VILNIAUS UNIVERSITETAS

JŪRATĖ ŠLIOGERĖ

SILPNAI PRIKLAUSOMŲ DYDŽIŲ SUMŲ DISKREČIOSIOS APROKSIMACIJOS

Daktaro disertacijos santrauka
Fiziniai mokslai, matematika (01 P)

Vilnius, 2016

Disertacija rengta 2011–2015 metais Vilniaus universitete

Mokslinis vadovas – prof. habil. dr. Vydas Čekanavičius (Vilniaus universitetas, fiziniai mokslai, matematika – 01 P)

Disertacija ginama Vilniaus universiteto Matematikos mokslo krypties mokslo taryboje:

Pirmininkas – prof. habil. dr. Remigijus Leipus (Vilniaus universitetas, fiziniai mokslai, matematika – 01 P)

Nariai:
Dr. Yang Yang (Nanjing audito universitetas, fiziniai mokslai, matematika – 01 P),
Prof. habil. dr. Vygantas Paulauskas (Vilniaus universitetas, fiziniai mokslai, matematika – 01 P),
Prof. habil. dr. Alfredas Račkauskas (Vilniaus universitetas, fiziniai mokslai, matematika – 01 P),
Prof. habil. dr. Jonas Kazys Sunklodas (Vilniaus universitetas, fiziniai mokslai matematika – 01 P)

Disertacija bus ginama viešame Matematikos mokslo krypties tarybos posėdyje 2016 m. liepos mėn. 4 d. 15 val. Vilniaus universiteto matematikos ir informatikos fakulteto 102 auditorijoje.
Adresas: Naugarduko g. 24, LT-03225 Vilnius, Lietuva

Disertacijos santrauka išsiuntinėta 2016 m. birželio mėn. 4 d.
Disertaciją galima peržiūrėti Vilniaus universiteto bibliotekoje ir VU interneto svetainėje adresu: www.vu.lt/lt/naujienos/ivykiu-kalendorius

VILNIUS UNIVERSITY

JŪRATĖ ŠLIOGERĖ

DISCRETE APPROXIMATIONS FOR SUMS OF WEAKLY DEPENDENT VARIABLES

Summary of doctoral dissertation
Physical sciences, mathematics (01 P)

Vilnius, 2016

Doctoral Dissertation was carried out during 2011–2015 at Vilnius University

Scientific supervisor – prof. habil. dr. Vydas Čekanavičius (Vilnius University, Physical Sciences, Mathematics – 01 P)

The dissertation will be defended at the Council of Scientific Field of Mathematics at Vilnius University:

Chairman – prof. habil. dr. Remigijus Leipus (Vilnius University, Physical Sciences, Mathematics – 01 P)

Members:
Dr. Yang Yang (Nanjing Audit University, Physical Sciences, Mathematics – 01 P),
Prof. habil. dr. Vygantas Paulauskas (Vilnius University, Physical Sciences, Mathematics – 01 P),
Prof. habil. dr. Alfredas Račkauskas (Vilnius University, Physical Sciences, Mathematics – 01 P),
Prof. habil. dr. Jonas Kazys Sunklodas (Vilnius University, Physical Sciences, Mathematics – 01 P)

The dissertation will be defended at the public meeting of the Council of Mathematics Science in the auditorium number 102 at the Faculty of Mathematics and Informatics of Vilnius University at 3 pm. on July 4, 2016.

Address: Naugarduko st. 24, LT-03225 Vilnius, Lithuania.
The summary of dissertation was distributed on June 4, 2016.
The dissertation is available at the Library of Vilnius University.
www.vu.lt/lt/naujienos/ivykiu-kalendorius

Mokslo problema

Atsitiktinių dydžių sumų ribinės teoremos atlieka pagrindinį vaidmenį tikimybių teorijoje ir plačiai taikomos statistikoje. Darbe nagrinėjami markoviškai priklausomi atsitiktiniai dydžiai, kurių reikšmės yra sveikieji skaičiai. Šioje srityje dauguma žinomų mokslinių tyrimų apriboti vadinamąja sekų schema, t.y., kai X_j nepriklauso nuo n ir $S_n = X_1 + X_2 + \cdots + X_n = S_{n-1} + X_n$. Tokiems dydžiams aproksimuoti buvo naudota centrinė ribinė teorema bei jos patikslinimai (Berry-Esseen tipo įverčiai ir Edgeworth tipo skleidimai). Atkreipiamas dėmesys, kad markoviškai priklausomų dydžių diskrečių sumų aproksimacija yra tolydus normalusis skirstinys. Dėl šios priežasties negalima taikyti nei pilnosios variacijos, nei taškinės metrikos, užtenka apsiriboti vien tolygiąja Kolmogorov'o metrika.

Bendresnė serijų schema, kai X_j priklauso nuo n, suma $S_n = X_{1,n} + X_{2,n} + \cdots + X_{n,n}$, $S_{n-1} = X_{1,n-1} + \cdots + X_{n-1,n-1}$ yra žymiai mažiau tirta. Šiems dydžiams taikytos Puasono ir sudėtinės Puasono aproksimacijos. Vis dėlto, daugumoje sudėtinių Puasono aproksimacijų yra stipri perėjimo tikimybių priklausomybė nuo dėmenų skaičiaus n, o aproksimacija sekų schemose yra trivialios $O(1)$ eilės. Taigi šių aproksimacijų negalime laikyti Gauso dėsnio diskrečiais analogais ar pakaitalais. Egzistuoja tiktai keletas aproksimacijų (daugiausia susijusių su markoviškai binominiu skirstiniu), kurios yra universalios ta prasme, kad gali efektyviai pakeisti normaliąją aproksimaciją esant visam perėjimo tikimybių spektrui: pradedant atveju, kai perėjimo tikimybės konverguoja į nulį eile $O(n^{-1})$, baigiant atveju, kai jos yra absoliutinės konstantos. Taip pat, kiek žinoma, Puasono tipo aproksimacijos simetrinei Markovo grandinei nebuvo tyrinėtos.

Šioje disertacijoje sudėtinę Puasono aproksimaciją taikome sumoms $\sum_{k=1}^{n} f(\xi_k)$, kur f – funkcija, įgyjanti reikšmes sveikųjų skaičių aibėje, $\xi_1, \xi_2, \cdots, \xi_n$ yra homogeninė Markovo grandinė su baigtiniu būsenų skaičiumi. Markovo grandinių serijų schema naudojama atvejais, kai f reikšmių sritis yra $\{0, 1\}$ arba $\{-1, 0, 1\}$. Specialus dėmesys skiriamas perėjimo tikimybių matricos simetrijai.

Tikslai ir uždaviniai

1. Parodyti, kad sekų schemose markoviškai priklausomų sveikareikšmių atsitiktinių dydžių aproksimacija paslinktu Puasono dėsniu gali būti tinkamesnė negu aproksimacija normaliuoju dėsniu.

2. Įrodyti Simons-Johnson teoremą markoviškai binominiam skirstiniui ir simetrinei trijų būsenų Markovo grandinei, t.y. pademonstruoti, kad konvergavimas į sudėtinį Puasono dėsnį yra stipresnis negu pilnosios variacijos metrikos atveju.

3. Įrodyti dalinį pirmosios tolygiosios Kolmogorov'o teoremos atvejį markoviškai binominiam skirstiniui konstruojant sudėtinį Puasono skirstinį, kuris aproksimuoja $O(n^{-2/3})$ eilės tikslumu.

4. Simetrinei trijų būsenų Markovo grandinei sukonstruoti sudėtinę Puasono aproksimaciją, kurios tikslumas būtų panašus į nepriklausomų simetrinių atsitiktinių dydžių sumos aproksimaciją lydinčiuoju Puasono dėsniu.

5. Simetrinei trijų būsenų Markovo grandinei sukonstruoti antros eilės sudėtines Puasono aproksimacijas.

6. Sukurti netolygius, lokalius įverčius bei įverčius iš apačios simetrinės trijų būsenų Markovo grandinės aproksimacijoms sudėtiniu Puasono dėsniu.

Aktualumas

Normaliosios aproksimacijos diskretūs analogai, kurie galioja stipresnėms metrikoms ir yra tikslesni serijų schemoms, turi teorinę ir praktinę reikšmę. Markoviškai priklausomų dydžių pradinis skirstinys gali būti labai sudėtingos struktūros. Tuo tarpu be galo dalios sudėtinės Puasono aproksimacijos yra aiškios struktūros ir daug patogesnės praktiniams skaičiavimams, ypač kai naudojamos Furjė transformacijos ar rekursiniai algoritmai.

Naujumas

Šioje disertacijoje pirmą kartą simetrinių markoviškai priklausomų atsitiktinių dydžių sumoms taikoma sudėtinė Puasono aproksimacija. Aproksimacijos tikslumas įvertinamas pilnosios variacijos, Wasserstein'o ir lokaliojoje metrikose. Pirmosios tolygiosios Kolmogorov'o teoremos dalinis atvejis įrodomas markoviškai binominiam skirstiniui. Taip pat markoviškai binominiam skirstiniui ir simetrinei trijų būsenų Markovo grandinei įrodoma Simons-Johnson teorema. Pademonstruota, kad sekų schemoms centrinės ribinės teoremos analoge paslinktas Puasono skirstinys gali būti naudojamas vietoje normaliojo dėsnio.

Metodai

Šioje disertacijoje naudojamas charakteristinių funkcijų metodas.

Ginamieji teiginiai

1. Markoviškai binominio skirstinio artumas be galo dalių dėsnių klasei pilnosios variacijos metrikoje yra ne blogesnis nei $O(n^{-2/3})$ eilės.
2. Simetrija pagerina markoviškai priklausomų dydžių aproksimaciją sudėtiniu Puasono dėsniu. Tritaškio skirstinio aproksimacijos tikslumas yra $O(n^{-1})$ eilės, toks pat kaip ir aproksimacijos sudėtiniu Puasono dėsniu nepriklausomų atsitiktinių dydžių atveju.
3. Sudėtinė Puasono aproksimacija gali būti taikoma serijų schemoje (kuomet kai kurios perėjimo tikimybės yra $o(1)$ eilės) ir sekų atveju (kai visos perėjimo tikimybės yra absoliutinės konstantos).
4. Įverčiai iš apačios pilnosios variacijos, lokalios ir Wasserstein'o metrikose patvirtina, kad sukonstruota sudėtinė Puasono aproksimacija simetrinei trijų būsenų Markovo grandinei yra tikslios eilės.
5. Markoviškai binominiam skirstiniui bei skirstiniui, paremtam simetrine trijų būsenų Markovo grandine, galioja Simons-Johnson teorema, t.y. į sudėtinį Puasono dėsnį konverguojama su eksponentiniais svoriais.
6. Simetrinei trijų būsenų Markovo grandinei antros eilės sudėtinės Puasono aproksimacijos pagerina konvergavimo greitį iki $O(n^{-2})$ eilės.
7. Paslinkta Puasono aproksimacija homogeninių Markovo grandinių schemoje yra to paties tikslumo kaip ir normalioji aproksimacija, tik stuktūriškai tinkamesnė.

