LATVIJAS UNIVERSITĀTE
FIZIKAS UN MATEMĀTIKAS FAKULTĀTE
MATEMĀTIKAS NODAĻA

JAUKTO PROCESU ANALĪZE UN TO PIELIETOJUMS STATISTIKĀ

MAĢISTRA DARBS

Autors: **Jolanta Rone**

Stud. apl. Mate030023

Darba vadītājs: doc. Dr.math. Jānis Valeinis

RĪGA 2009

Anotācija

Šajā darbā teorētiski tiek analizēti jauktie procesi pēc dažādiem vāji atkarīgiem jaukto procesu koeficientiem, kā arī apskatīts praktiskais pielietojums statistikā - jaukto procesu blīvuma un regresijas funkcijas novērtēšana ar kodolu gludināšanu, hipotēžu par jaukto procesu sadalījumu pārbaude (Bikela - Rozenblata tests un Neimaņa tests).

Atslēgas vārdi: Jauktie procesi, neparametriskā statistika, vājā atkarība.

Abstract

In this work mixing processes theoretically have been analysed using weak dependence coefficients. Application for mixing processes in nonparametric statistics are discussed specifically density and regression estimation using kernel smoothing method, goodness-of-fit tests (Bickel - Rosenblatt test and Neymann test).

Keywords: Mixing proceses, nonparametric statistics, weak dependence.

Saturs

Apzīmējumi

$\mathbb{Z}$ veselo skaitļu kopa,

$\mathbb{R}$ reālo skaitļu kopa,

$N(0, \sigma^2)$ normālais sadalījums ar matemātisko cerību 0 un dispersiju σ^2,

$U(0,1)$ vienmērīgais sadalījums segmentā $(0,1)$,

$C_{2,d}(b)$ telpa, kur ja funkcija $f \in C_{2,d}(b)$, d argumenta funkcija f ir nepārtraukta, divreiz atvasināma un $\| f \|_\infty \leq b$, $\| f^{(2)} \|_\infty \leq b$, kur $f^{(2)}$ ir otrās pakāpes parciālais atvasinājums un b konstante,

$I_{\{|x|\leq 1\}}$ ir indikātorfunkcija, tas ir, $I_{\{|x|\leq 1\}} = \begin{cases} 1, & ja \quad |x| \leq 1 \\ 0, & ja \quad |x| > 1 \end{cases}$,

$\rightarrow_d$ konverģence pēc sadalījuma,

$\rightarrow_p$ konverģence pēc varbūtības,

$\rightarrow_{g.d}$ gandrīz drošā konverģence,

$u_n \simeq v_n$ nozīmē to, ka eksistē tādas konstantes c_1 un c_2, ka $0 < c_1 v_n < u_n < c_2 v_n$ pietiekoši lieliem n.

Ievads

Klasiskā asimtotiskā teorija statistikā ir balstīta uz centrālo robežteorēmu un lielo skaitļu likumu neatkarīgu gadījuma lielumu virknēm. Pēc lielā skaitļa likuma izlases vidējā vērtība $\bar{X} = n^{-1}\sum_{i=1}^{n} X_i$, kur X_i $(i = 1, 2, ..., n)$ ir vienādi sadalīti neatkarīgi gadījuma lielumi un n - izlases apjoms, tiecas uz konstantu vērtību $\mu = EX$, kad $n \to \infty$. Savukārt pēc Centrālās robežteorēmas gadījumu lielumu summa ir asimptotiski normāli sadalīta. Precīzāk

$$\frac{\sqrt{n}(\bar{X} - \mu)}{\sigma} \to_d N(0, 1),$$

kur $\sigma = DX$.

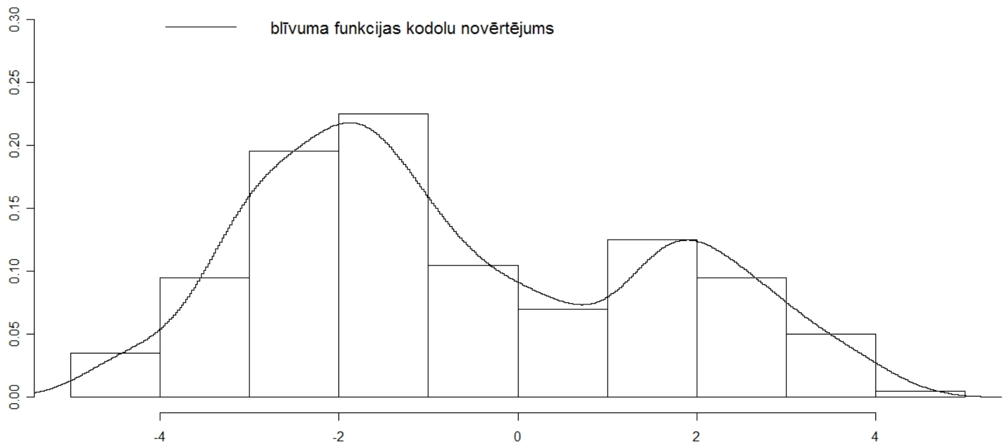
Tomēr praksē bieži dati ir atkarīgi, tādēļ nepieciešams analizēt lielo skaitļu likumu un centrālo robežteorēmu vispārējā gadījumā. Šajā darbā tiks analizēti *jauktie procesi*, kas vispārīgi pēta gadījuma lielumu atkarības struktūru. Jauktos procesus (vai jaukto gadījumu lielumu virknes) sauc arī par vāji atkarīgiem jeb īslaicīgās atmiņas procesiem (*Short-Memory proceses*), jo asimtotiski tie ir neatkarīgi. Pazīstamākie jaukto procesu piemēri ir lineārie un nelineārie ARIMA procesi, GARCH procesi, Markova ķēdes, Gausa procesi u.c.

Galvenais darba mērķis ir veikt jaukto procesu teorētisko analīzi, kā arī apskatīt to praktisko pielietojumu statistikā. Procesu atkarības struktūra tiek pētīta pēc dažādiem vāji atkarīgiem jaukto procesu koeficientiem, kurus 1956.gadā ieviesa amerikāņu matemātiķis M.Rozenblats. Visbiežāk apskatītais jaukto procesu koeficients ir α koeficients. Pieņemsim, ka dota varbūtību telpa $(\Omega, \mathcal{F}, P)$, kur Ω ir elementāro notikumu kopa, $\mathcal{F}$ - σ - algebra un P - varbūtību mērs. Starp divās σ-algebrām $\mathcal{A} \subset \mathcal{F}$ un $\mathcal{B} \subset \mathcal{F}$, α koeficients tiek definēts sekojoši

$$\alpha(\mathcal{A}, \mathcal{B}) = \sup_{A \in \mathcal{A}, B \in \mathcal{B}} \mid P(A \cap B) - P(A)P(B) \mid.$$

Ja σ algebras $\mathcal{A}$ un $\mathcal{B}$ ir neatkarīgas, tad $\alpha(\mathcal{A}, \mathcal{B}) = 0$.

Neparametriskās statistikas lekciju kursā tika apskatīts, kā neatkarīgiem vienādi sadalītiem datiem var novērtēt blīvuma funkciju un regresijas funkciju ar kodolu gludināšanu. Maģistra darba ietvaros tiks veikta neparametriskās statistikas analīze jauktajiem procesiem. 1956.gadā Rozenblats parādīja pie kādiem nosacījumiem strādā neparametriskā blīvuma funkcijas novērtēšana neatkarīgi vienādi sadalītiem datiem. Jaukto procesu gadījumā šos nosacījumus ir ļoti grūti pārbaudīt. Tomēr parasti dati tiek modelēti ar literatūrā pazīstamiem modeļiem, kā piemēram, ARIMA, GARCH modeļiem, kuriem literatūrā ir pierādīts, ka nosacījumi neparametriskājām metodēm izpildās. 1.attēlā ir attēlots jauktā procesa $X_t = (5X_{t-1})/(1+0.9X_{t-1}) + \varepsilon_t$, ($\varepsilon_t \sim N(0,1)$) blīvuma funkcijas kodolu novērtējums.



1. att.: Apjomā 200 ģenerēta izlase no procesa $X_t = (5X_{t-1})/(1 + 0.9X_{t-1}) + \varepsilon_t$, ($\varepsilon_t \sim N(0,1)$) ar blīvuma funkcijas novērtējumu, izmantojot kodolu gludināšanas metodi.

Neparametrisko regresijas funkcijas novērtējumu ar kodolu gludināšanas metodi neatkarīgiem vienādi sadalītiem datiem 1964.gadā ieviesa Nadaraja un Vatsons, bet samērā nesen 1991.gadā Bosq un 1994.gadā Rhomari parādīja to arī jauktajiem procesiem. Laikrindām viens no svarīgākajiem regresijas pielietojumiem ir prognozes veikšana.

Klasiskajā literatūrā pazīstamākie gadījuma lieluma sadalījuma hipotēžu pārbaudes testi ir χ^2 jeb Pīrsona, Kolmogorova - Smirnova, Šapiro - Vilka, Bikela - Rozenblata, Neimaņa testi, kas strādā neatkarīgiem vienādi sadalītiem gadījuma lielumiem. Turpretī jauktajiem procesiem samērā nesen tikai 2000.gadā tika atklāts Bikela - Rozenblata tests un 2009.gadā Neimaņa tests, kas arī darbā tiks teorētiski apskatīti.

Darbā tiks apskatīti arī ilglaicīgās atmiņas procesi (*Long-Memory proceses*) un salīdzināti ar jauktajiem jeb īslaicīgās atmiņas procesiem. Atšķirībā no jauktajiem procesiem, kam raksturīgs eksponenciāls autokovariāciju dilšanas ātrums, ilglaicīgās atmiņas procesiem autokovariācijas dilst ļoti lēni. 1989.gadā Diebolds un Rudebušs saskatīja ilglaicīgās atmiņas procesa īpašības datos, kas atspoguļoja ikceturkņa ASV nacionālo kopproduktu. Baillie, Chungs un Tieslau 1985.gadā atklāja, ka ilglaicīgās atmiņas procesa īpašības piemīt ASV patēriņa cenu indeksa inflācijai kādā noteiktā laika periodā.

Darbs sastāv no četrām nodaļām. Pirmajā nodaļā tiek definēti jaukto procesu koeficienti un to īpašības, kā arī tiek apskatīti jaukto procesu koeficienti piemēri, kad divdimensiju gadījuma lielums pieņem tikai divas vērtības. Otrajā nodaļā tiek apskatīta Ergodiskā teorija un Centrālā robežteorēma jauktajiem procesiem. Dažādas jaukto procesu laikrindas, kā arī ilglaicīgās atmiņas procesi tiek definēti trešajā nodaļā. Jaukto procesu pielietojums statistikā tiek apskatīts ceturtajā nodaļā. Darba nobeigumā ir apkopoti secinājumi.

1. Jauktie procesi

Jauktos procesus definē ar jaukto procesu koeficientiem, tādēļ šajā nodaļā tiks pētītas to galvenās īpašības, kā arī tiks apskatīti vairāki piemēri, kā šie jaukto procesu koeficienti tiek aprēķināti divdimensiāliem gadījuma lielumiem, kas pieņem tikai divas vērtības. Vispārīgā gadījumā jaukto procesu koeficientus aprēķināt ir samērā sarežģīti.

1.1. Jaukto procesu koeficienti un to īpašības

Definīcija 1. Par gadījuma procesu sauc gadījuma lieluma saimi $(X_t, t \in \mathbb{Z})$, kas uzdota varbūtību telpā $(\Omega, \mathcal{F}, P)$.

Gadījuma procesa visu iespējamo vērtību kopu S sauc par gadījuma procesa stāvokļa kopu, tas ir, $X : \Omega \mapsto S$. Ja kopa $S \in \mathbb{Z}_+$, tad procesu $(X_t, t \in \mathbb{Z})$ sauc par diskrētu, ja kopa $S \in \mathbb{R}_+$, tad process $(X_t, t \in \mathbb{Z})$ ir nepārtraukts.

Definīcija 2. Par gadījuma lielumu X, Y korelācijas koeficientu sauc lielumu

$$Corr(X,Y) = \frac{Cov(X,Y)}{\sqrt{D(X)}\sqrt{D(Y)}} = \frac{EXY - EX \cdot EY}{\sqrt{EX^2 - (EX)^2}\sqrt{EY^2 - (EY)^2}},$$

kur DX un DY ir dispersijas, $Cov(X,Y)$ ir kovariācijas koeficients.

Pieņemsim, ka dota varbūtību telpa $(\Omega, \mathcal{F}, P)$, $\mathcal{A}$ un $\mathcal{B}$ ir σ-algebras, $\mathcal{A}, \mathcal{B} \subset \mathcal{F}$. Definēsim atkarības koeficientus starp divām σ-algebrām $\mathcal{A}$ un $\mathcal{B}$ sekojošā veidā:

$$\alpha(\mathcal{A},\mathcal{B}) = \sup_{A\in\mathcal{A}, B\in\mathcal{B}} \mid P(A\cap B) - P(A)P(B) \mid; \tag{1.1.1}$$

$$\phi(\mathcal{A},\mathcal{B}) = \sup_{A\in\mathcal{A}, B\in\mathcal{B}} \mid P(B \mid A) - P(B) \mid, P(A) > 0; \tag{1.1.2}$$

$$\psi(\mathcal{A},\mathcal{B}) = \sup_{A\in\mathcal{A}, B\in\mathcal{B}} \left| \frac{P(A\cap B)}{P(A)P(B)} - 1 \right|, P(A) > 0, P(B) > 0; \tag{1.1.3}$$

$$\rho(\mathcal{A},\mathcal{B}) = \sup_{f\in L^2(\mathcal{A}),g\in L^2(\mathcal{B})} \mid Corr(f,g) \mid; \tag{1.1.4}$$

$$\beta(\mathcal{A},\mathcal{B}) = \sup_{A\in\mathcal{A},B\in\mathcal{B}} \frac{1}{2}\sum_{i=1}^{I}\sum_{j=1}^{J} \mid P(A_i \cap B_j) - P(A_i)P(B_j) \mid, \tag{1.1.5}$$

kur $\{A_1, A_2, ..., A_I\}$ un $\{B_1, B_2, ..., B_J\}$ ir nešķeļošas kopas, kas sadala Ω, $\forall i\ A_i \in \mathcal{A}$ un $\forall j\ B_j \in \mathcal{B}$.

Definīcija 3. Pieņemsim, ka dota varbūtību telpa $(\Omega, \mathcal{F}, P)$. σ-algebra $\mathcal{A}$ ir separabla, ja eksistē tāda bezgalīga vai sanumurējama virkne $A_1, A_2, ... \in \mathcal{A}$, ka $\mathcal{A}$ ir mazakā σ-algebra notikumu telpā Ω, kas satur visus šos notikumus $A_1, A_2, ...$

Apgalvojums 1. *Pieņemsim, ka dota varbūtību telpa $(\Omega, \mathcal{F}, P)$, $\mathcal{A} \subset \mathcal{F}$ un $\mathcal{B} \subset \mathcal{F}$. Ja $\mathcal{B}$ ir separabla un eksistē regulāra nosacītā varbūtība $P(B|\mathcal{A}, B \in \mathcal{B})$, tad*

$$\beta(\mathcal{A},\mathcal{B}) = E\left[\sup_{B\in\mathcal{B}} |P(B|\mathcal{A}) - P(B)|\right]. \tag{1.1.6}$$

Pierādījums. Skatīt [[1], 89.lpp].

Pieņemsim $(X_t, t \in \mathbb{Z})$ ir gadījumu lielumu virkne no varbūtību telpas $(\Omega, \mathcal{F}, P)$, $\forall J, L$ $(-\infty \leq J \leq L \leq \infty)$, definēsim σ-algebru $\mathcal{F}_J^L = \sigma(X_k, J \leq k \leq L)$. Jaukto procesu atkarības koeficienti $\forall n$ procesam $(X_t, t \in \mathbb{Z})$ tiek definēti sekojoši:

$$\begin{aligned}
\alpha(n) &= \sup_{J\in\mathbb{Z}} \alpha(\mathcal{F}_{-\infty}^J, \mathcal{F}_{J+n}^\infty);\\
\beta(n) &= \sup_{J\in\mathbb{Z}} \beta(\mathcal{F}_{-\infty}^J, \mathcal{F}_{J+n}^\infty);\\
\rho(n) &= \sup_{J\in\mathbb{Z}} \rho(\mathcal{F}_{-\infty}^J, \mathcal{F}_{J+n}^\infty);\\
\phi(n) &= \sup_{J\in\mathbb{Z}} \phi(\mathcal{F}_{-\infty}^J, \mathcal{F}_{J+n}^\infty);\\
\psi(n) &= \sup_{J\in\mathbb{Z}} \psi(\mathcal{F}_{-\infty}^J, \mathcal{F}_{J+n}^\infty).
\end{aligned}$$

Definīcija 4. Saka, ka process $(X_t, t \in \mathbb{Z})$ ir

1. α jauktais process, ja $\alpha(n) \to 0$, kad $n \to \infty$;

2. β jauktais process, ja $\beta(n) \to 0$, kad $n \to \infty$;

3. ρ jauktais process, ja $\rho(n) \to 0$, kad $n \to \infty$;

4. ϕ jauktais process, ja $\phi(n) \to 0$, kad $n \to \infty$;

5. ψ jauktais process, ja $\psi(n) \to 0$, kad $n \to \infty$.

Piezīme 2. *No šīs definīcijas kļūst skaidrs kāpēc jauktos procesus sauc par vāji atkarīgiem - asimtotiski atkarības koeficienti tiecas uz nulli. Kā zināms σ - algebras raksturo informāciju par procesu $(X_t, t \in \mathbb{Z})$. Jaukto procesu koeficienti, kas definēti tieši σ - algebrām tiecas uz nulli. Tas nozīmē, ka informācija (varbūtiskais procesa raksturs) par procesa pašreizējo uzvedību nav atkarīga no informācijas pēc liela laika perioda.*

Aplūkosim dažas jaukto procesu koeficientu īpašības, kas būs nepieciešamas turpmākajā analīzē.

Sekojoša diagramma ilustrē sakarības starp jaukto procesu koeficientiem.

$$\begin{array}{lcccr} & \nearrow & \beta \text{ jauktais} & \searrow & \\ \psi \text{ jauktais} \to \phi \text{ jauktais} & & & & \alpha \text{ jauktais process} \\ & \searrow & \rho \text{ jauktais} & \nearrow & \end{array}$$

Pamatosim šo diagrammu, jo tā parāda, ka jebkurš ψ, ϕ, β, ρ jauktais process ir arī α jauktais process. Tieši šī iemesla dēļ robežteorēmas, blīvuma funkcijas novērtēšana ar kodolu gludināšanas metodi apskatītas visbiežāk tieši α jauktajiem procesiem, jo tas ir vispārīgs process. Lai pamatotu šo diagrammu, vispirms nepieciešams izanalizēt dažas jaukto procesu īpašības.

