http://www.ttu.ee **TALLINNA TEHNIKAÜLIKOOL**
http://www.staff.ttu.ee/ math **MATEMAATIKAINSTITUUT**

http://www.staff.ttu.ee/ itammeraid **Ivar Tammeraid**

TÕENÄOSUSTEOORIA JA MATEMAATILINE STATISTIKA

Elektrooniline õppematerjal

http://www.tallinn.ee **TALLINN**
2005

Sisukord

Trükitud versioon: Ivar Tammeraid, Tõenäosusteooria ja matemaatiline statistika, TTÜ
Kirjastus,
Tallinn, 2003, 235 lk. ISBN 9985-59-366-9

Eessõna

Tõenäosusteooria on matemaatika osa, mis uurib juhuslike nähtuste üldisi seaduspärasusi sõltumatult nende nähtuste konkreetsest sisust ja annab meetodid nendele nähtustele mõjuvate juhuslike mõjude kvantitatiivseks hindamiseks. Nii looduses, tehnikas kui ka majanduses ei ole nähtusi, milles ei esineks juhuslikkuse mõju. Nähtuste kirjeldamisel tehakse vahet kahe lähenemisviisi, *deterministliku* (ettemääratusliku) ja *stohhastilise* (juhuslikkusel põhineva) vahel. Deterministliku käsitluse korral eraldatakse mõningad antud nähtust rohkem mõjutavad tegurid ja kirjeldatakse nähtust ainult nendest lähtudes, kusjuures vähem mõjutavaid tegureid ei arvestata. Juhuslikkusel põhineva käsitluse korral arvestatakse kõiki antud nähtust mõjutavaid tegureid, kusjuures vähem mõjuvate tegurite paljusus toob kaasa juhuslikkuse momendi. Juhuslikkusel põhinev lähenemine nõuab erilisi meetodeid, mida võimaldab tõenäosusteooria. *Matemaatiline statistika* on matemaatika osa, mis uurib statistiliste andmete kogumise, süstematiseerimise, töötlemise ja statistiliste järelduste tegemise meetodeid. Matemaatilise statistika eesmärgiks on statistiliste seaduspärasuste avastamine ja kirjeldamine.

Käesoleva õppevahendi aluseks on võetud Tallinna Tehnikaülikooli bakalaureuseõppe üliõpilastele peetud loengud tõenäosusteooriast ja matemaatilisest statistikast. Lisatud on mõningate väidete tõestused ja näiteülesanded, mille esitamiseks ei jätku loengul aega, kuid mis pakuvad lisavõimalusi üliõpilase iseseisvaks tööks. Õppija, keda ei huvita antud kursuse süvaõpe, võib osa keerukamatest tõestustest jätta vahele ja keskenduda näidetele. Õppevahend sobib kaugüliõpilastele. Iga peatüki lõppu on õpitud teooria kinnistamiseks lisatud harjutusülesanded, mis on varustatud vastustega.

Leidub palju tõenäosusteooria ja matemaatilise statistika täiendavaid õpikuid ja ülesannete kogusid ning statistikapakettide kirjeldusi. Pakume täiendavate õpikute valiku. Eestikeelsetest mainime õpikuid [5] , [20] , [8 − 11] , kõrgema matemaatika teatmikku [21] ja ülesannete kogu [12] . Venekeelsetest mainime õpikuid [25] , [29] , [32 − 35] , ülesannete kogusid [26] , [28] , käsiraamatut [31] ja entsüklopeediat [28] . Ingliskeelsetest mainime õpikuid [1] , [6 − 7] , [13 − 14] , [16] , [18] , [24] , ülesannete kogu [4] ja käsiraamatut [17] ning kirjandust tõenäosusteooria ja matemaatilise statistikapakettide kohta [2 − 3] , [22 − 23] . Enne võimsate statistikapakettide, nagu näiteks SAS, S, S-PLUS ja Stata, juurde asumist sobib põhjalikult tutvuda matemaatilise statistika võimalustega MS Excel keskkonnas [10] . Järgmisena tasub tutvuda statistika vabatarkvara paketiga R, mis kujutab endast paketi S-PLUS alamhulka.

Õppevahendi koostamisel on kasutatud paketti „Scientific WorkPlace 3.0", lühendatult SWP.

Tänan dotsente A. Lõhmust ja F. Vichmanni, kes abistasid autorit paljude sisuliste ja vormiliste märkustega käsikirja lõplikul viimistlemisel.

Autor

Peatükk 1

Juhuslikud sündmused

1.1 Sündmuse mõiste. Tehted sündmustega

Sündmuse mõiste on üks tõenäosusteooria põhimõiste. Lihtkäsitluses seda mõistet ei defineerita. Sündmusi tähistame tavaliselt ladina tähestiku algusosas paiknevate suurte tähtedega A, B, C, ... Vajaduse korral kasutame indekseid A_1, B_k, $C_{i,\,j}$, $D_{i,\,j,\,k}$, ... Meie käsitluse aluseks on *katse.* Katse seisneb teatud tingimuste komplekti realiseerimises. Katse käigus jälgitakse, kas teatud sündmused toimuvad või mitte. Sündmust, mis sellise katse käigus alati toimub, nimetatakse *kindlaks* sündmuseks. Sündmust, mis sellise katse käigus ei saa toimuda, nimetatakse *võimatuks* sündmuseks. Kasutame kindla sündmuse ja võimatu sündmuse tähistamiseks vastavalt tähti K ja V. Sündmust, mis sellise katse käigus võib toimuda või mitte toimuda, nimetatakse *juhuslikuks* sündmuseks.

Definitsioon 1. Sündmuse A *vastandsündmus* $\overline{A}$ on sündmus, mis katsel toimub parajasti siis, kui A ei toimu.

Definitsioon 2. Sündmusi A ja B nimetatakse *võrdseteks*, kui antud katse käigus sündmuse A toimumisega kaasneb sündmuse B toimumine ja sündmuse B toimumisega sündmuse A toimumine.

Sündmuste A ja B võrdsust tähistame $A = B$.

Definitsioon 3. Sündmuste A ning B *summaks* $A + B$ (ehk $A \cup B$) nimetatakse sündmust, mis seisneb neist vähemalt ühe toimumises (toimub kas A või B või mõlemad).

Definitsioon 4. Sündmuste A_k $(k = 1; \ldots ; n)$ summaks $\sum_{k=1}^{n} A_k$ (ehk $\bigcup_{k=1}^{n} A_k$) nimetatakse sündmust, mis seisneb neist vähemalt ühe toimumises.

Definitsioon 5. Sündmuste A ning B *korrutiseks* AB (ehk $A \cdot B$ või $A \cap B$) nimetatakse sündmust, mis seisneb neist mõlema toimumises (toimub A ja toimub B).

Definitsioon 6. Sündmuste A_k $(k = 1; \ldots ; n)$ korrutiseks $\prod_{k=1}^{n} A_k$ (ehk $\bigcap_{k=1}^{n} A_k$) nimetatakse sündmust, mis seisneb neist kõigi toimumises.

Tõestage Laused 1 ja 2 iseseisvalt.

Lause 1. Kehtivad seosed

$$A\overline{A} = V,\ A + \overline{A} = K,\ A + K = K,$$
$$A + V = A,\ AK = A,\ AV = V,\ A^2 = A.$$

Lause 2. Sündmuste korrutamine ja liitmine on kommutatiivsed, st

$$AB = BA, \quad A + B = B + A,$$

assotsiatiivsed

$$A\,(BC) = (AB)\,C, \quad A + (B + C) = (A + B) + C$$

ja distributiivne

$$A\,(B + C) = AB + AC.$$

Kehtivad *duaalsusseosed* (*Morgani seadused*)

$$\overline{A + B} = \overline{A} \cdot \overline{B},\ \overline{A \cdot B} = \overline{A} + \overline{B}. \tag{1.1.1}$$

Definitsioon 7. Sündmusi A ja B nimetatakse *teineteist välistavateks sündmusteks*, kui ühe toimumisel on teise toimumine samal katsel võimatu.

Definitsioon 8. Sündmuste süsteemi $\{A_1, A_2, \ldots, A_n\}$ nimetatakse *üksteist välistavate sündmuste süsteemiks*, kui iga kaks süsteemi kuuluvat sündmust on teineteist välistavad sündmused.

Definitsioon 9. Üksteist välistavate sündmuste süsteemi $\{A_1, A_2, \ldots, A_n\}$ nimetatakse *täielikuks* ehk *täissüsteemiks,* kui

$$A_1 + A_2 + \ldots + A_n = K.$$

Näide 1. Katse: münti visatakse üks kord. Selle katse käigus uuritakse järgmiste sündmuste toimumist: $A-$ tuleb kull; $B-$ tuleb kiri; C − tuleb kaks kulli.

Sündmused A ja B on juhuslikud sündmused. Sündmus C on võimatu sündmus. Kontrollige, et

$$\overline{A} = B,\ \overline{B} = A,\ A + B = K,\ AB = V,$$

st sündmuste süsteem $\{A, B\}$ on nii teineteist välistavate sündmuste süsteem kui ka täielik sündmuste süsteem. $\diamondsuit$

Näide 2. Katse: täringut visatakse üks kord. Selle katse käigus uuritakse järgmiste sündmuste toimumist: A_i −tuleb i silma $(1 \leq i \leq 7)$; B −tuleb paaris silmade arv; C − tuleb paaritu silmade arv. Veenduge, et

$$A_7 = V,\ BC = V,\ B + C = K,\ A_1 + A_3 + A_5 = C,\ A_2 + A_4 + A_6 = B$$

ja $\{A_1, A_3, A_5\}$ ja $\{A_2, A_4, A_6\}$ on üksteist välistavate sündmuste süsteemid ning $\{A_1, A_2, A_3, A_4, A_5, A_6\}$ on täielik süsteem. $\diamondsuit$

1.2 Sündmuse sagedus

Sooritatakse katse, mis seisneb teatud tingimuste komplekti realiseerimises. Selle katse korral uuritakse sündmuse A toimumise võimalikkust. Tavaliselt ühe katse põhjal saame kesise tulemuse sündmuse A toimumise võimalikkuse kohta. Parema hinnangu saamiseks sündmuse A toimumise võimalikkuse kohta sooritatakse veel samadel tingimustel $n-1$ katset. Eeldame, et katsed selles n-katselises seerias on sõltumatud, st ühe katse tulemus seerias ei mõjuta ülejäänud katsete tulemusi. Toimugu sündmus A selles n-katselises seerias n_A korda.

Definitsioon 1. Suurust $P^*(A) = \dfrac{n_A}{n}$ nimetatakse sündmuse A toimumise *sageduseks* ehk *suhteliseks sageduseks* selles n-katselises seerias.

Kui sündmus A on juhuslik, siis on juhuslik ka sündmuse A toimumiste arv n_A selle seeria jooksul ja on juhuslik ka sagedus $P^*(A)$.

Lause 1. Kehtivad seosed:

$$0 \le P^*(A) \le 1,\ P^*(K) = 1,\ P^*(V) = 0,\ P^*\left(\overline{A}\right) = 1 - P^*(A).$$

Tõestus. Et

$$\begin{aligned}
0 \le n_A \le n \ &\Rightarrow \ \frac{0}{n} \le \frac{n_A}{n} \le \frac{n}{n} \ \Leftrightarrow \ 0 \le P^*(A) \le 1,\\
n_K = n \ &\Rightarrow \ \frac{n_K}{n} = \frac{n}{n} = 1 \ \Leftrightarrow \ P^*(K) = 1,\\
n_V = 0 \ &\Rightarrow \ \frac{n_V}{n} = \frac{0}{n} = 0 \ \Leftrightarrow \ P^*(V) = 0,\\
n_{\overline{A}} = n - n_A \ &\Rightarrow \ \frac{n_{\overline{A}}}{n} = 1 - \frac{n_A}{n} \ \Leftrightarrow \ P^*\left(\overline{A}\right) = 1 - P^*(A),
\end{aligned}$$

siis väide on tõene. □

Vaatleme n-katselise seeria tulemusena sündmuste A, B, $A+B$, AB, $A|B$ ja $B|A$ toimumist. Sümboliga $A|B$ tähistame sündmust, mis seisneb sündmuse A toimumises eeldusel, et on toimunud sündmus B, ja sümboliga $B|A$ sündmust, mis seisneb sündmuse B toimumises eeldusel, et on toimunud sündmus A. Olgu $n_A, n_B, n_{A+B}, n_{AB}, n_{A|B}$ja $n_{B|A}$ vastavalt nende sündmuste toimumise kordade arvud selle seeria realiseerimisel.

Lause 2. Kehtivad seosed

$$P^*(A+B) = P^*(A) + P^*(B) - P^*(AB)$$

ja

$$P^*(AB) = P^*(A) \cdot P^*(B|A) = P^*(B) \cdot P^*(A|B).$$

Tõestus. Et

$$n_{A+B} = n_A + n_B - n_{AB} \quad \Rightarrow \frac{n_{A+B}}{n} = \frac{n_A}{n} + \frac{n_B}{n} - \frac{n_{AB}}{n} \Leftrightarrow$$

$$\Leftrightarrow P^*(A+B) = P^*(A) + P^*(B) - P^*(AB),$$

$$\frac{n_{AB}}{n} \overset{n_A \neq 0}{=} \frac{n_{AB}}{n_A} \cdot \frac{n_A}{n} \overset{n_B \neq 0}{=} \frac{n_{AB}}{n_B} \cdot \frac{n_B}{n} \Rightarrow$$

$$\Rightarrow \quad P^*(AB) = P^*(A) \cdot P^*(B|A) = P^*(B) \cdot P^*(A|B),$$

siis väide on tõene. □

Näide 1. Münti visati sada korda, kusjuures viiekümne seitsmel korral tuli kiri. Leiame kirja saamise sageduse selle seeria korral.

Katseks on mündi viskamine ja sündmuseks A kirja tulek sel katsel. Vastavalt definitsioonile saame $P^*(A) = 57/100 = 0.57$. ♢

Osutub, et teatud tingimustel on katsete arvu n suurendamisel sündmuse sagedusel tendents läheneda mingile kindlale arvule.

1.3 Tõenäosuse statistiline definitsioon

Definitsioon 1. Juhusliku sündmuse A *statistiliseks tõenäosuseks* nimetatakse arvu $P(A)$, millele selle sündmuse toimumise sagedusel $P^*(A)$ on tendents läheneda, kui samadel tingimustel sooritatud sõltumatute katsete arv n läheneb lõpmatusele.

Sündmus $P^*(A) \overset{n\to\infty}{\to} P(A)$ on juhuslik. Juhusliku sündmuse A statistilisel tõenäosusel $P(A)$ on tänu definitsioonile paljud omadused sarnased sündmuse toimumise sageduse $P^*(A)$ omadustega.

Lause 1. Kehtivad väited

$$0 \leq P(A) \leq 1, \quad P(K) = 1, \quad P(V) = 0, \quad P\left(\overline{A}\right) = 1 - P\left(A\right),$$

$$P(A+B) = P(A) + P(B) - P(AB),$$

$$P(AB) = P(A) \cdot P(B|A) = P(B) \cdot P(A|B).$$

Tõestus. Et

$$0 \leq P^*(A) \leq 1, \; P^*(K) = 1, \; P^*(V) = 0, \; P^*\left(\overline{A}\right) = 1 - P^*\left(A\right),$$

siis Definitsiooni 1 abil saame

$$0 \leq P(A) \leq 1, \quad P(K) = 1, \quad P(V) = 0, \quad P\left(\overline{A}\right) = 1 - P\left(A\right).$$

Kuna

$$P^*(A) = \frac{n_A}{n} \overset{n\to\infty}{\to} P(A) \ \wedge \ P^*(B) = \frac{n_B}{n} \overset{n\to\infty}{\to} P(B) \ \wedge$$
$$\wedge \ P^*(AB) = \frac{n_{AB}}{n} \overset{n\to\infty}{\to} P(AB) \ \wedge \ P^*(B|A) = \frac{n_{AB}}{n_A} \overset{n\to\infty}{\to} P(B|A) \ \Rightarrow$$
$$P^*(A+B) = \frac{n_{A+B}}{n} = \frac{n_A}{n} + \frac{n_B}{n} - \frac{n_{AB}}{n} \overset{n\to\infty}{\to} P(A) + P(B) - P(AB) \wedge$$
$$\wedge \ P^*(AB) = \frac{n_{AB}}{n_A} \cdot \frac{n_A}{n} = P^*(A) \cdot P^*(B|A) \overset{n\to\infty}{\to} P(A) \cdot P(B|A),$$

siis on tõene ka väite ülejäänud osa. □

1.4 Geomeetriline tõenäosus

Oletame, et katse käigus valitakse üks punkt hulgast $\Omega \subset \mathbf{R}^n$. Olgu Ω_A hulga Ω alamhulk, st $\Omega_A \subset \Omega$. Oletame, et oskame hulki Ω ja Ω_A mõõta, kusjuures $n = 1$ korral on selleks mõõduks μ pikkus, $n = 2$ korral on selleks mõõduks μ pindala, $n = 3$ korral ruumala V , jne. Olgu A sündmus, et katse käigus valitakse hulga Ω_A punkt. Eeldame, et punkti sattumise võimalikkus hulga Ω mingisse alamhulka sõltub vaid selle alamhulga mõõdust.

Definitsioon 1. Sündmuse A *geomeetriliseks tõenäosuseks* nimetatakse suurust

$$P(A) = \frac{\mu(\Omega_A)}{\mu(\Omega)}.$$

Tõestage, et Lause 1.3.1 kehtib ka geomeetrilise tõenäosuse korral.

Näide 1. Valitakse üks arv lõigust $[-1;3]$. Leiame tõenäosuse, et see arv: 1) on 2.3; 2) on suurem kui 0.5; 3) on väiksem kui 0.

Et $\Omega = [-1;3]$ on ruumi R^1 alamhulk, siis mõõduks on pikkus, kusjuures $\mu(\Omega) = 4$. Tähistame alaülesannetes esinevad sündmused vastavalt tähtedega A, B ja C. Olgu Ω_A, Ω_B ja Ω_C neile vastavad alamhulgad. Seega

$$\Omega_A = \{2.3\} \ \wedge \ \mu(\Omega_A) = 0 \quad \Rightarrow \ P(A) = \frac{\mu(\Omega_A)}{\mu(\Omega)} = \frac{0}{4} = 0,$$
$$\Omega_B = (0.5;3] \ \wedge \ \mu(\Omega_B) = 2.5 \quad \Rightarrow \ P(B) = \frac{\mu(\Omega_B)}{\mu(\Omega)} = \frac{2.5}{4} = 0.625,$$
$$\Omega_C = [-1;0) \ \wedge \ \mu(\Omega_C) = 1 \quad \Rightarrow \ P(C) = \frac{\mu(\Omega_C)}{\mu(\Omega)} = \frac{1}{4} = 0.25. \qquad \diamondsuit$$

Näitest 1 saame huvitava tähelepaneku, ka võimaliku sündmuse tõenäosus võib olla 0. Seega

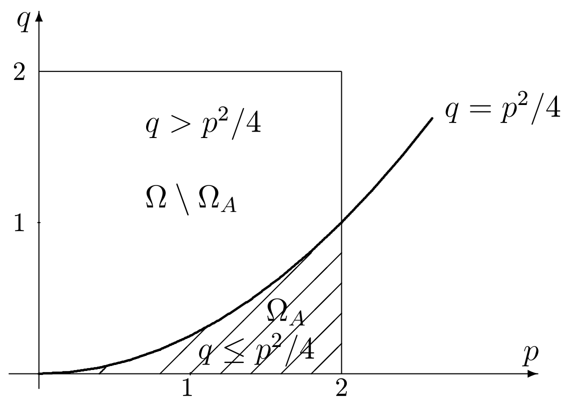
$$A = V \ \Rightarrow \ P(A) = 0, \quad P(A) = 0 \ \not\Rightarrow \ A = V.$$

Näide 2. Ruutvõrrandi $x^2 + px + q = 0$ kordajad p ja q kuuluvad lõiku $[0;2]$. Leiame tõenäosuse, et selle ruutvõrrandi lahendid on reaalsed.

Olgu A sündmus, et selle ruutvõrrandi lahendid on reaalsed. Et

$$x^2 + px + q = 0 \;\Rightarrow\; x_{1,2} = -p/2 \pm \sqrt{p^2/4 - q},$$

siis ruutvõrrandi lahendid on reaalsed parajasti siis, kui $p^2/4 - q \geq 0$, st $q \leq p^2/4$. Kuna $p, q \in [0; 2]$, siis Ω on ruut $[0; 2] \times [0; 2]$ pindalaga $\mu(\Omega) = 4$. Sündmus A toimub, kui selle ruudu punkt (p, q) on allpool parabooli $q = p^2/4$ (joonisel viirutatud osas), st $(p, q) \in \Omega_A$



Seega

$$\mu(\Omega_A) = \int_0^2 \frac{p^2}{4} dp = \left.\frac{p^3}{12}\right|_0^2 = \frac{2}{3}, \quad P(A) = \frac{\mu(\Omega_A)}{\mu(\Omega)} = \frac{\frac{2}{3}}{4} = \frac{1}{6}. \qquad \diamondsuit$$

1.5 Klassikaline tõenäosuse definitsioon

Definitsioon 1. Sündmuste süsteemi $\{A_1, A_2, \ldots, A_n\}$ nimetatakse *täielikuks*, kui:

$$1^\circ \; \sum\nolimits_{i=1}^{n} A_i = K; \quad 2^\circ \; A_i A_j = V \quad (i \neq j). \tag{1.5.1}$$

Definitsioon 2. Sündmuste süsteemi $\{A_1, A_2, \ldots, A_n\}$ nimetatakse *elementaarsündmuste süsteemiks,* kui:
1° see süsteem on täielik;
2° kõik selle süsteemi sündmused on *võrdvõimalikud*, st toimumise suhtes võrdväärsed.

Oletame, et meid huvitab katse tulemusena sündmuse A toimumine ja meil on võimalik selle katse jaoks leida selline elementaarsündmuste süsteem $\{A_1, A_2, \ldots, A_n\}$, mille korral

$$A = \sum\nolimits_{k=1}^{m} A_{i_k} \quad (m \leq n), \tag{1.5.2}$$

kusjuures süsteem $\{A_{i_1}, A_{i_2}, \ldots, A_{i_n}\}$ koosneb samadest sündmustest, mis $\{A_1, A_2, \ldots, A_n\}$. Iga sündmuse $A_{i_1}, A_{i_2}, \ldots, A_{i_m}$ toimumisega kaasnegu sündmuse A toimumine, kusjuures neid sündmusi A_{i_k} $(k = 1; \ldots; m)$ nimetatakse sündmuse A toimumiseks *soodsateks* elementaarsündmusteks.

Definitsioon 2. Suurust

$$P(A) = \frac{m}{n} \tag{1.5.3}$$

nimetatakse sündmuse A *klassikaliseks tõenäosuseks.*

Lause 1. Kehtivad väited

$$\begin{aligned} P(K) = 1, \quad P(V) = 0, \quad 0 \leq P(A) \leq 1, \quad P\left(\overline{A}\right) = 1 - P\left(A\right), \\ P(A+B) = P(A) + P(B) - P(AB), \quad P(AB) = P(A) \cdot P(B|A) = \\ = P(B) \cdot P(A|B). \end{aligned}$$

Tõestus. Olgu $\{A_1, A_2, \ldots, A_n\}$ elementaarsündmuste süsteem, mis sobib sündmuste A, $\overline{A}$, B, $A+B$ ja AB esitamiseks ning m_A, $m_{\overline{A}}$, m_B, m_{A+B} ja m_{AB} neile vastavate soodsate elementaarsündmuste arvud. Et

$$\begin{aligned} m_K = n \quad &\Rightarrow \quad \frac{m_K}{n} = \frac{n}{n} = 1 \quad \Leftrightarrow \quad P(K) = 1, \\ m_V = 0 \quad &\Rightarrow \quad \frac{m_V}{n} = \frac{0}{n} = 0 \quad \Leftrightarrow \quad P(V) = 0, \\ 0 \leq m_A \leq n \quad &\Rightarrow \quad \frac{0}{n} \leq \frac{m_A}{n} \leq \frac{n}{n} \quad \Leftrightarrow \quad 0 \leq P(A) \leq 1, \\ m_A + m_{\overline{A}} = n \quad &\Rightarrow \quad \frac{m_A}{n} = 1 - \frac{m_{\overline{A}}}{n} \quad \Leftrightarrow \quad P\left(\overline{A}\right) = 1 - P\left(A\right), \\ m_{A+B} = m_A + m_B - m_{AB} \quad &\Rightarrow \quad \frac{m_{A+B}}{n} = \frac{m_A}{n} + \frac{m_B}{n} - \frac{m_{AB}}{n} \Leftrightarrow \\ &\Leftrightarrow P(A+B) = P(A) + P(B) - P(AB), \end{aligned}$$

$$\begin{aligned} \frac{m_{AB}}{n} = \frac{m_{AB}}{n_A} \cdot \frac{n_A}{n} = \frac{m_{AB}}{n_B} \cdot \frac{n_B}{n} \quad \Leftrightarrow \\ \Leftrightarrow \quad P(AB) = P(A) \cdot P(B|A) = P(B) \cdot P(A|B), \end{aligned}$$

siis Lause 1 väide on tõene. $\square$

Näide 1. Täringut visatakse kaks korda. Leiame tõenäosuse, et mõlemal korral tuleb sama silmade arv.

Olgu A sündmus, et mõlemal korral tuleb sama silmade arv. Vaatleme kaht lahendust.

I Kasutame sündmuste süsteemi $\{A_{i,j}\}$ $(i = 1; \ldots; 6\ ,\ j = 1; \ldots; 6)$, kus $A_{i,j}$ on sündmus, et esimesel viskel tuleb i silma ja teisel viskel j silma. Kuna

$$\sum_{i=1}^{6} \sum_{j=1}^{6} A_{i,j} = K, \quad A_{i,j}\, A_{k,l} = V, \quad (i \neq k \vee j \neq l)$$

ja selle süsteemi sündmused on võrdvõimalikud, siis on tegemist elementaarsündmuste süsteemiga, milles on 36 elementaarsündmust, st $n = 36$. Et $A = A_{1,1} + A_{2,2} + A_{3,3} + A_{4,4} + A_{5,5} + A_{6,6}$, siis sündmuse A toimumiseks soodsate elementaarsündmuste arv $m = 6$ ja klassikalise tõenäosuse definitsiooni põhjal $P(A) = \dfrac{6}{36} = \dfrac{1}{6}$.

II Vaatleme sündmuste süsteemi $\{B_i\}$ $(i = 1; \ldots; 6)$, kus B_i on sündmus, et teisel viskel on silmede arv i. Veenduge, et süsteem $\{B_i\}$ on elementaarsündmuste süsteem. Samas on ka see süsteem sobiv sündmuse A kirjeldamiseks, $A = B_{i_k}$, st teisel viskel tuleb sama silmade arv mis esimesel viskel ja sündmuse A toimumiseks on soodne vaid üks süsteemi $\{B_i\}$ sündmus. Seega klassikalise tõenäosuse definitsiooni põhjal saame $P(A) = \dfrac{1}{6}$. $\diamondsuit$

Näide 2. Kaardipakist, milles on 52 kaarti, võetakse (huupi) üks kaart. Leiame järgmiste sündmuste tõenäosused: 1) A– võetu on äss; 2) B– võetud kaart on musta masti; 3) C– võetud kaart on pilt (soldat, emand, kuningas, äss).

Koostame sündmuste süsteemi $\{A_i\}_{(i=1;2;\ldots;52)}$, mille elementideks on iga konkreetse kaardi võtmine: A_1– võetud kaart on risti 2; $\ldots$; A_{52}– võetud kaart on poti äss (kasutame järjestust risti 2 ... risti äss $\rightarrow$ ruutu 2 ... ruutu äss $\rightarrow$ ärtu kaks ... ärtu äss $\rightarrow$ poti 2 ... poti äss). Kuna

$$\sum\nolimits_{i=1}^{52} A_i = K, \quad A_i A_j = V \quad (i \neq j),$$

siis tegemist on täieliku süsteemiga. Iga kaardi võtmine pakist on võrdvõimalik. Seega süsteem $\{A_i\}$ on elementaarsündmuste süsteem. Sündmuse A toimumiseks on neli soodsat elementaarsündmust: $A = A_{13} + A_{26} + A_{39} + A_{52}$. Seega klassikalise tõenäosuse definitsiooni põhjal

$$P(A) = \frac{4}{52} = \frac{1}{13}.$$

Sündmuse B toimumiseks on 26 soodsat elementaarsündmust:

$$B = \sum_{i=1}^{13} A_i + \sum_{i=40}^{52} A_i.$$

Seega

$$P(B) = \frac{26}{52} = \frac{1}{2}.$$

Sündmuse C toimumiseks on 16 soodsat elementaarsündmust:

$$B = \sum_{i=10}^{13} A_i + \sum_{i=23}^{26} A_i + \sum_{i=36}^{39} A_i + \sum_{i=49}^{52} A_i.$$

Seega

$$P(C) = \frac{16}{52} = \frac{4}{13}. \qquad \diamondsuit$$

Näide 3. Õpperühmas on 16 tudengit, neist 6 neidu ja 10 noormeest. Laboratoorseks tööks jaotatakse õpperühm huupi kaheks grupiks, mõlemas 8 tudengit. Leiame tõenäosuse, et 1) kõik 6 neidu on ühes grupis, 2) ühes grupis on täpselt 5 neidu, 3) ühes grupis on täpselt 4 neidu, 4) mõlemas grupis on 3 neidu.

Olgu A sündmus, et kõik 6 neidu on ühes grupis, B– ühes grupis on täpselt 5 neidu, C– ühes grupis on täpselt 4 neidu, D– mõlemas grupis on 3 neidu.

Lähtume tõsiasjast, et kui me neist 16-st tudengist 8 välja valime, siis grupid on sellega määratud. Piirdume järgnevas kaheksa tudengi väljavalimisega. Seega on katseks kaheksa tudengi väljavalimine kuueteistkümnest. Kuna ei ole oluline, mis järjekorras need valitud tudengid selles grupis on, siis need kaheksa tudengit moodustavad kombinatsiooni kuueteistkümnest tudengist kaheksa kaupa. Seda gruppi saab moodustada $C_{16}^8 = \binom{16}{8} = 12870$ erineval viisil. Käsitleme iga sellise kombinatsiooni valikut kui abisündmust sündmuste A, B, C ja D kirjeldamiseks. Moodustame kõigist sellistest abisündmustest süsteemi. Tegemist on elementaarsündmuste süsteemiga. Tõesti, täpselt üks neist abisündmustest leiab katse käigus aset, st kõigi nende abisündmuste summa on kindel sündmus ja need abisündmused on üksteist välistavad. Abisündmused on võrdvõimalikud. Sündmuse A toimumiseks on soodsad need elementaarsündmused, st need kombinatsioonid, milles on kas 6 neidu või ainult noormehed. Kombinatsiooni kuueteistkümnest tudengist kaheksa kaupa, milles on 6 neidu, saab moodustada nii, et kuuest neiust valime välja kõik kuus neidu (C_6^6 erinevat võimalust) ja kaheksase grupi saamiseks lisame 2 noormeest (C_{10}^2 erinevat võimalust). Iga neidude kuuikuga sobib suvaline noormeeste kombinatsioon kümnest kahe kaupa. Seega on $C_6^6 \cdot C_{10}^2$ sündmuse A toimumiseks soodsate elementaarsündmuste, millega kaasneb kuue neiu valik, arv. Sündmuse A toimumiseks on soodsad ka need elementaarsündmused, millega kaasneb kaheksa noormehe suvaline valik. Selliseid sündmuse A toimumiseks soodsaid elementaarsündmusi on $C_6^0 \cdot C_{10}^8$. Tõenäosuse klassikalise definitsiooni kohaselt

$$P(A) = \frac{C_6^6 \cdot C_{10}^2 + C_6^0 \cdot C_{10}^8}{C_{16}^8} = \left[\begin{array}{c} C_n^k = C_n^{n-k} \Rightarrow C_{10}^8 = C_{10}^2 \\ C_n^0 = 1, \quad C_n^n = 1 \end{array} \right] =$$

$$= \frac{1 \cdot \dfrac{10 \cdot 9}{1 \cdot 2} + 1 \cdot \dfrac{10 \cdot 9}{1 \cdot 2}}{\dfrac{16 \cdot 15 \cdot 14 \cdot 13 \cdot 12 \cdot 11 \cdot 10 \cdot 9}{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6 \cdot 7 \cdot 8}} = \frac{1}{143}.$$

Analoogiliste arutelude abil saame

$$P(B) = \frac{C_6^5 \cdot C_{10}^3 + C_6^1 \cdot C_{10}^7}{C_{16}^8} = \frac{2 \cdot C_6^1 \cdot C_{10}^3}{C_{16}^8} =$$

$$= \frac{2 \cdot \dfrac{6}{1} \cdot \dfrac{10 \cdot 9 \cdot 8}{1 \cdot 2 \cdot 3}}{\dfrac{16 \cdot 15 \cdot 14 \cdot 13 \cdot 12 \cdot 11 \cdot 10 \cdot 9}{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6 \cdot 7 \cdot 8}} = \frac{16}{143},$$

$$P(C) = \frac{C_6^4 \cdot C_{10}^4 + C_6^2 \cdot C_{10}^6}{C_{16}^8} = \frac{2 \cdot C_6^2 \cdot C_{10}^4}{C_{16}^8} =$$

$$= \frac{2 \cdot \dfrac{6 \cdot 5}{1 \cdot 2} \cdot \dfrac{10 \cdot 9 \cdot 8 \cdot 7}{1 \cdot 2 \cdot 3 \cdot 4}}{\dfrac{16 \cdot 15 \cdot 14 \cdot 13 \cdot 12 \cdot 11 \cdot 10 \cdot 9}{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6 \cdot 7 \cdot 8}} = \frac{70}{143}$$

ja

$$P(D)=\frac{C_6^3\cdot C_{10}^5}{C_{16}^8}=\frac{\dfrac{6\cdot5\cdot4}{1\cdot2\cdot3}\cdot\dfrac{10\cdot9\cdot8\cdot7\cdot6}{1\cdot2\cdot3\cdot4\cdot5}}{\dfrac{16\cdot15\cdot14\cdot13\cdot12\cdot11\cdot10\cdot9}{1\cdot2\cdot3\cdot4\cdot5\cdot6\cdot7\cdot8}}=\frac{56}{143}.$$

Kuna sündmused A, B, C ja D on üksteist välistavad ja nende summa on kindel sündmus, siis nende tõenäosuste summa peab tulema üks:

$$P(A)+P(B)+P(C)+P(D)=\frac{1}{143}+\frac{16}{143}+\frac{70}{143}+\frac{56}{143}=1. \qquad \diamondsuit$$

Näide 4. Bridžimängus (kaardipakis 52 kaarti) jagatakse igale mängijale 13 kaarti. Leiame tõenäosuse, et mängija saab (täpselt) k $(k=0;1;2;3;4)$ ässa.

Olgu katseks kolmeteist kaardi juhuslik võtmine viiekümne kahest kaardist. Vaatleme $k=0;1;2;3;4$ korral sündmusi A_k– mängija saab täpselt k ässa. Olgu abisündmuseks suvalise kolmeteist kaardi võtmine sellest pakist. Ilmselt (selles kontekstis) ei ole kaartide järjekord nende kolmeteistkümne hulgas oluline. Seega võime neid kolmeteist kaarti käsitleda kui kombinatsiooni viiekümnekahest elemendist kolmeteistkümne kaupa. Katse käigus täpselt üks neist abisündmustest leiab aset. Seega on need abisündmused üksteist välistavad ja nende kõigi summa on kindel sündmus. Lisaks on need abisündmused võrdvõimalikud. Seega võime selliste abisündmuste hulka vaadelda kui elementaarsündmuste süsteemi. Selles süsteemis on C_{52}^{13} elementaarsündmust. Sündmuse A_k toimumiseks on selles süsteemis soodsaid elementaarsündmusi $C_4^k\cdot C_{48}^{13-k}$. Tõesti selles kombinatsioonis on k ässa ja $13-k$ mitteässa. Selliseks ässade valikuks on C_4^k erinevat võimalust ja mitteässade valikuks C_{48}^{13-k} erinevat võimalust. Kuna iga sellise k ässa valikuga sobib suvaline $13-k$ mitteässa valik, siis $C_4^k\cdot C_{48}^{13-k}$ on erinevate soodsate elementaarsündmuste arv. Seega tõenäosuse klassikalise definitsiooni kohaselt

$$P(A_k)=\frac{C_4^k\cdot C_{48}^{13-k}}{C_{52}^{13}}=\frac{\binom{4}{k}\cdot\binom{48}{13-k}}{\binom{52}{13}} \qquad (k=0;1;2;3;4)$$

ja

$$P(A_0)=\frac{C_4^0\cdot C_{48}^{13}}{C_{52}^{13}}=\frac{\binom{4}{0}\cdot\binom{48}{13}}{\binom{52}{13}}=\ldots=\frac{6327}{20\,825}\approx 0.303\,8,$$

$$P(A_1)=\frac{C_4^1\cdot C_{48}^{12}}{C_{52}^{13}}=\frac{\binom{4}{1}\cdot\binom{48}{12}}{\binom{52}{13}}=\ldots=\frac{9139}{20\,825}\approx 0.438\,8,$$

$$P(A_2)=\frac{C_4^2\cdot C_{48}^{11}}{C_{52}^{13}}=\frac{\binom{4}{2}\cdot\binom{48}{11}}{\binom{52}{13}}=\ldots=\frac{4446}{20\,825}\approx 0.213\,5,$$

$$P(A_3)=\frac{C_4^3\cdot C_{48}^{10}}{C_{52}^{13}}=\frac{\binom{4}{3}\cdot\binom{48}{10}}{\binom{52}{13}}=\ldots=\frac{858}{20\,825}\approx 0.0\,412$$

ning

$$P(A_4) = \frac{C_4^4 \cdot C_{48}^9}{C_{52}^{13}} = \ldots = \frac{\binom{4}{4} \cdot \binom{48}{9}}{\binom{52}{13}} \approx \frac{11}{4165} \approx 0.0026\,.$$

Sündmused A_k $(k = 0; 1; 2; 3; 4)$ on üksteist välistavad ja nende summa on kindel sündmus. Saame kontrollida, kas nende tõenäosuste summa on 1 :

$$\sum_{k=0}^{4} P(A_k) = \frac{6327}{20\,825} + \frac{9139}{20\,825} + \frac{4446}{20\,825} + \frac{858}{20\,825} + \frac{11}{4165} = 1. \qquad \diamondsuit$$

1.6 Tõenäosusteooria aksioomid

Võrreldes eelneva käsitlusega on võimalik tõenäosuse mõistet defineerida rangemalt. Olgu Ω mingi hulk, mille elemente ω me nimetame *elementaarsündmusteks.* Olgu S hulga Ω mingi alamhulkade hulk. Hulga S elemente nimetame *juhuslikeks sündmusteks* ja hulka Ω *elementaarsündmuste ruumiks.*

Definitsioon 1. Hulka S nimetatakse hulga Ω *hulkade algebraks*, kui
1° $\Omega \in S$,
2° $A \in S \ \wedge\ B \in S \ \Rightarrow\ A \cup B \in S \ \wedge\ A \cap B \in S \ \wedge\ A \backslash B \in S.$

Vaatleme hulga Ω alamhulkade hulga S korral järgmisi *aksioome.*

I aksioom. S on hulga Ω hulkade algebra.

II aksioom. Igale hulgale $A \in S$ on vastavusse seatud mittenegatiivne reaalarv $P(A)$.

III aksioom. $P(\Omega) = 1$.

IV aksioom. $A \cap B = \emptyset \ \Rightarrow\ P(A \cup B) = P(A) + P(B)$.

Definitsioon 2. Kui elementaarsündmuste ruum Ω ja sellel antud alamhulkade hulk S rahuldavad aksioome I-IV, siis öeldakse, et on antud *tõenäosusruum* (Ω, S, P).

Suurust $P(A)$ nimetatakse juhusliku *sündmuse* A *tõenäosuseks.*

Sellel viisil esitatud seoste ühitamiseks eelnevatega tuleb $\overline{A} \overset{\text{def.}}{=} \Omega \backslash A$, $\cup \longleftrightarrow +$, $\cap \longleftrightarrow \cdot$, $\emptyset \longleftrightarrow V$.

Järeldus 1. Definitsioonist 1 järelduvad seosed

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \tag{1.6.1}$$

ja

$$P\left(\overline{A}\right) = 1 - P(A)\,, \tag{1.6.2}$$

Tõestus. Kuna

$$A \cup B = A \cup (B \backslash A) \ \wedge \ A \cap (B \backslash A) = \emptyset \ \Rightarrow$$
$$P(A \cup B) = P(A \cup (B \backslash A)) \overset{\mathbf{IV}}{=} P(A) + P(B \backslash A),$$
$$B = (B \backslash A) \cup (A \cap B) \ \wedge \ (B \backslash A) \cap (A \cap B) = \emptyset \overset{\mathbf{IV}}{\Rightarrow}$$
$$\Rightarrow \ P(B) = P(B \backslash A) + P(A \cap B) \ \Rightarrow$$
$$\Rightarrow \ P(B \backslash A) = P(B) - P(A \cap B),$$

siis

$$P(A \cup B) = P(A) + P(B \backslash A) = P(A) + P(B) - P(A \cap B).$$

Lisaks

$$A \cup \overline{A} = \Omega \ \wedge \ A \cap \overline{A} = \emptyset \overset{\mathbf{III, IV}}{\Rightarrow} P\left(A \cup \overline{A}\right) = P(A) + P\left(\overline{A}\right) = 1 \ \Rightarrow$$
$$\Rightarrow \ P\left(\overline{A}\right) = 1 - P(A). \qquad \square$$

Definitsioon 3. Kui $P(A) > 0$, siis tõenäosust $P(B|A) \overset{\text{def.}}{=} P(AB)/P(A)$ nimetatakse sündmuse B *tinglikuks tõenäosuseks* tingimusel A.

Näide 1. Olgu Ω üheelemendiline hulk, $\Omega = \{\omega\}$. Olgu $S = \{\Omega, \emptyset\}$ ja $P(\Omega) = 1$ ning $P(\emptyset) = 0$. Kontrollime, kas sel viisil saame tõenäosusruumi.

Esiteks kontrollime, kas selline S on hulga Ω hulkade algebra. Kuna $\Omega \in S$, siis tingimus 1° on täidetud. Et 1) $A = \Omega$ ja $B = \emptyset$ korral $\Omega \cup \emptyset = \Omega \in S \wedge \Omega \cap \emptyset = \emptyset \in S \ \wedge \ \Omega \backslash \emptyset = \Omega \in S$, 2) $A = \emptyset$ ja $B = \Omega$ korral $\Omega \cup \emptyset = \Omega \in S \wedge \Omega \cap \emptyset = \emptyset \in S \ \wedge \ \emptyset \backslash \Omega = \emptyset \in S$, 3) $A = \emptyset$ ja $B = \emptyset$ korral $\emptyset \cup \emptyset = \emptyset \in S \wedge \emptyset \cap \emptyset = \emptyset \ \wedge \ \emptyset \backslash \emptyset = \emptyset \in S$, 4) $A = \Omega$ ja $B = \Omega$ korral $\Omega \cup \Omega = \Omega \in S \wedge \Omega \cap \Omega = \Omega \in S \ \wedge \ \Omega \backslash \Omega = \emptyset \in S$, siis ka tingimus 2° on täidetud. Seega S on hulga Ω hulkade algebra, st aksioom I on täidetud. Teiseks, igale hulga S elemendile on vastavusse seatud mittenegatiivne reaalarv $P(A)$, st aksioom II on täidetud. Kolmandaks, tõesti $P(\Omega) = 1$, st aksioom III on täidetud. Neljandaks kontrollime, kas $A \cap B = \emptyset \ \Rightarrow \ P(A \cup B) = P(A) + P(B)$. Et 1) $A = \Omega$ ja $B = \emptyset$ korral $\Omega \cap \emptyset = \emptyset$ ja $P(\Omega \cup \emptyset) = P(\Omega) = P(\Omega) + P(\emptyset)$, 2) $A = \emptyset$ ja $B = \emptyset$ korral $\emptyset \cap \emptyset = \emptyset$ ja $P(\emptyset \cup \emptyset) = P(\emptyset) = P(\emptyset) + P(\emptyset)$, siis aksioom IV on täidetud. Seega saame tõenäosusruumi (Ω, S, P). Kuidas saadud tulemust tõlgendada? $\diamondsuit$

Näide 2. Olgu $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$. Olgu S kõigi hulga Ω osahulkade hulk ja $P(\omega_i) = p_i \geq 0 \ (i = 1; \ldots; n)$, kusjuures $\sum_{i=1}^{n} p_i = 1$. Iga hulga Ω osahulk A on esitatav kujul $A = \{\omega_{i_1}, \omega_{i_2}, \ldots, \omega_{i_m}\} \ (0 \leq m \leq n)$. Defineerime $P(A) = \sum_{j=1}^{m} p_{i_j}$. Kontrollige, kas nii saame tõenäosusruumi. Näidake, et $p_i = 1/n \ (i = 1; \ldots; n)$ korral saame erijuhuna klassikalise tõenäosuse definitsiooni. $\diamondsuit$

Märkus 1. Kui hulga Ω alamhulkade hulk S on loenduv, siis on otstarbekas aksioom IV asendada aksioomiga **IV'**:

$$A_i A_j = \varnothing \quad (i, j \in \mathbf{N}, i \neq j) \quad \Rightarrow \quad P\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} P(A_k). \qquad (1.6.3)$$

1.7 Tõenäosuste liitmis- ja korrutamislause

Järelduse 1.6.1 põhjal

$$P(A+B) = P(A) + P(B) - P(AB). \tag{1.7.1}$$

Kui sündmuste summas on kolm liidetavat, siis

$$\begin{aligned} P(A_1 + A_2 + A_3) &= \left[\begin{array}{c} \text{sündmuste liitmine} \\ \text{on assotsiatiivne} \end{array} \right] = P(A_1 + (A_2 + A_3)) = \\ &= \left[\begin{array}{c} \text{rakendame seost (1.7.1)} \\ A = A_1,\ B = A_2 + A_3 \end{array} \right] = \end{aligned}$$

$$\begin{aligned} &= P(A_1) + P(A_2 + A_3) - P(A_1(A_2 + A_3)) \overset{\text{distributiivsus}}{=} \\ &= P(A_1) + P(A_2 + A_3) - P(A_1A_2 + A_1A_3) \overset{(1.7.1)}{=} \\ &= P(A_1) + P(A_2) + P(A_3) - P(A_2A_3) - P(A_1A_2) - \\ &- P(A_1A_3) + P(A_1A_2A_1A_3) \overset{A_1 \cdot A_1 = A_1^2 = A_1}{=} \\ &= P(A_1) + P(A_2) + P(A_3) - P(A_2A_3) - \\ &- P(A_1A_2) - P(A_1A_3) + P(A_1A_2A_3) \end{aligned}$$

ehk

$$P\left(\sum_{i=1}^{3} A_i\right) = \sum_{i=1}^{3} P(A_i) - \sum_{i=1}^{2}\sum_{j>i}^{3} P(A_iA_j) + (-1)^{3+1} P\left(\prod_{i=1}^{3} A_i\right).$$

Lause 1. Kehtib väide

$$P\left(\sum_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i) - \sum_{i=1}^{n-1}\sum_{j>i}^{n} P(A_iA_j) + \ldots + (-1)^{n+1} P\left(\prod_{i=1}^{n} A_i\right). \tag{1.7.2}$$

Tõestus. Kasutame matemaatilise induktsiooni meetodit. Eelneva põhjal on baas olemas, $n = 2$ ja $n = 3$ korral väide kehtib. Induktsioonisammu lubatavuse tõestamiseks eeldame, et väide on tõene $n - 1$ korral, st

$$\begin{aligned} P\left(\sum_{i=1}^{n-1} A_i\right) &= \sum_{i=1}^{n-1} P(A_i) - \sum_{i=1}^{n-2}\sum_{j>i}^{n-1} P(A_iA_j) + \sum_{i=1}^{n-3}\sum_{j>i}^{n-2}\sum_{k>j}^{n-1} P(A_iA_jA_k) + \\ &+ \ldots + (-1)^{n} P\left(\prod_{i=1}^{n-1} A_i\right). \end{aligned} \tag{1.7.3}$$

Sel juhul saame

$$\begin{aligned}
P\left(\sum_{i=1}^{n} A_i\right) &= P\left(\left(\sum_{i=1}^{n-1} A_i\right) + A_n\right) \overset{(1.7.1)}{=} \\
&= P\left(\sum_{i=1}^{n-1} A_i\right) + P(A_n) - P\left(\left(\sum_{i=1}^{n-1} A_i\right) A_n\right) \overset{\text{distributiivsus}}{=} \\
&= P\left(\sum_{i=1}^{n-1} A_i\right) + P(A_n) - P\left(\sum_{i=1}^{n-1} A_i A_n\right) \overset{(1.7.3)}{=}
\end{aligned}$$

$$\begin{aligned}
&= \sum_{i=1}^{n-1} P(A_i) - \sum_{i=1}^{n-2}\sum_{j>i}^{n-1} P(A_iA_j) + \sum_{i=1}^{n-3}\sum_{j>i}^{n-2}\sum_{k>j}^{n-1} P(A_iA_jA_k) + \\
&+ \ldots + (-1)^n P\left(\prod_{i=1}^{n-1} A_i\right) + P(A_n) - \sum_{i=1}^{n-1} P(A_iA_n) + \\
&+ \sum_{i=1}^{n-2}\sum_{j>i}^{n-1} P(A_iA_nA_jA_n) - \sum_{i=1}^{n-3}\sum_{j>i}^{n-2}\sum_{k>j}^{n-1} P(A_iA_nA_jA_nA_kA_n) \\
&+ \ldots - (-1)^n P\left(\prod_{i=1}^{n-1} (A_iA_n)\right) \overset{A_n^m = A_n\ (m\geq 1)}{=} [\text{miks?}] = \\
&= \sum_{i=1}^{n} P(A_i) - \sum_{i=1}^{n-1}\sum_{j>i}^{n} P(A_iA_j) + \sum_{i=1}^{n-2}\sum_{j>i}^{n-1}\sum_{k>j}^{n} P(A_iA_jA_k) + \\
&+ \ldots + (-1)^{n+1} P\left(\prod_{i=1}^{n} A_i\right),
\end{aligned}$$

st kui Lause 1 väide on tõene $n-1$ korral, siis see väide on tõene ka n korral. Seega on induktsioonisamm lubatud ja Lause 1 väide on tõestatud matemaatilise induktsiooni meetodil. $\square$

Märkus 1. Kuna $\overline{\sum_{i=1}^{n} A_i} = \prod_{i=1}^{n} \overline{A}_i$, siis

$$P\left(\sum_{i=1}^{n} A_i\right) = 1 - P\left(\prod_{i=1}^{n} \overline{A}_i\right). \tag{1.7.4}$$

Definitsiooni 1.6.3 põhjal

$$P(AB) = P(A) \cdot P(B|A), \tag{1.7.5}$$

$$P(AB) = P(B) \cdot P(A|B), \tag{1.7.6}$$

kus $P(B|A)$ on sündmuse B toimumise tõenäosus eeldusel, et sündmus A on toimunud ja $P(A|B)$ on sündmuse A toimumise tõenäosus eeldusel, et sündmus B on toimunud.

Definitsioon 1. Sündmust B nimetatakse *sõltumatuks* sündmusest A, kui

$$P(B|A) = P(B). \tag{1.7.7}$$

Järeldus 1. Kui sündmus B on sõltumatu sündmusest A, siis sündmus A on sõltumatu sündmusest B, st

$$P(A|B) = P(A).$$

Seejuures on sündmused A ja B sõltumatud parajasti siis, kui

$$P(AB) = P(A) \cdot P(B). \tag{1.7.8}$$

Tõestus. Saame väidete ahela

$$\left\{ \begin{array}{c} P(AB) \overset{(1.7.5)}{=} P(A) \cdot P(B|A) \overset{(1.7.7)}{=} P(A) \cdot P(B) \\ \\ P(AB) \overset{(1.7.6)}{=} P(B) \cdot P(A|B) \end{array} \right. \Rightarrow$$

$$\Rightarrow\ P(A) \cdot P(B) = P(B) \cdot P(A|B)\ \Rightarrow\ P(A|B) = P(A). \qquad \square$$

Järeldus 2. Kui sündmused A ja B on sõltumatud, siis ka $\overline{A}$ ja $\overline{B}$ on sõltumatud.

Tõestus. Kuna

$$P(\overline{A} \cdot \overline{B}) \overset{(1.6.2)}{=} 1 - P\left(\overline{\overline{A} \cdot \overline{B}}\right) \overset{(1.1.1)}{=} 1 - P\left(A + B\right) =$$

$$= 1 - P\left(A\right) - P\left(B\right) + P\left(AB\right) \overset{(1.7.8)}{=} 1 - P\left(A\right) - P\left(B\right) + P\left(A\right) P\left(B\right) =$$

$$= \left(1 - P\left(A\right)\right)\left(1 - P\left(B\right)\right) = P\left(\overline{A}\right) P\left(\overline{B}\right),$$

siis Järelduse 1 põhjal on sündmused $\overline{A}$ ja $\overline{B}$ sõltumatud. $\square$

Kui sündmuste korrutises on kolm tegurit, siis

$$\begin{aligned} P(A_1A_2A_3) &= \left[\begin{array}{c} \text{sündmuste korrutamine} \\ \text{on assotsiatiivne} \end{array}\right] = P((A_1A_2)\,A_3) = \\ &= \left[\begin{array}{c} \text{rakendame seost } (1.7.5)\,, \\ A = A_1A_2,\ B = A_3 \end{array}\right] = P(A_1A_2) \cdot P(A_3|A_1A_2) = \\ &= \left[\begin{array}{c} \text{rakendame } (1.7.5)\,, \\ A = A_1,\ B = A_2 \end{array}\right] = P(A_1)P(A_2|A_1) \cdot P(A_3|A_1A_2), \end{aligned}$$

kus $P(A_3|A_1A_2)$ on sündmuse A_3 toimumise tinglik tõenäosus tingimusel, et sündmused A_1 ja A_2 on toimunud.

Lause 2. Kehtib väide

$$P\left(\prod_{i=1}^{n} A_i\right) = P\left(A_1\right) \prod_{j=2}^{n} P\left(A_j \left| \prod_{k=1}^{j-1} A_k\right.\right), \tag{1.7.9}$$

kus $P\left(A_j | \prod_{k=1}^{j-1} A_k\right)$ on sündmuse A_j toimumise tõenäosus tingimusel, et sündmused $A_1, \ldots, A_{j-1}$ on toimunud.

Tõestus. Kasutame matemaatilise induktsiooni meetodit. Eelneva põhjal on baas olemas, $n = 2$ ja $n = 3$ korral väide kehtib. Induktsioonisammu lubatavuse tõestamiseks eeldame, et väide on tõene $n - 1$ korral, st

$$P\left(\prod_{i=1}^{n-1} A_i\right) = P(A_1) \prod_{j=2}^{n-1} P\left(A_j \middle| \prod_{k=1}^{j-1} A_k\right). \tag{1.7.10}$$

Sel korral saame

$$P\left(\prod_{i=1}^{n} A_i\right) = P\left(\left(\prod_{i=1}^{n-1} A_i\right) A_n\right) = \left[\begin{array}{c}\text{rakendame seost (1.7.5)} \\ A = \left(\prod_{i=1}^{n-1} A_i\right),\ B = A_n\end{array}\right] =$$

$$= P\left(\prod_{i=1}^{n-1} A_i\right) P\left(A_n \middle| \prod_{i=1}^{n-1} A_i\right) \overset{(1.7.10)}{=}$$

$$= P(A_1)\left(\prod_{j=2}^{n-1} P\left(A_j \middle| \prod_{k=1}^{j-1} A_k\right)\right) P\left(A_n \middle| \prod_{i=1}^{n-1} A_i\right) =$$

$$= P(A_1) \prod_{j=2}^{n} P\left(A_j \middle| \prod_{k=1}^{j-1} A_k\right),$$

st kui Lause 2 väide on tõene $n - 1$ korral, siis see väide on tõene ka n korral. Seega on induksioonisamm lubatud ja Lause 2 väide on tõestatud matemaatilise induktsiooni meetodil. $\square$

Definitsioon 2. Sündmuste süsteemi $\{A_1, A_2, \ldots, A_n\}$ nimetatakse *sõltumatuks*, kui

$$P\left(A_k \middle| \prod_{i<k} A_i\right) = P(A_k) \quad (k = 2; 3; \ldots; n).$$

Järeldus 3. Kui sündmuste süsteem $\{A_1, A_2, \ldots, A_n\}$ on sõltumatu parajasti siis, kui

$$P\left(\prod_{i=1}^{n} A_i\right) = \prod_{i=1}^{n} P(A_i). \tag{1.7.11}$$

Märkus 1. Süsteemi $\{A_1, A_2, \ldots, A_n\}$ sündmuste paarikaupa sõltumatusest ei järeldu selle sündmuste süsteemi sõltumatus.

Näide 1. Münti visatakse kaks korda. Leiame $P(A)$ ja $P(B)$, kui A on sündmus, et saadakse kaks kulli, ja B on sündmus, et saadakse vähemalt üks kull.

Olgu A_k $(k = 1; 2)$ sündmus, et k-ndal viskel saadakse kull. Kuna $A = A_1A_2$ ja $B = A_1 + A_2$, siis

$$P(A) = P(A_1A_2) = [A_k\text{-d on sõltumatud}] \overset{(1.7.8)}{=} P(A_1)\,P(A_2) = 1/4,$$
$$P(B) = P(A_1 + A_2) \overset{(1.7.1)}{=} P(A_1) + P(A_2) - P(A_1A_2) = 3/4. \qquad \diamondsuit$$

Näide 2. Kaks poissi sooritavad kordamööda vabaviskeid. Mõlemal on 2 viset. Preemia saab see poistest, kes esimesena tabab. Esimesena viskaja tabamise tõenäosus on igal viskel 0.4 ja teisel 0.7. Leiame mõlema poisi preemia saamise tõenäosuse. Milline on tõenäosus, et preemiat ei saa kumbki poiss?

Olgu sündmus $A-$ preemia saab see poiss, kes viskab esimesena ja sündmus $B-$preemia saab teine poiss ning sündmus $C-$kumbki poiss ei saa preemiat. Kui A_i $(i = 1; 2)$ on sündmus, et esimesena viskaja tabab i-ndal viskel ja B_i $(i = 1; 2)$ on sündmus, et teisena viskaja tabab i-ndal viskel. Esimene poiss saab preemia, kui ta tabab esimesel viskel või ta esimesel viskel ei taba ja ka teine poiss ei taba esimesel viskel ja esimene poiss tabab teisel viskel, st

$$A = A_1 + \overline{A}_1\overline{B}_1A_2. \tag{1.7.12}$$

Analoogiliselt saame

$$B = \overline{A}_1B_1 + \overline{A}_1\overline{B}_1\overline{A}_2B_2 \tag{1.7.13}$$

ja

$$C = \overline{A}_1\overline{B}_1\overline{A}_2\overline{B}_2. \tag{1.7.14}$$

Seosest (1.7.12) järeldub

$$P(A) = P\left(A_1 + \overline{A}_1\overline{B}_1A_2\right) = \left[\begin{array}{c}\text{esimeses liidetavas on } A_1 \text{ ja teises}\\ \text{ühe tegurina } \overline{A}_1, \text{ st liidetavad on}\\ \text{teineteist välistavad}\end{array}\right] =$$
$$= P(A_1) + P\left(\overline{A}_1\overline{B}_1A_2\right) = P(A_1) + P\left(\overline{A}_1\right)P\left(\overline{B}_1|\overline{A}_1\right)P\left(A_2|\overline{A}_1\overline{B}_1\right)$$
$$= 0.4 + 0.6 \cdot 0.3 \cdot 0.4 = 0.\,472$$

ja seosest (1.7.13) järeldub

$$P(B) = P\left(\overline{A}_1B_1 + \overline{A}_1\overline{B}_1\overline{A}_2B_2\right) = P\left(\overline{A}_1B_1\right) + P\left(\overline{A}_1\overline{B}_1\overline{A}_2B_2\right) =$$
$$= P\left(\overline{A}_1\right)P\left(B_1|\overline{A}_1\right) + P\left(\overline{A}_1\right)P\left(\overline{B}_1|\overline{A}_1\right)P\left(\overline{A}_2|\overline{A}_1\overline{B}_1\right)P\left(B_2|\overline{A}_1\overline{B}_1\overline{A}_2\right) =$$
$$= 0.6 \cdot 0.7 + 0.6 \cdot 0.3 \cdot 0.6 \cdot 0.7 = 0.\,495\,6.$$

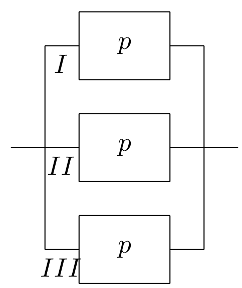
Seosest (1.7.14) järeldub

$$P(C) = P\left(\overline{A}_1\overline{B}_1\overline{A}_2\overline{B}_2\right) =$$
$$= P\left(\overline{A}_1\right)P\left(\overline{B}_1|\overline{A}_1\right)P\left(\overline{A}_2|\overline{A}_1\overline{B}_1\right)P\left(\overline{B}_2|\overline{A}_1\overline{B}_1\overline{A}_2\right) =$$
$$= 0.6 \cdot 0.3 \cdot 0.6 \cdot 0.3 = 0.0\,324.$$

Teostame kontrolli:

$$\begin{aligned} P(A+B+C) &= \left[\begin{array}{c} \text{süsteem } \{A,B,C\} \text{ on} \\ \text{täielik} \end{array} \right] = \\ &= P(A) + P(B) + P(C) = \\ &= 0.472 + 0.4956 + 0.0324 = 1.0 \qquad \diamondsuit \end{aligned}$$

Näide 3. Elemendi töökindlus on p. Töökindluse tõstmiseks dubleeritakse seda elementi paralleelselt kahe sama töökindlusega elemendiga. Saadakse süsteem



Leiame saadud süsteemi töökindluse. Töökindlus on tõenäosus, et vaadeldav süsteem peab aja T vastu. See süsteem on töökorras, kui on töökorras vähemalt üks paralleel. Elemendid lähevad rivist välja üksteisest sõltumatult.

Olgu sündmus $A-$ süsteem peab vastu aja T ja sündmus A_i- i-s element $(i = 1; 2; 3)$ peab vastu aja T. Saame

$$\begin{aligned} & P(A) = P(A_1 + A_2 + A_3) \overset{(1.7.2)}{=} \\ & = P(A_1) + P(A_2) + P(A_3) - P(A_2A_3) - P(A_1A_2) - \\ & - P(A_1A_3) + P(A_1A_2A_3) \overset{\text{sõltumatud}}{=} \\ & = P(A_1) + P(A_2) + P(A_3) - P(A_2)P(A_3) - P(A_1)P(A_2) - \\ & - P(A_1)P(A_3) + P(A_1)P(A_2)P(A_3) = \\ & = 3p - 3p^2 + p^3. \end{aligned}$$

Lihtsama lahenduse saame, kui kasutame valemit (1.7.4)

$$\begin{aligned} P(A_1 + A_2 + A_3) &= 1 - P\left(\overline{A}_1 \cdot \overline{A}_2 \cdot \overline{A}_3\right) = \\ &= 1 - P\left(\overline{A}_1\right) P\left(\overline{A}_2\right) P\left(\overline{A}_3\right) = \\ &= 1 - (1-p)^3 = 3p - 3p^2 + p^3. \qquad \diamondsuit \end{aligned}$$

1.8 Täistõenäosus. Bayesi valem

Sooritatakse katse, mille käigus jälgitakse sündmuse A toimumist. Olgu $\{H_1, H_2, \ldots, H_n\}$ täielik sündmuste süsteem selle katse korral, st

$$H_1 + H_2 + \ldots + H_n = K, \quad H_i H_j = V \ (i \neq j).$$

Nimetame seda süsteemi *hüpoteeside süsteemiks*. Kuna

$$A \overset{\text{Lause 1.1.1}}{=} K \cdot A = (H_1 + H_2 + \ldots + H_n) \cdot A \overset{\text{Lause 1.1.2}}{=}$$
$$= H_1 A + H_2 A + \ldots + H_n A,$$

siis

$$P(A) = P(H_1 A + H_2 A + \ldots + H_n A).$$

Tingimusest $H_i H_j = V \ (i \neq j)$ järeldub, et süsteemi $\{H_i A\}_{i=1,\ldots,n}$ sündmused on üksteist välistavad. Seega

$$P(A) = P(H_1 A + H_2 A + \ldots + H_n A) =$$
$$= P(H_1 A) + P(H_2 A) + \ldots + P(H_n A) \overset{(1.7.5)}{=}$$
$$= P(H_1) P(A|H_1) + P(H_2) P(A|H_2) + \ldots + P(H_n) P(A|H_n).$$

Oleme tõestanud järgmise väite.

