Vilniaus Universitetas

Dainius Dzindzalieta

TIKSLIOSIOS BERNULIO TIKIMYBIŲ NELYGYBĖS

Daktaro disertacijos santrauka

Fiziniai mokslai, matematika (01 P)

Vilnius, 2014

Disertacija rengta 2009-2013 metais Vilniaus universiteto Matematikos ir informatikos institute.

Moksliniai vadovai:

habil. dr. Vidmantas Kastytis Bentkus (Vilniaus universitetas, fiziniai mokslai, matematika — 01 P) (2009–2010),

prof. dr. Saulius Norvidas (Vilniaus universitetas, fiziniai mokslai, matematika — 01 P) (2010–2013).

Mokslinis konsultantas

prof. dr. Friedrich Götze (Bylefeldo universitetas, Vokietija, fiziniai mokslai, matematika — 01 P).

Disertacija ginama Vilniaus universiteto Matematikos mokslo krypties taryboje:

Pirmininkas

prof. habil. dr. Kęstutis Kubilius (Vilniaus universitetas, fiziniai mokslai, matematika — 01 P).

Nariai:

prof. dr. Kęstutis Dučinskas (Klaipėdos universitetas, fiziniai mokslai, matematika — 01 P),

prof. habil. dr. Vygantas Paulauskas (Vilniaus universitetas, fiziniai mokslai, matematika — 01 P),

prof. habil. dr. Donatas Surgailis (Vilniaus universitetas, fiziniai mokslai, matematika — 01 P),

prof. dr. Darius Šiaučiūnas (Šiaulių universitetas, fiziniai mokslai, matematika — 01 P).

Oponentai:

prof. habil. dr. Algimantas Jonas Bikelis (Vytauto Didžiojo universitetas, fiziniai mokslai, matematika — 01 P),

doc. dr. Marijus Radavičius (Vilniaus universitetas, fiziniai mokslai, matematika — 01 P).

Disertacija bus ginama viešame Matematikos mokslo krypties tarybos posėdyje 2014 m. balandžio 30 d. 13 val. Vilniaus universiteto Matematikos ir informatikos instituto 203 auditorijoje.

Adresas: Akademijos g. 4, LT-08663 Vilnius, Lietuva.

Disertacijos santrauka išsiuntinėta 2014 m. kovo 28 d.
Disertaciją galima peržiūrėti Vilniaus universiteto bibliotekoje.

Vilnius University

Dainius Dzindzalieta

TIGHT BERNOULLI TAIL PROBABILITY BOUNDS

Summary of Doctoral Dissertation

Physical Sciences, Mathematics (01 P)

Vilnius, 2014

Doctoral dissertation was prepared at the Institute of Mathematics and Informatics of Vilnius University in 2009–2013.

Scientific Supervisors:

Habil. Dr. Vidmantas Kastytis Bentkus (Vilnius University, Physical Sciences, Mathematics — 01 P) (2009–2010),

Prof. Dr. Saulius Norvidas (Vilnius University, Physical Sciences, Mathematics — 01 P) (2010–2013).

Scientific Adviser

Prof. Dr. Friedrich Götze (University of Bielefeld, Germany, Physical Sciences, Mathematics — 01 P).

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

Chairman

Prof. Dr. Habil. Kęstutis Kubilius (Vilnius University, Physical Sciences, Mathematics — 01 P).

Members:

Prof. Dr. Kęstutis Dučinskas (Klaipeda University, Physical Sciences, Mathematics — 01 P),

Prof. Dr. Habil. Vygantas Paulauskas (Vilnius University, Physical Sciences, Mathematics — 01 P),

Prof. Dr. Habil. Donatas Surgailis (Vilnius University, Physical Sciences, Mathematics — 01 P),

Prof. Dr. Darius Šiaučiūnas (Siauliai University, Physical Sciences, Mathematics — 01 P).

Opponents:

Prof. Dr. Habil. Algimantas Jonas Bikelis (Vytautas Magnus University, Physical Sciences, Mathematics — 01 P),

Prof. Dr. Marijus Radavičius (Vilnius University, Physical Sciences, Mathematics — 01 P).

The dissertation will be defended at the public meeting of the Council of Scientific Field of Mathematics in the auditorium No. 203 at the Institute of Mathematics and Informatics of Vilnius University, at 1 p.m. on the 30th of April, 2014.

Address: Akademijos st. 4, LT-08663, Vilnius, Lithuania.

The summary of the doctoral dissertation was distributed on the 28th of March, 2014.
A copy of the doctoral dissertation is available for review at the Library of Vilnius University.

1 Įžanga

Disertacijoje, kuri pristatoma šioje santrumpoje, nagrinėjame izoperimetrinį uždavinį

$$D(\mathcal{F}, I, n) = \sup_{S_n \in \mathcal{F}} \mathbb{P}\{S_n \in I\}, \tag{1}$$

kur $\mathcal{F}$ yra klasė priklausomų arba nepriklausomų Bernulio atsitiktinių dydžių sumų S_n, o I yra intervalas, kuris yra aprėžtas arba neaprėžtas. Visame darbe, jei $I = [x, \infty)$ arba $I = \{x\}$, mes vietoje $D(\mathcal{F}, I, n)$ rašome $D_n(x)$. Rašydami $\sup_{S_n}$, mes turime galvoje $\sup_{S_n \in \mathcal{F}}$, kur $\mathcal{F}$ yra klasė sumų $S_n = X_1 + \cdots + X_n$, kurios elementai X_i tenkina duotas sąlygas. Dėl šios priežasties mes formuluojame uždavinius pateikdami tik sąlygas (aprėžtumas, momentinės sąlygos, priklausomumas ir t.t.) atsitiktiniams dydžiams X_i. Disertacijoje mes randame optimalius viršutinius rėžius funkcijai $D(\mathcal{F}, I, n)$ ir pratęsiame rezultatus Lipšico funkcijų klasei.

