TÕENÄOSUSTEOORIA II

Loengukonspekt ja ülesanded

sügis 2019

Jüri Lember

Kirjandus:

1. P. Billingsley, "Probability and measure", Wiley 1995
2. A. Shirjajev, "Verojatnost", Nauka 1989
3. D. Williams, "Probability with martingales", Cambridge 1991
4. G. Grimmet, D. Stirzaker, "Probability and random processes", Oxford 1994, 2001
5. R. Dudley, "Real analysis and probability", Cambridge 2002
6. P. Pfeiffer, "Probability with applications", Springer 1990
7. R. Durrett, "Probality: theory and examples", Duxbury Press 1996
8. K. Chung, "A course in probability theory", AP 1974
9. Y. Chow, H.Teicher "Probability theory: independence, interchangeability, martingales", Springer 2003
10. S. Ross "Introduction to probability models", Elsevier/AP, 2007
11. $\cdots$

Konspekt: ÕIS

Contents

1 Algebra ja σ-algebra, nendega seotud mõisted

1.1 Põhimõisted

Olgu S hulk.

Def 1.1 *Hulga S alamhulkade süsteemi $\mathcal{I}$ nim* π-**süsteemiks** *hulgal S kui*

$$A, B \in \mathcal{I} \Rightarrow A \cap B \in \mathcal{I}.$$

Iga π-süsteem on kinnine lõplike ühisosade võtmise suhtes:

$$A, B, C \in \mathcal{I} \quad \Rightarrow A \cap B \in \mathcal{I} \text{ ja } (A \cap B) \cap C \in \mathcal{I}.$$

Kuid kinnisus loenduva ühisosa suhtes ei järeldu mitte!

Näited:

- $S = \mathbb{R}, \quad \pi(\mathbb{R}) := \{(-\infty, x], x \in \mathbb{R}\}$
- $S = \mathbb{R}, \quad \pi(\mathbb{Q}) := \{(-\infty, x], x \in \mathbb{Q}\}$
- $S = (0, 1], \quad \mathcal{I} := \{(a, b] : 0 \leq a \leq b \leq 1\}$
- $S = \mathbb{R}^k, \quad \pi(\mathbb{R}^k) := \{(-\infty, x_1] \times \cdots \times (-\infty, x_k] : (x_1, \ldots, x_k) \in \mathbb{R}^k\}$.
- S suvaline hulk, $\{S, \emptyset\}$ on π-süsteem.

Def 1.2 *Hulga S alamhulkade süsteemi Σ_0 nim.* **algebraks** *hulgal S kui*

1) $\emptyset, S \in \Sigma_0$

2) $A \in \Sigma_0 \Rightarrow A^c \in \Sigma_0$

3) $A, B \in \Sigma_0 \Rightarrow A \cup B \in \Sigma_0$.

Siin A^c on täiend:

$$A^c := \{s \in S : s \notin A\} = S \backslash A.$$

Märkused:

- Aksioomist **3)** järeldub, et algebra on kinnine lõplike ühendite suhtes:

$$A \cup B \cup C = (A \cup B) \cup C \in \Sigma_0.$$

- Aksioomi **2)** tõttu on algebra kinnine ka lõplike ühisosade suhtes:

$$A \cap B = (A^c \cup B^c)^c \in \Sigma_0.$$

  Kuid kinnisus lõpmatute ühisosade ja ühendite suhtes ei järeldu mitte!

- Algebra on π-süsteem.

- Algebra on kinnine elementide vahe ja sümmeetrilise vahe võtmise suhtes (ülesanne 1).

$$\begin{aligned} &A, B \in \Sigma_0 \Rightarrow A \backslash B \in \Sigma_0; \\ &A, B \in \Sigma_0 \Rightarrow A \triangle B = (A \backslash B) \cup (B \backslash A) \in \Sigma_0. \end{aligned}$$

- Tingimuse **1)** võib asendada tingimustega:

  1a) $\emptyset \in \Sigma_0$; **1b)** $S \in \Sigma_0$; **1c)** $\exists A \in \Sigma_0$.

Näited:

- Olgu S suvaline hulk ning koosnegu Σ_0 lõplikest hulkadest ning nende täienditest (kolõplikud hulgad), st.

$$A \in \Sigma_0 \Leftrightarrow A \text{ või } A^c \text{ on lõplik.}$$

  Klass Σ_0 on algebra (ülesanne 4).

- Olgu $S = (0, 1]$ ja

$$\mathcal{B}_0 = \{(a_1, b_1] \cup (a_2, b_2] \cup \cdots \cup (a_k, b_k], \quad 0 \leq a_1 \leq b_1 \leq a_2 \leq \cdots \leq b_k \leq 1, \quad k \geq 1\}. \tag{1.1}$$

  Klass $\mathcal{B}_0$ on algebra (ülesanne 7).

- Olgu S suvaline hulk; $\{S, \emptyset\}$ on algebra.

Def 1.3 *Hulga S alamhulkade süsteemi Σ nim* σ-**algebraks** *hulgal S kui*

1) $\emptyset, S \in \Sigma$

2) $A \in \Sigma \Rightarrow A^c \in \Sigma$

3) $A_1, A_2, \ldots \in \Sigma \Rightarrow \cup_{i=1}^{\infty} A_i \in \Sigma.$

Märkused:

- Aksioom **3)** tähendab, et σ-algebra on kinnine loenduvate ühendite võtmise suhtes. Kahtlemata on iga σ-algebra ka algebra (ning seetõttu ka π-süsteem).
- Aksioomi **2)** tõttu on **3)** ekvivalentne loenduvate ühisosade võtmise suhtes:

$$A_i \in \Sigma \Rightarrow \cap_i A_i = (\cup_i A_i^c)^c \in \Sigma.$$

Näited:

- Olgu S suvaline hulk. Klass $\{S, \emptyset\}$ on väikseim σ-algebra hulgal S.
- Kõikide alamhulkade hulk 2^S on suurim σ-algebra hulgal S.
- Olgu S suvaline hulk; koosnegu Σ ülimalt loenduvatest hulkades ja nende täienditest (ko-ülimalt loenduvad hulgad), s.t.

$$A \in \Sigma \Leftrightarrow A \text{ või } A^c \text{ on ülimalt loenduv.}$$

  Klass Σ on σ-algebra (ülesanne 5).

Näiteid π-süsteemide, algebrate ja σ-algebrate kohta vaata ülesannetest 4 - 7; σ-algebrate omadusi vaata ülesannetest 8 - 11.

Väide 1.1 *Kui $\{\Sigma_\theta\}$ on (suvalise võimsusega) σ-algebrate hulk, siis $\cap_\theta \Sigma_\theta$ on σ-algebra.*

Tõestus. Ülesanne. ■

Def 1.4 *Paari (S, Σ), kus S on mingi hulk ja Σ sellel antud σ-algebra nim.* **mõõtuvaks ruumiks** *(measurable space).*

1.2 Klassi tekitatud σ-algebra

Olgu $\mathcal{C}$ hulgal S antud alamhulkade klass.

Def 1.5 *Klassi* $\mathcal{C}$ **tekitatud σ-algebra** $\sigma(\mathcal{C})$ *on kõikide hulgal* S *olevate klassi* $\mathcal{C}$ *sisaldavate* σ*-algebrate ühisosa:*

$$\sigma(\mathcal{C}) := \cap\{\Sigma : \Sigma \textit{ on } \sigma\textit{-algebra hulgal } S \textit{ ning } \mathcal{C} \subset \Sigma\}.$$

Definitsioon on korrektne, sest

- hulgal S leidub vähemalt üks klassi $\mathcal{C}$ sisaldav σ-algebra: 2^S;
- väite 1.1 tõttu on suvaline σ-algebrate ühisosa on σ-algebra.

$\sigma(\mathcal{C})$ on (sisalduvusseose mõttes) väikseim σ-algebra (hulgal S), mis sisaldab klassi $\mathcal{C}$.

Definitsioonist järeldub, et $\mathcal{C} \subset \sigma(\mathcal{C})$ ning

$$\underline{\text{kui } \mathcal{C} \subset \Sigma \text{ ja } \Sigma \text{ on } \sigma\text{-algebra, siis } \sigma(\mathcal{C}) \subseteq \Sigma.}$$

1.3 Boreli σ-algebra

Olgu (S, τ) topoloogiline ruum s.t. τ on hulgal S antud topoloogia (lahtised hulgad).

Def 1.6 *Hulgal* S *antud lahtiste hulkade tekitatud* σ*-algebrat* $\sigma(\tau)$ *nimetatakse* **Boreli σ-algebraks**. *Seda tähistatakse* $\mathcal{B}(S)$.

Näited:

- $\mathcal{B}(\mathbb{R})$ (reaaltelje lahtiste hulkade tekitatud) ;
- $\mathcal{B}(\mathbb{R}^k)$ (ruumi $\mathbb{R}^k$ lahtiste hulkade tekitatud);
- $\mathcal{B}((0,1]) =: \mathcal{B}(0,1]$ (hulga $(0,1]$ lahtiste hulkade tekitatud).

Erinevad klassid võivad tekitada sama (Boreli) σ-algebra:

Lemma 1.1 *Järgmised klassid tekitavad $\mathcal{B}(\mathbb{R})$:*

1. $\{(-\infty, a] : a \in \mathbb{R}\} = \pi(\mathbb{R})$,
2. $\{(-\infty, a) : a \in \mathbb{R}\}$,
3. $\{[a, \infty) : a \in \mathbb{R}\}$,
4. $\{(a, \infty) : a \in \mathbb{R}\}$,
5. $\{(a, b] : a, b \in \mathbb{R}, a \leq b\}$,
6. $\{(a, b) : a, b \in \mathbb{R}, a \leq b\}$,
7. $\{[a, b] : a, b \in \mathbb{R}, a \leq b\}$,
8. $\{[a, b) : a, b \in \mathbb{R}, a \leq b\}$.

Tõestus. Tõestame väite 1. Ülejäänud väited tõestatakse analoogiliselt.
Iga reaalarvu a korral kehtib:

$$(-\infty, a] = \cap_{n=1}^{\infty}(-\infty, a + n^{-1}).$$

Seega $(-\infty, a] \in \mathcal{B}(\mathbb{R})$ (miks?), millest järeldub (kuidas?), et

$$\sigma(\pi(\mathbb{R})) \subset \mathcal{B}(\mathbb{R}).$$

Teistpidi. Sellest, et algebra on kinnine vahe võtmise suhtes järeldub, et

$$\{(a, b] |\, a, b \in \mathbb{R}, a \leq b\} \subset \sigma(\pi(\mathbb{R})),$$

millest omakorda järeldub, et

$$\sigma(\{(a, b] |\, a, b \in \mathbb{R}, a \leq b\}) \subset \sigma(\pi(\mathbb{R})).$$

Tõestame nüüd, et

$$\mathcal{B}(\mathbb{R}) \subset \sigma(\{(a, b] |\, a, b \in \mathbb{R}, a \leq b\}). \tag{1.2}$$

Olgu G lahtine. Iga $x \subset G$ korral leiduvad ratsionaalarvud $a_x < b_x$ nii, et $x \in (a_x, b_x] \subset G$. Järelikult $G = \cup_{x \in G}(a_x, b_x]$. Et aga selline ühend võetakse üle loenduva hulga, siis

$$G \in \sigma\Big(\{(a, b] |\, a, b \in \mathbb{R}, a \leq b\}\Big),$$

millest järeldub (1.2). ■

Muuhulgas järeldub lemmast, et kõik hulgad kujul $(b, a], (b, a), [b, a), [b, a]$, kus $b, a \in [-\infty, \infty]$ kuuluvad Boreli σ-algebrasse $\mathcal{B}(\mathbb{R})$, need on *Boreli hulgad.* Kõik "igapäevased" (praktikas ettetulevad) hulgad on Boreli hulgad, on raske (kuid mitte võimatu) defineerida hulka, mis pole Boreli hulk.

Märkused:

- Lemma 1 kehtib ka ruumis $\mathbb{R}^k$: $\sigma(\pi(\mathbb{R}^k)) = \mathcal{B}(\mathbb{R}^k)$ jne.;
- hulga $\mathbb{R}$ asemel võib kasutada ka hulka $\mathbb{Q}$: $\sigma(\pi(\mathbb{Q})) = \mathcal{B}(\mathbb{R})$ jne.

1.4 Boreli σ-algebra ja alamhulgad

Siiani vaatlesime lähemalt Boreli σ-algebrat $\mathcal{B}(\mathbb{R}^k)$. Analoogiliselt saab Boreli σ-algebra defineerida ka suvalisel alamhulgal $B \subset \mathbb{R}^k$. Tuletame meelde, et $A \subset B$ on alamhulga B topoloogias lahtine parajasti siis, kui $A = G \cap B$ ja G on lahtine ruumis $\mathbb{R}^k$. ("Alamhulga topoloogias" tähendab sisuliselt seda, et vaatleme hulka B kogu ruumina. Kogu ruum on alati lahtine. Seega peab B nüüd olema lahtine (isegi kui ta enne seda ei olnud) ja nii saab mõiste "lahtine hulk" hulga B "sees" uue sisu).

Edaspidises on oluline hulga $(0,1]$ lahtiste hulkade tekitatud σ-algebra $\mathcal{B}(0,1]$. Seega

$$\mathcal{B}(0,1] = \sigma\Big(\big\{(0,1] \cap G : G \text{ on lahtine ruumis } \mathbb{R}\big\}\Big). \tag{1.3}$$

Loomulikult on ka see σ-algebra tekitatud mitmete lihtsamate klasside poolt. Nii näiteks kehtib

$$\mathcal{B}(0,1] = \sigma(\{(a,b], a,b \in (0,1]\})$$

(ülesanne 15).

Kuidas on seotud $\mathcal{B}(0,1]$ ja $\mathcal{B}(\mathbb{R})$? Formaalselt on nad eraldi defineeritud. Kas erinevad definitsioonid on ka kooskõlas, s.t. kas σ-algebra $\mathcal{B}(0,1]$ elemendid on ka $\mathcal{B}(\mathbb{R})$ elemendid; kas $A \cap (0,1] \in \mathcal{B}(0,1]$, kus $A \in \mathcal{B}(\mathbb{R})$? Järgnev lemma näitab, et definitsioonid on kooskõlas.

Olgu $\mathcal{C}$ hulgal S antud alamhulkade klass, $B \subset S$. Tähistame

$$\mathcal{C} \cap B := \{A \cap B | A \in \mathcal{C}\}.$$

Näide: Toodud tähistuse abil esitub (1.3)

$$\mathcal{B}(0,1] = \sigma(\tau \cap (0,1]),$$

kus τ on lahtised hulgad reaalteljel.

Seega on $\mathcal{C} \cap B$ hulga B alamhulkade klass ning $\sigma(\mathcal{C} \cap B)$ on σ-algebra hulgal B. Teisest küljest aga on ka $\sigma(\mathcal{C}) \cap B$ on ka hulga B alamhulkade klass.

Lemma 1.2 $\sigma(\mathcal{C}) \cap B = \sigma(\mathcal{C} \cap B)$.

Tõestus. Aksioome kontrollides veendu järgmises: kui Σ on σ-algebra hulgal S ja $B \subset S$ on suvaline alamhulk, siis $\Sigma \cap B$ on σ-algebra hulgal B. Järelda sellest, et $\sigma(\mathcal{C}) \cap B$ on σ-algebra hulgal B, mis sisaldab klassi $\mathcal{C} \cap B$. See tähendab, et

$$\underline{\sigma(\mathcal{C}) \cap B \supseteq \sigma(\mathcal{C} \cap B)}.$$

Tõestame vastupidise seose: kui $A \in \sigma(\mathcal{C})$, siis $A \cap B \in \sigma(\mathcal{C} \cap B)$. Teisisõnu, defineerime hulga

$$\mathcal{G} := \{G \subset S : G \cap B \in \sigma(\mathcal{C} \cap B)\}$$

ja tõestame, et $\sigma(\mathcal{C}) \subset \mathcal{G}$. Kindlasti $\mathcal{C} \subset \mathcal{G}$ (miks?). Aksioome kontrollides veendu, et $\mathcal{G}$ on σ-algebra hulgal S. Seega $\sigma(\mathcal{C}) \subset \mathcal{G}$ ehk

$$\underline{\sigma(\mathcal{C} \cap B) \supseteq \sigma(\mathcal{C}) \cap B.}$$

■

Järeldus 1.1 $\mathcal{B}(0,1] = \{A : A \subseteq (0,1], A \in \mathcal{B}(\mathbb{R})\}$.

Tõestus. Kasuta eelmist lemmat, võttes $B = (0,1]$, $S = \mathbb{R}$, $\mathcal{C} = \tau$ on lahtised hulgad ruumis $\mathbb{R}$. Lemmast järeldub, et

$$\mathcal{B}(0,1] = \sigma(\tau \cap (0,1]) = \sigma(\tau) \cap (0,1] = \mathcal{B}(\mathbb{R}) \cap (0,1] = \{A : A \subseteq (0,1], A \in \mathcal{B}(\mathbb{R})\}.$$

■

1.5 Dynkini π-λ-teoreem

Def 1.7 *Hulga S alamhulkade süsteemi $\mathcal{D}$ nim d-**süsteemiks** hulgal S kui*

1) $S \in \mathcal{D}$

2) $A, B \in \mathcal{D},\ A \subseteq B \Rightarrow B \setminus A \in \mathcal{D}$

3) $A_1, A_2, \ldots \in \mathcal{D}, A_n \nearrow A \Rightarrow A \in \mathcal{D}.$

Tihti defineeritakse d-süsteem teisel kujul:

Def 1.8 *Hulga S alamhulkade süsteemi Λ nim λ-**süsteemiks** hulgal S kui*

1) $S \in \Lambda$

2) $A \in \Lambda \Rightarrow A^c \in \Lambda$

3) $A_1, A_2, \ldots \in \Lambda$ *on lõikumatud* $\Rightarrow \cup_n A_n \in \Lambda.$

λ-süsteem ja d-süsteem on ekvivalentsed mõisted (ülesanne 18).

Teoreem 1.9 Dynkini π-λ teoreem. *Olgu $\mathcal{I}$ hulgal S antud π-süsteem ja Λ hulgal S antud λ-süsteem. Kui $\mathcal{I} \subset \Lambda$, siis $\sigma(\mathcal{I}) \subset \Lambda$.*

(Arusaadavalt võib ülaltoodud teoreemis λ-süsteemi asendada d-süsteemiga. Tõestuse võib leida raamatutest: *Billingsley*, Section 3 või *Williams*, A1).

1.6 Ülesanded

1. Olgu Σ_0 algebra. Tõestada, et
1) $A, B \in \Sigma_0 \Rightarrow A \cap B \in \Sigma_0$ 2) $A, B \in \Sigma_0 \Rightarrow A \setminus B \in \Sigma_0$ 3) $A, B \in \Sigma_0 \Rightarrow A \triangle B \in \Sigma_0$.

2. Tõestada, et hulga S alamhulkade süsteem Σ_0 on algebra parajasti siis, kui
1) $S \in \Sigma_0$.
2) $A, B \in \Sigma_0 \Rightarrow A \setminus B \in \Sigma_0$.

3. Leida $\cup_{n=1}^{\infty} A_n$ ja $\cap_{n=1}^{\infty} B_n$, kus $A_n = [0, \frac{n}{n+1})$ ja $B_n = (0, n^{-1})$.

4. Olgu S lõplik hulk. Koosnegu Σ hulga S kõigist lõplikest ja kolõplikest alamhulkadest. Kas
a) Σ on π-süsteem?
b) Σ on algebra?
c) Σ on σ-algebra?
d) $\Sigma = 2^S$?
Vastata küsimustele a), b), c), d) juhul, kui S on lõpmatu hulk.

5. Olgu S suvaline hulk. Koosnegu Σ hulga S kõigist ülimalt loenduvatest ja ko-ülimalt loenduvatest alamhulkadest. Kas
a) Σ on π-süsteem?
b) Σ on algebra?
c) Σ on σ-algebra?
d) $\Sigma = 2^S$?

6. Olgu S mitteloenduv hulk. Koosnegu Σ hulga S kõigist ülimalt loenduvatest alamhulkadest. Kas
a) Σ on π-süsteem?
b) Σ on algebra?
c) Σ on σ-algebra?
d) $\Sigma = 2^S$?

7. Olgu $S = (0, 1]$, koosnegu alamhulkade hulk $\mathcal{B}_0$ lõikumatute poollõikude lõplikust ühendist, st $B \in \mathcal{B}_0$, kui $B = (a_1, b_1] \cup (a_2, b_2] \cup \cdots \cup (a_k, b_k]$, kus $0 \le a_1 \le b_1 \le a_2 \le \cdots \le b_k \le 1$.
a)kas $\mathcal{B}_0$ on algebra?
b)kas $\mathcal{B}_0$ on σ-algebra?

8. Olgu Σ_1 ja Σ_2 Hulgal S antud σ-algebrad. Kas järgmised hulga S alamhulkade süsteemid on σ-algebrad:
a) $\Sigma_1 \cap \Sigma_2$, b) $\Sigma_1 \cup \Sigma_2$, c) $\Sigma_1 \setminus \Sigma_2$, d) $\Sigma_1 \triangle \Sigma_2$.

9. Olgu $\Sigma_1, \Sigma_2, \ldots$ mittekahanev σ-algebrate jada (st $\Sigma_1 \subseteq \Sigma_2 \subseteq \cdots$). Kas $\cup_n \Sigma_n$ on
1)algebra?
2)σ-algebra?

10. Olgu (S, Σ) mõõtuv ruum, $\{A_1, A_2, \ldots\} \subset \Sigma$ lõikumatute mittetühjade hulkade jada. Tõestada, et Σ on vähemalt kontiiniumi võimsusega hulk.

11. Olgu (S, Σ) mõõtuv ruum. Tõestada, et kui Σ-s on lõpmata palju elemente, siis leidub lõikumatute mittetühjade hulkade jada $\{A_1, A_2, \ldots\} \subset \Sigma$. Järeldada, et ühegi σ-algebra võimsus ei saa olla loenduv.

Näpunäide: Olgu $B_1, B_2, \ldots$ erinevad Σ elemendid st $B_i \neq B_j$ iga $i \neq j$ korral. Defineeri hulgad $\{A_\alpha\}_{\alpha \in (0,1)}$ nii, et $A_\alpha = \cap_j C_j$, kus C_j on kas B_j või B_j^c. Veendu, et hulgad $\{A_\alpha\}$ on lõikumatud Σ elemendid. Lõpuks veendu, et nende seas on lõpmatu palju selliseid hulki, mis pole $\emptyset$. Olgu $\mathcal{A} \subset \{A_\alpha\}$ mittetühjade elementide klass (hulk). Vaja näidata, et $|\mathcal{A}| = \infty$. Selleks veendu, et iga hulk B_j esitub klassi $\mathcal{A}$ kuuluvate elementide ühendina. Et hulki B_j on lõpmata palju ja nad on kõik erinevad, et saa klass $\mathcal{A}$ olla lõplik.

12. Olgu $S = \{1, 2, 3\}$. Leida väikseim σ-algebra, mis sisaldab
a) hulka $\{1\}$
b) hulki $\{3, 1\}$ ja $\{1, 2\}$

13. Olgu $S = (0, 1]$. Leida $\sigma(\mathcal{C})$, kui
a)$\mathcal{C} = \{(0, 2/3], (1/3, 1]\}$
b)$\mathcal{C} = \{(0, 1/2], [1/2, 1]\}$
c)$\mathcal{C} = \{\{1\}\}$
d)$\mathcal{C} = \{[1/3, 1/2]\}$
e)$\mathcal{C} = \{\emptyset\}$
f)$\mathcal{C} = \{(0, 1]\}$
g)$\mathcal{C} = \{$ ratsionaalarvud hulgas $(0, 1]\}$

14. Olgu S mitteloenduv hulk. Kirjeldada σ-algebrat, mis on tekitatud
a)Kõikide ühepunktiliste hulkade poolt;
b)Kõikide loenduvate hulkade poolt;
c)Kõikide mitteloenduvate hulkade poolt;
d)Kõikide lõpmatute hulkade poolt;

15. Olgu $\mathcal{C} = \{(a, b] | 0 \leq a \leq b \leq 1\}$. Tõestada, et $\sigma(\mathcal{C}) = \mathcal{B}(0, 1]$ ja $\sigma(\mathcal{B}_0) = \mathcal{B}(0, 1]$.

16. Tõestada, et klass Σ on σ-algebra parajasti siis, kui ta on nii π-süsteem kui ka d-süsteem.

17. Tõestada, et klass Σ on σ-algebra parajasti siis, kui ta on nii π-süsteem kui ka

λ-süsteem.

18. Tõestada, et d-süsteem ja λ-süsteem on ekvivalentsed mõisted.

19. Olgu $\mathcal{A}$ selline hulga S alamhulkade mittetühi klass, et $\forall A \in \mathcal{A}$ korral A^c on klassi $\mathcal{A}$ elementide loenduv ühend. Näita, et $\sigma(\mathcal{A})$ langeb kokku väikseima klassi $\mathcal{A}$ sisaldava loenduva ühendi ja loenduva ühisosa võtmise suhtes kinnise klassiga.
Näpunäide: Olgu $\mathcal{K}(\mathcal{A})$ väikseim klassi $\mathcal{A}$ sisaldav loenduva ühendi ja loenduva ühisosa võtmise suhtes kinnise klass. Veendu, et $S \in \mathcal{K}(\mathcal{A})$ ja klass $\mathcal{B} = \{B \in \mathcal{K}(\mathcal{A}) | B^c \in \mathcal{K}(\mathcal{A})\} = \mathcal{K}(\mathcal{A})$.

20. Olgu $\mathcal{C}$ hulga S alamhulkadest moodustatud klass. Näidata, et kui $B \in \sigma(\mathcal{C})$, siis $\exists$ ülimalt loenduv alamklass $\mathcal{C}_B$ nii, et $B \in \sigma(\mathcal{C}_B)$.
Näpunäide: Veendu, et klass $\mathcal{B} = \{B \in \sigma(\mathcal{C}) | \exists$ ülimalt loenduv alamklass $\mathcal{C}_B$ nii, et $B \in \sigma(\mathcal{C}_B)\}$ on σ-algebra ja sisaldab esialgset klassi $\mathcal{C}$.

2 Mõõt, mõõdu ühesus ja jätkamine

2.1 Põhimõisted

Olgu Σ_0 hulgal S antud algebra.

Def 2.1 *Algebral* Σ_0 *antud funktsiooni* $\mu : \Sigma_0 \to [0, \infty]$ *nimetatakse* **mõõduks**, *kui ta on loenduvalt aditiivne:*

1) $\mu(\emptyset) = 0$

2) *Kui* $A_1,\, A_2, \ldots \in \Sigma_0,\ A_i \cap A_j = \emptyset\, \forall i \neq j$ *ja kui* $\cup_i A_i \in \Sigma_0$, *siis*

$$\mu(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i).$$

- Mõõtu μ nim. **lõplikuks**, kui $\mu(S) < \infty$.
- Mõõtu μ nim. **lõpmatuks**, kui $\mu(S) = \infty$.
- Mõõtu μ nim. σ-**lõplikuks**, kui leidub jada $A_1, A_2, \ldots \in \Sigma_0$ nii, et

  $$S = \cup A_i \text{ ja } \mu(A_i) < \infty\ \forall i = 1, 2, \ldots.$$

  Üldisust kitsendamata võime eeldada, et $A_1 \subseteq A_2 \subseteq A_3 \subseteq \cdots$ st $A_n \nearrow S$.
- Olgu $\mathcal{A} \subset \Sigma_0$ mingi alamklass. Mõõtu μ nim. σ-**lõplikuks klassil** $\mathcal{A}$, kui leidub jada $A_1, A_2, \ldots \in \mathcal{A}$ nii, et $A_n \nearrow S$ ja $\mu(A_i) < \infty\ \forall i = 1, 2, \ldots$. Seega σ-lõplik mõõt on σ-lõplik algebral Σ_0.
- Mõõtu μ nim. **tõenäosusmõõduks**, kui $\mu(S) = 1$. Tõenäosusmõõtu tähistatakse harilikult tähega $\mathbf{P}$.

Märkus (lõplik ja loenduv aditiivsus): Mõõt on algebral antud loenduvalt aditiivne mittenegatiivne (sest väärtused kuuluvad hulka $[0, \infty]$) funktsionaal. Öeldakse, et μ on (lõplikult) aditiivne, kui aksioom **2)** on asendatud nõudega: iga lõpliku mittelõikuvate Σ_0-elementide komplekti $A_1, \ldots, A_k$ korral (st $A_i \in \Sigma_0$ ja $A_i \cap A_j = \emptyset$, kui $i \neq j$) kehtib

$$\mu(\bigcup_{i=1}^{k} A_i) = \sum_{i=1}^{k} \mu(A_i).$$

Loenduvast aditiivsusest järeldub lõplik aditiivsus (kuidas?), kuid mitte vastupidi. Veendu, et $\mu :\ \Sigma_0 \to [0, \infty]$ on lõplikult aditiivne parajasti siis, kui ta rahuldab nõudeid:

1) $\mu(\emptyset) = 0$

2) kui $A, B \in \Sigma_0$ ja $A \cap B = \emptyset$, siis $\mu(A \cup B) = \mu(A) + \mu(B)$.

Algebral defineeritud aditiivset mittenegatiivset funktsionaali nimetatakse teinekord ka lõplikult aditiivseks mõõduks (mitte ajada segamini lõpliku mõõduga!).

Def 2.2 *Olgu Σ hulgal S antud σ-algebra, s.o. (S, Σ) olgu mõõtuv ruum, μ olgu σ-algebral Σ antud mõõt. Kolmikut (S, Σ, μ) nimetatakse* **mõõdu ruumiks** *(measure space). Σ elemente nimetatakse* **mõõtuvateks** *(alam)hulkadeks.*

Juhul, kui μ on tõenäosusmõõt, nimetatakse mõõdu ruumi **tõenäosusruumiks** ja tähistatakse harilikult $(\Omega, \mathcal{F}, \mathbf{P})$. $\mathcal{F}$ elemente nimetatakse **sündmusteks** (events).

Näited:

- Diraci mõõt. Olgu $x \in S$. Funktsioon

$$\delta_x(F) = \begin{cases} 1, & \text{kui } x \in F \\ 0, & \text{kui } x \notin F \end{cases}$$

on mõõt σ-algebral 2^S (ülesanne 8).

- Diskreetne e. atomaarne mõõt. Olgu $\{p_n\}$ positiivsed arvud $\{x_n\} \subset S$, $x_i \neq x_j$ kui $i \neq j$. Siis

$$\nu = \sum_{i=1}^{\infty} p_i \delta_{x_i} \tag{2.1}$$

mõõt σ-algebral 2^S (ülesanne 8).

Mõõt ν on σ-lõplik;
kui $\sum_i p_i < \infty$, on ν lõplik.

Elemente x_n nimetatakse mõõdu ν **aatomiteks**, arv p_i on aatomi x_i **mass**.

- Loendav mõõt. Olgu $(S, \Sigma) = (S, 2^S)$. Funktsioon

$$\mu : \Sigma \to [0, \infty], \quad \mu(A) = \begin{cases} |A|, & \text{kui } |A| < \infty \\ \infty, & \text{kui } A \text{ lõpmatu} \end{cases}$$

on mõõt;
μ on lõplik $\Leftrightarrow |S| < \infty$;
μ on σ-lõplik $\Leftrightarrow S$ ülimalt loenduv (ülesanne 9);
kui S on mitteloenduv, pole μ isegi σ-lõplik.

- Diskreetne tõenäosus. Olgu $\Omega = \{\omega_1, \omega_2, \dots\}$ loenduv elementaarsündmuste ruum, $\mathcal{F} = 2^\Omega$. Vaatleme funktsiooni $p : \Omega \mapsto [0, \infty]$ ning defineerime

$$\mu(A) := \sum_{\omega \in A} p(\omega), \quad A \subset \Omega. \tag{2.2}$$

Funktsioon μ on diskreetne mõõt, sest (veendu selles!)

$$\mu = \sum_{i=1}^{\infty} p_i \delta_{\omega_i}, \quad \text{kus } p_i := p(\omega_i).$$

Paneme tähele: et p_i võib olla ka 0, on mõõdu μ aatomid vaid need elementaarsündmused ω_i, mille korral $p(\omega_i) > 0$.

Mõõt μ on lõplik, kui $\sum_{\omega\in\Omega} p(\omega) < \infty$ ja μ on tõenäosusmõõt, kui $\sum_{\omega\in\Omega} p(\omega) = 1$. Seega iga lõplikul elementaarsündmuste ruumil $\Omega = \{\omega_1, \ldots, \omega_N\}$ defineeritud tõenäosusmõõt esitub vektorina $(p_1, \ldots, p_N)$, kus $\sum_i p_i = 1$.

2.2 Mõõdu omadused

Olgu (S, Σ, μ) mõõdu ruum.

1) μ on **monotoone**, s.o. $\mu(A) \le \mu(B)$, kui $A \subset B$, $A, B \in \Sigma$.

2) Kui $A, B \in \Sigma$, $A \subset B$ ja $\mu(B) < \infty$, siis $\mu(B \setminus A) = \mu(B) - \mu(A)$.

3) Kui μ on *lõplik* mõõt, siis

$$\mu(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} \mu(A_i) - \sum_{i<j}^{n} \mu(A_i \cap A_j) + \cdots + (-1)^{n-1} \mu(A_1 \cap \cdots \cap A_n).$$

4) **Loenduv subaditiivsus**: Kui $A_1, A_2, \ldots \in \Sigma$, siis

$$\mu(\bigcup_{i=1}^{\infty} A_i) \le \sum_{i=1}^{\infty} \mu(A_i).$$

5) **Pidevus alt**: Kui $A_1, A_2, \ldots \in \Sigma$ ja $A_i \nearrow A$, siis $\mu(A_i) \nearrow \mu(A)$.

6) **Pidevus ülalt**: Kui $A_1, A_2, \ldots \in \Sigma$, $\mu(A_k) < \infty$ mingi k korral ja $A_i \searrow A$, siis $\mu(A_i) \searrow \mu(A)$.

7) Olgu μ σ-lõplik. Olgu $\{A_\theta\}_{\theta\in\Theta}$ lõikumatute mõõtuvate hulkade kollektsioon, kusjuures $\mu(A_\theta) > 0$ iga θ korral. Siis Θ võimsus on ülimalt loenduv.

Tõestus. Omaduste **1)**, **2)** tõestus on ülesanne 2.

Omaduse **3)** tõestus on induktsiooniga n järgi:
$n = 2$: Tõestame

$$\mu(A \cup B) = \mu(A) + \mu(B) - \mu(A \cap B). \tag{2.3}$$

Selleks paneme tähele: $A\backslash B \cup (A \cap B) = A$ ja $B\backslash A \cup (B \cap A) = B$, millest

$$\mu(A) + \mu(B) = \mu(A\backslash B) + 2\mu(B \cap A) + \mu(B\backslash A). \tag{2.4}$$

Teisest küljest,

$$A \cup B = A \backslash B \cup (A \cap B) \cup (B \backslash A).$$

Seega $\mu(A \cap B) + \mu(A \cup B)$ võrdub seose (2.4) parema poolega. See tähendab, et

$$\mu(A \cap B) + \mu(A \cup B) = \mu(A) + \mu(B),$$

millest järeldub (2.3).

$n > 2:$

$$\mu(\cup_{i=1}^n A_i) = \mu(\cup_{i=1}^{n-1} A_i) + \mu(A_n) - \mu((\cup_{i=1}^{n-1} A_i) \cap A_n)) = \mu(\cup_{i=1}^{n-1} A_i) + \mu(A_n) - \mu(\cup_{i=1}^{n-1} (A_i \cap A_n))$$
$$= \sum_{i=1}^{n} \mu(A_i) - \sum_{i<j}^{n-1} \mu(A_i \cap A_j) - \sum_{i=1}^{n-1} \mu(A_i \cap A_n) + \cdots .$$

Omaduste **4)** ja **5)** tõestus on ülesanne 2.
(Näpunäide: defineerige hulgad

$$B_1 := A_1, \quad B_2 = A_2 \backslash A_1, \quad B_3 := A_3 \backslash (A_1 \cup A_2), \ldots .$$

Hulgad $B_1, B_2, \ldots$ on lõikumatud, $\cup_n B_n = \cup_n A_n$).

Omaduse **6)** tõestus on ülesanne 2.

Omaduse **7)** tõestame lõpliku μ korral. Olgu $\{A_\theta\}_{\theta \in \Theta}$ lõikumatute positiivse mõõduga hulkade kollektsioon. Defineerime

$$\Theta_m := \Big\{\theta : \mu(A_\theta) \geq \frac{1}{m}\Big\}.$$

Olgu $\Theta'_m \subset \Theta_m$ mingi ülimalt loenduv alamhulk. Sellisel juhul

$$\mu(S) \geq \mu\big(\cup_{\theta \in \Theta'_m} A_\theta\big) = \sum_{\theta \in \Theta'_m} \mu(A_\theta) \geq |\Theta'_m| \cdot \frac{1}{m}.$$

Seega $|\Theta'_m| \leq m\mu(S) < \infty$. Järelikult ka $|\Theta_m| \leq m\mu(S)$, sest vastasel juhul leiduks mingi alamhulk $|\Theta'_m|$ mille võimsus oleks rangelt suurem kui $m\mu(S)$. Nüüd veendu, et $\Theta = \cup_{m=1}^{\infty} \Theta_m$, st Θ on ülimalt loenduv.
Olgu nüüd μ σ-lõplik, st $S = \cup_n S_n$, kusjuures $\mu(S_n) < \infty$. Valime hulga S_n ja vaatleme lõikeid $\{B_\theta^n\}$, kus

$$B_\theta^n := S_n \cap A_\theta.$$

Fikseeritud n korral kasutame täpselt sama argumenti, mis lõpliku μ korralgi ja näeme, et $\{B_\theta^n\}$ on ülimalt loenduv. Sellest järeldub, et hulk $\cup_n \{B_\theta^n\}$ on ka ülimalt loenduv. Et iga kahe indeksi θ ja θ' korral on hulgad A_θ ja $A_{\theta'}$ lõikumatud, on seda ka iga n ja m korral B_θ^n ja $B_{\theta'}^m$. Et iga hulk A_θ on $\cup_n \{B_\theta^n\}$ elementide – lõigete – ühend ning et erinevatele hulkadele vastavad erinevad lõiked, saamegi, et ka $\{A_\theta\}$ on ülimalt loenduv hulk. ■

Märkus: Omadused **1), 2), 3)** kehtivad ka lõplikult aditiivse μ korral.

Järeldus 2.1 *Olgu* Σ_o *algebra,* $\mu : \Sigma_o \mapsto [0, \infty]$*. Siis*

1. μ *on loenduvalt additiivne* $\Leftrightarrow$ μ *on additiivne ja alt pidev;*

2. *kui* μ *on lõplik, siis* μ *on loenduvalt additiivne* $\Leftrightarrow$ μ *on additiivne ja ülalt pidev.*

Tõestus.

1. $\Rightarrow$: Omadus **5)**. $\Leftarrow$: Ülesanne.

2. $\Rightarrow$: Omadus **6)**. $\Leftarrow$: Ülesanne.

■

Tähelepanu: Kui Σ_0 on vaid algebra kuid mitte σ-algebra, siis kehtivad ülaltoodud omadused vaid siis, kui kõnealused loenduvad ühendid ja ühisosad algebrasse kuuluvad.

Järeldus 2.2 *Olgu* Σ_0 *algebra hulgal* S*. Olgu* $\mu : \Sigma_0 \to [0, \infty)$ *aditiivne lõplik funktsioon. Siis on järgmised omadused ekvivalentsed:*

1 μ *on loenduvalt aditiivne;*

2 μ *on ülalt pidev;*

3 *iga jada* $\Sigma_0 \ni A_n \searrow \emptyset$ *korral* $\lim_n \mu(A_n) = 0$*;*

4 *iga jada* $\Sigma_0 \ni A_n \searrow A$ *kus* $A \in \Sigma_0$ *ja* $\exists \epsilon > 0 : \mu(A_n) > \epsilon$, $\forall n$ *korral kehtib* $A \neq \emptyset$*.*

Tõestus.
1 $\Leftrightarrow$ **2** : Järeldus 2.1.
2 $\Leftrightarrow$ **3** :

$$A_n \searrow A \in \Sigma_0 \quad \Leftrightarrow \quad A_n \backslash A \searrow \emptyset.$$

2 $\Rightarrow$ **4** : Olgu $\Sigma_0 \ni A_n \searrow A$, kusjuures $\mu(A_n) > \epsilon > 0$. Mõõdu ülalt pidevuse tõttu

$$\mu(A) = \lim_n \mu(A_n) \geq \epsilon > 0,$$

millest

$$A = \cap_n A_n \neq \emptyset.$$

4 $\Rightarrow$ **3** : Olgu $\Sigma_0 \ni A_n \searrow \emptyset$. Oletame, et $\mu(A_n)$ ei koondu nulliks, s.t. $\exists \epsilon$ nii, et $\mu(A_n) > \epsilon > 0$ iga n korral. Seosest **4** järeldub

$$\cap A_n \neq \emptyset,$$

mis on aga vastuolus eeldusega, et $\Sigma_0 \ni A_n \searrow \emptyset$. Seega $\mu(A_n) \to 0$. ■

2.3 Mõõdu jätkamine ja ühesus

Olgu mingil σ-algebral Σ antud kaks mõõtu. Kontrollimaks, et need mõõdud on võrdsed, peaksime definitsioonist lähtudes veenduma, et mõõtude väärtused on võrdsed igal Σ elemendil. Mittetriviaalsete σ-algebrate (näiteks Boreli σ-algebra) korral on selle tingimuse kontrollimine pahatihti väga keeruline. Järgneb teoreem väidab aga, et kaks lõplikku mõõtu on võrdsed siis, kui nad langevad kokku mingil Σ tekitaval π-süsteemil. Selle tingimuse kontrollimine on tihti oluliselt lihtsam.

Teoreem 2.3 *Olgu $\mathcal{I}$ hulgal S antud π-süsteem, $\Sigma = \sigma(\mathcal{I})$. Kui μ_1 ja μ_2 on mõõtuval ruumil (S, Σ) antud sellised mõõdud, et $\mu_1(S) = \mu_2(S) < \infty$ ja $\mu_1(I) = \mu_2(I)$ $\forall I \in \mathcal{I}$, siis $\mu_1 = \mu_2$ (st. $\mu_1(A) = \mu_2(A)$ $\forall A \in \Sigma$).*

Tõestus. Tõestus põhineb Dynkini $\pi - \lambda$-teoreemil. Defineerime hulga

$$\Lambda := \{A \in \sigma(\mathcal{I}) : \mu_1(A) = \mu_2(A)\}, \quad \mathcal{I} \subseteq \Lambda \subseteq \sigma(\mathcal{I}).$$

Veendu, et Λ on λ-süsteem. Dynkini $\pi - \lambda$-teoreemist järeldub, et $\sigma(\mathcal{I}) \subset \Lambda$. ■

Järeldus 2.3 *Kui tõenäosusmõõdud $\mathbf{P}_1$ ja $\mathbf{P}_2$ langevad kokku π-süsteemil $\mathcal{I}$, siis nad langevad kokku ka σ-algebral $\sigma(\mathcal{I})$. Teisisõnu, iga σ-algebral $\sigma(\mathcal{I})$ antud tõenäosusmõõt on üheselt määratud π-süsteemiga $\mathcal{I}$.*

Näited:

- $S = \mathbb{R}$, $\Sigma = \mathcal{B}(\mathbb{R}), \mathcal{I} = \pi(\mathbb{R})$. Olgu $\mathbf{P}_1$ ja $\mathbf{P}_2$ σ-algebral $\mathcal{B}(\mathbb{R})$ antud tõenäosusmõõdud. Kui $\mathbf{P}_1(-\infty, x] = \mathbf{P}_2(-\infty, x]$, $\forall x \in \mathbb{R}$, siis $\mathbf{P}_1 = \mathbf{P}_2$.
- Oletame, et σ-algebral $\mathcal{B}(0,1]$ leidub mõõt μ nii, et $\mu(a,b] = b - a$ iga $(a,b] \subset (0,1]$ korral. Kui selline mõõt leidub, siis ta on ühene.

Teoreemi 2.3 võib laiendada ka σ-lõplike mõõtudeni.

Teoreem 2.4 *Olgu $\mathcal{I}$ hulgal S antud π-süsteem ning μ_1 ja μ_2 olgu σ-algebral $\sigma(\mathcal{I})$ antud mõõdud. Kui μ_1 ja μ_2 on π-süsteemil $\mathcal{I}$ σ-lõplikud ja $\mu_1(I) = \mu_2(I)$ $\forall I \in \mathcal{I}$, siis $\mu_1 = \mu_2$.*

Tõestus. Iga $B \in \mathcal{I}$ korral on $\Sigma \cap B$ σ-algebra hulgal B. Et $\mathcal{I} \subset \Sigma$, kuulub $B \in \Sigma$, millest

$$\Sigma \cap B = \{A \in \Sigma : A \subseteq B\}.$$

Seega on korrektselt defineeritud mõõdu μ_i ahend σ-algebrale $\Sigma \cap B$. Kui $\mu_i(B) < \infty$, on nimetatud ahend lõplik mõõt σ-algebral $\Sigma \cap B$. Et $\sigma(\mathcal{I}) \cap B = \sigma(\mathcal{I} \cap B)$, on $\mathcal{I} \cap B$ σ-algebrat $\Sigma \cap B$ tekitav π-süsteem. Et aga $\mu_1(A) = \mu_2(A)$ iga $A \in \mathcal{I}$ korral ning $A \cap B \in \mathcal{I}$ (miks?), kehtib ka, et

$$\mu_1(A) = \mu_2(A) \quad \forall A \in \mathcal{I} \cap B,$$

s.t. need mõõdud on võrdsed π-süsteemil $\mathcal{I} \cap B$. Et aga $\mu_1(B) < \infty$, siis järeldub teoreemist 2.3, et need mõõdud on võrdsed σ-algebral $\Sigma \cap B$.

Et μ_1 ja μ_2 on lõplikud π-süsteemil $\mathcal{I}$, leiduvad hulgad $I_n \in \mathcal{I}$ nii, et $I_n \nearrow S$ ja $\mu_1(I_n) = \mu_2(I_n)$ iga n korral. Seega on μ_1 ja μ_2 σ-algebral $\Sigma \cap I_n$ võrdsed:

$$\mu_1(A \cap I_n) = \mu_1(A \cap I_n), \quad \forall A \in \Sigma, \quad \forall n.$$

Pidevus alt: $\forall A \in \Sigma$, $\mu_1(A) = \lim_{n\to\infty} \mu_1(A \cap I_n) = \lim_{n\to\infty} \mu_2(A \cap I_n) = \mu_2(A)$. ■

Märkus: Teoreemist 2.4 järeldub teoreem 2.3.

Näide: Olgu μ σ-algebral $\mathcal{B}(\mathbb{R})$ antud mõõt, mis rahuldab $\mu(a, b] = b - a$. Et $\mathcal{B}(\mathbb{R}) = \sigma(\mathcal{I})$, kus $\mathcal{I} = \{(a, b] : a \leq b \in \mathbb{R}\}$, siis vastavalt teoreemile 2.4 on μ üheselt määratud.

Kas teoreemi 2.4 saab laiendada ka mõõtudeni, mis pole σ-lõplikud π-süsteemil $\mathcal{I}$?

Kontranäide: Vaatame ruumil $((0, 1], \mathcal{B}(0, 1])$ antud mõõte μ_1 ja μ_2, kus

$$\mu_1(F) = \begin{cases} 0 & \text{kui } F = \emptyset \\ \infty & \text{muidu.} \end{cases}, \quad \mu_2(F) = \begin{cases} |F| & \text{kui } |F| < \infty \\ \infty, & \text{muidu} \end{cases}$$

Need mõõdud on võrdsed algebral $\mathcal{B}_o(0, 1]$ aga mitte σ-algebral $\mathcal{B}(0, 1]$.

Teoreem 2.5 (Caratheodory teoreem) *Olgu Σ_0 hulgal S antud algebra, $\Sigma = \sigma(\Sigma_0)$. Kui μ_0 on algebral Σ_0 antud mõõt, siis mõõtuval ruumil (S, Σ) leidub mõõt μ nii, et $\mu = \mu_0$ algebral Σ_0, st. $\mu(A) = \mu_0(A)$ $\forall A \in \Sigma_0$.*

(Tõestuse võib leida raamatutest *Billingsley*, Thm 11.2; *Williams* A1).

Mõõtu μ nimetatakse mõõdu μ_o laiendiks. Charatheodory teoreemist järeldub, et iga algebral antud mõõtu saab laiendada selle algebra tekitatud σ-algebrale. Kui $\mu_0(S) < \infty$, siis vastavalt teoreemile 2.3 on saadud laiend ühene. Kui μ_0 on algebral Σ_0 σ-lõplik, siis vastavalt teoreemile 2.4 on laiend ühene.

2.4 Lebesgue'i mõõt

2.4.1 Lebesgue'i mõõt σ-algebral $\mathcal{B}(0, 1]$

Vaatleme hulka $(0, 1]$ ja sellel antud algebrat $\mathcal{B}_o$. Seega iga hulga $\mathcal{B}_o$ element avaldub

$$(a_1, b_1] \cup (a_2, b_2] \cup \cdots \cup (a_k, b_k], \text{ kus } b_i < a_{i+1}.$$

Defineerime funktsiooni (pikkus)

$$\mu_0 : \mathcal{B}_o \mapsto [0, 1], \quad \mu_o((a_1, b_1] \cup (a_2, b_2] \cup \cdots \cup (a_k, b_k]) = \sum_{i=1}^{k} (b_i - a_i).$$

Lemma 2.1 *Funktsioon μ_o on mõõt algebral $\mathcal{B}_o$.*

Tõestus. On selge, et μ on mittenegatiivne ja

$$\mu_o(\emptyset) = \mu_o(a, a] = a - a = 0.$$

Loenduv aditiivsus: kui $A_k \in \mathcal{B}_o$, $k = 1, 2, \dots$ on lõikumatud ja $\cup_k A_k \in \mathcal{B}_o$ (eeldus!), siis peab kehtima

$$\mu_o(\cup_k A_k) = \sum_k \mu_o(A_k).$$

On kerge veenduda, et μ_o on lõplikult aditiivne, s.t. iga kahe lõikumatu hulga $A, B \in \mathcal{B}_o$ korral kehtib

$$\mu_o(A \cup B) = \mu_o(A) + \mu_o(B).$$

Edaspidi kasutame järgmist omadust: olgu $A \in \mathcal{B}_o$, st A on kujul $(a_1, b_1] \cup \dots \cup (a_k, b_k]$. Et

$$\mu_o\big((x + \frac{1}{n}, y]\big) \nearrow \mu_o\big((x, y]\big),$$

siis iga $\delta > 0$ korral leidub hulk B, mis rahuldab järgmisi omadusi:

1. $B = (c_1, d_1] \cup \dots \cup (c_k, d_k]$, st $B \in \mathcal{B}_0$;

2. $c_i > a_i$ iga $i = 1, \dots, k$ korral;

3. $\mu_o(A \backslash B) \leq \delta$.

Omadustest 1. ja 2. järeldub, et hulga B sulundil $\bar{B}$ on järgmised omadused:

a) $\bar{B} = [c_1, d_1] \cup \dots \cup [c_k, d_k]$;

b) $\bar{B} \subset A$.

Järeldusest 2.2 (seos **4**) tõttu on μ_0 mõõt, kui tõestame, et järgmine implikatsioon kehtib:

$$A_n \in \mathcal{B}_o, \quad A_n \searrow A := \cap_n A_n \in \mathcal{B}_o, \quad \mu_o(A_n) \geq \epsilon > 0 \Rightarrow A \neq \emptyset. \tag{2.5}$$

Olgu $\{A_n\}$ selline jada. Ülaltoodust teame, et iga i korral leidub hulk $B_i \in \mathcal{B}_o$ nii, et kehtib **a)** ning

$$B_i \subset \bar{B}_i \subset A_i, \quad \mu_o(A_i \backslash B_i) \leq \frac{\epsilon}{2^{i+1}}.$$

Hulgad $C_n = \cap_{i=1}^n B_i$ on sellised, et ($C_i \subset B_i \subset A_i$ ja $A_n \subset A_i$ iga $i \leq n$ korral)

$$C_n \subset \cap_{i=1}^n \bar{B}_i \subset A_n, \quad \text{ja } A_n \backslash C_n = \cup_{i=1}^n (A_n \backslash B_i) \subset \cup_{i=1}^n (A_i \backslash B_i).$$

Et μ_o on aditiivne ning seetõttu monotoone ja lõplikult subaditiivne, saame

$$\mu_o(A_n \backslash C_n) \leq \sum_{i=1}^n \mu_o(A_i \backslash B_i) \leq \frac{\epsilon}{2} \quad \forall n.$$

Seega

$$\mu_o(C_n) = \mu_o(A_n) - \mu_o(A_n \backslash C_n) > \frac{\epsilon}{2}, \quad \forall n,$$

millest

$$\cap_{i=1}^n \bar{B}_i \supset C_n \neq \emptyset \quad \forall n. \tag{2.6}$$

Oletades vv.-lt, et $\cap_n A_n = \emptyset$, saame, et $\cap_{i=1}^\infty \bar{B}_i = \emptyset$. Defineerime iga i korral

$$D_i = [0,1] \backslash \bar{B}_i.$$

Et $\cap_{i=1}^\infty \bar{B}_i = \emptyset$, siis $\cup_{i=1}^\infty D_i = [0,1]$. Seosest **a)** teame, et D_i on lahtine ruumis $\mathbb{R}$. Seega hulgad D_i moodustavad kompaktse hulga $[0,1]$ lahtise katte. Heine-Boreli lemmast järeldub nüüd, et

$$\exists n : \cup_{i=1}^\infty D_i = \cup_{i=1}^n D_i = [0,1].$$

Et $D_i = [0,1] \backslash \bar{B}_i$, siis

$$\cup_{i=1}^n D_i = [0,1] \quad \Leftrightarrow \quad \cap_{i=1}^n \bar{B}_n = \emptyset,$$

mis on vastuolus seosega (2.6). ∎

Vastavalt Charatheodory teoreemile saame mõõtu jätkata Boreli σ-algebrale $\mathcal{B}(0,1]$. Saadud laiend on ühene (miks?), seda nimetatakse **Lebesgue**'i mõõduks ja tähistatakse tihti kas Leb või λ. Lebesgue'i mõõt formaliseerib lõigu pikkuse: see on ainus mõõt, mis igale intervallile seab vastavusse tema pikkuse.

Paneme tähele, et $((0,1], \mathcal{B}(0,1], \mathrm{Leb})$ on tõenäosusruum.

2.4.2 Lebesgue'i mõõt σ-algebral $\mathcal{B}(\mathbb{R})$

Defineerisime Lebesgue'i mõõdu σ-algebral $\mathcal{B}(0,1]$. Intervall $(0,1]$ pole mõõdu eksisteerimise ja ühesuse seisukohalt eriline, analoogiliselt saab Lebesgue'i mõõdu üheselt defineerida σ-algebral $\mathcal{B}(a,b]$, kus a ja b on lõplikud.

Defineerime Lebesgue'i mõõdu σ-algebral $\mathcal{B}(\mathbb{R})$ (seega reaalteljel). Olgu $(a_n, b_n]$ lõplikud ja lõikumatud lõigud nii, et $\mathbb{R} = \cup_n (a_n, b_n]$. (Kas sellised poollõigud leiduvad?) Igal σ-algebral $\mathcal{B}(a_n, b_n]$ defineerime Lebesgue'i mõõdu Leb_n. Defineerime funktsiooni

$$\mu : \mathcal{B}(\mathbb{R}) \mapsto [0, \infty], \quad \mu(A) = \sum_n \mathrm{Leb}_n(A \cap (a_n, b_n]). \tag{2.7}$$

Selline funktsioon on mõõt (ülesanne 11), mis igale reaaltelje poollõigule seab vastavusse tema pikkuse, st $\mu(c,d] = d - c$. Teame, et sellist tingimust rahuldav mõõt on ühene (miks?). Seega definitsioon on korrektne, s.o. definitsioon ei sõltu hulkade $(a_n, b_n]$ valikust. Seosega (2.7) defineeritud mõõtu nimetatakse (jälle) Lebesgue'i mõõduks (reaalteljel). Lebesgue'i mõõt reaalteljel on kõige tähtsam σ-lõplik mõõt.

Formaalselt oleme defineerinud mitu Lebesgue'i mõõtu. Need definitsioonid on kooskõlas, sest iga $a, b \in \mathbb{R}$ korral $\mathcal{B}(a,b] = \{A \in \mathcal{B}(\mathbb{R}) : A \subseteq (a,b]\}$ (miks?), mistõttu σ-algebral $\mathcal{B}(\mathbb{R})$ defineeritud Lebesgue'i mõõdul on ahend σ-algebrale $\mathcal{B}(a,b]$. Vastavalt teoreemile 2.3 ühtib see ahend σ-algebral $\mathcal{B}(a,b]$ defineeritud Lebesgue'i mõõduga.

2.5 Tõenäosusmõõdud reaalteljel

σ-algebral $\mathcal{B}(\mathbb{R})$ antud mõõtu nimetame mõõduks reaalteljel (näiteks Lebesgue'i mõõt reaalteljel).

2.5.1 Jaotusfunktsioon

Olgu $\mathbf{P}$ σ-algebral $\mathcal{B}(\mathbb{R})$ antud tõenäosusmõõt. Defineerime funktsiooni

$$F : \mathbb{R} \mapsto [0,1], \quad F(x) = \mathbf{P}(-\infty, x]. \tag{2.8}$$

Def 2.6 *Funktsiooni F nimetatakse mõõdu* **P jaotusfunktsiooniks**

Jaotusfunktsiooni omadused

1) F on mittekahanev (mõõdu monotoonsus)

2) $\lim_{x\to\infty} F(x) = 1$, $\lim_{x\to-\infty} F(x) = 0$ (mõõdu pidevus ülalt ja alt)

3) F on paremalt pidev:

$$\lim_{x_n \searrow x} F(x_n) = F(x), \quad \forall x$$

ja omab vasakpoolset piirväärtust igas punktis:

$$\exists F(x-) = \lim_{x_n \nearrow x} F(x_n) \quad \forall x.$$

(Paremalt pidevus. Olgu $x_n \searrow x$, defineeri $A_n := (-\infty, x_n]$, $A_n \searrow (-\infty, x]$, mõõdu pidevus ülalt: $F(x_n) = \mathbf{P}(A_n) \searrow \mathbf{P}(-\infty, x] = F(x)$.

Vasakpoolse piirväärtuse olemasolu: $x_n \nearrow x$, defineeri $A_n := (-\infty, x_n]$, $A_n \nearrow (-\infty, x)$, mõõdu pidevus alt: $F(x_n) = \mathbf{P}(A_n) \nearrow \mathbf{P}(-\infty, x) =: F(x-)$)

4) $F(x) - F(x-) = \mathbf{P}\{x\}$.

Omadusest **4)** järeldub, et punktis x on F pidev parajasti siis, kui $\mathbf{P}\{x\} = 0$. Kui $\mathbf{P}\{x\} > 0$, nimetame punkti x mõõdu **P aatom**iks. Mõõdu omadusest **7)** järelduvalt on tõenäosusmõõdul ülimalt loenduv hulk aatomeid. Kui mõõdu **P** jaotusfunktsioon on pidev, siis pole sel mõõdul ühtegi aatomit.

Järeldusest 2.3 järeldub, et σ-algebral $\mathcal{B}(\mathbb{R})$ ei saa olla teist mõõtu, millel oleks sama jaotusfunktsioon. Seega erinevatele mõõtudele vastavad erinevad jaotusfunktsioonid. Järgnev teoreem väidab sisuliselt, et igale jaotusfunktsiooni omadustega funktsioonile vastab üks tõenäosusmõõt.

Teoreem 2.7 *Olgu $F : \mathbb{R} \mapsto [0,1]$ tingimusi* **1)**, **2)**, **3)** *rahuldav funktsioon. Siis leidub σ-algebral $\mathcal{B}(\mathbb{R})$ parajasti üks tõenäosusmõõt* $\mathbf{P}$ *nii, et F on* $\mathbf{P}$ *jaotusfunktsioon.*

Teoreemi 2.7 saab tõestada otse mõõdu definitsiooni kasutades (vt. *Shirjajev*, pt. II, §3, T1). Käesolevas kursuses tõestame teoreemi 2.7 juhuslikke suurusi kasutades edaspidi.

Seega on σ-algebral $\mathcal{B}(\mathbb{R})$ antud tõenäosusmõõtude ja tingimusi **1)**, **2)**, **3)** rahuldavate funktsioonide vahel üks-ühene vastavus.

Teoreemist 2.7 järeldub Lebesgue'i mõõdu olemasolu σ-algebral $\mathcal{B}(0,1]$. Tõepoolest, vaatame jaotusfunktsiooni

$$F(x) = \begin{cases} 0, & \text{kui } x < 0 \\ x & \text{kui } 0 \leq x \leq 1 \\ 1 & \text{kui } x \geq 1 \end{cases}$$

Selline funktsioon rahuldab tingimusi **1)**, **2)** ja **3)**. Seega σ-algebral $\mathcal{B}(\mathbb{R})$ eksisteerib tõenäosusmõõt $\mathbf{P}$ jaotusfunktsiooniga (2.8). Iga intervalli $(a,b] \subset \mathbb{R}$ korral

$$\mathbf{P}(a,b] = F(b) - F(a)$$

(miks?). Kui $(a,b] \subset (0,1]$, siis $\mathbf{P}(a,b] = b - a$. Sellest järeldub, et $\mathbf{P}$ ahend σ-algebrale $\mathcal{B}(0,1]$ on Lebesgue'i mõõt σ-algebral $\mathcal{B}(0,1]$ (kas $\mathbf{P}$ on Lebesgue'i mõõt σ-algebral $\mathcal{B}(\mathbb{R})$?).

Analoogiliselt (kuidas?) võib jaotusfunktsiooni defineerida iga lõpliku mõõdu korral. Teatud tingimustel võib jaotusfunktsiooni analoogi defineerida ka σ-lõplike (kuid mitte ilmtingimata lõplike) mõõtude korral. See käib nii. Olgu μ σ-algebral $\mathcal{B}(\mathbb{R})$ antud mõõt, mis rahuldab tingimust: $\mu(a,b] < \infty, \quad \forall a,b \in \mathbb{R}$. Sellist mõõtu nimetatakse ka *Lebesgue'i-Stieltjesi* mõõduks. Igale Lebesgue'i-Stieltjesi mõõdule μ saab vastavusse seada funktsiooni $G : \mathbb{R} \mapsto \mathbb{R}$,

$$G(x) = \begin{cases} \mu(0,x], & \text{kui } x \geq 0 \\ -\mu(x,0], & \text{kui } x \leq 0 \end{cases}.$$

Funktsioon G on mittekahanev, paremalt pidev, $G(0) = 0$ ja rahuldab seost

$$\mu(a,b] = G(b) - G(a), \quad \forall a,b \in \mathbb{R}. \tag{2.9}$$

On selge (miks?), et ei leidu funktsiooni G, mis rahuldaks seost (2.9) kahe erineva mõõdu korral (vastupidine ei kehti: kui F rahuldab seost (2.9), siis teeb seda ka $F + c$, kuid sellega asi piirdubki). Küll aga kehtib teoreemi 2.7 üldistus.

Teoreem 2.8 *Olgu $G : \mathbb{R} \mapsto \mathbb{R}$ mittekahanev, paremalt pidev funktsioon. Siis leidub parajasti üks mõõt σ-algebral $\mathcal{B}(\mathbb{R})$, mis rahuldab (2.9).*

Kui lisada tingimus $G(0) = 0$, siis paremelt pidevate ja mittekahanevate funktsioonide ning Lebesgue'i -Stieltjesi mõõtude vahel on üks-ühene seos.

Teoreemist 2.8 Järeldub Lebesgue'i mõõdu olemasolu - võta $G(x) = x$.

2.5.2 Diskreetsed, absoluutselt pidevad ja singulaarsed tõenäosusmõõdud

Tõenäosusmõõdud σ-algebral $\mathcal{B}(\mathbb{R})$ on üks-üheses vastavuses jaotusfunktsioonidega. Kirjeldame olulisi jaotusfunktsioonide (tõenäosusmõõtude) klasse.

Diskreetsed mõõdud

Diskreetne tõenäosusmõõt on kontsentreeritud aatomitele. See tähendab, et leiduvad punktid (aatomid) $x_1, x_2, \ldots$ nii, et $\mathbf{P}\{x_i\} > 0$ ja $\sum_{i=1}^{\infty} \mathbf{P}\{x_i\} = 1$. Diskreetse mõõdu jaotusfunktsioon (diskreetne jaotusfunktsioon) teeb punktis x_i hüppe pikkusega

$$\Delta F(x_i) := F(x_i) - F(x_i-) = \mathbf{P}\{x_i\}.$$

Seega $\sum_{i=1}^{\infty} \Delta F(x_i) = 1$ ehk F on tükati konstantne.

Iga diskreetne mõõt on kirjeldatav aatomitega $\{x_i\}$ ja massidega $\{p_i\}$, kus $p_i = \mathbf{P}\{x_i\}$.

Näited:

- Binoomjaotus $B(n, \theta)$, $\mathbf{P}\{k\} = C_n^k \theta^k (1-\theta)^{n-k}$, $k = 0, 1, 2, \ldots, n$;
- Poissoni jaotus $P_o(\lambda)$, $\mathbf{P}\{k\} = \frac{e^{-\lambda}\lambda^k}{k!}$, $k = 0, 1, 2, \ldots$.

Järgnev lemma väidab, et iga jaotusfunktsioon esitub mingi diskreetse jaotusfunktsiooni ja pideva jaotusfunktsiooni kaalutud keskmisena.

Lemma 2.2 *Olgu F jaotusfunktsioon. Siis leidub diskreetne jaotusfunktsioon F_d, pidev jaotusfunktsioon F_c ja konstant $\lambda \in [0,1]$ nii, et*

$$F = \lambda F_d + (1-\lambda) F_c. \tag{2.10}$$

Tõestus. Olgu F suvaline jaotusfunktsioon. Olgu $x_1, x_2, \ldots$ funktsiooni F katkevuspunktide hulk (aatomid). Olgu $p_i := \Delta F(x_i)$ aatomi x_i mass. Olgu $\lambda := \sum_i p_i$. On selge, et $\lambda \in [0,1]$ (miks?). Kui $\lambda = 0$, on F pidev ja lemma tõestatud. Olgu $\lambda > 0$. Defineerime diskreetse tõenäosusmõõdu

$$\sum_i \frac{1}{\lambda} p_i \delta_{x_i}.$$

Tema jaotusfunktsioon on
$$F_d(x) = \frac{1}{\lambda}\sum_{x_i \le x} p_i.$$
Defineerime funktsiooni F_c
$$F_c(x) = \frac{F(x) - \lambda F_d(x)}{1-\lambda}. \tag{2.11}$$
Et
$$\lambda F_d(x) = \sum_{x_i \le x} p_i = \sum_{x_i \le x}(F(x_i) - F(x_i-)),$$
siis funktsiooni $F - \lambda F_d$ saame jaotusfunktsioonist F katkevuspunktide mahalahutamisel. Seega on $F - \lambda F_d$ pidev, mittekahanev, kusjuures $\lim_{x\to\infty} F(x) - \lambda F_d(x) = 1 - \lambda$ ja $\lim_{x\to-\infty} F(x) - \lambda F_d(x) = 0$. Seega rahuldab F_c tingimusi **1) 2)** ja **3)**. Teisisõnu, F_c on pidev jaotusfunktsioon. Seos (2.10) järeldub F_c definitsioonist (2.11). ■

Ülesanne:

Olgu
$$F(x) = \begin{cases} 0.5e^x & \text{kui } x < 0, \\ 0.7 & \text{kui } x \in [0,2), \\ 0.4x & \text{kui } x \in [2, 2.5]. \end{cases} \tag{2.12}$$
Esitada F kujul (2.10).

Absoluutselt pidevad mõõdud

Absoluutselt pideva mõõdu jaotusfunktsioon (absoluutselt pidev jaotusfunktsioon) esitub kujul
$$F(x) = \int_{-\infty}^{x} f(t)dt, \tag{2.13}$$
kus f on mittenegatiivne **tihedusfunktsioon**[1].

Näited:

- Normaaljaotus $N(\mu, \sigma^2)$;
- Ühtlane jaotus $U[a, b]$;
- Eksponentjaotus $E(\lambda)$.

Jaotusfunktsiooni F nimetatakse **singulaarseks** kui leidub hulk $N \in \mathcal{B}(\mathbb{R})$ nii, et $\text{Leb}(N) = 0$ ja $F'(x) = 0$ iga $x \in N^c$. Seega on singulaarse jaotusfunktsiooni tuletis 0 igas reaaltelje punktis välja arvatud punktihulk, mille Lebesgue' i mõõt on 0. Nendes punktides tuletist

[1] Definitsioonis (2.13) olev integraal on üldiselt Lebesgue'i integraal, millega tutvume hiljem

ei leidu. On selge, et diskreetne jaotusfunktsioon on singulaarne (miks?).

Järgnev lemma väidab, et iga jaotusfunktsioon esitub mingi absoluutselt pideva jaotusfunktsiooni ja singulaarse jaotusfunktsiooni kaalutud keskmisena.

Lemma 2.3 *Olgu F jaotusfunktsioon. Siis leidub absoluutselt pidev jaotusfunktsioon F_a, singulaarne jaotusfunktsioon F_s ja konstant $\gamma \in [0,1]$ nii, et*

$$F = \gamma F_a + (1-\gamma)F_s. \tag{2.14}$$

Seost (2.14) nimetatakse **Lebesgue'i dekompositsiooniks** (lahutuseks).

Ülesanne: Leida jaotusfunktsiooni (4.10) Lebesgue'i dekompositsioon.

Järeldus 2.4 *Olgu F jaotusfunktsioon. Siis leidub absoluutselt pidev jaotusfunktsioon F_a, diskreetne jaotusfunktsioon F_d, ning pidev kuid singulaarne jaotusfunktsioon F_s nii, et*

$$F = \alpha_1 F_d + \alpha_2 F_a + \alpha_3 F_s, \tag{2.15}$$

kus $\alpha_i \in [0,1]$ ja $\alpha_1 + \alpha_2 + \alpha_3 = 1$.

Tõestus. Olgu F jaotusfunktsioon. Lemmast 2.2 saame, et $F = \lambda F_d + (1-\lambda)F_c$, kus F_c on pidev jaotusfunktsioon. Rakendades jaotusfunktsioonile F_c Lebesgue'i dekompositsiooni, saame $F_c = \gamma F_a + (1-\gamma)F_s$, kus F_s on singulaarne. Et F_c ja F_a on pidevad funktsioonid (miks?), saame, et F_s on pidev ja singulaarne jaotusfunktsioon. Võttes $\alpha_1 = \lambda, \alpha_2 = (1-\lambda)\gamma, \alpha_3 = (1-\lambda)(1-\gamma)$, saame lahutuse (2.15). ∎

Seega saab iga jaotusfunktsiooni (tõenäosusmõõdu) lahutada diskreetseks, absoluutselt pidevaks ja pidev-singulaarseks komponendiks. Neist viimane on üks kummaline funktsioon: ta on pidev, mittekahanev (teinekord isegi rangelt kasvav), kuid tema tuletis on peaaegu kõikjal 0 (välja arvatud teatud 0-pikkusega hulk). Klassikaline näide singulaarsest funktsioonist on nn. Cantori funktsioon. Analüütiliselt selliseid funktsioone esitada ei saa ja praktikas seda komponenti enamasti ette ei tule (st tavaliselt $\alpha_3 = 0$).

2.6 Mõõdud ruumil $\mathbb{R}^k$

Olgu $\mathbf{P}$ tõenäosusmõõt σ-algebral $\mathcal{B}(\mathbb{R}^k)$.

Def 2.9 *Mõõdu $\mathbf{P}$ jaotusfunktsioon F on*

$$F : \mathbb{R}^k \mapsto [0,1], \quad F(x_1, \ldots, x_k) = \mathbf{P}\big((-\infty, x_1] \times \cdots \times (-\infty, x_k]\big). \tag{2.16}$$

Et $\sigma(\pi(\mathbb{R}^k)) = \mathcal{B}(\mathbb{R}^k)$, siis kahel erineval mõõdul ei saa olla ühte ja sama jaotusfunktsiooni.

Iga kaks punkti $a = (a_1, \ldots, a_k), b = (b_1, \ldots, b_k) \in \mathbb{R}^k$ nii, et $a_i \leq b_i$ defineerivad (tõkestatud) ristküliku:

$$A = \{(x_1, \ldots, x_k) \in \mathbb{R}^k : a_i < x_i \leq b_i\}. \tag{2.17}$$

Ristkülikul A on 2^k tippu: punktid $(c_1,\ldots,c_k)$, kus $c_i \in \{a_i, b_i\}$. Olgu $V(A)$ tippude hulk, iga tipu $c = (c_1,\ldots,c_k)$ korral olgu

$$\operatorname{sign}_A(c) = \begin{cases} 1 & \text{kui } (c_1,\ldots,c_k) \text{ sisaldab paarisarv } a_i\text{-si,} \\ -1 & \text{mujal.} \end{cases}$$

Iga ristküliku korral defineerime

$$\Delta_A F := \sum_{c \in V(A)} \operatorname{sign}_A(c) F(c).$$

Seega, kui $k = 1$, siis

$$\Delta_A F = F(b) - F(a);$$

kui $k = 2$, siis

$$\Delta_A F = F(b_1, b_2) - F(b_1, a_2) - F(a_1, b_2) + F(a_1, a_2)$$

jne. On selge, et iga ristküliku korral

$$\Delta_A F = \mathbf{P}(A). \tag{2.18}$$

Tähistame: $x^n \searrow x \in \mathbb{R}^k$ parajasti siis, kui $x_i^n \searrow x_i$ iga $i = 1, \ldots k$.

Jaotusfunktsiooni omadused

1) Iga ristküliku A korral

$$\Delta_A F \geq 0.$$

2) $\lim_{x\to\infty} F(x) = 1$, $\lim_{x \searrow y} F(x) = 0$, kui vähemalt üks kordinaat $y_1, \ldots, y_k$ on $-\infty$.

3) F on ülalt pidev:

$$\lim_{x^n \searrow x} F(x^n) = F(x), \quad \forall x.$$

Teoreem 2.7 üldistub ruumile $\mathbb{R}^k$:

Teoreem 2.10 *Olgu $F : \mathbb{R}^k \mapsto [0,1]$ tingimusi* **1)**, **2)**, **3)** *rahuldav funktsioon. Siis leidub σ-algebral $\mathcal{B}(\mathbb{R})$ parajasti üks tõenäosusmõõt* $\mathbf{P}$ *nii, et F on* $\mathbf{P}$ *jaotusfunktsioon.*

Tõestuse võib leida (*Shirjajev* II, §3).

Analoogiliselt võime defineerida lõpliku mõõdu jaotusfunktsiooni ning tõestada üks-ühese vastavuse lõplike mõõtude ja jaotusfunktsioonide (tingimusi **1)**, **2)**, **3)** rahuldavate funktsioonide) vahel.

Lebesgue'i-Stieltjesi mõõt σ-algebral $\mathcal{B}(\mathbb{R}^k)$ on iga selline mõõt μ, mis suvalisele ristkülikule seab vastavusse lõpliku arvu, s.t. $\mu(A) < \infty$ iga ristküliku A korral. Kehtib teoreemi 2.8 üldistus

Teoreem 2.11 *Olgu $F : \mathbb{R}^k \to \mathbb{R}$ ülalt pidev funktsioon, kusjuures iga ristküliku A korral kehtib $\Delta_A F \geq 0$. Siis leidub σ-algebral $\mathcal{B}(\mathbb{R}^k)$ parajasti üks mõõt μ nii, et iga risküliku A korral kehtib $\Delta_A F = \mu(A)$*

Mõõdu ühesus järeldub teoreemist 2.4, teoreemi olulisus seisneb selles, et seose (2.18) abil kõikide ristkülikute hulgal defineeritud hulgafunktsiooni võib laiendada Lebesgue-Stieltjes mõõduni. Teoreemi 2.11 tõestuse võib leida näiteks (*Billingsley*, Chapter 12, Thm.12.5).

2.6.1 Lebesgue'i mõõt σ-algebral $\mathcal{B}(\mathbb{R}^k)$

Vaatleme funktsiooni

$$F : \mathbb{R}^k \to \mathbb{R}, \quad F(x_1, \ldots, x_k) = x_1 \cdots x_k.$$

Selline funktsioon on ülalt pidev, kusjuures iga ristküliku (2.17) korral

$$\Delta_A F = (b_1 - a_1) \cdots (b_k - a_k) \geq 0.$$

Seega on teoreemi 2.11 eeldused täidetud ning leidub parajasti üks mõõt λ_k nii, et iga ristküliku korral

$$\lambda_k(A) = (b_1 - a_1) \cdots (b_k - a_k).$$

Seega seab λ_k igale ristkülikule vastavusse tema ruumala. Mõõtu λ_k nimetatakse (k-dimensionaalseks) **Lebesgue'i** mõõduks σ-algebral $\mathcal{B}(\mathbb{R}^k)$. Kui k on fikseeritud, siis tähistame ka k-dimensionaalset Lebesgue'i mõõtu lihtsalt Leb.

Lebesgue'i mõõt on nihke-invariantne:

Lemma 2.4 *Kui $A \in \mathcal{B}(\mathbb{R}^k)$, siis $A+x = \{a+x, a \in A\} \in \mathcal{B}(\mathbb{R}^k)$ ja $\mathrm{Leb}(A) = \mathrm{Leb}(A+x)$ $\forall x \ \in \mathbb{R}^k$.*

Tõestus. Olgu $x \in \mathbb{R}^k$ fikseeritud. Defineerime klassi (sõltub x-st)

$$\mathcal{C} = \{A \in \mathcal{B}(\mathbb{R}^k) : A + x \in \mathcal{B}(\mathbb{R}^k)\}.$$

Definitsioonist järeldub, et $\mathcal{C} \subseteq \mathcal{B}(\mathbb{R}^k)$. Veendu, et $\mathcal{C}$ on σ-algebra. Et $\mathcal{C}$ sisaldab kõiki ristkülikuid ja viimased on π-süsteem, mis tekitab $\mathcal{B}(\mathbb{R}^k)$, siis $\mathcal{B}(\mathbb{R}^k) = \mathcal{C}$ (miks?). Seega Boreli σ-algebra on nihke-invariantne, st iga Boreli hulga nihutamisel same ikka Boreli hulga.
Olgu endiselt $x \in \mathbb{R}$ fikseeritud ja defineerime Boreli σ-algebral $\mathcal{B}(\mathbb{R}^k)$ mõõdu $\mu(A) := \mathrm{Leb}(x + A)$ (selleks, et seda mõõtu defineerida, oligi meil eelnevalt vaja näidata, et $x + A$ on ka Boreli hulk. Kui mõne Boreli hulga A korral see nii poleks, poleks meil ka arvu $\mathrm{Leb}(x + A)$). Et μ ja Leb langevad kokku Boreli σ-algebrat tekitaval π-süsteemil (millisel?) ning on sellel π-süsteemil ka σ-lõplikud, siis teoreemi 2.4 põhjal $\mu = \mathrm{Leb}$. ■

Ütleme, et mõõt *normeerib ühikkuubi*, kui ta seab ühikkuubile vastavusse arvu 1. Saab näidata (vt ülesanne 17), et Lebesgue'i mõõt on ainus nihke-invariantne mõõt σ-algebral $\mathcal{B}(\mathbb{R}^k)$, mis normeerib ühikkuubi.

2.7 Tõenäosusruumi täielikustamine

Tihti on kasulik tõenäosusruumi $(\Omega, \mathcal{F}, \mathbf{P})$ laiendada nii, et kõikide 0-mõõduliste hulkade alamhulgad oleksid mõõtuvad. See toimub nii. Olgu $\mathcal{N} \subset 2^\Omega$ defineeritud järgmiselt

$$\mathcal{N} := \{N \in 2^\Omega : \quad \exists F \in \mathcal{F} \quad \text{nii, et} \quad N \subseteq F \quad \text{ja} \quad \mathbf{P}(F) = 0\}.$$

Seega kuuluvad hulka $\mathcal{N}$ kõik 0-mõõduliste hulkade alamhulgad.
Defineerime

$$\Sigma := \{F \in 2^\Omega : \quad \exists A, B \in \mathcal{F} \quad \text{nii, et} \quad A \subseteq F \subseteq B \quad \text{ja} \quad \mathbf{P}(B \backslash A) = 0\}.$$

Veendu, et Σ on σ-algebra ning $\Sigma = \sigma(\mathcal{N}, \mathcal{F})$. Seega Σ on väikseim σ-algebra, mis sisaldab $\mathcal{F}$ ja $\mathcal{N}$. Vaatleme nüüd ruumi (Ω, Σ) ja defineerime mõõdu

$$\mathbf{P}^* : \Sigma \mapsto [0,1], \quad \mathbf{P}^*(F) = \mathbf{P}(A) = \mathbf{P}(B).$$

Veendume, et $\mathbf{P}^*$ ei sõltu hulkade A ja B valikust, s.t. kui leiduvad A_1, A_2, B_1, B_2 nii, et

$$A_1 \subset F \subset B_1, \quad A_2 \subset F \subset B_2,$$

siis $\mathbf{P}(A_1) = \mathbf{P}(A_2)$. Tõepoolest, $A_i \subset B := B_1 \cap B_2, \quad i = 1,2$ ning

$$\mathbf{P}(B \backslash A_i) \leq \mathbf{P}(B_i \backslash A_i) = 0, \quad i = 1,2.$$

Seega

$$\mathbf{P}(A_1) = \mathbf{P}(B \backslash A_1) + \mathbf{P}(A_1) = \mathbf{P}(B) = \mathbf{P}(B \backslash A_2) + \mathbf{P}(A_2) = \mathbf{P}(A_2).$$

On kerge veenduda, et $\mathbf{P}^*$ on loenduvalt aditiivne. Seega on $\mathbf{P}^*$ korrektselt defineeritud mõõt.

Kokkuvõttes on ruum $(\Omega, \Sigma, \mathbf{P}^*)$ ruumi $(\Omega, \mathcal{F}, \mathbf{P})$ selline laiend, et iga null mõõdulise ($\mathbf{P}^*$ järgi) hulga alamhulk on mõõtuv (miks?). Ruumi $(\Omega, \Sigma, \mathbf{P}^*)$ nim. ruumi $(\Omega, \mathcal{F}, \mathbf{P})$ **täieldiks**.

Olgu $(\Omega, \mathcal{F}, \mathbf{P}) = ((0,1], \mathcal{B}(0,1], \text{Leb})$. Selle ruumi täield $((0,1], \Sigma, \text{Leb}^*)$ laiendab Lebesgue'i mõõtu nii, et iga 0-mõõdulise hulga alamhulk on mõõtuv. Mõõtu Leb^* nimetatakse endiselt Lebesgue'i mõõduks ja σ-algebrasse Σ kuuluvaid hulki nimetatakse **Lebesgue'i hulkadeks** või ka Lebesgue'i mõõtuvateks hulkadeks.

2.8 Miks σ-algebra? ehk mittemõõtuvad hulgad*

Lebesgue'i mõõt on pikkuse loomulik formulatsioon: $\text{Leb}((a,b]) = b - a$. Kehtib ka (ülesanne 4)

$$\text{Leb}([a,b]) = \text{Leb}([a,b)) = \text{Leb}((a,b)) = b - a.$$

Teame, et Lebesgue'i mõõt on ainus nihke-invariantne mõõt Boreli σ-algebral $\mathcal{B}(\mathbb{R})$ mis normeerib ühikkuubi. See kõik ühtib suurepäraselt intuitsiooniga. Paraku on Lebesgue'i

mõõt defineeritud vaid Boreli σ-algebral (laiendatult Lebesgue'i hulkadel). Miks ei võiks aga pikkus olla defineeritud igal reaaltelje alamhulgal? Teisisõnu: kas reaaltelje kõikide alamhulkade hulgal $2^{\mathbb{R}}$ saab defineerida mõõdu (loenduvalt aditiivse hulgafunktsiooni) λ mis on nihke-invariantne ja normeerib ühikkuubi? Selle küsimuse püstitas H. Lebesgue oma doktoritöös aastal 1902.
Vastus on eitav.

Kontranäide: (Vitali, 1905; Vitali konstruktsioon). Mõõdu λ nihke-invariantsus formaliseerub järgmiselt: iga A korral kehtib

$$\lambda(T_x A) = \lambda(A), \quad \forall x \in \mathbb{R},$$

kus

$$T_x : \mathbb{R} \to \mathbb{R}, \quad T_x(a) = x + a.$$

Kui λ on nihke-invariantne, siis iga $A \subset (0, 1]$ korral kehtib

$$\lambda(T_x^1 A) = \lambda(A), \quad \forall x \in \mathbb{R},$$

kus

$$T_x^1 : (0, 1] \to (0, 1], \quad T_x^1(a) = (a + x) \mod 1.$$

Vaatleme hulka $(0, 1]$ ühikringina tasandil (isomorfism $\alpha \mapsto (\sin(\alpha 2\pi), \cos(\alpha 2\pi))$). Funktsioon T_x^1 samastub siis pöördega. Seega: kui leidub nihke-invarantne ühikintervalli normaliseeriv mõõt kõikide reaaltelje alamhulkade hulgal, siis tasandi ühikringi S^1 kõikide alamhulkade hulgal on defineeritud pöörde-invariantne mõõt, kusjuures ringi mõõt 1. Veendume, et ühikringi kõikide alamhulkade hulgal pole võimalik defineeride pöörde-invariantset lõplikku mõõtu.
Selleks defineerime hulgal S^1 ekvivalentsisuhte järgmiselt $a \sim b$ parajasti siis kui a on saadud b-st pöörde $q2\pi$ abil, kus $q \in \mathbb{Q}$. (Intervallil $(0, 1]$ on see ekvivalentsiseos lihtsalt $(0, 1]/\mathbb{Q}$, s.t. $\alpha \sim \beta$, kui $\alpha - \beta \in \mathbb{Q}$.) Valikuaksioomi kasutades valime igast ekvivalentsiklassist ühe elemendi. Olgu A nende elementide hulk. Olgu $\{\rho_i\}$ pöörded $q2\pi$, kus $q \in \mathbb{Q} \cap (0, 1]$. Selliseid pöördeid on loenduv hulk. Hulgad $\rho_i(A)$ ja $\rho_j(A)$ on lõikumatud, kui $i \neq j$. Ekvivalentsiklasside definitsioonist järeldub, et $\cup_{i \in \mathbb{N}} \rho_i(A) = S^1$. Et λ on pöörde-invariantne, siis $\lambda(\rho_i(A)) = \lambda(A)$, millest

$$\lambda(S^1) = \sum_{i=1}^{\infty} \lambda(\rho_i(A)) = \sum_{i=1}^{\infty} \lambda(A) \quad !?$$

Seega hulgal A ei saa olla Lebesgue'i mõõtu. Et Lebesgue'i mõõt on korrektselt σ-algebral $\mathcal{B}(0, 1]$, siis $A \notin \mathcal{B}(0, 1]$.

Toodud näitest järeldub, et tihti pole mõõtu võimalik defineerida kõikide alamhulkade hulgal. Sellisel juhul defineeritakse mõõt σ-algebral, mille moodustavad hulgad, millistel on võimalik (vastuollu sattumata) meid huvitav mõõt defineerida. Lebesgue'i mõõtu on võimalik defineerida Boreli σ-algebral. Õnneks on viimane piisavalt rikas n.n. igapäevase matemaatika tarbeks.

