Univerzita Komenského v Bratislave
Fakulta matematiky, fyziky a informatiky

Metóda kvázi-maximálnej vierohodnosti v zovšeobecnených regresných modeloch

Diplomová práca

BRATISLAVA 2010

Bc. Jakub Wimmer

Univerzita Komenského v Bratislave

Fakulta matematiky, fyziky a informatiky

Metóda kvázi-maximálnej vierohodnosti v zovšeobecnených regresných modeloch

Diplomová práca

Ekonomická a finančná matematika

9.1.9 Aplikovaná matematika

Katedra aplikovanej matematiky a štatistiky

Vedúci práce: Doc. RNDr. Viktor Witkovský, CSc.

BRATISLAVA 2010

Bc. Jakub Wimmer

Poďakovanie

Touto cestou sa chcem poďakovať Doc. RNDr. Viktorovi Witkovskému, CSc. za odborné vedenie diplomovej práce, cenné pripomienky, ako aj za čas strávený pri konzultáciách.

Čestné prehlásenie

Prehlasujem, že som prácu vypracoval samostatne s použitím uvedenej literatúry.

V Bratislave 23. apríla 2010

podpis študenta

Abstrakt

Práca pojednáva o metóde kvázi-maximálnej vierohodnosti, ktorá vychádza z teórie zovšobecnených regresných modelov. Tieto nachádzajú uplatnenie aj v mnohých oblastiach ekonómie a poisťovníctva, predovšetkým v neživotnom poistení. Výhoda oproti štandardnej metóde maximálnej vierohodnosti je, že pri odhadovaní neznámych parametrov nie je nutné poznať rozdelenie pravdepodobnosti pozorovaných náhodných premenných (meraní). Potrebná je len znalosť vzťahu medzi strednou hodnotou a varianciou meraní. V takomto prípade môžeme použiť metódu kvázi-maximálnej vierohodnosti na odhad neznámych parametrov. Práca opisuje niektoré základné vlastnosti odhadu získaného metódou kvázi-maximálnej vierohodnosti a na simulovaných dátach vyšetruje jeho správanie sa s rastúcim počtom meraní.

Kľúčové slová: exponenciálna trieda rozdelení, zovšobecnené lineárne modely, kvázi-vierohodná funkcia.

Abstract

The topic of the diploma thesis is the quasi-maximum likelihood method of estimation. This method is based on generalized linear models, which are used in many areas of economics and insurance, especially in non-life insurance. An advantage of this method in comparison with the usual maximum likelihood method is no necessity of exact knowing the distribution of the observed random variables (measurements). For using the quasi-maximum likelihood method for estimation it is necessary only to know the relation between the mean and the dispersion of the measurement. In the thesis are described some properties of the quasi-maximum likelihood estimator and examined is its behavior with increasing number of measurement.

Keywords: exponential class of distribution, generalized linear models,
quasi-likelihood function.

Obsah

Kapitola 1

Úvod

S využitím modelov a metód regresnej analýzy pri vyhodnocovaní údajov sa nezriedka stretávame aj v rôznych oblastiach ekonómie. Napr. (pozri [5]) pri dopravných projektoch, kde sa pomocou nich modelujú a analyzujú priebežné náklady uvedeného projektu. Rozšírenejšie uplatnenie však majú v oblasti neživotného (ale aj životného) poistenia, kde sa takýmto spôsobom môže modelovať (pozri [1]) napr. závislosť motoristických poistných udalostí od veku vodiča či modelu auta.

Pri odhadovaní neznámych regresných parametrov regresných modelov metódou maximálnej vierohodnosti je potrebná konkrétna znalosť rozdelenia výstupnej (pozorovanej) náhodnej premennej. V praxi sa však mnohokrát stretávame s prípadmi, v ktorých rozdelenie výstupnej náhodnej premennej nepoznáme, respektíve sú známe len určité základné charakteristiky daného rozdelenia. Úplná znalosť rozdelenia je skôr výnimočná. Metóda kvázi-maximálnej vierohodnosti pre odhadovanie neznámeho regresného parametra, ktorá je hlavným obsahom tejto práce, zoslabuje predpoklady požadované v metóde maximálnej vierohodnosti na odhadovanie neznámeho regresného parametra ohľadne rozdelenia výstupnej náhodnej premennej, konkrétne nevyžaduje presnú znalosť rozdelenia výstupnej náhodnej premennej. Jediný predpoklad pre odhad neznámych parametrov je nezávislosť jednotlivých pozorovaní a to, že variancia výstupnej náhodnej premennej je známou funkciou jej strednej hodnoty. Toto je dôvod, prečo má metóda kvázi-maximálnej vierohodnosti veľký význam pri odhadovaní neznámych parametrov.

Cieľom predloženej diplomovej práce je štúdium zovšeobecnených lineárnych (regresných) modelov s možnostami využitia metódy kvázi-maximálnej vierohodnosti pri odhadovaní neznámych parametrov modelu. Na základe simulačnej štúdie budú preskúmané základné štatistické vlastnosti týchto odhadov pre malé rozsahy vektora pozorovaní. Na overenie vhodnosti a presnosti danej metódy bol ako súčasť diplomovej práce vytvorený program v programovacom jazyku *MATLAB*, slúžiaci k danému účelu.

V druhej kapitole, v časti 2.1, charakterizujeme exponenciálnu triedu rozdelení. Táto trieda tvorí základ zovšeobecnených lineárnych modelov, opísaných v časti 2.3. Ich dôležitou súčasťou je aj spájajúca funkcia definovaná v časti 2.2.

V tretej kapitole, v časti 3.1, sa venujeme metóde maximálnej vierohodnosti za predpokladu, že sú nám známe všetky informácie o danom rozdelení. V časti 3.2 sa zaoberáme odvodením odhadovacích rovníc pre celú exponenciálnu triedu.

V štvrtej kapitole, v časti 4.1, je definovaná funkcia kvázi-vierohodnosti a jej vlastnosti. Súčasťou tejto kapitoly, v časti 4.2, je určenie odhadovacích rovníc pre kvázi-maximálne vierohodný odhad neznámych parametrov modelu, ako aj opísanie asymptotických vlastností kvázi-maximálne vierohodného odhadu a testovanie hypotéz založených na týchto vlastnostiach v častiach 4.3 a 4.4.

V piatej kapitole, v časti 5.1, sa venujeme numerickým metódam k získaniu kvázi-maximálne vierohodného odhadu regresného parametra a to Newtonovmu-Raphsonovmu algoritmu a jeho modifikácii pomocou Fisherovje informačnej matice v časti 5.2.

V šiestej kapitole sú v simulačnej štúdii porovnané základné štatistické vlastnosti kvázi-maximálne vierohodného odhadu a adekvátnosť jeho použitia pre rôzne veľkosti počtu pozorovaní.

Kapitola 2

Zovšobecnené lineárne modely

2.1 Exponenciálna trieda rozdelení

Zovšeobecnené lineárne modely vychádzajú z exponenciálnej triedy rozdelení. Táto trieda rozdelení je charakterizovaná špeciálnym tvarom distribučnej funkcie. Hustota pravdepodobnostného rozdelenia, ktoré patrí do tejto triedy, sa dá zapísať v tvare

$$f(y;\theta,\phi) = e^{\frac{y\theta - b(\theta)}{a(\phi)}} e^{-c(y,\phi)}, \tag{2.1}$$

kde θ je parameter polohy, $a(\phi)$ je parameter škály a $c(y,\phi)$ je tzv. normalizačný faktor, funkčne nezávislý od parametra θ a slúžiaci nato, aby integrál hustoty bol rovný 1.

Združenú hustotu n nezávislých rovnako rozdelených (patriacich do exponenciálnej triedy) pozorovaní $Y_1, Y_2, \ldots, Y_n$ môžeme potom zapísať, ako

$$f(y_1,\ldots,y_n;\theta,\phi) = \prod_{i=1}^{n} f(y_i;\theta,\phi). \tag{2.2}$$

V prípade diskrétneho rozdelenia zodpovedá hustote pravdepodobnostná funkcia a (2.1) je v tvare

$$f(y_k;\theta,\phi) = e^{\frac{y_k\theta - b(\theta)}{a(\phi)}} e^{-c(y_k,\phi)}, \quad k = 1,2,\ldots\,. \tag{2.3}$$

Do exponenciálnej triedy rozdelení patrí aj Poissonove rozdelenie. Pre ukážku prepíšeme jeho pravdepodobnostnú funkciu do tvaru (2.3). Pravdepodobnostná funkcia Poissonového rozdelenia má tvar

$$f(y_k;\mu) = \frac{e^{\mu}\mu^{y_k}}{y_k!}, \tag{2.4}$$

kde $y_k = (k-1)$, $k = 1, 2, \dots$.

Stredná hodnota náhodnej veličiny Y, ktorá má Poissonovo rozdelenie (2.4) je rovná

$$E(Y) = \mu \in (0, \infty) \tag{2.5}$$

a jej variancia

$$var(Y) = \mu. \tag{2.6}$$

Krátkymi úpravami dostaneme

$$f(y_k;\mu) = e^{y_k \ln(\mu) - \mu} e^{-\ln(y_k!)} = \tag{2.7}$$

$$= e^{\frac{y_k \ln(\mu) - \mu}{1}} e^{-\ln\Gamma(y_k+1)}, \tag{2.8}$$

pričom $\Gamma(y)$ je gamma funkcia.

Poissonovo rozdelenie teda patrí do triedy exponenciálnych rozdelení s parametrami

$$\theta = \ln \mu \tag{2.9}$$

$$b(\theta) = \mu \tag{2.10}$$

$$a(\phi) = 1 \tag{2.11}$$

$$c(y_k;\phi) = \ln\Gamma(y_k+1). \tag{2.12}$$

Ďalšie rozdelenie patriace do tejto triedy je Bernoulliho rozdelenie s pravdepodobnostnou funkciou

$$f(y_k;\mu) = \mu^{y_k}(1-\mu)^{1-y_k}, \tag{2.13}$$

kde $y_k = (k-1)$, $k = 1, 2$.

Stredná hodnota náhodnej veličiny Y s Bernoulliho rozdelením pravdepodobnosti je rovná

$$E(Y) = \mu \in (0, 1) \tag{2.14}$$

a variancia

$$var(Y) = \mu(1-\mu) \in (0, 1). \tag{2.15}$$

Opäť, jednoduchými úpravami dostávame pravdepodobnostnú funkciu Bernoulliho rozdelenia v tvare (2.3)

$$f(y_k;\mu) = e^{y_k \ln \frac{\mu}{1-\mu} - (-\ln(1-\mu))}, \tag{2.16}$$

kde

$$\begin{aligned} \theta &= \ln \frac{\mu}{1-\mu}, \\ b(\theta) &= -\ln(1-\mu), \\ a(\phi) &= 1, \\ c(y_k, \phi) &= 0. \end{aligned} \tag{2.17}$$

Do exponenciálnej triedy rozdelení patria taktiež aj iné, často používané rozdelenia, ako sú napríklad:

