LATVIJAS UNIVERSITĀTE
FIZIKAS UN MATEMĀTIKAS FAKULTĀTE
MATEMĀTIKAS NODAĻA

ILGLAICĪGĀS ATMIŅAS PROCESI

KURSA DARBS

Autors: **Ilze Rēvalde**

Stud. apl. ir05016

Darba vadītājs: doc. Dr.math. Jānis Valeinis

RĪGA 2009

Anotācija

Darba tēma ir ilglaicīgā atmiņa un ar to saistītie procesi. Apskatīta šī jēdziena izcelsme un aprakstošie modeļi. Definēts Hursta fenomens un frakcionālā Brauna kustība. Aplūkota ilglaicīgās atmiņas modeļu klase FARIMA, to svarīgākās īpašības un parametru novērtēšanas metodes. Ar datorprogrammas R palīdzību simulēti procesi ARMA un ilglaicīgās atmiņas procesi FARIMA.

Atslēgas vārdi: FARIMA, Hursta parametrs, ilglaicīgā atmiņa

Abstract

The subject of this thesis is long memory and related processes. The origin of long memory notion and descriptive models are examined. Hurst phenomenon and fractional Brownian motion are defined. Long memory models FARIMA, the most important properties and estimation methods are considered. Processes ARMA and long memory processes FARIMA with program R are simulated.

Keywords: FARIMA, Hurst parameter, long memory

Saturs

Apzīmējumi

ACF autokorelāciju funkcija

AR autoregresīvs process

ARMA autoregresīvs slīdošā vidējā process

$(B(t))$ Brauna kustība

$\mathcal{C}[\cdot, \cdot]$ nepārtrauktu funkciju telpa

$D[0, 1]$ no labās puses nepārtrauktu, no kreisās puses ierobežotu funkciju telpa

FARIMA frakcionāls autoregresīvs slīdošā vidējā process

fBm frakcionālā Brauna kustība

FCLT funkcionālā centrālā robežteorēma jeb Donskera teorēma

fGn frakcionālais Gausa troksnis

FI frakcionāli integrēts

FN frakcionāls troksnis

H Hursta parametrs

LM, LRD ilglaicīgā atmiņa

MA slīdošā - vidējā process

ML vislielākā ticamība

PACF parciālā autokorelāciju funkcija

R/S statistika

$Var(X)$ dispersija no X

$\Rightarrow$ konverģē

Ievads

Darbā aplūkots ilglaicīgās atmiņas jēdziens un to raksturojoši modeļi. Tēma ir ļoti plaša, tādēļ sniegts ieskats tikai galvenajos raksturlielumos un īpašībās. Pēdējo dekāžu laikā, kopš atklāta ilglaicīgās atmiņas iedarbība, tā ieņem būtisku lomu laikrindu analīzē. Tapuši dažādi materiāli, kas pēta šo fenomenu un izstrādāti to aprakstoši modeļi.
Termins "ilglaicīgā atmiņa" kļuva aktuāls ap 1960. gadu. Protams, neizskaidroti empīriski novērojumi parādījās jau agrāk.
1906. gadā Kairā ieradās Harolds Edvīns Hursts, jauns angļu ierēdnis, kurš darbojās kā hidroloģiskais konsultants. Viņš pētīja sakarības Nīlas ūdens līmeņa svārstībās. Rezultātus publicēja apjomīgos krājumos "Nile Basin", kuru pielikumos bija dokumentēts nokrišņu daudzums, upes līmeņa un straumes izmaiņas 5 gadu ietvaros. Pēdējā no Hursta sējumiem - "The Future Conservation of the Nile"(1946), tikai ierosināts ikgadējos ūdenstilpņu datus "uzglabāt" plašā, ilglaicīgā fondā, kuru viņš nodēvēja par "gadsimta atmiņu". Tās koncepts balstās uz saistību starp amplitūdu vai maksimālo uzkrāto novirzi no vidējās vērtības, straumes standarta novirzi un novērojumu periodu, to aprakstīja ar statistiku R/S. Izrādījās, ka praksē datu uzvedība atšķīrās no teorētiski paredzamās, to dilšanas pakāpe H bija 0.74 sagaidāmās 0.5 vietā. Šī parādība kļuva pazīstama kā *Hursta fenomens*. Bija daudz nesekmīgu mēģinājumu fenomenu izskaidrot, līdz izrādījās, ka svarīgākais aspekts ir neparasti lēnā korelāciju dilšana jeb neparasti ilgstoša atmiņa. Drīz vien iznāca B. Mandelbrota un viņa kolēģu raksti par šo tēmu un citi pētījumi ilglaicīgās atmiņas sfērā.
Stacionārās laikrindās ilglaicīgā atmiņa jeb ilglaicīgā atkarība (turpmāk apzīmēsim - LRD) izpaužas, kā korelāciju tiekšanās uz nulli kā pakāpes funkcijai, turklāt tik lēni, ka to summas diverģē. Salīdzinājumā - laikrindu modeļiem ARMA korelācijas dilst eksponenciāli. LRD modeļu ieviešana nozīmīgi uzlaboja dažādu nozaru procesu analīzi. Mūsdienās tiem ir plašs lietojums hidroloģijā, ekonomikā, finansēs un telekomunikācijās, kā arī citur.
Pirmajā nodaļā plašāk izklāstīts Hursta fenomens, kā arī definēta Brauna kustība. Otrajā nodaļā apskatīti veidi, kā definēt ilglaicīgo atmiņu un definīciju savstarpējā saistība. Trešā nodaļa ir ieskats procesā, kas saistīts ar ilglaicīgo atmiņu un kuru var raksturot ar Hursta parametru - frakcionālajā Brauna kustībā. Ceturtajā nodaļā raksturota visplašāk izmantotā ilglaicīgās atmiņas modeļu klase FARIMA un to svarīgākās īpašības un lielumi. Piektajā nodaļā aplūkoti modeļu FARIMA parametru novērtēšanas veidi. Sestajā nodaļā,

ar datorprogrammas R palīdzību, simulēti FARIMA modeļi un iegūti to grafiskie attēli. Pielikumā A definēti darbā ietvertie, nepaskaidrotie termini un teorēmas.

1. Hursta fenomens

Apskatīsim Hursta fenomena teorētisko aspektu. Statistika, kuru Harolds Hursts piemēroja Nīlas datiem:

$$\frac{R}{S}\left(X_1, ..., X_n\right) = \frac{\max_{0\leq i\leq n}\left(S_i - \frac{i}{n}S_n\right) - \min_{0\leq i\leq n}\left(S_i - \frac{i}{n}S_n\right)}{\left(\frac{1}{n}\sum_{i=1}^{n}\left(X_i - \frac{1}{n}S_n\right)^2\right)^{\frac{1}{2}}}, \tag{1.0.1}$$

kur $X_1, X_2, ..., X_n$ ir izlase no n novērojumiem; $S_m = X_1 + ... + X_m$ ir parciālo summu virkne, visiem $m = 0, 1, \ldots$ un $S_0 = 0$. $\frac{S_n}{n}$ ir izlases vidējā vērība; savukārt starpība $\left(S_i - \frac{i}{n}S_n\right)$ parāda, par cik parciālās summas pārsniedz to izlases vidējo vērtību; maksimuma un minimuma starpība skaitītājā ir starpība starp zemākāko un augstāko paricālo summu pozīciju; dalītājs ir izlases standarta novirze.

Analizēsim statistikas (1.0.1) uzvedību, izmantojot materiālu no G. Samorodnitski grāmatas "Long Range Dependence"[1]. Vispirms definēsim teorēmu, kas būs nepieciešama.

Definīcija 1. [2] Stohastisku procesu no reālām vērtībām $\{B(t) : t \geq 0\}$ sauc par *Brauna kustību* ar sākumu punktā $x \in \mathbb{R}$, ja izpildās

1. $B(0) = x$,
2. procesam eksistē neatkarīgi pieaugumi, t.i., visiem $0 \leq t_1 \leq t_2 \leq \ldots \leq t_n$ pieaugumi $B(t_n) - B(t_{n-1}), B(t_{n-1}) - B(t_{n-2}), \ldots, B(2) - B(1)$ ir neatkarīgi gadījuma lielumi,
3. visiem $t \geq 0$ un $h > 0$ pieaugumi $B(t+h) - B(t)$ ir sadalīti normāli ar matemātisko cerību nulle un dispersiju h,
4. funkcija $t \mapsto B(t)$ ir nepārtraukta gandrīz visur.

Funkciju $\{B(t) : t \geq 0\}$ sauc par standarta Brauna kustību, ja $x = 0$.

Pieņemsim, ka virkne $\{X_n : n \geq 0\}$ ir neatkarīgi, vienādi sadalīti gadījuma lielumi ar

$E(X_n) = 0$ un $Var(X_n) = 1$. Gadījuma klejošana definēta ar $S_n = \sum_{k=1}^{n} X_k$ ir lineāri interpolējama veselās daļās, t.i.,

$$S(t) = S_{[t]} + (t - [t])(S_{[t]+1} - S_{[t]}).$$

Tādejādi ir uzdota gadījuma funkcija $S \in \mathcal{C}[0, \infty)$. Definējam gadījuma funkciju virkni $\{S_n^* : n \geq 1\}$ apgabalā $\mathcal{C}[0, 1]$ ar

$$S_n^*(t) = \frac{S(nt)}{\sqrt{n}}$$

visiem $t \in [0, 1]$.