Vaatleme veelkord ülaltoodud näidet. Olgu $A_i := \rho_i(A)$, $i \in \mathbb{N}$. Veendusime, et $S^1 = \cup_i A_i$. On selge, et suvalise j, i korral leidub pööre ρ_{ji} nii, et $\rho_{ji}(A_j) = \rho_{ji}(\rho_j(A)) = \rho_i(A) = A_i$. Et paarisarvude ja täisarvude hulga vahel on üks-ühene vastavus $\pi : 2\mathbb{N} \to \mathbb{N}$, siis iga $j \in 2\mathbb{N}$ korral leidub pööre $\rho_j := \rho_{j\pi(j)}$ nii, et

$$S^1 = \cup_{i\in\mathbb{N}} A_i = \cup_{j\in 2\mathbb{N}} \rho_j(A_j).$$

Seega leidub hulga S^1 alamhulk $\cup_{j\in 2\mathbb{N}} A_j$, mille saab lahutada loenduvaks arvuks lõikumatuteks hulkadeks nii, et igaühte neist pöörates saame kokku S^1. Sama argument kehtib ka paarituarvuliste hulkade $A_i, i \in \mathbb{N}\backslash 2\mathbb{N}$ korral:

$$S^1 = \cup_{j\in\mathbb{N}\backslash 2\mathbb{N}} \rho_j(A_j).$$

Teisisõnu: hulk S^1 sisaldab kahte lõikumatut alamhulka

$$\cup_{j\in 2\mathbb{N}} A_j \text{ ja } \cup_{j\in\mathbb{N}\backslash 2\mathbb{N}} A_j$$

nii, et kummalgi neist on järgmine omadus: leidub hulga loenduv lahutus nii, et lahutusse kuuluvaid alamhulki sobivalt pöörates saame hulga S^1 lahutuse. Minnes tagasi intervallile $(0, 1]$, võime saadud tulemuse esitada järbmiselt: intervalli $(0, 1]$ saab lahutada loenduvaks arvuks lõikumatuteks alamhulkadeks nii, et neist igaühte sobivalt nihutades saame hulga $(0, 2]$! Loomulikult välistab selline tulemus kõikide reaaltelje alamhulkade hulgal oleva ühikkera normeeriva nihke-invariantse mõõdu (pikkus) olemasolu, sest vastasel korral saaksime, et intervallidel $(0, 1]$ ja $(0, 2]$ on sama pikkus.
Ülaltoodud paradoks kehtib märksa üldisemalt. Bijektsiooni $T : \mathbb{R}^k \to \mathbb{R}^k$ nimetame *liigutamiseks*, kui T on isomeetria, s.t., $\|T(x) - T(y)\| = \|x - y\|$. Ruumil $\mathbb{R}^1$ on liigutamine sisuliselt nihe ja pööramine; ruumil $\mathbb{R}^2$ on iga liigutamine lahutatav kaheks osaks: pööramine (ümber originaali) või peegeldamine (originaali läbivast) sirgest ja nihe. Kui $k \geq 3$, on võimalike liigutuste rühm märksa laiem.

Teoreem 2.12 (Banach, Tarski, 1924) *Olgu A, B kaks suvalist mittetühja sisemusega ruumi $\mathbb{R}^k$ alamhulka. Leidub hulga A loenduv lahutus $A = \cup_i A_i$, $A_i \cap A_j = \emptyset$ $(i \neq j)$ ja liigutused T_i nii, et $T_i(A_i)$ on hulga B lahutus, s.t. $B = \cup_i T(A_i)$, $T(A_j) \cap T(A_i) = \emptyset$ $(i \neq j)$.*

Seega pole võimalik defineerida liigutamis-invariantset mõõtu (pindala, ruumala jne) ruumi $\mathbb{R}^k$ kõikide alamhulkade hulgal. See on võimalik Boreli σ-algebral, millest järeldub, et hulgad A_i ülaltoodud teoreemis ei ole Boreli (ega Lebesgue'i) hulgad. Kui A ja B on Boreli hulgad ja neil on sama mõõt, siis leidub ülaltoodud lahutus nii, et hulgad A_i on ka boreli hulgad. See aga ei vii vastuoluni.

Vitali konstruktsioon tõestab, et σ-algebral $2^{(0,1]}$ ei leidu nihke-invariantset tõenäosusmõõtu, sest nihke-invariantus viib vastuoluni. Kas aga hulgal $2^{(0,1]}$ üleüldse leidub mõnda tõenäosusmõõtu? Teame, et selliseid mõõte on: näiteks Diraci mõõt ja diskreetne tõenäosusmõõt. Need on aga diskreetsed mõõdud, s.t. neil on vähemalt üks aatom, s.t. leidub ühepunktiline hulk $\{x\}$ millel on positiivne mõõt. Nihke-invarantse mõõdu korral peab iga ühepunktilise hulga mõõt olema 0 (miks?). Selgub, et juba see nõue on liiast.

Teoreem 2.13 (Banach, Kuratowski, 1929) *Hulgal $2^{(0,1]}$ ei leidu ühtegi tõenäosusmõõtu P nii, et $\mathbf{P}(\{x\}) = 0 \ \forall x \in (0,1]$.*

Seega ei saa kõikide alamhulkade hulgal defineerida ühtegi tõenäosusmõõtu, mille jaotusfunktsioon oleks pidev.

2.8.1 Lõplik aditiivsus

Nägime, et ruumi $\mathbb{R}^k$ kõikide alamhulkade hulgal pole võimalik defineerida liigutamis-invariantset mõõtu (Lebesgue'i mõõtu), mis oleks mõnel tõkestatud hulgal lõplik. Kui see nii oleks, viiks mõõdu loenduv aditiivsus vastuoluni teoreemiga 2.12. Väljapääs sellest olukorrast on mittemõõtuvate hulkade olemasolu aksepteerimine ja mõõtuvate hulkade s.o. teatud σ-algebra kasutamine.

Teine võimalus olukorrast väljatulemiseks on mõõdu loenduva aditiivsuse asendamine (nõrgema) lõpliku aditiivsusega. Sellisel juhul ei vii liigutamis-invariantsus vastuoluni teoreemiga 2.12. Seega küsime: kas ruumi $\mathbb{R}^k$ kõigi alamhulkade hulgal on võimalik defineerida lõplikult aditiivset hulgafunktsiooni (mõõtu), mis oleks liigutamis-invariantne? Juba aastal 1914 tõestas F. Hausdorff (läbi n.n. Hausdorffi paradoksi), et juhul kui $k \geq 3$, siis sellist mõõtu ei leidu. Banach ja Tarski esitasid selle negatiivse tulemuse erakordselt selgel kujul:

Teoreem 2.14 (Banach, Tarski, 1924; Banach-Tarski paradoks) *Olgu $k \geq 3$ ja olgu $A, B \subset \mathbb{R}^k$ mittetühja sisemusega tõkestatud hulgad. Siis leidub Hulga A lõplik lahutus $A = A_1 \cdot \cdots \cdot A_n$ $A_i \cap A_j = \emptyset$, kui $i \neq j$ ja liigutused $T_1, \ldots, T_n$ nii, et $T_1(A_1), \cdots, T_n(A_n)$ on hulga B lahutus.*

Banach-Tarski paradoks kehtib juba ruumis $\mathbb{R}^3$. Seega on võimalik lahutada kera raadiusega 1 (hernes) lõplikuks arvuks tükkideks nii, et neist saab kokku kera raadiusega 10^{8000} (maakera). Loomulikult välistab Banach-Tarski paradoks lõplikult aditiivse liigutamis-invariantse mõõdu olemasolu ruumi $\mathbb{R}^k$ kõigil alamhulkadel, kui $k \geq 3$.

Kui $k = 1, 2$ on lõplikult aditiivse liigutamis-invariantse mõõdu ühikkera normeeriva mõõdu konstrueerimine võimalik:

Teoreem 2.15 (Banach, 1923) *Olgu $k = 1, 2$. Boreli σ-algebral defineeritud Lebesgue'i mõõdul eksisteerib liigutamis-invarantne lõplikult aditiivne laiend hulgal $2^{\mathbb{R}^k}$.*

2.9 Ülesanded

1. Olgu $\mu : \Sigma \to [0, \infty]$ σ-algebral defineeritud funktsioon. Tõestada, et kui μ rahuldab tingimusi:

1) $\exists A \in \Sigma : \mu(A) < \infty$

2) Kui $A_1,\ A_2, \ldots \in \Sigma,\ A_i \cap A_j = \emptyset\, \forall i \neq j$, siis $\mu(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i)$,

siis on μ mõõt.

2. Tõestada mõõdu omadused **1) 2) 4) 5) 6)**.

3. Olgu $(\Omega, \mathcal{F}, \mathbf{P})$ tõenäosusruum, $A_t \in \mathcal{F}$, $t > 0$. Tõestada

1) kui $\{A_t | t > 0\}$ on kasvav, st $A_s \subset A_t$, $s \leq t$, siis $A := \cup_{t>0} A_t \in \mathcal{F}$ ja $\mathbf{P}(A_t) \nearrow \mathbf{P}(A)$;

2) kui $\{A_t | t > 0\}$ on kahanev, st $A_s \supset A_t$, $s \leq t$, siis $A := \cap_{t>0} A_t \in \mathcal{F}$ ja $\mathbf{P}(A_t) \searrow \mathbf{P}(A)$.

4. Olgu $(S, \Sigma, \mu) = (\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathrm{Leb})$, $a, b \in \mathbb{R}, a < b$. Leida

$$\mathrm{Leb}(\{a\}), \mathrm{Leb}((-\infty, a]), \mathrm{Leb}((a, \infty)), \mathrm{Leb}((a, b)), \mathrm{Leb}([a, b]).$$

5. Olgu $(S, \Sigma, \mu) = (\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathrm{Leb})$, $H_n = (n, \infty)$. Veendu, et omadus 6 ei pea paika. Miks?

6. Olgu S lõpmatu hulk, Σ_o-lõplikud ja kolõplikud hulgad. Defineerime funktsiooni

$$\mu : \Sigma_0 \to [0, 1], \quad \mu(A) = \begin{cases} 1,\ A & \text{ko-lõplik} \\ 0,\ A & \text{lõplik} \end{cases}$$

Tõestada

1) μ pole korrektselt defineeritud, kui S on lõplik;

2) μ on lõplikult aditiivne;

3) Kui S on loenduv, siis μ pole loenduvalt aditiivne;

4) Kui S on mitteloenduv, siis μ on loenduvalt aditiivne.

7. Olgu S mitteloenduv hulk, Σ koosnegu ülimalt loenduvatest ja ko-ülimalt loenduvatest hulkadest. Defineerime funktsioon

$$\mu : \Sigma \to [0, 1], \quad \mu(A) = \begin{cases} 1,\ A & \text{ko - ülimalt loenduv} \\ 0,\ A & \text{ülimalt loenduv} \end{cases}$$

Tõestada

1) μ pole korrektselt defineeritud, kui S on loenduv;

2) μ on loenduvalt aditiivne.

8. Kuulugu $x \in S$.

1) Näidata, et funktsioon

$$\delta_x(F) = \begin{cases} 1, & \text{kui } x \in F \\ 0, & \text{kui } x \notin F \end{cases}$$

on mõõt σ-algebral $(S, 2^S)$

2) olgu $\{a_n\}$ positiivsete arvude jada, $\{x_n\} \subset S$. Näita, et funktsioon $\mu = \Sigma_1^\infty a_n \delta_{x_n}$ on ruumil $(S, 2^S)$ antud mõõt. Veendu, et μ on lõplik, kui $\sum_n a_n < \infty$. Millal on μ tõenäosusmõõt?

3) Olgu Σ selline, et $\{x_n\} \in \Sigma \, \forall n$. Tõestada, et μ on σ-lõplik, kui punktid x_n on erinevad.

9. Olgu $(S, \Sigma) = (S, 2^S)$. Defineerime kujutuse

$$\mu : \Sigma \to [0, \infty], \;\; \mu(A) = \begin{cases} |A|, & \text{kui } |A| < \infty \\ \infty, & \text{kui } A \text{ lõpmatu} \end{cases}$$

Näita, et

1) μ on mõõt ruumil $(S, 2^S)$;

2) μ on lõplik $\Leftrightarrow |S| < \infty$;

3) μ on σ-lõplik $\Leftrightarrow S$ on loenduv.

10. Olgu $S = \{1, 2, 3, \cdots\}, \Sigma = 2^S$. Defineerime kujutuse

$$\mu : \Sigma \to [0, \infty], \;\; \mu(A) = \begin{cases} \sum_{k \in A} 2^{-k} & \text{kui } |A| < \infty \\ \infty, & \text{kui } A \text{ lõpmatu} \end{cases}$$

Kas μ on lõplikult aditiivne? Kas μ on loenduvalt aditiivne?

11. Olgu (S, Σ) mõõtuv ruum, $S_n \subset S$, $n = 1, 2, \ldots$. Olgu μ_n σ-algebral $S_n \cap \Sigma$ defineeritud mõõt. Tõestada, et μ on mõõt, kus

$$\mu : \Sigma \to [0, \infty], \quad \mu(A) = \sum_{n=1}^{\infty} \mu_n(A \cap S_n).$$

12. Olgu (S_n, Σ_n, μ_n) mõõdu ruumid, $n = 1, 2, \ldots$. Olgu $S = \cup S_n, S_n \cap S_m = \emptyset, \forall n \neq m$. Defineerime hulga S alamhulkade süsteemi $\Sigma = \{A \subset S | A = \cup A_n, A_n \in \Sigma_n \, \forall n\}$. Näita, et

1) (S, Σ) on mõõtuv ruum;

2) funktsioon $\mu : \Sigma \to [0, \infty], \mu(A) = \sum_n \mu_n(A_n)$ on mõõt.

Millal μ on σ-lõplik?

13. Olgu $S = (0, 1]$, koosnegu alamhulkade hulk $\mathcal{B}_0$ lõikumatute poollõikude lõplikest ühenditest, st $B \in \mathcal{B}_0$, kui $B = (a_1, b_1] \cup (a_2, b_2] \cup \cdots \cup (a_k, b_k]$, kus $0 \leq a_1 \leq b_1 \leq a_2 \leq \cdots \leq b_k \leq 1$. Defineerime funktsiooni

$$\mu : \mathcal{B}_0 \to [0, 1], \;\; \mu(B) = \begin{cases} 1, & \text{kui } \exists \epsilon > 0 \,:\, (0.5, 0.5 + \epsilon] \subset B \\ 0 & \text{kui } \;\nexists \epsilon > 0 \,:\, (0.5, 0.5 + \epsilon] \subset B \end{cases}$$

Näidata, et μ on lõplikult, kuid mitte loenduvalt aditiivne.

14. Olgu (S, Σ, μ) mõõdu ruum, kuulugu $A_1, A_2, \ldots \in \Sigma$, kusjuures $\mu(A_n \cap A_m) = 0 \; \forall m \neq n$. Näita, et $\mu(\cup_n A_n) = \sum_n \mu(A_n)$.

15. Olgu (S, Σ, μ) mõõdu ruum, μ olgu lõplik mõõt. Defineerime $\rho(A, B) = \mu(A \triangle B)$. Tõestada, et ρ on poolmeetrika hulgal Σ.

16. Olgu (S, Σ, μ) mõõdu ruum, μ olgu σ-lõplik. Tõestada

1) kui $A \subset S$ ja $\forall B \in \Sigma$: $\mu(B) < \infty$ korral $A \cap B \in \Sigma$, siis $A \in \Sigma$;

2) kui $A \in \Sigma$ ja $\mu(A) > 0$, siis $\exists B \subset A : \; B \in \Sigma, 0 < \mu(B) < \infty$.

17. Olgu λ σ-algebral $\mathcal{B}(\mathbb{R})$ antud nihke-invariantne mõõt:

$$\lambda(x + A) = \lambda(A), \quad \forall x \in \mathbb{R}, \tag{2.19}$$

kusjuures $\lambda([0, 1]) = 1$. Tõestada, et $\lambda = \text{Leb}$.

3 Sündmused

Olgu (S, Σ) mõõtuv ruum, $\{A_n\}$ olgu sündmuste jada, st. $A_n \in \Sigma \; \forall \; n$ korral.

Def 3.1 *Sündmuste jada* A_n **ülemiseks piirväärtuseks** *nimetame hulka* $\limsup_n A_n$, *kus*

$$\begin{aligned}\limsup_n A_n &:= \bigcap_m^\infty \bigcup_{n\geq m}^\infty A_n \\ &= \{s | \forall \; m \; \exists n(s) \geq m \; : \; s \in A_{n(s)}\} \\ &= \{s | \; s \in A_n \text{ lõpmata palju kordi}\}\end{aligned}$$

Def 3.2 *Sündmuste jada* A_n **alumiseks piirväärtuseks** *nimetame hulka* $\liminf_n A_n$, *kus*

$$\begin{aligned}\liminf_n A_n &:= \bigcup_m^\infty \bigcap_{n\geq m}^\infty A_n \\ &= \{s | \; \exists \; m(s) : \; s \in A_n \; \forall n \geq m(s)\} \\ &= \{s | \; s \in A_n \; \forall n \, \text{korral mingist kohast alates}\}\end{aligned}$$

Def 3.3 *Kui* $\liminf_n A_n = \limsup_n A_n = A$, *ütleme, et jada* A_n *koondub hulgaks* A *ehk* $A_n \to A$ *ehk* $A = \lim A_n$.

Iga hulga $A \subset S$ **indikaatorfunktsioon** on

$$I_A(s) := \begin{cases} 1 & \text{kui } s \in A, \\ 0 & \text{kui } s \notin A. \end{cases}$$

Tähistus "$\limsup_n A_n$" ja "$\liminf_n A_n$" peegeldab fakti, et $\forall s \in S$ korral (ülesanne 7)

$$I_{\limsup_n A_n}(s) = \limsup_n I_{A_n}(s), \quad I_{\liminf_n A_n}(s) = \liminf_n I_{A_n}(s).$$

Lemma 3.1 (Esimene Boreli-Cantelli lemma (BC1)) *Olgu* (S, Σ, μ) *mõõdu ruum ja* $\{A_n\} \subset \Sigma$ *selline sündmuste jada, et* $\sum_n \mu(A_n) < \infty$. *Siis*

$$\mu(\limsup_n A_n) = \mu(\{s | \; s \in A_n \text{ lõpmata palju kordi}\}) = 0.$$

Tõestus. Mõõt on loenduvalt subaditiivne. Seega iga m korral kehtib:

$$\mu(\limsup_n A_n) \leq \mu(\cup_{n=m}^\infty A_n) \leq \sum_{n=m}^\infty \mu(A_n).$$

Kui $\sum_{n=1}^{\infty} \mu(A_n) < \infty$, siis $\lim_{m\to\infty} \sum_{n=m}^{\infty} \mu(A_n) = 0$. ■

BC 1 kehtib suvalise mõõdu korral, muuhulgas ka tõenäosusmõõdu korral. Tõenäosusmõõdu korral kehtivad võrratused

$$\mathbf{P}(\liminf_n A_n) \le \liminf_n \mathbf{P}(A_n) \le \limsup_n \mathbf{P}(A_n) \le \mathbf{P}(\limsup_n A_n). \tag{3.1}$$

Nende omaduste tõestus on ülesanded 8 ja 9. Võrratustest (3.1) järeldub, et kui $A_n \to A$, siis

$$\mathbf{P}(A_n) \to \mathbf{P}(A).$$

3.1 Ülesanded

1. Tõestada, et kui $\{A_n\} \subset \Sigma$, siis $\limsup_n A_n \in \Sigma$ ja $\liminf_n A_n \in \Sigma$.

2. Tõestada, et

$$\begin{aligned}
\liminf_n A_n &\subset \limsup_n A_n \\
(\limsup_n A_n)^c &= \liminf_n (A_n^c) \\
(\liminf_n A_n)^c &= \limsup_n (A_n^c).
\end{aligned}$$

3. Tõestada, et

$$\begin{aligned}
\limsup_n (A_n \cap B_n) &\subset (\limsup_n A_n) \cap (\limsup_n B_n) \\
\limsup_n (A_n \cup B_n) &= (\limsup_n A_n) \cup (\limsup_n B_n) \\
\liminf_n (A_n \cap B_n) &= (\liminf_n A_n) \cap (\liminf_n B_n) \\
\liminf_n (A_n \cup B_n) &\supset (\liminf_n A_n) \cup (\liminf_n B_n).
\end{aligned}$$

4. Tõestada, et

$$\limsup_n A_n \backslash \liminf_n A_n = \limsup_n (A_n \cap A_{n+1}^c) = \limsup_n (A_n^c \cap A_{n+1}).$$

5. Näidata, et kui $A_n \to A$ ja $B_n \to B$, siis $A_n \cup B_n \to A \cup B$ ja $A_n \cap B_n \to A \cap B$.

6. Olgu $A_i \in \Sigma$, $i = 1, \ldots, n$, $U_k = \cup(A_{i_1} \cap \cdots \cap A_{i_k})$ ja $I_k = \cap(A_{i_1} \cup \cdots \cup A_{i_k})$, kus ühendid ja ühisosad võetakse üle kõikide selliste k-elemendiliste alamhulkade, mis rahuldavad tingimust $1 \le i_1 < \cdots < i_k \le n$. Tõestada, et $U_k = I_{n-k+1}$.

7. Olgu $(\Omega, \mathcal{F}, \mathbf{P})$ tõenäosusruum, $\{A_n\} \subset 2^{\Omega}$. Tõestada, et

$$\begin{aligned}
I_{\limsup_n A_n}(\omega) &= \limsup_n I_{A_n}(\omega) \\
I_{\liminf_n A_n}(\omega) &= \liminf_n I_{A_n}(\omega).
\end{aligned}$$

$\forall \omega \in \Omega$ korral.

8. Olgu $(\Omega, \mathcal{F}, \mathbf{P})$ tõenäosusruum, $\{A_n\} \subset \mathcal{F}$. Tõestada, et

$$\mathbf{P}(\limsup_n A_n) = \lim_n \mathbf{P}(\cup_{i=n}^{\infty} A_i)$$
$$\mathbf{P}(\liminf_n A_n) = \lim_n \mathbf{P}(\cap_{i=n}^{\infty} A_i).$$

9. Olgu $(\Omega, \mathcal{F}, \mathbf{P})$ tõenäosusruum, $\{A_n\} \subset \mathcal{F}$. Tõestada, et

$$\mathbf{P}(\liminf_n A_n) \leq \liminf_n \mathbf{P}(A_n) \leq \limsup_n \mathbf{P}(A_n) \leq \mathbf{P}(\limsup_n A_n).$$

Järeldada, et kui $A_n \to A$, siis $\mathbf{P}(A_n) \to \mathbf{P}(A)$.

10. Olgu $(\Omega, \mathcal{F}, \mathbf{P})$ tõenäosusruum, $\{A_n\}, \{B_n\} \subset \mathcal{F}$. Olgu

$$\mathbf{P}(\limsup_n A_n) = \mathbf{P}(\liminf_n B_n) = 1.$$

Näidata, et

$$\mathbf{P}(\limsup_n (A_n \cap B_n)) = 1.$$

11. Olgu $(\Omega, \mathcal{F}, \mathbf{P})$ tõenäosusruum, $\{A_n\}, \{B_n\} \subset \mathcal{F}$, kusjuures $\mathbf{P}(B_n) \to 1$. Tõestada, et

$$\lim \mathbf{P}(A_n) = \lim \mathbf{P}(B_n \cap A_n),$$

kui vähemalt üks nendest piirväärtustest eksisteerib.

12. Tõestada, et

$$\limsup_n (\liminf_k (A_n \cap A_k^c)) = \emptyset.$$

13. Olgu $(\Omega, \mathcal{F}, \mathbf{P})$ tõenäosusruum, $\{A_n\} \subset \mathcal{F}$, $A^* = \limsup_n A_n$, $A_* = \liminf_n A_n$. Tõestada, et

a) $\mathbf{P}(A_n \backslash A^*) \to 0$ ja $\mathbf{P}(A_* \backslash A_n) \to 0$

b) kui $A_n \to A$, siis $\mathbf{P}(A \triangle A_n) \to 0$.

4 Mõõtuvad funktsioonid, juhuslikud suurused

4.1 Mõõtuvus

Olgu (S, Σ) mõõtuv ruum, $h : S \to \mathbb{R}$ funktsioon, $\mathcal{B} := \mathcal{B}(\mathbb{R})$.

Hulga $A \subset \mathbb{R}$ **originaaliks** $h^{-1}(A)$ nimetatakse hulka

$$h^{-1}(A) = \{s \in S | \, h(s) \in A\}.$$

Originaali omadused (ülesanne 1)

$$h^{-1}(\bigcap_\alpha A_\alpha) = \bigcap_\alpha h^{-1}(A_\alpha), \quad h^{-1}(\bigcup_\alpha A_\alpha) = \bigcup_\alpha h^{-1}(A_\alpha), \quad h^{-1}(A^c) = (h^{-1}(A))^c.$$

Def 4.1 *Funktsiooni h nimetatakse* Σ-**mõõtuvaks** *(ka $\Sigma\backslash\mathcal{B}$ mõõtuvaks), kui iga Boreli hulga $A \in \mathcal{B}$ korral $h^{-1}(A) \in \Sigma$.*

Kui σ-algebra Σ on fikseeritud, nimetatakse Σ-mõõtuvaid funktsioone ka lihtsalt mõõtuvateks. Kõikide Σ-mõõtuvate funktsioonide hulka tähistame $m\Sigma$.

Def 4.2 *Funktsiooni $h : S \to \mathbb{R}$ nim.* **Boreli funktsiooniks***, kui h on $\mathcal{B}(S)$-mõõtuv. Tihti on $S = \mathbb{R}$.*

Näited:

1. Olgu $A \subset S$. Hulga A **indikaatorfunktsioon**

$$I_A(s) = \begin{cases} 1 & \text{kui } s \in A, \\ 0 & \text{kui } s \notin A. \end{cases}$$

on Σ-mõõtuv parajasti siis, kui $A \in \Sigma$.

2. Funktsiooni $h : S \to \mathbb{R}$ nimetame **lihtsaks**, kui tal on lõplik hulk väärtusi. Seega h on lihtne parajasti siis, kui $h(S) = \{a_1, \dots, a_m\}$, kus $a_1, \dots, a_m \in \mathbb{R}$, $a_i \neq a_j$, kui $i \neq j$. Indikaatorfunktsioon on lihtne funktsioon (miks?). Võttes $A_i = \{s : h(s) = a_i\}$ näeme, et iga lihtsa funktsiooni saab esitada kujul

$$\sum_{i=1}^m a_i I_{A_i}, \quad \text{kus } A_i \subset S \text{ ja } A_i \cap A_j = \emptyset, \quad a_i \neq a_j, \text{ kui } i \neq j \text{ ja } \cup_i A_i = S. \quad (4.1)$$

Teisest küljest, igal kujul (4.1) antud funktsioonil on m väärtust, mistõttu ta on lihtne.

Millal on lihtne funktsioon (4.1) mõõtuv? Veendu, et iga Boreli hulga mittetühi originaal on kujul $\cup_j A_{i_j}$, kus $1 \leq i_1 \leq i_2 \leq \dots \leq i_k \leq m$ ning iga sellisel kujul olev hulk on mingi Boreli hulga originaal. Siit järeldub, et lihtne funktsioon on mõõtuv parajasti siis, kui $A_i \in \Sigma$ iga $i = 1, \dots, m$ korral. Et $A_i = \{s : h(s) = a_i\} = h^{-1}\{a_i\}$, saame et lihtne funktsioon h väärtuste hulgaga $\{a_1, \dots, a_m\}$ on mõõtuv parajasti siis, kui iga väärtuse originaal on mõõtuv (st kuulub hulka Σ):

$$h \in m\Sigma \quad \Leftrightarrow \quad h^{-1}\{a_i\} \in \Sigma, \quad i = 1, \dots, m.$$

4.1.1 Millal on funktsioon mõõtuv?

Definitsioonist lähtudes tuleb funktsiooni h mõõtuvuses veendumiseks kontrollida, et iga Boreli hulga originaal kuulub σ-algebrasse Σ. Järgnev lihtne kuid oluline lemma väidab aga, et h mõõtuvuseks piisab, kui mingi Boreli σ-algebrat $\mathcal{B}$ genereeriva klassi $\mathcal{C}$ korral kuulub σ-algebrasse Σ iga $\mathcal{C}$ elemendi originaal.

Lemma 4.1 *Olgu $\mathcal{C}$ reaalarvude hulkade klass, $h : S \to \mathbb{R}$ funktsioon. Kui $h^{-1}(A) \in \Sigma$, $\forall A \in \mathcal{C}$ ja $\sigma(\mathcal{C}) = \mathcal{B}$, siis $h \in m\Sigma$.*

Tõestus. Defineerime klassid

$$h(\Sigma) = \{A \subset \mathbb{R} | h^{-1}(A) \in \Sigma\}.$$

Originaali omadustest järeldub, et $h(\Sigma)$ on σ-algebra (ülesanne 4). Et $\mathcal{C} \subset h(\Sigma)$, siis $\sigma(\mathcal{C}) = \mathcal{B} \subset h(\Sigma)$. ■

Järeldus 4.1 *Olgu $h : S \to \mathbb{R}$ funktsioon. Järgmised väited on ekvivalentsed:*

- *h on Σ-mõõtuv;*
- $\{h \leq c\} := \{s \in S | h(s) \leq c\} \in \Sigma \quad \forall c \in \mathbb{R}$;
- $\{h \geq c\} \in \Sigma \quad \forall c \in \mathbb{R}$;
- $\{h > c\} \in \Sigma \quad \forall c \in \mathbb{R}$;
- $\{h < c\} \in \Sigma \quad \forall c \in \mathbb{R}$;
- $\{h \leq q\} \in \Sigma \quad \forall q \in \mathbb{Q}$;
- $\{h > q\} \in \Sigma \quad \forall q \in \mathbb{Q}$;
- $\{h \in O\} \in \Sigma \quad \forall O$ *lahtine;*
- $\{h \in F\} \in \Sigma \quad \forall F$ *kinnine;*
- $\{a < h \leq b\} \quad \forall a < b, \quad a, b \in \mathbb{R}$.

Järeldus 4.2 *Olgu S topoloogiline ruum ja $h : S \to \mathbb{R}$ pidev. Siis $h \in m\mathcal{B}(S)$. Seega kui $h : \mathbb{R} \to \mathbb{R}$ on pidev, on ta Boreli funktsioon.*

Tõestus. Topoloogilisel ruumil antud funktsioon on pidev parajasti siis kui iga lahtise hulga originaal on lahtine. Võta $\mathcal{C}$ rolli kõik lahtised hulgad ja rakenda lemmat 4.1. ■

Lemma 4.2 *Olgu $f : \mathbb{R} \to \mathbb{R}$ Boreli funktsioon. Kui $h \in m\Sigma$, siis $f \circ h \in m\Sigma$.*

Tõestus. Ülesanne 5. ■

Lemmast 4.2 ja järeldusest 4.2 järeldub, et kui h on mõõtuv, siis seda on ka iga funktsioon kujul $f \circ h$, kus f on pidev.

4.1.2 Mõõtuvad vektorfunktsioonid

Olgu

$$h = (h_1, \ldots, h_n) : S \to \mathbb{R}^n \tag{4.2}$$

Def 4.3 *Funktsioon (4.2) on mõõtuv, kui $h^{-1}(B) \in \Sigma$, iga Boreli hulga $B \in \mathcal{B}(\mathbb{R}^n)$ korral.*

Lemma 4.3 *Vektorfunktsioon h on mõõtuv parajasti siis, kui iga komponent $h_i : S \to \mathbb{R}$ on mõõtuv.*

Enne lemma tõestamist pane tähele, et Lemma 4.1 kehtib ka siis, kui ruum $\mathbb{R}$ asendada ruumiga $\mathbb{R}^n$ (tõestuses midagi ei muutu). **Tõestus.** Olgu $A = (-\infty, x_1] \times \cdots \times (-\infty, x_n]$ π-süsteemi $\pi(\mathbb{R}^n)$ element. On selge, et

$$h^{-1}(A) = \bigcap_{i=1}^{n} h_i^{-1}(-\infty, x_i]. \tag{4.3}$$

Kui h_i on mõõtuv iga $i = 1, \ldots, n$ korral siis (4.3) on Σ element ning et $\pi(\mathbb{R}^n)$ tekitab $\mathcal{B}(\mathbb{R}^n)$, siis lemmast 4.1 järeldub h mõõtuvus.
Teistpidi: fikseeri $i = 1, \ldots, n$ ning võta $x_j = m \ \forall j \neq i$, $x_i = x$. Saad hulgad A_m. Et h on mõõtuv, siis $h^{-1}(A_m) \in \Sigma$, $\forall m$. Et aga

$$h^{-1}(A_m) \nearrow h_i^{-1}(-\infty, x],$$

siis $h_i^{-1}(-\infty, x] \in \Sigma$, millest järeldub h_i mõõtuvus. ∎

Rakendus: Olgu $h_1, \ldots, h_n$ Σ-mõõtuvad funktsioonid ja olgu

$$f : \mathbb{R}^n \to \mathbb{R}$$

$\mathcal{B}(\mathbb{R}^n)$-mõõtuv (näiteks pidev). Lemmast 4.3 järeldub, et $(h_1, \ldots, h_n)$ on mõõtuv vektorfunktsioon. Lemmast 4.2 järeldub, et $f(h_1, \ldots, h_n)$ on Σ-mõõtuv.

Näide: Summa ja korrutis on ruumil $\mathbb{R}^n$ defineeritud mõõtuvad funktsioonid, seega on mõõtuvate funktsioonide h_i summa $\sum_{i=1}^{n} h_i$ ja korrutis $\prod_{i=1}^{n} h_i$ samuti mõõtuvad.

4.2 Piirväärtuste mõõtuvus

Laiendame mõõtuvate funktsioonide klassi $m\Sigma$ nii, et sinna kuuluksid ka laiendatud funktsioonid kujul $h : S \to [-\infty, \infty]$.

Def 4.4 *Funktsioon $h : S \to [-\infty, \infty]$ on Σ-mõõtuv, kui iga $A \in \mathcal{B}(\mathbb{R})$ korral $h^{-1}(A) \in \Sigma$ ning $h^{-1}\{-\infty\}, h^{-1}\{\infty\} \in \Sigma$.*

Lemma 4.4 *Funktisioon $h : S \to [-\infty, \infty]$ on Σ-mõõtuv parajasti siis, kui*

$$\{h \le c\} \in \Sigma \quad \forall c \in [-\infty, \infty]. \tag{4.4}$$

Tõestus. $\Rightarrow$: Olgu h mõõtuv definitsiooni 4.4 mõttes. On kerge näha, et siis kehtib (4.4). Tõepoolest, kui $c < \infty$, siis $\{h \le c\} = \{h = -\infty\} \cup \{-\infty < h \le c\} \in \Sigma$, sest $\{h = -\infty\} \in \Sigma$ ja $(-\infty, c] \in \mathcal{B}(\mathbb{R})$, millest $\{-\infty < h \le c\} = h^{-1}(-\infty, c] \in \Sigma$. Kui $c = \infty$, siis $\{h \le c\} = \{-\infty, \infty\} \cup h^{-1}(\mathbb{R}) \in \Sigma$.

$\Leftarrow$: Kehtigu (4.4). Tõestame, et h on mõõtuv definitsiooni 4.4 mõttes. Kõigepealt näitame:

$$h^{-1}\{\infty\}, h^{-1}\{-\infty\} \in \Sigma.$$

Et

$$\{h < \infty\} = \cup_n \{h \le n\} \in \Sigma,$$

siis

$$h^{-1}(\{\infty\}) = S \backslash \{h < \infty\} \in \Sigma.$$

Kui $c = -\infty$, siis

$$\{h \le -\infty\} = h^{-1}\{-\infty\} \in \Sigma.$$

Nüüd tõestame, et $h^{-1}(B) \in \Sigma$, kui $A \in \mathcal{B}$. Olgu

$$S_o := h^{-1}(\mathbb{R}), \quad h_o := h|_{S_o}, \quad \text{s.t. } h_o : S_o \to \mathbb{R}.$$

Veendume, et $\{h_o \le c\} \in \Sigma \; \forall c \in \mathbb{R}$. Et $\{h \le c\} \in \Sigma$ (eeldus) ja $h^{-1}\{-\infty\} \in \Sigma$, siis

$$\{h_o \le c\} = \{h \le c\} \backslash h^{-1}(-\infty) \in \Sigma.$$

Funktsioon h_o on harilik (st tema väärtused on lõplikud). Et $\{h_o \le c\} \in \Sigma$, siis järeldusest 4.1 saame, et h_o on Σ-mõõtuv ehk iga $A \in \mathcal{B}$ korral $h_o^{-1}(A) \in \Sigma$. Kuid $h^{-1}(A) = h_o^{-1}(A)$, mistõttu ka $h^{-1}(A) \in \Sigma$. $\blacksquare$

Märkus: Lemma 4.4 kehtib kõikide järelduses 4.1 toodud klasside korral.

Lemma 4.5 *Olgu h_n Σ-mõõtuvate funktsioonide jada, st. $h_n \in m\Sigma \quad n = 1, 2, \ldots$. Siis*

a) $\inf_n h_n \in m\Sigma$

b) $\sup_n h_n \in m\Sigma$

c) $\liminf_n h_n \in m\Sigma$

d) $\limsup_n h_n \in m\Sigma$

e) $\{s | \lim h_n(s) \text{ existeerib}\} \in \Sigma$.

Tõestus. Kasutame lemmat 4.4. Selleks paneme tähele, et iga $c \in [-\infty, \infty]$ korral

$$\begin{aligned}
&\{\sup_n h_n \le c\} = \cap_n \{h_n \le c\}; \quad \{\inf_n h_n \ge c\} = \cap_n \{h_n \ge c\};\\
&\liminf_n h_n = \sup_n \inf_{m \ge n} h_m, \quad \limsup_n h_n = \inf_n \sup_{m \ge n} h_m;\\
&\{s | \lim h_n \text{ eksisteerib}\} = \{s | \liminf_n h_n = \limsup_n h_n\}.
\end{aligned}$$

■

Järeldus 4.3 *Kui $h_n \in m\Sigma$ ja $h_n(s) \to h(s)$ $\forall s \in S$, siis $h \in m\Sigma$.*

Tõestus. Kui $h_n(s) \to h(s)$, $\quad \forall s \in S$, siis $h = \liminf_n h_n = \limsup_n h_n$. ■

Teoreem 4.5 *Funktsioon $h : S \to [-\infty, \infty]$ on Σ-mõõtuv parajasti siis, kui leidub lihtsate mõõtuvate funktsioonide jada h_n nii, et $h_n(s) \to h(s)$, $\forall s \in S$, kusjuures*

$$0 \le h_n(s) \nearrow h(s), \text{ kui } h(s) \ge 0 \quad \text{ja} \quad 0 \ge h_n(s) \searrow h(s), \text{ kui } h(s) \le 0.$$

Tõestus. $\Leftarrow$: Järeldus 4.3.
$\Rightarrow$: Olgu $h \in m\Sigma$. Teoreemis nõutavate omadustega lihtsad funktsioonid h_n defineerime

$$h_n(s) = \begin{cases} -n & \text{kui } -\infty \le h(s) \le -n \\ -(k-1)2^{-n} & \text{kui } -k2^{-n} < h(s) \le -(k-1)2^{-n} \ k = 1, \dots, n2^n \\ (k-1)2^{-n} & \text{kui } (k-1)2^{-n} \le h(s) < k2^{-n} \ k = 1, \dots, n2^n \\ n & \text{kui } n \le h(s) \le \infty. \end{cases}$$

Seega

$$h_n = \sum_{k=1}^{n2^n} \frac{k-1}{2^n} I_{A^+_{k,n}} + \sum_{k=1}^{n2^n} -\frac{(k-1)}{2^n} I_{A^-_{k,n}} + nI_{A^+_n} - nI_{A^-_n},$$

kus

$$\begin{aligned}
A^+_{k,n} &:= \Big\{s : \frac{(k-1)}{2^n} \le h(s) < \frac{k}{2^n}\Big\} = h^{-1}\Big(\big[(k-1)2^{-n}, k2^{-n}\big)\Big) \in \Sigma,\\
A^-_{k,n} &:= \Big\{s : -\frac{k}{2^n} < h(s) \le -\frac{(k-1)}{2^n}\Big\} = h^{-1}\Big(\big(-k2^{-n}, -(k-1)2^{-n}\big]\Big) \in \Sigma,\\
A^+_n &= \{s : h(s) \ge n\} = h^{-1}\big([n, \infty)\big) \in \Sigma\\
A^-_n &= \{s : h(s) \le -n\} = h^{-1}\big((-\infty, -n]\big) \in \Sigma,
\end{aligned}$$

mistõttu on h_n lihtne mõõtuv funktsioon.
Jada h_n koondub funktsiooniks h. Tõepoolest, olgu s selline, et $0 \le h(s) \le n$. Siis leidub k nii, et $s \in A^+_{k,n}$, s.t.

$$h_n(s) = (k-1)2^{-n}, \quad (k-1)2^{-n} \le h(s) < k2^{-n}$$

ja seetõttu

$$0 \le h(s) - h_n(s) \le \frac{1}{2^n}.$$

Intervallid $[(k-1)2^{-n}, k2^{-n}]$ moodustavad üksteisesse sisendatud tükelduse, millest $f(s) \ge 0$ korral $h_n(s) \le h_{n+1}(s)$ ning vastasel juhul $h_n(s) \ge h_{n+1}(s)$. ■

4.3 Juhuslikud suurused ja juhuslikud vektorid

Def 4.6 *Tõenäosusruumil* $(\Omega, \mathcal{F}, \mathbf{P})$ *antud* $\mathcal{F}$*-mõõtuvat funktsiooni* $X : \Omega \to \mathbb{R}$ *nimetatakse* **juhuslikuks suuruseks**.

Def 4.7 *Olgu* $X_1, \ldots, X_n$ *tõenäosusruumil* $(\Omega, \mathcal{F}, \mathbf{P})$ *antud juhuslikud suurused. Vektorfunktsiooni* $X = (X_1, \ldots, X_n) : \Omega \to \mathbb{R}^n$ *nimetatakse (n-dimensionaalseks)* **juhuslikuks vektoriks**.

Märkus: Lemma 4.3 tõttu on juhuslik vektor mõõtuv vektorfunktsioon $X : \Omega \to \mathbb{R}^n$.

Nagu eelpool näidatud, järeldub lemmadest 4.3 ja 4.2, et kui $f : \mathbb{R}^n \to \mathbb{R}$ on $\mathcal{B}(\mathbb{R}^n)$-mõõtuv funktsioon ja X on n-dimensionaalne juhuslik vektor, siis $f \circ X$ on juhuslik suurus. Seega on juhusliku vektori järgmised funktsioonid juhuslikud suurused:

$$\sum_{i=1}^{n} X_i, \quad \prod_{i=1}^{n} X_i, \quad X'X, \quad X_1 + X_2^2 + \cdots + X_n^n.$$

4.4 Funktsioonide tekitatud σ-algebrad

Olgu $\{h_\alpha |\, \alpha \in \Lambda\}$ funktsioonide hulk, s.t.

$$h_\alpha : S \to \mathbb{R} \quad \forall \alpha \in \Lambda.$$

Def 4.8 *Hulga* $\{h_\alpha |\, \alpha \in \Lambda\}$ *tekitatud* σ*-algebraks* $\sigma(h_\alpha |\, \alpha \in \Lambda)$ *nimetatakse väikseimat (sisalduvusseose mõttes)* σ*-algebrat* Σ^* *hulgal* S *nii, et* h_α *on* Σ^**-mõõtuv* $\forall \alpha \in \Lambda$.

Seega on $\sigma(h_\alpha |\, \alpha \in \Lambda)$ selline σ-algebra hulgal S, et

- h_α on $\sigma(h_\alpha |\, \alpha \in \Lambda)$-mõõtuv iga α korral;
- kui Σ on selline σ-algebra (hulgal S), et $\Sigma \subsetneqq \sigma(h_\alpha |\, \alpha \in \Lambda)$, siis leidub vähemalt üks funktsioon h_α, mis pole Σ-mõõtuv;
- kui $\{h_\alpha |\, \alpha \in \Lambda\}$ on Σ-mõõtuvate funktsioonide hulk, siis $\sigma(h_\alpha |\, \alpha \in \Lambda) \subseteq \Sigma$.

Definitsioonist järeldub

$$\sigma(h_\alpha | \alpha \in \Lambda) = \sigma(\{h_\alpha^{-1}(B) |\; B \in \mathcal{B}(\mathbb{R}), \alpha \in \Lambda\}).$$

Seega (ülesanne 9)

$$\sigma(h) = \{h^{-1}(B) |\, B \in \mathcal{B}(\mathbb{R})\}.$$

Viimane omadus üldistub ka juhuslikele vektoritele (tõestus on analoogiline):

$$\sigma(h_1, \ldots, h_n) = \Big\{(h_1, \ldots, h_n)^{-1}(B) \Big| \quad B \in \mathcal{B}(\mathbb{R}^n)\Big\}. \tag{4.5}$$

Olgu $\mathcal{C}$ mingi reaaltelje alamhulkade klass nii, et $\sigma(\mathcal{C}) = \mathcal{B}$. Klassi $\mathcal{C}$ kuuluvate elementide originaalid tekitavad $\sigma(h)$ (ülesanne 4):

$$\sigma(h) = \sigma\big(\{h^{-1}(C) : C \in \mathcal{C}\}\big).$$

Näide: Klass $\big\{\{s : h(s) \leq c\}, c \in \mathbb{R}\big\}$ tekitab $\sigma(h)$.

Tõenäosusruumil $(\Omega, \mathcal{F}, \mathbf{P})$ antud juhusliku suuruse X tekitatud σ-algebra $\sigma(X)$ on seega väikseim σ-algebra, mille suhtes X on mõõtuv st kui $X \in m\mathcal{G}$, siis $\sigma(X) \subseteq \mathcal{G}$.

Olgu $X = (X_1, \ldots, X_n)$ tõenäosusruumil $(\Omega, \mathcal{F}, \mathbf{P})$ antud juhuslik vektor. Seos (4.5):

$$\sigma(X) = \sigma(X_1, \ldots, X_n) = \{X^{-1}(B) |\, B \in \mathcal{B}(\mathbb{R}^n)\}.$$

Teoreem 4.9 *Olgu $X = (X_1, \ldots, X_n)$ tõenäosusruumil $(\Omega, \mathcal{F}, \mathbf{P})$ antud juhuslik vektor. Funktsioon $Y : \Omega \to \mathbb{R}$ on $\sigma(X)$-mõõtuv parajasti siis, kui leidub Boreli funktsioon $f : \mathbb{R}^n \to \mathbb{R}$ nii, et*

$$Y(\omega) = f(X_1(\omega), \ldots, X_n(\omega)), \quad \forall \omega \in \Omega.$$

Tõestus. Piisavus järeldub lemmast 4.2.
Tarvilikkus. Olgu $Y : \Omega \to \mathbb{R}$ mingi $\sigma(X)$-mõõtuv funktsioon.
1) Erijuht. Oletame, et Y on lihtne mõõtuv funktsioon, s.t.

$$Y = \sum_{i=1}^{k} y_i I_{A_i},$$

kus $y_i \neq y_j$ ja hulgad A_i on lõikumatud. Et Y on $\sigma(X)$-mõõtuv, siis $A_i \in \sigma(X), \forall i$. Seosest 4.5 järeldub, et leiduvad Boreli hulgad $B_i \in \mathcal{B}(\mathbb{R}^n)$ nii, et $A_i = X^{-1}(B_i)$. Defineerime funktsiooni

$$f : \mathbb{R}^n \to \mathbb{R}, \quad f := \sum_{i=1}^{k} y_i I_{B_i}.$$

On selge, et f on $\mathcal{B}(\mathbb{R}^n)$-mõõtuv (miks?). Iga $\omega \in \Omega$ korral leidub ainult üks B_i nii, et $X(\omega) \in B_i$. See tuleneb B_i definitsioonist ning asjaolust, et A_i-d on lõikumatud. Seega $X(\omega) \in B_i$ parajasti siis kui $\omega \in A_i$ ning sellisel juhul $f(X(\omega)) = y_i$. Seega $f(X(\omega)) = Y(\omega)$.
2) Üldine juht: Y on $\sigma(X)$-mõõtuv. Teoreemist 4.5 järeldub, et leidub lihtsate $\sigma(X)$-mõõtuvate funktsioonide jada Y_m nii, et

$$Y_m(\omega) \to Y(\omega), \quad \forall \omega \in \Omega.$$

Iga Y_m korral leidub $f_m : \mathbb{R}^n \to \mathbb{R}$ nii, et $f_m(X(\omega)) = Y_m(\omega)$. Olgu $M \subset \mathbb{R}^n$ defineeritud järgmiselt:

$$M := \{(x_1, \ldots, x_n) \in \mathbb{R}^n |\, \lim_m f_m(x_1, \ldots, x_n) \text{ eksisteerib}\}.$$

Teame, et $M \in \mathcal{B}(\mathbb{R}^n)$ (miks?). Defineerime

$$f : \mathbb{R}^n \to \mathbb{R}, \quad f := \begin{cases} \lim_m f_m(x) & \text{kui } x \in M \\ 0 & \text{mujal.} \end{cases}.$$

Et $f = \lim_m f_m I_M$ ning $f_m I_M$ on iga m korral Boreli funktsioon (miks?), siis vastavalt lemmale 4.5 on ka f Boreli funktsioon.
Funktsioonide f_m definitisioonist tulenevalt kehtib, et iga ω korral

$$Y(\omega) = \lim_m Y_m(\omega) = \lim_m f_m(X(\omega)).$$

Seega $X(\omega) \in M$ ning iga ω korral

$$Y(\omega) = \lim_m f_m(X(\omega)) = f(X(\omega)).$$

■

4.5 Mõõt μh^{-1}

Olgu (S, Σ, μ) mõõduga ruum, $h : S \to \mathbb{R}$, $h \in m\Sigma$.
Funktsioon h defineerib σ-algebral $\mathcal{B}(\mathbb{R})$ mõõdu μh^{-1}, (vaata ülesanne 10) kus

$$\mu h^{-1}(B) := \mu(h^{-1}(B)) = \mu\{s : h(s) \in B\}, \quad B \in \mathcal{B}.$$

Analoogiliselt defineerib Σ-mõõtuv vektorfunktsioon

$$h = (h_1, \ldots, h_n) : S \to \mathbb{R}^n$$

mõõdu μh^{-1} σ-algebral $\mathcal{B}(\mathbb{R}^n)$.

Def 4.10 *Olgu $X = (X_1, \ldots, X_n)$ tõenäosusruumil $(\Omega, \mathcal{F}, \mathbf{P})$ antud juhuslik vektor. Mõõtu $\mathbf{P}X^{-1}$ nimetatakse juhusliku vektori* **jaotuseks** *ja tähistatakse P_X. Kui $n = 1$, siis mõõtu $\mathbf{P}X^{-1}$ nimetatakse juhusliku suuruse jaotuseks.*

Olgu P σ-algebral $\mathcal{B}(\mathbb{R}^n)$ antud tõenäosusmõõt. Kui juhusliku vektori X jaotus on P, siis kirjutame

$$X \sim P.$$

Rõhutamaks, et P on juhusliku suuruse X jaotus, kirjutame teinekord ka P_X. Seega on tõenäosusruumil $(\Omega, \mathcal{F}, \mathbf{P})$ antud juhusliku vektori X jaotus P_X on selline tõenäosusmõõt (σ-algebral $\mathcal{B}(\mathbb{R}^n)$), mis rahuldab

$$P_X(B) = \mathbf{P}X^{-1}(B) = \mathbf{P}\{\omega : X(\omega) \in B\} =: \mathbf{P}(X \in B), \quad \forall B \in \mathcal{B}(\mathbb{R}^n).$$

Teame, et iga tõenäosusmõõtu σ-algebral $\mathcal{B}(\mathbb{R}^n)$ saab kirjeldada jaotusfunktsiooniga. Vastavalt definitsioonile avaldub mõõdu P_X jaotusfunktsioon

$$\begin{aligned} F(x_1, \ldots, x_n) &= P_X\Big((-\infty, x_1] \times \cdots \times (-\infty, x_n]\Big) \\ &= \mathbf{P}(X_i \in (-\infty, x_i],\ i = 1, \ldots, n) \\ &= \mathbf{P}(X_1 \leq x_1, \ldots, X_n \leq x_n). \end{aligned}$$

Juhusliku suuruse X korral

$$F(x) = P_X(-\infty, x] = \mathbf{P}(X \in (-\infty, x]) = \mathbf{P}(X \leq x).$$

Juhusliku vektori (suuruse) jaotusfunktsiooniks nimetame tema jaotuse jaotusfunktsiooni.

4.6 Skorohodi esitus

Igale juhuslikule suurusele vastab üks tõenäosusmõõt – tema jaotus. Iga jaotus on esitatav jaotusfunktsioonina. Iga jaotusfunktsioon rahuldab omadusi **1), 2), 3)** (jaotusfunktsiooni omadused). Kas aga iga omadusi **1), 2), 3)** rahuldav funktsioon F on mõne juhusliku suuruse jaotusfunktsioon, st kas leidub mingi tõenäosusruum ja sellel defineeritud juhuslik suurus nii, et selle juhusliku suuruse jaotus oleks F? Teoreemi 2.7 abil on sellele küsimusele kerge vastata. Tõepoolest, vastavalt teoreemile 2.7 leidub σ-algebral $\mathcal{B}(\mathbb{R})$ tõenäosusmõõt P nii, et $F(x) = P(-\infty, x]$, s.t. F on P jaotusfunktsioon. Olgu nüüd $(\Omega, \mathcal{F}, \mathbf{P}) = (\mathbb{R}, \mathcal{B}(\mathbb{R}), P)$. Jaotusega F juhusliku suuruse saab defineerida imelihtsalt:

$$X(\omega) = \omega.$$

Sellisel lihtsal konstruktsioonil on aga puudus. Nimelt sõltub ruumil $(\Omega, \mathcal{F})$ olev mõõt etteantud jaotusfunktsioonist. Mingi teise jaotusega juhuslik suurus peaks (selle konstruktsiooni põhjal) olema defineeritud mingil teisel tõenäosusruumil. Kas aga leidub selline universaalne tõenäosusruum $(\Omega, \mathcal{F}, \mathbf{P})$ millele saab konstrueerida suvalise jaotusfunktsiooniga juhusliku suuruse?
Järgnev teoreem näitab, et sellise tõenäosusruumina võib alati kasutada ruumi $((0,1), \mathcal{B}(0,1), \mathrm{Leb})$.

Teoreem 4.11 *Olgu F funktsioon, mis rahuldab jaotusfunktsiooni omadusi* **1), 2), 3)**. *Siis leidub ruumil $((0,1), \mathcal{B}(0,1), \mathrm{Leb})$ defineeritud juhuslik suurus X nii, et iga $x \in \mathbb{R}$ korral*

$$\mathrm{Leb}(X \leq x) = F(x).$$

Tõestus.
1) Erijuht: $F : \mathbb{R} \to [0,1]$ on pidev ja rangelt kasvav. Igal sellisel funktsioonil on pöördfunktsioon. Defineeri

$$X(\omega) = F^{-1}(\omega).$$

See funktsioon on pidev, mistõttu $\mathcal{B}$-mõõtuv. Kui $\omega \in (0,1)$, siis $X(\omega) = F^{-1}(\omega) \leq x$ parajasti siis, kui $\omega \leq F(x)$. Teisisõnu,

$$\{\omega : X(\omega) \leq x\} = \{\omega : \omega \leq F(x)\}. \tag{4.6}$$

Seega

$$\text{Leb}(X \leq x) = \text{Leb}\{\omega \in (0,1) : X(\omega) \leq x\} = \text{Leb}\{\omega \in (0,1) : \omega \leq F(x)\} = F(x). \quad (4.7)$$

2) Üldjuht. Kui F pole rangelt kasvav või pidev, siis iga $\omega \in (0,1)$ korral defineerime

$$X(\omega) := \inf\{x : \omega \leq F(x)\}. \quad (4.8)$$

Et F on paremalt pidev, on hulk $\{x : \omega \leq F(x)\}$ kujul $[t, \infty)$, mistõttu

$$\{x : \omega \leq F(x)\} = [X(\omega), \infty)$$

ja nii kehtib (4.6). Funktsiooni X mõõtuvus on ilmne, seosest (4.7) järeldub teoreemi väide. ■

Märkus: Tõestuses ei kasutatud teoreemi 2.7. Veel enam, teoreemist 4.11 järeldub teoreem 2.7. Tõepoolest, kui iga etteantud (jaotus)funktsiooni F korral leidub mingi tõenäosusruum $(\Omega, \mathcal{F}, \mathbf{P})$ ja sellel antud juhuslik suurus X nii, et $\mathbf{P}(X \leq x) = F(x)$, siis loomulikult leidub reaalteljel antud tõenäosusmõõt, mille jaotusfunktsioon on F: juhusliku suuruse jaotus P_X.
Seosega (4.8) defineeritud juhuslikku suurust nimetatakse *Skorohodi esituseks.*

Kokkuvõtteks: Iga mingil tõenäosusruumil $(\Omega, \mathcal{F}, \mathbf{P})$ antud juhuslik suurus X transformeerib (seose $\mathbf{P}X^{-1}$ kaudu) mõõdu $\mathbf{P}$ tõenäosusmõõduks reaalteljel P_X. Seda mõõtu nimetatakse X jaotuseks. Tõenäosusmõõdud reaalteljel on üksüheses vastavuses jaotusfunktsioonidega. Nii räägime juhusliku suuruse jaotusfunktsioonist.
Igale reaalteljel antud tõenäosusmõõdule P saab aga ruumil $((0,1), \mathcal{B}(0,1), \text{Leb})$ defineerida (Skorohodi esituse abil) juhusliku suuruse X nii, et $P_X = P$. Siit aga ei järeldu, et juhuslike suuruste ja jaotuste vahel oleks üksühene vastavus – mitmel juhuslikul suurusel võib olla sama jaotus.

Näited:

- $X \sim N(0, \sigma^2)$ ja $-X$ on sama jaotusega;
- Ruumil $((0,1), \mathcal{B}(0,1), \text{Leb})$ antud juhuslikud suurused I_A on kõik sama jaotusega:

  1 $A = (0, \frac{1}{2})$;
  2 $A = (0, \frac{1}{2}]$;
  3 $A = (\frac{1}{2}, 1)$;
  4 $A = (0, \frac{1}{8}) \cup (\frac{2}{8}, \frac{3}{8}) \cup (\frac{4}{8}, \frac{5}{8}) \cup (\frac{6}{8}, \frac{7}{8})$.

4.7 Ülesanded

1. Olgu $h : S \to \mathbb{R}$ funktsioon. Tõestada, et

$$h^{-1}(\bigcup_\alpha A_\alpha) = \bigcup_\alpha h^{-1}(A_\alpha), \quad h^{-1}(A^c) = (h^{-1}(A))^c$$

2. Olgu $(\Omega, \mathcal{F}, \mathbf{P})$ järgmine tõenäosusruum:
Ω- lõik $[0, 1]$;
$\mathcal{F}$- σ-algebra, mis on tekitatud kõikide ülimalt loenduvate hulkade ja nende täiendite poolt;
$\mathbf{P}$-Lebesque'i mõõt.
Kas järgmised funktsioonid on juhuslikud suurused? Miks?
a) $X(\omega) = \omega$;
b)

$$X(\omega) = \begin{cases} 1, & \text{kui } \omega = 0; \\ 2, & \text{kui } \omega = 0,5; \\ 3, & \text{kui } \omega = 1; \\ 0, & \text{mujal} \end{cases}$$

3. Olgu (S, Σ) mõõtuv ruum, h_1, $h_2 \in m\Sigma$. Tõestada, et
a)$A = \{s | h_1(s) < h_2(s)\} \in \Sigma$
b)$A = \{s | h_1(s) = h_2(s)\} \in \Sigma$
c)$A = \{s | h_1(s) \leq h_2(s)\} \in \Sigma$.

4. Olgu $h : S \to \mathbb{R}$ funktsioon ja Σ mingi hulgal S antud σ-algebra. Defineerime klassid

$$h^{-1}(\mathcal{B}) = \{h^{-1}(B) | \, B \in \mathcal{B}\},$$
$$h(\Sigma) = \{A \subset \mathbb{R} | \, h^{-1}(A) \in \Sigma\}$$

a) Tõestada, et $h^{-1}(\mathcal{B})$ ja $h(\Sigma)$ on σ-algebrad.

b) Tõestada, et $h^{-1}(\mathcal{B}) \subset \Sigma \Leftrightarrow \mathcal{B} \subset h(\Sigma)$;

c) Tõestada, et järgmised väited on samaväärsed

i) h on mõõtuv

ii) $h^{-1}(\mathcal{B}) \subset \Sigma$

iii) $\mathcal{B} \subset h(\Sigma)$.

d) Olgu $\mathcal{C}$ mingi reaaltelje alamhulkade klass, $\sigma(\mathcal{C}) = \mathcal{B}$. Tõestada, et

$$\sigma(h^{-1}(\mathcal{C})) = h^{-1}(\sigma(\mathcal{C})). \tag{4.9}$$

Järelda, et

$$\sigma(h) = \sigma(h^{-1}(\mathcal{C})).$$

[Näpunäide: seose $h^{-1}(\sigma(\mathcal{C})) \subseteq \sigma(h^{-1}(\mathcal{C}))$ näitamiseks võta $\Sigma = \sigma(h^{-1}(\mathcal{C}))$.]

5. Tõestada lemma 4.2.

6. Olgu $(\Omega, \mathcal{F}, \mathbf{P})$ tõenäosusruum. Kas funktsioon $X : \Omega \to \mathbb{R}$ on juhuslik suurus, kui järgmised funktsioonid on juhuslikud suurused:

$$X^2, \ |X|, \ \cos X, \ aX, \ \exp(X)?$$

7. Olgu $\{h_\alpha\}$ Σ-mõõtuvate funktsioonide mitteloenduv pere, st. $h_\alpha \in m\Sigma \quad \forall \alpha$ korral. Kas funktsioonid $\inf_\alpha h_\alpha$ ja $\sup_\alpha h_\alpha$ on Σ-mõõtuvad?

8. Olgu $(\Omega, \mathcal{F}, \mathbf{P}) = ([0{,}1], \mathcal{B}([0,1]), \text{Leb})$. Kirjeldada järgmiste juhuslike suuruste poolt tekitatud σ-algebraid:

- $X = 0.25 I_{[0.0.25)} + 0.5 I_{[0.25,0.75)} + I_{[0.75,1]}$;
- $X(\omega) = \frac{1}{2}\omega$;
- $X(\omega) = 0.5$.

9. Olgu X tõenäosusruumil $(\Omega, \mathcal{F}, \mathbf{P})$ antud juhuslik suurus. Tõestada, et

a) $\sigma(X) = \{\{\omega | X(\omega) \in B\} | \, B \in \mathcal{B}\} = \{X^{-1}(B) | \, B \in \mathcal{B}\} = X^{-1}(\mathcal{B})$;

b) $\sigma(X)$ on tekitatud järgmise π-süsteemi poolt:
$\pi(X) = \{\{\omega | \, X(\omega) < c\} | \, c \in \mathbb{R}\} = X^{-1}(\pi(\mathbb{R}))$.

10. Olgu (S, Σ, μ) mõõdu ruum, $h \in m\Sigma$. Tõestada, et μh^{-1} on mõõt ruumil $(\mathbb{R}, \mathcal{B})$.
Näita, et kui μ on tõenäosusmõõt, siis ka μh^{-1} on tõenäosusmõõt.
Näita, et kui μ on lõplik, siis on seda ka μh^{-1}.

11. Skorohodi esituse abil leida jaotusega P juhuslik suurus, kui

a) $P = E(\lambda)$ (eksponentjaotus parameetriga λ);

b) P jaotusfunktsioon on Olgu

$$F(x) = \begin{cases} 0.5e^x & \text{kui } x < 0, \\ 0.7 & \text{kui } x \in [0, 2), \\ 0.4x & \text{kui } x \in [2, 2.5]. \end{cases} \tag{4.10}$$

c $P = G(p)$ (geomeetriline jaotus parameetriga p), st

$$P = \sum_{k=1}^{\infty} p_k \delta_k, \quad p_k = (1-p)^{k-1} p.$$

5 Sõltumatus

5.1 Sündmuste sõltumatus

Olgu $(\Omega, \mathcal{F}, \mathbf{P})$ tõenäosusruum, $A_1, \dots, A_n$ olgu sündmused, s.t. $A_i \in \mathcal{F}$, $i = 1, \dots, n$.

Def 5.1

- *Sündmused $A_1 \dots, A_n$ on* **sõltumatud**, *kui*

$$\mathbf{P}(A_{k_1} \cap \ldots \cap A_{k_j}) = \mathbf{P}(A_{k_1}) \cdots \mathbf{P}(A_{k_j}),$$

  $2 \leq j \leq n$ ja $1 \leq k_1 < \cdots < k_j \leq n$ korral.

- *Lõpmatut (loenduvat või mitteloenduvat) hulka sündmusi $\{A_\alpha\}$ nim. sõltumatuteks, kui hulga $\{A_\alpha\}$ igasse lõplikusse alamhulka kuuluvad sündmused on sõltumatud.*

- *Sündmused $\{A_\alpha\}$ on* **paarikaupa sõltumatud**, *kui*

$$\mathbf{P}(A_{\alpha_i} \cap A_{\alpha_j}) = \mathbf{P}(A_{\alpha_i})\mathbf{P}(A_{\alpha_j}), \quad \forall \alpha_i \neq \alpha_j.$$

Definitsioonist järeldub, et neli sündmust A_1, A_2, A_3, A_4 on sõltumatud parajasti siis, kui kehtivad järgnevad võrdused:

$$\begin{aligned}
&\mathbf{P}(A_1 \cap A_2 \cap A_3 \cap A_4) = \mathbf{P}(A_1)\mathbf{P}(A_2)\mathbf{P}(A_3)\mathbf{P}(A_4),\\
&\mathbf{P}(A_1 \cap A_2 \cap A_3) = \mathbf{P}(A_1)\mathbf{P}(A_2)\mathbf{P}(A_3), \quad \mathbf{P}(A_2 \cap A_3 \cap A_4) = \mathbf{P}(A_2)\mathbf{P}(A_3)\mathbf{P}(A_4),\\
&\mathbf{P}(A_1 \cap A_3 \cap A_4) = \mathbf{P}(A_1)\mathbf{P}(A_3)\mathbf{P}(A_4), \quad \mathbf{P}(A_1 \cap A_2 \cap A_4) = \mathbf{P}(A_1)\mathbf{P}(A_2)\mathbf{P}(A_4),\\
&\mathbf{P}(A_1 \cap A_2) = \mathbf{P}(A_1)\mathbf{P}(A_2), \quad \mathbf{P}(A_1 \cap A_3) = \mathbf{P}(A_1)\mathbf{P}(A_3), \quad \mathbf{P}(A_1 \cap A_4) = \mathbf{P}(A_1)\mathbf{P}(A_4),\\
&\mathbf{P}(A_2 \cap A_4) = \mathbf{P}(A_2)\mathbf{P}(A_4), \quad \mathbf{P}(A_2 \cap A_3) = \mathbf{P}(A_2)\mathbf{P}(A_3), \quad \mathbf{P}(A_3 \cap A_4) = \mathbf{P}(A_3)\mathbf{P}(A_4).
\end{aligned}$$

Seega sündmuste A_1, A_2, A_3, A_4 sõltumatus tähendab, et sõltumatud on ka kõik sinna nelikusse kuuluvad paarid ja kolmikud. Nüüd on selge, et sõltumatud sündmused (olgu neid siis lõplik või lõpmatu hulk) on alati paarikaupa sõltumatud.

Paneme veel tähele, et loenduv hulk sündmusi $A_1, A_2, A_3, \ldots$ on sõltumatud parajasti siis, kui iga n korral on sõltumatud sündmused $A_1, \ldots, A_n$. Tõepoolest, definitsiooni kohaselt on sündmused $A_1, A_2, A_3, \ldots$ sõltumatud parajasti siis, kui suvaline lõplik alamhulk sündmusi on sõltumatud. Ja $A_1, \ldots, A_n$ on alati mingi lõplik alamhulk. Samas mitte kõik lõplikud n-elemendilised alamhulgad pole kujul $A_1, \ldots, A_n$. Kuidas jääb siis nende sõltumatusega? Kas näiteks sündmused A_3, A_5 ja A_{12} on ikka ka sõltumatud nagu definitsioon nõuab? Et aga kolmik A_3, A_5 ja A_{12} on sündmuste $A_1, A_2, \ldots, A_{12}$ alamkolmik, siis $A_1, A_2, \ldots, A_{12}$ sõltumatusest järeldub ka meid huvitava kolmiku sõltumatus. Seega loenduva hulga sündmuste $A_1, A_2, A_3, \ldots$ sõltumatuseks piisab tõepoolest, et sündmused $A_1, \ldots, A_n$ oleksid sõltumatud iga n korral.

5.2 Klasside sõltumatus

Olgu $\mathcal{A}_1, \ldots, \mathcal{A}_n$ σ-algebrasse $\mathcal{F}$ kuuluvad alamklassid, s.t. $\mathcal{A}_i \subset \mathcal{F}$, $i = 1, \ldots, n$ korral.

Def 5.2 *Klassid $\mathcal{A}_1, \ldots, \mathcal{A}_n$ on sõltumatud, kui iga $A_i \in \mathcal{A}_i$ korral sündmused $A_1, \ldots, A_n$ on sõltumatud.*

Kui klassid $\mathcal{A}_i$ on sellised, et $\Omega \in \mathcal{A}_i$, $i = 1, \ldots, n$ (näiteks, $\mathcal{A}_i$ on σ-algebra), siis on need klassid sõltumatud parajasti siis, kui iga $A_i \in \mathcal{A}_i$ korral kehtib

$$\mathbf{P}(A_1 \cap \cdots \cap A_n) = \mathbf{P}(A_1) \cdots \mathbf{P}(A_n). \tag{5.1}$$

Näitame seda. Sündmuste $A_1, \ldots, A_n$ sõltumatusest järeldub, (5.1). Teistpidi: olgu nüüd $A_{k_1}, \ldots, A_{k_j}$ sündmuste $A_1, \ldots, A_n$ mingi j-elemendiline ($j < n$) alamhulk. Meil on vaja näidata, et

$$\mathbf{P}(A_{k_1} \cap \cdots \cap A_{k_j}) = \mathbf{P}(A_{k_1}) \cdots \mathbf{P}(A_{k_j}).$$

Selleks defineerime sündmused $B_1, \ldots, B_n$ järgmiselt:

$$B_i = \begin{cases} A_i, & \text{kui } i = k_l \text{ mingi } l = 1, \ldots, j \text{ korral;} \\ \Omega, & \text{mujal.} \end{cases}$$

On selge, et iga $i = 1, \ldots, n$ korral $B_i \in \mathcal{A}_i$. Seosest (5.1) järeldub nüüd $\mathbf{P}(B_1 \cap \cdots \cap B_n) = \mathbf{P}(B_1) \cdots \mathbf{P}(B_n)$ ja seega

$$\mathbf{P}(A_{k_1} \cap \cdots \cap A_{k_j}) = \mathbf{P}(B_1 \cap \cdots \cap B_n) = \mathbf{P}(B_1) \cdots \mathbf{P}(B_n) = \mathbf{P}(A_{k_1}) \cdots \mathbf{P}(A_{k_j}).$$

Def 5.3 *Lõpmatu (loenduv või mitteloenduv) hulk klasse $\{\mathcal{A}_\alpha\}$ on sõltumatud, kui hulga $\{\mathcal{A}_\alpha\}$ iga lõplik alamhulk $\mathcal{A}_{\alpha_1}, \ldots, \mathcal{A}_{\alpha_n}$, $n > 1$ on sõltumatud.*

Seega on klassid $\{\mathcal{A}_\alpha\}$ sõltumatud parajasti siis, kui iga $A_\alpha \in \mathcal{A}_\alpha$ korral on hulgad $\{A_\alpha\}$ sõltumatud. Järelikult loenduv hulk klasse $\mathcal{A}_1, \mathcal{A}_2, \ldots$ on sõltumatud parajasti siis, kui iga valiku $A_i \in \mathcal{A}_i$ korral on sündmused $A_1, A_2, \ldots$ sõltumatud. Nagu alampeatükis 5.1 veendusime on viimane ekvivalentne sellega, et iga n korral on $A_1, A_2, \ldots, A_n$ sõltumatud. See aga tähendab, et iga n korral klassid $\mathcal{A}_1, \ldots, \mathcal{A}_n$ sõltumatud. Kui nüüd iga klass sisaldab hulka Ω, siis on viimane tingimus ekvivalentne sellega, et iga $n = 2, 3, \ldots$ ja iga sündmuste valiku $A_1 \in \mathcal{A}_1, \ldots, A_n \in \mathcal{A}_1$ korral kehtib (5.1). Sellised klassid, mis alati sisaldavad hulka Ω on näiteks σ-algebrad. Siit saame järelduse.

Järeldus 5.1 *Olgu $\mathcal{F}_1, \mathcal{F}_2, \ldots$ σ-algebra $\mathcal{F}$ alam-σ-algebrad (st $\mathcal{F}_i \subset \mathcal{F}$ iga i korral). Need σ-algebrad $\mathcal{F}_1, \mathcal{F}_2, \ldots$ on sõltumatud parajasti siis, kui iga $n = 2, 3, \ldots$ ja iga sündmuste valiku $A_i \in \mathcal{F}_i$, $i = 1, \ldots, n$ korral kehtib võrdus (5.1).*

5.2.1 σ-algebrate sõltumatus

Olgu $\mathcal{F}_1, \mathcal{F}_2, \ldots, \mathcal{F}_n$ σ-algebra $\mathcal{F}$ alam σ-algebrad. Teame, et nad on sõltumatud parajasti siis, kui võrdus (5.1) kehtib iga sündmuste valiku $A_i \in \mathcal{F}_i$ korral. Enamasti on aga σ-algebrad väga keerulised ja siis tuleb kontrollida väga palju võrdusi kujul (5.1). Järgnev lemma väidab aga, et $\mathcal{F}_1, \mathcal{F}_2, \ldots, \mathcal{F}_n$ on sõltumatud ka siis, kui (5.1) kehtib iga $A_i \in \mathcal{I}_i$ korral, kus $\mathcal{I}_i$ on π-süsteem, mis tekitab $\mathcal{F}_i$. Nagu ikka, kergendab see tulemus σ-algebrate sõltumatuse kontrollimist.

Lemma 5.1 *Olgu $\mathcal{I}_1, \ldots, \mathcal{I}_n$ σ-algebra $\mathcal{F}$ alam-π-süsteemid. Kui $\mathcal{I}_1, \ldots, \mathcal{I}_n$ on sõltumatud, siis alam-σ-algebrad $\sigma(\mathcal{I}_1), \ldots, \sigma(\mathcal{I}_n)$ on sõltumatud.*

Tõestus. Üldisust kitsendamata eeldame, et $\Omega \in \mathcal{I}_i$ (kui see nii pole, defineerime $J_i := \{\Omega, \mathcal{I}_i\}$. Süsteemid J_i on ka sõltumatud ja $\sigma(J_i) = \sigma(\mathcal{I}_i)$.)
1) Kõigepealt näitame, et

$$\sigma(\mathcal{I}_1), \mathcal{I}_2, \ldots, \mathcal{I}_n$$

on sõltumatud. Eelduse tõttu sisaldavad meie π-süsteemid hulka Ω ja seega piisab sõltumatuseks seosest (5.1). See tähendab seda, et iga $H \in \sigma(\mathcal{I}_1), I_2 \in \mathcal{I}_2, \ldots, I_n \in \mathcal{I}_n$ korral peame näitama, et kehtib

$$\mathbf{P}(H \cap I_2 \cap \cdots \cap I_n) = \mathbf{P}(H)\mathbf{P}(I_2) \cdots \mathbf{P}(I_n). \tag{5.2}$$

Fikseerime $I_j \in \mathcal{I}_j$, $j = 2, \ldots, n$ ning defineerime σ-algebral $\sigma(\mathcal{I}_1)$ mõõdud

$$\nu_1(H) = \mathbf{P}(H \cap I_2 \cap \cdots \cap I_n), \quad \mu_1(H) = \mathbf{P}(H)\mathbf{P}(I_2) \cdots \mathbf{P}(I_n).$$

(Veendu, et ν_1 ja μ_1 on ikka mõõdud. Kas nad on tõenäosusmõõdud?)
Kehtib:

$$\nu_1(\Omega) = \mu_1(\Omega) \text{ ja } \nu_1|_{\mathcal{I}_1} = \mu_1|_{\mathcal{I}_1}.$$

Viimane võrdus tähendab, et $\nu_1(I) = \mu_1(I)$ iga $I \in \mathcal{I}_1$ korral. Teoreemist 2.3 järeldub, et $\nu_1 = \mu_1$ ehk (5.2).
2) Veendume, et

$$\sigma(\mathcal{I}_1), \sigma(\mathcal{I}_2), \mathcal{I}_3, \ldots, \mathcal{I}_n$$

on sõltumatud. Vastavalt seosele (5.1) piisab, kui iga $H_1 \in \sigma(\mathcal{I}_1), H \in \sigma(\mathcal{I}_2), I_j \in \mathcal{I}_j, j = 3, \ldots, n$ korral kehtib

$$\mathbf{P}(H_1 \cap H \cap I_3 \cap \cdots \cap I_n) = \mathbf{P}(H_1)\mathbf{P}(H)\mathbf{P}(I_3) \cdots \mathbf{P}(I_n). \tag{5.3}$$

Fikseeri $H_1 \in \sigma(\mathcal{I}_1), I_3 \in \mathcal{I}_3, \ldots, I_n \in \mathcal{I}_n$ ja defineeri mõõdud

$$\nu_2(H) = \mathbf{P}(H_1 \cap H \cap I_3 \cap \cdots \cap I_n), \quad \mu_2(H) = \mathbf{P}(H_1)\mathbf{P}(H)\mathbf{P}(I_3) \cdots \mathbf{P}(I_n).$$

Et $\sigma(\mathcal{I}_1), \mathcal{I}_2, \ldots, \mathcal{I}_n$ on sõltumatud, siis

$$\nu_2(\Omega) = \mu_2(\Omega) \text{ ja } \nu_2|_{\mathcal{I}_2} = \mu_2|_{\mathcal{I}_2}.$$

Teoreemist 2.3 järeldub, et $\nu_2 = \mu_2$ ehk (5.3).
3) Analoogiliselt tõesta, et

$$\sigma(\mathcal{I}_1), \sigma(\mathcal{I}_2), \sigma(\mathcal{I}_3), \mathcal{I}_4, \ldots, \mathcal{I}_n$$

on sõltumatud jne. ■

Järeldus 5.2 *Olgu $\{\mathcal{I}_\alpha\}$ sõltumatud π-süsteemid. Siis $\{\sigma(\mathcal{I}_\alpha)\}$ on sõltumatud σ-algebrad.*

5.3 Juhuslike suuruste sõltumatus

Olgu $X_1, X_2, \ldots$ ruumil $(\Omega, \mathcal{F}, \mathbf{P})$ antud juhuslikud suurused.

Def 5.4

- *Juhuslikud suurused $X_1, \ldots, X_n$ on sõltumatud, kui σ-algebrad $\sigma(X_1), \ldots, \sigma(X_n)$ on sõltumatud.*
- *Juhuslikud suurused $X_1, X_2, X_3, \ldots$ on sõltumatud, kui σ-algebrad $\sigma(X_1)$, $\sigma(X_2), \sigma(X_3) \ldots$ on sõltumatud.*

Seega lõplik hulk juhuslikke suurusi $X_1, \ldots, X_n$ on sõltumatud parajasti siis, kui iga sündmuste valiku $A_i \in \sigma(X_i)$, $i = 1, \ldots, n$ korral kehtib (5.1) ning järelduse 5.1 tõttu on lõpmatu hulk juhuslikke suurusi $X_1, X_2, X_3, \ldots$ sõltumatud parajasti siis, kui iga n ja iga valiku $A_i \in \sigma(X_i)$, $i = 1, \ldots, n$ korral kehtib (5.1). Nüüd tuletame meelde, et

$$\sigma(X) = \{X^{-1}(B) : B \in \mathcal{B}(\mathbb{R})\}.$$

See tähendab, et kui $A \in \sigma(X)$, siis leidub Boreli hulk B nii, et $A = X^{-1}(B)$. Nüüd on kerge näha, et lõplik hulk juhuslikke suurusi $X_1, \ldots, X_n$ sõltumatud parajasti siis kui suvaliste Boreli hulkade $B_1, \ldots, B_n \in \mathcal{B}$ korral kehtib võrdus

$$\mathbf{P}(X_1 \in B_1, \ldots, X_n \in B_n) = \mathbf{P}(X_1 \in B_1) \cdots \mathbf{P}(X_n \in B_n). \tag{5.4}$$

Lõpmatu hulk juhuslikke suurusi X_1, X_2, ... on aga sõltumatud parajasti siis, kui (5.4) kehtib iga naturaalarvu n ja suvaliste Boreli hulkade $B_1, \ldots, B_n \in \mathcal{B}$ korral.

Pane samuti tähele, et lõpmatu hulk juhuslikke suurusi $X_1, X_2, \ldots$ on sõltumatud parajasti siis, kui iga n korral on $X_1, \ldots, X_n$ sõltumatud juhuslikud suurused.

Sõltumatus jaotusfunktsiooni kaudu. Olgu $\mathcal{I}$ reaaltelje alamhulkade π-süsteem mis tekitab $\mathcal{B}(\mathbb{R})$, s.t. $\sigma(\mathcal{I}) = \mathcal{B}(\mathbb{R})$. Iga juhusliku suuruse $X : \Omega \to \mathbb{R}$ korral klass

$$\{X^{-1}(I) : I \in \mathcal{I}\}$$

on π-süsteem, mis tekitab $\sigma(X)$ (peatükk 4, ülesanne 4 d)). Seega on juhuslikud suurused X ja Y sõltumatud, kui iga $I_1, I_2 \in \mathcal{I}$ korral kehtib

$$\mathbf{P}(X \in I_1, Y \in I_2) = \mathbf{P}(X^{-1}(I_1) \cap Y^{-1}(I_2)) = \mathbf{P}(X^{-1}(I_1))\mathbf{P}(Y^{-1}(I_2))) = \mathbf{P}(X \in I_1)\mathbf{P}(Y \in I_2). \tag{5.5}$$

Võttes nüüd

$$\mathcal{I} = \pi(\mathbb{R}) = \{(-\infty, x] : x \in \mathbb{R}\}$$

saame, et juhuslikud suurused X ja Y sõltumatud, kui iga $x, y \in \mathbb{R}$ korral

$$\begin{aligned} \mathbf{P}(X \leq x, Y \leq y) &= \mathbf{P}(X \in (-\infty, x], Y \in (-\infty, y]) \\ &= \mathbf{P}(X \in (-\infty, x])\mathbf{P}(Y \in (-\infty, y]) \\ &= \mathbf{P}(X \leq x)\mathbf{P}(Y \leq y). \end{aligned}$$

Olgu $Z = (X, Y)$ juhuslik vektor. Vektori Z jaotusfunktsioon

$$F_Z : \mathbb{R}^2 \to [0, 1], \quad F_Z(x, y) = \mathbf{P}(X \leq x, Y \leq y).$$

Seosest (5.5) järeldub, et X ja Y on sõltumatud parajasti siis, kui iga $x, y \in \mathbb{R}$ korral kehtib

$$F_Z(x, y) = F_X(x)F_Y(y).$$

Vaatleme nüüd enam kui kahte juhuslikku suurust $X_1, \ldots, X_n$. Täpselt nii nagu kahe juhusliku suuruse korral järeldub lemmast 5.1, et $X_1, \ldots, X_n$ sõltumatuseks piisab π-süsteemide

$$\{X_1^{-1}(-\infty, x] : x \in \mathbb{R}\}, \ldots, \{X_n^{-1}(-\infty, x] : x \in \mathbb{R}\}$$

sõltumatusest. Need klassid ei pruugi sisaldada hulka Ω, mistõttu definitsiooni kohaselt on nad sõltumatud parajasti siis, kui kui iga indeksite kombinatsiooni $1 \leq i_1 < i_2 < \ldots < i_k \leq n$ ja reaalarvude $x_{i_j} \in \mathbb{R}$, $j = 1, \ldots, k$ korral kehtib

$$\mathbf{P}(X_{i_1} \leq x_{i_1}, \ldots, X_{i_k} \leq x_{i_k}) = \mathbf{P}(X_{i_1} \leq x_{i_1}) \cdots \mathbf{P}(X_{i_k} \leq x_{i_k}). \tag{5.6}$$

Tegelikult on aga (5.6) samaväärne järgmise seosega: iga $x_n \in \mathbb{R}$ korral kehtib

$$\mathbf{P}(X_1 \leq x_1, \ldots, X_n \leq x_n) = \mathbf{P}(X_1 \leq x_1) \cdots \mathbf{P}(X_n \leq x_n). \tag{5.7}$$

Järeldus 5.3 *Juhuslikud suurused $X_1, X_2, \ldots$ on sõltumatud parajasti siis, kui iga naturaalarvu n ja $x_1, \ldots, x_n \in \mathbb{R}$ korral kehtib (5.7).*

Tõestus. Seosest (5.6) järeldub (5.7).
Vastupidise tõestamiseks paneme tähele, et iga $1 \leq i_1 < i_2 < \cdots < i_k \leq n$ korral kehtib $(n \to \infty)$

$$\{X_{i_j} \leq x_{i_j}, j = 1, \ldots, k, X_l \leq n, l = 1, \ldots, n, l \neq i_j, j = 1, \ldots, k\} \nearrow \{X_{i_1} \leq x_{i_1}, \ldots, X_{i_k} \leq x_{i_k}\}$$

Mõõdu pidevusest järeldub

$$\mathbf{P}(X_{i_j} \le x_{i_j}, j = 1, \ldots, k, X_l \le n, l = 1, \ldots, n, l \ne i_j, j = 1, \ldots, k) \nearrow \mathbf{P}(X_{i_1} \le x_{i_1}, \ldots, X_{i_k} \le x_{i_k}). \tag{5.8}$$

Teisest küljest, iga i korral $\mathbf{P}(X_i \le n) \nearrow 1$. Seega

$$\prod_{j=1}^{k} \mathbf{P}(X_{i_j} \le x_{i_j}) \prod_{l \ne i_j} \mathbf{P}(X_l \le n) \nearrow \prod_{j=1}^{k} \mathbf{P}(X_{i_j} \le x_{i_j}). \tag{5.9}$$

Seosest (5.7) järeldub, et iga n korral (5.8) võrdub (5.9). Kui koonduvate jadade elemendid on võrdsed, siis on võrdsed ka piirväärtused. Seega kehtib (5.6). ■

Järeldus 5.4 *Olgu F_X juhusliku vektori $X = (X_1, \ldots, X_n)$ jaotusfunktsioon, F_{X_i} olgu juhusliku suuruse X_i jaotusfunktsioon, $i = 1, \ldots, n$. Juhuslikud suurused $X_1, \ldots, X_n$ on sõltumatud parajasti siis kui suvaliste $x_1, \ldots, x_n \in \mathbb{R}$ korral kehtib*

$$F_X(x_1, \ldots, x_n) = F_{X_1}(x_1) \cdots F_{X_n}(x_n). \tag{5.10}$$

Tõestus. Järeldub seosest (5.7) ■

5.4 Juhuslike vektorite sõltumatus

Juhuslike vektorite $X = (X_1, \ldots, X_k)$ ja $Y = (Y_1, \ldots, Y_l)$ sõltumatus defineeritakse täpselt samamoodi kui juhuslike suuruste sõltumatuski: X ja Y on sõltumatud, kui $\sigma(X)$ ja $\sigma(Y)$ on sõltumatud. Ja nii defineeritakse ka suvalise hulga (lõpliku või lõpmatu) juhuslike vektorite sõltumatus: nende tekitatud σ-algebrad peavad olema sõltumatud.

Et vektori $X = (X_1, \ldots, X_k)$ korral $\sigma(X) = \{X^{-1}(B) : B \in \mathcal{B}(\mathbb{R}^k)\}$, siis analoogiliselt juhuslike suurustega saame, et vektorid $X = (X_1, \ldots, X_k)$ ja $Y = (Y_1, \ldots, Y_l)$ on sõltumatud parajasti siis, kui suvaliste Boreli hulkade valiku $B_1 \in \mathcal{B}(\mathbb{R}^k)$ ja $B_2 \in \mathcal{B}(\mathbb{R}^l)$ korral

$$\mathbf{P}(X \in B_1, Y \in B_2) = \mathbf{P}(X \in B_1)\mathbf{P}(Y \in B_2).$$

Lemmast 5.1 järeldub ka nüüd, et X ning Y on sõltumatud parajasti siis kui $(k+l)$-dimensionaalse juhusliku vektori (X, Y) jaotusfunktsioon on X jaotusfunktsiooni F_X ja ja Y jaotusfunktsiooni F_Y korrutis: iga $x = (x_1, \ldots, x_k)$ ja $y = (y_1, \ldots, y_l)$ korral

$$\begin{aligned} F_{X,Y}(x, y) &= \mathbf{P}(X_1 \le x_1, \ldots, X_k \le x_k, Y_1 \le y_1, \ldots Y_l \le y_l) \\ &= \mathbf{P}(X_1 \le x_1, \ldots, X_k \le x_k)\mathbf{P}(Y_1 \le y_1, \ldots Y_l \le y_l) = F_X(x)F_Y(y). \end{aligned}$$

Pane tähele, et vektorite X ja Y sõltumatus ei tähenda, et X komponendid, näiteks juhuslikud suurused X_1 ja X_2 oleksid sõltumatud. Küll aga tähendab see, et suvaline X komponent ja suvaline Y komponent on sõltumatud. Näiteks on X_2 ja Y_1 sõltumatud juhuslikud suurused (miks?)

5.5 Ülesanded

Olgu $(\Omega, \mathcal{F}, \mathbf{P})$ tõenäosusruum.

1. Olgu $A_1, \cdots, A_n \in \mathcal{F}$ sündmused. Vaatleme hulki $\{B_1, \cdots, B_n\}$, kus $B_i = A_i$ või $B_i = A_i^c$. Tõestada, et kui iga sellise hulga korral $\mathbf{P}(B_1 \cap \cdots \cap B_n) = \prod_{i=1}^n \mathbf{P}(B_i)$, siis sündmused $A_1, \cdots, A_n$ on sõltumatud.

