

Štatistická inferencia v časových radoch

DIPLOMOVÁ PRÁCA

Eduard Toman

Univerzita Komenského v Bratislave
Fakulta matematiky fyziky a informatiky
Katedra ekonomickej a finančnej matematiky, Bratislava 2007.

Školiteľ: PaedDr., RNDr. Stanislav Katina, PhD.

Bratislava 2007

Prehlasujem na svoju česť, že som diplomovú prácu vypracoval samostatne pod odborným vedením vedúceho diplomovej práce, s použitím uvedenej literatúry.

Za odbornú pomoc, rady a pripomienky pri písaní diplomovej práce ďakujem vedúcemu diplomovej práce PaedDr.,RNDr. Stanislavovi Katinovi, PhD.

Diplomová práca bola podporená grantami VEGA 1/30023/06 (KAMŠ FMFI UK) a zber a prvotné spracovanie údajov pre príklad zabezpečený grantom VEGA 2/6187/06 (UNPF SAV).

Abstrakt

Toman, Eduard: Štatistická inferencia v časových radoch. [Diplomová práca], Univerzita Komenského v Bratislave, Fakulta matematiky fyziky a informatiky, Katedra ekonomickej a finančnej matematiky, Bratislava 2006.

Školiteľ: PaedDr.,RNDr. Stanislav Katina, PhD.

V diplomovej práci skúmame testy používané v analýze časových radov. Permutačný test porovnáva dva nezávislé časové rady a Fisherov test overuje, či daný časový rad obsahuje skryté periódy. Ako alternatívu k permutačnému testu sme si zvolili *CP*-plot metódu. Ide o grafickú metódu, ktorá porovnáva dva nezávislé časové rady na základe ich normovaných kumulatívnych periodogramov (*NCP*).

Hlavným cieľom je preverenie vlastností chýb prvého a druhého druhu permutačného testu a Fisherovho testu pomocou simulačných štúdií a potom aplikovať tieto testy pri analýze časových radov lekárskeho údajov. Permutačný test je zovšeobecnením Kolmogorovho-Smirnovho testu dobrej zhody. Preto ďalším cieľom je urobiť aj simulačnú štúdiu vlastností chýb prvého a druhého druhu *KS*-testu a porovnať ich s permutačným testom.

Ukazuje sa, že permutačný a Fisherov test sú v praxi použiteľné len pre veľké (≥ 120) rozsahy súborov. Oba testy majú malú chybu prvého druhu, ale sila testov je pre v praxi často používané (≤ 50) rozsahy súborov slabá. U permutačného testu je sila rovná 0. U Fisherovho testu sa v sile testu vyskytujú veľké výkyvy.

Kľúčové slová: Časový rad, Periodogram, Simulácia.

Predhovor

Významnou časťou Matematickej štatistiky je analýza časových radov.

Existuje niekoľko testov porovnávajúcich dva nezávislé náhodné výbery. Najznámejším je Kolmogorov-Smirnov test dobrej zhody. Ak meriame alebo zisťujeme danú charakteristiku v postupných časových okamžikoch, bývajú pozorovania navzájom závislé a tvoria časový rad. V praxi potrebujeme vyriešiť niekoľko problémov, napr. môžeme mať dva nezávislé časové rady (napr. rady údajov z dvoch krajín) a chceme testovať hypotézu, že majú rovnaké rozdelenie pravdepodobnosti.

Alebo máme dva úseky rovnakého časového radu, napr. môže ísť o dva úseky pozorovaní meraných v dvoch po sebe idúcich obdobiach. Chceme testovať hypotézu, že medzi obdobiaami nenastala významná zmena, teda že rozdelenie pravdepodobnosti je stationárne po celé obdobie zisťovania.

Vidíme, že testy hypotéz o zhode dvoch časových radov majú pre prax veľký význam.

Autor

Obsah

1	Teoretické základy Kolmogorov-Smirnovho testu	
	dobrej zhody	6
1.1	Úvod	6
1.2	Základné pojmy	6
1.3	Kolmogorov-Smirnov test-princíp fungovania	6
2	Permutačný test	10
2.1	Úvod	10
2.2	Základné pojmy	10
2.3	Popis testu	11
2.4	Povaha testu	12
3	Teoretické základy Fisherovho testu	13
4	Simulačné štúdie	14
4.1	Úvod	14
4.2	Základné pojmy	14
4.3	Simulačná štúdia Kolmogorov-Smirnov testu	14
4.4	Simulačná štúdia permutačného testu	16
4.5	Simulačná štúdia Fisherovho testu	22
5	Aplikácia kváziperiodických časových radov	
	v kardiológii	27
5.1	Úvod	27
5.2	Materiál a metodika	28
5.2.1	Súbor jedincov	28
5.2.2	Štatistické metódy	28
5.2.3	Odhaľovanie významných frekvencií	30
5.2.4	Výsledky a diskusia	32

Úvod

V diplomovej práci sa zaoberáme časovými radmi. Na skúmanie časových radov využijeme regresnú analýzu a testovanie rôznych hypotéz.

Jedným z cieľov tejto práce je pomocou simulčných štúdií preveriť chyby prvého a druhého druhu permutačného testu porovnávajúceho dva nezávislé časové rady a porovnať výsledky s KS -testom dobrej zhody. Permutačný test je zovšeobecnením KS -testu dobrej zhody. Zovšeobecnenie spočíva v tom, že permutačný test porovnáva dva nezávislé časové rady, kým KS -test porovnáva dva nezávislé náhodné výbery. Jedna simulačná štúdia permutačného testu bola vykonaná v práci [2] (Diggle a Fisher, 1991). Zdalo sa nám však, že nebola urobená dostatočne podrobne. V práci chýbajú výsledky pre malé rozsahy výberov. Konkrétne boli preverené vlastnosti chýb prvého a druhého druhu pre rozsahy výberov $n = 64, 256, 1024$, čo je malý počet a nemá skoro žiadnu reálnu výpovednú hodnotu pre použitie testu v praxi. Ďalším cieľom je v prípade priaznivých výsledkov simulačnej štúdie aplikovať permutačný test na lekárske údaje. Ako alternatíva k permutačnému testu sa ponúka metóda tzv. CP -plotu. Ide o grafickú metódu, ktorá takisto porovnáva dva časové rady. V prípade, že simulačná štúdia nebude dávať uspokojivé výsledky použijeme metódu CP -plotu.

Pri popise časového radu pomocou regresného modelu sa snažíme modelovať okrem trendovej zložky aj periodickú zložku. Na modelovanie môžeme použiť kváziperiodický nelineárny regresný model. Medzi údaje vykazujúce periodické správanie patria napr. údaje lekárskeho alebo ekonomického charakteru. Naším cieľom bude aplikovať model na lekárske údaje. Trendovú zložku budeme modelovať pomocou lineárneho trendu a na odhalenie skrytých períód využijeme v prvej fáze periodogram skúmaného časového radu. Existuje Akaikeho informačné kritérium (AIC) a Fisherov test pomocou ktorých sa dajú vylúčiť z regresného modelu tie períódy, ktoré nie sú významné (druhá fáza určenia významných períód). Naším cieľom bude porovnať AIC a Fisherov test z hľadiska kvality výsledkov. Ako kritérium kvality si zvolíme reziduálnu sumu štvorcov modelu (RSS). To, či bude možné porovnávať AIC a Fisherov test sa rozhodneme na základe simulačnej štúdie Fisherovho testu. Doposiaľ nebola totiž urobená žiadna simulačná štúdia tohto testu v mne dostupnej literatúre.

Na záver zhrnieme naše ciele. Hlavným cieľom je preverenie vlastností chýb prvého a druhého druhu permutačného testu a Fisherovho testu pomocou simulačných štúdií a potom aplikovať tieto testy pri analýze časových radov lekárskeho údajov. Ďalším cieľom je urobiť aj simulačnú štúdiu vlastností chýb prvého a druhého druhu KS -testu a porovnať ich s permutačným testom.

1 Teoretické základy Kolmogorov-Smirnovho testu dobrej zhody

1.1 Úvod

Kolmogorov-Smirnov test dobrej zhody (v ďalšom používame skratku *KS*-test) je jedným zo základných testov porovnávajúcimi distribučné funkcie dvoch náhodných výberov. Je východiskom pre vznik iných testov založených na jeho princípe.

1.2 Základné pojmy

Definície pochádzajú z práce [10] (Zvára a Štěpán, 2002).

Definícia 1.2.1 Nech $F(\cdot)$ je distribučná funkcia. Nech X_1, X_2, \dots, X_n sú nezávislé náhodné veličiny, z ktorých každá má distribučnú funkciu $F(\cdot)$. Nech $x_1, x_2, \dots, x_n \in R$. Potom x_1, x_2, \dots, x_n nazývame náhodným výberom z distribučnej funkcie $F(\cdot)$.

Definícia 1.2.2 Majme náhodný výber x_1, x_2, \dots, x_n . Potom $F_n(x) = \frac{1}{n} \sum_i F(x_i)$, kde $F(x_i) = 1$ ak $x_i < x$, inak $F(x_i) = 0$, je empirická distribučná funkcia.

1.3 Kolmogorov-Smirnov test-princíp fungovania

Majme napr. realizácie x a y náhodných výberov X a Y . Môžeme sa pýtať, či tieto realizácie pochádzajú z toho istého rozdelenia pravdepodobnosti. Rozdelenie pravdepodobnosti náhodnej veličiny je reprezentované jej distribučnou funkciou. Zhodu rozdelenia pravdepodobnosti dvoch náhodných veličín chápeme teda ako zhodu ich distribučných funkcií. My však nevieme ako vyzerajú distribučné funkcie našich skúmaných náhodných veličín. Závery môžeme robiť len na základe odhadov empirických distribučných funkcií týchto náhodných veličín.

Ak porovnávame dve empirické distribučné funkcie, môžeme formulovať nasledovnú obojstrannú hypotézu

$$H_0 : F_n^X(x) = F_m^Y(x), \forall x,$$

$$H_1 : F_n^X(x) \neq F_m^Y(x)$$

pre aspoň jedno x .

Štatistika *KS*-testu dobrej zhody, pomocou ktorej overujeme danú hypotézu, je definovaná ako maximum rozdielu dvoch empirických distribučných funkcií [3] (Feller, 1948)

$$D = \max_{\forall x} |F_n^X(x) - F_m^Y(x)|,$$

kde X a Y predstavujú náhodné výbery, n a m sú rozsahy náhodných výberov.

Jednostranná hypotéza pre empirickú distribučnú funkciu a teoretickú distribučnú funkciu znie

$$H_0 : F_n^X(x) \geq F(x), \forall x,$$

$$H_1 : F_n^X(x) < F(x)$$

pre aspoň jedno x

alebo

$$H_0 : F_n^X(x) \leq F(x), \forall x,$$

$$H_1 : F_n^X(x) > F(x)$$

pre aspoň jedno x .

Tu má KS štatistika pomocou ktorej overujeme danú hypotézu nasledovný tvar

$$D^* = \max_{\forall x} |F_n^X(x) - F(x)|.$$

Z definície empirickej distribučnej funkcie je intuitívne zrejmé, že $F_n(x) \rightarrow F(x)$ pre $n \rightarrow \infty$. Teda čím väčší počet pozorovaní máme, tým by mal byť odhad distribučnej funkcie lepší.

Von Mises je autorom tzv. Von Misesho ω^2 -kritéria [3] (Von Misses, 1931 in Feller, 1948), ktoré poskytuje odhady pravdepodobnostných odchýliek $F_n(x)$ od $F(x)$ pre isté formy $F(x)$. Ešte so silnejším výsledkom prišiel A. Kolmogorov [3] (Kolmogorov, 1941 in Feller, 1948). Podarilo sa mu odvodiť rozdelenie štatistiky $D_n = \max |F_n^X(x) - F(x)|$. Vidíme, že rozdelenie štatistiky D_n je nezávislé na konkrétnej forme $F(x)$. Požadujeme iba, aby bola $F(x)$ spojitá. Presné rozdelenie štatistiky D_n nie je známe, ale Kolmogorov prišiel na to, že $n^{1/2}D_n$ má limitné rozdelenie.

Výsledok A. Kolmogorova je predmetom veľkého záujmu a skúmania v teórii neparametrického odhadovania.

Veta 1 *Nech $F(x)$ je spojitá a náhodná premenná D_n je definovaná ako*

$$D_n = \max_{\forall x} |F_n^X(x) - F(x)|.$$

Potom pre každé pevné $z \geq 0$ pre $n \rightarrow \infty$ platí

$$\Pr(D_n \leq zn^{1/2}) \rightarrow L(z),$$

kde $L(z)$ je kumulatívna distribučná funkcia, ktorá pre $z > 0$ je daná nasledujúcimi ekvivalentnými vťahmi

$$L(z) = 1 - 2 \sum_{\nu=1}^{\infty} (-1)^{\nu-1} e^{-\nu^2 z^2} = (2\pi)^{1/2} z^{-1} \sum_{\nu=1}^{\infty} e^{-(2\nu-1)^2 \pi^2 / 8z^2}.$$

Pre $z \leq 0$ máme $L(z) = 0$.

Rovnaké limitné rozdelenie má aj štatistika $\max_{\forall x} |F_n^X(x) - F_m^Y(x)|$, toto rozdelenie odvodil Smirnov [3] (Smirnov, 1939 in Feller, 1948).

Veta 2 *Nech (x_1, x_2, \dots, x_m) a (y_1, y_2, \dots, y_n) sú dva náhodné výbery s distribučnou funkciou $F(\cdot)$. Nech $F_m^X(x)$ a $F_n^Y(x)$ sú príslušné empirické distribučné funkcie a definujme novú náhodnú premennú $D_{m,n}$ ako*

$$D_{m,n} = \max_{\forall x} |F_m^X(x) - F_n^Y(x)|.$$

Nech

$$N = \frac{mn}{m+n}$$

a nech pre $m \rightarrow \infty, n \rightarrow \infty$ platí

$$\frac{m}{n} \rightarrow a,$$

kde a je konštanta. Potom pre každé fixné $z \geq 0$

$$\Pr(D_{m,n} \leq zN^{1/2}) \rightarrow L(z),$$

kde $L(z)$ je rovné

$$L(z) = 1 - 2 \sum_{\nu=1}^{\infty} (-1)^{\nu-1} e^{-\nu^2 z^2} = (2\pi)^{1/2} z^{-1} \sum_{\nu=1}^{\infty} e^{-(2\nu-1)^2 \pi^2 / 8z^2}.$$

Vidíme, že základom pre KS testovacie štatistiky sú rozdiely medzi napr. empirickými distribučnými funkciami z toho istého rozdelenia. Prvý krok je teda vytvoriť k porovnávaným skupinám pozorovaní ich empirické distribučné funkcie. Potom vypočítame $D = \max |F_n^X(x) - F_m^Y(x)|$ a na rozhodovanie použijeme približnú kritickú hodnotu

$$D_{n,m}(\alpha) = \left(\frac{n+m}{2nm} \log \frac{2}{\alpha} \right)^{\frac{1}{2}}.$$

Ak $D > D_{n,m}(\alpha)$, zamietame hypotézu H_0 [4] (Katina, 2006).