Teorēma 1. <u>*Donskera invariācijas princips*</u> *(Funkcionālā centrālā robežteorēma). [2] Nepārtrauktu funkciju telpā $\mathcal{C}[0, 1]$ virkne $\{S_n^* : n \geq 1\}$ pēc sadalījuma konverģē uz standarta Brauna kustību $\{B(t) : t \in [0, 1]\}$*

Pieņemsim, ka funkciju $f : D[0, 1] \mapsto \mathbb{R}$ definē kā

$$f(x) = \sup_{0 \leq t \leq 1} (x(t) - tx(1)) - \inf_{0 \leq t \leq 1} (x(t) - tx(1)). \tag{1.0.2}$$

Redzams, ka $x = (x(t), 0 \leq t \leq 1) \in D[0, 1]$ ir nepārtraukta. Pielietosim šo funkciju parciālo summu virknei ar vērtībām no apgabala $D[0, 1]$ - parciālo summu procesam. Pieņem, ka $X_1, X_2, \ldots$ ir stacionāra gadījumu lielumu virkne ar ierobežotu dispersiju un vidējo vērtību μ. Parciālo summu procesu definē

$$S^{(n)}(t) = S_{[nt]} - [nt]\mu, \tag{1.0.3}$$

visiem $0 \leq t \leq 1$. Pēc FCLT varam secināt, ka, ja $X_1, X_2, \ldots$ ir neatkarīgi, vienādi sadalīti gadījuma lielumi, tad

$$\frac{1}{\sqrt{n}} S^{(n)} \Rightarrow \sigma_* B \tag{1.0.4}$$

vājā nozīmē telpā $D[0, 1]$, kur σ_*^2 ir novērojumu dispersija σ^2 un B ir standarta Brauna kustība intervālā [0,1].

Tad formulas (1.0.1) skaitītāju varam izteikt kā $f\left(S^{(n)}\right)$. Tādēļ, ja FCLT izpildās, pēc nepārtrauktā attēlojuma teorēmas (Skat. pielikumu) iegūstam, ka

$$\frac{1}{\sqrt{n}} \left(\max_{0 \leq i \leq n} \left(S_i - \frac{i}{n} S_n \right) - \min_{0 \leq i \leq n} \left(S_i - \frac{i}{n} S_n \right) \right) = f\left(\frac{1}{\sqrt{n}} S^{(n)} \right)$$

$$\Rightarrow f(\sigma_* B) = \sigma_* \left[\sup_{0\le t\le 1} (B(t) - tB(1)) - \inf_{0\le t\le 1} (B(t) - tB(1)) \right]$$

$$:= \sigma_* \left[\sup_{0\le t\le 1} B_0(t) - \inf_{0\le t\le 1} B_0(t) \right],$$

kur B_0 apzīmē Brauna tiltu intervālā [0,1]. Turklāt, ja stacionāra virkne $X_1, X_2, \ldots$ ir ergodiska, tad izlases standratnovirze ir būtisks populācijas standartnovirzes novērtējums. Tādēļ

$$\left(\frac{1}{n} \sum_{i=1}^{n} \left(X_i - \frac{1}{n} S_n \right)^2 \right)^{1/2} \to \sigma$$

ar varbūtību 1. Visbeidzot varam secināt, ka

$$\frac{1}{\sqrt{n}} \frac{R}{S} (X_1, \ldots, X_n) \Rightarrow \frac{\sigma_*}{\sigma} \left[\sup_{0\le t\le 1} B_0(t) - \inf_{0\le t\le 1} B_0(t) \right]. \tag{1.0.5}$$

Tas nozīmē, ka R/S statistika pieaug kā kvadrātsakne no izlases apjoma. Tomēr, kad Harolds Hursts aprēķināja R/S statistiku Nīlas datiem, atklājās, ka empīriskais pieaugums ir tuvāks $n^{0.74}$. Šo fenomenu sauc par *Hursta fenomenu.* Izrādījās, ka atrast piemērotu stohastisku modeli, kas izskaidrotu šo parādību, nav vienkārši. Tika izvirzīta hipotēze, ka jāņem modelis, kuram korelācijas dilst lēni un tādēļ nav spēkā FCLT. Vienkāršākais no šādiem modeļiem ir frakcionālais Gausa troksnis, kuram izpildās

$$n^{-H} \frac{R}{S} (X_1, ..., X_n) \Rightarrow \frac{1}{\sigma} \left[\sup_{0\le t\le 1} (B_H(t) - tB_H(1)) - \inf_{0\le t\le 1} (B_H(t) - tB_H(1)) \right],$$

redzams, ka R/S statistika aug kā n^H. Tāpēc, izvēloties atbilstošu H, kuru dēvē par *Hursta parametru*, beidzot bija iespējams izskaidrot Hursta fenomenu. Nonāca pie secinājuma, ka pats svarīgākais faktors ir neparasti lēnā korelāciju dilšana, sevišķi pie lielām H vērtībām - neparasti ilglaicīgā atmiņa. Līdz ar to tika ieviesti tādi termini kā ilglaicīgā atmiņa un ilglaicīgā atkarība.

Parametrs H pieņem vērtības no 0.5 līdz 1. Jebkuram gadījuma procesam bez LRD $H = 0.5$, tomēr praksē daudziem reāliem procesiem $H \approx 0.73$. Jo Hursta parametrs tuvāks 1, jo lielāka LRD pakāpe un sarežģītāka procesa struktūra.

2. Ilglaicīgās atmiņas procesi

Šajā nodaļā apskatīsim veidus, kā definēt ilglaicīgo atmiņu un pierādīsim teorēmu par definīciju ekvivalenci pie īpašiem nosacījumiem, izmantojot W. Palmas grāmatu "Long - Memory Time Series: Theory and Methods"[3].

Definīcija 2. Pieņem, ka $\gamma(h) = \langle y_t, y_{t+h} \rangle$ ir autokovariāciju funkcija stacionāram procesam $\{y_t : t \in \mathbb{Z}\}$. Ilglaicīgo atmiņu definē

$$\sum_{h=-\infty}^{\infty} |\gamma(h)| = \infty. \tag{2.0.1}$$

Izmanto arī citas alternatīvas definīcijas. Lai tās saprastu, vispirms definēsim dažus jēdzienus.

Definīcija 3. Pozitīvu mērojamu funkciju, definētu kādā bezgalības apkārtnē $[a, \infty)$, sauc par *lēni variējošu* Karamata izpratnē tad un tikai tad, ja katram $c > 0$ izpildās $\dfrac{l(cx)}{l(x)} \to 1$, kad $x \to \infty$.

Lēni variējošas funkcijas piemēri ir $l(x) = \log(x)$ un $l(x) = b$, kur b ir pozitīva konstante. Funkcija $l(x) = x^\beta$ nav lēni variējoša nevienam reālam $\beta \neq 0$.

Teorēma 2. *Volda dekompozīcijas teorēma: Jebkuru stacionāru, nedeterminētu procesu var izteikt*

$$y_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j} = \psi(B)\varepsilon_t, \tag{2.0.2}$$

kur $\psi_0 = 1$, $\sum_{j=0}^{\infty} \psi_j^2 < \infty$, ε_t *ir baltais troksnis ar dispersiju* σ^2.

Ilglaicīgo atmiņu definē

$$\gamma(h) \sim h^{2d-1} l_1(h), \tag{2.0.3}$$

$$f(\lambda) \sim |\lambda|^{-2d} l_2 \left(\frac{1}{|\lambda|} \right), \tag{2.0.4}$$

$$\psi_j \sim j^{d-1} l_3(j), \tag{2.0.5}$$

kur $j > 0$, $\lambda \in 0$ apkārtnei, $l_1(\cdot), l_2(\cdot), l_3(\cdot)$ ir lēni variējošas funkcijas, d ir tā dēvētais *ilglaicīgās atmiņas parametrs* un apzīmējums $x_n \sim y_n$ jāsaprot kā $x_n/y_n \to 1$, kad $n \to \infty$. Parametru d un Hursta parametru saista sakarība $H = d + \frac{1}{2}$. Definīcija (2.0.3) balstās uz autokovariāciju dilšanu, kad $h \to \infty$, plaši izmantotā (2.0.4) izsaka LRD caur spektrālajām īpašībām un (2.0.5) balstās uz procesa Volda dekompozīciju.

Definīcijas (2.0.1) - (2.0.5) ir ekvivalentas tikai pie īpašiem nosacījumiem, kurus apskatīsim sekojošā teorēmā. Definēsim divas lemmas, kas nepieciešamas pierādījumam.

Lemma 3. *Pieņemsim, ka $r, s \in \mathbb{R}$ un $s < 1 < r + s; l_1, l_2$ ir lēni variējošas funkcijas un $\{f_k\}, \{g_k\}$ ir divas virknes, kuras pieder kopai $\mathbb{N}$, kuras apmierina $f_k \sim k^{-r} l_1(k), g_k \sim k^{-s} l_2(k)$, kad $k \to \infty$. Jebkuram $n \in \mathbb{N}$, virkne $\{f_{n+k} gk\}$ ir summējama un*

$$\sum_{k=0}^{\infty} f_{n+k} g_k \sim n^{-(r+s-1)} l_1(n) l_2(n) B(r+s-1, 1-s),$$

kad $n \to \infty$, kur $B(\cdot, \cdot)$ ir beta funkcija.

Lemma 4. *(a) Ja $l(\cdot)$ ir lēni variējoša un $\alpha > 0$, tad $x^{\alpha} l(x) \to \infty$, kad $x \to \infty$.*

(b) Ja $l(\cdot)$ ir lēni variējoša, m ir pietiekoši liels, tāds, ka $l(x)$ ir lokāli ierobežota apgabalā $[m, \infty)$ un $\alpha > -1$, tad

$$\sum_{k=m}^{n} k^{\alpha} l(k) \sim \frac{n^{\alpha+1} l(n)}{\alpha + 1},$$

kad $n \to \infty$.

(c) Ja $l(\cdot)$ir kvazi - monotona lēni variējoša, tad sekojošā rinda ir nosacīti konverģenta

$$\sum_{k=1}^{\infty} l(k) \cos(k\lambda) k^{-\mu} \sim \frac{\pi \lambda^{\mu-1} l(\frac{1}{\lambda})}{2\Gamma(\mu) \cos(\frac{\pi\mu}{2})},$$

visiem $\lambda \to 0_+$ un $0 < \mu < 1$.

Teorēma 5. *Pieņem, ka $\{y_t\}$ ir stacionārs process Volda izteiksmē (2.0.2), $0 < d < \frac{1}{2}$, tad*

(a) Ja procesam $\{y_t\}$ izpildās (2.0.5), tad izpildās arī (2.0.3).