2. Olgu $A_1, A_2, \ldots \in \mathcal{F}$ sõltumatud sündmused. Tõestada, et

$$\mathbf{P}(\bigcap_{n=1}^{\infty} A_n) = \prod_{n=1}^{\infty} \mathbf{P}(A_n).$$

3. Olgu $(\Omega, \mathcal{F}, \mathbf{P}) = ([0,1], \mathcal{B}, Leb)$. Vaatleme sündmusi

$$\begin{aligned}
A_1 &= (0, 2/8] \cup (6/8, 1], \\
A_2 &= (0, 1/8] \cup (2/8, 3/8] \cup (5/8, 6/8] \cup (7/8, 1], \\
A_3 &= (0, 1/8] \cup (3/8, 5/8] \cup (7/8, 1].
\end{aligned}$$

Tõestada, et A_1, A_2, A_3 on paarikaupa sõltumatud. Kas need sündmused on sõltumatud?

4. Olgu $\Omega = \{1, 2, 3, 4, 5, 6\}$, $p(i) = 1/6$. Vaatleme sündmusi: $A_1 = \{1, 2, 3, 4\}$, $A_2 = A_3 = \{4, 5, 6\}$. Kas need sündmused on paarikaupa sõltumatud? Kas need sündmused on sõltumatud?

5. Olgu $(\Omega, \mathcal{F}, \mathbf{P})$ selline tõenäosusruum, et iga sündmuse A jaoks kõigi temaga sõltumatute sündmuste hulk moodustab algebra. Olgu $\mathcal{D} \subset \mathcal{F}$ mingi paarikaupa sôltumatute sündmuste hulk. Tõestada, et klassi $\mathcal{D}$ kuuluvad sündmused on sõltumatud.

6. Tõestada, kui sündmuse A tõenäosus on 0 või 1, siis on ta sõltumatu mistahes teisest sündmusest. Järeldada, et tõenäosusruumi kõik sündmused on paarikaupa sõltumatud parajasti siis, kui iga sündmuse mõõt on 0 või 1.

7. Olgu A ja B sõltumatud sündmused. Tõestada, et järgmised sündmuste paarid on sõltumatud: $A^c, B \quad B^c, A \quad A^c, B^c$.

8. Tõestada, et sündmused $A_1, A_2, \ldots$ on sõltumatud parajasti siis, kui σ - algebrad $\sigma(A_1), \sigma(A_2), \ldots$ on sõltumatud.

9. Vaatleme ülesandes 3 konstrueeritud tõenäosusruumi ja sündmusi A_1, A_2, A_3. Moodustame klassid $\mathcal{A}_1 = \{A_1, A_2\}$ ja $\mathcal{A}_2 = \{A_3\}$. Tõestada, et
1) klassid $\mathcal{A}_1$ ja $\mathcal{A}_2$ on sõltumatud
2) klassid $\sigma(\mathcal{A}_1)$ ja $\sigma(\mathcal{A}_2)$ pole sõltumatud.

Kas tulemus 2) läheb vastuollu lemmaga 5.1? Miks?

10. Olgu X ja Y diskreetsed juhuslikud suurused, X väärtused olgu $\{x_1, x_2, \ldots\}$ ja Y väärtused olgu $\{y_1, y_2, \ldots\}$. Näita, et X ja Y on sõltumatud parajasti siis, kui iga paari (x_i, y_j) korral kehtib võrdus

$$\mathbf{P}(X = x_i, Y = y_j) = \mathbf{P}(X = x_i)\mathbf{P}(Y = y_j).$$

11. Tõestada, et kahe sõltumatu σ-algebra ühend ei saa olla σ-algebra, kui mõlemad esialgsed σ-algebrad sisaldavad sündmusi, mille mõõt erineb nullist või ühest.

12. Tõestada, et sündmused $A_1, A_2, \ldots$ on sõltumatud parajasti siis, kui juhuslikud suurused $I_{A_1}, I_{A_2}, \ldots$ on sõltumatud.

13. Tõestada, et juhuslikud suurused $X_1, X_2, \ldots$ on sõltumatud, kui $\sigma(X_1, \ldots, X_{n-1})$ ja $\sigma(X_n)$ on sõltumatud iga n korral.

14. Olgu $\Omega = \{1, 2, 3, 4\}$, $\mathcal{F} = 2^{\Omega}$, $\mathbf{P}(i) = 1/4$. Kas X ja Y on sõltumatud, kui

$$X(\omega) = \begin{cases} 0 & \text{kui } \omega = 1, 2 \\ 1 & \text{mujal.} \end{cases}, \quad Y(\omega) = \begin{cases} 1 & \text{kui } \omega = 1, 4 \\ 0 & \text{mujal.} \end{cases}$$

15. Tõestada, et juhuslik suurus ei sõltu iseendast parajasti siis kui ta on p.k. konstant.

16. Olgu X ja Y juhuslikud suurused. Kas nad on sõltumatud, kui X^2 ja Y^2 on sõltumatud?

17. Olgu $(\Omega, \mathcal{F}, \mathbf{P}) = ([0, 1], \mathcal{B}, Leb)$. Kas juhuslikud suurused X ja Y on sõltumatud, kui

1) $X(\omega) = \omega^2, \quad Y(\omega) = 1 - \omega^2$;
2) $X(\omega) = 1/2, \quad Y(\omega) = \omega$?

18. Olgu X ja Y sõltumatud juhuslikud suurused, f ja g olgu Boreli funktsioonid. Kas $f(X)$ ja $g(Y)$ on sõltumatud juhuslikud suurused?

19. Olgu $(\Omega, \mathcal{F}, \mathbf{P}) = ([0, 1], \mathcal{B}, Leb)$. Kas sellel tõenäosusruumil leidub mittetrivialne juhuslik suurus, mis oleks sõltumatu juhuslikust suurusest $X(\omega) = \omega$?

20. Olgu $X_1, X_2, \ldots, X_n$ sõltumatud ja sama jaotusega juhuslikud suurused. Olgu π hulga $\{1, 2, \ldots, n\}$ mingi permutatsioon. Tõestada, et vektorid $(X_{\pi 1}, X_{\pi 2} \ldots, X_{\pi n})$ ja $(X_1, \ldots, X_n)$ on sama jaotusega.

21. Olgu X ja Y sõltumatud sama jaotusega juhuslikud suurused. Näita, et iga Boreli funktsiooni $f : \mathbb{R}^2 \to \mathbb{R}$ korral $f(X, Y)$ ja $f(Y, X)$ on sama jaotusega. Järeldada, et juhusliku suuruse $X - Y$ jaotus on sümmeetriline, st $X - Y$ ja $Y - X$ on sama jaotusega.

6 Boreli-Cantelli lemmad ja Kolmogorovi 0-1 seadus

6.1 Boreli-Cantelli lemmad

Olgu $(\Omega, \mathcal{F}, \mathbf{P})$ tõenäosusruum, $A_n \in \mathcal{F}$ on sündmused.

Lemma 6.1 (Boreli-Cantelli 1. lemma (BC1)) *Kui*

$$\sum_{n=1}^{\infty} \mathbf{P}(A_n) < \infty,$$

siis

$$\mathbf{P}(\limsup_n A_n) = \mathbf{P}\{\omega : \omega \in A_n, \ \text{i.o.}\} = 0.$$

Lemma 6.2 (Boreli-Cantelli 2. lemma (BC2)) *Kui A_n on sõltumatud ja*

$$\sum_{n=1}^{\infty} \mathbf{P}(A_n) = \infty,$$

siis

$$\mathbf{P}(\limsup_n A_n) = \mathbf{P}\{\omega : \omega \in A_n, \ \text{i.o.}\} = 1.$$

Tõestus. $\mathbf{P}(\limsup_n A_n) = 1$ parajasti siis, kui

$$\mathbf{P}(\liminf_n A_n^c) = 0. \tag{6.1}$$

Tõestame (6.1). Definitsioon:

$$\liminf_n A_n^c = \cup_{n=1}^{\infty} \cap_{m=n}^{\infty} A_m^c.$$

Seega (6.1) kehtib, kui

$$\mathbf{P}(\cap_{m=n}^{\infty} A_m^c) = 0, \quad n = 1, 2, \ldots. \tag{6.2}$$

Sündmused $A_1^c, A_2^c, \ldots$ on sõltumatud. Seega iga n ja k korral

$$\mathbf{P}\Big(\bigcap_{m=n}^{n+k} A_m^c\Big) = \prod_{m=n}^{n+k} \mathbf{P}(A_m^c) = \prod_{m=n}^{n+k} \Big(1 - \mathbf{P}(A_m)\Big) \leq \exp[-\sum_{m=n}^{n+k} \mathbf{P}(A_m)],$$

sest $1 - x \leq e^{-x}$ iga x korral.
Et $\sum_{n=1}^{\infty} \mathbf{P}(A_n) = \infty$, siis $\sum_{m=n}^{n+k} \mathbf{P}(A_m) \nearrow \infty$, kui $k \to \infty$, millest

$$\exp[-\sum_{m=n}^{n+k} \mathbf{P}(A_m)] \to 0,$$

kui $k \to \infty$.
Teisest küljest

$$\mathbf{P}(\cap_{m=n}^{n+k} A_m^c) \searrow \mathbf{P}(\cap_{m=n}^{\infty} A_m^c),$$

millest järeldub (6.2). ■

Järeldus 6.1 *Kui A_n on sõltumatud ja $0 < \mathbf{P}(A_1) = \mathbf{P}(A_2) = \cdots$, siis*

$$\mathbf{P}(\limsup_n A_n) = \mathbf{P}\{\omega : \omega \in A_n, \text{ i.o.}\} = 1.$$

Seega, kui sündmused $A_1, A_2, \ldots$ on sõltumatud, siis sündmuse $\limsup_n A_n$ tõenäosus sõltub reast $\sum_n \mathbf{P}(A_n)$. Kui see rida koondub, on $\mathbf{P}(\limsup_n A_n) = 0$, vastasel juhul on see tõenäosus 1. Sõltuvate sündmuste puhul see muidugi nii ei ole: BC1 küll kehtib, kuid BC2 mitte. Näiteks kui A on selline sündmus, et $0 < \mathbf{P}(A) < 1$, $A_n = A$ iga n korral, siis $\sum_n \mathbf{P}(A_n) = \infty$ aga $\mathbf{P}(\limsup_n A_n) = \mathbf{P}(A) < 1$.

6.2 Näited Boreli-Cantelli lemmade kasutamisest

6.2.1 Kordse logaritmi seadus

Olgu $X_1, X_2, \ldots$ tõenäosusruumil $(\Omega, \mathcal{F}, \mathbf{P})$ antud sõltumatud juhuslikud suurused.
Olgu ψ_n, $n = 1, 2, \ldots$ mingi jada. Tihti huvitab meid tõenäosus

$$\mathbf{P}(X_n > \psi_n, \text{ i.o.}) = \mathbf{P}\{\omega : X_n(\omega) > \psi_n, \text{ i.o.}\}. \tag{6.3}$$

On selge, et

$$\{\omega : X_n(\omega) > \psi_n, \text{ i.o.}\} = \limsup_n \{\omega : X_n(\omega) > \psi_n\}.$$

Et X_n on mõõtuv, siis $\{X_n > \psi_n\} \in \mathcal{F}$.
Et $X_1, X_2, X_3, \ldots$ on sõltumatud, siis on sõltumatud ka sündmused $\{X_1 > \psi_1\}, \{X_2 > \psi_2\}, \ldots$.
Seega:

- BC1: $\mathbf{P}(X_n > \psi_n, \text{ i.o.}) = 0$, kui

$$\sum_{n=1}^{\infty} \mathbf{P}(X_n > \psi_n) = \sum_{n=1}^{\infty} \Big(1 - F_n(\psi_n)\Big) < \infty$$

- BC2: $\mathbf{P}(X_n > \psi_n, \text{ i.o.}) = 1$, kui

$$\sum_{n=1}^{\infty} \mathbf{P}(X_n > \psi_n) = \sum_{n=1}^{\infty} \Big(1 - F_n(\psi_n)\Big) = \infty.$$

Siin F_n on X_n jaotusfunktsioon.

Näide: Olgu $X_1, X_2, X_3, \ldots$ sõltumatud ja sama jaotusega (i.i.d.). Olgu $X_i \sim E(1)$, s.t.

$$F_n(t) = 1 - e^{-t}.$$

Meid huvitab selline jada ψ_n, et

$$\mathbf{P}(X_n > \psi_n, \text{ i.o.}) = 0 \text{ ehk } \mathbf{P}(X_n \leq \psi_n, \text{ ev.}) = 1. \tag{6.4}$$

Ülaltoodust on selge, et seost (6.4) rahuldav jada ei saa olla konstantne: iga $c \in \mathbb{R}$ korral $\mathbf{P}(X_n > c, \text{ i.o.}) = 1$ (ülesanne 3a).

Vaatleme jada $\psi_n = \alpha \ln n$. On kerge veenduda (ülesanne 3b), et

$$\mathbf{P}(X_n > \psi_n, \text{ i.o.}) = \mathbf{P}(X_n > \alpha \ln n, \text{ i.o.}) = \begin{cases} 0, & \text{kui } \alpha > 1; \\ 1, & \text{kui } \alpha \leq 1. \end{cases}$$

Seega jada $\psi_n = \alpha \ln n$ on ülemine piir iga $\alpha > 1$ korral, kuid $\psi_n = \ln n$ pole.

Veel enam. Selgub, et

$$\mathbf{P}(X_n > \ln n + \alpha \ln \ln n, \text{ i.o.}) = \begin{cases} 0, & \text{kui } \alpha > 1; \\ 1, & \text{kui } \alpha \leq 1. \end{cases}$$

Ülaltoodud seos täpsustab piiri.
Veel enam, kehtib

$$\mathbf{P}(X_n > \ln n + \ln \ln n + \alpha \ln \ln \ln n, \text{ i.o.}) = \begin{cases} 0, & \text{kui } \alpha > 1; \\ 1, & \text{kui } \alpha \leq 1. \end{cases}$$

jne.

6.2.2 Koondumine

Olgu $Y, X_1, X_2, \ldots$ juhuslikud suurused. Meid huvitab tõenäosus

$$\mathbf{P}(X_n \to Y) = \mathbf{P}\{\omega : \lim_n X_n(\omega) = Y(\omega)\}.$$

Kõigepealt paneme tähele, et hulk $\{\omega : \lim_n X_n(\omega) = Y(\omega)\}$ on sündmus (s.t. $\mathcal{F}$ element). See järeldub seostest

$$\{\omega : \lim_n X_n(\omega) = Y(\omega)\} = \{\omega : \liminf_n X_n(\omega) = Y(\omega)\} \cap \{\omega : \limsup_n X_n(\omega) = Y(\omega)\}.$$

Järgmine väide on oluline:
Kui $\forall \epsilon > 0$ kehtib

$$\sum_{n=1}^{\infty} \mathbf{P}(|X_n - Y| > \epsilon) < \infty,$$

siis BC1:

$$\mathbf{P}(|X_n - Y| > \epsilon, \text{ i.o.}) = 0,$$

millest

$$\mathbf{P}(X_n \to Y) = 1.$$

6.3 Kolmogorovi 0-1 seadus

Nägime, et sõltumatute juhuslike suuruste $X_1, X_2, \ldots$ korral sündmuse $(X_n > \psi_n,\ \text{i.o.})$ tõenäosus on kas 0 või 1. See pole juhus.

Def 6.1 *Olgu $X_1, X_2, \ldots$ juhuslike suuruste jada,*

$$\mathcal{T}_n := \sigma(X_n, X_{n+1}, \ldots) \text{ ja } \mathcal{T} := \cap_n \mathcal{T}_n.$$

σ-algebra $\mathcal{T}$ on jada $\{X_n\}$ **jääk σ-algebra**.

Sisuliselt kuuluvad jada jääk σ-algebrasse need sündmused, mis ei sõltu jada algusest. Näiteks

$$F := (X_n \in B,\ \forall n) \notin \mathcal{T}.$$

Näited sündmustest, mis kuuluvad jääk-σ-algebrasse:

- $F_1 := \{\omega | \exists \lim X_k(\omega)\} = (\lim X_k \quad \text{eksisteerib});$
- $F_2 := (\text{rida} \quad \sum X_k \quad \text{koondub});$
- $F_3 := (\lim \frac{X_1 + X_2 + \cdots + X_k}{k} \quad \text{eksisteerib});$
- $F_4 := (X_n > \psi_n, \quad \text{i.o.});$
- $F_5 := (\limsup_n X_n < \infty);$
- $F_6 := (\limsup_n \frac{X_1 + \cdots + X_n}{\sqrt{2n \ln n}} = 1).$

Veendume, et $F_1 \in \mathcal{T}$. Lemmast 4.5 e) järeldub, et

$$F_1 \in \sigma(X_1, X_2, \ldots) = \mathcal{T}_1.$$

Vaatleme jada $X_2, X_3, \ldots$. Sellel jadal on piirväärtus parajasti siis, kui jadal $X_1, X_2, \ldots$ on piirväärtus. Seega

$$F_1 \in \sigma(X_2, X_3, \ldots) = \mathcal{T}_2$$

jne. Seega $F_1 \in \mathcal{T}_n$ iga $n = 1, 2, \ldots$ korral, millest $F_1 \in \mathcal{T}$.

Veendume, et $F_4 \in \mathcal{T}$. Tõestus on analoogiline, sest

$$(X_n > \psi_n, \quad \text{i.o.},\ n \geq 1) = (X_n > \psi_n, \quad \text{i.o.},\ n \geq 2) = (X_n > \psi_n, \quad \text{i.o.},\ n \geq 3) = \cdots .$$

Iga $k \geq 1$ korral

$$(X_n > \psi_n, \quad \text{i.o.},\ n \geq k) \in \mathcal{T}_k$$

(miks?). Seega $F_4 \in \cap_k \mathcal{T}_k$.

Kolmogorovi 0-1 seadus ütleb, et juhul kui $X_1, X_2, \ldots$ on sõltumatud, koosneb jääk σ-algebra vaid sellistest elementidest, mille mõõt on vaid 0 või 1.

Teoreem 6.2 (Kolmogorovi 0-1 seadus). *Olgu $X_1, X_2, \ldots$ sõltumatute juhuslike suuruste jada, $\mathcal{T}$ olgu jada $\{X_n\}$ jääk σ-algebra. Kui $A \in \mathcal{T}$, siis $\mathbf{P}(A) = 0$ või $\mathbf{P}(A) = 1$.*

Tõestus. Defineerime klassid

$$\begin{aligned}
&\mathcal{K}_n := \{\{\omega | X_i(\omega) \leq x_i, \quad i = 1, \ldots, n-1\}, \, x_i \in (-\infty, \infty]\},\\
&\mathcal{I}_n := \{\{\omega | X_i(\omega) \leq x_i, \, i = n, \ldots, n+r\} | \; r \in \mathbf{N}, \, x_i \in (-\infty, \infty]\},\\
&\chi_n := \sigma(X_1, \ldots, X_{n-1}).
\end{aligned}$$

Klassi $\mathcal{K}_n$ tüüpiline element on kujul

$$\{X_1 \leq x_1\} \cap \{X_2 \leq x_2\} \cap \cdots \cap \{X_{n-1} \leq x_{n-1}\},$$

kusjuures x_i võib olla ∞. Sel juhul ülemine sündmus on kujul

$$\{X_1 \leq x_1\} \cap \{X_2 \leq x_2\} \cap \cdots \cap \{X_{i-1} \leq x_{i-1}\} \cap \{X_{i+1} \leq x_{i+1}\} \cap \cdots \cap \{X_{n-1} \leq x_{n-1}\}.$$

Klassi $\mathcal{I}_n$ tüüpiline element on kujul

$$\{X_n \leq x_n\} \cap \{X_{n+1} \leq x_{n+1}\} \cap \cdots \cap \{X_{n+r} \leq x_{n+r}\},$$

kusjuures x_{n+i} võib olla ∞.
Veendu (ülesanne 7), et

- $\mathcal{K}_n$ ja $\mathcal{I}_n$ on π-süsteemid
- $\sigma(\mathcal{K}_n) = \chi_n$ ja $\sigma(\mathcal{I}_n) = \mathcal{T}_n$

Et $X_1, X_2, \ldots$ on sõltumatud, siis iga $A \in \mathcal{K}_n$ ja $B \in \mathcal{I}_n$ korral $\mathbf{P}(A \cap B) = \mathbf{P}(A)\mathbf{P}(B)$, millest

$$\underline{\mathcal{K}_n \text{ ja } \mathcal{I}_n \text{ on sõltumatud.}}$$

Lemmast 5.1 järeldub, et

$$\underline{\chi_n \quad \text{ja } \mathcal{T}_n \text{ on sõltumatud.}}$$

Olgu $A \in \chi_n$ ja $B \in \mathcal{T}$. Siis $B \in \mathcal{T}_n$, millest A ja B on sõltumatud. Järelikult

$$\underline{\chi_n \quad \text{ja } \mathcal{T} \text{ on sõltumatud.}}$$

Olgu

$$\chi_\infty = \sigma\Big(\bigcup_{n=1}^{\infty} \chi_n\Big).$$

$\bigcup_{n=1}^{\infty} \chi_n$ on π-süsteem (miks?). Et χ_n ja $\mathcal{T}$ on sõltumatud iga n korral, saame

$$\underline{\bigcup_{n=1}^{\infty} \chi_n \text{ ja } \mathcal{T} \text{ on sõltumatud.}}$$

Lemmast 5.1 järeldub, et

$$\underline{\chi_\infty \quad \text{ja } \mathcal{T} \text{ on sõltumatud.}}$$

Tõestame, et

$$\chi_\infty = \sigma(X_1, X_2, \ldots).$$

Et $\chi_n \subset \sigma(X_1, X_2, \ldots)$, siis $\chi_\infty \subset \sigma(X_1, X_2, \ldots)$ (miks?).
Teistpidi. Olgu $B \in \mathcal{B}$, n suvaline. On selge et

$$X_n^{-1}(B) \in \chi_{n+1} \subset \bigcup_{n=1}^{\infty} \chi_n \subset \chi_\infty.$$

Seega $X_n \in m\chi_\infty$. Kehtib iga n korral, ehk $\sigma(X_1, X_2, \ldots) \subset \chi_\infty$. Kokkuvõttes

$$\underline{\sigma(X_1, X_2, \ldots) \quad \text{ja } \mathcal{T} \text{ on sõltumatud.}}$$

Teisest küljest $\mathcal{T} \subset \sigma(X_1, X_2, \ldots)$. Seega

$$\underline{\mathcal{T} \quad \text{ja } \mathcal{T} \text{ on sõltumatud.}}$$

Et $\mathcal{T}$ on sõltumatu iseendast, sisaldab ta ainult neid elemente, mille mõõt on vaid 0 või 1. ■

Väide 6.1 *Olgu* $(\Omega, \mathcal{F}, \mathbf{P})$ *selline tõenäosusruum, kus* $\mathcal{F}$ *koosneb vaid sellistest elementidest, mille mõõt on kas 0 või 1. Kui* X *on sellisel tõenäosusruumil antud juhuslik suurus, siis leidub* $c \in \mathbb{R}$ *nii, et* $\mathbf{P}(X = c) = 1$.

Tõestus. Iga $x \in \mathbb{R}$ korral $\mathbf{P}(X \leq x)$ on kas 0 või 1. Olgu F juhusliku suuruse X jaotusfunktsioon. Seega iga $x \in \mathbb{R}$ korral $F(x)$ on kas 1 või 0. Et F kui jaotusfunktsioon on paremalt pidev ning $\lim_{x\to\infty} F(x) = 1$ ja $\lim_{x\to-\infty} F(x) = 0$, on $F = I_{[c,\infty]}$ mingi konstandi $c \in \mathbb{R}$ korral. Seega $\mathbf{P}(X \leq c) = F(c) = 1$ ja $\mathbf{P}(X < c) = F(c-) = 0$. Järelikult $\mathbf{P}(X = c) = 1$. ■

Järeldus 6.2 *Olgu* $\mathcal{T}$ *sõltumatute juhuslike suuruste* $X_1, X_2, \ldots$ *jääk-*σ*-algebra. Kui* $X \in m\mathcal{T}$*, siis mingi konstandi* c *korral* $\mathbf{P}(X = c) = 1$.

6.4 Ülesanded

1. Olgu $\{A_n\} \subset \mathcal{F}$ sõltumatute sündmuste jada nii, et $\mathbf{P}(A_n) < 1 \; \forall n$. Tõestada, et

$$\mathbf{P}(\limsup A_n) = 1 \text{ parajasti siis, kui } \mathbf{P}(\cup_n A_n) = 1.$$

2. Olgu $\{A_n\} \subset \mathcal{F}$ sõltumatute sündmuste jada. Tõestada, et $\mathbf{P}(\limsup A_n) = 1$ parajasti siis, kui rida $\sum \mathbf{P}(A \cap A_n)$ hajub iga positiivse tõenäosusega A korral.

3. Olgu $X_1, X_2, \ldots$ sõltumatud, eksponentjaotusega juhuslike suuruste jada, $X_i \sim E(1)$, s.t. $\mathbf{P}(X_i > x) = e^{-x}$.

a) Tõestada, et suvalise $c \in \mathbb{R}$ korral $\mathbf{P}(X_n > c \quad i.\,o.) = 1$.

b) Tõestada, et

$$\mathbf{P}(X_n > \alpha \ln n \quad i.\,o.) = \begin{cases} 0 & \text{kui } \alpha > 1, \\ 1 & \text{kui } \alpha \leq 1.. \end{cases}$$

c) Olgu $L = \limsup_n (X_n / \ln n)$. Tõestada, et $L = 1$ p.k.

($\mathbf{P}(L \geq 1) = 1$ järeldub BC 2 -st, $\mathbf{P}(L > 1) = \mathbf{P}(\cup_k (L > 1 + k^{-1})$.)

4. Ahv tipib lõpmata kaua kirjutusmasinat. Selleks, et trükkida Shakespeare kogutud teosed, on tal vaja õiges järjestuses trükkida N sümbolit. Eeldades, et kirjutusmasinal on n sümbolit ja ahv trükib neid kõiki juhuslikult, kusjuures iga sümboli trükkimise tõenäosus on n^{-1} tõestada, et $\mathbf{P}$(ahv trükib lõpmata palju kordi Shakespeare kogutud teosed)=1.

(Olgu A_i sündmus, et ahv trükib kogutud teosed $N(i-1)$ kuni iN sümboliga ($i = 1, 2 \ldots$). Kasutada BC 2.)

5. Olgu $\{A_n\} \subset \mathcal{F}$ sõltumatute sündmuste jada. Ilma Boreli-Cantelli lemmasi kasutamata näidata, et $\mathbf{P}(\limsup_n A_n)$ ja $\mathbf{P}(\liminf_n A_n)$ on kas 0 või 1.

6. Olgu $X_1, X_2, \ldots$ juhuslike suuruste jada, $\mathcal{T}$ olgu vastav jääk-σ-algebra. Tõestada, et $F_1, F_2, F_3 \subset \mathcal{T}$, kus

$$\begin{aligned} F_1 &:= \{\omega | \exists \lim X_k(\omega)\} = (\lim X_k \text{eksisteerib}) \\ F_2 &:= (\text{rida} \sum X_k \text{koondub}) \\ F_3 &:= (\lim \frac{X_1 + X_2 + \cdots + X_k}{k} \text{eksisteerib}). \end{aligned}$$

7. Olgu $X_1, X_2, \ldots$ juhuslike suuruste jada. Defineerime klassid

$$\begin{aligned} \mathcal{K}_n &:= \{\{\omega | X_i(\omega) \leq x_i,\, i = 1, \ldots, n-1\} |\; x_i \in (-\infty, \infty]\} \\ \mathcal{I}_n &:= \{\{\omega | X_i(\omega) \leq x_i,\, i = n, \ldots, n+r\} |\; r \in \mathbf{N},\, x_i \in (-\infty, \infty]\}. \end{aligned}$$

Tõestada, et $\mathcal{K}_n$ ja $\mathcal{I}_n$ on π-süsteemid, mis tekitavad χ_n ja $\mathcal{T}_n$.

8. Olgu $X_1, X_2, \ldots$ juhuslike suuruste jada, $\mathcal{T}$ – jääk-σ-algebra,

$$\chi_n := \sigma(X_1, \cdots, X_{n-1}), \quad \chi_\infty := \sigma(X_1, X_2, \ldots).$$

Olgu $a \in \mathbb{R}$. Defineerime funktsioonid:

$$\begin{aligned} T(\omega) &= \inf\{n | X_n(\omega) > a\} \\ U(\omega) &= \inf\{n | \frac{X_1(\omega) + \cdots + X_n(\omega)}{n} > a\} \\ L(\omega) &= \sup\{n | \frac{X_1(\omega) + \cdots + X_n(\omega)}{n} > a\} \end{aligned}$$

Siin

$$\inf \emptyset = \infty \text{ ja } L(\omega) = \infty, \text{ kui } \frac{X_1(\omega) + \cdots + X_n(\omega)}{n} > a \quad \text{i.o..}$$

Kas

- T, U, L on (laiendatud) juhuslikud suurused (st. väärtuste hulgaga $[-\infty, \infty]$)
- Kas T, U, L on $\mathcal{T}_n$ mõõtuvad?
- Kas T, U, L on χ_n mõõtuvad?
- Kas T, U, L on χ_∞ mõõtuvad?
- Kas T, U, L on $\mathcal{T}$ mõõtuvad?

7 Lebesgue'i integraal

Olgu (S, Σ, μ) mõõduga ruum, $m\Sigma$ ruumil (S, Σ, μ) antud mõõtuvate laiendatud funktsioonide hulk, $m\Sigma^+$ mittenegatiivsete mõõtuvate laiendatud funktsioonide hulk.

7.1 Lihtsa mittenegatiivse mõõtuva funktsiooni integraal

Def 7.1 *Funktsioon* $f : S \mapsto [-\infty, \infty]$ *on* **lihtne** *kui*

$$f = \sum_{i=1}^{m} a_i I_{A_i}, \tag{7.1}$$

kus $a_i \in [-\infty, \infty]$ *ja* $A_i \cap A_j = \emptyset$ *ja* $a_i \neq a_j$, *kui* $i \neq j$.

Märkus: Lihtsa funktsiooni võib defineerida

D1 ka nii, et lisatud on nõue: $\cup A_i = S$. See tähendab, et hulgad A_i moodustavad ruumi S **tükelduse**;

D2 ka nii, et nõuded: $A_i \cap A_j = \emptyset$ ja $a_i \neq a_j$, kui $i \neq j$ puuduvad üldse,

D3 kui funktsiooni, millel on vaid lõplik arv väärtusi.

Need definitsioonid on kõik ekvivalentsed definitsiooniga 7.1 (ülesanne 1). Edaspidi, rääkides lihtsast funktsioonist $\sum_i^m a_i I_{A_i}$, peame silmas definitsiooni 7.1, s.t. $a_i \neq a_j$ ja $A_i \cap A_j = \emptyset$, kui $i \neq j$. Järgneva väite tõesuses oleme juba veendunud.

Väide 7.1 *Lihtne funktsioon (13.8) on mõõtuv parajasti siis, kui* $A_i \in \Sigma \quad i = 1, \ldots, m$.

Kui f on mittenegatiivne lihtne mõõtuv funktsioon, s.t. $a_i \in [0, \infty]$, siis kirjutame, et

$$f \in SF^+.$$

Def 7.2 *Kuulugu* $f \in SF^+$. *Funktsiooni* f **Lebesgue'i integraaliks** *(mõõdu* μ *järgi)* $\int f d\mu$ *nimetatakse summat*

$$\int f d\mu := \sum_{i=1}^{m} a_i \mu(A_i) \leq \infty.$$

(Siin $0 \cdot \infty = \infty \cdot 0 = 0$, $x \cdot \infty = \infty \cdot x = \infty$ *kui* $x > 0$ *ja* $\infty \cdot \infty = \infty$.*)*

Märkus: Toodud definitsioon ei sõltu lihtsa funktsiooni esitusest, s.t. ühe ja sama funktsiooni erinevad esitused (**D1**, **D2**, **D3**) (millal see võib nii olla?) annavad integraalile sama väärtuse (ülesanne 2).

Näide: Olgu $(S, \Sigma, \mu) = ((a,b], \mathcal{B}(a,b], \mathrm{Leb})$ ja

$$f = \sum_i^n a_i I_{A_i},$$

kus $A_i = (c_i, d_i]$. siis on Lebesgue'i integraal võrdne Riemanni integraaliga

$$\int f d\mu = \int_a^b f(x)dx.$$

7.1.1 Lihtsa mittenegatiivse funktsiooni integraali omadused

Kuulugu $f, g \in SF^+$.

A1 kui $\mu\{s : f(s) \neq g(s)\} = 0$, siis $\int f d\mu = \int g d\mu$;

A2 kui $f \leq g$, st $f(s) \leq g(s)\ \forall s \in S$, siis $\int f d\mu \leq \int g d\mu$ (**integraali monotoonsus**);

A3 kui $\alpha, \beta \in \mathbb{R}^+$, siis $\int (\alpha f + \beta g) d\mu = \alpha \int f d\mu + \beta \int g d\mu$ (**integraali lineaarsus**).

Tõestus algab tähelepanekust, et iga kahe lihtsa $f, g \in SF^+$ korral leidub hulga S mõõtuv tükeldus $\{C_1, \ldots, C_n\}$ nii, et

$$f = \sum_i a_i I_{C_i} \text{ ja } g = \sum_i b_i I_{C_i}$$

mingite a_i, b_i korral (mis muidugi ei pruugi rahuldada nõuet $a_i \neq a_j$ ja $b_i \neq b_j$ kui $i \neq j$). Edasine on lihtne.

7.2 Mittenegatiivse funktsiooni integraal

Def 7.3 *Mittenegatiivse mõõtuva funktsiooni $f \in m\Sigma^+$* **Lebesgue'i integraal** *(mõõdu μ järgi) $\int f d\mu$ defineeritakse järgmiselt*

$$\int f d\mu := \sup\Big\{\int h d\mu \Big| \ h \in SF^+, h \leq f\Big\} \leq \infty.$$

Alternatiivne definitsioon

Def 7.4 *Mittenegatiivse mõõtuva funktsiooni $f \in m\Sigma^+$* **Lebesgue'i integraal** *$\int f d\mu$ defineeritakse*

$$\int f d\mu := \sup \sum_i [\inf_{s \in A_i} f(s)] \mu(A_i),$$

kus supremum on võetud üle kõigi ruumi S Σ-mõõtuvate lõplike (st tükkide arv on lõplik) tükelduste $\{A_i\}$.

Need definitsioonid on ekvivalentsed (ülesanne 4).

Teoreem 7.5 *(***Monotoonse koondumise teoreem ehk Beppo-Levi teoreem***). Olgu $f_n, f \in m\Sigma^+$ ja $f_n(s) \nearrow f(s)$ iga $s \in S$ korral. Siis*

$$\int f_n d\mu \nearrow \int f d\mu \leq \infty.$$

Tõestuse võib leida (*Billingsley*, lk. 201; *Williams*, A5). See on väga oluline teoreem, sest koos teoreemiga 4.5 võimaldab monotoonse koondumise teoreem lihtsate funktsioonide integraalide omaduste ülekandmist mittenegatiivsete funktsioonide integraalidele. Samuti võimaldab see teoreem veel ühe definitsiooni mittenegatiivse funktsiooni Legesgue'i integraalile.

Def 7.6 *Olgu $f \in m\Sigma^+$ mõõtuv ja mittenegatiivne, olgu $h_n \in SF^+$ selline jada, et $h_n \nearrow f$. Funktsiooni f Lebesgue'i integraal on piirväärtus*

$$\int f d\mu := \sup_n \int h_n d\mu = \lim_{n\to\infty} \int h_n d\mu.$$

Teoreemi 4.5 tõttu selline jada h_n leidub. Monotoonse koondumise teoreemi tõttu ei sõltu integraal jada h_n valikust, s.t. definitsioon on korrektne.

7.2.1 Nullmõõdulised hulgad ja "peaaegu kõikjal"

Def 7.7 *Ütleme, et mingi omadus kehtib* **peaaegu kõikjal (p.k.)**, *kui*

$$\mu\{s: \text{ omadus ei kehti } \} = 0.$$

Näide: Olgu $f_n, f \in m\Sigma$. Siis

$$f_n \to f \text{ p.k. tähendab } \mu\{s : f_n(s) \not\to f(s)\} = 0.$$

Märkused:

- Kui $(S, \Sigma, \mu) = (\Omega, \mathcal{F}, \mathbf{P})$, s.t. vaadeldav mõõduga ruum on tõenäosusruum, siis termini "peaaegu kõikjal" asemel kasutatakse "peaaegu kindlasti".
- Omadus saab kehtida peaaegu kindlasti vaid siis, kui $\{s : \text{omadus ei kehti } \} \in \Sigma$.
- Loenduva hulga nullmõõduliste hulkade ühend on nullmõõduline hulk: kui $\mu(A_n) = 0\ \forall n$, siis
$$\mu(\cup_n A_n) \leq \sum_n \mu(A_n) = 0.$$

Näited nullmõõdulistest hulkadest:

1 Olgu $(S, \Sigma, \mu) = (\mathbb{R}, \mathcal{B}, \mu)$, kus μ on mingi Lebesgue'i-Stieltjesi mõõt, millel on pidev (üldistatud) jaotusfunktsioon G. Et G on pidev, on iga ühepunktiline hulk $\{x\}$ nullmõõduline hulk. Järelikult on ka kõik loenduvad hulgad nullmõõdulised hulgad. Seega on kõik ülimalt loendauvad hulgad nullmõõdulised, kui μ on

- Lebesgue'i mõõt
- pideva jaotusfunktsiooniga tõenäosusjaotus.

2 Olgu S suvaline hulk, μ olgu σ-algebral 2^S antud loendav mõõt. Siis ainus nullmõõduline hulk on $\emptyset$.

3 Olgu $\mu = \delta_x$ Diraci mõõt. Hulk A on nullmõõduline parajasti siis, kui $x \notin A$.

7.2.2 Mittenegatiivse funktsiooni integraali omadused

Olgu $f, g \in m\Sigma^+$. Siis

B1 $\int f d\mu = 0$ parajasti siis, kui $f = 0$ p.k.;

B2 kui $\int f d\mu < \infty$, siis $f < \infty$ p.k.;

B3 kui $f \leq g$ p.k., siis $\int f d\mu \leq \int g d\mu \leq \infty$;

B4 kui $f = g$ p.k., siis $\int f d\mu = \int g d\mu \leq \infty$;

B5 kui $\alpha, \beta \in [0, \infty)$, siis $\int (\alpha f + \beta g) d\mu = \alpha \int f d\mu + \beta \int g d\mu$.

Tõestus. B1: Kui $f = 0$ p.k., siis $\int f d\mu = 0$ (ülesanne 6).
Teistpidi. Olgu $f \in m\Sigma^+$ nii, et $\int f d\mu = 0$. Veendume, et $\mu\{s : f(s) > 0\} = 0$. Veendu, et $\{s : f(s) > 0\} = \lim_n \{s : f(s) > n^{-1}\}$. Oletades vastuväiteliselt, et $\mu\{s : f(s) > 0\} > 0$, saame, et leidub n_o nii, et

$$\mu\{s : f(s) > n_o^{-1}\} > 0.$$

Nüüd

$$f \geq n_o^{-1} I_{\{f(s) > n_o^{-1}\}},$$

millest integraali definitsiooni kohaselt

$$\int f d\mu \geq n_o^{-1} \mu\{s : f(s) > n_o^{-1}\} > 0.$$

B2: Ülesanne 6.

B3: Olgu $A = \{f(s) > g(s)\}$, $\mu(A) = 0$. Kuulugu $h \in \{h \in SF^+, h \leq f\}$. Defineerime $h^* = hI_{A^c} \in SF^+$. On selge, et $h = h^*$ p.k., vastavalt omadusele **A1** saame $\int hd\mu = \int h^* d\mu$ Et $h^* \in \{h \in SF^+ | h \leq g\}$, siis $\int hd\mu \leq \int gd\mu$. Seega,

$$\int fd\mu = \sup\{\int hd\mu | h \in SF^+, h \leq f\} \leq \int gd\mu.$$

B4: Järeldub eelmisest omadusest.

B5: Monotoonse koondumise teoreem ja teoreem 4.5 ja **A3**. ■

7.3 Integreeruvad funktsioonid

Olgu $f \in m\Sigma$. Funktsiooni f positiivseks ja negatiivseks osaks nimetatakse funktsioone f^+ ja f^-, kus

$$f^+(s) = \max(f(s), 0) \text{ ja } f^-(s) = \max(-f(s), 0).$$

On selge, et

$$f = f^+ - f^-, \quad |f| = f^+ + f^- \text{ ja } f^+, f^- \in m\Sigma^+.$$

Def 7.8 *Funktsiooni $f \in m\Sigma$ nimetatakse integreeruvaks ja tähistatakse $f \in \mathcal{L}_1(S, \Sigma, \mu)$ kui $\int f^+ d\mu < \infty$ ja $\int f^- d\mu < \infty$. Funktsiooni $f \in \mathcal{L}_1(S, \Sigma, \mu)$* **Lebesgue'i integraaliks** *(mõõdu μ järgi) $\int fd\mu$ nimetatakse vahet*

$$\int_S f(s)\mu(ds) =: \int fd\mu := \int f^+ d\mu - \int f^- d\mu.$$

Juhul, kui f ei ole integreeruv, kuid $\int f^+ d\mu < \infty$ või $\int f^- d\mu < \infty$, defineeritakse Lebesgue'i integraal

$$\int fd\mu := \begin{cases} \infty, & \text{kui } \int f^- d\mu < \infty \\ -\infty, & \text{kui } \int f^+ d\mu < \infty. \end{cases}$$

Kui $\int f^+ d\mu = \int f^- d\mu = \infty$, siis funktsiooni $f \in m\Sigma$ Lebesgue'i integraali ei defineerita.

Märkus: Kui $f \in m\Sigma$, siis $f \in \mathcal{L}_1(S, \Sigma, \mu)$ parajasti siis, kui $\int |f| d\mu < \infty$. (Ülesanne 7).

7.4 Integraali omadused

C1 kui $f \in \mathcal{L}_1(S, \Sigma, \mu)$, siis $|\int fd\mu| \leq \int |f| d\mu < \infty$;

C2 kui $f, g \in m\Sigma$, $f = g$ p.k. ja $\int gd\mu$ eksisteerib, siis $\int fd\mu = \int gd\mu$;

C3 kui $|f| \leq |g|$ p.k. ja $g \in \mathcal{L}_1(S, \Sigma, \mu)$, siis $f \in \mathcal{L}_1(S, \Sigma, \mu)$;

C4 **Monotoonsus:** kui $f, g \in \mathcal{L}_1(S, \Sigma, \mu)$ ja $f \leq g$ p.k., siis $\int fd\mu \leq \int gd\mu$;

C5 Lineaarsus: kui $f, g \in \mathcal{L}_1(S, \Sigma, \mu)$ ja $\alpha, \beta \in \mathbb{R}$, siis $\int(\alpha f + \beta g)d\mu = \alpha \int f d\mu + \beta \int g d\mu$.

Tõestus. C1: Ülesanne 7.

C2: Kui $f(s) = g(s)$, siis

$$f^+(s) = g^+(s) \text{ ja } f^-(s) = g^-(s),$$

mistõttu $f = g$ p.k. tähendab, et

$$f^+ = g^+ \text{ p.k. ja } f^- = g^- \text{ p.k. }.$$

Need on mittenegatiivsed funktsioonid, mistõttu

$$\int f^+ d\mu = \int g^+ d\mu \text{ ja } \int f^- d\mu = \int g^- d\mu.$$

C3: Järeldub definitsioonist 7.8 ja omadusest **B3**.

C4: Kui $f(s) \leq g(s)$, siis

$$f^+(s) \leq g^+(s) \text{ ja } -f^-(s) \leq -g^-(s).$$

Seega

$$f \leq g \text{ p.k. } \Rightarrow \quad f^+ \leq g^+ \text{ p.k. ja } g^- \leq f^- \text{ p.k.},$$

ning **C4** järeldub omadusest **B3**.

C5: Olgu $\alpha \in \mathbb{R}$, $f \in \mathcal{L}_1(S, \Sigma, \mu)$. Siis (**B5**)

$$\int |\alpha f| d\mu = \int |\alpha||f| d\mu = |\alpha| \int |f| d\mu < \infty,$$

millest $\alpha f \in \mathcal{L}_1(S, \Sigma, \mu)$. Kehtib

$$\int \alpha f d\mu = \alpha \int f d\mu,$$

sest iga $\alpha > 0$ korral

$$\int \alpha f d\mu = \int (\alpha f)^+ d\mu - \int (\alpha f)^- d\mu = \int \alpha f^+ d\mu - \int \alpha f^- d\mu = \alpha \Big(\int f^+ d\mu - \int f^- d\mu \Big) = \alpha \int f d\mu$$

ja

$$\int (-f) d\mu = \int f^- d\mu - \int f^+ d\mu = -\int f d\mu.$$

Seega peame tõestama, et

$$f, g \in \mathcal{L}_1(S, \Sigma, \mu) \Rightarrow f + g \in \mathcal{L}_1(S, \Sigma, \mu) \text{ ja } \int (f+g)d\mu = \int f d\mu + \int g d\mu. \tag{7.2}$$

Kui $f, g \in \mathcal{L}_1(S, \Sigma, \mu)$, siis omadustest **B3** ja **B5** saame

$$\int |f+g|d\mu \leq \int (|f| + |g|)d\mu = \int |f|d\mu + \int |g|d\mu < \infty.$$

Seega $f + g \in \mathcal{L}_1(S, \Sigma, \mu)$.

Olgu $h = f + g$. Tõestame, et

$$\int h d\mu = \int f d\mu + \int g d\mu.$$

Kehtib

$$h^+ - h^- = h = f + g = f^+ - f^- + g^+ - g^-,$$

millest

$$h^+ + f^- + g^- = h^- + g^+ + f^+$$

ja (**B4**)

$$\int (h^+ + f^- + g^-)d\mu = \int (h^- + g^+ + f^+)d\mu.$$

Kõik funktsioonid on mittenegatiivsed, seega (**B5**)

$$\int h^+ d\mu + \int f^- d\mu + \int g^- d\mu = \int h^- d\mu + \int g^+ d\mu + \int f^+ d\mu,$$

millest

$$\int h d\mu = \int h^+ d\mu - \int h^- d\mu = \int f^+ d\mu - \int f^- d\mu + \int g^+ d\mu - \int g^- d\mu = \int f d\mu + \int g d\mu.$$

■

7.5 Integraalide koondumisteoreemid

Olgu (S, Σ, μ) mõõduga ruum.

Olgu $f_n, f \in \mathcal{L}_1(S, \Sigma, \mu)$, kusjuures $f_n \to f$ p.k., s.t. $\mu\{s : f_n(s) \not\to f(s)\} = 0$. Kas integraalid koonduvad, s.t. kas kehtib

$$\int f_n d\mu \to \int f d\mu? \tag{7.3}$$

NB! Alati ei ole see nii!

Kui aga jada $\{f_n\}$ rahuldab teatud tingimusi, siis küll.

Teoreem 7.9 (Monotoonse koondumise teoreem: MON)
Kui $f_n, f \in m\Sigma$ ja $0 \le f_n \nearrow f$ p.k., siis

$$\int f_n d\mu \nearrow \int f d\mu \le \infty.$$

Tõestus. Olgu

$$A = \{s \in S : 0 \le f_n(s) \nearrow f(s)\}.$$

Kui $0 \le f_n \nearrow f$ p.k., siis $\mu(A^c) = 0$. Olgu

$$h_n := f_n I_A, \quad h = f I_A.$$

Kehtib:

$$h_n = f_n \text{ p.k.}, \quad h = f \text{ p.k.}, \quad h_n, h \in m\Sigma^+, \quad h_n(s) \nearrow h(s), \quad \forall s \in S.$$

MON: $\int h_n d\mu \nearrow \int h d\mu$ ja **C2**:

$$\int f_n d\mu = \int h_n d\mu \nearrow \int h d\mu = \int f d\mu.$$

■

Lemma 7.1 (Fatou lemma)
Kui $f_n \in m\Sigma$ ja $0 \le f_n$ p.k., siis

$$\int \liminf_n f_n d\mu \le \liminf_n \int f_n d\mu.$$

Tõestus. Olgu

$$g_n := \inf_{k \ge n} f_k.$$

Seega

$$g_n \nearrow \liminf_n f_n \text{ ja } g_n \ge 0 \text{ p.k.}$$

Monotoonse koondumise teoreemist järeldub:

$$\int g_n d\mu \nearrow \int \liminf_n f_n d\mu.$$

Teisest küljest aga $g_n \le f_k$ kui $k \ge n$ ning seega

$$\int g_n d\mu \le \inf_{k \ge n} \int f_k d\mu \nearrow \liminf_n \int f_n d\mu.$$

■

Lemma 7.2 (Pööratud Fatou lemma)
Kui $f_n \in m\Sigma$, kusjuures $\exists g \in m\Sigma$ nii, et $\int g d\mu < \infty$ ja $0 \le f_n \le g$ p.k., siis

$$\limsup_n \int f_n d\mu \le \int \limsup_n f_n d\mu.$$

Tõestus. Ülesanne 20. ■

Teoreem 7.10 ((Lebesgue'i) domineeritud koondumise teoreem: DOM)
Olgu $f_n, f \in m\Sigma$, $f_n \to f$ p.k.. Kui leidub domineeriv funktsioon $g \in \mathcal{L}_1(S, \Sigma, \mu)$ nii, et $|f_n| \le g$ p.k., siis $f, f_n \in \mathcal{L}_1(S, \Sigma, \mu)$ ja $\int |f_n - f| d\mu \to 0$, millest

$$\int f_n d\mu \to \int f d\mu.$$

Tõestus. Omaduse **C3** tõttu on f_n integreeruvad. Koondumisest $f_n \to f$ p.k. järeldub $|f| \le g$ p.k. ja $|f_n - f| \to 0$ p.k.. Võrratusest $|f| \le g$ p.k. saame, et f integreeruv. Kehtib $|f_n - f| \le 2g$ p.k.. Et $2g \in \mathcal{L}_1(S, \Sigma, \mu)$, saame kasutada lemmat 7.2, millest järeldub, et

$$0 \le \liminf_n \int |f_n - f| d\mu \le \limsup_n \int |f_n - f| d\mu \le \int \limsup_n |f_n - f| d\mu = 0.$$

Omaduse **C1** tõttu

$$\Big| \int f_n d\mu - \int f d\mu \Big| = \Big| \int (f_n - f) d\mu \Big| \le \int |f_n - f| d\mu \to 0.$$

■

Näide: Domineeriva funktsiooni g olemasolu on vajalik. Olgu

$$(S, \Sigma, \mu) = ((0,1], \mathcal{B}(0,1], \mathrm{Leb}), \quad f_n = n I_{(0, n^{-1})}.$$

Seega

$$f_n(s) = \begin{cases} n & \text{kui } s \le \frac{1}{n}, \\ 0 & \text{kui } s > \frac{1}{n}. \end{cases}$$

Kehtib: $f_n(s) \to 0, \quad \forall s$, kuid

$$\int f_n d\mathrm{Leb} = n \mathrm{Leb}((0, n^{-1})) = 1 \not\to \int 0 d\mathrm{Leb} = 0.$$

Järeldus 7.1 (Tõkestatud koondumise teoreem)
Olgu μ lõplik mõõt, $f_n, f \in m\Sigma$ ja $f_n \to f$ p.k. Kui $\exists K < \infty : |f_n(s)| < K \; \forall n$ p.k., siis $f, f_n \in \mathcal{L}_1(S, \Sigma, \mu)$ ja $\int |f_n - f| d\mu \to 0$, millest

$$\int f_n d\mu \to \int f d\mu.$$

Tõestus. Ülesanne 21. ■

Järgnev järeldus esitab piisava tingimuse selleks, et integreerimise (Lebesgue'i mõõdu järgi) ja diferentseerimise järjekorra võib vahetada.

Järeldus 7.2 *Olgu (S, Σ, μ) mõõduga ruum, kuulugu $a \in U \subset \mathbb{R}^n$, kusjuures U on lahtine. Olgu $f : S \times U \mapsto \mathbb{R}$ funktsioon ning olgu*

$$h : U \mapsto \mathbb{R}, h(x) = \int f(s, x)\mu(ds).$$

1) *Kui iga s korral on funktsioon $f(s, \cdot) : U \mapsto \mathbb{R}$ punktis a pidev ning kui leidub mingi integreeruv funktsiooni $g : S \mapsto \mathbb{R}$ nii, et*

$$|f(s, x)| \leq g(s), \quad \forall (s, x) \in S \times U,$$

siis on h punktis a pidev.

2) *Kui iga s korral on funktsioonil $f(s, \cdot)$ punktis a pidev osatuletis ning kui leidub mingi integreeruv funktsioon $g : S \to [0, \infty]$ nii, et*

$$|\frac{\partial f}{\partial x_i}(s, x)| \leq g(s), \quad \forall (s, x) \in S \times U, \quad 1 \leq i \leq n,$$

siis on funktsioonil h punktis a pidevad osatuletised, kusjuures kehtib

$$\frac{\partial h}{\partial x_i}(a) = \int \frac{\partial f}{\partial x_i}(s, a)\mu(ds).$$

Tõestus.

1) Tõestame: $x_n \to a \Rightarrow h(x_n) \to h(a)$.
Kui $x_n \to a$, siis $f(s, \cdot)$ pidevusest punktis a järeldub: $f(s, x_n) \to f(s, a)$, $\forall s$. Funktsioonid $f(s, x_n)$ rahuldavad tingimust $f(s, x_n) \leq g(s)$ $\forall s$, mistõttu DOM:

$$h(x_n) = \int f(s, x_n)d\mu \to \int f(s, a)d\mu = h(a).$$

2) Tõestame:

$$\lim_n \frac{h(a + t_n e_i) - h(a)}{t_n} = \int \lim_n \frac{f(s, a + t_n e_i) - f(s, a)}{t_n} d\mu,$$

kus $e_i = (0, \ldots, 0, 1, 0, \ldots, 0)$, ja 1 on i-s koordinaat. **C5**:

$$\lim_n \frac{h(a + t_n e_i) - h(a)}{t_n} = \lim_n \int \frac{f(s, a + t_n e_i) - f(s, a)}{t_n} d\mu$$

mistõttu piisab, kui tõestame, et

$$\lim_n \int \frac{f(s, a + t_n e_i) - f(s, a)}{t_n} d\mu = \int \lim_n \frac{f(s, a + t_n e_i) - f(s, a)}{t_n} d\mu \tag{7.4}$$

Kõik funktsioonid $f(s, \cdot)$ on pidevate osatuletistega. Et leidub $0 \leq u_n \leq t_n$ nii, et

$$\frac{f(s, a + t_n e_i) - f(s, a)}{t_n} = \frac{\partial f}{\partial x_i}(s, u_n e_i + a) \leq g(s),$$

siis domineeritud koondumise teoreemist järeldub koondumine (7.4).

■

Kontranäide: Olgu $S = [0, \infty)$, $\mu = \text{Leb}$. Vaatleme funktsiooni

$$f : [0, \infty) \times \mathbb{R} \to \mathbb{R}, \quad f(s, x) = x^3 e^{-sx^2}.$$

Veendu, et iga x korral (ka siis, kui $x = 0$)

$$h(x) = \int_0^\infty x^3 e^{-sx^2} ds = x.$$

Seega h on pidev ja diferentseeruv funktsioon, $h'(0) = 1$. Teisest küljest aga

$$\frac{\partial f(s, x)}{\partial x} = 3x^2 e^{-sx^2} - 2sx^4 e^{-sx^2}.$$

Saadud funktsioon on võrdne nulliga, kui $x = 0$, st

$$\int_0^\infty \frac{\partial f(s, 0)}{\partial x} ds = 0 \neq h'(0) = 1.$$

Seega punktis 0 ei saa tuletisega integraali märgi alla minna. Veendu, et juhul kui $a \neq 0$ on see võimalik. Veendu, et punktis 0 ei kehti teoreemi eeldused. Tõepoolest, võttes $x^2 = \frac{1}{s}$ saame, et

$$\exp[-sx^2](3x^2 - 2sx^4) = \frac{e^{-1}}{s}.$$

Võttes suvelise $\epsilon > 0$, valime s nii suure, et $s^{-1} < \epsilon^2$. Seega

$$\sup_{x \in (-\epsilon, \epsilon)} \exp[-sx^2](3x^2 - 2sx^4) \geq \frac{e^{-1}}{s},$$

mis omakorda tähendab, et domineeriv funktsioon g peab rahuldama tingimust

$$g(s) \geq \frac{e^{-1}}{s}.$$

Paraku pole aga s^{-1} integreeruv üle $[0, \infty)$.

7.6 Integreerimine üle hulga

Def 7.11 *Kuulugu $f \in \mathcal{L}_1(S, \Sigma, \mu)$ või $f \in m\Sigma^+$. Iga $A \in \Sigma$ korral defineerime integraali üle hulga*

$$\int_A f(s)\mu(ds) :=: \int_A f d\mu := \int I_A f d\mu.$$

Märkused:

- kui $\mu(A) = 0$, siis $\int_A f d\mu = 0$ (järeldub omadusest **B1**);
- $\int_S f d\mu = \int f d\mu$ (definitsioon).

Teoreem 7.12 *Olgu $f \in m\Sigma^+$ või $f \in \mathcal{L}_1(S, \Sigma, \mu)$. Kui $A_1, A_2, \dots \in \Sigma$ on lõikumatud, siis*

$$\int_{\cup A_n} f d\mu = \sum_n \int_{A_n} f d\mu.$$

Tõestus. Kui $f \in m\Sigma^+$, siis MON; kui $f \in \mathcal{L}_1(S, \Sigma, \mu)$, siis DOM. (Ülesanne 14.) ■

Teoreem 7.13

i) *Olgu $f, g \in m\Sigma^+$, μ olgu σ-lõplik. Siis*

$$\int_A f d\mu = \int_A g d\mu \quad \forall A \in \Sigma \quad \Leftrightarrow f = g \ p.k.$$

ii) *Olgu $f, g \in \mathcal{L}_1(S, \Sigma, \mu)$. Siis*

$$\int_A f d\mu = \int_A g d\mu \quad \forall A \in \Sigma \quad \Leftrightarrow f = g \ p.k.$$

Märkus: Väide **ii)** ei eelda mõõdu μ σ-lõplikkust.

Tõestus.

i) Olgu $f, g \in m\Sigma^+$, olgu μ σ-lõplik.
$\Leftarrow$: Definitsioon.
$\Rightarrow$:Tõestame: kui

$$\int_A f d\mu \leq \int_A g d\mu \quad \forall A \in \Sigma, \tag{7.5}$$

siis $f \leq g$ p.k.
Olgu

$$B_n = \{0 \leq g < f, g \leq n\}.$$

Mõõt μ on σ-lõplik, järelikult $\exists S_n \in \Sigma$ nii, et $S_n \nearrow S$ ja $\mu(S_n) < \infty$. Funktsioon g on tõkestatud hulgal B_n ning $\mu(S_n) < \infty$. Seega seosest (7.5):

$$\int_{S_n \cap B_n} f d\mu \leq \int_{S_n \cap B_n} g d\mu < \infty \quad \Rightarrow \int_{S_n \cap B_n} (g - f) d\mu \geq 0.$$

Teisest küljest, hulga B_n definitsioonist ning omadusest **B1**:

$$I_{B_n \cap S_n}(f - g) \geq 0 \quad \Rightarrow \quad \int_{S_n \cap B_n} (f - g) d\mu \geq 0 \quad \Rightarrow$$

$$\int_{S_n \cap B_n} (f - g) d\mu = 0 \quad \Rightarrow \quad I_{B_n \cap S_n}(f - g) = 0 \quad \text{p.k.} \quad \Rightarrow \quad \mu(B_n \cap S_n) = 0.$$

Et $B_n \nearrow \{0 \leq g < f\}$, siis (miks?)

$$\mu\{0 \leq g < f\} = 0,$$

seega $f \leq g$ p.k..

ii) Olgu $f, g \in \mathcal{L}_1(S, \Sigma, \mu)$. Eeldame (7.5). Siis

$$\int_{\{g<f\}} f d\mu \leq \int_{\{g<f\}} g d\mu,$$

mistõttu **C5**:

$$\int_{\{g<f\}} (f-g) d\mu \leq 0 \quad \Rightarrow \quad \int_{\{g<f\}} (f-g) d\mu = 0 \quad \Rightarrow \quad I_{\{g<f\}}(f-g) = 0 \quad \text{p.k.} \quad \Rightarrow f \leq g \text{ p.k.}$$

■

7.7 Sheffe lemma

Lemma 7.3 (Sheffe lemma)
Kui $f, f_n \in \mathcal{L}_1(S, \Sigma, \mu)$ ja $f_n \to f$ p.k., siis

$$\int |f_n - f| d\mu \to 0 \quad \Leftrightarrow \quad \int |f_n| d\mu \to \int |f| d\mu.$$

Tõestus. $\Rightarrow$: Triviaalne (miks?).
$\Leftarrow$:
I Olgu f_n, f mittenegatiivsed ning

$$\int f_n d\mu \to \int f d\mu < \infty. \tag{7.6}$$

Vaatleme funktsiooni $(f_n - f)^-$. Et $(f_n - f)^- \leq f$, siis DOM:

$$\int (f_n - f)^- d\mu \to 0. \tag{7.7}$$

Kehtib

$$\int (f_n - f)^+ d\mu = \int_{\{f_n \geq f\}} (f_n - f) d\mu = \int f_n d\mu - \int f d\mu - \int_{\{f_n < f\}} (f_n - f) d\mu. \tag{7.8}$$

Seosest (7.7) järeldub, et

$$\left| \int_{\{f_n<f\}} (f_n - f) d\mu \right| \leq \int_{\{f_n<f\}} |f_n - f| d\mu = \int (f_n - f)^- d\mu \to 0. \tag{7.9}$$

Seoste (7.6), (7.9) ja (7.8) tõttu

$$\int (f_n - f)^+ d\mu \to 0,$$

seosest (7.7) saame nüüd

$$\int |f_n - f| d\mu = \int (f_n - f)^+ d\mu + \int (f_n - f)^- d\mu \to 0.$$

II Olgu $f_n, f \in \mathcal{L}_1(S, \Sigma, \mu)$ (mitte ilmtingimata mittenegatiivsed).

$$\int f_n^+ d\mu + \int f_n^- d\mu = \int |f_n| d\mu \to \int |f| d\mu = \int f^+ d\mu + \int f^- d\mu. \tag{7.10}$$

Fatou lemma:

$$\int f^+ d\mu \leq \liminf_n \int f_n^+ d\mu, \quad \int f^- d\mu \leq \liminf_n \int f_n^- d\mu. \tag{7.11}$$

Koondumistest (7.10) ja (7.11):

$$\int f_n^+ d\mu \to \int f^+ d\mu, \quad \int f_n^- d\mu \to \int f^- d\mu.$$

Rakendades nüüd osa **I** funktsioonidele f_n^+ ja f_n^-, saame

$$\int |f_n - f| d\mu = \int |(f_n^+ - f^+) - (f_n^- - f^-)| d\mu \leq \int |f_n^+ - f^+| d\mu + \int |f_n^- - f^-| d\mu \to 0.$$

■

7.8 Näited integraalidest

7.8.1 Lebesgue'i integraal ja Riemanni integraal

Olgu $[a, b]$ lõplik intervall ning olgu $f : [a, b] \to \mathbb{R}$ mingi $\mathcal{B}[a, b]$-mõõtuv funktsioon. Olgu f Riemann mõttes integreeruv, s.t. eksisteerib Riemanni integraal $\int_a^b f(x)dx$. Kas f on integreeriv Lebesgue'i mõttes Lebesgue'i mõõdu järgi, s.t. kas Lebesgue'i integraal $\int_{[a,b]} f d\mathrm{Leb}$ eksisteerib? Kui ja, kas Riemanni ja Lebesgue'i integraal võivad olla erinevad?

Teoreem 7.14

1) *Olgu $f : [a, b] \mapsto \mathbb{R}$ Boreli mõõtuv. Kui f on integreeruv Riemanni mõttes, on ta integreeruv ka Lebesgue'i mõttes, kusjuures nimetatud integraalid on võrdsed:*

$$\int_{[a,b]} f d\mathrm{Leb} = \int_a^b f(x)dx.$$

2) *Olgu $f : \mathbb{R} \mapsto [0,\infty)$ Boreli mõõtuv. Olgu f igal intervallil $[-n,n]$ Riemanni mõttes integreeruv. Funktsioon f on Lebesgue mõttes integreeruv parajasti siis, kui lõpmatute rajadega Riemanni integraal*

$$\lim_n \int_{-n}^{n} f(x)dx$$

eksisteerib ning sellisel juhul on need integraalid võrdsed:

$$\int f d\mathrm{Leb} = \int_{-\infty}^{\infty} f(x)dx = \lim_n \int_{-n}^{n} f(x)dx.$$

Tõestus.

1) Kinnisel lõigul Riemanni mõttes integreeruv funktsioon on tõkestatud. Selline funktsioon on Lebesgue'i mõttes integreeruv (mõõdu Leb järgi, miks?). Seega on f Lebesgue'i mõttes integreeruv. Tõestame, et integraalid on võrdsed. Et f on Riemann mõttes integreeruv, siis iga $\epsilon > 0$ korral leidub hulga $[a,b]$ tükeldus $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ nii, et

$$\int_a^b f(x)dx - \epsilon \leq U \leq \int_a^b f(x)dx \leq O \leq \int_a^b f(x)dx + \epsilon,$$

kus U ja O on vastavad ülem ja alamsummad

$$U := \sum_{i=2}^{n} \big(\inf_{s \in [\alpha_{i-1},\alpha_i]} f(s) \big)(\alpha_i - \alpha_{i-1})$$
$$O := \sum_{i=2}^{n} \big(\sup_{s \in [\alpha_{i-1},\alpha_i]} f(s) \big)(\alpha_i - \alpha_{i-1}).$$

Need summad on aga järgmiste lihtsate funktsioonide Lebesgue'i integraalid

$$h = \sum_{i=2}^{n} \big(\inf_{s \in [\alpha_{i-1},\alpha_i]} f(s) \big) I_{[\alpha_{i-1},\alpha_i)}, \quad g = \sum_{i=2}^{n} \big(\sup_{s \in [\alpha_{i-1},\alpha_i]} f(s) \big) I_{[\alpha_{i-1},\alpha_i)},$$

kusjuures $h \leq f \leq g$. Järelikult

$$\int_a^b f(x)dx - \epsilon \leq \int_{[a,b]} h d\mathrm{Leb} = U \leq \int f d\mathrm{Leb} \leq O = \int_{[a,b]} g d\mathrm{Leb} \leq \int_a^b f(x)dx + \epsilon.$$

Seega

$$\Big| \int_{[a,b]} f d\mathrm{Leb} - \int_a^b f(x)dx \Big| \leq \epsilon.$$

2) Olgu

$$r_n = \int_{-n}^{n} f(x)dx$$

Riemanni integraal üle $A_n = [-n, n]$. Lõpmatu Riemanni integraal on

$$\int_{-\infty}^{\infty} f(x)dx = \lim_n r_n.$$

Osa **1)** tõttu $r_n = \int_{A_n} f d\text{Leb}$. MON: $I_{A_n} f \nearrow f$

$$\int_{-\infty}^{\infty} f(x)dx = \lim_n r_n = \lim_n \int_{A_n} f d\text{Leb} = \int f d\text{Leb}.$$

■

Seega on Lebesgue'i integraal (Lebesgue'i mõõdu järgi) Riemanni integraali üldistus: lõigul Riemanni mõttes integreeruv Boreli funktsioon on ka Lebesgue'i mõttes integreeruv, kusjuures need integraalid langevad kokku. Vastupidine ei pruugi kehtida: järgnev näide on funktsioonist, mis on küll Lebesgue'i mõttes integreeruv, kuid Riemanni mõttes pole.

Näide: Dirichlet funktsioon

$$f : [0,1] \mapsto \mathbb{R}, \quad f(s) = \begin{cases} 1 & \text{kui } s \text{ on ratsionaalarv,} \\ 0 & \text{kui } s \text{ on irratsionaalarv.} \end{cases}$$

Funktsioon f pole Riemanni mõttes integreeruv, kuid on integreeruv Lebesgue'i mõttes, sest

$$f = I_A, \text{ kus } A = \{\mathbb{Q} \cap [0,1]\}.$$

Et $A \in \mathcal{B}[0,1]$ ja $\text{Leb}(A) = 0$, siis $\int f d\text{Leb} = 0$.

Teoreemi 7.14 põhjal järeldub lõigul $[a, b]$ antud Boreli funktsiooni Riemanni mõttes integreeruvusest Lebesgue'i mõttes integreeruvus ja need integraalid langevad kokku. Kas aga iga lõigul $[a, b]$ antud Riemanni mõttes integreeruv funktsioon on Boreli funktsioon? Järgnev järeldus väidab, et iga Riemanni mõttes integreeruv funktsioon on Leb-p.k. võrdne mingi Boreli funktsiooniga ja seetõttu Lebesgue'i mõõtuv (tuletame meelde, et Lebesgue'i σ-algebra on ruumi $([a, b], \mathcal{B}[a, b], \text{Leb})$ täieldi σ-algebra). Küll aga ei pruugi Riemanni mõttes integreeruv funktsioon olla Boreli mõõtuv.

Järeldus 7.3 *Olgu $f : [a, b] \to \mathbb{R}$ Riemanni mõttes integreeruv. Siis leidub lihtsate mõõtuvate funktsioonide jada h_n nii, et*

$$\lim_n h_n = f \quad p.k..$$

Tõestus. Et f on Riemanni mõttes integreeruv, siis leidub lõigu $[a, b]$ tükelduste jada Λ_n nii, et Λ_{n+1} on Λ_n peenendus ning vastavad ülem- ja alamsummad (O_n, U_n) koonduvad monotoonselt üheks ja samaks piirväärtuseks, Riemanni integraaliks $\int_a^b f dx$. Olgu h_n ja g_n vastavad lihtsad mõõtuvad funktsioonid. On selge, et

$$h_n \leq h_{n+1} \leq \cdots, \quad g_n \geq g_{n+1} \geq \cdots.$$

Seega on $g_n - h_n$ kahanev funktsioonide jada ning seetõttu eksisteerib piirväärtus

$$q := \lim_n (g_n - h_n).$$

Fatou lemma:

$$\int q d\text{Leb} \leq \liminf_n \int (g_n - h_n) d\text{Leb} = \lim_n (U_n - O_n) = 0.$$

Seega $q = 0$ p.k. Et $h_n \leq f \leq g_n$, siis

$$\lim_n h_n = f, \quad \text{p.k.}$$

■

Lebesgue'i integraal ei sõltu funktsiooni väärtustest mingil 0-mõõdulisel hulgal (omadus **C2**). Seega võime (teatud reservatsiooniga) öelda, et iga Riemanni mõttes integreeruv funktsioon on Lebesgue'i mõttes integreeruv: integraaliks on antud funktsiooniga *p.k* võrdse Boreli funktsiooni (vastavalt järeldusele 7.3 selline alati eksisteerib) Lebesgue'i integraal.

7.8.2 Loendav mõõt: integreerimine on summeerimine

A Olgu S loenduv hulk, $\Sigma = 2^S$ ja μ olgu loendav mõõt σ-algebral 2^S, s.t.

$$\mu(A) = A \text{ võimsus, kui } A \text{ on lõplik, vastasel juhul } \infty.$$

Iga funktsioon $f : S \to \mathbb{R}$ avaldub kujul

$$\{a_i\}_{i=1}^{\infty}, \quad a_i = f(x_i), \quad S = \{x_1, x_2, \ldots\}.$$

Kehtib

- Kui $f \in m\Sigma^+$, siis iga $\forall A \in \Sigma$ korral

$$\int_A f d\mu = \sum_{x_i \in A} f(x_i) \leq \infty. \tag{7.12}$$

Tõepoolest, iga $x \in S$ korral

$$\int a I_{\{x\}} d\mu = a\mu(\{x\}) = a;$$

nüüd pane tähele (teoreem 7.12), et

$$\int_A f d\mu = \int_{\cup_{x_i \in A}\{x_i\}} f d\mu = \sum_{x_i \in A} \int_{\{x_i\}} f d\mu = \sum_{x_i \in A} \int I_{\{x_i\}} f(x_i) d\mu = \sum_{x_i \in A} f(x_i).$$

Seega ($A = S$),

$$\int f d\mu = \sum_{i=1}^{\infty} a_i \leq \infty. \tag{7.13}$$

- Kui f pole mittenegatiivne, siis $f \in \mathcal{L}_1(S, \Sigma, \mu)$ parajasti siis, kui $\sum_i |a_i| < \infty$ ning sellisel juhul eelmise osa põhjal
$$\int_A f d\mu = \int_A f^+ d\mu - \int_A f^- d\mu = \sum_{x_i \in A} f^+(x_i) - \sum_{x_i \in A} f^-(x_i). \tag{7.14}$$
Seega, võttes $a_i^+ = f^+(x_i), a_i^- = f^-(x_i)$ ja $A = S$ saame
$$\int f d\mu = \sum_{i=1}^{\infty} a_i^+ - \sum_{i=1}^{\infty} a_i^-.$$
Analüüsist on teada, et absoluutselt koonduv rida on tingimatult koonduv (iga ümberjärjestuse summa on sama), seega
$$\sum_{i=1}^{\infty} a_i^+ - \sum_{i=1}^{\infty} a_i^- = \sum_{i=1}^{\infty} a_i$$
ehk võrdus (7.13) kehtib.
[Viimast saab ka otse tõestada. Tõepoolest, defineeri $f_n = f I_{\{x_1,\ldots,x_n\}}$. On selge, et $f_n(x) \to f(x)$ iga $x \in S$ korral. Et $|f_n(x)| \leq |f(x)|$ ja $\int |f| d\mu < \infty$, siis DOM:
$$\sum_{i=1}^{n} a_i = \int f_n d\mu \to \int f d\mu.$$
Seega osasummade jada $\sum_{i=1}^n a_i$ on koonduv. Piirväärtus on $\sum_{i=1}^\infty a_i$, seega kehtib võrdus (7.13)].
Seega on ka iga alamhulga A korral rida
$$\sum_{x_i \in A} f(x_i)$$
absoluutselt koonduv ja nii saame
$$\int_A f d\mu = \sum_{x_i \in A} f^+(x_i) - \sum_{x_i \in A} f^-(x_i) = \sum_{x_i \in A} f(x_i). \tag{7.15}$$

B Olgu S suvaline hulk, olgu μ loendav mõõt σ-algebral 2^S.

Teoreem 7.15 *Funktsioon f on Lebesgue'i mõttes integreeruv (loendava mõõdu järgi) parajasti siis, kui*
$$\sum_{x \in S} |f(x)| := \sup\{\sum_{x \in A} |f(x)| : A \subset S \text{ lõplik}\} < \infty; \tag{7.16}$$
Sellisel juhul kehtib
$$\{x : f(x) \neq 0\} = \{x_1, x_2, \ldots\}$$
ja
$$\int f d\mu = \sum_{i=1}^{\infty} f(x_i) =: \sum_{x \in S} f(x). \tag{7.17}$$

Tõestus. Iga lõpliku A korral kehtib (miks?)

$$\int_A f d\mu = \sum_{x\in A} f(x). \tag{7.18}$$

Kui f on integreeruv, siis

$$\sum_{x\in A} |f(x)| = \int_A |f| d\mu \leq \int |f| d\mu < \infty,$$

millest järeldub (7.16) ehk supremum on lõplik.
Näitame vastupidist – seosest (7.16) (supremum on lõplik) järeldub, et f on integreeruv (st $\int |f| d\mu < \infty$) ja kehtib (7.17). Kui kehtib (7.16), siis $A_n := \{|f| \geq \frac{1}{n}\}$ on lõplik ja seetõttu on hulk

$$\{f \neq 0\} = \cup_n \{|f| \geq \frac{1}{n}\}$$

ülimalt loenduv, s.t. leidub loenduv alamhulk $\{x_1, x_2, \ldots\}$ nii, et

$$\{x : f(x) \neq 0\} \subseteq \{x_1, x_2, \ldots\}.$$

On kerge veenduda, et sellisel juhul

$$\sum_{x\in S} |f(x)| = \sum_{i=1}^{\infty} |f(x_i)|$$

(st supremum võrdub rea summa) ja supremumi definitsioonist ning koondumisest $A_n \nearrow \{x_1, x_2, \ldots\}$ järeldub koondumine

$$\sum_{x\in A_n} |f(x)| \nearrow \sum_{x\in S} |f(x)| \tag{7.19}$$

(kuidas?). Seosest (7.18) ja monotoonse koondumise teoreemist järeldub

$$\sum_{x\in A_n} |f(x)| = \int_{A_n} |f| d\mu \nearrow \int |f| d\mu. \tag{7.20}$$

Koondumistest (7.19) ja (7.20) aga saame, et

$$\int |f| d\mu = \sum_{x\in S} |f(x)| < \infty.$$

Seega f on integreeruv ja kui f on mittenegatiivne, siis kehtib (7.17).
Nägime, et kui f on integreeruv, siis on rida $\sum f(x_i)$ absoluutselt koonduv ning rea summa ei sõltu liidetavate järjekorrast. Seega

$$\sum_{x\in S} f(x) = \sum_i f(x_i) = \sum_{i:f(x_i)\geq 0} f(x_i) - \sum_{i:f(x_i)<0} -f(x_i) = \int f^+ d\mu - \int f^- d\mu = \int f d\mu.$$

■

Järelikult, kui f on integreeruv loendava mõõdu järgi, siis

$$f = \sum_{i=1}^{\infty} a_i I_{\{x_i\}}, \quad \text{kus } a_i = f(x_i).$$

Olgu μ loendav mõõt ruumil $(\mathbb{R}, \mathcal{B})$. Teoreemist 7.15 järeldub, et $\forall A \in \mathcal{B}$ korral on Lebesgue'i integraal loendava mõõdu järgi

$$\int_A f d\mu = \sum_{x_i \in A} f(x_i) = \sum_{x_i \in A} a_i. \tag{7.21}$$

Kuid funktsiooni

$$f = \sum_{i=1}^{\infty} a_i I_{\{x_i\}}$$

Lebesgue'i integraal Lebesgue'i mõõdu järgi:

$$\int f d\text{Leb} = 0.$$

7.9 Standardne meetod

Monotoonse koondumise teoreem võimaldab Lebesgue'i integraali omadusi tõestada nn. **standardset meetodi** kasutades:

1) Tõesta omadus indikaatori $f = I_A$ korral.

2) Lineaarsust kasutades tõesta omadus lihtsate mittenegatiivset funktsioonide korral ($f \in SF^+$).

3) Monotoonse koondumise teoreemi ja teoreemi 4.5 kasutades tõesta omadus mittenegatiivsete funktsioonide korral ($f \in m\Sigma^+$).

4) Seost $f = f^+ - f^-$ ja lineaarsust kasutades tõesta omadus integreeruvate funktsioonide korral.

7.10 Tihedus

Olgu $h \in m\Sigma^+$. Defineerime funktsiooni

$$\nu : \Sigma \to [0, \infty], \quad \nu(A) = \int_A h d\mu.$$

Väide 7.2 *Funktsioon ν on mõõt σ-algebral Σ.*

Tõestus. (ülesanne 15). ■

Def 7.16 *Kui* $\exists h \in m\Sigma^+$ *nii, et* $\nu(A) = \int_A h d\mu$, $\forall A \in \Sigma$, *siis ütleme, et* h *on mõõdu* ν **tihedus** *mõõdu* μ *suhtes. Seda tähistatakse ka*

$$h = \frac{d\nu}{d\mu} \text{ või } \nu = h\mu.$$

Märkused:

- Teoreemist 7.13 järeldub, et kui h_1 ja h_2 on kaks tihedust, s.t. $\nu = h_1\mu$ ja $\nu = h_2\mu$, siis $\mu\{h_1 \neq h_2\} = 0$. Ütleme, et tihedus μ-p.k. ühene. Teisest küljest, kui h_1 on tihedus, st $\nu = h_1\mu$ ning $\mu\{h_1 \neq h_2\} = 0$, siis on ka h_2 tihedus, st $\nu = h_2\mu$.

- Mõõt $\nu = h\mu$ on lõplik parajasti siis, kui $h \in \mathcal{L}_1(S, \Sigma, \mu)$.

- Kui $\nu = h\mu$ on lõplik, siis $\mu\{h = \infty\} = 0$. Sama kehtib, kui ν on σ-lõplik. Sellisel juhul leidub lõplik tihedus $h : S \to [0, \infty)$. Tõestuseks veendu, et kui $\mu\{h = \infty\} > 0$, siis ei saa $h\mu$ olla σ-lõplik.

- Kui $h : S \to [0, \infty)$ ja μ on σ-lõplik, on seda ka $\nu = h\mu$.

  Tõestuse idee: olgu (esialgu) μ lõplik. Defineerime hulgad: $S_n := \{x \in S : h(x) \leq n\}$. Et h on lõplik, siis $S_n \nearrow S$. Hulgal S_n on h tõkestatud ja $\nu(S_n) \leq n\mu(S_n) \leq n\mu(S) < \infty$. Seega ν on σ-lõplik.
  Kui μ on σ-lõplik, siis leiduvad hulgad A_i nii, et $\cup_i A_i = S$ ja $\mu(A_i) < \infty$. Olgu S_n nii nagu enne ja defineeri $B_{in} := S_n \cap A_i$. Hulgal B_{in} on ν lõplik.

7.10.1 Näited

Absoluutselt pidev tõenäosusmõõt. Olgu P reaalteljel (s.o. ruumil $(\mathbb{R}, \mathcal{B})$ antud absoluutselt pidev mõõt. Vastavalt peatükis 2.5 toodud definitisioonile tähendab see, et leidub Riemanni mõttes integreeruv funktsioon f nii, et iga x korral

$$F(x) = \int_{-\infty}^{x} f(t)dt. \tag{7.22}$$

Teoreemi 7.14 põhjal

$$\int_{-\infty}^{x} f(t)dt = \int_{(-\infty, x]} f d\text{Leb}.$$

Seega iga x korral

$$P(-\infty, x] = F(x) = \int_{-\infty}^{x} f(t)dt = \int_{(-\infty, x]} f d\text{Leb}.$$

Et f on ka Lebesque'i mõttes integreeruv (Lebesgue'i mõõdu järgi), siis leidub mõõt

$$\nu(A) = \int_A f d\text{Leb} \quad \forall A \in \mathcal{B},$$

kusjuures $\nu(-\infty, x] = F(x) = P(-\infty, x]$, $\forall x \in \mathbb{R}$. Teoreemist 2.3 järeldub, et $\nu = P$, seega on mõõdul P tihedus f Lebesgue'i mõõdu suhtes.

Laiendame peatükus 2.5 toodud definitsiooni:

Def 7.17 *Ütleme, et σ-algebral $\mathcal{B}(\mathbb{R})$ antud mõõt on absoluutselt pidev kui tal on tihedus definitsiooni 7.16 mõttes Lebesgue'i mõõdu suhtes.*

Riemanni integraali asendamine Lebesgue'i integraaliga on sisuline laiendamine, sest on tihedusfunktsioone millel pole ühtegi p.k. võrdset Riemanni mõttes integreeruvat funktsiooni. Seega leidub jaotusfunktsioone F nii, et mingi tiheduse g korral

$$F(x) = \int_{(-\infty,x]} g d\mathrm{Leb},$$

kuid pole ühtegi funktsiooni f nii, et

$$F(x) = \int_{-\infty}^{x} f(x) dx.$$

Jaotus F on ikkagi absoluutselt pidev.

Märkus: Jaotusfunktsiooni tuletisel ja tihedusfunktsiooni (Lebesgue'i mõõdu järgi) seos on üldiselt järgmine:

- Kui f on jaotuse (tõenäosusmõõdu) F tihedus, siis $F' = f$ p.k. (Lebesgue'i mõõdu suhtes). Kui tihedus on punktis x pidev, siis $F'(x) = f(x)$ (millest muuhulgas järeldub, et ei saa olemas olla kaht punktis x erinevat tihedust, mis mõlemad on selles punktis ka pidevad).

- Vastupidine üldiselt ei kehti: kui $F' = f$ p.k. (Lebesgue'i mõõdu suhtes), siis f ei pruugi olla mõõdu F tihedus (tooge näide).

- Tiheduse olemasolul peab F olema pidev. Kuid ka sellest ei piisa: ka siis kui $F' = f$ (Lebesgue'i mõõdu suhtes) p.k.ja F on pidev, ei pruugi f olla mõõdu F tihedus. Kontranäide: singulaarne jaotusfunktsioon.

- Küll aga on funktsioon f nii, et $F' = f$ p.k. tihedus, kui üks järgnevatest tingimustest on täidetud:

  1. $F'(x) = f(x)$ iga x korral.
     (Tuleneb sellest, et monotoonse funktsiooni tuletis on integreeruv (*Billingsley*, Thm 31.2, mittemonotoonse funktsiooni korral ei pruugi see alati nii olla); teisalt kui $F'(x) = f(x)$ iga x korral ja f on integreeruv, siis f on tihedus (see on nn *fundamental theorem of calculus*).
  2. Kui $\int f(s)ds = 1$.
     (Tuleneb Lebesgue'i dekompositsioonist: $F(x) = \gamma F_a(x) + (1-\gamma)F_s(x)$. Et $F_s' = 0$ p.k, siis $F_a' = F' = f$ p.k., millest $F_a(x) = \int_{(-\infty,x]} f(s)ds$ (sest absoluutselt pideva finktsiooni tuletis on p.k. võrdne tihedusega). Seega tingimus $\int f(s)ds = 1$ on ekvivalentne tingimusega $\lim_{x\to\infty} F_a(x) = 1$ ehk $\gamma = 1$.)

Diskreetne tõenäosusmõõt diskreetsel tõenäosusruumil. Olgu $(\Omega, \mathcal{F}, \mathbf{P})$ tõenäosusruum, kusjuures $\Omega = \{\omega_1, \omega_2, \ldots\}$, $\mathcal{F} = 2^\Omega$ ning $\mathbf{P}$ on diskreetne tõenäosusmõõt, s.t. leidub $p : \Omega \to [0,1]$ nii, et

$$\mathbf{P}(A) = \sum_{\omega_i \in A} p(\omega_i). \tag{7.23}$$

Paneme tähele, et seosest (7.12) saame

$$\sum_{\omega_i \in A} p(\omega_i) = \int_A p d\mu,$$

kus μ on loendav mõõt. Seega on mõõdul P tihedus loendava mõõdu suhtes.

Diskreetne (tõenäosus)mõõt reaalteljel. Olgu $(S,\Sigma)=(\mathbb{R},\mathcal{B})$ ning vaatleme mõõtu ν (2.1), s.t.

$$\nu=\sum_{i=1}^{\infty}p_i\delta_{x_i},\quad \{x_i\}\subset\mathbb{R},\quad p_i\geq 0 \text{ ja iga } A\in\mathcal{B},\quad \nu(A)=\sum_{x_i\in A}p_i.$$

Mõõt ν on tõenäosusmõõt, kui $\sum p_i=1$. Seose (7.21) tõttu iga $A\in\mathcal{B}$ korral

$$\nu(A)=\sum_{x_i\in A}p_i=\int_A f d\mu,$$

kus μ on loendav mõõt ning

$$f=\sum_{i=1}^{\infty}p_i I_{\{x_i\}}. \tag{7.24}$$

Seega on mõõdul ν tihedus (7.24) loendava mõõdu suhtes.

Märkus: Tihti esitatakse tihedust (7.24) ka kujul

$$\sum_i p_i\delta_{x_i}.$$

Sellisel juhul on mõõdul ν ning tihedusel (7.24) formaalselt sama kuju, kuid nad on aga erinevad matemaatilised objektid: üks on funktsioon $\mathcal{B}\to[0,\infty)$ (mõõt), teine aga on funktsioon $\mathbb{R}\to[0,\infty)$ (tihedus).

Kaks viimast näidet veensid meid, et ka **diskreetsel tõenäosusmõõdul on tihedus!** Kuid loendava mõõdu suhtes.