Lause 1 (täistõenäosuse valem). Kui $\{H_1, H_2, \ldots, H_n\}$ on hüpoteeside süsteem katse jaoks, mille korral uuritakse sündmuse A toimumist, siis

$$P(A) = P(H_1) P(A|H_1) + P(H_2) P(A|H_2) + \ldots + P(H_n) P(A|H_n)$$

ehk lühidalt

$$P(A) = \sum_{i=1}^{n} P(H_i) P(A|H_i). \tag{1.8.1}$$

Näide 1. Kohtunik valib ühe kahest korvpallurist sooritama vabaviset. Seejuures on esimesel neist vabaviske tabamise tõenäosus 0.7 ja teisel 0.9. Leiame tõenäosuse, et vise tabab.

Olgu sündmus $A-$ vise tabab. Koostame hüpoteeside süsteemi $\{H_1, H_2\}$, kus hüpotees H_1 – viskab esimene korvpallur ja hüpotees H_2 – viskab teine korvpallur. Kuna puudub täpsem info, siis $P(H_1) = P(H_2) = 1/2$. Seejuures $P(A|H_1) = 0.7$ ja $P(A|H_1) = 0.9$. Valemi (1.8.1) põhjal saame

$$P(A) = P(H_1) P(A|H_1) + P(H_2) P(A|H_2) = \frac{1}{2} \cdot 0.7 + \frac{1}{2} \cdot 0.9 = 0.8. \quad \diamondsuit$$

Näide 2. Karbis on 5 uut ja 4 kasutatud tennisepalli. Esimeseks mänguks võetakse karbist huupi 2 palli, mis pärast mängu pannakse karpi tagasi. Teiseks mänguks võetakse seejärel huupi 2 palli. Leiame tõenäosuse, et teiseks mänguks võetud pallide seas on täpselt $i \ \ (i = 0; 1; 2)$ uut palli.

Olgu A_k $(k = 0; 1; 2)$ sündmus, et teiseks mänguks võeti k uut palli. Koostame hüpoteeside süsteemi $\{H_0, H_1, H_2\}$, kus H_i on hüpotees, et esimeseks mänguks võeti täpselt i uut palli.

I lahendusvariant. Olgu B_i $(i = 1; 2)$ sündmus, et esimeseks mänguks i-ndana võetud pall on uus ja C_i $(i = 1; 2; 3)$ sündmus, et teiseks mänguks i-ndana võetud pall on uus. Sellise tähistuse korral saame

$$H_0 = \overline{B}_1\overline{B}_2,\ H_1 = B_1\overline{B}_2 + \overline{B}_1 B_2,\ H_2 = B_1 B_2$$

ja

$$\begin{aligned}
P(H_0) &= P\left(\overline{B}_1\overline{B}_2\right) = P\left(\overline{B}_1\right) P\left(\overline{B}_2|\overline{B}_1\right) = \frac{4}{9}\cdot\frac{3}{8} = \frac{1}{6},\\
P(H_1) &= P\left(B_1\overline{B}_2 + \overline{B}_1 B_2\right) = \begin{bmatrix}\text{liidetavad on teineteist}\\ \text{välistavad}\end{bmatrix} =\\
&= P\left(B_1\overline{B}_2\right) + P\left(\overline{B}_1 B_2\right) = P(B_1) P\left(\overline{B}_2|B_1\right) + P\left(\overline{B}_1\right) P\left(B_2|\overline{B}_1\right) =\\
&= \frac{5}{9}\cdot\frac{4}{8} + \frac{4}{9}\cdot\frac{5}{8} = \frac{5}{9},\\
P(H_2) &= P(B_1 B_2) = P(B_1) P(B_2|B_1) = \frac{5}{9}\cdot\frac{4}{8} = \frac{5}{18}.
\end{aligned}$$

Kontrollime, hüpoteeside tõenäosuste summa peab olema 1 :

$$P(H_0) + P(H_1) + P(H_2) = \frac{1}{6} + \frac{5}{9} + \frac{5}{18} = 1.$$

Et

$$A_0 = \overline{C}_1\overline{C}_2,\ A_1 = C_1\overline{C}_2 + \overline{C}_1 C_2,\ A_2 = C_1 C_2,$$

siis saame tinglikud tõenäosused

$$\begin{aligned}
P(A_0|H_0) &= P\left(\overline{C}_1\overline{C}_2|H_0\right) = P\left(\overline{C}_1|H_0\right)\cdot P\left(\overline{C}_2|H_0\overline{C}_1\right) = \frac{4}{9}\cdot\frac{3}{8} = \frac{1}{6},\\
P(A_0|H_1) &= P\left(\overline{C}_1\overline{C}_2|H_1\right) = P\left(\overline{C}_1|H_1\right)\cdot P\left(\overline{C}_2|H_1\overline{C}_1\right) = \frac{5}{9}\cdot\frac{4}{8} = \frac{5}{18},\\
P(A_0|H_2) &= P\left(\overline{C}_1\overline{C}_2|H_2\right) = P\left(\overline{C}_1|H_2\right)\cdot P\left(\overline{C}_2|H_2\overline{C}_1\right) = \frac{6}{9}\cdot\frac{5}{8} = \frac{5}{12},\\
P(A_1|H_0) &= P\left(\left(C_1\overline{C}_2 + \overline{C}_1 C_2\right)|H_0\right) \overset{\text{liidetavad on teineteist välistavad}}{=}\\
&= P\left(\left(C_1\overline{C}_2\right)|H_0\right) + P\left(\left(\overline{C}_1 C_2\right)|H_0\right) =\\
&= P(C_1|H_0) P\left(\overline{C}_2|(H_0 C_1)\right) + P\left(\overline{C}_1|H_0\right) P\left(C_2|H_0\overline{C}_1\right) =\\
&= \frac{5}{9}\cdot\frac{4}{8} + \frac{4}{9}\cdot\frac{5}{8} = \frac{5}{9},\\
P(A_1|H_1) &= P\left(\left(C_1\overline{C}_2 + \overline{C}_1 C_2\right)|H_1\right) \overset{\text{liidetavad on teineteist välistavad}}{=}\\
&= P\left(\left(C_1\overline{C}_2\right)|H_1\right) + P\left(\left(\overline{C}_1 C_2\right)|H_1\right) =\\
&= P(C_1|H_1) P\left(\overline{C}_2|(H_1 C_1)\right) + P\left(\overline{C}_1|H_1\right) P\left(C_2|H_1\overline{C}_1\right) =\\
&= \frac{4}{9}\cdot\frac{5}{8} + \frac{5}{9}\cdot\frac{4}{8} = \frac{5}{9},
\end{aligned}$$

$$
\begin{aligned}
P\left(A_1|H_2\right) &= P\left(\left(C_1\overline{C}_2+\overline{C}_1C_2\right)|H_2\right) \overset{\text{liidetavad on teineteist välistavad}}{=} \\
&= P\left(\left(C_1\overline{C}_2\right)|H_2\right)+P\left(\left(\overline{C}_1C_2\right)|H_2\right) = \\
&= P\left(C_1|H_2\right)P\left(\overline{C}_2|\left(H_2C_1\right)\right)+P\left(\overline{C}_1|H_2\right)P\left(C_2|H_2\overline{C}_1\right) = \\
&= \frac{3}{9}\cdot\frac{6}{8}+\frac{6}{9}\cdot\frac{3}{8}=\frac{1}{2},
\end{aligned}
$$

$$
\begin{aligned}
P\left(A_2|H_0\right) &= P\left(\left(C_1C_2\right)|H_0\right) = P\left(C_1|H_0\right)P\left(C_2|\left(H_0C_1\right)\right) = \frac{5}{9}\cdot\frac{4}{8}=\frac{5}{18}, \\
P\left(A_2|H_1\right) &= P\left(\left(C_1C_2\right)|H_1\right) = P\left(C_1|H_1\right)P\left(C_2|\left(H_1C_1\right)\right) = \frac{4}{9}\cdot\frac{3}{8}=\frac{1}{6}, \\
P\left(A_2|H_2\right) &= P\left(\left(C_1C_2\right)|H_2\right) = P\left(C_1|H_2\right)P\left(C_2|\left(H_2C_1\right)\right) = \frac{3}{9}\cdot\frac{2}{8}=\frac{1}{12}.
\end{aligned}
$$

Et

$$
P(A_k) \overset{(1.8.1)}{=} \sum_{i=0}^{2} P\left(H_i\right)P\left(A_k|H_i\right) \quad (k=0;1;2),
$$

siis

$$
\begin{aligned}
P(A_0) &= P\left(H_0\right)P\left(A_0|H_0\right)+P\left(H_1\right)P\left(A_0|H_1\right)+P\left(H_2\right)P\left(A_0|H_2\right) = \\
&= \frac{1}{6}\cdot\frac{1}{6}+\frac{5}{9}\cdot\frac{5}{18}+\frac{5}{18}\cdot\frac{5}{12}=\frac{193}{648},
\end{aligned}
$$

$$
\begin{aligned}
P(A_1) &= P\left(H_0\right)P\left(A_1|H_0\right)+P\left(H_1\right)P\left(A_1|H_1\right)+P\left(H_2\right)P\left(A_1|H_2\right) = \\
&= \frac{1}{6}\cdot\frac{5}{9}+\frac{5}{9}\cdot\frac{5}{9}+\frac{5}{18}\cdot\frac{1}{2}=\frac{175}{324}, \\
P(A_2) &= P\left(H_0\right)P\left(A_2|H_0\right)+P\left(H_1\right)P\left(A_2|H_1\right)+P\left(H_2\right)P\left(A_2|H_2\right) = \\
&= \frac{1}{6}\cdot\frac{5}{18}+\frac{5}{9}\cdot\frac{1}{6}+\frac{5}{18}\cdot\frac{1}{12}=\frac{35}{216}.
\end{aligned}
$$

Kuna sündmuste süsteem $\{A_0, A_1, A_2\}$ on täielik, siis

$$
P(A_0)+P(A_1)+P(A_2)=\frac{193}{648}+\frac{175}{324}+\frac{35}{216}=1.
$$

II lahendusvariant. Kasutame klassikalist tõenäosuse definitsiooni. Elementaarsündmuseks nii esimese kui ka teise mängu pallide võtmisel valime suvalise pallipaari väljavõtmise. Et pallide järjekord paaris ei ole oluline, siis on tegemist kombinatsiooniga üheksast elemendist kahe kaupa. Nende koguarv on $C_9^2=\binom{9}{2}=36$. Et hüpoteesi H_0 realiseerumiseks on neist soodsaid $C_4^2=6$, siis

$$
P\left(H_0\right)=\frac{6}{36}=\frac{1}{6}.
$$

Kuna hüpoteesi H_1 realiseerumiseks on soodsaid $C_4^1\cdot C_5^1=\binom{4}{1}\binom{5}{1}=20$, siis

$$
P\left(H_1\right)=\frac{20}{36}=\frac{5}{9}.
$$

Et hüpoteesi H_2 realiseerumiseks on soodsaid $C_5^2 = \binom{5}{2} = 10$, siis

$$P(H_2) = \frac{10}{36} = \frac{5}{18}.$$

Analoogiliselt leiame veel tinglikud tõenäosused:

$$\begin{aligned}
&P(A_0|H_0) = \frac{C_4^2}{C_9^2} = \frac{1}{6}, \ P(A_0|H_1) = \frac{C_5^2}{C_9^2} = \frac{5}{18}, \ P(A_0|H_2) = \frac{C_6^2}{C_9^2} = \frac{5}{12}, \\
&P(A_1|H_0) = \frac{C_5^1 \cdot C_4^1}{C_9^2} = \frac{5}{9}, \ P(A_1|H_1) = \frac{C_4^1 \cdot C_5^1}{C_9^2} = \frac{5}{9} \\
&P(A_1|H_2) = \frac{C_3^1 \cdot C_6^1}{C_9^2} = \frac{1}{2}, \ P(A_2|H_0) = \frac{C_5^2}{C_9^2} = \frac{5}{18}, \\
&P(A_2|H_1) = \frac{C_4^2}{C_9^2} = \frac{1}{6}, \ P(A_2|H_2) = \frac{C_3^2}{C_9^2} = \frac{1}{12}. \qquad \diamondsuit
\end{aligned}$$

Kuna

$$AH_k = H_k A \quad \Rightarrow \quad P(AH_k) = P(H_k A),$$

siis

$$P(A)\, P(H_k|A) = P(H_k)\, P(A|H_k).$$

Viimasest seosest saame avaldada tõenäosuse $P(H_k|A)$:

$$P(H_k|A) = \frac{P(H_k)\, P(A|H_k)}{P(A)} \overset{(1.8.1)}{=} \frac{P(H_k)\, P(A|H_k)}{\sum_{i=1}^{n} P(H_i)\, P(A|H_i)}.$$

Oleme tõestanud järgmise väite.

Lause 2 (Bayesi valem). Kui $\{H_1, H_2, \ldots, H_n\}$ on hüpoteeside süsteem katse jaoks, mille korral uuritakse sündmuse A toimumist, siis

$$P(H_k|A) = \frac{P(H_k)\, P(A|H_k)}{P(H_1)\, P(A|H_1) + P(H_2)\, P(A|H_2) + \ldots + P(H_n)\, P(A|H_n)}$$

ehk lühidalt

$$P(H_k|A) = \frac{P(H_k)\, P(A|H_k)}{\sum_{i=1}^{n} P(H_i)\, P(A|H_i)} = \frac{P(H_k)\, P(A|H_k)}{P(A)}. \tag{1.8.2}$$

Bayesi valemis esinevat suurust $P(H_k)$ nimetatakse hüpoteesi H_k *aprioorseks* ehk *katse-eelseks* tõenäosuseks ja suurust $P(H_k|A)$ nimetatakse hüpoteesi H_k *aposterioorseks* ehk *katsejärgseks* tõenäosuseks.

Näide 3. Leiame Näites 2 esitatud andmetel, teades katse tulemusena sündmuse A_2 toimumist, tõenäosuse, et esimene kord võeti kaks kasutatud palli.

Vaja on leida hüpoteesi H_0 aposterioorne tõenäosus, teades katse tulemust, sündmuse A_2 toimumist. Bayesi valemi põhjal saame

$$P(H_0|A_2) = \frac{P(H_0)\,P(A_2|H_0)}{P(H_0)\,P(A_2|H_0) + P(H_1)\,P(A_2|H_1) + P(H_2)\,P(A_2|H_2)} =$$
$$= \frac{\frac{1}{6}\cdot\frac{5}{18}}{\frac{1}{6}\cdot\frac{5}{18} + \frac{5}{9}\cdot\frac{1}{6} + \frac{5}{18}\cdot\frac{1}{12}} = \frac{\frac{5}{108}}{\frac{35}{216}} = \frac{2}{7}. \qquad \diamondsuit$$

Näide 4. Eksamil on viis piletit, igas kaks küsimust. Erinevates piletites on erinevad küsimused, st kokku 10 küsimust. Üliõpilane teab seitset küsimust. Ta sooritab eksami, kui teab mõlemat küsimust võetud piletist või täpselt üht küsimust võetud piletist ja lisaküsimust, mis antakse siis talle ühest teisest piletist. Leiame tõenäosuse, et üliõpilane sooritab eksami. Leiame tõenäosuse, et üliõpilane pidi vastama lisaküsimusele, kui on teada, et ta sooritas eksami.

Olgu A sündmus, et üliõpilane sooritab eksami, ja H_i $(i = 0; 1; 2)$ hüpotees, et üliõpilane teab võetud piletist täpselt i küsimust. Kui B_i $(i = 1; 2)$ on sündmus, et üliõpilane teab võetud pileti i-ndat küsimust, siis

$$H_0 = \overline{B}_1\overline{B}_2,\ H_1 = B_1\overline{B}_2 + \overline{B}_1 B_2,\ H_2 = B_1 B_2$$

ja

$$P(H_0) = P\left(\overline{B}_1\overline{B}_2\right) = P\left(\overline{B}_1\right) P\left(\overline{B}_2|\overline{B}_1\right) = \frac{3}{10}\cdot\frac{2}{9} = \frac{1}{15},$$
$$P(H_1) = P\left(B_1\overline{B}_2 + \overline{B}_1 B_2\right) \overset{\text{liidetavad teineteist välistavad}}{=}$$
$$= P\left(B_1\overline{B}_2\right) + P\left(\overline{B}_1 B_2\right) = P(B_1)\,P\left(\overline{B}_2|B_1\right) + P\left(\overline{B}_1\right) P\left(B_2|\overline{B}_1\right) =$$
$$= \frac{7}{10}\cdot\frac{3}{9} + \frac{3}{10}\cdot\frac{7}{9} = \frac{7}{15},$$
$$P(H_2) = P(B_1B_2) = P(B_1)\,P(B_2|B_1) = \frac{7}{10}\cdot\frac{6}{9} = \frac{7}{15}.$$

Kontrollime, hüpoteeside tõenäosuste summa peab olema 1:

$$P(H_0) + P(H_1) + P(H_2) = \frac{1}{15} + \frac{7}{15} + \frac{7}{15} = 1.$$

Leiame tinglikud tõenäosused:

$$P(A|H_0) = 0,\ P(A|H_1) = \frac{6}{8} = \frac{3}{4},\ P(A|H_2) = 1.$$

Täistõenäosusvalemi (1.8.1) abil saame

$$P(A) = P(H_0)\,P(A|H_0) + P(H_1)\,P(A|H_1) + P(H_2)\,P(A|H_2) = \frac{49}{60}.$$

Bayesi valemi (1.8.2) abil leiame hüpoteesi H_1 aposterioorse tõenäosuse, teades, et üliõpilane sai eksamil läbi,

$$P(H_1|A) = \frac{P(H_1)\,P(A|H_1)}{P(H_0)\,P(A|H_0) + P(H_1)\,P(A|H_1) + P(H_2)\,P(A|H_2)} = \frac{3}{7}. \quad \diamondsuit$$

1.9 Bernoulli valem

Vaatleme katseseeriat, milles on n sõltumatut katset samadel tingimustel. Oletame, et sündmus A toimub igal katsel tõenäosusega p. Sel korral kõneldakse katseseeria läbiviimisest *Bernoulli skeemi* järgi. Olgu $0 < p < 1$ ja $q \stackrel{\text{def}}{=} 1 - p$. Olgu $B_{n,m}$ sündmus, et A toimub selle n-katselise seeria korral täpselt m $(0 \le m \le n)$ korda. Meid huvitab tõenäosus $P_{n,m} \stackrel{\text{def.}}{=} P(B_{n,m})$.

Lause 1 (Bernoulli valem). Kui katseseerias on n sõltumatut katset ja sündmus A toimub igal katsel tõenäosusega p ning $B_{n,m}$ on sündmus, et A toimub selle seeria korral täpselt m $(0 \le m \le n)$ korda, siis

$$P_{n,m} = P(B_{n,m}) = C_n^m p^m q^{n-m}. \tag{1.9.1}$$

Tõestus. Kasutame matemaatilise induktsiooni meetodit. Olgu sündmus A_i – sündmus A toimub i-ndal $(1 \le i \le n)$ katsel. Seega

$$P(A_i) = p, \quad P\left(\overline{A}_i\right) = q \quad (1 \le i \le n).$$

Kui $n = 1$, siis

$$\begin{aligned} B_{1;0} = \overline{A}_1 \;&\Rightarrow\; P_{1;0} = P(B_{1;0}) = P\left(\overline{A}_1\right) = q = 1 \cdot p^0 q^1 = C_1^0 p^0 q^1, \\ B_{0;1} = A_1 \;&\Rightarrow\; P_{1;1} = P(B_{1;1}) = P(A_1) = p = 1 \cdot p^1 q^0 = C_1^1 p^1 q^0, \end{aligned}$$

st induktsiooni baas on olemas. Kasutades seoseid

$$C_n^m + C_n^{m-1} = C_{n+1}^m, \quad (n-m)\,C_n^m / (m+1) = C_n^{m+1},$$

tõestage iseseisvalt vastavalt induktsioonisammu $n \to n+1$ ja $m \to m+1$ lubatavus. $\square$

Uurime järgnevalt, millise m väärtuse korral on tõenäosus $P_{n,m}$ suurim. Sellise m saame määrata võrratuste süsteemist

$$\begin{cases} P_{n,m-1} \le P_{n,m} \\ P_{n,m+1} \le P_{n,m} \end{cases} \Leftrightarrow \begin{cases} C_n^{m-1} p^{m-1} q^{n-m+1} \le C_n^m p^m q^{n-m} \\ C_n^{m+1} p^{m+1} q^{n-m-1} \le C_n^m p^m q^{n-m}. \end{cases}$$

Saame võrratuste süsteemi

$$\begin{cases} \dfrac{n!}{(m-1)!\,(n-m+1)!} p^{m-1} q^{n-m+1} \le \dfrac{n!}{m!\,(n-m)!} p^m q^{n-m} \\ \dfrac{n!}{(m+1)!\,(n-m-1)!} p^{m+1} q^{n-m-1} \le \dfrac{n!}{m!\,(n-m)!} p^m q^{n-m}, \end{cases}$$

millest peale lihtsustamist leiame

$$\left\{\begin{array}{c} mq \leq (n-m+1)\,p \\ (n-m)\,p \leq (m+1)\,q \end{array}\right. \quad \Leftrightarrow \quad \left\{\begin{array}{c} mq+mp \leq (n+1)\,p \\ np-q \leq mq+mp \end{array}\right.$$

ehk

$$np-q \leq m \leq np+p. \tag{1.9.2}$$

Kuna

$$(np+p)-(np-q)=p+q=1,$$

siis (1.9.2) määrab lõigu pikkusega 1. Seega on vaadeldaval probleemil täisarvulise $np+p$ korral kaks lahendit $m=np+p$ ja $m=np-q$. Kui $np+p$ ei ole täisarv, siis on üks lahend $m=[np+p]$, kus $[np+p]$ on täisosa arvust $np+p$.

Seega oleme saanud järgmise tulemuse.

Lause 2. Kui katseteseerias on n sõltumatut katset ja sündmus A toimub igal katsel tõenäosusega p ning $B_{n,m}$ on sündmus, et A toimub selle seeria korral täpselt m $(0 \leq m \leq n)$ korda, siis täisarvulise $p\,(n+1)$ korral on sündmuse $B_{n,m}$ toimumise tõenäosus suurim, kui $m=np-q$ või $m=p\,(n+1)$. Kui $p\,(n+1)$ ei ole täisarv, siis suurim tõenäosus saadakse $m=[p\,(n+1)]$ korral.

Näide 1. Korvpallur sooritab 3 vabaviset. Igal viskel on tabamise tõenäosus 0.7. Leiame järgmiste sündmuste tõenäosused: $B-$ korvpallur tabab vaid ühel viskel; $C-$ korvpallur tabab vähemalt ühel viskel; $D-$ korvpallur tabab ülimalt ühel viskel; $E-$ korvpallur tabab täpselt kahel viskel; $F-$ korvpallur tabab vähemalt kahel viskel; $G-$ korvpallur tabab ülimalt kahel viskel. Milline on tõenäoseim tabamuste arv?

Kui A on sündmus, et korvpallur vabaviskel tabab, siis saame $p=0.7$ ja $q=1-p=0.3$. Olgu $B_{3,m}$ sündmus, et korvpallur tabab 3-viskelise seeria korral täpselt m $(0 \leq m \leq 3)$ korda ja $P_{3,m}=P\,(B_{3,m})$. Et

$$B=B_{3;1},\ C=B_{3;1}+B_{3;2}+B_{3;3},\ D=B_{3;0}+B_{3;1},$$
$$E=B_{3;2},\ F=B_{3;2}+B_{3;3},\ G=B_{3;0}+B_{3;1}+B_{3;2}$$

ja $B_{3,i}B_{3,j}=V \quad (i \neq j)$ ning Lause 1 põhjal

$$P\,(B_{3,m})=C_3^m p^m q^{3-m},$$

siis

$$P(B_{3;0})=C_3^0 \cdot 0.7^0 \cdot 0.3^{3-0}=0.0\,27,\ P(B_{3;1})=C_3^1 \cdot 0.7^1 \cdot 0.3^{3-1}=0.189,$$
$$P(B_{3;2})=C_3^2 \cdot 0.7^2 \cdot 0.3^{3-2}=0.441=\max_{0 \leq m \leq 3} P(B_{3;m}),$$
$$P(B_{3;3})=C_3^3 \cdot 0.7^3 \cdot 0.3^{3-3}=0.\,343.$$

Kontrollime

$$P(B_{3;0})+P(B_{3;1})+P(B_{3;2})+P(B_{3;3})=1.$$

Järelikult

$$\begin{aligned}
&P(B)=P(B_{3;1})=0.189,\ P(E)=P(B_{3;2})=0.441,\\
&P(C)=P(B_{3;1})+P(B_{3;2})+P(B_{3;3})=0.973,\\
&P(D)=P(B_{3;0})+P(B_{3;1})=0.216,\ P(F)=P(B_{3;2})+P(B_{3;3})=0.784,\\
&P(G)=P(B_{3;0})+P(B_{3;1})+P(B_{3;2})=0.657.
\end{aligned}$$

Tõenäoseima tabamuste arvu saame leida ka Lause 2 põhjal

$$[(n+1)\,p]=[4\cdot 0.7]=2. \qquad \diamondsuit$$

Näide 2. Karbis on 10 tühja disketti, kusjuures igat neist on eelnevalt kasutatud tõenäosusega 0.6. Leiame järgmiste sündmuste tõenäosused: $B-$ karbis on täpselt 2 eelnevalt kasutamata disketti; $C-$ karbis on ülimalt 2 eelnevalt kasutamata disketti; $D-$ karbis on vähemalt 2 eelnevalt kasutamata disketti. Leiame tõenäoseima kasutamata diskettide arvu karbis.

Kui A on sündmus, et disketti ei ole eelnevalt kasutatud, siis $p=0.4$ ja $q=0.6$. Kui $B_{10,m}$ on sündmus, et eelnevalt kasutamata diskettide arv selles karbis on m, siis

$$B=B_{10;2},\ C=B_{10;0}+B_{10;1}+B_{10;2},\ \overline{D}=B_{10;0}+B_{10;1}$$

ning Lause 1 põhjal $P(B_{10;m})=C_{10}^{m}\cdot 0.4^{m}\cdot 0.6^{10-m}$. Seega

$$\begin{aligned}
P(B)&=P(B_{10;2})=C_{10}^{2}\cdot 0.4^{2}\cdot 0.6^{8}\approx 0.120\,9,\\
P(C)&=P(B_{10;0}+B_{10;1}+B_{10;2})\overset{\text{liidetavad üksteist välistavad}}{=}\\
&=P(B_{10;0})+P(B_{10;1})+P(B_{10;2})=\\
&=C_{10}^{0}\cdot 0.4^{0}\cdot 0.6^{10}+C_{10}^{1}\cdot 0.4^{1}\cdot 0.6^{9}+C_{10}^{2}\cdot 0.4^{2}\cdot 0.6^{8}\approx 0.167\,3,\\
P(\overline{D})&=P(B_{10;0}+B_{10;1})\overset{\text{teineteist välistavad}}{=}P(B_{10;0})+P(B_{10;1})=\\
&=C_{10}^{0}\cdot 0.4^{0}\cdot 0.6^{10}+C_{10}^{1}\cdot 0.4^{1}\cdot 0.6^{9}\approx 0.04\,64,\\
P(D)&=1-P(\overline{D})\approx 1-0.04\,64=0.953\,6.
\end{aligned}$$

Lause 2 abil leiame tõenäoseima kasutamata diskettide arvu karbis

$$[(10+1)\,0.4]=[4.\,4]=4. \qquad \diamondsuit$$

Märkus 1. Kui katseseerias sooritatakse n sõltumatut katset ja igal katsel toimub täpselt üks sündmustest $A_1,\ldots,A_k$ vastavalt tõenäosustega $p_1,\ldots,p_k$, kus $\sum_{i=1}^{k}p_i=1$, siis sündmuse $B_{m_1,m_2,\ldots,m_k;n}-$sündmus A_i $(i=1;\ldots;k)$ toimub katseseerias täpselt m_i korda $\left(\sum_{i=1}^{k}m_i=n\right)$, tõenäosus on leitav valemi

$$P(B_{m_1,m_2,\ldots,m_k;n})=\frac{n!}{m_1!m_2!\ldots m_k!}p_1^{m_1}\,p_2^{m_2}\cdots p_k^{m_k} \qquad (1.9.3)$$

abil.

Näide 3. Lõik jagatakse neljaks võrdse pikkusega osalõiguks. Lõigul valitakse huupi kaheksa punkti. Leiame tõenäosuse, et igasse osalõiku satub kaks punkti.

Rakendame valemit (1.9.3), võttes $n = 8$, $k = 4$, $m_1 = m_2 = m_3 = m_4 = 2$ ja $p_1 = p_2 = p_3 = p_4 = 1/4$. Saame

$$P\left(B_{2,2,2,2;8}\right) = \frac{8!}{2!2!2!2!}\left(\frac{1}{4}\right)^2\left(\frac{1}{4}\right)^2\left(\frac{1}{4}\right)^2\left(\frac{1}{4}\right)^2 = \frac{315}{8192}. \qquad \diamondsuit$$

1.10 Ülesanded

1. Korvi suunas sooritatakse kolm viset. Olgu A_k sündmus, et k-ndal viskel $(k = 1; 2; 3)$ tabatakse. Avaldage sündmuste A_k abil järgmised sündmused: A – täpselt üks vise tabab, B – ülimalt üks vise tabab, C – vähemalt üks vise tabab, D – täpselt kaks viset tabavad, E – ülimalt kaks viset tabavad, F – vähemalt kaks viset tabavad. V: $A = A_1\overline{A}_2\overline{A}_3 + \overline{A}_1A_2\overline{A}_3 + \overline{A}_1\overline{A}_2A_3$,
$B = \overline{A}_1\overline{A}_2\overline{A}_3 + A_1\overline{A}_2\overline{A}_3 + \overline{A}_1A_2\overline{A}_3 + \overline{A}_1\overline{A}_2A_3$,
$C = A_1\overline{A}_2\overline{A}_3 + \overline{A}_1A_2\overline{A}_3 + \overline{A}_1\overline{A}_2A_3 + A_1A_2\overline{A}_3 + A_1\overline{A}_2A_3 + \overline{A}_1A_2A_3 + A_1A_2A_3$,
$D = A_1A_2\overline{A}_3 + A_1\overline{A}_2A_3 + \overline{A}_1A_2A_3$,
$E = \overline{A}_1\overline{A}_2\overline{A}_3 + A_1\overline{A}_2\overline{A}_3 + \overline{A}_1A_2\overline{A}_3 + \overline{A}_1\overline{A}_2A_3 + A_1A_2\overline{A}_3 + A_1\overline{A}_2A_3 + \overline{A}_1A_2A_3$,
$F = A_1A_2\overline{A}_3 + A_1\overline{A}_2A_3 + \overline{A}_1A_2A_3 + A_1A_2A_3$.

2. Üliõpilasel tuleb eksamisessioonil sooritada 4 eksamit. Olgu A_i $(i = 1; 2; 3; 4)$ sündmus, et üliõpilane saab läbi i-ndal eksamil. Avaldage sündmuste A_i ja eelnevalt avaldatud sündmuste ning vastandsündmuste abil järgmised sündmused: A– üliõpilane sooritab kõik eksamid; B– üliõpilane põrub (täpselt) ühel eksamil; C – üliõpilane saab läbi vähemalt ühel eksamil; D – üliõpilane saab läbi (täpselt) kahel eksamil; E – üliõpilane saab läbi vähemalt kahel eksamil; F – üliõpilane saab läbi ülimalt kahel eksamil. V: $A = A_1A_2A_3A_4$,
$B = \overline{A}_1A_2A_3A_4 + A_1\overline{A}_2A_3A_4 + A_1A_2\overline{A}_3A_4 + A_1A_2A_3\overline{A}_4$, $C = \overline{\overline{A}_1\overline{A}_2\overline{A}_3\overline{A}_4}$,
$D = A_1A_2\overline{A}_3\overline{A}_4 + A_1\overline{A}_2A_3\overline{A}_4 + \overline{A}_1A_2A_3\overline{A}_4 + A_1\overline{A}_2\overline{A}_3A_4 + \overline{A}_1A_2\overline{A}_3A_4 +$
$+\overline{A}_1\overline{A}_2A_3A_4$, $E = A + B + D$, $F = D + A_1\overline{A}_2\overline{A}_3\overline{A}_4 + \overline{A}_1A_2\overline{A}_3\overline{A}_4 +$
$+\overline{A}_1\overline{A}_2A_3\overline{A}_4 + \overline{A}_1\overline{A}_2\overline{A}_3A_4 + \overline{A}_1\overline{A}_2\overline{A}_3\overline{A}_4$.

3. Täringut visatakse kaks korda. Olgu A_k sündmus, et esimesel viskel tuleb k silma ja B_n sündmus, et teisel viskel tuleb n silma. Olgu C, D, ja E sündmused, et kahe viskega saadakse vastavalt kaheksa silma, vähemalt kümme silma ning ülimalt neli silma. Avaldage sündmused C, D, ja E sündmuste A_k ning B_n abil. V: $C = A_2B_6 + A_3B_5 + A_4B_4 + A_5B_3 + A_6B_2$,
$D = A_4B_6 + A_5B_5 + A_6B_4 + A_5B_6 + A_6B_5 + A_6B_6$,
$E = A_1B_1 + A_1B_2 + A_2B_1 + A_1B_3 + A_2B_2 + A_3B_1$.

4. Eksamipiletis on kolm küsimust. Olgu A_k sündmus, et üliõpilane teab oma pileti k-ndat küsimust. Vaatleme järgmisi sündmusi: A–üliõpilane teab oma pileti iga küsimust; B–üliõpilane teab (täpselt) kaht küsimust oma piletist; C–üliõpilane teab (täpselt) üht küsimust oma piletist; D–üliõpilane teab ülimalt kaht küsimust oma piletist; E–üliõpilane teab vähemalt üht küsimust oma pi-

letist; F – üliõpilane ei tea ühtki küsimust oma piletist. Avaldage sündmused A, B, C, D, E ja F sündmuste A_k kaudu. V: $A = A_1 A_2 A_3$,
$B = \overline{A}_1 A_2 A_3 + A_1 \overline{A}_2 A_3 + A_1 A_2 \overline{A}_3$, $C = A_1 \overline{A}_2 \overline{A}_3 + \overline{A}_1 A_2 \overline{A}_3 + \overline{A}_1 \overline{A}_2 A_3$,
$E = \overline{\overline{A}_1 \overline{A}_2 \overline{A}_3}$, $F = \overline{A}_1 \overline{A}_2 \overline{A}_3$, $D = F + C + B$.

5. Punktid $P(x_1)$ ja $Q(x_2)$ valitakse huupi x-telje lõigul $[0;1]$. Leidke tõenäosus, et punktide P ja Q vaheline kaugus on väiksem-võrdne ühest kahendikust.
V: $3/4$.

6. Arvud x, y, $z \in [0;3]$ valitakse juhuslikult. Leidke tõenäosus, et nende arvude täisosade summa on 3, st $[x] + [y] + [z] = 3$. V: $10/27 \approx 0.370$.

7. Võetakse huupi kaks positiivset arvu x ja y, mis mõlemad on väiksemad kümnest. Leidke tõenäosus, et nende korrutis xy on väiksem kümnest ja jagatis x/y on suurem kui 2.5. V: $(1 + \ln 4)/20 \approx 0.1193$.

8. Kaks sõpra lõunatavad samas kohvikus kella 12 ja 14 vahel. Mõlemal kulub selleks pool tundi. Leidke tõenäosus, et antud päeval sõbrad selles kohvikus kohtuvad. V: $5/9 \approx 0.5556$.

9. Viieliikmelisse uurimisrühma kandideerib 11 teadurit, kellest 5 on daamid. Kui suur on tõenäosus, et sellesse uurimisrühma võetakse: 1) ainult daamid; 2) täpselt 3 daami; 3) vähemalt 3 daami; 4) ülimalt 3 daami? V: 1/462, 25/77, 181/462, 431/462.

10. Õpperühm, milles on 20 üliõpilast, neist 4 neidu, jaotatakse keeleõppeks kaheks (à 10 üliõpilast). Leidke tõenäosus, et 1) mõlemas on kaks neidu, 2) ühes on neli neidu, 3) ühes on täpselt üks neiu. V: 135/323, 28/323, 160/323.

11. Kastis on 4 valget, 5 punast ja 6 musta kuuli. Üksteise järel võetakse huupi välja 10 kuuli, kusjuures võetud kuule tagasi ei panda. Leidke tõenäosus, et väljavõetute hulgas on täpselt 3 valget, 4 punast ja 3 musta kuuli. V: 400/3003.

12. Üheksakorruselise maja lifti siseneb esimesel korrusel 4 inimest. Leidke tõenäosus, et nad kõik väljuvad samal korrusel. V: 1/ 512.

13. Kasutatud autode müügipunkti toodi Saksamaalt 20 pruugitud autot, neist 8 Audit, 7 Opelit ja 5 BMW-d. Esimese kuuga õnnestus neist maha müüa 17. Leidke tõenäosus, et allesjäänud 3 autot on 1) kõik ühte marki, 2) kõik erinevat marki. V: 101/1140, 14/57.

14. Karbis on 10 pooljuhti, neist 7 hiljuti testitut. Karbist võetakse huupi 5 pooljuhti. Leidke tõenäosus, et nende hulgas on täpselt 3 hiljuti testitut.
V: 5/12.

15. Urnis on 10 kuuli, neist 6 valget ja 4 musta. Urnist võetakse järgemööda 3 kuuli. Leidke tõenäosus, et nad kõik on valged, kui 1) võetud kuul pannakse urni tagasi, 2) ei panda tagasi. V: 27/125, 1/6.

16. Neli jalgpallurit sooritavad igaüks ühe karistuslöögi. Nende tabamise tõenäosused on vastavalt 0.8, 0.7, 0.6 ja 0.5. Leidke tõenäosus, et tabamuste koguarv on 1) täpselt 3, 2) ülimalt 3, 3) vähemalt 3. V: 0.394, 0.832, 0.562.

17. Eksamipiletis on 4 küsimust, üks igast osast. Esimeses osas on 4 küsimust, teises 6, kolmandas 8 ja neljandas 6. Üliõpilane teab iga esimese osa küsimust, kolme küsimust teisest osast, viit küsimust kolmandast osast ja nelja küsimust neljandast osast. Leidke tõenäosus, et üliõpilane teab võetud piletist 1) kõiki küsimusi, 2) täpselt kahte küsimust, 3) vähemalt kahte küsimust 4) ülimalt

kahte küsimust. V: 5/24, 7/24, 15/16, 17/48.
18. Üks kolmandik loteriipiletitest võidavad. Mitu piletit tuleb osta, et tõenäosusega, mis on suurem kui 0.9, vähemalt üks neist võidaks. V: 6.
19. Urnist, milles on n valget ja m musta kuuli ($n \geq 2$, $m \geq 2$), võetakse huupi 2 kuuli. Leidke tõenäosus, et 1) nende hulgas on täpselt k valget kuuli ($k = 0; 1; 2$), 2) nende hulgas on ülimalt üks valge kuul, 3) nende hulgas on vähemalt üks valge kuul. V: $(m(m-1)) / ((n+m)(n+m-1))$,
$2mn/ ((n+m)(n+m-1))$, $(n(n-1)) / ((n+m)(n+m-1))$,
$(m(2n+m-1)) / ((n+m)(n+m-1))$,
$(n(n+2m-1)) / ((n+m)(n+m-1))$.
20. Urnis on n valget ja m musta kuuli ($m \geq 2$, $n \geq 2$). Võetakse huupi 2 kuuli. Kumb sündmustest on tõenäosem, kas $A-$ kuulid on sama värvi või $B-$ kuulid on erinevat värvi? V: kui $(n-m)^2 > n+m$, siis on tõenäosem, et kuulid on sama värvi.
21. Vabariigi korvpalli esiliigas esineb 8 võistkonda, neist 3 üliõpilasvõistkonda. Moodustatakse kaks alagruppi, à 4 võistkonda. Leidke järgmiste sündmuste tõenäosused: $A-$ kõik üliõpilasvõistkonnad on ühes alaprupis; $B-$ ühes alagrupis on 1 ja teises 2 üliõpilasvõistkonda. V: 1/7, 6/7.
22. *Buffoni* probleem. Joonelisele paberile, joonte vahega m cm, kukub huupi nõel pikkusega l. Olgu $l < m$. Leidke tõenäosus, et nõel lõikub ühega joontest. V: $2l/(m\pi)$.
23. Täringut visatakse kolm korda. Leidke järgmiste sündmuste tõenäosused: 1) kolmandal viskel tuleb rohkem silmi, kui tuli esimesel ja kui tuli teisel viskel; 2) kolmandal viskel tuleb rohkem silmi kui kahel esimesel viskel kokku. V: 55/216, 5/54.
24. Riiulile pannakse 10 raamatut, millest 3 on ingliskeelsed, juhuslikus järjekorras. Kui suur on tõenäosus, et ingliskeelsed raamatud satuvad kõrvuti? V: 1/15.
25. Ringjoonel raadiusega R valitakse huupi kolm punkti A, B ja C. Kui suur on tõenäosus, et kolmnurk ABC on teravnurkne? Täisnurkne? Nürinurkne? V: 1/4, 0, 3/4.
26. Sündmuse toimumise tõenäosus on igal katsel 0.2. Katseid sooritatakse järgemööda kuni sündmuse toimumiseni. Leidke tõenäosus, et tuleb sooritada 1) (täpselt) 3 katset; 2) vähemalt 3 katset; 3) ülimalt 3 katset. V: 0.128, 0.64, 0.488.
27. Üksteist välistavad neli sündmust võivad toimuda katsel vastavalt tõenäosustega 0.012, 0.01, 0.006, 0.002. Leidke tõenäosus, et katsel toimub 1) täpselt üks neist sündmustest, 2) vähemalt üks neist sündmustest, 3) ülimalt üks neist sündmustest, 4) täpselt kaks neist sündmustest, 5) vähemalt kaks neist sündmustest, 6) ülimalt kaks neist sündmustest. V: 0.03, 0.03, 1, 0, 0, 1.
28. Kolm laskurit tulistavad märklauda igaüks ühe lasu. Märklaua tabamise tõenäosus on esimesel laskuril 0.75, teisel 0.8 ja kolmandal 0.9. Kui suur on tõenäosus, et 1) ükski laskureist ei taba märklauda; 2) täpselt üks tabab; 3) kõik tabavad; 4) vähemalt üks neist tabab; 5) täpselt kaks tabavad; 6) vähemalt kaks tabavad; 7) ülimalt kaks tabavad? V: 0.005, 0.08, 0.54, 0.995, 0.375, 0.915, 0.46.

29. Laskur tabab igal lasul kas kümnesse või üheksasse. Tõenäosus, et ta saab ühe lasuga 10 silma, on 0.7. Üheksa silma saamise tõenäosus on 0.3. Leidke tõenäosus, et laskur saab kolme lasuga: 1) 30 silma; 2) täpselt 29 silma; 3) vähemalt 28 silma. V: 0.343, 0.441, 0.973.
30. Tõenäosus, et üliõpilane sooritab kontrolltöö esimesel katsel, on 0.7. Teisel katsel on selle sooritamise tõenäosus 0.6. Kolmandal katsel on selle sooritamise tõenäosus 0.5. Leidke tõenäosus, et üliõpilane sooritab kontrolltöö, kui talle võimaldatakse selleks ülimalt kolm katset. V: 47/50. 31. Leidke tõenäosus, et 10 dioodi hulgas pole ühtki mittekorras dioodi, kui huupi võetud 5 dioodi olid korras. Eeldatakse, et mittekorras dioodide arv 10 dioodi hulgas võib olla võrdse tõenäosusega kas 0, 1 või 2. V: 18/31.
32. Tõenäosus, et üliõpilane jõuab õigeaegselt loengule, on 0.9. Leidke tõenäosus, et neljast üliõpilasest jõuab õigeaegselt loengule 1) täpselt 3, 2) vähemalt 3, 3) ülimalt 3. V: 0.2916, 0.9477, 0.3439.
33. Täringut visatakse kolm korda. Leida tõenäosus, et iga kord tuleb 1) sama silmade arv, 2) kolmest suurem silmade arv, 3) kolmest väiksem silmade arv. V: 1/36, 1/8, 1/27.
34. Riiulil on 2 karpi diskettidega. Esimeses karbis on 5 uut ja 4 kasutatud disketti, teises 3 uut ja 5 kasutatud disketti. Esimesest karbist võetakse huupi 2 disketti ja pannakse teise. Pärast seda võetakse teisest karbist 4 disketti. Leida tõenäosus, et kõik teisest karbist võetud disketid on uued. V: 1/108.
35. Karbis on 9 uut tennisepalli. Igaks mänguks võetakse karbist huupi 3 palli, mis pärast mängu pannakse karpi tagasi. Leida tõenäosus, et pärast kolmandat mängu on karbis vähemalt üks uus pall. V: $1759/1764 \approx 0.9972$.
36. Kaardipakist (52 lehte) võetakse huupi 4 lehte. Leidke tõenäosus, et: 1) kõik on eri masti; 2) täpselt 2 on punased; 3) kõik on mustad; 4) kõik on ässad. V: $2197/20\,825 \approx 0.1055$, $325/833 \approx .3902$, $46/833 \approx 0.0552$, $1/270\,725 \approx 3.694 \times 10^{-6}$.
37. Urnis on 2 valget ja 3 musta kuuli. Kaks poissi võtavad urnist kordamööda huupi kuuli kuni esimese valge kuuli saamiseni. Võetud kuule urni tagasi ei panda. Leidke tõenäosus, et esimesena saab valge kuuli see, kes 1) alustas, 2) ei alustanud. V: 3/5, 2/5.
38. Urnis on 2 valget ja 3 musta kuuli. Kaks poissi võtavad urnist kordamööda huupi kuuli kuni esimese valge kuuli saamiseni. Enne järgmise kuuli võtmist pannakse kuul urni tagasi. Leidke tõenäosus, et esimesena saab valge kuuli see, kes 1) alustas, 2) ei alustanud. V: 5/8, 3/8.
39. On kolm urni. Esimeses on 1 valge ja 3 musta kuuli. Teises on 3 valget ja 2 musta kuuli. Kolmandas on ainult valged kuulid. Huupi valitakse urn ja sellest võetakse huupi 1 kuul. Leidke tõenäosus, et 1) võetud kuul on valge, 2) kuul võeti teisest urnist, kui on teada, et võeti valge kuul. V: 37/60, 12/37.
40. On n urni, igas a valget ja b musta kuuli. Esimesest urnist võetakse huupi kuul ja pannakse teise urni. Pärast seda võetakse teisest urnist huupi kuul ja pannakse kolmandasse jne. Leidke tõenäosus, et n-ndast urnist võetakse valge kuul. V: $a/(a+b)$.
41. Karbis on 3 uut ja 4 kasutatud tennisepalli. Esimeseks mänguks võetakse karbist huupi 2 palli, mis pärast mängu pannakse karpi tagasi. Teiseks mänguks

võetakse huupi 2 palli. Leidke tõenäosus, et 1) teiseks mänguks võeti uued pallid, 2) esimeseks mänguks võeti kasutatud pallid, kui selgus, et teiseks mänguks võeti uued pallid. V: 10/147, 3/5.

42. On kaks karpi. Esimeses on 2 uut ja 3 kasutatud tennisepalli. Teises on 3 uut ja 4 kasutatud palli. Esimesest võetakse huupi üks pall ja pannakse teise karpi. Pärast seda võetakse teisest karbist huupi üks pall. Leidke tõenäosus, et 1) teisest karbist võetud pall on uus, 2) esimesest karbist võetud pall oli uus, kui on täiendavalt teada, et teisest karbist võetud pall oli kasutatud. V: 17/40, 8/23.

43. Laual on kolm eksamipiletit. Tudeng võtab neist ühe pileti. Igas piletis on 3 küsimust, mis ei kordu ülejäänud piletites, st kokku on laual 9 erinevat küsimust. Üliõpilane oskab neist kuut küsimust. Üliõpilane saab eksamil läbi, kui ta teab: 1) vähemalt kaht küsimust oma piletist; 2) täpselt üht küsimust oma piletist ja teab kaht lisaküsimust, mis antakse talle allesjäänud kahest piletist. Leidke tõenäosus, et 1) üliõpilane sooritab eksami, 2) üliõpilane pidi vastama lisaküsimustele, kui on teada, et ta sai eksamil läbi. V: 11/12, 12/77.

44. Üliõpilane läks eksamile, olles 20 küsimusest selgeks õppinud 16. Talle esitati 3 küsimust. Kui suur on tõenäosus, et ta 1) teadis kõiki 3 küsimust, 2) oskas vastata täpselt kahele küsimusele kolmest, 3) ei osanud ühtki küsimust kolmest, 4) oskas vähemalt ühte küsimust? V: 28/57, 8/19, 1/285, 284/285.

45. Tudeng arvab, et teab 75% eksamiküsimustest. Ta teab neist, mida arvab teadvat, vaid 80%. Eksam sooritatakse testina, kusjuures igale küsimusele antakse 5 võimalikku valikut, millest vaid 1 on õige. Kui tudeng ei tea küsimust, valib ta vastuse huupi. Leidke tõenäosus, et ta 1) vastab saadud küsimusele õigesti, 2) vastab juhuslikult õigesti. V: 17/25, 2/17.

46. Rühmas on 10 üliõpilast, kellest 3 teavad materjali väga hästi, 2 teavad hästi, 4 rahuldavalt ja 1 halvasti. Kokku on 10 erinevat küsimust. Väga hästi valmistunud üliõpilane teab kõiki kümmet küsimust, hästi valmistunud kaheksat, rahuldavalt valmistunud kuut ja halvasti valmistunud nelja küsimust. Leidke tõenäosus, et esimesena sisenenud üliõpilane 1) teab kõiki kolme talle esitatud küsimust, 2) on materjali väga hästi teadev üliõpilane, kui ta teadis kõiki kolme esitatud küsimust, 3) on hästi valmistunud üliõpilane, kui ta teadis kõiki kolme küsimust, 4) on rahuldavalt valmistunud üliõpilane, kui ta teadis kolme küsimust, 5) on halvasti valmistunud üliõpilane, kui ta teadis kolme küsimust. V: 139/300, 90/139, 28/139, 20/139, 1/139.

47. Tõenäosus, et üliõpilane jõuab õigeaegselt loengule, on $0,9$. Leidke tõenäosus, et 1) neljast üliõpilasest täpselt 2 jõuab õigeaegselt loengule, 2) neljast üliõpilasest vähemalt 2 jõuab õigeaegselt loengule, 3) neljast üliõpilasest ülimalt 2 jõuab õigeaegselt loengule. V: 0.0 486, 0.9963, 0.0523.

48. Kumb on tõenäosem, kas võita võrdvõimelist vastast 1) täpselt kolmes partiis neljast või viies partiis kaheksast, 2) vähemalt kolmes partiis neljast või vähemalt viies partiis kaheksast? V: täpselt kolmes partiis neljast, vähemalt viies partiis kaheksast. Viik on välistatud.

49. Seade koosneb kümnest sõlmest. Iga sõlme töökindlus aja T jaoks on p. Leidke tõenäosus, et aja T jooksul ütleb üles 1) vähemalt 1 sõlm, 2) ülimalt üks sõlm, 3) täpselt 1 sõlm, 4) täpselt 2 sõlme, 5) vähemalt 2 sõlme, 6) ülimalt kaks

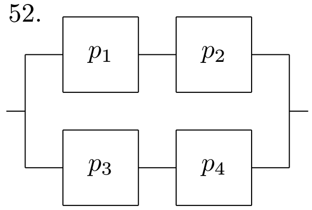
sõlme. V: $1-p^{10}$, $p^{10}+10p^9(1-p)$, $10p^9(1-p)$, $45\,(1-p)^2\,p^8$, $1+9p^{10}-10p^9$, $36p^{10}-80p^9+45p^8$.

50. Sooritatakse 3 lasku. Esimesel lasul on märklaua tabamise tõenäosus 0.2, teisel lasul 0.3 ja kolmandal 0.4. Leidke tõenäosus, et 1) vaid üks laskudest tabab, 2) ülimalt üks tabab, 3) vähemalt üks tabab, 4) ülimalt kaks tabab, 5) vähemalt kaks tabab. V: 0.452, 0.788, 0.664, 0.976, 0.212.

51. Seadmes on 6 kondensaatorit, millest üks läks rivist välja. Neid testitakse kordamööda kuni tõbise avastamiseni. Leidke tõenäosus, et tuli testida 1) täpselt kolm kondensaatorit, 2) rohkem kui kolm, 3) ülimalt kolm. V: 1/6, 1/3, 1/2.

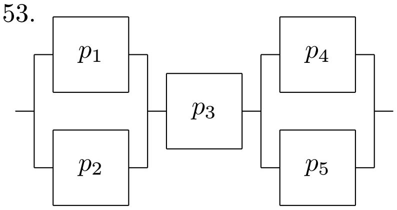
Ülesannetes 52-61 leidke skeemi töökindlus, kui elementide töökindlused on antud joonisel. Elemendid lähevad rivist välja üksteisest sõltumatult.

52.



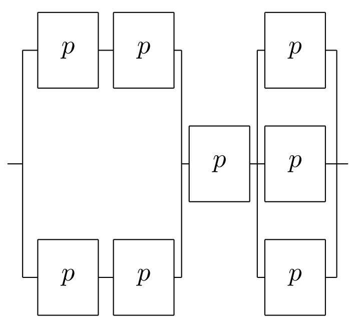
$V:\ p_1p_2+p_3p_4-p_1p_2p_3p_4.$

53.



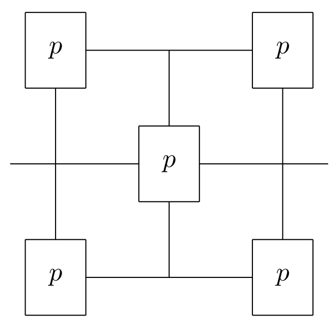
$V:\ (p_1+p_2-p_1p_2)p_3(p_4+p_5-p_4p_5).$

54.



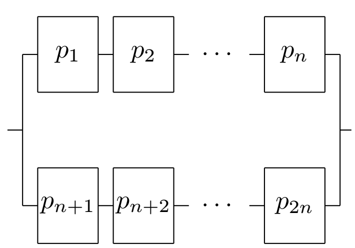
$V:\ p^4(2-p^2)(p^2-3p+3).$

55.



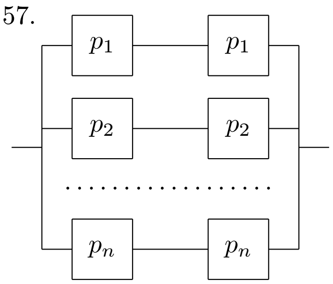
$V:\ p^5-p^4-2p^3+2p^2+p.$

56.

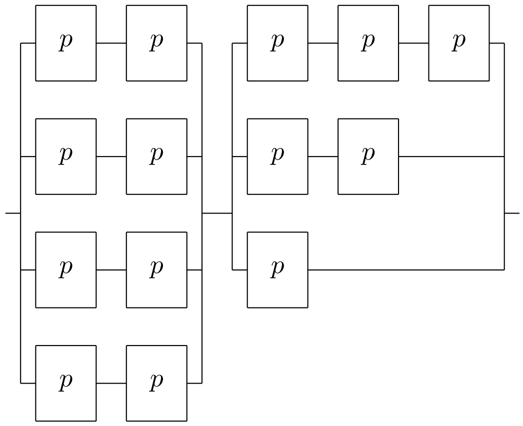


$V:\ \prod_{i=1}^n p_i+\prod_{i=n+1}^{2n} p_i-\prod_{i=1}^{2n} p_i.$

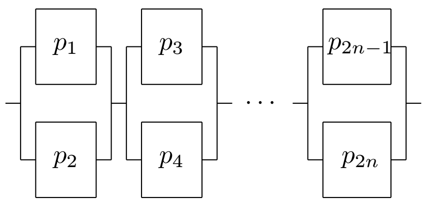
57.



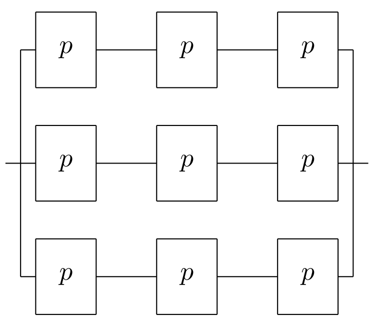
$V:\ 1-\prod_{i=1}^n(1-p_i^2).$

58. 

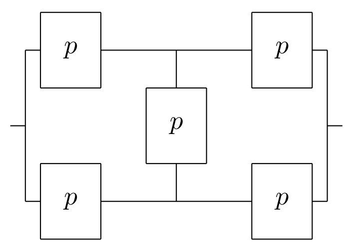
$V:\ (1-(1-p^2)^4)(1-(1-p^3)(1-p^2)(1-p))$.

59. 

$V:\ \prod_{i=1}^{n}(p_{2i-1}+p_{2i}-p_{2i-1}p_{2i})$.

60. 

$V:\ p^3(p^6-3p^3+3)$.

61. 

$V:\ 2p^5-5p^4+2p^3+2p^2$.

Peatükk 2

Juhuslikud suurused

2.1 Juhusliku suuruse mõiste. Jaotusfunktsioon

Sooritatakse katse.

Definitsioon 1. Suurust X, mille võrdumine katse käigus etteantud väärtusega x on juhuslik sündmus, nimetatakse *juhuslikuks suuruseks.*

Definitsioon 2. Väärtusi x, mida juhuslik suurus X võib katse käigus omandada, nimetatakse selle *juhusliku suuruse võimalikeks väärtusteks.*

Definitsioon 3. Hulka, mille elementide ja naturaalarvude hulga $\mathbf{N}$ elementide vahel saab korraldada üksühese vastavuse, nimetatakse *loenduvaks hulgaks.*

Ülesanne 1. Näidake, et loenduvad on kõigi kahega jaguvate naturaalarvude hulk $2\mathbf{N}$, kõigi täisarvude hulk $\mathbf{Z}$ ja kõigi ratsionaalarvude hulk $\mathbf{Q}$.

Osutub, et kõigi reaalarvude hulk $\mathbf{R}$ ei ole loenduv. Samuti ei ole loenduv lõigu $[a,b]$ punktide hulk.

Definitsioon 4. Juhuslikku suurust, mille võimalike väärtuste hulk on lõplik või loenduv, nimetatakse *diskreetseks juhuslikuks suuruseks.*

Definitsioon 5. Kui diskreetse juhusliku suuruse X korral on teada tema võimalike väärtuste hulk $\{x_i\}_{i\in I}$, kus I on lõplik või loenduv hulk, ja tõenäosused

$$p_i = P\left(X = x_i\right)_{i\in I} \quad \left(\sum_{i\in I} p_i = 1\right),$$

millega juhuslik suurus X iga neist võimalikest väärtustest omandab, siis öeldakse, et on antud diskreetse juhusliku suuruse X *jaotusseadus.*

Juhusliku suuruse X, mille võimalike väärtuste hulk $\{x_i\}_{i\in I}$ on lõplik, st $I = \{1; 2; \ldots; n\}$, ja $p_i = P\left(X = x_i\right)_{i\in I}$ $\left(\sum_{i=1}^{n} p_i = 1\right)$, jaotusseaduse saame esitada ka tabeli kujul

x_i	x_1	x_2	$\cdots$	x_n
p_i	p_1	p_2	$\cdots$	p_n

Definitsioon 6. Funktsiooni $F(x) = P(X < x)$ nimetatakse juhusliku suuruse X *jaotusfunktsiooniks.*

Seega juhusliku suuruse X jaotusfunktsioon $F(x)$ defineeritakse iga $x \in \mathbf{R}$ korral kui tõenäosus, et juhuslik suurus X omandab katse käigus väärtuse, mis on rangelt väiksem kui x. Olgu $F(-\infty) \overset{\text{def.}}{=} P(X < -\infty)$ ja $F(+\infty) \overset{\text{def.}}{=} P(X < +\infty)$.

Lause 1. Kui $F(x)$ on juhusliku suuruse X jaotusfunktsioon, siis:

1° $0 \le F(x) \le 1$; 2° $F(x) \nearrow$; 3° $F(-\infty) = 0$; 4° $F(+\infty) = 1$;

$$5^\circ\ P(\alpha \le X < \beta) = F(\beta) - F(\alpha) \quad (\alpha < \beta). \tag{2.1.1}$$

Tõestus. Et juhusliku suuruse X jaotusfunktsioon $F(x)$ defineeritakse kui tõenäosus, siis 1° on tõene. Kuna

$$P(X < \beta) \overset{\alpha\le\beta}{=} P((X < \alpha) + (\alpha \le X < \beta)) \overset{\text{liidetavad teineteist välistavad}}{=}$$
$$= P(X < \alpha) + P(\alpha \le X < \beta) \overset{P(\alpha\le X<\beta)\ge 0}{\Rightarrow}$$
$$\Rightarrow \begin{cases} P(\alpha \le X < \beta) = P(X < \beta) - P(X < \alpha) = F(\beta) - F(\alpha) \\ P(X < \alpha) \overset{\alpha<\beta}{\le} P(X < \beta) \ \Rightarrow \ F(x) \nearrow, \end{cases}$$

siis ka 2° ja 5° on tõesed. Et sündmus $X < -\infty$, st suurus X omandab katse käigus väärtuse, mis on väiksem kui $-\infty$, on võimatu, siis

$$F(-\infty) = P(X < -\infty) = P(V) = 0.$$

Kuna sündmus $X < +\infty$, st suurus X omandab katse käigus väärtuse, mis on väiksem kui $+\infty$, on kindel sündmus, siis

$$F(+\infty) = P(X < +\infty) = P(K) = 1.$$

Seega on tõesed ka 3° ja 4°. $\square$

Näide 1. Sooritatakse kaks vabaviset. Mõlemal viskel on tabamise tõenäosus 0.7. Visked on sõltumatud. Olgu X tabamuste koguarv selle kaheviskelise seeria korral. Kas suurus X on juhuslik? Kui jaa, siis kas see juhuslik suurus on diskreetne? Kui tegemist on diskreetse juhusliku suurusega, siis leiame juhusliku suuruse X jaotusseaduse ja jaotusfunktsiooni $F(x)$ ning selle graafiku.

Et sündmused $X = 0$, $X = 1$ ja $X = 2$ on juhuslikud, siis X on juhuslik suurus võimalike väärtustega 0, 1 ja 2. Seega $\{0; 1; 2\}$ on juhusliku suuruse X võimalike väärtuste hulk. Et võimalike väärtuste hulk on lõplik, siis on tegemist diskreetse juhusliku suurusega. Suuruse X jaotusseaduse kirjapanekuks leiame Bernoulli valemi ($n = 2,\ p = 0.7$) abil tõenäosused, millega X need võimalikud väärtused omandab:

$$p_0 = P(X = 0) = C_2^0 \cdot 0.7^0 \cdot 0.3^2 = 0.09,$$
$$p_1 = P(X = 1) = C_2^1 \cdot 0.7^1 \cdot 0.3^1 = 0.42,$$
$$p_2 = P(X = 2) = C_2^2 \cdot 0.7^2 \cdot 0.3^0 = 0.49.$$

Vormistame leitud jaotusseaduse tabelina

x_i	0	1	2
p_i	0.09	0.42	0.49

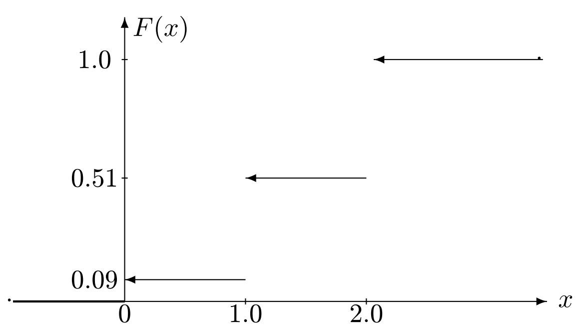
Et

$$
\begin{aligned}
&x \leq 0 \ \Rightarrow \ F(x) = P(X < x) = P(V) = 0,\\
&0 < x \leq 1 \ \Rightarrow \ F(x) = P(X < x) = P(X = 0) = 0.09,\\
&1 < x \leq 2 \ \Rightarrow \left\{ \begin{array}{c} F(x) = P(X < x) = P(X = 0) + P(X = 1) = \\ = 0.09 + 0.42 = 0.51, \end{array} \right.\\
&2 < x \ \Rightarrow \ \left\{ \begin{array}{c} F(x) = P(X < x) = P(X = 0) + P(X = 1) + P(X = 2) = \\ = 0.09 + 0.42 + 0.49 = 1, \end{array} \right.
\end{aligned}
$$

siis saame jaotusfunktsiooni $F(x)$ kuju

$$
F(x) = \left\{ \begin{array}{c} 0, \ \text{kui } x \leq 0, \\ 0.09, \ \text{kui } 0 < x \leq 1, \\ 0.51, \ \text{kui } 1 < x \leq 2, \\ 1, \ \text{kui } 2 < x. \end{array} \right.
$$

Skitseerime jaotusfunktsiooni $F(x)$ graafiku



Kui kasutada Heaviside'i funktsiooni

$$
\mathbf{1}(x) = \left\{ \begin{array}{c} 0, \ \text{kui } x \leq 0 \\ 1, \ \text{kui } x > 0, \end{array} \right.
$$

siis

$$
F(x) = 0.09 \cdot \mathbf{1}(x) + 0.42 \cdot \mathbf{1}(x-1) + 0.49 \cdot \mathbf{1}(x-2). \qquad \diamondsuit
$$

Kehtib järgmine väide.

