion see Hušková and Meintanis (2007), and Hušková and Meintanis (2009), respectively.

- The asymptotic null distribution is not distribution free. Therefore it does not provide an approximation for critical value. However, a properly chosen bootstrap method provide an approximation that leads to a consistent test.

- The tests in both setups are consistent for a large spectrum of alternatives. During the talk more detailed results will be presented.

- In the linear regression setup as the estimator of $\boldsymbol{\beta}$ can be used least squares estimators , more generally M-estimators.

- Notice that in the nonparametric setup the limit behavior of $T_{n,w}$ depends on the choice of neither the kernel $K(.)$ nor bandwidth h.

- The assumptions on the weight function w is quite mild. Quite often we choose $w(t) = e^{-a|t|}$, $t \in R$ or $w(t) = e^{-at^2}$, $t \in R$, with $a > 0$ being a tuning constant.Some particular choices leads to interesting versions of the test statistics. For instance, choosing $w(t) = \exp\{-t^2 a\}$, $t \in R^1$, $a > 0$ we get that

$$T_{n,a} = \frac{\pi}{2} \int_{-\infty}^{\infty} \left[f_n^{(+)}(x) - f_n^{(-)}(x) \right]^2 dx,$$

  where $f_n^{(+)}(x) = n^{-1} \sum_{j=1}^n (2\pi a)^{-1/2} e^{-(x-e_j)^2/2a}$ and $f_n^{(-)}(x) = f_n^{(+)}(-x)$ denote the densities of $\mathcal{P}_n * N(0, a)$ and $\mathcal{N}_n * N(0, a)$, respectively. Notice further that $f_n^{(+)}(\cdot)$ (resp. $f_n^{(-)}(\cdot)$) is the nonparametric density estimator of e_j (resp. $-e_j$) based on the Gaussian kernel $(2\pi)^{-1/2} e^{-x^2/2}$, with bandwidth equal to $\sqrt{a}$. Analogous interpretations arise for the test statistic with $w(t) = e^{-a|t|}$ and a nonparametric density estimator based on the Cauchy kernel; refer also to Henze et al. (2003).

References

[1] Bickel, P. (1982). On adaptive estimation. *Ann. Statist.* 10: 647 – 671.

[2] Dette, H., Kusi–Appiah, S. and Neumeyer, N. (2002). Testing symmetry in nonparametric regression models, *J. Nonparam. Statist.* 14: 477 – 4

[3] Fan, Y. and Gencay, R. (1995). A consistent nonparametric test of symmetry in linear regression models, *J. Amer. Statist. Assoc.* 90: 551–557.

[4] Feuerverger, A. and Mureika, R. (1977). The empirical characteristic function and its application, *Ann. Statist.* 5: 88 – 97.

[5] Ghosh, S. and Ruymgaart, F. (1992). Applications of empirical characteristic functions in some multivariate problems, *Canad. J. Statist.* 20: 429 – 440.

[6] Henze, N., Klar, B. and Meintanis, S.G. (2003). Invariant tests for symmetry about an unspecified point based on the empirical characteristic function, *J. Multivar. Anal.* 87: 275 – 297.

[7] Hettmansperger, T.P., McKean, J.W. and Seather, S.J. (2002). Finite sample performance of tests for symmetry of the errors in a linear model, *J. Statist. Comput. Simul.* 72: 863 – 879.

[8] Hušková, M. and Meintanis, S.G. (2007). Omnibus test for the error distribution in the linear regression model, *Statistics* 41: 363 – 376.

[9] Kappenman, R.F. (1988). Robust symmetric distribution location estimation and regression. *J. Statist. Plann. Infer.* 19: 55 – 72.

[10] Koutrouvelis, I.A. (1985). Distribution–free procedures for location and symmetry inference problems based on the empirical characteristic function. *Scand. J. Statist.* 12: 257 – 269.

[11] Neuhaus, G. and Zhu, L.-X. (1998). Permutation tests for reflected symmetry, *J. Multivar. Anal.* 67: 129 – 153.

[12] Neumeyer, N. and Dette, H. (2007). Testing for symmetric error distribution in nonparametric regression models, *Statistica Sinica* 17: 775 – 795.

[13] Neumeyer, N., Dette, H. and Nagel, E.–R. (2005). A note on testing symmetry of the error distribution in linear regression models, *J. Nonparam. Statist.* 17: 697 – 715.

[14] Newey, W.K. (1988). Adaptive estimation of regression models via moment restrictions. *J. Econometr.* 38: 301 – 339.

[15] Zayed, A.I. (1996) *Handbook of Function and Generalized Function Transformations.* CRC Press, New York

6th St.Petersburg Workshop on Simulation (2009) 737-741

L_p-type goodness-of-fit tests and their asymptotic comparison[1]

Norbert Henze[2], Yakov Nikitin[3], Bruno Ebner[4]

Abstract

We compare different integral statistics based on L_p-norms with respect to local approximate Bahadur efficiency. Simulation results corroborate the theoretical findings. Several examples illustrate that goodness-of-fit testing based on L_p-norms should receive more attention. We show how to determine the value of p giving the maximum efficiency.

1. Introduction

Let $X_1, \ldots, X_n$ be independent random variables with common continuous distribution function (df) F. A classical problem of statistics is testing the hypothesis $H_0 : F = F_0$, where F_0 is some known continuous df, against general alternatives.

Many celebrated distribution-free statistics for this goodness-of-fit testing problem are functionals of the empirical process $\xi_n(x) = \sqrt{n}\,(F_n(x) - F_0(x))\,,\ x \in \mathbb{R}$, where $F_n(x) = \frac{1}{n}\sum_{j=1}^{n} \mathbf{1}\{X_j \le x\}$ is the empirical df based on $X_1, \ldots, X_n$.

Prominent examples are the Kolmogorov-Smirnov statistic $D_n = \sup_x |\xi_n(x)|$ and the Cramér-von Mises statistic $\omega_n^2 = \int_{-\infty}^{\infty} \xi_n^2(x)\ dF_0(x)$. There is, however, a continuing interest in more general integral statistics based on the L_p-norm

$$\omega_{n,p} = \left(\int_{-\infty}^{\infty} |\xi_n(x)|^p\ dF_0(x) \right)^{1/p} = \sqrt{n} \left(\int_{-\infty}^{\infty} |F_n(x) - F_0(x)|^p\ dF_0(x) \right)^{1/p},$$

where $1 \le p < \infty$. For instance, the statistic $\omega_{n,1}$ was considered in [9]. Some properties of $\omega_{n,p}$-statistics were studied in [5]. In [1, 3, 4] weighted L_p-statistics were proposed. One-sided statistics $W_{n,k}$ for natural k, namely

$$W_{n,k} = \int_{-\infty}^{\infty} \xi_n^k(x) dF_0(x),$$

were studied in [8] and [6]. There is some evidence that the new statistics will have better properties than the classical ones, see, e.g., [10] and [6]. A formula for $\omega_{n,p}$

[1]This work was supported by grants RFBR 07-01-00159-a and NSh.638.2008.1.

[2]Karlsruhe University, E-mail: N.Henze@math.uni-karlsruhe.de

[3]St.Petersburg University, E-mail: yanikit47@gmail.com

[4]Karlsruhe University, E-mail: ebner@stoch.uni-karlsruhe.de

in terms of order statistics and null-distribution F_0 can be easily derived, hence, a computer routine for implementing the corresponding test is readily available.

Our aim is the comparison of the statistics $\omega_{n,p}$ for different values of p on the basis of local approximate Bahadur efficiency (ABE), see [2]. As these statistics are not asymptotically normal, Pitman efficiency is not applicable. The theoretical results are corroborated in a Monte Carlo Study.

2. ABE of L_p-statistics

Let $\{P_\theta : \theta \in \Theta\}$ be a set of probability measures on a measurable space $(S, \mathcal{S})$. Let $\Theta_0 \subset \Theta$, $\Theta_1 = \Theta \setminus \Theta_0$ and let $H_0 : \theta \in \Theta_0$. Assume $\{T_n\}$ is a sequence of test statistics defined on $(S, \mathcal{S})$, based on a sample of size n. $\{T_n\}$ is called a *standard sequence* if the following conditions hold:

a) There exists a continuous probability df G such that, for each $\theta \in \Theta_0$, $\lim_{n\to\infty} P_\theta(T_n \le x) = G(x)$ for every x.

b) $\ln(1 - G(x)) = -\frac{1}{2}ax^2 \cdot (1 + o(1))$ as $x \to \infty$ for some a, $0 < a < \infty$.

c) There exists a real-valued function $b(\theta)$ on Θ_1, with $0 < b(\theta) < \infty$, such that, for each $\theta \in \Theta_1$, $\frac{T_n}{\sqrt{n}} \to b(\theta)$ in $P_\theta -$ probability.

For any standard sequence of statistics $\{T_n\}$ define the *approximate Bahadur slope* $c_T^*(\theta)$ by $c_T^*(\theta) = a \cdot b^2(\theta)$, $\theta \in \Theta_1$. It is a measure of ABE.

The statistics $\{\omega_{n,p}\}$, $1 \le p \le \infty$, are standard statistics with limiting distribution $G_p(x) = P(\|B\|_p \le x)$, $x \in \mathbb{R}$, where B is the standard Brownian bridge on $[0, 1]$. We have (see [11]): $\lim_{x\to\infty} 2x^{-2} \cdot \ln(1 - G_p(x)) = -C(p)$, where $C(\infty) = 4$ and

$$C(p) = \frac{2p\pi}{(1 + \frac{p}{2})^{(p-2)/p}} \cdot \left(\frac{\Gamma(1 + 1/p)}{\Gamma(\frac{1}{2} + 1/p)}\right)^2, \quad 1 \le p < \infty.$$

Consider the Kullback-Leibler distance $K(\theta)$ between F_θ and F_0. It is usually true that $K(\theta) \sim \frac{1}{2} \cdot I_0 \cdot \theta^2$ as $\theta \to 0$, see [8], where I_0 is the Fisher information at the point $\theta = 0$. Take the ratio of the approximate slopes for our test and for the locally most powerful parametric test, and define the *local* ABE (e^B) as the limit of this ratio as $\theta \to 0$. Hence the expressions for the local ABE of the statistics $\omega_{n,p}$ are equal to

$$e^B(\omega_p; f_0) = \frac{C(p)}{I_0} \cdot \left\{\int_{-\infty}^{\infty} \left|\frac{d}{d\theta} F_\theta(x)\big|_{\theta=0}\right|^p dF_0(x)\right\}^{2/p}.$$

The behavior of the local ABE can be very different depending on the null distribution and the alternative. We have studied this behavior in the case of three different types of alternatives: shift, scale and skew.

3. Shift, scale and skew alternatives

For shift alternatives, the formula for ABE can be simplified, and $e^B(\omega_p; f_0, shift) = \frac{C(p)}{I_0} \cdot \left\{ \int_{-\infty}^{\infty} f_0^{p+1}(x)\, dx \right\}^{2/p}$, where f_0 is the density corresponding to F_0. Thus, the local ABE becomes a function of p showing which values of p yield the maximum efficiency and presumably the maximum power of the test.

Example 3.1: (Normal law). If $f_0 = \phi$ is the standard normal density, some algebra yields $e^B(\omega_p; \phi, shift) = \frac{C(p)}{2\pi(p+1)^{1/p}}$ which, as a function of p, is monotonically decreasing (see Fig. 1) with the limit $\lim_{p\to\infty} e^B(\omega_p; \phi, shift) = 2/\pi \approx 0.637$.

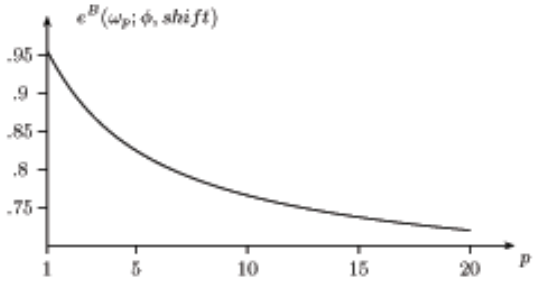


Figure 1: Local ABE (Perks' family: shift alternative

Example 3.2: (Perks' family of distributions). Consider the family of densities $p^{(s)}(x) = \frac{1}{B(1/2,1/s)} \cdot \cosh^{-2/s}(x)$, $x \in \mathbb{R}$, $s \geq 1$. This class of densities includes the logistic and the hyperbolic cosine distribution, which arise for $s = 1$ and $s = 2$, respectively. Here $e^B(\omega_p; p^{(s)}, shift) = C(p) \cdot \frac{s(s+2)\left(B\left(\frac{p+1}{s}, \frac{p+1}{s}\right)\right)^{2/p}}{(B(1/s,1/s))^{(2p+2)/p}}$. Fig. 2 shows the values of efiiciency as a function of p for $s = 3$ (left) and $s = 4$ (right). The maximum is 1 and is attained for $p = 3$ and $p = 4$, respectively.

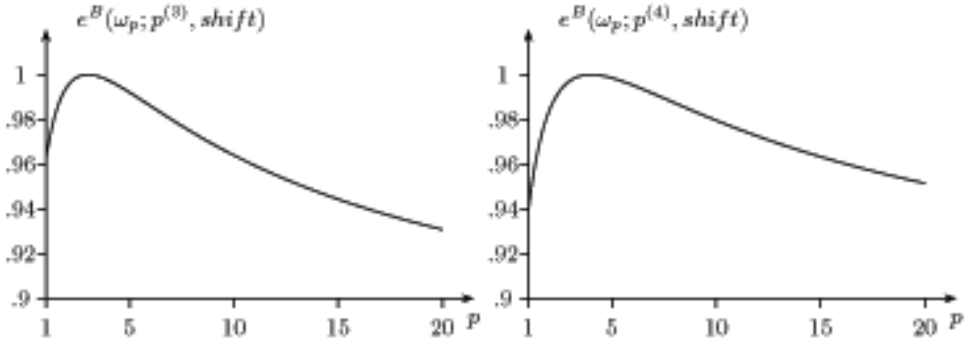


Figure 2: Local ABE (Normal distribution: shift alternative

Example 3.3: (Cauchy distribution). For the Cauchy distribution, we obtain $e^B(\omega_p; Cauchy, shift) = \frac{2C(p)}{\pi^{\frac{2p+2}{p}}} \left(\frac{\sqrt{\pi}\Gamma(p+1/2)}{\Gamma(p+1)}\right)^{2/p}$, see the plot in Fig. 3. The efficiency increases for small p and attains its maximum 0.876 at the unexpected point $p \approx 10.2$. After this maximum, the efficiency slowly decreases with p to the limiting value $8/\pi^2 \approx 0.814$ which is known as the ABE of the Kolmogorov test.

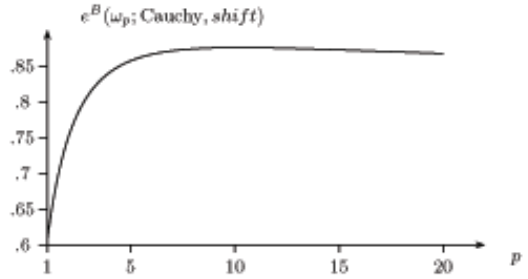


Figure 3: Local ABE (Cauchy distribution: shift alternative

We considered also scale alternatives of the form $F_\theta(x) = F_0(x\exp(-\theta))$, and skew alternatives with the alternative density $f_\theta(x) = 2\,f_0(x)\,F_0(\theta x)$, $x \in \mathbb{R}$, where f_0 is symmetric. The formulas for local ABE and corresponding plots are similar to the case of shift alternatives.

4. Independence testing

Similar reasoning can be applied in case of independence testing. Suppose we observe a sample of i.i.d. vectors $(X_1, Y_1), ..., (X_n, Y_n)$ with continuous joint df $F(x, y)$ and marginal df's $G(x)$ and $H(y)$. The problem is to test the independence hypothesis $H : F(x, y) = G(x)H(y)$ for all x, y. Denote by F_n, G_n and H_n corresponding empirical df's. Consider the statistics of L_p-type for $1 \le p < \infty$:

$$\Omega_{n,p} = \sqrt{n}\left\{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |F_n(x,y) - G_n(x)H_n(y)|^p\, dG_n(x)\, dH_n(y)\right\}^{1/p},$$

which are the generalizations of the celebrated Blum-Kiefer-Rosenblatt statistic.

The limiting null distribution of $\Omega_{n,p}$ is the distribution of $\|\mathbb{B}\|_p$ where $\mathbb{B}(x, y)$ is the Brownian pillow. The evaluation of the tail for this random variable follows from [7], namely $\lim_{x\to\infty} \frac{2}{x^2} \cdot \ln P(\|\mathbb{B}\|_p > x) = -C^2(p)$.

We assume that under the alternative $F(x, y) = G(x)H(y) + \theta \cdot L(G(x), H(y))$, where the so-called dependence function L is defined on the unit square, is smooth and is zero on the boundary. Such a model is well known (see [8], Ch. 5). The value of ABE for our statistics is

$$e^B(\Omega_p; L) = C^2(p) \cdot \left\{\int_0^1\int_0^1 |L(s,t)|^p\, ds\, dt\right\}^{2/p} / \int_0^1\int_0^1 \left(\frac{\partial^2 L}{\partial s \partial t}(s,t)\right)^2 dsdt.$$

For illustration we consider the special case of the dependence function $L_1(x,y) = \sin \pi x \cdot \sin \pi y$ which, by [8], Ch.6, should be locally optimal for the $\Omega_{n,2}$-test. We have $e^B(\Omega_p; L_1) = \frac{64\, C^2(p)}{\pi^{4(1+1/p)}} \cdot B\left(\frac{p+1}{2}, \frac{p+1}{2}\right)^{4/p}$. The plot is given in Fig. 4, and the maximum efficiency 1 is attained for $p = 2$, indeed.

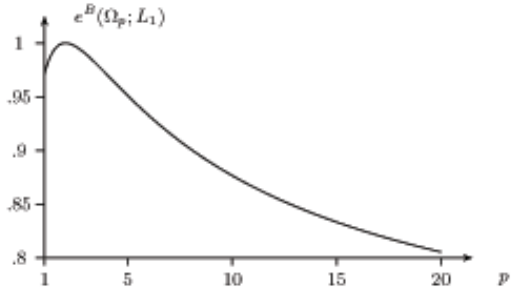


Figure 4: Local ABE (testing for independence

5. Simulations and conclusions

A Monte Carlo study corroborates the theoretical findings of previous sections. The simulations involved random samples of size $n = 100$, while critical values of statistics $\omega_{n,p}$ for each $p \in \{1, 1.5, \ldots, 20\}$ were based on 100000 replications. The nominal level of significance was 0.01. Our figures show relative frequencies of rejections of the null hypothesis based on L_p-statistics for each $p \in \{1, 1.5, \ldots, 20\}$. Among all plots we show only the plots for the Perks' Family in case of shift alternative with $\theta = 0.75$ which are rather typical.

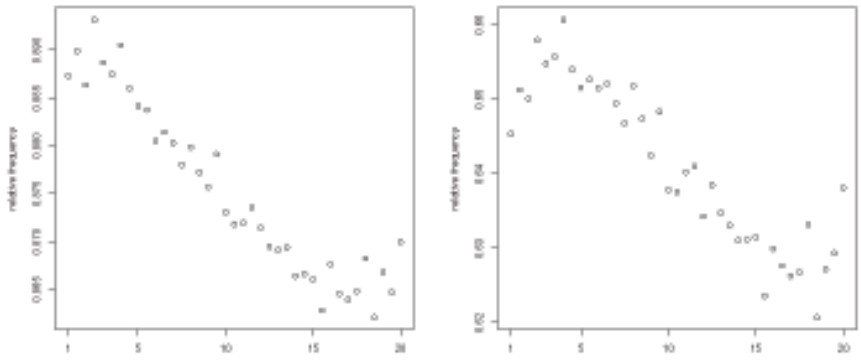


Figure 5: Percentages of rejection out of 100000 samples for shift alternatives from Perks' family of distributions, where $s = 3$ (left) and $s = 4$ (right)

In the figure 5 the left hand plot corresponds to $s = 3$ and the right hand plot

to $s = 4$, cf. Figure 2. The plots roughly resemble the curves of local ABE's.

We see that more flexibility with regard to power against specific alternatives may be gained by considering goodness-of-fit tests based on L_p-norms other than $p = 2$ and $p = \infty$. Of course, no universally best value of p may be expected even under a special setting of alternatives, like the shift model.

However, given a distribution function F_0 and a specific alternative, we can draw the plot of efficiency as a function of p and determine the value of p giving the maximum efficiency.

References

[1] Ahmad, I. (1997) Goodness-of-fit statistics based on weighted L_p-functionals. *Stat. & Prob. Lett.* **35**, 261–268.

[2] Bahadur, R.R. (1960) Stochastic comparison of tests. *Ann. Math. Stat.* **31**, 231–260.

[3] Berkes, I., Horvath, L., Shao, Q.-M., and Steinebach, J. (2000). Strong laws for L_p-norms of empirical and related processes. *Period. Math. Hung.* **41**, 35–69.

[4] Csörgő, M., and Horvath, L. (1993) *Weighted approximations in Probability and Statistics.* Wiley, New York.

[5] Fatalov, V.R. (1999) The Laplace method for computing exact asymptotics of distributions of integral statistics. *Math. Meth. of Stat.* **8**, 510–535.

[6] Ivanov, V., and Zrelov, P. (1997) Nonparametric Integral Statistics $\omega_n^k = n^{k/2} \int_{-\infty}^{\infty} [S_n(x) - F(x)]^k dF(x)$: Main Properties and Applications. *Comp. Math. Appl.* , **34**, 703–726.

[7] Nazarov, A.I., and Nikitin, Ya. (2000) Some extremal problems for Gaussian and empirical random fields (Russian). *Proceedings of St. Petersburg Math. Society,* **8**, p.214–230 .

[8] Nikitin, Ya. (1995) *Asymptotic efficiency of nonparametric tests.* Cambridge University Press.

[9] Schmid, F., and Trede, M. (1995) A distribution-free test for the two sample problem for general alternatives. *Comput. Stat. Data Anal.,* **20**, 409–419.

[10] Schmid, F., and Trede, M. (1996) An L_1-variant of the Cramér-von Mises test. *Stat. & Prob. Lett.* **26**, 91–96.

[11] Strassen, V. (1964) An invariance principle for the law of the iterated logarithm. *Zeitschr. Wahrsch. Verw. Geb.* **3**, 211–226.

6th St.Petersburg Workshop on Simulation (2009) 743-747

Inference Procedures for Stable–Paretian Stochastic Volatility Models

Simos G. Meintanis[1], Emanuele Taufer[2]

Abstract

A discrete stochastic volatility model is considered driven by a couple of stable–Paretian processes, one driving processs for the observations and the other for the scale parameter. Due to convolution properties of stable–Paretian laws, the unconditional distribution of the observations is also stable–Paretian, and therefore its characteristic function is expressed as a simple exponential–type function incorporating the parameters. Exploiting this feature of the stochastic volatility model considered, methods of estimation and testing goodness–of–fit are proposed employing the empirical characteristic function. The proposed procedures are applied with simulated data but also with some real data from the financial markets.

Keywords. Stable–Paretian distribution; Characteristic function; Estimation; Goodness-of-fit.

AMS 2000 classification numbers: 62G10, 62G20.

1. Introduction

Consider observations $y_t,\ (t = 1, ..., T)$, from the model

$$y_t = \delta + c_t^{1/\alpha_1} x_t,\ \delta \in \mathbb{R},\ 0 < \alpha_1 \leq 2, \tag{1}$$

with

$$c_t = \gamma c_{t-1} + \lambda z_{t-1},\ 0 < \gamma < 1,\ \lambda > 0, \tag{2}$$

where x_t follows a symmetric stable–Paretian distribution with characteristic function (CF) $\mathbf{E}(e^{iux_t}) = e^{-|u|^{\alpha_1}},\ u \in \mathbb{R}$, and z_t follows a positive stable–Paretian distribution with Laplace transform $\mathbf{E}(e^{-uz_t}) = e^{-u^{\alpha_2}},\ u > 0,\ 0 < \alpha_2 < 1$. Such model was first introduced by de Vries (1991) in an attempt to accomodate several stylized facts of returns of financial assets such as marginal distribution with

[1]Department of Economics, National and Kapodistrian University of Athens, Athens, Greece, E-mail: simosmei@econ.uoa.gr

[2]Department of Computer and Management Sciences, University of Trento, Trento, Italy, E-mail: emanuele.taufer@unitn.it

fat tails, volatility clusters, invariance property under additivity and existence of limiting laws for normalized sums. For further details see de Vries (1991).

From Theorem 1 in de Vries (1991), and as detailed in the next section, the CF of the unconditional distribution of $y_t,\ (t = 1, ..., T)$, for $\vartheta = (\delta, \gamma, \lambda, \alpha_1, \alpha_2)'$, is given by

$$\phi(u) = \phi(u; \vartheta) := e^{i\delta u - \sigma|u|^{\alpha}}, \tag{3}$$

where

$$\sigma = \frac{\lambda^{\alpha_2}}{1 - \gamma^{\alpha_2}},\ \alpha = \alpha_1\alpha_2,$$

In this paper procedures for estimation and testing are proposed for the model (1)–(2), based on the empirical CF

$$\hat{\phi}_T(u) = \frac{1}{T}\sum_{t=1}^{T} e^{iuy_t},\ i = \sqrt{-1},$$

of the observations $y_t,\ (t = 1, ..., T)$. In particular we consider the estimator $\hat{\vartheta}_T$ of ϑ minimizing the distance measure

$$\Delta_T(\vartheta) = \int_{-\infty}^{\infty} |\hat{\phi}_T(u) - \phi(u; \vartheta)|^2\, w(u)du, \tag{4}$$

where $w(u)$ denotes a suitable weight function. In addition, a goodness–of–fit statistic for the model (1)–(2) could be based on $\Delta_T(\hat{\vartheta}_T)$ where $\hat{\vartheta}_T$ denotes any consistent estimator of ϑ, and not necessarily the one minimizing (4). From (4) we have by simple algebra,

$$\Delta_T(\vartheta) = \frac{1}{T^2}\sum_{s,t=1}^{T} I(0, y_s - y_t) + I(2\sigma, 0) - \frac{2}{T}\sum_{t=1}^{T} I(\sigma, y_t - \delta), \tag{5}$$

where

$$I(\nu, z) = \int_{-\infty}^{\infty} \cos(uz)e^{-\nu|u|^{\alpha}} w(u)du,\ \nu > 0,\ z \in \mathbb{R}.$$

With this direct approach however it does not seem possible to avoid the need of special numerical techniques in calculating (5) because the computation of the integral $I(\nu, z)$ requires numerical quadrature. Alternatively, and since the CF in (3) satisfies the differential equation

$$f'(u) = i\delta f(u) - \alpha\sigma u^{\alpha-1} f(u),\ u > 0,$$

we could use the distance measure

$$\widetilde{\Delta}_T(\vartheta) = \int_{0}^{\infty} |D_T(\vartheta)|^2\, w(u)du, \tag{6}$$

where $D_T(\vartheta) = \hat{\phi}'_T(u) - i\delta\hat{\phi}_T(u) + \alpha\sigma u^{\alpha-1}\hat{\phi}_T(u)$. Then $\widetilde{\Delta}_T(\vartheta)$ may be written in closed form. In fact by straightforward computation we have

$$|D_T(\vartheta)|^2 = [C'_T(u)]^2 + [S'_T(u)]^2 + \left[\delta^2 + \alpha^2\sigma^2 u^{2(\alpha-1)}\right] |\hat{\phi}_T(u)|^2 \tag{7}$$

$$+2\delta\left[S_T(u)C_T^{'}(u)-S_T^{'}(u)C_T(u)\right]+2\alpha\sigma\left[C_T^{'}(u)C_T(u)+S_T^{'}(u)S_T(u)\right]u^{\alpha-1}$$

$$=\frac{1}{T^2}\sum_{s,t=1}^{T}y_sy_t\cos[(y_s-y_t)u]+\left[\delta^2+\alpha^2\sigma^2u^{2(\alpha-1)}\right]\frac{1}{T^2}\sum_{s,t=1}^{T}\cos[(y_s-y_t)u]$$

$$-2\delta\frac{1}{T^2}\sum_{s,t=1}^{T}y_t\cos[(y_s-y_t)u]-\alpha\sigma\frac{1}{T^2}\sum_{s,t=1}^{T}(y_s-y_t)\sin[(y_s-y_t)u],$$

where $C_T(u):=\Re[\hat{\phi}_T(u)]$ and $S_T(u):=\Im[\hat{\phi}_T(u)]$ and $|\hat{\phi}_T(u)|^2=C_T^2(u)+S_T^2(u)$. Letting $w(u)=e^{-au},\ a>0$, we obtain from (6) and (7),

$$\widetilde{\Delta}_T(\vartheta)=\frac{1}{T^2}\sum_{s,t=1}^{T}\Big\{\left[y_sy_t+\delta^2-2\delta y_t\right]\mathrm{K}(0,y_{st}) \tag{8}$$

$$+\alpha^2\sigma^2\mathrm{K}(2(\alpha-1),y_{st})-\alpha\sigma y_{st}\Lambda(\alpha-1,y_{st})\Big\},$$

where

$$\mathrm{K}(\nu,z)=$$
$$\int_0^\infty\cos(uz)u^\nu e^{-au}du=\frac{1}{a^{\nu+1}}(1+\frac{z^2}{a^2})^{-(1/2)(\nu+1)}\cos[(\nu+1)\tan^{-1}\left(\frac{z}{a}\right)]\Gamma(\nu+1),$$

$$\Lambda(\nu,z)=$$
$$\int_0^\infty\sin(uz)u^\nu e^{-au}du=\frac{1}{a^{\nu+1}}(1+\frac{z^2}{a^2})^{-(1/2)(\nu+1)}\sin[(\nu+1)\tan^{-1}\left(\frac{z}{a}\right)]\Gamma(\nu+1)$$

and $y_{st}=y_s-y_t,\ s,t=1,...,T$.

Remark 1. *Note that equation* (6) *actually provides a consistent estimating equation only if the first moment of the unconditional distribution of y_t, exists, i.e. if* $1\leq\alpha_1\leq 2$.

For testing purposes equations (5) and (8) can be further reduced. Specifically suppose that ϑ has been consistently estimated by $\hat{\vartheta}_T:=(\hat{\delta}_T,\hat{\sigma}_T,\hat{\lambda}_T,\hat{\alpha}_{1T},\hat{\alpha}_{2T})'$, and define the standardized observations

$$\hat{y}_t=\frac{y_t-\hat{\delta}_T}{\hat{\sigma}_T^{1/\hat{\alpha}_T}},\ t=1,...,T,$$

where

$$\hat{\sigma}_T=\frac{\hat{\lambda}_T^{\hat{\alpha}_{2T}}}{1-\hat{\gamma}_T^{\hat{\alpha}_{2T}}},\ \hat{\alpha}_T=\hat{\alpha}_{1T}\hat{\alpha}_{2T}.$$

Then the test statistic corresponding to $\Delta_T(\vartheta)$ is given by (5) with y_t replaced by $\hat{y}_t$, and $(\delta,\sigma)=(0,1)$. Likewise the test statistic corresponding to $\widetilde{\Delta}_T(\vartheta)$ is given by (8) with y_t replaced by $\hat{y}_t$, and with $(\delta,\sigma,\alpha)=(0,1,\hat{\alpha}_T)$.

Remark 2. *Note that the univariate CF depends on α_1, α_2 only through $\alpha = \alpha_1\alpha_2$ and on λ, γ through $\sigma = \frac{\lambda^{\alpha_2}}{1-\gamma^{\alpha_2}}$. Hence there is a problem of identifiability in using the univariate CF for estimation, i.e. only δ, σ and α are identifiable, but not α_1, α_2, λ, and γ. Very often this may not be a serious problem since one is interested in estimating the scale σ which can be obtained from the univariate CF and, in the spirit of Engle's original ARCH process one is lead to assume that $\alpha_1 = 2$ so that α_2 can be readily recovered.*

2. Multivariate Characteristic function

The problems of identifiability encountered with the univariate ch.f. can be solved by resorting to higher order CF. Here we produce the joint CF of $Y_{t,k} = (y_t, y_{t-1}, ..., y_{t-(k-1)})$, say $\varphi_{Y_{t,k}}(U_k)$, at $U_k = (u_0, u_1, ..., u_{k-1})$. In particular we have,

$$\begin{aligned}
\varphi_{Y_{t,k}}(U_k) = \mathrm{E}[e^{iU_k{}'Y_{t,k}}] &= \mathrm{E}[\exp\{i\sum_{j=0}^{k-1} iu_j y_{t-j}\}] \\
&= \mathrm{E}\left[\mathrm{E}[\exp\{i\sum_{j=0}^{k-1} iu_j y_{t-j}\}]\big|C_{t,k}\right] \\
&= \mathrm{E}\left[\mathrm{E}[\exp\{i\sum_{j=0}^{k-1} u_j(\delta + c_{t-j}^{1/\alpha_1} x_{t-j})\}]\right] \\
&= \exp\left\{i\delta\sum_{j=0}^{k-1} u_j\right\}\mathrm{E}\left[\prod_{j=0}^{k-1}\phi_{x_{t-j}}(u_j c_{t-j}^{1/\alpha_1})\}\right] \\
&= \exp\left\{i\delta\sum_{j=0}^{k-1} u_j\right\}\mathrm{E}\left[\exp\left\{-\sum_{j=0}^{k-1} c_{t-j}|u_j|^{\alpha_1}\right\}\right],
\end{aligned}$$

since x_{t-j} has CF $e^{-|u|^{\alpha_1}}$ and where $C_{t,k} = \big(c_t, c_{t-1}, ..., c_{t-(k-1)}\big)$, and $\phi_X(\cdot)$ denotes the CF of X. However repeated application of (2) yields

$$c_{t-j} = \gamma^{k-(j+1)} c_{t-(k-1)} + \lambda\sum_{m=j}^{k-2}\gamma^{m-j} z_{t-(m+1)},\ j = 0, 1, ..., k-2,$$

and due to the independence of $c_{t-(k-1)}$ and $z_{t-(m+1)}$, $m = j, j+1, ..., k-2$, $j = 0, 1, ..., k-2$, one obtains

$$\varphi_{Y_{t,k}}(U_k) = \exp\left\{i\delta\sum_{j=0}^{k-1} u_j\right\}\mathrm{E}\left[\exp\left\{-\sum_{j=0}^{k-1}\gamma^{k-(j+1)} c_{t-(k-1)}|u_j|^{\alpha_1}\right\}\right]$$

$$\times \mathrm{E}\left[\exp\left\{-\lambda\sum_{j=0}^{k-2}\sum_{m=j}^{k-2}\gamma^{m-j}z_{t-(m+1)}|u_j|^{\alpha_1}\right\}\right]$$

$$= \exp\left\{i\delta\sum_{j=0}^{k-1}u_j\right\}\mathrm{E}\left[\exp\left\{-\sum_{j=0}^{k-1}\gamma^{k-(j+1)}|u_j|^{\alpha_1}c_{t-(k-1)}\right\}\right]$$

$$\times \mathrm{E}\left[\exp\left\{-\lambda\sum_{m=0}^{k-2}\sum_{j=0}^{m}\gamma^{m-j}|u_j|^{\alpha_1}z_{t-(m+1)}\right\}\right]$$

$$= \exp\left\{i\delta\sum_{j=0}^{k-1}u_j\right\}\mathrm{E}\left[\exp\left\{-u_{k-1}c_{t-(k-1)}\right\}\right]\mathrm{E}\left[\exp\left\{-\lambda\sum_{m=0}^{k-2}u_m z_{t-(m+1)}\right\}\right]$$

$$= \exp\left\{i\delta\sum_{j=0}^{k-1}u_j\right\}L_{c_{t-(k-1)}}(u_{k-1})\prod_{m=0}^{k-2}L_{z_{t-(m+1)}}(\lambda u_m),$$

due to the independence of $Z_{t-(m+1)},\ m=0,1,...,k-2$, and where $L_X(\cdot)$ denotes the LT of X, and $v_m=\sum_{j=0}^{m}\gamma^{m-j}|u_j|^{\alpha_1},\ m=0,1,...,k-1$. However under the standing assumptions c_t has LT $e^{-\sigma u_2^{\alpha}}$ while z_t are i.i.d. with LT $e^{-u^{\alpha_2}}$. Hence we finally obtain the joint CF of $(y_t, y_{t-1}, ..., y_{t-(k-1)})$ as

$$\varphi_{Y_{t,k}}(U_k)=\exp\left\{i\delta\sum_{j=0}^{k-1}u_j\right\}\exp\left\{-\frac{\lambda^{\alpha_2}}{1-\gamma^{\alpha_2}}v_{k-1}^{\alpha_2}-\lambda^{\alpha_2}\sum_{j=0}^{k-2}v_j^{\alpha_2}\right\}. \tag{9}$$

Exploiting the above general formula one can produce estimating equations either by generalizing the argument of the previous section or by using standard CF estimation procedures, as nicely reviewed in Knight et al. (2002) and Yu (2004).

Considering the bivariate case if one wishes to apply the technique discussed in the previous section note that the mixed derivative of the above CF has the form

$$\frac{\partial P(u_1,u_2)}{\partial u_1\partial u_2}=P(u_1,u_2)f(u_1,u_2;\alpha_1,\alpha_2,\gamma,\lambda) \tag{10}$$

for some nonlinear function f of the parameters and u_1, u_2. Turning to the data problem, if we use empirical estimators of the bivariate CF, i.e.

$$\hat{\phi}_T(u_1,u_2)=\frac{1}{T-1}\sum_{j=2}^{T}\exp\{iu_1y_j+iu_2y_{j-1}\},\quad i=\sqrt{-1}, \tag{11}$$

it turns out that, after appropriate transformations, the bivariate analougue of the contrast (6) contains expressions of the form

$$\cos[u_1(y_j-y_k)]\cos[u_2(y_{1+j}-y_{1+k})],\quad \text{or}\quad \sin[u_1(y_j-y_k)]\sin[u_2(y_{1+j}-y_{1+k})]. \tag{12}$$

So, one can solve integration by iterated integrals and use again the expressions for $K(\nu,z)$ provided above. These approaches however turn out to be quite cumbersome to implement moreover Remark 1 applies to this case too.

We finally provide some evidence of the small sample properties of these procedures by simulations as well as an application to a real data problem. Result shows that the estimation procedures work well and the model adequately fits financial data.

References

[1] de Vries, Casper G. (1991). On the relation between GARCH and stable processes. *J. Econometrics* 48, 313–324.

[2] Knight, J.L., Satchell, S.E., Yu, Y. (2002). Efficient estimation of the stochastic volatility model by the empirical characteristic function method. *Austral. New Zealand J. Statist.* 44, 319-335.

[3] Yu, Y. (2004). Empirical Characteristic Function Estimation and Its Applications. *Econometric Reviews* 23, 93-123.

6th St.Petersburg Workshop on Simulation (2009) 749-753

A new method of evaluating the performance of bootstrap-based tests

James S. Allison[1], Jan W.H. Swanepoel[2]

Abstract

We discuss two methods to evaluate the performance of bootstrap-based tests: the first is one that is traditionally used in the literature, while the second is an alternative, more robust, method that we propose. We also present some theoretical properties regarding the bootstrap estimator of the critical value when testing for the mean in a univariate population. This will be based on the new method for evaluating the performance of a bootstrap-based test.

1. Introduction

When a bootstrap-based test is proposed, one would like to evaluate the performance of the test in order to assess how "good" the proposed test is. This evaluation can be done theoretically and/or by means of a Monte-Carlo simulation.

In Section 3 the evaluation method that is currently in use in the literature is discussed. We will refer to this evaluation method as *Method I*. In Section 4 we propose a new method of evaluating the performance of a bootstrap-based test. We will refer to this new evaluation method as *Method II*. Section 2 introduces the basic notation that will be used in discussing these two methods. The paper concludes with the application of this new evaluation method to a simple example.

2. Notation

Assume that observations $X_1, X_2, \ldots, X_n$ are available from some model with joint distribution function $F_{\boldsymbol{\theta},\boldsymbol{\nu}}(x_1, \ldots, x_n)$, depending on some unknown parameters $\boldsymbol{\theta}$ and $\boldsymbol{\nu}$. Let $\mathbf{X}_n = (X_1, X_2, \ldots, X_n)$ and denote by $\mathbf{x}_n = (x_1, x_2, \ldots, x_n)$ an observed realization of $\mathbf{X}_n$.

Consider the hypothesis

$$H_0 : \boldsymbol{\theta} \in \Theta_0 \qquad \text{vs.} \qquad H_A : \boldsymbol{\theta} \in \Theta_A,$$

[1]North-West University,South Africa E-mail: james.allison@nwu.ac.za
[2]North-West University, South Africa, E-mail: jan.swanepoel@nwu.ac.za

where Θ_0 and Θ_A are two disjoint subsets of some parameter space $\Theta = \Theta_0 \cup \Theta_A$.

Assume, without loss of generality, that the test procedure is of the form:

Reject H_0 if and only if

$$T_n(\mathbf{X}_n) \geq C_n(\alpha; \mathbf{X}_n),$$

where $T_n(\mathbf{X}_n)$ is an appropriate test statistic, $C_n(\alpha; \mathbf{X}_n)$ is a bootstrap critical value and α is the significance level of the test.

For testing H_0 we consider $\boldsymbol{\nu}$ as a nuisance parameter, which will be replaced by an estimator that is strongly consistent under both H_0 and H_A. These estimated nuisance parameters are therefore not indicated in the notation of $T_n(\mathbf{X}_n)$ and $C_n(\alpha; \mathbf{X}_n)$.

Furthermore, let $\boldsymbol{\theta}_T$ denote the true value of the parameter. If $\boldsymbol{\theta}_T \in \Theta_0$ it is denoted by $\boldsymbol{\theta}^0$, otherwise we write it as $\boldsymbol{\theta}^A$.

Denote the bootstrap random variables by $\mathbf{X}_n^* = (X_1^*, X_2^*, \ldots, X_n^*)$, where the components of $\mathbf{X}_n^*$ are i.i.d. drawn from F_n, the e.d.f. of $\mathbf{X}_n$. Let $\overline{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$, $S_n^2(\mathbf{X}_n) = \frac{1}{n}\sum_{i=1}^n (X_i - \overline{X}_n)^2$, $\overline{X}_n^* = \frac{1}{n}\sum_{i=1}^n X_i^*$ and $S_n^2(\mathbf{X}_n^*) = \frac{1}{n}\sum_{i=1}^n (X_i^* - \overline{X}_n^*)^2$.

3. Method I

In order to assess the accuracy of the bootstrap critical value $C_n(\alpha; \mathbf{X}_n)$, the following measure is currently in use in the literature:

$$P_{\boldsymbol{\theta}^0}(T_n(\mathbf{X}_n) \geq C_n(\alpha; \mathbf{X}_n)).$$

(See, e.g., [1] and [2].)

The power of the bootstrap-based test is evaluated similarly:

$$P_{\boldsymbol{\theta}^A}(T_n(\mathbf{X}_n) \geq C_n(\alpha; \mathbf{X}_n)).$$

It is important to note that $C_n(\alpha; \mathbf{X}_n)$ is a random variable which depends on the sample. In her PhD thesis, [3] remarked that *"this is exactly the problem in using the bootstrap to estimate the critical value. One property that a 'good' estimate of the critical value should satisfy is that the critical value estimate should be the same whether the null hypothesis is correct or not"*.

This remark by [3] led us to develop an alternative, more robust, evaluation method.

4. Method II

Suppose that $\mathbf{V}_n^0 = (V_1^0, V_2^0, \ldots, V_n^0)$ is a "pseudo random 'test' sample" with joint d.f. $F_{\boldsymbol{\theta}^0, \boldsymbol{\tau}}(\cdot)$ and assume that $\mathbf{V}_n^0$ is independent of the "training" sample $\mathbf{X}_n$. Here, $\boldsymbol{\tau}$ is a nuisance parameter which may differ from $\boldsymbol{\nu}$, the nuisance parameter defined in Section 2. It will also be replaced by some strongly consistent

estimator. In order to assess the accuracy of the bootstrap critical value $C_n(\alpha; \mathbf{X}_n)$, we propose the following measure:

$$P(T_n(\mathbf{V}_n^0) \geq C_n(\alpha; \mathbf{X}_n)).$$

Remark:
Let $\varphi^0(\mathbf{x}_n) = P(T_n(\mathbf{V}_n^0) \geq C_n(\alpha; \mathbf{x}_n))$, then

$$\begin{aligned} P(T_n(\mathbf{V}_n^0) \geq C_n(\alpha; \mathbf{X}_n)) &= E_{\boldsymbol{\theta}}\left(P(T_n(\mathbf{V}_n^0) \geq C_n(\alpha; \mathbf{X}_n)|\mathbf{X}_n)\right) \\ &= E_{\boldsymbol{\theta}}\left(\varphi^0(\mathbf{X}_n)\right), \end{aligned}$$

since

$$\begin{aligned} P(T_n(\mathbf{V}_n^0) \geq C_n(\alpha; \mathbf{X}_n)|\mathbf{X}_n = \mathbf{x}_n) &= P(T_n(\mathbf{V}_n^0) \geq C_n(\alpha; \mathbf{x}_n)|\mathbf{X}_n = \mathbf{x}_n) \\ &= P(T_n(\mathbf{V}_n^0) \geq C_n(\alpha; \mathbf{x}_n)) \\ &= \varphi^0(\mathbf{x}_n). \end{aligned}$$

The power of the bootstrap-based test can be evaluated similarly:

$$P(T_n(\mathbf{V}_n^A) \geq C_n(\alpha; \mathbf{X}_n)) = E_{\boldsymbol{\theta}}(\varphi^A(\mathbf{X}_n)),$$

where

$$\varphi^A(\mathbf{x}_n) = P(T_n(\mathbf{V}_n^A) \geq C_n(\alpha; \mathbf{x}_n)).$$

Now, $\mathbf{V}_n^A = (V_1^A, V_2^A, \ldots, V_n^A)$, has joint d.f. $F_{\boldsymbol{\theta}^A, \boldsymbol{\tau}}(\cdot)$ and is independent of the training sample $\mathbf{X}_n$.

5. Example: The mean in the univariate case

Let $\mathbf{X}_n = (X_1, X_2, \ldots, X_n)$ denote a random sample from an unknown univariate distribution F with finite mean μ and finite variance σ^2. In this section we will present some theoretical properties of the bootstrap estimate of the critical value for the test

$$H_0 : \mu = \mu_0 \quad \text{vs.} \quad H_A : \mu > \mu_0,$$

based on evaluation method II. This will be done using the "right" and "wrong" bootstrap critical value for the pivotal and non-pivotal test statistic. The term "right" refers to the case where resampling is done from data that are transformed in order to "mimic" H_0 in the bootstrap world, whereas the term "wrong" refers to resampling being done from the original sample. In what follows, the subscripts P and $N-P$ will refer to the pivotal and non-pivotal case, respectively, whereas the superscripts W and R will refer to the wrong and right critical value, respectively. We will consider the following four scenarios:

Scenario 1
The test rejects H_0 if and only if

$$T_{n,N-P}(X_n) = \sqrt{n}(\overline{X}_n - \mu_0) \geq C_{n,N-P}^W(\alpha; \mathbf{X}_n),$$

where $C^W_{n,N-P}(\alpha;\mathbf{X}_n)$ is defined by

$$P^*_{H^*_0}\left(\sqrt{n}(\overline{X}^*_n-\mu_0)\geq C^W_{n,N-P}(\alpha;\mathbf{X}_n)\right)\cong\alpha.$$

Denote this test procedure by (W,N-P).

Scenario 2

The test rejects H_0 if and only if

$$T_{n,P}(\mathbf{X}_n)=\frac{\sqrt{n}(\overline{X}_n-\mu_0)}{S_n(\mathbf{X}_n)}\geq C^W_{n,P}(\alpha;\mathbf{X}_n),$$

where $C^W_{n,P}(\alpha;\mathbf{X}_n)$ is defined by

$$P^*_{H^*_0}\left(\frac{\sqrt{n}(\overline{X}^*_n-\mu_0)}{S_n(\mathbf{X}^*_n)}\geq C^W_{n,P}(\alpha;\mathbf{X}_n)\right)\cong\alpha.$$

Denote this test procedure by (W,P).

Scenario 3

The test rejects H_0 if and only if

$$T_{n,N-P}(\mathbf{X}_n)=\sqrt{n}(\overline{X}_n-\mu_0)\geq C^R_{n,N-P}(\alpha;\mathbf{X}_n),$$

where $C^R_{n,N-P}(\alpha;\mathbf{X}_n)$ is defined by

$$P^*_{H^*_0}\left(\sqrt{n}(\overline{X}^*_n-\overline{X}_n)\geq C^R_{n,N-P}(\alpha;\mathbf{X}_n)\right)\cong\alpha,$$

Denote this test procedure by (R,N-P).

Scenario 4

The test rejects H_0 if and only if

$$T_{n,P}(\mathbf{X}_n)=\frac{\sqrt{n}(\overline{X}_n-\mu_0)}{S_n(\mathbf{X}_n)}\geq C^R_{n,P}(\alpha;\mathbf{X}_n),$$

where $C^R_{n,P}(\alpha;\mathbf{X}_n)$ is defined by

$$P^*_{H^*_0}\left(\frac{\sqrt{n}(\overline{X}^*_n-\overline{X}_n)}{S_n(\mathbf{X}^*_n)}\geq C^R_{n,P}(\alpha;\mathbf{X}_n)\right)\cong\alpha.$$

Denote this test procedure by (R,P).

The following theorems can now be derived:

Theorem 1. *Suppose that* $E(|X_1|^3)<\infty$, $E(X_1)=\mu_0$ *and* $\sigma^2=\sigma^2(V^0_1)$, *the variance of* V^0_1. *Then, as* $n\to\infty$,

$$P\left(\sqrt{n}(\overline{V}^0_n-\mu_0)\geq C^W_{n,N-P}(\alpha;\mathbf{X}_n)\right)\quad\to\quad 1-\Phi\left(\frac{z_{1-\alpha}}{\sqrt{2}}\right),$$

where $z_{1-\alpha}=\Phi^{-1}(1-\alpha)$ *and* $\overline{V}^0_n=\frac{1}{n}\sum_{i=1}^n V^0_i$.

Theorem 2. *Suppose that* $E(|X_1|^3) < \infty$, $E(X_1) = \mu_A$ *and* $\sigma^2 = \sigma^2(V_1^0)$, *the variance of* V_1^0. *Then, as* $n \to \infty$,

$$P\left(\sqrt{n}(\overline{V}_n^0 - \mu_0) \geq C_{n,N-P}^W(\alpha; \mathbf{X}_n)\right) \quad \to \quad 0.$$

Theorem 3. *Suppose that* $E(|X_1|^3) < \infty$ *and* $E(X_1) = \mu_0$. *Then, as* $n \to \infty$,

$$P\left(\frac{\sqrt{n}(\overline{V}_n^0 - \mu_0)}{S_n(\mathbf{V}_n^0)} \geq C_{n,P}^W(\alpha; \mathbf{X}_n)\right) \to 1 - \Phi\left(\frac{z_{1-\alpha}}{\sqrt{2}}\right),$$

where $S_n(\mathbf{V}_n^0)$ *is the sample standard deviation based on* $\mathbf{V}_n^0 = (V_1^0, V_2^0, \ldots, V_n^0)$.

Theorem 4. *Suppose that* $E(|X_1|^3) < \infty$ *and* $E(X_1) = \mu_A$. *Then, as* $n \to \infty$,

$$P\left(\frac{\sqrt{n}(\overline{V}_n^0 - \mu_0)}{S_n(\mathbf{V}_n^0)} \geq C_{n,P}^W(\alpha; \mathbf{X}_n)\right) \to 0.$$

Theorem 5. *Suppose* $E(|X_1|^5) < \infty$, $\sigma^2 = \sigma^2(V_1^0)$ *and that Cramér's continuity condition holds, i.e.,*

$$\limsup_{|t| \to \infty} |\chi(t)| < 1,$$

where $\chi(t)$ *is the characteristic function of* X_1. *Then, as* $n \to \infty$,

$$P\left(\sqrt{n}(\overline{V}_n^0 - \mu_0) \geq C_{n,N-P}^R(\alpha; \mathbf{X}_n)\right) = \alpha + C_n + O(n^{-3/2}), \text{ where}$$

$$C_n = \frac{z_{1-\alpha}\phi(z_{1-\alpha})\{z_{1-\alpha}^2(\mu_4 - \sigma^4) + (\mu_4 + 3\sigma^4)\}}{8\sigma^4 n} \text{ and } \mu_4 = E((X_1 - \mu)^4).$$

Theorem 6. *Suppose that* $E(X_1^6) < \infty$ *and that Cramér's continuity condition holds. Then, as* $n \to \infty$, *we have that*

$$P\left(\frac{\sqrt{n}(\overline{V}_n^0 - \mu_0)}{S_n(\mathbf{V}_n^0)} \geq C_{n,P}^R(\alpha; \mathbf{X}_n)\right) = \alpha + D_n + O(n^{-2}), \text{ where}$$

$$D_n = \frac{(1 + 2z_{1-\alpha}^2)\phi(z_{1-\alpha})\{\mu_3(21\sigma^4 + 15\mu_4) - 12\sigma^2\mu_5\}}{48\sigma^7 n^{3/2}} \text{ and } \mu_5 = E((X_1 - \mu)^5).$$

6. Concluding remarks

An extensive Monte-Carlo study was conducted for small to moderate sample sizes and various conclusions can be drawn from this:

(1) The estimated sizes of (W,N-P) and (W,P) converge to $1 - \Phi(\frac{1.645}{\sqrt{2}}) \approx 0.123$ as n increases, when the data $\mathbf{X}_n$ are generated from a distribution with the parameter specified by the *null hypothesis*. This agrees with the results of Theorem 1 and Theorem 3 ($\alpha = 0.05$).

(2) The estimated sizes of (W,N-P) and (W,P) converge to 0 as n increases, when the data $\mathbf{X}_n$ are generated from a distribution with the parameter specified by the *alternative hypothesis*. This agrees with the results of Theorem 2 and Theorem 4.

(3) The estimated sizes of (R,N-P) decreases monotonically to the nominal significance level as n increases. This is in accordance with the result of Theorem 5, where it was shown that $C_n \geq 0$. This test, therefore, tends to be "liberal".

(4) For *symmetrical distributions* the constant D_n (defined in Theorem 6) is equal to zero. The estimated sizes of (R,P) for symmetrical distributions attain the nominal significance level, even for *small* values of n.

(5) For *asymmetric distributions* we find that, while the estimated sizes of (R,P) are slightly more conservative than their symmetric counterparts, they are still close to the nominal significance level, even for small sample sizes.

(6) The Monte-Carlo study also showed that the estimated sizes of (R,P) appear to converge much quicker to the nominal significance level than the test (R,N-P). This agrees with the results of Theorem 5 and Theorem 6.

It is clear from the discussions in this section, that it is preferable to make use of a pivotal test statistic and the "right" bootstrap critical value. These findings are in agreement with the two guidelines proposed by [4]. For the purpose of this discussion we only considered a very simplistic problem to illustrate our findings more clearly. However, extensions to more complex tests can be found in [5].

References

[1] Fisher N.I., Hall P. (1990) On bootstrap hypothesis testing. *The Australian Journal of Statistics*, **32**, 177-190.

[2] Martin M.A. (2007) Bootstrap hypothesis testing for some common statistical problems: A critical evaluation of size and power properties. *Computational Statistics and Data Analysis*, **51**, 6321-6342.

[3] Sakov A. (1998) *Using the m out of n bootstrap in hypothesis testing.* PhD thesis, University of California, Berkeley.

[4] Hall P., Wilson S.R. (1991) Two guidelines for bootstrap hypothesis testing. *Biometrics*, **47**, 179-192.

[5] Allison J.S. (2008) *Bootstrap-based hypothesis tests*, PhD thesis, North-West University, Potchefstroom.

6th St.Petersburg Workshop on Simulation (2009) 755-759

On Robust Testing for Normality[1]

Luboš Střelec[2], Milan Stehlík[3]

Abstract

The aim of this paper is to introduce the general form of the robust Jarque-Bera test to systematize the results from some recent studies on variants of Jarque-Bera tests and give general guidelines for appropriate small sample testing for normality. Particularly, the special cases of this class are the classical Jarque-Bera test, the Jarque-Bera-Urzua test, the robust Jarque-Bera test introduced by Gel and Gastwirth (see [6]), the Geary's test and Uthoff's test. We prove the asymptotical normality of introduced robust measures of skewness and kurtosis, together with the consistence of given tests. The introduced test statistics have asymptotically χ^2_2 distribution, as does the Jarque-Bera statistic. Our tests are robust and have higher power than the medcouple tests and classical Jarque-Bera test. The introduced general class of robust tests of the normality is illustrated with selected datasets of financial time series.

1. Introduction

Many tests have been developed to check the validity of normality assumption ([4], [7], [8], among others). For thorough discussion on various normality tests see [13]. Today the most popular omnibus test for normality in general use is the Shapiro-Wilk (SW) test. The Jarque-Bera (JB) test is the most widely adopted omnibus test for normality in econometrics and related fields. The Lilliefors (Kolmogorov–Smirnov) ($L(KS)$) test is the best known omnibus test based on the empirical distribution function (EDF). Being omnibus procedures, the SW, JB and $L(KS)$ tests do not provide insight about the nature of the deviation from normality, e.g. skewness, heavy tails or outliers. Therefore, specialized tests directed at particular alternatives are desired in many practical situations. Interestingly enough (see [11]), the classical Jarque-Bera test has been known among statisticians since the work of [1]. They derived it after noting that, under normality, the asymptotic means of $\sqrt{b}_1$ and b_2 are 0 and 3, the asymptotic variances are $6/n$ and $24/n$, and the asymptotic covariance is 0. Yet, there are few instances in the statistics

[1]Research was supported by project AKTION Austria - Czech Republic, Nr. 51p7.

[2]Mendel University of Agriculture and Forestry in Brno, E-mail: xstrelec@node.mendelu.cz

[3]Johannes Kepler University in Linz, E-mail: Milan.Stehlik@jku.at

literature where the Bowman-Shenton-Jarque-Bera test has been studied. As one author states in a comprehensive survey of tests for normality "Due to the slow convergence of b_2 to normality this test is not useful" (see [4], p. 391). As pointed out by several authors (see for example [11]) the classical Jarque-Bera test behaves well in comparison with some other test for normality if the alternatives belong to the Pearson family. However, the Jarque-Bera test behaves very badly for distributions with short tails and bimodal shape, sometimes it is even biased (see [10]). On the other hand, the power of the modified tests is higher then the power of tests based on medcouple.

The aim of this paper is to systematize the robust normality testing recently addressed by several authors (see for example [6], [2], among others). For this reason we introduce a general class of robust tests, the so called RT class. It will be seen that the general robust class of tests will accommodate the alternatives which are problematic for the JB test: bimodal, Weibull and uniform alternatives. To maintain the continuity of explanation proofs are put into Appendix.

2. General approach to robust JB test

In this paper we use the following notation. Let $\mu_k = E(X - E(X))^k$ be a k-th theoretical moment of the distribution, let $\hat{\mu}_k = (1/n)\sum_{i=1}^{n}(X_i - \bar{X})^k$ be its empirical variant and let $m_{k,n} = (1/n)\sum_{i=1}^{n}(X_i - M_n)^k$ be its robust variant, where M_n denotes the sample median. Now let us introduce the RT class of Jarque-Bera tests. For that reason we relax the form of moment μ_j estimator by putting $M_{i,j} = (1/n)\sum_{m=1}^{n}\varphi_j(X_m - M_{(i)})$, where $i \in \{0,1\}$ denotes either the arithmetic mean $M_{(0)} = \bar{X}$ or median $M_{(1)} = M_n$, $j \in \{0,1,2,3,4\}$ and φ_j is tractable and continuous function $\varphi_0 = \sqrt{\pi/2}|x|$, $\varphi_1 = x$, $\varphi_2 = x^2$, $\varphi_3 = x^3$ and $\varphi_4 = x^4$. The RT class of Jarque-Bera test is defined by

$$RT = \frac{k_1(n)}{C_1}\left(\frac{M_{i_1,j_1}^{\alpha_1}}{M_{i_2,j_2}^{\alpha_2}}\right)^2 + \frac{k_2(n)}{C_2}\left(\frac{M_{i_3,j_3}^{\alpha_3}}{M_{i_4,j_4}^{\alpha_4}} - 3\right)^2. \tag{1}$$

The special cases of RT class are: the classical Jarque-Bera, the test of Urzua (see [11]), the robust Jarque-Bera test (see [6]), Geary's test a (see [5]) (originally denoted w_n') and Uthoff's test U (see [12]). The following theorems justify the feasibility of RT class.

Theorem 1. i) *Under the null hypothesis we have* $\lim_{n\to\infty} E(m_{k,n}) = \mu_k, k = 2,3,4$, *i.e.* $m_{k,n}$ *is an asymptotically unbiased estimator of* μ_k. *Here we consider only* $k = 2,3,4$ *since only these moments play roles in the Jarque-Bera test.*

ii) $M_{i,j}$ *are consistent estimators of* μ_j *for* $j = 0,1,2,3,4$. *Here* $\mu_0 = \sigma$.

Theorem 2. i) *Let* $X_1, ..., X_n$ *be iid* $N(\mu, \sigma^2)$, *i.e. null hypothesis holds. Then*

$$\sqrt{n}\begin{pmatrix} M_{3,i}/M_{2,j}^{3/2} \\ M_{4,k}/M_{2,l}^2 - 3 \end{pmatrix} \to N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} C_1 & 0 \\ 0 & C_2 \end{pmatrix}\right) \tag{2}$$

where $i, j, k, l \in \{0, 1\}$ denotes either to use arithmetic mean (by taking $\{0\}$) or median (by taking $\{1\}$).

ii) *Let $X_1, ..., X_n$ be iid $N(\mu, \sigma^2)$, i.e. null hypothesis holds. Then*

$$\sqrt{n}\begin{pmatrix} M_{3,i}/M_{0,j}^3 \\ M_{4,k}/M_{0,l}^4 - 3 \end{pmatrix} \to N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} C_1 & 0 \\ 0 & C_2 \end{pmatrix}\right) \tag{3}$$

where $i, j, k, l \in \{0, 1\}$ denotes using of arithmetic mean $\{0\}$ or median $\{1\}$.

Corollary 1 The *RT* test statistics from Theorem 2 asymptotically follow χ_2^2.

Remark: Choosing appropriate constants C_1 and C_2 is the hardest aspect of the RT variants of Jarque-Bera test. To obtain the constants C_1, C_2 we need to find the expressions for $E(M_{n_1,n_2}^k)$ for a finite sample size. Such calculations are very tedious and therefore we obtain these constants from Monte Carlo simulations (see Table 1 and Table 2).

Table 1: Monte Carlo simulations of C_1

skewness measure	n						asympt.
	25	50	100	200	500	1000	
$M_{3,0}/M_{2,0}^{3/2}$	4.71	5.40	5.69	5.91	6.00	5.98	6
$M_{3,0}/M_{2,1}^{3/2}$	4.32	5.15	5.55	5.83	5.97	5.97	6
$M_{3,1}/M_{2,0}^{3/2}$	15.66	16.36	16.67	16.87	17.17	17.23	18
$M_{3,1}/M_{2,0}^{3/2}$	13.51	15.18	16.01	16.53	17.03	17.16	18
$M_{3,0}/J_{n,0}^3$	5.73	6.04	6.05	6.10	6.07	6.00	6
$M_{3,0}/J_{n,1}^3$	6.44	6.40	6.22	6.18	6.10	6.02	6
$M_{3,1}/J_{n,0}^3$	16.14	16.69	16.88	16.96	17.19	17.22	18
$M_{3,1}/J_{n,1}^3$	18.85	18.06	17.57	17.30	17.33	17.29	18

Table 2: Monte Carlo simulations of C_2

kurtosis measure	n						*asympt.*
	25	50	100	200	500	1000	
$M_{4,0}/M_{2,0}^2 - 3$	13.02	18.46	21.08	23.07	23.91	23.32	24
$M_{4,0}/M_{2,1}^2 - 3$	12.92	18.27	20.95	23.02	23.84	23.31	24
$M_{4,1}/M_{2,0}^2 - 3$	19.09	22.05	23.18	24.10	24.42	23.54	24
$M_{4,1}/M_{2,1}^2 - 3$	15.28	19.98	22.04	23.55	24.15	23.42	24
$M_{4,0}/J_{n,0}^4 - 3$	52.58	59.32	58.81	59.68	59.26	57.67	58
$M_{4,0}/J_{n,1}^4 - 3$	59.37	62.97	60.46	60.42	59.63	57.84	58
$M_{4,1}/J_{n,0}^4 - 3$	61.95	64.63	61.78	60.99	59.89	57.95	58
$M_{4,1}/J_{n,1}^4 - 3$	75.82	71.16	64.60	62.23	60.45	58.22	58

In [9] we discuss power comparisons for normality testing against single alternative distributions - heavy tailed and light tailed alternatives, and also against mixtures alternative distributions, i.e. contaminated normal distribution and mixture of gamma and log-gamma distribution. As follows from our simulations the ranking of considered competing tests depends heavily on the type of tails.

3. Illustrative examples and conclusions

In what follows the RT_{JB} and RT_{RJB} tests will be used for normality testing of several datasets of financial time series. Source data include logarithmic returns of monthly average prices of stock exchange indexes PX and DJI and monthly average prices of CZK/EUR and CZK/USD exchange rate in the period from 1995 to 2008. Table 4 in [9] contains the p-values of used classical normality tests the Anderson-Darling test (AD), the Cramer-von Misses test (CM), the Lilliefors test (LT), the D'Agostino test (DT), the Jarque-Bera test (JB), the Jarque-Bera-Urzua test (JBU), the robust Jarque-Bera (RJB), the Shapiro-Wilk test (SW), the directed SJ test (SJ_{dir}), RT_{JB}, RT_{JBU}, RT_{RJB} and RT_{RJBU} classes and Medcouple tests ($MC1-3$) [2]. Figure 1 contains histograms and Q-Q plots of the most interesting time series of our analysis, i.e. CZK/USD and PX datasets. The null hypothesis of logarithmic returns of the Prague stock market PX is not rejected at the 1 % significance level in the case of the AD, CM, LT, directed SJ and $MC1-3$ tests and is rejected at the 1 % significance level in the case of the DT, JB, JBU, RJB, SW and RT class tests by reason of the asymmetric distribution and higher kurtosis than the normal distribution. The tests of the first class (AD, CM, LT, directed SJ and $MC1-3$ tests) are not based on skewness and kurtosis measures. On the other hand the null hypothesis of logarithmic returns of CZK/USD exchange rate is not rejected at the 5 % significance level for all analyzed normality tests. The most interesting are the ranges of RT tests. For instance p-values of RT_{JB} tests belong to the range$(0.10, 0.48)$, p-values of RT_{JBU} tests belong to the range $(0.09, 0.41)$, p-values of RT_{RJB} tests belong to the range $(0.13, 0.58)$ and p-values of RT_{RJBU} tests belong to the range $(0.13, 0.54)$. Note that p-values of the classical Jarque-Bera test, the Jarque-Bera-Urzua test and the robust Jarque-Bera test are on the lower bound of RT classes ranges. We can also see that the differences in the p-values of Jarque-Bera test and Urzua's modified Jarque-Bera test are very small but the differences in the p-values within RT_{JB} and RT_{JBU} classes are significant.

Conclusions This paper introduces the general class RT of robust tests for normality and discuss their properties. The further theoretical considerations of class RT will be of interest. Based upon our experience we can recommend the following general guidelines for normality testing.

1. Case by case approach

Different tests are appropriate in different situations. For instance, it is important to know whether the alternative belongs to the heavy tailed class or not.

2. Trade off between power and robustness

Two typical extremal behaviors occur in robust testing: the tests which are more robust have smaller power (since they are not affected by single outliers) and tests with higher power are typically less robust (because they are affected by single outliers). An example of the first extreme case is the Medcouple test and an example of the second extreme case is the RT test.

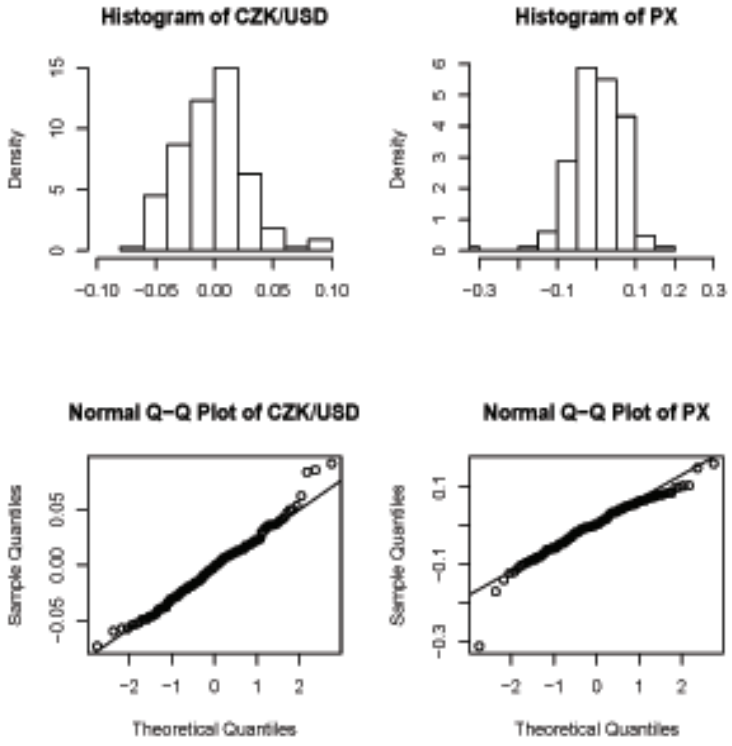


Figure 1: Histograms and Q-Q plots of several datasets of financial time series

4. Appendix

Proof of Theorem 1 i) Under the null, μ is the mean of Normal distribution and $\sqrt{n}(M_n - \mu) \sim N(0, 1/(4f(\mu)^2))$ (see [3], p. 484). Thus $E(\mu - M_n)^k \to 0$ for $n \to \infty$.

For $k = 2$ we have $E(X_i - M_n)^2 = E(X_i - \mu)^2 + E(\mu - M_n)^2$.

For $k = 3$ we have $E(X_i - M_n)^3 = 3E[(X_i - \mu)^2(\mu - M_n)] + E(\mu - M_n)^3$ Thus $E(\mu - M_n)^3 \to 0$ for $n \to \infty$ and from Cauchy-Schwarz inequality we have $|E(X_i - \mu)^2(\mu - M_n)| \leq \sqrt{E(X_i - \mu)^4 E(\mu - M_n)^2} \to 0$ for $n \to \infty$.

For $k = 4$ we have $E(X_i - M_n)^4 = E(X_i - \mu)^4 + 6E[(X_i - \mu)^2(\mu - M_n)^2] + E(\mu - M_n)^4$. Thus $E(\mu - M_n)^4 \to 0$ for $n \to \infty$ and from Cauchy-Schwarz inequality we have $|E(X_i - \mu)^2(\mu - M_n)^2| \leq \sqrt{E(X_i - \mu)^4 E(\mu - M_n)^4} \to 0$ for $n \to \infty$.

ii) M_n converges in probability to μ because $P(|M_n - \mu| \geq \epsilon) \sim 1/(4f(\mu)^2 n)$, $\forall \epsilon > 0$ and $\sqrt{n}(M_n - \mu) \sim N(0, 1/(4f(\mu)^2))$ (see [3], p. 484). Since $g_j(u) = (1/n)\sum_{i=1}^{n} \varphi_j(X_i - u)$ is continuous function for $j = 0, 1, 2, 3, 4$, $g_j(M_n)$ converges in probability to $g(\mu)$ which is the consistent estimator of μ_j. Therefore also $g_j(M_n)$ is the consistent estimator of μ_j, where $\mu_0 = \sigma$.

Proof of Theorem 2 i) From convergence in probability, we have the following convergences in distribution $M_{3,i} \to \hat{\mu_3}$, $M_{4,j} \to \hat{\mu_4}$, $M_{2,k}^{3/2} \to \sigma^3$ and $M_{2,l}^2 \to \sigma^4$ for $n \to \infty$. The proof is completed by employing of multivariate Slutsky's theorem.

ii) From convergence in probability, we have the following convergences in distribution $M_{3,i} \to \hat{\mu_3}$ and $M_{4,j} \to \hat{\mu_4}$. The rest of the proof follows from Theorem 1 and its proof in [6].

□

Acknowledgement The authors are grateful for the comments of anonymous referee which improved a quality of the paper. We are thankful for the support of Simos Meintanis.

References

[1] Bowman, K.O., Shenton, L.R. (1975). Omnibus contours for departures from normality based on $\sqrt{b}_1$ and b_2. *Biometrika*, 62, p. 243–250.

[2] Brys, G., Hubert, M., Struyf, A. (2004). A Robustification of the Jarque-Bera Test of Normality. COMPSTAT 2004, *Proceedings in Computational Statistics*, ed. J. Antoch, Springer, Physica Verlag. 753–760.

[3] Cassela, G., Berger, R.L. (2002). *Statistical Inference*, 2nd Edition, Duxbury Advanced Series, Thomson Learning.

[4] D'Agostino, R.B. (1986). Tests for normal distribution. In D'Agostino, R.B. and Shephens, M.A.: *Goodness of fit techniques*. New York: Marcel Dekker, 1986, 367–419.

[5] Geary, R.C. (1935). The ratio of the mean deviation to the standard deviation as a test of normality. *Biometrika*, 27, 310–332.

[6] Gel, Y.R., Gastwirth, J.L. (2008). A robust modification of the Jarque-Bera test of normality. *Economics Letters*, 99, 30–32.

[7] Jarque, C.M. and Bera, A.K. (1980). Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Economics Letters*, 6 (3), 255–259.

[8] Shapiro, S.S., Wilk, M.B. (1965). An analysis of variance test for normality. *Biometrika.* 52, 3 and 4, 591–611.

[9] Střelec, L., Stehlík, M. (2008). Some properties of robust tests for normality and their applications, IFAS Research report 2009.

[10] Thadewald, T., Bunning, H. (2004). Jarque-Bera test and its competitors for testing normality - a power comparison. *Volkswirtschaftliche Reihe*, der Freien Universitat Berlin, 2004/9. 52, 3 and 4, 591–611.

[11] Urzua, C.M. (1996). On the correct use of omnibus tests for normality. *Economics Letters*, 53, 1996, 247–251.

[12] Uthoff, V.A. (1973). The most powerful scale and location invariant test of the normal versus the double exponential. *Annals of Statistics*, 1, 1973, 170–174.

[13] Thode, H.C. (2002). *Testing for normality*, Statistics: textbooks and monographs. Dekker,New York.

6th St.Petersburg Workshop on Simulation (2009) 761-765

Tests of the exponentiality based on properties of order statistics[1]

Ksenia Volkova[2]

Abstract

We construct new tests of exponentiality based on the order statistic characterizations of exponential law. We calculate limiting distributions and local Bahadur efficiencies of new tests.

1. Introduction

Exponential law is one of central laws in Probability theory. Models with exponential observations often appear in practice, e.g., in reliability theory and survival analysis. Testing of exponentiality occupies the important place in Statistics.

Consider a sample $X_1, \ldots, X_n$ of i.i.d. rv's, with continuous df F. Let F_n be the usual empirical df. We are testing the hypothesis H_0 : F is exponential with the density $\lambda e^{-\lambda x}, x \geq 0, \lambda > 0$ against the alternative hypothesis H_1 : $F(x)$ is non-exponential df.

One of the well-known characterization of exponential distribution was obtained by Desu [2]: *Let X and Y - non-negative i.i.d. random variables with df F. Then X and $2\min(X, Y)$ are identically distributed iff F is exponential df.* Tests based on Desu's characterization were studied in [3] and [4].

We construct the test of exponentiality based on the generalization of Desu's characterization, which follows from [6]: *Let $X_1, \ldots, X_n$ be non-negative i.i.d. rv's and let natural k and m be such that $\frac{\log k}{\log m}$ is an irrational number. Then $k\min(X_1, \ldots, X_k)$ and $m\min(X_1, \ldots, X_m)$ are identically distributed iff the sample has the exponential distribution.*

In accordance with this characterization let construct for any natural $l \geq 1$ the $V-$statistical df $G_l(t)$ and $U-$statistical df $H_l(t)$

$$G_l(t) = n^{-l} \sum_{i_1,\ldots,i_l=1}^{n} \mathbf{1}\{l\min(X_{i_1}, \ldots, X_{i_l}) < t\},\ t \geq 0,$$

$$H_l(t) = \binom{n}{l}^{-1} \sum_{1\leq i_1<\ldots<i_l\leq n} \mathbf{1}\{l\min(X_{i_1}, \ldots, X_{i_l}) < t\},\ t \geq 0.$$

[1]This work was supported by grant RFBR 07-01-00159-a and NSh.638.2008.1

[2]St.-Petersburg State University, E-mail: efrksenia@mail.ru

We suggest two following statistics for testing of H_0 versus H_1:

$$\begin{aligned} R_{k,m} = &\quad \int_0^\infty (G_k(t) - G_m(t))\, dF_n(t), \\ S_{k,m} = &\quad \sup_{t\geq 0} |H_k(t) - H_m(t)|. \end{aligned}$$

These statistics generalize the statistics introduced in [3] and [4]. Note that $R_{k,m}$ is a V-statistic, while $S_{k,m}$ is a supremum of a family of U-statistics.

We consider the standard alternatives used in testing of exponentiality with the densities g_1, g_2, g_3 :

– Weibull with $g_1(x,\theta) = (1+\theta)x^\theta \exp(-x^{1+\theta}), x \geq 0$;

– Makeham with $g_2(x,\theta) = (1+\theta(1-e^{-x}))\exp(-x-\theta(e^{-x}-1+x)), x \geq 0$;

– linear failure rate with $g_3(x,\theta) = (1+\theta x)e^{-x-\frac{1}{2}\theta x^2}, x \geq 0$.

Now recall some notions from Bahadur theory [1]. Suppose we have a sequence of observations $s = (X_1, X_2, \ldots)$ with general distribution P_θ, where $\theta \in \Theta \subset R^1$. We are testing $H_0 : \theta \in \Theta_0 \subset \Theta$ versus $H_1 : \theta \in \Theta_1 = \Theta \setminus \Theta_0$. For that end we use the sequence of statistics $T_n(s) = T_n(X_1, \ldots X_n)$. Denote $F_n(t;\theta) = P_\theta(s : T_n(s) < t), H_n(t) = \inf\{F_n(t;\theta) : \theta \in \Theta_0\}$.

In typical cases one has $\lim_{n\to\infty} n^{-1}\ln[1-H_n(T_n(s))] = -\frac{1}{2}c_T(\theta)$, for $\theta \in \Theta_1$, in P_θ-probability, where $0 < c_T(\theta) < +\infty$; c_T is called Bahadur's exact slope of the sequence $\{T_n\}$. The following result shows how to compute $c_T(\theta)$ [1].

Theorem 1. *Let $T_n \to b(\theta)$ in P_θ-probability, for finite $b(\theta)$, $\theta \in \Theta_1$, and $\lim_{n\to\infty} n^{-1}\ln[1-H_n(t))] = -f(t)$, for $\forall t \in I$, where f is continuous on the open interval I and $\{b(\theta), \theta \in \Theta_1\} \subset I$. Then one has $c_T(\theta) = 2f(b(\theta))$, for $\theta \in \Theta_1$.*

Define the Kullback-Leibler information $K(\theta,\theta_0) = K(P_\theta, P_{\theta_0})$ by the formula

$$K(P_\theta, P_{\theta_0}) = \begin{cases} \int \ln \frac{dP_\theta}{dP_{\theta_0}}\, dP_\theta, P_\theta \ll P_{\theta_0} \\ +\infty, \text{ otherwise.} \end{cases} .$$

Put $K(\theta,\Theta_0) = \inf\{K(P_\theta, P_{\theta_0}) : \theta_0 \in \Theta_0\}$. It is well-known [1] that always $c_T(\theta) \leq 2K(\theta,\Theta_0)$, $\theta \in \Theta_1$. Hence it is natural to define the local Bahadur efficiency of the sequence $\{T_n\}$ as $eff(T) = \lim_{\theta\to\partial\Theta_0} c_T(\theta)/2K(\theta,\Theta_0)$.

We see from Theorem 1 that we should know the information on large deviation asymptotics of statistics under consideration. For any centred symmetric kernel Φ of degree m define $\psi(x) = \mathbb{E}\{\Phi(X_1, \ldots X_m) \mid X_1 = x\}$, $\Delta^2 = \mathbb{E}\psi^2(X_1)$. Next theorem was proved in [5]:

Theorem 2. *Let the kernel Φ of von Mises functional be bounded,centred and $\Delta^2 > 0$. Then $\lim_{n\to\infty} n^{-1}\ln \mathbb{P}\{V_n \geq a\} = \sum_{j=2}^\infty b_j a^j$, where the series with coefficients b_j converges for sufficiently small $a > 0$, moreover $b_2 = -(2m^2\Delta^2)^{-1}$.*

2. Statistics $R_{k,m}$

For the statistic $R_{k,m}$ we have $\Delta^2 = \Delta^2(k,m) = \frac{(m-k)^2(8mk-2k-2m+1)}{4(2k+2m-1)(4m-1)(4k-1)(m+1)^2}$ under H_0. Applying Hoeffding's Theorem, we obtain

Theorem 3. *For statistic $R_{k,m}$ under H_0 and for $n \to \infty$ we obtain*

$$\sqrt{n}R_{k,m} \xrightarrow{d} \mathcal{N}(0, (m+1)^2\Delta^2(k,m)).$$

To compute the local Bahadur efficiencies for our alternatives, we use Theorem 2 and obtain the following results:

$$eff(R_{k,m}; Weibull) = \frac{3\ln^2(\frac{m}{k})(2k+2m-1)(4m-1)(4k-1)}{2\pi^2(m-k)^2(8mk-2k-2m+1)};$$
$$eff(R_{k,m}, Makeham) = \frac{12(2k+2m-1)(4m-1)(4k-1)}{(2k+1)^2(2m+1)^2(8mk-2k-2m+1)};$$
$$eff(R_{k,m}, LFR) = \frac{(2k+2m-1)(4m-1)(4k-1)}{16m^2k^2(8mk-2k-2m+1)}.$$

The investigation of these expressions for local Bahadur efficiency for different k and m shows that the efficiency decreases when k and m increase, and the best case with maximal efficiency is $k = 1$ and $m = 2$ which was discussed in [3]. Next table presents the results for maximum local Bahadur efficiency for various alternatives when $k = 1$ and $m = 2$:

alternative	efficiency
Weibull	0.697
Makeham	0.509
Linear failure rate	0.149

Thus we showed that for the statistics $R_{k,m}$ the local Bahadur efficiency for standard alternatives is the best for Desu's case ($k = 1$ and $m = 2$) and its use for other k and m is hardly justified.

3. Statistics $S_{k,m}$

For the statistics $S_{k,m}$ under H_0 we obtain the variance function for the family of kernels depending on t in the form

$$\Delta^2(k,m,t) = \sup_{t>0}(1-2\frac{k}{m})e^{-2t+\frac{t}{m}} + \frac{k^2}{m^2}e^{-2t+\frac{t}{k}} - \frac{(m-k)^2}{m^2}e^{-2t}.$$

The limiting distribution for these statistics is non-normal and can be in principle obtained from [7]. In order to compute the local efficiency we use [4]. The results of computing local Bahadur efficiency for our alternatives are:

$$eff(S_{k,m}; Weibull) = \frac{6e^{-2}\ln^2(\frac{m}{k})}{\pi^2m^2\sup_{t>0}\Delta^2(k,m,t)};$$
$$eff(S_{k,m}; Makeham) = \frac{12(\sup_{t>0}[e^{-t}(k(e^{-\frac{t}{k}}-1)-m(e^{-\frac{t}{m}}-1))])^2}{m^2\sup_{t>0}\Delta^2(k,m,t)};$$
$$eff(S_{k,m}; LFR) = \frac{4e^{-4}(m-k)^2}{k^2m^4\sup_{t>0}\Delta^2(k,m,t)}.$$

Explicit computation of $\sup_{t>0} \Delta^2(k,m,t)$ is impossible, but we can use numerical methods using the package Maple 8. It turns out that for $k, m \geq 10$ the values of $\sup_{t>0} \Delta^2(k,m,t)$ are negligible. For remaining values of k, m we calculate approximate values of local Bahadur efficiency for standard alternatives. This enables us to find out the maximum local Bahadur efficiency.

The next table presents the maximum of local Bahadur efficiency for different alternatives and the values of k and m, where this maximum is attained.

alternative	efficiency	k and m
Weibull	0.2613	k=1,m=6
Makeham	0.2178	k=1,m=3
Linear failure rate	0.0732	k=1,m=2

For the Weibull alternative we improve the value of efficiency found in [4], namely $eff(S_{1,2}) = 0.158$. *Thus for the statistics $S_{k,m}$ we can increase efficiency selecting suitable k and m.*

We see that the value of maximum local Bahadur efficiency for $S_{k,m}$-statistics is less than for $R_{k,m}$-statistics. This drawback is partially compensated by the fact that the tests based on supremum-type statistics are consistent for any alternative.

References

[1] Bahadur R.R. (1971) *Some limit theorems in statistics.* Philadelphia: SIAM.

[2] Desu M.M. *A characterization of the exponential distribution by order statistics.* Ann. Math. Stat., 1971, 42 N 2, 837-838.

[3] Litvinova V.V. (2005) *Asymptotic properties for symmetry and goodness-of-fit tests based on characterization.* Ph.D.thesis, St.Petersburg.

[4] Nikitin Ya.Yu. (2009)*Large deviation for U-statistical analogues of criteria Kolmogorov-Smirnov.* Submitted to J. of Nonpar. Statistics.

[5] Nikitin Ya.Yu., Ponikarov E.V.(1999) *Rough large deviation asymptotics of Chernoff type for von Mises functionals and U-statistics.* Proc. of St. Petersburg Mathem. Society, 7, 124-167.

[6] Shimizu R. (1979) *A characterization of the exponential distribution.* Ann. Inst. Statist. Math., 31, No. 3, 367-372.

[7] Silverman B. W. (1983) *Convergence of a class of empirical distribution functions of dependent random variables.* Ann. Probab. 11, 745-751.

6th St.Petersburg Workshop on Simulation (2009) 765-769

Goodness-of-fit tests for the Weibull and Pareto distributions

Gennadi Martynov[1]

Abstract

In this thesis we will present the new class of the parametric distribution families such, that the limit distributions of the goodness-of-fit statistics based on the empirical process do not depend of unknown parameters. This is family $\{R((x/\beta)^{\alpha})),\ \alpha > 0,\ \beta > 0,\ x \in \mathcal{X} \subset [0,\infty)\}$, where α and β are unknown parameters. It was considered the Pareto and Weibull distribution families. The method was presented for calculation the eigenvalues of the corresponding covariance operator.

1. Introduction

Let $X^n = \{X_1, X_2, ..., X_n\}$ be the sample from the r.v. with the distribution function $F(x),\ x \in R_1$. We will test the hypothesis

$$H_0 :\ \ F(x) \in \mathcal{G} = \{G(x,\theta),\ \ \theta = (\theta_1, \theta_2, ...\theta_k)^{\top} \in \Theta \subset R_k\},$$

where θ is an unknown vector of parameters. We will consider the Cramér-von Mises statistic

$$\omega_n^2 = n \int_{-\infty}^{\infty} \psi^2(G(x,\theta_n))(F_n(x) - G(x,\theta_n))^2\, dG(x,\theta_n),$$

θ_n is an estimator of θ, $\psi(t)$ is the weight function, $F_n(x)$ is the empirical distribution function. The results below are applicable also to the Kolmogorov-Smirnov statistic

$$D_n = \sqrt{n} \sup_{-\infty<x<\infty} |\psi(G(x,\theta_n))(F_n(x) - G(x,\theta_n))|.$$

The exact methods for calculating the limit distribution are developed mostly for the Cramér-von Mises statistic (see [4], [8], [10], [11], [12]).
Let θ_n be the likelihood maximum estimator of θ. Under the certain number of the regularity conditions and under H_0 limit distribution of the statistics ω_n^2 coincide with the distribution of the functional

$$\omega^2 = \int_0^1 \psi^2(t)\xi^2(t,\theta_0)dt$$

[1]Institute for information transmission problems of RAS, E-mail: martynov@iitp.ru

of the Gauss process $\psi(t)\xi(t,\theta_0)$ with $E\psi^2(t)\xi(t,\theta_0)=0$, and with the covariance function

$$K(t,\tau)=E(\psi(t)\xi(t,\theta_0)\psi(\tau)\xi(\tau,\theta_0))$$
$$=\psi(t)\psi(\tau)(K_0(t,\tau)-q^{\top}(t,\theta_0)I^{-1}(\theta_0)q(\tau,\theta_0)),$$

where $K_0(t,\tau)=\min(t,\tau)-t\tau, \quad t,\tau\in(0,\ 1)$, θ_0 is an unknown value of the parameter θ,

$$q^{\top}(t,\theta)=\left(\partial G(x,\theta)/\partial\theta_1,...,\partial G(x,\theta)/\partial\theta_k\right)|_{t=G(x,\theta)}\,,$$

$I(\theta)$ is the Fisher information matrix,

$$I(\theta)=\left(E((\partial/\partial\theta_i)\log g(X,\theta)(\partial/\partial\theta_j)\log g(X,\theta))\right)_{1\le i,j\le k}\,,$$

$$g(x,\theta)=(\partial G(x,\theta)/\partial x).$$

The follow condition must be fulfilled:

$$\int_0^1\psi^2(t)K(t,t)dt<\infty.$$

The limit distribution for D_n coincides with the distribution of

$$D=\sup_{0<t<1}|\psi(t)\xi(t,\theta_0)|,$$

but the conditions on $\psi(t)$ and another conditions are different from the conditions for ω^2. They was studied in [3], [13]. The distribution of ω^2 depends generally from θ_0 and the distribution family $\mathcal{G}$. Khmaladze [9] has proposed the method of empirical process transformation for eliminate such dependance. Khmaladze and Haywood [7] has applied this method to exponentiality testing by the Cramér-von Mises statistic.

We will use here the traditional approach consistsing in using of the statistic ω_n^2. It is well known (see for example [8], [10]]) that the empirical process does not depend on unknown parameter θ_0 for the family of the form

$$\mathcal{G}=\{G((x-m)/\sigma),\ \ -\infty<x<\infty,\,\sigma>0\}.$$

Most known example of such family is the normal distribution family (see [5], [8)]. We will propose here another class of the distribution family with such property:

$$\mathcal{R}=\{R((x/\beta)^{\alpha}),\ \alpha>0,\ \beta>0,x\in\mathcal{X}\subset[0,\infty)\},$$

where $\mathcal{X}$ is the support of the distribution $R((x/\beta)^{\alpha})$. Here $R(z)$ is a distribution function with the support $\mathcal{Z}\subset[0,\infty)$. Particular cases of such families are Weibull and Pareto distributions. The limit distributions of Cramér-von Mises and Kolmogorov-Smirnov statistics do not depend on the unknown parameters in both families. Additionally, the limit distribution for Pareto family coincide with analogous distribution for exponential family. The goodness-of-fit tests is discussed now for the general Pareto distribution in many articles, particularly, in [1], [2], [6].

2. General results

Let $X^n = \{X_1, X_2, ..., X_n\}$ be the sample from the r.v. with a distribution function $F(x),\ x \in R_1$. We will test the hypothesis

$$H_0 :\ \ F(x) \in \mathcal{R} = \{R((x/\beta)^\alpha)),\ \alpha > 0,\ \beta > 0,\ x \in \mathcal{X} \subset [0, \infty)\},$$

where α and β are unknown parameters. The set of the alternative distributions contains all another distributions. Here $R(z)$ is the distribution function with the support $\mathcal{Z} \subset [0, \infty)$. We note the corresponding density function by $r(z)$. $\mathcal{R}$ is the family of Pareto distributions with $R(z) = 1 - 1/z,\ z > 1$ and $x > \beta$. The family $\mathcal{R}$ consists of Weibull distributions when $R(z) = 1 - \exp(-z),\ z > 0$, and $x > 0$. We will use the Cramér-von Mises and Kolmogorov-Smirnov tests. Both of them based on the empirical process $\xi_n(x) = \sqrt{n}(F_n(x) - R((x/\hat{\beta})^{\hat{\alpha}})))$, where $\hat{\alpha}$ and $\hat{\beta}$ are the ML estimates of α and β. If the regularity conditions are fulfilled for them we can write the follow covariance function for the transformed to $(0, 1)$ limit Gauss process $\xi(t)$:

$$K(t, \tau) = \min(t, \tau) - t\tau - (1/(B_{11}B_{22} - B_{12}^2))$$
$$\times(B_{22}s_1(t)s_1(\tau) - B_{12}(s_1(t)s_2(\tau) + s_2(t)s_1(\tau)) + B_{11}s_2(t)s_2(\tau)).$$

Here, $t, \tau \in (0,\ 1)$,

$$B_{11} = \int_{\mathcal{Z}} \left(\frac{z \log z\, r'(z)}{r(z)} + \log z + 1 \right)^2 r(z)dz,\ B_{22} = \int_{\mathcal{Z}} \left(\frac{z\, r'(z)}{r(z)} + 1 \right)^2 r(z)dz,$$

$$B_{12} = \int_{\mathcal{Z}} \left(\frac{z \log z\, r'(z)}{r(z)} + \log z + 1 \right) \left(\frac{z\, r'(z)}{r(z)} + 1 \right) r(z)dz$$

and

$$s_1(t) = r(R^{-1}(t))R^{-1}(t) \log(R^{-1}(t)),\ \ s_2(t) = r(R^{-1}(t))R^{-1}(t).$$

It follows from these formulae that the limit distributions of the considered statistics do not depend from the parameters α and β. Let β be known. Then the covariance function of the process $\xi(t)$is follow:

$$K(t, \tau) = \min(t, \tau) - t\tau - s_1(t)s_1(\tau)/B_{11}.$$

It does not depend of α in his turn. These results are used in the follow two sections.

3. Pareto distribution

We will consider the Pareto distribution in the form

$$F(x) = 1 - (x/\beta)^{-\alpha},\ x \geq \beta \geq 0,\ \alpha > 0.$$

For this distribution $R(z) = 1 - 1/z$ and $\mathcal{Z} = [\beta, \infty]$. It exists the supereffective unbiased estimate of β

$$\hat{\beta} = \frac{n\alpha - 1}{n\alpha} \min_{i=1,...,n} X_i.$$

We can transform the sample $X_1, ..., X_n$ to new sample $Y_1, ..., Y_n$, where $Y_i = X_i/\hat{\beta}$. The limit process $\psi(t)\xi(t)$ is equivalent to the process with $\beta = 1$. The MLE of parameter α is

$$\hat{\alpha} = n \,/ \sum_{i=1}^{n} \log\, X_i.$$

Hence the covariance function of $\xi(t)$ (without the pound function) is

$$K(t,\tau) = \min(t,\tau) - t\tau - (1-t)\log(1-t))(1-\tau)\log(1-\tau).$$

There

$$s_1(t) = -(1-t)\log(1-t), \quad B_{11} = 1.$$

This covariation function coincides with the corresponding covariance function for the exponential family

$$F(x) = 1 - \exp(-x/\beta),\ \beta \geq 0,\ x \geq 0.$$

It can be concluded that the limit distributions of the considered statistics for both families are the same one.

4. Weibull distribution

Consider the two parametric Weibull distribution family

$$F(x) = 1 - e^{-(x/\beta)^{-\alpha}},\ x \geq 0,\ \beta \geq 0,\ \alpha > 0.$$

We can note that $R(z) = 1 - e^{-z}$ and $\mathcal{Z} = [0, \infty]$. Maximum likelihood estimates $\hat{\beta}$ and $\hat{\alpha}$ for β and α can be found by numerical methods from the equation system

$$\hat{\beta} = \left(\frac{1}{n}\sum_{i=1}^{n} X_i^{\hat{\alpha}}\right)^{1/\hat{\alpha}}, \quad \frac{n}{\hat{\alpha}} + \log\left(\frac{X_1 \cdot ... \cdot X_n}{\hat{\beta}^n}\right) - \sum_{i=1}^{n}\left(\frac{X_i}{\hat{\beta}}\right)^{\hat{\alpha}} \log\left(\frac{X_i}{\hat{\beta}}\right) = 0.$$

The covariance function of $\xi(t)$ in this example has the follow elements:

$$s_1(t) = -(1-t)\log(1-t)\log(-\log(1-t)),$$

$$s_2(t) = -(1-t)\log(1-t),$$

$$B_{11}(t) = \int_0^\infty ((1-z)\log\ z - 1)^2\ e^{-z}dz = (1-C)^2 + \frac{\pi^2}{6},$$

$$B_{12}(t) = \int_0^\infty ((1-z)\log\ z - 1)(1-z)\ e^{-z}dz = 1 - C,$$

$$B_{22}(t) = \int_0^\infty (1-z)^2\ e^{-z}dz = 1,$$

$$B_{11}B_{22} - B_{12} = \pi^2/6,$$

where C is the Euler constant. It was found by simulation that the critical levels corresponding to the significance levels 0.1 and 0.05 are approximatively 0.10 and 0.12.

5. Eigenvalues

We will present briefly the method of calculation the eigenvalues of the covariance operator with the kernel (see [12])

$$K(t,\tau) = \psi(t)\psi(\tau)(\min(t,\tau) - t\tau) - (\psi(t)q^\top(t))I^{-1}(\psi(\tau)q(\tau)), \;\; t,\tau \in [0,1].$$

Let $\psi(t) = t^\alpha,\ \alpha > -2$. Then the equation for eigenvalues is as follow:

$$\det \begin{pmatrix} z_0(\gamma) & z_1(\gamma) & Z(\gamma) \\ 0 & \frac{1}{\pi}\Gamma(\nu)(\nu\sqrt{\gamma})^{-\nu} & \gamma q^\top(0) \\ J_\nu\left(2\nu\sqrt{\gamma}\right) & Y_\nu\left(2\nu\sqrt{\gamma}\right) & r^\top(\gamma,1) + \gamma q^\top(1) \end{pmatrix} = 0,$$

where

$$z_0(\gamma) = I^{-1}\int_0^1 t^{\alpha+1/2}q(t)J_\nu\left(2\nu\sqrt{\gamma}t^{1/2\nu}\right)dt,$$

$$z_1(\gamma) = I^{-1}\int_0^1 t^{\alpha+1/2}q(t)Y_\nu\left(2\nu\sqrt{\gamma}t^{1/2\nu}\right)dt,$$

$$Z(\gamma) = \int_0^1 [t^\alpha I^{-1}q(t)\rho^\top(\gamma,t) - E]dt,$$

$$r(\gamma,t) = \pi\gamma\nu\left(\sqrt{t}J_\nu\left(2\nu\sqrt{\gamma}t^{1/2\nu}\right)\int_0^t \sqrt{\tau}q''(\tau)Y_\nu\left(2\nu\sqrt{\gamma}\tau^{1/2\nu}\right)d\tau \right.$$
$$\left. -\sqrt{t}Y_\nu\left(2\nu\sqrt{\gamma}t^{1/2\nu}\right)\int_0^t \sqrt{\tau}q''(\tau)J_\nu\left(2\nu\sqrt{\gamma}\tau^{1/2\nu}\right)d\tau\right).$$

Here J_ν and Y_ν are the Bessel functions of the first and second order correspondingly. It was calculated in [12] with using this method exact tables for limit distribution of the Cramér-von-Mises statistic when the observations has the logistic distribution with one parameter.

References

[1] Beirlant, J., De Wet, T., Goegebeur, Y. (2006) A goodness-of-fit statistic for Pareto-type behaviour. *Journal of Computational and Applied Mathematics.*, **186**, 99–116.

[2] Choulakian, V. Stephens, M.A. (2001) Goodness-of-fit tests for the generalized Pareto distribution. *Technometrics.*, **43**, 478–484.

[3] Chibisov, D. M. (1965) An investigation of the asymptotic power of the test of fit. *Theory of Probability and Applications*, **10**, 421–437.

[4] Deheuvels, P., Martynov, G. (2003) Karhunen-Loève expansions for weighted Wiener processes and Brownian bridges via Bessel functions. *Progress in Probability.*, **55**, 57–93. Birkhäuser, Basel/Switzerland.

[5] Gikhman, I. I. (1954) One conception from the theory of ω^2-test. [in Ukrainian]. *Nauk. Zap. Kiiv Univ.*, **13**, 51–60.

[6] Gulati Sneh, Shapiro, S. (2008) Goodness of fit tests for the Pareto distribution. *Statistical Models and Methods for Biomedical and Technical Systems*, 263–277. Birkhäuser, Boston, (Vonta, F., Nikulin, M., Limnios, N., Huber, C., eds).

[7] Haywood, J., Khmaladze, E. (2008) On distribution-free goodness-of-fit testing of exponentiality. *Journal of Econometrics.*, **143**, 5–18.

[8] Kac, M., Kiefer, J., Wolfowitz, J. (1955) On tests of normality and other tests of goodness-of-fit based on distance methods. *Ann. Math. Statist.*, **30**, 420–447.

[9] Khmaladze, E.V. (1981) A martingale approach in the theory of parametric goodness-of-fit tests. " Theor. Prob. Appl.", **26**, 240–257.

[10] Martynov, G. V. The omega square tests. Moscow, "Nauka" , 1979, 80pp.

[11] Martynov, G. V. (1992) Statistical tests based on empirical processes and related questions. *J. Soviet. Math.*, **61**, 2195–2271.

[12] Martynov, G. V. (1994) Weighted Cramér-von-Mises test with estimated parameters. *LAD'2004: Longevty, Aging and Degradation Models*, StPeterburg, **2**, 207–222

[13] Neuhaus, G. (1974) Asymptotic properties of the Cramér-von Mises statistic when parameters are estimated. *Proc. Prague Symp. Asymptotic Stat.*, **2**, 1973, Prague, Charles Univ., 257–297

Session

Performance analysis of biological and queuing models

organized by J.R. Artalejo
(Spain)

6th St.Petersburg Workshop on Simulation (2009) 773-777

Analysis of the time to infection and disease extinction time in SIS epidemic models[1]

M.J. Lopez-Herrero[2]

Abstract

We consider an epidemic SIS model which can be viewed as a birth-death process with an absorbing state. We study the time until a non-infected individual becomes infected and also the extinction time of the epidemic or the time to absortion. Our objective is to get recursive schemes for computing their probability distributions and moments.

1. Introduction

The way that a disease spreads can give important insights to help in the fight against the disease itself. In this paper, we concentrate on two characteristics of this spread, namely the time until a tagged individual becomes infected and also the time till the end of the epidemic. The dynamics of the epidemic is described in terms of an SIS model. Our study relies on the birth-death process associated with the number of individuals of the population infected with the disease in which we are interested. The process has a unique absorbing state 0, corresponding to the end of the epidemic, while the rest of the states are transient (for details see Daley and Gani (1999) or Allen (2003)).

As far as we know, the time to infection has not been studied yet in the framework of stochastic biological models. The extincion time has attracted more attention but most of the papers concentrate only in the probability of having a finite extinction time and extinction time expectation (see, for instance Nasell (2001), Allen (2003) and Newman et al. (2004)). In the context of queueing theory the time to infection corresponds to the sojourn time given certain current state, while disease extinction time corresponds to an absorption time.

The organization of the paper is as follows. In Section 2 we introduce the model and the random variables representing the time to infection and the extinction time. We develop algorithmic schemes for their analysis, first focusing on their Laplace-Stieltjes transforms and then proceeding the study of the moments. In Section 3 we present numerical results.

[1]This work was supported by grant MTM2008-01121.

[2]University Complutense of Madrid, E-mail: lherrero@estad.ucm.es

2. Model description and analysis

An SIS epidemic model in continuous time is a closed population model of N individuals, where each individual is classified as either a susceptible or an infective. Individuals move from the susceptible to the infected group and then they recover returning to the susceptible pool. The stochastic model describing the evolution of the epidemic can be seen as a birth-death process $\{I(t) : t \geq 0\}$ with state space $S = \{0, 1, \ldots, N\}$, where $I(t)$ gives the number of infectives at time t. The birth rates, corresponding to infections, are denoted by λ_i and the death rates, corresponding to recuperations, are denoted by μ_i, $i = 0, 1, \ldots, N$. The infections are supposed to occur because of a contagious disease. Hence, when there are no infective, the process stays there for ever. The other states are assumed transient. More specifically we assume that $\lambda_0 = \lambda_N = \mu_0 = 0$, while $\mu_1, \mu_2, \ldots, \mu_N$ and $\lambda_1, \lambda_2, \ldots, \lambda_{N-1}$ are strictly positive. In the classical SIS model, it is assumed that $\lambda_i = \beta i(N - i)/N$ and $\mu_i = \gamma i$, where β is the contact rate and γ is the recovery rate per individual. However, in the present paper we will present the results for general birth and death rates. Indeed in several biological systems the data do not support the above classical rates.

3. Time to infection

Let us consider the population at an arbitrary time t and suppose that there are i, $0 \leq i \leq N - 1$, infected individuals at this time. We mark one of the $(N - i)$ non infected individuals and denote by S_i a random variable representing the time until the selected individual gets infected. Obviously, $S_0 = +\infty$ and the case $i = N$ has no sense. To study the variables S_i, $0 \leq i \leq N - 1$, we define

$$\begin{aligned} v_i &= P\{S_i < \infty\},\ 0 \leq i \leq N-1, \\ \psi_i(s) &= E[e^{-sS_i}\, \mathbf{1}_{\{S_i<\infty\}}],\ 0 \leq i \leq N-1,\ Re(s) \geq 0, \\ M_i^k &= E[S_i^k\, \mathbf{1}_{\{S_i<\infty\}}],\ 0 \leq i \leq N-1,\ k \geq 0, \end{aligned}$$

where $\mathbf{1}_{\{S_i<\infty\}}$ is the indicator random variable that takes the value 1 when the event $\{S_i < \infty\}$ occurs and is 0 otherwise.

Note that, the probabilities v_i are strictly between 0 and 1, for $1 \leq i \leq N - 1$. Indeed, $P\{S_i < \infty\} \geq \frac{\lambda_i}{\lambda_i+\mu_i} \cdots \frac{\lambda_{N-1}}{\lambda_{N-1}+\mu_{N-1}} > 0$ and $P\{S_i = +\infty\} \geq \frac{\mu_i}{\lambda_i+\mu_i} \cdots \frac{\mu_1}{\lambda_1+\mu_1} > 0$, hence $P\{S_i < \infty\} < 1$.

A first step argument conditioning the next infection, provides that the Laplace-Stieltjes transforms, $\psi_i(s)$, satisfy the following set of equations:

$$\psi_0(s) = 0, \tag{1}$$

$$\begin{aligned} \psi_i(s) &= \frac{\mu_i}{s+\lambda_i+\mu_i}\psi_{i-1}(s) + \frac{\lambda_i}{s+\lambda_i+\mu_i}\frac{N-i-1}{N-i}\psi_{i+1}(s) \\ &+ \frac{\lambda_i}{s+\lambda_i+\mu_i}\frac{1}{N-i},\ 1 \leq i \leq N-1. \end{aligned} \tag{2}$$

For every s the system of equations (1)-(2) is tri-diagonal and the coefficient matrix is strict diagonally dominant. We can solve the system by using a standard

forward-elimination-backward-substitution method. After some algebra, we obtain a stable recursive scheme which appears in the following theorem.

Theorem 1. *The Laplace-Stieltjes transforms $\psi_i(s)$, for $1 \leq i \leq N-1$, are computed by the equations*

$$\psi_{N-1}(s) = \frac{2\left(\lambda_{N-1}(s+g_{N-2}+\lambda_{N-2})+\mu_{N-1}D_{N-2}\right)}{2(s+\lambda_{N-1})(s+g_{N-2}+\lambda_{N-2})+\mu_{N-1}(2(s+g_{N-2})+\lambda_{N-2})}, \quad (3)$$

$$\begin{aligned}\psi_i(s) &= \sum_{k=i}^{N-2} \frac{D_k}{\lambda_k(\frac{N-k-1}{N-k})} \prod_{n=i}^{k} \frac{\lambda_n}{s+g_n+\lambda_n} \frac{N-n-1}{N-n} \\ &+ \psi_{N-1}(s) \prod_{k=i}^{N-2} \frac{\lambda_k}{s+g_k+\lambda_k} \frac{N-k-1}{N-k}, \quad 1 \leq i \leq N-2, \end{aligned} \quad (4)$$

where the coefficients g_i and D_i, for $1 \leq i \leq N-2$ are given by the recursive scheme

$$g_1 = \mu_1, \quad (5)$$

$$g_i = \mu_i \frac{s+g_{i-1}+\frac{\lambda_{i-1}}{N-i+1}}{s+g_{i-1}+\lambda_{i-1}}, \; 2 \leq i \leq N-2, \quad (6)$$

$$D_1 = \frac{\lambda_1}{N-1}, \quad (7)$$

$$D_i = \frac{\lambda_i}{N-i} + \frac{\mu_i D_{i-1}}{s+g_{i-1}+\lambda_{i-1}}, \; 2 \leq i \leq N-2. \quad (8)$$

Observe that $v_i = P\{S_i < \infty\} = \psi_i(0)$. Consequently, the set of probabilities v_i, $0 \leq i \leq N-1$, can be determined by putting $s=0$ in (3)-(8).

We now concentrate on the calculation of the moments $M_i^k = E[S_i^k \, \mathbf{1}_{\{S_i<\infty\}}]$, for $0 \leq i \leq N-1$. By differentiating k times equations (1)-(2) and evaluating at $s=0$, we find that

$$M_0^k = 0, \; k \geq 1, \quad (9)$$

$$(\lambda_i+\mu_i)\, M_i^k = \mu_i M_{i-1}^k + \lambda_i \frac{N-i-1}{N-i} M_{i+1}^k + k M_i^{k-1}, \; 1 \leq i \leq N-1, \; k \geq 1.$$

This system of equations provides, after some algebra, a stable recursive scheme for the computation of M_i^k, for $1 \leq i \leq N-1$, $k \geq 1$, having a similar structure as the one appearing on Theorem 1. Each iteration allows us to compute the unknowns moments of order k, in terms of the moments of one order less. Note that the moments of order $k=0$ are v_i. Hence, from them we obtain the moments for $k=1$ by solving the system (9), then we proceed for $k=2$, and so on.

4. Extinction time of a disease

Let us assume that at the initial time $t = 0$ the population has i infective individuals, and define a continuous random variable L_i to be the extinction time of the epidemic given the current population state. This variable can be seen as the absorption time by the state 0 given that $I(0) = i$. In a more general framework of CTMC on finite state space with rate probability matrix Q, the random variable L, the unconditional version of the absorption time, satisfies the following results (see Kulkarni (2005) and Latouche and Ramaswami (1999)):

- $P\{L \leq x\} = 1 - \alpha \exp\{Mx\}\mathbf{e}$,
- If M is invertible, L is finite with probability 1, $\varphi(s) = E[e^{-sL}] = -\alpha(sI - M)^{-1}M\mathbf{e}$ and $E[L^k] = k!\alpha(-M^{-1})^k\mathbf{e}$, $k \geq 1$,

where M is the submatrix of Q corresponding to the transient states $\{1, 2, \ldots, N\}$, α is row vector of dimension N containing the initial probabilities and $\mathbf{e}$ is a column vector with all coordinates equal to one.

Next we introduce some notation for absorption probabilities, Laplace-Stieltjes transforms and moments of L_i, for $0 \leq i \leq N$. Define

$$\begin{aligned} u_i &= P\{L_i < \infty\},\ 0 \leq i \leq N, \\ \varphi_i(s) &= E[e^{-sL_i}],\ 0 \leq i \leq N,\ Re(s) \geq 0 \\ \widetilde{M}_i^k &= E[L_i^k],\ 0 \leq i \leq N,\ k \geq 0. \end{aligned}$$

Coming back to the context of a birth-death process, when dealing with a finite population, M is an irreducible diagonally dominant matrix. Therefore, M is invertible and it is possible to analyze the behavior of $\varphi_i(s)$ and $\widetilde{M}_i^k$ in terms of the results for the unconditional absorption time, L. But the computation of the previous formulas implies to deal with powers and inverses of matrices having positive and negative entries. So the objective is to get stable recursive schemes for the computation of $\varphi_i(s)$ and $\widetilde{M}_i^k$ avoiding subtractions.

Using a first step analysis, we obtain that the functions $\varphi_i(s)$, for $0 \leq i \leq N$, satisfy

$$\varphi_0(s) = 1, \tag{10}$$

$$\varphi_i(s) = \frac{\mu_i}{s + \lambda_i + \mu_i}\varphi_{i-1}(s) + \frac{\lambda_i}{s + \lambda_i + \mu_i}\varphi_{i+1}(s),\ 1 \leq i \leq N. \tag{11}$$

We use a forward-elimination-backward-substitution algorithm to solve the system of equations (10)-(11). After some algebra, we get a stable recursive scheme from which the computation of Laplace-Stieltjes transforms can be done at a low computational cost. Theorem 2 summarizes this scheme.

Theorem 2. *The Laplace-Stieltjes transforms $\varphi_i(s)$, for $1 \leq i \leq N$, are computed*

by the equations

$$\varphi_N(s) = \frac{\mu_N D_{N-1}}{s\lambda_{N-1} + (s+\mu_N)(s+g_{N-1})}, \tag{12}$$

$$\varphi_i(s) = \sum_{k=i}^{N-1} \frac{D_k}{\lambda_k} \prod_{n=i}^{k} \frac{\lambda_n}{s+g_n+\lambda_n} + \varphi_N(s) \prod_{k=i}^{N-1} \frac{\lambda_k}{s+g_k+\lambda_k},\ 1 \le i \le N-1, \tag{13}$$

where the coefficients g_i and D_i, for $1 \le i \le N-1$ are given by the recursive scheme

$$g_1 = \mu_1, \tag{14}$$

$$g_i = \mu_i \frac{s+g_{i-1}}{s+g_{i-1}+\lambda_{i-1}},\ 2 \le i \le N-1, \tag{15}$$

$$D_1 = \mu_1, \tag{16}$$

$$D_i = \frac{\mu_i D_{i-1}}{s+g_{i-1}+\lambda_{i-1}},\ 2 \le i \le N-1. \tag{17}$$

Next we focus on the calculation of the moments $\{\widetilde{M}_i^k, 0 \le i \le N\}$, for any arbitrary non negative integer k. By differentiating equations (10)-(11) k times and evaluating at $s=0$, for $k \ge 1$, we find that

$$\begin{aligned} \widetilde{M}_0^k &= 0, \\ (\lambda_i+\mu_i)\,\widetilde{M}_i^k &= \mu_i \widetilde{M}_{i-1}^k + \lambda_i \widetilde{M}_{i+1}^k + k\widetilde{M}_i^{k-1},\ 1 \le i \le N. \end{aligned}$$

After some algebra, we can compute the kth moments of $\{L_i, 1 \le i \le N\}$ from the $(k-1)$th moments of the conditional absorption times, via a stable recursive scheme. Note that moments of order $k=0$ are $\widetilde{M}_i^0 = u_i = 1$, for $1 \le i \le N$ because $\{1, 2, \ldots, N\}$ is a non-decomposable set of states. Moreover, it is trivial that $u_0 = 1$. Therefore, starting from moments of order 0, we get the expected values for $k=1$ and we proceed for $k=2$, and so on.

5. Numerical illustration

In this section we present numerical results related to S_i and L_i. We consider a population having $N=30$ individuals, where the evolution on the number of infected individuals is represented by an SIS epidemic model. Birth and death rates, λ_i and μ_i, for $0 \le i \le N$, are defined as follows: $\lambda_i = \beta i(N-i)/N$ and $\mu_i = \gamma i$.

We want to point out that the recursive schemes apperaring in Theorems 1 and 2 provide, via numerical inversion, the probability distributions for S_i and L_i. This inversion can be efficiently addressed by using the EULER and POST-WIDDER algorithms (Abate and Whitt (1995)). However, due to lack of space,

Table 1: Characteristics for a population having one infected individual

$P\{S_1 < \infty\}$ $E[L_1]$	$\beta = 0.05$	$\beta = 0.5$	$\beta = 1.0$	$\beta = 5.0$	$\beta = 10.0$
$\gamma = 0.5$	0.00366 2.10311	0.12346 4.82039	0.48298 331.6434	0.89655 1.92×10^{17}	0.94827 2.17×10^{25}
$\gamma = 1.0$	0.00174 1.02494	0.02897 1.35564	0.12346 2.41019	0.79310 4.07×10^{9}	0.89655 9.62×10^{16}
$\gamma = 2.0$	0.00085 0.50613	0.01068 0.57173	0.02897 0.67782	0.58621 2113.9605	0.79310 2.03×10^{9}
$\gamma = 5.0$	0.00033 0.20097	0.00366 0.21031	0.00810 0.22211	0.12346 0.48203	0.48298 33.16434

we will display on Table 1 results only for the probabilities $P\{S_1 < \infty\}$ and $E[L_1]$, varying contact rate, β, and recovery rate per individual, γ.

We can observe that both quantities present an increasing behavior for increasing contact rates. However, when the recovery rate per individual increases we observe a decreasing behavior. It is not reported on the table but, numerical experiments show that both characteristics have an increasing behavior when the initial number of infective individuals in the population increases.

References

[1] Abate J. and Whitt W. (1995) Numerical inversion of Laplace transforms of probability distributions. *ORSA J. Comp.* **7**, 36-43.

[2] Allen L.J.S. (2003) *An Introduction to Stochastic Processes with Applications to Biology.* Prentice-Hall, New Jersey.

[3] Daley D.J. and Gani J. (1999) *Epidemic Modelling: An Introduction.* Cambridge University Press, Cambridge.

[4] Kulkarni V.G. (1995) *Modeling and Analysis of Stochastic Systems.* Chapman and Hall/CRC, Boca Raton.

[5] Latouche G. and Ramaswami V. (1999) *Introduction to Matrix Analytic Methods in Stochastic Modeling.* ASIA-SIAM Series on Statistics and Applied Probability. SIAM, Philadelphia.

[6] Nasell I. (2001) Extinction and quasi-stationarity in the Verhulst logistic model. *J. Theor. Biol.* **211**, 11-27.

[7] Newman T.J., Ferdy J.B. and Quince, C. (2004) Extinction times and moment closure in the stochastic logistic process. *Theor. Popul. Biol.* **65**, 115-126.

6th St.Petersburg Workshop on Simulation (2009) 779-783

Number of removals in the stochastic SIR model[1]

Jesus Artalejo[2]

Abstract

Algorithmic methods for the stochastic Susceptible-Infected-Removed (SIR) epidemic model are examined. We propose simple and efficient methods for the distribution of the number of removals. We investigate this descriptor both until the extinction of the epidemic and in transient regime. The methodology can also be used to study other stochastic epidemic models.

1. Introduction and model description

The existing literature on stochastic epidemic models is mainly focused on the extinction time and the quasi-stationary distribution of the epidemic size. In this paper, we concentrate on a discrete counterpart of the extinction time: the total number of individuals removed before the extinction (Artalejo et al., 2007). This descriptor has been largely ignored despite of the fact that it is helpful to understand the size of the epidemic.

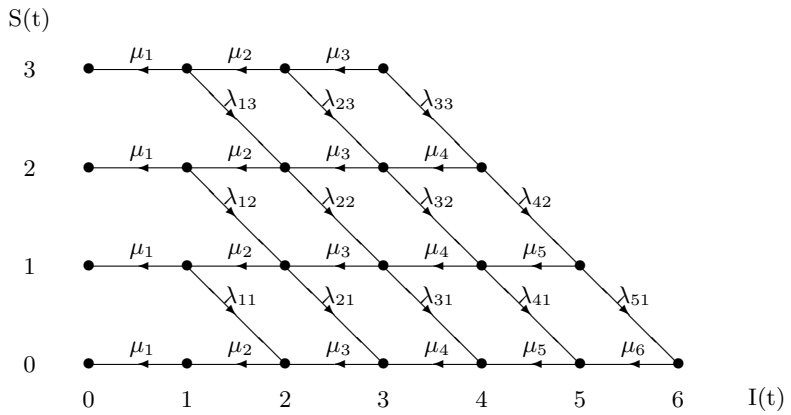


Figure 1. States and transitions of the SIR epidemic model

[1]This work was supported by grant MTM2008-01121.

[2]Complutense University of Madrid, E-mail: jesus_artalejo@mat.ucm.es

The model we consider is the stochastic SIR epidemic model. Let $\{(I(t), S(t)); t \geq 0\}$ be the bidimensional continuous-time Markov chain describing the model. At time t, the population consists of $I(t)$ infectives and $S(t)$ susceptibles. The initial condition is $(I(0), S(0)) = (m, n)$. When in state (i, j), for $i \geq 1$, the population state either moves to $(i+1, j-1)$ at rate λ_{ij} ($\lambda_{i0} = 0$) due to an infection, or to $(i-1, j)$ at rate μ_i ($\mu_0 = 0$) due to the removal of an infective. The state space of the SIR epidemic model is $S = \{(i, j);\ 0 \leq i \leq m+n,\ 0 \leq j \leq \min\{n, m+n-i\}\}$. We notice that states $(0, j)$, for $0 \leq j \leq n$, are absorbing states. Thus, the transition from $(0, j)$ to $(1, j-1)$ is not allowed. In other words, there are not replications of the epidemic. We assume that $m \geq 1$. A typical choice for the transition rates is $\lambda_{ij} = \frac{\beta}{N} ij$ and $\mu_i = \gamma i$, where N is the total population size (Allen, 2003). Figure 1 illustrates the state space and transitions for the case $(m, n) = (3, 3)$.

We are interested in the distribution of the number of individuals removed until the extinction of the epidemic, N_{mn}^R. We may also study the number of individuals infected before the extinction, N_{mn}^I. Since $N_{mn}^R = N_{mn}^I + m$, it is clear that we may reduce our study to N_{mn}^R. In what follows, we simply denote N_{mn}^R as N_{mn}. We also present an alternative transient analysis which is referred to any time interval $[0, t]$ rather than to the extinction time.

Our algorithmic solutions can be easily adapted to other stochastic epidemic models; for instance, the SIS epidemic model (Allen, 2003) and models with killing and catastrophes (Coolen-Schrijner and van Doorn, 2006).

The rest of the paper is organized as follows. In Section 2, we present an efficient algorithmic analysis for computing the distribution of N_{mn}. The corresponding transient distribution is investigated in Section 3. Finally, in Section 4, we present a numerical example.

2. Removals until the extinction of the epidemic

In this section we investigate the distribution of the number of removals before extinction of the epidemic. To this end, we first introduce some notation for generating functions and factorial moments:

$$\phi_{ij}(z) = E\left[z^{N_{ij}}\right],\ |z| \leq 1,\ (i, j) \in S,$$
$$M_{ij}^0 = P\{N_{ij} < \infty\} = 1,\ (i, j) \in S,$$
$$M_{ij}^k = E\left[N_{ij}(N_{ij}-1)\cdots(N_{ij}-k+1)\right],\ k \geq 1,\ (i, j) \in S.$$

A first-step argument yields the following system governing the generating functions $\phi_{ij}(z)$:

$$\phi_{0j}(z) = 1,\ 0 \leq j \leq n, \tag{1}$$
$$\phi_{ij}(z) = \frac{z\mu_i}{\lambda_{ij}+\mu_i}\phi_{i-1,j}(z) + \frac{\lambda_{ij}}{\lambda_{ij}+\mu_i}\phi_{i+1,j-1}(z),$$
$$0 \leq j \leq n,\ 1 \leq i \leq m+n-j. \tag{2}$$

For $j = 0$, we find that $\phi_{i0}(z) = z^i$, for $1 \leq i \leq m+n$. Then, for each fixed $j \in \{1, ..., n\}$, $\phi_{ij}(z)$ can be recursively computed from equations (1) and (2). Moreover, for $1 \leq i \leq m+n$, we have

$$M_{i0}^k = \begin{cases} 1, & \text{if } k = 0, \\ i(i-1)\cdots(i-k+1), & \text{if } 1 \leq k \leq i, \\ 0, & \text{if } k > i. \end{cases} \tag{3}$$

On the other hand, after appropriate differentiation over (1) and (2), we obtain

$$\begin{gathered} M_{ij}^0 = 1,\ (i,j) \in S, \\ M_{0j}^k = 0,\ k \geq 1,\ 0 \leq j \leq n, \\ M_{ij}^k = \frac{\mu_i}{\lambda_{ij}+\mu_i} M_{i-1,j}^k + \frac{\lambda_{ij}}{\lambda_{ij}+\mu_i} M_{i+1,j-1}^k + \frac{k\mu_i}{\lambda_{ij}+\mu_i} M_{i-1,j}^{k-1}, \\ k \geq 1,\ 1 \leq j \leq n,\ 1 \leq i \leq m+n-j. \end{gathered} \tag{4}$$

Formulas (3) and (4) provide an efficient recursive scheme for the computation of M_{ij}^k in the order $k \geq 0$, $0 \leq j \leq n$ and $0 \leq i \leq m+n-j$.

Our next objective is the computation of the probabilities

$$x_{ij}^k \equiv P\left\{N_{ij} = k\right\},\ (i,j) \in S,\ i \leq k \leq i+j.$$

First, we point out some trivial relationships

$$x_{0j}^k = \delta_{k0},\ 0 \leq j \leq n, \tag{5}$$

$$x_{ij}^0 = 0,\ 0 \leq j \leq n,\ 1 \leq i \leq m+n-j, \tag{6}$$

$$x_{i0}^k = \delta_{ki},\ 1 \leq i \leq m+n, \tag{7}$$

where δ_{ab} denotes Kronecker's 0-1 delta.

Now the fundamental equation governing x_{ij}^k is

$$\begin{gathered} x_{ij}^k = \frac{\mu_i}{\lambda_{ij}+\mu_i} x_{i-1,j}^{k-1} + \frac{\lambda_{ij}}{\lambda_{ij}+\mu_i} x_{i+1,j-1}^k, \\ k \geq 1,\ 1 \leq j \leq n,\ 1 \leq i \leq m+n-j. \end{gathered} \tag{8}$$

For each fixed $k \in \{1, ..., m+n\}$, the above equation (8) and the boundary conditions (5)-(7) can be combined to compute recursively x_{ij}^k in the natural order $0 \leq j \leq n$ and $0 \leq i \leq m+n-j$. In particular, we are interested in the sequence $\left\{x_{mn}^k;\ m \leq k \leq m+n\right\}$ associated to the initial state (m,n). The use of this recursive method avoids the numerical inversion of $\phi_{ij}(z)$.

2. Transient analysis

Our aim in this section is to deal with the transient analysis of the number of removals occurring in any time interval $[0,t]$, $N^R(t)$. The transient behavior provides updated information about the state of the infection at each time t. We

will concentrate on the study of $N^R(t)$ but the same methodology remains valid for the number of infected individuals in $[0,t]$.

We complete the epidemic model by adding the counting component $N^R(t)$; that is, we consider the transient probabilities

$$p_{ijk}^R(t) = P\left\{I(t)=i,\ S(t)=j,\ N^R(t)=k\right\},$$

for each $t \geq 0$, $(i,j) \in S$, $k_0 \leq k \leq k_0+m+n$, and the initial condition $\left(I(0),S(0),N^R(0)\right)=(m,n,k_0)$, for $m \geq 1$, $n \geq 0$, and $k_0 \geq 0$. The case $I(0)=0$ is trivial because then $p_{0nk_0}^R(t)=1$, for all $t \geq 0$.

Let $\widetilde{p}_{ijk}^R(s)$ be the Laplace transform of the probability $p_{ijk}^R(t)$; that is, $\widetilde{p}_{ijk}^R(s) = \int_0^\infty e^{-st}p_{ijk}^R(t)dt$, for $Re(s) \geq 0$.

Following the transform method, we next write down the Kolmogorov equations governing the dynamics of probabilities $p_{ijk}^R(t)$. They are as follows:

$$\begin{gathered}\frac{d}{dt}p_{ijk}^R(t) = -\left(\lambda_{ij}+\mu_i\right)p_{ijk}^R(t)\\ +(1-\delta_{i+j,m+n})(1-\delta_{kk_0})\mu_{i+1}p_{i+1,j,k-1}^R(t)\\ +(1-\delta_{i0})(1-\delta_{i1})(1-\delta_{jn})\lambda_{i-1,j+1}p_{i-1,j+1,k}^R(t),\\ (i,j)\in S,\ k_0 \leq k \leq k_0+m+n. \end{gathered} \tag{9}$$

Then, after taking Laplace transforms on (9), we find that

$$\begin{gathered}\left(s+\lambda_{ij}+\mu_i\right)\widetilde{p}_{ijk_0}^R(s) = \delta_{(i,j)(m,n)}\\ +(1-\delta_{i0})(1-\delta_{i1})(1-\delta_{jn})\lambda_{i-1,j+1}\widetilde{p}_{i-1,j+1,k_0}^R,\\ (i,j)\in S, \end{gathered} \tag{10}$$

$$\begin{gathered}\left(s+\lambda_{ij}+\mu_i\right)\widetilde{p}_{ijk}^R(s) = (1-\delta_{i+j,m+n})\mu_{i+1}\widetilde{p}_{i+1,j,k-1}^R(s)\\ +(1-\delta_{i0})(1-\delta_{i1})(1-\delta_{jn})\lambda_{i-1,j+1}\widetilde{p}_{i-1,j+1,k}^R(s),\\ k_0+1 \leq k \leq k_0+m+n,\ (i,j)\in S. \end{gathered} \tag{11}$$

Equations (10) and (11) can be solved in ascending order for k and descending order for indices j and i. Once the Laplace transforms have been computed, the marginal distribution $P\left\{N^R(t)=k\right\}$, for $k_0 \leq k \leq k_0+m+n$, can be obtained by inverting numerically $\sum_{j=0}^{n}\sum_{i=0}^{m+n-j}\widetilde{p}_{ijk}^R(s)$ (Cohen, 2007).

Finally, we turn our attention to the moments of $N^R(t)$. Define

$$\begin{aligned} m_{ij}^{R,p}(t) &= \sum_{k=k_0}^{k_0+m+n} k^p p_{ijk}^R(t),\ (i,j)\in S,\ p \geq 0,\\ \widetilde{m}_{ij}^{R,p}(s) &= \int_0^\infty e^{-st}m_{ij}^{R,p}(t)dt,\ (i,j)\in S,\ p \geq 0. \end{aligned}$$

For $p=0$ we obtain the transient probabilities $p_{ij}(t) = P\left\{I(t)=i,\ S(t)=j\right\}$, for $(i,j)\in S$. The initial conditions are $m_{ij}^{R,p}(0) = k_0^p\delta_{(i,j)(m,n)}$, for $(i,j)\in S$.

Summing over k on formula (9), using

$$\sum_{k=k_0+1}^{k_0+m+n} k^p p^R_{i+1,j,k-1}(t) = m^{R,p}_{i+1,j}(t) + (1-\delta_{p0}) \sum_{l=0}^{p-1} \binom{p}{l} m^{R,l}_{i+1,j}(t),$$

and taking Laplace transforms in the resulting equation, we get

$$\begin{aligned}
(s+\lambda_{ij}+\mu_i)\,\widetilde{m}^{R,p}_{ij}(s) &= k_0^p \delta_{(i,j)(m,n)} \\
+(1-\delta_{i+j,m+n})\mu_{i+1}(1-\delta_{p0}) &\sum_{l=0}^{p-1} \binom{p}{l} \widetilde{m}^{R,l}_{i+1,j}(s) \\
+(1-\delta_{i0})(1-\delta_{i1})(1-\delta_{jn})\lambda_{i-1,j+1}&\widetilde{m}^{R,p}_{i-1,j+1}(s) \\
+(1-\delta_{i+j,m+n})\mu_{i+1}&\widetilde{m}^{R,p}_{i+1,j}(s), \\
(i,j) \in S,\ p \geq 0.& \qquad (12)
\end{aligned}$$

Formula (12) provides the key for the recursive computation of $\widetilde{m}^{R,p}_{ij}(s)$ in the order $p \geq 0$, $j = n, ..., 0$ and $i = m+n-j, ..., 0$. As a result, the moments $m^{R,p}_{ij}(t)$ and $m^R_p(t) = \sum_{j=0}^{n} \sum_{i=0}^{m+n-j} m^{R,p}_{ij}(t)$, $p \geq 0$, are obtained by numerical inversion.

4. Numerical example

In order to evaluate the performance of the expected number of removals until the extinction of the epidemic, we next present a numerical example. A more detailed study including transient numerical results will be performed in an extended version of this paper.

Table 1: The expected number of removals until extinction of the epidemic

M^1_{mn}	$\beta = 0.05$	$\beta = 0.5$	$\beta = 1.0$	$\beta = 5.0$	$\beta = 10.0$
$\gamma = 0.5$	29.09199	29.61360	29.84612	29.99976	29.99999
	15.76390	22.63238	27.13318	29.99639	29.99999
	1.10610	3.89951	11.69815	26.97935	28.49683
$\gamma = 1.0$	29.04714	29.38081	29.61360	29.98855	29.99976
	15.37861	18.96128	22.63238	29.81723	29.99639
	1.05059	1.81790	3.89951	23.63078	26.97935
$\gamma = 2.0$	29.02386	29.21389	29.38081	29.90184	29.98855
	15.18842	16.95013	18.96128	28.23280	29.81723
	1.02472	1.30873	1.81790	15.34653	23.63078
$\gamma = 5.0$	29.00961	29.09199	29.17526	29.61360	29.84612
	15.07514	15.76390	16.55071	22.63238	27.13318
	1.00975	1.10610	1.23443	3.89951	11.69815

We display M^1_{mn} for different choices of β (infection rate) and γ (recovery rate). The population size is $N = 30$. Each cell is associated to a pair (β, γ)

and gives, from top to bottom, the value of M^1_{mn} for the initial states $(m,n) \in \{(29,1),(15,15),(1,29)\}$.

The main conclusion inferred from the Table is that M^1_{mn} increases as a function of β but it is a decreasing function of γ. As is to be expected the initial state (m,n) has a strong influence. In fact, the value of M^1_{mn} varies from m to $m+n$.

References

[1] Allen L.J.S. (2003) *An Introduction to Stochastic Processes with Applications to Biology.* Prentice-Hall, New Jersey.

[2] Artalejo J.R., Economou A. and Lopez-Herrero A. (2007) Evaluating growth measures in an immigration process subject to binomial and geometric catastrophes. *Math. Biosci. Eng.* **4**, 573-594.