Pôvodné dôkazy Kolmogorova a Smirnova sú veľmi zložité a sú založené na úplne odlišných metódach [3] (Feller, 1948). Kolmogorov dôkaz je založený na pomocnom tvrdení [3] (Kolmogorov, 1941 in Feller, 1948). Smirnov dokázal obe vety ako dôsledky

inej vety, hovoriacej o počte prienikov grafov $F_n^X(x)$ a $F(x) \pm \epsilon N^{-\frac{1}{2}}$, resp. $F_m^X(x)$ a $F_m^X(x) \pm \epsilon N^{-\frac{1}{2}}$, kde ϵ je náhodná premenná predstavujúca rozdiely medzi danými distribučnými funkciami [3] (Smirnov, 1939 in Feller, 1948).

S jednoduchším, hoci nie jednoduchým dôkazom prišiel W. Feller (1948) [3]. Podáva dôkaz oboch viet založený na rovnakých metódach.

2 Permutačný test

2.1 Úvod

Existuje niekoľko testov pre porovnanie dvoch nezávislých náhodných výberov. Jedným z nich je aj *KS*-test dobrej zhody predstavený v predošlej kapitole. Ak naše údaje tvoria časový rad, pozorovania už nie sú nezávislé a teda netvoria náhodný výber. Permutačný test predstavuje zovšeobecnenie *KS*-testu. Zistujeme, či dva nezávislé časové rady pochádzajú z toho istého náhodného procesu alebo nie. Zovšeobecnenie spočíva v tom, že namiesto náhodného výberu uvažujeme časový rad. Musí ísť ale o dva nezávislé časové rady. V tomto zmysle musí byť nezávislosť zachovaná.

2.2 Základné pojmy

Definície pochádzajú z prác [8] (Štulajter, 1990) a [2] (Diggle a Fisher, 1991).

Definícia 2.2.1 *Nech (Ω, F, P) je ľubovoľný pravdepodobnostný priestor a nech T je ľubovoľná množina - podmnožina E^1 . Nech $X(t)$ je pre každé $t \in T$ náhodná veličina definovaná na pravdepodobnostnom priestore (Ω, F, P) . Potom množinu $X = X(t); t \in T$ náhodných premenných $X(t)$ nazývame náhodný proces. Parameter t vystupujúci v definícii náhodného procesu budeme nazývať čas.*

Definícia 2.2.2 *Náhodný proces $X = \{X(t); t \in T\}$ nazývame stacionárny, ak jeho stredná hodnota je konštantná, teda ak platí $m(t) = E[X(t)] = c$ pre všetky $t \in T$, kde c je nejaké reálne číslo a jeho kovariančná funkcia závisí len od vzdialenosti argumentov, $\gamma_x(k) = \text{cov}\{X(t), X(t-k)\}$.*

Definícia 2.2.3 *Hovoríme, že náhodný proces $X = \{X(t); t \in T\}$ spĺňa lineárny regresný model, ak pre jeho strednú hodnotu $m(t) = E[X(t)]; t \in T$ platí:*

$$m(t) = \sum_{i=1}^k \beta_i f_i(t),$$

kde f_1, f_2, \dots, f_k sú známe funkcie definované na T a vektor $\beta = (\beta_1, \beta_2, \dots, \beta_k)^T$ je neznámy vektor regresných parametrov.

Definícia 2.2.4 *Funkciu $\lambda_x(\omega) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \gamma_x(k) e^{-ik\omega}$, $-\pi < \omega < \pi$ nazývame spektrálnou hustotou procesu $X = \{X(t), t \in T\}$.*

Definícia 2.2.5 Reálna funkcia $I_x(\omega)$; $\omega \in [-\pi, -\pi]$ definovaná predpisom

$$I_x(\omega) = \frac{1}{2\pi n} \left\{ \left[\sum_{t=1}^n x_t \cos(t\omega) \right]^2 + \left[\sum_{t=1}^n x_t \sin(t\omega) \right]^2 \right\},$$

sa nazýva periodogram pozorovania x_t , kde $\omega = \omega_j = \frac{2\pi j}{n}$, $j = 1, \dots, m$; $m = \frac{n-1}{2}$.

Veta 3 Ak je $\{X_t\}$ lineárny stacionárny proces, platí pre $n \rightarrow \infty$

$$I_x(\omega) \approx \frac{1}{2} \lambda_x(\omega) Z_2,$$

$0 < \omega < \pi$, kde Z_2 je náhodná veličina s χ^2 rozdelením pravdepodobnosti s 2 stupňami voľnosti a $I_x(\omega_j)$ a $I_x(\omega_k)$ sú asymptoticky nezávislé pre $j \neq k$.

2.3 Popis testu

Majme dva nezávislé časové rady $\{x_t, t = 1, \dots, n\}$ a $\{y_t, t = 1, \dots, n\}$, generované stacionárnymi procesmi $\{X_t\}$ a $\{Y_t\}$.

Rozdelenie pravdepodobností procesu je dané jeho spektrálnou hustotou. Hypotézu zhodného rozdelenia dvoch časových radov formulujeme ako zhodu ich spektrálnych hustôt nasledovne

$$H_0 : \lambda_x(\omega) = \lambda_y(\omega), 0 < \omega < \pi,$$

$$H_1 : \lambda_x(\omega) \neq \lambda_y(\omega), 0 < \omega < \pi.$$

Permutačný test je vlastne zovšeobecnením KS -testu dobrej zhody, teda testovacia štatistika bude tvaru [2] (Diggle a Fisher, 1991)

$$D_m = \max_{1 \leq j \leq m} |F_x(\omega_j) - F_y(\omega_j)| = d_1(\omega_1, \dots, \omega_m),$$

kde

$$F_x(\omega_j) = \frac{\sum_{i=1}^j I_x(\omega_i)}{\sum_{i=1}^m I_x(\omega_i)}$$

je normovaný kumulatívny periodogram (NCP) časového radu $\{x_t, t = 1, \dots, n\}$ (rovnako je definovaný aj $F_y(\omega_j)$).

Je dôležité mať na pamäti, že prvky periodogramu $I_x(\omega_1), I_x(\omega_2), \dots, I_x(\omega_m)$ sú nezávislé iba asymptoticky, to isté platí aj pre periodogram $I_y(\omega_j)$. To je hlavný rozdiel medzi KS -testom a permutačným testom. U KS -testu pochádzajú pozorovania z náhodného výberu a sú teda nezávislé.

Za predpokladu, že platí hypotéza H_0 , je rozdelenie pravdpodobnosti KS štatistiky D_m invariantné ku všetkým 2^m vzájomným zámenám $I_x(\omega_j)$ a $I_y(\omega_j)$, $j = 1, \dots, m$. V praxi vypočítame d nie pre všetkých 2^m zámien, ale iba pre ich dostatočne veľký počet. Vypočítané štatistiky označíme d_2, \dots, d_s , s je dostatočne veľké a vrátane štatistiky d_1 ich zoradíme od najmenej po najväčšiu, teda dostaneme $d_1 \leq d_2 \leq \dots \leq d_s$. P -hodnotu vypočítame ako

$$P = 1 - \left(\frac{r-1}{s+1} \right),$$

kde r je poradie štatistiky d_1 a s je počet všetkých štatistík.

2.4 Povaha testu

Všimnime si, že test je navrhnutý tak, že dokáže odhaliť iba prípadné odlišnosti v priebehu porovnávaných časových radov a nie absolútne rozdiely medzi nimi. Je tomu tak preto, že základ KS testovacej štatistiky tvorí normovaný kumulatívny periodogram a jeho hodnoty sa pohybujú vždy medzi nulou a jednotkou. Ak by porovnávané časové rady tvorili napríklad dva biele šumy s rozdielnym rozptylom, test by ich nedokázal rozlíšiť. Takto koncipovaný test môže mať široké uplatnenie v medicíne ale aj v ekonómii. Napríklad môžeme porovnávať ekonomické cykly Slovenska a Nemecka. Môžeme sa pýtať, či ekonomika na Slovensku a Nemecku zažíva súčasne boom alebo recesiú.

3 Teoretické základy Fisherovho testu

Základné pojmy sú prevzaté z práce [1] (Brockwell a Davis, 1991).

Fisherov test využíva periodogram a je postavený na dôsledku nasledovnej vety

Veta 4 Ak X_t je gaussovský biely šum, potom náhodné premenné

$$Y_i = \frac{\sum_{k=1}^i I(\lambda_k)}{\sum_{k=1}^q I(\lambda_k)},$$

$i = 1, \dots, q-1$, $q = \lfloor (n-1)/2 \rfloor$, sú nezávislé náhodné premenné a každá z nich je rovnomerne rozdelená na intervale $[0, 1]$

Dôsledok tejto vety môžeme formulovať ako nasledujúcu vetu.

Veta 5 Definujme $Y_0 := 0$, $Y_q := 1$ a

$$M_q := \max_{1 < i < q} (Y_i - Y_{i-1}) = \max_{1 < i < q} \frac{I(\lambda_i)}{\sum_{i=1}^q I(\lambda_i)}.$$

Potom

$$P(M_q \leq a) = \sum_{j=1}^q (-1)^j \binom{q}{j} (1 - ja)_+^{q-1},$$

kde $x_+ = \max(x, 0)$.

Pomocou tohto dôsledku môžeme formulovať nasledovnú hypotézu:

H_0 znie, že X_t je biely šum a alternatívna hypotéza H_1 znie, že X_t obsahuje periodický komponent neznámej frekvencie. Hypotézu H_0 zamietame ak periodogram X_t obsahuje hodnotu, ktorá je značne väčšia ako priemerná hodnota, teda ak hodnota

$$\xi_q = qM_q$$

je dostatočne veľká. Ak pravdepodobnosť

$$P(\xi \geq x) = 1 - \sum_{j=1}^q (-1)^j \binom{q}{j} (1 - jx)_+^{q-1}$$

je menšia ako α , tak zamietame hypotézu H_0 na hladine významnosti α .

4 Simulačné štúdie

4.1 Úvod

V nasledujúcich dvoch podkapitolách sa budeme zaoberať chybami prvého a druhého druhu KS -testu a permutačného testu. Tieto chyby nám hovoria o celkovej kvalite a teda použiteľnosti testu. Pokúsime sa porovnať výsledky pre tieto dva testy.

Chyby prvého a druhého druhu zisťujeme pomocou testovania hypotéz a preto si pripomenieme niektoré základné pojmy z teórie testovania hypotéz.

4.2 Základné pojmy

Základné pojmy sú prevzaté z práce [5] (Katina, 2006).

Nulová hypotéza $H_0 : \theta \in \Theta_0$, alternatívna hypotéza $H_1 : \theta \in \Theta_1$, $\Theta_0 + \Theta_1 = \Theta$, kde θ je parameter a Θ je výberový priestor.

Stretávame sa s nasledovnými štyrmi možnosťami: A) platí H_0 a naše rozhodnutie je prijať H_0 (správne), B) platí H_0 a naše rozhodnutie je zamietnuť H_0 (nesprávne), C) platí H_1 a naše rozhodnutie je prijať H_0 (nesprávne), D) platí H_1 a naše rozhodnutie je zamietnuť H_0 (správne).

V prípade A, D je naše rozhodnutie správne ($Pr_\theta(A) > 1 - \alpha$). V prípade B sa dopúšťame chyby prvého druhu ($CHPD, Pr_\theta(CHPD) \leq \alpha$), kde α je hladina významnosti testu, ktorú si na začiatku volíme. V prípade C sa dopúšťame chyby druhého druhu ($CHDD, Pr_\theta(CHDD) = \beta$). Doplnok ku $Pr_\theta(CHDD)$ je pravdepodobnosť $1 - \beta$ a nazýva sa sila testu (pri alternatíve H_1). Sila testu teda predstavuje pravdepodobnosť s akou zamietame hypotézu H_0 , ak H_0 neplatí (D). Silu testu môžeme znázorniť pomocou silofunkcie testu $1 - \beta(\theta) = Pr_\theta(X \in W)$, kde W je kritický obor testu. Ak náhodná premenná X padne do W , zamietame H_0 . θ je parameter, ktorý sa pri každom teste špecifikuje zvlášť. Pri KS -teste je parametrom silofunkcie rozdiel stredných hodnôt dvojice pozorovaní, pri permutačnom teste plní funkciu parametra rozsah výberov.

4.3 Simulačná štúdia Kolmogorov-Smirnov testu

Chyba prvého druhu

Na určenie chyby prvého druhu budeme simulovať dvojicu údajov z normálneho rozdelenia s identickou strednou hodnotou μ a rozptylom σ^2 . Ideme teda zisťovať, s akou pravdepodobnosťou sa KS -test pomýli a prehlási danú dvojicu realizácií náhodných výberov za dvojicu, ktorá nepochádza z toho istého rozdelenia pravdepodobnosti, ak z

neho v skutočnosti pochádza. Želáme si aby chyba prvého druhu bola čo najmenšia. Hypotéza znie

$$H_0 : F_n^X(x) = F_n^Y(x), \forall x,$$

$$H_1 : F_n^X(x) \neq F_n^Y(x)$$

pre aspoň jedno x .

Simulujeme $k = 1000$ -krát dvojicu údajov z normálneho rozdelenia s identickou strednou hodnotou μ a rozptylom σ^2 . Spočítame počet P-hodnôt menších ako 0,05 a vydelíme k . Dostaneme tak odhad chyby prvého druhu. Výpočet opakujeme pre rôzne rozsahy náhodných výberov $n = 10, 20, 30, 40, 50$.

Dochádzame k záveru, že chyba prvého druhu nezávisí na rozsahu náhodného výberu. Pohybuje sa medzi 0,02 a 0,04.