Norėdami iliustruoti mūsų sprendžiamų uždavinių tipą, mes pateikiame patį paprasčiausią atvejį. Pats paprasčiausias mūsų nagrinėjamas atvejis yra pateikiamas antrajame skyriuje, kuriame nagrinėjami simetriniai nepriklausomi aprėžti atsitiktiniai dydžiai X_i.

Tarkime $S_n = X_1 + \cdots + X_n$ yra suma simetrinių aprėžtų nepriklausomų atsitiktinių dydžių, kad $|X_i| \leq 1$. Tarkime $I = [x, \infty)$. Remiantis klasikiniu Hoeffdingo 1963 metų rezultatu mes gauname

$$D_n(x) = \mathbb{P}\{S_n \geq x\} \leq \exp\{-x^2/2n\}, \quad x \in \mathbb{R}. \tag{2}$$

Hoeffdingas 1963 parodė, kad ši nelygybė galioja ir tuo atveju, kai $S_0, S_1, \ldots$ sudaro martingalinę seką.

Jei vietoje S_n imtume sumą nepriklausomų vienodai pasiskirsčiusių Rademacherio dydžių $W_n = \varepsilon_1 + \cdots + \varepsilon_n$, t. y. atsitiktinių dydžių, kad $\mathbb{P}\{\varepsilon_i = \pm 1\} = \frac{1}{2}$, tada remiantis Centrine Ribine Teorema mes matome, kad eksponentinė funkcija dešinėje lygties (2) pusėje yra minimali. Visgi tam tikro daugiklio, kurio eilė yra x^{-1}, trūksta, nes standartinio normalaus dydžio uodega dideliems x yra apytiksliai lygi $(\sqrt{2\pi}x)^{-1}\exp\{-x^2/2\}$.

Be to, galima parodyti, kad suma S_n, kaip atsitiktinis dydis, yra subgausinis, t. y. galioja nelygybė

$$D_n(x) \leq c\mathbb{P}\{\sqrt{n}Z \geq x\}, \quad x \in \mathbb{R},$$

kur Z yra standartinis normalinis atsitiktinis dydis, o c yra tam tikra absoliuti konstanta.

Yra įvairių Hoeffdingo tipo nelygybės (2) pagerinimų ir apibendrinimų, keičiant sąlygas sumų dėmenims (simetriškumas, vienpusis aprėžtumas, dvipusis aprėžtumas, momentiniai apribojimai, dėmenų priklausomybė). Mūsų žiniomis, nėra nei vieno rezultato, kuriame būtų gauta tiksli funkcijos $D(\mathcal{F}, I, n)$ reikšmė. Mūsų darbo tikslas yra rasti tokias poras $(\mathcal{F}, I)$, kurioms galima nustatyti tikslias funkcijos $D(\mathcal{F}, I, n)$ reikšmes ir aprašyti ekstremalius elementus.

2 Disertacijos mokslinė problema ir tyrimo objektas

Disertacijos tyrimo objektas yra tikimybinės atsitiktinių dydžių sumų uodegos. Disertacijos mokslinė problema yra surasti tikslius viršutinius šių uodegų rėžius ir apibūdinti kiekvienos nagrinėjamos klasės ekstremalius elementus.

3 Pagrindiniai uždaviniai

Disertacijoje nagrinėjami šie uždaviniai:

- **Ekstremali kombinatorika ir tikimybinės nelygybės.** Antrajame skyriuje mes nagrinėjame nepriklausomų simetrinių vienodai aprėžtų dydžių sumas. Pagrindinis šio skyriaus uždavinys yra nustatyti, kokia nepriklausomų atsitiktinių dydžių suma maksimizuoja tikimybę patekti į intervalą I. Šiame skyriuje nagrinėsime tris skirtingus intervalo I tipus. Pirmu atveju intervalas I yra pustiesė $[x, \infty)$, antru atveju vientaškė aibė $\{x\}$, o trečiu antveju $[x-k, x+k)$, kur $x \in \mathbb{R}$ ir $k \in \mathbb{N}$. Pirmais dviem atvejais tarsime, kad atsitiktiniai dydžiai X_i yra aprėžti iš viršaus, t. y. $|X_i| \leq 1$. Trečiu atveju atsitiktiniai dydžiai yra aprėžti iš apačios, t. y. $X_i \geq 1$. Trečiu atveju uždavinys yra klasikinis Erdös 1945 išspręstas uždavinys. Mes pateikiame Erdös uždavio įrodymą nenaudojant Spernerio teorijos. Skyriaus pabaigoje mes pateikiame rezultatų apibendrinimą simetrinių Lipšico funkcijų klasei.
- **Svorinės Rademacherio sumos.** Trečiajame skyriuje mes vėl nagrinėjame simetrinių nepriklausomų atsitiktinių dydžių sumą, bet šiame skyriuje mes papildomai pareikalaujame, kad sumos dispersija būtų aprėžta iš viršaus, o nagrinėjami dydžiai būtų dvitaškiai. Šiame skyriuje mes randame optimalią subgausinę konstantą, bei pageriname klasikinę Čebyševo nelygybę $D_n(x) \leq 1/2x^2$ visiems $x > 1$, panaudodami apibendrintą Čebyševo nelygybę.
- **Sąlyginės nelygybės.** Ketvirtajame skyriuje mes nagrinėjame atsitiktinių dydžių sumas, kurios turi tuos pačius aprėžtumo ir momentinius apribojimus, kaip ir trečiajame skyriuje. Šiame skyriuje mes nagrinėjame, kas būtų, jei tartume, kad tikimybė patekti į tam tikrą intervalą yra fiksuota ir žinoma. Ar galima, ką nors pasakyti apie tokios sumos savybes to intervalo galuose? Mes parodome, kad prie tam tikrų reguliarumo sąlygų, sumos yra „beveik tolydžios“.
- **Martingalinės nelygybės.** Penktajame skyriuje mes nagrinėjame aprėžtų atsitiktinių dydžių sumas, kurios nariai yra priklausomi. Mūsų konkrečiu atveju nagrinėjame sumas, kurios sudaro martingalinę seką. Hoeffdingas 1963 parodė, kad martingalams galioja tos pačios eksponentinės nelygybės, kaip ir nepriklausomų dėmenų atveju. Iškyla natūralus klausimas, ar maksimizuojantys atsitiktiniai dydžiai turi panašias savybes. Šiame skyriuje parodome, kad ir maksimizuojantys martingalai turi panašias savybes kaip ir maksimizuojančios nepriklausomų atsitiktinių dydžių sumos.
- **Ekstremalios Lipšico funkcijos.** Šeštajame skyriuje mes apibendriname prieš tai buvusius rezultatus Lipšico funkcijų, apibrėžtų ant metrinių tikimybinių erdvių, klasei. Naudojant pirmuose skyriuose naudotus rezultatus, gaunami nuokrypio nuo vidurkio viršutiniai rėžiai rodo, kad maksimalios Lipšico funkcijos turėtų būti „labai panašios“ į tiesines funkcijas. Iškyla klausimas, ar maksimizuojančios Lipšico funkcijos yra tiesinės funkcijos. Šiame skyriuje parodome, kad maksimalios Lipšico funkcijos yra „beveik“ tiesinės ir lygios neigiamai atstumo nuo tam tikros ekstremalios aibės funkcijai.