7.10.2 Integreerimine tihedusega mõõdu järgi

Teoreem 7.18 *Olgu h mõõdu ν tihedus μ suhtes.*

i) *iga $f\in m\Sigma^+$ korral*

$$\int_A f d\nu=\int_A fh\,d\mu\quad \forall A\in\Sigma. \tag{7.25}$$

ii) *$f\in\mathcal{L}_1(S,\Sigma,\nu)$ parajasti siis, kui $fh\in\mathcal{L}_1(S,\Sigma,\mu)$ ja sellisel juhul kehtib (7.25).*

Tõestus. Tõestus koosneb kolmest osast.
i) Tõestame, et iga $f\in m\Sigma^+$ korral

$$\int f d\nu=\int fhd\mu. \tag{7.26}$$

Selleks kasuta standardset meetodi:
Kui $f = I_B$, siis

$$\int f d\nu = \nu(B) = \int_B h d\mu = \int h f d\mu.$$

Kui $f = \sum_i a_i I_{B_i}$, siis

$$\int f d\nu = \sum_i a_i \nu(B_i) = \sum_i a_i \int_{B_i} h d\mu = \int h(\sum_i a_i I_{B_i}) d\mu = \int h f d\mu.$$

Kui $f \in m\Sigma^+$, siis leidub lihtsate funktsioonide jada f_n nii, et $f_n \nearrow f$. Sellisel juhul ka $f_n h \nearrow fh$, ning monotoonse koondumise teoreemist järeldub (7.26).
ii) Tõestame, et $f \in \mathcal{L}_1(S, \Sigma, \nu)$ parajasti siis, kui $fh \in \mathcal{L}_1(S, \Sigma, \mu)$ ning sellisel juhul kehtib (7.26). Integreeruvuse tõestamiseks rakenda tõestuse eelmist osa funktsioonile $|f|$, seose (7.26) saamiseks kasuta asjaolu, et (7.26) kehtib mittenegatiivsete funktsioonide korral. Sellisel juhul

$$\int f d\nu = \int f^+ d\nu - \int f^- d\nu = \int f^+ h d\mu - \int f^- h d\mu = \int f h d\mu.$$

Kolmandaks: (7.25) saamiseks rakenda seost (7.26) funktsioonile fI_A. (ülesanne 18). ■

Järeldus 7.4 *Olgu μ, ν, ρ mõõdud σ-algebral Σ, kusjuures*

$$g = \frac{d\rho}{d\nu} \quad \text{ja} \quad h = \frac{d\nu}{d\mu}.$$

Siis

$$\frac{d\rho}{d\mu} = gh,$$

s.t.

$$g(h\mu) = (gh)\mu. \tag{7.27}$$

Tõestus. Ülesanne ■

7.10.3 Näiteid teoreemi 7.18 kasutamisest

1 Olgu P absoluutselt pidev tõenäosusmõõt Boreli σ-algebral $\mathcal{B}(\mathbb{R})$. Olgu f selle mõõdu tihedus. Kui

$$g : \mathbb{R} \mapsto \mathbb{R}$$

on P suhtes integreeruv funktsioon, siis seosest (7.25)saame

$$\int_A g dP = \int_A g f d\text{Leb}.$$

Kui f on Riemanni mõttes integreeruv ja g on Riemanni mõttes integreeruv, siis on seda ka gf ning võttes $A = [a, b]$, saame

$$\int_{[a,b]} g dP = \int_a^b g(x) f(x) dx. \tag{7.28}$$

Näiteks kui mõõdu P jaotusfunktsioon F on igas punktis diferentseeruv, siis P tihedus on $f = F'$ ning iga pideva g korral kehtib (7.28).

2 Olgu $(\Omega, \mathcal{F}, \mathbf{P})$ tõenäosusruum, kus $\Omega = \{\omega_1, \omega_2, \ldots\}$, $\mathcal{F} = 2^\Omega$ ja $\mathbf{P}$ olgu diskreetne tõenäosusmõõt (7.23). Seega on mõõdul $\mathbf{P}$ tihedus $\omega \mapsto p(\omega)$ loendava mõõdu μ suhtes.
Kui

$$g : \Omega \mapsto \mathbb{R}$$

on $\mathbf{P}$ suhtes mõõtuv, siis (7.26):

$$\int g d\mathbf{P} = \int g p d\mu$$

ning (7.14):

$$\int g d\mathbf{P} = \int g p d\mu = \sum_i g(\omega_i) p(\omega_i) = \sum_i g(\omega_i) p_i,$$

kus $p(\omega_i) = p_i$.

3 Olgu $(S, \Sigma) = (\mathbb{R}, \mathcal{B})$. Vaatleme mõõtu (2.1)

$$\nu = \sum_{i=1}^{\infty} p_i \delta_{x_i}, \quad \{x_i\} \subset S, \quad p_i \geq 0$$

tihedusega

$$f = \sum_{i=1}^{\infty} p_i I_{\{x_i\}}$$

loendava mõõdu μ suhtes.
kui $g : \mathbb{R} \mapsto \mathbb{R}$ on ν suhtes integreeruv, siis (7.25) ja (7.21):

$$\int_A g d\nu = \int_A g f d\mu = \sum_{x_i \in A} g(x_i) f(x_i) = \sum_{x_i \in A} g(x_i) p_i. \tag{7.29}$$

7.10.4 Scheffe teoreem koonduvatest tihedustest

Teoreem 7.19 (Scheffe) *Olgu ν_n ja ν sellised lõplikud mõõdud, et*

$$\nu_n(S) = \nu(S) < \infty, \quad \forall n. \tag{7.30}$$

Olgu $f = \frac{d\nu}{d\mu}$, $f_n = \frac{d\nu_n}{d\mu}$. Kui $f_n \to f$ p.k., siis

$$\sup_{A \in \Sigma} \Big| \nu(A) - \nu_n(A) \Big| \leq \int |f_n - f| d\mu \to 0. \tag{7.31}$$

Tõestus. Vastavalt funktsioonide f_n ja f definitsioonile ning integraali omadustele **C1, C5, C3**, iga $A \in \Sigma$ korral kehtib

$$\Big| \int_A f_n d\mu - \int_A f d\mu \Big| \leq \int_A |f_n - f| d\mu \leq \int |f_n - f| d\mu.$$

Seega kehtib võrratus (7.31).
Paneme tähele, et

$$\int |f_n| d\mu = \int f_n d\mu = \nu_n(S) = \nu(S) = \int f d\mu = \int |f| d\mu.$$

Seega kehtib

$$\int |f_n| d\mu \to \int |f| d\mu$$

ning Sheffe lemmast järeldub, et $\int |f_n - f| d\mu \to 0$. ∎

Näide: Scheffe teoreem on väga kasulik statistikas. Olgu antud tiheduste hulk $\{f(x, \theta) : \theta \in \Theta\}$, kus θ on jaotuse parameeter. Olgu θ^* meile tundmatu üldkogumi parameeter ning θ_n selline parameetrite jada, et $\theta_n \to \theta^*$. Parameeterid θ_n võib interpreteerida θ^* hinnangutena. Oletame nüüd, et tihedusfunktsioonidel kehtib järgmine omadus: kui $\theta_n \to \theta$, siis $f(x, \theta_n) \to f(x, \theta)$ iga x korral. Selline omadus on paljudel praktilistel mudelitel, näiteks normaaljaotusel, Poissoni jaotusel, eksponentjaotusel, geomeetrilisel jaotusel (veendu selles). Punktiviisilisest ehk kõikjal koondumisest järeldub loomulikult, et need funktsioonid koonduvad ka μ-p.k. (ükskõik, mis μ meil ka poleks). Scheffe teoreemist järeldub nüüd, et kui $\theta_n \to \theta^*$, siis iga Boreli hulga B korral

$$|P_n(B) - P(B)| \leq \sup_{B \in \mathcal{B}} |P_n(B) - P(B)| \leq \int |f(x, \theta_n) - f(x, \theta^*)| d\mu \to 0,$$

kus

$$P_n(B) = \int_B f(x, \theta_n) \mu(dx), \quad P(B) = \int_B f(x, \theta^*) \mu(dx).$$

Seega parameetrite koondumisest järeldub mistahes Boreli hulkade tõenäosuste (ühtlane) koondumine.

7.10.5 Radon-Nikodymi teoreem

Olgu h mõõdu ν tihedus μ suhtes. Sellisel juhul kehtib seos (miks?)

$$\mu(A) = 0 \quad \Rightarrow \quad \nu(A) = 0. \tag{7.32}$$

Def 7.20 *Kui (7.32) kehtib, siis ütleme, et ν on* **absoluutselt pidev** *μ suhtes ja kirjutame $\nu \ll \mu$.*

Kui μ on Lebesgue'i mõõt, siis ütleme lihtsalt, et ν (või P) on absoluutselt pidev.

Näited:

- Iga mõõt on absoluutselt pidev loendava mõõdu suhtes, sest $\emptyset$ on ainus nullmõõduline hulk (loendava mõõdu järgi);
- olgu δ_x Diraci mõõt. Siis $\nu \ll \delta_x$ parajasti siis, kui $\nu = a\delta_x$, kus $a \geq 0$.

Märkus: Saab näidata, et lõplik mõõt ν on absoluutselt pidev μ suhtes parajasti siis, kui iga $\epsilon > 0$ korral leidub $\delta > 0$ nii, et

$$\mu(A) \leq \delta \Rightarrow \nu(A) \leq \epsilon.$$

Olgu μ ja ν mõõdud σ-algebral Σ. Millal on mõõdul ν tihedus mõõdu μ suhtes? Ülaltoodust on selge, et mõõdu ν absoluutne pidevus μ suhtes on tarvilik. Selgub, et see on ka piisav.

Teoreem 7.21 (Radon-Nikodymi teoreem)
Olgu μ ja ν ruumil (S, Σ) antud σ-lõplikud mõõdud. Kui ν on absoluutselt pidev μ suhtes, siis mõõdul ν on tihedus h μ suhtes. Tihedus h on μ-p.k. ühene.

Näited:

1 Olgu S mitteloenduv, koosnegu Σ ülimalt loenduvatest hulkadest ja nende täienditest. Olgu μ loendav mõõt ning

$$\nu : \Sigma \to [0, \infty], \quad \nu(A) = \begin{cases} 0 & \text{kui } A \text{ on ülimalt loenduv,} \\ \infty & \text{kui } A^c \text{ on ülimalt loenduv.} \end{cases}$$

Kehtib $\nu \ll \mu$. Küll aga pole mõõdul ν tihedust μ suhtes. Kui leiduks $\nu = f\mu$, siis iga $s \in S$ korral

$$0 = \nu(\{s\}) = \int_{\{s\}} f d\mu = f(s)\mu(\{s\}) = f(s).$$

See saab aga olla vaid siis kui $f = 0$, millest $\nu = 0$?! Mõõdud μ ja ν pole σ-lõplikud.

2 Olgu $(S,\Sigma) = (\mathbb{R},\mathcal{B})$. Olgu μ loendav mõõt ja $\nu = \text{Leb}$. Iga mõõt on absoluutselt pidev μ suhtes, kuid $\text{Leb} = f\mu$ tähendaks teoreemi 7.15 põhjal, et $f = 0$-Leb p.k. ehk leidub ülimalt loenduv hulk $\{x_1, x_2, \ldots\}$ nii, et $f = \sum_{x_i} a_i I_{\{x_i\}}$. Järelikult iga $A \in \mathcal{B}$ korral $\text{Leb}(A) = \sum_{x_i \in A} a_i$. Seega ka $\text{Leb}\{x_i\} = a_i$, millest

$$\text{Leb}(A) = \sum_{x_i \in A} a_i = \sum_{x_i \in A} \text{Leb}(\{x_i\}) = 0 \quad ?!$$

3 Olgu ν_1 ja ν_2 σ-lõplikud mõõdud ruumil (S,Σ), olgu

$$\mu := \nu_1 + \nu_2.$$

Siis μ on σ-lõplik ja $\nu_i \ll \mu$, $i = 1,2$. Radon-Nikodymi teoreemi järgi on mõõdul ν_i tihedus mõõdu μ suhtes.

7.11 Muutuja vahetus

Olgu mõõduga ruumi (S,Σ,μ) kõrval veel teinegi mõõtuv ruum (S',Σ'). Defineerime mõõdu μh^{-1} üldistuse.

Funktsioon $T : S \to S'$ olgu selline, et

$$T^{-1}(B) \in \Sigma \quad \forall B \in \Sigma',$$

(ütleme, et T on Σ/Σ'-mõõtuv). Defineerime ruumil (S',Σ') mõõdu

$$\mu T^{-1} : \Sigma' \to [0,\infty], \quad \mu T^{-1}(A') = \mu(T^{-1}A').$$

Kui $(S',\Sigma') = (\mathbb{R},\mathcal{B})$ ja $T = h$, siis

$$\mu T^{-1} = \mu h^{-1}.$$

Kui $f \in m\Sigma'$, siis $f \circ T \in m\Sigma$ ja kehtib

Teoreem 7.22

i) *Kui $f \in m\Sigma'^{+}$, siis $\forall A' \in \Sigma'$ korral*

$$\int_{T^{-1}A'} f(Ts)\mu(ds) = \int_{A'} f(s')\mu T^{-1}(ds'); \tag{7.33}$$

ii) *funktsioon $f \in \mathcal{L}_1(S',\Sigma',\mu T^{-1})$ parajasti siis kui $f \circ T \in \mathcal{L}_1(S,\Sigma,\mu)$ ja siis kehtib (7.33).*

Tõestus. Tõestus on analoogiline teoreemi 7.18 tõestusega. See koosneb kolmest osast:
1. Standardset meetodi kasutades tõestada, et kui $f \in m\Sigma'^{+}$, siis

$$\int_S f(Ts)\mu(ds) = \int_{S'} f(s')\mu T^{-1}(ds'). \tag{7.34}$$

Algus: võta $f = I_{B'}$. Funktsioon $s \mapsto I_{B'}(Ts)$ võrdub funktsiooniga $s \mapsto I_{T^{-1}B'}(s)$, mistõttu

$$\int I_{B'}(Ts)\mu(ds) = \int I_{T^{-1}B'}(s)\mu(ds) = \mu T^{-1}(B') = \int I_{B'}(s')\mu T^{-1}(ds').$$

Oleme tõestanud, et kui $f = I_{B'}$ ja $A' = S'$, siis võrdus (7.33) tõepoolest kehtib.

2. Rakenda saadud tulemust funktsioonile $|f|$ ja tõesta, et $f \in \mathcal{L}_1(S', \Sigma', \mu T^{-1})$ parajasti siis kui $f \circ T \in \mathcal{L}_1(S, \Sigma, \mu)$ ja siis kehtib (7.34).
3. Asenda f funktsiooniga fI_A ning tõesta teoreem. (ülesanne 19). ■

Näide: Olgu $(S, \Sigma, \mu) = (\Omega, \mathcal{F}, \mathbf{P})$ tõenäosusruum. Ruum (S', Σ') olgu $(\mathbb{R}, \mathcal{B})$. Vaatleme juhuslikku suurust $X : \Omega \mapsto \mathbb{R}$. Juhuslik suurus X tekitab ruumil $(\mathbb{R}, \mathcal{B})$ mõõdu P_X - jaotuse.

- Olgu $g : \mathbb{R} \mapsto [0, \infty]$ mittenegatiivne Boreli funktsioon. Vastavalt teoreemi 7.22 osale **i)**

$$\int_{X^{-1}(B)} g(X(\omega))\mathbf{P}(d\omega) = \int_B g(x)P_X(dx), \quad \forall B \in \mathcal{B}. \tag{7.35}$$

  Kui mõõdul P_X on tihedus f (Boreli σ-algebral $\mathcal{B}$) Lebesgue'i mõõdu suhtes, siis teoreemi 7.18 **i)** tõttu

$$\int_{X^{-1}(B)} g(X(\omega))\mathbf{P}(d\omega) = \int_B g(x)P_X(dx) = \int_B g(x)f(x)\mathrm{Leb}(dx). \tag{7.36}$$

- Olgu nüüd $g : \mathbb{R} \mapsto [-\infty, \infty]$ Boreli funktsioon. Vastavalt teoreemi 7.22 osale **ii)**

$$g(X) \in \mathcal{L}_1(\Omega, \mathcal{F}, \mathbf{P}) \quad \Leftrightarrow \quad \int |g(X(\omega))|\mathbf{P}(d\omega) < \infty$$

  parajasti siis, kui

$$g \in \mathcal{L}_1(\mathbb{R}, \mathcal{B}, P_X) \quad \Leftrightarrow \quad \int |g(x)|P_X(dx) < \infty \tag{7.37}$$

  ning sellisel juhul kehtib (7.35). Kui mõõdul P_X on tihedus f, siis teoreemi 7.18 väitest **ii)** saame, et (7.37) kehtib parajasti siis, kui

$$gf \in \mathcal{L}_1(\mathbb{R}, \mathcal{B}, \mathrm{Leb}) \Leftrightarrow \int |g(x)f(x)|\mathrm{Leb}(dx) < \infty. \tag{7.38}$$

  ning siis kehtib ka (7.36). Erijuhul $B = \mathbb{R}$, saame seega

$$\int g(X)d\mathbf{P} = \int gf d\mathrm{Leb}. \tag{7.39}$$

7.12 Ülesanded

1. Tõestada, et lihtsa funktsiooni definitsioonid **D1**, **D2** ja **D3** on kõik ekvivalentsed definitsiooniga 7.1.

2. Olgu f mittenegatiivne mõõtuv funktsioon, kusjuures $fS) = \{b_1, \ldots, b_m\}$. Tõesta, et kui $A_i \in \Sigma$ ja $a_i \geq 0$ on sellised, et $f = \sum_{i=1}^n a_i I_{A_i}$, siis

$$\sum_{i=1}^n a_i \mu(A_i) = \sum_{j=1}^m b_j \mu(B_j), \text{ kus } B_j := \{s : f(s) = b_j\}. \tag{7.40}$$

Järeldada, et lihtsa mittenegatiivse mõõtuva funktsiooni Lebesgue'i integraal on korrektselt defineeritud, sest $\int fg\mu$ ei sõltu funktsiooni f esitusest.

Näpunäide: Hulgad $\{B_i\}$ moodustavad tükelduse. Olgu $\sum_{i=1}^n a_i I_{A_i}$ funktsiooni f mingi teine esitus. Siin hulgad A_i on kõik küll mõõtuvad, kuid ei pruugi moodustada tükeldust ja arvud a_i võivad olla korduvad. Vaja tõestada, et kehtib (7.40).
1) Kasutades asjaolu, et $\{B_i\}$ on tükeldus näita, et (7.40) kehtib, kui iga $j = 1, \ldots, m$ korral

$$\sum_{i=1}^n a_i \mu(A_i \cap B_j) = b_j \mu(B_j). \tag{7.41}$$

Samuti veendu, et $f = \sum_{i=1}^n a_i I_{A_i}$ on ekvivalentne seosega: iga $j = 1, \ldots, m$ korral

$$\sum_{i=1}^n a_i I_{A_i \cap B_j} = b_j I_{B_j}. \tag{7.42}$$

Järelda, et ülesande lahendamiseks piisab, kui iga j korral näitad, et võrdusest (7.42) järeldub võrdus (7.41).
2) Näitame nüüd, et võrdusest (7.42) järeldub võrdus (7.41). Fikseeri j, olgu $b := b_j$, $B := B_j$ ja $C_i := A_i \cap B$. Kehtigu $\sum_{i=1}^n a_i I_{C_i} = bI_B$. Vaja näidata, et

$$\sum_{i=1}^n a_i \mu(C_i) = b\mu(B). \tag{7.43}$$

Defineeri hulgad D_k, $k = 1, \ldots, 2^n$ järgmiselt: $D_k = \cap_i E_i$, kus E_i on kas C_i või $B \setminus C_i$. Veendu, et $\{D_1, \ldots, D_{2^n}\}$ on hulga B tükeldus ning et iga C_i ja D_k korral kas $D_k \subset C_i$ või $D_k \cap C_i = \emptyset$. Sellest järelda, et iga k korral

$$\Big(\sum_{i=1}^n a_i I_{C_i}\Big) I_{D_k} = \sum_{i=1}^n a_i I_{C_i \cap D_k} = \sum_{i=1}^n d_i^k I_{D_k}, \quad \text{kus } d_i^k = \begin{cases} a_i, & \text{kui } D_k \subset C_i \ ; \\ 0, & \text{mujal.} \end{cases}$$

Järelda, et iga k korral $\sum_{i=1}^n d_i^k = b$. Veendu, et $\sum_{i=1}^n a_i \mu(C_i \cap D_k) = \sum_{i=1}^n c_i^k \mu(D_k) = b\mu(D_k)$. Nüüd kasuta asjaolu, et $\{D_1, \ldots, D_{2^n}\}$ on hulga B tükeldus ja näita, et (7.43) kehtib.

3. Olgu $(S, \Sigma, \mu) = ([0,1], \mathcal{B}([0,1]), Leb)$, $f = I_A$, kus A-kõik ratsionaalarvud lõigus $[0,1]$.

- Kas $f \in m\Sigma$?
- Kas $f \in SF^+$?
- Kas $f \in \mathcal{L}_1(S, \Sigma, \mu)$?
- Leida $\int f\, d\mu$.
- Kas f on Riemanni mõttes integreeruv?

4. Tõestada omadused **A1**,**A2**, **A3**.

5.

a) Tõestada definitsioonide 7.3 ja 7.4 ekvivalentsus.

b) Defineerime

$$\int^* f d\mu = \inf\{\Sigma_i(\sup_{s \in A_i} f(s))\mu(A_i))\},$$

kus alumine raja on võetud üle hulga S kõigi lõplike Σ-mõõtuvate tükelduste.

- Tõestada, et kui $\mu\{s|f(s) > 0\} = \infty$, siis $\int^* f d\mu = \infty$.
- Tõestada, et kui $\mu\{s|f(s) > a\} > 0 \ \forall a > 0$, siis $\int^* f d\mu = \infty$.
- Kas $\int^* f d\mu = \int f d\mu$?

6. Kuulugu $f \in m\Sigma^+$. Tõestada

1) kui $f = 0$ p.k., siis $\int f\, d\mu = 0$;

2) kui $\int f d\mu < \infty$, siis $\mu\{f = \infty\} = 0$.

7. Tõestada

1) $f \in \mathcal{L}_1(S, \Sigma, \mu)$ parajasti siis, kui $\int |f| d\mu < \infty$

2) Integraali omadused **C1, C2, C3, C4**.

8. Olgu $(S, \Sigma, \mu) = ([0,1], \mathcal{B}([0,1]), Leb)$.

- Defineerime funktsioonid $f_n = n^2 I_{(0,n^{-1})}$.
  Leida $\int \liminf_n f_n d\mu$ ja $\liminf_n \int f_n d\mu$.
  Kas Fatou lemma kehtib?
  Kas $f_n \to 0$ p.k. ?
  Kas kehtivad domineeritud koondumise teoreemi eeldused ja väide?

- Defineerime funktsioonid $f_n = nI_{(0,n^{-1})}$.

  Kas kehtivad domineeritud koondumise teoreemi eeldused ja väide?

9. Olgu $f_n \in \mathcal{L}_1(S, \Sigma, \mu)$, kusjuures $\sup_n \int f_n d\mu < \infty$. Monotoonse koondumise teoreemi kasutades näidata, et kui $f_n \nearrow f$, siis $f \in \mathcal{L}_1(S, \Sigma, \mu)$ ja $\int f_n d\mu \to \int f d\mu$.

10. Olgu $a_n, f_n, b_n \in m\Sigma$, kusjuures $a_n \to a$ p.k. $b_n \to b$ p.k. ja $f_n \to f$ p.k. Oletame, et $a_n, b_n, a, b \in \mathcal{L}_1(S, \Sigma, \mu)$, kusjuures $\int a_n d\mu \to \int a d\mu$ ja $\int b_n d\mu \to \int b d\mu$. Lõpuks oletame, et $a_n \geq f_n \geq b_n$ p.k. Tõestada, et $\int f_n d\mu \to \int f d\mu$. Järelda domineeritud koondumise teoreem. (Fatou lemmat kasutades näidata, et $\limsup_n \int f_n d\mu \leq \int f d\mu \leq \liminf_n \int f_n d\mu$).

11. Olgu $f_n \in m\Sigma^+$. Tõestada, et

$$\int \sum_n f_n d\mu = \sum_n \int f_n d\mu \leq \infty.$$

12. Olgu $f_n \in m\Sigma$, $n = 1, 2, \ldots$. Tõestada, et kui $\sum_n f_n$ koondub p.k. ja leidub $g \in \mathcal{L}_1(S, \Sigma, \mu)$ nii, et $|\sum_{k=1}^n f_k| \leq g$, siis $f_n \in \mathcal{L}_1(S, \Sigma, \mu)$ $\forall n$ ja $\sum f_n \in \mathcal{L}_1(S, \Sigma, \mu)$, kusjuures $\int \sum_n f_n d\mu = \sum_n \int f_n d\mu$.(Kasutada domineeritud koondumise teoreemi).

13. Olgu $f_n \in m\Sigma$, $n = 1, 2, \ldots$ ja $\sum_n \int |f_n| d\mu < \infty$. Veenduda, et rida $\sum_n f_n$ koondub p.k., $\sum_n f_n \in \mathcal{L}_1(S, \Sigma, \mu)$ ja

$$\int \sum_n f_n d\mu = \sum_n \int f_n d\mu.$$

(Rakendada ülesannet 11, integraali omadusi ja ülesannet 12, võttes $g = \sum_n |f_n|$).

14. Eelmisi ülesandeid kasutades tõestada teoreem 7.12.
(rakendada ülesannet 11 juhul, kui $f \in m\Sigma^+$ ja ülesannet 13 juhul, kui f in integreeruv).

15. Olgu $f \in m\Sigma^+$. Tõestada, et $\nu(A) = \int_A f d\mu$ on mõõt. Millal ν on lõplik mõõt? Näidata, et $\nu(A) = 0$, kui $\mu(A) = 0$.

16. Olgu $f \in \mathcal{L}_1(S, \Sigma, \mu)$. Tõestada, et iga $\epsilon > 0$ korral $\mu\{s||f(s)| \geq \epsilon\} < \infty$.

17. Olgu $f, g \in \mathcal{L}_1(S, \Sigma, \mu)$, kusjuures $\int f d\mu = \int g d\mu$. Tõestada, et kui $\int_A f d\mu = \int_A g d\mu$ iga $A \in \pi$ korral, kus π on selline π-süsteem, et $\sigma(\pi) = \Sigma$, siis $f = g$ p.k.
(Kui $f, g \in m\Sigma^+$,siis rakendada teoreemi 7.13 ja mõõdu jätkamise teoreemi. Üldisel juhul tõestada, et $f^+ + g^- = f^- + g^+$ p.k.).

18. Olgu mõõdul ν on tihedus h mõõdu μ suhtes. Standardset meetodi kasutades tõestada

1) iga $f \in m\Sigma^+$ korral $\int f d\nu = \int fh\, d\mu$;

2) $f \in \mathcal{L}_1(S, \Sigma, \nu)$ parajasti siis, kui $fh \in \mathcal{L}_1(S, \Sigma, \mu)$ ja sellisel juhul $\int f d\nu = \int fh\, d\mu$
(Rakendada osa 1) funktsioonile $|f|$);

3) teoreem 7.18.
(Järeldub eelmisest kahest väitest, kui võtta fI_A.)

19. Standardset meetodi kasutades tõestada

1) Kui $f \in m\Sigma^+$, siis kehtib (7.34). (Tõestada, et $I_{A'}(Ts) = I_{T^{-1}A'}$, edasi standardne);

2) Funktsioon $f \in \mathcal{L}_1(S', \Sigma', \mu T^{-1})$ parajasti siis kui $f \circ T \in \mathcal{L}_1(S, \Sigma, \mu)$ ja sellisel juhul kehtib (7.34)
(Rakendada osa 1) funktsioonile $|f|$);

3) Tõestada teoreem 7.22.
(Järeldub eelmisest kahest väitest, kui võtta fI_A).

20. Tõestada pööratud Fatou lemma.

21. Tõestada tõkestatud koondumise teoreem.

22. Olgu P_1 ja P_2 reaalteljel (so ruumil $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ antud tõenäosusmõõdud, μ mingi reaalteljel antud mõõt. Tõestada, et kui mõõdul P_i on tihedus f_i mõõdu μ suhtes, st $\frac{dP_i}{d\mu} = f_i$, $i = 1, 2$ ja kui $P_2 \ll P_1$, siis

$$\frac{dP_2}{dP_1} = \frac{f_2}{f_1}.$$

23. Olgu ν ja μ reaalteljel (so ruumil $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ antud lõplikud mõõdud.

A kas $\nu \ll \mu$?

B kui jah, siis leida $\frac{d\nu}{d\mu}$, kui

1. $\nu = a\mu$, $a > 0$.
2. $\nu = Po(4)$, $\mu = Po(2)$. Siin $Po(\lambda)$ on Poissoni jaotus parameetriga λ.
3. Olgu $P_1(k) = \frac{1}{4}$, $k = 1, 2, 3, 4$ ja $P_2(k) = \frac{1}{6}$, $k = 1, 2, 3, 4, 5, 6$.
    - $\nu = P_2$, $\mu = P_1$;
    - $\nu = P_1$, $\mu = P_2$.
4. $\nu = Exp[1]$ ja $\mu = Exp[2]$. Siin $Exp[\lambda]$ on eksponentjaotus parameetriga λ.
5. Olgu $U = U[-1, 1]$ (ühtlane), $\mu = \frac{1}{3}\delta_0 + \frac{2}{3}U$ ning ν olgu: a) $\nu = \delta_0$ b) $\nu = U$.

24. Olgu $(S, \sigma) = ([0,1], \mathcal{B}[0,1])$. Olgu funktsioon f järgmine

$$f(s) = \begin{cases} 5, & \text{kui } s = \frac{k}{20}, \quad k = 1, 2, \ldots, 10 \\ 10, & \text{kui } s = \frac{k}{20}, \quad k = 11, 12, \ldots, 10 \\ 0, & \text{mujal} \end{cases}$$

Leida $\int f d\mu$, kui

1. μ on Lebesgue' i mõõt (Leb);
2. μ on loendav mõõt;
3. $\mu = 8\delta_{\frac{1}{2}}$;
4. $\mu = \frac{1}{2}\text{Leb} + \frac{1}{4}\delta_{\frac{1}{2}} + \frac{1}{4}\delta_1$;

8 Juhuslike suuruste keskväärtus, momendid, võrratused

Olgu $(\Omega, \mathcal{F}, \mathbf{P})$ tõenäosusruum, X juhuslik suurus, P tema jaotus, $P = \mathbf{P}X^{-1}$.

Def 8.1 *Integreeruva juhusliku suuruse $X \in \mathcal{L}_1(\Omega, \mathcal{F}, \mathbf{P})$* **keskväärtuseks** *$EX$ nim. funktsiooni X Lebesgue'i integraali:*

$$EX := \int X \, d\mathbf{P} = \int_\Omega X(\omega) \, \mathbf{P}(d\omega).$$

Kui X ei ole integreeruv, kuid tal on defineeritud Lebesgue'i integraal (st kas $E(X^-) < \infty$ või $E(X^+) < \infty$), siis juhusliku suuruse X keskväärtus EX on ikka tema Lebesgue integraal. Seega mittenegatiivsel juhuslikul suuruse on keskväärtus on alati defineeritud $EX := \int X d\mathbf{P} \leq \infty$.

Seega on juhusliku suuruse keskväärtusel kõik Lebesgue'i integraali omadused. Muuhulgas kehtivad ka kõik koondumisteoreemid.

8.1 Koondumisteoreemid

Olgu X_n, X juhuslikud suurused, $X_n \to X$ p.k., s.t. $\mathbf{P}\{\omega | \, X_n(\omega) \to X(\omega)\} = 1$. Kehtivad järgmised tulemused:

Teoreem 8.2 (Monotoonse koondumise teoreem)
Kui $0 \leq X_n \nearrow X$ p.k., siis $EX_n \nearrow EX$.

Teoreem 8.3 (Fatou lemma)
Kui $0 \leq X_n$ p.k., siis $EX = E \liminf_n X_n \leq \liminf_n EX_n$.

Teoreem 8.4 (Pööratud Fatou lemma)
Kui $0 \leq X_n \leq Y$ p.k. ja $EY < \infty$, siis $EX = E \limsup_n X_n \geq \limsup_n EX_n$.

Teoreem 8.5 (Domineeritud koondumise teoreem)
Kui $|X_n| \leq Y$ p.k. ja $EY < \infty$, siis $E|X_n - X| \to 0$, millest $EX_n \to EX$.

Teoreem 8.6 (Tõkestatud koondumise teoreem)
Kui $|X_n| \leq K < \infty$ p.k., siis $E|X_n - X| \to 0$, millest $EX_n \to EX$.

Teoreem 8.7 (Sheffe lemma)
$E|X_n| \to E|X| < \infty$ parajasti siis, kui $E|X_n - X| \to 0$.

8.2 Juhusliku suuruse funktsiooni keskväärtus

Olgu $g : \mathbb{R} \to \mathbb{R}$ Boreli funktsioon. Teoreemist 7.22 (seos (7.37)) teame, et juhuslik suurus $g \circ X$ on integreeruv, s.o. $g(X) \in \mathcal{L}_1(\Omega, \mathcal{F}, \mathbf{P})$ parajasti siis, kui g on jaotuse P järgi integreeruv, s.o. $g \in \mathcal{L}_1(\mathbb{R}, \mathcal{B}, P)$. Sel juhul kehtib (7.35)

$$Eg(X) = \int g(X) d\mathbf{P} = \int_\Omega g(X(\omega))\mathbf{P}(d\omega) = \int_\mathbb{R} g(x)\, P(dx) = \int g\, dP. \qquad (8.1)$$

Juhul, kui g on samasusfunktsioon: $g(x) = x$, saame

$$EX = \int X d\mathbf{P} = \int_\Omega X(\omega)\mathbf{P}(d\omega) = \int_\mathbb{R} x\, P(dx).$$

Olgu mõõdul P tihedus f mõõdu μ suhtes. Siis vastavalt teoreemile 7.18 eksisteerib $\int g\, dP$ parajasti siis, kui eksisteerib $\int g f d\mu$ ning need integraalid on võrdsed. Seega

$$Eg(X) = \int g\, dP = \int g f d\mu. \qquad (8.2)$$

8.2.1 Pidev juhuslik suurus

Def 8.8 *Funktsiooni f nim. juhusliku suuruse X* **tihedusfunktsiooniks***, kui ta on jaotuse P tihedus Lebesgue'i mõõdu suhtes. Sellisel juhul nimetame juhuslikku suurust X (ja jaotust) P* **pidevaks***.*

Seega on f selline mittenegatiivne Boreli funktsioon, et $\forall B \in \mathcal{B}$ korral

$$\mathbf{P}\{\omega | X(\omega) \in B\} = \mathbf{P}X^{-1}(B) = P(B) = \int_B f(x)\mathrm{Leb}(dx) =: \int_B f(x)\, dx.$$

Kui juhuslikul suurusel X (mõõdul P) on tihedusfunktsioon f, siis $g \in \mathcal{L}_1(\mathbb{R}, \mathcal{B}, P)$ parajasti siis, kui $fg \in \mathcal{L}_1(\mathbb{R}, \mathcal{B}, \mathrm{Leb})$ (seos (7.38)) ja kehtib (7.36), s.t.

$$Eg(X) = \int_\mathbb{R} g(x)\, P(dx) = \int g\, dP = \int g\, f\ d\mathrm{Leb} = \int_\mathbb{R} g(x) f(x)\mathrm{Leb}(dx) =: \int_\mathbb{R} g(x) f(x) dx.$$

Juhul, kui g on samasusfunktsioon, saame

$$EX = \int_\mathbb{R} x\, P(dx) = \int_\mathbb{R} x\, f(x)\mathrm{Leb}(dx) =: \int_\mathbb{R} x\, f(x) dx.$$

8.2.2 Diskreetne juhuslik suurus

Tuletame meelde, et Boreli σ-algebral antud tõenäosusmõõt P on diskreetne, kui ta on kujul

$$P = \sum_{i=1}^\infty p_i \delta_{x_i}, \quad \{x_i\} \subset \mathbb{R}, \quad p_i \geq 0, \quad \sum_i p_i = 1. \qquad (8.3)$$

Vastavalt definitsioonile iga $A \in \mathcal{B}$ korral

$$P(A) = \sum_{x_i \in A} p_i.$$

Def 8.9 *Juhuslikku suurust X nim.* **diskreetseks** *kui tema jaotus P on diskreetne mõõt.*

Seega on juhuslik suurus X diskreetne parajasti siis, kui tema jaotus P on kujul (8.3), kus

$$p_i := P\{x_i\} = \mathbf{P}X^{-1}(x_i) = \mathbf{P}\{\omega : X(\omega) = x_i\} = \mathbf{P}(X = x_i).$$

Et funktsioon

$$f = \sum_i p_i I_{\{x_i\}}$$

on mõõdu P tihedus loendava mõõdu suhtes, siis $Eg(X)$ eksisteerib parajasti siis kui eksisteerib $\int gfd\mu$, kus μ on loendav mõõt ning ning sellisel juhul saame seosest (7.29), et

$$Eg(X) = \int_{\mathbb{R}} g(x)\, P(dx) = \int_{\mathbb{R}} g(x) f(x) \mu(dx) = \sum_i g(x_i) p_i.$$

Juhul kui f on samasusfunktsioon, saame

$$EX = \int_{\mathbb{R}} x\, P(dx) = \sum_i x_i p_i.$$

8.3 Lemma keskväärtusest

Lemma 8.1 *Olgu X mittenegatiivne juhuslik suurus. Sellisel juhul*

$$EX = \int_{[0,\infty)} \mathbf{P}(X \geq t) dt = \int_{[0,\infty)} \mathbf{P}(X > t) dt = \int_{[0,\infty)} (1 - F(t)) dt. \tag{8.4}$$

Enne tõestust paneme tähele

- Monotoone funktsioon on alati Boreli-mõõtuv, seega integraal $\int \mathbf{P}(X \geq t)dt \leq \infty$ on defineeritud.
- Monotoonsel funktsioonil on vaid ülimalt loenduv hulk katkevuspunkte, mistõttu $\mathbf{P}(X = t)$ võib olla positiivne vaid loenduva arvu t korral ehk

  $$\text{Leb}\{t | \mathbf{P}(X \geq t) \neq \mathbf{P}(X > t)\} = 0.$$

  Seega

  $$\int \mathbf{P}(X \geq t) dt = \int \mathbf{P}(X > t) dt.$$

Tõestus. Standardne:
Olgu $X \geq 0$ lihtne juhuslik suurus. Siis (8.4) kehtib. Selle tõestus on ülesanne 5.
Olgu X mittenegatiivne juhuslik suurus, $X_n \nearrow X$, kus $X_n \geq 0$ on lihtsate juhuslike suuruste jada. Monotoonse koondumise teoreem (MON):

$$EX_n \nearrow EX.$$

Teisest küljest, iga t korral $\{X_n > t\} \nearrow \{X > t\}$ (miks?). Mõõdu pidevusest järeldub, et

$$\mathbf{P}(X_n > t) \nearrow \mathbf{P}(X > t)$$

iga t korral. MON: $\int \mathbf{P}(X_n > t)dt \nearrow \int \mathbf{P}(X > t)dt$. Iga n korral (8.4) kehtib, seega kehtib ta ka piiril. ■

Järeldus 8.1 *Olgu $X \in \mathcal{L}_1(\Omega, \mathcal{F}, \mathbf{P})$. Siis*

$$EX = \int_{[0,\infty)} \mathbf{P}(X > t)dt - \int_{(-\infty,0]} \mathbf{P}(X < t)dt = \int_0^\infty (1 - F(t))dt - \int_{-\infty}^0 F(t)dt.$$

Tõestus. Ülesanne 5. ■

8.4 Momendid

Olgu $0 < p < \infty$.

Def 8.10 *Keskväärtust $E|X|^p$ nim. juhusliku suuruse X* **absoluutseks p-momendiks**.

Absoluutne p-moment on alati defineeritud, kuid ei pruugi olla alati lõplik. Kui $E|X|^p < \infty$, kirjutame $X \in \mathcal{L}_p(\Omega, \mathcal{F}, \mathbf{P})$.

Def 8.11 *Olgu $X \in \mathcal{L}_p(\Omega, \mathcal{F}, \mathbf{P})$, $1 \leq p$. Juhusliku suuruse $X - EX$ absoluutset p-momenti $E|X - EX|^p$ nim. juhusliku suuruse X* **tsentreeritud absoluutseks p-momendiks**.

Def 8.12 *Olgu $k = 1, 2, \ldots$ täisarv, $X \in \mathcal{L}_k(\Omega, \mathcal{F}, \mathbf{P})$. (Lõplikku) keskväärtust EX^k nimetatakse juhusliku suuruse X* **k-momendiks**.

Def 8.13 *Olgu $k = 1, 2, \ldots$ täisarv, $X \in \mathcal{L}_k(\Omega, \mathcal{F}, \mathbf{P})$. Juhusliku suuruse $X - EX$ k-momenti $E(X - EX)^k$ nim. juhusliku suuruse X* **tsentreeritud k-momendiks**.

Def 8.14 *Juhusliku suuruse teist tsentreeritud momenti nim. juhusliku suuruse* **dispersiooniks** *ja tähistatakse $DX := E(X - EX)^2$.*

8.5 Sõltumatus

Def 8.15 *Juhuslike suuruste* X, $Y \in \mathcal{L}_2(\Omega, \mathcal{F}, \mathbf{P})$ **kovariatsiooniks** *nimetatakse keskväärtust*

$$\operatorname{cov}(YX) := E(X - EX)(Y - EY) < \infty.$$

Teoreem 8.16 *Olgu* X, Y *sõltumatud juhuslikud suurused, kusjuures* $X, Y \in \mathcal{L}_1(\Omega, \mathcal{F}, \mathbf{P})$. *Siis* $XY \in \mathcal{L}_1(\Omega, \mathcal{F}, \mathbf{P})$ *ja*

$$E(XY) = E(X)E(Y). \tag{8.5}$$

Erijuhul, kui $X, Y \in \mathcal{L}_2(\Omega, \mathcal{F}, \mathbf{P})$, *siis*

$$\operatorname{cov}(X, Y) = 0 \quad \text{ja} \quad D(X + Y) = D(X) + D(Y).$$

Tõestus. Standardne:
1) Olgu $X = I_A$, $Y = I_B$; X ja Y on sõltumatud parajasti siis, kui seda on A ja B (miks?). Seos (8.5) on nüüd ilmne.
2) Olgu

$$X = \sum_{i=1}^{m} a_i I_{A_i}, \quad Y = \sum_{i=1}^{n} b_i I_{B_i} \in SF^+,$$

kus hulgad A_i on lõikumatud, hulgad B_i on lõikumatud, $a_i \neq a_j$ ning $b_i \neq b_j$ kui $j \neq i$. Et $A_i \in \sigma(X)$ ja $B_j \in \sigma(Y)$, siis I_{A_i} ja I_{B_j} on sõltumatud. Integraali lineaarsusest järeldub nüüd (8.5):

$$\begin{aligned} EXY &= E\Big(\sum_{i=1}^{m}\sum_{j=1}^{n} a_i b_j I_{A_i \cap B_j}\Big) = \sum_{i=1}^{m}\sum_{j=1}^{n} a_i b_j \mathbf{P}(A_i \cap B_j) \\ &= \sum_{i=1}^{m}\sum_{j=1}^{n} a_i b_j \mathbf{P}(A_i)\mathbf{P}(B_j) = \big(\sum_{i=1}^{m} a_i \mathbf{P}(A_i)\big)\big(\sum_{j=1}^{n} b_i \mathbf{P}(B_j)\big) = EXEY. \end{aligned}$$

On ka selge, et kui $EX < \infty$ ja $EY < \infty$, siis $E(XY) < \infty$.
3) Olgu $X, Y \in m\mathcal{F}^+$. Olgu $X_n \nearrow X$, $Y_n \nearrow Y$ lihtsate mittenegatiivsete funktsioonide jadad nagu teoreemis 4.5. Et $X_n \in m\sigma(X)$ ja $Y_n \in m\sigma(Y)$ (miks?), siis iga n korral on X_n ja Y_n sõltumatud (miks?). Monotoonse koondumise teoreemist järeldub (8.5) kusjuures seostest $EX < \infty$ ja $EY < \infty$ järeldub $E(XY) < \infty$.
4) Olgu $X, Y \in \mathcal{L}_1(\Omega, \mathcal{F}, \mathbf{P})$. Seega $E|X| < \infty$ ja $E|Y| < \infty$ ehk $EX^+ < \infty$, $EX^- < \infty$, $EY^+ < \infty$, $EY^- < \infty$. Nüüd $|X|$ ja $|Y|$ on sõltumatud (miks?) mittenegatiivsed juhuslikud suurused, mistõttu $E|XY| = E|X||Y| = E|X|E|Y| < \infty$ ehk $XY \in \mathcal{L}_1(\Omega, \mathcal{F}, \mathbf{P})$. Mittenegatiivsed funktsioonid $X^+, X^- \in m\sigma(X)$, $Y^+, Y^- \in m\sigma(Y)$ (teoreem 4.9). Seega on kõik paarid X^+, Y^+, X^+, Y^-, X^-, Y^+, X^-, Y^- sõltumatud. Keskväärtuse lineaarsusest saame

$$\begin{aligned} E(XY) &= E\big((X^+ - X^-)(Y^+ - Y^-)\big) = E\big(X^+Y^+ - X^-Y^+ - X^+Y^- + X^-Y^-\big) \\ &= EX^+EY^+ - EX^-EY^+ - EX^+EY^- + EX^-EY^- \\ &= (EX^+ - EX^-)(EY^+ - EY^-) = EXEY. \end{aligned}$$

Kui $X \in \mathcal{L}_1(\Omega, \mathcal{F}, \mathbf{P})$, siis iga konstandi $a \in \mathbb{R}$ korral $X - a \in \mathcal{L}_1(\Omega, \mathcal{F}, \mathbf{P})$, sest $|X - a| \leq |X| + |a| \in \mathcal{L}_1(\Omega, \mathcal{F}, \mathbf{P})$. Seega, kui $X, Y \in \mathcal{L}_1(\Omega, \mathcal{F}, \mathbf{P})$, siis $(X - EX), (Y - EY) \in \mathcal{L}_1(\Omega, \mathcal{F}, \mathbf{P})$. On ka selge, et kui X, Y on sõltumatud, siis $(X - EX), (Y - EY)$ on sõltumatud (miks?). Seega $E|(X - EX)(Y - EY)| < \infty$ ning kehtib (8.5):

$$E(X - EX)(Y - EY) = E(X - EX)E(Y - EY) = 0.$$

■

Märkus. Teoreemist järeldub, et sõltumatute $X, Y \in \mathcal{L}_1(\Omega, \mathcal{F}, \mathbf{P})$ korral $XY \in \mathcal{L}_1(\Omega, \mathcal{F}, \mathbf{P})$. Sõltuvate juhuslike suuruste $X, Y \in \mathcal{L}_1(\Omega, \mathcal{F}, \mathbf{P})$ korral ei pruugi aga XY olla integreeruv (tooge kontranäide).

8.6 Võrratused

X on juhuslik suurus.

Teoreem 8.17 (Markovi võrratus)
Olgu $g : \mathbb{R} \to [0, \infty)$, mittekahanev Boreli funktsioon, $c > 0$. Siis

$$Eg(X) \geq g(c)\mathbf{P}(X \geq c).$$

Tõestus.

$$\begin{aligned} Eg(X) &= \int g(X(\omega))\mathbf{P}(d\omega) \geq \int_{\{\omega : X(\omega) \geq c\}} g(X(\omega))\mathbf{P}(d\omega) \\ &\geq \int_{\{\omega : X(\omega) \geq c\}} g(c)\mathbf{P}(d\omega) = g(c)\mathbf{P}\{\omega : X(\omega) \geq c\}. \end{aligned}$$

■

Järeldus 8.2 (Tŝebõŝevi võrratus)
Olgu $c > 0$, $X \in \mathcal{L}_2(\Omega, \mathcal{F}, \mathbf{P})$. Siis

$$\mathbf{P}(|X - EX| \geq c) \leq \frac{DX}{c^2}.$$

Tõestus. Järeldub Markovi võrratusest, kui võtta

$$g(x) = \begin{cases} x^2, & \text{kui } x > 0; \\ 0, & \text{mujal.} \end{cases}$$

ja kui X rolli võtta $|X - EX|$. ■

Teoreem 8.18 (Jenseni võrratus).
Olgu g kumer funktsioon, kusjuures $E|g(X)| < \infty$ ja $E|X| < \infty$. Siis

$$Eg(X) \geq g(EX).$$

Tõestus. Tuleta meelde kumera funktsiooni definitisioon. Kumeral funktsioonil g on omadus:

$$\forall y \in \mathbb{R} \quad \exists m(y) \in \mathbb{R}: \quad g(x) - g(y) \geq m(y)(x-y), \quad \forall x \in \mathbb{R}.$$

($m(y) = g'(y)$, kui viimane eksisteerib).
Olgu $y = EX \in \mathbb{R}$. Iga $\omega \in \Omega$ korral

$$g(X(\omega)) - g(EX) \geq m(EX)(X(\omega) - EX).$$

Võta keskväärtus ja järelda, et $Eg(X) - g(EX) \geq 0$. ■

Teoreem 8.19 (Hölderi võrratus)
Kuulugu $X \in \mathcal{L}_p(\Omega, \mathcal{F}, \mathbf{P})$, $Y \in \mathcal{L}_q(\Omega, \mathcal{F}, \mathbf{P})$, *kusjuures* $p > 1$, $q > 1$, $p^{-1} + q^{-1} = 1$. *Siis* $XY \in \mathcal{L}_1(\Omega, \mathcal{F}, \mathbf{P})$, *kusjuures*

$$E|XY| \leq \left(E|X|^p\right)^{\frac{1}{p}} \left(E|Y|^q\right)^{\frac{1}{q}} < \infty.$$

Tõestus. Kasutame asjaolu, et kui $a, b \geq 0$, siis

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q}$$

(E. Oja, P. Oja "Funktsionaalanalüüs" lk. 11).
Seega iga $\omega \in \Omega$ korral

$$\frac{|X(\omega)Y(\omega)|}{\left(E|X|^p\right)^{\frac{1}{p}} \left(E|Y|^q\right)^{\frac{1}{q}}} = \frac{|X(\omega)|}{\left(E|X|^p\right)^{\frac{1}{p}}} \frac{|Y(\omega)|}{\left(E|Y|^q\right)^{\frac{1}{q}}} \leq \frac{|X(\omega)|^p}{\left(E|X|^p\right)p} + \frac{|Y(\omega)|^q}{\left(E|Y|^q\right)q}.$$

Keskväärtuse monotoonsus:

$$\frac{E|XY|}{\left(E|X|^p\right)^{\frac{1}{p}} \left(E|Y|^q\right)^{\frac{1}{q}}} \leq \frac{E|X|^p}{p(E|X|^p)} + \frac{E|Y|^q}{q(E|Y|^q)} = \frac{1}{p} + \frac{1}{q} = 1.$$

■

Järeldus 8.3 (Cauchy-Bunjakovski-Schwarzi võrratus)
Olgu $X \in \mathcal{L}_2(\Omega, \mathcal{F}, \mathbf{P})$ *ja* $Y \in \mathcal{L}_2(\Omega, \mathcal{F}, \mathbf{P})$. *Siis* $XY \in \mathcal{L}_1(\Omega, \mathcal{F}, \mathbf{P})$, *kusjuures*

$$E\,XY \leq E|XY| \leq \sqrt{E|X|^2}\sqrt{E|Y|^2} < \infty$$

Teoreem 8.20 (Tšebõšev-Cantelli võrratus)
Olgu $c > 0$, $X \in \mathcal{L}_2(\Omega, \mathcal{F}, \mathbf{P})$. *Siis*

$$\mathbf{P}(X - EX \geq c) \leq \frac{DX}{c^2 + DX}.$$

Tõestus. Üldisust kitsendamata eeldame, et $EX = 0$ (veendu, et üldisust pole kitsendatud). Siis

$$c = E(c - X) \leq EI_{\{X<c\}}(c - X).$$

Cauchy-Bunjakovski-Schwartz:

$$c \leq E\big(I_{\{X<c\}}(c - X)\big) \leq \sqrt{E(I_{\{X<c\}})^2}\sqrt{E(c - X)^2},$$

millest

$$c^2 \leq E(c - X)^2 EI^2_{\{X<c\}} = E(c - X)^2\mathbf{P}(X < c) = (DX + c^2)\mathbf{P}(X < c).$$

Seega

$$1 - \mathbf{P}(X < c) \leq 1 - \frac{c^2}{c^2 + DX}.$$

■

Teoreem 8.21 (Ljapunovi võrratus)
Kui $0 < p < r$ *ja* $X \in \mathcal{L}_r(\Omega, \mathcal{F}, \mathbf{P})$, *siis* $X \in \mathcal{L}_p(\Omega, \mathcal{F}, \mathbf{P})$ *ja*

$$\big(E|X|^p\big)^{\frac{1}{p}} \leq \big(E|X|^r\big)^{\frac{1}{r}} < \infty.$$

Tõestus. Piisab, kui näitame, et

$$\big(E|X|^p\big)^{\frac{r}{p}} \leq E|X|^r < \infty. \tag{8.6}$$

Kui (8.6) kehtib, siis $E|X|^p < \infty$ ja võrratus kehtib.
Defineerime

$$Y_n(\omega) := (\min\{|X(\omega)|, n\})^p.$$

Seega $E|Y_n| \leq n^p < \infty$. Vaatleme kumerat funktsiooni $g(x) = |x|^{\frac{r}{p}}$. Seega $g(Y_n(\omega)) = Y_n(\omega)^{\frac{r}{p}} \leq n^r$ millest $Eg(Y_n) \leq n^r < \infty$.
Rakendame Jenseni võrratust:

$$(EY_n)^{\frac{r}{p}} = g(EY_n) \leq Eg(Y_n) = E(Y_n)^{\frac{r}{p}} = E(\min\{|X|, n\}^r) \leq E|X|^r.$$

Et $Y_n \nearrow |X|^p$, siis MON: $EY_n \nearrow E|X|^p$. Järelikult $(EY_n)^{\frac{r}{p}} \nearrow (E|X|^p)^{\frac{r}{p}} \leq E|X|^r$. ■

Teoreem 8.22 (Minkowski võrratus)
Kui $X, Y \in \mathcal{L}_p(\Omega, \mathcal{F}, \mathbf{P})$, $p \geq 1$, *siis* $X + Y \in \mathcal{L}_p(\Omega, \mathcal{F}, \mathbf{P})$, *kusjuures*

$$\big(E|X + Y|^p\big)^{\frac{1}{p}} \leq \big(E|X|^p\big)^{\frac{1}{p}} + \big(E|Y|^p\big)^{\frac{1}{p}} < \infty.$$

Tõestus. Kõigepealt paneme tähele, et suvaliste reaalarvude x ja y korral

$$|x + y|^p \leq \big(|x| + |y|\big)^p \leq \big(2(|x| \vee |y|)\big)^p \leq 2^p(|x|^p \vee |y|^p) \leq 2^p(|x|^p + |y|^p),$$

millest saame, et, kui $X, Y \in \mathcal{L}_p(\Omega, \mathcal{F}, \mathbf{P})$, siis $E|X+Y|^p \leq 2^p E(|X|^p + |Y|^p) < \infty$ ehk $X + Y \in \mathcal{L}_p(\Omega, \mathcal{F}, \mathbf{P})$.
Olgu $p^{-1} + q^{-1} = 1$, s.t. $(p-1)q = p$. Olgu

$$A := \Big(E|X+Y|^p\Big)^{\frac{1}{q}}.$$

Kasutame Hölderi võrratust:

$$\begin{aligned}
E|X+Y|^p \leq & E(|X||X+Y|^{p-1} + |Y||X+Y|^{p-1}) \\
\leq & \big(E|X|^p\big)^{\frac{1}{p}} \big(E|X+Y|^{(p-1)q}\big)^{\frac{1}{q}} + \big(E|Y|^p\big)^{\frac{1}{p}} \big(E|X+Y|^{(p-1)q}\big)^{\frac{1}{q}} \\
= & \big[(E|X|^p)^{\frac{1}{p}} + (E|Y|^p)^{\frac{1}{p}}\big] A.
\end{aligned}$$

Kui $A = 0$, siis $X + Y = 0$ p.k. (miks?) ja võrratus kehtib triviaalselt. Seega eeldame, et $A > 0$, millest

$$\Big(E|X+Y|^p\Big)^{\frac{1}{p}} = \frac{E|X+Y|^p}{A} \leq (E|X|^p)^{\frac{1}{p}} + (E|Y|^p)^{\frac{1}{p}}.$$

■

Märkus Minkowski võrratusest ja integraali omadustest järeldub, et

$$\|X\|_p := (E|X|^p)^{\frac{1}{p}}$$

on poolnorm ruumis $\mathcal{L}_p(\Omega, \mathcal{F}, \mathbf{P})$. Faktoriseerides selle ruumi ekvivalentsisieose p.k. järgi (s.t. peaaegu kindlasti võrdsed juhuslikud suurused on üks ja sama element), saame normeeritud ruumi, mida harilikult tähistatakse $L_p(\Omega, \mathcal{F}, \mathbf{P})$. Saab näidata, et $L_p(\Omega, \mathcal{F}, \mathbf{P})$ on täielik ruum, seega on ta Banachi ruum.

8.7 Kontsentratsioonivõrratused

Olgu $X_1, \ldots, X_n$ sõltumatud juhuslikud suurused, $X_i \in \mathcal{L}_2(\Omega, \mathcal{F}, \mathbf{P})$, $D(X_i) = \sigma_i^2$,

$$S_n := \sum_{i=1}^{n} X_i.$$

Tŝebõŝevi võrratusest järeldub, et iga $c > 0$ korral

$$\mathbf{P}\big(|S_n - ES_n| > c\big) \leq \frac{\sum_{i=1}^n \sigma_i^2}{c^2} \tag{8.7}$$

(ülesanne 23).

8.7.1 Kolmogorovi võrratus

Järgnev teoreem täpsustab võrratust (8.7).

Teoreem 8.23 (Kolmogorovi võrratus)
Olgu $X_1, \dots, X_n$ sõltumatud juhuslikud suurused, $EX_i = 0$, $DX_i = \sigma_i^2 < \infty$, $i = 1, \dots, n$. Defineerime $S_k := X_1 + \ldots + X_k$. Kui $c > 0$, siis

$$\mathbf{P}(\max_{1\le k\le n} |S_k| \ge c) \le \frac{D(S_n)}{c^2} = \frac{\sum_i \sigma_i^2}{c^2}.$$

Tõestus. $A_k := \{|S_1| < c\} \cap \{|S_2| < c\} \cap \cdots \cap \{|S_{k-1}| < c\} \cap \{|S_k| \ge c\}$, $k = 1, \dots, n$. Hulgad A_i on lõikumatud.

$$\begin{aligned}D(S_n) = ES_n^2 = \int_\Omega S_n^2 d\mathbf{P} \ge \sum_{k=1}^n \int_{A_k} S_n^2 d\mathbf{P} = \sum_{k=1}^n \int_{A_k} (S_k + (S_n - S_k))^2 d\mathbf{P}\\ = \sum_{k=1}^n \int_{A_k} [S_k^2 + 2(S_n - S_k)S_k + (S_n - S_k)^2] d\mathbf{P} \ge \sum_{k=1}^n \int_{A_k} [S_k^2 + 2(S_n - S_k)S_k] d\mathbf{P}.\end{aligned}$$

Paneme tähele, et $A_k \in \sigma(X_1, \dots, X_k)$. Samuti $S_k \in m\sigma(X_1, \dots, X_k)$. Teisest küljest aga $S_n - S_k = X_{k+1} + \cdots + X_n$ on $\sigma(X_{k+1}, \dots, X_n)$-mõõtuv. Et $X_1, \dots, X_n$ on sõltumatud juhuslikud suurused, siis $\sigma(X_1, \dots, X_k)$ ja $\sigma(X_{k+1}, \dots, X_n)$ on sõltumatud σ-algebrad (vaata Kolmogorovi 0-1 seaduse tõestus). Seega teoreemist 8.16 järeldub

$$\int_{A_k} (S_n - S_k) S_k d\mathbf{P} = \int_{A_k} S_k d\mathbf{P} \int (S_n - S_k) d\mathbf{P} = 0$$

(ülesanne 7). Kokkuvõttes,

$$D(S_n) \ge \sum_{k=1}^n \int_{A_k} S_k^2 d\mathbf{P} \ge \sum_{k=1}^n c^2 \mathbf{P}(A_k) = c^2 \mathbf{P}(A),$$

kus

$$A = \{\omega : \max_{1\le k\le n} |S_k(\omega)| > c\} = \cup_{k=1}^n A_n.$$

■

8.7.2 Höffdingi võrratus

Järgnevalt üldistame võrratust (8.7):

$$\mathbf{P}\big(|S_n - ES_n| > c\big) \le \frac{\sum_{i=1}^n \sigma_i^2}{c^2}$$

teises suunas. Selleks defineerime

$$\sigma^2 := \frac{\sum_{i=1}^n \sigma_i^2}{n} \tag{8.8}$$

ja esitame(8.7) kujul

$$\mathbf{P}\Big(|\frac{S_n}{n} - \frac{ES_n}{n}| > \epsilon\Big) \leq \frac{\sigma^2}{\epsilon^2}\frac{1}{n}$$

Erijuhul, kui $EX_i = 0$ ja $DX_i = \sigma^2$ iga $i = 1, \ldots, n$ korral, saame

$$\mathbf{P}\Big(|\frac{S_n}{n}| > \epsilon\Big) \leq \frac{\sigma^2}{\epsilon^2}\frac{1}{n} \tag{8.9}$$

Seosest (8.9) tuleneb, et $\mathbf{P}\big(\big|\frac{S_n}{n}\big| > \epsilon\big) = O(n^{-1})$, kuid tihti jääb sellest väheks. Järgnevast teoreemist järeldub, et kui juhuslikud suurused X_i on *tõkestatud*, siis

$$\mathbf{P}\big(|\frac{S_n}{n}| > \epsilon\big)$$

kahaneb eksponentsiaalselt. See on väga kasulik rakendustes. Alustame abitulemusest.

Lemma 8.2 *Olgu X selline juhuslik suurus, et $EX = 0$ ja $a \leq X(\omega) \leq b$ $\forall\omega$. Siis iga $s > 0$ korral*

$$E\big(e^{sX}\big) \leq e^{\frac{s^2(b-a)^2}{8}}.$$

Tõestus. Funktsioon $x \mapsto e^{sx}$ on kumer, st iga $\lambda \in [0,1]$ korral kehtib võrratus

$$\exp[s(\lambda a + (1-\lambda)b)] \leq \lambda \exp[sa] + (1-\lambda)\exp[sb],$$

millest $\lambda = \frac{b-x}{b-a}$ korral saame $(x = \lambda a + (1-\lambda)b)$

$$e^{sx} \leq \frac{b-x}{b-a}e^{sa} + \frac{x-a}{b-a}e^{sb}, \quad a \leq x \leq b.$$

Et $EX = 0$, siis

$$E\big(e^{sX}\big) \leq \frac{b}{b-a}e^{sa} - \frac{a}{b-a}e^{sb} = (1-p+pe^{s(b-a)})e^{-ps(b-a)} =: e^{\phi(u)},$$

kus $p = -\frac{a}{b-a}$, $u = s(b-a)$ ja $(1-p = \frac{b}{b-a})$

$$\phi(u) = -pu + \ln(1-p+pe^u).$$

Nüüd

$$\phi'(u) = -p + \frac{p}{p+(1-p)e^{-u}},$$

millest $\phi'(0) = \phi(0) = 0$. Teine tuletis avaldub

$$\phi''(u) = \frac{p(1-p)e^{-u}}{(p+(1-p)e^{-u})^2} \leq \frac{1}{4}.$$

Taylori valemist saame, et mingi $\theta \in [0,u]$ korral kehtib

$$\phi(u) = \phi(0) + u\phi'(0) + \frac{u^2}{2}\phi''(\theta) \leq \frac{u^2}{8} = \frac{s^2(b-a)^2}{8}.$$

■

Teoreem 8.24 (Höffdingi võrratus)
Olgu $X_1, \dots, X_n$ sõltumatud tõkestatud juhuslikud suurused, kusjuures $a_i \le X_i \le b_i$ p.k. Siis iga $c > 0$ korral kehtivad võrratused

$$\mathbf{P}\big(S_n - ES_n \ge c\big) \le \exp\Big[-\frac{2c^2}{\sum_{i=1}^n (b_i - a_i)^2}\Big] \tag{8.10}$$

$$\mathbf{P}\big(S_n - ES_n \le -c\big) \le \exp\Big[-\frac{2c^2}{\sum_{i=1}^n (b_i - a_i)^2}\Big]. \tag{8.11}$$

Seega

$$\mathbf{P}\big(|S_n - ES_n| \ge c\big) \le 2\exp\Big[-\frac{2c^2}{\sum_{i=1}^n (b_i - a_i)^2}\Big].$$

Tõestus. Tõestus põhineb nn. *Tŝernovi meetodil.* Olgu $s > 0$. Markovi võrratusest tuleneb, et iga juhusliku suuruse X korral

$$\mathbf{P}(X \ge c) = \mathbf{P}\Big(e^{sX} \ge e^{sc}\Big) \le \frac{Ee^{sX}}{e^{sc}}.$$

Tŝernovi meetodi idee on leida selline s, et võrratuse parem pool oleks võimalikult väike. Rakendame ülaltoodud võrratust juhuslikule suurusele $S_n - ES_n$ ja saame

$$\begin{aligned}
\mathbf{P}\Big(S_n - ES_n \ge c\Big) &\le e^{-sc} E\left(\exp\Big[s\sum_i^n (X_i - EX_i)\Big]\right) \\
&= e^{-sc} E\left(\prod_{i=1}^n \exp\Big[s(X_i - EX_i)\Big]\right) \\
&= e^{-sc} \prod_{i=1}^n E\Big(e^{s(X_i - EX_i)}\Big) \\
&\le e^{-sc} \prod_{i=1}^n e^{\frac{s^2(b_i-a_i)^2}{8}} \\
&= e^{-sc} e^{\frac{s^2\sum_i (b_i-a_i)^2}{8}} \\
&= \exp\big[-sc + s^2 \frac{\sum_i (b_i - a_i)^2}{8}\big].
\end{aligned}$$

Teine võrdus tuleneb sõltumatusest, teine võrratus lemmast 8.2. Võttes nüüd

$$s = \frac{4c}{\sum (b_i - a_i)^2},$$

saame

$$\mathbf{P}\Big(S_n - ES_n \ge c\Big) \le \exp\big[-\frac{2c^2}{\sum_i (b_i - a_i)^2}\big].$$

Teine võrratus järeldub esimesest võrratusest, kui X_i rolli võtta $-X_i$ (ülesanne 25). ∎

Järeldus 8.4 *Olgu $X_1, \ldots, X_n$ sõltumatud juhuslikud suurused, kusjuures $-a \leq X_i \leq a$ ja $EX_i = 0$. Siis*

$$\mathbf{P}\Big(\Big|\frac{S_n}{n}\Big| > \epsilon\Big) \leq 2\exp\big[-n\frac{\epsilon^2}{2a^2}\big].$$

Paneme tähele, et Höffdingi võrratuses ei kasutata juhuslike suuruste dispersioone. Intuitiivselt on aga selge, et kui S_n dispersioon on väike, siis peaks seda olema ka $\mathbf{P}\big(|\frac{S_n}{n}| > \epsilon\big)$. Järgnev võrratus annab Höffdingi võrratusest parema tõkke juhul, kui S_n dispersioon on väike.

Teoreem 8.25 (Bernsteini võrratus)
Olgu $X_1, \ldots, X_n$ sõltumatud juhuslikud suurused, kusjuures $X_i \leq a$ p.k. ja $EX_i = 0$. Olgu σ^2 defineeritud seosega (8.8). Siis

$$\mathbf{P}\Big(\frac{S_n}{n} > \epsilon\Big) \leq \exp\big[-n\frac{\epsilon^2}{2\sigma^2 + \frac{2}{3}a\epsilon}\big]. \tag{8.12}$$

Bernsteini võrratust on ka võimalik tõestada Tŝernovi meetodil.

Võrratusest (8.12) järeldub, et kui $X_i \geq -a$ p.k., siis

$$\mathbf{P}\Big(\frac{S_n}{n} < -\epsilon\Big) \leq \exp\big[-n\frac{\epsilon^2}{2\sigma^2 + \frac{2}{3}a\epsilon}\big].$$

Seega, kui $-a \leq X_i \leq a$ p.k., siis

$$\mathbf{P}\Big(|\frac{S_n}{n}| > \epsilon\Big) \leq 2\exp\big[-n\frac{\epsilon^2}{2\sigma^2 + \frac{2}{3}a\epsilon}\big].$$

Näide: Olgu $B \sim B(p, n)$ (binoomjaotusega). Siis Höffdingi võrratusest järeldub

$$\mathbf{P}\Big(\frac{B}{n} - p > \epsilon\Big) \leq \exp[-2\epsilon^2 n] \tag{8.13}$$

ning Bernsteini võrratusest järeldub

$$\mathbf{P}\Big(\frac{B}{n} - p > \epsilon\Big) \leq \exp[-\frac{\epsilon^2 n}{2(1-p)(p+\frac{\epsilon}{3})}], \tag{8.14}$$

mis väikse p ja ϵ korral võib olla parem. Tegelikult pole see veel kõik: kui $p \geq \frac{1}{2}$, kehtib veel paremgi võrratus (ja ka seda saab tõestada Tŝernovi meetodil):

$$\mathbf{P}\Big(\frac{B}{n} - p > \epsilon\Big) \leq \exp[-\frac{\epsilon^2 n}{2p(1-p)}], \quad \text{kui } p \geq \frac{1}{2}. \tag{8.15}$$

Märgime veel, et kõige parema võrratuse annab nn *suurte hälvete printsiip:*

$$\mathbf{P}\Big(\frac{B}{n} - p > \epsilon\Big) \leq \exp\Big[-n\Big((p+\epsilon)\ln\frac{p+\epsilon}{p} + (1-(p+\epsilon))\ln\frac{1-(p+\epsilon)}{1-p}\Big)\Big]. \tag{8.16}$$

Võrratus (8.16) on kõige täpsem, seda parandada ei saa.

8.7.3 McDiarmidi võrratus

Selgub, et Höffdingi võrratust saab üldistada ka selles suunas, et summa on asendatud mingi teise funktsiooniga, mis rahuldab tingimust: i-nda argumendi muutmine (määramispiirkonna raames) ei tohi muuta funktiooni väärtust rohkem kui konstandi võrra. Sisuliselt tähendab see tingimus, et üksikul argumendil pole liiga suurt mõju.

Teoreem 8.26 (McDiarmidi võrratus)
Olgu $X_1, \ldots, X_n$ sõltumatud juhuslikud suurused, kusjuures juhuslik suurus X_i võtab väärtusi hulgas A_i. Oletame, et funktsioon $f : A_1 \times \cdots \times A_n \mapsto \mathbb{R}$ rahuldab tingimust

$$\sup_{x_1,\ldots,x_n,x_i' \in A_i} |f(x_1, \ldots, x_n) - f(x_1, \ldots, x_{i-1}, x_i', x_{i+1}, \ldots, x_n)| \leq \Delta_i, \quad i = 1, \ldots, n.$$

Siis iga $c > 0$ korral

$$\mathbf{P}\Big(|f(X_1, \ldots, X_n) - Ef(X_1, \ldots, X_n)| \geq c\Big) \leq 2 \exp\big[- \frac{2c^2}{\sum_{i=1}^n \Delta_i^2}\big].$$

Kui võtta $f(x_1, \ldots, x_n) = \sum_{i=1}^n x_i$, siis McDiarmidi võrratusest järeldub Höffdingi võrratus (ülesanne 28).

Näited:

1. (Kohvrite pakkimine (*bin packing*)) Olgu $x_1, \ldots, x_n$ arvud vahemikust $[0, 1]$ (st $x_i \in [0, 1]$). Olgu k minimaalne täisarv nii, et arvud $x_1, \ldots, x_n$ saab jagada k gruppi ja igas grupis olevate arvude summa pole suurem kui 1. Arve $x_1, \ldots, x_n$ võib käsitleda pakitavate asjade ruumaladena ja k on minimaalne pakkimiseks vajalik kohvrite arv kui iga kohvri ruumala on 1. Olgu nüüd asjade ruumalad juhuslikud, st $X_1, \ldots, X_n$ on $[0, 1]$-väärtuselised sõltumatud (kuid mitte ilmtingimata sama jaotusega) juhuslikud suurused. Olgu K_n minimaalne kohvrite arv. On selge, et K_n on $X_1, \ldots, X_n$ funktsioon ja seega juhuslik suurus. Samuti on selge, et ühte argumenti (ühe asja ruumala) muutes muutub kohvrite arv maksimaalselt ühe võrra. Seega $\Delta_i = 1$ ja McDiarmidi võrratusest saame kohe: kui $\epsilon > 0$, siis
$$\mathbf{P}(|K_n - EK_n| \geq \epsilon n) \leq 2 \exp[-2\epsilon^2 n].$$

2. (Pikim ühisjada (*longest common subsequence*)) Olgu $\mathcal{A}$ diskreetne hulk – tähestik – ja $x_1, \ldots, x_n$, $y_1, \ldots, y_n$ kaks teksti, st $x_i, y_i \in \mathcal{A}$. Pikim ühisjada on suurim täisarv k, mis rahuldab tingimust: leiduvad indeksid $i_1 < i_2 < \ldots < i_k$ ja $j_1 < j_2 < \ldots < j_k$ nii, et $x_{i_l} = y_{i_l}$, $l = 1, \ldots, k$. On selge, et muutes tekstis $x_1, \ldots, x_n$ ühte tähte, muutub ühisjada pikkus maksimaalsest ühe võrra. Sama kehtib arusaadavalt teise teksti kohta. Olgu $L(x_1, \ldots, x_n; y_1, \ldots, y_n)$ pikima ühisjada pikkus. Oletades nüüd, et tähed $X_1, \ldots, X_n, Y_1, \ldots, Y_n$ on sõltumatud juhuslikud suurused, saame, et ka pikima ühisjada pikkus, olgu see $L_n := L(X_1, \ldots, X_n; Y_1, \ldots, Y_n)$, on juhuslik suurus. Pane tähele, et L_n on $2n$ sõltumatu juhusliku suuruse funktsioon. McDiarmidi võrratusest saame aga, et kui $c > 0$, siis
$$\mathbf{P}(|L_n - EL_n| \geq c) \leq 2 \exp[-\frac{c^2}{n}]. \tag{8.17}$$

Kontsentratsioonivõrratuste abil saame hinnata ülalt juhuslike suurusete tsentraalseid momente isegi siis, kui keskväärtus pole teada. Näiteks võrratusest (8.17) saame iga $r > 0$ korral

$$\mathbf{P}(|L_n - EL_n|^r \geq c^r) \leq 2\exp[-\frac{c^2}{n}],$$

millest ($c^r = x$)

$$\mathbf{P}(|L_n - EL_n|^r \geq x) \leq 2\exp[-\frac{x^{\frac{2}{r}}}{n}].$$

Lemma keskväärtusest annab nüüd hinnangu

$$E|L_n - EL_n|^r = \int_0^\infty \mathbf{P}(|L_n - EL_n|^r \geq x)dx \leq 2\int_0^\infty \exp[-\frac{x^{\frac{2}{r}}}{n}]dx.$$

Muutuja vahetus $x^{\frac{2}{r}}n^{-1} = u$ annab

$$2\int_0^\infty \exp[-\frac{x^{\frac{2}{r}}}{n}]dx = r\ n^{\frac{r}{2}}\int_0^\infty e^{-u}u^{\frac{r}{2}-1}du = rn^{\frac{r}{2}}\Gamma\Big(\frac{r}{2}\Big).$$

Võttes $r = 2$, saame $DL_n \leq 2n$.

8.7.4 Kontsentratsioonivõrratused ja vahemikhinnangud

Kontsentratsioonivõrratusi saab kasutada mitteasümptootiliste vahemikhinnangute leidmiseks jaotuse keskväärtusele. Tõepoolest, olgu $X_1, \ldots, X_n$ sõltumatud ja sama jaotusega juhuslikud suurused (juhuslik valim), soovime leida kahepoolset usaldusintervalli keskväärtusele $EX_i = \mu$. Olgu meil nüüd kontsentratsioonivõrratus kujul

$$\mathbf{P}\Big(|\frac{S_n}{n} - \mu| \geq \epsilon\Big) \leq \delta(\epsilon, n), \tag{8.18}$$

kus võrratuse parem pool $\delta(\epsilon, n)$ ei sõltu keskväärtusest μ ja on iga n ja ϵ korral leitav. Olgu $1 - \alpha$ etteantud usaldusnivoo. Võttes $\epsilon(\alpha, n)$ nii, et

$$\delta(\epsilon(\alpha, n), n) = \alpha,$$

saame usaldusintervalli kujul

$$\frac{S_n}{n} \pm \epsilon(\alpha, n),$$

sest (8.18) on nüüd

$$\mathbf{P}\Big(\frac{S_n}{n} - \epsilon(\alpha, n) \leq \mu \leq \frac{S_n}{n} + \epsilon(\alpha, n)\Big) \geq 1 - \delta(\epsilon(\alpha, n), n) = 1 - \alpha. \tag{8.19}$$

Juhul, kui nende juhuslike suuruste dispersioon on teada või ülalt hinnatav, annab lihtsa kuid väga jämeda hinnangu Tšebõševi võrratus. Tõepoolest, olgu $D(X_i) = \sigma^2$. Siis (8.18) on

$$\mathbf{P}\Big(|\frac{S_n}{n} - \mu| \geq \epsilon\Big) \leq \frac{\sigma^2}{n\epsilon^2}, \tag{8.20}$$

millest

$$\epsilon(\alpha, n) = \sqrt{\frac{\sigma^2}{\alpha n}}. \tag{8.21}$$

Kui aga on teada, et juhuslikud suurused X_i on tõkestatud, st leiduvad $-\infty < a < b < \infty$ such that $a \leq X_i \leq b$, siis on mõttekas kasutada Höffdingi võrratust. Sellisel juhul (8.18) on

$$\mathbf{P}\Big(|\frac{S_n}{n} - \mu| \geq \epsilon\Big) \leq 2\exp[-\frac{2n\epsilon^2}{(b-a)^2}],$$

millest

$$\epsilon(\alpha, n) = (b-a)\sqrt{\frac{\ln\frac{2}{\alpha}}{2n}}. \tag{8.22}$$

Näide: Olgu valim Bernoulli jaotusega, st $X_i \sim B(1, p)$ ja hinnatav keskväärtus seega p. Sellisel juhul $\sigma^2 = p(1-p) \leq \frac{1}{4}$ ning seega (8.20) on

$$\mathbf{P}\Big(\Big|\frac{S_n}{n} - \mu| \geq \epsilon\Big) \leq \frac{\sigma^2}{n\epsilon^2} \leq (n\epsilon^2 4)^{-1},$$

millest (8.21) on

$$\epsilon(\alpha, n) = \sqrt{\frac{1}{4\alpha n}}.$$

Bernoulli jaotusega juhuslikud suurused on tõkestatud, st $0 \leq X_i \leq 1$, mistõttu võime kasutada Höffdingi võrratust võttes $a = 0$ ja $b = 1$. Seega (8.22) on

$$\epsilon(\alpha, n) = \sqrt{\frac{\ln\frac{2}{\alpha}}{2n}}.$$

Paneme tähele, et saadud vahemikhinnangud on *mitteasümptootilised*, st kehtvad iga n korral. See on sellepärast, et seos (8.19) kehtib ka kuitahes võikese n korral. Tõsi, liiga väikese n korral ei pruugi hinnangul olla sügavat mõtet, st vahemik on liiga lai, kuid ta on alati korrektne. Asümptootilised vahemikhinnangud kehtivad alati vaid "piisavalt suure" n korral. Lihtne, normaaljaotuse lähendil põhinev asümptootiline vahemikhinnang Bernoulli jaotuse parameetrile on kujul

$$\frac{S_n}{n} \pm \lambda_{\frac{\alpha}{2}}\sqrt{\frac{\sigma^2}{n}} \leq \frac{S_n}{n} \pm \lambda_{\frac{\alpha}{2}}\sqrt{\frac{1}{4n}},$$

kus $\lambda_{\frac{\alpha}{2}}$ on normaaljaotuse $\frac{\alpha}{2}$-kvantiil ning oleme jälle kasutanud dispersiooni hinnangut $\sigma^2 \leq \frac{1}{4}$. Võrdleme saadud hinnanguid olulisuse nivool 0.95, st $\alpha = 0.05$. Siis $\lambda_{\frac{\alpha}{2}} = 1.96$ ja $\ln\frac{2}{0.05} \approx 3.68$ mistõttu asümptootiline, Höffdingi võrratusel ja Tŝebõŝevi võrratusel põhinevad vahemikhinnangud on vastavalt

$$\frac{S_n}{n} \pm 0.98\sqrt{\frac{1}{n}}, \quad \frac{S_n}{n} \pm 1.36\sqrt{\frac{1}{n}}, \quad \frac{S_n}{n} \pm 2.23\sqrt{\frac{1}{n}}.$$

Näeme, et asümptootiline vahemikhinnang on kõige kitsam ja arusaadavalt ei saa see kuidagi olla teisiti. Höffdingi võrratus annab ligikaudu 1.4 korda laiema vahemiku, kuid Tŝebõŝevi võrratus annab juba üle kahe korra laiema vahemiku. Samas Höffdingi (loomulikult ka Tŝebõŝevi) võrratusel põhinev vahemik kehtib iga n korral.
Loomulikult saab vahemikhinnangute konstrueerimiseks kasutada ka teisi kontsentratsioonivõrratusi. Väga tihti, eriti tehisõppes (*machine learning*) kasutatakse McDiarmidi võrratusel põhinevaid usalduspiire.

8.8 Ülesanded

1. Olgu $(\Omega, \mathcal{F}, \mathbf{P}) = ([0,1], \mathcal{B}, \mathrm{Leb})$. Leida EX, kus

1) $X(\omega) = \omega^2$

2) $X(\omega) = \omega - 0,5$.

2. Olgu $X \in \mathcal{L}_1(\Omega, \mathcal{F}, \mathbf{P})$, kusjuures $EX = a$ ja $E|X| = b$.

1) Millist tingimust peavad konstandid a ja b rahuldama, et see saaks nii olla?

2) Leida $E\max\{0, X\}$ ja $E\min\{0, X\}$.

3. Olgu X, Y sõltumatud diskreetsed juhuslikud suurused, mille väärtused on mittenegatiivsed täisarvud. Tõestada, et

1) $EX = \sum_{i=1}^{\infty} \mathbf{P}(X \geq i)$.

2) $E(\min\{X, Y\}) = \sum_{i=1}^{\infty} \mathbf{P}(X \geq i)\mathbf{P}(Y \geq i)$.

4.

1) Kuulugu $X \in \mathcal{L}_1(\Omega, \mathcal{F}, \mathbf{P})$. Domineeritud koondumise teoreemi kasutades tõestada, et

$$\lim_{t\to\infty} \int_{\{\omega||X(\omega)|>t\}} |X| d\mathbf{P} = 0$$

2) Kuulugu $X \in \mathcal{L}_p(\Omega, \mathcal{F}, \mathbf{P}) \quad (p > 0)$. Tõestada, et

$$\lim_{t\to\infty} t^p \mathbf{P}\{\omega||X(\omega)| > t\} = 0.$$

5.

1) Tõestada lemma 8.1 lihtsate juhuslike suuruste korral.

2) Monotoonse koondumise teoreemi ja tulemust 1) kasutades tõestada lemma 8.1.

3) Kuulugu $X \in \mathcal{L}_1(\Omega, \mathcal{F}, \mathbf{P})$. Lemmat 8.1 kasutades tõestada, et

$$EX = \int_{[0,\infty)} \mathbf{P}(X > t)dt - \int_{(-\infty,0]} \mathbf{P}(X < t)dt = \int_0^{\infty} (1 - F(t))dt - \int_{-\infty}^{0} F(t)dt.$$

6. Olgu $X_1, \ldots, X_n$ sõltumatud juhuslikud suurused, kusjuures $X_i \in \mathcal{L}_1(\Omega, \mathcal{F}, \mathbf{P})$. Tõestada, et $X_1 \cdots X_n \in \mathcal{L}_1(\Omega, \mathcal{F}, \mathbf{P})$ ja

$$E(X_1 \cdots X_n) = EX_1 \cdots EX_n.$$

7. Olgu X, Y lõplike keskväärtustega sõltumatud juhuslikud suurused, $A \in \sigma(X)$. Tõestada, et $\int_A XY d\mathbf{P} = \int_A X d\mathbf{P} EY$. Järeldada, et

$$\int_A Y d\mathbf{P} = \mathbf{P}(A) EY.$$

8. Olgu X selline positiivne juhuslik suurus, et $EX^p < \infty$, kus $p < 0$. Tõestada, et

$$\lim_{t \to 0} t^p F(t) = 0.$$

9. Olgu $X \in \mathcal{L}_1(\Omega, \mathcal{F}, \mathbf{P})$. Tõestada, et iga X korral

$$\max\{x, EX\} \leq E \max\{x, X\} < \infty.$$

10. Olgu $X, Y \in \mathcal{L}_1(\Omega, \mathcal{F}, \mathbf{P})$ jaotusfunktsioonidega vastavalt F ja G. Olgu $x \geq 0$. Tõestada, et

$$\int_x^\infty (1 - F(t)) dt \leq \int_x^\infty (1 - G(t)) dt$$

parajasti siis, kui $E \max\{x, X\} \leq E \max\{x, Y\}$.

11. Olgu X mittenegatiivne lihtne juhuslik suurus väärtuste hulgaga $\{a_1, \ldots, a_m\}$. Tõestada, et

1) $\lim_n \frac{E(X^{n+1})}{E(X^n)} = \max\{a_1, \ldots, a_m\}$,

2) $\lim_n \big(E(X^n)\big)^{\frac{1}{n}} = \max\{a_1, \ldots, a_m\}$.

12. Olgu X, Y sõltumatud, kusjuures $EX = 1$, $EY = 2$, $DX = 1$, $DY = 4$. Leida

1) $E(X^2 + 2Y^2 - XY - 4X + Y + 4)$.

2) $E(X + Y + 1)^2$.

13. Kuulugu $X, Y \in \mathcal{L}_2(\Omega, \mathcal{F}, \mathbf{P})$. Cauchy-Bunjakovski võrratust kasutades tõestada, et

$$(\sqrt{DX} - \sqrt{DY})^2 \leq D(X + Y) \leq (\sqrt{DX} + \sqrt{DY})^2.$$

14.

1) Olgu X selline juhuslik suurus, et $0 < X < 1$ p.k. Tõestada, et $X \in \mathcal{L}_2(\Omega, \mathcal{F}, \mathbf{P})$, näidata, et $DX < EX$.

2) Olgu X selline juhuslik suurus, et $|X| \leq c$ p.k.. Tõestada, et $DX \leq cE|X|$.

15. Olgu $X_1, \ldots, X_n$ paarikaupa sõltumatud juhuslikud suurused. Tõestada, et $D(\sum_n X_n) = \sum D(X_n)$.

16. Kuulugu $X, Y \in \mathcal{L}_2(\Omega, \mathcal{F}, \mathbf{P})$, kusjuures X ja Y on sõltumatud.

1) Tõestada, et $DXY \geq DXDY$.

2) Milliseid tingimusi peavad juhuslikud suurused X ja Y rahuldama, et kehtiks võrdus $DXY = DXDY$.

17. Olgu X selline juhuslik suurus, et $EX = 0$. Tõestada, et $E|X| \leq \frac{1}{2}(D(X) + 1)$. (Leida $E(|X| - 1)^2$).

18. Olgu $X \in \mathcal{L}_p(\Omega, \mathcal{F}, \mathbf{P})$, $p > 0$. Tõestada, et $\lim_{a\to\infty} E|X + a|^p = \infty$. (Kasuta Fatou lemmat).

19. Olgu $p \geq 2$ paarisarv, $a_1, \ldots, a_n \in \mathbb{R}$ ja $X_1, \ldots, X_n$ olgu sellised juhuslikud suurused, et $E(X_i)^2 = \sigma^2 < \infty$.

1) Jenseni võrratust kasutades tõestada, et $(\sum a_i)^p \leq n^{p-1} \sum a_i^p$.

2) Järeldada, et $E(X_1 + \cdots + X_n)^2 \leq n^2\sigma^2$.

20. Olgu $X, Y \in L_p(\Omega, \mathcal{F}, \mathbf{P})$. Tõestada, et $aX + bY \in L_p(\Omega, \mathcal{F}, \mathbf{P})$, kus $a, b \in \mathbb{R}$.

21. Tõestada, et $DX = \min_a E(X - a)^2$.

22. Tõestada Tŝebõŝevi võrratus.

23. Tõestada (8.7).

24. Seosest (8.10) järeldada (8.11).

25. Tõestada (8.13) ja (8.14).

26. Olgu P, Q reaalteljel antud tõenäosusjaotused, μ olgu samuti reaalteljel antud σ-lõplik mõõt. Olgu f ja g mõõtude P ja Q tihedused mõõdu μ suhtes, st $\frac{dP}{d\mu} = f$ ja $\frac{dQ}{d\mu} = g$. Tõestada

$$D(P\|Q) := \int \ln\Big(\frac{f(x)}{g(x)}\Big) f(x)\mu(dx) \geq 0.$$

27. McDiarmidi võrratusest järeldada Höffdingi võrratus.

28. Olgu N juhuslik suurus, nii, et $\mathbf{P}(N \in \mathbb{N}) = 1$. Olgu $X_1, X_2, \ldots$ iid juhuslikud suurused, $EX_1 = 0$ ja $DX_1 = \sigma^2$. Tõesta, et iga $m \in \mathbb{N}$ ja $\epsilon > 0, \delta > 0$ korral $(S_k = X_1 + \cdots + X_k)$

$$\mathbf{P}\Big(|S_N - S_m| > \epsilon\Big) \leq \frac{\sigma^2 \delta m}{\epsilon^2} + \mathbf{P}\big(|\frac{N}{m} - 1| > \delta\big).$$

9 Juhuslike jadade koondumised

Olgu $X \sim P, X_1 \sim P_1, X_2 \sim P_2, \ldots$ tõenäosusruumil $(\Omega, \mathcal{F}, \mathbf{P})$ antud juhuslikud suurused jaotusfunktsioonidega vastavalt $F, F_1, F_2, \ldots$.