1. binomické rozdelenie s pravdepodobnostnou funkciou

$$f(y;n,p) = \binom{n}{y} p^y (1-p)^{n-y}, \quad y = 0,1,2,\ldots,n, \quad p \in (0,1), \tag{2.18}$$

so strednou hodnotou

$$E(Y) = np \tag{2.19}$$

a varianciou

$$var(Y) = np(1-p); \tag{2.20}$$

2. gamma rozdelenie s hustotou

$$f(y;k,\phi) = y^{k-1} \frac{e^{-\frac{y}{\phi}}}{\phi^k \Gamma(k)}, \quad y,k,\phi > 0, \tag{2.21}$$

so strednou hodnotou

$$E(Y) = k\phi \tag{2.22}$$

a varianciou

$$var(Y) = k\phi^2; \tag{2.23}$$

3. inverzné Gaussove rozdelenie s hustotou

$$f(y;\mu,\lambda) = \left(\frac{\lambda}{2\pi y^3}\right)^{\frac{1}{2}} e^{\frac{-\lambda(y-\mu)^2}{2\mu^2 y}}, \quad y,\mu,\lambda > 0, \tag{2.24}$$

so strednou hodnotou

$$E(Y) = \mu \tag{2.25}$$

a varianciou

$$var(Y) = \frac{\mu^3}{\lambda}; \tag{2.26}$$

4. geometrické rozdelenie s pravdepodobnostnou funkciou

$$f(y;p) = (1-p)^y p, \quad y = 0,1,2,\ldots, \quad p \in (0,1), \tag{2.27}$$

so strednou hodnotou

$$E(Y) = \frac{1-p}{p} \tag{2.28}$$

a varianciou

$$var(Y) = \frac{1-p}{p^2}; \tag{2.29}$$

5. negatívne binomické rozdelenie s pravdepodobnostnou funkciou

$$f(y;p,r) = \binom{y+r-1}{r-1}(1-p)^r p^y, \quad y = 0,1,2,\ldots, \quad r > 0, p \in (0,1), \tag{2.30}$$

so strednou hodnotou

$$E(Y) = r\frac{p}{1-p} \tag{2.31}$$

a varianciou

$$var(Y) = r\frac{p}{(1-p)^2}; \tag{2.32}$$

6. normálne rozdelenie s hustotou

$$f(y;\mu,\sigma^2) = \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(y-\mu)^2}{2\sigma^2}} \quad y \in R, \mu \in R, \sigma^2 > 0 \tag{2.33}$$

so strednou hodnotou

$$E(Y) = \mu \tag{2.34}$$

a varianciou

$$var(Y) = \sigma^2. \tag{2.35}$$

2.2 Spájajúca funkcia

Veľmi dôležitá súčasť zovšeobecnených lineárnych modelov je aj spájajúca funkcia (v anglickej literatúre označovaná, ako link function) $g(\cdot)$, ktorá je vo všeobecnosti nejakou vhodnou diferencovateľnou a invertovateľnou funkciou strednej hodnoty náhodnej veličiny Y, pričom platí

$$g(E(Y)) = g(\mu) = \eta. \tag{2.36}$$

Táto funkcia zabezpečuje zahrnutie informácie z tzv. kovariátov (vysvetľujúcich premenných) do modelu a je vyjadrená, ako lineárna kombinácia týchto kovariátov a neznámeho vektora parametrov modelu $\boldsymbol{\beta}$ v tvare

$$\eta = \mathbf{x}'\boldsymbol{\beta}. \tag{2.37}$$

Tu $\mathbf{x} = (x_1, x_2, \dots, x_p)'$ je známy vektor kovariátov. V prípade n pozorovaní (meraní), teda náhodných veličín $Y_1, Y_2, \dots, Y_n$ dostávame n rozmerný vektor

$$\boldsymbol{\eta} = (\eta_1, \eta_2, \dots, \eta_n)' = \begin{pmatrix} g(E(Y_1)) \\ g(E(Y_2)) \\ \vdots \\ g(E(Y_n)) \end{pmatrix} = \mathbf{X}\boldsymbol{\beta}, \tag{2.38}$$

kde známa $n \times p$ matica $\mathbf{X}$ je tzv. matica plánu a $\mathbf{x_i}' = (x_{i1}, x_{i2}, \dots, x_{ip})$ je i-ty riadok matice $\mathbf{X}$.

Každé zo spomenutých rozdelení má tzv. kanonickú spájajúcu funkciu. Táto opisuje "priamy" vzťah medzi η a strednou hodnotou μ. Kanonická spájajúca funkcia sa dá jednoducho odvodiť pri prepise hustoty jednotlivých rozdelení do všeobecného tvaru v exponenciálnej triede rozdelení. Jej tvar je potom nasledovný

$$\theta = g(\mu) = \eta \tag{2.39}$$

Vidíme, že napr. pri Bernoulliho rozdelení je

$$\theta = \ln \frac{\mu}{1 - \mu} \tag{2.40}$$

a teda kanonický tvar spájajúcej funkcie je v tomto prípade

$$g(\mu) = \ln \frac{\mu}{1-\mu} = \eta. \tag{2.41}$$

Táto funkcia sa nazýva aj logitová (logit function) spájajúca funkcia. V prípade Poissonovho rozdelenia je tvar kanonickej spájajúcej funkcie (tzv. log link) nasledovný

$$g(\mu) = \ln \mu = \eta. \tag{2.42}$$

Spájajúca funkcia je teda nejakou jednoduchou funkciou strednej hodnoty. V ďalšej analýze však zväčša budeme využívať inverziu spájajúcej funkcie v tvare

$$\mu = g^{-1}(\eta). \tag{2.43}$$

2.3 Zovšeobecnené lineárne modely - GLM

Zovšeobecnené lineárne modely, v anglickej literatúre nazývané generalized linear models (GLM), prvý krát predstavené Nelderom a Wedderburnom v roku 1972 v článku [10]. Boli zavedené s cieľom zjednotiť rôzne regresné modely do jednotného celku. V skutočnosti predstavujú zovšeobecnenie klasických lineárnych modelov, v ktorých sa parametre odhadujú metódou maximálnej vierohodnosti, teda ak predpokladáme, že vektor nameraných hodnôt $\mathbf{y} = (y_1, y_2, \ldots, y_n)'$ je realizácia náhodného vektora meraní (pozorovaní) $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_n)'$, ktorého jednotlivé zložky sú navzájom nezávislé, normálne rozdelené so strednou hodnotou $E(\mathbf{Y}) = \boldsymbol{\mu} = (\mu_1, \mu_2, \ldots, \mu_n)'$ a konštantnou disperziou σ^2. Ďalej sa predpokladá, že vektor stredných hodnôt $\boldsymbol{\mu}$ sa dá vyjadriť ako lineárna kombinácia menšieho počtu neznámych parametrov $\beta_1, \beta_2, \ldots, \beta_p$, kde $p < n$, a síce

$$\boldsymbol{\mu} = \mathbf{X}\boldsymbol{\beta},$$

kde $\mathbf{X}$ je známa $(n \times p)$-rozmerná matica a $\boldsymbol{\beta} = (\beta_1, \beta_2, \ldots, \beta_p)'$. Práve upustenie od predpokladu normality výstupnej premennej $\mathbf{Y}$ spolu so skutočnosťou, že lineárna kombinácia parametrov $\beta_1, \beta_2, \ldots, \beta_p$ nemusí byť rovná strednej hodnote $\boldsymbol{\mu}$, ale len nejakej známej funkcii tejto strednej hodnoty, konkrétne

$$\mathbf{x}_i'\boldsymbol{\beta} = g(\mu_i), \quad \mathrm{i} = 1,\, 2,\, \ldots,\, \mathrm{n}$$

kde $\mathbf{x}_i'$ je i-ty riadok matice $\mathbf{X}$, charakterizuje rozšírenie klasických lineárnych modelov na celú triedu zovšeobecnených lineárnych modelov. Zovšeobecnené lineárne modely sú potom určené nasledujúcimi vlastnosťami:

1. jednotlivé pozorovania $y_1, y_2, \ldots, y_n$ sú realizáciami navzájom nezávislých náhodných premenných $Y_1, Y_2, \ldots, Y_n$, kde každé Y_i, $i = 1, 2, \ldots, n$, pochádza z exponenciálnej triedy rozdelení definovanej v časti 2.1

2. vzťah medzi lineárnou kombináciou parametrov $\beta_1, \beta_2, \ldots, \beta_p$ a strednou hodnotou $E(Y_i)$, $i = 1, 2, \ldots, n$ je určený známou spájajúcou funkciou definovanou v časti 2.2.

Takéto zjednotenie mnohých lineárnych modelov pod jeden celok nám následne umožňuje vytvoriť spoločný postup pri analýze údajov pochádzajúcich z rôznych rozdelení spĺňajúcich vyššie uvedené predpoklady. Tento postup má mnohé výhody najmä pri tvorbe algoritmov na odhadovanie neznámej hodnoty parametrov $\beta_1, \beta_2, \ldots, \beta_p$, ale taktiež pri štatistickom odvodzovaní záverov pre takto získané odhady.

Kapitola 3

Odhad parametrov GLM pomocou funkcie vierohodnosti

3.1 FIML odhadovacie rovnice pre zovšeobecnené lineárne modely

Na úvod uvedieme príklad odhadu neznámeho parametra v prípade, že máme úplnú informáciu o výstupnej premennej. Tento prístup sa nazýva aj FIML (full information maximum likelihood). V takomto prípade má metóda maximálnej vierohodnosti nasledujúce kroky:

1. *zvoliť rozdelenie výstupnej premennej - meranej (pozorovanej) náhodnej veličiny* Y*;*
2. *napísať združené rozdelenie pre množinu navzájom nezávislých náhodných veličín* $Y_1, Y_2, \ldots, Y_n$*;*
3. *určiť funkciu vierohodnosti;*
4. *parametrizovať funkciu vierohodnosti v zmysle lineárnej kombinácie kovariátov* $\mathbf{X}$ *a s nimi združenými koeficientmi neznámeho parametra* $\boldsymbol{\beta}$*;*

5. *určiť rozsah reštrikcií pre koeficienty neznámeho parametra* $\boldsymbol{\beta}$ *v strednej hodnote a variancii pre zvolené rozdelenie;*

6. *napísať odhadovacie rovnice pre získanie odhadov neznámych parametrov.*

Bohužiaľ, odhadovacie rovnice sú vo väčšine prípadov ťažko analyticky riešiteľné, a preto sa k získaniu realizácie odhadov neznámych parametrov používajú rôzne numerické metódy (napr. Newtonova-Raphsonova, alebo tzv. Fischerova-skórovacia (Fischer scoring)).

Pre výstupnú premennú, ktorá spĺňa Bernoulliho rozdelenie vyzerajú uvedené kroky nasledovne

1. zvoliť rozdelenie výstupnej premennej -

Y má Bernoulliho rozdelenie, ak nadobúda hodnoty $y_1 = 0$ a $y_2 = 1$ s pravdepodobnosťami

$$P\{Y = y_1\} = P\{Y = 0\} = p$$
$$P\{Y = y_2\} = P\{Y = 1\} = 1 - p. \tag{3.1}$$

Y má teda pravdepodobnostnú funkciu

$$f(y_k; p) = p^{y_k}(1-p)^{1-y_k}, \quad y_k = k-1, k = 1, 2, \tag{3.2}$$

kde $p \in (0, 1)$ a

$$E(Y) = p = \mu$$
$$var(Y) = p(1 - p). \tag{3.3}$$

2. napísať združené rozdelenie pre množinu navzájom nezávislých meraní, pozorovaní, teda náhodných veličín $Y_1, Y_2, \dots, Y_n$ *-*

majme $Y_1, Y_2, \dots, Y_n$, kde Y_i má Bernoulliho rozdelenie pravdepodobnosti s parametrom p_i, ktoré sú nezávislé. Združená pravdepodobnostná funkcia je

$$f(y_1, y_2, \dots, y_n; p_1, p_2, \dots, p_n) = \prod_{i=1}^{n} e^{y_i ln(\frac{p_i}{1-p_i}) + ln(1-p_i)}. \tag{3.4}$$

kde $p_i \in (0,1) \quad i = 1, 2, \dots n$.

3. určiť funkciu vierohodnosti -

funkcia vierohodnosti je združená pravdepodobnostná funkcia (3.4), pričom $y_1, y_2, \dots, y_n$ považujeme za dané čísla a parametre (v tomto prípade $p_1, p_2, \dots, p_n$) za neznáme veličiny

$$L(p_1, p_2, \dots, p_n; y_1, y_2, \dots, y_n) = \prod_{i=1}^{n} e^{y_i ln(\frac{p_i}{1-p_i}) + ln(1-p_i)}. \tag{3.5}$$

V tomto prípade môžeme jednoducho nahradiť $\mu_i = p_i$ a dostávame

$$L(\mu_1, \mu_2, \dots, \mu_n; y_1, y_2, \dots, y_n) = \prod_{i=1}^{n} e^{y_i ln(\frac{\mu_i}{1-\mu_i}) + ln(1-\mu_i)}. \tag{3.6}$$

V praktických výpočtoch sa však častejšie využíva logaritmus funkcie vierohodnosti daný

$$l(\mu_1, \mu_2, \dots, \mu_n; y_1, y_2, \dots, y_n) = \sum_{i=1}^{n} \left[y_i ln(\frac{\mu_i}{1-\mu_i}) + ln(1-\mu_i) \right]. \tag{3.7}$$