Apgalvojums 3. *Ja $\mathcal{A}$, $\mathcal{A}_0$, $\mathcal{B}$ un $\mathcal{B}_0$ ir σ-algebras tādas, ka $\mathcal{A}_0 \subset \mathcal{A}$ un $\mathcal{B}_0 \subset \mathcal{B}$, tad izpildās sekojoši apgalvojumi:*

$$\alpha(\mathcal{A}_0, \mathcal{B}_0) \leq \alpha(\mathcal{A}, \mathcal{B});$$

$$\beta(\mathcal{A}_0, \mathcal{B}_0) \leq \beta(\mathcal{A}, \mathcal{B});$$

$$\rho(\mathcal{A}_0, \mathcal{B}_0) \leq \rho(\mathcal{A}, \mathcal{B});$$

$$\phi(\mathcal{A}_0, \mathcal{B}_0) \leq \phi(\mathcal{A}, \mathcal{B});$$

$$\psi(\mathcal{A}_0, \mathcal{B}_0) \leq \psi(\mathcal{A}, \mathcal{B}).$$

Pierādījums. Skatīt [[1], 27.lpp un 68.lpp].

Teorēma 4. *Pieņemsim, ka ir dota varbūtību telpa* $(\Omega, \mathcal{F}, P)$ *un* $\mathcal{A}, \mathcal{B} \subset \mathcal{F}$*, tad*

1. *izpildās sekojošas nevienādības:*

   a) $0 \leq \alpha(\mathcal{A}, \mathcal{B}) \leq 4^{-1}$*; b)* $0 \leq \beta(\mathcal{A}, \mathcal{B}) \leq 1$*; c)* $0 \leq \rho(\mathcal{A}, \mathcal{B}) \leq 1$*;*

   d) $0 \leq \phi(\mathcal{A}, \mathcal{B}) \leq 1$*; e)* $0 \leq \psi(\mathcal{A}, \mathcal{B}) \leq \infty$.

2. *ekvivalenti sekojoši apgalvojumi:*

   a) σ*-algebras* $\mathcal{A}, \mathcal{B}$ *ir neatkarīgas; b)* $\alpha(\mathcal{A}, \mathcal{B}) = 0$*; c)* $\beta(\mathcal{A}, \mathcal{B}) = 0$*;*

   d) $\rho(\mathcal{A}, \mathcal{B}) = 0$*; e)* $\phi(\mathcal{A}, \mathcal{B}) = 0$*; f)* $\psi(\mathcal{A}, \mathcal{B}) = 0$.

Pierādījums. Skatīt [[1], 68.lpp].

Lemma 5. *Ir spēkā divi apgalvojumi*

a) $\forall i \in \{1, 2, ...I\}$ *izpildās nevienādība*

$$\sum_{j=1}^{J} | P(A_i \cap B_j) - P(A_i)P(B_j) | \leq 2P(A_i)\phi(\mathcal{A}, \mathcal{B});$$

b) $\forall j \in \{1, 2, ...J\}$ *izpildās nevienādība*

$$\sum_{i=1}^{I} | P(A_i \cap B_j) - P(A_i)P(B_j) | \leq 2P(B_j)\phi(\mathcal{A}, \mathcal{B}).$$

Pierādījums.

Pierādīsim a) apgalvojumu. Pieņemsim, ka $i = 1$.

Ja $P(A_1) = 0$, a) apgalvojums izpildās, tādēļ pierādīsim gadījumam, ka $P(A_1) > 0$.

Definēsim kopu S, kas sastāv no visiem $j \in \{1, 2, ..., J\}$ tā, $P(B_j \mid A_1) - P(B_j) \geq 0$, kopu T, kas sastāv no visiem $j \in \{1, 2, ..., J\}$ tā, ka $P(B_j \mid A_1) - P(B_j) < 0$, un notikumus $C = \cup_{j \in S} B_j$ un $D = \cup_{j \in T} B_j$.

Tad

$$\begin{aligned}
\sum_{j=1}^{J} | P(A_1 \cap B_j) - P(A_1)P(B_j) | &= \sum_{j=1}^{J} | P(A_1)P(B_j \mid A_1) - P(B_j) | \\
&= P(A_1) \left[\sum_{j \in S} [P(B_j \mid A_1) - P(B_j)] \right. \\
&- \left. \sum_{j \in T} [P(B_j \mid A_1) - P(B_j)] \right] \\
&= P(A_1)[[P(C \mid A_1)] - [P(D \mid A_1)] \\
&\leq P(A_1)[\phi(\mathcal{A}, \mathcal{B}) + \phi(\mathcal{A}, \mathcal{B})] \\
&= 2P(A_1)\phi(\mathcal{A}, \mathcal{B}).
\end{aligned}$$

a) apgalvojums ir pierādīts gadījumam, kad $i = 1$, bet tādā pašā veidā varam pierādīt $\forall i$. b) apgalvojumu var pierādīt līdzīgi a) apgalvojumam (skatīt [[1], 74.lpp]).

Tagad tiks pamatota augstāk minētā diagramma.

Apgalvojums 6. *Pieņemsim, ka ir dota varbūtību telpa* $(\Omega, \mathcal{F}, P)$ *un* $\mathcal{A}, \mathcal{B} \subset \mathcal{F}$*, tad izpildās sekojošas nevienādības:*

1. $2\alpha(\mathcal{A},\mathcal{B}) \leq \beta(\mathcal{A},\mathcal{B}) \leq \phi(\mathcal{A},\mathcal{B}) \leq \frac{1}{2}\psi(\mathcal{A},\mathcal{B})$*;*

2. $4\alpha(\mathcal{A},\mathcal{B}) \leq \rho(\mathcal{A},\mathcal{B}) \leq \psi(\mathcal{A},\mathcal{B})$*.*

Pierādījums. Pieņemsim, ka A un B ir divas notikumu kopas. Tā kā

$$P(A \cap B) + P(A^c \cap B) = P(B) = P(A)P(B) + P(A^c)P(B),$$

tad

$$P(A^c \cap B) + P(A^c)P(B) = -\left[P(A \cap B) + P(A)P(B)\right]. \tag{1.1.7}$$

Ja $P(A) > 0$, tad

$$P(B \mid A) - P(B) = -\left[P(B^c \mid A) - P(B^c)\right]. \tag{1.1.8}$$

Pierādīsim nevienādību $2\alpha(\mathcal{A},\mathcal{B}) \leq \beta(\mathcal{A},\mathcal{B})$.

Pieņemsim, ka $A \in \mathcal{A}$ un $B \in \mathcal{B}$. Sadalām Ω nešķēļošās kopās $\{A_1, A_2\}$ un $\{B_1, B_2\}$ tā, ka $A_1 = A$, $A_2 = A^c$, $B_1 = B$, $B_2 = B^c$, tad pēc (1.1.5) un (1.1.7) vienādības

$$\begin{aligned}
\beta(\mathcal{A},\mathcal{B}) &\geq \frac{1}{2}\sum_{i=1}^{2}\sum_{j=1}^{2} \mid P(A_i \cap B_j) - P(A_i)P(B_j) \mid \\
&= \frac{1}{2}(\mid P(A_1 \cap B_1) - P(A_1)P(B_1) \mid \\
&+ \mid P(A_1 \cap B_2) - P(A_1)P(B_2) \mid \\
&+ \mid P(A_2 \cap B_1) - P(A_2)P(B_1) \mid \\
&+ \mid P(A_2 \cap B_2) - P(A_2)P(B_2) \mid) \\
&= \frac{1}{2}(4 \mid P(A \cap B) - P(A)P(B) \mid) \\
&= 2 \mid P(A \cap B) - P(A)P(B) \mid \\
&= 2\alpha(\mathcal{A},\mathcal{B}).
\end{aligned}$$

Pierādīsim nevienādību $\beta(\mathcal{A},\mathcal{B}) \leq \phi(\mathcal{A},\mathcal{B})$.

Pieņemsim, ka $\{A_1, ..., A_I\}$ un $\{B_1, ..., B_J\}$ veido notikumu kopu Ω tā, ka $\forall i \ A_i \in \mathcal{A}$ un $\forall j \ B_j \in \mathcal{B}$.

Tākā 5.lemma apgalvo, ka $\forall i \in \{1, 2, ...I\}$

$$\sum_{j=1}^{J} \mid P(A_i \cap B_j) - P(A_i)P(B_j) \mid \leq 2P(A_i)\phi(\mathcal{A},\mathcal{B}),$$

tad katru pusi, summējot pa $\forall i$, iegūsim, ka

$$\sum_{i=1}^{I}\sum_{j=1}^{J} \mid P(A_i \cap B_j) - P(A_i)P(B_j) \mid \leq 2\phi(\mathcal{A},\mathcal{B})$$

un, ja iegūstās nevienādības katru pusi vēl pareizinās ar 2^{-1}, iegūsim, ka

$$\beta(\mathcal{A},\mathcal{B}) \leq \phi(\mathcal{A},\mathcal{B}).$$

Pierādīsim nevienādību $\phi(\mathcal{A},\mathcal{B}) \leq (1/2)\psi(\mathcal{A},\mathcal{B})$.

Ir zināms, ka, ja $P(B) \geq 1/2$, tad $P(B^c) \leq 1/2$. No (1.1.8) vienādības seko, ka

$$\phi(\mathcal{A},\mathcal{B}) = \sup_{A\in\mathcal{A},B\in\mathcal{B}} \mid P(B \mid A) - P(B) \mid, P(A) > 0, P(B) \leq 1/2.$$

Pieņemsim, ka $A \in \mathcal{A}$, $B \in \mathcal{B}$, $P(A) > 0$ un $0 < P(B) \leq 1/2$, tad

$$\frac{1}{2}\psi(\mathcal{A},\mathcal{B}) = \frac{1}{2}\sup_{A\in\mathcal{A},B\in\mathcal{B}} \mid \frac{P(A\cap B)}{P(A)P(B)} - 1 \mid \geq P(B) \mid \frac{P(A\cap B)}{P(A)P(B)} - 1 \mid =$$
$$= |P(B \mid_A) - P(B)| = \phi(\mathcal{A},\mathcal{B}).$$

Tākā $P(B \mid_A) - P(B) = -\left[P(B^c \mid_A) - P(B^c)\right]$, tad vienādība

$$\phi(\mathcal{A},\mathcal{B}) \leq (1/2)\psi(\mathcal{A},\mathcal{B})$$

būs spēkā arī gadījumam, kad $P(B) > 1/2$.

Pierādīsim nevienādību $4\alpha(\mathcal{A},\mathcal{B}) \leq \rho(\mathcal{A},\mathcal{B})$.

Pieņemsim, ka $A \in \mathcal{A}$ un $B \in \mathcal{B}$.

$$\alpha(\mathcal{A},\mathcal{B}) = \sup_{A\in\mathcal{A},B\in\mathcal{B}} \mid P(A\cap B) - P(A)P(B) \mid = \sup_{A\in\mathcal{A},B\in\mathcal{B}} Cov(I_A, I_B) =$$
$$= \sqrt{D(I_A)}\sqrt{D(I_B)} \sup_{A\in\mathcal{A},B\in\mathcal{B}} Corr(I_A, I_B) \leq \sqrt{D(I_A)}\sqrt{D(I_B)}\rho(\mathcal{A},\mathcal{B}),$$

kur I_A un I_B ir indikātorfunkcija.

Tākā $\forall x \in [0,1]$ $(x \in R)$ $x - x^2 \leq 1/4$, tad

$$D(I_A) = (P(A)(1-P(A))) \leq \frac{1}{4},\ D(I_B) = (P(B)(1-P(B))) \leq \frac{1}{4},$$

tādejādi $\alpha(\mathcal{A},\mathcal{B}) = (1/4)\rho(\mathcal{A},\mathcal{B})$.

Nevienādības $\rho(\mathcal{A},\mathcal{B}) \leq \psi(\mathcal{A},\mathcal{B})$ pierādījumu skatīt [[1], 76.lpp].

1.2. Jaukto procesu analīze un piemēri, kad gadījuma lielums pieņem tikai divas vērtības

Ja gadījuma lielums pieņem tikai divas vērtības, tad jaukto procesu koeficientus ir daudz vienkāršāk aprēķināt. Šajā nodaļā apskatīsim, kā no vispārējiem atkarības koeficientiem, kas definēti σ - algebrām, pāriet uz samērā vienkāršiem matemātiskiem modeļiem un aprēķiniem.

Apgalvojums 7. *Pieņemsim, ka X un Y ir gadījuma lielumi, kur katrs pieņem tikai divas vērtības, tad*

$$\alpha(\sigma(X),\sigma(Y)) = \frac{1}{2}\beta(\sigma(X),\sigma(Y)) = \mid P(X=s, Y=t) - P(X=s)P(Y=t) \mid, \quad (1.2.1)$$

kur s (vai t) ir viena no iespējamām vērtībām, kuru pieņem gadījuma lielums X (vai Y) un

$$\rho(\sigma(X),\sigma(Y)) = \mid Corr(X,Y) \mid . \quad (1.2.2)$$

Pierādījums. Definēsim notikumus $A_0 = \{X = s\}$ un $B_0 = \{Y = t\}$, tad

$$\sigma(X) = \{\Omega, \emptyset, A_0, A_0^c\},\ \sigma(Y) = \{\Omega, \emptyset, B_0, B_0^c\}.$$

Pēc vienādības (1.1.7)

$$\mid P(A \cap B) - P(A)P(B) \mid = \begin{cases} \mid P(A_0 \cap B_0) - P(A_0)P(B_0) \mid, \\ \qquad \text{ja } A = A_0 \text{ vai } A_0^c,\ B = B_0 \text{ vai } B_0^c; \\ 0, \text{ ja } A = \Omega \text{ vai } \emptyset; \\ 0, \text{ ja } B = \Omega \text{ vai } \emptyset. \end{cases}$$

Tādēļ

$$\alpha(\sigma(X),\sigma(Y)) = \mid P(A_0 \cap B_0) - P(A_0)P(B_0) \mid.$$

Līdzīgi var aprēķināt (skatīt [[1], 87.lpp])

$$\beta(\sigma(X),\sigma(Y)) = 2 \mid P(A_0 \cap B_0) - P(A_0)P(B_0) \mid.$$

Pierādīsim vienādību (1.2.2). Pieņemsim, ka V un W ir patvaļīgi gadījuma lielumi, kas ir mērojami attiecīgi ar σ-algebrām $\sigma(X)$ un $\sigma(Y)$.

Pieņemsim, ka s un s' ir divas iespējamās vērtības, ko pieņem gadījumu lielums X. Vērtības u un u' pieņem gadījuma lielums V notikumam

$$A_0 = \{X = s\} \text{ vai } A_0^c = \{X = s^{'}\}.$$

Pieņemsim, ka q un r ir tādi reāli skaitļi, ka līnija $y = qY + r$ iet caur punktiem (s, u) un $(s^{'}, u^{'})$. Tad $V = qX + r$.

Līdzīgi $W = cY + d$, kur arī c un d ir reāli skaitļi.

Tātad

$$Corr(V, W) = \begin{cases} \text{Corr(X,Y)}, & \text{ja qc>0;} \\ \text{-Corr(X,Y)}, & \text{ja qc<0;} \\ 0, & \text{ja qc=0.} \end{cases}$$

Tādēļ $Corr(V, W) \leq \mid Corr(X, Y) \mid$. Tākā V un W ir patvaļīgas un $\sigma(X)$ un $\sigma(Y)$ mērojamas, tad $Corr(V, W) = \mid Corr(X, Y) \mid = \rho(\sigma(X), \sigma(Y))$.

Definīcija 5. Par gadījuma lielumu X, Y marginālajiem sadalījumiem sauc gadījuma lielumu atsevišķu komponenšu varbūtību sadalījumi.

$$P(X = x) = \sum_y P(Y = y, X = x),$$

kur $P(Y = y, X = x)$ ir gadījuma lielumu X, Y kopējais sadalījums.

Apskatīsim dažus piemērus, kuros parādīts, kā tiek aprēķināti jaukto procesu koeficienti divdimensiju gadījuma lielumam, kas pieņem tikai divas vērtības 0 un 1.

Piemērs 1. Tika izmantota literatūra [[1], 100.lpp].

Pieņemsim, ka $0 \leq \varepsilon \leq 1$ un X, Y ir gadījuma lielumi ar sekojošiem gadījuma lielumu X, Y kopējiem sadalījumiem:

$$P(X = Y = 0) = P(X = Y = 1) = \frac{1 + \varepsilon}{4},$$

$$P(X = 0, Y = 1) = P(X = 1, Y = 0) = \frac{1 - \varepsilon}{4}.$$

Iegūsim gadījumu lielumu X un Y marginālos sadalījumus

$$P(X = 0) = P(X = 0, Y = 0) + P(X = 0, Y = 1) = \frac{1 + \varepsilon}{4} + \frac{1 - \varepsilon}{4} = \frac{1}{2}.$$

Līdzīgi

$$P(X = 1) = P(Y = 0) = P(Y = 1) = \frac{1}{2}.$$

Iegūsim gadījumu lielumu X un Y varbūtību nosacītos sadalījumus

$$P(X=0 \mid Y=0)=\frac{P(X=0,Y=0)}{P(Y=0)}=\frac{2(1+\varepsilon)}{4}=\frac{1+\varepsilon}{2}.$$

Līdzīgi

$$P(X=1 \mid Y=1)=P(Y=0 \mid X=0)=P(Y=1 \mid X=1)=\frac{1+\varepsilon}{2};$$

$$\begin{aligned} P(X=1 \mid Y=0)=P(Y=0 \mid X=1)=P(Y=1 \mid X=0) \\ =P(X=0 \mid Y=1)=\frac{1+\varepsilon}{2}. \end{aligned}$$

Tad iegūstam, ka

$$\alpha(\sigma(X),\sigma(Y)) =\mid P(X=0,Y=1)-P(X=0)P(Y=1) \mid= \left|\frac{1-\varepsilon}{4}-\frac{1}{2}\cdot\frac{1}{2}\right|=\frac{\varepsilon}{4};$$

$$\begin{aligned} \beta(\sigma(X),\sigma(Y)) &= \mid P(X=0,Y=1)-P(X=0)P(Y=1) \mid \\ &+ \mid P(X=1,Y=0)-P(X=1)P(Y=0) \mid \\ &+ \mid P(X=0,Y=0)-P(X=0)P(Y=0) \mid \\ &+ \mid P(X=1,Y=1)-P(X=1)P(Y=1) \mid \\ &= 2\left|\frac{1-\varepsilon}{2}-\frac{1}{4}\right|+2\left|\frac{1+\varepsilon}{2}-\frac{1}{4}\right|=\frac{\varepsilon}{2}; \end{aligned}$$

$$\begin{aligned} \phi(\sigma(X),\sigma(Y)) = \sup_{i,j} \mid P(Y=j|X=i)-P(Y=j) \mid \\ =\mid P(Y=1|X=1)-P(Y=j) \mid = \frac{2(1+\varepsilon)}{4}-\frac{1}{2}=\frac{\varepsilon}{2}; \end{aligned}$$

$$\begin{aligned} \psi(\sigma(X),\sigma(Y)) = \left|\sup_{i,j}\frac{P(X=i,Y=j)}{P(X=i)P(Y=j)}-1\right| \\ =\left|\frac{P(X=1,Y=1)}{P(X=1)P(Y=1)}-1\right|=\frac{4(1+\varepsilon)}{4}-1=\varepsilon; \end{aligned}$$

$$\rho(\sigma(X),\sigma(Y)) \;=\mid\; Corr(X,Y) \;\mid= \; \frac{E(XY)-E(X)E(Y)}{\sqrt{E(X^2)-(E(X))^2}\sqrt{E(Y^2)-(E(Y))^2}} \;=\; \varepsilon,$$

jo

$$E(XY)=\frac{1+\varepsilon}{4},\ EX=EY=EX^2=EY^2=\frac{1}{2}.$$

Varam redzēt, ka izpildās arī 6.apgalvojums.