9.1 Koondumistüübid

Def 9.1 *Juhuslike suuruste jada $X_1, X_2, \ldots$ koondub juhuslikuks suuruseks X* **peaaegu kindlasti***, kui*

$$\mathbf{P}\{\omega | X_n(\omega) \to X(\omega)\} = 1.$$

Peaaegu kindlasti koondumist tähistame: $X_n \to X$ p.k..

Def 9.2 *Juhuslike suuruste jada $X_1, X_2, \ldots$ koondub juhuslikuks suuruseks X* **tõenäosuse järgi***, kui $\forall \epsilon > 0$ korral*

$$\mathbf{P}\{\omega | \, |X_n(\omega) - X(\omega)| > \epsilon\} \to 0, \quad n \to \infty.$$

Tõenäosuse järgi koondumist tähistame: $X_n \xrightarrow{P} X$.

Def 9.3 *Juhuslike suuruste jada $X_1, X_2, \ldots$ koondub juhuslikuks suuruseks X* **p-keskmise mõttes** *$(p > 0)$, kui $E|X_n - X|^p \to 0$.*

Koondumist p-keskmise mõttes nimetatakse ka koondumiseks ruumis $\mathcal{L}_p(\Omega, \mathcal{F}, \mathbf{P})$ ja tähistatakse: $X_n \xrightarrow{p} X$. Olgu öeldud, et koondumine $E|X_n - X|^p \to 0$ ei pruugi ilmtingimata tähendada. et $X_n, X \in \mathcal{L}_p(\Omega, \mathcal{F}, \mathbf{P})$, küll aga kuulub sinna $X_n - X$, vähemalt mingist n-st alates.

Def 9.4 *Juhuslike suuruste jada $X_1, X_2, \ldots$ koondub juhuslikuks suuruseks X* **jaotuse järgi***, kui*

$$F_n(x) \to F(x)$$

iga funktsiooni F pidevuspunkti x korral.

Jaotuse järgi koondumist tähistame: $X_n \Rightarrow X$ või $X_n \xrightarrow{D} X$.

Seega $X_n \Rightarrow X$ parajasti siis, kui $P_n(-\infty, x] \to P(-\infty, x]$ iga sellise x korral, kui $P\{x\} = 0$. Erinevalt peaaegu kindlasti koondumisest, p-keskmise mõttes koondumisest ja tõenäosuse järgi koondumisest omab jaotuse järgi koondumine mõtet ka siis, kui juhuslikud suurused X_n on defineeritud erinevatel tõenäosusruumidel.

9.2 Peaaegu kindlasti koondumine

Teoreem 9.5 *Järgmised väited on samaväärsed*

1) $X_n \to X$ *p.k.*

2) *iga* $\epsilon > 0$ *korral* $\mathbf{P}(|X_n - X| > \epsilon \ i.o.) = 0$

3) *iga* $\epsilon > 0$ *korral* $\mathbf{P}(\cup_{n\geq m}\{|X_n - X| > \epsilon\}) \to 0$, *kui* $m \to \infty$

4) *iga* $\epsilon > 0$ *korral* $\mathbf{P}(\sup_{n\geq m} |X_n - X| > \epsilon) \to 0$, *kui* $m \to \infty$.

Tõestus. Defineerime hulgad

$$\begin{aligned}
A_n(\epsilon) &:= \{\omega| \, |X_n(\omega) - X(\omega)| > \epsilon\}, \\
A(\epsilon) &:= \limsup_n A_n(\epsilon) = \cap_{m=1}^{\infty} \cup_{n\geq m} A_n(\epsilon) = \{\omega| \, |X_n(\omega) - X(\omega)| > \epsilon \ i.o.\} \\
A &:= \{\omega | X_n(\omega) \to X(\omega)\}.
\end{aligned}$$

Kui $\epsilon_1 \leq \epsilon_2$, siis $A(\epsilon_1) \supset A(\epsilon_2)$, millest $\cup_{\epsilon>0} A(\epsilon) = \cup_{i=1}^{\infty} A(\frac{1}{i})$.

1)⇒ 2): $A^c = \cup_{\epsilon>0} A(\epsilon)$. Seega $\mathbf{P}(A^c) = 0 \Rightarrow \mathbf{P}(A(\epsilon)) = 0 \ \forall \epsilon > 0$.
2)⇒ 1): $A^c = \cup_{\epsilon>0} A(\epsilon) = \cup_{i=1}^{\infty} A(\frac{1}{i}) \in \mathcal{F}$. Kui $\mathbf{P}(A(\epsilon)) = 0$, $\forall \epsilon > 0$, siis $\mathbf{P}(A^c) = 0$.
2)⇔ 3): ülesanne 2.
3)⇔ 4): Pane tähele, et iga n korral

$$\cup_{n\geq m}\{|X_n - X| > \epsilon\} = \{\sup_{n\geq m} |X_n - X| > \epsilon\}.$$

■

Järeldus 9.1 *Kui* $\sum_n \mathbf{P}(|X_n - X| > \epsilon) < \infty$, $\forall \epsilon > 0$, *siis* $X_n \to X$ *p.k.*

Tõestus. BC 1 ja Teoreem 9.5. (ülesanne 4.). ■

Järeldus 9.2 *Kui* $X_n \to X$ *p.k., siis* $X_n \xrightarrow{P} X$.

Tõestus. Teoreemist 9.5 järeldub, et $\forall \epsilon > 0$ korral $\mathbf{P}(A_n(\epsilon)) \to 0$ (ülesanne 5). ■

9.2.1 Cauchy jada

Arvjada $\{x_n\}$ on koonduv parajasti siis, kui ta on Cauchy jada (fundamentaaljada):

$$\forall \epsilon > 0, \quad \exists N : |x_k - x_l| < \epsilon \quad \text{kui vaid} \quad k, l > N.$$

Olgu

$$B := \Big\{\omega | \{X_n(\omega)\} \text{ on Cauchy jada}\Big\}.$$

Teame, et $B \in \mathcal{F}$ (lemma 4.5 e)). Et $A \subset B$, siis koondumisest $X_n \to X$ p.k. järeldub, et $\mathbf{P}(B) = 1$.

Teisest küljest, kui $\mathbf{P}(B) = 1$, siis $\exists$ juhuslik suurus X nii, et $X_n \to X$ p.k. Tõepoolest, defineerime $X : \Omega \to \mathbb{R}$ järgmiselt

$$X(\omega) = \begin{cases} \lim_n X_n(\omega) & \text{kui } \omega \in B, \\ 0 & \text{mujal.} \end{cases}$$

Järelduse 4.3 põhjal X on juhuslik suurus, sest $X_n I_B(\omega) \to X(\omega)$ iga ω korral; definitsiooni kohaselt $X_n \to X$ p.k. Seega juhuslike suuruste jada $\{X_n\}$ koondub peaaegu kindlasti (mingiks juhuslikuks suuruseks X) parajasti siis, kui ta on peaaegu kindlasti Cauchy jada (s.t. $\mathbf{P}(B) = 1$).

Järgnev teoreem annab mõne piisava ja tarviliku kriteeriumi selleks, et $\mathbf{P}(B) = 1$.

Teoreem 9.6 *Järgmised väited on samaväärsed*

1) $\{X_n\}$ *on p.k. Cauchy jada*

2) *iga* $\epsilon > 0$ *korral* $\mathbf{P}(\cup_{k,l \geq m}\{|X_k - X_l| > \epsilon\}) \to 0$*, kui* $m \to \infty$

3) *iga* $\epsilon > 0$ *korral* $\mathbf{P}(\sup_{k,l \geq m} |X_k - X_l| > \epsilon) \to 0$ *, kui* $m \to \infty$

4) *iga* $\epsilon > 0$ *korral* $\mathbf{P}(\sup_{k \geq 0} |X_{m+k} - X_m| > \epsilon) \to 0$*, kui* $m \to \infty$

Tõestus. Defineerime hulgad

$$B_{kl}(\epsilon) := \{\omega | \, |X_k(\omega) - X_l(\omega)| > \epsilon\},$$
$$B(\epsilon) := \cap_{m=1}^{\infty} \cup_{k,l \geq m} B_{k,l}(\epsilon).$$

Kui $\epsilon_1 \leq \epsilon_2$, siis $B(\epsilon_1) \supset B(\epsilon_2)$, millest $\cup_{\epsilon > 0} B(\epsilon) = \cup_{i=1}^{\infty} B(\frac{1}{i})$.
Kui arvjada $\{x_n\}$ pole Cauchy jada, siis leidub $\epsilon > 0$ nii, et iga n korral leiduvad sellised $k, l \geq n$, et $|x_k - x_l| > \epsilon$. Seega $B^c = \cup_{\epsilon > 0} B(\epsilon) = \cup_{i=1}^{\infty} B(\frac{1}{i})$, millest järeldub 1) $\Leftrightarrow$ 2).
2)$\Leftrightarrow$ 3): ülesanne 6.
3)$\Leftrightarrow$ 4): Iga arvujada $\{x_n\}$ korral kehtib

$$\sup_{k \geq 0} |x_{n+k} - x_n| \leq \sup_{k,l \geq 0} |x_{n+k} - x_{n+l}| \leq 2 \sup_{k \geq 0} |x_{n+k} - x_n|.$$

Toodud võrratuste ahel tõestab, et 3)$\Leftrightarrow$ 4) (ülesanne 7). $\blacksquare$

9.3 Tõenäosuse järgi koondumine

Meeldetuletus: Olgu $\{x_n\}$ arvjada. Kehtivad väited:

1. Kui jada koondub piirväärtuseks x, siis selle jada iga alamjada koondub samaks piirväärtuseks.

2. Kui jada $\{x_n\}$ *igal* alamjadal leidub *vähemalt üks* koonduv alamjada (seega esialgse jada alam-alamjada) ja iga sellise koonduva jada piirväärtus on x, siis koondub ka esialgne jada piirväärtuseks x.

Tõestame väite 2. Oletame vastuväiteliselt, et jada $\{x_n\}$ igal alamjadal on arvuks x koonduv alamjada, kuid $x_n \not\to x$. Seega $|x_n - x| \not\to 0$. Vastavalt arvjada koondumise definitsioonile leidub $\epsilon > 0$ nii, et $|x_n - x| > \epsilon$ i.o.. Teisisõnu, leidub mingi alamjada $\{x_{n_k}\}$ nii, et $|x_{n_k} - x| > \epsilon$ iga $k = 1, 2, 3\ldots$ korral. Vastavalt eeldusele aga leidub alamjadal $\{x_{n_k}\}$ omakorda arvuks x koonduv alamjada $\{x_{n_{k_l}}\}$. See aga tähendab, et leidub l_o nii, et $|x_{n_{k_l}} - x| < \epsilon$, kui $l > l_o$. Saime vastuolu.

Võttes ülaltoodud kaks väidet kokku saame, et arvjada koondub $\{x_n\}$ piirväärtuseks x parajasti siis, kui igal $\{x_n\}$ alamjadal leidub omakorda almjada, mis koondub piirväärtuseks x.

Tõenäosuse järgi koonduva jada alamjadad. Veendume, et ülaltoodud väited kehtivad ka tõenäosuse järgi koondumisel. Kasutades teoreemi 9.5 tõestuses toodud tähistusi, saame: $X_n \stackrel{P}{\to} X$ parajasti siis, kui $\epsilon > 0$ korral $\mathbf{P}(A_n(\epsilon)) \to 0$. Jadad $\{\mathbf{P}(A_n(\epsilon))\}$ on aga arvjadad ja nende korral kehtivad ülaltoodud väited. Seega, kui $X_n \stackrel{P}{\to} X$ ja $\{X_{n_k}\}$ on mingi alamjada, siis iga $\epsilon > 0$ korral $\mathbf{P}(A_{n_k}(\epsilon)) \to 0$ ehk $X_{n_k} \stackrel{P}{\to} X$. Seega tõenäosuse järgi koonduva jada iga alamjada koondub samaks piirväärtuseks ehk kehtib väide 1.

Näitame, et kehtib ka väide 2. Olgu $\{X_n\}$ juhuslike suuruste jada, millel on järgmine omadus: leidub juhuslik suurus X nii, et mistahes alamjadal $\{X_{n_k}\}$ leidub omakorda alamjada $\{X_{n_{k_l}}\}$ nii, et $X_{n_{k_l}} \stackrel{P}{\to} X$. Veendume, et sellisel juhul ka $X_n \stackrel{P}{\to} X$. Oletades vastuväiteliselt, et $X_n \stackrel{P}{\not\to} X$, peab leiduma $\epsilon > 0$ ja $\delta > 0$ ja alamjada $\{X_{n_k}\}$ nii, et $\mathbf{P}(A_{n_k}(\epsilon)) > \delta$ iga k korral. Vastavalt eeldusele aga peab alamjadal $\{X_{n_k}\}$ leiduma omakorda alamjada $\{X_{n_{k_l}}\}$ nii, et $X_{n_{k_l}} \stackrel{P}{\to} X$. See aga tähendab, et ka $\mathbf{P}(A_{n_{k_l}}(\epsilon)) \to 0$. Vastuolu.

Kokkuvõttes: $X_n \stackrel{P}{\to} X$ parajasti siis, kui mistahes alamjadal $\{X_{n_k}\}$ leidub vähemalt üks alamjada $\{X_{n_{k_l}}\}$, mis koondub tõenäosuse järgi piirväärtuseks X.

Märkus: Ülaltoodud omadus iseloomustab koondumisi, mis on *metriseeruvad*: st leidub meetrika (kaugus) nii, et uuritav koondumine on ekvivalentne koondumisega selles meetrikas. Tõenäosuse järgi koondumine on metriseeruv, sest kõikide juhuslike suuruste hulgal on võimalik defineerida kaugus (neid on rohkem kui üks) d nii, et $X_n \stackrel{P}{\to} X$ parajasti siis, kui $d(X_n, X) \to 0$. Et aga $\{d(X_n, X)\}$ on arvjada, on selge, et väited 1 ja 2 peavad kehtima.

Järgnev lemma näitab aga, et igal tõenäosuse järgi koonduval jadal leidub vähemalt üks alamjada, mis koondub veelgi tugevamas mõttes: peaaegu kindlasti.

Lemma 9.1 *Kui $X_n \xrightarrow{P} X$, siis leidub alamjada $\{X_{n_k}\}$, $k = 1, 2, \ldots$ nii, et $X_{n_k} \to X$ p.k.*

Tõestus. Iga $\epsilon > 0$ ja $\delta > 0$ korral $\exists N$ nii, et $\mathbf{P}(A_n(\epsilon)) \leq \delta$, kui $n \geq N$. Seega leidub kasvav jada $n_1, n_2, \ldots$ nii, et $\mathbf{P}(A_{n_k}(\frac{1}{k})) \leq \frac{1}{k^2}$. Olgu $\epsilon > 0$ suvaline. Vali $k_o > \frac{1}{\epsilon}$. Nüüd

$$\sum_k \mathbf{P}(A_{n_k}(\epsilon)) = \sum_{k=1}^{k_o} \mathbf{P}(A_{n_k}(\epsilon)) + \sum_{k>k_o} \mathbf{P}(A_{n_k}(\epsilon)) \leq k_o + \sum_{k>k_o} \mathbf{P}(A_{n_k}(\frac{1}{k})) < \infty,$$

sest $A_{n_k}(\frac{1}{k}) \supset A_{n_k}(\epsilon)$, kui $k > k_o$ ja $\sum_k \frac{1}{k^2} < \infty$. Järelduse 9.1 tõttu nüüd $X_{n_k} \to X$, p.k. kui $k \to \infty$. ■

Järeldus 9.3 *$X_n \xrightarrow{P} X$ parajasti siis, kui iga alamjada $\{X_{n_k}\}$ sisaldab omakorda alamjada $\{X_{n_{k_l}}\}$ nii, et $X_{n_{k_l}} \to X$ p.k. $(l \to \infty)$.*

Tõestus. Ülesanne 11. ■

Märkus: Järeldus 9.3 on mõnevõrra üllatav, sest analoogia põhjal võiks oletada järgmist: kui leidub X nii, et juhuslike suuruste jada $\{X_n\}$ igal alamjadal leidub vähemalt üks piirväärtuseks X *peaaegu kindlasti* koonduv alamjada, siis ka esialgne jada $\{X_n\}$ koondub piirväärtuseks X *peaaegu kindlasti*. Paraku pole see nii: esialgne jada küll koondub piirväärtuseks X kuid mitte ilmtingimata peaaegu kindlasti vaid nõrgemas mõttes – tõenäosuse järgi. Sellest järeldub, et peaaegu kindlasti koondumise pole metriseeruv, st kõikide juhuslike suuruste hulgal ei saa leida meetrikat nii, et peaaegu kindlasti koondumine oleks samaväärne koondumisega selles meetrikas.

Järelduse 9.3 abil saab lihtsalt tõestada, et tõenäosuse järgi koonduvate juhuslike suuruste pidevad funktsioonid koonduvad ka tõenäosuse järgi:

Järeldus 9.4 *Olgu $X_n \xrightarrow{P} X$ ja $g : \mathbb{R} \to \mathbb{R}$ pidev funktsioon. Siis $g(X_n) \xrightarrow{P} g(X)$.*

Tõestus. Ülesanne 27. ■

9.4 Koondumisliikide omavahelised seosed

Lemma 9.2 *Kui $X_n \xrightarrow{P} X$, siis $X_n \Rightarrow X$.*

Tõestus. Olgu x funktsiooni F pidevuspunkt. Et iga $\epsilon > 0$ korral kehtivad võrdused

$$\{X_n \leq x\} = \{X_n \leq x, X \leq x + \epsilon\} \cup \{X_n \leq x, X > x + \epsilon\}$$

ja

$$\{X \leq x - \epsilon\} = \{X \leq x - \epsilon, X_n \leq x\} \cup \{X \leq x - \epsilon, X_n > x\},$$

siis

$$F_n(x) = \mathbf{P}(X_n \leq x, X \leq x + \epsilon) + \mathbf{P}(X_n \leq x, X > x + \epsilon) \leq F(x + \epsilon) + \mathbf{P}(A_n(\epsilon))$$
$$F(x - \epsilon) = \mathbf{P}(X \leq x - \epsilon, X_n \leq x) + \mathbf{P}(X \leq x - \epsilon, X_n > x) \leq F_n(x) + \mathbf{P}(A_n(\epsilon)).$$

Et $\mathbf{P}(A_n(\epsilon)) \to 0$, siis $\limsup_n F_n(x) \leq F(x + \epsilon)$, $\liminf_n F_n(x) \geq F(x - \epsilon)$. Et x on F pidevuspunkt, siis protsessis $\epsilon \searrow 0$ saame

$$F(x) = \liminf_n F_n(x) = \limsup_n F_n(x) = \lim_n F_n(x).$$

■

Teoreem 9.7 *Kehtivad implikatsioonid*

- $(X_n \to X p.k.) \Rightarrow (X_n \xrightarrow{P} X) \Rightarrow (X_n \Rightarrow X)$
- $(X_n \xrightarrow{p} X) \Rightarrow (X_n \xrightarrow{P} X)$
- $(X_n \xrightarrow{p} X) \Rightarrow (X_n \xrightarrow{r} X)$, kui $p \geq r > 0$.

Tõestus. Järeldus 9.2, Markovi võrratus, Lemma 9.2, Ljapunovi võrratus (ülesanne 12).
■

9.4.1 Kontranäited

Üldiselt ükski teine implikatsioon ei kehti.

- Olgu $(\Omega, \mathcal{F}, \mathbf{P}) = ([0,1], \mathcal{B}, Leb)$. Olgu $A_n^i := [\frac{i-1}{n}, \frac{i}{n}]$, $X_n^i := I_{A_n^i}$. Vaatleme jada

$$\{X_1^1, X_2^1, X_2^2, X_3^1, X_3^2, X_3^3, \ldots\} \tag{9.1}$$

  Et $E|X_n^i|^p = \mathbf{P}(|X_n^i| > \epsilon) = \mathbf{P}(A_n^i) = n^{-1}$, kui $\epsilon < 1$, $p > 0$ siis jada (9.1) koondub p-keskmise mõttes ja seega ka tõenäosuse järgi konstantseks juhuslikuks suuruseks 0. On aga selge, et jada (9.1) ei koondu mitte ühegi argumendi $\omega \in [0,1]$ korral. Seega (9.1) ei koondu peaaegu kindlasti.

  Toodud näide tõestab, et tõenäosuse järgi koondumisest ei järeldu peaaegu kindlasti koondumine, samuti ei järeldu peaaegu kindlasti koondumine p-keskmise mõttes koondumisest.

- Olgu $(\Omega, \mathcal{F}, \mathbf{P}) = ([0,1], \mathcal{B}, Leb)$, $X_n := e^n I_{[0,n^{-1}]}$. On lihtne veenduda, et $X_n \to 0$ p.k., millest järeldub, et $X_n \xrightarrow{P} 0$ ja $X_n \Rightarrow 0$. Teisest küljest aga ei kehti $X_n \xrightarrow{p} 0$ suvalise $p > 0$ korral (ülesanne 19).

- Olgu juhuslikud suurused X_n jaotusega $\mathbf{P}(X_n = 0) = 1 - n^{\frac{-(r+p)}{2}}$, $\mathbf{P}(X_n = n) = n^{\frac{-(r+p)}{2}}$, kusjuures $(p > r > 0)$. Kehtib $X_n \xrightarrow{r} 0$, kuid mitte $X_n \xrightarrow{p} 0$ (ülesanne 15). Seega jada koondub r-keskmise mõttes, kuid mitte p-keskmise mõttes $(p > r)$.

- Lõpuks veendume, et jaotuse järgi koondumisest ei järeldu tõenäosuse järgi koondumine.
  Olgu $X \sim N(0,1)$, $X_n := -X$, $n = 1, 2, \ldots$. On selge, et $X_n \Rightarrow X$, kuid
  $$\mathbf{P}(|X_n - X| > \epsilon) = \mathbf{P}(|X| > \frac{\epsilon}{2}) > 0.$$
  Seega koondumine $X_n \xrightarrow{P} X$ ei kehti.

9.4.2 Erandid

Erijuhtudel on erinevad koondumisliigid ekvivalentsed.
Kui piirjaotus on p.k. konstantne, siis tõenäosuse järgi koondumine on ekvivalentne jaotuse järgi koondumisega.

Lemma 9.3 *Kui $X_n \Rightarrow c$, kus c on konstant, siis $X_n \xrightarrow{P} c$.*

Tõestus. Ülesanne 16. ■

Järgmine lemma üldistab domineeritud koondumise teoreemi ja annab tarviliku tingimuse, et koondumine p-keskmise mõttes oleks ekvivalentne koondumisega tõenäosuse järgi. Lemma tõestuseks vajame abitulemust.

Väide 9.1 *Kuulugu $X \in \mathcal{L}_1(\Omega, \mathcal{F}, \mathbf{P})$, $A_n \in \mathcal{F}$. Kui $\mathbf{P}(A_n) \to 0$, siis $E(|X| I_{A_n}) \to 0$.*

Tõestus. Vastuväiteliselt oletades saame, et $\exists \epsilon > 0$ ja alamjada $\{A_{n_k}\}$ nii, et $\mathbf{P}(A_{n_k}) < 2^{-k}$, kuid $E(|X| I_{A_{n_k}}) > \epsilon$. Olgu $A := \limsup_k A_{n_k}$. BC 1 põhjal $\mathbf{P}(A) = 0$. Rakendades aga jadale $\{|X| I_{A_{n_k}}\}$ pööratud Fatou lemmat (milline on domineeriv funktsioon?), saame

$$E(\limsup_k |X| I_{A_{n_k}}) = E(|X| I_A) \geq \limsup_k E(|X| I_{A_{n_k}}) \geq \epsilon > 0.$$

See on vastuolus integraali omadustega. ■

Teoreem 9.8 *Kui $X_n \xrightarrow{P} X$ ja leidub $Y \in \mathcal{L}_p(\Omega, \mathcal{F}, \mathbf{P})$ $p > 0$ nii, et $|X_n| \leq Y$ p.k., $n = 1, 2, \ldots$, siis $X_n \xrightarrow{p} X$.*

Tõestus. Lemma 9.1 põhjal leidub alamjada $X_{n_k} \to X$ p.k. Et iga $k = 1, 2, \ldots$ korral $|X_{n_k}| \le Y$ p.k., siis $|X| \le Y$ p.k. Olgu $\epsilon > 0$. Nüüd $|X_n - X|^p \le (|X_n| + |X|)^p \le 2^p Y^p$ p.k., millest

$$E|X_n - X|^p = E(|X_n - X|^p I_{A_n(\epsilon)}) + E(|X_n - X|^p I_{A_n^c(\epsilon)}) \le E(2^p Y^p I_{A_n(\epsilon)}) + \epsilon^p.$$

Et $\mathbf{P}(A_n(\epsilon)) \to 0$, siis väite 9.1 põhjal $E(2^p Y^p I_{A_n(\epsilon)}) \to 0$, kui $n \to \infty$. Et ϵ oli suvaline, siis $E|X_n - X|^p \to 0$. ■

Järeldus 9.5 *Kui $X_n \xrightarrow{P} X$ ja leidub $K < \infty$ nii, et $\mathbf{P}(|X_n| \le K) = 1$, $n = 1, 2, \ldots$, siis iga $p > 0$ korral $X_n \xrightarrow{p} X$.*

9.5 Koondumiste omadusi

Olgu $x_n \to x$ ja $y_n \to y$ koonduvad arvjadad, $a \in \mathbb{R}$. Siis $x_n + y_n \to x + y$, $x_n y_n \to xy$ ja $ax_n \to ax$. Seda arvestades on võimalik tõestada järgmine teoreem.

Teoreem 9.9 *Olgu $a \in \mathbb{R}$. Kehtivad väited*

1) *kui $X_n \to X$ p.k. ja $Y_n \to Y$ p.k., siis $X_n + Y_n \to X + Y$ p.k., $X_n Y_n \to XY$ p.k. ja $aX_n \to aX$ p.k.;*

2) *kui $X_n \xrightarrow{P} X$ ja $Y_n \xrightarrow{P} Y$, siis $X_n + Y_n \xrightarrow{P} X + Y$, $X_n Y_n \xrightarrow{P} XY$ ja $aX_n \xrightarrow{P} aX$;*

3) *kui $X_n \xrightarrow{p} X$ ja $Y_n \xrightarrow{p} Y$, siis $X_n + Y_n \xrightarrow{p} X + Y$ ja $aX_n \xrightarrow{p} aX$;*

4) *kui $X_n \Rightarrow X$, siis $aX_n \Rightarrow aX$.*

Tõestus.

1) järeldub vahetult arvjada analoogilistest omadustest (ülesanne 1);

2) kasuta järeldust 9.3 (ülesanne 8 a));

3) kasuta seost $(a + b)^p \le 2^p(a^p + b^p)$, $a, b \ge 0$ (ülesanne 20);

4) kasuta definitsiooni (ülesanne 23).

■

On lihtne veenduda (aga veendu!), et kui $X_n \Rightarrow X$ ja $Y_n \Rightarrow Y$, siis EI KEHTI koondumine $X_n + Y_n \Rightarrow X + Y$. Erand vaid siis, kui üks piirväärtustest on konstant. Selle postuleerib järgmine lemma.

Lemma 9.4 *(Slutsky) Kui $X_n \Rightarrow X$ ja $Y_n \Rightarrow c$, kus $c \in \mathbb{R}$ on konstant, siis kehtivad koondumised:*

1. *$X_n + Y_n \Rightarrow X + c$;*

2. $Y_n X_n \Rightarrow cX$*;*

3. $Y_n^{-1} X_n \Rightarrow c^{-1} X$*, kui* $c \neq 0$*.*

Tõestus. 1) Olgu $X_n \Rightarrow X$ ja $Y_n \xrightarrow{P} 0$. Tõestame, et $X_n + Y_n \Rightarrow X$. Olgu $x \in \mathbb{R}$ ja $\epsilon > 0$. Olgu F juhusliku suuruse X jaotusfunktsioon. Leiame

$$\begin{aligned}\mathbf{P}(X_n + Y_n \le x) &= \mathbf{P}(X_n + Y_n \le x, |Y_n| > \epsilon) + \mathbf{P}(X_n + Y_n \le x, |Y_n| \le \epsilon) \\ &\le \mathbf{P}(|Y_n| > \epsilon) + \mathbf{P}(X_n \le x + \epsilon) \\ \mathbf{P}(X_n + Y_n > x) &= \mathbf{P}(X_n + Y_n > x, |Y_n| > \epsilon) + \mathbf{P}(X_n + Y_n > x, |Y_n| \le \epsilon) \\ &\le \mathbf{P}(|Y_n| > \epsilon) + \mathbf{P}(X_n > x - \epsilon).\end{aligned}$$

Viimane võrratus on

$$\mathbf{P}(X_n + Y_n \le x) \ge 1 - \big(\mathbf{P}(|Y_n| > \epsilon) + \mathbf{P}(X_n > x - \epsilon)\big) = -\mathbf{P}(|Y_n| > \epsilon) + \mathbf{P}(X_n \le x - \epsilon).$$

Kui ϵ on selline, et $x + \epsilon$ ja $x - \epsilon$ on F pidevuspunktid, siis

$$\mathbf{P}(X_n \le x + \epsilon) \to F(x + \epsilon), \quad \mathbf{P}(X_n \le x - \epsilon) \to F(x - \epsilon).$$

Et $Y \xrightarrow{P} 0$, siis $\mathbf{P}(|Y_n| > \epsilon) \to 0$, millest

$$\limsup_n \mathbf{P}(X_n + Y_n \le x) \le F(x + \epsilon), \quad \liminf_n \mathbf{P}(X_n + Y_n \le x) \ge F(x - \epsilon).$$

Et x on F pidevuspunkt, siis iga $\delta > 0$ korral leidub ϵ nii, et $x + \epsilon$ ja $x - \epsilon$ on F pidevuspunktid ning kehtivad võõratused $F(x + \epsilon) - F(x) \le \delta$ ja $F(x) - F(x - \epsilon) \le \delta$. Seega $\mathbf{P}(X_n + Y_n \le x) \to F(x)$.
Veendu, et kui $X_n \Rightarrow X$, siis $X_n + c \Rightarrow X + c$ ning kui $Y_n \xrightarrow{P} c$, siis $Y_n - c \xrightarrow{P} 0$. Nendest järeldub, et $X_n + Y_n = X_n + c + Y_n - c \Rightarrow X + c$.

2) Olgu $X_n \Rightarrow X$ ja $Y_n \xrightarrow{P} 0$. Tõestame, et $X_n Y_n \xrightarrow{P} 0$. Paneme tähele, et iga $x > 0$ korral

$$\begin{aligned}\mathbf{P}(|X_n Y_n| \ge \epsilon) = \mathbf{P}(|X_n||Y_n| \ge \epsilon) &= \mathbf{P}(|X_n||Y_n| \ge \epsilon, |X_n| \le x) + \mathbf{P}(|X_n||Y_n| \ge \epsilon, |X_n| > x) \\ &\le \mathbf{P}(|Y_n| \ge \frac{\epsilon}{x}) + \mathbf{P}(|X_n| > x).\end{aligned}$$

Valides etteantud $\delta > 0$ korral x (sõltub δ-st) nii suure, et $1 - F(x) + F(-x) < \delta$ ja x on F pidevuspunkt, siis koondumisest $X_n \Rightarrow X$ järeldub, et

$$\begin{aligned}\mathbf{P}(|X_n| > x) &= \mathbf{P}(X_n > x) + \mathbf{P}(X_n < -x) \le 1 - \mathbf{P}(X_n \le x) + \mathbf{P}(X_n \le -x) \to \\ &\to 1 - F(x) + F(-x) < \delta,\end{aligned}$$

millest saame, et $\limsup_n \mathbf{P}(|X_n| > x) \le \delta$. Järelikult $\limsup_n \mathbf{P}(|X_n Y_n| \ge \epsilon) \le \delta$ ehk $\mathbf{P}(|X_n Y_n| \ge \epsilon) \to 0$.
Kui $Y_n \xrightarrow{P} c$, siis $Y_n - c \xrightarrow{P} 0$, millest osa 1) tõttu

$$X_n Y_n = X_n(Y_n - c + c) = X_n(Y_n - c) + X_n c \Rightarrow Xc,$$

sest $X_n(Y_n - c) \xrightarrow{P} 0$ ja $X_n c \Rightarrow Xc$.

Tõestamaks 3) veendume, et kui $Y_n \xrightarrow{P} c$ ja $c \neq 0$, siis $Y_n^{-1} \xrightarrow{P} \frac{1}{c}$. Tõepoolest, järelduse 9.3 tõttu piisab kui näitame, et igal Y_n^{-1} alamjadal $Y_{n_k}^{-1}$ on omakorda alamjada $Y_{n_{k_l}}^{-1}$ mis koondub konstandiks c^{-1} p.k. Et $Y_n \xrightarrow{P} c$, siis jadal Y_{n_k} on alamjada $Y_{n_{k_l}}$, mis koondub konstandiks c p.k. Kuid see tähendab, et $Y_{n_{k_l}}^{-1} \to c^{-1}$ p.k. Seos 3) järeldub nüüd seosest 2). ■

9.6 Juhuslike vektorite koondumine

Vaatleme k-dimensionaalseid juhuslikke vektoreid. Jada $x_n \in \mathbb{R}^k$ koondub vektoriks $x \in \mathbb{R}^k$ parajasti siis, kui $\|x_n - x\| \to 0$, kus $\|x\|$ on eukleidiline norm, st $x = (x^1, \ldots, x^k) \in \mathbb{R}^k$ korral

$$\|x\| = \|(x^1, \ldots, x^k)\| = \left((x^1)^2 + \cdots + (x^k)^2\right)^{\frac{1}{2}}.$$

Nüüd on selge, et asendades absoluutväärtuse $|\cdot|$ normiga $\|\cdot\|$ üldistuvad ülaltoodud koondumised väga loomulikult juhuslikele vektoritele.
Olgu $X_n = (X_n^1, \ldots X_n^k)$ juhuslike vektorite jada. Ütleme, et jada koondub juhuslikuks vektoriks $X = (X_1, \ldots, X_k)$

- peaaegu kindlasti, kui

$$\mathbf{P}\{\omega : X_n(\omega) \to X(\omega)\} = \mathbf{P}\{\omega : \|X_n(\omega) - X(\omega)\| \to 0\} = 1.$$

  Peaaegu kindlasti koondumist tähistatakse $X_n \to X$ p.k.

- tõenäosuse järgi, kui

$$\mathbf{P}\{\omega : \|X_n(\omega) - X(\omega)\| > \epsilon\} \to 0, \quad \forall \epsilon > 0.$$

  Tõenäosuse järgi koondumist tähistatakse $X_n \xrightarrow{P} X$.

- p-keskmise mõõtes ($p > 0$) või ka ruumis L_p kui

$$E\|X_n - X\|^p \to 0.$$

  Koondumist p-keskmise mõttes tähistatakse $X_n \xrightarrow{p} X$.

Juhuslike vektorite jaotuse järgi koondumise defineerime hiljem.
Toodud definitsioonidest ja juhuslike suuruste vastavatest omadustest (teoreem 9.7) järeldub vahetult (defineeri juhuslik suurus $Y_n := \|X_n - X\|$), et p.k. koondumisest järeldub tõenäosuse järgi koondumine ja p-keskmise mõttes koondumisest järeldub ka tõenäosuse järgi koondumine.

Komponentide koondumine. Ülaltoodud kolme koondumisliigi korral on juhusliku vektori koondumine ekvivalentne kõigi komponentide koondumisega:

$$\begin{array}{lll}
(X_n^1,\dots,X_n^k)\to(X^1,\dots,X^k)\quad \text{p.k.} & \Leftrightarrow & X_n^i\to X^i\quad \text{p.k.}\quad i=1,\dots,k\\
(X_n^1,\dots,X_n^k)\xrightarrow{p}(X^1,\dots,X^k) & \Leftrightarrow & X_n^i\xrightarrow{p} X^i\quad i=1,\dots,k\\
(X_n^1,\dots,X_n^k)\xrightarrow{P}(X^1,\dots,X^k) & \Leftrightarrow & X_n^i\xrightarrow{P} X^i\quad i=1,\dots,k
\end{array}$$

Esimene järeldub vahetult definitsioonist. Teise tõestamiseks paneme tähele, et iga $x=(x^1,\dots,x^k)\in\mathbb{R}^k$ korral $\max_i|x^i|\leq\|x\|\leq|x^1|+\dots+|x^k|$, millest

$$\max_i|x^i|^p\leq\|x\|^p\leq(|x^1|+\dots+|x^k|)^p.$$

Seega iga i korral

$$|X_n^i-X^i|^p\leq\|X_n-X\|^p\leq\big(|X_n^1-X^1|+\dots+|X_n^k-X^k|\big)^p,$$

millest iga i korral

$$E|X_n^i-X^i|^p\leq E\|X_n-X\|^p\leq E\big(|X_n^1-X^1|+\dots+|X_n^k-X^k|\big)^p.$$

Minkowski võrratusest saame nüüd, et

$$\max_i\|X_n^i-X^i\|_p\leq\|X_n-X\|_p\leq\sum_{i=1}^k\|X_n^i-X^i\|_p.$$

Kolmanda samasuse tõestuseks piisab kui paneme tähele, et

$$\max_i\mathbf{P}\Big((X_n^i)^2>\epsilon^2\Big)\leq\mathbf{P}\Big(\sum_{i=1}^k(X_n^i)^2>\epsilon^2\Big)=\mathbf{P}(\|X_n\|>\epsilon)\leq\sum_{i=1}^k\mathbf{P}\Big((X_n^i)^2>\frac{\epsilon^2}{k}\Big).$$

9.7 Stohhastilised O ja o

Olgu $\{x_n\}$ ja $\{\epsilon_n\}$ reaalarvude jadad, $\epsilon_n>0$ iga n korral. Tuletame meelde, et $x_n=O(\epsilon_n)$, kui leidub $K<\infty$ nii, et $|x_n|\leq K\epsilon_n$ ehk $\frac{|x_n|}{\epsilon_n}<K$ iga n korral. Näiteks $x_n=O(1)$ tähendab jada $\{x_n\}$ tõkestatust ning $x_n=O(n^{-1})$ tähendab, et jada $\{nx_n\}$ on tõkestatud (jada x_n ise peab sel juhul koonduma nulliks).

Tuletame samuti meelde, et $x_n=o(\epsilon_n)$ tähendab, et $\frac{x_n}{\epsilon_n}\to 0$. Seega $x_n=o(1)$ tähendab, et $\{x_n\}$ on nulliks koonduv jada ning $x_n=o(n^{-1})$ tähendab, et x_n koondub nullks nii kiiresti, et isegi nx_n koondub nulliks.

Kui $\{x_n\}$ on $\mathbb{R}^k$ elementidest moodustatud jada, siis $x_n=O(\epsilon_n)$ tähendab, et leidub $K<\infty$ nii, et $\|x_n\|\leq K\epsilon_n$ ehk $\frac{\|x_n\|}{\epsilon_n}<K$ iga n korral ning $x_n=o(\epsilon_n)$ tähendab, et $\frac{x_n}{\epsilon_n}\to 0$. Viimane on arusaadavalt ekvivalentne $\frac{\|x_n\|}{\epsilon_n}\to 0$.

Olgu nüüd X_n k-dimensionaalsed juhuslikud vektorid. Ütleme, et $X_n = O_P(\epsilon_n)$, kui iga $\delta > 0$ korral leiduvad $K(\delta) < \infty$ ja $n_o(\delta)$ nii, et

$$\mathbf{P}\big(\|X_n\| > K\epsilon_n\big) < \delta, \quad \text{kui} \quad n > n_o(\delta). \tag{9.2}$$

Pane tähele, et $\mathbf{P}(\|X_n\| < \infty) = 1$, siis seosest (9.2) järeldub, et leidub $K' < \infty$ nii, et

$$\mathbf{P}\big(\|X_n\| > K'\epsilon_n\big) < \delta, \quad n = 1, 2, \ldots. \tag{9.3}$$

Seega sellisel juhul võib $O_P(\epsilon_n)$ definitisioonis nõude $n > n_o(\delta)$ asendada tugevama nõudega $\forall n$. Kui juhuslikud suurused X_n koonduvad nõrgalt, st $X_n \Rightarrow X$, siis $X_n = O_P(1)$ (veendu selles). Näiteks, kui T_n on parameetri μ selline hinnang, et $\sqrt{n}(T_n - \mu) \Rightarrow X$, siis $T_n - \mu = O_P(n^{-\frac{1}{2}})$. Tsentraalse piirteoreemi tõttu on $O_P(n^{-\frac{1}{2}})$ statistikas levinud kiirus.

Ütleme, et $X_n = o_P(\epsilon_n)$, kui $\frac{X_n}{\epsilon_n} \xrightarrow{P} 0$ ehk iga $\delta > 0$ korral $\mathbf{P}(\|X_n\| > \delta\epsilon_n) \to 0$. Seega $X_n = o_P(1)$, kui $X_n \xrightarrow{P} 0$.

Kirjandus:

Juhuslike suuruste koondumistüüpide kohta loe *Grimmet, Stirzaker* osad 7.2 ja 7.3; *Ŝirjajev* osa II 10.

9.8 Ülesanded

1. Tõestada teoreemi 9.9 väide 1)

2. Tõestada teoreemi 9.5 implikatsioon 2)$\Leftrightarrow$ 3).

3. Tõestada teoreemi 9.5 implikatsioon 3)$\Leftrightarrow$ 4).

4. Tõestada järeldus 9.1.

5. Tõestada järeldus 9.2.

6. Tõestada teoreemi 9.6 implikatsioon 2)$\Leftrightarrow$ 3).

7. Tõestada teoreemi 9.6 implikatsioon 3)$\Leftrightarrow$ 4).

8. a) Tõestada teoreemi 9.7 väide 2);
b) tõestada, et kui $X_n \xrightarrow{P} X$, siis $|X_n| \xrightarrow{P} |X|$.

9. Olgu $\{X_n\}$ juhuslike suuruste jada, kusjuures juhusliku suuruse X_n jaotus on järgmine: $\mathbf{P}(X_n = e^{-an}) = 1 - e^{-bn}$ ja $\mathbf{P}(X_n = e^{an}) = e^{-bn}$. Millistel a ja b väärtustel $X_n \xrightarrow{P} 0$.

10. Olgu $\{X_n\}$ juhuslike suuruste jada, kusjuures $\mathbf{P}(|X_n| \geq c > 0) \geq \delta > 0 \quad \forall n$ korral. Olgu $\{a_n\}$ reaalarvude jada, kusjuures $a_n X_n \xrightarrow{P} 0$. Tõestada, et $a_n \to 0$.

11. Tõestada järeldus 9.3

12. a) tõestada, et koondumisest p-keskmise mõttes ($p > 0$) järeldub tõenäosuse järgi koondumine;
b) tõestada, et koondumisest p-keskmise mõttes jareldub koondumine r-keskmise mõttes ($p \geq r > 0$).

13. Olgu juhuslik suurus X jaotusega $\mathbf{P}(X = 0) = \mathbf{P}(X = 1) = 0.5$. Defineerime juhuslikud suurused $X_n := X \; \forall \, n$ korral, $Y := 1 - X$. Kas $X_n \Rightarrow Y$, $X_n \xrightarrow{P} Y$, $X_n \to Y$ p.k., $X_n \xrightarrow{p} Y$ ($p > 0$)?

14. Olgu juhuslikud suurused X_n jaotusega $\mathbf{P}(X_n = n^3) = n^{-2}$, $\quad \mathbf{P}(X_n = 0) = 1 - n^{-2}$. Kas $X_n \Rightarrow 0$, $X_n \xrightarrow{P} 0$, $X_n \xrightarrow{p} 0$ ($p \geq 1$) ja $X_n \to 0$ p.k.?

15. Olgu juhuslikud suurused X_n jaotusega $\mathbf{P}(X_n = 0) = 1 - n^{\frac{-(r+p)}{2}}$, $\mathbf{P}(X_n = n) = n^{\frac{-(r+p)}{2}}$, ($p > r > 0$). Kas $X_n \xrightarrow{p} 0$? Kas $X_n \xrightarrow{r} 0$?

16. Tõestada lemma 9.3.
(Tõestada, et $\forall \, \epsilon > 0$ korral $\mathbf{P}(|X_n - c| > \epsilon) = \mathbf{P}(X_n > c + \epsilon) + \mathbf{P}(X_n < c - \epsilon) \to 0$.)

17. Olgu sõltumatud juhuslikud suurused X_n jaotusega $\mathbf{P}(X_n = 1) = n^{-1}$, $\mathbf{P}(X_n = 0) = 1 - n^{-1}$. Kas $X_n \Rightarrow 0$, $X_n \xrightarrow{P} 0$, $X_n \to 0$ p.k., $X_n \xrightarrow{p} 0$ ($p > 0$)?

18. Olgu $(\Omega, \mathcal{F}, \mathbf{P}) = ([0,1], \mathcal{B}, Leb)$, $X_n := I_{[0,n^{-1}]}$. Kas $X_n \Rightarrow 0$, $X_n \xrightarrow{P} 0$, $X_n \to 0$ p.k., $X_n \xrightarrow{p} 0$ ($p > 0$)?

19. Olgu $(\Omega, \mathcal{F}, \mathbf{P}) = ([0,1], \mathcal{B}, Leb)$, $X_n := e^n I_{[0,n^{-1}]}$. Kas $X_n \Rightarrow 0$, $X_n \xrightarrow{P} 0$, $X_n \to 0$ p.k., $X_n \xrightarrow{p} 0$ ($p > 0$)?

20. Tõestada teoreemi 9.9 väide 3).

21. Olgu $E|X_n|^p < \infty$, $E|X|^p < \infty$
1) Olgu $p \geq 1$. Minkowski võrratust kasutades tõestada, et koondumisest $X_n \xrightarrow{p} X$ järeldub koondumine $E|X_n|^p \to E|X|^p$. Koondumisest $X_n \xrightarrow{2} X$ järelda $D(X_n) \to D(X)$.
2) Olgu $p \in (0,1)$. Võrratust $(|x| + |y|)^p \leq |x|^p + |y|^p$ kasutades näita, et koondumisest $X_n \xrightarrow{p} X$ järeldub koondumine $E|X_n|^p \to E|X|^p$.

22. Olgu $X_n, X \in \mathcal{L}_2(\Omega, \mathcal{F}, \mathbf{P})$. Tõesta või lükka ümber:

1. $X_n \xrightarrow{2} X \quad \Leftrightarrow \quad DX_n \to DX, \quad EX_n \to EX$;

2. $X_n \xrightarrow{2} a \quad \Leftrightarrow \quad DX_n \to 0, \quad EX_n \to a$, $a \in \mathbb{R}$;

23. Olgu $X_1 \leq X_2 \leq \cdots$ juhuslikud suurused, kusjuures $X_n \xrightarrow{P} X$. Tõestada, et $X_n \to X$. p.k.

24. Tõestada teoreemi 9.9 väide 4).

25. Olgu X_n p.k. konstantne, s.t. $X_n = a_n$ p.k., $n = 1, 2, \ldots$. Tõestada, et $X_n \Rightarrow a \in \mathbb{R}$ parajasti siis, kui $a_n \to a$.

26. Tõestada, et kui $X_n - a_n \xrightarrow{P} 0$ ja $X_n - b_n \xrightarrow{P} 0$, kus $\{a_n\}$ ja $\{b_n\}$ on reaalarvude jadad, siis $a_n - b_n \to 0$.

27. Olgu $X_1, X_2, \ldots$ juhuslike suuruste jada, mis koondub konstandiks a p.k., st $X_n \to a$, p.k.. Olgu $Y_1, Y_2, \ldots$ teine juhuslike suuruste jada, kusjuures iga n korral X_n ja Y_n on sama jaotusega.

1. Näita, et $Y_n \xrightarrow{P} a$.

2. Kontranäite abil veendu, et koondumine $Y_n \to a$ p.k. ei pruugi kehtida (kasuta näiteks jada (9.1)).

3. Näita, et kui iga n korral juhuslike vektorite $(X_1, \ldots, X_n)$ ja $(Y_1, \ldots, Y_n)$ jaotused on võrdsed, siis $Y_n \to a$ p.k.(kasuta näiteks teoreemi 9.5 väidet 3)).

28. Tõesta järeldus 9.4

29. (Anscombe teoreem, erijuht) Olgu N_n mittenegatiivsete täisarvuliste juhuslike suuruste jada nii, et $\frac{N_n}{m_n} \xrightarrow{P} 1$, kus $m_n \nearrow \infty$ on täisarvude jada. Olgu $X_1, X_2, \ldots$ iid juhuslikud suurused nii, et $EX_i = 0$ ja $EX_i^2 = \sigma^2$. Olgu $S_k = X_1 + \cdots + X_k$. Näita, et

$$\frac{S_{N_n}}{\sqrt{m_n}} - \frac{S_{m_n}}{\sqrt{m_n}} \xrightarrow{P} 0. \tag{9.4}$$

(Kasuta eelmise peatüki ülesannet nr 28.) Olgu Z on selline juhuslik suurus, et

$$\frac{S_{m_n}}{\sqrt{m_n}} \Rightarrow Z \tag{9.5}$$

Järelda, et

$$\frac{S_{N_n}}{\sqrt{m_n}} \Rightarrow Z, \quad \frac{S_{N_n}}{\sqrt{N_n}} \Rightarrow Z. \tag{9.6}$$

30. Üldista eelmist teoreemi juhule, kui m_n pole ilmtingimata täisarv. Sellisel juhul võrratus (9.4) ja koondumine (9.7) loomulikult ei kehti, sest S_{m_n} pole defineeritud. Küll kehtib järgmine: kui kehtivad ülesande **29.** eeldused, kuid m_n pole täisarv ja kui Z on selline juhuslik suurus, et

$$\frac{S_{\lceil m_n \rceil}}{\sqrt{\lceil m_n \rceil}} \Rightarrow Z, \tag{9.7}$$

siis kehtivad võrratused (9.6). Näita seda.

10 Juhuslike ridade koondumine

Olgu $X_1, X_2, \ldots$ tõenäosusruumil $(\Omega, \mathcal{F}, \mathbf{P})$ antud *sõltumatud* juhuslikud suurused, $X_n \sim P_n$. Uurime juhusliku rea $\sum_n X_n$ koondumist.

Sündmus

$$F := \Big\{ \sum_n X_n \text{ koondub } \Big\}$$

on jada $\{X_n\}$ jääk σ-algebra element (miks?). Et $\{X_n\}$ on sõltumatud, siis Kolmogorovi 0-1 seaduse kohaselt sündmuse F mõõt on kas 1 või 0, s.t. kõnealune rida koondub p.k. või hajub p.k. Järgnevad teoreemid annavad mõne piisava ja tarviliku tingimuse selleks, et $\mathbf{P}(F) = 1$. Tõestame abitulemuse.

Lemma 10.1 *Olgu $X_1, \ldots, X_n$ sõltumatud juhuslikud suurused, kusjuures $|X_i| \leq K < \infty$ p.k., $EX_i = 0$, $DX_i = \sigma_i^2 < \infty$, $i = 1, \ldots, n$. Defineerime $S_k := X_1 + \cdots + X_k$. Kui $c > 0$, siis*

$$\mathbf{P}(\max_{1 \leq k \leq n} |S_k| \geq c) \geq 1 - \frac{(K+c)^2}{D(S_n)} = 1 - \frac{(K+c)^2}{\sum_{i=1}^n \sigma_i^2}.$$

Tõestus. Kasutame Kolmogorovi võrratuse tõestusest tuttavaid hulki

$$A := \{ \max_{1 \leq k \leq n} |S_k| \geq c \}, \quad A_k := \{ |S_i| < c,\, i = 1, \ldots, k-1, |S_k| \geq c \}, \quad k = 1, \ldots, n.$$

Me teame, et $A = \cup_{k=1}^n A_k$ ja $\mathbf{P}(A) = \sum_{k=1}^n \mathbf{P}(A_k)$ (miks ?).

Kui $\omega \in A^c$, siis $|S_n(\omega)| \leq \max_{1 \leq k \leq n} |S_k(\omega)| \leq c$, millest $E(S_n^2 I_{A^c}) \leq c^2 \mathbf{P}(A^c)$. Seega

$$E(S_n^2 I_A) = ES_n^2 - E(S_n^2 I_{A^c}) \geq ES_n^2 - c^2 \mathbf{P}(A^c) \geq ES_n^2 - c^2 + c^2 \mathbf{P}(A). \tag{10.1}$$

Kui $\omega \in A_k$, siis $|S_k(\omega)| \leq |S_{k-1}(\omega)| + |X_k(\omega)| < c + |X_k(\omega)|$, millest $S_k I_{A_k} \leq (c+K) I_{A_k}$ p.k. ja nii

$$E(S_k^2 I_{A_k}) \leq (c+K)^2 \mathbf{P}(A_k), \quad k = 1, \ldots, n.$$

Et juhuslikud suurused I_{A_k} ja $X_{k+1} + \cdots + X_n$ on sõltumatud (miks?), siis

$$E(I_{A_k}(S_n - S_k)^2) = \mathbf{P}(A_k) \sum_{i=k+1}^n \sigma_i^2 \leq \mathbf{P}(A_k) ES_n^2, \quad k = 1, \ldots, n.$$

Kasutades Kolmogorovi võrratuse tõestusest tuttavat arutelu, saame

$$E(S_n^2 I_A) = \sum_{k=1}^n E(S_k^2 I_{A_k}) + \sum_{k=1}^n E((S_n - S_k)^2 I_{A_k}) \leq \mathbf{P}(A)[(c+K)^2 + ES_n^2]. \tag{10.2}$$

Seostest (10.1) ja (10.2) saame seose $(c+K)^2 - c^2 \geq 0$ tõttu

$$\mathbf{P}(A) \geq \frac{ES_n^2 - c^2}{(c+K)^2 + ES_n^2 - c^2} \geq 1 - \frac{(c+K)^2}{ES_n^2}.$$

■

Seega, lemma 10.1 eeldustel kehtivad võrratused

$$1 - \frac{(K+c)^2}{\sum_{i=1}^n \sigma_i^2} \le \mathbf{P}(\max_{1\le k\le n} |S_k| \ge c) \le \frac{\sum_{i=1}^n \sigma_i^2}{c^2}.$$

10.1 Ridade koondumisteoreemid

Esimene teoreem annab piisava ja tarviliku tingimuse juhusliku rea koondumiseks juhul, kui $EX_n = 0$, $\forall n$.

Teoreem 10.1 (Kolmogorov, Hintŝin). *Olgu X_n, $n = 1, 2, \ldots$ sõltumatud juhuslikud suurused, $EX_n = 0$, $n = 1, 2, \ldots$.*

1) *Kui $\sum_n D(X_n) < \infty$, siis rida $\sum_n X_n$ koondub p.k.;*

2) *kui $\exists\ K < \infty$ nii, et $|X_n| \le K$ p.k. $\forall n$ siis $\sum_n X_n$ koondub p.k. parajasti siis, kui $\sum_n D(X_n) < \infty$.*

Tõestus. 1) Rea $\sum_n X_n$ koondumine tähendab osasummade jada S_n, $n = 1, 2, \ldots$ koondumist. Teame, et jada $\{S_n\}$ koondub parajasti siis, kui ta on Cauchy jada. Vastavalt teoreemile 9.6 on jada $\{S_n\}$ Cauchy jada p.k. parajasti siis, kui iga $\epsilon > 0$ korral kehtib koondumine

$$\lim_{m\to\infty} \mathbf{P}(\sup_{k\ge 0} |S_{m+k} - S_m| > \epsilon) \to 0.$$

Fikseerime $\epsilon > 0$. Iga m korral

$$\{\max_{0\le k\le n} |S_{m+k} - S_m| > \epsilon\} \nearrow \{\sup_{k\ge 0} |S_{m+k} - S_m| > \epsilon\},$$

millest mõõdu pidevuse ja Kolmogorovi võrratuse tõttu saame nüüd

$$\begin{aligned}\mathbf{P}(\sup_{k\ge 0} |S_{m+k} - S_m| > \epsilon) &= \lim_{n\to\infty} \mathbf{P}(\max_{0\le k\le n} |S_{m+k} - S_m| > \epsilon)\\ &\le \lim_{n\to\infty} \frac{\sum_{k=m+1}^n D(X_k)}{\epsilon^2} = \frac{\sum_{k=m+1}^\infty D(X_k)}{\epsilon^2}.\end{aligned}$$

Kui $\sum_n D(X_n) < \infty$, siis $\lim_m \sum_{n=m}^\infty D(X_n) = 0$. Seega $\{S_n\}$ on Cauchy jada p.k.

2) Oletame, et rida $\sum_n X_n$ koondub p.k. Siis $\{S_n\}$ on Cauchy jada p.k. ja iga ϵ korral

$$\mathbf{P}(\sup_{k\ge 0} |S_{m+k} - S_m| > \epsilon) < \frac{1}{2}, \qquad (10.3)$$

kui m on piisavalt suur. Teisest küljest, lemma 10.1 tõttu aga iga m korral

$$\mathbf{P}(\sup_{k\ge 0} |S_{m+k} - S_m| > \epsilon) \ge 1 - \frac{(K+\epsilon)^2}{\sum_{k=m+1}^\infty D(X_k)}.$$

(ülesanne 1). Seega, kui $\sum_n D(X_n) = \infty$, siis $\mathbf{P}(\sup_{k\geq 0} |S_{m+k} - S_m| > \epsilon) = 1$, mis on vastuolus võrratusega (10.3). ■

Näide 1: Rida $\sum_n \frac{1}{n}$ hajub, vahelduvate märkidega rida $\sum_n (-1)^n \frac{1}{n}$ koondub. Kas sõltumatute liikmetega rida $\sum_n \frac{X_n}{n}$ koondub või hajub, kui

$$\mathbf{P}(X_n = 1) = \mathbf{P}(X_n = -1) = 0.5?$$

Et $E(\frac{X_n}{n}) = 0$, ja $\sum_n E(\frac{X_n}{n})^2 < \infty$, siis teoreemi 10.1 põhjal rida $\sum_n \frac{X_n}{n}$ koondub p.k. (vaata ka ülesanne 2).

Järgnev teoreem üldistab teoreemi 10.1, sest tingimust $EX_n = 0$, $\forall n$ ei eeldata.

Teoreem 10.2 (kahe rea teoreem). *Olgu X_n, $n = 1, 2, \ldots$ sõltumatud juhuslikud suurused.*

1) *kui $\sum_n D(X_n) < \infty$ ja rida $\sum_n EX_n$ koondub, siis rida $\sum_n X_n$ koondub p.k.;*

2) *kui $\exists\, K < \infty$ nii, et $|X_n| \leq K$ p.k. $\forall n$, siis $\sum_n X_n$ koondub p.k. parajasti siis, kui $\sum_n D(X_n) < \infty$ ja rida $\sum_n EX_n$ koondub.*

Tõestus. 1) Järeldub teoreemist 10.1(ülesanne 3).

2) Tarvilikkuse tõestamiseks kasutame nn. sümmetriseerimist. Olgu Y_n, $n = 1, 2, \ldots$ tõenäosusruumil $(\Omega, \mathcal{F}, \mathbf{P})$ antud selline juhuslike suuruste jada, et $Y_n \sim P_n$ $\forall n$ ja hulka $\{X_n, Y_n\}$ kuuluvad juhuslikud suurused on sõltumatud.
Oletame, et rida $\sum_n X_n$ koondub p.k. Seega ka rida $\sum_n Y_n$ koondub p.k., seega ka $\sum_n (X_n - Y_n)$ koondub p.k. Et $E(X_n - Y_n) = 0$ ja $|X_n - Y_n| \leq 2K$ p.k., siis teoreemi 10.1 **2)** põhjal $\sum_n D(X_n - Y_n) < \infty$. Et $DX_n = \frac{1}{2} D(X_n - Y_n)$ (miks?), siis $\sum_n D(X_n) = \frac{1}{2} \sum_n D(X_n - Y_n) < \infty$. Seega teoreemi 10.1 **1)** põhjal rida $\sum_n (X_n - EX_n)$ koondub p.k. Vastavalt eeldusele $\sum_n X_n$ koondub p.k.. Järelikult rida $\sum_n EX_n$ koondub.
■

Märkus: Olgu X_n mittenegatiivsed juhuslikud suurused. Kui $\sum_n EX_n < \infty$, siis $\sum_n X_n$ koondub p.k. (ülesanne 4). Seega võib $\sum_n X_n$ koonduda ka siis, kui $\sum_n D(X_n) = \infty$. Teoreemi 10.2 tõttu võib see juhtuda vaid tõkestamata juhuslike suuruste korral.

Näide 2: Olgu X_n sõltumatud, kusjuures $n^{-1} X_n \sim B(1, n^{-3})$. Et juhuslikud suurused on mittenegatiivsed, ja $\sum_n EX_n < \infty$, siis rida $\sum_n X_n$ koondub p.k. On kerge veenduda, et $\sum_n D(X_n) = \infty$, st teoreemi 10.2 väide **2)** ei kehti (ülesanne 5).

Järgnev, kõige üldisem teoreem annab piisava ja tarviliku tingimuse sõltumatute liikmetega juhusliku rea koondumiseks. Juhusliku suuruse X ja konstandi $K < \infty$ korral defineerime

$$X^K(\omega) = (X I_{\{|X| \leq K\}})(\omega) = \begin{cases} X(\omega) & \text{kui } |X(\omega)| \leq K, \\ 0 & \text{kui } |X(\omega)| > K. \end{cases}$$

Teoreem 10.3 (Kolmogorovi kolme rea teoreem). *Olgu X_n, $n = 1, 2, \ldots$ sõltumatud juhuslikud suurused.*

1) *Kui $\exists K \geq 0$ nii, et*

$$\sum_n D(X_n^K) < \infty, \quad \sum_n \mathbf{P}(|X_n| > K) < \infty \quad \text{ja} \quad \sum_n E(X_n^K) \,\text{koondub}, \tag{10.4}$$

siis rida $\sum_n X_n$ koondub p.k.;

2) *kui $\sum_n X_n$ koondub p.k., siis (10.4) kehtib iga $0 < K < \infty$ korral.*

Tõestus.
1) Et $\sum_n \mathbf{P}(|X_n| > K) < \infty$, siis BC1 põhjal $\mathbf{P}(X_n = X_n^K \text{ ev }) = 1$ (miks?). Seega rida $\sum_n X_n$ koondub p.k. parajasti siis, kui $\sum_n X_n^K$ koondub p.k. Viimane koondumine aga järeldub teoreemist 10.2.

2) Olgu $0 < K < \infty$ suvaline ja oletame, et $\sum_n X_n$ koondub p.k. Siis $X_n \to 0$ p.k., millest $\mathbf{P}(|X_n| \leq K \text{ ev}) = 1$. Et sündmused $\{|X_n| > K\}$, $n = 1, 2, \ldots$ on sõltumatud, siis BC2 põhjal $\sum_n \mathbf{P}(|X_n| > K) < \infty$. Rea $\sum_n X_n$ koondumisest p.k. järeldub rea $\sum_n X_n^K$ koondumine p.k.(miks?). Teoreemi 10.2 tõttu nüüd koonduvad ka read $\sum_n E(X_n^K)$ ja $\sum_n D(X_n^K)$. ∎

Märkus: Teoreemist 10.3 järeldub, et kui $\exists K \geq 0$ nii, et (10.4) kehtib, siis kehtib see ka suvalise $0 < K < \infty$ korral.

Näide 3: Olgu $\sum_n X_n$ näites 2 vaadeldud juhuslik rida. Et $\sum_n X_n < \infty$ p.k., siis vastavalt teoreemile 10.3 iga $K > 0$ korral $\sum_n \mathbf{P}(|X| > K) < \infty$, $\sum_n E(X_n^K)$ koondub ja $\sum_n D(X_n^K) < \infty$. On lihtne veenduda, et see tõepoolest nii on (ülesanne 5).

10.2 Ülesanded

1. Olgu $X_1, X_2, \ldots$ konstandiga K tõkestatud sõltumatud juhuslikud suurused, $EX_n = 0$ $\forall n$. Tõestada, et

$$\mathbf{P}(\sup_{k\geq 1} |S_{m+k} - S_m| \geq \epsilon) \geq \begin{cases} 1 & \text{kui } \sum_n D(X_n) = \infty, \\ 1 - \frac{(K+\epsilon)^2}{\sum_{k=m+1}^{\infty} D(X_k)} & \text{kui } \sum_n D(X_n) < \infty. \end{cases}$$

2. Olgu X_n sõltumatud, $P_n(-1) = P_n(1) = \frac{1}{2}$. Leida piisav ja tarvilik tingimus rea $\sum_n a_n X_n$ koondumiseks ($a_n \in \mathbb{R}$).

3. Tõestada piisavus kahe rea teoreemis.

4. Olgu $X_1, X_2, \ldots$ mittenegatiivsed juhuslikud suurused (võivad olla sõltuvad). Tõestada, et kui $\sum_n EX_n < \infty$, siis $\sum_n X_n < \infty$ p.k. Tõestada, et kui $\{X_n\}$ on sõltumatud ja $|X_n| \leq K$ p.k., siis tingimusest $\sum_n EX_n = \infty$ järeldub, et $\sum_n X_n = \infty$ p.k.

5. Olgu $X_1, X_2, \ldots$ sõltumatud juhuslikud suurused, $n^{-1}X_n \sim B(1, n^{-3})$.
1) Leida $\sum_n EX_n$, $\sum_n DX_n$. Kas $\sum_n X_n < \infty$ p.k.?
2) Olgu $K < \infty$ suvaline. Kas koonduvad: $\sum_n \mathbf{P}(|X_n| > K)$, $\sum_n E(X_n^K)$, $\sum_n D(X_n^K)$?

6. Olgu $X_1, X_2, \ldots$ iid, $E(X_i) = 0$, $D(X_i) < \infty$. Kas rida $\sum_n \frac{X_n}{n}$ koondub?

7. Olgu $X_1, X_2, \ldots$ sõltumatud juhuslikud suurused. Tõestada väited
1) Kui $\exists K \geq 0$ nii, et $\sum_n \mathbf{P}(|X_n| > K) < \infty$ ja $\sum_n E(|X_n^K|) < \infty$, siis $\sum_n |X_n| < \infty$ p.k..
2) Kui $\sum_n |X_n| < \infty$ p.k., siis iga $\infty > K > 0$ korral

$$\sum_n \mathbf{P}(|X_n| > K) < \infty \text{ ja } \sum_n E(|X_n^K|) < \infty.$$

3) Kui read $\sum_n \mathbf{P}(|X_n| > K)$, $\sum_n E(X_n^K)$ ja $\sum_n D(X_n^K)$ koonduvad, ent $\sum_n E(|X_n^K|) = \infty$ mingi $K > 0$ korral, siis $\mathbf{P}($ rida $\sum_n X_n$ koondub tingimisi kuid mitte absoluutselt$)=1$.

8. Olgu $X_1, X_2, \ldots$ sõltumatud, $EX_n = 0$ iga n korral ja $\sum_n D(X_n) < \infty$. Ühe rea teoreemist teame, et piirväärtus $S := \sum_n X_n$ leidub. Fatou lemmat kasutades näidata, et $ES^2 \leq \sum_n D(X_n)$. Veenduda, et $S_n \xrightarrow{1} S$ ja $S_n \xrightarrow{2} S$. Järeldada, et $DS = \sum_n D(X_n)$.

9. Olgu $X_1, X_2, \ldots$ sõltumatud.
1) Tõestada, et kui $\sum_n X_n^2 < \infty$ p.k., siis $\sum_n X_n$ koondub parajasti siis kui koondub rida $\sum_n E(X_n I_{\{|X_n| \leq 1\}})$.
2) Tõestada, et kui $\sum_n X_n$ koondub p.k., siis $\sum_n X_n^2 < \infty$ p.k. parajasti siis, kui $\sum_n (E(|X_n| I_{\{|X_n| \leq 1\}}))^2 < \infty$.

11 Suurte arvude seadused

Olgu $X_1, X_2, \ldots$ tõenäosusruumil $(\Omega, \mathcal{F}, \mathbf{P})$ antud *sõltumatud* juhuslikud suurused, kusjuures $E|X_i| < \infty$. Tähistame

$$S_n := X_1 + \cdots + X_n, \quad n = 1, 2, \ldots.$$

Def 11.1 *Koondumist*

$$\frac{S_n - E(S_n)}{n} \xrightarrow{P} 0$$

nimetatakse **nõrgaks suurte arvude seaduseks** *(NSAS).*
Koondumist

$$\frac{S_n - E(S_n)}{n} \to 0 \quad p.k. \tag{11.1}$$

nimetatakse **tugevaks suurte arvude seaduseks** *(TSAS).*

Kui $EX_n = \mu$ (näiteks X_1, X_2, ... on sõltumatud ja sama jaotusega juhuslikud suurused), siis NSAS ja TSAS on vastavalt

$$\frac{X_1 + \cdots + X_n}{n} \xrightarrow{P} \mu \text{ ja } \frac{X_1 + \cdots + X_n}{n} \to \mu \text{ p.k.}$$

Suurte arvude seadused on tõenäosusteoorias ja matemaatilises statistikas keskel kohal. Teadaolevalt esimese suurte arvude seaduse tõestas juba 1698 a. Jakob Bernoulli.

Tihti põhinavad suurte arvude seaduste tõestused Cesaro ja Kroneckeri lemmal. Sõnastame need.

Lemma 11.1 (Cesaro lemma) *Olgu $\{a_n\}$ mittenegatiivsete reaalarvude jada, kusjuures $a_1 > 0$ ja $\sum_n a_n = \infty$. Tähistame $b_n := \sum_{i=1}^n a_i$. Olgu $x_n \to x$ suvaline koonduv jada. Siis*

$$\frac{1}{b_n} \sum_{i=1}^n a_i x_i \to x, \quad \text{kui} \quad n \to \infty.$$

Juhul, kui $a_n = 1$, saame

$$\frac{x_1 + \ldots + x_n}{n} \to x.$$

Lemma 11.2 (Kroneckeri lemma) *Olgu $\{b_n\}$ positiivsete reaalarvude jada, kusjuures $b_n \nearrow \infty$. Olgu $\{x_n\}$ selline reaalarvude jada, et $\sum_n \frac{x_n}{b_n}$ koondub. Siis*

$$\frac{1}{b_n} \sum_{i=1}^n x_i \to 0 \quad \text{kui} \quad n \to \infty.$$

Seega, kui

$$\sum_n \frac{x_n}{n}$$

koondub, kehtib

$$\frac{x_1 + \ldots + x_n}{n} \to 0.$$

Lemmade tõestusi võib leida: *Ŝirjajev* IV pt osa 3; *Williams* 12.6, 12.7.

11.1 NSAS

Alljärgnev on üks suhteliselt lihtsesti tõestatav NSAS.

Teoreem 11.2 *Olgu $X_1, X_2, \ldots$ lõpliku dispersiooniga sõltumatud juhuslikud suurused. Kui*

$$\sum_{n=1}^{\infty} \frac{D(X_n)}{n^2} < \infty,$$

siis

$$\frac{S_n - E(S_n)}{n} \xrightarrow{2} 0, \quad \frac{S_n - E(S_n)}{n} \xrightarrow{P} 0. \tag{11.2}$$

Tõestus. Tõestuseks piisab, kui näitame, et

$$E\big(\frac{S_n - E(S_n)}{n}\big)^2 \to 0,$$

kui $n \to \infty$ (miks?). Et juhuslikud suurused $X_1, X_2, \ldots$ on sõltumatud, siis

$$E\Big(\frac{S_n - E(S_n)}{n}\Big)^2 = \frac{1}{n^2} \sum_{i=1}^{n} D(X_i)$$

(ülesanne 1). Kroneckeri lemmast aga järeldub, et

$$\frac{1}{n^2} \sum_{i=1}^{n} D(X_i) \to 0,$$

kui $n \to \infty$. Teoreem on tõestatud. ■

Järeldus 11.1 *Olgu $X_1, X_2, \ldots$ lõpliku dispersiooniga sõltumatud juhuslikud suurused. Kui $\exists\ K < \infty$ nii, et $\forall n$ korral $D(X_n) \leq K$, siis kehtib (11.2).*

Järeldus 11.2 *Kui $X_1, X_2, \ldots$ on lõpliku dispersiooniga iid juhuslikud suurused ja $\mu := EX_n$, siis*

$$\frac{S_n}{n} \xrightarrow{2} \mu \text{ ja } \frac{S_n}{n} \xrightarrow{P} \mu.$$

Järelduste 11.1 ja 11.2 tõestus on ülesanne 2.

11.2 TSAS

Üldiselt koondumisest 2-keskmise mõttes ei järeldu koondumine p.k.. Selgub aga, et teoreemi 11.2 eeldustel kehtib ka TSAS.

Teoreem 11.3 (Kolmogorovi I teoreem). *Olgu* $X_1, X_2, \ldots$ *lõpliku dispersiooniga sõltumatud juhuslikud suurused. Kui*

$$\sum_{n=1}^{\infty} \frac{D(X_n)}{n^2} < \infty,$$

siis

$$\frac{S_n - E(S_n)}{n} \to 0, \quad \text{p.k.}$$

Tõestus. Et juhuslikud suurused $\{\frac{X_n - EX_n}{n}\}$ on sõltumatud ja

$$\sum_{n=1}^{\infty} D\Big(\frac{X_n - EX_n}{n}\Big) = \sum_{n=1}^{\infty} \frac{D(X_n)}{n^2} < \infty,$$

siis teoreemi 10.1 tõttu rida

$$\sum_{n=1}^{\infty} \frac{(X_n - EX_n)}{n}$$

koondub p.k.. Kroneckeri lemmast saame, et

$$\frac{1}{n}\sum_{i=1}^{n}(X_i - EX_i) = \frac{S_n - E(S_n)}{n} \to 0 \text{ p.k.}$$

■

Järeldus 11.3 *Järelduse 11.1 eeldustel kehtib TSAS.*

Järeldus 11.4 *Järelduse 11.2 eeldustel kehtib TSAS.*

11.3 IID

Olgu $X_1, X_2, \ldots$ *sõltumatud ja sama jaotusega* (iid) juhuslikud suurused, $X_n \sim P$. Järeldustest 11.1 ja 11.2 teame, et kui $D(X_n) < \infty$, st $X_n \in \mathcal{L}_2(\Omega, \mathcal{F}, \mathbf{P})$, siis kehtivad koondumised

$$\frac{S_n}{n} \to \mu \text{ p.k. (TSAS)}, \ \frac{S_n}{n} \xrightarrow{2} \mu,$$

millest järelduvad

$$\frac{S_n}{n} \xrightarrow{P} \mu$$

(NSAS) ja

$$\frac{S_n}{n} \Rightarrow \mu.$$

Alltoodud teoreem 11.4 üldistab järeldust 11.4 ja väidab, et TSAS kehtib ka juhul, kui $D(X_n) = \infty$. Loomulikult peab juhuslike suuruste keskväärtus olema defineeritud.

Tõestame kaks abitulemust.

Väide 11.1 *Kui X on mittenegatiivne juhuslik suurus, siis*

$$\sum_{n=1}^{\infty} \mathbf{P}(X \geq n) \leq EX \leq \sum_{n=1}^{\infty} \mathbf{P}(X \geq n) + 1. \tag{11.3}$$

Tõestus. Defineerime täisarvuliste väärtustega juhusliku suuruse $Y : \Omega \to \mathbb{N}$

$$Y := \sum_{n=1}^{\infty} I_{\{X \geq n\}}.$$

Definitsioonist tulenevalt, $\forall \omega \in \Omega, Y(\omega) \leq X(\omega) \leq Y(\omega)+1$, millest $EY \leq EX \leq EY+1$. Võrdusest

$$EY = \sum_{n=1}^{\infty} E(I_{\{X \geq n\}}) = \sum_{n=1}^{\infty} \mathbf{P}(X \geq n)$$

(miks?) järeldub (11.3). ■

Märkus: Funktsioon $\mathbf{P}(X \geq t)$ on mittekasvav. Seetõttu võrratus $\sum_{n=1}^{\infty} \mathbf{P}(X \geq n) \leq EX$ järeldub vahetult lemmast 8.1 (ülesanne 3).

Lemma 11.3 *Olgu $X_1, X_2, \dots$ iid juhuslikud suurused, kusjuures $E|X_n| < \infty$, $E(X_n) = \mu$. Defineerime*

$$Y_n := X_n I_{\{|X_n| \le n\}}, \quad n = 1, 2, \dots.$$

Kehtivad väited

1) $EY_n \to \mu$;

2) $\mathbf{P}(X_n = Y_n \,\text{ev}) = 1$;

3) $\sum_{n=1}^{\infty} \frac{DY_n}{n^2} < \infty$.

Tõestus. 1)
Paneme tähele, et

$$X_1 I_{\{|X_1| \le n\}} \to X_1 \text{ p.k.},$$

millest domineeritud koondumise tõttu (milline on domineeriv funktsioon?)

$$E(X_1 I_{\{|X_1| \le n\}}) \to EX_1 = \mu.$$

Et juhuslikud suurused X_n on sama jaotusega P, saame

$$EY_n = E(X_n I_{\{|X_n| \le n\}}) = \int_{\{|x| \le n\}} xP(dx) = E(X_1 I_{\{|X_1| \le n\}}) \to EX_1 = \mu.$$

2) Väitest 11.1 saame

$$\sum_{n=1}^{\infty} \mathbf{P}(X_n \ne Y_n) = \sum_{n=1}^{\infty} \mathbf{P}(|X_n| > n) = \sum_{n=1}^{\infty} \mathbf{P}(|X_1| > n) \le \sum_{n=1}^{\infty} \mathbf{P}(|X_1| \ge n) \le E|X_1| < \infty.$$

Seega $\mathbf{P}(Y_n \ne X_n \text{ i.o.}) = 0$ (BC1).

3) Paneme tähele, et iga $z \ge 0$ korral

$$\sum_{n=1}^{\infty} \frac{z^2 I_{\{z \le n\}}}{n^2} = z^2 f(z), \quad \text{kus} \quad f(z) = \sum_{n \ge \max\{z,1\}} \frac{1}{n^2}.$$

Et

$$\frac{1}{n^2} \le \frac{2}{n(n+1)} = 2(\frac{1}{n} - \frac{1}{n+1}),$$

siis

$$z^2 f(z) \le 2 \frac{z^2}{\max\{z, 1\}} = \begin{cases} 2z, & \text{kui } z \ge 1; \\ 2z^2 \le 2z, & \text{kui } z < 1. \end{cases}$$

Nüüd

$$\sum_{n=1}^{\infty} \frac{DY_n}{n^2} \le \sum_{n=1}^{\infty} \frac{EY_n^2}{n^2} = \sum_{n=1}^{\infty} \frac{E(X_n I_{\{|X_n| \le n\}})^2}{n^2} = \sum_{n=1}^{\infty} \frac{E(X_1 I_{\{|X_1| \le n\}})^2}{n^2}$$
$$= E\Big(\sum_{n=1}^{\infty} \frac{X_1^2 I_{\{|X_1| \le n\}}}{n^2}\Big) \le 2E|X_1| < \infty.$$

■

Lemma 11.3 abil tõestame kõige üldisema suurte arvude seaduse iid juhuslike suuruste korral.

Teoreem 11.4 (Kolmogorovi II teoreem). *Olgu $X_1, X_2, \ldots$ iid juhuslikud suurused, $E|X_n| < \infty$, $EX_n = \mu$. Siis*

$$\frac{S_n}{n} \to \mu \text{ p.k. ja ruumis } \mathcal{L}_1.$$

Tõestus. Olgu (sõltumatud) juhuslikud suurused Y_n defineeritud nii nagu lemmas 11.3. Väite **3)** tõttu

$$\sum_{n=1}^{\infty} \frac{DY_n}{n^2} < \infty.$$

Rakendades teoreemi 11.3, saame, et juhuslike suurusete $Y_1, Y_2, \ldots$ korral kehtib TSAS, millest

$$\frac{1}{n}\sum_{i=1}^{n}(Y_i - EY_i) \to 0 \text{ p.k..}$$

Et $EY_n \to \mu$, siis Cesaro lemma tõttu

$$\frac{1}{n}\sum_{i=1}^{n} EY_i \to \mu.$$

Seega, kui $n \to \infty$, siis

$$\frac{1}{n}\sum_{i=1}^{n} Y_i = \frac{1}{n}\sum_{i=1}^{n}(Y_i - EY_i) + \frac{1}{n}\sum_{i=1}^{n} EY_i \to \mu. \quad \text{p.k..}$$

Lõpuks kasutame asjaolu, et $\mathbf{P}(X_n = Y_n \text{ ev}) = 1$ ja järeldame, et

$$\frac{S_n}{n} \to \mu \text{ p.k.}$$

Tõestame koondumise ruumis $\mathcal{L}_1$. Erijuhul, kui X_n on lõpliku keskväärtusega mittenegatiivsed juhuslikud suurused, siis koondumine

$$\frac{S_n}{n} \xrightarrow{1} \mu$$

järeldub koondumisest $\frac{S_n}{n} \to \mu$, p.k. ja Sheffe lemmast (miks?).
Üldjuhul kehtib aga

$$\frac{\sum_{i=1}^n X_i^+}{n} \xrightarrow{1} EX_1^+ \text{ ja } \frac{\sum_{i=1}^n X_i^-}{n} \xrightarrow{1} EX_1^-.$$

Järelikult

$$E\Big|\frac{S_n}{n} - \mu\Big| \leq E\Big|\frac{\sum_{i=1}^n X_i^+}{n} - EX_1^+\Big| + E\Big|\frac{\sum_{i=1}^n X_i^-}{n} - EX_1^-\Big| \to 0.$$

■

Mitteintegreeruvad juhuslikud suurused. Vastavalt definitsioonile 11.1 on TSAS-e kehtimiseks tarvilik, et juhuslikel suurusetel oleks lõplik keskväärtus, st iga i korral $E|X_i| < \infty$. Ainult siis on ES_n iga n korral lõplik arv ja nii on definitsioonil 11.1 mõte. Sama kehtib ka nõrga suurte arvude seaduse korral. Seega kui $X_1, X_2, \dots$ on sellise jaotusega i.i.d. juhuslikud suurused, et ainult üks mittenegatiivsetest funktsioonidest X_i^+ või X_i^- omab lõplikku integraali, siis vastavalt definitsioonile TSAS ja NSAS kehtida ei saa, kuigi neil juhuslikel suurustel on keskväärtus defineeritud (see on kas $+\infty$ või $-\infty$). Küll aga võime küsida, kuidas käitub sellisel juhul valimi keskmine $\frac{S_n}{n}$. Järeldusest 11.4 selgub, et sellisel juhul valimi keskmine koondub *p.k.* üldkogumi lõpmatuks keskväärtuseks, st $\frac{S_n}{n}$ käitub nii nagu suurte arvude seaduse korral. Seega kui $EX_1 = \infty$ või $EX_i = -\infty$, siis TSAS ja NSAS ei kehti, kuid $\frac{S_n}{n} \to EX_1$ p.k. Siin on täielik analoog integreeruvusega: kui $EX = \infty$ või $EX = -\infty$, siis X pole integreeruv, kuid tal on integraal.

Järeldus 11.5 *Olgu $X_1, X_2, \dots$ iid juhuslikud suurused, $EX_n^- < \infty$ ja $EX_n^+ = \infty$ (või $EX_n^+ < \infty$ ja $EX_n^- = \infty$). Siis*

$$\frac{S_n}{n} \to \infty \quad \text{p.k.} \quad \Big(\frac{S_n}{n} \to -\infty \quad \text{p.k.}\Big).$$

Tõestus. Kui $\mathbf{P}(X_1 = \infty) > 0$, siis $\sum_i \mathbf{P}(X_i = \infty) = \infty$, millest BC2:

$$\mathbf{P}(X_i = \infty \text{ i.o.}) = 1.$$

Seega, kui juhuslikud suurused võivad positiivse tõenäosusega võtta väärtusi $+\infty$, siis toimub see varem või hiljem p.k. Kui aga vähemalt üks juhuslikest suurustest $X_1, \dots, X_n$ on $+\infty$ (ja ükski neist pole $-\infty$), on seda ka $\frac{S_n}{n}$ ja väide kehtib.

Seega edaspidi vaateleme olukorda, kus

$$\mathbf{P}(X_1 = \infty) = 0. \tag{11.4}$$

Olgu $K < \infty$ ja defineerime

$$Y_n^K := X_n I_{\{X_n \leq K\}}.$$

Juhuslikud suurused Y_n^K, $n = 1, 2 \dots$ on iid (miks?) ning

$$E|Y_n^K| \leq K + EX_n^- < \infty.$$

Teoreemi 11.4 põhjal kehtib TSAS. Seda arvestades saame

$$\liminf_n \frac{S_n}{n} \geq \liminf_n \frac{1}{n} \sum_{i=1}^n Y_i^K = E(X_1 I_{\{X_1 \leq K\}}) \quad \text{p.k.}.$$

Monotoonse koondumise põhjal

$$E(X_1 I_{\{X_1 \leq K\}}) \nearrow \infty,$$

kui $K \nearrow \infty$. Siin kasutasime lisaeeldust (11.4) ja eeldust $EX^+ = \infty$. Nüüd on lihtne veenduda (ülesanne 4), et

$$\liminf_n \frac{S_n}{n} = \infty \text{ p.k.}$$

■

Järeldus 11.5 ütleb, et juhul, kui $EX = \infty$ või $EX = -\infty$, siis valimi keskmine koondub ikkagi keskväärtusesks EX_1. Pane tähele, et siis

$$\limsup_n |\frac{S_n}{n}| = \infty \text{ p.k.},$$

s.t. jada $\{\frac{S_n}{n}\}$ on p.k. tõkestamata. Viimane omadus kehtib ka juhul, kui juhuslike suuruste X_n keskväärtus pole defineeritud, st $EX^+ = EX^- = \infty$.

Järeldus 11.6 *Olgu $X_1, X_2, \ldots$ iid juhuslikud suurused, $E|X_n| = \infty$. Siis*

$$\limsup_n |\frac{S_n}{n}| = \infty \quad \text{p.k.}.$$

Tõestus. Olgu $K < \infty$. Juhuslikud suurused $K^{-1}X_n$, $n = 1, 2, \ldots$ on sõltumatud, $E|K^{-1}X_n| = \infty$, millest väite 11.1 tõttu

$$\sum_{n=1}^{\infty} \mathbf{P}\Big(\frac{|X_n|}{n} \geq K\Big) = \infty.$$

Kasutades BC2, saame

$$\mathbf{P}\big(\limsup_n \frac{|X_n|}{n} \geq K\big) \geq \mathbf{P}\big(\frac{|X_n|}{n} \geq K \text{ i.o.}\big) = 1.$$

Järelikult $\limsup_n \frac{|X_n|}{n} = \infty$ p.k.(ül. 5). Seostest

$$\frac{|S_n|}{n} \geq \frac{|X_n|}{n} - \frac{|S_{n-1}|}{(n-1)} \frac{(n-1)}{n}$$

järeldub, et $\limsup_n \frac{|S_n|}{n} = \infty$ p.k. (ülesanne 5.) ■

Viimasest järeldusest saame muuhulgas ka, et iid juhuslike suuruste korral kehtib koondumine

$$\frac{S_n}{n} \to c \in \mathbb{R} \text{ p.k.}$$

parajasti siis, kui $E|X_n| < \infty$ ka sellisel juhul $c = \mu$, st kehtiv koondumine on TSAS. Sõnastame selle järeldusena.

Järeldus 11.7 *Olgu $X_1, X_2, \ldots$ iid juhuslikud suurused. Siis*

$$\frac{S_n}{n} \to c \in \mathbb{R} \text{ p.k.}$$

parajasti siis, kui $E|X_n| < \infty$ ja sellisel juhul $c = \mu := EX_n$.