4. parametrizovať funkciu vierohodnosti v zmysle lineárnej kombinácie kovariátov $\mathbf{X}$ *a s nimi združených koeficientov neznámeho parametra* $\boldsymbol{\beta}$ -

ako sme spomenuli v časti 2.2, kanonická spájajúca funkcia pre Bernoulliho rozdelenie je

$$g(\mu) = ln\frac{\mu}{1-\mu}.$$

Pre zjednodušenie výpočtov a pre ukážku odvodenia odhadovacích rovníc pomocou inej, ako kanonickej spájajúcej funkcie zavedieme tzv. identickú spájajúcu funkciu tvaru

$$g(\mu) = \mu. \tag{3.8}$$

Vo všobecnosti platí

$$g(\mu_i) = \eta_i = \mathbf{x_i'}\boldsymbol{\beta}, \quad i = 1, 2, \dots, n\,, \tag{3.9}$$

kde

$$\mathbf{x_i} = \begin{pmatrix} x_{i1} \\ x_{i2} \\ \vdots \\ x_{ip} \end{pmatrix} \tag{3.10}$$

je známy i-ty riadok matice $\mathbf{X}$ a

$$\boldsymbol{\beta} = (\beta_1, \beta_2, \ldots, \beta_p)' \tag{3.11}$$

je neznámy vektor parametrov modelu. Po zavedení kovariátov a neznámeho vektora parametrov do modelu dostávame

$$l(\beta_1, \beta_2, \ldots, \beta_p; y_1, y_2, \ldots, y_n) = \sum_{i=1}^{n} \left[y_i ln(\mathbf{x_i'}\boldsymbol{\beta}) + (1 - y_i) ln(1 - \mathbf{x_i'}\boldsymbol{\beta}) \right]. \tag{3.12}$$

5. určiť rozsah reštrikcií pre koeficienty $\boldsymbol{\beta}$ v strednej hodnote a variancii pre zvolené rozdelenie -

pre spájajúcu funkciu (3.9) však nesmieme zabudnúť na ďalšiu podmienku, kladenú na odhadovaný parameter $\boldsymbol{\beta}$, a to, že

$$\mathbf{x_i'}\boldsymbol{\beta} \in (0, 1) \qquad i = 1, 2, \ldots, n. \tag{3.13}$$

6. napísať rovnice pre získanie odhadov neznámych parametrov -

deriváciou l podľa jednotlivých β_k, $k = 1, 2, \ldots, p$ a ich položením rovným 0. Riešením týchto rovníc je realizácia maximálne vierohodného odhadu parametra $\boldsymbol{\beta}$, ak riešenie padne do parametrického priestoru určeného v (3.13). Počítajme

$$\frac{\partial l}{\partial \beta_1} = \frac{\partial}{\partial \beta_1} y_1 ln(x_{11}\beta_1 + x_{12}\beta_2 \ldots + x_{1p}\beta_p) + (1 - y_1) ln(1 - x_{11}\beta_1 - x_{12}\beta_2 \ldots - x_{1p}\beta_p) +$$

$$+ y_2 ln(x_{21}\beta_1 + x_{22}\beta_2 \ldots + x_{2p}\beta_p) + (1 - y_2) ln(1 - x_{21}\beta_1 - x_{22}\beta_2 \ldots - x_{2p}\beta_p) +$$

$$\vdots$$

$$+ y_n ln(x_{n1}\beta_1 + x_{n2}\beta_2 \ldots + x_{np}\beta_p) + (1 - y_n) ln(1 - x_{n1}\beta_1 - x_{n2}\beta_2 \ldots - x_{np}\beta_p) \tag{3.14}$$

$$= y_1 \frac{x_{11}}{\mathbf{x_1}'\boldsymbol{\beta}} + (1 - y_1) \frac{-x_{11}}{1 - \mathbf{x_1}'\boldsymbol{\beta}} +$$

$$+ y_2 \frac{x_{21}}{\mathbf{x_2}'\boldsymbol{\beta}} + (1 - y_2)\frac{-x_{21}}{1 - \mathbf{x_1}'\boldsymbol{\beta}} +$$
$$\vdots$$
$$+ y_n \frac{x_{n1}}{\mathbf{x_n}'\beta} + (1 - y_n)\frac{-x_{n1}}{1 - \mathbf{x_n}'\boldsymbol{\beta}} \tag{3.15}$$
$$= \sum_{i=1}^{n} \left(\frac{y_i}{\mathbf{x_i}'\boldsymbol{\beta}} - \frac{1 - y_i}{1 - \mathbf{x_i}'\boldsymbol{\beta}} \right) x_{i1} = 0. \tag{3.16}$$

Rovnako potom

$$\frac{\partial l}{\partial \beta_j} = \sum_{i=1}^{n} \left(\frac{y_i}{\mu_i} - \frac{1 - y_i}{1 - \mu_i} \right) x_{ij} = 0 \qquad j = 1, 2, \ldots, p. \tag{3.17}$$

Rovnice (3.17) sú odhadovacie rovnice pre parameter $\boldsymbol{\beta}$ v Bernoulliho rozdelení. Takto získaný odhad (ak patrí do patričného parametrického priestoru) sa nazýva maximálne vierohodný odhad (FIML metóda).

3.2 LIML odhadovacie rovnice pre zovšeobecnené lineárne modely

Ako sme spomínali, FIML metóda predpokladá úplnú znalosť rozdelenia výstupnej premennej, ktorú však, v mnohých prípadoch, nemáme k dispozícií. Preto je vhodné uvažovať metódu maximálnej vierohodnosti s obmedzenou informáciou o výstupnej premennej. Táto metóda nepredpokladá znalosť úplnej informácie o rozdelení observovanej náhodnej veličiny Y, teda jej hustoty, respektíve pravdepodobnostnej funkcie. V literatúre (napr. [3]) je označovaná ako limited information maximum likelihood (LIML) a predpokladá len, že rozdelenie pravdepodobnosti pozorovanej náhodnej veličiny Y patrí do exponenciálnej triedy rozdelení.

Exponenciálna trieda rozdelení má už spomínanú hustotu tvaru

$$f(y; \theta, \phi) = e^{\frac{y\theta - b(\theta)}{a(\phi)}} e^{-c(y,\phi)},$$

pričom (pozri napr. [3], [2], resp. [6])

$$E(Y) = \frac{\partial b(\theta)}{\partial \theta} = \mu \tag{3.18}$$

a

$$var(Y) = \frac{\partial^2 b(\theta)}{\partial \theta^2} a(\phi). \tag{3.19}$$

Pre lepšiu prehľadnosť zavádzame

$$\frac{\partial^2 b(\theta)}{\partial \theta^2} = V(\mu), \tag{3.20}$$

kde $V(\mu)$ je známa funkcia. Variancia náhodnej premennej Y je teda známou funkciou jej strednej hodnoty μ a parametra škály $a(\phi)$, ktorý považujeme za konštantný, pričom môže byť aj neznámy.

Združená hustota pre n nezávislých pozorovaní $Y_1, Y_2, \ldots, Y_n$ vyzerá nasledovne

$$f(y_1, y_2, \ldots, y_n; \theta_1, \theta_2, \ldots, \theta_n, \phi) = \prod_{i=1}^{n} e^{\frac{y_i\theta_i - b(\theta_i)}{a(\phi)} - c(y_i, \phi)}, \tag{3.21}$$

a funkcia vierohodnosti je

$$L(\theta_1, \theta_2, \ldots, \theta_n, \phi; y_1, y_2, \ldots, y_n) = \prod_{i=1}^{n} e^{\frac{y_i\theta_i - b(\theta_i)}{a(\phi)} - c(y_i, \phi)}. \tag{3.22}$$

Je to funkcia $\theta_1, \theta_2, \ldots, \theta_n$ a ϕ pri známych realizáciách $y_1, y_2, \ldots, y_n$ náhodných veličín $Y_1, Y_2, \ldots, Y_n$.

Zavedenie lineárnej kombinácia kovariátov a neznámeho vektora parametrov $\mathbf{x_i}'\boldsymbol{\beta}$, $i = (1, 2, \ldots, n)$ do tohto modelu tak, ako v predošlej časti by značne komplikovalo ďalšie výpočty, a preto ich zavedieme neskôr, priamo do odhadovacích rovníc.

Logaritmus funkcie vierohodnosti má teraz tvar

$$l(\theta_1, \theta_2, \ldots, \theta_n, \phi; y_1, \ldots, y_n) = \sum_{i=1}^{n} \left[\frac{y_i\theta_i - b(\theta_i)}{a(\phi)} + c(y_i, \phi) \right]. \tag{3.23}$$

Označme

$$\frac{\partial l}{\partial \theta_i} = \frac{y_i - b'(\theta_i)}{a(\phi)}. \tag{3.24}$$

Využitím (3.18), teda toho že $b'(\theta_i) = \mu_i$ dostávame

$$\frac{\partial l}{\partial \theta_i} = \frac{y_i - \mu_i}{a(\phi)}. \tag{3.25}$$

Po zavedení lineárnej kombinácie kovariátov a príslušného vektora neznámych parametrov $\boldsymbol{\beta}$ do modelu získame odhadovacie rovnice nasledovne

$$\frac{\partial l}{\partial \beta_r} = 0 \quad r = 1, 2, \ldots, p\,, \tag{3.26}$$

kde použitím "reťazového pravidla" pre derivovanie dostávame

$$\frac{\partial l}{\partial \beta_r} = \sum_{i=1}^{n} \left(\frac{\partial l}{\partial \theta_i}\right)\left(\frac{\partial \theta_i}{\partial \mu_i}\right)\left(\frac{\partial \mu_i}{\partial \eta_i}\right)\left(\frac{\partial \eta_i}{\partial \beta_r}\right), \quad r = 1, 2, \ldots, p\,. \tag{3.27}$$

Nahradením

$$\frac{\partial l}{\partial \theta_i} = \frac{y_i - \mu_i}{a(\phi)}, \tag{3.28}$$

$$\frac{\partial \theta_i}{\partial \mu_i} = \frac{\partial \theta_i}{\frac{\partial b(\theta_i)}{\partial \theta_i}} = \frac{\partial \theta_i^2}{\partial^2 b(\theta_i)} =$$

$$= \frac{1}{\frac{\partial^2 b(\theta_i)}{\partial \theta_i^2}} = \frac{1}{V(\mu_i)}, \tag{3.29}$$

$$\frac{\partial \eta_i}{\partial \beta_j} = x_{ij} \tag{3.30}$$

dostávame

$$\frac{\partial l}{\partial \beta_j} = \sum_{i=1}^{n} \frac{y_i - \mu_i}{a(\phi)V(\mu_i)}\left(\frac{\partial \mu_i}{\partial \eta_i}\right) x_{ij} \qquad\qquad j = 1, 2, \ldots, p. \tag{3.31}$$

Všeobecné LIML odahodvacie rovnice pre exponenciálnu triedu rozdelení teda vyzerajú nasledovne

$$\frac{\partial l}{\partial \beta_j} = \sum_{i=1}^{n} \frac{y_i - \mu_i}{a(\phi)V(\mu_i)}\left(\frac{\partial \mu_i}{\partial \eta_i}\right) x_{ij} = 0 \quad j = 1, 2, \ldots, p. \tag{3.32}$$

Získali sme všeobecné odhadovacie rovnice (3.32) pre celú exponenciálnu triedu rozdelení na odhad $\tilde{\boldsymbol{\beta}}$ neznámeho parametra $\boldsymbol{\beta}$. Stredná hodnota μ je parametrizovaná pomocou $\eta = \mathbf{x}'\boldsymbol{\beta}$. Ďalšou podmienkou je znalosť variancie odhadu, ktorá je známou funkciou strednej hodnoty. V nej sa objavuje aj parameter $a(\phi)$, ktorý nie je priamo potrebný na analýzu, a teda ho môžeme považovať za doplnkový (v GLM literatúre sa niekedy nazýva aj ako disperzia). Autori v [7] navrhujú odhadovať tento parameter pomocou momentovej metódy, ako

$$a(\tilde{\phi}) = \frac{1}{n-p} \sum_{i=1}^{n} \frac{(y_i - \tilde{\mu}_i)^2}{V(\tilde{\mu}_i)}, \tag{3.33}$$

kde n je počet pozorovaní, p je dimenzia neznámeho vektora parametrov $\boldsymbol{\beta}$ a $\tilde{\mu}_i = g^{-1}(\tilde{\eta}_i) = \mathbf{x_i'}\tilde{\boldsymbol{\beta}}$.

Okrem tejto variancie zovšeobecnené lineárne modely (GLM) charakterizuje už spomínaná spájajúca funkcia. Zvolením správnej spájajúcej funkcie, t.j. takej, ktorá zabezpečí správne obmedzenia na strednú hodnotu a varianciu, nám hľadanie optimálneho riešenia (správny odhad parametra $\boldsymbol{\beta}$) ,podľa [4], nebude robiť žiadne väčšie numerické problémy. Voľba inej, nie vhodnej spájajúcej funkcie, ktorá priamo nezabezpečí správne ohraničenia na strednú hodnotu a varianciu, nám do určitej miery skomplikuje odhadovanie neznámeho parametra $\boldsymbol{\beta}$. Naše riešenie môže v takomto prípade pochádzať z oblasti, kde je variancia záporná, respektíve nedefinovaná. Ak si napr. zvolíme Poissonov model, kde $E(y) = Var(y) = \mu$ sú kladné a zavedieme identickú spájajúcu funkciu

$$\mathbf{x_i'}\boldsymbol{\beta} = \eta_i = g(\mu_i) = \mu_i \qquad i = 1, 2 \ldots, n, \tag{3.34}$$

nebudeme mať zabezpečené správne ohraničenia (pretože $\mathbf{x_i'}\boldsymbol{\beta} \in (-\infty, \infty)$, $i = 1, 2 \ldots, n$). Avšak zavedenie ďalšieho ohraničenia (v tomto prípade $\mathbf{x_i'}\boldsymbol{\beta} > 0$, $i = 1, 2 \ldots, n$) nám zabezpečí správnosť nájdeného riešenia.