[3] Cohen, A.M. (2007) *Numerical Methods for Laplace Transform Inversion.* Springer, New York.

[4] Coolen-Schrijner, P. and van Doorn E.A. (2006) Quasi-stationary distributions for a class of discrete-time Markov chains. *Comput. Appl. Probab.* **8**, 449-465.

6th St.Petersburg Workshop on Simulation (2009) 785-789

A sufficient condition for stability of fluid limit models[1]

Rosario Delgado[2]

Abstract

We prove that in the subcritical case, a fluid limit model is stable if *state space collapse* with a "lifting" matrix that verifies a restriction holds.

1. Introduction

We consider a fluid model which consists of J stations, with a single server and an infinite-capacity buffer at each one, that process K different fluid classes, with $K \geq J \geq 1$. Each fluid class can be processed at only one station and feedback is allowed. We assume a work-conserving service discipline. This fluid model can be considered as the fluid approximation of an associated queueing network that works under any head-of-the-line work-conserving service discipline and with inter-arrival and service times not necessarily exponential. It is known that the stability of the queueing network (the *positive Harris recurrence* of the underlying Markov process describing the network dynamics) is ensured if the fluid model is stable (see Theorem 4.2 [2]). *Stability* of a fluid limit model means that the queue process reaches zero in finite time and stays there regardless of the initial fluid levels. It is known that sub-criticality (traffic intensity strictly less than one at each station) is not a sufficient condition, although necessary, for stability.

In this work we establish a sufficient condition for the stability of the fluid limit model (in the subcritical case): it is a kind of *state space collapse* assumption with a "lifting" matrix that verifies a technical restriction. *State space collapse* condition has turned out to be a key ingredient in the proof of *heavy-traffic* limits for multi-class queueing networks in the light-tailed as well as in the heavy-tailed environment. See for instance [5], [4], [1], [7] and [3]. As far as we know, this is the first time that this kind of condition has been related with the study of stability (*light-traffic*).

[1]This work was supported by grant MEC-FEDER ref. MTM2006-06427 and 2005SGR-01043

[2]Universitat Autònoma de Barcelona (Spain), E-mail: delgado@mat.uab.cat

2. The fluid limit model

Consider a fluid model consisting of $J \geq 1$ single-server stations with a single server and an infinite-capacity buffer at each one. There are K fluid classes with $K \geq J$, each one processed at only one station (but at each station more than one fluid class can be processed), $s(k)$ being the station where class k fluid is processed, and $s^{-1}(j)$ the set of fluid classes served at station j. We introduce the $J \times K$ (deterministic) *constituency matrix* $C = \left(C_{jk}\right)_{j,k}$ by $C_{jk} = 1$ if $j = s(k)$ and 0 otherwise.

Let $\alpha_k \geq 0$ be the exogenous inflow rate and $\mu_k > 0$ be the potential outflow rate, for class k fluid, and define $m_k \stackrel{\text{def}}{=} \frac{1}{\mu_k}$ and matrix $M \stackrel{\text{def}}{=} diag(m_1, \ldots, m_k)$. Upon being processed at station $s(k)$, a proportion $P_{k\ell}$ of class k fluid leaving station $s(k)$ goes next to station $s(\ell)$ to be processed there as a class ℓ fluid. The *"flow-transfer" matrix* $P = (P_{k\ell})_{k,\ell=1}^{K}$ is assumed to be sub-stochastic and to have spectral radius strictly less than one. Hence, $Q \stackrel{\text{def}}{=} \left(I - P^T\right)^{-1}$ is well defined. We assume that fluid at each station is processed following a *work-conserving* service discipline by arrival order into each class.

The fluid model is described by elements $\alpha = (\alpha_1, \ldots, \alpha_k)^T$, M, C, P and $z = (z_1, \ldots, z_K)^T \geq 0$, z_k being the initial amount of class k fluid in the system. We refer to it by (α, M, C, P, z).

We define λ to be the unique $K-$dimensional vector solution to the *traffic equation* $\lambda = \alpha + P^T \lambda$, that is, $\lambda = Q\,\alpha$, and introduce the *fluid traffic intensity* at station j as $\rho_j \stackrel{\text{def}}{=} \sum_{k \in s^{-1}(j)} \lambda_k m_k$ (in matrix form, $\rho = C\,M\,\lambda$). We will assume throughout the paper that $\rho < e$, with $e = (1, \ldots, 1)^T$ (sub-criticality).

Processes A, D, T, Z, W and Y will be used to measure the performance of the fluid network: $A(t)$ is the cumulative amount of fluid arrived (from outside and by feedback) by time t (to each fluid class) and $D(t)$ is the cumulative amount of fluid departing from each class (to other classes or to outside). $T(t)$ is the cumulative amount of processing time spent on each fluid class by time t. $Z(t)$ is the amount of fluid of any class in the system at time t. All the above processes are $K-$dimensional and the rest are $J-$dimensional: $W(t)$ denotes the workload or amount of time required by any server to complete processing of all fluid in queue, at time t, and $Y(t)$ is the cumulative amount of time that the server at each station has been idle in the interval $[0, t]$. By definition, T and Y are nondecreasing processes which depend on the specific service discipline, and $A(0) = D(0) = T(0) = Y(0) = 0$.

These processes are related by means of the following *fluid model equations*:

$$A(t) = \alpha\, t + P^T\, D(t)\,, \tag{1}$$

$$Z(t) = z + A(t) - D(t) = z + \alpha\, t - (I - P^T)\, M^{-1}\, T(t)\,, \tag{2}$$

$$D(t) = M^{-1}\, T(t)\,, \tag{3}$$

$$C\,T(t) + Y(t) = e\,t\,, \tag{4}$$

$$\int_0^\infty W_j(t)\,dY_j(t) = 0 \quad \text{for all } j = 1, \ldots, J\,, \tag{5}$$

$$W(t) = C\,M\,(z + A(t)) - C\,T(t)\,, \tag{6}$$

Note that equation (5) expresses that for any station j, idle time Y_j can only increase when workload W_j is zero, that is exactly the meaning of a work-conserving discipline.

Let $\Psi(\cdot) \stackrel{\text{def}}{=} \big(A(\cdot), D(\cdot), T(\cdot), Z(\cdot), W(\cdot), Y(\cdot)\big)$ be any solution of the fluid model equations (1)-(6), which may not have in general a unique solution.

Definition 1 (Stability of the fluid limit model)**.** *We say that the fluid limit model* (α, M, C, P, z) *is stable if there exists* $t_0 > 0$ *such that for any solution* $\Psi(\cdot)$ *of the fluid model equations,* $Z(t) = 0 \;\; \forall t \geq t_0\,|z|$ *, where* $|z| \stackrel{\text{def}}{=} \sum_{k=1}^{K} z_k$ *.*

3. The main result

Note that from (6), (2) and (3) we can express the workload in terms of the queue process by means of

$$W(t) = C\,M\,\big(z + A(t) - M^{-1}\,T(t)\big) = C\,M\,Z(t)\,, \tag{7}$$

that is, for any j, $W_j(t) = \sum_{k \in s^{-1}(j)} m_k\,Z_k(t)$, which expresses workload at station j in terms of fluid amount for each class processed at that station. Next definition introduces a condition establishing that Z, in its turn, can be expressed in terms of W by means of a *"lifting"* deterministic matrix.

Definition 2. *Given a solution* $\Psi(\cdot)$ *of the fluid model equations associated to a fluid limit model* (α, M, C, P, z)*, we say that the fluid limit model satisfies* **state space collapse** *with "lifting" matrix* Δ *if*

$$\boxed{Z = \Delta\,W}$$

where $\Delta = \big(\Delta_{kj}\big)_{kj}$ *with* $\Delta_{kj} = \delta_k > 0$ *if* $k \in s^{-1}(j)$ *and* 0 *otherwise. And we say that the "lifting" matrix* Δ *is* **regular** *if accomplishes the following technical restriction:* $C\,M\,Q\,\Delta$ *is invertible and matrix* R *defined by* $R \stackrel{\text{def}}{=} \big(C\,M\,Q\,\Delta\big)^{-1}$ *verifies assumption* **(HR)***:* R *can be expressed as* $I + \Theta$*, with* Θ *a square matrix such that the matrix obtained from* Θ *by replacing its elements by their absolute values, has spectral radius strictly less than* 1*.*

Roughly speaking, *state space collapse* assumption expresses that any fluid class k contributes a fixed portion δ_k to the workload at station $s(k)$. That is, the fluid classes processed at the same station are mixed in a fixed way in the station's queue.

Remark 1. *In the particular case $K = J$, if we assume for convenience (and without loss of generality) that $s(j) = j$ for any $j = 1, \ldots, J$, then $C = I$, (7) becomes $W = M\,Z$ and we trivially obtain state space collapse with regular "lifting" matrix $\Delta = M^{-1}$.*

Now we establish our main result. Recall that we assume $\rho < e$.

Theorem 1. *The fluid limit model is stable if verifies state space collapse with a regular "lifting" matrix Δ.*

The proof of the theorem is based on two lemmas formulated below. For the sake of completeness we introduce a known definition:

Definition 3 (R-regularization or Skorokhod problem)**.** *Let $\tilde{X}$ be a $J-$dim. stochastic process with continuous paths, defined on some probability space, with $\tilde{X}(0) \geq 0$, and $\tilde{R}$ a $J \times J$ matrix. We say that the pair $(\tilde{W}, \tilde{Y})$ of $J-$dim. stochastic processes defined on the same probability space and with continuous paths is a solution of the $\tilde{R}-$Skorokhod problem of $\tilde{X}$ in the first orthant $\mathbb{R}^J_+$ if:*

$$\tilde{W}(t) \in \mathbb{R}^J_+ \quad \text{for all } t \geq 0\,, \quad \tilde{W} = \tilde{X} + \tilde{R}\,\tilde{Y} \text{ a.s.}$$

$$\tilde{Y} \text{ has non} - \text{decreasing paths}\,,\ \tilde{Y}(0) = 0 \text{ and for any } j,\ \tilde{Y}_j \text{ only increases}$$

$$\text{if } \tilde{W} \text{ is on face } \{w \in \mathbb{R}^J_+ \,:\, w_j = 0\}\,,\ \text{that is}\,,\ \int_0^\infty \tilde{W}_j(t)\,d\,\tilde{Y}_j(t) = 0\,.$$

Remark 2. *Proposition 4.2 [6] shows that condition* **(HR)** *on a matrix $\tilde{R}$ is sufficient to ensure strong path-wise uniqueness of the solution.*

Lemma 1 (Lemma 5.1 [2])**.** *Assume $\rho < e$. Let $(\tilde{W}, \tilde{Y})$ be the (unique) solution of the $\tilde{R}-$Skorokhod problem on the first orthant of a process $\tilde{X}$, with $\tilde{R}$ verifying assumption* **(HR)**. *If*

$$\tilde{W}(s) + \tilde{X}(t+s) - \tilde{X}(s) \geq \theta\, t \quad \text{for all } s,\, t \geq 0\,,$$

with $\theta = \tilde{R}\,(\rho - e)$, then we have that $\tilde{Y}(t+s) - \tilde{Y}(s) \leq (e-\rho)\,t$ for all $s,\, t \geq 0$, and hence $\tilde{Y}'(s) \leq (e-\rho)$ if $\tilde{Y}(\cdot)$ is differentiable at s and $\tilde{Y}'(\cdot)$ denotes its derivative.

Lemma 2 (Lemma 5.2 [2])**.** *Let $f : [0, +\infty) \longrightarrow [0, +\infty)$ be a nonnegative function that is absolutely continuous and let $\kappa > 0$ be a constant. Suppose that for almost surely (with respect to the Lebesgue measure on $[0, +\infty)$) all regular points t, $f'(t) \leq -\kappa$ whenever $f(t) > 0$. Then f is nonincreasing and $f(t) \equiv 0$ for $t \geq \frac{f(0)}{\kappa}$.*

Proof of Theorem 1: Consider a *fluid limit model* (α, M, C, P, z) with $\rho < e$ and satisfying *state space collapse* with a regular "lifting" matrix Δ. We want to prove the existence of some $t_0 > 0$ such that for any solution of the fluid model equations, $\Psi(\cdot) = \big(A(\cdot), D(\cdot), T(\cdot), Z(\cdot), W(\cdot), Y(\cdot)\big),\ Z(t) = 0\ \ \forall t \geq t_0\,|z|$.

Step 1: We will see that (W, Y) is the unique solution of the $R-$Skorokhod problem of X on the first orthant, X being defined by $X(t) \stackrel{\text{def}}{=} W(0) + \theta\, t$, and

$R = (C\,M\,Q\,\Delta)^{-1}$. Indeed, from (2) we obtain $D(t) = z + A(t) - Z(t)$, which can be substituted in (1) giving

$$A(t) = Q\,\alpha\,t + Q\,P^T\,z - Q\,P^T\,Z(t)\,. \tag{8}$$

By *state space collapse* assumption with *regular* "lifting" matrix Δ, we can replace in (8) Z by $\Delta\,W$, and by substituting into (6) obtain

$$W(t) = W(0) + C\,M\left(Q\,\alpha\,t + Q\,P^T\,\Delta\,W(0) - Q\,P^T\,\Delta\,W(t)\right) - e\,t + Y(t)\,,$$

by using (4) and the fact that $W(0) = C\,M\,z$. By isolating $W(t)$ in its turn from this expression and taking into account the definition of R and the fact that $I + C\,M\,Q\,P^T\,\Delta = C\,M\,Q\,\Delta$, and that $\rho = C\,M\,Q\,\alpha$, we finally have that

$$W(t) = W(0) + R\,(\rho - e)\,t + R\,Y(t)\,. \tag{9}$$

If we denote $R\,(\rho - e)$ by θ as in Lemma 1, we have by (9) and (5) that $(W,\,Y)$ is a solution of the $R-$Skorokhod problem of X on the first orthant. Assumption **(HR)** on matrix R given by the regularity of Δ, ensures the uniqueness of the solution. Therefore we can apply Lemma 1 because

$$W(s) + X(t+s) - X(s) \geq \theta\,t \quad \text{for all } s,\,t \geq 0\,,$$

which is easy to check since

$$W(s) + X(t+s) - X(s) = W(s) + (W(0) + \theta\,(t+s)) - (W(0) + \theta\,s) = W(s) + \theta\,t\,,$$

and $W \geq 0$. As a consequence, if Y is differentiable at point s,

$$Y'(s) \leq e - \rho\,. \tag{10}$$

Step 2: Take the Lyapunov function

$$g(t) = e^T\,R^{-1}\,W(t)\,, \tag{11}$$

to which we will apply Lemma 2. By substituting (9) into (11),

$$g(t) = e^T\,R^{-1}\,W(0) + e^T\,(\rho - e)\,t + e^T\,Y(t) = g(0) + \sum_{j=1}^{J}\left((\rho_j - 1)\,t + Y_j(t)\right).$$

Then, the points of differentiability of $Y_j(\cdot)$ coincide with those of $g(\cdot)$, and if t is one of these points,

$$g'(t) = \sum_{j=1}^{J}\left((\rho_j - 1) + Y_j'(t)\right), \tag{12}$$

$g'(t)$ being non positive by (10). We finish the proof using Lemma 2. To this end, let $t \geq 0$ be a point such that $g(t) > 0$ (if any). By definition of g and

nonnegativity of all elements of R^{-1}, there exists i such that $W_i(t) > 0$. Then, by (5), $Y_i'(t) = 0$, and by (12),

$$g'(t) \leq \rho_i - 1 \leq \max_{j=1,\ldots,J} \rho_j - 1 \quad (< 0 \quad \text{because } \rho < e\ ). \tag{13}$$

Thus, we have proved that $g'(t) \leq -\kappa$, with $\kappa = 1 - \max_{j=1,\ldots,J} \rho_j > 0$, at any point t of differentiability of $g(\cdot)$ such that $g(t) > 0$. Lemma 2 ensures that, in this situation, $g(\cdot)$ is non-increasing and that $g(t) \equiv 0$ for $t \geq \frac{g(0)}{\kappa}$.

Finally we have that

$$g(0) = e^T R^{-1} W(0) = e^T R^{-1} C M z \leq e^T R^{-1} C M e\, |z|\,,$$

and therefore, by Lemma 2,

$$g(t) \equiv 0 \ \text{ for any } t \geq t_0\, |z|\,, \quad \text{with } t_0 = \frac{e^T R^{-1} C M e}{1 - \max_{j=1,\ldots,J} \rho_j} > 0\,.$$

On account of (11) and the nonnegativity of the elements of R^{-1} we also obtain that $W(t) \equiv 0$ for any $t \geq t_0\,|z|$, and the same applies for Z. $\square$

Remark 3. *In the particular case $J = 1$ (a $\bigvee -$system), (12) becomes $g'(t) = (\rho_1 - 1) + Y_1'(t)$, and (13) in its turn, $g'(t) = \rho_1 - 1 < 0$. The rest of the proof follows similarly with $\kappa = 1 - \rho_1 > 0$. Note that we do not use (10) in this situation, so actually we do not need Lemma 1. As a consequence, if $J = 1$ we have that $\rho_1 < 1$ is sufficient to ensure stability (see Theorem 6.1 [2]).*

References

[1] Bramson, M.: State space collapse with application to heavy traffic limits for multiclass queueing networks. Queueing Syst. **30**, 89-148 (1998).

[2] Dai, J. G.: On positive Harris recurrence of multiclass queueing networks: a unified approach via fluid limit models. Ann. Appl. Prob. **5(1)**, 49-77 (1995).

[3] Delgado, R.: State space collapse for asymptotically critical multi-class fluid networks. Queueing Syst. **59**, 157-184 (2008).

[4] Peterson, W. P.: Diffusion approximations for networks of queues with multiple customers types. Math. Oper. Res. **16**, 90-118 (1991).

[5] Reiman, M. I.: Open queueing networks in heavy traffic. Math. Oper. Res. **9(3)**, 441-458 (1984).

[6] Williams, R. J.: An invariance principle for semimartingale reflecting Brownian motions in an orthant. Queueing Syst. **30**, 5-25 (1998).

[7] Williams, R. J.: Diffusion approximations for open multi-class queueing networks: sufficient conditions involving state space collapse. Queueing Syst. **30**, 27-88 (1998).

6th St.Petersburg Workshop on Simulation (2009) 791-795

Importance sampling simulation of the fork-join queue

Ad Ridder[1]

Abstract

In this paper we consider a rare-event problem in the fork-join queue for which we develop an efficient importance sampling algorithm.

1. Introduction

We consider the discrete-time Markov chain on the two-dimensional positive quadrant of integers that results by embedding a two-dimensional fork-join queue [5]. This is queueing system with a single Poisson arrival process with rate λ, and where any arriving job splits itself in two subjobs each joining a single server queue. These two queues act as independent $M/M/1$ queues with service rates μ_1, μ_2, respectively. For stability we demand $\lambda < \min(\mu_1, \mu_2)$. The folklore application of this queueing model is a system with two bathrooms, one for men and one for women, with arrivals of heterosexual couples only. The original motivation was to study a machine with parallel coupled processors, and an inventory control problem of data base systems.

The associated discrete-time Markov chain that results by embedding at jump times, is denoted by $(S(k))_{k=0}^{\infty}$ and has state space $\mathbb{Z}_+^2$, representing the backlogs at the two queues (including the servers). We are interested in transient probabilities of large backlogs of at least one of the queues:

$$\gamma_n(x, y, T) = \mathbb{P}(S_1(nT) \geq ny_1 \text{ or } S_2(nT) \geq ny_2 | S(0) = nx), \tag{1}$$

for fixed scaled initial state $x = (x_1, x_2) \in \mathbb{R}_+^2$, fixed scaled threshold $y = (y_1, y_2) \in \mathbb{R}_+^2$, fixed scaled horizon $T > 0$, and parameter $n \to \infty$. The set of interest is scaled by n and then called the rarity set:

$$D = \{\eta \in \mathbb{R}_+^2 : \eta_1 \geq y_1 \text{ or } \eta_2 \geq y_2\}.$$

The difficulty here is twofold: (i) the rarity set is not convex which may cause troubles for developing an efficient importance sampling scheme [4]; and (ii) the set D cannot decomposed in two disjoint sets such that the separate probabilities are estimated by efficient importance sampling estimators [6].

[1]Vrije University Amsterdam, E-mail: aridder@feweb.vu.nl

In this paper we investigate the method of universal simulation distributions for finding an efficient importance sampling estimator. This method has been introduced originally in [9], and further developed in [1], see also [2, Chapter 10]. The method is based on large deviations for sequences of random variables, however, we need to adapt it to process level because we deal with sample path large deviations. In Section we shall briefly review the method of universal simulation distributions, and the sample path large deviations for the folk-join queue. In Section we give our algorithm and we conclude with a few numerical results in Section .

2. Preliminaries

Universal simulation distributions.

Suppose that we can write the target of our problem to be

$$\gamma_n(x, y, T) = \mathbb{P}\left(f_n(Y_n)/n \in E\right),$$

where Y_n is a random variable on some appropriate space $\mathcal{S}_n$, and f_n a function $\mathcal{S}_n \to \mathbb{R}^d$. Let P_n be the probability measure on $\mathcal{S}_n$ induced by Y_n. Under some regularity conditions the Gärtner-Ellis theorem holds [2, Section 3.2],

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(f_n(Y_n)/n \in E\right) = -I(E).$$

The large deviations rate function $I(v), v \in \mathbb{R}^d$, is the convex conjugate of the asymptotic log moment generating function of $(f_n(Y_n))$, with optimising argument θ_v. Suppose that there is also a sequence $(Z_n)_n$ of $\mathcal{S}_n$-valued random variables, with induced probability measure Q_n, such that $P_n \ll Q_n$. The importance sampling estimator of the target probability γ_n is defined by

$$\widehat{\gamma_n} = \frac{dP_n}{dQ_n}(Z_n)\, 1\{f_n(Y_n)/n \in E\}. \tag{2}$$

The universal simulation distribution method says that if there are $m < \infty$ points $v_1, \ldots, v_m \in \mathbb{R}^d$, such that

$$\text{(i) } I(v_i) \geq I(E) \quad (i = 1, \ldots, m); \quad \text{(ii) } E \subset \cup_{i=1}^m \mathcal{H}(v_i), \tag{3}$$

where $\mathcal{H}(v)$ is the half-space $\{w \in \mathbb{R}^d : \langle \theta_v, w - v\rangle \geq 0\}$, then for any probability vector $\pi = (\pi_i)$ with positive elements, the change of measure

$$dQ_n(s) = \left(\sum_{i=1}^m \pi_i \exp\left(\langle \theta_{v_i}, f_n(s)\rangle - \psi_n(\theta_{v_i})\right)\right) dP_n(s)$$

gives an asymptotically optimal importance sampling estimator (2), see [1, 9].

The random walk associated with the fork-join queue.

The discrete-time Markov chain $(S(k))_{k=0}^{\infty}$ representing the backlogs at the queues at their jump times is a face-homogeneous random walk on the positive quadrant $\mathbb{Z}_+^2$, which means the following [8]: for any $s \in \mathbb{R}_+^2$, let $\Lambda(s)$ be the set of indices i for which $s_i > 0$. For any subset $\Lambda \subset \{1,2\}$, define face F_Λ by

$$F_\Lambda = \{s \in \mathbb{R}_+^2 : \Lambda(s) = \Lambda\}.$$

Notice that $F_\emptyset = \{0\}$. The transition probabilities $p_{ss'}$ of the chain are the same for all s in the same face, and depend only the jump $s' - s$:

$$p_{ss'} = p_{\Lambda(s)}(s - s').$$

Thus, for our two-dimensional fork-join queue, there are four random variables X_Λ that represent the jumps.

Sample path large deviations.

Since the chain is ergodic, a sample path large deviations hold for continuous piecewise linear paths [3, 8]. Firstly, we define the scaled continuous-time processes $(S^{[n]}(t))_{0\le t\le T}$, $n = 1, 2, \ldots$, by $S^{[n]}(t) = S(nt)/n$ for $t = 0, 1/n, 2/n \ldots, T$ and linear interpolation in the other points. Secondly, consider the affine function $\phi(t) = x + vt$ for $t \in [0, T]$ with $x \in \mathbb{R}_+^d$ and $v \in \mathbb{R}^d$, such that $\phi(t) \in F_\Lambda$ for all $0 < t < T$. We call the gradient v the constant speed of the path. Then for the scaled processes $S^{[n]}$, starting in $S^{[n]}(0) = x$, it holds that there exists a locate rate $\ell_\Lambda(v)$ such that

$$\lim_{\epsilon \downarrow 0} \lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(\sup_{0\le t\le T} \left| S^{[n]}(t) - \phi(t) \right| < \epsilon \right) = -T\ell_\Lambda(v). \tag{4}$$

For continuous piecewise linear paths ϕ, such that each piece lies entirely in some face, we may apply (4) to each piece and add these 'costs'.

The main issue remains to identify the local rate functions $\ell_\Lambda(v)$. A general method has been developed in [7] for face-homogeneous random walks which can be applied to our two-dimensional fork-join queue. Clearly $\ell_\emptyset = 0$ because the process is ergodic. All the other local rate functions are convex conjugates of (adapted) log moment generating functions:

$$\ell_\Lambda(v) = \sup_{\theta\in\mathbb{R}} \big(\langle \theta, v \rangle - \psi(\theta) \big). \tag{5}$$

For the interior face $F_{\{1,2\}}$ it is just the function associated with the jump variable $X_{\{1,2\}}$, for the boudary faces $F_{\{1\}}$ and $F_{\{2\}}$ we have to consider both boundary and interior. The optimisation program (5) is solved numerically and the optimiser denoted by θ_v.

3. Importance sampling algorithm

We assume that the random walk starts off at state 0. Then we rewrite the target probability (1) as

$$\gamma_n(0, y, T) = \mathbb{P}\left(S(nT)/n \in D | S(0) = 0\right) = \mathbb{P}\left(S^{[n]} \in E\right),$$

where E is an appropriate set of absolute continuous paths $\phi : [0,T] \to \mathbb{R}^2_+$ with specifically $\phi(0) = 0$ and $\phi(T) \in D$. Let $\tilde{E} \subset E$ be the subset of piecewise linear paths of the following form.

> $\phi = \phi_{\tau,v}$: it stays in 0 until time τ $(0 \leq \tau < T)$ and then it goes straight at constant speed v to the point $\eta = (T-\tau)v \in D$.

The theory of sample path large deviations [3, 7, 10] says that

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(S^{[n]} \in E\right) = \lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}\left(S^{[n]} \in \tilde{E}\right) = -I(\tilde{E}).$$

So, we are in the business of the universal simulation distributions by finding $m < \infty$ pairs $\left(\tau, v^{(i)}\right)$ $(i = 1, \ldots, m)$ (same τ!) such that (see (3)),

$$\text{(i) } I(\phi^{(i)}) \geq I(\tilde{E}) \quad (i = 1, \ldots, m); \quad \text{(ii) } V_\tau \subset \cup_{i=1}^m \mathcal{H}(v^{(i)}),$$

where

(a.) $\phi^{(i)} = \phi_{\tau,v^{(i)}} \in \tilde{E}$ with cost $I(\phi^{(i)}) = \tau \ell_{\Lambda(v^{(i)})}(v^{(i)})$;

(b.) $V_\tau = \{v \in \mathbb{R}^2_+ : (T-\tau)v \in D\}$, set of speed vectors to the rarity set;

(c.) $\mathcal{H}(v)$ is the half space $\{w \in \mathbb{R}^d : \langle \theta_v, w - v\rangle \geq 0\}$, where θ_v is the optimiser for the local rate function $\ell_{\Lambda(v)}(v)$, see (5).

Hence, the remarkable observation is that we apply the method to drift vectors of affine paths of the (limiting scaled) random walk. From (b) and (c) we deduce that two pairs suffice, one on each of the two boundaries:

$$\begin{cases} v_1^{(1)} = y_1/(T-\tau), \quad 0 \leq v_2^{(1)} < y_2/(T-\tau); \\ 0 \leq v_1^{(2)} < y_1/(T-\tau), \quad v_2^{(2)} = y_2/(T-\tau); \\ (\theta_{v^{(1)}})_2 = 0, \quad (\theta_{v^{(2)}})_1 = 0. \end{cases} \tag{6}$$

However, in most examples there are no two such pairs, but there are two pairs with different zero-sojourn times $\tau^{(i)}$ that satisfy (6). For each of the drift vectors $v^{(i)}$ we determine the associated jump probabilities $q_\Lambda^{(i)}$ of the jump variables X_Λ by an exponential change of measure given by the shift parameters $\theta_{v^{(i)}}$.

4. Example

Let $\lambda = 1, \mu_1 = 1.5, \mu_2 = 2, x = (0.0), y = (1, 1.2), T = 10$. The optimal path $\phi^* = \arg\inf_{\phi\in E} I(\phi)$, has a zero-sojourn time of $\tau = 2$ time units and then runs at a constant speed $v = (1/8, 0)$ along face $F_{\{1\}}$ to state $(1, 0)$.

The naive importance sampling algorithm is based on simulating the random walk with the original jump probabilities p_Λ until time $n\tau$, and next with the exponentially twisted jump probabilities obtained by the parameter θ_v until time nT. It can be shown that the associated estimator is not efficient. Its behaviour is

illustrated by plotting the estimated target probability, and the estimated variance of the estimator as functions of the sample size k. The estimates are irregular overestimates of the true value, and the variances get a shock upward whenever the rare event is hit by a 'wrong' path.

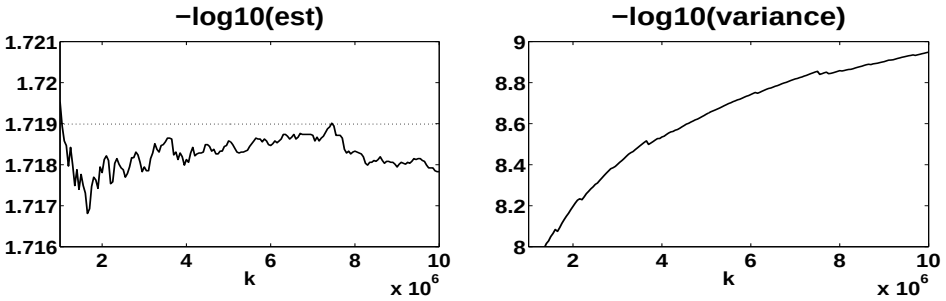


Figure 1. *Estimates from the naive importance sampling algorithm for sample sizes $k = 10^5$–10^7. Scaling $n = 10$. The dotted line is the exact probability.*

We find numerically a solution of the system (6) with different zero-sojourn times. The first pair is $\left(\tau^{(1)}, v^{(1)}\right) = (2, (1/8, 0))$, which gives the same path ϕ^* as above, the other pair is $\left(\tau^{(2)}, v^{(2)}\right) = (4.6, (1/9, 2/9))$, which gives a path running in the interior face $F_{\{1,2\}}$. In the simulations we used mixing probabilties $(0.8, 0.2)$.

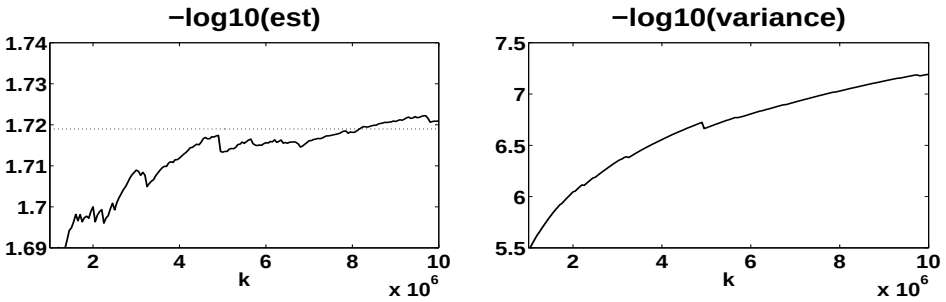


Figure 2. *Estimates from the importance sampling algorithm with universal simulation distribution for sample sizes $k = 10^5$–10^7. Scaling $n = 10$. The dotted line is the exact probability.*

We experimented with various scalings $n = 10$–200 with sample size $k = 50000$ and collected the relative half width of the 95% confidence interval RHW (efficient estimators have RHW that grows at most polynomially), and the ratio RAT $= \log \mathbb{E}[(\widehat{\gamma_n})^2] / \log \mathbb{E}[\widehat{\gamma_n}]$ (efficient estimators have RAT that converge to 2).

The conclusion is that the mixed importance sampling estimator improves the naive one, and is asymptotically optimal, although its variance is still irregular. Further investigations are needed, also for other starting points, and other algorithms, for instance with mixing probailities that depend on state and time [4].

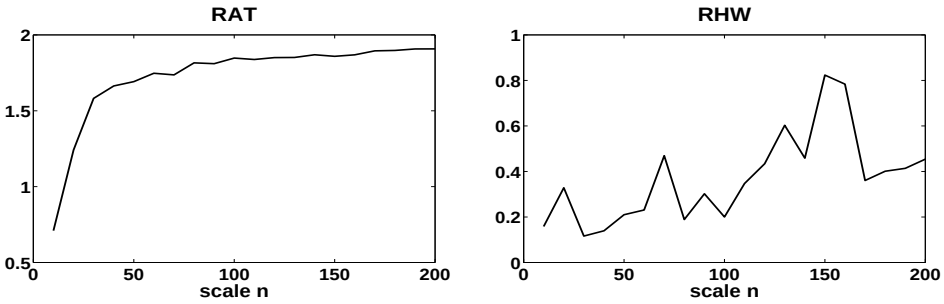


Figure 3. *Performance of the importance sampling estimator with universal simulation distribution for scalings* $n = 10$–200.

References

[1] Bucklew, J.A., Nitinarawat, S., Wierer, J., 2004. Universal simulation distributions, *IEEE Transactions on Information Theory* 50, pp. 2674-2685.

[2] Bucklew, J.A., 2004. *Introduction to Rare Event Simulation*. Springer, New York.

[3] Dupuis, P., and Ellis, R., 1995. The large deviation principle for a general class of queueing systems, *Transactions of the American Mathematical Society* 347, pp. 2689-2751.

[4] Dupuis, P., and Wang, H., 2007. Subsolutions of an Isaacs equation and efficient schemes for importance sampling. *Mathematics of Operations Research* 32, pp. 723-757.

[5] Flatto, L., and Hahn, S., 1984. Two parallel queues created by arrivals with two demands, *SIAM Journal on Applied Mathematics* 44, pp. 1041-1053.

[6] Glassermann, P., and Wang, Y., 1997. Counterexamples in importance sampling for large deviations probabilities, *Annals of Applied Probability* 7, pp. 731-746.

[7] Ignatiouk-Robert, I., 2001. Sample path large deviations and convergence parameters, *Annals of Applied Probability* 11, pp. 1292-1329.

[8] Ignatiouk-Robert, I., 2005. Large deviations for processes with discontinous statistics, *Annals of Probability* 33, pp. 1479-1508.

[9] Sadowsky, J.S., and Bucklew, J.A., 1990. On large deviations theory and asymptotically efficient Monte Carlo estimation. *IEEE Transactions on Information Theory* 36, pp. 579-588.

[10] Shwartz, A., and Weiss, A., 1995. *Large Deviations for Performance Analysis*, Chapman & Hall.

6th St.Petersburg Workshop on Simulation (2009) 797-801

A discrete-time queueing system with total renewal discipline[1]

Ivan Atencia[2], Alexander V. Pechinkin[3]

Abstract

This paper analyses a discrete-time queueing system with geometrical arrivals and total renewal service discipline, that is, once a service is completed all the customers in the queue get service simultaneously through a geometric process. We study the underlying Markov chain and the joint generating function of the number of customers in the service and in the queue. We also derive the mean number of the customers in the system and in the queue and the mean queueing time and sojourn time. Finally, we give numerical examples to illustrate the effect of the parameters on several performance characteristics.

1. Introduction

The interest in discrete systems has experimented a spectacular growth with the advent of the new technologies. A fundamental incentive to study discrete-time queues is that these systems are more appropriate than their continuous-time counterparts for modelling computer and telecommunication systems, since the basic units in these systems are digital such as a machine cycle time, bits and packets, etc. Indeed, much of the usefulness of discrete-time queues derives from the fact that they can be used in the performance analysis of Broadband Integrated Services Digital Network (BISDN), Asynchronous Transfer Mode (ATM) and related computer communication technologies wherein the continuous-time models do not adapt [1], [2].

In many real telecommunication systems, it is frequently observed that the server processes the packets in groups. In such bulk-service systems, jobs that arrive one at a time must wait in the queue until a sufficient number of jobs get accumulated. A variety of bulk-service queues with infinite waiting space have been studied by many researchers e.g. [3], [4] and [5]. This service discipline is closely related to other disciplines described in the queueing literature like G-networks, clearing systems, catastrophes, etc, see for example [6]– [11].

[1]This work was supported by grants MTM2008-01121 and 08-07-00152 of the Russian Foundation for Basic Research.

[2]Málaga University, E-mail: iatencia@ctima.uma.es

[3]Institute of Informatics Problems RAS, E-mail: apechinkin@ipiran.ru

In this paper we consider a discrete-time single-server queueing system with infinite buffer and bulk-service with the peculiarity that when a service is completed all the customers present in the queue get service simultaneously through a geometric process.

2. The Mathematical model

We consider a single-server discrete-time queueing system where the time axis is divided into a sequence of equal time intervals (called slots) and it is assumed that all queueing activities (arrivals and departures) take place at the slot boundaries. For mathematical convenience, we will suppose that the departures occur at the moment immediately before the slot boundaries and the arrivals occur at the moment immediately after the slot boundaries.

Customers arrive according to a Bernoulli arrival process with rate a, that is, a is the probability that a customer arrives at a slot. If, upon arrival, the service is idle, the service of the arriving customer begins immediately.

Once a service has finished, all the customers of the queue get service simultaneously according to a geometric process of parameter b.

3. The Markov chain

At time m^+ (the instant immediately after time slot m), the system can be described by the process $X_m = (H_m, Q_m)$ where H_m and Q_m represent the number of customers in the service and in the queue respectively.

It can be shown that $\{X_m,\ m \in \mathbb{N}\}$ is the two-dimensional Markov chain of our queueing system, whose state space is:

$$\chi = \{(0,0),\ (i,j),\ \ i \geq 1,\ \ j \geq 0\}.$$

Our next objective is to study the stationary distribution of the Markov chain $\{X_m,\ m \in \mathbb{N}\}$, which will be denoted by

$$p_{ij} = \lim_{m \to \infty} P[X_m = (i,j)],\ \ (i,j) \in \chi.$$

The Kolmogorov equations for the stationary distribution of the system are:

$$p_{00} = \overline{a}\, p_{00} + \overline{a}\, b \sum_{i=1}^{\infty} p_{i0} \Leftrightarrow a\, p_{00} = \overline{a}\, b \sum_{i=1}^{\infty} p_{i0}, \tag{1}$$

$$p_{10} = a\, p_{00} + \overline{a}\, \overline{b}\, p_{10} + a\, b \sum_{i=1}^{\infty} p_{i0} + \overline{a}\, b \sum_{i=1}^{\infty} p_{i1}, \tag{2}$$

$$p_{i0} = \overline{a}\, \overline{b}\, p_{i0} + \overline{a}\, b \sum_{l=1}^{\infty} p_{li},\ \ i \geq 2, \tag{3}$$

$$p_{i1} = a\,\overline{b}\,p_{i0} + \overline{a}\,\overline{b}\,p_{i1} + a\,b\sum_{l=1}^{\infty} p_{li},\ \ i \geq 1, \tag{4}$$

$$p_{ij} = a\,\overline{b}\,p_{i,j-1} + \overline{a}\,\overline{b}\,p_{ij},\ \ i \geq 1,\ \ j \geq 2, \tag{5}$$

where $\overline{a} = 1 - a$ and $\overline{b} = 1 - b$.

The normalizing condition is

$$\sum_{i=0}^{\infty} p_{i0} + \sum_{i=1}^{\infty} p_{i1} + \sum_{i=1}^{\infty}\sum_{j=2}^{\infty} p_{ij} = 1.$$

With the aim of solving (1)–(5) we introduce the following generating functions:

$$P_0(z) = \sum_{i=0}^{\infty} p_{i0}\,z^i, \qquad P_0^*(z) = \sum_{i=1}^{\infty} p_{i0}\,z^i,$$

$$P_1(z) = \sum_{i=1}^{\infty} p_{i1}\,z^i, \qquad P(z_1, z_2) = \sum_{i=1}^{\infty}\sum_{j=2}^{\infty} p_{ij}\,z_1^i\,z_2^i.$$

From Eq. (1), (2) and (3) we readily obtain:

$$P_0^*(1) = \frac{a\,p_{00}}{\overline{a}\,b}, \qquad P_1(1) = \frac{a\,\overline{b}}{\overline{a}\,(1 - \overline{a}\,\overline{b})}\,P_0^*(1).$$

Multiplying Eq. (5) by $z_1^i\,z_2^j$ and summing over i and j we obtain after some algebra:

$$P(z_1, z_2) = \frac{a\,\overline{b}\,z_2^2}{1 - \overline{a}\,\overline{b} - a\,\overline{b}\,z_2}\,P_1(z_1). \tag{6}$$

In order to obtain the value of p_{00} we note that the normalizing condition can be written as

$$p_{00} + P_0^*(1) + P_1(1) + P(1,1) = 1.$$

Henceforth taking into account (6) and the expressions of $P_0^*(1)$ and $P_1(1)$ we have that

$$p_{00} = \frac{(\overline{a}\,b)^2}{(\overline{a}\,b)^2 + a\,(a\,\overline{b} + \overline{a}\,b)}.$$

From Eqs. (2) and (3) we obtain the expression of $P_0^*(z)$:

$$P_0^*(z) = \frac{1 + \nu\,(1 - z)}{1 - \nu\,z} \cdot \frac{\nu\,z}{\overline{a}\,\overline{b}}\,p_{00},$$

and from Eqs. (4) we have:

$$P_1(z) = \frac{1 + \overline{a}\,\nu\,(1 - z)}{1 - \nu\,z} \cdot \frac{\nu^2\,z}{\overline{a}^2\,\overline{b}}\,p_{00}$$

where $\nu = a\,\overline{b}/(1-\overline{a}\,\overline{b})$.

Note that if we call

$$P_j(z) = \sum_{i=1}^{\infty} p_{ij}\, z^j,\ \ j \geq 1.$$

Using (5) we have that

$$P_j(z) = \nu\, P_{j-1}(z),\ \ j \geq 2,$$

and applying recursively the former relation we obtain that

$$P_j(z) = \nu^{j-1}\, P_1(z),\ \ j \geq 1.$$

Then, we can express the generating function $P(z)$ of the number of customers in the system as

$$P(z) = P_0(z) + \sum_{j=1}^{\infty} z^j\, P_j(z) = P_0(z) + z\, P_1(z) \sum_{j=0}^{\infty} (\nu\, z)^j$$

$$= P_0(z) + \frac{z}{1-\nu\, z}\, P_1(z) = \left[1 + \frac{a\,\nu\, z + \overline{a}\,(1+\nu\,(1-z))}{(1-\nu\, z)^2} \cdot \frac{\nu\, z}{\overline{a}^2\,\overline{b}}\right] p_{00}.$$

The mean number of customers in the system is given by

$$N = P'(1) = \frac{\overline{a}\,(1-\nu)^2 + 2\,\nu}{(1-\nu)^3} \cdot \frac{\nu}{\overline{a}^2\,\overline{b}}\, p_{00}.$$

The generating function $Q(z)$ of the number of customers in the queue is

$$Q(z) = P_0(1) + z\, P_1(1) + P(1,z) = \left[1 + \frac{a\,(\overline{a} + a\,\nu\, z)}{\overline{a}^2\, b\,(1-\nu\, z)}\right] p_{00},$$

and the mean length of the queue has the expression

$$Q = Q'(1) = \frac{a\,\nu}{\overline{a}^2\, b\,(1-\nu)^2}\, p_{00}.$$

The generating function of the waiting time that spends a customer before obtaining service is

$$w(z) = p_0^* + (1 - p_0^*)\,\beta(z)$$

where p_0^* is the probability that a customer arrived finds an empty system and $\beta(z) = b\, z/(1-\overline{b}\, z)$ is the generating function of the service time. To find the constant p_0^*, observe that $p_0^* = p_{00} + b\, P_0^*(1) = p_{00}/\overline{a}$.

The generating function of the sojourn time is given by

$$\varphi(z) = w(z)\,\beta(z).$$

The mean time w of waiting for service and the sojourn time v obey the formulas:

$$w = w'(1) = \frac{1 - p_0^*}{b}, \qquad v = \varphi'(1) = \frac{2 - p_0^*}{b}.$$

We note that this system verifies the Little's principle, that is:

$$Q = a\, w \ \text{ and } \ N = a\, v.$$

4. Numerical examples

In this section, we present some numerical examples on the performance measures.

Firstly in Figure 1 the mean number of customers in the system is plotted against the arrival rate a for diferent values of the service parameter ($b = 0.1, 0.3, 0.5, 0.7$).

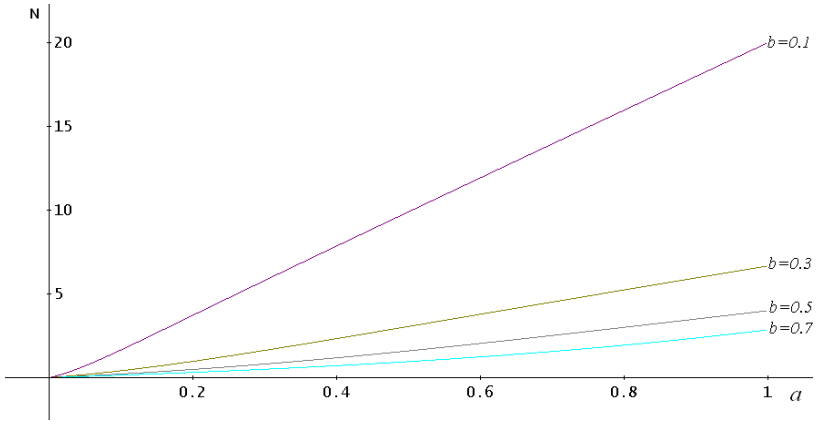


Figure 1: Mean number of customers in the system versus parameter a.

As it was expected, the mean number of customers in the system increases as the arrival rate increases. In the graphic it is also shown that for the parameter $b = 0.1$ the jump in terms of demands in the system is considerably higher than in the rest ones. This difference is caused by the small rate of service.

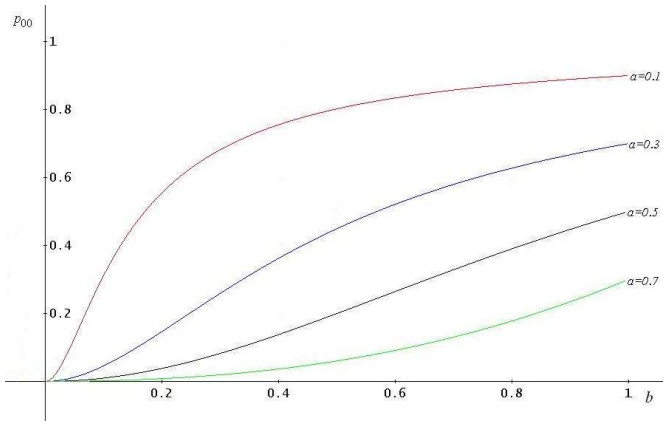


Figure 2: Probability that the system is empty versus parameter b.

In Figure 2 the probability that the system is empty is plotted against the service rate b for different values of a ($a = 0.1, 0.3, 0.5, 0.7$). The graphic 2 corro-

borates that as the service rate increases the probability that the system is empty increases, although this growth depends on the arrival rate a.

References

[1] Bruneel H., Kim B.G. (1993) *Discrete-time Models for Communication Systems Including ATM*. Kluwer Academic Publishers, Boston.

[2] Woodward M.E. (1994) *Communication and Computer Networks: Modelling with Discrete-time Queues*. IEEE Computer Society Press, California.

[3] Chaudhry M.L., Templeton J.G.C. (1983) *A First Course in Bulk Queues*. John Wiley & Sons, New York.

[4] Medhi J. (1984) *Recent Development in Bulk Queueing Models*. Wiley Eastern Limited.

[5] Medhi J. (1991) *Stochastic Models in Queueing Theory*. Academic Press Inc.

[6] Artalejo J.R. (2000) *G-networks: A versatile approach for work removal in queueing networks*, Eur. J. Oper. Res., **126**, 233-249.

[7] Atencia I., Moreno P. (2004) *The discrete-time Geo/Geo/1 queue with negative customers and disasters*, Comput. Oper. Res. **31**, 1537-1548.

[8] Bocharov P.P., Zaryadov I.S. (2007) *Stationary probability distribution of a queueing system with renovation*, Vestnik RUDN series Mathematics, I. Technology, Phisics. **1–2**, 15–25. (in Russian).

[9] Zaryadov I.S. (2008) *Stationary service characteristics in a $G/M/n/r$ system with generalized renovation*, Vestnik RUDN series Mathematics, I. Technology, Phisics. **2**, 3–10. (in Russian).

[10] Kreinin A. (1997) *Queueing Systems with Renovation*, Journal of Applied Math. Stochast. Analysis. **10**, 431–443.

[11] Towsley D., Tripathi S.K. (1999) *A single server priority queue with server failure and queue flushing*, Oper. Res. Lett.

[12] Bocharov P.P., D'Apice C., Pechinkin A.V., Salerno S. (2004) *Queueing Theory*. Brill Academic Pub.

6th St.Petersburg Workshop on Simulation (2009) 803-807

A discrete-time retrial queueing model with multiple servers

Rein Nobel[1]

A multi-server queueing model with retrials is considered in discrete time. The time-axis is divided into slots and time is counted only in slots. During every slot primary customers arrive in batches, the batch size follows a general probability distribution and the numbers of primary arrivals in different time slots are mutually independent. Each customer requires from a server a geometrically distributed number of slots for his service, and the service times of the different customers are independent. Customers arriving in a slot can start their service only at the beginning of the next slot [Delayed Access]. When upon arrival customers find all servers busy the system either sends all the incoming customers into orbit, or one or more new servers are activated to serve some or all incoming customers. When upon arrival customers find one or more servers idle, then all incoming customers who can be assigned to an idle server start their service at the beginning of the next slot, whereas for the remaining incoming customers in that slot, if any, a choice must be made between sending them into orbit or activating one or more servers to enable the start of their service at the beginning of the next time slot. During each slot customers in the orbit try to reenter the system individually, independent of each other, with a given retrial probability. So, the total number of arrivals in a given slot is the sum of the number of primary customers and the number of customers from the orbit arriving in that slot. Departures take place at the end of a slot, so arrivals have precedence over departures [Late Arrivals]. At departure epochs a choice must be made between deactivating one or more of the servers who have become idle, or keeping them active to serve future arrivals. The number of servers available is unlimited.

Keeping idle servers active goes with a cost, and serving customers brings with it a reward (both per server and per unit time). Activating a new server requires a set-up cost, and finally, for customers in the orbit linear holding costs are incurred.

The problem is to find an optimal activating/deactivating policy for the servers in order to obtain a total minimal cost per unit time. By modeling the system as a discrete-time Markov decision model we show how this policy can be calculated efficiently. As a side result we obtain several performance measures for the multi-server retrial model in discrete time with a fixed number of servers.

[1]Department of Econometrics, Vrije University, Amsterdam. E-mail: rnobel@feweb.vu.nl

6th St.Petersburg Workshop on Simulation (2009) 804-808

The maximum number of infectives in SIS epidemic models: Computational techniques and quasi-stationary distributions[1]

Antonis Economou[2]

Abstract

We study the maximum number of infectives (MNI) for a Susceptible-Infective-Susceptible (SIS) model which corresponds to a birth-death process with an absorbing state. We develop computational schemes for the corresponding transient and steady-state distributions. Some connections with quasi-stationary distributions of the model are also discussed.

1. Introduction

The study of the MNI during an epidemic is of great importance for assessing its impact and the possibilities of intervention for controlling it. In this paper we study the MNI in the framework of a generalized SIS model. This epidemic is modeled as a birth-death process that counts the number of infectives in a finite population. State 0 is the unique absorbing state that corresponds to the end of the epidemic and all the other states are transient (for details see Allen (2003)). Therefore, the stationary distribution of the model is degenerate and the main computations concern the so-called quasi-stationary distribution and some transient distributions associated with its evolution.

Similar questions have been studied in the framework of queueing theory, where the maximum number of customers during a busy period has been used for assessing the level of congestion. Neuts (1964) studied the maximum number of customers during a busy period in a basic queueing model, while Serfozo (1988) introduced an asymptotic approach for the study of the extreme values of a birth-death process. These works have been further generalized and extended for certain structured multidimensional Markov chains (see e.g. Artalejo et al. (2007), Artalejo and Chakravarthy (2007) and Artalejo (2008)).

The paper is organized as follows. In Section 2 we introduce the model and we present an efficient algorithm for computing the transient distribution of the MNI. The corresponding steady-state distribution is also derived in closed form. In Section 3 we study the connections of the MNI with the quasi-stationary distribution of the model. Section 4 presents a numerical example.

[1]This work was supported by the Ministry of Science and Innovation of Spain (grant MTM2008-01121) and by the University of Athens (grant Kapodistrias 70/4/6415).

[2]University of Athens, E-mail: aeconom@math.uoa.gr

2. Transient and steady-state analysis of the MNI

A SIS epidemic model in continuous-time is a closed population model of N individuals, in which the population consists only of susceptibles and infectives. The evolution of such a model can be described by a birth-death process $\{I(t), t \geq 0\}$ with state space $S = \{0, 1, \ldots, N\}$, where $I(t)$ records the number of infectives at time t. The birth rates, corresponding to infections, are denoted by λ_i and the death rates, corresponding to recuperations, are denoted by μ_i, $i = 0, 1, \ldots, N$. The infections are supposed to occur because of a contagious disease. Therefore, when there are no infectives, the process stays there for ever. The other states are assumed transient. More specifically we assume that $\lambda_0 = \lambda_N = \mu_0 = 0$, while $\mu_1, \mu_2, \ldots, \mu_N > 0$ and $\lambda_1, \lambda_2, \ldots, \lambda_{N-1} > 0$. In the classical SIS model, it is assumed that $\lambda_i = \beta i(N-i)/N$ and $\mu_i = \gamma i$, where β is the contact rate and γ is the recovery rate per customer. However, in the present paper we will present the results for general birth and death rates. Indeed in several biological systems the data do not support the above classical rates.

We are interested in the distribution of the MNI $M(t) = \max\{I(s) : 0 \leq s \leq t\}$. Let $I(0) = i_0$ be the number of infectives at the beginning of the observation period and $M(0) = k_0$ be the maximum number observed till that time, $0 < i_0 \leq k_0$. We are interested in computing the transient distribution $p_{i,k}(t) = \Pr[I(t) = i, M(t) = k | I(0) = i_0, M(0) = k_0]$. It is clear that $p_{i,k}(t) = 0$ for $k < k_0$ or $k < i$, so we have to compute $p_{i,k}(t)$ for $k_0 \leq k \leq N$ and $0 \leq i \leq k$ (equivalently for $\max(k_0, i) \leq k \leq N$ and $0 \leq i \leq N$). The forward Kolmogorov differential equations of the process $\{(I(t), M(t)), t \geq 0\}$ assume the form

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} p_{i,k}(t) &= -(\lambda_i + \mu_i) p_{i,k}(t) + (1 - \delta_{i,0}) \lambda_{i-1} p_{i-1,k}(t) \\ &\quad + (1 - \delta_{i,k})(1 - \delta_{i,N}) \mu_{i+1} p_{i+1,k}(t) \\ &\quad + (1 - \delta_{k,k_0}) \delta_{i,k} \lambda_{i-1} p_{i-1,k-1}(t),\ k_0 \leq k \leq N,\ 0 \leq i \leq k, \end{aligned} \tag{1}$$

with $\delta_{i,k}$ being Kronecker's 0-1 function and initial conditions

$$p_{i,k}(0) = \delta_{i,i_0} \delta_{k,k_0},\ k_0 \leq k \leq N,\ 0 \leq i \leq k. \tag{2}$$

These equations can be solved by employing Laplace transforms. To this end we introduce the Laplace transforms $\tilde{p}_{i,k}(s) = \int_0^\infty e^{-st} p_{i,k}(t) \mathrm{d}t$. By transforming the system of equations (1), taking into account (2), we obtain the linear system

$$\begin{aligned} s\tilde{p}_{i,k}(s) &= \delta_{i,i_0} \delta_{k,k_0} - (\lambda_i + \mu_i) \tilde{p}_{i,k}(s) + (1 - \delta_{i,0}) \lambda_{i-1} \tilde{p}_{i-1,k}(s) \\ &\quad + (1 - \delta_{i,k})(1 - \delta_{i,N}) \mu_{i+1} \tilde{p}_{i+1,k}(s) \\ &\quad + (1 - \delta_{k,k_0}) \delta_{i,k} \lambda_{i-1} \tilde{p}_{i-1,k-1}(s),\ k_0 \leq k \leq N,\ 0 \leq i \leq k. \end{aligned} \tag{3}$$

For every fixed k with $k_0 \leq k \leq N$ the system (3) is tridiagonal and can be solved using the standard forward-elimination-backward-substitution method at a low computational cost. Moreover the simplicity of the constant term $\delta_{i,i_0} \delta_{k,k_0} + (1 - \delta_{k,k_0}) \delta_{i,k} \lambda_{i-1} \tilde{p}_{i-1,k-1}(s)$ which is 0 for many pairs (i, k) allows further simplifications. After some algebra, we conclude with a stable recursive scheme that we summarize below.

Theorem 1. *The Laplace transforms $\tilde{p}_{i,k}(s)$, $k_0 \le k \le N$, $0 \le i \le k$ are computed by the equations*

$$\tilde{p}_{k_0,k_0}(s) = \frac{(s+g_{k_0-1}+\lambda_{k_0-1})\delta_{k_0,i_0}+\lambda_{k_0-1}D_{k_0-1}}{(s+\lambda_{k_0})(s+g_{k_0-1})+(s+\lambda_{k_0})\lambda_{k_0-1}+\mu_{k_0}(s+g_{k_0-1})}, \quad (4)$$

$$\tilde{p}_{i,k_0}(s) = \sum_{j=i}^{k_0-1}\frac{D_j}{\mu_{j+1}}\prod_{n=i}^{j}\frac{\mu_{n+1}}{s+g_n+\lambda_n} + \tilde{p}_{k_0,k_0}(s)\prod_{j=i}^{k_0-1}\frac{\mu_{j+1}}{s+g_j+\lambda_j},\ 0 \le i \le k_0-1, \quad (5)$$

$$\tilde{p}_{k,k}(s) = \frac{(s+g_{k-1}+\lambda_{k-1})\lambda_{k-1}\tilde{p}_{k-1,k-1}(s)}{(s+\lambda_k)(s+g_{k-1})+(s+\lambda_k)\lambda_{k-1}+\mu_k(s+g_{k-1})},\ k_0+1 \le k \le N, \quad (6)$$

$$\tilde{p}_{i,k}(s) = \tilde{p}_{k,k}(s)\prod_{j=i}^{k-1}\frac{\mu_{j+1}}{s+g_j+\lambda_j},\ k_0+1 \le k \le N,\ 0 \le i \le k-1, \quad (7)$$

where the coefficients $g_i, 0 \le i \le N-1$ and $D_i, 0 \le i \le k_0-1$ are given by the recursive scheme

$$g_0 = 0, \quad (8)$$

$$g_i = \mu_i\frac{s+g_{i-1}}{s+g_{i-1}+\lambda_{i-1}},\ 1 \le i \le N-1, \quad (9)$$

$$D_i = \delta_{i,i_0},\ 0 \le i \le i_0, \quad (10)$$

$$D_i = \prod_{j=i_0}^{i-1}\frac{\lambda_j}{s+g_j+\lambda_j},\ i_0+1 \le i \le k_0-1. \quad (11)$$

Let now M denote the MNI till absorption. We are interested in the distribution of M, given that $(I(0), M(0)) = (i_0, k_0)$. We set $y_{i,k,m} = \Pr[M = m | I(0) = i, M(0) = k]$, $0 \le i \le k \le m \le N$. Then by conditioning on the first transition out of the initial state for the process $\{(I(t), M(t)), t \ge 0\}$ (first-step analysis), we obtain the linear system

$$y_{0,k,m} = \delta_{k,m},\ 0 \le k \le m \le N, \quad (12)$$

$$y_{i,k,m} = \frac{\mu_i}{\lambda_i+\mu_i}y_{i-1,k,m}+\frac{\lambda_i}{\lambda_i+\mu_i}y_{i+1,k,m},\ 2 \le k \le m \le N,\ 1 \le i < k, \quad (13)$$

$$y_{k,k,m} = \frac{\mu_k}{\lambda_k+\mu_k}y_{k-1,k,m}+\frac{\lambda_k(1-\delta_{k,m})}{\lambda_k+\mu_k}y_{k+1,k+1,m},\ 1 \le k \le m \le N. \quad (14)$$

For any fixed pair (k, m) with $0 \le k \le m$ the system is tridiagonal and can be solved explicitly. After some algebraic manipulations we derive a closed form for the probabilities $y_{i,k,m}$.

Theorem 2. *The probabilities* $y_{i,k,m} = \Pr[M = m | I(0) = i, M(0) = k]$, $0 \le i \le k \le m \le N$, *are given by the formulas*

$$\begin{aligned} y_{i,k,m} &= \rho_{i-1}/\rho_{m-1} - \rho_{i-1}/\rho_m,\ 1 \le i \le k \le m-1,\ 1 \le m \le N-1, && (15)\\ y_{i,m,m} &= 1 - \rho_{i-1}/\rho_m,\ 1 \le i \le m \le N-1, && (16)\\ y_{i,k,N} &= \rho_{i-1}/\rho_{N-1},\ 1 \le i \le k \le N-1, && (17)\\ y_{i,N,N} &= 1,\ 1 \le i \le N, && (18) \end{aligned}$$

where

$$\rho_0 = 1,\ \ \rho_i = \sum_{s=0}^{i} \prod_{j=1}^{s} \frac{\mu_j}{\lambda_j},\ 1 \le i \le N-1. \quad (19)$$

3. Quasi-stationarity and MNI

The continuous-time Markov chain $\{(I(t), M(t)), t \ge 0\}$ given that $(I(0), M(0)) = (i_0, k_0)$ is absorbing. Indeed, with probability 1, the chain will be absorbed in some state in $A = \{(0, k_0), (0, k_0 + 1), \ldots, (0, N)\}$ while the other states in $S = \{(i, k) : k_0 \le k \le N, 1 \le i \le k\}$ are transient. Therefore, its stationary and limiting distributions give positive probabilities only to states in A. However, to quantify the behavior of the process, it is important to study what happens given that the absorption has not yet occurred, i.e. as long as the process remains in S. To this end we use the notion of quasi-stationarity.

Van Doorn and Pollett (2008) have recently studied various results concerning the quasi-stationarity issue for absorbing continuous-time Markov chains with a finite state of transient states S which is reducible. Their results generalize previously reported results concerning the case where S is irreducible (see e.g. Nasell (2001)). The general framework of van Doorn and Pollett is indispensable for our study since the set of transient states of the process of interest $\{(I(t), M(t)), t \ge 0\}$ of our model is reducible.

Observe that the subset of transient states in our case consists of $N-k_0+1$ communicating classes: $S_j = \{(i, N-j+1) : 1 \le i \le N-j+1\}$, $j = 1, 2, \ldots, N-k_0+1$. Note that with this ordering we have that S_j is accessible from $S_{j'}$ if and only if $j \le j'$. The submatrix A_Q of the transition rate matrix of $\{(I(t), M(t))\}$ corresponding to the transient states in S, under the ordering $\{(1, N), (2, N), \ldots, (N, N), (1, N-1), (2, N-1), \ldots, (N-1, N-1), \ldots, (1, k_0), (2, k_0), \ldots, (k_0, k_0)\}$, is lower block didiagonal. The $k \times k$ diagonal block corresponding to transitions within the set $\{(1, k), (2, k), \ldots, (k, k)\}$ is tridiagonal and it is identical to the corresponding submatrix of the transition rate matrix of $\{I(t)\}$. The $k \times (k+1)$ subdiagonal block corresponding to transitions from $\{(1, k), (2, k), \ldots, (k, k)\}$ to $\{(1, k+1), (2, k+1), \ldots, (k+1, k+1)\}$ has only one non-zero entry, the entry $(k, k+1)$ with value λ_k.

Since A_Q is block didiagonal, its eigenvalues are given as the union of the eigevalues of the matrices $D_N, D_{N-1}, \ldots, D_{k_0}$. Moreover, the matrix D_k is the part of D_N corresponding to its left-upper $k \times k$ submatrix. Using the so-called Given's method (see Ciarlet (1989) Theorem 6.2-2) we have that the matrix D_k

has k distinct real eigenvalues for every k with $k_0 \leq k \leq N$. Moreover, the k eigenvalues of D_k separate the $k+1$ eigenvalues of D_{k+1}, for every k with $k_0 \leq k \leq N-1$. As a corollary, we have in particular that the maximal eigenvalue of A_Q is given by the maximal eigenvalue of D_N and has algebraic multiplicity 1. Using these facts we can now apply the results of van Doorn and Pollett (2008).

Theorem 3. *Consider the submatrix of the transition rate matrix of* $\{I(t), t \geq 0\}$ *corresponding to the transient states and let* $-\alpha < 0$ *be its unique maximal eigenvalue and* $(q_1, q_2, \ldots, q_N)$ *the corresponding unique left eigenvector (the quasi-stationary distribution of* $\{I(t), t \geq 0\}$*). We also define*

$$T = \sup\{t \geq 0 : I(t) \neq 0\}, \tag{20}$$

the absorption time. Then

a) The maximal eigenvalue of the submatrix A_Q *of the transition rate matrix of* $\{(I(t), M(t), t \geq 0\}$ *corresponding to the transient states is* $-\alpha$*.*

b) There exists a unique row vector $\bar{q} = (q_{1,N}, q_{2,N}, \ldots, q_{N,N}, q_{1,N-1}, q_{2,N-1}, \ldots, q_{1,k_0}, q_{2,k_0}, \ldots, q_{k_0,k_0})$*, such that* $\bar{q} \geq \bar{0}$*,* $\bar{q}A_Q = -\alpha\bar{q}$ *and* $\bar{q}\bar{1}^T = 1$ *(with* $\bar{0}$ *and* $\bar{1}$ *being row vectors of 0s and 1s of appropriate dimensions and* $\bar{v}^T$ *denoting the transpose of a vector* $\bar{v}$*). The only non-zero entries of* $\bar{q}$ *correspond to the states in* $\{(1,N),(2,N),\ldots,(N,N)\}$*. The vector* $(q_{1,N}, q_{2,N}, \ldots, q_{N,N})$ *is exactly the quasi-stationary distribution* $(q_1, q_2, \ldots, q_N)$ *of* $\{I(t)\}$*.*

c) For any initial distribution of $(I(0), M(0))$*, where absorption has not yet occurred (i.e. there is zero probability on states with* $I(0) = 0$*) we have*

$$\lim_{t\to\infty} \Pr[T > t + s | T > t, I(0), M(0)] = e^{-\alpha s}, s > 0, \tag{21}$$

$$\lim_{t\to\infty} \Pr[(I(t), M(t)) = (i,k) | T > t, I(0), M(0)] = q_{i,k}, (i,k) \in S. \tag{22}$$

The above theorem assures that for $t \to \infty$ and given that the absorption has not yet occurred, it is certain that the MNI has reached N. Moreover, the process $\{I(t)\}$ resides in state i with probability q_i. Similar results can be derived regarding the limits (21)-(22) for T replaced by $T_m = \sup\{t \geq 0 : M(t) \leq m\}$. The corresponding limits quantify the limiting behavior of the process $\{(I(t), M(t)), t \geq 0\}$, given that the absorption has not yet occurred, nor has the MNI exceeded m.

4. A numerical study

In this section we present some numerical results for an instance of the classical SIS Model with population size $N = 50$, i.e. birth rates $\lambda_i = \beta\frac{i(50-i)}{50}$ and death rates $\mu_i = \gamma i$, $i = 0, \ldots, 50$. We consider contact rates $\beta \in \{0.05, 0.5, 1.0, 5.0, 10.0\}$ and recovery rate per customer $\gamma \in \{0.5, 1.0, 2.0, 5.0\}$. We suppose that initially there exist $I(0) = 20$ infectives. We observe the system at three different epochs $t = 0.5, 5.0$ and 50.0 and we provide in each cell, from top to bottom,the corresponding expected MNI observed up to time t.

$E[M(t)]$	β =0.05	β =0.5	β =1.0	β =5.0	β =10.0
γ =0.5	20.062795	20.956955	22.645022	40.089593	47.856359
	20.063830	21.489649	29.415663	49.226208	49.998893
	20.063830	21.563272	36.631472	49.996687	50.000000
γ =1.0	20.030921	20.418540	21.191152	36.118887	45.940009
	20.030928	20.428938	21.530456	46.481532	49.638600
	20.030928	20.428938	21.564466	48.511584	49.999990
γ =2.0	20.015228	20.176475	20.427863	29.376866	41.744306
	20.015228	20.176487	20.428938	39.183956	47.290524
	20.015228	20.176487	20.428938	42.803358	48.888843
γ =5.0	20.006036	20.063830	20.136371	21.489649	29.415663
	20.006036	20.063830	20.136371	21.563272	36.631472
	20.006036	20.063830	20.136371	21.564490	40.174717

$E[M(t)]$ is non-decreasing function of t. Moreover, for $\beta < \gamma$ the function $E[M(t)]$ is nearly constant while for $\beta > \gamma$, we observe a drastically increasing behavior. For a fixed epoch, we see that the greater the contact rates the larger the expected MNI, while longer recovery rates imply smaller expected MNI.

References

[1] Allen L.J.S. (2003) *An Introduction to Stochastic Processes with Applications to Biology*. Prentice-Hall, New Jersey.

[2] Artalejo J.R. (2008) On the transient behavior of the maximum level length in structured Markov chains. *Preprint*.

[3] Artalejo J.R., Chakravarthy S.R. (2007) Algorithmic analysis of the maximal level length in general-block two-dimensional Markov processes. *Math. Probl. Eng.* Article ID 53570, 1-15.

[4] Artalejo J.R., Economou A. and Gomez-Corral A. (2007) Applications of maximum queue lengths to call center management. *Comp. Oper. Res.* **34**, 983-996.

[5] Ciarlet P.G. (1989) *Introduction to Numerical Linear Algebra and Optimization*. Cambridge University Press, Cambridge.

[6] Nasell I. (2001) Extinction and quasi-stationarity in the Verhulst logistic model. *J. Theor. Biol.* **211**, 11-27.

[7] Neuts M.F. (1964) The distribution of the maximum length of a Poisson queue during a busy period. *Oper. Res.* **12**, 281-285.

[8] Serfozo R.F. (1988) Extreme values of birth and death processes and queues. *Stoch. Proc. Appl.* **27**, 291-306.

[9] van Doorn E.A. and Pollett P.K. (2008) Survival in a quasi-death process. *Lin. Alg. Appl.* **429**, 776-791.

Part III
Section reports

Section

Queuing and other discrete systems

6th St.Petersburg Workshop on Simulation (2009) 815-819

The $BMAP/PH/N$ queue operating in random environment

Valentina Klimenok[1][2], Valentina Khramova[3], Alexander Dudin[4], Hee Yeol Eom[5], Chesoong Kim[6]

Abstract

We consider the $BMAP/PH/N$ queueing system operating in a finite state space Markovian random environment. Sojourn time in the system is analyzed. Illustrative numerical examples are presented.

1. Introduction

Classical queueing theory assumes that the characteristics of the arrival and service process are fixed and do not change during evolution of the system. However, in many real world objects, which can be modeled in terms on the queueing theory, these characteristics can randomly vary due to some factors. This motivates investigation of so called queues operating in a *random environment* (RE). It means that that there are a queueing system and an external finite state space stochastic process called the RE. Under the fixed state of the RE, the queueing system operates as a classic queueing system of the correspondent type. However, when the RE jumps into another state, the parameters of the queueing system (inter-arrival times distribution or arrival rate, service times distribution or service rate, number of servers, retrial rate, etc) can immediately change their values.

Short overview of the recent literature devoted to investigation of queues operating in the RE can be found in [1]. In [1], the quite general $BMAP/PH/N/0$ type queueing system operating in a Markovian RE was investigated. In the present paper, we supplement analysis given in [1] for the system with losses, i.e., system having no buffer, to the case of the system with infinite buffer. Essential novelty of the present results is twofold. First, the system with losses is always stable (under reasonable assumptions on the system parameters) while it is necessary to get stability condition for the system with the infinite buffer. We prove criterion

[1]This work was supported by the Korean Research Foundation Grant Funded by the Korean Government (MOEHRD) (KRF-2008-313-D01211)

[2]Belarusian State University, E-mail: klimenok@bsu.by

[3]Belarusian State University, E-mail:valyakhramova@inbox.ru

[4]Belarusian State University, E-mail: dudin@bsu.by

[5]Sangji University, E-mail: hyeom11@hanmail.net

[6]Sangji University, E-mail: dowoo@sangji.ac.kr

for ergodicity in a nice intuitively tractable form. Second, customers in the system with a buffer can be queued and the problem of calculation of a sojourn time distribution for an arbitrary customer arises. This problem is solved in our paper.

The obtained formulas are quite involved. To demonstrate possibility of their effective realization on computer and to show importance of investigation of queues operating in the *RE*, we present numerical results in the concluding section.

2. The mathematical model

We consider N-server queueing system with an infinite waiting space. The system behavior depends on the state of the stochastic process (Random Environment -*RE*) r_t, $t \geq 0$, which is assumed to be an irreducible continuous time Markov chain with the state space $\{1, \ldots, R\}$, $R \geq 2$, and the infinitesimal generator Q.

The input flow into the system is the following modification of the Batch Markov Arrival Process ($BMAP$). In this input flow, the batch arrivals are directed by the process ν_t, $t \geq 0$, (underlying process) with the state space $\{0, 1, \ldots, W\}$. Under the fixed state r of the random environment, this process behaves as an irreducible continuous time Markov chain. Transitions of the chain ν_t, $t \geq 0$, which are accompanied by arrival of k-size batch, are described by the matrices $D_k^{(r)}$, $k \geq 0$, $r = \overline{1, R}$, with the generating function $D^{(r)}(z) = \sum\limits_{k=0}^{\infty} {D_k}^{(r)} z^k$, $|z| \leq 1$. The service process is defined by the modification of the phase (PH)-type service time distribution. Under the fixed value r of the random environment, the service process is defined by the irreducible continuous time Markov chain m_t, $t \geq 0$, with the state space $\{1, \ldots, M+1\}$ and irreducible representation $(\boldsymbol{\beta}^{(r)}, S^{(r)})$.

Our aim is to analyze the described queueing model in steady-state.

For the use in the sequel, let us introduce the following notation:

I_n ($\mathbf{e}_n$) is an identity matrix (column vector consisting of 1's) of size n.

$\otimes$ and $\oplus$ are symbols of the Kronecker product and sum of matrices;

$\Omega^{\otimes l} = \underbrace{\Omega \otimes \ldots \otimes \Omega}_{l}$, $l \geq 1$, $\Omega^{\otimes 0} = 1$, $\Omega^{\oplus l} = \sum\limits_{m=0}^{l-1} I_{n^m} \otimes \Omega \otimes I_{n^{l-m-1}}$, $l \geq 1$;

$D(z) = \sum\limits_{k=0}^{\infty} diag\{D_k^{(r)},\ r = \overline{1, R}\} z^k$;

$\mathcal{D}_k^{(n)} = diag\{D_k^{(r)} \otimes I_{M^n},\ r = \overline{1, R}\},\ n = \overline{0, N},\ k \geq 0$;

$\mathcal{D}^{(n)}(z) = \sum\limits_{k=0}^{\infty} \mathcal{D}_k^{(n)} z^k,\ n = \overline{0, N}$;

$\mathcal{B}_l^{(n)} = diag\{I_{\bar{W}} \otimes I_{M^n} \otimes \left(\boldsymbol{\beta}^{(r)}\right)^{\otimes l},\ r = \overline{1, R}\},\ n = \overline{1, N},\ \bar{W} = W + 1$;

$\mathcal{S}^{(n)} = diag\{I_{\bar{W}} \otimes \left(S^{(r)}\right)^{\oplus n},\ r = \overline{1, R}\},\ n = \overline{1, N}$;

$\mathcal{S} = diag\{S^{(r)},\ r = \overline{1, R}\}$;

$\mathcal{S}_0^{(n)} = diag\{I_{\bar{W}} \otimes \left(\boldsymbol{S}_0^{(r)}\right)^{\oplus n},\ r = \overline{1, R}\},\ n = \overline{1, N},\ \boldsymbol{S}_0^{(r)} = -S^{(r)}\mathbf{e}$;

$\bar{\mathcal{S}}_0^{(N)} = diag\{I_{\bar{W}} \otimes \left(\mathbf{S}_0^{(r)} \beta^{(r)}\right)^{\oplus N},\ r = \overline{1, R}\}$;

$\mathcal{C}^{(n)} = Q \otimes I_{\bar{W}} \otimes I_{M^n} + \mathcal{D}_0^{(n)} + \mathcal{S}^{(n)},\ n = \overline{0, N}$.

3. Stationary state distribution

It is easy to see that operation of the considered queueing model is described in terms of the regular irreducible continuous-time Markov chain

$$\xi_t = \{n_t, r_t, \nu_t, m_t^{(1)}, \dots, m_t^{(\min\{n_t, N\})}\},\ t \geq 0,$$

where n_t is the number of customers in the system, $n_t \geq 0$; r_t is the state of random environment, $r_t = \overline{1, R}$; ν_t is the state of the $BMAP$ process, $\nu_t = \overline{0, W}$; $m_t^{(n)}$ is the phase of PH service process in the nth busy server, $m_t^{(n)} = \overline{1, M},\ n = \overline{1, N}$, at epoch $t,\ t \geq 0$.

Lemma 1. *Infinitesimal generator A of the Markov chain $\xi_t,\ t \geq 0$, has the block structure, $A = (A_{i,j})_{i,j \geq 0}$, where*

$$A_{i,i} = \begin{array}{ll} \mathcal{C}^{(i)}, & i = \overline{0, N-1}, \\ \mathcal{C}^{(N)}, & i \geq N; \end{array} \qquad A_{i,i-1} = \begin{cases} \mathcal{S}_0^{(i)}, & i = \overline{1, N}, \\ \bar{\mathcal{S}}_0^{(N)}, & i \geq N+1; \end{cases}$$

$$A_{i,j} = \begin{cases} \mathcal{D}_{j-i}^{(i)} \mathcal{B}_{min\{N-i, j-i\}}^{(i)}, & j > i,\ i = \overline{0, N-1}, \\ \mathcal{D}_{j-i}^{(N)}, & j > i,\ i \geq N. \end{cases}$$

It follows from Lemma that the chain ξ_t belongs to the class of continuous time multi-dimensional Markov chain ($QTMC$) investigated in [2] Thus, in the steady state analysis of the model under study we use the results from [2] for $QTMC$.

Theorem 1. *The necessary and sufficient condition for existence of the Markov chain $\xi_t, t \geq 0$, stationary distribution is the fulfillment of the inequality*

$$\rho = \lambda / \bar{\mu} < 1,$$

where $\lambda = \mathbf{x}_1 D(z)|'_{z=1}\mathbf{e}$, $\bar{\mu} = \mathbf{x}_2 diag\{\left(\boldsymbol{S}_0^{(r)}\right)^{\oplus N}, r = \overline{1, R}\}\mathbf{e}$, and the vectors $\mathbf{x}_n, n = 1, 2$, are defined as the unique solutions to the following systems:

$$\mathbf{x}_1 (Q \otimes I_{\bar{W}} + D(1)) = \mathbf{0}, \mathbf{x}_1 \mathbf{e} = 1,$$

$$\mathbf{x}_2 (Q \otimes I_{M^N} + diag\{\left(\boldsymbol{S}^{(r)} + \boldsymbol{S}_0^{(r)} \boldsymbol{\beta}^{(r)}\right)^{\oplus N}, r = \overline{1, R}\}) = \mathbf{0},\ \ \mathbf{x}_2 \mathbf{e} = 1.$$

Let us enumerate the states of the chain $\xi_t,\ t \geq 0$, in the lexicographic order and form the row vectors $\mathbf{p}_i$ of probabilities corresponding to the state i of the first component of the process ξ_t. To compute the stationary probability vectors $\mathbf{p}_i,\ i \geq 0$, we use the effective stable procedure [2] based on the special structure of the matrix A. Having these vectors been computed, we can calculate a number of performance measures of the system and the distribution of the sojourn time in the system.

4. Sojourn time distribution

Theorem 2. *The Laplace-Stieltjes transform $v_a(s)$ of the sojourn time distribution of an arbitrary customer in the system is calculated as follows*

$$v_a(s) = \frac{1}{\lambda} \{ \sum_{i=0}^{N-1} \mathbf{p}_i \sum_{k=1}^{\infty} min\{k, N-i\} \mathcal{D}_k^{(i)} (I_R \otimes \mathbf{e}_{\bar{W} M^i}) +$$

$$\sum_{i=0}^{\infty}\mathbf{p}_i\sum_{k=\varphi_i+1}^{\infty}\mathcal{D}_k^{(min\{i,N\})}\mathcal{B}^{(\varphi_i)}\sum_{l=\varphi_i+1}^{k}(\mathcal{F}(s))^{i-N+l}(I_R\otimes\mathbf{e}_{M^N})\}\mathcal{H}(s)\mathbf{e}_R,$$

where

$$\mathcal{H}(s)=diag\{\boldsymbol{\beta}^{(r)},r=\overline{1,R}\}(sI-(Q\otimes I_M+\mathcal{S}))^{-1}diag\{\mathbf{S}_0^{(r)},r=\overline{1,R}\}dt,$$

$$\mathcal{F}(s)=(sI-(Q\otimes I_{M^N}+diag\{(S^{(r)})^{\oplus N},r=\overline{1,R}\}))^{-1}diag\{(\mathbf{S}_0^{(r)}\boldsymbol{\beta}^{(r)})^{\oplus N},r=\overline{1,R}\},$$

$$\mathcal{B}^{(n)}=diag\{\mathbf{e}_{\bar{W}}\otimes I_{M^{N-n}}\otimes(\boldsymbol{\beta}^{(r)})^{\oplus n},r=\overline{1,R}\},n=\overline{0,N},\ \ \varphi_i=max\{0,N-i\}.$$

Theorem 3. *The mean sojourn time $\bar{v}_a$ of an arbitrary customer in the system is calculated from the formula*

$$\bar{v}_a=-\frac{1}{\lambda}\{\sum_{i=0}^{N-1}\mathbf{p}_i\sum_{k=1}^{\infty}min\{k,N-i\}\mathcal{D}_k^{(i)}(I_R\otimes\mathbf{e}_{\bar{W}M^i})\mathcal{H}^{'}(0)+$$

$$\sum_{i=0}^{\infty}\mathbf{p}_i\sum_{k=\varphi_i+1}^{\infty}\mathcal{D}_k^{(min\{i,N\})}\mathcal{B}^{(\varphi_i)}\times$$

$$\sum_{l=\varphi_i+1}^{k}[\sum_{m=0}^{i+l-N-1}(\mathcal{F}(0))^m\mathcal{F}^{'}(0)(I_R\otimes\mathbf{e}_{M^N})+(\mathcal{F}(0))^{i-N+l}(I_R\otimes\mathbf{e}_{M^N})\mathcal{H}^{'}(0)]\}\mathbf{e}.$$

5. Numerical examples

The goal of the numerical experiments is to demonstrate the feasibility of the proposed algorithm and to give insight into behavior of the considered queueing system. We consider the $BMAP/PH/N$ systems operating in the RE which has two states ($R=2$). The generator of the RE is $Q=\begin{matrix}-5 & 5\\ 15 & -15\end{matrix}$. The number of servers is $N=3$.

We use two $BMAP$s for description of the arrival process. Both they have fundamental rate 3.488, squared coefficient of variation 4, probability that the batch consists of k customers is equal to $q^{k-1}\frac{1-q}{1-q^K},\ k=\overline{1,K},\ q=0.9,\ K=5$. These $BMAP$s have different coefficients of correlation of inter-arrival times. The $BMAP_1, BMAP_2$ have coefficients of correlation 0.2, 0.3 respectively. We use three service processes $PH_r,\ r=\overline{1,3}$, having the mean rates of service $\mu^{(1)}=2$, $\mu^{(2)}=14,\ \ \mu^{(3)}=28$. The squared coefficients of variation of the service time distribution are 0.5, 1.24, 1.29 respectively.

In the **first experiment** we compare the value of mean number of customers and mean sojourn time of an arbitrary customer in the $BMAP/PH/N$ system operating in the RE and in simpler queueing system which can be considered as its simplified analog. This analog is the $M^X/PH/N$ system in the RE where, under the fixed value of the RE, the input flow is a group stationary Poisson with the same batch size distribution and intensity equal to fundamental rate of the corresponding $BMAP$ in the original system. Input flow is described by the $BMAP_1$ and the $BMAP_2$. Service processes are the PH_1 and PH_2.

Table 1 shows the values of the mean numbers of customers L and the mean sojourn time $\bar{v}_a$ in the original system $BMAP/PH/N$ in the RE and its analog for different value of ρ. Here and in the sequel the variation of the system load ρ is performed by means the scaling the value of the fundamental rate λ.

Table 1: Mean number of customers and mean sojourn time

		$BMAP/PH/N$ in RE		$M^X/PH/N$ in RE	
λ	ρ	L	V_a	L	V_a
0.74785	0.05016	0.22534	0.30131	0.22859	0.30566
3.01889	0.20247	1.39632	0.46253	1.02846	0.34067
5.22197	0.35022	4.44618	0.85144	2.0578	0.39407
7.53409	0.50528	11.27939	1.49712	3.68104	0.48858
9.67354	0.64877	23.67705	2.44761	6.33356	0.65473
12.0344	0.80711	59.18059	4.91761	13.6235	1.13205
14.1111	0.94638	262.6847	18.6168	55.0359	3.90018

Two conclusions follow from the Table:
(i) The considered approximation can be quite bad.
(ii) Little's formula is valid. Note that this fact is confirmed by the results of all numerical experiments that were performed for the system under study.

Thus, we see that an approximation of the system performance measures by means of their values in simpler queueing system can be bad. However, it is intuitively clear the following. If the random environment is "very slow" (the rate of the RE is essentially less then the rates of the input flow and the service processes), an approximation called below as "mixed system" can be applied successfully. This approximation consists of calculation of the performance measures under the fixed states of the RE and their averaging by the RE distribution. If the random environment is "very fast", approximation called below as "mixed parameters" can be successfully applied. This approximation consists of averaging parameters of the arrival and service processes by the distribution of the RE and calculation of the performance measures in the system with averaged arrival and service rates.

In the **second experiment**, we show numerically that these approximations make sense. However, in situations when environment is neither "very slow" or "very fast", these approximations can be very poor. We consider the REs with different rates which are characterized by the generators of the form $Q^{(k)} = Q \cdot 10^k$.

We vary the parameter k from -7 to 4 what corresponds to the variation of the RE rate from "very slow" to "very fast". The input flow is described by the $BMAP_1$ and $BMAP_2$ and the service process is defined by the PH_1 and PH_3. The results are presented on the left picture of *Figure 1.* This picture confirms the hypothesis that the first type approximation "mixed system" is good in case of "very slow" RE and the second one can be applied to case of "very fast" RE. But, there is an interval for RE rate (interval [-3,0]) where we can not use the estimates for mean sojourn time calculated by the considered approximating models. The utilization of such estimates can lead to large value of relative error. *Figure 1*

confirms the importance of investigation implemented in this paper. Simple engineering approximations can lead to unsatisfactory performance evaluation and capacity planning in real world systems.

The idea of the **third experiment** is the following. Let us assume that the RE has two states. One state corresponds to the peak traffic periods, the second one corresponds to the normal traffic periods. Service times during these periods are defined by the PH_1 and PH_2 distributions. Arrivals during these periods are defined by the stationary Poisson flow with the rates λ_1 and λ_2 correspondingly and initially we assume that $\lambda_1 >> \lambda_2$. It is intuitively clear that if it is possible to redistribute the arrival processes (i.e., to reduce the arrival rate during the peak periods and to increase it correspondingly during the normal traffic periods) without changing the average arrival rate, the mean numbers of call in the system can be reduced. In real life system such a redistribution is sometimes possible, e.g., by means of controlling tariffs during the peak traffic periods. The goal of this experiment is to show that this intuitive consideration is correct and to illustrate the effect of the redistribution.

We assume that the averaged arrival rate λ should be 12.5 and consider four different situations: a huge difference of arrival rates $\lambda_1 = 50\lambda_2$, a very big difference $\lambda_1 = 10\lambda_2$, a big difference $\lambda_1 = 3\lambda_2$ and equal arrival rates $\lambda_1 = \lambda_2$. The generator of the random environment is as follows $Q = \begin{matrix} -15 & 15 \\ 5 & -5 \end{matrix}$.

It can be seen from the right picture of *Figure 1* that smoothing the peak rates can cause essential decrease of the mean numbers of call in the system.

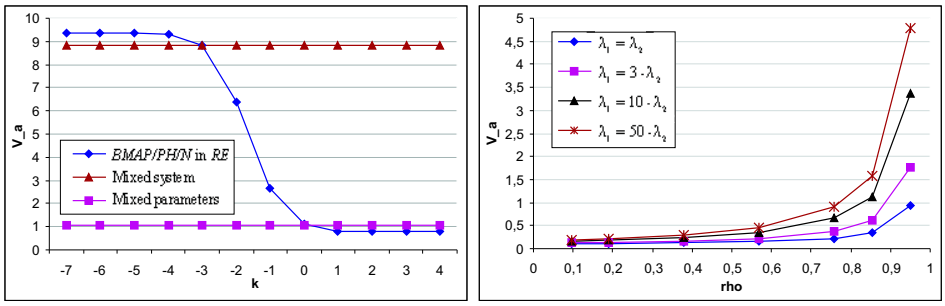


Figure 1: Mean sojourn time

References

[1] Kim CS, Dudin A, Klimenok V, Khramova V. Erlang loss queueing system with batch arrivals operating in a random environment. Computers and Operations Research. 2009; 36: 674-697.

[2] Klimenok VI, Dudin AN. Multi-dimensional asymptotically quasi-Toeplitz Markov chains and their application in queueing theory. Queueing Systems 2006;54: 245-259.

6th St.Petersburg Workshop on Simulation (2009) 821-825

Steady-state analysis of a dual tandem queue with arbitrary inter-arrival time distribution

Valentina Klimenok[1,2], Che Soong Kim[3], Olga Taramin[4]

Abstract

We consider a tandem queue consisting of two single-server stations. The first station is represented by a single-server queue with waiting room and arrival flow forming a renewal process. After service at the first station a customer proceeds to the second station. No queue is allowed between the servers, so that a customer having completed processing at the first station and meeting the busy second server is forced to wait at the first station occupying the server space until the second server becomes available. Thus, the first server becomes blocked or not available for service of incoming customers.

We derive the stationary performance measures of the system and analyze them by means of a numerical experiment.

1. Introduction

Tandem queueing system can be used for modeling real-life two-node networks as well as for the validation of general decomposition algorithms in networks. Thus, tandem queueing systems have found much interest in the literature. An extensive survey of early papers on tandem queues can be seen in [4]. Most of these papers are devoted to exponential queueing models. In the past decade tandem queues with a batch Markovian arrival process have attracted a considerable interest among researches. The relevant references can be found in [2]. The main feature of the service process in tandem queues is blocking after service. One can found a number of papers devoted to tandem queues with blocking and Markovian input flow. At the same time we can mention only the papers by Avi-Itzhak and co-authors, see, e.g., [3], dealing with tandem queues with blocking and non-Markovian input. In those papers tandem queues with arbitrary input and regular service time are considered. To the best of our knowledge tandem queues with arbitrary input, blocking and Markovian service process have not studied in the literature even in the case of exponential service time distribution.

In the present paper we consider a dual tandem queue with blocking, renewal input and PH (Phase-type) service time distribution at both stations.

[1]This work was supported by the Korean Research Foundation Grant Funded by the Korean Government (MOEHRD) (KRF-2008-313-D01211)

[2]Belarusian State University, E-mail: klimenok@bsu.by

[3]Sangji University, E-mail: dowoo@sangji.ac.kr

[4]Belarusian State University, E-mail: taramin@mail.ru

2. The mathematical model

The first station of the system under consideration is represented by the $GI/PH/1$ queue. The inter-arrival times at the station are independent random variables with general distribution $A(t)$ and the finite first moment $a_1 = \int\limits_0^\infty t dA(t)$.

After service at the first station a customer proceeds to the second station that is represented by a single-server queue without a buffer. A customer having completed processing at the first station and meeting the busy second server is forced to wait at the first station occupying the server space until the second server becomes available. Thus, the first server becomes blocked or not available for service of incoming customers. We assume that a customer caused the blocking stays at the first server until the end of the blocking.

The service times of a customer at both stations have PH distributions. Service time having PH distribution with an irreducible representation $(\boldsymbol{\beta}, S)$ can be interpreted as a time until an underlying Markov process $m_t, t \geq 0$, with a finite state space $\{1, \ldots, M, *\}$ reaches the single absorbing state $*$ conditional the initial state of this process is selected among the states $\{1, \ldots, M\}$ according to probabilistic vector $\boldsymbol{\beta}$. Transition rates of the process m_t within the set $\{1, \ldots, M\}$ are defined by the sub-generator S and transition rates into the absorbing state are given by the entries of the column vector $\boldsymbol{S}_0 = -S\mathbf{e}$. Hereinafter $\mathbf{e}$ is a column vector of 1's. For more information about PH distributions see, e.g.,[4].

We assume that service process at the rth, $r = 1, 2$, server has PH_r distribution with an irreducible representation $(\boldsymbol{\beta}_r, S^{(r)})$ and is governed by the Markov chain $m_t^{(r)}$, $t \geq 0$, with the state space $\{1, \ldots, M_r, *^{(r)}\}$ where the state $*^{(r)}$ is an absorbing one.

Our aim is to investigate of the proposed queueing system in steady-state.

3. Stationary distribution. Performance measures

Let t_n be the time of the nth arrival at the first station and i_n be the number of customers at this station (including the blocked customer, if any) at the epoch $t_n - 0$. It is easy to see that the process i_n, $n \geq 1$, is non-Markovian.

To be able to construct an embedded Markov chain we need to know the distribution of the number of customers proceeding to the second station during the inter-arrival time. In contrast to the ordinary $GI/PH/1$ queue where the similar problem is solved by introducing the additional component m_n that denotes the phase of service process at the epoch $t_n + 0$, $n \geq 1$, in the queue under consideration the calculation of such a distribution is hard problem due to the blocking phenomenon. It is explained by the fact that we do not know a priory the real sojourn time of a tagged customer at the first server since it depends on whether or not the second server is free at the epoch of the service completion at the first server.

To solve the problem we introduce into consideration the notion of generalized service time. The generalized service time of a tagged customer is just the service

time at the first server if the server of the second station is free at the service completion epoch at the first server. In opposite case the blocking occurs and the generalized service time consists of the service time of the tagged customer at the first server and the time during which the first server is blocked (waits until the second server becomes free).

Lemma 1. *The generalized service time has the PH-type distribution with an irreducible representation* $(\boldsymbol{\beta}, S)$, *where*

$$\boldsymbol{\beta} = (\boldsymbol{\beta}_1 \otimes \boldsymbol{\beta}_2, \mathbf{s}0_{M_1+M_2}),$$

$$S = \begin{pmatrix} S^{(1)} \oplus S^{(2)} & I_{M_1} \otimes S_0^{(2)} & S_0^{(1)} \otimes I_{M_2} \\ 0 & S^{(1)} & 0 \\ 0 & 0 & S^{(2)} \end{pmatrix}, \boldsymbol{S}_0 = -S\mathbf{e}.$$

Here $\oplus$ *and* $\otimes$ *are symbols of Kronecker's sum and product respectively,* I_a *is an identity matrix of size* a.

Now we are able to construct the embedded Markov chain describing the queue under consideration. Let m_n *be the phase of generalized service time at the epoch* $t_n + 0$, $n \geq 1$. *It is easy to see that the process* $\xi_n = \{i_n, m_n\}$, $n \geq 1$, *is an irreducible Markov chain with the state space* $\{(0,m), m = 1, \ldots, K_0; (i,m), i > 0, m = 1, \ldots, K\}$ *where* $K = M_1M_2 + M_1 + M_2$ *and* $K_0 = M_1M_2 + M_1$. *In the following we will assume that the states of the chain* ξ_n *are enumerated in the lexicographic order.*

Lemma 2. *The transition probability matrix of the chain* ξ_n, $n \geq 1$, *has the following block structure*

$$P = \begin{pmatrix} \tilde{B}_0 & C_0 & 0 & 0 & \ldots \\ B_1 & A_1 & A_0 & 0 & \ldots \\ B_2 & A_2 & A_1 & A_0 & \ldots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix},$$

where

$$A_n = \int_0^\infty P(n,t)dA(t), \; B_n = \int_0^\infty \int_0^t P(n,x)\boldsymbol{S}_0 dx \boldsymbol{\beta}^*(t-x)dA(t), \; n \geq 0,$$

$$\tilde{B}_0 = \tilde{I}B_0, \; C_0 = \tilde{I}A_0, \; \boldsymbol{\beta}^*(y) = (\boldsymbol{\beta}_1 \otimes \boldsymbol{\beta}_2 e^{S^{(2)}y}, \; \boldsymbol{\beta}_1(1 - \boldsymbol{\beta}_2 e^{S^{(2)}y}\mathbf{e})),$$

$$\tilde{I} = \begin{pmatrix} I_{M_1M_2} & 0 & 0 \\ 0 & I_{M_1} & 0_{M_1 \times M_2} \end{pmatrix},$$

and $P(n,t)$, $n \geq 0$, *are defined by* $\sum\limits_{n=0}^{\infty} P(n,t)z^n = e^{(S+\boldsymbol{S}_0\boldsymbol{\beta}z)t}$.

Corollary 1. *The process* ξ_n *is of the* $GI/M/1$ *type Markov chain, see [4].*

Theorem 1. *Stationary distribution of the Markov chain* ξ_n, $n \geq 1$, *exists if and only if the inequality*

$$\rho = a_1^{-1} b_1^{(g)} < 1$$

is fulfilled. Here $b_1^{(g)} = -\boldsymbol{\beta} S^{-1} \mathbf{e}$ *is the mean value of generalized service time.*

Denote the stationary state probabilities of the chain ξ_n by $\pi(0, m,)$, $i \geq 0$, $m = \overline{1, K_0}$, $\pi(i, m,)$, $i > 0$, $m = \overline{1, K}$. Introduce the notation for row vectors of these probabilities

$$\boldsymbol{\pi}_0 = (\pi(0,1), \pi(0,2), \ldots, \pi(0,K_0)), \ \boldsymbol{\pi}_i = (\pi(i,1), \pi(i,2), \ldots, \pi(i,K)), \ i > 0.$$

Theorem 2. *The stationary probability vectors* $\boldsymbol{\pi}_i$, $i \geq 0$, *are calculated as follows:*

$$\boldsymbol{\pi}_i = \boldsymbol{\pi}_1 \mathcal{R}^{i-1}, \ i \geq 2,$$

where the matrix $\mathcal{R}$ *is the minimal non-negative solution of the matrix equation*

$$\mathcal{R} = \sum_{j=0}^{\infty} \mathcal{R}^j A_j,$$

and the vector $(\boldsymbol{\pi}_0, \boldsymbol{\pi}_1)$ *is the unique solution of the system*

$$(\boldsymbol{\pi}_0, \boldsymbol{\pi}_1) = (\boldsymbol{\pi}_0, \boldsymbol{\pi}_1) T, \ \boldsymbol{\pi}_0 \mathbf{e} + \boldsymbol{\pi}_1 (I - \mathcal{R})^{-1} \mathbf{e} = 1,$$

where

$$T = \begin{pmatrix} \tilde{B}_0 & C_0 \\ \sum\limits_{j=1}^{\infty} \mathcal{R}^{j-1} B_j & \sum\limits_{j=1}^{\infty} \mathcal{R}^{j-1} A_j \end{pmatrix}.$$

Having the stationary distributions $\boldsymbol{\pi}_i$, $i \geq 0$, been calculated we can find a number of stationary performance measures of the system under consideration. Some of them are calculated as follows:

• Mean number of customers at the first station at the arrival epoch

$$L = \boldsymbol{\pi}_1 (I - \mathcal{R})^{-2} \mathbf{e};$$

• Probability that both servers are busy at the arrival epoch

$$P_{busy}^{(1,2)} = \boldsymbol{\pi}_1 (I - \mathcal{R})^{-1} diag\,\{I_{M_1 M_2}, 0_{M_1 + M_2}\} \mathbf{e}.$$

• Probability that the server of the first station is busy and the server of the second station is idle at the arrival epoch

$$P_{busy}^{(1)} = \boldsymbol{\pi}_1 (I - \mathcal{R})^{-1} diag\,\{0_{M_1 M_2}, I_{M_1}, 0_{M_2}\} \mathbf{e}.$$

• Probability that the first server is blocked at the arrival epoch

$$P_{block} = \boldsymbol{\pi}_1 (I - \mathcal{R})^{-1} diag\,\{0_{M_1 M_2}, 0_{M_1}, I_{M_2}\} \mathbf{e}.$$

Remark. Using the stationary distributions $\boldsymbol{\pi}_i$, $i \geq 0$, we have also found the steady state distribution of the system states at an arbitrary time and the correspondent performance measures $\tilde{L}$, $\tilde{P}_{busy}^{(1,2)}$, $\tilde{P}_{busy}^{(1)}$, $\tilde{P}_{block}$. These results are not presented here due to limitation on the paper space.

4. Sojourn time

Theorem 3. *The Laplace-Stieltjes transform of the stationary distribution of the actual sojourn time has the following form:*
(*i*) *at the first station:*

$$v_1(\theta) = [\boldsymbol{\pi}_0\tilde{I} + \boldsymbol{\pi}_1(I - \mathcal{R}\varphi_g(\theta))^{-1}\varphi_g(\theta)](\theta I - S)^{-1}\boldsymbol{S}_0;$$

(*ii*) *in the whole system:*

$$v(\theta) = v_1(\theta)\varphi_2(\theta)).$$

Here $\varphi_g(\theta)) = \boldsymbol{\beta}(\theta I - S)^{-1}\boldsymbol{S}_0$ and $\varphi_2(\theta)) = \boldsymbol{\beta}_2(\theta I - S^{(2)})^{-1}\boldsymbol{S}_0^{(2)}$ are the Laplace-Stieltjes transforms of distribution of generalize service time and service time at the second station respectively.

Corollary 2. *The mean actual sojourn time is calculated as follows:*
(*i*) *at the first station:*

$$\bar{v}_1 = [\boldsymbol{\pi}_0\tilde{I} + \boldsymbol{\pi}_1(I - \mathcal{R})^{-1}](-S)^{-1}\mathbf{e} + Lb_1^{(g)};$$

(*ii*) *in the whole system:*

$$\bar{v} = \bar{v}_1 + b_1^{(2)},$$

where $b_1^{(2)} = -\boldsymbol{\beta}_2(S^{(2)})^{-1}\mathbf{e}$ *is the mean value of service time at the second station.*

5. Numerical examples

In presented numerical examples we investigate an impact of variation in the input process on the system performance measures.

Consider five input processes with different coefficients of variation (c_{var}) and the same value $a_1 = 0.1$ of an inter-arrival time. The *first process* is coded as D and corresponds to the deterministic distribution ($c_{var} = 0$). The *second process* is coded as U and corresponds to the uniform distribution in interval $[0.05, 0.15]$ ($c_{var} = 0.28$). The *third process* is coded as E and corresponds to the Erlangian distribution of order 4 with parameter 40 ($c_{var} = 0.5$). The *fourth process* is coded as M and corresponds to the exponential distribution ($c_{var} = 1$). The *fifth service process* is coded as HM_2. It corresponds to hyper-exponential distribution of order 2 defined by the vector (0.05, 0.95) and the intensities (0.62, 49) ($c_{var} = 5$).

Phase-type service time distributions are defined as follows:

$$PH_1: \quad \boldsymbol{\beta}_1 = (1,\, 0), \quad S^{(1)} = \begin{pmatrix} -20 & 20 \\ 0 & -20 \end{pmatrix}; \quad b_1 = 0.1,\, c_{var} = 0.71;$$

$$PH_2: \quad \boldsymbol{\beta}_2 = (0.2,\, 0.8), \quad S^{(2)} = \begin{pmatrix} -9 & 2.7 \\ 4.5 & -18 \end{pmatrix}; \quad b_1 = 0.1,\, c_{var} = 1.08.$$

Figure 1 shows the system characteristics $\bar{v}_1$, P_{block}, $\tilde{P}_{busy}^{(1)}$, $\tilde{P}_{busy}^{(1,2)}$ as functions of the system load ρ for five input flows introduced above. The value of ρ is varied by means of change of the mean value a_1 of inter-arrival time in the interval $[0.145, 1.2]$ by scaling the time. Note that the coefficients of variation do not change

under such a scaling. It is seen from the figure that variation in input flow has a great impact on the characteristics of the system. The mean value of sojourn time $\bar{v}_1$ and the probabilities P_{block}, $\tilde{P}_{busy}^{(1,2)}$ increase when the variation increases while the probability $\tilde{P}_{busy}^{(1)}$ decreases.

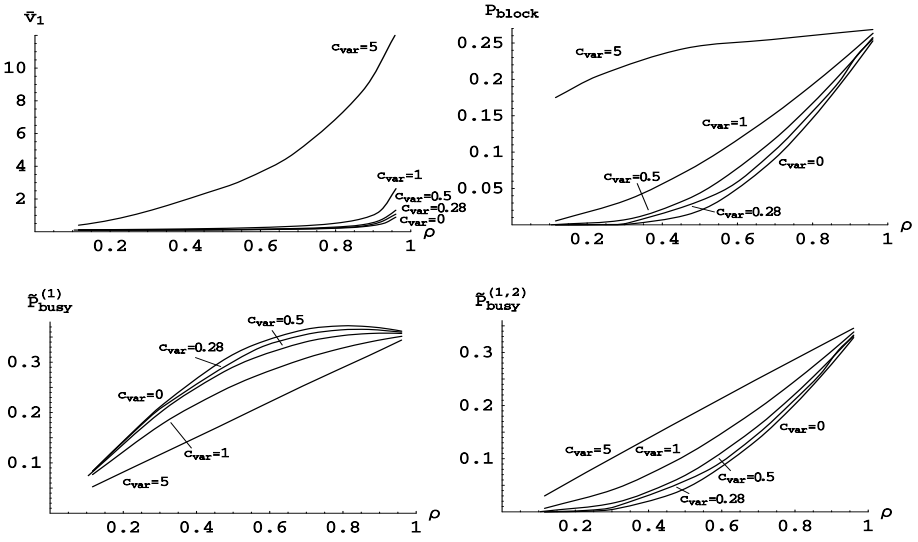


Figure 1: The system performance measures as functions of the system load ρ for different variation in the input process

6. Conclusion

In this paper, the $GI/PH/1 \rightarrow \bullet/PH/1/0$ tandem queue with blocking is studied. The condition for the existence of the stationary distribution is derived and the algorithms for calculating the steady state probabilities are presented. The Laplace-Stieltjes transforms of the distribution of the actual sojourn time at both stations as well as in the whole system are derived. Formulas for the mean values of these times are presented. The results of this paper can be applied to areas such as capacity planning, performance evaluations, and optimization of real-life production lines, two-node networks, etc.

References

[1] Gnedenko BW, Konig D. Handbuch der Bedienungstheorie. Berlin: Akademie Verlag, 1983.

[2] Gomez-Corral A, Martos ME. Performance of two-station tandem queues with blocking: The impact of several flows of signals. Performance Evaluation 2006; 63: 910-938.

[3] Avi-Itzhak B., Levy H. A sequence of servers with arbitrary input and regular service time revisited. Management science. 1995; 41: 1039-1047.

[4] Neuts MF. Matrix-Geometric Solutions in Stochastic Models - An Algorithmic Approach. Johns Hopkins University Press, 1981.

6th St.Petersburg Workshop on Simulation (2009) 827-831

Analysis of a Multi-Class Discrete-time Queueing System under the Slot-Bound Priority rule

Sofian De Clercq[1], Bart Steyaert[1], Herwig Bruneel[1]

Abstract

In this paper, we consider the dual-class discrete-time GI/G/1 queue with slot-bound priority service order, meaning that class-1 customers have priority over class-2 customers that arrive during the same slot; customers that arrive during different slots however are served on a FCFS basis. We demonstrate that the introduction of the concept of *groups* allows us to analyse the system, and as a result, the joint probability generating function (pgf) of both types of customers in the queue at random slot marks is derived.

1. Introduction

Multi-class queueing systems, or queues buffering multiple types of customers, have been widely adopted in queueing theory to model non-identical behaviour of customers. In a multiclass environment, virtually any combination of features with respect to the arrival characteristics, service requirements, buffer management rules that pertain to the individual classes (Fiems[1]) can be considered.

In this paper we study a two-class discrete-time $GI_2/G_2/1$ queueing system with infinite waiting room, under the slot-bound priority rule. The slot-bound priority rule dictates that no customer can get served before any other customer that joined the queue prior to the former customer's arrival slot. In addition, high priority (i.e. class 1) customers receive preferential treatment over low (class 2) priority customers that have arrived during the same slot. In this sense the high-priority class has limited priority over the lower one. The purpose of this study is to calculate the joint pgf of the number of type-j ($j = 1, 2$) customers in the queue at the beginning of an arbitrary slot.

In the related literature, multi-class queueing systems in discrete-time have been studied under various priority-type rules. Some of these studies include non-preemptive priority scheduling (e.g. Walraevens[2], Fiems[1], Ndreca[3]), gated priority (e.g. Stavrakakis[4]), or even simple FCFS (e.g. VanHoudt[5]). The latter studied an n-class FCFS discrete-time queueing system for the specific case of *MMAP[K]/PH[K]/1*. To the best of our knowledge however, nothing quite like the slot-bound priority rule presented here, has been studied before.

[1]SMACS Res. Group, Ghent University, Belgium, E-mail: sdclercq@telin.ugent.be

For the system under investigation one encounters the same intricate problem as Takine[6] mentioned having in a multi-class FCFS continuous-time setting: "It is widely recognized that the queue length distribution in a FIFO queue with multiple non-Poissonian arrival streams having different service time distributions is very hard to analyze, since we have to keep track of the complete order of customers in the queue to describe the queue length dynamics". We will see that our approach will suffice to deliver a discrete-time solution to this problem.

2. Definitions and Concepts

In a single server system the slot-bound priority rule states that customers get served in order of their respective arrival slots. In addition, when multiple customers of different types enter the system during the same slot, those of class 1 are served first. Now we define $a_{j,n}$ as the number of type-j customers ($j = 1, 2$) that enter the system during slot n. For all n, the discrete random variables (drv's) $a_{1,n}$ and $a_{2,n}$ may be correlated. On the other hand, $\{(a_{1,n}, a_{2,n})|n \geq 0\}$ is assumed to be a set of i.i.d. discrete random vectors with joint pgf equal to $A(z_1, z_2) = E[z_1^{a_1} z_2^{a_2}]$, which is independent of n. All service times are independent and the drv s_j, $(j = 1, 2)$ will be used to indicate a generic service time of a type-j customer, and $S_j(z) = E[z^{s_j}]$ for its pgf. Consequently, if we define μ_X as the expected value of the drv X, then $\rho_j = \mu_{a_j}\mu_{s_j}$ represents the fraction of slots during which a type-j customer is served. Hence, the workload ρ of the system satisfies $\rho = \rho_1 + \rho_2$. Due to the i.i.d. nature of the arrival and departure process described above, this system will reach a stochastic equilibrium if $\rho < 1$, which we assume to be satisfied in the remainder.

In a first part of the analysis we will use the concept of *customer groups*, much like the groups introduced in Stavrakakis[4]. If during a certain slot one or more customers enter the system, we collect them into one group (of customers), and call the event a grouparrival. So at most one group can enter the system each slot. $A_g(z)$, pgf of $a_{g,n}$, the number of groups entering the system during slot n, is naturally independent of n because of the i.i.d. property of the number of customer arrivals during consecutive slots. Furthermore a group's service time, s_g (with pgf $S_g(z)$), is the combined service time of all customers making up that group, and a group leaves the queue when its last customer does. The number of groups left behind behind by an arbitrary leaving group is denoted by d (with pgf $D(z)$). The number of type-j costumers of each type making up an arbitrary group will be denoted by b_j, and b_1 and b_2 have a joint pgf given by $B(z_1, z_2)$. For $A_g(z)$, $B(z_1, z_2)$ and $S_g(z)$, we find the following relations:

$$A_g(z) = A(0,0) + (1 - A(0,0))z \tag{1}$$

$$B(z_1, z_2) = \frac{A(z_1, z_2) - A(0,0)}{1 - A(0,0)} \tag{2}$$

$$S_g(z) = B(S_1(z), S_2(z)). \tag{3}$$

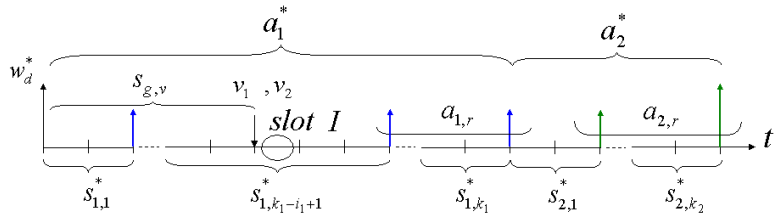


Figure 1: Service time of group K. We assumed $d^* > 0$ and $b_j^* = k_j$.

Notice that, by grouping customers like this, we effectively aggregate the customers that are affected by the priority rule, leaving the groups to be served in FCFS order. We can now rely on the results of Bruneel[7] to derive that $D(z)$ satisfies

$$D(z) = (1-\rho)\left(1 + z\,\frac{S_g A_g(z) - 1}{z - S_g A_g(z)}\right), \tag{4}$$

where $A_g(z)$ and $S_g(z)$ satisfy (1) and (3) for the system with slot-bound priority, and where we will adopt the notation $XY(z) \equiv X(Y(z))$ in the remainder.

3. Analysis

The purpose of this analysis is to determine the equilibrium distribution of (v_1, v_2), the number of type-1 and type-2 customers in the queue at the beginning of a random slot, in the form of their joint pgf $V(z_1, z_2) = E[z_1^{v_1} z_2^{v_2}]$. To this end we will relate d to v_j, because we already know the pgf of the former. We use an auxiliary drv to represent the number of groups in the queue just after the group departure preceding an arbirtrary slot (d^* for short). Clearly d and d^* do not necessarily have the same distribution, but what we do find is that they are related by

$$\Pr[d = i] = \Pr[d^* = i | v_1 + v_2 > 0], \quad \forall i \geq 0, j \in \{1, 2\}, \tag{5}$$

since, if we omit the idle periods of the system, selecting an arbitrary slot is equivalent to selecting the group departure that precedes it in a random fashion (as d was selected). We can now express v_j in terms of d^* as follows. Let us select an arbitrary slot during which the system is nonempty. We name this slot, slot I, and let us denote by group K, the customer group that is being served during slot I. Based on the above conventions and assumptions, we may then write that (see Fig. 1)

$$v_j = \sum_{i=1}^{(d^*-1)^+} b_{j,i} + \sum_{i=1}^{f} a_{j,i} + r_j, \quad j \in \{1, 2\}, v_1 + v_2 > 0. \tag{6}$$

In the right-hand side of (6), $b_{j,i}$ represents the number of type-j customers in the i-th group left behind in the queue by the previous group departure (hence,

the joint pgf of $(b_{1,i}, b_{2,i})$ is given by $B(z_1, z_2)$). The first sum evaluates all type-j customers in the buffer from groups present in the queue just after aforementioned group departure, except for group K, whose remaining type-j customers at the beginning of slot I are represented by the drv r_j (with $x^+ = max(0, x)$). Furthermore type-j customers present in the system at the beginning of slot I can also have arrived during the already elapsed service time of group K (here denoted by the drv f). With $a_{j,i}$ we mean to indicate the number of type-j arrivals during the i-th slot of this service time (with $A(z_1, z_2)$ as the joint pgf of $(a_{1,i}, a_{2,i})$).

Note that all drvs that appear in the right-hand of (6) are statistically independent, except for (r_1, r_2, f), and the main part of the analysis is devoted to calculating their joint pgf:

$$H(x_1, x_2, z) \stackrel{\triangle}{=} E[x_1^{r_1} x_2^{r_2} z^f] = H_1(x_1, x_2, z) + H_2(x_1, x_2, z),$$

where the sum portrays a conditioning on the type of customer being served during slot I. The index j in the partial joint pgf's $H_j(x_1, x_2, z)$ reflects that customer's type. The order in which customers of the same group are served will give birth to an asymmetry between these functions.

The set (r_1, r_2, f) is stochastically dependent on the number of customers of each type in group K, denoted by (b_1^*, b_2^*), their respective service times, denoted by $s_{j,i}^*$ (for the service time of the i-th type -j customer in group K), and the type of customer being served during slot I (see Fig. 1). Observe that b_j^* and $s_{j,i}^*$ do not necessarily correspond to the number of type-j arrivals in an arbitrary slot and the service time of a random type-j customer, since a randomly selected slot has a tendency to belong to longer group service times. We can reason that, in view of the slot-bound priority rule, upon further conditioning on these parameters, the following relation holds in case we assume the customer under service during slot I is of type 1.

$$\begin{aligned}
&\Pr[r_1 = i_1, r_2 = i_2, f = m, b_1^* = k_1, b_2^* = k_2, s_{1,1}^* = j_{1,1}, \ldots, s_{2,k_2}^* = j_{2,k_2}] \\
&\quad = \frac{1 - A(0,0)}{\rho_1} \Pr[b_1 = k_1, b_2 = k_2] \prod_{l=1}^{2} \prod_{n=1}^{k_l} s_l(j_{l,n}), \\
&\quad \text{if } \sum_{p=1}^{k_1 - i_1} j_{1,p} \leq m < \sum_{p=1}^{k_1 - i_1 + 1} j_{1,p},\ 1 \leq i_1 \leq k_1 \text{ and } i_2 = k_2.
\end{aligned}$$

Consequently, combining these expressions leads to a closed-form formula for $H_1(x_1, x_2, z)$:

$$H_1(x_1, x_2, z) = x_1 \, \frac{S_1(z) - 1}{S_1(z) - x_1} \, \frac{B(S_1(z), x_2) - B(x_1, x_2)}{S_g'(1)(z - 1)}, \tag{7}$$

We can find $H_2(x_1, x_2, z)$ in an analogous way. Do note that the latter is not really a function of x_1, since all type-1 customers from group K have already been served in that case. This leads to

$$H_2(x_1, x_2, z) = x_2 \, \frac{S_2(z) - 1}{S_2(z) - x_2} \, \frac{S_g(z) - B(S_1(z), x_2)}{S_g'(1)(z - 1)}, \tag{8}$$

Next, when we convert (6) to the z-domain, we can write $V(z_1, z_2)$ as

$$\frac{V(z_1,z_2)-V(0,0)}{1-V(0,0)} = \frac{D(0)(B(z_1,z_2)-1)+DB(z_1,z_2)}{B(z_1,z_2)} H(z_1,z_2,A(z_1,z_2)),$$

One last substitution in the latter formula of $H_1(x_1, x_2, z)$ by (7), $H_2(x_1, x_2, z)$ by (8) and $D(z)$ by (4) yields after some tedious calculations

$$\begin{aligned}\frac{V(z_1,z_2)}{1-\rho} = 1 + z_1 \; & \frac{S_1A(z_1,z_2)-1}{z_1-S_1A(z_1,z_2)} \; \frac{B(z_1,z_2)-B(S_1A(z_1,z_2),z_2)}{B(z_1,z_2)-S_gA(z_1,z_2)} \\ & + z_2 \; \frac{S_2A(z_1,z_2)-1}{z_2-S_2A(z_1,z_2)} \; \frac{B(S_1A(z_1,z_2),z_2)-S_gA(z_1,z_2)}{B(z_1,z_2)-S_gA(z_1,z_2)}. \end{aligned} \quad (9)$$

Using this pgf we can calculate the average total queue length, as well as the average number of type-1 and type-2 customers in the queue at arbitrary slot boundaries represented by $E[v_1]$ and $E[v_2]$ in Fig. 2. To demonstrate the slot-bound priority mechanism we will assume an arrival distribution and service time distributions with pgf's equal to

$$A(z_1,z_2) = e^{\lambda(z_1z_2-1)}; \qquad S_1(z) = S_2(z) = z/(2-z).$$

Hence, in this specific example, an equal number of type-1 and type-2 customers arrive during any slot. The marginal distribution of a_1 and a_2 is Poisson with parameter λ. The effects of the slot-bound priority rule on $E[v_1]$ and $E[v_2]$ is boosted by this strong correlation. Furthermore, we have chosen the same geometrically distributed service time distribution for the two customer classes. We choose the marginal arrival distributions and service time distributions of the two types of customers the same so that any difference between $E[v_1]$ and $E[v_2]$ portrays the workings of the operated priority rule. In Fig. 2 we have also included results for the average number of type-j customers, if instead of slot-bound priority, non-preemptive HoL-priority (Walraevens[2]) were operated, and we have marked the resulting graphs with $E[v_{p,j}]$.

On fig. 2 we can see that, not surprisingly, $E[v_2] > E[v_1]$ for all values of the load. For high loads, $E[v_1]$ gets relatively closer to $E[v_2]$ while for low loads $E[v_j]$ gets relatively closer and closer to $E[v_{p,j}]$ as expected. For really low loads the buffer will almost always be empty and thus when one group is in the queue, it will most likely be the only one in the queue. Consequently, the slot-bound priority mechanism converges to the HoL non-preemptive priority paradigm. For higher loads, although the difference between $E[v_2]$ and $E[v_1]$ grows, due to more arrivals (we increase λ), their relative difference shrinks because the queue performance is increasingly dominated by the numbers of customer arrivals over multiple slots, rather than the details of the service order of the customers that arrive during a single slot.

References

[1] Fiems D., Walraevens J., Bruneel H. (2007) *Performance of a partially shared priority buffer with correlated arrivals.* Proceedings of the 20th International

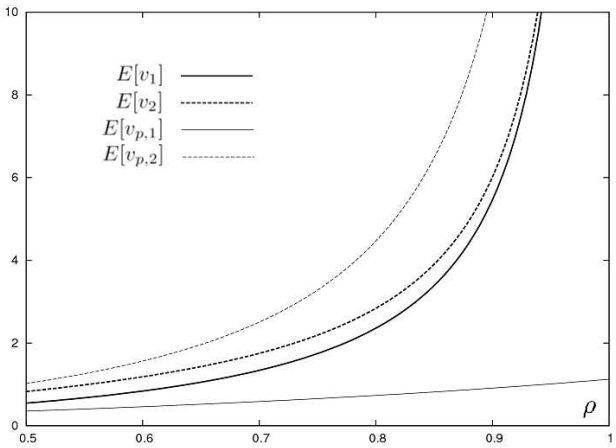


Figure 2: type-1 and type-2 population for different loads.

Teletraffic Congress (ITC20), Lecture Notes in Computer Science 4516, pp. 582-593

[2] Walraevens J., Steyaert B., Bruneel H. (2001) *Performance analysis of the system contents in a discrete-time non-preemptive priority queue with general service times*, Belgian Journal of Operations Research, Statistics and Computer Science (JORBEL) vol. 40, no. 1-2, pp. 91-103

[3] Ndreca S., Scoppola B. (2008) *Discrete-time GI/Geom/1 queueing system with priority*, European Cournal of operational research, vol. 189, no. 3, pp. 1403-1408.

[4] Stavrakakis I. (1994) *Delay bounds on a queueing system with consistent priorities*. IEEE Transactions on Communications, vol. 42, no. 2-4, part 1, pp. 615-624.

[5] Van Houdt B., Blondia C. (2002) *The delay distribution of a type-k customer in a first-come-first-served MMAP[K]/PH[K]/1 queue*. Journal of Applied Probability, vol. 39, no. 1, pp. 213-223.

[6] Takine T. (2001) *Queue length distribution in a fifo single-server queue with multiple arrival streams having different service time distributions*. Queueing Systems, vol. 39, nr. 4, pp. 349-375.

[7] Bruneel H. (1993) *Performance of discrete-time queueing systems*. Operations Research, vol. 20, nr. 3, pp.303-320.

6th St.Petersburg Workshop on Simulation (2009) 833-837

Approximate Method for QoS Analysis of Multi-Threshold Queuing Model of Multi-Service Wireless Networks

Hee Yeol Eom[1], Che Soong Kim[1], Agassi Melikov[2], Mehriban Fattakhova[2]

Abstract

A new effective approximate method for calculation of quality-of-service (QoS) metrics of the multi-threshold queuing model of a voice/data wireless network is developed. Results of numerical experiments are demonstrating high accuracy of the proposed formulas.

Keywords: network, queuing model, QoS metrics, calculation algorithm

1. Introduction

In an integrated voice/data wireless networks dropping of an arrived voice calls (handover or new) is more undesirable than blocking of data ones. A number of call admission strategies for multi-traffic wireless networks have been proposed in literature. So, in the [1], chapter 3, the queuing model is investigated with assumption that both voice and data calls are identical in terms of bandwidth requirements and channels holding times.The system is described with one-dimensional Markov chain and authors managed to find relatively simple formulas for calculating QoS metrics of the system.

However data traffic usually requires more bandwidth than voice. A queuing model with different bandwidth requirements for voice and data call as well as service differentiation between handover data calls and new data calls is developed in [2]. In this paper the given model was investigated by the method which is based on recursive technique for solving a large system of balance equations. But this method faces with known computational difficulties for case large scale networks. Here for solving the given problem new approach is introduced. Our approach is based on the principles of theory of phase merging of stochastic systems [3]. This approach was used earlier for solving similar problems in wireless networks model with one threshold in paper [4].

This paper is organized as follows. In Section 2, we describe the model and provide a simple algorithm to calculate QoS metrics. Numerical results are given in Section 3. In Section 4, we provide some conclusion remarks.

[1]Sangji University, Wonju, Kangwon, Korea, E-mail: dowoo@sangji.ac.kr

[2]Institute of Cybernetics of National Academy of Sciences of Azerbaijan, E-mail: agassi@science.az

2. The Model And Calculation Method

An isolated cell of multi-service wireless network contains $N > 1$ channels, which divided into four segments by three thresholds N_1, N_2, and N_3. It is assumed that N_1 and N_2 are multiplies of b, the number of channels needed to service a data call and $0 < N_1 \leq N_2 \leq N_3 \leq N$. In the cell Poisson flows of new voice calls (with intensity λ_{ov}), handover voice calls (with intensity λ_{hv}), new data calls (with intensity λ_{od}), and handover data calls (with intensity λ_{hd}) are handled. The following restricted admission strategy for heterogeneous calls is defined [1]:

- If admission of a new data call will increase the number of busy channels to a number less than or equal to N_1, then the new data call will be accepted; otherwise it will be blocked.
- If admission of a handover data call will increase the number of busy channels to a number less than or equal to N_2, then the handover data call will be accepted; otherwise it will be blocked.
- If admission of a new voice call will increase the number of busy channels to a number less than or equal to N_3, then the new voice call will be accepted; otherwise it will be blocked.
- If upon arrival of a handover voice call, there is at least one free channel, then the handover voice call will be accepted; otherwise it will be blocked.

Distribution functions of channel occupancy time of both types of calls (i.e. voice and data) are exponential, but their parameters are different, namely intensity of handling of voice call (new or handover) calls equals μ_v and intensity of handling of voice (new or handover) calls equals μ_d, generally speaking $\mu_v \neq \mu_d$

For the sake of simplicity below we assume that $b = 1$. The extension of any value of b is straightforward. The state of the system at any time is described by two-dimensional vector $\mathbf{n} = (n_d, n_v)$ where n_d and n_v denote the total number of data and voice calls in the cell, respectively. Then state space of appropriate Markov chain (MC) is given by

$$S := \{\mathbf{n} : n_d = 0, N_2, n_v = 0, N, n_d + n_v \leq N\} \tag{1}$$

Elements of generating matrix of this MC $q(\mathbf{n}, \mathbf{n}'), \mathbf{n}, \mathbf{n}' \in S$ are determined as follows:

$$q(\mathbf{n}, \mathbf{n}') = \begin{cases} \lambda_d & \text{if } n_d + n_v \leq N_1 - 1, \mathbf{n}' = \mathbf{n} + \mathbf{e}_1, \\ \lambda_{hd} & \text{if } N_1 \leq n_d + n_v \leq N_2 - 1, \mathbf{n}' = \mathbf{n} + \mathbf{e}_1, \\ \lambda_v & \text{if } n_d + n_v \leq N_3 - 1, \mathbf{n}' = \mathbf{n} + \mathbf{e}_2, \\ \lambda_{hv} & \text{if } N_3 \leq n_d + n_v \leq N - 1, \mathbf{n}' = \mathbf{n} + \mathbf{e}_2, \\ n_d\mu_d & \text{if } \mathbf{n}' = \mathbf{n} - \mathbf{e}_1, \\ n_v\mu_v & \text{if } \mathbf{n}' = \mathbf{n} - \mathbf{e}_2, \\ 0 & \text{otherwise,} \end{cases} \tag{2}$$

where $\lambda_d := \lambda_{od} + \lambda_{hd}, \lambda_v := \lambda_{ov} + \lambda_{hv}, e_1 = (1, 0), e_2 = (0, 1)$.

Major QoS metrics of the model are probability of blocking of calls of each type $P_x, x \in \{hv, ov, hd, od\}$ and an average number of busy channels, N_{av}. These quantities are determined via stationary distribution of the initial model:

$$P_{hv}: = \sum_{\mathbf{n}\in S} p(\mathbf{n}) I(n_d + n_v = N), \tag{3}$$

$$P_{ov}: = \sum_{\mathbf{n}\in S} p(\mathbf{n}) I(n_1 + n_2 \geq N_3), \tag{4}$$

$$P_{hd}: = \sum_{\mathbf{n}\in S} p(\mathbf{n}) I(n_1 + n_2 \geq N_2), \tag{5}$$

$$P_{od}: = \sum_{\mathbf{n}\in S} p(\mathbf{n}) I(n_1 + n_2 \geq N_1), \tag{6}$$

$$N_{av}: = \sum_{k=1} k \sum_{\mathbf{n}\in S} p(\mathbf{n}) I(n_1 + n_2 \geq N_3), \tag{7}$$

where $p(\mathbf{n})$ - stationary probability of state $\mathbf{n} \in S$, $I(A)$ – indicator function of event A.

New method for calculation of QoS metrics of the given model is suggested below.

Assumption: It is assumed that $\lambda_v \gg \lambda_v$, $\mu_v \gg \mu_v$. This assumption is the most suitable regime in integrated voice/data wireless networks. Indeed, in such networks voice traffic is characterized by very short durations in comparison with data one while real time traffics (i.e. voice calls in our case) constitute the greater part of the total traffic. The following splitting of state space (1) is examined:

$$S = \bigcup_{k=0}^{N_2} S_k, \; S_k \cap S_{k'} = \emptyset, \; k \neq k', \tag{8}$$

where $S_k := \{\mathbf{n} \in S : n_d = k\}$.

State classes S_k combine into separate merged states $< k >$ and the following merge function in state space S is introduced:

$$U(\mathbf{n}) =< k >, if \mathbf{n} \in S_k, k = \overline{0, N_2} \tag{9}$$

Function (9) determines merged model which is one-dimensional MC with the state space $\widetilde{S} := \{< k >: \; k = \overline{0, N_2}\}$. Then stationary distribution of the initial model approximately equals:

$$p(k.i) \approx \rho_k(i)\pi(< k >), \; (k, i) \in S_k, \; k = \overline{0, N_2}, \tag{10}$$

where $\{\rho_k(i) : (k, i) \in S_k\}$ and $\left\{\pi(< k >) :< k >\in \widetilde{S}\right\}$ are stationary distributions within class S_k and merged model, respectively. Stationary distribution within class S_k $\rho_k(i), (k, i) \in S_k$ are determined as follows:

$$\rho_k(i) = \begin{cases} \frac{v_v^i}{i!}\rho_k(0) & \text{if } 1 \le i \le N_3 - k, \\ \left(\frac{v_v}{v_{hv}}\right)^{N_3-k} \frac{v_{hv}^i}{i!}\rho_k(0) & \text{if } N3 - k + 1 \le i \le N - k, \end{cases} \tag{11}$$

Here

$$\rho_k(0) = \left(\sum_{i=0}^{N_3-k} \frac{v_v^i}{i!} + \left(\frac{v_v}{v_{hv}}\right)^{N_3-k} \sum_{i=N_3-k+1}^{N-k} \frac{v_{hv}^i}{i!} \right)^{-1}, v_v := \lambda_v/\mu_v, v_{hv} := \lambda_{hv}/\mu_v$$

So elements of generating matrix of a merged model $q\left(< k >, < k' >\right), < k >$ $, < k' > \in \widetilde{S}$ are

$$q\left(< k >, < k' >\right) = \begin{cases} \lambda_d \sum_{i=0}^{N_1-k-1} \rho_k(i) + \sum_{i=N_1-1}^{N_2-k-1} \rho_k(i) & \text{if } 0 \le k \le N_1 - 1, k = k + 1, \\ \lambda_{hd} \sum_{i=0}^{N_2-k-1} \rho_k(i) & \text{if } N_1 \le k \le N_2 - 1, k = k + 1, \\ k\mu_d & \text{if } k = k - 1, \\ 0 & \text{otherwise} \end{cases} \tag{12}$$

Consequently, stationary distribution of a merged model is determined as:

$$\pi(< k >) = \frac{\pi(< 0 >)}{k!\mu_d^k} \prod_{i=1}^{k} q(< k-1 >, < k >), k = \overline{1, N_2}, \tag{13}$$

where

$$\pi(< k >) = \left(1 + \sum_{k=1}^{N_2} \frac{1}{k!\mu_d^k} \prod_{i=1}^{k} q(< k-1 >, < k >)\right)^{-1}.$$

Finally, using (10)-(13) we find the following approximate formulas for calculating the QoS metrics of the system (3)-(7):

$$P_{hv} \approx \sum_{k=0}^{N_2} \pi(< k >)\rho_k(N - k); \tag{14}$$

$$P_{ov} \approx \sum_{k=0}^{N_2} \pi(< k >) \sum_{i=N_3-k}^{N-k} \rho_k(i); \tag{15}$$

$$P_{hd} \approx \sum_{k=0}^{N_2} \pi(< k >) \sum_{i=N_2-k}^{N-k} \rho_k(i); \tag{16}$$

$$P_{od} :\approx \sum_{k=0}^{N_1-1} \pi(< k >) \sum_{i=N_1-k}^{N-k} \rho_k(i) \sum_{k=N_1}^{N_2} (< k >); \tag{17}$$

$$N_{av} := \sum_{k=1}^{N} k \sum_{i=0}^{f(k)} \pi(< i >)\rho_i(k-i). \quad (18)$$

Here $f(k) = \begin{cases} k & \text{if } 1 \le k \le N_2 \\ N_2 & \text{if } N_2 + 1 \le k \le N \end{cases}$

3. Numerical Results

Large volume of computational experiments in broad range of changes of structural and load parameters of the system is carried out. Due to limitation of volume of work only high accuracy of suggested formulas are shown for loss probabilities of heterogeneous calls where we set $N = 16, N_3 = 14, N_2 = 10, \lambda_{ov} = 10, \lambda_{hv} = 6, \lambda_{od} = 4, \lambda_{hd} = 3, \mu_v = \mu_d = 2$. Thus, maximal difference between our results and those in [1], pages 131-135, in the case $b = 1$ and $\mu_v = \mu_d$ (it is very case covered in [1] and the formulae given there are considered accurate ones) for voice calls does not exceed 0.01in worst case maximal difference comprises around 9the comparison by Table 1 and 2, where EV - exact value, AV - approximates value.