Piemērs 2. Tika izmantota literatūra [[1], 100.lpp].

Pieņemsim, ka $0 < \varepsilon \leq 1/2$ un X, Y ir gadījuma lielumi ar sekojošiem gadījuma lielumu X, Y kopējiem sadalījumiem:

$$P(X = Y = 0) = 1 - \varepsilon,$$

$$P(X = 1, Y = 1) = \varepsilon.$$

Tad iegūstam, ka

$$\alpha(\sigma(X), \sigma(Y)) = \mid P(X = 1, Y = 1) - P(X = 1)P(Y = 1) \mid = \mid \varepsilon - \varepsilon^2 \mid = \varepsilon(1 - \varepsilon);$$

$$\beta(\sigma(X), \sigma(Y)) = 2\alpha(\sigma(X), \sigma(Y)) = 2\varepsilon(1 - \varepsilon);$$

$$\begin{aligned}\phi(\sigma(X), \sigma(Y)) &= \sup_{i,j} \mid P(Y = j|X = i) - P(Y = j) \mid \\ &= \sup_{i,j} \left| \frac{P(Y = j, X = i)}{P(X = i)} - P(Y = j) \right| \\ &= \left| \frac{P(Y = 1, X = 1)}{P(X = 1)} - P(Y = 1) \right| \\ &= \left| \frac{\varepsilon}{\varepsilon} - \varepsilon \right| = 1 - \varepsilon;\end{aligned}$$

$$\begin{aligned}\psi(\sigma(X), \sigma(Y)) = \left| \sup_{i,j} \frac{P(X = i, Y = j)}{P(X = i)P(Y = j)} - 1 \right| \\ = \left| \frac{P(X = 1, Y = 1)}{P(X = 1)P(Y = 1)} - 1 \right| = \frac{1 - \varepsilon}{\varepsilon};\end{aligned}$$

$$\rho(\sigma(X), \sigma(Y)) = \mid Corr(X, Y) \mid = \frac{E(XY) - E(X)E(Y)}{\sqrt{E(X^2) - (E(X))^2}\sqrt{E(Y^2) - (E(Y))^2}} = 1,$$

jo $E(XY) = EX = EY = EX^2 = EY^2 = \varepsilon$.

Ja ε ir mazs, tad arī $\alpha(\sigma(X), \sigma(Y))$ un $\beta(\sigma(X), \sigma(Y))$ būs mazi, bet $\psi(\sigma(X), \sigma(Y))$ būs liels.

Piemērs 3. Tika izmantota literatūra [[1], 101.lpp].

Pieņemsim, ka $0 < \varepsilon \leq 1/2$ un X, Y ir gadījuma lielumi ar sekojošiem gadījuma lielumu X, Y kopējiem sadalījumiem:

$$P(X = Y = 0) = 1 - \varepsilon,\ P(X = 0, Y = 1) = 0,$$

$$P(X = 1, Y = 0) = \varepsilon - \varepsilon^2 \text{ un } P(X = 1, Y = 1) = \varepsilon^2.$$

Iegūsim gadījumu lielumu X un Y marginālos sadalījumus

$$P(X = 0) = 1 - \varepsilon;\ P(Y = 0) = 1 - \varepsilon^2;\ P(X = 1) = \varepsilon;\ P(Y = 1) = \varepsilon^2.$$

Tad iegūstam, ka

$$\alpha(\sigma(X), \sigma(Y)) = \mid P(X = 1, Y = 1) - P(X = 1)P(Y = 1) \mid = \varepsilon^2 - \varepsilon^3;$$

$$\beta(\sigma(X), \sigma(Y)) = 2(\varepsilon^2 - \varepsilon^3);$$

$$\begin{aligned}
\phi(\sigma(X), \sigma(Y)) &= \sup_{i,j} \mid P(Y = j | X = i) - P(Y = j) \mid \\
&= \sup_{i,j} \left| \frac{P(Y = j, X = i)}{P(X = i)} - P(Y = j) \right| \\
&= \left| \frac{P(Y = 1, X = 1)}{P(X = 1)} - P(Y = 1) \right| \\
&= \left| \frac{\varepsilon}{\varepsilon} - \varepsilon \right| = \varepsilon - \varepsilon^2;
\end{aligned}$$

$$\begin{aligned}
\psi(\sigma(X), \sigma(Y)) = \left| \sup_{i,j} \frac{P(X = i, Y = j)}{P(X = i)P(Y = j)} - 1 \right| \\
= \left| \frac{P(X = 1, Y = 1)}{P(X = 1)P(Y = 1)} - 1 \right| = \frac{1 - \varepsilon}{\varepsilon};
\end{aligned}$$

$$\begin{aligned}
\rho(\sigma(X), \sigma(Y)) = \mid Corr(X, Y) \mid = \frac{E(XY) - E(X)E(Y)}{\sqrt{E(X^2) - (E(X))^2}\sqrt{E(Y^2) - (E(Y))^2}} \\
= \frac{\varepsilon^2 - \varepsilon^3}{\sqrt{\varepsilon - \varepsilon^2}\sqrt{\varepsilon^2 - \varepsilon^4}},
\end{aligned}$$

jo $E(XY) = \varepsilon^2\ EX = EX^2 = \varepsilon\ EY = EY^2 = \varepsilon^2$.

Ja ε ir ļoti mazs, tad $\alpha(\sigma(X), \sigma(Y))$, $\beta(\sigma(X), \sigma(Y))$ un $\phi(\sigma(X), \sigma(Y))$ arī ir mazi, bet $\psi(\sigma(X), \sigma(Y))$ ir liels.

2. Ergodiskā teorēma un Centrālā robežteorēma

Šajā nodaļā tiks apskatīta Ergodiskā un Centrālā robežteorēma stacionāriem α jauktajiem procesiem. Lai attēlotu bieži jaukto procesu gadījumā lietoto bloku tehniku, īsi tiks ieskicēts pierādījums Centrālajai robežteorēmai. Dinamiskām sistēmām ergodiskā teorēma apraksta to uzvedību ilgā laika periodā. Šī īpašība ir vājāka kā jaukto procesu īpašība. Ergodisko teorēmu var uzskatīt kā Lielo skaitļu likuma vispārinājumu.

2.1. Ergodiskā teorēma

Lai apskatītu Ergodisko teorēmu, sākumā nepieciešams apskatīt dažas definīcijas.

Pieņemsim, ka ir dota varbūtību telpa $(\Omega, \mathcal{F}, P)$.

Definīcija 6. $T : \Omega \to \Omega$ ir mēru saglabājoša transformācija, ja tā ir $\mathcal{F}$-mērojama un $\forall A \in \mathcal{F}\ P(T^{-1}A) = P(A)$.

No 6.definīcijas seko, ka $\forall n\ P(T^{-n}A) = P(A)$.

Definīcija 7. Kopa $A \in \mathcal{F}$ ir stacionāra saskaņā ar T, ja $T^{-1}A = A$. $0 < P(A) < 1$.

Pieņemsim, ka $X(\omega) = (X_k(\omega), k \in \mathbb{Z})$, $X(\omega) = \omega$, tad $f(X(\omega)) = f(\omega)$. $\forall j$ un $\forall \omega \in \Omega$ definēsim $T^j(X(\omega)) = (X_{j-k}, k \in \mathbb{Z})$ $(j = 1, 2, ...)$.

Definīcija 8. Mērojama funkcija f ir stacionāra, ja $\forall \omega\ f(T\omega) = f(\omega)$. Kopa A ir stacionāra tad un tikai tad, ja indikātorfunkcija I_A ir stacionāra.

Definīcija 9. T ir ergodiska, ja σ-algebra $\mathcal{F}$ sastāv no netukšām stacionārām kopām.

Teorēma 8. *[Ergodiskā teorēma]*

Pieņemsim, ka T ir mēru saglabājoša transormācija varbūtību telpā $(\Omega, \mathcal{F}, P)$ un f ir mērojama un integrējama, T ir ergodiska. Tad

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} f(T^{k-1}\omega) = Ef(\omega)$$

ar varbūtību 1.

Šī ir klasiskā teorēma (Georga Birkhofa [1932]). Tai ir vairākas formas, bieži šo teorēmu sauc arī par "Birkhofa Ergodisko teorēmu".

Pierādījums. Skatīt [[2], 314.lpp]

Piezīme 9. *Ja $f = I_A$, tad*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} I_A(T^{k-1}\omega) = E(I_A(\omega))$$

un ergodiskā gadījumā

$$I_A(\omega) = P(A).$$

ar varbūtību 1.

2.2. Centrālās robežteorēmas

Definīcija 10. Procesu $(X_t, t \in \mathbb{Z})$ sauc par stacionāru, ja procesa galīgi dimensionālie sadalījumi paliek nemainīgi pie patvaļīgas laika nobīdes, t.i,

$$F_{X_{t_1+s},\ldots,X_{t_n+s}}(X_1, \ldots, X_n) = F_{X_{t_1},\ldots,X_{t_n}}(X_1, \ldots, X_n)\ \forall n,\ \forall s.$$

Pieņemsim, ka $(X_t, t \in \mathbb{Z})$ ir α-jauktais process ar stacionāru sadalījumu. Definēsim $S_n = X_1 + X_2 + \ldots + X_n$ un pieņemsim, ka eksistē $\gamma(k) = Corr(X_{t+k}, X_t)$, kur $k \in \mathbb{Z}$.

Teorēma 10. *Pieņemsim, ka izpildās vismaz viens no sekojošiem apgalvojumiem*

1. $\mid EX_t \mid^{\sigma} < \infty$ *un* $\sum_{j\geq 1} \alpha(j)^{1-2/\sigma} < \infty$ *kādām konstantēm* $\sigma > 2$*;*

2. $P(\mid X_t \mid < C) = 1$ *kādām kosntantēm* $C > 0$ *un* $\sum_{j\geq 1} \alpha(j) < \infty$*,*

tad

$$\sum_{j\geq 1} \mid \gamma(j) \mid < \infty$$

un

$$\frac{1}{n}D(S_n) \to \gamma(0) + 2\sum_{j=1}^{\infty}\gamma(j),$$

kad $n \to \infty$.

Pierādījums. Skatīt [[3], 74.lpp].

Apgalvojums 11. *Pieņemsim, ka* $(X_t, t \in \mathbb{Z})$ α *jauktais process ar stacionāru sadalījumu un* $EX_t = 0$. *Pieņemsim, ka* α *jauktais procesu koeficients ir formā (1.1.1) un* $q \geq 2$. *Tad*

$$|E(S_n^q)| \leq Cn^{q/2},$$

kur C *ir konstante, ja izpildās viens no sekojošiem apgalvojumiem*

1. $E|X_t| < \infty$ *kādam* $\sigma > q$ *un* $\alpha(n) = O(n^{-\frac{\sigma q}{2(\sigma - q)}})$;

2. $P(|X_t| < C_1) = 1$ *kādām konstantēm* C_1 *un* $\alpha(n) = O(n^{-\frac{q}{2}})$.

Pierādījums. Skatīt [[3], 72.lpp].

Apgalvojums 12. *Pieņemsim, ka* α *jauktais procesu koeficients ir formā (1.1.1),* ξ_1,...,ξ_k *ir gadījuma lielumi, kuri pieņem kompleksas vērtības* ($k \geq 2$), *ar* σ*-algebrām* $\mathcal{F}_{i_1}^{j_1}, ..., \mathcal{F}_{i_k}^{j_k}$. *Pieņemsim, ka* $\forall l = 1, ..., k-1$ $i_{l+1} - j_l \geq k$ *un* $\forall l = 1, ..., k$ $j_l \geq i_l$ *un* $P(|\xi_l| \leq 1) = 1$. *Tad*

$$|E(\xi_1, ..., \xi_k) - E(\xi_1)...E(\xi_k)| \leq 16(k-1)\alpha(n)$$

Pierādījums. Skatīt [[1], 32.lpp].

Nodefinēsim **Lindeberga nosacījumu**. Pieņemsim, ka ir dota gadījumu lielumu virkne X_1, X_2, ..., X_n, $E(X_i) = 0$, ${\sigma_i}^2(X_i)$ ir ierobežota,

$$S_n \equiv X_1 + X_2 + \ldots + X_n$$

un

$${s_n}^2 \equiv \sum_{i=1}^{n} {\sigma_i}^2.$$

Pieņemsim, ka

$$\Lambda_n(\epsilon) \equiv \sum_{k=1}^{n} \int_{|X_k| \geq s_n \epsilon} \left(\frac{X_k}{s_n}\right)^2 dP$$

$\forall \epsilon > 0$, tad Lindeberga nosacījums ir

$$\lim_{n \to \infty} \Lambda_n(\epsilon) = 0$$

$\forall \epsilon > 0$.

Teorēma 13. *[Lindeberga-Levi teorēma]*

Pieņemsim, ka dota gadījuma lielumu virkne $X_1, X_2, ..., X_n$, ${\sigma_i}^2(X_i)$ ir ierobežota, $E(X_i) = 0$,

$$S_n \equiv X_1 + X_2 + \ldots + X_n,$$

$${s_n}^2 \equiv \sum_{i=1}^{n} {\sigma_i}^2,$$

un izpildās Lindeberga nosacījums $\forall \epsilon > 0$, tad

$$\frac{S_n}{s_n} \to N(0, \sigma^2),$$

kad $n \to \infty$.

Pierādījums. Skatīt [[2], 359.lpp].

Tagad tiks apskatīta un pierādīta centrālā robežteorēma α jauktajiem procesiem.

Teorēma 14. *Pieņemsim, ka $EX_t = 0$ un*

$$\sigma^2 = \gamma(0) + 2 \sum_{j=1}^{\infty} \gamma(j) \ (\sigma^2 \geq 0),$$

tad

$$\frac{S_n}{\sqrt{n}} \to_d N(0, \sigma^2),$$

ja izpildās vismaz viens no sekojošajiem apgalvojumiem

1. *$\mid EX_t \mid^\sigma < \infty$ un $\sum_{j \geq 1} \alpha(j)^{1-2/\sigma} < \infty$ kādām konstantēm $\sigma > 2$;*

2. *$P(\mid X_t \mid < C) = 1$ kādām kosntantēm $C > 0$ un $\sum_{j \geq 1} \alpha(j) < \infty$.*

Pierādījums. Pierādīsim tikai 2.apgalvojumu. 1.apgalvojuma pierādījumu skatīt [[4], 34.lpp].

Lai izmantotu mazo un lielo bloku argumentus, sadalīsim kopu $\{1, ..., n\}$ $2k_n + 1$ apakškopās ar lielo bloku argumentiem apjomā l_n un ar mazo bloku argumentiem apjomā s_n un beidzamā atlikusī kopa apjomā $n - k_n(l_n + s_n)$, kur l_n un s_n ir tādi, ka

$$s_n \to \infty,\ \frac{s_n}{l_n} \to 0,\ \frac{l_n}{n} \to 0 \text{ un } k_n = \left[\frac{n}{l_n + s_n}\right] = O(s_n).$$

Piemēram, ja $\forall r > 2$ mēs izvēlēsimies $l_n = O(n^{\frac{r-1}{r}})$ un $s_n = O(n^{\frac{1}{r}})$,

tad $k_n = O(n^{\frac{1}{r}}) = O(s_n)$. $\forall j = 1, ..., k_n$ definēsim

$$\xi_j = \sum_{i=(j-1)(l_n+s_n)+1}^{jl_n+(j-1)s_n} X_i,\ \eta_j = \sum_{i=jl_n+(j-1)s_n}^{j(l_n+s_n)} X_i \text{ un } \zeta = \sum_{i=k_n(l_n+s_n)+1}^{n} X_i$$

ξ_j $(j = 1, 2, ..., l_n)$ ir lielais bloks un η_j $(j = 1, 2, ..., s_n)$ ir mazais bloks, bet ζ ir atlikušais lielums, kas netika iekļauts ne mazajā ne lielajā blokā.

Atzīmēsim, ka $\alpha(n) = o(n^{-1})$ un $((k_n s_n)/n) \to 0$.