(b) Ja procesam $\{y_t\}$ izpildās (2.0.3), tad izpildās arī (2.0.1).

(c) Ja $l_1(\cdot)$ ir kvazi - monotona lēni variējoša funkcija, tad (2.0.3) nozīmē (2.0.4).

Pierādījums. a) No Volda dekompozīcijas (2.0.2) seko, ka procesa $\{y_t\}$ autokovariāciju funkciju var uzrakstīt formā

$$\gamma(h) = \sigma^2 \sum_{k=0}^{\infty} \psi_k \psi_{k+|h|}.$$

Tādēļ, pēc nosacījuma (2.0.5) un Lemmas 3, mēs varam secināt, ka

$$l_1(h) = \sigma^2 l_3(h)^2 B(1-2d, d).$$

b) Jebkuram veselam skaitlim $0 < m < n$

$$\sum_{h=-\infty}^{\infty} |\gamma(h)| \geq \sum_{h=m}^{n} |\gamma(h)|.$$

Tagad, no un Lemmas 4 (b) pie lieliem n iegūsim

$$\sum_{h=m}^{n} |\gamma(h)| \sim \frac{n^{2d} l_1(n)}{2d}.$$

Tā kā $d > 0$, tad no Lemmas 4 seko, ka $n^{2d} l_1(n) \to \infty$, kad $n \to \infty$. Tas nozīmē, ka nosacījums (2.0.3) apmierina nosacījumu (2.0.1).

c) Tā kā l_1 ir kvazi - monotona lēni variējoša, tad no Lemmas 4 punkta (c) seko, ka (2.0.4):

$$l_2(\lambda) = \frac{l_1(\lambda)}{2\Gamma(1-2d)\sin(\pi d)} = \frac{\Gamma(1-d)\Gamma(d)}{2\pi\Gamma(1-2d)} l_1(\lambda)$$

visiem $\lambda > 0$. $\square$

3. Frakcionālā Brauna kustība

Saistībā ar LRD un Hursta parametru, apskatīsim procesu, kuru dēvē par frakcionālo Brauna kustību (turpmāk lietosim saīsinājumu fBm). Materiāls no grāmatas "Theory and Applications of Long Range Dependence" [4].

Definīcija 4. Stohastisku procesu $Z = \{Z(t)\}_{t\in\mathbb{R}}$, kas pieņem reālas vērtības, sauc par *sev līdzīgu* ar indeksu $H > 0$, visiem $a > 0$, ja

$$\{Z(at)\}_{t\in\mathbb{R}} =^d \left\{a^H Z(t)\right\}_{t\in\mathbb{R}}, \tag{3.0.1}$$

kur ar $=^d$ apzīmē galīgi dimensionālu sadalījumu ekvivalenci. Procesu (3.0.1) turpmāk dēvēsim par H - se procesu. Kaut arī tie nav stacionāri, starp sev līdzīgiem un satcionāriem procesiem pastāv zināma sakarība.

Apgalvojums 6. *Ja process $\{X(t)\}_{t>0}$ ir H - se, tad process $Y(t) = e^{-tH}X(e^t), t \in \mathbb{R}$, ir stacionārs. Kā arī otrādi, ja $\{Y(t)\}_{t\in\mathbb{R}}$ ir stacionārs, tad $X(t) = t^H Y(\ln t), t > 0$, ir H - se process.*

Definīcija 5. Saka, ka procesam $Z = \{Z(t)\}_{t\in\mathbb{R}}$ ir *stacionāri pieaugumi*, ja, visiem $h \in \mathbb{R}$,

$$\{Z(t+h) - Z(h)\}_{t\in\mathbb{R}} =^d \{Z(t) - Z(0)\}_{t\in\mathbb{R}}.$$

Procesus H - se plaši izmanto, jo no tiem var izveidot virknes ar svarīgām īpašībām. Ja $\{Z(t)\}_{t\in\mathbb{R}}$ ir H - ss process, tad

$$X_j = \triangle Z(j) = Z(j+1) - Z(j), j \in \mathbb{Z},$$

ir stacionāra virkne, saukta par laikrindām. H - se procesam $Z = \{Z(t)\}_{t\in\mathbb{R}}$ ar ierobežotu dispersiju piemīt sekojošas īpašības:

- $Z(0) = 0$.

- Ja $H \neq 1$, tad $EZ(t) = 0$ visiem $t \in \mathbb{R}$.

- $Z(-t) = -Z(t)$.

- $EZ^2(t) = EZ^2(|t| \text{ sign} t) = |t|^{2H} EZ^2(\text{sign} t) = |t|^{2H} EZ^2(1) = |t|^{2H}\sigma^2$. Ja $\sigma^2 = EZ^2(1) = 1$, tad procesu $Z = \{Z(t)\}_{t\in\mathbb{R}}$ sauc par standarta.

- H-ss procesa Z kovariāciju funkcija ir

$$\begin{aligned} Cov(Z(s), Z(t)) &= \frac{1}{2}\left\{EZ^2(s) + EZ^2(t) - E(Z(t) - Z(s))^2\right\} \\ &= \frac{\sigma^2}{2}\left\{|t|^{2H} + |s|^{2H} - |t-s|^{2H}\right\}. \end{aligned} \tag{3.0.2}$$

- Procesa Z parametrs $H \leq 1$.

Ja $0 < H \leq 1$, tad kovariāciju funkcija ir nenegatīvi definita.

Definīcija 6. Gausa H - se procesu $\{B_H(t)\}_{t\in\mathbb{R}}$, kur $0 < H < 1$, sauc par *frakcionālo Brauna kustību.* To dēvē par standarta fBm, ja $\sigma^2 = Var B_H(1) = 1$.

Piezīme 7. *Ja* $H = \frac{1}{2}$, *tad fBm* $\left\{B_{\frac{1}{2}}(t)\right\}_{t\in\mathbb{R}}$ *ir parastā Brauna kustība (Definīcija 1) ar* $Cov(B(s), B(t)) = EB_{\frac{1}{2}}(s)B_{\frac{1}{2}}(t) = \min(|s||t|)$, *ja* s *un* t *ir ar vienādām zīmēm un* $Cov(B(s), B(t)) = 0$, *ja zīmes ir pretējas.*

Apgalvojums 8. *Pieņemsim, ka* $\{X(t)\}_{t\in\mathbb{R}}$

1. *ir Gausa process ar vidējo vērtību* 0, $X(0) = 0$,

2. $EX^2(t) = \sigma^2|t|^{2H}$ *kādam* $\sigma > 0$ *un* $0 < H < 1$,

3. *ir stacionāri pieaugumi;*

tad $\{X(t)\}_{t\in\mathbb{R}}$ *ir frakcionālā Brauna kustība.* Frakcionāla Brauna kustība ir vienīgais Gausa H - se process.

Vēl viens labi zināms LRD process ir *frakcionālais Gausa troksnis* (turpmāk lietosim saīsinājumu fGn), kas būtībā ir diskrētu soļu pieaugumu process no fBm.

Definīcija 7. Ja $\{Z(t)\}_{t\in\mathbb{R}}$ ir fBm, tad procesu $\{X_k\}_{k\in\mathbb{Z}}$ sauc par frakcionālo Gausa troksni, kuram izpildās

$$X_k = Z(k+1) - Z(k), k \in \mathbb{Z}.$$

Apgalvojums 9. *Augošai virknei $\{X_k\}_{k\in\mathbb{Z}}$ (Definīcija 7) piemīt sekojošas īpašības:*

1. *$\{X_k\}_{k\in\mathbb{Z}}$ ir stacionāra.*

2. *$EX_k = 0$.*

3. *$EX_k^2 = \sigma^2 = EZ(1)^2$.*

4. *Procesa $\{X_k\}_{k\in\mathbb{Z}}$ ACF ir dota ar*

$$\gamma(k) = EX_iX_{i+k} = \frac{\sigma^2}{2}(|k+1|^{2H} - 2|k|^{2H} + |k-1|^{2H}) = \frac{\sigma^2}{2}\triangle^2|k|^{2H},$$

   kur $\triangle^2$ apzīmē otrās kārtas diferenci.

5. *Pieņem, ka $k \neq 0$. Tad $\gamma(k) = 0$, ja $H = \frac{1}{2}$, $\gamma(k) < 0$, ja $0 < H < \frac{1}{2}$ un $\gamma(k) > 0$, ja $\frac{1}{2} < H < 1$.*

6. *Ja $H \neq \frac{1}{2}$, tad*

$$\gamma(k) \sim \sigma^2 H(2H-1)|k|^{2H-2},$$

   kad $k \to \infty$.

4. FARIMA modeļi

Plaši pazīstama ilglaicīgās atmiņas modeļu klase ir frakcionālie autoregresīvie slīdošā - vidējā jeb *FARIMA* (dažkārt arī - *ARFIMA*) modeļi. Būtībā tie ir ARMA procesu vispārinājums, tikai FARIMA var piemist arī ilglaicīgā atmiņa.