Table 1: Comparison with exact results for voice calls

N	P_{ov}		P_{hv}	
	EV	AV	EV	AV
1	0.03037298	0.03465907	0.00092039	0.00119181
2	0.03037774	0.03469036	0.00092054	0.00119309
3	0.03040249	0.03482703	0.00092129	0.00119878
4	0.03048919	0.03521813	0.00092392	0.00121521
5	0.03072036	0.03604108	0.00093092	0.00125021
6	0.03122494	0.03741132	0.00094621	0.00130942
7	0.03217389	0.03932751	0.00097497	0.00139396
8	0.03377398	0.04168754	0.00102345	0.00150073
9	0.03627108	0.04432985	0.00109912	0.00162373
10	0.03997025	0.04706484	0.00121112	0.00175503

4. Conclusion

Effective approximate method for calculation of queuing model with multi - thresholds scheme for call admission control in isolated cell of an integrated voice/data wireless network are given. Performed numerical results demonstrate high efficiency (with regard to the degree of the complexity) and accuracy of the developed method. The suggested algorithm for calculation of QoS metrics allows optimal choose (in some sense) the values of threshold parameters. It is important to note that the proposed approach also may be applied for studying models in which either finite or infinite queue of heterogeneous calls is allowed. These problems are subjects to separate investigation.

Table 2: Comparison with exact results for data calls

N	P_{od}		P_{hd}	
	EV	AV	EV	AV
1	0.99992793	0.99985636	0.39177116	0.35866709
2	0.99925564	0.99855199	0.39183255	0.35886135
3	0.99612908	0.99271907	0.39215187	0.35969536
4	0.98645464	0.97565736	0.39327015	0.36203755
5	0.96398536	0.93891584	0.39625194	0.36685275
6	0.92198175	0.87621832	0.40276033	0.37462591
7	0.85564333	0.78660471	0.41500057	0.38506671
8	0.76370389	0.67487475	0.43563961	0.39731190
9	0.64880652	0.55004348	0.46784883	0.41028666
10	0.51556319	0.42295366	0.51556319	0.42295366

References

[1] Chen H., Huang L., Kumar S., Kuo C.C. (2004) *Radio resource management for multimedia QoS support in wireless networks*. Kluwer Academic Publishers, Boston.

[2] Ogbonmwan S.E., Wei L. (2006) *Multi-threshold bandwidth reservation scheme of an integrated voice/data wireless network*. Computer Communications, 29 (9), 1504–1515.

[3] Korolyuk V.S., Korolyuk V.V. (1999) *Stochastic models of systems*. Kluwer Academic Publishers, Boston.

[4] Melikov A.Z, Babaev A.T. (2006) *Refined approximations for performance analysis and optimization of queuing model with guard channels for handovers in cellular networks*. Computer Communications, 29 (8), 1386–1392

6th St.Petersburg Workshop on Simulation (2009) 839-841

Managing Uncertainty in Complex Stochastic Models: Design and Emulation of a Rabies Model[1]

Alexis Boukouvalas[2], Dan Cornford[3], Alexander Singer[4]

Abstract

In this paper we present a novel method for emulating a stochastic, or random output, computer model and show its application to a complex rabies model. The method is evaluated both in terms of accuracy and computational efficiency on synthetic data and the rabies model. We address the issue of experimental design and provide empirical evidence on the effectiveness of utilizing replicate model evaluations compared to a space-filling design. We employ the Mahalanobis error measure to validate the heteroscedastic Gaussian process based emulator predictions for both the mean and (co)variance. The emulator allows efficient screening to identify important model inputs and better understanding of the complex behaviour of the rabies model.

1. Introduction

In many scientific and engineering problems complex simulators, based on mechanistic and physical process driven models, are routinely used to solve complex problems. Such simulators are often computationally expensive, and full uncertainty analysis, sensitivity analysis or other probabilistic analysis becomes extremely time consuming, effectively being computationally intractable. The most commonly applied solution is to create a meta-model for the simulator [5], often referred to as an *emulator* [3]. The role of the emulator can be seen to be approximating the simulator. In most existing work emulator methods are applied to deterministic models, of the form $\mathbf{y} = \mathbf{f}(\mathbf{x})$ where $\mathbf{x}$ represents the inputs to the simulator, $\mathbf{y}$ represents the outputs of the simulator, or some summary of these, and $\mathbf{f}$ represents the mapping imposed by the simulator evaluation. The probabilistic nature of the emulator, which is typically modelled as a Gaussian Process (GP) [3], arises from the *approximation* of the simulator due to having a finite number of simulator runs. In this paper we develop novel methods for the emulation of a stochastic simulator, a relatively new field [5].

[1]This research was funded as part of the Managing Uncertainty in Complex Models project by EPSRC grant D048893/1.

[2]Aston University, E-mail: boukouva@aston.ac.uk

[3]Aston University, E-mail: D.Cornford@aston.ac.uk

[4]Central Science Laboratories, E-mail: alexssinger@googlemail.com

A GP is defined as a collection of random variables, any finite subset of which has a joint Gaussian distribution [8]. It is completely defined by a mean and a covariance function, the specification of which allows the incorporation of prior knowledge in the emulation analysis such as the smoothness and differentiability of the approximated function, that is the simulator.

Another issue commonly occurring in the context of complex datasets is that of experimental design [7]. We assess the efficiency of different designs, examining the effect of replicate model evaluations, where the simulator is evaluated repeatedly for a single design point, against a more traditional space filling design. Utilizing the moments of the replicate evaluations allows for computationally efficient inference, and we empirically show that it also increases the accuracy of the heteroscedastic emulator, especially the (co)variance estimates.

2. Stochastic emulation

Relatively little work has addressed the question of the emulation of stochastic simulators. In this work we consider a stochastic simulator to be a mapping that produces random output given a fixed set of inputs. A recent review of the application of 'Kriging' (or GP regression) to emulation can be found in [5].

Kleijnen and co-workers [5] have studied the problem of stochastic emulation closely, investigating queuing models. In the work of Kleijnen the emulator of stochastic simulators uses m repetitions of the simulator at each of the i design points. From this the mean response $\bar{y}_i = \frac{1}{m}\sum_{j=1}^{m} \mathbf{y}_{i,j}$ and the variance of the response $\mathbf{S}^2{}_i = \frac{1}{m-1}\sum_{j=1}^{m}(\bar{y}_i - \mathbf{y}_{i,j})^2$ are computed, where $\mathbf{y}_{i,j}$ is the j'th realisation from the stochastic simulator, at the i'th design point. The main concern in [6] is modelling the mean response of the stochastic simulator. The variance estimates, $\mathbf{S}_i$ are used to 'Studentize' the output with the transformation $\tilde{\mathbf{y}}_i = \bar{y}_i / \sqrt{\mathbf{S}_i / m^2}$, where they assume $\mathbf{y}$ has had any 'large scale' trend removed. A standard GP regression of the transformed output, $\tilde{\mathbf{y}}_i$, is then applied. The allowance for heteroscedastic, i.e. input dependent, variance is limited to a small number of simple parametric models. In all the work on stochastic emulation very little attention is paid to the treatment of heterogeneity of the output variance. In this paper we extend the recent work of [4] to enable improved stochastic emulation of more complex models and test it on a rabies disease simulator.

3. Heteroscedastic Modelling

In this section we briefly describe our method. The reader is referred to [2] for a detailed description. Following [4], we define a GP on the mean model output $\mathbf{G}_\mu$ and a second GP on the log variance of the model output, $\mathbf{G}_{\mathbf{\Sigma}}$. We do not present the full GP inference framework here but note that in all experiments maximum marginal likelihood estimation was used for the covariance hyper-parameters. The notation used is: N the number of design points used during inference, $D = \{\mathbf{x}_i, y_i\}$ the training dataset, n_i the number of replicate model evaluations at each design point location $\mathbf{x}_i$ $i \in [1, \ldots, N]$ and $diag$ signifies a diagonal matrix.

The algorithm is initialized by estimating a homoscedastic GP which is fitted on the empirical mean values. This is treated as our initial estimate of $\mathbf{G}_\mu$. We proceed by estimating the variance GP $\mathbf{G}_{\mathbf{\Sigma}}$. Where no replicate model evaluations are available for a design point $\mathbf{x}_i$, the predictive distribution of the mean GP $\mathbf{G}_\mu$ is sampled to estimate the noise levels of the data [4]. In the case of replicate evaluations at $\mathbf{x}_i$ the empirical variance $\mathbf{S}_i$ is estimated directly. To correct for the biased estimate of the variance due to the log transformation we apply the correction: $r_i = log(\mathbf{S}_i) + (d_i + d_i \log(2) - \Psi(d_i/2))^{-1}$, where r_i is the true log variance, $d_i = n_i - 1$, and Ψ the digamma function.

Finally the heteroscedastic GP $\mathbf{G}_\mu$ is estimated to jointly predict the mean and variance. The predictive distribution equations for $\mathbf{G}_\mu$ for M test points $\mathbf{x}_*$ are:

$$\begin{aligned} E[\mathbf{y}_*|\mathbf{x}_*, D] &= K^*(K + RP^{-1})^{-1}\mathbf{y} + E^T\bar{\beta}, \\ Var[\mathbf{y}_*|\mathbf{x}_*, D] &= K^{**} + R^* - K^{*^T}(K + R)^{-1}K^* + E^T(H(K + R)^{-1}H^T)^{-1}E, \end{aligned}$$

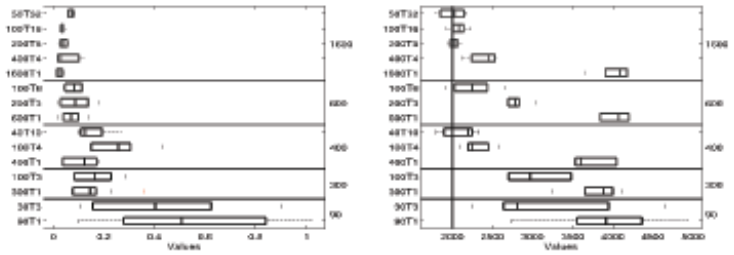
where $\mathbf{y} = [y_1 \dots y_N]$ is the vector of outputs in the training set D, K is the covariance of training points, K^* the cross-covariance between training and test points, K^{**} the covariance of test points, H a set of fixed basis functions, $\bar{\beta} = (H(K + R)^{-1}H^T)^{-1}H(K + R)^{-1}\mathbf{y}$ the regression coefficients, $E = H_* - H(K + R)^{-1}K^*$, $P = diag(n_1 \dots n_N)$ the number of replicates at each training point, $R = diag[r(x_1) \dots r(x_N)]$ and $R^* = diag[r(x_{*1}) \dots r(x_{*M})]$ the variance estimate from $\mathbf{G}_{\mathbf{\Sigma}}$ at the training and test points respectively. We note that the non-standard RP^{-1} term in the predictive mean arises from the use of replicate evaluations. The algorithm is repeated until convergence.

4. Experimental design analysis using synthetic data

In this section we utilize our framework to assess the efficacy of different experimental design towards emulation accuracy on a synthetic dataset [10]. Our chief validation measure is the Mahalanobis error $D_{MD} = (\mathbf{y} - \mathbf{t})'\mathbf{\Sigma^{-1}}(\mathbf{y} - \mathbf{t})$, where $\mathbf{t}$ the vector of model outputs, $\mathbf{y}$ and Σ the predictive GP mean and covariance respectively. The Mahalanobis error assesses the goodness of the joint fit, both of the mean and covariance prediction [1].

In this experiment the total number of model evaluations is kept fixed and we contrast a space-filling design with only single model evaluations against a more widely-spaced replicate design that has the same number of evaluations for all design points.

The benefits of a replicate design can be seen in Figure 1 where the Mean Squared Error (MSE) and Mahalanobis error are shown for the different designs. There is little difference in terms of MSE signifying similar performance with regards to the prediction of the mean. The Mahalanobis error however reveals significant gains when replicate designs are used, reflecting an improvement in variance prediction. The replicate designs are also substantially faster to use from a computational perspective, i.e. inference time.



(a) Mean Squared Error (b) Mahalanobis

Figure 1: Comparison of emulator fit where the total number of model evaluations is fixed at different levels. Notation is: 30T3 = 30 design points each with 3 replicates. Results shown for a total of 90, 300, 400, 600 and 1600 total number of model evaluations.

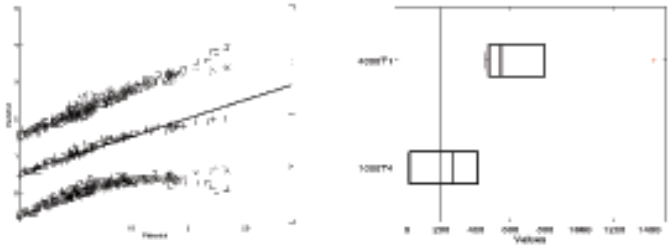
4.1. Stochastic Rabies Model

Although wildlife rabies was eradicated from large parts of Europe, there is a remaining risk of disease re-introduction. The situation is aggravated by an invasive species, the raccoon dog (Nyctereutes procyonoides) that can act as a second rabies vector in addition to the red fox (Vulpes vulpes). The purpose of our rabies model is to analyse the risk of rabies spread in this new type of vector community [9]. The individual-based, non-spatial, time-discrete model incorporates population and disease dynamical processes such as host reproduction and mortality as well as disease transmission. These processes are modelled stochastically to reflect natural variability (e.g. demographic stochasticity). Thus model analysis (e.g. sensitivity analysis) has to deal with stochastic, indeed heteroscedastic, model output.

The model output investigated in this study is the number of time steps to disease extinction. This output is important in deciding on the response to a potential rabies outbreak. This output has a rather complex, non-Gaussian, distribution for a fixed input; in this paper we emulate the first two moments of the log extinction time, which is more approximately Gaussian, as evidenced from visual inspection of Q-Q plots.

In Figure 2 we show the validation results of a single instance of our GP framework. The GPs were trained using a 1000 point Latin Hypercube design with a mixture of single and replicate model evaluations. A total of 4000 rabies model evaluations were used. In Figure **??**, estimates of the 'correct' mean and standard deviation response (using 1000 repetitions) are plotted against the corresponding predicted values from $\mathbf{G}_\mu$.

We finally explore the question of how the replicate framework compares to approximations often applied within GP inference. The projected process method utilizes all N training points but it only represents $m < N$ latent function values, called support points, as an approximation to the full GP posterior [8]. In Figure **??** the Mahalanobis error of applying the approximation on [4] using a 4000



(a) Estimated (square) vs Predicted (x) deviation and mean (+).

(b) Mahalanobis Error

Figure 2: (a) Emulating the rabies model using 1000 design points with a replicate design. (b) Projected process 'Kersting' (4000) vs replicated design (1000).

point space-filling design with $m = 1000$ support points is contrasted against the replicate method on a 1000 point space-filling design with 4 replicate observations at each design point. Both methods require approximately the same amount of computational resource, but the replicate observation method gives substantially better results, over 10 repetitions.

4.2. Screening of the rabies model

Lastly we consider using the replicate framework to perform screening which is often used as a preliminary stage in sensitivity analysis to remove clearly unimportant factors. In our framework, screening can be accomplished quite intuitively by looking at the posterior values of regression coefficients and correlation length scales. Furthermore these effects can be decomposed for the mean process ($\mathbf{G}_\mu$) and variance process ($\mathbf{G}_{\mathbf{\Sigma}}$).

The three dominant factors (out of 14 model inputs) on the variance response of the rabies model in terms of linear effects and correlation length scales are shown in Table 1. We observe that density and mortality rates of raccoon dogs have strong linear effects (significantly higher regression coefficients than other parameters). With regards to correlation length scales which reveal non-linear and interaction effects, factors related to disease in the vector species appear influential.

Table 1: Interpreting the variance emulator ($\mathbf{G}_{\mathbf{\Sigma}}$) by looking at the regression coefficients (Coeff) and correlation length scales (Scale).

Factor	Coeff	Factor	Scale
Rac Density	0.1608	Rac Rabid	1.4281
Rac Death	0.0633	Fox Inf	1.4594
Rac Birth	0.0200	Fox Rabid	1.5047

5. Conclusions

In this paper we have presented a new approach to the emulation of stochastic models which improves upon existing methods both in terms of accuracy and computational efficiency. Our framework allows further analysis to be carried out in a straight-forward and efficient manner using the emulator as a proxy for the simulator. Examples of such analyses include screening and uncertainty analysis, where we have included a demonstration of the former on a rabies model. Furthermore the computer model parameter space can be explored without the necessity of a large number of (computationally demanding) simulator runs. In combination with a discrepancy model and real-world observations, this method could facilitate the efficient statistical calibration of stochastic models.

References

[1] L. S. Bastos and A. O'Hagan. Diagnostics for Gaussian process emulators. Technical report, University of Sheffield, 2008.

[2] A. Boukouvalas and D. Cornford. Learning heteroscedastic Gaussian processes for complex datasets. Technical report, NCRG, Aston University, Aston University, Aston Triangle, Birmingham, B4 7ET, 2009.

[3] M.C. Kennedy and A. O'Hagan. Bayesian calibration of computer models (with discussion). *Journal of the Royal Statistical Society*, B63:425–464, 2001.

[4] K. Kersting, C. Plagemann, P. Pfaff, and W. Burgard. Most likely heteroscedastic Gaussian process regression. In Zoubin Ghahramani, editor, *Proc. 24th International Conf. on Machine Learning*, pages 393–400. Omnipress, 2007.

[5] J P C Kleijnen. Kriging metamodeling in simulation: a review. *European Journal of Operational Research*, 2007.

[6] J.P.C. Kleijnen and W.C.M. van Beers. Robustness of Kriging when interpolating in random simulation with heterogeneous variances: Some experiments. *European Journal of Operational Research*, 165(3):826–834, 2005.

[7] Andreas Krause, Ajit Singh, and Carlos Guestrin. Near-optimal sensor placements in Gaussian processes: Theory, efficient algorithms and empirical studies. *J. Mach. Learn. Res.*, 9:235–284, 2008.

[8] C E Rasmussen and C K I Williams. *Gaussian Processes for Machine Learning*. MIT Press, 2006.

[9] A. Singer, F. Kauhala, K. Holmala, and G.C. Smith. Rabies risk in raccoon dogs and foxes. *Developments in Biologicals*, 131:213–222, 2008.

[10] M. Yuan and G. Wahba. Doubly penalized likelihood estimator in heteroscedastic regression. *Statistics and Probability Letters*, 69:11–20, 2004.

6th St.Petersburg Workshop on Simulation (2009) 845-849

Using latent variables to account for heterogeneity in exponential family random graph models

Johan Koskinen[1]

Abstract

We consider relaxing the homogeneity assumption in exponential family random graph models (ERGMs) using binary latent class indicators. This may be interpreted as combining a posteriori blockmodelling with ERGMs, relaxing the independence assumptions of the former and the homogeneity assumptions of the latter. We propose a Markov chain Monte Carlo algorithm for drawing from the joint posterior of the model parameters and latent class indicators.

1. Introduction

Researchers in social science have long recognised the potential of using graph theory for study the social interaction among social units [13]. A social network may be conceptualised as consisting of a set of vertices $\mathcal{V} = \{1, \ldots, n\}$, representing the social units, e.g. people, that are pairwise relationally connected by ties, represented by an edge set $E \subseteq \mathcal{N}$, where $\mathcal{N} = \binom{\mathcal{V}}{2}$, for undirected relations and $\mathcal{N} = \mathcal{V}^{(2)}$ for directed relations. We define the adjacency matrix $x = \{x_e : e \in \mathcal{N}\}$ as the collection of edge indicators $x_e = \mathbf{1}_E(e)$, $x \in \mathcal{X} = \{0,1\}^{\mathcal{N}}$. White et al. [19] proposed summarising the structural information by reducing the graph to a blockmodel, a schematic representation of positions $\rho : \mathcal{V} \to \mathcal{P} = \{0, \ldots, P-1\}$ and a blockmodel that is a graph $\mathcal{G}(\mathcal{P}, \mathcal{B})$. Strictly speaking $\mathcal{G}(\mathcal{P}, \mathcal{B})$ is such that $x_{ij} = \mathbf{1}_{\mathcal{B}}(\{\rho(i), \rho(j)\})$, for $i \neq j$. A priori blockmodels and a posteriori blockmodels are now well established practice in social network analysis [1]. Stochastic a priori blockmodels [2], [7] have since been extended to stochastic a posterior blockmodels, where the target of inference is to infer the positions of the vertices from their observed relations [21], [16], [23]. For stochastic blockmodels the (pairs of) tie-indicators are independent but [6] sought to relax this usually unrealistic assumption.

Exponential random graph models (ERGMs) take dependencies between the elements of x into account. Dependencies are specified according to a dependence graph $\mathcal{D} = \mathcal{G}(\mathcal{N},\mathcal{E})$, for example $\{\{i,j\},\{k,\ell\}\} \in \mathcal{E}$, iff $\{i,j\} \cap \{k,\ell\} \neq \varnothing$ in the case of Markov graphs [3]. This specifies a log-linear model $p(x)$ with parameters

[1]University of Melbourne, E-mail: johank@unimelb.edu.au

α_A and sufficient statistics $\prod_{e\in A} x_e$ for A that are cliques of $\mathcal{D}$. This model formulation has too many parameters wherefore a homogeneity assumption is usually imposed meaning that p is invariant to permutations of the labels of elements of $\mathcal{V}$. In general

$$p(x|\theta) = \exp\{\theta^{\mathrm{T}} z(x) - \varkappa(\theta)\},$$

where where z is a $p \times 1$ vector valued function of x and the normalising constant $\varkappa(\theta) = \log(c(\theta))$, $c(\theta) = \sum_{y\in\mathcal{X}} \exp\{\theta^{\mathrm{T}} z(y)\}$, is a function of the parameter vector $\theta \in \Theta \subseteq R^p$. For Markov graphs z consists of counts of the number of edges, stars and triangles (for non-Markov specifications see [22]).

The homogeneity assumption may be relaxed if attributes of the nodes are observed [18] but unexplained heterogeneity has detrimental effects on estimation.

2. Bayesian inference for ERGM

Given a fixed x we may perform inference for θ by exploring a sample from the posterior distribution $\pi(\theta|x) \propto p(x|\theta)\pi(\theta)$, using the linked importance sampler auxiliary variable (LISA) Metropolis-Hastings proposed in [10]. LISA combines the auxiliary variable MCMC [15] with the linked importance sampler [17] (a similar algorithm is proposed in [14]). In brief, LISA relies on the introduction of an auxiliary variable ω defined on the space $\prod_{i=1}^{m} \mathcal{X}^K \times [K] \times [K]$, constructed such that there are two pmfs, a forward distribution $P^F_{\theta,\psi}(\omega)$ and a backward distribution $P^B_{\psi,\theta}(\omega)$, each indexed by a pair (θ, ψ) of parameter vectors. Furthermore, the normalising constants of these distributions are $c(\theta)$ and $c(\psi)$ respectively. In LISA, instead of producing draws from $\pi(\theta|x)$ directly, we produce draws from the joint distribution

$$\pi(\theta, \omega|x) \propto p(x|\theta)P^B_{\psi,\theta}(\omega) = \exp\{\theta^{\mathrm{T}} z(x)\}/c(\theta)Q^B_{\psi,\theta}(\omega)/c(\psi)$$

This may be accomplished for a Metropolis-Hastings sampler that in each iteration, with present state $(\theta^{(t)}, \omega^{(t)})$, proposes a move to (θ^*, ω^*) drawn from the proposal distribution $f(\theta^*, \omega^*|\theta^{(t)}, \omega^{(t)})$, and this move is accepted with probability $\min\{1, H\}$, where

$$H = \frac{\pi(\theta^*, \omega^*|x)}{\pi(\theta^{(t)}, \omega^{(t)}|x)} \frac{f(\theta^{(t)}, \omega^{(t)}|\theta^*, \omega^*)}{f(\theta^*, \omega^*|\theta^{(t)}, \omega^{(t)})}.$$

Evaluating this expression involves a ratio $c(\theta^*)/c(\theta^{(t)})$ of undefined normalising constants but if f is chosen to be $f(\theta^*, \omega^*|\theta^{(t)}, \omega^{(t)}) = f(\omega^*|\theta^*)f(\theta^*|\theta^{(t)})$, where $\theta^*|\theta^{(t)} \sim N(\theta^{(t)}, \Sigma)$ and $\omega^*|\theta^* \sim Q^F_{\theta^*,\psi}(\omega^*)/c(\theta^*)$, the ratio H simplifies to $e^{\theta^{*\mathrm{T}} z(x) - \theta^{(t)\mathrm{T}} z(x)} \Pi_\psi(\omega^*, \theta^*)/\Pi_\psi(\omega^{(t)}, \theta^{(t)})$, where $\Pi_\psi(\omega, \theta) = Q^B_{\psi,\theta}(\omega)/Q^F_{\theta,\psi}(\omega)$. Note that H has the interpretation that $\mathrm{E}[\Pi_\psi(\omega^*, \theta^*)] = c(\psi)/c(\theta^*)$, where the expectancy is with respect to $P^F_{\theta^*,\psi}(\omega^*)$. This is also why we may characterise the algorithm as a Metropolis-Hastings that accepts a proposal with a Hastings ratio where $\Pi_\psi(\omega, \theta) = \lambda_{LIS}(\theta, \psi; \omega)$, and where λ_{LIS} is the linked importance

sampler (LIS) estimator [17] of the ratio of normalising constants. Given a sample $\omega = (y, \mu, \nu)$, the LIS estimate of $\lambda(\theta, \psi)$ is given by

$$\lambda_{LIS}(\theta, \psi; \omega) = \prod_{j=1}^{m-1} \frac{\sum_{i=1}^{K} w(y_j^{(i)}; \theta(j), \theta(j+1))}{\sum_{i=1}^{K} w(y_{j+1}^{(i)}; \theta(j+1), \theta(j))},$$

where the weights w are derived from the sampling process that we now proceed to describe.

The LIS estimator is based on K sample points from m Markov chains $y_j = (y_j^{(i)})_{i=1}^K$ with different target distributions, drawn using Metropolis-Hasting transition probabilities $T_{\theta(t)}$ and $\underline{T}_{\theta(t)}$, for a smooth mapping connecting θ and ψ as in path sampling [4].

The m samples are connected in points $\mu_1, \ldots, \mu_m$ and $\nu_1, \ldots, \nu_m$, such that given μ_j and $(y_j^{(1)})_{i=1}^K$, we set $y_{j+1}^{(\nu_{j+1})} := y_j^{(\mu_j)}$. Given ν_j and $y_j^{(\nu_j)}$ we create the chain $(y_j^{(1)})_{i=1}^K$ by simulating forward from $y_j^{(\nu_j)}$ using $T_{\theta(j)}(y_j^{(\nu_j)}, y_j^{(\nu_j+1)})$, $T_{\theta(j)}(y_j^{(\nu_j+1)}, y_j^{(\nu_j+2)})$, etc., until we have produced $y_j^{(K)}$. We also simulate backwards from $y_j^{(\nu_j)}$ using the reversed transition kernels $\underline{T}_{\theta(j)}(y_j^{(i)}, y_j^{(i-1)})$, until we have produced $y_j^{(1)}$. The implied pmf of a chain $y_j = (y_j^{(i)})_{i=1}^K$ conditional on the insertion point and the linking state is

$$P(y_j | \nu_j, y_j^{(\nu_j)}) = \prod_{i=1}^{\nu_j - 1} \underline{T}_{\theta(j)}(y_j^{(i+1)}, y_j^{(i)}) \prod_{i=\nu_j}^{K} T_{\theta(j)}(y_j^{(i)}, y_j^{(i+1)}).$$

To choose which of the K sample points that should provide the link to the next chain, we choose μ_j with probabilities

$$\eta(\mu_j | y_j) = w(y_j^{(\mu_j)}; \theta(j), \theta(j+1)) / \sum_{i=1}^{K} w(y_j^{(i)}; \theta(j), \theta(j+1)),$$

where $w(y; \theta, \theta^*) = q(y; \theta)^{-1/2} q(y; \theta^*)^{1/2}$ and insertion points ν_j are chosen uniformly on $\{1, \ldots, K\}$. The initial state $y_1^{(\nu_1)}$ of the first chain is chosen according to $p(y_1^{(\nu_1)} | \theta)$.

3. Latent variable

For $P = 2$, assuming $a = (a_i)_{i \in \mathcal{V}}$ is a collection of indicators $a_i = \rho(i)$, including the statistics $z_L(x) = \deg(x)$, $z_M(x; a) = \sum_{i<j} x_{ij}(a_i + a_j)$, and $z_H(x; a) = \sum_{i<j} x_{ij} a_i a_j$, parameters $(\theta_L, \theta_M, \theta_H)^{\mathrm{T}} \in R^3$, defines a Bernoulli blockmodel (BBM) where the probability of a tie $\{i, j\} \in E$ is $p_{a_i a_j}$, with $\mathrm{logit}(p_{a_i a_j}) = \theta_L + \theta_M(a_i + a_j) + \theta_H a_i a_j$. If a is unobserved estimation may be done as in [16] but when we introduce parameters that correct for lack of independence this is no longer possible (nor is the ML-approach of [20] as $\varkappa$ becomes a non-trivial

function of both θ and the binary a). Without loss of generality we may assume that the dependence is described by the alternating k-triangle statistic [22] $z_T(x;\alpha) = (1+e^{-\alpha})^{-1}\{\deg(x) - \sum_{i<j} x_{ij} e^{-\alpha S_{ij}}\}$, for a smoothing constant $\alpha > 0$, and a parameter θ_T, where $S_{ij} = \sharp\{k : \{i,k\},\{j,k\} \in E \backslash \{i,j\}\}$. For the purpose of estimation we assume that a is another parameter to be estimated, defining $\eta = (\theta^{\mathrm{T}}, a^{\mathrm{T}})^{\mathrm{T}}$, where $\theta = (\theta_L, \theta_M, \theta_H, \theta_T)^{\mathrm{T}}$, as our target of inference for the model $p(x|\eta) = \exp\{g(x;\eta) - \varkappa(\eta)\}$. This now defines a curved ERGM [8] with natural parametrisation $g(x;\eta) = \beta(\eta)^{\mathrm{T}} h(x)$, for $\beta(\eta) = (\theta_L, \theta_T, \theta_{1,2}, \ldots, \theta_{n,n(n-1)})^{\mathrm{T}}$ and $h(x) = (z(x), z_T(x;\alpha), x_{1,2}, \ldots, x_{n,n(n-1)})^{\mathrm{T}}$, where $\theta_{ij} = \theta_M(a_i + a_j) + \theta_H a_i a_j$. LISA only requires that we may evaluate $p(x|\eta)c(\eta)$ for fixed values of η, in which case $p(x|\eta)$ reduces to a regular ERGM with parameter θ and statistics $z(x;a)$.

The model is not fully identified as θ^* may be set so that $p(x;\eta) = p(x;\eta^*)$, $a^* = 1 - a$. This may lead to label switching [12] which may be solved in different ways for BBMs [16] but here there is the additional issue that the model may be "separated" (c.p. [5]) for some realisation of a so that the posterior may be improper for improper priors [11]. To counter these two issues we assume the following partially informative priors: $\pi(\theta_L, \theta_T) \propto c$; $\theta_M|\theta_H, \theta_L, \theta_T \sim N(0, \lambda)$; and $\theta_H|\theta_M, \theta_L, \theta_T \sim N(0, \lambda)$ truncated to the left in θ_L, for shrinkage factor λ.

To illustrate the application of the algorithm we fit the three models in Table 1 to the well-known Kapferer's tailor data that has $n = 39$ actors [9] ($\lambda = 10$ in all models). Model II only includes z_L and $z_T(x;\alpha)$ and is fitted according to [10]; to set $\psi = (\hat{\theta}^{\mathrm{T}}, \hat{a}^{\mathrm{T}})^{\mathrm{T}}$ for Model III, we have used the predicted $\hat{a} = (\hat{a}_i)$ from Model I, and $\hat{\theta}$ is obtained as $\hat{\theta}_{MLE}(\hat{a})$ assuming $\hat{a}$ to be true. A proposal distribution that is consistent with the form of the prior is to draw, in iteration j, $\theta_M^*, \theta_L^*, \theta_T^* \sim N(\theta_M^{(j)}, \theta_L^{(j)}, \theta_T^{(j)}, \Sigma_{124})$, and $\theta_H^* \sim N(\theta_H^{(j)}, \sigma_3)$ truncated to the left in θ_M^*. To set Σ_{124} and σ_3 we have used the rescaled information matrix $I(\hat{\theta}_{MLE}(\hat{a}))^{-1}$. A nearest neighbour proposal is used for $a^*|a^{(j)}$, where $a_i^* := 1 - a_i^{(j)}$ for a number (usually one) of $i \in \mathcal{V}$. In the LIS part we have used a linear map $\eta(t) = \eta t + (1-t)\psi$ (as described in [10]). In other words, for the purposes of implementing LIS, a_i is allowed to be continuous on $[0,1]$. To improve mixing, θ and a are updated in separate blocks which give satisfactory mixing, with the caveats in [10] regarding drawing $p(y_1^{(\nu_1)}|\theta)$ (in lieu of perfect sampling a burn-in of 50,000 is used for model III and for Model II the pseudo perfect sampling scheme of [10] was used; details of the performance may be obtained from the author).

The allocations for the vertices are stable and the measure $H = \frac{8}{n(n-1)} \sum_{i<j} \hat{\pi}_{ij}(1 - \hat{\pi}_{ij})$ [16], where $\hat{\pi}_{ij}$ is the MCMC estimator of $\Pr(a_i = a_j|x)$, is 0.21 and 0.05 for Model I and III respectively. The estimates in Table 1 differ mostly in magnitude between models, most notably for θ_L, but not substantively. The correlation of $\Pr(a_i = 1|x)$ between Model I and III is .976.

4. Summary

We have proposed an algorithm for performing Bayesian inference for ERGMs with latent variables meant to capture unexplained heterogeneity. We have illustrated the application of the algorithm to a well-known data set assuming two classes.

Table 1: Summaries of posteriors for three models fitted to Kapferer's tailors [9]

	Model I		Model II		Model III	
	Mean	Std	Mean	Std	Mean	Std
θ_L	0.83	0.28	-4.24	0.38	-1.58	0.50
θ_M	-2.13	0.25			-2.38	0.26
θ_H	2.04	0.39			2.65	0.34
θ_T			1.37	0.17	1.10	0.20

Further work needs to be done in order to investigate the performance of the algorithm but the principled solution also opens up for further research into how to asses the choice of P; how to make full use of the posterior predictive distributions; and if latent classes may improve the fit of the ERGM.

References

[1] Doreian, P., Batagelj, V., and Ferligoj, A. (2004) *Generalized blockmodeling*. Cambridge University Press, Cambridge.

[2] Fienberg, S.E. and S. Wasserman (1981) Categorical data analysis of single sociometric relations, in: S. Leinhardt (ed.), *Sociological Meth.*. Jossey-Bass, San Francisco, 156–192.

[3] Frank O. and Strauss D. (1986) Markov Graphs. *J. Am. Statist. Association*, **81**, 832–842.

[4] Gelman A., and Meng X.L. (1998) Simulating Normalizing Constants: From Importance Sampling to Bridge Sampling to Path Sampling. *Statistical Science*, **13**, 163–185.

[5] Handcock M.S. (2003) Assessing degeneracy in statistical models of social networks. Working Paper no. 39, Center for Statist. & the Social Sci., Uni Washington. (Av. from http://www.csss.washington.edu/Papers/wp39.pdf).

[6] Handcock M.S., Raftery A.E., and Tantrum J.M. (2007) Model-based clustering for social networks. *J. Roy. Statist. Soc. A*, **170**, 301–354.

[7] Holland, P.W., K.B. Laskey and S. Leinhardt (1983) Stochastic blockmodels: first steps. *Social Networks*, **5**, 109–137.

[8] Hunter D.R., and Handcock M.S. (2006) Inference in Curved Exponential Family Models for Networks. *J. Comp. & Graph. Statistics*, **15**, 565–583.

[9] Kapferer B. (1972) *Strategy and transaction in an African factory*. Manchester University Press, Manchester.

[10] Koskinen, J.H. (2008) The Linked Importance Sampler Auxiliary Variable Metropolis Hastings Algorithm for Distributions with Intractable Normalising Constants. MelNet Tech. Report 08–01, Dep Psych, Uni Melbourne. (Av. from http://www.sna.unimelb.edu.au/publications/MelNet_Techreport_08_01.pdf)

[11] Koskinen J.H., Robins G., & Pattison P. (2008) Analysing Exponential Random Graph (p-star) Models with Missing Data Using Bayesian Data Augmentation. MelNet Tech Rep 08–04, Dep Psych, Uni Melbourne. (Av. from http://www.sna.unimelb.edu.au/publications/MelNet_Techreport_08_04.pdf)

[12] McCulloch R. and Rossi P.E. (1994) An exact likelihood analysis of the multinomial probit model. *J. Econometrics*, **64**, 207–240.

[13] Moreno J.L. (1934) *Who Shall Survive? Foundations of Sociometry, Group Psychotherapy and Sociodrama*. Nervous and Mental Disease Publishing Co, Washington, D.C.

[14] Murray I., Ghahramani Z., and MacKay D.J.C. (2006) MCMC for doubly intractable distributions. Proc. of the 22nd Annual Conference on Uncertainty in Artificial Intelligence (UAI).

[15] Møller J., Pettitt A.N., Berthelsen K.K., and Reeves R.W. (2005) An Efficient Markov Chain Monte Carlo Method for Distributions with Intractable Normalising Constants. *Biometrika*, **93**, 451 – 458.

[16] Nowicki K. and Snijders T.A.B (2001) Estimation and prediction for stochastic blockstructures. *J. Am. Statist. Association*, **96**, 1077–1087.

[17] Neal R. M. (2005) Estimating Ratios of Normalizing Constants Using Linked Importance Sampling. Technical Report No. 0511, Dep of Statistics, Uni Toronto. (available from http://arxiv.org/abs/math.ST/0511216).

[18] Robins G., Elliott P., & Pattison P. (2001) Network models for social selection processes. *Social Networks*, **23**, 1–30.

[19] White H.C., Boorman, S., & Breiger, R.L. (1976) Social structure from multiple networks, I. Blockmodels of roles and positions. *Am. J. Sociology*, **81**, 730–780.

[20] Snijders, T.A.B. (2002) Markov chain Monte Carlo estimation of exponential random graph models. *J. Social Structure*, **3**, April.

[21] Snijders T.A.B. and Nowicki K. (1997) Estimation and prediction for stochastic blockmodels for graphs with latent block structure. *J. Classification*, **14**, 75–100.

[22] Snijders T.A.B., Pattison P.E., Robins, G.L., and Handcock M.S. (2006) New Specifications for Exponential Random Graph Models. *Sociological Meth.*, **36**, 99–153.

[23] Tallberg C. (2005) A Bayesian approach to modeling stochastic blockstructures with covariates. *J. Math. Sociology*, **29**, 1–23.

6th St.Petersburg Workshop on Simulation (2009) 851-855

An adaptive backoff protocol with Markovian contention window control

Andrey Lukyanenko[1], Andrei Gurtov[2], Evsey Morozov[3]

Abstract

Binary Exponential Backoff (BEB) is widely used for sharing a common resource among several stations in communication networks. A general backoff protocol can improve the system throughput but increases the capture effect, permitting one station to seize the channel. In this paper we analyze adaptive backoff protocol, where a station dynamically reduces its contention window after a successful transmission. We derive a solution that will enable computing an optimal reduction for the contention window. Preliminary simulation results indicate that the adaptive backoff protocol can reduce capture effect in Ethernet and wireless networks.

1. Introduction

Binary Exponential Backoff (BEB) is used in many scenarios when sharing of a resource among several stations is needed. When two stations attempt to transmit a packet simultaneously, resulting collision leads to data loss and subsequent need to delay the transmission by one of the stations.

Perhaps the most prominent application of BEB is Medium Access Control (MAC) in Ethernet and Wireless LANs. BEB is also used by transport protocols in the Internet, including TCP, during timeouts. In summary, even a small improvement in backoff performance could have significant impact on real-life applications.

Most systems nowadays implement BEB rather than a generic backoff algorithm for several reasons. BEB offers a simple and quite efficient resource allocation behavior. It is simple to implement in computers with a registry shift operation. However, BEB does not perform optimally in all scenarios as we have showed [5].

BEB has been analyzed extensively in the related work [**?**, 2, 4]. Several researchers attempted developing a generic model of backoff behavior. However, no explicit solution has been obtained due to complexity of the analysis. We made

[1]Helsinki Institute for Information Technology, E-mail: `firstname.secondname@hiit.fi`

[2]Helsinki Institute for Information Technology, E-mail: `gurtov@hiit.fi`

[3]Institute of Applied Mathematical Research, KRC, RAS, E-mail: `emorozov@karelia.ru`; research is supported by RFBR grant 07-07-00888.

a simplifying assumption that a probability of collision p_c in each state is fixed. It simplifies the task significantly and allows to derive optimal parameters for a generic backoff. This assumption is also used in related work [2]. However, this simplification should be validated by measurements in real networks.

Introducing a general backoff where stations increase waiting time before a next transmission attempt by other factor than two significantly improves the performance especially in scenarios with many stations. Unfortunately, it also increases the capture effect, where a station can hog the medium after a successful transmission. Therefore, although general backoff can increase overall system throughput, it does not achieve fair channel allocation among stations.

In this paper, we attempt to develop a model of adaptive backoff where the stations after a successful transmission do not reset their backoff counters. In other words, in case of a collision the station waits for several timeslots instead of two. Such approach eliminates the capture effect while retaining the benefits of higher throughput provided by general backoff.

The rest of the paper is organized as follows. In Section 2, we describe the general backoff protocol, and introduce its adaptive extension in Section 3. In Section 4, we describe an analytical model of the contention window of the new protocol as an irreducible, aperiodic Markov chain. In Section 5, initial evaluation of adaptive backoff is given through simulations. Section 6 presents a summary of main results.

2. Background

The binary exponential backoff protocol (BEB) was introduced in Ethernet [6] and later adopted for several wireless protocols (e.g., IEEE 802.11 [3]). Using the backoff protocol, a station transmits a message depending on the current contention window (CW). CW is a set of successive timeslots; during one random uniformly distributed slot of CW, a station attempts transmitting a message. The message transmissions can collide. After a collision, CW should be increased to decrease probability of further collisions. The message is sent in one of the CW slots or discarded after $M + 1$ unsuccessful transmission attempts. After a successful transmission, CW is reduced back to initial value CW_0.

Backoff protocols differ in how the CW changes depending on the success of transmissions. In BEB, CW is doubled upon a collision. Most backoff protocols reduce the contention window CW to initial window CW_0 upon a successful transmission. Previous work concentrated on studying the constant initial window [**?**, 2, 4, 5]. In this paper, we focus on the dynamic initial window.

3. Adaptive backoff protocol

Consider the following modification of a standard backoff protocol (BP) with $M + 1$ states $0, 1, \ldots, M$ implemented in a communication network. Initially, after $M + 1 < \infty$ unsuccessful transmission attempts (collisions), a packet is discarded. Later on, if an unsuccessful transmission is attempted in state M then the

message is discarded. The probability of collision p_c is assumed to be stationary and independent of the state of the network. A new aspect of the model is that after a successful transmission of a packet in state i backoff restarts in state $j < i$ with a given probability $p_{i,j}$. (We put $p_{0,0} = 1$ and $p_{i,j} = 0$ for $j \geq i$, otherwise.) Thus, we obtain a random walk with jump-up transition p_c and given jump-down transition probabilities $p_{i,j}$ for $j < i$. It is clear that the states of the backoff constitute an irreducible, aperiodic, finite Markov chain Y_n, $n \geq 0$, where Y_n is the state of the backoff after the n-th attempt (successful or not).

To analyze this protocol, in general it is enough to study an embedded Markov chain X formed by the states just after a successful transmission (or discarding), or the Markov chain X^* formed by the states before jump-down. Of course, these chains are strongly connected. It follows that these embedded Markov chains are also aperiodic and irreducible. Thus the corresponding stationary distributions $\pi^* = \{\pi_0^*, \ldots, \pi_M^*\}$ (of the chain X^*) and $\pi = \{\pi_0, \pi_1, \ldots, \pi_M\}$ (of the chain X) exist.

4. Analysis

First, we construct the transition $(M+1) \times (M+1)$ matrix $P = ||q_{i,j}||$ connecting the starting state of window extension and the final state when the first successful transmission occurs (exception is $q_{i,M}$, where successful transmission or discarding occurs). Obviously, $q_{i,j} = (1 - p_c)p_c^{j-i}$ for $0 \leq i \leq j < M$ (and $q_{i,j} = 0$ if $j < i$). Moreover, $q_{i,M} = p_c^{M-i}$. Hence,

$$P = \begin{pmatrix} (1-p_c) & (1-p_c)p_c & (1-p_c)p_c^2 & \ldots & p_c^M \\ 0 & (1-p_c) & (1-p_c)p_c & \ldots & p_c^{M-1} \\ 0 & 0 & (1-p_c) & \ldots & p_c^{M-2} \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & 1 \end{pmatrix}. \tag{1}$$

Introduce also $(M+1) \times (M+1)$ transition matrix $\bar{P} = ||p_{i,j}||$:

$$\bar{P} = \begin{pmatrix} 1 & 0 & 0 & 0 & \ldots & 0 \\ 1 & 0 & 0 & 0 & \ldots & 0 \\ p_{2,0} & p_{2,1} & 0 & 0 & \ldots & 0 \\ p_{3,0} & p_{3,1} & p_{3,2} & 0 & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \\ p_{M,0} & p_{M,1} & p_{M,2} & p_{M,3} & \ldots & 0 \end{pmatrix}. \tag{2}$$

It is obvious that vectors π and π^* are connected as

$$\pi^* = \pi P, \;\; \pi = \pi^* \bar{P}, \tag{3}$$

or $\pi^* = \pi^* \bar{P}P$. In order to solve the equation, we need to find a kernel of matrix $(\bar{P}P - I)^T$, where I is identity matrix and $(\ldots)^T$ is an operation of transposition.

The matrix becomes of the following form

$$\begin{pmatrix} K_{0,0}-1 & K_{1,0} & K_{2,0} & \dots & K_{M-1,0} & K_{M,0} \\ K_{0,1} & K_{1,1}-1 & K_{2,1} & \dots & K_{M-1,1} & K_{M,1} \\ K_{0,2} & K_{1,2} & K_{2,2}-1 & \dots & K_{M-1,2} & K_{M,2} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ K_{0,M} & K_{1,M} & K_{2,M} & \dots & K_{M-1,M} & K_{M,M}-1+p_c \end{pmatrix}, \quad (4)$$

where $K_{i,j} = (1-p_c)\sum_{k=0}^{j} p_c^{j-k} p_{i,k}$. Using the property that $K_{i,j+1} - p_c K_{i,j} = (1-p_c)p_{i,j+1}$ and some algebra, the kernel of the matrix above can be written as the following system of equations:

$$\begin{cases} \pi_0^* = (1-p_c)\sum_{k=0}^{M} p_{k,0}\pi_k^*, \\ \pi_i^* = p_c\pi_{i-1}^* + (1-p_c)\sum_{k=i+1}^{M} p_{k,i}\pi_k^*, \quad 1 \le i \le M-1, \\ \pi_M^* = p_c\pi_{M-1}^* + p_c\pi_M^*. \end{cases} \quad (5)$$

To find the distribution π^* in an explicit form, we denote $a_{i,i} = -1/p_c$ for all i and $a_{i,j} = dp_{i,j}$, where $d = (1-p_c)/p_c$ for all $i > j$. Also let $\alpha_0(i) = 1$, $\alpha_1(i) = \frac{1}{p_c} = -a_{i,i}$, and define recursively (for $0 \le i \le M-2$)

$$\alpha_k(i) = -\sum_{j=0}^{k-1} a_{i+k-1,i+j}\alpha_j(i) \text{ for } 1 \le k \le M-2.$$

After some algebra we obtain that

$$\pi_i^* + \pi_{M-1}^* \sum_{j=i+1}^{M-2} \alpha_{j-i-1}(i+1)a_{M-1,j} + \pi_M^* \sum_{j=i+1}^{M-2} \alpha_{j-i-1}(i+1)a_{M,j} = 0, \quad (6)$$

or (because $\pi_{M-1}^* = d\,\pi_M^*$, see (5))

$$\pi_i^* = -\pi_M^* \sum_{j=i+1}^{M-2} (d\,a_{M-1,j} + a_{M,j})\,\alpha_{j-i-1}(i+1), \ 0 \le i \le M-2. \quad (7)$$

The normalization condition $\sum_{i=0}^{M} \pi_i^* = 1$ allows us to obtain π_M^* in an explicit form:

$$\pi_M^* = \frac{1}{1+d-\sum_{j=0}^{M-2}(d\,a_{M-1,j}+a_{M,j})\sum_{i=0}^{j-1}\alpha_{j-i-1}(i+1)}. \quad (8)$$

Finally we obtain for $0 \le k \le M-2$

$$\pi_k^* = \frac{-\sum_{j=i+1}^{M-2}(d\,a_{M-1,j}+a_{M,j})\,\alpha_{j-i-1}(i+1)}{1+d-\sum_{j=0}^{M-2}(d\,a_{M-1,j}+a_{M,j})\sum_{i=0}^{j-1}\alpha_{j-i-1}(i+1)}, \quad (9)$$

and

$$\pi_{M-1}^* = \frac{d}{1+d-\sum_{j=0}^{M-2}(d\,a_{M-1,j}+a_{M,j})\sum_{i=0}^{j-1}\alpha_{j-i-1}(i+1)}. \quad (10)$$

We also know that $\pi_i = \sum_{k=i+1}^{M} \pi_k^* p_{k,i}$, thus, we can find distribution π from π^* as following

$$\begin{cases} \pi_0^* = (1-p_c)\pi_0, \\ \pi_i^* = p_c\pi_{i-1}^* + (1-p_c)\pi_i, \quad 1 \le i \le M-1, \\ \pi_M^* = p_c\pi_{M-1}^* + p_c\pi_M^*. \end{cases} \tag{11}$$

Thus, we obtain distributions π and π^* in an explicit form. This is only a preliminary analysis but nevertheless it allows to calculate various stationary performance measures describing the new adaptive backoff protocol. Note that conditions ensuring positiveness of π_k require further analysis.

The analysis of stationary distribution of states is required to study the adaptive backoff protocol. In previous work, we studied general backoff protocol [5], with geometrically distributed states. It was based on the fact that starting from the initial state, the current state increased with probability p_c and decreased to the initial state with probability $1-p_c$. Adaptive backoff does not have such property and we need to find the distribution explicitly. Using the distribution and knowing the holding time in each state, we can obtain the average service time for a message (before transmission or discarding) as in previous work [5].

5. Simulations

In this section, we describe simulations of adaptive backoff protocol. Although the study is incomplete, we provide intermediate results and scenarios for future simulations. For simulations, we use the ns-2 simulator, with necessary modifications in the backoff protocol.

Service time for a packet is defined as the difference between the time when the packet is on the top of the MAC layer queue ready to be sent, and the time when it successfully leaves the MAC layer. All simulations were carried out for 10 seconds over a 10 Mbps link.

We simulated the standard backoff protocol with different ratio of increase of CWs (in BEB CW doubles after each collision, the ratio for standard BP is 2). We simulated backoff protocols with different ratios $1.1, \ldots, 2.9$, with step 0.1. Our goal is to decrease the service time for a station. The simulation showed that in addition to reduction of service times, a well-known problem of capture effect is strengthened. The stations are behaving heterogeneously, some sending plenty of messages, while others unable to send even 50 messages. With a truncated BEB protocol, stations send about 200 messages each.

It appears that the adaptive backoff protocol can deal better with the capture effect. The reason for a capture effect is that after a successful transmission over a heavily loaded channel, the station returns to the initial window CW_0 and with high probability will get access to the channel again. If an adaptive backoff protocol were used, then station would not return to CW_0 (which corresponds to state 0 in the standard model) but instead returns to some intermediate state. In our simulation, we have tested returning to CW which is multiplication of the previous CW by $\frac{1}{3}$, $\frac{1}{2}$, and $\frac{2}{3}$.

The number of dropped packets has been greatly reduced. The number of dropped packets was less than one hundred in any of these simulations. A protocol with $\frac{CW}{3}$ behaves better than others. Although the protocol $\frac{2CW}{3}$ has much less service time, the deviation (one station sends a lot, while another cannot send) is high.

We are going to simulate the adaptive backoff. In particular, we are interested to consider a case where, for some l, $p_{i,i-l} = 1$ for $i \geq l$, and $p_{i,0} = 1$ for $i < l$. Another example is dynamically changing returning states when $p_{i,j} = 1$ if $j = \lfloor i/k \rfloor$ for some $k \geq 2$.

6. Conclusion

We suggested a new model to describe the behavior of a contention window in general (not necessary exponential) backoff as a irreducible, aperiodic, finite Markov chain. To analyze this adaptive protocol, we studied the corresponding random walk describing the dynamics of the contention window.

The original Markov chain is replaced by an embedded Markov chain, and the stationary distribution of the latter chain is obtained in an explicit form. The result enables computation of the optimal contention window after a successful transmission. Other stationary characteristics require further research.

Preliminary simulation results indicate that the adaptive backoff protocol can improve throughput and reduce capture effect in Ethernet and wireless networks.

References

[1] Aldous D. (1987). Ultimate instability of exponential back-off protocol for acknowledgment-based transmission control of random access communication channels. *Information Theory, IEEE Transactions on*, 33(2), 219–223.

[2] Bianchi G. (2000). Performance analysis of the IEEE 802.11 distributed coordination function. *IEEE Journal on Selected Areas in Communications*, 18(3), 535–547.

[3] IEEE 802.11 LAN/MAN Wireless LANS. http://standards.ieee.org/getieee802/802.11.html

[4] Hastad J., Leighton T., Rogoff B. (1987). Analysis of backoff protocols for multiple access channels. In: *STOC '87: Proceedings of the nineteenth annual ACM symposium on Theory of computing*, 241–253.

[5] Lukyanenko A., Gurtov A. (2008). Performance analysis of general backoff protocols. *Journal of Communications Software and Systems*, 4(1).

[6] Metcalfe R.M., Boggs D.R. (1976). Ethernet: distributed packet switching for local computer networks. *Commun. ACM*, 19(7).

6th St.Petersburg Workshop on Simulation (2009) 857-861

On estimations of periodically non-stationary stochastic automaton behavior under fuzzy conditions[1]

E.N. Mosyagina [2]

Abstract

Optimal upper and lower estimations of a fuzzy given goal reaching degree by non-stationary stochastic automaton with periodically varying structure under fuzzy reacting surroundings were determined.

1. Basic definition

Periodically non-stationary generalized stochastic finite automaton is a system

$$\mathcal{A}_{pv} = \langle X^{(\tau)}, A^{(\tau)}, Y^{(\tau)}, a_{i_0}, \{\mathbf{P}^{(\tau)}(x_s, y_l)\}, t_0, T, t_p \rangle, \tag{1}$$

where t_0 is a preperiod length; T is a period of automaton structure parameters repetition; t_p is a postperiod length; $\tau = \tau(t)$ is a structural tack number, defined through current tact number $t = 0, 1, 2 \ldots$ in the following way:

$$\tau = \begin{cases} t, & 0 < t \le t_0, \\ (t - t_0)(mod T) + t_0, & t_0 + 1 \le t \le t_0 + nT, \\ (t - t_0)(mod T) + t_0 + T - 1, & t_0 + nT + 1 \le t \le t_0 + nT + t_p, \end{cases} \tag{2}$$

where following contingencies were accepted $0 < t_0 < T$, $0 < t_p < T$, n is an integer of period T repetitions, $n \ge 1$; $X^{(\tau)} = X^{(\tau(t))}$, $Y^{(\tau)} = Y^{(\tau(t))}$ are the alphabets of input and output symbols admissible at the τth tact, $|X^{(\tau)}| = n_\tau$, $|Y^{(\tau)}| = k_\tau$, $\tau = \overline{1, t_0 + T + t_p - 1}$; $A^{(\tau)} = A^{(\tau(t))}$ is the alphabet of automaton states at the τth tact, $|A^{(\tau)}| = m_\tau$, $\tau = \overline{0, t_0 + T + t_p - 1}$, $a_{i_0} \in A^{(0)}$ is an initial state; $\{\mathbf{P}^{(\tau)}(x_s, y_l)\}$ is a system of $(m_{\tau(t-1)} \times m_{\tau(t)})$-matrices, $x_s \in X^{(\tau)}$, $y_l \in Y^{(\tau)}$, $\tau = \overline{1, t_0 + T + t_p - 1}$, whose elements define conditional transition probabilities $P^{(\tau(t))}_{i_{t-1} i_t}(x_{s_t}, y_{l_t}) = P^{(\tau(t))}(a_{i_t} y_{l_t} | a_{i_{t-1}} x_{s_t})$ at the tact t to the state a_{i_t} with distribution of the symbol y_{l_t} for every pare $(a_{i_{t-1}}, x_{s_t})$, where $a_{i_{t-1}} \in A^{(\tau(t-1))}$, $a_{i_t} \in A^{(\tau(t))}$, $x_{s_t} \in X^{(\tau(t))}$, $y_{l_t} \in Y^{(\tau(t))}$.

[1]This work was supported by RFBR (grant 07-01-00355)

[2]Math. Department, St.Petersburg State University.

Arrange to denote *fuzzy surroundings* following system

$$\mathcal{C} = \langle \mathbf{C}^{\tau}, \tau = \overline{1, t_0 + T + t_p - 1} \rangle, \tag{3}$$

where $\mathbf{C}^{\tau} = (C^{\tau(t)}_{l_{\tau(t-1)}}(x_{s_{\tau(t)}}))$ is $(k_{\tau-1} \times n_{\tau})$-matrix of fuzzy restriction, being set by surroundings on input symbols $x_{s_t} \in X^{(\tau(t))}$ of automaton $\mathcal{A}_{pv}$ at the tact t, if automaton acted upon surroundings by output symbol $y_{l_{t-1}} \in Y^{(\tau(t-1))}$ at the previous tact. Elements of matrices (3) define fuzzy sets in $X^{(\tau(t))}$ at different $y_{l_{t-1}} \in Y^{(\tau(t-1))}$, representing a value of membership function $\mu^{(\tau(t))}_{l_{t-1}}(x_{s_t})$, $x_{s_t} \in \in X^{(\tau(t))}$, $y_{l_{t-1}} \in Y^{(\tau(t-1))}$, $\tau = \overline{1, t_0 + T + t_p - 1}$, taking values from the interval [0,1].

2. Problem setting

The automaton $\mathcal{A}_{pv}$ is in interaction with some surroundings $\mathcal{C}$, which is described by restrictions matrices $\mathbf{C}^{\tau(t)}$ on input symbols. At each current tact $t-1$ the automaton (1) returns a symbol $y_{l_{t-1}}$ acting on surroundings $\mathcal{C}$, imposing in turn fuzzy restrictions $C^{\tau(t)}_{l_{t-1}}(x_{s_t})$ at the tact t on input control symbols $x_{s_t} \in X^{(\tau(t))}$ of automaton (1) against taken at the previous tact output symbol $y_{l_{t-1}}$. Besides, for the fixed structural automaton tacts $\tau_N = N$ and $\tau_M = M$ such, that

$$\tau_N = N = t_0, \qquad \tau_M = M = t_0 + T + t_p - 1,$$

fuzzy goals are given. *Fuzzy goals* is fuzzy sets G^N and G^M, being defined by the membership functions $\mu_{G^N}(a_i, y_l)$, $a_i \in A^{(N)}$, $y_l \in Y^{(N)}$, and $\mu_{G^M}(a_i, y_l)$, $a_i \in A^{(M)}$, $y_l \in Y^{(M)}$.

Conventionally speaking, exterior "observer" controls on automaton $\mathcal{A}_{pv}$, interacting with some surroundings C, with help of feeding on automaton of input symbols sequence from the alphabets $X^{(\tau(t))}$. Whereas the automaton state is known to him only at the initial moment $t = 0$ and at the moment $t = t_N$, where $\tau(t_N) = N$. At remaining current tacts $t > 0$ the "observer" and surroundings C take only exterior reaction y_{l_t} of automaton $\mathcal{A}_{pv}$ on input symbol x_{s_t}, but they are unaware of state a_{i_t}, in witch automaton has transited. The problem is to find an optimal input actions at structural tacts $\tau = \overline{1, t_0 + T + t_p - 1}$, permitting to take upper and lower estimates of given goal grade of membership at the tacts t_N and t_M. Whereas every actions sequence will be formed for maximum possible estimates of grade of membership of regular given goal result reaching under optimal solution getting condition on a previous stage.

3. Method of solution

Divide automaton structural tacts into three parts $\tau = 0, 1, \ldots, N$, $\tau = N, N+$ $+1, \ldots, N + T - 1, N$, $\tau = N, N + T, \ldots, M$ (i. e. present, in fact, automaton $\mathcal{A}_{pv}$ as a sequence of three automata $\mathcal{A}'_{pv}$, $\mathcal{A}''_{pv}$, $\mathcal{A}'''_{pv}$) and find an optimal input actions for every part. Begin the setting problem solution with optimal sequences

determination for automaton $\mathcal{A}'_{pv}$ with initial states $a_{i_0} \in A^{(0)}$ and fixed process termination time $t_N(n) = N + nT$ for $n = 0$, i. e. $t_N(0) = N$. Whereas fuzzy restrictions, being set by surroundings $\mathcal{C}$ at $t = \tau = \overline{0, N}$, and fuzzy goal at the moment $t_N(0) = N$ are known.

Present a solution as [1]

$$D = C^1 \cap \ldots \cap C^N \cap G^N,$$

where $C^1 = C^1_{l_0}(x_{s_1})$ is a restriction, being imposed on input symbol x_{s_1}, independently of output symbol, as at the tact $t = 0$ single output symbol is an "empty" symbol $y_{l_0} = e$.

Consider the fuzzy goal G^N as a fuzzy event in space $A^{(N)} \times Y^{(N)}$. Conditional probability of this event at fixed $a_{i_{N-1}}$, x_{s_N} is expressed by the formula

$$Pr(G^N | a_{i_{N-1}} x_{s_N}) = E\mu_{G^N} = \sum_{(a_{i_N}, y_{l_N})} P^{(N)}(a_{i_N} y_{l_N} | a_{i_{N-1}} x_{s_N}) \mu_{G^N}(a_{i_N}, y_{l_N}). \quad (4)$$

Let control action and automaton reaction be $w = x_{s_1}, \ldots x_{s_N}$ and $v = y_{l_1}, \ldots y_{l_{N-1}}$ accordingly, then

$$\mu_D(w|v) = \min(\mu^{(1)}_e(x_{s_1}), \ldots, \mu^{(N)}_{l_{N-1}}(x_{s_N}), E\mu_{G^N}). \quad (5)$$

It is necessary to estimate magnitude μ_D against input sequences w. Introduce following notation: w^{opt} is an input symbols sequence, enabling to take maximum lower and upper estimates μ_D. Present a solution as $x_{s_t} = \pi_{\tau(t)}(y_{l_{t-1}})$, $t = \overline{1, N}$, where $\pi_{\tau(t)}$ – accepted choice rule x_{s_t} against $y_{l_{t-1}}$, being called *optimal strategy.* Apply dynamic programming method [1-3] for forming required sequence, take into account, whereas, that magnitude $\mu_D(w^{opt}|v)$ will be obtained "fuzzily" as a lower and upper estimates interval, since optimal choice of input symbol x_{s_t} at each tact t (except for initial tact $t = 1$) only executed against automaton reaction $y_{l_{t-1}}$ at the previous tact without taking its states $a_{i_{t-1}} \in A^{(\tau(t-1))}$ into consideration. According to expression (5) it is possible to note

$$\mu_D(w^{opt}|v) = \max_{x_{s_1}, \ldots, x_{s_N}} \min(\mu^{(1)}_e(x_{s_1}), \ldots, \mu^{(N)}_{l_{N-1}}(x_{s_N}), E\mu_{G^N}), \quad (6)$$

where according to the formula (4) the magnitude $E\mu_{G^N}$ is a function of $a_{i_{N-1}}$, x_{s_N} . Designate, using (6)

$$\mu_{G^{N-1}}(y_{l_{N-1}}) = \min(\mu^{(N)}_{l_{N-1}}(x_{s_N}), E\mu_{G^N}), \quad (7)$$

$$\mu_{G^{N-1}} = \mu_{G^{N-1}}(a_{i_{N-1}}, y_{l_{N-1}}) = \max_{x_{s_N}} \mu_{G^{N-1}}(y_{l_{N-1}}), \quad (8)$$

where $\mu_{G^{N-1}}(y_{l_{N-1}})$ is a function of $a_{i_{N-1}}$, x_{s_N} at fixed $y_{l_{N-1}} \in Y^{(\tau(N-1))}$, and value x_{s_N}, and estimates of magnitudes $\mu_{G^{N-1}}$ in (8) are defined with help of following procedure.

Form the tables of magnitudes $\mu_{G^{N-1}}(y_{l_{N-1}})$, $y_{l_{N-1}} \in Y^{(\tau(N-1))}$, which lines correspond to various states $a_{i_{N-1}} \in A^{(\tau(N-1))}$, and columns – input actions $x_{s_N} \in$

$\in X^{(\tau(N))}$. Then in each line $a_{i_{N-1}}$, $a_{i_{N-1}} \in A^{(\tau(N-1))}$, of the table $\mu_{G^{N-1}}(y_{l_{N-1}})$ choose largest element and single out columns, containing even one chosen element. Every such chosen in the table $\mu_{G^{N-1}}(y_{l_{N-1}})$ column x_{s_N} will correspond to possible value $\pi_N(y_{l_{N-1}}) = x_{s_N}$ and estimate $\mu_{G^{N-1}} \in [\mu^{min}_{G^{N-1}}, \mu^{max}_{G^{N-1}}]$, where

$$\mu^{min}_{G^{N-1}} = \min_{a_{i_{N-1}}} \mu_{G^{N-1}}(y_{l_{N-1}}),\ \mu^{max}_{G^{N-1}} = \max_{a_{i_{N-1}}} \mu_{G^{N-1}}(y_{l_{N-1}}). \tag{9}$$

Having taken one by one chosen column in each table $\mu_{G^{N-1}}(y_{l_{N-1}})$, get one of possible strategy variant $\pi_N(y_{l_{N-1}}) = x_{s_N},\ y_{l_{N-1}} \in Y^{(\tau(N-1))},\ x_{s_N} \in X^{(\tau(N))}$, at that, if in each table is singled out only one column, at once constructed strategy is optimal. Even though there are tables where is singled out more then one column, different columns combinations, taken from the tables one by one, determine a possible strategies set. In the latter case optimal strategy is chosen on basis of corresponding estimates analysis (9). Whereas, if estimates comparison at various $y_{l_{N-1}} \in Y^{(\tau(N-1))}$ allows to choose single optimal strategy at the structural tact τ_N, remaining strategies further are not considered. Otherwise the best are left and also not comparable by estimates strategies, and optimal strategy is chosen by the following iteration results.

For every abandoned strategy from corresponding to it chosen in the tables $\mu_{G^{N-1}}(y_{l_{N-1}})$, $y_{l_{N-1}} \in Y^{(\tau(N-1))}$ columns is formed table of functions $\mu_{G^{N-1}} = = \mu_{G^{N-1}}(a_{i_{N-1}}, y_{l_{N-1}})$, whose lines correspond to various states , and columns – output symbols. Then equality (6) is reduced to

$$\mu_D(w^{opt}|v) = \max_{x_{s_1},\dots x_{s_{N-1}}} \min(\mu^{(1)}_e(x_{s_1}), \dots, \mu^{(N-1)}_{l_{N-2}}(x_{s_{N-1}}), \mu_{G^{N-1}}(a_{i_{N-1}}, y_{l_{N-1}})).$$

Get a recurring equations system, reiterating inverse iterations process

$$\mu_{G^{N-\nu}}(y_{l_{N-\nu}}) = \min(\mu^{(N-\nu+1)}_{l_{N-\nu}}(x_{s_{N-\nu+1}}), E\mu_{G^{N-\nu+1}}), \tag{10}$$

$$\mu_{G^{N-\nu}} = \mu_{G^{N-\nu}}(a_{i_{N-\nu}}, y_{l_{N-\nu}}) = \max_{x_{s_{N-\nu+1}}} \mu_{G^{N-\nu}}(y_{l_{N-\nu}}),$$

$$E\mu_{G^{N-\nu+1}} = \sum_{(a_{i_{N-\nu+1}}, y_{l_{N-\nu+1}})} P^{(N-\nu+1)}(a_{i_{N-\nu+1}} y_{l_{N-\nu+1}} | a_{i_{N-\nu}} x_{s_{N-\nu+1}}) \times \tag{11}$$

$$\times \mu_{G^{N-\nu+1}}(a_{i_{N-\nu+1}}, y_{l_{N-\nu+1}}).$$

Required sequences w^{opt} for various reactions v are got by sequential strategy values determination $\pi_{N-\nu+1}(y_{l_{N-\nu}}) = x_{s_{N-\nu+1}},\ \nu = \overline{1, N}$, corresponding to the best estimates $\mu_{G^{N-\nu}}$.

Consider now periodical part of the automaton $\mathcal{A}_{pv}$ (1), as a new automaton $\mathcal{A}''_{pv}$. Lock in beginning time $t_N = N$ and termination time of the process $t_N = = N + T$. As an initial conditions of the automaton $\mathcal{A}''_{pv}$ choose states and output symbols y_{l_N} of the automaton $\mathcal{A}'_{pv}$ at the moment $t_N = N$. Optimal strategies for the automaton $\mathcal{A}''_{pv}$ are constructed by the same method.

For the automaton $\mathcal{A}'''_{pv}$ with a fix beginning time $t_N = N + nT$ and termination time of the process $t_M = M$, initial state and output symbol, corresponding to $\mathcal{A}''_{pv}$, $t_N = N + nT$, required optimal actions sequences are formed as well as for the automaton $\mathcal{A}'_{pv}$ and $\mathcal{A}''_{pv}$.

As a result for all parts of the automaton $\mathcal{A}_{pv}$ will be formed corresponding optimal strategies.

4. Example

As an example consider the automaton A_{pv} (1), where $t_0 = 2$, $T = 3$, $t_p = 2$, $X^{(1)} = X^{(3)} = X^{(5)} = \{x_2, x_4\}$, $X^{(2)} = X^{(6)} = \{x_1, x_2, x_3\}$, $X^{(4)} = \{x_2, x_5, x_6\}$, $A^{(0)} = A^{(1)} = A^{(3)} = A^{(4)} = A^{(5)} = A^{(6)} = \{a_1, a_2\}$, $A^{(2)} = \{a_1, a_2, a_3\}$, $Y^{(1)} = = Y^{(3)} = Y^{(4)} = Y^{(5)} = \{y_1, y_2, y_3\}$, $Y^{(2)} = \{y_2, y_3\}$, $Y^{(6)} = \{y_1, y_2\}$ and matrices of transition probabilities $P^{(\tau)}(x_s, y_l)$, $\tau = \overline{1,6}$ are specified in the following way:

$\mathbf{P}^{(1)}$	y_1		y_2		y_3	
x_2	0,2	0,3	0,1	0,2	0,2	0
	0,3	0	0	0,1	0,5	0,1
x_4	0,4	0	0,4	0	0,1	0,1
	0,1	0,2	0	0,2	0,3	0,2

$\mathbf{P}^{(2)}$	y_2			y_3		
x_1	0,5	0	0,1	0,2	0,1	0,1
	0	0,3	0	0	0,6	0,1
x_2	0,3	0,1	0,1	0,2	0,1	0,2
	0,4	0	0	0,4	0,1	0,1
x_3	0,4	0	0,2	0	0	0,4
	0,1	0,1	0	0,2	0,3	0,3

$\mathbf{P}^{(3)}, \mathbf{P}^{(5)}$	y_1		y_2		y_3	
x_2	0,3	0,2	0,1	0	0,4	0
	0	0,4	0,1	0,2	0	0,3
	0,7	0	0	0,2	0,1	0
x_4	0,1	0,1	0,1	0,4	0,1	0,2
	0,2	0	0	0,2	0,3	0,3
	0,1	0	0,3	0	0,2	0,4

$\mathbf{P}^{(4)}$	y_1		y_2		y_3	
x_2	0,3	0,1	0,2	0,1	0,3	0
	0,1	0,2	0	0,4	0	0,3
x_5	0	0,5	0,1	0,1	0	0,3
	0,3	0,1	0,3	0	0,1	0,2
x_6	0,2	0,3	0,1	0,1	0,1	0,2
	0,6	0	0,2	0,1	0,1	0

$\mathbf{P}^{(6)}$	y_1		y_2	
x_1	0,1	0,2	0,6	0,1
	0,1	0,7	0	0,2
x_2	0,4	0,3	0,1	0,2
	0	0,3	0,4	0,3
x_3	0	0,7	0,2	0,1
	0,3	0,1	0	0,6

For structural tacts $\tau_N = N = 2$ and $\tau_M = M = 6$ fuzzy goals are given, being defined by the membership functions μ_{G^2}, μ_{G^6} with values

μ_{G^2}	y_2	y_3
a_1	0,5	0,7
a_2	0,8	0,7
a_3	0,6	0,9

μ_{G^6}	y_1	y_2
a_1	0,9	0,9
a_2	0,7	0,5

Automaton $\mathcal{A}_{pv}$ interacts with fuzzy surroundings $\mathcal{C}$, which, taking output symbols of the automaton $y_{l_{\tau(t-1)}}$, applies following fuzzy restrictions $\mathbf{C}^{\tau}$ on input symbols $x_{s_{\tau(t)}}$ at structural tacts $\tau = \overline{1,6}$. Let $\mu^{(\tau)} = \mu^{(\tau(t))}_{l_{t-1}}(x_{s_t})$.

$\mu^{(1)}$	x_2	x_4
e	0,4	0,8

$\mu^{(2)}$	x_1	x_2	x_3
y_1	0,9	0,5	0,6
y_2	0,4	0,7	0,6
y_3	0,7	0,8	0,3

$\mu^{(3)}$	x_2	x_4
y_2	0,5	0,3
y_3	0,4	0,7

$\mu^{(4)}$	x_2	x_5	x_6
y_1	0,7	0,6	0,6
y_2	0,5	0,8	0,7
y_3	0,6	0,7	0,5

$\mu^{(5)}$	x_2	x_4
y_2	0,5	0,3
y_3	0,4	0,7

$\mu^{(6)}$	x_1	x_2	x_3
y_1	0,8	0,7	0,5
y_2	0,3	0,4	0,6
y_3	0,8	0,9	0,9

It is required to choose controlling actions $x_{s_t} = \pi_{\tau(t)}(y_{l_{t-1}})$, $\tau = \overline{1,6}$ on automaton $\mathcal{A}_{pv}$ at various output reactions $y_{l_{\tau(t-1)}} \in Y^{(\tau(t-1))}$.

Represent given optimal control strategies for cases $n = 0$, $t = 1, 2$, and $n = 1$, $t = 3, 4, 5$, correspondingly leading to the first and the second hit at the structural tact $\tau_N = 2$, as following tables:

t	$\tau(t)$	$\pi_{\tau(t)}(y_{l_{t-1}}) = x_{s_t}$	$\mu_{G^{t-1}}$
2	2	$\pi_2(y_1) = x_1$	$\mu_{G^1} \in [0,61; 0,75]$
		$\pi_2(y_2) = x_2$	$\mu_{G^1} \in [0,64; 0,68]$
		$\pi_2(y_3) = x_2$	$\mu_{G^1} \in [0,64; 0,68]$
1	1	$\pi_1(e) = x_4$	$\mu_{G^0} \in [0,648; 0,671]$

t	$\tau(t)$	$\pi_{\tau(t)}(y_{l_{t-1}}) = x_{s_t}$	$\mu_{G^{t-1}}$
5	2	$\pi'_2(y_1) = x_1$	$\mu_{G^4} \in [0,61; 0,75]$
		$\pi'_2(y_2) = x_2$	$\mu_{G^4} \in [0,64; 0,68]$
		$\pi'_2(y_3) = x_1$	$\mu_{G^4} \in [0,61; 0,7\ ]$
4	4	$\pi_4(y_1) = x_2$	$\mu_{G^3} \in [0,641; 0,677]$
		$\pi_4(y_2) = x_5$	$\mu_{G^3} \in [0,663; 0,717]$
		$\pi_4(y_3) = x_5$	$\mu_{G^3} \in [0,663; 0,7]$
3	3	$\pi_3(y_2) = x_2$	$\mu_{G^2} \in [0,5; 0,5]$
		$\pi_3(y_3) = x_4$	$\mu_{G^2} \in [0,6697; 0,6844]$

For the postperiod and tacts $t = 6, 7$, leading the automaton to the structural tact $\tau_M = 6$, we get

t	$\tau(t)$	$\pi_{\tau(t)}(y_{l_{t-1}}) = x_{s_t}$	$\mu_{G^{t-1}}$
7	6	$\pi_6(y_1) = x_1$	$\mu_{G^6} \in [0,68; 0,8]$
		$\pi_6(y_2) = x_3$	$\mu_{G^6} \in [0,6; 0,6]$
		$\pi_6(y_3) = x_2$	$\mu_{G^6} \in [0,72; 0,76]$
6	5	$\pi_5(y_2) = x_2$	$\mu_{G^5} \in [0,5; 0,5]$
		$\pi_5(y_3) = x_4$	$\mu_{G^5} \in [0,668; 0,7]$

References

[1] Bellman R.E., Zadeh L.A. (1970) *Decision-Making in a Fuzzy Environment*// Management Science, 17, 4, 141-164.

[2] Mosyagina E.N., Tchirkov M.K. (2006) *Optimal strategies of action on periodically nonstationary stochastic automaton in the fuzzy set conditions*// Stochastic optimization in computer science. vol.2. St.Petersburg, 134-146.(in Russian)

[3] Bellman R.E. (1957) *Dynamic Programming*. Princeton.

6th St.Petersburg Workshop on Simulation (2009) 863-867

On Decomposition of Fuzzy Automata Models[1]

Khokhulina V.A.[2], Tchirkov M.K.[2]

Abstract

In this article it is shown, that every stationary fuzzy automaton model (generalized fuzzy automaton) may put in correspondence with a fuzzy set of generalized non-deterministic automata prescribed above the Boolean lattice, which union is a decomposition of initial generalized fuzzy automaton on various fuzziness levels, taking into account the grades of their membership to this fuzzy set.

1. Introduction

In the theory of fuzzy sets [1] it is shown the possibility of the representation of any finite fuzzy set as a union of level clear sets "multiplied" by the corresponding fuzziness level. In the theory of stochastic automata [2] there exist a method of synthesis of any stochastic automaton as a deterministic automaton with random input which is actually a union of the deterministic automata finite number to each of which corresponds with some probability of its choice at every tact in dependence on the input random symbol. Therefore, the research of a similar problem for the case of fuzzy automata models [3, 4] is obviously justified. Exactly to this problem with the reference to so called generalized fuzzy automata [4] the given article is devoted to.

2. Basic definitions

Let's consider the complete distributive lattice $\mathcal{L} = \langle [0,1],\ \max, \min,\ \geqslant \rangle$, i.e. a closed interval [0,1] with operations (where $a, b \in [0,1]$)

$$a + b = \max(a, b), \qquad ab = \min(a, b), \tag{1}$$

are conventionally called "addition", "multiplication", and usual ordering. Also agree to designate the set of all $(m \times n)$-matrices above $\mathcal{L}$ as $\mathcal{L}^{m,n}$.

A generalized fuzzy finite automation is called [4] a system

$$\mathcal{A}_f = \langle X,\ A,\ Y,\ \mathbf{r},\ \{\mathbf{F}(x,y)\},\ \mathbf{q} \rangle \tag{2}$$

[1]This work was supported by RFBR (grant 07-01-00355).

[2]Math.Department, St.Petersburg State University, E-mail: vakh08@mail.ru

where X, A, Y are alphabets of inputs, states and outputs, $|X| = n$, $|A| = m$, $|Y| = k$; $\mathbf{r} \in \mathcal{L}^{1,m}$ is an initial row vector (grades of the states membership to the initial states set); $\mathbf{q} \in \mathcal{L}^{m,1}$ is a final column vector (the grades of the states membership to the finale states set) and $\{\mathbf{F}(x,y)\}$ is a set of transition matrices (the grades of a transition states membership to a set of various transitions):

$$\{\mathbf{F}(x,y)\} = \{\mathbf{F}(x,y)|\mathbf{F}(x,y) \in \mathcal{L}^{m,m}, \quad x \in X, \quad y \in Y\},$$

setting the mapping $X \times Y \to \mathcal{L}^{m,m}$.

A fuzzy mapping $\Phi_f : X^d \times Y^d \to [0,1], \quad d = 0,1,2,\ldots,$ which is induced by the fuzzy automaton $\mathcal{A}_f$, is determined by the membership function (where, later on, "addition" and "multiplication" are considered as the operations (1) and $w = x_{s_1}x_{s_2}\ldots x_{s_d}$, $v = y_{l_1}y_{l_2}\ldots y_{l_d}$ are the input and output words)

$$\Phi_f(w,v) = \mathbf{r}\prod_{t=1}^{d}\mathbf{F}(x_{s_t}, y_{l_t})\mathbf{q} \tag{3}$$

for every $(w,v) \in X^d \times Y^d, d = 0,1,2,\ldots,$ and thus it sets a fuzzy subset of pairs of input-output words with the same, but arbitrary large length with the membership function (3).

In particular, when the distributive lattice $\mathcal{L}$ is *the Boolean lattice*

$$\mathcal{L}_1 = \langle\{0,1\},\ \vee,\ \&,\ 1 \geqslant 0\rangle,$$

the system (2) is called *the generalized non-deterministic automaton* [5]

$$\mathcal{A}_{nd} = \langle X,\ A,\ Y,\ \mathbf{d},\ \{\mathbf{D}(x,y)\},\ \mathbf{g}\rangle, \tag{4}$$

where the row vector of initial states $\mathbf{d} \in \mathcal{L}_1^{1,m}$, the column vector of final states $\mathbf{g} \in \mathcal{L}_1^{m,1}$ and

$$\{\mathbf{D}(x,y)\} = \{\mathbf{D}(x,y)|\mathbf{D}(x,y) \in \mathcal{L}_1^{m,m},\ x \in X,\ y \in Y\}$$

is a set of transition matrices. The non-deterministic automaton $\mathcal{A}_{nd}$ induces *a non-deterministic mapping* $\Phi_{nd} : X^d \times Y^d \to \{0,1\}$, $d = 0,1,2,\ldots,$ determined by the expression

$$\Phi_{nd}(w,v) = \mathbf{d}\prod_{t=1}^{d}\mathbf{D}(x_{s_t}, y_{l_t})\mathbf{g}, \tag{5}$$

where under "addition" and "multiplication" the operations $\vee$ and & are meant.

3. Decomposition of fuzzy matrices product

Let's consider a finite set of fuzzy $(m \times m)$-matrices $\mathbf{F}(x,y) \in \mathcal{L}^{m,m}$, $x \in X$, $y \in Y$, $\mathbf{F}(x,y) = (F_{ij}(x,y))_{m,m}$, where $F_{ij}(x,y) \in [0,1]$. Taking into account the finiteness of the alphabets X, Y and the values of indexes i, j, the matrices $\mathbf{F}(x,y)$,

$x \in X$, $y \in Y$, contain a finite number values $F_{ij}(x,y)$. Let's denote a set of such various values, which we agree to call *fuzziness levels*, as

$$\mu_f = \{\mu_f^{(1)}, \mu_f^{(2)}, \ldots, \mu_f^{(q)}\}, \tag{6}$$

having arranged $\mu_f^{(\nu)}$ in order of their value increase, i.e. it will be considered that $\mu_f^{(\nu-1)} < \mu_f^{(\nu)}$ for all ν. Using (6) for each matrix $\mathbf{F}(x,y)$ and $\mu_f^{(\nu)}$ let's enter $(m \times m)$-matrix $\mathbf{D}^{(\nu)}(x,y) = (D_{ij}^{(\nu)})_{m,m}$, $D_{ij}^{(\nu)} \in \{0,1\}$, such, as

$$D_{ij}^{(\nu)}(x,y) = \begin{cases} 1 & \text{when } F_{ij}(x,y) \geqslant \mu_f^{(\nu)}, \\ 0 & \text{when } F_{ij}(x,y) < \mu_f^{(\nu)}. \end{cases} \tag{7}$$

In this case, considering (1) and (7), it is possible to represent each fuzzy matrix $\mathbf{F}(x,y)$ as an "addition" of matrices "product" $\mathbf{D}^{(\nu)}(x,y)$ on the corresponding fuzziness levels $\mu_f^{(\nu)}$, $\nu = \overline{1,q}$,

$$\mathbf{F}(x,y) = \sum_{\nu=1}^{q} \mathbf{D}^{(\nu)}(x,y)\mu_f^{(\nu)}. \tag{8}$$

Let's agree to call the expression (8) *a decomposition of fuzzy matrix* $\mathbf{F}(x,y)$ *on various fuzziness levels* $\mu_f^{(\nu)}$, $\nu = \overline{1,q}$.

Now let (w,v) be any pair of words with the length $d > 0$ in the alphabets X, Y. Consider matrices "product"

$$\mathbf{F}(w,v) = \prod_{t=1}^{d} \mathbf{F}(x_{s_t}, y_{l_t}). \tag{9}$$

For the matrix $\mathbf{F}(x_{s_t}, y_{l_t})$ the decomposition (8) is trule for any $t = \overline{1,d}$

$$\mathbf{F}(x_{s_t}, y_{l_t}) = \sum_{\nu=1}^{d} \mathbf{D}^{(\nu)}(x_{s_t}, y_{l_t})\mu_f^{(\nu)}, \quad t = \overline{1,d}. \tag{10}$$

For the matrices "product" (9) using the method of mathematical induction according to the words length d the next statement may be proved.

Theorem 1. *Let* $\mathbf{F}(x_{s_t}, y_{s_t}), x_{s_t} \in X$, $y_{s_t} \in Y$, *be fuzzy* $(m \times m)$*-matrices, each of which corresponds to its decomposition of view* (10) *on various fuzziness levels* (6)*, then for any pair of words* (w,v), $w = x_{s_1}x_{s_2}\ldots x_{s_d}$, $v = y_{l_1}y_{l_2}\ldots y_{l_d}$, $d = 1,2,\ldots$, *this fuzzy matrices "product"* (9) *may be represented as its decomposition on various fuzziness levels* (6)

$$\mathbf{F}(w,v) = \sum_{\nu=1}^{d} \mathbf{D}^{(\nu)}(w,v)\mu_f^{(\nu)}, \tag{11}$$

where

$$\mathbf{D}^{(\nu)}(w,v) = \prod_{t=1}^{d} \mathbf{D}^{(\nu)}(x_{s_t}, y_{l_t}). \tag{12}$$

4. Decomposition of fuzzy automata

On the basis of Theorem 1 the truth of the statements dealing with decomposition of fuzzy mappings and fuzzy automata may be proved. As, for the generalized fuzzy automaton (2) $\mathbf{r} \in \mathcal{L}^{1,m}$, $\mathbf{q} \in \mathcal{L}^{m,1}$, and if we consider, that a fuzziness levels set (6) takes into account vectors $\mathbf{r}$ and $\mathbf{q}$ elements also, these vectors are possible to be presented as decompositions

$$\mathbf{r} = \sum_{\nu=1}^{q} \mathbf{d}^{(\nu)} \mu_f^{(\nu)}, \qquad \mathbf{q} = \sum_{\nu=1}^{q} \mathbf{g}^{(\nu)} \mu_f^{(\nu)}, \tag{13}$$

where $\mathbf{d}^{(\nu)}$ is a m-measured row vector, and $\mathbf{g}^{(\nu)}$ is a m-measured column vector, $\nu = \overline{1,q}$, with elements

$$d_i^{(\nu)} = \begin{cases} 1 & \text{when } \ r_i \geqslant \mu_f^{(\nu)}, \\ 0 & \text{when } \ r_i < \mu_f^{(\nu)}, \end{cases} \qquad g_i^{(\nu)} = \begin{cases} 1 & \text{when } \ q_i \geqslant \mu_f^{(\nu)}, \\ 0 & \text{when } \ q_i < \mu_f^{(\nu)}. \end{cases} \tag{14}$$

In this case the following statements turn out to be true according to the expressions (3)-(5) and Theorem 1.

Theorem 2. *A fuzzy mapping Φ_f, induced by generalized fuzzy automaton* (2)*, may be represented as the decomposition*

$$\Phi_f = \sum_{\nu=1}^{q} \Phi_{nd}^{(\nu)} \mu_f^{(\nu)},$$

where $\Phi_{nd}^{(\nu)}$, $\nu = \overline{1,q}$, are non-deterministic mappings $\Phi_{nd}^{(\nu)} : X^d \times Y^d \to [0,1]$, $d = 0, 1, \ldots$, such, that

$$\Phi_{nd}^{(\nu)}(w,v) = \begin{cases} \mathbf{d}^{(\nu)} \mathbf{D}^{(\nu)}(w,v) \mathbf{g}^{(\nu)} & \text{when } \ d > 0, \\ \mathbf{d}^{(\nu)} \mathbf{g}^{(\nu)} & \text{when } \ d = 0. \end{cases} \tag{15}$$

and $\mathbf{d}^{(\nu)}$, $\mathbf{D}^{(\nu)}(w,v)$, $\mathbf{g}^{(\nu)}$ are determined by the expressions (10)-(12), (14).

Theorem 3. *A generalized fuzzy automaton $\mathcal{A}_f$* (2) *may be represented as the decomposition*

$$\mathcal{A}_f = \sum_{\nu=1}^{q} \mathcal{A}_{nd}^{(\nu)} \mu_f^{(\nu)}, \tag{16}$$

where $\mathcal{A}_{nd}^{(\nu)}, \nu = \overline{1,q}$, are generalized non-deterministic automata

$$\mathcal{A}_{nd}^{(\nu)} = \langle X,\ A,\ Y,\ \mathbf{d}^{(\nu)},\ \{\mathbf{D}^{(\nu)}(x,y)\},\ \mathbf{g}^{(\nu)} \rangle, \tag{17}$$

and the "summation" in the expression (16) *means, that vectors $\mathbf{r}, \mathbf{q}$ and matrices $\mathbf{F}(x,y)$ of the automaton $\mathcal{A}_f$ are determined by the formulas* (13), (8).

5. Example

Let the generalized fuzzy automaton $\mathcal{A}_f$ (2), where $X = \{x_1, x_2\}$, $A = \{a_1, a_2, a_3\}$, $Y = \{y_1, y_2\}$, be given (hereinafter $\mathbf{F}(x_s, y_l) = \mathbf{F}(s, l)$)

$$\mathbf{r} = (0,6;\ 0,1;\ 0), \qquad \mathbf{q}^T = (0;\ 0,2;\ 0,7)^T,$$

$$\mathbf{F}(1,1) = \begin{pmatrix} 0,4 & 0 & 0,8 \\ 0 & 0 & 0,6 \\ 0 & 0,8 & 0 \end{pmatrix}, \qquad \mathbf{F}(1,2) = \begin{pmatrix} 0 & 0,4 & 0 \\ 0,1 & 0 & 0,4 \\ 0 & 0,8 & 0 \end{pmatrix},$$

$$\mathbf{F}(2,1) = \begin{pmatrix} 0 & 0 & 0,4 \\ 0,4 & 0 & 0 \\ 0 & 0,3 & 0,8 \end{pmatrix}, \qquad \mathbf{F}(2,2) = \begin{pmatrix} 0,7 & 0 & 0 \\ 0 & 0 & 0,8 \\ 0,2 & 0,4 & 0 \end{pmatrix}. \tag{18}$$

In accordance to (18) get

$$\mu_f = (0,1;\ 0,2;\ 0,3;\ 0,4;\ 0,6;\ 0,7;\ 0,8).$$

Taking into account the expressions (14), (15), (18), for non-deterministic automata $\mathcal{A}_{nd}^{(6)}$ and $\mathcal{A}_{nd}^{(7)}$ we have $\mathbf{d}^{(6)} = \mathbf{d}^{(7)} = (0,0,0)$ and the mappings, induced by them, certainly are zero for all $(w, v) \in X^d \times Y^d, d = 0, 1, 2, \ldots$, and, consequently, the decomposition (16) of the given generalized fuzzy automaton (18) will have the view $\mathcal{A}_f = \sum_{\nu=1}^{5} \mathcal{A}_{nd}^{(\nu)} \mu_f^{(\nu)}$, where non-deterministic automata $\mathcal{A}_{nd}^{(\nu)}, \nu = \overline{1,5}$, are determined by the expressions

$$\mathbf{d}^{(1)} = (1,1,0), \quad \mathbf{d}^{(2)} = \ldots = \mathbf{d}^{(5)} = (1,0,0);$$

$$\mathbf{g}^{(1)} = \mathbf{g}^{(2)} = (0,1,1)^T, \quad \mathbf{g}^{(3)} = \ldots = \mathbf{g}^{(5)} = (0,0,1)^T.$$

$$\mathbf{D}^{(1)}(1,1) = ... = \mathbf{D}^{(4)}(1,1) = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad \mathbf{D}^{(5)}(1,1) = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix},$$

$$\mathbf{D}^{(1)}(1,2) = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad \mathbf{D}^{(2)}(1,2) = ... = \mathbf{D}^{(4)}(1,2) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix},$$

$$\mathbf{D}^{(5)}(1,2) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \qquad \mathbf{D}^{(4)}(2,1) = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

$$\mathbf{D}^{(1)}(2,1) = ... = \mathbf{D}^{(3)}(2,1) = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix},$$

$$\mathbf{D}^{(5)}(2,1) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{D}^{(5)}(2,2) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix},$$

$$\mathbf{D}^{(1)}(2,2) = \mathbf{D}^{(2)}(2,2) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}, \quad \mathbf{D}^{(3)}(2,2) = \mathbf{D}^{(4)}(2,2) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

6. Conclusion

So, it has been proved, that any generalized fuzzy finite automaton $\mathcal{A}_f$ may be represented as a union ("sum") of generalized non-deterministic finite automata $\mathcal{A}_{nd}^{(\nu)}$, $\nu = \overline{1,q}$ corresponding to various fuzziness levels $\mu_f^{(\nu)}$. This result, in particular, has a great importance for the future development of special methods abstract analysis and synthesis fuzzy automata models solving, by its reduction to the analysis and the synthesis non-deterministic automata models for various levels of fuzziness.

References

[1] Kaufmann A. (1982) *Introduction to the Theory of Fuzzy Subsets.* Moscow, 432. (in Russian translation from French).

[2] Tchirkov M.K, Ponomareva A.Yu. (2008) *Stationary Deterministic and Stochastic Automata (Theory of Automata Models).* St. Petersburg, 248. (in Russian)

[3] A. Kandel, S.C. Lee. (1979) *Fuzzy Switching and Automata: Theory and Applications.* New York, Crane Russak & Comp. Inc., 303.

[4] Skorikova Ya.I., Tchirkov M.K. (2005) *Abstract analysis of generalized fuzzy automata.* Mathematical Models. Theory and Applications, vol. 6, St. Petersburg, 110-122. (in Russian)

[5] Ponomareva A.Yu., Sandrykina N.V., Tchirkov M.K. (2003) *On minimal forms of generalized nondeterministic automata.* Mathematical Models. Theory and Applications, vol. 3, St. Petersburg, 94-102. (in Russian)

6th St.Petersburg Workshop on Simulation (2009) 869-873

On simulation of stochastic languages by nondeterministic automata[1]

A.S. Shevchenko[2], M.K. Tchirkov[2]

Abstract

The possibilities of optimal simulation stochastic languages represented by finite stochastic automata are considered by the instrumentality of nondeterministic automata with additional stochastic input. The problems of analysis and synthesis such kind of automata are solved.

1. Introduction

It is known that a method of synthesis any stochastic automaton as a deterministic automaton with a special stochastic input exists in the theory of stochastic automata [1]-[3]. This representation of stochastic automaton can be very difficult in case of big quantity of states. Therefore a searching of more optimal methods of simulation of stochastic automata and represented they languages is on the agenda. Just this problem in regard to so-called nondeterministic automata with stochastic input is considered in this article.

2. Basic definitions

A stochastic finite automaton $\mathcal{A}_{pr}$ is a system

$$\mathcal{A}_{pr} = \langle X,\, A,\, Y,\, \mathbf{p}, \{\mathbf{P}(x_s, y_l)\}\rangle, \tag{1}$$

where X is an input alphabet, $|X| = n$, A is an alphabet of states, $|A| = m$, Y is an output alphabet, $|Y| = k$, $\mathbf{p} = (p_0,\, p_1, \ldots, p_{m-1})$ is start distribution of states probabilities and $\{\mathbf{P}(x_s, y_l)\}$ is a system $(m \times m)$-matrices of transitions probabilities $\mathbf{P}(x_s, y_l) = (P_{ij}(x_s, y_l))_{m,m}$, $P_{ij}(x_s, y_l) = P(a_j y_l | a_i x_s)$, where $a_i, a_j \in A$, $x_s \in \in X$, $y_l \in Y$. If an output alphabet is an alphabet of states then this automaton is called *an abstract stochastic finite automaton* $\mathcal{A}_{pr} = \langle X,\, A,\, \mathbf{p},\, \{\mathbf{P}(x_s)\}\rangle$.

A deterministic automaton with stochastic input $\mathcal{A}_{inp}$ is deterministic finite automaton, which has two input channels, and symbols come to one of them from sampling machine with specified probabilities. So it is a system

$$\mathcal{A}_{inp} = \langle Z \times X,\, A,\, Y,\, a_0,\, \Psi,\, \mu\rangle, \tag{2}$$

[1]This work was supported by RFBR grant 07-01-00355.

[2]Petersburg University, E-mail: annion@yandex.ru

where $Z = \{z_0, z_1, \ldots, z_{q-1}\}$, $\Psi\colon A \times Z \times X \to A \times Y$, and $\mu = (\mu_0, \mu_1, \ldots, \mu_{q-1})$ is the distribution of probabilities of symbols from Z effect to automaton. If an output alphabet is an alphabet of states then this automaton is called *an abstract deterministic finite automaton with stochastic input* $\mathcal{A}_{inp} = \langle Z \times X, A, a_0, \Psi, \mu\rangle$.

Let $\mathcal{R}$ be the Boolean lattice $\mathcal{R} = \langle\{0,1\}, \vee, \&, \geq\rangle$ and $\mathcal{R}^{m,m}$ designates the set of all $(m \times m)$-matrices above $\mathcal{R}$.

A generalized nondeterministic finite automaton $\mathcal{A}_{nd}$ is a system

$$\mathcal{A}_{nd} = \langle X, A, Y, \mathbf{r}, \{\mathbf{D}(x_s, y_l)\}, \mathbf{q}\rangle,$$

where $\mathbf{r} \in \mathcal{R}^{1,m}$ – row vector of start states, $\mathbf{q} \in \mathcal{R}^{m,1}$ – column vector of finite states and $\{\mathbf{D}(x_s, y_l)\} = \{\mathbf{D}(x_s, y_l) |\ \mathbf{D}(x_s, y_l) \in \mathcal{R}^{m,m}, x_s \in X, y_l \in Y\}$ – family of nk $(m \times m)$ – matrices of transitions and outputs. If an output alphabet is an alphabet of states then this automaton is called *an abstract nondeterministic finite automaton*$\mathcal{A}_{nd} = \langle X, A, \mathbf{r}, \{\mathbf{D}(x_s)\}, \mathbf{q}\rangle$.

A generalized nondeterministic mapping induced by automaton $\mathcal{A}_{nd}$ is a mapping $X^* \times Y^*$ (where X^*, Y^* – the sets of all words any size in X, Y) into $\{0,1\}$ defined by expression (for $w = x_{s_1}x_{s_2}\ldots x_{s_t}$, $v = y_{l_1}y_{l_2}\ldots y_{l_t}$)

$$\Phi_{nd}(w,v) = \begin{cases} \mathbf{r}\prod_{\nu=1}^{t}\mathbf{D}(x_{s_\nu}, y_{l_\nu})\mathbf{q}, & |w| = |v| = t, \\ 0, & |w| \neq |v|, \end{cases}$$

where the operations $\vee$ and & are meant under "addition" and "multiplication".

A generalized nondeterministic finite automaton with stochastic input is a system

$$\mathcal{A}_{snd} = \langle Z \times X, A, Y, \mathbf{r}, \{\mathbf{D}(z_g, x_s, y_l)\}, \mathbf{q}, \mu\rangle, \tag{3}$$

which presents nondeterministic automaton, which has two input channels, and symbols come to one of them from sampling machine with specified probabilities, where

$$\{\mathbf{D}(z_g, x_s, y_l)\} =$$
$$= \{\mathbf{D}(z_g, x_s, y_l) | \mathbf{D}(z_g, x_s, y_l) \in \mathcal{R}^{m,m}, z_g \in Z, x_s \in X, y_l \in Y\}$$

is a family of qnk transitions and outputs matrices and μ is defined similarly (2). If an output alphabet is an alphabet of states then this automaton is called *an abstract nondeterministic finite automaton with stochastic input* $\mathcal{A}_{snd} = \langle Z \times \times X, A, \mathbf{r}, \{\mathbf{D}(z_g, x_s)\}, \mathbf{q}, \mu\rangle$.

A language $\mathcal{Z}$ in the alphabet X is any subset of words $\mathcal{Z} \subseteq X^*$ in this alphabet, which is represented by characteristic function $\chi_{\mathcal{Z}}(w) \in \{0, 1\}$ for each $w \in X^*$ and therefore $w \in \mathcal{Z} \subseteq X^* \Leftrightarrow \chi_{\mathcal{Z}}(w) = 1$.

A "fuzzy" language $\mathcal{Z}$ in alphabet X is a "fuzzy" set of words from X^*, which is specified by generalized characteristic function (*membership function*) $\chi_{\mathcal{Z}} : X^* \to \to [0, 1]$. Let $\mathcal{A}_{pr} = \langle X, A, Y, \mathbf{p}, \{\mathbf{P}(x_s, y_l)\}\rangle$ is stochastic finite automaton, $Y^{(K)} \subseteq Y$ is dedicated set of its output symbols and $P(y_l|w)$ is probability of outputting the letter y_l in tact t by inputting the word $w = x_{s_1}x_{s_2}\ldots x_{s_t}$. Then *the stochastic language* $\mathcal{Z}$, represented in $\mathcal{A}_{pr}$ by subset of output symbols $Y^{(K)}$, is the "fuzzy" set of words from X, when

$$\mathcal{Z} = \{w, \chi(w)\}_{w \in X^*}, \qquad \chi(w) = \sum_{y_l \in Y^{(K)}} P(y_l|w).$$

If the automaton $\mathcal{A}_{pr} = \langle X, A, \mathbf{p}, \{\mathbf{P}(x_s)\}\rangle$ is stochastic abstract automaton with dedicated subset of finite states $A^{(K)} \subseteq A$ and $P(a_i|w)$ is the probability of staying $\mathcal{A}_{pr}$ in tact t in state a_i by inputting the word w then *the stochastic language* $\mathcal{Z}$, represented in $\mathcal{A}_{pr}$ by subset of finite states $A^{(K)}$ is the "fuzzy" set of words $\mathcal{Z} = \{w, \chi(w)\}_{w\in X^*}$ in X that $\chi(w) = \sum_{a_i\in A^{(K)}} P(a_i|w)$, for each $w \in X^*$.

Let the generalized nondeterministic automaton $\mathcal{A}_{nd}$ is given and $Y^{(K)} \subseteq Y$ is dedicated subset of its output symbols. Then the language $\mathcal{Z}$ specified by characteristic function $\Phi_{\mathcal{Z}} : X^* \to \{0, 1\}$, is represented in $\mathcal{A}_{nd}$ by subset $Y^{(K)}$ when

$$\bigvee_{v\in Y^{t-1}} \bigvee_{y_l\in Y^{(K)}} \Phi_{nd}(w, vy_l) = \Phi_{\mathcal{Z}}(w),$$

for each $w \in X^*$, $|w| = |vy_l| = t$, $t = 0, 1, \ldots$

Let $Y^{(K)}$ is dedicated subset of output symbols in the generalized nondeterministic automaton with stochastic input $\mathcal{A}_{snd}$ (3).Then the "fuzzy" language $\mathcal{Z}$ specified by characteristic function $\Phi_{\mathcal{Z}} : X^* \to [0, 1]$, is represented in $\mathcal{A}_{snd}$ by subset $Y^{(K)}$ if

$$\Phi_{\mathcal{Z}}(w) = \sum_{y_l\in Y^{(K)}} P(y_l|w),$$

for each $w \in X^*$, $|w| = t$, $t = 0, 1, \ldots.$

If it is considering the abstract nondeterministic automaton $\mathcal{A}_{snd} = \langle Z \times \times X, A, \mathbf{r}, \{\mathbf{D}(z_g, x_s), \mathbf{q}, \mu\}\rangle$ then the "fuzzy" language $\mathcal{Z}$ represented in automaton is specified by characteristic function $\Phi_{\mathcal{Z}} : X^* \to [0, 1]$, where

$$\Phi_{\mathcal{Z}}(w) = \sum_i P(a_i|w)q_i$$

for each $w \in X^*$, $|w| = y$, $t = 0, 1, \ldots.$

Let there be given two automata – a stochastic finite automaton $\mathcal{A}_{pr}$ (1) and a generalized nondeterministic finite automaton with stochastic input $\mathcal{A}_{snd}$ (3), which represent languages $\mathcal{Z}_1$ and $\mathcal{Z}_2$ respectively. These automata are called equivalent if $\mathcal{Z}_1 = \mathcal{Z}_2$, i.e. they represent the same language.

3. The problem of analysis

At first it is necessary to show that the nondeterministic finite automaton $\mathcal{A}_{snd}$ (3) is equivalent by represented language to a stochastic finite automaton $\mathcal{A}_{pr}$ (1). Two next statements may be proved.

Theorem 1. *For each generalized nondeterministic finite automaton with stochastic input $\mathcal{A}_{snd}$, which has $|A| = m$, $|X| = n$, $|Y| = k$, and represents the "fuzzy" language $\mathcal{Z}$, may be construct an abstract finite stochastic automaton $\mathcal{A}_{pr}$ which has $\tilde{m} \leq 2^{mk}$ states and is equivalent initial automaton by represented language.*

Theorem 2. *The language may be represented by the generalized nondeterministic automaton if and only if this language is stochastic language.*

4. The problem of synthesis

Then the inverse problem is considering. Let the abstract stochastic finite automaton $\mathcal{A}_{pr}$ is given, it is necessary to construct the equivalent by represented stochastic language abstract nondeterministic automaton with stochastic input having minimal number of states. Besides it is necessary to characterize the range of values number of states received automaton.

The following statement is true:

Theorem 3. *For each abstract stochastic finite automaton*

$$\mathcal{A}_{pr} = \langle X, A, a_0, \{\boldsymbol{P}(x_s)\}, \mathbf{e}^{(K)}\rangle,$$

which has M states and represents stochastic language $\mathcal{Z}$, may be construct abstract nondeterministic finite automaton with stochastic input

$$\mathcal{A}_{snd} = \langle Z \times X,\, B,\, \tilde{\boldsymbol{r}},\, \{\tilde{\boldsymbol{D}}(z,\, x)\},\, \hat{\boldsymbol{q}},\, \mu\rangle,$$

equivalent initial automaton and it has m states at most, where $m = \lceil \log_2 M \rceil + \delta$, $\delta \in [0, M - \lceil \log_2 M \rceil]$, and $\lceil \log_2 M \rceil$ is nearest integer number not less or equal $\log_2 M$.

Algorithm of task solution involves next phases:

- synthesis of the abstract deterministic automaton with stochastic input equivalent to given abstract stochastic automaton [1]-[3];
- replacement of the pairs of input symbols by one symbol;
- optimization of the abstract deterministic automaton under the condition of conservation of presented regular language [3];
- synthesis of the abstract nondeterministic automaton equivalent to given abstract deterministic automaton, for it at first $m = \lceil \log_2(M) \rceil$ may be selected, then it is necessary to match a coding of states that with assignable elements of matrices $\mathbf{D}(x_s),\ s = \overline{1, n}$, will obey:

  $$(a_1^i a_2^i \dots a_m^i)\mathbf{D}(x_s) = (a_1^j a_2^j \dots a_m^j)$$

  i.e.

  $$\bigvee_{\nu} a_\nu^i d_{\nu\eta}^s = a_\eta^j,$$

  where $\nu, \eta = \overline{1, m},\ i, j = \overline{1, M},\ \mathbf{D}(x_s) = (d_{ij}^s)_{m,m},\ d_{ij}^s \in \{0, 1\},\ c_i \longmapsto \longmapsto (a_1^i a_2^i \dots a_m^i)$, $c_j = f(a_i, x_s) \longmapsto (a_1^j a_2^j \dots a_m^j)$, and if this conditions are incompatibility, than $m = m+1$ and the process of matching of states coding repeats;
- conversion of the nondeterministic abstract automaton into the nondeterministic abstract automaton with stochastic input by the instrumentality of inverse interchange of symbols into the pairs of symbols.

There are two simple examples to show extreme values of δ. Let the automaton

$$\mathcal{A}_{pr} = \langle X,\, A,\, \mathbf{p}, \{\mathbf{P}(s),\, \mathbf{e}^{(K)}\}\rangle$$

be given, where $X = \{x_0,\, x_1\}$, $A = \{a_0,\, a_1, \ldots,\, a_7\}$, $\mathbf{p} = (0\,1\,0\ldots0)$, $\mathbf{e}^{(K)} = = (0\,1\;0\,1\,0\,1\,0\,1)^T$, and matrices $\mathbf{P}(x_s)$ have the forms:

$$\mathbf{P}(x_0) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0,8 & 0 & 0 & 0 & 0 & 0 & 0,2 & 0 \\ 0 & 0,2 & 0 & 0 & 0,8 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0,8 & 0 & 0 & 0,2 \\ 0 & 0,8 & 0,2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0,8 & 0 & 0 & 0 & 0 & 0,2 & 0 \\ 0 & 0 & 0 & 0,2 & 0 & 0,8 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0,8 & 0 & 0,2 \end{pmatrix},$$

$$\mathbf{P}(x_1) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0,8 & 0 & 0,2 & 0 & 0 & 0 \\ 0 & 0 & 0,2 & 0 & 0,8 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0,8 & 0 & 0 & 0 & 0,2 & 0 & 0 & 0 \\ 0 & 0 & 0,8 & 0 & 0,2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0,8 & 0 & 0,2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

It is necessary to construct the abstract nondeterministic automata with stochastic input with minimal number of states. It is synthesized the automaton $\mathcal{A}_{snd}$ using the algorithm described above:

$$\mathcal{A}_{snd} = \langle Z \times X,\, B,\, \mathbf{r},\, \{\mathbf{D}(z_g,\, x_s)\},\, \mathbf{q},\, \mu\rangle,$$

where $Z = \{z_0, z_1\}$, $X = \{x_0, x_1\}$, $B = \{b_0, b_1, b_2\}$, $\mathbf{r} = \mathbf{q}^T = (0\,0\,1)$, $\mu = = (0,2;\, 0,8)$, and matrices $\{\mathbf{D}(z_g,\, x_s)\}$ have the forms:

$$\mathbf{D}(z_0, x_0) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}, \qquad \mathbf{D}(z_1, x_0) = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$\mathbf{D}(z_0, x_1) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \qquad \mathbf{D}(z_1, x_1) = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

Thereby this example shows us the possibility of minimal value $\delta = 0$.

Let the automaton $\mathcal{A}_{pr}$ be given, where $X = \{x_0,\, x_1\}$, $A = \{a_0,\, a_1, \ldots,\, a_4\}$, $\mathbf{p} = (1\,0\,0\,0\,0)$, $\mathbf{e}^{(K)} = (0\,0\,0\,0\,1)^T$, and matrices $\mathbf{P}(x_s)$ have the forms:

$$\mathbf{P}(x_0) = \begin{pmatrix} 0 & 0,2 & 0,8 & 0 & 0 \\ 0 & 0 & 0,2 & 0,8 & 0 \\ 0 & 0 & 0 & 0,2 & 0,8 \\ 0,8 & 0 & 0 & 0 & 0,2 \\ 0,2 & 0,8 & 0 & 0 & 0 \end{pmatrix},$$

$$\mathbf{P}(x_1) = \begin{pmatrix} 0 & 0 & 0 & 0,2 & 0,8 \\ 0,8 & 0 & 0 & 0 & 0,2 \\ 0,2 & 0,8 & 0 & 0 & 0 \\ 0 & 0,2 & 0,8 & 0 & 0 \\ 0 & 0 & 0,2 & 0,8 & 0 \end{pmatrix}.$$

It is necessary to construct the abstract nondeterministic automata with stochastic input with minimal number of states.

The number of states can not be minimized in this case. It is synthesised the automaton $\mathcal{A}_{snd} = \langle Z \times X,\, A,\, \mathbf{r},\, \{\mathbf{D}(z_g,\, x_s)\},\, \mathbf{q},\, \mu\rangle$, where $Z = \{z_0, z_1\}$, $X = \{x_0, x_1\}$, $B = \{b_0, b_1, \ldots b_4\}$, $\mathbf{r} = \mathbf{p}$, $\mathbf{q} = \mathbf{e}^{(K)}$, $\mu = (0,2;\, 0,8)$, and matrices $\{\mathbf{D}(z_g,\, x_s)\}$ have the forms:

$$\mathbf{D}(z_0, x_0) = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad \mathbf{D}(z_1, x_0) = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{pmatrix},$$

$$\mathbf{D}(z_0, x_1) = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix}, \quad \mathbf{D}(z_1, x_1) = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

Thereby this example shows the possibility of maximal value $\delta = M - \lceil \log_2 M \rceil$.

References

[1] Buharaev R.G. (1985) *Foundation of Stochastic Automata Theory.* Moscow, 288 (in Russian).

[2] Chentsov V.M. (1985) *Synthesis of Stochastic Automaton.*// Problems of Digital Automata Synthesis. Moscow, 135-144 (in Russian).

[3] Tchirkov M.K., Ponomareva A.Yu. (2008) *Stationary Deterministic and Stochastic Automata (Theory of Automata Models).* St.Petersburg, 248 (in Russian).

6th St.Petersburg Workshop on Simulation (2009) 875-879

Evaluation of the Lyapunov exponent for generalized linear second-order exponential systems[1]

Nikolai Krivulin[2]

Abstract

We consider generalized linear stochastic dynamical systems with second-order state transition matrices. The entries of the matrix are assumed to be either independent and exponentially distributed or equal to zero. We give an overview of new results on evaluation of asymptotic growth rate of the system state vector, which is called the Lyapunov exponent of the system.

1. Introduction

The evolution of actual systems that occur in management, engineering, computer sciences, and other areas can frequently be represented through stochastic dynamic equations of the form

$$\mathbf{m}z(k) = A(k)\mathbf{m}z(k-1),$$

where $A(k)$ is a random state transition matrix, $\mathbf{m}z(k)$ is a system state vector, and matrix-vector multiplication is thought of as defined in terms of a semiring with the operations of taking maximum and addition [1, 2, 3].

In many cases, the analysis of a system involves evaluation of asymptotic growth rate of the system state vector $\mathbf{m}z(k)$, which is normally referred to as the Lyapunov exponent [4, 5].

Evaluation of the Lyapunov exponent typically appears to be a difficult problem even for quite simple systems. Related results include the solutions obtained in [6, 5] for systems with matrices of the second order with independent and exponentially distributed entries. In [6], the Lyapunov exponent is obtained in the case that all entries of the matrix are identically distributed with unit mean.

Further results are given in [5] under the condition that the diagonal entries have one common distribution, whereas the off-diagonal entries do follow another distribution. A system with a matrix such that its diagonal entries are distributed with unit mean, and the off-diagonal entries are equal to zero is also examined.

[1]The work was partially supported by the Russian Foundation for Basic Research under Grant #09-01-00808.

[2]St.Petersburg State University, E-mail: Nikolai.Krivulin@pobox.spbu.ru

The purpose of this paper is to give an overview of new results which are related to evaluation of the Lyapunov exponent in generalized linear systems that have matrices of the second order with exponentially distributed entries (second-order exponential systems).

2. Stochastic Linear Dynamical System

Consider a dynamical system that can be represented through the linear equation in the semiring with the operations of maximum and addition

$$\mathbf{m}z(k) = A(k)\mathbf{m}z(k-1),$$

where

$$A(k) = \begin{pmatrix} \alpha_k & \beta_k \\ \gamma_k & \delta_k \end{pmatrix}, \qquad \mathbf{m}z(k) = \begin{pmatrix} x(k) \\ y(k) \end{pmatrix}, \qquad \mathbf{m}z(0) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

With ordinary notation, the above vector equation can be written as

$$\begin{aligned} x(k) &= \max(x(k-1) + \alpha_k, y(k-1) + \beta_k), \\ y(k) &= \max(x(k-1) + \gamma_k, y(k-1) + \delta_k). \end{aligned}$$

The Lyapunov exponent for the system is given by

$$\lambda = \lim_{k\to\infty} \frac{1}{k} \max(x(k), y(k)).$$

Suppose the sequences $\{\alpha_k\}$, $\{\beta_k\}$, $\{\gamma_k\}$, and $\{\delta_k\}$ each involve independent and identically distributed random variables; α_k, β_l, γ_m, and δ_n are independent for any k, l, m, n. Finally, we assume that α_k, β_k, γ_k, and δ_k have the exponential probability distributions with respective parameters μ, ν, σ, and τ.

3. Systems With Matrices Having Zero Entries

We start with results obtained in [7, 8] for systems with state transition matrices having one or two nonrandom entries that are equal to zero. The reduced number of random entries in the matrices allows one to simplify the evaluation of the Lyapunov exponent and usually gives quite compact results.

The solution method is based on construction of a sequence of probability distribution functions. The convergence of the sequence is examined and the limiting distribution is derived as the solution of an integral equation. The Lyapunov exponent is then evaluated as the expected value of a random variable determined through the limiting distribution function.

3.1. Matrix With Zero Off-Diagonal Entries

The system state transition matrix together with its related result take the form

$$A(k) = \begin{pmatrix} \alpha_k & 0 \\ 0 & \delta_k \end{pmatrix}, \qquad \lambda = \frac{\mu^4 + \mu^3\tau + \mu^2\tau^2 + \mu\tau^3 + \tau^4}{\mu\tau(\mu+\tau)(\mu^2+\tau^2)}.$$

With $\tau = \mu = 1$ we have the result $\lambda = 1.25$ which coincides with that in [5].

3.2. Matrix With Zero Diagonal

In this case, the matrix and the Lyapunov exponent are represented as

$$A(k) = \begin{pmatrix} 0 & \beta_k \\ \gamma_k & 0 \end{pmatrix}, \qquad \lambda = \frac{4\nu^2 + 7\nu\sigma + 4\sigma^2}{6\nu\sigma(\nu+\sigma)}.$$

3.3. Matrix With Zero Row or Column

Provided that the second row in the matrix has only zero entries, we arrive at

$$A(k) = \begin{pmatrix} \alpha_k & \beta_k \\ 0 & 0 \end{pmatrix}, \qquad \lambda = \frac{2\mu^4 + 7\mu^3\nu + 10\mu^2\nu^2 + 11\mu\nu^3 + 4\nu^4}{\mu\nu(\mu+\nu)^2(3\mu+4\nu)}.$$

When the entries of the second column are zero, we have

$$A(k) = \begin{pmatrix} \alpha_k & 0 \\ \gamma_k & 0 \end{pmatrix}, \qquad \lambda = \frac{2\mu^4 + 7\mu^3\sigma + 10\mu^2\sigma^2 + 11\mu\sigma^3 + 4\sigma^4}{\mu\sigma(\mu+\sigma)^2(3\mu+4\sigma)}.$$

3.4. Matrix With Zero Entry on Diagonal

Consider a system with the state transition matrix

$$A(k) = \begin{pmatrix} \alpha_k & \beta_k \\ \gamma_k & 0 \end{pmatrix}.$$

Whereas evaluation of the Lyapunov exponent for this system in the general case leads to rather cumbersome algebraic manipulations, there are two main particular cases which offer their related results in a relatively compact form. Under the condition that $\sigma = \mu$, we have

$$\lambda = \frac{48\mu^5 + 238\mu^4\nu + 495\mu^3\nu^2 + 581\mu^2\nu^3 + 326\mu\nu^4 + 68\nu^5}{2\mu\nu(36\mu^4 + 147\mu^3\nu + 215\mu^2\nu^2 + 130\mu\nu^3 + 28\nu^4)}.$$

Provided that $\sigma = \nu$, the value of the Lyapunov exponent is given by

$$\lambda = P(\mu,\nu)/Q(\mu,\nu),$$

where

$$\begin{aligned} P(\mu,\nu) &= 15\mu^8 + 152\mu^7\nu + 624\mu^6\nu^2 + 1382\mu^5\nu^3 + 1838\mu^4\nu^4 + 1592\mu^3\nu^5 + \\ &\quad + 973\mu^2\nu^6 + 384\mu\nu^7 + 64\nu^8, \\ Q(\mu,\nu) &= \mu\nu(\mu+\nu)^2(12\mu^5 + 97\mu^4\nu + 286\mu^3\nu^2 + 397\mu^2\nu^3 + 256\mu\nu^4 + 64\nu^5). \end{aligned}$$

3.5. Matrix With Zero Entry Below Diagonal

Suppose that there is a system with the state transition matrix defined as

$$A(k) = \begin{pmatrix} \alpha_k & \beta_k \\ 0 & \delta_k \end{pmatrix}.$$

Consider two particular cases. With the condition $\nu = \mu$, we have the result

$$\lambda = P(\mu, \tau)/Q(\mu, \tau),$$

where

$$\begin{aligned} P(\mu,\tau) &= 288\mu^8 + 1048\mu^7\tau + 1936\mu^6\tau^2 + 2688\mu^5\tau^3 + 3012\mu^4\tau^4 + \\ &\quad + 2226\mu^3\tau^5 + 941\mu^2\tau^6 + 204\mu\tau^7 + 17\tau^8, \\ Q(\mu,\tau) &= 2\mu\tau(144\mu^7 + 524\mu^6\tau + 968\mu^5\tau^2 + 1200\mu^4\tau^3 + 910\mu^3\tau^4 + \\ &\quad + 387\mu^2\tau^5 + 84\mu\tau^6 + 7\tau^7). \end{aligned}$$

Provided that $\tau = \mu$, the solution takes the form

$$\lambda = P(\mu, \nu)/Q(\mu, \nu),$$

where

$$\begin{aligned} P(\mu,\nu) &= 256\mu^{10} + 2112\mu^9\nu + 8044\mu^8\nu^2 + 19355\mu^7\nu^3 + 32167\mu^6\nu^4 + \\ &\quad + 36887\mu^5\nu^5 + 28709\mu^4\nu^6 + 14854\mu^3\nu^7 + 4912\mu^2\nu^8 + 944\mu\nu^9 + 80\nu^{10}, \\ Q(\mu,\nu) &= 2\mu\nu(\mu+\nu)(192\mu^8 + 1344\mu^7\nu + 4047\mu^6\nu^2 + 6770\mu^5\nu^3 + \\ &\quad + 6799\mu^4\nu^4 + 4216\mu^3\nu^5 + 1600\mu^2\nu^6 + 344\mu\nu^7 + 32\nu^8). \end{aligned}$$

4. General Second-Order Exponential System

Consider a general second-order exponential system which has the matrix

$$A(k) = \begin{pmatrix} \alpha_k & \beta_k \\ \gamma_k & \delta_k \end{pmatrix}$$

with its entries α_k, β_k, γ_k, and δ_k assumed to be independent random variables that are exponentially distributed with the respective parameters μ, ν, σ, and τ.

To get the Lyapunov exponent, a computational technique developed in [9] can be implemented which reduces the problem to the solution of a system of linear equations, accompanied by the evaluation of a linear functional of the system solution. The technique leans upon construction and examination of a sequence of probability density functions. It is shown that there is a one-to-one correspondence between the density functions and vectors in a vector space. The correspondence is then exploited to provide for the solution in terms of algebraic computations.

Based on the above technique, the Lyapunov exponent can be evaluated as follows. First we introduce the vectors

$$\mathbf{m}\omega_1 = (\omega_{10}, \omega_{11}, \omega_{12}, \omega_{13})^T,\ \mathbf{m}\omega_2 = (\omega_{20}, \omega_{21}, \omega_{22}, \omega_{23})^T,\ \mathbf{m}\omega = (\mathbf{m}\omega_1^T, \mathbf{m}\omega_2^T)^T.$$

Furthermore, we define the matrices

$$U_1 = \begin{pmatrix} 1 & 1 & 1 \\ \frac{\mu}{\mu+\nu} & \frac{1}{2} & \frac{\mu+\nu}{\mu+2\nu} \\ \frac{\mu}{\mu+\tau} & \frac{\nu}{\nu+\tau} & \frac{\mu+\nu}{\mu+\nu+\tau} \\ \frac{\mu}{\mu+\nu+\tau} & \frac{\nu}{2\nu+\tau} & \frac{\mu+\nu}{\mu+2\nu+\tau} \end{pmatrix}, \qquad U_2 = \begin{pmatrix} 1 & 1 & 1 \\ \frac{\sigma}{\mu+\sigma} & \frac{\tau}{\mu+\tau} & \frac{\sigma+\tau}{\mu+\sigma+\tau} \\ \frac{1}{2} & \frac{\tau}{\sigma+\tau} & \frac{\sigma+\tau}{2\sigma+\tau} \\ \frac{\sigma}{\mu+2\sigma} & \frac{\tau}{\mu+\sigma+\tau} & \frac{\sigma+\tau}{\mu+2\sigma+\tau} \end{pmatrix},$$

$$V_{11} = \begin{pmatrix} \frac{\sigma}{\mu+\sigma} & 0 & -\frac{\mu\sigma}{(\mu+\tau)(\mu+\sigma+\tau)} & 0 \\ 0 & \frac{\sigma}{\nu+\sigma} & 0 & -\frac{\nu\sigma}{(\nu+\tau)(\nu+\sigma+\tau)} \\ 0 & -\frac{\sigma}{\mu+\nu+\sigma} & 0 & \frac{\sigma(\mu+\nu)}{(\mu+\nu+\tau)(\mu+\nu+\sigma+\tau)} \end{pmatrix},$$

$$V_{12} = \begin{pmatrix} 0 & \frac{\tau}{\mu+\tau} & 0 & -\frac{\mu\tau}{(\mu+\sigma)(\mu+\sigma+\tau)} \\ \frac{\tau}{\nu+\tau} & 0 & -\frac{\nu\tau}{(\nu+\sigma)(\nu+\sigma+\tau)} & 0 \\ 0 & -\frac{\tau}{\mu+\nu+\tau} & 0 & \frac{\tau(\mu+\nu)}{(\mu+\nu+\sigma)(\mu+\nu+\sigma+\tau)} \end{pmatrix},$$

$$V_{21} = \begin{pmatrix} \frac{\mu}{\mu+\sigma} & -\frac{\mu\sigma}{(\nu+\sigma)(\mu+\nu+\sigma)} & 0 & 0 \\ 0 & 0 & \frac{\mu}{\mu+\tau} & -\frac{\mu\tau}{(\nu+\tau)(\mu+\nu+\tau)} \\ 0 & 0 & -\frac{\mu}{\mu+\sigma+\tau} & \frac{\mu(\sigma+\tau)}{(\nu+\sigma+\tau)(\mu+\nu+\sigma+\tau)} \end{pmatrix},$$

$$V_{22} = \begin{pmatrix} 0 & 0 & \frac{\nu}{\nu+\sigma} & -\frac{\nu\sigma}{(\mu+\sigma)(\mu+\nu+\sigma)} \\ \frac{\nu}{\nu+\tau} & -\frac{\nu\tau}{(\mu+\tau)(\mu+\nu+\tau)} & 0 & 0 \\ 0 & 0 & -\frac{\nu}{\nu+\sigma+\tau} & \frac{\nu(\sigma+\tau)}{(\mu+\sigma+\tau)(\mu+\nu+\sigma+\tau)} \end{pmatrix}.$$

Suppose that the vector $\mathbf{m}\omega$ is the solution of the system

$$(I - W)\mathbf{m}\omega = \mathbf{m}0,$$
$$\omega_{10} + \omega_{20} = 1,$$

where I represents identity matrix, and

$$W = \begin{pmatrix} U_1V_{11} & U_1V_{12} \\ U_2V_{21} & U_2V_{22} \end{pmatrix}.$$

The value of the Lyapunov exponent is then given by

$$\lambda = \mathbf{m}q_1^T\,\mathbf{m}\omega_1 + \mathbf{m}q_2^T\,\mathbf{m}\omega_2,$$

where $\mathbf{m}q_1$ and $\mathbf{m}q_2$ are vectors such that

$$\mathbf{m}q_1 = \begin{pmatrix} \dfrac{\mu^2+\mu\sigma+\sigma^2}{\mu\sigma(\mu+\sigma)} \\ \dfrac{\mu\sigma(\mu+2\nu+\sigma)}{\nu(\mu+\nu)(\nu+\sigma)(\mu+\nu+\sigma)} \\ \dfrac{\mu\sigma(\mu+\sigma+2\tau)}{\tau(\mu+\tau)(\sigma+\tau)(\mu+\sigma+\tau)} \\ -\dfrac{\mu\sigma(\mu+2\nu+2\tau+\sigma)}{(\nu+\tau)(\mu+\nu+\tau)(\nu+\sigma+\tau)(\mu+\nu+\sigma+\tau)} \end{pmatrix},$$

$$\mathbf{m}q_2 = \begin{pmatrix} \dfrac{\nu^2+\nu\tau+\tau^2}{\nu\tau(\nu+\tau)} \\ \dfrac{\nu\tau(2\mu+\nu+\tau)}{\mu(\mu+\nu)(\mu+\tau)(\mu+\nu+\tau)} \\ \dfrac{\nu\tau(\nu+2\sigma+\tau)}{\sigma(\nu+\sigma)(\sigma+\tau)(\nu+\sigma+\tau)} \\ -\dfrac{\nu\tau(2\mu+\nu+2\sigma+\tau)}{(\mu+\sigma)(\mu+\nu+\sigma)(\mu+\sigma+\tau)(\mu+\nu+\sigma+\tau)} \end{pmatrix}.$$

Suppose that $\tau = \mu$ and $\sigma = \nu$. Implementation of the above technique gives the solution

$$\lambda = P(\mu,\nu)/Q(\mu,\nu),$$

where

$$\begin{aligned} P(\mu,\nu) &= 160\mu^{10} + 1776\mu^9\nu + 8220\mu^8\nu^2 + 21378\mu^7\nu^3 + 35595\mu^6\nu^4 + \\ &+ 41566\mu^5\nu^5 + 35595\mu^4\nu^6 + 21378\mu^3\nu^7 + 8220\mu^2\nu^8 + \\ &+ 1776\mu\nu^9 + 160\nu^{10}, \\ Q(\mu,\nu) &= 16\mu\nu(\mu+\nu)(8\mu^8 + 80\mu^7\nu + 321\mu^6\nu^2 + 690\mu^5\nu^3 + 880\mu^4\nu^4 + \\ &+ 690\mu^3\nu^5 + 321\mu^2\nu^6 + 80\mu\nu^7 + 8\nu^8). \end{aligned}$$

Note that the obtained solution coincides with that in [5].

The author is grateful to the anonymous reviewer for valuable comments and suggestions.

References

[1] Kolokoltsov V.N., and Maslov V.P. (1997) *Idempotent Analysis and Its Applications*. Kluwer Academic Publishers, Dordrecht. (Mathematics and Its Applications, Vol. 401)

[2] Litvinov G.L., Maslov V.P., and Sobolevskii A.N. (1998) *Idempotent mathematics and interval analysis*. ESI, Vienna. (Preprint ESI 632)

[3] Heidergott B., Olsder G.J., and van der Woude J. (2006) *Max-Plus at Work: Modeling and Analysis of Synchronized Systems*. Princeton University Press, Princeton.

[4] Heidergott B. (2006) *Max-Plus Linear Stochastic Systems and Perturbation Analysis*. Springer, New York.

[5] Jean-Marie A. (1994) Analytical computation of Lyapunov exponents in stochastic event graphs. // Performance Evaluation of Parallel and Distributed Systems. Solution Methods: Proc. 3rd QMIPS Workshop. CWI, Amsterdam, 309–341. (CWI Tracts, Vol. 106.)

[6] Olsder G.J., Resing J.A.C., De Vries R.E., Keane M.S., Hooghiemstra G. (1990) Discrete event systems with stochastic processing times. // IEEE Transactions on Automatic Control, **35**, 3, 299–302.