Pagrindiniai rezultatai

Žymenys

Šioje disertacijoje naudojamos keturios metrikos. Pilnosios variacijos metrika tarp dviejų tikimybinių matų P ir Q, kurie koncentruoti sveikuose skaičiuose $\mathbb{Z}$, yra apibrėžiama taip:

$$\|P-Q\| = \sum_{k=-\infty}^{\infty} |P\{k\} - Q\{k\}| = 2 \sup_{A\subset\mathbb{Z}} |P(A) - Q(A)|.$$

Taip pat galima išreikšti ekvivalenčia forma:

$$\|P-Q\| \quad = \quad \sup_{f\in\mathcal{F}} \Big| \sum_{k=-\infty}^{\infty} f(k)P\{k\} - \sum_{k=-\infty}^{\infty} f(k)Q\{k\} \Big|.$$

Čia supremumas imamas pagal visas funkcijas $f : \mathbb{Z} \to \mathbb{R}$, kurios yra aprėžtos 1. Wasserstein'o metrika (taip pat žinoma kaip Dudley, Fortet-Mourier arba Kantorovich metrika) apibrėžiama taip:

$$\begin{aligned} \|P - Q\|_W &= \sum_{k=-\infty}^{\infty} |P\{(-\infty, k]\} - Q\{(-\infty, k]\}| \\ &= \sup_{f \in \mathcal{F}_1} \Big| \sum_{k=-\infty}^{\infty} f(k)P\{k\} - \sum_{k=-\infty}^{\infty} f(k)Q\{k\} \Big|. \end{aligned}$$

Čia supremumas imamas pagal visas funkcijas $f : \mathbb{Z} \to \mathbb{R}$, tenkinančias $\sup_k |f(k) - f(k-1)| \leqslant 1$. Lokali bei Kolmogorov'o metrikos atitinkamai yra užrašomos taip:

$$\|P - Q\|_\infty = \sup_{k \in \mathbb{Z}} |P\{k\} - Q\{k\}|, \qquad \|P - Q\|_K = \sup_{k \in \mathbb{Z}} |P\{(-\infty, k]\} - Q\{(-\infty, k]\}|.$$

Toliau pateikiame naudojamas santrumpas. Tegul I_k žymi pasiskirstymą, koncentruotą sveikajame taške $k \in \mathbb{Z}$, $I \equiv I_0$. Tegul U ir V yra du baigtiniai matai, koncentruoti sveikųjų skaičių aibėje $\mathbb{Z}$. Matų U bei V sandaugos ir laipsniai suprantami sąsūkos prasme, t.y. $U * V\{A\} = \sum_{k=-\infty}^{\infty} U\{A - k\} V\{A\}$ aibei $A \subseteq \mathbb{Z}$; taip pat $V^0 = I$. Mato V eksponentė išreiškiama $\exp\{V\} = \sum_{j=0}^{\infty} V^{*j}/j!$, o Fourier-Stieltjes transformacija žymima $\widehat{V}(t)$. Taip pat $\widehat{\exp\{V\}}(t) = \exp\{\widehat{V}(t)\}$, $\widehat{I}(t) = 1$, $\widehat{I}_a(t) = \mathrm{e}^{\mathrm{i}ta}$.

Taip pat apibrėžiame $U(x) = U\{(-\infty, x]\}$. Visiems $y \in \mathbb{R}$ ir $j \in \mathbb{N} = \{1, 2, 3, \dots\}$,

$$\binom{y}{j} = \frac{1}{j!} y(y-1)\dots(y-j+1), \qquad \binom{y}{0} = 1.$$

Visas teigiamas absoliutines konstantas žymime C. Tam tikrais atvejais, kad išvengtume galimų nesusipratimų, konstantas C pateikiame su indeksais. Žymėjimas $C(\cdot)$ naudojamas konstantoms, priklausančioms nuo nurodytų parametrų.

Pagrindinė disertacijos aproksimacija yra sudėtinis Puasono skirstinys su geometriniu skirstiniu eksponentėje. Pateikiame jo charakteristinę funkciją:

$$\exp\left\{ \frac{\lambda(\mathrm{e}^{\mathrm{i}t} - 1)}{1 - (1-p)\mathrm{e}^{\mathrm{i}t}} \right\}, \qquad 0 < p < 1.$$

Markoviškai priklausomų dydžių sekos: paslinkta Puasono aproksimacija

Sirazhdinov [15] pirmasis ištyrė homogeninių Markovo grandinių su baigtiniu būsenų skaičiumi konvergavimo į normalųjį dėsnį greitį. Nagaev [11] tęsė panašių homogeninių Markovo grandinių su begaliniu būsenų skaičiumi tyrimą. Pateiksime vieną pagrindinių Nagaev rezultatų. Tegul $\mathcal{Y}$ yra taškų ω aibė, $\mathcal{B}_{\mathcal{Y}}$ - jos poaibių σ-algebra, $p(\omega, \mathcal{A})$ - perėjimo tikimybių funkcija. Tegul $p(\omega, \mathcal{A})$ tenkina sąlygą: egzistuoja toks teigiamas sveikasis skaičius k_0, kad

$$\sup_{\omega, \tau, \mathcal{A}} |p^{(k_0)}(\omega, \mathcal{A}) - p^{(k_0)}(\tau, \mathcal{A})| = \delta < 1, \quad \mathcal{A} \in \mathcal{B}_{\mathcal{Y}}, \omega, \tau \in \mathcal{Y}, \tag{1}$$

kur $p^{(k_0)}(\omega, \mathcal{A})$ žymi k_0 žingsnių perėjimo tikimybių funkciją. Jei sąlyga (1) tenkinama, tuomet egzistuoja toks stacionarus skirstinys $p(\mathcal{A})$, kad

$$\sup_{\omega \in \mathcal{Y}, \mathcal{A} \in \mathcal{B}_{\mathcal{Y}}} |p(\mathcal{A}) - p^{(n)}(\omega, \mathcal{A})| \leqslant \delta^{\lfloor n/k_0 \rfloor} < \delta^{-1} \rho^n.$$

Čia $\rho = \delta^{1/k_0}$ ir $\lfloor \cdot \rfloor$ žymi sveikąją skaičiaus dalį.

Tarkime, kad atsitiktinių dydžių seka $\mathcal{Y}_1, \mathcal{Y}_2, \cdots, \mathcal{Y}_n, \cdots$ tenkina sąlygas:

$$\begin{aligned} \mathrm{P}(\mathcal{Y}_1 \in \mathcal{A}) &= \pi(\mathcal{A}), \\ \mathrm{P}(\mathcal{Y}_n \in \mathcal{A}) &= \int_{\mathcal{Y}} p^{(n-1)}(\omega, \mathcal{A})\pi(d\omega). \end{aligned}$$

Čia $\pi(\cdot)$ žymi pradinį skirstinį.

Tarkime, kad funkcija $f(\omega)$, apibrėžta aibėje $\mathcal{Y}$, yra mati sigma algebros $\mathcal{B}_{\mathcal{Y}}$ atžvilgiu. Jei $\pi(\mathcal{A}) = p(\mathcal{A})$ ir $\int_{\mathcal{Y}} f^2(\omega)p(d\omega) < \infty$, tuomet egzistuoja

$$\lim_{n\to\infty} \mathrm{E}\Big[\frac{1}{\sqrt{n}}\sum_{k=1}^{n}\Big(f(\mathcal{Y}_k) - \int_{\mathcal{Y}} f(\tau)p(d\tau)\Big)\Big]^2 = \sigma^2 > 0.$$

Žemiau pateiktos normuotos sumos pasiskirstymo funkciją žymima $F_{n\pi}(x)$.

$$\begin{aligned} \mathcal{S}_n &= \frac{1}{\sigma\sqrt{n}}\sum_{k=1}^{n+1}\Big(f(\mathcal{Y}_k) - \int_{\mathcal{Y}} f(\tau)p(d\tau)\Big) \\ &= \frac{1}{\sigma\sqrt{n}}\sum_{k=1}^{n+1} f(\mathcal{Y}_k) - \frac{n+1}{\sigma\sqrt{n}}\int_{\mathcal{Y}} f(\tau)p(d\tau) = \frac{\tilde{\mathcal{S}}_n}{\sigma\sqrt{n}} - \frac{\mathcal{A}}{\sigma\sqrt{n}}, \end{aligned}$$

čia $\tilde{\mathcal{S}}_n = \sum_{k=1}^{n+1} f(\mathcal{Y}_k)$ bei $\mathcal{A} = (n+1)\int_{\mathcal{Y}} f(\tau)p(d\tau)$. Standartinio normalaus pasiskirstymo funkciją žymime $\Phi(x)$. Sumos $\tilde{\mathcal{S}}_n$ pasiskirstymo funkciją žymime $\tilde{F}_{n\pi}(x)$. Toliau formuluojame sąlygas, naudojamas [11].

Sąlyga (H_k): realių skaičių aibėje egzistuoja tokia funkcija $g(x)$, kad $\lim_{x\to\infty} g(x) = \infty$ ir $\sup_\omega \int_X |f(\tau)|^k g(|f(\tau)|)p(\omega, d\tau) < \infty$.

Sąlyga $(\bar{H})$: tegul $\mathcal{Y}$ yra skaiti būsenų ω_i aibė, formuojanti teigiamą klasę, $f(\omega_i) = m + k_i h$, kur k_i yra sveikasis skaičius, m - bet koks realus skaičius, $h > 0$, ir pasirinktiems i ir j įmanoma rasti tokį indeksą k, kad $p_{ik} > 0$ ir $p_{jk} > 0$ $(p_{ik} = p(\omega_i, \omega_k))$.