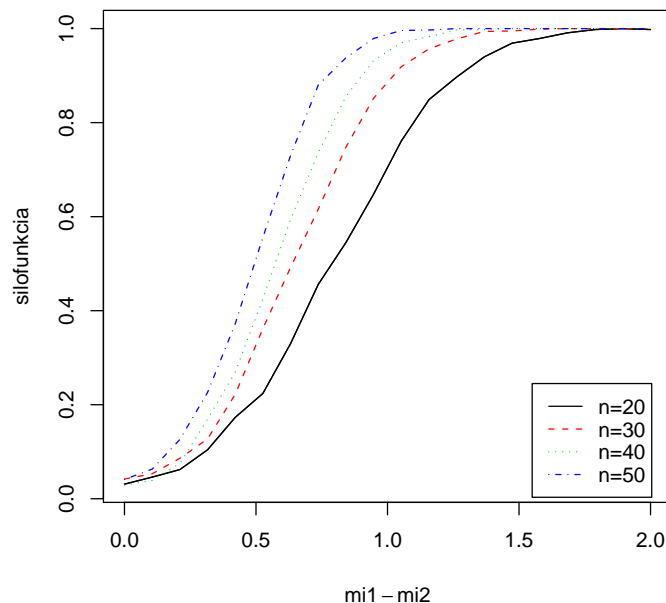
Sila testu

Majme tie isté hypotézy ako pri určovaní vlastností chýb prvého druhu. Silu testu preveríme tak, že naša dvojica náhodných výberov sa bude líšiť v stredných hodnotách μ_1 a μ_2 ($\sigma_1^2 = \sigma_2^2 = \sigma^2$). Postupne zväčšujeme rozdiel stredných hodnôt $\mu_1 - \mu_2$ pre fixný rozsah náhodných výberov.

Presný postup je taký, že rozdiel v stredných hodnotách je $\mu_1 - \mu_2 = (0, 0.2, 0.4, \dots, 2)$ pričom $\mu_2 = 0$ a fixné rozsahy náhodných výberov sú postupne $n = 20, 30, 40, 50$.

Vieme, že hypotéza H_0 určite neplatí, pretože naša dvojica náhodných výberov sa líši v stredných hodnotách. Zisťujeme s akou pravdepodobnosťou zamietne KS -test hypotézu H_0 ak neplatí. Želáme si, aby bola sila testu čo najväčšia a teda chyba druhého druhu čo najmenšia. Výpočet prebieha tak isto ako pri zisťovaní chyby prvého druhu.

Ukazuje sa, že na silu testu vplývajú dva parametre. Prvým parametrom je rozsah náhodného výberu. So zväčšujúcim sa rozsahom náhodného výberu rastie aj sila testu. Druhým parametrom je rozdiel stredných hodnôt náhodných výberov. Pri narastajúcom rozdieli stredných hodnôt sa sila testu zvyšuje. Silofunkcie testu sú zachytené na obrázku č. 1. Na obrázku je znázornená iba polovica intervalu, druhá polovica je párovo identická, keďže ide o obojstrannú hypotézu. Z tohoto obrázku je zrejmé, že rozdiely medzi silofunkciami sú veľmi malé pre $n \geq 50$.



Obrázok č. 1: Silofunkcia KS -testu.

4.4 Simulačná štúdia permutačného testu

Diggle a Fisher(1991) [2] prezentujú výsledky simulačnej štúdie, ktorá sa zaoberala chybami prvého a druhého druhu nimi navrhnutého testu KS -typu. Simulované boli procesy $Ar(p)$ a $Ma(q)$. Dochádzajú k záverom, že test môže konkurovať semiparametrickým a parametrickým testom pre niektoré dvojice procesov. Chyby prvého a druhého druhu preverovali na simulovaných náhodných výberoch dĺžky $n = 64, 256$ a 1024 . Z ich štúdie nie je zrejmé, aký je minimálny rozsah výberu, aby mal test ešte dobré vlastnosti (chyba prvého druhu, chyba druhého druhu). Autori aplikujú test na reálne lekárske údaje o rozsahu $n = 44$, čo je menej ako $n = 64$, najkratší rozsah, pre ktorý skúmali chyby prvého a druhého druhu tohoto ich KS -testu. To nás viedlo k podrobnejšiemu prevereniu chýb prvého a druhého druhu testu.

Chyba prvého druhu

Definujme hypotézu nasledovne

$$H_0 : \lambda_x(\omega) = \lambda_y(\omega), 0 < \omega < \pi,$$

$$H_1 : \lambda_x(\omega) \neq \lambda_y(\omega), 0 < \omega < \pi,$$

kde $\lambda(\omega)$ boli definované v *def.2.2.3*.

Simulujeme $k=1000$ -krát napr. dvojicu $Ar(1)$ procesov, aby sme dostali vždy dvojice nezávislých realizácií. Simulácie robíme pre rôzne dĺžky dvojíc procesov. Jednotlivé dĺžky dvojíc procesov sú $n = 20, 30, \dots, 100, 120, \dots, 200, 300, \dots, 800$ (po $k = 1000$ simulácií). Pre každú z týchto dvojíc robíme $s = 100$ permutácií a vypočítame príslušnú P -hodnotu. Zvolíme si hladinu významnosti $\alpha = 0,05$ a spočítame počet tých P -hodnôt, ktoré sú menšie ako $0,05$ a vydáme tento počet počtom simulácií k . Dostaneme tak odhad chyby prvého druhu.

Pre dvojice zhodných procesov nám vyšlo, že bez ohľadu na rozsah výberov je pravdepodobnosť zamietnutia hypotézy približne $0,001$. Test teda prakticky nerobí chybu prvého druhu. Podobne sa správa aj KS -test, taktiež nezáleží na rozsahu výberov, len pravdepodobnosť nastatia je vyššia (pozri Kap.3.3).

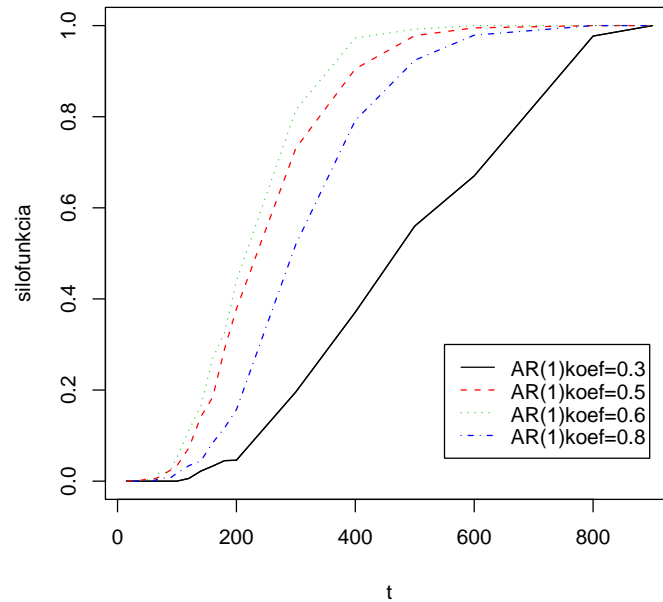
Sila testu

Majme tie isté hypotézy ako pri zisťovaní vlastností chyby prvého druhu. Zisťovanie sily testu prebieha presne tak ako pri zisťovaní chyby prvého druhu s tým rozdielom, že simulujeme dvojicu rôznych procesov, napr. $Ar(1)$ a biely šum.

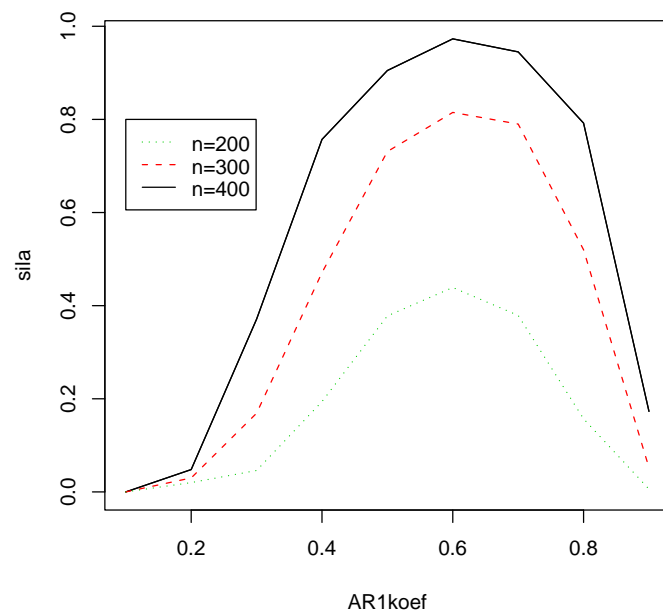
Záver je, že test má dobré vlastnosti len pre veľké rozsahy údajov vo všeobecnosti okolo $n = 500$ a viac.

Pre korelačný koeficient veľkosti 0.1 a 0.2 nie je test schopný rozlíšiť $Ar(1)$ proces od bieleho šumu ani pre veľké rozsahy výberov. Silofunkcie testu sú znázornené na obrázku č. 2 pre $Ar(1)$ proces s rôznymi korelačnými koeficientami. Z tohto obrázku je vidno, že pri korelačnom koeficiente rovnom 0.6 je silofunkcia "najstrmšia" pre ostatné korelačné koeficienty sa strmosť silofunkcie znižuje na obe strany od korelačného koeficientu rovného 0.6 . Teda napr. pre korelačný koeficient rovný 0.6 je silofunkcia strmšia ako pre korelačný koeficient rovný 0.8 , čo je nečakané. Vidíme, že vzťah medzi korelačným koeficientom $Ar(1)$ procesu a silou testu je prekvapujúco nelineárny. Teda s narastajúcim korelačným koeficientom najprv sila testu rastie, svoje maximum dosahuje pre korelačný koeficient rovný 0.6 a pre korelačné koeficienty väčšie ako 0.6 zase sila testu klesá. Tento vzťah je zrejmy už z obrázka č. 2 a je zobrazený na obrázku č. 3, kde fixujeme rozsah výberov a postupne zväčšujeme korelačný koeficient.

V porovnaní s KS -testom dobrej zhody sa správa permutačný test odlišne. Pri KS -teste dobrej zhody pri určovaní sily testu sme postupne zväčšovali rozdiely medzi strednými hodnotami náhodných výberov s tým, že sila testu nám rástla. Pri permutačnom teste robíme čosi podobné, keď postupne zväčšujeme korelačný koeficient $Ar(1)$ procesu a tým sa akoby "vzdľahujeme od bieleho šumu. Tu nám ale sila testu najprv rastie a neskôr začína klesať.



Obrázok č. 2: Silofunkcia permutačného testu pre rôzne rozsahy výberov pre $AR(1)$ procesy s rôznym korelačným koeficientom

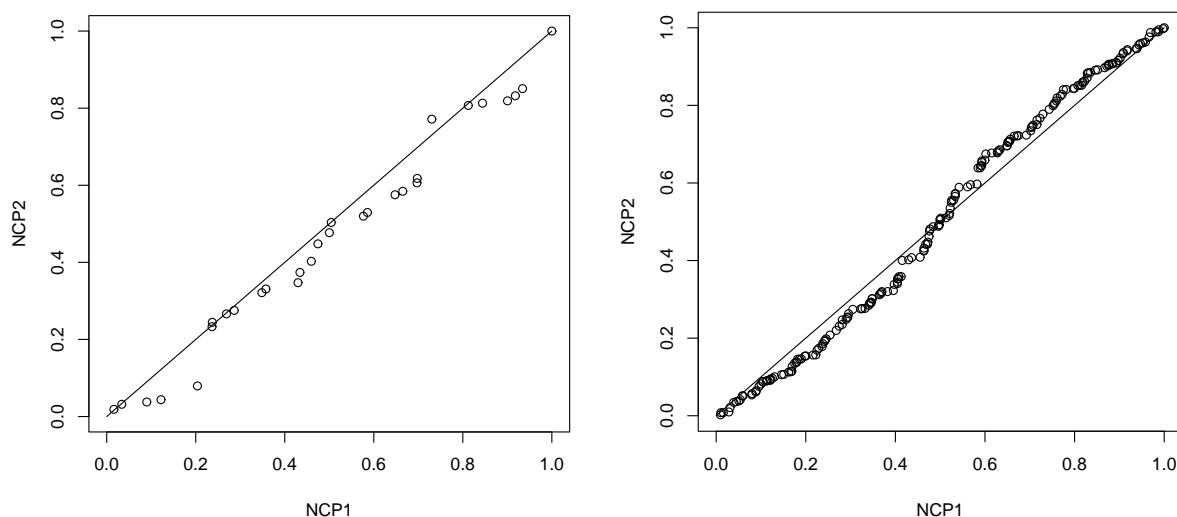


Obrázok č. 3: Silofunkcia permutačného testu pre fixný rozsah výberov a meniaci sa korelačný koeficient $AR(1)$ procesu

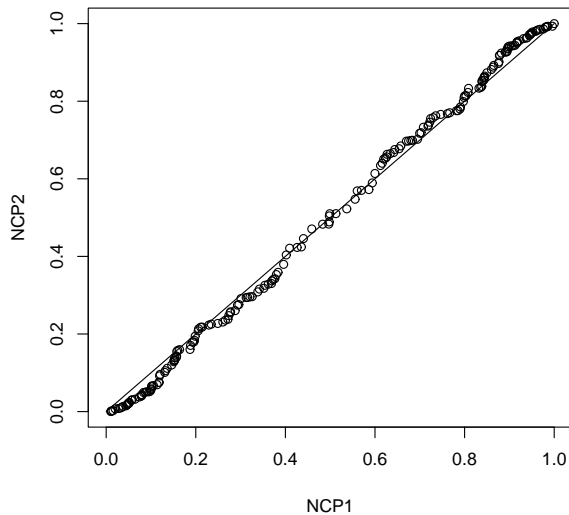
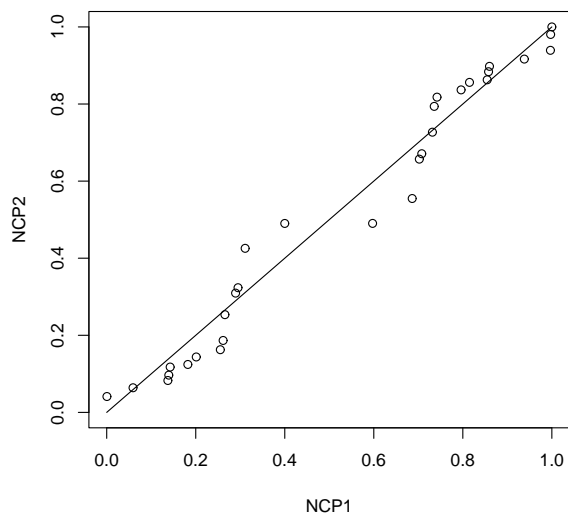
Poznámka 4.4.1 Všimnime si, že sme pri chybe prvého druhu ani pri sile testu neuvažovali rozptyl σ^2 náhodného procesu. Z teórie totiž vyplýva, že permutačný test je schopný zachytávať iba odlišnosti v priebehu daných časových radov. Teóriu podporuje aj výsledok simulčnej štúdie, kde sme simulovali dvojicu procesov presne tak isto ako pri zisťovaní vlastností chyby prvého druhu s tým rozdielom, že sa dané $Ar(1)$ procesy líšili v rozptyle. Rozptyl jedného procesu sme zvolili $\sigma^2 = 1$ a rozptyl druhého procesu sme zvolili $\sigma^2 = 2$. Teda meriame vlastne silu testu. Simulácia preukázala, že permutačný test nedokáže rozlíšiť časové rady ak sa odlišujú iba v rozptyle. Sila testu vyšla bez ohľadu na rozsah výberov 0.