[7] Krivulin N.K. (2007) The growth rate of state vector in a generalized linear stochastic system with symmetric matrix. // Journal of Mathematical Sciences, **142**, 4, 6924–6928.

[8] Krivulin N.K. (2009) Evaluation of the Lyapunov exponent for generalized linear systems with exponential distribution of elements of transition matrix. // Vestnik St.Petersburg University: Mathematics, **42**.

[9] Krivulin N.K. (2008) Evaluation of the growth rate of the state vector in a second-order generalized linear stochastic system. // Vestnik St.Petersburg University: Mathematics, **41**, 1, 28–38.

Section

Monte-Carlo methods and stochastic modeling

6th St.Petersburg Workshop on Simulation (2009) 885-889

Limiting Distributions for Randomly-Shifted Lattice Rules[1]

Pierre L'Ecuyer[2], Bruno Tuffin[3]

Abstract

Quasi-Monte Carlo (QMC) methods approximate the integral of a function f over the unit cube by the average of evaluations of f at n points having a highly-uniform distribution. In Randomized QMC (RQMC), the points are also randomized so that the average provides an unbiased estimator. Error analysis has been done in terms of worst-case error bounds and variance bounds, for various classes of functions and QMC or RQMC point sets, but limit theorems for the average (when $n \to \infty$) have been obtained only for a few special cases of RQMC methods In this presentation, we examine the distribution of the average when the RQMC point set is a randomly-shifted lattice rule and the integrand is smooth. This distribution function, properly standardized, typically converges to a spline of degree equal to the dimension. Thus, the limiting distribution is not normal, but piecewise-polynomial. There are also special cases where convergence is faster due to error cancellation.

1. Randomized quasi-Monte Carlo

Randomized quasi-Monte Carlo (RQMC) methods estimate the integral of some function $f : [0,1)^s \to \mathbb{R}$,

$$\mu = \mu(f) = \int_{[0,1)^s} f(\mathbf{u})\, d\mathbf{u} = \mathbb{E}[f(\mathbf{U})], \tag{1}$$

by the average of integrand evaluations at n highly-uniform randomized points $\mathbf{U}_0, \ldots, \mathbf{U}_{n-1}$:

$$\hat{\mu}_{n,\mathrm{rqmc}} = \frac{1}{n} \sum_{i=0}^{n-1} f(\mathbf{U}_i). \tag{2}$$

Here, $[0,1)^s$ is the s-dimensional unit hypercube, $\mathbf{u} = (u_1, \ldots, u_s)$ represents a point in this cube, $\mathbf{U}, \mathbf{U}_0, \ldots, \mathbf{U}_{n-1}$ are random points uniformly distributed

[1]This work was supported by an NSERC-Canada Grant and a Canada Research Chair to the first author, EuroNF Network of Excellence to the second author, and INRIA's associated team MOCQUASIN to both authors.

[2]Université de Montréal, `http://www.iro.umontreal.ca/~lecuyer`

[3]INRIA Rennes – Bretagne Atlantique, `bruno.tuffin@irisa.fr`

in this cube, and the (randomized) point set $P_n = \{\mathbf{U}_0, \ldots, \mathbf{U}_{n-1}\}$ is assumed to be highly-uniform over $[0,1)^s$. Thus, each point of the RQMC point set $\{\mathbf{U}_0, \ldots, \mathbf{U}_{n-1}\}$ has the uniform distribution over the hypercube, but the points are stochastically dependent. This framework covers most stochastic simulations whose aim is to estimate a mean; the dimension s represents the number of calls to the uniform random number generator that drives the simulation. For background on QMC and RQMC, and more precise definitions of "highly uniform," we refer the reader to [4, 6, 9, 11, 12].

The RQMC estimator has mean $\mathbb{E}[\hat{\mu}_{n,\mathrm{rqmc}}] = \mu$ and variance

$$\mathrm{Var}[\hat{\mu}_{n,\mathrm{rqmc}}] = \mathbb{E}[(\hat{\mu}_{n,\mathrm{rqmc}} - \mu)^2]. \tag{3}$$

One way to estimate this variance and compute a confidence interval on μ is to obtain m independent realizations of $\hat{\mu}_{n,\mathrm{rqmc}}$, say $X_1, \ldots, X_m$, based on independent randomizations of P_n, and compute their sample mean $\bar{X}_m$ and their sample variance and $S^2_{x,m}$. One has $\mathbb{E}[\bar{X}_m] = \mu$ and $\mathbb{E}[S^2_{x,m}] = m\mathrm{Var}[\bar{X}_m]$ [5, 6]. By assuming that $\bar{X}_m$ is approximately normally distributed, one can then compute a confidence interval on μ in a standard way.

However, we cannot assume that the distribution of $\hat{\mu}_{n,\mathrm{rqmc}}$ is approximately normal in general. Central-limit theorems (for $n \to \infty$) have been proved for the special cases of Latin hypercube sampling (LHS) [10] and digital nets with full nested scrambling [7]. In the LHS case, even a bound on the total variation convergence to the normal distribution is available. See [8] for a survey. However, LHS is not among the most powerful RQMC methods, because it ensures good uniformity only for the one-dimensional projections, and nested scrambling is not very popular in practice because it is very time-consuming. The purpose of this presentation is to shed light on the distribution of $\hat{\mu}_{n,\mathrm{rqmc}}$ for the special case of a randomly-shifted lattice rule, a popular RQMC method whose definition is recalled below.

2. Randomly-shifted lattice rules

An *integration lattice* is a discrete vector space of the form

$$L_s = \left\{ \mathbf{v} = \sum_{j=1}^{s} h_j \mathbf{v}_j \text{ such that each } h_j \in \mathbb{Z} \right\},$$

where $\mathbf{v}_1, \ldots, \mathbf{v}_s \in \mathbb{R}^s$ are linearly independent over $\mathbb{R}$ and where L_s contains $\mathbb{Z}^s$, the set of integer vectors. The QMC approximation of μ with the lattice point set $P_n = L_s \cap [0,1)^s$ is a *lattice rule* [5, 12, 13]. The *dual lattice* to L_s is

$$L_s^* = \{\mathbf{h} \in \mathbb{R}^s : \mathbf{h} \cdot \mathbf{v} \in \mathbb{Z} \text{ for all } \mathbf{v} \in L_s\},$$

where the "$\cdot$" denotes the ordinary scalar product. All its vectors have integer coordinates and all coordinates of vectors $\mathbf{v} \in L_s$ are multiples of $1/n$. Lattice

rules typically are of rank 1, which means that $\mathbf{v}_j = \mathbf{e}_j$ (the jth unit vector in s-dimensions) for $j = 2, \ldots, s$. Then we can write

$$P_n = \{\mathbf{v} = i\mathbf{v}_1 \bmod 1,\, i = 0, \ldots, n-1\} = \{(i\mathbf{a}_1 \bmod n)/n,\, i = 0, \ldots, n-1\},$$

where $\mathbf{a}_1 = (a_1, \ldots, a_s)$ and $\mathbf{v}_1 = \mathbf{a}_1/n$. We usually take $a_1 = 1$ and $\gcd(a_j, n) = 1$ for each j, so that each one dimensional projection of P_n is the set $\{0, 1, \ldots, n-1\}$. The uniformity of P_n depends on the choice of $\mathbf{a}_1$ and can be measured in various ways [4, 12].

To apply a *random shift modulo 1* [2, 5, 12], we generate a single point $\mathbf{U}$ uniformly over $(0,1)^s$ and add it to each point of P_n, modulo 1, coordinate-wise. For a lattice rule, this is the same as randomly shifting L_s and then taking the intersection with the unit hypercube $[0,1)^s$. This gives an RQMC point set: The lattice structure is preserved (we obtained a shifted lattice) and each point $\mathbf{U}_i = (i\mathbf{v}_1 + \mathbf{U}) \bmod 1$ of the randomized version of P_n is uniformly distributed over $[0,1)^s$.

For a randomly-shifted lattice rule, the integration error is a function of the random point $\mathbf{U}$, say $g_n(\mathbf{U}) = \hat{\mu}_{n,\mathrm{rqmc}} - \mu$. We are interested in the distribution of this random variable $g_n(\mathbf{U})$ when n is large. Let

$$f(\mathbf{u}) = \sum_{\mathbf{h} \in \mathbb{Z}^s} \hat{f}(\mathbf{h}) \exp(2\pi \mathrm{i}\, \mathbf{h} \cdot \mathbf{u}), \tag{4}$$

be the Fourier expansion of f, with Fourier coefficients

$$\hat{f}(\mathbf{h}) = \int_{(0,1)^s} f(\mathbf{u}) \exp(-2\pi \mathrm{i}\, \mathbf{h} \cdot \mathbf{u})\, d\mathbf{u},$$

where $\mathrm{i} = \sqrt{-1}$. Then one can show that the Fourier coefficients of g are $\hat{g}_n(\mathbf{h}) = \hat{f}(\mathbf{h})$ if $\mathbf{0} \neq \mathbf{h} \in L_s^*$, and $\hat{g}_n(\mathbf{h}) = 0$ otherwise [5]. Therefore, the Fourier expansion of g_n can be written as

$$g_n(\mathbf{u}) = \sum_{\mathbf{0} \neq \mathbf{h} \in L_s^*} \hat{f}(\mathbf{h}) \exp(2\pi \mathrm{i}\, \mathbf{h} \cdot \mathbf{u}) \tag{5}$$

and by the Parseval equality the variance is

$$\mathrm{Var}[\hat{\mu}_{n,\mathrm{rqmc}}] = \mathrm{Var}[g_n(\mathbf{U})] = \mathbb{E}[g_n^2(\mathbf{U})] = \sum_{\mathbf{0} \neq \mathbf{h} \in L_s^*} |\hat{f}(\mathbf{h})|^2$$

whenever f is square integrable [5]. It seems difficult to figure out the distribution of $g_n(\mathbf{U})$ directly from (5), so we will take a different path.

By making assumptions on how fast the Fourier coefficients converge when the size of $\mathbf{h}$ increases, we can obtain asymptotic bounds on the worst-case error $\sup_{\mathbf{u} \in [0,1)^s} |g_n(\mathbf{u})|$ or on the variance $\mathrm{Var}[g_n(\mathbf{U})]$. For example, let $\alpha > 1/2$, take some non-negative constants $\gamma_1, \ldots, \gamma_s$, and consider the class of functions $f : [0,1)^s \to \mathbb{R}$ such that for all $\mathbf{h} = (h_1, \ldots, h_s) \in \mathbb{Z}^s$,

$$|\hat{f}(\mathbf{h})|^2 \leq w(\mathbf{h}) \stackrel{\mathrm{def}}{=} \prod_{\{j : h_j > 0\}} \gamma_j |h_j|^{-2\alpha}.$$

It is known that there exist a sequence of rank-1 lattice rules, indexed by n, such that for any $\delta > 0$, the worst-case error and the variance converge as $O(n^{-\alpha+\delta})$ and $O(n^{-2\alpha+\delta})$, respectively [3, 12]. When α is an integer, the above condition can be written in terms of square integrability of a collection of partial derivatives: For every subset of coordinates, the partial derivative of order α with respect to these coordinates must be square integrable [3, 4].

3. Error distribution

To see what the distribution of $g(\mathbf{U})$ looks like, we start in one dimension ($s = 1$). The randomly-shifted lattice can then be written as $\{U/n, (1+U)/n, \ldots, (n-1+U)/n)$ where U has the uniform distribution over $[0, 1)$, because the random shift can be generated equivalently over $[0, 1/n)$ in this case. If f is sufficiently smooth, we can write its Taylor expansion around the center $x_i = (i + 1/2)/n$ of each interval $[i/n, (i+1)/n)$, which gives

$$f(u) = a_i + (u - x_i)\, b_i + (u - x_i)^2\, \frac{c_i}{2} + e_i$$

for $i/n \le u < (i+1)/n$ and $i = 0, \ldots, n-1$, where $\sup_i |e_i| = O(n^{-3})$. Then it is easily seen that

$$g(U) = \frac{1}{n} \sum_{i=0}^{n-1} \left[\frac{(U - 1/2) b_i}{n} + O(n^{-2}) \right] = (U - 1/2) \frac{B_n}{n} + O(n^{-2}),$$

where $B_n = \sum_{i=0}^{n-1} b_i/n$ can be interpreted as the approximate average slope of f between 0 and 1, and converges to $\int_0^1 f'(u)du = f(1) - f(0)$ with error $O(1/n)$ when $n \to \infty$, under mild conditions, for example if we take $b_i = f'(x_i)$ and if f' exists and is Rieman-integrable. Thus,

$$W_n = 1/2 + \frac{n g(U)}{f(1) - f(0)}$$

converges to a $U(0, 1)$ random variable when $n \to \infty$, if $f(1) \ne f(0)$.

When $f(1) = f(0)$, the periodic continuation of f is continuous and the integration errors on the linear pieces cancel out asymptotically: we obtain $B_n = O(1/n)$. Convergence then turns out to be faster and we have

$$\begin{aligned} g(U) &= \frac{1}{n} \sum_{i=0}^{n-1} \left[\frac{(U - 1/2) b_i}{n} + \frac{(U - 1/2)^2 c_i}{2n^2} + O(n^{-3}) \right] \\ &= (U - 1/2) \frac{B_n}{n} + (U - 1/2)^2 \frac{C_n}{2n^2} + O(n^{-3}), \end{aligned}$$

where $C_n = \sum_{i=0}^{n-1} c_i/n$, which converges to $f'(1) - f'(0)$ when $n \to \infty$ if f'' is Rieman-integrable. If $f'(1) \ne f'(0)$, using the fact that nB_n converges to

$-1/24$, we have that $2n^2 g(U)$ converges in distribution to the random variable $[(U-1/2)^2 - 1/12](f'(1) - f'(0))$, and therefore

$$W_n = \frac{2n^2 g(U)}{[f'(1) - f'(0)]} + \frac{1}{12}$$

converges in distribution to $(U-1/2)^2$, whose density is $1/\sqrt{x}$ for $0 < x < 1/4$. This means that $2\sqrt{W_n}$ converges to a uniform random variable over $(0,1)$.

An interesting situation where $f(1) = f(0)$ is when f is symmetric with respect to $1/2$ and n is even. In that case, we have $b_i = -b_{n-i}$, so the errors on the linear parts in the intervals $[i/n, (i+1)/n)$ and $[(n-i-1)/n, (n-i)/n)$ cancel each other exactly. Note that a non-symmetric integrand f can be transformed into a symmetric integrand $\tilde{f}$ having the same integral by defining $\tilde{f}(1-u) = \tilde{f}(u) = f(2u)$ for $0 \le u \le 1/2$. Equivalently, one can keep f unchanged and transform the randomized points U_i via $\tilde{U}_i = 2U_i$ if $U_i < 1/2$ and $\tilde{U}_i = 2(1-U_i)$ if $U_i \ge 1/2$. This is known as the baker's transformation; it stretches all the points by a factor of two and then folds back those that exceed 1. It is easily seen that after applying this transformation, the lattice points become locally antithetic in each interval of the form $[i/n, (i+2)/n]$, in the sense that they are at equal distance from the center of the interval, on each side. As a result, they integrate exactly any linear function over this interval. Because this holds for every such interval, a piecewise-linear approximation which is linear over each interval is integrated exactly.

In s dimensions, the limiting distribution is more complicated. We consider the unit hypercube as a torus, for which all coordinates are reduced modulo 1. Every basis $\{\mathbf{v}_1, \ldots, \mathbf{v}_s\}$ of L_s determines a parallelepiped of volume $1/n$, defined as the set of vectors of the form $\mathbf{v} = u_1\mathbf{v}_1 + \cdots + u_s\mathbf{v}_s$ where $0 \le u_j < 1$ for each j, and the torus is the union of exactly n shifted copies of this parallelepiped. Moreover, the randomly-shifted lattice has exactly one point uniformly distributed in each of these n parallelepipeds, at position $\mathbf{U}_i = (U_1\mathbf{v}_1 + \cdots + U_s\mathbf{v}_s) \bmod 1$, where $\mathbf{U} = (U_1, \ldots, U_s)$ is the random shift. The shifted points are at the same relative position for all the parallelepipeds. The latter correspond to the intervals of length $1/n$ in the one-dimensional case.

If f has enough smoothness we can approximate it by the linear term of its Taylor expansion over each parallelepiped. The integration error for this linear approximation is a linear combination of $U_1 - 1/2, \ldots, U_s - 1/2$, which determine the position of the shift in the parallelepiped. Thus, $g(\mathbf{U})$ can also be written as such a linear combination, plus some terms of lesser order as a function of n. If this linear part does not vanish, by applying a theorem of Barrow and Smith [1], it follows that a properly standardized version of $g(\mathbf{U})$ has bounded support and that its cumulative distribution function is approximately a spline of degree s, with $s-1$ continuous derivatives. However, in general the function f can be discontinuous at the "boundaries", i.e., when a coordinate jumps from 1 to 0, so the Taylor approximation no longer works for the parallelepipeds that are split across a boundary. Thus the previous argument holds only if we assume that the contribution of these discontinuities becomes negligible when $n \to \infty$. This is certainly not obvious, but in some empirical experiments, we have observed that the spline of degree s was indeed a good approximation. In the case where f is

smooth on the torus, that is, if its periodic continuation is smooth, then the error on the linear part vanishes and the distributional behavior is determined by the higher-order terms of the Taylor expansion. Further details and examples will be given in the presentation.

References

[1] D. L. Barrow and P. W. Smith. Spline notation applied to a volume problem. *The American Mathematical Monthly*, 86:50–51, 1979.

[2] R. Cranley and T. N. L. Patterson. Randomization of number theoretic methods for multiple integration. *SIAM Journal on Numerical Analysis*, 13(6):904–914, 1976.

[3] J. Dick, I. H. Sloan, X. Wang, and H. Wozniakowski. Good lattice rules in weighted Korobov spaces with general weights. *Numerische Mathematik*, 103:63–97, 2006.

[4] P. L'Ecuyer. Quasi-Monte Carlo methods with applications in finance. *Finance and Stochastics*, 2008. To appear.

[5] P. L'Ecuyer and C. Lemieux. Variance reduction via lattice rules. *Management Science*, 46(9):1214–1235, 2000.

[6] P. L'Ecuyer and C. Lemieux. Recent advances in randomized quasi-Monte Carlo methods. In M. Dror, P. L'Ecuyer, and F. Szidarovszky, editors, *Modeling Uncertainty: An Examination of Stochastic Theory, Methods, and Applications*, pages 419–474. Kluwer Academic, Boston, 2002.

[7] W.-L. Loh. On the asymptotic distribution of scramble nets quadratures. *Annals of Statistics*, 31:1282–1324, 2003.

[8] W.-L. Loh. On the asymptotic distribution of some randomized quadrature rules. In C. Stein, A. D. Barbour, and L. H. Y. Chen, editors, *Stein's Method and Applications*, volume 5, pages 209–222. World Scientific, 2005.

[9] H. Niederreiter. *Random Number Generation and Quasi-Monte Carlo Methods*, volume 63 of *SIAM CBMS-NSF Regional Conference Series in Applied Mathematics*. SIAM, Philadelphia, PA, 1992.

[10] A. B. Owen. A central limit theorem for Latin hypercube sampling. *Journal of the Royal Statistical Society B*, 54(2):541–551, 1992.

[11] A. B. Owen. Latin supercube sampling for very high-dimensional simulations. *ACM Transactions on Modeling and Computer Simulation*, 8(1):71–102, 1998.

[12] I. H. Sloan and S. Joe. *Lattice Methods for Multiple Integration*. Clarendon Press, Oxford, 1994.

[13] B. Tuffin. Variance reduction order using good lattice points in Monte Carlo methods. *Computing*, 61:371–378, 1998.

6th St.Petersburg Workshop on Simulation (2009) 891-895

Kinetic approach and method of influence function to modelling of polarized radiation transfer[1]

Tamara Sushkevich, Sergey Strelkov, Svetlana Maksakova[2]

Abstract

The general vectorial boundary-value problem for the integro-differential kinetic Boltzmann's equation, which describes the polarized radiation transfer in a heterogeneous plane layer with a horizontally non-homogeneous and anisotropically reflective boundary, can not be solved by the finitely-difference methods. A mathematical model is proposed and accounted for that gives an asymptotically precise solution of the boundary-value problem in the class of the slow growth functions. The new model is constructed by the influence function method and efficient for the algorithm's of the parallel calculations.

1. Introduction

The mathematical models of the matrix-vector transfer operators are suggested to compute Stoke's parameters vector as the solution of the general boundary-value problem in polarized radiation transfer theory for the 1d, 2d, 3d plane layers and spherical shell's. Space and angular distributions of polarized radiation inside the relevant layer of these media as well as reflected and passed through the layer radiation are formed as a result of multiple scattering and absorption, polarization and depolarization. A new approach is proposed to radiation transfer modelling in optically thick layers that are represented as a heterogeneous system with multi-layers each layer of which is described by different radiation conditions.

The kernels of the matrix-vector transfer operators are the tensors of the influence functions. The tensors of the influence functions of each layer are determined by solutions of the radiation transfer first boundary vector problem with an vector external source function [1]. The boundary vector problem for each layer can be solved depending on optical thickness, scattering and absorption characteristics by one of the following techniques:

i) as a solution of transfer equation with an azimuth dependence;
ii) as a solution of the problem with azimuth symmetry;
iii) as a solution in two-flux approach;
iv) as an approximate solution in asymptotic approach.

[1]This work was supported by RFBR grants 08-01-00024 and 09-01-00071.
[2]Keldysh Institute of Applied Mathematics of RAS, E-mail:tamaras@keldysh.ru

The matrix operators of radiation transmittance and reflectance on the boundaries between the layers are formulated based on the collision integrals and the separate layers are united in a system by these operators. The representation of the solution to boundary-value vector problem as a functional is the transfer operator of the radiation transfer system which establishes the explicit relationship between the recording radiation and the "scenarios" (the optical image) at the dividing boundaries of media. In turn, by the use of the tensors of the influence functions, the "scenarios" is described clearly through the characteristics of the reflection and transmission of the dividing boundaries at the given its illumination. The tensors of the influence functions are invariant about the conditions of the illumination and the properties of the dividing boundaries.

2. Mathematical statement of the polarized radiation transfer problem

The vector of the Stoke's parameters $\mathbf{\Phi}$ and the first of all its component I, called as the radiation intensity (or the radiance), being considered as the most complete characteristics of the quasi-monochromatic electromagnetic field, is in fact a complex functional from the parameters of the medium and the boundaries as well as from the sources. In SP-representation (Stoke's-Poincaré) the components of any vector-column $\mathbf{\Phi} = (I, Q, U, V)^T$ have the norm of the intensity:

$$Q = Ip\cos 2\chi\cos 2\beta, \quad U = Ip\sin 2\chi\cos 2\beta, \quad V = Ip\sin 2\beta,$$

where χ — the azimuth of the polarization plane, β — the ellipticity degree, $0 \le p \le 1$ — the polarization degree.

The Stoke's vector can be found as a solution of the general vector boundary-value problem of the transfer theory (GVBP at $\hat{R} \not\equiv 0$)

$$\hat{K}\mathbf{\Phi} = \mathbf{F}, \quad \mathbf{\Phi}\,|_{\,t} = \mathbf{F}^0, \quad \mathbf{\Phi}\,|_{\,b} = \varepsilon\hat{R}\mathbf{\Phi} + \mathbf{F}^H \tag{1}$$

with the linear operators: the transfer operator

$$\hat{D} \equiv (s, grad) + \sigma(z) = \hat{D}_z + \left(s_\perp, \frac{\partial}{\partial r_\perp}\right), \quad \hat{D}_z \equiv \mu\frac{\partial}{\partial z} + \sigma(z)\ ;$$

the collisions integral

$$\hat{S}\mathbf{\Phi} \equiv \sigma_s(z)\int_\Omega \hat{P}(z, s, s')\mathbf{\Phi}(z, r_\perp, s')\,ds', \quad ds' = d\mu'\,d\varphi'\ ;$$

the uniformly restricted reflectivity operator

$$[\hat{R}\mathbf{\Phi}](H, r_\perp, s) \equiv \int_{\Omega^+} \hat{q}(r_\perp, s, s^+)\mathbf{\Phi}(H, r_\perp, s^+)\,ds^+\ ; \tag{2}$$

the integro-differential operator $\hat{K} \equiv \hat{D} - \hat{S}$; the one-dimensional operator $\hat{K}_z \equiv \hat{D}_z - \hat{S}$; $\hat{P}(z, s, s')$ is the phase matrix of the scattering; $\sigma(z)$ and $\sigma_s(z)$ are the vertical profiles of the extinction and scattering coefficients; $\hat{q}(r_\perp, s, s^+)$ is the phase

matrix of reflectivity; the parameter $0 \le \varepsilon \le 1$ fix the act of the radiation interaction with the underlying surface; $\mathbf{F}(z,s)$, $\mathbf{F}^0(r_\perp,s)$, $\mathbf{F}^H(r_\perp,s)$ are the sources of the insolation. If any from the functions $\mathbf{F}^0$, $\mathbf{F}^H$, $\hat{q}$ depends on $r_\perp$, than the solution of the problem (1)–(2) is found in the 5D-phase space $(x,y,z,\vartheta,\varphi)$, and if no dependence is from $r_\perp$, than in the 3D-phase space (z,ϑ,φ). The phase matrix of the scattering

$$\hat{P}(z,s,s') \;=\; \hat{L}(\alpha)\hat{\gamma}(z,\vartheta_s)\hat{L}(\alpha')$$

is determined through the rotary matrix $\hat{L}(\alpha)$ and the scattering matrix $\hat{\gamma}(z,\vartheta_s)$ which is the function of the scattering angle ϑ_s between the directions of the incident s' and scattered s light beams.

3. Mathematical model of polarized radiation transfer in multi-media heterogeneous system

Let us consider the radiation transfer system, consisting from M layers with the borders h_m, $\;m = 1 \div M+1$:

$$z \in [0,H], \quad H = \bigcup_{m=1}^{m=M} [h_m, h_{m+1}], \quad h_1 \;=\; 0\;, \quad h_m < h_{m+1}, \quad h_{M+1} = H\;.$$

The phase areas are introduced to write the boundary conditions:

$$d\downarrow, m = \{z,s : z = h_m, s \in \Omega^\downarrow\}\,; \;\; d\uparrow, m = \{z,s : z = h_m, s \in \Omega^\uparrow\}\,; \;\; \Omega = \Omega^\downarrow \cup \Omega^\uparrow\,;$$

$$\mu^\downarrow = \cos\vartheta^\downarrow\,, \;\; \vartheta^\downarrow \in [0,\pi/2)\;; \;\; \Omega^\downarrow = \left\{s^\downarrow = (\mu^\downarrow,\varphi) \;:\; \mu^\downarrow \in (0,1]\,,\; \varphi \in [0,2\pi]\right\}\;;$$

$$\mu^\uparrow = \cos\vartheta^\uparrow\,, \;\; \vartheta^\uparrow \in (\pi/2,\pi]\;; \;\; \Omega^\uparrow = \left\{s^\uparrow = (\mu^\uparrow,\varphi) \;:\; \mu^\uparrow \in [-1,0)\,,\; \varphi \in [0,2\pi]\right\}\;.$$

The total Stoke's vector of the radiation $\mathbf{\Phi}_\lambda(r,s)$ where the index λ is the wavelength (is omitted below) is found as a solution of the general boundary-value problem of the transfer theory for multi-layers heterogeneous system

$$\hat{K}\mathbf{\Phi} \;=\; \mathbf{F}^{in}, \quad \mathbf{\Phi}|_{\,t\downarrow} \;=\; \mathbf{F}^\downarrow_{\,t}\;, \quad \mathbf{\Phi}|_{\,b\uparrow} \;=\; \hat{R}^\uparrow_{\,b}\mathbf{\Phi} + \mathbf{F}^\uparrow_{\,b} \tag{3}$$

with the boundary conditions on the internal boundaries of the layers for $m \;=\; 2 \div M$:

$$\mathbf{\Phi}|_{\,d\uparrow,m} \;=\; \varepsilon(\hat{R}^\uparrow_m\mathbf{\Phi} \,+\, \hat{T}^\uparrow_m\mathbf{\Phi}) + \mathbf{F}^\uparrow_{m-1}\;, \quad \mathbf{\Phi}|_{\,d\downarrow,m} \;=\; \varepsilon(\hat{R}^\downarrow_m\mathbf{\Phi} \,+\, \hat{T}^\downarrow_m\mathbf{\Phi}) + \mathbf{F}^\downarrow_m \tag{4}$$

and for the external boundaries of the system

$$\mathbf{F}^\downarrow_1 \;=\; \mathbf{F}^\downarrow_{\,t}\;; \quad \mathbf{F}^\uparrow_M \;=\; \mathbf{F}^\uparrow_{\,b}\;; \quad d\downarrow, 1 \;=\; t\downarrow\;; \quad d\uparrow, M+1 \;=\; b\uparrow\;. \tag{5}$$

We search a solution in the form of the regular perturbation series

$$\mathbf{\Phi} \;=\; \sum_{n=0}^{\infty} \varepsilon^n \mathbf{\Phi}^{(n)}\;.$$

We introduce the algebraic vectors with dimension $2M$:
the complete solution

$$\mathbf{\Phi} \;=\; \{\mathbf{\Phi}_1^{\downarrow}, \mathbf{\Phi}_1^{\uparrow}, \mathbf{\Phi}_2^{\downarrow}, \mathbf{\Phi}_2^{\uparrow}, \ldots, \mathbf{\Phi}_m^{\downarrow}, \mathbf{\Phi}_m^{\uparrow}, \ldots, \mathbf{\Phi}_M^{\downarrow}, \mathbf{\Phi}_M^{\uparrow}\} \;;$$

n-approximation of the sources

$$\mathbf{F}^{(n)} \;=\; \{\mathbf{F}_1^{\downarrow(n)}, \mathbf{F}_1^{\uparrow(n)}, \mathbf{F}_2^{\downarrow(n)}, \mathbf{F}_2^{\uparrow(n)}, \ldots, \mathbf{F}_m^{\downarrow(n)}, \mathbf{F}_m^{\uparrow(n)}, \ldots, \mathbf{F}_M^{\downarrow(n)}, \mathbf{F}_M^{\uparrow(n)}\} \;;$$

n-approximation of the solution

$$\mathbf{\Phi}^{(n)} \;=\; \{\mathbf{\Phi}_1^{\downarrow(n)}, \mathbf{\Phi}_1^{\uparrow(n)}, \mathbf{\Phi}_2^{\downarrow(n)}, \mathbf{\Phi}_2^{\uparrow(n)}, \ldots, \mathbf{\Phi}_m^{\downarrow(n)}, \mathbf{\Phi}_m^{\uparrow(n)}, \ldots, \mathbf{\Phi}_M^{\downarrow(n)}, \mathbf{\Phi}_M^{\uparrow(n)}\} \;;$$

the initial approximation of the sources

$$\mathbf{E} \;=\; \{\mathbf{E}_1^{\downarrow}, \mathbf{E}_1^{\uparrow}, \mathbf{E}_2^{\downarrow}, \mathbf{E}_2^{\uparrow}, \ldots, \mathbf{E}_m^{\downarrow}, \mathbf{E}_m^{\uparrow}, \ldots, \mathbf{E}_M^{\downarrow}, \mathbf{E}_M^{\uparrow}\} \;;$$

the "scenario" at the boundaries

$$\mathbf{Z} \;=\; \{\mathbf{Z}_1^{\downarrow}, \mathbf{Z}_1^{\uparrow}, \mathbf{Z}_2^{\downarrow}, \mathbf{Z}_2^{\uparrow}, \ldots, \mathbf{Z}_m^{\downarrow}, \mathbf{Z}_m^{\uparrow}, \ldots, \mathbf{Z}_M^{\downarrow}, \mathbf{Z}_M^{\uparrow}\} \;;$$

the vectorial influence functions of the layers

$$\mathbf{\Theta} \;=\; \{\mathbf{\Theta}_1^{\downarrow}, \mathbf{\Theta}_1^{\uparrow}, \mathbf{\Theta}_2^{\downarrow}, \mathbf{\Theta}_2^{\uparrow}, \ldots, \mathbf{\Theta}_m^{\downarrow}, \mathbf{\Theta}_m^{\uparrow}, \ldots, \mathbf{\Theta}_M^{\downarrow}, \mathbf{\Theta}_M^{\uparrow}\} \;;$$

the influence functions tensors of the layers

$$\hat{\Pi} \;=\; \{\hat{\Pi}_1^{\downarrow}, \hat{\Pi}_1^{\uparrow}, \hat{\Pi}_2^{\downarrow}, \hat{\Pi}_2^{\uparrow}, \ldots, \hat{\Pi}_m^{\downarrow}, \hat{\Pi}_m^{\uparrow}, \ldots, \hat{\Pi}_M^{\downarrow}, \hat{\Pi}_M^{\uparrow}\} \;.$$

We decompose the original problem (3)–(5) into $2M$ problems with their own boundary conditions. The initial approach is the radiation from the sources without the radiation exchange between the layers ($\mathbf{F}_1^{\downarrow} \;=\; \mathbf{F}_t^{\downarrow}$; $\mathbf{F}_M^{\uparrow} \;=\; \mathbf{F}_b^{\uparrow}$) for $m \;=\; 1 \div M$:

$$\hat{K}\mathbf{\Phi}_m^{\downarrow(0)} \;=\; \mathbf{F}_m^{\downarrow in} \;, \quad \mathbf{\Phi}_m^{\downarrow(0)}\Big|_{d\downarrow,m} \;=\; \mathbf{F}_m^{\downarrow} \;, \quad \mathbf{\Phi}_m^{\downarrow(0)}\Big|_{d\uparrow,m+1} \;=\; 0 \;;$$

$$\hat{K}\mathbf{\Phi}_m^{\uparrow(0)} \;=\; \mathbf{F}_m^{\uparrow in} \;, \quad \mathbf{\Phi}_m^{\uparrow(0)}\Big|_{d\downarrow,m} \;=\; 0 \;, \quad \mathbf{\Phi}_m^{\uparrow(0)}\Big|_{d\uparrow,m+1} \;=\; \mathbf{F}_m^{\uparrow} \;.$$

The approaches with $n \geq 1$ are described by the system of $2M$ equations for layers $m = 1 \div M$:

$$\hat{K}\mathbf{\Phi}_m^{\downarrow(n)} \;=\; 0 \;, \quad \mathbf{\Phi}_m^{\downarrow(n)}\Big|_{d\downarrow,m} \;=\; \mathbf{F}_m^{\downarrow(n-1)} \;, \quad \mathbf{\Phi}_m^{\downarrow(n)}\Big|_{d\uparrow,m+1} \;=\; 0 \;;$$

$$\hat{K}\mathbf{\Phi}_m^{\uparrow(n)} \;=\; 0 \;, \quad \mathbf{\Phi}_m^{\uparrow(n)}\Big|_{d\downarrow,m} \;=\; 0 \;, \quad \mathbf{\Phi}_m^{\uparrow(n)}\Big|_{d\uparrow,m+1} \;=\; \mathbf{F}_m^{\uparrow(n-1)}$$

with the sources at the inner boundaries h_m with $m = 2 \div M$:

$$\mathbf{F}_m^{\downarrow(n)} \;=\; \hat{T}_m^{\downarrow}\mathbf{\Phi}_{m-1}^{\downarrow(n)} \;+\; \hat{T}_m^{\downarrow}\mathbf{\Phi}_{m-1}^{\uparrow(n)} \;+\; \hat{R}_m^{\downarrow}\mathbf{\Phi}_m^{\downarrow(n)} \;+\; \hat{R}_m^{\downarrow}\mathbf{\Phi}_m^{\uparrow(n)} \;;$$

$$\mathbf{F}_m^{\uparrow(n)} \;=\; \hat{R}_{m+1}^{\uparrow}\mathbf{\Phi}_m^{\downarrow(n)} \;+\; \hat{R}_{m+1}^{\uparrow}\mathbf{\Phi}_m^{\uparrow(n)} \;+\; \hat{T}_{m+1}^{\uparrow}\mathbf{\Phi}_{m+1}^{\downarrow(n)} \;+\; \hat{T}_{m+1}^{\uparrow}\mathbf{\Phi}_{m+1}^{\uparrow(n)}$$

and at the external boundaries h_m with $m = 1$ and $m = M$:

$$\mathbf{F}_1^{\downarrow(n)} \;=\; 0\;; \quad \mathbf{F}_M^{\uparrow(n)} \;=\; \hat{R}^{\uparrow}{}_b \mathbf{\Phi}_M^{\downarrow(n)} \;+\; \hat{R}^{\uparrow}{}_b \mathbf{\Phi}_M^{\uparrow(n)} \;.$$

The solutions are found in the form of the vectorial linear functionals for each layer with $m \;=\; 1 \div M$:

$$\mathbf{\Phi}_m^{\downarrow(n)} \;=\; \left(\hat{\Pi}_m^{\downarrow}, \mathbf{F}_m^{\downarrow(n-1)}\right)\;; \quad \mathbf{\Phi}_m^{\uparrow(n)} \;=\; \left(\hat{\Pi}_m^{\uparrow}, \mathbf{F}_m^{\uparrow(n-1)}\right)\;.$$

The kernels of the functionals are the influence functions tensors $\hat{\Pi}_m^{\downarrow} = \{\mathbf{\Theta}_m^{\downarrow}\}$, $\hat{\Pi}_m^{\uparrow} = \{\mathbf{\Theta}_m^{\uparrow}\}$ of layers and theirs elements are determined from the boundary-value problems $m \;=\; 1 \div M$

$$\hat{K}\mathbf{\Theta}_m^{\downarrow} \;=\; 0, \quad \mathbf{\Theta}_m^{\downarrow}\big|_{\,d\downarrow,m} \;=\; \mathbf{f}_{\delta,m}^{\downarrow}, \quad \mathbf{\Theta}_m^{\downarrow}\big|_{\,d\uparrow,m+1} \;=\; 0\;;$$

$$\hat{K}\mathbf{\Theta}_m^{\uparrow} \;=\; 0, \quad \mathbf{\Theta}_m^{\uparrow}\big|_{\,d\downarrow,m} \;=\; 0, \quad \mathbf{\Theta}_m^{\uparrow}\big|_{\,d\uparrow,m+1} \;=\; \mathbf{f}_{\delta,m}^{\uparrow}\;.$$

In the vectorial form n-approach of the solution

$$\mathbf{\Phi}^{(n)} \;=\; \left(\hat{\Pi}, \mathbf{F}^{(n-1)}\right)\;.$$

The source in the $(n-1)$-approach

$$\mathbf{F}^{(n-1)} \;=\; \hat{P}\mathbf{\Phi}^{(n-1)}\;.$$

The matrix $\hat{P}$ is that of a tape type with the characteristics of the reflectance and transmittance of the boundaries:

The matrix-vectorial operation describes the only act of the radiation interaction on the boundaries and takes into account the multiple scattering, absorption and polarization in the layers through their influence functions tensors:

$$\hat{G}\mathbf{F} \;=\; \hat{P}(\hat{\Pi}, \mathbf{F}) \;=$$

$$\left[\begin{array}{l}
0 \\
\hat{R}_2^{\uparrow}(\hat{\Pi}_1^{\downarrow}, \mathbf{F}_1^{\downarrow}) \;+\; \hat{R}_2^{\uparrow}(\hat{\Pi}_1^{\uparrow}, \mathbf{F}_1^{\uparrow}) \;+\; \hat{T}_2^{\uparrow}(\hat{\Pi}_2^{\downarrow}, \mathbf{F}_2^{\downarrow}) \;+\; \hat{T}_2^{\uparrow}(\hat{\Pi}_2^{\uparrow}, \mathbf{F}_2^{\uparrow}) \\
\vdots \\
\hat{T}_m^{\downarrow}(\hat{\Pi}_{m-1}^{\downarrow}, \mathbf{F}_{m-1}^{\downarrow}) \;+\; \hat{T}_m^{\downarrow}(\hat{\Pi}_{m-1}^{\uparrow}, \mathbf{F}_{m-1}^{\uparrow}) \;+\; \hat{R}_m^{\downarrow}(\hat{\Pi}_m^{\downarrow}, \mathbf{F}_m^{\downarrow}) \;+\; \hat{R}_m^{\downarrow}(\hat{\Pi}_m^{\uparrow}, \mathbf{F}_m^{\uparrow}) \\
\hat{R}_{m+1}^{\uparrow}(\hat{\Pi}_m^{\downarrow}, \mathbf{F}_m^{\downarrow}) \;+\; \hat{R}_{m+1}^{\uparrow}(\hat{\Pi}_m^{\uparrow}, \mathbf{F}_m^{\uparrow}) \;+\; \hat{T}_{m+1}^{\uparrow}(\hat{\Pi}_{m+1}^{\downarrow}, \mathbf{F}_{m+1}^{\downarrow}) \;+\; \hat{T}_{m+1}^{\uparrow}(\hat{\Pi}_{m+1}^{\uparrow}, \mathbf{F}_{m+1}^{\uparrow}) \\
\vdots \\
\hat{T}_M^{\downarrow}(\hat{\Pi}_{M-1}^{\downarrow}, \mathbf{F}_{M-1}^{\downarrow}) \;+\; \hat{T}_M^{\downarrow}(\hat{\Pi}_{M-1}^{\uparrow}, \mathbf{F}_{M-1}^{\uparrow}) \;+\; \hat{R}_M^{\downarrow}(\hat{\Pi}_M^{\downarrow}, \mathbf{F}_M^{\downarrow}) \;+\; \hat{R}_M^{\downarrow}(\hat{\Pi}_M^{\uparrow}, \mathbf{F}_M^{\uparrow}) \\
\hat{R}^{\uparrow}{}_b(\hat{\Pi}_M^{\downarrow}, \mathbf{F}_M^{\downarrow}) \;+\; \hat{R}^{\uparrow}{}_b(\hat{\Pi}_M^{\uparrow}, \mathbf{F}_M^{\uparrow})
\end{array}\right]\;.$$

Two successive n-approaches are connected by the recurrent relation where $\mathbf{E}$ is an initial approach:

$$\mathbf{\Phi}^{(n)} \;=\; \left(\hat{\Pi}, \hat{P}\mathbf{\Phi}^{(n-1)}\right) \;=\; \left(\hat{\Pi}, \hat{G}^{n-1}\mathbf{E}\right) \;.$$

Asymptotically precise solution is obtained in the form of the matrix-vectorial linear functional, i.e. the matrix transfer operator:

$$\mathbf{\Phi} \;=\; \left(\hat{\mathbf{\Theta}}, \mathbf{Z}\right) \;. \tag{6}$$

"Scenario" is given by the vector $\mathbf{Z}$ of the radiance distributions on the boundaries

$$\mathbf{Z} \;\equiv\; \hat{Z}\mathbf{E} \;\equiv\; \sum_{n=0}^{\infty} \hat{G}^n\mathbf{E} \;=\; \mathbf{E} \;+\; \sum_{n=1}^{\infty} \hat{P}\mathbf{\Phi}^{(n)} \;=\; \mathbf{E} + \sum_{n=1}^{\infty} \mathbf{F}^{(n)}$$

and is the Neumann's series sum on the multiplicity of the radiation transfer through the boundaries taking into account the multiple scattering impact using the influence functions tensors of each layer.

The stages of the calculation realization:

1. Calculation of the vectorial influence functions with parametric dependence for each layer is realized by algorithms of parallel computation on multi-processor computers and writing them down into archives of solutions. The computational method is selected in each layer depending on the radiation regime of this layer. Two algorithms of the parallel computations are realized: on the layers ("domain-decomposition" of the system) and on parameters of the influence functions.

2. Calculation of the "scenario" vector on the boundaries of the layers through the matrix-vector procedure.

3. Calculation of the angular and spatial distributions of the radiation inside the system and on its boundaries using the matrix transfer operator (6).

References

[1] Sushkevich T.A. (2005) *Mathematical models of radiation transfer*. Moscow, BINOM. Laboratory of Knowledge Publishers.

6th St.Petersburg Workshop on Simulation (2009) 897-901

A Semi-Markov Model for Patients Following a Clinically Isolated Syndrome Event Prior to Progression to Clinically Definite Multiple Sclerosis

John H. Walker[1] and Michael Iskedjian[2]

Abstract

This paper presents the methodology that was used for developing a Semi-Markov model for modelling the progression of patients who have had a single demyelinating event [Clinically Isolated Syndrome (CIS)] that is suggestive of Multiple Sclerosis (MS). This model can be used for comparing CIS and MS modifying agents in terms of costs and outcomes.

1. Introduction

Decision analytic models can be used for estimating the cost effectiveness of health-care interventions. These models allow for the synthesis of data from various sources as well as the extrapolation of data from primary data sources [1]. When using a Markov model for a chronic disease, the disease is divided into a number of distinct, mutually exclusive, health states. Transition probabilities are assigned to movements among these states during a discrete time frame, called a Markov Cycle. The Markov Cycle length is chosen to represent a clinically meaningful time length [2]. If the transition probabilities among the states as well as health outcomes change with time, then we have a semi–Markov process [3].

2. Background–Multiple Sclerosis

Multiple Sclerosis is a progressive disease of the central nervous system that has serious long-term consequences. MS is characterized by areas of demyelination and axon injury [4]. It has an onset of about 30 years [5] and is more common in women than in men [6]. For patients to be diagnosed with Clinically Definite Multiple Sclerosis (CDMS), they are required to have experienced at least two neurological demyelinating events separated in both time and space [7]. Revised diagnostic criteria for MS diagnosis as reported by McDonald et al. [8] suggest the

[1]Faculty of Business, Brock University, 500 Glenridge Avenue, St. Catharines, Ontario, Canada L2S 3A1. E-mail: jowalker@brocku.ca

[2]PharmIdeas Research and Consulting Inc., 1175 North Service Road West, Oakville, Ontario, Canada L6M 2W1. E-mail: miskedjian@pharmideas.com

integration of magnetic resonance imaging (MRI) and clinical diagnostic methods to facilitate the diagnosis of MS. The authors also suggested that, after clinical evidence of one lesion, MRI based evidence of a second demyelinating event should fulfill the criteria of a proper MS diagnosis. However, MRI-based evidence alone is not sufficient as put forth by O'Connor and Uitdehaag [4,9]. Individuals who have had a single clinical attack suggestive of MS should be classified as having CIS. Within CIS, patients are categorized as monofocal (signs and symptoms could only be attributed to a single lesion) or multifocal (signs and symptoms could be attributed to multiple lesions).

At the time of this work, two products were available in Canada for treating CIS patients, Interferon beta-1a (Avonex®) and Interferon beta-1b (Betaseron®) [10,11]. The two clinical trials, the Controlled High Risk Subjects Avonex Multiple Sclerosis Prevention Study (CHAMPS) and the BEtaseron in Newly Emerging Multiple Sclerosis for Initial Treatment (BENEFIT), examined the treatment of patients with CIS. The CHAMPS study determined that the two year probabilities of progression to CDMS were 0.211 and 0.386 for the Avonex® and placebo groups, respectively [12]. Furthermore, Avonex was found to impact the progression of MS patients, and also reduce the number of relapses [12,13]. The BENEFIT study found that the two year probabilities of progression to CDMS were 0.280 and 0.450 for the Betaseron® and placebo groups, respectively [14]. Furthermore, clinical studies also found that Betaseron® reduced relapse rates and reduced CDMS progression [15,16].

3. Model

A Markov model was developed using the TreeAge Pro Suite 2006 decision analysis software package (TreeAge Software Inc.) [17]. The model horizon was set at 15 years. This was based on the median time to progress to CDMS from the Avonex® clinical trial (about 6 years) [12] and adding the median time to DSS 3 (approximately 7 years) from the natural history data [19].

Transitional Probabilities

The length of each cycle was set as 1 year. The CIS health state was used as the entry point for all patients following a CIS; thus, the probability of being in the CIS health state during the first cycle was 1.0. At the end of the first year, patients could remain in the CIS state or experience an event and transition into CDMS at EDSS levels 1–6.

The probability of transitioning out of the CIS state into the various EDSS levels was derived using the proportion of patients who reached different CDMS levels in the CHAMPS study as reported by Jacobs et al. [18] The CHAMPS study results for all patients was used since there are no significant differences the placebo and Avonex arms of the clinical trial. We derived the probabilities of transitioning into the different EDSS levels for Avonex® and for Betaseron® by multiplying the annualized rate $= 1 - \sqrt{[1 - (Two - YearPate)]}$ of transitioning out of the CIS state by the proportion of patients who reached a specific EDSS level. For the Best Supportive Care (BSC) rate, we used a weighted average of the annualized

Table 1: Summary of transitional probabilities from the CIS health state

	Transitional Probability (%)		
Transition	Best Supportive Care	Avonex®	Betaseron®
CIS to EDSS1	0.091[1] (38.6[2])	0.043[3] (38.6[2])	0.058[4] (38.6[2])
CIS to EDSS2	0.073[1] (30.7[2])	0.034[3] (30.7[2])	0.047[4] (30.7[2])
CIS to EDSS3	0.052[1] (21.9[2])	0.024[3] (21.9[2])	0.033[4] (21.9[2])
CIS to EDSS4	0.004[1] (1.8[2])	0.002[3] (1.8[2])	0.003[4] (1.8[2])
CIS to EDSS5	0.010[1] (4.4[2])	0.005[3] (4.4[2])	0.007[4] (4.4[2])
CIS to EDSS6	0.006[1] (2.6[2])	0.003[3] (2.6[2])	0.004 (2.6[2])
Total	**0.237 (100)**	**0.112 (100)**	**0.152 (100)**

[1]Jacobs et al. and Kappos et al. [12,14], [2]Jacobs et al. [18], [3]Jacobs et al. [12], and [4]Kappos et al.[14]

placebo rates for the two trials since there has not been a clinical trial examining both interferon therapies with a common placebo group. These probabilities are presented in Table 1 below with the proportion of patients reaching specific CDMS levels.

In the second cycle of the model, (i.e., Year 2), patients could again transition from the CIS state to different CDMS levels, denoted as Extended Disability Status Scale (EDSS) levels developed by Kurtzke [21]. However, patients who started Year 2 in an EDSS level had their transitioning ability restricted to the same EDSS level, or progress to the next EDSS level. EDSS 6 was considered to be an absorbing state. Transitions within the model are depicted in Figure 1.

The probabilities associated with transitioning through the various EDSS stages of the model were time-dependent [19]. Tracker variables were used to account for the number of years spent in CIS as well as each CDMS level. All outcomes were determined using a 10,000-iteration Monte-Carlo simulation. The probabilities for transitioning through the various EDSS levels for Best Supportive Care were determined from Weinshenker et al. [20] because the data are based on natural history and have not been compromised by any therapeutic intervention for the treatment of MS.

The EDSS transitional probabilities for the therapy treatments were determined by modifying the transitional EDSS probabilities of the Best Supportive Care by the reduction in EDSS progression percentages from the pivotal clinical trials, which were 37%, and 29%, for Avonex®, and Betaseron®, respectively [13,15].

4. Model Validation

The model was validated for both the estimated time spent in the CIS state and for the progression through CDMS. Jacobs reported a median time of progression from the CIS state to CDMS of 36 months for the placebo group and an estimated

median time of 6 years to progress from the clinically isolated syndrome state to CDMS in the Avonex® group was determined from the Kaplan–Meier curve in the Jacobs study [12]. A similar approach was taken for Betaseron® using the BENEFIT study data [14]. We estimated that the median times were 3.8 years and 2.1 years, respectively, for the Betaseron® group and the placebo group. Results are presented in Table 2 below. Acknowledgements

This work was supported, in part, by a grant from Biogen Idec Canada Inc.

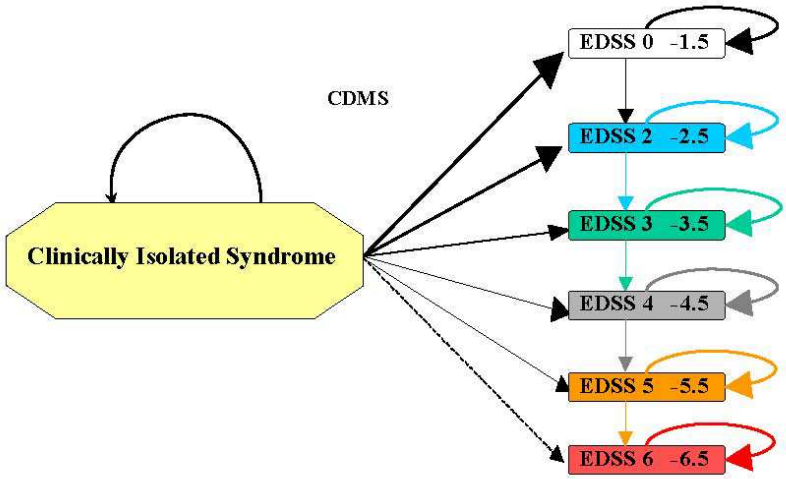


Figure 1: Possible transitions within the Markov model

References

[1] Briggs, A. and Sculpher, M. An Introduction to Markov Modelling for Economic Evaluation. *Pharmacoeconomics* (1998); 13(4): 397–409.

[2] Sonnenberg, FA and Beck, JR. Markov Models in Medical Decision Making: A Practical Guide. *Medical Decision Making* (1993); 13(4): 322–338.

[3] Ross, S. *Introduction to Probability Models*. New York, New York; Academic Press, (1997).

[4] O'Connor P. Key issues in the diagnosis and treatment of multiple sclerosis. An overview. *Neurology* (2002); 59(6 Suppl 3): S1–S33.

[5] Vukusic S, Confavreux C. The natural history of multiple sclerosis. In Cook S, ed. *Handbook of Multiple Sclerosis*, New York, New York: Marcel Dekker, Inc., (2001).

Table 2: Results of the model validation

Parameter	Median time (years) according to criterion	Model estimated median time (years)	Relative Percentage Difference (Model-Standard)/ Standard*100%
CIS→ CDMS			
Avonex®*	6.0	5.8	3.33%
BSC*	3.0	2.9	10.00%
CIS → CDMS			
Betaseron®**	3.8	4.2	10.53%
BSC**	2.1	2.4	14.29%
Time			
To reach EDSS3***	7.7	7.4	3.65%
To reach EDSS6***	14.9	15.4	2.67%

*Jacobs et al. [12]; **BENEFIT [14]; ***Weinshenker et al. [19]

[6] Pryse-Phillips W, Costello F. The epidemiology of multiple sclerosis. In Cook S, ed. *Handbook of Multiple Sclerosis*, New York, New York: Marcel Dekker, Inc., (2001).

[7] Poser C, Paty D, Scheinberg L, et al. New diagnostic criteria for multiple sclerosis: guidelines for research protocols. *Annals of Neurology* (1983); 13(3): 227–31.

[8] McDonald W, Compston A, Edan G, Goodkin D, Hartung H, Lublin F et al. Recommended diagnostic criteria for Multiple Sclerosis: guidelines from the international panel on the diagnosis of multiple sclerosis. *Annals of Neurology* (2001); 50(1): 121–7.

[9] Uitdehaag B, Kappos L, Bauer L, Freedman M, Miller D, Sandbrink R et al. Discrepencies in the interpretation of clinical symptoms and signs in the diagnosis of multiple sclerosis. A proposal for standardization. *Multiple Sclerosis* (2005); 11(2): 227–31.

[10] Health Canada. Notices of Compliance (NOC) Biologic Products for Human Use, January 1- December 31, 1998: Ottawa, Accessed 12-21 (2002), http://www.hc-sc.gc.ca/dhp-mps/alt_formats/hpfb-dgpsa/txt/prodpharma/bio98et_e.txt.

[11] Health Canada. Notice of Compliances, Biologics and Radiopharmaceuticals for Human Use, January - December, 1995: Ottawa, Ac-

cessed 2006-12-(2006), http://www.hc-sc.gc.ca/dhp-mps/alt_formats/hpfb-dgpsa/txt/prodpharma/bio95et_e.txt.

[12] Jacobs L, Beck R, Simon J, Kinkel R, Brownscheidle C, Murray T et al. Intramuscular Interferon Beta-1a Therapy Initiated During a First Demyelinating Event in Multiple Sclerosis. *The New England Journal of Medicine* (2000); 343(13): 898–904.

[13] Rudick R, Goodkin D, Jacobs L, Cookfair D, Herndon R, Richert J et al. Impact of interferon beta-1a on neurologic disability in relapsing multiple sclerosis. *Neurology* (1997); 49(2): 358–63.

[14] Kappos L, Polman C, Freedman M, Edan G, Hartung H, Miller D et al. Treatment with interferon beta-1b delays conversion to clinically definite and McDonald MS in patients with clinically isolated syndromes. *Neurology* (2006); 67 (7): 1242–9.

[15] IFNB. Interferon beta-1b is effective in relapsing-remitting multiple sclerosis: I. Clinical results of a multicenter, randomized, double-blind, placebo-controlled trial. *Neurology* (1993); 43(4): 655–61.

[16] IFNB. Interferon beta-1b in the treatment of multiple sclerosis: final outcome of the randomized controlled trial. *Neurology* (1995); 45(7): 1277–85.

[17] *TreeAge Software*. Inc. Williamstown, Massachusetts, (2007).

[18] Jacobs, LD, Beck, RW, Simon, JH, Kinkel, RP, Brownscheidle, CM, Murray TJ, and et al. Intramuscular interferon beta-1a therapy initiated during a first demyelinating event in multiple sclerosis. *Internal Study Report*, Tables 16.2.6.1 and 16.2.6.2. Cambridge, MA: Biogen Idec USA Inc., 2000.

[19] Weinshenker B, Bass B, Rice G, Noseworthy J, Carriere W, Baskerville J et al. The natural history of multiple sclerosis: A geographically based study. *I.Clinical course and disability*. Brain (1989); 112(1): 133–46.

[20] Weinshenker B, Rice G, Noseworthy J, Carriere W, Baskerville J, Ebers G. The natural history of multiple sclerosis: A geographically based study. *4. Applications to planning and interpretation of clinical therapeutic trials*. Brain (1991); 114(2): 1057-67.

[21] Kurtzke J. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). *Neurology* (1983); 33(11): 1444–52.

6th St.Petersburg Workshop on Simulation (2009) 903-907

Modelling of tensile strength of fiber and composite using MinMaxDM distribution family

Yuri Paramonov[1], Janis Andersons[2], Martinsh Kleinhofs[3]

Abstract

Generalization of extended family of weakest-link distributions with application to the composite specimen strength analysis is presented. Composite specimen for tensile strength is modeled as series system but every "link" of this system is modeled as parallel system. Results of enough successful attempts of using some specific distribution from this family for fitting of experimental dataset of strength of some carbon fiber reinforced specimens are presented.

1. Introduction

We consider a composite specimen for test of tensile strength as bundle of n_C longitudinal items (fibers or bundles) immersed into *composite matrix* (CM). We consider CM as composition of the matrix itself and all the layers with stackings different from the longitudinal one. We make very simplified assumption that only longitudinal items (LI) carry the longitudinal load but matrix only redistributes the loads after the failure of some longitudinal items. We part the composite into n_L parts of the same length l_1(approximately, this length can be interpreted as the interval inside of which the load of failed LI will be fully transmitted to the neighbor LI intact; the stronger the CM the smaller l_1). The total length of the composite specimens is equal to $l = n_L l_1$. We suppose that a development of process of fracture of specimen takes place in one or in several these parts ("links"). For the purpose of simplicity of subsequent text we call these links as "cross sections" (CS). So using this term we describe the composite as *series system of* CS. For description of the development of the fracture process of the series system it is appropriate to use the ideas, on which the extended weakest link distribution family, described in the authors papers [1-4], is based. Let the process of loading (i.e. the process of nominal stress (or mean load of one LI) increase in the specimen cross section) is described by an ascending (up to infinity) sequence $\{x_1, x_2, ..., x_t, ...\}$, and let $K_{Ci}(t)$, $0 \le K_{Ci} \le n_C$, is the number of failure of LI in ith CS with n_C initial number LI at the load x_t. Then the strength of ith CS

$$X_i^* = \max(x_t : n_C - K_{Ci}(t) > 0), \tag{1}$$

[1]Riga Technical University, E-mail: alexprm@svnets.lv
[2]Latvia University, E-mail: Janis.Andersons@pmi.lv
[3]Riga Technical University, E-mail: Martins.Kleinhofs@rtu.lv

but the ultimate strength of the specimen (which is the sequence of n_L CS) is

$$X = \min_{l \le i \le n_L} X_i^* = \min_{l \le i \le n_L} \max(x_t : n_C - K_{Ci}(t) > 0). \quad (2)$$

2. Model of failure of parallel system with redistribution of load after failure of some LI

Let $(X_1, ..., X_n)$ are $n = n_C - K_C$ random strengths of LI intact in some CS and X_j is jth order statistics in this CS. If there is uniform distribution of load between n LI, and load increases uninterruptedly, then the ultimate strength of this CS

$$X^* = \max_{l \le j \le n} X_j(n - j + 1)/n. \quad (3)$$

Daniels (1945,1983) studied the case when $K_C = 0$. In general case for random value K_C, (technological) failure number, we suppose existence of some a priori distribution $\pi_C = (\pi_1, \pi_2, ..., \pi_{n_C+1})$ (here $\pi_k = P(K_C = k - 1)$) . Then

$$F_{X^*}(x) = \pi_C \overrightarrow{F(x)}, \quad (4)$$

where vector column $\overrightarrow{F}(x) = (F_1(x), ..., F_{n_{C+1}}(x))'$, $F_k(x)$ is cdf of X^* if $n = n_C + 1 - k$, $k = 1, ..., n_C$; $F_{n_C+1}(x)$ is identical with unity (there are not any LI intact).

Much more reach spectrum of models of considered process can be developed using theory of Markov chains. We consider the process of accumulation of failures as an inhomogeneous finite Markov chain (MC) with finite state space $I = \{i_1, i_2, ..., i_{n_c+1}\}$. We say that MC is in state i if there are failures of $(i - 1)$ LI, $i = 1, ..., n_{C+1}$. State n_{C+1} is an absorbing state corresponding to the fracture of CS (fracture of all LI in this CS). The process of MC state change and the corresponding process $K_{C_i}(t)$ are described by transition probabilities matrix P. At the tth-step of MC matrix P is a function of t, $t = 1, 2, ...$. The cdf of strength of CS is defined on the sequence $\{x_1, x_2, ..., x_t, ...\}$ by equation

$$F_{X^*}(x_t) = \pi_C \big(\prod_{j=1}^{t} P(j) \big) u, \quad (5)$$

where $P(j)$ is the transition matrix for $t = j$, column vector $u = (0, ..., 0, 1)'$. We consider three main versions (hypotheses) of the matrix P structure. In first simplest version we assume that in one step of MC only failure of one LI can take place. And for corresponding matrix P_a we define $p_{ii} = 1 - F_C(x_t)$, where $F_C(x_t) = (F_0(x_t) - F_0(x_{t-1}))/(1 - F_0(x_{t-1}))$ is conditional cdf of strength of one LI under condition that the failure of it did not take place under load x_{t-1}, $F_0(x)$ is initial cdf of strength of one LI; $p_{i(i+1)} = 1 - p_{ii}$, $i = 1, ..., n_C$, $p_{(n_C+1)(n_C+1)} = 1$, but all the other p_{ij} are equal to zero. In the second version we assume that the number of failures in one step of MC has binomial distribution. Then for corresponding matrix P_b we have $p_{i(i+r)} = b(r; p, k)$, $p = F_C(x_t)$, $k = n_C + 1 - i$, $r = 0, ..., k$,

$i = 1, ..., n_C$; and again $p_{(n_C+1)(n_C+1)} = 1$, but all the other $p_{i,j}$ are equal to zero. The third version corresponds to minor cross crack growth. We suppose that the first failure appears in the boundary of CS and all following failures can appear only in the next LI. Let now j is ordered number of LI in some CS (and, for example, for the left-hand boundary LI $j = 1$). In this case it is easy enough to take into account the stress concentration next to the peak of the crack. Let the redistribution of CS load $x(t)$ between LI intact is defined by a "stress concentration" function $h(j; i, n_C)$. Then $p_{ij} = \prod_{k=i+1}^{j} F_C(x_{ik}(t)) \prod_{k=j+1}^{n_C+1}(1 - F_C(x_{ik}(t))$ for $j = i + 1, ..., n_C$; $p_{i(n_C+1)} = \prod_{k=i+1}^{n_C+1} F_C(x_{ik}(t))$; $p_{ii} = 1 - \sum_{k=i+1}^{n_C+1} p_{ik}$, $p_{ij} = 0$ for $j < i$, $i = 1, ..., n_C + 1$; where $x_{ij}(t) = h(j; i, n_C)x(t)n_C/(n_C + 1 - i)$ describes the stress in jth order LI after failure of ith order LI, $j = i + 1, ..., n_C + 1$.

3. Models of failure of series system with defected items

Let us denote by Z_i the ultimate strength of ith CS in which there is not any defect and we say it is Z-type CS. The strength of another type of CS we denote by Y_i and we say it is Y-type CS. And let random variable K_L, $0 \le K_L \le n_L$, denotes the number of Y-type CS. We suppose that the difference between the strength of two types of CS is defined only by the difference in a priori distribution of failure number before beginning of loading. Therefore the cdf of the strength of CS is defined for both CS types by equation (4) or (5) but for Z-type CS we use specific a priori distribution $\pi_C = (1, 0, ..., 0, 0)$, corresponding to absence of failure before beginning of loading. We suppose (see [4]) that the failure process of considered system has two stages. In first stage, the process develops along the specimen and K_L CS of Y-type can appear, $0 \le K_L \le n_L$. Then the second stage takes place: the process of accumulation of elementary damages in crosswise direction up to failure of both CS and specimen. We consider three levels of accuracy (minuteness) of description of the second stage and three corresponding probability models (probability structures). Level A: the development of fracture process takes place in every CS (containing or not some initial defects) and the strength of the weakest CS (weakest link) defines the strength of the specimen. Level AB: the strength of the CS without defects can be (relatively) so high and probability of its fructure before fructure of the defected CS is so small that independence of such probability on n_L can be assumed . Level B: development of fracture process takes place only in one, critical CS . Then only the probability that $K_L > 0$ depends on the number of CS , n_L, but the strength distribution of the critical CS does not depend on this number also (of course, the limit of satisfiability of the hypotheses AB and B should be find in special test; but it seems acceptable in considered in the paper numerical example). Correspondingly we have three probability structures.

$$A: \quad X = \min(Y_1, ..., Y_{K_L}, Z_1, ..., Z_{n_L - K_L});$$

$$AB: \qquad X = \begin{cases} \min(Y_1, ..., Y_{K_L}, Z) & K_L > 0, \\ Z, & K_L = 0; \end{cases} \qquad B: \qquad X = \begin{cases} Y, & K_L > 0, \\ Z, & K_L = 0. \end{cases}$$

Two different versions of the first stage can be considered also. First version: (technological) defects appear before the loading and their number does not depend on the subsequent loading. Second version: defects appear during loading (instantly or gradually) and their number depends on the load. **For "instant fructure" version** for structures A, AB, B we have correspondingly

$$F(x) = 1 - (1 - F_Z(x))^{n_L} \sum_{k=0}^{n_L} p_k \delta^k, \;\; \delta(x) = (1 - F_Y(x))/(1 - F_Z(x)). \quad (6)$$

$$F(x) = 1 - (1 - F_Z(x)) \sum_{k=0}^{n_L} p_k (1 - F_Y(x))^k, \quad (7)$$

$$F(x) = p_Y F_Y(x) + (1 - p_Y) F_Z(x), \quad (8)$$

where $p_k = b(k; p_L, n_L)$ is binomial probability mass function (pmf), $p_Y = 1 - p_0^{n_L}$. For description of random number of cross section of Y-type, K_L, binomial or Poisson pmf can be used. In last case equations (6, 7) (approximately, if n_L is large enough) can be written in the following way

$$F(x) = 1 - (1 - F_Z(x))^{n_L} exp(-\lambda(1 - \delta(x)) \quad (9)$$

$$F(x) = 1 - (1 - F_Z(x)) exp(-\lambda F_Y(x)) \quad (10)$$

where $\lambda = n_L p_L$ or it is just independent parameter of Poisson pmf. If initiation of the defects depends on the applied load, then it can be assumed that $p_L = F_K(x)$, where $F_K(x)$ is some cdf of defect initiation load.

For the case when defects appear during loading the process of accumulation of defects along the chain of n_L CS again can be considered as Markov chain (MC). In this case MC is in state i if there are $(i - 1)$ of Y-type CS, $i = 1, ..., n_L + 1$. State i_{n_L+2} is an absorbing state corresponding to the fracture of specimen. The initial distribution is represented by some row vector $\pi_L = (\pi_{L1}, \pi_{L2}, ..., \pi_{L,n+1}, \pi_{L,n+2})$. In the new approach the number of CS of Y-type and the strength of specimens are random functions of time, $K_L(t)$ and $X(t)$. Now three corresponding probability structures we denote by MA, MAB and MB. For example for the MA we have $X(t) = \min(Y_1, Y_2, ..., Y_{K_L(t)}, Z_1, Z_2, ..., Z_{n_L - K_L(t)})$. In similar way $X(t)$ is defined for another structures. Ultimate strength of specimen and cdf of it are defined again by equations (2) and (5). If there is only one LI ($n_C = 1$) we should talk about defected links with specific $F_Y(x)$ instead of Y-type CS. Different assumptions about the $F_Y(x)$, $F_Z(x)$ and examples of formulae for calculation of elements of matrix $P(j)$ are given in [4]. In this paper we assume that $F_Y(x)$ and $F_Z(x)$ are cdf of CS strength of Y-type or Z-type correspondingly. In the following numerical example we suppose that logarithm of strength of one LI (in one CS) without defect has smallest extreme value (sev) distribution: $F_0(x) = 1 - exp(-exp((x - \theta_{0Z1})/\theta_{1Z1})$. So we use the logarithm scale and in this case the cdf of specimen strength also has location and scale parameters θ_0 and θ_1: $F_X(x) = F_{X}^{0}((x - \theta_0)/\theta_1))$.

4. MinMaxDM distribution family

Different assumptions about the distribution of strength of strand (fiber) within the limits of one "link", a priori distribution of initial (technological) defects, influence of length and width of specimens compose a family of the distributions of ultimate composite tensile strength. Taking into account (2) and (3) we denote this family by abbreviation MinMaxD (in memory of Daniels) if the cdf $F_{X^*}(x)$ is defined by equation (4) and by abbreviation MinMaxM (because of connection with Markov chain theory), if it is defined by equation (5), and for unified family we offer to use abbreviation MinMaxDM.

5. Processing of test data

In [5] there are the test results of both 64 carbon fiber strands with length 20 mm ($data_1$) and the same number of stripes of 10 strands of the same length ($data_2$). We try to get description of $data_2$ using results of processing $data_1$. Let x_i be ith order statistic, $i = 1, 2, ..., n$, n is the sample size; $E(X_i)$ is the expected value of ith order statistic, $E\binom{0}{X_i}$ is the same but for $\theta_0 = 0$ and $\theta_1 = 1$. Then for estimation of θ_0 and θ_1, if all the other parameters are fixed, we have the following linear regression model: $E(X_i) = \theta_0 + \theta_1 E\binom{0}{X_i}$. We make fitting of the $data_1$ and get linear regression parameter estimates $\hat{\theta}_0 = 6.554$ and $\hat{\theta}_1 = 0.1243$ assuming that sev distribution takes place (here X is logarithm of strength).Then we make fitting of the $data_2(+)$ again assuming the sev distribution (Fig. 1a). In Fig. 1b we see fitting of the same $data_2$ using $E\binom{0}{X_i}$ of cdf corresponding to MinMaxMa.sev-B structure model (for Pa type of matrix P, $F_0(x)$ is sev distribution, B structure (see equation (8)); $n_C = 5$; π_C is binomial a priori distribution of K_C with $p_C = 0.01$; $p_Y = 0.9048$). "Regression prediction" $(*)$, $\hat{x}_i = \hat{\theta}_0 + \hat{\theta}_1 E\binom{0}{X_i}$, using estimates $\hat{\theta}_0$ and $\hat{\theta}_1$ obtained processing $data_1$ is shown also (but we increase $\hat{\theta}_1$ up to 0.2912 taking into account variation of Young's modulus ($Var(E) = 0.03$)). The statistic $OSPPt = (\sum_{i=1}^{n}(x_i - \hat{x}_i)^2 / \sum_{i=1}^{n}(x_i - (\sum_{i=1}^{n} x_i)/n)^2)^{1/2}$ [4], as the measure of fitting for Fig.1a is equal to 0.267 (for sev distribution) and as the measure of fitting and prediction quality for Fig. 1b (for MinMaxMa.sev-B structure model) is equal to 0.161 and 0.192 correspondingly.

6. Conclusions

We see that MinMaxMa.sev-B structure model provides better fitting of results of tensile strength test of carbon fiber stripe of 10 strands but only if we assume that in CS there are only 5 strands instead of 10 and taking into account variation of Young's modulus. It seems, that MinMaxDM distribution family deserves to be studied much more thoroughly using much more test data.

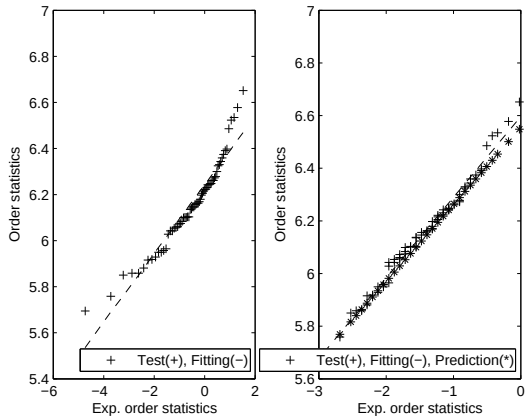


Figure 1: Fitting of results (+) of tensile strength test of carbon fiber stripe of 10 threads (see explanation in text).

References

[1] Paramonov Yu., Andersons J. (2006) *A new model family for the strength distribution of fibers in relation to their length.* Mechanics of Composite Material, **42(2)**, 179–192.

[2] Paramonov Yu., Andersons J. (2007) *Modified weakest link family for tensile strength distribution.*. Proceeding of Fifth International Conference on Mathematical Methods in Reliability Methodology and Practice (MMR 2007), 1-4 July, Glasgow, UK. - 8 pp.

[3] Paramonov Yu. (2008) *Extended weakest link distribution family and analysis of fiber strength dependence on length.* Composites: Part A, 39950-955.

[4] Paramonov Yu., and Andersons J.(2008) *Analysis of fiber strength dependence on its length by weakest-link approach.Part 1. Weakest link distribution family.* Mechanics of Composite Material, **44(5)**, 479–486.

[5] Kleinhofs M.(1983) *Investigation of static strength and fatigue of composite material used in aircraft structure.* Candidate degree thesis, Riga.

6th St.Petersburg Workshop on Simulation (2009) 909-913

An efficient Multilevel Splitting scheme[1]

Denis Miretskiy[2], Werner Scheinhardt[3], Michel Mandjes[4]

1. Introduction

Rare event analysis has been attracting continuous and growing attention over the past decades. It has many possible applications in different areas, e.g., queueing theory, insurance, engineering etc. As explicit expressions are hard to obtain, and asymptotic approximations often lack error bounds, one often applies simulation methods to obtain performance measures of interest.

Obviously, the use of standard Monte Carlo simulation for estimating rare event probabilities has an inherent problem: it is extremely time consuming to obtain reliable estimates since the number of samples needed to obtain an estimate of a certain predefined accuracy is inversely proportional to the probability of interest. Two important techniques to speed up simulations are *Importance Sampling (IS)* and *Multilevel Splitting (MS)*.

IS prescribes to simulate the system under a *new* probability measure such that the event of interest occurs more frequently, and corrects the simulation output by means of likelihood ratios to retain unbiasedness. The likelihood ratios essentially capture the likelihood of the realization under the old measure with respect to the new measure. The choice of a 'good' new measure is rather delicate; in fact only measures that are *asymptotically efficient* are worthwile to consider. We refer to [3] for more background on IS and its pitfalls.

The other technique, multilevel splitting (MS), is conceptually easier, in the sense that one can simulate under the normal probability measure. When a sample path of the process is simulated, this is viewed as the path of a 'particle'. When the particle approaches the target set to a certain distance, the particle splits into a number of new particles, each of which is then simulated independently of each other and of the past. This process may repeat itself several times, hence the term *multilevel* splitting. Typically, the states where particles should be split are determined by selecting a number of level sets of an *importance function* f. Every time a particle (sample path) crosses the next level set of the importance function f, it is split. The splitting factor (i.e. the number of particles that replaces the original particle) may depend on the current level.

[1]Part of this research has been funded by the Dutch BSIK/BRICKS project.

[2]University of Twente, E-mail: d.miretskiy@math.utwente.nl

[3]University of Twente, E-mail: w.r.w.scheinhardt@math.utwente.nl

[4]University of Amsterdam, E-mail: mmandjes@science.uva.nl

The challenge in MS is to choose an importance function that will ensure that the probability of reaching the target set is roughly the same for all states that belong to the same level. Moreover, choosing the splitting factors appropriately is also important. Sample paths will hardly ever end up in the rare set if this factor is too small, while the number of particles (and consequently the simulation effort) will grow fast if this factor is too large. For an overview of the MS method see [5].

There are not many examples of *asymptotically efficient* MS schemes for estimating general types of rare events in the present literature. Most articles deal either with effective heuristics for particular (queueing) models, usually providing good estimates without rigorous analysis, see e.g. [6]; or with restrictive models, see e.g. [2]. The recent work in [1] does enable one to construct an asymptotically efficient MS scheme for estimating the probability of first entrance to a rare set, when the decay rate of the probability is known for all starting states. The authors used control-theoretic techniques to derive and prove their results.

In this work we also provide a simple and asymptotically efficient MS scheme for estimating the probability of first entrance to some rare set. The scheme can be seen as part of the class of asymptotically efficient MS schemes developed in [1]. However, since we are only interested in easy-to-implement (but still efficient) schemes, we use a fixed, pre-specified splitting factor R, to be used for all levels. This is in contrast to the setting in [1] where the splitting factor may vary between levels and is usually noninteger (which is then implemented by using a randomization procedure). We accompany the scheme with a proof of its asymptotic efficiency which is relatively easy, in the sense that it only uses probabilistic arguments and some simple bounds, thereby giving insight into why the scheme works so well.

The rest of the paper's structure is as follows. In Section 2 we first describe the model of interest and, after a brief review of the MS method, we provide the MS scheme itself. A sketch of the proof of asymptotic efficiency of the scheme is given in Section 3. Supporting numerical results for a two-node tandem model are presented in Section 4 and compared with results from IS on the same model; in fact it turns out that MS can be a good alternative to IS for certain parameter settings.

2. Model and Preliminaries

2.1. Model

We consider some Markov process $\{Q_k\}$ that lives in a domain D^B and has a finite number of possible jump directions v_i with corresponding transition probabilities ν_i. Although this is not essential, we will assume $\{Q_k\}$ to be a random walk for ease of exposition. We are interested in the probability that $\{Q_k\}$ hits the (rare) target set T^B before the 'tabu' set A^B, starting from some state $s \notin T^B \cup A^B$.

To clarify the situation we provide a simple queueing example, in which $\{Q_k\}$ is the joint-queue length after the k-th transition of the Markov chain that describes a two-node Jackson tandem network. Then we may be interested in the event where, starting from some state, the queue of the second node reaches a level B before

the entire system becomes empty. Then obviously, B is the 'rarity parameter' (in the sense that the event becomes more rare as we choose larger values for B), and we have $D^B = \mathbb{R}^2_+$; $T^B = \{x \in D : x_2 \geq B\}$ and $A^B = (0,0)$.

It is convenient to scale the process $\{Q_k\}$ with the parameter B. The scaled process $X_k = Q_k/B$ then makes jumps of size v_i/B, and lives in the domain D, which is the scaled version of D^B. The target and tabu sets T^B and A^B are scaled in the same manner, their scaled versions being given by T and A.

For such (disjoint) sets A and T and some state $s \in D$, such that $s \notin A \cup T$, we define the stopping time

$$\tau_B^s = \inf\{k > 0 : X_k \in T,\ X_j \notin A\ \forall j = 1, \ldots, k-1,\ X_0 = s\},$$

where $\tau_B^s = \infty$ if $\{X_k\}$ hits the set A before T. The probability of interest is now as follows:

$$p_B^s = \mathbb{P}\left(X_{\tau_B^s} < \infty\right). \tag{1}$$

Importantly, we will assume that this probability decays exponentially in B, with decay rate

$$-\lim_{B\to\infty} B^{-1} \log p_B^s = \gamma(s).$$

In fact we will even assume that this convergence is uniform in s:

Assumption 1. For any $\epsilon > 0$, some $B^* > 0$ exists such that for all $s \notin A \cup T$ we have $B^{-1} \log p_B^s + \gamma(s) < \epsilon$ for $B > B^*$.

2.2. Multilevel Splitting

To apply MS, one first needs to define a family of nested sets $\{L_k\}$, $k = 0, \ldots, m$ such that $s \in \ell_0 = \partial L_0$ and

$$T = L_m \subset L_{m-1} \subset \ldots \subset L_1 \subset L_0 \subset D.$$

This family $\{L_k\}$ should be chosen such that every state that belongs to the boundary of L_k has similar importance, i.e., the probability of reaching T before A should be approximately equal for every state $x \in \ell_k = \partial L_k$. We will require

$$-\lim_{B\to\infty} B^{-1} \log p_B^s = c_k, \quad \forall s \in \ell_k, \quad k = 0, \ldots, m,$$

where the c_k are positive constants. Given this family, we start at the initial state s (which belongs to ℓ_0) with exactly R_0 particles. We continue to simulate each of them until they either cross level ℓ_1 or hit the tabu set A. All particles that end up in A are to be terminated without any replacement. Every particle that reaches level ℓ_1 is to be replaced by R_1 independent replicas. We continue to simulate all the (new) particles until they cross the next level ℓ_2 or hit the tabu set A, and so on. At stage k we start with some number of particles in level ℓ_{k-1} and simulate them until they reach ℓ_k or A. Then each particle that crossed ℓ_k is replaced by R_k independent copies, while all particles in A are terminated. We

stop the procedure when the m-th level (i.e., the target set T) is reached. Now we construct the estimator as follows:

$$\hat{p}_B = \frac{X}{R_0 \cdot R_1 \cdot \ldots \cdot R_{m-1}}, \tag{2}$$

where X is the number of particles that eventually reaches the target set T before the tabu set A. The estimate of p_B^s is constructed by averaging a number of independent replications of $\hat{p}_B$.

We now describe the Multilevel Splitting scheme we propose:

1. Choose some integer R to be the splitting factor for all levels.
2. Compute the number of levels $n_B := \lfloor B\gamma(s)/\log R \rfloor$.
3. Define levels $\ell_k := \left(x \in D : \gamma(s) - \gamma(x) = \frac{k}{B} \log R \right)$, $k = 0, \ldots, n_B$.
4. Define $R' := \lfloor e^{B\gamma(s) - n_B \log R} \rfloor$, to be used as splitting factor at level n_B only.

(3)

The idea of the scheme is as follows: different states x in the same level have the same decay rate for their corresponding probabilities p_B^x, and the different levels are defined such that the total decay rate $\gamma(s)$ is 'evenly spread'; in other words, the distances between consecutive levels are equal in terms of decay rate. The corresponding probability of reaching the next level is roughly equal to $1/R$ due to the choice of n_B in step 2, so that on average only one particle out of R will reach the next level. Finally, since level n_B is in general not the boundary of the target set T (due to the rounding in step 2), and the probability to reach T from this level is larger than $1/R$, we can do with the lower splitting factor R' at level n_B.

3. Asymptotic Efficiency

In this section we provide a sketch of the proof of asymptotic efficiency of our MS scheme; we will call an estimator asymptotically efficient if

$$\limsup_{B \to \infty} B^{-1} \log \left(w(B) \mathbb{E} \hat{p}_B^2 \right) \leq -2\gamma(s), \tag{4}$$

where $w(B)$ represents the expected computational effort per replication of $\hat{p}_B$. For the specific form of $w(B)$ we can make various choices. Here we assume that the required time effort increases linearly in the starting level. That is, we assume it takes $k+1$ time units to simulate a sample path of a particle starting from level k, since with high probability it will reach A before ℓ_{k+1}; see also [2] for the motivation of this choice.

In order to simplify notation we omit the dependence on B in the notation n_B for the number of levels. Also we rewrite the estimator in (2) as follows:

$$\hat{p}_B = \frac{1}{R^n R'} \sum_{i=1}^{R^n R'} I_i. \tag{5}$$

Here we used that we have the same splitting factor R at each level, except the last one which is R', and the I_i are indicator random variables for each of the $R^n R'$ *possible* particles that may be simulated; $I_i = 1$ if the i-th particle hits the target set T before the tabu set A, and $I_i = 0$ otherwise. At first sight, it may seem that the number of particles needed to obtain this estimator grows exponentially in n, and consequently in B. However this is not the case, since we only need to simulate a few of all possible $R^n R'$ particles till the end. Suppose for instance that from the initial R particles only one reaches ℓ_1 before A, then the maximum number of possible particles to be simulated further is already reduced from $R^n R'$ to $R^{n-1} R'$.

In order to prove that (4) holds for our scheme, we first analyze the second moment of the estimator, for whichwe have:

$$B^{-1} \log \mathbb{E}\hat{p}_B^2 = B^{-1} \log \left(\frac{1}{R^{2n} R'^2} \right) + B^{-1} \log \mathbb{E} \left[\sum_{i=1}^{R^n R'} I_i \right]^2 . \tag{6}$$

It is not difficult to see that the first term in the right-hand side of (6) converges to $-2\gamma(s)$ as B grows to ∞, thanks to line 2 in (3). By applying some combinatorial methods and Assumption 1, we can show that the last term in (6) converges to zero when B grows to ∞, leading to

$$\lim_{B \to \infty} B^{-1} \log \mathbb{E}\hat{p}_B^2 \leq -2\gamma(s). \tag{7}$$

Also for the expected computational effort our analysis can be based on some combinatorics and Assumption 1, which leads to

$$\lim_{B \to \infty} B^{-1} \log w(B) = 0. \tag{8}$$

Combining the statements in (7) and (8) now immediately leads to the main result:

Theorem 1. *The Multilevel Splitting algorithm (3) is asymptotically efficient.*

4. Numerical Results

In this section we illustrate the efficiency of the MS scheme by applying it to a two-node tandem Jackson network; we consider the rare event in which the second queue collects some large number of jobs B before the entire system empties. We provide some estimates for the corresponding probability p_B^s using our MS scheme (3) and compare its performance with that of the (also asymptotically efficient) IS scheme developed in [4]. There, we always performed a fixed number of 10^6 simulation runs, while the relative error and the actual simulation time (in seconds) were important indicators of the efficiency of the scheme. Here we use the computer time from [4] as a time budget for the current MS scheme in order to make a fair comparison.

Table 1: Simulation results

(λ, μ_1, μ_2)	s	B	p_B^s	RE	$RE(IS)$	time
$(0.3, 0.36, 0.34)$	$(0,0)$	20	$5.98 \cdot 10^{-2} \pm 2.57 \cdot 10^{-4}$	$2.19 \cdot 10^{-3}$	$3.12 \cdot 10^{-3}$	28
		50	$1.52 \cdot 10^{-3} \pm 1.72 \cdot 10^{-5}$	$5.77 \cdot 10^{-3}$	$3.94 \cdot 10^{-3}$	80
		100	$2.91 \cdot 10^{-6} \pm 5.80 \cdot 10^{-8}$	$10.1 \cdot 10^{-3}$	$4.74 \cdot 10^{-3}$	168
$(0.1, 0.55, 0.35)$	$(0.6B, 0)$	20	$1.99 \cdot 10^{-5} \pm 4.09 \cdot 10^{-7}$	$1.04 \cdot 10^{-2}$	$1.32 \cdot 10^{-3}$	7
		50	$3.19 \cdot 10^{-12} \pm 2.09 \cdot 10^{-13}$	$5.10 \cdot 10^{-2}$	$1.58 \cdot 10^{-3}$	18
		100	$1.87 \cdot 10^{-23} \pm 3.54 \cdot 10^{-24}$	$9.65 \cdot 10^{-2}$	$1.81 \cdot 10^{-3}$	35
$(0.3, 0.33, 0.37)$	$(0,0)$	20	$3.29 \cdot 10^{-2} \pm 2.59 \cdot 10^{-4}$	$4.01 \cdot 10^{-3}$	$3.79 \cdot 10^{-2}$	28
		50	$7.00 \cdot 10^{-5} \pm 2.07 \cdot 10^{-6}$	$1.50 \cdot 10^{-2}$	$7.90 \cdot 10^{-2}$	84
		100	$1.92 \cdot 10^{-9} \pm 1.29 \cdot 10^{-10}$	$3.42 \cdot 10^{-2}$	$13.4 \cdot 10^{-2}$	189

In Table 1 we present estimates of p_B^s for different starting states s and parameter settings, accompanied by their 95% confidence intervals and relative errors, as well as the relative errors obtained using the IS scheme from [4].

Clearly, the MS scheme (3) gives good results. In fact the relative error is lower than that of the IS scheme when the parameters λ, μ_1, μ_2 are close to each other. Indeed, it was known that IS performs relatively poorly in such scenarios, and it is interesting to see that MS provides a good alternative. On the other hand, when the parameters are not close to each other, MS is outperformed by IS. This may be understood from the fact that simulating under the normal measure (as is done in MS) is difficult for such cases, since the number of jobs in the second queue has a strong downward drift.

References

[1] T. Dean and P. Dupuis. Splitting for rare event simulation: a large deviations approach to design and analysis. *Preprint*, 2008.

[2] P. Glasserman, P. Heidelberger, P. Shahabuddin, and T. Zajic. Multilevel splitting for estimating rare event probabilities. *IBM Research Report*, RC 20478, 1996.

[3] P. Heidelberger. Fast simulation of rare events in queueing and reliability models. *ACM Transactions on Modeling and Computer Simulation*, 5(1):43–85, 1995.

[4] D.I. Miretskiy, W.R.W. Scheinhardt, and M.R.H. Mandjes. Rare-event simulation for tandem queues: a simple and efficient importance sampling scheme. *Preprint*, 2008.

[5] P. Shahabuddin. Rare event simulation in stochastic models. In *Proceedings of the 27th conference on Winter simulation*, 178–185. IEEE Computer Society, 1995.

[6] M. Villén-Altamirano and J. Villén-Altamirano. On the efficiency of RESTART for multidimensional state systems. *ACM Transactions on Modeling and Computer Simulation*, 16(3):251–279, 2006.

6th St.Petersburg Workshop on Simulation (2009) 915-919

Simulation-based multi-craft workforce scheduling

Hesham K. Alfares[1]

Abstract

A simulation model is used for stochastic days-off scheduling of maintenance crews. In order to determine optimum employee work schedules, the model considers limited employee availability, stochastic workload demand, and labor scheduling regulations. The stochastic simulation model was implemented for actual days-off scheduling of a multi-craft pipeline maintenance workforce in a large oil company. Alternative employee days-off schedules generated by the model are expected to improve the productivity of the existing maintenance workforce by an average of 25%.

1. Introduction and Literature Review

A simulation model is presented for stochastic scheduling of a multi-craft maintenance workforce of an oil and gas pipelines department. The objective is to minimize the average throughput (waiting plus processing) time of maintenance work orders. The pipelines maintenance crews are composed of 19 technicians belonging to five different maintenance crafts. Current labor regulations allow only three possible days-off schedules for the maintenance workforce. The workload for each craft is stochastic because the majority of maintenance jobs are unscheduled and require multi-craft crews.

A simulation model was constructed to represent and analyze the maintenance work order and workforce scheduling system. This model was used to evaluate several days-off scheduling alternatives for the pipelines maintenance workforce. The model suggested alternative schedules for the five maintenance crafts that are expected to reduce the throughput time on average by 25%.

There is a lot of literature on employee scheduling, but here we focus on simulation-based approaches for employee scheduling. Such approaches have initially been applied in manufacturing facilities. Davis and Mabert (2000) use simulation to evaluate and compare two mathematical modeling techniques for order dispatching and labor assignment decisions in two alternative cellular manufacturing (CM) arrangements. Yang et al. (2002) use simulation to study the impact of several flexible workday policies for maximizing the flexibility and responsiveness of a job shop by adjusting the length of workdays. Zülch et al. (2004) employ the

[1]Systems Engineering Department, King Fahd University of Petroleum & Minerals, Dhahran - 31261, Saudi Arabia, E-mail: hesham@ccse.kfupm.edu.sa

personnel-oriented simulation tool $ESPE$ to evaluate three techniques for planning and re-assigning personnel in manufacturing.

Simulation-based workforce scheduling has also been utilized in service facilities. Smith et al. (2002) use simulation with integer programming (IP) to staff geographically distributed service facilities. Chong et al. (2003) apply simulation to model an airline's flight and staff schedules and to generate staff rosters for check-in agents. Guttkuhn et al. (2003) use simulation to assign train crews in order to meet train traffic and labor rules. Gupta et al. (2003) combine simulation with optimization to schedule the aircraft line maintenance employees of Continental Airlines. Bazargan and McGrath (2003) use a similar approach to allocate maintenance mechanics to various shifts. Li and Li (2000) combine simulation and goal programming to investigate the costs and benefits of staff flexibility in a Chinese clinic. Centeno et al. (2003) integrate simulation with an IP model to determine staffing requirements and optimal schedules for ER staff.

2. The current maintenance work order process

The pipelines maintenance workforce is responsible for scheduled and emergency repairs and maintenance of all pipelines throughout a designated area. It consists of 19 employees divided into five crafts: 2 air conditioning (AC) technicians, 6 digital (DG) technicians, 5 electrical (EL) technicians, 3 machinist (MA) technicians, and 3 metal (ME) technicians. The maintenance employees can work for a maximum of 12 hours per day, and they can be assigned to only three types of days-off schedules:

a) The (5/2) schedule: 5 consecutive workdays followed by 2 consecutive off days (weekend) per week. Initially, 13 technicians (1 AC, 6 DG, 3 EL, 1 MA, and 2 ME) are on this schedule.

b) The (14/7) schedule: 14 consecutive workdays followed by 7 consecutive off days per three-week cycle. Initially, four technicians (1 AC, 1 EL, and 2 MA) are on this schedule.

c) The (7/3-7/4) schedule: two work stretches each of 7 consecutive workdays separated by two breaks of 3 and 4 consecutive off days, per three-week cycle. Initially, only two technicians (1 EL and 1 ME) are on this schedule.

After a maintenance work order (W/O) is initiated, materials needed as well as labor requirements of each craft type are listed. Next, the maintenance cost is estimated and the originator's approval is obtained. Subsequently, the W/O is prioritized and scheduled and then work is started. After finishing the work, a report is sent to the originator for either comment or approval in order to close the W/O. A simplified flowchart of the system is shown in Figure 1. From the preceding description, each W/O must pass through the following phases:

a) Hold (HLD) phase: waiting to receive the materials or approval to start.

b) Work (WRK) phase: being processed and undergoing maintenance work.

c) Finish (FIN) phase: completed, but waiting for approval to close.

d) Close (CLS) phase: completed, approved, and entered into the database.

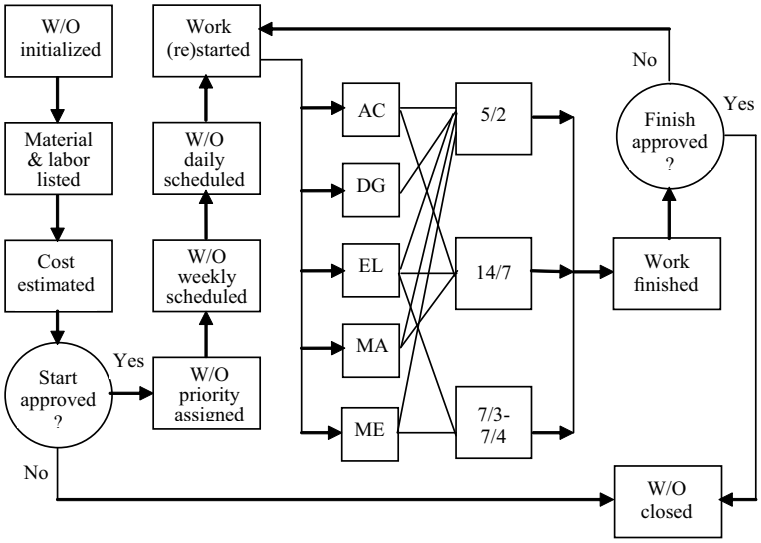


Figure 1: Simplified flowchart of the maintenance W/O process.

Data covering a period of 7 months was collected and analyzed. For each W/Os during the given period, data were collected on durations and inter-arrival times in hours, and also the required number of employees of each craft. To fit probability distributions, the data was plotted and the relevant statistics were calculated, and then the Chi-square goodness-of-fit test was used with $\alpha = 0.05$. The probability distribution for the W/O inter-arrival time was found to be EXPON (9.79). For each craft, Table 1 shows the fitted probability distributions for the service times (time spent on each work order per technician) and other relevant statistics. For each W/O, the number of men required of each craft has a discrete empirical distribution, and it must be fully available for starting the given W/O.

3. Simulation model

The simulation model's assumptions include the following:

a) Any W/O requiring several crafts is processed by each craft in parallel, and consequently considered as multiple W/Os, each requiring a single craft type.

b) The number of technicians from each craft type assigned to each W/O is a random variable calculated from empirical probability distributions.

c) During their work times, the pace of work of maintenance craft employees is represented by the service time probability distributions given in Table 1.

Table 1: Service time distributions and statistics for the five craft types

Craft type	No. of men available	Avg. no. of men needed	Probability of craft (%)	Service time probability distribution (hours)
AC	2	1.3	18.5	WEIBL (0.83, 19.03)
DG	6	1.49	24.9	WEIBL (0.76, 22.77)
EL	5	1.59	21.3	LOGNR (2.48, 1.07)
MA	3	1.35	23.1	EXPON (13.78)
ME	3	1.11	12.2	WEIBL (0.89, 37.92)

The AweSim! simulation software was used. The program was run for 210 simulated days (7 work months), which is well into steady state. To validate the model, actual and simulation output values of W/O throughput times in hours were compared. For each craft, confidence intervals were constructed of the differences between the averages of five simulation runs and five randomly chosen actual system observations. Since the confidence intervals of the difference (error) for all five crafts contain zero, the simulation model can be accepted as a valid representation of the real system. The number of replications was determined to be 10, which is the sample size that gives the smallest confidence interval for the average throughput time.

For each craft type, the simulation output gives information about the number of completed W/Os, average waiting time of W/Os, average number of orders waiting to be served, and average utilization of employees. However, the main performance measure is the average W/O throughput time for each craft, which is illustrated in Table 2.

Table 2: Current W/O throughput time for each craft in hours

Craft type	Average	Stand. dev.	95% Confidence interval
AC	9.07	2.24	[7.77, 10.37]
DG	16.85	1.70	[15.86, 17.84]
EL	7.47	1.43	[6.64, 8.3]
MA	24.43	3.29	[22.53, 26.33]
ME	9.14	2.65	[7.61, 10.68]

3.1. Alternative MA days-off schedules

As Table 2 shows that the throughput time is highest for the machinist (MA) craft, alternative days-off schedules were first tried for the MA technicians. Keeping the number of MA technicians as 3, but considering all their possible assignments

to the 3 feasible days-off schedules, there are 10 possible scheduling scenarios (alternatives) shown in Table 3 (alternative 10 is the current schedule). For each scenario, the model was modified according to that scenario and then run again for 10 replications. The new average throughput time for the MA craft W/Os under each scenario is shown in Table 3.

Table 3: Days-off scheduling alternatives for 3 MA technicians

Alternative	No. assigned to days-off schedules			Ave. throughput
number	5/2	14/7	7/3-7/4	time (hrs)
1	3			30.68
2		3		43.58
3	2	1		23.14
4	1	1	1	20.41
5	2		1	23.6
6		2	1	24.54
7		1	2	24.39
8	1		2	24.33
9			3	37.22
10	1	2		24.43

As can be seen from Table 3, scenario number 4 (1 man on each of the 3 days-off schedules) is the best. Under this scenario, the average throughput time for MA work orders will be reduced by 16.4% from 24.43 hours to 20.41 hours. The hiring cost will not change since the pay is the same for all 3 days-off schedules.

3.2. Alternative schedules for other crafts

Following the same procedure, the best days-off schedules were determined for AC, DG, EL, and ME technicians. Since the work orders are processed in parallel, the inter-dependence among the different crafts is minimal. Therefore, the different scenarios (days-off schedules) for each craft were run while fixing the other crafts at their current schedules. The most efficient days-off schedules and corresponding reductions in work order throughput times for all crafts are summarized in Table 4. The reductions in throughput times range from 4% to 67%, with an average of 25%.

4. Conclusions

A simulation model for stochastic workforce days-off scheduling has been presented. This approach has been applied to real-life days-off scheduling of a pipelines maintenance workforce consisting of five types of crafts. The simulation model determined the optimum allocation of technicians of each craft type to three applicable days-off schedules. A 25% increase in productivity is expected due to reduction

Table 4: Summary table of best days-off schedules for all crafts

Craft	No. of schedules	5/2	14/7	7/3-7/4	From(hr)	To(hr)	%Reduction
AC	6	1		1	9.07	8.71	4
DG	27	1	2	3	16.85	5.57	66.9
EL	21	2	2	1	7.47	7.44	0.39
MA	10	1	1	1	24.43	20.41	16.4
ME	10	1	1	1	9.14	5.59	38.8

in average work order throughput times. This improvement can be obtained simply by changing the employee scheduling assignments, without increasing either the size or the cost of the maintenance workforce.

Acknowledgment: The author thanks King Fahd University of Petroleum & Minerals for the support.

References

a) D.J. Davis, V.A. Mabert, Order dispatching and labour assignment in cellular manufacturing systems, *Decision Sciences*, *31*(4), 2000, 745-771.

b) K.K. Yang, S. Webster, R.A. Ruben, An evaluation of flexible workday policies in job shops, *Decision Sciences*, *33*(2), 2002, 223-249.

c) G. Zlch, S. Rottinger, & T. Vollstedt, A simulation approach for planning and re-assigning personnel in manufacturing, *International Journal of Production Economics*, *90*(2), 2004, 265-277.

d) N. Li, L.X. Li, Modelling staffing flexibility: A case of China, *European Journal of Operational Research*, *124*(2), 2000, 255-266.

e) M.A. Centeno, R. Giachetti, R. Linn, A.M. Ismail, A simulation-ILP based tool for scheduling ER staff, *Proceedings of the 2003 Winter Simulation Conference*, *2*, New Orleans, LA, 2003, 1930-1938.

f) L.D. Smith, Staffing geographically distributed service facilities with itinerant personnel, *Computers & Operations Research*, *29*(14), 2002, 2023-2041.

g) K-L Chong, M. Grewal, J. Loo, S.L. Oh, A simulation-enabled DSS for allocating check-in agents, *INFOR Journal*, *41*(3), 2003, 259-273.

h) R. Guttkuhn, T. Dawson, & U. Trutschel, A discrete event simulation for the crew assignment process in North American freight railroads, *Proceedings of 2003 Winter Simulation Conference*, *2*, New Orleans, LA, 1686-1692.

i) P. Gupta, M. Bazargan, & R.N. McGrath, Simulation model for aircraft line maintenance planning, *Proceedings of the 2003 Annual Reliability & Maintainability Symposium*, Tampa, FL, 2003, 387-391.

j) M. Bazargan, R.N. McGrath, Discrete event simulation to improve aircraft availability and maintainability, *Proceedings of the 2003 Annual Reliability and Maintainability Symposium*, Tampa, FL, 2003, 63-67.

6th St.Petersburg Workshop on Simulation (2009) 922-926

On two new Monte-Carlo methods for solving nonlinear equations[1]

Sergej M. Ermakov[2], Konstantin A. Timofeev[3]

Abstract

A new Monte-Carlo method for solving systems of algebraic equations with quadratic nonlinearity is suggested. It is proved to be more efficient than the Newton's method in several cases. It is shown that mutual comparative efficiency of deterministic methods does not guarantee the efficiency of corresponding stochastic analogues.

A new Monte-Carlo methods for solving nonlinear evolutional differential equations is presented. The method is applicable to solving numerically the discretized Navier-Stokes equation. The rate of convergence of this method and its properties of parallelism are examined. Numerical research of the method is performed by the example of difference analogue of the two- and three-dimensional Navier-Stokes equation.

1. A Monte-Carlo method for solving systems of algebraic equations with quadratic nonlinearity

In the paper [1] it was shown that while Newton's method has quadratic rate of convergence, it's stochastic analogue has linear rate of convergence only, but lower computational complexity. Consequently, other deterministic methods with linear rate of convergence could exist, which stochastic analogues would have the same rate of convergence and the same computational complexity as the stochastic Newton's method.

Let's consider a system of equations with unknown $x \in \mathbb{R}^r$

$$x_i = f_i + \sum\nolimits_{j=1}^{r} a_{ij} x_j + \sum\nolimits_{j,k=1}^{r} b_{ijk} x_j x_k, \quad i = 1, \ldots, r. \tag{1}$$

By $\overline{x} = (\overline{x}_1, \ldots, \overline{x}_r) \in \mathbb{R}^r$ we denote its solution, which is supposed to be unique in some neighborhood. Multiplying left and right parts of the system (1) by x_l, $l = 1, \ldots, r$, we receive system

$$y_{il} = f_i x_l + \sum\nolimits_{j=1}^{r} a_{ij} y_{jl} + x_l \sum\nolimits_{j,k=1}^{r} b_{ijk} y_{jk}, \quad y_{il} = x_i x_l, \quad i, l = 1, \ldots, r.$$

[1]This work was supported by grant of RFBR 08-01-00194

[2]Saint Petersburg State University, E-mail: sergej.ermakov@gmail.com

[3]Saint Petersburg State University, E-mail: k.timofeev@gmail.com

If in this equation $x = \overline{x}$, then in a number of cases solution $\overline{y} = \|\overline{y}_{ij}\|_{i,j=1}^{r}$ of the constructed system corresponds to solution $\overline{x}$ of system (1). As a simple example, let us assume that $x_i \geq 0$, $i = 1, \ldots, r$, then the following equation takes place

$$x_i = \sqrt{y_{ii}}, \quad i = 1, \ldots, r.$$

Below we present the algorithm of the method of artificial chaos (AC) which uses this idea. The name for the method was given according to existing analogies with the Bird's method from gas dynamics (see [2]).

a) One chooses initial approximation x^0, assumes $n = 0$;

b) linear with respect to parameter y^{n+1} equation is solved

$$y_{il}^{n+1} = f_i x_l^n + \sum\nolimits_{j=1}^{r} a_{ij} y_{jl}^{n+1} + x_l^n \sum\nolimits_{j,k=1}^{r} b_{ijk} y_{jk}^{n+1}; \tag{2}$$

c) one assume $x^{n+1} = \psi(y^{n+1})$, where $\psi = (\psi_1, \ldots, \psi_r) : \mathbb{R}^{r^2} \to \mathbb{R}^r$ is some mapping;

d) variable n is increased by one and steps 2-3 are repeated until some stopping criterion conditions are fulfiled.

There are infinite number of ways of calculating x^{n+1} with given value y^{n+1}. For example, in case when one knows that $\sum\limits_{i=1}^{r} \overline{x}_i > 0$, one can take

$$\psi_i(y) = \frac{1}{2} \sum\nolimits_{j=1}^{r} (y_{ij} + y_{ji}) / \Big(\sum\nolimits_{j,k=1}^{r} y_{jk} \Big)^{-1/2}, \quad i = 1, \ldots, r. \tag{3}$$

Theorem 1. *If the following conditions are satisfied:*

a) for $y \in \mathbb{R}^{r^2}$ we construct $\varepsilon(y) \in \mathbb{R}^{r^2}$ in the following way: $\varepsilon_{ij}(y) = y_{ij} - \overline{x}_i \overline{x}_j$, $i, j = 1, \ldots, r$. Let's assume that there is constant $c_0 > 0$, such that for all y, that $\|\varepsilon(y)\| < c_0$, function $\psi : \mathbb{R}^{r^2} \to \mathbb{R}^r$ (see step 3 of the algorithm) can be written as $\psi(y) = \overline{x} + L\varepsilon(y) + O(\|\varepsilon(y)\|^2)$, where $L : \mathbb{R}^{r^2} \to \mathbb{R}^r$ — is some linear mapping;

b) there exists operator $(I - K_1 - K_2)^{-1}$, where operators $K_1 : \mathbb{R}^{r^2} \to \mathbb{R}^{r^2}$ and $K_2 : \mathbb{R}^{r^2} \to \mathbb{R}^{r^2}$ are determined by formulas

$$(K_1 y)_{ij} = \sum\nolimits_{k=1}^{r} a_{ik} y_{kj}, \quad (K_2 y)_{ij} = \sum\nolimits_{k,l=1}^{r} b_{ikl} \overline{x}_j y_{kl}, \quad i, j = 1, \ldots, r.$$

c) $w := \|(I - K_1 - K_2)^{-1}\| \|L\| \|(I - K_3)\overline{x}\| < 1$, where L is the linear operator from the first condition, operator $K_3 : \mathbb{R}^r \to \mathbb{R}^r$ is determined by formula $(K_3 x)_i = \sum\limits_{k=1}^{r} a_{ik} x_k$, $i = 1, \ldots, r$.

Then for each δ, $0 < \delta < 1 - w$, there is small value $\rho > 0$, such that for each vector $x^0 \in \mathbb{R}^r$ satisfying condition $\|\overline{x} - x^0\| < \rho$, statement $\|x^n - \overline{x}\| \to 0$ holds true when $n \to \infty$. At the same time the rate of convergence is linear: $\|x^n - \overline{x}\| \leq (w + \delta)\|x^{n-1} - \overline{x}\|$.

Scheme of the theorem proof is the following. One constructs mapping $G : \mathbb{R}^r \to \mathbb{R}^r$ such that $x^{n+1} = G(x^n)$ if x^n lies in a neighbourhood of $\overline{x}$. Further a sequence γ^n, $n = 0, 1, \ldots$, is introduced by expression $x^n = \overline{x} + \gamma^n$, and using Taylor series the expression $\gamma^{n+1} = H\gamma^n + Q(\gamma^n)$ is derived, where H is a matrix and Q is a mapping from $\mathbb{R}^r$ to $\mathbb{R}^r$. After that it is shown that $\|H\| \leq w < 1$ and for each δ, $0 < \delta < 1 - w$, there is constant $\rho > 0$, such that for each γ, $\|\gamma\| < \rho$, inequality $\|Q(\gamma)\| \leq \delta\|\gamma\|$ takes place. Thus the sufficient conditions of fulfillment of inequality $\|\gamma^{n+1}\| < (w + \delta)\|\gamma^n\|$ are found ant the theorem is proved.

It can be shown that expression (3) fulfils the first condition of Theorem 1. This expression will be used in the subsequent text for clearness.

Large dimension (r^2) of linear equation (2) hinders from its solving with deterministic methods. As for using expression (3) it is enough to calculate $r + 1$ linear functionals of solution of equation (2), application of Monte-Carlo methods makes the AC method comparable with stochastic Newton's method by computational complexity.

The method, based on the AC method and which uses direct or conjugate estimates by collisions or adsorption (see [3]) for solving linear equation (2), is named stochastic AC method. There are special techniques of application of these Monte-Carlo methods which require storing only coefficients of initial equation. Similar to sequential method suggested in [1], it is an iterative Monte-Carlo method, based on linearization procedure.

The authors proved a theorem, which shows that if some conditions are met, for estimates $\xi^{(n)} \in \mathbb{R}^r$ of stochastic AC method (where n is the number of iterations) the following inequality is valid $\|\mathbb{E}(\xi^{(n)} - \overline{x})(\xi^{(n)} - \overline{x})^T\| \leq c\|\mathbb{E}(\xi^{(n-1)} - \overline{x})(\xi^{(n-1)} - \overline{x})^T\|$, where $0 < c < 1$. This theorem is not presented in the paper due to its complexity. The authors are going to publish it on the Internet.

The randomized AC method is based on solving linear systems (2) with the use of the Monte-Carlo method, that permits the use of parallelism - one can average estimates after several iterations.

Let us consider computational complexity of the suggested method when estimates by collisions are used in auxiliary computations. Coefficients of the equation under consideration are supposed rarefied (most are zero). By n we denote a number of iterations of the randomized AC method, r is dimension of equation (1), m is a number of Markov chains that are used on every iteration, k is their average length, M is a number of available processors. Then computational complexity of the randomized AC method has order $n\lceil m/M \rceil (1 + k)O(r \ln(r))$.

Comparing received computational complexity with computational complexity of the Newton's method, and assuming the computational complexity of exact matrix inversion to be $O(r^2)$, then the randomized AC method has smaller computational complexity if $(n \ln(r))/(\ln(n)r) < O(1)$, where the right part of inequality is defined by coefficients of the equation and properties of methods realizations. This condition is satisfied if the problem's dimension is great and required accuracy is low.

Figure 1 demonstrates dependence of discrepancy norm on number of averaged estimates of linear system solution at every iteration for the randomized AC method (the conjugate estimate by collisions is used) and the randomized Newton's

method (computational complexities of two methods are comparable).

Figure 2 presents dependency of discrepancy norm on number of iterations in cases, when for solving auxiliary systems of linear equations at every iteration one uses $m = 70$ and 500 Markov chains. The flat part on the second graph corresponds to limitation of computational accuracy.

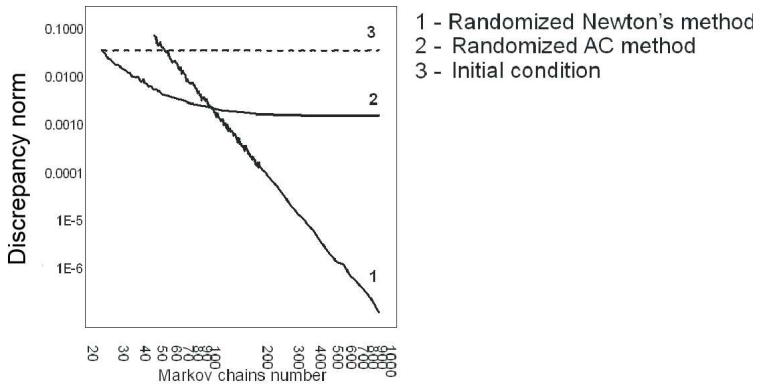


Figure 1: Dependence of discrepancy's norm from number of trajectories.

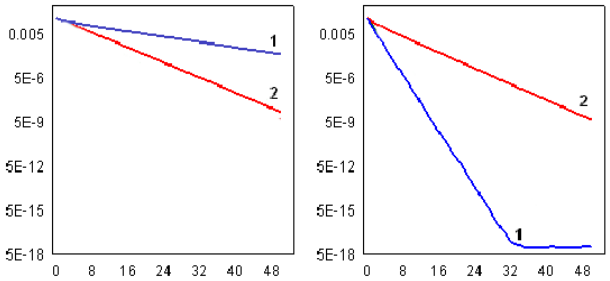


Figure 2: Dependence of discrepancy norm on iteration number (70, 500 estimates per iteration)

The figures demonstrate that the randomized AC method is more effective than the randomized Newton's method in a number of cases. It requires lower number of averages of auxiliary estimates at each iteration. One can also conclude that efficiency of the Newton's method in its convergence rate does not provide efficiency of its randomized analogue.

2. A Monte-Carlo Method for solving evolutional differential equations

One of the widespread methods of solving evolutional differential equations is by reduction to systems of ordinary differential equations of the following kind

$$\frac{\partial v(t)}{\partial t} = f(t, v), \quad v(0) = v^0, \quad t \geq 0, \tag{4}$$

where $v(t) = (v_1(t), \ldots, v_m(t))^T$, $f(t, v) = (f_1(t, v), \ldots, f_m(t, v))^T$. The mapping f can be finite-difference approximation of differential operator with partial derivatives. In practically interesting cases the dimension m can surpass 10^6.

If one uses Euler's method or Runge-Kutta method for solving system (4), then each step requires one or more calculations of vector $f(t, v)$. In the case of functions f which are difficult to calculate and in the case of big dimension m, this makes significant demands on calculating resources.

The technique, described bellow (randomization), allows us to decrease computational work in some cases and to show an easy way of parallelisation of the computational process.

Euler's method for solving equation (4) consists in construction of successive estimates v^n of values $v(\triangle tn)$, $n \in \mathbb{N}$, with the use of expression

$$v^{n+1} = v^n + \triangle t f(\triangle tn, v^n). \tag{5}$$

One can calculate random estimates $\widehat{v}^{n+1}$ in the form

$$\widehat{v}^{n+1} = \widehat{v}^n + \triangle t \widehat{f}(\triangle tn, \widehat{v}^n), \quad n \geq 0, \tag{6}$$

where $\widehat{v}^0 = v^0$, $\widehat{f}(t, v)$ is a family of random vectors, for which in some sense $\widehat{f}(t, v) \approx f(t, v)$. This approach was investigated earlier by Nekrutkin V., Golyandina N, Tur N., Potapov P. and others (see [4] and [5]).

If the family of random vectors $\widehat{f}$ is such, that its modeling time is lower than time of calculation of f, then profit in computational speed is achieved. As a consequence of using the stochastic procedure the additional stochastic error appears, which requires observation of covariance matrix of the estimate.

The process of modeling the estimates is easy to parallelize. If one uses L processors with independent pseudo-random numbers generators, one can decrease dispersions of components of estimates by L times by averaging L independent estimates $\widehat{v}^n$, calculated for fixed n on these processors (big-grained parallelism). At the same time one can calculate confidence intervals for estimates.

It is also possible to average several estimates on each step n (small- grained parallelism).

Thus, up to 100% of computational resources of parallel processors can be used.

In the present report an estimate, applicable to solving wide class of nonlinear differential evolutional equations is suggested. Let us consider a family of random vectors with parameters z, t and v

$$\xi^{(z)}(t, v) = \frac{g_z(\|v\|) f_\alpha(t, v)}{p_\alpha(t, v)} e_\alpha, \tag{7}$$

where vector $p(t,v) = (p_1(t,v), \ldots, p_m(t,v))$ with fixed $t \in [0,T]$ and $v \in \mathbb{R}^m$ specifies a distribution on $\{1,\ldots,m\}$, α is random variable with distribution $p(t,v)$, e_i — i-th m-dimensional unit vector, function $g_z(y) : [0,\infty) \to [0,\infty)$ with fixed $z > 0$ is determined by expression

$$g_z(y) = \begin{cases} 1 & , y < z; \\ 2(y-z)^3 - 3(y-z)^2 + 1 & , z \le y < z+1; \\ 0 & , y \ge z. \end{cases}$$

Theorem 2. *Assume that equation (4) has unique solution $v(t)$ on interval $t \in [0,T]$. Let $c_1 := \|v(t)\|_{C[0,T]}$. Let at the same time for each $t \in [0,T]$ function $f(t,v)$ is twice continually differentiable by the second variable for all values v, such that $\|v\| \le c_1$.*

Assume that there is constant c_2, such that for each $i = 1,\ldots,m$ one of the following equations takes place: $p_i(t,v) \ge c_2 g_{2c_1}(\|v\|)|f_i(t,v)| / \sum\limits_{j=1}^{m} |f_j(t,v)|$ or $p_i(t,v) \ge c_2 g_{2c_1}(\|v\|)|f_i(t,v)|$.

Consider the following sequence of random vectors ν^n, $n \ge 0$:

$$\nu^{n+1} = \nu^n + \tau \xi^{(2c_1)}(\triangle t n, \nu^n), n \ge 0, \quad \nu^0 = v^0.$$

Then while $\triangle t$ tends to zero the following states hold true uniformly by $t \in [0,T]$

a) $\|\mathbb{E}\nu^{\lfloor t/\triangle t \rfloor} - v(t)\| = O(\triangle t)$;

b) $\|cov\ \nu^{\lfloor t/\triangle t \rfloor}\| = O(\triangle t)$ *(where cov denotes covariance matrix of random vector);*

c) $\|\mathbb{E}(\nu^{\lfloor t/\triangle t \rfloor} - v(t))(\nu^{\lfloor t/\triangle t \rfloor} - v(t))^T\| = O(\triangle t)$.

This theorem shows in particular, that the proposed method has the same rate of convergence as Euler's method.

The scheme of the theorem proof is following. Using finite dimensionality of considering spaces the proof is performed for Euclid norm ($\|\cdot\| = \|\cdot\|_2$). First, using Taylor series one proves that there are constants c_3 and c_4, such that for each random vector ν with finite covariance matrix cov ν the inequality $\|\mathbb{E}(\xi^{2c1}(t,\nu) - g_{2c1}(\|\nu\|)f(t,\nu))(\xi^{2c1}(t,\nu) - g_{2c1}(\|\nu\|)f(t,\nu))^T\| \le c_3 + c_4\|R\nu\|$ holds true. Using the obtained inequality and Taylor series one proves that $\|\text{cov } \nu^{n+1}\| \le (1 + c_5\triangle t + c_6(\triangle t)^2)\|\text{cov } \nu^n\| + c_7(\triangle t)^2$ for all $n < \lfloor t/\triangle t \rfloor$, where c_5, c_6 and c_7 are constant values. Using this inequality one can show that the second assertion of the theorem holds true. Further, using fulfillment of the second theorem's assertion, one can prove that there are constants c_8 and c_9 such that $\|\mathbb{E}(\nu^{n+1} - v^{n+1})\| \le (1 + \triangle t(c_8 + c_9\|\mathbb{E}(\nu^n - v^n)\|))\|\mathbb{E}(\nu^n - v^n)\| + O((\triangle t)^2)$, where sequence v^n is determined by (5). Then it can be shown by induction that $\|\mathbb{E}\nu^{\lfloor t/\tau \rfloor} - v^{\lfloor t/\tau \rfloor}\| = O(\tau)$, which leads to validity of the first theorem's assertion. Finally, the third assertion of the theorem follows from the first two.

For the proposed random estimate one can show theoretical method of selection of locally (for the fixed time-step number) optimal parameters.

The presented method is effective for solving equations, which require small time step values. As an example one can point out the Navier-Stokes equation with rather big Reynolds number (Re). For testing the method authors solved difference analogues of two- and three-dimensional Navier-Stokes equation (circulation flow of liquid in a square cavern with a movable lid). The dimension of solved equations was up to $3 \cdot 10^6$. The profit in computational time comparing to Euler's method was up to 30 times (a small time step was selected).

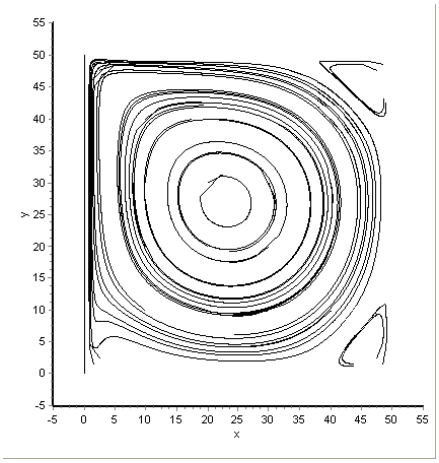


Figure 3: Example of solution of 2d Navier-Stokes equation

References

[1] Halton Jh. (2006) *Sequential Monte Carlo Techniques for Solving Non-Linear Systems* // Monte Carlo Methods and Appl. Vol. 12. P. 113–141.

[2] Bird G.A. (1994) *Molecular Gas Dynamics and The Direct Simulation of Gas Flows*, Claderon Press. Oxford.

[3] Ermakov S. M. (1971) *Monte Carlo Method and Related Questions*, Nauka, Moscow.

[4] Golyandina N., Nekrutkin V. (1999) *Homogeneous Balance equations for measures: errors of the stochastic solution* // Monte Carlo Methods and Appl., V.5, No 3, pp.1-67.

[5] Nekrutkin V., Potapov P. (2004) *Two variants of a stochastic Euler method for homogeneous balance differential equations.* // Monte Carlo Methods Appl. 10, No. 3-4, 469-479.

6th St.Petersburg Workshop on Simulation (2009) 929-933

Application of the Monte-Carlo and Quasi Monte-Carlo methods to solving systems of linear equations[1]

Sergej M. Ermakov[2], Anna I. Rukavishnikova[3]

Abstract

This report suggests and explains stochastic method for experimental error estimation for the Quasi Monte-Carlo method. New modification of the Monte-Carlo and Quasi Monte-carlo methods is suggested and applied to solving systems of linear equations. This modification permits to weaken condition of dominated convergence and to decrease constructive dimension of estimated integrals for the Quasi Monte-Carlo method. In that way the modification provides higher rate of convergence than the Quasi Monte-Carlo method. Optimization of estimates which are used in the modified Quasi Monte-Carlo method is performed.

1. Introduction

The Quasi Monte-Carlo(QMC) method is known to be a method of numerical integration, which set of nodes is a sequel of points with property of "good" uniformity. The uniformity criteria is value of "star discrepancy" ([1]) :

Definition 1. *For set of s-dimensional points* $Y = \{Y_0, ..., Y_{N-1}\}$, $Y_i \in [0,1]^s$, $i = 0, \ldots, N-1$, *and for measurable subset* $Q \subset [0,1]^s$ *local "discrepancy" is defined as*

$$D(Q,Y) := \frac{1}{N} \sum_{0 \le p \le N} \chi_Q(Y_p) - \int\limits_{[0,1]^s} \chi_Q(X) dX,$$

where χ_Q *is indicator of subset* Q.

Definition 2. *"Star discrepancy" of set of s-dimensional points* Y *is defined as*

$$D^*(Y) := \sup_{Q^*} |D(Q^*, Y)|,$$

where Q^* *runs through all intervals* $[0, u_0) \times \ldots \times [0, u_s)$, $0 \le u_i \le 1$, $i = 1, \ldots, s$.

[1]This work was supported by grant 08-01-00194 RFBR 2009

[2]Petersburg University, E-mail: Sergej.Ermakov@pobox.spbu.ru

[3]Petersburg University, E-mail: anyaruk@mail.ru

The QMC error can be theoretically estimated by the Koksma-Hlavka inequality

Theorem 1. *For arbitrary sequence of s-dimensional points* $Y = \{Y_0, ..., Y_{N-1}\} \subset [0,1]^s$ *and for every function f, which variation is bounded in terms of Hardi-Krause, the QMC error can be estimated by*

$$\Big|\frac{1}{N}\sum_{n=0}^{N-1} f(Y_n) - \int\limits_{[0,1]^s} f(Y)dY\Big| \le V(f)D^*(Y),$$

where $V(f)$ *– is variation of function* f *in terms of Hardi-Krause.*

The quasi random sequences of Halton and Sobol are known to provide asymptotically optimal decrease of "star discrepancy". They give the following error in the QMC method: $R_N[f] \sim O(ln^s N/N)$.

2. Stochastic estimation of the QMC error

One of the well-known problems in the QMC application is difficulty of the error estimation. This difficulty can be explained by obvious impossibility of direct computation of the right part of the Koksma-Hlavka inequality, especially component $V(f)$.

The following stochastic method of the QMC error estimation is suggested.

Let's denote the QMC estimate by J_0

$$J_0 = \frac{1}{N}\sum_{j=1}^{N} f(X_j),$$

where $X_j \in \mathbb{R}^s$ are quasirandom vectors.

We construct a set of stochastic estimates of the integral under consideration, which use nodes of the QMC method (vectors X_j). In order to receive one estimate J_i all nodes X_j, $j = 1, ..., N$ are shifted by random vector $\Xi^i = (\alpha_1^i, ..., \alpha_s^i)$, which components are independent and uniformly distributed in $[0,1]$, then, so as to stay in boundaries of cube $[0,1]^s$, we consider fractional part of the result. The following lemma is valid.

Lemma 1. *Let's consider estimate* $J = \frac{1}{N}\sum\limits_{j=1}^{N} f(\{X_j + \Xi\})$, *where* $\Xi = (\alpha_1, ..., \alpha_s) \in [0,1]^s$ *is a random vector, which components are independent and uniformly distributed in* $[0,1]$, $X_j = (x_1^j, ..., x_s^j) \in [0,1]^s$ *is an arbitrary vector. Estimate* J *is an unbiased estimate of integral* $\int\limits_{[0,1]^s} f(X)dX$, *i.e.* $\mathbb{E}J = \overline{J}$.

One should expect that dispersion σ^2 of estimates J in lemma 1 is small, due to uniform nature of quasirandom nodes X_j, $j = 1, ..., N$ and to the fact that operations of shift and of fractional part $\{x_j + \alpha^{(i)}\}$ preserve distances between the nodes (in metric of torus).

Thus, the described method could turn out to be effective especially due to the property of uniformity of quasirandom numbers.

The research is confirmed by a lot of conducted experiments.

3. Modification of the Monte-Carlo method

Idea of the modification belongs to Ermakov S.M. and Wagner W. (see [2]), it was discussed in view of stochastic stability of the Monte-Carlo. But application of the modification to solving systems of linear equations is a new result obtained by the authors of report.

Main goal of this modification is to weaken the condition of dominated convergence, which is a natural restriction and stems from necessity of solution's representation as an integral over trajectories.

Algorithm of the modification.

Consider a system of linear equations $X = AX + F$, where $A = \|a_{ij}\|_{i,j=1}^{n}$ is system's matrix, $F = \{f_i\}_{i=1}^{n}$ is a vector in the right part. We assume that $\rho(A) < 1$ (ρ is a spectral radius)

Let's fix some rather great number $k \in \mathbb{N}$ of computated components in the Newmann series, then

$$X = F + AF + \ldots + A^k F + \varepsilon_k = X_k + \varepsilon_k,$$

where ε_k is such systematic error, that $\|\varepsilon_k\| \leq \|A\|^{k+1} \frac{\|F\|}{1-\|A\|} \to 0$ when $k \mapsto \infty$.

Initial system $X = AX + F$ is equivalent to $(X - X_0) = A(X - X_0) + (AX_0 - X_0 + F)$, where X_0 is an initial approximation of the systems solution. With the use of the Monte-Carlo method, one calculates estimates $\overline{AY}$ of product of matrix A and vector Y, where Y differs from step to step:

- $Z \leftarrow AX_0 + F$;
- at the first step one calculates estimate $\xi_1 = \overline{AY}$, where $Y = AX_0 - X_0 + F$, $Z \leftarrow Z + \xi_1$;
- at the second step one calculates estimate $\xi_2 = \overline{AY}$, where $Y = \xi_1$ is estimate, received at the first step, $Z \leftarrow Z + \xi_2$;
- at step k one calculates estimate $\xi = \overline{AY}$, where $Y = \xi_{k-1}$ is estimate, received at step k-1, $Z \leftarrow Z + \xi_k$.

Final estimate of system's solution X is Z.

One variant of estimates for estimation of $\overline{AY}$ is presented below.

Let's fix initial distribution $P^0 = \|p_i^0\|_{i=1}^{n}$ and stochastic matrix of transition probability $P = \|p_{ij}\|_{i,j=1}^{n}$ (its lines determine distributions). Let α be a random value with distribution P^0, β — a random value with distribution $P(\beta = j|\alpha = i) = p_{ij}$, $i, j = 1, \ldots, n$.

Lemma 2. *Let's consider estimate* ξ*:*

$$\xi = \frac{y_\alpha}{p^0_\alpha}\frac{a_{\beta\alpha}}{p_{\alpha\beta}}e_\beta,$$

where e_i *is unit* n*-dimensional vector with component* i *equal to 1. We assume the following fitting conditions: if* $a_{ij}y_j \neq 0$*, then* $p_{ji} > 0$ *and if* $y_j \sum\limits_{i=1}^{n} a_{ij} \neq 0$*, then* $p^0_j > 0$*. In that case* $\mathbb{E}\xi = AY$*.*

Problem of optimization of parameters P^0 and P arises naturally.

If one assume that optimality is minimization of expression $\sum\limits_{i=1}^{n} \mathbb{D}\xi_i$, then the following theorem suggest optimal parameters.

Theorem 2. *Under the following fitting conditions expresion* $\sum\limits_{i=1}^{n} \mathbb{D}\xi_i^{(2)}$ *reaches minimum with parameters*

$$p^0_j = |y_j|\sum_{i=1}^{n}|a_{ij}|/\sum_{i,k=1}^{n}|a_{ik}y_k|, \quad p_{ji} = |a_{ij}|/\sum_{k=1}^{n}|a_{kj}|.$$

4. Modification of the Quasi Monte-Carlo method

Modification described in the previous section can be easily applied to the Quasi Monte-Carlo method due to similarities in algorithms of MC and QMC. The reason for this application is following.

The rate of error reduction for the QMC is $O(\log^s N/N)$, where N is a number of auxiliary estimates, s is constructive dimension of the estimated integral. Thus the QMC can be more effective than the MC, which rate of error reduction is $O(1/\sqrt{N})$.

But if s is rather great (as is the case of linear systems solving), then the rate of QMC error reduction is low and this is a great disadvantage of the QMC.

The suggested modification settle this problem, since with the use of the modification constructive dimension of estimated integrals is constant and equal to 1.

Let's estimate the rate of error convergence for the modified QMC.

We assume that $X_0 = 0$. The modified QMC implies calculation of sequence of estimates $\Xi_{n+1} = \widehat{A\Xi_n}$, where Ξ_n is estimate of $D_n = X_n - X_{n-1}$, $X_{n+1} = AX_n + F$, $\Xi_1 = F$.

Suppose $\Xi_n = D_n + \varepsilon_n$, then $\varepsilon_1 = 0$.

Proposition 1. *For the error of one step of the modified QMC the following recurrent equation takes place* $\varepsilon_{n+1} = (A + \delta_n)\,\varepsilon_n + \delta_n D^n$*, where* δ_n *is a matrix which norm is* $\|\delta_n\| = O(\frac{\ln N}{N})$*, due to the rate of error reduction of the QMC method; and*

$$\|\varepsilon_{n+1}\| \leq O\left(\frac{\ln N}{N}\right)\|A\|^{n-1}\|F\|n.$$

Let's denote the error of the modified QMC by Δ_k, where k is a number of estimated components in the Newman series.

Proposition 2. *The following recurrent expression takes place:*
$\|\Delta_k - X_k\| \leq O\left(\frac{\ln N}{N}\right)$.

5. Examples

First example. Let's solve a system $X = AX + F$ of linear equations, A is a matrix $[100 \times 100]$. Components of the systems matrix $A = \{a_{ij}\}_{ij=1}^{n}$ and of the vector $F = \{f_i\}_{i=1}^{n}$ are chosen randomly, but on condition that $\sum\limits_{j=1}^{n} |a_{ij}| = 0.9$, thus $\|A\| < 1$. In order to compare MC, QMC and modified MC and QMC methods we denote by N a number of "random" numbers (pseudorandom for MC and modified MC, quasirandom for QMC and modified QMC) used in every method.