Definīcija 8. Procesu $\{y_t\}$ sauc par FARIMA (p, d, q) procesu, ja

$$\phi(B)y_t = \theta(B)(1-B)^{-d}\varepsilon_t, \tag{4.0.1}$$

kur $\phi(B) = 1 + \phi_1 B + \ldots + \phi_p B^p$ un $\theta(B) = 1 + \theta_1 B + \ldots + \theta_q B^q$ ir attiecīgi autoregresīvā un slīdošā vidējā operatori,$(1-B)^{-d}$ ir frakcionālais diferenču operators definēts ar binominālo izvirzījumu

$$(1-B)^{-d} = \sum_{j=0}^{\infty} \eta_j B^j = \eta(B),$$

kur

$$\eta_j = \frac{\Gamma(j+d)}{\Gamma(j+1)\Gamma(d)}, \tag{4.0.2}$$

visiem $d < \frac{1}{2}$, $d \neq 0, -1, -2, \ldots$, un $\{\varepsilon_t\}$ ir baltais troksnis ar ierobežotu dispersiju. Vispārīgāk varam definēt

$$\eta(z) = (1-z)^{-d},$$

tad šī funkcija ir analītiska vaļējā riņķī $\{z : |z| < 1\}$, kā arī slēgtā vienības riņķī $\{z : |z| \leq 1\}$ negatīviem d. Šo apstākļu dēļ varam uzrakstīt $\eta(\cdot)$ ar Teilora rindas izvirzījumu

$$\eta(z) = \sum_{j=0}^{\infty} \eta_j z^j.$$

Teorēma 10. *Pieņemam, ka process FARIMA definēts ar* (4.0.1)*; polinomiem* $\phi(\cdot)$ *un* $\theta(\cdot)$ *nav kopīgu sakņu un* $d \in \left(-1, \frac{1}{2}\right)$*. Tad*

(a) Ja $\phi(\cdot)$ *saknes atrodas ārpus vienības riņķa* $\{z : |z| = 1\}$*, tad eksistē* *viens vienīgs* (4.0.1) *atrisinājums*

$$y_t = \sum_{j=-\infty}^{\infty} \psi_j \varepsilon_{t-j},$$

kur $\psi_z = (1-z)^{-d}\dfrac{\theta(z)}{\phi(z)}$.

(b) Ja $\theta(\cdot)$ *saknes atrodas ārpus slēgta vienības riņķa* $\{z : |z| \leq 1\}$, *tad atrisinājums ir apgriežams.*

Piezīme 11. *Procesu FARIMA* (p, d, q) *mēdz uzdot arī kā*

$$\phi(B)(1-B)^d y_t = \theta(B)\varepsilon_t. \tag{4.0.3}$$

Tomēr vienādojuma (4.0.3) atrisinājums var nebūt viens vienīgs. Piemēram, ja $\{y_t\}$ ir (4.0.3) stacionārs atrisinājums ar $d > 0$ un pieņem, ka v ir gadījuma lielums ar ierobežotu dispersiju, tad stacionārs process $x_t = y_t + v$ arī ir (4.0.3) atrisinājums.

Bezgalīga AR un MA reprezentācija

Saskaņā ar Teorēmu 11, ņemot vērā, ka polinomu $\phi(B)$ un $\theta(B)$ saknes atrodas ārpus slēgta vienības riņķa $\{z : |z| \leq 1\}$ un $d \in \left(-1, \frac{1}{2}\right)$, process FARIMA (p, d, q) ir stacionārs un apgriežams. Tādā gadījumā varam pierakstīt

$$y_t = (1-B)^{-d}\phi(B)^{-1}\theta(B)\varepsilon_t = \psi(B)\varepsilon_t,$$

un

$$\varepsilon_t = (1-B)^{d}\phi(B)\theta(B)^{-1}y_t = \pi(B)y_t.$$

MA (∞) koeficienti, ψ_j, un AR (∞) koeficienti, π_j, apmierina sekojošas asimptotiskas sakarības:

$$\psi_j = \frac{\theta(1)j^{d-1}}{\phi(1)\Gamma(d)} + \mathcal{O}(j^{-1}), \tag{4.0.4}$$

$$\pi_j = \frac{\phi(1)j^{-d-1}}{\theta(1)\Gamma(-d)} + \mathcal{O}(j^{-1}), \tag{4.0.5}$$

kad $j \to \infty$.

Frakcionāliem trokšņa procesiem ar ilglaicīgās atmiņas parametru d attiecīgie koeficienti ir doti ar vienādojumiem

$$\psi_j = \prod_{t=1}^{j} \frac{t-1+d}{t} = \frac{\Gamma(j+d)}{\Gamma(d)\Gamma(j+1)}, \tag{4.0.6}$$

$$\pi_j = \prod_{t=1}^{j} \frac{t-1-d}{t} = \frac{\Gamma(j-d)}{\Gamma(-d)\Gamma(j+1)}, \tag{4.0.7}$$

visiem $j \geq 1$ un $\psi_0 = \pi_0 = 1$.

Spektrālais blīvums

Ja izpildās Teorēmas 11 nosacījumi, tad procesa (4.0.1) spektrālo blīvumu var pierakstīt sekojošā veidā

$$f(\lambda) = \frac{\sigma^2}{2\pi}|1 - e^{-i\lambda}|^{-2d}\frac{|\theta(e^{-i\lambda})|^2}{|\phi(e^{-i\lambda})|^2} = \frac{\sigma^2}{2\pi}\left[2\sin\frac{\lambda}{2}\right]^{-2d}\frac{|\theta(e^{-i\lambda})|^2}{|\phi(e^{-i\lambda})|^2}, \tag{4.0.8}$$

un tas ir

$$f(\lambda) \sim \frac{\sigma^2}{2\pi}\frac{|\theta(1)|^2}{|\phi(1)|^2}|\lambda|^{-2d}, \tag{4.0.9}$$

visiem $|\lambda| \to 0$.

Autokovariāciju funkcija

Procesam FARIMA $(0, d, 0)$ autokovariāciju funkcija ir

$$\gamma_0(h) = \sigma^2\frac{\Gamma(1-2d)}{\Gamma(1-d)\Gamma(d)}\frac{\Gamma(h+d)}{\Gamma(1+h-d)}, \tag{4.0.10}$$

savukārt autokorelāciju funkcija izsakāma

$$\rho_0(h) = \frac{\Gamma(1-d)}{\Gamma(d)}\frac{\Gamma(h+d)}{\Gamma(1+h-d)}. \tag{4.0.11}$$

Vispārīgā gadījumā procesam FARIMA (p, d, q) ACF

$$\gamma(h) = \sigma^2\sum_{i=-q}^{q}\sum_{j=1}^{p}\psi(i)\xi_j C(d, p+i-h, \rho_j), \tag{4.0.12}$$

kur

$$\psi(i) = \sum_{k=\max(0,i)}^{\min(q,q+i)}\theta_k\theta_{k-i},$$

$$\xi_j = \left[\rho_j\prod_{i=1}^{p}(1-\rho_i\rho_j)\prod_{m\neq j}(\rho_j-\rho_m)\right]^{-1},$$

$$C(d, h, \rho) = \frac{\gamma_0(h)}{\sigma^2}[\rho^{2p}\beta(h) + \beta(-h) - 1],$$

$$\beta(h) = F(d+h, 1, 1-d+h, \rho)$$

un $F(a, b, c, x) = 1 + \frac{a\cdot b}{\gamma\cdot 1}x + \frac{a\cdot(a+1)\cdot b\cdot(b+1)}{\gamma\cdot(\gamma+1)\cdot 1\cdot 2}x^2 + \ldots$.
To varam izteikt

$$\gamma(h) \sim c_\gamma|h|^{2d-1}, \tag{4.0.13}$$

kad $|h| \to \infty$ un $c_\gamma = \frac{\sigma^2}{\pi} \frac{|\theta(1)|^2}{|\phi(1)|^2} \Gamma(1-2d)\sin(\pi d)$.

Izlases vidējā vērtība

Pieņam, ka $y_1, y_2, \ldots, y_n$ ir izlase no procesa FARIMA(p,d,q) un $\bar{y}$ ir izlases vidējā vērtība. Tad lieluma $\bar{y}$ dispersija ir

$$Var(\bar{y}) = \frac{1}{n}\left[2\sum_{j=1}^{n-1}\left(1-\frac{j}{n}\right)\gamma(j)+\gamma(0)\right].$$

No (4.0.13) seko, ka $\gamma(j) \sim c_\gamma j^{2d-1}$ pie lieliem j. Tādēļ, pie lielām n vērtībām, iegūsim

$$\begin{aligned}
Var(\bar{y}) &\sim 2c_\gamma n^{2d-1}\sum_{j=1}^{n-1}\left(1-\frac{j}{n}\right)\left(\frac{j}{n}\right)^{2d-1}\frac{1}{n} \\
&\sim 2c_\gamma n^{2d-1}\int_0^1 (1-t)t^{2d-1}dt \\
&\sim \frac{c_\gamma}{d(2d+1)}n^{2d-1}.
\end{aligned}$$

Iegūtais rezultāts parāda FARIMA procesu izlases vidējās vērtības asimptotisko uzvedību.

Parciālās autokorelācijas

Precīzi formulētas parciālās autokorelācijas funkcijas (PACF) vispārīgam FARIMA modelim atrast ir sarežģīti. Labākās lineārās prognozes

$$\hat{y}_{n+1} = \phi_{n1}y_n + \ldots + \phi_{nn}y_1,$$

koeficienti frakcionālam trokšņa procesam ir

$$\phi_{nj} = -\begin{pmatrix} n \\ j \end{pmatrix}\frac{\Gamma(j-d)\Gamma(n-d-j+1)}{\Gamma(-d)\Gamma(n-d+1)}$$

visiem $j = 1, \ldots, n$. Tādēļ varam vienkārši izteikt parciālās autokorelācijas

$$\phi_{nn} = \frac{d}{n-d},$$

un $\phi_{nn} \sim \frac{d}{n}$ pie lieliem n. Neskatoties uz precīzas PACF formulas trūkumu vispārīgā FARIMA gadījumā, Inoue(2002) ir pierādījis, ka parciālo autokorelāciju absolūtā vērtība asimptotiski uzvedas līdzīgi kā frakcionālā trokšņa gadījumā: Ja ϕ_{nn} ir procesa FARIMA, kur $d \in \left(0, \frac{1}{2}\right)$, parciālās autokorelācijas, tad

$$|\phi_{nn}| \sim \frac{d}{n},$$

kad $n \to \infty$.