Nagaev įrodė, kad jei tenkinamos sąlygos (H_3) ir $(\bar{H})$, didžiausias bendras daliklis k_i lygus 1 ir $\sum_{k=1}^{\infty} |f(\omega_k)|\pi(\omega_k) < \infty$, tuomet

$$F_{n\pi}(x) - \Phi(x) = \frac{1}{\sqrt{2\pi}}\mathrm{e}^{-x^2/2}\Big(\frac{Q_{1\pi}(x)}{\sqrt{n}} + \frac{S_1(x)}{\sqrt{n}}\Big) + o\Big(\frac{1}{\sqrt{n}}\Big), \tag{2}$$

tolygiai x-o atžvilgiu, kur

$$\begin{aligned} S_1(x) &= \frac{h}{\sigma}S_0\Big(\frac{(x+m_n)\sigma\sqrt{n}}{h}\Big), \qquad S_0(x) = \lfloor x \rfloor - x + \frac{1}{2}, \qquad m_n = -\frac{\sqrt{n}m}{\sigma}; \\ Q_{1\pi}(x) &= \frac{\lambda_{3p}}{6}(1-x^2) + \lambda_\pi, \qquad \lambda_\pi = \frac{1}{\sigma}\sum_{k=1}^{\infty}\mathrm{E}_\pi \tilde{f}(\mathcal{Y}_k), \\ \lambda_{3p} &= \frac{1}{\sigma^3}\Big(\mathrm{E}_p\tilde{f}^3(\mathcal{Y}_1) + 3\sum_{k=1}^{\infty}\mathrm{E}_p\tilde{f}^2(\mathcal{Y}_1)\tilde{f}(\mathcal{Y}_{k+1}) + 3\sum_{k=1}^{\infty}\mathrm{E}_p\tilde{f}(\mathcal{Y}_1)\tilde{f}^2(\mathcal{Y}_{k+1}) \\ &\quad +6\sum_{k,j=1}^{\infty}\mathrm{E}_p\tilde{f}(\mathcal{Y}_1)\tilde{f}(\mathcal{Y}_{k+1})\tilde{f}(\mathcal{Y}_{k+j+1})\Big) \\ \tilde{f}(\omega) &= f(\omega) - \int_{\mathcal{Y}} f(\tau)p(d\tau). \end{aligned}$$

Matome, kad diskretus skirstinys aproksimuojamas tolydžiu, todėl aproksimacijoje atsiranda papildomas glodinantis narys $S_1(x)/\sqrt{n}$. Normalioji aproksimacija pilnojoje variacijoje yra trivialios eilės, t.y. $\|F_{n\pi} - \Phi\| = O(1)$.

Suformuluosime ir lokalų įvertį, pateiktą [11]. Jei tenkinamos sąlygos (H_k), $(\bar{H})$ ir $\sum_{i=1}^{\infty} |f(\omega_i)|^{k-2}\pi(\omega_i) < \infty$ tenkinamos, pradinis skirstinys yra $\pi_i = \pi(\omega_i)$, tuomet

$$\mathrm{P}\Big(\sum_{1}^{n} f(\mathcal{Y}_i) = mn + sh\Big) = \frac{h}{\sigma\sqrt{n}} \cdot \frac{\mathrm{e}^{-z_{ns}/2}}{\sqrt{2\pi}}\left(1 + \frac{1}{\sqrt{n}}\left(\frac{\lambda_{3p}}{6}(z_{ns}^3 - 3z_{ns}) - z_{ns}\lambda_\pi\right)\right) + o(n^{-1/2}), \tag{3}$$

čia $z_{ns} = \Big(m(n+1) + sh - (n+1)\sum_{1}^{\infty} f(\omega_i)p(\omega_i)\Big)/\sigma\sqrt{n}$.

(2) ir (3) lygčių įrodymai paremti charakteristinių funkcijų skirtumu nulio aplinkoje.

Skirtingai negu Nagaev [11], kuris aproksimavo centruotos ir normuotos sumos $\mathcal{S}_n$ skirstinį $F_{n\pi}$, šioje disertacijoje aproksimuojame pradinės sumos $\tilde{\mathcal{S}}_n$ skirstinį $\tilde{F}_{n\pi}$. Toks pasirinkimas atrodo natūralesnis gardelinėms aproksimacijoms.

Tegul $\mathcal{G}$ yra matas su tokia charakteristine funkcija:

$$\begin{aligned}\widehat{\mathcal{G}}(t) \;&=\; \exp\{\mathrm{i}t\lfloor\mathcal{A} - n\sigma^2\rfloor + (n\sigma^2 + \{\mathcal{A} - n\sigma^2\})(\mathrm{e}^{\mathrm{i}t} - 1)\} \times \\ &\quad \left[1 + n\sigma^2\left(\sigma\alpha - \frac{1}{6}\right)(\mathrm{e}^{\mathrm{i}t} - 1)^3 + \sigma\lambda_\pi(\mathrm{e}^{\mathrm{i}t} - 1)\right].\end{aligned}$$

Suformuluosime pagrindinį trečio skyriaus pirmo poskyrio rezultatą.

1 Teorema. *Jei galioja sąlygos (H_k), $(\bar{H})$ ir $\sum_{i=1}^{\infty} |f(\omega_i)|\pi(\omega_i) < \infty$, tuomet visiems $n = 1, 2, \ldots$,*

$$\|\tilde{F}_{n\pi} - \mathcal{G}\|_\infty = o(n^{-1}), \tag{4}$$

$$\|\tilde{F}_{n\pi} - \mathcal{G}\|_K = o(n^{-1/2}). \tag{5}$$

Teorema įrodoma remiantis trikampio nelygybe bei Nagaev gautais rezultatais [11]. Teoremos įrodymas pateiktas disertacijoje.

Lyginant (4) su (3) bei (5) su (2), matome, kad aproksimacijų tikslumas išlieka toks pat. Iš kitos pusės, aproksimuojantysis matas $\mathcal{G}$ yra daug logiškesnės struktūros, nes jis apibrėžtas toje pačioje gardelėje kaip ir $\tilde{F}_{n\pi}$. Dėl šios priežasties, papildomas glodinantis narys $S_1(x)/\sqrt{n}$ tampa nebereikalingu.

MB skirstinio aproksimacijos

Markoviškai binominio skirstinio apibrėžimas skirtinguose šaltiniuose truputį skiriasi. Šioje disertacijoje pasirinkome bendriausią apibrėžimą iš [7]. Tegul $\xi_0, \xi_1, \ldots, \xi_n, \ldots$ yra nestacionari Markovo grandinė su pradiniu skirstiniu $\mathrm{P}(\xi_0 = 1) = p_0$, $\mathrm{P}(\xi_0 = 0) = 1 - p_0$, $p_0 \in [0, 1]$ ir perėjimo tikimybėmis

$$\begin{aligned}&\mathrm{P}(\xi_i = 1 \,|\, \xi_{i-1} = 1) = p_{11}, \qquad \mathrm{P}(\xi_i = 0 \,|\, \xi_{i-1} = 1) = p_{10},\\ &\mathrm{P}(\xi_i = 1 \,|\, \xi_{i-1} = 0) = p_{01}, \qquad \mathrm{P}(\xi_i = 0 \,|\, \xi_{i-1} = 0) = p_{00},\\ &p_{11} + p_{10} = p_{01} + p_{00} = 1, \qquad p_{11}, p_{01} \in (0, 1), \qquad i \in \mathbb{N}.\end{aligned}$$

Kitaip tariant, perėjimo tikimybių matrica atrodo taip:

$$\mathcal{P} = \begin{pmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{pmatrix}.$$

Sumos $S_n = \xi_1 + \cdots + \xi_n$ $(n \in \mathbb{N})$ skirstinys $\mathcal{L}(S_n)$ vadinamas markoviškai binominiu skirstiniu. Iš tikro S_n parodo sėkmingų bandymų skaičių iš n markoviškai priklausomų bandymų. Kiti autoriai į apibrėžimą dar įtraukia stacionarumo prielaidą arba pasirenka konkretų pradinį skirstinį. Pavyzdžiui, Dorbushin'o darbe

[7] tariama, kad $p_0 = 1$ ir nagrinėjama $S_{n-1} + 1$. Vienas pirmųjų markoviškai binominį skirstinį analizavo Koopman [9]. Jis perėjimo tikimybes susiejo su bandymų skaičiumi n ir tyrė sumos S_n ribinį skirstinį.

Markoviškai binominis skirstinys laikomas binominio skirstinio apibendrinimu. Jei $p_{00} = p_{10}$, tuomet markoviškai binominis skirstinys sutampa su binominiu skirstiniu. Žinoma, kad atitinkamai normuotas binominis skirstinys turi 2 neišsigimusius ribinius dėsnius - normalųjį ir Puasono. Tuo tarpu markoviškai binominis skirstinys turi 7 ribinius dėsnius.

Pateikiame keletą žinomų rezultatų, susijusių su markoviškai binominio skirstinio aproksimacijomis ir binominiais dėsniais. Pradedame nuo klasikinės Prokhorov'o Puasono aproksimacijos binominiams skirstiniui. Tegul $p \in (0,1]$ ir tegul $\mathrm{Pois}(np)$ žymi Puasono skirstinį su parametru np, t.y. $\mathrm{Pois}(np) = \exp\{np(I - I_1)\}$. Prokhorov pirmasis įrodė, kad jei $np \leqslant 1$, tuomet aproksimacijos tikslumas yra $O(p)$ eilės, žr. [13]. Įvertis su geriausia žinoma konstanta

$$\|((1-p)I + pI_1)^n - \mathrm{Pois}(np)\| \leqslant 2\min(p, np^2)$$

įrodytas [3].

Markov [10] parodė, kad, jei perėjimo tikimybės yra nenulinės konstantos, atitinkamai normuotas markoviškai binominis skirstinys yra asimptotiškai normalus. Jei tenkinamos sąlygos $p_0 = 1$, $p_{10}n \to \infty$, $p_{01}n \to \infty$ ir $\max\{np_{00}, np_{11}\} \to \infty$, tuomet markoviškai binominiam skirstiniui galioja

$$P\Big\{\frac{(S_{n-1}+1) - n\mathcal{M}}{\sqrt{n\mathcal{N}}} < x\Big\} \quad \to \quad \int_{-\infty}^{x} \mathrm{e}^{-t^2/2}\,\mathrm{d}t,$$

kur $\mathcal{M} = p_{01}/(p_{10} + p_{01})$ ir $\mathcal{N} = (p_{10}p_{01}(p_{11} + p_{00}))/(p_{10} + p_{01})^2$, žr. [7].