Kapitola 4

Odhad parametrov metódou kvázi-maximálnej vierohodnosti

4.1 Funkcia kvázi-vierohodnosti

Kým v metóde FIML sa predpokladá konkrétna znalosť rozdelenia, v metóde LIML je len predpoklad toho, že pozorovania pochádzajú z exponenciálnej triedy rozdelení. Toto sú vo všeobecnosti silné predpoklady, pričom ešte aj naša domnienka o rozdelení výstupnej premennej môže byť častokrát mylná. Preto autori [11] vo svojich úvahách upúšťajú od znalosti rozdelenia výstupnej premennej a na základe postupov uvedených v predošlých sekciách definujú tzv. funkciu kvázi-vierohodnosti (v anglickej literatúre označovaná ako quasi-likelihood function). Jedinými podmienkami sú

- predpoklad vzájomnej nezávislosti jednotlivých pozorovaní
- varianca je známou funkciou strednej hodnoty (až na konštantu).

Toto sú jednoznačne slabšie predpoklady, ako pri predošlých odvodzovaniach odhadovacích rovníc.

Majme teda nezávislé pozorovania $Y_1, Y_2, \ldots, Y_n$ s príslušnými strednými hodnotami μ_i, $i = 1, 2, \ldots, n$, a známou variančnou funkciou $V(\mu_i)$, kde $var(Y_i) = a(\phi)V(\mu_i)$.

Ďalej predpokladajme, že

$$\mathbf{X}'\boldsymbol{\beta} = \boldsymbol{\eta} = \begin{pmatrix} g(\mu_1) \\ g(\mu_2) \\ \vdots \\ g(\mu_n) \end{pmatrix}, \tag{4.1}$$

kde $g(\cdot)$ je známa spájajúca funkcia. Potom, pre každé pozorovanie definujeme funkciu kvázi-vierohodnosti $Q_i(y_i, \mu_i)$ vzťahom (pozri [10])

$$\frac{\partial Q_i(y_i, \mu_i)}{\partial \mu_i} = \frac{y_i - \mu_i}{V(\mu_i)a(\phi)}, \tag{4.2}$$

respektíve (pozri [3])

$$Q_i(y_i, \mu_i) = \int_{y_i}^{\mu_i} \frac{y_i - \mu^*}{V(\mu^*)a(\phi)} d\mu^* + f(y_i). \tag{4.3}$$

Združená funkcia kvázi-vierohodnosti pre n nezávislých pozorovaní $Y_1, Y_2, \ldots, Y_n$ je potom daná vzťahom

$$Q = Q_1(Y_1, \mu_1) + Q_2(Y_2, \mu_2) + \ldots + Q_n(Y_n, \mu_n) = \sum_{i=1}^{n} Q_i. \tag{4.4}$$

Zavedením kovariátov a príslušného vektora neznámych parametrov $\boldsymbol{\beta}$ do modelu získame funkciu kvázi-vierohodnosti pre parameter $\boldsymbol{\beta}$, ktorej derivácia podľa β_k je pre každé pozorovanie $i = 1, 2, \ldots, n$ daná vzťahom

$$\frac{\partial}{\partial \beta_k} Q_i = \frac{y_i - \mu_i}{V(\mu_i)a(\phi)} \frac{\partial \mu_i}{\partial \beta_k} \quad k = 1, 2, \ldots, p, \tag{4.5}$$

a teda derivácia združenej funkcie kvázi-vierohodnosti pre n nezávislých pozorovaní $Y_1, Y_2, \ldots, Y_n$ podľa parametra β_k je

$$\frac{\partial}{\partial \beta_k} \sum_{i=1}^{n} Q_i = \sum_{i=1}^{n} \frac{y_i - \mu_i}{V(\mu_i)a(\phi)} \frac{\partial \mu_i}{\partial \beta_k} \quad k = 1, 2, \ldots, p. \tag{4.6}$$

Teraz je zrejmé, že ak predpokladáme, že hustota, resp. pravdepodobnostná funkcia pozorovanej výstupnej premennej patrí do exponenciálnej triedy rozdelení, tak (4.6) prechádza na (3.31).

4.1.1 Vlastnosti funkcie kvázi-vierohodnosti

Nech $Q_i = Q_i(y_i, \mu_i)$ je daná v (4.2), resp. (4.3). Predpokladajme, že μ_i je vyjadrená ako funkcia parametrov $\beta_1, \beta_2, \ldots, \beta_p$. Potom Q_i má nasledujúce vlastnosti (viď [11])

$$(i) \qquad E\left(\frac{\partial Q_i}{\partial \mu_i}\right) = 0, \tag{4.7}$$

čo vyplýva priamo z definície funkcie kvázi-maximálnej vierohodnosti;

$$(ii) \qquad E\left(\frac{\partial Q_i}{\partial \beta_s}\right) = 0, \tag{4.8}$$

čo dostávame využitím "reťazového pravidla" pre derivovanie

$$\frac{\partial Q_i}{\partial \beta_s} = \left(\frac{\partial Q_i}{\partial \mu_i}\right)\left(\frac{\partial \mu_i}{\partial \beta_s}\right) = 0;$$

$$(iii) \qquad E\left(\frac{\partial Q_i}{\partial \beta_s}\frac{\partial Q_i}{\partial \beta_t}\right) = -E\left(\frac{\partial^2 Q_i}{\partial \beta_s \partial \beta_t}\right) = \frac{1}{a(\phi)V(\mu_i)}\frac{\partial \mu_i}{\partial \beta_s}\frac{\partial \mu_i}{\partial \beta_t}, \tag{4.9}$$

pretože platí

$$\begin{aligned} E\left(\frac{\partial Q_i}{\partial \beta_s}\frac{\partial Q_i}{\partial \beta_t}\right) &= E\left(\frac{\partial Q_i}{\partial \mu_i}\right)^2 \frac{\partial \mu_i}{\partial \beta_s}\frac{\partial \mu_i}{\partial \beta_t} = E\left[\frac{(y-\mu_i)^2}{(a(\phi)V(\mu_i))^2}\right]\frac{\partial \mu_i}{\partial \beta_s}\frac{\partial \mu_i}{\partial \beta_t} \\ &= \frac{1}{a(\phi)V(\mu_i)}\frac{\partial \mu_i}{\partial \beta_s}\frac{\partial \mu_i}{\partial \beta_t} \end{aligned}$$

a tiež

$$\begin{aligned} -E\left(\frac{\partial^2 Q_i}{\partial \beta_s \partial \beta_t}\right) &= -E\left[\frac{\partial}{\partial \beta_t}\left\{\frac{y-\mu_i}{a(\phi)V(\mu_i)}\frac{\partial \mu_i}{\partial \beta_s}\right\}\right] \\ = -\frac{1}{a(\phi)}E&\left[(y-\mu_i)\frac{\partial}{\partial \beta_t}\left\{\frac{1}{V(\mu_i)}\frac{\partial \mu_i}{\partial \beta_s}\right\} - \frac{1}{V(\mu_i)}\frac{\partial \mu_i}{\partial \beta_s}\frac{\partial \mu_i}{\partial \beta_t}\right] \\ &= \frac{1}{a(\phi)V(\mu_i)}\frac{\partial \mu_i}{\partial \beta_s}\frac{\partial \mu_i}{\partial \beta_t}, \end{aligned} \tag{4.10}$$

z čoho ľahko dostávame (4.9);

$$(iv) \qquad E\left(\frac{\partial Q_i}{\partial \mu_i}\right)^2 = -E\left(\frac{\partial^2 Q_i}{\partial \mu_i^2}\right) = \frac{1}{a(\phi)V(\mu_i)}, \tag{4.11}$$

lebo (4.11) je iba špeciálnym prípadom (4.9).

Poznamenajme, že všetky spomínané vlastnosti funkcie kvázi-vierohodnosti sú zhodné s vlastnosťami funkcie vierohodnosti v prípade známeho rozdelenia výstupnej premennej.

4.2 Kvázi-maximálne vierohodný odhad

Odhadovacie rovnice na získanie odhadu $\tilde{\boldsymbol{\beta}}$ neznámeho regresného parametra $\boldsymbol{\beta}$ metódou kvázi-maximálnej vierohodnosti dostaneme rovnakým postupom, ako pri metódach LIML a FIML, a teda zo združenej funkcie kvázi-vierohodnosti (4.6) pre n nezávislých pozorovaní $Y_1, Y_2, \dots, Y_n$, a to

$$\frac{\partial}{\partial \beta_k} \sum_{i=1}^{n} Q_i \Bigg|_{\boldsymbol{\beta}=\tilde{\boldsymbol{\beta}}} = 0 \qquad k = 1, 2, \dots, p, \tag{4.12}$$

čo je ekvivalentné s výrazom

$$\sum_{i=1}^{n} \frac{y_i - \mu_i}{V(\mu_i)a(\phi)} \frac{\partial \mu_i}{\partial \beta_k} \Bigg|_{\boldsymbol{\beta}=\tilde{\boldsymbol{\beta}}} = 0 \quad k = 1, 2, \dots, p. \tag{4.13}$$

Opäť, rovnako ako v kapitole 2, odhadovacie rovnice (4.13) sú v mnohých prípadoch analyticky ťažko spočítateľné, a preto sa k ich riešeniu využívajú numerické metódy, z ktorých dve (najčastejšie používané) sú popísané v kapitole 5.

Odhad $a(\tilde{\phi})$ môžeme dosiahnuť rovnakým spôsobom ako pri metóde FIML, respektíve LIML pomocou vzorca (3.33).

4.3 Asymptotická nevychýlenosť a asymptotická normalita kvázi-maximálne vierohodného odhadu

Pre odhady regresného parametra $\boldsymbol{\beta}$ získaného pomocou funkcie kvázi-maximálnej vierohodnosti platia rovnaké asymptotické vlastnosti ako v prípade odhadov získaných metódou maximálnej vierohodnosti. Sú to asymptotická nevychýlenosť a asymptotická normalita (pozri [9], alebo [6]). Teda

$$\tilde{\boldsymbol{\beta}} \sim N\left[\boldsymbol{\beta}, (\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}\right], \tag{4.14}$$

kde

$$\{W\}_{ii} = \frac{1}{V(\mu_i)a(\phi)} \left(\frac{\partial \mu_i}{\partial \eta_i}\right)^2 \qquad \{W\}_{ij} = 0 \quad i \neq j. \tag{4.15}$$

Dá sa ukázať (pozri [4]), že (s,t)-ty prvok asymptotickej kovariančnej matice kvázi-maximálne vierohodného odhadu parametra $\boldsymbol{\beta}$ je rovný

$$var(\tilde{\boldsymbol{\beta}})_{s,t} = (\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}_{st} = \left[E\left(\frac{\partial Q}{\partial\beta_s}\frac{\partial Q}{\partial\beta_t}\right)\right]^{-1} = \left[-E\left(\frac{\partial^2 Q}{\partial\beta_s\partial\beta_t}\right)\right]^{-1} =$$

$$= a(\phi)\left[\sum_{i=1}^{n}\left(\frac{\partial\mu_i}{\partial\eta_i}\right)^2_i \frac{1}{V(\mu_i)}x_{is}x_{it}\right]^{-1}. \tag{4.16}$$

Táto matica je zrejme funkciou parametra ϕ, ktorý je vo väčšine praktických úloh neznámy. Navyše v prípade, ak spájajúca funkcia $g(\cdot)$ nie je identická, je táto kovariančná matica aj funkciou neznámeho vektora parametrov $\boldsymbol{\beta}$. Preto je často potrebné odhadovať i túto kovariančnú maticu. V kapitole 5 sú uvedené niektoré štandardne používané odhady kovariančnej matice (4.16).