Ilglaicīgās atmiņas procesu aproksimācija

Jebkuru ilglaicīgās atmiņas procesu, kas uzdots ar (2.0.4), var labi aproksimēt ar kādu FARIMA modeli, kā redzams sekojošā teorēmā:

Teorēma 12. *Pieņem, ka* $\{y_t : t \in \mathbb{Z}\}$ *ir lineārs process, kurš apmierina* (2.0.2). f_y *ir stingri pozitīvs spektrālais blīvums, kas apmierina* (2.0.4). *Tad eksistē kāds FARIMA process ar spektrālo blīvumu* f, *tādu, ka visiem* $\varepsilon > 0$

$$|\frac{f_y(\lambda)}{f(\lambda)} - 1| < \varepsilon,$$

vienmērīgi visiem $\lambda \in [-\pi, \pi]$.

5. Parametru novērtēšanas metodes

Šajā nodaļā gūsim vispārīgu priekšstatu par dažām metodēm, ar kurām iespējams novērtēt LRD modeļu parametrus, kā arī noteikt, vai modelis satur ilglaicīgo atmiņu.

5.1. Maksimālās ticamības novērtējums

Pieņem, ka $\{y_t\}$ ir Gausa process ar vidējo vērtību nulle. Tad procesa log - ticamības funkcija ir dota

$$\mathcal{L}(\theta) = -\frac{1}{2}\log\det\Gamma_\theta - \frac{1}{2}y'\Gamma_\theta^{-1}y, \tag{5.1.1}$$

kur $y = (y_1, \ldots, y_n)'$, $\Gamma_\theta = \mathrm{Var}(y)$ un θ ir parametru vektors. Rezultātā vislielākās ticamības (ML) novērtējumu $\hat{\theta}$ iegūst maksimizējot $\mathcal{L}(\theta)$. Izmantojot log - ticamības funkciju (5.1.1), nepieciešams aprēķināt dispersiju - kovariāciju inverso matricu Γ_θ un determinantu, to var izdarīt ar Čoleskija dekompozīcijas metodi. Apakšnodaļas turpinājumā apskatīsim metodes, ar kurām aprēķināt funkciju (5.1.1).

Teorēma 13. *Pieņemsim, ka $\hat{\theta}_n$ ir vērtība, kas maksimizē log - ticamības funkciju, kur*

$$\theta = (\phi_1, \ldots, \phi_p, \theta_1, \ldots, \theta_q, d)'$$

ir $p + q + 1$ dimensionāls parametru vektors un ka θ_0 ir parametra īstā vērtība. Tad
(a) $\hat{\theta}_n \to \theta_0$ pēc varbūtības, kad $n \to \infty$.
(b) Centrālā robežteorēma: $\sqrt{n}(\hat{\theta}_n - \theta_0) \to N(0, \Gamma^{-1}(\theta_0))$, kad $n \to \infty$, kur $\Gamma(\theta) = (\Gamma_{ij}(\theta))$ ar

$$\Gamma_{ij}(\theta) = \frac{1}{4}\int_{-\pi}^{\pi}\left[\frac{\partial \log f_\theta(\lambda)}{\partial \theta_i}\right]\left[\frac{\partial \log f_\theta(\lambda)}{\partial \theta_j}\right]d\lambda,$$

kur f_θ ir procesa spektrālais blīvums.
(c) $\hat{\theta}_n$ ir efektīvs parametra θ_0 novērtējums.

Čoleskija dekompozīcijas metode

Tā kā matrica Γ_θ ir simetriski pozitīvi definita, varam rakstīt

$$\Gamma_\theta = U'U,$$

kur U ir augšējā trijstūra matrica. Atsaucoties uz iepriekšminēto Čoleskija dekompozīciju, matricas Γ_θ determinantu izsaka $\det \Gamma_\theta = (\det U)^2 = \prod_{j=1}^{n} u_{jj}^2$, kur u_{jj} ir matricas U j - tās diagonāles elements. Savukārt inverso matricu no Γ_θ var iegūt $\Gamma_\theta^{-1} = U^{-1}\left(U^{-1}\right)'$, kur U inverso matricu aprēķina no vidējām vērtībām pēc kādas ļoti vienkāršas procedūras. Čoleskija algoritma aprēķinu kārta ir $\mathcal{O}\left(n^3\right)$.

Durbina - Levina algoritms

Čoleskija dekompozīcijas metode pie garām laikrindām var būt neefektīva, tādēļ tika attīstītas ātrākas metodes log - ticamības funkcijas aprēķināšanai.

Pieņemsim, ka $\hat{y}_1 = 0$ un $\hat{y}_{t+1} = \phi_{t1}y_t + \ldots + \phi_{tt}y_1, t = 1, \ldots, n-1$ ir procesa $\{y_t\}$ prognoze vienam solim uz priekšu, balstīta uz galīga skaita pagātni $\{y_1, \ldots, y_{t-1}\}$, kur regresijas koeficienti ϕ_{tj} doti ar vienādojumiem

$$\phi_{tt} = [\nu_{t-1}]^{-1}\left[\gamma(t) - \sum_{i=1}^{t-1}\phi_{t-1,i}\gamma(t-i)\right],$$

$$\phi_{tj} = \phi_{t-1,j} - \phi_{tt}\phi_{t-1,t-j},$$

$$\nu_0 = \gamma(0),$$

$$\nu_t = \nu_{t-1}\left[1 - \phi_{tt}^2\right],$$

kur $j = 1, \ldots, t-1$. Turklāt, ja $e_t = y_t - \hat{y}_t$ ir prognozes kļūda un $e = (e_1, \ldots, e_n)'$, tad $e = Ly$, kur L ir apakšējā trijstūra matrica:

$$L = \begin{pmatrix} 1 & & & & & \\ -\phi_{11} & 1 & & & & \\ -\phi_{22} & -\phi_{21} & 1 & & & \\ -\phi_{33} & -\phi_{32} & -\phi_{31} & & & \\ \vdots & \vdots & & & 1 & \\ -\phi_{n-1,n-1} & -\phi_{n-1,n-2} & -\phi_{n-1,n-3} & \cdots & -\phi_{n-1,1} & 1 \end{pmatrix}$$

Tādēļ varam rakstīt $\Gamma_\theta = LDL'$, kur D =diag$(\nu_0, \dots, \nu_{n-1})$. Tāpēc, $\det \Gamma_\theta = \prod_{j=1}^{n} \nu_{j-1}$ un $y'\Gamma_\theta^{-1} y = e'D^{-1}e$. Rezultātā, log - ticamības funkciju (5.1.1) var izteikt

$$\mathcal{L}(\theta) = -\frac{1}{2}\sum_{t=1}^{n} \log \nu_{t-1} - \frac{1}{2}\sum_{t=1}^{n} \frac{e_t^2}{\nu_{t-1}}.$$

Durbina algoritma sarežģītības pakāpe lineāram stacionāram procesam ir ar kārtu $\mathcal{O}(n^2)$. Dažiem Markova procesiem, piemēram ARMA modeļiem,Durbina - Levina algoritms ir izpildāms ar $\mathcal{O}(n)$ darbībām. Diemžēl, FARIMA modeļi nav Markova procesi.

Autokovariāciju aprēķināšana

Čoleskija un Durbina - Levina algoritmi balstās uz FARIMA procesa ACF aprēķināšanu. Viena no pieejām, kā to izdarīt, ir tā sauktā *sašķelšanas (splitting) metode*, kura balstās uz principu sadalīt procesu FARIMA tā ARMA un frakcionāli integrētajā (FI) daļās. Pieņem, ka $\gamma_1(\cdot)$ ir ACF no komponentes ARMA, savukārt $\gamma_2(\cdot)$ ir ACF no frakcionālā trokšņa. Tad procesa FARIMA autokorelāciju funkcija

$$\gamma(h) = \sum_{j=-\infty}^{\infty} \gamma_1(j)\gamma_2(j-h).$$

Ja bezgalīgo summu nošķeļam līdz m, tad iegūstam tuvinājumu

$$\gamma(h) \approx \sum_{j=-m}^{m} \gamma_1(j)\gamma_2(j-h).$$

No iepriekšējās izteiksmes varam efektīvi aprēķināt ACF $\gamma(\cdot)$ ar augstu precizitātes līmeni.

5.2. Autoregresīvie novērtējumi

Lai paātrinātu parametru novērtējumu skaitļošanas procesu, daudzi autori aplūko autoregresīvos novērtējumus. Pieņemsim, ka $\{y_t : t \in \mathbb{Z}\}$ ir ilglaicīgās atmiņas process, kas definēts ar

$$y_t = \varepsilon_t + \pi_1(\theta)y_{t-1} + \pi_2(\theta)y_{t-2} + \pi_3(\theta)y_{t-3} + \dots,$$

kur $\pi_j(\theta)$ ir $\phi(B)\theta^{-1}(B)(1-B)^d$ koeficienti. Tā kā praktiski pieejami ir tikai ierobežots skaits novērojumu - $\{y_1, \dots, y_n\}$, tad varam apskatīt saīsinātu modeli

$$y_t = \tilde{\varepsilon}_t + \pi_1(\theta)y_{t-1} + \pi_2(\theta)y_{t-2} + \dots + \pi_m(\theta)y_{t-m}, \tag{5.2.1}$$

visiem $m < t \leq n$. Aptuvenu ML novērtējumu $\hat{\theta}_n$ iegūst, minimizējot funkciju

$$\mathcal{L}_1(\theta) = \sum_{t=m+1}^{n} \left[y_t - \pi_1(\theta)y_{t-1} - \pi_2(\theta)y_{t-2} - \ldots - \pi_m(\theta)y_{t-m}\right]^2. \tag{5.2.2}$$

Apakšnodaļas turpinājumā apskatīsim uzlabotas metodes, kas, pamatojoties uz šo struktūru, dod precīzākus novērtējumus.