Existuje aj menej formálny spôsob, ako overiť, či dvojica časových radov pochádza z toho istého stochastického procesu alebo nie. Pred testovaním môžeme urobiť tzv. *CP*-plot. Na os x nanesieme normovaný kumulatívny periódogram (*NCP*) jedného časového radu a na os y *NCP* druhého časového radu. Ak dostaneme graf v tvare $y = x$, znamená to, že časové rady pochádzajú z toho istého stochastického procesu. Ak sa rady líšia, bude sa graf vychýľovať od funkcie $y = x$.

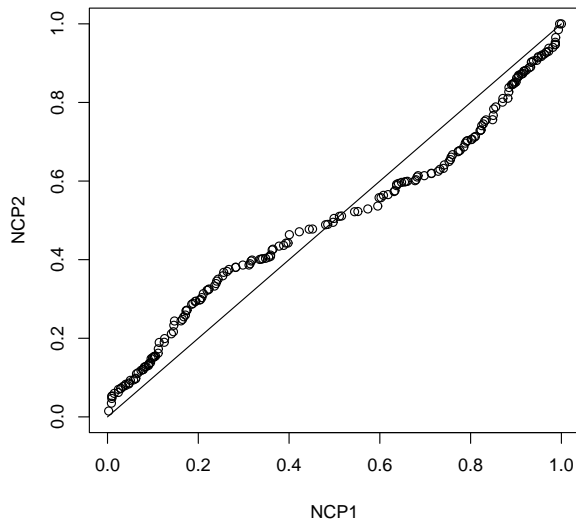
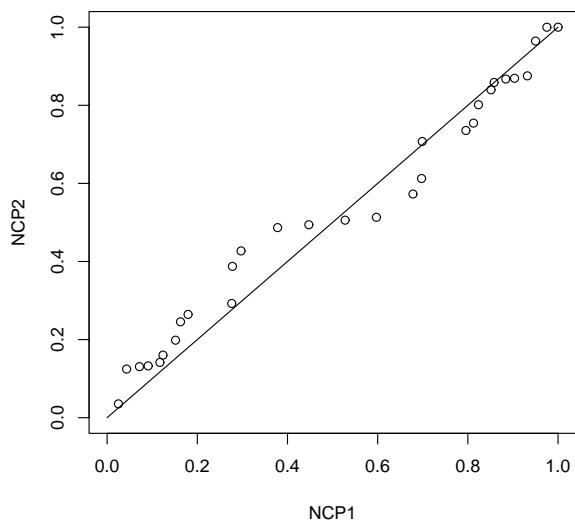
CP-plot je vhodný najmä v počiatočnej fáze analýzy dát. Na ilustráciu uvádzame niekoľko *CP*-plotov (obrázky č. 4 - 11).



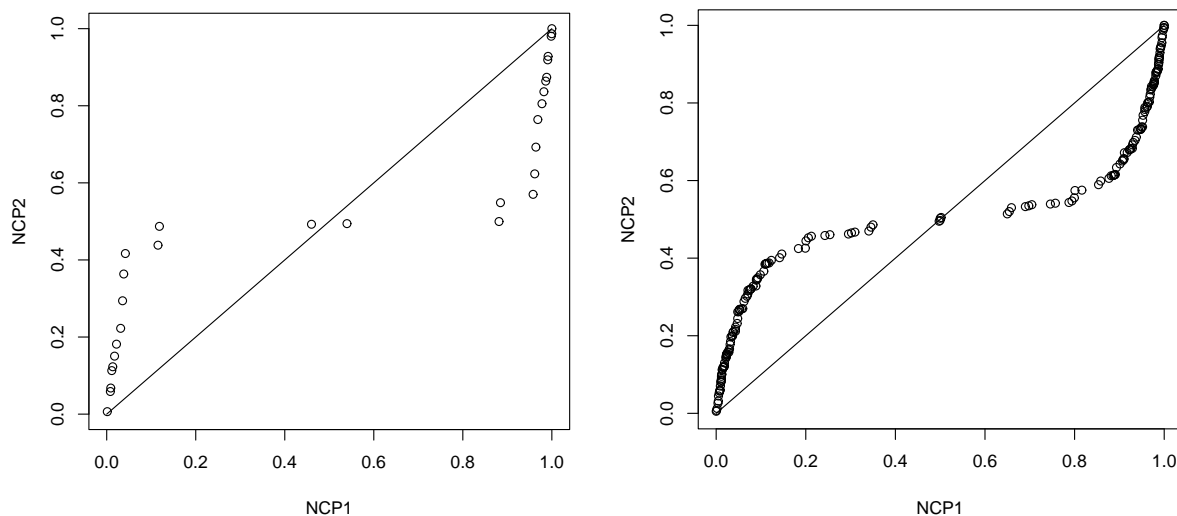
Obrázok č. 4 a 5: *CP*-plot dvojice bielych šumov s rozsahom výberov $n = 30$, resp. $n = 200$



Obrázok č. 6 a 7: *CP*-plot dvojice $AR(1)$ procesov s korelačným koeficientom 0.3 s rozsahom výberov $n = 30$, resp. $n = 200$



Obrázok č. 8 a 9: *CP*-plot dvojice $AR(1)$ procesu s korelačným koeficientom 0.3 a bieleho šumu s rozsahom výberov $n = 30$, resp. $n = 200$



Obrázok č. 10 a 11: CP -plot dvojice $AR(1)$ procesu s korelačným koeficientom 0.8 a bieleho šumu s rozsahom výberov $n = 30$, resp. $n = 200$

Prvú dvojicu tvoria časové rady bieleho šumu (obrázky č. 4 a č. 5). Vidíme, že CP -plot je dobrým ukazovateľom zhodnosti časových radov. Plot je tvaru $y = x$ pre výber rozsahu $n = 30$, ako aj pre $n = 200$, čo znázorňuje už asymptotické správanie.

Druhú dvojicu tvoria časové rady $Ar(1)$ procesu s korelačným koeficientom 0.3 (obrázky č. 6 a č. 7). Pre výber rozsahu $n = 30$ by sme sa na základe obrázka č. 6 mohli domnievať, že ide o dva rôzne časové rady. Graf má tvar esovitej krivky teda nie je tvaru $y = x$. Asymptotické správanie je na rozdiel od rozsahu výberu $n = 30$, lepšie (obrázok č. 7).

Tretiu dvojicu tvoria dva rôzne časové rady, $Ar(1)$ proces s korelačným koeficientom 0.3 a biely šum (obrázky č. 8 a č. 9). CP -plot v tomto prípade iba mierne naznačuje odlišnosť časových radov (obrázok č. 8). Rozdiel je vidno iba asymptoticky (obrázok č. 9).

Poslednú dvojicu tvoria dva rôzne časové rady, $Ar(1)$ proces s korelačným koeficientom 0.8 a biely šum (obrázky č. 10 a č. 11). Tu je CP -plot dobrým ukazovateľom odlišnosti časových radov.

Záver

Je vidno, že rozhodovať iba na základe CP -plotu sa nedá. CP -plot môže slúžiť ako pomôcka, niekedy môže aj zavádzať, pre malý rozsah výberov sa CP -plot správa nepredvídateľne. (obrázky č. 6 a č. 7 a obrázky č. 6 a č. 8). Na základe obrázkov č. 8 - č. 11

by sme sa mohli domnievať, že so zväčšujúcim sa korelačným koeficientom $Ar(1)$ procesu sa CP -plot odchyľuje od grafu tvaru $y = x$ čoraz viac.

Môžeme povedať, že pre samotný CP -plot platí podobný záver ako pre permutačný test. Teda, že pre malý rozsah výberov je táto metóda nespoľahlivá.

4.5 Simulačná štúdia Fisherovho testu

Chyba prvého druhu

Hypotéza znie

H_0 : náhodný proces X_t je biely šum,

H_1 : náhodný proces X_t obsahuje periodický člen neznámej periódy.

Simulujeme $k = 1000$ -krát biely šum. Simulácie robíme pre biely šum postupne pre rozsahy $n = 10, 20, 30, \dots, 300$ (po $k = 1000$ simulácií pre každý rozsah). Zvolíme si hladinu významnosti $\alpha = 0.05$ a vypočítame počet tých P -hodnôt, ktoré sú menšie ako $\alpha = 0.05$ a vydelíme tento počet počtom simulácií k . Dostaneme tak odhad chyby prvého druhu.

Dochádzame k záveru, že Fisherov test prakticky nerobí chybu prvého druhu. Nezávisle na veľkosti rozsahu výberu je chyba prvého druhu rovná približne 0.001. Rovnako spoľahlivo sa správa aj permutačný test (pozri Kap.3.4).

Sila testu

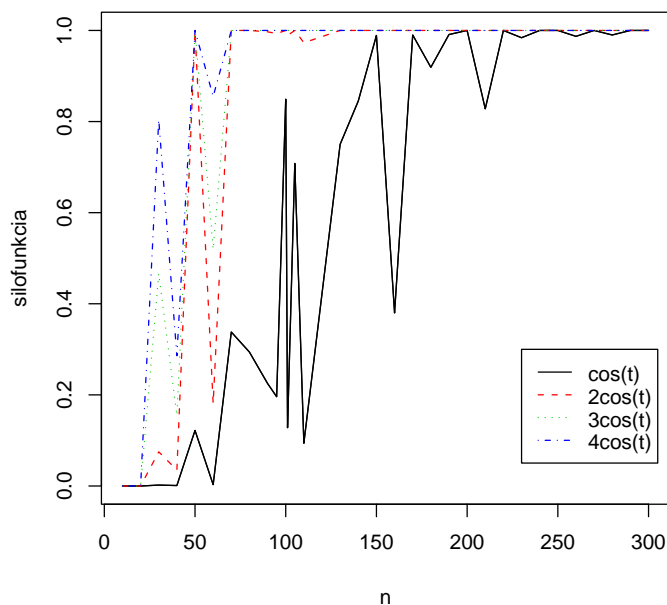
Majme tie isté hypotézy ako pri zisťovaní vlastností chyby prvého druhu. Budeme simulovať $k = 1000$ -krát náhodný proces pozostávajúci z rôznych zmesí kosínusov a sínusov. Simulácie robíme pre rozsahy výberov $n = 10, 20, 30, \dots, 300$ a ešte aj pre $n = 95, 101$ kvôli lepšej názornosti (po $k = 1000$ simulácií pre každý rozsah). Teraz si rozoberieme konkrétne skúmané prípady.

A) Simulujeme náhodný proces $\{X_t\}$, $X_t = K \cos(t) + Z_t$, kde Z_t je biely šum a K je konštanta.

Postupujeme tak, že pre $K = 1, 2, 3, 4$ simulujeme $k = 1000$ -krát proces $\{X_t\}$ pre každé K zvlášť. Zvolíme si hladinu významnosti $\alpha = 0.05$ a vypočítame počet tých P -hodnôt, ktoré sú menšie ako $\alpha = 0.05$ a vydelíme tento počet počtom simulácií k . Výsledok simulácie je znázornený na obrázku č. 12. Z tohto obrázka je zrejmé, že silofunkcia Fisherovho testu nie je neklesajúca na celom intervale rozsahu výberu. Neklesajúca je iba na istých podintervaloch rozsahov výberov.

Vidíme, že na silu testu vplývajú dva parametre. Jedným parametrom je rozsah výberu. So zväčšujúcim sa rozsahom výberu sa "výkyvy" v hodnotách silofunkcie znižujú. Pre rozsahy výberov nad $n = 200$ je sila testu bez výkyvov pre všetky skúmané K a je dostatočne veľká.

Druhým parametrom je konštanta K . Je vidno, že pre fixný rozsah výberov so zväčšujúcim sa K sila testu rastie pre všetky fixné rozsahy výberov. Dôvod prečo hrá konštanta K takú významnú úlohu pri sile testu je ten, že Fisherov test zamietá hypotézu H_0 , ak je maximálna hodnota periodogramu daného pozorovania značne väčšia ako priemerná hodnota periodogramu. So zväčšujúcim sa K nám tento rozdiel narastá a preto sila testu rastie.



Obrázok č. 12: Silofunkcia Fisherovho testu pre náhodný proces $X_t = K \cos(t) + Z_t$, $K = 1, 2, 3, 4$

B) Simulujeme náhodné procesy X_t^1, X_t^2, X_t^3 , kde

$$X_t^1 = \cos(t) + \cos(2t) + \sin(t) + \sin(2t) + Z_t,$$

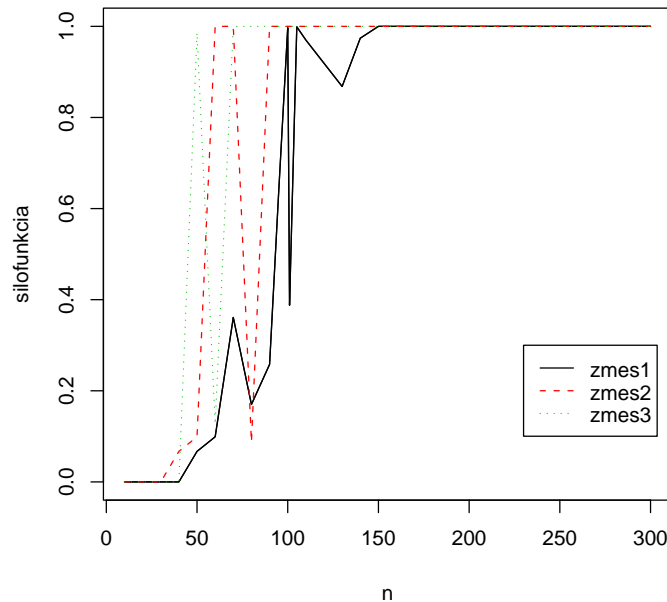
$$X_t^2 = \cos(t) + 2 \cos(2t) + 2 \sin(t) + 3 \sin(2t) + Z_t,$$

$$X_t^3 = 3 \cos(t) + 2 \cos(2t) + 3 \sin(t) + 3 \sin(2t) + Z_t,$$

Z_t je biely šum.