4.4 Testovanie hypotéz

Na základe spomenutých asymptotických vlastnostiach kvázi-maximálne vierohodného odhadu (4.14) môžeme teraz testovať hypotézy ohľadne regresného parametra $\boldsymbol{\beta}$, konkrétne hypotézy o lineárnej kombinácii $\mathbf{K}'\boldsymbol{\beta}$, kde $\mathbf{K}$ je známa $p \times k$ matica plnej hodnosti. K tomu nám slúži tzv. test waldovského typu (pozri napr. [8]), v ktorom na testovanie hypotézy

$$H_0 : \mathbf{K}'\boldsymbol{\beta} = \mathbf{m}, \text{ kde } \mathbf{m} \text{ je známy } k\text{-rozmerný vektor}$$

môžeme použiť testovaciu štatistiku v tvare

$$Wald = (\mathbf{K}'\tilde{\boldsymbol{\beta}} - \mathbf{m})'(\mathbf{K}'(\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}\mathbf{K})^{-1}(\mathbf{K}'\tilde{\boldsymbol{\beta}} - \mathbf{m}). \tag{4.17}$$

Táto štatistika má za platnosti nulovej hypotézy chi kvadrát rozdelenie s k stupňami voľnosti. Testovanú hypotézu H_0 potom zamietame na hladine významnosti α, ak

$$Wald \geq \chi^2_k(1-\alpha), \tag{4.18}$$

kde $\chi^2_k(1-\alpha)$ je $(1-\alpha)$ kvantil chi kvadrát rozdelenia s k stupňami voľnosti.

Žiaľ, ako bolo poznamenané už v predchádzajúcej kapitole, kovariančná matica $(\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}$ je často neznáma a preto sa vo vzťahu (4.17) nahrádza jej odhadom. Vhodnosťou použitia takejto aproximácie sa zaoberá kapitola 6.

Kapitola 5

Algoritmy na výpočet kvázi-maximálne vierohodného odhadu

5.1 Newtonova-Raphsonova metóda

Jedna z najužívanejších metód na získanie odhadu parametra $\boldsymbol{\beta}$ je Newtonova-Raphsonova metóda. Táto metóda nájdenia nulových bodov danej funkcie je založená na rozvinutí funkcie do Taylorovho radu. V našom prípade rozvíjame funkciu kvázi-maximálnej vierohodnosti. V ďalšom postupe budeme uvažovať p-rozmerný vektor

$$Q_{\boldsymbol{\beta}} = \frac{\partial Q}{\partial \boldsymbol{\beta}}, \quad Q_{\boldsymbol{\beta}} = \left(\frac{\partial Q}{\partial \beta_1}, \frac{\partial Q}{\partial \beta_2}, \dots, \frac{\partial Q}{\partial \beta_p} \right)' \tag{5.1}$$

a $(p \times p)$-rozmernú maticu

$$Q_{\boldsymbol{\beta}\boldsymbol{\beta}'} = \frac{\partial^2 Q}{\partial \boldsymbol{\beta} \partial \boldsymbol{\beta}'}, \quad \left\{ Q_{\boldsymbol{\beta}\boldsymbol{\beta}'} \right\}_{ij} = \frac{\partial^2 Q}{\partial \beta_i \partial \beta_j}. \tag{5.2}$$

Našim cieľom je riešenie rovnice

$$Q_{\boldsymbol{\beta}}(\tilde{\boldsymbol{\beta}}) = 0, \tag{5.3}$$

čo sú odhadovacie rovnice (4.13) pre odhad parametra $\boldsymbol{\beta}$. Rozvinutím (5.3) do Taylorovho radu okolo (vhodnej) štartovacej hodnoty $\beta^{(0)}$ dostávame

$$\mathbf{0} = Q_{\boldsymbol{\beta}}(\boldsymbol{\beta}^{(0)}) + (\boldsymbol{\beta} - \boldsymbol{\beta}^{(0)}) Q_{\boldsymbol{\beta}\boldsymbol{\beta}'}(\boldsymbol{\beta}^{(0)}) + \dots . \tag{5.4}$$

Zanedbaním členov druhého a vyšších rádov Taylorovho radu dostaneme rovnosť

$$\mathbf{0} \approx Q_{\boldsymbol{\beta}}(\boldsymbol{\beta}^{(0)}) + (\boldsymbol{\beta} - \boldsymbol{\beta}^{(0)})Q_{\boldsymbol{\beta}\boldsymbol{\beta}}(\boldsymbol{\beta}^{(0)}). \tag{5.5}$$

Riešením (5.5) pre danú štartovaciu hodnotu $\boldsymbol{\beta}^{(0)}$ dostávame

$$\boldsymbol{\beta} \approx \boldsymbol{\beta}^{(0)} - \left\{Q_{\boldsymbol{\beta}\boldsymbol{\beta}}(\boldsymbol{\beta}^{(0)})\right\}^{-1} Q_{\boldsymbol{\beta}}(\boldsymbol{\beta}^{(0)}). \tag{5.6}$$

V r-tej iterácii získame odhad vektora parametrov $\boldsymbol{\beta}$ (teda $\tilde{\boldsymbol{\beta}} = \boldsymbol{\beta}^{(r)}$) Newtonovou-Raphsonovou metódou, a síce

$$\boldsymbol{\beta}^{(r)} = \boldsymbol{\beta}^{(r-1)} - \left\{Q_{\boldsymbol{\beta}\boldsymbol{\beta}}(\boldsymbol{\beta}^{(r-1)})\right\}^{-1} Q_{\boldsymbol{\beta}}(\boldsymbol{\beta}^{(r-1)}) =$$

$$\boldsymbol{\beta}^{(r)} = \boldsymbol{\beta}^{\mathbf{(r-1)}} + \left\{\mathbf{S}(\boldsymbol{\beta}^{(r-1)})\right\}^{-1} Q_{\boldsymbol{\beta}}(\boldsymbol{\beta}^{(0)}). \tag{5.7}$$

$\mathbf{S}$ je daná vzťahom

$$\{\mathbf{S}\}_{s,t} = -\frac{\partial^2 Q}{\partial\beta_s \partial\beta_t}, \tag{5.8}$$

pričom Q je definované v (4.4).

Odhadom kovariančnej matice takto získaného odhadu $\tilde{\boldsymbol{\beta}}$ je podľa [3]

$$\widetilde{var(\tilde{\boldsymbol{\beta}})} = \left\{\mathbf{S}(\tilde{\boldsymbol{\beta}})\right\}^{-1}, \tag{5.9}$$

v ktorom $a(\phi)$ odhadujeme pomocou vzťahu (3.33).

Upravujme ešte ďalej

$$\frac{\partial^2 Q}{\partial\beta_s \partial\beta_t} = \sum_{i=1}^{n} \frac{1}{a(\phi)} \left(\frac{\partial}{\partial\beta_t}\right) \left\{\frac{y_i - \mu_i}{V(\mu_i)} \left(\frac{\partial\mu_i}{\partial\eta_i}\right) x_{is}\right\} =$$

$$= \sum_{i=1}^{n} \frac{1}{a(\phi)} \left[\frac{-\frac{\partial\mu_i}{\partial\beta_t} V(\mu_i) - (y_i - \mu_i)\frac{\partial V(\mu_i)}{\partial\beta_t}}{V(\mu_i)^2} \frac{\partial\mu_i}{\partial\eta_i} + \frac{y_i - \mu_i}{V(\mu_i)} \frac{\partial}{\partial\beta_t} \left(\frac{\partial\mu_i}{\partial\eta_i}\right)\right] x_{is} =$$

$$= \sum_{i=1}^{n} \frac{1}{a(\phi)} \left[-\frac{1}{V(\mu_i)} \frac{\partial\mu_i}{\partial\beta_t} \frac{\partial\mu_i}{\partial\eta_i} - \right.$$

$$\left. -(y_i - \mu_i) \left\{\frac{1}{V(\mu_i)^2} \frac{\partial V(\mu_i)}{\partial\beta_t} \frac{\partial\mu_i}{\partial\eta_i} - \frac{1}{V(\mu_i)} \frac{\partial}{\partial\beta_t} \left(\frac{\partial\mu_i}{\partial\beta_t}\right)\right\}\right] x_{is} =$$

$$= -\sum_{i=1}^{n} \frac{1}{a(\phi)} \left[\frac{1}{V(\mu_i)} \left(\frac{\partial\mu_i}{\partial\eta_i}\right)^2 - \right.$$

$$-(\mu_i - y_i)\left\{\frac{1}{V(\mu_i)^2}\left(\frac{\partial\mu_i}{\partial\eta_i}\right)^2\frac{\partial V(\mu_i)}{\partial\mu_i} - \frac{1}{V(\mu_i)}\left(\frac{\partial^2\mu_i}{\partial\eta_i^2}\right)\right\}\Bigg] x_{is}x_{it}. \qquad (5.10)$$

Najťažším krokom v (5.7) je podľa [4] výpočet matice druhých derivácií (5.2). Táto matica sa nazýva hessián. Na výpočet odhadov $\tilde{\boldsymbol{\beta}}$ však môžeme použiť aj inú metódu, pričom v nej nepoužijeme priamo hessián, ale tzv. Fischerovu skórovaciu maticu (4.9). V literatúre sa táto metóda nazýva aj IRLS (Iteratively Reweighted Least Squares). Hoci za určitých (dosť všeobecných) podmienok sú tieto dva algoritmy na rekuretný odhad parametra ekvivalentné, aj tak sa môžu medzi týmito odhadmi objaviť malé nezrovnalosti, spôsobené rozdielmi v štartovacej hodnote a rozdielnou mierou konvergencie. Obe metódy sa však asymptoticky rovnajú (pozri [4]).

5.2 IRLS metóda

Majme Newtonov-Raphsonov algoritmus na rekuretný výpočet kvázi-maximálne vierohodného odhadu parametra $\boldsymbol{\beta}$

$$\boldsymbol{\beta}^{(r)} = \boldsymbol{\beta}^{(r-1)} + \left\{\mathbf{S}(\boldsymbol{\beta}^{(r-1)})\right\}^{-1}\frac{\partial Q}{\partial\boldsymbol{\beta}^{(r-1)}}, \qquad (5.11)$$

pozri (5.7). Iná metóda na výpočet tohto odhadu je tzv. Fischerová skórovacia metóda, v ktorej sa člen $\mathbf{S}(\beta^{(r-1)})$ nahradí jeho strednou hodnotou $E\left(\mathbf{S}(\beta^{(r-1)})\right)$. Dostávame

$$\boldsymbol{\beta}^{(r)} = \boldsymbol{\beta}^{(r-1)} + \left\{E\left(\mathbf{S}(\boldsymbol{\beta}^{(r-1)})\right)\right\}^{-1}\frac{\partial Q}{\partial\boldsymbol{\beta}^{(r-1)}}. \qquad (5.12)$$

Využitím vzťahu $var(Y_i) = E\left\{(Y_i-\mu_i)^2\right\} = V(\mu_i)a(\phi)$ dostávame (pozri(4.16))

$$\{\mathbf{F}\}_{s,t} = \{E(\mathbf{S})\}_{s,t} = -E\left(\frac{\partial^2 Q}{\partial\beta_s\partial\beta_t}\right) = E\left(\frac{\partial Q\partial Q}{\partial\beta_s\partial\beta_t}\right) =$$

$$= \sum_{i=1}^{n}\frac{1}{a(\phi)V(\mu_i)}\frac{\partial\mu_i}{\partial\beta_s}\frac{\partial\mu_i}{\partial\beta_t} = \sum_{i=1}^{n}\left(\frac{\partial\mu_i}{\partial\eta_i}\right)_i^2\frac{1}{V(\mu_i)a(\phi)}x_{is}x_{it}. \qquad (5.13)$$

Odhadom kovariančnej matice odhadu $\tilde{\boldsymbol{\beta}}$ je podľa [3]

$$\widetilde{var(\tilde{\boldsymbol{\beta}})} = \left\{\mathbf{F}(\tilde{\boldsymbol{\beta}})\right\}^{-1}, \qquad (5.14)$$

ktorej (s,t)-ty prvok je definovaný v (5.13) a $a(\phi)$ odhadujeme pomocou vzťahu (3.33).

Kapitola 6

Simulačná štúdia

V tejto časti práce sa zameriavame na charakterizáciu základných štatistických vlastností kvázi-vierohodného odhadu neznámeho vektora parametrov $\boldsymbol{\beta}$ vzhľadom na zvyšujúci sa rozsah počtu pozorovaní. Konkrétne sú na základe simulovaných hodnôt spočítané empirické hladiny významnosti testu popísaného v časti 4.4, pričom kovariančnú maticu odhadu $\tilde{\boldsymbol{\beta}}$ považujeme za neznámu a vo vzťahu (4.17) je nahradená skutočná kovariančná matica odhadu $\tilde{\boldsymbol{\beta}}$ jej odhadom (5.14).