4 Tyrimų metodika

Pagrindinis mūsų darbe naudojamas metodas yra matematinė indukcija pagal dimensiją. Kadangi šis metodas yra labai paprastas, tai didžiausias iššūkis yra sugalvoti, kaip tiksliai konkrečiu atveju ją reikia panaudoti. Tiesioginis bandymas ją pritaikyti priveda prie griozdiškų struktūrų, su kuriomis dirbti tampa neįmanoma. Trumpai aptarsime, kaip pritaikėme matematinę indukciją kiekviename skyriuje. Antrajame – penktajame skyriuose nagrinėjamos tokios klasės, kuriose maksimizuojantys atsitiktiniai dydžiai yra dvitaškiai. Integruojant pagal vieną iš kintamųjų nelygybę $\mathbb{P}\{S_n \geq x\}$, mes gauname du dėmenis, kurių dimensija jau yra lygi $n-1$, todėl galime pritaikyti indukcinę prielaidą. Kai kuriais atvejais indukcinės prielaidos negalima taikyti tiesiogiai, nes gautų dėmenų suma gaunasi didesnė, nei sakoma indukciniame teiginyje. Tada į dviejų tikimybių sumą žiūrima, kaip į dviejų indikatorių sumos vidurkį. Kelios aritmetinės gudrybės ir vidurkio adityvumas padeda gauti norimą rezultatą. Šeštajame skyriuje indukcijos jau negalima pritaikyti, nes jame nagrinėjamos ne tik erdvės gaunamos sudauginus vienmates erdves, todėl dimensijos sąvoka neturi prasmės. Šiuo atveju mes parodome, kad mūsų nagrinėjamas uždavinys yra ekvivalentus izoperimetriniam aibių teorijos uždaviniui, kuris yra jau gana plačiai išnagrinėtas.

5 Moksliniai rezultatai

5.1 Simetrinių dydžių sumos.

Antrajame skyriuje mes nagrinėjame nepriklausomų simetrinių atsitiktinių (aprėžtų ir neaprėžtų) dydžių sumas. Mūsų tikslas yra gauti tikslius viršutinius funkcijos $D(\mathcal{F}, I, n)$ įverčius. Iš pateiktų rezultatų matysis, kad pateikti rėžiai yra nepagerinami.

Tarkime $S_n = X_1 + \cdots + X_n$ yra tokia nepriklausomų simetrinių dydžių suma, kad $|X_i| \leq 1$. Hoeffding 1963 parodė, kad $\mathbb{P}\{S_n \geq x\} \leq \exp\{-x^2/2n\}$ visiems $x \in \mathbb{R}$. Nuo to laiko buvo gauta daug šios nelygybės pagerinimų ir apibendrinimų, bet, mūsų žiniomis, nebuvo gautas nė vienas tikslus tokio tipo rezultatas. Mes antrajame skyriuje pateikiame tikslų viršutinį funkcijos $\mathbb{P}\{S_n \geq x\}$ rėžį. Tarkime $W_n = \varepsilon_1 + \cdots + \varepsilon_n$ yra suma nepriklausomų vienodai pasiskirsčiusių Rademacherio dydžių, t. y. atsitiktinių dydžių tokių, kad $\mathbb{P}\{\varepsilon_i = \pm 1\} = 1/2$. Tada galioja ši teorema.

Theorem 1 (Dzindzalieta, Juškevičius, Šileikis 2012)**.** *Visiems $x > 0$ turime*

$$\sup_{S_n} \mathbb{P}\{S_n \geq x\} = \begin{cases} \mathbb{P}\{W_n \geq x\}, & \textit{jei } \lceil x \rceil + n \in 2\mathbb{Z}, \\ \mathbb{P}\{W_{n-1} \geq x\}, & \textit{jei } \lceil x \rceil + n \in 2\mathbb{Z}+1. \end{cases}$$

Mes taip pat nagrinėjame dydį $\sup_{S_n} \mathbb{P}\{S_n = x\}$, kurį galima laikyti netolygiu sumos S_n koncentracijos taške x rėžiu.