Hasleta - Rafterija metode

Procesa FARIMA, kas definēts $y_t = \frac{\theta(B)}{\phi(B)}(1-B)^{-d}\varepsilon_t = \sum_{j=0}^{\infty}\psi_j\varepsilon_{t-j}$, aptuvena $\{y_t\}$ prognoze vienam solim ir

$$\hat{y}_t = \phi(B)\theta(B)^{-1}\sum_{j=1}^{t-1}\phi_{tj}y_{t-j}, \tag{5.2.3}$$

ar prognozes kļūdas dispersiju

$$v_t = Var(y_t - \hat{y}_t) = \sigma_y^2 k \prod_{j=1}^{t-1}(1-\phi_{jj}^2),$$

kur $\sigma_y^2 = Var(y_t)$, k ir jaunās dispersijas attiecība pret ARMA(p,q) procesa dispersiju un

$$\phi_{tj} = -\begin{pmatrix} t \\ j \end{pmatrix} \frac{\Gamma(j-d)\Gamma(t-d-j+1)}{\Gamma(-d)\Gamma(t-d+1)}, \tag{5.2.4}$$

visiem $j = 1, \ldots, t$.

Lai skaitļošanā izvairītos no liela skaita koeficientu ϕ_{tj}, novērtējam

$$\sum_{j=1}^{t-1}\phi_{tj}y_{t-j} \approx \sum_{j=1}^{M}\phi_{tj}y_{t-j} - \sum_{j=M+1}^{t-1}\pi_j y_{t-j}, \tag{5.2.5}$$

pie lieliem j $\phi_{tj} \sim -\pi_j$, vienkāršības dēļ $\pi_j(\theta)$ apzīmēts ar π_j.

Tālāk aproksimējam

$$\sum_{j=M+1}^{t-1}\pi_j y_{t-j} \approx M\pi_M d^{-1}\left[1 - \left(\frac{M}{t}\right)^d\right]\bar{y}_{M+1,t-1-M},$$

kur $\bar{y}_{M+1,t-1-M} = \frac{1}{t-1-2M}\sum_{j=M+1}^{t-1-M} y_j$. Novērtējumu $\hat{\theta}_n$ iegūst, maksimizējot

$$\mathcal{L}_2(\theta) = konstante - \frac{1}{2}n\log\left[\hat{\sigma}^2(\theta)\right],$$

kur

$$\hat{\sigma}^2(\theta) = \frac{1}{n}\sum_{t=1}^{n}\frac{(y_t-\hat{y}_t)^2}{v_t}.$$

Hastleta - Rafterija algoritma sarežģītības pakāpe ir ar kārtu $\mathcal{O}(nM)$. Ja parametrs M ir fiksēts, tad metodes kārta ir $\mathcal{O}(n)$, kas liecina par to, ka tā ir daudz ātrāka kā Čoleskija vai Durbina - Levina pieeja. Pie $M = 100$ algoritms lielākoties strādā labi, turklāt, ja $M = n$, iegūstam ML novērtējumu frakcionālajam troksnim. Tādā gadījumā gan skaitļošanas pakāpe ir $\mathcal{O}(n^2)$.

Berana pieeja

Citu autoregresīvo aproksimācijas tehniku piedāvāja Berans. Aplūkosim sekojošu Gausa virkni

$$\varepsilon_t = y_t - \sum_{j=1}^{\infty} \pi_j(\theta) y_{t-j}.$$

Tā kā netiek novērotas vērtības $\{y_t, t \leq 0\}$, varam pieņemt, ka $y_t = 0$ visiem $t \leq 0$ un uzdot virkni

$$u_t = y_t - \sum_{j=1}^{t-1} \pi_j(\theta) y_{t-j},$$

visiem $j = 2, \ldots, n$. Pieņemsim, ka $r_t(\theta) = \dfrac{u_t(\theta)}{\sigma}$ un $\theta = (\sigma, \phi_1, \ldots, \phi_p, \theta_1, \ldots, \theta_q, d)$.Tad θ novērtējumu iegūsim, minimizējot

$$\mathcal{L}_2(\theta) = 2n \log(\sigma) + \sum_{t=2}^{n} r_t^2(\theta).$$

Ņemot parciālos atvasinājumus pēc θ, minimizācijas problēma ir ekvivalenta nelineāra vienādojumu atrisināšanai

$$\sum_{t=2}^{n} \{r_t(\theta)\dot{r}_t(\theta) - E\left[r_t(\theta)\dot{r}_t(\theta)\right]\} = 0, \tag{5.2.6}$$

kur $\dot{r}_t(\theta) = \left(\frac{\partial r_t(\theta)}{\partial \theta_1}, \ldots, \frac{\partial r_t(\theta)}{\partial \theta_r}\right)'$.

Berana pieejas aritmētisko darbību komplicētība, līdzīgi kā Durbina - Levina algoritmam, ir ar kārtu $\mathcal{O}(n^2)$.

5.3. Citas metodes parametru novērtēšanai

Parametru novērtēšanai izmanto arī citas pieminēšanas vērtas metodes, kuras sīkāk šajā darbā neapskatīsim:

- Slīdošā - vidējā (MA) apoksimācija, kura izpaužas kā Volda LRD procesa pieraksta nošķelšana no MA(∞) uz MA(m). Tādā gadījumā iespējams algoritma kārtu no $\mathcal{O}(n^3)$ reducēt uz $\mathcal{O}(n)$. Tajā izmanto procesa MA(m) *state space* reprezentāciju.

- Plaši pazīstama ir Vitla metode, kura balstās uz periodigrammu aprēķiniem, kas ir ļoti ātrs algoritms.

- Semiparametriskās metodes, kurās var izmantot datus bez īpašiem nosacījumiem, tās balstītas uz pieņēmumu par laikrindas spektrālā blīvuma veidu.

- Mainītās amplitūdas metode, kura būtībā ir pirmajā nodaļā apskatītā statistika R/S (1.0.1). Ar šīs pieejas palīdzību, iegūst novērtējumu LRD parametram d.

- Dispersiju shēmas. Metode paredzēta d novērtēšanai. Izlasi no n novērojumiem sadala m izmēra k blokos, $n = m \times k$, rezultātā iegūst:

$$\hat{d} = \frac{1}{2} - \frac{\sum_{j=1}^{k} (\log j - a) \left[\log Var(\bar{y}_j) - b\right]}{2 \sum_{j=1}^{k} (\log j - a)^2},$$

kur $a = \left(\frac{1}{k}\right) \sum_{j=1}^{k} \log j$, $b = \left(\frac{1}{k}\right) \sum_{j=1}^{k} \log Var(\bar{y}_j)$ un $\bar{y}_j = \frac{1}{m} \sum_{t=(j-1)\times m+1}^{j \times m} y_t$ pie $j = 1, \ldots, k$.

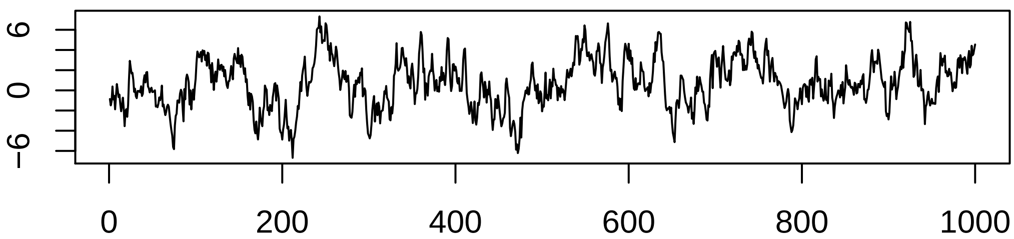
- Detrendētā svārstību, kā arī citas metodes.

6. Piemēri

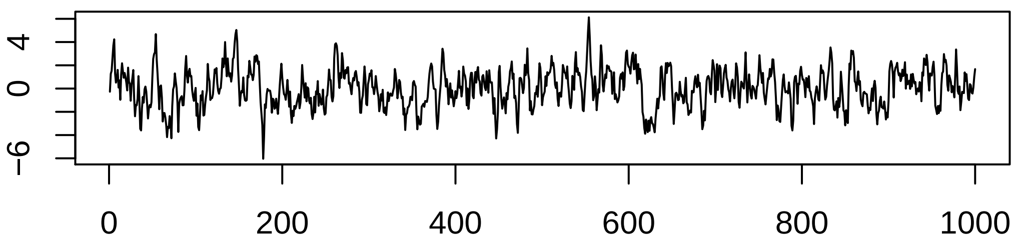
Šajā nodaļā, ar datorprogrammas R palīdzību, simulēsim procesus ARMA un ilglaicīgās atmiņas procesus FARIMA ar dažādiem parametriem ϕ un θ. Apskatīsim šo modeļu grafiskos attēlus un to ACF un PACF attēlus. Lai darbotos ar ilglaicīgās atmiņas procesiem, ir izstrādāta speciāla R pakete "fArma". Tajā iebūvētas dažādas komandas LRD modeļu simulācijai un parametru novērtēšanai.

Procesu FARIMA simulācijai, lieto funkciju "farimaSim", norādot vēlamo apjomu n un MA, AR un d koeficientu vērtības. Lai gūtu priekšstatu par ilglaicīgās atmiņas izpausmēm, simulēsim 3 veidu FARIMA(p, d, q) un ARMA(p, q):