Lause 2. Kui $\{x_k\}_{k \in I}$ on diskreetse juhusliku suuruse X võimalike väärtuste hulk ja $p_k = P(X = x_k)_{k \in I}$, kus $\sum_{k \in I} p_k = 1$, siis selle suuruse jaotusfunktsioon on esitatav kujul

$$
F(x) = \sum_{k \in I} p_k \cdot \mathbf{1}(x - x_k). \tag{2.1.2}
$$

Kontrollige, et kehtib järgmine väide.

Lause 3. Diskreetse juhusliku suuruse X jaotusfunktsioon $F(x)$ on igas punktis $x \in \mathbf{R}$ pidev vasakult.

Definitsioon 7. Juhuslikku suurust, mille jaotusfunktsioon on pidev hulgal $\mathbf{R}$, nimetatakse *pidevaks juhuslikuks suuruseks.*

Järeldus 1. Pidev juhuslik suurus omandab iga väärtuse tõenäosusega 0.

Tõestus. Kui X on pidev juhuslik suurus, siis

$$\begin{aligned}
P(X = x) &= \lim_{\Delta x \to 0+} P(x \le X < x + \Delta x) \overset{(2.1.1)}{=} \\
&= \lim_{\Delta x \to 0+} (F(x + \Delta x) - F(x)) = \\
&= \lim_{\Delta x \to 0+} F(x + \Delta x) - \lim_{\Delta x \to 0+} F(x) \overset{F(x) \text{ on pidev punktis } x}{=} \\
&= F(x) - F(x) = 0. \qquad \square
\end{aligned}$$

Järeldus 2. Kui X on pidev juhuslik suurus ja $\alpha < \beta$, siis

$$P(\alpha \le X < \beta) = P(\alpha < X \le \beta) = P(\alpha \le X \le \beta) = P(\alpha < X < \beta). \tag{2.1.3}$$

Tõestus. Et

$$P(\alpha \le X \le \beta) \ge P(\alpha \le X < \beta) \ge P(\alpha < X < \beta)$$

ja

$$P(\alpha \le X \le \beta) \ge P(\alpha < X \le \beta) \ge P(\alpha < X < \beta)$$

ning

$$\begin{aligned}
P(\alpha \le X \le \beta) &= P((X = \alpha) + (\alpha < X < \beta) + (X = \beta)) = \\
&= P(X = \alpha) + P(\alpha < X < \beta) + P(X = \beta) = \\
&= 0 + P(\alpha < X < \beta) + 0 = P(\alpha < X < \beta),
\end{aligned}$$

siis Järelduse 2 väide kehtib. $\square$

Järgnevas on pideva juhusliku suuruse võimalike väärtuste hulgaks tavaliselt lõik, vahemik või poollõik, kusjuures vahemik ja poollõik võivad olla ka lõpmatud.

Definitsioon 8. Juhuslikku suurust, millel on nii pideva kui ka diskreetse juhusliku suuruse omadusi, nimetatakse *segatüüpi juhuslikuks suuruseks.*

Definitsioon 9. Öeldakse, et juhuslik suurus X allub *binoomjaotusele* parameetritega n ja p $(0 < p < 1)$, kui $\{0; 1; 2; \ldots; n\}$ on selle juhusliku suuruse võimalike väärtuste hulk ja

$$P(X = k) = C_n^k p^k q^{n-k} \quad (k \in \{0; 1; 2; \ldots; n\}),$$

kusjuures $q = 1 - p$.

Lause 4. Kui juhuslik suurus X allub binoomjaotusele parameetritega n ja p, siis selle juhusliku suuruse jaotusfunktsioon on esitatav kujul

$$F(x) = \sum_{k=0}^{n} C_n^k p^k q^{n-k} \cdot \mathbf{1}\,(x-k)\,. \tag{2.1.4}$$

Paketis SWP saab kasutada funktsiooni

$$\text{BinomialDist}\,(m; n, p) \stackrel{\text{def.}}{=} \sum_{k=0}^{m} C_n^k p^k q^{n-k} \quad (m = 0; 1; \ldots ; n)\,.$$

Näide 2. Täringut visatakse 1000 korda. Olgu juhuslikuks suuruseks X kordade arv, mille korral saadakse 6 silma. Millisele jaotusele allub suurus X? Leiame suuruse X jaotusfunktsiooni $F(x)$. Avaldame jaotusfunktsiooni abil tõenäosuse, et kuute koguarv selle tuhandese seeria korral kuulub lõiku $[150; 200]$.

Kuute koguarvu X kui juhusliku suuruse võimalike väärtuse hulk on $\{0; 1; 2; \ldots ; 1000\}$, kusjuures Bernoulli valemi põhjal

$$P\,(X = k) = C_{1000}^k \left(\frac{1}{6}\right)^k \left(\frac{5}{6}\right)^{1000-k} \qquad (k \in \{0; 1; 2; \ldots ; 1000\})\,.$$

Seega allub X binoomjaotusele parameetritega 1000 ja 1/6. Valemi (2.1.2) abil saame

$$F(x) = \sum_{k=0}^{1000} C_{1000}^k \left(\frac{1}{6}\right)^k \left(\frac{5}{6}\right)^{1000-k} \cdot \mathbf{1}\,(x-k)\,.$$

Seega

$$P\,(150 \le X \le 200) = P\,(150 \le X < 201) \stackrel{(2.1.1)}{=} F(201) - F(150) =$$

$$= \sum_{k=0}^{1000} C_{1000}^k \left(\frac{1}{6}\right)^k \left(\frac{5}{6}\right)^{1000-k} \cdot (\mathbf{1}\,(201-k) - \mathbf{1}\,(150-k))\,. \quad \diamondsuit$$

Definitsioon 10. Öeldakse, et juhuslik suurus X allub *Poissoni jaotusele* parameetriga λ $(\lambda > 0)$, kui $\mathbf{N}_0 \stackrel{\text{def.}}{=} \mathbf{N} \cup \{0\}$ on selle juhusliku suuruse võimalike väärtuste hulk ja

$$P\,(X = k) = \frac{\lambda^k}{k!} e^{-\lambda} \quad (k \in \mathbf{N}_0)\,.$$

Et

$$(\lambda > 0) \ \wedge \ (k \in \mathbf{N}_0) \ \Rightarrow \ \frac{\lambda^k}{k!} e^{-\lambda} \ge 0$$

ja

$$\sum_{k=0}^{\infty} P\,(X = k) = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{-\lambda} e^{\lambda} = 1,$$

siis Definitsiooni 5 tingimused on täidetud.

Lause 5. Kui juhuslik suurus X allub Poissoni jaotusele parameetriga λ, siis selle juhusliku suuruse jaotusfunktsioon on esitatav kujul

$$F(x) \overset{(2.1.2)}{=} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} \cdot \mathbf{1}\left(x-k\right). \tag{2.1.5}$$

Paketis SWP saab kasutada funktsiooni

$$\mathrm{PoissonDist}(m;\lambda) \overset{\text{def.}}{=} \sum_{k=0}^{m} \frac{\lambda^k}{k!} e^{-\lambda} \ \left(m \in \mathbf{N}_0\right).$$

Näide 3. Juhuslik suurus X allub Poissoni jaotusele parameetriga $\lambda = 0.2$. Leiame selle juhusliku suuruse jaotusfunktsiooni $F(x)$. Leiame tõenäosuse, et X omandab katse käigus väärtuse, mis on suurem kui 2.

Valemi (2.1.5) abil saame

$$F(x) = \sum\nolimits_{k=0}^{\infty} \frac{0.2^k}{k!} e^{-0.2} \cdot \mathbf{1}\left(x-k\right).$$

Et

$$(X > 2) = \overline{(X \leq 2)},$$

siis

$$\begin{aligned} P\left(X>2\right) &= 1 - P\left(X \leq 2\right) = 1 - P\left((X=0)+(X=1)+(X=2)\right) = \\ &= 1 - \frac{0.2^0}{0!}e^{-0.2} - \frac{0.2^1}{1!}e^{-0.2} - \frac{0.2^2}{2!}e^{-0.2} = 1 - 1.22e^{-0.2} \approx \\ &\approx 0.0011485. \qquad \diamondsuit \end{aligned}$$

2.2 Juhusliku suuruse jaotustihedus.

Definitsioon 1. Funktsiooni

$$f(x) = \lim_{\Delta x \to 0+} \frac{P\left(x \leq X < x+\Delta x\right)}{\Delta x} \quad \left(x \in \mathbf{R}\right) \tag{2.2.1}$$

nimetatakse juhusliku suuruse X *jaotustiheduseks*.

Märkus 1. Pideva juhusliku suuruse X korral võime Järelduse 2.1.2 põhjal Definitsioonis 1 suuruse $P\left(x \leq X < x+\Delta x\right)$ asendada ühega suurustest $P\left(x \leq X \leq x+\Delta x\right)$, $P\left(x < X \leq x+\Delta x\right)$ või $P\left(x < X < x+\Delta x\right)$.

Seega on suuruse X jaotustihedus $f(x)$ piirväärtus suuruse X poollõiku $[x, x+\Delta x)$ sattumise tõenäosuse ja selle poollõigu pikkuse suhtest.

Lause 1. Kehtivad järgmised seosed

$$f(x) = F'(x), \tag{2.2.2}$$

$$f(x) \geq 0, \tag{2.2.3}$$

$$F(x) = \int_{-\infty}^{x} f(t)dt, \tag{2.2.4}$$

$$\int_{-\infty}^{+\infty} f(x)dx = 1, \tag{2.2.5}$$

$$P(\alpha \le X < \beta) = \int_{\alpha}^{\beta} f(x)dx \quad (\alpha < \beta). \tag{2.2.6}$$

Tõestus. Seos (2.2.2) järeldub võrduste ahelast

$$f(x) \overset{\text{Def. 1}}{=} \lim_{\Delta x \to 0+} \frac{P(x \le X < x + \Delta x)}{\Delta x} \overset{(2.1.1)}{=}$$
$$= \lim_{\Delta x \to 0+} \frac{F(x + \Delta x) - F(x)}{\Delta x} \overset{\text{tuletise def.}}{=} F'(x).$$

Kuna seoses (2.2.1) esinevas murrus on nii lugeja kui ka nimetaja mittenegatiivsed, siis on seda ka jagatis ja selle piirväärtus ning seos (2.2.3) on tõene. Et funktsiooni $f(x)$ üheks algfunktsiooniks on $F(x)$ ja funktsiooni algfunktsiooniks on määratud integraal ülemise raja funktsioonina $\int_{-\infty}^{x} f(t)dt$, siis

$$F(x) = \int_{-\infty}^{x} f(t)dt + C,$$

kusjuures konstandi C määrame tingimustest $F(-\infty) = 0$. Seega $C = 0$ ja kehtib (2.2.4). Et $F(+\infty) = 1$, siis seosest (2.2.4) järeldub seos (2.2.5). Kuna $\alpha < \beta$ korral

$$P(\alpha \le X < \beta) \overset{(2.1.1)}{=} F(\beta) - F(\alpha) \overset{(2.2.4)}{=} \int_{-\infty}^{\beta} f(t)dt - \int_{-\infty}^{\alpha} f(t)dt =$$

$$= \int_{-\infty}^{\alpha} f(t)dt + \int_{\alpha}^{\beta} f(t)dt - \int_{-\infty}^{\alpha} f(t)dt = \int_{\alpha}^{\beta} f(t)dt,$$

siis kehtib seos (2.2.6). $\square$

Suvalist hulgal $\mathbf{R}$ määratud funktsiooni $f(x)$, mis rahuldab tingimusi (2.2.3) ja (2.2.5), võime käsitleda kui mingi juhusliku suuruse X jaotustihedust.

Definitsioon 2. Öeldakse, et juhuslik suurus X allub lõigul $[a, b]$ *ühtlasele jaotusele,* kui selle juhusliku suuruse jaotustihedus avaldub kujul

$$f(x) = \begin{cases} \dfrac{1}{b-a}, \text{ kui } x \in [a, b], \\ 0, \text{ kui } x \notin [a, b]. \end{cases} \tag{2.2.7}$$

Seostest (2.2.4) ja (2.2.7) järeldub, et

$$F(x) = \int_{-\infty}^{x} f(t)dt = \begin{cases} 0, \text{ kui } x < a, \\ \dfrac{x-a}{b-a}, \text{ kui } a \le x \le b, \\ 1, \text{ kui } x > b. \end{cases} \tag{2.2.8}$$

Paketis SWP on nende funktsioonide tähistuseks vastavalt UniformDen $(x; a, b)$ ja UniformDist$(x; a, b)$. Kontrollige, et seosega (2.2.7) määratud funktsioon $f(x)$ rahuldab tingimusi (2.2.3) ja (2.2.5). Kontrollige, et saadud $F(x)$ on pidev funktsioon. Seega on lõigul $[a, b]$ ühtlasele jaotusele alluv X Definitsiooni 2.1.7 põhjal pidev juhuslik suurus. Veenduge, et ka vahemikus (a, b) või poollõikudes $[a, b)$ ja $(a, b]$ ühtlasele jaotusele alluva juhusliku suuruse X jaotusfunktsioon on kujul (2.2.8).

Näide 1. Juhuslik suurus X allub lõigul $[-1; 2]$ ühtlasele jaotusele. Leiame $f(x)$ ja $F(x)$ nii analüütiliselt kui ka graafiliselt. Leiame $P(X \in [-2; 1.3])$.

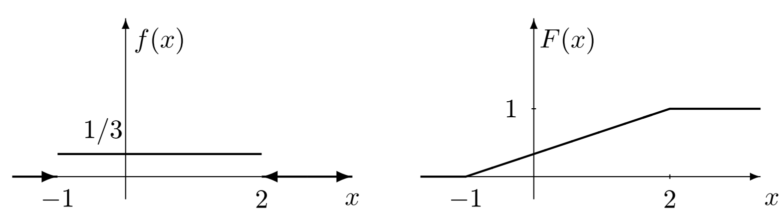
Valemi (2.2.7) abil saame

$$f(x) = \begin{cases} \dfrac{1}{2-(-1)} = \dfrac{1}{3}, \text{ kui } x \in [-1; 2], \\ 0, \text{ kui } x \notin [-1; 2]. \end{cases}$$

Jaotusfunktsiooni $F(x)$ leiame valemi (2.2.8) abil

$$F(x) = \int_{-\infty}^{x} f(t)dt = \begin{cases} \int_{-\infty}^{x} 0 \cdot dt = 0, \text{ kui } x < -1, \\ \int_{-\infty}^{-1} 0 \cdot dt + \int_{-1}^{x} \dfrac{1}{3} dt = \dfrac{x+1}{3}, \text{ kui } -1 \le x \le 2, \\ \int_{-\infty}^{-1} 0 \cdot dt + \int_{-1}^{3} \dfrac{1}{3} dt + \int_{3}^{x} 0 \cdot dt = 1, \text{ kui } x > 2. \end{cases}$$

Skitseerime funktsioonide $f(x)$ ja $F(x)$ graafikud



Leiame

$$P(X \in [-2; 1.3]) = F(1.3) - F(-2) = \frac{1.3+1}{3} - 0 = \frac{23}{30}. \qquad \diamondsuit$$

Definitsioon 3. Öeldakse, et juhuslik suurus X allub *normaaljaotusele* parameetritega a ja σ $(\sigma > 0)$, kui selle juhusliku suuruse jaotustihedus avaldub kujul

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\dfrac{(x-a)^2}{2\sigma^2}} \quad (x \in \mathbf{R}). \tag{2.2.9}$$

Kontrollime, et seosega (2.2.9) määratud funktsioon $f(x)$ rahuldab tingimusi (2.2.3) ja (2.2.5). Et $\sigma > 0$ ja eksponentfunktsiooni väärtused on mittenegatiivsed hulgal $\mathbf{R}$, siis tingimus (2.2.3) on rahuldatud. Kuna

$$\int_{-\infty}^{+\infty} f(x)dx = \int_{-\infty}^{+\infty} \frac{e^{-(x-a)^2/\left(2\sigma^2\right)}}{\sigma\sqrt{2\pi}}dx =$$
$$= \left[\begin{array}{c} t = (x-a)/\sigma,\ x = a+\sigma t,\ dx = \sigma dt, \\ \\ x = -\infty \overset{\sigma>0}{\leftrightarrow} t = -\infty,\ x = +\infty \overset{\sigma>0}{\leftrightarrow} t = +\infty \end{array} \right] =$$
$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-t^2/2}dt = \frac{2}{\sqrt{2\pi}} \int_{0}^{+\infty} e^{-t^2/2}dt$$

ja

$$\int_0^{+\infty} e^{-t^2/2}dt = \sqrt{\int_0^{+\infty} e^{-x^2/2}dx \int_0^{+\infty} e^{-y^2/2}dy} = \sqrt{\int_0^{+\infty}\int_0^{+\infty} e^{-(x^2+y^2)/2}dxdy} =$$
$$= \left[\begin{array}{c} \text{läheme üle polaar-} \\ \text{koordinaatidesse} \end{array} \right] = \sqrt{\int_0^{\pi/2} d\varphi \int_0^{+\infty} \rho e^{-\rho^2/2}d\rho} =$$
$$= \sqrt{\frac{\pi}{2} \lim_{A\to+\infty} \left(-e^{-\rho^2/2}\right)\Big|_0^A} = \sqrt{\frac{\pi}{2} \lim_{A\to+\infty} \left(-e^{-A^2/2}+1\right)} = \sqrt{\frac{\pi}{2}},$$

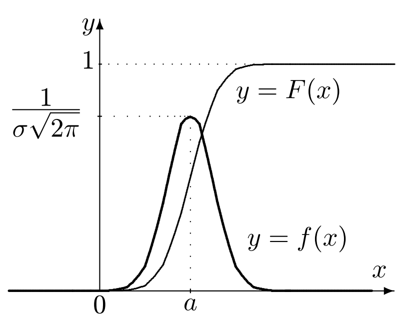
siis on rahuldatud ka tingimus (2.2.5).

Asjaolu, et juhuslik suurus X allub normaaljaotusele parameetritega a ja σ, tähistame lühidalt $X \sim N(a,\sigma)$. Kui $X \sim N(0;1)$, siis kõneldakse *standardsest normaaljaotusest* ehk *tsentreeritud ja normeeritud normaaljaotusest.*

Kui juhuslik suurus X allub normaaljaotusele parameetritega a ja σ, siis selle juhusliku suuruse jaotusfunktsioon on

$$F(x) = \int_{-\infty}^{x} \frac{1}{\sigma\sqrt{2\pi}} e^{-(t-a)^2/\left(2\sigma^2\right)}dt. \tag{2.2.10}$$

Skitseerime $X \sim N(a,\sigma)$ jaotustiheduse $f(x)$ ja jaotusfunktsiooni $F(x)$ graafikud vastavalt jämeda ja peene joonega



Definitsioon 4. *Deltafunktsiooniks* ehk *nullindat järku Diraci impulssfunktsiooniks* nimetatakse funktsiooni

$$\delta(x) \overset{\text{def.}}{=} \lim_{h \to 0+} \frac{\mathbf{1}(x) - \mathbf{1}(x-h)}{h} = \mathbf{1}'(x). \tag{2.2.11}$$

Seega $\delta(x) = \mathbf{1}'(x)$ ja

$$\delta(x) = \begin{cases} +\infty, & \text{kui } x = 0, \\ 0, & \text{kui } x \neq 0. \end{cases} \tag{2.2.12}$$

Järelikult $\delta(x)$ ei ole funktsioon tavalises mõttes. Tegemist on *üldistatud funktsiooniga* ehk *distributsiooniga.* Seosest (2.2.12) ei piisa deltafunktsiooni määramiseks. Kui $\varphi(x)$ on pidev funktsioon punkti a mingis ümbruses, siis formaalselt kehtib seos

$$\int_{-\infty}^{+\infty} \varphi(x)\delta(x-a)dx = \begin{bmatrix} \text{kasutame} \\ \text{Definitsiooni 4} \end{bmatrix} =$$

$$= \int_{-\infty}^{+\infty} \varphi(x) \lim_{h \to 0+} \frac{\mathbf{1}(x-a) - \mathbf{1}(x-a-h)}{h} dx =$$

$$= \begin{bmatrix} \text{vahetame formaalselt integreerimise ja} \\ \text{piirväärtuse võtmise järjekorra} \end{bmatrix} =$$

$$= \lim_{h \to 0+} \int_{-\infty}^{+\infty} \varphi(x) \frac{\mathbf{1}(x-a) - \mathbf{1}(x-a-h)}{h} dx =$$

$$= \lim_{h \to 0+} \frac{1}{h} \int_{a}^{a+h} \varphi(x)dx = \begin{bmatrix} \text{kasutame integraali keskväärtus-} \\ \text{teoreemi pideva } \varphi(x) \text{ korral} \end{bmatrix} =$$

$$= \lim_{h \to 0+} \frac{1}{h} \varphi(a+\theta h)h = \lim_{h \to 0+} \varphi(a+\theta h) = \begin{bmatrix} \varphi(x) \text{ on pidev} \\ \text{punktis } a \end{bmatrix} = \varphi(a).$$

Seega iga funktsiooni $\varphi(x)$ korral, mis on pidev punkti a mingis ümbruses, kehtib seos

$$\int_{-\infty}^{+\infty} \varphi(x)\delta(x-a)dx = \varphi(a). \tag{2.2.13}$$

Formaalselt võib defineerida deltafunktsiooni ka seose (2.2.13) abil, nõudes selle seose täidetust suvalise pideva funktsiooni $\varphi(x)$ korral. Deltafunktsiooni korral kasutatakse tihti formaalseid seoseid

$$\delta(-x) = \delta(x); \ \delta(cx) = |c|^{-1} \delta(x) \quad (c = \text{konstant})\,; \ x\delta(x) = 0,$$

mille tegelik sisu avaldub vaid (2.2.13) korral. Valides $\varphi(x) = 1$, saame seosest (2.2.13) väite

$$\int_{-\infty}^{+\infty} \delta(x-a)dx = 1,$$

millest juhul $a = 0$ järeldub, et

$$\int_{-\infty}^{+\infty} \delta(x)dx = 1.$$

Seega $\delta(x)$ rahuldab formaalselt tingimusi (2.2.3) ja (2.2.5), st tegemist on jaotustihedusega. Kuna $\mathbf{1}'(x) = \delta(x)$, siis võime deltafunktsiooni kasutada diskreetse juhusliku suuruse jaotustiheduse kirjapanekuks.

Lause 2. Kui $\{x_k\}_{k \in I}$ on diskreetse juhusliku suuruse X võimalike väärtuste hulk ja $p_k = P\left(X = x_k\right)_{k \in I}$, siis selle suuruse jaotustihedus on esitatav kujul

$$f(x) = \sum_{k \in I} p_k \cdot \delta\left(x - x_k\right). \tag{2.2.14}$$

Tõestus järeldub Lausest 2.1.2 ja seosest (2.2.2) ning (2.2.11). □

Järeldus 1. Kui juhuslik suurus X allub binoomjaotusele parameetritega n ja p, siis selle juhusliku suuruse jaotustihedus on esitatav kujul

$$f(x) = \sum_{k=0}^{n} C_n^k p^k q^{n-k} \cdot \delta\left(x - k\right). \tag{2.2.15}$$

Tõestus järeldub Lausetest 2.1.4 ja 2. □

Järeldus 2. Kui juhuslik suurus X allub Poissoni jaotusele parameetriga λ, siis selle juhusliku suuruse jaotustihedus on esitatav kujul

$$f(x) = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} \cdot \delta\left(x - k\right). \tag{2.2.16}$$

Tõestus järeldub Lausetest 2.1.5 ja 2. □

Analoogiliselt funktsiooniga $\delta(x)$ võib defineerida selle funktsiooni tuletised $\delta^{(k)}(x)$ (*k-ndat järku Diraci impulssfunktsioonid)*

$$\int_{-\infty}^{+\infty} \varphi(x)\delta^{(k)}(x-a)dx = (-1)^k \varphi^{(k)}(a),$$

kus $\varphi(x)$ on suvaline funktsioon, mis omab pidevaid tuletisi kuni järguni k.

2.3 Juhusliku suuruse keskväärtus

Definitsioon 1. Kindlat suurust

$$\mathrm{E}X \stackrel{\text{def.}}{=} \int_{-\infty}^{+\infty} x\, f(x)dx \tag{2.3.1}$$

nimetatakse juhusliku suuruse X *keskväärtuseks.*

Seega juhusliku suuruse X keskväärtus $\mathrm{E}X$ kui kindel suurus on arv.

Märkus 1. Kuna

$$\exists \int_{-\infty}^{+\infty} f(x)dx \;\not\Rightarrow\; \exists \int_{-\infty}^{+\infty} x\, f(x)dx,$$

siis mõningatel juhuslikel suurustel ei eksisteeri keskväärtust.

Näide 1. Leiame juhusliku suuruse X, mis allub lõigul $[a;b]$ ühtlasele jaotusele, keskväärtuse.

Selle juhusliku suuruse jaotustihedus on määratud seosega (2.2.7). Seega saame vastavalt Definitsioonile 1

$$\begin{aligned}\mathrm{E}X &= \int_{-\infty}^{+\infty} x\, f(x)dx = \int_a^b \frac{x}{b-a}dx = \left.\frac{x^2}{2\,(b-a)}\right|_a^b = \\ &= \frac{b^2-a^2}{2\,(b-a)} = \frac{a+b}{2}. \qquad \diamondsuit\end{aligned}$$

Näide 2. Leiame juhusliku suuruse $X \sim N(a,\sigma)$ keskväärtuse.

Selle juhusliku suuruse jaotustihedus on määratud seosega (2.2.9). Seega saame vastavalt Definitsioonile 1

$$\begin{aligned}\mathrm{E}X &= \int_{-\infty}^{+\infty} x\, f(x)dx = \int_{-\infty}^{+\infty} \frac{x}{\sigma\sqrt{2\pi}} e^{-(x-a)^2/\left(2\sigma^2\right)}dx = \\ &= \left[\begin{array}{c} t = \dfrac{x-a}{\sigma},\ x = a+\sigma t,\ dx = \sigma dt, \\ x = -\infty \overset{\sigma>0}{\leftrightarrow} t = -\infty,\ x = +\infty \overset{\sigma>0}{\leftrightarrow} t = +\infty \end{array}\right] = \\ &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} (a+\sigma t)\, e^{-t^2/2}dt = \\ &= a\cdot\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{-t^2/2}dt + \sigma\cdot\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} te^{-t^2/2}dt.\end{aligned}$$

Et $\exp\left(-t^2/2\right)/\sqrt{2\pi}$ on suuruse $T \sim N(0;1)$ jaotustihedus, siis (2.2.5) põhjal summa esimene liidetav on a. Kuna $t\exp\left(-t^2/2\right)$ on paaritu funktsioon ja rajad on sümmeetrilised nullpunkti suhtes, siis saadud summa teine liidetav on 0. Seega $\mathrm{E}X = a$. $\diamondsuit$

Lause 1. Kui $\{x_k\}_{k\in I}$ on diskreetse juhusliku suuruse X võimalike väärtuste hulk ja $p_k = P\left(X = x_k\right)_{k\in I}$, siis

$$\mathrm{E}X = \sum_{k\in I} x_k p_k. \tag{2.3.2}$$

Tõestus. Seose (2.2.14) ja Definitsiooni 1 põhjal saame

$$\mathrm{E}X = \int_{-\infty}^{+\infty} xf(x)dx = \int_{-\infty}^{+\infty} x\sum_{k\in I} p_k\cdot\delta\left(x-x_k\right)dx =$$

$$= \sum_{k\in I} p_k \int_{-\infty}^{+\infty} x\cdot\delta\left(x-x_k\right)dx =$$
$$= \left[\begin{array}{c}\text{rakendame seost (2.2.13) valiku}\\ \varphi(x)=x \text{ ja } a=x_k \text{ korral}\end{array}\right] = \sum_{k\in I} p_k\cdot x_k. \qquad \square$$

Näide 3. Leiame juhusliku suuruse X, mis allub Poissoni jaotusele parameetriga λ, keskväärtuse.

Poissoni jaotusele (vt Definitsiooni 2.1.10), mille parameeter on λ $(\lambda>0)$, alluva juhusliku suuruse korral

$$x_k = k,\ p_k = P\left(X=k\right) = \frac{\lambda^k}{k!}e^{-\lambda} \quad \left(k\in\mathbf{N}_0\right),$$

Seega seose (2.3.2) abil saame

$$\mathrm{E}X = \sum_{k=0}^{\infty} k\cdot\frac{\lambda^k}{k!}e^{-\lambda} = \sum_{k=1}^{\infty} k\cdot\frac{\lambda^k}{k!}e^{-\lambda} = \lambda e^{-\lambda}\sum_{k=1}^{\infty}\frac{\lambda^{k-1}}{(k-1)!} =$$
$$= \left[\begin{array}{c} m=k-1,\\ k=m+1\end{array}\right] = \lambda e^{-\lambda}\sum_{m=0}^{\infty}\frac{\lambda^m}{m!} = \lambda e^{-\lambda}e^{\lambda} = \lambda. \qquad \diamondsuit$$

Definitsioon 2. Kaht juhuslikku suurust nimetatakse *sõltumatuteks,* kui ühe jaotus ei sõltu sellest, millise võimaliku väärtuse omandab katse käigus teine suurus.

Lause 2. Juhusliku suuruse X keskväärtusel $\mathrm{E}X$ on järgmised omadused:

1° $X = C$ (C on kindel suurus) $\Rightarrow$ $\mathrm{E}C = C$;
2° $\mathrm{E}\left(X+Y\right) = \mathrm{E}X + \mathrm{E}Y$;
3° $\mathrm{E}\left(X\cdot Y\right) \overset{X \text{ ja } Y \text{ on sõltumatud}}{=} \mathrm{E}X\cdot\mathrm{E}Y$;
4° $\mathrm{E}\left(C\cdot X\right) = C\cdot\mathrm{E}X$.

Tõestus. Kindlat suurust C võime käsitleda kui diskreetse juhusliku suuruse, millel on vaid üks võimalik väärtus C, erijuhtu. Seega saame $X = C$ jaoks jaotusseaduse

x_k	C
p_k	1

Seose (2.3.2) abil leiame

$$\mathrm{E}C = C\cdot 1 = C.$$

Kuigi omadus 2° kehtib suvalise juhuslike suuruste paari korral, piirdume tehnilistel kaalutlustel tõestusega vaid diskreetsete lõpliku arvu võimalike väärtustega X ja Y korral. Olgu X ja Y jaotusseadused antud tabelite abil

x_i	x_1	x_2	$\ldots$	x_n
p'_i	p'_1	p'_2	$\ldots$	p'_n

y_j	y_1	y_2	$\ldots$	y_m
p''_j	p''_1	p''_2	$\ldots$	p''_m

.

Kui $Z = X + Y$, siis ka Z on diskreetne juhuslik suurus, kusjuures

$$\{x_i + y_j\}_{i=1;2;\ldots;n \,\wedge\, j=1;2;\ldots;m}$$

on suuruse Z võimalike väärtuste hulk. Märgime, et nii kirjapandud suuruse Z võimalike väärtuste hulgas on mõningad selle hulga elemendid kirja pandud mitu korda, st erinevate paaride x_i ja y_j summa võib anda sama suuruse Z võimaliku väärtuse. Leiame, et

$$P\left(Z = x_i + y_j\right) = P\left(\left(X = x_i\right)\left(Y = y_j\right)\right).$$

Seose (2.3.2) abil saame

$$\begin{aligned}\mathrm{E}Z &= \sum_{i=1}^{n}\sum_{j=1}^{m}\left(x_i + y_j\right) P\left(Z = x_i + y_j\right) = \\ &= \sum_{i=1}^{n}\sum_{j=1}^{m}\left(x_i + y_j\right) P\left(\left(X = x_i\right)\left(Y = y_j\right)\right) =\end{aligned}$$

$$\begin{aligned}
&= \sum_{i=1}^{n}\sum_{j=1}^{m} x_i P\left(\left(X = x_i\right)\left(Y = y_j\right)\right) + \sum_{i=1}^{n}\sum_{j=1}^{m} y_j P\left(\left(X = x_i\right)\left(Y = y_j\right)\right) = \\
&= \sum_{i=1}^{n} x_i \sum_{j=1}^{m} P\left(\left(X = x_i\right)\left(Y = y_j\right)\right) + \sum_{j=1}^{m} y_j \sum_{i=1}^{n} P\left(\left(X = x_i\right)\left(Y = y_j\right)\right) = \\
&= \left[\begin{array}{c}\text{sündmused } \left(X = x_i\right)\left(Y = y_j\right) \text{ on erinevate} \\ \text{indeksipaaride } (i,j) \text{ korral teineteist välistavad}\end{array}\right] = \\
&= \sum_{i=1}^{n} x_i P\left(\sum_{j=1}^{m}\left(X = x_i\right)\left(Y = y_j\right)\right) + \sum_{j=1}^{m} y_j P\left(\sum_{i=1}^{n}\left(X = x_i\right)\left(Y = y_j\right)\right) = \\
&= [\text{kasutame distributiivsuse omadust}] = \\
&= \sum_{i=1}^{n} x_i P\left(\left(X = x_i\right)\sum_{j=1}^{m}\left(Y = y_j\right)\right) + \sum_{j=1}^{m} y_j P\left(\left(Y = y_j\right)\sum_{i=1}^{n}\left(X = x_i\right)\right) = \\
&= \left[\sum_{j=1}^{m}\left(Y = y_j\right) \text{ ja } \sum_{i=1}^{n}\left(X = x_i\right) \text{ on kindlad sündmused}\right] = \\
&= \sum_{i=1}^{n} x_i P\left(\left(X = x_i\right)\cdot K\right) + \sum_{j=1}^{m} y_j P\left(\left(Y = y_j\right)\cdot K\right) = \\
&= \sum_{i=1}^{n} x_i P\left(X = x_i\right) + \sum_{j=1}^{m} y_j P\left(Y = y_j\right) = \\
&= \sum_{i=1}^{n} x_i p_i' + \sum_{j=1}^{m} y_j p_j'' = \mathrm{E}X + \mathrm{E}Y.
\end{aligned}$$

Tõestame omaduse 3° diskreetsete lõpliku arvu võimalike väärtustega X ja Y korral. Olgu juhuslike suuruste X ja Y jaotusseadused antud eelnevalt esitatud tabelite abil. Kui $Z = X \cdot Y$, siis ka Z on diskreetne juhuslik suurus, kusjuures

$$\{x_i \cdot y_j\}_{i=1;2;\ldots;n \,\wedge\, j=1;2;\ldots;m}$$

on suuruse Z võimalike väärtuste hulk ja

$$P(Z = x_i \cdot y_j) = P((X = x_i)(Y = y_j))$$

ning

$$\begin{aligned}
\mathrm{E}Z &= \sum_{i=1}^{n}\sum_{j=1}^{m} (x_i \cdot y_j)\, P(Z = x_i \cdot y_j) = \\
&= \sum_{i=1}^{n}\sum_{j=1}^{m} x_i y_j P((X = x_i)(Y = y_j)) \overset{X \text{ ja } Y \text{ on sõltumatud}}{=} \\
&= \sum_{i=1}^{n}\sum_{j=1}^{m} x_i y_j P(X = x_i)\, P(Y = y_j) = \\
&= \left(\sum_{i=1}^{n} x_i P(X = x_i)\right)\left(\sum_{j=1}^{m} y_j P(Y = y_j)\right) = \\
&= \left(\sum_{i=1}^{n} x_i p_i'\right)\left(\sum_{j=1}^{m} y_j p_j''\right) = \mathrm{E}X \cdot \mathrm{E}Y.
\end{aligned}$$

Omadus 4° on järeldus omadusest 3°, sest kindel suurus C ja juhuslik suurus X on alati sõltumatud. Miks? □

Näide 4. Leiame binoomjaotusele, mille parameetrid on n ja p, alluva juhusliku suuruse X keskväärtuse.

Et Definitsiooni 2.1.9 põhjal on $x_k = k \quad (k \in \{0; 1; 2; \ldots; n\})$ ja

$$p_k = P(X = k) = C_n^k p^k q^{n-k},$$

siis Lause 1 vahetul kasutamisel saame tulemuseks

$$\mathrm{E}X = \sum_{k \in I} p_k \cdot x_k = \sum_{k=0}^{n} k C_n^k p^k q^{n-k}.$$

Kuna viimase avaldise lihtsustamine on keerukas, siis esitame juhusliku suuruse X kujul

$$X = X_1 + X_2 + \ldots + X_n, \tag{2.3.3}$$

kusjuures X_i $(i = 1; 2; \ldots ; n)$ on sündmuse A esinemiste arv i-ndal katsel. Suurused X_i alluvad binoomjaotusele, mille parameetrid on 1 ja p ning mille jaotusseadused on kujul

x_k	1	0
p_k	p	q

Lause 1 abil saame

$$\mathrm{E}X_i = 1 \cdot p + 0 \cdot q = p \quad (i = 1; 2; \ldots ; n) .$$

Seega

$$\mathrm{E}X = \mathrm{E}\left(X_1 + X_2 + \ldots + X_n\right) = \left[\begin{array}{c} \text{rakendame Lause 2 omaduse } 2^\circ \\ \text{üldistust } n \text{ liidetava jaoks} \end{array} \right] =$$
$$= \mathrm{E}X_1 + \mathrm{E}X_2 + \ldots + \mathrm{E}X_n = np. \qquad \diamondsuit$$

Üldistame keskväärtuse mõistet suvalise juhusliku argumendiga reaalse või kompleksse funktsiooni jaoks.

Definitsioon 3. Arvu

$$\mathrm{E}\left(h\left(X\right)\right) \overset{\text{def.}}{=} \int_{-\infty}^{+\infty} h\left(x\right) f(x) dx \tag{2.3.4}$$

nimetatakse *juhusliku argumendiga X funktsiooni $h\left(X\right)$ keskväärtuseks.*

Näide 5. Allugu juhuslik suurus X ühtlasele jaotusele lõigul $[a, b]$. Leiame juhusliku suuruse $Y = 2X - 1$ keskväärtuse.

Vastavalt Definitsioonile 3 saame

$$\mathrm{E}Y = \mathrm{E}\left(2X - 1\right) = \int_a^b \left(2x - 1\right) \frac{1}{b-a} dx = \left. \frac{x^2 - x}{b-a} \right|_a^b =$$
$$= \frac{b^2 - b}{b-a} - \frac{a^2 - a}{b-a} = \frac{b^2 - a^2 - b + a}{b-a} =$$
$$= \frac{\left(b-a\right)\left(b+a-1\right)}{b-a} = a + b - 1. \qquad \diamondsuit$$

Lause 3. Kui $\{x_k\}_{k \in I}$ on diskreetse juhusliku suuruse X võimalike väärtuste hulk ja $p_k = P\left(X = x_k\right)_{k \in I}$ ning $Y = h\left(X\right)$, kusjuures $h(x)$ on pidev punktide x_k $(k \in I)$ mingis ümbruses, siis

$$\mathrm{E}Y = \mathrm{E}\left(h\left(X\right)\right) = \sum_{k \in I} h\left(x_k\right) p_k. \tag{2.3.5}$$

Tõestus. Et $f(x) \overset{(2.2.14)}{=} \sum_{k \in I} p_k \cdot \delta\left(x - x_k\right)$, siis juhusliku suuruse Y korral

saame

$$
\begin{aligned}
\mathrm{E}Y &= \mathrm{E}\left(h\left(X\right)\right) \overset{(2.3.4)}{=} \int_{-\infty}^{+\infty} h\left(x\right) \sum_{k \in I} p_k \cdot \delta\left(x - x_k\right) dx = \\
&= \sum_{k \in I} p_k \int_{-\infty}^{+\infty} h\left(x\right) \cdot \delta\left(x - x_k\right) dx = \\
&= \left[\begin{array}{c} \text{rakendame seost (2.2.13) valiku} \\ \varphi(x) = h\left(x\right) \text{ ja } a = x_k \text{ korral} \end{array} \right] = \\
&= \sum_{k \in I} h\left(x_k\right) p_k. \qquad \square
\end{aligned}
$$

2.4 Dispersioon

Definitsioon 1. Arvu

$$
\mathrm{D}X \overset{\text{def.}}{=} \mathrm{E}\left(X - \mathrm{E}X\right)^2 \tag{2.4.1}
$$

nimetatakse juhusliku suuruse X *dispersiooniks.*

Definitsioon 2. Arvu

$$
\sigma_X \overset{\text{def.}}{=} \sqrt{\mathrm{D}X} \tag{2.4.2}
$$

nimetatakse juhusliku suuruse X *standardhälbeks.*

Kui juhusliku suuruse X keskväärtus $\mathrm{E}X$ kujutab endast selle suuruse võimalike väärtuste kaalutud keskmist, siis nii $\mathrm{D}X$ kui ka σ_X on juhusliku suuruse X hajuvuse mõõdud.

Olgu $f(x)$ juhusliku suuruse X jaotustihedus. Seoste (2.4.1) ja (2.3.4) abil saame

$$
\mathrm{D}X = \int_{-\infty}^{+\infty} \left(x - \mathrm{E}X\right)^2 f(x) dx. \tag{2.4.3}
$$

Kui X on diskreetne juhuslik suurus ja $\{x_k\}_{k \in I}$ on selle suuruse võimalike väärtuste hulk ning $p_k = P\left(X = x_k\right)_{k \in I}$, siis seoste (2.4.3) ja (2.2.14) abil saame

$$
\begin{aligned}
\mathrm{D}X &= \int_{-\infty}^{+\infty} \left(x - \mathrm{E}X\right)^2 \sum_{k \in I} p_k \cdot \delta\left(x - x_k\right) dx = \\
&= \sum_{k \in I} p_k \int_{-\infty}^{+\infty} \left(x - \mathrm{E}X\right)^2 \cdot \delta\left(x - x_k\right) dx = \\
&= \left[\begin{array}{c} \text{rakendame seost (2.2.13) valiku} \\ \varphi(x) = \left(x - \mathrm{E}X\right)^2 \text{ ja } a = x_k \text{ korral} \end{array} \right] = \\
&= \sum_{k \in I} p_k \cdot \left(x_k - \mathrm{E}X\right)^2.
\end{aligned}
$$

Lause 1. Kui X on diskreetne juhuslik suurus ja $\{x_k\}_{k\in I}$ on selle suuruse võimalike väärtuste hulk ning $p_k = P\left(X = x_k\right)_{k\in I}$, siis

$$\mathrm{D}X = \sum_{k\in I} \left(x_k - \mathrm{E}X\right)^2 p_k. \tag{2.4.4}$$

Lause 2. Juhusliku suuruse X dispersioonil $\mathrm{D}X$ on järgmised omadused:

1° $\mathrm{D}C = 0$ (C – kindel suurus);
2° $\mathrm{D}\left(CX\right) = C^2 \cdot \mathrm{D}X$ (C – kindel suurus);
3° $\mathrm{D}X = \mathrm{E}\left(X^2\right) - \left(\mathrm{E}X\right)^2 \overset{\text{lühidalt}}{=} \mathrm{E}X^2 - \left(\mathrm{E}X\right)^2$;
4° $\mathrm{D}\left(X+Y\right) = \mathrm{D}X + \mathrm{D}Y + 2\mathrm{E}\left(\left(X - \mathrm{E}X\right)\left(Y - \mathrm{E}Y\right)\right)$;
5° $\mathrm{D}\left(X+Y\right) \overset{X \text{ ja } Y \text{ on sõltumatud}}{=} \mathrm{D}\left(X\right) + \mathrm{D}\left(Y\right)$.

Tõestus. Seosest (2.4.1) ja keskväärtuse omadustest, vt Lauset 2.3.2, järeldub

$$\begin{aligned}
\mathrm{D}C &= \mathrm{E}\left(C - \mathrm{E}C\right)^2 = \mathrm{E}\left(C - C\right)^2 = \mathrm{E}0 = 0,\\
\mathrm{D}\left(C\cdot X\right) &= \mathrm{E}\left(C\cdot X - \mathrm{E}\left(C\cdot X\right)\right)^2 = \mathrm{E}\left(C\cdot X - C\cdot \mathrm{E}X\right)^2 =\\
&= \mathrm{E}\left(C\cdot\left(X - \mathrm{E}X\right)\right)^2 = C^2\cdot \mathrm{E}\left(X - \mathrm{E}X\right)^2 = C^2\cdot \mathrm{D}X,\\
\mathrm{D}X &= \mathrm{E}\left(X - \mathrm{E}X\right)^2 = \mathrm{E}\left(X^2 - 2\cdot X\cdot \mathrm{E}X + \left(\mathrm{E}X\right)^2\right) =\\
&= \mathrm{E}X^2 - \mathrm{E}\left(2\cdot X\cdot \mathrm{E}X\right) + \mathrm{E}\left(\mathrm{E}X\right)^2 =\\
&= \mathrm{E}X^2 - 2\cdot \mathrm{E}X\cdot \mathrm{E}X + \left(\mathrm{E}X\right)^2 = \mathrm{E}X^2 - \left(\mathrm{E}X\right)^2
\end{aligned}$$

ja

$$\begin{aligned}
\mathrm{D}\left(X+Y\right) &= \mathrm{E}\left(\left(X+Y\right) - \mathrm{E}\left(X+Y\right)\right)^2 = \mathrm{E}\left(\left(X - \mathrm{E}X\right) + \left(Y - \mathrm{E}Y\right)\right)^2 =\\
&= \mathrm{E}\left(\left(X - \mathrm{E}X\right)^2 + 2\left(X - \mathrm{E}X\right)\left(Y - \mathrm{E}Y\right) + \left(Y - \mathrm{E}Y\right)\right)^2 =\\
&= \mathrm{E}\left(X - \mathrm{E}X\right)^2 + \mathrm{E}\left(2\left(X - \mathrm{E}X\right)\left(Y - \mathrm{E}Y\right)\right) + \mathrm{E}\left(Y - \mathrm{E}Y\right)^2 =\\
&= \mathrm{D}X + 2\mathrm{E}\left(\left(X - \mathrm{E}X\right)\left(Y - \mathrm{E}Y\right)\right) + \mathrm{D}Y.
\end{aligned}$$

Kuna

$$\begin{aligned}
\mathrm{E}\left(\left(X - \mathrm{E}X\right)\left(Y - \mathrm{E}Y\right)\right) &= \left[\begin{array}{c} X \text{ ja } Y \text{ on sõltumatud} \Rightarrow \\ \Rightarrow X - \mathrm{E}X \text{ ja } Y - \mathrm{E}Y \text{ on sõltumatud} \end{array}\right] =\\
&= \mathrm{E}\left(X - \mathrm{E}X\right)\cdot \mathrm{E}\left(Y - EY\right) =\\
&= \left(\mathrm{E}X - \mathrm{E}\left(\mathrm{E}X\right)\right)\left(\mathrm{E}Y - \mathrm{E}\left(EY\right)\right) =\\
&= \left(\mathrm{E}X - \mathrm{E}X\right)\left(\mathrm{E}Y - \mathrm{E}Y\right) = 0,
\end{aligned}$$

siis omadusest 4° järeldub omadus 5°. □

Näide 1. Leiame lõigul $[a, b]$ ühtlasele jaotusele alluva juhusliku suuruse dispersiooni ja standardhälbe.

Näites 2.3.1 leidsime, et $\mathrm{E}X = (a+b)/2$. Leiame

$$\mathrm{E}X^2 = \int_{-\infty}^{+\infty} x^2 f(x)dx \overset{(2.2.7)}{=} \int_a^b x^2 \frac{1}{b-a} dx = \left.\frac{x^3}{3(b-a)}\right|_a^b =$$
$$= \frac{b^3 - a^3}{3(b-a)} = \frac{(b-a)\left(a^2+ab+b^2\right)}{3(b-a)} = \frac{a^2+ab+b^2}{3}.$$

Kasutame Lause 2 kolmandat väidet

$$\mathrm{D}X = \mathrm{E}\left(X^2\right) - (\mathrm{E}X)^2 = \frac{a^2+ab+b^2}{3} - \left(\frac{a+b}{2}\right)^2 =$$
$$= \frac{4a^2+4ab+4b^2-3a^2-6ab-3b^2}{12} = \frac{(b-a)^2}{12},$$

st

$$\mathrm{D}X = \frac{(b-a)^2}{12}, \quad \sigma_X = (b-a)/\left(2\sqrt{3}\right). \qquad \diamondsuit$$

Näide 2. Leiame juhusliku suuruse X, mis allub binoomjaotusele parameetritega n ja p, dispersiooni ja standardhälbe.

Kasutame juhusliku suuruse X esitust Näites 2.3.4 vaadeldud kujul

$$X = X_1 + X_2 + \ldots + X_n,$$

kus sõltumatud juhuslikud suurused X_i alluvad binoomjaotusele, mille parameetrid on 1 ja p. Seejuures $\mathrm{E}X_k = p$. Valiku $h(x) = x^2$ korral saame Lause 2.3.3 abil

$$\mathrm{E}\left(X_k^2\right) = \sum_{k \in I} x_k^2 p_k = 1^2 \cdot p + 0^2 \cdot q = p.$$

Lause 2 omaduse 3° abil leiame

$$\mathrm{D}X_k = \mathrm{E}X_k^2 - (\mathrm{E}X_k)^2 = p - p^2 = p(1-q) = pq.$$

Kasutades suuruste X_k sõltumatust ja Lause 2 omaduse 5° üldistust n liidetava jaoks, saame

$$\mathrm{D}X = \mathrm{D}(X_1 + X_2 + \ldots + X_n) =$$
$$= \mathrm{D}X_1 + \mathrm{D}X_2 + \ldots + \mathrm{D}X_n = npq,$$

st

$$\mathrm{D}X = npq, \quad \sigma_X = \sqrt{npq} \qquad \diamondsuit.$$

Näide 3. Leiame juhusliku suuruse X, mis allub Poissoni jaotusele parameetriga λ, dispersiooni ja standardhälbe.

Näites 2.3.3 leidsime, et Poissoni jaotusele parameetriga λ alluva juhusliku suuruse X korral $\mathrm{E}X = \lambda$. Valiku $h(x) = x^2$ korral saame Lause 2.3.3 abil

$$
\begin{aligned}
\mathrm{E}X^2 &= \sum_{k=0}^{\infty} k^2 \frac{\lambda^k}{k!} e^{-\lambda} = \sum_{k=1}^{\infty} k \frac{\lambda^k}{(k-1)!} e^{-\lambda} = [m = k-1,\ k = m+1] = \\
&= \sum_{m=0}^{\infty} (m+1) \frac{\lambda^{m+1}}{m!} e^{-\lambda} = \sum_{m=0}^{\infty} m \frac{\lambda^{m+1}}{m!} e^{-\lambda} + \sum_{m=0}^{\infty} \frac{\lambda^{m+1}}{m!} e^{-\lambda} = \\
&= \lambda^2 \sum_{m=1}^{\infty} \frac{\lambda^{m-1}}{(m-1)!} e^{-\lambda} + \lambda \sum_{m=0}^{\infty} \frac{\lambda^m}{m!} e^{-\lambda} = \left[\begin{array}{c} \text{esimese rea jaoks} \\ \nu = m-1,\ m = \nu+1 \end{array} \right] = \\
&= \lambda^2 e^{-\lambda} \sum_{\nu=0}^{\infty} \frac{\lambda^\nu}{\nu!} + \lambda e^{-\lambda} \sum_{m=0}^{\infty} \frac{\lambda^m}{m!} = \left[\sum_{m=0}^{\infty} \frac{\lambda^m}{m!} = e^{\lambda} \right] = \lambda^2 + \lambda.
\end{aligned}
$$

Lause 2 omaduse 3° põhjal saame

$$
\mathrm{D}X = \mathrm{E}X^2 - (\mathrm{E}X)^2 = \lambda^2 + \lambda - \lambda^2 = \lambda,
$$

st

$$
\mathrm{D}X = \lambda, \quad \sigma_X = \sqrt{\lambda}. \qquad \diamondsuit
$$

Näide 4. Leiame juhusliku suuruse $X \sim N(a, \sigma)$ dispersiooni ja standardhälbe.

Et

$$
f(x) \overset{(2.2.9)}{=} \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-a)^2/\left(2\sigma^2\right)} \quad (x \in \mathbf{R}),
$$

siis valiku $h(x) = x^2$ korral saame valemi (2.3.4) abil

$$
\mathrm{E}X^2 = \int_{-\infty}^{+\infty} x^2 f(x)dx = \int_{-\infty}^{+\infty} \frac{x^2}{\sigma\sqrt{2\pi}} e^{-(x-a)^2/\left(2\sigma^2\right)} dx =
$$

$$
\begin{aligned}
&= \left[\begin{array}{c} t = \dfrac{x-a}{\sigma},\ x = a + \sigma t,\ dx = \sigma dt, \\ x = -\infty \overset{\sigma > 0}{\leftrightarrow} t = -\infty,\ x = +\infty \overset{\sigma > 0}{\leftrightarrow} t = +\infty \end{array} \right] = \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} (a + \sigma t)^2 e^{-t^2/2} dt = \\
&= \frac{a^2}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-t^2/2} dt + \frac{2a\sigma}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} t e^{-t^2/2} dt + \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} t^2 e^{-t^2/2} dt.
\end{aligned}
$$

Et $\exp\left(-t^2/2\right)/\sqrt{2\pi}$ on suuruse $T \sim N(0;1)$ jaotustihedus, siis seose (2.2.5) abil saame, et saadud summa esimene liidetav on a^2. Kuna $t \exp\left(-t^2/2\right)$ on paaritu funktsioon ja rajad on sümmeetrilised nullpunkti suhtes, siis saadud summa teine liidetav on 0. Et

$$\int_0^{+\infty} t^2 e^{-t^2/2} dt = \left[\begin{array}{c} du = te^{-t^2/2}dt,\ u = -e^{-t^2/2} \\ v = t, \qquad dv = dt \end{array} \right] =$$

$$= -\lim_{A\to+\infty} te^{-t^2/2}\Big|_0^A + \int\limits_0^{+\infty} e^{-t^2/2} dt = 0 + \frac{\sqrt{2\pi}}{2},$$

siis saadud summa kolmas liidetav on σ^2. Seega

$$\mathrm{E}X^2 = a^2 + 0 + \sigma^2 = a^2 + \sigma^2$$

ja

$$\mathrm{D}X = \mathrm{E}X^2 - (\mathrm{E}X)^2 = a^2 + \sigma^2 - a^2 = \sigma^2 \ \Rightarrow\ \mathrm{D}X = \sigma^2,$$

st

$$\sigma_X = \sigma. \qquad \diamondsuit \tag{2.4.5}$$

2.5 Juhusliku suuruse momendid ja teised arvkarakteristikud

Juhusliku suuruse iseloomustamiseks kasutatakse teatud kindlaid suurusi, arve, mida nimetatakse juhuslike suuruste *arvkarakteristikuteks.* Kaht neist, keskväärtust ja dispersiooni, uurisime eelnevalt.

Definitsioon 1. Arvkarakteristikut (arvu)

$$\nu_n = \mathrm{E}\left(X^n\right) \quad (n \in \mathbf{N}) \tag{2.5.1}$$

nimetatakse juhusliku suuruse X *$n-$järku algmomendiks.*

Et vältida probleemi 0^0, defineeritakse täiendavalt $\nu_0 \overset{\text{def.}}{=} 1$. Kui vaatluse all on mitu juhuslikku suurust, siis segaduste vältimiseks kasutame suuruse X korral ν_n asemel tähistust $\nu_n(X)$.

Definitsioon 2. Arvu

$$\mu_n = \mathrm{E}\left(X - \mathrm{E}X\right)^n \quad (n \in \mathbf{N}) \tag{2.5.2}$$

nimetatakse juhusliku suuruse X *n-järku kesk-* ehk *tsentraalmomendiks.*

Olgu täiendavalt $\mu_0 \overset{\text{def.}}{=} 1$. Mitme juhusliku suuruse vaatlemisel kasutame suuruse X korral μ_n asemel tähistust $\mu_n(X)$. Kui $f(x)$ on juhusliku suuruse X jaotustihedus, siis seoste (2.5.1), (2.5.2) ja (2.3.4) abil saame

$$\nu_n = \int_{-\infty}^{+\infty} x^n f(x) dx \tag{2.5.3}$$

ning

$$\mu_n = \int_{-\infty}^{+\infty} (x - \mathrm{E}X)^n f(x) dx. \tag{2.5.4}$$

Lause 1. Kui X on diskreetne juhuslik suurus ja $\{x_k\}_{k\in I}$ on selle suuruse võimalike väärtuste hulk ning $p_k = P\left(X = x_k\right)_{k\in I}$, siis

$$\nu_n = \sum_{k\in I} x_k^n \cdot p_k \tag{2.5.5}$$

ja

$$\mu_n = \sum_{k\in I} \left(x_k - \mathrm{E}X\right)^n \cdot p_k. \tag{2.5.6}$$

Tõestus. Seoste (2.5.3) ja (2.5.4) abil saame vastavalt

$$\begin{aligned}
\nu_n = \mathrm{E}\left(X^n\right) &= \int_{-\infty}^{+\infty} x^n \sum_{k\in I} p_k \cdot \delta\left(x - x_k\right) dx = \\
&= \sum_{k\in I} p_k \int_{-\infty}^{+\infty} x^n \cdot \delta\left(x - x_k\right) dx = \\
&= \left[\begin{array}{c} \text{rakendame seost (2.2.13) valiku} \\ \varphi(x) = x^n \text{ ja } a = x_k \text{ korral} \end{array} \right] = \\
&= \sum_{k\in I} p_k \cdot x_k^n = \sum_{k\in I} x_k^n \cdot p_k
\end{aligned}$$

ja

$$\begin{aligned}
\mu_n = \mathrm{E}\left(X - \mathrm{E}X\right)^n &= \int_{-\infty}^{+\infty} \left(x - \mathrm{E}X\right)^n \sum_{k\in I} p_k \cdot \delta\left(x - x_k\right) dx = \\
&= \sum_{k\in I} p_k \int_{-\infty}^{+\infty} \left(x - \mathrm{E}X\right)^n \cdot \delta\left(x - x_k\right) dx = \\
&= \left[\begin{array}{c} \text{rakendame seost (2.2.13) valiku} \\ \varphi(x) = \left(x - \mathrm{E}X\right)^n \text{ ja } a = x_k \text{ korral} \end{array} \right] = \\
&= \sum_{k\in I} p_k \cdot \left(x_k - \mathrm{E}X\right)^n = \sum_{k\in I} \left(x_k - \mathrm{E}X\right)^n \cdot p_k,
\end{aligned}$$

st kehtivad seosed (2.5.5) ja (2.5.6). $\square$

Lause 2. Juhusliku suuruse kesk- ja algmomentide vahel on seos

$$\mu_n = \sum_{k=0}^{n} C_n^k \left(-1\right)^{n-k} \nu_1^{n-k} \nu_k. \tag{2.5.7}$$

Tõestus. Saame järgmise võrduste ahela

$$\begin{aligned}
\mu_n &= \mathrm{E}\left(X - \mathrm{E}X\right)^n = \mathrm{E}\left(\textstyle\sum_{k=0}^{n} C_n^k X^k \left(-\mathrm{E}X\right)^{n-k}\right) = \\
&= \textstyle\sum_{k=0}^{n} \mathrm{E}\left(C_n^k X^k \left(-1\right)^{n-k} \left(\mathrm{E}X\right)^{n-k}\right) = \textstyle\sum_{k=0}^{n} C_n^k \left(-1\right)^{n-k} \left(\mathrm{E}X\right)^{n-k} \mathrm{E}\left(X^k\right) = \\
&= \sum_{k=0}^{n} C_n^k \left(-1\right)^{n-k} \nu_1^{n-k} \nu_k. \qquad \square
\end{aligned}$$

Seose (2.5.7) erijuht $n = 2$ korral on eelnevalt tuttav Lausest 2.4.2:

$$\begin{aligned}\mu_2 &= \sum_{k=0}^{2} C_2^k (-1)^{2-k} \nu_1^{2-k} \nu_k = \\ &= C_2^0 (-1)^{2-0} \nu_1^{2-0} \nu_0 + C_2^1 (-1)^{2-1} \nu_1^{2-1} \nu_1 + C_2^2 (-1)^{2-2} \nu_1^{2-2} \nu_2 = \\ &= \nu_1^2 - 2\nu_1^2 + \nu_2 = \nu_2 - (\nu_1)^2 ,\end{aligned}$$

st

$$\mu_2 = \nu_2 - (\nu_1)^2 . \tag{2.5.8}$$

Näidake, et $n = 3$ ja $n = 4$ korral saame seosest (2.5.7) vastavalt

$$\mu_3 = \nu_3 - 3\nu_1\nu_2 + 2\nu_1^3 \tag{2.5.9}$$

ja

$$\mu_4 = \nu_4 - 4\nu_1\nu_3 + 6\nu_1^2\nu_2 - 3\nu_1^4. \tag{2.5.10}$$

Näide 1. Avaldame juhusliku suuruse X, mille keskväärtus on a, algmomendid keskmomentide kaudu.

Saame

$$\begin{aligned}\nu_n &= \mathrm{E}(X^n) = \mathrm{E}(((X - \mathrm{E}X) + \mathrm{E}X)^n) = \mathrm{E}\left(\textstyle\sum_{k=0}^{n} C_n^k (X - \mathrm{E}X)^k (\mathrm{E}X)^{n-k}\right) \\ &= \textstyle\sum_{k=0}^{n} C_n^k a^{n-k} \mathrm{E}(X - \mathrm{E}X)^k = \sum_{k=0}^{n} C_n^k a^{n-k} \mu_k,\end{aligned}$$

kusjuures $\mu_0 = 1$. $\diamondsuit$

Definitsioon 3. Arvu x_p, mis on määratud tingimusega

$$P(X < x_p) = p \quad (0 < p < 1), \tag{2.5.11}$$

nimetatakse pideva juhusliku suuruse X *p-kvantiiliks.*

Definitsioon 4. Juhusliku suuruse X 0.5-kvantiili nimetatakse selle suuruse *mediaaniks.*

Juhusliku suuruse mediaani tähistatakse sümboliga $\mathrm{Me}\, X$. Seega

$$\mathrm{Me}\, X = x_{0.5}.$$

Millised probleemid tekiks p-kvantiili määramisel diskreetse juhusliku suuruse korral?

Definitsioon 5. Juhusliku suuruse X jaotust nimetatakse *sümmeetriliseks,* kui iga $x \in \mathbf{R}$ korral

$$P(X < \mathrm{Me}\, X - x) = P(X > \mathrm{Me}\, X + x). \tag{2.5.12}$$

Tingimus (2.5.12) on pideva juhusliku suuruse X korral samaväärne tingimustega

$$F(\mathrm{Me}\, X - x) = 1 - F(\mathrm{Me}\, X + x)$$

ja

$$f\left(\operatorname{Me} X-x\right)=f\left(\operatorname{Me} X+x\right). \tag{2.5.13}$$

Definitsioon 6. Juhusliku suuruse X kvantiile $x_{0.25}$ ja $x_{0.75}$ nimetatakse vastavalt selle suuruse *alumiseks* ja *ülemiseks kvartiiliks.*

Definitsioon 7. Diskreetse juhusliku suuruse *moodiks* nimetatakse selle suuruse sellist võimalikku väärtust, mille omandamise tõenäosus on suurim.

Definitsioon 8. Pideva juhusliku suuruse *moodiks* nimetatakse selle suuruse sellist võimalikku väärtust, milles selle suuruse jaotustihedusel on lokaalne maksimum.

Juhusliku suuruse X moodi tähistatakse sümboliga $\operatorname{Mo} X$. Kui juhuslik suurus X allub binoomjaotusele parameetritega n ja p, siis suuruse $\operatorname{Mo} X$ saame määrata võrratuste ahela (1.9.2) abil.

Ülesanne 1. Näidake, et juhusliku suuruse X, mis allub normaaljaotusele parameetritega a ja σ, korral $\operatorname{Mo} X=a$.

Definitsioon 9. Kindlat suurust

$$\operatorname{As} X \stackrel{\text{def.}}{=} \mu_3/\sigma^3 \tag{2.5.14}$$

nimetatakse juhusliku suuruse X *asümmeetriakordajaks.*

Lause 3. Sümmeetrilise jaotuse korral $\operatorname{As} X=0$.

Tõestus. Et sümmeetrilise jaotuse korral

$$\exists\, \mathrm{E}X \stackrel{\text{tõestage!}}{\Rightarrow} \mathrm{E}X=\operatorname{Me} X,$$

siis

$$\begin{aligned}
\mu_3 &= \int_{-\infty}^{+\infty}\left(x-\mathrm{E}X\right)^3 f(x)dx=\int_{-\infty}^{+\infty}\left(x-\operatorname{Me} X\right)^3 f(x)dx= \\
&= [t=x-\operatorname{Me} X,\ x=t+\operatorname{Me} X,\ dt=dx]= \\
&= \int_{-\infty}^{+\infty} t^3 f(t+\operatorname{Me} X)dt= \\
&= \int_{-\infty}^{0} t^3 f(t+\operatorname{Me} X)dt+\int_{0}^{+\infty} t^3 f(t+\operatorname{Me} X)dt= \\
&= [\text{teostame esimeses integraalis muutujate vahetuse } t=-u]= \\
&= -\int_{0}^{+\infty} u^3 f(\operatorname{Me} X-u)du+\int_{0}^{+\infty} t^3 f(t+\operatorname{Me} X)dt \stackrel{(2.5.13)}{=} \\
&= -\int_{0}^{+\infty} u^3 f(\operatorname{Me} X+u)du+\int_{0}^{+\infty} t^3 f(t+\operatorname{Me} X)dt=0
\end{aligned}$$

ja $\operatorname{As} X=\mu_3/\sigma^3=0$. $\square$

Definitsioon 10. Arvu

$$\operatorname{Ex} X \stackrel{\text{def.}}{=} \mu_4/\sigma^4-3 \tag{2.5.15}$$

nimetatakse juhusliku suuruse X *ekstsessiks.*

Juhusliku suuruse X ekstsess mõõdab juhusliku suuruse X jaotuse erinevust sama keskväärtuse ja dispersiooniga normaaljaotusest, kusjuures normaaljaotusega juhusliku suuruse ekstsess on null.