6.1 Generovanie údajov

Predpokladajme, že

$$\mu_i = e^{\mathbf{x}_i'\boldsymbol{\beta}}, \tag{6.1}$$

kde $i = (1, 2, \dots n)$, $\mathbf{x}_i$ je známy vektor kovariátov a $\boldsymbol{\beta}$ je neznámy vektor parametrov. Nech

$$Y_i^* \sim Po(\mu_i), \quad i = (1, 2, \dots n), \tag{6.2}$$

sú navzájom nezávislé. Položme

$$Y_i = cY_i^*, \tag{6.3}$$

kde c je nejaká, vopred určená konštanta. V takomto prípade rozdelenie Y_i nepatrí medzi štandardné známe rozdelenia, avšak vieme, že

$$E(Y_i) = c\mu_i \tag{6.4}$$

a

$$var(Y_i) = c^2 var(Y_i^*) = c^2(\mu_i). \tag{6.5}$$

Týmto sú splnené podmienky uvedené v časti 4.1 a to, že jednotlivé pozorovania sú navzájom nezávislé a varianca každého pozorovania je známou funkciou jeho strednej hodnoty (až na konštantu). Odhad neznámeho vektora parametrov $\boldsymbol{\beta}$ z realizácii y_i náhodných premenných Y_i môžeme v tomto prípade získať metódou kvázi-maximálnej vierohodnosti.

Program ne generovanie takýchto údajov, skonštruovaný v programovacom jazyku *MATLAB* je uvedený v prílohe.

6.2 Empirické hladiny významnosti

V programovacom jazyku *MATLAB* bol vytvorený aj program k získaniu odhadu vektora parametrov $\boldsymbol{\beta}$ pomocou IRLS metódy popísanej v časti 5.2. Tento program je uvedený v prílohe. Jeho výstupom je okrem už spomínaného kvázi-maximálne vierohodného odhadu $\tilde{\boldsymbol{\beta}}$ aj odhad kovariančnej matice $\tilde{\boldsymbol{\beta}}$ definovaný vzťahom (5.14). Následne boli pre rôzne počty pozorovaní n spočítané na základe 10000 opakovaní (metódou Monte Carlo simulácií) empirické hladiny významnosti testu hypotézy

$$H_0 : \mathbf{K}'\boldsymbol{\beta} = \mathbf{K}'\boldsymbol{\beta}_0,$$

pomocou testovacej štatistiky

$$Wald_K = (\mathbf{K}'\tilde{\boldsymbol{\beta}} - \mathbf{K}'\boldsymbol{\beta}_0)'(\mathbf{K}'\widetilde{(\mathbf{X}'\mathbf{W}\mathbf{X})}^{-1}\mathbf{K})^{-1}(\mathbf{K}'\tilde{\boldsymbol{\beta}} - \mathbf{K}'\boldsymbol{\beta}_0), \tag{6.6}$$

kde $\widetilde{(\mathbf{X}'\mathbf{W}\mathbf{X})}^{-1} = \left\{\mathbf{F}(\tilde{\boldsymbol{\beta}})\right\}^{-1}$ definované vzťahom (5.14). Tu sme podľa časti 4.4 predpokladali, že za platnosti hypotézy má $Wald_K$ chi kvadrát rozdelenie s k stupňami voľnosti a teda testovanú hypotézu zamietame na hladine významnosti α ak $Wald_K \geq \chi_k^2(1-\alpha)$. Získané výsledky sú uvedené v nasledujúcom texte. Test sme realizovali tak, že za $\boldsymbol{\beta}_0$ sme položili "spávnu" hodnotu parametra $\boldsymbol{\beta}$.

6.3 Výsledky simulačnej štúdie

Príklad 1.

Podľa (6.2) sme vygenerovali vektor pozorovaní $\mathbf{y}^*$ pre parameter $\boldsymbol{\mu}_0 = (1, 2, \dots, n)'$. Pre známu strednú hodnotu $\boldsymbol{\mu}_0$ sme na základe (6.1) určili maticu $\mathbf{X}$ pre $\boldsymbol{\beta}_0 = (1, 2)'$ tak, aby $\mu_{0i} = e^{\mathbf{x}_i' \boldsymbol{\beta}_0}$, $i = (1, 2, \dots n)$. Ďalej sme podľa (6.3) pre konštantu $c = 2$ určili vektor pozorovaní $\mathbf{y}$. Na základe takto získanej realizácie $\mathbf{y}$ sme pomocou programu uvedeného v prílohe odhadli vektor parametrov $\tilde{\boldsymbol{\beta}}$ a následne testovali hypotézu H_0 z časti 6.2 na hladine významnosti $\alpha = 0,05$ v prípade rôznych tvarov matice $\mathbf{K}$. Uvažovali sme

$$\mathbf{K}_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \mathbf{K}_2 = \begin{pmatrix} 1 & 0 \end{pmatrix}, \quad \mathbf{K}_3 = \begin{pmatrix} 0 & 1 \end{pmatrix}, \quad \mathbf{K}_4 = \begin{pmatrix} 1 & 1 \end{pmatrix}, \quad \mathbf{K}_5 = \begin{pmatrix} 1 & -1 \end{pmatrix}.$$

V tabuľke 6.1 sú uvedené empirické hladiny významnosti testovanej hypotézy H_0 z 6.2 z 10000 opakovaní.

matica $\mathbf{K}$	$\mathbf{K}_1$	$\mathbf{K}_2$	$\mathbf{K}_3$	$\mathbf{K}_4$	$\mathbf{K}_5$
počet pozorovaní					
$n = 5$	0,1764	0,1344	0,1350	0,1347	0,1336
$n = 10$	0,0958	0,0759	0,0781	0,0792	0,0781
$n = 20$	0,0762	0,0635	0,0665	0,0657	0,0673
$n = 50$	0,0615	0,0536	0,0526	0,0518	0,0524
$n = 100$	0,0553	0,0518	0,0520	0,0511	0,0521
$n = 1000$	0,0468	0,0471	0,0470	0,0471	0,0461

Tabuľka 6.1: Empirické hladiny významnosti pre test hypotézy H_0 o regresnom parametri danej v 6.2 s nominálnou hladinou významnosti $\alpha = 0,05$ pre hodnoty zadané v *Príklade 1.*

Príklad 2.

Podľa (6.2) sme vygenerovali vektor pozorovaní $\mathbf{y}^*$ pre parameter $\boldsymbol{\mu}_0 = (0,1; 0,2; \ldots; 0,1n)'$. Pre známu strednú hodnotu $\boldsymbol{\mu}_0$ sme na základe (6.1) určili maticu $\mathbf{X}$ pre $\boldsymbol{\beta}_0 = (1,2)'$ tak, aby $\mu_{0i} = e^{\mathbf{x}_i' \boldsymbol{\beta}_0}$, $i = (1, 2, \ldots n)$. Ďalej sme podľa (6.3) pre konštantu $c = 2$ určili vektor pozorovaní $\mathbf{y}$. Na základe takto získanej realizácie $\mathbf{y}$ sme pomocou programu uvedeného v prílohe odhadli vektor parametrov $\tilde{\boldsymbol{\beta}}$ a následne testovali hypotézu H_0 z časti 6.2 na hladine významnosti $\alpha = 0,05$ v prípade rôznych tvarov matice $\mathbf{K}$. Uvažovali sme

$$\mathbf{K}_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \mathbf{K}_2 = \begin{pmatrix} 1 & 0 \end{pmatrix}, \quad \mathbf{K}_3 = \begin{pmatrix} 0 & 1 \end{pmatrix}, \quad \mathbf{K}_4 = \begin{pmatrix} 1 & 1 \end{pmatrix}, \quad \mathbf{K}_5 = \begin{pmatrix} 1 & -1 \end{pmatrix}.$$

V tabuľke 6.2 sú uvedené empirické hladiny významnosti testovanej hypotézy H_0 z 6.2 z 10000 opakovaní.

matica $\mathbf{K}$	$\mathbf{K}_1$	$\mathbf{K}_2$	$\mathbf{K}_3$	$\mathbf{K}_4$	$\mathbf{K}_5$
počet pozorovaní					
$n = 5$	0,2086	0,0768	0,0763	0,0728	0,1286
$n = 10$	0,1120	0,0919	0,0939	0,0905	0,0748
$n = 20$	0,0780	0,0654	0,0652	0,0649	0,0617
$n = 50$	0,0595	0,0514	0,0569	0,0589	0,0562
$n = 100$	0,0571	0,0545	0,0567	0,0550	0,0532
$n = 1000$	0,0508	0,0484	0,0491	0,0501	0,0491

Tabuľka 6.2: Empirické hladiny významnosti pre test hypotézy H_0 o regresnom parametri danej v 6.2 s nominálnou hladinou významnosti $\alpha = 0,05$ pre hodnoty zadané v *Príklade 2.*

Príklad 3.

Podľa (6.2) sme vygenerovali vektor pozorovaní $\mathbf{y}^*$ pre parameter $\boldsymbol{\mu}_0 = (2, 4, \ldots, 2n)'$. Pre známu strednú hodnotu $\boldsymbol{\mu}_0$ sme na základe (6.1) určili maticu $\mathbf{X}$ pre $\boldsymbol{\beta}_0 = (1, 2)'$ tak, aby $\mu_{0i} = e^{\mathbf{x}_i' \boldsymbol{\beta}_0}, \quad i = (1, 2, \ldots n)$. Ďalej sme podľa (6.3) pre konštantu $c = 2$ určili vektor pozorovaní $\mathbf{y}$. Na základe takto získanej realizácie $\mathbf{y}$ sme pomocou programu uvedeného v prílohe odhadli vektor parametrov $\tilde{\boldsymbol{\beta}}$ a následne testovali hypotézu *H0* z časti 6.2 na hladine významnosti $\alpha = 0,05$ v prípade rôznych tvarov matice $\mathbf{K}$. Uvažovali sme

$$\mathbf{K}_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \mathbf{K}_2 = \begin{pmatrix} 1 & 0 \end{pmatrix}, \quad \mathbf{K}_3 = \begin{pmatrix} 0 & 1 \end{pmatrix}, \quad \mathbf{K}_4 = \begin{pmatrix} 1 & 1 \end{pmatrix}, \quad \mathbf{K}_5 = \begin{pmatrix} 1 & -1 \end{pmatrix}.$$

V tabuľke 6.3 sú uvedené empirické hladiny významnosti testovanej hypotézy *H0* z 6.2 z 10000 opakovaní.

matica $\mathbf{K}$	$\mathbf{K}_1$	$\mathbf{K}_2$	$\mathbf{K}_3$	$\mathbf{K}_4$	$\mathbf{K}_5$
počet pozorovaní					
n = 5	0,1873	0,1396	0,1423	0,1395	0,1426
n = 10	0,1032	0,0857	0,0833	0,0889	0,0827
n = 20	0,0774	0,0698	0,0717	0,0661	0,0694
n = 50	0,0608	0,0544	0,0545	0,0565	0,0529
n = 100	0,0530	0,0519	0,0513	0,0519	0,0520

Tabuľka 6.3: Empirické hladiny významnosti pre test hypotézy *H0* o regresnom parametri danej v 6.2 s nominálnou hladinou významnosti $\alpha = 0,05$ pre hodnoty zadané v *Príklade 3.*

Príklad 4.

Podľa (6.2) sme vygenerovali vektor pozorovaní $\mathbf{y}^*$ pre parameter $\boldsymbol{\mu}_0 = (1, 2, \ldots, n)'$. Pre známu strednú hodnotu $\boldsymbol{\mu}_0$ sme na základe (6.1) určili maticu $\mathbf{X}$ pre $\boldsymbol{\beta}_0 = (1, 2)'$ tak, aby $\mu_{0i} = e^{\mathbf{x}_i' \boldsymbol{\beta}_0}$, $i = (1, 2, \ldots n)$. Ďalej sme podľa (6.3) pre konštantu $c = 20$ určili vektor pozorovaní $\mathbf{y}$. Na základe takto získanej realizácie $\mathbf{y}$ sme pomocou programu uvedeného v prílohe odhadli vektor parametrov $\tilde{\boldsymbol{\beta}}$ a následne testovali hypotézu *H₀* z časti 6.2 na hladine významnosti $\alpha = 0,05$ v prípade rôznych tvarov matice $\mathbf{K}$. Uvažovali sme

$$\mathbf{K}_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \mathbf{K}_2 = \begin{pmatrix} 1 & 0 \end{pmatrix}, \quad \mathbf{K}_3 = \begin{pmatrix} 0 & 1 \end{pmatrix}, \quad \mathbf{K}_4 = \begin{pmatrix} 1 & 1 \end{pmatrix}, \quad \mathbf{K}_5 = \begin{pmatrix} 1 & -1 \end{pmatrix}.$$

V tabuľke 6.4 sú uvedené empirické hladiny významnosti testovanej hypotézy H_0 z 6.2 z 10000 opakovaní.

matica $\mathbf{K}$	$\mathbf{K}_1$	$\mathbf{K}_2$	$\mathbf{K}_3$	$\mathbf{K}_4$	$\mathbf{K}_5$
počet pozorovaní					
$n = 5$	0,1748	0,1294	0,1335	0,1339	0,1293
$n = 10$	0,1010	0,0818	0,0820	0,0816	0,0838
$n = 20$	0,0807	0,0664	0,0700	0,0683	0,0687
$n = 50$	0,0538	0,0511	0,0537	0,0530	0,0525
$n = 100$	0,0540	0,0558	0,0559	0,0536	0,0558

Tabuľka 6.4: Empirické hladiny významnosti pre test hypotézy H_0 o regresnom parametri danej v 6.2 s nominálnou hladinou významnosti $\alpha = 0,05$ pre hodnoty zadané v *Príklade 4.*

Príklad 5.