No 11.apgalvojuma seko, ka

$$\frac{1}{n} E\left(\sum_{j=1}^{k_n} \eta_j\right)^2 \to 0.$$

Tādā veidā

$$\frac{1}{\sqrt{n}} S_n = \frac{1}{\sqrt{n}}\left(\sum_{j=1}^{k_n} \xi_j + \sum_{j=1}^{k_n} \eta_j + \zeta\right) = \frac{1}{\sqrt{n}} \sum_{j=1}^{k_n} \xi_j + o_p(1).$$

No 12.apgalvojuma seko, ka

$$\left| E\left\{exp\left(\frac{it}{\sqrt{n}} \sum_{j=1}^{k_n} \xi_j\right)\right\} - \prod_{j=1}^{k_n} E\left\{exp\left(\frac{it\xi_j}{\sqrt{n}}\right)\right\}\right| \le 16(k_n - 1)\alpha(s_n) \to 0.$$

Izmantojot 10.teorēmas 2.nosacījumu, iegūstam, ka $l^{-1}E\xi_1^2 \to \sigma^2$ un

$$\frac{1}{n}\sum_{j=1}^{k_n} E\xi_j^2 = \frac{k_n l_n}{n} \cdot \frac{1}{l_n} E\xi_1^2 \to \sigma^2.$$

No 11.apgalvojuma seko, ka

$$E\left\{\xi_1^2 I\left(|\xi_1| \ge \varepsilon\sigma n^{\frac{1}{2}}\right)\right\} \le E\left\{\xi_1^4\right\} P\left\{|\xi_1| \ge \varepsilon\sigma n^{\frac{1}{2}}\right\} \le C l_n \frac{1}{n\varepsilon^2\sigma^2} E\xi_1^2 = O\left(\frac{l_n^2}{n}\right).$$

Varam redzēt, ka izpildās Lindeberga nosacījums

$$\frac{1}{n}\sum_{j=1}^{k_n} E\left\{\xi_j^2 I\left(|\xi_j| \ge \varepsilon\sigma n^{\frac{1}{2}}\right)\right\} = O\left(\frac{k_n l_n^2}{n^2}\right) = O\left(\frac{l_n}{n}\right) \to 0.$$

Un no 13.teorēmas seko, ka

$$\prod_{j=1}^{k_n} E\left(e^{\frac{it\xi_j}{\sqrt{n}}}\right) \to e^{-\frac{t^2\sigma^2}{2}}.$$

3. Jauktie procesi - stacionāras lineāras un nelineāras laikrindas

Praksē ir ļoti grūti pārbaudīt vai dati ir jauktais process, bet ja zināms modelis pēc kāda tie ir modelēti, tad jaukto procesu nosacījumus iesējams pārbaudīt. Tādēļ šajā nodaļā tiks apskatītas stacionāras lineāras un nelineāras jaukto procesu laikrindas, kā arī tiks analizēti ilglaicīgās atmiņas procesi. Pazīstamākie jaukto procesu modeļi ir lineāri un nelineāri ARIMA, ARCH, GARCH modeļi, Markova ķēdes u.c. (skat. [5]).

3.1. Markova ķēdes

Teorētiskajā analīzē tiks izmantota literatūra [[1], 7.nodaļa, [6] un [2], 8.nodaļa].

Definīcija 11. Virkne $\{X_n, n \in \mathbb{Z}\}$ veido Markova ķēdi vai Markova procesu, ja

$$P(X_n = j \mid X_0 = k_0, X_1 = k_1, ..., X_{n-2} = k_{n-2}, X_{n-1} = i)$$
$$= P(X_n = j \mid X_{n-1} = i) = p_{ij}^n,$$

kur $p_{i,j}^n$ ir pārejas varbūtība no i un j viena soļa laikā n-tajā solī, $\sum_i p_{i,j}^n = 1$, $i, j, k_l \in S$, S- procesa stāvokļu kopa.

Definīcija 12. Pārejas varbūtību matrica P Markova ķēdēm ir $n \times n$ matrica , kuras (i, j) elements ir p_{ij}. Matricu P sauc par stohastisku matricu, kur

a) $0 \leq p_{ij} \leq 1$, kur $i = 1, 2, ...n$ un $j = 1, 2, ..., n$,

b) $\displaystyle\sum_{j=1}^{n} p_{i,j} = 1$, kur $i = 1, 2, ..., n$.

Definīcija 13. Markova ķēde ir nereducējama, ja $\forall i, j$ $p_{i,j}^n > 0$, pretējā gadījumā Markova ķēde ir reducējama.

Definīcija 14. Par stāvokļa i periodu $d = d_i$ sauc lielāko kopīgo dalītāju izteiksmei $\{n \geq 0 : p_{i,i}^n > 0\}$. Markova ķēdi sauc par neperiodisku, ja $\forall i \; d_i = 1$.

Pieņemsim, ka Markova ķēdes sākotnējā varbūtība ir π_i, kas apmierina vienādību

$$\sum_{i \in S} \pi_i p_{i,j} = \pi_j, \text{ kur } j \in S,$$

tad pēc indukcijas seko, ka

$$\sum_{i \in S} \pi_i p_{i,j}^n = \pi_j, \text{ kur } j \in S \text{ un } n = 0, 1, 2, ...$$

Ja $\pi_i = P(X_0 = i)$, tad

$$\sum_{i \in S} \pi_i p_{i,j}^n = P(X_n = j).$$

Definīcija 15. Varbūtību vektoru π sauc par stacionāru sadalījumu pārejas matricai P, ja $\pi P = \pi$.

Teorēma 15. *Pieņemsim, ka $X = \{X_n, n \in \mathbb{Z}\}$ ir Markova ķēde ar stacionāru sadalījumu, tad $\forall n$ ir spēkā sekojoši apgalvojumi:*

a) $\alpha(n) = \alpha(\mathcal{F}_0^0, \mathcal{F}_n^n)$; *b)* $\beta(n) = \beta(\mathcal{F}_0^0, \mathcal{F}_n^n)$; *c)* $\rho(n) = \rho(\mathcal{F}_0^0, \mathcal{F}_n^n)$;

d) $\phi(n) = \phi(\mathcal{F}_0^0, \mathcal{F}_n^n)$; *e)* $\psi(n) = \psi(\mathcal{F}_0^0, \mathcal{F}_n^n)$.

Pierādījums. Skatīt [[1], 206.lpp].

Var redzēt, ka jaukto procesu koeficienti ir atkarīgi tikai no gadījuma lieluma sākotnējā un beidzamā stāvokļa.

Teorēma 16. *Pieņemsim, ka $X = \{X_n, n \in \mathbb{Z}\}$ ir Markova ķēde ar stacionāru sadalījumu un sanumurējamu stāvokļā kopu S, μ ir (invariants) marginālais sadalījums. Pieņemsim, ka $\forall i \in S \; \mu_i = P(X_0 = i) > 0$. Ja X ir nereducējama un neperiodiska, tad izpildās sekojoši apgalvojumi:*

1. $\forall j \in S \; \lim_{n \to \infty} p_{j,i}^n = \lim_{n \to \infty} P(X_n = i \mid X_0 = j) = \mu_i$;

2. $\forall j \in S \; \lim_{n \to \infty} \sum_{i \in S} \mid P(X_n = i \mid X_0 = j) - \mu_i \mid = 0$.

Pierādījums. Teorēmas 1.daļas pierādījumu skatīt [[2], 125.lpp] un teorēmas 2.daļas pierādījums seko no teorēmas 1.daļas.

Teorēma 17. *[Dominējošā konverģence]*

Pieņemsim, ka mums dota varbūtību telpa $(\Omega, \mathcal{F}, P)$ *un gadījumu lielumu virkne -* $(X_n, n \in \mathbb{Z})$ $(X_n \in \mathcal{F}, n \in \mathbb{Z})$ *un gadījuma lielums* Y $(Y \in \mathcal{F})$ *Ja* $|X_n| \leq Y$*, kur gadījuma lielums* Y *ir tāds, ka* $\forall n$ $EY < \infty$ *un* $X_n \to_p X$*, kad* $n \to \infty$*, tad arī* $EX_n \to EX$*.*

Pierādījums. Skatīt [[2], 77.lpp]

Definīcija 16. Pieņesim, ka T ir mēru-saglabājoša transormācija varbūtību telpā $(\Omega, \mathcal{F}, P)$. Stacionāra gadījumu lielumu virkne $X = (X_n, n \in \mathbb{Z})$ sauc par jaukto procesu ergodiskā nozīmē, ja

$$\lim_{n \to \infty} P(X \in A \cap T^{-n}B) = P(X \in A)P(X \in B)\ ,\ \forall A, B \in \mathcal{F}$$

Teorēma 18. *Pieņemsim, ka* $X = (X_n, n \in \mathbb{Z})$ *ir Markova ķēde ar stacionāru sadalījumu un sanumurējamu stāvokļa kopu* S*. Tad sekojošie apgalvojumi ir ekvivalenti:*

a) X *ir nereducējama un neperiodiska;*

b) X *ir jauktais process ergodiskā nozīmē;*

c) $\alpha(n) \to 0$ *, kad* $n \to \infty$*;*

d) $\beta(n) \to 0$ *, kad* $n \to \infty$*.*

Pierādījums. Pieņemsim, ka μ ir marginālais sadalījums. $\forall i \in S$ $\mu_i > 0$. Acīmredzami ir $d) \Rightarrow c) \Rightarrow b)$, tādēļ jāpierāda tikai, ka $b) \Rightarrow a)$ un $a) \Rightarrow d)$.
Pierādīsim, ka $b) \Rightarrow a)$. Pieņemsim, ka b) apgalvojums izpildās, $i, j \in S$, tad no b) apgalvojuma seko, ka

$$\mu_i P(X_n = j \mid X_0 = i) = P(X_n = j, X_0 = i) \to \mu_i \mu_j > 0,$$

kad $n \to \infty$. Tādejādi

$$\forall n P(X_n = j \mid X_0 = i) = p_{i,j}^n > 0, \tag{3.1.1}$$

kas parāda, ka Markova ķēde X ir nereducējama (13.definīcija). Pierādīsim, ka Markova ķēde X ir arī neperiodiska no pretējā.
Pieņemsim, ka Markova ķēde X ir periodiska ar periodu $d > 1$. Paņemsim patvaļīgu stāvokli $i \in S$. Stāvoklis i ir ar periodu d. Ir redzams, ka $P(X_n = i \mid X_0 = i) = 0$, kas ir pretrunīga ar (3.1.1) nevienādību. Tādēļ mūsu pieņēmums, ka Markova ķēde X ir periodiska, ir aplams. Markova ķēde X ir neperiodiska.

Pierādīsim, ka a) $\Rightarrow$ d). Pieņemsim, ka a) agalvojums izpildās, tad pēc 16.teorēmas seko, ka

$$\forall j \ \mu_j \sum_{i\in S} \mid P(X_n = i \mid X_0 = j) - \mu_i \mid \to 0,$$

kad $n \to \infty$. Arī $\forall j \in S$

$$\mu_j \sum_{i\in S} \mid P(X_n = i \mid X_0 = j) - \mu_i \mid \leq \mu_j \sum_{i\in S} \mid P(X_n = i \mid X_0 = j) + \mu_i \mid = 2\mu_j.$$

Kamēr $\sum_{j\in S} 2\mu_j = 2 < \infty$, tad pēc dominējošās konverģences (17.teorēma) izriet, ka

$$\sum_{j\in S} \mu_j \sum_{i\in S} \mid P(X_n = i \mid X_0 = j) - \mu_i \mid \to 0, \tag{3.1.2}$$

kad $n \to \infty$. Izmantojot vienādību (1.1.5) un 15.teorēmas b) apgalvojumu, iegūstam, ka

$$\begin{aligned}
\beta(n) &= \beta(\sigma(X_0), \sigma(X_n)) \\
&= \frac{1}{2} \sum_{j\in S} \sum_{i\in S} \mid P(X_0 = j, X_n = i) - P(X_0 = j) P(X_0 = i) \mid \\
&= \frac{1}{2} \sum_{j\in S} \sum_{i\in S} \mid P(X_0 = j, X_n = i) - \mu_j \mu_i \mid \\
&= \frac{1}{2} \sum_{j\in S} \mu_j \sum_{i\in S} \mid P(X_n = i \mid X_0 = j) - \mu_i \mid .
\end{aligned}$$

Tad no (3.1.2) vienādības izriet, ka $\beta(n) \to 0$, kad $n \to \infty$, kas pēc 4.definīcijas 2.apgalvojuma nozīmē, ka process X ir β jauktais process.

Teorēma 19. *Pieņemsim, ka $X = \{X_n, n \in \mathbb{Z}\}$ ir Markova ķēde ar stacionāru sadalījumu un bezgalīgu stāvokļa kopu S. Tad sekojošie apgalvojumi ir ekvivalenti:*

a) *X ir nereducējama un neperiodiska;*

b) *X ir jauktais process ergodiskā nozīmē;*

c) *$\alpha(n) \to 0$, kad $n \to \infty$;*

d) *$\psi \to 0$, kad $n \to \infty$*

Pierādījums. Skatīt lieratūru [[1], 220.lpp].

No 18. un 19.teorēmas izriet, ka jebkura nereducējama, neperiodiska Markova ķēde ar sanumurējamu vai bezgalīgu stāvokļa kopu ir α jauktais process.

Apskatīsim vienu piemēru, kur parādīts, kā aprēķina jaukto procesu koeficientus Markova ķēdēm.

Piemērs 4. Tika izmantota literatūra [[7], 176.lpp un [1], 215.lpp].

Pieņemsim, ka $0 \leq \varepsilon \leq 1$ un $X = (X_n, n \in \mathbb{Z})$ ir Markova ķēde ar stacionāru sadalījumu, ar stāvokļa kopu $S = \{0, 1\}$, ar marginalajiem sadalījumiem

$$P(X_0 = 0) = P(X_0 = 1) = \frac{1}{2}$$

un viena soļa pārejas varbūtību matricu

$$P = \begin{pmatrix} (1/2)(1+\varepsilon) & (1/2)(1-\varepsilon) \\ (1/2)(1-\varepsilon) & (1/2)(1+\varepsilon) \end{pmatrix}.$$

$\forall n > 0$ n-soļu pārejas varbūtību matrica

$$P^n = \begin{pmatrix} (1/2)(1+\varepsilon^n) & (1/2)(1-\varepsilon^n) \\ (1/2)(1-\varepsilon^n) & (1/2)(1+\varepsilon^n) \end{pmatrix}.$$

Aprēķināsim kopējos varbūtību sadalījumus

$$\begin{aligned} P(X_0 = 0, X_1 = 0) &= P(X_0 = 1, X_1 = 1) \\ &= P(X_1 = 1 \mid X_0 = 1)P(X_0 = 1) = \frac{1+\varepsilon^n}{2} \cdot \frac{1}{2} = \frac{1+\varepsilon^n}{4}; \end{aligned}$$

$$P(X_0 = 0, X_1 = 1) = P(X_0 = 1, X_1 = 0) = \frac{1+\varepsilon^n}{4}.$$

Tad iegūstam, ka

$$\begin{aligned} \alpha(n) = \alpha(\mathcal{F}_0^0, \mathcal{F}_n^n) &= |P(X_0 = 0, X_1 = 1) - P(X_0)P(X_1 = 1)| \\ &= \left| \frac{1-\varepsilon^n}{4} - \frac{1}{2} \cdot \frac{1}{2} \right| = \frac{\varepsilon^n}{4}; \end{aligned}$$

$$\begin{aligned} \beta(n) &= \beta(\mathcal{F}_0^0, \mathcal{F}_n^n) = \mid P(X_0 = 0, X_1 = 1) - P(X_0 = 0)P(X_1 = 1) \mid \\ &+ \mid P(X_0 = 1, X_1 = 0) - P(X_0 = 1)P(X_1 = 0) \mid \\ &+ \mid P(X_0 = 0, X_1 = 0) - P(X_0 = 0)P(X_1 = 0) \mid \\ &+ \mid P(X_0 = 1, X_1 = 1) - P(X_0 = 1)P(X_1 = 1) \mid \\ &= 2\left| \frac{1-\varepsilon^n}{2} - \frac{1}{4} \right| + 2\left| \frac{1+\varepsilon^n}{2} - \frac{1}{4} \right| = \frac{\varepsilon^n}{2}; \end{aligned}$$

$$\begin{aligned} \phi(n) = \phi(\mathcal{F}_0^0, \mathcal{F}_n^n) &= \sup_{i,j} \mid P(X_1 = j | X_0 = i) - P(X_1 = j) \mid \\ &= \mid P(X_1 = 1 | X_0 = 1) - P(X_1 = j) \mid = \frac{2(1+\varepsilon^n)}{4} - \frac{1}{2} = \frac{\varepsilon^n}{2}; \end{aligned}$$

$$\psi(n) = \psi(\mathcal{F}_0^0, \mathcal{F}_n^n) = \left| \sup_{i,j} \frac{P(X_0 = i, X_1 = j)}{P(X_0 = i)P(X_1 = j)} - 1 \right|$$
$$= \left| \frac{P(X_0 = 1, X_1 = 1)}{P(X_0 = 1)P(X_1 = 1)} - 1 \right| = \frac{4(1+\varepsilon^n)}{4} - 1 = \varepsilon^n;$$

$$\rho(n) = \rho(\mathcal{F}_0^0, \mathcal{F}_n^n) =\mid Corr(X_0, X_1) \mid$$
$$= \frac{E(X_0X_1) - E(X_0)E(X_1)}{\sqrt{E(X_0^2) - (E(X_0))^2}\sqrt{E(X_0^2) - (E(X_1))^2}} = \varepsilon^n,$$

jo

$$E(X_0X_1) = \frac{1+\varepsilon^n}{4}, \; EX_0 = EX_1 = EX_0^2 = EX_1^2 = \frac{1}{2}.$$

Ja $n \to 0$, tad $\alpha(n) \to 0$, $\beta(n) \to 0$, $\phi(n) \to 0$, $\rho(n) \to 0$ un $\psi(n) \to 0$. Markova ķēde X ir jauktais process.

3.2. ARMA modeļi

Definīcija 17. Laikrinda $(X_t, t \in \mathbb{Z})$ ir modelēta pēc ARMA(p,q) (*autoregresīvā vidēji slīdošā*) modeļa, ja tā ir stacionāra un

$$X_t = b_1X_{t-1} + ... + b_pX_{t-p} + \varepsilon_t + a_1\varepsilon_{t-1} + ... + a_q\varepsilon_{t-q}$$

ar $b_p \neq 0$, $a_q \neq 0$ un $\sigma_\varepsilon^2 > 0$, kur $\{\varepsilon_t, t \in \mathbb{Z}\}$ neatkarīgi vienādi sadalīti gadījuma lielumi.

Definīcija 18. Ja $q = 0$, tad modeli sauc par AR(p) (*autoregresīvo*) modeli un

$$X_t = b_1X_{t-1} + ... + b_pX_{t-p} + \varepsilon_t, \tag{3.2.1}$$

No vienādības formā (3.2.1) var redzēt, ka $(X_t, t \in \mathbb{Z})$ ir korelēts un X_t $(t \in \mathbb{Z})$ ir atkarīgs no visām iepriekšējām vērtībām.