Comparing the errors of MC, QMC and modified MC and QMC, one can see that the modified QMC is the most efficient with growth of computational complexity(N).

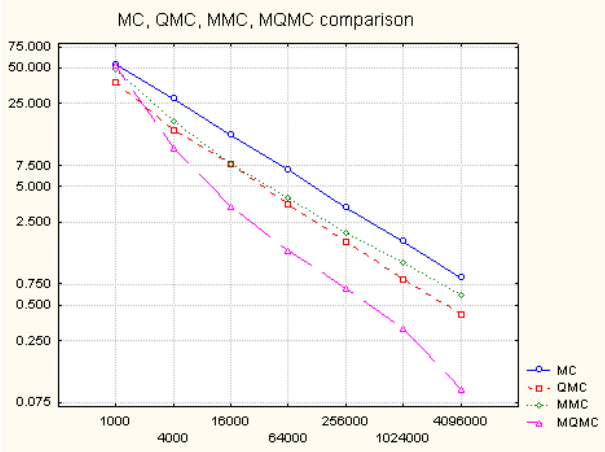


Figure 1: Errors of MC, QMC and modified MC and QMC methods with growth of N.

Second example. Let's solve a system of linear equations $X = AX + F$, A is a matrix $[100 \times 100]$. Components of the matrix A and vector $F = \{f_i\}_{i=1}^{n}$ are chosen randomly. In this case we apply a relaxation parameter

$$X_{n+1} = X_n - \tau(X_n - AX_n - F). \tag{1}$$

The distribution law of the coefficients of matrix A was chosen in such way, that the dominated convergence condition for new system is broken, while still $\rho(A) < 1$ (ρ denotes spectral radius).

Let's compare modified MC and QMC methods, N is a number of "random" numbers (pseudorandom for modified MC, quasirandom modified QMC) used in every method.

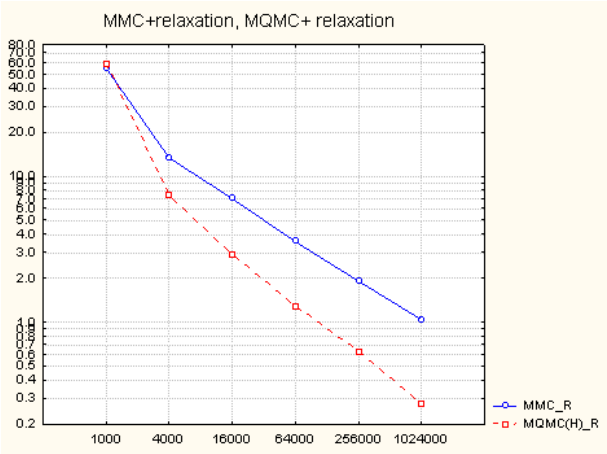


Figure 2: Errors of the modified MC and QMC methods with growth of N; using relaxation parameter.

One can see that the modified QMC is more efficient efficient than the modified MC, especially with growth of computational complexity(N).

References

[1] Kuipers and Niederreiter (1974) *Uniform Distribution of Sequences.* Wiley-Interscience, New-York.

[2] Wagner W., Ermakov S.M. (2001) *Stochastic stability and parallelism of the Quasi Monte-Carlo method.* Report to Academy of Sciences.

6th St.Petersburg Workshop on Simulation (2009) 935-939

Stochastic Modeling in Chromatography[1]

Boris P. Harlamov[2]

Abstract

A stochastic model of gas chromatography is considered. A semi-Markov process of diffusion type is used to describe movement of particles of analyzable matter through a porous medium. In order to obtain local parameters of the process we propose to use macroscopic parameters of the gas-eluent (carrier) in a chromatography column. The problem is set to find a distribution of the first exit time for a particle leaving the column. This distribution in practice corresponds to a form of curve peak on the chromatogram. Hence it can serve as a basis for estimating parameters of the model. Some moments of the distribution and other derivative characteristics of the chromatography process are found.

Introduction We consider a stochastic model for movement of particles of analyzable matter through a chromatography column with flow of an eluent. For chromatography process modeling we use a continuous semi-Markov process. Its connection with the process of chromatography separation firstly was mentioned in work [4], although the first work where actually a monotone continuous semi-Markov process was applied for analysis in liquid chromatography is a work of Gut and Ahlberg [3]. In the present paper we discuss some results on semi-Markov modeling in gas chromatography taking into account gas-dynamical phenomena in the column. The proposed model gives understanding of mechanism of chromatography separation under some ideal conditions. Hence it never replaces substitutes all the important engineering tricks which were made for practical aims.

Continuous semi-Markov processes. A one-dimensional process (X_t) $(t \geq 0)$ with continuous trajectories is said to be continuous semi-Markov if it posesses Markov property with respect to the first exit time from any open interval $(a, b) \subset (-\infty, \infty)$. This process is determined by a consistent semi-Markov family of probability measures which may be non-Markovian. In what follows the term semi-Markov refers only to continuous semi-Markov processes. A semi-Markov process is said to be non-decreasing if $P_x(X_{\sigma_{(a,b)}} = a) = 0$ for any $a < x < b$, here $\sigma_{(a,b)}$ is the first exit time from (a, b). Such a class of processes describes phenomena in liquid column chromatography. Semi-Markov processes of diffusion type are used for analysis of gas chromatography.

[1]This work was supported by grant RFBR NSh-4222.2006.1

[2]Institute of Problems of Mechanical Engineering, E-mail: boris.harlamov@ipme.ru

Monotone semi-Markov processes and liquid column chromatography Let us consider the first exit time of a non-decreasing semi-Markov process from the interval (a, b) as a random variable with respect to the measure P_x. Write $\tau_b = \sigma_{(a,b)}$. Then (τ_b) as a function of b is a process with independent increments. For liquid chromatography it means that random time delays of a particle on non-overlapped column segments are mutually independent. Hence the process (X_t) itself is the inverse process with independent positive increments. Its properties are well-known. In particular, except a trivial case a trajectory of such a process contains intervals of constancy of random value and position. If the process X_t is interpreted as movement of a particle of a absorbable matter through a chromatography column, then intervals of constancy represent periods of the particle being in absorbed state. The deterministic component of a trajectory corresponds to velocity of the eluent which carries the particle away from points of sorbtion up to the next stopping and so on. The semi-Markov model of liquid chromatography is the simplest part of the theory which requires minimum of assumptions. It successfully permits to evaluate main parameters of chromatography separation in terms of few parameters of the model.

Semi-Markov processes of diffusion type and gas chromatography The model of gas chromatography is more complex since it must take into account diffusion of the gas-carrier. Besides in gas chromatography we have to consider acceleration of the carrier while moving along the column. In this case the movement of a particle is well described by a diffusion semi-Markov process.

One-dimensional semi-Markov process is said to be a process of diffusion type if for any $a < x < b$ both probabilities $P_x(X_{\sigma_{(a,b)}} = a)$ and $P_x(X_{\sigma_{(a,b)}} = b)$ are positive. We will consider a case when a particle does leave of any bounded interval: $P_x(X_{\sigma_{(a,b)}} = a) + P_x(X_{\sigma_{(a,b)}} = b) = 1$. One-dimensional semi-Markov process of diffusion type is completely determined by a family of semi-Markov transition generating functions:

$$g_{(a,b)}(\lambda; x) = E_x(\exp(-\lambda\sigma_{(a,b)}); X_{\sigma_{(a,b)}} = a),$$

$$h_{(a,b)}(\lambda; x) = E_x(\exp(-\lambda\sigma_{(a,b)}); X_{\sigma_{(a,b)}} = b),$$

where $\lambda \geq 0$. Both these functionals as functions of x satisfy the differential equation [5]

$$\frac{1}{2}\,u'' + b(x)u' - c(\lambda, x)u = 0 \tag{1}$$

with boundary conditions

$$g_{(a,b)}(\lambda; a) = h_{(a,b)}(\lambda; b) = 1, \quad g_{(a,b)}(\lambda; b) = h_{(a,b)}(\lambda; a) = 0,$$

where $b(x)$ is a continuous differentiable function, $c(\lambda, x)$ is continuous in x and infinitely differentiable with respect to λ, it is positive for $\lambda > 0$ and vanishes at $\lambda = 0$, moreover for any x the function $c(\lambda, x)$ has a positive completely monotone first derivative with respect to λ. According to the famous Bernstein theorem this

function can be represented in the form

$$c(\lambda, x) = \lambda\,\gamma(x) + \int_{0+}^{\infty} (1 - \exp(-\lambda\,u))\,\eta(du|\,x),$$

where $\gamma(x)$ is some positive function, $\eta(du|\,x)$ being a family of measures on $(0, \infty)$. The semi-Markov family of measures (P_x) is a strong-Markov family (more narrow class), if all the measures η are equal to zero.

An important part of our model is the assumption that a strong-Markov process, controlled by equation

$$\frac{1}{2}\,u'' + b(x)u' - \lambda\,\gamma(x)u = 0, \tag{2}$$

describes movement of the gas-carrier in the column. This process is called a supporting Markov process for the semi-Markov one. The non-linear in λ part of the function $c(\lambda, x)$ determines delay of sorbable particle of the gas mixture with respect to eluent.

Let $(\overline{P}_x)$ be a family of measures of the supporting Markov process; $\overline{g}_{(a,b)}(\lambda, x)$, $\overline{h}_{(a,b)}(\lambda, x)$ be corresponding transition generating functions of this process, which satisfy equation (2). Analysis of this equation permits to interprete coefficients $b(x)$ and $\gamma(x)$ in terms of local Kolmogorov's parameters of the diffusion Markov process. Namely, $1/\gamma(x)$ is a local diffusion coefficient, and $b(x)/\gamma(x)$ is a local drift coefficient. One can estimate these coefficients in macroscopic terms of the gas-eluent movement along the column.

We assume that 1) $\gamma(x)$ and $\eta(du|\,x)$ do not depend on x (at least for a termostated column), 2) $V(x) \equiv b(x)/\gamma(x)$ is a group velocity of the gas-eluent along the column.

Velocity of gas Evaluation of gas velocity in cylindrical vessel is based on the Boyle-Mariott and Gagen-Poiseil equations [2, c. 71], they entail that velocity of gas at the cross-section x of the vessel is equal to

$$V(x) = \frac{Cm_0/S}{\sqrt{p^2(b) + (b - x)\,2Cm_0/(kS)}}, \tag{3}$$

where $p(b)$ is a gas pressure; for usual operation $p(b)$ is equal to atmospheric pressure; $C = cT$, T is the absolute temperature, c is a coefficient depending on chemical composition of gas (in present work we do not consider dependence of the parameters on T), m_0 is a normed expenditure of gas mass, which is constant at every cross-section of the vessel under stationarity condition, k is a coefficient depending on a quality of the column filling and chemical composition of the gas, S is the area of a cross-section. Denominator of this expression represents a pressure at cross-section x of the column, in particular, $p^2(0) = p^2(b) + b\,2Cm_0/(kS)$.

Now equation (1) can be rewritten in the form

$$\frac{1}{2}\,f'' + V(x)\,\gamma\,f' - c(\lambda)\,f = 0, \tag{4}$$

where $\gamma = \gamma(0)$, $c(\lambda) = c(\lambda, 0)$ (do not depend on x). We do not know an analytical form of solution $h_{(-\infty,b)}(\lambda; x)$, the most interesting one. However, in order to find moments of the random variable τ_b we need not have this analytical form because from equation (4) we can derive a differential equation for any moments $M_{k,a}(x)$ ($k \geq 1$), where $M_{k,a}(x) = E_x((\sigma_{(a,b)})^k)$.

Approximate solution Before analyzing equation (4) with variable coefficient $V = V(x)$ we consider an approximate solution, which is a solution of a similar equation but with constant V, for example, with $V(x)$ averaged over the interval $(0, b)$. The solution we are interested in is

$$f(\lambda|\, x) = \exp\left((b-x)(V\gamma - \sqrt{(V\gamma)^2 + 2c(\lambda)})\right).$$

The moments can be found by differentiating the function $f(\lambda|\, x)$ with respect to λ at the point $\lambda = 0$. We have

$$M_1(x) \equiv E_x(\tau_b) = -\left.\frac{\partial}{\partial\lambda} f(\lambda|\, x)\right|_{\lambda=0} = (b-x)\frac{\gamma+\delta_1}{V\gamma},$$

with

$$\gamma + \delta_1 = \left.\frac{\partial c(\lambda)}{\partial\lambda}\right|_{\lambda=0}, \quad \delta_1 = \int_0^\infty u\,\eta(du),$$

$$M_2(x) \equiv E_x((\tau_b)^2) = \left.\frac{\partial^2}{(\partial\lambda)^2} f(\lambda|\, x)\right|_{\lambda=0} =$$

$$= \left(\frac{(b-x)(\gamma+\delta_1)}{V\gamma}\right)^2 + (b-x)\frac{\delta_2}{V\gamma} + (b-x)\frac{(\gamma+\delta_1)^2}{(V\gamma)^3},$$

and

$$\delta_2 = -\left.\frac{\partial^2 c(\lambda)}{(\partial\lambda)^2}\right|_{\lambda=0} = \int_0^\infty u^2\,\eta(du).$$

Whence:

$$m_1(b) \equiv M_1(0) = b\frac{\gamma+\delta_1}{V\gamma}, \quad \mu_2(b) \equiv M_2(0) - m_1^2(b) = b\frac{\delta_2}{V\gamma} + b\frac{(\gamma+\delta_1)^2}{(V\gamma)^3}.$$

Recall a useful characteristic of chromatograph, a so called height equivalent to a theoretical plate (HETP). According to Giddings [1] it is defined as follows

$$H(b) \equiv \frac{b\mu_2(b)}{m_1^2(b)}.$$

From our model with constant velocity we obtain

$$H(b) = \frac{\delta_2}{(\gamma+\delta_1)^2}\,V\gamma + \frac{1}{V\gamma}.$$

This is just the same expression as in famous van Deemter formula in gas chromatography.

First moment In case of a variable coefficient $V(x)$ we can derive from equation (4) new equations for all the moments of the first exit time τ_b. The equation for the first moment is

$$\frac{1}{2}\,M''_{1,a} + V(x)\,\gamma\,M'_{1,a} + (\gamma+\delta_1) = 0, \tag{5}$$

with boundary conditions $M_{1,a}(a) = M_{1,a}(b) = 0$. This gives

$$M_1(x) = \lim_{a\to-\infty} M_{1,a}(x) =$$

$$= (b-x)(\gamma+\delta_1)/(4A^2\gamma^2) + \frac{2}{3}\left((B-x)^{3/2} - (B-b)^{3/2}\right)(\gamma+\delta_1)/(A\gamma) =$$

$$= \frac{\gamma+\delta_1}{A\gamma}\left(\frac{b-x}{4A\gamma} + \frac{2}{3}((B-x)^{3/2} - (B-b)^{3/2})\right), \tag{6}$$

where according to formula (3)

$$V(x) = A/\sqrt{B-x}, \quad A = \sqrt{\frac{Cm_0k}{2S}}, \quad B = \frac{p^2(0)kS}{2Cm_0}.$$

Thus,

$$m_1(b) \equiv M_1(0) = \frac{\gamma+\delta_1}{A\gamma}\left(\frac{b}{4A\gamma} + \frac{2}{3}(B^{3/2} - (B-b)^{3/2})\right).$$

Second moment Twice differentiating of equation (4) with respect to λ at the point $\lambda = 0$ gives

$$\frac{1}{2}\,M''_{2,a} + \gamma V M'_{2,a} + \delta_2 + 2(\gamma+\delta_1)M_{1,a} = 0.$$

Getting solution of this equation under zero boundary conditions we find $M_2(x)$ as a limit of $M_{2,a}(x)$ when $a \to -\infty$. Making elementary but tedious transformations, we obtain

$$\mu_2(b) = \delta_2\left(\frac{b}{4A^2\gamma^2} + \frac{2(B^{3/2} - (B-b)^{3/2})}{3A\gamma}\right) +$$

$$+\frac{(\gamma+\delta_1)^2}{A^2\gamma^2}\left(\frac{11b}{64A^4\gamma^4} + \frac{11(B^{3/2} - (B-b)^{3/2})}{16A^3\gamma^3} +\right.$$

$$\left.+\frac{5(B^2 - (B-b)^2)}{8A^2\gamma^2} + \frac{2(B^{5/2} - (B-b)^{5/2})}{5A\gamma}\right).$$

Derivative characteristics

Dependence of HETP on velocity can be expressed with the help of local HETP, which is defined as an integral characteristic of a short column, beginning at the point x:

$$H(x) = \frac{\psi_2(x)}{(\psi_1(x))^2}, \quad \psi_1(x) = -\frac{\partial}{\partial x}M_1(x), \quad \psi_2(x) = -\frac{\partial}{\partial x}\widetilde{\mu}_2(x),$$

where $\widetilde{\mu}_2(x) \equiv M_2(x) - M_1^2(x)$. For a model with a constant velocity of gas the local HETP coincides with the integral one. In general case we have

$$\psi_1(x) = \frac{\gamma + \delta_1}{A\gamma}\left(\frac{1}{4A\gamma} + (B-x)^{1/2}\right),$$

$$\psi_2(x) = \frac{\delta_2}{A\gamma}\left(\frac{1}{4A\gamma} + (B-x)^{1/2}\right) + \left(\frac{\gamma+\delta_1}{A\gamma}\right)^2\left(\frac{11}{64}\,\frac{1}{A^4\gamma^4} + \right.$$
$$\left. + \frac{11}{16}\,\frac{(B-x)^{1/2}}{A^3\gamma^3} + \frac{5}{4}\,\frac{B-x}{A^2\gamma^2} + \frac{(B-x)^{3/2}}{A\gamma}\right),$$

where according to accepted notations we have $(B-x)^{1/2} = A/V(x)$. From analysis of these expressions it follows that for any x there exists a minimum of $H(x)$ as a function of velocity. One can consider this formula as refined variant of van Deemter formula for gas chromatography taking into account different velocities of eluent at different cross-sections of the column.

References

[1] *J.C. Giddings* Dynamics of chromatography. New York: Marcel Decker. Inc., 1965. Vol. 1.

[2] *Golbert C.A., Vigdergauz M.S.*, Course of Gas Chromatography, Chemistry, Moscow, 1974 (in Russian).

[3] *A. Gut, P. Ahlberg.* On the theory of chromatography based upon renewal theory and central limit theorem for randomly indexed partial sums of random variables. Chemica Scripta, v.18, 5, 1981, 248–255.

[4] *B.P. Harlamov.* Continuous semi-Markov processes and their application to chromatography. Proceedings of the 2-nd International Symposium on Semi-Markov Models: Theory and Applications, Compiegne, Dec. 9-11, Université de Compiègne, 1998, 251–255.

[5] *B.P. Harlamov.* Continuous semi-Markov processes. ISTE & Wiley, London, 2008.

6th St.Petersburg Workshop on Simulation (2009) 941-945

Artificial Monte Carlo interactions for linear problems[1]

Vladimir Nekrutkin[2], Nikolay Rumyantsev[3]

Abstract

Two variants of interacting multidimensional Markov processes are proposed for Monte Carlo solution of linear problems associated with backward Kolmogorov equations for one-dimensional jump-wise Markov processes. Theoretical results on complexities of the corresponding estimates are presented. The results of computational experiments are discussed.

1. Introduction

Let (D, ρ) be a metric space equipped with Borel σ-algebra $\mathcal{B}$. Denote H the set of all distributions defined on $(D, \mathcal{B})$ and consider the equation

$$\frac{d\mu_t}{dt} = \int_D T(\,\cdot\,; u)\mu_t(du) - \mu_t, \tag{1}$$

where $T(\,\cdot\,; u) \in H$ for any $u \in D$ and $\mu_t|_{t=0} = \mu \in H$. Assume that our aim is to find the value of linear functional $\psi(\mu_t) = \int_D g d\mu_t$ for fixed $g \in \mathbf{C}(D)$ and $t > 0$.

Of course, this problem has a linear nature. The standard Monte Carlo solution of this problem (briefly, *the* $\mathbf{S}$*-method*) is also linear and can be described as follows.

Let $\zeta(t) \in D$ be a jump-wise Markov process, determined by an initial distribution μ, a jump frequency $\lambda = 1$ and a jump law $T(\,\cdot\,, u)$. If $\zeta_1(t), \ldots, \zeta_n(t)$ stand for independent copies of $\zeta(t)$ and $w_n(g, t) = \sum_{i=1}^n g(\zeta_i(t))/n$, then $\mathrm{E}w_n(g, t) = \psi(\mu_t)$ and $\mathrm{D}w_n(g, t) = \sigma^2_{\mathbf{S}}/n$ with $\sigma^2_{\mathbf{S}} = \int_D g^2 d\mu_t - \psi^2(\mu_t)$.

In this paper we consider two alternative methods for Monte Carlo estimation of $\psi(\mu_t)$. These methods are based on the techniques and results published in [1] – [4] (especially on [1, sect. 8], where the ideas of artificial interactions for linear problems are briefly discussed).

Consider the first method and turn to the description of the corresponding $\mathbf{B}^{(k)}_{\varphi}$*-algorithm.*

Firstly, we fix some $k \in \mathbb{N}$. Then we suppose that there exists a random variable ω and a function φ such that $\mathcal{L}(\varphi(\omega, u)) = T(\,\cdot\,; u)$ for any u. (Here and

[1]This work was supported by the RFFI grant N° 08-01-00194.

[2]St. Petersburg University, Russia. E-mail: vnekr@mail.ru

[3]St. Petersburg University, Russia. E-mail: nikolay_rumyanzv@mail.ru

further $\mathcal{L}(\xi)$ stands for the distribution of the random variable ξ.) Naturally, this means nothing but a concrete simulation method for the distribution $T(\,\cdot\,;u)$.

To produce consistent estimate of $\psi(\mu_t)$ we use a sort of jump-wise (n,k)-particle Markov process $\zeta_n(t) = \left(\zeta_n^{(1)}(t),\dots,\zeta_n^{(n)}(t)\right) \in D^n$ with a jump frequency n/k and an initial distribution $\mu^{\otimes n}$. As for a jump law of the process $\zeta_n(t)$, it can be described in the following manner using the language of "particles".

Let $(u_1,\dots,u_n)$ be the position of the process before a jump. We choose a random subset of $\{i_1,\dots,i_k\}$ of the set $\{1,\dots,n\}$ (each subset is chosen with the same probability) and simulate the random variable ω. Then coordinates $u_{i_1},\dots,u_{i_k}$ of the chosen particles are modified by the rule $u'_{i_j} \leftarrow \varphi(\omega,u_{i_j})$. All other particles stand still.

Of course, random variables ω are independent for different jumps of the process. The estimate has the form $w_n^{(k,\varphi)}(g,t) = \sum_{i=1}^n g(\zeta_n^{(i)}(t))/n$.

The second method (and the corresponding $\mathbf{S}_\varphi^{(k)}$-*algorithm*) has similar but different structure. For fixed k we consider m independent copies of a (k,k)-particle process, analogous to the (n,k)-particle process described above. Thus we come to m independent random variables $w_{k,j}^{(k,\varphi)}(g,t)$ analogous to $w_n^{(k,\varphi)}(g,t)$ and get the final estimate as the sample average of $w_{k,j}^{(k,\varphi)}(g,t)$.

2. Asymptotical variances and complexities

As it follows from [2], asymptotical as $n\to\infty$ features of the $\mathbf{B}_\varphi^{(k)}$-method can be investigated analytically.

Naturally, the solution μ_t of the equation (1) depends of the initial condition μ. Therefore we write down $\mu_t = \mu_t(\mu)$. Denote $\psi(\nu) = \int_D g d\nu$, $\psi(\tau,v) = \psi(\mu_\tau(\delta_v))$ (here δ_v stands for the distribution concentrated at point v), and $\Delta_\varphi(\tau,\omega,u) = \psi(\tau,\varphi(\omega,u)) - \psi(\tau,u)$.

Theorem 1. *1. The variance of the standard estimate has the representation*

$$n\mathbb{D}w_n(g,t) = \sigma_{\mathbf{S}}^2 = V_0(\psi,t,\mu) + V_{\mathbf{S}}(\psi,t,\mu),$$

where

$$V_0(\psi,t,\mu) = \int_D \psi^2(t,v)\mu(dv) - \left(\int_D \psi(t,v)\mu(dv)\right)^2$$

and $V_{\mathbf{S}}(\psi,t,\mu) = \int_0^t s_{\mathbf{S}}^2(\psi,\tau,\mu_{t-\tau}(\mu))d\tau$ *with*

$$s_{\mathbf{S}}^2(\psi,\tau,\nu) = \int_{D^2} \Big(\psi(\tau,v) - \psi(\tau,u)\Big)^2 T(dv;u)\nu(du).$$

2. In the case of the $\mathbf{B}_\varphi^{(k)}$*-estimate,* $\mathbb{E}w_n^{(k,\varphi)}(g,t) = \psi(\mu_t)$ *and under restrictions of [1, sect. 8],*

$$n\left(\mathbb{D}w_n^{(k,\varphi)}(g,t) - \mathbb{D}w_n(g,t)\right) \to (k-1)V_{\mathbf{B}_\varphi} \tag{2}$$

as $n \to \infty$, where

$$V_{\mathbf{B}_\varphi}(\psi, t, \mu) = \int_0^t d\tau \int_{D^2} \mu_{t-\tau}(du_1)\mu_{t-\tau}(du_2)\, \mathbb{E}\, \Delta_\varphi(\tau, \omega, u_1)\Delta_\varphi(\tau, \omega, u_2).$$

Remark 1. *1. The proof of this proposition in based on [1, th. 6.2].*
2. Since $0 < V_{\mathbf{B}_\varphi} \le V_{\mathbf{S}}$, the variance of $\mathbf{B}_\varphi^{(k)}$-estimate asymptotically in n exceeds the variance of the standard estimate not more that for k times. If $k = 1$, then these variances coincide.

Note that the average number of jumps of $\mathbf{B}_\varphi^{(k)}$-process is k times less than the number of jumps necessary to get the standard estimate $w_n(g, t)$. Therefore we get a hope that the Monte Carlo complexity of the estimate $w_n^{(k,\varphi)}(g, t)$ can be less than the complexity of the standard estimate.

Let us give the formal definition of the (asymptotical) complexity of $\mathbf{B}_\varphi^{(k)}$-algorithm. As usually, this complexity must be proportional to the variance of $w_n^{(k,\varphi)}(g, t)$ as well to the average time necessary to calculate this estimate. Let t be fixed. Denote $c(\omega)$, $c(\varphi)$, $c(p \,|\, k, n)$ and $c(\lambda)$ the complexities of a) simulation of the random variable ω, b) calculation of the function φ, c) simulation of the random subset $\{i_1, \dots, i_k\}$ of the set $\{1, \dots, n\}$, and d) simulation of the exponential distribution $\mathrm{EXP}(\lambda)$ correspondingly.

Additionally we suppose that $c(p \,|\, k, n)$ is proportional to k and does not depend on n at least for big n. In other words, we suppose that $c(p \,|\, k, n) = kc(p)$.

Finally, let $\Im_0$ denote the complexity of simulation of the initial distribution μ and let $n\Im_g$ stands for the calculation complexity of the function $\sum_{i=1}^n g(x_i)/n$.

Definition 1. *The complexity of $\mathbf{B}_\varphi^{(k)}$-algorithm (reduced to the number of particles) is defined as*

$$L_{\mathbf{B}_\varphi}(k) = \left(\frac{c(\omega) + kc(\varphi) + kc(p) + c(\lambda)}{k}\, t + \Im_0 + \Im_g \right) \Big(V_0 + V_{\mathbf{S}} + (k-1)V_{\mathbf{B}_\varphi} \Big).$$

Analogously, the complexity of the standard algorithm is defined by the formula $L_{\mathbf{S}} = \big((c(\omega) + c(\varphi) + c(\lambda))t + \Im_0 + \Im_g\big)(V_0 + V_{\mathbf{S}})$.

Remark 2. *1. The minimal complexity of the $\mathbf{B}_\varphi^{(k)}$-algorithm corresponds to*

$$k_{opt}^2 \approx \frac{c(\omega) + c(\lambda)}{c(\varphi) + c(p) + \Im_0/t + \Im_g/t} \left(\frac{V_0 + V_{\mathbf{S}}}{V_{\mathbf{B}_\varphi}} - 1 \right). \tag{3}$$

2. Note that V_0, $V_{\mathbf{S}}$, and $V_{\mathbf{B}_\varphi}$ depend on t.
3. It can be easily seen that $L_{\mathbf{B}_\varphi}(1) > L_{\mathbf{S}}$.

3. Examples

Consider the equation (1) with $D = \mathbb{R}$, $T(\,\cdot\,; u) = \mathrm{N}(u, 1)$ and the initial distribution $\mu = \mathrm{N}(0, 4)$. Additionally, we take two variants of the function g for the linear functional ψ: $g_1(u) = u^2$ and $g_2(u) = \cos u$. Thus we get two linear functionals ψ_1 and ψ_2 dependent on t.

3.1. $\mathbf{B}_{\varphi}^{(k)}$-algorithm

Simple calculations show that for $g(u) = u^2$ we obtain $\psi_1(\mu_t) = t + 4$, $V_0^{(1)} = 32$, $V_{\mathbf{S}}^{(1)} = 2t^2 + 19t$, and $V_{\mathbf{B}_\varphi}^{(1)} = 3t$.

If $g(u) = \cos u$, then $\psi_2(\mu_t) = e^{-(1-1/\sqrt{e})t-2}$, $V_0^{(2)} = e^{-2\left(1-1/\sqrt{e}\right)t}\left(1-e^{-4}\right)^2/2$, $V_{\mathbf{B}_\varphi}^{(2)} = t\, e^{-2\left(1-1/\sqrt{e}\right)t-4}\left(3-4e^{-1/2}+e^{-2}\right)/2$, and

$$V_{\mathbf{S}}^{(2)} = \left(e^{-(1-1/e^2)t-8} - e^{-2(1-1/\sqrt{e})t}\left(1+e^{-8}\right)+1\right)/2.$$

Note that formal technical restrictions of [1, sect. 8] are not totally satisfied for these examples. Still computational experiments show that the results of Theorem 1 are still valid here.

Table 1: Quotients $L_{\mathbf{S}}/L_{\mathbf{B}_\varphi}(k)$ and τ_S/τ_{B_φ} for ψ_1 and ψ_2

Functional ψ_1				Functional ψ_2		
t	5	10	15	t	3	7
k	5	6	7	k	23	82
$L_{\mathbf{S}}/L_{\mathbf{B}_\varphi}$	1.36	1.51	1.62	$L_{\mathbf{S}}/L_{\mathbf{B}_\varphi}$	1.95	2.58
τ_S/τ_{B_φ}	1.33	1.39	1.46	τ_S/τ_{B_φ}	1.92	2.34

Table 2: The quotient τ_S/τ_{B_φ}. Timing results for ψ_1, ψ_2 and different k, t

Functional ψ_1											
k	1	2	3	4	5	6	7	8	9	10	25
$t=5$	0.79	1.12	1.25	1.31	1.33	1.31	1.30	1.26	1.24	1.20	0.79
$t=10$	0.80	1.13	1.28	1.35	1.40	1.39	1.37	1.34	1.32	1.29	0.89
$t=15$	0.72	1.08	1.27	1.37	1.43	1.46	1.46	1.44	1.43	1.42	1.03
Functional ψ_2											
k	1	10	15	20	23	25	30	35	40	60	90
$t=3$	0.84	1.85	1.91	1.92	1.92	1.91	1.91	1.89	1.88	1.79	1.67
k	1	30	60	75	85	90	100	120	150	200	250
$t=5$	0.76	2.24	2.31	2.34	2.34	2.34	2.35	2.35	2.33	2.31	2.29

Since $T(\,\cdot\,,\,,u) = \mathrm{N}(u,1)$, then $\varphi(x,y) = x+y$ and $\omega \in \mathrm{N}(0,1)$. In our computational experiments the normal distribution is simulated by the modified polar method. Also, $c(\lambda)$ corresponds to the standard simulation of the exponential distribution and the choice of a subset $\{i_1,\dots,i_k\}$ of the set $\{1,\dots,n\}$ is produced by the swapping method (e.g., [5, ch. 12]). The corresponding parameters $c(\varphi)$, $c(\omega)$, $c(p)$ and $c(\lambda)$ are estimated by separate computational experiments. (Of course, these parameters depend on the features of the concrete computer, etc.)

Having these characteristics at hand we can calculate the complexity $L_{\mathbf{B}_\varphi}(k)$ (as well as $L_{\mathbf{S}}$) and find $k_{opt} = k_{opt}(t)$. It follows from (3) that $k_{opt}(5) = 5$, $k_{opt}(10) = 6$, and $k_{opt}(15) = 7$ for the functional ψ_1. Analogously, $k_{opt}(3) = 22$ and $k_{opt}(7) = 82$ for ψ_2. The quotients $L_{\mathbf{S}}/L_{\mathbf{B}_\varphi}(k)$ for these $k = k_{opt}$ are placed in Table 1.

The last row of Table 1 presents the analogous timing results for the same k and t. The average time necessary to calculate the standard estimate corresponds to τ_S, while τ_{B_φ} has the same meaning for $\mathbf{B}_\varphi^{(k)}$-estimate. Note that τ_S and τ_{B_φ} are measured under the restriction that the variances of both estimates coincide. This restriction is provided by the choice of the appropriate parameter $n = n(k,t)$ for $\mathbf{B}_\varphi^{(k)}$-algorithm. To get values of τ_S and τ_{B_φ}, we used $5 \cdot 10^4$ trajectories of $\mathbf{B}_\varphi^{(k)}$-process and approximately $5 \cdot 10^7$ trajectories of $\mathbf{S}$-process.

Table 1 shows that the quotients $L_{\mathbf{S}}/L_{\mathbf{B}_\varphi}(k)$ and τ_S/τ_{B_φ} are in good (though not ideal) correspondence. Note that in all cases both quotients exceed 1. This means that $\mathbf{B}_\varphi^{(k)}$-estimate gives better result than the standard estimate.

Table 2 is analogous to the last row of Table 1 but displays wide intervals of the parameter k. It is important to mention that maximal values of τ_S/τ_{B_φ} are next to theoretically optimal k.

3.2. $\mathbf{S}_\varphi^{(k)}$-algorithm

For $\mathbf{S}_\varphi^{(k)}$-estimate, we have no theoretical results analogous to (2). On the other hand this estimate is unbiased and we can calculate its sample variance in the same manner as for the standard estimate. (Generally, this is impossible for $\mathbf{B}_\varphi^{(k)}$-estimates.)

Table 3: The quotient τ_S/τ_{S_φ}. Timing results for ψ_1, ψ_2 and different k, t

Functional ψ_1									
k	2	3	4	5	6	7	8	9	10
$t = 5$	1.29	1.34	1.31	1.27	1.22	1.17	1.11	1.05	1.01
$t = 15$	1.17	1.18	1.13	1.10	1.05	1.01	0.98	0.94	0.89
Functional ψ_2									
k	5	10	15	20	25	30	40	60	120
$t = 3$	2.52	3.01	3.02	3.01	2.88	2.81	2.57	2.22	1.53
k	10	30	45	60	70	80	100	120	150
$t = 7$	4.14	4.14	3.77	3.40	3.25	3.00	2.72	2.45	2.11

Table 3 is quite analogous to Table 2 and presents timing results for $\mathbf{S}_\varphi^{(k)}$-algorithm. These results show that the choice of the appropriate k gives the preference to the $\mathbf{S}_\varphi^{(k)}$-method against the standard method. Of course, here optimal values of k can differ from those of Table 2.

4. Summary

Examples show that both variants of artificial Monte Carlo interactions can be more effective than the standard **S**-method. In view of (3), this effectiveness enlarges if simulation complexity of the random variable ω is big and if φ is a "simple" function. Besides, $\mathbf{S}_{\varphi}^{(k)}$-algorithm allows to calculate confidence intervals for the functional under estimation.

References

[1] Golyandina, N. and Nekrutkin, V. (1999), Homogeneous balance equations for measures: errors of the stochastic solution, *Monte Carlo Methods and Appl.*, V.5, 3, 1 – 67.

[2] Golyandina, N. and Nekrutkin, V. (2000), Estimation errors for functionals on measure spaces, In: *Stochastic simulation methods*, Eds.: N. Balakrishnan, S. Ermakov and V. Melas, Birkhauser, Boston-Basel-Berlin, 29 – 46.

[3] Golyandina, N. (2003), Central Limit Theorem for (n, k)-particle processes solving balance equations, *Monte Carlo Methods and Appl.*, V.9, 1, 1 – 11.

[4] Nekrutkin, V. and Potapov, P. (2004), Two variants of a stochastic Euler method for homogeneous balance differential equations, *Monte Carlo Methods and Appl.*, V.10, 3-4, 469 – 480.

[5] Devroy, L. (1986), *Non-Uniform Random Variate Generation*, Springer-Verlag, New York-Berlin-Heidelberg-Tokyo.

6th St.Petersburg Workshop on Simulation (2009) 947-951

Monte-Carlo Simulations for the Multi-Armed Bandit Problem

A.V. Kolnogorov[1], T.N. Shelonina[2]

Abstract

The multi-armed bandit problem is considered in minimax setting on the finite sufficiently large horizon N. An effective threshold strategy is proposed which provides asymptotically unimprovable in order of magnitude maximal values of the function of losses. The invariance property of the strategy allows to improve Monte-Carlo simulations for determination the optimal values of thresholds.

1. Introduction

The multi-armed bandit is a slot machine with K arms ($K \geq 2$) (see, e.g. [1]). If the player chooses the ℓ-th arm of the slot at the point of time n ($n = 1, 2, \ldots, N$) he gets random income ξ_n which value depends on the chosen arm only and is equal to 1, 0 with probabilities p_ℓ, q_ℓ, $\ell = 1, \ldots, K$. Hence, the multi-armed bandit can be described by a vector parameter $\theta = (p_1, \ldots, p_K)$. This parameter is unknown to the player. It is assumed that it can take any value from the set $\Theta = \{\theta : 0 \leq p_\ell \leq 1, \ell = 1, \ldots, K\}$ which is geometrically equivalent to the K-dimensional unit cube.

The goal of the player is to maximize (in some sense) his total expected income. To achieve the goal, the player uses the strategy σ which at each point of time n can use all current history of the process $y^{n-1}, \xi^{n-1} = (y_1, \xi_1, \ldots, y_{n-1}, \xi_{n-1})$, i.e.

$$\Pr\{y_n = \ell | y^{n-1}, \xi^{n-1}\} = \sigma_\ell(y^{n-1}, \xi^{n-1}), \quad \ell = 1, \ldots, K.$$

If parameter θ is known to the player his optimal strategy is always to use the alternative corresponding to the largest value of $p_1, \ldots, p_K$, and his expected total income is thus $N \max_{\ell=1,\ldots,K} p_\ell$. Since actually parameter is unknown his total expected income is $E_{\sigma,\theta}\left(\sum_{n=1}^{N} \xi_n\right)$ where $E_{\sigma,\theta}$ stands for mathematical expectation over the measure generated by strategy σ and parameter θ. The difference

$$L_N(\sigma, \theta) = N \max_{\ell=1,\ldots,K} p_\ell - E_{\sigma,\theta}\left(\sum_{n=1}^{N} \xi_n\right)$$

[1]Novgorod State University, E-mail: Alexander.Kolnogorov@novsu.ru
[2]Novgorod State University, E-mail: Tatyana.Shelonina@novsu.ru

describes losses of the player due to the lack of complete information. The corresponding minimax risk is defined as

$$R_N(\Theta) = \inf_{\{\sigma\}} \sup_{\Theta} L_N(\sigma, \theta).$$

A well known asymptotic estimates for the minimax risk in case $K = 2$ are given in [2] and state that asymptotically (as $N \to \infty$)

$$0.265N^{1/2} \leq R_N(\Theta) \leq 0.370N^{1/2}.$$

The upper estimate was proved in [2] for the following threshold strategy. The alternatives should be applied by turn until the time of control exceeds the horizon N or the absolute difference of total current incomes exceeds the magnitude $0.29N^{1/2}$. If this value of difference is achieved and the total time of control does not expire the alternative corresponding to the larger total income should be applied up to the end of control.

In this paper the threshold strategy is generalized for the multi-armed bandit problem with $K > 2$. The simplified threshold strategy in this case was proposed in [3]. In [4] some estimates are given which show that the maximal value of the function of losses has the order of $N^{1/2}$. The main objective of this paper is to give an effective algorithm of numerical determination of the optimal values of thresholds.

There are several other approaches to the problem depending on its possible applications. Accordingly, the problem is interpreted as the adaptive control problem or the problem of expedient behavior in random media. For example, asymptotically optimal procedure can be used, which makes estimates of probabilities $p_1, \ldots, p_K$, and then preferably applies the alternative corresponding to the current largest value of these estimates [5]. To describe the expedient behavior of the simplest organisms in random media, models based on finite automata are used [6]. In [1] the Bayesian approach to the problem is considered. Many famous examples for particular prior distributions are considered there.

The structure of the paper is the following. In section 2 formal description of the threshold strategy is given. In section 3 an invariance property of the threshold strategy is considered. In section 4 application of the invariance property to effective numerical calculation of the optimal values of thresholds is discussed,

2. Threshold strategy for the multi-armed bandit with $K > 2$

First, let's define a procedure of rejecting "one the worst alternative" from the group of M alternatives which is applied on the horizon T. In words it can be described as follows. Suppose that some initial incomes corresponding to all alternatives are given. Alternatives should be applied by turn until the total time of control T expires or the difference of maximal and minimal total current incomes exceeds the magnitude a_M. If the time of control does not expire the alternative

corresponding to the minimal total current income is assumed to be "the worst" one and rejected.

Formally, assume that there are M alternatives ($2 \le M \le K$) numbered as $\ell_1, \dots, \ell_M$ and corresponding initial incomes $S_{\ell_1}, \dots, S_{\ell_M}$. Let's apply alternatives by turn, i.e. $y(n) = \ell_i$ at $n = M\tau + i$, accumulating total current incomes

$$S_{\ell_1}(t) = S_{\ell_1} + \sum_{\tau=0}^{t-1} \xi_{M\tau+1}, \dots, S_{\ell_M}(t) = S_{\ell_M} + \sum_{\tau=0}^{t-1} \xi_{M\tau+M},$$

until $t_0 = \min(t^*, [T/M])$ where t^* is the first number when inequality holds true

$$\max_{\ell_k} S_{\ell_k}(t^*) - \min_{\ell_k} S_{\ell_k}(t^*) \ge a_M > 0.$$

Here $\{\xi_{M\tau+i}\}$ are incomes which were got for the ℓ_i-th alternative application ($i = 1, \dots, M$). If $Mt^* < T$ the ℓ_{i_0}-th alternative

$$\ell_{i_0} = \mathrm{argmin}_{\ell_i} S_{\ell_i}(t_0)$$

is removed from the initial group of alternatives. If $M = 1$ we'll treat that procedure applies the single alternative all the time T.

Second, let's define a sequential procedure of rejecting "all the worst alternatives" from the group of M alternatives numbered as $\ell_1, \dots, \ell_M$ on the horizon T. This procedure at first applies previously described procedure of rejecting "one the worst alternative". If the total time of control does not expire and the ℓ_{i_0}-th alternative has been rejected then on the rest horizon $T - Mt_0$ this sequential procedure is applied to the remaining $M - 1$ alternatives with new initial incomes $\{S_{\ell_i} = S_{\ell_i}(t_0)\}$, $i \ne i_0$.

The threshold strategy can be defined as a procedure of sequential rejecting "all the worst alternatives" applied to initial K alternatives on the horizon N. Note, that the threshold strategy does not change if all total incomes vary by the same value. In the sequel it would be convenient to think that the threshold strategy depends on the differences $X_\ell(t) = p \cdot t - S_\ell(t)$, $\ell = 1, \dots, K$ for some p. If $p \ge p_\ell$, $\ell = 1, \dots, K$, this differences can be considered as losses with respect to incomes $p \cdot t$.

In the sequel we use notations $a_{i..j}$ for the interval of sequence $a_i, \dots, a_j$ ($i \le j$) and $a_{i..j}^{-k}$ for this interval of sequence without a_k, i.e. $a_i, \dots, a_{k-1}, a_{k+1} \dots, a_j$ ($i \le k \le j$). It is assumed that $a_i, \dots, a_{k-1} = \emptyset$ if $i = k$ and $a_{k+1}, \dots, a_j = \emptyset$ if $k = j$. Let's denote by $\sum_{n=1}^{N} E(\xi_n | a_{2..K}; p_{1..K}; Y_{1..K})$ the expected total income for threshold strategy provided the magnitudes of thresholds are $a_{2..K}$ and initial values of differences $X_{1..K}$ are $Y_{1..K}$. We'll consider the function of losses comparatively to the total income Np for some p ($p \ge p_\ell$, $\ell = 1, \dots, K$) which is equal to

$$L_N(a_{2..K}; p, p_{1..K}; Y_{1..K}) = \sum_{n=1}^{N} E(p - \xi_n | a_{2..K}; p_{1..K}; Y_{1..K}). \qquad (1)$$

This definition is convenient for analyzing the losses of defined procedures in case the alternative corresponding to the largest value of $p_1, \dots, p_K$ has been already rejected.

3. An invariance property of the function of losses

Denote $D = p(1-p)$, $0 < p < 1$, and let $p_\ell = p - \beta_\ell (D/N)^{1/2}$, $\beta_\ell \geq 0$, $\ell = 1, \ldots, K$. It is assumed that N is sufficiently large, so that $p_\ell \geq 0$, $\ell = 1, \ldots, K$. Let $a_\ell = \alpha_\ell (DN)^{1/2}$ with $\alpha_\ell > 0$, $\ell = 2, \ldots, K$ and $Y_\ell = \gamma_\ell (DN)^{1/2}$, $\ell = 1, \ldots, K$. An invariance property of the strategy is the following.

Theorem 1. *Under above assumptions for* $0 < \rho \leq 1$ *the limiting equality holds*

$$\lim_{N\to\infty} (DN)^{-1/2} L_{\rho N}(a_{2..K}; p, p_{1..K}; Y_{1..K}) = \mathcal{L}^K(\alpha_{2..K}; \beta_{1..K}; \gamma_{1..K}; \rho). \quad (2)$$

As a result of the definition (1) and the invariance property (2) the following corollary holds.

Corollary 1. *The equalities hold*

$$\mathcal{L}^K(\alpha_{2..K}; \beta_{1..K}; \gamma_{1..K}; \rho) = \quad (3)$$

$$= \rho^{1/2} \mathcal{L}^K(\alpha_2 \rho^{-1/2}, \ldots, \alpha_K \rho^{-1/2}; \beta_1 \rho^{1/2}, \ldots, \beta_K \rho^{1/2}; \gamma_1 \rho^{-1/2}, \ldots, \gamma_K \rho^{-1/2}; 1),$$

$$\mathcal{L}^K(\alpha_{2..K}; \beta_{1..K}; \gamma_{1..K}; 1) = \mathcal{L}^K(\alpha_{2..K}; \beta_1 - \beta, \ldots, \beta_K - \beta; \gamma_{1..K}; 1) + \beta, \quad (4)$$

if $\beta \leq \min_\ell \beta_\ell$.

$$\mathcal{L}^K(\alpha_{2..K}; \beta_{1..K}; \gamma_{1..K}; 1) = \mathcal{L}^K(\alpha_{2..K}; \beta_{1..K}; \gamma_1 - \gamma, \ldots, \gamma_K - \gamma; 1) + K\gamma. \quad (5)$$

Equalities (3), (4) and (5) allow to calculate $\mathcal{L}^K(\alpha_{2..K}; \beta_{1..K}; \gamma_{1..K}; \rho)$ at $\rho = 1$, $\beta_1 = 0$, $\gamma_1 = 0$ only. The idea of the proof of Theorem 1 is the following. Obviously, (2) is true at $K = 1$ and

$$L^1(\cdot; \beta_1; \gamma_1, \rho) = \beta_1 \rho + \gamma_1.$$

Then (2) can be checked by induction. If $K \geq 2$ and $\rho = 1$ then recurrent equation holds

$$L_N(a_{2..K}; p, p_{1..K}; Y_{1..K}) = \quad (6)$$

$$= E\left(\sum_{t=1}^{t_0} \sum_{\ell=1}^{K} X_\ell(t) + L_{N-Kt_0}(a_{2..K-1}; p, p^{-j}_{1..K}; X^{-j}_{1..K}(N - Kt_0)) \right),$$

where j is random alternative removed at the point of time Kt_0. Denote by $P(X_{1..K}; t)$ the joint distribution of losses $X_1, \ldots, X_K$ at the point of time Kt. Let $x_\ell = X_\ell (DN)^{-1/2}$, $\ell = 1, \ldots, K$, $\tau = tN^{-1}$. Then on the initial stage of strategy application (i.e. before "the worst alternative" is rejected) $N^{-1} P(X_{1..K}; t)$ weakly converges to the joint distribution density $f(x_{1..K}; \alpha_{2..K}; \beta_{1..K}; \gamma_{1..K}; \tau)$ which satisfies partial differential equation

$$\frac{\partial f}{\partial \tau} = -\sum_{\ell=1}^{K} \beta_\ell \frac{\partial f}{\partial x_\ell} + \frac{1}{2} \sum_{\ell=1}^{K} \frac{\partial^2 f}{\partial x_\ell^2} \quad (7)$$

with initial condition

$$f(x_{1..K};\alpha_{2..K};\beta_{1..K};\gamma_{1..K};0)=\delta(x_1-\gamma_1)\cdot\dots\cdot\delta(x_K-\gamma_K) \tag{8}$$

at $\tau=0$ and boundary conditions

$$f(x_{1..K};\alpha_{2..K};\beta_{1..K};\gamma_{1..K};\tau)=0,\quad (x_1,\dots,x_K)\in G_{ij}, \tag{9}$$

at $0<\tau\le 1/K$, where $G_{ij}=\{(x_1,\dots,x_K): x_i=\min_\ell x_\ell, x_j=\max_\ell x_\ell, x_j-x_i=\alpha_K\}$; $i,j=1,\dots,K$; $i\ne j$.

Equation (7) with initial condition (8) and boundary conditions (9) describes K-dimensional continuous time random walking starting from the source in $(\gamma_1,\dots,\gamma_K)$ and absorbing at bounds $\{G_{ij}\}$. Denote by $G=\{(x_1,\dots,x_K): |x_i-x_j|<\alpha_K; i,j=1,\dots,K; i\ne j\}$ a domain in which random walking continues without absorbtion. Then (6) results to the following expression

$$\mathcal{L}^K(\alpha_{2..K};\beta_{1..K};\gamma_{1..K};1)=$$

$$=\int_G (x_1+\dots+x_K) f\left(x_{1..K};\alpha_{2..K};\beta_{1..K};\gamma_{1..K};\frac{1}{K}\right)(dx)_{1..K}+$$

$$+\sum_{i=1}^{K}\sum_{j\ne i}\int_0^{\frac{1}{K}} I_{ij}(\alpha_{2..K};\beta_{1..K};\gamma_{1..K};\tau)d\tau,$$

where $(dx)_{1..K}=dx_1\dots dx_K$. Here

$$I_{ij}(\alpha_{2..K};\beta_{1..K};\gamma_{1..K};\tau)=\int_{G_{ij}} r_{ij}(x_{1..K};\alpha_{2..K};\beta_{1..K};\gamma_{1..K};\tau)\cdot$$

$$\cdot\left(x_1+\dots+x_K+\mathcal{L}^{K-1}(\alpha_{2..K-1};\beta^{-j}_{1..K};x^{-j}_{1..K};1-K\tau)\right)\Big|_{x_j=x_i+\alpha_K}(dx)^{-j}_{1..K},$$

where $(dx)^{-j}_{1..K}=dx_1\dots dx_{j-1}dx_{j+1}\dots dx_K$ and $r_{ij}(x_{1..K};\alpha_{2..K};\beta_{1..K};\gamma_{1..K};\tau)$ is the density of absorbtion at the bound G_{ij} at the point of time $K\tau$.

4. Simulations

Simulations were based on recurrent equation (6) and invariance property (2). Probability p was chosen equal to 0.5 since this value corresponds to maximal losses according to (2). Random losses were simulated and accumulated on the initial stage of threshold strategy only (i.e. $n\le Kt_0$) and then expected losses on the remaining horizon $N-Kt_0$ were added to accumulated ones using the property (2). If $K=2$ then optimal magnitude of threshold is $\alpha_2\approx 0.29$. If $K=3$ then optimal magnitudes of thresholds are $\alpha_3\approx 0.20$, $\alpha_2\approx 0.28$.

References

[1] Berry D.A., Fristedt B. (1985) *Bandit Problems: Sequential Allocation of Experiments*. Chapman and Hall, London, New York.

[2] Vogel W. (1960) An asymptotic minimax theorem for the two-armed bandit problem. *Ann. Math. Stat.*, **31**, 444–451.

[3] Kolnogorov A.V. (1991) Behavior strategy in a stationary environment with an unimprovable guaranteed mean-income convergence bound. *Automation and Remote Control*, **52**, No 5, 183–186. (Translated from Russian).

[4] Kolnogorov A.V., Shelonina T,N. (2007) A threshold control strategy in a random environment. *Proc. 6th Intern. Conf. SICPRO'07*. Moscow. Institute of Control Sciences. Russian Academy of Sciences. ISBN - 5-201-14992-8, 1500–1512. (In Russian)

[5] Sragovich V. G. (2006) *Mathematical Theory of Adaptive Control. Interdisciplinary Mathematical Ssciences – Vol. 4*. World Scientific. New Jersey, London, Singupore, Beijing, Shanghai, Hong Kong, Taipei, Chennai.

[6] Tsetlin M.L. (1973) *Automation Theory and Modeling of Biological Systems*. Academic Press, New York. (Translated from Russian).

6th St.Petersburg Workshop on Simulation (2009) 953-957

The generalized geometric distribution and the associated Galton-Watson model with a medical statistical application

Nina Alexeyeva[1]

Abstract

This article contains a comparison of two types of generalized geometric distributions based on their association with the multi-type Galton-Watson model of branching processes. Difference between them is interpreted in the view of learnable and trained systems by the example of a cold test data before coronary artery bypass grafting (CABG) in patients with a radial artery (RA) as a conduit with its intraoperative spasm and without it.

1. Introduction

The generalized binomial positive and negative distributions were introduced by means of partially inversion of functions [1], [2] and associated [2] with the multi-type Galton-Watson model of branching processes [4]. These distributions were supposed to be used in the study of regulations in medical and biological systems but in practice can be rarely seen. In this article a new modification of the generalized geometrical distribution is introduced and used successfully in statistical analysis of the medical data. This modification is connected with the round-off error in generalized binomial distributions and resulted in different number of particles for the associate Galton-Watson model and another interpretations.

2. The generalized geometrical distribution

Consider a random number of successes in n Bernoulli tests with probability p, denoted by $k = \xi(n)$, as a function of n. The left and right inverses of ξ are correspondingly [1]

$$\begin{aligned} \xi_0^-(k) &= min\{n : \xi(n) \geq k\}, \\ \xi_1^-(k) &= max\{n : \xi(n) \leq k\}. \end{aligned}$$

[1]Saint Petersburg State University, E-mail: ninaalexeyeva@mail.ru

At $k = 0$ the left inverse $\xi_0^-(0)$ is trivial and equals 0 with probability 1 and the right inverse $\xi_1^-(0)$ means the maximal number of failures before the first success and has the geometric distribution

$$P\{\xi_1^-(0) = j\} = pq^j = q^j - q^{j+1}, \quad j = 0, 1, 2, \ldots.$$

The Bart's partially inverse [1],[2],[2] with the parameter $\alpha \in (0, 1]$ was introduced as a function which possesses the value c between the values of the right b and the left a inverses and $(c - a) = \alpha(b - a)$.

Particularly at $k = 0$ the partially inverse $\alpha\xi_1^-(0)$ is not integer mainly and so it is necessary to rounding off it. The random variables $\lceil \alpha\xi_1^-(0) \rceil$ and $\lfloor \alpha\xi_1^-(0) \rfloor$ are considered the right and left partially inverses. They have the generalized geometric distribution of A-type and of B-type correspondingly, denoted by $\beta_-^A(\cdot|p, \alpha)$ and $\beta_-^B(\cdot|p, \alpha)$, $0 < \alpha \leq 1$. One of them, $\beta_-^B(\cdot|p, \alpha)$, was presented in [1]:

$$P\{\lfloor \alpha\xi_1^-(0) \rfloor = j\} = q^{\lceil \frac{j}{\alpha} \rceil} - q^{\lceil \frac{j+1}{\alpha} \rceil}, \quad j = 0, 1, 2, \ldots.$$

At $\alpha = 1/m$, $m = 1, 2, \ldots$, we have the geometric distribution $\beta_-^B(\cdot|p, 1/m) = \beta_-(\cdot|1 - (1 - p)^m)$.

The distribution $\beta_-^B(\cdot|p, \alpha)$ underlies the generalized positive $\beta_+(\cdot|n, p, \alpha)$ and negative $\beta_-(\cdot|k, p, \alpha)$ distributions [2]. $\beta_-(\cdot|k, p, \alpha)$ was entered as a distribution of the sum $\sum_{i=1}^{k} \xi_i$ where ξ_i has $\beta_-^B(\cdot|p, \alpha)$. $\beta_+(\cdot|n, p, \alpha)$ was defined as fiducial distribution: $\beta_+^<(k|n, p, \alpha) = 1 - \beta_-^\leq(n|k, p, \alpha)$. An explicit form of these distributions is complicated but they have quite simple interpretation by means of the random walk on integer squares originating from $(0, 0)$.

A failure is represented by increasing x by 1 (probability q) and a success by increasing y by 1 (probability $p = 1 - q$). The number of iterations until the first contact to the line $x + y = n$ has the binomial distribution and the number iterations before the first contact to the line $y = 1$ has the geometric distribution. Probabilities of failures before the first success generate homogeneous sequence $q, q, q, q, \ldots$. Evidently this random walk can be described by means of the recurrence relation of generating functions $f_n(\nu)$ of binomial distribution $\beta_+(\cdot|n, p, \alpha = 1)$:

$$f_{n+1}(\nu) = (p\nu + q)f_n(\nu) = p\nu f_n(\nu) + qf_n(\nu), \quad f_0(\nu) = 1.$$

At $\alpha = \frac{s}{m}$ the generalized binomial distribution $\beta_+(\cdot|n, p, \alpha)$ corresponds another random walk [2]. Consider for example $\alpha = \frac{2}{2r+1}$ which is used in the application. The basic distribution $\beta_-^B(\cdot|p, \frac{2}{2r+1})$ has a periodic property:

$$\begin{aligned} P\{\lfloor \alpha\xi_1^-(0) \rfloor &= 2k\} = q^{k(2r+1)}(1 - q^{r+1}), \\ P\{\lfloor \alpha\xi_1^-(0) \rfloor = 2k + 1\} &= q^{k(2r+1)}q^{r+1}(1 - q^r), \end{aligned}$$

where $k = 0, 1, 2, \ldots$. Hence probabilities of failure generate a periodical sequence $q_1, q_2, q_1, q_2, \ldots$ (the success probability is equal $p_i = 1 - q_i$, $i = 1, 2$) where $q_1 =$

q^{r+1}, $q_2 = q^r$ (Fig.1). If the structure of the random walk is repeated after the success then it can be described by means of two generating functions:

$$\begin{cases} f_{n+1}^{(1)}(\nu) = p_1 \nu f_n^{(1)}(\nu) + q_1 f_n^{(2)}(\nu), \\ f_{n+1}^{(2)}(\nu) = p_2 \nu f_n^{(1)}(\nu) + q_2 f_n^{(1)}(\nu), \end{cases} \tag{1}$$

where $f_0^{(1)}(\nu) = f_0^{(2)}(\nu) = 1$, $f_n^{(1)}(\nu)$ and $f_n^{(2)}(\nu)$ are the generating functions of the random number of successes until the first contact to the line $x + y = n$ from the point (0,0) with probabilities of failure $q_1, q_2, q_1, q_2, \ldots$ and $q_2, q_1, q_2, q_1, \ldots$ correspodingly. But in second case the sequence $q_1, q_2, q_1, q_2, \ldots$ is reconstructed after the success. According to [2] $f_n^{(1)}(\nu)$ is the generating function of $\beta_+(\cdot|n, p, \alpha = \frac{2}{2r+1})$.

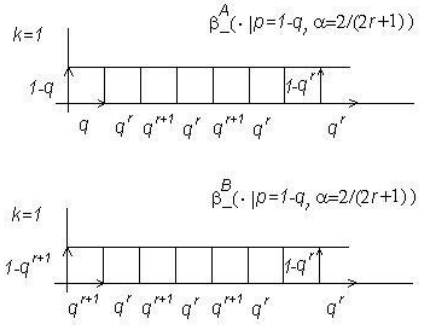


Figure 1: An interpretation of the generalized geometric distributions as random walk on integer squares.

Consider the right partially inverse $\lceil \alpha \xi_1^-(0) \rceil$ detailed.

Theorem 1. $P\{\lceil \alpha \xi_1^-(0) \rceil = 0\} = 1 - q$, $P\{\lceil \alpha \xi_1^-(0) \rceil = j\} = q^{\lfloor \frac{j-1}{\alpha} \rfloor + 1} - q^{\lfloor \frac{j}{\alpha} \rfloor + 1}$, $j = 1, 2, \ldots$.

Proof. Write the distribution function of $\xi_1^-(0)$

$$P\{\xi_1^-(0) \le x\} = \sum_{j=0}^{\lfloor x \rfloor} P\{\xi_1^-(0) = j\} = 1 - q^{\lfloor x \rfloor + 1}.$$

If $0 < \alpha \le 1$ then $P\{\lceil \alpha \xi_1^-(0) \rceil = j\} = P\left\{\frac{j-1}{\alpha} < \xi_1^-(0) \le \frac{j}{\alpha}\right\} = q^{\lfloor \frac{j-1}{\alpha} \rfloor + 1} - q^{\lfloor \frac{j}{\alpha} \rfloor + 1}$. Separately $P\{\lceil \alpha \xi_1^-(0) \rceil = 0\} = P\{\xi_1^-(0) = 0\} = 1 - q$. □

Corollary 1. *Denote* $\zeta = \lceil \alpha \xi_1^-(0) \rceil$ *at* $\alpha = \frac{2}{2r+1}$. *Then by* $k = 1, 2, \ldots$

$$\begin{aligned} P\{\zeta = 2k-1\} &= q^{k(2r+1)-2r}(1 - q^r), \\ P\{\zeta = 2k\} &= q^{k(2r+1)-2r} q^r (1 - q^{r+1}) \end{aligned}$$

and probabilities of failures generate a sequence: $q_0, q_1, q_2, q_1, q_2, \ldots$, *where* $q_0 = q$, $q_1 = q^r$ *and* $q_2 = q^{r+1}$.

Consider the random walk on integer squares (Fig.1) which is described by means of three generating functions:

$$\begin{cases} f_{n+1}^{(0)}(\nu) = p_0 \nu f_n^{(0)}(\nu) + q_0 f_n^{(1)}(\nu), \\ f_{n+1}^{(1)}(\nu) = p_1 \nu f_n^{(0)}(\nu) + q_1 f_n^{(2)}(\nu), \\ f_{n+1}^{(2)}(\nu) = p_2 \nu f_n^{(0)}(\nu) + q_2 f_n^{(1)}(\nu), \end{cases} \tag{2}$$

where $f_0^{(0)}(\nu) = f_0^{(1)}(\nu) = f_0^{(2)}(\nu) = 1$ and $f_n^{(0)}(\nu)$, $f_n^{(1)}(\nu)$, $f_n^{(2)}(\nu)$ are the generating functions of the random number of successes until the first contact to the line $x + y = n$ from the point (0,0) with probabilities of failure $q_0, q_1, q_2, q_1, q_2, \ldots$ $q_1, q_2, q_1, q_2, \ldots$ and $q_2, q_1, q_2, q_1, \ldots$ correspodingly. By analogy in second and third cases the sequence $q_0, q_1, q_2, q_1, q_2, \ldots$ is reconstructed after the success.

Thus at $\beta_-^B(\cdot|p, \alpha = \frac{2}{2r+1})$ we have $q_1 < q_2$ but at $\beta_-^A(\cdot|p, \alpha = \frac{2}{2r+1})$ conversely $q_1 > q_2$, at that $q \geq q_1$ and $q > q_2$. These inequalities are significant for interpretations.

3. The Galton-Watson model and the generalization geometric distribution

Denote by T the set of k-dimensional vectors with nonnegative integer components and by $\mathbf{e}_i$ $(1 \leq i \leq k)$ the k-dimensional vector which i-component is 1 and other components are 0. The Galton-Watson branching process is a Markov process $\mathbf{Z_0}, \mathbf{Z_1}, \mathbf{Z_2}, \ldots$, which states are vectors from T. Let $\mathbf{Z_0}$ is nonrandom. i-component Z_n^i of vector $\mathbf{Z_n}$ is equal a number of particles of i-type in n-generation.

The transition law is given in the following way. If $\mathbf{Z_0} = \mathbf{e_i}$ then $\mathbf{Z_1}$ has the generating function

$$f^i(\nu_1, \ldots, \nu_k) = \sum_{r_1, \ldots, r_k = 0}^{\infty} p^i(r_1, \ldots, r_k) \nu_1^{r_1} \ldots \nu_k^{r_k},$$

where i-particle generates r_1 new particles of type 1, ..., r_k new particles of type k with probability $p^i(r_1, \ldots, r_k)$. If $\mathbf{Z}_n = (r_1, \ldots, r_k) \in T$, then $\mathbf{Z}_{n+1}$ is the sum of N independent vectors among which r_i vectors have the generating function f^i, $i = 1, \ldots, k$, $N = r_1 + \ldots + r_k$.

In [2] the generalized binomial distribution $\beta_+(\cdot|n, p, \alpha = \frac{s}{m})$ was described by means of the associated multi-type Galton-Watson model with particles of $s + 1$ types. At $\alpha = \frac{2}{2r+1}$ the transition law was given in the following way:

$$\begin{cases} f^1(\nu_1, \nu_2, \nu_3) = \nu_1 \nu_3 (1 - q^{r+1}) + \nu_2 q^{r+1}, \\ f^2(\nu_1, \nu_2, \nu_3) = \nu_1 \nu_3 (1 - q^r) + \nu_1 q^r, \\ f^3(\nu_1, \nu_2, \nu_3) = \nu_3. \end{cases} \tag{3}$$

It was shown that the random number of 3-particles has the generalized binomial distribution $\beta_+(\cdot|n, p, \alpha = \frac{2}{2r+1})$. Hence the random number of 1- and 2- particles

before the first appearance of 3-particle has the generalized geometric distribution $\beta_-^B(\cdot|p, \alpha = \frac{2}{2r+1})$.

Theorem 2. *Consider the multi-type Galton-Watson model with four types of particles and the generating functions:*

$$\begin{cases} f^0(\nu_0,\nu_1,\nu_2,\nu_3) = \nu_0\nu_3(1-q) + \nu_1 q, \\ f^1(\nu_0,\nu_1,\nu_2,\nu_3) = \nu_0\nu_3(1-q^r) + \nu_2 q^r, \\ f^2(\nu_0,\nu_1,\nu_2,\nu_3) = \nu_0\nu_3(1-q^{r+1}) + \nu_1 q^{r+1}, \\ f^3(\nu_0,\nu_1,\nu_2,\nu_3) = \nu_3. \end{cases} \tag{4}$$

Then the random number of 0-, 1- and 2- particles before the first appearance of a 3-particle has the generalized geometric distribution $\beta_-^A(\cdot|p, \alpha = \frac{2}{2r+1})$.

Proof. The recurrent system of generating functions (4) is given by:

$$\begin{cases} f^0_{n+1}(\bar\nu) = f^0_n(\bar\nu) f^3_n(\bar\nu)(1-q) + f^1_n(\bar\nu) q, \\ f^1_{n+1}(\bar\nu) = f^0_n(\bar\nu) f^3_n(\bar\nu)(1-q^r) + f^2_n(\bar\nu) q^r, \\ f^2_{n+1}(\bar\nu) = f^0_n(\bar\nu) f^3_n(\bar\nu)(1-q^{r+1}) + f^1_n(\bar\nu) q^{r+1}, \\ f^3_{n+1}(\bar\nu) = f^3_n(\bar\nu). \end{cases}$$

where $\bar\nu = (\nu_0,\nu_1,\nu_2,\nu_3)$. At $\nu_0 = \nu_1 = \nu_2 = 1, \nu_3 = \nu$ and $\mathbf{Z_0} = (1,0,0,0)$ we have [4] the system of the generating functions for the random number of 3-particles as (2) where $q_0 = q$, $q_1 = q^r$ and $q_2 = q^{r+1}$, $p_i = 1 - q_i$, $i = 0,1,2$. Hence probabilities of i-particles ($i = 0,1,2$) before the first appearance of 3-particle generate a sequence q, q^r, q^{r+1}, q^r, q^{r+1}, q^r ,... which corresponds to the generalized geometric distribution $\beta_-^A(\cdot|p, \alpha = 2/(2r+1))$ according to corrolary 1. □

4. The medical application of the generalized geometric distribution

For example consider a cold test data before CABG (coronary artery bypass grafting). We have the first group of patients with the RA (radial artery) graft vasospasm during the operation and the second reference group ($i = 1,2$). Let n^i_j denote number of patients from i-th group with a RA postcoldtest recovery in j minutes.

Table 1: The postcoldtest recovery time data

	Time (minutes) j	0	1	2	3	4	≥ 5
Number of patients	n^1_j	7	3	10	6	2	12
	n^2_j	34	9	3	2	2	1

Patients of the first group have a significantly higher maximally registered level of systolic blood pressure (BP) and a bigger Kernogan index. In other words their

RA is worse functionally and morphologically and the patient's organism is more vulnerable.

In this group the post cold test recovery time positively correlates with an artery intima thickness and negatively with the maximally registered level of systolic BP. The distribution of the recovery time fits the generalized geometric distribution $\beta_{-}^{A}(\cdot|p=0.16,\alpha=2/3)$, P-level equals 0.23. Probability of the instantaneous RA postcoldtest recovery is very small and equals $p=0.16$ ($q=0.84$). The parameter $\alpha=2/(2r+1)$ is interpreted as the fast time parameter, in our case $r=1$, $\alpha=2/3$. At $p=0.15$ and $\alpha=\frac{r}{2r+1}=0.35$ by $r=1.16$ distance between distributions is decreasing but over the degree of freedom P-level equals 0.14. Consider last estimations as interpretations for simplicity.

Let the 3-type particle corresponds to a success (the RA diameter is recovered during one minute) then particles of the other type correspond to failures. The 0-type acquaintance particle corresponds to a failure in the initial test with probability $q_0=1-p=0.85$ which is typical for the state of an organ (RA). The 1-type particle is interpreted as a failure which is connected with a local defensive factor and its probability $q_1=q^r=0.83$. Possibly the nearly equal probabilities q_0 and q_1 are explained by worse functional and morphological state of RA. The 2-type particle is interpreted as a failure which is connected with a humoral defensive factor and its probability is $q_2=q^{r+1}=0.7$. This model describes a learnable system with a local and humoral reserve which interchange.

Hence in the discussed model p is responsible for the RA state and α corresponds to the maximally registered level of systolic BP which accelerates a RA postcoldtest recovery. The parameter r can be called as a defensive scale.

The distribution of the recovery time from the second reference group fits the geometric distribution $\beta_{-}^{B}(\cdot|p,\alpha=1)$ with probability $p=0.58$ of the instantaneous recovery after the cold test, P-level equals 0.4. This model describes a trained system because the 0-type acquaintance particle does not exist and the success probability can be presented as $1-q^r$. For example there are possible following combinations: $p=0.19, r=4$; $p=0.25, r=3$; $p=0.35, r=2$ or $p=0.58, r=1$. According to distribution it is impossible to conclude if p small and r is high or p is high. In a trained system the defensive scale r is presented a priori and the humoral and local defensive factors are indiscernible. Patients of this group are more physically active and their morphological RA changes are less.

References

[1] Bart A., Klochkova (Alexeyeva) N., Kozhanov V. (1993) The Universal Scheme of Regulations in Biosystems for the Analysis of Neuron Junctions as an example. *Model-Oriented Data Analysis, W.G.Muller, H.P.Wynn, A.A.Zhigljavsky (Eds.) Physica-Verlag, Heidelberg.* 167-177.

[2] Bart A., Alexeyeva N., Bochkina N. (2000) Partially Inversion of Functions for Statistical Modeling of Regulatory Systems. *Advances in Stochastic Simulation Methods (Series Statistics for Industry and Technology) eds.*

N.Balakrishan, V.B., S.M.Ermakov, Burkhaser. Boston-Basel-Berlin. 355-371.

[3] Bart A. G. (2003) *The analysis of medical and biological systems. The method of partially inverse functions.* Sankt-Petersburg. 280p. (in Russian).

[4] Harris E. (1925) *The Theory of branching processes.* Springen-Verlag, Berlin-Goettingen-Heidelberg. 356p.

6th St.Petersburg Workshop on Simulation (2009) 960-964

Choice of the reference set in randomization tests in the presence of missing values - a simulation study[1]

Nicole Heussen[2], Ralf-Dieter Hilgers[3] , Diana Ackermann[4]

Abstract

The connection between the randomization procedure and analysis is often ignored in the statistical analysis of clinical trials although regulatory guidelines explicitly call upon (ICH E8). Randomization tests comply with this requirements, as they use the applied random allocation to determine the distribution of the test statistic. However, the impact of missing values on randomization tests is not studied up to now. We explore the use of a randomization test when analyzing data of randomized clinical trials with missing values. We will consider two scenarios leading to a reduced or non-reduced reference set depending on handling of missing data. A simulation study shows that the number of liberal test decisions are slightly smaller for the reduced reference set.

1. Introduction

Randomization tests become more popular in medical and biological sciences, not only to check for robustness of distributional assumptions, but also as a valid analytical method. This analysis meets the requirements of the ICH guideline [1], which recommends that the statistical analysis of the study takes into account the special characteristics of the randomization procedure.

As the usual parametric model needs the assumption of random sampling and a special distribution, the so-called 'population model' is often not appropriate. Especially in clinical trials a population model is often not appropriate, as non-random sampling occurs, a randomization model is rather suitable in this case. Samples that are drawn at random from a population of patients are uncommon in clinical trials. Patients are rather recruited from nonrandom choices of hospitals and places and narrow inclusion criteria are often specified. In fact a 'randomization model' is needed as described for instance in [2] or [3].

[1]The research was financially supported by the BMBF within the joint research project SKAVOE (Foerderkennzeichen: 03HIPAB4).

[2]RWTH Aachen University, E-mail: nheussen@ukaachen.de

[3]RWTH Aachen University, E-mail: rhilgers@ukaachen.de

[4]Boehringer Ingelheim, E-mail: diana.ackermann@ing.boehringer-ingelheim.com

The randomization model uses randomization tests to investigate differences in treatments. There is no need for random sampling, the only random component is the allocation of subjects to treatments, which seems to be an adequate model for a clinical trial where a randomized allocation is used. Additionally the assumption of distribution is needless, when conducting a randomization test.

In this paper the problem of dealing with missing values, when conducting randomization tests, is investigated. Therefore we consider two different approaches to the set of all permutations of allocation. The two approaches lead to reducing the reference set by deleting the allocated treatment corresponding to the missing observation or keeping the originally reference set by assigning zero to the missing observation. It has to be noted that dealing with missing values in standard statistical software like SAS favors the approach of deleting missing values. Which corresponds to reducing the reference set in the context of exact rank tests.

However in the light of randomization tests Edgington [4] pointed out that both above approaches mentioned may lead to valid results.

The aim of this paper is to explore randomization tests in case of occurrence of missing data using the two reference sets with respect to their analytical properties and conservativeness. After an introduction to randomization rank tests in section 2 we explore two approaches of dealing with missing values in randomization tests in section 3. In section 4 asymptotic results will be discussed. A simulation study is described in section 5 to compare the two approaches with respect to the p-values and test decisions.

2. Randomization rank tests

Let us consider a random allocation rule (RAR) where $N_A = N/2$ patients are randomly allocated to treatment A and $N_B = N/2$ patients to treatment B. The set of all possible random allocation sequences under RAR is called reference set.

The Wilcoxon rank sum statistic $T = \sum_{i=1}^{N} \delta_i i$ will be used to proof the null hypothesis of no treatment effect. Hereby δ_i is an index variable, belonging to the observations from group A ($\delta_i = 1$) or group B ($\delta_i = 0$), in the combined and ordered sample.

For $1 \leq i \leq N$ the expectation of δ_i is $E(\delta_i) = N/2$ and the variance of δ_i is $Var(\delta_i) = 1/4$. Furthermore it holds that $Cov(\delta_i, \delta_j) = -1/(4(N-1)), 1 \leq i, j \leq N, i \neq j$.

Moreover

$$E(T) = (N(N+1))/4 \tag{1}$$

$$Var(T) = (N^2(N+1))/48. \tag{2}$$

In the case of ties and the usage of midranks the expectation (1) persists, but the variance gets smaller. Although in practical situations tied observations will occur, we will restrict our consideration to the case of not ties.

A recursive formula for the exact distribution of T is given in [5]. Algorithms to compute the exact distribution can be found in various articles (see [6], [7], [8] or [9]).

The exact distribution of the test statistic of this randomization test results by computing the Wilcoxon rank sum statistic T for each allocation sequence of the reference set. The p-value is given as proportion of allocation sequences in the reference set that leads to values of the test statistic greater than or equal to the observed value of T based on the applied allocation sequence.

Note that the number of allocation schemes under random allocation rule equals $\binom{N}{N/2}$.

3. Handling /Dealing with missing values in randomization tests

Up to now there exists no recommendation for the choice of the reference set, when dealing with missing values in randomization tests. Obviously there are two options. Firstly it is possible to keep the original reference set according to the randomization procedure (CRS) or secondly to reduce the reference set according to the number and location of the missing value(s) (RRS). Edgington pointed out that both approaches are valid, but a investigation of the differences is not given.[4]. Thus the question which of the two approaches lead to a more conservative or liberal test decisions remains unanswered. Table illustrates the different reference sets for $N = 6$. While the complete reference set consists by $6!/3!^2 = 20$ allocation sequences, the reduced set is only composed by $5!/3!2! = 10$ sequences. The elements of the reduced reference set depend on the location of the missing value (group A or B).

4. Asymptotic investigation

Our considerations starts with deriving of the expectation and variance of the test statistic T for the two approaches CRS and RRS, if $k_A \leq k < N/2$ where k_A is the number of missing values in group A and k is the total number of missing values.

In the case of RRS, the allocated treatment corresponding to the missing value is skipped. Then the reference set consists of all permutations of the remaining reduced allocation sequence.

Denote by T_{RRS,k,k_A} the corresponding test statistic, then the expectation and variance of T_{RRS,k,k_A} are given by

$$E(T_{RRS,k,k_A}) = \frac{(N/2 - k_A)(N - k + 1)}{2} \tag{3}$$

$$Var(T_{RRS,k,k_A}) = \frac{(N - 2k_A)(N - 2(k - k_A))(N - k + 1)}{48} \, . \tag{4}$$

On the other hand, denote by $T_{CRS,k}$ the test statistic while keeping the reference set and omitting the missing ranks. Let l_j be the localization of the jth missing value within the rank vector $(R_1, ..., R_N)'$, $j = 1, ..., k$ and $l_j = N + 1$ for $j > k$ and $k_0 = 0$.

Table 1: Different reference sets according to the two approaches

complete reference set (CRS)	**reduced reference set (RRS)** missing value in A	**reduced reference set (RRS)** missing value in B
AAABBB	AABBB	BBAAA
AABABB	ABABB	BABAA
AABBAB	ABBAB	BAABA
AABBBA	ABBBA	BAAAB
ABAABB	BABAB	ABABA
ABABBA	BAABB	ABBAA
ABABAB	BABBA	ABAAB
ABBAAB	BBABA	AABAB
ABBABA	BBAAB	AABBA
ABBBAA	BBBAA	AAABB
BAAABB		
BAABAB		
BAABBA		
BABABA		
BABAAB		
BABBAA		
BBAAAB		
BBAABA		
BBABAA		
BBBAAA		

According to the consideration in section 2 the Wilcoxon rank sum statistic for k missing values is given by

$$\begin{aligned} T_{CRS,k} &= \sum_{i=1}^{l_1-1} \delta_i i + \sum_{i=l_1+1}^{l_2-1} \delta_i (i-1) + \sum_{i=l_2+1}^{l_3-1} \delta_i (i-2) + \ldots + \sum_{i=l_k+1}^{N} \delta_i (i-k) \\ &= \sum_{j=0}^{k} \sum_{i=l_j+1}^{l_{j+1}-1} \delta_i (i-j), \quad l_0 = 0; l_k + 1 = N + 1 \, . \end{aligned} \tag{5}$$

Hence the expectation of the rank sum statistic with k missing values is

$$E(T_{CRS,k}) = \sum_{j=0}^{k} \sum_{i=l_j+1}^{l_{j+1}-1} E(\delta_i)(i-j) = \frac{1}{2} \sum_{i=1}^{N-k} i = \tfrac{1}{4}(N-k)(N-k+1) \tag{6}$$

.

The variance of $T_{CRS,k}$ can be computed in an analogously using the properties of δ_i. After straight forward computation it results

$$Var(T_{CRS,k}) = \frac{(N-k)(N-k+1)}{48(N-1)}\,[N(N-1)+k(2N-3k+3)] \qquad (7)$$

Comparison of the two approaches

A straightforward computation shows that $Var(T_{RRS,k,k_A}) \leq Var(T_{CRS,k})$ if $k_A \leq k \leq N/2$.

One reason for the smaller variance of T in the case of the RRS approach is the minor set of values of the test statistic due to the additional information that k_A gives.

To study the asymptotic performance of the Wilcoxon rank sum test it might be acceptable to consider the z-score $z = (T_0 - E(T))/(\sqrt{Var(T)})$ or $z = (T_0 - E(T))^2/(Var(T))$.

If $k_A = k/2$, meaning half of the missing values occur in the A-group, it can be shown that $E(T_{RRS,k,k_A}) = E(T_{CRS,k})$ and thus $z_{RRS} \leq z_{CRS}$.

Thus the reduced set is anticonservative in the sense that given T_0 the p-value will be smaller. The same trend can be shown if $k_A = k = 1, 2$ and $N > 10$.

There is no unique trend for the relation between z_{RRS} and z_{CRS} for other constellation of k_A, k and N. To give some illustrations we refer to our simulation study in the next section.

5. A simulation study

We compare the p-values of the Wilcoxon rang sum test resulting from the two different kinds of reference set in a numerical simulation study (SAS System V 9.1., Windows XP).

The data of a subset (34 observations) of a randomized controlled clinical trial [10] is used to investigate the impact of missing values on the test decision of randomisation tests. In the SPR Study two ophthalmic surgical techniques are compared.

Here $N_A = N_B = 17$. To get the simulated distribution, one million allocation sequences are randomly generated. Uniformly distributed random numbers are generated using a prime modulus multiplicative generator (cf [11]). Then the value of the test statistic is calculated, respectively. This is done for the number of missing values $k = 0$ to $k = 10$ and all possible values of k_A, respectively. The simulated arithmetic means and standard deviations were compared to their theoretical equivalents and only small differences were found (data not shown).

Examples of the resulting simulated distributions of the test statistic are shown in figure 1, for $k = 0$ and $k = 4$. The differences between the two approaches CRS and RRS are apparent. If the RRS approach is used, the distribution depends on the location of missing values and therefore on k_A, the number of missing values in group A . In figure 1 it is also seen that the variance of T is smaller, when the reference set is reduced.

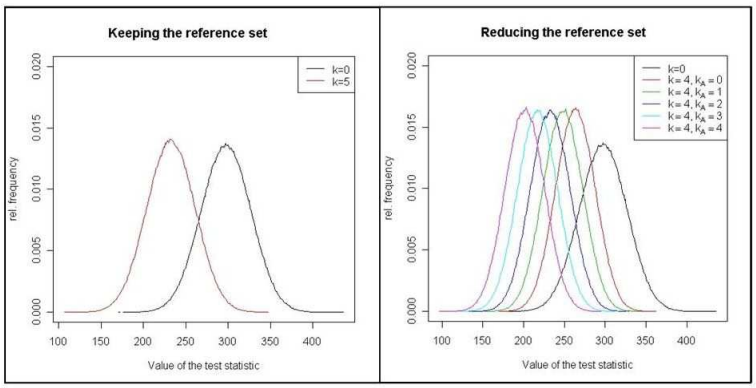


Figure 1: Simulated distributions of the two approaches

The p-value under the two approaches CRS and RRS is computed and compared to the p-value of the case without missing values. The p-value can be equal (rounded to three decimal places), smaller (liberal test decision) or greater (conservative test decision).For this comparison there are additional 10000 allocation sequences generated.

Up to four missing values ($k = 4$) this is done for all possible missing values. To save computing time, from $k = 5$ only a random sample of size 10000 is chosen. This is marked with * in the following table.

The resulting proportions in per cent for both approaches CRS and RRS are shown in table . When using the reduced reference set (the RRS approach) the proportion of liberal test decisions is always smaller than when using the complete reference set, except if $k = 1$. However, the differences are rather slight. Common for both approaches is that the proportion of equal p-values decreases when the number of missing values increases.

It is also seen for $k = 4$ that the proportions are very similar, if all possible missing values are used and if a sample of size 10000 is drawn. That reflects the quality of the simulation study.

Table provides means, standard deviation, maxima and minima of the absolute difference between the p-value of the nonmissing case and with missing observations. Thus negative (positive) values of the difference represent that the nonmissing p-values are smaller (greater) than the p-values with missing observations. It is seen that the difference of p-values is in average negative for both approaches RRS and CRS, which means that the nonmissing p-values are in average smaller than the p-values with missing values (conservative test decision) . The maximum difference in p-values is higher when using the complete reference set (CRS).

From tabel it is seen, that both approaches provide conservative test decisions in average and the difference is increasing with increasing number of missing values.

Table 2: Results of the simulation study (number of test decisions in per cent)

	reduced ref. set (RRS)			**complete ref. set (CRS)**		
k	equal	conservative	liberal	equal	conservative	liberal
		test decision			test decision	
1	1.16	49.96	48.88	3.54	49.22	47.24
2	0.80	50.12	49.08	0.29	50.03	49.68
3	0.77	50.10	49.12	0.46	50.02	49.52
4	0.73	50.17	49.10	0.33	50.02	49.65
4	0.73*	50.17*	49.10*	0.33*	50.02*	49.65*
5	0.67*	50.34*	48.98*	0.39*	50.28*	49.34*
6	0.94*	50.29*	48.77*	0.32*	50.20*	49.48*
7	0.58*	50.48*	48.94*	0.30*	50.20*	49.50*
8	0.57*	50.65*	48.79*	0.27*	50.35*	49.38*
9	0.48*	50.79*	48.72*	0.17*	50.24*	49.59*
10	0.43*	50.83*	48.73*	0.19*	50.39*	49.42*

Table 3: Results of the simulation study (differences of p-values)

	reduced ref. set (RRS)				**complete ref. set (CRS)**			
k	**MEAN**	**STD**	**MIN**	**MAX**	**MEAN**	**STD**	**MIN**	**MAX**
1	-0.00046	0.05	-0.12	0.12	-0.00051	0.10	-0.23	0.23
2	-0.00084	0.08	-0.24	0.24	-0.00056	0.14	-0.43	0.43
3	-0.00087	0.09	-0.36	0.36	-0.00061	0.17	-0.61	0.60
4	-0.00193	0.11	-0.48	0.48	-0.00262	0.19	-0.74	0.74
5	-0.00281	0.12	-0.56	0.59	-0.00465	0.21	-0.82	0.82
6	-0.00323	0.13	-0.63	0.63	-0.00418	0.22	-0.89	0.89
7	-0.00387	0.14	-0.69	0.70	-0.00454	0.24	-0.92	0.93
8	-0.00461	0.16	-0.75	0.72	-0.00574	0.25	-0.95	0.96
9	-0.00520	0.17	-0.81	0.78	-0.00508	0.26	-0.96	0.97
10	-0.00557	0.17	-0.84	0.81	-0.00642	0.26	-0.99	0.98

6. Discussion

Applying a randomization test is a valid analytical method and consistent with the ICH guideline [1] for statistical analyses. Edginton describes a variety of applications of randomization tests including ANOVA, trend tests and correlation tests [9]. Randomization test are not only used in clinical trials, but can also be applied to biological or environmental data [12].

Our simulation study indicates that the choice of the reference set affects the test decision of a randomization test, when dealing with missing values. The reduced reference set shows less liberal test decisions; the probability distribution is in the special case of using the random allocation rule implemented via the

exact Wilcoxon rank sum test. Hence, the use of the reduced reference set is suggested, as it meets the criteria of conservativeness, that are specified in the CPMP-guideline 'Points to consider on missing values' [13]. We provided formulas for the expectation and variances of the test statistic for both approaches and have seen that both approaches have the same expectation when the missing values are 'balanced' across the groups. In this case the variance of the test statistic is smaller when using the RRS approach than when using the CRS approach.

There are some further investigations needed due to the following topics: The impact of ties in the above considered model and the long run effect on the differences of p-values and the test decisions. One possible assumption is that the proportion of liberal and conservative test decisions converge to 0.5 with increasing number of observations and decreasing number of missing values.

Also the use of other allocation methods as for instance permuted block randomization and their impact on the test decision of randomization tests needs further investigation.

References

[1] ICH, E8 (1997) *General Considerations for Clinical Trials*, http://www.emea.europa.eu/pdfs/human/ich/029195en.pdf.

[2] Rosenberger, W.F. and Lachin, J.M.(2002).*Randomization in Clinical Trials. Theory and Praxis.* Wiley, New York.

[3] Lachin, J.M. (1988). Statistical Properties of Randomization in Clinical Trials. *Controlled Clinical Trials* 9:289-311.

[4] Edgington, E.S.(1995). *Randomization tests.* 3rd edition, Marcel Dekker, New York.

[5] Hajek, J. and Sidak, Z. and Sen, P.K. (1999) *Theory of Rank Tests.*, 2nd edition, Academic Press, New York .

[6] Pagano, M. and Tritchler, D. (1983) On Obtaining Permutation Distributions in Polynomial Time , *Journal of the American Statistical Association*, 78: 435-440.

[7] Mehta, C.R. and Patel, N. R. (1983) A network algorithm for Performing Fisher's Exact Test in r $\times$ c Contingency Tables, *Journal of the American Statistical Association*,78: 427-434.

[8] Mehta, C.R. and Patel, N. R. and Wei, L.J.(1988)Constructing exact significance tests with restricted randomization rules , *Biometrika*, 75: 295-302.

[9] Streitberg, B., Rohmel, J., (1986). Exact distributions for permutation and rank tests: An introduction to some recently published algorithms. *Statistical Software Newsletter* 12, 10-18.

[10] Heimann, H.; Hellmich, M.; Bornfeld,N.; Bartz-Schmidt, K.U.; Hilgers, R.D. and Foerster, M.H. (2001) Scleral buckling versus primary vitrectomy in rhematogeneous retinal detachment (SPR Study): Design issues and implications. *Graefe's Arch Clin Exp Ophthalmol*, 239: 567-574

[11] SAS 9.1.3 (2006) *Language Reference: Dictionary, Fifth Edition.* Cary, NC: SAS Institute Inc.

[12] Manly, B.F.J.(2007). *Randomization,Bootstrap and Monte Carlo Methods in Biology.* Chapman and Hall, London.

[13] CPMP (2001) *Points to Consider on Missing Data*, http://www.emea.eu.int/pdfs/human/ewp/177699EN.pdf.

6th St.Petersburg Workshop on Simulation (2009) 969-973

On Discrepancy in Connection with the Residual of the Cubature Formulae[1]

Elfia G. Burnaeva[2]

,

Abstract

Discrepancy of the given point set A in s-dimensional hypercube is well-known equidistribution criteria. Besides discrepancy is quality criteria for numerical integration error (Quasi Monte-Carlo). It is studied in this article some simple generalization of discrepancy connected with introduction of weights for the points of A.

The residuals of the cubature formulae under the integration of the functions of small smoothness are connected closely to the uniform distributions of the integration knots. The measure of the divergence from the uniformity is the so-called discrepancy. The well-known Koksma-Hlavka inequality determines the connection between the norm of the residual functional in the limited variation function (variation in the sense of Hardy-Krause) and discrepancy [1]. If

$$R_N[f] = \int_{\mathrm{K}_d} f(X)dX - \frac{1}{N}\sum_{j=1}^{N} f(X_j),$$

than we have

$$\big|R_N[f]\big| \le \frac{1}{N} V(f) D^*(X_1, \ldots, X_N), \tag{1}$$

Where $V(f)$ is the variation of f, K_d is a unit d-measured hypercube,

$$D^*(X_1, \ldots, X_N) = \sup_{X \in \mathrm{K}_d} \big||\mathcal{A}_N \cap [0, X)| - NL([0, X))\big|, \tag{2}$$

$\mathcal{A}_N = \{X_1, \ldots, X_N\}$, $X = (x_1, \ldots, x_d)$,
$[0, X)$ is a parallelepiped with sides $[0, x_l)$, $l = 1, \ldots, d$.
$|S|$ is a number of points of S, if S is a point set while L is Lebesgue measure. The inequality (1) is the justification of using Quasi Monte Carlo method under solving different application problems. Discrepancy and its generalizations are being widely investigated in the analysis and the number theory [3].

[1]This work was supported by grant of RFBR 08-01-00194 RFBR
[2]Saint Petersburg State University, E-mail: burnaeva@mail.ru

The papers [9] and [4] are devoted to the case of cubature sums of the general form $\sum\limits_{j=1}^{N} A_j f(X_j)$, where A_j are constant coefficients if the formula. The discrepancy analog has also been studied in these papers. Further we will show some new results on the correlation of the independent criterions of a cubature formulas quality for the simple case of $d = 2$. Let us remind that $f(x, y)$ that has corresponding derivatives, can be represented in K_2. In the form

$$f(x,y) = f(1,1) + \int\limits_1^x f_x^{'}(u,1)\,du + \int\limits_1^y f_y^{'}(1,v)\,dv + \int\limits_1^x \int\limits_1^y f_{xy}^{''}(u,v)\,dudv. \qquad (3)$$

Out of this presentation we get (see for example [5])

$$R_N[f] = \int\limits_0^1 f_x^{'}(u,1)\big[u - \sum_{j=1}^{N} A_j\Theta(x_j - u)\big]\,du +$$
$$\int\limits_0^1 f_y^{'}(1,v)\big[v - \sum_{j=1}^{N} A_j\Theta(y_j - v)\big]\,dv +$$
$$\int\limits_0^1\int\limits_0^1 f_{xy}^{''}(u,v)\big[uv - \sum_{j=1}^{N} A_j\Theta(x_j - u)\Theta(y_j - v)\big]\,dudv,$$

$$\Theta(t) = \begin{cases} 0, & t > 0 \\ 1, & t < 0 \\ \frac{1}{2}, & t = 0. \end{cases}$$

That is

$$R_N[f] = \int\limits_0^1 f_x^{'}(u,1)\mathcal{K}_1(u)\,du + \int\limits_0^1 f_y^{'}(1,v)\mathcal{K}_2(v)\,dv + \int\limits_0^1\int\limits_0^1 f_{xy}^{''}(u,v)\mathcal{K}_3(u,v)\;dudv, \qquad (4)$$

where

$$\mathcal{K}_1(u) = u - \sum_{j=1}^{N} A_j\Theta(x_j - u), \mathcal{K}_2(u) = v - \sum_{j=1}^{N} A_j\Theta(y_j - v),$$

$$\mathcal{K}_3(u) = uv - \sum_{j=1}^{N} A_j\Theta(x_j - u)\Theta(y_j - v).$$

As $f_x^{'}(x,1)$, $f_y^{'}(1,y)$ and $f_{xy}^{''}(u,v)$ are independent, each of these functions can have any value from the linear normalized spaces F_1, F_2 and F_3 correspondently. For $F_1 = F_2 = L_1$, $F_3 = L_1$, where L_1 is space of integrated functions. We have

$$\big|R_N[f]\big| \le V(f)\cdot\max\big(\sup_u|\mathcal{K}_1(u)|, \sup_v|\mathcal{K}_2(v)|, \sup_{u,v}|\mathcal{K}_3(u,v)|\big), \qquad (5)$$

$$V(f) = \int\limits_0^1 |f_x^{'}(u,1)|\,du + \int\limits_0^1 |f_y^{'}(1,v)|\,dv + \int\limits_0^1\int\limits_0^1 |f_{xy}^{''}(u,v)|\,dudv$$

Is a variation of the function f. But as it is shown in [1]

$$\begin{array}{l}\sup\limits_{u,v}|\mathcal{K}_3(u,v)| \geq \sup\limits_{u}|\mathcal{K}_1(u)|,\\ \sup\limits_{u,v}|\mathcal{K}_3(u,v)| \geq \sup\limits_{v}|\mathcal{K}_2(v)|\\ \text{and correspondently } \left|R_N[f]\right| \leq V(f)\sup\limits_{u,v}|\mathcal{K}_3(u,v)|.\end{array} \tag{6}$$

It is easy to see that in the case when the coefficients A_j are equal,(6) coincides with Koksma-Hlavka inequality. Thus if F_l, $l = 1, 2, 3$ are spaces of the integrated functions, then the quality of the cubature formula is determined by the only criterion than coincides with star discrepancy for $A_j = 1/N$. We state the following:
1. In the case $F_l = L_2$ the inequalities of $\|\mathcal{K}_3\|_{L_2} \geq \|\mathcal{K}_l\|_{L_2}$ form do not actually exist, that is the quality of the cubature formula is not defined by an only criterion. It is of interest to find out at least for the presentation case (4), if the other spaces exist except the spaces of the integrated functions and the functions of the limited variation for which the quality of the cubature formula is not defined by one criterion.
2. The criterion $\sup\limits_{u,v}|\mathcal{K}_3(u,v)|$ defines also the quality of a formula that is precise for the polynomial of 1 degree in the case when the second derivatives exist. The proof of this statement can be received from the representation of the residual B in the form

$$R_N[f] = \int\limits_0^1 f''_{x^2}(u,1)\widetilde{\mathcal{K}}_1(u)\,du + \int\limits_0^1 f''_{y^2}(1,v)\widetilde{\mathcal{K}}_2(v)\,dv + \int\limits_0^1\int\limits_0^1 f''_{xy}(u,v)\mathcal{K}_3(u,v)\,dudv,$$

where

$$\widetilde{\mathcal{K}}_1(u) = \left|\frac{u^2}{2} - \sum a_i(x_i - u)\Theta(x_i - u)\right|,\ \widetilde{\mathcal{K}}_2(v) = \left|\frac{v^2}{2} - \sum a_i(y_i - v)\Theta(y_i - v)\right|.$$

Further we should verify the inequalities

$$\sup_{u,v}|\mathcal{K}_3(u,v)| \geq \sup_{u}|\mathcal{K}_1(u)| \qquad \sup_{u,v}|\mathcal{K}_3(u,v)| \geq \sup_{v}|\mathcal{K}_2(v)|$$

with all A_j, x_j, $j = 1, 2, \ldots, N$.

Thus the discrepancy analog (with unequal A_j) serves as a criterion of quality also for the formulae that are precise for the polynomials of the 1-summarized degree.
3. The results shown above (including Koksma-Hlavka inequality) are actual also for the wide class of the integration convex domains.

References

[1] Niederreiter H. (1992) Random Number Generation and Quasi-Monte Carlo Methods. SIAM, Philadelphia

[2] Sobol I.M. (1969) Multidimensional Quadrature Formulae and Haar Functions. (in Russian), Nauka ed., Moskow.

[3] Ermakov S.M, Burnaeva E.G. (2005) *On multicriterion problems in theory of cubature functions* Vestnik of S-Petersburg University, ser.1, pub.1.

[4] Ermakov S.M, Burnaeva E.G., Sidorovskaya M.V.(2006) *Multicriterial problems of error analisis for cubatire formalas* Mathematical models. Theory and Applications, pub.7, NIIH ed., S-Petersburg

[5] Ermakov S.M (2009) *Monte-Carlo method in calculus mathematics.* Piter ed., S-Petersburg (in russian)

Section
Probabilistic models

6th St.Petersburg Workshop on Simulation (2009) 975-979

A remark to large deviation probabilities for sums of i.i.d. random variables in the domain of attraction of a stable law[1]

Leonid Rozovsky[2]

Abstract

We study a logarithmic asymptotics of large deviation probabilities for sums of i.i.d. random variables in the domain of attraction of a stable law.

1. Introduction and Results

Let $X_1, X_2, \dots$ be a sequence of independent random variables with a common d.f. $V(x)$ and let $S_n = X_1 + \cdots + X_n$.

By [1, Theorem 2.1 and Proposition 2.1] the following result holds:

Theorem 1. *Assume that a nonnegative random variable X_1 is in the domain of attraction of a stable law of index α, $0 < \alpha < 1$. Let $x_n \to \infty$ and $\lambda_n \to 0$ as $n \to \infty$ so that $\alpha\Gamma(1-\alpha)(1 - V(1/\lambda_n)) \sim \lambda_n x_n$ and $n\lambda_n x_n \to 0$. Then*

$$-\log \mathbf{P}(S_n \le n x_n) \sim \tfrac{1-\alpha}{\alpha}\, n\lambda_n x_n, \quad n \to \infty.$$

Much later the same result was obtained in [2] again by applying the different method. Below we present more general version of (1) and a similar one for the case $1 < \alpha < 2$.

In what follows we assume that a positive differentiable function $l(x)$ is slowly varying near infinity. For $\alpha \in (0, 2)$ put

$$k_\alpha = (\Gamma(2-\alpha)(|\alpha - 1|/\alpha)^\alpha)^{1/(\alpha-1)}, \quad v_\alpha(x) = x^{-\alpha} l(x), \quad \tau_\alpha(x) = x v_\alpha(x)$$

and let $v_\alpha^{-1}(x)$, $\tau_\alpha^{-1}(x)$ denote the inverse functions of $v_\alpha(x), \tau_\alpha(x)$, respectively, which exist by the properties of $l(x)$.

Theorem 2. *Let $1 - V(x) \sim v_\alpha(x)$ as $x \to \infty$, where a constant $\alpha \in (0, 1)$.*

If $\mathbf{E}e^{-\lambda X_1} < \infty$ for some $\lambda > 0$ then

$$-\log \mathbf{P}(S_n \le nx) \sim k_\alpha\, nx/\tau_\alpha^{-1}(x), \quad n \to \infty,$$

uniformly with respect to x, $1/\epsilon_n < x < \epsilon_n v_\alpha^{-1}(1/n)/n$, where a positive $\epsilon_n \to 0$ arbitrarily slowly.

[1]This work was supported by grant 638.2008.1 of Leading Scientific Schools.

[2]St. Petersburg Chemical Pharmaceutical Academy, E-mail: L-Rozovsky@mail.ru

Theorem 1 follows directly from Theorem 2 with $x = x_n$.

Now consider the case $\alpha \in (1, 2)$.

Theorem 3. *Let $V(-x) \sim v_\alpha(x)$ as $x \to \infty$, where a constant $\alpha \in (1, 2)$. Assume that $\mathbf{E}e^{\lambda X_1} < \infty$ for some $\lambda > 0$ and $\mathbf{E}X_1 = 0$. Then*

$$-\log \mathbf{P}(S_n \geq nx) \sim k_\alpha \, nx/\tau_\alpha^{-1}(x), \quad n \to \infty,$$

uniformly with respect to x, $v_\alpha^{-1}(1/n)/(\epsilon_n n) < x < \epsilon_n$, where a positive $\epsilon_n \to 0$ arbitrarily slowly.

2. Proof

Let $L(\lambda) = \mathbf{E}X_1^2 e^{\lambda X_1} < \infty$ for some $\lambda > 0$. For $0 < u \leq \lambda$ define

$$L(u) = \mathbf{E}e^{uX_1}, \quad m(u) = \big(\log L(u)\big)', \quad \sigma^2(u) = m'(u), \quad Q(u) = um(u) - \log L(u).$$

Lemma([3, Corollary 2]). *Assume that a positive sequence $h_n \to 0$ so that $n\mathbf{E}(1 \wedge h_n^2 X^2) \to \infty$. Then*

$$-\log \mathbf{P}(S_n \geq nx) \sim nQ(h), \quad n \to \infty,$$

uniformly in x, $m(h_n) \leq x \leq m(\lambda)$, where $h = h(x)$ is the unique solution of the equation $m(h) = x$.

To prove Theorems 2 and 3 we apply Lemma (see also [3, Remark 3 and Corollaries 3 and 4)]. Besides that we estimate the asymptotic behavior of the functions $Q(h)$ and $m(h)$ as $h \to 0$ by [1, Proposition 2.1](the case $\alpha \in (0, 1)$) and by [4, (41)-(43)] for the case $\alpha \in (1, 2)$.

References

[1] Jain N.C., Pruitt W.E. (1987) *Lower tail probability estimates for subordinators and nondecreasing random walks.* Ann. Probab., **15**, No. 1, 76–101.

[2] Kasahara Y., Kosugi N. (2000) *Large deviation around the origin for sums of nonnegative i.i.d. random variables.* Natural Science Report, Ochanomizu University, **51**, No. 1, 27–31.

[3] Rozovsky L.V. (2001) *On a lower bound of large-deviation probabilities for a sample mean under the Cramer condition.* Zapiski nauch.semin.POMI(in Russian, translated in Journal of Math. Sci.), **278**, 208–224.

[4] Rozovsky L.V. (1996) *Large deviations of sums of independent random variables from the domain of attraction of a stable law.* Zapiski nauch.semin.POMI(in Russian, translated in Journal of Math. Sci.), **228**, 262–283.

6th St.Petersburg Workshop on Simulation (2009) 977-981

On the strong law of large numbers for sequences of dependent random variables

V.V. Petrov[1], V.M. Korchevsky[2]

Abstract

We present some sufficient condition for the applicability of the strong law of large numbers to sequences of random variables without the independence condition.

1.

Following [3], we denote by Ψ_c (or, respectively, Ψ_d) the set of functions $\psi(x)$ such that $\psi(x)$ is positive and non-decreasing in the interval $x > x_0$ for some x_0 and the series $\sum \frac{1}{n\psi(n)}$ converges (respectively, diverges). The value x_0 need not be the same for different functions ψ.

We consider a sequence of non-negative random variables $\{X_n\}$ with finite absolute moments of some order $p \geqslant 1$ and put $S_n = \sum_{k=1}^n X_k$.

Theorem 1. *Let $\{w_n\}$ be a sequence of positive numbers,*

$$W_n = \sum_{k=1}^n w_k, \quad T_n = \sum_{k=1}^n w_k X_k. \tag{1}$$

Suppose that $W_n \to \infty$ $(n \to \infty)$,

$$\sum_{k=m}^n w_k EX_k \leqslant C \sum_{k=m}^n w_k \tag{2}$$

for all sufficiently large $n - m$ where C is a constant,

$$E\,|T_n - ET_n|^p = O\left(\frac{W_n^p}{\psi(W_n)}\right) \text{ for some function } \psi \in \Psi_c. \tag{3}$$

Then

$$\frac{T_n - ET_n}{W_n} \to 0 \ \ a.s. \tag{4}$$

[1]St.Petersburg State University, E-mail: petrov2v@mail.ru
[2]St.Petersburg State University, E-mail: valery_ko@list.ru

Theorem 1 is a generalization of some results in [4] and [5] corresponding to the case $p = 2$. We remove the additional condition $\frac{w_n}{W_n} \to 0$ $(n \to \infty)$ of Theorem 2 in [5]. Let us indicate two more consequences of Theorem 1.

Theorem 2. *Suppose that*

$$E(S_n - S_m) \leqslant C(n - m) \tag{5}$$

for all sufficiently large $n - m$,

$$E\,|S_n - ES_n|^p = O\left(\frac{n^p}{\psi(n)}\right) \text{ for some function } \psi \in \Psi_c. \tag{6}$$

Then

$$\frac{S_n - ES_n}{n} \to 0 \ \ a.s. \tag{7}$$

We get this result using Theorem 1 for $w_n = 1$ $(n = 1, 2, \ldots)$.

Theorem 3. *If* $ES_n \to \infty$ $(n \to \infty)$,

$$E\,|S_n - ES_n|^p = O\left(\frac{(ES_n)^p}{\psi(ES_n)}\right) \text{ for some function } \psi \in \Psi_c, \tag{8}$$

then

$$\frac{S_n}{ES_n} \to 1 \ \ a.s. \tag{9}$$

We arrive at this proposition applying Theorem 1 to the sequence of random variables $\{Y_n\}$ where $Y_n = \frac{X_n}{EX_n}$ (assuming without loss of generality that $EX_n > 0$ for all n) and putting $w_n = EX_n$. Then we get $T_n = S_n$, $W_n = ES_n = ET_n$, and (4) reduces to (9).

Theorems 2 and 3 are generalizations of some results from [4] and [5] corresponding to the case $p = 2$. Other sets of conditions sufficient for (4), (7) or (9), were given by Etemadi [1], [2].

It was shown in [4] and [5] that in theorems of these papers it is impossible to replace conditions (6) or (8) for $p = 2$ by the weaker conditions that correspond to the replacement of $\psi \in \Psi_c$ by some function $\psi \in \Psi_d$. It follows, for example, that the condition $VarS_n = O((ES_n)^2)$ or even $VarS_n = O(\frac{(ES_n)^2}{\log ES_n})$ together with $ES_n \to \infty$ $(n \to \infty)$ does not guarantee the relation (9). According to Theorem 3, the conditions $ES_n \to \infty$ and $VarS_n = O(\frac{(ES_n)^2}{(\log ES_n)^{1+\delta}})$ for some $\delta > 0$ are sufficient for the relation (9).

2.

Now we consider a sequence of non-negative random variables $\{X_n\}$ with finite second moments.

Theorem 4. *Suppose that*

$$VarS_n \leqslant C\sum_{k=1}^{n} VarX_k \tag{10}$$

for all n where C is a constant,

$$\sum_{n=1}^{\infty} \frac{VarX_n}{n^2} < \infty \tag{11}$$

and the condition (5) *is satisfied. Then the relation* (7) *holds.*

The following result is a consequence of Theorem 4.

Theorem 5. *Suppose that $X_1, X_2, \ldots$ are pairwise independent random variables (not necessarily non-negative). If*

$$\sum_{k=m+1}^{n} E\,|X_k - EX_k| \leqslant C(n-m)$$

for all sufficiently large $n-m$ where C is a constant and the condition (11) *is satisfied, then the relation* (7) *holds.*

Theorems 4 and 5 are generalizations of some results of [1]. Conditions of [1] include the uniform boundedness of EX_n.

Theorem 6. *Let $\{w_n\}$ be a sequence of positive numbers. Using notation* (1) *suppose that*

$$W_n \to \infty, \quad \frac{w_n}{W_n} \to 0 \quad (n \to \infty),$$

$$\sum_{n=1}^{\infty} \frac{w_n^2 VarX_n}{W_n^2} < \infty,$$

$$VarT_n \leqslant C\sum_{k=1}^{n} w_k^2 VarX_k$$

for all n. Let the condition (2) *be satisfied. Then the relation* (4) *holds.*

Theorem 7. *If*

$$ES_n \to \infty, \quad \frac{EX_n}{ES_n} \to 0 \quad (n \to \infty),$$

$$\sum_{n=1}^{\infty} \frac{VarX_n}{(ES_n)^2} < \infty$$

and the condition (10) *is satisfied, then the relation* (9) *holds.*

This result is a consequence of Theorem 6. Theorem 7 implies that the relation (9) holds for a sequence of identically distributed random variables $X_1, X_2, \ldots$ satisfying the conditions $EX_1 > 0$ and $VarS_n \leqslant Cn$ for all n.

References

[1] Etemadi N. (1983) On the law of large numbers for non-negative random variables. *J. Multivariate Analysis*, **13**, 187–193.

[2] Etemadi N. (1983) Stability of sums of weighted non-negative random variables. *J. Multivariate Analysis*, **13**, 361–365.

[3] Petrov V.V. (1975) *Sums of independent random variables*. Springer, New York.

[4] Petrov V.V. (2008) On the strong law of large numbers for a sequence of non-negative random variables. *Theory Probability Appl.*, **53**, N 2, 379–382.

[5] Petrov V.V. (2008) On stability of sums of non-negative random variables. *Notes of Scient. Seminars of St.Petersburg Dept. of Steklov Math. Institute*, **361**, 78–82.

6th St.Petersburg Workshop on Simulation (2009) 981-985

On one new model of records[1]

Valery Nevzorov[2], Vahagn Saghatelyan[3]

Abstract

[1] A new model of records (the so called confirmed records) is discussed. Some definitions and properties of the corresponding record times and values are given.

1. Introduction

Let $X_1, X_2, \ldots$ be a sequence of independent random variables with a common distribution function (d.f.) F. The classical definition of record values is the following. We say that X_j is an upper record value (simply speaking , X_j is a record value), if $X_j > \max\{X_1, \ldots, X_{j-1}\}, \quad j = 2, 3, \ldots$. Note that X_1 is considered as the first record value. We also say that $1 = L(1) < L(2) < \ldots$ are record times if

$$X(j) = X_{L(j)}, \ j = 1, 2, \ldots,$$

are the corresponding record values.

The theory of the classical records is developed well (see, for example, [1] and [2]). There are some record schemes also for sequences of nonidentically distributed X's. We suggest one more record model. The matter is that the classical record scheme is rather sensitive to a presence of outliers. The only outlier amongst X's can change a record sequence in a great rate. Hence we suppose one new definition of records, which can be characterized by lack of sensitivity to outliers.

2. Definitions of confirmed records

Let $X_1, X_2, \ldots$ be a sequence of independent identically distributed (i.i.d.) random variables (r.v.) with a common continuous d.f. F. Fix some k , $k = 1, 2, \ldots$. The first record vector with the k-fold confirmation $\mathbf{X}(1, k) = (X_1(1), X_2(1), \ldots, X_k(1))$ coincides with $(X_{1,k}, X_{2,k}, \ldots, X_{k,k})$, where $X_{1,k}, X_{2,k}, \ldots, X_{k,k}$ are the order statistics based on the observations

$$X_1, X_2, ..., X_k,$$

[1]This work was supported by grants RFBR 07-01-00688 and #09-01-00808

[2]St. Petersburg state university, E-mail: valnev@mail.ru

[3]St. Petersburg state university, E-mail: vahagn_s@yahoo.com

and the first record time $L(1,k)$ coincides with k. Then we are waiting for an appearance of k first values

$$X_{\alpha(1)}, X_{\alpha(2)}, \ldots, X_{\alpha(k)},\ k < \alpha(1) < \alpha(2) < \ldots < \alpha(k),$$

which are greater than $X_{k,k}$, and generate order statistics

$$X_{1,k}(2), X_{2,k}(2), \ldots, X_{k,k}(2),$$

based on the given observations . Thus we define the second record vector

$$\mathbf{X}(2,k) = (X_{1,k}(2), X_{2,k}(2), \ldots, X_{k,k}(2))$$

and the second record time $L(2,k)$, which coincides with $\alpha(k)$, and so on. Hence we get a sequence of the corresponding record times

$$k = L(1,k) < L(2,k) < \ldots < L(n,k) < L(n+1,k) < \ldots$$

and a sequence of record vectors

$$\mathbf{X}(n,k) = (X_{1,k}(n), X_{2,k}(n), \ldots, X_{k,k}(n)),\quad n = 1, 2, \ldots.$$

If $k = 1$ then this record scheme coincides with the classical record model. For the confirmed records (as well as for the classical records) there are two initial distributions (exponential and uniform), for which distributions of elements of $X(n,k)$ have a rather simple form. Indeed, distributions of $L(n,k)$ doesn't depend on the initial continuous d.f. F. We will use

$$\mathbf{Z}(n,k) = (Z_1(n), Z_2(n), \ldots, Z_k(n)) \quad \text{and} \quad \mathbf{U}(n,k) = (U_1(n), U_2(n), \ldots, U_k(n))$$

instead of

$$\mathbf{X}(n,k) = (X_{1,k}(n), X_{2,k}(n), \ldots, X_{k,k}(n))$$

in the situations, when $F(x) = 1 - \exp(-x),\ x > 0$, and $F(x) = x,\ 0 < x < 1$, correspondingly.

3. Distributions of record times

Consider a random sample $X_1, X_2, \ldots, X_n$ of size n from a population with a continuous d.f. F. It is easy to find ([4]) that the distribution of the minimal number $N(n)$ of additional observations, which are needed to order to get the first value exceeding $X_{n,n} = \max\{X_1, \ldots, X_n\}$, is given as follows:

$$P\{N(n) > m\} = \frac{n}{(n+m)}. \tag{1}$$

Let now $N_k(n)$ denote the minimal number of observations appearing until the first value exceeding $X_{n-k+1,n}$ occurs. Then (see, for example, [2])

$$P\{N_k(n) > m\} = \frac{n!(n-k+m)!}{(n-k)!(n+m)!}. \tag{2}$$

In our case let $N(n,k)$ denote the number of observations, which are needed to get exactly k values exceeding $X_{n,n}$. It appears that

$$P\{N(n,k) > m\} = 1 - \frac{m!(n+m-k)!}{(m-k)!(n+m)!}, \quad m \geq k. \tag{3}$$

Note that if $k = 1$, then (3) and (2) coincide with (1). It follows from (3) that

$$P\{L(n+1,k) = j | L(n,k) = i\} = \frac{ki(j-k-1)!(j-i-1)!}{j!(j-i-k)!} \tag{4}$$

for any $n = 1, 2, \ldots$ and $j \geq i + k$.

One gets from (4) that the joint distribution of record times $L(1,k), \ldots, L(n,k)$ is given as follows:

$$\begin{aligned} &P\{L(1,k) = k, L(2,k) = \alpha(2), \ldots, L(n,k) = \alpha(n)\} \\ &= \prod_{r=1}^{n-1} P\{L(r+1,k) = \alpha(r+1) | L(r,k) = \alpha(r)\}, \end{aligned}$$

where $\alpha(1) = k$ and $\alpha(r+1) - \alpha(r) \geq k, \quad r = 1, 2, \ldots, n-1$.

4. Presentations for exponential and uniform confirmed record values

The following relations are valid for the exponential records

$$\begin{aligned} &\mathbf{Z}(n,k) = (Z_1(n), Z_2(n), \ldots, Z_k(n)) : \\ &\mathbf{Z}(1,k) = (Z_{1,k}, Z_{2,k}, \ldots, Z_{k,k}) \stackrel{d}{=} \left(\tfrac{\nu_1}{k}, \tfrac{\nu_1}{k} + \tfrac{\nu_2}{k-1}, \ldots, \tfrac{\nu_1}{k} + \tfrac{\nu_2}{k-1} + \cdots + \nu_n \right), \end{aligned} \tag{5}$$

$$\begin{aligned} &\mathbf{Z}(n+1,k) \stackrel{d}{=} Z(n,k) + \\ &\left(\tfrac{\nu_{kn+1}}{k}, \tfrac{\nu_{kn+1}}{k} + \tfrac{\nu_{kn+2}}{k-1}, \ldots, \tfrac{\nu_{kn+1}}{k} + \tfrac{\nu_{kn+2}}{k-1} + \cdots + \nu_{k(n+1)} \right), \\ &n = 1, 2, \ldots, \end{aligned} \tag{6}$$

$$Z_r(n+1) \stackrel{d}{=} Z_k(n) + \left(\frac{\nu_{kn+1}}{k} + \frac{\nu_{kn+2}}{k-1} + \cdots + \frac{\nu_{kn+r}}{k+1-r} \right), \ 1 \leq r \leq k, \ n = 1, 2, \ldots, \tag{7}$$

$$\begin{aligned} Z_k(n) \stackrel{d}{=} \Big(&\tfrac{\nu_1 + \nu_2 + \cdots + \nu_n}{k} + \tfrac{\nu_{n+1} + \nu_{n+2} + \cdots + \nu_{2n}}{k-1} + \cdots + \\ &\nu_{k(n-1)+1} + \cdots + \nu_{kn} \Big) \end{aligned} \tag{8}$$

etc, where ν's in each of the given relations are independent random variables with a common d.f. $F(x) = 1 - \exp(-x), \ x > 0$.

The analogous representations take place for the uniform records $\mathbf{U}(n,k) = (U_1(n), U_2(n), \ldots, U_k(n))$. One of them is given below:

$$\mathbf{U}_k(n) \stackrel{d}{=} 1 - (u_1^{(1)})^{\frac{1}{k}} (u_2^{(1)})^{\frac{1}{k}} \cdot \ldots \cdot (u_n^{(1)})^{\frac{1}{k}} (u_1^{(2)})^{\frac{1}{k-1}} \cdot \ldots \cdot (u_n^{(2)})^{\frac{1}{k-1}} \cdot \ldots \cdot (u_1^{(k)}) \cdot \ldots \cdot (u_n^{(k)}), \tag{9}$$

where random variables $u_r^{(m)},\ r = 1, 2, \ldots;\ m = 1, 2, \ldots, k$, are independent and have the common uniform $U([0, 1])$ distribution.

Using the representations given above and some analogous (see, for example, [3]) one can obtain that for any continuous d.f. F the following equalities are valid:

$$F(X_r(n)) \stackrel{d}{=} U_r(n) \stackrel{d}{=} 1 - \exp(-Z_r(n)),\ r = 1, 2, \ldots, k;\ n = 1, 2, \ldots.$$

5. Remarks

a) Looking at (5) and (8) it isn't difficult to see that

$$Z_k(n) \stackrel{d}{=} Z_{k,k}^{(1)} + Z_{k,k}^{(2)} + \cdots + Z_{k,k}^{(n)}, \tag{10}$$

where random variables $Z_{k,k}^{(r)}, r = 1, 2, \ldots$, are independent and have the same distribution as $\max\{\nu_1, \nu_2, \ldots, \nu_k\}$, that is the RHS of (10) unites properties of extremes and sums of independent identically distributed random variables at the same time. Hence, it is interesting to study the asymptotic behavior of $Z_k(n)$ and $X_k(n)$ as $n \to \infty$ and $k = k(n) \to \infty$. The corresponding limit laws for $X_k(n)$ under the suitable normalizing and centering were obtained by Saghatelyan in the situation when $n \to \infty$ for fixed k.

b) It is interesting to study distributions and properties of spacings of the confirmed records. In this model as spacings one can consider differences

$$X_r(n) - X_s(n),\ n = 1, 2, \ldots,\ k \geq r > s \geq 1,$$

and

$$X_r(n) - X_r(m),\ r = 1, 2, \ldots, k,\ 1 \leq m < n,$$

as well.

References

[1] Arnold B.C., Balakrishnan N., Nagaraja H.N. (1998) Records. Wiley, New York.

[2] Nevzorov V.B. (2001) Records: Mathematical Theory. American Mathematical Society, Providence, Rhode Island.

[3] Saghatelyan V.K. (2008) On one new model of record values. Vestnik of St-Petersburg university, Ser.1, n.3, 144-147.

[4] Wilks S.S. (1959) Recurrence of extreme observations. J. Amer. Math. Soc., 1, n.1, 106-112.

6th St.Petersburg Workshop on Simulation (2009) 985-989

Limit theorems for the counting process of near-records[2]

Raúl Gouet[3], F. Javier López[4], Gerardo Sanz[5]

Abstract

Near-records of a sequence are observations lying within a distance a of the current record. In this paper we study the asymptotic behaviour of the number of near-records among the first n observations in a sequence of independent identically distributed continuous random variables. We give conditions for the finiteness of the total number of near-records as well as laws of large numbers and central limit theorems for their counting process.

1. Introduction

Given a sequence of observations, a near-record, as defined in [1], is an observation which is not a record but is within a distance $a > 0$ of the current record value. In [1], the authors study the asymptotic behaviour of the random variable $\xi_n(a)$, defined as the number or near-records associated to the $n-$th record value in a sequence of independent identically distributed (iid) random variables with common continuous distribution function F. In the case of $F(x) < 1$ for every $x > 0$, assuming that $\beta(a) = \lim_{x\to\infty}(1 - F(x))/(1 - F(x - a))$ exists, they show that $\xi_n(a) \xrightarrow{P} \infty$ if $\beta(a) = 0$ (i.e., when F is light-tailed), $\xi_n(a) \xrightarrow{D} Geom(\beta(a))$ if $0 < \beta(a) < 1$ (medium-tailed) and $\xi_n(a) \xrightarrow{P} 0$ if $\beta(a) = 1$ (heavy-tailed). In [10], the author gives limit theorems for $\xi_n(a)$ for continuous distributions in a maximal domain of attraction when a is allowed to depend on n and, for fixed a, limit theorems for $\log \xi_n(a)$ for some classes of light-tailed distributions.

The aim of this paper is to deepen the knowledge of the asymptotic behaviour of near-records by establishing limit theorems for their counting process, that is, for the number of near-records observed up to the $n-$th observation. Let $a > 0$ be fixed and $(X_n)_{n\geq 1}$ a sequence of iid nonnegative random variables with

[1]The work was partially supported by the Russian Foundation for Basic Research under Grant .

[2]This work was supported by by FONDAP and BASAL-CMM projects, Fondecyt grant 1060794 and project MTM2007-63769 of MEC.

[3]Universidad de Chile, E-mail: rgouet@dim.uchile.cl

[4]Universidad de Zaragoza, E-mail: javier.lopez@unizar.es

[5]Universidad de Zaragoza, E-mail: gerardo.sanz@unizar.es

common continuous distribution function F and let $I_n = 1_{\{X_n \in (M_{n-1}-a, M_{n-1}]\}}$, where $M_n = \max\{X_1, \ldots, X_n\}$ ($M_0 = -\infty$ by convention), and $D_n = \sum_{k=1}^{n} I_k$. That is, I_n is the indicator of observation n being a near-record and D_n is the number of near-records among the first n observations.

In this article we give conditions for $D_\infty < \infty$ (that is, there is a finite number of near-records along the whole sequence) and, in the case $D_\infty = \infty$, laws of large numbers and central limit theorems for D_n. In order to do so, we use a martingale approach which relates our process D_n with a sum of minima of iid random variables. This approach has been successfully applied for the study of the counting process of record and record-like statistics both in discrete and continuous settings (see [3]-[8]).

Note that, letting $N_n^\delta = \sum_{k=1}^{n} 1_{\{X_k > M_{k-1}+\delta\}}$, the number of $\delta-$records as defined in [5],[8], we have

$$D_n = N_n^{-a} - N_n^0. \tag{1}$$

N_n^0 is the number of usual records and it is well known that $N_n^0/\log n \xrightarrow{a.s.} 1$ and $(N_n^0 - \log n)/\sqrt{\log n} \xrightarrow{D} N(0,1)$. On the other hand, laws of large numbers for N_n^δ are given in [8]. This can be used in some cases (e.g. for light-tailed distributions) to directly obtain laws of large numbers for D_n since in that case $N_n^{-a}/N_n^0 \xrightarrow{a.s.} \infty$ so $D_n \sim N_n^{-a}$. However, for heavy-tailed distributions, we have $N_n^{-a}/\log n \xrightarrow{a.s.} 1$ so laws of large numbers for D_n cannot be deduced from the respective laws of N_n^{-a} and N_n^0. Moreover, central limit theorems for N_n^{-a} have not appeared in the literature so we cannot use (1) to obtain central limit theorems for D_n.

We will use the following notation. F is the common distribution function of the random variables (X_n), with density f, survival function $\overline{F} = 1 - F$ and hazard function $\lambda = f/\overline{F}$. The quantile function is defined as $m(t) = \sup\{x \geq 0 : \overline{F}(x) \geq 1/t\}$.

In the next section we give the main results of the paper; sketches of the proofs of the results are given in Section . An example is presented in Section .

2. Main results

Since we are dealing with upper extremes, there is no loss of generality in considering nonnegative random variables. Moreover, if $r_F = \sup\{x \geq 0 : F(x) < 1\} < \infty$, the asymptotic behaviour of D_n is immediately obtained:

$$\frac{D_n}{n} \xrightarrow{a.s.} \overline{F}(r_F - a); \qquad \frac{D_n - \overline{F}(r_F - a)n}{\sqrt{F(r_F - a)\overline{F}(r_F - a)n}} \xrightarrow{D} N(0,1),$$

so the only interesting case in the study of the asymptotic behaviour of D_n is $r_F = \infty$. Therefore, we consider F concentrated on $(0, \infty)$, with $r_F = \infty$. Besides, we assume that the density f is ultimately decreasing.

The first result gives conditions for the finiteness of the total number of $\delta-$records.

Proposition 1. *Under mild conditions*[1] *on the hazard function* λ*, if* $\int_0^\infty \lambda(x)^2 dx < \infty$*, then* $D_\infty < \infty$ *a.s.*

For the rest of this section we assume $\int_0^\infty \lambda(x)^2 dx = \infty$. The next result gives laws of large numbers for D_n.

Theorem 1. *Under mild conditions on λ, we have*

(a) If λ is bounded above, then

$$\frac{D_n}{\int_0^{m(n)} (\overline{F}(x-a) - \overline{F}(x))\lambda(x)/\overline{F}(x)dx} \xrightarrow{a.s.} 1.$$

(b) If $\lambda(x) \to \infty$ with λ' bounded above then

$$\frac{D_n}{\int_0^{m(n)} \overline{F}(x-a)\lambda(x)/\overline{F}(x)dx} \xrightarrow{P} 1;$$

moreover, if $|\lambda'(x)| < x^{-r}$ for some $r > 1/2$ and all x large enough, the convergence holds in the a.s. sense.

Last we state the central limit theorem for D_n (recall that we assume $\int_0^\infty \lambda(x)^2 dx = \infty$ as otherwise $D_\infty < \infty$).

Theorem 2. *Under mild conditions on λ and letting*

$$\phi(t) = \int_t^\infty (f(x-a) - f(x)) \left(\frac{2\overline{F}(x-a)}{\overline{F}(x)} - 1 \right) dx, \tag{2}$$

if any of the following conditions hold (a) λ is bounded above or (b) $\lambda(x) \to \infty$ with λ' bounded above, then

$$\frac{D_n - \int_0^{m(n)} (f(x-a) - f(x))/\overline{F}(x)dx}{\left(\int_0^{m(n)} \phi(x)\lambda(x)/\overline{F}(x)dx \right)^{1/2}} \xrightarrow{D} N(0,1). \tag{3}$$

3. Sketches of proofs

In this section we give some ideas of the proofs of the results in Section . The key result for Proposition 1 and Theorem 1(a) is the following proposition, which connects D_n with the sum of minima of iid random variables (Theorem 1(b) is directly deduced from the behaviour of usual records and $\delta-$records studied in [8] using identity (1)). We denote $\mathcal{F}_k = \sigma(X_1, \ldots, X_k)$.

Proposition 2. *Let f be decreasing and $g(t) = \overline{F}(t-a) - \overline{F}(t)$. Then*

(a) $E[I_k \mid \mathcal{F}_{k-1}] = g(M_{k-1})$.

[1]The statements of the results include the hypothesis "under mild conditions on λ". These conditions refer to smoothness conditions of the hazard function λ and vary from statement to statement; they have not been explicited for a matter of space. The main hypotheses for the results are written explicitly in their statements.

(b) $D_n \sim \sum_{k=1}^{n} \min\{Y_1, \ldots, Y_k\}$ *a.s., where the random variables* Y_n *are iid with distribution function* $G(t) = \overline{F}(g^{-1}(t))$.

(c) If $\int_0^\infty \lambda^2(x)dx < \infty$, *then* $\sum_{n=1}^{\infty} \min\{Y_1, \ldots, Y_n\} < \infty$ *a.s.; otherwise, under the conditions of Theorem 1(a),*

$$\frac{\sum_{k=1}^{n} \min\{Y_1, \ldots, Y_k\}}{\int_0^{m(n)} (\overline{F}(x-a) - \overline{F}(x))\lambda(x)/\overline{F}(x)dx} \xrightarrow{a.s.} 1. \tag{4}$$

Proof. Part (a) is a simple calculation. For part (b), note that, since f is decreasing, g is also decreasing, so

$$\sum_{k=1}^{n} E[I_k \mid \mathcal{F}_{k-1}] = \sum_{k=1}^{n} g(M_{k-1}) = \sum_{k=1}^{n} \min\{g(X_1), \ldots, g(X_{k-1})\}.$$

It is immediate that the distribution function of $g(X_k)$ is G and the conditional Borel-Cantelli lemma gives the result.

(c) The sum of minima of iid random variables is studied in [2], which asserts that if $\int_0^\infty G^{-1}(1/x)dx < \infty$ then $\sum_{n=1}^{\infty} \min\{Y_1, \ldots, Y_n\} < \infty$ a.s. and it can be shown that $\int_0^\infty \lambda^2(x)dx < \infty$ implies $\int_0^\infty G^{-1}(1/x)dx < \infty$. Otherwise, when $\int_0^\infty \lambda^2(x)dx = \infty$, conditions in [2] for the strong convergence of $\sum_{k=1}^{n} \min\{Y_1, \ldots, Y_k\}$ are equivalent to

$$\sum_{n=2}^{\infty} \frac{ng(m(n))^2}{\left(\sum_{k=2}^{n} g(m(k))\right)^2} < \infty; \qquad \lim_{t\to\infty} \frac{\int_0^{m(t\log t)} g(u)\lambda(u)/\overline{F}(u)du}{\int_0^{m(t)} g(u)\lambda(u)/\overline{F}(u)du} = 1,$$

which hold under the conditions of Theorem 1(a). □

The following proposition gives the main steps in the proof of Theorem 2.

Proposition 3. *Let* f *be decreasing and* $\psi(t) = \int_0^t (f(x-a) - f(x))/\overline{F}(x)dx$. *Under the conditions of Theorem 2 we have:*

(a) The process $Z_n = D_n - \psi(M_n)$ *is a martingale.*

(b) The martingale Z_n *is cubic integrable and the increments of the process of conditional variances are*

$$E[(Z_k - Z_{k-1})^2 \mid \mathcal{F}_{k-1}] = \phi(M_{k-1}), \tag{5}$$

where ϕ *is defined in* (2).