Tõestus. Piisavus järeldub teoreemist 11.4.
Oletame, et $E|X_n| = \infty$. Järelduse 11.6 tõttu $\limsup_n \frac{|S_n|}{n} = \infty$ p.k. mis on vastuolus eeldusega $\frac{S_n}{n} \to c$ p.k.. Järelikult $E|X_n| < \infty$ ja $\frac{S_n}{n} \to \mu$ p.k.. See aga tähendab, et $\mu = c$.
■

Järeldus 11.7 annab seega piisava ja tarviliku tingimuse selleks, et jada $\frac{S_n}{n}$ koonduks lõplikuks keskväärtuseks p.k. Paneme tähele, et esialgu ehk vaid formaalsusena paistnud lõpliku keskväärtuse olemasoluks piisav ja tarvilik tingimus $E|X| < \infty$ omab nüüd märksa laiema sisu. Tõepoolest, kui $E|X| = \infty$, siis $\frac{S_n}{n}$ ei koondu. Järelikult ei saa sellisel juhul mingit "teoreetilist keskmist" olla ka sisuliselt. Olgu näiteks X jaotus sümmeetriline 0 suhtes, st X ja $-X$ on sama jaotusega. Kui $E|X| < \infty$, siis sümmeetria tõttu $EX = 0$ ja TSAS: $\frac{S_n}{n} \to 0$ p.k.. Kui aga sümmeetrilise jaotusega X on selline, et $E|X| = \infty$ (selline jaotus on näiteks Cauchy jaotus), siis võib tekkida kahtlus, et vaatamata Lebesgue'i integraali formaalsele mitteeksisteerimisele jaotuse "sisuline keskmine" on ikkagi 0 ja $\frac{S_n}{n} \to 0$ p.k. Järeldus 11.6 välistab selle ning vädab, et $\limsup_n |\frac{S_n}{n}| = \infty$ p.k.. Kuidas aga sellisel sümmeetrilisel juhul käitub jada $\frac{S_n}{n}$? Sümmeetria tõttu on sellele küsimusele lihtne vastata. Kõigepealt paneme tähele, et sündmused $\{\limsup_n \frac{S_n}{n} = \infty\}$ ja $\{\liminf_n \frac{S_n}{n} = -\infty\}$ on mõlemad jääk-σ-algebra elemendid, mistõttu Kolmogorovi 0-1 seaduse tõttu nende tõenäosused on kas 0 või 1. Sümmeetria tõttu on selge, et nimetatud tõenäosused on võrdsed. Kui nad mõlemad oleksid 0, siis kehtiks

$$\mathbf{P}(\limsup_n \frac{S_n}{n} < \infty) = 1,$$

mis oleks vastuolus järeldusega 11.7. Seega peavad need tõenäosused võrduma ühega. Seega sümmeetrilise jaotuse korral

$$\limsup_n \frac{S_n}{n} = \infty \text{ p.k ja } \liminf_n \frac{S_n}{n} = -\infty \text{ p.k.},$$

mis sisuliselt tähendab jada $\frac{S_n}{n}$ lõpmatut fluktueerimist järjest suureneva amplituudiga.

11.4 Empiirilised mõõdud

Meeldetuletus: Olgu $\nu = \sum_i p_i \delta_{x_i}$ reaalteljel antud diskreetne mõõt. Seega reaalarvud $x_i, x_2, \ldots$ on aatomid ja $p_1, p_2, \ldots$ nende tõenäosused. Tuletame meelde, et hulga A mõõt avaldub

$$\nu(A) = \sum_{i: x_i \in A} p_i$$

ning suvalise funktsiooni $f : \mathbb{R} \to \mathbb{R}$ korral integraal mõõdu ν järgi on (valem (7.29))

$$\int f d\nu = \sum_i f(x_i) p_i. \tag{11.5}$$

Empiirilise mõõdu definitsioon. Olgu $X_1, X_2, \ldots$ iid juhuslikud suurused, $X_n \sim P$. Iga $n = 1, 2, \ldots$ ja $\omega \in \Omega$ korral vaatleme n esimese juhusliku suuruse realisatsiooni $X_1(\omega), \ldots, X_n(\omega)$ (valim) ja defineerime ruumil $(\mathbb{R}, \mathcal{B})$ tõenäosusmõõdu P_n^ω järgmiselt

$$P_n^\omega := \frac{1}{n} \sum_{i=1}^n \delta_{X_i(\omega)}.$$

Mõõtu P_n nimetatakse **empiiriliseks mõõduks**. Seega $\forall A \in \mathcal{B}$ korral

$$P_n^\omega(A) = \frac{1}{n} \sum_{i=1}^n I_A(X_i(\omega)).$$

Definitsioonist saame, et hulga A empiiriline mõõt $P_n^\omega(A)$ on võrdne sinna hulka kuuluvate valimi elementide $X_i(\omega)$ arvu ja valimi mahu n suhtega. Paneme veel tähele, et suvalise funktsiooni f korral (valem (11.5)):

$$\int f dP_n^\omega = \frac{1}{n} \sum_{i=1}^n f(X_i(\omega)).$$

Koondumisest. Fikseeritud ω korral on P_n^ω fikseeritud diskreetne tõenäosusmõõt ning seega iga A korral $P_n^\omega(A)$ mingi arv. Vaadeldes aga valimit juhuslikuna saame, et iga A korral on $P_n(A)$ juhuslik suurus:

$$P_n(A) = \frac{1}{n} \sum_{i=1}^n I_A(X_i).$$

Meid huvitab selle juhusliku suuruse käitumine valimi mahu n kasvades. Et $I_A(X_1), I_A(X_2), \ldots$ on iid juhuslikud suurused keskväärtusega

$$EI_A(X_1) = \int_A P(dx) = P(A),$$

siis teoreemist 11.4 järeldub

$$P_n(A) = \frac{1}{n} \sum_{i=1}^n I_A(X_i) \to P(A) \quad \text{p.k.}$$

ehk $\forall A \in \mathcal{B}$ korral

$$\mathbf{P}\{\omega | P_n^\omega(A) \to P(A)\} = 1.$$

Olgu nüüd F_n^ω empiirilisele mõõdule P_n^ω vastav jaotusfunktsioon, F olgu P jaotusfunktsioon. Võttes hulgaks $A = (-\infty, x]$, saame ülaltoodud koondumisest, et $\forall x \in \mathbb{R}$ korral kehtib $F_n(x) \to F(x)$ p.k. ehk

$$\mathbf{P}(F_n(x) \to F(x)) = 1. \tag{11.6}$$

Üldiselt, olgu $f : \mathbb{R} \to \mathbb{R}$ selline Boreli funktsioon, et $\int |f| dP = E|f(X_1)| < \infty$. Sellisel juhul $f(X_1), f(X_2), \ldots$ on iid juhuslikud suurused, kusjuures $E|f(X_1)| < \infty$. Teoreemist 11.4 järeldub nüüd, et

$$\int f dP_n = \frac{1}{n} \sum_{i=1}^n f(X_i) \to Ef(X_1) = \int f dP. \quad \text{p.k..} \tag{11.7}$$

Loomulikult järeldub koondumisest (11.7) nii koondumine $P_n(A) \to P(A)$ p.k. (võttes $f = I_A$) kui ka $F_n(x) \to F(x)$ p.k.

Glivenko-Cantelli teoreem. Uurime lähemalt koondumist (11.6) ehk $F_n(x) \to F(x)$ p.k. See koondumine tähendab, et $\mathbf{P}(\Omega(x)) = 1$, kus

$$\Omega(x) = \{\omega | F_n^\omega(x) \to F(x)\}.$$

Et hulk $\Omega(x)$ sõltub argumendist x, ei saa me vahetult järeldada (miks?), et

$$\mathbf{P}\{\omega | F_n^\omega(x) \to F(x), \quad \forall x \in \mathbb{R}\} = 1. \tag{11.8}$$

Selgub aga, et (11.8) kehtib. Olgu $\mathbf{Q}$ kõigi ratsionaalarvude hulk ja J olgu funktsiooni F katkevuspunktide hulk. Et hulk $\mathbf{Q} \cup J$ on ülimalt loenduv (miks?), siis $\mathbf{P}(\Omega^*) = 1$, kus

$$\Omega^* = \cap_{\mathbf{Q} \cup J} \Omega(x).$$

Olgu $\omega \in \Omega^*$ suvaline. Pole raske veenduda, et $\forall x \in \mathbb{R}$ korral kehtib $F_n^\omega(x) \to F(x)$. Tõepoolest: olgu x funktsiooni F pidevuspunkt. Siis $\forall \epsilon > 0$ korral leiduvad $z, y \in \mathbf{Q}$ nii, et $z \leq x \leq y$ ja $F(y) - F(z) \leq \epsilon$. Et F_n ja F on mittekahanevad, saame

$$\limsup_n F_n^\omega(x) \leq \limsup_n F_n^\omega(y) = F(y), \quad \liminf_n F_n^\omega(x) \geq \liminf_n F_n^\omega(z) = F(z)$$

ehk

$$\limsup F_n^\omega(x) - \liminf F_n^\omega(x) \leq \epsilon.$$

Et ϵ oli suvaline, saame

$$\limsup_n F_n^\omega(x) = \liminf_n F_n^\omega(x) = \lim_n F_n^\omega(x).$$

Samas ülaltoodud argumendist järeldub, et iga $\epsilon > 0$ korral

$$\limsup_n F_n^\omega(x) \leq \limsup_n F_n^\omega(y) = F(y) \leq F(x) + \epsilon$$
$$\liminf_n F_n^\omega(x) \geq \liminf_n F_n^\omega(z) = F(z) \geq F(x) - \epsilon.$$

Seega $\lim_n F_n^\omega(x) = F(x)$ ja see ongi (11.8).
Järgmine teoreem näitab, et koondumine (11.8) on ühtlane.

Teoreem 11.5 (Glivenko-Cantelli teoreem) *Olgu F suvaline jaotusfunktsioon, F_n vastav empiiriline jaotusfunktsioon. Koondumine (11.8) on ühtlane, s.t.*

$$\mathbf{P}\{\omega | \sup_{x\in\mathbb{R}} |F_n^\omega(x) - F(x)| \to 0\} = 1. \tag{11.9}$$

Tõestus. Tõestus koosneb kahest osast. Esimeses osas näitame, et (11.9) kehtib, kui $F = G$, kus G on ühtlase jaotuse jaotusfunktsioon, st

$$G(x) = \max\{0, \min\{x, 1\}\}.$$

Olgu G_n vastav empiiriline jaotusfunktsioon. Näitame, et

$$\mathbf{P}\{\omega | \sup_{x\in\mathbb{R}} |G_n^\omega(x) - G(x)| \to 0\} = 1. \tag{11.10}$$

Eelkõige paneme tähele, et iga n korral on supremum maksimum. Tõepoolest, et $G(X_k) = X_k$, siis

$$\sup_{x\in\mathbb{R}} |G_n(x) - G(x)| = \max\big\{ \max_{k=1,\dots,n} |n^{-1}\sum_{i=1}^n I_{(-\infty,X_k]}(X_i) - X_k|, \max_{k=1,\dots,n} |n^{-1}\sum_{i=1}^n I_{(-\infty,X_k)}(X_i) - X_k| \big\}.$$

Seega on

$$\omega \mapsto \sup_x |G_n^\omega(x) - G(x)|$$

mõõtuv.

Argument (11.10) tõestamiseks on sisuliselt meile juba tuttav. Olgu $\epsilon > 0$ suvaline. Näitame, et leidub hulk $\Omega(\epsilon)$ nii, et $\mathbf{P}(\Omega(\epsilon)) = 1$ ja iga $\omega \in \Omega(\epsilon)$ korral $\exists n_o$ (võib sõltuda ω-st) nii, et

$$\sup_x |G_n^\omega(x) - G(x)| \leq \epsilon, \quad \text{kui } n \geq n_o. \tag{11.11}$$

Sellest järeldub (11.10) (kuidas?).

Olgu m nii suur, et $m^{-1} \leq \frac{\epsilon}{2}$. Defineerime

$$E := \{\frac{k}{m}, k = 0, 1, 2, \dots, m\} \quad \text{ja} \quad \Omega(\epsilon) := \{\omega | G_n^\omega(x) \to G(x), \quad \forall x \in E\}.$$

Et E on lõplik, siis $\mathbf{P}(\Omega(\epsilon)) = 1$. Olgu $\omega \in \Omega(\epsilon)$. Et E on lõplik, siis $\exists n_o$ nii suur, et iga $n > n_o$ korral $|G_n^\omega(x) - G(x)| \leq \frac{\epsilon}{2}$ iga $x \in E$ korral. Olgu $x \in \mathbb{R}$ suvaline. Näitame, et

$$|G_n^\omega(x) - G(x)| \leq \epsilon, \quad \text{kui } n > n_o. \tag{11.12}$$

Sellest järeldub (11.11).

Leiduvad $z, y \in E$ nii, et $z \leq x \leq y$ ja $y - z = G(y) - G(z) \leq m^{-1}$. Seega

$$G_n^\omega(x) \geq G_n^\omega(z) \geq G(z) - \frac{\epsilon}{2} = z - \frac{\epsilon}{2} \geq x - \epsilon = G(x) - \epsilon$$

ning analoogiliselt $G_n(x) \le G(x) + \epsilon$, millest saame (11.12).

Koondumise (11.9) tõestamiseks üldisema jaotuse korral kasutame Skorohodi esitust. Olgu F suvaline ja

$$X_F(t) := \inf\{x : F(x) \ge t\}.$$

Teame, et $X_F(t)$ on ruumil $((0,1), \mathcal{B}, \mathrm{Leb})$ defineeritud juhuslik suurus, mille jaotus on F. Seega, kui $Y \sim G$, siis $X_F(Y) \sim F$, sest kehtib

$$\{t : X_F(t) \le x\} = \{t : t \le F(x)\}$$

(vaata Skorohodi esitust), millest

$$\mathbf{P}\{\omega : X_F(Y(\omega)) \le x\} = \mathbf{P}\{\omega : Y(\omega) \le F(x)\} = F(x).$$

Järelikult, kui $Y_1, Y_2, \ldots$ on i.i.d ühtlase jaotusega $U(0,1)$, siis $X_F(Y_1), X_F(Y_2), \ldots$ on jaotusega F i.i.d juhuslikud suurused. Olgu G_n juhuslike suuruste $Y_1, Y_2, \ldots, Y_n$ põhjal konstrueeritud empiiriline jaotusfunktsioon G_n. Seega

$$\begin{aligned} G_n^\omega(F(x)) &= n^{-1} \sum_{i=1}^n I_{(-\infty, F(x)]}(Y_i(\omega)) = n^{-1} \sum_{i=1}^n I_{\{\omega : Y_i(\omega) \le F(x)\}}(\omega) \\ &= n^{-1} \sum_{i=1}^n I_{\{\omega : X_F(Y_i(\omega)) \le x\}}(\omega) = n^{-1} \sum_{i=1}^n I_{(-\infty, x]}(X_F(Y_i(\omega))). \end{aligned}$$

Seega iga x korral

$$G_n(F(x)) = n^{-1} \sum_{i=1}^n I_{(-\infty, x]}(X_F(Y_i))$$

ehk funktsioon $x \mapsto G_n^\omega(F(x))$ on jaotuse F empiiriline jaotusfunktsioon. Samas iga $\omega \in \Omega$ korral kehtib

$$\sup_x |G_n^\omega(F(x)) - G(F(x))| \le \sup_z |G_n^\omega(z) - G(z)|, \tag{11.13}$$

kusjuures võrratus on võrdus, kui F on pidev. Sellest järeldub, et

$$\sup_x |G_n(F(x)) - G(F(x))| = \sup_x |F_n(x) - F(x)| \to 0 \quad \text{p.k.}$$

siis, kui $\sup_z |G_n(z) - G(z)| \to 0$ p.k.. Selle tõestasime aga eespool. ∎

Dvoretzki-Kiefer-Wolfowitzi võrratus. Peaaegu kindlasti koondumisest $F_n(x) \to F(x)$ p.k. järeldub tõenäosuse järgi koondumine. Seega

$$\mathbf{P}\big(|F_n(x) - F(x)| \ge \epsilon\big) \to 0.$$

Viimase koondumise kiirust on võimalik hinnata Höffdingi võrratusest (ülesanne 9.)

$$\mathbf{P}\big(F_n(x) - F(x) > \epsilon\big) \le e^{-2n\epsilon^2}, \quad \mathbf{P}\big(|F_n(x) - F(x)| \ge \epsilon\big) \le 2e^{-2n\epsilon^2}. \tag{11.14}$$

Samuti on selge, et koondumisest (11.14) järeldub $F_n(x) \to F(x)$ p.k.

Samuti järeldub peaaegu kindlasti koondumisest $\sup_x |F_n(x) - F(x)| \to 0$ tõenäosuse järgi koondumine:

$$\mathbf{P}\big(\sup_x |F_n(x) - F(x)| > \epsilon\big) \to 0. \tag{11.15}$$

Glivenko-Cantelli teoreemi tõestusest järeldub, et viimane koondumine on *ühtlane üle jaotuste F*. Tõepoolest, võrratuse (11.13) parem pool ei sõltu jaotusest F. Seega iga $\epsilon > 0$ ja iga jaotuse F korral kehtib

$$\mathbf{P}\big(\sup_x |F_n(x)-F(x)| > \epsilon\big) = \mathbf{P}\big(\sup_x |G_n(F(x))-F(x)| > \epsilon\big) \le \mathbf{P}\big(\sup_z |G_n(z)-G(z)| > \epsilon\big), \tag{11.16}$$

Analoogiliselt saame ühepoolse võrratuse

$$\mathbf{P}\big(\sup_x \big(F_n(x) - F(x)\big) > \epsilon\big) \le \mathbf{P}\big(\sup_z \big(G_n(z) - G(z)\big) > \epsilon\big), \tag{11.17}$$

kusjuures pideva F korral on mõlemad võrratused võrdused. Seosest (11.15) järeldub, et võrratuste (11.16) ja (11.17) paremad pooled koonduvad nulliks. Kas aga selle koondumise kiirust saab hinnata? Tuletame meelde võrratused (11.14): iga x ja $\epsilon > 0$ korral

$$\mathbf{P}\big(F_n(x) - F(x) > \epsilon\big) \le e^{-2n\epsilon^2}, \quad \mathbf{P}\big(|F_n(x) - F(x)| > \epsilon\big) \le 2e^{-2n\epsilon^2}.$$

Selgub, et piisavalt suure n korral kehtivad need võrratused ka supreemumi korral.

Teoreem 11.6 (P. Massart, 1990) *Olgu F_n jaotuse F empiiriline jaotusfunktsioon. Kui ϵn on nii suur, et $e^{-2n\epsilon^2} \le \frac{1}{2}$, siis*

$$\mathbf{P}\big(\sup_x \big(F_n(x) - F(x)\big) > \epsilon\big) \le e^{-2n\epsilon^2}. \tag{11.18}$$

Iga n ja $\epsilon > 0$ korral kehtib aga

$$\mathbf{P}\big(\sup_z |F_n(z) - F(z)| > \epsilon\big) \le 2e^{-2n\epsilon^2}. \tag{11.19}$$

Pane tähele, et võrratusest (11.18) järeldub vahetult võrratus (11.19). Võrratusi (11.19) ja (11.18) nimetatakse **Dvoretzki-Kiefer-Wolfowitzi (DKW) võrratusteks**. Seoste (11.16) ja (11.17) piisab kui nad on tõestatud vaid ühtlase jaotuse korral. On selge, et järelduse 9.1 tõttu järeldub hinnangust (11.19) Glivenko-Cantelli teoreem ehk koondumine (11.9).

Kolmogorov-Smirnovi test. DKW võrratused võimaldavad konstrueerida statistiku kontrollimaks hüpoteesi, kas empiiriline jaotus F_n pärineb jaotusest F. Tõepoolest, täpselt samamoodi kui peatükis 8.7.4, saame *iga n korral* ühepoolsed usalduspiirid statistikule

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

järgmiselt: iga $\alpha > 0$ korral

$$\mathbf{P}\Big(\sqrt{n}D_n > \sqrt{(1/2)\ln(2/\alpha)}\Big) \leq \alpha. \tag{11.20}$$

(ülesanne 10). Seega, kui statistik $\sqrt{n}D_n$ on suurem kui $\sqrt{(1/2)\ln(2/\alpha)}$, lükkame hüpoteesi (et F_n pärineb jaotusest F) nivool α ümber. Analoogiliselt saab leida usalduspiiri ka ühepoolsele statistikule (ülesanne 10)

$$D_n^+ := \sup_x \big(F_n(x) - F(x)\big).$$

Asümptootilised usalduspiirid. Aastal 1933 tõestas A. Kolmogorov, et

$$\mathbf{P}\big(\sqrt{n}\sup_x |G_n(x) - G(x)| > \epsilon\big) \to 2\sum_{k=1}^{\infty}(-1)^{k-1}e^{-2k^2\epsilon^2}.$$

Seega piisavalt suure n korral

$$\mathbf{P}(\sqrt{n}D_n > \epsilon\big) \leq \mathbf{P}\big(\sqrt{n}\sup_x |G_n(x) - G(x)| > \epsilon\big) \approx 2\sum_{k=1}^{\infty}(-1)^{k-1}e^{-2k^2\epsilon^2},$$

kusjuures pideva F korral on võrratus võrdus. Toodud seosest saame *asümptootilised* usalduspiirid: olgu $\epsilon(\alpha)$ võrrandi

$$2\sum_{k=1}^{\infty}(-1)^{k-1}e^{-2k^2\epsilon^2} = \alpha \tag{11.21}$$

lahend. Siis $\mathbf{P}(\sqrt{n}D_n > \epsilon(\alpha)) \approx \alpha$ ehk kui $\sqrt{n}D_n > \epsilon(\alpha)$ võime (piisavalt suure n korral) hüpoteesi (et F_n pärineb jaotusest F) nivool α ümber lükata. On selge, et iga n ja α korral $\epsilon(\alpha) < \sqrt{(1/2)\ln(2/\alpha)}$ (miks?), kuid piiri $\epsilon(\alpha)$ kasutamine eeldab piisavalt suurt valimimahtu n. Et valemis (11.21) oleva rea summa pole analüütiliselt teada pole võrrandi (11.21) lahendamine lihtne, kuid 1948. aastal publitseeris N. Smirnov tabeli, mis tegi saadud testi kasutamise praktiliselt võimalikuks.

Asümptootilise testi ühepoolse statistiku D_n^+ kaudu saame järgmisest koondumisest (N. Smirnov, 1944):

$$\mathbf{P}\big(\sqrt{n}\sup_x(G_n(x) - G(x)) > \epsilon\big) \to e^{-2\epsilon^2}. \tag{11.22}$$

Seega piisavalt suure n korral (ülesanne 10)

$$\mathbf{P}\big(\sqrt{n}D_n^+ > \sqrt{1/2\ln(1/\alpha)}\big) \le \mathbf{P}\big(\sqrt{n}\sup_x(G_n(x)-G(x)) > \sqrt{1/2\ln(1/\alpha)}\big) \approx \alpha, \quad (11.23)$$

kusjuures pideva F korral on võrratus võrdus. Pane tähele, et antud juhul asümptootilised usalduspiirid langevad kokku nendega, mis saadud DKW võrratuse abil. Statistikuid D_n ja D_n^+ nimetatakse **Kolmogorov-Smirnovi statistikuteks** ja nende abil saadud teste **Kolmogorov-Smirnovi testideks**.

Loglog-seadus. Võttes

$$\alpha = \frac{2}{\ln n},$$

saame seosest (11.20)

$$\mathbf{P}\Big(\sqrt{\frac{n}{2\ln\ln n}}D_n > \frac{1}{2}\Big) = \mathbf{P}\Big(\sqrt{n}D_n > \sqrt{(1/2)\ln\ln n}\Big) \le \frac{2}{\ln n} \to 0.$$

Et rida $\sum_n \frac{2}{\ln n}$ ei koondu, ei saa me vahetult järeldada, et

$$\mathbf{P}\Big(\limsup_n \sqrt{\frac{n}{2\ln\ln n}}D_n \le \frac{1}{2}\Big) = 1. \quad (11.24)$$

Saab näidata, et tegelikult kehtib ka (11.24). Seega iga $a > \frac{1}{2}$ korral $\sqrt{\frac{n}{2\ln\ln n}}D_n \le a$ mingist kohast alates, p.k.. Selgub, et kehtib ka teistpidi tulemus, mis väidab, et empiiriline jaotisfunktsioon ei saa koonduda liiga kiiresti (siin F ei ole aatomi jaotusfunktsioon):

$$\liminf_n \sqrt{2\ln\ln n}D_n = \frac{\pi}{2}, \quad \text{p.k..}$$

Gliveko-Cantelli teoreem ja selle üldistused on tähtsal kohal matemaatilise statistika mitmetes harudes. Nii põhineb suuresti just Gliveko-Cantelli teoreemi üldistusel n.n. *Vapnik-Cervonenkise* teooria, mis omakorda on *masinõppe (statistical learning)* matemaatiline baas.
Juhuslikku funktsiooni

$$x \mapsto \sqrt{n}(F_n(x) - F(x))$$

nimetatakse **empiiriliseks protsessiks**. Empiirilistest protsessidest on (viimastel aastakümnetel) saanud väga oluline matemaatiline vahend asümptootilises statistikas. Nende eduka kasutamise teeb aga võimalikuks *tõenäosusmõõtude nõrga koondumise* teooria. Järgnevas tutvume selle teooria põhimõistetega.

Kirjandus. Suurte arvude seaduste kohta loe: *Ŝirjajev* IV osa 3; *Williams* 12.6-12.10; *Grimmett, Stirzaker* 7.4.

11.5 Ülesanded

1. Olgu $X_1, X_2, \ldots$ sõltumatud juhuslikud suurused. Tõestada, et $\forall n$ korral

$$D\Big(\frac{S_n}{n}\Big) = \frac{1}{n^2}\sum_{i=1}^{n} D(X_i).$$

2. Tõestada järeldused 11.1 ja 11.2.

3. Olgu X mittenegatiivne juhuslik suurus. Lemmat 8.1 kasutades tõestada, et

$$\sum_{n=1}^{\infty} \mathbf{P}(X \geq n) \leq EX.$$

4. Tõestada, et kui $EX^+ = \infty$, $EX^- < \infty$ ja kui $\forall K$ korral kehtib

$$\liminf_n \frac{S_n}{n} \geq E(X_1 I_{\{X_1 \leq K\}}) \quad \text{p.k.},$$

siis $\liminf_n \frac{S_n}{n} = \infty$ p.k..

5. Tõestada järeldus 11.6.

6. Olgu $X_1, X_2, \ldots$ iid juhuslikud suurused, $X_n \sim P$. Olgu $\{P_n\}$ vastavate empiiriliste mõõtude jada. Näidata, et $\forall A \in \mathcal{B}$ korral $P_n(A) \to P(A)$ p.k.

7. Olgu $X_1, X_2, \ldots$ lõplike dispersioonidega sõltumatud juhuslikud suurused, $\{b_n\}$ positiivsete reaalarvude jada, kusjuures $b_n \nearrow \infty$. Tõestada, et kui $\sum_{n=1}^{\infty} \frac{D(X_n)}{b_n^2} < \infty$, siis kehtivad koondumised

$$\frac{S_n - E(S_n)}{b_n} \xrightarrow{2} 0 \quad \frac{S_n - E(S_n)}{b_n} \to 0 \quad \text{p.k.}$$

8. Olgu $X_1, X_2, \ldots$ sõltumatud juhuslikud suurused, juhusliku suuruse X_n jaotus olgu vastavalt:

1) $\mathbf{P}(X_n = -\sqrt{n}) = \mathbf{P}(X_n = \sqrt{n}) = \frac{1}{2n}$, $\mathbf{P}(X_n = 0) = 1 - \frac{1}{n}$;
2) $\mathbf{P}(X_n = -n) = \mathbf{P}(X_n = n) = \frac{1}{2n^2}$, $\mathbf{P}(X_n = 0) = 1 - \frac{1}{n^2}$;
3) $\mathbf{P}(X_n = -2^n) = \mathbf{P}(X_n = 2^n) = \frac{1}{2^{2n+1}}$, $\mathbf{P}(X_n = 0) = 1 - \frac{1}{2^{2n}}$;
4) $\mathbf{P}(X_n = -2^n) = \mathbf{P}(X_n = 2^n) = \frac{1}{2}$;
5) $\mathbf{P}(X_n = -\phi(n)) = \mathbf{P}(X_n = \phi(n)) = \frac{1}{\psi(n)}$, $\mathbf{P}(X_n = 0) = 1 - \frac{2}{\psi(n)}$, kusjuures $\psi(n) \geq 2$ ja $\frac{\phi^2(n)}{\psi(n)} < K < \infty$

Kas kehtivad koondumised: $\frac{S_n}{n} \xrightarrow{2} 0$, TSAS, TSAS?

9. Tõestada (11.14).

10.

1. Tõesta (11.20).

2. Võrratuse (11.18) abil leida täpne α-usalduspiir statistikule D_n^+. (Ära unusta, et (11.18) kehtib vaid siis, kui $n\epsilon$ on piisavalt suur).

3. Tõesta (11.23).

11.

1. Olgu $f \in C[0,1]$, st lõigul $[0,1]$ määratud pidev funktsioon. Olgu iga n korral $B_n(f)$ samuti lõigul $[0,1]$ määratud pidev funktsioon, mis on defineeritud järgmiselt:

$$B_n(f)(x) := \sum_{k=0}^{n} f\Big(\frac{k}{n}\Big)\binom{n}{k} x^k(1-x)^{n-k}.$$

Siin $0^0 \equiv 1$. Tõesta, et iga x korral

$$\lim_n B_n(f)(x) = f(x).$$

Tõesta, et kui f on kumer, siis

$$B_n(f) \geq f.$$

2. Olgu f lõigul $(0,\infty)$ antud pidev tõkestatud funktsioon ning olgu iga n korral

$$S_n(f)(x) := \sum_{k=0}^{\infty} f\Big(\frac{k}{n}\Big) e^{-nx}\frac{(nx)^k}{k!}, \quad x > 0.$$

Tõesta, et iga x korral

$$\lim_n S_n(f)(x) = f(x).$$

Tõesta, et kui f on kumer (mitte ilmtingimata tõkestatud), siis

$$S_n(f) \geq f.$$

12. Olgu $X_1, X_2, \ldots$ i.i.d. juhuslikud suurused, $E|X_i| < \infty$, $EX_1 = \mu$. Olgu $m_n \to \infty$ mittenegatiivsete täisarvude jada. Näita, et ($S_k = X_1 + \cdots + X_k$)

1. $\frac{S_{m_n}}{m_n} \to \mu$ p.k.,

2. $\frac{X_{m_n}}{m_n} \to 0$, p.k.,

3. $\frac{X_{m_n+1}}{m_n} \to 0$, p.k.,

4. kui lisaks $\frac{m_n}{n} \to a > 0$, siis $\frac{S_{m_n}}{n} \to \mu \cdot a$, p.k..

Olgu nüüd $N_1 \leq N_2 \leq N_3 \leq \cdots$ täisarvuliste juhuslike suuruste jada nii, et $\lim_n N_n = \infty$ p.k. Näita

1. $\frac{S_{N_n}}{N_n} \to \mu \quad$ p.k.,

2. $\frac{X_{N_n}}{N_n} \to 0, \quad$ p.k.,

3. $\frac{X_{N_n+1}}{N_n} \to 0, \quad$ p.k.,

4. kui lisaks $\frac{N_n}{n} \to a > 0$, p.k., siis $\frac{S_{N_n}}{n} \to \mu \cdot a, \quad$ p.k..

13. Olgu $X_1, X_2, \ldots$ mittenegetiivsed iid juhuslikud suurused $EX_i = \mu > 0$. Olgu iga $t \geq 0$ korral

$$N(t) = \max\{n : S_n \leq t\},$$

kus $S_0 = 0$ ja $S_k = X_1 + \cdots + X_k$. Näita:

1. $\mathbf{P}(N(t) < \infty) = 1$, iga t korral;

2. $\mathbf{P}(\lim_{t\to\infty} N(t) = \infty) = 1$;

3. $\lim_{t\to\infty} \frac{S_{N(t)}}{N(t)} = \mu$, p.k.;

4. $\lim_{t\to\infty} \frac{N(t)}{t} = \frac{1}{\mu}$, p.k (järelda eelmisest koondumisest, kasutades asjaolu, et $S_{N(t)} \leq t \leq S_{N(t)+1}$);

5. eeldusel, et leidub juhuslik suurus Z nii, et protsessis $k \to \infty$ kehtib

$$\frac{S_k - \mu k}{\sqrt{k}} \Rightarrow Z,$$

siis protsessis $t \to \infty$,

$$\frac{S_{N(t)} - \mu N(t)}{\sqrt{\frac{t}{\mu}}} \Rightarrow Z, \quad \frac{S_{N(t)} - \mu N(t)}{\sqrt{N(t)}} \Rightarrow Z. \tag{11.25}$$

(Kasuta ptk 9 ülesannet 30).

12 Tõenäosusmõõtude nõrk koondumine

Olgu $X_n \sim P_n$, $n = 1, 2, \ldots$, $X \sim P$ juhuslikud suurused. F_n, F olgu vastavas jaotusfunktsioonid. Teame, et jada X_n koondub jaotuse järgi juhuslikuks suuruseks X ($X_n \Rightarrow X$) parajasti siis, kui

$$P_n(-\infty, x] \to P(-\infty, x] \quad \text{iga sellise } x \text{ korral, et} \quad P\{x\} = 0. \tag{12.1}$$

Arusaadavalt on tingimus (12.1) ekvivalentne tingimusega

$$F_n(x) \to F(x) \quad \text{iga sellise } x \text{ korral, kus} \quad F \quad \text{on pidev.} \tag{12.2}$$

Seega koondumine $X_n \Rightarrow X$ on tegelikult tõenäosusmõõtude P_n ja P omadus. Kui mõõtudel P_n ja P on see omadus, siis ütleme, et jada P_n koondub **nõrgalt** mõõduks P. Sõnastame selle definitsioonina.

Def 12.1 *Olgu P_n, $n = 1, 2, \ldots$ ja P σ-algebral $\mathcal{B}(\mathbb{R})$ antud tõenäosusmõõdud. Jada P_n koondub nõrgalt mõõduks P, kui kehtib (12.1). Nõrka koondumist tähistame $P_n \Rightarrow P$.*

Seega juhuslikud suurused koonduvad jaotuse järgi parajasti siis, kui nende jaotused koonduvad nõrgalt.

12.1 Skorohodi esitus (veelkord)

Et mingi juhuslike suuruste jada jaotuse järgi koondumine on vaid nende jaotuste omadus, ei saa jaotuse järgi koondumisest teha mingeid järeldusi nende juhuslike suuruste kui funktsioonide omavahelise seose kohta. Võib isegi nii olla, et iga juhuslik suurus selles jadas on defineeritud eraldi tõenäosusruumil.
Teame aga, et iga etteantud jaotuse P korral saab tõenäosusruumil $((0,1), \mathcal{B}(0,1), \mathrm{Leb})$ konstrueerida juhusliku suuruse Z, mille jaotus on P, jaotusfunktsiooniga F. Üks võimalus sellise juhusliku suuruse konstrueerimiseks on Skorohodi esitus (vt. alajaotus 4.11), kus Z on sisuliselt F pöördfunktsioon. Täpsemalt,

$$Z(\omega) = \inf\{x : \omega \leq F(x)\}. \tag{12.3}$$

Seega, kui $P_n \Rightarrow P$ on nõrgalt koonduvate mõõtude jada, siis saab ruumil $((0,1), \mathcal{B}(0,1), \mathrm{Leb})$ seose (12.3) abil defineerida juhuslikud suurused Z_n, Z nii, et $Z_n \Rightarrow Z$. Järgnev teoreem näitab, et need juhuslikud suurused koonduvad p.k..

Teoreem 12.2 *Olgu $P_n \Rightarrow P$ σ-algebral $\mathcal{B}(\mathbb{R})$ antud nõrgalt koonduvad tõenäosusmõõdud. Kui $Z_n \sim P_n$ ja $Z \sim P$ on Skorohodi esituse (12.3) kaudu defineeritud juhuslikud suurused, siis $Z_n \to Z$ p.k..*

Tuletame meelde, et Skorohodi esituse definitsioonist järeldub:

$$\{\omega : \omega \leq F(x)\} = \{\omega : Z(\omega) \leq x\} \quad \Leftrightarrow \quad \{\omega : \omega > F(x)\} = \{\omega : Z(\omega) > x\}.$$

Tõestus. Näitame, et $Z_n \to Z$ p.k. Olgu $\omega \in (0,1)$ selline, et Z on pidev punktis ω. Näitame
1) $\liminf_n Z_n(\omega) \geq Z(\omega)$.
Olgu $\epsilon > 0$ suvaline. Valime x nii, et $Z(\omega) - \epsilon < x < Z(\omega)$ ja $P\{x\} = 0$ (selline x alati leidub). Seega $F(x) < \omega$ ja koondumisest $F_n(x) \to F(x)$ saame $F_n(x) < \omega$ kui n on piisavalt suur. See aga tähendab, et $Z_n(\omega) > x$ (miks?), kui n on piisavalt suur, mistõttu $\liminf_n Z_n(\omega) \geq x > Z(\omega) - \epsilon$. Et ϵ oli suvaline, saame nüüd, et $\liminf_n Z_n(\omega) \geq Z(\omega)$.
2) $\limsup_n Z_n(\omega) \leq Z(\omega)$.
Olgu $\omega' > \omega$, $\epsilon > 0$ suvaline. Vali y nii, et $P\{y\} = 0$ ja $Z(\omega') < y < Z(\omega') + \epsilon$ (selline y alati leidub). Seega $\omega < \omega' \leq F(y)$ ja koondumisest $F_n(y) \to F(y)$ saame $F_n(y) \geq \omega$ kui n on piisavalt suur. See aga tähendab, et $Z_n(\omega) \leq y < Z(\omega') + \epsilon$, kui n on piisavalt suur, millest $\limsup_n Z_n(\omega) \leq Z(\omega')$. Et Z on punktis ω pidev, siis protsessis $\omega' \searrow \omega$ saame $\limsup_n Z_n(\omega) \leq Z(\omega)$.
Funktsioon Z on lõigus $(0,1)$ mittekahanev, mistõttu on tal ülimalt loenduv hulk katkevuspunkte. Olgu see hulk A; $\mathbf{P}(A) = 0$. Et $Z_n(\omega) \to Z(\omega)$ $\forall \omega \in A^c$, siis $Z_n \to Z$ p.k. ■

Järeldus 12.1 (Pideva kujutise teoreem) *Olgu $X_n \Rightarrow X$, $g : \mathbb{R} \to \mathbb{R}$ olgu pidev. Siis $g(X_n) \Rightarrow g(X)$ ja $P_n g^{-1} \Rightarrow P g^{-1}$.*

Tõestus. Olgu Z_n, Z sellised juhuslikud suurused nagu teoreemis 12.2. Seega $Z_n \to Z$ p.k., millest $g(Z_n) \to g(Z)$ p.k.. Seega $g(Z_n) \Rightarrow g(Z)$. Et $g(Z_n)$ ja $g(X_n)$ ning $g(Z)$ ja $g(X)$ on sama jaotusega, siis $g(X_n) \Rightarrow g(X)$.
Et $Z \sim P$, siis iga Boreli hulga B korral

$$Pg^{-1}(B) = P(g^{-1}(B)) = \mathbf{P}(Z \in g^{-1}(B)) = \mathbf{P}(g(Z) \in B)$$

ehk $g(Z) \sim Pg^{-1}$. Analoogiliselt, $g(Z_n) \sim P_n g^{-1}$ ja koondumine $g(Z_n) \Rightarrow g(Z)$ on sama mis $P_n g^{-1} \Rightarrow P g^{-1}$. ■

12.2 Koondumiskriteeriumid

Teoreemi 12.2 abil on kerge tõestada erinevaid koondumiskriteeriume.

Def 12.3 *Boreli hulka $A \subset \mathcal{B}$ nim P-**pidevaks**, kui $P(\partial A)$=0, kus ∂A on hulga A raja.*

Suvalise alamhulga $A \subset \mathbb{R}$ korral on ∂A kinnine, mistõttu $\partial A \in \mathcal{B}$. Näiteks $(-\infty, x]$ on P-pidev parajasti siis, kui $P\{x\} = 0$ ehk P jaotusfunktsioon F on pidev punktis x.

Teoreem 12.4 *Olgu P ja P_n, $n = 1, 2, \ldots$ ruumil $(\mathbb{R}, \mathcal{B})$ antud tõenäosusmõõdud. Järgmised väited on samaväärsed:*

1) $P_n \Rightarrow P$;

2) $\int f dP_n \to \int f dP$ *iga pideva tõkestatud f korral;*

3) $P_n(A) \to P(A)$ *iga P-pideva Boreli hulga A korral.*

Tõestus. **1)** $\Rightarrow$ **2)**: Olgu $Z_n \sim P_n$ ja $Z \sim P$ sellised ruumil $(\Omega, \mathcal{F}, \mathbf{P})$ defineeritud juhuslikud suurused nagu teoreemis 12.2; f olgu tõkestatud ja pidev funktsioon. Et $Z_n \to Z$ p.k. ja f on pidev, siis $f(Z_n) \to f(Z)$ p.k. Tõkestatud koondumise teoreemist ja Lebesgue'i integraali muutuja vahetusest järeldub nüüd, et

$$\int f dP_n = Ef(Z_n) \to Ef(Z) = \int f dP.$$

1) $\Rightarrow$ **3)**: Üldistame eeltoodud arutelu. Olgu f tõkestatud Boreli funktsioon, D_f olgu tema katkevuspunktide hulk (kui f on pidev, siis $D_f = \emptyset$). Pidevuse asemel eeldame, et $P(D_f) = 0$. Seega $f(Z_n(\omega)) \to f(Z(\omega))$, kui $Z_n(\omega) \to Z(\omega)$ ja $Z(\omega) \in {D_f}^c$ ehk $\omega \in Z^{-1}({D_f}^c)$. Et aga $\mathbf{P}(Z^{-1}({D_f}^c)) = P({D_f}^c) = 1$, siis $f(Z_n) \to f(Z)$ p.k.. Nüüd kasutame taas tõkestatud koondumise teoreemi ja saame $Ef(Z_n) \to Ef(Z)$ ehk $\int f dP_n \to \int f dP$. Olgu nüüd A mingi P-pidev Boreli hulk. Võtame $f = I_A$. Funktsioon f on tõkestatud Boreli funktsioon, $D_f = \partial A$, mistõttu $P(D_f) = 0$. Järelikult, $P_n(A) = \int f dP_n \to \int f dP = P(A)$.
3) $\Rightarrow$ **1)**: Ilmne (miks?).
2) $\Rightarrow$ **1)**: Olgu F_n ja F mõõtude P_n ja P jaotusfunktsioonid. Olgu x funktsiooni F pidevuspunkt. Väide on tõestatud, kui me näitame, et $F_n(x) \to F(x)$.
Valime $y > x$ ja defineerime

$$f(t) = \begin{cases} 1, & \text{kui } t \leq x; \\ \frac{(y-t)}{(y-x)}, & \text{kui } x \leq t \leq y; \\ 0, & \text{kui } t \geq y. \end{cases}$$

Definitsioonist saame, et $F_n(x) = \int_{(-\infty,x]} dP_n \leq \int f dP_n$ ja $\int f dP \leq \int_{(-\infty,y]} dP = F(y)$. Seosest **2)** järeldub $\limsup_n F_n(x) \leq F(y)$. Protsessis $y \searrow x$ saame $\limsup_n F_n(x) \leq F(x)$.
Olgu $u < x$. Analoogiliselt arutledes saame $F(u) \leq \liminf_n F_n(x)$. Et F on punktis x pidev, siis protsessis $u \nearrow x$ saame $F(x) \leq \liminf_n F_n(x)$. Järelikult $F_n(x) \to F(x)$. ■

Teoreemis 12.4 toodud koondumiskriteeriume võib kasutada (ja tihti kasutataksegi) tõenäosusmõõtude nõrga koondumise definitsioonina.

Teoreem 12.4 vastab muuhulgas ka järgmisele väga olulisele küsimusele: olgu B mingi Boreli hulk ja $P_n \Rightarrow P$ mingi nõrgalt koonduvate tõenäosusmõõtude jada. Kas siis kehtib koondumine $P_n(B) \to P(B)$, st kas piisavalt suure n korral võib kasutada lähendit $P_n(B) \approx P(B)$? Teoreemist jäeldub, et kui B on P-pidev hulk, siis küll. Järgmistest näidetest selgub aga, et üldiselt see nii pole.

Näited:

1. Aatomite koondumine. $P_n := \delta_{x_n} \Rightarrow P := \delta_x$ parajasti siis, kui $x_n \to x$. Sellisel juhul on hulk A P-pidev parajasti siis, kui tema raja ei sisalda punkti x. On kerge veenduda,

et iga sellise hulga A puhul $P_n(A) \to P(A)$, hulga $A = \{x\}$ korral ei pruugi see aga nii olla.

2. Olgu P_n ühtlane üle aatomite $\{0, \frac{1}{n}, \dots \frac{n-1}{n}\}$, P olgu ühtlane jaotus hulgal $[0,1]$. On lihtne veenduda, et iga x korral $F_n(x) \to F(x)$. Kuid ikkagi eksisteerib hulk A nii, et $P_n(A) \not\to P(A)$. Loomulikult ei saa A sellisel juhul olla P-pidev.

Juhuslikud suurused. Olgu $X_n \sim P_n$ ja $X \sim P$ juhuslikud suurused. Tuletame veelkord meelde, et juhuslike suuruste jada X_n koondub nõrgalt juhuslikuks suuruseks X, kui nende jaotused koonduvad nõrgalt, st $P_n \Rightarrow P$. Esitame ülaltoodud koondumiskriteeriumid juhuslike suuruste kaudu. Tingimus (12.1) (ja ka tingimus (12.2)) avaldub juhuslike suuruste kaudu järgmiselt

$$\mathbf{P}(X_n \leq x) \to \mathbf{P}(X \leq x) \quad \text{iga sellise } x \text{ korral, kus} \quad \mathbf{P}(X = x) = 0. \tag{12.4}$$

Teoreemi 12.4 tingimus **2)** avaldub juhuslike suuruste kaudu järgmiselt

$$E\big(f(X_n)\big) \to E\big(f(X)\big) \text{ iga pideva ja tõkestatud } f \text{ korral.} \tag{12.5}$$

NB! Polünoomid pole tõkestatud funktsioonid, seega ei järeldu seosest (12.5) keskväärtuste, dispersioonide ja muude momentide koondumine.

Teoreemi 12.4 tingimus **3)** avaldub juhuslike suuruste kaudu järgmiselt

$$\mathbf{P}(X_n \in A) \to \mathbf{P}(X \in A) \quad \text{iga sellise } A \in \mathcal{B}(\mathbb{R}) \text{ korral, kus} \quad \mathbf{P}(X \in \partial A) = 0. \tag{12.6}$$

12.3 Suhteline kompaktsus

Kui arvjada $\{x_n\}$ on tõkestatud, siis iga tema alamjadal leidub omakorda koonduv alamjada. Matemaatikas nimetatakse sellist omadust *suhteliseks kompaktsuseks*. Seega arvjada $\{x_n\}$ on suhteliselt kompaktne, kui igal alamjadal leidub koonduv alamjada. Arvjadade korral on tõkestatus ja suhteline kompaktsus ekvivalentsed mõisted.
Teame juba, et arvjada $\{x_n\}$ koondub piirväärtuseks x parajasti siis, kui jada $\{x_n\}$ igal alamjadal leidub omakorda piirväärtuseks x koonduv alamjada.

Seega $x_n \to x$ parajasti siis, kui tal on kaks omadust:

- $\{x_n\}$ on suhteliselt kompaktne;
- iga koonduva alamjada piirväärtus on x.

(Veendu selles!)

Iga suhteliselt kompaktne jada pole veel koonduv; samuti võib tuua näiteid mittekoonduvast jadast, mille igal koonduval alamjadal on üks ja seesama piirväärtus.

Analoogilised omadused on nõrgalt koonduval tõenäosusmõõtude jadal.

Def 12.5 *Olgu $\{P_n\}$ σ-algebral $\mathcal{B}(\mathbb{R})$ antud tõenäosusmõõtude jada. Jada $\{P_n\}$ nimetatakse* **suhteliselt kompaktseks**, *kui tema igal alamjadal leidub mingisuguseks tõenäosusmõõduks nõrgalt koonduv (alam)alamjada.*

Lemma 12.1 *Olgu $\{P_n\}$ σ-algebral $\mathcal{B}(\mathbb{R})$ antud tõenäosusmõõtude jada. Jada $\{P_n\}$ koondub piirväärtuseks P parajasti siis, kui tema igal alamjadal leidub mõõduks P koonduv alam-alamjada.*

Tõestus. $\Rightarrow$: Oletame, et $P_n \Rightarrow P$. Olgu P_{n_k} suvaline alamjada. Definitsioonist (12.1) järeldub, et ka $P_{n_k} \Rightarrow P$ (miks?). Sellega on tarvilikkus tõestatud.
$\Leftarrow$: Oletame vastuväiteliselt, et $P_n \not\Rightarrow P$. See tähendab, et $\exists x$ nii, et $P\{x\} = 0$, kuid

$$P_n(-\infty, x] \not\to P(-\infty, x].$$

Seega $\exists \epsilon > 0$ ja alamjada $\{P_{n_k}\}$ nii, et $|P_{n_k}(-\infty, x] - P(-\infty, x]| \geq \epsilon$, $k = 1, 2, \ldots$. Vastavalt eeldusele aga jadal $\{P_{n_k}\}$ leidub mõõduks P koonduv alamjada $\{P_{n_{k_l}}\}$, mis tähendab, et $|P_{n_{k_l}}(-\infty, x] - P(-\infty, x]| \to 0$, kui $l \to \infty$. See on aga vastuolu. ■

Seega kehtib koondumine $P_n \Rightarrow P$ parajasti siis, kui

1 jada $\{P_n\}$ on suhteliselt kompaktne;

2 iga nõrgalt koonduva alamjada piirväärtus on P, st kui $P_{n_k} \Rightarrow Q$ on mingiks tõenäosusmõõduks Q nõrgalt koonduv alamjada, siis $P = Q$.

Tihti tuleb ette olukordi, mil mingi antud jada $\{P_n\}$ korral on suhteliselt lihtne kontrollida omadust **2**. Veendumaks, et tõenäosusmõõtude jada koondub nõrgalt, on meil sellisel juhul vaja veel kontrollida selle jada suhtelist kompaktsust (omadust **1**). See aga ei olegi üldjuhul eriti lihtne.

Järgnev teoreem annab praktikas kasuliku kriteeriumi jada $\{P_n\}$ suhtelise kompaktsuse kontrollimiseks. Defineerime tõenäosusmõõtude jada tiheduse.

Def 12.6 *Olgu $\{P_n\}$ σ-algebral $\mathcal{B}(\mathbb{R})$ antud tõenäosusmõõtude jada. Jada $\{P_n\}$ nimetatakse* **tihedaks** *kui $\forall \epsilon > 0$ korral leidub lõplik poollõik $(a, b]$ nii, et iga n korral $P_n(a, b] > 1 - \epsilon$.*

Jaotusfunktsioonide kaudu antud tõenäosusmõõtude jada $\{F_n\}$ on tihe parajasti siis, kui $\forall \epsilon > 0$ korral leiduvad $a < b$ nii, et $F_n(b) - F_n(a) > 1 - \epsilon$ iga n korral.
Seega pole $\{F_n\}$ tihe, kui $\exists \epsilon > 0$ nii, et iga a ja b korral leidub n nii, et $F_n(b) - F_n(a) \leq 1 - \epsilon$.

Tõenäosusmõõtude jada tihedus üldistab tõkestatuse mõistet. Sisuliselt tähendab tihedus ju seda, et iga $\epsilon > 0$ korral saame kõikide mõõtude sabad korraga ära lõigata nii, et ühegi lõigatud saba mass pole suurem kui ϵ. Arvajadade korral on tõkestatus ja suhteline kompaktsus samaväärsed. Tõenäosusmõõtude korral on tihedus ja suhteline kompaktsus samaväärsed. See tulemus on Prohorovi teoreem 12.8. Prohorovi teoreemi tõestus põhineb järgmisle teoreemil.

Teoreem 12.7 (Helly) *Olgu $\{F_n\}$ jaotusfunktsioonide jada. Siis leidub alamjada $\{F_{n_k}\}$ ja mittekahanev paremalt pidev funktsioon F nii, et $\forall x \in \mathbb{R}$ korral $0 \leq F(x) \leq 1$ ja iga funktsiooni F pidevuspunkti x korral kehtib koondumine $F_{n_k}(x) \to F(x)$.*

Tõestuse võib leida raamatutest: *Billingsley*, lk. 336; *Ŝirjajev*, lk. 340; *Williams* lk. 183.

Märkus: Piirfunktsioon F Helly teoreemis rahuldab jaotusfunktsiooni omadusi, välja arvatud asjaolu, et võib juhtuda, et $\lim_{x\to\infty} F(x) < 1$ ja $\lim_{x\to-\infty} F(x) > 0$. Seega ei pruugi piirfunktsioon olla jaotusfunktsioon.
Et aga F rahuldab jaotusfunktsiooni teisi omadusi, järeldub teoreemist 2.8, et sellisele funktsioonile saab seose (2.9) abil σ-algebral $\mathcal{B}(\mathbb{R})$ defineerida mingi mõõdu, mis ei pruugi küll olla tõenäosusmõõt.

Näited:

1. Olgu $P_n = \delta_{x_n}$, kus $x_n \to \infty$. Siis $F_n(x) = I_{[x_n,\infty)}$ ning $F_n(x) \to 0$ iga x korral. Seega Helly teoreem küll kehtib, kuid piirfunktsioon on samaselt 0, s.t. $F \equiv 0$ ja pole seega jaotusfunktsioon. On lihtne veenduda, et $\{P_n\}$ pole tihe.

2. Olgu G selline pidev jaotusfunktsioon, et $G(0) = 0.5$. Defineerime jaotusfunktsioonide jada

$$F_n(x) = G(x)G(\frac{x}{n}).$$

(F_n on $\max\{X, nY\}$ jaotus, kus X, Y on sõltumatud ja jaotusega G.)
Iga x korral $F_n(x) \to \frac{1}{2}G(x)$. Seega Helly teoreem küll kehtib, kuid piirfunktsioon pole jaotusfunktsioon.
Jällegi pole raske veenduda, et jada $\{F_n\}$ pole tihe (aga veenduge!).

Järgnev teoreem väidab sisuliselt, et kui jada $\{P_n\}$ on tihe, siis piirfunktsioon Helly teoreemis on jaotusfunktsioon ja sellele vastav mõõt on tõenäosusmõõt.

Teoreem 12.8 (Prohorov) *Olgu $\{P_n\}$ σ-algebral $\mathcal{B}(\mathbb{R})$ antud tõenäosusmõõtude jada. Jada $\{P_n\}$ on suhteliselt kompaktne parajasti siis, kui ta on tihe.*

Tõestus. $\Rightarrow$: Olgu $\{P_n\}$ suhteliselt kompaktne. Näitame, et see jada on tihe. Vastuväiteliselt oletades $\exists\epsilon > 0$ nii, et iga lõpliku poollõigu $(a, b]$ korral $\exists n$ nii, et $P_n(a, b] \leq 1-\epsilon$. Olgu n_k selline, et $P_{n_k}(-k, k] \leq 1-\epsilon$. Suhtelise kompaktsuse tõttu saame leida alamjada $\{P_{n_{k_l}}\}$ ja mingi tõenäosusmõõdu P nii, et $P_{n_{k_l}} \Rightarrow P$, kui $l \to \infty$. Olgu nüüd a, b sellised, et $P\{a\} = P\{b\} = 0$ ja $P(a, b] > 1-\epsilon$ (sellised a ja b leiduvad). Kui l on piisavalt suur, siis $(a, b] \subset (-k_l, k_l]$, millest $1-\epsilon \geq P_{n_{k_l}}(-k_l, k_l] \geq P_{n_{k_l}}(a, b] \to P(a, b]$. Järelikult $P(a, b] \leq 1-\epsilon$ - vastuolu.
$\Leftarrow$: Olgu $\{P_n\}$ tihe. Näitame, et see jada on suhteliselt kompaktne. Olgu $\{P_{n_k}\}$ suvaline alamjada, $\{F_{n_k}\}$ vastavad jaotusfunktsioonid. Helly teoreemi abil leiame (alam-)alamjada $\{F_{n_{k_l}}\}$ ja piirfunktsiooni F nii, et $F_{n_{k_l}}(x) \to F(x)$ iga funktsiooni F pidevuspunkti korral.

Implikatsioon on tõestatud, kui näitame, et F on jaotusfunktsioon, so $\lim_{x\to\infty} F(x) = 1$ ja $\lim_{x\to-\infty} F(x) = 0$. Et $0 \leq F(x) \leq 1$ ja F on mittekahanev, siis selleks piisab, kui näitame, et iga $\epsilon > 0$ korral $\exists a < b$ nii, et $F(b) - F(a) \geq 1 - \epsilon$.
Olgu $\epsilon > 0$ suvaline. Valime a, b nii, et $P_n(a, b] > 1 - \epsilon$, $n = 1, 2 \ldots$. Suurendades b-d ja vähendades a-d kui vaja, võime eeldada, et punktid a ja b on funktsiooni F pidevuspunktid. Seega $F(b) - F(a) = \lim_l F_{n_{k_l}}(b) - \lim_l F_{n_{k_l}}(a) = \lim_l P_{n_{k_l}}(a, b] \geq 1 - \epsilon$.
■

Järeldus 12.2 *Olgu tõenäosusmõõtude jada $\{P_n\}$ tihe. Kui leidub tõenäosusmõõt P nii, et P on jada $\{P_n\}$ iga nõrgalt koonduva alamjada piirväärtus, siis $P_n \Rightarrow P$.*

12.3.1 Juhuslikud vektorid ja tõenäosusmõõdud ruumis $\mathbb{R}^k$

Olgu $X_n := (X_n^1, \ldots, X_n^k)$ ja $X := (X^1, \ldots, X^k)$ k-dimensonaalsed juhuslikud vektorid. Nende jaotused olgu vastavalt P_n ja P. Jaotused P_n ja P on tõenäosusmõõdud σ-algebral $\mathcal{B}(\mathbb{R}^k)$, neile vastavad jaotusfunktsioonid F_n ja F on nüüd defineeritud ruumil $\mathbb{R}^k$.

Ruumil $\mathbb{R}^k$ (st σ-algebral $\mathcal{B}(\mathbb{R}^k)$) antud tõenäosusmõõtude koondumine $P_n \Rightarrow P$ defineeritakse põhimõtteliselt sama moodi kui reaalteljel. Tihti tehakse seda läbi jaotusfunktsioonide: ütleme, et $P_n \Rightarrow P$ kui kehtib (12.2) (jaotusfunktsiooni argument on nüüd muidugi ruumi $\mathbb{R}^k$ element). Samuti kehtib teoreemi 12.4 üldistus: $P_n \Rightarrow P$ parajasti siis kui

$$\int f dP_n \to \int f dP \text{ iga pideva ja tõkestatud } f : \mathbb{R}^k \to \mathbb{R} \text{ korral} \tag{12.7}$$

ning $P_n \Rightarrow P$ parajasti siis kui

$$P_n(A) \to P(A) \quad \text{iga } P\text{-pideva hulga } A \in \mathcal{B}(\mathbb{R}^k) \text{ korral.} \tag{12.8}$$

Seose (12.1) üldistamisel mitmemõõtmelisele juhule tuleb olla tähelepenelik. Punkt $x = (x^1, \ldots, x^k) \in \mathbb{R}^k$ on jaotusfunktsiooni pidevuspunkt parajasti siis, kui

$$P\big((-\infty, x^1] \times \cdots \times (-\infty, x^k]\big) - P\big((-\infty, x^1) \times \cdots \times (-\infty, x^k)\big) = 0.$$

Juhuslike vektorite nõrk ehk jaotuse järgi koondumine defineeritakse analoogiliselt: juhuslikud vektorid koonduvad nõrgalt, kirjutame $X_n \Rightarrow X$, parajasti siis, kui nende jaotused koonduvad nõrgalt, st $P_n \Rightarrow P$. Seosed (12.7) ja (12.8) avalduvad juhuslike vektrite kaudu järgmiselt:

$$E\big(f(X_n)\big) \to E\big(f(X)\big) \text{ iga pideva ja tõkestatud } f : \mathbb{R}^k \to \mathbb{R} \text{ korral;} \tag{12.9}$$

$$\mathbf{P}(X_n \in A) \to \mathbf{P}(X \in A) \quad \text{iga sellise } A \in \mathcal{B}(\mathbb{R}^k) \text{ korral, kus} \quad \mathbf{P}(X \in \partial A) = 0. \tag{12.10}$$

Märgime veel, et ka juhuslike vektorite korral kehtib teoreem 12.2, st leiduvad sama jaotusega p.k. koonduvad juhuslikud vektorid. Aga erinevalt ühedimensionaalsest juhust, nende peaaegu kindlalt koonduvate sama jaotusega vektorite konstrueerimine pole enam jautusfunktsiooni (üldistatud) pöördfunktsiooni s.o. Skorohodi esituse abil. Sellest teoreemist järeldub pideva kujutise teoreem: kui $X_n \Rightarrow X$, siis $g(X_n) \Rightarrow g(X)$, kus $g : \mathbb{R}^k \to \mathbb{R}$ on pidev (mitte ilmtingimata tõkestatud).

Ruumil $\mathbb{R}^k$ (st σ-algebral $\mathcal{B}(\mathbb{R}^k)$) antud tõenäosusmõõtude jada $\{P_n\}$ tihedus defineeritakse analoogiliselt ühemõõtmelise juhuga: $\{P_n\}$ on tihe, kui iga $\epsilon > 0$ korral leidub ristkülik $A = (a^1, b^1] \times \cdots \times (a^k, b^k]$ nii, et $P_n(A) > 1 - \epsilon$ iga n korral. Ka juhuslike vektorite korral kehtib Prohorovi teoreem: jada $\{P_n\}$ on suhteliselt kompaktne parajasti siis kui ta on tihe.

Pideva kujutise teoreemist järeldub, et kui $X_n \Rightarrow X$ ja $t \in \mathbb{R}^k$, siis $t'X_n \Rightarrow t'X$ (veendu). Seega, kui $X_n \Rightarrow X$, siis koonduvad nõrgalt ka kõik komponendid: $X_n^i \Rightarrow X^i$ iga $i = 1, \ldots, k$ korral (miks?). Vastupidine üldiselt ei kehti: komponentide nõrgast koondumisest ei järeldu vektorite nõrk koondumine (leia kontranäide). Üks erijuht, mil juhuslike vektorite nõrk koondumine on ekvivalentne komponentide nõrga koondumisega on siis, kui komponendid on iga n korral sõltumatud: kui $X_n^1, \ldots, X_n^k$ on iga n korral sõltumatud juhuslikud suurused ja $X_n^i \Rightarrow X^i$ iga i korral, siis

$$(X_n^1, \ldots, X_n^k) \Rightarrow (X_1, \ldots, X_k),$$

kus vektori $(X_1, \ldots, X_k)$ komponendid on sõltumatud (ülesanne 7). Küll aga kehtib järgmine tulemus: kui komponendid (marginaaljaotused) koonduvad nõrgalt, on juhuslike vektorite jaotuste jada tihe ning seega (Prohorovi teoreem) suhteliselt kompaktne.

Lemma 12.2 *Olgu $X_n = (X_n^1, \ldots, X_n^k) \sim P_n$ juhuslikud vektorid, mille komponendid koonduvad nõrgalt, st leiduvad juhuslikud suurused X^i, $i = 1, \ldots, k$ nii, et $X_n^i \Rightarrow X^i$ iga i korral, Siis jada $\{P_n\}$ on tihe.*

Tõestus. Olgu $X_n^i \sim P_n^i$. Et $X_n^i \Rightarrow X^i$, siis iga i korral on jada $\{P_n^i\}$ tihe (miks?). Seega iga $\epsilon > 0$ korral leidub $(a^i, b^i]$ nii, et $P_n^i\big((a^i, b^i]\big) > 1 - \frac{\epsilon}{k}$. Olgu $A := (a^1, b^1] \times \cdots \times (a^k, b^k]$. Veendume, et $\{P_n\}$ on tihe:

$$P^n(A^c) = \mathbf{P}(X_n^1 \notin (a^1, b^1], \cdots, X_n^k \notin (a^k, b^k]) \leq \sum_{i=1}^k \mathbf{P}(X_n^i \notin (a^i, b^i]) = \sum_{i=1}^k P_n^i\big((a^i, b^i]^c\big) \leq \epsilon.$$

■

Järgmine oluline teoreem ütleb, et tegelikult on koondumine $t'X_n \Rightarrow t'X$ iga $t \in \mathbb{R}^k$ korral ekvivalentne koondumisega $X_n \Rightarrow X$.

Teoreem 12.9 *Juhuslikud vektorid X_n koonduvad nõrgalt juhuslikuks vektoriks X, st $X_n \Rightarrow X$ parajasti siis, kui iga $t \in \mathbb{R}^k$ korral $t'X_n \Rightarrow t'X$.*

Teoreemi 12.9 abil on kerge üldistada Slutsky lemmat juhuslikele vaktoritele,

Lemma 12.3 (Slutsky lemma) *Olgu X_n, Y_n ja X k-dimensionaalsed juhuslikud vektorid. Kui $X_n \Rightarrow X$ ja $Y_n \Rightarrow c$, kus $c \in \mathbb{R}^k$, siis kehtivad koondumised:*

1. $X_n + Y_n \Rightarrow X + c$;
2. $Y_n' X_n \Rightarrow c' X$.

Tõestus. 1) Olgu $X_n = (X_n^1, \ldots, X_n^k)$, $X = (X^1, \ldots, X^k)$ ja $Y_n = (Y_n^1, \ldots, Y_n^k)$ juhuslikud vektorid, kusjuures $X_n \Rightarrow X$ ja $Y_n \xrightarrow{P} c$ (ekvivalentselt $Y_n \Rightarrow c$). Veendume, et $X_n + Y_n \Rightarrow X + c$. Selleks piisab, kui näitame, et iga vektori $t \in \mathbb{R}^k$ korral

$$t'(X_n + Y_n) = t'X_n + t'Y_n \Rightarrow t'X + t'c = t'(X + c). \tag{12.11}$$

Et $X_n \Rightarrow X$, siis $t'X_n \Rightarrow t'X$. Samuti $t'Y_n \Rightarrow t'c \in \mathbb{R}$. Ühemõõtmelisest Slutsky lemmast (seos 1) saame nüüd, et $t'X_n + t'Y_n \Rightarrow t'X + t'c$ ja see on (12.11).
2) Oletame, et $c = 0$, st $X_n \Rightarrow X$ ja $Y_n \xrightarrow{P} 0$. See tähendab, et suvaliste komponentide i ja j korral $X_n^i \Rightarrow X^i$ ja $Y_n^i \xrightarrow{P} 0$, millest (Slutsky lemma) $X_n^i Y_n^i \xrightarrow{P} 0$. Seega kõik korrutised $X_n^i Y_n^i$ koonduvad tõenäosuse järgi nulliks. Liidetavate tõenäosuse järgi koondumisest järeldub summa tõenäosuse järgi koondumine ehk

$$Y_n' X_n = \sum_{i=1}^{k} Y_n^i X_n^i \xrightarrow{P} 0.$$

Kui $c \neq 0$, siis $Y_n - c \Rightarrow 0$ (miks?), millest $(Y_n - c)' X_n \xrightarrow{P} 0$. Et $c' X_n \Rightarrow c' X$, siis osa 1) (või ühemõõtmelise Slutsky lemma) tõttu

$$(Y_n - c + c)' X_n = (Y_n - c)' X_n + c' X_n \Rightarrow c' X.$$

■

Kirjandus. *Billingsley* 25; *Williams* 17; *Ŝirjajev* III osad 1 ja 2; *Grimmett, Stirzaker* 7.2.

12.4 Ülesanded

1.

1. Olgu iga $t \in \mathbb{R}$ korral
$$G_\lambda(t) := \exp[-e^{-\lambda t}], \quad \lambda > 0.$$
Tõesta, et G_λ on jaotusfunktsioon
2. Olgu $X_1, X_2, \ldots$ iid juhuslikud suurused eksponentjaotusega $E(\lambda)$. Seega
$$\mathbf{P}(X_i > t) = \exp[-\lambda t],$$

kui $t \geq 0$. Olgu $M_n := \max\{X_1, \dots, X_n\}$. Tõestada, et

$$M_n - \frac{\ln n}{\lambda} \Rightarrow Y,$$

kus $Y \sim G_\lambda$.

2. Olgu $\{P_n\}$ absoluutselt pidevate mõõtude jada (Lebesgue'i mõõdu suhtes), f_n olgu mõõdu P_n tihedus. Olgu f mõõdu P tihedus (Lebesgue'i mõõdu suhtes). Tõesta, et kui $f_n \to f$ Leb-p.k., siis $P_n \Rightarrow P$.

3. Olgu $x_1, x_2, \dots$ ülimalt loenduv reaalarvude hulk, $x_i \neq x_j$, kui $j \neq j$. Olgu iga n korral $P_n := \sum_i p_i^n \delta_{x_i}$ tõenäosusmõõt (st $\sum_i p_i^n = 1$). Olgu $P = \sum_i p_i \delta_{x_i}$ samuti tõenäosusmõõt. Tõesta, et kui iga i korral $\lim_n p_i^n = p_i$, siis $P_n \Rightarrow P$.

4. Too näide nõrgalt koonduvatest mõõtudest $P_n \Rightarrow P$ ja tõkestatud funktsioonist f nii, et

$$\lim_n \int f dP_n \neq \int f dP.$$

5. Olgu $\{X_n\}$ suvaline juhuslike suuruste jada. Tõestada, et alati leiduvad konstandid a_n nii, et $a_n > 0$ iga n korral, kuid $a_n X_n \Rightarrow 0$.

6. Olgu $X_n \Rightarrow X$ ja $a_n \to 0$. Tõestada, et $a_n X_n \Rightarrow 0$.

7. Olgu $X_1, \dots$ ja $, Y_1, \dots$ sellised juhuslikud suurused, et iga $n = 0, 1, \dots$ korral X_n ja Y_n on sõltumatud. Olgu X, Y sõltumatud juhuslikud suurused ja kehtigu koondumised: $X_n \Rightarrow X$ ja $Y_n \Rightarrow Y$. Tõesta või lükka ümber:

- $(X_n, Y_n) \Rightarrow (X, Y)$;
- $X_n - Y_n \Rightarrow X - Y$.
- $X_n^2 + Y_n^2 \Rightarrow X^2 + Y^2$.
- $X_n Y_n \Rightarrow XY$.

8. Olgu $X_n \Rightarrow X$ ja $Y_n \Rightarrow c$. Tõesta või lükka ümber: $(X_n, Y_n) \Rightarrow (X, c)$.

9. Olgu X, Y sõltumatud juhuslikud suurused jaotustega P_X ja P_Y. Olgu (X_n, Y_n) sellised juhuslikud vektorid, et iga P_X-pideva A ja P_Y-pideva B korral

$$\mathbf{P}(X_n \in A, Y_n \in B) \to P_X(A) P_Y(B).$$

Näita, et

$$(X_n, Y_n) \Rightarrow (X, Y).$$

10. Kas $\{P_n\}$ on tihe, kui:

1. P_n on $E(n)$?
2. P_n on $E(n^{-1})$?
3. P_n on $U(1,n)$?
4. P_n on $U(0,n^{-1})$?
5. P_n on $\mathcal{N}(n,1)$?
6. P_n on $\mathcal{N}(0,n))$?
7. P_n on $\mathrm{Po}(n)$?
8. P_n on järgmine: $P_n(n) = n^{-1}$ ja $P_n(0) = 1 - n^{-1}$?

11. Olgu P tõenäosusmõõt. Näita, et P-pidevate sündmuste hulk moodustab algebra.

12. Olgu X_n, Y_n, Z_n $(n = 1, 2, \ldots)$ ja X juhuslikud suurused, kusjuures $X_n \Rightarrow X$ ja $Z_n \xrightarrow{P} 0$. Tõestada, et $Y_n \Rightarrow X$, kui

a) $|X_n - Y_n| \leq Z_n |X_n|$.

b) $|X_n - Y_n| \leq Z_n |Y_n|$.

[b) tõestamisel kasutage osa a) ja seost: kui $\epsilon < \frac{1}{2}$, siis võrratusest $|x - y| \leq \epsilon |y|$ järeldub, et $|x - y| \leq 2\epsilon |x|$. Tõesta see võrratus.]

13. Olgu $X_1, X_2, \ldots$ jaotusege P iid juhuslikud suurused, P_n vastav empiiriline mõõt. Kas kehtivad väited:

a) $\mathbf{P}(P_n \Rightarrow P) = 1$;

b) $\mathbf{P}\big(\sup_x |P_n(-\infty, x] - P(-\infty, x]| \to 0\big) = 1$;

c) $\mathbf{P}\big(\sup_{B \in \mathcal{B}} |P_n(B) - P(B)| \to 0\big) = 1$.

13 Momente genereeriv funktsioon ja suured hälbed

Olgu $X \sim P$ tõenäosusruumil $(\Omega, \mathcal{F}, \mathbf{P})$ antud juhuslik suurus.

Iga reaalarvu $t \in \mathbb{R}$ korral on e^{tX} mittenegatiivne juhuslik suurus, mille keskväärtus $E(e^{tX}) \leq \infty$ on alati defineeritud.

Def 13.1 *Laiendatud funktsiooni*

$$M : \mathbb{R} \to (0, \infty], \quad M(t) = E(e^{tX})$$

nimetatakse juhusliku suuruse X (mõõdu P) **momente genereerivaks funktsiooniks**.

Lebesgue'i integraal muutuja vahetuse kohaselt saame $\forall t \in \mathbb{R}$ korral

$$M(t) = E(e^{tX}) = \int e^{tx} P(dx).$$

Seega funktsioon M sõltub vaid juhusliku suuruse X jaotusest ehk sama jaotusega juhuslikel suurustel on sama momente genereeriv funktsioon.

Kui P on pidev jaotus tihedusfunktsiooniga f, siis

$$M(t) = \int_{-\infty}^{\infty} e^{tx} f(x) dx,$$

diskreetse jaotuse $P = \sum_i p_i \delta_{x_i}$ korral

$$M(t) = \sum_i e^{tx_i} p_i.$$

13.1 Piirkond T

Olgu $T := \{t \in \mathbb{R} | M(t) < \infty\}$. Arusaadavalt pakub funktsioon M huvi vaid hulgal T. Uurime seda (jaotusest P sõltuvat) piirkonda. Eelkõige paneme tähele, et iga jaotuse P korral $0 \in T$.

Olgu X p.k. mittenegatiivne. Sellisel juhul $M(t) = \int_{[0,\infty)} e^{tx} P(dx)$ ja on kerge veenduda, et M on mittekahanev. Et $0 \in T$, siis mittenegatiivse jaotuse korral on piirkond T kujul $T = (-\infty, t_2]$ või $T = (-\infty, t_2)$, kus $t_2 \geq 0$.

Olgu X p.k. mittepositiivne. Sellisel juhul $M(t) = \int_{(-\infty,0]} e^{tx} P(dx)$ ja M on mittekasvav. Et $0 \in T$, siis mittepositiivse jaotuse korral on piirkond T kujul $T = [t_1, \infty)$ või $T = (t_1, \infty)$, kus $t_1 \leq 0$.

Üldiselt aga $M(t) = \int_{(\infty,0]} e^{tx} P(dx) + \int_{(0,\infty)} e^{tx} P(dx)$, mistõttu on neli võimalust:

$$T = [t_1, t_2] \text{ või } T = [t_1, t_2) \text{ või } T = (t_1, t_2] \text{ või } T = (t_1, t_2),$$

kus $t_1 \leq 0$, $t_2 \geq 0$ kuid alati $0 \in T$. Kui $t_1 < 0$ ja $t_2 > 0$, siis $\exists t_o > 0 : (-t_o, t_o) \in T$.

Näide Jaotus P võib olla ka selline, et $T = \{0\}$. Selline jaotus on näiteks

$$P = \sum_{n=1}^{\infty} \frac{C}{n^2}(\delta_n + \delta_{-n}),$$

kus $\frac{1}{2C} = \sum_{n=1}^{\infty} \frac{1}{n^2}$.
Kui $t > 0$, siis rida $\sum_n \frac{e^{tn}}{n^2}$ hajub ning rida $\sum_n \frac{e^{-tn}}{n^2}$ koondub, millest $M(t) = \infty$. Analoogiliselt saame, et $M(t) = \infty$ iga $t < 0$ korral. Seega $T = \{0\}$.

13.2 Momente genereeriva funktsiooni arendamine astmeritta

Tuletame meelde, et kui (analüütiline) funktsioon f on arendatav astmeritta, st $f(x) = \sum_{n=0}^{\infty} a_n x^n$, siis $a_n = \frac{f^{(n)}(0)}{n!}$, $n = 0, 1, 2, \ldots$ (*G.Kangro* lk.73). Sellist rida nimetatakse Taylori reaks (ka Maclaureni reaks). Eksponentfunktsioon on arendatav Taylori ritta; et $\frac{d^n e^x}{dx^n} = e^x$, siis toodud valemist saame

$$e^x = 1 + x + \frac{x^2}{2} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots = \sum_{n=0}^{\infty} \frac{x^n}{n!}.$$

Selle rea koondumisvahemik on kogu arvsirge.

Uurime momente genereeriva funktsiooni arendamist astmeritta. Arusaadavalt on see võimalik vaid siis, kui M on lõplik punkti 0 ümbruses, st $\exists t_0 > 0$ nii, et $(-t_0, t_0) \subset T$. Eeldamegi sellist olukorda.
Olgu $t \in (-t_0, t_0)$. Arendades funktsiooni $e^{|tx|}$ astmeritta, saame

$$e^{|tx|} = \sum_{n=0}^{\infty} \frac{|tx|^n}{n!} = \sum_{n=0}^{\infty} \Big| \frac{(tx)^n}{n!} \Big|.$$

Et $t \in (-t_0, t_0)$, siis $M(t) < \infty$ ja $M(-t) < \infty$, millest saame

$$\int \sum_{n=0}^{\infty} \Big| \frac{(tx)^n}{n!} \Big| P(dx) = \int e^{|tx|} P(dx) \leq \int e^{tx} P(dx) + \int e^{-tx} P(dx) = M(t) + M(-t) < \infty.$$

Lebesgue'i integraali omadustest on teada (vt ülesanne 17), et kui summa $\int \sum_n |f_n| d\mu < \infty$, siis $\int \sum_n f_n d\mu = \sum_n \int f_n d\mu$. Seda tulemust kasutades saame $\forall t \in (-t_0, t_0)$ korral

$$M(t) = \int \sum_{n=0}^{\infty} \frac{(xt)^n}{n!} P(dx) = \sum_{n=0}^{\infty} \int x^n P(dx) \frac{t^n}{n!} = \sum_{n=0}^{\infty} E(X^n) \frac{t^n}{n!}. \tag{13.1}$$

Seega vaadeldaval juhul eksisteerivad jaotuse P kuitahes kõrget järku momendid, astmerea kordajad aga avalduvad tuletiste kaudu

$$\frac{E(X^n)}{n!} = \frac{M^{(n)}(0)}{n!} \quad \text{ehk} \quad M^{(n)}(0) = E(X^n) = \int x^n P(dx). \tag{13.2}$$

Kokkuvõtteks: kui momente genereeriv funktsioon M on lõplik punkti 0 mingis ümbruses, siis on tal kuitahes kõrget järku tuletised; jaotusel P (juhuslikul suurusel X) eksisteerivad aga kuitahes kõrged momendid, kusjuures seose (13.2) abil on kõik momendid määratud funktsiooni M tuletiste kaudu.
Nüüd on selge, millest tuleb nimetus – momente genereeriv funktsioon.

13.3 Näited

1. Normaaljaotus Olgu $P = N(\mu, \sigma^2)$. Vastavalt definitsioonile

$$M(t) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^{tx} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \frac{1}{\sqrt{2\pi}} e^{\mu t} \int_{-\infty}^{\infty} e^{\sigma u t - \frac{u^2}{2}} du,$$

kus $u = \frac{x-\mu}{\sigma}$. Et $\sigma u t - \frac{u^2}{2} = \frac{-(u-\sigma t)^2}{2} + \frac{\sigma^2 t^2}{2}$, saame

$$M(t) = e^{\mu t + \frac{\sigma^2 t^2}{2}} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\frac{-(u-\sigma t)^2}{2}} du = e^{\mu t + \frac{\sigma^2 t^2}{2}}.$$

Leiame tuletised. Induktsiooni abil on kerge veenduda, et

$$\begin{aligned}
M'(t) &= M(t)(\mu + \sigma^2 t), \\
M''(t) &= M'(t)(\mu + \sigma^2 t) + \sigma^2 M(t), \\
M'''(t) &= M''(t)(\mu + \sigma^2 t) + 2\sigma^2 M'(t), \\
&\cdots \\
M^{(n)}(t) &= M^{(n-1)}(t)(\mu + \sigma^2 t) + (n-1)\sigma^2 M^{(n-2)}(t).
\end{aligned}$$

Normaaljaotuse momendid avalduvad seega

$$\begin{aligned}
EX &= M'(0) = \mu, \quad EX^2 = M''(0) = \mu^2 + \sigma^2, \\
EX^3 &= M'''(0) = \mu^3 + 3\sigma^2\mu, \quad EX^4 = M^{(4)}(0) = \mu^4 + 6\sigma^2\mu^2 + 3\sigma^4, \\
&\cdots \\
EX^n &= M^{(n)}(0) = \mu EX^{n-1} + (n-1)\sigma^2 EX^{n-2}.
\end{aligned}$$

Juhul, kui $\mu = 0$, $\sigma^2 = 1$, saame $M(t) = e^{\frac{t^2}{2}}$ ja $EX^n = (n-1)EX^{n-2}$, $n = 1, 2, \ldots$. Et $EX = 0$, siis $EX^n = 0$ iga paarituarvulise n korral, paarisarvulised momendid aga avalduvad

$$EX^2 = 1,\ EX^4 = 3,\ EX^6 = 15,\ EX^8 = 7 \cdot 15,\ \cdots, EX^{2n} = 1 \cdot 3 \cdot 5 \cdots (2n-1).$$

2. Eksponentjaotus $P = E(\lambda)$. Vastavalt definitsioonile

$$M(t) = \int_0^{\infty} e^{tx} \lambda e^{-\lambda x} dx = \lambda \int_0^{\infty} e^{-(\lambda - t)x} dx = \frac{\lambda}{\lambda - t},$$

kui $t < \lambda$. Seega $T = (-\infty, \lambda)$ ja sellel piirkonnal $M(t) = \frac{\lambda}{\lambda - t}$.
Leiame tuletised

$$M'(t) = \frac{\lambda}{(\lambda - t)^2},\ M''(t) = \frac{2\lambda}{(\lambda - t)^3}, \ldots, M^{(n)}(t) = \frac{n!\lambda}{(\lambda - t)^{n+1}}.$$

Seega $EX^n = M^{(n)}(0) = n!\lambda^{-n}$, $n = 1, 2, \ldots$.

3. Poissoni jaotus $P = P_o(\lambda)$. Vastavalt definitsioonile

$$M(t) = \sum_{n=0}^{\infty} e^{nt} e^{-\lambda} \frac{\lambda^n}{n!} = e^{-\lambda} \sum_{n=0}^{\infty} e^{tn} \frac{\lambda^n}{n!} = e^{-\lambda} \sum_{n=0}^{\infty} \frac{(\lambda e^t)^n}{n!} = e^{-\lambda + e^t \lambda} = \exp[\lambda(e^t - 1)].$$

Seega $T = (-\infty, \infty)$ ja $M(t) = \exp[\lambda(e^t - 1)]$.
Leiame tuletised

$$\begin{aligned} M'(t) &= M(t)\lambda e^t \\ M''(t) &= M'(t)\lambda e^t + M(t)\lambda e^t = \lambda e^t[M'(t) + M(t)] \\ M'''(t) &= M''(t)\lambda e^t + M'(t)\lambda e^t + M(t)\lambda e^t = \lambda e^t[M''(t) + 2M'(t) + M(t)]. \end{aligned}$$

Esimesed momendid avalduvad seega

$$EX = \lambda,\ EX^2 = \lambda^2 + \lambda,\ EX^3 = \lambda^3 + 3\lambda^2 + \lambda.$$

13.4 Sõltumatute juhuslike suuruste summa momente genereeriv funktstioon

Olgu $X_1, \ldots, X_k$ tõenäosusruumil $(\Omega, \mathcal{F}, \mathbf{P})$ antud *sõltumatud* juhuslikud suurused, nende momente genereerivad funktsioonid olgu $M_1, \ldots, M_k$.
Oletame, et funktsioonid M_i on lõplikud punkti 0 ümbruses. Seega $\exists t_0 > 0$ nii, et iga $t \in (-t_0, t_0)$ korral $M_i(t) < \infty$, $i = 1, \ldots, k$.
Juhuslikud suurused $e^{tX_1}, \ldots, e^{tX_k}$ on sõltumatud, mistõttu teoreemi 8.16 põhjal $\forall t \in (-t_0, t_0)$ korral

$$E(e^{t \sum_{i=1}^k X_i}) = E(e^{tX_1} \cdots e^{tX_k}) = E(e^{tX_1}) \cdots E(e^{tX_k}) = M_1(t) \cdots M_k(t) < \infty.$$

Seega on juhusliku suuruse $X_1 + \cdots + X_k$ summa momente genereeriv funktsioon M lõplik lõigul $(-t_0, t_0)$ ja iga t korral sellest vahemikust kehtib seos

$$M(t) = M_1(t) \cdots M_k(t).$$

13.5 Momentide probleem

Olgu m_1, m_2, ... juhusliku suuruse X (jaotuse P) momendid. **Momentide probleem** on järgmine:

Kas momendid $\{m_n\}_{n=1}^{\infty}$ määravad üheselt jaotuse P?

Teisisõnu: kas leidub jaotus Q nii, et $P \neq Q$ kuid

$$\int x^n dP = \int x^n dQ, \quad n = 1, 2, \ldots?$$

Üldiselt momendid üheselt jaotust ei määra. Teatud tingimustel aga küll.

Teoreem 13.2 *Olgu P selline, et astmereal*

$$\sum_n \frac{m_n t^n}{n!} \tag{13.3}$$

on positiivne koondumisraadius. Siis määravad momendid $\{m_n\}_{n=1}^{\infty}$ üheselt jaotuse P.

Selle teoreemi tõestus põhineb karakteristlikel funktsioonidel ja selle tõestuse võib leida raamatust (*Billingsley* trm 30.1).
Olgu $\mu_1, \mu_2, \ldots$ jaotuse P absoluutsed momendid, st

$$\mu_n = \int |x|^n dP.$$

Teoreemist 13.2 järeldub järgmine teoreem (*Shirjajev* II osa 12 (teoreem 1).)

Teoreem 13.3 *Olgu P selline, et astmereal*

$$\sum_n \frac{\mu_n t^n}{n!} \tag{13.4}$$

on positiivne koondumisraadius. Siis momendid määravad $\{m_n\}_{n=1}^{\infty}$ üheselt jaotuse P.

Tõestus. Iga n korral $|m_n| \leq \mu_n$, mistõttu rea (13.4) koondumisest järeldub rea

$$\sum_n \frac{|m_n| t^n}{n!}$$

koondumine ja sellest järeldub rea (13.3) koondumine. ■

Järeldus 13.1 *Kui*

$$\limsup_n \frac{\mu_n^{\frac{1}{n}}}{n} < \infty, \tag{13.5}$$

siis real (13.4) on positiivne koondumisraadius ja momendid $\{m_n\}$ määravad üheselt P.

Tõestus. Olgu $t_o > 0$ nii väike, et

$$\limsup_n \frac{(\mu_n)^{\frac{1}{n}}}{n} < (et_o)^{-1},$$

millest

$$\limsup_n \frac{(\mu_n t_o^n)^{\frac{1}{n}}}{n} < e^{-1}.$$

Stirlingi valemist:

$$\lim_n \frac{n!}{\sqrt{2\pi n} e^{-n} n^n} = 1$$

järeldub (kui jada $a_n \to 1$, siis $(a_n)^{\frac{1}{n}} \to 1$), et

$$\lim_n \frac{(n!)^{\frac{1}{n}}}{(2\pi n)^{\frac{1}{2n}} e^{-1} n} = 1,$$

millest saame

$$\lim \frac{(n!)^{\frac{1}{n}}}{n} = e^{-1}.$$

Seega

$$\limsup_n \Big(\frac{\mu_n t_o^n}{n!}\Big)^{\frac{1}{n}} < 1.$$

Cauchy koondumiskriteeriumist (positiivne rida $\sum_n u_n$ koondub, kui $\limsup_n (u_n)^{\frac{1}{n}} < 1$) saame, et

$$\sum_n \frac{\mu_n t_o^n}{n!} < \infty.$$

See aga tähendab, et iga $|t| \leq t_o$ korral

$$\sum_n \frac{\mu_n |t^n|}{n!} \leq \sum_n \frac{\mu_n t_o^n}{n!} < \infty$$

ehk astmerea (13.4) koondumisraadius on vähemalt t_o. ■

Paneme tähele – kui $X \sim P$ on juhuslik suurus, siis (13.5) on

$$\limsup_n \frac{\|X\|_n}{n} < \infty.$$

See tingimus on täidetud näiteks siis kui jaotuse P kandja on tõkestatud.

Järeldus 13.2 *Olgu M jaotuse P momente genereeriv funktsioon. Kui M on lõplik punkti 0 mingis ümbruses, siis jaotus P on üheselt määratud oma momentidega. Et momendid on määratud funktsiooniga M, määrab viimane üheselt jaotuse.*

Tõestus. Olgu $t_o > 0$ selline, et $M(t) < \infty\ \forall |t| \leq t_o$ korral. Teame aga, et

$$\sum_n \frac{\mu_n |t|^n}{n!} = \sum_n \int \Big|\frac{(tx)^n}{n!}\Big| P(dx) = \int \sum_{n=0}^{\infty} \Big|\frac{(tx)^n}{n!}\Big| P(dx) \leq M(t) + M(-t) < \infty.$$

Teoreemist 13.3 järeldub nüüd, et momendid määravad jaotuse üheselt. ∎

Järeldus 13.3 *Kui*

$$\limsup_n \frac{(m_{2n})^{\frac{1}{2n}}}{2n} < \infty,$$

siis määravad momendid jaotuse üheselt.