Podľa (6.2) sme vygenerovali vektor pozorovaní $\mathbf{y}^*$ pre parameter $\boldsymbol{\mu}_0 = (1, 2, \ldots, n)'$. Pre známu strednú hodnotu $\boldsymbol{\mu}_0$ sme na základe (6.1) určili maticu $\mathbf{X}$ pre $\boldsymbol{\beta}_0 = (1, 2)'$ tak, aby $\mu_{0i} = e^{\mathbf{x}_i'\boldsymbol{\beta}_0}$, $i = (1, 2, \ldots n)$. Ďalej sme podľa (6.3) pre konštantu $c = 200$ určili vektor pozorovaní $\mathbf{y}$. Na základe takto získanej realizácie $\mathbf{y}$ sme pomocou programu uvedeného v prílohe odhadli vektor parametrov $\tilde{\boldsymbol{\beta}}$ a následne testovali hypotézu H_0 z časti 6.2 na hladine významnosti $\alpha = 0,05$ v prípade rôznych tvarov matice $\mathbf{K}$. Uvažovali sme

$$\mathbf{K}_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \mathbf{K}_2 = \begin{pmatrix} 1 & 0 \end{pmatrix}, \quad \mathbf{K}_3 = \begin{pmatrix} 0 & 1 \end{pmatrix}, \quad \mathbf{K}_4 = \begin{pmatrix} 1 & 1 \end{pmatrix}, \quad \mathbf{K}_5 = \begin{pmatrix} 1 & -1 \end{pmatrix}.$$

V tabuľke 6.5 sú uvedené empirické hladiny významnosti testovanej hypotézy H_0 z 6.2 z 10000 opakovaní.

matica $\mathbf{K}$	$\mathbf{K}_1$	$\mathbf{K}_2$	$\mathbf{K}_3$	$\mathbf{K}_4$	$\mathbf{K}_5$
počet pozorovaní					
$n = 5$	0,1715	0,1305	0,1284	0,1303	0,1239
$n = 10$	0,1040	0,0767	0,0801	0,0854	0,0783
$n = 20$	0,0724	0,0656	0,0658	0,0668	0,0659
$n = 50$	0,0597	0,0544	0,0566	0,0515	0,0560
$n = 100$	0,0563	0,0542	0,0526	0,0543	0,0542

Tabuľka 6.5: Empirické hladiny významnosti pre test hypotézy H_0 o regresnom parametri danej v 6.2 s nominálnou hladinou významnosti $\alpha = 0,05$ pre hodnoty zadané v *Príklade 5.*

Príklad 6.

Podľa (6.2) sme vygenerovali vektor pozorovaní $\mathbf{y}^*$ pre parameter $\boldsymbol{\mu}_0 = (1, 2, \ldots, n)'$. Pre známu strednú hodnotu $\boldsymbol{\mu}_0$ sme na základe (6.1) určili maticu $\mathbf{X}$ pre $\boldsymbol{\beta}_0 = (0,01; 0,002)'$ tak, aby $\mu_{0i} = e^{\mathbf{x}_i'\boldsymbol{\beta}_0}$, $i = (1, 2, \ldots n)$. Ďalej sme podľa (6.3) pre konštantu $c = 2$ určili vektor pozorovaní $\mathbf{y}$. Na základe takto získanej realizácie $\mathbf{y}$ sme pomocou programu uvedeného v prílohe odhadli vektor parametrov $\tilde{\boldsymbol{\beta}}$ a následne testovali hypotézu *H0* z časti 6.2 na hladine významnosti $\alpha = 0,05$ v prípade rôznych tvarov matice $\mathbf{K}$. Uvažovali sme

$$\mathbf{K}_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \mathbf{K}_2 = \begin{pmatrix} 1 & 0 \end{pmatrix}, \quad \mathbf{K}_3 = \begin{pmatrix} 0 & 1 \end{pmatrix}, \quad \mathbf{K}_4 = \begin{pmatrix} 1 & 1 \end{pmatrix}, \quad \mathbf{K}_5 = \begin{pmatrix} 1 & -1 \end{pmatrix}.$$

V tabuľke 6.6 sú uvedené empirické hladiny významnosti testovanej hypotézy *H0* z 6.2 z 10000 opakovaní.

matica $\mathbf{K}$	$\mathbf{K}_1$	$\mathbf{K}_2$	$\mathbf{K}_3$	$\mathbf{K}_4$	$\mathbf{K}_5$
počet pozorovaní					
n = 5	0,1718	0,1199	0,1298	0,1196	0,1198
n = 10	0,1020	0,0816	0,0832	0,0816	0,0816
n = 20	0,0756	0,0597	0,0617	0,0598	0,0597
n = 50	0,0646	0,0570	0,0575	0,0568	0,0570
n = 100	0,0563	0,0529	0,0518	0,0529	0,0529

Tabuľka 6.6: Empirické hladiny významnosti pre test hypotézy *H0* o regresnom parametri danej v 6.2 s nominálnou hladinou významnosti $\alpha = 0,05$ pre hodnoty zadané v *Príklade 6.*

Príklad 7.

Podľa (6.2) sme vygenerovali vektor pozorovaní $\mathbf{y}^*$ pre parameter $\boldsymbol{\mu}_0 = (1, 2, \ldots, n)'$. Pre známu strednú hodnotu $\boldsymbol{\mu}_0$ sme na základe (6.1) určili maticu $\mathbf{X}$ pre $\boldsymbol{\beta}_0 = (1, 2)'$ tak, aby $\mu_{0i} = e^{\mathbf{x}_i'\boldsymbol{\beta}_0}, \quad i = (1, 2, \ldots n)$. Ďalej sme podľa (6.3) pre konštantu $c = 2$ určili vektor pozorovaní $\mathbf{y}$. Na základe takto získanej realizácie $\mathbf{y}$ sme pomocou programu uvedeného v prílohe odhadli vektor parametrov $\tilde{\boldsymbol{\beta}}$ a následne testovali hypotézu *H₀* z časti 6.2 na hladine významnosti $\alpha = 0,1$ v prípade rôznych tvarov matice $\mathbf{K}$. Uvažovali sme

$$\mathbf{K}_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \mathbf{K}_2 = \begin{pmatrix} 1 & 0 \end{pmatrix}, \quad \mathbf{K}_3 = \begin{pmatrix} 0 & 1 \end{pmatrix}, \quad \mathbf{K}_4 = \begin{pmatrix} 1 & 1 \end{pmatrix}, \quad \mathbf{K}_5 = \begin{pmatrix} 1 & -1 \end{pmatrix}.$$

V tabuľke 6.7 sú uvedené empirické hladiny významnosti testovanej hypotézy H_0 z 6.2 z 10000 opakovaní.

matica $\mathbf{K}$	$\mathbf{K}_1$	$\mathbf{K}_2$	$\mathbf{K}_3$	$\mathbf{K}_4$	$\mathbf{K}_5$
počet pozorovaní					
$n = 5$	0,2198	0,1850	0,1757	0,1860	0,1735
$n = 10$	0,1598	0,1318	0,1371	0,1390	0,1349
$n = 20$	0,1292	0,1165	0,1135	0,1158	0,1134
$n = 50$	0,1158	0,1083	0,1100	0,1129	0,1091
$n = 100$	0,1079	0,1039	0,1064	0,1019	0,1042

Tabuľka 6.7: Empirické hladiny významnosti pre test hypotézy H_0 o regresnom parametri danej v 6.2 s nominálnou hladinou významnosti $\alpha = 0,1$ pre hodnoty zadané v *Príklade 7.*

Príklad 8.

Podľa (6.2) sme vygenerovali vektor pozorovaní $\mathbf{y}^*$ pre parameter $\boldsymbol{\mu}_0 = (1, 2, \ldots, n)'$. Pre známu strednú hodnotu $\boldsymbol{\mu}_0$ sme na základe (6.1) určili maticu $\mathbf{X}$ pre $\boldsymbol{\beta}_0 = (1, 2)'$ tak, aby $\mu_{0i} = e^{\mathbf{x}_i'\boldsymbol{\beta}_0}$, $i = (1, 2, \ldots n)$. Ďalej sme podľa (6.3) pre konštantu $c = 2$ určili vektor pozorovaní $\mathbf{y}$. Na základe takto získanej realizácie $\mathbf{y}$ sme pomocou programu uvedeného v prílohe odhadli vektor parametrov $\tilde{\boldsymbol{\beta}}$ a následne testovali hypotézu H_0 z časti 6.2 na hladine významnosti $\alpha = 0,01$ v prípade rôznych tvarov matice $\mathbf{K}$. Uvažovali sme

$$\mathbf{K}_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \mathbf{K}_2 = \begin{pmatrix} 1 & 0 \end{pmatrix}, \quad \mathbf{K}_3 = \begin{pmatrix} 0 & 1 \end{pmatrix}, \quad \mathbf{K}_4 = \begin{pmatrix} 1 & 1 \end{pmatrix}, \quad \mathbf{K}_5 = \begin{pmatrix} 1 & -1 \end{pmatrix}.$$

V tabuľke 6.8 sú uvedené empirické hladiny významnosti testovanej hypotézy H_0 z 6.2 z 10000 opakovaní.

matica $\mathbf{K}$	$\mathbf{K}_1$	$\mathbf{K}_2$	$\mathbf{K}_3$	$\mathbf{K}_4$	$\mathbf{K}_5$
počet pozorovaní					
$n = 5$	0,1086	0,0736	0,0736	0,0725	0,0750
$n = 10$	0,0437	0,0319	0,0335	0,0286	0,0333
$n = 20$	0,0214	0,0177	0,0176	0,0180	0,0179
$n = 50$	0,0163	0,0144	0,0147	0,0136	0,0150
$n = 100$	0,0119	0,0118	0,0110	0,0120	0,0114

Tabuľka 6.8: Empirické hladiny významnosti pre test hypotézy H_0 o regresnom parametri danej v 6.2 s nominálnou hladinou významnosti $\alpha = 0,01$ pre hodnoty zadané v *Príklade 8.*

Príklad 9.

Podľa (6.2) sme vygenerovali vektor pozorovaní $\mathbf{y}^*$ pre parameter $\boldsymbol{\beta}_0 = (0,1; 0,2; 0,3)'$ a danú maticu plánu $\mathbf{X}$, kde $\mathbf{x}_i' = (1, \frac{i}{100}, \frac{i^2}{100})'$. Hodnota $\boldsymbol{\mu}_0$ bola dopočítaná na základe (6.1). Ďalej sme podľa (6.3) pre konštantu $c = 2$ určili vektor pozorovaní $\mathbf{y}$. Na základe takto získanej realizácie $\mathbf{y}$ sme pomocou programu uvedeného v prílohe odhadli vektor parametrov $\tilde{\boldsymbol{\beta}}$ a následne testovali hypotézu H_0 z časti 6.2 na hladine významnosti $\alpha = 0,05$ v prípade rôznych tvarov matice $\mathbf{K}$. Uvažovali sme

$$\mathbf{K}_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{K}_2 = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix}, \quad \mathbf{K}_3 = \begin{pmatrix} 0 & 1 & 0 \end{pmatrix},$$
$$\mathbf{K}_4 = \begin{pmatrix} 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{K}_5 = \begin{pmatrix} 1 & 1 & 1 \end{pmatrix}.$$

V tabuľke 6.9 sú uvedené empirické hladiny významnosti testovanej hypotézy H_0 z 6.2 z 10000 opakovaní.

matica $\mathbf{K}$	$\mathbf{K}_1$	$\mathbf{K}_2$	$\mathbf{K}_3$	$\mathbf{K}_4$	$\mathbf{K}_5$
počet pozorovaní					
$n = 5$	0,2021	0,1298	0,1317	0,1395	0,1291
$n = 10$	0,0907	0,0690	0,0751	0,0760	0,0753
$n = 20$	0,0803	0,0632	0,0648	0,0640	0,0653
$n = 50$	0,0605	0,0576	0,0562	0,0535	0,0556
$n = 100$	0,0558	0,0518	0,0510	0,0510	0,0510

Tabuľka 6.9: Empirické hladiny významnosti pre test hypotézy H_0 o regresnom parametri danej v 6.2 s nominálnou hladinou významnosti $\alpha = 0,05$ pre hodnoty zadané v *Príklade 9.*

Na základe výsledkov simulácií uvedených tabuľkách 6.1 - 6.9 môžeme konštatovať, že v uvažovaných prípadoch sa zo zvyšujucím počtom pozorovaní n empirická hladina významosti blíži k svojej nominálnej hodnote α, pričom, už pre $n \approx 50$ pozorovaní dosahuje predom stanovenú nominálnu hodnotu.