Theorem 2 (Dzindzalieta, Juškevičius, Šileikis 2012)**.** *Jei $x > 0$ ir $k = \lceil x \rceil$ tada*

$$\sup_{S_n} \mathbb{P}\{S_n = x\} \leq \mathbb{P}\{W_m = k\},$$

kur

$$m = \begin{cases} \min\{n, k^2\}, & \textit{jei } n + k \in 2\mathbb{Z}, \\ \min\{n-1, k^2\}, & \textit{jei } n + k \in 2\mathbb{Z}+1. \end{cases}$$

Lygybė galioja, kai $S_n = \frac{x}{k} W_m$.

Mes savo rezultatus įrodome naudodami du visiškai skirtingus metodus. Vienas būdas yra tiesiogiai taikyti matematinę indukciją. Antras būdas yra taikyti ekstremalios kombinatorikos rezultatus. Darbe pateikiame ir klasikinės Erdös 1945 teoremos paprastą įrodymą naudojant matematinę indukciją. Šio uždavinio, kurį pirmasis įrodė Erdös 1945 metais naudodamas Spernerio teoremą, formuluotė yra tokia.

Theorem 3 (Erdös 1945). *Tegul $S_n = a_1\varepsilon_1 + \cdots + a_n\varepsilon_n$, kur a_i yra tokie realūs skaičiai, kad $|a_i| \geq 1$. Tada*

$$\max_{x\in\mathbb{R}} \sup_{S_n} \mathbb{P}\{S_n \in (x-k, x+k]\} = \mathbb{P}\{W_n \in (-k, k]\}.$$

Nors mūsų uždavinyje nagrinėjami atsitiktiniai dydžiai nebūtinai yra dvitaškiai, mes parodome, kad maksimizuojantys atsitiktiniai dydžiai yra dvitaškiai. Tai rodo tiesioginį mūsų rezultatų ryšį su Erdös 1945 uždaviniu.

Lemma 4 (Dzindzalieta, Juškevičius, Šileikis 2012). *Tegul $g : \mathbb{R}^n \to \mathbb{R}$ yra aprėžta išmatuojama funkcija. Tada*

$$\sup_{X_1,\ldots,X_n} \mathbb{E}\, g(X_1, \ldots, X_n) = \sup_{a_1,\ldots,a_n} \mathbb{E}\, g(a_1\varepsilon_1, \ldots a_n\varepsilon_n),$$

kur supremumas kairėje lygybės pusėje imamas pagal visus tokius nepriklausomus simetrinius atsitiktinius dydžius $X_1, \ldots, X_n$, kad $|X_i| \leq 1$, o dešinėje supremumas imamas pagal skaičius $-1 \leq a_1, \ldots, a_n \leq 1$.

Norėdami įrodyti 2 Teoremą, mes įrodome šias lemas. Tegul $S_n = a_1\varepsilon_1 + \cdots + a_n\varepsilon_n$. Naudodami indukcijos metodą įrodome, kad galioja šis rezultatas.

Lemma 5 (Dzindzalieta, Juškevičius, Šileikis 2012). *Tegul $x > 0$, $k = \lceil x \rceil$. Tarkime $0 < a_1 \leq \cdots \leq a_n \leq 1$. Tada*

$$\mathbb{P}\{S_n = x\} \leq \begin{cases} \mathbb{P}\{W_n = k\}, & \text{jei} \quad n+k \in 2\mathbb{Z}, \\ \mathbb{P}\{W_{n-1} = k\}, & \text{jei} \quad n+k \in 2\mathbb{Z}+1. \end{cases} \tag{3}$$

Kitas pagalbinis rezultatas duoda tikslią funkcijos $\sup_{S_n} \mathbb{P}\{S_n = x\}$ reikšmę ir todėl pagerina 5 lemą.

Lemma 6 (Dzindzalieta, Juškevičius, Šileikis 2012). *Tegul $0 < a_1, \ldots, a_n \leq 1$, $x > 0$, $k = \lceil x \rceil$. Tada*

$$\sup_{S_n} \mathbb{P}\{S_n = x\} = \begin{cases} \mathbb{P}\{W_n = k\}, & \text{jei} \quad n+k \in 2\mathbb{Z}, \\ \mathbb{P}\{W_n = k+1\}, & \text{jei} \quad n+k \in 2\mathbb{Z}+1. \end{cases}$$

Ši lema įrodoma naudojant ekstremalios kombinatorikos rezultatus.

Antrojo skyriaus pabaigoje mes pratęsiame rezultatus nelyginių Lipšico funkcijų klasei. Tarkime $C_n = [-1, 1]^n$ yra kubas su ℓ^1 metrika d. Mes sakome, kad funkcija $f : C_n \to \mathbb{R}$ yra K-Lipšico su $K > 0$, jei

$$|f(x) - f(y)| \leq K d(x, y), \qquad x, y \in C_n. \tag{4}$$

Mes sakome, kad funkcija $f : C_n \to \mathbb{R}$ yra nelyginė, jei $f(-x) = -f(x)$ visiems $x \in C_n$.

Naudojant ekstremalios kombinatorikos rezultatus mes įrodome šią teoremą.