Dobrushin [7] parodė, kad normalusis skirstinys yra tik vienas iš septynių galimų ribinių dėsnių markoviškai binominiam skirstiniui. Pavyzdžiui, jei $p_{11} \to \tilde{p}$, $np_{01} \to \lambda$ ir $p_0 = 1$, tuomet sumos $S_{n-1} + 1$ ribinis dėsnis yra sąsūka $G_0 * \exp\{\lambda(G_0 - I)\}$, kur G_0 yra geometrinis skirstinys su parametru $\tilde{p}$, t.y. $\hat{G}_0(t) = (1-\tilde{p})\mathrm{e}^{\mathrm{i}t}/(1 - \tilde{p}\mathrm{e}^{\mathrm{i}t})$. Šis ribinis skirstinys buvo tirtas ir kitų autorių. Pvz., [6] įrodyta, kad, jei $p_0 = p_{01}/(p_{10} + p_{01})$ ir $\lambda > 0$, tuomet

$$\|\mathcal{L}(S_n) - \exp\{\lambda(G - I)\}\| \quad \leqslant \quad 2|np_{01} - \lambda| + \frac{2p_{01}(1 + p_{11} + np_{01}(2 - p_{11}))}{p_{10} + p_{01}},$$

čia G yra geometrinis skirstinys ir $\hat{G}(t) = p_{10}\mathrm{e}^{\mathrm{i}t}/(1 - p_{11}\mathrm{e}^{\mathrm{i}t})$. Aproksimacijos tikslumas yra ne geresnis negu $O(np_{01}^2)$ eilės. Panašus įvertis buvo gautas ir [19].

Serijų schema: pirma tolygioji Kolmogorov'o teorema markoviškai binominiam skirstiniui

Beveik prieš šešiasdešimt metų Kolmogorov [8] įrodė, kad tolygiojoje metrikoje bet koks nepriklausomų atsitiktinių dydžių sumos skirstinys mažai skiriasi nuo be galo dalių dėsnių aibės $\mathbb{D}$. Optimalius aproksimacijos tikslumo greičius nustatė Arak ir Zaĭtsev, išsamią šios problemos istoriją galima rasti [1]. Arak įrodė, kad vienodai pasiskirsčiusių atsitiktinių kintamųjų sumai (taip pat vadinama pirma tolygioji Kolmogorov'o teorema) galioja

$$C_1 n^{-2/3} \leqslant \sup_{F \in \mathcal{F}} \inf_{\mathcal{D} \in \mathbb{D}} \|F^{*n} - \mathcal{D}\|_K \leqslant C_2 n^{-2/3}. \tag{6}$$

Čia supremumas imamas pagal visų skirstinių aibę. Nors skirstiniui F nėra jokių prielaidų, aproksimacijos tikslumas vis tiek yra daug geresnis negu žymiojoje Berry-Esseen teoremoje. Be to, geriausia be galo dali aproksimacija nėra tiesiogiai susijusi su aproksimuojamosios sumos ribiniu eglesiu.

Apskritai, pirmoji tolygioji Kolmogorov'o teorema negalima pilnosios variacijos metrikoje. Iš kitos pusės, gardeliniams atsitiktiniams dydžiams, turintiems pakankamą momentų skaičių, tam tikras (6) analogas pilnojoje variacijoje galioja; žr. [2] teoremą 4.1. Iš tikro aproksimacija bet kokio F aibe $\mathbb{D}$ gali būti susiaurinta iki aproksimacijos, kai F yra gerai žinomas parametrinis skirstinys. Tuomet Kolmogorov'o problemą suprantame kaip F^{*n} artumo $\mathbb{D}$ klasei tolygų įvertinimą pagal visus galimus F parametrus. Pavyzdžiui, binominiam skirstiniui Kolmogorov'o problema yra visiškai išspręsta pilnojoje variacijoje, ž. [1] and [12], IV ir VIII skyrius. Tegul $\mathrm{Bi}(n,p)$, $p \leqslant 1/2$ žymi binominį skirstinį. Tuomet

$$C_3\varepsilon_{n,p} \leqslant \inf_{\mathcal{D}\in\mathbb{D}} \|\mathrm{Bi}(n,p)-\mathcal{D}\| \leqslant C_4\varepsilon_{n,p}. \tag{7}$$

Čia $\varepsilon_{n,p} = \min\big(np^2, p, \max\big((np)^{-2}, n^{-1}\big)\big)$. Toliau galima formuluoti tolygiąją Kolmogorov'o teoremą binominiam skirstiniui:

$$C_5 n^{-2/3} \leqslant \sup_{p\leqslant 1/2} \inf_{\mathcal{D}\in\mathbb{D}} \|\mathrm{Bi}(n,p)-\mathcal{D}\| \leqslant C_6 n^{-2/3}. \tag{8}$$

Pastebima, kad lygtyje (8) aproksimacijos tikslumo eilė yra ta pati kaip ir įvertyje (6), nors naudojama stipresne metrika.

Priklausomų atsitiktinių dydžių sumos aproksimacijos aibe $\mathbb{D}$ problema dar nėra išspręsta, nes sumos elgesys stipriai priklauso nuo priklausomybės pobūdžio. Netgi silpnai priklausomų Bernulio dydžių suma gali turėti savybes, kurios labai skirsis nuo binominio skirstinio savybių. Markoviškai binominis skirstinys yra tiesioginis binominio skirstinio apibendrinimas. Dėl šios priežasties aproksimuojant galime tikėtis kažkokio lygties (7) analogo. Iš teoremos 3.1 [6] išplaukia, kad, jei $p_{11} \leqslant 1/2$, tuomet

$$\inf_{\mathcal{D}\in\mathbb{D}} \|\mathcal{L}(S_n)-\mathcal{D}\| \leqslant Cp_{01}(p_{11}+p_{01})\min\Big(1, \frac{1}{\sqrt{np_{01}}}\Big) + C\min(p_{01}, np_{01}^2) + C(p_{11}+p_{01})\mathrm{e}^{-C_7 n}. \tag{9}$$

Čia $C_7 = \ln 30/19 = 0.4567\ldots$. Pastebime, kad esant $p_{01} = O(1)$, lygties (9) įvertis yra tos pačios trivialios eilės $O(1)$. Iš teoremos 1.3 [5] išplaukia, kad, jei $p_0 = 1$, $p_{11} \leqslant 1/20$, $p_{01}/(p_{10}+p_{01}) \leqslant 1/30$ and $p_{11} \leqslant p_{01}$, tuomet

$$\inf_{\mathcal{D}\in\mathbb{D}} \|\mathcal{L}(S_n)-\mathcal{D}\| \leqslant C(n^{-1}+(np_{01})^{-2}). \tag{10}$$

Prielaida $p_{11} \leqslant p_{01}$ yra labai ribojanti. Ji neapima sudėtinio Puasono ribos tuo atveju, kai $p_{11} \to \tilde{p} > 0$, $np_{01} \to \lambda$.

Šioje disertacijoje įrodomas (10) lygties analogas visiems $p_{11} \leqslant 1/4$ ir $p_{01} \leqslant 1/30$. Kitaip tariant, pasirenkame mažą p_{01} ir ne tokį mažą p_{11}, kurie pagal prielaidas reikalingi, kad ribinis dėsnis būtų sudėtinis Puasono skirstinys. Iš vienos pusės abi tikimybės mažos, iš kitos, tikimybės p_{01} ir p_{11} gali būti konstantos eilės, t.y. apimama ir centrinės ribinės teoremos galiojimo sritis.

Toliau įvedame įvairias S_n charakteristikas. Tegul

$$\begin{aligned}
\gamma_1 &= \frac{p_{10}p_{01}}{p_{10}+p_{01}}, \qquad \gamma_2 = -\frac{p_{10}p_{01}^2}{(p_{10}+p_{01})^2}\Big(p_{11}+\frac{p_{10}}{p_{10}+p_{01}}\Big) - \frac{\gamma_1^2}{2},\\
\gamma_3 &= \gamma_1^2\bigg\{\frac{\gamma_1}{3} + \frac{1}{p_{10}(p_{10}+p_{01})}\bigg\{p_{11}^2p_{01} + \frac{p_{11}p_{10}(2p_{01}-p_{10})}{p_{10}+p_{01}} + \frac{2p_{01}p_{10}^2}{(p_{10}+p_{01})^2}\bigg\}\\
&\quad + \frac{p_{01}}{p_{10}+p_{01}}\Big(p_{11}+\frac{p_{10}}{p_{10}+p_{01}}\Big)\bigg\},\\
\varkappa_1 &= \gamma_1\Big(\frac{p_{01}-p_{11}}{p_{10}+p_{01}} - p_0\Big), \quad \varkappa_2 = p_0\frac{p_{11}p_{10}}{p_{10}+p_{01}},\\
\gamma &= -6n(\gamma_2+\gamma_3)+\beta, \quad 0 \leqslant \beta < 1, \text{ ir } \beta \text{ yra toks, kad } \gamma \in \mathbb{Z},\\
\lambda_1 &= n(\gamma_1+4\gamma_2+3\gamma_3)-\beta, \quad \lambda_2 = \frac{\beta}{6}, \quad \lambda_{-1} = -n(2\gamma_2+3\gamma_3)+\frac{\beta}{3}.
\end{aligned}$$

Taip pat apibrėšime matus:

$$\begin{aligned}
\tilde{G} &= p_{10}I_1 * \sum_{j=0}^{\infty} p_{11}^j I_j \quad \left(\widehat{\tilde{G}}(t) = \frac{p_{10}\mathrm{e}^{\mathrm{i}t}}{1-p_{11}\mathrm{e}^{\mathrm{i}t}}\right), \quad \mathcal{H} = I + \varkappa_2(\tilde{G}-I),\\
\mathcal{U} &= \tilde{G}^{*\gamma} * \exp\{\lambda_1(\tilde{G}-I) + \lambda_2(\tilde{G}^{*2}-I) + \lambda_{-1}(I_{-1}-I)/p_{10}\}.
\end{aligned}$$

Suformuluosime pagrindinius trečio skyriaus antro poskyrio rezultatus.