Kapitola 7

Záver

V diplomovej práci sme sa venovali metóde kvázi-maximálnej vierohodnosti, ktorá vychádza zo zovšeobecnených lineárnych (regresných) modelov. Tieto modely sú uvedené v druhej kapitole. Sú založené na exponenciálnej triede rozdelení, ktorá je tiež opísaná v druhej kapitole.

Metóda kvázi-maximálnej vierohodnosti, ktorú sme zadefinovali a jej vlastnosti podrobnejšie opísali v štvrtej kapitole, má veľký význam pri odhadovaní neznámych regresných parametrov. Jej najväčšou výhodou je, že pri odhadovaní neznámych parametrov, nie je nutné poznať konkrétne rozdelenie pravdepodobnosti náhodnej premennej (meraní) $\mathbf{Y}$. V štvrtej kapitole sme tiež odvodili odhad získaný metódou kvázi-maximálnej vierohodnosti a ukázali jeho asymptotické vlastnoti.

V kapitole 5 práce sme uviedli dva algoritmy (Newtonov-Raphsonov a IRLS) na rekurentný výpočet kvázi-maximálne vierohodného odhadu. IRLS algoritmus je využitý pri tvorbe programu v programovacom jazyku *MATLAB* na výpočet kvázi-maximálne vierohodného odhadu ako aj na výpočet odhadu jeho kovariančnej matice.

V 6 kapitole boli pomocou nami vytvoreného programu overené niektoré asymptotické vlastnosti kvázi-maximálne vierohodného odhadu v prípade, že kovariančná matica odhadu nie je známa a je potrebné ju tiež odhadnúť. Na základe dosiahnutých výsledkov sa dá usudzovať, že za predpokladu platnosti uvažovaného modelu môžeme v praktických úlohách považovať odhad vektora regresných parametrov za postačujúci v prípade, že máme k dispozícii aspoň 50 hodnôt meranej veličiny.

Literatúra

[1] Cerchiara R. R., Edwards M., Gambini A., 2008: *Generalized linear models in life insurance: decrements and risk factor analysis under Solvency II.* In: 18th International AFIR Colloquium.

[2] Dobson A. J., 2002: *An introduction to generalized linear models.* Chapman & Hall/CRC, London.

[3] Hardin J.W., Hilbe J. M., 2003: *Generalized estimating equations.* Chapman & Hall/CRC, London.

[4] Hardin J.W., Hilbe J. M., 2007: *Generalized linear models and extensions, 2nd ed.*. Stata press, Texas.

[5] Chou J.S., 2009: *Generalized linear model-based expert system for estimating the cost of transportation projects.* Expert Systems with Applications, 36, 4253–4267.

[6] McCullagh P., 1983: *Quasi-likelihood functions.* The annals of statistics, 11, 59-67.

[7] McCullagh P., Nelder J. A., 1989: *Generalized linear models, 2nd ed.*. Chapman & Hall/CRC, London.

[8] McCulloch CH. E., Searle S. R., Neuhaus J. M., 2008: *Generalized, linear, and mixed models, 2nd ed.*. John Wiley & sons, New Jersey.

[9] Mittelhammer R. C., Judge G. G., Miller G. J., 2000: *Econometric foundations.* Cambridge university press, Cambridge.

[10] Nelder J. A., Wedderburn R. W. M., 1972: *Generalized linear models.* Journal of the Royal Statistical Society, 135, 370-447.

[11] Wedderburn R. W. M., 1974: *Quasi-likelihood functions, generalized linear models, and the Gauss-Newton method*. Biometrica, 61, 439-447.

Kapitola 8

Príloha

Generovanie údajov

```
% Zadanie skutočnej hodnoty regresného parametra
beta_skutocne = [1; 2];
% Ak je vektor_strednych_hodnot = 1, vygeneruje vektor stredných hodnôt,
% viď nižšie. Može sa použiť v prípade 2-rozmerného parametra beta_skutocne
vektor_strednych_hodnot = 0;
% Ak je matica_X = 1, vygeneruje maticu plánu X, viď nižšie.
% Može sa použiť v prípade ľubovoľného rozmeru parametra beta_skutocne
matica_X = 1;
% Ak sú obe predchádzajúce premenné rovné 0, je nutné zadať buď maticu X
% alebo vektor stredných hodnôt
% mi_skutocne = [1; 3; 2; 4; 2; 4];
% X = 0;
pocet_pozorovani = 1000;
nasobok = 2;
n = pocet_pozorovani;
k = size(beta_skutocne, 1);
%Vytvorenie matice X, ak nie je zadaná
if (matica_X == 1)
    X = zeros(n, k);
    X(:, 1) = 1;
    for i = 1 : n
        for j = 2 : k
            X(i, j) = i^(j - 1)/100;
        end
    end
% inverzia link funkcie
    mi_generovane = zeros(n, 1);
    for i = 1 : n
        mi_generovane(i) = exp(X(i, :)*beta_skutocne);
    end
end
% Vytvorenie stredných hodnôt a matice X pre 2-rozmerný regresný parameter
if (vektor_strednych_hodnot == 1)
    mi_generovane = (1 : n)';
    X = zeros(n, k);
    X(:, 1) = 1;
    for i = 1 : n
        X(i, 2) = (log(mi_generovane(i)) - X(i, 1)*beta_skutocne(1))/...
```

```
            (beta_skutocne(2));
    end
end
mi_skutocne = (nasobok)*mi_generovane;
% Generovanie y-nov
y_Po = zeros(n, 1);
kontrola_y = 0;
while (sum(kontrola_y) < k)
    for i = 1 : n
        y_Po(i) =  poissrnd(mi_generovane(i));
    end
    kontrola_y = (y_Po ~= 0);
end
y = nasobok*y_Po;
```

Odhad parametrov

```
% Zadanie vstupných údajov
pocet_pozorovani = 1000;
nasobok = 2;
n = pocet_pozorovani;
k = size(beta_skutocne, 1);
Max_NO_Iteration = 1000;
No_of_iteration = 0;
L = eye(size(beta_skutocne, 1));
chi2_kvantil = chi2inv(0.95, rank(L));
% Počiatočná hodnota odhadhovaného parametra
beta = zeros(k, 1);
beta(1) = log(mean(y));
beta_old = beta + 1;
% Odhad parametrov - iteracna procedura
while((norm(beta-beta_old) > 1000*eps) && ...
        No_of_iteration < Max_NO_Iteration)
    No_of_iteration = No_of_iteration + 1;
    beta_old = beta;
    mi = zeros(n, 1);
    for i = 1 : n
        mi(i) = nasobok*exp(X(i, :)*beta);
    end
    derivacia = diag(mi);
    V_mi = mi;
    V_mi_inv = diag(1./V_mi);
    grad = ((y - mi)'*V_mi_inv*derivacia*X)';
    hes = X'*derivacia*V_mi_inv*derivacia*X;
    hes_inverzia = inv(hes);
    beta = beta_old + hes_inverzia*grad;
end
mi = zeros(n, 1);
for i = 1 : n
    mi(i) = nasobok*exp(X(i, :)*beta);
end
derivacia = diag(mi);
V_mi = mi;
V_mi_inv = diag(1./V_mi);
grad = ((y - mi)'*V_mi_inv*derivacia*X)';
```

```
hes = X'*derivacia*V_mi_inv*derivacia*X;
phi = 0;
for i = 1 : n
    phi = phi + (((y(i) - mi(i))^2)/V_mi(i));
end
phi = (1/(n - k))*phi;
Var_beta_odhad = phi*inv(hes);
% Wald test, ktorej výstupm je premenná test. Ak je jej hodnota
% rovná 0, tak zamietame testovanú hypotézu
Wald = (beta - beta_skutocne)'*L'*inv(L*Var_beta_odhad*L')*L*...
    (beta - beta_skutocne);
test = (Wald <= chi2_kvantil);
```

Výpočet empirických hladín významnosti

Program je spojením (a malým rozšírením) predchádzajúcich dvoch procedúr.

```
beta_skutocne = [1; 2];
vektor_strednych_hodnot = 1;
matica_X = 0;
% mi_skutocne = [1; 3; 2; 4; 2; 4];
% X = 0;
pocet_pozorovani = 100;
nasobok = 2;
n = pocet_pozorovani;
k = size(beta_skutocne, 1);
Max_NO_Iteration = 1000;
poc_opakovani_sim1 = 10000;
pocet_iteracii = zeros(poc_opakovani_sim1, 1);
L = eye(size(beta_skutocne, 1));
chi2_kvantil = chi2inv(0.95, rank(L));
if (matica_X == 1)
    X = zeros(n, k);
    X(:, 1) = 1;
    for i = 1 : n
        for j = 2 : k
            X(i, j) = i^(j - 1)/100;
        end
    end
    mi_generovane = zeros(n, 1);
    for i = 1 : n
        mi_generovane(i) = exp(X(i, :)*beta_skutocne);
    end
end
if (vektor_strednych_hodnot == 1)
    mi_generovane = (1 : n)';
    X = zeros(n, k);
    X(:, 1) = 1;
    for i = 1 : n
        X(i, 2) = (log(mi_generovane(i)) - X(i, 1)*beta_skutocne(1))/...
            (beta_skutocne(2));
    end
end
mi_skutocne = (nasobok)*mi_generovane;
pocet_uspesnych = 0;
for poc_opakovani_sim = 1 : poc_opakovani_sim1
% Generovanie y-nov
```

```
    y_Po = zeros(n, 1);
    kontrola_y = 0;
    while (sum(kontrola_y) < k)
        for i = 1 : n
            y_Po(i) =  poissrnd(mi_generovane(i));
        end
        kontrola_y = (y_Po ~= 0);
    end
    y = nasobok*y_Po;
% Odhad parametrov - výpočet štartovacej hodnoty
    beta = zeros(k, 1);
    beta(1) = log(mean(y));
    beta_old = beta + 1;
    No_of_iteration = 0;
% Odhad parametrov - iteračná procedúra
    while((norm(beta-beta_old) > 1000*eps) && ...
            No_of_iteration < Max_NO_Iteration)
        No_of_iteration = No_of_iteration + 1;
        beta_old = beta;
        mi = zeros(n, 1);
        for i = 1 : n
            mi(i) = nasobok*exp(X(i, :)*beta);
        end
        derivacia = diag(mi);
        V_mi = mi;
        V_mi_inv = diag(1./V_mi);
        grad = ((y - mi)'*V_mi_inv*derivacia*X)';
        hes = X'*derivacia*V_mi_inv*derivacia*X;
        hes_inverzia = inv(hes);
        beta = beta_old + hes_inverzia*grad;
    end
    mi = zeros(n, 1);
    for i = 1 : n
        mi(i) = nasobok*exp(X(i, :)*beta);
    end
    derivacia = diag(mi);
    V_mi = mi;
    V_mi_inv = diag(1./V_mi);
    grad = ((y - mi)'*V_mi_inv*derivacia*X)';
    hes = X'*derivacia*V_mi_inv*derivacia*X;
    phi = 0;
    for i = 1 : n
        phi = phi + (((y(i) - mi(i))^2)/V_mi(i));
    end
    phi = (1/(n - k))*phi;
    Var_beta_odhad = phi*inv(hes);
    pocet_iteracii(poc_opakovani_sim) = No_of_iteration;
% Wald test
    if(No_of_iteration < Max_NO_Iteration && norm(beta) > 1000*eps)
        pocet_uspesnych = pocet_uspesnych + 1;
        phi_odhad(pocet_uspesnych) = phi;
        Wald = (beta - beta_skutocne)'*L'*inv(L*Var_beta_odhad*L')*L*...
            (beta - beta_skutocne);
        test(pocet_uspesnych) = (Wald <= chi2_kvantil);
     end
end
empiricka_hladina_vyznamnosti = sum(test)/size(test, 2);
```