Definīcija 19. Ja $p = 0$, tad modeli sauc par MA(q) (*vidēji slīdošo*) modeli un

$$X_t = \varepsilon_t + a_1\varepsilon_{t-1} + ... + a_q\varepsilon_{t-q},$$

Korelācija starp X_t un X_{t-h} pastāv, tā var būt atkarīga no ε_{t-h} $(t, h \in \mathbb{Z})$. X_t un X_{t-h} nav korelēti gadījumā, kad $h > q$ $(t, h \in \mathbb{Z})$. q nosaka atkarību no pagātnes, bet ε_t $(t \in \mathbb{Z})$ atjauno informāciju, ko satur σ-algebra $\sigma(\varepsilon_{t-1}, \varepsilon_{t-2}, ..., \varepsilon_{t-q})$.

ARMA(p,q) procesus pielieto finansanšu matemātikā, ķīmijā, meterioloģijā.

Dati, kas modelēti pēc ARMA(p,q) modeļa, var būt

- akciju cenas biržā;
- valūtas apmaiņas kursi;
- ikgadējie nokrišņu daudzumi;
- vidējā mēneša temperatūra, kas fiksēta dažādās vietās;
- kādas konkrētas izejvielas stundas koncentrācijas;
- bezdarbnieku skaits dažādos mēnešos u.c.

Apgalvojums 20. *ARMA process ir jauktais process ar eksponenciāli dilstošu jaukto procesu koeficientu, ja* $b_i < 1$ *(*$i = 1, 2, ..., p$*).*

Pierādījums. Skat. [8]

Apskatīsim piemēru, kurā parādīts, ka AR(1) process $(X_t, t \in \mathbb{Z})$ nav jauktais process.

Piemērs 5. Tika izmantota literatūra [[7], 180.lpp un [9], 8.lpp].

Pieņemsim, ka ir dots neatkarīgs vienādi sadalīts process $(Z_t, t \in \mathbb{Z})$ ar

$$P(Z_t = 0) = P(Z_t = 1) = \frac{1}{2}.$$

Definēsim virkni $(X_t, t \in \mathbb{Z})$ kā

$$X_t = \frac{1}{2}Z_t + \frac{1}{4}Z_{t-1} + \frac{1}{8}Z_{t-2} + \frac{1}{16}Z_{t-3} + ... \tag{3.2.2}$$

Tad $(X_t, t \in \mathbb{Z})$ ir stacionārs AR(1) process, jo

$$X_t = \frac{1}{2}X_{t-1} + \frac{1}{2}Z_t.$$

No (3.2.2) vienādības viegli var redzēt, ka

$$X_t = 2X_{t+1} - Z_{t+1},$$

kas nozīmē, ka $\sigma(X_t) \subset \sigma(X_{t+1})$.

Pēc indukcijas iegūstam, ka $\sigma(X_t) \subset \sigma(X_s, s \geq t + k)$.

$$\begin{aligned}
\alpha(t) &\geq \alpha(\sigma(X_0), \sigma(X_t)) \geq \alpha(\sigma(X_0), \sigma(X_0)) \\
&\geq (P(X_0 < 1/2, X_0 < 1/2) - P(X_0 < 1/2)P(X_0 < 1/2)) \\
&= (1/2) - (1/2)(1/2) = 1/4
\end{aligned}$$

Tākā 4.teorēmas a) agalvojums apgalvo, ka $0 \leq \alpha(\sigma(X_0), \sigma(X_t)) \leq (1/4)$, tad procesam X_t $\alpha(\sigma(X_0), \sigma(X_t)) = 1/4$, tādeļ process X_t nav α jauktais process.

Lai process būtu jauktais process, procesam ir jābūt nepārtrauktam.

3.3. ARCH un GARCH procesi

Definīcija 20. Stacionāru laikrindu $(X_t, t \in \mathbb{Z})$ sauc par ARCH(q) (*autoregresīvo nosacīto heteroskedastisko*) procesu, ja

$$X_t = \sigma_t \varepsilon_t, \tag{3.3.1}$$

kur $\sigma_t^2 = a_0 + b_1 X_{t-1}^2 + ... + b_q X_{t-q}^2$, $a_0 \geq 0$, $b_j \geq 0$ $(j = 1, ..., q)$, ε_t neatkarīgi vienādi sadalīti gadījuma lielumi.

ARCH procesus definēja Engels 1982.gadā. Tie bieži tiek izmantoti finanšu matemātikā, kad laikrinda ir ar lielām vērtībām un dispersiju, kas ir heteroskedastiska (mainīga). ARCH procesa $(X_t, t \in \mathbb{Z})$ sadalījums ir atkarīgs no ε_t $(t \in \mathbb{Z})$ sadalījuma. Piemēram, ja $\varepsilon_t \sim N(0,1)$, tad procesa $(X_t, t \in \mathbb{Z})$ sadalījums ir $N(0, \sigma_t^2)$. Ja $\sum_{j=1}^{p} b_j < 1$, tad

$$X_t \sim N\left(0, \frac{a_0}{1 - \sum_{j=1}^{q} b_j}\right)$$

Definīcija 21. Procesu sauc par GARCH (*ģenerēto autoregresīvo nosacīto heteroskedastisko*) procesu, ja tas ir formā (3.3.1), kur

$$\sigma_t^2 = a_0 + a_1\sigma_{t-1}^2 + ... + a_p\sigma_{t-p}^2 + b_1 X_{t-1}^2 + ... + b_q X_{t-q}^2$$

un $a_j \geq 0$ un $b_j \geq 0$. Dispersija ir

$$\frac{a_0}{1 - \sum_{i=1}^{p} a_i - \sum_{j=1}^{q} b_j}$$

GARCH proecsus Bolerelevs definēja 1986.gadā.

Apgalvojums 21. *GARCH ir α jauktie procesi ar eksponenciāli dilstošu jaukto procesu koeficientu, ja*

a) $\sum_{1 \leq i \leq p} a_i + \sum_{1 \leq j \leq q} b_j < 1$*;*

b) ε_t *$(t \in \mathbb{Z})$ blīvuma funkcija ir pozitīva un nepārtraukta.*

Pierādījums. Skatīt [10].

3.4. Ilglaicīgās atmiņas procesi

Procesa autokorelācijas $\forall h$ apzīmēsim ar $\rho(h)$. Ja vājās atkarības procesiem $\rho(h)$ dilst eksponenciāli ātri, kad $h \to \infty$, tad ilglaicīgās atmiņas procesiem $\rho(h)$ dilst ļoti lēni, kad $h \to \infty$ (skat. [5],[11] un [12]).

Definīcija 22. $(X_t, t \in \mathbb{Z})$ ir ilglaicīgās atmiņas process, ja

$$\rho(h) \sim Ch^{2d-1},$$

kad $h \to \infty$, kur $C \neq 0$ un $d < 0.5$

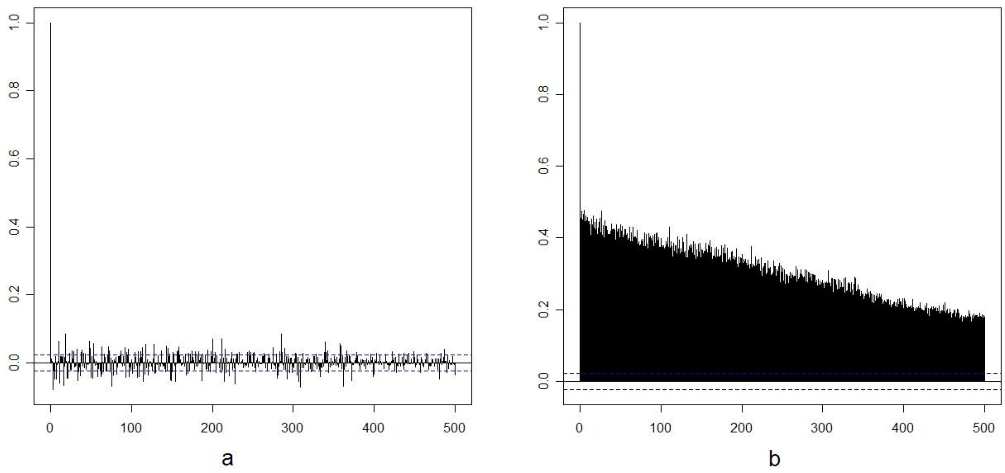
Ilglaicīgās atmiņas procesiem autokovariācijas tiecas uz nulli lēnāk, nekā h^{2d-1} un

$$\sum_{h=0}^{\infty} |\rho(h)| = \infty,$$

kad $d \in (0, 0.5)$.

Ilglaicīgās atmiņas procesi ir sastopami hidroloģijā, finanšu matemātikā un citur. Apskatām vienu piemēru, kur salīdzināta vājās atkarības procesa autokorelācijas funkcija ar ilglaicīgās atmiņas procesa autokorelācijas funkciju.

Piemērs 6. Apskatām datus $(Y_t, t \in \mathbb{Z})$ (literatūras avots [13]), kas balstās uz 500 lielāko Amerikas uzņēmumu akciju svērtajām vidējām cenām (aptuveni 70% no Amerikas tirgus), laika posmā no 1972.gada 3.janvāra līdz 2000.gada 19.janvārim (katras dienas slēgšanas vērtības). 3.1.attēlā attēlotas autokorelācijas funkcijas $Z_t = Y_t - Y_{t-1}$, kas ir vājās atkarības process, un $R_t = |Y_t - Y_{t-1}|$, kas ir ilglaicīgās atmiņas process. Var redzēt, ka ilglaicīgās atmiņas procesa autokorelācijas funkcija dilst daudz lēnāk, nekā vājās atkarības procesam.



3.1. att. Aukorelācijas funkcija a) procesam $\{Z_t, t \in \mathbb{Z}\}$ un b)procesam $\{R_t, t \in \mathbb{Z}\}$

Ilglaicīgās atmiņas process ir arī Frakcionālais integrētais (*fractionally integrated*) - ARMA process (FARIMA(p,d,q) process).

Pieņemsim, ka ir dots process $(X_t, t \in \mathbb{Z})$, tad definēsim pārbīdes operatoru, kam izpildās

$$BX_t = X_{t-1}$$

un $B^2X_t = B(BX_t) = BX_{t-1} = X_{t-2}$,...,$B^kX_t = X_{t-k}$. Acīmredzams, ka

$$\bigtriangledown X_t = (1-B)X_t$$

$\forall d > -1$ definēsim diferences operātoru

$$\bigtriangledown^d \equiv (1-B)^d = \sum_{j=0}^{\infty} \varphi_j B^j,$$

kur $\varphi_0 = 1$ un $j \geq 1$

$$\varphi_j = \frac{\Gamma(j-d)}{\Gamma(j+1)\Gamma(-d)} = \prod_{0 \leq h \leq j} \frac{h-1-d}{h},$$

kur

$$\Gamma(x) = \begin{cases} \int_0^\infty t^{x-1}e^{-t}dt, & ja \quad x > 0; \\ \infty, & ja \quad x = 0; \\ x^{-1}\Gamma(1+x), & ja \quad -1 < x < 0. \end{cases}$$

Definīcija 23. Stacionāru procesu $(X_t, t \in \mathbb{Z})$ sauc par ARIMA(p,d,q) procesu ar $d \in (-0.5, 0.5)$ un $d \neq 0$, ja

$$\bigtriangledown^d X_t = \varepsilon_t,$$

kur ε neatkarīgs vienādi sadalīts gadījuma lielums.

Definīcija 24. Stacionāru procesu $(X_t, t \in \mathbb{Z})$ sauc par FARIMA(p,d,q) procesu ar $d \in (-0.5, 0.5)$ un $d \neq 0$, ja

$$\bigtriangledown^d X_t \sim ARMA(p, q),$$

kas nozīmē, ka

$$b(B) \bigtriangledown^d X_t = a(B)\varepsilon_t,$$

kur ε_t neatkarīgs vienādi sadalīts gadījuma lielums,

$$b(z) = 1 - b_1 z - ... - b_p z^p \text{ un } a(z) = 1 - a_1 z - ... - a_q z^q.$$

4. Jaukto procesu pielietojumi statistikā

Šajā nodaļā tiks apskatīts procesa blīvuma un regresijas funkcijas novērtēšana ar kodolu gludināšanas metodi (skat. [3],[14] un [4]), kā arī tiks apskatīti Rozenblata - Bikela un Neimaņa testi, kas tiek izmantots hipotēžu par jaukto procesu sadalījumu pārbaudei. Apskatot simulēto jaudu, Neimaņa tests tiks salīdzināts ar Kolmogorova - Smirnova testu.

4.1. Procesa blīvuma funkcijas novērtēšana ar kodolu gludināšanu

Pieņemsim, ka dati X_1,...,X_n ir gadījumu lielumu virkne ar blīvuma funkciju f.

Visvienkāršākais blīvuma funcijas novērtējums ir histogramma. Pieņemsim, ka funkcija f definēta intervālā $[0,1]$. Sadalot intervālu $[0,1]$ m vienādās daļās, definēsim

1. sadalījuma intervālus - Binus

$$B_1 = \left[0, \tfrac{1}{m}\right),\ B_2 = \left[\tfrac{1}{m}, \tfrac{2}{m}\right),\ ...,\ B_m = \left[\tfrac{m-1}{m}, 1\right]$$

$\hat{p}_j = \frac{Y_j}{n}$, kur Y_j ir novērtojumu skaits intervālā (binā) B_j, un $p_j = \int_{B_j} f(u)du$;

2. soli h - joslas platumu

$$h = \frac{1}{m}.$$

Tad **histogrammas novērtējums** tiks definēts ar funkciju

$$\hat{f}_n(x) = \sum_{i=1}^{n} \frac{\hat{p}_j}{h} I_{\{x \in B_j\}},$$

kur $I_{\{x \in B_j\}}$ ir indikātorfunkcija.

Definīcija 25. Par kodolu sauc jebkuru gludu funkciju K, tādu, ka

1. $K(x) \geq 0 \ \forall x \in \mathbb{R}$;
2. $\displaystyle\int K(x)dx = 1$ un K - blīvuma funkcija;
3. $\displaystyle\int xK(x)dx = 0$ un K ir simetriska funkcija;
4. $\sigma_k^2 = \displaystyle\int x^2 K(x)dx > 0$.

Definīcija 26. Ja dots kodols K un pozitīvs skaitlis h, ko sauc par joslas platumu, tad procesa $(X_t, t \in \mathbb{Z})$ blīvuma novērtējums ar kodolu gludināšanas metodi ir

$$\hat{f}_n(x) = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{h}K\left(\frac{x - X_i}{h}\right). \tag{4.1.1}$$

Biežāk izmantotie kodoli:

1. Boxcar kodols
$$K(x) = \frac{1}{2}I_{\{|x|\leq 1\}},$$
kur $I_{\{|x|\leq 1\}}$ ir indikātorfunkcija;
2. Gausa kodols
$$K(x) = \frac{1}{\sqrt{2\pi}}\exp^{-\frac{x^2}{2}};$$
3. Jepaņicņekova kodols
$$K(x) = \frac{3}{4}\left(1 - x\right)^2 I_{\{|x|\leq 1\}}.$$

Skaidrs, ka soļa h izvēle ir ļoti nozīmīga, lai novērtētu blīvuma funkciju f ar kodolu gludināšanas metodi. Tipiski nosacījumi neparametriskām metodēm ir $h \to +0$ un $nh \to +\infty$, kad $n \to +\infty$. Visbiežāk izmantotā metode, lai noteiktu joslas platumu h, ir krosvalidācijas metode.

Definīcija 27. Krosvalidācijas metode ir tāda, ka nepieciešams minimizēt krosvalidācijas funkciju

$$J(h) = \int \left(\hat{f}_n(x)\right)^2 dx - \frac{2}{n}\sum_{i=1}^{n}\hat{f}_{-i}(X_i),$$

kur $\hat{f}_{-i}(X_i)$ ir blīvuma funkcijas novērtējums, izlaižot X_i novērojumu.

Specifiskā gadījumā krosvalidācijas funkcija $\forall h > 0$ tiek definēta formā

$$J(h) = \frac{1}{hn^2} \sum_i \sum_j K^* \left(\frac{X_i - X_j}{h} \right) + \frac{2}{nh} K(0) + O\left(\frac{1}{n^2}\right),$$

kur

$$K^*(x) = K^{(2)}(x) - 2K(x)$$

un

$$K^{(2)}(z) = \int K(z-y)K(y)dy$$

Par optimālo h tiek uzskatīts h^*, kas ir šīs minimizējošās krosvalidācijas funkcijas atrisinājums.

Apskatīsimies pie kādiem nosacījumiem jauktajiem procesiem var novērtēt blīvuma funkciju ar kodolu gludināšanas metodi.

Teorēma 22. *Pieņemsim, ka $(X_t, t \in \mathbb{Z})$ ir stacionārs jauktais process ar $\alpha(k) \leq c_0 \rho^k$, kur $c_0 > 0$, $k \geq 1$ un $\rho \in [0,1]$, f ir procesa $(X_t, t \in \mathbb{Z})$ blīvuma funkcija. Ja f ir nepārtraukts punktā x un*

1. *ja*

$$\frac{nh}{(ln^2 n)} \to +\infty,$$

tad

$$f_n(x) \to_{g.d} f(x),$$

2. *ja $f \in C_{2,d}(b)$ kādam b un ja*

$$h = c_n \left(\frac{ln(n)}{n} \right)^{\frac{1}{d+4}},$$

kur $c_n \to c$ $(c > 0)$, tad $\forall x \in \mathbb{R}^d$ un $\forall k$, kur $k \in \mathbb{Z}$

$$\frac{1}{log_k n} \left(\frac{n}{ln(n)} \right)^{\frac{2}{d+4}} (f_n(x) - f(x)) \to_{g.d} 0,$$

Pierādījums. Skatīt [[4], 45.lpp]. Teorēma parāda, ka novērtētā blīvuma funkcija tiecas uz īsto blīvuma funkciju pie lieliem izlašu apjomiem.

4.2. Regresijas novērtēšana ar kodolu gludināšanu

Pieņemsim, ka ir dots stacionārs process $Z_t = (X_t, Y_t), t \in \mathbb{Z}$. Apskatīsim regresijas vienādojumu

$$Y_t = r(X_t) + \varepsilon_t$$

ar $E(\varepsilon_t) = 0$, kur r - regresijas funkcija, Y - rezultējošais jeb atbildes mainīgais, X - skaidrojošais mainīgais, prediktors un ε_t neatkarīgs vienādi sadalīts gadījuma lielums.

Neparametriskajā regresijā, kodoli tiek lietoti, lai noteiktu lokālo vidējo vērtību.