Definitsioon 11. Diskreetse juhusliku suuruse X *entroopiaks* $\mathrm{H}\,(X)$ nimetatakse arvu, mis avaldub võimalike väärtuste x_k $(k \in I)$ omandamise tõenäosuste $p_k = P\,(X = x_k)$ kaudu kujul

$$\mathrm{H}\,(X) = -\textstyle\sum_{k \in I} p_k \ln p_k. \tag{2.5.16}$$

Näide 2. Leiame jaotusseadusele $P\,(X = k) = 1/n \quad (k = 1; \ldots ; n)$ alluva juhusliku suuruse X entroopia.

Saame

$$\mathrm{H}\,(X) = -\textstyle\sum_{k=1}^{n} \dfrac{1}{n} \ln \dfrac{1}{n} = -n\dfrac{1}{n}\,(\ln 1 - \ln n) = \ln n. \qquad \diamondsuit$$

Märkus 1. Saab tõestada, et kui diskreetsel juhuslikul suurusel X on n erinevat võimalikku väärtust, siis $\mathrm{H}(X) \leq \ln n$.

Definitsioon 12. Pideva juhusliku suuruse X *entroopiaks* nimetatakse arvu $\mathrm{H}\,(X)$, mis avaldub selle juhusliku suuruse X jaotustiheduse $f(x)$ kaudu kujul

$$\mathrm{H}(X) = -\int_{-\infty}^{+\infty} f(x) \ln f(x) dx. \tag{2.5.17}$$

Entroopia on juhusliku suuruse määramatuse ja tema võimalike väärtuste varieeruvuse mõõt.

Näide 3. Leiame juhusliku suuruse X, mille jaotustihedus on kujul

$$f(x) = 3x^2\,(\mathbf{1}\,(x) - \mathbf{1}\,(x-1))\,,$$

järgmised arvkarakteristikud: $\nu_k,\ \mu_k\ \ (k = 1; 2; 3; 4)\,,\ \sigma,\ \mathrm{Me}\,X,\ x_{0.25},\ x_{0.75},\ \mathrm{Mo}\,X,$ $\mathrm{As}\,X,\ \mathrm{Ex}\,X$ ja $\mathrm{H}(X)$.

Seose (2.5.3) abil saame

$$\nu_1 = \int_0^1 x \cdot 3x^2 dx = \frac{3}{4},\ \nu_2 = \int_0^1 x^2 \cdot 3x^2 dx = \frac{3}{5},$$
$$\nu_3 = \int_0^1 x^3 \cdot 3x^2 dx = \frac{1}{2}, \nu_4 = \int_0^1 x^4 \cdot 3x^2 dx = \frac{3}{7}.$$

Veenduge, et $\mu_1 = 0$. Valemite (2.5.8), (2.5.9) ja (2.5.10) abil saame vastavalt

$$\mu_2 = \nu_2 - (\nu_1)^2 = \frac{3}{5} - \left(\frac{3}{4}\right)^2 = \frac{3}{80} \quad \Rightarrow\ \sigma = \sqrt{\mu_2} = \sqrt{\frac{3}{80}},$$
$$\mu_3 = \nu_3 - 3\nu_1\nu_2 + 2\nu_1^3 = \frac{1}{2} - 3 \cdot \frac{3}{4} \cdot \frac{3}{5} + 2\left(\frac{3}{4}\right)^3 = -\frac{1}{160},$$
$$\mu_4 = \nu_4 - 4\nu_1\nu_3 + 6\nu_1^2\nu_2 - 3\nu_1^4 =$$
$$= \frac{3}{7} - 4 \cdot \frac{3}{4} \cdot \frac{1}{2} + 6 \cdot \left(\frac{3}{4}\right)^2 \cdot \frac{3}{5} - 3 \cdot \left(\frac{3}{4}\right)^4 = \frac{39}{8960}.$$

Et $F(x_p) = P(X < x_p)$ ja $F(x) = x^3 \cdot (\mathbf{1}(x) - \mathbf{1}(x-1)) + \mathbf{1}(x-1)$, siis seose (2.5.11) abil saame

$$\begin{aligned}
(\mathrm{Me}\,X)^3 = \frac{1}{2} \quad &\Rightarrow \quad \mathrm{Me}\,X = \sqrt[3]{\frac{1}{2}} = \frac{1}{2}\sqrt[3]{4}, \\
(x_{0.25})^3 = \frac{1}{4} \quad &\Rightarrow \quad x_{0.25} = \sqrt[3]{\frac{1}{4}} = \frac{1}{2}\sqrt[3]{2}, \\
(x_{0.75})^3 = \frac{3}{4} \quad &\Rightarrow \quad x_{0.75} = \sqrt[3]{\frac{3}{4}} = \frac{1}{2}\sqrt[3]{6}.
\end{aligned}$$

Et selle juhusliku suuruse jaotustihedus on rangelt kasvav lõigul $[0;1]$, siis $\mathrm{Mo}\,X = 1$. Valemi (2.5.14) abil leiame suuruse $\mathrm{As}\,X$:

$$\mathrm{As}\,X = \mu_3/\sigma^3 = \frac{-\dfrac{1}{160}}{\left(\sqrt{\dfrac{3}{80}}\right)^3} = -\frac{2}{9}\sqrt{15}.$$

Juhusliku suuruse X ekstsessi saame valemi (2.5.15) abil

$$\mathrm{Ex}\,X = \mu_4/\sigma^4 - 3 = \frac{39/8960}{\left(\sqrt{3/80}\right)^4} - 3 = \frac{2}{21}.$$

Juhusliku suuruse X entroopia $\mathrm{H}(X)$ saame valemi (2.5.17) abil

$$\mathrm{H}(X) = -\int_0^1 3x^2 \ln\left(3x^2\right) dx = \frac{2}{3} - \ln 3 \approx -0.431\,95. \qquad \diamondsuit$$

Näide 4. Leiame juhusliku suuruse X, mis allub Poissoni jaotusele parameetriga 2, moodi ja entroopia.

Et antud juhul $p_k = \dfrac{2^k}{k!}e^{-2}$, siis selle suuruse mood $\mathrm{Mo}\,X$ rahuldab võrratuste süsteemi

$$\begin{cases} \dfrac{2^{k-1}}{(k-1)!} \leq \dfrac{2^k}{k!} \\ \dfrac{2^{k+1}}{(k+1)!} \leq \dfrac{2^k}{k!} \end{cases} \Leftrightarrow \begin{cases} k \leq 2 \\ 1 \leq k \end{cases} \Rightarrow \mathrm{Mo}\,X = 1 \ \vee \ \mathrm{Mo}\,X = 2.$$

Entroopia $\mathrm{H}(X)$ avaldame valemi (2.5.16) abil

$$\mathrm{H}(X) = -\sum\nolimits_{k=0}^{\infty} \frac{2^k}{k!}e^{-2} \ln\left(\frac{2^k}{k!}e^{-2}\right) \overset{\mathrm{SWP}}{\approx} 1.704\,88. \qquad \diamondsuit$$

2.6 Juhusliku suuruse karakteristlik funktsioon

Definitsioon 1. Funktsiooni

$$g_X(\omega) \overset{\text{def.}}{=} \mathrm{E}e^{i\omega X} \tag{2.6.1}$$

nimetatakse juhusliku suuruse X *karakteristlikuks funktsiooniks.*

Seejuures tuleb suurust mõista järgnevalt:

$$\mathrm{E}e^{i\omega X} = \mathrm{E}\left(\cos(\omega X) + i\sin(\omega X)\right) = \mathrm{E}\left(\cos(\omega X)\right) + i\,\mathrm{E}\left(\sin(\omega X)\right).$$

Kui antud kontekstis on tegemist vaid ühe juhusliku suurusega X ja mingit segiminekut ei ole karta, siis tähistuse $g_X(\omega)$ asemel kasutatakse tihti lühendatud kirjapilti $g(\omega)$.

Kui $f(x)$ on juhusliku suuruse X jaotustihedus, siis valemi (2.3.4) põhjal saame

$$g_X(\omega) = \int_{-\infty}^{+\infty} e^{i\omega x} f(x)dx \tag{2.6.2}$$

ehk täpsemalt

$$g_X(\omega) = \mathrm{E}\left(\cos(\omega X)\right) + i\,\mathrm{E}\left(\sin(\omega X)\right) =$$

$$= \int_{-\infty}^{+\infty} \cos(\omega x)\, f(x)dx + i \int_{-\infty}^{+\infty} \sin(\omega x)\, f(x)dx = \int_{-\infty}^{+\infty} e^{i\omega x} f(x)dx.$$

Seega on juhusliku suuruse X karakteristlik funktsioon $g_X(\omega)$ selle juhusliku suuruse jaotustiheduse $f(x)$ Fourier' teisend. Teatud tingimustel on seos (2.6.2) pööratav

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-i\omega x} g_X(\omega)d\omega, \tag{2.6.3}$$

st juhusliku suuruse X jaotustihedus $f(x)$ on leitav kui suuruse X karakteristliku funktsiooni $g_X(\omega)$ Fourier' pöördteisend.

Kui X on diskreetne juhuslik suurus ja $\{x_k\}_{k\in I}$ on selle suuruse võimalike väärtuste hulk ning $p_k = P(X = x_k)_{k\in I}$, siis seostest (2.2.14) ja (2.6.2) järeldub

$$\begin{aligned}
g_X(\omega) &= \int_{-\infty}^{+\infty} e^{i\omega x} f(x)dx = \\
&= \int_{-\infty}^{+\infty} e^{i\omega x} \sum_{k\in I} p_k \cdot \delta(x - x_k)\, dx = \\
&= \sum_{k\in I} p_k \int_{-\infty}^{+\infty} e^{i\omega x} \cdot \delta(x - x_k)\, dx = \\
&= \left[\begin{array}{c} \text{rakendame seost (2.2.13) valiku} \\ \varphi(x) = e^{i\omega x} \text{ ja } a = x_k \text{ korral} \end{array} \right] = \\
&= \sum_{k\in I} e^{i\omega x_k} p_k.
\end{aligned}$$

Lause 1. Kui X on diskreetne juhuslik suurus ja $\{x_k\}_{k\in I}$ on selle suuruse võimalike väärtuste hulk ning $p_k = P\left(X = x_k\right)_{k\in I}$, siis

$$g_X\left(\omega\right) = \sum_{k\in I} e^{i\omega x_k} p_k. \tag{2.6.4}$$

Lause 2. Kui $X = cY + b$, kus c ja b on arvud, siis

$$g_X\left(\omega\right) = e^{i\omega b}\cdot g_Y\left(c\omega\right).$$

Tõestus. Leiame seose (2.6.1) abil, et

$$\begin{aligned} g_X\left(\omega\right) &= \mathrm{E}e^{i\omega X} = \mathrm{E}e^{i\omega(cY+b)} = \\ &= \mathrm{E}\left(e^{i\omega cY}e^{i\omega b}\right) = \begin{bmatrix}\text{tegurid}\\ \text{sõltumatud}\end{bmatrix} \overset{\text{Lause 2.3.2}}{=} \\ &= \mathrm{E}\left(e^{i(\omega c)Y}\right)\cdot \mathrm{E}\left(e^{i\omega b}\right) = e^{i\omega b}\, g_Y\left(c\omega\right). \qquad \Box \end{aligned}$$

Lause 3. Kui $X_1, \dots, X_n$ on sõltumatud juhuslikud suurused ja

$$X = \sum_{k=1}^{n} X_k,$$

siis

$$g_X\left(\omega\right) = \textstyle\prod_{k=1}^{n} g_{X_k}\left(\omega\right). \tag{2.6.5}$$

Tõestus. Seose (2.6.1) abil saame

$$\begin{aligned} g_X\left(\omega\right) &= \mathrm{E}e^{i\omega X} = \mathrm{E}e^{i\omega\sum_{k=1}^{n}X_k} = \mathrm{E}\left(e^{i\omega X_1}\cdots e^{i\omega X_n}\right) = \\ &= \begin{bmatrix}\text{suurused } X_k & \Rightarrow & \text{suurused } e^{i\omega X_k}\\ \text{on sõltumatud} & & \text{on sõltumatud}\end{bmatrix} = \\ &= \mathrm{E}e^{i\omega X_1}\cdots\mathrm{E}e^{i\omega X_n} = g_{X_1}\left(\omega\right)\cdots g_{X_n}\left(\omega\right) = \textstyle\prod_{k=1}^{n} g_{X_k}\left(\omega\right). \qquad \diamondsuit \end{aligned}$$

Lause 4. Kui $g_X\left(\omega\right)$ on juhusliku suuruse X karakteristlik funktsioon, siis:

1° $\exists\, g_X\left(\omega\right) \;\Rightarrow\; g_X\left(0\right) = 1 \;\wedge\; g_X\left(-\omega\right) = \overline{g_X\left(\omega\right)} \;\wedge\; \left|g_X\left(\omega\right)\right| \le 1;$

2° $\exists\, \mathrm{E}\left(X^k\right) \;\wedge\; \exists\, g_X^{(k)}\left(\omega\right) \;\Rightarrow\; \mathrm{E}\left(X^k\right) = g_X^{(k)}\left(0\right)/i^k.$

Tõestus. Valemi (2.6.2) abil saame

$$g_X\left(0\right) = \int_{-\infty}^{+\infty} e^{i0x} f(x)dx = \int_{-\infty}^{+\infty} f(x)dx \overset{(2.2.5)}{=} 1.$$

Et

$$g_X(-\omega) = \int_{-\infty}^{+\infty} e^{i(-\omega)x} f(x)dx = \int_{-\infty}^{+\infty} e^{-i\omega x} f(x)dx$$

ja

$$\overline{g_X(\omega)} = \overline{\int_{-\infty}^{+\infty} e^{i\omega x} f(x)dx} = \int_{-\infty}^{+\infty} \overline{e^{i\omega x} f(x)}dx =$$

$$= \int_{-\infty}^{+\infty} \overline{e^{i\omega x}} \cdot \overline{f(x)}dx = \left[\begin{array}{c} \overline{e^{i\omega x}} = e^{-i\omega x}, \\ \overline{f(x)} \overset{f(x)\geq 0}{=} f(x) \end{array} \right] = \int_{-\infty}^{+\infty} e^{-i\omega x} f(x)dx,$$

siis $g_X(-\omega) = \overline{g_X(\omega)}$. Kuna

$$\begin{aligned} |g_X(\omega)| &= \left| \int_{-\infty}^{+\infty} e^{i\omega x} f(x)dx \right| \leq \int_{-\infty}^{+\infty} \left| e^{i\omega x} f(x) \right| dx = \\ &= \int_{-\infty}^{+\infty} \left| e^{i\omega x} \right| |f(x)| \, dx = \left[\left| e^{i\omega x} \right| = 1, \; f(x) \geq 0 \right] = \\ &= \int_{-\infty}^{+\infty} f(x)dx \overset{(2.2.5)}{=} 1, \end{aligned}$$

siis saame hinnangu $|g_X(\omega)| \leq 1$. Et

$$\begin{aligned} \frac{d^k}{d\omega^k} g_X(\omega) &= \frac{d^k}{d\omega^k} \int_{-\infty}^{+\infty} e^{i\omega x} f(x)dx = \int_{-\infty}^{+\infty} \frac{d^k}{d\omega^k} \left(e^{i\omega x} f(x) \right) dx = \\ &= \int_{-\infty}^{+\infty} i^k x^k e^{i\omega x} f(x)dx, \end{aligned}$$

siis

$$g_X^{(k)}(0) = \int_{-\infty}^{+\infty} i^k x^k e^{i0x} f(x)dx = i^k \int_{-\infty}^{+\infty} x^k f(x)dx \overset{(2.5.3)}{=} i^k \mathrm{E}\left(X^k\right)$$

ja

$$\mathrm{E}\left(X^k\right) = g_X^{(k)}(0) / i^k. \qquad \square \tag{2.6.6}$$

Näide 1. Leiame juhusliku suuruse X, mis allub lõigul $[a, b]$ ühtlasele jaotusele, karakteristliku funktsiooni ja selle abil keskväärtuse EX.

Et selle juhusliku suuruse jaotustihedus on kujul

$$f(x) = \begin{cases} 1/(b-a), \text{ kui } x \in [a, b], \\ 0, \text{ kui } x \notin [a, b], \end{cases}$$

siis vastavalt valemile (2.6.2) saame

$$g_X(\omega) = \int\limits_{-\infty}^{+\infty} e^{i\omega x} f(x)dx = \int\limits_a^b e^{i\omega x} \frac{1}{b-a} dx = \left. \frac{e^{i\omega x}}{i\omega(b-a)} \right|_a^b = \frac{e^{i\omega b} - e^{i\omega a}}{i\omega(b-a)}.$$

Kuna

$$\nexists\ g_X(0) \ \Rightarrow\ \nexists\, g_X^{(k)}(0) \quad (k \in \mathbf{N}),$$

siis valem (2.6.6) ei ole vahetult rakendatav. Osutub, et kehtib selle valemi üldistus

$$\mathrm{E}\left(X^k\right) = \lim_{\omega\to 0} g_X^{(k)}(\omega)/i^k. \tag{2.6.7}$$

Saame

$$\begin{aligned} g_X'(\omega) &= \frac{d}{d\omega}\frac{e^{i\omega b}-e^{i\omega a}}{i\omega(b-a)} = \\ &= \frac{1}{i(b-a)}\frac{\left(e^{i\omega b}ib-e^{i\omega a}ia\right)\omega-\left(e^{i\omega b}-e^{i\omega a}\right)}{\omega^2}. \end{aligned}$$

Valemi (2.6.7) ning L'Hospitali reegli abil leiame

$$\mathrm{E}X = \lim_{\omega\to 0} g_X'(\omega)/i = \frac{1}{i^2(b-a)}\lim_{\omega\to 0}\frac{\left(e^{i\omega b}ib-e^{i\omega a}ia\right)\omega-\left(e^{i\omega b}-e^{i\omega a}\right)}{\omega^2} =$$

$$\begin{aligned} &= \frac{1}{i^2(b-a)}\lim_{\omega\to 0}\frac{\left(e^{i\omega b}(ib)^2-e^{i\omega a}(ia)^2\right)\omega}{2\omega} = \\ &= \frac{1}{i^2(b-a)}\lim_{\omega\to 0}\frac{\left(e^{i\omega b}(ib)^2-e^{i\omega a}(ia)^2\right)}{2} = \frac{b^2-a^2}{2(b-a)} = \frac{a+b}{2}. \quad \diamondsuit \end{aligned}$$

Näide 2. Leiame juhusliku suuruse X, mis allub binoomjaotusele parameetritega n ja p, karakteristliku funktsiooni. Kasutades karakteristlikku funktsiooni, leiame suurused $\mathrm{E}X$ ning $\mathrm{D}X$.

Rakendame valemit (2.6.4), kusjuures $x_k = k$, $p_k = C_n^k p^k q^{n-k}$ ja $I = \{0;1;2;\ldots;n\}$. Saame

$$g_X(\omega) = \sum_{k=0}^{n} e^{i\omega k} C_n^k p^k q^{n-k} = \sum_{k=0}^{n} C_n^k \left(pe^{i\omega}\right)^k q^{n-k} = \left(pe^{i\omega}+q\right)^n.$$

Seega

$$g_X(\omega) = \left(pe^{i\omega}+q\right)^n. \tag{2.6.8}$$

Tõestage valem (2.6.8) ka seoste (2.3.3) ja (2.6.5) abil.

Kuna

$$\begin{aligned} g_X'(\omega) &= n\left(pe^{i\omega}+q\right)^{n-1} pie^{i\omega} \Rightarrow g_X'(0) = n\left(pe^{i0}+q\right)^{n-1} pie^{i0} = npi \Rightarrow \\ &\Rightarrow \mathrm{E}(X) = g_X'(0)/i = npi/i = np \end{aligned}$$

ja

$$\begin{aligned}
g''_X(\omega) &= n(n-1)\left(pe^{i\omega}+q\right)^{n-2}\left(pie^{i\omega}\right)^2 + n\left(pe^{i\omega}+q\right)^{n-1}pi^2e^{i\omega} \Rightarrow \\
&\Rightarrow g''_X(0) = n(n-1)\left(pe^{i0}+q\right)^{n-2}\left(pie^{i0}\right)^2 + n\left(pe^{i0}+q\right)^{n-1}pi^2e^{i0} = \\
&= n(n-1)p^2i^2 + npi^2 \;\Rightarrow \\
&\Rightarrow \mathrm{E}\left(X^2\right) = g''_X(0)/i^2 = \left(n(n-1)p^2i^2+npi^2\right)/i^2 = \\
&= n(n-1)p^2+np,
\end{aligned}$$

siis

$$\begin{aligned}
\mathrm{D}X &= \mathrm{E}\left(X^2\right) - (\mathrm{E}X)^2 = n(n-1)p^2 + np - n^2p^2 = \\
&= np - np^2 = np(1-p) = npq. \qquad \diamondsuit
\end{aligned}$$

Näide 3. Sõltumatud juhuslikud suurused X ja Y alluvad binoomjaotustele vastavalt parameetritega 2 ja 1/3 ning 3 ja 1/2. Leiame juhusliku suuruse $Z = X + Y$ karakteristliku funktsiooni ja jaotusseaduse ning jaotusfunktsiooni.

Valemi (2.6.8) abil saame

$$g_X(\omega) = \left(e^{i\omega}/3 + 2/3\right)^2,\ g_Y(\omega) = \left(e^{i\omega}/2 + 1/2\right)^3$$

ja valemi (2.6.5) abil

$$\begin{aligned}
g_Z(\omega) &= \left(e^{i\omega}/3 + 2/3\right)^2\left(e^{i\omega}/2 + 1/2\right)^3 = \\
&= \frac{1}{72}e^{5i\omega} + \frac{7}{72}e^{4i\omega} + \frac{19}{72}e^{3i\omega} + \frac{25}{72}e^{2i\omega} + \frac{2}{9}e^{i\omega} + \frac{1}{18}.
\end{aligned}$$

Seega

z_k	0	1	2	3	4	5
p_k	1/18	2/9	25/72	19/72	7/72	1/72

on juhusliku suuruse Z jaotusseadus ja

$$\begin{aligned}
F(z) &= \frac{1}{18}\mathbf{1}(z) + \frac{2}{9}\mathbf{1}(z-1) + \frac{25}{72}\mathbf{1}(z-2) + \frac{19}{72}\mathbf{1}(z-3) + \frac{7}{72}\mathbf{1}(z-4) + \\
&+ \frac{1}{72}\mathbf{1}(z-5). \qquad \diamondsuit
\end{aligned}$$

Näide 4. Leiame juhusliku suuruse X, mis allub Poissoni jaotusele parameetriga λ, karakteristliku funktsiooni. Kasutades karakteristlikku funktsiooni, leiame suurused $\mathrm{E}X$ ning $\mathrm{D}X$.

Sel korral $x_k = k$, $p_k = e^{-\lambda}\lambda^k/k!$ ja $I = \mathbf{N}_0$ ning valemi (2.6.4) abil saame

$$g_X(\omega) = \sum_{k=0}^{\infty}\frac{e^{i\omega k}\lambda^k}{k!}e^{-\lambda} = e^{-\lambda}\sum_{k=0}^{\infty}\frac{\left(\lambda e^{i\omega}\right)^k}{k!} = e^{-\lambda}e^{\lambda e^{i\omega}} = e^{\lambda\left(e^{i\omega}-1\right)}.$$

Seega

$$g_X(\omega) = e^{\lambda\left(e^{i\omega}-1\right)}. \tag{2.6.9}$$

Kuna

$$g_X'(\omega) = e^{\lambda\left(e^{i\omega}-1\right)}\lambda e^{i\omega} i \Rightarrow g_X'(0) = e^{\lambda\left(e^{i0}-1\right)}\lambda e^{i0} i = \lambda i \Rightarrow$$
$$\Rightarrow \mathrm{E}(X) = g_X'(0)/i = \lambda i/i = \lambda$$

ja

$$g_X''(\omega) = e^{\lambda\left(e^{i\omega}-1\right)}\left(\lambda e^{i\omega} i\right)^2 + e^{\lambda\left(e^{i\omega}-1\right)}\lambda e^{i\omega} i^2 \Rightarrow$$
$$\Rightarrow g_X''(0) = e^{\lambda\left(e^{i0}-1\right)}\left(\lambda e^{i0} i\right)^2 + e^{\lambda\left(e^{i0}-1\right)}\lambda e^{i0} i^2 = \lambda^2 i^2 + \lambda i^2 \Rightarrow$$
$$\Rightarrow \mathrm{E}\left(X^2\right) = g_X''(0)/i^2 = \left(\lambda^2 i^2 + \lambda i^2\right)/i^2 = \lambda^2 + \lambda,$$

siis

$$\mathrm{D}X = \mathrm{E}\left(X^2\right) - (\mathrm{E}X)^2 = \lambda^2 + \lambda - \lambda^2 = \lambda. \qquad \diamondsuit$$

Näide 5. Leiame juhusliku suuruse $X \sim N(a, \sigma)$ karakteristliku funktsiooni. Kasutades karakteristlikku funktsiooni, leiame suurused $\mathrm{E}X$ ning $\mathrm{D}X$.

Olgu $Y \sim N(0;1)$. Et juhusliku suuruse Y jaotustihedus on

$$f(y) = e^{-y^2/2}/\sqrt{2\pi},$$

siis valemi (2.6.2) abil saame

$$g_Y(\omega) = \int_{-\infty}^{+\infty} e^{i\omega y} f(y)dy = \int_{-\infty}^{+\infty} e^{i\omega y}\cdot\frac{1}{\sqrt{2\pi}}e^{-y^2/2}dy = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{i\omega y - y^2/2}dy =$$
$$= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{-(y-i\omega)^2/2-\omega^2/2}dy = e^{-\omega^2/2}\cdot\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{-(y-i\omega)^2/2}dy =$$
$$= \left[\begin{array}{c}\text{saab näidata, et}\\ \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{-(y-i\omega)^2/2}dy = 1\end{array}\right] = e^{-\omega^2/2}.$$

Seega

$$Y \sim N(0;1) \;\Leftrightarrow\; g_Y(\omega) = e^{-\omega^2/2}. \tag{2.6.10}$$

Kui $Y \sim N(0;1)$ ja $X = \sigma Y + a \quad (c \neq 0)$, siis

$$F_X(x) = P(X < x) = P(\sigma Y + a < x) = P\left(Y < \frac{x-a}{\sigma}\right) =$$
$$= [Y \sim N(0;1)] = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{(x-a)/\sigma} e^{-y^2/2}dy$$

ning

$$f_X(x) = \frac{dF_X(x)}{dx} = \frac{dF_X(x)}{d\left((x-a)/\sigma\right)}\cdot\frac{d\left((x-a)/\sigma\right)}{dx} =$$
$$= \frac{1}{\sqrt{2\pi}}e^{-((x-a)/\sigma)^2/2}\cdot\frac{1}{\sigma} = \frac{1}{\sigma\sqrt{2\pi}}e^{-(x-a)^2/\left(2\sigma^2\right)},$$

st

$$Y \sim N(0;1) \ \wedge \ X = \sigma Y + a \ \ \Leftrightarrow \ X \sim N(a,\sigma).$$

Suuruse $X \sim N(a,\sigma)$ karakteristliku funktsiooni leidmiseks kasutame Lauset 2, valides $c = \sigma,\ b = a$. Saame

$$g_X(\omega) = e^{i\omega a} \cdot g_Y(\sigma\omega) = e^{i\omega a} \cdot e^{-\sigma^2\omega^2/2} = e^{i\omega a - \sigma^2\omega^2/2}.$$

Seega

$$X \sim N(a,\sigma) \ \Leftrightarrow \ g_X(\omega) = e^{i\omega a - \sigma^2\omega^2/2}. \tag{2.6.11}$$

Kuna

$$\begin{aligned} g'_X(\omega) &= e^{i\omega a - \sigma^2\omega^2/2}\left(ia - \sigma^2\omega\right) \Rightarrow g'_X(0) = e^{i0a - \sigma^2 0^2/2}\left(ia - \sigma^2 0\right) = ia \Rightarrow \\ &\Rightarrow \mathrm{E}X = g'_X(0)/i = ia/i = a \end{aligned}$$

ja

$$\begin{aligned} g''_X(\omega) &= e^{i\omega a - \sigma^2\omega^2/2}\left(ia - \sigma^2\omega\right)^2 + e^{i\omega a - \sigma^2\omega^2/2}\left(-\sigma^2\right) \Rightarrow \\ &\Rightarrow g''_X(0) = e^{i0a - \sigma^2 0^2/2}\left(ia - \sigma^2 0\right)^2 + e^{i0a - \sigma^2 0^2/2}\left(-\sigma^2\right) = \\ &= -a^2 - \sigma^2 \Rightarrow \\ &\Rightarrow \mathrm{E}\left(X^2\right) = g''_X(0)/i^2 = \left(-a^2 - \sigma^2\right)/i^2 = a^2 + \sigma^2, \end{aligned}$$

siis

$$\mathrm{D}X = \mathrm{E}\left(X^2\right) - (\mathrm{E}X)^2 = a^2 + \sigma^2 - a^2 = \sigma^2. \qquad \diamondsuit$$

Lause 5. Kui $g_X(\omega)$ on juhusliku suuruse X karakteristlik funktsioon ja eksisteerivad selle juhusliku suuruse k-ndat järku keskmoment ning

$$\left.\left(\frac{d^k}{d\omega^k}\left(e^{-i\omega \mathrm{E}X} g_X(\omega)\right)\right)\right|_{\omega=0},$$

siis

$$\mu_k = \frac{1}{i^k}\left.\left(\frac{d^k}{d\omega^k}\left(e^{-i\omega \mathrm{E}X} g_X(\omega)\right)\right)\right|_{\omega=0}. \tag{2.6.12}$$

Tõestus. Kuna

$$\frac{d^k}{d\omega^k}\left(e^{-i\omega \mathrm{E}X} g_X(\omega)\right) = \frac{d^k}{d\omega^k}\left(e^{-i\omega \mathrm{E}X} \int_{-\infty}^{+\infty} e^{i\omega x} f(x)dx\right) =$$

$$= \frac{d^k}{d\omega^k}\left(\int_{-\infty}^{+\infty} e^{i\omega x - i\omega \mathrm{E}X} f(x)dx\right) = \int_{-\infty}^{+\infty} \frac{d^k}{d\omega^k}\left(e^{i\omega(x - \mathrm{E}X)} f(x)\right)dx =$$

$$= i^k \int_{-\infty}^{+\infty} (x - \mathrm{E}X)^k e^{i\omega(x-\mathrm{E}X)} f(x)dx,$$

siis

$$\left(\frac{d^k}{d\omega^k}\left(e^{-i\omega \mathrm{E}X} g_X(\omega)\right)\right)\Big|_{\omega=0} = i^k \int_{-\infty}^{+\infty} (x-\mathrm{E}X)^k e^{i0(x-\mathrm{E}X)} f(x)dx =$$

$$= i^k \int_{-\infty}^{+\infty} (x-\mathrm{E}X)^k f(x)dx = i^k \mu_k.$$

Seega väide (2.6.12) on tõene. □

Järeldus 1. Kui eksisteerivad DX ja $\left(\frac{d^2}{d\omega^2}\left(e^{-i\omega \mathrm{E}X} g_X(\omega)\right)\right)\Big|_{\omega=0}$, siis

$$DX = \frac{1}{i^2}\left(\frac{d^2}{d\omega^2}\left(e^{-i\omega \mathrm{E}X} g_X(\omega)\right)\right)\Big|_{\omega=0}.$$

2.7 Juhusliku suuruse genereeriv funktsioon

Definitsioon 1. Funktsiooni

$$G_X(z) \stackrel{\text{def.}}{=} \mathrm{E}z^X \tag{2.7.0.1}$$

nimetatakse juhusliku suuruse X *genereerivaks funktsiooniks.*

Märkus 1. Kehtib seos

$$G_X\left(e^{i\omega}\right) = g_X(\omega). \tag{2.7.0.2}$$

Tõestus. Saame

$$G_X\left(e^{i\omega}\right) = \mathrm{E}\left(\left(e^{i\omega}\right)^X\right) = \mathrm{E}\left(e^{i\omega X}\right) = g_X(\omega). \qquad \square$$

Kui $f(x)$ on juhusliku suuruse X jaotustihedus, siis vastavalt Definitsioonile 2.3.3 saame

$$G_X(z) = \int_{-\infty}^{+\infty} z^x f(x)dx. \tag{2.7.3}$$

Näide 1. Juhuslik suurus X allub ühtlasele jaotusele lõigul $[a,b]$. Leiame selle juhusliku suuruse genereeriva funktsiooni valemi (2.7.3) abil.

Saame

$$G_X(z) = \int_a^b z^x \frac{1}{b-a}dx = \left.\frac{z^x}{(b-a)\ln z}\right|_a^b = \frac{z^b - z^a}{(b-a)\ln z}. \qquad \diamondsuit$$

Lause 1. Kui X on diskreetne juhuslik suurus ja $\{x_k\}_{k\in I}$ on selle suuruse võimalike väärtuste hulk ning $p_k = P(X = x_k)_{k\in I}$, siis

$$G_X(z) = \sum\nolimits_{k\in I} z^{x_k} p_k. \tag{2.7.4}$$

Tõestus. Seostest (2.7.3) ja (2.2.14) järeldub

$$\begin{aligned}
G_X(z) &= \int_{-\infty}^{+\infty} z^x f(x)dx = \\
&= \int_{-\infty}^{+\infty} z^x \textstyle\sum_{k\in I} p_k \cdot \delta(x-x_k)\,dx = \\
&= \textstyle\sum_{k\in I} p_k \displaystyle\int_{-\infty}^{+\infty} z^x \cdot \delta(x-x_k)\,dx = \\
&= \left[\begin{array}{c}\text{rakendame seost (2.2.13) valiku}\\ \varphi(x)=z^x \text{ ja } a=x_k \text{ korral}\end{array}\right] = \\
&= \textstyle\sum_{k\in I} z^{x_k} p_k. \qquad \square
\end{aligned}$$

Märkus 2. Tavaliselt vaadeldakse genereerivat funktsiooni vaid diskreetse juhusliku suuruse, mille võimalike väärtuste hulk on $\{0; 1; \ldots; n\}$ või $\mathbf{N}_0$, korral. Sel korral on genereerivaks funktsiooniks vastavalt polünoom

$$G_X(z) = \textstyle\sum_{k=0}^{n} z^k P(X=k) \tag{2.7.0.3}$$

või astmerida

$$G_X(z) = \textstyle\sum_{k=0}^{\infty} z^k P(X=k). \tag{2.7.0.4}$$

Näide 2. Leiame juhusliku suuruse X, mis allub Poissoni jaotusele parameetriga λ, genereeriva funktsiooni.

Valemi (2.7.6) abil saame

$$G_X(z) = \textstyle\sum_{k=0}^{\infty} z^k \dfrac{\lambda^k}{k!} e^{-\lambda} = \textstyle\sum_{k=0}^{\infty} \dfrac{(z\lambda)^k}{k!} e^{-\lambda} = e^{z\lambda}\cdot e^{-\lambda} = e^{\lambda(z-1)}. \qquad \diamondsuit$$

Lause 2. Kui $G_X(z)$ on juhusliku suuruse X karakteristlik funktsioon, siis:

1° $G_X(1) = 1$;

2° $\exists \mathrm{E}X \ \wedge\ \exists \mathrm{E}\left(X^2\right) \ \wedge\ \exists G_X'(1) \ \wedge\ \exists G_X''(1) \ \Rightarrow$
$\Rightarrow\ \mathrm{E}X = G_X'(1) \ \wedge\ \mathrm{E}\left(X^2\right) = G_X''(1) + G_X'(1).$

Tõestus. Seose (2.7.3) abil saame

$$\begin{aligned}
G_X'(z) &= \frac{d}{dz}\int_{-\infty}^{+\infty} z^x f(x)dx = \int_{-\infty}^{+\infty} \frac{d}{dz}\left(z^x f(x)\right)dx = \\
&= \int_{-\infty}^{+\infty} x z^{x-1} f(x)dx \Rightarrow \\
&\Rightarrow\ G_X'(1) = \int_{-\infty}^{+\infty} x 1^{x-2} f(x)dx = \\
&= \int_{-\infty}^{+\infty} x f(x)dx = \mathrm{E}X \quad\Rightarrow\quad \mathrm{E}X = G_X'(1)
\end{aligned}$$

ja

$$\begin{aligned}
G''_X(z) &= \frac{d^2}{dz^2}\int_{-\infty}^{+\infty} z^x f(x)dx = \int_{-\infty}^{+\infty} \frac{d^2}{dz^2}\left(z^x f(x)\right)dx = \\
&= \int_{-\infty}^{+\infty} x(x-1)\,z^{x-2} f(x)dx \;\Rightarrow \\
&\Rightarrow\; G''_X(1) = \int_{-\infty}^{+\infty} x(x-1)\,1^{x-2} f(x)dx = \\
&= \int_{-\infty}^{+\infty} x(x-1)\,f(x)dx = \mathrm{E}\left(X^2\right) - \mathrm{E}X \;\Rightarrow \\
&\Rightarrow\; \mathrm{E}\left(X^2\right) = G''_X(1) + G'_X(1). \qquad \square
\end{aligned}$$

Lause 3. Kui $X_1, \ldots, X_n$ on sõltumatud juhuslikud suurused ja

$$X = \sum_{k=1}^{n} X_k, \tag{2.7.0.5}$$

siis

$$G_X(z) = \prod\nolimits_{k=1}^{n} G_{X_k}(z). \tag{2.7.0.6}$$

Tõestus. Definitsiooni 1 abil saame

$$\begin{aligned}
G_X(z) &= \mathrm{E}z^X \overset{(2.7.7)}{=} \mathrm{E}z^{\sum_{k=1}^{n} X_k} = \mathrm{E}\left(z^{X_1} z^{X_2} \cdots z^{X_n}\right) = \begin{bmatrix} \text{tegurite} \\ \text{sõltumatus} \end{bmatrix} = \\
&= \left(\mathrm{E}z^{X_1}\right)\left(\mathrm{E}z^{X_2}\right)\cdots\left(\mathrm{E}z^{X_n}\right) = \prod\nolimits_{k=1}^{n} G_{X_k}(z). \qquad \square
\end{aligned}$$

Näide 3. Sooritatakse n sõltumatut katset, kusjuures k-ndal ($k = 1; \ldots; n$) katsel toimub sündmus A tõenäosusega p_k. Olgu X sündmuse A toimumiste koguarv selles katseseerias. Leiame juhusliku suuruse X genereeriva funktsiooni.

Kui X_k on sündmuse A toimumiste arv k-ndal katsel, siis kehtib seos (2.7.7) ja tänu katsete sõltumatusele selles seerias Lause 3 põhjal ka seos (2.7.8). Et

$$G_{X_k}(z) = z^1 p_k + z^0 q_k = p_k z + q_k \quad (k = 1; 2; \ldots; n),$$

kus $q_k = 1 - p_k$, siis valemi (2.7.8) abil saame

$$G_X(z) = \prod\nolimits_{k=1}^{n} \left(p_k z + q_k\right). \qquad \diamondsuit \tag{2.7.0.7}$$

Viimase näite erijuhuna $p_k = p \;\; (k = 1; 2; \ldots; n)$ saame binoomjaotusele alluva juhusliku suuruse genereeriva funktsiooni.

Järeldus 1. Kui juhuslik suurus X allub binoomjaotusele parameetritega n ja p, siis

$$G_X(z) = (pz + q)^n,$$

kus $q = 1 - p$.

Lause 4. Kui $X_1, \ldots, X_n$ on sõltumatud diskreetsed juhuslikud suurused vastavalt võimalike väärtuste hulkadega

$$\{0; 1; \ldots; m_r\} \quad (r = 1; 2; \ldots; n), \tag{2.7.10}$$

siis

$$\{0; 1; \ldots; m\} \quad (m = \textstyle\sum_{r=1}^{n} m_r) \tag{2.7.11}$$

on seosega (2.7.7) määratud juhusliku suuruse X võimalike väärtuste hulk ja tõenäosus $P(X = k)$ on leitav kui astme x^k kordaja juhuslike suuruste X_r genereerivate funktsioonide korrutise

$$\textstyle\prod_{r=1}^{n} G_{X_r}(z) \tag{2.7.0.8}$$

arenduses muutuja x astmete järgi.

Tõestus. Neil eeldustel saame

$$G_{X_r}(z) = \textstyle\sum_{\nu=0}^{m_r} z^{\nu} P(X_r = v)$$

ja

$$G_X(z) = \textstyle\sum_{k=0}^{m} z^k P(X = k) = (\sum_{\nu=0}^{m_1} z^{\nu} P(X_1 = v)) \cdots (\sum_{\nu=0}^{m_n} z^{\nu} P(X_n = v)).$$

Et kaks polünoomi

$$\textstyle\sum_{k=0}^{m} z^k P(X = k)$$

ja

$$\textstyle\prod_{r=1}^{n} G_{X_r}(z) = (\sum_{\nu=0}^{m_1} z^{\nu} P(X_1 = v)) \cdots (\sum_{\nu=0}^{m_n} z^{\nu} P(X_n = v))$$

on võrdsed parajasti siis, kui neis vastavate astmete kordajad on võrdsed, siis tõenäosus $P(X = k)$ on leitav kui astme x^k kordaja juhuslike suuruste X_r genereerivate funktsioonide korrutise (2.7.12) arenduses muutuja x astmete järgi. □

Näide 4. Mäetippu ründab üksteisest sõltumatult 5 alpinisti, kusjuures neil alpinistidel on mäetipu vallutamise tõenäosus vastavalt 0.4; 0.8; 0.6; 0.7 ja 0.5. Olgu X tippu jõudvate alpinistide koguarv. Leiame juhusliku suuruse X genereeriva funktsiooni ja Lause 4 abil sündmuste $X = k \quad (k = 0; 1; 2; 3; 4; 5)$ tõenäosused.

Valemi (2.7.9) abil saame

$$\begin{aligned}
&G_X(z) = \textstyle\prod_{r=1}^{5} (p_r z + q_r) = \\
&= (0.4x + 0.6)(0.8x + 0.2)(0.6x + 0.4)(0.7x + 0.3)(0.5x + 0.5) = \\
&= 0.0\,672x^5 + 0.\,258\,4x^4 + 0.\,364\,4x^3 + 0.\,234\,4x^2 + 0.0\,684x + 0.00\,72.
\end{aligned}$$

Lause 4 põhjal on genereeriva funktsiooni $G_X(z)$ arenduses muutuja x astmete järgi astme x^5 kordajaks $P(X = 5)$, astme x^4 kordajaks tõenäosus $P(X = 4)$, astme x^3 kordajaks $P(X = 3)$, astme x^2 kordajaks $P(X = 2)$, astme x^1 kordajaks $P(X = 1)$ ja astme x^0 kordajaks $P(X = 0)$. Seega

$$P(X = 5) = 0.0\,672,\ P(X = 4) = 0.\,258\,4,\ P(X = 3) = 0.\,364\,4,$$
$$P(X = 2) = 0.\,234\,4,\ P(X = 1) = 0.0\,684,\ P(X = 0) = 0.00\,72.$$

Teostame kontrolli, kõigi nende tõenäosuste summa peab olema 1:

$$\begin{gathered}\textstyle\sum_{k=0}^{5} P\,(X=k) = \\ = 0.0\,672 + 0.\,258\,4 + 0.\,364\,4 + 0.\,234\,4 + 0.0\,684 + 0.00\,72 = 1.0. \qquad \diamondsuit\end{gathered}$$

2.8 Normaaljaotus

Normaaljaotusele parameetritega a ja σ $(\sigma > 0)$ alluva juhusliku suuruse $X \sim N(a, \sigma)$ jaotustiheduse defineeerisime seose (2.2.9) abil, kusjuures $a = \mathrm{E}X$ ja $\sigma = \sigma_X = \sqrt{\mathrm{D}X}$. Paketis SWP on $X \sim N(a, \sigma)$ jaotustiheduse ja jaotusfunktsiooni tähisteks vastavalt $\mathrm{NormalDen}(x, a, \sigma)$ ja $\mathrm{NormalDist}(x, a, \sigma)$ ning

$$\mathrm{NormalDist}(x) \overset{\text{def.}}{=} \mathrm{NormalDist}(x, 0, 1).$$

Lisaks saab paketis SWP $p \in (0; 1)$ korral kasutada funktsiooni

$$p \longmapsto x_p = \mathrm{NormalInv}\,(p) : \ \mathrm{NormalDist}(x_p) = p.$$

Näide 1. Leiame juhusliku suuruse $X \sim N(a, \sigma)$ jaotustiheduse $y = f(x)$ graafiku käänupunktid.

Kuna

$$\begin{aligned}\frac{d^2}{dx^2} f(x) &= \frac{d^2}{dx^2}\left(\frac{1}{\sigma\sqrt{2\pi}} e^{-(x-a)^2/\left(2\sigma^2\right)}\right) = \\ &= \frac{1}{\sigma^5\sqrt{2\pi}} e^{-(x-a)^2/\left(2\sigma^2\right)}\left(-\sigma^2 + x^2 - 2xa + a^2\right),\end{aligned}$$

siis

$$\frac{d^2}{dx^2} f(x) = 0 \ \Leftrightarrow \ x^2 - 2xa + a^2 - \sigma^2 = 0 \ \Rightarrow \ x_{1;2} = a \pm \sigma,$$

st punktid

$$\left(a - \sigma, \frac{1}{\sigma\sqrt{2e\pi}}\right), \ \left(a + \sigma, \frac{1}{\sigma\sqrt{2e\pi}}\right)$$

on jaotustiheduse graafiku käänupunktid. Miks? $\diamondsuit$

Näide 2. Veendume, et juhusliku suuruse $X \sim N(a, \sigma)$ korral

$$\mu_{2k-1} = 0 \quad (k \in \mathbf{N}). \tag{2.8.0.9}$$

Tõesti

$$\begin{aligned}\mu_{2k-1} &= \int_{-\infty}^{+\infty} \frac{(x-a)^{2k-1}}{\sigma\sqrt{2\pi}} e^{-(x-a)^2/\left(2\sigma^2\right)} dx = \left[\begin{array}{c} t = \dfrac{x-a}{\sigma}, \\ x = a + \sigma t \end{array}\right] = \\ &= \frac{\sigma^{2k-1}}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} t^{2k-1} e^{-t^2/2} dt = \left[\begin{array}{c} \text{paaritu funktsioon, rajad} \\ \text{sümmeetrilised nulli suhtes} \end{array}\right] = 0. \ \diamondsuit\end{aligned}$$

Saab tõestada, et

$$X \sim N(a,\sigma) \quad \Rightarrow \quad \mu_{2k} = (2k-1)!!\sigma^{2k} \quad (k \in \mathbf{N})\,. \tag{2.8.2}$$

Suuruse X jaotusfunktsioon on kujul

$$F(x) = \int_{-\infty}^{x} \frac{1}{\sigma\sqrt{2\pi}} e^{-(t-a)^2/\left(2\sigma^2\right)} dt.$$

Tõestasime, et $g_X(\omega) = e^{i\omega a - \sigma^2\omega^2/2}$ on suuruse $X \sim N(a,\sigma)$ karakteristlik funktsioon. Olgu

$$\Phi(x) \overset{\text{def.}}{=} \frac{1}{\sqrt{2\pi}} \int_0^x e^{-t^2/2} dt. \tag{2.8.0.10}$$

Uurime tõenäosuse $P(\alpha \le X \le \beta)$ arvutamist *Laplace'i funktsiooni* $\Phi(x)$ abil. Et

$$\Phi(-x) = \frac{1}{\sqrt{2\pi}} \int_0^{-x} e^{-t^2/2} dt = -\frac{1}{\sqrt{2\pi}} \int_{-x}^{0} e^{-t^2/2} dt =$$

$$= \left[\begin{array}{c} e^{-t^2/2} \text{ on paaris-} \\ \text{funktsioon} \end{array} \Rightarrow \int_{-x}^{0} e^{-t^2/2} dt = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-t^2/2} dt \right] =$$

$$= -\frac{1}{\sqrt{2\pi}} \int_0^x e^{-t^2/2} dt = -\Phi(x),$$

siis funktsioon $\Phi(x)$ on paaritu. Saame

$$\begin{aligned} P(\alpha \le X \le \beta) &= \frac{1}{\sigma\sqrt{2\pi}} \int_\alpha^\beta e^{-(x-a)^2/\left(2\sigma^2\right)} dx = \\ &= \left[\begin{array}{c} t = \dfrac{x-a}{\sigma},\ dx = \sigma dt \\ x = a + \sigma t, \end{array} \right] = \frac{1}{\sqrt{2\pi}} \int\limits_{(\alpha-a)/\sigma}^{(\beta-a)/\sigma} e^{-t^2/2} dt = \\ &= \frac{1}{\sqrt{2\pi}} \int_0^{(\beta-a)/\sigma} e^{-t^2/2} dt - \frac{1}{\sqrt{2\pi}} \int_0^{(\alpha-a)/\sigma} e^{-t^2/2} dt \overset{(2.8.3)}{=} \\ &= \Phi\left(\frac{\beta-a}{\sigma}\right) - \Phi\left(\frac{\alpha-a}{\sigma}\right). \end{aligned}$$

Lisas 2 esitame lühikese tabeli funktsiooni $\Phi(x)$ ümardatud väärtustega.

Lause 1. Kui $X \sim N(a,\sigma)$, siis

$$P(\alpha \le X \le \beta) = \Phi\left(\frac{\beta-a}{\sigma}\right) - \Phi\left(\frac{\alpha-a}{\sigma}\right). \tag{2.8.4}$$

Järeldus 1. Kui $X \sim N(a,\sigma)$, siis kehtib valem

$$P(|X-a| \le \gamma) = 2 \cdot \Phi\left(\frac{\gamma}{\sigma}\right). \tag{2.8.5}$$

Tõestus. Et

$$
\begin{aligned}
P\left(|X-a| \leq \gamma\right) &= P\left(-\gamma \leq X-a \leq \gamma\right) = P\left(a-\gamma \leq X \leq a+\gamma\right) = \\
&= \left[\begin{array}{c}\text{rakendame valemit (2.8.4) juhul} \\ \alpha = a-\gamma \text{ ja } \beta = a+\gamma\end{array}\right] = \\
&= \Phi\left(\frac{a+\gamma-a}{\sigma}\right) - \Phi\left(\frac{a-\gamma-a}{\sigma}\right) = \Phi\left(\frac{\gamma}{\sigma}\right) - \Phi\left(\frac{-\gamma}{\sigma}\right) = \\
&= \Phi\left(\frac{\gamma}{\sigma}\right) + \Phi\left(\frac{\gamma}{\sigma}\right) = 2\cdot\Phi\left(\frac{\gamma}{\sigma}\right). \qquad \square
\end{aligned}
$$

Kontrollige, et suuruse $X \sim N(a,\sigma)$ korral

$$F(x) = \frac{1}{2} + \Phi\left(\frac{x-a}{\sigma}\right).$$

Kui funktsiooni $\Phi\,(x)$ asemel kasutada funktsiooni

$$\Phi^*(x) \overset{\text{def.}}{=} \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-t^2/2}dt, \tag{2.8.6}$$

siis

$$\Phi(x) = \Phi^*(x) - 0.5$$

ja

$$
\begin{aligned}
P\left(\alpha \leq X \leq \beta\right) &= \Phi\left(\frac{\beta-a}{\sigma}\right) - \Phi\left(\frac{\alpha-a}{\sigma}\right) = \\
&= \left(\Phi^*(\frac{\beta-a}{\sigma}) - 0.5\right) - \left(\Phi^*(\frac{\alpha-a}{\sigma}) - 0.5\right) = \\
&= \Phi^*\left(\frac{\beta-a}{\sigma}\right) - \Phi^*\left(\frac{\alpha-a}{\sigma}\right).
\end{aligned}
$$

Lause 2. Kui $X \sim N(a,\sigma)$, siis

$$P\left(\alpha \leq X \leq \beta\right) = \Phi^*\left(\frac{\beta-a}{\sigma}\right) - \Phi^*\left(\frac{\alpha-a}{\sigma}\right). \tag{2.8.7}$$

Funktsioon $\Phi^*(x)$ ei ole paaris ega paaritu. Tõestage, et $\Phi^*(x)$ rahuldab seost $\Phi^*(-x) = 1 - \Phi^*(x)$.

Kui abifunktsiooni $\Phi\,(x)$ asemel kasutada nn *veafunktsiooni*

$$\operatorname{erf}(x) \overset{\text{def.}}{=} \frac{2}{\sqrt{\pi}}\int_{0}^{x} e^{-t^2}dt \tag{2.8.8}$$

ja arvestada, et $\operatorname{erf}(x)$ on paaritu funktsioon ning

$$\Phi(x) = \frac{1}{2}\operatorname{erf}\left(x/\sqrt{2}\right),$$

siis saame

$$
\begin{aligned}
P(\alpha \leq X \leq \beta) &= \Phi\left(\frac{\beta - a}{\sigma}\right) - \Phi\left(\frac{\alpha - a}{\sigma}\right) = \\
&= \frac{1}{2}\operatorname{erf}\left(\frac{\beta - a}{\sigma}/\sqrt{2}\right) - \frac{1}{2}\operatorname{erf}\left(\frac{\alpha - a}{\sigma}/\sqrt{2}\right) = \\
&= \frac{1}{2}\left(\operatorname{erf}\left((\beta - a)/\left(\sigma\sqrt{2}\right)\right) - \operatorname{erf}\left((\alpha - a)/\left(\sigma\sqrt{2}\right)\right)\right).
\end{aligned}
$$

Lause 3. Kui $X \sim N(a, \sigma)$, siis

$$
P(\alpha \leq X \leq \beta) = \frac{1}{2}\left(\operatorname{erf}\left((\beta - a)/\left(\sigma\sqrt{2}\right)\right) - \operatorname{erf}\left((\alpha - a)/\left(\sigma\sqrt{2}\right)\right)\right). \tag{2.8.9}
$$

Näide 3. Disketi tõrgeteta tööiga X on keskmiselt 24 kuud. Olgu $\sigma_X = 4$. Eeldame, et X allub normaaljaotusele. Leiame tõenäosuse

$$
P(20 \leq X \leq 32).
$$

Valemi (2.8.3) abil saame

$$
\begin{aligned}
P(20 \leq X \leq 32) &= \Phi\left(\frac{32 - 24}{4}\right) - \Phi\left(\frac{20 - 24}{4}\right) = \Phi(2) - \Phi(-1) = \\
&= \Phi(2) + \Phi(1) \overset{\text{Tabel 1}}{\approx} 0.477\,2 + 0.341\,3 = 0.818\,5. \qquad \diamondsuit
\end{aligned}
$$

Näide 4 (k sigma reegel)**.** Avaldame juhusliku suuruse X, mis allub normaaljaotusele parameetritega a ja σ, korral $P(|X - a| \leq k\sigma)$ $(k = 1; 2; 3; 4)$.

Leiame valemi (2.8.4) abil

$$
\begin{aligned}
P(|X - a| \leq \sigma) &= 2 \cdot \Phi\left(\frac{\sigma}{\sigma}\right) = 2 \cdot \Phi(1) \approx 2 \cdot 0.341\,3 = 0.682\,6, \\
P(|X - a| \leq 2\sigma) &= 2 \cdot \Phi\left(\frac{2\sigma}{\sigma}\right) = 2 \cdot \Phi(2) \approx 2 \cdot 0.477\,2 = 0.954\,4, \\
P(|X - a| \leq 3\sigma) &= 2 \cdot \Phi\left(\frac{3\sigma}{\sigma}\right) = 2 \cdot \Phi(3) \approx 2 \cdot 0.498\,65 = 0.997\,3, \\
P(|X - a| \leq 4\sigma) &= 2 \cdot \Phi\left(\frac{4\sigma}{\sigma}\right) = 2 \cdot \Phi(4) \approx 2 \cdot 0.499\,96 \approx 0.999\,9. \qquad \diamondsuit
\end{aligned}
$$

2.9 Markovi ja Tšebõšovi võrratused

Vaatleme järgnevalt Markovi ja Tšebõšovi võrratusi, mida kasutatakse paljude väidete tõestamisel tõenäosusteoorias. Nende võrratuste vahetud rakendused ülesannete lahendamisel annavad suhteliselt jämedad hinnangud.

Lause 1 (*Markovi!võrratus*). Kui juhusliku suuruse X võimalike väärtuste hulk sisaldub hulgas $R^+ \cup \{0\}$ ja $\exists \mathrm{E}X$ ning $\alpha > 0$, siis

$$
P(X \geq \alpha) \leq \frac{\mathrm{E}X}{\alpha} \tag{2.9.1}
$$

ja

$$P\left(X<\alpha\right)>1-\frac{\mathrm{E}X}{\alpha}. \tag{2.9.2}$$

Tõestus. Kuna kehtib järgmine võrduste ja võrratuste ahel

$$\begin{aligned}\mathrm{E}X &= \int_0^{+\infty} x\,f(x)\,dx = \int_0^{\alpha} x\,f(x)\,dx + \int_{\alpha}^{+\infty} x\,f(x)\,dx \geq \\ &\geq \int_{\alpha}^{+\infty} x\,f(x)\,dx \geq \left[\begin{array}{c} x\geq\alpha\,\wedge\,f(x)\geq 0\,\Rightarrow\,xf(x)\geq\alpha f(x)\,\Rightarrow \\ \Rightarrow\int_{\alpha}^{+\infty}x\,f(x)\,dx\geq\alpha\int_{\alpha}^{+\infty}f(x)\,dx\end{array}\right]\geq \\ &\geq \alpha\int_{\alpha}^{+\infty} f(x)\,dx = \alpha P\left(X\geq\alpha\right),\end{aligned}$$

siis kehtib võrratus (2.9.1) ning seega ka võrratus (2.9.2). □

Näide 1. Tunni aja jooksul ületab silda keskmiselt 600 autot. Leiame tõenäosuse, et järgmise tunni jooksul ületab silda vähemalt 1000 autot.

Olgu X järgmise tunni jooksul silda ületavate autode arv. Juhusliku suuruse X võimalike väärtuste hulk sisaldub hulgas $\mathbf{R}^+\cup\{0\}$. Olgu $\alpha>0$. Kuna $\exists\ \mathrm{E}X$, siis on täidetud Lause 1 eeldused. Rakendame Markovi võrratust (2.9.1)

$$P\left(X\geq 1000\right)\leq\frac{600}{1000}=0.6. \qquad \diamondsuit$$

Lause 2 (*Tšebõšovi võrratus*). Kui juhusliku suuruse X korral $\exists\,\mathrm{E}X$ ja $\exists\,\mathrm{D}X$ ning $\alpha>0$, siis

$$P\left(|X-\mathrm{E}X|\geq\alpha\right)\leq\frac{\mathrm{D}X}{\alpha^2} \tag{2.9.3}$$

ja

$$P\left(|X-\mathrm{E}X|<\alpha\right)>1-\frac{\mathrm{D}X}{\alpha^2}. \tag{2.9.0.11}$$

Tõestus. Moodustame juhusliku suuruse $(X-\mathrm{E}X)^2$. Selle suuruse võimalikud väärtused on mittenegatiivsed. Et $\alpha>0\;\Rightarrow\;\alpha^2>0$ ja $\exists\,\mathrm{D}X\;\Leftrightarrow\;\exists\,\mathrm{E}\left(X-\mathrm{E}X\right)^2$, siis juhusliku suuruse $(X-\mathrm{E}X)^2$ korral on rahuldatud Lause 1 eeldused. Rakendame Markovi võrratust (2.9.1) juhusliku suuruse $(X-\mathrm{E}X)^2$ korral

$$P\left(\left(X-\mathrm{E}X\right)^2\geq\alpha^2\right)\leq\frac{\mathrm{E}\left(X-\mathrm{E}X\right)^2}{\alpha^2}. \tag{2.9.5}$$

Kuna sündmused $(X-\mathrm{E}X)^2\geq\alpha^2$ ja $|X-\mathrm{E}X|\geq\alpha$ on võrdsed ning $\mathrm{D}X=\mathrm{E}\left(X-\mathrm{E}X\right)^2$, siis väited (2.9.5) ja (2.9.3) ühtivad. Võrratus (2.9.4) järeldub võrratusest (2.9.3). □

Näide 2. Pere tarbib keskmiselt 300 kWh elektrit kuus. Pere kuise elektritarbe kui juhusliku suuruse X standardhälve ei ületa 60 kWh. Hindame tõenäosust, et sel kuul pere tarbib elektrit vähem kui 500 kWh.

Juhusliku suuruse X korral $\sigma_X \leq 60$ ja $\mathrm{E}X = 300$. Seega $\exists\,\mathrm{D}X \leq 3600$. Olgu $\alpha > 0$. Lause 2 tingimused on rahuldatud. Et

$$|X - \mathrm{E}X| < \alpha \Leftrightarrow -\alpha < X - \mathrm{E}X < \alpha \Leftrightarrow \mathrm{E}X - \alpha < X < \mathrm{E}X + \alpha,$$

st $\mathrm{E}X + \alpha = 500$, siis valime $\alpha = 200$. Kasutame võrratust (2.9.4)

$$P\,(300 - 200 < X < 300 + 200) > 1 - \frac{\mathrm{D}X}{40000}.$$

Kuna $\mathrm{D}X \leq 3600$, siis viimasest võrratusest järeldub

$$P\,(100 < X < 300 + 200) > 1 - \frac{3600}{40000},$$

st

$$P\,(100 < X < 500) > 0.91.$$

Et

$$\begin{aligned} P\,(X < 500) = P\,(0 \leq X \leq 100) + P\,(100 < X < 500) > \\ > P\,(100 < X < 500)\,, \end{aligned}$$

siis

$$P\,(X < 500) > 0.91.$$

Markovi võrratuse (2.9.2) põhjal saame

$$P\,(X < 500) > 1 - \frac{300}{500},$$

st

$$P\,(X < 500) > 0.4.$$

Markovi võrratuse abil saadud kesisema hinnangu põhjus on selles, et Tšebõšovi võrratuse rakendamisel oleme kasutanud täiendavat informatsiooni juhusliku suuruse X kohta, standardhälvet σ_X. $\diamondsuit$

2.10 Tšebõšovi ja Bernoulli piirteoreemid

Lause 1 (*Tšebõšovi piirteoreem*). Kui juhuslikud suurused X_k $(k \in \mathbf{N})$ on sõltumatud ja $\exists\,\mathrm{E}X_k$ ning $\exists\mathrm{D}X_k$, kusjuures $\mathrm{D}X_k \leq M$ $(k \in \mathbf{N})$, siis $\forall\varepsilon > 0$ korral

$$\lim_{n\to\infty} P\left(\left|\frac{X_1 + \ldots + X_n}{n} - \frac{\mathrm{E}X_1 + \ldots + \mathrm{E}X_n}{n}\right| < \varepsilon\right) = 1. \qquad (2.10.1)$$

Tõestus. Kui

$$Y_n \overset{\text{def.}}{=} \frac{1}{n}\sum_{k=1}^{n} X_k,$$

siis

$$\mathrm{E}Y_n = \mathrm{E}\left(\frac{1}{n}\sum_{k=1}^{n} X_k\right) = \frac{1}{n}\sum_{k=1}^{n} \mathrm{E}X_k,$$
$$\mathrm{D}Y_n = \mathrm{D}\left(\frac{1}{n}\sum_{k=1}^{n} X_k\right) = \frac{1}{n^2}\sum_{k=1}^{n} \mathrm{D}X_k \leq \frac{n\cdot M}{n^2} = \frac{M}{n}$$

ja Tšebõšovi võrratuse (2.9.4) rakendamisel Y_n korral saame

$$P\left(|Y_n - \mathrm{E}Y_n| < \alpha\right) > 1 - \frac{\mathrm{D}Y_n}{\alpha^2} \overset{\alpha=\varepsilon}{\Rightarrow}$$

$$\Leftrightarrow P\left(\left|\frac{1}{n}\sum_{k=1}^{n} X_k - \frac{1}{n}\sum_{k=1}^{n} \mathrm{E}X_k\right| < \varepsilon\right) > 1 - \frac{M}{n\varepsilon^2},$$

st

$$\forall\varepsilon > 0:\ P\left(|Y_n - \mathrm{E}Y_n| < \varepsilon\right) > 1 - \frac{M}{n\varepsilon^2}. \qquad (2.10.2)$$

Viimane võrratus on samaväärne väitega (2.10.1). □

Rõhutame Tšebõšovi teoreemi sisu. Kui $\mathrm{D}X_k \leq M\quad (k \in \mathbf{N})$, siis juhuslike sõltumatute suuruste X_k suure arvu n korral on $Y_n = \left(\sum_{k=1}^{n} X_k\right)/n$ praktiliselt mittejuhuslik suurus. Täpsemini, sündmus, et juhuslik suurus Y_n erineb kindlast suurusest $\left(\sum_{k=1}^{n} \mathrm{E}X_k\right)/n$ kuitahes vähe, on peaaegu kindel sündmus.