(c) $\sum_{k=1}^{n} E[(Z_k - Z_{k-1})^2 \mid \mathcal{F}_{k-1}] \sim \sum_{k=1}^{n} \min\{R_1, \ldots, R_k\}$ *a.s., where* R_n *are iid random variables with distribution function* $G(t) = \overline{F}(\phi^{-1}(t))$.

(d)

$$\frac{\sum_{k=1}^{n} E[(Z_k - Z_{k-1})^2 \mid \mathcal{F}_{k-1}]}{\int_0^{m(n)} \phi(x)\lambda(x)/\overline{F}(x)dx} \xrightarrow{P} 1.$$

(e)

$$\frac{D_n - \psi(M_n)}{\left(\int_0^{m(n)} \phi(x)\lambda(x)/\overline{F}(x)dx\right)^{1/2}} \xrightarrow{D} N(0,1).$$

(f)

$$\frac{\psi(M_n) - \psi(m(n))}{\left(\int_0^{m(n)} \phi(x)\lambda(x)/\overline{F}(x)dx\right)^{1/2}} \xrightarrow{P} 0$$

Proof. Parts (a) and (b) are simple calculations. For (b) recall that $Z_k - Z_{k-1} = I_k + \psi(M_k) - \psi(M_{k-1})$, $I_k^2 = I_k$ and $I_k(\psi(M_k) - \psi(M_{k-1})) = 0$.

(c) Since ϕ is decreasing, we have $\phi(M_{k-1}) = \min\{\phi(X_1), \ldots, \phi(X_{k-1}))\}$. It is a matter of simple checking that the distribution of $\phi(X_k)$ is $G(t) = \overline{F}(\phi^{-1}(t))$.

(d) From (c) it suffices to show that

$$\frac{\sum_{k=1}^n \min\{R_1, \ldots, R_k\}}{\int_0^{m(n)} \phi(x)\lambda(x)/\overline{F}(x)dx} \xrightarrow{P} 1.$$

Following [2], this amounts to showing:

$$\int_1^t \phi(m(x))dx = \int_0^{m(t)} \frac{\phi(x)\lambda(x)}{\overline{F}(x)}dx; \qquad \lim_{n\to\infty} \frac{\sum_{k=2}^n k\phi(m(k))^2}{\left(\sum_{k=2}^n \phi(m(k))\right)^2} = 0$$

and the existence of an increasing sequence v_n such that

$$\lim_{n\to\infty} \frac{\int_0^{m(nv_n)} \phi(x)\lambda(x)/\overline{F}(x)dx}{\int_0^{m(n)} \phi(x)\lambda(x)/\overline{F}(x)dx} = 1.$$

Their proofs differ substantially between conditions (a) and (b) of Theorem 2.

(e) It follows from (a-d) and a Lyapunov version of the martingale central limit theorem [9] (the conditional Lyuapunov condition needs some additional results not explicited here).

(f) Results on the difference between the maxima of the sequence M_n and the quantile function $m(n)$ are used to show the convergence. As in (d), the proof differs substantially between conditions (a) and (b) of Theorem 2. □

Theorem 2 follows immediately from parts (e) and (f) of this proposition.

4. Example

Our results can be applied to any continuous distribution having an ultimately decreasing density function. We present here a class which includes heavy, medium and light tails distributions. Let $\alpha > 0$, $r \in [-1, 1]$ and let the distribution F have hazard function $\lambda(x) = \alpha x^r$ for $x > 0$ (in the case $r = -1$, let $\lambda(x) = \alpha x^{-1}$ for $x > x_0 > 0$). Then F is heavy-tailed for $r < 0$, has an exponential tail for $r = 0$ and is light-tailed for $r > 0$. The following proposition summarizes the asymptotic behaviour of D_n which, as can be expected, depends heavily on the value of r.

Proposition 4. *(a) If $r < -1/2$ then $N_\infty < \infty$ a.s.*

(b) If $r = -1/2$ then $D_n/a_n \xrightarrow{a.s.} 1$ and $(D_n - a_n)/\sqrt{a_n} \xrightarrow{D} N(0,1)$ with $a_n = 2\alpha^2 a \log\log n$.

(c) If $r \in (-1/2, 0)$ then $D_n/a_n \xrightarrow{a.s.} 1$ and $(D_n - \psi(m(n))/\sqrt{a_n} \xrightarrow{D} N(0,1)$ where $a_n = \frac{a\alpha^2}{2r+1}\left(\frac{r+1}{\alpha}\log n\right)^{\frac{2r+1}{r+1}}$ and $m(n) = \left(\frac{r+1}{\alpha}\log n\right)^{1/(r+1)}$.

(d) If $r = 0$ then $D_n/a_n \xrightarrow{a.s.} 1$ and $(D_n - a_n)/\sqrt{(2e^{a\alpha} - 1)a_n} \xrightarrow{D} N(0,1)$ with $a_n = (e^{a\alpha} - 1)\log n$.

(e) If $r \in (0,1)$ then $D_n/a_n \xrightarrow{P\ (a.s.)} 1$ (the convergence holds in probability for $r \in (0,1)$ and a.s. for $r \in (0,1/2)$) and $(D_n - \psi(m(n)))/\sqrt{c_n} \xrightarrow{D} N(0,1)$, with $a_n = \frac{m(n)}{ar}e^{\alpha a m(n)^r}$, $c_n = \frac{m(n)}{ar}e^{2\alpha a m(n)^r}$ and $m(n) = \left(\frac{r+1}{\alpha}\log n\right)^{1/(r+1)}$.

(f) If $r = 1$ then $D_n/a_n \xrightarrow{P} 1$ and $(D_n - b_n)/\sqrt{c_n} \xrightarrow{D} N(0,1)$, with $a_n = e^{-\alpha a^2/2} m(n) e^{\alpha a m(n)}/a$, $b_n = e^{-\alpha a^2/2} e^{\alpha a m(n)}\left(m(n) - (a + \frac{1}{\alpha a})\right)/a$, $c_n = e^{-\alpha a^2} m(n) e^{2\alpha a m(n)}/a$ and $m(n) = \sqrt{(2/\alpha)\log n}$.

References

[1] Balakrishnan N., Pakes A.G., Stepanov A. (2005) On the number and sum of near-record observations. *Adv. in Appl. Probab.*, **37**, 765–780.

[2] Deheuvels P. (1974) Valeurs extrémales d'échantillons croissants d'une variable aléatoire réelle. *Ann. Inst. H. Poincaré*, **X**, 89–114.

[3] Gouet R., López F.J., San Miguel M. (2001) A martingale approach to strong convergence of the number of records. *Adv. in Appl. Probab.*, **33**, 864–873.

[4] Gouet R., López F.J., Sanz G. (2005) Central limit theorems for the number of records in discrete models. *Adv. in Appl. Probab.*, **37**, 781–800.

[5] Gouet R., López F.J., Sanz G. (2007) Asymptotic normality for the counting process of weak records and δ-records in discrete models. *Bernoulli*, **13**, 754–781.

[6] Gouet R., López F.J., Sanz G. (2008) Laws of large numbers for the number of weak records. *Statist. Probab. Lett.*, **78**, 2010–2017.

[7] Gouet R., López F.J., Sanz G. (2009) Limit laws for the cumulative number of ties for the maximum in a random sequence. *J. Statist. Plann. Inference.* doi:10.1016/j.jspi.2009.02.001

[8] Gouet R., López F.J., Sanz G. (2009) Laws of large numbers for the counting process of δ-records in general distributions. *Submitted.*

[9] Hall P., Heyde C.C. (1980) *Martingale Limit Theory and its Application.* Academic Press, New York.

[10] Pakes A.G. (2007) Limit theorems for numbers of near-records. *Extremes*, **10**, 207–224.

6th St.Petersburg Workshop on Simulation (2009) 991-995

Random motions in hyperbolic spaces

Enzo Orsingher[2], Valentina Cammarota[3]

Abstract

We consider the Poincaré half-plane $\mathbb{H}_2^+$ and a random motion at finite velocity on its geodesic lines.

A particle starting from the origin O of $\mathbb{H}_2^+$ moves at hyperbolic finite velocity c on the geodesic line (with probability $\frac{1}{2}$ on either directions). At Poisson spaced times it moves on the orthogonal line (in one of the two possible directions) until a second Poisson event occurs, then it deviates orthogonally with respect to the half-circle through O and the current position.

After $N(t)$ Poisson events the current position has hyperbolic distance equal to

$$\cosh \eta(t) = \prod_{k=1}^{N(t)+1} \cosh c(t_k - t_{k-1}) \tag{1}$$

where t_k are the random instants where the deviations of motion happen $t_{N(t)+1} = t$, $t_0 = 0$. Formula (1) is obtained by successively applying the hyperbolic Pythagorean theorem and we are able to obtain that

$$E \cosh \eta(t) = e^{-\frac{\lambda t}{2}} \quad \cosh \frac{t\sqrt{\lambda^2 + 4c^2}}{2} + \frac{\lambda}{\sqrt{\lambda^2 + 4c^2}} \sinh \frac{t\sqrt{\lambda^2 + 4c^2}}{2}$$

We study also the case where a branching process is associated with the random motion on $\mathbb{H}_2^+$. At each deviation the particle splits into two splinters of equal size, one of which deviates orthogonally and the other continues its motion on the previous geodesic line.

At time t the hyperbolic distance of the center of mass of the $N(t) + 1$ moving particles created in this process equals

$$E\{\cosh \eta_{cm}(t) | N(t) = n\} \quad =$$

$$= \quad \frac{n!}{t^n} \sum_{k=0}^{n-1} \frac{1}{2^{k+1}} \int_0^t \mathrm{d}t_1 \cdots \int_{t_{n-1}}^t \mathrm{d}t_n \prod_{j=1}^{k+1} \cosh c(t_j - t_{j-1})$$

$$+ \frac{n!}{t^n} \frac{1}{2^n} \int_0^t \mathrm{d}t_1 \cdots \int_{t_{n-1}}^t \mathrm{d}t_n \prod_{j=1}^{n+1} \cosh c(t_j - t_{j-1}).$$

[2]Sapienza University of Rome, E-mail: enzo.orsingher@uniroma1.it
[3]Sapienza University of Rome, E-mail: valentina.cammarota@uniroma1.it

We are able to show in two different and independent ways that

$$E\{\cosh \eta_{cm}(t)\} = \frac{2^3 c^2 e^{-\frac{3}{2^2}\lambda t}}{\sqrt{\lambda^2+2^4c^2}} \left\{ \frac{e^{-\frac{t}{2^2}\sqrt{\lambda^2+2^4c^2}}}{3\sqrt{\lambda^2+2^4c^2}+5\lambda} + \frac{e^{\frac{t}{2^2}\sqrt{\lambda^2+2^4c^2}}}{3\sqrt{\lambda^2+2^4c^2}-5\lambda} \right\} + \frac{\lambda+2c}{2(\lambda+3c)} e^{ct} + \frac{\lambda-2c}{2(\lambda-3c)} e^{-ct}.$$

All results above are examined also on the Poincaré disc and the corresponding dynamics illustrated.

References

[1] Cammarota, V., E. Orsingher: Cascades of Particles Moving at Finite Velocity in Hyperbolic Spaces, *J. Stat. Phys.*, **133**, 1137–1159 (2008)

[2] Cammarota, V., Orsingher, E.: Travelling randomly on the Poincaré halfplane with a Pythagorean compass. *J. Stat. Phys.*, **130**, 455–482 (2008)

[3] Faber, R. L.: Foundations of Euclidean and Non-Euclidean Geometry. Dekker, New York (1983)

[4] Gertsenshtein, M. E., Vasiliev, V. B.: Waveguides with random inhomogeneities and Brownian motion in the Lobachevsky plane. *Theory Probab. Appl.*, **3**, 391–398 (1959)

[5] Getoor, R. K.: Infinitely divisible probabilities on the hyperbolic plane. *Pacific J. Math.*, **11**, 128–1308 (1961)

[6] Gruet, J. C.: Semi-groupe du mouvement Brownien hyperbolique. *Stochast. Stochast. Rep.*, **56**, 53–61 (1996)

[7] Gruet, J. C.: A note on hyperbolic von Mises distributions. *Bernoulli*, **6**, 1007–1020 (2000)

[8] Kelbert, M., Suhov, Yu. M.: Branching diffusions on H^d with variable fission: the Hausdorff dimension of the limiting set. *Theory Probab. Appl.*, **51**, 155–167 (2007)

[9] Kelbert, M., Suhov, Yu.: Large-time behaviour of a branching diffusion on e hyperbolic space. *Manuscript*

[10] Lao, L., Orsingher, E.: Hyperbolic and fractional hyperbolic Brownian motion. *Stochastics*, **79**, 505–522 (2007)

[11] Lalley, S. P., Sellke, T.: Hyperbolic branching Brownian motion. *Probab. Theory Relat. Fields*, **108**, 171–192 (1997)

[12] Meschkowski, H.: Non-Euclidean Geometry. Academic Press, New York (1964)

[13] Orsingher, E., De Gregorio, A.: Random motions at finite velocity in a non-Euclidean space. *Adv. Appl. Prob.*, **39**, 588–611 (2007)

[14] Rogers, L. C. G., Williams, D.: Diffusions, Markov Processes, and Martingales. Wiley, Chirchester (1987)

[15] Yor, M.: On some exponential functionals of Brownian motion. *Adv. Appl. Prob.*, **24**, 509–53 (1992)

Section

Applied stochastic procedures

6th St.Petersburg Workshop on Simulation (2009) 995-999

Wavelet-based detection of outliers in volatility models[1]

Aurea Grané[2], Helena Veiga[3]

Abstract

Outliers in financial data can lead to model parameter estimation biases, invalid inferences and poor volatility forecasts. Therefore, their detection should be taken seriously when modeling financial data. This paper focuses on these issues and proposes a general detection method based on wavelets that can be applied to a large class of volatility models. The effectiveness of our proposal is tested by an intensive Monte Carlo study for six well known volatility models and compared to alternative proposals in the literature.

1. Introduction

Financial time series typically exhibit excess of kurtosis and volatility clustering, which consists of periods of high (low) volatility followed by periods of high (low) volatility. Several models had been proposed in the literature with the aim to capture these features. The ARCH model by Engle (1982) and the GARCH model by Bollerslev (1986) became benchmarks models in finance, specially due to their easy applicability and flexibility in allowing for simple extensions that better fit the empirical facts of financial data. Indeed, since the estimated standardized residuals computed from the GARCH model often have excess of kurtosis, Bollerslev (1987) introduced a t-distributed GARCH model by allowing the error term to follow a Student's t distribution. This slight modification allows the model to reach levels of kurtosis more comparable to the ones observed in the data. However, it can be observed that the estimated residuals from this extension still register excess of kurtosis (see Baillie and Bollerslev 1989; Teräsvirta, 1996). One possible reason for this to occur is that some observations on returns, which are called additive outliers (AO), are not fitted by a gaussian GARCH model and even by a t-distributed GARCH model. The additive outliers can be level outliers (ALO) in the sense that they have effects on the level of the series but not on the evolution of the underlying volatility or volatility outliers (AVO) (see Hotta and Tsay, 1998; Sakata and White, 1998). This last type of additive outliers also affects the conditional

[1]This work was supported by grants MTM2006-09920, SEJ2007-64500 and SEJ2006-03919 (Spanish Ministry of Education and Science and FEDER).

[2]Universidad Carlos III de Madrid, E-mail: aurea.grane@uc3m.es

[3]Universidad Carlos III de Madrid, E-mail: mhveiga@est-econ.uc3m.es

variance. Neglecting the existence of these outliers leads to biased parameter estimates (see, for example, Fox, 1972; Van Dijk et al.1999), undesirable effects on the tests of conditional homoskedasticity (see Carnero et al., 2007) and to biased out-of-sample forecasts (see for instance Ledolter, 1989; Chen and Miu, 1993a, Franses and Ghijsels, 1999).

This paper focuses mainly on additive (level and volatility) outliers. The effects of innovative outliers on the dynamic properties of the series are less important because they are propagated by the same dynamics like in the rest of the series (see for example Peña, 2001). Our approach is inspired by Bilen and Huzurbazar (2002) who proposed an outlier detection method based on wavelets, but our method departs from theirs in the way the threshold limits are obtained. Our proposal deals with the estimated model residuals because we are interested in detecting if an "abnormal" observation is an outlier for a particular volatility model. Our method is applied to several volatility models, such as: the GARCH, the GJR-GARCH (see Glosten et al., 1993) and the autoregressive stochastic volatility model (ARSV) by Taylor (1986) with errors following a gaussian or a Student's t distribution. The Monte Carlo results show that our proposal is not only as good as that of Bilen and Huzurbazar (2002) in detecting outliers, whenever both methods can be applied, but also it is much reliable since it detects a significant smaller number of false outliers.

2. Additive outliers in volatility models

The GARH(p, q) models proposed by Bollerslev (1986) and (1987) are given by:

$$y_t = \mu + \varepsilon_t = \mu + \sigma_t \, \epsilon_t,$$

where μ is the conditional mean, ε_t is the prediction error, σ_t^2 is the variance of y_t given information at time $t-1$, $\sigma_t > 0$, $\epsilon_t \sim NID(0,1)$ or follows a Student's t distribution and $\sigma_t^2 = \omega + \theta(L)\,\varepsilon_t^2$, where $\theta(L) = 1 - \frac{\alpha^*(L)}{\beta(L)}$, with $\alpha^*(L) = 1 - \sum_{i=1}^{q} \alpha_i^* \, L^i$, $\beta(L) = 1 - \sum_{i=1}^{p} \beta_i \, L^i$ and $\omega > 0$. In a GARCH(1,1) model with $0 \leq \beta_1 < 1$, the conditional variance equation can be written as

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2,$$

where $\alpha_0 = \omega(1 - \beta_1)$.

The GJR(1,1) model differs from the GARCH(1,1) since it introduces the possibility that positive and negative shocks might affect differently the conditional variance σ_t^2. In fact, the conditional variance equation is given by

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \gamma_1 S_{t-1}^- \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2,$$

where S_t^- is a dummy variable that assumes the value 1 when ε_t is negative and 0 otherwise. Once more, $\epsilon_t \sim NID(0,1)$ or follows a Student's t distribution.

In the context of stochastic volatility, a natural competitor to the GARCH and GJR models is the autoregressive stochastic volatility model (denoted ARSV(1))

by Taylor (1986). The ARSV model is given by the following expressions:

$$\begin{aligned} y_t &= \mu + \sigma\,\epsilon_t\,\exp\left(\frac{h_t}{2}\right), \\ (1-\phi L)\,h_t &= \eta_t. \end{aligned} \tag{1}$$

In equation (1), μ is the mean of y_t, σ denotes a scale parameter, $\sigma_t = \exp(h_t/2)$ is the volatility of y_t (the return at time t), ϵ_t is $NID(0,1)$ or follows a Student's t distribution and η_t is $NID(0,\sigma_\eta^2)$.

2.1. Additive level outliers (ALO)

The conditional mean equations of the GARCH(1,1) and the GJR(1,1) models with an additive level outlier are defined as:

$$y_t = \mu + \omega_{AO} I_T(t) + \varepsilon_t = \mu + \omega_{AO} I_T(t) + \sigma_t\,\epsilon_t,$$

where ω_{AO} represents the magnitude of the additive level outlier and $I_T(t) = 1$ for $t \in T$ and 0 otherwise, representing the presence of the outlier at a set of times T. The equations of the conditional variances for the two models remain the same since this type of outlier only affects the level of the series.

In the context of stochastic volatility, the additive level outlier is defined as:

$$\begin{aligned} y_t &= \mu + \omega_{AO} I_T(t) + \sigma\,\epsilon_t\,\exp\left(\frac{h_t}{2}\right), \\ (1-\phi L)\,h_t &= \eta_t, \end{aligned}$$

where ω_{AO} and $I_T(t)$ are defined as before. Examples of additive level outliers may be an institutional change or a market correction that does not affect volatility.

2.2. Additive volatility outliers (AVO)

An additive volatility outlier for the GARCH(1,1) model is defined as:

$$\begin{aligned} y_t = \mu + \varepsilon_t^* &= \mu + \sigma_t^*\,\epsilon_t, \\ \varepsilon_t^* &= \omega_{AO} I_T(t) + \varepsilon_t, \\ \sigma_t^{*2} &= \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^{*2} + \alpha_1 \left(2\omega_{AO}\varepsilon_{t-1} + \omega_{AO}^2\right) I_T(t-1), \end{aligned}$$

and for the GJR(1,1) model:

$$\begin{aligned} y_t = \mu + \varepsilon_t^* &= \mu + \sigma_t^*\,\epsilon_t, \\ \varepsilon_t^* &= \omega_{AO} I_T(t) + \varepsilon_t, \\ \sigma_t^{*2} &= \alpha_0 + \left(\alpha_1 + \gamma_1 S_{t-1}^-\right)\varepsilon_{t-1}^2 \\ &+ \left(\alpha_1 + \gamma_1 S_{t-1}^-\right)\left(2\omega_{AO}\varepsilon_{t-1} + \omega_{AO}^2\right) I_T(t-1) + \beta_1 \sigma_{t-1}^{*2}. \end{aligned}$$

where ω_{AO} represents the magnitude of the additive level outlier and $I_T(t) = 1$ for $t \in T$ and 0 otherwise, representing the presence of the outlier at a set of times T as in subsection 2.1 and S_t^- is a dummy variable that assumes the value 1 when ε_t is negative and 0 otherwise. Note that in both cases, $\epsilon_t \sim NID(0,1)$ or follows a Student's t distribution. Therefore, the effect of the AVO outlier affects not only the volatility but also the series level. Its effect in the original series is similar to a patch of ALO outliers with decreasing magnitudes when $\beta_1 < 1$.

3. Wavelet-based detection procedure

3.1. The procedure

The procedure we propose is based on the detail coefficients resulting from the discrete wavelet transform (DWT) of the series of residuals ($\hat{\epsilon}_t$), obtained after fitting a particular model. The outliers are identified as those observations in the original series whose detail coefficients are greater (in absolute value) than a certain threshold. In the context of financial time series, it is very common to assume that if the fitted model has captured the structure of the data, then the residuals are supposed to be independent and identically distributed (iid) random variables following either a standard normal or a Student's t distribution. Using a Monte Carlo scheme, we have obtained, for different sample sizes, the distribution of the maximum of the detail coefficients (in absolute value) resulting from the DWT of iid random variables following either a standard normal or a Student's t distribution. We have taken the threshold to be the 95%-percentile of the distribution of this maximum. In practice, we have found that in order to detect isolated additive level outliers (ALOs) it is enough to work with the first level detail wavelet coefficients. Using the inverse discrete wavelet transform, the procedure identifies the outliers recursively, one by one, avoiding the masking effect. However, if there are patches of ALOs or isolated additive volatility outliers (AVOs) it is necessary to use the first and second level detail wavelet coefficients to identify the influential observations.

3.2. Performance of the procedure: A simulation study

The study involves single, multiple and patches of additive level outliers (ALOs) and single additive volatility outliers (AVOs) observed in different financial models, such as GARCH, GJR and ARSV with errors following either a gaussian or a Student's t distribution. The outliers are placed randomly along the time series and the simulation study involves 1000 replications for each scenario. The measures used in the performance study are the proportion of times that the location of the outliers is correctly detected jointly with the average number of false detections and their standard errors.

We have considered magnitudes of isolated ALOs of $\omega_{AO} = 5\sigma_y, 10\sigma_y, 15\sigma_y$ and sample sizes of $n = 500, 1000, 5000$, where σ_t is the standard deviation of y_t. The frequency of the simulations is daily and the parameters used are: $\{\alpha_0 = 0.0126, \alpha_1 = 0.0757, \beta = 0.9122\}$ for the GARCH model, $\{\alpha_0 = 0.0000, \alpha_1 = 0.0139, \beta = 0.9139, \gamma_1 = 0.1106\}$ for the GJR model and $\{\phi = 0.98, \sigma_\eta^2 = 0.05, \sigma = 1\}$ for the ARSV model, which have been chosen by fitting the models to real return series. We have considered patches of three ALOs in the same volatility models as before with magnitudes of $\omega_{A0} = 10\sigma_y, 15\sigma_y$ and sample sizes of $n = 500, 1000, 5000$. The beginning of the patch was placed randomly in the time series. Concerning additive volatility outliers, we have simulated single AVOs of magnitude $\omega_{AO} = 15\sigma_y, 25\sigma_y, 50\sigma_y$ in GARCH and GJR models with errors following either a gaussian or a Student's t distribution for a sample size of $n = 1000$.

Here we only reproduce the results for the single and multiple ALOs. Tables 1 and 2 contain these results.

Table 1: Percentage of correct detection of additive level outliers in 1000 replications of size n for various volatility models with errors following a normal or a Student's t distribution.

		$N(0,1)$						$t(7)$		
		GARCH		ARSV		GJR		GARCH	ARSV	GJR
	n	G&V	B&H	G&V	B&H	G&V	B&H			
1 outlier	500	66.3	93.4	64.1	87.1	25.8	95.5	39.9	17.1	30.3
of size	1000	66.0	90.2	63.6	85.3	20.7	92.6	41.9	17.5	37.1
$\omega_{AO} = 5$	5000	59.6	86.1	60.5	80.1	10.7	93.1	88.5	9.1	56.9
1 outlier	500	98.4	100	96.9	99.7	92.1	98.5	68.1	68.5	78.4
of size	1000	98.9	99.9	96.0	99.0	92.9	99.0	70.3	66.1	77.9
$\omega_{AO} = 10$	5000	98.4	99.7	94.0	97.6	97.1	99.8	93.5	52.5	78.2
1 outlier	500	99.0	100	99.5	99.9	91.0	97.9	79.3	93.1	92.2
of size	1000	99.7	100	99.6	100	94.8	98.3	79.8	88.5	91.4
$\omega_{AO} = 15$	5000	99.8	99.9	98.6	99.8	99.6	100	95.8	80.7	87.5
3 outliers	500	63.3	91.8	71.3	95.2	63.7	92.5	35.9	40.8	51.8
of sizes	1000	71.4	92.0	76.9	94.9	64.9	92.1	48.8	47.6	61.3
$\omega_{AO}=$ 5,10,15	5000	77.6	90.3	81.3	92.5	65.1	91.9	80.9	45.7	73.2

(*) G&V stands for our method and B&H for Bilen and Huzurbazar's.

From Table 1 we see that when the magnitude of the outliers is $\omega_{AO} = 10\sigma_y, 15\sigma_y$, the procedure detects more than 90% of single and multiple outliers, for models with gaussian errors. When the errors follow a Student's t distribution, the detection rate goes from 52% to 95%, being around 80% in mean. Additionally, the average number of false detections is no greater than 1, (note from Table 2 that it is no greater than 0.1 in practically all cases). Moreover, we observe from the outlier detection results that the ARSV and GJR models are more robust to outliers of small size in the sense that they can not be distinguished from the observations generated by the two specifications.

Concerning to the detection of patches of additive outliers, we have seen that, in general, the detection rate is greater for models with gaussian errors, going from 41% to nearly 98%, whereas the average number of false detections is always no greater than 0.03. Regarding to single additive volatility outliers, we have seen that the detection rate is greater for models with gaussian errors, whereas the average number of false detections is always no greater than 0.004. In all situations, the sensitivity of the method increases as the magnitude ω_{AO} increases.

4. Conclusion

The existing outlier procedures in financial time series are based on the proposal by Chen and Miu (1993b) that consists in an iterative outlier detection and adjustment method to jointly estimate the model parameters and the outlier effects. However, along the iterative process they have to estimate the model several times and the estimates of the parameters can be affected by the presence of remaining outliers.

On the contrary, our outlier detection proposal is based on applying wavelets to the residuals of some volatility models. It does not need subsequential estimations

Table 2: Average number of false detections (standard deviation) of additive level outliers in 1000 replications of size n for various volatility models with errors following a normal or a Student's t distribution.

		$N(0,1)$						$t(7)$		
		GARCH		ARSV		GJR		GARCH	ARSV	GJR
	n	G&V	B&H	G&V	B&H	G&V	B&H			
1		0.02	1.96	0.14	2.50	0.02	3.64	0.001	0.01	0.01
	500									
outlier		(0.15)	(7.99)	(0.37)	(1.97)	(0.13)	(2.77)	(0.03)	(0.11)	(0.11)
of		0.05	1.91	0.24	3.43	0.01	5.02	0.01	0.02	0.02
size	1000									
ω_{AO}		(0.22)	(1.58)	(0.50)	(2.20)	(0.11)	(2.91)	(0.10)	(0.13)	(0.14)
=5		0.05	2.63	0.65	7.17	0.01	10.52	0.03	0.03	0.02
	5000									
		(0.21)	(1.73)	(0.78)	(2.92)	(0.11)	(3.74)	(0.18)	(0.18)	(0.12)
1		0.03	3.85	0.09	2.52	0.03	3.92	0.01	0.01	0.01
	500									
outlier		(0.20)	(19.25)	(0.30)	(1.97)	(0.17)	(2.90)	(0.08)	(0.08)	(0.12)
of		0.03	2.21	0.15	3.31	0.02	5.15	0.01	0.01	0.02
size	1000									
ω_{AO}		(0.16)	(1.91)	(0.40)	(2.15)	(0.13)	(3.00)	(0.10)	(0.11)	(0.13)
=10		0.04	2.71	0.57	7.16	0.01	10.54	0.03	0.03	0.02
	5000									
		(0.19)	(1.82)	(0.73)	(2.90)	(0.11)	(3.79)	(0.17)	(0.17)	(0.12)
1		0.04	5.07	0.04	2.51	0.06	4.32	0.01	0.005	0.01
	500									
outlier		(0.20)	(22.16)	(0.21)	(2.03)	(0.26)	(3.34)	(0.08)	(0.07)	(0.11)
of		0.04	4.35	0.10	3.41	0.03	5.45	0.01	0.01	0.02
size	1000									
ω_{AO}		(0.21)	(27.28)	(0.31)	(2.20)	(0.18)	(3.17)	(0.10)	(0.08)	(0.13)
=15		0.03	2.84	0.49	7.32	0.01	10.50	0.03	0.02	0.01
	5000									
		(0.17)	(1.96)	(0.71)	(2.96)	(0.12)	(3.83)	(0.17)	(0.13)	(0.12)
3		0.03	5.00	0.02	2.72	0.09	5.10	0.001	0.002	0.01
	500									
outliers		(0.19)	(8.84)	(0.15)	(2.10)	(0.33)	(4.45)	(0.03)	(0.04)	(0.11)
of		0.04	7.34	0.04	3.25	0.07	5.80	0.004	0.003	0.01
sizes	1000									
ω_{AO}		(0.24)	(41.28)	(0.22)	(2.11)	(0.29)	(3.42)	(0.06)	(0.05)	(0.11)
=5,10		0.03	3.15	0.35	7.14	0.02	10.70	0.04	0.01	0.01
	5000									
and		(0.17)	(2.09)	(0.62)	(3.00)	(0.12)	(4.03)	(0.19)	(0.12)	(0.11)
15										

(*) G&V stands for our method and B&H for Bilen and Huzurbazar's.

of the model parameters and therefore, it is not susceptible to the previous criticism. The method uses the discrete wavelet transform and detects changes in the wavelet coefficients by using thresholds based on the distribution of the maximum of the detailed coefficients (in absolute value) obtained by Monte Carlo. In this way, our method can be applied to the estimated residuals of different volatility models with errors following any known distribution.

The advantages of our proposal are several: first, it applies when the location and the number of outliers are unknown, second, the data can be generated by any known distribution; third, it is well suited for one outlier or multiple outlier detections; fourth, it is the one, as far as we know, that detects patches of outliers in different volatility models; and finally, the method is easy and quick to apply

which converts it in an attractive tool to be used by academic communities and/or by practitioners.

The effectiveness of our method is tested with simulated and it is compared with other outlier detection methods. The simulations report evidence that our proposal is not only as good as that of Bilen and Huzurbazar (2002), whenever both methods can be applied, but also much reliable since it detects a significant smaller number of false outliers. Moreover, since Bilen and Huzurbazar (2002) showed that their outlier detection procedure performed better that the ones based on likelihood ratio tests like the method by Chen and Miu (1993b), we may conclude that our detection method is better than the existent proposals in financial time series, with the advantage that we can test for patches of additive level outliers and data generated from different known distributions.

References

[1] Baillie, R. and T. Bollerslev (1989). The message in daily exchange rates: a conditional variance tale. *Journal of Business and Economic Statistics 7*, 297–309.

[2] Bilen, C. and S. Huzurbazar (2002). Wavelet-based detection of outliers in time series. *Journal of Computational and Graphical Statistics 11*, 311–327.

[3] Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics 31*, 307–327.

[4] Bollerslev, T. (1987). A conditionally heteroskedastic time series model for speculative prices and rates of return. *Review of Economic and Statistics 69*, 542–547.

[5] Carnero, M., D. Peña, and E. Ruiz (2007). Effects of outliers on the identification and estimation of GARCH models. *Journal of Time Series Analysis 28*, 471–497.

[6] Chen, C. and L. Miu (1993a). Forecasting time series with outliers. *Journal of Forecasting 12*, 13–35.

[7] Chen, C. and L. Miu (1993b). Joint estimation of model parameters and outlier effects. *Journal of American Statistical Association 88*, 284–297.

[8] Engle, R. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation. *Econometrica 50*, 987–1008.

[9] Fox, A. (1972). Outliers in time series. *Journal of Royal Statistical Society B 34*, 350–363.

[10] Franses, P. and H. Ghijsels (1999). Additive outliers, GARCH and forecasting volatility. *International Journal of Forecasting 15*, 1–9.

[11] Glosten, L., R. Jagannathan, and D. Runkle (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *Journal of Finance 48*, 1779–1801.

[12] Hotta, L. and R. Tsay (1998). Outliers in GARCH processes. *Manuscript. Graduate School of Business, University of Chicago*.

[13] Ledolter, J. (1989). The effect of additive outliers on the forecasts from ARIMA models. *International Journal of Forecasting 5*, 231–240.

[14] Peña, D. (2001). Outliers, influential observations and missing data. In D. Peña, G. Tiao, and R. Tsay (Eds.), *A Course in Time Series*, New York, pp. 136–170. Wiley.

[15] Sakata, S. and H. White (1998). High breakdown point conditional dispersion estimation with aplication to S&P500 daily returns volatility. *Econometrica 66*, 529–567.

[16] Taylor, S. (1986). *Modelling Financial Time Series*. Wiley, New York.

[17] Teräsvirta, T. (1996). Two stylized facts and the GARCH(1,1) model. *Working Paper 96, Stockholm School of Economics*.

[18] Van Dijk, D., P. Franses, and A. Lucas (1999). Testing for ARCH in the presence of additive outliers. *Journal of Applied Econometrics 14*, 539–562.

6th St.Petersburg Workshop on Simulation (2009) 1003-1007

Comparative Study of Effective Bandwidth Estimators: Batch Means and Regenerative Cycles

Irina Dyudenko[1], Evsey Morozov[2], Michele Pagano[3], Werner Sandmann[4]

Abstract

We consider the estimation of effective bandwidths in single-server queueing networks with finite buffers and regenerative input process. Drawbacks of batch means estimators in simulation practice are discussed and a new regenerative estimator is suggested.

1. Introduction

In order to characterize the quality of service (QoS) offered by a communication network, one of the most relevant parameters is the packet loss *ratio*, which can be estimated as the buffer overflow probability where the buffer size b of the bottleneck router along the path of the traffic is considered as threshold. The minimum capacity C_Γ that guarantees a mean overflow probability of at most Γ is called the effective bandwidth (EB) of the incoming traffic where in practice Γ is given as a QoS constraint for the maximum acceptable packet loss rate [1, 2, 3].

Effective bandwidth estimation can be treated on the basis of large deviations theory (LDT).

If the queueing system is stable, the weak limit $W_n \Rightarrow W$ of the queue size (workload) process $(W_n)_{n\in\mathbb{N}}$ exists and the stationary workload W *(under mild assumptions (see [10]))* satisfies a large deviations principle (LDP) such that the overflow probability has an asymptotically exponential form [10]. More precisely, let

$$\Lambda(\theta) = \lim_{n\to\infty} \frac{1}{n} \log \mathbb{E} e^{\theta \sum_{i=1}^{n} X_i} \tag{1}$$

the logarithmic scaled cumulant generating function (LSCGF) of the arrival process where X_i denotes the number of arrivals during the ith time unit. Define

$$\delta(C) := \sup\{\theta > 0 : \Lambda(\theta) \le C\theta\}. \tag{2}$$

The work of two first authors is supported by RFBR, project 07-07-00088.

[1]Institute of Applied Mathematical Research, Russian Academy of Sciences, E-mail: dyudenko@krc.karelia.ru

[2]Institute of Applied Mathematical Research, Russian Academy of Sciences, E-mail: emorozov@karelia.ru

[3]University of Pisa, Italy, E-mail: m.pagano@iet.unipi.it

[4]University of Bamberg, Germany, E-mail: werner.sandmann@uni-bamberg.de

Then, provided that the limit (1) exists,

$$\lim_{b\to\infty} \frac{1}{b} \log P(W > b) = -\delta(C), \tag{3}$$

which in turn implies the approximation $P(W > b) \approx e^{-\delta(C)b}$ for the overflow probability. The required effective bandwidth C_Γ for a given maximum acceptable packet loss rate Γ is given by $C_\Gamma = \min\{C : e^{-\delta(C)b} \leq \Gamma\}$. As the equation $e^{-\delta(C)b} = \Gamma$ has a unique root $\theta^* := \delta(C_\Gamma) = -\log\Gamma/b$, it follows from (2) that C_Γ can be expressed as

$$C_\Gamma = \frac{\Lambda(\theta^*)}{\theta^*}.$$

Thus, estimation C_Γ is reduced to estimation of $\Lambda(\theta^*)$.

In this paper, we consider two different approaches to the estimation of $\Lambda(\theta^*)$, the batch means method and the regenerative method. The general framework addressed by both methods is the construction of confidence intervals for the steady-state mean of a covariance-stationary discrete-time stochastic process $(X_n)_{n\in\mathbb{N}}$. In the setting of queueing networks, typical covariance-stationary processes of interest include the arrival and the service process, the workload process, and the waiting time process, amongst others. The simplest approach to steady-state simulation is the replication-deletion approach where multiple independent realizations (replications) of the stochastic process under consideration are generated and for each realization an initial transient phase must be deleted, which causes an enormous overhead if a lot of replications are needed. In contrast, the batch means method and the regenerative method provide confidence intervals based on one single realization of the process. Before turning to effective bandwidth estimation we briefly outline the general underlying theory and some key properties of the methods.

2. Batch means method

With the batch means method the data from one single simulation run of length n is grouped into k batches of size m such that $n = km$. Hence, for $i = 1, \ldots, k$ the i-th batch mean and the sample variance from k batches are given by

$$Y_i = \frac{1}{m}\sum_{j=1}^{m} X_{(i-1)m+j}, \qquad S_k^2 = \frac{1}{k-1}\sum_{i=1}^{k} \left(Y_i - \bar{Y}_k\right)^2, \tag{4}$$

where the batch means sample mean equals the overall sample mean of $X_1, \ldots, X_n$:

$$\bar{Y}_k = \frac{1}{k}\sum_{i=1}^{k} Y_i = \frac{1}{km}\sum_{i=1}^{k}\sum_{j=1}^{m} X_{(i-1)m+j} = \frac{1}{n}\sum_{i=1}^{n} X_i = \bar{X}_n. \tag{5}$$

Based on the batch means, a $100(1-\alpha)\%$ confidence interval is constructed by

$$I = \left[\bar{Y}_k - t_{k-1,1-\alpha/2}\frac{S_k}{\sqrt{k}}\ ,\ \bar{Y}_k + t_{k-1,1-\alpha/2}\frac{S_k}{\sqrt{k}}\ \right], \tag{6}$$

where $t_{k-1,1-\alpha/2}$ denotes the $1-\alpha/2$ quantile of the Student t-distribution with $k-1$ degrees of freedom.

The justification of constructing confidence intervals based on batch means relies on central limit theorems for covariance-stationary processes, according to which the batch means are asymptotically normal as the batch size approaches infinity. Furthermore, it is shown in [7] that if

$$0 < \sum_{i=-\infty}^{\infty} |Cov(X_j, X_{j+i})| < \infty, \tag{7}$$

then all lag-autocorrelations in the batch means process vanish as the batch size approaches infinity. These results seem to indicate that everything becomes fine when the batch size is chosen sufficiently large. However, asymptotic results do not strictly apply in practice and can be misleading for finite simulation run lengths.

The crucial point is to find a reasonable balance between the batch size and the number of batches. On the one hand, we have to assure by a sufficiently large batch size that the batch means are at least approximately i.i.d. normal in order to achieve a coverage probability close to the nominal value given by the confidence level. On the other hand, we need sufficiently many batches in order to construct reliable confidence intervals that are reasonably narrow and stable.

The study that probably most influenced the choice of the number of batches in practical applications of the batch means method is [8], which is often summarized overly simplified by just citing the recommendation of $10 \leq k \leq 30$ batches. We believe that it is important to know the framework and some of the details that led to this recommendation. In [8], the existence of a maximum number of batches $k^* \geq 2$ and a corresponding minimum batch size $m^* = n/k^*$ is assumed such that for all $k \leq k^*$, the dependency and the nonnormality of the batch means are "negligible" (in an intuitive sense, where a formalization remains open), and only the effects of batch sizes $m \geq m^*$ are studied. In this setting more batches imply a smaller expected confidence interval width but also a smaller coverage probability. According to [8], for all confidence levels the expected width of the confidence interval monotonically increases but the decrease rate quickly decays with increasing number of batches. The standard deviation and the coefficient of variation are much more sensitive to choice of k. Consequently, more than 30 batches can be reasonable if the confidence interval stability is important.

Another important point to note is that in practice we usually do not know suitable k^* or m^* as assumed in [8]. In most simulations some nonnormality or dependencies are actually present and cause biased estimators. Moreover, guidelines that are useful in a classical queueing setting may break down when considering realistic Internet traffic models. In particular, condition (7) hardly holds in the presence of long range dependencies. Despite a great deal work on modifications has been carried out, no generally satisfactory choices of k and m are available.

3. Regenerative method

The process $(X_n)_{n\in\mathbb{N}}$ is called (zero-delayed) classically regenerative if an infinite sequence $0 = T_0 < T_1 < \dots$ of *regeneration instants* exists such that the distribution of X_{n+T_k} is the same for each $k \geq 1$ and independent of the pre-history X_n, $n < T_k$, $n \geq 1$. The i.i.d. regeneration cycles are defined as $G_n = (X_k,\, T_{n-1} \leq k < T_n)$, and the cycle periods $\tau_n = T_n - T_{n-1}$ are also i.i.d., $n \geq 1$.

$$T_0 = 0,\ T_{n+1} = \min(k > T_n : X_k = 0),\ n \geq 0. \tag{8}$$

We assume the regenerative process to be positive recurrent, that is $\mathbb{E}\tau < \infty$. (Throughout the paper we suppress an index to denote a generic element.)

To estimate a stationary characteristic $\gamma = \mathbb{E}f(X)$ of the process for a measurable function f, assuming the weak limit $f(X_n) \Rightarrow f(X)$ exists, we define the i.i.d. variables

$$Y_i = \sum_{k=T_{i-1}}^{T_i-1} f(X_k),\ i \geq 1. \tag{9}$$

If $\mathbb{E}|Y| < \infty$ and the cycle period τ is aperiodic, then with probability 1,

$$\gamma_n = \frac{1}{n}\sum_{k=1}^{n} f(X_k) \to \frac{\mathbb{E}Y}{\mathbb{E}\tau} = \mathbb{E}f(X),\ n \to \infty. \tag{10}$$

Moreover, assuming that $\sigma^2 \equiv Var(Y - \tau\gamma) \in (0, \infty)$, a (regenerative) central limit theorem states that the $100(1-\alpha)\%$ confidence interval for γ is

$$\left[\gamma_n - \frac{z_\alpha\sqrt{v_n}}{\sqrt{n}},\, \gamma_n + \frac{z_\alpha\sqrt{v_n}}{\sqrt{n}}\right], \tag{11}$$

where z_α satisfies $P(N(0,1) \leq z_\alpha) = 1 - \alpha/2$, and the empirical variance

$$v_n = \frac{1}{n}\frac{\sum_{i=1}^{n}(Y_i - \gamma_n\tau_i)^2}{\overline{\tau}_n^2} \Rightarrow \sigma^2. \tag{12}$$

(Here $\overline{\tau}_n$ stands for the sample mean cycle period.) If the number of regenerative cycles $k \leq 30$ then the $1 - \alpha/2$ quantile of the Student t-distribution with $k-1$ degrees of freedom is more appropriate to use instead of z_α. A minimal sufficient condition for v_n to be weakly consistent is $\mathbb{E}(Y - \gamma\tau)^2 < \infty$, while under stronger assumptions, $\mathbb{E}Y^2 < \infty$, $\mathbb{E}\tau^2 < \infty$, the estimate is strongly consistent [11].

4. Application to effective bandwidth estimation

With regard to effective bandwidth estimation based on the batch means method, assuming that the batch means Y_i according to (4) are i.i.d. as a random variable Y, we obtain

$$\log \mathbb{E}e^{\theta^* \sum_{i=1}^{n} X_i} = \log \mathbb{E}e^{\theta^* \sum_{i=1}^{k} mY_i} = \log \mathbb{E}e^{\theta^* mkY} = \log \mathbb{E}e^{\theta^* nY} = n \log \mathbb{E}e^{\theta^* Y}, \tag{13}$$

which suggests the estimator

$$\hat{\Lambda}(\theta^*) = \frac{1}{n} n \log \frac{1}{k} \sum_{i=1}^{k} e^{\theta^* Y_i} = \log \frac{1}{k} \sum_{i=1}^{k} e^{\theta^* Y_i}. \tag{14}$$

Alternatively, if we consider the batch sums $\hat{X}_i = mY_i$ and assume that they are i.i.d. as a random variable $\hat{X}$, we obtain

$$\log \mathbb{E} e^{\theta^* \sum_{i=1}^{n} X_i} = \log \mathbb{E} e^{\theta^* \sum_{i=1}^{k} \hat{X}_i} = \log \mathbb{E} e^{\theta^* k \hat{X}} = k \log \mathbb{E} e^{\theta^* \hat{X}}, \tag{15}$$

which suggests the estimator [13]

$$\hat{\Lambda}(\theta^*) = \frac{1}{n} k \log \frac{1}{k} \sum_{i=1}^{k} e^{\theta^* \hat{X}_i} = \frac{1}{m} \log \frac{1}{k} \sum_{i=1}^{k} e^{\theta^* \hat{X}_i} \tag{16}$$

For both versions, an estimator of the effective bandwidth is given by

$$\hat{C}_\Gamma = \frac{\hat{\Lambda}(\theta^*)}{\theta^*}. \tag{17}$$

However, the problems with appropriately choosing the batch size m and the number k of batches as outlined in Section 2 for the general framework carries over to effective bandwidth estimation and the Y_i or $\hat{X}_i$, respectively, are only approximately i.i.d. even with a good choice. Therefore, we supposed in [13] not to group into batches of fixed size but to consider regenerative cycles. Indeed, if the arrival process has regeneration instants T_k, then the variables

$$\hat{X}_k = \sum_{i=T_k}^{T_{k+1}-1} X_i, \; k \geq 1 \tag{18}$$

are really i.i.d., not only approximately as with the batch means method. Grouping in such a way we form an alternative estimator of the LSCGF function:

$$\hat{\Lambda}(\theta^*) = \frac{k}{T_k} \log \frac{1}{k} \sum_{i=1}^{k} e^{\theta^* \hat{X}_i}, \tag{19}$$

where k is the number of regeneration cycles. Assuming $\mathbb{E} e^{\theta^* \hat{X}} < \infty$, we obtain with probability 1

$$\lim_{n \to \infty} \hat{\Lambda}_k(\theta^*) = \frac{1}{\mathbb{E}\tau} \log \mathbb{E} e^{\theta^* \hat{X}}. \tag{20}$$

Preliminary analysis shows that the desired equality $\Lambda(\theta) = \log \mathbb{E} e^{\theta^* \hat{X}} / \mathbb{E}\tau$ seems plausible [12, 13].

5. Summary of simulation results

Due to lack of space, we do not present excessive simulation results but rather briefly describe our simulation setup and summarize our findings. To compare the properties of the two estimators, we consider a two-station tandem network with Poisson input to the first station, a (desired) constant service rate C_2 at the second one, and with the finite buffers b_1, b_2, respectively. In such a system the arrival process to the second station regenerates when an arriving customer sees an empty first station. This allows to construct the regenerative estimator by an evident way. Our goal is to find (by estimation) the required constant rate C_2 which guarantees given loss probability Γ.

The simulation has revealed an advantage of the regenerative estimator of $\Lambda(\theta)$ over the batch mean one (in the terms of variance reduction) when exponential or constant service time at the first station is used [12, 13]. On the other hand, simulation shows that the batch mean estimator, being optimistic, has an advantage when regenerative period has a large variance. Also we consider the state-dependent service rate at the first station: it is C_1^1 until queue size exceeds a threshold L, and becomes $C_1^2 (> C_1^1)$ when buffer exceeds level L.

References

[1] Chang C.-S., Thomas J.A. (1995). Effective Bandwidths on High-Speed Digital Networks. *IEEE Journal on Selected Areas in Communications*, 13(6), 1091–1100.

[2] Gibbens R.J. (1996). Traffic Characterisation and Effective Bandwidths for Broadband Network Traces. In *Stochastic Networks: Theory and Applications*. Ed. by F.P.Kelly, S.Zachary and I. B. Ziedins. Oxford University Press, 169–181.

[3] Kelly F.P. (1996). Notes on Effective Bandwidths. In *Stochastic Networks: Theory and Applications*. Ed. by F.P.Kelly, S.Zachary and I. B. Ziedins. Oxford University Press, 141–168.

[4] Wischik D. (1999). The Output of a Switch, or, Effective Bandwidths for Networks. *Queueing Systems*, 32(4), 383–396.

[5] Asmussen S. (2003). *Applied Probability and Queues*, 2nd, Springer, NY.

[6] Bischak D.P., Kelton W.D. and Pollock S.M. (1993). Weighted Batch Means for Confidence Intervals in Steady-state Simulations. *Management Science*, 39(8), 1002–1019.

[7] Law A.M. and Carson J.S. (1979). A Sequential Procedure for Determining the Length of a Steady-State Simulation. *Operations Research*, 27(5), 1011–1025.

[8] Schmeiser B. (1982). Batch Size Effects in the Analysis of Simulation Output. *Operations Research*, 30(3), 556–568.

[9] Crosby S., Leslie I., Huggard M., Lewis J.T., McGurk B., and Russel R. (1996). Predicting bandwidth requirements of ATM and Ethernet traffic. In: Proc. of IEE UK Teletraffic Symposium, Glasgow, UK.

[10] Ganesh A., O'Connell N. and Wischik D. (2004). *Big Queues*, Springer-Verlag, Berlin.

[11] Glynn P.W. and Iglehart D.L. (1993). Conditions for the applicability of the regenerative method. *Management Science*, 39, 1108–1111.

[12] Morozov E., Dyudenko I., and Pagano M. (2008). Regenerative estimator of the overflow probability in a tandem network. In. *Proc. of the 7th International Workshop on Rare Event Simulation*, Rennes, France, 283–287.

[13] Vorobieva I., Morozov E., Pagano M., and G. Procissi (2008). A New Regenerative Estimator for Effective Bandwidth Prediction. *Proc. of AMICT 2007*, 175–186, Petrozavodsk, Russia.

6th St.Petersburg Workshop on Simulation (2009) 1010-1014

Efficient Estimation for Incomplete Multivariate Data[1]

Bent Jørgensen[2], Hans Chr. Petersen[3]

Abstract

We compare the Fisher scoring and EM algorithms for incomplete multivariate data, and investigate the corresponding estimating functions under second-moment assumptions. We propose a hybrid algorithm, where Fisher scoring is used for the mean vector and the EM algorithm for the covariance matrix. A bias-corrected estimate for the covariance matrix is obtained.

1. Introduction

Incomplete data are a major concern in applied areas such as osteology and paleontology, where it is common to deal with data having a large proportion of missing values. Especially in connection with multivariate data, it is important to deal with incomplete data as efficiently as possible, in terms of both statistical and computational efficiency.

The current standard for estimation in the k-variate normal distribution with incomplete data is the EM algorithm (Dempster, Laird and Rubin, 1977), see e.g. Schafer (1997, Ch. 5), Johnson and Wichern (1998, pp. 268–273) and Little and Rubin (2002, Ch. 11). In the multivariate normal case, this method goes back to Orchard and Woodbury (1972) and Beale and Little (1975), and the use of the linear predictor for the purpose of imputing missing values in multivariate normal data dates back at least as far as Anderson (1957).

It is interesting to note, however, that already Trawinski and Bargmann (1964) and Hartley and Hocking (1971) developed the Fisher scoring algorithm for incomplete multivariate normal data. The scoring algorithm is usually more efficient than the EM algorithm in terms of the number of iterations, and this algorithm also has the advantage that a suitable bias-corrected estimate for the covariance matrix may be obtained. On the other hand the information matrix for the covariance matrix $\mathbf{\Sigma}$ has size $[k(k+1)/2]^2$, which may limit the usefulness of the scoring algorithm for large k. It is hence useful to compare the performance of these two

[1]This work was supported by the Danish Natural Science Research Council

[2]University of Southern Denmark, E-mail: bentj@stat.sdu.dk

[3]University of Southern Denmark, E-mail: hcpetersen@stat.sdu.dk

types of algorithms, and compare with a hybrid algorithm, where Fisher scoring is used for the mean vector and the EM algorithm is used for the covariance matrix.

We phrase the discussion of these issues in terms of estimating functions under second-moment assumptions, and develop a suitable matrix representation for the results.

2. Incomplete data

Consider a k-vector of data $\boldsymbol{Y}$, partitioned into observed data $\boldsymbol{Y}_r$ and missing data $\boldsymbol{Y}_m$ as follows:

$$\begin{bmatrix} \boldsymbol{Y}_r \\ \boldsymbol{Y}_m \end{bmatrix} = \begin{bmatrix} \boldsymbol{R} \\ \boldsymbol{M} \end{bmatrix} \boldsymbol{Y}. \tag{1}$$

Here $\begin{bmatrix} \boldsymbol{R} \\ \boldsymbol{M} \end{bmatrix}$ is an orthogonal permutation matrix of zeroes and ones ($\boldsymbol{R}$ = retain; $\boldsymbol{M}$ = missing). It follows that the inverse of the mapping (1) is

$$\boldsymbol{Y} = \begin{bmatrix} \boldsymbol{R} \\ \boldsymbol{M} \end{bmatrix}^\top \begin{bmatrix} \boldsymbol{Y}_r \\ \boldsymbol{Y}_m \end{bmatrix} = \boldsymbol{R}^\top \boldsymbol{Y}_r + \boldsymbol{M}^\top \boldsymbol{Y}_m. \tag{2}$$

This matrix representation of the missing data structure is very useful both theoretically and practically, and we may think of (1) and (2) as giving the relation between the rectangular ($\boldsymbol{Y}$) and ragged ($\boldsymbol{Y}_r$) representation of the data. The use of matrix algebra in connection with incomplete data is by no means new, see e.g. Trawinski and Bargmann (1964), but it turns out to be important to explore the linearity of (1) and (2) along with the orthogonality of the permutation matrix. In particular, this orthogonality implies the following useful relations

$$\boldsymbol{R}\boldsymbol{R}^\top = \boldsymbol{I}, \quad \boldsymbol{M}\boldsymbol{M}^\top = \boldsymbol{I}, \ \boldsymbol{R}\boldsymbol{M}^\top = \boldsymbol{0},$$

and

$$\boldsymbol{R}^\top \boldsymbol{R} + \boldsymbol{M}^\top \boldsymbol{M} = \boldsymbol{I}. \tag{3}$$

Let us introduce the notation $\boldsymbol{Y} \sim [\boldsymbol{\mu}; \boldsymbol{\Sigma}]$, which means that $\boldsymbol{Y}$ has mean vector $\boldsymbol{\mu}$ and variance matrix $\boldsymbol{\Sigma}$. We consider estimation of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ based on this second-moment assumption. Let $\boldsymbol{\mu}_r$ and $\boldsymbol{\mu}_m$ denote the mean vectors of $\boldsymbol{Y}_r$ and $\boldsymbol{Y}_m$, respectively. Then (1) and (2) imply

$$\begin{bmatrix} \boldsymbol{\mu}_r \\ \boldsymbol{\mu}_m \end{bmatrix} = \begin{bmatrix} \boldsymbol{R} \\ \boldsymbol{M} \end{bmatrix} \boldsymbol{\mu} \ \text{ and } \ \boldsymbol{\mu} = \boldsymbol{R}^\top \boldsymbol{\mu}_r + \boldsymbol{M}^\top \boldsymbol{\mu}_m,$$

respectively. Also, (1) implies that

$$\operatorname{Var} \begin{bmatrix} \boldsymbol{Y}_r \\ \boldsymbol{Y}_m \end{bmatrix} = \begin{bmatrix} \boldsymbol{R}\boldsymbol{\Sigma}\boldsymbol{R}^\top & \boldsymbol{R}\boldsymbol{\Sigma}\boldsymbol{M}^\top \\ \boldsymbol{M}\boldsymbol{\Sigma}\boldsymbol{R}^\top & \boldsymbol{M}\boldsymbol{\Sigma}\boldsymbol{M}^\top \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Sigma}_{rr} & \boldsymbol{\Sigma}_{rm} \\ \boldsymbol{\Sigma}_{mr} & \boldsymbol{\Sigma}_{mm} \end{bmatrix},$$

say. Similarly, calculating the covariance matrix on both sides of (2) yields the following relationship:

$$\boldsymbol{\Sigma} = \boldsymbol{R}^\top \boldsymbol{\Sigma}_{rr} \boldsymbol{R} + \boldsymbol{M}^\top \boldsymbol{\Sigma}_{mm} \boldsymbol{M} + \boldsymbol{R}^\top \boldsymbol{\Sigma}_{rm} \boldsymbol{M} + \boldsymbol{M}^\top \boldsymbol{\Sigma}_{mr} \boldsymbol{R}. \tag{4}$$

Let us now predict $\boldsymbol{Y}_m$ from $\boldsymbol{Y}_r$ using the BLUP (Best Linear Unbiased Predictor), defined by

$$\begin{aligned}\widehat{\boldsymbol{Y}}_m &= \boldsymbol{\mu}_m + \boldsymbol{\Sigma}_{mr}\boldsymbol{\Sigma}_{rr}^{-1}\left(\boldsymbol{Y}_r - \boldsymbol{\mu}_r\right) \\ &= \boldsymbol{M}\boldsymbol{\mu} + \boldsymbol{M}\boldsymbol{\Sigma}\boldsymbol{R}^\top\left(\boldsymbol{R}\boldsymbol{\Sigma}\boldsymbol{R}^\top\right)^{-1}\boldsymbol{R}\left(\boldsymbol{Y} - \boldsymbol{\mu}\right).\end{aligned}$$

We may expand this to a predictor for $\boldsymbol{Y}$. From (2) and (3), we obtain

$$\begin{aligned}\widehat{\boldsymbol{Y}} &= \boldsymbol{R}^\top\boldsymbol{Y}_r + \boldsymbol{M}^\top\widehat{\boldsymbol{Y}}_m \\ &= \boldsymbol{\mu} + \boldsymbol{\Sigma}\boldsymbol{R}^\top\left(\boldsymbol{R}\boldsymbol{\Sigma}\boldsymbol{R}^\top\right)^{-1}\boldsymbol{R}\left(\boldsymbol{Y} - \boldsymbol{\mu}\right). \qquad (5)\end{aligned}$$

3. The score function for $\boldsymbol{\mu}$

Now, let us consider i.i.d. data $\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_n$ from $[\boldsymbol{\mu}; \boldsymbol{\Sigma}]$. Under multivariate normality, the score function for $\boldsymbol{\mu}$ using the complete data is

$$\boldsymbol{U} = \sum_{i=1}^{n} \boldsymbol{\Sigma}^{-1}\left(\boldsymbol{Y}_i - \boldsymbol{\mu}\right).$$

Let the incomplete data be given by $\boldsymbol{Y}_{ri} = \boldsymbol{R}_i\boldsymbol{Y}_i$ for $i = 1, \ldots, n$, say, where subscript i refers to the ith data case. Now $\boldsymbol{Y}_{ri}$ follows the linear model

$$\boldsymbol{Y}_{ri} \sim \left[\boldsymbol{R}_i\boldsymbol{\mu}; \boldsymbol{R}\boldsymbol{\Sigma}\boldsymbol{R}^\top\right].$$

The corresponding quasi score function for $\boldsymbol{\mu}$ is the BLUP of $\boldsymbol{U}$ given the observed data,

$$\begin{aligned}\widehat{\boldsymbol{U}} &= \sum_{i=1}^{n} \boldsymbol{\Sigma}^{-1}\left(\widehat{\boldsymbol{Y}}_i - \boldsymbol{\mu}\right) \\ &= \sum_{i=1}^{n} \boldsymbol{R}_i^\top\left(\boldsymbol{R}_i\boldsymbol{\Sigma}\boldsymbol{R}_i^\top\right)^{-1}\left(\boldsymbol{Y}_{ri} - \boldsymbol{R}_i\boldsymbol{\mu}\right). \qquad (6)\end{aligned}$$

Under multivariate normality, this is the score function for $\boldsymbol{\mu}$. The estimating equation $\widehat{\boldsymbol{U}} = \boldsymbol{0}$ has an explicit solution, namely

$$\hat{\boldsymbol{\mu}} = \mathcal{I}^{-1}\sum_{i=1}^{n} \boldsymbol{R}_i^\top\left(\boldsymbol{R}_i\boldsymbol{\Sigma}\boldsymbol{R}_i^\top\right)^{-1}\boldsymbol{Y}_{ri}, \qquad (7)$$

where

$$\mathcal{I} = \sum_{i=1}^{n} \boldsymbol{R}_i^\top\left(\boldsymbol{R}_i\boldsymbol{\Sigma}\boldsymbol{R}_i^\top\right)^{-1}\boldsymbol{R}_i \qquad (8)$$

is the information matrix, so that in fact $\hat{\boldsymbol{\mu}} \sim \left[\boldsymbol{\mu}; \mathcal{I}^{-1}\right]$.

We note that the estimating function $\widehat{\boldsymbol{U}}$ is $\boldsymbol{\Sigma}$-insensitive, in the sense of Jørgensen and Knudsen (2004), meaning that $\mathrm{E}\left(\partial\widehat{\boldsymbol{U}}/\partial\boldsymbol{\Sigma}\right) = \boldsymbol{0}$. This follows from the fact that $\mathrm{E}\left(\partial\widehat{\boldsymbol{Y}}_i/\partial\boldsymbol{\Sigma}\right) = \boldsymbol{0}$, which in turn is a consequence of (5). Insensitivity implies that the asymptotic variance of $\hat{\boldsymbol{\mu}}$ remains the same whether or not $\boldsymbol{\Sigma}$ is known, and independently of which estimate for $\boldsymbol{\Sigma}$ is used. In any case, the asymptotic variance for $\hat{\boldsymbol{\mu}}$ is hence $\mathcal{I}^{-1}$, the inverse of (8).

4. The score function for $\boldsymbol{\Sigma}$

Let $\sigma_{\ell m}$ denote the ℓmth element of $\boldsymbol{\Sigma}$ for $1 \leq \ell \leq m \leq k$. The corresponding elements of the score function for $\boldsymbol{\Sigma}$ under multivariate normality are given by

$$\psi_{\ell m} = \sum_{i=1}^{n} \mathrm{tr}\left\{\boldsymbol{W}_{i\ell m}\left(\boldsymbol{r}_i\boldsymbol{r}_i^\top - \boldsymbol{\Sigma}_{rri}\right)\right\},$$

where

$$\boldsymbol{r}_i = \boldsymbol{Y}_{ri} - \boldsymbol{R}_i\boldsymbol{\mu}, \quad \boldsymbol{\Sigma}_{rri} = \boldsymbol{R}_i\boldsymbol{\Sigma}\boldsymbol{R}_i^\top, \quad \boldsymbol{W}_{i\ell m} = \boldsymbol{\Sigma}_{rri}^{-1}\frac{\partial\boldsymbol{\Sigma}_{rri}}{\partial\sigma_{\ell m}}\boldsymbol{\Sigma}_{rri}^{-1},$$

and

$$\frac{\partial\boldsymbol{\Sigma}_{rri}}{\partial\sigma_{\ell m}} = \boldsymbol{R}_i\frac{\partial\boldsymbol{\Sigma}}{\partial\sigma_{\ell m}}\boldsymbol{R}_i^\top.$$

The matrices $\partial\boldsymbol{\Sigma}_{rri}/\partial\sigma_{\ell m}$ are null in case $\boldsymbol{Y}_{ri}$ does not contain information about $\sigma_{\ell m}$, and otherwise contain a diagonal 1 when $\ell = m$ or two symmetrically placed ones when $\ell < m$, the remaining entries being zero.

The sensitivity matrix $\boldsymbol{S}$ for $\boldsymbol{\Sigma}$, defined as the expected derivative of the estimating function $\boldsymbol{\psi}$ with respect to the entries $\sigma_{\ell m}$, has elements given by

$$\boldsymbol{S}_{(\ell m)(\ell' m')} = -\sum_{i=1}^{n} \mathrm{tr}\left(\boldsymbol{\Sigma}_{rri}^{-1}\frac{\partial\boldsymbol{\Sigma}_{rri}}{\partial\sigma_{\ell m}}\boldsymbol{\Sigma}_{rri}^{-1}\frac{\partial\boldsymbol{\Sigma}_{rri}}{\partial\sigma_{\ell' m'}}\right),$$

and dimension K^2, where $K = k(k+1)/2$.

The Newton scoring algorithm for $\boldsymbol{\Sigma}$, in the sense of Jørgensen and Knudsen (2004), is hence given by means of the update

$$\boldsymbol{\Sigma}^* = \boldsymbol{\Sigma} - \boldsymbol{S}^{-1}\,\boldsymbol{\psi}, \tag{9}$$

where $\boldsymbol{S}^{-1}\boldsymbol{\psi}$ is understood as a k^2 symmetric matrix with lower triangle defined by symmetry, and $\boldsymbol{S}$ and $\boldsymbol{\psi}$ are evaluated at the value $\boldsymbol{\Sigma}$. When (9) is combined with (5), we obtain the chaser algorithm of Jørgensen and Knudsen (2004), which, due to the $\boldsymbol{\Sigma}$-insensitivity of $\widehat{\boldsymbol{U}}$, is equivalent to the full Newton scoring algorithm for the parameter $(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ or, under multivariate normality, the Fisher scoring algorithm.

5. The EM algorithm

Let us summarize the EM algorithm, with a slight modification of Johnson and Wichern (1998, pp. 268–273). Let the current values of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ be given. The $\boldsymbol{\mu}$ part of the algorithm is obtained by first calculating $\widehat{\boldsymbol{Y}}_i$ using (5), and then calculating the update $\boldsymbol{\mu}^*$ as follows:

$$\boldsymbol{\mu}^* = \frac{1}{n}\sum_{i=1}^{n}\widehat{\boldsymbol{Y}}_i = \boldsymbol{\mu} + \frac{1}{n}\sum_{i=1}^{n}\boldsymbol{\Sigma}\boldsymbol{R}_i^\top\left(\boldsymbol{R}_i\boldsymbol{\Sigma}\boldsymbol{R}_i^\top\right)^{-1}\left(\boldsymbol{Y}_{ri} - \boldsymbol{R}_i\boldsymbol{\mu}\right). \tag{10}$$

Note that, upon convergence, $\boldsymbol{\mu}^* = \boldsymbol{\mu}$, in which case $\widehat{\boldsymbol{U}}$ in (6) is zero.

The $\boldsymbol{\Sigma}$ part of the algorithm involves the conditional variance given $\boldsymbol{Y}_{ri}$,

$$\mathrm{Var}\left(\boldsymbol{Y}_i|\boldsymbol{Y}_{ri}\right) = \boldsymbol{M}_i^\top\left(\boldsymbol{\Sigma}_{mmi} - \boldsymbol{\Sigma}_{mri}\boldsymbol{\Sigma}_{rri}^{-1}\boldsymbol{\Sigma}_{rmi}\right)\boldsymbol{M}_i.$$

We then update $\boldsymbol{\Sigma}$ as follows:

$$\boldsymbol{\Sigma}^* = \frac{1}{n}\sum_{i=1}^{n}\left[\left(\widehat{\boldsymbol{Y}}_i - \boldsymbol{\mu}\right)\left(\widehat{\boldsymbol{Y}}_i - \boldsymbol{\mu}\right)^\top + \mathrm{Var}\left(\boldsymbol{Y}_i|\boldsymbol{Y}_{ri}\right)\right]. \tag{11}$$

Since (10) and (11) involve only first and second moments, they may be used as estimating equations for $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ under the second-moment assumption.

6. Comparing algorithms

Using both real and simulated data, we compare the following three algorithms for estimating $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ from incomplete data under the second-moment assumption:

- The full EM algorithm given by (10) and (11).
- The full scoring algorithm given by (7) combined with (9). A good starting value for this algorithm is based on taking $\boldsymbol{\Sigma} = \boldsymbol{I}$, which makes $\hat{\boldsymbol{\mu}}$ in the first step equal to the available cases average.
- A hybrid algorithm combining (7) and (11), with the same starting values as for the full scoring algorithm.

The hybrid algorithm has a more efficient $\boldsymbol{\mu}$ part than the EM algorithm, but requires the additional inversion of the $k \times k$ matrix $\mathcal{I}$ in each step, which in turn yields the asymptotic variance of the estimate $\hat{\boldsymbol{\mu}}$ as a by-product. The $\boldsymbol{\Sigma}$ part of the hybrid algorithm, given by (11), is likely to be the bottleneck, and it will be interesting to see how much the hybrid algorithm can improve on the full EM algorithm.

An advantage of the scoring algorithm is that a bias-corrected estimate for the covariance matrix may be obtained, using the method developed by Holst and Jørgensen (2008). This is particularly important for incomplete data, where the effective degrees of freedom can be small for variables with a large proportion of missing values. This is illustrated by simulation.

References

[1] Anderson T. (1957) Maximum likelihood estimates for a multivariate normal distribution when some observations are missing. *J. Amer. Statist. Assoc.,* **52**, 200–203.

[2] Beale E.M.L., Little R.J.A. (1975) Missing values in multivariate analysis. *J. Roy. Statist. Soc. B*, **37**, 129–145.

[3] Dempster A.P., Laird N.M., Rubin D.B. (1977) Maximum likelihood from incomplete data via the EM algorithm (with discussion). *J. Roy. Statist. Soc. B*, **39**, 1–38.

[4] Hartley H.O., Hocking R.R. (1971) The analysis of incomplete data. *Biometrics*, **27**, 783–823.

[5] Holst R., Jørgensen B. (2008). Efficient and robust estimation for generalized linear longitudinal mixed models. Unpublished manuscript.

[6] Johnson R.A., Wichern D. (1998) *Applied Multivariate Statistical Analysis.* Prentice Hall, Englewood Cliffs, New Jersey.

[7] Jørgensen B., Knudsen S.J. (2004) Parameter orthogonality and bias adjustment for estimating functions. *Scand. J. Statist.,* **31**, 93–114.

[8] Little R.J.A., Rubin D. (2002) *Statistical Analysis with Missing Data.* Wiley Interscience, New York.

[9] Orchard T., Woodbury M.A. (1972) A missing information principle: theory and applications. *Proceedings of the Sixth Berkeley Symposium*, Vol. 1, 697–715.

[10] Schafer J.L. (1997) *Analysis of Incomplete Multivariate Data.* Chapman & Hall/CRC Press, London.

[11] Trawinski I.M., Bargmann R.E. (1964) Maximum likelihood estimation with incomplete multivariate data. *Ann. Math. Statist.,* **35**, 647–657.

6th St.Petersburg Workshop on Simulation (2009) 1016-1020

Maximal likelihood estimates for modified gravitation model by aggregated data[1]

Alexander Andronov[2]

Abstract

A nonlinear regression model for a forecasting of passenger flow between various points (towns) is described. Unknown parameters are estimated using aggregated data when only the information on the departed from each town passengers quantity is available.

1. Introduction

We have n corresponding points (towns) with numbers $i = 1, 2, ..., n$. For the point i, one is known inhabitants (citizen) number h_i and m numerical characteristics (categorical data) $c_{i,j}, j = 1, 2, ..., m$, those are known constants. For all pairs of the points (i, l) the distance $d_{i,l}$ between them is kown as well. In addition, we known the quantity of the departed passengers Y_i from the point i during considered time interval, that is a random variable.
Our aim is to estimate correspondence value $Y_{i,l}$ for all pairs of points (i, l), precisely the quantity of the departed passengers from the point i to the point l. The matrix of $Y_{i,l}$ is to be said *the correspondence matrix.* Let us denote an estimate of $Y_{i,l}$ by $Y^*_{i,l}$. It requests that all $Y^*_{i,l}$ are positive $Y^*_{i,l} > 0$ for $i \neq l$, $Y^*_{i,i} = 0$ and $Y^*_{i,l} = Y_{l,i}$. As a model for the concrete correspondence (i, l) for $i \neq l$ we use

$$Y_{i,l} = \frac{(h_i h_l)^v}{(d_{i,l})^\tau} exp\big(a + (c_{(i)} + c_{(l)})\alpha + g_{(i,l)}\beta + V_{i,l}\big), \tag{1}$$

where a, $\alpha = (\alpha_1\ \alpha_2\ ...\ \alpha_m)^T$ and $\beta = (\beta_1\ \beta_2\ ...\ \beta_m)^T$ are unknown regression parameters, τ and v are unknown form parameters,
$c_{(i)} = (c_{i,1}\ ...\ c_{i,m})$ and $g_{(i,l)} = (c_{i,1}c_{l,1}\ ...\ c_{i,m}c_{l,m})$ be m-vector-rows,
$\{V_{i,l}\}$ are independent identically distributed random variables with zero mean and unknown variance σ^2.
In addition we set $Y_{i,i} = 0$. Note that a case $v = 1$ and $\tau = 2$ corresponds to the so called gravitation model.

[1]This work was supported by grant 7326 of Latvian Ministry of Science and Education.
[2]Riga Technical University, E-mail: Aleksandrs.Andronovs@rtu.lv

As a corollary of this model we get the following presentation for the quantity of the departed passengers from the point i:

$$Y_i = \sum_{\substack{l=1 \\ i \neq l}}^{n} Y_{i,l} = \sum_{\substack{l=i \\ i \neq l}}^{n} \frac{(h_i h_l)^v}{(d_{i,l})^\tau} exp\big(a + (c_{(i)} + c_{(l)})\alpha + g_{(i,l)}\beta + V_{i,l}\big). \tag{2}$$

Now, we must estimate unknown parameters on the basis of fixed values $\{Y_i\}$. Such a problem was considered earlier in literature. Besides, usually the entropy approach is used for that. But there are many received estimates $Y^*_{i,l}$ equal to zero that is inaccessible. We use maximal likelihood method [2, 3]. For that we need to investigate the distribution and the expectation of $Y_{i,l}$.

2. Distribution analysis

If $V_{i,l}$ has normal distribution, then $Z_{i,l} = exp(V_{i,l})$ has the log-normal distribution [1] with characteristics

$$E(Z_{i,l}) = E(exp(V_{i,l})) = exp\big(\frac{1}{2}\sigma^2\big),$$

$$D(Z_{i,l}) = D(exp(V_{i,l})) = exp(\sigma^2)(exp(\sigma^2) - 1).$$

Therefore, for $i \neq l$

$$E(Y_{i,l}) = \frac{(h_i h_l)^v}{(d_{i,l})^\tau} exp(a + (c_{(i)} + c_{(l)})\alpha + g_{(i,l)}\beta) exp\big(\frac{1}{2}\sigma^2\big), \tag{3}$$

$$D(Y_{i,l}) = \frac{(h_i h_l)^{2v}}{(d_{i,l})^{2\tau}} exp(2(a + (c_{(i)} + c_{(l)})\alpha + g_{(i,l)}\beta)) exp(\sigma^2)(exp(\sigma^2) - 1) =$$

$$= (exp(\sigma^2) - 1)(E(Y_{i,l}))^2. \tag{4}$$

Analogous formulae have place for $\{Y_i\}$:

$$E(Y_i) = \sum_{l=1}^{n} E(Y_{(i,l)}) = exp\big(\frac{1}{2}\sigma^2\big) \sum_{\substack{l=1 \\ i \neq l}}^{n} \frac{(h_i h_l)^v}{(d_{i,l})^\tau} exp(a + (c_{(i)} + c_{(l)})\alpha + g_{(i,l)}\beta), \tag{5}$$

$$D(Y_i) = (exp(\sigma^2) - 1) \sum_{\substack{l=1 \\ i \neq l}}^{n} (E(Y_{i,l}))^2. \tag{6}$$

Further we need to use covariance $W_{i,l} = Cov(Y_i, Y_l)$ between Y_i and Y_l. We have for $i \neq l$

$$W_{i,l} = Cov(Y_i, Y_l) = D(Y_{i,l}) = (exp(\sigma^2) - 1)(E(Y_{i,l}))^2. \tag{7}$$

3. A normal approximation

With respect to the central limit theorem we suppose that Y_i has the normal distribution. Then the log-likelihood function for the sample $Y = (Y_1, Y_2, ..., Y_n)$ is as follows (without constants):

$$l(a, \alpha, \beta, \sigma, v, \tau) = -\frac{1}{2} ln(|W|) - \frac{1}{2}(Y - E(Y))^T W^{-1}(Y - E(Y)), \tag{8}$$

where the values of $E(Y) = (E(Y_1), E(Y_2), ..., E(Y_n))^T$ and the matrix W are calculated with respect to (5), (6) and (7).
So, our model contains the unknown parameters $a, \{\alpha_j\}, \{\beta_j\}, \sigma^2, v$ and τ. We get the corresponding estimates minimizing the log-likelihood function (8). For that we use a gradient method. Let $\theta = (a \;\; \alpha \;\; \beta \;\; \sigma^2 \;\; v \;\; \tau)^T$,

$$\nabla l(\theta) = \nabla l(a, \alpha, \beta, \sigma^2, v, \tau) = \big(\frac{\partial}{\partial a} l \;\; \frac{\partial}{\partial \alpha} l \;\; \frac{\partial}{\partial \beta} l \;\; \frac{\partial}{\partial \sigma^2} l \;\; \frac{\partial}{\partial v} l \;\; \frac{\partial}{\partial \tau} l\big)^T \tag{9}$$

be the score vector for (8).
To avoid cumbersome expressions, we write down the score vector by parts.
a) *The gradient of* $E(Y_{i,l})$ is the $2(m+2)$-vector: for $i \neq l$

$$\nabla E(Y_{i,l}) = \frac{\partial}{\partial \theta} E(Y_{i,l}) = E(Y_{i,l})(1 \;\; (c_{(i)} + c_{(l)}) \;\; g_{(i,l)} \;\; \frac{1}{2} \;\; ln(h_i h_l) \;\; - ln(d_{(i,l)}))^T. \tag{10}$$

b) *The gradient of* $E(Y_i)$:

$$\frac{\partial}{\partial \theta} E(Y_i) = \sum_{l=1}^{n} \nabla E(Y_{i,l}). \tag{11}$$

c) *The derivative of* $E(Y) = (E(Y_1) \; E(Y_2) \; ... \; E(Y_n))$ is the $2(m+2) \times n$-matrix :

$$\frac{\partial}{\partial \theta} E(Y) = \big(\frac{\partial}{\partial \theta} E(Y_1) \; \frac{\partial}{\partial \theta} E(Y_2) \; ... \; \frac{\partial}{\partial \theta} E(Y_n)\big). \tag{12}$$

d) *The gradient of* $W_{i,l} = Cov(Y_i, Y_l)$ for $i \neq l$:

$$\frac{\partial}{\partial \theta} W_{i,l} = \frac{\partial}{\partial \theta} Cov(Y_i, Y_l) = \frac{\partial}{\partial \theta} D(Y_{i,l}) =$$

$$= 2(exp(\sigma^2) - 1)(E(Y_{i,l}))\nabla E(Y_{i,l}) + exp(\sigma^2)(E(Y_{i,l}))^2 e_{2m+1}, \tag{13}$$

where $e_i = (0 \; ... \; 0 \; 1 \; 0 \; ... \; 0)^T$ is the i-th column of the $n \times n$ identity matrix.
e) *The gradient of* $W_{i,i} = D(Y_i)$:

$$\frac{\partial}{\partial \theta} W_{i,i} = \frac{\partial}{\partial \theta} D(Y_i) = \sum_{\substack{l=i \\ i \neq l}}^{n} \frac{\partial}{\partial \theta} D(Y_{i,l}) = \sum_{\substack{l=i \\ i \neq l}}^{n} \frac{\partial}{\partial \theta} W_{i,l}. \tag{14}$$

f) *The gradient of* $ln|(W)|$. The chain rule gives us (see Tables 4.1 and 4.7 in [4]):

$$\nabla ln|(W)| = \frac{\partial |W|}{\partial \theta} = \frac{\partial vecW}{\partial \theta} \frac{\partial ln|W|}{\partial vecW} = \frac{\partial vecW}{\partial \theta} vec(W^{-1}), \tag{15}$$

where $\frac{\partial vecW}{\partial \theta}$ is the $2(m+2)\times n^2$-matrix which columns are determined by formulas (13) and (14).

g) *The derivatives of* $vec(W^{-1})$. The chain rule gives us the $2(m+2)\times n^2$-matrix

$$\nabla vec(W^{-1}) = \frac{\partial vec(W^{-1})}{\partial \theta} = \frac{\partial vecW}{\partial \theta}\frac{\partial vec(W^{-1})}{\partial vecW} = -\frac{\partial vecW}{\partial \theta}(W^{-1}\otimes W^{-1}). \tag{16}$$

h) *The gradient of* $a^TW^{-1}b$. If a and b are vectors of constants, then (see Table 4.3 in [1]):

$$\nabla(a^TW^{-1}b) = \frac{\partial}{\partial\theta}a^TW^{-1}b = \frac{\partial vecW}{\partial\theta}\frac{\partial}{\partial vecW}a^TW^{-1}b = -\big(\frac{\partial}{\partial\theta}vecW\big)(W^{-1}b\otimes W^{-1}a).$$

In partial,

$$\frac{\partial}{\partial vecW}((Y-E(Y))^TW^{-1}(Y-E(Y))) = -(W^{-1}(Y-E(Y))\otimes W^{-1}(Y-E(Y))). \tag{17}$$

Assembling obtained results together we get the final expression for *the score vector* (9) *of log-likelihood function* (8):

$$\nabla l(\theta) = -\nabla\frac{1}{2}(ln(|W|) + (Y-E(Y)^TW^{-1}(Y-E(Y)) = -\frac{1}{2}\frac{\partial vecW}{\partial\theta}vec(W^{-1}) +$$

$$+\big(\frac{\partial}{\partial\theta}E(Y)\big)W^{-1}(Y-E(Y) + \frac{1}{2}\big(\frac{\partial vecW}{\partial\theta}\big)(W^{-1}(Y-E(Y))\otimes W^{-1}(Y-E(Y)))\big). \tag{18}$$

4. Information matrix

We define the information matrix as

$$I(\theta) = -\frac{1}{n}E\big(\frac{\partial^2}{\partial\theta\partial\theta^T}l(\theta)\big) = -\frac{1}{n}E\big(\frac{\partial}{\partial\theta}\nabla l(\theta)\big). \tag{19}$$

With respect to (18) we have

$$\frac{\partial}{\partial\theta}\nabla l(\theta) = -\frac{\partial}{\partial\theta}\big(\frac{1}{2}\frac{\partial vecW}{\partial\theta}vec(W^{-1})\big) + \frac{\partial}{\partial\theta}\big(\big(\frac{\partial}{\partial\theta}E(Y)\big)W^{-1}(Y-E(Y))\big) +$$

$$+\frac{1}{2}\frac{\partial}{\partial\theta}\big(\big(\frac{\partial vecW}{\partial\theta}\big)(W^{-1}(Y-E(Y))\otimes W^{-1}(Y-E(Y)))\big).$$

Let us consider some necessary expressions.

a) *The Hessian matrix of* $W_{i,l} = Cov(Y_i, Y_l)$ for $i \neq l$:

$$\frac{\partial}{\partial\theta}\big(\frac{\partial}{\partial\theta}W_{i,l}\big) = exp(\sigma^2)(E(Y_{i,l}))^2e_{2m+1}e_{2m+1}^T + 2exp(\sigma^2)E(Y_{i,l})\nabla E(Y_i,l)e_{2m+1}^T +$$

$$+2exp(\sigma^2)E(Y_{i,l})e_{2m+1}\nabla E(Y_{i,l})^T + 4(exp(\sigma^2)-1)\nabla E(Y_{i,l})(\nabla E(Y_{i,l}))^T.$$

b) *The product rule* (see Lemma 4.3 in [3]) gives:

$$\frac{\partial}{\partial\theta}\big(\frac{\partial vecW}{\partial\theta}vec(W^{-1})\big) = \frac{\partial}{\partial\theta}\big(\frac{\partial vecW}{\partial\theta}vec(W^{-1})\big)\big|_{\substack{vec(W^{-1})\\ constant}} +$$

$$+\frac{\partial}{\partial\theta}\big(\frac{\partial vecW}{\partial\theta}vec(W^{-1})\big)\big|_{\substack{\partial vecW/\partial\theta\\ constant}} = \big(\frac{\partial}{\partial\theta}vec\big(\frac{\partial vecW}{\partial\theta}\big)\big)(vec(W^{-1})\otimes I_{2(m+2)}) +$$

$$+\frac{\partial vec(W^{-1})}{\partial\theta}\big(\frac{\partial vecW}{\partial\theta}\big)^T,$$

where the derivatives are calculated by (14), (16) and the previous formula. Now we are able to write down

$$I(\theta) = -\frac{1}{n}E\big(\frac{\partial}{\partial\theta}\nabla l(\theta)\big) = \frac{1}{2n}\big(\frac{\partial}{\partial\theta}vec\big(\frac{\partial vecW}{\partial\theta}\big)\big)(vec(W^{-1})\otimes I_{2(m+2)}) +$$

$$+\frac{1}{2n}\frac{\partial vec(W^{-1})}{\partial\theta}\big(\frac{\partial vecW}{\partial\theta}\big)^T + \frac{1}{n}\big(\frac{\partial}{\partial\theta}E(Y)\big)W^{-1}\big(\frac{\partial}{\partial\theta}E(Y)\big)^T. \qquad (20)$$

Note that we are able to use alternative expression here as well:

$$\frac{\partial}{\partial\theta}\big(\frac{\partial vecW}{\partial\theta}vec(W^{-1})\big)\big|_{\substack{vec(W^{-1})\\ constant}} = \sum_{i=1}^{n}\sum_{l\neq i}(W^{-1})_{i,l}\frac{\partial^2}{\partial\theta^2}W_{i,l}.$$

Now we able to estimate the asymptotical covariance matrix of θ^*.

5. Example

Our example concerns seven ($n = 9$) largest towns of Latvia: 1.Riga, 2.Daugavpils, 3.Jekabpils 4.Jelgava, 5.Jurmala, 6.Liepaja, 7.Rezekne, 8.Valmiera, 9.Ventspils. The population sizes (in ten thousand people) are represented by vector h:

$$h = (h_1\ \ h_2\ \ ...\ \ h_n)^T = (76.6\ \ 11.6\ \ 0.32\ \ 6.4\ \ 5.6\ \ 9.0\ \ 3.9\ \ 0.28\ \ 4.4)^T.$$

As numerical characteristics of the i-th town, the two of them have been chosen: $c_{i,1}$ - significance as a rail junction and $c_{i,2}$ - significance as a seaport. The corresponding numerical values are presented by 7×2 matrix:

$$c = \begin{pmatrix} 1 & 2 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 5 & 0 & 0 & 2 \end{pmatrix}^T.$$

The distances between towns ase used as well.
Factual departed passenger quantity (in billion passengers) in some year is the following:

$$Y = (Y_1\ \ Y_2\ \ ...\ \ Y_n)^T = (15.0\ \ 2.46\ \ 0.09\ \ 0.16\ \ 8.40\ \ 4.87\ \ 0.37\ \ 0.08\ \ 1.45)^T.$$

The above described estimation procedure gives the following estimates:

$$\begin{pmatrix} a^* \\ \alpha_1^* \\ \alpha_2^* \end{pmatrix} = \begin{pmatrix} -4.389 \\ 0.623 \\ 0.352 \end{pmatrix}, \begin{pmatrix} \beta_1^* \\ \beta_2^* \end{pmatrix} = \begin{pmatrix} 0.322 \\ 0.21 \end{pmatrix}, \sigma^{2*} = 0.127, v^* = 1.114, \tau^* = 1.79.$$

Now we can calculate maximal likelihood estimates of the arbitrary correspondence (1).

The considered approach is used successfully within the Scientific Project "Creation of mathematical models, algorithms and computer programs for Latvia's transport system's analysis, development prognosis and optimization".

References

[1] Sleeper A. (2007) *Six Sigma Distribution Modeling.* McGraw-Hill, New York.

[2] Srivastava M.S. (2002) *Methods of Multivariate Statistics.* John Wiley & Sons, USA.

[3] Turkington D.A. (2002) *Matrix Calculus & Zero-One Matrices. Statistical and Econometric Applications.* Cambridge University Press, Cambridge.

6th St.Petersburg Workshop on Simulation (2009) 1022-1026

On estimation of the special form nonlinear regression parameters[1]

S.M. Ermakov[2], M.L. Nechaeva[3]

Abstract

The subject of the present research is a numerical experiment associated with estimation of nonlinear regression parameters, which has appearance of a product of exponents with linear combinations of harmonic functions. Parameters are estimated by the least square method. The report provides an analysis of results obtained by different numerical methods.

1. Problem setting

Consider the problem of approximation of data with random error with function of the following form:

$$\eta(x|\theta) = \eta(x|\lambda, \omega, \alpha, \beta) = \sum_{i=1}^{p} e^{\lambda_i x}\big(\alpha_i \cos(\omega_i x) + \beta_i \sin(\omega_i x)\big), \tag{1}$$

where $\theta = \{\lambda, \omega, \alpha, \beta\}$, $\lambda = \{\lambda_i\}_{i=1}^p$, $\omega = \{\omega_i\}_{i=1}^p$, $\alpha_i = \{\alpha_i\}_{i=1}^p$, $\beta = \{\beta_i\}_{i=1}^p$ —the sets of parameters, or which is equivalent, with functions

$$\tilde{\eta}(x|\lambda_i^{'}, \omega_i^{'}, \alpha_i^{'}, \beta_i^{'}, i = 1, \ldots, p) = \sum_{i=1}^{p} \alpha_i^{'} e^{\lambda_i^{'} x} \cos(\omega_i^{'} x - \beta_i^{'}).$$

Functions $y = A\sin(Twx + \varphi)$ often appear in mechanics, when we consider systems connected with various oscillations and a signal transmission. The quantity of the radioactive isotopes y, remaining after the x seconds of radioactive decay, is defined by equation $y = e^{-\lambda x}$, which is a particular case of our model. It is not uncommon to find many similar examples in different domains of science.

In this report a numerical experiment is considered related to parameter estimation in classical scheme of regression. In case of absence of a systematic error,

[1]This work was supported by grant 08-01-00194.

[2]Mathematics and Mechanics Department, St.Petersburg State University, E-mail: Ermakov@math.spbu.ru

[3]Mathematics and Mechanics Department, St.Petersburg State University, E-mail: Maria.Nechaeva@statmod.ru

the problem can be formulated in the following way. Let in $N+1$ points of the set X $x_0 \ldots, x_N$ the values of random variables $y_j = y(x_j)$ are observed, which we can represent as:

$$y_j = \eta(x_j|\theta) + \varepsilon_j, \quad j = 0, \ldots, N, \quad x_j \in X, \tag{2}$$

where $\eta(x|\theta)$ — the function of regression, it has appearance (1) — the function on $X \times \Theta$, given with accuracy to unknown parameters θ from some parametric set $\Theta \subseteqq \mathbb{R}^{4p}$, these parameters require estimation; ε_j, $j = 0, \ldots, N$, — random error of measurements.

Let us take the interval $[0, 1]$ as the set X and consider equidistant points in this interval; let's restrict to mentioning just a case of one component for function (1), i.e. $\eta(x|\theta) = \eta(x|\lambda, \omega, \alpha, \beta) = e^{\lambda x}\big(\alpha\cos(\omega x) + \beta\sin(\omega x)\big)$; let's assume that errors of measurements ε_j, $j = 0, \ldots, N$ are aligned, not correlated, have equal final second moments ($\sigma^2 < +\infty$) and also are distributed under the normal law, i.e. $\varepsilon \sim N(0, \sigma^2 I)$, where $\varepsilon = (\varepsilon_0, \ldots, \varepsilon_N)^T$ — vector of errors.

Parameters are estimated by least square method. Thus the required estimation is:

$$\hat{\theta} = \arg\min_{\theta\in\Theta} F(\theta), \quad \hat{\theta} = \left\{\hat{\lambda},\ \hat{\omega},\ \hat{\alpha},\ \hat{\beta}\right\}, \tag{3}$$

where

$$F(\theta) = \sum_{j=0}^{N} \big[y_j - \eta(x_j|\theta)\big]^2.$$

As the dependence on parameters in regression (1) is nonlinear, the algorithm of least square method appears difficult. Though theoretically inthe general case the problem was investigated in details (for example [2]), numerical experiment reveals a number of important features.

2. Numerical experiments

A number of numerical experiments on finding estimations (3) has been organised. Also, the dependence of the error in the received estimations of parameters on the error in measurements (i.e. from the value of random error ε_j) was constructed.

The initial data was computed by modelling. Values (2) for concrete set of parameters $\tilde{\theta}$ (a true set) were calculated:

$$y_j = \eta\big(x_j|\tilde{\theta}\big) + \varepsilon_j \quad j = 0, \ldots, N. \tag{4}$$

Our goal is to observe how various numerical methods work with reference to formulated task.

For testing some methods of minimization of function were chosen. Among them gradient methods were considered: the quasi-Newton's method, the simplex method of Nelder-Mead, the Newton-Gauss's method and the method of Levenberg-Marquardt.

The sets of parameters $\tilde{\theta}$ were chosen as follows: at first a few random sets were taken, then one parameter $\tilde{\alpha} = 1$ was fixed, values for $\tilde{\lambda}$, $\tilde{\omega}$, $\tilde{\beta}$ were randomly selected from the cube $[0.5, 5]^3$.

After the analysis of results it was noticed, that for some values of parameters all the methods listed above are not quite satisfactory. The considered sets of parameters $\tilde{\theta}$ we have conditionally split into "good" and "bad" sets of parameters.

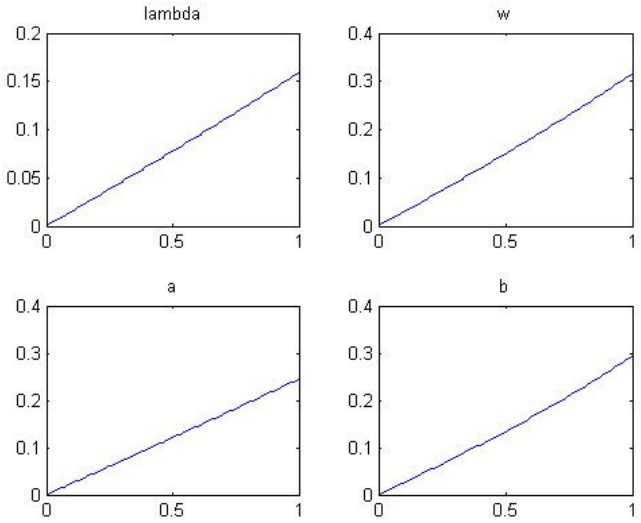


Figure 1: Dependence between $\hat{\sigma}(\theta)$ and $\sigma(\varepsilon)$ for parameters $\{\tilde{\lambda} = 3;\ \tilde{\omega} = 3;\ \tilde{\alpha} = 1;\ \tilde{\beta} = 1\}$ with step $h = 0.01$ ($lambda = \tilde{\lambda}$; $w = \tilde{\omega}$; $a = \tilde{\alpha}$; $b = \tilde{\beta}$).

Fig. 1 and 2 represent the dependence of estimates of the standard deviation of errors in parameters estimation $\big($let's denote them: $\hat{\sigma}(\theta) = \{\hat{\sigma}(\lambda),\ \hat{\sigma}(\omega),\ \hat{\sigma}(\alpha),\ \hat{\sigma}(\beta)\}\big)$ on standard deviation for a random error in measurements, i.e. from value $\sigma = \sigma(\varepsilon)$, for the following sets of parameters: $\{\tilde{\lambda} = 3;\ \tilde{\omega} = 3;\ \tilde{\alpha} = 1;\ \tilde{\beta} = 1\}$ and $\{\tilde{\lambda} = \tilde{\omega} = \tilde{\alpha} = \tilde{\beta} = 1\}$.

In the case of "good" sets all the four methods converge fairly quickly and give the same result; the standard error deviations of estimation of parameters $\hat{\sigma}(\theta)$ for all parameters are sufficiently small.

In the case of "bad" sets the results of the methods differ; for some of estimated parameters standard deviations of error of estimations strongly increase with growth of the error of the initial data.

Certainly, such separation is associated with a Hessian, which in the case of "bad" parameters was small. As the case of small values of Hessian is the case of functions of ravine type, it is possible to expect, that methods of adaptive random search can appear comprehensible to such bad cases.

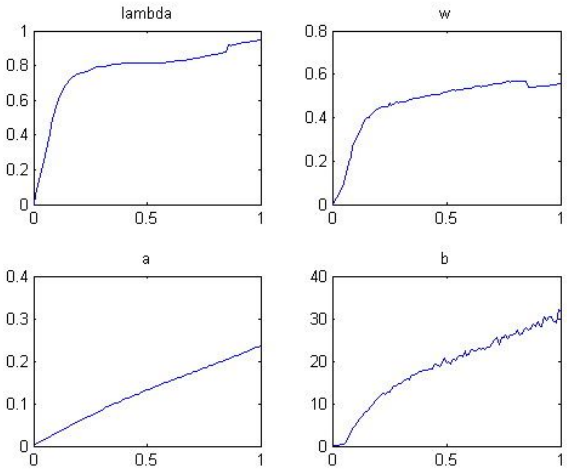


Figure 2: Dependence between $\hat{\sigma}(\theta)$ and $\sigma(\varepsilon)$ for parameters $\{\tilde{\lambda} = 1;\ \tilde{\omega} = 1;\ \tilde{\alpha} = 1;\ \tilde{\beta} = 1\}$ with step $h = 0.01$; parameters were estimate with method Nelder-Mead ($lambda = \tilde{\lambda};\ w = \tilde{\omega};\ a = \tilde{\alpha};\ b = \tilde{\beta}$).

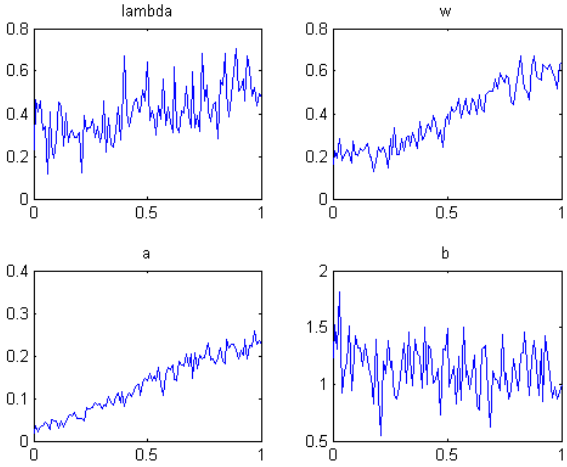


Figure 3: Dependence between $\hat{\sigma}(\theta)$ and $\sigma(\varepsilon)$ for parameters $\{\tilde{\lambda} = 1;\ \tilde{\omega} = 1;\ \tilde{\alpha} = 1;\ \tilde{\beta} = 1\}$ with step $h = 0.01$; parameters were estimate with random search ($lambda = \tilde{\lambda};\ w = \tilde{\omega};\ a = \tilde{\alpha};\ b = \tilde{\beta}$).

Such methods were considered. In particular, some of the genetic algorithms, described in [4], and method of search, based on estimation of covariance matrix, presented in [5], were used. Fig. 3 describes the considered dependencies of errors

Table 1: Separation of some considered values of parameters.

"good" parameters				"bad" parameters			
$\tilde{\lambda}$	$\tilde{\omega}$	$\tilde{\alpha}$	$\tilde{\beta}$	$\tilde{\lambda}$	$\tilde{\omega}$	$\tilde{\alpha}$	$\tilde{\beta}$
2	3	1	4	1	1	1	1
3	3	1	1	0.5	0.5	1	5
5	5	1	0.5	0.5	0.5	1	0.5
0.5	5	1	5	5	0.5	1	5
0.8	4.4	1	2.6	4.3	1.4	1	1.1
0.5	5	1	5	1.71	0.8	1	0.9

for one of the cases of "bad" sets of the parameters, estimated by means of the second of these methods.

Indeed, it appears, that genetic methods are more useful for considered problem, than gradient type methods, and the method, based on an estimation of covariance matrix, gives result even better. In general, for these methods it is possible to notice, that for "bad" sets they give a small error in estimation, the result is more stable subject to error growth of the initial data. However, these methods have larger computation complexity than the gradient type algorithms. Authors intend to construct further special algorithms like an adaptive random search, using combinations of methods (from [4] and [5]), which will be less complex.

References

[1] Ermakov S.M., Zhigljavsky L.A. (1987) *Mathematical theory of optimal experiment.* Nauka, Moscow.

[2] Demidenko E.Z. (1981) *Linear and nonlinear regression.* Finances and Statistics, Moscow.

[3] Bard J. (1979) *Nonlinear estimation of parameters.* Statistics, Moscow.

[4] Kenneth V. Price, Rainer M. Storn, Jouni A. Lampinen (2005) *Differential evolution. A practical approach to global optimization.* Springer, Berlin.

[5] Vladimirova L.V., (1977) *On one method of search of extremum, based on an estimation of covariance matrix.* Automatics and computer science, Riga.

6th St.Petersburg Workshop on Simulation (2009) 1027-1031

On the Consistency of ML-estimates for the Special Model of Survival Curves with Incomplete Data

Anton Korobeynikov[1]

Abstract

We study the estimation of the parameters for a certain special parametric model of survival curves. It is assumed, that survival variable cannot be observed directly and mixed case interval censoring model is used instead. The data consist of a sequence of times of inspection events and a mark variable. Mark variable indicates the endpoints of inspection interval where variable of interest is located. We derive the properties of the maximum likelihood estimates of the parameters and prove their consistency.

1. Introduction

Let positive random variable X denote the time until some event, usually called "failure". *Survival function* $S(\cdot)$ shows the probability of survival beyond a specified time:

$$S(t) = \mathbf{P}(X > t).$$

We consider a special parametric model of survival function given by

$$S_{\eta,\tau}(x) = 1 - F_{\eta,\tau}(x) = \exp(-\eta x)\cos\left(\frac{\pi}{2\tau}x\right), \quad \eta > 0,\ 0 < x < \tau,$$

where $F_{\eta,\tau}(x) = \mathbf{P}(X \leqslant t)$.

This parametric model has been originally introduced in [1] and was successfully used later on to describe the survival dynamics of the chronic glomerulonephritis patients [1], wound processes [2], hypertension [3], generalized severe periodontitis [6].

In practice, survival data can hardly be directly observed. One usually can know only the interval, determined by several observations, where survival variable X is located. For example, consider the study, when X is an onset time of some disease. As it is impossible to provide continuous monitoring we can infer whether disease was developed or not only at certain inspection times. Let us introduce *censoring model* which defines the observable variable Y.

Let K be a positive random integer. Denote by T a triangular array of random variables $\{T_{k,j}, j = 1 \dots k, k = 1 \dots, +\infty\}$ such that $0 = T_{k,0} < T_{k,1} <$

[1]Saint Petersburg State University, E-mail: asl@math.spbu.ru

$T_{k,2} < \cdots < T_{k,k} < T_{k,k+1} = +\infty$. Assume throughout, that variables X and (K, T) are independent. Now introduce random vector $Y = (\Delta_K, T_K, K)$, where T_k is k-th row of the triangular array T, $\Delta_k = (\Delta_{k,1}, \ldots, \Delta_{k,k+1})$ with $\Delta_{k,j} = 1_{(T_{k,j-1}, T_{k,j}]}(X)$. In other words, Y describes the partition of time semi-axis $[0, +\infty)$ on $K + 1$ (random) intervals and specifies the interval containing X.

This censoring model is known as *mixed case interval censoring model* [7] and often used in clinical treats. If $K \equiv 1$, then the model reduces to *case 1 interval censoring model* [4] which is usually described as $Y' = (T, \delta)$, where T is "inspection time" and $\delta = 1_{[X \leqslant T]}$. For $K = k = const > 1$ we come to *case k interval censoring model* [5], where exactly k "inspection times" are allowed.

The typical example of mixed case interval censoring model in clinical studies is the situation when an examination is performed at the start of study and follow-ups are scheduled one at a time till the end of the study. If Z_i denote the times between consecutive follow-ups and L the total duration of the study, then

$$T_{k,j} = \sum_{i=1}^{j-1} Z_i, \quad K = \sup_{j \geqslant 1} \left\{ \sum_{i=1}^{j-1} Z_i < L \right\}. \tag{1}$$

2. Maximum Likelihood Estimates

Suppose we observe n i.i.d. copies of Y: $Y_1, \ldots, Y_n$ with $Y_i = \left(\Delta_{K^{(i)}}, T_{K^{(i)}}, K^{(i)}\right)$. Denote by $\mathcal{L}(\xi)$ the distribution of random variable ξ. One can easily see that conditionally on $K = k$ and $T_K = t_k = (t_{k,1}, \ldots, t_{k,k})$, vector Δ_K has a multinomial distribution:

$$\mathcal{L}(\Delta_K | K = k, T_K = t_k) \sim \mathrm{Mult}(1, h_1, \ldots, h_{k+1}) \tag{2}$$

with $h_j = F_{\eta,\tau}(t_{k,j}) - F_{\eta,\tau}(t_{k,j-1})$ (we will interpret $t_{k,0} = 0$ and $t_{k,k+1} = +\infty$ later on). Denote by G_k the distribution of $(T_K | K = k)$ and assume that the family $\{G_k\}$ is dominated. Let g_k be the density of G_k and $p_k = \mathbf{P}(K = k)$. Then the distribution of Y is also dominated and its density is given by

$$p(y;\, \eta, \tau) = p(\delta_k, t_k, k;\, \eta, \tau) = \prod_{j=1}^{k+1} \left[F_{\eta,\tau}(t_{k,j}) - F_{\eta,\tau}(t_{k,j-1})\right]^{\delta_{k,j}} g_k(t_k) p_k. \tag{3}$$

Here δ_k denotes the indicator vector $(\delta_{k,1}, \ldots, \delta_{k,k+1})$ with one nonzero element.

For sake of brevity we use the following operator notation for taking expectations. We write P for $\mathcal{L}(Y)$ and $\mathbb{P}_n$ for empirical measure induced by observations $Y_1, \ldots, Y_n$. Furthermore, we use the abbreviation Qf for $\int f\, dQ$ for given function f and measure Q.

Denote by $m_{\eta,\tau}$ the log of the density (3) dropping the terms not depending on (η, τ):

$$m_{\eta,\tau}(\delta_k, t_k, k) = \sum_{j=1}^{k+1} \delta_{k,j} \log\left[F_{\eta,\tau}(t_{k,j}) - F_{\eta,\tau}(t_{k,j-1})\right].$$

Then the normalized log-likelihood function for (η, τ) is given by

$$l_n(\eta,\tau) = \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{K^{(i)}+1}\Delta_{K,j}^{(i)}\log\left[F_{\eta,\tau}(T_{K,j}^{(i)}) - F_{\eta,\tau}(T_{K,j-1}^{(i)})\right] = \mathbb{P}_n m_{\eta,\tau}. \quad (4)$$

We will study large sample properties of *approximate maximum likelihood estimates* [8] $(\hat{\eta}_n, \hat{\tau}_n)$, that is,

$$l_n(\hat{\eta}_n, \hat{\tau}_n) \geqslant c \sup_{\eta,\tau} l_n(\eta,\tau), \quad 0 < c \leqslant 1. \quad (5)$$

Later on we use notation (η_0, τ_0) for the true values of unknown parameters and $F_0(x)$ for distribution function $F_{\eta_0,\tau_0}(x)$.

3. Consistency of Maximum Likelihood Estimates

We will start with studying of asymptotic properties of the likelihood function (4) itself. By law of large numbers $l_n(\eta,\tau) = \mathbb{P}_n m_{\eta,\tau} \xrightarrow{\mathbf{P}} P m_{\eta,\tau}$ for each (η,τ), provided that the latter exists.

Proposition 1. *Suppose that* $\mathbf{E}(K) < \infty$. *Then* $P m_{\eta,\tau}$ *is finite for each* (η,τ).

Proof. Our proof is analogous to one in [7]. Using (2) $P m_{\eta,\tau}$ can be written as

$$P m_{\eta,\tau} = \sum_{k=1}^{\infty}\mathbf{P}(K=k)\mathbf{E}(f_{\eta,\tau,k}(T_{k,1},\ldots,T_{k,k})|K=k),$$

where for $0 = t_0 < t_1 < \cdots < t_k < t_{k+1} = +\infty$

$$f_{\eta,\tau,k}(t_1,\ldots,t_k) = \sum_{j=1}^{k+1}[F_0(t_j) - F_0(t_{j-1})]\log[F_{\eta,\tau}(t_j) - F_{\eta,\tau}(t_{j-1})]. \quad (6)$$

We interpret $0\log 0$ as zero and put $\log 0 = -\infty$ throughout. Let $f_{F,k}$ denote the same expression as in (6) but with $F_{\eta,\tau}$ substituted for arbitrary distribution function F. Then for each positive integer k and each set of numbers $t_1 < \cdots < t_k$: $\sup_{\eta,\tau} f_{\eta,\tau,k}(t_1,\ldots,t_k) \leqslant \sup_{\mathcal{F}} f_{F,k}(t_1,\ldots,t_k)$, where $\mathcal{F}$ denotes the set of all probability distribution functions. However, it is easy to check, that the latter supremum attains its maximum on function $\tilde{F} \in \mathcal{F}$ iff $F_0(t_j) = \tilde{F}(t_j)$ for $j = 1,\ldots,k$. Thus, for each k and $0 < t_1 < \cdots < t_k < +\infty$ we have

$$\sup_{\eta,\tau} f_{\eta,\tau,k}(t_1,\ldots,t_k) = \sum_{j=1}^{k+1}[F_0(t_j) - F_0(t_{j-1})]\log[F_0(t_j) - F_0(t_{j-1})].$$

Since $[F_0(t_j) - F_0(t_{j-1})]\log[F_0(t_j) - F_0(t_{j-1})] \leqslant \max_{0\leqslant p\leqslant 1}|p\log p|$ we deduce

$$P m_{\eta,\tau} \leqslant \sum_{k=1}^{\infty}\mathbf{P}(K=k)\mathbf{E}(\sup_{\eta,\tau} f_{\eta,\tau,k}(T_{k,1},\ldots,T_{k,k})|K=k) \leqslant$$
$$\leqslant \sum_{k=1}^{\infty} p_k \sum_{j=1}^{k+1}\max_{0\leqslant p\leqslant 1}|p\log p| \leqslant C\sum_{k=1}^{\infty}(k+1)p_k = C\,(\mathbf{E}(K)+1) < \infty. \quad (7)$$

□

Although it seems reasonable to expect that approximate maximizer $(\hat{\eta}_n, \hat{\tau}_n)$ of likelihood function $\mathbb{P}_n m_{\eta,\tau}$ converges to maximizer of asymptotic likelihood function $Pm_{\eta,\tau}$, in order to obtain consistent estimates we had to prove that $Pm_{\eta,\tau}$ has unique point of maximum and, even more, that the maximum is attained on (η_0, τ_0), the true values of estimated parameters.

Let μ denote the Schick measure [7] on the Borel σ-field $\mathcal{B}$ of subsets of $\mathbb{R}$:

$$\mu(B) = \sum_{k=1}^{+\infty} \mathbf{P}(K=k) \sum_{j=1}^{k} \mathbf{P}\left(T_{k,j} \in B \,|\, K=k\right), \quad B \in \mathcal{B}.$$

Proposition 2. *Suppose that $\mu\left((0,\tau_0) \setminus \{\delta\}\right) > 0$ for any point $\delta \in (0,\tau_0)$. Then* $\operatorname{argmax}_{\eta,\tau} Pm_{\eta,\tau} = (\eta_0, \tau_0)$.

Proof. It follows from proposition 1, that (η_0, τ_0) maximizes the asymptotic likelihood function $Pm_{\eta,\tau}$. Also, any other point of maximum (η', τ') satisfies the equality $F_{\eta',\tau'} = F_0$ μ-a.e. Denote $\Omega_0 = \{x : F_{\eta,\tau}(x) \neq F_0(x)\}$, then $\mu(\Omega_0) = 0$. This means that $\sum_{j=1}^{k'} \mathbf{P}\left(T_{k',j} \in \Omega_0\right) = 0$ for any k' such that $\mathbf{P}(K=k') > 0$. Therefore $\mathbf{P}\left(T_{k',j} \in \Omega_0\right) = 0$ for $j = 1, \ldots, k'$.

It is easy to check that the equality $F_0(x) = F_{\eta',\tau'}(x)$ can take place for at most one $x = x_0 \in (0, \max(\tau', \tau_0))$. Then the set Ω_0 can be written as follows (set $\{x_0\}$ can be empty): $\Omega_0 = (0, \max(\tau', \tau_0)) \setminus \{x_0\}$, hence for $j = 1, \ldots, k'$:

$$\mathbf{P}\left(T_{k',j} \in (0, \max(\tau', \tau_0)) \setminus \{x_0\} \,|\, K=k'\right) = 0,$$

which implies $\mathbf{P}\left(T_{k',j} \in (0,\tau_0) \setminus \{x_0\} \,|\, K=k'\right) = 0$. But then we obtain

$$\mu\left((0,\tau_0) \setminus \{x_0\}\right) = \sum_{k=1}^{+\infty} \mathbf{P}\left(K=k\right) \sum_{j=1}^{k} \mathbf{P}\left(T_{k,j} \in (0,\tau_0) \setminus \{x_0\} \,|\, K=k\right) = 0.$$

The latter equation contradicts the conditions of the proposition and therefore we deduce $(\eta', \tau') = (\eta_0, \tau_0)$. □

Remark 1. *The conditions of the proposition can be slightly weakened when additional information about the distribution of (T, K) is available. Informally speaking, to force the model to be identifiable we need to have at least two distinct observation points on the support of X.*

Now we are ready to state our main result, namely the consistency of approximate maximum likelihood estimates (5).

Theorem 1. *Let $\mathbf{E}(K) < \infty$. Assume that conditions of proposition 2 are satisfied. Then approximate maximum likelihood estimates* (5) *are consistent.*

To prove the theorem we apply van der Vaart's extension to Wald's general consistency theorem (see [8], section 5.2.1):

Theorem 2 (Wald-van der Vaart)**.** *Denote Θ the parameter set. Assume that Θ is metric space. Let the mapping $\theta \mapsto m_\theta(x)$ be upper-semicontinuos for almost all x. Assume that for every sufficiently small ball $U \subset \Theta$ the function $x \mapsto \sup_{\theta \in U} m_\theta(x)$ is measurable and $P \sup_{\theta \in U} m_\theta$ is finite. Denote $\Theta_0 = \{\theta_0 \in \Theta : Pm_{\theta_0} = \sup_\theta Pm_\theta\}$ and consider the sequence of estimators $\hat{\theta}_n$ such that $\mathbb{P}_n m_{\hat{\theta}_n} \geqslant c\mathbb{P}_n m_{\theta_0}$ for some $\theta_0 \in \Theta_0$ and $0 < c \leqslant 1$. Then*

$$\mathbf{P}\left(dist\left(\hat{\theta}_n, \Theta_0\right) \geqslant \varepsilon,\, \hat{\theta}_n \in K\right) \to 0 \quad as \quad n \to \infty$$

for any $\varepsilon > 0$ and any compact $K \subset \Theta$:

Proof of theorem 1. The natural parameter space $\Theta = (0, +\infty) \times (0, +\infty)$ is not compact, therefore, in order to use Wald's method and obtain consistent estimates some compactification of parameter space is needed. It seems, that the most convenient way to do so is to enlarge Θ to compact subset $[0, +\infty] \times [0, +\infty]$ of $\bar{\mathbb{R}} \times \bar{\mathbb{R}}$. Thus it is necessary to extend the definition of likelihood to the case $\eta, \tau \in \{0, +\infty\}$, which we perform by continuity (ignoring the case where $T_{K,1} = 0$, which has probability zero):

$$l_n(+\infty, \cdot) = l_n(\cdot, 0) = -\infty,$$

$$l_n(0, \tau) = \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{K^{(i)}+1} \Delta_{K,j}^{(i)} \log \left[\cos\left(\frac{\pi}{2\tau} T_{K,j-1}^{(i)}\right) - \cos\left(\frac{\pi}{2\tau} T_{K,j}^{(i)}\right)\right],$$

$$l_n(\eta, +\infty) = \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{K^{(i)}+1} \Delta_{K,j}^{(i)} \log \left[\exp\left(-\eta T_{K,j-1}^{(i)}\right) - \exp\left(-\eta T_{K,j}^{(i)}\right)\right],$$

hence the map $(\eta, \tau) \mapsto l_n(\eta, \tau)$ is continuous at every (η, τ).

Note, that theorem 2 implicitly assumes, that at least Pm_{η_0, τ_0} is finite, which is guaranteed by proposition 1 given that $\mathbf{E}(K) < \infty$ (all results of propositions 1 and 2 still take place after the compactification). Even more, the same proposition implies, that $P \sup_{\eta,\tau} m_{\eta,\tau} < \infty$. Identifiability condition due to proposition 2 immediately gives $\Theta_0 = \{(\eta_0, \tau_0)\}$.

Then, all conditions of Wald's theorem are satisfied and taking $K = \bar{\mathbb{R}} \times \bar{\mathbb{R}}$ we obtain that

$$(\hat{\eta}_n, \hat{\tau}_n) \xrightarrow[n \to \infty]{\mathbf{P}} (\eta_0, \tau_0)\,.$$

□

4. Acknowledgement

Author would like to thank V.V. Nekrutkin for thoughtful comments that led to significantly improved presentation of the paper.

References

[1] Bart, A.G., Bondarenko, B.B and Boiko, B.I. (1980, in Russian) Mathematical analysis of the evolution of chronic glomerulonephritis. In: Riabov, S.I (Ed.), *Glomerulonephritis*. Leningrad: Medicine, 213-225.

[2] Bart, A.G. (2003, in Russian) *Analysis of the Medical and Biological Systems (The Inverse Functions Approach)*. St. Petersburg: St. Petersburg University Press.

[3] Bart, A.G., Bart, V.A., Steland, A. and Zaslavskiy M.L. (2005) Modelling disease dynamics and survivor function by sanogenesis curves. *J. of Stat. Planning and Inference*, **132**, 33-51.

[4] Groeneboom, P. and Wellner, J.A. (1992) *Information Bounds and Nonparametric Maximum Likelihood Estimation*. Basel: Birkhäuser.

[5] Huang, J. and Wellner, J.A. (1997) Interval censored survival data: a review of recent progress. In *Procs. of the First Seattle Symposium in Biostatistics: Survival Analysis*, **123**, 123-169.

[6] Madai, D.Yu., Bart, A.G., Korobeynikov, A.I. et all. (2006, in Russian) *Reparative efficiency of 'Vilon' for cure of aged patients with generalized severe periodontitis*. Novgorod: Novgorod University Press.

[7] Schick, A. and Yu, Q. (2000) Consistency of the GMLE with mixed case interval censored data. *Scand. J. Stat.*, **27**, 45-55.

[8] van der Vaart, A.W. (1998) *Asymptotic Statistics*. Cambridge: Cambridge University Press.

6th St.Petersburg Workshop on Simulation (2009) 1033-1037

Fuzzy Fisher test for contingency tables with application to genomics data

Natalia Bochkina[1], Alex Lewin [2]

We address the problem of testing for independence between factors in a 2 by 2 contingency table where there is an uncertainty about classification of observations into the factor levels. Such problem often arises in genomics studies where variables (genes) are classified as associated with disease A or disease B using noisy data, and it is of interest to test whether there is an association between the diseases at the genetic level. Given the probability of classification of each variable into the factor levels for two factors of interest, we propose a test of no association between the factors, derive its asymptotic distribution and study its power. This test generalises Fisher's test for contingency tables to the case of fuzzy classification. We also propose a Bayesian framework to address this problem. Performance of both methods is compared to other approaches using simulated and genomics data.

[1]University of Edinburgh, UK, E-mail: N.Bochkina@ed.ac.uk
[2]Imperial College London, UK, E-mail: A.M.Lewin@ic.ac.uk

6th St.Petersburg Workshop on Simulation (2009) 1034-1038

Bootstrap from moderate deviation viewpoint

Mikhail Ermakov[1]

Abstract

The bootstrap became the standard tool for solving problems of confidence estimation and hypothesis testing. The significant levels in confidence estimation and type I error probabilities in hypothesis testing have usually small values. By this reason the bootstrap analysis in moderate deviation zone causes the natural interest. In the talk we show that the distributions of statistics and their bootstrap versions admit the same normal approximation in moderate deviation zones. However the moderate deviation zones for these two normal approximations are different. The zones of normal approximation of statistics are narrower than the corresponding zones of their bootstrap versions. Thus the problem of bootstrap adequacy deserves more serious study in the applications. The results established in terms of Moderate Deviation Principles for empirical probability measure and empirical bootstrap measure.

1. Introduction

Let S be a Hausdorff space, $\Im$ the σ-field of Borel sets in S and Λ the space of all probability measures (pms) on $(S, \Im)$. Let $X_1, \ldots, X_n$ be i.i.d.r.v.'s taking values in S according to a pm $P_0 \in \Lambda$ and let $\hat{P}_n$ be the empirical measure of $X_1, \ldots, X_n$. The distributions of statistics depending on the sample $X_1, \ldots, X_n$ are often analyzed on the base of the bootstrap procedure. For given statistics $V(X_1, \ldots, X_n)$, we simulate independent samples $X_1^*, \ldots, X_n^*$ with the distribution $\hat{P}_n$ and consider the empirical distribution of $V(X_1^*, \ldots, X_n^*)$ as an estimator of the distribution of $V(X_1, \ldots, X_n)$.

What is of special interest, are the estimates of large and moderate deviation probabilities of $V(X_1, \ldots, X_n)$. Such problems constantly emerge in confidence estimation and hypothesis testing. The significant levels in the confidence estimation and the p-values in the hypothesis testing usually have small values and therefore often can be more correctly described by the theorems on large and moderate deviations. From this viewpoint it is natural to compare the probabilities of large and

[1]Institute of Problems of Mechanical Engineering, E-mail: ermakov@random.ipme.ru

moderate deviations of $V(X_1,\dots,X_n)$ and $V(X_1^*,\dots,X_n^*)$. In the paper we carry out such a comparison in a slightly different setting. The statistics $V(X_1,\dots,X_n)$ usually can be represented as a functional $T(\hat{P}_n)$ of the empirical measure $\hat{P}_n$, that is, $V(X_1,\dots,X_n)=T(\hat{P}_n)$. Similarly, $V(X_1^*,\dots,X_n^*)=T(P_n^*)$, where P_n^* is the empirical measure of $X_1^*,\dots,X_n^*$. Thus, we can reduce the problem to the study of large and moderate deviations probabilities of statistical functionals $T(\hat{P}_n)-T(P)$ and $T(P_n^*)-T(\hat{P}_n)$. In this paper, we carry out such a type of comparison with the help of Moderate Deviation Principle (MDP). The LDP - MDP analysis for i.i.d. random objects is well known from Sanov [5], Borovkov and Mogulskii [2], Dembo and Zeitouni [3], Arcones [1] , where main results are obtained under rather general assumptions.

Our goal is twofold.

1. We develop MDP technique from the above-mentioned papers for

$$(P_n^*-\hat{P}_n)\times(\hat{P}_n-P).$$

and implement the above result for the MDP comparison of

$$T(\hat{P}_n)-T(P) \text{ and } T(P_n^*)-T(\hat{P}_n).$$

2. We establish the MDP for a conditional distribution of an empirical bootstrap measure P_n^* given empirical probability measure $\hat{P}_n$.

It turns out that MDP for the joint

"empirical bootstrap + empirical probability"

measures is valid in "smaller time zone" than the MDP for empirical measure only. On the other hand, the time zone for the above-mentioned conditional MDP is essentially larger with probability close to one. The first statement shows instability of a bootstrap procedure provided that the empirical measure belongs to the MDP zone. The second statement shows that the bootstrap statistics have more stable properties.

2. Main Results

Bellow MDP will be given only for empirical probability measures and conditional empirical bootstrap measures. This allows to compare the results. MDP for empirical measure follows straightforwardly from Arcones [1]. All above mentioned results are discussed in my talk.

Through the paper, the following notations are implemented:

- $Q_2\times Q_1$ the Cartesian product of probability measures $Q_2,Q_1\in\Lambda$;
- $\Lambda^2=\Lambda\times\Lambda$ denotes the set of all measures $Q_2\times Q_1$ with $Q_2,Q_1\in\Lambda$;
- C,c are generic positive constants;
- $\chi(A)$ is the indicator of event A;
- $[t]$ is the integral part of real number t;
- $\int$ always denotes $\int_S$.

The results are given in terms of $\boldsymbol{\tau_\Phi}$-topology.

Let us fix a decreasing sequence of positive numbers $(b_n)_{n\geq 1}$ with properties:

$$\left.\begin{array}{l} b_n \to 0 \\ nb_n^2 \to \infty \\ \frac{b_n}{b_{n+1}} \to 1 \end{array}\right\} n \to \infty. \tag{1}$$

Denote Φ the set of measurable functions $f : S \to \mathbb{R}$ with the following property:

$$\lim_{n\to\infty} \frac{1}{nb_n^2} \log(nP(|f(X)| > b_n^{-1})) = -\infty. \tag{2}$$

Let

$$\Lambda_\Phi = \Big\{ P \in \Lambda : \int |f(X)| dP < \infty,\ \forall\ f \in \Phi \Big\}.$$

The the **coarsest** topology in Λ_Φ providing the continuous mapping

$$Q \Rightarrow \int f dQ,\ \forall\ f \in \Phi,\ Q \in \Lambda_\Phi$$

is known as τ_Φ-topology (henceforth, all topological notions are related to the τ_Φ-topology).

For any set $\Omega \subset \Lambda_\Phi$ notations: $\mathfrak{clo}(\Omega)$ and $\mathfrak{int}(\Omega)$ are used for the closure and interior of Ω respectively.

For $P, Q \in \Lambda$ and $P, Q \in \Lambda_\Phi$, we define the sets Λ_0 and $\Lambda_{0\Phi}$ respectively of differences $P - Q$. The τ_Φ-topologies in $\Lambda_{0\Phi}$ and $\Lambda_{0\Phi}^2$ are defined in a standard way as well as $\mathfrak{clo}(\bar{\Omega}_0)$ and $\mathfrak{int}(\bar{\Omega}_0)$ the closure and interior of $\bar{\Omega}_0 \subset \Lambda_{0\Phi}^2$.

For $G \in \Lambda_0$, let

$$\rho_0^2(G|P) = \begin{cases} \frac{1}{2} \int \left(\frac{dG}{dP}\right)^2 dP, & G \ll P \\ \infty, & \text{otherwise} \end{cases}$$

is the rate function (in statistical terms, $2\rho_0^2(G|P)$ is the Fisher information) which arises naturally in the MDP analysis of empirical measures $\hat{P}_n$ (see Borovkov and Mogulskii [2]; Ermakov [4] and Arcones [1])

Bellow the MDP for empirical probability measures is given.

Define the set Φ of measurable functions $f : S \to R^1$ such that

$$\lim_{n\to\infty} (nd_n^2)^{-1} \log(nP(|f(X)| > nd_n)) = -\infty \tag{3}$$

where $d_n \to 0, nd_n^2 \to \infty, d_{n+1}/d_n \to 1$ as $n \to \infty$.

Let the charges $H, H_n \in \Lambda_{0\Phi}$ satisfy the following assumptions.

A. There hold $P_n = P + b_n H_n \in \Lambda_\Phi, P + b_n H \in \Lambda_\Phi$ and $H_n \to H$ as $n \to \infty$ in the τ_Φ-topology.

B. For any $f \in \Phi$

$$\lim_{n\to\infty} (nd_n^2)^{-1} \sup_{m\geq n} \log \left(nd_n \int \chi(|f(x)| > nd_n)\, d|H_m| \right) = -\infty. \tag{4}$$

Theorem 1. *Assume A and B. Let $\Omega_0 \subset \Lambda_{0\Phi}$. Then the MDP holds*

$$\liminf_{n\to\infty}(nd_n^2)^{-1}\log P_n(\hat{P}_n \in P + d_n\Omega_0) \geq -\rho_0^2(\mathfrak{int}(\Omega_0 - H), P_0) \tag{5}$$

and

$$\limsup_{n\to\infty}(nd_n^2)^{-1}\log P_n(\hat{P}_n \in P + d_n\Omega_0) \leq -\rho_0^2(\mathfrak{cl}(\Omega_0 - H), P_0) \tag{6}$$

Theorem 2 given below shows that the MDP holds with probability $1+o(1)$ for the conditional distribution of the empirical bootstrap measure given the empirical probability measure. This version of MDP we call the conditional MDP. In this model we allow the sample size $k = k_n$ of bootstrap procedures has the values different from n.

Let $X_1^*, \ldots, X_{k_n}^*$ be i.i.d.r.v.'s having pm $\hat{P}_n$. Denote $P_{k_n}^*$ the empirical probability measure of $X_1^*, \ldots, X_{k_n}^*$. Suppose that $\frac{k_n}{n} < c < \infty$ and $k_n \to \infty$ as $n \to \infty$.

Theorem 2. *Let a sequence $a_n > 0, a_n \to 0, a_{n+1}/a_n \to 1, k_n a_n^2 \to \infty$ as $n \to \infty$ be given. Let function $h : R_+^1 \to R_+^1$ be such that*

$$\lim_{n\to\infty} nh\left(\frac{a_n}{c}\right) = 0 \tag{7}$$

for any $c > 0$. Suppose we are given a set Θ_h of real functions such that for any $f \in \Theta_h$

$$P(|f(X)| > s^{-1}) < h(s), \quad s > 0. \tag{8}$$

Let there exists $t > 2$ such that

$$E|f^2(X) - Ef^2(X)|^t < \infty \tag{9}$$

for all $f \in \Theta_h$.

Then for any $\Omega_0 \subset \Lambda_{0\Theta_h}$, for any $\epsilon > 0$ and $n > n_0(\epsilon, \{k_i\}_{i=1}^\infty)$ there hold

$$(k_n a_n^2)^{-1}\log \hat{P}_n(P_{k_n}^* \in \hat{P}_n + a_n\Omega_0) \geq -\rho_0^2(\mathfrak{int}(\Omega_0), P) - \epsilon \tag{10}$$

and

$$(k_n a_n^2)^{-1}\log \hat{P}_n(P_{k_n}^* \in \hat{P}_n + a_n\Omega_0) \leq -\rho_0^2(\mathfrak{cl}(\Omega_0), P) + \epsilon \tag{11}$$

with probability $\kappa_n = \kappa_n(\epsilon, \Omega_0) = 1 - C(\epsilon, \Omega_0)[\beta_{1n} + \beta_{2n}]$ where $\beta_{1n} = nC(\Omega_0)h(\frac{a_n}{\epsilon C_1(\Omega_0)})$ and

$$\beta_{2n} = C_1(\Omega_0)n^{1-t}\epsilon^{-t} + \exp\{-C_2(\Omega_0)\epsilon^2 n\}.$$

Example. Let $E[\exp\{cY_1^\gamma\}] < \infty$ with $\gamma > 0$. Then we have the following asymptotics

$$b_n = o\left(n^{-\frac{1}{1+\gamma}}\right), \quad a_n = o\left(|\log n|^{-\gamma}\right).$$

Thus the conditional MDP for empirical bootstrap measure holds for wider zone than the usual MDP for the empirical measure.

References

[1] M.A. Arcones. Moderate deviations of empirical processes. In *Stochastic Inequalities and Applications,* 189-212 ed.E.Gine, C.Houdre and D.Nualart. Birkhauser Boston, 2003.

[2] A.A. Borovkov and A.A. Mogulskii. On probabilities of large deviations in topological spaces. II. *Siberian Mathematical Journal,* 21(5):12-26, 1980.

[3] A. Dembo and O. Zeitouni. *Large Deviations Techniques and Applications.* Jones and Bartlett, Boston, 1993.

[4] M.S. Ermakov. Asymptotic minimaxity of tests of Kolmogorov and omega-square type. *Theory Probability and Applications,* 40:54-67, 1995.

[5] I. Sanov. On the probability of large deviations of random variables. *Mathematicheskii Sbornik*, 42: 70-95 1957(In Russian) English translation:*Selected Translations in Mathematical Statistics and Probability*, 1: 213-244, 1961.

Section

Experimental design

6th St.Petersburg Workshop on Simulation (2009) 1041-1045

On Capacity of Screening Experiments under Linear Programming Analysis

Mikhail Malyutov[1], Hanai Sadaka[2]

Abstract

Screening experiments (SE) deal with finding a small number s of significant factors out of a vast total number t of inputs in a regression model. Of special interest in the SE theory is finding the so-called maximal rate (**capacity**) defined as $\log t/N(s,t,\gamma)$ such that a *random* $(N \times t)$*-design matrix with* $N \geq N(s,t,\gamma)$ *enables identifying* s *randomly chosen significant variables out of* t *with the probability exceeding* $1-\gamma$. The capacity was found asymptotically as $t \to \infty$ in a very general setting for the '**brute force**' analysis of experiments in (Malyutov(1979)) and its relation to the capacity region of Multiple Access Communication Channel was outlined. In this paper, we use a simple tractable linear programming relaxation instead of the brute force analysis, and we use simulations to approximate the quantity $N^*(s,t,\gamma)$ such that the same property as above holds, if $N \geq N^*(s,t,\gamma)$ for analysis of experiments using linear programming relaxation. We find that the linear programming relaxation is often successful in finding the significant variables, but the hypothesis $N^*(s,t,\gamma) = N(s,t,\gamma)$ is not supported by our simulation, i.e. it turns out that the capacity of screening under this practical method of analysis is less than that for the brute force analysis.