Theorem 7 (Dzindzalieta, Juškevičius, Šileikis 2012). *Tarkime funkcija $f : C_n \to \mathbb{R}$ yra 1-Lipšico ir nelyginė. Tegul $X_1, \ldots, X_n$ yra tokie simetriniai nepriklausomi atsitiktiniai dydžiai, kad $|X_i| \leq 1$. Tada visiems $x > 0$ mes gauname*

$$\sup_{X_1,\ldots,X_n} \mathbb{P}\{f(X_1, \ldots, X_n) \geq x\} = \begin{cases} \mathbb{P}\{W_n \geq x\}, & \text{jei } n + \lceil x \rceil \in 2\mathbb{Z}, \\ \mathbb{P}\{W_{n-1} \geq x\}, & \text{jei } n + \lceil x \rceil \in 2\mathbb{Z}+1. \end{cases} \tag{5}$$

5.2 Svorinės Rademacherio a. d. sumos.

Trečiajame skyriuje mes vėl nagrinėjame sumą $S_n = a_1\varepsilon_1 + \cdots + a_n\varepsilon_n$, kur a_i yra realūs skaičiai, o ε_i yra nepriklausomi Rademacherio atsitiktiniai dydžiai. Šiame skyriuje mes papildomai pareikalaujame, kad šios sumos dispersija būtų aprėžta iš viršaus, t. y. $\mathbb{E}\, S_n^2 \le 1$. Ši sąlyga yra ekvivalenti sąlygai $\sum_{i=1}^n a_i^2 \le 1$. Kadangi mūsų rėžiai šiame skyriuje nepriklauso nuo dimensijos, tai vietoje S_n mes taip pat rašome S.

I. Pinelis 1994 metais parodė, kad suma S yra subgausinė, t. y. egzistuoja tokia absoliuti konstanta $c > 0$, kad $\mathbb{P}\{S \ge x\} \le c\mathbb{P}\{\eta \ge x\}$ visiems $x \in \mathbb{R}$, kur $\eta \sim N(0,1)$ žymi standartinį normalinį atsitiktinį dydį.

Bentkus 2007 metais iškėlė hipotezę, kad optimali konstanta, kurią žymėsime raide c_*, yra tokia, kad $\mathbb{P}\left\{\frac{\varepsilon_1+\varepsilon_2}{\sqrt{2}} \ge \sqrt{2}\right\} = c_*\mathbb{P}\{N \ge \sqrt{2}\}$, t. y. $c_* \approx 3.178$. Mūsų svarbiausias šio skyriaus rezultatas yra šios hipotezės įrodymas.

Theorem 8 (Bentkus, Dzindzalieta 2014)**.** *Tegul $\eta \sim N(0,1)$ yra standartinis normalinis atsitiktinis dydis, tada*

$$\mathbb{P}\{S_n \ge x\} \le c\mathbb{P}\{\eta \ge x\} \quad \textit{visiems } x \in \mathbb{R} \tag{6}$$

su konstanta c lygia

$$c_* := (4\mathbb{P}\{\eta \ge \sqrt{2}\})^{-1} \approx 3.178.$$

Kitas svarbus šio skyriaus rezultatas yra viršutinis rėžis funkcijai $\mathbb{P}\{S_n \ge x\}$ visiems $x \in (1, \sqrt{2}]$.

Theorem 9 (Dzindzalieta 2013)**.**

$$\mathbb{P}\{S \ge x\} \le \frac{1}{4} + \frac{1}{8}\left(1 - \sqrt{2 - \frac{2}{x^2}}\right), \qquad \textit{kai } x \in (1, \sqrt{2}).$$

Šis rezultatas yra ypač svarbus algoritmų teorijos specialistams, kuriems reikia kuo tikslesnių rėžių, kai x yra „mažas“.

Norėdami įrodyti šias teoremas mes naudojome Berry-Esseen nelygybę ir pagalbinę lemą, kuri yra Čebyševo nelygybės apibendrinimas.

Lemma 10 (Bentkus, Dzindzalieta 2014)**.** *Tegul $s > 0$ ir $0 \le a \le b$, tada kiekvienam atsitiktiniam dydžiui Y galioja*

$$a^s\mathbb{P}\{|Y| \ge a\} + (b^s - a^s)\mathbb{P}\{|Y| \ge b\} \le \mathbb{E}\,|Y|^s.$$

Jei Y yra simetrinis, tada

$$a^s\mathbb{P}\{Y \ge a\} + (b^s - a^s)\mathbb{P}\{Y \ge b\} \le \mathbb{E}\,|Y|^s/2.$$

Naudojant šią lemą mes galime gauti, pavyzdžiui, kad

$$\mathbb{P}\{S_n \ge 1\} + \mathbb{P}\{S_n \ge \sqrt{2}\} + \mathbb{P}\{S_n \ge \sqrt{3}\} + \cdots \le 1/2.$$

5.3 Sąlyginės uodegų tikimybės

Ketvirtajame skyriuje mes vėl nagrinėjame svorines nepriklausomų Rademacherio a. d. sumas $S_n = a_1\varepsilon_1 + \cdots + a_n\varepsilon_n$, su tokiomis pačiomis sąlygomis kaip ir trečiajame skyriuje. Trečiajame skyriuje įrodymuose naudojama Berry-Esseen nelygybė sako, kad jei visi svoriai (skaičiai a_i) yra maži (lokalus atvejis), tai sumos S_n uodegos tikimybė yra "netoli" standartinio normalinio dydžio uodegos tikimybės. Šiame skyriuje mes parodysime, kad jei pareikalausime tam tikros koncentracijos sąlygos, tai suma S_n turi ir panašias koncentracijos savybes, kaip ir normalusis dydis. Skyriaus rezultatai yra taikomi Būlio funkcijų, su Būlio reikšmėmis, diskrečiojoje Furje analizėje.

Mūsų pagrindinis skyriaus rezultatas yra ši teorema.