Simulácia prebieha presne tak ako v prípade A. Výsledok simulácie je zobrazený na obrázku č. 13. Z tohto obrázka je vidno, že v porovnaní s prípadom A dochádza k

pretínaniu jednotlivých siločiar napriek tomu, že koeficienty pri kosínusoch a sínusoch narastajú pri skúmaných procesoch (u procese X_t^1 sú najmenšie a u procese X_t^3 najväčšie). Teda sila testu je napr. pre rozsah výberu $n = 70$ väčšia pre proces X_t^1 ako pre proces X_t^3 , pritom proces X_t^3 má pri kosínusoch a sínusoch väčšie koeficienty ako proces X_t^1 . Ďalej je badateľné, že rozdiely v sile testu nie sú také veľké v porovnaní s prípadom A, keď zmes kosínusov a sínusov tvoril iba jeden člen $K \cos(t)$ v procese a zvyšovali sme postupne koeficient K . Teraz nám vystupujú v procese 4 členy, taktiež zvyšujeme koeficienty, ale rozdiely v sile testu sú menšie ako v prípade A.



Obrázok č. 13: Silofunkcia Fisherovho testu pre náhodné procesy X_t^1, X_t^2, X_t^3

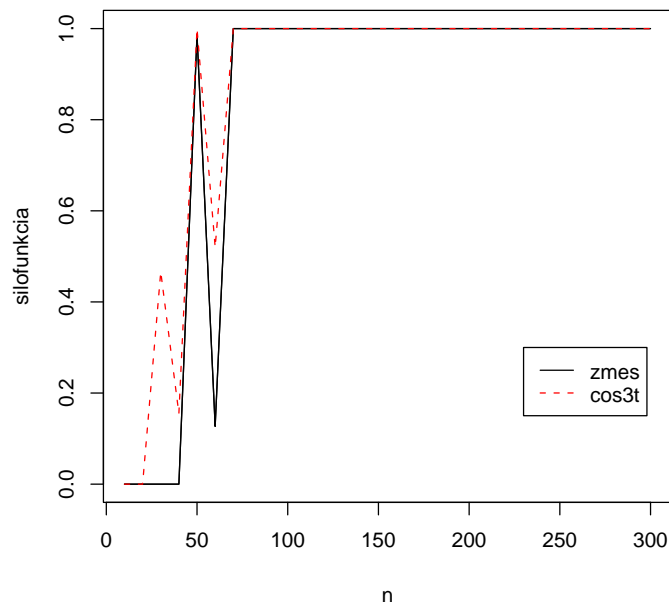
C) Simulujeme náhodné procesy X_t^1, X_t^2 , kde

$$X_t^1 = 3 \cos(t) + Z_t,$$

$$X_t^2 = 3 \cos(t) + 2 \cos(2t) + 3 \sin(t) + 3 \sin(2t) + Z_t,$$

Z_t je biely šum.

Simulácia prebieha presne tak isto ako v prípadoch A a B. Výsledok simulácie je zobrazený na obrázku č. 14.



Obrázok č. 14: Silofunkcia Fisherovho testu pre náhodné procesy

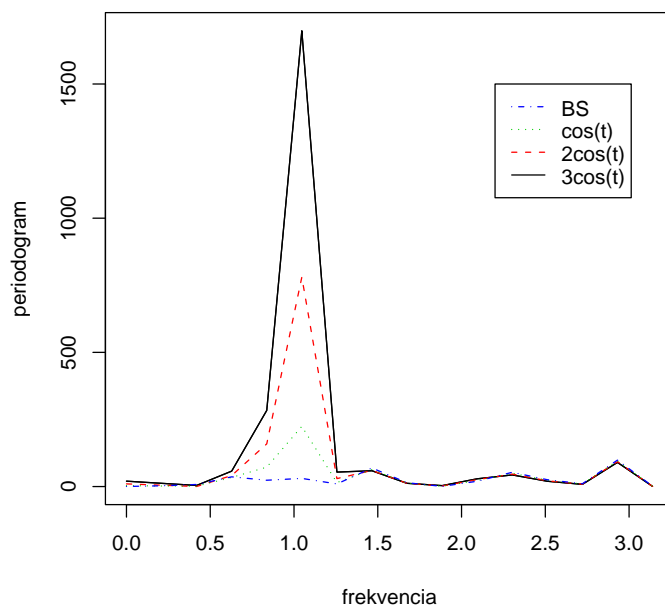
Z tohto obrázka je vidno, že silofunkcie pre procesy X_t^1, X_t^2 sú veľmi podobné. Zaujímavé je, že pre proces X_t^1 nadobúda silofunkcia väčšie hodnoty ako silofunkcia procesu X_t^2 . Pritom v procese X_t^2 tvoria zmes kosínusov a sínusov 4 členy, kým proces X_t^1 obsahuje iba jeden.

Záver

Môžeme povedať, že s narastajúcim rozsahom výberu nám rastie sila testu i keď s istými "výkyvmi". Pre výbery menších rozsahov sú výkyvy väčšie, so zväčšujúcim sa rozsahom výberu okrem zväčšujúcej sa sily testu dochádza aj k zmenšovaniu jej výkyvov.

Okrem rozsahu výberu má významný vplyv na silu testu aj veľkosť koeficientov pri sínusoch a kosínusoch, ktoré vystupujú v náhodnom procese. Menší vplyv má už počet a veľkosť periódy týchto funkcií. Je tomu tak preto, že s narastajúcim koeficientom K_i pri niektorom z členov (sínusu alebo kosínusu) náhodného procesu narastá aj maximum jeho periodogramu. Tým sa nám zväčšuje rozdiel medzi maximálnou a priemernou hodnotou periodogramu. Táto situácia je zachytená na obrázku č. 15. Z tohto obrázka vidíme, že pre narastajúce koeficienty K nadobúda periodogram procesu $\{X_t\}$ maximum v tej istej frekvencii, ale veľkosť maxima sa zväčšuje. To je aj dôvod prečo nemá až taký vplyv na silu testu počet členov vystupujúcich v náhodnom procese. Pre zväčšujúci sa počet

členov nám nerastie maximum periodogramu, ak sú koeficienty pri pridávaných členoch menšie alebo rovné ako tie pôvodné. Pre menšie rozsahy výberov sa nám môže dokonca stať, že pridaním ďalších členov sa nám zväčší priemerná hodnota periodogramu (pretože pridaním každej novej funkcie sinus alebo kosínus nám v periodograme pribudne nové lokálne maximum) a sila testu sa tak môže zmenšiť.



Obrázok č. 15: Periodogram náhodných procesov $X_t = K \cos(t) + Z_t$, $K = 0, 1, 2, 3$

Najdôležitejším poznatkom je ale fakt, že pre malé rozsahy výberov je Fisherov test nepoužiteľný. Z obrázkov č. 12 - 14 je vidno, že sila testu je pre rozsahy výberov $n \leq 50$ veľmi malá. Navyše sú prítomne veľké výkyvy v silofunkciách.

5 Aplikácia kváziperiodických časových radov v kardiológii

Táto aplikácia vychádza z práce [6] (Katina, Štulajter et al, 2005). Z tejto práce je prevzatý úvod, základné pojmy a časť teórie.

5.1 Úvod

Matematická štatistika má široké uplatnenie v biomedicínskych vedách. Jedným z príkladov je dizajn experimentu s opakovanými meraniami v ekvidištantných časových bodoch. Práve vyššie spomenutý dizajn vedie k použitiu kváziperiodických časových radov.

S narastajúcimi možnosťami ambulantného monitorovania krvného tlaku (TK) u dospelých sa potvrdzuje, že významnou zložkou intraindividuálnej variability TK je jeho periodcké kolísanie. V literatúre najskôr a najviac (rôznymi metódami pre analýzu periodických procesov) dokumentovaným biologickým rytmom (TK) je jeho cirkadiálne (CD) kolísanie s nočným minimom. Najčastejšou štatistickou metódou jeho dôkazu a amplitúdovo-časovej kvantifikácie je Halbergom zavedená kosínorová analýza (Halberg et al. 1967). Svoje miesto v diagnostike hypertenzie, resp. v chronofarmakológii má hľadanie charakteru cca 24-hodinového periodického výkyvu jeho amplitúdy okolocelodenného priemeru, časovanie maxima a minima a taktiež prípadná neprítomnosť nočného poklesu TK .

Konštrukcia vlastného ultrazvukového prístroja na meranie TK u novorodencov (Kellerová et al. 1978) nám umožnila prioritný dôkaz na existencie CD kolísania TK už u novorodencov v druhom postnatálnom dni (Kellerová et al. 1978, 1981). Rozptyl jednotlivých hodnôt TK okolo odhadu strednej hodnoty s 24 hodinovou periódou (s použitím len kosínusovej funkcie) si vyžiadala podrobnejšiu analýzu na prítomnosť superponovaných ultradiálnych (UD) periodicít. Táto ukázala, že kým CD rytmus TK a pulzovej frekvencie sa vyskytuje u 100 percent zdravých dospelých, približne po 2 pomalé UD rytmy skoro u každého a rýchle UD len asi z 30 percent, v porovnaní s tým u novorodencov vyšetovaných v kontrolovanom režime novorodeneckého oddelenia, sa CD a pomalé UD cykly pozorovali u 60-57 percent a rýchlych UD periód bolo približne 4-násobne viac (Kellerová et al. 1989).

5.2 Materiál a metodika

5.2.1 Súbor jedincov

Súbor jedincov tvorilo $p = 20$ fyziologických novorodencov s postnatálnym vekom 45 ± 11 hodín, pôrodnou hmotnosťou 3551 ± 172 gramov (aritmetický priemer \pm smerodajná odchylka), vo voľnom režime "rooming in" s matkou. Na týchto jedincoch sme merali TK v hodinových intervaloch automatickým oscilometrickým tlakomerom Nippon-Collin. Z ďalšej analýzy sme vyradili 3, pretože jedno alebo viac chýbajúcich pozorovaní, teda $p = 17$.

5.2.2 Štatistické metódy

Analýza kváziperiodických časových radov je postavená na báze analýzy frekvencií asociovaných s cyklami, ktoré sa v dátach nachádzajú. Najčastejšie používanou množinou matematických funkcií je zmes kosínusov a sínusov.

Majme nelineárny regresný model [7] (*NRM*, Štulajter, 2002)

$$X(t) = \beta_1 + \beta_2 t + \sum_{j=1}^k [\beta_{1j} \cos \lambda_j t + \beta_{2j} \sin \lambda_j t] + \varepsilon(t); t \in T, \quad (1)$$

kde $\beta_{full} = (k, \beta, \lambda)^T = (k, \beta_1, \beta_2, \beta_{11}, \dots, \beta_{1k}, \beta_{21}, \dots, \beta_{2k}, \lambda_1, \dots, \lambda_k)^T$ je neznámy vektor regresných parametrov (počet parametrov modelu (1) je $3k + 3 < n$, kde k je potrebné tiež odhadnúť), $\varepsilon(t)$ je biely šum (nekorelované náhodné premenné s nulovou strednou hodnotou a s rovnakým rozptylom σ^2). Neznáme parametre modelu (1) odhadneme z dát $X(t); t = 1, 2, \dots, n$ tak, že najprv vypočítame odhady lineárneho trendu $\hat{\beta}_1$ a $\hat{\beta}_2$ a to obyčajnou metódou najmenších štvorcov (*MNS*), potom vypočítame $\widehat{X}(t) = X(t) - \hat{\beta}_1 - \hat{\beta}_2 t, t = 1, 2, \dots, n$ a tieto použijeme na výpočet periodogramu podľa vstahu (Christensen, 2003)

$$I_n(\lambda) = \frac{1}{2\pi n} \left\{ \left[\sum_{t=1}^n \widehat{X}(t) \cos(\lambda t) \right]^2 + \left[\sum_{t=1}^n \widehat{X}(t) \sin(\lambda t) \right]^2 \right\}, \lambda \in [0, \pi], t \in T. \quad (2)$$

Poznámka 5.2.1 Periodogram sa počíta len v niektorých frekvenciách. Najčastejšie sa používajú, ak n je párne, Furierove frekvencie $\lambda_j = \frac{2\pi}{n} j, j = 1, 2, \dots, n/2$. Frekvenciám λ zodpovedajú periódy $T_\lambda = \frac{2\pi}{\lambda}$.

Hodnotu k a frekvencie $\hat{\lambda}_j$, odhady frekvencií $\lambda_j, j = 1, 2, \dots, k$, určíme s periodogramu, je to počet všetkých lokálnych maxím funkcie (2), resp. frekvencie, v ktorých sú tieto

lokálne maximá. Potom uvažujeme lineárny regresný model (*LRM*)

$$X(t) = \hat{\beta}_1 + \hat{\beta}_2 t + \sum_{j=1}^k [\beta_{1j} \cos \hat{\lambda}_j t + \beta_{2j} \sin \hat{\lambda}_j t] + \varepsilon(t); t \in T, \quad (3)$$

v ktorom odhadneme neznáme parametre $\beta_1, \beta_2, \beta_{11}, \dots, \beta_{1k}, \beta_{21}, \dots, \beta_{2k}$ pomocou obyčajnej *MNŠ*.

Model (3) je aditívnym saturovaním modelom bez interakcií. Vychádzajúc z tohoto modelu, pomocou Akaikeho informačného kritéria (*AIC*, Akaike, 1974)

$$AIC(\widehat{\mathbf{F}}\hat{\beta}, \hat{\sigma}^2) = -2l(\widehat{\mathbf{F}}\hat{\beta}, \hat{\sigma}^2) - 2k,$$

kde $l(\cdot)$ je maximum funkcie vierohodnosti, $\widehat{\mathbf{F}}\hat{\beta}$ je maticový zápis odhadu strednej hodnoty, nájdeme optimálny submodel použitím spätného krokového mechanizmu (Vanables a Ripley, 2002).

Teda iterovaním za predpokladu konvergenencie vyradujeme pomocou slučiek tie regresné parametre, ktoré spôsobujú pokles *AIC*. Iterácie zastanú vtedy, keď už nenastane ďalší pokles *AIC*.