In Appendix we review results on the capacity of screening in most general models for readers interested in extending the L_1 analysis to the general case.

1. Capacity of Screening Experiments

The first models of screening that were studied include the Boolean sum model, and the forged coin model (FC-model). The Boolean sum model, or simply $\cup$-model, was used in Dorfman (1943) for modeling the pooled testing of patients blood for the presence of a certain antigen. Dorfman's innovative design was successful in reducing experimental costs by an order of magnitude! Another popular combinatorial example is the search for a subset of forged coins of identical weight distinct from the weight of genuine coins by weighing the minimal number of subsets of coins described by P. Erdös and A. Renyí in the sixties.

Let us formalize these two examples to show their interrelation. Introduce binary variables (inputs) $x_i(a)$ as indicators of participation of a-th coin (patient)

[1]Northeastern University, E-mail: mltv@neu.edu

[2]Northeastern University, E-mail: h.sadaka@neu.edu

in the i-th trial. Let A be the subset of forged coins (or, respectively, of sick patients) and denote the corresponding subset of **significant** variables in the i-th experiment by $x_i(A)$. Given the $N \times t$ design matrix $X = \{x_i(a)\}, i = 1, \ldots, N$, and $a = 1, \ldots, t$, the output y_i of the i-th trial can then be described by the formula:

$$y_i = g(x_i(A)), \tag{1}$$

where in our first example g represents a boolean sum, (we use notation $\cup$ for boolean sums)

$$g(x(A)) = \cup_{a \in A} x(a), \tag{2}$$

and in the second example g represents an ordinary summation:

$$g(x(A)) = r \sum_{a \in A} x(a), \tag{3}$$

where r is the weight difference of forged and genuine coins. Static designs (we only deal with static designs in this paper) for the $\cup$-model were successfully applied in various settings, e.g. for quality improvement of complex circuits in Quality Center of London City University (H.P. Wynn et al (1991)), for trouble-shooting of large circuits with redundancy (Malyutov et al (1976)), etc. The FC-model was applied for transmitting information packets via multi-access communication channels.

Of special interest in the SE theory is finding the so-called maximal rate or (**capacity**) defined as $\log t / N(s, t, \gamma)$ such that a *randomly chosen $N \times t$ design matrix with $N \geq N(s, t, \gamma)$ enables* **identifying** *s randomly chosen significant variables out of t with the probability exceeding $1 - \gamma$*. Here all the logarithms are binary. We call 'random' a matrix with I.I.D. entries determined by $\beta = P(x_i(a) = 0)$. The upper bounds below presume the optimal choice of $\beta = 2^{-1/s}$ for $\cup$-model, and $\beta = 1/2$ for FC-model to minimize $N(s, t, \gamma)$. While the latter choice seems natural for the FC-model, the proof is not trivial.

Due to the symmetry our randomized designs, $N(s, t, \gamma)$ has an alternative combinatorial interpretation, namely that a *'weakly separating'* $(N \times t)$ matrix **exists for** $N \geq N(s, t, \gamma)$ such that it allows identifying the true s-tuple of significant inputs with probability $1 - \gamma$ under uniform distributions of significant t-tuples. The mathematical beauty of the weakly separating property lies in the possibility of finding the *asymptotics* for $N(s, t, \gamma)$ as $t \to \infty$ in the most general model with nonparametric response function and arbitrary unknown measurement noise – we review this in the next section.

The theory of *'strongly separating designs' (SSD) for sure identification of every significant s-tuple* is far from being complete: existing lower and upper bounds for the minimal number $N(s, t)$ in SSD differ several times in *all models* including the two elementary ones dealt with in this paper. Apparently, this is because randomly generated strongly separating designs require asymptotically larger N than the best combinatorial ones.

The upper bounds $\bar{N}(s, t, \gamma)$ for $N(s, t, \gamma)$ in many elementary models including the above two can be found in a survey paper Malyutov (1977). An early upper bound for $\cup$-model $N(s, t, \gamma) \leq s \log t + 2s \log s - s \log \gamma$ was obtained in Malyutov

(1976) and strengthened in Malyutov (1978). For the FC-model the asymptotic capacity expression is $\bar{N}(s,t,\gamma) = s \log t / H(B_s(1/2))(1+o(1))$ as $t \to \infty$, where $B_s(1/2)$ is Binomial with parameter 1/2, and $H(B_s(1/2))$ is its binary entropy.

'Identifying' in the above definition means the unique restoration by the **'brute force' analysis**, which involves searching through all possible subsets of all variables. Of course, this type of analysis becomes extremely computationally intensive for even moderate values of s as $t \to \infty$. The analysis problem becomes even more critical in more general models with noise and nuisance parameters introduced in our Appendix.

Instead of using brute force analysis we use linear programming relaxations for both of the problems. For the linear version of the problem we use the popular ℓ_1-norm relaxation of sparsity. The problem in (3) can be represented as

$$\min |A|, \text{ such that } y_i = \sum_{a \in A} x_i(a), \ \forall i \tag{4}$$

Here $|A|$ represents the number of elements in A. Instead, we define the indicator vector I_A such that $I_A(a) = 1$ if $a \in A$, and focus on the ℓ_1-norm of I_A, i.e. on $\sum_a I_A(a)$. Note that in our case $I_A \in \{0,1\}$, so it is always positive, so instead of $\sum_a |I_A(a)|$ we can use $\sum_a I_A(a)$. We solve the relaxed problem

$$\min \sum_a I_A(a), \text{ such that } y_i = \sum_{a \in A} x_i(a), \ \forall i \tag{5}$$

This type of relaxation has received considerable amount of attention in many fields, including statistics (Lasso regression [**?**]), and in signal processing (basis pursuit, [**?**]). While the much simpler linear programming relaxation is not guaranteed to solve the original combinatorial problem, theoretical conditions were developed showing that if the unknown sparse signal is 'sparse enough', then the linear programming relaxation exactly recovers the unknown signal [2]. This problem has also been studies in the random setting, and bounds were developed on recovery of the unknown signal with high probability [**?**].

For the non-linear problem we are forced to relax not only the sparsity of the indicator vector I_A, but also the measurement model. Since $y_i = \cup_{a \in A} x_i(a)$, then it must hold that $y_i \leq \sum_{a \in A} x_i(a)$. Hence, our first relaxation is

$$\min \sum_a I_A(a) \text{ such that } y_i \leq \sum_{a \in A} x_i(a), \quad 0 \leq I_A(a) \leq 1 \tag{6}$$

We also note that if $y_i = 0$, then it must hold that all $x_i(a) = 0$ and the inequality $y_i \leq \sum_{a \in A} x_i(a)$ holds with equality. Hence, a stronger relaxation is obtained by enforcing this equality constraint.

$$\min \sum_a I_A(a) \text{ such that } 0 \leq I_A(a) \leq 1, \text{ and } y_i \leq \sum_{a \in A} x_i(a), \quad \text{if } y_i \neq 0 \text{ and} \tag{7}$$

$$y_i = \sum_{a \in A} x_i(a) \quad \text{if } y_i = 0 \text{ and} \tag{8}$$

To our knowledge, bounds for the performance of this linear programming relaxation of the nonlinear screening problem have not been studied in the literature.

2. N^* simulation for $\cup$-model

a) Set $t = 100(200, 400, \dots)$, $s = 2$, $N = 20$.

b) Create a random vector $\mathbf{x}$ using:
$r_{\text{perm}} = \text{randperm}(t)$;
$\mathbf{x}(r_{\text{perm}}(1:s)) = 1$;

c) Solve 100 trials to find N^*:

(a) Generate the binary random matrix A, with $p(0) = \sqrt[s]{\frac{1}{2}}$

(b) Solve min $\| \mathbf{x} \|_1$ such that $A\mathbf{x} \geq \mathbf{y}$, where $\mathbf{y} = (A\mathbf{x} > \mathbf{0})$.

(c) Compare if $\| \hat{\mathbf{x}} - \mathbf{x} \|_1 \geq 10^{-6}$, then failure, otherwise success, where $\hat{\mathbf{x}}$ is the solution to step (3b).

(d) If $p(\text{failure}) > 0.05$, increase N, otherwise stop.

d) set $s = s + 1$, then go to step 1.

s/t	100	200	400	800	1600	3200	6400	12800
2	49	64	70	82	94	109	123	138
3	75	94	114	142	168	197	231	256
4	96	126	157	192	237	281	325	372
5		145	190	242	296	365	428	496
6		175	220	283	358	431	518	602
7		190	248	321	405	501	607	779

If we replace the boolean sum $[A\mathbf{x}]$ by the regular sum $A\mathbf{x}$ for the binary matrices A and binary vectors $\mathbf{x}$ and solve

$$\begin{aligned} & \min ||\mathbf{x}||_1 \\ \text{s.t.} \quad & A\mathbf{x} \geq \mathbf{y}, \end{aligned}$$

where $\mathbf{y} = [A\mathbf{x}]$. We reduce $N*$ as it shown in the table below

s/t	100	200	400	800	1600	3200	6400	12800
2	23	26	28	30	32	34	35	38
3	36	39	43	45	48	51	54	56
4	49	52	55	58	63	66	68	74
5	61	67	72	76	80	85	87	91
6	74	80	86	89	93	100	105	111
7	91	96	101	108	113	118	124	129

The table below represents the early upper bound (Malyutov (1976)) for $N(s, t, \gamma) \leq s \log_2 t + 2s \log_2 s - s \log_2(0.05)$.

s/t	100	200	400	800	1600	3200	6400	12800
2	16	17	18	20	22	24	27	28
3	22	24	25	29	31	33	35	38
4	28	32	33	35	39	41	44	49
5	35	36	40	44	47	52	57	62
6	44	44	50	51	53	58	64	69
7	51	53	55	59	62	65	73	79

Malyutov (1978), Remark 3, p. 166, gives a more accurate upper bound $N(s,t,\gamma) \leq \log \binom{t-s}{s} - \log(\gamma - t^{-c_{17}})$, where constant c_{17} depending only on s is obtained as a result of a transformation chain consisting of 17 steps. For $\gamma = 0.05$, replacing $-\log(\gamma - t^{-c_{17}})$ for our big $t's$ with a larger value 5, we get the following table.

s/t	100	200	400	800	1600	3200	6400	12800
2	17.21	19.25	21.27	23.28	25.28	27.29	29.29	31.29
3	22.17	25.26	28.30	31.32	34.34	37.34	40.34	43.35
4	26.66	30.83	34.91	38.95	42.97	46.98	50.99	54.99
5	30.79	36.06	41.19	46.25	51.28	56.30	61.30	66.31
6	34.60	41.00	47.19	53.28	59.33	65.35	71.36	77.37
7	38.14	45.69	52.95	60.08	67.14	74.18	81.19	88.20

Comparing the last and the first two tables we conclude that the capacity of $\cup$-screening is **smaller** than the simulated N^* under the Linear Programming analysis.

3. N^* simulation for the FC model

The linear code is similar, we replace $\mathbf{y} = (A\mathbf{x} > \mathbf{0})$ by $\mathbf{y} = A\mathbf{x}$ and the starting point for N is 15.

s/t	100	200	400	800	1600	3200	6400	12800
2	16	17	18	20	22	24	27	28
3	22	24	25	29	31	33	35	38
4	28	32	33	35	39	41	44	49
5	35	36	40	44	47	52	57	62
6	44	44	50	51	53	58	64	69
7	51	53	55	59	62	65	73	79

We know only an *asymptotic* upper bound for the FC-model capacity obtained in Malyutov (1977), Theorem 5.4.1, (following also from our general result (Malyutov and Mateev (1980)) on *'ordinarity'* of symmetrical models) which use a non-trivial entropy $H(B_s(p))$ of binomial distribution maximization result by P. Mateev: $\max_{0<p<1} H(B_s(p)) = H(B_s(1/2)) := a_s$ for all $s \geq 1$: If for some $0 < \beta < 1$

$$N \geq s \log t/a_s + \kappa(\log t/a_s)^{(1+\beta)/2}$$

then $\gamma \leq (s \log t/a_s)^{\beta}$. Thus $logt/N(s,t,\gamma) \to a_s/s$ as $t \to \infty$.

$s : H(B_s)$	2:1.81	3:2.03	4:2.20	5:2.33	6:2.45	7:2.54

Using the preceding $H(B_s)$ table, we prepare the $\log \binom{t}{s}/a_s$ table which is asymptotically equivalent to $N(s,t,\gamma)$ (as $t \to \infty$) for any $0 < \gamma < 1$.

s/t	100	200	400	800	1600	3200	6400	12800
2	8.18	9.52	10.86	12.19	13.52	14.86	16.19	17.53
3	9.55	11.22	12.88	14.54	16.20	17.86	19.51	21.17
4	10.79	12.78	14.76	16.73	18.71	20.68	22.65	24.62
5	11.90	14.21	16.50	18.79	21.06	23.34	25.62	27.89
6	12.92	15.54	18.14	20.72	23.30	25.87	28.44	31.02
7	13.85	16.78	19.67	22.55	25.42	28.28	31.15	34.01

As in $\cup$-model, $N^*(s,t)$ exceed the brute force approximated $N(s,t,\gamma)$ significantly (more than two times). The asymptotic nature of the last table cannot influence this comparison. We observe that for large t, (when the accuracy of our approximation is better), their ratio becomes even larger.

4. Discussion

Comparison of our simulation results with theoretical upper bounds under the brute force analysis for the two popular elementary models suggests: the vaguely formulated statement that the L_1 analysis solves screening problems as good as the brute force analysis is at best an exaggeration. Of course, only a formal proof can firmly establish this. Moreover, running several independent series of 100 experiments for each s,t to estimate variability of N^* (which is in work now) would be a more reliable argument.

Also, running, say 100 series of 100 experiments for each s,t such that only 5 per cent of the series contain at least one misidentification of the s-tuple SI's would estimate the Linear Programming analogue of capacity for strongly separating random designs to compare with the upper bound for the brute force capacity established theoretically. We expect similar discrepancy between these two capacities to be true and plan to show this in our future simulations.

References

[1] Csiszar, I. and Körner, J. (1981) *Information Theory: Coding Theorems for Discrete Memoryless Systems*, Academic Press and Akadémiai Kiadó, Budapest.

[2] Donoho, D. and Elad, M. (2003) Maximal Sparsity Representation via L_1 Minimization, Proc. Nat. Aca. Sci., Vol. 100, pp. 2197-2202, March 2003.

[3] Erdös, P. and Renyi, A. (1963) On two Problems of Information Theory, Publ. Math. Inst. of Hung. Acad.of Sc., **8,** 229-243.

[4] Malioutov, D.M., Cetin, M. and Willsky, A.S.(2004) Optimal Sparse Representations in general overcomplete Bases, IEEE International Conference on Acoustics, Speech, and Signal Processing, May 2004, Montreal, Canada, pp. II-793-796 vol.2.

[5] Malyutov, M.B. and Tsitovich, I.I. (2000) Non-parametric Search for Significant Inputs of Unknown System, In N. Callaos editor, Proceedings of SCI'2000/ISAS 2000 World. Multiconference on Systemics, Cybernetics and Informatics, July 23-26 2000, Orlando, FL, vol. XI, 75-83.

[6] Malyutov, M.B. and Sadaka, H. (1998) Jaynes Principle in Testing Significant Variables of Linear Model, Random Operators and Stochastic Equations, **6**, 311-330.

[7] Malyutov, M.B. and Mateev, P.S. (1980) Screening Designs for Non-Symmetric Response Function. Mat. Zametki, **27**, l09-127.

[8] Malyutov, M.B. (1979) On the maximal rate of screening designs. Theory Probab. and Appl., XXIV no.3.

[9] Malyutov, M.B. (1978) Separating Property of Random Matrices. Mat. Zametki, **23**, 155-167.

[10] Malyutov, M.B. (1977) Mathematical Models and Results in Theory of Screening Experiments. In Theoretical Problems of Experimental Design, ed. by Malyutov M.B., 5-69, Soviet Radio, Moscow (In Russian).

[11] Malyutov, M.B. (1976) On Planning of Screening Experiments. In Proceedings of 1975. IEEE-USSR Workshop on Inform. Theory, N.Y. ,IEEE Inc., 1976, 144-147.

6th St.Petersburg Workshop on Simulation (2009) 1048-1052

Adaptive Designs for Dose Escalation Studies[1]

Katrin Roth[2]

Abstract

In Phase I dose escalation studies the 3+3 design is a commonly used method. We suggest a modification of this method to include a parametric model and the concepts of optimal design theory, with the aim of improving this method. The performance of the original and the new approach are investigated by simulations.

1. Introduction

Phase I dose escalation studies are part of the clinical drug development process. At that stage of the development, little knowledge about how the drug and the human body interact is available. The primary goal of these studies is to find the maximum tolerated dose (MTD). This is the dose that induces an intolerable toxic event (dose limiting toxicity, DLT) with a probability less than $\frac{1}{3}$. Due to safety issues, dose escalation studies are performed adaptively. An approach widely used is the 3+3 design. Following this methods, subjects are assigned in cohorts of three to one out of a sequence of specified doses. The first cohort is assigned to the lowest dose. The following cohorts are assigned to the next higher, the same or the next lower dose depending on the observed number of toxicities in the previous cohort or cohorts. Details can be found in Ting [5]. The 3+3 design is safe in the sense that only few patients experience DLTs or are treated with toxic doses. However, the probability of finding the actual MTD can be quite low and in most cases, the MTD is underestimated. Additionally the number of subjects needed gets large if the true MTD is much larger than the starting dose. These properties of the 3+3 design have been investigated in various simulation studies, among others in Gerke and Siedentop [4].

The purpose of this work is to improve the designs for dose escalation studies. In section 2 we will introduce the suggested methods, which we will apply in a simulation study presented in section 3. We will conclude with a discussion of the results.

[1]This work was supported by a PhD scholarship granted by Bayer Schering Pharma.

[2]Bayer Schering Pharma AG, Berlin and Otto-von-Guericke-University, Magdeburg, E-mail: katrin.roth@bayerhealthcare.com

2. Methods

2.1. General Idea

In order to combine the 3+3-Design with optimal design theory, some modifications are suggested. Given a class of models appropriate for the dose-response-relationship (e.g. logistic, proportional odds or E_{max}-model), the 3+3 design is conducted until parameter estimation is possible in the above class of models. Then a design that is optimal according to a specified optimality criterion and conditioned on the previous observations and estimated parameters is determined. The next cohort of patients is treated according to this design. The parameter estimation is repeated after each cohort and the design is adjusted. This procedure is repeated until a stopping rule is met.

This procedure is quite flexible since it can be applied to different underlying models, using various optimality criteria, flexible cohort sizes and several stopping rules.

A similar approach for adaptive D-optimal designs is suggested in Dragalin, Fedorov and Wu [2], but we will also allow for other optimality criteria and explicitly incorporate the 3+3 design as part of the method.

2.2. The Conditional Locally Optimal Design

To determine the optimal design at some stage of the trial, we will follow the common concepts of optimal design theory (cf. Fedorov and Hackl [3]). We will especially use notation based on information matrices. The optimal design at any stage of the trial is given by the design that maximizes the overall information of the experiment by allocating n additional subjects to doses within the design region $\mathcal{X}$.

Denote the vector of design points used so far by $\boldsymbol{x}_{obs}$, the estimated parameters for the model by $\hat{\boldsymbol{\theta}}$, the information matrix for design points $\boldsymbol{x}$ and parameters $\boldsymbol{\theta}$ by $\boldsymbol{M}(\boldsymbol{x}, \boldsymbol{\theta})$. Then the conditional information matrix for the observed design points $\boldsymbol{x}_{obs}$, estimated parameters $\hat{\boldsymbol{\theta}}$ and fixed cohort size n is given by

$$\boldsymbol{M}(\boldsymbol{x}_{obs}, \hat{\boldsymbol{\theta}}) + \sum_{i=1}^{n} \boldsymbol{M}(x_i, \hat{\boldsymbol{\theta}}), \quad x_i \in \mathcal{X}.$$

Maximizing a function of this matrix according to the chosen design criterion yields the conditional optimal design.

2.3. Variance and Confidence Intervals

Given the information matrix, we can derive results on the precision of the parameter estimates and we can construct confidence intervals.

Let $\boldsymbol{M}(\xi, \boldsymbol{\theta})$ denote the information matrix under a given model for a design ξ and parameter vector $\boldsymbol{\theta}$ (cf. Fedorov and Hackl [3]).

We will make use of the fact that

$$\mathrm{Cov}(\hat{\boldsymbol{\theta}}) \approx \boldsymbol{M}^{-1}(\xi, \hat{\boldsymbol{\theta}})$$

(cf. Cox and Hinkley [1]).

Assume that the dose-toxicity-relationship is described by a logistic model in the way that

$$P(DLT(x)) = \frac{\exp(\frac{x-\mu}{\sigma})}{1+\exp(\frac{x-\mu}{\sigma})}$$

with x being the dose. Let $\boldsymbol{\theta} = \begin{pmatrix} \mu \\ \sigma \end{pmatrix}$ and the MTD is the dose for which $P(DLT) = \frac{1}{3}$. Then

$$\mathrm{Var}(\widehat{MTD}) \approx \begin{pmatrix} 1 & -\log 2 \end{pmatrix} \boldsymbol{M}^{-1}(\xi, \hat{\boldsymbol{\theta}}) \begin{pmatrix} 1 \\ -\log 2 \end{pmatrix}.$$

Assuming an asymptotic normal distribution of $\log(MTD)$, an approximate $(1-\alpha)$ confidence interval for the MTD is given by

$$\left[\frac{\widehat{MTD}}{\exp\left(q_{1-\frac{\alpha}{2}} \frac{1}{\widehat{MTD}} \sqrt{\mathrm{Var}(\widehat{MTD})} \right)} \; , \; \widehat{MTD} \cdot \exp\left(q_{1-\frac{\alpha}{2}} \frac{1}{\widehat{MTD}} \sqrt{\mathrm{Var}(\widehat{MTD})} \right) \right]$$

with $q_{1-\frac{\alpha}{2}}$ being the $1-\frac{\alpha}{2}$ quantile of the normal distribution.

These results can be transferred easily to any other model with known information matrix.

3. Simulation Study

3.1. Settings

To compare the performance of the new approach with the traditional 3+3 design, a simulation study was conducted. We will only present some of the results in this paper. Originally six different dose toxicity scenarios were considered. Here we will focus on only one. A sequence of twelve doses, $d = (0.6, 1.2, 2, 3, 4, 5.3, 7, 9, 12.4, 16.53, 22, 29.4)$ is used for the simulation. All doses are given in mg. The underlying true dose toxicity relationship is given by a logistic model with parameters $\mu = 30$ and $\sigma = 7.67$, which gives a true MTD of 24.68 mg. Thus the goal of any adaptive procedure would be to pick the dose 22 mg as the MTD.

Within this setting, the course of a study performed with the 3+3 design was simulated 100 000 times.

The modified 3+3 design was simulated with the following settings. A logistic model and a proportional odds model with four categories were considered as the model for analysis. Parameter estimation was done using the maximum likelihood method. The stopping rule was defined by a maximum sample size that was given by the median sample size needed in the simulated 3+3 design. The cohort size was one. The optimality criteria considered were c- and D-optimality, where the c-criterion was specified to maximize the precision of the estimated MTD. If we

denote the conditional information matrix as defined above by $\boldsymbol{M}$, then $\det\boldsymbol{M}$ was maximized in case of the D-criterion, and $\boldsymbol{c}'\boldsymbol{M}^{-1}\boldsymbol{c}$ was minimized in case the c-criterion was considered, where $\boldsymbol{c}$ depends on the underlying model. Using the logistic model,

$$\boldsymbol{c} = \begin{pmatrix} 1 \\ -\log(2) \end{pmatrix},$$

whereas with the proportional odds model

$$\boldsymbol{c} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ -\log(2) \end{pmatrix}.$$

Two different settings for the design region were investigated. In both cases, the lower boundary of the design region was 0, which is a natural choice, whereas the upper boundary was chosen adaptively. In the first case ('design region 1'), the upper boundary was the dose out of the sequence that was one step above the maximum dose used so far. In the second case ('design region 2'), the design region was bounded by the dose above the currently estimated MTD. However, in any case the upper bound did never exceed the prespecified range of doses. These boundaries were introduced as a safety precaution. In total, eight different settings were investigated, the ones based on the logistic model with 100 000 simulation runs, the ones based on the proportional odds model with 10 000 simulation runs.

3.2. Results

The key results are displayed in Tables 1 and 2. The setting for the modified 3+3 design is described by two letters and a number being abbreviations for the underlying model ('L' for the logistic and 'P' for the proportional odds model), the optimality criterion ('D' or 'c') and the design region ('1' or '2', as defined above).

In Table 1, the percentage of correctly estimated MTDs, slightly over- and underestimated MTD and strongly underestimated MTDs are shown, as well as the percentage of simulation runs where the method failed to give an estimate of the MTD. The latter can be the case if the 3+3 design stops at the initial dose level due to several observed DLTs while the MLE in the considered underlying model does not exist. It can be seen that the modified 3+3 design with underlying logistic model performs slightly better than the traditional 3+3 design for one of the design regions, whereas the modified 3+3 design with underlying proportional odds model uniformly performs better. This can also be seen from the reduction of the MSE presented in Table 2

In Table 2 some more aspects are displayed. The average sample size does not differ much between the methods, because they were set to be comparable (c.f. Section 3.1.). Since also the safety of the procedure is of major interest, we looked at the average number of DLTs observed during the course of a study as well as the number of patients treated with doses above the MTD. Here it becomes

Table 1: Results of the Simulation Study

Dose	$\leq$ 12.4 mg	16.53 mg	22 mg	29.4 mg	no MTD
3+3-Design	33.99	29.03	28.92	7.56	0.5
mod.3+3(L,D,1)	28.18	25.85	33.49	11.98	0.50
mod.3+3(L,D,2)	51.54	24.67	18.10	5.19	0.50
mod.3+3(L,c,1)	28.71	26.21	32.54	12.02	0.52
mod.3+3(L,c,2)	51.56	24.66	18.11	5.18	0.49
mod.3+3(P,D,1)	10.28	24.91	49.18	15.25	0.38
mod.3+3(P,D,2)	11.70	26.10	47.72	14.10	0.38
mod.3+3(P,c,1)	9.94	23.74	50.53	15.41	0.38
mod.3+3(P,c,2)	11.63	25.64	48.26	14.09	0.38

Table 2: Results of the Simulation Study (cont.)

	sample size (mean)	DLTs (mean)	patients treated above MTD (mean)	MSE of est. MTD
3+3-Design	38.43	**3.44**	**1.61**	109.16
mod.3+3(L,D,1)	35.35	3.81	3.86	99.59
mod.3+3(L,D,2)	35.09	3.75	3.36	145.15
mod.3+3(L,c,1)	35.25	3.69	3.51	100.98
mod.3+3(L,c,2)	35.09	3.76	3.37	145.07
mod.3+3(P,D,1)	37.94	5.61	6.18	55.85
mod.3+3(P,D,2)	37.94	5.53	5.40	58.58
mod.3+3(P,c,1)	37.94	5.81	6.90	54.65
mod.3+3(P,c,2)	37.94	5.63	5.75	58.18

obvious that the higher percentage of correctly estimated MTDs comes at the cost of treating more patients at toxic doses and thus having more patients experiencing DLTs. However this moderately increased risk for the patients can be justified by the potential of reducing the MSE of the estimated MTD by approximately 50 per cent and the possibility of making statements about the precision of the estimates.

For all the settings of the modified 3+3 design, confidence intervals for the MTD were calculated and the ratios of the upper and lower 95% confidence limits were considered. The median ratio varied only slightly and was between 1.94 and 1.96 for the eight different settings. This shows that the precision of the estimated MTD is not influenced much by the different settings, but taking also a look at the MSEs leads to the conclusion that the bias differs notably.

4. Discussion

The results of the simulation study show that the modified 3+3 design has the potential to perform better than the traditional 3+3 design, especially when it is applied with the more complex proportional odds model. As opposed to the traditional 3+3 design, the parametric approach allows for meaningful conclusions on the precision of the estimates. However, we have to carefully take into consideration the increased risk for the patients. It could still be investigated if this risk can be reduced by choosing the design regions differently or by applying a different optimality criterion. It might also be worth taking a look at different models for the analysis of the dose toxicity relationship, e.g. the E_{max}-model.

The parametric modification of the 3+3 design thus is a promising alternative to the traditional 3+3 design that might be worth investigating further.

Acknowledgement The author is very grateful to Rainer Schwabe (University of Magdeburg) and Thomas Schmelter (Bayer Schering Pharma AG) for their support.

References

[1] Cox D.R., Hinkley D.V. (2000) *Theoretical Statistics.* Chapman & Hall / CRC, Boca Raton.

[2] Dragalin V., Fedorov V.V., Wu Y. (2006) *Optimal Designs for Bivariate Probit Model.* Technical Report, GlaxoSmithKline Pharmaceuticals, Collegeville, PA.

[3] Fedorov V.V., Hackl P. (1997) *Model-Oriented Design of Experiments,* vol. 125 of *Lecture Notes in Statistics.* Springer, New York.

[4] Gerke O., Siedentop H. (2007) Optimal phase I dose-escalation trial designs in oncology - A simulation study. Statistics in Medicine 27, 5329-5344

[5] Ting N. (Editor) (2006) *Dose Finding in Drug Development.* Springer, New York.

6th St.Petersburg Workshop on Simulation (2009) 1054-1058

Adaptive Covariate Adjusted Designs for Clinical Trials that Seek Balance

Anthony C. Atkinson[1]

Abstract

A new "doubly-adaptive" rule is introduced for adaptive treatment allocation in randomized sequential clinical trials with normal responses and covariate adjustment. The rule is shown to have excellent properties for repairing damaged designs.

1. Introduction

The paper introduces a new treatment allocation rule for adaptive clinical trials with normal responses. Simulations show that the rule has good properties for repairing designs that have a history of biased allocations.

In the trial, patients arrive sequentially and are to be allocated one of t treatments. Adjustment for covariates is by least squares regression. The inferential aim of the trial is to obtain treatment estimates with minimum variance. The ethical aim of the trial is to ensure that as few patients as possible receive inferior treatments. We formalize this by seeking to allocate given proportions $p^* = (p_1^*, \ldots, p_t^*)^T$ of patients to the ranked treatments, with the best treatment receiving the largest allocation. Of course, the ranking of the treatments is initially unknown and has to be estimated.

The sequential allocation of treatments using optimum design theory yields a deterministic allocation rule from which the allocation can be guessed, with the risk of bias. We therefore employ biased-coin rules that bring some randomization into treatment allocation. In the usage of Chapter 9 of Hu and Rosenberger (2006) our rule is a covariate-adjusted response-adaptive (CARA) randomization procedure.

2. Background

Models. The vector of t unknown treatment effects is α and the patient presents with a vector x_i of covariates. The data, perhaps after data transformation, are analysed using the regression model

$$E(y_i) = f_i^T \beta = h_i^T \alpha + z_i^T \theta, \tag{1}$$

[1]London School of Economics, E-mail: a.c.atkinson@lse.ac.uk

with additive independent errors of constant variance. Here h_i is a vector of t indicator variables, the one non-zero element indicating which treatment the patient received. The $v \times 1$ vector z_i contains those covariates, including any powers or interactions of the elements of x_i, which will be used to adjust the responses when estimating α.

The model (1) for n patients in matrix form is

$$E(Y_n) = F_n\beta = H_n\alpha + Z_n\theta, \tag{2}$$

where Y_n is the $n \times 1$ vector of responses for the n patients. As the effects of the variables z_i are not of interest in themselves, the parameters θ in (1) and (2) become nuisance parameters.

Rather than all elements of β our interest will be in only a single linear combination

$$a^T\beta = l^T\alpha + w^T\theta. \tag{3}$$

As θ is a nuisance parameter, the v elements of w are zero. Since we estimate only one linear combination of the $t+v$ parameters, the nuisance parameters will be in a space of dimension $q = t + v - 1$.

The variance of the estimated combination of coefficients is

$$\text{var}\,\{a^T\widehat{\alpha}\} = \sigma^2 a^T(F_n^TF_n)^{-1}a, \tag{4}$$

where $\hat{\alpha}$ is the least squares estimate of α and σ^2 is the variance of the errors, assumed additive in (1). Minimisation of this variance is central to our adaptive designs for clinical trials.

Deterministic Sequential Design: Rule D. In optimum design theory (Silvey 1980) designs minimising the variance (4) are a special case of D_A-optimality. In sequential trials the extended design matrix F_n is known. Patient $n+1$ arrives with a vector of covariates x_{n+1}, a function of which forms the last row of Z_{n+1} in (2). If the vector of allocation and prognostic factors for the $(n+1)$st patient when treatment j is allocated is $f_{n+1,j}$, $F_{n+1,j}$ is formed by adding the row $f_{n+1,j}^T$ to F_n.

In the sequential construction of D_A-optimum designs minimising (4) Atkinson (1982) shows we allocate treatment j for which

$$\begin{aligned} d_A(j, n, x_{n+1}) &= f_{n+1,j}^T(F_n^TF_n)^{-1}A\{A^T(F_n^TF_n)^{-1}A\}^{-1}A^T(F_n^TF_n)^{-1}f_{n+1,j}. \\ &\quad (j = 1, \ldots, t), \end{aligned} \tag{5}$$

is a maximum. Here x_{n+1} is the vector of covariates for the new patient that are included in the vector $f_{n+1,j}$. In this general formulation, $A^{(t+v)\times s}$ is a matrix of s linear combinations that are of interest. In our case $s = 1$ and $A = a$.

Randomization and Skewing. There is no randomness in such an allocation rule. Atkinson (2002) compares a number of "biased-coin" rules that introduce randomness, but aim for equal allocation over all treatments. For example, in his Rule A the probability of allocating treatment j is proportional to (5).

Unequal, or "skewed", allocations are obtained through choice of the vector l. To obtain such allocations combined with efficient parameter estimation we find

designs for estimation of the linear combination with

$$l^T\alpha = \pm p_1\alpha_1 \mp \ldots \pm p_t\alpha_t, \tag{6}$$

where the coefficients $p_j, j = 1, \ldots, t$ are such that $0 < p_j < 1$ and $\sum p_j = 1$. It is straightforward to show that the variance of $l^T\hat{\alpha}$, in the absence of covariates, is minimised when the proportion of patients receiving treatment j, is p_j, as it is when the design is balanced across treatments.

Adaptive Designs. There is a large literature on response-adaptive designs for binomial responses. For the normal responses of interest here, the problem is how to convert the observed treatment effects $\hat{\alpha}_1$ into probabilities. Bandyopadhyay and Biswas (2001) use a normal link function, in which the degree of skewing depends on the treatment differences. Here we follow Atkinson and Biswas (2009) and find designs for given target allocations p_j^* that depend on the ranking $R(j)$ of treatment j. For an operational rule we use $\widehat{R}(j)$ the estimated rank of treatment j, so that $p_j = p^*_{\widehat{R}(j)}$.

Regularisation. There is a risk that, early in the trial, an adaptive design may select a non-optimum treatment as best and thereafter allocate only to that treatment. To avoid this, when simulating adaptive designs we regularise to ensure that each treatment is allocated throughout the trial, although with a decreasing frequency for non-optimum treatments. The exception is when all treatments are the same. With two treatments five of the first ten patients are allocated to each treatment. Thereafter, if the number allocated to either treatment is below $\sqrt{n}$, that treatment is allocated when n is an integer squared. For an 800 trial design the first regularisation could occur when $n = 16$ and the last when $n = 784$. The dependence on $\sqrt{n}$ is arbitrary, but it is desirable for full efficiency to use a rule that asymptotically vanishes.

3. Comparisons

1. Random Allocation Rule R. Treatment j is allocated with probability $p^*_{\widehat{R}(j)}$.

2. Rule G. Treatment j is allocated with probability

$$\pi(j|x_{n+1}) = \frac{\{1 + d_A(j, n, x_{n+1})\}^{1/\gamma} p^*_{\widehat{R}(j)}}{\sum_{s=1}^{t} \{1 + d_A(s, n, x_{n+1})\}^{1/\gamma} p^*_{\widehat{R}(s)}}. \tag{7}$$

For small n and γ Rule G behaves like Rule D, becoming increasingly like Rule R for larger values of γ and as n increases.

3. "Doubly-adaptive" Rule H. Hu and Zhang (2004) investigate the properties of a doubly-adaptive allocation rule that ignores covariates. For our purposes let $r_{j,n}$ be the proportion of allocations to treatment j. Then with $b_j = r_{j,n}$ and $c_j = p^*_{\widehat{R}(j)}$, the probability of allocating treatment j is

$$\pi_H(j, n+1) = \frac{c_j(c_j/b_j)^\delta}{\sum_{s=1}^{t} c_s(c_s/b_s)^\delta}. \tag{8}$$

In (8) δ is a non-negative constant which determines the strength of forcing balance between the allocated proportions and the targets. As δ increases the probability of allocation to give this balance becomes larger.

4. Forcing Rule Z. The new rule combines the covariate adaptivity of Rule G with the strong targeting of the p_j^* from Rule H by replacing $p^*_{\widehat{R}(j)}$ in (7) by $\pi_H(j, n+1)$ in (8). We obtain

$$\pi_Z(j|x_{n+1}) = \frac{\{1 + d_A(j, n, x_{n+1})\}^{1/\gamma} c_j (c_j/b_j)^\delta}{\sum_{s=1}^{t} \{1 + d_A(s, n, x_{n+1})\}^{1/\gamma} c_s (c_s/b_s)^\delta}, \tag{9}$$

since the summation in the denominator of (8) cancels. We thus obtain a series of allocation rules with two variable parameters: δ and γ.

Some general properties of these five rules are clear. Rules H and R do not respond to the covariate pattern. Rule D is unrandomized. Only Rules G and Z both respond to the covariate pattern and target a given probability.

Loss and the Assessment of Designs. Let the treatments be correctly ranked. Then the linear combination of the parameters corresponding to the proportions p_j^* is $l_*^T \alpha$. From (4) the variance of the estimated linear combination has a minimum value of var $\{l_*^T \hat{\alpha}_*\} = \sigma^2/n$, where $\hat{\alpha}_*$ is the estimate from the optimum design with treatment proportions p_j^* when there is balance over the covariates.

For other designs we find the variance of the same linear combination from (4). Comparisons can use either the ratio of variances, that is the efficiency E_n, or the loss (Burman 1996), calculated by Atkinson (2002) for eleven rules for unskewed treatment allocation. The loss L_n is defined by writing the variance (4) as

$$\text{var}\,\{l_*^T \widehat{\alpha}\} = \sigma^2/(n - L_n). \tag{10}$$

With a random element in treatment allocation, the loss L_n is a random variable, the value of which depends upon the particular trial and pattern of covariates. Let $E(L_n) = \mathcal{L}_n$. For random allocation of two treatments $\mathcal{L}_n \to q$, the number of nuisance parameters, as n increases. Other designs that force more balance have lower values of $\mathcal{L}_n$. The loss can be interpreted as the number of patients on whom information is lost due to the lack of optimality of the design.

We begin with part of the simulation results from comparing Rules G and Z for skewing the allocation. The assumption is that the ordering of the treatments is known. As a consequence $\widehat{R}(j)$ is replaced by $R(j)$ in the definition of the rules. The designs are thus not response adaptive.

The left-hand panel of Figure 1 shows the average loss $\bar{L}_n$ for n up to 200. We have taken $p_1^* = 0.8$ and $\gamma = 0.1$. The value of p_1^* has little effect on the value of loss. The three curves of loss for Rule Z are for three very different values of δ, 0.1, 1 and 10. As δ increases the rule forces closer balance to the target proportion of allocations to the treatments. As can be seen, this has virtually no effect on the average losses, which decrease to values of around 1.3 at $n = 200$. The average loss for Rule G has a rather different shape for this value of γ. Initially Rule G behaves like deterministic allocation, forcing balance over the covariates and the

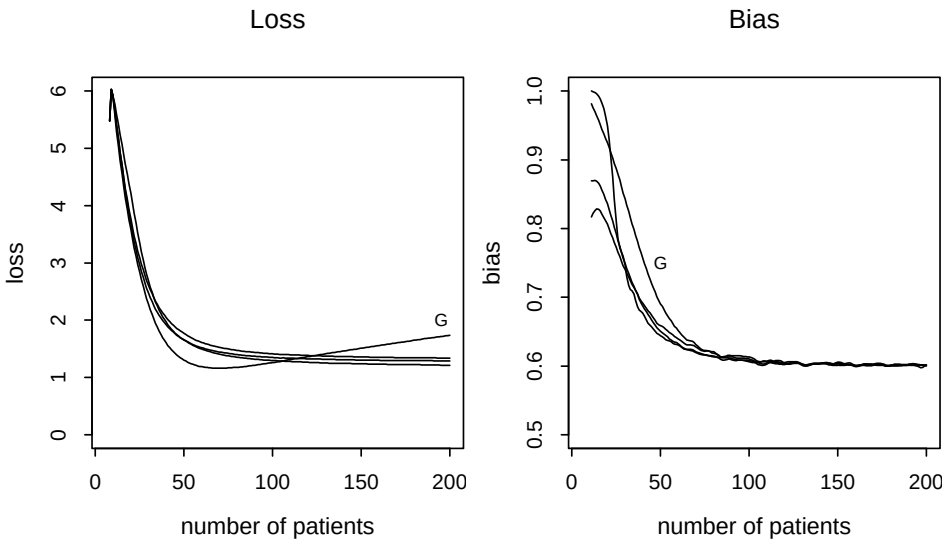


Figure 1: Rules G and Z ($\delta = 10$, 1 and 0.1) skewing. Left-hand panel: average loss $\bar{L}_n$. Right-hand panel: smoothed average bias $\bar{B}_n$; $p_1^* = 0.8$, $q = 5$, $\gamma = 0.01$, 100,000 simulations.

loss decreases steeply. However, as n increases, the rule becomes more like random allocation and so the loss increases tending towards q, here five.

Bias. Selection bias occurs when the clinician is able correctly to guess the next treatment to be allocated. For two treatments it can be estimated from n_{sim} simulations by $\bar{B}_n =$ (number of correct guesses of allocation to patient n − number of incorrect guesses)$/n_{\text{sim}}$. The non-randomized sequential construction of optimum designs gives a value of one for B_n. For random allocation with $p_1^* = 0.5$ the value is 0. For skewed random allocation the value for Rule R when guessing the treatment more likely to be allocated is $2p_1^* - 1$. With $p_1^* = 0.8$, this value is 0.6.

The right-hand panel of Figure 1 shows the values of $\bar{B}_n$ for Rules G and Z. For Rule Z and small n the three curves reading downwards are for $\delta = 10$, 1 and 0.1. For the largest value of δ allocation is forced to balance the design and guessing the under-represented treatment has a high probability of success. As n increases, the designs become more balanced for all three values of δ and the bias tends to the asymptotic value of 0.6. The bias for Rule G is also high at the beginning, decreasing slightly more slowly than that for Rules Z.

Recovery from Misallocation. To demonstrate the ability of the various rules to rebalance the trial, suppose that, up to $n = 100$, treatment 2 is allocated whenever the sum of the four explanatory variables is greater than 2. Thereafter the rules are applied correctly. The average losses $\bar{L}_{100}$ & $\bar{L}_{200}$ are in Table 1. The results show that Rule R is least affected by this lack of balance, followed by Rule Z. Although its value of $\bar{L}_{100}$ is only slightly less than that of those for the other three rules, Rule Z provides a much smaller value of $\bar{L}_{200}$.

Table 1: Recovery from Misallocation: Average losses at $n = 100$ and 200 and average bias at $n = 200$

Criterion	$\bar{L}_{100}$	$\bar{L}_{200}$	$\bar{B}_{200}$
R	22.46	15.44	0.608
G	18.82	12.05	0.619
H	20.92	12.60	0.655
Z	15.13	3.41	0.628
D	9.32	0.52	1.00

The last column of the Table gives the values of average bias $\bar{B}_{200}$, calculated assuming that the rule for excess allocation to treatment 2 is known. Rule D is deterministic and so the bias is one. Of the other rules, all values of bias are tending towards 0.6. This satisfactory bias, combined with the small value of loss, suggests that Rule Z should be seriously considered for use in trials of the kind discussed here.

References

[1] Atkinson, A. C. (1982). Optimum biased coin designs for sequential clinical trials with prognostic factors. *Biometrika 69*, 61–67.

[2] Atkinson, A. C. (2002). The comparison of designs for sequential clinical trials with covariate information. *Journal of the Royal Statistical Society, Series A 165*, 349–373.

[3] Atkinson, A. C. and A. Biswas (2009). Adaptive designs for clinical trials that maximize utility. (Submitted.).

[4] Bandyopadhyay, U. and A. Biswas (2001). Adaptive designs for normal responses with prognostic factors. *Biometrika 88*, 409–419.

[5] Burman, C.-F. (1996). *On Sequential Treatment Allocations in Clinical Trials*. Göteborg: Department of Mathematics.

[6] Hu, F. and W. F. Rosenberger (2006). *The Theory of Response-Adaptive Randomization in Clinical Trials*. New York: Wiley.

[7] Hu, F. and L.-X. Zhang (2004). Asymptotic properties of doubly adaptive biased coin designs for multitreatment clinical trials. *Annals of Statistics 32*, 268–301.

[8] Silvey, S. D. (1980). *Optimum Design*. London: Chapman and Hall.

6th St.Petersburg Workshop on Simulation (2009) 1060-1064

On the functional approach to studying optimal Bayesian design[1]

Viatcheslav B. Melas[2]

Abstract

In this paper we discuss opportunities of the functional approach developed in Melas (2006) to studying Bayesian efficient design for nonlinear regression models.

1. Introduction

The optimal design concept has been rapidly developed in the last fifty years. Theoretical results are well documented in the book Pukelsheim (2006). For a more practical guide we refer to the recent monograph Atkinson, Donev and Tobias (2007). In these and other books and papers optimal designs are determined as discrete probability measures giving an extremal value to a functional of the information matrix that assures some attractive statistical properties of the designs.

In order to construct such a measure two main approaches are typically applied. One approach consists of finding the measure explicitly or reducing the problem to a classical mathematical problem. For example, the support points of D-optimal designs for one dimensional polynomial models are found to be the roots of some orthogonal polynomials. Another approach is merely numerical evaluation of the designs. However, both approaches have some limitations: the analytical approach is rarely available and the numerical one does not allow sufficiently full study of the design's properties.

A different approach that can be a promising alternative to the approaches described above was called "the functional approach". It goes back to the paper Melas (1978) and has been thoroughly elaborated in the monograph Melas (2006). The idea of this approach consists of considering optimal design support points as implicit functions of certain auxiliary parameters. In Dette, Melas and Pepelyshev (2004) some general recurrence formulae were introduced for expanding such functions into Taylor series. This approach allows us not only to construct optimal designs but also to study their dependence on the auxiliary parameters.

In the present paper we give an outline of this approach for Bayesian optimal designs in nonlinear regression models following Melas, Staroselsky (2008).

[1]This work was partly supported by RFBR, project No 09-01-00508.

[2]St.Petersburg State University, Faculty of Mathematics & Mechanics, Russia. E-mail: v.melas@pochta.tvoe.tv

2. Statement of the problem

Let the experimental results be described by the nonlinear regression model:

$$y(x_j) = \eta(x_j, \theta) + \varepsilon_j, \quad j = 1, \ldots, N,$$

where $x_1, \ldots, x_N$ are the experimental conditions for the N observations, $\theta = (\theta_1, \ldots, \theta_m)^T$ are unknown parameters, $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_N$ are independent identically distributed random values with zero mean and unknown common variance $\sigma^2 > 0$. In many practical applications (see, e.g., Seber, Wild, 1989) it can be assumed that $x_1, \ldots, x_N \in [d_1, d_2]$, a design interval. A discrete probability measure given by the table

$$\xi = \begin{pmatrix} x_1 & \ldots & x_n \\ w_1 & \ldots & w_n \end{pmatrix},$$

where $x_i \neq x_j$ $(i \neq j)$, $w_i > 0$, $\sum_{i=1}^{n} w_i = 1$ will, as usual, be called an (approximate) experimental design, $x_1, \ldots, x_n$ are called support points and $w_1, \ldots, w_n$ are weight coefficients.

Let r_i observations be performed in points x_i, $r_i = \lfloor N w_i \rfloor + \delta$, $\delta = 0$ or 1, so that $\sum r_i = N$. Denote by $\widehat{\theta}(N)$ the least squares estimate of θ and let θ^* denote the true value of θ. Note that under the known regularity conditions (see Jennrich, 1969) the value $\sqrt{N}(\widehat{\theta}(N) - \theta^*)$ with $N \to \infty$ is asymptotically normally distributed with zero mean and covariance matrix $\sigma^2 M^{-1}(\xi, \theta)$, where $\theta = \theta^*$,

$$M(\xi, \theta) = \left(\sum_{s=1}^{n} f_i(x_s, \theta) f_j(x_s, \theta) w_s \right)_{i,j=1}^{m}$$

— the Fisher information matrix — and

$$f_i(x, \theta) = \frac{\partial}{\partial \theta_i} \eta(x, \theta), \ i = 1, \ldots, m.$$

In order to avoid the dependence of the design criteria on unknown parameters we will apply the Bayesian approach (see Chaloner, Larntz, 1989 or Dette, Braess, 2007 for a detailed explanation of the approach).

Following Melas, Staroselsky (2008) we will consider designs minimizing the value

$$\int \left(\frac{\det M(\xi^*(\theta), \theta)}{\det M(\xi, \theta)} \right)^{1/m} dP(\theta),$$

where $\xi^*(\theta)$ is a locally D-optimal design, that is a design maximizing $\det M(\xi, \theta)$ under the fixed value θ, $P(\theta)$ is a prior distribution for unknown parameters. Such designs will be called D-efficient Bayesian designs.

The value shows how many times more observations will be needed for a given design with respect to a design that is locally D-optimal (for the true value of parameters) in order to estimate the parameters with the same accuracy.

In the next section we describe how the functional approach can be applied for the construction and study of D-efficient Bayesian designs.

3. The functional approach

Remember that a real function is called real analytic on a set if in a vicinity of each point of this set it can be expanded into convergent Taylor series.

For any (scalar, vector or matrix) function $\psi(z)$ and a point $z_0 \in R^1$ write

$$\psi_{(0)} = \psi(z_0),\ \psi_{(n)} = \frac{1}{n!}\psi^{(n)}(z_0),\ \psi_{<n>}(z) = \sum l_{s=0}^{n}\psi_{(s)}(z - z_0)^s,\ n = 1, 2, \ldots.$$

We will need the following auxiliary result.

Theorem 1. *Let H be an open set in R^{m-1}, $\tau \in H$, $z \in [z_1, z_2]$, $z_1 < z_2$ be arbitrary real numbers, $\varphi(\tau, z)$ be a real analytical function on the set $H \times [z_1, z_2]$. Further, for notational simplicity, let*

$$g(\tau, z) = \left(\frac{\partial}{\partial\tau_i}\varphi(\tau, z)\right)_{i=1}^{m-1},$$

$$V(\tau, z) = \left(\frac{\partial^2}{\partial\tau_i\partial\tau_j}\varphi(\tau, z)\right)_{i,j=1}^{m-1}.$$

Under these conditions if, for a point $(\tau_{(0)}, z_0) \in H \times [z_1, z_2]$ the equality

$$g(\tau_{(0)}, z_0) = 0$$

holds and $\det V(\tau_{(0)}, z_0) \neq 0$, then:

I *In a vicinity of z_0 there exists a unique continuous function $\widehat{\tau}(z)$ such that $\widehat{\tau}(z_0) = \tau_0$ and $g(\widehat{\tau}(z), z) = 0$. This function is real analytic in a vicinity of z_0;*

II *The following recurrence formulae take place*

$$\tau_{(n+1)} = -[V^{-1}](g(\tau_{<n>}(z), z))_{(n+1)},\ n = 0, 1, \ldots,$$

where $V = V(\tau_{(0)}, z_0)$.

Part (I) of the Theorem 1 is the well known Implicit Function Theorem (see, e.g., Gunning, Rossi, 1965). Part (II) is proved in Melas (2006, Ch.2).

For applying this theorem to studying Bayesian efficient designs let us consider a special class of designs. Designs with the number of support points equal to the number of parameters ($n = m$) and equal weight coefficients will be called saturated designs. Locally D-optimal designs are often saturated designs (see Melas, 2006). For this reason the class of saturated designs is of particular interest. Moreover, usually one of the support points coincides with a bound. Assume that

$$d_1 < x_1 < x_2 < \ldots < x_m = d_2.$$

The case $x_1 = d_1$ can be considered in a similar way. Denote by τ the vector of design points not coinciding with a bound,

$$\tau = (x_1, \ldots, x_{m-1})$$

when the corresponding saturated design will be denoted as

$$\xi_\tau = \begin{pmatrix} x_1 & \dots & x_{m-1} & d_2 \\ 1/m & \dots & 1/m & 1/m \end{pmatrix}.$$

Let the regression function be of the form

$$\eta(x,\theta) = \sum_{i=1}^{k} a_i \eta_i(x,b),\ a_i \neq 0,\ i = 1,\dots,k, \tag{1}$$

where $b = (b_1,\dots,b_{m-k})^T,\ \theta = (a_1,\dots,a_k,b_1,\dots,b_{m-k})^T,\ \eta_i(x,b),\ i = 1,\dots,k$ are given functions.

Many popular models have the form (1) and it is easy to check that locally D-optimal designs do not depend on parameters $a_1,\dots,a_k$ linearly included in the model (see Melas, 2006). It is evident that Bayesian D-efficient designs also do not depend on $a_1,\dots,a_k$ and without loss of generality we can assume that $a_1 = a_2 = \dots = a_k = 1$. Let us now re-denote the information matrix as

$$M(\xi,b) = M(\xi,\theta),\ \theta = (1,\dots,1,b_1,\dots,b_{m-k})^T.$$

Assume also that all parameters b_i are positive

$$0 < b_i < \infty,\ i = 1,\dots,m-k.$$

This assumption holds true for many practically important nonlinear models. Besides, it can be seen that it is not a strict restriction on our methodology. Introduce the set

$$\Omega = \Omega_z = \Omega_z(c) = \{b = (b_1\dots,b_{m-k})^T : (1-z)c_i \le b_i \le c_i/(1-z),\ i = 1,\dots,m-k\},$$

where the point $c = (c_1,\dots,c_{m-k})^T$ can be considered as an approximation for the vector b^*, and z can be considered as the value of relative error. With $z = 0$ we have $\Omega_z(c) = \{c\}$, and with $z = 1$ we have $\Omega_z = [0,\infty]^{m-k}$.

Let $P_z(b)$ be the uniform distribution on Ω_z,

$$\varphi(\tau,0) = \left(\frac{\det M(\xi_{\tau^*(c)},c)}{\det M(\xi_\tau,c)}\right)^{1/m},$$

$$\varphi(\tau,z) = \int \left(\frac{\det M(\xi_{\tau^*(b)},b)}{\det M(\xi_\tau,b)}\right)^{1/m} dP_z(b), \quad z \in [0,1], \tag{2}$$

$$\varphi(\tau,-z) = \varphi(\tau,z),\ z \in (-1,0],$$

where $\xi_{\tau^*(b)}$ is a locally D-optimal design in the class of saturated designs with $x_m = d_z$. Such a choice of the prior distribution seems to be rather natural. At the same time, the approach can be applied to different prior distributions, for example, for the beta distribution.

Under some natural conditions it can be proved (see Melas, Staroselsky, 2008) that for the function (2) Theorem 1 holds and with $\tau_{(0)} = \tau^*(c)$, $0 \in (z_1,z_2)$. the

design $\xi_{\widehat{\tau}_{(z)}}$ is a Bayesian D-efficient design for the uniform distribution on Ω_z for a sufficiently small positive z.

Let z^* be the maximal value of z such that the design $\xi_{\widehat{\tau}(z)}$ is a D-efficient Bayesian design (in the class of all designs, not necessarily saturated) for the uniform distribution on Ω_z. The value z^* can be estimated numerically on the basis of the following theorem that is an obvious analog of the well known Kiefer-Wolfowitz equivalence theorem (see Pukelsheim, 2006).

Theorem 2. *Let the regression equation be of the form (1), where the function η is continuously differentiable w.r.t. the parameters. Then a design ξ is a D-efficient Bayesian design for an arbitrary distribution $P(\theta)$ if and only if*

$$\int (f^T(x,\theta)M^{-1}(\xi,\theta)f(x,\theta) - m)\left(\frac{detM(\xi^*(\theta),\theta)}{detM(\xi,\theta)}\right)^{1/m} dP(\theta) \le 0$$

for arbitrary $x \in [d_1, d_2]$, where $f(x,\theta) = \left(\frac{\partial\eta(x,\theta)}{\partial\theta_1}, \ldots, \frac{\partial\eta(x,\theta)}{\partial\theta_m}\right)^T$.

4. Example: A rational model

Let us consider the model given by the equation:

$$E(y|x) = \frac{\theta_0 x(\theta_1 + x)}{\theta_2 + \theta_3 x + x^2},\ x \in [0, T].$$

It can be checked that the Bayesian efficient designs do not depend on the parameters θ_0 and θ_1. Thus, setting $\theta_0 = \theta_1 = 1$ we will take $\Omega(z)$ of the form:

$$\Omega = \Omega(z) = \{b = (\theta_2, \theta_3) = (b_1, b_2); (1-z)c_i \le b_i \le \frac{1}{1-z}c_i,\ i = 1, 2\}.$$

Taking $c = (0.079, 5.05)$ we construct Taylor expansions of the support points of Bayesian designs that allow us to find values of these points with a high precision for given z. Certainly, any other initial vector c can be chosen.

Using Theorem 2 it is easy to determine the value z^* numerically. In the case considered $z^* \approx 0.305$. Thus we will set $z = z^*$ and compare the D-efficient Bayesian design for $z = 0.305$:

$$\begin{pmatrix} 0.013 & 0.226 & 2.86 & 10 \\ 1/4 & 1/4 & 1/4 & 1/4 \end{pmatrix}$$

and the most efficient equidistant design (constructed numerically):

$$\begin{pmatrix} 0.01 & 0.8325 & 1.675 & \ldots & 10 \\ 1/13 & 1/13 & 1/13 & \ldots & 1/13 \end{pmatrix}.$$

From Figure 1 we can see that the Bayesian designs are much more efficient than the best equidistant designs.

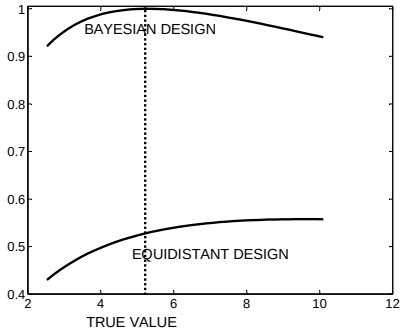


Figure 1: The rational model. Dependency of efficiency of Bayesian and equidistant designs on the parameter θ_3

References

[1] Atkinson A. C., Donev A. N., and R. D. Tobias, 2007. *Optimum Experimental Designs, with SAS*. Oxford University Press, Oxford.

[2] Dette H., Braess D., 2007. On the number of support points of maximin and Bayesian D-optimal designs in nonlinear regression models. *Annals of Statistics* 35(2), 772-792.

[3] Dette H., Melas V., and A. Pepelyshev, 2004. Optimal designs for estimating individual coefficients in polynomial regression – a functional approach. *J. of Stat. Planning and Inference*, 118, 2004, 201-219.

[4] Chaloner K. and Larntz K., 1989. Optimal Bayesian design applied to logistic regression experiments, *J. Statist. Plann. Inf.* 21, 191-208.

[5] Gunning R.C., Rossi H., 1965. *Analytical Functions of Several Complex Variables*. Prentice-Hall, New York.

[6] Jennrich R.I., 1969. Asymptotic properties of non-linear least squares estimators. *Ann. Math. Statist.* 40, 633-643.

[7] Melas V.B., 1978. Optimal designs for exponential regression. *Mathematische Operationsforschung und Statistik*, Ser. Statistics, 9, 45-59.

[8] Melas V.B., 2006. *Functional Approach to Optimal Experimental Design*. Lecture Notes in Statistics, 184. Springer Science+Business Media, Inc., Heidelberg.

[9] Melas V.B. and Yu. Staroselsky, 2008. D-efficient Bayesian designs for a class of nonlinear models. *J. of Statistical Theory and Practice*, 2 (4), 568–587.

[10] Pukelsheim F., 2006. *Optimal Design of Experiments*. Wiley, New York.

[11] Seber G.A.F., Wild C.J., 1989. Nonlinear regression. Wiley, New York.

6th St.Petersburg Workshop on Simulation (2009) 1066-1070

Sparse Sampling D-Optimal Designs in Quadratic Regression With Random Effects[1]

Tobias Mielke[2]

Abstract

In mixed effect models the variability of the regression parameters has substantial influence on the choice of the optimal design. If less observations per individual are possible than parameters are to be estimated, the optimality results of single-group designs no longer hold.

1. Introduction

In population pharmacokinetic studies the blood samples of individuals are evaluated together in one model, assuming that the same regression function can be used for all subjects, with slightly different parameters for the different individuals. These differences from the population mean are modeled by random variables. The purpose of this article is to study the design for quadratic regression in mixed effect models, in the case of two allowed observations per individual. Taking many blood samples of one individual is costly, unethical and in some cases even not possible. If less observations are being made than parameters are to be estimated, the occuring D-optimal individual design will lead to a singular information matrix. The use of population designs with different observation groups helps to construct estimates for the population parameter vector.
Cheng[2] and Atkins and Cheng[1] provide D-optimal designs for quadratic regression with random intercept, considering two observations per subject. In this article we generalize these results to polynomial regression with random slope and random curvature.
In section 2 we introduce the mixed effects model. In section 3, the structure of the D-optimal information matrix for quadratic regression with two observations per individual and the D-optimal designs for random slope and random curvature will be introduced. We will show results on the efficiency of the D-optimal designs compared to a trivial three-group design.

[1]This work was supported by the BMBF grant SKAVOE 03SCPAB3
[2]Otto-von-Guericke University Magdeburg, E-mail: tobias.mielke@ovgu.de

2. The model

In the considered mixed model, the j-th observation of individual i, taken at the experimental setting x_{ij} in a design region X, is modeled by

$$\begin{aligned} Y_{ij} &= f(x_{ij})^T\beta_i + \epsilon_{ij}, \text{ where} \\ \epsilon_{ij} &\sim N(0,\sigma^2) \\ \beta_i &= \beta + b_i \text{ and } b_i \sim N_p(0,\sigma^2 D). \end{aligned}$$

The vector β is the p-dimensional vector of population parameters, and b_i is the vector of individual effects. It is assumed that the covariance matrix of the individual effects is a diagonal matrix:

$$D = diag(d_1,d_2,d_3) = \begin{pmatrix} d_1 & 0 & 0 \\ 0 & d_2 & 0 \\ 0 & 0 & d_3 \end{pmatrix}.$$

To interpret the matrix an example may help: If we assume $d_2 = d_3 = 0$ the regression functions of the individuals would differ in the intercept β^1 only.

The observation errors ϵ_i and the individual effects b_i are assumed to be independent. The vector of the m_i observations taken from individual i is then described by

$$Y_i = F_i\beta_i + \epsilon_i,$$

where $F_i = (f(x_{i1}),...,f(x_{im_i}))^T$ is the design matrix for the observations of individual i. Then the individual discrete design can be represented by

$$\xi_i = (x_{i1},...,x_{ik_i},m_{i1},...,m_{ik_i}) \text{ with } \sum_{j=1}^{k_i} m_{ij} = m_i.$$

The integer m_{ij} represents the number of replicated measurements in the experimental setting x_{ij}. For discrete designs m_{ij} are integers. For arbitrary $m_{ij} \in \mathbb{R}^+$ with $\sum_{j=1}^{k_i} m_{ij} = m_i$ we call ξ_i an approximate individual design.
With the normality of the random error term and the individual effects, the observation vector Y_i has the marginal distribution

$$Y_i \sim N_{m_i}(F_i\beta, \sigma^2(F_iDF_i^T + I_{m_i})).$$

Let the matrices F, G, V_i and V be defined as

$$\begin{aligned} F &:= (F_1^T,...,F_n^T)^T, \\ G &:= diag((F_1,...,F_n)), \\ V_i &:= (F_iDF_i^T + I_{m_i}) \text{ and} \\ V &:= diag(V_1,...,V_n). \end{aligned}$$

Then the model for all observations is described by

$Y = F\beta + Gb + \epsilon \sim N_{\sum m_i}(F\beta, \sigma^2 V)$, where $b = (b_1^T, ..., b_n^T)$, $\epsilon = (\epsilon_1^T, ..., \epsilon_n^T)^T$.

If we consider the case that the matrix F has full column rank and that the parameter variance matrix D is known, the population parameter vector β can be estimated using the weighted least squares estimator:

$$\hat{\beta} = (F^T V^{-1} F)^{-1} F^T V^{-1} Y.$$

Target of the design optimization is to minimize in some sense the covariance of the estimator $\hat{\beta}$:

$$cov(\hat{\beta}) = \sigma^2 (F^T V^{-1} F)^{-1}.$$

This optimization task is equivalent to maximizing the information matrix, which is of the form:

$$\mathfrak{M}_{pop} := \sum_{i=1}^{n} F_i^T V_i^{-1} F_i.$$

D-optimal designs minimize the volume of the confidence ellipsoid, what is equivalent to maximizing the determinant of the information matrix.
To turn the discrete optimization problem in a continuous one, we allow approximate more-group designs:

$$\zeta := (\xi_1, ..., \xi_k, \omega_1, ..., \omega_k) \text{ with } \sum_{i=1}^{k} \omega_i = 1.$$

This means that $100 \times \omega_i\%$ of the population will be observered unter the discrete individual design ξ_i. For approximate individual designs Schmelter[4] proved that D-optimal approximate single-group designs retain their optimality even if more-group designs are allowed.
Optimal approximate individual designs can be realized only in few cases. For the special case of two observations per individual in quadratic regression, single-group designs would lead to singular information matrices. It is obivous that they lose their optimality. In the next section we will construct D-optimal designs for quadratic regression under the assumption of random effects with variance matrices $D_1 = (d_1, 0, 0)$, $D_2 = (0, d_2, 0)$ and $D_3 = (0, 0, d_3)$ and that the individual designs consist of two observations only, whereas the population designs are assumed to be approximate more-group designs.

3. Optimal design for quadratic regression with random parameters

It is obvious that approximate single-group designs lose their optimality, if less observation per individual can be taken than parameters are to be estimated.

According to Cheng[2] and Atkins and Cheng[1], the D-optimal design for the random intercept case with two observations per individual, will be an approximate three-group design of the form

$$\zeta_{d_1}^* = \begin{pmatrix} \binom{1}{-1} & \binom{1}{\alpha_{d_1}} & \binom{-1}{-\alpha_{d_1}} \\ \omega_{d_1} & 1-\frac{\omega_{d_1}}{2} & 1-\frac{\omega_{d_1}}{2} \end{pmatrix},$$

where $\alpha_{d_1} \in (-1,1)$ and $\omega_{d_1} \in (0,1)$ depend on the variance of the intercept.
Due to invariance considerations, the D-optimal information matrix in the case with random effects and two observations per individual will be of the form

$$\mathfrak{M}_{pop}(\zeta) = \begin{pmatrix} a & 0 & b \\ 0 & c & 0 \\ b & 0 & d \end{pmatrix}$$

with

$$\begin{aligned} a &= \sum_{i=1}^{k} \frac{\omega_i}{detV_i}(2 + d_2(x_i - y_i)^2 + d_3(x_i^2 - y_i^2)^2) \\ b &= \sum_{i=1}^{k} \frac{\omega_i}{detV_i}(x_i^2 + y_i^2 - d_2 x_i y_i (x_i - y_i)^2) \\ c &= \sum_{i=1}^{k} \frac{\omega_i}{detV_i}(x_i^2 + y_i^2 + d_1(x_i - y_i)^2 + d_3 x_i^2 y_i^2 (x_i - y_i)^2) \\ d &= \sum_{i=1}^{k} \frac{\omega_i}{detV_i}(x_i^4 + y_i^4 + d_1(x_i^2 - y_i^2)^2 + d_2 x_i^2 y_i^2 (x_i - y_i)^2) \end{aligned}$$

and an arbitrary more-group design

$$\zeta := \begin{pmatrix} \binom{x_1}{y_1} & \dots & \binom{x_k}{y_k} \\ \omega_1 & \dots & \omega_k \end{pmatrix}.$$

To construct D-optimal designs and to prove the optimality, we can use the multivariate form of the Equivalence Theorem[3] with the sensitivity function g_ζ:

$$g_\zeta(x,y) := TrF(x,y)\mathfrak{M}_{pop}(\zeta)^{-1}F(x,y)^T V(x,y)^{-1},$$

for a design ζ and

$$F(x,y) := \begin{pmatrix} 1 & x & x^2 \\ 1 & y & y^2 \end{pmatrix}, \quad V(x,y) := (F(x,y)DF(x,y)^T + I_2).$$

Theorem 1. *Assume $X = [-1,1]$, $D = D_2$ and $F_i = F(x_i, y_i)$ with $x_i, y_i \in X$. The D-optimal design $\zeta_{d_2}^*$ for estimating the population parameter vector β is of the form*

$$\zeta_{d_2}^* = \begin{pmatrix} \binom{1}{-1} & \binom{\alpha_{d_2}}{-\alpha_{d_2}} \\ \omega_{d_2} & 1-\omega_{d_2} \end{pmatrix}$$

with α_{d_2} and ω_{d_2} being dependent of the variance in the parameter β^2.

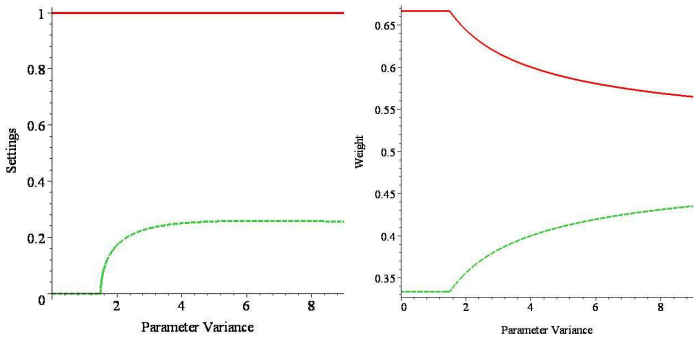


Figure 1: The figure shows the optimal experimental settings in dependence of the variance for the case D_2 and the according optimal weights.

Theorem 2. *Assume $X = [-1, 1]$, $D = D_3$ and $F_i = F(x_i, y_i)$ with $x_i, y_i \in X$. The optimal design $\zeta^*_{d_3}$ for estimating the population parameter vector β is of the form*

$$\zeta^*_{d_3} = \begin{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} & \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\ 2/3 & 1/3 \end{pmatrix},$$

independent of the variance in the parameter β^3.

The designs for $D = D_2$ and $D = D_3$ can be explicitly constructed. The D-optimal design for random slopes, is for $d_2 \leq 1.5$ equal to the optimal design in the case of random curvatures. For $d_2 > 1.5$ the point α_{d_2} and the weights ω_{d_2} can be easily calculated.

It is obvious that the two-observation more-group designs are less D-efficient than the D-optimal single-group designs with approximate individual designs. Already more interesting is the efficiency of a trivial design

$$\zeta_B = \begin{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} & \begin{pmatrix} 1 \\ 0 \end{pmatrix} & \begin{pmatrix} -1 \\ 0 \end{pmatrix} \\ 1/3 & 1/3 & 1/3 \end{pmatrix},$$

compared to the optimal designs in the cases $D = D_2$ and $D = D_3$. Cheng[2] shows similar results for random intercepts, which were included in the table. For random intercepts the decrease in efficiency is much slower than for random slope or random curvature. This can be explained by the optimal design structure, which is for $D = D_1$ similar to the trivial design.

Table 1: D-efficiency of the trivial design ζ_B compared to the D-optimal two-observation designs

$\rho = \frac{d_k}{d_k+1}$	efficiency for $k=1$	efficiency for $k=2$	efficiency for $k=3$
0	1	1	1
0.1	0.99950	0.99755	0.99755
0.2	0.99820	0.99000	0.99000
0.3	0.99631	0.97745	0.97745
0.4	0.99397	0.95995	0.95995
0.5	0.99126	0.93750	0.93750
0.6	0.98826	0.91000	0.91000
0.7	0.98500	0.86360	0.87750
0.8	0.98153	0.78661	0.84000
0.9	0.97788	0.67671	0.79750

4. Discussion

In the present simple model of quadratic regression, the optimal design in cases when only two observations are allowed, strongly depends on the variance in the parameters. For isolated variances in the slope or the curvature of the regression function, the optimal design could be explicitly calculated. If more than one parameter has random effects, the model complicates. In examples for this case could be seen that one of the design structures derivated for the cases $D = D_1$, $D = D_2$ or $D = D_3$ leads to the D-optimal design. This leads to the assumption that one parameter dominates the design and that the domination is dependent of the variance ratios. In relation to this, it will be very interesting, to analyze the design for cases, if two or more parameters at the same time influence the design.

Acknowledgment The author is very grateful for the help by the project SKAVOE of the BMBF, to Rainer Schwabe (Univ. of Magdeburg) and to Thomas Schmelter (Bayer Schering Pharma AG).

References

[1] Atkins, J.E., Cheng, C.S.(1998) *Optimal regression in the presence of random block effects.* Journal of Statistical Planning and Inference 77, 321-335.

[2] Cheng, C.S. (1995) *Optimal regression designs under random block effects.* Statist. Sinica 5, 485-497

[3] Fedorov, V.V. (1972) *Theory of optimal experiments*, Academic Press, New York.

[4] Schmelter, T. (2007) *The optimality of single-group designs for certain mixed models.* Metrika (2007) 65:183-193

6th St.Petersburg Workshop on Simulation (2009) 1072-1076

Multiple-Objective response-adaptive repeated measurement designs for clinical trials with dichotomous outcomes[1]

Yuanyuan Liang[2], Keumhee Chough Carriere[3]

Abstract

The development and use of adaptive design methods in clinical trials are in great demand. Liang and Carriere [1] proposed a new adaptive allocation to improve current strategies for building response-adaptive designs to construct multiple-objective repeated measurement designs (RMDs). This new rule is designed to increase estimation precision and treatment benefit by assigning more patients to a better treatment sequence. In their paper, they demonstrate that the designs constructed under the new proposed allocation rule for studies with normally distributed outcomes can be nearly as efficient as fixed optimal designs in terms of the mean squared error, while leading to improved patient care. In this paper, we study the properties of this adaptive allocation rule on dichotomous outcomes.

1. Introduction

Drug development is complex and costly, requiring the testing of numerous chemical compounds for their potential to treat disease. Before a new drug can be marketed in the United States, a new drug application (NDA), which includes scientific and clinical data, must be approved by the Food and Drug Administration (FDA). In the past several decades, it is recognized that increasing spending on biomedical research does not reflect an increase of the success rate of pharmaceutical (clinical) development. In 2006, the FDA released a Critical Path Opportunities List that outlines initial project to assist the sponsors in identifying the scientific challenges underlying the medical product pipeline problem. Among these initial projects, the FDA calls for advancing innovative trial designs, especially for the use of prior experience or accumulated information in trial design. The development and use of adaptive design methods in clinical trials are in great demand.

[1]This work was supported in part by grants from the Alberta Heritage Foundation for Medical Research and the Natural Sciences and Engineering Council of Canada.

[2]University of Texas Health Science Center at San Antonio, E-mail: liangy@uthscsa.edu

[3]University of Alberta, E-mail: kccarrie@ualberta.ca

Liang and Carriere [1] proposed a new adaptive allocation to improve current strategies for building response-adaptive designs to construct multiple-objective repeated measurement designs. This new rule is designed to increase estimation precision and treatment benefit by assigning more patients to a better treatment sequence. In their paper, they demonstrate that the designs constructed under the new proposed allocation rule for studies with normally distributed outcomes can be nearly as efficient as fixed optimal designs in terms of the mean squared error, while leading to improved patient care.

In this paper, we study the properties of this adaptive allocation rule for dichotomous outcomes.

2. Allocation Rule

We discuss the procedure of constructing a multiple-objective response-adaptive RMD with a pre-specified N (total number of subjects) and λ (percentage weight given to the objective of maximizing the information matrix). Basically, we adopted the usual optimal design construction methods [2-5] to determine the allocation rule for solving multiple objectives.

- Step 1: The first m ($m < N$) patients are randomly assigned to all possible treatment sequences or using some desired subset of them.
- Step 2: To allocate the l^{th} patient, $(m+1) \leq l \leq N$, calculate the expected Fisher information matrix, $\hat{\mathbf{A}}_l^{\,k}(\mathbf{H}_{l-1})$ and the evaluation function $g_{l-1,k}(\mathbf{H}_{l-1})$ for each treatment sequence k, where the domain of k is all possible treatment sequences; $\mathbf{H}_{l-1}$ is the actual data observed from the first $(l-1)$ patients; $\hat{\mathbf{A}}_l^{\,k}(\mathbf{H}_{l-1})$ is the information matrix with the first l patients on the basis of $\mathbf{H}_{l-1}$ and the assumption that the l^{th} patient will be treated by treatment sequence k; and $g_{l-1,k}(\mathbf{H}_{l-1})$ is a suitably chosen evaluation function defined based on $\mathbf{H}_{l-1}$ for treatment sequence k. For simplicity, $\hat{\mathbf{A}}_l^{\,k}(\mathbf{H}_{l-1})$ and $g_{l-1,k}(\mathbf{H}_{l-1})$ will be written as $\hat{\mathbf{A}}_l^{\,k}$ and $g_{l-1,k}$, respectively. Without loss of generality, we assume that a higher value of $g_{l-1,k}$ indicates a better treatment sequence.
- Step 3: Choose the treatment sequence k^* from the s possible treatment sequences for the l^{th} patient, such that $\Lambda(l,k^*) = max_{k=1\ldots s}\Lambda(l,k)$, where

$$\Lambda(l,k) = \lambda\frac{\Theta(\hat{\mathbf{A}}_l^{\,k})}{\Theta(\hat{\mathbf{A}}_l^{\,k_l^{(O)}})} + (1-\lambda)\frac{g_{l-1,k}}{g_{l-1,k_{l-1}^{(B)}}}, \tag{1}$$

  Equation (1) is called the selection criterion, which is designed to achieve two objectives by creating a balance between them. When $\lambda = 1$, the selection criterion becomes the traditional criterion of a response-adaptive design, that is, maximizing the information matrix under a given optimality criterion Θ such as A-, D- or E-optimality. When $\lambda = 0$, the efficacy of treatment

sequences measured by the pre-defined evaluation function is the only concern. At this time, out of all s treatment sequences, one of the treatment sequences $k_l^{(O)}$ maximizes the optimality criterion, Θ, and possibly another treatment sequence $k_{l-1}^{(B)}$ has the best value on the evaluation function, $g_{l-1,k}$. In situations where more than one treatment sequence achieves the maximum criterion score, one can randomly assign the l^{th} patient to one of them.

- Step 4: Repeat steps 2 to 3 until all N patients have been allocated.

3. Adaptive Two-Treatment p-Period Repeated Measurement Design

Let A and B denote the two different treatments. In a p-period repeated measurement design, there are 2^p possible treatment sequences. Suppose that N patients are randomly selected from a well-defined population. The first l patients are assigned to some specific treatment sequences and the responses are collected. The $(l+1)^{th}$ patient is ready for assignment. The traditional design methodology [6-8] can be applied for this binary design. In this section, we give the detailed steps for implementing the new proposed binary adaptive design.

When $r = 1, 2, ..., p$, let π_r be the success probability of treatment A in the r^{th} period; when $r = p+1, p+2, ..., 2p$, let π_r be the success probability of treatment B in the $(r-p)^{th}$ period. In this paper, for simplicity, we assume there are no covariates and $\pi_r s$ are fixed but unknown parameters of interest.

Let N_{kl} denote the number of subjects receiving treatment sequence k up to the l^{th} patient, where $k = 1, 2, ..., 2^p$, and $\sum_{k=1}^{2^p} N_{kl} = l$. Let $S_l = (S_{1Al}, ..., S_{pAl}, S_{1Bl}, ..., S_{pBl})^T$, where S_{itl} denotes the number of successes of treatment t in the i^{th} period, where $i = 1, 2, ..., p$ and $t = A$ or B. For example, S_{1Al} represents the number of successes of A in the first period among the first l patients. Let $S_l[r]$ be the r^{th} element of S_l. Also, let $N_{L_l} = (N_{L_l}[1], N_{L_l}[2], ..., N_{L_l}[2p])^T$. Up to the l^{th} patient, $N_{L_l}[r]$ denotes the total number of patients receiving treatment A in the r^{th} period for $r = 1, 2, ..., p$; and $N_{L_l}[r]$ denotes the total number of patients receiving treatment B in the $(r-p)^{th}$ period for $r = p+1, p+2, ..., 2p$.

We let ℓ_l denote the log-likelihood function up to the l^{th} patient so that:

$$\ell_l \propto \sum_{r=1}^{2p} (S_l[r] log \pi_r + (N_{L_l}[r] - S_l[r]) log(1-\pi_r)).$$

The Fisher information matrix up to the l^{th} patient, $\mathbf{A}_l$ becomes a $2p \times 2p$ diagonal matrix with the r^{th} element in the diagonal of $E[\frac{S_l[r]}{\pi_r^2} + \frac{N_{L_l}[r]-S_l[r]}{(1-\pi_r)^2}]$, that is

$$\mathbf{A}_l = diag\{E[\frac{S_l[r]}{\pi_r^2} + \frac{N_{L_l}[r] - S_l[r]}{(1-\pi_r)^2}]\}_{2p \times 2p}.$$

The maximum likelihood estimation of the unknown parameter π_r up to the l^{th} patient is obtained as

$$\hat{\pi}_r = \frac{S_l[r]}{N_{L_l}[r]}, \tag{2}$$

where $r = 1, 2, ..., 2p$.

After some algebraic manipulations, we can show that the conditional expected information matrix given the history of the first l patients and the assumption that the $(l+1)^{th}$ patient will receive treatment sequence k becomes

$$\mathbf{A}_l^k = diag\{E[\frac{S'_{l+1}[r]}{\pi_r^2} + \frac{N'_{L_{l+1}}[r] - S'_{l+1}[r]}{(1-\pi_r)^2}]\}_{2p\times 2p}, \tag{3}$$

where $S'_{l+1}[r] = S_l[r] + \alpha_r$, and $N'_{L_{l+1}} = N_{L_l}[r] + \beta_r$, α_r is the r^{th} element of the diagonal of the matrix $\pi \times \mu_k$, β_r is the r^{th} element of μ_k, $\pi = (\pi_1, \pi_2, ..., \pi_{2p})^T$ and $\mu_k = (d(1,k), d(2,k), ..., d(2p,k))$ is a $1 \times 2p$ row vector of zeros and ones. For $1 \le r \le p$, $d(r,k) = 1$ if treatment A is used in the r^{th} period of the treatment sequence k; and $d(r,k) = 0$, otherwise. For $p+1 \le r \le 2p$, $d(r,k) = 1$ if treatment B is used in the $(r-p)^{th}$ period of the treatment sequence k; and $d(r,k) = 0$, otherwise.

The unknown parameters in Equation (3), $\pi_r s$, are estimated using the maximum likelihood method (see Equation (2)).

In the spirit of the play-the-winner rule, an evaluation function for treatment sequence k up to the l^{th} patient is defined as the average number of successes for treatment sequence k up to the l^{th} patients, that is

$$g_{lk} = \frac{\mu_k \times S_l}{N_{kl}}.$$

Given a value of λ and an optimality function Θ, we will assign the treatment sequence k^* to the $(l+1)^{th}$ patient such that the selection criterion (Equation (1)) will be maximized.

4. Simulation Study

In this section, we apply the allocation rule to construct two-treatment two- and three-period response-adaptive RMDs. To assess the efficiency of an adaptive design, the matrix of mean squared error (MSE) for θ is computed

$$MSE = E[(\hat{\theta} - \theta)(\hat{\theta} - \theta)^T]$$

where θ is a vector of parameters of interest, and $\hat{\theta}$ is an estimate of θ.

In simulation studies, MSE is estimated by

$$MSE = \frac{\sum_{b=1}^{B} (\hat{\theta}^{(b)} - \theta)(\hat{\theta}^{(b)} - \theta)^T}{B}$$

where $\hat{\theta}^{(b)}$ is the maximum likelihood estimator of θ obtained in the b^{th} simulation run for the total of B simulations.

Denote MSE_1 as the matrix of MSE for a proposed adaptive design and MSE_0 for a reference design. Based on A-, D-, or E- optimality criteria [9], the relative efficiency (RE) of the adaptive design compared with the reference design is defined as follows (respectively):

$$RE_A = \frac{trace(MSE_0)}{trace(MSE_1)}; RE_D = \frac{|MSE_0|}{|MSE_1|}; RE_E = \frac{maxeigenvalue(MSE_0)}{maxeigenvalue(MSE_1)}.$$

When $RE = a > 1$, the adaptive design is $(a-1) \times 100\%$ more efficient than the reference design. When $a < 1$, the adaptive design is only $a \times 100\%$ as efficient as the reference design. The results under the E-optimality are similar to those under the A-optimality and therefore we will not discuss it further.

Figure 1 (upper row) illustrates the relative efficiency between the adaptive designs and the reference design (the fixed design $AB/BA/AA/BB$ with equal number of subjects per treatment sequence) under the A- and D-optimality, respectively. We can see that the design with $\lambda = 1$ is nearly as efficient as the fixed design and it produces a more precise estimation than designs with $\lambda < 1$. This is as expected, because the weight is placed on the efficacy of treatments rather than on the precision of the estimates, as the value of λ decreases. Designs with a small value of λ are not recommended since they have a considerably low design efficiency compared to the fixed design. Figure 1 (bottom row) illustrates the similar comparison under unequal success probability situation. Although the design with $\lambda = 1$ has the highest efficiency in terms of MSE, the design with $\lambda < 1$ takes treatment benefits into account and offers relatively high estimation precision as well.

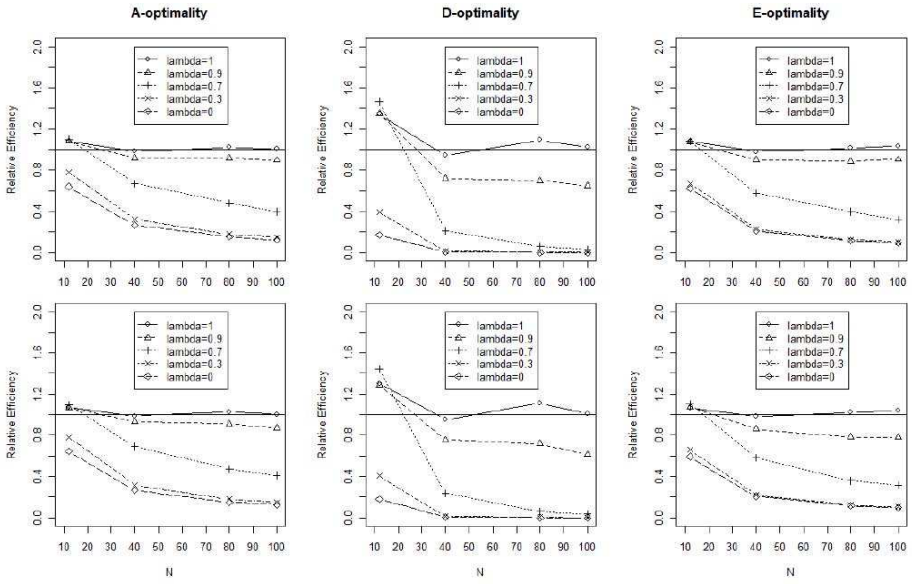
We also considered the case of three-period adaptive design in reference to the fixed ABB/BAA design. Similar conclusion as in the two-period design is observed.

To summarize the above, when $\lambda = 1$, an approximately equal number of subjects is assigned to a treatment sequence and its dual treatment sequence, and slightly more patients were given AA/BB and ABB/BAA in a two-period and three-period design, respectively. However, when $\lambda < 1$ and as it decreases, adaptive designs assign more patients to a better performing treatment sequence, and less subjects to a worse treatment sequence. In addition, when the total number of patients involved in the study increases, the precision of the estimators increases accordingly. The design with $\lambda = 1$ has the highest efficiency in terms of MSE which is almost as efficient as the fixed optimal design. However, the designs with $\lambda < 1$ take the treatment advantage into account. In practice, these two objectives should be balanced out.

5. Conclusion

In this paper, we utilized the allocation strategy proposed by [1] to construct adaptive repeated measurement designs with dichotomous responses/outcomes. We

provide the detailed allocation rule for constructing adaptive two-treatment two- or three-period repeated measurement designs, and then extend it to two-treatment p-period repeated measurement designs. In simulation studies, we demonstrate that, as expected, the adaptive designs constructed under the new proposed allocation rule are not as efficient as the fixed designs in terms of the mean squared error, but they successfully allocate more patients to better treatment sequences. The value of λ, which is to balance the two objectives of increasing the estimation precision and decreasing the proportion of patients receiving inferior treatments, can be pre-determined by researchers. A large value of λ will place more emphasis on the estimation precision. When $\lambda = 1$, the allocation rule becomes the usual response adaptive design as considered by other researchers [4]. A small value of λ will emphasize the performance/benefit of the treatment. When $\lambda = 0$, the allocation rule becomes a typical play-the-winner rule [10]. In addition, simulation studies show that the design with a high value of $\lambda < 1$ significantly favors the allocation results toward more effective treatment sequences without much loss of estimation precision.



Note: $\boldsymbol{\theta} = (\pi_1, \pi_2, \pi_3, \pi_4)^T$. The fixed design $AB/BA/AA/BB$ is the reference design. Relative efficiencies are calculated based on 5,000 simulations with π_r=0.5, r = 1,2,3,4 (upper row) and $\pi_1 = 0.5$, $\pi_2 = 0.6$, $\pi_3 = 0.4$, $\pi_4 = 0.5$ (bottom row).

Figure 1: Relative Efficiency of Estimation of θ for Two-Period Two-Treatment Designs

References

[1] Liang Y., Carriere K.C. (2009) Multiple-objective response-adaptive repeated measurement designs for clinical trials. *Journal of Statistical Planning and*

Inference, **139**, 1134-1145.

[2] Kershner R.P. (1986) Optimal 3-period 2-treatment crossover designs with and without baseline measurements. *Proceedings of the Biopharmacerutical Section of the American Statistical Association*, 152-156.

[3] Atkinson A.C., Donev A.N., Tobias R.D. (2007) *Optimum Experimental Designs, with SAS*. Oxford: Oxford University Press.

[4] Kushner H.B. (2003) Allocation rules for adaptive repeated measurements designs. *Journal of Statistical Planning and Inference*, **113**, 293-313.

[5] Laska E.M., Meisner M., Kushner H.B. (1983) Optimal crossover designs in the presence of carryover effects. *Biometrics*, **39**, 1087-1091.

[6] Fedorov V.V., Hackl P. (1997) *Model-Oriented Design of Experiments*. Springer, New York.

[7] Fedorov V.V., Leonov S. (2005) Response driven designs in drug development. In: Berger M.P.F., Wong W.K. (Eds.), *Applied Optimal Designs*. Wiley, Chichester, pp. 103-136.

[8] Dragalin V., Fedorov V., Wu Y. (2008) Adaptive designs for selecting drug combinations based on efficacy-toxicity response. *Journal of Statistical Planning and Inference*, **138**, 352-373.

[9] Kiefer J. (1975) Construction and optimality of generalized Youden designs. In *A Survey of Statistical Designs and Linear Models* (J. N. Srivastava, Ed.) 333-353. North-Holland, Amsterdam.

[10] Zelen M. (1969) Play the winner rule and the controlled clinical trial. *Journal of the American Statistical Association*, **64**, 131-146.

6th St.Petersburg Workshop on Simulation (2009) 1079-1083

Unbiased procedures in time-dependent regression experiments

E.V. Sedunov[1], A.N. Sedunova[1]

Abstract

The paper deals with the problem of optimal unbiased design of dynamic regression experiment under the continuous observation over the object using a priori information of the Bayesian type.

1. Introduction

On practice we deal with some regression problems, where the dependence under research $\eta(x,t)$ includes unoperate variable t along with the operate one x. We call t time although in physical sense it may be any monotone variable. According to conditions of the experiment, under a fixed value of x we can continuously obtain observation results without sufficient material expenditures, as performing observation under another value of x corresponds to starting the process once again and thus requires some additional resources. An a priori information may lead to various approaches to stating the problems of the kind [1,2,3]. In this paper as well as in [4,5], we consider using a priori information of the Bayesian type.

2. Measurement scheme

Let $\mathcal{F} \subset L^2(\mathcal{X},\mu)$ and $\mathcal{G} \subset L^2(\tau,\nu)$ be finite-dimensional Hilbert spaces, $\mathcal{X}$ be a compact in R^k, τ be an interval, μ be a finite measure on $\mathcal{X}$, and ν be a Lebesgue measure on τ. Let us introduce the space $\mathcal{H} = \mathcal{F} \otimes \mathcal{G}$, which consists of functions $h(x,t)$, such that $h(\cdot,t) \in \mathcal{F}$ (**mod** ν) and $h(x,\cdot) \in \mathcal{G}$ (**mod** μ). Let us define the following scalar product:

$$\langle h_1, h_2 \rangle_{\mathcal{H}} = \int_{\mathcal{X}} \int_{\tau} h_1(x,t) h_2(x,t)\, d\mu(x)\, d\nu(t)$$

in this space. The dependence under research $\eta(x,t)$, $x \in \mathcal{X}$, $t \in \tau$ belongs to the Hilbert space $\mathcal{H}$, and the domain of experiment design appears to be a bounded

[1]St. Petersburg State University of Technology and Design

subset $\mathcal{U}$ of the linear restricted operators $\mathcal{Q} : \mathcal{H} \to \mathcal{G}$. The experiment designs take the form of the following discrete probability masses ξ on $\mathcal{U}$:

$$\xi = (\mathcal{Q}_1, \ldots, \mathcal{Q}_n; p_1, \ldots, p_n), \quad \mathcal{Q}_j \in \mathcal{U}, \quad \mathcal{Q}_j : \mathcal{H} \to \mathcal{G} \tag{1}$$

where p_j is the weight of the operator $\mathcal{Q}_j$, $p_j > 0$, $j = 1, \ldots, n$, $\sum_{j=1}^{n} p_j = 1$.

At design ξ, the measurement scheme for unknown member η of the space $\mathcal{H}$ takes the following form:

$$y_j(t, \omega) = (\mathcal{Q}_j \eta)(t, \omega) + \epsilon_j(t, \omega), \quad j = 1, \ldots, n, \tag{2}$$

where $\mathcal{Q}_j \in \mathbf{supp}\ \xi$, $t \in \tau$, $y_j(t, \omega)$ and $\epsilon_j(t, \omega)$ are random observation results and errors of j-th experiment, respectively, and ω is a member of the set of random events, such that

$$E\epsilon_j(t, \omega) = 0, \quad E\epsilon_j(t, \omega)\epsilon_i(t, \omega) = 0, \quad j, i = 1, \ldots, n, \quad j \neq i, \quad t \in \tau$$

and covariation operator $\mathcal{D}[\epsilon_j] : \mathcal{G} \to \mathcal{G}$ is defined and it is invertible, $j = 1, \ldots, n$.

The boundedness of resources is stated in the following form:

$$\mathbf{card}(\mathbf{supp}\ \xi) = n < n_{\mathbf{max}} = \mathbf{dim}\ \mathcal{F}, \tag{3}$$

that leads to necessity of taking into account the bias approximation error.

Let us consider a particular case of experiment (1)–(2), namely the following measurement scheme over the set $\mathcal{U}_e$ of elementary operators $\mathcal{Q}_z$:

$$\begin{aligned}
y_j(t, \omega) &= \left(\mathcal{Q}_{z_j} \eta\right)(t, \omega) + \epsilon_j(t, \omega) = \eta(z_j, t) + \epsilon_j(t, \omega), \quad j = 1, \ldots, n, \\
\mathcal{U} &= \mathcal{U}_e := \{\mathcal{Q}_z, z \in \mathcal{X}\}, \quad \mathcal{Q}_z : \mathcal{H} \to \mathcal{G}, \quad z \in \mathcal{X}, \\
(\mathcal{Q}_z h)(t) &= h(z, t), \quad \forall\, h \in \mathcal{H}, \quad t \in \tau, \\
\xi &= (\mathcal{Q}_{z_1}, \ldots, \mathcal{Q}_{z_n}, p_1, \ldots, p_n).
\end{aligned} \tag{4}$$

Covariation operator $\mathcal{D}[\epsilon_j]$ is defined by its matrix $\mathcal{D}(\epsilon_j)$ in some ν-orthonormal basis of the space $\mathcal{G}$,

$$\mathcal{D}(\epsilon_j) = \left(\sigma_j^{(i,k)} p_j^{-1}\right)_{i,k=1}^{\nu}, j = 1, \ldots, n.$$

Another particular case is spectroscopic measurements on the set $\mathcal{U}_\Delta$ of integral operators of the following form:

$$\begin{aligned}
y_j(t, \omega) &= \left(\mathcal{Q}_{a_j} \eta\right)(t) + \epsilon_j(t, \omega) \\
&= \int_{\mathcal{X}} a_j(x) \eta(x, t) \mu(dx) + \epsilon_j(t, \omega), \quad j = 1, \ldots, n, \\
\mathcal{U} &= \mathcal{U}_\Delta := \{\mathcal{Q}_a, a \in \mathcal{A}_\Delta\}, \quad \mathcal{Q}_a : \mathcal{H} \to \mathcal{G}, \\
\Delta &\subset \mathcal{R}, \\
\mathcal{A}_\Delta &:= \left\{a \in L^2(\mathcal{X}, \mu) : a(x) \in \Delta\,(\mathbf{mod}\ \mu)\right\}, \\
(Q_a h)(t) &= \int_{\mathcal{X}} a(x) h(x, t) \mu(dx), \quad \forall\, h \in \mathcal{H},
\end{aligned} \tag{5}$$

where $a(x)$ is the device's instrument function.

3. Optimality criterion

By force of condition (3), an estimate $\hat{\eta}$ of the set member $\eta \in \mathcal{H}$ belongs to the finite-dimensional affine subspace $\mathcal{H}_d \subset \mathcal{H}$. Let us find it using the affine operator $S_d : L^2(\mathcal{U}, \xi) \to \mathcal{H}_d$:

$$\hat{\eta} = S_d y \tag{6}$$

where $\mathcal{H}_d = d + \mathcal{H}_0$, $d \in \mathcal{H}_d^{\perp}$, $\mathcal{H}_0$ is a linear subspace of $\mathcal{H}$, $\mathcal{H}_0^{\perp}$ is the orthogonal complement to $\mathcal{H}_0$ in $\mathcal{H}$; $S_d = d + S_0$, $S_0 : L^2(\mathcal{U}, \xi) \to \mathcal{H}_0$ is a linear operator.

According to paper [6], let us call a collection $\pi = (d, \mathcal{H}_0, S_0, \xi)$, satisfying following condition of unbiasedness of estimate (6):

$$E\hat{\eta} = P_d \eta, \quad \forall\, \eta \in \mathcal{H}, \tag{7}$$

where $P_d : \mathcal{H} \to \mathcal{H}_d$ is an affine orthogonal projector, an affine unbiased in metric of the space $\mathcal{H}$ procedure of recovery of $\eta \in \mathcal{H}$. Let us designate the collection π of procedures, satisfying condition (7), with $\tilde{\Pi}$. Two-criteria extremum problem (where criteria take into account bias and random errors, respectively) is considered with priority given to the systematic error:

$$B_\gamma(\pi) := \int_{\mathcal{H}} \mathbf{dist}^2(\eta, E\hat{\eta}) d\gamma(h), \tag{8}$$

averaged via a priori defined measure γ.

The random approximation error is characterized with continuous monotone functional Φ of covariation operator $\mathcal{D}[\hat{\eta}] : \mathcal{H}_0 \to \mathcal{H}_0$ of the estimate $\hat{\eta}$.

Thus, let us pass to the following problem statement:

$$\pi^* = \arg \inf_{\pi \in \tilde{\Pi}} \beta_\gamma(\pi), \tag{9}$$

$$\pi^* = \arg \inf_{\pi \in \tilde{\Pi}^*} \Phi(\pi), \tag{10}$$

where $\tilde{\Pi}^* \subseteq \tilde{\Pi}$ is the set of solutions of problem (9).

4. Basic analytical results

The results from papers [4,5] can be applied to the considered problem statement.

Theorem 1. *While solving problem (9) for measurement scheme (4), it is sufficient to consider the procedures of the following form:*

$$\pi_\xi = (d_\xi, \mathcal{H}_\xi, S_\xi, \xi) \tag{11}$$

where $\xi \in \Xi_n := \{\xi \in \Xi : \boldsymbol{supp}\ \xi \subset \mathcal{U}_e, \boldsymbol{card}(\boldsymbol{supp}\ \xi) = n\}$; $d = d_\xi = P_\xi^{\perp}(E\gamma)$, $P_\xi^{\perp} : \mathcal{H} \to \mathcal{H}_\xi^{\perp}$ *is an orthogonal projector;*
$\mathcal{H}_\xi = \boldsymbol{Lin}\{h_{z_j}(x)g(t), g \in \mathcal{G}, \mathcal{Q}_{z_j} \in \boldsymbol{supp}\ \xi \subset \mathcal{U}_e, j = 1, \ldots, n\}$, $h_z(x)$ *is the reproducing kernel of the space* $\mathcal{F}$;

$S_\xi = M_\xi^{-1} J_\xi^*, \mathcal{D}^{-1}[\epsilon]$, $S_\xi : L^2(\mathcal{U}, \xi) \to \mathcal{H}_\xi$, J_ξ *is a reduction of the measurement operator* $I_\xi : \mathcal{H} \to L^2(\mathcal{U}_e, \xi)$ *to the subspace* $\mathcal{H}_\xi$, $I_\xi \eta = \mathcal{Q}\eta$, $\eta \in \mathcal{H}$, $\mathcal{Q} \in$ ***supp*** ξ, $M_\xi = J_\xi^* \mathcal{D}^{-1}[\epsilon] J_\xi : \mathcal{H}_\xi \to \mathcal{H}_\xi$ *is the information operator of the estimate* $\hat{\eta}(x,t)$; $S_{d_\xi} = d_\xi + S_\xi : L^2(\mathcal{U}_e, \xi) \to \mathcal{H}_{d_\xi} = d_\xi + \mathcal{H}_\xi$. *Here the value of estimate* $\hat{\eta}(x,t)$ *from the procedure* π_ξ *is calculated by the following formula:*

$$\hat{\eta}(x,t) = (E\gamma)(x,t) + \sum_{j,i=1}^{n} \left[H^{-1}\right]_{ji} \left\{\bar{y}_i(t,\omega) - (E\gamma)(z_i,t)\right\} h_{z_j}(x), \qquad (12)$$

where $H = \left(h_{z_j}(z_i)\right)_{j,i=1}^n$, $\bar{y}_j(t,\omega)$ *is the average value of the observations at* j*-th experiment with respect to the weights of the observations.*

Theorem 2. *If there exists a design* ξ^* *from* Ξ_n*, which satisfies condition (3), and* $\mathcal{H}_{\xi^*} = \mathcal{H}^{(nv)}$*, which is a subspace of* $\mathcal{H}$*, generated by the eigenvectors of the covariation operator* $\mathcal{D}[\gamma]$ *of measure* γ*, corresponding to its* nv *largest eigenvalues, then procedure* π_{ξ^*} *of form (11) solves problem (9). At that for bias error (8),* $B_\gamma(\pi_{\xi^*}) = \sum_{i=nv+1}^{m} \lambda_i$*, where* $\lambda_1 \geq \ldots \geq \lambda_{nv} \geq \ldots \geq \lambda_m \geq 0$ *is the sequence of eigenvalues of the operator* $\mathcal{D}[\gamma]$.

Corollary 1. *Let the spectrum* $\{z_1^*, \ldots, z_n^*\}$ *of design* ξ^* *meet the following conditions:*

$$\begin{aligned} \det\left(\psi_i(z_j^*,t)\right)_{i,j=1}^n &\neq 0, \quad \forall t \in \tau, \\ \psi_i(z_j^*,t) &= 0, \quad i = nv+1, \ldots, m; \quad j = 1, \ldots, n, \quad \forall\, t \in \tau, \end{aligned} \qquad (13)$$

where $\psi_i(x,t)$, $i = 1, \ldots, m$ *are orthonormal eigenfunctions of the operator* $\mathcal{D}[\gamma]$*, ordered via decrease of corresponding eigenvalues. Then* π_{ξ^*} *is a solution of problem (9).*

Here and further let us assume that the support of design ξ is defined and fixed, and the weights $\{p_j\}_{j=1}^n$ of design $\xi(p)$ can be evaluated as the solutions of problem (10) in the following statement:

$$p^* = \arg\inf_{p \in \mathcal{P}} \Phi[\mathcal{D}_{\xi(p)}], \qquad (14)$$

where $\mathcal{P} := \left\{p = (p_1, \ldots, p_n) : 0 < p_j \leq 1, j = 1, \ldots, n, \sum_{j=1}^n p_j = 1\right\}$,
$\Phi[\mathcal{D}_{\xi(p)}]$ is a given differentiable convex functional of covariant operator $\mathcal{D}_{\xi(p)} = M_{\xi(p)}^{-1}$ of the estimate $\hat{\eta}(x,t)$, which matrix in the orthonormal basis of the space $\mathcal{H}_\xi$ assumes the following form:

$$\mathcal{D}(p) = A\Sigma^{1/2} P^{-1} \Sigma^{1/2} A',$$

where $A = \left(\left\langle f_j, h_{z_1}^+ \right\rangle_{\mathcal{F}} I_v, \ldots, \left\langle f_j, h_{z_n}^+ \right\rangle_{\mathcal{F}} I_v\right)_{j=1}^n$,
$P = \mathrm{diag}(p_1 I_v, \ldots, p_n I_v)$, $\Sigma = \mathrm{diag}(\Sigma_1, \ldots, \Sigma_n)$,

$\{f_l(x)\}_{l=1}^n$ is the orthonormal basis of the space $\mathcal{F}_\xi = \mathbf{Lin}\left\{h_{z_j}(x), j = 1, \ldots, n\right\}$, $\Sigma_j = (\sigma_j^{u,w})_{u,w=1}^v$, $\left\{h_{z_j}^+(x)\right\}_{j=1}^n$ is a system, which is biorthogonal to the linearly independent system $\left\{h_{z_j}(x)\right\}_{j=1}^n$, I_v is the unit diagonal matrix of dimension $(v \times v)$.

Theorem 3. *Under the above assumptions:*

a) Problem (14) has a solution.

b) Validity of the following equalities is the necessary and sufficient condition for the optimal choice of weights $p_1^*, \ldots, p_n^*$ in problem (14):

$$\varphi_j(\xi(p^*)) = \mathrm{tr}\left(\mathcal{D}\frac{\partial \Phi}{\partial \mathcal{D}}(\xi(p^*))\right), \quad j = 1, \ldots, n,$$

where

$$\begin{aligned} \varphi_j(\xi(p)) &= p_j^{-2}\mathrm{tr}\left(A_j'\frac{\partial \Phi}{\partial \mathcal{D}}(\xi(p))A_j\Sigma_j\right), \quad j = 1, \ldots, n, \\ A_j &= \left(\left\langle f_1, h_{z_j}^+\right\rangle_{\mathcal{F}} I_v, \ldots, \left\langle f_n, h_{z_j}^+\right\rangle_{\mathcal{F}} I_v\right)', \quad j = 1, \ldots, n. \end{aligned}$$

c) The optimal choice of the weights via the D-criterion $(\Phi[\mathcal{D}(p)] = \ln\det\mathcal{D}(p))$ and linear criteria $(\Phi[\mathcal{D}(p)] = \mathrm{tr}(L\mathcal{D}(p))$, where L is a given positively semidefinite matrix) is defined with the following formulas:

$$p^* = 1/n, j = 1, \ldots, n;$$
$$p_j^* = \frac{\left(\mathrm{tr}\left(A_j'LA_j\Sigma_j\right)\right)^{1/2}}{\sum_{r=1}^n \left(\mathrm{tr}(A_r'LA_r)\right)^{1/2}}.$$

The above results can be easily applied to problem (9), (10) in experiment (5). In this view, we are just to re-define the estimate space as follows:

$$\mathcal{H}_\xi = \mathbf{Lin}\left\{P_{\mathcal{F}}a_j(x)g(t) : g \in \mathcal{G}, a_j(x) \in \mathcal{A}_\Delta, j = 1, \ldots, n\right\},$$

where $P_{\mathcal{F}} : L^2(\mathcal{X}, \mu) \to \mathcal{F}$ is the orthogonal projector, and re-define the estimation formula as well:

$$\hat{\eta}(x,t) = (E\gamma)(x,t) + \sum_{j,i=1}^n \left[H^{-1}\right]_{ji}\left\{\bar{y_i}(t) - \langle E\gamma(\cdot,t)a_i(\cdot)\rangle_{\mathcal{F}}\right\}P_{\mathcal{F}}a_j(x)$$

where $H = (\langle P_{\mathcal{F}}a_j, P_{\mathcal{F}}a_i\rangle)_{j,i=1}^n$.

5. Example

Let us consider the problem of observing the spectro-temporal distributions $\eta(x,t)$ of density of molecular structures' luminescence while exciting them with an impulse source. The standard methods of analysis assumes expanding the luminescent emission in a monochromator by the wave lengths x. After each j-th

excitation pulse, we are to measure the decay curves of the luminescence $\eta(x,t)$, $x \in \left[x_j - \frac{\Delta x}{2}, x_j + \frac{\Delta x}{2}\right]$ in some small intervals of wave length Δx. The pulse number of the excitation source should be restricted with some value, which depends on molecular structure under research and the source intensity.

Let $\mathcal{F} = \mathbf{Lin}\left\{1, x, x^2\right\}$, $x \in \mathcal{X} = [-1,1]$, $\mu(dx) = 0.5dx$, $G = \mathbf{Lin}\{t\}$, $t \in \tau = [0,1]$; $\mathcal{U} = \mathcal{U}_e$, $\mathbf{dim}\ \mathcal{F} = 3$, $n = 2 < 3$, $\gamma = (t, xt, x^2t, ; 1/3, 1/3, 1/3)$, $\Sigma = \mathbf{diag}(0,1;0,5)$, and $\Phi[\mathcal{D}(p)] = \mathrm{tr}\mathcal{D}(p)$ (A-optimality) be the criterion.

Let us introduce an orthonormal basis $(\sqrt{3}\ t, 3xt, (\sqrt{15}/2)(3x^2-1)t)$ in the space $\mathcal{H}$ and expand the matrix of operator $\mathcal{D}[\gamma]$ by it. The calculations demonstrate, that eigenvector $\bar{q}_3 = (1, \sqrt{3}, \sqrt{5})'$ and eigenfunction $\psi_3(x,t) = 3\sqrt{3}t(5x^2 + 2x - 1)$ correspond to the minimal eigenvalue $\lambda_3 = 0$ of the operator $\mathcal{D}[\gamma]$. Thus, it follows from (13), that $z_1^* = -(\sqrt{6}+1)/5$ and $z_2^* = -(\sqrt{6}-1)/5$ correspond to operators $\mathcal{Q}_{z_j} \in \mathbf{supp}\ \xi^*$, $j = 1,2$. By theorem 3, the A-optimal weights are $p_1^* = 0.227$ and $p_2^* = 0.723$. Finally derive from (12) the following expression for estimate $\hat{\eta}(x,t)$:

$$\hat{\eta}(x,t) = (5x^2 + 2x - 1)t/6 + [\sqrt{6}(5x^2 - 3x - 2)(\bar{y}_1(t,\omega) - \bar{y}_2(t,\omega)) - \\ -(5x^2 + 2x - 7)(\bar{y}_1(t,\omega) + \bar{y}_2(t,\omega))]/12.$$

References

[1] Dubova I.S., Fedorov V.V., Fedorova G.S.: *Selection of optimal paths under time-dependent response, regression experiments*, Moscow, MGU, (1977), 30-38.

[2] Kozlov V.P., Sedunov E.V., Beletsky V.E., Jakhno V.V.: *Optimization of dynamic experiments at switching measurement circuits*, Zavodskaya Laboratoriya, **7**, (1989), 94-99.

[3] Sedunov E.V., Sedunova E.A.: *Unbiased experimental design in inverse problems of mathematical physics*, Proceedings of the 5th St. Petersburg Workshop in Simulation, St. Petersburg, (2005), 611-614.

[4] Sedunov E.V., Sidorenko N.G.: *Unbiased experimental design in Hilbert spaces*, Probability theory and its applications, vol. 32, **4**, (1987), 804-808.

[5] Sidorenko N.G.: *Design and analysis of experiments with vector response under shortage of resources*, Synopsis of Candidates's of PMS dissertation, Leningrad, LGU, (1988), 16.

[6] Ermakov S.M.: *Of optimal unbiased designs of regression experiments*, LOMI's Works, Vol. 111, (1970), 252 - 257.

6th St.Petersburg Workshop on Simulation (2009) 1085-1089

Optimal designs for estimating the linear combination of the coefficients in trigonometric regression models

Holger Dette[1], Viatcheslav B. Melas[2], Petr Shpilev[3]

Abstract

In the common Fourier regression model we investigate the optimal design problem for the estimation linear combination of the coefficients, where the explanatory variable varies in the interval $[-\pi, \pi]$. In a paper Dette et. al. (2008) determined optimal designs for estimating certain pairs of the coefficients in the model. The optimal design problem corresponds to a linear optimality criterion for a specific matrix L. Here we extend these results to more general matrices L. By our results the optimal design problem for a Fourier regression of large degree can be reduced to a design problem in a model of lower degree, which allows the determination of L-optimal designs in many important cases. The results are illustrated by several examples.

1. Introduction

Consider the common Fourier or trigonometric regression model

$$y = \beta^T f(t) = \beta_0 + \sum_{j=1}^{m} \beta_{2j-1} \sin(jt) + \sum_{j=1}^{m} \beta_{2j} \cos(jt) + \varepsilon, \tag{1}$$

where $\beta = (\beta_0, \ldots, \beta_{2m})^T$ denotes the vector of unknown parameters,

$$f(t) = (f_0(t), \ldots, f_{2m}(t))^T = (1, \sin t, \cos t, \ldots, \sin(mt), \cos(mt))^T.$$

is the vector of regression functions. The explanatory variable t varies in the compact interval $[-\pi, \pi]$ and observations under different experimental conditions are assumed to be independent. An approximate design is defined as a probability measure ξ on the design space $[-\pi, \pi]$ with finite support [see Kiefer (1974)]. Note, that for a symmetric design ξ after an appropriate permutation $P \in R^{2m+1 \times 2m+1}$

[1]Ruhr-Universit, E-mail: holger.dette@rub.de
[2]St. Petersburg State University, E-mail: v.melas@pobox.spbu.ru
[3]St. Petersburg State University, E-mail: pitshp@hotmail.com

of the order of the regression functions the information matrix will be block diagonal, that is

$$\widetilde{M}(\xi) = PM(\xi)P = \begin{pmatrix} M_c(\xi) & 0 \\ 0 & M_s(\xi) \end{pmatrix}, \tag{2}$$

where the blocks are given by

$$M_c(\xi) = \left(\int_{-\pi}^{\pi} \cos(it)\cos(jt) d\xi(t) \right)_{i,j=0}^{m}, \tag{3}$$

$$M_s(\xi) = \left(\int_{-\pi}^{\pi} \sin(it)\sin(jt) d\xi(t) \right)_{i,j=1}^{m}. \tag{4}$$

For a given matrix

$$L = \sum_{i=0}^{2m} l_i l_i^T, \tag{5}$$

with vectors $l_i \in \mathbb{R}^{2m+1}$ the class Ξ_L is defined as the set of all approximate designs for which linear combinations of the parameters $l_i^T\beta, i = 0, \ldots, 2m$ are estimable, that is $l_i \in \text{Range}(M(\xi))$; $(i = 0, \ldots, 2m)$. We say that an approximate design η belongs to class Ξ_L^* if $\eta \in \Xi_L$ and for any approximate design ξ such limit is fulfilled

$$\lim_{\alpha \to 0} f^T(t) M^+(\xi_\alpha) L M^+(\xi_\alpha) f(t) = f^T(t) M^+(\eta) L M^+(\eta) f(t),$$

where $\xi_\alpha = (1-\alpha)\eta + \alpha\xi,\ \alpha \in [0,1]$. Finally a designs ξ^* is called L-optimal if

$$\xi^* = \arg\min_{\xi \in \Xi_L} trLM^+(\xi),$$

where L is a fixed and nonnegative definite matrix and for a given matrix A the matrix A^+ is the Moore-Penrose inverse of the Matrix A [see Rao (1968)]. The following result gives a characterization of L-optimal designs, which is particularly useful for determining L-optimal designs with a singular information matrix. The theorem is stated for a general regression model $J = \beta^T f(t) + \varepsilon$ with $2m+1$ regression functions.

Theorem 1. *Let $L \in R^{(2m+1)\times(2m+1)}$ denote a given and nonnegative definite matrix of the form (5) and assume that exist the optimal design $\xi^* \in \Xi_L^*$.*

1) The design ξ is an element of the class Ξ_L if and only if

$$l_i^T M^-(\xi) M(\xi) = l_i^T,\ i = 0, \ldots, 2m.$$

2) The design $\xi \in \Xi_L^*$ is L-optimal if and only if

$$\max_{t \in \chi} \varphi(t, \xi^*) = trLM^+(\xi^*), \tag{6}$$

where $\varphi(t,\xi) = f^T(t) M^+(\xi) L M^+(\xi) f(t)$. Moreover, the equality

$$\varphi(t_i, \xi^*) = trLM^+(\xi^*) \tag{7}$$

holds for any $t_i \in supp(\xi^*)$.

3) Assume that a design $\xi \in \Xi_L$ *but* $\xi \notin \Xi_L^*$ *and these exist an interval* $[x_0; b)$ *and a family of designs* $\{\xi(x)\}$ *such that*

$$\xi(x) \in \Xi_L^* \; for \; x \in (x_0, b), \; \lim_{x \to x_0} \xi(x) = \xi and \lim_{x \to x_0} \max_{t \in \chi} \varphi(t, \xi(x)) = trLM^+(\xi)$$

Then the design is L-optimal.

In general an analytical determination of L-optimal designs is very difficult. However, below we demonstrate that Theorem 1 can be used to check the optimality of a given design. The next theorem is our main result which characterizes the structure of the Moore-Penrose inverse of the information matrix for several symmetric L-optimal design. Also it can be very useful for constructing optimal design for estimating any linear combination of the parameters in the model.

Consider the trigonometric regression model (1) of degree $m = Nk$ and the L-optimal design problem for nonnegative definite matrices $L^{(k)} \in R^{(2Nk+1)\times(2Nk+1)}$ of the form

$$\tilde{L}^{(k)} = P\, L^{(k)} P = \begin{pmatrix} L_{\cos}^{(k)} & 0 \\ 0 & L_{\sin}^{(k)} \end{pmatrix}, \tag{8}$$

where P is the permutation defined by (2) and the matrices $L_{\cos}^{(k)} \in \mathbb{R}^{Nk+1 \times Nk+1}$ and $L_{\sin}^{(k)} \in \mathbb{R}^{Nk \times Nk}$ are matrices with elements $L_{\cos,ij}^{(k)}$ and $L_{\sin,ij}^{(k)}$ are given by

$$L_{\cos,uv}^{(k)} = \begin{cases} L_{\cos,ij}^{(1)} & \text{if} \quad u = (i-1)k+1, \; v = (j-1)k+1 \\ 0 & \text{else} \end{cases} \tag{9}$$

$(u, v = 1, \ldots, Nk+1; \quad i, j = 1, \ldots, N+1)$ and

$$L_{\sin,uv}^{(k)} = \begin{cases} L_{\sin,ij}^{(1)} & \text{if} \quad u = ik; \; v = jk \\ 0 & \text{else} \end{cases} \tag{10}$$

$(u, v = 1, \ldots, Nk)$, respectively. Here $L_{\sin}^{(1)} = (L_{\sin,u,v}^{(1)})_{u,v=1}^{N}$, $L_{\cos}^{(1)} = (L_{\cos,u,v}^{(1)})_{u,v=1}^{N+1}$ are given $N \times N$ and $(N+1) \times (N+1)$ matrices, respectively.

Theorem 2. *Consider trigonometric regression model (1) with* $m = Nk$, $N, k \in \mathbb{N}$, $N \geq 2$, *and let* L *be given by (8) with blocks* $L_{\cos}^{(k)}$ *and* $L_{\sin}^{(k)}$ *defined by (9) and (10), respectively.*

1) *Define the design*

$$\xi_m^{sin} = \begin{pmatrix} -t_m & -t_{m-1} & \ldots & -t_1 & t_1 & \ldots & t_m \\ \omega_m & \omega_{m-1} & \ldots & \omega_1 & \omega_1 & \ldots & \omega_m \end{pmatrix}, \tag{11}$$

where

$$\omega_1 = \frac{z_1}{k}, \; \ldots, \; \omega_{\frac{N}{2}-1} = \frac{z_{\frac{N}{2}-1}}{k}, \; \omega_{\frac{N}{2}} = \frac{z_{\frac{N}{2}}}{k}, \; \omega_{\frac{N}{2}+1} = \frac{z_{\frac{N}{2}}}{k}, \; \ldots, \; \omega_N = \frac{z_1}{k},$$

$$\sum_{j=1}^{\frac{N}{2}} z_j = \frac{1}{4}, \; t_1 = \frac{x_1}{k}, \; \ldots, t_{\frac{N}{2}} = \frac{x_{\frac{N}{2}}}{k}, t_{\frac{N}{2}+1} = \frac{\pi - x_{\frac{N}{2}}}{k}, \ldots, \; t_N = \frac{\pi - x_1}{k},$$

$$\omega_i = \omega_{i-N}, \; t_i = t_{i-N} + \frac{\pi}{k}, \; i = N+1, \ldots, m$$

if N is even , and

$$\omega_1 = \frac{z_1}{k},\ \ldots,\ \omega_{\frac{N-1}{2}} = \frac{z_{\frac{N-1}{2}}}{k},\ \omega_{\frac{N+1}{2}} = \frac{1}{2k} - 2\sum_{j=1}^{\frac{N-1}{2}} \frac{z_j}{k},\ \omega_{\frac{N+3}{2}} = \frac{z_{\frac{N-1}{2}}}{k}, \ldots,$$

$$\omega_N = \frac{z_1}{k};\ t_1 = \frac{x_1}{k},\ \ldots, t_{\frac{N-1}{2}} = \frac{x_{\frac{N-1}{2}}}{k},\ t_{\frac{N+1}{2}} = \frac{\pi}{2k},\ t_{\frac{N+3}{2}} = \frac{\pi - x_{\frac{N-1}{2}}}{k}, \ldots,$$

$$t_N = \frac{\pi - x_1}{k};\ \omega_i = \omega_{i-N},\ t_i = t_{i-N} + \frac{\pi}{k},\ i = N+1, \ldots, m$$

if N is odd. If the matrix L is of the form (8) with $L_{\cos}^{(k)} = 0$, then for the design ξ_n^{sin}
the quantities $trL_{\sin}^{(k)} M_s^+(\xi_n^{sin})$ and the coefficients of the function

$$\varphi(t, \xi_n^{sin}) = f_s^T(t) M_s^+(\xi_n^{sin}) L_{\sin}^{(k)} M_s^+(\xi_n^{sin}) f_s(t)$$

are independent of the value k for any matrix $L_{\sin}^{(k)} \in \mathbb{R}^{m \times m}$.
2) *Define the design*

$$\xi_n^{cos} = \begin{pmatrix} -\pi & -t_{n-1} & \ldots & -t_1 & 0 & t_1 & \ldots & t_{n-1} & \pi \\ \omega_n - \alpha & \omega_{n-1} & \ldots & \omega_1 & \omega_0 & \omega_1 & \ldots & \omega_{n-1} & \alpha \end{pmatrix}, \qquad (12)$$

where $\alpha \in [0, \omega_n],\ n = (N+1)k$ and

$$t_0 = 0,\ t_1 = \frac{x_1}{k},\ \ldots, t_{\frac{N}{2}} = \frac{x_{\frac{N}{2}}}{k}, t_{\frac{N}{2}+1} = \frac{\pi - x_{\frac{N}{2}}}{k}, \ldots,\ t_N = \frac{\pi - x_1}{k},$$

$$\omega_0 = \frac{1 - 4\sum_{i=1}^{\frac{N}{2}} z_i}{2k},\ \omega_1 = \frac{z_1}{k},\ \ldots,\ \omega_{\frac{N}{2}-1} = \frac{z_{\frac{N}{2}-1}}{k},\ \omega_{\frac{N}{2}} = \frac{z_{\frac{N}{2}}}{k},$$

$$\omega_{\frac{N}{2}+1} = \frac{z_{\frac{N}{2}}}{k},\ \ldots,\ \omega_N = \frac{z_1}{k},\ \ \omega_{i+1} = \omega_{i-N},\ t_{i+1} = t_{i-N} + \frac{\pi}{k},\ i = N, \ldots, n-1$$

if N is even , and

$$t_0 = 0,\ t_1 = \frac{x_1}{k},\ \ldots, t_{\frac{N-1}{2}} = \frac{x_{\frac{N-1}{2}}}{k},\ t_{\frac{N+1}{2}} = \frac{\pi}{2k},$$

$$t_{\frac{N+3}{2}} = \frac{\pi - x_{\frac{N-1}{2}}}{k}, \ldots,\ t_N = \frac{\pi - x_1}{k},\ t_{i+1} = t_{i-N} + \frac{\pi}{k},\ i = N, \ldots, n-1,$$

$$\omega_0 = z_0,\ \omega_1 = z_1,\ \ldots,\ \omega_{\frac{N-1}{2}} = z_{\frac{N-1}{2}},\ \omega_{\frac{N+1}{2}} = \frac{1}{2k} - \sum_{j=0}^{\frac{N-1}{2}} \frac{z_j}{k},$$

$$\omega_{\frac{N+3}{2}} = z_{\frac{N-1}{2}}, \ldots,\ \omega_N = z_1,\ \ \omega_{i+1} = \omega_{i-N},\ i = N, \ldots, m$$

if N is odd . If the matrix L is of the form (8) with $L_{\sin}^{(k)} = 0$, then for the design ξ_n^{cos} the quantities $trL_{\cos}^{(k)} M_c^+(\xi_n^{cos})$ and the coefficients of the function $\varphi(t, \xi_n^{cos}) = f_c^T(t) M_c^+(\xi_n^{cos}) L_{\cos}^{(k)} M_c^+(\xi_n^{cos}) f_c(t)$ are independent of the value k for any matrix $L_{\cos}^{(k)} \in \mathbb{R}^{m+1 \times m+1}$.

The proof of Theorem 2 and the technique of the expansion $\theta^*(r,\omega)$ into Taylor series can be found in [6].

Note that the designs in (11) and (12) are determined by the parameters $x_1, x_2, \ldots$ and $z_1, z_2, \ldots$, which usually have to be determined numerically. Theorem 2 is a very useful instrument for finding L-optimal designs, because it allows to reduce the optimal design problem for the trigonometric regression model (1) to a design problem in a model of substantially smaller degree. As a consequence the L-optimal design problem simplifies sufficiently. We will now illustrate its application in a concrete example.

Example 1. *L-optimal design for estimating the coefficients of* $\sin(40t)$ *and* $\sin(50t)$ *in the Fourier regression model of degree* $m = 50$.

We consider the trigonometric regression model (1) of degree $m = Nk$ and use Theorem 2 to determine the L-optimal design for estimating the pair of coefficients $\beta_{2(N-1)k-1}$ and β_{2Nk-1}, which correspond to the terms $\sin((N-1)kt)$ and $\sin(Nkt)$). Exemplarily we will consider the case $m = 50$, $N = 5$, $k = 10$. Theorem 2 allows to reduce the problem of constructing L-optimal design for estimating the pair of coefficients β_{79} and β_{99} for the model degree $m = 50$ to the problem of constructing L-optimal design for estimating the pair of coefficients β_7 and β_9 for the model degree $m = 5$. Now we can find this L-optimal design as a solution of the system:

$$\begin{cases} \frac{\partial tr L_{sin}^{(1)} M_s^+(\xi_5^{sin})}{\partial x_i} = 0 \\ \frac{\partial tr L_{sin}^{(1)} M_s^+(\xi_5^{sin})}{\partial z_i} = 0 \end{cases},$$

where ξ_5^{sin} is a symmetric design with a structure is defined in Theorem 1. Thus we have

$$\xi_5^{sin} = \begin{pmatrix} -\pi + x_1 & -\pi + x_2 & -\frac{\pi}{2} & -x_2 & -x_1 & x_1 & \ldots & \pi - x_1 \\ z_1 & z_2 & \frac{1-4z_1-4z_2}{2} & z_2 & z_1 & z_1 & \ldots & z_1 \end{pmatrix},$$

where x_1=0.3519978036, x_2=1.020599385, z_1=0.1112542423, z_2=0.09865639359.

Now from Theorem 2 follows that the L-optimal design for estimating coefficients $\beta_{2(N-1)k-1}$ and β_{2Nk-1} (i.e. β_{79} and β_{99} in our case) is given by

$$\xi_{Nk}^{sin} = \xi_{50}^{sin} = \begin{pmatrix} -t_{50} & -t_{49} & \ldots & -t_1 & t_1 & \ldots & t_{50} \\ \omega_{50} & \omega_{49} & \ldots & \omega_1 & \omega_1 & \ldots & \omega_{50} \end{pmatrix},$$

$$t_1 = \frac{x_1}{10},\ t_2 = \frac{x_2}{10},\ t_3 = \frac{\pi}{20},\ t_4 = \frac{\pi - x_2}{10},\ t_5 = \frac{\pi - x_1}{10}, \tag{13}$$

$$t_i = t_{i-5} + \frac{\pi}{10},\ i = 6, \ldots, 50, \tag{14}$$

$$\omega_1 = \frac{z_1}{10},\ \omega_2 = \frac{z_2}{10},\ \omega_3 = \frac{1}{20} - \frac{z_1}{5} - \frac{z_2}{5},\ \omega_4 = \frac{z_2}{10},\ \omega_5 = \frac{z_1}{10}, \tag{15}$$

$$\omega_i = \omega_{i-N},\ i = N+1, \ldots, m \tag{16}$$

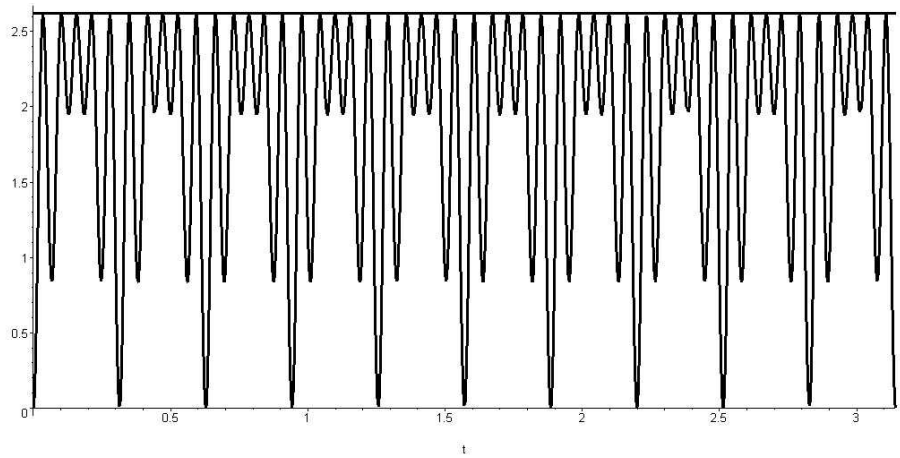


Figure 1: *The function $\varphi(t, \xi^*_{79,99})$ defined in the equivalence Theorem 1 for the L-optimal design problem discussed in Example 1.*

where x_1, x_2, z_1, z_2 is defined above. We finally note that a straightforward calculation yields for the function $\varphi(t, \xi^*_{79,99})$ in the equivalence Theorem 1

$$\begin{aligned}\varphi(t, \xi^*_{79,99}) &= f^T(t) M^+(\xi^*_{79,99}) L M^+(\xi^*_{79,99}) f(t) = \\ &= 0.2 \sin^2(30t) - 1.047213603 \sin(50t) \sin(30t) \\ &\quad + 1.370820397 \sin^2(50t) + 2.094427187 \sin^2(40t).\end{aligned}$$

This function is depicted in Figure 1.

References

[1] *Kiefer, J.C.* (1974). General equivalence theory for optimum designs (approximate theory). The Annals of Statistics 2, 849-879.

[2] *Rao, C.R.* (1968). Linear statistical inference and its applications. Wiley, New York.

[3] *Pukelsheim, F.* (2006). Optimal Design of Experiments. Wiley, New York.

[4] *Dette, H. and Melas, V.B.* (2003). Optimal designs for estimating individual coefficients in Fourier regression models. The Annals of Statistics, Vol. 31, 1669-1692.

[5] *Dette, H., Melas, V.B. and Shpilev, P.V.* (2008) Optimal designs for estimating the pairs of coefficients in Fourier regression models. To appear in: Statistica Sinica.

[6] *Dette, H., Melas, V.B. and Shpilev, P.V.* (2009) Optimal designs for trigonometric regression models. Submitted to JSPI.

6th St.Petersburg Workshop on Simulation (2009) 1091-1095

Improvement of random LHD for high dimensions[1]

Andrey Pepelyshev[2]

Abstract

Designs of experiments for multivariate case are reviewed. Fast algorithm of construction of good Latin hypercube designs is developed.

1. Introduction

The mathematical theory for designing experiments was started to develop by Sir Ronald A. Fisher who pioneered the design principles in his studies of analysis of variance originally in agriculture. The theory of experimental designs have received considerable development further in the middle of the twentieth century in works by G.E.P Box, J. Kiefer and many others. Computer experiments have become available with the appearance of computer engineering. Mathematical computer models are a replacement for natural (physical, chemical, biological) experiments which are too time consuming or too costly. Moreover, mathematical models may describe phenomena which can not be reproduced, for example, weather modeling.