2 Teorema. *Tegul* $p_{11} \leqslant 1/4$, $p_{01} \leqslant 1/30$, $np_{01} \geqslant 3$. *Tuomet*

$$\|\mathcal{L}(S_n) - \mathcal{H} * \exp\{\varkappa_1(\tilde{G}-I)\} * \mathcal{U}\| \leqslant C_8 \max\left(\frac{1}{n}, \frac{1}{(np_{01})^2}\right), \tag{11}$$

$$\|\mathcal{L}(S_n) - \mathcal{H} * \exp\{\varkappa_1(\tilde{G}-I)\} * \mathcal{U}\|_\infty \leqslant \frac{C_9}{\sqrt{np_{01}}} \max\left(\frac{1}{n}, \frac{1}{(np_{01})^2}\right). \tag{12}$$

1 Pastaba. *(i) Jeigu 2 teoremoje tikimybės* p_{11} *ir* p_{01} *yra absoliutinės konstantos, tuomet galioja centrinė ribinė teorema, žr. [7]. Šiuo atveju, (11) įvertis yra* $O(n^{-1})$ *eilės, t.y. daug geresnis negu* $O(n^{-1/2})$ *eilės, kurios būtų galima tikėtis iš normaliosios aproksimacijos.*

(ii) Lemoje 5.3 [6] įrodyta, kad $\mathcal{H} \in \mathbb{D}$. *Taip pat šioje disertacijoje parodoma, kad* $\exp\{\varkappa_1(\mathcal{Q}-I)\}*\mathcal{U} \in \mathbb{D}$, *todėl aproksimacija 2 teoremoje yra be galo dalus skirstinys (o ne ženklą keičiantis matas).*

Tolygiosios Kolmogorov'o problemos kontekste prielaida $np_{01} \geqslant 3$ nėra labai ribojanti. Jei $np_{01} < 3$, tuomet teoremoje gautas aproksimacijos tikslumas gali būti pasiektas su bet kokiu be galo daliu skirstiniu. Suformuluosime tolygiosios Kolmogorov'o teoremos versiją markoviškai binominiam skirtiniui.

3 Teorema. *Egzistuoja tokios absoliutinės teigiamos konstantos* C_{10} *ir* C_{11}, *kad visiems* $n \in \mathbb{N}$

$$\sup_{p_{11}\leqslant 1/4, p_{01}\leqslant 1/30} \inf_{\mathcal{D}\in\mathbb{D}} \|\mathcal{L}(S_n) - \mathcal{D}\| \leqslant C_{10} n^{-2/3}, \tag{13}$$

$$\sup_{p_{11}\leqslant 1/4, p_{01}\leqslant 1/30} \inf_{\mathcal{D}\in\mathbb{D}} \|\mathcal{L}(S_n) - \mathcal{D}\|_\infty \leqslant C_{11} n^{-5/6}. \tag{14}$$

Matome, kad aproksimacijos tikslumas įvertyje (13) yra toks pats kaip ir (6) bei (8). Lyginant su (10), čia nenaudojama labai ribojanti prielaida $p_{11} \leqslant p_{01}$. Klausimas apie galimą nelygybės (13) išplėtimą, kai $p_{01} \to 1$, lieka atviras.

Serijų schema: Simons-Johnson teorema markoviškai binominiam skirstiniui

1971 Simons ir Johnson [16] įrodė, kad binominio skirstinio $\mathrm{Bi}(n,p)$ konvergavimas į ribinį Puasono dėsnį $\mathrm{Pois}(\lambda)$ gali būti daug stipresnis negu pilnosios variacijos metrikoje. Jie įrodė, kad, jeigu $p = \lambda/n$ ir $g(x)$ tenkina $\sum_0^\infty g(k)\mathrm{Pois}(\lambda)\{k\} < \infty$, tuomet

$$\sum_{k=0}^{\infty} g(k)|\mathrm{Bi}(n,p)\{k\} - \mathrm{Pois}(\lambda)\{k\}| \to 0, \quad n \to \infty.$$

Rezultatas buvo išplėstas nepriklausomų gardelinių kintamųjų bendresniems atvejams. Panašūs rezultatai galioja m-priklausomiems dydžiams.

Taip pat [4] buvo įrodyta, kad, jei $h > 0$, $np_{01} = \lambda + o(1)$, $p_{11} = o(1)$, tuomet

$$\sum_{j=0}^{\infty} \mathrm{e}^{hj}|MB\{j\} - \mathrm{Pois}(\lambda)\{j\}| \to 0, \qquad n \to \infty. \tag{15}$$

Įvertis (15) yra nepatenkinamas dviem aspektais: a) jame naudojamas Puasono ribinis skirstinys vietoj bendresnio sudėtinio Puasono dėsnio; b) nėra konvergavimo greičio įverčio. Net aproksimuojant nepriklausomų atsitiktinių dydžių sumą, sudėtinis Puasono dėsnis kaip aproksimacija iki šiol tiriama esant labai ribojančioms prielaidoms, žr. [18]. Wang [18] taip pat pabrėžė, kad jo rezultatas yra netinkamas sudėtiniam Puasono dėsniui su sudėtiniu geometriniu skirstiniu.

Įrodysime, kad jei markoviškai binominis skirstinys konverguoja į sudėtinį Puasono ribinį dėsnį, konvergavimo greitis yra labai didelis yra jokia kita aproksimacija nėra reikalinga. Jei $p_{11} \to \tilde{p}$, $np_{01} \to \lambda$, tuomet markoviškai binominis skirstinys konverguoja į sudėtinį Puasono dėsnį su geometriniu skirstiniu, t.y. turime atvejį, kuris nebuvo išspręstas netgi esant nepriklausomiems dydžiams. Nepaisant to, įrodysime, kad galioja Simons-Johnson teoremos analogas. Tarę, kad p_{11} bei np_{01} yra arti jų tikėtinų reikšmių $\tilde{p}$ ir λ, įvertinsime konvergavimo greitį. Tegul

$$|p_{11} - \tilde{p}| \leqslant \frac{\tilde{p}}{3}, \quad |np_{01} - \lambda| \leqslant \frac{\lambda}{2}, \quad p_{01}(\mathrm{e}^h + 1)^2 \leqslant \frac{1}{25}, \quad \tilde{p}\mathrm{e}^h \leqslant \frac{1}{4}. \tag{16}$$

Tegul $Z = (1-\tilde{p})I_1 * \sum_{j=0}^{\infty} \tilde{p}^j I_j$ ir $\Psi = (I + p_0\tilde{p}(Z - I)) * \exp\{\lambda(Z - I)\}$. Suformuluojame pagrindinius trečio skyriaus trečio poskyrio rezultatus.

4 Teorema. *Tegul $h > 0$, $\lambda > 0$ ir $0 \leqslant \tilde{p} < 1$ yra absoliutinės konstantos. Jei sąlygos (16) galioja, tuomet*

$$\sum_{k=0}^{\infty} \mathrm{e}^{hk} |\mathcal{L}(S_n)\{k\} - \Psi\{k\}| \leqslant C_{12}(|p_{11} - \tilde{p}| + |np_{01} - \lambda| + n^{-1}). \tag{17}$$

1 Išvada. *Tegul $h > 0$, $0 \leqslant \tilde{p} < 1$, $\tilde{p}\mathrm{e}^h \leqslant 1/4$. Jei $np_{01} \to \lambda$, $p_{11} \to \tilde{p}$, tuomet*

$$\sum_{k=0}^{\infty} \mathrm{e}^{hk} |\mathcal{L}(S_n)\{k\} - \Psi\{k\}| \to 0.$$

2 Pastaba. *(i) Nėra aišku, kada būtų galima atsisakyti prielaidos $\tilde{p}\mathrm{e}^h \leqslant 1/4$. Šioje disertacijoje ji atlieka techninį vaidmenį, t.y. esant šiai sąlygai, eilutė $\sum_1^{\infty} (\tilde{p}\exp\{\mathrm{i}t + h\})^j$ konverguoja absoliučiai.*

(ii) Pastebima, kad jei $p_0 < 1$, tuomet $\Psi \in \mathbb{D}$, žr. lemą 5.3, [6].

(iii) Sprendimas pasirinkti eksponentinius svorius e^{hj} buvo pasirinktas remiantis įrodymo metodu.

Simetrinės trijų būsenų Markovo grandinės aproksimacija SP dėsniu

Gerai žinoma, kad esant pakankamai silpnoms sąlygoms, nepriklausomų atsitiktinių dydžių tinkamai normuota suma silpnai konverguoja į be galo dalų atsitiktinį kintamąjį. Po ribinės teoremos natūralu įvertinti konvergavimo greitį. Klasikiniu ribinės teoremos patobulinimu laikome konvergavimo greičio įvertinimą kokioje nors metrikoje. Tipiniai pavyzdžiai yra centrinė ribinė ir Berry-Esseen teoremos. Tipinė situacija, kai ribinis skirstinys naudojamas galimai aproksimacijai statistikoje.

Šiuolaikinėje tikimybių teorijoje aproksimacija kokiu nors skirstiniu reiškia, kad jo artėjimas į aproksimuojamą sumą yra tiriamas esant kuo silpnesnėms sąlygoms lyginant su tomis, kurios yra reikalingos ribinio skirstinio su panašia struktūra egzistavimui. Pavyzdžiui, (6) binominio ir Puasono skirstinių artumas įvertinamas esant visiems binominio skirstinio parametrams n ir p, nors Puasono riba atsiranda tik kai $p = O(1/n)$.

Vienas iš šios disertacijos tikslų yra išnagrinėti simetrijos efektą aproksimacijos tikslumui, kai aproksimuojami markoviškai priklausomi dydžiai sudėtiniu Puasono dėsniu. Simetrija pradiniame skirstinyje iš esmės pagerina nepriklausomų atsitiktinių dydžių sumų aproksimacijos tikslumą. Pavyzdžiui, vienas iš bendresnių rezultatų teigia, jei $\widehat{F}(t) \geqslant 0$ visiems $t \in \mathbb{R}$, tuomet

$$\|F^{*n} - \exp\{n(F - I)\}\|_K \leqslant C_{13} n^{-1}, \tag{18}$$

žr. teoremą 5.5, [1]. Panašūs rezultatai galioja ir atsitiktinių dydžių, kurie įgyja reikšmę 0 su daug didesne tikimybe, skirstiniams, žr. teoremą 7.1., [1]. Bendru atveju, įvertis (18) negalioja pilnajai variacijai. Kitavertus, jei $F\{1\} = F\{-1\} = b$, $F\{0\} = 1 - 2b$, $b \leqslant 1/4$, tuomet bet kuriam $n \in \mathbb{N}$

$$\|F^{*n} - \exp\{n(F - I)\}\| \leqslant C_{14} \min(nb^2, n^{-1}). \tag{19}$$

Išplėtimą žr. [12]. Franken'o sąlygą tenkinantiems skirstiniams įverčio (19) išpėtimą galima rasti [17].