1. FARIMA$(1, d, 1)$ (attēls 6.1.) un ARMA $(1, 1)$ (attēls 6.2.), kur $d = 0.4, \phi = 0.7, \theta = 0.2$.

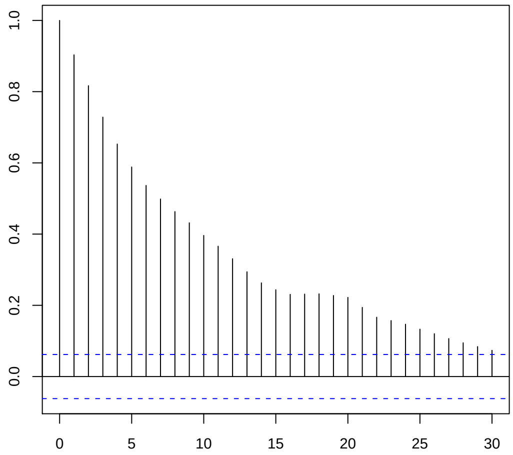


6.1. att. FARIMA$(1, d, 1)$

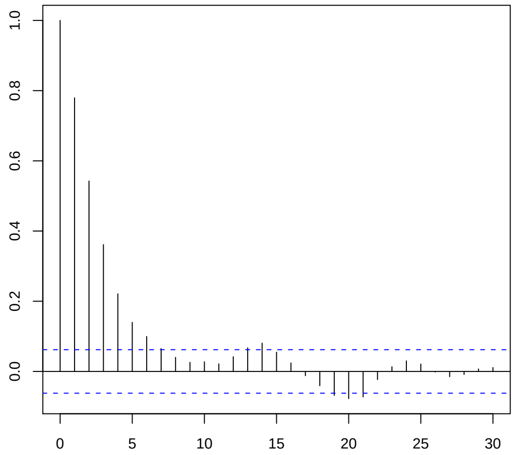


6.2. att. ARMA$(1, 1)$

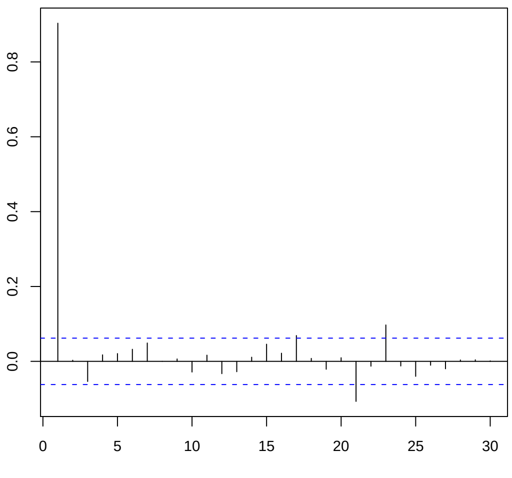
Attēls 6.3. ir procesa FARIMA$(1, d, 1)$ ACF funkcija, savukārt attēls 6.4. - ARMA$(1, 1)$ ACF. Ļoti uzskatāmi izpaužas lēnā korelāciju dilšana 6.3. salīdzinājumā ar 6.4..



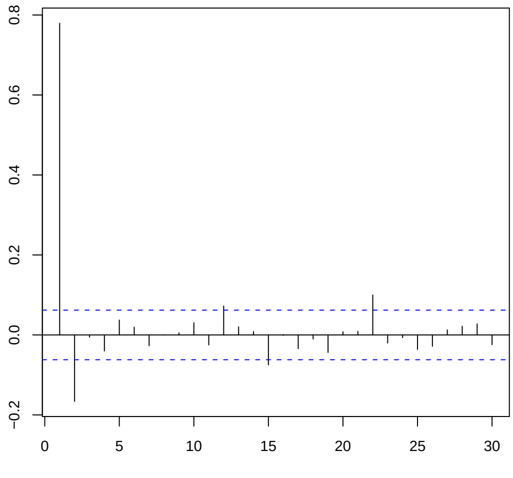
6.3. att. Procesa FARIMA$(1, d, 1)$ ACF



6.4. att. Procesa ARMA$(1, 1)$ ACF

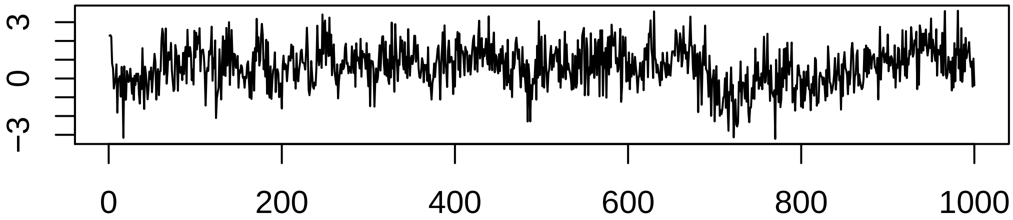


6.5. att.: Procesa FARIMA$(1, d, 1)$ PACF

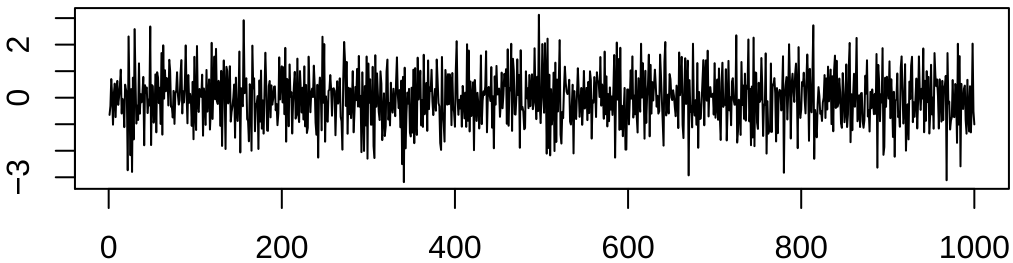


6.6. att. Procesa ARMA$(1, 1)$ PACF

2. FARIMA$(1, d, 0)$ (attēls 6.7.) un ARMA $(1, 0)$ (attēls 6.8.), kur $d = 0.4, \phi = -0.3, \theta = 0$.

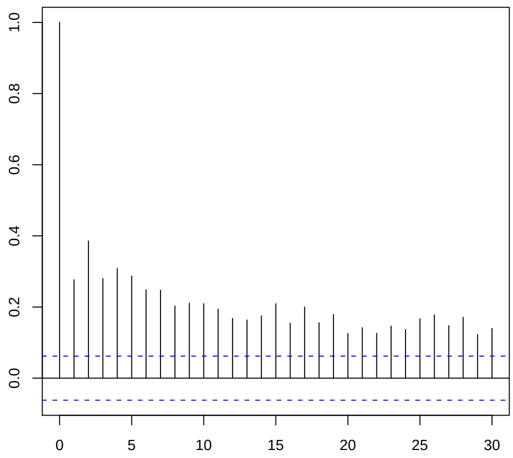


6.7. att. FARIMA$(1, d, 0)$

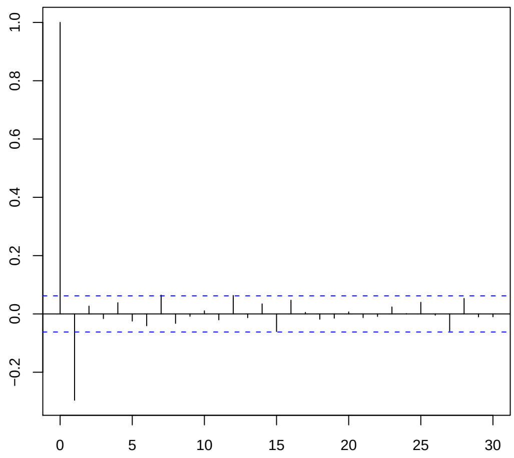


6.8. att. ARMA$(1, 0)$

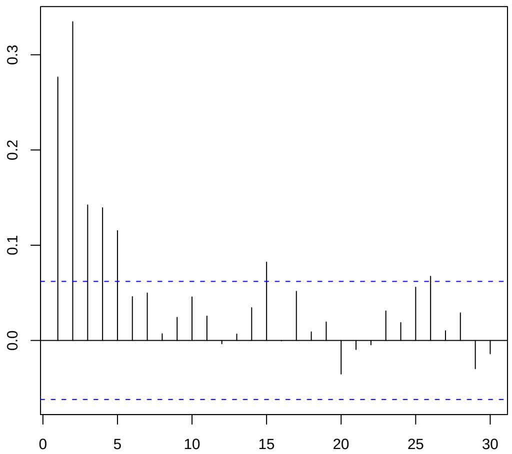
Attiecīgi procesu 6.7. un 6.8. ACF un PACF



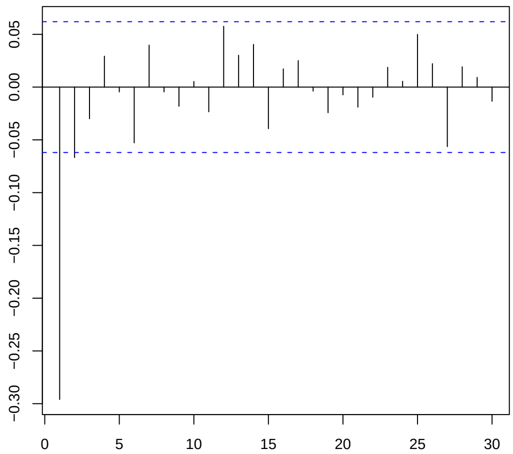
6.9. att. Procesa FARIMA$(1, d, 0)$ ACF



6.10. att. Procesa ARMA$(1, 0)$ ACF

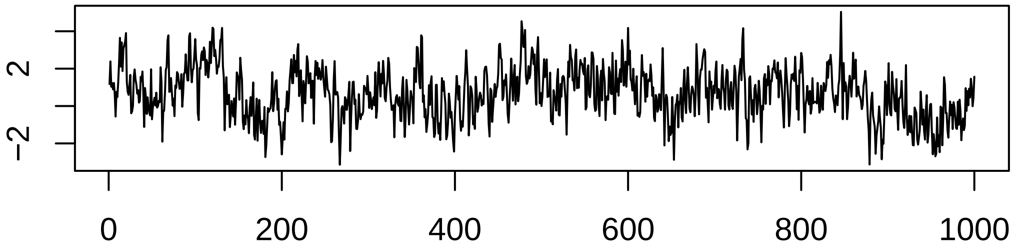


6.11. att.: Procesa FARIMA$(1, d, 0)$ PACF

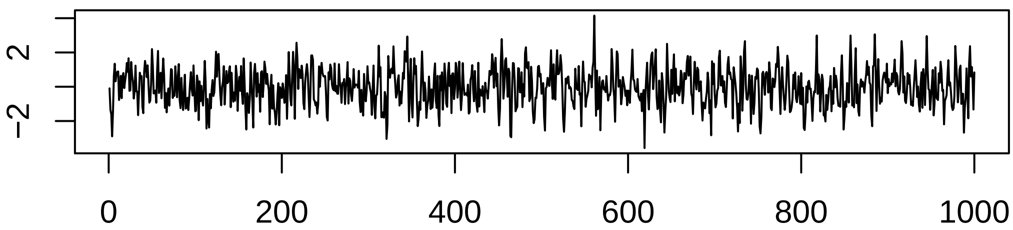


6.12. att. Procesa ARMA$(1, 0)$ PACF

3. FARIMA$(2, d, 1)$ (attēls 6.13.) un ARMA $(2, 1)$ (attēls 6.14.), kur $\phi_1 = 0.3, \phi_2 = -0.2, d = 0.4, \theta = 0.1$.