Theorem 11 (Dzindzalieta, Götze 2013). *Tegul $\|a\|_2 = 1$, $\|a\|_\infty \leq \tau$ kažkokiam $\tau > 0$ ir*

$$\mathbb{P}\{|S_n - \tau| < \theta\} = 1 - \epsilon \tag{7}$$

kažkokiam $0 \leq \epsilon \leq 1$ ir $\theta \in \mathbb{R}$. Tada egzistuoja tokia absoliuti konstanta $c > 0$, kad

$$\mathbb{P}\{|S_n - \theta| \leq \tau\} \geq \frac{c\epsilon\tau}{\sqrt{1+\tau^2}}. \tag{8}$$

Optimali konstanta šioje nelygybėje yra ne didesnė nei $3\sqrt{2}/64 \approx 0.066$.

Šios teoremos įrodyme naudojamas kitas svarbus rezultatas.

Lemma 12 (Dzindzalieta, Götze 2013). *Tegul $\|a\|_2 \leq 1$ and $\|a\|_\infty \leq \tau$. Tada egzistuoja tokia absoliuti konstanta $c > 0$, kad*

$$\mathbb{P}\{|S_n| \leq \tau\} \geq \frac{c\tau}{\sqrt{1+\tau^2}}. \tag{9}$$

Optimali konstanta šioje nelygybėje yra ne didesnė nei $3\sqrt{2}/16 \approx 0.26$.

5.4 Tikimybinės nelygybės martingalams

Penktajame skyriuje mes nagrinėjame atsitiktinius n žingsnių priklausomus klajojimus, tarkime $W_n = (M_0, M_1, \dots, M_n)$, prasidedančius taške 0 ir apibrėžtus ant martingalų sekos $M_k = X_1 + \cdots + X_k$ (tarkime $M_0 = 0$) su skirtumais $X_m = M_m - M_{m-1}$. Tegul $\mathcal{M}$ yra klasė martingalų su tokiais aprėžtais skirtumais, kad $|X_m| \leq 1$ ir $\mathbb{E}(X_m|\mathcal{F}_{m-1}) = 0$ didėjančios σ-algebrų sekos $\emptyset \subset \mathcal{F}_0 \subset \cdots \subset \mathcal{F}_n$ atžvilgiu. Jei atsitiktinis klajojimas W_n yra apibrėžtas ant martingalų sekos iš klasės $\mathcal{M}$, tada mes simboliškai rašome $W_n \in \mathcal{M}$. Mes tariame, kad klasė yra tokia didelė, kad į ją patenka visi martingalai su skirtumais, kurie yra sąlyginai Bernulio atsitiktiniai dydžiai.

Mes išsprendžiame izoperimetrinį uždavinį

$$D_n(x) = \sup_{W_n \in \mathcal{M}} \mathbb{P}\left\{W_n \text{ patenka į intervalą } [x, \infty)\right\}, \quad x \in \mathbb{R}. \tag{10}$$

Šis uždavinys yra atskiras Hoeffding 1963 antros teoremos atvejis.

Mes parodome, kokie atsitiktiniai klajojimai yra ekstremalūs ir pateikiame tikslią išreikštinę funkcijos $D_n(x)$ reikšmę.

Pirma mes įrodome bendrą savybę martingalams.

Theorem 13 (Dzindzalieta 2013)**.** *Tarkime* $\tau_x = \min\{k : M_k \geq x\}$ *yra stabdymo laikas. Jeigu martingalų klasė* $\mathcal{M}$ *yra uždara stabdymo atžvilgiu, tada*

$$\sup_{W_1,\dots,W_n \in \mathcal{M}} \mathbb{P}\left\{\max_{1\leq k\leq n} M_k \geq x\right\} = \sup_{M_n\in\mathcal{M}} \mathbb{P}\{M_n \geq x\}.$$

Tai reiškia, kad (10) *tipo problemų sprendiniai yra nehomogeniniai Markovo procesai.*

Mes darbe apibūdiname maksimizuojantį atsitiktinį klajojimą RW_n. Įsivaizduokite klajojimą, kaip atsitiktinį dalelės judėjimą. Jos kiekvienas žingsnis priklauso tik nuo to, kiek jau buvo padaryta žingsnių, tarkime k, ir nuo atstumo nuo aibės $A = [x, \infty)$, tarkime ρ_k. Jei $\rho_k = 0$, tada dalelė nejuda. Jei $0 < \rho_k < 1$ ir $n-k$ yra nelyginis, tada dalelė juda ρ_k link aibės A arba $1 - \rho_k$ tolyn nuo aibės A. Jei $\rho_k + n - k$ sveikoji dalis yra nelyginė ir $\rho_k \geq 1$, tada dalelė juda per 1 link aibės A arba $1 - \{\rho_k\}$ tolyn nuo aibės A. Jei $\rho_k + n - k$ sveikoji dalis yra lyginė, tada dalelė juda per $\{\rho_k\}$ link aibės A arba 1 tolyn nuo aibės A.

Mūsų pagrindinis rezultatas yra ši teorema.

Theorem 14 (Dzindzalieta 2013)**.** *Atsitiktinis klajojimas* RW_n *maksimizuoja tikimybę patekti į intervalą* $[x, \infty)$ *per pirmus* n *žingsnių, t. y. galioja ši lygybė*

$$D_n(x) = \mathbb{P}\{RW_{n,x} \text{ patenka į intervalą } [x,\infty)\} = \mathbb{P}\{M_{n,x}^x \geq x\} \tag{11}$$

visiems $x \in \mathbb{R}$ *ir* $n = 0, 1, 2, \dots$.