Näide 1. Mitu korda tuleb kondensaatori mahtuvust mõõta, et tõenäosusega vähemalt 0.9 garanteerida nende sõltumatute mõõtmiste aritmeetilise keskmise kõrvalekalle mitte rohkem kui 2 μF. On teada, et mõõtmisel puudub süstemaatiline viga ja mõõtmistulemuste X_k kui juhuslike suuruste standardhälbed σ_{X_k} ei ületa 10 μF.

Seega suurused X_k on sõltumatud, $Y_n = \left(\sum_{k=1}^{n} X_k\right)/n$ ja

$$\exists\sigma_{X_k} \wedge \sigma_{X_k} \leq 10 \Rightarrow \exists\,\mathrm{E}X_k \wedge \exists\mathrm{D}X_k \wedge \mathrm{D}X_k \leq 100 \quad (k \in \mathbf{N})$$

ning Lause 1 on rakendatav. Hinnangu (2.10.2) põhjal saame $\varepsilon = 2$ ja $M = 100$ korral hinnangu

$$P\left(|Y_n - \mathrm{E}Y_n| < 2\right) > 1 - \frac{100}{n\cdot 2^2}.$$

Seega saame mõõtmiste arvu n määrata:

$$1 - \frac{100}{n\cdot 2^2} \geq 0.9 \quad \Leftrightarrow \quad \frac{100}{n\cdot 2^2} \leq 0.1 \quad \Leftrightarrow \quad n \geq 250. \qquad \diamondsuit$$

Järeldus 1. Kui juhuslikud suurused $X_k\quad (k \in \mathbf{N})$ on sõltumatud ja neil on ühine keskväärtus a ning $\exists\mathrm{D}X_k$, kusjuures $\mathrm{D}X_k \leq M\quad (k \in \mathbf{N})$, siis $\forall\varepsilon > 0$ korral

$$\lim_{n\to\infty} P\left(\left|\frac{X_1 + \ldots + X_n}{n} - a\right| < \varepsilon\right) = 1. \qquad (2.10.3)$$

Definitsioon 1. Öeldakse, et juhuslike suuruste jada $\{Z_n\}$ *koondub tõenäosuse järgi* arvuks γ, kui suvalise $\varepsilon > 0$ korral

$$\lim_{n\to\infty} P\left(|Z_n - \gamma| < \varepsilon\right) = 1.$$

Lause 2 (*Bernoulli piirteoreem*). Kui sõltumatud katsed n-katselises seerias toimuvad ühesugustes tingimustes ja sündmus A toimub selles seerias n_A korda, siis koondub katsete arvu piiramatul suurendamisel sündmuse A toimumise sagedus $P^*(A) = n_A/n$ tõenäosuse järgi sündmuse A toimumise tõenäosuseks $P(A) = p$, st suvalise $\varepsilon > 0$ korral

$$\lim_{n\to\infty} P\left(\left|\frac{n_A}{n} - p\right| < \varepsilon\right) = 1. \tag{2.10.4}$$

Tõestus. Näitame, et Lause 2 väide tuleneb Järeldusest 1. Kui X_k on sündmuse A esinemiste arv k-ndal katsel, siis $\sum_{k=1}^{n} X_k = n_A$ ja $\mathrm{E}X_k = a = p$ ning $\mathrm{D}X_k = pq \leq 1$. Seega on täidetud Järelduse 1 tingimused ja (2.10.4) on väite (2.10.3) erijuht. $\square$

2.11 Tsentraalne piirteoreem

Definitsioon 1. Kui normaaljaotusega $N(0;1)$ juhusliku suuruse jaotusfunktsioon on $n \to \infty$ korral juhuslike suuruste $(X_n - EX_n)/\sqrt{DX_n}$ $(n \in \mathbf{N})$ jaotusfunktsioonide jada piirväärtus, siis öeldakse, et jada $\{X_n\}$ on *asümptootiliselt normaalne.*

Olgu sõltumatud juhuslikud suurused X_k $(k = 1, \ldots, n)$ ühesuguse jaotusega ja eksisteerigu neil karakteristlikud funktsioonid, millel on punkti null ümbruses pidevad tuletised kuni kolmanda järguni. Saab näidata, et neil eeldustel $\exists \mathrm{E}X_k$ ja $\exists \mathrm{E}\left(X_k^2\right)$. Et piirteoreemi oleks lihtsam tõestada, eeldame täiendavalt, et suurused X_k on tsentreeritud, st $\mathrm{E}X_k = 0$. Tegelikult ei ole see kitsendus, sest juhul $\mathrm{E}X_k \neq 0$ võiks teha suuruste vahetuse $W_k = X_k - \mathrm{E}X_k$ ja uurida piirteoreemi tsentreeritud suuruste W_k korral. Kirjutame välja suuruse X_k karakteristliku funktsiooni $g_{X_k}(\omega)$ korral teist järku Maclaurini valemi

$$g_{X_k}(\omega) = g_{X_k}(0) + g'_{X_k}(0)\omega + \frac{g''_{X_k}(0)}{2!}\omega^2 + \frac{g'''_{X_k}(0+\theta\omega)}{3!}\omega^3 =$$

$$= \left[g_{X_k}(0) = 1,\ g'_{X_k}(0) = i\,\mathrm{E}X_k = 0,\ \ g''_{X_k}(0) = i^2\mathrm{E}X_k^2 \overset{\mathrm{E}X_k=0}{=} -\sigma^2\right] =$$

$$= 1 - \frac{\sigma^2\omega^2}{2!} + \frac{g'''_{X_k}(\theta\omega)}{3!}\omega^3 \quad (0 < \theta < 1).$$

Olgu $Y_n = \sum_{k=1}^{n} X_k$. Siis

$$\mathrm{E}Y_n = \mathrm{E}\sum_{k=1}^{n} X_k = \sum_{k=1}^{n} \mathrm{E}X_k = \sum_{k=1}^{n} 0 = 0$$

ja

$$\mathrm{D}Y_n = \mathrm{D}\sum_{k=1}^{n} X_k = [X_k \text{ sõltumatud}] = \sum_{k=1}^{n} \mathrm{D}X_k = \sum_{k=1}^{n} \sigma^2 = n\sigma^2,$$
$$\sigma_{Y_n} = \sqrt{\mathrm{D}Y_n} = \sqrt{n\sigma^2} = \sigma\sqrt{n}$$

ning

$$g_{Y_n}(\omega) = \mathrm{E}e^{i\omega Y_n} = \mathrm{E}e^{i\omega\sum_{k=1}^{n} X_k} = \begin{bmatrix} X_k \text{ sõltumatud } \Leftrightarrow \\ \Leftrightarrow e^{i\omega X_k} \text{ sõltumatud} \end{bmatrix} =$$

$$= \mathrm{E}e^{i\omega X_1} \cdots Ee^{i\omega X_n} = \Pi_{k=1}^{n} g_{X_k}(\omega) =$$

$$= \begin{bmatrix} \text{ühesugused jaotused} \Rightarrow \\ \Rightarrow \text{samad karakterist-} \\ \text{likud funktsioonid} \end{bmatrix} =$$
$$= \left(1 - \frac{\sigma^2\omega^2}{2!} + \frac{g'''_{X_k}(0+\theta\omega)}{3!}\omega^3\right)^n.$$

Kui

$$Z_n \stackrel{def}{=} \frac{Y_n - \mathrm{E}Y_n}{\sigma_{Y_n}} = \frac{Y_n}{\sigma\sqrt{n}},$$

siis Z_n on tsentreeritud ja normeeritud juhuslik suurus, sest

$$\mathrm{E}Z_n = \mathrm{E}\frac{Y_n}{\sigma\sqrt{n}} = 0, \quad \mathrm{D}Z_n = \mathrm{D}\frac{Y_n}{\sigma\sqrt{n}} = \frac{1}{\sigma^2 n}\mathrm{D}Y_n = \frac{\sigma^2 n}{\sigma^2 n} = 1.$$

Lisaks leiame, et

$$g_{Z_n}(\omega) = \mathrm{E}e^{i\omega Z_n} = \mathrm{E}e^{i\omega\frac{Y_n}{\sigma\sqrt{n}}} = \mathrm{E}e^{i\frac{\omega}{\sigma\sqrt{n}}Y_n} = g_{Y_n}(\tfrac{\omega}{\sigma\sqrt{n}}) =$$

$$= \left(1 - \frac{\sigma^2\omega^2}{2\sigma^2 n} + \frac{g'''_{X_k}(0+\theta\frac{\omega}{\sigma\sqrt{n}})}{3!}\left(\tfrac{\omega}{\sigma\sqrt{n}}\right)^3\right)^n =$$

$$= \left[\left(1 - \frac{\omega^2}{2n} + \frac{g'''_{X_k}(0+\theta\frac{\omega}{\sigma\sqrt{n}})}{3!}\frac{\omega^3}{\sigma^3 n\sqrt{n}}\right)^{\frac{1}{-\frac{\omega^2}{2n} + \frac{g'''_{X_k}(0+\theta\frac{\omega}{\sigma\sqrt{n}})}{3!}\frac{\omega^3}{\sigma^3 n\sqrt{n}}}}\right]^{\alpha(n)},$$

kus

$$\alpha(n) = n\left(-\frac{\omega^2}{2n} + \frac{g'''_{X_k}(0+\theta\frac{\omega}{\sigma\sqrt{n}})}{3!}\frac{\omega^3}{\sigma^3 n\sqrt{n}}\right) =$$
$$= -\frac{\omega^2}{2} + \frac{g'''_{X}(0+\theta\frac{\omega}{\sigma\sqrt{n}})}{3!}\frac{\omega^3}{\sigma^3\sqrt{n}}.$$

Et

$$\left(1-\frac{\sigma^2\omega^2}{2n}+\frac{g'''_{X_k}(0+\theta\frac{\omega}{\sigma\sqrt{n}})}{3!}\frac{\omega^3}{\sigma^3 n\sqrt{n}}\right)^{\frac{1}{-\frac{\sigma^2\omega^2}{2n}+\frac{g'''_{X_k}(0+\theta\frac{\omega}{\sigma\sqrt{n}})}{3!}\frac{\omega^3}{\sigma^3 n\sqrt{n}}}}\overset{n\to\infty}{\to} e$$

ja

$$\alpha(n)=-\frac{\omega^2}{2}+\frac{g'''_X(0+\theta\frac{\omega}{\sigma\sqrt{n}})}{3!}\frac{\omega^3}{\sigma^3\sqrt{n}}\overset{n\to\infty}{\to}-\frac{\omega^2}{2},$$

siis

$$g_{Z_n}(\omega)\overset{n\to\infty}{\to}e^{-\omega^2/2}.$$

Kui juhuslik suurus on normaaljaotusega $N(a,\sigma)$, siis selle suuruse karakteristlik funktsioon on kujul $e^{ia\omega-\frac{\sigma^2\omega^2}{2}}$. Seega läheneb piirprotsessis $n\to\infty$ juhusliku suuruse Z_n karakteristlik funktsioon $g_{Z_n}(\omega)$ standardse normaaljaotusega suuruse karakteristlikule funktsioonile. Seega on jadad Y_n ja Z_n asümptootiliselt normaalsed.

Lause 1 (*tsentraalne piirteoreem*)**.** Kui sõltumatud juhuslikud suurused X_k $(k=1,\ldots,n)$ on ühesuguse jaotusega ja neil eksisteerivad karakteristlikud funktsioonid, millel on punkti null ümbruses pidevad tuletised kuni kolmanda järguni, siis juhuslikud jadad Y_n ja Z_n, kus $Y_n=\sum_{k=1}^n X_k$ ja $Z_n=\dfrac{Y_n-\mathrm{E}Y_n}{\sigma_{Y_n}}$, on asümptootiliselt normaalsed.

2.12 Moivre-Laplace'i piirteoreem

Rakendame tsentraalset piirteoreemi binoomjaotuse korral. Olgu X_k sündmuse A toimumiste arv k-ndal katsel ja $X=\sum_{k=1}^n X_k$. Juhuslik suurus X allub binoomjaotusele parameetritega n ja p. Tsentraalse piirteoreemi põhjal on X asümptootiliselt normaalne ja suurus $Z=\dfrac{X-np}{\sqrt{npq}}$ käitub suure katsete arvu n korral ligikaudu nagu normaaljaotusele $N(0;1)$ alluv juhuslik suurus. Seega

$$P(\alpha\le Z<\beta)\approx\Phi(\beta)-\Phi(\alpha).$$

Et

$$(k_1\le X<k_2)\Leftrightarrow(k_1-np\le X-np<k_2-np)\Leftrightarrow$$
$$\Leftrightarrow\left(\frac{k_1-np}{\sqrt{npq}}\le Z<\frac{k_2-np}{\sqrt{npq}}\right),$$

siis

$$P(k_1\le X<k_2)=P\left(\frac{k_1-np}{\sqrt{npq}}\le Z<\frac{k_2-np}{\sqrt{npq}}\right)\approx$$
$$\approx\Phi\left(\frac{k_2-np}{\sqrt{npq}}\right)-\Phi\left(\frac{k_1-np}{\sqrt{npq}}\right),$$

kus

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-t^2/2} dt. \qquad (2.12.0.12)$$

Lause 1 (*Moivre-Laplace'i piirteoreem*). Kui juhuslik suurus X allub binoomjaotusele parameetritega n ja p, siis

$$P(k_1 \le X < k_2) \approx \Phi\left(\frac{k_2 - np}{\sqrt{npq}}\right) - \Phi\left(\frac{k_1 - np}{\sqrt{npq}}\right). \qquad (2.12.2)$$

Märkus 1. Lause 1 väidet kasutatakse tihti kujul

$$P(k_1 \le X \le k_2) \approx \Phi\left(\frac{k_2 - np}{\sqrt{npq}}\right) - \Phi\left(\frac{k_1 - np}{\sqrt{npq}}\right), \qquad (2.12.3)$$

sest binoomjaotuse asendamisel normaaljaotusega tehtav viga on tavaliselt suurem kui $P(X = k_2)$.

Näide 1. Münti visatakse 100 korda. Leiame tõenäosuse, et kullide arv X tuleb viiest viiekümne viieni.

Suurus X on juhuslik ja allub binoomjaotusele parameetritega $n = 100$ ja $p = 0.5$. On vaja leida tõenäosus $P(5 \le X \le 55)$. Kasutame Moivre-Laplace'i piirteoreemi, st lähendame binoomjaotust sama keskväärtust ja dispersiooni omava normaaljaotusega. Valemi (2.12.3) abil saame

$$\begin{aligned} P(5 \le X \le 55) &\approx \Phi\left(\frac{55 - 100 \cdot 0.5}{\sqrt{100 \cdot 0.5 \cdot 0.5}}\right) - \Phi\left(\frac{5 - 100 \cdot 0.5}{\sqrt{100 \cdot 0.5 \cdot 0.5}}\right) = \\ &= \Phi\left(\frac{5}{5}\right) - \Phi\left(\frac{-45}{5}\right) = \Phi(1) - \Phi(-9) = \\ &= \left[\begin{array}{c} \Phi(x) \text{ on} \\ \text{paaritu} \end{array}\right] = \Phi(1) + \Phi(9) \approx \\ &\approx 0.3413 + 0.5000 = 0.841\,3. \qquad \diamondsuit \end{aligned}$$

Näide 2. Seeneline leiab metsast 192 kuuseriisikat. Rästiku andmeil on sel päeval 75% kuuseriisikatest ussitanud. Olgu X seenelise poolt leitud ussitanud kuuseriisikate arv. Leiame $P(140 \le X \le 157)$.

Kui X_k omandab väärtuse 1 ussitanud ja väärtuse 0 mitteusitanud k-nda seene korral, siis $X = \sum_{k=1}^{n} X_k$. Et heal seeneaastal on seeni metsas palju, siis võime eeldada, et ühe ussitanud seene leidmine ei mõjuta teise ussitanud seene leidmise tõenäosust (tegelikult mõjutab, kuid seda väga väikest mõju me ei arvesta), st sündmused $X_k = \nu$ ja $X_i = \varrho$ $(k \ne i)$ on sõltumatud. Seega on suurused X_k sõltumatud ja ussitanud seente arv X allub binoomjaotusele parameetritega $n = 192$ ja $p = 0.75$. Valemi (2.12.3) abil saame

$$\begin{aligned} P(140 \le X \le 157) &= \Phi\left(\frac{157 - 192 \cdot 0.75}{\sqrt{192 \cdot 0.75 \cdot 0.25}}\right) - \Phi\left(\frac{140 - 192 \cdot 0.75}{\sqrt{192 \cdot 0.75 \cdot 0.25}}\right) \approx \\ &\approx \Phi(2.1667) - \Phi(-0.666\,7) \approx \\ &\approx 0.484\,9 + 0.247\,5 = 0.732\,4. \qquad \diamondsuit \end{aligned}$$

Lauset 1 nimetatakse ka *Moivre-Laplace'i integraalseks piirteoreemiks.* Leiame Lause 1 abil tõenäosuse

$$P(X=k)\equiv P\left(k\le X<k+1\right)\approx\Phi\left(\frac{k+1-np}{\sqrt{npq}}\right)-\Phi\left(\frac{k-np}{\sqrt{npq}}\right)=$$

$$=\frac{1}{\sqrt{2\pi}}\int_0^{\frac{k+1-np}{\sqrt{npq}}}e^{-t^2/2}dt-\frac{1}{\sqrt{2\pi}}\int_0^{\frac{k-np}{\sqrt{npq}}}e^{-t^2/2}dt=\frac{1}{\sqrt{2\pi}}\int_{\frac{k-np}{\sqrt{npq}}}^{\frac{k+1-np}{\sqrt{npq}}}e^{-t^2/2}dt\approx$$

$$\approx\frac{1}{\sqrt{2\pi}}e^{-\left(\frac{k-np}{\sqrt{npq}}\right)^2/2}\left(\frac{k+1-np}{\sqrt{npq}}-\frac{k-np}{\sqrt{npq}}\right)=\frac{1}{\sqrt{npq}}\frac{1}{\sqrt{2\pi}}e^{-\left(\frac{k-np}{\sqrt{npq}}\right)^2/2}=$$

$$=\frac{1}{\sqrt{npq}}\varphi\left(\frac{k-np}{\sqrt{npq}}\right),$$

kus

$$\varphi\left(t\right)=\frac{1}{\sqrt{2\pi}}e^{-t^2/2}. \tag{2.12.4}$$

Lause 2. (*Moivre-Laplace'i lokaalne piirteoreem*). Kui juhuslik suurus X allub binoomjaotusele parameetritega n ja p, siis

$$P(X=k)\approx\frac{1}{\sqrt{npq}}\varphi\left(\frac{k-np}{\sqrt{npq}}\right), \tag{2.12.5}$$

kus funktsioon $\varphi\left(t\right)$ on määratud eeskirjaga (2.12.4).

Näide 3. Loterii piletitest üks kümnendik võidab. Mängur ostab 100 piletit. Leiame tõenäosuse, et neist sajast (täpselt) 12 võidab.

Kui X on mänguri võitvate piletite arv, siis X allub binoomjaotusele parameetritega $n=100$ ja $p=0.1$. Valemi (2.12.5) abil saame

$$P(X=12)\approx\frac{1}{\sqrt{100\cdot 0.1\cdot 0.9}}\varphi\left(\frac{12-100\cdot 0.1}{\sqrt{100\cdot 0.1\cdot 0.9}}\right)=$$
$$=\frac{1}{3}\cdot\varphi\left(0.6667\right)\overset{(2.12.4)}{=}\frac{1}{3}\cdot\frac{1}{\sqrt{2\pi}}e^{-0.6667^2/2}\approx 0.1065.$$

Võrdluseks toome SWP abil leitud tulemuse

$$P(X=12)=\mathrm{BinomialDen}\left(12;100,0.1\right)\approx 0.098788 \qquad \diamondsuit$$

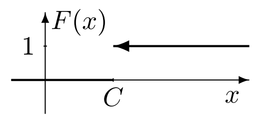
2.13 Ülesanded

1. Kindlat suurust C vaadeldakse kui juhusliku suuruse X erijuhtu. Leidke suuruse C jaotusseadus, jaotusfunktsioon $F(x)$ analüütiliselt ja graafiliselt, jaotustihedus $f(x)$, karakteristlik funktsioon $g(\omega)$, genereeriv funktsioon $G(z)$, keskväärtus EC ja dispersioon DC ning standardhälve σ. V:

x_k	C
p_k	1

,

$$F(x) = \mathbf{1}(x-C) = \begin{cases} 0, \ x \le C, \\ 1, \ x > C, \end{cases}$$



$f(x) = \delta(x-C)$, $g(\omega) = e^{Ci\omega}$, $\mathrm{E}C = C$, $\mathrm{D}C = \sigma = 0$.

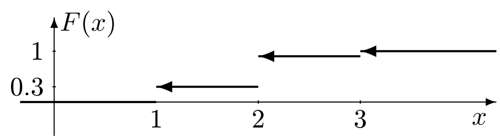
2. Eksamiküsimusi on viis. Tudeng teab neist kolme. Talle esitatakse kolm küsimust. Olgu X küsimuste arv, mida tudeng neist teab. Leidke suuruse X jaotusseadus, $F(x)$ analüütiliselt ja graafiliselt, $f(x)$, $g(\omega)$, $G(z)$, EX, DX ning σ.

V:

x_k	1	2	3
p_k	0.3	0.6	0.1

, $F(x) = 0.3 \cdot \mathbf{1}(x-1) + 0.6 \cdot \mathbf{1}(x-2) + 0.1 \cdot \mathbf{1}(x-3) =$

$$= \begin{cases} 0, \ x \le 1, \\ 0.3, \ 1 < x \le 2, \\ 0.9, \ 2 < x \le 3, \\ 1, \ x > 3, \end{cases}$$



$f(x) = 0.3\delta(x-1) + 0.6\delta(x-2) + 0.1\delta(x-3)$, $\mathrm{E}X = 1.8$, $\mathrm{D}X = 0.36$,
$g(\omega) = 0.3e^{i\omega} + 0.6e^{2i\omega} + 0.1e^{3i\omega}$, $G(z) = 0.3z + 0.6z^2 + 0.1z^3$, $\sigma = 0.6$.

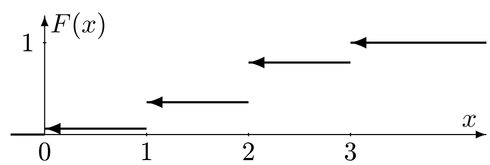
3. Tudeng teab kolme viiendikku eksamiküsimustest. Talle esitatakse kolm küsimust. Olgu X küsimuste arv, mida tudeng neist teab. Leidke juhusliku suuruse X jaotusseadus, $F(x)$ analüütiliselt ja graafiliselt, $f(x)$, $g(\omega)$, $G(z)$, EX, DX, σ ning MoX. V:

x_k	0	1	2	3
p_k	0.064	0.288	0.432	0.216

,

$F(x) = 0.064 \cdot \mathbf{1}(x) + 0.288 \cdot \mathbf{1}(x-1) + 0.432 \cdot \mathbf{1}(x-2) + 0.216 \cdot \mathbf{1}(x-3) =$

$$= \begin{cases} 0, \ x \le 0, \\ 0.064, \ 0 < x \le 1, \\ 0.352, \ 1 < x \le 2, \\ 0.784, \ 2 < x \le 3, \\ 1, \ x > 3, \end{cases}$$



$f(x) = 0.064\delta(x) + 0.288\delta(x-1) + 0.432\delta(x-2) + 0.216\delta(x-3)$,
$g(\omega) = 0.064 + 0.288e^{i\omega} + 0.432e^{2i\omega} + 0.216e^{3i\omega}$, $\mathrm{E}X = 1.8$, $\mathrm{D}X = 0.72$,
$G(z) = 0.064 + 0.288z + 0.432z^2 + 0.216z^3$, $\sigma \approx 0.85$, $\mathrm{Mo}X = 2$.

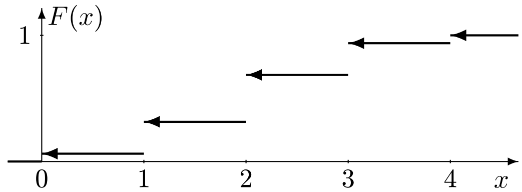
4. Münti visatakse 4 korda. Olgu X saadud kullide koguarv. Leidke juhusliku suuruse X jaotusseadus, $F(x)$ analüütiliselt ja graafiliselt, $f(x)$, $g(\omega)$, $G(z)$, EX, DX, σ ning MoX. V:

x_k	0	1	2	3	4
p_k	1/16	1/4	3/8	1/4	1/16

,

$F(x) = \mathbf{1}(x)/16 + \mathbf{1}(x-1)/4 + 3\cdot\mathbf{1}(x-2)/8 + \mathbf{1}(x-3)/4 + \mathbf{1}(x-4)/16 =$

$$= \begin{cases} 0,\ x \le 0, \\ 1/16,\ 0 < x \le 1, \\ 5/16,\ 1 < x \le 2, \\ 11/16,\ 2 < x \le 3, \\ 15/16,\ 3 < x \le 4 \\ 1,\ x > 4, \end{cases}$$



$f(x) = \delta(x)/16 + \delta(x-1)/4 + 3\cdot\delta(x-2)/8 + \delta(x-3)/4 + \delta(x-4)/16,$
$g(\omega) = 1/16 + e^{i\omega}/4 + 3e^{2i\omega}/8 + e^{3i\omega}/4 + e^{4i\omega}/16,\ \mathrm{E}X = 2,\ \mathrm{D}X = 1,$
$G(z) = 1/16 + z/4 + 3z^2/8 + z^3/4 + z^4/16, \quad \sigma = 1,\ \mathrm{Mo}X = 2.$

5. Märklaua suunas sooritatakse 3 sõltumatut lasku. Igal lasul on tabamise tõenäosus 0.7. Olgu X tabamuste koguarv. Leidke juhusliku suuruse X jaotusseadus, $F(x)$, $f(x)$, $g(\omega)$, $G(z)$, EX, DX ja σ ning $P(X \le 2)$.

V:

x_k	0	1	2	3
p_k	0.0 27	0. 189	0. 441	0. 343

, $\mathrm{E}X = 2.1$, $\mathrm{D}X = 0.63$, $\sigma \approx .794$,
$F(x) = 0.0\,27\cdot\mathbf{1}(x) + 0.\,189\cdot\mathbf{1}(x-1) + 0.\,441\cdot\mathbf{1}(x-2) + 0.\,343\cdot\mathbf{1}(x-3),$
$f(x) = 0.0\,27\cdot\delta(x) + 0.\,189\cdot\delta(x-1) + 0.\,441\cdot\delta(x-2) + 0.\,343\cdot\delta(x-3),$
$g(\omega) = \left(0.7e^{i\omega} + 0.3\right)^3,\ G(z) = (0.7z + 0.3)^3,\ P(X \le 2) = 0.657.$

6. Üliõpilane läheb eksamile, olles 20 küsimusest selgeks õppinud 16. Talle esitatakse 3 küsimust. Olgu X küsimuste arv, mida ta neist kolmest teab. Leidke suuruse X jaotusseadus, EX, DX, σ, $F(x)$ ja $G(z)$.

V:

x_k	0	1	2	3
p_k	1/285	8/95	8/19	28/57

, $\mathrm{E}X = 12/5$, $\mathrm{D}X = 204/475$,
$\sigma \approx 0.\,655$, $F(x) = \mathbf{1}(x)/285 + 8\cdot\mathbf{1}(x-1)/95 + 8\cdot\mathbf{1}(x-2)/19 + 28\cdot\mathbf{1}(x-3)/57,$
$G(z) = 1/285 + 8z/95 + 8z^2/19 + 28z^3/57.$

7. Mängija saab kaardipakist (52 kaarti) 13 kaarti. Olgu X saadud ässade arv (vt Näidet 1.5.4). Leidke suuruse X jaotusseadus, $F(x)$, EX, DX, σ, μ_3, μ_4, As X, Ex X, H (X) ja Mo X.

V:

x_k	0	1	2	3	4
p_k	$\frac{6327}{20\,825}$	$\frac{9139}{20\,825}$	$\frac{4446}{20\,825}$	$\frac{858}{20\,825}$	$\frac{11}{4165}$

,

$F(x) = \frac{6327}{20\,825}\mathbf{1}(x) + \frac{9139}{20\,825}\mathbf{1}(x-1) + \frac{4446}{20\,825}\mathbf{1}(x-2) + \frac{858}{20\,825}\mathbf{1}(x-3) +$
$+\frac{11}{4165}\mathbf{1}(x-4),\ \mathrm{E}X = 1,\ \mathrm{D}X = 12/17,\ \sigma \approx 0.\,840,\ \mu_3 \approx 0\ .311,$
$\mu_4 \approx 1.\,390,\ \mathrm{As}\,X \approx 0\ .524,\ \mathrm{Ex}\,X \approx -0.\,210,\ \mathrm{H}(X) \approx 1.\,200,\ \mathrm{Mo}X = 1.$

8. Juhuslik suurus X allub binoomjaotusele, mille keskväärtus on 6 ja standardhälve 2. Leidke suuruse X jaotusseadus, $F(x)$, $f(x)$, $g(\omega)$ ja $G(z)$.
V: $P(X = k) = C_{18}^k (1/3)^k (2/3)^{18-k} \quad (k = 0; 1; 2; \ldots; 18),$
$F(x) = \sum_{k=0}^{18} C_{18}^k (1/3)^k (2/3)^{18-k}\,\mathbf{1}(x-k),\ g(\omega) = \left(e^{i\omega}/3 + 2/3\right)^{18},$
$f(x) = \sum_{k=0}^{18} C_{18}^k (1/3)^k (2/3)^{18-k}\,\delta(x-k),\ G(z) = (z/3 + 2/3)^{18}.$

9. Sooritatakse n-katseline seeria Bernoulli skeemi järgi. Sündmus A toimub igal katsel tõenäosusega p. Olgu X sündmuse A toimumiste koguarv selle seeria jooksul ja suurus $Y = X/n$. Leidke juhusliku suuruse Y jaotusseadus, $F(y)$, $f(y)$,

$g_Y(\omega)$ ja EY, DY. V: $P\,(Y = k/n) = C_n^k p^k q^{n-k} \quad (k = 0; 1; 2; \ldots ; n)$,
$F\,(y) = \sum_0^n C_n^k p^k q^{n-k}\mathbf{1}(x - k/n)$, $f\,(y) = \sum_0^n C_n^k p^k q^{n-k}\delta(x - k/n)$,
$g_Y(\omega) = \left(pe^{i\omega/n} + q\right)^n$, $G_Y(z) = \left(p\sqrt[n]{z} + q\right)^n$, E$Y = p$, D$Y = pq/n$.

10. Korvpallur sooritab kaks sõltumatut vabaviset. Esimesel viskel on tabamise tõenäosus 0.6 ja teisel 0.8. Olgu $X = X_1 + X_2$, kus X_1 ja X_2 on vastavalt tabamuste arvud esimesel ja teisel viskel. Leidke suuruste X_1 ja X_2 ning X genereerivad funktsioonid, suuruse X jaotusseadus, $F(x)$, $g_X(z)$, EX ja DX.
V: $G_X\,(z) = (0.6z + 0.4)\,(0.8z + 0.2) = 0.\,48z^2 + 0.\,44z + 0.0\,8$,

x_k	0	1	2
p_k	0.08	0, 44	0.48

, $F(x) = 0.0\,8\cdot\mathbf{1}(x) + 0.\,44\cdot\mathbf{1}(x-1) + 0.\,48\cdot\mathbf{1}(x-2)$,
$g_X\,(\omega) = 0.\,48e^{2i\omega} + 0.\,44e^{i\omega} + 0.0\,8$, E$X = 1.4$, D$X = 0.4$.

11. Korvpallur sooritab kaks sõltumatut vabaviset. Esimesel viskel on tabamise tõenäosus p_1 ja teisel p_2. Olgu $X = X_1 + X_2$, kus X_1 ja X_2 on vastavalt tabamuste arvud esimesel ja teisel viskel. Leida $g_X(\omega)$, suuruse X jaotusseadus, $G_X(z)$, $F(x)$, EX ja DX. V: $g_X(\omega) = q_1q_2 + (q_1p_2 + q_2p_1)\,e^{i\omega} + p_1p_2e^{2i\omega}$,

x_k	0	1	2
p_k	q_1q_2	$q_1p_2 + q_2p_1$	p_1p_2

, $G_X(z) = q_1q_2 + (q_1p_2 + q_2p_1)\,z + p_1p_2z^2$,
$F\,(x) = q_1q_2\mathbf{1}(x) + (q_1p_2 + q_2p_1)\,\mathbf{1}(x-1) + p_1p_2\mathbf{1}(x-2)$, E$X = p_1 + p_2$,
D$X = p_1q_1 + p_2q_2$.

12. Igaüks kolmest jalgpallurist sooritab ühe vabalöögi. Esimesel lööjal on tabamise tõenäosus 0.4, teisel 0.6 ja kolmandal 0.8. Löögid sooritatakse üksteisest sõltumatult. Olgu $X = X_1 + X_2 + X_3$, kus X_k on tabamuste arv k-ndal lööjal. Leidke $g_X\,(\omega)$, $G_X(z)$, jaotusseadus, $F(x)$, $f(x)$, EX, DX ja σ.
V: $g_X\,(\omega) = 0\ .0\,48 + 0.\,296e^{i\omega} + 0.\,464e^{2i\omega} + 0.\,192e^{3i\omega}$, $G_X(z) = 0\ .0\,48+$
$+0.\,296z + 0.\,464z^2 + 0.\,192z^3$,

x_k	0	1	2	3
p_k	0.0 48	0. 296	0. 464	0. 192

,
$F(x) = 0\ .0\,48\cdot\mathbf{1}(x) + 0.\,296\cdot\mathbf{1}(x-1) + 0.\,464\cdot\mathbf{1}(x-2) + 0.\,192\cdot\mathbf{1}(x-3)$,
$f(x) = 0\ .0\,48\cdot\delta(x) + 0.\,296\cdot\delta(x-1) + 0.\,464\cdot\delta(x-2) + 0.\,192\cdot\delta(x-3)$,
E$X = 1.8$, D$X = 0.64$, $\sigma = 0.8$.

13. Olgu X sündmuse A toimumiste arv kolmekatselise seeria jooksul. Esimesel katsel on sündmuse A toimumise tõenäosus p_1, teisel p_2 ja kolmandal p_3. Katsed on sõltumatud. Leidke $G(z)$, EX, DX ja suuruse X jaotusseadus.
V: $G(z) = q_1q_2q_3 + (p_1q_2q_3 + q_1p_2q_3 + q_1q_2p_3)\,z + (p_1p_2q_3 + p_1q_2p_3 + q_1p_2p_3)\,z^2 +$
$+p_1p_2p_3z^3$, E$X = p_1 + p_2 + p_3$. D$X = p_1q_1 + p_2q_2 + p_3q_3$,

x_k	0	1	2	3
p_k	$q_1q_2q_3$	$p_1q_2q_3 + q_1p_2q_3 + q_1q_2p_3$	$p_1p_2q_3 + p_1q_2p_3 + q_1p_2p_3$	$p_1p_2p_3$

.

14. Olgu X sündmuse A toimumiste arv n-katselise seeria jooksul ja p_k $(1 \le k \le n)$ sündmuse A toimumise tõenäosus k-ndal katsel. Katsed on sõltumatud. Leidke EX, DX, $P\,(X = 1)$ ja $P\,(X = n-1)$.
V: E$X = \sum_{k=1}^n p_k$, D$X = \sum_{k=1}^n p_kq_k$, $P\,(X = 1) = \sum_{k=1}^n p_k\Pi_{m=1,\,m\neq k}^n q_m$,
$P\,(X = n-1) = \sum_{k=1}^n q_k\Pi_{m=1,\,m\neq k}^n p_m$.

15. Kaks korvpallurit sooritavad üksteisest sõltumatult mõlemad ühe vabaviske. Esimesel on tabamise tõenäosus p_1 ja teisel p_2. Olgu X esimese ja Y teise tabamuste arv. Leidke juhusliku suuruse $Z = Y - X$ jaotusseadus, EZ, DZ ja

$F(z)$. V:

z_k	-1	0	1
p_k	p_1q_2	$p_1p_2+q_1q_2$	q_1p_2

, $\mathrm{E}Z = p_2 - p_1$, $\mathrm{D}Z = p_1q_1 + p_2q_2$,
$F(z) = p_1q_2\mathbf{1}(z+1) + (p_1p_2+q_1q_2)\mathbf{1}(z) + q_1p_2\mathbf{1}(z-1)$.

16. Juhuslik suurus X allub Poissoni jaotusele parameetriga $\lambda = 3$. Leidke juhusliku suuruse X jaotusseadus, $F(x)$, $f(x)$, $g(\omega)$, $G(z)$, $\mathrm{E}X$, $\mathrm{D}X$ ja σ. Leidke tõenäosus, et X omandab katse käigus keskväärtusest väiksema väärtuse. V: $P(X=k) = e^{-3}3^k/k!\ \ (k \in \mathbf{N}_0)$, $F(x) = e^{-3}\sum_{k=0}^{\infty} 3^k/k! \cdot \mathbf{1}(x-k)$, $f(x) = e^{-3}\sum_{k=0}^{\infty} 3^k/k! \cdot \delta(x-k)$, $g(\omega) = e^{3\left(e^{i\omega}-1\right)}$, $G(z) = e^{3(z-1)}$, $\mathrm{E}X = 3$, $\mathrm{D}X = 3$, $\sigma = \sqrt{3}$, $P(X<3) \approx 0.423$.

17. Autojuhti karistatakse keskeltläbi kaks korda aastas kiiruse ületamise eest. Leidke tõenäosus, et teda: 1) karistatakse sel aastal kiiruse ületamise eest täpselt 3 korda; 2) ülimalt 3 korda; 3) vähemalt 3 korda; 4) sel aastal enam ei karistata, kui esimese kuuga karistati juba 2 korda. Karistamiste arv kiiruse ületamise eest allugu Poissoni jaotusele. V: 0.180, ≈ 0.857, ≈ 0.323, ≈ 0.160.

18. Sõltumatud juhuslikud suurused X_1 ja X_2 alluvad Poissoni jaotusele vastavalt keskväärtustega λ_1 ja λ_2. Uurige juhusliku suuruse $X = X_1 + X_2$ jaotust, leides $g_X(\omega)$. Leidke suuruse X jaotusseadus, $G_X(z)$, $F(x)$, $f(x)$, $\mathrm{E}X$ ja $\mathrm{D}X$.
V: $g_X(\omega) = e^{(\lambda_1+\lambda_2)\left(e^{i\omega}-1\right)}$, $P(X=k) = e^{-\lambda_1-\lambda_2}(\lambda_1+\lambda_2)^k/k!\ \ (k \in \mathbf{N}_0)$,
$G_X(z) = e^{(\lambda_1+\lambda_2)(z-1)}$, $F(x) = \sum_{k=0}^{\infty} e^{-\lambda_1-\lambda_2}(\lambda_1+\lambda_2)^k\mathbf{1}(x-k)/k!$,
$f(x) = \sum_{k=0}^{\infty} e^{-\lambda_1-\lambda_2}(\lambda_1+\lambda_2)^k\delta(x-k)/k!$, $\mathrm{E}X = \mathrm{D}X = \lambda_1+\lambda_2$.

19. Korvpallur sooritab pealeviskeid esimese möödaviskeni. Igal viskel on tabamise tõenäosus 0.7. Olgu X sooritatud visete koguarv. Leidke juhusliku suuruse X jaotusseadus, $F(x)$, $f(x)$, $g(\omega)$, $G(z)$, $\mathrm{E}X$, $\mathrm{D}X$ ja σ.
V: $P(X=k) = 0.7^{k-1}\cdot 0.3\ (k \in \mathbf{N})$, $F(x) = 0.3\sum_{k=1}^{\infty} 0.7^{k-1}\mathbf{1}(x-k)$,
$f(x) = 0.3\sum_{k=1}^{\infty} 0.7^{k-1}\delta(x-k)$, $g(\omega) = 0.3e^{i\omega}/\left(1-0.7e^{i\omega}\right)$, $\mathrm{E}X = 10/3$,
$G(z) = 0.3z/(1-0.7z)$, $\mathrm{D}X = 70/9$, $\sigma = \sqrt{70/9} \approx 2.789$.

20. Olgu $P(X=k) = a^k(1+a)^{-k-1}\quad (k \in \mathbf{N}_0,\ a>0)$ juhusliku suuruse X jaotusseadus. Leidke $F(x)$, $f(x)$, $g(\omega)$, $G(z)$, $\mathrm{E}X$, $\mathrm{D}X$, σ ja mood $\mathrm{Mo}\,X$.
V: $F(x) = \sum_{k=0}^{\infty} a^k(1+a)^{-k-1}\mathbf{1}(x-k)$, $f(x) = \sum_{k=0}^{\infty} a^k(1+a)^{-k-1}\delta(x-k)$, $g(\omega) = 1/\left(a+1-ae^{i\omega}\right)$, $G(z) = 1/(a+1-az)$, $\mathrm{E}X = a$, $\mathrm{D}X = a^2+a$, $\sigma = \sqrt{a^2+a}$, $\mathrm{Mo}\,X = 0$.

21. Juhuslik suurus X allub geomeetrilisele jaotusele, st $P(X=k) = pq^{k-1}$ $(k \in \mathbf{N})$, kus $0<p<1$ ja $q = 1-p$. Leidke $F(x)$, $f(x)$, $\mathrm{E}X$, $\mathrm{D}X$, $g(\omega)$ ja $G(z)$.
V: $F(x) = \sum_{k=1}^{\infty} pq^{k-1}\mathbf{1}(x-k)$, $f(x) = \sum_{k=1}^{\infty} pq^{k-1}\delta(x-k)$, $\mathrm{E}X = 1/p$, $\mathrm{D}X = q/p^2$, $g(\omega) = pe^{i\omega}/\left(1-qe^{i\omega}\right)$, $G(z) = pz/(1-qz)$.

22. Diskreetsel juhuslikul suurusel X on vaid kaks võimalikku väärtust x_1 ja x_2 $(x_2 > x_1)$. On teada, et $P(X=x_1) = 0.6$, $\mathrm{E}X = 1.3$ ja $\mathrm{D}X = 0.96$. Leida suuruse X jaotusseadus, $F(x)$, $f(x)$, $g(\omega)$ ja $G(z)$.

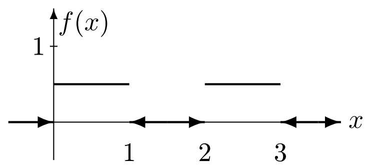
V:

x_k	0.5	2.5
p_k	0.6	0.4

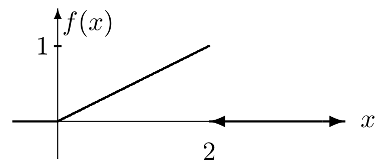
, $F(x) = 0.6\cdot\mathbf{1}(x-0.5) + 0.4\cdot\mathbf{1}(x-2.5)$,
$f(x) = 0.6\cdot\delta(x-0.5) + 0.4\cdot\delta(x-2.5)$, $g(\omega) = 0.6e^{0.5i\omega} + 0.4e^{2.5i\omega}$,
$G(z) = 0.6\sqrt{z} + 0.4\sqrt{z^5}$.

Ülesannetes 23-26 on suuruse X jaotustihedus $f(x)$ antud graafiliselt. Leidke $f(x)$ analüütiliselt, $F(x)$ analüütiliselt ja graafiliselt, $g(\omega)$, $P(0.4 \le X \le 0.8)$, EX, DX ja σ.

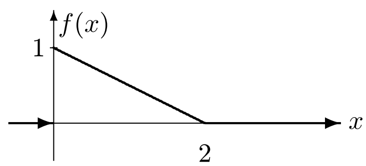
23.



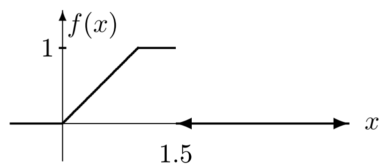
24.



25.

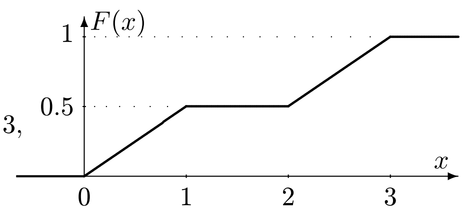


26.



23. V: $f(x) = \begin{cases} 0,\ x < 0 \vee 1 < x < 2 \vee x > 3, \\ 0.5,\ 0 \le x \le 1 \vee 2 \le x \le 3, \end{cases}$ $P(0.4 \le X \le 0.8) = 0.2,$

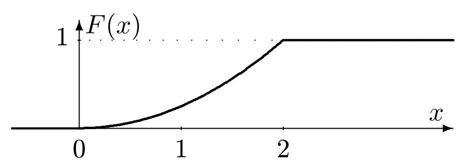
$$F(x) = \begin{cases} 0,\ x < 0, \\ 0.5x,\ 0 \le< x \le 1, \\ 0.5,\ 1 < x < 2, \\ -0.5 + 0.5x,\ 2 \le x \le 3, \\ 1,\ x > 3, \end{cases}$$



$g(\omega) = 0.5i(1 - e^{i\omega} + e^{2i\omega} - e^{3i\omega})/\omega,\ \mathrm{E}X = 1.5,\ \mathrm{D}X = 13/12,\ \sigma \approx 1.041.$

24. V: $f(x) = \begin{cases} 0,\ x < 0 \vee x > 2, \\ 0.5x,\ 0 \le x \le 2 \end{cases}$ $P(0.4 \le X \le 0.8) = 0.12,$

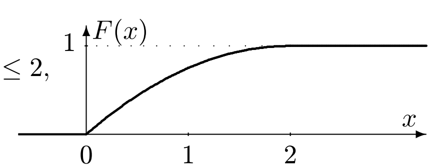
$$F(x) = \begin{cases} 0,\ x < 0, \\ 0.25x^2,\ 0 \le x \le 2, \\ 1,\ x > 2, \end{cases}$$



$g(\omega) = 0.5(e^{2i\omega}(1 - 2i\omega) - 1)/\omega^2,\quad \mathrm{E}X = 4/3,\ \mathrm{D}X = 2/9,\quad \sigma \approx 0.471.$

25. V: $f(x) = \begin{cases} 0,\ x < 0 \vee x > 2, \\ 1 - 0.5x,\ 0 \le x \le 2 \end{cases}$ $P(0.4 \le X \le 0.8) = 0.28,$

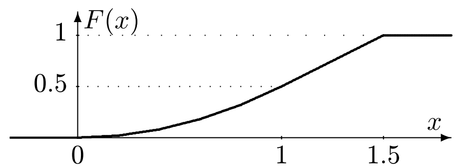
$$F(x) = \begin{cases} 0,\ x < 0, \\ 1 - 0.25(x-2)^2,\ 0 \le x \le 2, \\ 1,\ x > 2, \end{cases}$$



$g(\omega) = 0.5(1 + 2i\omega - e^{2i\omega})/\omega^2,\qquad \mathrm{E}X = 2/3,\ \mathrm{D}X = 2/9,\quad \sigma \approx 0.471.$

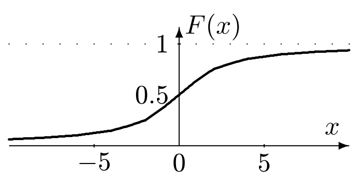
26. V: $f(x) = \begin{cases} 0, \ x < 0 \lor x > 1.5, \\ x, \ 0 \leq x \leq 1, \\ 1, \ 1 < x \leq 1.5, \end{cases}$ $\quad P(0.4 \leq X \leq 0.8) = 0.24,$

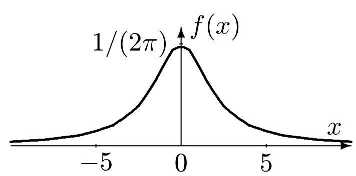
$$F(x) = \begin{cases} 0, \ x < 0, \\ 0.5x^2, \ 0 \leq< x \leq 1, \\ x - 0.5, \ 1 < x < 1.5, \\ 1, \ x > 1.5, \end{cases}$$



$g(\omega) = (e^{i\omega} - 1 - i\omega e - e^{3i\omega/2})/\omega^2, \ \mathrm{E}X = 23/24, \ \mathrm{D}X = 71/576, \ \sigma \approx 0.351.$

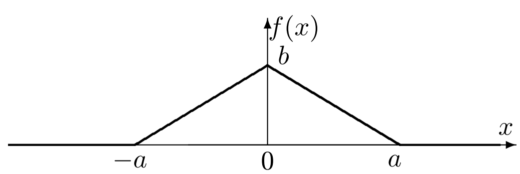
27. Olgu juhusliku suuruse X jaotusfunktsiooniks $F(x) = a + b \arctan{(x/2)}$. Leidke a, b, funktsiooni $F(x)$ graafik, $f(x)$ koos graafikuga, EX ja DX.
V: $a = 1/2, \ b = 1/\pi, \ f(x) = 2/\left(\pi\left(x^2+4\right)\right), \ \mathrm{E}X \stackrel{?}{=} 0, \ \nexists \mathrm{D}X,$





28. Juhuslik suurus X allub Cauchy jaotusele, st $f(x) = \lambda/\left(1+(x-a)^2\right)$. Leidke λ, $F(x)$, EX, DX ja $P\left(-1+a < X < 1+a\right)$. V: $\lambda = 1/\pi$, $\mathrm{E}X \stackrel{?}{=} a$, $\nexists \mathrm{D}X$, $P\left(-1+a < X < 1+a\right) = 0.5$, $F(x) = 0.5 + (1/\pi)\arctan(x-a)$.

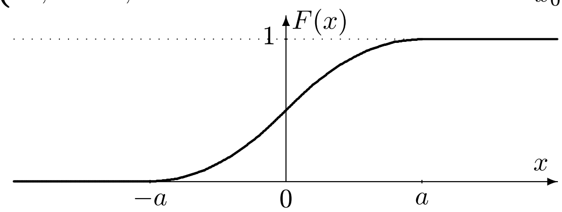
29. Juhuslik suurus X allub Simpsoni jaotusele tihedusega



Leidke $b = b(a)$, $f(x)$ analüütiliselt, $F(x)$ analüütiliselt ja graafiliselt, 0.2-kvantiil, EX, DX, σ ja $P(-a/3 \leq X \leq a/4)$. V: $b = 1/a$,

$$f(x) = \begin{cases} 0, \ x < -a \vee x > a, \\ 1/a + x/a^2, \ -a \le x \le 0, \\ 1/a - x/a^2, \ 0 < x \le a, \end{cases} \quad P(-a/3 \le X \le a/4) = 143/288,$$

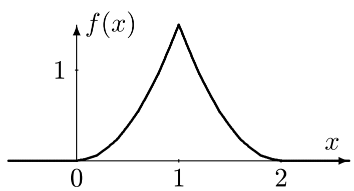
$$F(x) = \begin{cases} 0, \ x < -a, \\ (x+a)/a + 0.5\left(x^2 - a^2\right)/a^2, \ -a \le x \le 0, \\ 0.5 + x/a - 0.5x^2/a^2, \ 0 < x \le a, \\ 1, \ x > a, \end{cases} \quad \begin{array}{r} \mathrm{E}X = 0, \\ \mathrm{D}X = a^2/6, \\ \sigma \approx 0.408a, \\ x_{0.2} \approx -0.368a. \end{array}$$

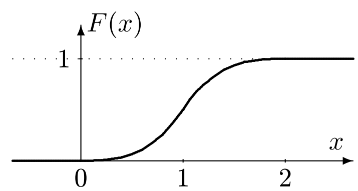


30. Juhusliku suuruse X jaotustihedus on

$$f(x) = \begin{cases} 0, \ \text{kui } x \notin [0;2]\,, \\ ax^2, \ \text{kui } x \in [0;1)\,, \\ a\,(2-x)^2\,, \ \text{kui } x \in [1;2]\,. \end{cases}$$

Leidke a, $f(x)$ graafik, $F(x)$ ja selle graafik, EX, DX, σ ja $P(0 < X < 0.5)$.
V: $a = 3/2$, E$X = 1$, D$X = 1/10$, $\sigma \approx 0.316$, $P(0 < X < 0.5) = 0.0625$, $F(x) =$
$= 0.5x^3\left(\mathbf{1}(x) - \mathbf{1}(x-1)\right) + \left(1 + 0.5\,(x-2)^3\right)\left(\mathbf{1}(x-1) - \mathbf{1}(x-2)\right) + \mathbf{1}(x-2)$,

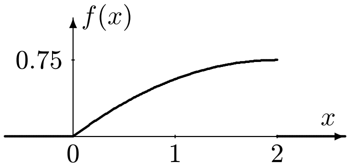


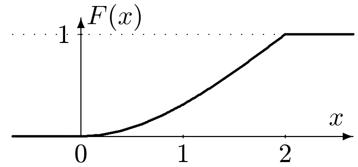


31. Juhusliku suuruse X jaotustihedus on

$$f(x) = \begin{cases} 0, \ \text{kui } x \notin [0;2]\,, \\ \lambda\left(4x - x^2\right), \ \text{kui } x \in [0;2]\,. \end{cases}$$

Leidke λ ja funktsiooni $f(x)$ graafik. Leidke $F(x)$ ja selle graafik, EX, DX ning σ. V: $\lambda = 3/16$, E$X = 5/4$, $DX = 19/80$, $\sigma \approx 0.487$,
$F(x) = \begin{cases} 0, \ x \notin [0;2]\,, \\ 3x^2/8 - x^3/16, \ x \in [0;2]\,, \end{cases}$

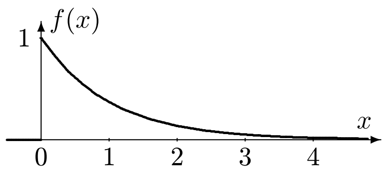


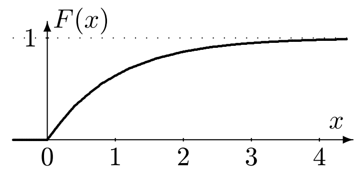


32. Suuruse X jaotusfunktsioon on

$$F(x) = \begin{cases} 0, \text{ kui } x \leq -2, \\ 0.5 + (2/\pi)\arctan(x/2), \text{ kui } -2 < x \leq 2, \\ 1, \text{ kui } x > 2. \end{cases}$$

Leidke $P\left(-1 \leq X \leq \sqrt{3}\right)$, $f(x)$, $\mathrm{Me}\,X$, $\mathrm{Mo}X$, $\mathrm{E}X$, $\mathrm{D}X$ σ ja $\mathrm{H}(X)$.
V: $P\left(-1 \leq X \leq \sqrt{3}\right) \approx 0.750$, $f(x) = 4/\left(\pi\left(4+x^2\right)\right)$, $\mathrm{Me}\,X = \mathrm{Mo}\,X = 0$, $\mathrm{E}\,X = 0$, $\mathrm{D}X = 4(4-\pi)/\pi \approx 1.093$, $\sigma \approx 1.045$, $\mathrm{H}(X) \approx 1.3648$.
33. Suurus X allub ühtlasele jaotusele lõigul $[\alpha, \beta]$ ja Y lõigul $[\beta, \gamma]$. Seejuures on X ja Y sõltumatud. Leidke $\mathrm{D}(XY)$.
V: $\mathrm{D}(XY) = \left(\alpha^2 + \alpha\beta + \beta^2\right)\left(\beta^2 + \gamma\beta + \gamma^2\right)/9 - (\alpha+\beta)^2(\beta+\gamma)^2/16$.
34. Öeldakse, et juhuslik suurus X allub eksponentjaotusele parameetriga $\lambda > 0$, kui $f(x) = \lambda(\exp(-\lambda x))\mathbf{1}(x)$. Leidke $\mathrm{E}X$, $\mathrm{D}X$, σ, $F(x)$, $P(0.5\lambda < X < 2\lambda)$, $\mathrm{Me}\,X$ ja $g(\omega)$. Skitseerige funktsioonide $f(x)$ ja $F(x)$ graafikud $\lambda = 1$ korral.
V: $\mathrm{E}X = 1/\lambda$, $\mathrm{D}X = 1/\lambda^2$, $\sigma = 1/\lambda$, $F(x) = \left(1 - e^{-\lambda x}\right)\mathbf{1}(x)$, $\mathrm{Me}\,X = (\ln 2)/\lambda$, $P(0.5\lambda < X < 2\lambda) = e^{-0.5\lambda^2} - e^{-2\lambda^2}$, $g(\omega) = \left(\lambda^2 + i\lambda\omega\right)/\left(\lambda^2 + \omega^2\right)$,



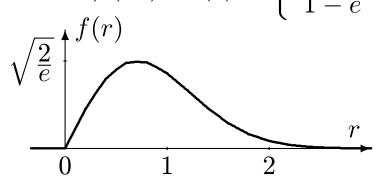


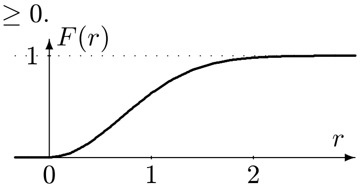
35. Öeldakse, et juhuslik suurus X allub Laplace'i jaotusele parameetritega $\lambda > 0$ ja a, kui $f(x) = (\lambda/2)\exp(-\lambda|x-a|)$. Leidke $\mathrm{E}X$, $\mathrm{D}X$, σ, $\mathrm{Mo}X$ ja $F(x)$. V: $\mathrm{E}X = a$, $\mathrm{D}X = 2/\lambda^2$, $\sigma = \sqrt{2}/\lambda$, $\mathrm{Mo}X = a$,
$F(x) = \begin{cases} 0.5\exp(\lambda(x-a)), \ x < a, \\ 1 - 0.5e^{\lambda(a-x)}, \ x \geq a. \end{cases}$

36. Juhuslik suurus R allub Rayleigh' jaotusele tihedusega

$$f(r) = \begin{cases} Ar\exp\left(-h^2r^2\right), \text{ kui } r \geq 0, \\ 0, \text{ kui } r < 0, \end{cases}$$

kus $h > 0$. Leidke konstant A, $\mathrm{E}R$, $\mathrm{E}R^2$, $\mathrm{D}R$, $\mathrm{Mo}R$ ja $F(r)$. Skitseerige funktsioonide $f(r)$ ja $F(r)$ graafikud $h = 1$ korral.
V : $A = 2h^2$, $\mathrm{E}R = \sqrt{\pi}/(2h)$, $\mathrm{E}R^2 = 1/h^2$, $\mathrm{D}R = (4-\pi)/\left(4h^2\right)$,
$\mathrm{Mo}R = \sqrt{2}/(2h)$, $F(r) = \begin{cases} 0, \ r < 0, \\ 1 - e^{-h^2r^2}, \ r \geq 0. \end{cases}$





37. Nooruki pikkus L on juhuslik suurus, mis allub normaaljaotusele. Politseikooli kursandiks sobib nooruk, kelle pikkus kuulub lõiku $[l_1, l_2]$. On teada, et $\mathrm{E}L = (l_1 + l_2)/2$ ja $\mathrm{D}L = (l_2 - l_1)^2/4$. Leidke tõenäosus, et huupi valitud nooruk sobib pikkuse poolest politseikooli kursandiks. V: ≈ 0.683.
38. On teada, et laserkaugusmõõtja viga on juhuslik suurus, mis allub normaaljaotusele keskväärtusega 2 m ja standardhälbega 4 m. Leidke tõenäosus, et mõõtmisel tekkiv viga kuulub lõiku $[-2; 10]$. V: ≈ 0.819.
39. Juhuslik suurus X allub normaaljaotusele keskväärtusega 1. Suuruse X lõiku $[0; 2]$ sattumise tõenäosus on $0,5$. Leidke σ, $f(x)$ ja $F(x)$. V: $\sigma \approx 1.481$, $f(x) \approx 0.269 \exp\left(-0.228\,(x-1)^2\right)$, $F(x) \approx 0.269 \int_{-\infty}^{x} \exp\left(-0.228\,(t-1)^2\right) dt$.
40. Kuullaagri kuuli diameeter D on juhuslik suurus, kusjuures $\mathrm{E}D = 5$ mm ja $\sigma = 0.05$ mm. Praagitakse kuulid, mille diameeter erineb ettenähtud viiest millimeetrist rohkem kui 0.1 mm. Mitu protsenti kuulidest praagitakse mittesobiva diameetri tõttu? V: $\approx 4.6\%$.
41. Juhuslik suurus allub normaaljaotusele keskväärtusega 0. Olgu $0 \notin [\alpha, \beta]$. Millisel standardhälbe σ väärtusel on lõiku $[\alpha, \beta]$ sattumise tõenäosus suurim? V: $\sqrt{0.5\,(\beta^2 - \alpha^2)\,/\,(\ln(\beta/\alpha))}$.
42. Korvpallur tabab vabaviske tõenäosusega 0.6. Nädala jooksul on ta sooritanud 2400 vabaviset. Leidke tõenäosus, et neist on ta tabanud romkem kui 999 ja vähem kui 1465. V: Moivre-Laplace'i valemi abil ≈ 0.841, Bernoulli valem + SWP ≈ 0.846.
43. Tudeng jõuab hommikul õigeaegselt loengule tõenäosusega 0.8. Leida tõenäosus, et tuhande kuuesajast tudengist jõuab õigeaegselt loengule vähemalt 1300. V: Moivre-Laplace'i valemi abil ≈ 0.106, Bernoulli valem + SWP ≈ 0.111.

Peatükk 3

Juhuslikud vektorid

3.1 Juhusliku vektori jaotusfunktsioon ja jaotustihedus

Definitsioon 1. Vektorit $(X_1, X_2, \ldots, X_n)$, mille komponentideks on juhuslikud suurused X_k $(k = 1; 2; \ldots; n)$, nimetatakse *juhuslikuks vektoriks.*

Katse tulemusena omandab juhuslik vektor $(X_1, X_2, \ldots, X_n)$ realisatsiooni ehk väärtuse, milleks on kindel vektor $(x_1, x_2, \ldots, x_n)$. Vektorit $(x_1, x_2, \ldots, x_n)$ nimetatakse juhusliku vektori $(X_1, X_2, \ldots, X_n)$ *võimalikuks väärtuseks.* Kui juhusliku vektori võimalike väärtuste hulk on kas lõplik või loenduv, siis nimetatakse seda juhuslikku vektorit *diskreetseks.*

Definitsioon 2. Funktsiooni

$$F(x_1, x_2, \ldots, x_n) \overset{\text{def.}}{=} P((X_1 < x_1)(X_2 < x_2) \cdots (X_n < x_n)) \tag{3.1.1}$$

nimetatakse juhusliku vektori $(X_1, X_2, \ldots, X_n)$ *jaotusfunktsiooniks.*

Definitsioon 3. Vektorit $(X_1, X_2, \ldots, X_n)$ nimetatakse *pidevaks juhuslikuks vektoriks,* kui tema jaotusfunktsioon $F(x_1, x_2, \ldots, x_n)$ on pidev funktsioon hulgal $\mathbf{R}^n$.

Kuna juhusliku vektori jaotusfunktsioon defineeritakse kui sündmuse tõenäosus, siis $0 \leq F(x_1, x_2, \ldots, x_n) \leq 1$. Et $\Delta x_1 \geq 0$ korral

$$F(x_1 + \Delta x_1, x_2, \ldots, x_n) = P((X_1 < x_1 + \Delta x_1)(X_2 < x_2) \cdots (X_n < x_n)) =$$

$$= P(((X_1 < x_1) + (x_1 \leq X_1 < x_1 + \Delta x_1))(X_2 < x_2) \cdots (X_n < x_n)) =$$

$$= P((X_1 < x_1)(X_2 < x_2) \cdots (X_n < x_n)) +$$

$$+ P((x_1 \leq X_1 < x_1 + \Delta x_1)(X_2 < x_2) \cdots (X_n < x_n)) \geq$$

$$\geq P((X_1 < x_1)(X_2 < x_2) \cdots (X_n < x_n)) = F(x_1, x_2, \ldots, x_n),$$

siis jaotusfunktsioon $F(x_1, x_2, \ldots, x_n)$ on muutuja x_1 järgi monotoonselt kasvav. Analoogiliselt saab näidata, et $F(x_1, x_2, \ldots, x_n)$ on ka muutujate x_k $(k = 2; 3; \ldots; n)$ järgi monotoonselt kasvav funktsioon. Kuna sündmus

$$(X_1 < -\infty)(X_2 < x_2)\cdots(X_n < x_n)$$

on võimatu, siis $F(-\infty, x_2, \ldots, x_n) = 0$. Analoogiliselt saame

$$F(x_1, -\infty, \ldots, x_n) = \ldots = F(x_1, x_2, \ldots, -\infty) = 0.$$

Seosest $(X_1 < x_1)(X_2 < +\infty)\cdots(X_n < +\infty) = (X_1 < x_1)$ järeldub

$$F(x_1, +\infty, \ldots, +\infty) = F_1(x_1),$$

kus $F_1(x_1)$ on vektori $(X_1, X_2, \ldots, X_n)$ esimese komponendi X_1 jaotusfunktsioon. Analoogiliselt saame

$$F(+\infty, x_2, +\infty, \ldots, +\infty) = F_2(x_2), \ldots, F(+\infty, +\infty, \ldots, +\infty, x_n) = F_n(x_n)$$

ja

$$F(+\infty, +\infty, \ldots, +\infty) = 1.$$

Sõnastame saadud tulemused.