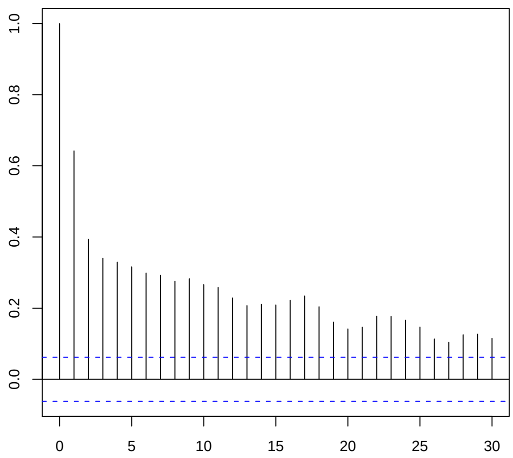


6.13. att. FARIMA$(2, d, 1)$

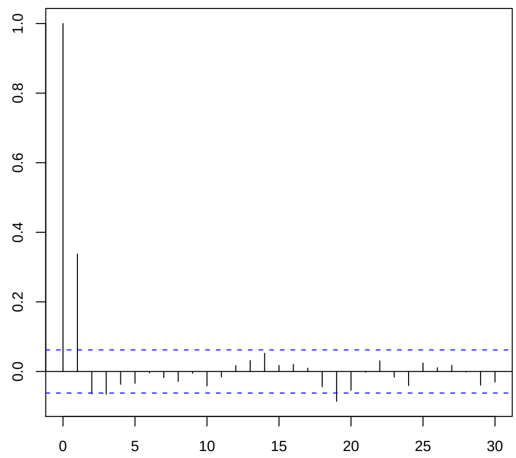


6.14. att. ARMA$(2, 1)$

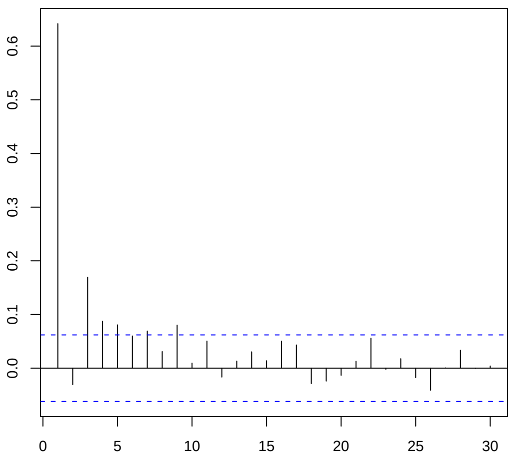
Attiecīgi procesu 6.13. un 6.14. ACF un PACF



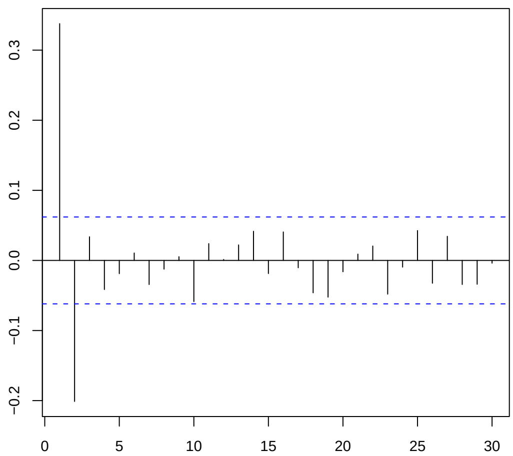
6.15. att.: Procesa FARIMA$(2, d, 1)$ ACF



6.16. att. Procesa ARMA$(2, 1)$ ACF



6.17. att.: Procesa FARIMA$(2, d, 1)$ PACF



6.18. att. Procesa ARMA$(2, 1)$ PACF

No iepriekš apskatītajiem attēliem 6.3. - 6.18. būtiskas atšķirības redzamas gan pašu modeļu FARIMA un ARMA, gan to ACF grafikos. Pamanāms, ka procesa FARIMA novērojumos pat pēc liela laika intervāla (1000 novērojumu) ir daudz izlecēju vērtību, kamēr procesam ARMA novērojumi vairāk tiecas izlīdzināties ap vidējo vērtību. Savukārt ACF attēlos no FARIMA redzama nenoliedzama ilglaicīgās atmiņas klātbūtne.

Secinājumi

Darbā izdevies gūt priekšstatu par to, kas ir ilglaicīgā atmiņa un kādēļ tā ir aktuāla. Teorētiskajā daļā vispārīgi apskatīti procesi, kas saistīti ar ilglaicīgo atmiņu - frakcionālā Brauna kustība, frakcionālais Gausa troksnis, sev līdzīgi procesi un pazīstamākā no LRD modeļu klasēm FARIMA. Sniegts ieskats FARIMA parametru novērtēšanas metodēs. Piemēros simulēti dažādi FARIMA procesi, kas satur ilglaicīgo atmiņu un ARMA procesi, kas to nesatur. Grafiski labi redzams, kā izpildās ilglaicīgās atmiņas pamatiezīme - neparasti lēnā korelāciju dilšana, salīdzinot ar ARMA procesiem.

Tēma ir plaša un prasa padziļinātas zināšanas statistikā, tādēļ raksturota ļoti vispārīgi. Turpināšu šo tēmu pētīt diplomdarbā, kura mērķis būs gūt detalizētāku izpratni.

Izmantotā literatūra un avoti

[1] Samorodnitsky G. *Long Range Dependence*, volume 99. Now Publishers,INC., Hanover, 2007.

[2] Peres Y. Morters P. *Brownian Motion*, volume 399. 2006.

[3] Palma W. *Long - Memory Time Series: Theory and Methods*, volume 285. John Wiley & Sons,INC., New Jersey, 2007.

[4] Taqqu M.S. Doukhan P., Oppenheim G. *Theory and Applications of Long Range Dependence*, volume 719. Birkhauser Boston,INC., New York, 2003.

A Pielikums

Teorēma 14. *Nepārtraukts attēlojums. Pieņem, ka $\{X_n\}_{n=1}^{\infty}$ ir gadījuma lielumu virkne un X ir gadījuma lielums, definēti vienā metriskā telpā $\mathcal{X}$. Pieņem, ka $X_n \xrightarrow{D} X$. Pieņem, ka $\mathcal{Y}$ ir metriska telpa un $g : \mathcal{X} \to \mathcal{Y}$. Definē $C_g = \{x : g$ ir nepārtraukts uz $x\}$. Pieņem, ka $P(X \in C_g) = 1$. Tad $g(X_n) \xrightarrow{D} g(X)$.*

Definīcija 9. [3] Stohastisku procesu $\{y_t : t \in \mathbb{Z}\}$ sauc par *otrās kārtas* jeb *stacionāru vājā nozīmē*,ja eksistē simetriska autokovariāciju funkcija $\gamma(\cdot)$, kurai $\gamma(0) = Var(y_t) < \infty$, tāda, ka $\langle y_t, y_{t+h}\rangle = \gamma(h)$ visiem $t, h \in \mathbb{Z}$.

Definīcija 10. [3] Otrās kārtas procesa autokovariāciju funkcija $\gamma(\cdot)$

$$\gamma(h) = \int_{-\pi}^{\pi} e^{ih\lambda} dF(\lambda),$$

kur funkcija F ir no labās puses nepārtraukta, nedilstoša, ierobežota $[-\pi, \pi]$ un izpildās nosacījums $F(-\pi) = 0$. F sauc par $\gamma(\cdot)$ *spektrālo sadalījumu*. Turpmāk, ja

$$F(\lambda) = \int_{-\pi}^{\lambda} f(\omega) d(\omega),$$

tad $f(\cdot)$ sauc par $\gamma(\cdot)$ *spektrālo blīvumu.*

Definīcija 11. Pieņem, ka $\{X_t : t \in T\}$ ir stohastisks process, kuram $Var(X_t) < \infty$ visiem $t \in T$. Tad $\{X_t\}$ *autokovariāciju funkcija* ir

$$\gamma_X(r, s) := Cov(X_r, X_s) = E\left[(X_r - E(X_r))(X_s - E(X_s))\right]$$

visiem $r, s \in T$.

Definīcija 12. [4] Funkciju $\Gamma(p)$ sauc par *gamma funkciju*, ja

$$\Gamma(p) = \begin{cases} \displaystyle\int_0^{\infty} t^{p-1} e^{-t} dt, & p > 0, \\ \infty, & p = 0, \\ p^{-1}\Gamma(1+p), & p < 0 \end{cases}$$

Definīcija 13. Funkciju $B(x, y)$ sauc par *beta funkciju*, ja

$$B(x, y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}.$$

Kursa darbs “Ilglaicīgās atmiņas procesi” izstrādāts LU Fizikas un Matemātikas fakultātē.

Ar savu parakstu apliecinu, ka pētījums veikts patstāvīgi, izmantoti tikai tajā norādītie informācijas avoti un iesniegtā darba elektroniskā kopija atbilst izdrukai.

Autors: Ilze Rēvalde

____________________ ________________

(paraksts) (datums)

Rekomendēju darbu aizstāvēšanai.

Vadītājs: doc. Dr.math. Jānis Valeinis

____________________ ________________

(paraksts) (datums)

Recenzents:

____________________ ________________

(paraksts) (datums)

Darbs iesniegts Matemātikas nodaļā ________________

(datums)

____

(darbu pieņēma)

Darbs aizstāvēts kursa darbs gala pārbaudījuma komisijas sēdē

________________ prot. Nr. ________, vērtējums________

(datums)

Komisijas sekretārs/-e: ________________ ________________

(Vārds, Uzvārds) (paraksts)