Tarkime $B(n, k)$ yra normalizuota $n - k + 1$ mažiausių binominių koeficientų suma, t. y.

$$B(n,k) = 2^{-n} \sum_{i=0}^{n-k} \binom{\lfloor \frac{i}{2} \rfloor}{n}.$$

Tarkime $x = m + \alpha$, $m \in \mathbb{Z}$ ir $0 \leq \alpha < 1$. Išreikštinė funkcijos $D_n(x)$ priklauso nuo $m + n$ lyginumo. Jei $m + n$ yra nelyginis, tada

$$D_n(x) = \sum_{i=0}^{h} a_i\, B(n-i-1, m+i), \quad a_i = \frac{\alpha^i}{(1+\alpha)^{i+1}}, \tag{12}$$

kur $h = (n - m - 1)/2$.

Jei $m + n$ yra lyginis, tada

$$D_n(x) = \sum_{i=0}^{m+1} b_i\, B(n-i-1, m-i+1), \tag{13}$$

kur $b_i = \frac{(1-\alpha)^i}{(2-\alpha)^{i+1}}$, kai $i < m$, $b_m = \alpha \left(\frac{1-\alpha}{2-\alpha}\right)^m$ ir $b_{m+1} = (1-\alpha)\left(\frac{1-\alpha}{2-\alpha}\right)^m$.

Šie rezultatai taip pat galioja, jei martingalų klasę $\mathcal{M}$ pakeisime į supermartingalų klasę.

5.5 Ekstremalios Lipšico funkcijos

Šeštajame skyriuje mes nagrinėjame Lipšico funkcijas ant tikimybinių metrinių erdvių (*TME*) (V, d, μ). Mes sakome, kad (V, d, μ) yra *TME*, jei matas μ yra normalizuotas ($\mu(V) = 1$) Borelio matas. Duotai netuščiai aibei $A \in V$ mes apibrėžiame atstumo funkciją $d(A, u) = \min\{d(u, v), v \in A\}$. Tegul $\mathcal{F} = \mathcal{F}(V, d, \mu)$ yra klasė integruojamų, t. y. $f \in L_1(V, d, \mu)$, 1-Lipšico funkcijų $f : V \to \mathbb{R}$, kad $|f(u) - f(v)| \le d(u, v)$ visiems $u, v \in V$. Sakome, kad klasė $\mathcal{F}$ yra *pilna*, jei atstumo funkcija priklauso tai klasei.

Mes nagrinėjame izoperimetrinę problemą

$$D(x) = \sup_{f \in \mathcal{F}} \mu\{f - \mathbb{E}_\mu f \ge x\} \qquad \text{visiems } x \in \mathbb{R}. \tag{14}$$

Mūsų pagrindinis šio skyriaus rezultatas yra tai, kad ši problema yra ekvivalenti kitai izoperimetrinei problemai.

Duotam $t \ge 0$ ir $h \ge 0$,

$$\text{minimizuoti } \mu(A^h) \text{ , kai } A \subset V \text{ toks, kad } \mu(A) \ge t, \tag{15}$$

kur $A^h = \{u \in V : d(u, A) \le h\}$ yra aibės A h-plėtinys.

Mes sakome, kad TME (V, d, μ) yra *izoperimetrinė*, jei visiems $t \ge 0$ egzistuoja (15) sprendinys, žymime A_{opt}, kuris nepriklauso nuo h.

Mūsų pirmas šio skyriaus rezultatas tinka visoms pilnoms TME. Tarkime $\mathcal{F}^*$ yra $\mathcal{F}$ poaibis visų neigiamų atstumo funkcijų pavidalo $-d(\cdot, A)$, $A \subset V$.

Theorem 15 (Dzindzalieta 2013)**.** *Jei $\mathcal{F}(V, d, \mu)$ yra pilna, tada* sup *problemoje* (14) *yra pasiekiamas ant neigiamų atstumo funkcijų klasės*

$$\sup_{f \in \mathcal{F}} \mu\{f - \mathbb{E}_\mu f \ge x\} = \sup_{f \in \mathcal{F}^*} \mu\{f - \mathbb{E}_\mu f \ge x\}, \qquad x \in \mathbb{R}.$$

Kita teorema galioja visoms izoperimetrinėms TME.

Theorem 16 (Dzindzalieta 2013)**.** *Jei (V, d, μ) yra izoperimetrinė ir $\mathcal{F}$ yra pilna, tada*

$$D(x) = \mu\{f^*_{opt} - \mathbb{E}_\mu f^*_{opt} \ge x\}, \qquad \textit{kai } x \in \mathbb{R},$$

kur $f^*_{opt}(u) = -d(A_{opt}, u)$ yra neigiama atstumo funkcija nuo ekstremalios aibės A_{opt}. Be to, $\mu(A_{opt}) = D(x)$.

6 Darbo mokslinis naujumas ir aktualumas

Dauguma disertacijoje pristatytų rezultatų yra originalūs ir daugiau ar mažiau atitinka šešių mokslinių publikacijų (žr. skyrių „Pagrindinės publikacijos“) turinį. Disertacijoje gauti rezultatai buvo sėkmingai pristatyti tarptautinėse konferencijose ir seminaruose srities specialistams (žr. skyrių „Rezultatų sklaida“).

7 Darbo struktūra ir apimtis

Disertacija parašyta anglų kalba. Disertacija sudaro 7 skyriai: įvadas, 5 moksliniams rezultatams paskirti skyriai, išvados ir literatūros sąrašas. Bendra darbo apimtis yra 78 puslapiai.

8 Pagrindinės publikacijos

Pristatomos disertacijos rezultatai iš 6 mokslinių publikacijų:

[1] V. Bentkus and D. Dzindzalieta. A tight Gaussian bound for weighted sums of Rademacher random variables. *Accepted to Bernoulli*, 2014.