Disertacijoje mūsų tikslas yra įrodyti įverčio (19) analogą markoviškai priklausomiems atsitiktiniams dydžiams. Tiriama simetrinė trijų būsenų Markovo grandinė. Įrodome, kad tokios Markovo grandinės suma gali būti aproksimuojama sudėtiniu Puasono dėsniu $O(1/n)$ eilės tikslumu. Nors riboje sudėtinis Puasono skirstinys atsiranda tik tuo atveju, kai kurios nors iš perėjimo tikimybių yra $O(1/n)$ eilės, aproksimacijos tikslumas įvertinamas esant daug silpnesnėms sąlygoms, t.y. kai visos perėjimo tikimybės yra $O(1)$ eilės. Įverčiai iš apačios parodo, kad esant tam tikroms sąlygoms, gauta aproksimacijos tikslumo eilė yra gera.

Tegul $\xi_0, \xi_1, \dots \xi_n, \dots$ yra nestacionari trijų būsenų $\{a_1, a_2, a_3\}$ Markovo grandinė. F_n žymi sumos $\mathcal{S}_n = f(\xi_1) + \cdots + f(\xi_n)$ $(n \in \mathbb{N})$ skirstinį, t.y. $P(\mathcal{S}_n = m) = F_n\{m\}$, kur $m \in \mathbb{Z}$. Čia $f(a_1) = -1$, $f(a_2) = 0$, $f(a_3) = 1$. Tegul pradinis skirstinys žymimas $P(\xi_0 = a_1) = \pi_1$, $P(\xi_0 = a_2) = \pi_2$ ir $P(\xi_0 = a_3) = \pi_3$. Pateikiamos perėjimo tikimybės:

$$\begin{gathered}
\mathrm{P}(\xi_i = a_1 \,|\, \xi_{i-1} = a_1) = a, \qquad \mathrm{P}(\xi_i = a_2 \,|\, \xi_{i-1} = a_1) = 1 - 2a, \qquad \mathrm{P}(\xi_i = a_3 \,|\, \xi_{i-1} = a_1) = a, \\
\mathrm{P}(\xi_i = a_1 \,|\, \xi_{i-1} = a_2) = b, \qquad \mathrm{P}(\xi_i = a_2 \,|\, \xi_{i-1} = a_2) = 1 - 2b, \qquad \mathrm{P}(\xi_i = a_3 \,|\, \xi_{i-1} = a_2) = b, \\
\mathrm{P}(\xi_i = a_1 \,|\, \xi_{i-1} = a_3) = a, \qquad \mathrm{P}(\xi_i = a_2 \,|\, \xi_{i-1} = a_3) = 1 - 2a, \qquad \mathrm{P}(\xi_i = a_3 \,|\, \xi_{i-1} = a_3) = a, \\
a, b \in [0, 0.5].
\end{gathered}$$

Taigi perėjimo tikimybių matrica yra tokia:

$$\begin{pmatrix} a & 1-2a & a \\ b & 1-2b & b \\ a & 1-2a & a \end{pmatrix}.$$

Įvedame tolimesnius žymėjimus:

$$
\begin{aligned}
L &= \frac{1}{2}(I_{-1}+I_1), \qquad X = L - I, \qquad D = (1-2(a-b)-2aX)^{*2} * (I+\Delta),\\
P_1 &= \pi_1 I + \frac{\Lambda_1 - 2aL}{1-2a}\pi_2 + \pi_3 I, \qquad P_2 = \pi_1 I + \frac{\Lambda_2 - 2aL}{1-2a}\pi_2 + \pi_3 I,\\
\Delta &= \frac{8b}{(1+2b)^2} X * \bigg(\sum_{j=0}^{\infty}\Big(\frac{2a}{1+2b}L\Big)^{*j}\bigg)^{*2}, \quad \Delta_1 = \frac{8(b-a)}{(1-2b)^2} L * \bigg(\sum_{j=0}^{\infty}\Big(\frac{-2a}{1-2b}\Big)^{j} L^{*j}\bigg)^{*2},\\
\Lambda_{1,2} &= \frac{1}{2}\bigg(I + 2(a-b)I + 2aX \pm (I - 2(a-b)I - 2aX) * \sum_{j=0}^{\infty}\binom{1/2}{j}\Delta^{*j}\bigg)\\
&= \frac{1}{2}\bigg((1-2b)I + 2aL \pm ((1-2b)I + 2aL) * \sum_{j=0}^{\infty}\binom{1/2}{j}\Delta_1^{*j}\bigg),\\
W_{1,2} &= \frac{1}{2}\bigg(I \pm \Big(\frac{2aX + (1-2a+2b)I}{(1+2b)^2}\Big) * \bigg(\sum_{j=0}^{\infty}\Big(\frac{2a}{1+2b}L\Big)^{*j}\bigg) * \sum_{j=0}^{\infty}\binom{-1/2}{j}\Delta^{*j}\bigg),\\
G &= \exp\Big\{\frac{2b(1-2a)}{1-2a+2b}(H-I)\Big\}, \qquad H = (1-2a)L * \sum_{j=0}^{\infty}(2aL)^{*j},\\
M &= \frac{1-2(a-b)}{1+2b}\sum_{j=0}^{\infty}\Big(\frac{2a}{1+2b}L\Big)^{*j}, \qquad E = \Big(1 - \frac{2a\pi_2}{1-2a}\Big)I + \frac{2a\pi_2}{1-2a}L,\\
A_1 &= \frac{-2b^2(1-2a)}{(1-2a+2b)^2}\Big((1+2a)I + \frac{2(1-2a)M}{1-2a+2b}\Big) * (H-I)^{*2},\\
E_1 &= \frac{2b(H-I)}{1-2a+2b}\pi_2, \qquad M_1 = \frac{-2b(1-2a)}{(1-2a+2b)^2}(2M-I)*(H-I).
\end{aligned}
$$

Skirstinys F_n konverguoja į sudėtinį Puasono dėsnį, kai $nb \to \tilde{b}_0$, $a \to \tilde{a}_0$ (kai $\tilde{a}_0 = 0$, turime Puasono dėsnį). Taigi sudėtinės Puasono ribos egzistavimui reikia labai ribojančios prielaidos $b = O(1/n)$. Kiek ši prielaida gali būti susilpninta, kad egzistuotų tinkama priešribinė sudėtinė Puasono aproksimacija? Kaip išplaukia iš žemiau pateiktų rezultatų, geras aproksimacijos tikslumas pasiekiamas net kai $b = O(1)$. Visi rezultatai gaunami esant sąlygai

$$0 \leqslant a \leqslant \frac{1}{30}, \quad 0 \leqslant b \leqslant \frac{1}{30}. \tag{20}$$

Įrodymo metodas nulemia konstantų mažumą. Formuluojame pagrindinius trečio skyriaus ketvirto poskyrio rezultatus.

5 Teorema. *Tegul galioja sąlyga (20). Tuomet visiems $n = 1, 2, \ldots$,*

$$
\begin{aligned}
\|F_n - E * M * G^{*n}\| &\leqslant C\bigg(\min\Big\{\frac{1}{n}, b\Big\} + 0.2^n|a-b|\bigg), &(21)\\
\|F_n - E * M * G^{*n}\|_\infty &\leqslant C\bigg(\min\Big\{\frac{1}{n\sqrt{nb}}, b\Big\} + 0.2^n|a-b|\bigg),\\
\|F_n - E * M * G^{*n}\|_W &\leqslant C\bigg(\min\Big\{\frac{\sqrt{nb}}{n}, b\Big\} + 0.2^n|a-b|\bigg).
\end{aligned}
$$

3 Pastaba. *Aproksimuojantysis dėsnis turi tris atskiras komponentes.*

(i) Sudėtinis binominis skirstinys E susijęs su pradiniu ξ_0 skirstiniu ir nėra be galo dalus. Vis dėlto, jei tarsime, kad $\pi_2 = 0$, tuomet $E = I$.

(ii) M yra sudėtinis geometrinis skirstinys su sudėtiniu simetriniu skirstiniu L. Taigi M yra tam tikras sudėtinio Puasono skirstinio atvejis. Iš tikrųjų,

$$M = \exp\Big\{\sum_{j=1}^{\infty}\Big(\frac{2a}{1+2b}\Big)^{j}\frac{L^{*j} - I}{j}\Big\}.$$

(iii) Pagrindinė aproksimuojančio dėsnio dalis yra n-kartė G sąsūka. Akivaizdu, kad G yra sudėtinis Puasono skirstinys su sudėtiniu skirstiniu H. Pastebime, kad H taip pat yra sudėtinis geometrinis skirstinys su sudėtiniu simetriniu skirstiniu L.

2 Išvada. *Tegul tenkinama sąlyga (20). Tuomet visiems $n = 1, 2, \dots,$*

$$\|F_n - E * M * G^{*n}\| \leqslant Cn^{-1}. \tag{22}$$

Lyginant (21) ir (19) įverčius matome, kad pastarasis yra tikslesnis kai $b < 1/n$. Iš esmės įmanoma gauti įvertį nb^2 sudėtingesnės aproksimacijos struktūros kaina. Kita išvada parodo šį faktą.

3 Išvada. *Tegul galioja sąlyga (20). Tuomet visiems $n = 1, 2, \dots,$*

$$\|F_n - (E + E_1) * M * (I + M_1) * G^{*n}\| \leqslant C\big(\min(nb^2, n^{-1}) + 0.2^n|a - b|\big).$$

4 Pastaba. *Jei $a = b$, tuomet F_n tampa F^{*n}, tirtu (19). Tuo tarpu $E*M*G^{*n}$ skiriasi nuo $\exp\{n(F-I)\}$. Skirtumas atsiranda, nes $E*M*G^{*n}$ atsižvelgia į galimą skirstinio F_n artumą sudėtiniui Puasono skirstiniui su sudėtiniu geometriniu dėsniu. Panašūs ribiniai skirstiniai yra neįmanomi nepriklausomiems tritaškiams atsitiktiniams dydžiams.*

Parodysime, kaip aproksimacijos tikslumas gali būti pagerintas antros eilės aproksimacija.

6 Teorema. *Tegul galioja sąlyga (20). Tuomet visiems $n = 1, 2, \dots,$*

$$\begin{aligned}
\|F_n - (E + E_1) * G^{*n} * (I + nA_1 + M_1)\| &\leqslant C\bigg(\min\bigg\{\frac{1}{n^2}, b^2\bigg\} + 0.2^n|a - b|\bigg),\\
\|F_n - (E + E_1) * G^{*n} * (I + nA_1 + M_1)\|_\infty &\leqslant C\bigg(\min\bigg\{\frac{1}{n^2\sqrt{nb}}, b^2\bigg\} + 0.2^n|a - b|\bigg),\\
\|F_n - (E + E_1) * G^{*n} * (I + nA_1 + M_1)\|_W &\leqslant C\bigg(\min\bigg\{\frac{\sqrt{nb}}{n^2}, b^2\bigg\} + 0.2^n|a - b|\bigg).
\end{aligned}$$

Aptarsime, kokiu mastu teoremos įverčiai yra nepagerinami. Pateiksime įverčius iš apačios.