Poznámka 5.2.2 Pôvodne sme plánovali použiť okrem *AIC* kritéria aj Fisherov test a porovnať kvalitu výsledkov pomocou veľkosti reziduálnej sumy štvorcov (*RSS*). Simulačná štúdia Fisherovho testu ukázala, že pre malé rozsahy výberov je nespoľahlivý a tak sme ho z analýzy vynechali.

Načrtnutý postup hľadania optimálneho submodelu zhrnieme v nasledovných bodoch:

1. najprv vyberieme všetky lokálne maximá funkcie $I_n(\cdot)$ v (2), ich počet označíme $k < n$,

2. potom frekvencie k nim prislúchajúce použijeme v *LRM* (3), kde regresné parametre odhadneme *MNŠ*,

3. ďalej po spätnej krokovej procedúre ostane v modeli (3) $l \leq 2k$ regresných parametrov, kedy tento model budeme považovať za optimálny submodel.

Majme p jedincov v náhodnom výbere, na ktorých opakovane meriame n -krát nejakú premennú v čase za rovnakých podmienok. Pre každého i -teho jedinca máme realizácie $x_i(t); i = 1, 2, \dots, p; t = 1, 2, \dots, n$ a pre každého z nich uvažujeme *NRM* (1), periodogram $I_n^i(\lambda)$ s $\lambda \in [0, \pi]$ a *LRM* (3). Hlavnou úlohou je identifikovať najčastejšie sa vyskytujúce periódny za predpokladu, že realizácie $x_i(t)$ môžu byť teoreticky u každého jedinca posunuté o inú fázu a navyiac výskyt hľadaných periód môže byť rôznych.

Ďalej sa môžeme pýtať, či periodické zmeny v systolickom krvnom tlaku i -teho jedinca pochádzajú z toho istého stochastického procesu ako periodické zmeny j -teho jedinca, pre $i = 1, 2, \dots, p, i \neq j$. Pod periodickými zmenami systolického krvného tlaku i -teho jedinca rozumieme údaje $x_i(t) - \hat{\beta}_1 - \hat{\beta}_2 t; t = 1, 2, \dots, 24$, kde $\hat{\beta}_1$ a $\hat{\beta}_2$ sú odhady lineárneho trendu realizácií $x_i(t); t = 1, 2, \dots, 24$. Tieto údaje jednotlivých jedincov tvoria nezávislé stacionárne časové rady. Nezávislosť je zabezpečená tým, že realizácie $x_i(t); t = 1, 2, \dots, 24$ a $x_j(t); t = 1, 2, \dots, 24$ pre $i \neq j$ sú nezávislé, pretože sú merané na rôznych jedincoch v tých istých časových odstupoch (hodinových).

Teda hypotéza H_0 znie, že periodické zmeny v systolickom krvnom tlaku i -teho a j -teho jedinca pochádzajú z toho istého stochastického procesu.

Hypotéza H_1 znie, že periodické zmeny v systolickom krvnom tlaku i -teho a j -teho jedinca pochádzajú z rôznych stochastických procesov. Ide teda o obojstranné hypotézy.

Máme dve možnosti ako overiť dané hypotézy. Prvou možnosťou je použiť permutačný test (pozri kap.2). Simulačná štúdia tohto testu však ukázala, že test je pre malé rozsahy výberov nepoužiteľný (pozri podkap. 4.4). Druhou možnosťou je použitie grafickej metódy, tzv CP -plot (pozri podkap. 4.4).

5.2.3 Odhaľovanie významných frekvencií

V našej analýze zmien systolického krvného tlaku sa snažíme identifikovať významné frekvencie, ktoré by mali byť zahrnuté do regresného modelu. Kľúčovú úlohu pri tom zohráva periodogram, kde za významné frekvencie považujeme frekvencie, v ktorých nadobúda svoje lokálne maximá (pozri postup v predchádzajúcej podkap. 5.2.2). Teraz si objasníme prečo je to tak.

Majme kváziperiodický nelineárny regresný model

$$X(t) = \gamma_1 + \sum_{i=1}^k (\beta_i^1 \cos \lambda_i t + \beta_i^2 \sin \lambda_i t) + \varepsilon(t); t \in T,$$

kde $\gamma = (\gamma_1, \beta_1^1, \dots, \beta_k^1, \beta_1^2, \dots, \beta_k^2, \lambda_1, \dots, \lambda_k)^T = (k, \gamma_1, \beta, \lambda)^T$ je vektor neznámych regresných parametrov, ktoré je potrebné odhadnúť. Na odhad parametrov použijeme periodogram. Nech $X(t) = m(t) + \varepsilon(t); t \in T$ je časový rad. Nech $X_n = (X(1), \dots, X(n))^T$ je jeho konečným pozorovaním.

Definícia 5.2.1 *Reálna funkcia $I_n(\lambda); \lambda \in [-\pi, \pi]$ definovaná predpisom*

$$I_n(\lambda) = \frac{1}{2\pi n} \left\{ \left[\sum_{t=1}^n x_t \cos(t\lambda) \right]^2 + \left[\sum_{t=1}^n x_t \sin(t\lambda) \right]^2 \right\},$$

sa nazýva *periódogram pozorovania* X_n , kde $\omega = \omega_j = \frac{2\pi j}{n}$, $j = 1, \dots, m$, $m = \frac{n-1}{2}$ (Štula-
jter, 1990).

Periódogram sa počíta väčšinou pre Furiérove frekvencie $\lambda_j = (\frac{2\pi}{n})j$, $j = 1, 2, \dots, \frac{n}{2}$.

Periódogram môžeme zapísať ako

$$I_n(\lambda) = \frac{1}{2\pi n} \left\{ \left[\sum_{t=1}^n [m(t) + \varepsilon(t)] \cos(t\lambda) \right]^2 + \left[\sum_{t=1}^n [m(t) + \varepsilon(t)] \sin(t\lambda) \right]^2 \right\}.$$

V prípade, že časový rad ideme modelovať kváziperiodickým nelineárnym regresným modelom, teda keď

$$m(t) = m_\gamma(t) = \gamma_1 + \sum_{i=1}^k (\beta_i^1 \cos \lambda_i t + \beta_i^2 \sin \lambda_i t); t = 1, 2, \dots, n,$$

periodogram nadobúda tvar

$$I_n(\lambda) = \frac{1}{2\pi n} \left[\sum_{t=1}^n \left(\gamma_1 + \sum_{i=1}^k (\beta_i^1 \cos \lambda_i t + \beta_i^2 \sin \lambda_i t) \right) \cos(\lambda_j t) \right]^2 + \frac{1}{2\pi n} \left[\sum_{t=1}^n \left(\gamma_1 + \sum_{i=1}^k (\beta_i^1 \cos \lambda_i t + \beta_i^2 \sin \lambda_i t) \right) \sin(\lambda_j t) \right]^2.$$

Nech teraz $\lambda_j = \lambda_i$ pre nejaké $i = 1, 2, \dots, k$, potom

$$I_n(\lambda_i) = \frac{1}{n} (\beta_i^1 \sum_{t=1}^n \cos^2 \lambda_i(t) + \sum_{t=1}^n \varepsilon(t) \cos^2 \lambda_i(t))^2 + \frac{1}{n} (\beta_i^2 \sum_{t=1}^n \sin^2 \lambda_i(t) + \sum_{t=1}^n \varepsilon(t) \sin^2 \lambda_i(t))^2.$$

Vidíme, že výraz sa nám značne zjednodušil, platia totiž nasledujúce vzťahy

$$\begin{aligned} \sum_{t=1}^n \sin \lambda_i t &= \sum_{t=1}^n \cos \lambda_j t, \\ \sum_{t=1}^n \cos \lambda_i t \cos \lambda_j t &= 0, \\ \sum_{t=1}^n \sin \lambda_i t \sin \lambda_j t &= 0, \\ \sum_{t=1}^n \sin \lambda_i t \cos \lambda_j t &= 0. \end{aligned}$$

$$\begin{aligned} I_n(\lambda) &= \frac{1}{n} \left[\left(\frac{n}{2} \beta_i^1 + \sum_{t=1}^n \varepsilon(t) \cos^2 \lambda_i(t) \right)^2 + \left(\frac{n}{2} \beta_i^2 + \sum_{t=1}^n \varepsilon(t) \sin^2 \lambda_i(t) \right)^2 \right] \\ &= n \left[\left(\frac{1}{2} \beta_i^1 + \frac{1}{n} \sum_{t=1}^n \varepsilon(t) \cos^2 \lambda_i(t) \right)^2 + \left(\frac{1}{2} \beta_i^2 + \frac{1}{n} \sum_{t=1}^n \varepsilon(t) \sin^2 \lambda_i(t) \right)^2 \right]. \end{aligned}$$

$\varepsilon(t)$ je biely šum so strednou hodnotou $\mu = 0$ a rozptylom σ^2 , teda platí

$$E \left[\frac{1}{n} \sum_{t=1}^n \varepsilon(t) \cos \lambda_i(t) \right] = E \left[\frac{1}{n} \sum_{t=1}^n \varepsilon(t) \sin \lambda_i(t) \right] = 0,$$

$$D \left[\frac{1}{n} \sum_{t=1}^n \varepsilon(t) \cos \lambda_i(t) \right] = D \left[\frac{1}{n} \sum_{t=1}^n \varepsilon(t) \sin \lambda_i(t) \right] = \frac{\sigma^2}{2n}.$$

Dostávame

$$\lim_{n \rightarrow \infty} I_n(\lambda_i) = +\infty$$

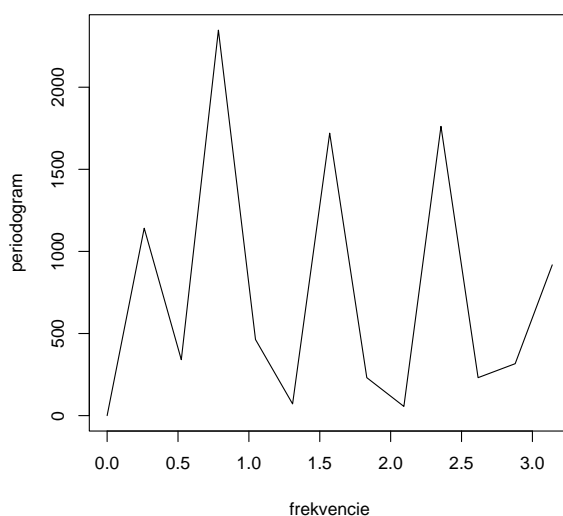
pre všetky $i = 1, 2, \dots, k$.

Vidíme, že ak je nejaká z furiérových frekvencií významná pre náš regresný model, pôjde o frekvenciu, v ktorej dosahuje periodogram lokálne maximum. Teda všetky lokálne maximá periodogramu by mali byť zahrnuté ako významné frekvencie do regresného modelu. Existuje Fisherov test (kap. 3), pomocou ktorého sa dá určiť či je daná frekvencia v ktorej nadobúda periodogram lokálne maximum významná pre regresný model.

5.2.4 Výsledky a diskusia

Identifikácia významných periód v systolickom krvnom tlaku

V našom prípade meriame t v hodinách ($n = 24$), $t = 1$ pre 7. hodinu ráno a $t = 24$ je 6. hodina ráno nasledujúceho dňa. Na ilustráciu sme si vybrali časové zmeny systolického krvného tlaku novorodenca č. 3. Z periodogramu (obrázok č. 16) vyberieme všetky lokálne maximá, ktoré spolu s odhadnutým lineárnym trendom vstupujú do iteračných procedúr.



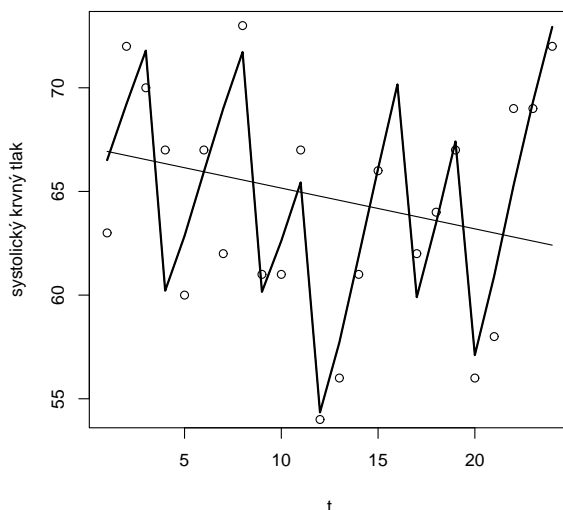
Obrázok č. 16: Periodogram (novorodenec č. 3)

Optimálny submodel bude mať potom nasledovný tvar

$$\hat{x}(t) = 67.128 - 0.196t + 2.879 \cos(0.083\pi t) + 4.078 \cos(0.25\pi t) + 3.503 \cos(0.75\pi t) - 3.611 \sin(0.5\pi t),$$

kde sú zahrnuté nasledovné frekvencie $\hat{\lambda}_{(1)} = 0.083\pi$, $\hat{\lambda}_{(2)} = 0.25\pi$, $\hat{\lambda}_{(3)} = 0.75\pi$, $\hat{\lambda}_{(4)} = 0.083\pi$ a k nim prislúchajúce periódy $T_{0.083\pi}^{(1)} = 3.831$, $T_{0.25\pi}^{(2)} = 1.273$, $T_{0.75\pi}^{(3)} = 0.424$, $T_{0.5\pi}^{(4)} = 0.636$ hodiny (tu $k = 4$).

Odhadnutú strednú hodnotu optimálneho submodelu časového radu znázorňujeme ako spojitú funkciu času. Do grafu zakresľujeme aj skutočné (namerané) hodnoty sledovaného parametra v diskretných časoch a lineárny trend (obrázok č. 17).