[2] D. Dzindzalieta. On random signs. *Submitted*, 2013.

[3] D. Dzindzalieta and F. Götze. Halfspaces with influential variable. *Submitted*, 2013.

[4] D. Dzindzalieta, T. Juškevičius, and M. Šileikis. Optimal probability inequalities for random walks related to problems in extremal combinatorics. *SIAM Journal on Discrete Mathematics*, 26(2):828–837, 2012.

[5] Dainius Dzindzalieta. Extremal lipschitz functions in the deviation inequalities from the mean. *Electron. Commun. Probab.*, 18:no. 66, 1–5, 2013.

[6] Dainius Dzindzalieta. Random walks maximizing the probability to visit an interval. *arXiv preprint arXiv:1305.6735*, 2013.

9 Rezultatų sklaida

Disertacijoje gauti rezultatai buvo pristatyti šiose mokslinėse konferencijose ir seminaruose:

- Matematikos seminaras. Bielefeldo universitetas, Bielefeldas, Vokietija, 2014.
- Tikimybių teorijos ir statistikos seminaras. Vilniaus universitetas, Matematikos ir Informatikos institutas, Vilnius, 5 kartus per studijų laikotarpį.
- The last two papers by V. Bentkus on inequalities for sums of independent random variables. (su M. Šileikiu). 10 Tarptautinė Vilniaus konferencija, 2010.
- Lietuvos matematikos draugijos konferencija. 2010, 2012 metais.
- Extremal Lipschitz functions on isoperimetric spaces. Matematikos seminaras, Minesotos Matematikos institutas, JAV, lapkričio 29, 2011.
- Extremal Lipschitz functions in the deviation from the mean inequalities. Konferencijoje High Dimensional Phenomena, Minesota, JAV, rugsėjo 26-30, 2011.

10 Stažuotės

- Bielefeldo universitetas, Vokietija Gruodžio 1, 2009 – Gruodžio 17, 2009. Vadovas Friedrich Götze.
- Bielefeldo universitetas, Vokietija Liepos 1, 2011 – Rugpjūčio 31, 2011. Vadovas Friedrich Götze.
- Minesotos Matematikos institutas, JAV Rugsėjo 1, 2011 – Gruodžio 24, 2011. Vadovas Sergey Bobkov.

- Bielefeldo universitetas, Vokietija Rugsėjo 1, 2012 – Lapkričio 30, 2012. Vadovas Friedrich Götze.
- Bielefeldo universitetas, Vokietija Sausio 6, 2014 – Sausio 30, 2014. Vadovas Friedrich Götze.

11 Summary

In our thesis we investigate tight bounds for the tail probabilities for sums of independent or dependent random variables. The results are of theoretic nature and are divided into 5 chapters.

In the second chapter we consider a class $\mathcal{F}$ of sums of independent symmetric random variables. We find the maximizing random variables which gives the largest probability that the sum is in an interval. We also give a simple solution of a famous Littlewood-Offord problem.

In the third chapter we consider a sum of weighted independent Rademacher's random variables with bounded variance. We find an optimal subgaussian constant for the sum and find almost optimal bound for the tail probability for small arguments. We show some applications to Student statistics and Self-normalized sums.

In the fourth chapter we present an application of the results from the third chapter to investigate single coordinate influence of Boolean valued half-space functions on the Boolean cube. We reformulate the problem in probabilistic terms and obtain conditional small ball inequality for the sum of weighted independent Rademacher's random variables. Furthermore we confirm a conjecture by Matulef, O'Donnell, Rubinfeld and Servedio 2010 saying that the threshold function associated to a linear function with some large coefficient, if balanced, must have a large influence.

In the fifth chapter we assume that Bernoulli random variables are only bounded and have a martingale type dependence. We find tight bounds for the probability that a random walk based on a martingale sequence visits an interval. We also show that the maximizing random walk is an inhomogeneous Markov chain. We present a full description of the maximizing random walk and give explicit expression for the maximal probability. We extend the results to random walks based on supermartingale sequences.

In the sixth chapter we find an upper bound $D(x) = \sup_{f\in\mathcal{F}} \mu\{f - \mathbb{E}_\mu f \geq x\}$, for all $x \in \mathbb{R}$, where $\mathcal{F}$ is the complete class of integrable, Lipschitz functions on probability metric (product) spaces. As corollaries we obtain $D(x)$ for Euclidean unit sphere S^{n-1} with a geodesic distance function and a normalized Haar measure, for $\mathbb{R}^n$ equipped with a Gaussian measure and for the multidimensional cube, rectangle, torus or Diamond graph equipped with uniform measure and Hamming distance function. We also prove that in general probability metric spaces extremal Lipschitz functions are from a family of negative distance functions.

The thesis also contains an introduction, conclusions and bibliography. Additionally to the thesis, an extensive summary in Lithuanian is provided. Additionally to the summary in Lithuanian, a summary of the summary is provided in English within the summary in Lithuanian.

12 Trumpos žinios apie autorių

Išsilavinimas

Šiaulių Stasio Šalkauskio vidurinė mokykla (su pagyrimu),

Vilniaus universiteto Magna cum laude statistikos magistras.

Mokslinio darbo patirtis

2005 - Vilniaus universiteto matematikos ir informatikos instituto laborantas, po to specialistas, po to jaunesnysis mokslinis darbuotojas.

Kita veikla

Lietuvos 6-8 klasių jaunųjų matematikų olimpiados užduočių rengėjas ir vertinimo komisijos narys,

Lietuvos 9-12 klasių jaunųjų matematikų olimpiados vertinimo komisijos narys.