Definīcija 28. Pieņemsim, ka $h > 0$ ir pozitīvs skaitlis, kuru sauc par soli.
Regresijas novērtējums ar kodolu gludināšanu ir formā:

$$\hat{r}_n(x) = \sum_{i=1}^{n} l_i(x) Y_i,$$

kur K ir kodols un svari

$$l_i(x) = \begin{cases} \dfrac{K\left(\dfrac{x-X_i}{h}\right)}{\sum\limits_{j=1}^{n} K\left(\dfrac{x-X_j}{h}\right)} & ja \quad \sum\limits_{j=1}^{n} K\left(\dfrac{x-X_j}{h}\right) \neq 0 \\ \dfrac{1}{n} & ja \quad \sum\limits_{j=1}^{n} K\left(\dfrac{x-X_j}{h}\right) = 0 \end{cases}. \tag{4.2.1}$$

Bieži h izvēlei izmanto krosvalidācijas metodi.

Definīcija 29. Krosvalidācijas metode ir tāda, ka nepieciešams minimizēt novērtēto risku

$$\hat{R}(h) = \sum_{i=1}^{n} (Y_i - \hat{r}_{-i}(X_i))^2,$$

kur $\hat{r}_{-i}(x)$ ir novērtējums, kas iegūstams izlaižot pāri (X_i, Y_i)

$$\hat{r}_{-i}(x) = \sum_{i=1}^{n} Y_l l_{j,(-i)}(x), \text{ kur } l_{j,(-i)}(x) = \begin{cases} 0, & ja \quad j = 1 \\ \dfrac{l_j(x)}{\sum\limits_{k \neq i} l_k(x)}, & ja \quad j \neq i \end{cases}$$

Piezīme 23. $E\left(Y_i - \hat{r}_{-i}\left(X_i\right)\right)^2 \approx \sigma^2 + \underbrace{E\left(r\left(X_i\right) - \hat{r}_{-i}\left(X_i\right)\right)}_{\approx R}$ *un* $E\left(Y_i - \hat{r}_{-i}\left(X_i\right)\right) = 0$

Tādejādi $E\left(\hat{R}\right) \approx R + \sigma^2$*, līdz ar to iegūstam, ka krosvalidācijas metode dod gandrīz nenovirzītu novērtējumu.*

Definīcija 30. $\hat{R}$ ir nenovirzīts R novērtējums, ja $E\left(\hat{R}\right) = R$.

Apskatīsim pie kādiem nosacījumiem jauktajiem procesiem var novērtēt regresijas funkciju ar kodolu gludināšanas metodi.

Teorēma 24. *Pieņemsim, ka* $Z_t = (X_t, Y_t), t \in \mathbb{Z}$ *ir stacionārs jauktais process ar* $\alpha(k) \leq c_0\rho^k$*, kur* $c_0 > 0$*,* $k \geq 1$ *un* $\rho \in [0, 1]$*,* f *ir procesa* Z_t *blīvuma funkcija,* $f \in C_{2,d}(b)$ *kādam* b *un* $E(exp(a|Y_0|^\tau)) < \infty$ *kādam* $a > 0$ *un kādam* $\tau > 0$*.*

Ja

$$\frac{nh}{(ln(n))^{2+1/\tau}} \to \infty$$

un S *ir kompakta kopa, tāda, ka* $\inf_{x \in S} f(x) > 0$*, tad*

$$\sup_{x \in S} |\hat{r}_n(x) - r(x)| \to 0$$

Turklāt, ja

$$h \simeq \left(\frac{(ln(n))^{2-1/\tau}}{n} \right)^{1/(d/4)},$$

tad

$$\frac{n^{2/(d+4)}}{log_k n(ln(n))^{1/\tau+(1-1/\tau)(2/(d+4))}} \sup_{x \in S} |\hat{r}_n(x) - r(x)| \to 0.$$

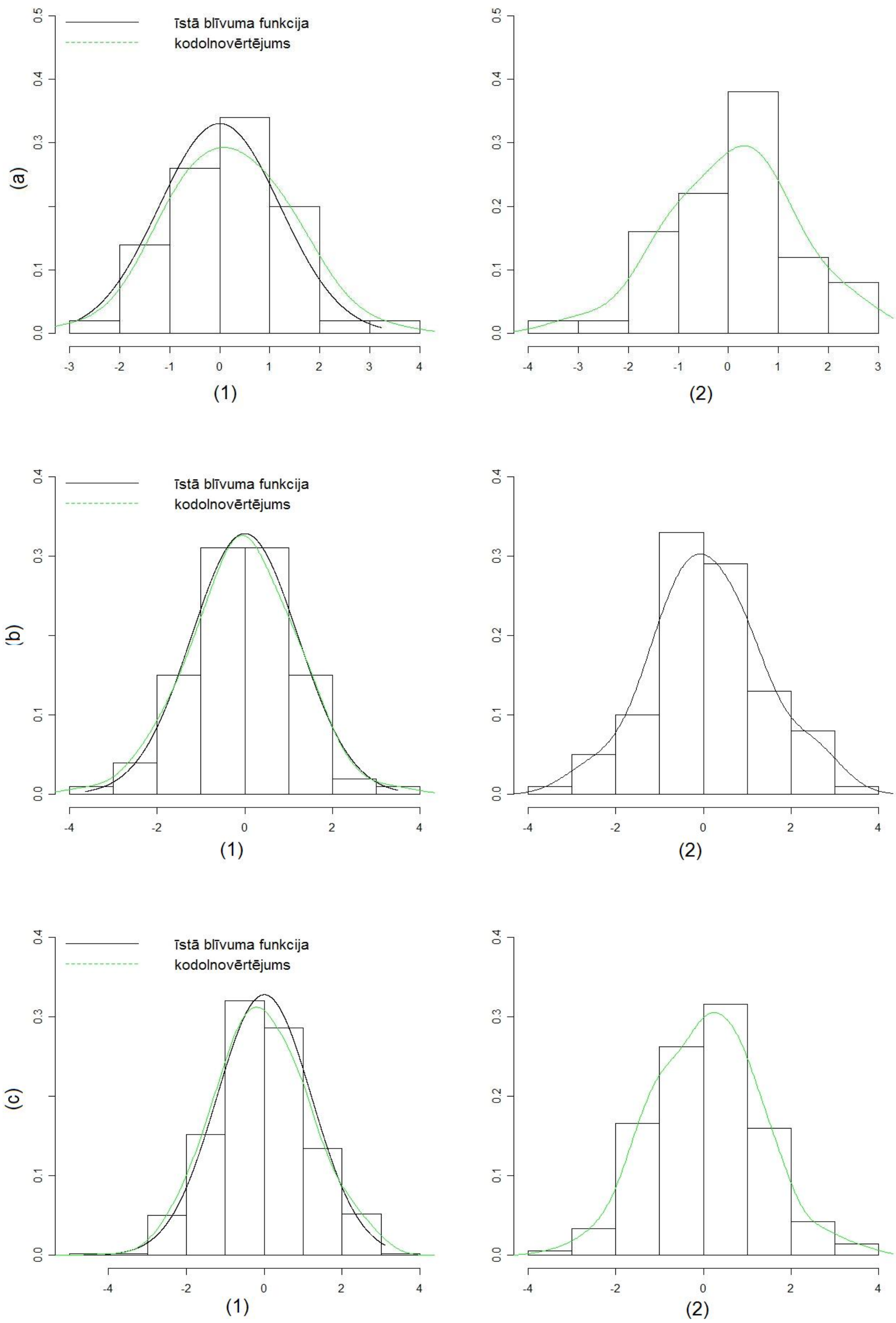
Pierādījums. Skatīt [[4], 71.lpp]. Teorēma parāda, ka novērtētā regresijas funkcija tiecas uz īsto regresijas funkciju katrā datu punktā.

4.3. Piemēri par jaukto procesu blīvuma un regresijas funkcijas novērtēšanu ar kodolu gludināšanu

Piemērs 7. Tika ģenerēta izlase apjomā $n = 50$, 100 un 500 novērojumiem pēc ARMA(1,1) modeļa

$$X_t = aX_{t-1} + \varepsilon_t + b\varepsilon_{t-1},$$

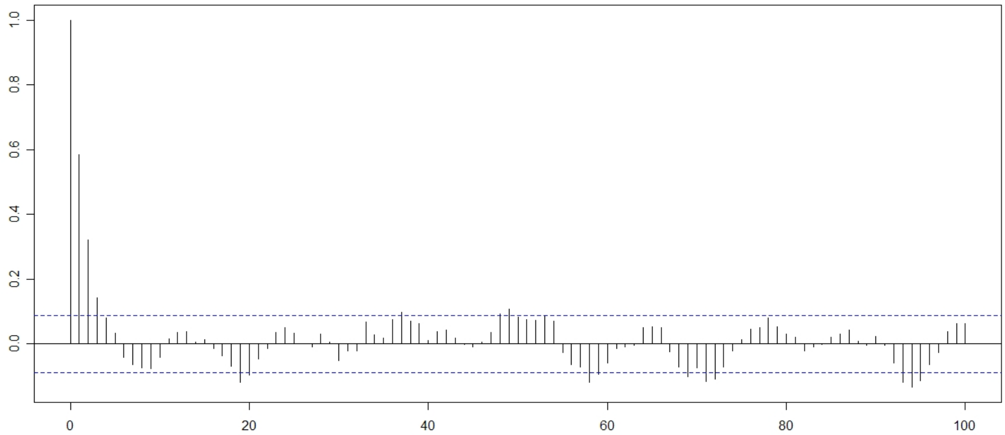
kur $\varepsilon_t \sim N(0, 1)$. Konstruēsim kodola blīvuma funkcijas novērtējumu formā (4.1.1) un salīdzināsim ar gadījumu, kad ģenerētā izlase ir ar vienādi sadalītiem neatkarīgiem novērojumiem un ar normālo sadalījuma funkciju.



4.1. att.: Konstruētais kodola funkcijas novērtējums (1) - izlase ar atkarīgiem novērojumiem, ja $a = 0.5$ un $b = 0.1$; (2) - izlase ar vienādi sadalītiem neatkarīgiem novērojumiem; (a) - $n = 50$; (b) - $n = 100$; (c) - $n = 500$.

Apskatot 4.1.attēlu var redzēt, ka izlases blīvuma funkcijas novērtējums ar kodolu gludināšanas metodi darbojas līdzīgi gan izlasei ar atkarīgiem novērojumiem gan izlasēm ar vienādi sadalītiem novērojumiem.

Tākā koeficienti a, b ir mazāki par viens un autokorelācijas funkcija strauji dilst (skat. 4.2.attēlu), process ir jauktais process, kuram var novērtēt blīvuma funkciju.



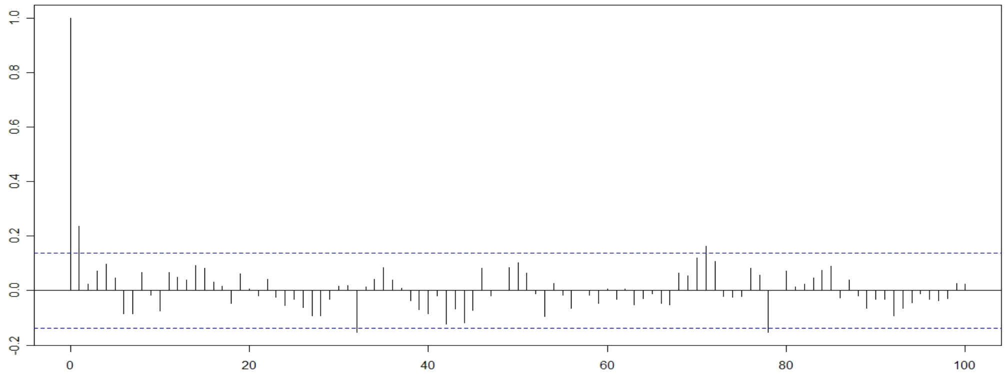
4.2. att. Autokorelācijas funkcija procesam $(X_t, t \in \mathbb{Z})$, ja $a = 0.5$ un $b = 0.1$.

Piemērs 8. Apskatīsim piemēru ar nelineāru laikrindu formā

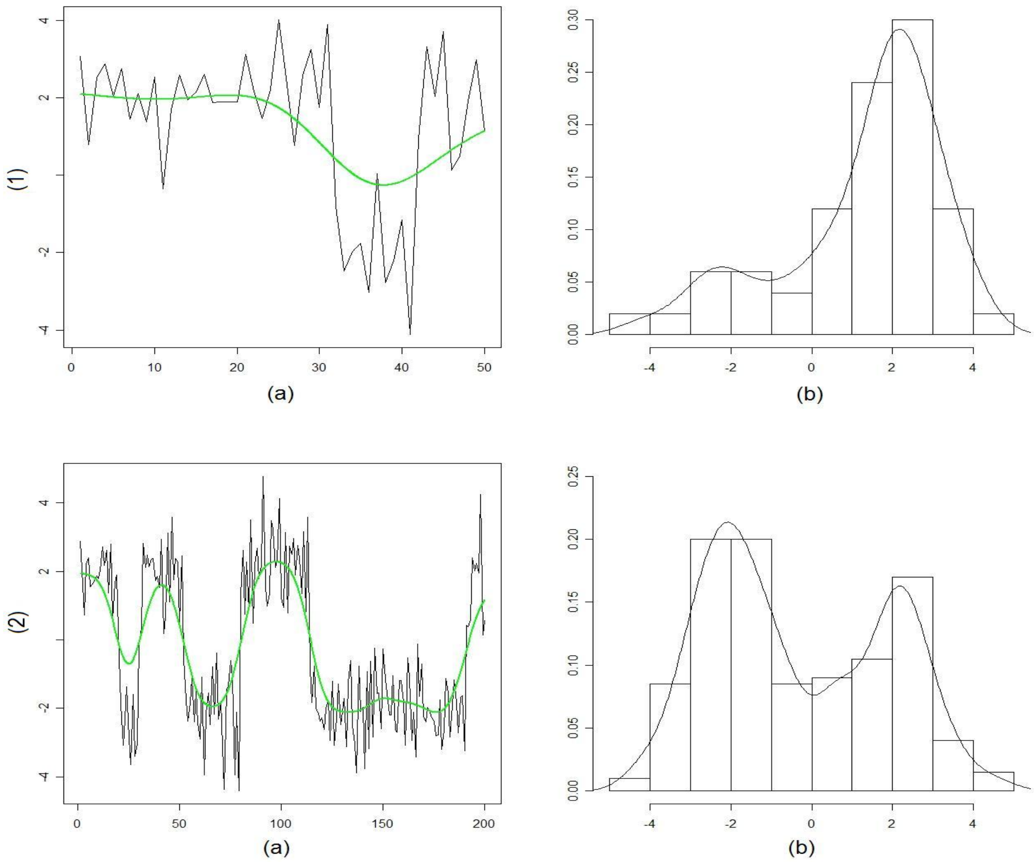
$$X_t = \frac{aX_{t-1}}{1 + bX_{t-1}^2} + \varepsilon_t,$$

Autokorelācijas funkcija strauji dilst (skat. 4.3.attēlu), varētu secināt, ka šis process ir jauktais process. Apskatīsimies, kā darbojas procesa blīvuma funkcijas novērtējums ar kodolu gludinašanas metodi.

25teorēmas nosacījumus pārbaudīt ir ļoti grūti, tādēļ nevar īsti pateikt, vai šajā gadījumā var novērtēt blīvuma funkciju ar kodolu gludināšanas metodi. Apskatot 4.4.attēlu, var redzēt, ka blīvuma un regresijas funkcijas novērtējums ar kodolu gludināšanas metodi darbojas labi.

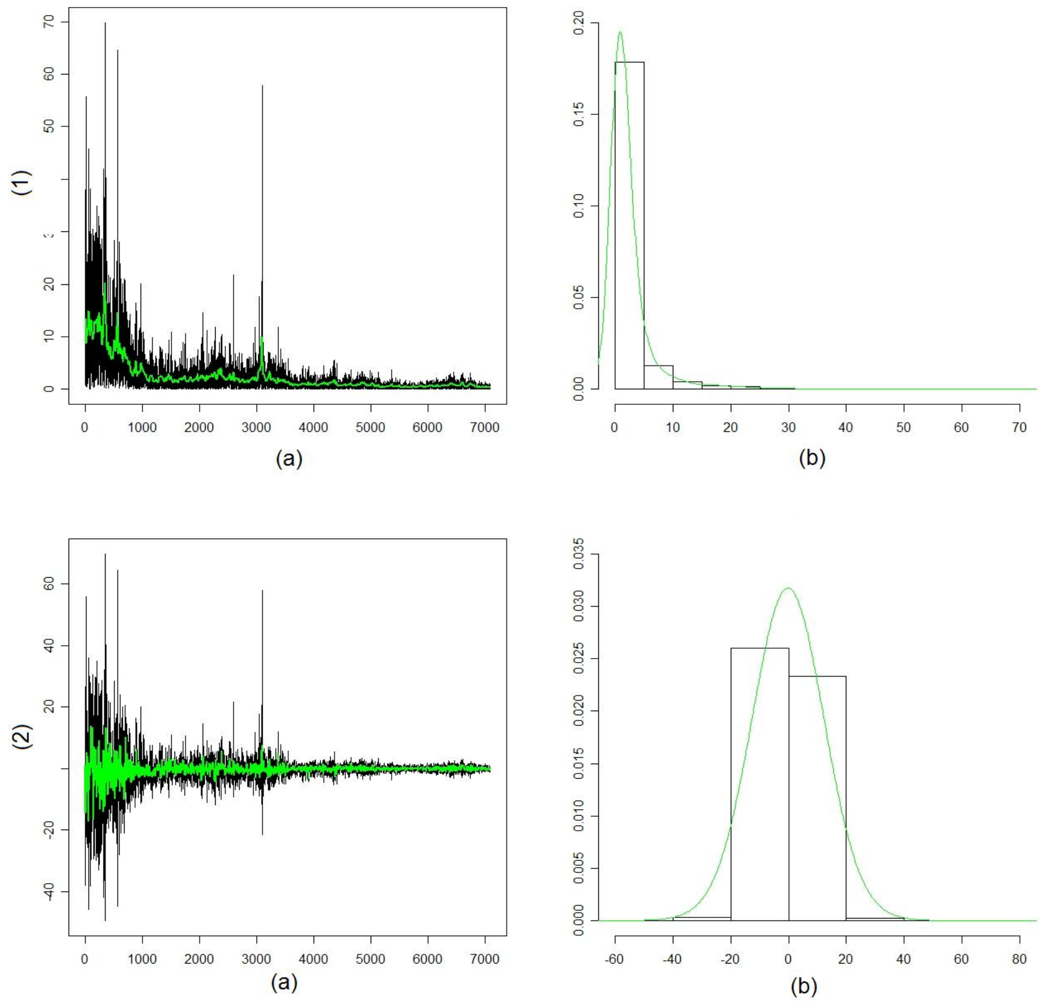


4.3. att. Autokorelācijas funkcija procesam $(X_t, t \in \mathbb{Z})$, kad $a = 5$ un $b = 0.9$.