Lause 1. Kui $F(x_1, x_2, \ldots, x_n)$ on juhusliku vektori $(X_1, X_2, \ldots, X_n)$ jaotusfunktsioon, siis:
1° $0 \le F(x_1, x_2, \ldots, x_n) \le 1$;
2° $F(x_1, x_2, \ldots, x_n)$ on iga muutuja x_k järgi monotoonselt kasvav funktsioon;
3° $F(-\infty, x_2, \ldots, x_n) = F(x_1, -\infty, \ldots, x_n) = \ldots = F(x_1, x_2, \ldots, -\infty) = 0$;
4° $F(x_1, +\infty, \ldots, +\infty) = F_1(x_1)$, $F(+\infty, x_2, +\infty, \ldots, +\infty) = F_2(x_2)$, $\ldots$, $F(+\infty, +\infty, \ldots, +\infty, x_n) = F_n(x_n)$;
5° $F(+\infty, +\infty, \ldots, +\infty) = 1$.

Iga Lause 1 tingimusi rahuldavat funktsiooni $F(x_1, \ldots, x_n)$ võime käsitleda kui mingi juhusliku vektori $(X_1, \ldots, X_n)$ jaotusfunktsiooni.

Kui diskreetse juhusliku vektori $(X_1, X_2, \ldots, X_n)$ korral on teada selle vektori võimalike väärtuste hulk

$$\{(x_{i_1}, \ldots, x_{i_n})\}_{i_k \in I_k}, \tag{3.1.2}$$

kus $k = 1; 2; \ldots; n$ ja hulgad I_k on kas lõplikud või lõpmatud naturaalarvude hulga $\mathbf{N}$ osahulgad, ning on teada tõenäosused

$$P((X_1, \ldots, X_n) = (x_{i_1}, \ldots, x_{i_n})) = p_{i_1, \ldots, i_n} \quad \left(\sum_{i_1 \in I_1} \cdots \sum_{i_n \in I_n} p_{i_1, i_2, \ldots, i_n} = 1\right), \tag{3.1.3}$$

millega juhuslik vektor need võimalikud väärtused omandab, siis öeldakse, et on teada selle *juhusliku vektori jaotusseadus.* Kehtib järgmine väide. Miks?

Lause 2. Jaotusseadusega

$$P((X_1, \ldots, X_n) = (x_{i_1}, \ldots, x_{i_n})) = p_{i_1, \ldots, i_n} \ \ (i_k \in I_k \ \ (k = 1; \ldots; n)) \tag{3.1.4}$$

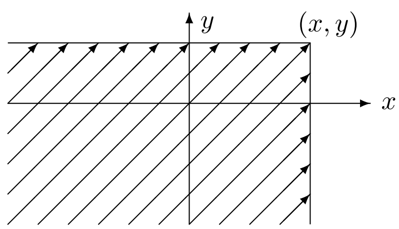
antud diskreetse juhusliku vektori $(X_1, \ldots, X_n)$ jaotusfunktsioon on esitatav kujul

$$F(x_1, \ldots, x_n) = \sum_{i_1 \in I_1} \cdots \sum_{i_n \in I_n} p_{i_1, i_2, \ldots, i_n} \mathbf{1}(x_1 - x_{i_1}) \cdots \mathbf{1}(x_n - x_{i_n}) . \tag{3.1.5}$$

Me piirdume järgnevalt tehnilistel kaalutlustel põhiliselt juhuga $n = 2$. Sel korral on otstarbekas kasutada juhusliku vektori jaoks tähistust (X, Y), kusjuures

$$F(x, y) = P((X < x)(Y < y)), \tag{3.1.6}$$

st jaotusfunktsiooni väärtus $F(x, y)$ annab tõenäosuse, et vektor (X, Y) omandab katsel väärtuse viirutatud piirkonnast



Juhu $n = 2$ jaoks saame Lausele 1 järgmise kuju.

Järeldus 1. Kui $F(x, y)$ on juhusliku vektori (X, Y) jaotusfunktsioon, siis:
1° $0 \leq F(x, y) \leq 1$;
2° $F(x, y)$ on nii muutuja x kui ka muutuja y järgi monotoonselt kasvav funktsioon;
3° $F(-\infty, y) = F(x, -\infty) = 0$;
4° $F(x, +\infty) = F_1(x)$, $F(+\infty, y) = F_2(y)$;
5° $F(+\infty, +\infty) = 1$.

Kui diskreetse juhusliku vektori (X, Y) korral on teada selle vektori jaotusseadus

$$P((X, Y) = (x_i, y_j)) = p_{i,j} \quad (i \in I, j \in J), \tag{3.1.7}$$

kus $\{(x_i, y_j)\}_{i \in I, j \in J}$ on selle vektori võimalike väärtuste hulk (hulgad I ja J on kas lõplikud või lõpmatud naturaalarvude hulga $\mathbf{N}$ osahulgad), siis selle vektori (X, Y) jaotusfunktsioon on esitatav kujul

$$F(x, y) = \sum_{i \in I} \sum_{j \in J} p_{i,j} \mathbf{1}(x - x_i) \mathbf{1}(y - y_j) . \tag{3.1.8}$$

Juhul

$$I = \{1; 2; \ldots; n\}, \ J = \{1; 2; \ldots; m\} \tag{3.1.9}$$

saame jaotusseaduse (3.1.7) esitada tabelina

$x_i \backslash y_j$	y_1	y_2	$\cdots$	y_m
x_1	$p_{1,1}$	$p_{1,2}$	$\cdots$	$p_{1,m}$
x_2	$p_{2,1}$	$p_{2,2}$	$\cdots$	$p_{2,m}$
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
x_n	$p_{n,1}$	$p_{n,2}$	$\cdots$	$p_{n,m}$

(3.1.10)

Näide 1. Sooritatakse üks vabavise. Tabamise tõenäosus on 0.7. Olgu X tabamuste arv ja Y möödavisete arv. Moodustame juhusliku vektori (X, Y). Leiame selle vektori jaotusseaduse ja jaotusfunktsiooni.

Selle juhusliku vektori võimalikeks väärtusteks on $(1; 0)$ ja $(0; 1)$, kusjuures $P((X, Y) = (1; 0)) = 0.7$ ning $P((X, Y) = (0; 1)) = 0.3$. Selle vektori võimalike väärtuste hulgaks on lõplik hulk $\{(1; 0), (0; 1)\}$. Selle jaotusseaduse esitamiseks tabeli kujul laiendame vektori võimalike väärtuste hulka „väärtustega" $(0; 0)$ ja $(1; 1)$, mida see vektor omandab katse käigus tõenäosusega null. Seega on (X, Y) diskreetne juhuslik vektor jaotusseadusega

$x_i \backslash y_j$	0	1
0	0	0.3
1	0.7	0

Valemi (3.1.8) abil saame

$$F(x, y) = \sum_{i=1}^{2} \sum_{j=1}^{2} p_{i,j} \mathbf{1}(x - x_i) \mathbf{1}(y - y_j) = 0 \cdot \mathbf{1}(x - 0) \mathbf{1}(y - 0) +$$

$$+0.3 \cdot \mathbf{1}(x - 0) \mathbf{1}(y - 1) + 0.7 \cdot \mathbf{1}(x - 1) \mathbf{1}(y - 0) + 0 \cdot \mathbf{1}(x - 1) \mathbf{1}(y - 1) =$$

$$= 0.3 \cdot \mathbf{1}(x) \mathbf{1}(y - 1) + 0.7 \cdot \mathbf{1}(x - 1) \mathbf{1}(y)$$

ehk

$$F(x, y) = \begin{cases} 0, \text{ kui } (x \leq 0) \vee (y \leq 0) \vee (x \leq 1 \wedge y \leq 1), \\ 0.3, \text{ kui } (0 < x \leq 1) \wedge (y > 1), \\ 0.7, \text{ kui } (0 < y \leq 1) \wedge (x > 1), \\ 1, \text{ kui } (x > 1) \wedge (y > 1). \qquad \diamondsuit \end{cases}$$

Lause 3. Kehtivad seosed

$$P((X < x)(y_1 \leq Y < y_2)) = F(x, y_2) - F(x, y_1), \tag{3.1.11}$$

$$P((x_1 \leq X < x_2)(Y < y)) = F(x_2, y) - F(x_1, y) \tag{3.1.12}$$

ja

$$\begin{gathered} P((x_1 \leq X < x_2)(y_1 \leq Y < y_2)) = \\ = F(x_2, y_2) - F(x_2, y_1) - F(x_1, y_2) + F(x_1, y_1). \end{gathered} \tag{3.1.13}$$

Tõestus. Et $y_1 < y_2$ korral

$$\begin{aligned} F(x, y_2) = P((X < x)(Y < y_2)) = \\ = P((X < x)((Y < y_1) + (y_1 \leq Y < y_2))) = \\ = P((X < x)(Y < y_1) + (X < x)(y_1 \leq Y < y_2)) = \\ = P((X < x)(Y < y_1)) + P((X < x)(y_1 \leq Y < y_2)) = \\ = F(x, y_1) + P((X < x)(y_1 \leq Y < y_2)), \end{aligned}$$

st

$$F(x, y_2) = F(x, y_1) + P((X < x)(y_1 \le Y < y_2)),$$

siis kehtib valem (3.1.11). Kuna $x_1 < x_2$ korral

$$\begin{aligned} F(x_2, y) &= P((X < x_2)(Y < y)) = \\ &= P(((X < x_1) + (x_1 \le X < x_2))(Y < y)) = \\ &= P((X < x_1)(Y < y) + (x_1 \le X < x_2)(Y < y)) = \\ &= P((X < x_1)(Y < y)) + P((x_1 \le X < x_2)(Y < y)) = \\ &= F(x_1, y) + P((x_1 \le X < x_2)(Y < y)), \end{aligned}$$

siis kehtib valem (3.1.12). Et $(x_1 < x_2) \wedge (y_1 < y_2)$ korral

$$F(x_2, y_2) = P((X < x_2)(Y < y_2)) =$$

$$= P(((X < x_1) + (x_1 \le X < x_2))((Y < y_1) + (y_1 \le Y < y_2))) =$$

$$= P((X < x_1)(Y < y_1)) + P((X < x_1)(y_1 \le Y < y_2)) +$$

$$+P((x_1 \le X < x_2)(Y < y_1)) + P((x_1 \le X < x_2)(y_1 \le Y < y_2)) =$$

$$= [\text{kasutame valemeid (3.1.6), (3.1.11) ja (3.1.12)}] =$$

$$= F(x_1, y_1) + (F(x_1, y_2) - F(x_1, y_1)) + (F(x_2, y_1) - F(x_1, y_1)) +$$

$$+P((x_1 \le X < x_2)(y_1 \le Y < y_2)) =$$

$$= F(x_1, y_2) + F(x_2, y_1) - F(x_1, y_1) + P((x_1 \le X < x_2)(y_1 \le Y < y_2)). \square$$

Järeldus 2. Pidev juhuslik vektor (X, Y) omandab väärtuse hulgast

$$L = \{(x, y) | x = x_1 \wedge y_1 \le y \le y_2\}$$

tõenäosusega 0.

Tõestus. Kasutame sirglõigu L korral valemit (3.1.13). Kui $x_2 = x_1 + \Delta x$, siis

$$P((x_1 \le X < x_1 + \Delta x)(y_1 \le Y < y_2)) \overset{(3.1.13)}{=}$$

$$= F(x_1 + \Delta x, y_2) - F(x_1 + \Delta x, y_1) - F(x_1, y_2) + F(x_1, y_1)$$

ja

$$P((X, Y) \in L) = \lim_{\Delta x \to 0+} P((x_1 \le X < x_1 + \Delta x)(y_1 \le Y < y_2)) =$$

$$= \lim_{\Delta x \to 0+} (F(x_1 + \Delta x, y_2) - F(x_1 + \Delta x, y_1) - F(x_1, y_2) + F(x_1, y_1)) \overset{F \text{ pidev}}{=}$$
$$= F(x_1, y_2) - F(x_1, y_1) - F(x_1, y_2) + F(x_1, y_1) = 0. \quad \square$$

Märkus 1. Järeldus 1 jääb kehtima, kui sirglõik L asendada suvalise tükati sileda joonega L.

Märkus 2. Kui (X, Y) on pidev juhuslik vektor ja D on sidus hulk xy-tasandil, siis

$$P((X,Y) \in D) = P\left((X,Y) \in \overline{D}\right),$$

kus $\overline{D}$ on hulga D sulund, st hulga D ja tema rajapunktide ühend.

3.2 Juhusliku vektori jaotustihedus

Definitsioon 1. Funktsiooni

$$f(x,y) \stackrel{\text{def.}}{=} \lim_{(\Delta x, \Delta y) \to (0+;0+)} \frac{P((x \le X < x + \Delta x)(y \le Y < y + \Delta y))}{\Delta x \, \Delta y} \tag{3.2.1}$$

nimetatakse *juhusliku vektori* (X, Y) *jaotustiheduseks.*

Et Definitsioonis 1 esineva murru lugeja ja nimetaja on mittenegatiivsed, siis on murd mittenegatiivne ning ka piirväärtuse vastava omaduse põhjal

$$f(x,y) \ge 0.$$

Definitsiooni 1 põhjal võib väita, et suurus $f(x,y)\,dxdy$ määrab tõenäosuse dP, et juhuslik vektor (X, Y) omandab väärtuse lõpmata väikesest ristkülikust

$$[x, x + dx] \times [y, y + dy].$$

Seega $dP = f(x,y)\,dxdy$. Summeerides tõenäosused, saame xy-tasandi suvalise piirkonna D korral

$$\iint_D dP = \iint_D f(x,y)\,dxdy.$$

Kuna $\iint_D dP = P((X,Y) \in D)$, siis

$$P((X,Y) \in D) = \iint_D f(x,y)\,dxdy. \tag{3.2.2}$$

Lause 1. Kehtib seos

$$f(x,y) = \frac{\partial^2 F(x,y)}{\partial x \partial y}. \tag{3.2.3}$$

Tõestus. Lähtudes Definitsioonist 1, saame

$$f(x,y) = \lim_{(\Delta x, \Delta y) \to (0+;0+)} \frac{P((x \le X < x + \Delta x)(y \le Y < y + \Delta y))}{\Delta x \, \Delta y} \stackrel{(3.1.13)}{=}$$

$$= \lim_{\Delta y \to 0+} \lim_{\Delta x \to 0+} \frac{F(x+\Delta x, y+\Delta y) - F(x+\Delta x, y) - F(x, y+\Delta y) + F(x,y)}{\Delta x \Delta y}$$

$$= \lim_{\Delta y \to 0+} \frac{1}{\Delta y} \left(\lim_{\Delta x \to 0+} \frac{F(x+\Delta x, y+\Delta y) - F(x, y+\Delta y)}{\Delta x} - \right.$$

$$\left. - \lim_{\Delta x \to 0+} \frac{F(x+\Delta x, y) - F(x,y)}{\Delta x} \right) =$$

$$= \lim_{\Delta y \to 0+} \frac{\dfrac{\partial F(x, y+\Delta y)}{\partial x} - \dfrac{\partial F(x,y)}{\partial x}}{\Delta y} = \frac{\partial^2 F(x,y)}{\partial x \partial y}. \qquad \square$$

Lause 2. Kehtib seos

$$F(x,y) = \int_{-\infty}^{x} du \int_{-\infty}^{y} f(u,v)\, dv. \tag{3.2.4}$$

Tõestus. Kuna

$$\frac{\partial^2 \left(\int_{-\infty}^{x} du \int_{-\infty}^{y} f(u,v)\, dv \right)}{\partial x \partial y} = \frac{\partial}{\partial y} \left(\frac{\partial}{\partial x} \int_{-\infty}^{x} du \int_{-\infty}^{y} f(u,v)\, dv \right) =$$
$$= \frac{\partial}{\partial y} \left(\int_{-\infty}^{y} f(x,v)\, dv \right) = f(x,y)$$

ja valemiga (3.2.4) määratud $F(x,y)$ rahuldab kõiki Järelduse 3.1.1 tingimusi, siis valem (3.2.4) määrab jaotustihedusele $f(x,y)$ vastava jaotusfunktsiooni $F(x,y)$. $\square$

Märkus 1. Juhusliku vektori $(X_1, X_2, \ldots, X_n)$ korral kehtivad seosed

$$f(x_1, \ldots, x_n) = \frac{\partial^n F(x_1, \ldots, x_n)}{\partial x_1 \cdots \partial x_n}$$

ja

$$F(x_1, \ldots, x_n) = \int_{-\infty}^{x_1} du_1 \ldots \int_{-\infty}^{x_n} f(u_1, \ldots, u_n)\, du_n,$$

kus

$$f(x_1, \ldots, x_n) \stackrel{\text{def.}}{=} \lim_{(\Delta x_1, \ldots, \Delta x_n) \to (0+; \ldots; 0+)} \frac{P\left(\prod_{k=1}^{n} (x_k \le X_k < x_k + \Delta x_k) \right)}{\Delta x_1 \cdots \Delta x_n}.$$

Lause 3. Kehtivad seosed

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x,y)\, dx dy = 1, \tag{3.2.5}$$

$$\int_{-\infty}^{+\infty} f(x,y)\,dy = f_1(x) \tag{3.2.6}$$

ja

$$\int_{-\infty}^{+\infty} f(x,y)\,dx = f_2(y), \tag{3.2.7}$$

kus $f_1(x)$ ja $f_2(y)$ on vektori komponentide jaotustihedused, nn *marginaalsed jaotustihedused*.

Tõestus. Seose (3.2.4) ja Järelduse 3.1.1 abil saame

$$\begin{aligned}
&F(x,y) = \int_{-\infty}^{x} du \int_{-\infty}^{y} f(u,v)\,dv \ \wedge \ F(+\infty,+\infty) = 1 \ \Rightarrow \ (3.2.5)\,,\\
&F(x,y) = \int_{-\infty}^{x} du \int_{-\infty}^{y} f(u,v)\,dv \ \wedge \ F(x,+\infty) = F_1(x) \ \Rightarrow\\
&\qquad \Rightarrow \ F_1(x) = F(x,+\infty) = \int_{-\infty}^{x} du \int_{-\infty}^{+\infty} f(u,y)\,dy \ \Rightarrow\\
&\qquad \Rightarrow \ f_1(x) = \frac{d}{dx}F_1(x) = \frac{d}{dx}\int_{-\infty}^{x} du \int_{-\infty}^{+\infty} f(u,y)\,dy =\\
&\qquad = \int_{-\infty}^{+\infty} f(x,y)\,dy \ \Rightarrow \ (3.2.6)
\end{aligned}$$

ja

$$\begin{aligned}
&F(x,y) = \int_{-\infty}^{x} du \int_{-\infty}^{y} f(u,v)\,dv \ \wedge \ F(+\infty,y) = F_2(y) \ \Rightarrow\\
&\qquad \Rightarrow \ F_2(y) = F(+\infty,y) = \int_{-\infty}^{y} dv \int_{-\infty}^{+\infty} f(x,v)\,dx \ \Rightarrow\\
&\qquad \Rightarrow \ f_2(y) = \frac{d}{dy}F_2(y) = \frac{d}{dy}\int_{-\infty}^{y} dv \int_{-\infty}^{+\infty} f(x,v)\,dx =\\
&\qquad = \int_{-\infty}^{+\infty} f(x,y)\,dx \ \Rightarrow \ (3.2.7)\,. \qquad \square
\end{aligned}$$

Definitsioon 2. Kui

$$f(x,y) = f_1(x)\ f_2(y)\,, \tag{3.2.8}$$

siis öeldakse, et juhusliku vektori (X,Y) komponendid X ja Y on *sõltumatud.* Kui tingimus (3.2.8) ei ole täidetud, siis vektori (X,Y) komponente X ja Y nimetatakse *sõltuvateks.*

Definitsioon 3. Kui juhusliku vektori (X,Y) jaotustihedus on kujul

$$f(x,y) = \begin{cases} \dfrac{1}{S_D}, \ \text{kui} \ (x,y) \in D, \\[2ex] 0, \ \text{kui} \ (x,y) \notin D, \end{cases}$$

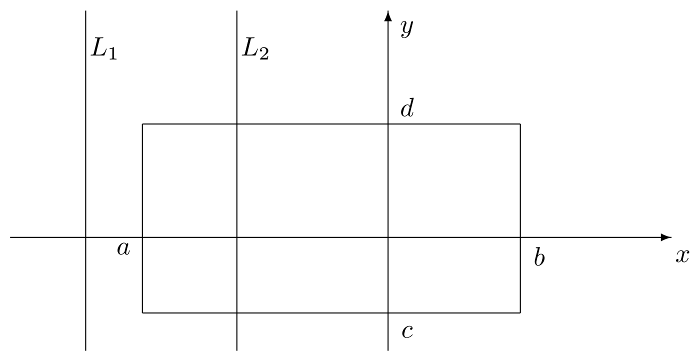
siis öeldakse, et juhuslik vektor (X,Y) allub hulgal D *ühtlasele jaotusele.*

Näide 1. Allugu vektor (X,Y) ristkülikus $[a,b]\times[c,d]$ ühtlasele jaotusele. Leiame selle juhusliku vektori marginaalsed jaotustihedused. Kas selle vektori komponendid on sõltuvad või sõltumatud?

Selle vektori jaotustihedus on kujul

$$f(x,y)=\begin{cases}\dfrac{1}{(b-a)\,(d-c)}, \text{ kui } (x,y)\in[a,b]\times[c,d]\,,\\[2ex] 0, \text{ kui } (x,y)\notin[a,b]\times[c,d]\,.\end{cases} \tag{3.2.9}$$

Teeme joonise



Vektori (X,Y) esimese komponendi X jaotustiheduse $f_1(x)$ leiame valemi (3.2.6) abil. Märgime, et valemis (3.2.6) esinevas integraalis $\int_{-\infty}^{+\infty} f\,(x,y)\,dy$ toimub integreerimine xy-tasandil fikseeritud x korral piki sirget, mis on risti x-teljega. Kui $x\notin[a,b]$, siis selle sirge, näiteks sirge L_1, kõigis punktides $f\,(x,y)=0$ ja seega $f_1(x)\overset{x\notin[a,b]}{=}0$. Kui $x\in[a,b]$, siis selle sirge, näiteks sirge L_2, kui integraali $\int_{-\infty}^{+\infty} f\,(x,y)\,dy$ integreerimispiirkonna saame jaotada kolme ossa

$$f_1(x)=\int_{-\infty}^{+\infty} f\,(x,y)\,dy=\int_{-\infty}^{c} 0dy+\int_{c}^{d}\frac{1}{(b-a)\,(d-c)}dy+\int_{d}^{+\infty} 0dy=$$
$$=\int_{c}^{d}\frac{1}{(b-a)\,(d-c)}dy=\frac{1}{b-a}.$$

Saame

$$f_1(x)=\begin{cases}\dfrac{1}{b-a}, \text{ kui } x\in[a,b]\,,\\[2ex] 0, \text{ kui } x\notin[a,b]\,.\end{cases}$$

Analoogiliselt saame näidata, et

$$f_2(y)=\begin{cases}\dfrac{1}{d-c}, \text{ kui } y\in[c,d]\,,\\[2ex] 0, \text{ kui } y\notin[c,d]\,.\end{cases}$$

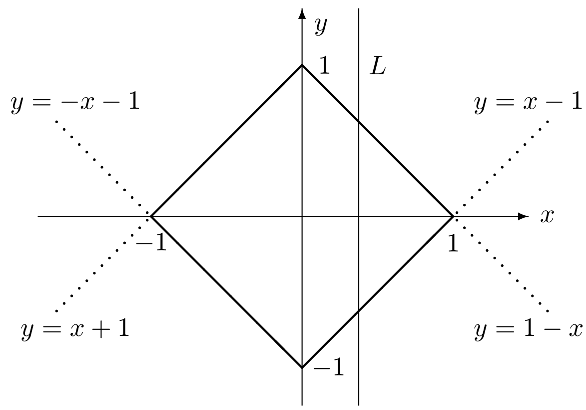
Seega alluvad selle vektori mõlemad komponendid ühtlasele jaotusele, vastavalt lõigul $[a, b]$ ja lõigul $[c, d]$. Veenduge, et selle vektori (X, Y) korral on täidetud tingimus (3.2.8), st vektori (X, Y) komponendid X ja Y on sõltumatud. $\diamondsuit$

Näide 2. Allugu vektor (X, Y) ühtlasele jaotusele ruudus

$$|x| + |y| \leq 1.$$

Leiame selle juhusliku vektori marginaalsed jaotustihedused. Kas selle vektori komponendid on sõltuvad või sõltumatud?

Teeme joonise



Selle ruudu pindala on 2. Seega

$$f(x, y) = \begin{cases} 0.5, \text{ kui } |x| + |y| \leq 1, \\ 0, \text{ kui } |x| + |y| > 1 \end{cases}$$

on vektori (X, Y) jaotustihedus. Vektori esimese komponendi X jaotustiheduse $f_1(x)$ leiame valemi (3.2.6) abil. Kui $x \notin [-1; 1]$, siis sirge, piki mida toimub integreerimine, kõigis punktides $f(x, y) = 0$ ja seega $f_1(x) \overset{x \notin [-1;1]}{=} 0$. Kui $x \in [0; 1]$, siis selle sirge, näiteks sirge L, kui integraali $\int_{-\infty}^{+\infty} f(x, y)\, dy$ integreerimispiirkonna saame jaotada kolme ossa. Saame

$$\begin{aligned} f_1(x) &= \int_{-\infty}^{+\infty} f(x, y)\, dy = \int_{-\infty}^{x-1} 0 dy + \int_{x-1}^{1-x} 0.5 dy + \int_{1-x}^{+\infty} 0 dy = \\ &= \int_{x-1}^{1-x} 0.5 dy = 0.5\,(1 - x - (x - 1)) = 1 - x \overset{x \geq 0}{=} 1 - |x|\,. \end{aligned}$$

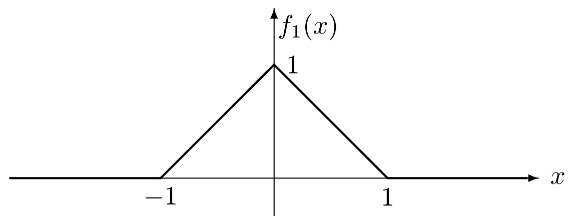
Analoogiliselt saame $x \in [-1; 0]$ korral

$$\begin{aligned} f_1(x) &= \int_{-\infty}^{+\infty} f(x, y)\, dy = \int_{-\infty}^{-x-1} 0 dy + \int_{-x-1}^{x+1} 0.5 dy + \int_{x+1}^{+\infty} 0 dy = \\ &= \int_{-x-1}^{x+1} 0.5 dy = 0.5\,(x + 1 - (-x - 1)) = 1 + x \overset{x \leq 0}{=} 1 - |x|\,. \end{aligned}$$

Seega

$$f_1(x) = \begin{cases} 1 - |x|\,, \text{ kui } x \in [-1;1]\,, \\ \\ 0, \text{ kui } x \notin [-1;1]\,, \end{cases}$$

st vektori (X,Y) esimene komponent X allub Simpsoni jaotusele *kandjaga* (hulgaga, millel jaotustihedus on nullist erinev) $[-1;1]$



Analoogiliselt saab näidata, et vektori (X,Y) komponent Y allub Simpsoni jaotusele jaotustihedusega

$$f_2(y) = \begin{cases} 1 - |y|\,, \text{ kui } y \in [-1;1]\,, \\ \\ 0, \text{ kui } y \notin [-1;1]\,. \end{cases}$$

Veenduge, et vektori (X,Y) korral ei ole täidetud tingimus (3.2.8), st vektori (X,Y) komponendid X ja Y on sõltuvad. $\diamondsuit$

Lause 4. Kui

$$P\left((X,Y) = (x_i, y_j)\right) = p_{i,\,j} \ \ (i \in I,\, j \in J) \tag{3.2.10}$$

on diskreetse juhusliku vektori (X,Y) jaotusseadus, kus $\{(x_i, y_j)\}_{i\in I,\, j\in J}$ on selle vektori võimalike väärtuste hulk, siis selle vektori (X,Y) jaotustihedus on esitatav kujul

$$f\,(x,y) = \textstyle\sum_{i\in I}\sum_{j\in J} p_{i,\,j}\,\delta\,(x - x_i)\,\delta\,(y - y_j)\,. \tag{3.2.11}$$

Tõestus. Lause 4 väide on järeldus seostest (3.1.8) ja (3.2.3). Tõesti,

$$f\,(x,y) = \frac{\partial^2 F\,(x,y)}{\partial x \partial y} = \frac{\partial^2}{\partial x \partial y} \textstyle\sum_{i\in I}\sum_{j\in J} p_{i,\,j}\,\mathbf{1}\,(x - x_i)\,\mathbf{1}\,(y - y_j) =$$

$$= \textstyle\sum_{i\in I}\sum_{j\in J} p_{i,\,j} \dfrac{\partial^2}{\partial x \partial y}\left(\mathbf{1}\,(x - x_i)\,\mathbf{1}\,(y - y_j)\right) =$$
$$= \textstyle\sum_{i\in I}\sum_{j\in J} p_{i,\,j} \dfrac{\partial}{\partial y}\left(\delta\,(x - x_i)\,\mathbf{1}\,(y - y_j)\right) =$$
$$= \textstyle\sum_{i\in I}\sum_{j\in J} p_{i,\,j}\left(\delta\,(x - x_i)\,\delta\,(y - y_j)\right). \qquad \square$$

Näide 3. Karbist, milles on 7 uut ja 3 kasutatud disketti, võetakse huupi 2 disketti. Olgu X ja Y vastavalt uute ja kasutatud diskettide arv nende kahe võetu hulgas. Moodustame juhusliku vektori (X,Y). Leiame selle vektori jaotusseaduse, jaotusfunktsiooni, jaotustiheduse ja marginaalsed jaotustihedused ning uurime, kas selle vektori komponendid on sõltuvad.

Et

$$P((X,Y)=(0;2))=P(X=0)=\frac{C_3^2\cdot C_7^0}{C_{10}^2}=\frac{\frac{3\cdot 2}{2!}}{\frac{10\cdot 9}{2!}}=\frac{3\cdot 2}{10\cdot 9}=\frac{1}{15},$$
$$P((X,Y)=(1;1))=P(X=1)=\frac{C_3^1\cdot C_7^1}{C_{10}^2}=\frac{3\cdot 7}{\frac{10\cdot 9}{2!}}=\frac{3\cdot 7}{45}=\frac{7}{15},$$
$$P((X,Y)=(2;0))=P(X=2)=\frac{C_3^0\cdot C_7^2}{C_{10}^2}=\frac{\frac{7\cdot 6}{2!}}{\frac{10\cdot 9}{2!}}=\frac{7\cdot 6}{10\cdot 9}=\frac{7}{15},$$

siis

$x_i \backslash y_j$	0	1	2
0	0	0	1/15
1	0	7/15	0
2	7/15	0	0

on vektori (X,Y) jaotusseadus,

$$F(x,y)\overset{(3.1.8)}{=}\sum\nolimits_{i=0}^{2}\sum\nolimits_{j=0}^{2}p_{i,j}\mathbf{1}(x-x_i)\mathbf{1}(y-y_j)=$$
$$=\frac{1}{15}\mathbf{1}(x-0)\mathbf{1}(y-2)+\frac{7}{15}\mathbf{1}(x-1)\mathbf{1}(y-1)+\frac{7}{15}\mathbf{1}(x-2)\mathbf{1}(y-0)$$

on selle vektori jaotusfunktsioon,

$$f(x,y)\overset{(3.2.11)}{=}\sum\nolimits_{i=0}^{2}\sum\nolimits_{j=0}^{2}p_{i,j}\delta(x-x_i)\delta(y-y_j)=$$
$$=\frac{1}{15}\delta(x)\delta(y-2)+\frac{7}{15}\delta(x-1)\delta(y-1)+\frac{7}{15}\delta(x-2)\delta(y)$$

on selle vektori jaotustihedus. Valemite (3.2.6) ja (3.2.7) abil saame vastavalt komponentide X ja Y marginaalsed jaotustihedused

$$f_1(x) = \int_{-\infty}^{+\infty} f(x,y)\,dy =$$

$$= \int_{-\infty}^{+\infty} \left(\frac{1}{15}\delta(x)\,\delta(y-2) + \frac{7}{15}\delta(x-1)\,\delta(y-1) + \frac{7}{15}\delta(x-2)\,\delta(y) \right) dy =$$

$$= \frac{1}{15}\delta(x) \int_{-\infty}^{+\infty} \delta(y-2)\,dy + \frac{7}{15}\delta(x-1) \int_{-\infty}^{+\infty} \delta(y-1)\,dy +$$

$$+\frac{7}{15}\delta(x-2) \int_{-\infty}^{+\infty} \delta(y)\,dy = \left[\int_{-\infty}^{+\infty} \delta(y-a)\,dy \overset{(2.2.12)}{=} 1 \right] =$$

$$= \frac{1}{15}\delta(x) + \frac{7}{15}\delta(x-1) + \frac{7}{15}\delta(x-2)$$

ning

$$\begin{aligned} f_2(y) &= \int_{-\infty}^{+\infty} f(x,y)\,dx = [\text{analoogiline arutelu }] = \\ &= \frac{1}{15}\delta(y-2) + \frac{7}{15}\delta(y-1) + \frac{7}{15}\delta(y)\,. \end{aligned}$$

Veenduge, et selle vektori (X,Y) korral ei ole täidetud tingimus (3.2.8), st vektori (X,Y) komponendid X ja Y on sõltuvad. Kontrollige lineaarse seose $X+Y=2$ olemasolu! $\diamondsuit$

3.3 Juhusliku vektori tinglikud jaotustihedused

Juhusliku vektori tingliku jaotustiheduse mõiste sissetoomisel lähtume seosest (3.2.1)

$$f(x,y) = \lim_{(\Delta x,\Delta y)\to(0+;0+)} \frac{P\left((x \le X < x+\Delta x)\,(y \le Y < y+\Delta y)\right)}{\Delta x\,\Delta y} =$$

$$= \lim_{\Delta y\to 0+} \lim_{\Delta x\to 0+} \frac{P(x \le X < x+\Delta x)}{\Delta x} \frac{P(y \le Y < y+\Delta y \,|\, x \le X < x+\Delta x)}{\Delta y} =$$

$$= \lim_{\Delta y\to 0+} f_1(x) \frac{P(y \le Y < y+\Delta y | X = x)}{\Delta y} = f_1(x) f_2(y\,|\,x),$$

kus

$$f_2(y\,|\,x) \overset{\text{def.}}{=} \lim_{\Delta y\to 0+} \frac{P(y \le Y < y+\Delta y \,|\, X = x)}{\Delta y}. \qquad (3.3.1)$$

Definitsioon 1. Funktsiooni $f_2(y\,|\,x)$, mis on määratud seosega (3.3.1), nimetatakse juhusliku vektori (X,Y) *komponendi Y tinglikuks jaotustiheduseks.*

Analoogiliselt defineeritakse juhusliku vektori (X,Y) komponendi X tinglik jaotustihedus

$$f_1(x\,|\,y) \overset{\text{def.}}{=} \lim_{\Delta x\to 0+} \frac{P\,(x \le X < x+\Delta x\,|\,Y=y)}{\Delta x}. \tag{3.3.2}$$

Lause 1. Juhusliku vektori (X,Y) korral kehtib võrduste ahel

$$f\,(x,y) = f_1(x)f_2(y\,|\,x) = f_2(y)f_1(x\,|\,y). \tag{3.3.3}$$

Järeldus 1. Kehtivad seosed

$$f_1(x\,|\,y) = \frac{f\,(x,y)}{f_2(y)} \quad (f_2(y) \ne 0) \tag{3.3.4}$$

ja

$$f_2(y\,|\,x) = \frac{f\,(x,y)}{f_1(x)} \quad (f_1(x) \ne 0)\,. \tag{3.3.5}$$

Lause 2. Juhusliku vektori (X,Y) komponendid X ja Y on sõltumatud parajasti siis, kui

$$f_1(x\,|\,y) = f_1(x) \tag{3.3.6}$$

või

$$f_2(y\,|\,x) = f_2(y). \tag{3.3.7}$$

Tõestus. Juhusliku vektori (X,Y) komponendid X ja Y on sõltumatud, kui

$$f\,(x,y) \overset{(3.2.8)}{=} f_1\,(y)\; f_2\,(y)\,. \tag{3.3.8}$$

Seostest (3.3.3) ja (3.3.8) järeldub Lause 2 väide. $\diamondsuit$

Näide 2. Allugu vektor (X,Y) ruudus $|x|+|y| \le 1$ ühtlasele jaotusele. Leiame selle juhusliku vektori tinglikud jaotustihedused.

Näites 3.2.2 leidsime selle vektori jaotustiheduse ja tema komponentide jaotustihedused. Tingliku jaotustiheduse $f_1(x\,|\,y)$ saame seose (3.3.4) abil

$$f_1(x\,|\,y) = \frac{f\,(x,y)}{f_2(y)} = \begin{cases} \dfrac{0.5}{1-|y|}, \text{ kui } |x| < 1-|y|\,, \\ 0, \text{ kui } |x| \ge 1-|y| \end{cases} \Rightarrow \; f_1(x\,|\,y) \ne f_1(x).$$

Märgime, et juhul $Y=y$ $(|y|<1)$ allub vektori (X,Y) komponent X ühtlasele jaotusele vahemikus $(|y|-1;1-|y|)\,.$ Analoogiliselt saame seose (3.3.5) abil

$$f_2(y\,|\,x) = \frac{f\,(x,y)}{f_1(x)} = \begin{cases} \dfrac{0.5}{1-|x|}, \text{ kui } |y| < 1-|x|\,, \\ 0, \text{ kui } |y| \ge 1-|x| \end{cases} \Rightarrow \; f_2(y\,|\,x) \ne f_2(y). \quad \diamondsuit$$

Märkus 1. Juhusliku vektori $(X_1,X_2,\ldots,X_n)$ korral kehtib seos

$$f\,(x_1,x_2,\ldots,x_n) = f_1(x_1)f_2(x_2\,|\,x_1)f_3(x_3\,|\,x_1,x_2)\cdots f_n(x_n\,|\,x_1,x_2,\ldots,x_{n-1}), \tag{3.3.9}$$

kus $f_k(x_k \mid x_1, x_2, \ldots, x_{k-1}) \quad (1 < k \leq n)$ on selle vektori k-nda komponendi tinglik jaotustihedus eeldustel $X_1 = x_1, \ldots, X_{k-1} = x_{k-1}$.

Definitsioon 2. Juhusliku vektori $(X_1, X_2, \ldots, X_n)$ komponente nimetatakse *sõltumatuteks*, kui

$$f(x_1, x_2, \ldots, x_n) = f_1(x_1) f_2(x_2) f_3(x_3) \cdots f_n(x_n). \tag{3.3.10}$$

3.4 Juhusliku vektori momendid

Juhusliku argumendiga funktsiooni $h(X)$ keskväärtus $\mathrm{E}(h(X))$ on defineeritud seose (2.3.4) abil. Üldistame selle mõiste juhusliku vektori jaoks.

Definitsioon 1. Kui $f(x_1, \ldots, x_n)$ on juhusliku vektori $(X_1, \ldots, X_n)$ jaotustihedus, siis kindlat suurust

$$\mathrm{E}(h(X_1, \ldots, X_n)) \stackrel{\text{def.}}{=} \int_{-\infty}^{+\infty} \ldots \int_{-\infty}^{+\infty} h(x_1, \ldots, x_n) f(x_1, \ldots, x_n) dx_1 \cdots dx_n \tag{3.4.1}$$

nimetatakse *funktsiooni* $h(X_1, \ldots, X_n)$ *keskväärtuseks.*

Järeldus 1. Kui $f(x, y)$ on juhusliku vektori (X, Y) jaotustihedus, siis

$$\mathrm{E}(h(X, Y)) \stackrel{\text{def.}}{=} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} h(x, y) f(x, y) dx dy \tag{3.4.2}$$

on funktsiooni $h(X, Y)$ keskväärtus.

Definitsioon 2. Arvu

$$\nu_{k_1, \ldots, k_n} \stackrel{\text{def.}}{=} \mathrm{E}\left(X_1^{k_1} \cdots X_n^{k_n}\right) \tag{3.4.3}$$

nimetatakse *juhusliku vektori* $(X_1, \ldots, X_n)$ $(k_1 + \ldots + k_n)$*-järku algmomendiks.*

Lause 1. Kui $f(x, y)$ on vektori (X, Y) jaotustihedus, siis selle juhusliku vektori (X, Y) $(k + m)$-järku algmoment $\nu_{k,m}$ avaldub kujul

$$\nu_{k,m} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x^k y^m f(x, y) dx dy. \tag{3.4.4}$$

Tõestus. Definitsiooni 2 põhjal $\nu_{k,m} = \mathrm{E}\left(X^k Y^m\right)$. Lause 1 väite saame Järeldusest 1 valiku $h(x, y) = x^k y^m$ korral. $\square$

Juhusliku vektori (X, Y) algmomentidega $\nu_{k,0}$ ja $\nu_{0,m}$ oleme eelnevalt tuttavad. Tõesti,

$$\nu_{k,0} \stackrel{(3.4.4)}{=} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x^k y^0 f(x, y) dx dy = \int_{-\infty}^{+\infty} x^k dx \int_{-\infty}^{+\infty} f(x, y) dy \stackrel{(3.2.6)}{=}$$

$$= \int_{-\infty}^{+\infty} x^k f_1(x) dx = \mathrm{E}\left(X^k\right) = \nu_k(X).$$

Analoogiliselt saab näidata, et $\nu_{0,m} = \mathrm{E}(Y^m) = \nu_m(Y)$.

Kui vaatluse all on rohkem kui kaks juhuslikku suurust, siis kasutatakse suuruse $\mathrm{E}\left(X^k Y^m\right)$ tähistamiseks $\nu_{k,m}$ asemel tähistust $\nu_{k,m}(X,Y)$.

Lause 2. Kui

$$P\left((X,Y)=(x_i,y_j)\right)=p_{i,j}\ \ (i\in I,\, j\in J)$$

on diskreetse juhusliku vektori (X,Y) jaotusseadus, siis

$$\nu_{k,m}=\textstyle\sum_{i\in I}\sum_{j\in J}x_i^k y_j^m p_{i,j}. \tag{3.4.5}$$

Tõestus. Saame

$$\nu_{k,m} \overset{(3.4.4),\,(3.2.11)}{=} \textstyle\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x^k y^m \sum_{i\in I}\sum_{j\in J} p_{i,j}\,\delta(x-x_i)\,\delta(y-y_j)\,dxdy=$$

$$=\textstyle\sum_{i\in I}\sum_{j\in J}p_{i,j}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}x^k y^m\delta(x-x_i)\,\delta(y-y_j)\,dxdy=$$

$$=\textstyle\sum_{i\in I}\sum_{j\in J}p_{i,j}\int_{-\infty}^{+\infty}x^k\delta(x-x_i)\,dx\int_{-\infty}^{+\infty}y^m\delta(y-y_j)\,dy=$$

$$=\textstyle\sum_{i\in I}\sum_{j\in J}x_i^k y_j^m p_{i,j}. \qquad \square$$

Definitsioon 3. Kindlat suurust

$$\mu_{k_1,\ldots,k_n}\overset{\text{def.}}{=}\mathrm{E}\left((X_1-\mathrm{E}X_1)^{k_1}\cdots(X_n-\mathrm{E}X_n)^{k_n}\right) \tag{3.4.6}$$

nimetatakse *juhusliku vektori* $(X_1,\ldots,X_n)$ $(k_1+\ldots+k_n)$-*järku keskmomendiks.* Kui kasutada tsentreeritud suurusi $X_k^o=X_k-\mathrm{E}X_k$, siis on valem (3.4.6) esitatav kujul

$$\mu_{k_1,\ldots,k_n}=\mathrm{E}\left((X_1^o)^{k_1}\cdots(X_n^o)^{k_n}\right).$$

Lause 3. Kui $f(x,y)$ on vektori (X,Y) jaotustihedus, siis selle juhusliku vektori (X,Y) $(k+m)$-järku keskmoment $\mu_{k,m}$ avaldub kujul

$$\mu_{k,m}=\int_{-\infty}^{+\infty}(x-\mathrm{E}X)^k\,(y-\mathrm{E}Y)^m\,f(x,y)dxdy. \tag{3.4.7}$$

Tõestus. Definitsiooni 3 põhjal $\mu_{k,m}=\mathrm{E}\left((X-\mathrm{E}X)^k\,(Y-\mathrm{E}Y)^m\right)$. Lause 3 väite saame Järeldusest 1 valiku

$$h(x,y)=(x-\mathrm{E}X)^k\,(y-\mathrm{E}Y)^m$$

korral. $\square$

Näidake, et vektori (X,Y) korral

$$\mu_{k;0}=\mathrm{E}\left(X^k\right)=\mu_k(X),\quad \mu_{0;m}=\mathrm{E}\left(Y^m\right)=\mu_m(Y).$$

Kui vaatluse all on rohkem kui kaks juhuslikku suurust, siis kasutame suuruse $\mathrm{E}\left((X-\mathrm{E}X)^k\,(Y-\mathrm{E}Y)^m\right)$ tähistamiseks $\mu_{k,m}$ asemel tähistust $\mu_{k,m}(X,Y)$.

Tõestage järgmise väite tõesus.

Lause 4. Kui $P\left((X,Y)=(x_i,y_j)\right)=p_{i,\,j}$ $(i\in I,\, j\in J)$ on diskreetse juhusliku vektori (X,Y) jaotusseadus, siis

$$\mu_{k,m}=\sum\nolimits_{i\in I}\sum\nolimits_{j\in J}\left(x_i-\mathrm{E}X\right)^k\left(y_j-\mathrm{E}Y\right)^m p_{i,\,j}. \qquad (3.4.8)$$

Näide 1. Allugu vektor (X,Y) ühtlasele jaotusele ruudus

$$|x|+|y|\le 1.$$

Leiame selle juhusliku vektori korral algmomendi $\nu_{2;4}$ ja keskmomendid $\mu_{1;1}$ ning $\mu_{2;2}$.

Näites 3.2.2 leidsime

$$f(x,y)=\begin{cases}0.5, \text{ kui } |x|+|y|\le 1,\\ 0, \text{ kui } |x|+|y|>1.\end{cases}$$

Komponentide X ja Y alluvusest Simpsoni jaotusele lõigul $[-1;1]$ järeldub $\mathrm{E}X=0$ ja $\mathrm{E}Y=0$. Saame

$$\nu_{2;4}\overset{(3.4.4)}{=}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}x^2y^4f(x,y)dxdy=0.5\int_{-1}^{1}x^2dx\int_{-1+|x|}^{1-|x|}y^4dy=$$

$$=\int_{-1}^{1}x^2dx\int_{0}^{1-|x|}y^4dy=\frac{1}{5}\int_{-1}^{1}x^2\left(y^5\right)\Big|_0^{1-|x|}dx=$$

$$=\frac{1}{5}\int_{-1}^{1}x^2\left(1-|x|\right)^5dx=\frac{2}{5}\int_{0}^{1}x^2\left(1-x\right)^5dx=\frac{1}{210}$$

ja

$$\mu_{1;1}\overset{(3.4.7)}{=}\int_{-\infty}^{+\infty}\left(x-\mathrm{E}X\right)\left(y-\mathrm{E}Y\right)f(x,y)dxdy=0.5\int_{-1}^{1}xdx\int_{-1+|x|}^{1-|x|}ydy=0$$

ning

$$\mu_{2;2}\overset{(3.4.7)}{=}\int\limits_{-\infty}^{+\infty}\left(x-\mathrm{E}X\right)^2\left(y-\mathrm{E}Y\right)^2f(x,y)dxdy=$$

$$=0.5\int\limits_{-1}^{1}x^2dx\int\limits_{-1+|x|}^{1-|x|}y^2dy=\int\limits_{-1}^{1}x^2dx\int\limits_{0}^{1-|x|}y^2dy=\frac{1}{3}\int\limits_{-1}^{1}x^2dx\cdot y^3\Big|_0^{1-|x|}=$$

$$=\frac{1}{3}\int\limits_{-1}^{1}x^2\left(1-|x|\right)^3dx=\frac{2}{3}\int\limits_{0}^{1}x^2\left(1-x\right)^3dx=\frac{1}{90}. \qquad \diamondsuit$$

Näide 2. Karbist, milles on 7 uut ja 3 kasutatud disketti, võetakse huupi 2 disketti. Olgu X ja Y vastavalt uute ja kasutatud diskettide arv nende kahe võetu hulgas. Leiame vektori (X,Y) algmomendi $\nu_{1;1}$ ja keskmomendi $\mu_{1;1}$.

Näites 3.2.3 leidsime selle vektori jaotusseaduse

$x_i \backslash y_j$	0	1	2
0	0	0	1/15
1	0	7/15	0
2	7/15	0	0

Näidake, et $\mathrm{E}X = \dfrac{7}{5}$ ja $\mathrm{E}Y = \dfrac{3}{5}$. Valemite (3.4.5) ja (3.4.8) abil saame vastavalt

$$\nu_{1;1} = \textstyle\sum_{i=1}^{3}\sum_{j=1}^{3} x_i y_j p_{i,\,j} = 2 \cdot 0 \cdot \dfrac{7}{15} + 1 \cdot 1 \cdot \dfrac{7}{15} + 0 \cdot 2 \cdot \dfrac{1}{15} = \dfrac{7}{15}$$

ja

$$\mu_{1;1} = \sum_{i=1}^{3}\sum_{j=1}^{3} (x_i - \mathrm{E}X)(y_j - \mathrm{E}Y)\, p_{i,\,j} = \left(2 - \frac{7}{5}\right) \cdot \left(0 - \frac{3}{5}\right) \cdot \frac{7}{15} +$$

$$+ \left(1 - \frac{7}{5}\right) \cdot \left(1 - \frac{3}{5}\right) \cdot \frac{7}{15} + \left(0 - \frac{7}{5}\right) \cdot \left(2 - \frac{3}{5}\right) \cdot \frac{1}{15} = -\frac{28}{75}. \qquad \diamondsuit$$

Definitsioon 4. Funktsiooni

$$g(\omega_1, \omega_2 \ldots, \omega_n) \overset{\text{def.}}{=} \mathrm{E}\, e^{i\omega_1 X_1 + i\omega_2 X_2 + \ldots + i\omega_n X_n}$$

nimetatakse *juhusliku vektori* $(X_1, X_2, \ldots, X_n)$ *karakteristlikuks funktsiooniks.*

Lause 5. Kui $g(\omega_1, \omega_2 \ldots, \omega_n)$ on juhusliku vektori $(X_1, X_2, \ldots, X_n)$ karakteristlik funktsioon ja eksisteerivad $\nu_{k_1,\ldots,k_n}$ ja $\mu_{k_1,\ldots,k_n}$ ning

$$\exists \left(\frac{\partial^{k_1+\ldots+k_n} g(\omega_1, \ldots, \omega_n)}{\partial\omega_1^{k_1} \cdots \partial\omega_n^{k_n}}\right)_{(\omega_1,\ldots,\omega_n)=(0;\ldots;0)}$$

ja

$$\exists \left(\frac{\partial^{k_1+\ldots+k_n} \left(e^{-i\omega_1 \mathrm{E}X_1 \ldots - i\omega_n \mathrm{E}X_n} g(\omega_1, \ldots, \omega_n)\right)}{\partial\omega_1^{k_1} \cdots \partial\omega_n^{k_n}}\right)_{(\omega_1,\ldots,\omega_n)=(0;\ldots;0)},$$

siis

$$\nu_{k_1,\ldots,k_n} = \frac{1}{i^{k_1+\ldots+k_n}} \left(\frac{\partial^{k_1+\ldots+k_n} g(\omega_1, \ldots, \omega_n)}{\partial\omega_1^{k_1} \cdots \partial\omega_n^{k_n}}\right)_{(\omega_1,\ldots,\omega_n)=(0;\ldots;0)}$$

ja

$$\mu_{k_1,\ldots,k_n} = \frac{1}{i^{k_1+\ldots+k_n}} \cdot$$

$$\cdot \left(\frac{\partial^{k_1+\ldots+k_n} \left(e^{-i\omega_1 \mathrm{E}X_1 \ldots - i\omega_n \mathrm{E}X_n} g(\omega_1, \ldots, \omega_n)\right)}{\partial\omega_1^{k_1} \cdots \partial\omega_n^{k_n}}\right)_{(\omega_1,\ldots,\omega_n)=(0;\ldots;0)}.$$

3.5 Komponentide korreleeruvus. Regressioon

Definitsioon 1. Kindlat suurust

$$\operatorname{cov}(X,Y) \stackrel{\text{def.}}{=} \mathrm{E}\left((X-\mathrm{E}X)\left(Y-\mathrm{EY}\right)\right) \tag{3.5.1}$$

nimetatakse juhuslike suuruste X ja Y *kovariatsiooniks.*

Seega $\operatorname{cov}(X,Y) = \mu_{1;1}(X,Y) = E\left(X^o Y^o\right).$ Suurust $\operatorname{cov}(X,Y)$ nimetatakse ka juhusliku vektori (X,Y) kovariatsiooniks, sest $\operatorname{cov}(X,Y)$ määramiseks on vaja teada vektori (X,Y) jaotust.

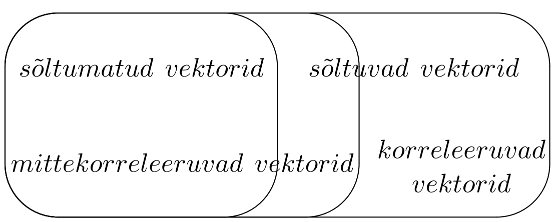
Definitsioon 2. Juhuslikke suurusi X ja Y nimetatakse *korreleeruvateks,* kui $\operatorname{cov}(X,Y) \neq 0$, ja *mittekorreleeruvateks,* kui $\operatorname{cov}(X,Y) = 0$.

Lause 1. Juhuslike suuruste X ja Y sõltumatusest järeldub nende mittekorreleeruvus.

Tõestus. Saame

$$\begin{aligned}
\operatorname{cov}(X,Y) &= \mathrm{E}\left((X-\mathrm{E}X)\left(Y-\mathrm{E}Y\right)\right) = \\
&= \left[\begin{array}{c} X \text{ ja } Y \text{ on sõltumatud} \Rightarrow X-\mathrm{E}X \text{ ja } Y-\mathrm{E}Y \\ \text{on sõltumatud} \Rightarrow \text{korrutise keskväärtus on} \\ \text{keskväärtuste korrutis} \end{array}\right] \\
&= \mathrm{E}\left(X-\mathrm{E}X\right)\mathrm{E}\left(Y-\mathrm{E}Y\right) = \left(\mathrm{E}X-\mathrm{E}X\right)\left(\mathrm{E}Y-\mathrm{E}Y\right) = 0. \qquad \square
\end{aligned}$$

Lause 1 ei ole pööratav, st juhuslike suuruste X ja Y mittekorreleeruvusest ei järeldu nende sõltumatus. Tõesti, Näites 3.4.1 leidsime ruudus $|x|+|y| \leq 1$ ühtlasele jaotusele alluva vektori (X,Y) korral, et $\mu_{1;1} = 0$, st $\operatorname{cov}(X,Y) = 0$. Seega on komponendid X ja Y mittekorreleeruvad. Näite 3.3.2 põhjal on selle vektori komponendid X ja Y sõltuvad. Järgneval joonisel on kujutatud kõigi juhuslike vektorite (X,Y) hulga jagunemine



Lause 2. Kehtib võrratus

$$|\operatorname{cov}(X,Y)| \leq \sigma_X\, \sigma_Y. \tag{3.5.2}$$

Tõestus. Kui X ja Y on juhuslikud suurused, siis $\mathrm{D}X$, $\mathrm{D}Y$ ja $\operatorname{cov}(X,Y)$ on kindlad suurused. Iga $\lambda \in \mathbf{R}$ korral kehtib väidete ahel

$$\begin{aligned}
\mathrm{E}\left(\left(\lambda X^0 - Y^0\right)^2\right) \geq 0 \Leftrightarrow \lambda^2 \mathrm{E}\left(X^0\right)^2 - 2\lambda \mathrm{E}\left(\left(X^0\right)\left(Y^0\right)\right) + \mathrm{E}\left(Y^0\right)^2 \geq 0 \Leftrightarrow \\
\Leftrightarrow \lambda^2 \cdot \mathrm{D}X - 2\lambda \cdot \operatorname{cov}(X,Y) + \mathrm{D}Y \geq 0,
\end{aligned}$$

st ruutpolünoomi väärtused on mittenegatiivsed. See on võimalik, kui vastava ruutvõrrandi diskriminant on mittepositiivne, $(\mathrm{cov}(X,Y))^2 - (\mathrm{D}X)\cdot(\mathrm{D}Y) \le 0$. Seega

$$(\mathrm{cov}(X,Y))^2 \le (\mathrm{D}X)\cdot(\mathrm{D}Y) \quad \Leftrightarrow \quad |\mathrm{cov}(X,Y)| \le \sqrt{\mathrm{D}X}\cdot\sqrt{\mathrm{D}Y} = \sigma_X\,\sigma_Y\,. \square$$

Lause 3. Kehtib seos

$$\mathrm{cov}(X,Y) = \mathrm{E}\,(XY) - (\mathrm{E}X)\,(\mathrm{E}Y)\,. \tag{3.5.3}$$

Tõestus. Saame

$$\begin{aligned}\mathrm{cov}(X,Y) &= \mathrm{E}\,((X-\mathrm{E}X)\,(Y-\mathrm{E}Y)) = \\ &= \mathrm{E}\,(XY - X\,(\mathrm{E}Y) - (\mathrm{E}X)\,Y + (\mathrm{E}X)\,(\mathrm{E}X)) = \\ &= \mathrm{E}\,(XY) - (\mathrm{E}X)\,(\mathrm{E}Y) - (\mathrm{E}X)\,(\mathrm{E}Y) + (\mathrm{E}X)\,(\mathrm{E}X) = \\ &= \mathrm{E}\,(XY) - (\mathrm{E}X)\,(\mathrm{E}X)\,. \qquad \square\end{aligned}$$

Definitsioon 3. Kindlat suurust

$$\mathrm{r}(X,Y) \overset{\mathrm{def.}}{=} \frac{\mathrm{cov}(X,Y)}{\sigma_X\,\sigma_Y} \tag{3.5.4}$$

nimetatakse juhuslike suuruste X ja Y *korrelatsioonikordajaks.*

Korrelatsioonikordaja tähistamiseks kasutatakse $\mathrm{r}(X,Y)$ asemel veel tähistusi r_{xy} ja r_{XY}.

Definitsioonist 3 ja võrratusest (3.5.2) järeldub järgmine tulemus.

Järeldus 1. Kehtib võrratus $|\mathrm{r}(X,Y)| \le 1$.

Lausest 3.4.3 saame järgmise järelduse.

Järeldus 2. Kui $f(x,y)$ on vektori (X,Y) jaotustihedus, siis

$$\mathrm{cov}(X,Y) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}(x-\mathrm{E}X)\,(y-\mathrm{E}Y)\,f(x,y)dxdy. \tag{3.5.5}$$

Lausest 3.4.4 saame järgmise järelduse.

Järeldus 3. Kui $P\,((X,Y)=(x_i,y_j)) = p_{i,\,j}$ $(i\in I,\, j\in J)$ on diskreetse juhusliku vektori (X,Y) jaotusseadus, siis

$$\mathrm{cov}(X,Y) = \textstyle\sum_{i\in I}\sum_{j\in J}(x_i-\mathrm{E}X)\,(y_j-\mathrm{E}Y)\,p_{i,\,j}. \tag{3.5.6}$$

Näites 3.4.2 uurisime karbist, milles on 7 uut ja 3 kasutatud disketti, huupi kahe disketi võtmist, kusjuures X ja Y on vastavalt uute ja kasutatud diskettide arv nende kahe võetu hulgas. Leidsime $\mathrm{cov}(X,Y) = -28/75$. Näidake, et $\mathrm{D}X = \mathrm{D}Y = 28/75$. Seega $\mathrm{r}(X,Y) = -1$.

Erijuhul $Y = X$ saame

$$\mathrm{cov}(X,X) = \mathrm{E}\,((X-\mathrm{E}X)\,(X-\mathrm{E}X)) = \mathrm{E}\left((X-\mathrm{E}X)^2\right) = \mathrm{D}X,$$

st juhusliku suuruse kovariatsioon iseendaga on tema dispersioon

$$\operatorname{cov}(X, X) = \mathrm{D}X. \tag{3.5.7}$$

Definitsioon 4. Ruutmaatriksit

$$\mathrm{K} = \begin{pmatrix} \operatorname{cov}(X_1, X_1) & \operatorname{cov}(X_1, X_2) & \cdots & \operatorname{cov}(X_1, X_n) \\ \operatorname{cov}(X_2, X_1) & \operatorname{cov}(X_2, X_2) & \cdots & \operatorname{cov}(X_2, X_n) \\ \cdots & \cdots & \cdots & \cdots \\ \operatorname{cov}(X_n, X_1) & \operatorname{cov}(X_n, X_2) & \cdots & \operatorname{cov}(X_n, X_n) \end{pmatrix} \tag{3.5.8}$$

nimetatakse juhusliku vektori $(X_1, \ldots, X_n)$ *kovariatsioonimaatriksiks.*

Kovariatsioonimaatriksit tähistatakse ka sümbolitega (K_{ij}), $\left(\mathrm{K}_{x_i,\, x_j}\right)$ ja $\left(\mathrm{K}_{X_i,\, X_j}\right)$.

Tõestage, et kovariatsioonimaatriks on sümmeetriline peadiagonaali suhtes ja selle maatriksi peadiagonaalil on vektori komponentide dispersioonid.

Definitsioon 5. Ruutmaatriksit

$$\mathrm{R} = \begin{pmatrix} r(X_1, X_1) & r(X_1, X_2) & \cdots & r(X_1, X_n) \\ r(X_2, X_1) & r(X_2, X_2) & \cdots & r(X_2, X_n) \\ \cdots & \cdots & \cdots & \cdots \\ r(X_n, X_1) & r(X_n, X_2) & \cdots & r(X_n, X_n) \end{pmatrix} \tag{3.5.9}$$

nimetatakse juhusliku vektori $(X_1, \ldots, X_n)$ *korrelatsioonimaatriksiks.*

Korrelatsioonimaatriksit R tähistatakse sümbolitega (r_{ij}), $\left(\mathrm{r}_{x_i,\, x_j}\right)$, $\left(\mathrm{r}_{X_i,\, X_j}\right)$. Veenduge, et korrelatsioonimaatriks on sümmeetriline peadiagonaali suhtes ja selle maatriksi peadiagonaalil on ühed. Kuna nii kovariatsiooni- kui ka korrelatsioonimaatriksid on sümmeetrilised peadiagonaali suhtes, siis nende maatriksite diagonaali alla jäävaid elemente tavaliselt välja ei kirjutata.

Näide 1. Olgu $D = \{(x, y) \mid 0 \le x \le 1 \ \wedge\ 0 \le y \le x\}$ ja

$$f(x, y) = \begin{cases} 8xy, \text{ kui } (x, y) \in D, \\ \\ 0, \text{ kui } (x, y) \in D. \end{cases}$$

Leiame vektori (X, Y) kovariatsiooni- ja korrelatsioonimaatriksi.

Saame

$$f_1(x) = \begin{cases} \int_0^x 8xydy = 4x^3, \text{ kui } x \in [0; 1], \\ \\ 0, \text{ kui } x \notin [0; 1], \end{cases}$$

$$\mathrm{E}X = \int_0^1 x \cdot 4x^3 dx = \frac{4}{5},\ \mathrm{E}X^2 = \int_0^1 x^2 \cdot 4x^3 dx = \frac{2}{3},$$

$$\mathrm{D}X = \mathrm{E}X^2 - (\mathrm{E}X)^2 = \frac{2}{3} - \left(\frac{4}{5}\right)^2 = \frac{2}{75},\ \sigma_X = \sqrt{\frac{2}{75}},$$

$$f_2(y) = \begin{cases} \int_y^1 8xydx = 4y - 4y^3, \text{ kui } y \in [0; 1], \\ \\ 0, \text{ kui } y \notin [0; 1], \end{cases}$$

$$\mathrm{E}Y = \int_0^1 y \cdot \left(4y - 4y^3\right) dy = \frac{8}{15},\ \mathrm{E}Y^2 = \int_0^1 y^2 \cdot \left(4y - 4y^3\right) dy = \frac{1}{3},$$
$$\mathrm{D}Y = \mathrm{E}Y^2 - (\mathrm{E}Y)^2 = \frac{1}{3} - \left(\frac{8}{15}\right)^2 = \frac{11}{225},\ \sigma_Y = \sqrt{\frac{11}{225}}$$

ja

$$\begin{aligned}\mathrm{cov}(X,Y) &= \int_0^1 dx \int_0^x \left(x - \frac{4}{5}\right)\left(y - \frac{8}{15}\right) 8xydy = \\ &= \int_0^1 \left(\frac{8}{3}x^5 - \frac{64}{15}x^4 + \frac{128}{75}x^3\right) dx = \frac{4}{225}, \\ \mathrm{r}(X,Y) &= \frac{\mathrm{cov}(X,Y)}{\sigma_X\, \sigma_Y} = \frac{\frac{4}{225}}{\sqrt{\frac{2}{75}}\sqrt{\frac{11}{225}}} = \frac{2}{33}\sqrt{66}\end{aligned}$$

ning kovariatsioonimaatriksi ja korrelatsioonimaatriksi

$$\mathrm{K} = \begin{pmatrix} 2/75 & 4/225 \\ & 11/225 \end{pmatrix}, \quad \mathrm{R} = \begin{pmatrix} 1 & 2\sqrt{66}/33 \\ & 1 \end{pmatrix}. \qquad \diamondsuit$$

Definitsioon 6. Joont võrrandiga

$$y = \mathrm{E}\left(Y \mid x\right) \tag{3.5.10}$$

nimetatakse vektori (X, Y) komponendi Y *regressioonijooneks* komponendi X suhtes ja joont võrrandiga

$$x = \mathrm{E}\left(X \mid y\right) \tag{3.5.11}$$

vektori (X, Y) komponendi X regressioonijooneks komponendi Y suhtes, kusjuures

$$\mathrm{E}\left(Y \mid x\right) = \int_{-\infty}^{+\infty} y\, f_2\left(y|x\right) dy \tag{3.5.12}$$

ning

$$\mathrm{E}\left(X \mid y\right) = \int_{-\infty}^{+\infty} x\, f_1\left(x|y\right) dx. \tag{3.5.13}$$

Võrrandiga (3.5.10) antud regressioonijoone $y = \mathrm{E}\left(Y|\, x\right)$ korral leitakse tinglik keskväärtus $\mathrm{E}\left(Y|\, x\right)$ eeldusel, et komponent X omandas katse käigus väärtuse x. Nii seatakse igale juhusliku suuruse X võimalikule väärtusele x vastavusse arv $\mathrm{E}\left(Y|\, x\right)$ ja saadakse regressioonijoon (3.5.10). Kui regressioonijooneks on sirge, siis nimetatakse seda sirget *regressioonisirgeks.*

Näide 2. Leiame Näites 1 esitatud juhusliku vektori regressioonijooned.