7 Teorema. *Tegul galioja sąlyga (20), $nb \geqslant 1$ ir $\pi_2 = 0$. Tuomet egzistuoja tokios absoliutinės konstantos C_i, $(i = 14, \dots, 19)$ kad visiems $n = 1, 2, \dots,$*

$$\begin{aligned}
\|F_n - E * M * G^{*n}\| &\geqslant \|F_n - E * M * G^{*n}\|_K \geqslant \frac{C_{14}}{n}(1 - C_{15}\,0.2^n|a - b|), \qquad (23)\\
\|F_n - E * M * G^{*n}\|_\infty &\geqslant \frac{C_{16}}{n\sqrt{nb}}(1 - C_{17}\,0.2^n|a - b|),\\
\|F_n - E * M * G^{*n}\|_W &\geqslant \frac{C_{18}\sqrt{b}}{\sqrt{n}}(1 - C_{19}\,0.2^n|a - b|).
\end{aligned}$$

4 Išvada. *Tegul galioja sąlyga (20), $nb \geqslant 1$ ir $\pi_2 = 0$. Tuomet egzistuoja tokia absoliutinė konstanta C_{20}, kad visiems $n \geqslant C_{20}$,*

$$\|F_n - E * M * G^{*n}\| \geqslant \frac{C_{20}}{n}.$$

Taigi eilė $O(1/n)$, gauta (21), negali būti pagerinama.

Toliau suformuluosime netolygius įverčius. Jie reikalingi norint nustatyti aproksimacijos tikslumą, kai m yra toli nuo vidurkio.

8 Teorema. *Tegul $nb \geqslant 1$ ir tenkinama sąlyga (20). Tuomet visiems $m = 0, 1, 2, \dots,$*

$$\begin{aligned} |(F_n - E * M * G^{*n})\{m\}|\left(1 + \frac{|m|}{\sqrt{nb}}\right) &\leqslant \frac{C}{n\sqrt{nb}}, \\ |(F_n - E * M * G^{*n})\{(-\infty, m]\}|\left(1 + \frac{|m|}{\sqrt{nb}}\right) &\leqslant \frac{C}{n}. \end{aligned}$$

Suformuluosime Simons-Johnson teoremos analogą trijų būsenų Markovo grandinės skirstiniui. Įvesime tokius žymėjimus. Tegul $0 \leqslant \tilde{a} < 1$ ir

$$\mathcal{E} = \pi_1 I + \left(I - \frac{2\tilde{a}}{1 - 2\tilde{a}} X\right)\pi_2 + \pi_3 I, \qquad \mathcal{M} = (1 - 2\tilde{a})\sum_{j=0}^{\infty}(2\tilde{a}L)^{*j},$$

$$B = \exp\left\{\frac{2\lambda}{n}(L - I) * \sum_{j=0}^{\infty}(2\tilde{a}L)^{*j}\right\}, \qquad \mathcal{G} = \mathcal{E} * \mathcal{M} * B^{*n}.$$

Tegul

$$|a - \tilde{a}| \leqslant \frac{\tilde{a}}{10}, \qquad |nb - \lambda| \leqslant \frac{\lambda}{2}, \qquad b(e^h + 1)^2 \leqslant \frac{1}{25}, \qquad \tilde{a}e^h \leqslant \frac{1}{11}. \tag{24}$$

9 Teorema. *Tegul $h > 0$, $\lambda > 0$ ir $0 \leqslant \tilde{a} < 1$ yra kažkokios absoliutinės konstantos. Jei galioja prielaidos (24), tuomet*

$$\sum_{k=-\infty}^{\infty} e^{h|k|}|F_n\{k\} - \mathcal{G}\{k\}| \leqslant C(|a - \tilde{a}| + |nb - \lambda| + n^{-1}).$$

5 Išvada. *Tegul $h > 0$, $0 \leqslant \tilde{a} < 1$, $\tilde{a}e^h \leqslant 1/4$. Jei $nb \to \lambda$, $a \to \tilde{a}$, tuomet*

$$\sum_{k=-\infty}^{\infty} e^{h|k|}|F_n\{k\} - \mathcal{G}\{k\}| \to 0.$$

Visų suformuluotų teoremų įrodymai pateikti disertacijoje.

Išvados

Homogeninei Markovo grandinei su baigtiniu būsenų skaičiumi Nagaev [11] įvertino normaliosios aproksimacijos tikslumą. Šioje disertacijoje tokiai pat grandinei taikyta paslinkta Puasono aproksimacija. Lyginant su normaliąja aproksimacija, paslinkta Puasono aproksimacija turi šiuos privalumus:

- sukoncentruota ant tos pačios gardelės kaip ir pradinis skirstinys;
- kitaip nei normaliosios aproksimacijos atveju, nebereikia papildomo glodinančio nario;
- paslinktos Puasono aproksimacijos tikslumą galima įvertinti pilnosios variacijos metrikoje.

Markoviškai binominiam skirstiniui aproksimuoti sukonstruota geriausia be galo dali aproksimacija. Parodyta, kad

- aproksimacijos tikslumo greitis pilnosios variacijos metrikoje yra $O(n^{-1})$, t.y. daug geresnis negu normaliosios aproksimacijos, kurios eilė būtų $O(n^{-1/2})$;
- pirmosios tolygiosios Kolmogorov'o teoremos analogas turi tą patį tikslumą kaip ir binominio skirstinio atveju.

Įrodytas Simons-Johnson teoremos analogas markoviškai binominiam skirstiniui bei simetrinės trijų būsenų Markovo grandinės sumos skirstiniui. Parodyta, kad

- abiem atvejais aproksimuojama sukonstruotu sudėtiniu Puasono dėsniu su geometriniu skirstiniu;
- konvergavimas galioja sumoms su eksponentiniais svoriais;

Simetrinei trijų taškų Markovo grandinei sukonstruota sudėtinio Puasono dėsnio su sudėtiniu geometriniu skirstiniu aproksimacija. Parodyta, kad

- aproksimacija pilnojoje variacijoje yra $O(n^{-1})$ eilės, t.y. tokios pat kaip ir simetrizuoto binominio dydžio atveju;
- antros eilės sudėtinė Puasono aproksimacija pagerina konvergavimo įvertį iki $O(n^{-2})$ eilės;
- tam tikrais atvejais įvertis iš apačios irgi yra $O(n^{-1})$ eilės;
- gauti netolygūs įverčiai.

Aprobacija

Publikacijų disertacijos tema sąrašas

1. J. Šliogerė, V. Čekanavičius. Two limit theorems for Markov binomial distribution. *Lithuanian Math. J.* 55(3), 451–463, 2015.
2. J. Šliogerė, V. Čekanavičius. Approximation of symmetric three-state Markov chain by compound Poisson law. *priimtas, pasirodys Lithuanian Math. J.*

Pranešimų konferencijose sąrašas

1. J. Šliogerė, V. Čekanavičius. The approximation of Markov binomial distribution by negative binomial law. *11th International Vilnius Conference on Probability Theory and Mathematical Statistics : Abstracts of Communication*, p. 273, birželio 30 – Liepos 4, 2014, Vilnius, Lietuva.
2. J. Šliogerė, The approximation of Markov binomial distribution by compound Poisson distribution with compounding Geometric Distribution. *8th Nordic Econometric Meeting*, gegužės 28–30, 2015, Helsinkis, Suomija.
3. J. Šliogerė, V. Čekanavičius. Approximation of symmetric Markov chain by compound Poisson law. *Lietuvos matematikų draugijos LVI konferencija*, birželio 16 – 17, 2015, Kaunas.

Trumpa informacija apie autorę

Gimimo data ir vieta	1986 gruodžio 25 d., Ukmergė
Išsilavinimas	
2011−2015	Vilniaus universiteto matematikos doktorantūra
2009−2011	Baigta Vilniaus universiteto ekonometrijos studijų programa, įgytas statistikos magistro kvalifikacinis laipsnis
2005−2009	Baigta Vilniaus universiteto ekonometrijos studijų programa, įgytas statistikos bakalauro kvalifikacinis laipsnis
Darbo patirtis	
nuo 2014	Asistentė, Vilniaus universitetas, Matematikos ir Informatikos fakultetas
nuo 2011	Informacinių sistemų analitikė, TEO LT, AB

SUMMARY

Research topic

Limit Theorems for sums of random variables (rv's) play a central role in Probability Theory and have a wide range of applications in Statistics. In this thesis, we consider sums of Markov dependent integer-valued rv's. The majority of known research in this field is restricted to the scheme of sequences, i.e., when X_j does not depend on n and $S_n = X_1 + X_2 + \cdots + X_n = S_{n-1} + X_n$. The Central Limit Theorem (CLT) and its improvements (Berry-Esseen type estimates and Edgeworth type expansions) are investigated in many papers. Notably, the main approximation to the discrete sum of Markov dependent variables is continuous Normal distribution. Consequently, only uniform Kolmogorov metric can be applied. The more general scheme of triangular arrays ($S_n = X_{1,n} + X_{2,n} + \cdots + X_{n,n}$, $S_{n-1} = X_{1,n-1} + \cdots + X_{n-1,n-1}$) is considerably less explored. For this setup Poisson and Compound Poisson (CP) approximations are applied. However, the majority of CP approximations indirectly assume strong dependency of the transition probabilities on the number of summands n, and are of the trivial order $O(1)$ for the scheme of sequences. Thus, they can not be viewed as complete discrete analogues or the replacements of the Gaussian law. There are very few approximations (mainly related to the Markov Binomial (MB) distribution), which are universal in a sense that they can effectively replace the normal approximation for the whole spectrum of transition probabilities. In other words, the same approximation can be applied when certain transition probabilities converge to zero with the order $O(n^{-1})$ and to the case, when all transition probabilities are absolute constants. As far as we know, the case of Poisson type approximations to the symmetric Markov chain was not investigated.