4.4. att.: (a) Nelineāra laikrinda $(X_t, \in \mathbb{Z})$ ar konstruētu regresijas funkcijas novērtējumu ar kodolu gludināšanu ($a = 5$,$b = 0.9$ un $\varepsilon_t \sim N(0,1)$); (b) Nelineāra laikrinda $(X_t, t \in \mathbb{Z})$ histogramma ar konstruētu blīvuma funkcijas novērtējumu ar kodolu gludināšanas metodi($a = 5$, $b = 0.9$ un $\varepsilon_t \sim N(0,1)$); (1)$n = 50$; (2)$n = 200$.

Piemērs 9 (Praktisks piemērs). 6.piemērā apskatītajam ilglaicīgās atmiņas procesam kostruēsim procesa $Z_t = Y_t - Y_{t-1}$, kas ir vājās atkarības process, un $R_t = |Y_t - Y_{t-1}|$, kas ilglaicīgās atmiņas process, blīvuma un regresijas funkcijas novērtējumu ar kodolu gludināšanu.

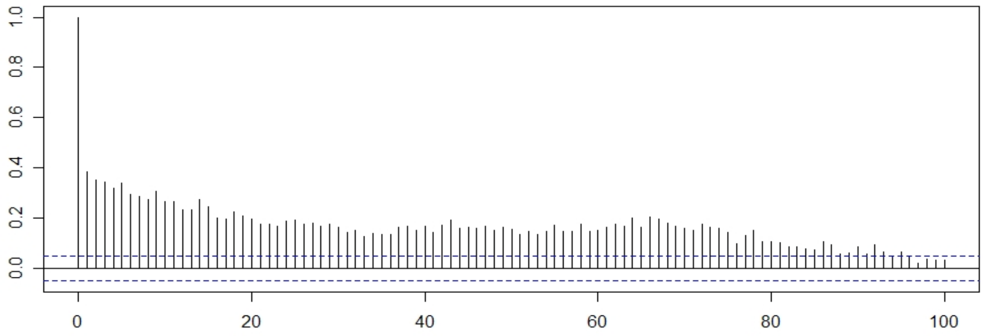


4.5. att.: (1) Procesa $(R_t, t \in \mathbb{Z})$ (2) Procesa $(Z_t, t \in \mathbb{Z})$ a) regresijas funkcijas novērtējums ar kodolu gludināšanu ($h = 7$), b) blīvuma funkcijas novērtējums ar kodolu gludināšanu ($h = 1.5$).

Konstruējot procesa $(R_t, t \in \mathbb{Z})$ blīvuma un regresijas funkcijas novērtējumu ar kodolu gludināšanu, nedarbojās krosvalidācijas metode soļa h izvēlei, tādēļ h tika izvēlēts tāds, lai blīvuma un regresijas funkcijas novērtējums ar kodolu gludināšanu būtu pēc iespējas labāks (skat. 4.5.attēlu). Procesam $(Z_t, t \in \mathbb{Z})$ novērtējot blīvuma un regresijas funkcijas novērtējumu ar kodolu gludināšanu h tika izvēlēts pēc krosvalidācijas metodes. Autokorelācijas funkcija procesam $(Z_t, t \in \mathbb{Z})$ dilsts straujāk nekā procesam $(R_t, t \in \mathbb{Z})$ (skat.

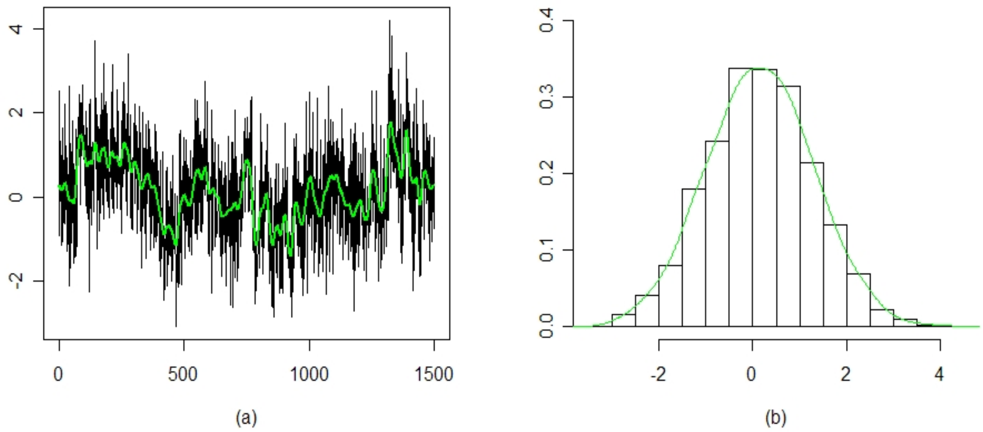
3.1.attēlu).

Piemērs 10. Ģenerēsim FARIMA(1,0.4,1) procesu apjomā 1500. Tākā autokorelācijas funkcija dilst lēni, tad tas ir ilglaicīgās atmiņas process (skat. 4.6.attēlu).



4.6. att. FARIMA(1,0.4,1) procesa autokorelācijas funkcija

Arī šeit konstruējot procesa blīvuma un regresijas funkcijas novērtējumu nedarbojās krosvalidācijas metode, tādēļ arī šeit h tika brīvi izvēlēts (skat. 4.7.attēlu).



4.7. att.: FARIMA(1,0.4,1) procesa konstruētais a) regresijas funkcijas novērtēums ar kodolu gludināšanu (h = 7), b) blīvuma funkcijas novērtējums ar kodolu gludināšanu ($h = 0.3$).

4.4. Sadalījuma noteikšanas testi atkarīgiem datiem

Šajā nodaļā teorētiski tiks apskatīti Bikela Rozenblata un Neimaņa testi, ar kuru var pārbaudīt hipotēzes par jaukto procesu sadalījumu. Sākotnēji šos testus izmantoja, lai pārbaudītu hipotēzes par neatkarīgu vienādi sadalītu gadījuma lielumu sadalījumu. Neimaņa tests tiks salīdzināts ar Kolmogorova - Smirnova testu, apskatot empīrisko jaudu.

4.4.1. Bikela Rozenblata tests

Pieņemsim, ka ir dots stacionārs process $(X_t, t \in \mathbb{Z})$. Pārbaudīsim vienkāršo hipotēzi

$$H_0 : f = f_0,$$

$$H_1 : f \neq f_0,$$

kur f ir procesa $(X_t, t \in \mathbb{Z})$ blīvuma funkcija.

Pieņemsim, ka

$$\hat{f}_n(x) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{x - X_i}{h}\right),$$

kur $\hat{f}_n(x)$ ir blīvuma funkcijas f kodolu novērtējums, $h = h(n)$, $h \to 0$, kad $n \to \infty$. Salīdzināsim $\hat{f}_n$ ar nogludnāto f_0 ar statistiku

$$T_n = nh^{d/2} \int \left(\hat{f}_n(x) - (K_h * f_0)(x)\right)^2 dx,$$

kur

$$(K_h * f_0)(x) = \frac{1}{h} \int K\left(\frac{x - z}{h}\right) f_0(z) dz \tag{4.4.1}$$

un d ir procesa $(X_t, t \in \mathbb{Z})$ dimensija.

Pirms definēsim teorēmu par statistikas T_n robežsadalījumu, kuru 2000.gadā savā darbā parādīja Neimanis un Paparoditis, definēsim pieņēmumus.

Pieņēmums 1. Absolūti regulārais $\beta(n)$ jaukto procesu koeficients ir definēts formā (1.1.6) un tas dilst eksponenciāli ātri, t.i.,

$$\beta(n) \leq Ce^{-Ck},$$

kur C ir konstante.

Pieņemsim, ka f ir stacionāra procesa $(X_t, t \in \mathbb{Z})$ blīvuma funkcija un $f_{X_{i_1},\ldots,X_{i_m}}$ kopējā blīvuma funkcija.

Pieņēmums 2. Blīvuma funkcija f ir nepārtraukta un

$$\sup_{x_1,...,x_m} \{f_{X_{i_1},...,X_{i_m}}(x_1,...,x_m)\} < \infty$$

$\forall m$, kur $i_1 < i_2 < ... < i_m$

Pieņēmums 3. $h = o\left([ln(n)]^{-3}\right)$ un $h^{-d} = o(n)$, kur d ir procesa $(X_t, t \in \mathbb{Z})$ dimensija.

Teorēma 25. *Pieņemsim, ka (1)-(3) pieņēmumi izpildās, tad, ja hipotēze H_0 ir spēkā,*

$$T_n - h^{\frac{d}{2}} \int K^2(u)du \to_d N(0, \sigma^2),$$

kur

$$\sigma^2 = 2 \int f^2(x)dx \cdot \int \left(\int K(u)K(u+v)du\right)^2 dv$$

Pierādījums. Skatīt [[15], 144.lpp]. Līdzīgus rezultātus 1987.gadā parādīja arī Takahata un Joshihara.

Apskatīsim salikto hipotēzi, tad

$$H_0 : f \in \mathcal{F},$$

$$H_1 : f \text{ nepieder } \mathcal{F},$$

kur $\mathcal{F}$ ir blīvuma funkciju klase. Šo gadījumu ir iespējams reducēt uz vienkāršās hipotēzes gadījumu. Praksē tiek pārbaudītas parametriskas hipotēzes, tad $\mathcal{F} = \mathcal{F}^\theta = \{f_\theta, \theta \in \Theta\}$, kur $\Theta \subseteq \mathbb{R}^p$. Iepriekš, kad apskatījām vienkāršo hipotēzi, tad $\hat{f}_n$ salīdzinājām ar f_0, tad tagad $\hat{f}_n$ salīdzina ar $f_{\hat{\theta}}$, kur $f_{\hat{\theta}}$ ir blīvuma funkcijas f parametriski novērtētā funkcija ar statistiku

$$T_{n,\hat{\theta}} = nh^{\frac{d}{2}} \int \left(\hat{f}_n(x) - (K_h * f_{\hat{\theta}})(x)\right),$$

kur $(K_h * f_{\hat{\theta}})(x)$ ir formā (4.4.1).

Pieņēmums 4.

1. $\int \left((K_h * f_{\hat{\theta}})(x) - (K_h * f_{\theta_0})(x)\right)^2 dx = o_p(n^{-1}h^{-\frac{d}{2}})$;
2. $\hat{\theta} - \theta_0 = o_p(n^{-\frac{1}{2}}h^{-\frac{d}{2}})$;
3. $\sup_x\{|f'_{\theta_0}|\} < \infty$.

Teorēma 26. *Pieņemsim, ka izpildās (1)-(4) pieņēmumi un hipotēze H_0 ir spēkā, tad*

$$T_{n,\hat{\theta}} - h^{-\frac{d}{2}} \int K^2(u)du \to_d N(0, \sigma^2).$$

Pierādījums ir līdzīgs 25.teorēmas pierādījumam.

4.4.2. Neimaņa tests

Pieņemsim, ka ir dots stacionārs α jauktais process $(X_t, t \in \mathbb{Z})$. Pārbaudīsim vienkāršo hipotēzi

$$H_0 : F = F_0,$$

$$H_1 : F \neq F_0,$$

kur F ir procesa $(X_t, t \in \mathbb{Z})$ sadalījuma funkcija.

Hipotēzes pārbaudei Valeinis piedāvāja izmantot sekojošu statistiku

$$T_k = \frac{1}{12\sigma^2} \sum_{j=1}^{k} \left(n^{1/2} \sum_{i=1}^{n} \phi_j(X_i) \right)^2,$$

kur $k = 1, 2, ..., d(n)$, $\sigma^2 = \sum_{t=-\infty}^{\infty} Cov(X_0, X_t)$, ϕ_j, ir ortonormāli Lažandra polinomi. Pirmie Lažandra polinomi ir sekojoši:

$$\phi_0(x) = 1;\ \phi_1(x) = \sqrt{3}(2x - 1);$$
$$\phi_2(x) = \sqrt{5}\,(6x^2 - 6x + 1);$$
$$\phi_3(x) = \sqrt{7}\,(20x^3 - 30x^2 + 12x - 1);$$
$$\phi_4(x) = 3\,(70x^4 - 140x^3 + 90x^2 - 20x + 1).$$

Piezīme 27. *Ja $1/(12\sigma^2) = 1$, tad iegūstam statistiku, ko izmanto hipotēzes par neatkarīgu vienādi sadalītu procesu sadalījumu pārbaudei, ko ieviesa 1937.gadā Neimanis (skat. [16]).*

Hipotēze H_0 par noteiktu sadalījuma likumu tiek noraidīta izvēlētajā nozīmības līmenī α_0 ja testa statistika, T_k, ir lielāka par tabulēto kritisko vērtību.

Nozīmīga ir komponenšu skaita k izvēle. Sākotnēji, pirms Ledvina ieviesa 1994.gadā ieviesa švarca izvēles likumu (skat. [17]), literatūrā rekomendēja izvēlēties $k = 2, 3$. Komponenšu skaita k izvēle ietekmē jaudu. Tādēļ apskatīsim Ledvina Švarca izvēles likumu.

Pieņemsim, ka

$$Y_n = (\bar{\phi}_1, ..., \bar{\phi}_n),$$

kur

$$\bar{\phi}_j = n^{-1} \textstyle\sum_{i=1}^{n} \phi_j(X_i),\ (j = 1, 2, ..., k),$$

tad Švarca izvēles likums ir

$$S \quad = \quad \min\left\{k : 1 \leq k \leq d(n), n|Y_n|_k^2 - kln(n) \geq n|Y_n|_j^2 - jln(n), j = 1, 2, ..., d(n)\right\}.$$

Pirms definēsim teorēmu par statistikas T_k robežsadalījumu ir nepieciešami četri pieņēmumi.

Pieņēmums 5. $\alpha(k) \leq a\rho^k$ kādām $a > 0$, $0 < \rho < 1$.

Pieņēmums 6. $E|X_t|^\gamma < \infty$ kādiem $\gamma > 2$.

Pieņēmums 7. $\sigma^2 = \sum_{t=-\infty}^{\infty} Cov(X_0, X_t) > 0$.

Pieņēmums 8. $d(n) = o\left(\frac{ln(n)}{lnln(n)}\right)$.

Teorēma 28. *Pieņemsim, ka $(X_t, t \in \mathbb{Z})$ ir stacionārs α jauktais process, izpildās 5. un 8.pieņēmums.*

a) *Tad*

$$\lim_{n\to\infty} P(S = 1) = 1.$$

b) *Ja hipotēze H_0 ir spēkā un izpildās 6.,7.pieņēmums, tad*

$$N_k \to_d \chi_1^2.$$

Pierādījums. Skat. [18].

Piezīme 29. *Ja hipotēze H_1 ir spēkā, tad $N_k \to \infty$.*

Sekas 30. *Pieņemsim, ka izpildās 5.-8.pieņēmumi un $\hat{\sigma}^2$ ir σ^2 novērtējums, tad*

$$\frac{1}{12\hat{\sigma}^2} \sum_{j=1}^{k} \left(n^{1/2} \sum_{i=1}^{n} \phi_j(X_i) \right)^2 \to_d \chi_1^2.$$

Lai novērtētu stacionāra procesa $(X_t, t \in \mathbb{Z})$ σ^2, pieņemsim, ka $\forall t, h$ $(t \in \mathbb{Z}, h \in \mathbb{Z})$ $\gamma(h) = Cov(X_t, X_{t+h})$. Ja novērtēsim kovariācijas funkciju $\gamma(h)$ ar $\hat{\gamma}(h)$, tad

$$\hat{\gamma}(h) = (n-h)^{-1} \sum_{t-1}^{n-h} (X_t - \bar{X})(X_{t+h} - \bar{X}),$$

kur $0 \leq h \leq n-1$. Tagad novērtēsim σ^2 ar

$$\hat{\sigma}^2 = \hat{\gamma}(0) + 2\sum_{j=1}^{q} \hat{\gamma}(j),$$

kur q tiek izvēlēts diezgan mazs, noapaļojot autokovariācijas ar trim komatiem aiz zīmes, jo procesa autokovariācijas funkcija dilst eksponenciāli ātri.

Piezīme 31. *Jauktajiem procesiem Neimaņa tests saliktajām hipotēzēm nav vēl pierādīts.*

Apskatot empīriski simulētās jaudas, salīdzināsim Neimaņa testu ar Kolmogorova - Smirnova testu, kas balstās uz statistiku

$$D_n = \sup_x |F(x) - F_n(x)|,$$

kur $F_n(x)$ apzīmē empīrisko sadalījuma funkciju.