Kasutame Näites 1 saadud tulemusi. Leiame $x \in (0;1)$ korral

$$f_2(y\,|\,x) \overset{(3.3.5)}{=} \frac{f(x,y)}{f_1(x)} = \begin{cases} \dfrac{8xy}{4x^3} = \dfrac{2y}{x^2}, \text{ kui } y \in (0,x)\,, \\ \\ 0, \text{ kui } y \notin (0,x)\,, \end{cases}$$

$$\mathrm{E}\,(Y\,|\,x) \overset{(3.5.12)}{=} \textstyle\int_0^x y \dfrac{2y}{x^2} dy = \dfrac{2}{3}x,\; y \overset{(3.5.9)}{=} \dfrac{2}{3}x$$

ja $y \in (0;1)$ korral

$$f_1(x\,|\,y) \overset{(3.3.4)}{=} \frac{f(x,y)}{f_2(y)} = \begin{cases} \dfrac{8xy}{4y-4y^3} = \dfrac{2x}{1-y^2}, \text{ kui } x \in (y,1)\,, \\ \\ 0, \text{ kui } x \notin (y,1)\,, \end{cases}$$

$$\mathrm{E}\,(X\,|\,y) \overset{(3.5.13)}{=} \textstyle\int_y^1 x \dfrac{2x}{1-y^2} dx = \dfrac{2}{3}\dfrac{y^2+y+1}{y+1},\; x \overset{(3.5.11)}{=} \dfrac{2}{3}\dfrac{y^2+y+1}{y+1}. \qquad \diamondsuit$$

Definitsioon 7. Pindasid võrranditega

$$z = \mathrm{E}\,(Z\,|\,x,y)\,, \tag{3.5.14}$$

$$y = \mathrm{E}\,(Y\,|\,x,z) \tag{3.5.15}$$

ja

$$x = \mathrm{E}\,(X\,|\,y,z)\,, \tag{3.5.16}$$

kus

$$\mathrm{E}\,(Z\,|\,x,y) = \int_{-\infty}^{+\infty} z\, f_3\,(z|x,y)\,dz, \tag{3.5.17}$$

$$\mathrm{E}\,(Y\,|\,x,z) = \int_{-\infty}^{+\infty} y\, f_2\,(y|x,z)\,dy, \tag{3.5.18}$$

$$\mathrm{E}\,(X\,|\,y,z) = \int_{-\infty}^{+\infty} x\, f_1\,(x|y,z)\,dx \tag{3.5.19}$$

ning

$$f_1(x|y,z) = f(x,y,z)/f(y,z), \tag{3.5.20}$$

$$f_2(y|x,z) = f(x,y,z)/f(x,z), \tag{3.5.21}$$

$$f_3(z|x,y) = f(x,y,z)/f(x,y), \tag{3.5.22}$$

nimetatakse juhusliku vektori (X,Y,Z) *regressioonipindadeks.*

Võrrandiga (3.5.14) antud regressioonipinna korral leitakse tinglik keskväärtus $\mathrm{E}\,(Z|\,x,y)$ eeldusel, et komponent X omandas katse käigus väärtuse x ja Y väärtuse y. Nii seatakse igale juhusliku vektori (X,Y) võimalikule väärtusele (x,y) vastavusse arv $\mathrm{E}\,(Z|\,x,y)$ ja saadakse regressioonipind võrrandiga (3.5.14).

Näide 3. Olgu $\Omega = \{(x, y, z) \mid x \geq 0 \wedge y \geq 0 \wedge z \geq 0 \wedge 4x + 2y + z \leq 4\}$ ja allugu vektor (X, Y, Z) ühtlasele jaotusele hulgal Ω. Leiame $f(x, y, z)$, vektori (X, Y) jaotustiheduse $f(x, y)$, vektori (Y, Z) jaotustiheduse $f(y, z)$, vektori (X, Z) jaotustiheduse $f(x, z)$. Leiame marginaalsed jaotustihedused $f_1(x)$, $f_2(y)$ ja $f_3(z)$ ning tinglikud jaotustihedused $f_2(y|x)$, $f_3(z|y)$ ja $f_3(z|x, y)$. Leiame vektori (X, Y, Z) kovariatsioonimaatriksi K ja korrelatsioonimaatriksi R ning regressioonipinna $z = \mathrm{E}(Z \mid x, y)$ ja regressioonijooned $y = \mathrm{E}(Y \mid x)$ ning $z = \mathrm{E}(Z \mid y)$.

Et Ω on püramiid ruumalaga $V = 1 \cdot 2 \cdot 4/6 = 4/3$, siis

$$f(x, y, z) = \begin{cases} 3/4, \text{ kui } (x, y, z) \in \Omega, \\ 0, \text{ kui } (x, y, z) \notin \Omega. \end{cases}$$

Püramiidi ristprojektsioon xy-tasandile on

$$\mathrm{pr}_{xy}\Omega = \{(x, y) \mid 0 \leq x \leq 1 \wedge 0 \leq y \leq 2 - 2x\}.$$

Analoogiliselt saame

$$\mathrm{pr}_{yz}\Omega = \{(y, z) \mid 0 \leq y \leq 2 \wedge 0 \leq z \leq 4 - 2y\}$$

ja

$$\mathrm{pr}_{xz}\Omega = \{(x, z) \mid 0 \leq x \leq 1 \wedge 0 \leq z \leq 4 - 4x\}.$$

Et $f(x, y) = \int_{-\infty}^{+\infty} f(x, y, z)dz$, siis saame leida kahedimensionaalse jaotustiheduse

$$f(x, y) = \begin{cases} \int_0^{4-4x-2y} \frac{3}{4} dz = 3 - 3x - \frac{3}{2}y, \text{ kui } (x, y) \in \mathrm{pr}_{xy}\Omega, \\ 0, \text{ kui } (x, y) \notin \mathrm{pr}_{xy}\Omega \end{cases}$$

ja analoogiliselt kaks ülejäänud kahedimensionaalset jaotustihedust

$$f(y, z) = \begin{cases} \int_0^{1-y/2-z/4} \frac{3}{4} dx = \frac{3}{4} - \frac{3}{8}y - \frac{3}{16}z, \text{ kui } (y, z) \in \mathrm{pr}_{yz}\Omega, \\ 0, \text{ kui } (y, z) \notin \mathrm{pr}_{yz}\Omega \end{cases}$$

ja

$$f(x, z) = \begin{cases} \int_0^{2-2x-z/2} \frac{3}{4} dy = \frac{3}{2} - \frac{3}{2}x - \frac{3}{8}z, \text{ kui } (x, z) \in \mathrm{pr}_{xz}\Omega, \\ 0, \text{ kui } (x, z) \notin \mathrm{pr}_{xz}\Omega. \end{cases}$$

Leiame marginaalsed jaotustihedused

$$f_1(x) \overset{(3.2.6)}{=} \begin{cases} \int_0^{2-2x} \left(3 - 3x - \frac{3}{2}y\right) dy = 3(x-1)^2, \text{ kui } x \in [0; 1], \\ 0, \text{ kui } x \notin [0; 1] \end{cases}$$

ja

$$f_2(y) \overset{(3.2.7)}{=} \begin{cases} \int_0^{1-y/2} \left(3 - 3x - \frac{3}{2}y\right) dx = \frac{3}{2}\left(\frac{y}{2} - 1\right)^2, \text{ kui } y \in [0; 2], \\ 0, \text{ kui } y \notin [0; 2]. \end{cases}$$

Analoogiliselt leiame

$$f_3(z) = \begin{cases} \int_0^{1-z/4} \left(\frac{3}{2} - \frac{3}{2}x - \frac{3}{8}z\right) dx = \frac{3}{4}\left(\frac{z}{4} - 1\right)^2, \text{ kui } z \in [0; 4], \\ 0, \text{ kui } z \notin [0; 4]. \end{cases}$$

Leiame soovitud ühedimensionaalsed tinglikud jaotustihedused

$$f_2(y|x) \overset{(3.3.5)}{=} \frac{f(x,y)}{f_1(x)} = \begin{cases} \frac{2 - 2x - y}{2(x-1)^2}, \text{ kui } y \in [0; 2-2x], \\ 0, \text{ kui } y \notin [0; 2-2x] \end{cases}$$

ja

$$f_3(z|y) = \frac{f(y,z)}{f_2(y)} = \begin{cases} \frac{4 - 2y - z}{2(y-2)^2}, \text{ kui } z \in [0; 4-2y], \\ 0, \text{ kui } z \notin [0; 4-2y]. \end{cases}$$

Leiame soovitud tingliku kahedimensionaalse jaotustiheduse

$$f_3(z|x,y) \overset{(3.5.22)}{=} \frac{f(x,y,z)}{f(x,y)} = \begin{cases} \frac{1}{2(2 - 2x - y)}, \text{ kui } z \in [0; 4-4x-2y], \\ 0, \text{ kui } z \notin [0; 4-4x-2y]. \end{cases}$$

Leiame vajalikud algmomendid, keskmomendid ja standardhälbed

$$\begin{aligned} \mathrm{E}X &= \int_0^1 x3(x-1)^2\,dx = 1/4,\ \mathrm{E}X^2 = \int_0^1 x^2 3(x-1)^2\,dx = 1/10, \\ \mathrm{E}Y &= \int_0^2 y\frac{3}{2}\left(\frac{y}{2}-1\right)^2 dy = 1/2,\ \mathrm{E}Y^2 = \int_0^2 y^2\frac{3}{2}\left(\frac{y}{2}-1\right)^2 dy = 2/5, \\ \mathrm{E}Z &= \int_0^4 z\frac{3}{4}\left(\frac{z}{4}-1\right)^2 dz = 1,\ \mathrm{E}Z^2 = \int_0^4 z^2\frac{3}{4}\left(\frac{z}{4}-1\right)^2 dz = \frac{8}{5}, \\ \mathrm{E}(XY) &= \int_0^1 xdx \int_0^{2-2x} \left(3 - 3x - \frac{3}{2}y\right) ydy = 1/10, \\ \mathrm{E}(XZ) &= \int_0^1 xdx \int_0^{4-4x} \left(\frac{3}{2} - \frac{3}{2}x - \frac{3}{8}z\right) zdz = 1/5, \\ \mathrm{E}(YZ) &= \int_0^2 ydy \int_0^{4-2y} \left(\frac{3}{4} - \frac{3}{8}y - \frac{3}{16}z\right) zdz = 2/5, \end{aligned}$$

$$\begin{aligned}
\mathrm{D}X &= \mathrm{E}\left(X^2\right) - (\mathrm{E}X)^2 = \frac{1}{10} - \left(\frac{1}{4}\right)^2 = \frac{3}{80},\ \sigma_x = \sqrt{3/80},\\
\mathrm{D}Y &= \mathrm{E}\left(Y^2\right) - (\mathrm{E}Y)^2 = \frac{2}{5} - \left(\frac{1}{2}\right)^2 = \frac{3}{20},\ \sigma_y = \sqrt{3/20},\\
\mathrm{D}Z &= \mathrm{E}\left(Z^2\right) - (\mathrm{E}Z)^2 = \frac{8}{5} - 1 = \frac{3}{5},\ \ \sigma_z = \sqrt{3/5},\\
\mathrm{cov}(X,Y) &= \mathrm{E}\,(XY) - (\mathrm{E}X)\,(\mathrm{E}Y) = \frac{1}{10} - \frac{1}{4\cdot 2} = -\frac{1}{40},\\
\mathrm{cov}(X,Z) &= \mathrm{E}\,(XZ) - (\mathrm{E}X)\,(\mathrm{E}Z) = \frac{1}{5} - \frac{1}{4\cdot 1} = -\frac{1}{20},\\
\mathrm{cov}(Y,Z) &= \mathrm{E}\,(YZ) - (\mathrm{E}Y)\,(\mathrm{E}Z) = \frac{2}{5} - \frac{1}{2} = -\frac{1}{10}.
\end{aligned}$$

Kirjutame (3.5.8) põhjal välja vektori (X, Y, Z) kovariatsioonimaatriksi

$$\mathrm{K} = \begin{pmatrix} \mathrm{cov}(X,X) & \mathrm{cov}(X,Y) & \mathrm{cov}(X,Z)\\ & \mathrm{cov}(Y,Y) & \mathrm{cov}(X,Z)\\ & & \mathrm{cov}(Z,Z) \end{pmatrix} = \begin{pmatrix} 3/80 & -1/40 & -1/20\\ & 3/20 & -1/10\\ & & 3/5 \end{pmatrix}.$$

Et valemi (3.5.4) põhjal

$$\begin{aligned}
\mathrm{r}_{xy} &= \frac{\mathrm{cov}(X,Y)}{\sigma_X\sigma_Y} = -\frac{1}{40}/\left(\sqrt{3/80}\sqrt{3/20}\right) = -\frac{1}{3},\\
\mathrm{r}_{xz} &= -\frac{1}{20}/\left(\sqrt{3/80}\sqrt{3/5}\right) = -\frac{1}{3},\\
\mathrm{r}_{yz} &= -\frac{1}{10}/\left(\sqrt{3/20}\sqrt{3/5}\right) = -\frac{1}{3},
\end{aligned}$$

siis korrelatsioonimaatriks on kujul

$$\mathrm{R} \overset{(3.5.9)}{=} \begin{pmatrix} \mathrm{r}(X,X) & \mathrm{r}(X,Y) & \mathrm{r}(X,Z)\\ & \mathrm{r}(Y,Y) & \mathrm{r}(X,Z)\\ & & \mathrm{r}(Z,Z) \end{pmatrix} = \begin{pmatrix} 1 & -1/3 & -1/3\\ & 1 & -1/3\\ & & 1 \end{pmatrix}.$$

Leiame soovitud regressioonipinna. Et

$$\mathrm{E}\,(Z\,|\,x,y) \overset{(3.5.17)}{=} \int_{-\infty}^{+\infty} z f_3(z\,|\,x,y)dz \overset{(3.5.22)}{=}$$

$$= \begin{cases} \int_0^{4-4x-2y} \dfrac{z}{4-4x-2y}dz = 2-2x-y,\ \text{kui}\ \ (x,y)\in \mathrm{pr}_{xy}\Omega,\\[2ex] 0,\ \text{kui}\ \ (x,y)\notin \mathrm{pr}_{xy}\Omega, \end{cases}$$

siis

$$z = \begin{cases} 2-2x-y,\ \text{kui}\ \ (x,y)\in \mathrm{pr}_{xy}\Omega,\\[2ex] 0,\ \text{kui}\ \ (x,y)\notin \mathrm{pr}_{xy}\Omega \end{cases}$$

on regressioonipinna võrrand. Leiame regressioonijooned. Et

$$\mathrm{E}(Y \mid x) \overset{(3.5.12)}{=} \begin{cases} \displaystyle\int_0^{2-2x} \frac{(2-2x-y)\,y}{2(x-1)^2} dy = \frac{2}{3} - \frac{2}{3}x, \text{ kui } x \in [0;1], \\ 0, \text{ kui } x \notin [0;1], \end{cases}$$

$$z = \mathrm{E}(Z \mid y) = \begin{cases} \displaystyle\int_0^{4-2y} \frac{1}{2} \frac{4-2y-z}{(y-2)^2} z dz = 4/3 - 2y/3, \text{ kui } y \in [0;2], \\ 0, \text{ kui } y \notin [0;2], \end{cases}$$

siis

$$y = \begin{cases} 2/3 - 2x/3, \text{ kui } x \in [0;1], \\ 0, \text{ kui } x \notin [0;1] \end{cases}$$

ja

$$z = \begin{cases} 4/3 - 2y/3, \text{ kui } y \in [0;2], \\ 0, \text{ kui } y \notin [0;2] \end{cases}$$

on soovitud regressioonijoonte võrrandid. $\diamondsuit$

3.6 Juhusliku vektori normaaljaotus

Olgu juhusliku vektori $(X_1, \ldots, X_n)$ komponendid $X_i \sim N(a_i, \sigma_i)$ sõltumatud. Et sõltumatute komponentidega vektori jaotustihedus $f(x_1, \ldots, x_n)$ avaldub komponentide jaotustiheduste

$$f_i(x_i) = \frac{1}{\sqrt{2\pi}\sigma_i} e^{-\frac{(x_i - a_i)^2}{2\sigma_i^2}}$$

korrutisena, siis

$$f(x_1, \ldots, x_n) = \frac{1}{\left(\sqrt{2\pi}\right)^n \sigma_1 \cdots \sigma_n} e^{-\frac{(x_1-a_1)^2}{2\sigma_1^2} - \ldots - \frac{(x_n-a_n)^2}{2\sigma_n^2}} =$$

$$= \frac{1}{\left(\sqrt{2\pi}\right)^n \sigma_1 \cdots \sigma_n} e^{-\frac{1}{2}\left(\frac{(x_1-a_1)^2}{\sigma_1^2} + \ldots + \frac{(x_n-a_n)^2}{\sigma_n^2}\right)}.$$

Vektori jaotustiheduse nivoopindu võrranditega

$$\frac{(x_1-a_1)^2}{\sigma_1^2} + \ldots + \frac{(x_n-a_n)^2}{\sigma_n^2} = k^2 \qquad (k \geq 0)$$

nimetatakse vektori *hajuvusellipsoidideks*.

Järgnevalt piirdume juhuga $n = 2$.

Lause 1. Kui juhusliku vektori (X, Y) komponendid $X \sim N(a_x, \sigma_x)$ ja $Y \sim N(a_y, \sigma_y)$ on sõltumatud, siis

$$f(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} e^{-\frac{1}{2}\left(\frac{(x-a_x)^2}{\sigma_x^2} + \frac{(y-a_y)^2}{\sigma_y^2}\right)} \tag{3.6.1}$$

on vektori (X, Y) jaotustihedus ning

$$\frac{(x-a_x)^2}{\sigma_x^2} + \frac{(y-a_y)^2}{\sigma_y^2} = k^2 \tag{3.6.2}$$

on selle vektori *hajuvusellipsi* võrrand $k \geq 0$ korral.

Näide 1. Olgu vektori (X, Y) jaotustihedus antud seosega (3.6.1). Leiame tõenäosuse, et juhuslik vektor (X, Y) omandab katse käigus väärtuse hajuvusellipsiga (3.6.2) piiratud piirkonnast. Leiame parameetri k väärtuse, mille korral on vektori (X, Y) hajuvusellipsisse sattumise tõenäosus 0.5.

Saame

$$P\left(\frac{(X-a_x)^2}{\sigma_x^2} + \frac{(Y-a_y)^2}{\sigma_y^2} \leq k^2\right) =$$

$$= \iint\limits_{\frac{(x-a_x)^2}{\sigma_x^2} + \frac{(y-a_y)^2}{\sigma_y^2} \leq k^2} \frac{1}{2\pi\sigma_x\sigma_y} e^{-\frac{1}{2}\left(\frac{(x-a_x)^2}{\sigma_x^2} + \frac{(y-a_y)^2}{\sigma_y^2}\right)} dxdy =$$

$$= \left[x = a_x + k\sigma_x\rho\cos\varphi,\ y = a_y + k\sigma_y\rho\sin\varphi,\ J = k^2\sigma_x\sigma_y\rho\right] =$$

$$= \frac{k^2}{2\pi}\int_0^{2\pi} d\varphi \int_0^1 \rho e^{-\frac{k^2\rho^2}{2}} d\rho = -\frac{1}{2\pi}\int_0^{2\pi} d\varphi \left(e^{-\frac{k^2\rho^2}{2}}\right)\Bigg|_0^1 = 1 - e^{-\frac{k^2}{2}}.$$

Leiame parameetri k väärtuse, mille korral on vektori (X, Y) hajuvusellipsisse sattumise tõenäosus 0.5. Selleks lahendame võrrandi

$$1 - e^{-k^2/2} = 1/2.$$

Saame

$$e^{-k^2/2} = 1/2 \ \Leftrightarrow\ \ln e^{-k^2/2} = \ln\frac{1}{2} \ \Leftrightarrow\ -k^2/2 = \ln 1 - \ln 2$$
$$\Leftrightarrow\ k = \sqrt{2\ln 2}. \qquad \diamondsuit$$

Kehtib väide.

Lause 2. Kui r_{xy} on juhuslike suuruste $X \sim N(a_x, \sigma_x)$ ja $Y \sim N(a_y, \sigma_y)$ korrelatsioonikordaja, siis

$$f(x,y) = \frac{e^{-\frac{1}{2\left(1-\mathrm{r}_{xy}^2\right)}\left[\frac{(x-a_x)^2}{\sigma_x^2} - \frac{2\mathrm{r}_{xy}(x-a_x)(y-a_y)}{\sigma_x\sigma_y} + \frac{(y-a_y)^2}{\sigma_y^2}\right]}}{2\pi\sigma_x\sigma_y\sqrt{1-\mathrm{r}_{xy}^2}} \tag{3.6.3}$$

on vektori (X,Y) jaotustihedus.

Definitsioon 1. Vektori (X,Y) jaotust nimetatakse *kahemõõtmeliseks normaaljaotuseks*, kui

$$f(x,y) = \frac{e^{-\frac{1}{2(1-\mathrm{r}^2)}\left[\frac{(x-a_1)^2}{\sigma_1^2} - \frac{2\mathrm{r}(x-a_1)(y-a_2)}{\sigma_1\sigma_2} + \frac{(y-a_2)^2}{\sigma_2^2}\right]}}{2\pi\sigma_1\sigma_2\sqrt{1-\mathrm{r}^2}} \tag{3.6.4}$$

on vektori (X,Y) jaotustihedus, kusjuures $\sigma_1, \sigma_2 > 0$, $|\mathrm{r}| < 1$ ja $a_1, a_2 \in \mathbf{R}$.

Kehtib järgmine väide.

Lause 3. Kui vektor (X,Y) on kahemõõtmelise normaaljaotusega, st selle vektori jaotustihedus on kujul (3.6.4), siis $X \sim N(a_1, \sigma_1)$, $Y \sim N(a_2, \sigma_2)$ ja $\mathrm{r}_{xy} = \mathrm{r}$ ning

$$y = a_2 + \mathrm{r}\frac{\sigma_2}{\sigma_1}(x - a_1) \tag{3.6.5}$$

ja

$$x = a_1 + \mathrm{r}\frac{\sigma_1}{\sigma_2}(y - a_2) \tag{3.6.6}$$

on vektori (X,Y) regressioonisirgete võrrandid.

Seostest (3.6.1), (3.6.3) ja (3.6.4) järeldub väide.

Järeldus 1. Kahemõõtmelise normaaljaotusega vektori (X,Y) komponendid on sõltumatud parajasti siis, kui nad on mittekorreleeruvad.

3.7 Juhusliku argumendiga funktsioonid

Olgu X juhuslik suurus, mille jaotustihedus on $f_1(x)$, ja $\varphi(x)$ kindel (determineeritud) funktsioon.

Definitsioon. Funktsiooni $Y = \varphi(X)$ nimetatakse *juhusliku argumendiga funktsiooniks*. Juhusliku argumendiga funktsiooni $\varphi(X)$ väärtused Y on juhuslikud suurused. Vaatleme järgnevalt, kuidas leida suuruse Y jaotustihedust $f_2(y)$, kui on teada $f_1(x)$ ja $\varphi(x)$. Esitame kolm võimalust funktsiooni $f_2(y)$ leidmiseks.

I Moodustame juhusliku vektori (X, Y). Leiame selle vektori jaotustiheduse $f(x, y)$. Teame, et

$$f(x, y) = f_1(x)\, f_2(y|x).$$

Kui on teada, et juhuslik suurus X omandab katse käigus väärtuse x, siis juhuslikul suurusel Y on vaid üks võimalik väärtus $\varphi(x)$, mille ta omandab tõenäosusega 1. Seega

$$f_2(y|x) = 1 \cdot \delta(y - \varphi(x)) \Rightarrow f(x, y) = f_1(x)\delta(y - \varphi(x)).$$

Teades juhusliku vektori (X, Y) jaotustihedust, saame leida selle vektori teise komponendi Y jaotustiheduse

$$f_2(y) = \int_{-\infty}^{+\infty} f(x, y)dx = \int_{-\infty}^{+\infty} f_1(x)\delta(y - \varphi(x))dx.$$

Lause 1. Kui pideva juhusliku suuruse X jaotustihedus on $f_1(x)$ ja $Y = \varphi(X)$, kus $\varphi(x)$ on kindel funktsioon, siis juhusliku suuruse Y jaotustihedus avaldub kujul

$$f_2(y) = \int_{-\infty}^{+\infty} f_1(x)\delta(y - \varphi(x))dx. \tag{3.7.1}$$

Näide 1. Allugu juhuslik suurus X ühtlasele jaotusele lõigul $[a, b]$. Leiame juhusliku suuruse $Y = kX + c$ $(k > 0)$ jaotustiheduse.

Et

$$f_1(x) = \begin{cases} \dfrac{1}{b-a}, \text{ kui } x \in [a, b]\,, \\ 0, \text{ kui } x \notin [a, b]\,, \end{cases}$$

siis valemi (3.7.1) abil saame

$$\begin{aligned} f_2(y) &= \int_{-\infty}^{+\infty} f_1(x)\delta(y - \varphi(x))dx = \int_a^b \frac{1}{b-a}\delta(y - (kx + c))dx = \\ &= \left[u = kx + c,\ x = \frac{u-c}{k}, \quad dx = \frac{du}{k}\right] = \\ &= \frac{1}{b-a}\int_{ka+c}^{kb+c} \delta(y - u)\frac{du}{k} = [\text{deltafunktsioon on paarisfunktsioon}] = \\ &= \frac{1}{k\,(b-a)}\int_{ka+c}^{kb+c} 1 \cdot \delta(u - y)du = \\ &= \left[\int_\beta^\gamma h(u)\delta(u-\alpha)du = \begin{cases} h(\alpha), \text{ kui } \alpha \in (\beta, \gamma) \wedge\ h(u) \in C(\alpha), \\ 0, \text{ kui } \alpha \notin (\beta, \gamma) \end{cases}\right] = \\ &= \begin{cases} \dfrac{1}{k\,(b-a)}, \text{ kui } y \in (ka + c, kb + c) \\ 0, \text{ kui } y \notin (ka + c, kb + c)\,, \end{cases} \end{aligned}$$

st suurus Y allub ühtlasele jaotusele vahemikus $(ka+c, kb+c)$. $\diamondsuit$

Lause 2. Kui
1° pideva juhusliku suuruse X võimalike väärtuste hulk on vahemik või poollõik või lõik otspunktidega a ja b, kusjuures poollõigu korral võib üks otspunktidest olla lõpmatu ja vahemiku korral kas üks või mõlemad olla lõpmatud,
2° $f_1(x)$ on suuruse X jaotustihedus ning $Y=\varphi(X)$,
3° $\varphi(x)$ on suuruse X võimalike väärtuste hulgal diferentseeruv rangelt monotoonne funktsioon, siis juhusliku suuruse Y jaotustihedus avaldub kujul

$$f_2(y)=\begin{cases} f_1(\psi(y))\left|\psi'(y)\right|, \text{ kui } y\in(\min\{\varphi(a),\varphi(b)\},\max\{\varphi(a),\varphi(b)\}), \\ \\ 0, \text{ kui } y\notin(\min\{\varphi(a),\varphi(b)\},\max\{\varphi(a),\varphi(b)\}), \end{cases} \tag{3.7.2}$$

kusjuures funktsioon $x=\psi(y)$ on funktsiooni $y=\varphi(x)$ pöördfunktsioon ja

$$\begin{aligned} &\varphi(a)=-\infty \vee \varphi(b)=-\infty \ \Rightarrow\ \min\{\varphi(a),\varphi(b)\}=-\infty, \\ &\varphi(a)=+\infty \vee \varphi(b)=+\infty \ \Rightarrow\ \max\{\varphi(a),\varphi(b)\}=+\infty. \end{aligned}$$

Tõestus. Olgu funktsioon $\varphi(x)$ rangelt kasvav juhusliku suuruse X võimalike väärtuste hulgal. Et funktsioon $y=\varphi(x)$ on rangelt kasvav ja diferentseeruv sel hulgal, siis eksisteerib tal diferentseeruv pöördfunktsioon $\psi(y)$ vahemikus $(\min\{\varphi(a),\varphi(b)\},\max\{\varphi(a),\varphi(b)\})$. Lähtume seosest (3.7.1). Saame

$$\begin{aligned} f_2(y) &= \int_{-\infty}^{+\infty} f_1(x)\delta(y-\varphi(x))dx = \\ &= [u=\varphi(x)\Leftrightarrow x=\psi(u)\Rightarrow dx=\psi'(u)du] = \\ &= \int_{\varphi(a)}^{\varphi(b)} f_1(\psi(u))\delta(y-u)\psi'(u)du = \int_{\varphi(a)}^{\varphi(b)} f_1(\psi(u))\psi'(u)\delta(u-y)du = \\ &= \begin{cases} f_1(\psi(y))\psi'(y), \text{ kui } y\in(\varphi(a),\varphi(b)), \\ 0, \text{ kui } y\notin(\varphi(a),\varphi(b)), \end{cases} = \\ &= \begin{cases} f_1(\psi(y))\left|\psi'(y)\right|, \text{ kui } y\in(\min\{\varphi(a),\varphi(b)\},\max\{\varphi(a),\varphi(b)\}), \\ 0, \text{ kui } y\notin(\min\{\varphi(a),\varphi(b)\},\max\{\varphi(a),\varphi(b)\}). \end{cases} \end{aligned}$$

Kui funktsioon $\varphi(x)$ on rangelt kahanev, siis $\varphi(a)>\varphi(b)$, $\psi'(y)<0$ ja $\left|\psi'(u)\right|=-\psi'(u)$ ning

$$\begin{aligned} f_2(y) &= \ldots = -\int_{\varphi(a)}^{\varphi(b)} f_1(\psi(u))\left|\psi'(u)\right|\delta(u-y)du = \\ &= \int_{\varphi(b)}^{\varphi(a)} f_1(\psi(u))\left|\psi'(u)\right|\delta(u-y)du = \\ &= \int_{\min\{\varphi(a),\varphi(b)\}}^{\max\{\varphi(a),\varphi(b)\}} f_1(\psi(u))\left|\psi'(u)\right|\delta(u-y)du = \\ &= \begin{cases} f_1(\psi(y))\left|\psi'(y)\right|, \text{ kui } y\in(\min\{\varphi(a),\varphi(b)\},\max\{\varphi(a),\varphi(b)\}), \\ 0, \text{ kui } y\notin(\min\{\varphi(a),\varphi(b)\},\max\{\varphi(a),\varphi(b)\}). \end{cases} \end{aligned}$$

Seega on Lause 2 tõestatud. □

Märkus 1. Lause 2 põhjal on funktsiooni $f_2(y)$ kandja, st hulk, millel $f_2(y) \neq 0$, lahtine. See on tingitud kasutatud tõestusmetoodikast. Et pideva juhusliku suuruse korral on iga võimaliku väärtuse omandamise tõenäosus 0, siis võime vahemiku ülesande sisust lähtudes asendada kas lõigu või poollõiguga.

Sõnastage Näite 1 vastus, arvestades Märkust 1.

Näide 2. Olgu

$$f_1(x) = \begin{cases} 2x, \text{ kui } x \in [0;1], \\ 0, \text{ kui } x \notin [0;1] \end{cases}$$

ja $Y = X^2$. Leiame $f_2(y)$.

Et funktsioon $\varphi(x) = x^2$ on rangelt kasvav lõigul $[0;1]$, siis on rakendatav Lause 2, kusjuures $\psi(y) = \sqrt{y}$. Valemi (3.7.2) abil saame

$$f_2(y) = \begin{cases} 2\sqrt{y} \cdot \dfrac{1}{2\sqrt{y}} = 1, \text{ kui } y \in (0;1), \\ \\ 0, \qquad \text{kui } y \notin (0;1). \end{cases}$$

Märkuse 1 põhjal võime suuruse Y jaotustiheduse esitada kujul

$$f_2(y) = \begin{cases} 2\sqrt{y} \cdot \dfrac{1}{2\sqrt{y}} = 1, \text{ kui } y \in [0;1], \\ \\ 0, \qquad \text{kui } y \notin [0;1]. \end{cases}$$

Seega allub juhuslik suurus Y lõigul $[0;1]$ ühtlasele jaotusele. ♢

Näide 3. Allugu juhuslik suurus X lõigul $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ ühtlasele jaotusele. Leiame suuruse $Y = \cos X$ jaotustiheduse.

Funktsioon $\cos x$ ei ole rangelt monotoonne antud lõigul. Lõik tuleb jagada osalõikudeks $\left[-\frac{\pi}{2}, 0\right]$, millel $\cos x$ on rangelt kasvav, ja $\left[0, \frac{\pi}{2}\right]$, millel on rangelt kahanev. Kasutame Lauset 1. Seose (3.7.1) abil saame

$$f_2(y) = \int_{-\infty}^{+\infty} f_1(x)\delta(y - \varphi(x))dx =$$

$$= \int_{-\pi/2}^{0} \frac{1}{\pi}\delta(y - \cos x)dx + \int_{0}^{\pi/2} \frac{1}{\pi}\delta(y - \cos x)dx =$$

$$= \left[\begin{array}{cc} \text{I integraal} & \text{II integraal} \\ u = \cos x & u = \cos x \\ x = -\arccos u & x = \arccos u \\ dx = \dfrac{du}{\sqrt{1-u^2}} & dx = -\dfrac{du}{\sqrt{1-u^2}} \end{array} \right] =$$

$$= \int_0^1 \frac{1}{\pi}\delta(y-u)\frac{du}{\sqrt{1-u^2}} - \int_1^0 \frac{1}{\pi}\delta(y-u)\frac{du}{\sqrt{1-u^2}} =$$

$$= 2\int_0^1 \frac{1}{\pi}\delta(y-u)\frac{du}{\sqrt{1-u^2}} = \int_0^1 \frac{2}{\pi\sqrt{1-u^2}}\delta(u-y)du =$$

$$= \begin{cases} \dfrac{2}{\pi\sqrt{1-y^2}}, \text{ kui } y \in (0;1)\,, \\ 0, \text{ kui } y \notin (0;1)\,. \qquad \diamondsuit \end{cases}$$

II Leiame juhusliku suuruse $Y = \varphi(X)$ karakteristliku funktsiooni

$$g_2(\omega) = \mathrm{E}e^{i\omega Y} = \mathrm{E}e^{i\omega\varphi(X)} = \int_{-\infty}^{+\infty} e^{i\omega\varphi(x)} f_1(x)\, dx.$$

Järgmisena saame Fourier' pöördteisenduse abil jaotustiheduse

$$f_2(y) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-i\omega y} g_2(\omega)\, d\omega.$$

Lause 3. Kui $f_1(x)$ on juhusliku suuruse X jaotustihedus ja $Y = \varphi(X)$, siis

$$f_2(y) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-i\omega y} g_2(\omega)\, d\omega, \tag{3.7.3}$$

kus

$$g_2(\omega) = \int_{-\infty}^{+\infty} e^{i\omega\varphi(x)} f_1(x)\, dx. \tag{3.7.4}$$

Lahendame Lause 3 abil Näite 2. Valemit (3.7.4) rakendades saame

$$g_2(\omega) = \int_0^1 e^{i\omega x^2} 2x dx = \left.\frac{e^{i\omega x^2}}{i\omega}\right|_0^1 = \frac{e^{i\omega}-1}{i\omega}.$$

Kuigi juhusliku suuruse Y jaotustiheduse saame avaldada valemi (3.7.3) abil

$$f_2(y) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-i\omega y}\frac{e^{i\omega}-1}{i\omega}d\omega,$$

on selle integraali leidmine keerukas probleem. Näites 2 leidsime, et Y allub lõigul $[0;1]$ ühtlasele jaotusele. Kontrollime, et $\left(e^{i\omega}-1\right)/(i\omega)$ on lõigul $[0;1]$ ühtlasele jaotusele alluva juhusliku suuruse Y karakteristlik funktsioon

$$\int_0^1 1\cdot e^{i\omega x}dx = \left.\frac{e^{i\omega x}}{i\omega}\right|_0^1 = \frac{e^{i\omega}-1}{i\omega}.$$

Seega võib väita, et Lause 3 väide on teoreetilist laadi ja tehniliselt keerukas konkreetsete ülesannete lahendamisel.

III Leiame juhusliku suuruse $Y = \varphi(X)$ jaotusfunktsiooni

$$F_2(y) = P\left(Y < y\right) = P\left(\varphi(X) < y\right) = \int\limits_{\varphi(x)<y} f_1(x)dx$$

abil jaotustiheduse

$$f_2(y) = \frac{dF_2(y)}{dy} = \frac{d}{dy}\int\limits_{\varphi(x)<y} f_1(x)dx.$$

Märgime, et juhuslik sündmus $\varphi(X) < y$ toimub parajasti siis, kui X omandab katse käigus sellise võimaliku väärtuse x, mille korral $\varphi(x) < y$. Seega sündmus $\varphi(X) < y$ toimub parajasti siis, kui juhuslik suurus X omandab katse käigus väärtuse võimalike väärtuste hulga sellest alamhulgast, mille elemendid rahuldavad tingimust $\varphi(x) < y$. Järelikult tuleb sündmuse $\varphi(X) < y$ tõenäosuse leidmisel suuruse X jaotustihedust $f_1(x)$ integreerida üle suuruse X võimalike väärtuste hulga tingimusega $\varphi(x) < y$ määratud alamhulga.

Lause 4. Kui $Y = \varphi(X)$ ja $f_1(x)$ on juhusliku suuruse X jaotustihedus, siis suuruse $Y = \varphi(X)$ jaotustihedus avaldub kujul

$$f_2(y) = \frac{d}{dy}\int\limits_{\varphi(x)<y} f_1(x)dx. \tag{3.7.5}$$

Lahendame Näite 2 Lause 4 abil. Rakendame valemit (3.7.5). Et

$$\int\limits_{x<\sqrt{y}} f_1(x)dx =$$

$$= \int\limits_{(x\in[0;1])\wedge\left(x<\sqrt{y}\right)} 2xdx = \begin{cases} \int_0^{\sqrt{y}} 2x\,dx = y, \text{ kui } 0 \le y \le 1, \\ 0, \text{ kui } y < 0, \\ 1, \text{ kui } y > 1, \end{cases}$$

siis saame

$$\frac{d}{dy}\int\limits_{\varphi(x)<y} f_1(x)dx = \begin{cases} 1, \text{ kui } y \in (0;1)\,, \\ 0, \text{ kui } y \notin (0;1)\,. \end{cases}$$

Lahendame ka Näite 3 Lause 4 abil. Märgime, et

$$X \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \quad \Rightarrow \quad Y = \cos X \in [0;1]\,.$$

Rakendame valemit (3.7.5). Et

$$\int\limits_{\left(x\in\left[-\frac{\pi}{2},\frac{\pi}{2}\right]\right)\wedge(\cos x\,<\,y)} f_1(x)dx =$$

$$= \left[\begin{array}{c} \text{tingimus } \cos x \ < \ y \quad (0 < y < 1) \text{ on rahuldatud} \\ X \text{ võimalike väärtuste hulga alamhulkadel} \\ \left[-\frac{\pi}{2}, -\arccos y\right) \text{ ja } \left(\arccos y, \frac{\pi}{2}\right] \end{array} \right] =$$

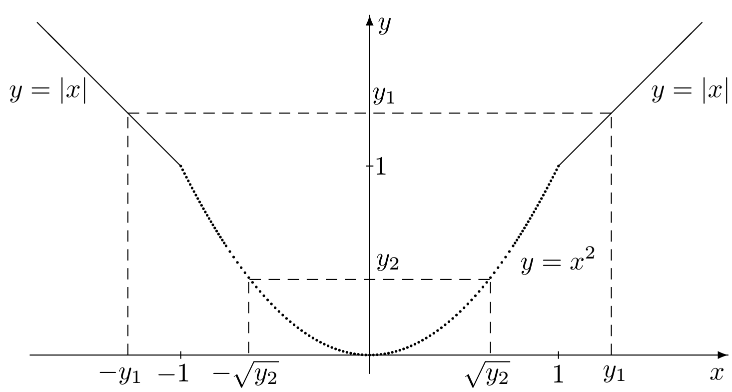
$$= \begin{cases} 2\int_{\arccos y}^{\pi/2} \frac{1}{\pi}\, dx = \dfrac{\pi - 2\arccos y}{\pi}, \text{ kui } 0 \le y \le 1, \\ 0, \text{ kui } y < 0, \\ 1, \text{ kui } y > 1, \end{cases}$$

siis

$$f_2(y) = \frac{d}{dy} \int\limits_{\varphi(x)<y} f_1(x)dx = \begin{cases} \dfrac{2}{\pi\sqrt{1-y^2}}, \text{ kui } y \in [0;1)\,, \\ 0, \text{ kui } y \notin [0;1)\,. \end{cases}$$

Näide 4. Olgu juhusliku suuruse X jaotustihedus $f_1(x)$ pidev hulgal $\mathbf{R}$. Leiame juhusliku suuruse $Y = \min\left\{X^2, |X|\right\}$ jaotustiheduse $f_2(y)$.

Skitseerime funktsiooni $y = \min\left\{x^2, |x|\right\}$ graafiku



Olgu $y_1 > 1$. Sel korral

$$P(Y < y_1) = P\left(\min\left\{X^2, |X|\right\} < y_1\right) = [\text{vt joonist}] = P(X \in (-y_1, y_1))\,,$$

st (Definitsioon 1.1.2) neist sündmustest ühe toimumisega kaasneb ülejäänud kahe sündmuse toimumine. Seega saame $y > 1$ korral

$$P(Y < y) = P(X \in (-y, y)) = \int_{-y}^{y} f_1(x)dx$$

ja valemi (3.7.5) põhjal

$$\begin{aligned} f_2(y) &= \frac{d}{dy}\int_{-y}^{y} f_1(x)dx = \frac{d}{dy}\left(\int_{-y}^{0} f_1(x)dx + \int_{0}^{y} f_1(x)dx\right) = \\ &= \frac{d}{dy}\left(\int_{0}^{y} f_1(x)dx - \int_{0}^{-y} f_1(x)dx\right) = \\ &= \frac{d}{dy}\int_{0}^{y} f_1(x)dx - \frac{d}{d(-y)}\int_{0}^{-y} f_1(x)dx \cdot \frac{d(-y)}{dy} = \\ &= f_1(y) - f_1(-y)(-1) = f_1(y) + f_1(-y). \end{aligned}$$

Olgu $0 < y_2 < 1$. Sel korral

$$(Y < y_2) = \left(\min\left\{X^2, |X|\right\} < y_2\right) = \left(X \in (-\sqrt{y_2}, \sqrt{y_2})\right).$$

Seega saame $0 < y < 1$ korral

$$P(Y < y) = P\left(X \in (-\sqrt{y}, \sqrt{y})\right) = \int_{-\sqrt{y}}^{\sqrt{y}} f_1(x)dx$$

ja valemi (3.7.5) põhjal

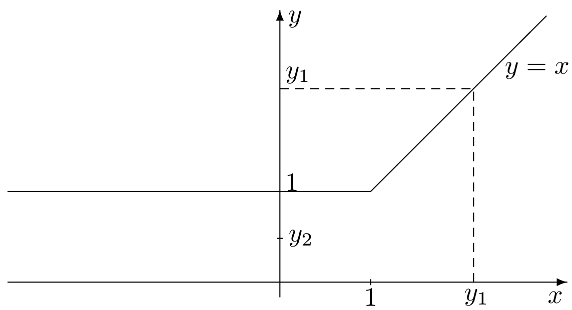
$$\begin{aligned} f_2(y) &= \frac{d}{dy}\int_{-\sqrt{y}}^{\sqrt{y}} f_1(x)dx = \frac{d}{dy}\left(\int_{-\sqrt{y}}^{0} f_1(x)dx + \int_{0}^{\sqrt{y}} f_1(x)dx\right) = \\ &= \frac{d}{dy}\left(\int_{0}^{\sqrt{y}} f_1(x)dx - \int_{0}^{-\sqrt{y}} f_1(x)dx\right) = \\ &= \frac{d}{d\sqrt{y}}\int_{0}^{\sqrt{y}} f_1(x)dx \cdot \frac{d\sqrt{y}}{dy} - \frac{d}{d\left(-\sqrt{y}\right)}\int_{0}^{-\sqrt{y}} f_1(x)dx \cdot \frac{d\left(-\sqrt{y}\right)}{dy} = \\ &= f_1(\sqrt{y})\frac{1}{2\sqrt{y}} - f_1(-\sqrt{y})\left(-\frac{1}{2\sqrt{y}}\right) = \frac{f_1(\sqrt{y}) + f_1(-\sqrt{y})}{2\sqrt{y}}. \end{aligned}$$

Kui $y < 0$, siis sündmus $(Y < y) = \left(\min\left\{X^2, |X|\right\} < y\right)$ on võimatu ja $F_2(y) = 0$. Et juhusliku suuruse X pidevusest järeldub suuruse Y pidevus (miks?), siis

$$f_2(y) = \begin{cases} 0, \text{ kui } y \le 0, \\ \\ \dfrac{f_1(\sqrt{y}) + f_1(-\sqrt{y})}{2\sqrt{y}}, \text{ kui } 0 < y < 1, \\ \\ f_1(y) + f_1(-y), \text{ kui } y \ge 0. \end{cases} \qquad \diamondsuit$$

Näide 5. Olgu juhusliku suuruse X jaotustihedus $f_1(x)$ pidev hulgal $\mathbf{R}$. Leiame segatüüpi juhusliku suuruse $Y = \max\{X, 1\}$ jaotustiheduse $f_2(y)$.

Skitseerime funktsiooni $y = \max\{x, 1\}$ graafiku



Olgu $y_1 > 1$. Sel korral

$$\begin{aligned} P(Y < y_1) &= P(\max\{X, 1\} < y_1) = [\text{vt joonist}] = P(X \in (-\infty, y_1)) \\ &= \int_{-\infty}^{y_1} f_1(x)dx. \end{aligned}$$

Seega saame $y > 1$ korral

$$P(Y < y) = P(X \in (-\infty, y)) = \int_{-\infty}^{y} f_1(x)dx$$

ja valemi (3.7.5) põhjal

$$f_2(y) = \frac{d}{dy}\int_{-\infty}^{y} f_1(x)dx = f_1(y).$$

Olgu $y_2 < 1$ (vt joonist). Sel korral on sündmus $Y < y_2$ võimatu ja

$$P(Y < y_2) = 0.$$

Seega $y < 1$ korral saame $P(Y < y) = 0$. Kuna lisaks

$$P(Y = 1) = P(X \in (-\infty, 1)) = \int_{-\infty}^{1} f_1(x)dx,$$

siis $y \in \mathbf{R}$ korral

$$F_2(y) = P(Y < y) = \left(\int_{-\infty}^{y} f_1(x)dx\right)\mathbf{1}(y-1)$$

ja

$$\begin{aligned} f_2(y) = \frac{dF_2(y)}{dy} &= \frac{d}{dy}\left(\left(\int_{-\infty}^{y} f_1(x)dx\right)\mathbf{1}(y-1)\right) = \\ &= f_1(y)\mathbf{1}(y-1) + \left(\int_{-\infty}^{y} f_1(x)dx\right)\delta(y-1) \overset{\text{miks?}}{=} \\ &= f_1(y)\mathbf{1}(y-1) + \left(\int_{-\infty}^{1} f_1(x)dx\right)\delta(y-1). \qquad \diamondsuit \end{aligned}$$

3.8 Hii-ruut-jaotus

Definitsioon 1. Öeldakse, et juhuslik suurus

$$Y_n = X_1^2 + X_2^2 + \ldots + X_n^2 \tag{3.8.1}$$

allub χ^2-*jaotusele* ehk *hii-ruut-jaotusele,* kui $X_k \sim N(0;1)$ $(k = 1; 2; \ldots; n)$ on sõltumatud juhuslikud suurused. Arvu n nimetatakse χ^2-jaotusele alluva juhusliku suuruse Y_n *vabadusastmete arvuks.*

Leiame suuruse $Y_1 = X^2$, kus $X \sim N(0;1)$, karakteristliku funktsiooni. Olgu $Y \sim N(0;1)$. Et $g_{Y^2}(\omega) = g_{X^2}(\omega)$ ja

$$g_{X^2}(\omega) = \mathrm{E}e^{i\omega X^2} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{i\omega x^2} e^{-x^2/2} dx,$$

siis

$$\begin{aligned}
g_{X^2}(\omega) &= \sqrt{\left(\mathrm{E}e^{i\omega X^2}\right) \cdot \left(\mathrm{E}e^{i\omega X^2}\right)} = \sqrt{\left(\mathrm{E}e^{i\omega X^2}\right) \cdot \left(\mathrm{E}e^{i\omega Y^2}\right)} = \\
&= \sqrt{\left(\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{i\omega x^2} e^{-x^2/2} dx\right) \cdot \left(\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{i\omega y^2} e^{-y^2/2} dy\right)} = \\
&= \sqrt{\frac{1}{2\pi} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{(x^2+y^2)(i\omega - 1/2)} dx dy} = \\
&= \sqrt{\frac{1}{2\pi} \int_0^{2\pi} d\varphi \int_0^{+\infty} \rho e^{\rho^2 (i\omega - 1/2)} d\rho} = \\
&= \sqrt{\lim_{A \to +\infty} \int_0^A \frac{e^{\rho^2(i\omega - 1/2)}}{2i\omega - 1} d\left(\rho^2 \left(i\omega - \frac{1}{2}\right)\right)} = \\
&= \sqrt{\lim_{A \to +\infty} \left. \frac{e^{\rho^2(i\omega - 1/2)}}{2i\omega - 1} \right|_0^A} = \sqrt{\lim_{A \to +\infty} \left(\frac{e^{A^2(i\omega - 1/2)}}{2i\omega - 1} - \frac{1}{2i\omega - 1}\right)} = \\
&= \frac{1}{\sqrt{1 - 2i\omega}} = (1 - 2i\omega)^{-1/2}
\end{aligned}$$

on juhusliku suuruse $Y_1 = X^2$ karakteristlik funktsioon. Leiame suuruse Y_n karakteristliku funktsiooni

$$\begin{aligned}
g_{Y_n}(\omega) &= Ee^{i\omega Y_n} = \mathrm{E}e^{i\omega(X_1^2 + X_2^2 + \ldots + X_n^2)} = \mathrm{E}\left(e^{i\omega X_1^2} e^{i\omega X_2^2} \cdots e^{i\omega X_n^2}\right) = \\
&= \left[\begin{array}{c} X_k \ (k = 1; 2; \ldots; n) \text{ on sõltumatud} \Rightarrow \\ \Rightarrow e^{i\omega X_k^2} \quad (k = 1; 2; \ldots; n) \text{ on sõltumatud} \end{array}\right] = \\
&= \left(\mathrm{E}e^{i\omega X_1^2}\right)\left(\mathrm{E}e^{i\omega X_2^2}\right) \cdots \left(\mathrm{E}e^{i\omega X_n^2}\right) = \left(\mathrm{E}e^{i\omega X^2}\right)^n = \\
&= \left((1 - 2i\omega)^{-1/2}\right)^n = (1 - 2i\omega)^{-n/2}.
\end{aligned}$$

Avaldame Fourier' pöördteisenduse abil suuruse Y_n jaotustiheduse

$$f_{Y_n}(y) = \frac{1}{2\pi} \int\limits_{-\infty}^{+\infty} (1-2i\omega)^{-n/2} e^{-i\omega y} d\omega.$$

Saab näidata, et

$$f_{Y_n}(y) = \begin{cases} \dfrac{1}{2^{n/2}\Gamma(n/2)} y^{n/2-1} e^{-y/2}, \text{ kui } y > 0, \\ 0, \text{ kui } y \le 0, \end{cases} \tag{3.8.2}$$

kus

$$\Gamma(\alpha) \overset{\text{def.}}{=} \int_0^{+\infty} t^{\alpha-1} e^{-t} dt \quad (\alpha > 0) \tag{3.8.3}$$

on *gammafunktsioon*. Seejuures

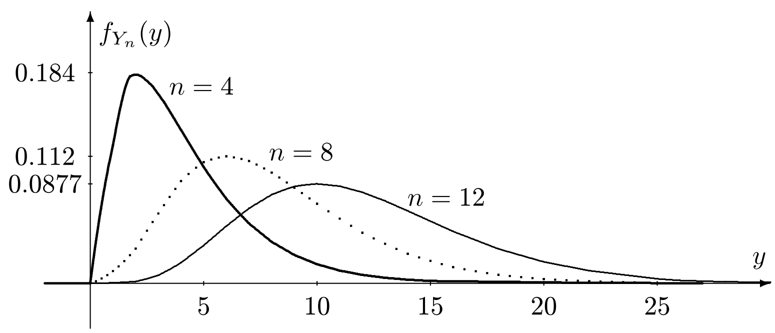
$$\Gamma(n+1) = n! \quad (n \in \mathbf{N}_0),$$
$$\Gamma\left(n + \frac{1}{2}\right) = \frac{1 \cdot 3 \cdots (2n-1)}{2^n} \sqrt{\pi}.$$

Seose

$$\Gamma(\alpha) = \frac{\Gamma(\alpha+1)}{\alpha} \quad (-\alpha \notin \mathbf{N}_0)$$

abil saab laiendada gammafunktsiooni määramispiirkonda.

Skitseerime juhuslike suuruste Y_4, Y_8 ja Y_{12} jaotustiheduste graafikud



Lause 1. Kui juhuslik suurus Y_n allub χ^2-jaotusele vabadusastmete arvuga n, siis

$$g_{Y_n}(\omega) = (1-2i\omega)^{-n/2} \tag{3.8.4}$$

on suuruse Y_n karakteristlik funktsioon ja jaotustihedus on esitatav kujul (3.8.2).

Et

$$\frac{d\, g_{Y_n}(\omega)}{d\omega} = -\frac{n}{2}(1-2i\omega)^{-\frac{n}{2}-1}(-2i) = n\, i\, (1-2i\omega)^{-\frac{n}{2}-1}$$

ja

$$\frac{d^2 g_{Y_n}(\omega)}{d\omega^2} = n \cdot i \left(-\frac{n}{2} - 1\right)(1-2i\omega)^{-\frac{n}{2}-1}(-2i) = -n\,(n+2)\,(1-2i\omega)^{-\frac{n}{2}-1}$$

ning

$$\left.\frac{d\, g_{Y_n}(\omega)}{d\omega}\right|_{\omega=0} = ni, \qquad \left.\frac{d^2 g_{Y_n}(\omega)}{d\omega^2}\right|_{\omega=0} = -n\,(n+2)\,,$$

siis

$$\nu_1 = \mathrm{E}Y_n = \frac{\left.\frac{d\, g_{Y_n}(\omega)}{d\omega}\right|_{\omega=0}}{i} = \frac{n\,i}{i} = n$$

ja

$$\nu_2 = \mathrm{E}Y_n^{\,2} = \frac{\left.\frac{d^2 g_{Y_n}(\omega)}{d\omega^2}\right|_{\omega=0}}{i^2} = \frac{-n\,(n+2)}{-1} = n\,(n+2)$$

ning

$$\mathrm{D}Y_n = \mathrm{E}Y_n^{\,2} - (\mathrm{E}Y_n)^2 = n\,(n+2) - n^2 = 2n.$$

Lause 2. Kui juhuslik suurus Y_n allub χ^2-jaotusele vabadusastmete arvuga n, siis $\mathrm{E}Y_n = n$ ja $\mathrm{D}Y_n = 2n$ ning $\sigma_{Y_n} = \sqrt{2n}$.

Leiame lisaks suuruste Y_1 ja Y_2 jaotustihedused vastavate jaotusfunktsioonide $F_{Y_1}(y)$ ja $F_{Y_2}(y)$ abil. Leiame suuruse $Y_1 = X_1^2$ jaotusfunktsiooni

$$\begin{aligned} F_{Y_1}(y) &= P\,(Y_1 < y) = P\left(X_1^2 < y\right) = P\,(|X_1| < \sqrt{y}) = \\ &= \begin{cases} \dfrac{1}{\sqrt{2\pi}} \int_{-\sqrt{y}}^{\sqrt{y}} e^{-\frac{x_1^2}{2}}\,dx_1 = \dfrac{2}{\sqrt{2\pi}} \int_0^{\sqrt{y}} e^{-\frac{x_1^2}{2}}\,dx_1 & \text{kui } y > 0, \\ 0, \ \text{kui } y \le 0. \end{cases} \end{aligned}$$

Seega on suuruse Y_1 jaotustihedus kujul

$$\begin{aligned} f_{Y_1}(y) = \frac{d}{dy} F_{Y_1}(y) &= \begin{cases} \dfrac{d}{dy} \dfrac{2}{\sqrt{2\pi}} \int_0^{\sqrt{y}} e^{-\frac{x_1^2}{2}}\,dx_1, & \text{kui } y > 0, \\ \dfrac{d}{dy} 0, & \text{kui } y \le 0 \end{cases} = \\ &= \begin{cases} \dfrac{d}{d\sqrt{y}} \left(\dfrac{2}{\sqrt{2\pi}} \int_0^{\sqrt{y}} e^{-\frac{x_1^2}{2}}\,dx_1 \right) \cdot \dfrac{d\sqrt{y}}{dy}, & \text{kui } y > 0, \\ 0, \ \text{kui } y \le 0 \end{cases} = \\ &= \begin{cases} \dfrac{2}{\sqrt{2\pi}} e^{-\frac{y}{2}} \cdot \dfrac{1}{2\sqrt{y}} = \dfrac{1}{2} \dfrac{\sqrt{2}}{\sqrt{\pi}} \dfrac{e^{-\frac{1}{2}y}}{\sqrt{y}}, & \text{kui } y > 0, \\ 0, \ \text{kui } y \le 0. \end{cases} \end{aligned}$$

Leiame suuruse $Y_2 = X_1^2 + X_2^2$ jaotusfunktsiooni

$$F_{Y_2}(y) = P(Y_2 < y) = P(X_1^2 + X_2^2 < y) = P(\sqrt{X_1^2 + X_2^2} < \sqrt{y}) =$$

$$= \begin{cases} \displaystyle\iint\limits_{\sqrt{x_1^2+x_2^2}<\sqrt{y}} \frac{1}{2\pi} e^{-\frac{x_1^2+x_2^2}{2}} dx_1 dx_2, \text{ kui } y > 0, \\ 0, \text{ kui } y \leq 0 \end{cases} =$$

$$= \begin{cases} \displaystyle\frac{1}{2\pi} \int_0^{2\pi} d\varphi \int_0^{\sqrt{y}} \rho e^{-\frac{\rho^2}{2}} d\rho, \text{ kui } y > 0, \\ 0, \text{ kui } y \leq 0 \end{cases} =$$

$$= \begin{cases} \left. -e^{-\frac{\rho^2}{2}} \right|_0^{\sqrt{y}} = 1 - e^{-\frac{y}{2}}, \quad \text{kui } y > 0, \\ 0, \quad \text{kui } y \leq 0 \end{cases}$$

ja jaotustiheduse

$$f_{Y_2}(y) = \frac{d}{dy} F_{Y_2}(y) = \begin{cases} \dfrac{d}{dy}\left(1 - e^{-\frac{y}{2}}\right) = \dfrac{1}{2} e^{-\frac{y}{2}}, \quad \text{kui } y > 0, \\ \dfrac{d}{dy} 0 = 0, \quad \text{kui } y \leq 0. \end{cases}$$

Lause 3. Kui juhuslik suurus Y_n allub χ^2-jaotusele vabadusastmete arvuga n, siis juhusliku suuruse

$$\sqrt{2Y_n} - \sqrt{2n-1}$$

jaotus läheneb asümptootiliselt normaaljaotusele, kusjuures

$$\mathrm{E}\left(\sqrt{2Y_n} - \sqrt{2n-1}\right) \overset{n\to\infty}{\to} 0, \quad \mathrm{D}\left(\sqrt{2Y_n} - \sqrt{2n-1}\right) \overset{n\to\infty}{\to} 1.$$

Märkus 1. Matemaatilises statistikas kasutatav funktsioon $(\alpha, n) \mapsto q_{\alpha,n}$, kus

$$P(Y_n > q_{\alpha,n}) = \int_{q_{\alpha,n}}^{+\infty} f_{Y_n}(y) dy = \alpha,$$

on tabuleeritud Lisas 3 (χ^2-jaotuse, mille vabadusastmete arv on n, *täiendkvantiilid*).

Näide 1. Leiame tõenäosuse $P(Y_5 > \mathrm{E}Y_5)$.

Kuna Lause 2 põhjal $\mathrm{E}Y_5 = 5$, siis tuleb leida tõenäosus

$$P(Y_5 > 5).$$

Leiame tõenäosuse $P(Y_5 \in (-\infty; 5])$. Kasutame paketist SWP jaotusfunktsiooni

$$\text{ChiSquareDist}(y, n) \equiv F_{Y_n}(y) = \begin{cases} \dfrac{1}{2^{n/2}\Gamma(n/2)} \int\limits_0^y t^{n/2-1} e^{-t/2} dt \quad (t > 0), \\ 0 \quad (t \leq 0). \end{cases}$$

Saame

$$P(Y_5 \in [0;5]) \overset{\text{miks?}}{=} P(Y_5 \in (-\infty;5]) = \text{ChiSquareDist}(5,5) = 0.\,584\,12\,.$$

Seega

$$P(Y_5 > 5) = 1 - P(Y_5 \in [0;5]) = 0.415\,88.$$

Millise tulemuseni jõuame Lisas 3 esitatud χ^2-jaotuse täiendkvantiilide tabeli abil? $\diamondsuit$

Näide 2. Leiame Lisas 3 esitatud tabeli abil juhusliku suuruse Y_8 sellised võimalikud väärtused y_1 ja y_2, et

$$P(y_1 \le Y_8 \le y_2) = 0.8, \quad P(Y_8 \le y_1) = P(Y_8 \ge y_2)\,.$$

Kuna Y_8 on pidev juhuslik suurus, siis

$$P(Y_8 \le y_1) = P(Y_8 \ge y_2) = (1 - P(y_1 \le Y_8 \le y_2))/2 = 0.1,$$

$$P(Y_8 \ge y_2) = P(Y_8 > y_2) = \int_{y_2}^{+\infty} f_{Y_8}(y)dy = 0.1 \overset{\text{Lisa 3}}{\Rightarrow} y_2 \approx 13.4.$$

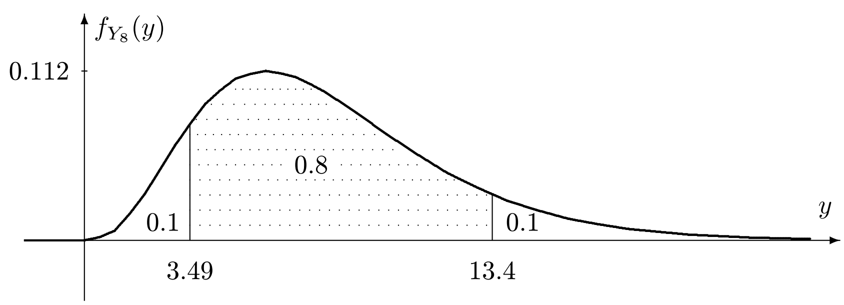
ja

$$P(Y_8 \le y_1) = 1 - P(Y_8 > y_1) = 0.1 \Rightarrow P(Y_8 > y_1) = 0.9 \overset{\text{Lisa 3}}{\Rightarrow} y_1 \approx 3.49.$$

Seega

$$P(3.49 \le Y_8 \le 13.4) \approx 0.8, \quad P(Y_8 \le 3.49) \approx P(Y_8 \ge 13.4)\,.$$

Teeme joonise



$\diamondsuit$

3.9 Studenti jaotus

Definitsioon 1. Öeldakse, et juhuslik suurus

$$T_n = \frac{X}{\sqrt{Y_n}}\sqrt{n} \tag{3.9.1}$$

allub *Studenti jaotusele* ehk *t-jaotusele* vabadusastmete arvuga n, kui $X \sim N(0;1)$ ja Y_n on sõltumatud juhuslikud suurused ning Y_n on χ^2-jaotusega vabadusastmete arvuga n.