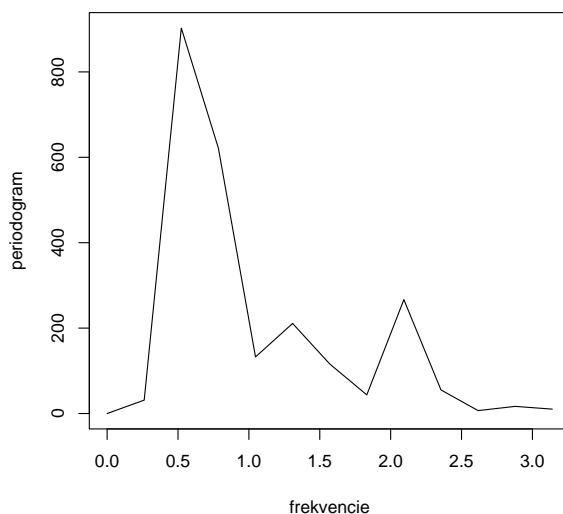


Obrázok č.17: Kváziperiodický časový rad systolického krvného tlaku (novorodenec č. 3), odhadnutá stredná hodnota a lineárny trend

Globálny pohľad na periodické zmeny systolického krvného tlaku novorodencov v nami sledovanom súbore môžeme vyjadriť pomocou periodogramu vypočítaného z nasledovných údajov

$$\bar{x}(t) - \bar{x} = \frac{1}{17} \sum_{i=1}^{17} x_i(t) - \frac{1}{24} \sum_{t=1}^{24} \bar{x}(t); t = 1, 2, \dots, 24.$$

Z periodogramu vypočítaného z týchto údajov vyplýva, že v systolickom krvnom tlaku pozorovanej skupiny detí sú významné periódy. Periodogram je znázornený na obrázku č. 18.



Obrázok č. 18: Periodogram z časového radu priemerného systolického krvného tlaku sledovanej skupiny detí.

Optimálny submodel bude mať potom nasledovný tvar

$$\begin{aligned}\hat{x}(t) = & 69.739 - 0.282t \\ & + 1.736 \cos(0.166\pi t) \\ & + 1.949 \sin(0.166\pi t) + 0.973 \sin(0.328\pi t) - 1.117 \sin(0.666\pi t).\end{aligned}$$

Teda najvýznamnejšie frekvencie sú $\hat{\lambda}_{(1)} = 0.166\pi$, $\hat{\lambda}_{(2)} = 0.328\pi$, $\hat{\lambda}_{(3)} = 0.666\pi$ a k nim prislúchajúce periódy $T_{0.166\pi}^{(1)} = 1.912$, $T_{0.328\pi}^{(2)} = 0.97$, $T_{0.666\pi}^{(3)} = 0.477$ hodiny (tu $k = 3$).

Testovanie hypotéz

Zoberme časové rady periodických zmien systolického krvného tlaku prvého a druhého jedinca. Chceme vedieť, či tieto časové rady pochádzajú z toho istého stochastického procesu alebo nie.

Hypotéza znie

Prípad A)

H_0 : periodické zmeny jedinca č.3 = periodické zmeny jedinca č.4,

H_1 : periodické zmeny jedinca č.3 \neq periodické zmeny jedinca č.4.

Prípád B)

H_0 : periodické zmeny jedinca č.2 = periodické zmeny jedinca č.3,

H_1 : periodické zmeny jedinca č.2 \neq periodické zmeny jedinca č.3.

Ide teda o obojstranné hypotézy.

Na overenie danej hypotézy použijeme permutačný test a CP -plot.

Metóda CP -plotu

Na obrázkoch č. 19 a č. 20 sú znázornené CP -ploty jedincov č. 3 a č. 4 (prípád A) resp. č. 2 a č. 3 (prípád B). Vidíme, že CP -plot v prípade A je približne tvaru $y = x$, kým CP -plot v prípade B nie. Na obrázkoch je zachytený aj 95 % interval spoľahlivosti.

Interval spoľahlivosti konštruujeme tak, že simulujeme $k = 1000$ -krát biely šum a ten pripočítame k periodickej časti odhadnutej strednej hodnoty jedinca č.3 (strednú hodnotu jedinca č.3 sme odhadovali v "Identifikácia významných periód v systolickom krvnom tlaku"). Simulácie teda vyzerajú nasledovne

$$X_t^i = 2.879 \cos(0.083\pi t) + 4.078 \cos(0.25\pi t) + 3.503 \cos(0.75\pi t) - 3.611 \sin(0.5\pi t) + Z_t^i,$$

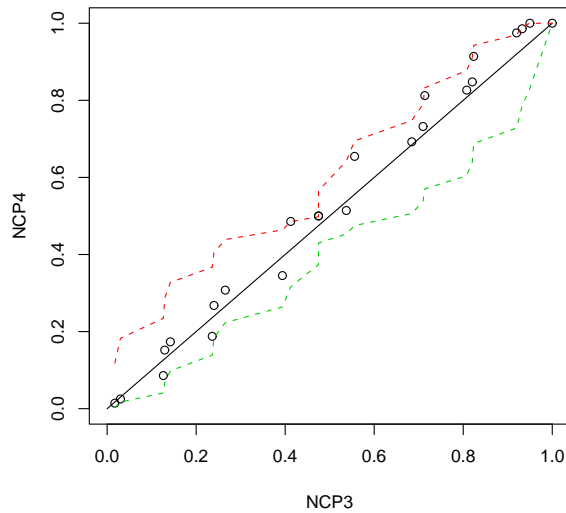
kde Z_t^i je biely šum pre $t = 1, 2, \dots, 24; i = 1, 2, \dots, 1000$. Dostaneme tak $k = 1000$ procesov X_t so strednou hodnotou rovnou periodickej časti jedinca č.3, pre každý z týchto procesov vypočítame NCP (normovaný kumulatívny periodogram) a vypočítame 5% a 95% kvantil NCP v každej jeho frekvencii. Periodogram počítame vo furiérových frekvenciách.

Poznámka 5.2.3 Všimnime si, že nikde nespomínáme rozptyl simulovaného bieleho šumu ani rozptyl skúmaných pozorovaní. Veľkosť rozptylu nemusíme uvažovať vďaka vlastnosti NCP , ktorá nám hovorí, že podstatný je priebeh časového radu a nie jeho absolútna hodnota. Teda či má napr. biely šum rozptyl $\sigma^2 = 1$, alebo $\sigma^2 = 0.5$, jeho NCP bude rovnaký.

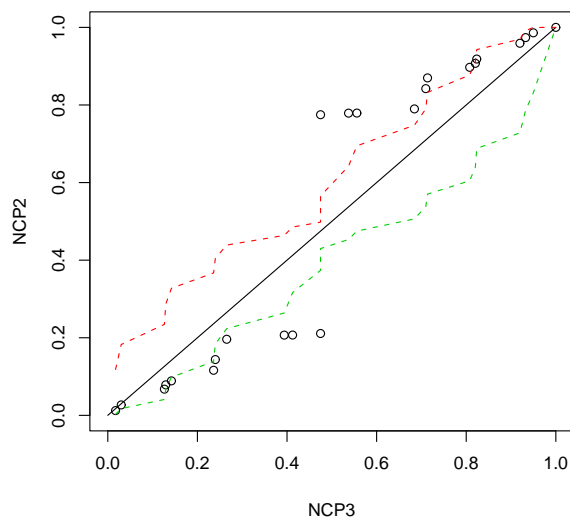
Rovnakú vlastnosť má aj permutačný test (pozri podkap.2.4).

Vidíme, že CP -plot v prípade A sa do intervalu spoľahlivosti vošiel, kým v prípade B značná časť CP -plotu prekračuje interval spoľahlivosti. Mohli by sme teda urobiť záver, že v prípade A nezamietame hypotézu H_0 na hladine významnosti $\alpha = 0.05$, kým v prípade B zamietame H_0 . Potrebujeme ale poznať, aká veľká je chyba druhého druhu tejto metódy, aby naše závery boli správne. Chybu druhého druhu budeme zisťovať na dvojici bieleho šumu a pozorovania jedinca č. 3. Vieme, že pozorovania jedinca č. 3 obsahujú skryté

periódy a teda sa určite líšia od bieleho šumu. Zaujímá nás s akou pravdepodobnosťou zamietne táto metóda H_0 ak vieme, že H_0 je nepravdivá.



Obrázok č. 19: CP plot jedincov č. 3 a č. 4



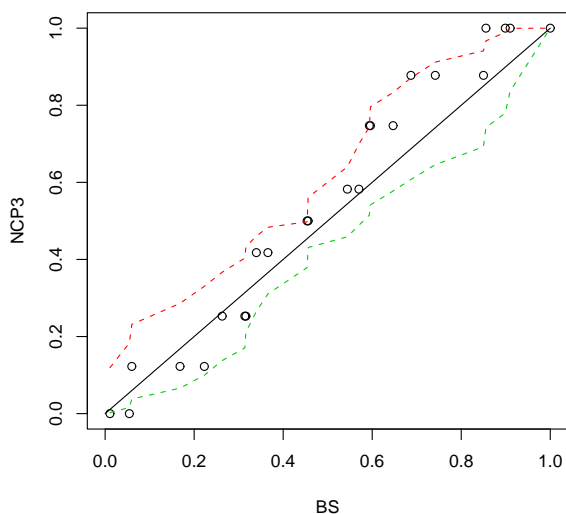
Obrázok č. 20: CP plot jedincov č. 2 a č. 3

Na obrázku č. 21 je CP -plot bieleho šumu a jedinca č. 3. Vidíme, že graf má charakter esovitej krivky, čo naznačuje, že ide o dva rôzne procesy.

Interval spoľahlivosti konštruujeme tak, že simulujeme $k = 1000$ -krát biely šum, vypočítame pre každý z nich NCP a vypočítame 5 % a 95 % kvantil NCP v každej jeho

frekvencii. Vidíme že skoro celý CP -plot leží buď na hranici intervalu spoľahlivosti alebo v ňom. Chyba druhého druhu tejto metódy je teda veľká, ak porovnávame biely šum a jedinca č. 3. Dôvod je ten, že ak sa pozrieme na periodogram jedinca č. 3 (obrázok č. 16) vidíme, že lokálne maximá periodogramu sú približne rovnako veľké. Z vlastností (NCP) periodogramu potom vyplýva, že (NCP) jedinca č. 3 sa nebude veľmi odlišovať od (NCP) bieleho šumu.

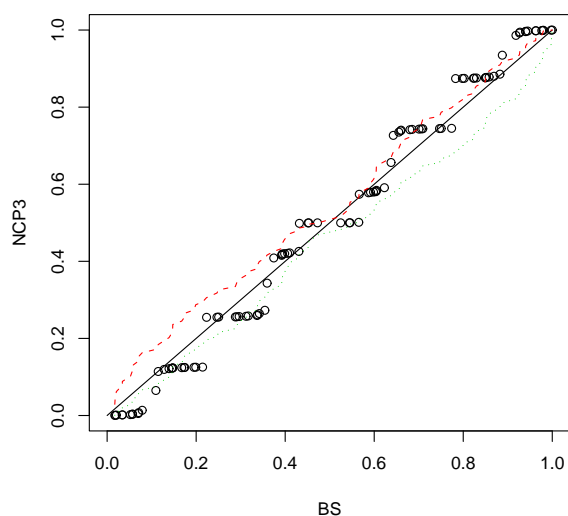
Aby sme presne poznali veľkosť chýb druhého druhu tejto metódy, trebalo by preveriť rôzne kombinácie dvojíc procesov obsahujúce rôzne zmesi sínusov a kosínusov a preveriť chyby pre rôzne veľkosti rozsahov výberov. Totiž pre každú z týchto možností môžeme dostávať rôzne veľkosti chýb druhého druhu. To isté platí aj pre chybu prvého druhu, ktorá sa môže líšiť pre rovnaké dvojice porovnávaných procesov, tvorené rôznymi zmesami sínusov a kosínusov.



Obrázok č. 21: CP plot jedinca č. 3 a bieleho šumu

Na obrázku č. 22 je znázornený ten istý prípad ako na obrázku č. 21, len rozsah výberu sme zväčšili na $n = 24$ na $n = 100$. Z obrázka vidíme jasne, že ide o dva rôzne procesy. Graf nie je tvaru $y = x$. Vidíme, že väčšina bodov CP plotu prekračuje interval spoľahlivosti, teda pravdepodobnosť zamietnutia hypotézy H_0 je dostatočná. Ak porovnáme tento prípad s prípadom na obrázku č. 21 tak zistíme, že rozsah výberu má na túto metódu značný vplyv.

Podstatné ale je, že pre rozsah našich skúmaných údajov veľkosti $n = 24$, je táto metóda nevyhovujúca, pretože má malú silu (veľkú chybu druhého druhu).



Obrázok č. 22: CP plot stochastického procesu so strednou hodnotou jedinca č. 3 a bieleho šumu s rozsahom výberu $n = 100$

Záver teda je, že bez podrobnej analýzy chýb prvého a druhého druhu môže mať táto metóda len informatívny charakter. Teda pozeráme sa na CP -plot bez zahrnutia intervalu spoľahlivosti s tým, že sledujeme, či je CP -plot tvaru $y = x$ alebo nie je.

Skúsime ešte otestovať dané hypotézy pomocou permutačného testu.

Permutačný test

Pre jedinco č. 3 a č. 4 je P -hodnota testu rovná 1. Teda na ľubovolnej hladine významnosti α nezamietame hypotézu H_0 . Vieme však, že sila testu je veľmi malá (podľa simulačnej štúdie je pre náš rozsah výberu rovná 0) a tak treba brať tento výsledok s veľkou rezervou.

Pre jedinco č. 2 a č. 3 je P -hodnota testu rovná 0.089. Teda na hladine významnosti $\alpha = 0.05$ nezamietame hypotézu H_0 . V porovnaní s predošlým prípadom síce tiež nezamietame hypotézu H_0 , ale P -hodnota je oveľa menšia ($0.089 < 1$). Ak sa pozrieme na predošlé výsledky dané grafickou metódou CP -plotu, vidíme istú analógiu. Pre jedinco č. 3 a č. 4 je CP -plot tvaru $y = x$ a P -hodnota testu je rovná 1. Pre jedinco č. 2 a č. 3 nemá CP -plot tvar $y = x$ a P -hodnota testu je rovná 0.089. Samozrejme metóda CP -plotu ani permutačný test nám v ani v jednom z dvoch skúmaných prípadov nezamietli hypotézu H_0 , ale spolu nám dávajú lepší obraz o situácii. Môžeme povedať, že v prípade jedinco č. 3 a č. 4 nám metóda CP -plotu aj permutačný test silnejšie nezamietajú hypotézu H_0 ako v prípade jedinco č. 2 a č. 3.

Záver

Na základe periodogramu urobeného z priemerov sledovaného systolického krvného tlaku detí môžeme povedať, že v zmenách tohto tlaku sa vyskytujú významné periódy. Celkovo nám tlak klesá (trendová zložka je v regresnom modeli klesajúca) s kolísaniami periodického charakteru (spôsobenými identifikovanými významnými periódami).