Tõestus. Ülesanne. ∎

13.6 Kumulante genereeriv funktsioon

13.6.1 Kumulante genereeriv funktsioon, kumulandid

Def 13.4 *Olgu $M : \mathbb{R} \to [0, \infty]$ juhusliku suuruse X (või jaotuse P) momente genereeriv funktsioon. Funktsiooni*

$$\Lambda : \mathbb{R} \to (-\infty, \infty], \quad \Lambda(t) = \ln M(t)$$

nimetatakse juhusliku suuruse X (jaotuse P) **kumulante genereerivaks funktsiooniks**.

Seega kumulante genereeriv funktsioon on momente genereeriva funktsiooni logaritm. Kumulante genereeriv funktsioon on lõplik parajasti siis, kui momente genereeriv funktsioon on lõplik, seega piirkonnal T. On selge, et $\Lambda(0) = 0$.

Väide 13.1 *Piirkonnas T on kumulante genereeriv funktsioon kumer.*

Tõestus. Tuletame meelde Hölderi võrratuse: kui $E|Z|^p < \infty$ ja $E|Y|^q < \infty$, kusjuures $\frac{1}{p} + \frac{1}{q} = 1$, siis

$$EZY \leq \big(E|Z|^p\big)^{\frac{1}{p}} \big(E|Y|^q\big)^{\frac{1}{q}}.$$

Olgu $t_1, t_2 \in T$. See tähendab, et $M(t_1) = Ee^{t_1 X} < \infty$ ja $M(t_2) = Ee^{t_2 X} < \infty$. Valime $\lambda \in (0,1)$ ja defineerime $p := \frac{1}{\lambda}$ ja $q := \frac{1}{1-\lambda}$. Nüüd on selge, et $\frac{1}{p} + \frac{1}{q} = 1$. Defineerime juhuslikud suurused

$$Z = e^{\lambda t_1 X}, \quad Y = e^{(1-\lambda) t_2 X}.$$

Veendu, et $E|Z|^p < \infty$ ja $E|Y|^q < \infty$. Hölderi võrratusest saame nüüd, et

$$M(\lambda t_1 + (1-\lambda) t_2) = E\big(e^{\lambda t_1 X + (1-\lambda) t_2 X}\big) = EZY \leq \big(E(Z^{\frac{1}{\lambda}})\big)^{\lambda} \big(E(Y^{\frac{1}{1-\lambda}})\big)^{1-\lambda} = M(t_1)^{\lambda} M(t_2)^{1-\lambda},$$

millest

$$\Lambda(\lambda t_1 + (1-\lambda) t_2) = \ln M(\lambda t_1 + (1-\lambda) t_2) \leq \lambda \ln M(t_1) + (1-\lambda) \ln M(t_2). \tag{13.6}$$

∎

Seega piirkonnal T on Λ kumer, sest (13.6) kehtib. Väljaspool piirkonda T (st, kui t_1 või t_2 ei kuulu hulka T), kehtib (13.6) triviaalselt (miks?), seega võib öelda, et Λ on kogu reaalteljel kumer funktsioon.

Teame, et kui momente genereeriv funktsioon M on lõplik punkti 0 mingis ümbruses, on funktsioonil M punktis 0 kõik tuletised mis võrduvad juhusliku suuruse X momentidega (sellest järeldub, et tarvilik tingimus M lõplikkuseks punkti 0 ümbruses on kõikide momentide olemasolu.) Kui M on lõplik punkti 0 mingis ümbruses, on seda ka Λ ning M lõpmatust diferentseeruvusest punktis 0 järeldub Λ lõpmatu diferentseeruvus. Seega, kui Λ on lõplik punkti 0 mingis ümbruses on tal punktis 0 kõik tuletised. Sellisel juhul nimetatakse tuletist $\Lambda^{(m)}(0)$ juhusliku suuruse X (või jaotuse P) m-**ndaks kumulandiks** ja tähistatakse harilikult $\kappa_m(X)$.

Ülesanne: Olgu juhusliku suuruse kumulante genereeriv funktsioon lõplik punkti 0 mingis ümbruses. Leida juhusliku suuruse esimene ja teine kumulant, $\kappa_1(X)$ ja $\kappa_2(X)$.

13.7 Suurte hälvete printsiip

Olgu $X_1, X_2, \ldots$ jaotusega P iid juhuslikud suurused, $EX_i = \mu < \infty$. $S_n = X_1 + \cdots + X_n$. Uurime tõenäosust $\mathbf{P}(\frac{S_n}{n} \geq c) = \mathbf{P}(S_n \geq cn)$, kus $c > \mu$. Höffdingi võrratusest teame, et kui juhuslikud suurused X_i on tõkestatud, st leiduvad konstandid $a < b$ nii, et $a \leq X_i \leq b$ p.k., siis

$$\mathbf{P}(\frac{S_n}{n} \geq c) = \mathbf{P}(S_n \geq cn) = \mathbf{P}(S_n - n\mu \geq (c-\mu)n) \leq \exp[-\frac{2(c-\mu)^2 n}{(b-a)^2}],$$

teatud tingimustel annab parema hinnangu Bernsteini võrratus. Milline on aga parim võimalik hinnang?
Olgu M jaotuse P momente genereeriv funktsioon. Hindame tõenäosust $\mathbf{P}(S_n \geq cn)$ Tŝernovi meetodil. Olgu $t > 0$.

$$\mathbf{P}(S_n \geq cn) = \mathbf{P}(e^{tS_n} \geq e^{tcn}) \leq \frac{E(e^{tS_n})}{e^{tcn}} = \frac{M^n(t)}{e^{tcn}} = \exp[n(\ln M(t) - ct)] = \exp[-n(ct - \Lambda(t))].$$

Seega iga $t \geq 0$ korral

$$\mathbf{P}(S_n \geq cn) \leq \exp[-n(ct - \Lambda(t))]. \tag{13.7}$$

Võrratuse (13.7) vasak pool ei sõltu t-st. Meid huvitab t, mille korral võrratuse parem pool on võimalikult väike. Seega maksimiseerime funktsiooni $t \mapsto (ct - \Lambda(t))$ üle $(0, \infty)$. On selge, et antud juhul on maksimiseerimine üle hulga $(0, \infty)$ on ekvivalentne maksimiseerimisega üle hulga $[0, \infty)$ (miks?). Järgnevas veendume, et antud juhul võime maksimiseerida üle kogu reaaltelje.

Def 13.5 *Funktsiooni*

$$\Lambda^* : \mathbb{R} \to [0, \infty), \quad \Lambda^*(x) = \sup_{t \in \mathbb{R}} (xt - \Lambda(t)) \tag{13.8}$$

nimetatakse funktsiooni Λ **Legendre-Fencheli** *teisenduseks.*

Funktsiooni Λ^* omadused:

- Definitsioonis (13.8) võib maksimiseerimise üle hulga $\mathbb{R}$ asendada maksimiseerimisega üle hulga T. Tõepoolest, kui $t \notin T$, siis $xt - \Lambda(t) = -\infty$, millest

$$\Lambda^*(x) = \sup_{t \in \mathbb{R}}(xt - \Lambda(t)) = \sup_{t \in T}(xt - \Lambda(t)). \tag{13.9}$$

- Λ^* on mittenegatiivne (miks?).

- Λ^* on kumer. Tõepoolest, iga $\lambda \in (0,1)$ korral

$$\begin{aligned}
\Lambda^*(\lambda x + (1-\lambda)y) &= \sup_{t \in \mathbb{R}} \big(\lambda xt - \lambda\Lambda(t) + (1-\lambda)yt - (1-\lambda)\Lambda(t)\big) \\
&= \sup_{t \in \mathbb{R}} \big(\lambda(xt - \Lambda(t)) + (1-\lambda)(yt - \Lambda(t))\big) \\
&\leq \lambda \sup_{t \in \mathbb{R}}(xt - \Lambda(t)) + (1-\lambda) \sup_{t \in \mathbb{R}}(yt - \Lambda(t)) \\
&= \lambda\Lambda^*(x) + (1-\lambda)\Lambda^*(y).
\end{aligned}$$

- Funktsioon $t \mapsto xt - \Lambda(t)$ on nõgus. See tähendab, et kui leidub t_x nii, et $\Lambda'(t_x) = x$, saavutab $t \mapsto xt - \Lambda(t)$ selles punktis maksimumi, siis

$$\Lambda^*(x) = xt_x - \Lambda(t_x). \tag{13.10}$$

  See valem aitab funktsiooni $\Lambda^*(x)$ leida, (vt näited allpool).

- Kui Λ (ja ka M) on lõplik punkti 0 mingis ümbruses (st $\exists t_o > 0$ nii, et $(-t_o, t_o) \subset T$), siis $\Lambda'(0) = \mu$, st $t_\mu = 0$, millest valemi (13.10) kaudu

$$\Lambda^*(\mu) = 0 - \Lambda(0) = 0.$$

  Seega Λ^* on kumer mittenegatiivne funktsioon, mis saavutab oma miinimumi (väärtuse 0) punktis μ. Selgub, et μ on ainus Λ^* miinimumkoht (vt järelduse 13.4 tõestus).

- Funktsioon $\Lambda(t)$ on kumer, kusjuures $\Lambda(0) = 0$. Kui ta on lõplik punkti 0 ümbruses, siis $\Lambda'(0) = \mu$ ja kumera funktsiooni omaduste tõttu iga t korral

$$\Lambda(t) - \Lambda(0) \geq \mu(t - 0) \quad \Leftrightarrow \quad \Lambda(t) \geq \mu t.$$

  Kui nüüd $c > \mu$, siis $ct - \Lambda(t) = (c-\mu)t + (\mu t - \Lambda(t)) \leq (c-\mu)t$. Kui aga $t < 0$, siis ülaltoodud võrratustest järeldub, et $ct - \Lambda(t) < 0$. Järelikult sellisel juhul ei saa funktsiooni $ct - \Lambda(t)$ maksimiseerida negatiivne t ehk: kui $c > \mu$, siis

$$\sup_{t \in \mathbb{R}}(ct - \Lambda(t)) = \sup_{t \geq 0}(ct - \Lambda(t)) \tag{13.11}$$

Väide 13.2 *Olgu $X_1, X_2, \dots$ iid jaotusega P juhuslikud suurused. Olgu P selline, et tema momente genereeriv funktsioon on lõplik punkti 0 mingis ümbruses. Sellisel jaotusel on lõplik keskväärtus μ. Olgu $c > \mu$. Siis kehtib võrratus*

$$\mathbf{P}(\frac{S_n}{n} \geq c) \leq \exp[-n\Lambda^*(c)]. \tag{13.12}$$

Tõestus. Nagu just nägime, antud tingimistel kehtib (13.11). Väide (13.12) järeldub nüüd väitest (13.7). ■

Järeldus 13.4 *Väite 13.2 eeldustel on iga $c > \mu$ korral $\Lambda^*(c) > 0$, st tõenäosus $\mathbf{P}(\frac{S_n}{n} \geq c)$ koondub nulliks eksponentsiaalselt.*

Tõestus. Kumerate funktsioonide teooriast on teada, et kui leidub arv m nii, et $\Lambda(t) - \Lambda(0) \geq m(t-0)$ iga t korral, siis $m \in [\Lambda'(0-), \Lambda'(0+)]$. Seega, kui Λ on diferentseeruv punktis 0 ja $\Lambda(t) \geq mt$ iga t korral, siis $m = \Lambda'(0)$. Antud eeldustel $\Lambda'(0) = \mu$, mistõttu $\Lambda(t) \geq \mu t$ iga $t \in \mathbb{R}$ korral ning iga $c \neq \mu$ korral leidub t nii, et $\Lambda(t) < ct$. ■

Märkus: Väide 13.2 kehtib ka selliste jaotuste korral, mille momente genereeriv funktsioon pole lõplik nulli ümbruses. See on nii sellepärast, et võrratus (13.7) kehtib iga jaotuse korral ning sellest järeldub, et

$$\mathbf{P}(\frac{S_n}{n} \geq c) \leq \exp\big[-n\big(\sup_{t \geq 0}(ct - \Lambda(t))\big)\big].$$

Selgub, et kui $c > \mu$, siis võrratus (13.11) kehtib ka üldisemal juhul. Veendume selles. Tõepoolest, kui $T \subset [0, \infty)$, siis (13.11) järeldub seosest (13.9). Märgime, et kui leidub $0 < t_o \leq \infty$ nii, et $T = [0, t_o)$ või $T = [0, t_o]$ st M on lõplik punkti 0 parempoolses ümbruses, siis pole raske näidata, et $M'(0+) = \Lambda(0+) = \mu$. Järelduse 13.4 tõestuses toodud argumendist järeldub, et $\Lambda^*(c) > 0$, kui $c > \mu$. Kui M on lõplik punkti 0 vasakpoolses ümbruses, st leidub $0 < t_o \leq \infty$ nii, et $T = (-t_o, 0]$ või $T = [t_o, 0]$, siis pole raske näidata, et $M'(0-) = \Lambda(0-) = \mu$, millest $\Lambda(t) \geq \mu t$ iga $t \in \mathbb{R}$ korral. Kui nüüd $c > \mu$ ja $t < 0$, siis $ct - \Lambda(t) = (c-\mu)t + (\mu t - \Lambda(t)) \leq (c-\mu)t < 0$, millest järeldub (13.11). Sellisel juhul aga $\Lambda^*(c) = 0$, sest iga $t > 0$ korral $\Lambda(t) = \infty$ ning võrratuse (13.12) parem pool on 1 (võrratus on triviaalne). Seega võrratus (13.11) kehtib tõepoolest iga T korral. Veel enam, et $\Lambda^*(c) > 0$ vaid siis kui T on lõplik punkti 0 parempoolses ümbruses, saame järeldust 13.4 üldistada järgmiselt: kui T on lõplik punkti 0 parempoolses ümbruses (st leidub $0 < t_o \leq \infty$ nii, et $T = [0, t_o)$ või $T = [0, t_o]$), siis $\mathbf{P}(\frac{S_n}{n} \geq c)$ koondub nulliks eksponentsiaalselt.

Näited (L-F teisendus):

- Olgu P standardne normaaljaotus $\mathcal{N}(0,1)$. Teame, et siis $M(t) = e^{\frac{t^2}{2}}$ ning seega $\Lambda(t) = \frac{t^2}{2}$. Et $\Lambda'(t) = t$, siis $\Lambda'(t) = x$ parajasti siis kui $t = x$, millest

$$\Lambda^*(x) = \sup_t (xt - \Lambda(t)) = x^2 - \frac{x^2}{2} = \frac{x^2}{2}.$$

- Olgu P Bernoulli jaotus $B(1,p)$ Tema momente genereeriv funktsioon on $M(t) = pe^t + (1-p)$ (tõesta!), millest $\Lambda(t) = \ln(pe^t + (1-p))$. Pane tähele, et iga t korral

$$0 < \Lambda'(t) = \frac{pe^t}{pe^t + (1-p)} < 1.$$

see tähendab, et kui $x \notin (0,1)$, siis ei leidu sellist reaalarvu t_x, et $\Lambda'(t) = x$. Sellest järeldub, et $\Lambda^*(x) = \infty$, kui $x \notin (0,1)$. Olgu nüüd $x \in (0,1)$. Iga sellise x korral leidub reaalarv t_x nii, et $\Lambda'(t_x) = x$ (miks?). Seega

$$x = \frac{pe^{t_x}}{pe^{t_x} + (1-p)}.$$

Lahendades selle võrrandi, saame

$$pe^{t_x} = \frac{x(1-p)}{1-x}, \quad t_x = \ln\big(\frac{x(1-p)}{(1-x)p}\big).$$

Seega, kui $x \in (0,1)$, siis (veendu selles!)

$$\Lambda^*(x) = xt_x - \ln(pe^{t_x} + (1-p)) = x\ln\big(\frac{x}{p}\big) + (1-x)\ln\big(\frac{1-x}{1-p}\big).$$

Kui $p = 0.5$, siis $\Lambda^*(0.7) \approx 0.0822$. Seega, kui ausat münti visatakse 1000 korda, siis tõenäosusele, et vähemalt 700 korda tuleb kiri saame Väite põhjal järgmise ülemise hinnangu (veendu!)

$$\mathbf{P}(S_{1000} \geq 700) = \mathbf{P}(S_{1000} \geq 0.7 \cdot 1000) \leq e^{-1000\Lambda^*(0.7)} \approx e^{-82}.$$

Leiame analoogilise hinnangu Höffdingi võrratuse abil:

$$P(S_{1000} \geq 700) = \mathbf{P}\Big(\frac{S_{1000}}{1000} - \frac{1}{2} \geq 0.2\Big) \leq e^{-2\cdot 1000\cdot(0.2)^2} = e^{-80}.$$

Viimases näites nägime, et hinnang (13.12) oli vaadeldava näite korral natuke parem kui Höffdingi võrratuse abil leitud hinnang. On see alati nii? Kas neid hinnanguid saab parandada?
Paneme tähele, et võrratuse (13.12) võime kirjutada kujul

$$\frac{1}{n}\ln\mathbf{P}(\frac{S_n}{n} \geq c) \leq -\Lambda^*(c).$$

Seega iga $c > \mu$ korral

$$\limsup_n \frac{1}{n} \ln \mathbf{P}(\frac{S_n}{n} \geq c) \leq -\Lambda^*(c).$$

Järgmine väga oluline teoreem väidab, et $-\Lambda^*(c)$ on tegelikult jada $\frac{1}{n} \ln \mathbf{P}(S_n \geq c)$ piirväärtus. See tähendav sisuliselt, et võrratus (13.7) ei kehti suurema konstandiga kui $\Lambda^*(c)$.

Teoreem 13.6 *(***suurte hälvete printsiip***) Olgu $X_1, X_2, \ldots$iid juhuslikud suurused jaotusega P. Olgu jaotusel P lõplik keskväärtus $\mu < \infty$. Siis iga $c > \mu$ korral*

$$\lim_n \frac{1}{n} \ln \mathbf{P}(\frac{S_n}{n} \geq c) = -\Lambda^*(c) \tag{13.13}$$

Tõsestuse võib leida (juhul kui M on lõplik) *Grimmet, Stirzaker* trm 5.11, üldisel juhul *Kallenberg*, trm 27.3.

Teoreem 13.6 uurib tõenäosust

$$\mathbf{P}(\frac{S_n}{n} \geq c) = \mathbf{P}\big(\frac{S_n}{n} \in [c, \infty)\big).$$

Tihti pakub aga huvi üldisem tõenäosus $\mathbf{P}\big(\frac{S_n}{n} \in B\big)$, kus B on mingi Boreli hulk. Selgub, et teoreemi 13.6 on võimalik üldistada ka selles suunas. Järgnevas olgu B^o hulga sisemus ning $\bar{B}$ selle hulga sulund.

Teoreem 13.7 (suurte hälvete printsiip) *Olgu $X_1, X_2, \ldots$ iid juhuslikud suurused jaotusega P. Olgu P kumulante genereeriv funktsioon Λ lõplik reaalteljel (st $T = \mathbb{R}$). Siis iga $B \in \mathcal{B}(\mathbb{R})$ korral*

$$-\inf_{x \in B^o} \Lambda^*(x) \leq \liminf \frac{1}{n} \ln \mathbf{P}\big(\frac{S_n}{n} \in B\big) \leq \limsup \frac{1}{n} \ln \mathbf{P}\big(\frac{S_n}{n} \in B\big) \leq -\inf_{x \in \bar{B}} \Lambda^*(x). \tag{13.14}$$

Tõestus: *Kallenberg, 27.*

Pane tähele, et kui $B = [c, \infty)$, siis $B^o = (c, \infty)$ ja $\bar{B} = [c, \infty)$. Teoreemi eeldustel on Λ lõplik ja siis $\Lambda^*(\mu) = 0$, mistõttu iga $c > \mu$ korral Λ^* on hulgal $[c, \infty)$ mittekahanev. Samuti on Λ^* kui kumer funktsioon pidev. Seega

$$\inf_{x \in (c,\infty)} \Lambda^*(x) = \lim_{x \searrow c} \Lambda^*(x) = \Lambda^*(c) = \inf_{x \in [c,\infty)} \Lambda^*(x)$$

ehk $B = [c, \infty)$ korral (13.14) on (13.13).

Kirjandus. Momente genereerivate funktsioonide kohta loe *Billingsley*, 21. Suurte hälvete seadusest loe *Grimmet ja Stirzaker*, 5 ning *Kallenberg*, 27.

14 Karakteristlikud funktsioonid

14.1 Kompleksfunktsioonid

Olgu (S, Σ, μ) mõõdu ruum, $f : S \to \mathbb{C}$ kompleksväärtuseline funktsioon (kompleksfunktsioon). Funktsiooni f võib alati esitada kujul $f = g + ih$, kus $g, h : S \to \mathbb{R}$ on reaalväärtuselised funktsioonid. Ütleme, et kompleksfunktsioon f on Σ-mõõtuv, kui funktsioonid g, h on Σ-mõõtuvad. Sellisel juhul defineerime kompleksfunktsiooni f Lebesgue'i integraali järgmiselt:

$$\int f d\mu = \int (g + ih) d\mu := \int g d\mu + i \int h d\mu.$$

Seega f on integreeruv, kui funktsioonid h ja g on integreeruvad. Seosest $\max\{|g|, |h|\} \leq |f| \leq |g| + |h|$ on aga kerge näha, et funktsioonid g ja h on integreeruvad parajasti siis, kui (mittenegatiivne ja mõõtuv) funktsioon $|f|$ on integreeruv, st $\int |f| d\mu < \infty$.

Üldiselt on kompleksfunktsiooni integraalil samasugused omadused mis reaalfunktsiooni integraalil:

- $\int (\alpha f + \beta g) d\mu = \alpha \int f d\mu + \beta \int g d\mu$, kui $\alpha, \beta \in \mathbb{C}$ (miks?)
- $|\int f d\mu| \leq \int |f| d\mu$ (vt *Billingsley* lk. 219)
- $\int |f| d\mu = 0 \Leftrightarrow f = 0$ p.k. (miks?).

Ka kehtib kompleksfunktsiooni integraalide korral domineeritud koondumise teoreem. Tihti kasutame (komplekset) domineeritud koondumise teoreemi kujul:

Teoreem 14.1 (DOM.) *Olgu $f_h : S \to \mathbb{C}$, $h \in \mathbb{R}$ kompleksfunktsioonide pere, kusjuures iga $s \in S$ korral $f_h(s) \to f_0(s)$, kui $h \to 0$. Kui leidub $g : S \to \mathbb{C}$ nii, et iga $h \in \mathbb{R}$ ja iga $s \in S$ korral $|f_h(s)| \leq |g(s)|$ ja $\int |g| d\mu < \infty$, siis*

$$\lim_{h \to 0} \int f_h(s) \mu(ds) = \int \lim_{h \to 0} f_h(s) \mu(ds) = \int f_0(s) \mu(ds).$$

Teoreem järeldub harilikust (reaalsest) domineeritud koondumise teoreemist, kui seda eraldi rakendada f_h reaal- ja imaginaarosale.

Funktsiooni $f : \mathbb{C} \to \mathbb{C}$ nimetatakse **kompleksmuutuja funktsiooniks**. Analüütilise (astmeritta arendatava) funktsiooni laiend kompleksatasandile defineeritakse Taylori reaksarenduse kaudu. Seega kompleksne astmefunktsioon $\exp : \mathbb{C} \to \mathbb{C}$ defineeritakse $e^z := \sum_{n=0}^{\infty} \frac{z^n}{n!}$. Erijuhul, kui $z = ix$, saame

$$e^{ix} := 1 + (ix) + \frac{(ix)^2}{2!} + \frac{(ix)^3}{3!} + \cdots . \tag{14.1}$$

Vastavalt Euleri valemile kehtib seos

$$e^{ix} = \cos x + i \sin x.$$

Funktsioonid $\cos x$ ja $\sin x$ on pidevad ja seetõttu Boreli mõõtuvad. Seega ka kompleksfunktsioon e^{ix} on mõõtuv. Seosest $|e^{ix}| = \sqrt{\cos^2 x + \sin^2 x} = 1$ järeldub, et iga reaalteljel antud tõenäosusmõõdu P korral $\int |e^{ix}| P(dx) = 1$ ehk kompleksfunktsioon $f(x) = e^{ix}$ on integreeruv.

Märgime veel, et suvalise kompleksarvu $z = a + ib$ korral

$$e^z = e^a e^{ib} = e^a(\cos b + i \sin b).$$

Trigonomeetriavalemeid kasutades on lihtne veenduda, et suvaliste $z_1, z_2 \in \mathbb{C}$ korral kehtib

$$e^{z_1+z_2} = e^{z_1} e^{z_2}. \tag{14.2}$$

14.2 Karakteristlikud funktsioonid

Olgu $X \sim P$ tõenäosusruumil $(\Omega, \mathcal{F}, \mathbf{P})$ antud juhuslik suurus, $t \in \mathbb{R}$. Vaatleme kompleksfunktsiooni

$$e^{itX} : \Omega \to \mathbb{C}.$$

Euleri valemist saame

$$e^{itX} = \cos tX + i \sin tX,$$

mistõttu vaadeldav funktsioon on mõõtuv ja integreeruv.

Def 14.2 *Juhusliku suuruse X (jaotuse P) karakteristlikuks funktsiooniks nimetatakse funktsiooni*

$$\varphi : \mathbb{R} \to \mathbb{C}, \quad \varphi(t) = E(e^{itX}) = E(\cos tX) + iE(\sin tX).$$

Juhul, kui on tarvis näidata ka juhuslikku suurust, kasutatakse tähistust φ_X.

Lebesgue'i integraali muutuja vahetuse kohaselt

$$\varphi(t) = \int \cos tx P(dx) + i \int \sin tx P(dx) = \int e^{itx} P(dx).$$

Nagu iga keskväärtus, sõltub ka karakteristlik funktsioon vaid juhusliku suuruse X jaotusest.

Karakteristlik funktsioon on momente genereeriva funktsiooni analoog imaginaartasandil, erinevalt momente genereerivast funktsioonist on aga karakteristlik funktsioon lõplik (defineeritud) iga $t \in \mathbb{R}$ korral. Matemaatilises analüüsis nimetatakse karakteristlikku funktsiooni ka Fourier' teisenduseks.

Karakteristlikud funktsioonid on väga kasulik abivahend tõenäosusteoreetiliste probleemide lahendamisel; eelkõige põhineb nende kasulikkus järgmistel omadustel:

- sõltumatute juhuslike suuruste summa karakteristlik funktsioon on liidetavate karakteristlike funktsioonide korrutis;

- karakteristlikud funktsioonid on üksüheses vastavuses jaotustega;
- karakteristlike funktsioonide punktiviisilisest koondumisest järeldub vastavate jaotuste nõrk koondumine.

14.3 Elementaarsed omadused

Teoreem 14.3 *Karakteristlikel funktsioonidel on omadused*

1) $\varphi(0) = 1$;

2) $|\varphi(t)| \leq 1$;

3) φ *on ühtlaselt pidev;*

4) $\varphi(-t) = \overline{\varphi}(t) \quad \forall t \in \mathbb{R}$;

5) *kui* $a, b \in \mathbb{R}$, *siis* $\varphi_{aX+b}(t) = e^{itb}\varphi_X(at) \quad \forall t \in \mathbb{R}$.

Tõestus. Omadused **1)** ja **2)** on triviaalsed.
3) Iga $t, h \in \mathbb{R}$ korral (vt kompleksfunktsiooni integreerimine ja (14.2))

$$|\varphi(t+h) - \varphi(t)| = \Big| \int (e^{(t+h)ix} - e^{tix})P(dx) \Big| \leq \int |e^{(t+h)ix} - e^{tix}|P(dx) =$$
$$\int |e^{tix}(e^{ihx} - 1)|P(dx) = \int |e^{ihx} - 1|P(dx) =: \psi(h).$$

Kui $h \to 0$, siis

$$f_h(x) := |e^{ihx} - 1| \to 0.$$

Et $|f_h(x)| \leq |e^{ihx}| + 1 = 2$, siis tõkestatud koondumise teoreemi põhjal integraal

$$\psi(h) = \int f_h dP \to 0$$

protsessis $h \to 0$. Seega φ on pidev; et aga $\psi(h)$ ei sõltu t-st, on φ ühtlaselt pidev.

4)

$$\varphi(-t) = \int \cos(-tx)P(dx) + i\int \sin(-tx)P(dx) = \int \cos tx P(dx) - i\int \sin tx P(dx) = \overline{\varphi}(t).$$

5)

$$\varphi_{aX+b}(t) = E(e^{it(aX+b)}) = e^{itb}E(e^{iatX}) = e^{itb}\varphi(at).$$

■

14.4 Karakteristlik funktsioon ja momendid

Teame, et teatud juhtudel avalduvad juhusliku suuruse (jaotuse) kõik momendid momente genereeriva funktsiooni tuletiste kaudu. Järgnevalt näeme, et kui juhuslikul suurusel X on n-järku lõplik moment ($X \in L_n$), siis tema karakteristlikul funktsioonil on n-järku tuletis, kusjuures n-järku moment avaldub karakteristliku funktsiooni n-järku tuletise kaudu.

Teeme ühe kasuliku tähelepaneku. Euleri valemi abil on kerge veenduda, et iga $x < y \in \mathbb{R}$ korral kehtib

$$\int_x^y ie^{it}dt = e^{iy} - e^{ix},$$

millest

$$|e^{iy} - e^{ix}| = \Big|\int_x^y ie^{it}dt\Big| \leq \int_x^y |ie^{it}|dt = y - x.$$

Erijuhul saame

$$|e^{ix} - 1| \leq |x|. \tag{14.3}$$

Lemma 14.1 *Olgu $X \in L_n(\Omega, \mathcal{F}, \mathbf{P})$, $n \geq 1$, $X \sim P$. Siis iga $k \leq n$ korral leidub tuletis $\varphi^{(k)}$, kusjuures*

$$\varphi^{(k)}(t) = E[(iX)^k e^{itX}] = \int (ix)^k e^{itx} P(dx). \tag{14.4}$$

Tõestus. Olgu $X \in L_n(\Omega, \mathcal{F}, \mathbf{P})$, $n \geq 1$.
Definitsioonist saame

$$\frac{\varphi(t+h) - \varphi(t)}{h} = \int e^{itx}(\frac{e^{ihx} - 1}{h})P(dx).$$

Näitame:

1) funktsioonid

$$f_h(x) := e^{itx}(\frac{e^{ihx} - 1}{h})$$

on domineeritud integreeruva funktsiooniga;

2) $f_h(x) \to ixe^{itx}$, kui $h \to 0$.

Domineeritus järeldub hinnangust (14.3):

$$|f_h(x)| = \Big|e^{itx}(\frac{e^{ihx} - 1}{h})\Big| = \Big|\frac{e^{ihx} - 1}{h}\Big| \leq |x|.$$

Vastavalt eeldusele $E|X| = \int |x|P(dx) < \infty$ (miks?).
Euleri valemit kasutades on kerge veenduda, et

$$\frac{d}{dt}e^{ixt} = ixe^{ixt},$$

millest

$$\lim_{h\to 0} f_h(x) = \lim_{h\to 0} \frac{e^{i(t+h)x} - e^{itx}}{h} = xie^{itx}.$$

Seega võime kasutada domineeritud koondumise teoreemi:

$$\lim_{h\to 0} \int f_h(x)P(dx) = \int \lim_{h\to 0} f_h(x)P(dx),$$

millest

$$\varphi'(t) = \lim_{h\to 0} \frac{\varphi(t+h) - \varphi(t)}{h} = \lim_{h\to 0} \int f_h(x)P(dx) = \int e^{itx}(xi)P(dx).$$

Seega juhul, kui $k = 1$, valem (14.4) kehtib.
Edasi kasutame induktsiooni. Oletame, et $k < n$ ja valem (14.4) kehtib. Tõestame, et see kehtib ka $k+1$ korral. Tõepoolest,

$$\frac{\varphi^{(k)}(t+h) - \varphi^{(k)}(t)}{h} = \int (ix)^k e^{itx}\Big(\frac{e^{ihx} - 1}{h}\Big)P(dx) = \int (ix)^k f_h(x)P(dx).$$

Iga h korral

$$|(ix)^k f_h(x)| \le |x|^k |f_h(x)| \le |x|^{k+1}$$

ja $\int |x|^{k+1}P(dx) < \infty$ (miks?). Kui $h \to 0$, siis $(ix)^k f_h(x) \to (ix)^{k+1}e^{itx}$. Nüüd kasutame jälle domineeritud koondumise teoreemi, millest saame

$$\varphi^{(k+1)}(t) = \lim_{h\to 0} \int (ix)^k f_h(x)P(dx) = \int (ix)^{k+1} \lim_{h\to 0} f_h(x)P(dx) = \int (ix)^{k+1}e^{itx}P(dx).$$

Lemma on tõestatud. ■

Kokkuvõtteks: juhusliku suuruse lõpliku n-järku momendi korral võib selle leida tema karakteristliku funktsiooni n-järku tuletisest järgneva seose abil

$$EX^n = \int x^n P(dx) = \frac{\varphi^{(n)}(0)}{i^n}. \tag{14.5}$$

Tõestuseta märgime, et vastupidine väide kehtib vaid paarisarvulise n-korral: kui karakteristlikul funktsioonil φ on punktis 0 $2n$-järku tuletis, siis mõõdul P on lõplik $2n$-järku moment (vt. *Shirjajev, II, §12*). Seega võib juhtuda, et $\varphi'_X(0) = ia$, kus $a \in \mathbb{R}$, kuid $E|X| = \infty$.

Järeldus 14.1 *Olgu $X \in L_n(\Omega, \mathcal{F}, \mathbf{P})$, $n \ge 1$, $X \sim P$. Siis punkti 0 ümbruses kehtib järgmine valem*

$$\varphi(t) = \sum_{k=0}^{n} \frac{EX^k}{k!}(it)^k + o(|t|^n). \tag{14.6}$$

Tõestus. Lemmast teame, et karakteristlikul funktsioonil φ eksisteerivad kuni n-järku tuletised. Kompleksmuutujate funktsioonide teooriast on teada, et Taylori valem kehtib ka kompleksmuutuja funktsioonide korral. Seega

$$\varphi(t) = \sum_{k=0}^{n} \frac{\varphi^{(k)}(0)}{k!} t^k + o(|t|^n). \tag{14.7}$$

Seose (14.4) tõttu $\varphi^{(k)}(0) = i^k E X^k$, millest järeldub (14.6). ■

Märkus: Valemis (14.6) olevat jääkliiget on võimalik hinnata:

$$\Big|\varphi(t) - \sum_{k=0}^{n} \frac{EX^k}{k!}(it)^k\Big| \le E\min\Big(\frac{|tX|^{n+1}}{(n+1)!}, \frac{2|tX|^n}{n!}\Big).$$

Näiteks, kui $X \in L_2$, siis

$$\varphi(t) = 1 + itEX - t^2\frac{EX^2}{2} + o,$$

kus

$$|o| \le t^2 E\min\Big(|t|\frac{|X|^3}{6}, X^2\Big).$$

Kui $t \to 0$, siis $\min(|t|\frac{|X|^3}{6}, X^2) \to 0$. Et aga miinimum on tõkestatud funktsiooniga X^2, siis DOM tõttu

$$\lim_{t\to 0} E\min\Big(|t|\frac{|X|^3}{6}, X^2\Big) = 0$$

ehk jääkliige on tõepoolest järku $o(t^2)$.

14.5 Sõltumatute juhuslike suuruste summa karakteristlik funktsioon

Olgu X_1 ja X_2 sõltumatud. Olgu $t \in \mathbb{R}$ suvaline ja tähistame $Y_j = \cos tX_j$ ja $Z_j := \sin tX_j$, kus $j = 1, 2$. On selge, et juhuslikud suurused Y_i ja Z_j on sõltumatud, kui $i \ne j$.
Leiame $\varphi_{X_1+X_2}$.

$$\begin{aligned}
&\varphi_{X_1+X_2}(t) = E[e^{it(X_1+X_2)}] = E[e^{itX_1}e^{itX_2}] = \\
&E[(Y_1 + iZ_1)(Y_2 + iZ_2)] = E[(Y_1Y_2 - Z_1Z_2 + i(Z_1Y_2 + Y_1Z_2)] = \\
&E(Y_1)E(Y_2) - E(Z_1)E(Z_2) + i[E(Z_1)E(Y_2) + E(Y_1)E(Z_2)] = \\
&[E(Y_1) + iE(Z_1)][E(Y_2) + iE(Z_2)] = E[e^{itX_1}]E[e^{itX_2}] = \varphi_{X_1}(t)\varphi_{X_2}(t).
\end{aligned}$$

Seega $\varphi_{X_1+X_2} = \varphi_{X_1}\varphi_{X_2}$.

Juhul, kui $X_1, \ldots, X_k$ on sõltumatud, saame analoogiliselt

$$\varphi_{X_1+\cdots+X_k} = \varphi_{X_1}\cdots\varphi_{X_k}.$$

14.6 Ühesus

Alljärgnevas näeme, et tõenäosusmõõtude ja nende karakteristlike funktsioonide vahel on üksühene vastavus – erinevate jaotuste karakteristlikud funktsioonid on erinevad. See asjaolu põhineb järgmisel fundamentaalsel teoreemil, mis näitab kuidas jaotusfunktsiooni saab konstrueerida karakteristliku funktsiooni järgi.

Teoreem 14.4 (Levy' pöördvalem). *Olgu φ tõenäosusmõõdu P karakteristlik funktsioon. Kui $a < b \in \mathbb{R}$ on sellised, et $P\{a\} = P\{b\} = 0$, siis*

$$P(a, b] = \lim_{T\to\infty} \frac{1}{2\pi} \int_{-T}^{T} \frac{e^{-ita} - e^{-itb}}{it} \varphi(t)dt. \tag{14.8}$$

Teoreemi tõestuse võib leida: *Billingsley*, lk. 346, *Ŝirjajev* osa II §12, *Williams* 16.6., *Durrett* 3.2.

Märkused:

- Teoreemist järeldub, et (14.8) parem pool on reaalarv.
- Karakteristlik funktsioon φ on pidev, samuti on pidev funktsioon e^{ix}. Seega on valemis (14.8) olev integraalialune (kompleks)funktsioon pidev (tema reaal-ja imaginaarosa on pidevad). Kinnises lõigus $[-T, T]$ pidev funktsioon on Riemanni mõttes integreeruv. Seega on integraali (14.8) Riemanni integraal (reaal-ja imaginaarosa on integraalid on mõlemad Riemanni integraalid). Sellest ka vastav tähustus.
- Kinnises lõigus pidev funktsioon on alati tõkestatud. Veendume, et integraalialune funktsioon on tervel reaalteljel tõkestatud. Selleks kasutame hinnangut (14.3) ja karakteristlikke funktsioonide omadust **2)**. Iga $t \in \mathbb{R}$ korral saame

  $$\left|\frac{e^{-ita} - e^{-itb}}{it}\varphi(t)\right| \leq \left|\frac{e^{-ita} - e^{-itb}}{it}\right| = \frac{|e^{-itb}(e^{-ita+itb} - 1)|}{|t|} = \frac{|e^{it(b-a)} - 1|}{|t|} \leq |a - b|. \tag{14.9}$$

  Tõkestatusest aga ei järeldu, et valemis (14.8) olev integraal eksisteeriks ka üle kogu reaaltelje ($T = \infty$). Erijuhul (sõltuvalt φ-st) võib see küll nii olla (vt järeldus 14.4). Sellisel juhul ühtib vaadeldav piirväärtus integraaliga üle reaaltelje.
- Kui F on mõõdu P jaotusfunktsioon, saame

  $$F(b) - F(a) = \lim_{T\to\infty} \frac{1}{2\pi} \int_{-T}^{T} \frac{e^{-ita} - e^{-itb}}{it} \varphi(t)dt.$$

  Levy' pöördvalemist järeldub kergesti jaotuste ja karakteristlike funktsioonide vaheline üksühene vastavus.

Järeldus 14.2 *Olgu P ja Q σ-algebral $\mathcal{B}(\mathbb{R})$ antud sellised tõenäosusmõõdud, et $\varphi_P = \varphi_Q$. Siis $P = Q$.*

Tõestus. Olgu $T_P = \{x \in \mathbb{R} | P\{x\} = 0\}$. Seega on T_P mõõdu P jaotusfunktsiooni pidevuspunktide hulk. See hulk on reaalteljel kõikjal tihe (miks?). Analoogiliselt defineerime hulga T_Q. Loomulikult on ka hulk $T_P \cap T_Q$ kõikjal tihe. See aga tähendab, et π-süsteem $\pi := \{(a, b] | a, b \in T_P \cap T_Q\}$ tekitab Boreli σ-algebra $\mathcal{B}$ (miks?).
Et $\varphi_P = \varphi_Q$, siis (14.8) põhjal $P(a, b] = Q(a, b]\ \forall (a, b] \in \pi$. Teoreemist 2.3 järeldub nüüd, et $P = Q$. ∎

Järeldus 14.3 *Juhusliku suuruse X karakteristlik funktsioon φ_X on reaalne parajasti siis, kui X jaotus P on sümmeetriline, s.o. $\forall B \in \mathcal{B}$ korral $P(B) = P(-B)$, kus $-B := \{-x : x \in B\}$.*

Tõestus. $\Leftarrow$: Olgu P sümmeetriline. Et $x \mapsto \sin(tx)$ on paaritu funktsioon (see tähendab, et $\sin t(-x) = -\sin tx$), siis $\int \sin(tx) P(dx) = 0$. Seega $\varphi_X(t) = \int \cos tx P(dx) \in \mathbb{R}$.

$\Rightarrow$: Olgu φ_X reaalne. Omadusest **4)** järeldub, et $\varphi_X(-t) = \overline{\varphi_X(t)}$. Omadusest **5)** saame $\varphi_{-X}(t) = \varphi_X(-t) = \varphi_X(t)$. Seega $\varphi_X = \varphi_{-X}$, millest järelduse 14.2 tõttu X ja $-X$ on sama jaotusega P. Seega $\forall B \in \mathcal{B}$ korral

$$P(-B) = \mathbf{P}(X \in -B) = \mathbf{P}(-X \in B) = P(B).$$

∎

Juhul, kui φ on integreeruv, on (üle reaaltelje) integreeruv ka valemis (14.8) olev integraalialune funktsioon. Järgnevalt näeme, et sellisel juhul on mõõdul P pidev tihedusfunktsioon, mis avaldub üheselt karakteristliku funktsiooni kaudu.

Järeldus 14.4 *Olgu φ mõõdu P karakteristlik funktsioon. Kui $\int_{-\infty}^{\infty} |\varphi(t)| dt < \infty$, siis mõõdul P on pidev tihedusfunktsioon f, mis avaldub karakteristliku funktsiooni kaudu järgmiselt*

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \varphi(t) dt, \quad x \in \mathbb{R}. \tag{14.10}$$

Tõestus. Olgu $\varphi \in L_1$. Veendume, et sellisel juhul pole mõõdul P aatomeid ($T_P = \mathbb{R}$). Olgu $x \in \mathbb{R}$ suvaline. Valime jada a_n ja b_n nii, et $a_n < x < b_n$, $a_n \nearrow x$, $b_n \searrow x$, $P\{a_n\} = P\{b_n\} = 0$ (sellised jadad leiduvad, sest T_P on kõikjal tihe).
Seosest (14.8) ja hinnangust (14.9) saame $\forall n$ korral

$$\begin{aligned} P\{x\} \le P(a_n, b_n] &= \lim_{T \to \infty} \frac{1}{2\pi} \int_{-T}^{T} \frac{e^{-ita_n} - e^{-itb_n}}{it} \varphi(t) dt \\ &\le \lim_{T \to \infty} \frac{1}{2\pi} \int_{-T}^{T} \Big| \frac{e^{-ita_n} - e^{-itb_n}}{it} \varphi(t) \Big| dt \\ &\le \frac{b_n - a_n}{2\pi} \lim_{T \to \infty} \int_{-T}^{T} |\varphi(t)| dt = \frac{b_n - a_n}{2\pi} \int_{-\infty}^{\infty} |\varphi(t)| dt. \end{aligned}$$

Et $b_n - a_n \to 0$, siis $P\{x\} = 0$.
Järelikult, iga $x, x+h \in \mathbb{R}$ korral kehtib (14.8); et integraalialune funktsioon on (üle reaaltelje) integreeruv, saame

$$\frac{F(x+h)-F(x)}{h} = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{e^{-itx}-e^{-it(x+h)}}{ith}\varphi(t)dt.$$

Kui $h \to 0$, siis $\forall x \in \mathbb{R}$ korral

$$\frac{e^{-itx}-e^{-it(x+h)}}{ith} \to -\frac{1}{it}\frac{d}{dx}e^{-itx} = e^{-itx}.$$

Domineeritud koondumise teoreemist saame seega (milline on domineeriv funktsioon?)

$$F'(x) = \lim_{h\to 0}\frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{e^{-itx}-e^{-it(x+h)}}{ith}\varphi(t)dt = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx}\varphi(t)dt.$$

Tähistame $f := F'$. Lõpuks kasutame asjaolu, et kui jaotusfunktsioonil on iga x korral tuletis, siis see on tema tihedusfunktsioon. Funktsiooni f pidevus tõestatakse samamoodi kui karakteristliku funktsiooni pidevusega. ■

Märkus: Nägime, et kui $\varphi \in L_1$, siis on mõõdul P pidev tihedus f, kusjuures kehtivad duaalsed valemid

$$\varphi(t) = \int_{-\infty}^{\infty} e^{itx} f(x)dx$$
$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx}\varphi(t)dt.$$

Lisame veel, et vastupidine implikatsioon ei kehti: leidub jaotusi, millel on küll tihedus, kuid $\varphi \notin L_1$ (Ülesanne 5).

14.7 Lisaks *

Levy pöördvalem kehtib ka siis, kui a või b on aatomid. Nimelt kehtib seose (14.8) üldistus:

$$P(a,b) + \frac{1}{2}P\{a,b\} = \lim_{T\to\infty}\frac{1}{2\pi}\int_{-T}^{T}\frac{e^{-ita}-e^{-itb}}{it}\varphi(t)dt. \tag{14.11}$$

Juhul kui $P\{a,b\} = 0$, järeldub siit (14.8).

Aatomi massi saab leida järgmisest pöördvalemist:

$$P\{a\} = \lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T} e^{-ita}\varphi(t)dt. \tag{14.12}$$

Siit on kerge näha, et kui $\lim_{|t|\to\infty}\varphi(t) = 0$, siis $P\{a\} = 0$ iga $a \in \mathbb{R}$, st mõõdul pole aatomeid. Tõepoolest, et $\lim_{|t|\to\infty}\varphi(t) = 0$, siis iga $\epsilon > 0$ korral leidub $t_o < \infty$ nii, et $\varphi(t) < \epsilon$, kui $t > t_o$. Siis aga iga reaalarvu a korral

$$\begin{aligned}
P\{a\} &\leq \lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}|e^{-ita}\varphi(t)|dt = \lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}|\varphi(t)|dt \\
&= \lim_{T\to\infty}\frac{1}{2T}\Big(\int_{-t_o}^{t_o}|\varphi(t)|dt + \int_{t_o}^{T}|\varphi(t)|dt + \int_{-T}^{-t_o}|\varphi(t)|dt\Big) \\
&\leq \lim_{T\to\infty}\frac{1}{2T}\int_{-t_o}^{t_o}|\varphi(t)|dt + \epsilon = \epsilon.
\end{aligned}$$

Siit järeldub, et kui $\varphi \in L_1$, siis vastaval mõõdul pole aatomeid, mille tõestasime ka ülalpool.
Tegelikult piisab mõõdu mitteatomaarsuseks isegi sellest, et $\lim_{t\to\infty}\varphi(t) = 0$. Tõepoolest, olgu φ juhusliku suuruse X karakteristlik funktsioon. Olgu Y sama jaotusega ning juhuslikust suurusest X sõltumatu. Siis $X - Y$ karakteristlik funktsioon on $|\varphi(t)|^2$, mis on paarisfunktsioon, st. $|\varphi(-t)|^2 = |\varphi(t)|^2$ (miks?). Seega $\lim_{t\to-\infty}|\varphi(t)|^2 = \lim_{t\to\infty}|\varphi(t)|^2$. Kui see piirväärtus on 0, siis, nagu eelpool veendusime, kehtib

$$\lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}|\varphi(t)|^2dt = \mathbf{P}(X - Y = 0) = 0.$$

Olgu $a \in \mathbb{R}$. Et $P(X - Y = 0) \geq \mathbf{P}(X = a, Y = a) = P^2\{a\}$, saame, et kui $\lim_{t\to\infty}\varphi(t) = 0$, siis ($X$ ja Y ühisel) jaotusel P pole aatomeid. Vastupidine üldiselt ei kehti: on jaotusi, millistel pole aatomeid, kuid $\varphi(t) \not\to 0$ kui $t \to \infty$. Sellisel jaotusel ei saa olla tihedust, sest on võimalik näidata, et kui jaotusel on tihedus, siis $\lim_{t\to\infty}\varphi(t) = 0$ (*Billingsley* Thm 26.1).

Kokkuvõtteks: Olgu φ P karakteristlik funktsioon. Kehtivad implikatsioonid

$$\varphi \in L_1 \quad\Rightarrow\quad P \text{ abs. pidev} \quad\Rightarrow\quad \lim_{t\to\infty}\varphi(t) = 0 \quad\Rightarrow\quad P \text{ pole aatomeid.}$$

14.8 Näiteid karakteristlikest funktsioonidest

Aatom : $P = \delta_a$, $a \in \mathbb{R}$.

$$\varphi(t) = e^{ita}.$$

Erijuht $a = 0$, $\varphi(t) = 1$.

Sümmeetriline Bernoulli : $P = \frac{1}{2}(\delta_1 + \delta_{-1})$.

$$\varphi(t) = e^{it}\frac{1}{2} + e^{-it}\frac{1}{2} = \frac{e^{it} + e^{-it}}{2} = \cos t.$$

Bernoulli $B(1,p)$: $P = q\delta_0 + p\delta_1$.

$$\varphi(t) = e^{it0}q + e^{it}p = q + pe^{it}.$$

Binoomjaotus $B(n,p)$: $P=\sum_{k=0}^{n} C_n^k q^{n-k} p^k \delta_k$.

Binoomjaotusega juhuslik suurus on n sõltumatu Bernoulli jaotusega juhusliku suuruse summa. Seega

$$\varphi(t)=\Pi_{k=0}^{n}(q+pe^{it})=(q+pe^{it})^n.$$

Poissoni jaotus $Po(\lambda)$: $P=\sum_{n=0}^{\infty} e^{-\lambda}\frac{\lambda^n}{n!}\delta_n$.

$$\varphi(t)=e^{-\lambda}\sum_{n=0}^{\infty} e^{itn}\frac{\lambda^n}{n!}=e^{-\lambda}\sum_{n=0}^{\infty}\frac{[e^{it}\lambda]^n}{n!}=e^{-\lambda}\exp[e^{it}\lambda]=e^{\lambda[e^{it}-1]}.$$

Paneme tähele, et formaalselt saame Poissoni jaotuse karakteristliku funktsiooni, kui asendame tema momente genereeriva funktsioonis argumendi t argumendiga it. See EI OLE tõestus, sest imaginaarühikut ei saa käsitleda kui reaalarvu. Ülaltoodud tõestuses kasutasime kompleksmuutuja funktsiooni definitsiooni.

Ühtlane jaotus $U(0,1)$: tihedus $f(x)=1$, $x\in[0,1]$

$$\varphi(t)=\int_0^1 e^{itx}dx=\int_0^1 \cos tx dx+i\int_0^1 \sin tx dx=\frac{i\sin t+\cos t-1}{it}=\frac{e^{it}-1}{it}.$$

Eksponentjaotus $E(1)$: tihedus $f(x)=e^{-x}$, $x\in[0,\infty)$

$$\varphi(t)=\int_0^{\infty} e^{itx}e^{-x}dx=\int_0^{\infty}\cos(tx)e^{-x}dx+i\int_0^{\infty}\sin(tx)e^{-x}dx.$$

Ositi integreerides saame

$$\int_0^{\infty}\cos(tx)e^{-x}dx=-\cos(tx)e^{-x}|_0^{\infty}-t\int_0^{\infty}\sin(tx)e^{-x}dx=1-t\int_0^{\infty}\sin(tx)e^{-x}dx$$
$$\int_0^{\infty}\sin(tx)e^{-x}dx=-\sin(tx)e^{-x}|_0^{\infty}+t\int_0^{\infty}\cos(tx)e^{-x}dx=t\int_0^{\infty}\cos(tx)e^{-x}dx.$$

Lahendades selle võrrandisüsteemi, saame

$$\int_0^{\infty}\cos(tx)e^{-x}dx=\frac{1}{1+t^2},\quad \int_0^{\infty}\sin(tx)e^{-x}dx=\frac{t}{1+t^2},$$

millest

$$\varphi(t)=\frac{1+it}{1+t^2}=\frac{1}{1-it}.$$

Paneme tähele, et formaalselt saame eksponentjaotuse karakteristliku funktsiooni, kui asendame tema momente genereeriva funktsioonis argumendi t argumendiga it. See EI OLE tõestus.

Analoogiliselt saab näidata, et $E(\lambda)$ karakteristlik funktsioon on

$$\frac{\lambda}{\lambda - it} = \frac{\lambda^2}{\lambda^2 + t^2} + i\frac{t\lambda}{\lambda^2 + t^2}$$

Kahepoolne eksponentjaotus : tihedus $f(x) = \frac{1}{2}e^{-|x|}$, $x \in \mathbb{R}$. Olgu φ_e ühepoolse eksponentjaotuse karakteristlik funktsioon.

$$\begin{aligned}\varphi(t) &= \frac{1}{2}\int_{-\infty}^{\infty} e^{itx}e^{-|x|}dx = \frac{1}{2}\Big[\int_{-\infty}^{0} e^{itx}e^{-x}dx + \int_{0}^{\infty} e^{itx}e^{x}dx\Big] \\ &= \frac{1}{2}\Big(\varphi_e(-1) + \varphi_e(1)\Big) = \frac{1}{2}\Big(\frac{1}{1-it} + \frac{1}{1+it}\Big) = \frac{1}{1+t^2}.\end{aligned}$$

Cauchy jaotus : tihedus $f(x) = \frac{1}{\pi}\frac{1}{1+x^2}$, $x \in \mathbb{R}$.

Kahepoolse eksponentjaotuse karakteristlik funktsioon on integreeruv. Seega kehtib (14.10), millest

$$\frac{1}{2}e^{-|x|} = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx}\frac{dt}{1+t^2}. \tag{14.13}$$

On selge, et iga $x \in \mathbb{R}$ korral $\int_{-\infty}^{\infty} \sin tx(1+t^2)^{-1}dt = 0$, millest

$$\int_{-\infty}^{\infty} e^{-itx}\frac{dt}{1+t^2} = \int_{-\infty}^{\infty} \frac{\cos tx}{1+t^2}dt = \int_{-\infty}^{\infty} e^{itx}\frac{dt}{1+t^2}.$$

Kasutades seost (14.13), saame

$$\varphi(t) = e^{-|t|}.$$

Normaaljaotus $N(\mu, \sigma^2)$: tihedus $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}}e^{\frac{(x-\mu)^2}{\sigma^2}}$, $x \in \mathbb{R}$.

Olgu f standardse normaaljaotuse tihedus. Standardne normaaljaotus on sümmeetriline, st

$$\int_{-\infty}^{\infty} e^{itx}f(x)dx = \int_{-\infty}^{\infty} \cos(tx)f(x)dx.$$

Et rida

$$\cos(tx) = \sum_{n=o}^{\infty}(-1)^n\frac{(xt)^{2n}}{(2n)!} \tag{14.14}$$

on absoluutselt koonduv, siis integreerime seda liikmeti ning saame

$$\int_{-\infty}^{\infty} \cos(tx)f(x)dx = \sum_{n=o}^{\infty}(-1)^n\frac{t^{2n}}{(2n)!}\int x^{2n}f(x)dx.$$

Standardse normaaljaotuse momente genereeriva funktsiooni tuletistest saame, et

$$\int x^{2n} f(x)dx = 1 \cdot 3 \cdot \dots \cdot (2n-1) = \frac{(2n)!}{2^n n!}$$

ning sellest järeldub, et

$$\sum_{n=o}^{\infty} (-1)^n \frac{t^{2n}}{(2n)!} \int x^{2n} f(x)dx = \sum_{n=o}^{\infty} (-\frac{t^2}{2})^n \frac{1}{n!} = \exp[-\frac{t^2}{2}].$$

Karakteristlikku funktsioonide omaduse **5)** põhjal saame üldise normaaljaotuse karakteristlikuks funktsiooniks

$$\varphi(t) = e^{it\mu}\varphi_o(\sigma t) = e^{it\mu} e^{\frac{-t^2\sigma^2}{2}} = e^{it\mu - \frac{t^2\sigma^2}{2}}.$$

Polya kriteerium Olgu $\varphi(t)$ reaalväärtuseline, mittenegatiivne, hulgal $(0, \infty)$ mittekasvav ja kumer, kusjuures $\lim_{t\to 0} \varphi(t) = \varphi(0) = 1, \varphi(-t) = \varphi(t)$ ja $\lim_{t\to\infty} \varphi(t) = 0$. Siis φ on karakteristlik funktsioon. (Tõestuse võib leida näiteks *Durrett* ptk 2). Nii näiteks on karakteristlikud funktsioonid

$$\varphi(t) = \exp(-|t|^\alpha), \quad 0 < \alpha \leq 1. \tag{14.15}$$

(Ülesanne 8). Veel enam, saab näidata, et (14.15) on karakteristlik funktsioon, kui $1 < \alpha \leq 2$, kuid mitte siis kui $\alpha > 2$.

14.9 Koondumisteoreem

Olgu P_n ja P tõenäosusmõõdud, φ_n ja φ vastavad karakteristlikud funktsioonid, $n = 1, 2, \dots$. Selgub, et $P_n \Rightarrow P$ parajasti siis, kui $\varphi_n(t) \to \varphi(t)$, $\forall t \in \mathbb{R}$ ehk tõenäosusmõõtude nõrk koondumine on ekvivalentne nende karakteristlike funktsioonide punktiviisilise koondumisega.

Tõestus põhineb abilemmal.

Lemma 14.2 *Olgu P σ-algebral $\mathcal{B}(\mathbb{R})$ antud tõenäosusmõõt, φ olgu tema karakteristlik funktsioon. Siis iga $K < \infty$ korral*

$$1 - P[-2K, 2K] \leq K \int_{-K^{-1}}^{K^{-1}} (1 - \varphi(t))dt.$$

Tõestus.

$$\int_{-K^{-1}}^{K^{-1}} (1-\varphi(t))dt = \int_{-K^{-1}}^{K^{-1}} \Big[1 - \int e^{itx} P(dx)\Big] dt = \int_{-K^{-1}}^{K^{-1}} \Big[\int (1 - e^{itx}) P(dx)\Big] dt.$$

Saab näidata, et ülaltoodud integraalis võib integreerimisjärjekorra ära vahetada (Fubini teoreem). Seega, võttes arvesse, et iga $u > 0$ korral

$$\int_{-u}^{u}\big(1-\cos(tx)-i\sin(tx)\big)dt = \int_{-u}^{u}\big(1-\cos(tx)\big)dt = 2u - 2\frac{\sin ux}{x} = 2u\Big(1-\frac{\sin ux}{ux}\Big),$$

saame (kasutades seoseid $\frac{\sin x}{x} \leq 1$ ja $\frac{\sin x}{x} \leq |\frac{\sin x}{x}| \leq \frac{1}{|x|}$),

$$K\int_{-K^{-1}}^{K^{-1}}(1-\varphi(t))dt = 2\int\Big(1-\frac{\sin K^{-1}x}{K^{-1}x}\Big)P(dx) \geq 2\int_{\{x:|x|\geq 2K\}}\Big(1-\frac{\sin K^{-1}x}{K^{-1}x}\Big)P(dx)$$
$$\geq 2\int_{\{x:|x|\geq 2K\}}\Big(1-\frac{1}{|K^{-1}x|}\Big)P(dx) \geq P\{x:|x|\geq 2K\}.$$

■

Lemma 14.2 abil tõestame põhiteoreemi.

Teoreem 14.5 (Levy, Cramer) *Olgu P_n ja P σ-algebral $\mathcal{B}(\mathbb{R})$ antud tõenäosusmõõdud, φ_n, φ vastavad karakteristlikud funktsioonid. Siis $P_n \Rightarrow P$ parajasti siis, kui iga $t \in \mathbb{R}$ korral $\varphi_n(t) \to \varphi(t)$.*

Tõestus. $\Rightarrow$: Oletame, et $P_n \Rightarrow P$. Olgu $t \in \mathbb{R}$ suvaline. Funktsioonid sin ja cos on pidevad ja tõkestatud, mistõttu tõkestatud koondumise teoreemist saame

$$\int \cos txP_n(dx) \to \int \cos txP(dx), \quad \int \sin txP_n(dx) \to \int \sin txP(dx).$$

Seega $\varphi_n(t) \to \varphi(t)$.
$\Leftarrow$: Oletame, et $\varphi_n(t) \to \varphi(t)$, $\forall t \in \mathbb{R}$.
Eelmisest lemmast järeldub, et iga $h > 0$ ja iga karakteristliku funktsioon φ korral on $\int_{-h}^{h}\big(1-\varphi_n(t)\big)dt$ reaalarv. Iga $h > 0$ korral

$$\frac{1}{2h}\int_{-h}^{h}\big(1-\varphi_n(t)\big)dt \leq \frac{1}{2h}\int_{-h}^{h}\big|1-\varphi_n(t)\big|dt \leq \frac{1}{2h}\int_{-h}^{h}\big|1-\varphi(t)\big|dt + \frac{1}{2h}\int_{-h}^{h}\big|\varphi(t)-\varphi_n(t)\big|dt. \tag{14.16}$$

1) Tõestame, et

$$\lim_{h\to 0}\frac{1}{2h}\int_{-h}^{h}\big|1-\varphi(t)\big|dt = 0. \tag{14.17}$$

Karakteristlike funktsioonide omadustest teame, et φ on pidev ja $\varphi(0) = 1$, millest $\forall\delta > 0$ korral $\exists h > 0$ nii, et $|1-\varphi(t)| \leq \delta$, kui $t \in (-h, h)$. Iga sellise h korral aga

$$\frac{1}{2h}\int_{-h}^{h}\big|1-\varphi(t)\big|dt \leq \delta.$$

2) Näitame, et iga $h > 0$ korral

$$\frac{1}{2h}\int_{-h}^{h}|\varphi_n(t)-\varphi(t)|dt \to 0, \quad \text{kui} \quad n \to \infty. \tag{14.18}$$

Et $\varphi_n(t) \to \varphi(t)$, siis iga fikseeritud h korral (14.18) järeldub domineeritud koondumise teoreemist ($|\varphi_n(t) - \varphi(t)| \le 2$).
3) Näitame, et jada $\{P_n\}$ on tihe.
Olgu $\epsilon > 0$ suvaline. Valime $h(\epsilon)$ nii väikese, et integraal (14.17) oleks väiksem kui $\frac{\epsilon}{2}$. Selle h korral valime $n_o(h)$ nii suure, et integraal (14.18) oleks väiksem kui $\frac{\epsilon}{2}$. Seega võrratuse (14.16)) parem pool pole suurem kui ϵ. Lemmast 14.2 saame iga $n > n_o$ korral

$$1 - P_n[-2h^{-1}, 2h^{-1}] \le \frac{1}{h}\int_{-h}^{h}\big(1-\varphi_n(t)\big)dt \le 2\epsilon. \tag{14.19}$$

Iga tõenäosusmõõt (omaette võetuna) on tihe. Sellest järeldub, et iga lõplik hulk tõenäosusmõõte on tihe (miks?). Seega, vähendades h-d kui vaja, saame, et võrratus (14.19) kehtib iga n korral. Et ϵ oli suvaline, järeldub sellest jada $\{P_n\}$ tihedus.
4) Näitame, et $P_n \Rightarrow P$. Selleks piisab kui näitame, et jada $\{P_n\}$ iga nõrgalt koonduva alamjada piirväärtus on P (järeldus 12.2). Olgu $P_{n_k} \Rightarrow Q$ suvaline nõrgalt koonduv alamjada, ϕ olgu mõõdu Q karakteristlik funktsioon. Seega $\varphi_{n_k}(t) \to \phi(t)$ forall $t \in \mathbb{R}$, millest $\phi = \varphi$. Järelduse 14.2 põhjal $P = Q$. ■

Teoreemis 14.5 eeldasime, et karakteristlike funktsioonide jada (punktiviisiline) piirväärtus on mingi mõõdu P karakteristlik funktsioon. Järgnevas järelduses näeme, et see eeldus on täidetud, kui karakteristlike funktsioonide jada (punktiviisiline) piirväärtus on pidev punktis 0.

Järeldus 14.5 *Olgu P_n σ-algebral $\mathcal{B}(\mathbb{R})$ antud tõenäosusmõõdud, φ_n vastavad karakteristlikud funktsioonid. Oleteme, et iga $t \in \mathbb{R}$ korral kehtib koondumine $\varphi_n(t) \to g(t)$, kus piirfunktsioon g on pidev punktis 0. Siis leidub tõenäosusmõõt P nii, et $P_n \Rightarrow P$ ja g on P karakteristlik funktsioon.*

Tõestus. Vaatame veel kord piisavuse tõestust teoreemis 14.5. Omaduse $\{P_n\}$ tiheduse tõestamisel kasutasime seosed (14.17) ja (14.18). Omaduse (14.18) tõestamisel kasutasime vaid punktiviisilist koondumist; see tõestus kehtib ka käesoleval juhul. Seose (14.17) tõestamisel kasutasime vaid asjaolu, et φ on pidev punktis 0. Seega seose (14.17) tõestus kehtib ka vaadeldaval juhul. Järelikult ka praegusel juhul on $\{P_n\}$ tihe, mistõttu tal on nõrgalt koonduv alamjada $P_{n_k} \Rightarrow P$, millest teoreemi 14.5 tõttu $\varphi_{n_k}(t) \to \varphi(t)$ $\forall t \in \mathbb{R}$, kus φ on P karakteristlik funktsioon. Seega $\varphi = g$ ja g on P karakteristlik funktsioon. ■

Järelduse tõestusest nägime, et omadus – g on pidev punktis 0 – oli vajalik vaid jada $\{P_n\}$ tiheduse tõestamiseks (vt ka ülesanne 9). Seega kehtib järeldus.

Järeldus 14.6 *Olgu P_n ruumil $(\mathbb{R}, \mathcal{B})$ antud tõenäosusmõõdud, φ_n vastavad karakteristlikud funktsioonid. Oletame, et iga $t \in \mathbb{R}$ korral kehtib koondumine $\varphi_n(t) \to g(t)$. Kui $\{P_n\}$ on tihe, siis leidub tõenäosusmõõt P nii, et $P_n \Rightarrow P$ ja g on P karakteristlik funktsioon.*

14.10 Juhusliku vektori karakteristlik funktsioon

Olgu $X = (X_1, \ldots, X_k)$ k-dimensionalne juhuslik vektor. Juhusliku vektori karakteristlik funktsioon $\varphi : \mathbb{R}^k \to \mathbb{C}$ defineeritakse analoogiliselt

$$\varphi(t_1, \ldots, t_k) = \varphi(t) = Ee^{it'X} = E(\cos t'X) + iE(\sin t'X).$$

Pea meeles, et $t = (t_1, \ldots, t_k)$ on nüüd k-dimensionaalne vektor. Juhuslike vektorite karakteristlikul funktsioonil on sisuliselt samad omadused, mis juhuslike suuruste karakteristlikel funktsioonidel. Näiteks sõltumatute juhuslike vektorite summa karakteristlik funktsioon on nende vektorite karakteristlike funktsioonide korrutis: kui X ja Y on sõltumatud, siis $\varphi_{X+Y}(t) = \varphi_X(t)\varphi_Y(t)$ (tõesta!).

14.10.1 Segamomendid ja osatuletised

Hölder: kui $X \in L_p$ ja $Y \in L_q$, kus $p^{-1} + q^{-1} = 1$, siis $E|XY| \le \|X\|_p \|Y\|_q$. Seda võrratust on kerge üldistada suvalise lõpliku arvu juhuslike suurusteni: olgu $X_i \in L_{p_i}$, kus $\sum_{i=1}^k p_i^{-1} = 1$. Siis (ülesanne 12)

$$E\Big(\prod_{i=1}^k |X_i|\Big) \le \prod_{i=1}^k \|X_i\|_{p_i}. \tag{14.20}$$

Kuulugu juhuslik vektor X ruumi L_m, st $E\|X\|^m < \infty$. Võrratuse (14.20) abil on kerge veenduda, et siis leiduvad kõik segamomendid $E\big((X_1)^{l_1} \cdots (X_k)^{l_k}\big)$, kus $l_1 + \cdots + l_k \le m$, sest

$$E\big|(X_1)^{l_1} \cdots (X_k)^{l_k}\big| \le \prod_{i=1}^k \|X_i\|_m^{l_i} < \infty.$$

(ülesanne 12). Sellisel juhul leiduvad karakterstlikul funktsioonil osatuletised

$$\frac{\partial^{l_1+\cdots+l_k}}{\partial(t_1)^{l_1} \cdots \partial(t_k)^{l_k}} \varphi(t_1, \ldots, t_k) \tag{14.21}$$

ja kehtib seos

$$m_{l_1 \cdots l_k}(X) := E\big((X_1)^{l_1} \cdots (X_k)^{l_k}\big) = \frac{1}{i^{l_1+\cdots+l_k}} \frac{\partial^{l_1+\cdots+l_k} \varphi(t_1, \ldots, t_k)}{\partial(t_1)^{l_1} \cdots \partial(t_k)^{l_k}}\Big|_{t=0}. \tag{14.22}$$

(Ühemõõtmelisel juhul on see lemma 14.4). Järelduse 14.1 üldistus mitmemõõtmelisele juhule on seega järgmine: kui $E\|X\|^m < \infty$, siis protsessis $t \to 0$ kehtib järgmine valem

$$\varphi(t_1, \ldots, t_k) = \sum_{(l_1, \ldots, l_k) : |l| \le m} \frac{m_{l_1 \cdots l_k}(X)}{l_1! \cdots l_k!} i^{|l|} (t_1)^{l_1} \cdots (t_k)^{l_k} + o(|t|^m), \tag{14.23}$$

kus $l = (l_1, \ldots, l_k) \in \mathbb{N}^k$, $|l| = l_1 + \cdots + l_k$ ja $|t| = |t_1| + \cdots + |t_k|$.

14.10.2 Ühesus

Levi pöördvalemi (14.8) üldistus on järgmine. Olgu P σ-algebral $\mathcal{B}(\mathbb{R}^k)$ antud tõenäosusmõõt (k-dimensionaalse juhusliku vektori jaotus). Olgu $A = (a_1, b_1] \times \cdots \times (a_k, b_k]$ P-pidev ristkülik. Siis

$$P(A) = \lim_{T\to\infty} \frac{1}{(2\pi)^k} \int_{-T}^{T} \cdots \int_{-T}^{T} \prod_{j=1}^{k} \frac{e^{-it_j a_j} - e^{-it_j b_j}}{it_j} \varphi(t_1, \ldots, t_k) dt_1 \cdots dt_k. \tag{14.24}$$

(tõestust vaata näiteks *Billingsley* (29.3)). Valemist (14.24) järeldub tõenäosusmõõtude ja karakteristlike funktsioonide vaheline üks-ühesus: kui P ja Q on σ-algebral $\mathcal{B}(\mathbb{R}^k)$ antud sellised tõenäosusmõõdud, et $\varphi_P = \varphi_Q$, siis $P = Q$. Selle tõestus on analoogiline järelduse 14.2 tõestusega. See tähendab, et kahel erineva *jaotusega* juhuslikul vektoril (mitte kahel erineval juhuslikul vektoril!) on erinev karakteristlik funktsioon. Viimasest järeldub aga, et lineaarkombinatsioonide $t'X$ (ühemõõtmelise) jaotused määravad üheselt X jaotuse:

Järeldus 14.7 *Juhuslikud vektorid X ja Y on erineva jaotusega parajasti siis, kui leidub $t \in \mathbb{R}^k$ nii, et juhuslikud suurused $t'X$ ja $t'Y$ on erineva jaotusega.*

Tõestus. Kui X ja Y on erineva jaotusega, siis peavad nende karakteristilikud funktsioonid, olgu need φ_X ja φ_Y, olema erinevad. See aga tähendab, et leidub vähemalt üks $t \in \mathbb{R}^k$ nii, et $\varphi_X(t) \neq \varphi_Y(t)$. Ent $\varphi_X(t)$ on ühemõõtmelise juhusliku suuruse $t'X$ karakteristliku funktsiooni väärtus kohal 1 ja $\varphi_Y(t)$ on ühemõõtmelise juhusliku suuruse $t'Y$ karakteristliku funktsiooni väärtus kohal 1. Need väärtused on erinevad, millest saame, et $t'X$ ja $t'Y$ on erineva jaotusega (kuidas?).
Teistpidi: kui leidub $t \in \mathbb{R}^k$ nii, et $t'X$ ja $t'Y$ on erineva jaotusega, on nende (ühemõõtmeliste juhuslike suuruste) karakteristlikud funktsioonid erinevad. See tähendab, et leidub reaalarv s nii, et

$$\varphi_X(st) = Ee^{is(t'X)} \neq Ee^{is(t'Y)} = \varphi_Y(st).$$

Seega $\varphi_X \neq \varphi_Y$, millest saame, et X ning Y on erineva jaotusega. ■

Järeldus 14.8 *Juhuslikud suurused $X_1, \ldots X_k$ on sõltumatud parajasti siis, kui iga $t_1, \ldots, t_k \in \mathbb{R}$ korral on juhusliku suuruse $t_1X_1 + \cdots + t_kX_k$ karakteristlik funktsioon $\varphi_{t_1X_1+\cdots+t_kX_k}(s)$ võrdne korrutisega*

$$\varphi_{X_1}(st_1) \cdot \cdots \cdot \varphi_{X_k}(st_k).$$

Tõestus. Kui $X_1, \ldots X_k$ on sõltumatud, siis on sõltumatud ka $t_1X_1, \ldots, t_kX_k$ ja nende summa karakteristlik funktsioon kohal s on korrutis

$$\prod_i \varphi_{t_iX_i}(s) = \prod_i \varphi_{X_i}(st_i).$$

Teistpidi: vaatleme juhuslikku vektorit $X = (X_1, \ldots, X_k)$. Summa $t_1X_1+\cdots+t_kX_k$ karakteristlik funktsioon kohal $s = 1$ on vektori X karakteristlik funktsioon kohal $(t_1, \ldots, t_k)$ ja eeldusest saame

$$\varphi_X(t_1, \ldots, t_k) = \varphi_{t_1X_1+\cdots+t_kX_k}(1) = \prod_i \varphi_{X_i}(t_i).$$

Olgu $\tilde{X} = (\tilde{X}_1, \ldots, \tilde{X}_k)$ selline juhuslik vektor, et $\tilde{X}_i$ on sama jaotusega kui X_i, kuid $\tilde{X}_1, \ldots, \tilde{X}_k$ on sõltumatud juhuslikud suurused. Seega

$$\varphi_{\tilde{X}}(t_1, \ldots, t_k) = \prod_i \varphi_{\tilde{X}_i}(t_i) = \prod_i \varphi_{X_i}(t_i) = \varphi_X(t_1, \ldots, t_k).$$

Seega on vektoritel X ja $\tilde{X}$ sama karakteristlik funktsioon ja järeldusest 14.7 tulenevalt ka sama jaotus. Siit aga järeldub, et $X_1, \ldots X_k$ on sõltumatud. ■

Järelduse 14.8 rakendamisel pane tähele, et seos

$$\varphi_{t_1 X_1 + \cdots + t_k X_k}(s) = \prod_i \varphi_{X_i}(s t_i)$$

peab kehtima iga komplekti $(t_1, \ldots, t_k)$ korral. Olgu näiteks X_1 Cauchy jaotusega juhuslik suurus ja defineerime $X_2 = X_1$. Võttes $t_1 = t_2 = 1$, saame

$$\varphi_{X_1+X_2}(s) = \varphi_{2X_1}(s) = \varphi_{X_1}(2s) = \exp[-2|s|] = \big(\exp[-|s|]\big)^2 = \varphi_{X_1}(s)\varphi_{X_2}(s).$$

Seega paari $(1, 1)$ korral võrdus kehtib, kuid kindlasti pole X_1 ja X_2 sõltumatud.

Et Järelduse 14.8 teise poole tõestuses kasutasime vaid argumenti $s = 1$ (näitasime: kui $\varphi_X(t_1, \ldots, t_k) = \varphi_{X_1}(t_1) \cdots \varphi_{X_k}(t_k)$, siis on komponendid sõltumatud) siis võime Järelduse 14.8 ümberr sõnastada järgmiselt.

Järeldus 14.9 *Juhusliku vektori X komponendid on sõltumatud parajasti siis, kui iga $(t_1, \ldots, t_k)$ korral*

$$\varphi_X(t_1, \ldots, t_k) = \varphi_{X_1}(t_1) \cdots \varphi_{X_k}(t_k).$$

14.10.3 Juhuslike vektorite nõrk koondumine ja karakteristlikud funktsioonid

Veendume, et juhuslike vektorite nõrk koondumine (ehk σ-algebral $\mathbb{B}^k$ antud tõenäosusmõõtude nõrk koondumine) on ekvivalentne karakteristlike funktsioonide punktiviisilise koondumisega.

Teoreem 14.6 *Olgu $X_n \sim P_n$ ja $X \sim P$ k-dimensionaalsed juhuslikud vektorid. Olgu φ_n, φ vastavad karakteristlikud funktsioonid. Siis $X_n \Rightarrow X$ (ekvivalentselt $P_n \Rightarrow P$) parajasti siis, kui iga $t \in \mathbb{R}^k$ korral $\varphi_n(t) \to \varphi(t)$.*

Tõestus. Kui $X_n \Rightarrow X$, ekvivalentselt $P_n \Rightarrow P$, siis (pideva kujutise teoreem) iga $t \in \mathbb{R}^k$ korral $t'X_n \Rightarrow t'X$ ja $t \in \mathbb{R}^k$ järeldub nüüd teoreemist 14.5.
Kui iga $t \in \mathbb{R}^k$ korral $\varphi_n(t) \to \varphi(t)$, siis suvalise $t \in \mathbb{R}^k$ ja iga $s \in \mathbb{R}$ korral

$$Ee^{is(t'X_n)} = \varphi_n(st) \to \varphi(st) = Ee^{is(t'X)}, \tag{14.25}$$

millest teoreemi 14.5 põhjal järeldub $t'X_n \Rightarrow t'X$. Et t oli suvaline, siis iga i korral $X_n^i \Rightarrow X^i$. Lemmast 12.2 saame, et jada $\{P_n\}$ on tihe. Olgu $P_{n_j} \Rightarrow Q$ koonduv alamjada.

Seega iga t korral $\varphi_{n_j}(t) \to \varphi_Q(t)$, millest $\varphi = \varphi_Q$ ning $P = Q$. ■

Teoreemist 14.6 järeldub teoreem 12.9, mis väidab, et $X_n \Rightarrow X$ parajasti siis, kui iga $t \in \mathbb{R}^k$ korral $t'X_n \Rightarrow t'X$. Tõepoolest, kui $X_n \Rightarrow X$, siis $t'X_n \Rightarrow t'X$ järeldub pideva kujutise teoreemist. Teistpidi, kui $t'X_n \Rightarrow t'X$, siis teoreemi 14.5 põhjal koonduvad vastavad karakteristlikud funktsioonid, st kehtib (14.25). Võttes $s = 1$, saame koondumise $\varphi_n(t) \to \varphi(t)$. Et viimane kehtib iga t korral, saame teoreemist 14.6, et $X_n \Rightarrow X$.

Kirjandus. *Billingsley* 26, 29; *Ŝirjajev*, II §12, III §3; *Williams* 16; *Durrett*, ptk. 2.

14.11 Ülesanded

1. Olgu φ juhusliku suuruse X karakteristlik funktsioon, olgu Y sama jaotusega kui X, X ja Y olgu sõltumatud. Leida $X - Y$ karakteristlik funktsioon.

2. Olgu $X \sim N(\mu_1, \sigma_1^2)$, $Y \sim N(\mu_2, \sigma_2^2)$. Olgu X ja Y sõltumatud. Tõesta, et $X + Y \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$.

3. **a)** Olgu $\mathbf{P}(X \in \mathbb{Z}) = 1$, φ olgu X karakteristlik funktsioon. Tõesta, et φ on perioodiline perioodiga 2π, st $\varphi(t + 2\pi) = \varphi(t)$ iga t korral.

   b) Olgu $\mathbf{P}(X \in h\mathbb{Z}) = 1$, kus $h > 0$; φ olgu X karakteristlik funktsioon. Tõesta, et φ on perioodiline perioodiga $\frac{2\pi}{h}$.

4. **a)** Tõesta, et $U(-c, c)$ karakteristlik funktsioon on $\varphi(t) = \frac{\sin ct}{ct}$.

   b) Tõesta, et kui $X_1, \ldots, X_n$ on iid $U(-1, 1)$ jaotusega juhuslikud suurused ja $S_n = X_1 + \ldots + X_n$, siis S_n tihedus on

$$f(x) = \frac{1}{\pi} \int_0^\infty \left(\frac{\sin t}{t}\right)^n \cos(tx) dt.$$

5. Leida absoluutselt pidev jaotus, mille karakteristlik funktsioon pole integreeruv.

6. **a)** Tõesta, et kui $X_1, \ldots, X_n$ on iid Cauchy jaotusega juhuslikud suurused ja $S_n = X_1 + \ldots + X_n$, siis $\frac{S_n}{n}$ on samuti Cauchy jaotusga. Kas kehtib SAS?

   b) Tõesta, et kui $X_1, \ldots, X_n$ on karakteristliku funktsiooniga $\exp[-|t|^\alpha]$ iid jaotused, kus $\alpha \in (0, 2]$, siis $\frac{S_n}{n^{\frac{1}{\alpha}}}$ on sama jaotusega.

7. Olgu $X_n \Rightarrow X$ and $Y_n \Rightarrow Y$, kusjuures X_n and Y_n on iga n korral sõltumatud. Tõestada, et kui X ha Y on sõltumatud, siis $X_n + Y_n \Rightarrow X + Y$.

8. Polya kriteeriumi abil tõesta, et $\exp[-|t|^\alpha]$ on karakteristlik funktsioon, kui $\alpha \in (0, 1]$.

9. Olgu $X_n \sim N(0, n)$. Leida X_n karakteristlik funktsioon φ_n. Näita, et leidub funktsioon $g(t)$ nii, et $g(0) = 1$ ja $\varphi_n(t) \to g(t)$ iga t korral. Näita, et jada X_n ei koondu nõrgalt.

10. Olgu $X_n \sim N(0, \sigma_n^2)$, $X_n \Rightarrow X$. Tõesta, et sellisel juhul $\sigma_n^2 \to \sigma^2 \in [0, \infty)$.

11. Arvestades, et
$$\lim_{T\to\infty} \int_0^T \frac{\sin t}{t} dt = \frac{\pi}{2},$$
tõesta Levy pöördvalem karakterisliku funktsiooni $\varphi(t) \equiv 1$ korral.

12.
    - Tõestada (14.20) (kasuta Hölderi võrratust kahe juhusliku suuruse korral)
    - Tõestada, et kui $X = (X_1, \ldots, X_k)$ on selline juhuslik vektor, et $\|X\|^m < \infty$, siis $X_i \in L_m$ iga i korral. Tõestada, et kui $l_1, \ldots, l_k$ on sellised mittenegatiivsed reaalarvud, et $l_1 + \ldots + l_k = m$, siis
    $$E\big|(X_1)^{l_1} \cdots (X_k)^{l_k}\big| \le \prod_{i=1}^{k} \big(E|X_i|^m\big)^{\frac{l_i}{m}} < \infty.$$
    - Järeldada, et üleltoodud võrratused kehtivad ka siis, kui $l_1 + \ldots + l_k \le m$.

15 Tsentraalsed piirteoreemid

Karakteristlikud funktsioonid on väga kasulikud mitmesuguste piirteoreemide tõestamisel, eelkõige kasutatakse neid jaotuste nõrga koondumise tõestamisel. Järgnevas näeme, kuidas karakteristlike funktsioonide abil on võimalik lihtsalt tõestada matemaatilises statistikas ja tõenäosusteooria väga olulised teoreemid – **tsentraalsed piirteoreemid.**

Lepime kokku: kui $X_1 \sim P_1, X_2 \sim P_2, \dots$ on juhuslike suuruste jada ja P mingi reaalteljel defineeritud tõenäosusmõõt, siis kirjaviis $X_n \Rightarrow P$ tähendab, et $P_n \Rightarrow P$.

15.1 Tsentraalne piirteoreem iid juhuslike suuruste korral

Tõestame klassikalise tsentraalse piirteoreemi iid juhuslike suuruste korral.

Märgime, et ka kompleksarvude korral kehtib järgmine elementaarne tulemus:
kui $x_n \to x$ on koonduv kompleksarvude jada, siis (vt näiteks *Durrett, 2.4*)

$$\lim_n \Big(1 + \frac{x_n}{n}\Big)^n = e^x.$$

Teoreem 15.1 (Tsentraalne piirteoreem (CLT)). *Olgu $X_1, X_2, \dots$ iid juhuslikud suurused, $S_n = \sum_{i=1}^n X_i$. Kui $\sigma^2 = D(X_n) < \infty$, siis*

$$\frac{S_n - n\mu}{\sigma\sqrt{n}} \Rightarrow N(0,1), \tag{15.1}$$

kus $\mu = EX_n$.

Tõestus. Üldisust kitsendamata võime eeldada, et $\mu = 0$ (vastasel juhul vaatleme iid juhuslikke suurusi $X_n - \mu$).
Olgu $X_n \sim P$; φ olgu P karakteristlik funktsioon. Leiame juhusliku suuruse $\frac{S_n}{\sigma\sqrt{n}}$ karakteristliku funktsiooni. Karakteristlike funktsioonide omaduse **5)** põhjal $\forall i$ korral

$$\varphi_{\frac{X_i}{\sigma\sqrt{n}}}(t) = \varphi\Big(\frac{t}{\sigma\sqrt{n}}\Big),$$

millest sõltumatuse ja valemi (14.6) tõttu saame $\forall t \in \mathbb{R}$ korral

$$\varphi_{\frac{S_n}{\sigma\sqrt{n}}}(t) = \Big[\varphi\Big(\frac{t}{\sigma\sqrt{n}}\Big)\Big]^n = \Big[1 + \frac{\sigma^2}{2}\Big(\frac{it}{\sigma\sqrt{n}}\Big)^2 + o\Big(\frac{t^2}{\sigma^2 n}\Big)\Big]^n = \Big[1 - \frac{t^2}{2n} + o\Big(\frac{1}{n}\Big)\Big]^n.$$

Seega protsessis $n \to \infty$ iga $t \in \mathbb{R}$ korral

$$\varphi_{\frac{S_n}{\sigma\sqrt{n}}}(t) \to e^{-\frac{t^2}{2}}.$$

Et $e^{-\frac{t^2}{2}}$ on jaotuse $N(0,1)$ karakteristlik funktsioon, siis teoreemi 14.5 tõttu

$$\frac{S_n}{\sigma\sqrt{n}} \Rightarrow N(0,1).$$

∎

Ülesanne: Olgu $X \in L_1$. Karakteristlikke funktsioone kasutades tõestada nõrk SAS.

Näited:

1. **De Moivre-Laplace teoreem.** Olgu $X_1, X_2, \ldots, X_n$ iid Bernoulli jaotusega juhuslikud suurused parameetriga p. Siis $S_n \sim B(n,p)$ ja
$$\frac{S_n - np}{\sqrt{np(1-p)}} \Rightarrow N(0,1).$$
Seega kui $Y \sim B(n,p)$ siis (suure n korral) Y jaotus on lähendatav jaotusega $N(np, \sqrt{np(1-p)})$.
See lähendus annab väikese vea isegi suhteliselt väikese n korral. Tõepoolest, oletame, et meid huvitab tõenäosus, et 16 ausa mündi viskest täpelt pooled on kiri, st $\mathbf{P}(Y=8)$, kus $Y \sim B(16, 0.5)$. Binoomjaotusest saame
$$\mathbf{P}(Y=8) = C_{16}^{8}\Big(\frac{1}{2}\Big)^{16} = 0.1964.$$
Teisest küljest, vaadeldes Y summana S_{16}, saame
$$\mathbf{P}(S_{16} \in [7.5, 8.5]) = \mathbf{P}(|S_{16} - 8| \le 0.5) = \mathbf{P}\Big(|S_n - n\mu| \le 0.5\Big) = \mathbf{P}\Big(\frac{|S_n - n\mu|}{\sqrt{n}\sigma} \le 0.25\Big),$$
sest $\mu n = 8$ ja $\sqrt{n}\sigma = 4\sigma = 2$. Normaaljaotusega lähendades saame, et
$$\mathbf{P}\Big(\frac{|S_n - n\mu|}{\sqrt{n}\sigma} \le 0.25\Big) \approx \mathbf{P}(|X| \le 0.25) = 2(0.5987 - 0.5) = 0.1974,$$
kus $X \sim N(0,1)$.