In this thesis, we apply CP type approximations to the sums $\sum_{k=1}^{n} f(\xi_k)$, where f is integer-valued function and $\xi_1, \xi_2, \cdots, \xi_n$ form a homogeneous Markov chain with a finite state space. The triangular array of Markov chains is used for the case when the value domain of f is $\{0, 1\}$ or $\{-1, 0, 1\}$. In the letter case, a special emphasis is on the symmetry of transition matrix.

Actuality

Discrete analogues of the normal approximation, that hold for stronger metrics and are more precise in the scheme of triangular arrays are of theoretical and practical importance. The initial distribution of Markov dependent variables can be of a very complicated structure. Meanwhile, infinitely divisible CP approximations have explicit structures and are much more convenient for practical calculations, especially when combined with Fourier transforms or recursive algorithms.

Aims and goals

1. To show that translated Poisson approximation can be preferable to the Normal approximation in the scheme of sequences of Markov dependent integer-valued rv's.
2. To prove Simons-Johnson theorem for the MB distribution and symmetric three-state Markov chain, that is, to demonstrate that convergence to the CP limit holds in stronger than total variation metric.
3. To prove a partial case of the first uniform Kolmogorov theorem for the MB distribution by constructing CP distribution which approximates MB with the accuracy $O(n^{-2/3})$.
4. To construct a CP approximation to symmetric three-state Markov chain with the accuracy comparable to that of accompanying Poisson approximation to the sum of symmetric independent rv's.
5. To construct a second-order CP approximations for symmetric three-state Markov chain.
6. To obtain non-uniform, local and lower bound estimates for CP approximations for symmetric three-state Markov chain.

Novelty

In this thesis, CP approximation is for the first time applied to the sum of symmetric Markov dependent rv's. Its accuracy of approximation is estimated in the total variation, Wasserstein and local metrics. The partial case of the first uniform Kolmogorov theorem is proved for the MB distribution. The Simons-Johnson theorem is proved for MB and symmetric three-state Markov chain. It is demonstrated that translated Poisson distribution can replace the normal law in the analogue of the CLT for the scheme of sequences.

Main results

Translated Poisson approximation to Markov dependent integer-valued rv's distribution is applied (theorem 3.1 in the thesis). MB distribution is approximated by the set of infinitely divisible laws in the total variation and local metrics (theorems 3.2, 3.3 in the thesis). Simons-Johnson theorem for MB distribution and for symmetric three-state Markov chain is obtained (theorems 3.4, 3.9 in the thesis). Upper and lower bound estimates for approximation of a symmetric three-state Markov chain by the accompanying CP law are established in the total variation, local and Wasserstein metrics (theorems 3.5, 3.7 in the thesis). The approximation of symmetric three-state Markov chain is improved by the second-order approximations (theorem 3.6 in the thesis). Non-uniform local estimates are proved (theorem 3.8 in the thesis).

Statements presented for defence

1. The closeness of the MB distribution to the set of all infinitely divisible distributions in total variation is of the order $O(n^{-2/3})$.

2. Symmetry improves the approximation of Markov dependent variables by CP law. The accuracy of approximation for the sum of a three point distribution is of the order $O(n^{-1})$ and is equivalent to the accuracy of CP approximation in the case of independent rv's.

3. The CP approximations used for the symmetric three-state Markov chain are universal in the sense, that they can be applied for the case of triangular arrays (when some of the transition probabilities are of the order $o(1)$) and for the case of sequences (when all transition probabilities are absolute constants).

4. Lower bound estimates in total variation, local and Wasserstein metrics for constructed CP type approximation to symmetric three-state Markov chain confirm that upper bound estimates are of the right order.

5. The Simons-Johnson theorem holds for the MB distribution and the distribution based on a symmetric three-state Markov chain, that is the convergence to a CP limit holds with exponential weights.

6. Second order CP type approximations for symmetric three-state Markov chain improve the rate of accuracy to $O(n^{-2})$.

7. In the scheme of homogeneous Markov chains the translated Poisson approximation has the same accuracy as the Normal approximation, but is structurally more adequate.

Methods

In this thesis, the characteristic function method is used.

Structure of the thesis

In Chapter 2, the necessary notation is introduced and an overview of known results is given. Chapter 3 contains formulations of the results. All proofs and auxiliary results are given in Chapter 4. Finally the conclusions and bibliography are presented Chapter 5.

General conclusions

Nagaev [11] estimated the accuracy of normal approximation for homogeneous Markov chain with an finite number of states. In this thesis a translated Poisson approximation is constructed for the same Markov chain. In comparison to the normal approximation the translated Poisson approximation has the following advantages:

- it is concentrated on the same lattice as the initial distribution;
- unlike the normal case, no additional smoothing summands are needed;
- the accuracy of a translated Poisson approximation can be estimated in the total variation metric.

The best infinitely divisible approximation is constructed for MB distribution. It is shown that

- the accuracy of the approximation is of the order $O(n^{-1})$ in total variation, i.e. much better than $O(n^{-1/2})$, the order that can be expected from the normal approximation;
- the first uniform Kolmogorov theorem for the MB distribution has the same accuracy as for the case for binomial distribution.

The Simons-Johnson theorem is proved for the MB distribution and for distribution of sum of symmetric three-state Markov chain. It is proved that

- in both cases constructed approximation is a CP distribution with the compounding geometric distribution;
- the convergence holds for sums with exponential weights.

It is proved the limit law of symmetric three-state Markov chain is a CP distribution with the compounding geometric distribution. It is also proved that

- the approximation is of the order $O(n^{-1})$, i.e. the same order as for the case of symmetrized binomial rv;
- the second order CP approximation improves the rate of convergence and is of the order $O(n^{-2})$;
- in some cases the lower-bound estimates are also of the order $O(n^{-1})$;
- non-uniform estimates are constructed.

Publications

1. J. Šliogerė, V. Čekanavičius. Two limit theorems for Markov binomial distribution. *Lithuanian Math. J.* 55(3), 451–463, 2015.
2. J. Šliogerė, V. Čekanavičius. Approximation of symmetric three-state Markov chain by compound Poisson law. *to appear in Lithuanian Math. J.*

Conferences

1. J. Šliogerė, V. Čekanavičius. The approximation of Markov binomial distribution by negative binomial law. *11th International Vilnius Conference on Probability Theory and Mathematical Statistics : Abstracts of Communication*, p. 273, June 30 – July 4, 2014, Vilnius, Lithuania.

2. J. Šliogerė, The approximation of Markov binomial distribution by compound Poisson distribution with compounding geometric distribution. *8th Nordic Econometric Meeting.* May 28–30, 2015, Helsinki, Finland.

3. J. Šliogerė, V. Čekanavičius. Approximation of symmetric Markov chain by compound Poisson law. *Lietuvos matematikų draugijos LVI konferencija*, June 16 – 17, 2015, Kaunas.

Short information about the author

Birth date and place	1986 December 25, Ukmergė
Education	
2011−2015	PhD studies in Mathematics, Vilnius University
2009−2011	Master's Degree in Econometrics, Vilnius University
2005−2009	Bachelor's Degree in Econometrics, Vilnius University
Experience	
since 2014	Assistant at Vilnius University, Faculty of Mathematics and Informatics
since 2011	Information Systems Analyst at TEO LT, AB

Cituota literatūra

[1] T.V. Arak and A.Yu. Zaĭtsev, Uniform limit theorems for sums of independent random variables, *Proc. Steklov Inst. Math.* **174**, 1-222, 1988.

[2] A.D. Barbour and V. Čekanavičius, Total variation asymptotics for sums of independent integer random variables, *Ann. Probab.,* **30**, 509-545, 2002.

[3] A.D. Barbour, L. Holst, S. Janson. Poisson Approximation, *Clarendon Press, Oxford*, 2002.

[4] V. Čekanavičius. On the convergence of Markov binomial to Poisson distribution. *Stat. Probab. letters.* **58**, 83–91, 2002.

[5] V. Čekanavičius and M. Mikalauskas. Local theorems for the Markov binomial distribution. *Lithuanian Math. J.* **41,** 219–231, 2001.

[6] V. Čekanavičius and P. Vellaisamy. Compound Poisson and signed compound Poisson approximations to the Markov binomial law. *Bernoulli*, **16**(4), 1114–1136, 2010.

[7] R.L. Dobrushin. Limit theorems for a Markov chain of two states. *Izv. Akad. Nauk. USSR Ser. mat.* **17,** 291–330, 1953 (Russian). English translation in *Select. Transl. Math. Stat. and Probab.* **1,** Inst. Math. Stat. and Amer. Math. Soc. 97–134, 1961.

[8] A.N. Kolmogorov. Two uniform limit theorems for sums of independent random variables. *Theory Prob. Appl.* **1**, 384-394, 1956.

[9] B.O. Koopman. A generalization of Poisson's distribution for Markoff chains. *Proc. Nat. Acad. Sci. USA* 36, 202–207, 1950.

[10] A.A. Markov. Investigations of general experiments connected by a Markov chain. *Zap. Akad. Nauk Fiz. Mat. Otdel., VIII Ser.* **2 5**(3), 1910.

[11] S.V. Nagaev. More exact statements of limit theorems for homogeneous Markov chains. *Theory of probability and its applications* **6**, 62-–81, 1961.

[12] É.L. Presman. Approximation of binomial distributions by infinitely divisible ones. *Theory Probab. Appl.* **28**, 393–403, 1983.

[13] Yu.V. Prokhorov. Asymptotic behaviour of the binomial distribution. *Uspekhi Mat. Nauk.* **8**, 135-142 (Russian). English translation in *Select. Transl. Math. Stat. and Probab.* **1**, *Inst. Math. Stat. and Amer. Soc.* (1961) 87-95, 1953.

[14] A. Röllin. Translated Poisson approximation using exchangeable pair couplings. *Ann. Probab.* **17**, 1596-–1614, 2007.

[15] S.H. Sirazdinov, S.K. Formanov. On the estimates of the rate of convergence in the central limit theorem for homogeneous Markov chains. *Teor. Veroyatnost. i Primenen.* **28**(2) , 219–228, 1983.

[16] G. Simons and N.L. Johnson. On the convergence of binomial to Poisson distributions. *Ann. Math. Statist.* **42**, 1735-1736, 1971.

[17] J. Šiaulys and V. Čekanavičius. Approximation of distributions of integer-valued additive functions by discrete charges. I. *Lithuanian Math. J.* **28**(4), 392–401, 1988.

[18] Y.H. Wang. A compound Poisson convergence theorem. *Ann. Probab.* **19**, 452-455, 1991.

[19] Y.H. Wang. Approximating kth-order two-state Markov chains. *J. Appl. Probab.* **29,** 861–868, 1992.