Experimental designs for deterministic computer models was studied first by McKay et al. (1979). The theoretical principles of analysis of deterministic computer models were determined in Sacks et al. (1989) and the analysis of simulation models (deterministic computer codes with stochastic output) in Kleijnen (1987). During the last decade the Bayesian approach to computer experiments was extensively developed, see Kennedy, O'Hagan (2001), Conti, O'Hagan (2008) and references within. The technique used in the Bayesian approach is close to Kriging in a manner that a special construction is used to interpolate the values of the output of the deterministic code rather than the values of a random field and uncertainty intervals for untried values of inputs are calculated; see Koehler, Owen (1996), Kennedy, O'Hagan (2001). One run by a computer model may require considerable time. Thus the main problem is to reduce an uncertainty of inferences on a computer model by making only a few runs. Consequently, we are faced with the problem of optimal choice of experimental conditions.

The present paper is organized as follows. In Section 2 we review experimental designs for a multivariate case in order to choose the most appropriate criteria of optimality. In Section 3 we propose a fast algorithm for constructing good optimal designs for computer experiments.

[1]This work was partly supported by RFBR, project No 09-01-00508.

[2]University of Sheffield, E-mail: a.pepelyshev@sheffield.ac.uk, St.Petersburg State University

2. Comparison of natural and computer experiments

Basic features of natural and computer experiments are presented in the following two-column style.

Natural experiments	Computer experiments
The response is observed with errors which may be correlated.	The output is deterministic. The running of a computer code at the same inputs gives the same output.
The response is described either by known regression function with unknown parameters or by multivariate linear or quadratic model which is valid at a design subspace.	A computer code is considered to be like as a black box. The main assumption is factor sparsity, that is the output depends in nonlinear way only on a few number of inputs[1].
A primary objective is to estimate parameters or find conditions which maximize response.	A primary objective is to fit a cheaper unbiased low uncertain predictor.
Other aims are identifying variables which have a significant effect, etc.	Other aims are calibration of model parameters to physical data, optimization of output, etc.
Optimal design is typically to minimize the (generalized) variance of estimated characteristics.	Optimality criteria is the minimization of mean square error over design space or the maximization of entropy.
Optimal designs are, for example, factorial, incomplete block, orthogonal, central composite, screening and D-optimal designs.	Optimal design is space-filling design. Latin hypercube design is recommended in many papers.

Note that optimal designs for natural experiments mostly have two or three points in projection on each coordinate, e.g. the 2^{k-p} block and orthogonal designs have two points in projection, central composite design has three points in projection. This fact is a consequence of the multivariate linear or quadratic model which is assumed to be valid. Such designs are not suitable for computer experiments since we assume that the output may be highly nonlinear in several variables. Due to the objectives of computer experiments, optimal design should minimize mean square error between the prediction of response at untried inputs and the true output. This criterion leads to the optimal design which should fill an entire design space uniformly at the initial stage of computer experiments. The examples of space-filling design are Latin hypercube design, sphere packing design, distance based design, uniform design, design based on random or pseudo-random sequences, see Santner et al. (2003), Fang et al. (2006). The optimal design should be a dense set in projection to each coordinate and should be a dense set in entire design space. Each of the above space-filling designs has attractive properties and satisfies some useful criterion. As far as is known, the best design should optimize a compound criterion.

[1]Without sparsity assumption we need a lot of runs to construct an unbiased low uncertain predictor.

3. Latin Hypercube Designs

At first, we need to recall an algorithm for construction of LH designs, which was introduced in McKay et al. (1979). The algorithm generates n points in dimension d in the following manner. 1) Generate n uniform equidistant points $x_1^{(s)}, \ldots, x_n^{(s)}$ in the range of each input, $s = 1, \ldots, d$. 2) Generate a matrix $(p_{i,j})$ of size $d \times n$ such that each row is a random permutation of numbers $1, \ldots, n$ and these permutations are independent. 3) Each column of the matrix $(p_{i,j})$ corresponds to a design point, that is $(x_{p_{1,j}}^{(1)}, \ldots, x_{p_{d,j}}^{(d)})^T$ is jth point of LHD.

Without loss of generality, we assume that the range of each input is $[0, 1]$ and $x_j^{(s)} \in \mathcal{R} = \{0, 1/(n-1), 2/(n-1), \ldots, 1\}$.

By construction, LHD has the best filling of range in projection on each coordinate. Unfortunately, LHD may have a poor filling of entire hypercube. Several criteria of optimality are introduced in order to choose a good LHD in a class of all LHD. Maximin criterion is a maximization of minimal distance

$$\Psi_p(L) = \min_{i,j=1,\ldots,n} ||x_i - x_j||_p = \min_{i,j=1,\ldots,n} \left(\sum_{s=1}^{d} |x_{s,i} - x_{s,j}|^p \right)^{1/p}$$

usually used with $p = 2$ where $x_i = (x_{1,i}, \ldots, x_{d,i})^T$ is ith point of design L. An LHD which maximize $\Psi_p(L)$ is called by maximin LHD. Audze-Eglais criterion introduced in Audze, Eglais (1977) is a sum of forces between charged particles and is a minimization of

$$\Psi_{AE}(L) = \sum_{i=1}^{n} \sum_{j=i+1}^{n} \frac{1}{||x_i - x_j||_2^2}.$$

Others criteria of uniformity are star L_2-discrepancy, centered L_2-discrepancy, wrap-around L_2-discrepancy which are motivated by quasi-Monte-Carlo methods and the Koksma-Hlawka inequality, see Hickernell (1998), Fang et al. (2000). Algorithms of optimization are studied in a number of papers, the local search algorithm in Grosso et al. (2008), the enhanced stochastic evolutionary algorithm in Jin et al. (2005), the simulated annealing algorithm in Morris, Mitchell, (1995) the columnwise-pairwise procedure in Ye et al. (2000), the genetic algorithm in Liefvendahl, Stocki, (2006) and Bates et al. (2003), the collapsing method in Fang, Qin (2003). Cited authors concentrate on the case of low dimensions.

Basing on an analysis of papers on computer experiments, we can say that the size of LHD is approximately equal to the input dimension multiplied by 10, that is $n \approx 10d$. Further we propose a fast algorithm of constructing good LHD for the case of high dimensions which is not studied, to the best of our knowledge.

First, we need to study features of random LHD generated by the above algorithm. Let $L = \{x_1, \ldots, x_n\}$ be a LHD. Let r_i be a minimal distance between x_i and other points of L; that is $r_i = \min_{j \neq i} ||x_i - x_j||_2$ (further we consider euclidian distances). These distances characterize a design L. Let Q_α denote an α-percentile for sample $r_1, \ldots, r_n$. Averaged values of low and upper quartiles, $Q_{0.25}$ and $Q_{0.75}$, are presented in table 1. We see that the inter-point distances are varied and the quarter of distances are quite small. Also note that distances between points is increased as the dimension is increased since $n = 10d$.

Table 1: Low and upper quartiles of distances between points of n-point random LHD for different dimensions, $n = 10d$.

d	2	3	4	5	6	7
$Q_{0.25}$	0.108	0.167	0.232	0.305	0.368	0.434
$Q_{0.75}$	0.175	0.270	0.347	0.431	0.502	0.573
d	8	9	10	14	20	
$Q_{0.25}$	0.494	0.554	0.610	0.821	1.096	
$Q_{0.75}$	0.636	0.699	0.757	0.972	1.249	

For construction of n-point LHD with a given inter-point distance r at dimension d, we propose the following heuristic algorithm.

Algorithm.

a) Let L_k is a k-point design at kth step. Let $L_1 = \{x_1\}$ where x_1 is a random point in the middle of $\mathcal{R}^d$ such that its coordinates are unequal to each other.

b) Compute a boolean matrix $B = \{B_{i,j}\}$ of size $d \times n$ upon L_k such that $b_{i,j} = 1$ ('used') if there exists a point in L_k with ith coordinate which equals $(j-1)/(n-1)$, and 0 ('unused') otherwise.

c) Generate a random point $z = (q_1, \ldots, q_d)^T/(n-1) \in \mathcal{R}^d$ such that each coordinate is unused; that is $b_{i,q_i} = 0$, $i = 1, \ldots, d$ except one random coordinate which should be taken nearby 0.5.

d) Create a set C of candidate points in $\mathcal{R}^d$ with unused coordinates which are approximate points which are the closest and the furthest point from z lied on spheres $S_r(x_j)$ with centers $x_j \in L_k$ and radius r, $j = 1, \ldots, k$.

e) Find a point $x^* \in C$ such that x^* lies outside of all $S_r(x_j)$, that is $||x^* - x_j|| > r$, $j = 1, \ldots, k$. If there exist several such points, choose a point which minimizes $\#\{s : \sum_{i=1}^{d} B_{i,s} = m^*\}$, where $m^* = \min_{j=1,\ldots,n} \sum_{i=1}^{d} B_{i,j}$ and $B = B(L_k \bigcup \{x^*\})$.

f) Add x^* to design, that is $L_{k+1} = L_k \bigcup \{x^*\}$. Stop at nth step.

g) If we could not find x^* at step 5, go to step 3. If we could find x^* after several trials, we should decrease r since it is impossible to find a point which is far from L_k at given distance r.

Let a design obtained by Algorithm be called SLHD. Numerical results show that Algorithm is fast and work well for any dimension. It requires 40 seconds to compute 100-point SLHD at dimension $d = 10$ and 60 seconds for 140-point SLHD at $d = 14$ and 200 seconds for 200-point SLHD at $d = 20$ on PC 2.1GHz. The choice of r should be smaller than r^* where r^* is the minimal distance between points of exact maximin LHD. Since r^* is unknown, we recommend the running Algorithm with different r, say, start with $Q_{0.75}$ for random LHD and increase it by small increment. The decreasing of r at step 7 does not mean that SLHD does not exist for given r and is a consequence a poor placement of points at previous iterations.

Table 2: Low and upper quartiles and $Q_{0.1}$ of distances between points of n-point SLHD for different dimensions, $n = 10d$. The value r^* of maximin LHD is given.

d	2	3	4	5	6	7
$Q_{0.1}$	0.217	0.310	0.351	0.393	0.512	0.584
$Q_{0.25}$	0.217	0.312	0.363	0.476	0.535	0.617
$Q_{0.75}$	0.217	0.323	0.409	0.486	0.552	0.626
r^*	0.223	0.360	0.476	0.589	0.687	0.779
d	8	9	10	14	20	
$Q_{0.1}$	0.679	0.763	0.823	1.035	1.268	
$Q_{0.25}$	0.694	0.765	0.824	1.037	1.271	
$Q_{0.75}$	0.706	0.774	0.836	1.045	1.281	
r^*	0.867	0.950	1.021	-	-	

Features of SLHD are presented in Table 2. Values of r^* are taken from website http://www.spacefillingdesigns.nl/. We see that 90% of inter-point distances of SLHD are higher than the most of distances at random LHD. Thus SLHD has a better filling of entire design space. Figure 1 display points of SLHD for $d = 2$ and $d = 20$. We can see a quite uniform filling of square. Further improvement of experimental design can be done by applying the local search or the simulated annealing algorithm.

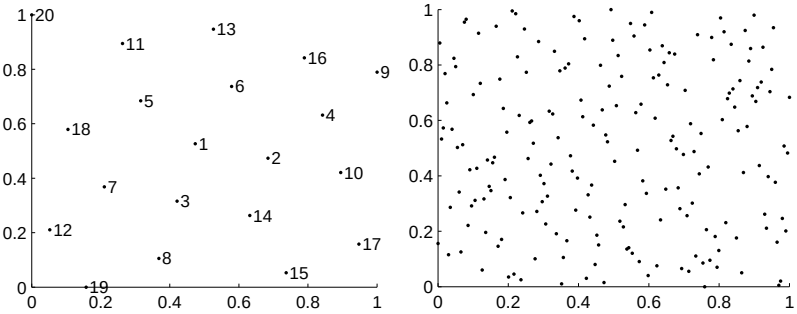


Figure 1: The points of SLHD for $d = 2$ with the order of including (left) and two coordinates of points of SLHD for $d = 20$ (right).

4. Conclusion

The algorithm of construction of LHD with given inter-point distance is constructed and studied. By the algorithm we can quickly compute LHD such that the most of inter-point distances are larger than distances at random LHD. The proposed algorithm is more efficient than simply generate many random LHDs and choose the best one.

References

[1] Bates S.J., Sienz J., Langley D.S. Formulation of the AudzeEglais Uniform Latin Hypercube design of experiments. Advances in Engineering Software 34, (2003) 493–506.

[2] Conti S., O'Hagan A. Bayesian emulation of complex multi-output and dynamic computer models. J. Statist. Plan. Infer. (2008) To appear.

[3] Jin R., Chen, W., Sudjianto A. An efficient algorithm for constructing optimal design of computer experiments. J. Statist. Plann. Inf. 134 (2005), 268–287.

[4] Grosso A., Jamali A., Locatelli M. Finding maximin latin hypercube designs by Iterated Local Search heuristics. accepted to European J. Operational Research. (2008).

[5] Fang K.-T., Qin H. A note on construction of nearly uniform designs with large number of runs. Statist. Probab. Lett. 61 (2003), no. 2, 215–224.

[6] Fang K.-T., Lin D.K.J., Winker P., Zhang Y. Uniform design: theory and application. Technometrics 42 (2000), no. 3, 237–248.

[7] Fang K.-T. Li R., Sudjianto A. Design and modeling for computer experiments. Chapman & Hall/CRC, (2006).

[8] Fang K.-T., Ma C.-X., Winker P. Centered L_2-discrepancy of random sampling and Latin hypercube design, and construction of uniform designs. Math. Comp. 71 (2002), no. 237, 275–296.

[9] Hickernell F.J. A generalized discrepancy and quadrature error bound. Math. Comp. 67 (1998), no. 221, 299–322.

[10] Kennedy M.C., O'Hagan A. Bayesian calibration of computer models. J. R. Stat. Soc. Ser. B 63 (2001), no. 3, 425–464.

[11] Kleijnen J.P.C. Statistical tools for simulation practitioners. (1986)

[12] Koehler J.R., Owen A.B. Computer experiments. In Handbook of Statistics, (1996), 261–308.

[13] Liefvendahl M., Stocki R. A study on algorithms for optimization of Latin hypercubes. J. Statist. Plann. Inference 136 (2006), 3231–3247.

[14] McKay M. D., Beckman R. J., Conover W. J. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. Technometrics 21 (1979), no. 2, 239–245.

[15] Morris M.D., Mitchell T.J. Exploratory designs for computer experiments J. Stat. Plan. Inf., 43 (1995), 381–402.

[16] Sacks J., Welch W.J., Mitchell T.J., Wynn H.P. Design and analysis of computer experiments. With comments and a rejoinder by the authors. Statist. Sci. 4 (1989), no. 4, 409–435.

[17] Santner T.J., Williams B.J., Notz W. The Design and Analysis of Computer Experiments. (2003).

[18] Ye K.Q., Li W., Sudjiantoc A. Algorithmic construction of optimal symmetric Latin hypercube designs. J. Stat. Plan. Inf. 90, (2000), 145–159.

6th St.Petersburg Workshop on Simulation (2009) 1097-1101

Optimality of Equidistant Sampling Designs for a Nonstationary Ornstein-Uhlenbeck Process[1]

Radoslav Harman[2], František Štulajter[3]

Abstract

We show that the equidistant sampling designs are optimal for parametric estimation as well as for prediction in the case of a nonstationary Ornstein-Uhlenbeck process with one unknown parameter.

1. Introduction

Assume the Ornstein-Uhlenbeck (OU) process $\tilde{X}(\cdot) = \left\{\tilde{X}(t) : t \geq 0\right\}$ given by

$$\tilde{X}(t) = v_0 e^{-\lambda t} + \bar{v}(1 - e^{-\lambda t}) + \sigma \int_0^t e^{-\lambda(t-s)} dB(s), \tag{1}$$

where $B(\cdot)$ is the Brownian motion. The process (1) can be interpreted as the velocity of a moving particle, where v_0 is the initial velocity, $\bar{v}$ is the asymptotic mean velocity, and λ, σ are positive coefficients related to the physical characteristics of the environment and the particle (see, e.g., [5], [2]). In this paper we will assume that the constants λ and σ^2 are known, while exactly one of the values v_0 and $\bar{v}$ is unknown and needs to be estimated from observations of the process taken in $n \geq 2$ time instants. The problem with both v_0 and $\bar{v}$ unknown is more complex and will be covered by a separate paper.

It is a well known fact that the mean value and the covariance function of the process $\tilde{X}(\cdot)$ is given by (see, e.g., [3])

$$\begin{aligned} E[\tilde{X}(t)] &= v_0 e^{-\lambda t} + \bar{v}\left(1 - e^{-\lambda t}\right) \text{ for } 0 \leq t, \text{ and} \\ R(s,t) &= \frac{\sigma^2}{2\lambda}\left(e^{-\lambda(t-s)} - e^{-\lambda(t+s)}\right) \text{ for } 0 \leq s \leq t. \end{aligned} \tag{2}$$

Note that as the limit case $t, s \to \infty$ we obtain the standard stationary OU process with the mean value $\bar{v}$ and the covariance function

$$R_\infty(s,t) = \frac{\sigma^2}{2\lambda} e^{-\lambda(t-s)} \text{ for } 0 \leq s \leq t. \tag{3}$$

[1]This work was supported by the Slovak VEGA-Grant No. 1/0077/09.

[2]Faculty of Mathematics, Physics and Informatics, Comenius University Bratislava, E-mail: harman@fmph.uniba.sk

[3]Faculty of Mathematics, Physics and Informatics, Comenius University Bratislava, E-mail: stulajter@fmph.uniba.sk

Suppose that we can observe the process $\tilde{X}(\cdot)$ in n distinct times chosen in the experimental domain $[T_*, T^*]$, where $0 < T_* < T^*$. By an n-point design τ we will call an n-dimensional vector $\tau = (t_1, ..., t_n)'$ of strictly increasing values ("sampling times") from $[T_*, T^*]$, i.e., $T_* \leq t_1 < ... < t_n \leq T^*$. We will denote the set of all such designs τ by $\mathcal{T}_n$. In the paper we focus on the problem of choosing the n-point sampling design in an optimal way with respect to estimation of the unknown parameter or prediction of the future value of the process.

For a fixed design $\tau = (t_1, ..., t_n)' \in \mathcal{T}_n$ and for a function $h : [T_*, T^*] \to \mathbb{R}$, we will use the notation $\mathbf{h}(\tau)$ to denote the vector $(h(t_1), ..., h(t_n))'$. Specifically, by $e^{-\lambda\tau}$ we will denote the vector $(e^{-\lambda t_1}, ..., e^{-\lambda t_n})'$ and by $\mathbf{1}_n$ we will denote the vector $(1, 1, ..., 1)' \in \mathbb{R}^n$.

The main aim of the paper is to study the model (1) if one of the parameters v_0 and $\bar{v}$ is unknown. Thus, we will analyze the regression model with a one-dimensional parameter β, in which the vector of observations under the design $\tau \in \mathcal{T}_n$ satisfies

$$\mathbf{X}(\tau) = (a\mathbf{1}_n + be^{-\lambda\tau})\beta + \varepsilon(\tau);\ \beta \in \mathbb{R}, \tag{4}$$

where a and b are given constants not simultaneously equal to 0. Notice that if the initial velocity v_0 is known, $a = 1$, $b = -1$ and $\beta = \bar{v}$, then $\mathbf{X}(\tau)$ corresponds to the vector of observations of the process $\{\tilde{X}(t) - v_0 e^{-\lambda t} : t \geq 0\}$. Similarly, if the asymptotic mean velocity $\bar{v}$ is known, $a = 0$, $b = 1$, and $\beta = v_0$, then $\mathbf{X}(\tau)$ is the vector of observations of the process $\{\tilde{X}(t) - \bar{v}(1 - e^{-\lambda t}) : t \geq 0\}$. Moreover, the stationary (limit) case with $a = 1$, $b = 0$ specifies to the standard OU process that has recently been extensively studied (see [1], [4], and [7]).

Assume a linear regression model with a parameter β and a regression function $f : [0, \infty) \to \mathbb{R}$ where, under the design $\tau = (t_1, ..., t_n)' \in \mathcal{T}^n$, the n-dimensional random vector $\mathbf{X}(\tau)$ of observations satisfies

$$\mathbf{X}(\tau) = \mathbf{f}(\tau)\beta + \epsilon(\tau); \beta \in \mathbb{R}, \tag{5}$$

and $\epsilon(\tau)$ has zero mean value and a regular covariance matrix $\Sigma(\tau)$.

Under the design $\tau \in \mathcal{T}_n$, the weighted least squares estimator of β is

$$\begin{aligned} \beta^*(\tau) &= M^{-1}(\tau)\mathbf{f}'(\tau)\Sigma^{-1}(\tau)\mathbf{X}(\tau), \text{ where} \\ M(\tau) &= \mathbf{f}'(\tau)\Sigma^{-1}(\tau)\mathbf{f}(\tau) \end{aligned} \tag{6}$$

is the information about the parameter β. Since $M^{-1}(\tau) = \mathrm{Var}[\beta^*(\tau)]$, a design $\tau^*_{n,\beta}$ is said to be optimal for estimating the parameter β, if it maximizes $M(\tau)$ among all designs from the set $\mathcal{T}_n$.

An alternative aim of inference can be prediction of the process to a future time $T_d = T^* + d$, where $d > 0$. The best linear unbiased predictor (BLUP) of the process in the time T_d is given by (see, e.g., [8])

$$\begin{aligned} X^*(\tau, d) &= f(T_d)\beta^*(\tau) + \mathbf{r}(\tau, d)\Sigma^{-1}(\tau)\left(\mathbf{X}(\tau) - \mathbf{f}(\tau)\beta^*(\tau)\right), \text{ where} \\ \mathbf{r}(\tau, d) &= (R(t_1, T_d), ..., R(t_n, T_d))'. \end{aligned}$$

The mean squared error (MSE) of $X^*(\tau, d)$ is

$$MSE[X^*(\tau, d)] = \mathrm{Var}\left[X(T_d)\right] - \mathbf{r}'(\tau, d)\Sigma^{-1}(\tau)\mathbf{r}(\tau, d) + \frac{c^2(\tau, d)}{M(\tau)}, \tag{7}$$

where

$$c(\tau, d) = f(T_d) - \mathbf{f}'(\tau)\Sigma^{-1}(\tau)\mathbf{r}(\tau, d).$$

Therefore, a design $\tau^*_{n,d} \in \mathcal{T}_n$ will be called an n-point prediction optimal design if it minimizes $MSE[X^*(\tau, d)]$ among all designs from $\mathcal{T}_n$.

Let $t_1 \in [T_*, T^*)$. By $\tau_n(t_1)$ we denote the n-point equidistant sampling design with the initial design point t_1 and the final design point T^*, i.e., $(\tau_n(t_1))_i = t_1 + (T^* - t_1)(i-1)/(n-1)$ for all $i = 1, ..., n$. In this paper we will show that for the OU model (4) the class $\{\tau_n(t_1) : T^* \leq t_1 < T^*\}$ is optimal in a very broad sense; see Theorem 1.

2. Consequences of the product covariance structure

Notice that the covariance function (2) has the product form

$$R(s, t) = u(s)v(t) \text{ for all } 0 \leq s \leq t < \infty, \tag{8}$$

where

$$u(s) = (e^{\lambda s} - e^{-\lambda s})/2, \text{ and } v(t) = \sigma^2\lambda^{-1}e^{-\lambda t}. \tag{9}$$

It turns out that the product structure of a covariance function, which has already been used in the context of optimal design in the paper [6], simplifies calculation of the information corresponding to a design τ, as well as the MSE of the BLUP. In this section we will derive an expression for the information and the MSE corresponding to a fixed design, and a general covariance structure (8). Let $\tau = (t_1, ..., t_n)' \in \mathcal{T}_n$ and let the covariance matrix $\Sigma(\tau)$ be derived from a covariance function with a product structure (8), that is

$$(\Sigma(\tau))_{i,j} = u(t_i)v(t_j);\ i \leq j. \tag{10}$$

Denote $\Sigma_{(n)} = \Sigma(\tau)$ and let $\Sigma_{(n)}$ as well as its upper-left $m \times m$ submatrices $\Sigma_{(m)}$ be regular. Denote $\mathbf{u}_{(n)} = \mathbf{u}(\tau)$ and let $\mathbf{u}_{(m)}$ be the vector formed by the first m components of $\mathbf{u}_{(n)}$. Also, let $v_i = v(t_i)$ and $u_i = u(t_i)$ for all $i = 1, ..., n$. Using this notation we can write:

$$\Sigma_{(n)} = \begin{pmatrix} \Sigma_{(n-1)} & v_n\mathbf{u}_{(n-1)} \\ v_n\mathbf{u}'_{(n-1)} & u_nv_n \end{pmatrix}. \tag{11}$$

Equation (11) and the formula for the inverse of a partitioned matrix yield:

$$\Sigma^{-1}_{(n)} = \begin{pmatrix} \Sigma^{-1}_{(n-1)} + s_n^{-1}v_n^2\mathbf{h}_{(n-1)}\mathbf{h}'_{(n-1)} & -s_n^{-1}v_n\mathbf{h}_{(n-1)} \\ -s_n^{-1}v_n\mathbf{h}'_{(n-1)} & s_n^{-1} \end{pmatrix}, \tag{12}$$

where

$$\mathbf{h}_{(n-1)} = \Sigma^{-1}_{(n-1)}\mathbf{u}_{(n-1)} = \left(0, ..., 0, v_{n-1}^{-1}\right)', \text{ and}$$

$$s_n = u_nv_n - v_n^2\mathbf{u}'_{(n-1)}\Sigma^{-1}_{(n-1)}\mathbf{u}_{(n-1)} = u_nv_n - v_n^2\frac{u_{n-1}}{v_{n-1}} = v_n^2\left(\frac{u_n}{v_n} - \frac{u_{n-1}}{v_{n-1}}\right).$$

Let $n \geq 2$. Let $\mathbf{x}_{(n-1)} = (x_1, ..., x_{n-1})' \in \mathbb{R}^{n-1}$ be the subvector of a vector $\mathbf{x}_{(n)} = (x_1, ..., x_{n-1}, x_n)' \in \mathbb{R}^n$, and let $\mathbf{y}_{(n-1)} = (y_1, ..., y_{n-1})' \in \mathbb{R}^{n-1}$ be the subvector of a vector $\mathbf{y}_{(n)} = (y_1, ..., y_{n-1}, y_n)' \in \mathbb{R}^n$. Using (12) it is straightforward to obtain the recurrent equation

$$\mathbf{x}'_{(n)}\Sigma^{-1}_{(n)}\mathbf{y}_{(n)} = \mathbf{x}'_{(n-1)}\Sigma^{-1}_{(n-1)}\mathbf{y}_{(n-1)} + \frac{\left(\frac{x_n}{v_n} - \frac{x_{n-1}}{v_{n-1}}\right)\left(\frac{y_n}{v_n} - \frac{y_{n-1}}{v_{n-1}}\right)}{\left(\frac{u_n}{v_n} - \frac{u_{n-1}}{v_{n-1}}\right)},$$

which can be repeatedly applied to obtain

$$\mathbf{x}'_{(n)}\Sigma^{-1}_{(n)}\mathbf{y}_{(n)} = \frac{x_1 y_1}{u_1 v_1} + \sum_{i=2}^{n} \frac{\left(\frac{x_i}{v_i} - \frac{x_{i-1}}{v_{i-1}}\right)\left(\frac{y_i}{v_i} - \frac{y_{i-1}}{v_{i-1}}\right)}{\left(\frac{u_i}{v_i} - \frac{u_{i-1}}{v_{i-1}}\right)}. \tag{13}$$

The formulas (6) and (13) entail that the information of a design $\tau = (t_1, ..., t_n) \in \mathcal{T}^n$ is

$$M(\tau) = \frac{f^2(t_1)}{u(t_1)v(t_1)} + \sum_{i=2}^{n} \frac{\left(\frac{f(t_i)}{v(t_i)} - \frac{f(t_{i-1})}{v(t_{i-1})}\right)\left(\frac{f(t_i)}{v(t_i)} - \frac{f(t_{i-1})}{v(t_{i-1})}\right)}{\left(\frac{u(t_i)}{v(t_i)} - \frac{u(t_{i-1})}{v(t_{i-1})}\right)}. \tag{14}$$

Therefore, under the product structure of the covariance function the position of the design point t_i contributes to the information only via its relation with the neighboring design points t_{i-1} (if $i \geq 2$) and t_{i+1} (if $i \leq n-1$).

In the rest of this section we will use the product structure of the covariance matrix to derive simple formulas for the MSE of the BLUP. First, notice that $\Sigma_{(n)}\mathbf{e}_n = v_n\mathbf{u}_{(n)}$, where $\mathbf{e}_n = (0, ..., 0, 1)' \in \mathbb{R}^n$. Therefore $v_n\Sigma^{-1}_{(n)}\mathbf{u}_{(n)} = \mathbf{e}_n$, implying that the terms appearing in (7) can be expressed in the form

$$\begin{aligned} \mathbf{r}'(\tau, d)\Sigma^{-1}_{(n)}\mathbf{r}(\tau, d) &= v^2(T_d)\mathbf{u}'_{(n)}\Sigma^{-1}_{(n)}\mathbf{u}_{(n)} = v^2(T_d)\frac{u(t_n)}{v(t_n)}, \\ c(\tau, d) &= v(T_d)\left(\frac{f(T_d)}{v(T_d)} - \frac{f(t_n)}{v(t_n)}\right). \end{aligned}$$

Therefore, for general product covariance structure, a design $\tau \in \mathcal{T}_n$ and $d > 0$, the BLUP and its MSE are given by the formulas

$$\begin{aligned} X^*(\tau, d) &= f(T_d)\beta^*(\tau) + \frac{v(T_d)}{v(t_n)}\left[X(t_n) - f(t_n)\beta^*(\tau)\right], \\ MSE[X^*(\tau, d)] &= v^2(T_d)\left(\frac{u(T_d)}{v(T_d)} - \frac{u(t_n)}{v(t_n)} + \frac{\left(\frac{f(T_d)}{v(T_d)} - \frac{f(t_n)}{v(t_n)}\right)^2}{M(\tau)}\right). \end{aligned} \tag{15}$$

Note that if the last point t_n of τ is fixed, then the $MSE[X^*(\tau, d)]$ is minimized if τ maximizes $M(\tau)$.

3. Optimal sampling designs for the OU process

Assume the regression model of the form (4) and a fixed $\tau = (t_1, ..., t_n)'$. Using the notation of the general model (5) we have $\mathbf{f}(\tau) = a\mathbf{1}_n + be^{-\lambda\tau}$, and the covariance matrix of the errors $\epsilon(\tau)$ is given by (9) and (10).

For $i = 1, ..., n-1$ let $\delta_i = t_{i+1} - t_i$ be the intersampling distances. The equation (14) give the following form of the information:

$$M(\tau) = \frac{2\lambda}{\sigma^2}\left(\frac{\left(a + be^{-\lambda t_1}\right)^2}{1 - e^{-2\lambda t_1}} + a^2 \sum_{i=1}^{n-1} \tanh\left(\frac{\lambda\delta_i}{2}\right)\right).$$

The hyperbolic tangent $\tanh(z) = (e^{2z} - 1)/(e^{2z} + 1)$ is concave on $[0, \infty)$, which implies $\sum_{i=1}^{n-1} \tanh\left(\frac{\lambda\delta_i}{2}\right) \leq \sum_{i=1}^{n-1} \tanh\left(\frac{\lambda\bar{\delta}}{2}\right)$, where $\bar{\delta} = \sum_{i=1}^{n-1} \delta_i/(n-1)$ is the average intersampling distance. Hence, given a fixed initial time $t_1 \in [T_*, T^*)$, the information is maximized for the equidistant design $\tau_n(t_1)$ and the corresponding value of the information is

$$M(\tau_n(t_1)) = \frac{2\lambda}{\sigma^2}\left(\frac{\left(a + be^{-\lambda t_1}\right)^2}{1 - e^{-2\lambda t_1}} + a^2(n-1)\tanh\left(\frac{\lambda(T^* - t_1)}{2n-2}\right)\right). \tag{16}$$

Therefore to find the optimal design $\tau^*_{n,\beta} = \tau_n(t_*)$ for estimating the parameter β, we only need to calculate the optimal initial time $t_1 = t_*$, which is a relatively simple one-dimensional optimization problem

$$t_* = \text{argmax}_{T_* \leq t_1 < T^*} M(\tau_n(t_1)). \tag{17}$$

Note that the interval in (17) for possible values of t_1 is not closed, which means that in principle the optimal design needs not exist; it can "degenerate" to T^*. Nevertheless, it is simple to check that $dM(\tau_n(t_1))/dt_1|_{t_1=T^*} < 0$, which implies that the function $M(\tau_n(t_1))$ is strictly decreasing in the point $t_1 = T^*$. Thus the maximum in (17) is attained at some $t_* < T^*$, i.e., the optimal design does exist.

To find the optimal prediction design $\tau^*_{n,d}$, note that $f(\tau) = a\mathbf{1}_n + b\lambda\sigma^{-2}\mathbf{v}(\tau)$, which applied to the general formula (15) gives

$$MSE[X^*(\tau, d)] = \frac{\sigma^2}{2\lambda}\left(1 - e^{-2\lambda(T_d - t_n)}\right) + a^2\left(1 - e^{-\lambda(T_d - t_n)}\right)^2 M^{-1}(\tau). \tag{18}$$

Let $\tau_n(t_*)$ be the design that maximizes $M(\tau)$. Since the last time point of $\tau_n(t_*)$ is T^*, it is clear that $\tau_n(t_*)$ minimizes $MSE[X^*(\tau, d)]$. The equation (18) thus implies that for all choices of the constants a, b and d the design $\tau_n(t_*)$, which is optimal for parametric estimation, is also optimal for prediction. That is, we have $\tau^*_{n,d} = \tau^*_{n,\beta} = \tau_n(t_*)$.

If $ab \geq 0$, then the information $M(\tau_n(t_1))$ given by (16) is obviously decreasing as a function of t_1. Consequently, the equidistant design with the initial sampling time T_* is optimal, i.e., $\tau^*_{n,d} = \tau^*_{n,\beta} = \tau_n(T_*)$.

Summarizing the obtained results we can formulate the following theorem.

Theorem 1. *Assume the model* (4) *such that the constants* a *and* b *are not simultaneously equal to* 0 *and the covariance matrix of* $\epsilon(\tau)$ *is given by* (9) *and* (10). *Then there exists* $t_* \in [T_*, T^*)$ *such that the equidistant design* $\tau_n(t_*)$ *is optimal for both parametric estimation and prediction. Moreover, if* $ab \geq 0$, *then* $t_* = T_*$.

As a direct consequence of Theorem 1 we see that the equidistant design $\tau_n(T_*)$ is optimal for the special case $a = 1, b = 0$, i.e., for the OU process with a constant mean value. As a limit stationary case where the covariance function is given by (3) we obtain a theorem that has been recently proved in the papers [4] and [7].

A next corollary of Theorem 1 is the optimality of $\tau_n(T_*)$ for the case $a = 0, b = 1$, that is for the case that the asymptotic average velocity $\bar{v}$ is known and the aim is estimating the initial velocity v_0.

The situation $a = 1, b = -1$, i.e., if we know v_0 and the aim is estimating the asymptotic average velocity $\bar{v}$, does not conform to the assumptions of Theorem 1. Nonetheless, in this model we can easily verify that $M(\tau_n(t_1))$, as a function of t_1, is increasing on $[0, \frac{T^*}{n})$ and decreasing on $(\frac{T^*}{n}, T^*]$. Consequently, the optimal design for the model (4) with $a = 1$ and $b = -1$ is $\tau_n(t_*)$, where $t_* = T_*$ if $T_* > \frac{T^*}{n}$, and $t_* = \frac{T^*}{n}$ if $T_* \leq \frac{T^*}{n}$. We see that that in general the initial optimal sampling design point t_* needs not coincide with the beginning T_* of the experimental domain.

References

[1] Dette H., Kunert J., Pepelyshev A. (2008) Exact optimal designs for weighted least squares analysis with correlated errors. Statistica Sinica, 18:135–154.

[2] Karatzas I., Schreve S.E. (2008) Brownian Motion and Stochastic Calculus. Springer Verlag.

[3] Karlin S., Taylor H.M. (1998) An Introduction to Stochastic Modeling, Third Edition. Academic Press.

[4] Kiseľák J., Stehlík M. (2008) Equidistant and D-optimal designs for parameters of Ornstein-Uhlenbeck process. Statistics and Probability Letters, 78:1388–1396.

[5] Lemons D.S. (2002) An Introduction to Stochastic Processes in Physics. The Johns Hopkins University Press.

[6] Mukherjee B. (2003) Exactly Optimal Sampling Designs for Processes with a Product Covariance Structure. The Canadian Journal of Statistics / La Revue Canadienne de Statistique, 31:69–87.

[7] Zagoraiou M., Antognini A.B. (2009) Optimal designs for parameter estimation of the Ornstein-Uhlenbeck process. Appl. Stochastic Models Bus. Ind. (to appear)

[8] Štulajter F. (2002) Predictions in Time Series Using Regression Models. Springer Verlag.

6th St.Petersburg Workshop on Simulation (2009) 1103-1107

A-Optimal Designs for Two-Color Microarray Experiments for interesting contrasts[1]

Katharina Schiffl[2], Ralf-Dieter Hilgers[3]

Abstract

Two-color microarray experiments form an important tool in gene expression analysis. Due to the high costs of microarray experiments it is fundamental to plan these experiments carefully. Therefore, in this paper we propose a method to construct A-optimal microarray designs which ensure unbiased and precise estimation of treatment control comparisons. This method can also be applied to derive optimal designs for other contrast settings. For practical applications recommendations for the choice of efficient experimental layouts can be derived from the constructed designs. We will show that our designs are more efficient than the designs currently used in praxis.

1. Introduction

In recent years microarray technology has become one of the most prominent tools in gene expression analysis due to the fact that thousands of genes can be screened simultaneously. Two-color microarray experiments are able to compare two samples on one single array, one sample is colored in green and the other one in red. After the processing of the experiment we get dye intensities for each sample, which are connected to the corresponding activities of the observed genes. However, microarrays are very expensive, so it is fundamental to use appropriate designs to get precise results with minimal resources. Optimal experimental designs estimate the parameters of the underlying statistical model unbiased and with minimal variance and give instructions which treatments should be combined on the same microarray. Design issues for microarray experiments have been investigated intensively in the last years, see for example [3] [8]. Kerr et al. first analysed two-color microarray data by analysis of variance (ANOVA) and recommended a model describing the logarithm of the measured intensities dependent on the array-, variety-, dye- and gene-effect including appropriate interactions [4]. Their work has been extended by many authors. For instance, Wolfinger et al. modeled the array effect

[1]The research was financially supported by the BMBF within the joint research project SKAVOE (Foerderkennzeichen: 03HIPAB4).

[2]RWTH Aachen University, E-mail: kschiffl@ukaachen.de

[3]RWTH Aachen University, E-mail: rdhilgers@ukaachen.de

as random [7]. Landgrebe et al. analysed the ratios of the logarithmic dye intensities separately for each gene [5]. Although in medical applications scientists are very often interested in comparing several treatments to a control treatment, only few authors considered design problems for treatment-control comparisons in microarray experiments [6]. In this work we will propose a method to derive exact and approximate A-optimal designs for microarray experiments. We will apply this method to construct A-optimal designs for estimating treatment-control comparisons and we will show that these designs are more efficient than the designs currently used in praxis.

2. Preliminaries

The statistical analysis is based on the gene specific model

$$y_{ijk} = \tau_i + \alpha_j + \delta_k + \epsilon_{ijk} \tag{1}$$

where y_{ijk} describes the logarithm of the intensity of treatment i, $i \in \{0, \ldots, t\}$ on array j, $j \in \{1, \ldots, a\}$ colored in dye k, $k \in \{\text{green}, \text{red}\}$ [1]. The error terms ϵ_{ijk} are assumed to be iid with mean zero and variance σ^2, w.l.o.g. $\sigma^2 = 1$. In our special case τ_0 indicates the control-treatment and τ_i, $i \in \{1, \ldots, t\}$ indicate the other treatments. We will ignore the dye effects in a first step, an approach to allocate the dyes afterwards in a suitable way is given in [1]. Assuming model (1) we will derive A-optimal designs for treatment-control comparisons, which minimize the sum of the variances of the considered estimators. Specifically we have to minimize $\sum_{i=1}^{t} \mathrm{Var}(\hat{\tau}_0 - \hat{\tau}_i)$, if we are interested in estimating all treatment-control contrasts $\tau_i - \tau_0$ for $i = 1, \ldots, t$. This sum is in particular dependent on the number of used arrays a and on the two treatments combined on each array as we can see in section . Every microarray design can be illustrated by a graph with $t+1$ vertices and a edges, so that every two treatments connected by an edge in the graph are tested on the same array in the experiment.

3. Results

3.1. Exact designs

In this section we will give a formula to derive exact optimal designs for practical situations with a given quantity of arrays and small numbers of treatments, since in many applications the number of treatments will not exceed a known limit. As mentioned above, the optimal designs have to minimize the term $\sum_{i=1}^{t} \mathrm{Var}(\hat{\tau}_0 - \hat{\tau}_i)$. We can calculate the exact values $\mathrm{Var}(\hat{\tau}_0 - \hat{\tau}_i)$ with results from combinatorics and physical networks. Theorem 1 gives the corresponding value of $\mathrm{Var}(\hat{\tau}_i - \hat{\tau}_j)$.

Theorem 1. *Let $V = \{0, \ldots, t\}$ be the set of vertices (treatments) and $E = \{x_{01}, x_{02}, \ldots, x_{(t-1)t}\}$ the set of edges of a given graph with $|E| = \binom{t+1}{2}$. The function $b : x_{ij} \longmapsto b(x_{ij}) : E \longrightarrow \mathbb{N}_0$.*

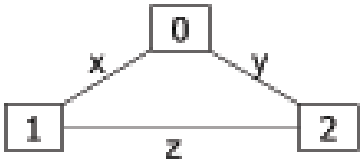


Figure 1: Graph representation of a microarray experiment with $t = 2$

$b(x_{ij})$ specifies for each treatment pair (i, j) the number of arrays comparing treatments i and j. Then

$$Var(\hat{\tau}_i - \hat{\tau}_j) = \frac{\sum\limits_{\substack{A \subset E \setminus \{x_{ij}\}: |A| = t-1: \\ (V, A \cup \{x_{ij}\}) \text{ has no circles}}} b(a_1)b(a_2)\cdots b(a_{L-2})}{\sum\limits_{\substack{A \subset E: |A| = t: \\ (V,A) \text{ has no circles}}} b(a_1)b(a_2)\cdots b(a_{L-1})} \tag{2}$$

with $i, j \in \{0, \ldots, t\}$, $i \neq j$ and $A = \{a_1, a_2 \ldots\} \subset E$.

Theorem 1 can be proven with results from physical networks, especially with the help of resistance matrices. For a detailed description see Bailey and Cameron [2]. We apply formula (2) for small values of t and calculate exact optimal designs for given numbers of arrays a. We will demonstrate the application of theorem 1 by means of the following example.

Example 1: If we consider two treatments and one control treatment the experiment with $a = x+y+z$ arrays is illustrated in Fig. 1. Let $b(x_{01}) = x$, $b(x_{02}) = y$, $b(x_{12}) = z$, i.e. treatment 1 is combined with the control-treatment on x arrays etc. Consequently we get $\mathrm{Var}(\hat{\tau}_0 - \hat{\tau}_1) = \frac{y+z}{xy+xz+yz}$ and $\mathrm{Var}(\hat{\tau}_0 - \hat{\tau}_2) = \frac{x+z}{xy+xz+yz}$. To achieve A-optimality we have to minimize

$$\mathrm{Var}(\hat{\tau}_0 - \hat{\tau}_1) + \mathrm{Var}(\hat{\tau}_0 - \hat{\tau}_2) = \frac{x + y + 2z}{xy + xz + yz}$$

under the constraints $x + y + z = a$ and $x, y, z \in \{0, \ldots, a\}$. The results of this minimization obtained for $a \in \{6, 8, 10, 12, 15\}$ arrays are displayed in Table 1 (using Mathematica, Wolfram Research). In Table 2 and Table 3 we listed similar results for $t \in \{3, 4\}$, because these values are often used in practical situations. For higher values of t the computation of $\mathrm{Var}(\hat{\tau}_i - \hat{\tau}_j)$ increases extremely. Therefore, we will propose approximate optimal designs for these situations in section . The corresponding approximate optimality results can be used to construct nearly optimal designs for all values of t and a.

3.2. Approximate optimal designs

The derivation of the approximate optimal designs is also based on the results of section , i.e. we will use the same approach to minimize $\sum_{i=1}^{t} \mathrm{Var}(\hat{\tau}_0 - \hat{\tau}_i)$, but we

will consider continuous values for $B_{ij} \in [0,1]$. Therefore we will minimize

$$\sum_{i=1}^{t} \mathrm{Var}(\hat{\tau}_0 - \hat{\tau}_i)$$

under the constraints

$$B_{01} + B_{02} + \cdots + B_{(t-1)t} = 1$$

$$B_{ij} \in [0,1]$$

for all $0 \leq i, j \leq t$, $i \neq j$. As a result we get for $t \neq 3$

$$B_{0i} = \frac{2((t-1)\sqrt{t+1} - (t+1))}{t(t+1)(t-3)}$$

for $1 \leq i \leq t$ and

$$B_{ij} = \frac{2(t - 2\sqrt{t+1} + 1)}{t(t+1)(t-3)}$$

for $1 \leq i, j \leq t$ and $i \neq j$. For $t = 3$ we get $B_{0i} = \frac{1}{4}$ and $B_{ij} = \frac{1}{12}$. The according weights B_{0i} and B_{ij} computed for $t \in \{3,4,5,6,7\}$ are summarized in Table 4.

3.3. Efficiency Considerations

In practical situations researchers often use the star designs. Star designs allocate the control treatment with another treatment on every array. Using the designs proposed in this paper instead of the star designs we observe a gain in efficiency of at least 10% for $(t,a) \in \{(3,9),(4,16)\}$ (see Table 5), we get similar results for other values of t and a.

4. Conclusion

Kunert et al. also proposed approximate optimal designs for treatment-control comparisons in [6]. In our paper we extend their work by the construction of exact optimal designs. Exact designs are very important for practical situations, because they provide optimal solutions for a given number of arrays. Approximate designs may be used for recommendations, when the computing of exact designs is difficult. They do not consider settings with infinite arrays like the approximate theory. The recommended exact designs exist for all numbers of arrays. Moreover the presented approach can be used to develop optimal designs for different contrasts. Consequently the target function in the minimization will change to $\min \sum_{i=1}^{m} \mathrm{Var}(c_i\hat{\tau})$ for the contrasts c_i, $i = 1,\ldots,m$ and the parameter vector $\hat{\tau} = (\hat{\tau}_1,\ldots,\hat{\tau}_t)^T$.

Table 1: Exact designs for $t = 2$ treatments and a arrays

	$a = 6$	$a = 8$	$a = 10$	$a = 12$	$a = 15$
B_{01}	3	4	4	5	6
B_{02}	2	3	4	5	6
B_{12}	1	1	2	2	3

Table 2: Exact designs for $t = 3$ treatments and a arrays

	$a = 9$	$a = 11$	$a = 12$	$a = 15$	$a = 20$	$a = 25$—
B_{01}	2	3	3	4	5	7
B_{02}	2	2	3	4	5	6
B_{03}	2	3	3	4	5	6
B_{12}	1	1	1	1	1	2
B_{13}	1	1	1	1	2	2
B_{23}	1	1	1	1	2	2

Table 3: Exact designs for $t = 4$ treatments and a arrays

	$a = 14$	$a = 15$	$a = 16$	$a = 20$	$a = 25$
B_{01}e	2	3	2	4	5
B_{02}	2	2	2	3	4
B_{03}	2	3	3	4	5
B_{04}	2	2	3	3	4
B_{12}	1	1	1	1	1
B_{13}	1	0	1	1	1
B_{14}	1	1	1	1	1
B_{23}	1	1	1	1	1
B_{24}	1	1	1	1	2
B_{34}	1	1	1	1	1

Table 4: Weights of the approximate designs

	B_{0i}	B_{ij}
$t = 2$	$\frac{1}{3}(3 - \sqrt{3})$	$\frac{1}{3}(-3 + 2\sqrt{3})$
$t = 3$	$\frac{1}{4}$	$\frac{1}{12}$
$t = 4$	$\frac{1}{10}(-5 + 3\sqrt{5})$	$\frac{1}{10}(5 - 2\sqrt{5})$
$t = 5$	$\frac{1}{15}(-3 + 2\sqrt{6})$	$\frac{1}{15}(3 - \sqrt{6})$
$t = 6$	$\frac{1}{63}(-7 + 5\sqrt{7})$	$\frac{1}{63}(7 - 2\sqrt{7})$

Table 5: Comparison with the star design

	Star $t=3$ $a=9$	New $t=3$ $a=9$	Star $t=4$ $a=16$	New $t=4$ $a=16$
B_{01}	3	2	4	2
B_{02}	3	2	4	2
B_{03}	3	2	4	3
B_{04}			4	3
B_{12}	0	1	0	1
B_{13}	0	1	0	1
B_{14}			0	1
B_{23}	0	1	0	1
B_{24}			0	1
B_{34}			0	1
Average	1	0.9	1	0.87
Variance		90%		87%

References

[1] R. A. Bailey *Designs for two-colour microarray experiments.* Journal of the Royal Statistical Society: Series C (Applied Statistics) 56, 2007

[2] R. A. Bailey and P. J. Cameron *Combinatorics of optimal designs.* www.maths.qmw.ac.uk/ pjc/preprints/optimal.pdf, preprint, 2008.

[3] M. K. Kerr and G. A. Churchill *Experimental design for gene expression microarrays.* Biostatistics, 2, 2001.

[4] M. K. Kerr, M. Martin and G. A. Churchill *Analysis of variance for gene expression microarray data.* J. Computnl Biol., 7, 2000.

[5] J. Landgrebe, F. Bretz and E. Brunner *Efficient design and analysis of two colour factorial microarray experiments.* Computational Statistics and Data Analysis, 2006.

[6] J. Kunert, R.J. Martin and S. Rothe *Optimal designs for treatment-control comparisons in microarray experiments* in B. Schipp and W. Kraemer, *Statistical Inference, Econometric Analysis and Matrix Algebra*, 1rd ed. Dortmund, Springer-Verlag, 2009.

[7] R. D. Wolfinger, G. Gibson, E. D. Wolfinger, L. Bennett, H. Hamadeh, P. Bushel, C. Afshari and R. S.Paules *Assessing gene significance from cDNA microarray expression data via mixed models.* J. Computnl Biol., 8, 2001.

[8] H. Y. Yang and T. Speed *Design issues for cDNA microarray experiments.* Nat. Rev. Genet., 3, 2002.

Section

Optimization techniques

6th St.Petersburg Workshop on Simulation (2009) 1111-1115

On some properties of the simulated annealing algorithm

Alexey Tikhomirov[1]

Abstract

The paper is devoted to the theoretical comparison of the simulated annealing algorithm with the so-called Markov monotonous search. It is shown that the Markov monotonous search can be considered as a limiting case of the simulated annealing algorithm. This result is used to construct fast variants of the simulated annealing algorithm. It is shown that the asymptotic rate of convergence of the simulated annealing algorithm may be just marginally worse than the rate of convergence of a standard descent algorithm (e.g., steepest descent).

1.Introduction

Let X be a feasible region and ρ be a metric on X. We shall call the pair (X, ρ) the *optimization space*. We shall use the notation $S_r(x) = \{y \in X$, such that $\rho(x, y) \leq r\}$ for the closed ball.

Consider a optimization space (X, ρ) and suppose that f is a certain objective function defined on X. Let f have a unique minimizer $x_0 = \arg\min_{x \in X} f(x)$ and assume that our aim is to find x_0 with a given accuracy $\varepsilon > 0$. To estimate x_0, we use random search sequences of a special kind.

Let $\{\xi_i\}_{i \geq 0}$ be any sequence (either finite or infinite) of random points in X. If this sequence forms a Markov chain (that is, if for all i the distribution of ξ_{i+1} conditional on $\xi_0, \ldots, \xi_i$ coincides with the distribution of ξ_{i+1} conditional on ξ_i only), then we say that $\{\xi_i\}_{i \geq 0}$ is a Markov (random) search. If, in addition, for all $i \geq 0$ we have $f(\xi_{i+1}) \leq f(\xi_i)$ with probability 1, then we shall say that $\{\xi_i\}_{i \geq 0}$ is a *Markov monotonous search*.

It is useful to present a general algorithmic scheme for simulation of the Markov random sequence $\{\xi_i\}_{i \geq 0}$ with $\xi_0 = x \in X$.

Algorithm 1 (A general scheme of Markov algorithms).

1. Set $\xi_0 = x$ and the iteration number $i = 1$.

2. Obtain a point ζ_i by sampling from the distribution $P_i(\xi_{i-1}, \cdot\,)$.

3. Set $\xi_i = \begin{cases} \zeta_i & \text{with probability } p_i, \\ \xi_{i-1} & \text{with probability } 1 - p_i. \end{cases}$

Here $p_i = p_i\big(\zeta_i, \xi_{i-1}, f(\zeta_i), f(\xi_{i-1})\big)$ is the acceptance probability.

[1]St. Petersburg University, Novgorod University, E-mail: Tikhomirov.AS@mail.ru

4. Set $i = i + 1$ and go to Step 2.

Here $P_i(x, \cdot\,)$ are Markov transition probabilities; that is, $P_i(x, \cdot\,)$ is a probability measure for all $i \geq 1$ and $x \in X$, and $P_i(\,\cdot\,, A)$ is $\mathcal{B}_X$-measurable for all $i \geq 1$ and $A \in \mathcal{B}_X$ (where $\mathcal{B}_X$ is the Borel σ-algebra of subsets of X). Due to the structure of Algorithm 1, distributions $P_i(x, \cdot\,)$ can be called *trial transition functions*.

Particular choices of the transition probability $P_i(x, \cdot\,)$ and acceptance probabilities p_i lead to specific Markov global random search algorithms. The most well-known among them is the celebrated 'simulated annealing' which is considered in this paper.

A general simulated annealing algorithm is Algorithm 1 with acceptance probabilities

$$p_i = \begin{cases} 1 & \text{if } \Delta_i \leq 0, \\ \exp(-\beta_i \Delta_i) & \text{if } \Delta_i > 0, \end{cases} \tag{1}$$

where $\Delta_i = f(\zeta_i) - f(\xi_{i-1})$ and $\beta_i \geq 0$ $(i = 1, 2, \ldots)$.

The choice (1) for the acceptance probability p_i means that any 'promising' new point ζ_i (for which $f(\zeta_i) \leq f(\xi_{i-1})$) is accepted unconditionally; a 'non-promising' point (for which $f(\zeta_i) > f(\xi_{i-1})$) is accepted with probability $p_i = \exp(-\beta_i \Delta_i)$. As the probability of acceptance of a point which is worse than the preceding one is always greater than zero, the search trajectory may leave a neighbourhood of a local and even a global minimizer. Note however that the probability of acceptance decreases if the difference $\Delta_i = f(\zeta_i) - f(\xi_{i-1})$ increases. This probability also decreases if β_i increases. In the limiting case, where $\beta_i = +\infty$ for all i, the simulated annealing algorithm becomes the Markov monotonous search.

Below, we shall consider the case where the parameteres β_i do not depend on the iteration number i (that is, $\beta_i = \beta$).

Markov monotonous search can be considered as Algorithm 1 where the acceptance probabilities p_i are

$$p_i = \begin{cases} 1 & \text{if } \Delta_i \leq 0, \\ 0 & \text{if } \Delta_i > 0. \end{cases}$$

For simplicity of references, let us now formulate the algorithm for generating a Markov monotonous search.

Algorithm 2 (Markov monotonous search).

1. Set $\xi_0 = x$ and the iteration number $i = 1$.

2. Obtain a point ζ_i by sampling from the distribution $P_i(\xi_{i-1}, \cdot\,)$.

3. If $f(\zeta_i) \leq f(\xi_{i-1})$ then set $\xi_i = \zeta_i$. Alternatively, set $\xi_i = \xi_{i-1}$.

4. Set $i = i + 1$ and go to Step 2.

Below, we shall write P_x and E_x for the probabilities and expectations related to the search of Algorithm 1 starting at a point $x \in X$.

We use the random search for finding a minimizer x_0 with a given accuracy ε (approximation with respect to the argument). It can happen, however, that after reaching the set $S_\varepsilon(x_0)$ at iteration i, a search algorithm leaves it at a subsequent

iteration. In order to avoid complications related to this phenomenon, we introduce the sets

$$M_r = \{x \in S_r, \text{ such that } f(x) < f(y) \text{ for any } y \notin S_r\}.$$

It is easy to see that the sets M_r have the following properties:
a) if $r_1 < r_2$, then $M_{r_1} \subset M_{r_2}$, and
b) if $x \in M_r$ and $y \notin M_r$, then $f(x) < f(y)$.

Thus, any monotonous search does not leave the set M_ε after reaching it.

In a general case we shall use $\xi_i^* = \arg\min\{f(\xi_0), \ldots, f(\xi_i)\}$. We shall assume that $\arg\min\{f(\xi_0), \ldots, f(\xi_i)\} = \xi_j$, where $j = \max\big\{k \in \{0, \ldots, i\}$, such that $f(\xi_k) = \min\{f(\xi_0), \ldots, f(\xi_i)\}\big\}$.

If the simulated annealing hits the set M_ε, ξ_i^* never leaves it.

To study the approximation with respect to the argument, we shall study the moment the algorithm reaches the set M_ε for the first time; as above, ε is the required precision with respect to the argument. Thus we come to a random variable $\tau_\varepsilon = \min\{i \geq 0, \text{ such that } \xi_i \in M_\varepsilon\}$.

Since we always assume that in order to generate the transition probabilities P_i we do not need to evaluate the objective function $f(\cdot)$, we only need one function evaluation at each iteration $\xi_{i-1} \mapsto \xi_i$ of Algorithm 1. Hence the distribution of the random variable τ_ε provides us with very useful information about the quality of a particular random search algorithm. Indeed, in τ_ε iterations of Algorithm 1 the objective function $f(\cdot)$ is evaluated $\tau_\varepsilon + 1$ times.

The quantity $\mathsf{E}_x\tau_\varepsilon$ can be interpreted as the average number of iterations of a search algorithm required to reach the set M_ε.

The other characteristic of τ_ε considered in this paper is defined as the minimal number of steps $N = N(x, \varepsilon, \gamma)$ at which the hit of M_ε is ensured with the probability greater than γ; in other words

$$\begin{aligned} N(x, \varepsilon, \gamma) &= \min\big\{i \geq 0, \text{ such that } \mathsf{P}_x(\xi_i^* \in M_\varepsilon) > \gamma\big\} = \\ &= \min\big\{i \geq 0, \text{ such that } \mathsf{P}_x(\tau_\varepsilon \leq i) > \gamma\big\}. \end{aligned}$$

We shall study $\mathsf{E}_x\tau_\varepsilon$ and $N(x, \varepsilon, \gamma)$ as functions of the required precision ε, as $\varepsilon \to 0$. Note that for many local optimization algorithms (such as steepest descent) the number of iterations has the order $O(|\ln \varepsilon|)$, $\varepsilon \to 0$. In global optimization problems the order for the number of iterations is typically worse; it is $O(1/\varepsilon^\alpha)$ for some $\alpha > 0$. If $X \subset \mathbb{R}^d$, then (see, e.g., [2]) the so-called *pure random search* methods need in average $O(1/\varepsilon^d)$ calculations of an objective function to hit $S_\varepsilon(x_0)$ for any reasonable ρ and any f. The conclusion concerning the simulated annealing algorithm made in [2] is as follows: "The theoretical rate of convergence of the simulated annealing is very slow; this convergence is based on the convergence of the pure random search which is contained within the simulated annealing algorithms ...". Thus, there is a large gap between 'fast' local methods and 'slow' stochastic global optimization methods.

Below, we shall indicate versions of the Markov search such that $\mathsf{E}_x\tau_\varepsilon$ (as well as $N(x, \varepsilon, \gamma)$) has the order $O(\ln^2 \varepsilon)$. This is achieved by means of a clever choice of the trial transition probabilities P_i. Certainly, some restrictions on an objective function $f(\cdot)$ are required.

2. Relation to a Markov monotonous search

The next theorem shows that a Markov monotonous search can be considered as the limiting case of a simulated annealing algorithm.

Theorem 1. *Suppose that the objective function $f(\cdot)$ attains the global minimum at a single point x_0. Consider a simulated annealing algorithm $\{\xi_i\}_{i\geq 0}$ with initial point $\xi_0 = x$, trial transition functions P_i, and acceptance probabilities p_i defined by (1) where the parameteres β_i do not depend on the iteration number i (that is, $\beta_i = \beta$). Consider a Markov monotonous search $\{\eta_i\}_{i\geq 0}$ with initial point $\eta_0 = x$ and trial transition functions P_i. The following assertions are true.*
1. *Let $n \in \mathbb{N}$ and $A_1, \ldots, A_n \in \mathcal{B}_X$ (where $\mathcal{B}_X$ is the Borel σ-algebra of subsets of X). Then*

$$\lim_{\beta\to+\infty} \mathsf{P}_x(\xi_1 \in A_1, \ldots, \xi_n \in A_n) = \mathsf{P}_x(\eta_1 \in A_1, \ldots, \eta_n \in A_n).$$

2. *Suppose that $\varepsilon > 0$, $\gamma \in (0,1)$ and $x \in X$. Set*

$$N(x,\varepsilon,\gamma) = \min\bigl\{i \geq 0, \text{ such that } \mathsf{P}_x(\xi_i^* \in M_\varepsilon) > \gamma\bigr\},$$
$$N_*(x,\varepsilon,\gamma) = \min\bigl\{i \geq 0, \text{ such that } \mathsf{P}_x(\eta_i \in M_\varepsilon) > \gamma\bigr\}.$$

Then there exists a $\beta_0 = \beta_0(x,\varepsilon,\gamma) > 0$ such that the inequality

$$N_*(x,\varepsilon,\gamma) \geq N(x,\varepsilon,\gamma)$$

holds for all $\beta \geq \beta_0$.
3. *Suppose that $\varepsilon > 0$ and, for all $i \in \mathbb{N}$ and $x \in X$, probabilities $P_i(x, M_\varepsilon)$ satisfy the inequalities $P_i(x, M_\varepsilon) \geq c_\varepsilon > 0$. Set $\tau_\varepsilon = \min\bigl\{i \geq 0 : \xi_i \in M_\varepsilon\bigr\}$, $\tau_\varepsilon^* = \min\bigl\{i \geq 0 : \eta_i \in M_\varepsilon\bigr\}$. Then it holds that*

$$\lim_{\beta\to+\infty} \mathsf{E}_x\tau_\varepsilon = \mathsf{E}_x\tau_\varepsilon^*.$$

3. Examples of fast random search methods

Let us take $X = \mathbb{I}^d = (0,1]^d$ and equip X with the d-dimensional Lebesgue measure μ and the metric

$$\rho(x,y) = \rho_\infty(x,y) = \max_{1\leq i\leq d} \min\bigl\{|x_i - y_i|, 1 - |x_i - y_i|\bigr\},$$

where $x = (x_1, \ldots, x_d)$ and $y = (y_1, \ldots, y_d)$. Metric space $(\mathbb{I}^d, \rho_\infty)$ can be treated as the d-dimensional torus. Note that the volume of the ball $S_r(x)$ does not depend on x. Set $\varphi(r) = \mu(S_r(x))$. Evidently, $\operatorname{diam}\mathbb{I}^d = 0.5$, and $\varphi(r) = (2r)^d$ for $r \leq 0.5$.

We shall assume that the objective function $f\colon \mathbb{I}^d \mapsto \mathbb{R}$ is measurable, bounded from below and satisfies the following conditions.

CF1. Function f attains its minimal value at a single point $x_0 \in X$.

CF2. Function f is continuous at the point x_0.

CF3. Inequality $\inf\{f(x),\ \text{such that } x \notin S_r(x_0)\} > f(x_0)$ holds for any $r > 0$.

Condition CF3 provides that the convergence $\rho(x_n, x_0) \to 0$ follows from the convergence $f(x_n) \to f(x_0)$. Note that the functions of the class thus specified can be multimodal in any neighborhood of the global minimum.

The main information used about the objective function $f(\cdot)$ will be contained in the so-called *asymmetry coefficient*

$$F^f(r) = \mu\big(M_r\big)/\mu\big(S_r(x_0)\big) = \mu(M_r)/\varphi(r).$$

This coefficient 'compares' the behaviour of $f(\cdot)$ with the F-ideal uniextremal function which has an asymmetry coefficient $F^f \equiv 1$. In particular, the asymmetry coefficient codes the information about the local minima of $f(\cdot)$.

The conditions imposed on $f(\cdot)$ guarantee that $F^f(r) > 0$ for all $r > 0$. The functions $f(\cdot)$ such that $\liminf F^f(r) > 0$ as $r \to 0$, will be called *non-degenerate*.

In particular, assume that

$$c_1\rho^t(x_0, x) \le f(x) - f(x_0) \le c_2\rho^t(x_0, x),$$

in a neighborhood of x_0, where t, c_1, $c_2 > 0$. Then, $\liminf F^f(r) \ge (c_1/c_2)^{d/t} > 0$ as $r \to 0$. If $f(x) = f(x_0) + \phi\big(\rho(x_0, x)\big)$, where $\phi(\cdot)$ is a monotonically increasing nonnegative function with $\phi(0) = 0$, then $F^f \equiv 1$.

Other facts concerning the properties of the function $F^f(\cdot)$ can be found in [2, 6, 8].

Below, we assume that the objective function $f(\cdot)$ satisfies the following condition.

CF4. Function f is non-degenerate.

We shall consider the case where the trial transition probabilities $P_i(x, \mathrm{d}y)$ do not depend on the iteration number i (that is, $P_i = P$) and have a symmetric density $p(x, y)$ of the form

$$p(x, y) = g\big(\rho(x, y)\big), \tag{2}$$

where ρ is the metric and g is a non-increasing nonnegative function defined on $(0, 0.5]$. The function g is called the *search form* and also the form of the transition density p. In order for the function $p(x, y)$ defined in (2) to be a density, the search form g must satisfy the normalization condition

$$\int_{(0,0.5]} g(r)\,\mathrm{d}\varphi(r) = 1. \tag{3}$$

Without loss of generality, we will assume that g is continuous from the left.

Markov monotonous search with transition densities of the form (2) will be called *Markov symmetric monotonous search*.

Below, we shall consider Markov searches with transition densities of a special kind. Let us fix a nonincreasing left-continuous function $h\colon (0, 0.5] \mapsto (0, \infty)$. Additionally, let $h(r)r^d \to 1$ as $r \to 0$. For a certain $a > 0$ and $0 < \varepsilon < 1/(2a)$ consider a search form

$$g_\varepsilon(r) = \frac{1}{\Lambda_\varepsilon}\begin{cases} h(a\varepsilon) & \text{if } 0 < r \le a\varepsilon, \\ h(r) & \text{if } a\varepsilon < r \le 0.5, \end{cases} \tag{4}$$

where Λ_ε is a normalizing coefficient providing the equality (3).

As it is proved in [7, 8, 9] (see also [2]), for the Markov symmetric monotonous search determined by search forms (4), relations $\mathsf{E}_x\tau_\varepsilon = O(\ln^2 \varepsilon)$ and $N(x, \varepsilon, \gamma) = O(\ln^2 \varepsilon)$ hold true. All conditions of Theorem 1 are satisfied. Applying this theorem we obtain that for the simulated annealing algorithm determined by search forms (4), relations $\mathsf{E}_x\tau_\varepsilon = O(\ln^2 \varepsilon)$ and $N(x, \varepsilon, \gamma) = O(\ln^2 \varepsilon)$ hold true. Therefore, a asymptotic rate of convergence of the simulated annealing algorithm is just marginally worse than the rate of convergence of a standard descent algorithm (e.g., steepest descent) for an ordinary local optimization problem.

References

[1] Ermakov S.M., Zhigljavsky A.A. (1983) On a random search of a global extremum // Probability theory and its applications, No 1, pp. 129–136.

[2] Zhigljavsky A., Zilinskas A. (2008) Stochastic Global Optimization. Berlin: Springer-Verlag.

[3] Spall J.C. (2003) Introduction to stochastic search and optimization: estimation, simulation, and control, Wiley, New Jersey.

[4] Spall J.C., Hill S.D., Stark D.R. (2006) Theoretical framework for comparing several stochastic optimization approaches // Probabilistic and Randomized Methods for Design Under Uncertainty, London: Springer-Verlag, pp. 99–117.

[5] Yin G. (1999) Rates of convergence for a class of global stochastic optimization algorithms // SIAM Journal on Optimization, Vol. 10. No. 1, pp. 99–120.

[6] Nekrutkin V.V., Tikhomirov A.S. (1993) Speed of convergence as a function of given accuracy for random search methods // Acta Applicandae Mathematicae, Vol. 33, pp. 89–108.

[7] Tikhomirov A., Stojunina T., Nekrutkin V. (2007) Monotonous random search on a torus: Integral upper bounds for the complexity // Journal of Statistical Planning and Inference, Vol. 137. Issue 12, pp. 4031–4047.

[8] Tikhomirov A.S. (2006) On the Markov Homogeneous Optimization Method // Computational Mathematics and Mathematical Physics, Vol. 46, No. 3, pp. 361-375.

[9] Tikhomirov A.S. (2007) On the Convergence Rate of the Markov Homogeneous Monotone Optimization Method // Computational Mathematics and Mathematical Physics, Vol. 47, No. 5, pp. 780-790.

[10] Tikhomirov A.S. (2007) On the complexity of the simulated annealing algorithm, Available from VINITI, 2007, no. 230.

6th St.Petersburg Workshop on Simulation (2009) 1117-1121

Optimization of Real-Time Systems

Joseph Kreimer[1], Edward Ianovsky[1,2]

Abstract

We consider a real-time multiserver system with identical servers (e.g. unmanned aerial vehicles, manufacturing controllers, etc.) that provide service for requests of real-time jobs arriving via several different channels (e.g. surveillance areas, assembly lines, etc.) working under maximum load regime. There are fixed numbers of jobs in each channel at any instant. There are ample identical maintenance teams available to repair all servers simultaneously, if needed. After maintenance servers are distributed between channels according to assignment probabilities. We compute analytically (for exponentially distributed service and maintenance times) and via simulation using Cross-Entropy method optimal assignment probabilities which maximize system availability.

1. Introduction

Real-time systems (RTS) are defined as those for which correctness depends not only on the logical properties of the computed results, but also on the temporal properties of these results. In RTS a calculation that uses temporally invalid data or an action performed too early/late, may be useless, and sometimes harmful – even if such a calculation or action is functionally correct.

We will focus on RTS with a zero deadline for the beginning of job processing. The particular interest in such RTS was aroused by military intelligence problems (see [1]) involving unmanned aerial vehicles (UAV).

We can summarize the main characteristics of RTS under consideration as follows (see [1]):

(i) Data/jobs acquisition and processing are as fast as the data arrival rate.

(ii) Between data arrival and their acquisition and processing, delays are negligible. Thus, jobs arriving are processed immediately (conditional on system availability) in real time.

(iii) Storage of nonprocessed data is impossible. That part of the job which is not processed immediately is lost forever since it cannot be served later.

[1]Department of Industrial Engineering and Management, Ben-Gurion University of the Negev, P.O.Box 653, Beer-Sheva 84105, Israel, E-mail: kremer@bgu.ac.il

[2]The Institute of Evolution, University of Haifa, Mt. Carmel, Haifa 31905, Israel.

(iv) Tasks are arriving to the RTS, and their times are expiring, *either being processed or lost* (partly or completely), without any connection to servers operation.

Thus, by their very nature, queues of jobs cannot exist in these RTS, nevertheless the use of queueing theory is possible via dual approach of changing the roles between servers and jobs.

2. The Model

We consider a multiserver RTS consisting of N identical servers that provide service for requests of real-time jobs arriving via r *different* channels required to be under nonstop surveillance. There are exactly r_i ($r_i \geq 1$) jobs in ith channel at *any* instant (there are no additional job arrivals to the busy channel), and therefore r_i servers at most are used to serve the ith channel (with others being on stand-by or in maintenance or providing the service to another channel) at any given time.

Each channel has its own specifications and conditions, etc., and therefore different kinds of equipment and inventory are needed to serve different channels.

A server is operative for a period of time S_i before requiring R_i hours of maintenance. S_i and R_i are independent identically distributed random values respectively.

It is assumed that there are ample identical maintenance teams available to repair (with repair times R_i being i.i.d.r.v.) all N servers simultaneously, if needed. Thus, each server coming back from a mission enters maintenance facilities without delay. This server is assigned to the ith channel with probability p_i ($i = 1, \ldots, r$). It receives the appropriate kind of maintenance (equipment, programming, etc.), and therefore cannot be sent to another channel. Assignment probabilities p_i may depend upon inventory conditions. They also can be used as control parameters. The duration R_i of repair does not depend on the channel. After maintenance, the server will either be on stand-by or serving the region it was assigned to.

The system works under a maximum load (worst case) of nonstop data arrival to each one of r channels. This kind of operation is typical in high performance data acquisition and control systems, such as self-guided missiles, space stations, satellites, etc.

If, during some period of time of length T, there is no available server to serve the job in one of the channels, we will say that the part of this job of length T is lost.

3. Main Analytical Results

Suppose that both server's operation and maintenance times S_i and R_i are independent exponentially distributed random variables with parameters μ_i ($i = 1, \ldots, r$) and λ respectively.

Denote: $\lambda_i = \lambda p_i$ and $\rho_i = \lambda_i/\mu_i$. Here p_i is the probability that server going to maintenance is assigned to the ith channel.

Denote $\bar{n} = (n_1, \ldots, n_r)$ the state of the system, where n_i $(i = 1, \ldots, r)$ is a number of fixed servers assigned to the ith channel (obviously $\sum_{i=1}^{r} n_i \leq N$), and $p_{\bar{n}} = p_{n_1,\ldots,n_r}$ the corresponding steady state probability. There are $\binom{N+r}{r}$ states in total. Steady-state probabilities $p_{n_1,n_2,\ldots,n_r}$ are given (see [2]) by the following

Theorem 1. *A real-time system with N servers, ample maintenance facilities, r $(r \geq 2)$ different channels operating under a maximum load regime with exactly r_i $(r_i \geq 1)$ jobs in the ith channel at any instant, and exponentially distributed operating and maintenance times (with parameters μ_i (for the ith channel $(i = \overline{1,r})$) and λ respectively) has steady state probabilities given by the following formulae:*

$$p_{n_1,n_2,\ldots,n_r} = \frac{N!}{\left(N - \sum\limits_{i=1}^{r} n_i\right)!} \prod_{j=1}^{r} \frac{\rho_j^{n_j}}{\min(n_j, r_j)! r_j^{\max(n_j - r_j, 0)}} p_{0,\ldots,0},$$

$$p_{0,\ldots,0} = \left[\sum_{\substack{n_1,\ldots,n_r \geq 0 \\ n_1+\ldots+n_r \leq N}} \frac{N!}{\left(N - \sum\limits_{i=1}^{r} n_i\right)!} \prod_{j=1}^{r} \frac{\rho_j^{n_j}}{\min(n_j, r_j)! r_j^{\max(n_j - r_j, 0)}} \right]^{-1}.$$

Definition 1. *The availability of the RTS with N servers and r channels is*

$$A\nu_N(\rho_1, \ldots, \rho_r) = \sum_{\bar{n}} \left[\sum_{i=1}^{r} \min(n_i, r_i) / \sum_{i=1}^{r} r_i \right] p_{\bar{n}}.$$

Then we have the following (see [2])

Theorem 2. *A RTS with large number of servers $N \to \infty$, ample maintenance facilities, r $(r \geq 2)$ different channels operating under a maximum load regime with exactly r_i $(r_i \geq 1)$ jobs in the ith channel at any instant, and exponentially distributed operating and maintenance times (with parameters μ_i (for the ith channel $(i = \overline{1,r})$) and λ respectively) has following limiting values of average system availability*

$$Av(\rho_1, \ldots, \rho_r) = \lim_{N \to \infty} A\nu_N(\rho_1, \ldots, \rho_r) = \frac{r_u}{\rho_u} \frac{\sum\limits_{n=1}^{r} \rho_n}{\sum\limits_{n=1}^{r} r_n},$$

where $\frac{\rho_u}{r_u} = \max_{i=1,\ldots,r} \frac{\rho_u}{r_u}$ is a maximal ratio.

Finally optimal assignment probabilities are given by the following (see [2])

Theorem 3. *In RTS with large number of servers $N \to \infty$, ample maintenance facilities, r $(r \geq 2)$ different channels operating under a maximum load regime with exactly r_k $(r_k \geq 1)$ jobs in the kth channel at any instant, and exponentially distributed operating and maintenance times (with parameters μ_k (for the kth channel $(k = \overline{1,r})$) and λ respectively) optimal assignment probabilities, which maximize system availability as $N \to \infty$ are $p_k^* = \mu_k r_k / \sum\limits_{i=1}^{r} \mu_i r_i$ for $k = \overline{1,r}$.*

Corresponding optimal limiting value of system availability is $Av\,(\rho_1^*, \ldots, \rho_r^*) = 1$ *and* $(\rho_k^* = \frac{\lambda p_k^*}{\mu_k},\ k = \overline{1, r})$*, i.e. a maximal possible value of system availability at all.*

Full proofs of Theorems 1-3 can be found in [2].

Note 1. *Paradox.*
Consider again results of Theorems 2 and 3. It can be easily seen that if $p_j \neq p_j^*$ for some j (see [3]), then $Av\,(\rho_1, \ldots, \rho_r)$ does not achieve *its maximal* possible value 1. Thus the fact that total number of servers (and maintenance teams) $N \to \infty$, and, therefore, numbers of servers assigned to each one of channels ($Np_j \to \infty$ for $p_j > 0$), does not guarantee *maximal availability.* Why?! The explanation can be found in the following

Example 1. Let $N = 100$, $r = 2$, $r_1 = r_2 = 1$; $\mu_1 = 3$, $\mu_2 = 6$, $\lambda = 10$, then $p_1^* = 1/3$, $p_2^* = 2/3$. Let $p_1 = p = 0.4$, $p_2 = q = 0.6$ and assume that at very beginning numbers of servers, assigned for each one of channels are 50.

After one time unit, $9 = \mu_1 + \mu_2$ (3 routed to the 1st channel and 6 routed to the 2nd channel) servers *on average* will be reassigned after providing service: $9p = 3.6$ to the 1st channel and $9q = 5.4$ to the 2nd channel. Thus, after one time unit *on average*, $6 - 5.4 = 0.6$ servers will be reassigned from the 2^{nd} channel to the 1st. It will take *on average* $50/0.6 = 84$ time units to reroute almost *all* $N = 100$ servers to the 1st channel and to leave the 2nd almost empty.

Maintenance performs here like a kind of time *delay* of $1/\lambda$ on average, since there are ample maintenance facilities.

The following graph presents numbers of fixed servers routed to different channels in the process of simulation.

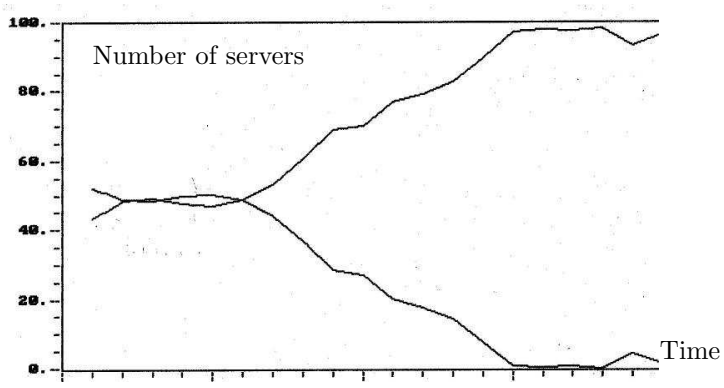


Figure 1: Numbers of servers at different channels.

The upper and lower curves correspond to numbers of servers in the first l and in the second channels respectively.

4. Simulation Results

In this Section we consider again a model presented in Section 2, while waiving the requirement of exponentially distributed operating/service times considered in Section 3. In that case we can not obtain analytically system states distributions, as it has been done in Theorem 1 of Section 3. We will, therefore, use for this purpose, as well for optimization, the Cross Entropy (CE) simulation method.

In that case, at any given moment the state of such a RTS is completely determined by vector $\bar{n} = (n_1, \ldots, n_r)$, where n_k $(k = 1, \ldots, r)$ is a number of fixed servers assigned to the kth channel (see Section 3) and by vector C of clocks associated with each one of servers to indicate the completion of the service in progress (see [3]). Nevertheless, the system availability, as defined by Definition 1 in Section 3, is still completely determined by vector $\bar{n} = (n_1, \ldots, n_r)$ alone.

In order to maximize system availability and to find corresponding optimal assignment probabilities we use CE simulation method (see [4], [5]) and its modification (see [6])

CE Algorithm for continuous optimization

1. Choose an initial set of assignment probabilities $\hat{\mu} = (p_1, \ldots, p_r)$, and an initial set of standard deviations $\bar{\sigma}_0 = \widehat{\sigma}_0 = (1, \ldots, 1)$ (diagonal covariance matrix), i.e. in other words construct initial parameter vector $v_0 = \hat{v}$. Set $t = 1$ (level counter)
2. Generate a sample $\bar{p}^{(m)} = (p_1^{(m)}, \ldots, p_r)^{(m)}$, $m = 1, \ldots, M$, from the Normal densities $N(\hat{\mu}_{t-1}, \hat{\sigma}^2_{t-1})$, normalize vectors $\bar{p}^{(m)}$ so that $\sum_{k=1}^{r} p_k^{(m)} = 1$, $m = 1, \ldots, M$, (*reject and don't take into account* invalid sets with negative $p_k^{(m)}$) and compute the sample $(1-\rho)$-quantile $\hat{\gamma}_t = S_{\lceil (1-\rho)M \rceil}$ of system availability.
3. Update the parameter vector $\hat{v}_t$, as follows: $\hat{\mu}_{t,k} = \sum_{\bar{p}^{(m)} \in E_t} p_k^{(m)} \Big/ |E_t|$ and $\hat{\sigma}_{t,k} = \sqrt{\sum_{\bar{p}^{(m)} \in E_t} (p_k^{(m)} - \hat{\mu}_{t,k}) \Big/ |E_t|}$, where E_t denote the best (largest) $100(1-\rho)$ % of system availabilities estimates.
4. If the stopping criterion is met, STOP; otherwise set $t = t + 1$ and go to Step 2.

To be more specific, we used in Step 1, as an initial set of assignment probabilities, the set of *optimal* assignment probabilities $\hat{\mu}_0 = (p_1^{opt}, \ldots, p_r^{opt})$ obtained deterministically for (*auxiliary*) Jackson network considered in Section 3 with *exponentially* distributed service times having expectations equal to those of models studied in this Section, and all other parameters: N, r, r_k, λ being exactly the same. It proved to be a good guess.

Next, at Step 2 the sample size was chosen to be M=20 and $1 - \rho = 0.25$. Finally, at Step 4 the stopping criterion was chosen to be $\hat{\sigma}_{t,k} = 0$ for any $k = 1, \ldots, r$.

We consider the following

Example 2. RTS with $N = 10$ servers, $r = 3$ channels, $r_1 = 3$, $r_2 = 2$, $r_3 = 1$, maintenance rate $\lambda = 2$, $\mu_1 = 1$, $\mu_2 = 3$, $\mu_3 = 5$, optimal assignment probabilities

$p_1^{opt} = 0.327274$, $p_2^{opt} = 0.49051$ and $p_3^{opt} = 0.182216$. Now we assume that service times are distributed non-exponentially, e.g. uniformly $U(0, 2\mu_k^{-1})$, $k = 1, 2, 3$; normally $N(\mu_k^{-1}, \sigma_k = 1/(4\mu_k))$, $k = 1, 2, 3$; *Erlang* $(2, 1/(2\mu_k))$; but with the same expectations $\mu_1^{-1} = 1$, $\mu_2^{-1} = 1/3$ and $\mu_3^{-1} = 1/5$ as corresponding exponential distributions. Corresponding simulation results are presented in Table 1.

Table 1: Values of p_i^{opt} $(i = \overline{1, r-1})$ and $A\nu_N^{opt} = \max_{\substack{0 \le p_i \le 1 \\ i=\overline{1,r}}} A\nu_N(\rho_1, \ldots, \rho_r)$ obtained from CE and Crude Monte Carlo (CMC) simulations

p.d.f	p_1^{opt} CE/CMC	p_2^{opt} CE/ CMC	p_3^{opt} CE/ CMC	$A\nu_N^{opt}$ CE/ CMC	Nu.of CE iterations
Uniform	.3140/.3140	.5264/.5264	.1596/.1596	.7154/.7154	12
Normal	.3100/.3091	.4779/.4791	.2121/.2118	.7220/.7204	14
Erlang	.3603/.3603	.4564/.4564	.1833/.1833	.6451/.6451	11

For similar model with exponentially distributed service times (see Section 3) we used $(p_1 = 0.2, p_2 = 0.4, p_3 = 0.4)$ as an initial set of assignment probabilities (initial parameter vector $\hat{\mu}_0 = \hat{\nu}_0$), and obtained following results:
$p_1^{opt} = 0.3273$, $p_2^{opt} = 0.4905$, $p_3^{opt} = 0.1822$, $A\nu_N^{opt} = 0.6891$ for deterministic gradient descent algorithm (we have analytical expressions in Theorem 1 for state probabilities);
$p_1^{opt} = 0.3122$, $p_2^{opt} = 0.4646$, $p_3^{opt} = 0.2232$, $A\nu_N^{opt} = 0.6893$ for CE, (7 iterations);
$p_1^{opt} = 0.3077$, $p_2^{opt} = 0.4721$, $p_3^{opt} = 0.2202$, $A\nu_N^{opt} = 0.6871$ for CMC.

Note 2. We can also use optimal limiting ($N \to \infty$, see Theorem 3) assignment probabilities $(p_1^*, p_2^*, p_3^*) = (3/14, 6/14, 5/14)$ obtained in Section 4, as initial set of assignment probabilities (initial parameter vector $\hat{\mu}_0 = \hat{\nu}_0$).

Thus for queueing networks under consideration with four queues and total number of states $\bar{n} = (n_1, \ldots, n_r)$ equal to 286 the CE method works good.

It is worthwhile to note that queueing networks in this work have specific structure. These networks are centralized with maintenance shop node as a central and therefore we have only r assignment probabilities, instead of $(r+1)^2$ such probabilities in general queueing network with $(r+1)$ queues/nodes. This fact makes the treatment of these networks much easier.

Acknowledgement. This research was supported by the Paul Ivanier Center for Robotics Research and Production Management at Ben-Gurion University.

References

[1] J. Kreimer and A. Mehrez, (1994) Optimal Real-Time Data Acquisition and Processing by a Multiserver Stand-By System, *Operations Research*, vol. 42, no. 1, pp. 24-30, January-February.

[2] E. Ianovsky, *Analysis and Optimization of Real-Time Systems*, Ph. D. Thesis, Ben-Gurion University of the Negev, Beer-Sheva, Israel.

[3] P-T. de Boer, (2005) Rare Event Simulation of Non-Markovian Queueing Networks Using a State-Dependent Change of Measure Determined Cross-Entropy, *Annals of Operations Research,* **134**, 69-100.

[4] R.Y. Rubinstein and D.P. Kroese, (2004) *The Cross-Entropy Method: A Unified Approach to Combinatorial Optimization, Monte-Carlo Simulation, and Machine Learning,* Springer-Verlag.

[5] R.Y. Rubinstein and D.P. Kroese, (2008) *Simulation and the Monte Carlo Method*, 2nd edition, John Wiley & Sons, New York.

[6] E. Ianovsky and J.Kreimer, (2009) An Optimal Routing Policy for Unmanned Aerial Vehicles (Analytical and Cross-Entropy Simulation Approach) *Annals of Operations Research* (to appear).

6th St.Petersburg Workshop on Simulation (2009) 1124-1128

Robust Lagrange multiplier test with forward search simulation envelopes

Fabrizio Laurini[1], Luigi Grossi[2], Giorgio Gozzi[3]

Abstract

The Lagrange multiplier test, often adopted to detect heteroscedasticity, suffers from severe size distortion and has low power. An existing robust test based on a forward search algorithm has shown better performance than many existing robust methods. Nevertheless, such a forward robust test relies on confidence bands based on the Student's-t distribution which hold only approximately. The robust forward weighted Lagrange multiplier test can be improved through extensive simulation of forward search confidence bands, which are set up under the hypothesis of no outlier in the data.

1. Introduction

Engle (1982) derived a test for the identification of heteroscedastic components, based on Lagrange multipliers (LM), through an auxiliary regression of residuals of a conditional mean fit $\hat{\eta}_t = y_t - E_y$, where E_y is an ARMA fit for the conditional mean of the process y_t. The auxiliary regression is then:

$$\hat{\eta}_t^2 = \alpha_0 + \alpha_1 \hat{\eta}_{t-1}^2 + \ldots + \alpha_q \hat{\eta}_{t-q}^2 + \varepsilon_t, \qquad t = q+1, \ldots, T, \tag{1}$$

where ε_t is the innovation of the auxiliary regression. The null hypothesis of conditional homoscedasticity can be formulated as $H_0 : \alpha_1 = \alpha_2 = \ldots = \alpha_q = 0$. This way of testing for ARCH components can be computed using $LM = TR^2$, where R^2 is the squared multiple correlation coefficient of model (1). Under the null hypothesis, the LM test statistic has an asymptotic $\chi^2(q)$ distribution. The auxiliary regression (1) can be used to detect GARCH effects.

The LM test is sensitive to outliers, leading to two different wrong conclusions. Van Dijk et al. (1999) show that, in some cases, outliers lead the LM test to reject the null hypothesis of conditional homoscedasticity too often and, in other cases, outliers may hide true heteroscedasticity. These findings motivated the use of a weighted regression LM test (Van Dijk et al., 1999), studied also by

[1]University of Parma, E-mail: fabrizio.laurini@unipr.it
[2]University of Verona, E-mail: luigi.grossi@univr.it
[3]University of Parma, E-mail: giorgio.gozzi@unipr.it

Carnero et al. (2007), that exhibits size distortion in large samples. A different approach involves the robust estimation of autocorrelations of the squared residuals (Duchesne, 2004), combined with the one-sided test for ARCH effects developed by Hong (1997).

The LM test can identify ARCH effects even when the true generating process is an ARCH(1) or a Gaussian *iid* sequence (see Grossi and Laurini, 2004). The robust test of Van Dijk et al. (1999) has low power when data are not contaminated, as it may wrongly detect outliers when the generating process is a Gaussian *iid* sequence.

Grossi and Laurini (2008) address the problems of size distortion and power, proposing a new way of weighting observations for the robust LM test. The robust LM test, computed with the new system of weights, has lower size distortion than both the LM test and the test of Van Dijk et al. (1999), and also has higher power than the test of Van Dijk et al. (1999).

The computation of weights proposed by Grossi and Laurini (2008) exploits the iterative forward search algorithm, and use these weights in the auxiliary regression (1). Such weights measure the distance of each standardized residual from a Student-t confidence band, for all the steps of the iteration. However, the confidence bands do not have a Student's-t distribution for all the forward search steps, but a closed analytical form is difficult to derive. We tackle the problem by simulation of forward search confidence bands under the assumption of no outlier in the data. Some preliminary results show a significant improvement over the Student-t approximation.

2. Robust LM test through the forward search

Our target is to compute weights $w_t \in [0, 1]$, for each observation in the time series y_t, $t = 1, \ldots, T$, with the forward search method. The weights will be used in a weighted linear regression, and they are such that the most outlying observations get small weight.

The forward search is an iterative algorithm which combines robust and efficient estimators. Here we only discuss:

1) the use of very robust methods to sort and split the data into a clean outlier-free small subset, generally called the "clean data set" (CDS), and a bigger subset of potential outliers;

2) the definition of a criterion by which new observations are introduced into the CDS until all observations are included.

In the following we focus on the LM test-statistic linked to the alternative hypothesis of ARCH components, but similar steps will be used when testing for the presence of other effects, with the only difference being the initial auxiliary regression (1).

2.1. Identification of the clean data set

Observations belonging to the CDS are identified by applying least median of squares.

If data are independent, then residuals can be take in random order, but if, as in a time series, there is dependence in data, then the forward search starts by taking blocks of contiguous residuals. In all cases, the number of initial observations in (1) is $k = q + 1$, for which the auxiliary regression model is fully identified.

Let $z_t' = (1, \hat{\eta}_t^2, \ldots, \hat{\eta}_{t-q}^2)$, with $t = k, \ldots, T$, be the t-th row of the matrix Z of dimension $(T - k) \times k$. If T is not too large and $k \ll T$ we can find the best outlier-free subset, through an exhaustive search of all possible C_{T-k}^k distinct k-tuples $S_{t_j,\ldots,t_{j+k}}^{(k)} \equiv \{z_{t_j}, \ldots, z_{t_{j+k}}\}$ where $z_{t_j}', \ldots, z_{t_{j+k}}'$ is a block of k rows of Z, with $k \leq t_j, \ldots, t_{j+k} \leq T$. In the special case of time series such rows will be consecutive.

Define the set of time indices $\tau = \{t_j, \ldots, t_{j+k}\}$ and let $e_{t,S_\tau^{(k)}}$ be the least squares residual of model (1) for observation t given observations in $S_\tau^{(k)}$. The CDS is a subset satisfying the LMS criterion. Namely, among all blocks of size k, we select the subset of k observations $S_*^{(k)}$ which minimizes the median of the squared residuals, i.e.

$$e^2_{[med],S_*^{(k)}} = \min_\tau \{e^2_{[med],S_\tau^{(k)}}\},$$

where $e^2_{[i],S_\tau^{(k)}}$ is the i-th ordered squared residual among $e^2_{i,S_\tau^{(k)}}$, $i = k, \ldots, T$ and med is the sample median, i.e. the integer part of $(T + 1)/2$. If C_{T-k}^k is too large, we select the CDS from a large number of samples, perhaps 10000 or 50000. We remark that the final results are not affected by the choice of the number of samples used to define the initial CDS.

2.2. Including observations into the CDS

Given the best subset $S_*^{(m)}$ of dimension $m \geq k$ detected at step m, we move to a subset of size $m + 1$ sorting all observations non-decreasingly according to the squared OLS residuals $e^2_{t,S_*^{(m)}}$, $t = k, \ldots, T$. The new subset is given by the $m+1$ observations with the smallest squared OLS residuals. The procedure is repeated until $m = T$, i.e. until all observations are included in the CDS. In the first steps, the forward search avoids the inclusion of outliers.

In most moves from m to $m + 1$ just one new observation joins the subset. However, it may be the case that l or more observations join the subset as $l - 1$ leave it. This behavior is called "interchange" and it is frequent in the first steps of the forward search. A very important feature of this forward approach is that it involves, at the same time, a robust method with high breakdown point (selection of the CDS) and the fully efficient least squares estimators (inclusion of new observations into the CDS).

2.3. Computation of weights

Let $y_t = e_t^2$ and $x_t' = (1, e_{t-1}^2, \ldots, e_{t-q}^2)$, with $t = k, \ldots, T$ and $\alpha' = (\alpha_0, \alpha_1, \ldots, \alpha_q)$ be a vector of length $k = q + 1$ with the parameters of model (1). Now define

$$\tilde{e}_t = \frac{y_t - x_t' \hat{\alpha}_m^*}{\hat{\sigma}_{S_*^{(m)}} \sqrt{1 + x_t' (X_{S_*^{(m)}}' X_{S_*^{(m)}})^{-1} x_t}}, \tag{2}$$

where $X_{S_*^{(m)}}$ is the $m \times k$ matrix with units belonging to the CDS, whereas $\hat{\alpha}_m^*$ and $\hat{\sigma}^2_{S_*^{(m)}} = \sum_{i=1}^m e^2_{i,S_*^{(m)}}/(m)$ are, respectively, the least squares estimator and the residual mean square estimate, based on the m observations belonging to the CDS.

The residuals $\tilde{e}$ are defined similarly to the externally studentized residuals used also by Atkinson and Riani (2000). Externally studentized residuals have a Student-t distribution with $T-k-1$ degrees of freedom when all the observations belong to the CDS. For intermediate steps such a result is no longer true, and we take, as the reference distribution, the confidence bands derived from the forward search simulation of outlier-free data. Such simulated confidence bands depend, among other things, on the sample size, the step of the forward search and the order of the auxiliary regression. This creates new computational problems that have to be tackled efficiently.

2.4. Comparison of trajectories

Following Atkinson (1994), we compare the trajectories of $\tilde{e}_t$ during the forward search with confidence bands built from simulated envelopes.

At each step of the forward search, we measure the degree of outlyingness of each observation $t = 1, \ldots, T$, as the squared Euclidean distance, π, between the residuals (2) lying outside a confidence band and the boundaries of the band itself.

For a fixed step of the forward search m, we record the distance of the t-th observation from the quantile of the confidence band, provided that the t-th observation is outside such a confidence band.

If the t-th residual lies between the simulated quantiles, then, at the m-th step, it will get zero distance. At the next step $m+1$, the weight of the t-th observation will be increased by an amount which is induced by the squared Euclidean distance from the t-th residual and the simulated quantile of the confidence band, provided that at the step $m+1$ the t-th observation exceeds the observed empirical level. If at the step $(m+1)$the t-th observation lies between the simulated quantiles, then a zero will be added to the distance computed at step m.

The overall degree of outlyingness for the t-th observation is given by the sum of all squared Euclidean distances, computed only when the observation exceeds the confidence bands. Such an overall sum is scaled by the number of steps of the forward search $T - k$.

Formally, at the nominal level $1 - \delta$, we define $\pi_m^{(t)}$ to be the distance between $\tilde{e}_{t,S_*^{(m)}}$ and the simulated quantile $z_{\delta/2}$ for unit t at step m, i.e. the squared

Euclidean distance will be

$$\pi_m^{(t)} = \begin{cases} 0 & \text{if } \tilde{e}_{t,S_*^{(m)}} \in [-z_{\delta/2}, +z_{\delta/2}], \\ (\tilde{e}_{t,S_*^{(m)}} - z_{\delta/2})^2 & \text{if } \tilde{e}_{t,S_*^{(m)}} > z_{\delta/2}, \\ (\tilde{e}_{t,S_*^{(m)}} + z_{\delta/2})^2 & \text{if } \tilde{e}_{t,S_*^{(m)}} < -z_{\delta/2}, \end{cases}$$

and we consider the overall distance of the t-th observation as the sum of such distances during the forward search, i.e.

$$\pi_t = \frac{\sum_{m=k}^{T} \pi_m^{(t)}}{T-k}. \tag{3}$$

2.5. Mapping the squared Euclidean distance

The squared Euclidean distance measures the degree of outlyingness of each observation through the computation of a weight, in the interval $[0, 1]$, obtained with the following mapping of (3)

$$w_t = \exp(-\pi_t). \tag{4}$$

Hence, influential observations that have large squared residuals (large Euclidean distance) during many steps of the forward search, will get a small weight w_t.

2.6. Weighted regression

The last step of the procedure is based on a weighted version of (1), where the vector of weighted parameter estimates α^* is given by

$$\alpha^* = (X'WX)^{-1}X'Wy. \tag{5}$$

Here, W represents the weight matrix and it is a $(T-k)\times(T-k)$ diagonal matrix having, in the main diagonal, the weights w_t for all observations computed using (4). Hence, the robust version of the LM test is simply given by

$$LM^* = (T-k)R_*^2,$$

where R_*^2 is the coefficient of determination of the linear model with parameter estimates obtained from (5).

3. Preliminary results

Our simulations show that the weighted forward test does not suffer from size distortion, even when data are contaminated by several individual outliers or by patches of outliers. Moreover, the forward weighted test has higher power than the backward weighted test of Van Dijk et al. (1999). The difference in their performances relies on the capability of the forward search method to avoid the so called "masking" and "swamping" effects.

The weights computed with the forward search, with the transformation induced by $\exp(-x)I_{x>0}$, can be generalized to some flexible parametric model, measuring the squared Euclidean distance in different ways. Such weights can eventually be used to provide robust estimates of GARCH parameters.

References

Atkinson, A. C. (1994) Fast very robust methods for the detection of multiple outliers. *Journal of the American Statistical Association*, 89, 1329–1339.

Atkinson, A. C. and Riani, M. (2000) *Robust Diagnostic Regression Analysis*. Springer–Verlag, New York.

Carnero, M.A., Peña, D. and Ruiz, E. (2007) Effects of outliers on the identification and estimation of GARCH models. *Journal of Time Series Analysis*, 28, 471–497.

Duchesne, P. (2004) On robust testing for conditional heteroscedasticity in time series models. *Computational Statistics and Data Analysis*, 46, 227–256.

Engle, R. F. (1982) Autoregressive conditional heteroscedasticity with estimates of the variance of UK inflation. *Econometrica*, 50, 987–1007.

Grossi, L. and Laurini, F. (2004) Analysis of economic time series: effects of extremal observations on testing heteroscedastic components. *Applied Stochastic Models in Business and Industry*, 20, 115–130.

Grossi, L. and Laurini, F. (2008) A robust forward weighted Lagrange multiplier test for conditional heteroscedasticity. *Computational Statistics and Data Analysis*, to appear.

Hong, Y. (1997) One-sided testing for conditional heteroscedasticity in time series models. *Journal of Time Series Analysis*, 18, 253–277.

Van Dijk, D., Franses, P. H. and Lucas, A. (1999) Testing for ARCH in the presence of additive outliers. *Journal of Applied Econometrics*, 14, 539–562.

6th St.Petersburg Workshop on Simulation (2009) 1130-1134

Parameter Estimation and Tracking For the Rosenbrock Function Using Simultaneous Perturbation Stochastic Approximation Algorithm

Oleg Granichin, Lev Gurevich, Alexander Vakhitov[1]

Abstract

In this paper application of the Kiefer-Wolfowitz algorithm with randomized differences to the minimum tracking problem for the non-constrained optimization is considered. The upper bound of mean square estimation error is determined in case of once differentiable functional and almost arbitrary noises. Numerical simulation of the estimates stabilization for the multidimensional optimization with non-random noise is provided.

1. Introduction

Non-stationary optimization problems can be described in discrete or continuous time. In our paper we consider only discrete time model. Let $f(x, n)$ be a functional we are optimizing at the moment of time n ($n \in \mathbf{N}$). In book [2] Newton method and gradient method are applied to problems like that, but they are applicable only in case of two times differentiable functional and $l < \nabla^2 f_k(x) < L$. Both methods require possibility of direct measurement of gradient in arbitrary point.

Algorithms of the SPSA-type with one or two measurements per each iteration appeared in papers of different researchers in the end of the 1980s [5, 6, 7, 8]. These algorithms are known for their applicability to problems with almost arbitrary noise [4]. Moreover, the number of measurements made on each iteration is only one or two and is independent from the number of dimensions of the state space d. This property sufficiently increases the rate of convergence of the algorithm in multidimensional case ($d >> 1$), comparing to algorithms, that use direct estimation of gradient, that requires $2d$ measurements of function in case if direct measurement of function gradient is impossible. Detailed review of development of such methods is provided in [4, 9].

Stochastic approximation algorithms were initially proven in case of the stationary functional. The gradient algorithm for the case of minimum tracking is

[1]Saint Petersburg State University, E-mail: oleg_granichin@mail.ru gurevich.lev@gmail.com alexander.vakhitov@gmail.com

provided in [2], however the stochastic setting is not discussed there. Further development of these ideas could be found in paper [1], where conditions of drift pace were relaxed.

In this paper we consider application of SPSA algorithm to the problem of tracking of the functional minimum. The closest case was studied in [10], but we do not use the ODE approach and we establish more wide conditions for the estimates stabilization. In the following section we will give the problem statement that is more general than in [1, 2], in the third section we provide the algorithm and prove its estimates mean squared stabilization. In the last section we illustrate the algorithm, applying it to minimum tracking in a particular system.

2. Problem Statement

Consider the problem of minimum tracking for average risk functional:

$$f(x,n) = E_w\{F(x,w,n)\} \to \min_x, \tag{1}$$

where $x \in \mathbf{R}^d$, $w \in \mathbf{R}^p$, $n \in \mathbf{N}$, $E_w\{\cdot\}$ — mean value conditioned on the minimal σ-algebra in which w is measurable.

The goal is to estimate θ_n — minimum point of functional $f(x,n)$, changing over time: $\theta_n = argmin_x f(x,n)$.

Let us assume that on the iteration we can measure:

$$y_n = F(x_n, w_n, n) + v_n, \tag{2}$$

where x_n — arbitrary measurement point chosen by algorithm, w_n — random values, that are non-controlled uncertainty and v_n — observation noise.

Time in our model is discrete and implemented in number of iteration n.

To define the quality of estimates we will use the following definition:

Definition [11] A random matrix (or vector) sequence $\{A_k, k \geq 0\}$ defined on the basic probability space $\{\Omega, \mathcal{F}, P\}$ is called L_P-stable ($p > 0$) if

$$\sup_{k \geq 0} E\|A_k\|^p < \infty.$$

We will use the definition in case of $p = 2$.

Further we will consider estimates $\{\hat{\theta}_n\}$ for the problem (1), satisfying the definition for p=2, in following conditions.

(A) Drift satisfies: $\|\theta_n - \theta_{n-1}\| \leq A$.

(B) Function $f(\cdot, n)$ is strictly convex for each n: $\langle \nabla f(x,n), x - \theta_n \rangle \geq \mu \|x - \theta_n\|^2$.

(C) Gradient $\nabla F(\cdot, w, n)$ is Lipschitz $\forall n, \forall w$:
$\|\nabla F(x,w,n) - \nabla F(y,w,n)\| \leq B\|x - y\|$.

(D) Average difference of function $F(x,\cdot,n)$ in any point x for moments n and $n+1$ is limited:
$E_{w_1,w_2}|F(x,w_1,n+1) - F(x,w_2,n)| \leq C\|x - \theta_n\| + D$.

(E) Local Lebesgue property for the function $\nabla F(w,x)$: $\forall x \in \mathbb{R}^d \exists$ neigbourhood U_x such that $\forall x' \in U_x \ \|\nabla F(w,x)\| < \Phi_x(w)$ where $\Phi_x(w) : \mathbb{R}^p \to \mathbb{R}$ is integrable by w: $\int_{\mathbb{R}^q} \Phi_x(w)dw < \infty$

(F) The observation noise v_n satisfies:$|v_{2n} - v_{2n-1}| \leq \sigma_v$, or if it has statistical nature then: $E\{|v_{2n} - v_{2n-1}|^2\} \leq \sigma_v^2$.

Here we should make several notes:

1). Sequence $\{v_n\}$ could be of non-statistical but unknown deterministic nature. 2). Constraint (**A**) allows both random and deterministic drift. In certain cases Brownian motion could be described without tracking. Tracking is needed when there is both determined and non-determined aspects of drift. Similar condition is introduced in [2], it is slightly relaxed in [1]. For example it could be relaxed in a following way:

$$(\mathbf{A}') \quad \theta_n \leq A_1\theta_{n-1} + A_2 + \xi_n,$$

where ξ_n is random value.

In this paper we will only consider drift constraints in form **(A)**. Mean square stabilization of estimation under condition **(A)** implies its applicability to wide variety of problems.

3. Algorithm and Stabilization Of Estimates

In this section we are introducing a modification of SPSA algorithm provided by Chen et al [12], which takes one perturbed and one non-perturbed measurement on each step.

Let perturbation sequence $\{\Delta_n\}$ be an independent sequence of Bernoulli random vectors, with component values $\pm 1/\sqrt{d}$ with probability $\frac{1}{2}$. Let vector $\hat{\theta}_0 \in \mathbb{R}^d$ be the initial estimation. We will estimate a sequence of minimum points $\{\theta_n\}$ with sequence $\{\hat{\theta}_n\}$ which is generated by the algorithm with fixed stepsize, which is applied to the observations model (2) :

$$\begin{cases} x_{2n} = \hat{\theta}_{2n-2} + \beta\Delta_n,\ x_{2n-1} = \hat{\theta}_{2n-2}, \\ \\ \hat{\theta}_{2n} = \hat{\theta}_{2n-2} - \frac{\alpha}{\beta}\Delta_n(y_{2n} - y_{2n-1}), \hat{\theta}_{2n-1} = \hat{\theta}_{2n-2}. \end{cases} \tag{3}$$

(**G**) We will further assume that random values Δ_n generated by algorithm are not dependent on $\hat{\theta}_k$, w_k , $\hat{\theta}_0$ and on v_k (if they are assumed to have random nature) for $k = 1, 2, \ldots, 2n$.

Let us define $H = (\alpha^2 B + \frac{\alpha^2}{\beta^2}D + \frac{\alpha^2}{\beta^2}\sigma_v)(\beta B + C) + \alpha AB + \alpha\beta B + A$. Let K and $\delta > 0$ be constants satisfying following condition: $K = 1 - 2\alpha\mu + \frac{\alpha^2}{\beta}B + \frac{\alpha^2}{\beta^2}C + \delta < 1$. Denote $L = \frac{H^2}{\delta} + A^2 + 2\alpha\beta AB + + \frac{\alpha^2}{\beta^2}((\beta^2 B + D)^2 + \sigma_v^2 + 2(\beta^2 B + D)\sigma_v)$.

Theorem 1 Assume that conditions **(A)**–**(G)** on functions f, F and ∇F and values θ_n, $\hat{\theta}_n$, v_n, w_n, y_n and Δ_n are satisfied. Let us further assume that constants $\alpha, \beta > 0$ satisfy the inequality:

$$0 < \delta < 2\alpha\mu - \frac{\alpha^2}{\beta}B - \frac{\alpha^2}{\beta^2}C. \tag{4}$$

Then estimates provided by the algorithm (3) stabilize in mean squares and following inequality holds:

$$E\{\|\theta_n - \hat{\theta}_n\|^2\} \leq K^n \|\theta_0 - \hat{\theta}_0\|^2 + \frac{L(1-K^n)}{1-K}. \quad (5)$$

Note that Theorem 1 provides asymptotically effective value for the estimates: $\bar{L} = L/(1-K)$.

Conditions **(A)**–**(C)**,**(E)**–**(G)** are standard for SPSA algorithms [4]. Earlier the proof of the similar theorem was given in [13] with more strict conditions. See the proof of Theorem 1 in appendix.

To build the upper bound of the algorithm's estimates error using Theorem 1, it is needed to find α and β satisfying the condition (4), then to find δ which gives minimum value of the fraction $\frac{L}{1-K}$. The following theorem provides the basis for adjustment of the step-size α depending on the drift and noise parameters.

Theorem 2 Assume the conditions of the Theorem 1 hold. Then, for small $\alpha > 0$ and expressions $C_1(\beta, \mu, B, C, D, \sigma_v)$ and $C_2(\beta, \mu, A, B, C, D)$ the following inequality holds:

$$E\|\hat{\theta}_n - \theta_n\|^2 \leq O((1-\alpha\mu)^n) + O(\alpha^{-2}\frac{A^2}{\mu^2} + \alpha^{-1}\frac{A^2}{\mu} + \\ + A * C_1(\beta, \mu, B, C, D) + \alpha(C_2(\beta, \mu, A, B, C, D) + \sigma_v). \quad (6)$$

The proof of the Theorems 1 and 2 follows.

The result of the Theorem 2 shows that without drift (in case of $A = 0$) α should be made smaller as noise level σ_v grows. In the same time, the bigger is the drift norm, the larger should be the step size α. This tradeoff was demonstrated also in case of the linear system [14].

4. Example

Simple practical application of the algorithm (3) is estimation of the multidimensional moving point coordinates when only information about distance from arbitrary point to the moving point is available with additive noise. As a result of Theorem 1, the algorithm (3) provides the point estimate in case of limited drift of the point and somehow limited observation noise. In [13]the drift with formula $\theta_n = \theta_{n-1} + \zeta$, where ζ is uniformly distributed on the sphere: $\|\zeta\| = 1$. The function $\|x - \theta_n\|^2$ was measured with additional non-random noise sequence $\|v_n\| < 1$. The estimates have shown convergence to the theoretically proven interval.

Here we would like to consider a Rosenbrock test function [15], modified so that the minimum of it is changing with changes of the parameter T:

$$F(x) = \sum_{i=1}^{5} 100(x_{2i} - x_{2i-1}^2)^2 + (T_n - t_{2i-1})^2, \quad (7)$$

where the minimum value is 0 while $\theta = argminF = (T, T^2, T, T^2, ..., T^2)$. The function has many local minima, that is why we test the tracking behavior of

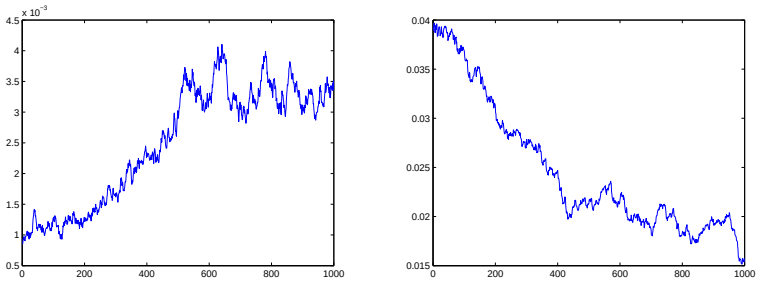


Figure 1: Estimation error $\|\hat{\theta}_n - \theta_n\|^2$ (left) and loss function values $F(\hat{\theta}_n)$ (right)

the algorithm proposed in the small neighbourhood of the optimal point θ. The function does not fully satisfy the conditions of the theorem above, however it is a well-known test function for the optimization algorithms [15]. The T parameter in our experiment performs random walk starting from $T_0 = 1$ with Bernoulli steps of size 0.002. In the same time, the noise v added to the observations has non-random nature and is of the nearly same scale as the minimum drift: $v_{2n} = 0.002(1 - (n \bmod 3))$, $v_{2n-1} = 0.002(1 - 3 * (n \bmod 7))$. In this setting, we used $\beta = 0.0001$, $\alpha = 0.000005$. At the Fig. 1 the estimation error in the form $\|\hat{\theta}_n - \theta_n\|^2$ (left), and the loss function $F(\cdot)$ values $F(\hat{\theta}_n)$ (right) are presented. From the figures it can be seen that the estimates stabilize at the certain distance from the minimum, and the loss function values diminish.

4. Conclusion

In our work we apply the SPSA-type algorithm to the problem of extreme point tracking with almost arbitrary noise. Drift is only assumed to be limited, which includes random and directed drift. It was proven that the estimation error of this algorithm is limited with constant value. The modeling was performed on a multidimensional case.

The authors want next to prove more precise boundaries of the estimation error. The stabilization of estimates for arbitrary p rather than for $p = 2$ (as in this paper) could be considered. It could be also interesting to modify the algorithm to work with unknown polynomial drift, using the technique of polynomial fitting demonstrated in [16].

5. Proof

Theorem 2.

Let us denote $\bar{L} = L - \frac{H^2}{\delta}$, $\bar{K} = K - \delta$. Then, the asymptotic boundary from the Theorem 1 $\frac{L}{1-K} = \frac{\bar{L}+\frac{H^2}{\delta}}{1-\bar{K}-\delta}$. We will firstly optimize this expression by δ. The derivative's nominator is here $-\frac{H^2}{\delta^2}(1-\bar{K}-\delta)+\bar{L}+\frac{H^2}{\delta} = -\frac{H^2(1-\bar{K})}{\delta^2}+\frac{2H^2}{\delta}+\bar{L} =$

$P(1/\delta)$. The expression got is a quadratic function of $1/\delta$ with negative sign near the second argument degree, therefore its maximum value is achieved in the point where $1/\delta = 1/(1-\bar{K})$, $P(1/(1-\bar{K})) = H^2/(1-\bar{K}) + \bar{L} > 0$. Because of that, greater root of the equation $P(1/\delta) = 0$ is a maximum point while smaller one is a minimum point. Solving the equation, we get $1/\delta_{opt} = 1/(1-\bar{K}) \pm \sqrt{H^4 + \bar{L}H^2(1-\bar{K})}/(H^2(1-\bar{K}))$, and $\delta_{opt} = (1-\bar{K})/(1+\sqrt{1+\frac{\bar{L}(1-\bar{K})}{H^2}})$.

From this formula we find that $\delta = O(\alpha\mu)$ for $\alpha \to 0$. Then $K^n = O((1-\alpha\mu)^n)$.

Let us substitute this expression for δ into formulas for L and K. Then, $\frac{L}{1-K} = \frac{\bar{L}}{(1-\bar{K})(1-\frac{1}{1+\sqrt{1+\frac{\bar{L}(1-\bar{K})}{H^2}}})} + \frac{H^2(-1+\sqrt{1+\frac{\bar{L}(1-\bar{K})}{H^2}})}{(1-\bar{K})^2(1-\frac{1}{1+\sqrt{1+\frac{\bar{L}(1-\bar{K})}{H^2}}})}$. From these formulas the coefficients for -1 and -2 degrees of α in the expansion can be easily obtained. The coefficients for the degrees 0 and 1 were found using the Maxima symbolic math system (*http://maxima.sourceforge.net*).

References

[1] Popkov A. Yu. (2005) Gradient methods for nonstationary unconstrained optimization problems. *Automat. Remote Control*, No. 6, pp. 883–891.

[2] Polyak B. T. (1987) *Introduction to Optimization.* New York: Optimization Software.

[3] Kushner H., Yin G. (2003) *Stochastic Approximation and Recursive Algorithms and Applications.* Springer.

[4] Granichin O. N., Polyak B. T. (2003) *Randomized Algorithms of an Estimation and Optimization Under Almost Arbitrary Noises.* Moscow: Nauka.

[5] Granichin O. N. (1992) Procedure of stochastic approximation with disturbances at the input," *Automat. Remote Control*, vol. 53, No 2, part 1, pp. 232–237.

[6] Granichin O. N. (1989) A stochastic recursive procedure with dependent noises in the observation that uses sample perturbations in the input. *Vestnik Leningrad Univ. Math.*, vol. 22, No 1(4), pp. 27–31.

[7] Polyak B. T., Tsybakov A. B. (1990) Optimal order of accuracy for search algorithms in stochastic optimization. *Problems of Information Transmission*, vol. 26, No 2, pp. 126–133.

[8] Spall J. C. (1992) Multivariate stochastic approximation using a simultaneous perturbation gradient approximation. *IEEE Trans. Automat. Contr.*, vol. 37, pp. 332–341.

[9] Spall J. C. (1994) Developments in stochastic optimization algorithms with gradient approximations based on function measurements. *Proc. of the Winter Simulation Conf.*, pp. 207–214.

[10] Borkar V. S. (2008) *Stochastic Approximation. A Dynamical Systems Viewpoint.* Cambridge University Press.

[11] Guo L. (1994) Stability of recursive stochastic tracking algorithms *SIAM J. Control and Optimization*, vol. 32, No 5, pp. 1195–1225.

[12] Chen H.-F., Duncan T. E., Pasik-Duncan B. A (1999) Kiefer-Wolfowitz algorithm with randomized differences. *IEEE Trans. Automat. Contr.*, vol. 44, No 3, pp. 442–453.

[13] Gurevich L., Vakhitov A. (2008) SPSA Algorithm for Tracking. *In Proc. 12th International Student Olympiad on Automatic Control (Baltic Olympiad)*, pp. 52–57.

[14] Granichin O. N. (1999) Linear Regression and Filtering Under Nonstandard Assumptions (Arbitrary Noise). *IEEE Trans. Automat. Contr.*, vol. 44, No 3, 442–453.

[15] Spall J.C. (2003) *Introduction to Stochastic Search and Optimization.* Wiley & Sons, New Jersey.

[16] Katkovnik V. Ya., Khejsin V. E. (1979) Dynamic stochastic approximation of polynomial drifts *Automat. Remote Control*, vol. 40, No 5, pp. 700–708.

6th St.Petersburg Workshop on Simulation (2009) 1137-1141

Estimation of the proportion of true null hypotheses among dependent tests

Chloé Friguet[1], David Causeur[2]

Abstract

Whole genome microarray experiments has markedly contributed to the development of the statistical methodology for multiple testing in high dimensional data. In this context, the impact of dependence on error control is especially crucial. Many recent papers (see for instance [2, 5]) mention that dependence produces high variability and bias for the parameters of the procedures, and consequently misleading values for the estimated error rates. In many realms of application, particularly for the study of microarray experiments, data contain large scale correlation structures, which can be handled by specific models. We propose a methodology based on a Factor Analysis model, which shows improvements with respect to existing Multiple Testing Procedures [6], and provides a general framework for multiple testing dependence. In this presentation, we focus on the estimation of the number of true null hypotheses which is a key parameter to manage with power in MTP.

Key words *Multiple testing, Factor analysis model, Dependence, Null hypotheses proportion*

1. Introduction

In multiple testing, a challenging issue is to provide an accurate estimation of the proportion of true null hypotheses, hereafter denoted π_0, among the whole set of tests. Besides a biological interpretation, this parameter is also involved in the control of error rates such as the False Discovery Rates (FDR) (see [1]). Improving its estimation can result in more powerful/less conservative methods of differential analysis (see [2]).

Two kinds of estimation methods for π_0 have been developed in the literature, briefly presented in the first part of the presentation: either based on Schweder and Spjøtvoll's approximation (see [9]) or on nonparametric maximum likelihood

[1]Agrocampus Ouest - Applied mathematics department, Rennes, France, E-mail: chloe.friguet@agrocampus-ouest.fr

[2]Agrocampus Ouest - Applied mathematics department, Rennes, France, E-mail: david.causeur@agrocampus-ouest.fr

estimation of the p-values' density function (see [8]). Both rely on the assumption of a two-component mixture model for the distribution of the p-values and furthermore a uniform distribution is assumed for p-values associated to true null hypotheses. Therefore, the estimators of π_0 are all derived under the assumption of independent test statistics.

However, dependence among variables, which is observed in microarray data for example, leads to high variability of the estimations. We propose a general framework to deal with dependence, considering a Factor Analysis model for the correlation structure [6]. After a recall about the Factor Analysis settings for MTP, a conditional estimator of π_0 is introduced. At last, a study of this new estimator's performance is presented, through comparisons with other estimators on simulated data, considering a range of scenarios for the correlation structure (from low to high levels of dependence). We show that taking advantage of conditional independence on the factors yields more accurate estimation of π_0.

2. Estimating the proportion of null hypotheses

Most existing estimators of π_0 rely on the following condition (see [3]), that ensures the identifiability of the parameter:

$$\exists t \in]0;1]/\forall k \in \{1;\ldots;m\} \qquad p_k \in [t;1] \Rightarrow p_k \sim \mathcal{U}[t;1], \tag{1}$$

where $p_1, p_2, \ldots, p_m$ are the p-values for the m simultaneous tests.

If this assumption holds, then the p-values associated to true null hypotheses are uniformly distributed. Broadly speaking, two kinds of methods are found in the literature about π_0 estimation: the first one is based on an intuitive reasoning proposed by Schweder and Spjøtvoll (see [9]) and the second one makes use of dedicated algorithms (see [8]) to estimate the density function of the p-values.

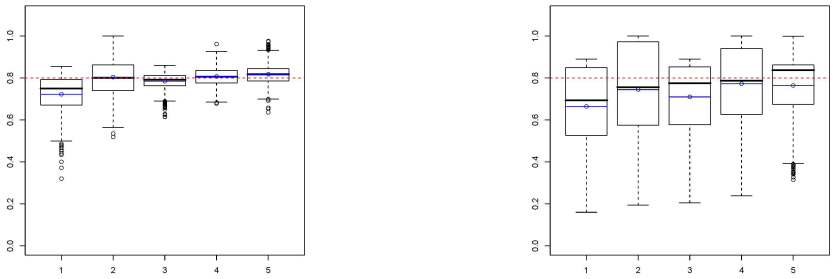
Intuitive estimator. Let us define W_t as the number of p-values greater than a threshold t: $W_t = \#\{k \in [1;m],\ p_k > t\}$. If t is not too small, a large majority of the p-values in $[t;1]$ should be associated to true null hypotheses (see condition (1)). Then, $\mathbb{E}(W_t) = m_0(1-t)$ and the following estimator of π_0 is deduced (see [9]):

$$\hat{\pi}_0(t) = \frac{W_t}{m(1-t)} \tag{2}$$

A relevant choice of t should result from a bias-variance trade-off for $\hat{\pi}_0(t)$. Storey's bootstrap method (see [10]) or Storey and Tibshirani's smoothing method (see [11]) can be used to achieve a good compromise.

Estimators based on estimating the p-values' density function. As suggested in [3] or [4], the distribution of the p-values, which density function is denoted f, can be modeled by a mixture of two components :

$$f(p) = \pi_0.g(p) + (1-\pi_0).h(p) \tag{3}$$



Independent simulated data | Highly correlated data

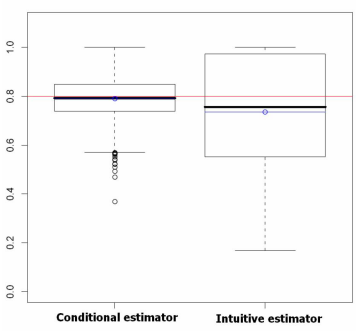
Figure 1: Estimation of π_0 from p-values of t-tests with different methods for 1000 simulated datasets ($\pi_0 = 0.8$)
Method 1 : intuitive estimator with splines smoothing choice of t [9, 11]
Method 2 : intuitive estimator with bootstrap choice of t [9, 10]
Method 3 : Convex estimation of p-value density [8]
Method 4 : Kernel estimation of p-value density [8]
Method 5 : Grenander estimator of p-value density + longest constant interval [8]

Assuming $h(1) = 0$ to achieve identifiability of the model's parameters, an estimation of π_0 is obtained for $p = 1$. Various dedicated algorithm are derived to estimate f with this intent (see for example [8]).

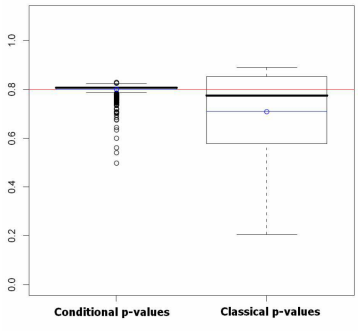
Impact of dependence on the estimation of π_0. The violation of the independence assumption leads to high variability in the parameter estimation. Figure 1 illustrates this phenomenon on independent **??** and highly correlated **??** simulated data.

3. Estimation of the proportion of null hypotheses under dependence

Factor Analysis: a general framework to deal with dependence in MTP. We propose to model the common information shared by all the variables using Factor Analysis (FA). FA is mostly used by social scientists or psychometricians as a dimension reduction technique and has only appeared recently as an interesting tool to investigate the dependence structure in high dimensional microarray datasets (see [7]). This model describes the correlation between the observed variables by a small number of latent variables, called the common factors. The covariance matrix Σ can be expressed as $\Sigma = B.B' + \Psi$ where B is a matrix of loadings associated to the common variability and Ψ a diagonal matrix for variable specific variance also called "uniqueness". The parameters are estimated thanks to an EM-algorithm.



Comparison between the intuitive estimator and the conditional estimator



Method based on density function estimate: comparison between unconditional and conditional p-values

Figure 2: Conditional approach on simulated data with a high correlation scheme ($\pi_0 = 0.8$)

Within this framework, a new test statistics is defined (see [6]), taking into account the factor structure to improve the power of multiple testing procedures, while still controlling the type-I error rate.

Conditional estimation of π_0. First, let us denote Z the factors extracted from the data. $\mathbb{E}\left(W_t|Z\right) = m_0 - \sum_{k\in\mathcal{M}_0} t_Z^{(k)}$ where $t_Z^{(k)} = \mathbb{P}(p_k < t|Z)$. Inspired from [9], a conditional estimator of π_0 is $\hat{\pi}_0^{(c)} = W_t/m(1-\bar{t}_Z)$. The cut-off t is then chosen so as to have an optimal trade-off between bias and variance. This estimator is shown to correct the effect of dependence observed on the intuitive estimator presented before (see **??**). Another way to estimate π_0 is to consider the density of conditional p-values derived from factor-adjusted test statistics. Conditioning on the factor structure yields independence between the p-values, which satisfies the assumption for the estimator based on a density function estimate. Therefore again, the variability in the estimation is strongly reduced compared with p-values of classical t-tests (see **??**).

4. Conclusion

Considering a general framework for dependence in multiple testing by means of a Factor Analysis model helps reducing the negative effect of dependence on the variability of error rates. Moreover, this framework allows the definition of conditional methods of estimation for the proportion of true null hypotheses, which are more accurate than classical ones in presence of dependence between the data.

References

[1] Benjamini Y. and Hochberg Y. (1995) *Controlling the False Discovery Rate: a practical and powerful approach to multiple testing* JRSS B, 57:289-300

[2] Black M. A. (2004) *A note on the adaptative control of false discovery rates* JRSS B, 66:297-304

[3] Celisse A. and Robin S. (2008) *Nonparametric density estimation by explicit leave-p-out cross-validation* Comput. Statist. Data Analysis, 52:2350-2368

[4] Efron B., Tibshirani R., Storey J.D. and Tusher V.(2001) *Empirical bayes analysis of a microarray experiment* JASA, 96:1151-1160

[5] Efron B. (2007) *Correlation and large-scale simultaneous testing* JASA, 102:93-103

[6] Friguet C., Kloareg M. and Causeur D. (2009) *A factor model approach to multiple testing under dependence* JASA, to appear

[7] Kustra R., Shioda R. and Zhu M.(2006) *A factor analysis model for functional genomics* BMC Bioinformatics, 7

[8] Langaas M., Lindqvist B. H., and Ferkingstad, E. (2005) *Estimating the proportion of true null hypotheses, with application to DNA microarray data* JRSS B, 67:555-572

[9] Schweder T. and Spjøtvoll E.(1982) *Plots of p-values to evaluate many tests simultaneously* Biometrika, 69:493-502

[10] Storey J. D.(2002) *A direct approach to false discovery rates* JRSS B, 64:479-498

[11] Storey J. D. and Tibsirani R. (2003) *Statistical significance for genomewide studies* PNAS B, 100:9440-9445

Index

Ackermann D., 960
Alexeyeva N., 953
Alfares H.K., 915
Allison J., 749
Amo-Salas M., 595
Andersons J., 903
Andronov A., 1016
Artalejo J., 779
Atencia I., 797
Atkinson A.C., 589, 1054

Bochkina N., 1033
Bogacka B., 589
Boukouvalas A., 839
Broniatowsky M., 726
Bruneel H., 827
Brunner E., 605
Burnaeva E., 969

Cammarota V., 991
Cancela H., 703
Caroni C., 651
Carriere KC., 1072
Causeur D., 628, 1137
Cornford C., 839

Darkhovsky B., 611
De Clercq S., 827
Delgado R., 785
Dette H., 1085
Dey D., 623
DuClos C., 693
Dudin A., 815
Dyudenko I., 1003

Ebner B., 737
Economou A., 804
El Khadiri M., 703
Eom H., 815, 833
Ermakov M., 1034
Ermakov S.M., 922, 929, 1022

Fattakhova M., 833
Friguet C., 628, 1137

Garel B., 634
Ghosh S., 623
Gotway C.A., 693
Gouet R., 985
Gozzi G., 1124
Grané A., 995
Granichin O., 1130
Grossi L., 1124
Guan R., 640
Guo L., 640
Gurevich L., 1130
Gurtov A., 851
Gut A., 667

Harlamov B., 935
Harman R., 1097
Henze N., 737
Heussen N., 960
Hilgers R.D., 960, 1103
Hollander M., 687
Horton G., 709
Hu J., 656
Huber C., 649
Huskova M., 673, 731

Ianovsky E., 1117
Iskedjian M, 897

Jørgensen B., 1010

Kearney G., 693
Khokhulina V.A., 863
Khramova V., 815
Kim C.S., 815, 821, 833
Kleinhov M., 903
Klimenok V., 815, 821
Kodia B., 634
Kolnogorov A.V., 947
Korchevsky V.M., 977
Korobeynikov A., 1027
Koskinen J., 845
Kreimer J., 1117
Krivulin N., 875
Krull C., 709

L'Ecuyer P., 885
López-Fidalgo J., 595
López-Ríos V., 595
López F.J., 985
Lagnoux A., 721
Laurini F., 1124
Lee M-L T., 650
Lewin A., 1033
Liang Y., 1072
Lopez-Herrero M.J., 773
Lukyanenko A., 851

Maksakova S., 891
Malyutov M., 1041
Mandjes M., 909
Martynov G., 765
Marusiakova M., 673
McGee D., 687
Meintanis S.G., 731, 743
Melas V.B., 1060, 1085
Melikov A., 833
Mielke T., 1066
Miretskiy D., 909
Morozov E., 851, 1003
Mosyagina E.N., 857

Nechaeva M.L., 1022
Nekrutkin V., 941
Nevzorov V., 981
Nikitin Y., 737
Nikulin M., 657
Nobel R., 803

Olkin I., 699
Orsingher E., 991

Pagano M., 1003
Paramonov Y., 903
Patan M., 589
Pechinkin A.V., 797
Pepelyshev A., 1091
Petersen H.C., 1010
Petrov V.V., 977

Ridder A., 791
Roth K., 1048
Rozovsky L., 975
Rubino G., 703
Rukavishnikova A.I., 929
Rumyantzev N., 941
Ryabko B., 617

Saaidia N., 657
Sadaka H., 1041
Saghatelyan V., 981
Sandmann W., 715, 1003
Sanz G., 985
Scheinhardt W., 909
Schiffl K., 1103
Sedunov E.V., 1079
Sedunova A.N., 1079
Shelonina T.N., 947
Shevchenko A.S., 869
Shpilev P., 1085
Simino J., 687
Singer A., 839
Střelec L., 755
Stehlík M., 755
Steinebach J.G., 667
Steland A., 679
Steyaert B., 827
Strelkov S., 891
Stulajter F., 1097
Sushkevich T., 891
Swanepoel J., 749

Taramin O., 821
Taufer E., 743
Tchirkov M.K., 863, 869
Tikhomirov A., 1111
Timofeev K.A., 922
Tommasi C., 583
Tuffin B., 703, 885

Vakhitov A., 1130
Veiga H., 995
Volkova K., 761

Walker J., 897
Whitmore G.A., 650

Young L.J., 693