2. Ruletil on 18 punast ja 18 musta numbrit, kaks on rohelised. Värvi õigesti ära arvates võidab mängija 1 rahaühiku, vastasel juhul kaotab ta samuti ühe ühiku (-1). Olgu $X_1, X_2, ...$ mängija võit (kaotus) i-ndal mängul. Mängud on sõltumatud, seega on $X_1, X_2, \ldots$ iid $\mu = EX_i = -\frac{1}{19}$, $DX_i = 1 - (\frac{1}{19})^2 = 0.9972$. Keskmine võit ühel mängul on seega $-\frac{1}{19} = -0.05263$. Mängides $19^2 = 361$ korda, on keskmine kaotus -19. Kui suur on aga tõenäosus, et pärast 361 mängu on mängija plussis? See tõenäosus on ($n = 361$),
$$\mathbf{P}(S_n \ge 0) = \mathbf{P}\Big(\frac{S_n - n\mu}{\sigma\sqrt{n}} \ge -\frac{n\mu}{\sigma\sqrt{n}}\Big) \approx P\Big(X \ge -\frac{n\mu}{\sigma\sqrt{n}}\Big),$$
kus $X \sim N(0,1)$. Võtame $\sigma \approx 1$, siis
$$-\frac{n\mu}{\sigma\sqrt{n}} = 1$$
ja seega otsitav tõenäosus on ligikaudu $\mathbf{P}(X \ge 1) = 1 - 0.8413 = 0.1587$.

3. Olgu $Z_\lambda \sim P_o(\lambda)$. Veendume, et suure λ korral

$$Po(\lambda) \approx N(\lambda, \lambda). \tag{15.2}$$

Oletame, et $X_1, \ldots, X_n$ on iid $Po(1)$ (seega $EX_i = DX_i = 1$). Siis $S_n \sim Po(n)$. Teisest küljest, tsentraalsest piirteoreemist saame, et

$$\frac{S_n - n}{\sqrt{n}} \Rightarrow N(0,1). \tag{15.3}$$

Seega täisarvulise parameetri korral (15.2) kehtib.
Olgu nüüd λ suvaline (mitte ilmtingimata täisarv). Olgu N_1, N_2, N_3 sõltumatud Poissoni jaotusega juhuslikud suurused parameetritega $\lfloor\lambda\rfloor$, $\lambda - \lfloor\lambda\rfloor$ ja $\lfloor\lambda\rfloor + 1 - \lambda$. Olgu $S_{\lfloor\lambda\rfloor} := N_1$, $Z_\lambda := N_1 + N_2$ ja $S_{\lfloor\lambda\rfloor+1} := N_1 + N_2 + N_3$. Sõltumatute Poissoni jaotusega juhuslike suuruste summa on Poissoni jaotusega, parameetrid liituvad. Seega $Z_\lambda \sim Po(\lambda)$ ja $S_{\lfloor\lambda\rfloor+1} \sim Po(\lfloor\lambda\rfloor + 1)$, Poissoni jaotusega juhuslikud suurused on mittenegatiivsed, seega

$$S_{\lfloor\lambda\rfloor} \leq Z_\lambda \leq S_{\lfloor\lambda\rfloor+1}.$$

Tähistades $n := \lfloor\lambda\rfloor$, st $S_n = S_{\lfloor\lambda\rfloor}$, järeldub sellest, et $S_n - (n+1) \leq Z_\lambda - \lambda \leq S_{n+1} - n$ ja

$$\frac{S_n - n}{\sqrt{n}} \frac{\sqrt{n}}{\sqrt{\lambda}} - \frac{1}{\sqrt{\lambda}} \leq \frac{Z_\lambda - \lambda}{\sqrt{\lambda}} \leq \frac{S_{n+1} - (n+1)}{\sqrt{n+1}} \frac{\sqrt{n+1}}{\sqrt{\lambda}} + \frac{1}{\sqrt{\lambda}}.$$

Tähistades $a_\lambda := \sqrt{\frac{n}{\lambda}}$ ja $b_\lambda := \sqrt{\frac{n+1}{\lambda}}$

$$\mathbf{P}\Big(\frac{S_{n+1} - (n+1)}{\sqrt{n+1}} \leq b_\lambda^{-1}\big(t - \frac{1}{\sqrt{\lambda}}\big)\Big) \leq \mathbf{P}\Big(\frac{Z_\lambda - \lambda}{\sqrt{\lambda}} \leq t\Big) \leq \mathbf{P}\Big(\frac{S_n - n}{\sqrt{n}} \leq a_\lambda^{-1}\big(t - \frac{1}{\sqrt{\lambda}}\big)\Big) \quad \forall t.$$

Et protsessis $\lambda \to \infty$, $a_\lambda \to 1$ ja , $b_\lambda \to 1$, saame (15.3) tõttu, et äärmised tõenäosused ülaltoodud reas koonduvad λ kasvamisel suuruseks $\Phi(t)$ (siin Φ on standardse normaaljaotuse jaotusfunktsioon), seega koondub ka keskmine. Äärmiste tõenäosuste koondumine järeldub ühtlasest koondumisest:

$$P(T_n \leq t) \to \Phi(t), \quad \forall t \quad \Rightarrow \quad \sup_t |P(T_n \leq t) - \Phi(t)| \to 0, \tag{15.4}$$

mis (sarnaselt Glivenko-Cantelli teoreemi tõestusega) järeldub Φ pidevusest ja jaotusfunktsioonide monotoonsusest.

4. **Taastuv protsess.** Olgu $T_1, T_2, \ldots$ mittenegatiivsed iid juhuslikud suurused, $ET_i = \mu > 0$ ja $D(T_i) = \sigma^2 < \infty$. Näiteks T_1 on mingi sündmuse (kliendi või bussi saabumine, õnnetus, õnnestumine, elektripirni läbipõlemine) toimumise aeg, T_2 on esimese ja teise sündmuse toimumise vaheline aeg jne, olgu $S_0 = 0$ ja $S_n = T_1 + \cdots + T_n$. Defineerime iga $t \geq 0$ korral

$$N(t) = \max\{n : S_n \leq n\}.$$

Seega $N(t)$ on juhuslik suurus, mis loeb, mitu sündmust ajavahemikus $[0,t]$ toimus. Funktsioon $t \mapsto N(t)$ on tükati konstantne, paremalt pidev, mittekahanev ja minimaalne hüppe pikkus on 1. Suurte arvude seaduse tõttu protsessis $t \to \infty$ kehtib:

$$N(t) \to \infty \text{ p.k. }, \frac{S_{N(t)}}{N(t)} \to \mu \text{ p.k.}, \frac{N(t)}{t} \to \frac{1}{\mu} \text{ p.k..}$$

(vt Peatükk 11, ülesanne 12). Tsentraalse piirteoreemi tõttu protsessis $k \to \infty$

$$\frac{S_k - \mu k}{\sqrt{k}} \Rightarrow \mathcal{N}(0, \sigma^2),$$

mistõttu (ülesanne 13, ptk 11) protsessis $t \to \infty$

$$\frac{S_{N(t)} - \mu N(t)}{\sqrt{t/\mu}} \Rightarrow \mathcal{N}(0, \sigma^2), \quad \frac{S_{N(t)} - \mu N(t)}{\sqrt{N(t)}} \Rightarrow \mathcal{N}(0, \sigma^2)$$

(need on koondumised (11.25).) Nüüd kasutame asjaolu, et iga t korral $N(t)$ definitsioonist järeldub

$$S_{N(t)} \leq t \leq S_{N(t)+1},$$

millest

$$\frac{S_{N(t)} - \mu N(t)}{\sqrt{N(t)}} \leq \frac{t - \mu N(t)}{\sqrt{N(t)}} \leq \frac{S_{N(t)} - \mu N(t)}{\sqrt{N(t)}} + \frac{T_{N(t)+1}}{\sqrt{N(t)}}.$$

Veendu, et

$$\frac{T_{N(t)+1}}{\sqrt{N(t)}} \xrightarrow{P} 0$$

ja järelda (Slutsky), et

$$\frac{t - \mu N(t)}{\sqrt{N(t)}} \Rightarrow \mathcal{N}(0, \sigma^2).$$

Et $\frac{N(t)}{t} \to \frac{1}{\mu}$, p.k., siis ülaltoodud koondumisest järeldub (jälle Slutsky)

$$\frac{t - \mu N(t)}{\sqrt{t/\mu}} \Rightarrow \mathcal{N}(0, \sigma^2)$$

ja (normaaljaotus on sümmeetriline)

$$\frac{\mu N(t) - t}{\sqrt{t/\mu}} = \frac{N(t) - t/\mu}{\sqrt{t/\mu^3}} \Rightarrow \mathcal{N}(0, \sigma^2).$$

Seega suure t korral sündmuste toimumise arvu ajahetkeni t saab lähendada $\mathcal{N}(t/\mu, \sigma^2 t/\mu^3)$ jaotusega.

15.1.1 Tsentraalne piirteoreem, nõrk suurte arvude seadus ja kordse logaritmi seadus

Tsentraalsest piirteoreemist järeldub nõrk suurte arvude seadus. Olgu

$$T_n = \frac{S_n - n\mu}{\sigma\sqrt{n}}.$$

Tsentraalne piirteoreem: $T_n \Rightarrow \mathcal{N}(0,1)$. Slutsky lemmast järeldub, et iga nulliks koonduva reaalarvude jada $a_n \to 0$ korral $a_n T_n \Rightarrow 0$ ehk $a_n T_n \xrightarrow{P} 0$. Võttes

$$a_n = \sigma n^{\frac{1}{2}-\alpha},$$

saame: kui $\alpha > \frac{1}{2}$, siis

$$a_n T_n = \sigma n^{\frac{1}{2}-\alpha} \frac{S_n - n\mu}{\sigma n^{\frac{1}{2}}} = \frac{S_n - n\mu}{n^{\alpha}} \xrightarrow{P} 0.$$

Jada T_n ise muidugi tõenäosuse järgi nulliks koonduda ei saa, sest ta koondub jaotuse järgi millekski muuks. Kas ta üldse koondub tõenäosuse järgi? Selgub, et mitte: kuigi $T_n \Rightarrow \mathcal{N}(0,1)$, ei koondu juhuslike suuruste jada T_n tõenäosuse järgi. Veendume selles. Lihtsuse mõttes võtame $\mu = 0$ ja $\sigma = 1$, siis $T_n = \frac{S_n}{\sqrt{n}}$. Oletades vastuväiteliselt, et leidub juhuslik suurus X (mis jaotusega ta peaks olema?) nii, et $T_n \xrightarrow{P} X$, saame, et ka $T_{2n} \xrightarrow{P} X$ ja seetõttu (kahe tõenäosuse järgi koonduva jada vahe koondub) $T_{2n} - T_n \xrightarrow{P} 0$. Nüüd aga pane tähele, et

$$T_{2n} - T_n = \frac{S_{2n}}{\sqrt{2n}} - \frac{S_n}{\sqrt{n}} = \frac{1}{\sqrt{2}}\Big(\frac{S_{2n} - S_n}{\sqrt{n}}\Big) - \big(1 - \frac{1}{\sqrt{2}}\big)\frac{S_n}{\sqrt{n}}.$$

Juhuslikud suurused $T_n = \frac{S_n}{\sqrt{n}}$ ja $\frac{S_{2n}-S_n}{\sqrt{n}}$ on sõltumatud, sama jaotusega ning mõlemad koonduvad jaotuse järgi standardseks normaaljaotuseks. Seega leiduvad sõltumatud standardse normaaljaotusega juhuslikud suurused X ja Y nii, et

$$\frac{S_n}{\sqrt{n}} \Rightarrow X, \quad \frac{S_{2n} - S_n}{\sqrt{n}} \Rightarrow Y.$$

Sellest järeldub aga, et

$$T_{2n} - T_n = \frac{1}{\sqrt{2}}\Big(\frac{S_{2n} - S_n}{\sqrt{n}}\Big) - \big(1 - \frac{1}{\sqrt{2}}\big)\frac{S_n}{\sqrt{n}} \Rightarrow \frac{1}{\sqrt{2}}Y - \big(1 - \frac{1}{\sqrt{2}}\big)X.$$

See aga on vastuolus eeldusega $T_{2n} - T_n \xrightarrow{P} 0$ (miks?)

Kuidas aga käitub jada $T_n(\omega)$? Eelnevast teame (kuidas?), et ei leidu sellist (normaaljaotusega) juhuslikku suurust X nii, et peaaegu iga ω korral $T_n(\omega) \to X(\omega)$ (p.k. koondumine). Tsentraalsest piirteoreemist järeldub

$$\limsup_n T_n = \infty, \quad \liminf_n T_n = -\infty, \quad \text{p.k.} \tag{15.5}$$

Tõestus põhineb järgmisel ülesandel.

Ülesanne. Olgu $T_1, T_2, \ldots$ suvaline juhuslike suuruste jada. Tõesta, et iga reaalarvu t korral

$$\mathbf{P}(\limsup_n T_n \geq t) \geq \limsup_n \mathbf{P}(T_n > t). \tag{15.6}$$

Kui $T_n \Rightarrow \mathcal{N}(0,1)$, siis

$$\limsup_n \mathbf{P}(T_n > t) = \lim_n \mathbf{P}(T_n > t) = 1 - \Phi(t) > 0,$$

kus Φ on standardse normaaljaotuse jaotusfunktsioon. Sündmus $\{\limsup_n T_n > t\}$ on jääk σ-algebra element, seega tema tõenäosus on kas 1 või 0 (Kolmogorovi 0-1 seadus), antud juhul on see tõenäosus positiivne, seega 1. Rakendades sama argumenti juhuslikele suurustele $-X_i$, saame

$$\mathbf{P}(\limsup_n -T_n > t) = \mathbf{P}(\liminf_n T_n < -t) > 0$$

ja Kolmogorovi 0-1 seadusest järeldub jällegi, et $\mathbf{P}(\liminf_n T_n = -\infty) = 1$. Seega peaaegu iga ω-korral jada $T_n(\omega)$ kõigub järjest suureneva amplituudiga, tema ülemine piirväärtus on ∞ ja alumine piirväärtus on $-\infty$. Sellest omakorda järeldub, et jada $S_n - n\mu$ kõigub kiiremini kui $\sqrt{n}$, st kuitahes suure K korral on $S_n - n\mu$ lõpmata palju kordi suurem kui $K\sqrt{n}$ (st $S_n - \mu n \geq K\sqrt{n}$ i.o.) ja lõpmata palju kordi väiksem kui $-K\sqrt{n}$ (st $S_n - \mu n \leq -K\sqrt{n}$ i.o.). Selgub, et selle kõikumise kiirust saab täpsustada, sest kehtib järgmine kuulus teoreem. Järgnevas teoreemis on üldisust kitsendamata eeldatud, et $\mu = 0$.

Teoreem 15.2 (Kordse logaritmi seadus). *Olgu $X_1, X_2, \ldots$ iid juhuslikud suurused $EX_i = 0$ ja $EX_i^2 = \sigma^2$. Siis*

$$\mathbf{P}\Big(\limsup_n \frac{S_n}{\phi(n)} = 1\Big) = 1, \quad \mathbf{P}\Big(\liminf_n \frac{S_n}{\phi(n)} = -1\Big) = 1, \tag{15.7}$$

kus

$$\phi(n) := \sqrt{2\sigma^2 \cdot n \cdot \ln\ln n}.$$

Teoreemist järeldub, et kuitahes väikese $\epsilon > 0$ ja peaaegu iga ω korral

$$\begin{aligned} &S_n(\omega) \geq (1-\epsilon)\phi(n), \quad \text{i.o.} \quad \text{ja} \quad S_n(\omega) \leq (1+\epsilon)\phi(n), \quad \text{ev} \\ &S_n(\omega) \leq (-1+\epsilon)\phi(n), \quad \text{i.o.} \quad \text{ja} \quad S_n(\omega) \geq (-1+\epsilon)\phi(n), \quad \text{ev.} \end{aligned}$$

Seega kuitahes väikese $\epsilon > 0$ ja peaaegu iga ω korral $T_n = \frac{S_n}{\sigma\sqrt{n}}$ käitub järgmiselt

$$\begin{aligned} &T_n(\omega) \geq (1-\epsilon)\sqrt{2\ln\ln n}, \quad \text{i.o.} \qquad \text{ja} \quad T_n(\omega) \leq (1+\epsilon)\sqrt{2\ln\ln n}, \quad \text{ev} \\ &T_n(\omega) \leq (-1+\epsilon)\sqrt{2\ln\ln n}, \quad \text{i.o.} \qquad \text{ja} \quad T_n(\omega) \geq (-1-\epsilon)\sqrt{2\ln\ln n}, \quad \text{ev.} \end{aligned}$$

Neis võrratustest järeldub muuhulgas, et $\limsup_n T_n = \infty$ p.k ja $\liminf_n T_n = -\infty$ p.k. (kuidas?). Kõige lõpuks paneme aga tähele, et

$$\frac{S_n}{\phi(n)} \xrightarrow{P} 0$$

(miks?). Kas see koondumine on vastuolus seostega (15.7)?

15.2 Tsentraalsed piirteoreemid üldistel eeldustel

15.2.1 Normeeritud seeriad

Olgu $X_1, X_2, \ldots$ lõpliku dispersiooniga iid juhuslikud suurused, justnagu teoreemi 15.1 eelduses. Tähistame

$$Y_{kn} := \frac{X_k - \mu}{\sqrt{n\sigma^2}}, \quad n = 1, 2, \ldots, \quad k = 1, 2, \ldots, n. \tag{15.8}$$

Sellisel juhul

$$EY_{kn} = 0, \quad \sum_{k=1}^{n} DY_{kn} = 1 \tag{15.9}$$

ja

$$\frac{S_n - n\mu}{\sqrt{n\sigma^2}} = Y_{1n} + \ldots + Y_{nn} =: T_n. \tag{15.10}$$

Seega tsentraalse piirteoreemi 15.1 väide toodud tähistustes on $T_n \Rightarrow N(0,1)$.

Seeriate skeem. Üldistame ülalkirjeldatud (iid) situatsiooni ja vaatleme sõltumatuid juhuslikke suurusi, nn *seeriaid*

$$Y_{1n}, \ldots, Y_{nn},$$

mis rahuldavad seoseid (15.9) ning uurime tingimusi, mis garanteerivad koondumise $T_n \Rightarrow N(0,1)$. Paneme tähele: meid huvitab jada $\{T_n\}$ jaotuse järgi koondumine. Selline koondumine sõltub aga T_n jaotustest. Seega võib olla, et iga n korral on seeria $Y_{1n}, \ldots, Y_{nn}$ defineeritud omaette tõenäosusruumil. Seose (15.8) abil defineeritud seeriate korral on juhuslikud suurused $Y_{1n}, \ldots, Y_{nn}$ iid, nüüd aga loobume eeldusest, et seeriasse kuuluvad juhuslikud suurused on sama jaotusega (kuid ikka sõltumatud!) ja vaatame seeriaid üldisemalt. Tsentraalsete piirteoreemide esitust (võib olla, et erinevatel tõenäosusruumidel defineeritud) seeriate abil nimetatakse *seeriate skeemiks*.

Esitame seeriate skeemi abil ühe väga üldise tsentraalse piirteoreemi. Olgu F_{kn} juhusliku suuruse Y_{kn} jaotusfunktsioon, Φ olgu standardse normaaljaotuse jaotusfunktsioon. Tähistame $\sigma_{kn}^2 = DY_{kn}$ ja $\Phi_{kn}(x) = \Phi(\frac{x}{\sigma_{nk}})$. Veendu, et Φ_{kn} on $N(0, \sigma_{nk}^2)$ jaotusfunktsioon.

Teoreem 15.3 *Olgu $T_n := Y_{1n} + \cdots + Y_{nn}$. Selleks et kehtiks koondumine*

$$T_n \Rightarrow N(0,1) \tag{15.11}$$

on tarvilik ja piisav, et iga $\epsilon > 0$ korral kehtib

$$\sum_{k=1}^{n} \int_{\{|x|>\epsilon\}} |x||F_{kn}(x) - \Phi_{kn}(x)|dx \to 0. \tag{15.12}$$

Piisavuse tõestuse ja viite tarvilikkuse tõestusele võib leida *Shirjajev, III,* §6.

Lindebergi tingimus. Tingimus (15.12) on väga üldine ja ehk raskesti kontrollitav. Järgnev teoreem näitab, et (15.12) on aga täidetud, kui seeriad $Y_{1n}, \ldots, Y_{nn}$ rahuldavad nn **Lindebergi tingimust**: iga $\epsilon > 0$ korral

$$\lim_n \sum_{k=1}^{n} \int_{\{|Y_{kn}|\geq\epsilon\}} Y_{kn}^2 d\mathbf{P} = 0. \tag{15.13}$$

Lemma 15.1 *Kui seeriad $Y_{1n}, \ldots, Y_{nn}$ rahuldavad Lindebergi tingimust (15.13), siis*

$$\max_{1\leq k\leq n} EY_{kn}^2 \to 0. \tag{15.14}$$

Tõestus. Iga $k = 1, \ldots, n$ ja iga $\epsilon > 0$ korral

$$EY_{kn}^2 \leq \epsilon^2 + \int_{\{|Y_{kn}|\geq\epsilon\}} Y_{kn}^2 d\mathbf{P},$$

millest

$$\max_{1\leq k\leq n} EY_{kn}^2 \leq \epsilon^2 + \sum_{k=1}^{n} \int_{\{|Y_{kn}|\geq\epsilon\}} Y_{kn}^2 d\mathbf{P}.$$

■

Teoreem 15.4 **1.** *Kui seeriad $Y_{1n}, \ldots, Y_{nn}$ rahuldavad Lindebergi tingimust (15.13), siis kehtib ka (15.12) ning seega $T_n \Rightarrow N(0,1)$.*

2. *Kui seeriad $Y_{1n}, \ldots, Y_{nn}$ rahuldavad tingimust (15.14), siis on Lindebergi tingimus (15.13) ja tingimus (15.12) ekvivalentsed.*

Selle teoreemi tõestuses kasutame Lebesgue'i integraali ositi integreerimise valemit. Sõnastame selle lemmana. Tõestuse võib leida *Shirjajev*, II § 6, *Billingsley*, 18.

Lemma 15.2 (Ositi integreerimine). *Olgu P reaalteljel olev tõenäosusmõõt, F olgu P jaotusfunktsioon. Olgu h pidevalt diferentseeruv funktsioon, kusjuures $\int |h|dP < \infty$. Sellisel juhul kehtib ositi inegreerumise valem*

$$\int_{(a,b]} hdP =: \int_{(a,b]} hdF = h(b)F(b) - h(a)F(a) - \int_{(a,b]} F(x)h'(x)dx,$$

kus $-\infty \leq a \leq b \leq \infty$ ning $F(-\infty) = \lim_{x\to-\infty} F(x)$, $F(\infty) = \lim_{x\to\infty} F(x)$, $h(-\infty) = \lim_{x\to-\infty} h(x)$, $h(\infty) = \lim_{x\to\infty} h(x)$.

On lihtne veenduda, et iga konstandi c korral

$$h(b)F(b)-h(a)F(a)-\int_{(a,b]} F(x)h'(x)dx = h(b)(c+F(b))-h(a)(c+F(a))-\int_{(a,b]} (c+F(x))h'(x)dx.$$

Viimast asjaolu võib kirjutada

$$\int hd(c+F) = \int hdF.$$

Olgu G ja F kaks jaotusfunktsiooni, kusjuures $\int |h|dF < \infty$ ja $\int |h|dG < \infty$. Defineerime

$$\int hd(F+G) := \int hdF + \int hdG, \quad \int hd(F-G) := \int hdF - \int hdG.$$

On kerge näha, et ositi integreerimise valem kehtib st iga $-\infty \leq a \leq b \leq \infty$ korral

$$\int_{(a,b]} hd(F+G) = h(b)(F(b)+G(b)) - h(a)(F(a)+G(a)) - \int_{(a,b]} h'd(F+G),$$
$$\int_{(a,b]} hd(F-G) = h(b)(F(b)-G(b)) - h(a)(F(a)-G(a)) - \int_{(a,b]} h'd(F-G).$$

Teoreemi 15.4 tõestus **1** Kehtigu (15.13). Näitame, et kehtib (15.12). Hindame: iga $k = 1, \ldots, n$ ja iga $\epsilon > 0$ korral

$$\int_{\{x:|x|\geq\epsilon\}} x^2 d\Phi_{kn} = \sigma_{kn}^2 \int_{\{\sigma_{kn}|Y|>\epsilon\}} Y^2 d\mathbf{P} \leq \sigma_{kn}^2 \int_{\left\{|Y|>\frac{\epsilon}{\sqrt{\max_k \sigma_{kn}^2}}\right\}} Y^2 d\mathbf{P},$$

kus Y on standardse normaaljaotusega juhuslik suurus. Et $\sum_{k=1}^n \sigma_{kn}^2 = 1$, ja lemma 15.1 tõttu kehtib (15.14), saame

$$\sum_{k=1}^n \int_{\{x:|x|\geq\epsilon\}} x^2 d\Phi_{kn} \leq \int_{\left\{|Y|>\frac{\epsilon}{\sqrt{\max_k \sigma_{kn}^2}}\right\}} Y^2 d\mathbf{P} \to 0. \tag{15.15}$$

Lindebergi tingimuse (15.13) võib formuleerida

$$\sum_{k=1}^n \int_{\{x:|x|\geq\epsilon\}} x^2 dF_{kn} = \sum_{k=1}^n \int_{\{|Y_{kn}|\geq\epsilon\}} Y_{kn}^2 d\mathbf{P} \to 0.$$

Seega (15.15) ning (15.13) annavad

$$\sum_{k=1}^n \int_{\{x:|x|\geq\epsilon\}} x^2 d[F_{kn} + \Phi_{kn}] \to 0. \tag{15.16}$$

Olgu h selline paarisfunktsion, et

$$h(x) = \begin{cases} x^2, & \text{kui } |x| > 2\epsilon; \\ 0, & \text{kui } |x| \leq \epsilon. \end{cases}$$

ning $|h(x)| \leq x^2$ ja $|h'(x)| \leq 4x$.
Funktsiooni h definitsioonist ja (15.16)

$$\sum_{k=1}^{n} \int_{\{x:|x|\geq\epsilon\}} h(x)d[F_{kn} + \Phi_{kn}] \leq \sum_{k=1}^{n} \int_{\{x:|x|\geq\epsilon\}} x^2 d[F_{kn} + \Phi_{kn}] \to 0.$$

Integreerime ositi

$$\int_{[\epsilon,\infty)} h(x)d(1 - F_{kn}) = \lim_{x\to\infty}(1 - F_{kn}(x))h(x) - h(\epsilon)(1 - F_{kn}(\epsilon)) - \int_{[\epsilon,\infty)} h'(x)(1 - F_{kn}(x))dx,$$
$$\int_{[\epsilon,\infty)} h(x)d(1 - \Phi_{kn}) = \lim_{x\to\infty}(1 - \Phi_{kn}(x))h(x) - h(\epsilon)(1 - \Phi_{kn}(\epsilon)) - \int_{[\epsilon,\infty)} h'(x)(1 - \Phi_{kn}(x))dx.$$

Et $\int x^2 dF_{kn} < \infty$ ja $\int x^2 d\Phi_{kn} = \sigma_{nk}^2 EY^2 \leq EY^2 < \infty$ siis

$$\lim_{x\to\infty}(1-F_{kn}(x))h(x) = \lim_{x\to\infty}(1-F_{kn}(x))x^2 = 0, \quad \lim_{x\to\infty}(1-\Phi_{kn}(x))h(x) = \lim_{x\to\infty}(1-\Phi_{kn}(x))x^2 = 0.$$

Funktsiooni h definitsioonist järeldub, et $h(\epsilon)(1-F_{kn}(\epsilon)) = h(\epsilon)(1-\Phi_{kn}(\epsilon)) = 0$. Kokkuvõttes saame

$$\int_{[\epsilon,\infty)} h(x)dF_{kn} = \int_{[\epsilon,\infty)} h'(x)(1 - F_{kn}(x))dx, \quad \int_{[\epsilon,\infty)} hd\Phi_{kn} = \int_{[\epsilon,\infty)} h'(x)(1 - \Phi_{kn}(x))dx, \tag{15.17}$$

millest

$$\sum_{k=1}^{n} \int_{\{x\geq\epsilon\}} h'(x)[(1-F_{kn}(x))+(1-\Phi_{kn}(x))]dx = \sum_{k=1}^{n} \int_{\{x:x\geq\epsilon\}} h(x)d[F_{kn}+\Phi_{kn}] \to 0. \tag{15.18}$$

Analoogiliselt

$$\int_{(-\infty,-\epsilon]} h(x)dF_{kn} = h(-\epsilon)F_{kn}(-\epsilon) - \lim_{x\to-\infty} F_{kn}(x)h(x) - \int_{(-\infty,-\epsilon]} h'(x)F_{kn}(x)dx,$$
$$\int_{(-\infty,-\epsilon]} h(x)d\Phi_{kn} = h(-\epsilon)\Phi_{kn}(-\epsilon) - \lim_{x\to-\infty} \Phi_{kn}(x)h(x) - \int_{(-\infty,-\epsilon]} h'(x)\Phi_{kn}(x)dx.$$

Et $\int x^2 dF_{kn} < \infty$ ja $\int x^2 d\Phi_{kn} = \sigma_{nk}^2 EY^2 \leq EY^2 < \infty$ siis

$$\lim_{x\to-\infty} \Phi_{kn}(x)h(x) = \lim_{x\to-\infty} \Phi_{kn}(x)x^2 = 0, \quad \lim_{x\to-\infty} \Phi_{kn}(x)h(x) = \lim_{x\to-\infty} \Phi_{kn}(x)x^2 = 0.$$

Seega

$$\int_{(-\infty,-\epsilon]} h(x)dF_{kn} = -\int_{(-\infty,\epsilon]} h'(x)F_{kn}(x)dx, \quad \int_{(-\infty,\epsilon]} hd\Phi_{kn} = -\int_{(-\infty,\epsilon]} h'(x)\Phi_{kn}(x)dx. \tag{15.19}$$

Nüüd

$$\sum_{k=1}^{n}\int_{\{x\leq -\epsilon\}} h'(x)[F_{kn}(x)+\Phi_{kn}(x)]dx = -\sum_{k=1}^{n}\int_{\{x:x\leq -\epsilon\}} h(x)d[F_{kn}+\Phi_{kn}] \to 0. \quad (15.20)$$

Et $h'(x)=2x$, kui $x\geq 2\epsilon$, saame seostest (15.18) ja (15.20)

$$\sum_{k=1}^{n}\int_{\{|x|\geq 2\epsilon\}} |x||F_{kn}(x)-\Phi_{kn}(x)|dx \to 0.$$

2 Kehtigu (15.14) ja (15.12). Tõestame, et kehtib (15.13). Olgu h ülaldefineeritud funktsioon. Et $\max_k \sigma_{kn}^2 \to 0$, siis koondumisest (15.15) saame

$$\sum_{k=1}^{n}\int_{\{|x|\geq\epsilon\}} h(x)d\Phi_{kn} \leq \sum_{k=1}^{n}\int_{\{|x|\geq\epsilon\}} x^2 d\Phi_{kn} \to 0.$$

Seostest (15.17) ja (15.19) ja võttes arvesse, et $|h'(x)|\leq 4x$, saame

$$\Big|\sum_{k=1}^{n}\int_{\{|x|\geq\epsilon\}} h(x)d[F_{kn}-\Phi_{kn}]\Big| = \Big|\sum_{k=1}^{n}\int_{\{x\geq\epsilon\}} h(x)d[F_{kn}-\Phi_{kn}]\Big| + \Big|\sum_{k=1}^{n}\int_{\{x\leq-\epsilon\}} h(x)d[F_{kn}-\Phi_{kn}]\Big|$$

$$= \Big|\sum_{k=1}^{n}\int_{\{x\geq\epsilon\}} h'(x)[(1-F_{kn}(x))-(1-\Phi_{kn}(x))]dx\Big| + \Big|\sum_{k=1}^{n}\int_{\{x\leq-\epsilon\}} h'(x)[F_{kn}(x)-\Phi_{kn}(x)]dx\Big|$$

$$\leq \sum_{k=1}^{n}\int_{\{x\geq\epsilon\}} |h'(x)||(1-F_{kn}(x))-(1-\Phi_{kn}(x))|dx + \sum_{k=1}^{n}\int_{\{x\leq-\epsilon\}} |h'(x)||F_{kn}(x)-\Phi_{kn}(x)|dx$$

$$\leq 4\sum_{k=1}^{n}\int_{\{|x|\geq\epsilon\}} |x||F_{kn}(x)-\Phi_{kn}(x))|dx \to 0.$$

Kokkuvõttes,

$$\sum_{k=1}^{n}\int_{\{|x|\geq 2\epsilon\}} x^2 dF_{kn} \leq \sum_{k=1}^{n}\int_{\{|x|\geq\epsilon\}} h(x)dF_{kn} = \sum_{k=1}^{n}\int_{\{|x|\geq\epsilon\}} h(x)d[F_{kn}-\Phi_{kn}] + \sum_{k=1}^{n}\int_{\{|x|\geq\epsilon\}} h(x)d\Phi_{kn} \to 0.$$

15.2.2 Normeerimata seeriad

Vaatleme normeerimata seeriaid: iga n korral on juhuslikud suurused

$$X_{1n},\ldots,X_{nn}$$

sõltumatud, kuid ei rahulda tingimusi (15.9) (keskväärtused on null ja dispersioonide summa üks). Kasutamaks eelmises peatükis kasutatud teoreeme, tuleb igasse seeriasse kuuluvad juhuslikud suurused $X_{1n},\ldots,X_{nn}$ tsentreerida ja normeerida nii, et dispersioonide summa oleks üks. Tähistame

$$\mu_{kn}=EX_{kn}, \quad \sigma_{kn}^2 = DX_{kn}, \quad s_n^2=\sum_{k=1}^{n}\sigma_{kn}^2.$$

Defineerides iga n korral

$$Y_{kn} := \frac{X_{kn} - \mu_{kn}}{s_n}, \quad k = 1, \dots, n,$$

saame, et juhuslikud suusused $Y_{1n}, \dots, Y_{nn}$ on endiselt sõltumatud, kuid rahuldavad nüüd seoseid (15.9). Tsentraalne piirteoreem Lindebergi tingimustel on normeerimata reeriate korral selline:

Järeldus 15.1 *Kui iga $\epsilon > 0$ korral*

$$\lim_{n\to\infty} \frac{1}{s_n^2} \sum_{k=1}^{n} \int_{\{|X_{kn}-\mu_{kn}|>\epsilon s_n\}} (X_{kn} - \mu_{kn})^2 d\mathbf{P} = 0, \tag{15.21}$$

siis kehtib tsentraalne piirteoreem

$$\frac{\sum_{k=1}^{n}(X_{kn} - \mu_{kn})}{s_n} \Rightarrow N(0,1). \tag{15.22}$$

Juhul, kui

$$\max_{k=1,\dots,n} \frac{\sigma_{kn}^2}{s_n^2} \to 0, \tag{15.23}$$

on (15.22) ja (15.21) ekvivalentsed.

Näide. Kui (15.23) ei kehti, pole (15.21) tsentraalseks piirteoreemiks (15.22) tarvilik. Olgu näiteks $X_1, X_2, \dots$ sõltumatud normaaljaotusega juhuslikud suurused, kusjuures $EX_i = 0$,

$$DX_1 = 1, DX_2 = \frac{1}{2}, \dots, DX_k = 2^{1-k}.$$

$$X_{11} = X_1, \quad X_{12} = X_1, \quad X_{22} = X_2, \quad \dots \quad X_{kn} = X_k.$$

Seega $S_n = X_1 + \dots + X_n$ ning

$$\sum_{k=1}^{n} \sigma_{kn}^2 = \sum_{k=1}^{n} DX_k = 1 + \frac{1}{2} + \left(\frac{1}{2}\right)^2 + \dots + \left(\frac{1}{2}\right)^{n-2} \leq 2.$$

Tähistame $s_n^2 = \sum_{k=1}^{n} DX_k$.
Seega

$$\max_{k=1,\dots,n} \frac{\sigma_{kn}^2}{s_n^2} > \frac{1}{2}$$

ehk (15.23) ei kehti. Et (15.23) on tarvilik Lindebergi tingimuse kehtimiseks, näeme, et see tingimus ei kehti. Lindebergi tingimuse (15.21) mittekehtivust saab ka vahetult kontrollida. Tõepoolest, et

$$\sum_{k=1}^{n} \int_{\{|X_{kn}-\mu_{kn}|>\epsilon s_n\}} X_{kn}^2 d\mathbf{P} = \sum_{k=1}^{n} \int_{\{|X_k|>\epsilon s_n\}} X_k^2 d\mathbf{P} \geq \sum_{k=1}^{n} \int_{\{|X_k|>2\epsilon\}} X_k^2 d\mathbf{P},$$

näeme, et n kasvades on summa mittekahanev. Küll aga on triviaalselt kehtiv (15.22), sest

$$\frac{\sum_{k=1}^{n}(X_{kn}-\mu_{kn})}{s_n} = \frac{\sum_{k=1}^{n}(X_k-\mu_k)}{\sqrt{\sum_{k=1}^{n} DX_n}}$$

on standardse normaaljaotusega.
Vastavalt teoreemile 15.3 peab aga kehtima (15.12), kus

$$Y_{kn} = \frac{X_k}{s_n}.$$

Veendume selles: et $X_k \sim N(0, DX_k)$, siis

$$F_{kn}(x) = \mathbf{P}(Y_{kn} \le x) = \mathbf{P}(X_k \le xs_n), \quad \Phi_{kn}(x) = \mathbf{P}\Big(Y \le \frac{s_n}{\sqrt{DX_k}}x\Big) = \mathbf{P}(X_k \le xs_n),$$

kus $Y \sim N(0,1)$. See tähendab, et $F_{kn} = \Phi_{kn}$ iga k ja n korral.

Ljapunovi tingimus. Veendume, et Lindebergi tingimus (15.21) on täidetud, kui seeriad $X_{1n},\dots,X_{nn}$ rahuldavad nn. **Ljapunovi tingimusi**

$$\exists\delta > 0: \quad \lim_{n\to\infty}\frac{1}{s_n^{2+\delta}}\sum_{k=1}^{n} E|X_{kn}-\mu_{kn}|^{2+\delta} = 0. \tag{15.24}$$

Et hulgal $\{|X_{kn}-\mu_{kn}| \ge \epsilon s_n\}$ kehtib võrratus

$$|X_{kn}-\mu_{kn}|^{2+\delta} \ge (X_{kn}-\mu_{kn})^2(\epsilon s_n)^\delta,$$

saame

$$(\epsilon s_n)^\delta \int_{\{|X_{kn}-\mu_{kn}|\ge\epsilon s_n\}} (X_{kn}-\mu_{kn})^2 d\mathbf{P} \le \int_{\{|X_{kn}-\mu_{kn}|\ge\epsilon s_n\}} |X_{kn}-\mu_{kn}|^{2+\delta} d\mathbf{P} \le E|X_{kn}-\mu_{kn}|^{2+\delta}.$$

Seega

$$\frac{1}{s_n^2}\sum_{k=1}^{n}\int_{\{|X_{kn}-\mu_{kn}|>\epsilon s_n\}} (X_{kn}-\mu_{kn})^2 d\mathbf{P} \le \frac{1}{\epsilon^\delta s_n^{2+\delta}}\sum_{k=1}^{n} E|X_{kn}-\mu_{kn}|^{2+\delta}$$

ning (15.24) garanteerib (15.21). Seega kehtib

Järeldus 15.2 *Olgu $X_{1n},\dots,X_{nn}$ sõltumatute juhuslike suuruste seeriad, mis rahuldavad Ljapunovi tingimusi (15.24). Siis kehtib tsentraalne piirteoreem (15.22).*

Järeldus 15.3 *Olgu $X_{1n},\dots,X_{nn}$ sõltumatute juhuslike suuruste seeriad, kusjuures leidub jada M_n nii, et*

$$|X_{kn}-\mu_{kn}| \le M_n, \quad \text{p.k.} \quad \frac{M_n}{s_n} \to 0. \tag{15.25}$$

Siis kehtib kehtib tsentraalne piirteoreem (15.22).

Tõestus. Võta $\delta = 1$ ja hinda

$$\sum_{k=1}^{n} E|X_{kn} - \mu_{kn}|^3 \leq M_n \sum_{k=1}^{n} \sigma_{kn}^2 = M_n s_n^2,$$

millest saame tingimuse (15.24)

$$\frac{\sum_{k=1}^{n} E|X_{kn} - \mu_{kn}|^3}{s_n^3} \leq \frac{M_n}{s_n} \to 0.$$

■

Näide: juhuslikud permutatsioonid. Olgu iga n korral elemetaarsündmus ω hulga $\{1, \ldots, n\}$ permutatsioon. Kokku $n!$ elementaarsündmust. Olgu iga permutatsiooni (elementaarsündmuse) tõenäosus $\frac{1}{n!}$. Defineerime juhuslikud suurused X_{kn} järgmiselt: $X_{kn}(\omega) = j$ kui permutatsioonis ω täpselt j elementi hulgast $\{1, 2, \ldots k-1\}$ on peale arvu k. Seega $X_{1n} = 0$,

$$X_{2n}(\omega) = \begin{cases} 0, & \text{kui permutatsioonis } \omega \text{ on 1 enne kui 2;} \\ 1, & \text{kui permutatsioonis } \omega \text{ on 2 enne kui 1.} \end{cases}$$

$$X_{3n}(\omega) = \begin{cases} 0, & \text{kui permutatsioonis } \omega \text{ on 1 ja 2 enne kui 3;} \\ 1, & \text{kui permutatsioonis } \omega \text{ on üks hulga } \{1,2\} \text{ element enne kui 3;} \\ 2, & \text{kui permutatsioonis } \omega \text{ on 3 enne kui 1 ja 2:} \end{cases},$$

Seega juhusliku suuruse X_{kn} võimalikud väärtused on $0, 1, \ldots, k-1$. Selgub, et need juhuslikud suurused on sõltumatud ja ühtlase jaotusega, st

$$\mathbf{P}(X_{kn} = j) = \frac{1}{k}, \quad j = 0, 1, \ldots, k-1.$$

Näiteks, kui $n = 3$, siis võimalikud elementaarsündmused on

$$\omega_1 = (1,2,3), \quad \omega_2 = (1,3,2) \quad \omega_3 = (3,1,2) \quad \omega_4 = (3,2,1) \quad \omega_5 = (2,1,3) \quad \omega_6 = (2,3,1).$$

Nüüd $X_{23} = 0$ siis kui permutatsioonis on 1 enne kui 2 ja selliseid on 3 ehk $\mathbf{P}(X_{23} = 0) = 1/2$. Vaatleme juhuslikku suurust X_{33}. Tal on kolm võimalikku väärtust: 0,1,2. Permutatsioone, mille korral $X_{33} = 1$ on täpselt 2 ((2,3,1) ja (1,3,2)) ehk $\mathbf{P}(X_{33} = 1) = 1/3$. Permutatsioone, mille korral $X_{33} = 2$ on samuti täpselt 2 ((3,2,1) ja (3,1,2)). Seega $X_{3,3}$ on ühtlase jaotusega. Veendume, et $X_{2,3}$ ja $X_{3,3}$ on sõltumatud:

$$\mathbf{P}(X_{23} = 0, X_{33} = 1) = \mathbf{P}(\omega_2) = \frac{1}{6} = \mathbf{P}(X_{23} = 0)\mathbf{P}(X_{33} = 1),$$
$$\mathbf{P}(X_{23} = 0, X_{33} = 2) = \mathbf{P}(\omega_3) = \frac{1}{6} = \mathbf{P}(X_{23} = 0)\mathbf{P}(X_{33} = 2).$$

Leiame

$$\mu_{kn} = EX_{kn} = \frac{1}{k}\sum_{j=1}^{k-1} j = \frac{k-1}{2}, \quad \sum_{k=1}^{n} \mu_{kn} = \frac{n(n-1)}{4}.$$

Dispersiooni arvutamiseks kasutame *Faulhaberi valemit*

$$\sum_{j=1}^{n} j^2 = \frac{2n^3+3n^2+n}{6}.$$

Seega

$$\sigma_{kn}^2 = DX_{kn} = \frac{1}{k}\sum_{j=1}^{k-1} j^2 - \frac{(k-1)^2}{4} = \frac{(k-1)(2k-1)}{6} - \frac{(k-1)^2}{4} = \frac{k^2-1}{12}.$$

Seega

$$s_n^2 = \sum_{k=1}^{n} \sigma_{kn}^2 = \frac{1}{12}\Big(\sum_{k=1}^{n} k^2 - n\Big) = \frac{1}{6\cdot 12}\Big(2n^3+3n^2+n-6n\Big) = \frac{n^3}{36} + \frac{n^2}{24} - \frac{5n}{72}.$$

Nüüd kontrollime järelduse eeldusi (15.25). Iga ω korral kehtib

$$|X_{kn}(\omega) - \mu_{kn}| = |X_{kn}(\omega) - \frac{(k-1)}{2}| \le \frac{k-1}{2} \le \frac{n-1}{2} < n =: M_n.$$

Et

$$\frac{M_n}{s_n} = \frac{n}{\sqrt{\frac{n^3}{36} + \frac{n^2}{24} - \frac{5n}{72}}} \to 0,$$

näeme, et järelduse tingimused on täidetud ja

$$\frac{S_n - \frac{n(n-1)}{4}}{\sqrt{\frac{n^3}{36} + \frac{n^2}{24} - \frac{5n}{72}}} \Rightarrow N(0,1).$$

Sellest järeldub (kuidas?), et

$$\frac{S_n - \frac{n^2}{4}}{\frac{n^{\frac{3}{2}}}{6}} \Rightarrow N(0,1).$$

Summa $S_n = X_1 + \cdots + X_n$ on teatavas mõttes juhusliku permutatsiooni erinevus permutatsioonist $1,\ldots,n$ ja õigesti skaleerituna on see arv suure n korral ligikaudu standardse normaaljaotusega.

15.2.3 Sõltumatute juhuslike suuruste jada

Olgu $X_1, X_2, \ldots$ sõltumatute juhuslike suuruste jada, $\mu_k = EX_k$ ja $\sigma_k^2 = DX_k$ ja

$$s_n^2 = \sum_{k=1}^{n} \sigma_k^2.$$

Kui defineerida seeriad

$$X_{11} = X_1, \quad X_{12} = X_1,\, X_{22} = X_2, \quad X_{13} = X_1,\, X_{23} = X_2,\, X_{33} = X_3, \ldots$$

siis ülatoodud tulemused annavad tsentraalse piirteoreemi juhuslike suuruste osasummade jadale $S_n = X_1 + \cdots + X_n$.
Tsentraalne piirteoreem Lindebergi tingimustel on siis selline.

Järeldus 15.4 *Kui*

$$\lim_{n\to\infty} \frac{1}{s_n^2} \sum_{k=1}^{n} \int_{\{|X_k-\mu_k|>\epsilon s_n\}} (X_k - \mu_k)^2 d\mathbf{P} = 0, \tag{15.26}$$

siis kehtib tsentraalne piirteoreem

$$\frac{\sum_{k=1}^{n}(X_k - \mu_k)}{s_n} \Rightarrow N(0,1). \tag{15.27}$$

Juhul, kui

$$\max_{k=1,\ldots,n} \frac{\sigma_k^2}{s_n^2} \to 0, \tag{15.28}$$

on (15.27) ja (15.26) ekvivalentsed.

Märkus: Paneme tähele, et kui leidub vähemalt üks $\sigma_k^2 > 0$ (vastasel juhul on kõik X_i-d p.k. konstandid ja CLT on triviaalne), siis seosest (15.28) järeldub, et $s_n \to \infty$. Teisõnu, kui dispersioonide summa on tõkestatud, st $s_n^2 \to s^2 < \infty$ (monotoone jada alati koondub), siis Lindebergi tingimus kehtida ei saa. Kui (15.28) ei kehti, siis Lindebergi tingimuse mittekehtimisest ei järeldu küll, et CLT ei kehti, kuid saab näidata, et praegusel juhul kehtib järgmine kasulik väide: kui $s_n^2 \to s^2 < \infty$, siis CLT kehtib vaid siis, kui kõik juhuslikud suurused $X_1, X_2, \ldots$ on normaaljaotusega (nagu eelmises näites).

Tsentraalne piirteoreem Ljapunovi tingimustel on sellisel juhul järgmine

Järeldus 15.5 *Kui*

$$\exists \delta > 0 : \quad \lim_{n\to\infty} \frac{1}{s_n^{2+\delta}} \sum_{k=1}^{n} E|X_k - \mu_k|^{2+\delta} = 0, \tag{15.29}$$

siis kehtib (15.27).

Järeldus 15.6 *Olgu $X_1, X_2, \ldots$ sõltumatud tõkestatud juhuslikud suurused, s.t. $\exists K < \infty$ nii, et $|X_k| \leq K$ iga k korral. Kui dispersioonide summa hajub, st*

$$\sum_{k=1}^{\infty} \sigma_k^2 = \infty,$$

siis kehtib tsentraalne piirteoreem (15.27).

Tõestus. Tõkestatud juhuslikul suurusel eksisteerivad kõik momendid. Seega

$$\frac{1}{s_n^3} \sum_{k=1}^{n} E|X_k - \mu_k|^3 \leq \frac{1}{s_n^3} \sum_{k=1}^{n} E|X_k - \mu_k|^2 2K = \frac{2K}{s_n} \to 0,$$

sest dispersioonide summa hajumine tähendab, et $s_n^2 \to \infty$. Seega kehtib Ljapunovi tingimus (15.29), kus $\delta = 1$. ■

Lõpuks näitame, et tsentraalne piirteoreem i.i.d. juhuslike suuruste korral, teoreem 15.1, järeldub väga lihtsalt üldisematest tulemustest.

Järeldus 15.7 *Teoreem 15.1 järeldub järeldusest 15.4.*

Tõestus. Olgu $X_1, X_2, \ldots$ i.i.d. juhuslikud suurused. Näitame, et kehtivad Lindebergi tingimused. Et $s_n^2 = \sigma^2 n$ ja juhuslikud suurused on sama jaotusega, saame tingimusele (15.26) kuju

$$\frac{1}{n\sigma^2} \sum_{k=1}^{n} \int_{\{|X_k - \mu_k| > \epsilon s_n\}} (X_k - \mu_k)^2 d\mathbf{P} = \frac{1}{\sigma^2} \int_{\{|X_1 - \mu_1| > \epsilon\sqrt{n\sigma^2}\}} (X_1 - \mu_1)^2 d\mathbf{P} \to 0,$$

sest n kasvades $\mathbf{P}\{|X_1 - \mu_1| > \epsilon\sqrt{n\sigma^2}\} \searrow 0$. Nüüd kasuta väidet 9.1 veendumaks, et parempoolne integraal koondub nulliks. ■

15.3 Ühtlane koondumine

Olgu S_n selliste seeriate osasumma, et kehtib tsentraalne piirteoreem:

$$\frac{S_n - ES_n}{\sqrt{DS_n}} \Rightarrow N(0,1).$$

Vastavalt nõrga koondumise definitsioonile, $\forall x \in \mathbb{R}$

$$\mathbf{P}\Big(S_n \leq \sqrt{DS_n}x + ES_n\Big) = \mathbf{P}\Big(\frac{S_n - ES_n}{\sqrt{DS_n}} \leq x\Big) \to \Phi(x), \tag{15.30}$$

kus Φ on $N(0,1)$ jaotusfunktsioon. Tähistades $y_n = \sqrt{DS_n}x + ES_n$, on (15.30) ekvivalentne koondumisega

$$\Big|\mathbf{P}(S_n \leq y_n) - \Phi\big(\frac{y_n - ES_n}{\sqrt{DS_n}}\big)\Big| \to 0.$$

Kas aga kehtib koondumine

$$\Big|\mathbf{P}(S_n \le y) - \Phi\Big(\frac{y - ES_n}{\sqrt{DS_n}}\Big)\Big| \to 0? \tag{15.31}$$

Teisisõnu, kas võime öelda, et piisavalt suure n korral

$$\mathbf{P}(S_n \le y) \approx \Phi\Big(\frac{y - ES_n}{\sqrt{DS_n}}\Big)?$$

Pelgalt seosest (15.30) ülaltoodud koondumine (15.31) ei järeldu. Tõepoolest, tähistades

$$x_n = \frac{y - ES_n}{\sqrt{DS_n}},$$

saame, et (15.31) on ekvivalentne koondumisega

$$\Big|\mathbf{P}\Big(\frac{S_n - ES_n}{\sqrt{DS_n}} \le x_n\Big) - \Phi(x_n)\Big| \to 0 \tag{15.32}$$

Koondumine (15.32) ei järeldu otseselt punktiviisilisest koondumisest, küll aga järeldub see *ühtlasest koondumisest*

$$\sup_x \Big|\mathbf{P}\Big(\frac{S_n - ES_n}{\sqrt{DS_n}} \le x\Big) - \Phi(x)\Big| \to 0. \tag{15.33}$$

(vt ka (15.4)). Veendu, et (15.33) kehtib (selleks kasuta Φ pidevust ja jaotusfunktsioonide monotoonsust).

Teatud tingimustel saab koondumiste (15.30) ja (15.33) kiirusi hinnata. Olgu $X_1, X_2, \ldots$ i.i.d juhuslikus suurused, kusjuures $E|X_1|^3 < \infty$, ja (üldisust kitsendamata) $EX_1 = 0$. Olgu F_n seeriate summa $\frac{S_n}{\sigma\sqrt{n}}$ jaotusfunktsioon. Sellisel juhul kehtib punktiviislilne hinnang

$$|F_n(x) - \Phi(x)| \le \frac{CE|X_1|^3}{\sigma^3\sqrt{n}} \frac{1}{(1+|x|)^3}, \tag{15.34}$$

kus C on mingi konstant. Seosest (15.34) järeldub nn **Berry-Esseeni hinnang**

$$\sup_x |F_n(x) - \Phi(x)| \le \frac{CE|X_1|^3}{\sigma^3\sqrt{n}}. \tag{15.35}$$

Saab näidata, et konstant C hinnangus (15.35) rahuldab võrratusi $(2\pi)^{-\frac{1}{2}} \le C < 0.8$.

Näide: Näitame, et hinnangut (15.35) ei saa parandada. Olgu $X_1, X_2, \ldots$ i.i.d juhuslikud suurused, kusjuures $\mathbf{P}(X_1 = -1) = \mathbf{P}(X_1 = 1) = \frac{1}{2}$. Siis $E|X_1|^3 = \sigma = 1$.

Olgu F_n juhusliku suuruse $\frac{S_n}{\sqrt{n}}$ jaotusfunktsioon. Hindame funktsiooni F_{2n} hüpet punktis 0.

$$\begin{aligned} F_{2n}(0) - F_{2n}(0-) &= 1 - \mathbf{P}\Big(\frac{S_{2n}}{\sqrt{2n}} > 0\Big) - \mathbf{P}\Big(\frac{S_{2n}}{\sqrt{2n}} < 0\Big) \\ &= 1 - \mathbf{P}(S_{2n} > 0\Big) - \mathbf{P}(S_{2n} < 0) = \mathbf{P}(S_{2n} = 0) = C_{2n}^n(\frac{1}{2})^{2n}. \end{aligned}$$

Stirlingi valemist $n! \sim \sqrt{2\pi n}(n)^n e^{-n}$ saame

$$C_{2n}^n(\frac{1}{2})^{2n} \sim \frac{1}{\sqrt{\pi n}},$$

millest

$$F_{2n}(0) - F_{2n}(0-) = \frac{1}{\sqrt{\pi n}}(1 + o(1)).$$

Funktsioon Φ on aga punktis 0 pidev, mistõttu

$$\sup_x \Big|F_{2n}(x) - \Phi(x)\Big| \geq \frac{1}{\sqrt{2\pi 2n}}(1 + o(1)).$$

Seega ei saa konstant C hinnangus (15.35) olla väiksem kui $\frac{1}{\sqrt{2\pi}}$.

15.4 Lokaalsed piirteoreemid

15.4.1 Pidev jaotus

Oletame, et $X_1, X_2, \dots$ i.i.d. lõpliku teist järku momendiga pidevad juhuslikud suurused. Sellisel juhul on pidevad ka juhuslikud suurused

$$\frac{S_n - ES_n}{\sqrt{DS_n}}. \tag{15.36}$$

Olgu f_n (15.36) tihedus. Tsentraalne piirteoreem väidab, et

$$\sup_x \Big|F_n(x) - \Phi(x)\Big| = \sup_x \Big| \int_{(-\infty,x]} f_n(t)dt - \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{\frac{-t^2}{2}} dt\Big| \to 0. \tag{15.37}$$

Sellest aga ei järeldu, et ühtlaselt koonduksid tihedused

$$\sup_x \Big|f_n(x) - \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}}\Big| \to 0. \tag{15.38}$$

Tõepoolest, jaotuse järgi koondumine tähendab ju jaotusfunktsioonide punktiviisilist koondumist, mitte aga ilmtingimata tihedusfunktsioonide punktiviisilit koondumist (sellest, et $F_n(t) \to F(t)$ iga t korral ei järeldu ju ka, et $F_n'(t) \to F(t) =$ iga t korral). Kontranäitena vaatleme järgmist ülesannet.

Ülesanne: Olgu

$$F_n(x) = \begin{cases} 1, & \text{kui } x \geq 1; \\ x - \frac{\sin(2n\pi x)}{2n\pi}, & \text{kui } x \in [0,1]; \\ 0, & \text{kui } x \leq 0 \end{cases}$$

- Veendu, et F_n on iga n korral mingi jaotuse, olgu see P_n, jaotusfunktsioon.
- Veendu, et $P_n \Rightarrow P$, kus P on mingi jaotus, leia P.
- Veendu, et jaotustel P_n ja P on tihedused, olgu need f_n ja f, (Lebesgue'i mõõdu suhtes).
- Veendu, et $\text{Leb}\{x \in [0,1] : f_n(x) \to f(x)\} = 0$.

Küll aga järeldub koondumisest (15.38) koondumine (15.37) (Scheffe teoreem 7.19).

Teatud tingimustel saab aga tõestada koondumise (15.38). Selliseid koondumisi nimetatakse **lokaalseteks (tsentraalseteks) piirteoreemideks**.

Teoreem 15.5 (lokaalne piirteoreem). *Olgu $X_1, X_2, \dots$ lõpliku teist järku momendiga i.i.d. juhuslikud suurused. Oletame, et juhuslike suuruste X_i jaotuse karakteristlik funktsioon φ rahuldab tingimust*

$$\exists r \geq 1 : \quad \int_{-\infty}^{\infty} |\varphi(t)|^r dt < \infty.$$

Siis kehtib (15.38).

Teoreemi tõestuse võib leida *Grimmet& Stirzaker* lk. 177.

Teoreemi (15.5) eeldustel on $|\varphi(t)|^n$ integreeruv iga piisavalt suure n korral. Et S_n karakteristlik funktsioon on punktis t on $\varphi^n(t)$, millest seose $|\varphi^n(t)| \leq |\varphi(t)|^n$ tõttu on S_n karakteristlik funktsioon integreeruv. Järelduse 14.4 tõttu on summal S_n pidev tihedusfunktsioon f_n. Pidev tihedusfunktsioon on Riemanni mõttes integreeruv ja koondumisel (15.37) on seega kuju

$$\sup_x \Big| \int_{-\infty}^x f_n(t)dt - \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{\frac{-t^2}{2}} dt \Big| \to 0.$$

15.4.2 Võrejaotus

Def 15.6 *Ütleme, et diskreetne jaotus P on* **võrejaotus** *kui $\exists a \in \mathbb{R}, h > 0$ nii, et kõik P aatomid sisalduvad hulgas*

$$\{a, a \pm h, a \pm 2h, a \pm 3h, \dots\}. \tag{15.39}$$

Edaspidi eeldame, et h on suurim selline konstant (*span*). Näiteks, kui $P\{-1\} = P\{1\} = \frac{1}{2}$, siis $h = 2$, sest P aatomid sisalduvad hulgas $\{-1, -1 \pm 2, -1 \pm 2h, -1 \pm 3h, \dots\}$. Võrejaotused on näiteks Poissoni jaotus, binoomjaotus, geomeetriline jaotus (kõigil $h = 1$). Saab näidata, et P on võrejaotus parajasti siis, kui leidub $t \neq 0$ nii, et $\varphi(t) = 1$, kuid $|\varphi(u)| < 1$ iga $u \in (0, t)$ korral. Siin φ on P karakteristlik funktsioon (vt *Durrett*, thm 3.5.1). Sel juhul on $h = \frac{2\pi}{t}$ ja φ on perioodiline. Triviaalne erijuht võrejaotusest on Diraci jaotus, sellisel juhul on karakteristlik funktsioon konstantselt 1.

Olgu $X_1, X_2, \ldots$ võrejaotusega i.i.d. juhuslikud suurused. Üldisust kitsendamata eeldame, et $EX_i = 0$. Olgu $DX_i = \sigma^2 < \infty$. Et X_i väärtused sisalduvad hulgas

$$\{a, a \pm h, a \pm 2h, a \pm 3h, \ldots\} = \{a + hz : z \in \mathbb{Z}\},$$

siis S_n väärtused sisalduvad hulgas $\{an + hz : z \in \mathbb{Z}\}$ ja $\frac{S_n}{\sigma\sqrt{n}}$ väärtused sisalduvad hulgas

$$\mathcal{L}_n := \Big\{\frac{(an + hz)}{\sigma\sqrt{n}} : z \in \mathbb{Z}\Big\}.$$

Tehtud eeldustel kehtib järgmine lokaalne piirteoreem. Siin ϕ on standardse normaaljaotuse tihedusfunktsioon.

Teoreem 15.7 (Gnedenko, 1948)

$$\sup_{x \in \mathcal{L}_n} \Big| \frac{\sigma\sqrt{n}}{h} \mathbf{P}\Big(\frac{S_n}{\sigma\sqrt{n}} = x\Big) - \phi(x)\Big| \to 0 \tag{15.40}$$

Teoreemi tõestust vaata näiteks (*Durrett*, thm 3.5.2). Punktid x moodustavad $\frac{S_n}{\sigma\sqrt{n}}$ väärtuste piirkonna, kahe järjestikuse punkti vaheline kaugus on $\frac{h}{\sigma\sqrt{n}}$. Seega koondumist (15.40) võib interpreteerida nii: piisavalt suure n korral

$$\mathbf{P}\big(\frac{S_n}{\sigma\sqrt{n}} = x\big) \approx \int_{x-\frac{1}{2}\frac{h}{\sigma\sqrt{n}}}^{x+\frac{1}{2}\frac{h}{\sigma\sqrt{n}}} \phi(s)ds \approx \frac{h}{\sigma\sqrt{n}}\phi(x).$$

Seos (15.40): $x = \frac{an+hz}{\sigma\sqrt{n}}$, millest

$$\sup_{k \in \mathbb{Z}} \Big| \frac{\sigma\sqrt{n}}{h} \mathbf{P}\Big(\frac{S_n}{\sigma\sqrt{n}} = \frac{an + kh}{\sigma\sqrt{n}}\Big) - \frac{1}{\sqrt{2\pi}} e^{-\frac{(kh+an)^2}{2\sigma^2 n}} \Big| =$$
$$\sup_{k \in \mathbb{Z}} \Big| \frac{\sigma\sqrt{n}}{h} \mathbf{P}\big(S_n = kh + an\big) - \frac{1}{\sqrt{2\pi}} e^{-\frac{(kh+an)^2}{2\sigma^2 n}} \Big| \to 0.$$

Teisisõnu, leidub jada $\delta_n \to 0$ nii, et

$$\sup_{k \in \mathbb{Z}} \Big| \mathbf{P}(S_n = kh + an) - \frac{h}{\sqrt{2\pi\sigma^2 n}} \exp(-\frac{(kh + an)^2}{2\sigma^2 n})\Big| = \frac{\delta_n}{\sqrt{n}} = o(\frac{1}{\sqrt{n}}). \tag{15.41}$$

Näide: Olgu $X_1, X_2, \ldots$ iid $B(1,p)$-jaotusega. Siis $X_1 - p, X_2 - p, \ldots$ on tsentreeritud (keskväärtusega 0) juhuslikud suurused, mille väärtused sisalduvad hulgas

$$\{-p, -p \pm 1, -p \pm 2, -p \pm 3, \ldots\},$$

st $a = -p$ ja $h = 1$. Olgu $S_n = X_1 + \cdots + X_n$ (binoomjaotusega juhuslik suurus). Seos (15.41) kehtib tsentreeritud juhuslike suuruste summa $S_n - np$ korral. Et aga $an = -pn$, siis saame, et iga mittenegatiivse täisarvu $k \in \mathbb{N}$ korral $\mathbf{P}(S_n = k) = \mathbf{P}(S_n - pn = k + an)$ ja seos (15.41) on nüüd

$$\sup_{k\in\mathbb{N}} \Big| \mathbf{P}(S_n = k) - \frac{1}{\sqrt{2\pi p(1-p)n}} \exp[-\frac{(k-np)^2}{2np(1-p)}] \Big| \leq \frac{\delta_n}{\sqrt{n}}. \tag{15.42}$$

Tähistame

$$\phi_n(k) := \frac{1}{\sqrt{2\pi p(1-p)n}} \exp[-\frac{(k-np)^2}{2np(1-p)}].$$

Siis (15.42) on

$$\sup_{k\in\mathbb{N}} \Big| \mathbf{P}(S_n = k) - \phi_n(k) \Big| \leq \frac{\delta_n}{\sqrt{n}} \tag{15.43}$$

ehk iga k korral

$$\Big| \frac{\mathbf{P}(S_n = k)}{\phi_n(k)} - 1 \Big| \leq \frac{\delta_n}{\phi_n(k)\sqrt{n}}. \tag{15.44}$$

Kui leidub konstant $A < \infty$ nii, et $|k - np| \leq A\sqrt{n}$ ehk $|k - np| = O(\sqrt{n})$, siis leidub konstant $b > 0$ nii, et $\sqrt{n}\phi_n(k) \geq b$ (miks?) ja seetõttu kehtib koondumine

$$\sup_{k:|k-np|\leq A\sqrt{n}} \Big| \frac{\mathbf{P}(S_n = k)}{\phi_n(k)} - 1 \Big| \to 0. \tag{15.45}$$

Koondumine (15.45) on tuntud kui **DeMoivre-Laplace'i lokaalne piirteoreem** ja selle tõestuse esitas P.- S. Laplace juba aastal 1795. Selgub, et ühtlane koondumine kehtib ka üle laiema hulga. Nimelt kui g_n on reaalarvude jada, mis rahuldab tingimust $n^{-\frac{2}{3}} g_n \to 0$ ehk $g_n = o(n^{\frac{2}{3}})$, siis kehtib koondumine

$$\sup_{k:|k-np|\leq g_n} \Big| \frac{\mathbf{P}(S_n = k)}{\phi_n(k)} - 1 \Big| \to 0 \tag{15.46}$$

(tõestust vaata näiteks *Shirjajev, lk. 68*). Veendu, et $A\sqrt{n} = o(n^{\frac{2}{3}})$, seega koondumisest (15.46) järeldub (15.45), kuid mitte vastupidi.

15.5 Poissoni teoreem

Olgu $X_{1n}, \ldots, X_{nn}$ normeerimata seeriaid. Eelpool vaatlesime tingimusi mis garanteerivad tsentraalse piirteoreemi ehk koondumise

$$\frac{S_n - \mu_n}{s_n} \Rightarrow N(0,1), \tag{15.47}$$

kus

$$S_n = \sum_{k=1}^{n} X_{kn}, \quad \mu_n = \sum_{k=1}^{n} \mu_{kn}, \quad s_n^2 = \sum_{k=1}^{n} \sigma_{kn}^2.$$

Oletame, et kehtivad koondumised $\mu_n \to \mu$ ja $s_n \to \sigma$. Sellisel juhul ülaltoodud koondumisest järeldub, et $S_n \Rightarrow N(\mu, \sigma)$ (miks?).

Vaatleme seeriaid $X_{1n}, \ldots, X_{nn}$, kus X_{kn} on Bernoulli p_{kn}-jaotusega juhuslik suurus, s.t.

$$X_{kn} \sim B(1, p_{kn})$$

ning

$$\max_k p_{kn} \to 0 \quad \sum_{k=1}^{n} p_{kn} \to \lambda. \tag{15.48}$$

Seega

$$\mu_n = \sum_{k=1}^{n} p_{kn} \to \lambda, \quad s_n^2 = \sum_{k=1}^{n} p_{kn}(1 - p_{kn}) \to \lambda.$$

(pole raske veenduda, et kui $\max_k p_{kn} \to 0$, siis koondumised $\sum_{k=1}^{n} p_{kn}(1 - p_{kn}) \to \lambda$ ja $\sum_{k=1}^{n} p_{in} \to \lambda$ on ekvivalentsed). Seega kehtib tsentraalne piirteoreem parajasti siis, kui $S_n \Rightarrow N(\lambda, \lambda)$.

Veendume, et toodud seeriad rahuldavad tingimust (15.23):

$$\max_k \frac{\sigma_{kn}^2}{s_n^2} = \max_k \frac{p_{kn}(1 - p_{kn})}{s_n^2} \to 0,$$

sest $\max_k p_{kn} \to 0$. Seega on tsentraalne piirteoreem ehk koondumine $S_n \Rightarrow N(\lambda, \lambda)$ ekvivalentne Lindebergi tingimusega (15.21). On aga kerge veenduda, et kui $\lambda < \infty$, siis Lindebergi tingimus pole täidetud. Olgu ϵ nii väike, et $\frac{3}{2}\epsilon\sqrt{\lambda} < 1$. Olgu n nii suur, et

$$s_n > \frac{\sqrt{\lambda}}{2}, \quad \max_k \mu_{kn} < \epsilon \frac{\sqrt{\lambda}}{2} < \epsilon s_n.$$

Sellisel juhul $1 - \mu_{kn} > 1 - \epsilon\frac{\sqrt{\lambda}}{2} > \epsilon\sqrt{\lambda} > \epsilon s_n$, millest

$$\int_{\{|X_{kn} - \mu_{kn}| > \epsilon s_n\}} (X_{kn} - \mu_{kn})^2 d\mathbf{P} = (1 - \mu_{kn})^2 p_{kn} = (1 - p_{kn})^2 p_{kn}$$

ja

$$\frac{1}{s_n^2} \sum_{k=1}^{n} \int_{\{|X_{kn} - \mu_{kn}| > \epsilon s_n\}} (X_{kn} - \mu_{kn})^2 d\mathbf{P} = \frac{\sum_{k=1}^{n} (1 - p_{kn})^2 p_{kn}}{\sum_{k=1}^{n} (1 - p_{kn}) p_{kn}} \to 1.$$

Seega tsentraalne piirteoreem ei kehti ja S_n ei koondu normaaljaotuseks. Selgub, et osasummade jada S_n koondub nõrgalt hoopis Poissoni jaotuseks.

Teoreem 15.8 (Poissoni teoreem) *Olgu X_{in}, $i = 1, \ldots, n$ sõltumatud seeriad, kusjures $X_{in} \sim B(1, p_{in})$ ja p_{in} rahuldavad tingimusi (15.48). Siis $S_n \Rightarrow \mathrm{Po}(\lambda)$ ehk*

$$\mathbf{P}(S_n = m) \to e^{-\lambda} \frac{\lambda^m}{m!}, \quad \forall\, m = 0, 1, 2, \ldots. \tag{15.49}$$

Tõestus. Tõestus põhineb asjaolul: kui $x_{in} \in \mathbb{C}$, $i = 1, \ldots, n$ ning $\lim_{n\to\infty} \sum_{i=1}^n x_{in} = x$ ja $\lim_{n\to\infty} \max_i x_{in} = 0$, siis

$$\prod_{i=1}^{n} \Big(1 + x_{in}\Big) \to e^x, \tag{15.50}$$

sest $\sum_{i=1}^n \ln(1 + x_{in}) \sim \sum_{i=1}^n x_{in}$.
Juhusliku suuruse X_{kn} karakteristlik funktsioon on $p_{kn}e^{it} + (1 - p_{kn})$, millest summa S_n karaktersitlik funktsioon

$$\varphi_{S_n}(t) = \prod_{k=1}^{n} \Big(p_{kn}e^{it} + (1 - p_{kn})\Big) = \prod_{k=1}^{n} \Big(1 + p_{kn}(e^{it} - 1)\Big).$$

Seosest (15.50) saame, et iga $t \in \mathbb{R}$ korral

$$\varphi_{S_n}(t) = \prod_{k=1}^{n} \Big(1 + p_{kn}(e^{it} - 1)\Big) \to \exp[\lambda(e^{it} - 1)].$$

Funktsioon $t \mapsto \exp[\lambda(e^{it} - 1)]$ on aga Poissoni jaotuse $\mathrm{Po}(\lambda)$ karakteristlik funktsioon. Teoreemist 14.5 saame

$$S_n \Rightarrow Po(\lambda).$$

Nii S_n kui ka Poissoni jaotuse on aatomid sisalduvad mittenegatiivsete täisarvude hulgas. Sellisel juhul on nõrk koondumine ekvivalente aatomite masside koondumisega (15.49) (veendu selles). ■

Seega, kui $X_{1,n}, \ldots, X_{n,n}$ on iid $B(1, \frac{\lambda}{n})$, siis $X_{1,n} + \cdots + X_{n,n} \Rightarrow Po(\lambda)$. Teisest küljest, $X_{1,n} + \cdots + X_{n,n} \sim B(n, \frac{\lambda}{n})$. Siit tuleneb praktikas kasulik järeldus: kui n on "piisavalt suur" ja np "piisavalt väike", siis binoomjaotus $B(n, p)$ on lähedane Poissoni jaotusele $Po(np)$. (Kui np ei ole "piisavalt väike", võib binoomjaotust lähendada nii Poissoni kui ka normaljaotusega, sest nagu teame, parameetri (np) kasvades Poissoni jaotus läheneb normaaljaotusele.)

Näited:

1. Olgu grupis 400 inimest, X olgu nende inimeste arv, kellel on sünnipäev 16. detsembril. Ignoreerides liigaastaid, saame $X \sim B(400, \frac{1}{365})$. Poissoni lähend sellele jaotusele on $Po(\frac{400}{365})$. Seega tõenäosus, et ühelgi neist pole sünnipäev 16. detsembril, on $\mathbf{P}(X = 0) = \left(\frac{364}{365}\right)^{400} \approx 0,334$, Poissoni lähend on $\exp[-\frac{400}{365}] \approx 0.334$.

2. Kahte täringut visatakse 36 korda. Olgu X nende visete arv, mis annab paari 6 ja 6; $X \sim B(36, \frac{1}{36})$. Poissoni lähend sellele jaotusele on $Po(1)$. Järgnev tabel näitab, et tegemist on hea lähendiga:

k	0	1	2	3
$Po(1)$	0.3678	0.3678	0.1839	0.0613
$B(36, \frac{1}{36})$	0.3627	0.3730	0.1865	0.0604

3. Radioaktiivse aatomi lagunemisel kiirgub α-osake. Olgu teada, et 1 grammi radioaktiivse aine lagunemisel kiirgub keskmiselt λ osakest sekundis. Vaatleme 1 grammi ainet kui n aatomi kogumit. Igaüks neist laguneb ja kiirgab α-osakese 1 sekundi jooksul tõenäosusega p. Seega keskmiselt laguneb $np = \lambda$ aatomit, millest $p = \frac{\lambda}{n}$. Seega 1 sekundi jooksul lagunenud aatomite arv, X, on jaotusega $B(n, \frac{\lambda}{n})$. Selle jaotuse Poissoni lähend on $Po(\lambda)$. Oletame, et meid huvitab tõenäosus, et 1 sekundi jooksul ei kiirgu üle 2 osakese. Kasutades Poissoni lähendit, saame
$$\mathbf{P}(X \leq 2) \approx e^{-\lambda} + \lambda e^{-\lambda} + \frac{\lambda^2}{2} e^{-\lambda}.$$
1920.-l aastal viisis Rutherford, Chadwick ja Ellis läbi järgmise katse. Nad registreerisid 7.5 sekundi jooksul kiirgunud α-osakesi. Selleks jagasid nad pika ajalõigu 7.5 sekundi pikkusteks intervallideks ja lugesid kokku igas ajaintervallis registreeritud osakeste arvu. Kokku registreeriti 10094 α-osakest $N := 2608$ intervalli jooksul. Seega keskmiselt $\frac{10094}{2608} \approx 3.87$ osakest intervallis. Seega ajaintervallis registreeritud osakeste arv on ligikaudu $Po(3.87)$ jaotusega. Järgnevas tabelis on N_k nende intervallide arv, kus mõõdeti k osakest. Võrdluseks on
$$N \exp[-3.87] \frac{3.87^k}{k!} =: Np_k.$$

k	0	1	2	3	4	5	6	7	8	9	≥ 10
N_k	57	203	383	525	532	408	273	139	45	27	16
Np_k	54.4	210.5	407.3	525.5	508.4	393.5	253.8	140.3	67.8	29.2	17

4. Olgu $\xi_1, \ldots, \xi_n$ $U(a,b)$ jaotusega juhuslikud suurused. Jagame intervalli $[a,b]$ $k(n)$ võrdse pikkusega alamlõiguks, kusjuures $\frac{n}{k(n)} \to \lambda$. Olgu Δ_n esimene alamlõik (pikkusega $\frac{b-a}{k(n)}$). Defineerime
$$X_{in} := I_{\Delta_n}(\xi_i).$$
On selge, et $X_{in} \sim B(1, \frac{1}{k(n)})$, $\sum_{i=1}^n p_{in} = \frac{n}{k(n)} \to \lambda$. Seega suure n korral alamlõiku Δ_n sattunud ξ_i-de arv on ligikaudu Poissoni jaotusega $Po(\lambda)$. Üldistub mitmemõõtmelisele juhule.
Kui teise maailmasõja aegne London jagada 0.25 ruutkilomeetri suurusteks ruutudeks, saame 576 ruutu. Olgu N_k nende ruurude arv, kuhu sõja ajal langes k pommi. Kokku loeti 537 pommi, seega keskmiselt $\frac{537}{575} =: \lambda$ pommi ruudu kohta. Tabel

k	0	1	2	3	4	≥ 5
N_k	229	211	93	35	7	1
Np_k	226.7	211.3	98.5	30.6	7.14	1.6

15.6 Lõpmatult jagunevad jaotused

Vaatleme seeriaid $Y_{1,n}, \ldots, Y_{n,n}$, $n = 1, 2, \ldots$ ning eeldame, et seeriasse kuuluvad juhuslikud suurused on lisaks sõltumatusele ka sama jaotusega (kuid ei pruugi olla normeeritud

ja tsentreeritud). Vaatleme seeriate summasi $T_n = Y_{1,n} + \ldots + Y_{n,n}$ ning uurime neid jaotusi, mis avalduvad jada T_n piirjaotustena.
Tsentraalsetest piirteoreemidest teame, et võimalike piirjaotuste hulka kuuluvad kõik normaaljaotused.
Poissoni teoreemist saame, et kui $Y_{i,n} \sim B(1, \frac{\lambda}{n})$, siis $T_n \Rightarrow P_o(\lambda)$, s.t. võimalike piirjaotuste hulka kuuluvad kõik Poissoni jaotused. Kas aga iga jaotus P avaldub piirjaotusena $T_n \Rightarrow P$, kus T_n sõltumatute ja sama jaotusega juhuslike suuruste seeria summa?
Selgub, et jaotus P avaldub jada T_n piirjaotusena parajsti siis, kui ta on *lõpmatult jagunev.*

Def 15.9 *Jaotust P (juhuslikku suurust X, karakteristlikku funktsiooni φ) nimetatakse* **lõpmatult jagunevaks**, *kui iga $k \geq 1$ korral leiduvad sõltumatud ja sama jaotusega juhuslikud suurused $Z_1, \ldots, Z_k$ nii, et summa $Z_1 + \cdots + Z_k$ jaotus on P (summa jaotus on X jaotus, summa karakteristlik funktsioon on φ).*

Definitsioonist järeldub, et karaktersitlik funktsioon φ on lõpmatult jagunev parajasti siis, kui iga k korral leiduvad karakteristlikud funktsioonid $\varphi_k(t)$ nii, et

$$\varphi(t) = [\varphi_k(t)]^k. \tag{15.51}$$

Näited:

1. Normaaljaotuse $N(\mu, \sigma)$ karakteristlik funktsioon on

$$e^{it\mu - \frac{t^2}{2}\sigma^2} = \left[e^{i\frac{t\mu}{k} - \frac{t^2\sigma^2}{2k}}\right]^k = \left[e^{it\frac{\mu}{k} - \frac{t^2}{2}\frac{\sigma^2}{k}}\right]^k.$$

Funktsioon

$$e^{it\frac{\mu}{k} - \frac{t^2}{2}\frac{\sigma^2}{k}}$$

on $N(\frac{\mu}{k}, \frac{\sigma}{\sqrt{k}})$ karakteristlik funktsioon. Seega kehtib (15.51) ehk $N(\mu, \sigma)$ lõpmatult jagunev jaotus.

2. Poissoni jaotuse $\mathrm{Po}(\lambda)$ karakteristlik funktsioon on $\exp[\lambda(e^{it} - 1)]$, mis avaldub kujul

$$\exp[\lambda(e^{it} - 1)] = \Big(\exp[\frac{\lambda}{k}(e^{it} - 1)]\Big)^k.$$

Funktsioon $\exp[\frac{\lambda}{k}(e^{it} - 1)]$ on $\mathrm{Po}(\frac{\lambda}{k})$ karakteristlik funktsioon. Seega kehtib (15.51) ehk $\mathrm{Po}(\lambda)$ on lõpmatult jagunev jaotus.

3. Gamma-jaotuse tihedusfunktsioon on

$$f(x) = \frac{x^{\alpha-1}e^{\frac{-x}{\beta}}}{\Gamma(\alpha)\beta^\alpha}, \quad \text{kui } x \geq 0.$$

Siin α, β on parameetrid. Selle jaotuse karakteristlik funktsioon on

$$\varphi(t) = \frac{1}{(1 - i\beta t)^\alpha} = [\frac{1}{(1 - i\beta t)^{\frac{\alpha}{k}}}]^k.$$

Seega on gamma-jaotus lõpmatult jagunev.

Paneme tähele, et kahe sõltumatu lõpmatult jaguneva jaotusega juhusliku suuruse summa jaotus on lõpmatult jagunev. Nii näiteks on sõltumatute normaal- ja Poissoni jaotusega juhusliku suuruse summa jaotus lõpmatult jagunev.

Teoreem 15.10 *Jaotus P avaldub jada T_n piirjaotusena parajasti siis, kui ta on lõpmatult jagunev.*

Tõestus. $\Leftarrow$ Olgu P lõpmatult jagunev. Vastavalt definitsioonile leidub iga n korral sõltumatute ja sama jaotusega juhuslike suuruste seeria, nii, et seeria summa T_n jaotus on P. Sellisel juhul, muidugi, $T_n \Rightarrow P$.

$\Rightarrow$ Leidugu sellised (sõltumatud ja sama jaotusega) seeriad, et nende summad T_n koonduvad nõrgalt jaotuseks P. Veendume, et iga k korral leiduvad sellised sõltumatud ja sama jaotusega juhuslikud suurused $Z_1, \ldots, Z_k$ nii, et $Z_1 + \cdots + Z_k \sim P$. Siis on P lõpmatult jagunev.
Fikseerime $k \geq 1$ ja vaatame seeriaid $k, 2k, 3k, \ldots$ jne. Jagame seeriasse mk kuuluvad juhuslikud suurused k-sse ossa ja defineerime

$$\begin{aligned}
Z_m^{(1)} &:= Y_{1,(mk)} + \cdots + Y_{m,(mk)}, \\
Z_m^{(2)} &:= Y_{(m+1),(mk)} + \cdots + Y_{2m,(mk)}, \\
&\cdots \\
Z_m^{(k)} &:= Y_{((k-1)m+1),(mk)} + \cdots + Y_{(km),(mk)}.
\end{aligned}$$

Juhuslikud suurused $Z_m^{(1)}, \ldots, Z_m^{(k)}$ on sõltumatud ja sama jaotusega. Seetõttu iga $K > 0$ korral

$$[\mathbf{P}(Z_m^{(1)} > K)]^k = \mathbf{P}(Z_m^{(1)} > K, \ldots, Z_m^{(k)} > K) \leq \mathbf{P}(Z_m^{(1)} + \cdots + Z_m^{(k)} > kK) = \mathbf{P}(T_{(km)} > kK)$$
$$[\mathbf{P}(Z_m^{(1)} < -K)]^k = \mathbf{P}(Z_m^{(1)} < -K, \ldots, Z_m^{(k)} < -K) \leq \mathbf{P}(Z_m^{(1)} + \cdots + Z_m^{(k)} < -kK) = \mathbf{P}(T_{(km)} < -kK)$$

Et $T_{km} \Rightarrow P$, kui $m \to \infty$, on jada $\{T_{km}\}$ suhteliselt kompaktne. Seega iga $\epsilon > 0$ korral leidub $kK < \infty$ nii, et iga m korral

$$\mathbf{P}(T_{km} > kK) \leq \left(\frac{\epsilon}{2}\right)^k, \quad \mathbf{P}(T_{km} < -kK) \leq \left(\frac{\epsilon}{2}\right)^k,$$

millest ülaltoodud seoste tõttu $\mathbf{P}[|Z_m^{(1)}| > K] \leq \epsilon$ ehk $\{Z_m^{(1)}\}$ on suhteliselt kompaktne. Analoogiliselt saame, et $\{Z_m^{(i)}\}$ on suhteliselt kompaktne iga $i = 1, \ldots, k$ korral. Seega leidub alamjada m' (miks?) nii, et $Z_{m'}^{(i)} \Rightarrow P_i$ iga $i = 1, \ldots, k$ korral. Siin P_i on mingid tõenäosusmõõdud. Olgu P_i karakteristlik funktsioon $\varphi^{(i)}$.

Olgu $\varphi_m^{(i)}$ juhusliku suuruse $Z_m^{(i)}$ karakteristlik funktsioon. Seega $T_{km} = Z_m^{(1)} + \cdots + Z_m^{(k)}$ karakteristlik funktsioon avaldub

$$\varphi_{T_{km}}(t) = \prod_{i=1}^{k} \varphi_m^{(i)}(t)$$

ning koondumise $Z_{m'}^{(i)} \Rightarrow P_i$ tõttu iga t korral

$$\varphi_{T_{km'}}(t) = \prod_{i=1}^{k} \varphi_{m'}^{(i)}(t) \to \prod_{i=1}^{k} \varphi^{(i)}(t).$$

Teisest küljest, koondumisest $T_{km'} \Rightarrow P$ saame, et

$$\varphi_{T_{km'}}(t) \to \varphi_P(t), \quad \forall t,$$

kus φ_P on P karakteristlik funktsioon. Järelikult

$$\varphi_P = \prod_{i=1}^{k} \varphi^{(i)}. \tag{15.52}$$

Olgu $Z_1 \sim P_1, \ldots, Z_k \sim P_k$ sõltumatud juhuslikud suurused. Sellisel juhul $Z_1 + \cdots + Z_k$ karakteristlik funktsioon avaldub $t \mapsto \prod_{i=1}^{k} \varphi^{(i)}(t)$, millest (15.52) tõttu $Z_1 + \cdots + Z_k$ jaotus on P. ■

Selgub, et lõpmatult jagunevatel karakteristlikel funktsioonidel on kindel kuju.

Teoreem 15.11 (Levy-Hintŝini esitus) *Karakteristlik funktsioon φ on lõpmatult jagunev parajasti siis, kui*

$$\ln \varphi(t) = it\beta - \frac{t^2\sigma^2}{2} + \int_{\mathbb{R}} \Big(e^{itx} - 1 - \frac{itx}{1+x^2}\Big) \frac{1+x^2}{x^2} \mu(dx), \tag{15.53}$$

kus $\beta \in \mathbb{R}$, $\sigma^2 \geq 0$ ja μ on mingi lõplik mõõt, kusjuures $\mu\{0\} = 0$.

Veendu, et normaaljaotuse karakteristlik funktsioon on kujul (15.53), mõõt μ on siis konstantselt 0. Poissoni jaotuse $\mathrm{Po}(\lambda)$ korral on $\mu = \frac{\lambda}{2}\delta_1$.

Tsentraalsetest piirteoreemidest loe: *Billingsley* 27; *Ŝirjajev*, III § 4,5; *Williams* 18.
Poissoni teoreemist loe: *Ŝirjajev*, I § 6, III § 3, *Billingsley* 23.
Lõpmatult jagunevatest jaotustest loe: *Ŝirjajev* III § 6, *Billingsley* 28.