Piemērs 11. Tika ģenerēta gadījuma izlases apjomā $n = 20$, 50 un 100 pēc AR(1) modeļa formā

$$X_t = 0.5X_{t-1} + \varepsilon_t$$

un pēc GARCH(1,1) modeļa formā

$$X_t = \sigma_t \varepsilon_t,$$

kur $\sigma_t = 0.9 + 0.2X_{t-1} + 0.6\sigma_{t-1}$, pie nozīmības līmeņa $\alpha = 5\%$ empīriski tika simulēta jauda:

1. Neimaņa testam;

2. Kolmogorova-Smirnova testam.

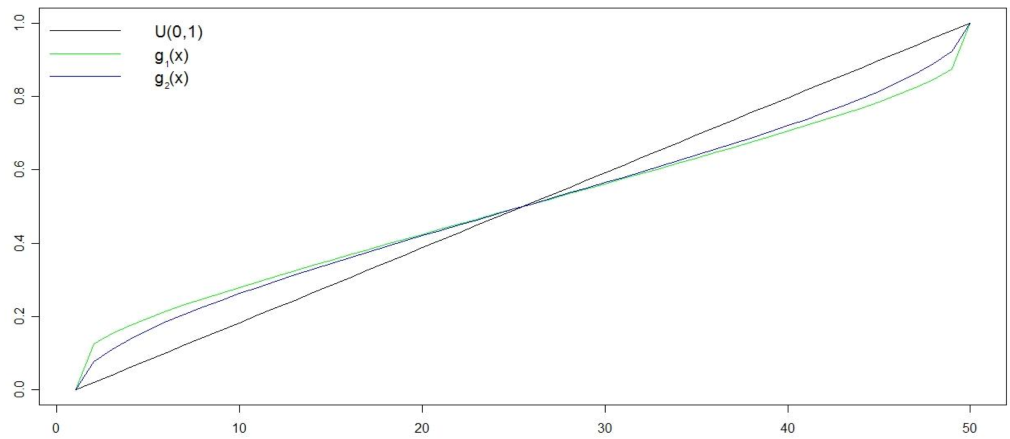
Apskatīsim divas alternatīvas (tika izmantots [19].literatūras avots)

$$\begin{aligned} g_1(x) &= \frac{1}{2} + \frac{1}{20}\left(x^{-4/5} + (1-x)^{-4/5}\right); \\ g_2(x) &= \frac{1}{4}\left(x^{-1/2} + (1-x)^{-1/2}\right), \end{aligned}$$

jo šo alternatīvu blīvuma funkcijas tiecas uz vienmērīgo blīvuma funkciju segmentā $[0, 1]$ (skat 4.8.attēlu). Pāriesim uz jauniem mainīgajiem Z_t, izmantojot datus un alternatīvas, un izvirzīsim hipotēzi

$$\begin{aligned} H_0 &: Z_t \sim U(0, 1); \\ H_1 &: Z_t \nsim U(0, 1), \end{aligned}$$

Veicot 10000 reižu simulācijas, aprēķināsim empīrisko jaudu datiem dažādiem izlašu apjomiem, kas ģenerētas no alternatīvām $g_1(x)$ un $g_2(x)$. Empīriski simulētā Neimaņa testam ir lielāka nekā Kolmogorova-Smirnova testam (skat. 4.1.tabulu).



4.8. att. Vienmērīgā blīvuma funkcija segmentā [0,1] ar alternatīvām $g_1(x)$ un $g_2(x)$.

4.1. tabula: Pie nozīmības līmeņa $\alpha = 5\%$ 10000 reižu simulētā jauda jauktajiem procesam

	$n = 20$	$n = 50$	$n = 100$
ar alternatīvu $g_1(x)$ AR(1) procesam			
Neimaņa tests	0.553	0.819	0.97
Kolmogorova-Smirnova tests	0.294	0.468	0.706
ar alternatīvu $g_2(x)$ AR(1) procesam			
Neimaņa tests	0.326	0.523	0.783
Kolmogorova-Smirnova tests	0.261	0.397	0.604
ar alternatīvu $g_1(x)$ GARCH(1,1) procesam			
Neimaņa tests	0.995	1	1
Kolmogorova-Smirnova tests	0.616	0.955	0.997
ar alternatīvu $g_2(x)$ GARCH(1,1) procesam			
Neimaņa tests	0.982	1	1
Kolmogorova-Smirnova tests	0.549	0.918	0.998

Tagad Apskatīsimies gadījumu, kad dati $(X_t, t \in \mathbb{Z})$ ir neatkarīgi vienādi sadalīti gadījuma lielumi.

Veicot 10000 reižu simulācijas, aprēķināsim empīriski simulēto jaudu, ģenerējot datus no alternatīvas $g_1(x)$.

4.2. tabula: Pie nozīmības līmeņa $\alpha = 5\%$ 10000 reižu simulētā jauda neatkarīgiem vienādi sadalītiem gadījuma lielumiem

	$n = 20$	$n = 50$	$n = 100$	$n = 500$
$X_t \sim U(0, 1)$				
Neimaņa tests	0.762	0.953	0.998	1
Kolmogorova-Smirnova tests	0.214	0.466	0.8	0.997
$X_t \sim N(0, 1)$				
Neimaņa tests	0.773	0.954	0.995	0.998
Kolmogorova-Smirnova tests	0.215	0.471	0.799	0.996

Pie maziem izlašu apjomiem Neimaņa tests darbojas labāk nekā Kolmogorova-Smirnova tests, jo simulētā jauda ir lielāka (skat. 4.2.tabulu).

Kolmogorova-Smirnova tests nav pamatots jauktajiem procesiem. Apskatītajām alternatīvām Neimaņa tests darbojas labāk vienkāršo hipotēžu gadījumā.

Nobeigums

Maģistra darbā tika aplūkoti jauktie procesi, kas tiek definēti ar jaukto procesu koeficientiem, un apskatīti to pielietojumi statistikā. Jaukto procesu koeficientus vispārīgā veidā ir ļoti grūti aprēķināt, vieglāk tos ir aprēķināt tad, kad divdimensionāla gadījuma lielums pieņem tikai divas vērtības. Viens no mērķiem bija analizēt, kā aprēķināt jaukto procesu koeficientus praksē novērotiem datiem. Tomēr, nepieņemot kādu modeli vai struktūru par datiem, jaukto procesu koeficientus, manuprāt, noteikt praktiski ir neiespējami.

Praktiski apskatot dažādu procesu blīvuma un regresijas funkcijas novērtējumus ar kodolu gludināšanu, var secināt, ka šīs metodes darbojas līdzīgi kā neatkarīgu datu gadījumā, kas tika apskatīti neparametrikās statistikas lekciju kursā. Nosacījumus, uz kuriem balstās neparemetriskās statistikas metodes, arī nav viegli pārbaudīt

Kaut gan blīvuma un regresijas funkcijas novērtējumi ar kodolu gludināšanu teorētiski strādā jauktajiem procesiem, pagaidām literatūrā vēl īsti nav pamatota šī metode ilglaicīgās atmiņas procesiem. Praktisko datu analizē Ilglaicīgās atmiņas procesiem nedarbojas krovalidācijas metode joslas platuma h izvēlei.

Neimaņa tests darbojas labāk par Kolmogorova - Smirnova testu gan neatkarīgiem datiem, gan jauktajiem procesies, kaut arī Kolmogorova - Smirnova tests jauktajiem procesiem nav pamatots. Saliktām hipotēzēm Neimaņa tests vēl nav pierādīts jauktajiem procesiem. Būtu nepieciešams veikt smalkāku Bikela - Rozenblata testa analīzi jauktajiem procesiem, jo tas pagaidām ir vienīgais tests jaukto procesu saliktām hipotēzēm par sadalījumu.

Izmantotā literatūra un avoti

[1] R.C. Bradley. *Introduction to Strong Mixing Conditions, Volume 1.* Kendrick Press: Department of Mathematics, Indiana University, Heber City, USA, 2007.

[2] P. Billingsley. *Probability And Measure, third edition.* Wiley, New York, USA, 1995.

[3] Q. Fan, J. Yao. *Nonlinear Time Series Nonparametric And Parametric Methods.* Springer-Verlag, New York, USA, 2003.

[4] D. Bosq. *Nonparametric Statistcs for Stohastic Processes. Estimation and Prediction, Second edtion.* Springer, New York, USA, 1998.

[5] D.S. Shumway, R.H. Stoffer. *Time Series Analysis And Its Applications With R Examples, second edition.* Springer Science + Business Media, LLC, New York, USA, 2006.

[6] V. Carkova. *Markova ķēdes.* Latvijas Universitāte, Rīga, 2001.

[7] R.C. Bradley. Basic properties of strong mixing conditions. *In: Dependence in Probability and Statistics,(E.Eberlein and M.S.Taqqu,eds.), Progress in Probability and Statistics*, 11, 1986.

[8] L.T. PHAM, T.D. TRAN. Some mixing properties of time series models. *Stohastic Processes and Their Applications*, 19:297–303, 1995.

[9] P. Lang G. Leon R. J.R. Louhicni S. Prieur C. Dedecker, J. Doukhan. *Weak dependence with examples and applications.* Springer Science + Business Media, LLC, New York, USA, 2007.

[10] R.A. Basrak, B. Davis and T. Mikosch. Regular variation of garch processes. *Stohastic Processes and Their Applications*, 99:95–115, 2002.

[11] J.D. Lillo, F. Farmer. The long memory of the efficient market. *Studies in Nonlinear Dynamics and Econometrics*, 8, 2004.

[12] G. Doukhan, P. Oppenhelm and M.S. Taqqu. *Theory and Application of Long Range Dependence*. Springer-Verlag, New York, USA, 2003.

[13] F. Jianqing. mājas lapa $http : //www.orfe.princeton.edu/ \sim jqfan/fan/nls/datasets.html$.

[14] L. Wasserman. *All of Nonparametric Statistics*. Springer Science + Business Media, New York, USA, 2006.

[15] E. Neumann, M.H. Paparoditis. On bootstrapping l_2-type statistics in density testing. *Statistic and Probability Letter*, 50:137–147, 2000.

[16] D.J. Rayner, J.C.W. Best. *Smooth Tests of Goodness of Fit*. Oxford University, New York, USA, 1989.

[17] W.C.N. Inglot, T. Kallenberg and T. Ledvina. Data driven smooth tests for composite hypotheses. *The Annals of Statistics*, 25:1223–1249, 1997.

[18] J. Valeinis. Neyman smooth goodness-of-fit tests for the marginal distribution of dependence data. Iesniegta publicēšanai, 2009.

[19] T. Ledvina. Data - driven version of neyman's smooth test of fit. *American Statistical Association*, 89:1000–1005, 1994.

A Izveidoto programmu kodi

1. programmas kods

```
#### Arma process
n1<-1000
AR<-0.5
MA<-0.1
par(mfrow=c(3,2))
n<-100
k<-n1*n
ts.sim <- arima.sim(list(ar = AR,ma=MA), k)
vid<-mean(ts.sim)
var<-var(ts.sim)
xx<-1:n
dati<-ts.sim[(n1+1-n):n1]
plot(dati, type="l",main="",xlab="(a)",ylab="" )
nad.fun<-function(x)
{
sum(dnorm((x-xx)/h)*dati)/
sum(dnorm((x-xx)/h))
}
d<-seq(min(xx),max(xx),len=1000)
lines(d,sapply(d,nad.fun),lwd=2,col="green")
legend("topright", paste(c("ista blivuma funkcija",
"kodolnovertejums")),
col=c("black","green"), lty=1:2, pch = "", ncol =1,
cex = 1.3,bty="n")
```

```
hist(dati,prob=T,main="",ylim=c(0,0.6),xlab="(1)",
ylab="(b)")
xx<-seq(min(dati),max(dati),by=0.001)
points(xx,dnorm(xx,0,sqrt(var)),type="l")
library(sm)
h1<-hcv(dati)
points(density(dati,bw=h1),type="l",col="green")

### gadijuma lielums ir neatkarigs vienadi sadalits
iid<-rnorm(n,0,sqrt(var))
h4<-hcv(iid)
hist(iid,prob=T,ylim=c(0,0.6),xlab="(2)",ylab="",main="")
points(density(iid,bw=h4),col="green",type="l" )
```

2. programmas kods

```
## nelineara laikrinda
c.t<-c()
n1<-1000
par(mfrow=c(3,2))
eps.t<-rnorm(n1)
x.t<-0
for (j in 2:n1)
{
x.t[j]<-0.1*x.t[j-1]/(1+0.9*(x.t[j-1])^2)+eps.t[j]
}
n<-200
dati<-x.t[(n1+1-n):n1]
plot(dati, type="l",main="",ylab="(1)",xlab="(a)")
acf(dati,lag.max=100,main="",xlab="(2)",ylab="")
hist(dati,prob=T,xlab="",ylab=""
,main="",ylim=c(0,0.3),main="",xlab="(2)",ylab="")
library(sm)
```

```
h1<-hcv(dati)
points(density(dati,bw=h1),type="l",col="green")
```

3. programmas kods

```
#### process  Z_t
par(mfrow=c(1,2))
y.data<-read.table(file="S&P1.txt",dec=",")[,5]
dati<-c()
for (i in 1:n-1)
{
dati[i]<-y.data[i+1]-y.data[i]
}
n<-length(dati)
plot(dati,xlab="",ylab="",type="l",main="")
acf(dati,lag.max=500,main="",xlab="a",ylab="")
hist(dati,prob=T,breaks = 7,ylim = c(0, 0.2),
xlab="",ylab="",main="")
xx<-seq(min(dati1),max(dati1),by=(max(dati1)
-min(dati1))/(n-1))
nad.fun<-function(x)
{
sum(dnorm((x-xx)/h)*dati/
sum(dnorm((x-xx)/h)))
}
library(sm)
h1<-hcv(dati)
points(density(dati,bw=h1),type="l",col="green")

#### long memory process  R_t
par(mfrow=c(1,2))
dati1<-abs(dati)
n<-length(dati1)
plot(dati1,xlab="(a)",ylab="",type="l")
```

```
xx<-1:n
h<-7
nad.fun<-function(x)
{
sum(dnorm((x-xx)/h)*dati1)/
sum(dnorm((x-xx)/h))
}
d<-seq(min(xx),max(xx),len=1000)
lines(d,sapply(d,nad.fun),lwd=2,col="green")
acf(dati1,lag.max=500,main="",xlab="b",ylab="")
hist(dati1,prob=T,ylim = c(0, 0.2),main="",
ylab="",xlab="(b)")
xx<-seq(min(dati1),max(dati1),by=(max(dati1)
-min(dati1))/(n-1))
h1<-1.5
points(density(dati1,bw=h1),type="l",col="green")

## FARIMA process
library(fracdiff)
dati<-fracdiff.sim(1500, ar = .2, ma = .4, d = .4)
acf(as.ts(dati$series),lag.max=100,xlab="",ylab=""
,main="")
par(mfrow=c(1,2))
plot(as.ts(dati$series),xlab="(a)",ylab="",type="l")
xx<-1:1500
h<-7
nad.fun<-function(x)
{
sum(dnorm((x-xx)/h)*as.ts(dati$series))/
sum(dnorm((x-xx)/h))
}
d<-seq(min(xx),max(xx),len=1000)
```

```
lines(d,sapply(d,nad.fun),lwd=2,col="green")
acf(dati1,lag.max=500,main="",xlab="b",ylab="")
hist(as.ts(dati$series),prob=T,ylim = c(0, 0.4),
main="",ylab="",xlab="(b)")
h1<-0.3
points(density(as.ts(dati$series),bw=h1),type="l",
col="green")
```

4. programmas kods

```
#tiek izveidoti atkarigie dati
n<-100
sim<-10000
#AR<-0.5
#var<-1-(AR)^2
#ts.sim <- arima.sim(list(ar = AR,ma=c()), sim*n, sd = sqrt(var))
library("fSeries")
library(fGarch)
a0<-0.9
a1<-0.2
b1<-0.6
ts.sim = garchSim(model = list(omega = a0,alpha =a1,beta=b1), sim*n)
dati<-pnorm(ts.sim,0,1)
w<-20
autocov<-acf(dati,type = c("covariance"))[[1]][2:w]
s.sum<-sum(autocov[autocov>0.001])
sigma_kv<-acf(dati,type = c("covariance"))[[1]][1]+2*s.sum
#definejam polinomus
p1<-function(x)sqrt(3)*(2*x-1)
p2<-function(x)sqrt(5)*(6*x^2-6*x+1)
p3<-function(x)sqrt(7)*(20*x^3-30*x^2+12*x-1)
p4<-function(x)3*(70*x^4-140*x^3+90*x^2-20*x+1)
#definejam alternativi
#alt<-function(x){1/2+1/20*(x^(-4/5)+(1-x)^(-4/5))}
```

```
alt<-function(x){1/4*(x^(-1/2)+(1-x)^(-1/2))}
#altelnativas integralis
#g2<-function(y){1/2*y+1/4*y^(1/5)-1/4*(1-y)^(1/5)+1/4}
g2<-function(y){1/2*(1+y^(1/2)-(1-y)^(1/2))}
# 10000 reizu rekinam  statistiku
h<-log(n)
s<-c() # uzkraj statistiku
l<-c() # paligmainigais, lai izrekinatu  statistiku
k<-c() # uzkraj k vertibas
# precizitate, kas nepieciesama, lai rekinatu polinoma saknes
eps<-0.05
#uzkraj KS testa p  vertibas
kom.pval<-c()
for (i in 1:sim)  {
x<-dati[(n*(i-1)+1):(n*i)]
# rekina datus, kas veidojas ar alternativi
y<-c()
for (o in 1:n){
a<-0
b<-1
c<-(a+b)/2
h11<-abs(g2(c)-x[o])
while(h11>eps){
c<-(a+b)/2
h11<-abs(g2(c)-x[o])
h1<-(g2(a)-x[o])*(g2(c)-x[o])
if (h1<0) b<-c else a<-c
}
y[o]<-c
}
#Ar KolmogorovasSmirnova testu tiek rekinata  p-vertibaa
kom.pval[i]<-ks.test(y,"punif",0,1,alternative="greater")$p.value
```

```
d1<-sum(p1(y))
d2<-sum(p2(y))
d3<-sum(p3(y))
d4<-sum(p4(y))
d11<-n*(((1/n)*d1)^2)-1*h
d12<-n*(((1/n)*d1)^2+((1/n)*d2)^2)-2*h
d13<-n*(((1/n)*d1)^2+((1/n)*d2)^2+((1/n)*d3)^2)-3*h
d14<-n*(((1/n)*d1)^2+((1/n)*d2)^2+((1/n)*d3)^2+((1/n)*d4)^2)-4*h
d<-c(d11,d12,d13,d14)
k1<-order(d)[4] #  nosaka k  vertibu
k[i]<-k1  #  saglaba k vertibas
l[1]<-(((1/n)^(1/2))*d1)^2
l[2]<-(((1/n)^(1/2))*d1)^2+(((1/n)^(1/2))*d2)^2
l[3]<-(((1/n)^(1/2))*d1)^2+(((1/n)^(1/2))*d2)^2+(((1/n)^(1/2))*d3)^2
l[4]<-(((1/n)^(1/2))*d1)^2+(((1/n)^(1/2))*d2)^2+(((1/n)^(1/2))*d3)^2
+(((1/n)^(1/2))*d4)^2
s[i]<-l[k1]
}
s<-s*(1/12/sigma_kv)   ###  statistika tiek rekinata
ticam<-0.95 # ticamibas limenis
krit.vert<-qchisq(ticam,1)
u<-length(s[s>krit.vert])
p<-u/sim # jauda neimana testam
p
p1<-length(kom.pval[kom.pval<0.05])/sim #jauda KS testam
p1
```

Maģistra darbs “Jaukto procesu analīze un to pielietojums statistikā” 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.

Autore: Jolanta Rone

____________________ ________________

(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 maģistra gala pārbaudījuma komisijas sēdē

________________ prot. Nr. ________, vērtējums________

(datums)

Komisijas sekretārs/-e: ________________ ________________

(Vārds, Uzvārds) (paraksts)