Nepodarilo sa nám určiť, či periodické zmeny v systolickom krvnom tlaku detí pochádzajú z toho istého stochastického procesu. Dôvodom je slabá sila permutačného testu pre malé rozsahy výberov. Naše rozsahy výberov sú rozsahu $n = 24$. Pre tento rozsah výberu je sila permutačného testu nulová (pozri podkap. 4.4), čiže nemá zmysel výsledky daného testu uvažovať. Ako alternatíva sa ponúkala metóda založená na *CP*-plote. Ani tu však neboli preukázané uspokojivé výsledky. Môže za to malá schopnosť metódy zamietnuť hypotézu H_0 ak je nepravdivá. Preto táto metóda môže slúžiť iba na počiatočnú analýzu pozorovaní.

Záver

Jeden zo základných cieľov, ktoré sme si na začiatku stanovili bol určiť veľkosť chýb prvého a druhého druhu permutačného testu porovnávajúceho dva nezávislé časové rady. Výsledky simulačnej štúdie permutačného testu ukázali, že tento test je použiteľný len pre veľké rozsahy výberov. Ide o rozsahy výberov veľkosti $n = 200$ a viac, pre malé rozsahy je sila testu nulová. Metóda *CP*-plotu dáva podobné výsledky ako permutačný test. Pre malé rozsahy výberov sa *CP*-plot správa pomerne nevyspytateľne, ak neberieme do úvahy aj 95 % interval spoľahlivosti. Tu nebola urobená dôkladná simulačná štúdia, na základe ktorej by sme mali vedomosti o chybách prvého a druhého druhu. Odporúčame pri analýze časových radov používať permutačný test aj metódu *CP*-plotu spolu. Ani jedna metóda síce nie je spoľahlivá, ale spolu nám môžu slúžiť na vytvorenie predstavy o správaní sa daných časových radov. Simulačná štúdia Fisherovho testu preukázala, že test je podobne ako permutačný test použiteľný len pre veľké rozsahy výberov. Pre malé rozsahy výberov sú badať veľké výkyvy v sile testu. S narastajúcim rozsahom výberu sa výkyvy postupne znižujú, ale k ustáleniu dôjde až pri veľkom rozsahu výberu. Preto neodporúčame používať v praxi tento test pre rozsahy výberov menšie ako $n = 120$, kedy je sila testu dostatočná a výkyvy sú zanedbateľné. V našom prípade to znamená merať systolický krvný tlak *SKT* aspoň 4 dni v hodinových intervaloch alebo každých 15 minút po dobu minimálne 1.5 dňa. *KS*-test, permutačný test a Fisherov test majú jedno spoločné. Chyba prvého druhu je pri týchto testoch veľmi malá a nezávisí na veľkosti rozsahu výberu. U *KS*-testu sa pohybuje na úrovni okolo 0.04 a u zvyšných dvoch je približne rovná 0.001. Teda dá sa povedať, že permutačný a Fisherov test dobre kontrolujú chybu prvého druhu ale chyba druhého druhu je pre malé rozsahy výberov príliš veľká. *KS*-test má o niečo väčšiu chybu prvého druhu, ale zato je chyba druhého druhu oveľa menšia.

V analýze *SKT* sme predstavili metódu *CP*-plotu. Oplatilo by sa spraviť podrobnú simulačnú štúdiu vlastností chýb prvého a druhého druhu tejto metódy, rovnako aj *AIC* kritéria. Ďalej z analýzy *SKT* vyplynulo, že u novorodencov sa prejavuje nočný pokles krvného tlaku, čo by bolo zaujímavé vyhodnotiť.

Literatúra

- [1] Brockwell, P. a Davis, R.: Time Series: Theory and Methods, Second edition, New York, Springer-Verlag, 1991, pp. 328-329.
- [2] Diggle, P. a Fisher, N.: Nonparametric Comparison of Cumulative Periodograms, Appl. Statist., 1991, 40, No. 3, pp. 423-434.
- [3] Feller, W.: On the Kolmogorov-Smirnov Limit Theorems for Empirical Distributions, The Annals of Mathematical Statistics, Vol. 19, No. 2. (Jun, 1948), pp. 177-189.
- [4] Katina, S.: Vybrané kapitoly z počítačovej štatistiky I, elektronické študijné materiály, 18/01/2006
- [5] Katina, S.: Štatistická analýza plánovaných experimentov, XVII Summer School of Biometrics, 2006 (Lednice na Morave, Czech Republic).
- [6] Katina, S., Štulajter, F. a Kellerová, E.: Aplikácia kváziperiodických časových radov v kardiológii, Konferencia PRASTAN, 2005, Tajov, SR.
- [7] Štulajter, F.: Predictions in Time Series Using Regression Models, New York, Springer-Verlag, 2002.
- [8] Štulajter, F.: Odhady v náhodných procesoch, Alfa, Bratislava, 1990.
- [9] Zvára, K. a Štěpán, J.: Pravděpodobnost a matematická statistika, Veda, Bratislava, 2002.

Príloha

V prílohe sú funkcie naprogramované v štatistickom programe *R*, ktorý sme používali pri výpočtoch. *R*-ko nemá tieto funkcie zabudované. Simulačné štúdie permutačného a Fisherovho testu sme nenavrhlí ako funkcie. V oboch prípadoch uvádzame ukážky konkrétnych simulácií.

Funkcia *KS*-testu

vstup: rozsah výberov dlzka, stredná hodnota x

výstup: P-hodnota

```
KS.test <-function(dlzka,x){  
  
  a<-0  
  for ( i in 1:1000 ){  
      x1<-rnorm(dlzka,x,1)  
      x2<-rnorm(dlzka,0,1)  
      z<-ks.test(x1,x2)  
  if ( z[2]<=0.05 ) a<-a+1  
      }  
  pvalue <- a/1000  
  pvalue  
      }
```

Simulácia *KS*-testu

vstup: rozsah výberov dlzka, vektor strednych hodnôt X

výstup: vektor P-hodnôt

Funkcia “KS.testvela”využíva funkciu “KS.test”.

```
KS.testvela <-function(dlzka,X){  
  
  n <- length(X)  
  Pval <- rep(0,n)  
  for (i in 1:n){  
      Pval[i] <- KStest(dlzka,X[i])  
  }  
  Pval  
      }
```

Funkcia počítajúca periodogram

vstup: nahodný proces X

výstup: periodogram náhodného procesu X

```
periodogram<-function(X) {  
  
n<-length(X)  
Z<-diag(seq(-3.14+6.28/n,3.14, 6.28/n))  
U<-array(1:n, dim=c(n,n))  
T<-t(U)  
c<-cos(Z%*%T)%*%X  
e<-c*c  
d<-sin(Z%*%T)%*%X  
f<-d*d  
vysledok<-e+f  
vysledok  
  
}
```

Funkcia počítajúca permutácie medzi periodogramami porovnávaných procesov.

vstup: periodogramy procesov X1, X2 (vstup treba vypočítať pomocou funkcie “periodogram”)

výstup: štatistika

```
permutacie<-function( X1, X2 ){  
k<-0  
maximum<-0  
kumulat1<-numeric()  
kumulat2<-numeric()  
rozdiel<-numeric()  
X<-matrix(0, 2, length(X1))  
for ( i in 1:length(X1) ){  
X[(i+k):(i+k+1)]<-sample(c(X1[i],X2[i]))  
}  
kumulat1<-cumsum(t(X)[1:length(X1),1])/sum(t(X)[1:length(X1),1])  
kumulat2<-cumsum(t(X)[1:length(X1),2])/sum(t(X)[1:length(X1),2])  
rozdiel<-abs(kumulat1-kumulat2)  
maximum<-max(rozdiel)  
maximum  
  
}
```

Funkcia permutačného testu

vstup: stacionárne náhodné procesy Y1, Y2 (rovnakého rozsahu)

výstup: P-hodnota

Funkcia `Permutacny.test` využíva funkciu “`periodogram`”.

```
Permutacny.test<-function(Y1, Y2){  
  
  pocet<-0  
  period1<-periodogram(Y1)  
  period2<-periodogram(Y2)  
  statistika<-max(abs(cumsum(period1)/sum(period1)-cumsum(period2)/sum(period2)))  
  for ( i in 1:100 ){  
    a[i]<-permutacie(period1,period2)  
    if (a[i]>=statistika) pocet=pocet+1  
  }  
  pvalue<-1-(100-pocet)/101  
  pvalue  
}
```

Simulácia permutačného testu

Simulácia nie je navrhnutá ako funkcia, využíva funkcie periodogram a Permutácie.

Ukážka simulácie je urobená pre AR(1)proces s koef.=“0.8”a biely šum oba procesy rozsahu “n=100”(nastavenia vo funkcii “arima.sim”), počet simulácií “k=1000”(nastavenie v prvom “for”cykle).

výstup: P-hodnota testu

```
a<-numeric()
pvalue<-numeric()
po<-0
for ( k in 1:1000 ){
  ts1=arima.sim(n=100, model = list(ar = c(0.8), ma = c(0)), sd=1)
  ts2=arima.sim(n=100, model = list(ar = c(0), ma = c(0)), sd=1)
  period1<-periodogram(ts1)
  period2<-periodogram(ts2)
  pocet<-0
  statistika<-max(abs(cumsum(period1)/sum(period1)-cumsum(period2)/sum(period2)))
  for ( i in 1:100 ){
    a[i]<-permutacie(period1,period2)
    if (a[i]>=statistika) pocet=pocet+1
  }
  pvalue[k]<-1-(100-pocet)/101
  if (pvalue[k]<=0.05) po<-po+1
}
po/1000
```

Funkcia počítajúca kombinačné číslo.

vstup: prvý člen kombinačného čísla n, druhý člen kombinačného čísla k

výstup: hodnota kombinačného čísla

```
Kombin<-function(n,k){
x<-numeric()
if(k==0||k==n) x=1
else{
  x<-prod(seq(n,1,-1))/(prod(seq(n-k,1,-1))*prod(seq(k,1,-1)))
}
x
}
```

Funkcia Fisherovho testu

vstup: stacionárny náhodný proces X

výstup: P-hodnota

Funkcia "Fisher.Test" používa funkciu "Kombin".

```
Fisher.Test<-function(X){
n<-length(X)
q<-numeric()
q<-floor((n-1)/2)
x<-numeric()
x<-(max(X)/sum(X))*q
a<-numeric()
b<-numeric()
Prob<-0
for ( i in 1:(q+1) ){
  a[i]<-(max(c((1-(i-1)*x/q),0)))^(q-1)
  b[i]<-(-1)^(i-1)*Kombin(q,i-1)*a[i]
}
Prob<-1-sum(b)
Prob
}
```


Simulácia Fisherovho testu

Simulácia nie je navrhnutá ako funkcia

Simulácia používa funkcie "Fisher.Test" a "periodogram".

Ukážka simulácie je urobená pre proces rozsahu "n=90" (nastavenie vo funkcii "arima.sim") so strednou hodnotou $\cos(t)$, počet simulácií "i=1000" (nastavenie vo "for" cykle)

Výstup: P-hodnota

```
F<-numeric()
po<-0
for (i in 1:1000){
  ts1=arima.sim(n=90, model = list(ar = c(0), ma = c(0)), sd=1)
  ts2<-ts1+4*cos(c(1:length(ts1)))
  P<-periodogram(ts2,length(ts2))
  F[i]<-Fisher_Test(P)
  if (F[i]<=0.05) po<-po+1
}
po/1000
```

Funkcia simulujúca kvantily bieleho šumu

vstup: rozsah bieleho šumu t , pravdepodobnosť daného kvantilu Pr , počet simulácií je pevne daný v tele funkcie "i=1000"(prvý"for"cyklus). výstup: kvantil bieleho šumu.

Funkcia "kvantil"využíva funkciu "periodogram".

```
kvantil<-function(t,pr){
  X<-numeric()
  for ( i in 1:1000 ){
    X[(i*t-(t-1)):i*t]<-(ts1=arima.sim(n=t, model = list(ar = c(0), ma = c(0)), sd=1))
      }
  Y<- matrix(data = X, nrow = 1000, ncol = t, byrow = FALSE, dimnames = NULL)
  P<-numeric()
  kump1<-numeric()
  for ( i in 1:1000 ){
    P[(i*t-(t-1)):i*t]<-periodogram(Y[(i*t-(t-1)):i*t],t)
    kump1[(i*t-(t-1)):i*t]<-cumsum(P[(i*t-(t-1)):i*t])/sum(P[(i*t-(t-1)):i*t])
      }
  Ykump1<-matrix(data = kump1, nrow = 1000, ncol = t, byrow = TRUE, dimnames = NULL)
  Q<-numeric()
  for ( i in 1:t ){
    Q[i]<-quantile(Ykump1[((i*1000)-999):(i*1000)], probs =pr, na.rm = FALSE)
      }
  Q
  }
```

Funkcia simulujúca kvantily skúmaných jedincov v aplikácii

vstup: rozsah bieleho šumu t , pravdepodobnosť daného kvantilu Pr , počet simulácií je pevne daný v tele funkcie "i=1000"(prvý "for"cyklus) okrem toho si treba v tele funkcie vždy nastaviť aj danú strednú hodnotu (premenná "jedinec3"). výstup: kvantil daného jedinca.

Funkcia "kvantil.jedinec"využíva funkciu "periodogram".

```
kvantil.jedinec<-function(t,pr){
X<-numeric()
for ( i in 1:1000 ){
X[(i*t-(t-1)):i*t]<-(ts1=arima.sim(n=t, model = list(ar = c(0), ma = c(0)), sd=1))
}
Y<- matrix(data = X, nrow = 1000, ncol = t, byrow = TRUE, dimnames = NULL)
jedinec3<- 2.879*cos(0.26166*c(1:24))+4.0708*cos(0.785*c(1:24))
-3.611*sin(1.57*c(1:24))+3.503*cos(2.355*c(1:24))
viacjedinec3<-rep(jedinec3,1000)
X2<-X+viacjedinec3
Y2<-matrix(data = X2, nrow = 1000, ncol = t, byrow = TRUE, dimnames = NULL)
P<-numeric()
kump1<-numeric()
for ( i in 1:1000 ){
P[(i*t-(t-1)):i*t]<-periodogram(Y2[(i*t-(t-1)):i*t],t)
kump1[(i*t-(t-1)):i*t]<-cumsum(P[(i*t-(t-1)):i*t])/sum(P[(i*t-(t-1)):i*t])
}
Ykump1<-matrix(data = kump1, nrow = 1000, ncol = t, byrow = TRUE, dimnames = NULL)
Q<-numeric()
for ( i in 1:t ){
Q[i]<-quantile(Ykump1[((i*1000)-999):(i*1000)], probs =pr, na.rm = FALSE)
}
Q
}
```