Seega

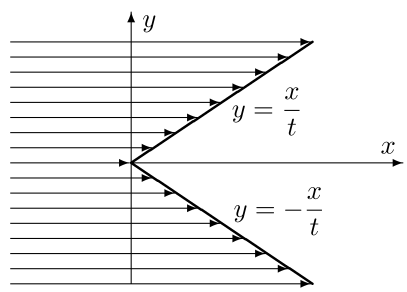
$$T_n = \frac{X}{\sqrt{X_1^2 + X_2^2 + \ldots + X_n^2}}\sqrt{n}, \tag{3.9.2}$$

kus suurused $X \sim N(0;1)$ ja $X_k \sim N(0;1)$ $(k = 1, \ldots, n)$ on sõltumatud juhuslikud suurused.

Vaatleme juhtu $n = 1$. Kasutame tähistust $X_1 \equiv Y$. Leiame, et

$$F_{T_1}(t) = P(T_1 < t) = P\left(\frac{X}{\sqrt{X_1^2}} < t\right) = \\ = P\left(\frac{X}{\sqrt{Y^2}} < t\right) = P\left(\frac{X}{|Y|} < t\right).$$

Joon $|y| = \dfrac{x}{t}$ jagab xy-tasandi kahte ossa. Uurime kaht juhtu $t > 0$ ja $t < 0$. Kui $t > 0$, siis saame piirkonna rajajoone, mis koosneb kahest kiirest, ja võrratust $|y| > \dfrac{x}{t}$ rahuldava piirkonna (viirutatu)



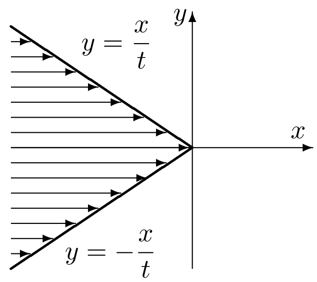
Saame

$$F_{T_1}(t) = P\left(|Y| > \frac{X}{t}\right) = \frac{1}{2\pi}\iint_{|y| > \frac{x}{t}} e^{-\frac{x^2+y^2}{2}}dxdy = \\ = \frac{1}{2\pi}\int\limits_{\arctan(1/t)}^{2\pi - \arctan(1/t)} d\varphi \int\limits_0^{+\infty} \rho e^{-\frac{\rho^2}{2}}d\rho = \\ = \frac{1}{2\pi}\left(2\pi - 2\arctan\frac{1}{t}\right) = 1 - \frac{1}{\pi}\arctan\frac{1}{t}$$

ja

$$f_{T_1}(t) = \frac{d}{dt}\left(1 - \frac{1}{\pi}\arctan\frac{1}{t}\right) = -\frac{1}{\pi}\frac{1}{1+\left(\frac{1}{t}\right)^2}\left(-\frac{1}{t^2}\right) = \frac{1}{\pi}\frac{1}{1+t^2}.$$

Kui $t < 0$, siis saame piirkonna rajajoone, mis koosneb kahest kiirest, ja võrratust $|y| < \dfrac{x}{t}$ rahuldava piirkonna (viirutatu)



Saame

$$F_{T_1}(t) = P\left(|Y| < \frac{X}{t}\right) = \frac{1}{2\pi} \iint\limits_{|y| < \frac{x}{t}} e^{-\frac{x^2+y^2}{2}} dxdy =$$

$$= \frac{1}{2\pi} \int\limits_{\pi+\arctan(1/t)}^{\pi-\arctan(1/t)} d\varphi \int_0^{+\infty} \rho e^{-\frac{\rho^2}{2}} d\rho = \frac{1}{2\pi}\left(-2\arctan\frac{1}{t}\right) = -\frac{1}{\pi}\arctan\frac{1}{t}$$

ja

$$f_{T_1}(t) = \frac{d}{dt}\left(-\frac{1}{\pi}\arctan\frac{1}{t}\right) = -\frac{1}{\pi}\frac{1}{1+\left(\frac{1}{t}\right)^2}\left(-\frac{1}{t^2}\right) = \frac{1}{\pi}\frac{1}{1+t^2}.$$

Seega

$$f_{T_1}(t) = \frac{1}{\pi}\frac{1}{1+t^2} \quad (t \neq 0).$$

Kuna T_1 on pidev juhuslik suurus, siis

$$f_{T_1}(t) = \frac{1}{\pi}\frac{1}{1+t^2} \quad (t \in \mathbf{R}).$$

Seega allub suurus T_1 Cauchy jaotusele. Saab näidata, et kehtib järgnev väide.

Lause 1. Kui juhuslik suurus T_n allub Studenti jaotusele vabadusastmete arvuga n, siis

$$f_{T_n}(t) = \frac{\Gamma\left((n+1)/2\right)}{\sqrt{\pi n}\Gamma\left((n/2)\right)}\left(1+\frac{t^2}{n}\right)^{-(n+1)/2} \quad (t \in \mathbf{R},\ n = 1; 2; \ldots) \tag{3.9.3}$$

ja

$$\mathrm{E}T_n = 0 \quad (n \geq 2) \tag{3.9.4}$$

ning

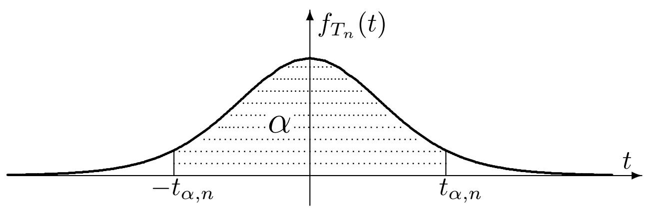
$$\mathrm{D}T_n = \frac{n}{n-2} \quad (n \geq 3). \tag{3.9.5}$$

Jaotustihedus $f_{T_n}(t)$ on paarisfunktsioon. Studenti jaotus läheneb asümptootiliselt normaaljaotusele.

Märkus 1. Matemaatilises statistikas kasutatav funktsioon (vt [4], [24], [31]) $(\alpha, n) \mapsto t_{\alpha,n}$, kus

$$P(|T_n| \le t_{\alpha,n}) = P(-t_{\alpha,n} \le T_n \le t_{\alpha,n}) = \int\limits_{-t_{\alpha,n}}^{t_{\alpha,n}} f_{T_n}(t)dt = 2\int\limits_{0}^{t_{\alpha,n}} f_{T_n}(t)dt = \alpha, \tag{3.9.6}$$

st



on tabuleeritud Lisas 4. Matemaatilises statistikas kasutatakse sageli (vt [19], [27]) hüpoteeside kontrollimisel kriitiliste punktide määramiseks nn täiendkvantiile, funktsiooni $(\alpha, n) \mapsto q_{\alpha,n}$, kus

$$P(|T_n| > q_{\alpha,n}) = \int\limits_{-\infty}^{-q_{\alpha,n}} f_{T_n}(t)dt + \int\limits_{q_{\alpha,n}}^{+\infty} f_{T_n}(t)dt = 2\int\limits_{q_{\alpha,n}}^{+\infty} f_{T_n}(t) = \alpha. \tag{3.9.7}$$

Seega $t_{\alpha,n} = q_{1-\alpha,n}$.

Näide 1. Leiame seosest

$$P(-a < T_5 < a) = 0.8$$

suuruse a Lisas 4 esitatud t-jaotuse tabeli abil ja SWP abil.

1° Kuna T_n on pidev juhuslik suurus, siis $P(T_5 = -a) = P(T_5 = a) = 0$ ja piisab uurida juhtu

$$P(|T_5| \le a) = 0.8.$$

Kasutame seoseid (3.9.6), täpsemalt Lisa 4 t-jaotuse funktsiooni $(\alpha, n) \mapsto t_{\alpha,n}$ tabelit. Saame $(0.8; 5) \mapsto t_{0.8\,;\,5} \approx 1.476 = a$.

2° Kasutame suuruse T_n korral paketist SWP jaotusfunktsiooni

$$\mathrm{TDist}(t,n) \equiv F_{T_n}(t) = \frac{\Gamma((n+1)/2)}{\sqrt{\pi n}\Gamma(n/2)} \int\limits_{-\infty}^{t} \left(1 + \frac{u^2}{n}\right)^{-(n+1)/2} du = p$$

pöördfunktsiooni $\mathrm{TInv}(p, n)$. Kuna suuruse T_n jaotustihedus $f_{T_n}(t)$ on paarisfunktsioon, siis

$$P(-\infty < T_5 \le -a) = P(a \le T_5 < \infty) = (1 - P(|T_5| \le a))/2$$

ja

$$\begin{aligned}P(-\infty < T_5 < a) &= P((-\infty < T_5 \le -a) + (-a < T_5 < a)) = \\ &= P(-\infty < T_5 \le -a) + P(-a < T_5 < a) = \\ &= (1 - 0.8)/2 + 0.8 = 0.9\end{aligned}$$

ning $a = \mathrm{TInv}(0.9, 5) = 1.4759$. $\diamondsuit$

3.10 Fisheri jaotus

Definitsioon 1. Öeldakse, et juhuslik suurus

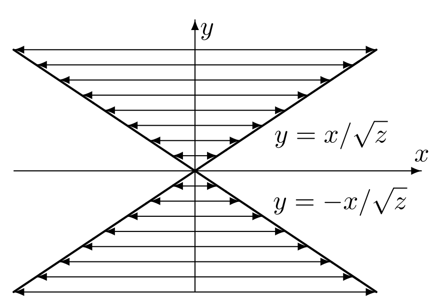
$$Z_{n,m} = \frac{X_n/n}{Y_m/m} \tag{3.10.1}$$

allub *Fisheri jaotusele* (*F-jaotusele, Snedecori jaotusele*) vabadusastmetega n ja m, kui sõltumatud juhuslikud suurused X_n ja Y_n alluvad χ^2-jaotusele vastavalt vabadusastmete arvudega n ja m.

Uurime erijuhtu $n = 1$ ja $m = 1$. Leiame, et

$$\begin{aligned}F_{Z_{1,1}}(z) &= P(Z_{1,1} < z) = P\left(\frac{X_1/1}{Y_1/1} < z\right) = \left[\begin{array}{c}\text{tähistame}\\ X_1 = X^2,\ Y_1 = Y^2\end{array}\right] = \\ &= P\left(\frac{X^2}{Y^2} < z\right) = P\left(\frac{|X|}{|Y|} < \sqrt{z}\right) = P\left(|Y| > \frac{|X|}{\sqrt{z}}\right),\end{aligned}$$

kus $X \sim N(0;1)$ ja $Y \sim N(0;1)$ on sõltumatud juhuslikud suurused. Kujutame xy-tasandil tingimust $|y| > |x|/\sqrt{z}$ rahuldava piirkonna



Saame $z > 0$ korral

$$F_{Z_{1,1}}(z) = \frac{1}{2\pi} \iint\limits_{|y|>\left(|x|/\sqrt{z}\right)} e^{-\left(x^2+y^2\right)/2} dxdy =$$

$$= \frac{4}{2\pi} \int\limits_{\arctan\left(1/\sqrt{z}\right)}^{\pi/2} d\varphi \int\limits_0^{+\infty} \rho e^{-\rho^2/2} d\rho =$$

$$= \frac{2}{\pi}\left(\pi/2 - \arctan\left(1/\sqrt{z}\right)\right) \lim_{A\to+\infty}\left(1 - e^{-A^2/2}\right) =$$

$$= \frac{2}{\pi}\left(\pi/2 - \arctan\left(1/\sqrt{z}\right)\right) = 1 - \frac{2}{\pi}\arctan\left(1/\sqrt{z}\right).$$

Seega

$$f_{Z_{1,1}}(z) = \frac{d}{dz}F_{Z_{1,1}}(z) = \begin{cases} \dfrac{d}{dz}\left[\dfrac{2}{\pi}\left(\pi/2 - \arctan\left(1/\sqrt{z}\right)\right)\right], \text{ kui } z > 0, \\ 0, \text{ kui } z \le 0 \end{cases} =$$

$$= \begin{cases} 1/\left(\pi\sqrt{z}\,(z+1)\right), \text{ kui } z > 0 \\ 0, \text{ kui } z \le 0. \end{cases}$$

Osutub, et

$$f_{Z_{n,m}}(z) = \begin{cases} \dfrac{\Gamma\left((n+m)/2\right)}{\Gamma(n/2)\,\Gamma(m/2)}\left(\dfrac{n}{m}\right)^{n/2} z^{(n-2)/2}\left(1 + \dfrac{n}{m}z\right)^{-(n+m)/2}, \text{ kui } z > 0 \\ 0, \text{ kui } z \le 0 \end{cases} \tag{3.10.2}$$

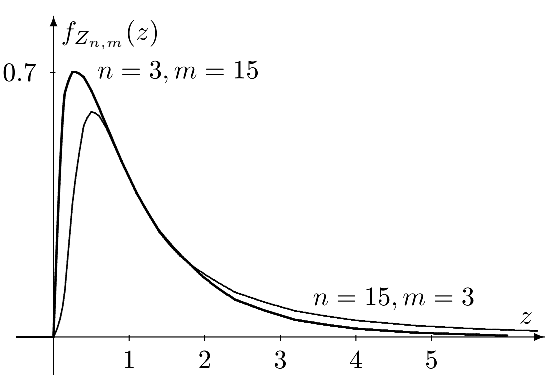
ja

$$\mathrm{E}Z_{n,m} = \frac{m}{m-2} \quad (m > 2) \tag{3.10.3}$$

ning

$$\mathrm{D}Z_{n,m} = \frac{2m^2(n+m-2)}{n(m-2)^2(m-4)} \quad (m > 4). \tag{3.10.4}$$

Skitseerime suuruse $Z_{3;15}$ ja $Z_{15;3}$ jaotustihedused $f_{Z_{3;15}}(z)$ ja $f_{Z_{15;3}}(z)$ vastavalt jämeda ja peenikese joonega



Näide 1. Leiame tõenäosuse

$$P\left(|Z_{3;15} - \mathrm{E}Z_{3;15}| < \sigma_{Z_{3,15}}\right).$$

Kuna

$$\mathrm{E}Z_{3;15} \overset{(3.10.3)}{=} \frac{15}{15-2} = \frac{15}{13} \approx \quad 1.154,$$

$$\mathrm{D}Z_{3;15} \overset{(3.10.4)}{=} \frac{2\,(15)^2\,(3+15-2)}{3\,(15-2)^2\,(15-4)} = \frac{2400}{1859} \approx 1.291$$

ja

$$\sigma_{Z_{3;15}} = \sqrt{DZ_{3;15}} \approx \sqrt{1.291} \approx 1.136\,,$$

siis

$$P\left(|Z_{3;15} - \mathrm{E}Z_{3;15}| < \sigma_{Z_{3;15}}\right) \approx P\left(|Z_{3;15} - 1.154| < 1.136\right) =$$

$$P\,(0.0\,18 < Z_{3;15} < 2.29) = \mathrm{FDist}\,(2.29; 3, 15) - \mathrm{FDist}\,(0.0\,18; 3, 15) \approx 0.877,$$

kus FDist $(z; n, m)$ on suuruse $Z_{n,m}$ jaotusfunktsioon paketis SWP. ♢

Lisas 5 (F-jaotuse täiendkvantiilid) on $\alpha = 0.05$ ja $\alpha = 0.01$ korral tabuleeritud funktsioon $(\alpha, n, m) \mapsto z_{\alpha,n,m}$, kus

$$P\,(Z_{n,m} > z_{\alpha,n,m}) = \int_{z_{\alpha,n,m}}^{+\infty} f_{Z_{n,m}}(z)dz = \alpha.$$

Kehtib seos

$$z_{\alpha,n,m} = \frac{1}{z_{1-\alpha,m,n}} \tag{3.10.5}$$

Näide 2. Leiame Lisas 5 esitatud tabeli abil juhusliku suuruse $Z_{15;3}$ sellised võimalikud väärtused z_1 ja z_2, et

$$P\,(z_1 \le Z_{15;3} \le z_2) = 0.9, \quad P\,(Z_{15;3} \le z_1) = P\,(Z_{15;3} \ge z_2)\,.$$

Kuna $Z_{15;3}$ on pidev juhuslik suurus, siis

$$P\,(Z_{15;3} \le z_1) = P\,(Z_{15;3} \ge z_2) = (1 - P\,(z_1 \le Y_8 \le z_2))\,/2 = 0.05,$$

$$P\,(Z_{15;3} \ge z_2) = P\,(Z_{15;3} > z_2) = \int_{z_2}^{+\infty} f_{Z_{15;3}}(z)dz = 0.05 \overset{\text{Lisa 5}}{\Rightarrow} z_2 \approx 8.70.$$

ja

$$z_1 = z_{0.95;15;3} \overset{(3.10.5)}{=} \frac{1}{z_{0.05;3;15}} \overset{\text{Lisa 5}}{=} \frac{1}{3.29} \approx 0.304.$$

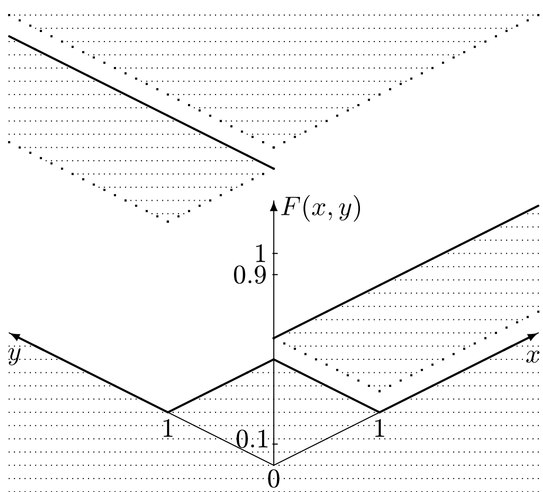
Seega

$$P\,(0.304 \le Z_{15;3} \le 8.70) \approx 0.9, \quad P\,(Z_{15;3} \le 0.304) \approx P\,(Z_{15;3} \ge 8.70)\,. \qquad \diamondsuit$$

3.11 Ülesanded

1. Poiss ostab ühe loteriipileti. Võidu tõenäosus on 0.1. Olgu X poisi võitnud piletite arv ja Y tema võiduta piletite arv. Leidke vektori (X,Y) jaotusseadus, jaotusfunktsioon $F(x,y)$ nii analüütiliselt kui ka graafiliselt ja jaotustihedus $f(x,y)$. Leidke EX, EY, DX, DY, σ_X, σ_Y, vektori (X,Y) kovariatsioonimoment $\operatorname{cov}(X,Y)$, korrelatsioonikordaja $\mathrm{r}(X,Y)$, kovariatsioonimaatriks K ja korrelatsioonimaatriks R. Kas selle vektori komponendid on sõltuvad? Korreleeruvad?

V: $F(x,y)=0.1\cdot\mathbf{1}(x-1)\,\mathbf{1}(y)+0.9\cdot\mathbf{1}(x)\,\mathbf{1}(y-1)$, $f(x,y)=0.1\cdot\delta(x-1)\,\delta(y)+$ $+0.9\cdot\delta(x)\,\delta(y-1)$, E$X=0.1$, E$Y=0.9$, D$X=$ D$Y=0.09$, $\sigma_X=\sigma_Y=0.3$,



$x_i \backslash y_j$	0	1
0	0	0.9
1	0.1	0

$\operatorname{cov}(X,Y)=-0.09$,

$\mathrm{r}(X,Y)=-1$,

$\mathrm{K}=\begin{pmatrix}0.09 & -0.09\\ & 0.09\end{pmatrix}$,

$\mathrm{R}=\begin{pmatrix}1 & -1\\ & 1\end{pmatrix}$,

sõltuvad, korreleeruvad.

2. Kaks poissi sooritavad mõlemad 2 vabaviset. Esimesel poisil on mõlemal viskel tabamise tõenäosus 0.6 ja teisel poisil 0.7. Olgu X esimese poisi tabamuste koguarv ja Y teise poisi tabamuste koguarv. Leidke vektori (X,Y) jaotusseadus, $F(x,y)$, $f(x,y)$, EX, EY, DX, DY, σ_X, σ_Y, $\operatorname{cov}(X,Y)$, $\mathrm{r}(X,Y)$, K ja R. Kas selle vektori komponendid on sõltuvad? Korreleeruvad?

V: $F(x,y)=0.0144\cdot\mathbf{1}(x)\,\mathbf{1}(y)+0.0672\cdot\mathbf{1}(x)\,\mathbf{1}(y-1)+0.0784\cdot\mathbf{1}(x)\,\mathbf{1}(y-2)+$ $+0.0432\cdot\mathbf{1}(x-1)\,\mathbf{1}(y)+0.2016\cdot\mathbf{1}(x-1)\,\mathbf{1}(y-1)+0.2352\cdot\mathbf{1}(x-1)\,\mathbf{1}(y-2)+$ $+0.0324\cdot\mathbf{1}(x-2)\,\mathbf{1}(y)+0.1512\cdot\mathbf{1}(x-2)\,\mathbf{1}(y-1)+0.1764\cdot\mathbf{1}(x-2)\,\mathbf{1}(y-2)$,
$f(x,y)=0.0144\cdot\delta(x)\,\delta(y)+0.0672\cdot\delta(x)\,\delta(y-1)+0.0784\cdot\delta(x)\,\delta(y-2)+$ $+0.0432\cdot\delta(x-1)\,\delta(y)+0.2016\cdot\delta(x-1)\,\delta(y-1)+0.2352\cdot\delta(x-1)\delta(y-2)+$ $+0.0324\cdot\delta(x-2)\,\mathbf{1}(y)+0.1512\cdot\delta(x-2)\,\delta(y-1)+0.1764\cdot\delta(x-2)\,\delta(y-2)$,

$x_i \backslash y_j$	0	1	2
0	0.0144	0.0672	0.0784
1	0.0432	0.2016	0.2352
2	0.0324	0.1512	0.1764

E$X=1.2$, E$Y=1.4$, D$X=0.48$,
D$Y=0.42$, $\sigma_X\approx 0.693$, $\sigma_Y\approx 0.648$,
$\operatorname{cov}(X,Y)=0$, $\mathrm{r}(X,Y)=0$,
$\mathrm{K}=\begin{pmatrix}0.48 & 0\\ & 0.42\end{pmatrix}$, $\mathrm{R}=\begin{pmatrix}1 & 0\\ & 1\end{pmatrix}$,
sõltumatud, mittekorreleeruvad.

3. Märklaua suunas sooritatakse kaks lasku. Mõlemal lasul on tabamise tõenäosus p. Olgu X tabamuste arv ja Y möödalaskude arv. Leidke vektori (X,Y) jaotusseadus, $F(x,y)$, $f(x,y)$, EX, EY, DX, DY, σ_X, σ_Y, $\mathrm{cov}\,(X,Y)$, $\mathrm{r}\,(X,Y)$, K ja R. Kas selle vektori komponendid on sõltuvad? Korreleeruvad?
V: $F(x,y) = p^2 \cdot \mathbf{1}(x-2)\,\mathbf{1}(y) + 2pq \cdot \mathbf{1}(x-1)\,\mathbf{1}(y-1) + q^2 \cdot \mathbf{1}(x)\,\mathbf{1}(y-2)$,
$f(x,y) = p^2 \cdot \delta(x-2)\,\delta(y) + 2pq \cdot \delta(x-1)\,\delta(y-1) + q^2 \cdot \delta(x)\,\delta(y-2)$,

$x_i \setminus y_j$	0	1	2
0	0	0	q^2
1	0	$2pq$	0
2	p^2	0	0

$$\mathrm{E}X = 2p,\ \ \mathrm{E}Y = 2q,\ \mathrm{D}X = \mathrm{D}Y = 2pq,$$
$$\sigma_X = \sigma_Y = \sqrt{2pq},$$
$$\mathrm{cov}\,(X,Y) = -2pq, \qquad \mathrm{r}\,(X,Y) = -1,$$
$$\mathrm{K} = \begin{pmatrix} 2pq & -2pq \\ & 2pq \end{pmatrix}, \quad \mathrm{R} = \begin{pmatrix} 1 & -1 \\ & 1 \end{pmatrix},$$
sõltuvad, korreleeruvad.

4. Diskreetne juhuslik vektor (X,Y) on antud jaotusseadusega

$x_i \setminus y_j$	-1	2	3
0	0.2	0.1	0.1
1	0.1	0.3	0.2

Leidke $F(x,y)$, $f(x,y)$, EX, EY, DX, DY, σ_X, σ_Y, $\mathrm{cov}\,(X,Y)$, $\mathrm{r}\,(X,Y)$ ja R. Kas selle vektori komponendid on korreleeruvad? Leidke regressioonijooned $y = \mathrm{E}\,(Y\,|x)$ ja $x = \mathrm{E}\,(X\,|y)$.
V: $F(x,y) = 0.2 \cdot \mathbf{1}(x)\,\mathbf{1}(y+1) + 0.1 \cdot \mathbf{1}(x)\,\mathbf{1}(y-2) + 0.1 \cdot \mathbf{1}(x)\,\mathbf{1}(y-3) +$
$+ 0.1 \cdot \mathbf{1}(x-1)\,\mathbf{1}(y+1) + 0.3 \cdot \mathbf{1}(x-1)\,\mathbf{1}(y-2) + 0.2 \cdot \mathbf{1}(x-1)\,\mathbf{1}(y-3)$,
$f(x,y) = 0.2 \cdot \delta(x)\delta(y+1) + 0.1 \cdot \delta(x)\,\delta(y-2) + 0.1 \cdot \delta(x)\,\delta(y-3) +$
$+ 0.1 \cdot \delta(x-1)\,\delta(y+1) + 0.3 \cdot \delta(x-1)\,\delta(y-2) + 0.2 \cdot \delta(x-1)\,\delta(y-3)$,
$\mathrm{E}X = 0.6$, $\mathrm{E}Y = 1.4$, $\mathrm{D}X = 0.24$, $\mathrm{D}Y = 2.64$, $\sigma_X = \sqrt{0.24} \approx 0.490$,
$\sigma_Y = \sqrt{2.64} \approx 1.625$, $\mathrm{cov}\,(X,Y) = 0.26$, $\mathrm{r}\,(X,Y) \approx 0.327$, korreleeruvad,
$$\mathrm{R} \approx \begin{pmatrix} 1 & 0.327 \\ & 1 \end{pmatrix}, \ y = \begin{cases} 3/4, & x = 0, \\ 11/6, & x = 1, \end{cases} \quad x = \begin{cases} 1/3, & y = -1, \\ 3/4, & y = 2, \\ 2/3, & y = 3. \end{cases}$$

5. Sooritatakse 2 katset. Mõlemal on sündmuse A toimumise tõenäosus 0.6. Leidke juhusliku vektori (X,Y), kus X on sündmuse A toimumiste arv ja Y sündmuse A toimumiste arvu ja A mittetoimumiste arvu vahe, jaotusseadus, EX, EY, DX, DY, σ_X, σ_Y, $\mathrm{cov}\,(X,Y)$, $\mathrm{r}\,(X,Y)$, K ja R. Kas selle vektori komponendid on sõltuvad? Korreleeruvad? V:

$x_i \setminus y_j$	-2	0	2
0	0.16	0	0
1	0	0.48	0
2	0	0	0.36

$$\mathrm{E}X = 1.2,\ \mathrm{E}Y = 0.4,\ \mathrm{D}X = 0.48,$$
$$\mathrm{D}Y = 1.92,\ \sigma_X \approx 0.693,\ \sigma_Y \approx 1.386,$$
$$\mathrm{cov}\,(X,Y) = 0.96, \qquad \mathrm{r}\,(X,Y) = 1,$$
$$\mathrm{K} = \begin{pmatrix} 0.48 & 0.96 \\ & 1.92 \end{pmatrix}, \quad \mathrm{R} = \begin{pmatrix} 1 & 1 \\ & 1 \end{pmatrix},$$
sõltuvad, korreleeruvad.

6. On antud diskreetse juhusliku suuruse X jaotusseadus

x_k	-1	0	1	2
p_k	0.1	0.3	0.4	0.2

ja $Y = \left[2^X\right]$, kus $\left[2^X\right]$ on suuruse 2^X täisosa, ning $Z = X^2$. Leidke Y ja Z jaotusseadused, EX, EY, EZ, DX, DY, EZ, σ_X, σ_Y, ja σ_Z. Leidke (X, Y, Z) korral K ja R. V:

y_k	0	1	2	4
p_k	0.1	0.3	0.4	0.2

,

z_k	0	1	4
p_k	0.3	0.5	0.2

,

E$X = 0.7$, E$Y = 1.49$, E$Z = 1.3$, D$X = 0.81$, D$Y = 1.49$, D$Z = 2.01$, $\sigma_X = 0.9$, $\sigma_Y \approx 1.221$ $\sigma_Z \approx 1.418$,

$$\mathrm{K} = \begin{pmatrix} 0.81 & 1.07 & 0.99 \\ & 1.49 & 1.53 \\ & & 2.01 \end{pmatrix}, \ \mathrm{R} \approx \begin{pmatrix} 1 & 0.974 & 0.776 \\ & 1 & 0.884 \\ & & 1 \end{pmatrix}.$$

7. Olgu (X, Y) ühtlase jaotusega ristkülikus, mis on määratud võrratustega $-1 \le y \le 2$ ja $-3 \le x \le 3$. Leidke $f(x,y)$, $f_1(x)$, $f_2(y)$, $f_1(x|y)$, $f_2(y|x)$ ja $\mathrm{cov}\,(X, Y)$. Kas selle vektori komponendid on sõltuvad? Korreleeruvad? Leidke regressioonijooned $y = \mathrm{E}\,(Y \mid x)$ ja $x = \mathrm{E}\,(X \mid y)$.

V: $f(x,y) = \begin{cases} 1/18, \ (x,y) \in [-1;2] \times [-3;3], \\ 0, \ (x,y) \notin [-1;2] \times [-3;3], \end{cases}$ $\quad f_1(x) = \begin{cases} 1/3, \ x \in [-1;2], \\ 0, \ x \notin [-1;2], \end{cases}$

$f_2(y) = \begin{cases} 1/6, \ y \in [-3;3], \\ 0, \ y \notin [-3;3], \end{cases}$ $\quad f_1(x|y) = f_1(x)$, $f_2(y|x) = f_2(y)$, $\mathrm{cov}\,(X, Y) = 0$,

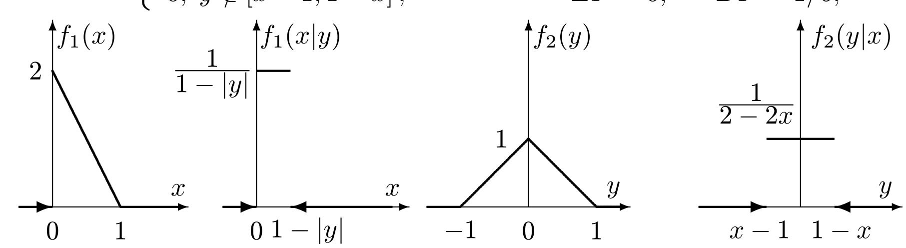
sõltumatud, mittekorreleeruvad, $y = 0 \ (x \in [-1;2])$, $x = 0 \ (y \in [-3;3])$.

8. Vektor (X, Y) allub ühtlasele jaotusele võrratustega $|x| + |y| \le 1$ ja $x \ge 0$ määratud kolmnurgas. Leidke $f(x,y)$, $f_1(x)$, $f_2(y)$, $f_1(x|y)$, $f_2(y|x)$ EX, EY, DX, DY, $\mathrm{cov}\,(X, Y)$, K ja R. Skitseerige funktsioonide $f_1(x)$, $f_2(y)$, $f_1(x|y)$ ja $f_2(y|x)$ graafikud. Kas selle vektori komponendid on sõltuvad? Korreleeruvad? Leidke regressioonijooned $y = \mathrm{E}\,(Y \mid x)$ ja $x = \mathrm{E}\,(X \mid y)$.

V: $f(x,y) = \begin{cases} 1, \ |x| + |y| \le 1 \ \wedge \ x \ge 0, \\ 0, \ |x| + |y| > 1 \ \vee \ x < 0, \end{cases}$ $\quad f_1(x) = \begin{cases} 2 - 2x, \ x \in [0;1], \\ 0, \ x \notin [0;1], \end{cases}$

$f_2(y) = \begin{cases} 1 - |y|, \ y \in [-1;1], \\ 0, \ y \notin [-1;1], \end{cases}$ $\quad f_1(x \mid y) = \begin{cases} 1/(1 - |y|), \ x \in [0; 1 - |y|], \\ 0, \ x \notin [0; 1 - |y|], \end{cases}$

$f_2(y \mid x) = \begin{cases} 1/(2 - 2x), \ y \in [x - 1, 1 - x], \\ 0, \ y \notin [x - 1, 1 - x], \end{cases}$ $\quad$ E$X = 1/3$, D$X = 1/18$, E$Y = 0$, D$Y = 1/6$,



$\mathrm{cov}\,(X, Y) = 0$, $K = \begin{pmatrix} 1/18 & 0 \\ & 1/6 \end{pmatrix}$, $R = \begin{pmatrix} 1 & 0 \\ & 1 \end{pmatrix}$, sõltuvad, mittekorreleeruvad.

$y = 0 \ (0 \le x \le 1)$, $x = (1 - |y|)/2 \ (-1 \le y \le 1)$.

9. Vektor (X, Y) allub ühtlasele jaotusele võrratustega $|x| + |y| \le 1$ ja $y \ge 0$

määratud kolmnurgas. Leidke $f(x,y)$, $f_1(x)$, $f_2(y)$, $f_1(x|y)$, $f_2(y|x)$ EX, EY, DX, DY, cov(X,Y), K ja R. Kas selle vektori komponendid on sõltuvad? Korreleeruvad? Leidke regressioonijooned $y = \mathrm{E}(Y \mid x)$ ja $x = \mathrm{E}(X \mid y)$

V: $f(x,y) = \begin{cases} 1, \ |x| + |y| \le 1 \ \wedge \ y \ge 0, \\ 0, \ |x| + |y| > 1 \ \vee \ y < 0, \end{cases}$ $f_1(x) = \begin{cases} 1 - |x|, \ x \in [-1;1], \\ 0, \ x \notin [-1;1], \end{cases}$

$f_2(y) = \begin{cases} 2 - 2y, \ y \in [0;1], \\ 0, \ y \notin [0;1], \end{cases}$ $f_1(x \mid y) = \begin{cases} 1/(2-2y), \ x \in [y-1, 1-y], \\ 0, \ x \notin [y-1, 1-y], \end{cases}$

$f_2(y \mid x) = \begin{cases} 1/(1-|x|), \ y \in [0; 1-|x|], \\ 0, \ y \notin [0; 1-|x|], \end{cases}$ $\mathrm{E}X = 0,\quad \mathrm{D}X = 1/6,$ $\mathrm{E}Y = 1/3,\quad \mathrm{D}Y = 1/18,$

$\mathrm{cov}(X,Y) = 0$, $\mathrm{K} = \begin{pmatrix} 1/6 & 0 \\ & 1/18 \end{pmatrix}$, $\mathrm{R} = \begin{pmatrix} 1 & 0 \\ & 1 \end{pmatrix}$, sõltuvad, mittekorreleeruvad.

$y = (1-|x|)/2 \ (-1 \le x \le 1), \ \ x = 0 \ (0 \le y \le 1).$

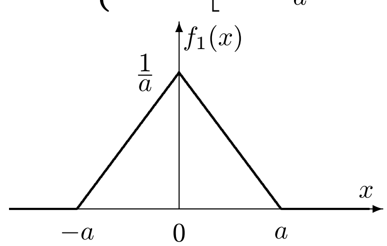
10. Vektor (X,Y) allub ühtlasele jaotusele rombis $|x|/a + |y|/b \le 1$. Leidke $f(x,y)$, $f_1(x)$, $f_2(y)$, $f_1(x|y)$, $f_2(y|x)$ EX, EY, DX, DY, cov(X,Y), K ja R. Skitseerige funktsioonide $f_1(x)$ ja $f_1(x|y)$ graafikud. Kas selle vektori komponendid on sõltuvad? Korreleeruvad? Leidke regressioonijooned $y = \mathrm{E}(Y \mid x)$ ja $x = \mathrm{E}(X \mid y)$.

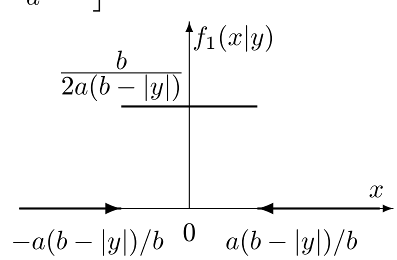
V: $f(x,y) = \begin{cases} \dfrac{1}{2ab}, \ \dfrac{|x|}{a} + \dfrac{|y|}{b} \le 1, \\ 0, \ \dfrac{|x|}{a} + \dfrac{|y|}{b} > 1, \end{cases}$ $f_1(x) = \begin{cases} \dfrac{a-|x|}{a^2}, \ x \in [-a;a], \\ 0, \ x \notin [-a;a], \end{cases}$

$f_2(y) = \begin{cases} \dfrac{b-|y|}{b^2}, \ y \in [-b;b], \\ 0, \ y \notin [-b;b], \end{cases}$ $\mathrm{E}X = 0,\quad \mathrm{D}X = a^2/6,$ $\mathrm{E}Y = 0,\quad \mathrm{D}Y = b^2/6,$

$$f_1(x \mid y) = \begin{cases} \dfrac{b}{2a(b-|y|)}, \ x \in \left[-\dfrac{a(b-|y|)}{b}, \dfrac{a(b-|y|)}{b}\right], \\ 0, \ x \notin \left[-\dfrac{a(b-|y|)}{b}, \dfrac{a(b-|y|)}{b}\right], \end{cases}$$

$$f_2(y \mid x) = \begin{cases} \dfrac{a}{2b(a-|x|)}, \ y \in \left[-\dfrac{b(a-|x|)}{a}, \dfrac{b(a-|x|)}{a}\right], \\ 0, \ y \notin \left[-\dfrac{b(a-|x|)}{a}, \dfrac{b(a-|x|)}{a}\right], \end{cases}$$





$\mathrm{cov}(X,Y) = 0$, $\mathrm{K} = \begin{pmatrix} a^2/6 & 0 \\ & b^2/6 \end{pmatrix}$, $\mathrm{R} = \begin{pmatrix} 1 & 0 \\ & 1 \end{pmatrix}$, sõltuvad, mittekorreleeruvad.

$y = 0 \ \ (-a \le x \le a), \ \ x = 0 \ \ (-b \le y \le b).$

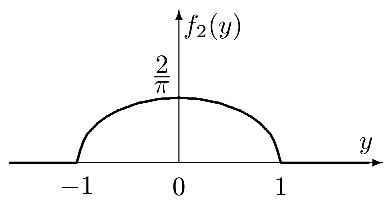
11. Juhusliku vektori (X,Y) jaotustihedus on $f(x,y) = a \cdot \mathbf{1}\left(1^2 - x^2 - y^2\right)$, kus a on konstant. Leidke $a = ?$ Leidke $f_1(x)$, $f_2(y)$, $f_1(x \mid y)$, $f_2(y \mid x)$ ja K. Skitseerige funktsioonide $f_2(y)$ ja $f_2(y \mid x)$ graafikud. Kas selle vektori komponendid

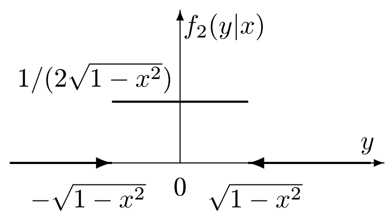
on sõltuvad? Korreleeruvad? V: $a=1/\pi$,

$$f_1(x)=\begin{cases}(2/\pi)\sqrt{1-x^2}, & x\in[-1;1],\\ 0, & x\notin[-1;1],\end{cases}\quad f_2(y)=\begin{cases}(2/\pi)\sqrt{1-y^2}, & y\in[-1;1],\\ 0, & y\notin[-1;1],\end{cases}$$

$$f_1(x\,|\,y)=\begin{cases}1/\left(2\sqrt{1-y^2}\right), & x\in\left[-\sqrt{1-y^2},\sqrt{1-y^2}\right],\\ 0, & x\notin\left[-\sqrt{1-y^2},\sqrt{1-y^2}\right],\end{cases}$$

$$f_2(y\,|\,x)=\begin{cases}1/\left(2\sqrt{1-x^2}\right), & y\in\left[-\sqrt{1-x^2},\sqrt{1-x^2}\right],\\ 0, & y\notin\left[-\sqrt{1-x^2},\sqrt{1-x^2}\right],\end{cases}\quad \mathrm{K}=\begin{pmatrix}1/4 & 0\\ & 1/4\end{pmatrix},$$





12. Juhusliku vektori (X,Y) jaotustihedus on

$$f(x,y)=a\cdot\left(1-x^2-y^2\right)\mathbf{1}\left(1-x^2-y^2\right),$$

kus a on konstant. Leidke a, $f_1(x)$, $f_2(y)$, $f_1(x\,|\,y)$, $f_2(y\,|\,x)$, $\operatorname{cov}(X,Y)$, K ja R. Kas selle vektori komponendid on sõltuvad? Korreleeruvad? Leidke regressioonijooned $y=\mathrm{E}(Y\,|\,x)$ ja $x=\mathrm{E}(X\,|\,y)$. V: $a=2/\pi$, $\operatorname{cov}(X,Y)=0$,

$$\begin{aligned}f_1(x)&=8\left(1-x^2\right)\sqrt{1-x^2}/(3\pi), \quad x\in[-1;1],\\ f_2(y)&=8\left(1-y^2\right)\sqrt{1-y^2}/(3\pi), \quad y\in[-1;1],\end{aligned}\quad \mathrm{K}=\begin{pmatrix}1/6 & 0\\ & 1/6\end{pmatrix},$$

$$f_1(x\,|\,y)=\begin{cases}\dfrac{3\left(1-x^2-y^2\right)}{\sqrt{\left(1-y^2\right)^3}}, & x\in\left[-\sqrt{1-y^2},\sqrt{1-y^2}\right],\\ 0, & x\notin\left[-\sqrt{1-y^2},\sqrt{1-y^2}\right],\end{cases}\quad \mathrm{R}=\begin{pmatrix}1 & 0\\ & 1\end{pmatrix},$$

$$f_2(y\,|\,x)=\begin{cases}\dfrac{3\left(1-x^2-y^2\right)}{\sqrt{\left(1-x^2\right)^3}}, & y\in\left[-\sqrt{1-x^2},\sqrt{1-x^2}\right],\\ 0, & y\in\left[-\sqrt{1-x^2},\sqrt{1-x^2}\right],\end{cases}$$

sõltuvad, mittekorreleeruvad, $y=0$ $(x\in[-1;1])$, $x=0$ $(y\in[-1;1])$.

13. Vektor (X,Y) allub tsentraalsümmeetrilisele normaaljaotusele tihedusega $f(x,y)=\left(2\pi\sigma^2\right)^{-1}\exp\left(-\left(x^2+y^2\right)/\left(2\sigma^2\right)\right)$. Leidke $f_1(x)$, $f_2(y)$, $f_1(x\,|\,y)$, $f_2(y\,|\,x)$, $y=\mathrm{E}(Y\,|\,x)$, $x=\mathrm{E}(X\,|\,y)$, $P(|X|+|Y|\le\sigma)$, $\operatorname{cov}(X,Y)$, K ja R. Kas selle vektori komponendid on sõltuvad? Korreleeruvad?

V: $f_1(x)=\left(\sqrt{2\pi}\sigma\right)^{-1}\exp\left(-x^2/\left(2\sigma^2\right)\right)$, $f_2(y)=\left(\sqrt{2\pi}\sigma\right)^{-1}\exp\left(-y^2/\left(2\sigma^2\right)\right)$, $f_1(x|y)=f_1(x)$, $f_2(x|y)=f_2(y)$, $y=0$, $x=0$, $\operatorname{cov}(X,Y)=0$, ≈ 0.271, $\mathrm{K}=\begin{pmatrix}\sigma^2 & 0\\ & \sigma^2\end{pmatrix}$, $\mathrm{R}=\begin{pmatrix}1 & 0\\ & 1\end{pmatrix}$, sõltumatud, mittekorreleeruvad.

14. Juhusliku vektori (X,Y) komponendid X ja Y on sõltumatud, kusjuures X allub normaaljaotusele parameetritega a ja σ ning Y ühtlasele jaotusele lõigul $[0;1]$. Leidke $f(x,y)$ ja $F(x,y)$.

V: $f(x,y)=\begin{cases}\left(\sqrt{2\pi}\sigma\right)^{-1}\exp\left(-(x-a)^2/\left(2\sigma^2\right)\right),\ y\in[0;1],\\ 0,\ \ y\notin[0;1],\end{cases}$
$F(x,y)=\begin{cases}0,\ y<0,\\ y\left(0.5+\Phi\left((x-a)/\sigma\right)\right),\ 0\le y\le 1,\\ 0.5+\Phi\left((x-a)/\sigma\right),\ y>1.\end{cases}$

15. Vektori (X,Y) jaotustihedus on $f(x,y)=a/(1+x^2+y^2+x^2y^2)$. Leidke a, $F(x,y)$, $f(x,y)$, $f_1(x)$, $f_2(y)$, $f_1(x\,|\,y)$, $f_2(y\,|\,x)$, $\operatorname{cov}(X,Y)$. Kas selle vektori komponendid on sõltuvad? Korreleeruvad? V: $a=1/\pi^2$,
$F(x,y)=(\arctan x+\pi/2)(\arctan y+\pi/2)/\pi^2$, $f_1(x)=(\arctan x+\pi/2)/\pi$, $f_2(y)=(\arctan y+\pi/2)/\pi$, $f_1(x|y)=f_1(x)$, $f_2(y|x)=f_2(y)$, $\operatorname{cov}(X,Y)\overset{?}{=}0$, sõltumatud, mittekorreleeruvad.

16. On antud juhusliku suuruse X jaotustihedus

$$f_1(x)=\begin{cases}\lambda\exp(-\lambda x),\ \text{kui } x>0,\\ 0,\ \text{kui } x\le 0,\end{cases}$$

kus $\lambda>0$, ja $Y=\exp(-X)$. Leidke $f_2(y|x)$, $f(x,y)$, $f_2(y)$, $f_1(x|y)$, $\operatorname{cov}(X,Y)$, K ja R. Kas vektori (X,Y) komponendid on sõltuvad? Korreleeruvad?
V: $f_2(y|x)=\delta\left(y-\exp(-x)\right)$,
$f(x,y)=\begin{cases}\lambda\exp(-\lambda x)\,\delta\left(y-\exp(-x)\right),\ x>0,\\ 0,\ x\le 0,\end{cases}$ $f_2(y)=\begin{cases}\lambda y^{\lambda-1},\ y\in(0;1),\\ 0,\ y\notin(0;1),\end{cases}$
$f_1(x|y)=\begin{cases}y^{1-\lambda}\exp(-\lambda x)\,\delta\left(y-\exp(-x)\right),\ y\in(0;1)\ \wedge\ x>0,\\ 0,\ y\notin(0;1)\ \vee\ x>0,\end{cases}$,
$\operatorname{cov}(X,Y)=-1/(\lambda+1)^2$, sõltuvad, korreleeruvad,
$\mathrm{K}=\begin{pmatrix}1/\lambda^2 & -1/(\lambda+1)^2\\ & \lambda/\left((\lambda+2)(\lambda+1)^2\right)\end{pmatrix}$, $\mathrm{R}=\begin{pmatrix}1 & -\sqrt{\lambda(\lambda+2)}/(\lambda+1)\\ & 1\end{pmatrix}$.

17. Suurus X allub ühtlasele jaotusele lõigul $[-1;1]$ ja $Y=X^2$. Leidke $f_2(y)$.
V: $f_2(y)=\begin{cases}1/\left(2\sqrt{y}\right),\ y\in(0;1),\\ 0,\ y\notin(0;1).\end{cases}$

18. Olgu $f_1(x)=\begin{cases}x/2,\ x\in[0;2],\\ 0,\ x\notin[0;2]\end{cases}$ ja $Y=\exp\left(-X^2/2\right)$. Leidke $f_2(y)$.
V: $f_2(y)=\begin{cases}1/(2y),\ y\in\left(e^{-2};1\right),\\ 0,\ y\notin\left(e^{-2};1\right).\end{cases}$

19. Olgu $f_1(x)=\begin{cases}1/(2a),\ x\in[-a;a],\\ 0,\ x\notin[-a;a]\end{cases}$ ja $Y=X^4$. Leidke $f_2(y)$.
V: $f_2(y)=\begin{cases}1/\left(4a\sqrt[4]{y^3}\right),\ y\in\left(0;a^4\right),\\ 0,\ y\notin\left(0;a^4\right).\end{cases}$

20. Olgu $f_1(x)=\begin{cases}nx^{n-1},\ x\in(0;1),\\ 0,\ x\notin(0;1)\end{cases}$ ja $Y=X^n$ $(n\in\mathbf{N})$. Leidke $f_2(y)$.
V: $f_2(y)=\begin{cases}1,\ y\in(0;1),\\ 0,\ y\notin(0;1).\end{cases}$

21. Olgu $f_1(x)=\begin{cases}\sin x,\ x\in(0;\pi/2),\\ 0,\ x\notin(0;\pi/2)\end{cases}$ ja $Y=\cos X$. Leidke $f_2(y)$.

V: $f_2(y) = \begin{cases} 1,\ y \in (0;1)\,, \\ 0,\ y \notin (0;1)\,. \end{cases}$

22. Olgu $f_1(x) = \begin{cases} (\tan x)\,/\left(\ln\sqrt{2}\right),\ x \in (0;\pi/4)\,, \\ \quad 0,\ x \notin (0;\pi/4) \end{cases}$ ja $Y = \ln(\cos X)$. Leidke $f_2(y)$. V: $f_2(y) = \begin{cases} 1/\left(\ln\sqrt{2}\right),\ y \in \left(\ln\left(\sqrt{2}/2\right);0\right), \\ 0,\ y \notin \left(\ln\left(\sqrt{2}/2\right);0\right). \end{cases}$

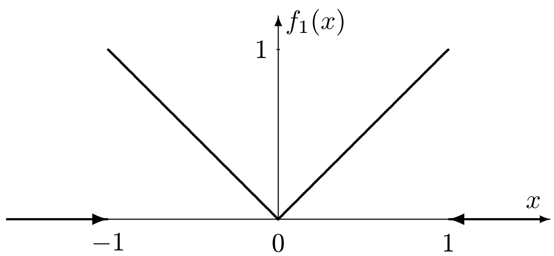
23. On antud pideva juhusliku suuruse X jaotustihedus $f_1(x)$. Avaldage suuruse $Y = |1 - X|$ jaotustihedus $f_2(y)$ tiheduse $f_1(x)$ abil.
V: $f_2(y) = \begin{cases} f_1(1-y) + f_1(1+y),\ y \in (0;+\infty)\,, \\ 0,\ y \in (-\infty;0]\,. \end{cases}$

24. Ringi raadius R allub Rayleigh' jaotusele tihedusega

$$f_1(r) = \begin{cases} \left(r/\sigma^2\right)\exp\left(-r^2/\left(2\sigma^2\right)\right),\ \text{kui } r > 0, \\ \quad 0,\ \text{kui } r \le 0. \end{cases}$$

Leidke suuruse $S = \pi R^2$ jaotustihedus $f_2(s)$.
V: $f_2(s) = \begin{cases} 1/\left(2\pi\sigma^2\right)\exp\left(-s/\left(2\pi\sigma^2\right)\right),\ s \in (0;+\infty)\,, \\ 0,\ s \in (-\infty;0]\,. \end{cases}$

25. Juhusliku suuruse X jaotustihedus $f_1(x)$ on antud graafiliselt



Leidke suuruse $Y = 1 - X^2$ jaotustihedus $f_2(y)$ ja $f(x,y)$.
V: $f_2(y) = \begin{cases} 2 - 2y,\ , \\ 0,\ y \notin (0;1)\,, \end{cases}$ $f(x,y) = \begin{cases} |\,x\,|\ \delta\left(y - 1 + x^2\right),\ x \in [-1;1]\,, \\ 0,\ x \notin [-1;1]\,. \end{cases}$

26. Juhuslik suurus X allub Cauchy jaotusele tihedusega $f_1(x) = 1/\left(\pi\left(1 + x^2\right)\right)$. Leidke suuruse $Y = 1/X$ jaotustihedus $f_2(y)$. V: $f_2(y) = 1/\left(\pi\left(y^2 + 1\right)\right)$.

27. On antud juhusliku suuruse X jaotustihedus

$$f_1(x) = \begin{cases} (\cos x)\,/2,\ \text{kui } x \in (-\pi/2, \pi/2)\,, \\ 0,\ \text{kui } x \notin (-\pi/2, \pi/2) \end{cases}$$

ja $Y = \sin X$. Leidke $f_2(y|x)$, $f(x,y)$, $f_2(y)$ ja K. Kas vektori (X,Y) komponendid on sõltuvad? Korreleeruvad?
V: $f_2(y|x) = \delta\,(y - \sin x)\,,\ f(x,y) = \begin{cases} 0.5\cos x\,\delta\,(y - \sin x)\,,\ x \in (-\pi/2, \pi/2)\,, \\ 0,\ x \notin (-\pi/2, \pi/2)\,, \end{cases}$
$f_2(y) = \begin{cases} 0.5,\ y \in (-1;1)\,, \\ 0,\ y \notin (-1;1)\,, \end{cases}$ $\quad \mathrm{K} = \begin{pmatrix} \pi^2/4 + 2 & 1 \\ & 1/3 \end{pmatrix},$
sõltuvad, korreleeruvad.

28. Olgu $f_1(x)$ pideva suuruse X jaotustihedus. Leidke juhusliku suuruse $Y = \min\left\{X,\ X^2\right\}$ jaotustihedus $f_2(y)$. V: $f_2(y) = \begin{cases} f_1(y),\ y \notin (0;1)\,, \\ f_1(y)/\left(2\sqrt{y}\right),\ y \in (0;1)\,. \end{cases}$

29. Olgu suuruse X jaotustihedus $f_1(x)$. Leidke suuruse $Y = \max\left\{X,\, X^2\right\}$ jaotustihedus $f_2(y)$. V: $f_2(y) = \begin{cases} f_1(y) + f_1(-\sqrt{y})/\left(2\sqrt{y}\right),\ y \in (0;1)\,, \\ \left(f_1(\sqrt{y}) + f_1(-\sqrt{y})\right)/\left(2\sqrt{y}\right),\ y \in (1;+\infty)\,, \\ 0,\ y \in (-\infty;0] \ \vee\ y = 1. \end{cases}$

30. Olgu suuruse X jaotustihedus $f_1(x)$. Leidke suuruse $Y = \min\{X,\, 1\}$ jaotustihedus $f_2(y)$. V: $f_2(y) = \begin{cases} f_1(y),\ y \in (-\infty, 1)\,, \\ \delta\,(y-1)\int_1^{+\infty} f_1(x)dx,\ y \in (0;1)\,. \end{cases}$

31. Olgu vektor (X, Y) ühtlase jaotusega ruudus $[0;1] \times [0;1]$ ja S ristküliku $[0;X] \times [0;Y]$ pindala. Leidke suuruse S jaotustihedus $f(s)$. V: $f(s) = -\ln s$ $(s \in (0;1])\,.$

32. Tõestage, et kahe mittekorreleeruva juhusliku suuruse korrutise keskväärtus on tegurite keskväärtuste korrutis.

33. Juhusliku suuruse X keskväärtus on m_x ja dispersioon D_x. Leidke suuruste $Y = -X$, $Z = X + 2Y - 1$ ja $U = 3X - Y + 2Z - 3$ keskväärtused ja dispersioonid. V: $\mathrm{E}Y = -m_x$, $\mathrm{E}Z = -m_x - 1$, $\mathrm{E}U = 2m_x - 5$, $\mathrm{D}Y = D_x$, $\mathrm{D}Z = D_x$, $\mathrm{D}U = 4D_x$.

34. Juhusliku vektori (X, Y, Z) kovariatsioonimaatriks on

$$\begin{pmatrix} 4 & -2 & 8 \\ & 9 & 3 \\ & & 16 \end{pmatrix}.$$

Leidke selle vektori korrelatsioonimaatriks R. Leidke juhusliku suuruse

$$U = 2X - 3Y + 4Z - 1$$

keskväärtus $\mathrm{E}U$ ja dispersioon $\mathrm{D}U$, kui $\mathrm{E}X = 1$, $\mathrm{E}Y = -2$ ning $\mathrm{E}Z = 3$.

V: $\mathrm{R} = \begin{pmatrix} 1 & -1/3 & 1 \\ & 1 & 1/4 \\ & & 1 \end{pmatrix}$, $\quad \mathrm{E}U = 19,\ \mathrm{D}U = 433.$

Peatükk 4

Juhuslikud funktsioonid

4.1 Juhusliku funktsiooni jaotusfunktsioonid ja jaotustihedused

Olgu T lõpmatu hulk.

Definitsioon 1. *Juhuslikuks funktsiooniks* $X(t)$ $(t \in T)$ nimetatakse funktsiooni, mille väärtus argumendi iga väärtuse t korral on juhuslik suurus.

Kui juhusliku funktsiooni $X(t)$ argument t on aeg, siis seda funktsiooni nimetatakse ka *juhuslikuks protsessiks* või *stohhastiliseks protsessiks.* Kui T on loenduv hulk, siis juhuslikku protsessi $X(t)$ nimetatakse *juhuslikuks jadaks* ehk *aegreaks.* Juhusliku funktsiooni $X(t)$ väärtus argumendi fikseeritud väärtusel t on juhuslik suurus, mida nimetatakse argumendi väärtusele t vastavaks *juhusliku funktsiooni lõikeks.* Juhuslikku funktsiooni on teataval määral võimalik iseloomustada selle funktsiooni lõigetest $X(t_k)$ $(k = 1; \ldots ; n,\ t_k \in T)$ moodustatud juhusliku vektori $(X(t_1), \ldots , X(t_n))$ abil.

Definitsioon 2. Kindlat funktsiooni $x(t)$, mille juhuslik funktsioon $X(t)$ omandab katse käigus, nimetatakse *juhusliku funktsiooni* $X(t)$ *realisatsiooniks.*

Näide 1. Olgu $X(t)$ $(t \in T)$ voolutugevus vooluahelas ajahetkel t. Tegemist on juhusliku funktsiooniga.

Näide 2. Olgu $X(t)$ $(t \in T)$ temperatuur fikseeritud punktis ajahetkel t. Tegemist on juhusliku funktsiooniga.

Näide 3. Olgu $X(t)$ $(t \in T)$ raketi kõrvalekalle meetrites ettenähtud trajektoorist ajahetkel t. Tegemist on juhusliku funktsiooniga.

Definitsioon 3. Funktsiooni

$$F_n\left(x_1, \ldots , x_n; t_1, \ldots , t_n\right) \stackrel{\text{def.}}{=} P\left((X(t_1) < x_1) \cdots (X(t_n) < x_n)\right) \qquad (4.1.1)$$

nimetatakse juhusliku funktsiooni $X(t)$ *n-mõõtmeliseks jaotusfunktsiooniks.*

Tegemist on selle funktsiooni lõigetest koostatud juhusliku vektori $(X(t_1), \ldots , X(t_n))$ jaotusfunktsiooniga. Juhuslike vektorite korral uurisime, kuidas avaldub juhusliku vektori jaotustihedus jaotusfunktsiooni kaudu. Saame juhusliku funktsiooni $X(t)$ *n-mõõtmelise jaotustiheduse*

$$f_n(x_1,\ldots,x_n;t_1,\ldots,t_n)=\frac{\partial^n F_n(x_1,\ldots,x_n;t_1,\ldots,t_n)}{\partial x_1\ldots\partial x_n}. \tag{4.1.2}$$

Kehtib seos

$$\begin{gathered}F_n(x_1,\ldots,x_n;t_1,\ldots,t_n)=\\ =\int\limits_{-\infty}^{x_1}d\xi_1\int\limits_{-\infty}^{x_2}d\xi_2\ldots\int\limits_{-\infty}^{x_n}f_n(\xi_1,\ldots,\xi_n;t_1,\ldots,t_n)\,d\xi_n.\end{gathered} \tag{4.1.3}$$

Kasutame juhusliku vektori jaotustiheduse omadust

$$\begin{gathered}f_{n-1}(x_1,\ldots,x_{n-1};t_1,\ldots,t_{n-1})=\int\limits_{-\infty}^{+\infty}f_n(x_1,\ldots,x_n;t_1,\ldots,t_n)\,dx_n,\\ f_{n-2}(x_1,\ldots,x_{n-2};t_1,\ldots,t_{n-2})=\int\limits_{-\infty}^{+\infty}f_n(x_1,\ldots,x_{n-1};t_1,\ldots,t_{n-1})\,dx_{n-1},\\ \ldots\ldots\ldots\ldots\ldots\ldots\\ f_1(x_1;t_1)=\int\limits_{-\infty}^{+\infty}f_2(x_1,x_2;t_1,t_2)\,dx_2.\end{gathered}$$

Seega, teades juhusliku funktsiooni $X(t)$ n-mõõtmelist jaotustihedust $f_n(x_1,\ldots,x_n;t_1,\ldots,t_n)$, saame leida kõik madalama mõõtmega jaotustihedused integreerimise teel.

Definitsioon 4. Öeldakse, et juhuslik funktsioon allub normaaljaotusele, kui kõik tema mitmemõõtmelised lõiked alluvad normaaljaotusele.

Näide 4. Vaatleme konstantset juhuslikku funktsiooni $X(t)=U$, kus $F(u)$ on juhusliku suuruse U jaotusfunktsioon. Leiame funktsiooni $X(t)$ ühemõõtmelise ja kahemõõtmelise jaotusfunktsiooni

$$\begin{aligned}F_1(x_1;t_1)&=P(X(t_1)<x_1)=P(U<x_1)=F(x_1),\\ F_n(x_1,x_2;t_1,t_2)&=P((X(t_1)<x_1)(X(t_2)<x_2))=\\ &=P((U<x_1)(U<x_2))=\\ &=P(U<\min\{x_1,x_2\})=\\ &=F(\min\{x_1,x_2\}).\qquad\diamondsuit\end{aligned}$$

Näide 5. Olgu $X(t)=t\,U+b$, kus b on kindel suurus ja juhuslik suurus $U\sim N(a,\sigma)$. Leiame funktsiooni $X(t)$ ühe- ja kahemõõtmelise jaotusfunktsiooni ja ühemõõtmelise jaotustiheduse.

Saame

$$\begin{aligned}
F_1(x_1;t_1) &= P(X(t_1)<x_1) = P(t_1U+b<x_1) = [t_1>0] = \\
&= P\left(U<\frac{x_1-b}{t_1}\right) = \frac{1}{\sigma\sqrt{2\pi}}\int\limits_{-\infty}^{(x_1-b)/t_1}\exp\left(-\frac{(u-a)^2}{2\sigma^2}\right)du, \\
F_2(x_1,x_2;t_1,t_2) &= P((X(t_1)<x_1)(X(t_2)<x_2)) = \\
&= P((t_1U+b<x_1)(t_2U+b<x_2)) = [t_1,t_2>0] = \\
&= P\left(\left(U<\frac{x_1-b}{t_1}\right)\left(U<\frac{x_2-b}{t_2}\right)\right) = \\
&= P\left(U<\min\left\{\frac{x_1-b}{t_1},\frac{x_2-b}{t_2}\right\}\right) =
\end{aligned}$$

$$= \frac{1}{\sigma\sqrt{2\pi}}\int\limits_{-\infty}^{\min\{(x_1-b)/t_1,(x_2-b)/t_2\}}\exp\left(-(u-a)^2/\left(2\sigma^2\right)\right)du$$

ja

$$\begin{aligned}
f_1(x_1;t_1) &= \frac{\partial F_1(x_1;t_1)}{\partial x_1} = \\
&= \frac{\partial}{\partial x_1}\frac{1}{\sigma\sqrt{2\pi}}\int\limits_{-\infty}^{(x_1-b)/t_1}\exp\left(-(u-a)^2/\left(2\sigma^2\right)\right)du = \\
&= \frac{1}{\sigma\sqrt{2\pi}}\exp\left(-((x_1-b)/t_1-a)^2/\left(2\sigma^2\right)\right)\cdot\frac{1}{t_1} = \\
&= \frac{1}{\sigma t_1\sqrt{2\pi}}\exp\left(-(x_1-(at_1+b))^2/\left(2\sigma^2t_1^2\right)\right). \qquad \diamondsuit
\end{aligned}$$

4.2 Juhusliku funktsiooni keskväärtus, dispersioon ja kovariatsioon

Uurime järgnevalt juhusliku funktsiooni $X(t)$ $(t\in T)$ iseloomustamist teatud kindlate funktsioonide abil. Kui fikseerime argumendi t väärtuse, saame tulemuseks argumendi väärtusele t vastava juhusliku funktsiooni $X(t)$ lõike. See lõige on juhuslik suurus. Eeldame, et sel lõikel leidub keskväärtus $\mathrm{E}X(t)$. Nii saame meie poolt fikseeritud argumendi t väärtusel seada argumendi sellele väärtusele vastavusse kindla suuruse $\mathrm{E}X(t)$. Lihtsuse mõttes eeldame, et selline vastavusse seadmine on võimalik iga $t\in T$ korral. Tulemuseks saame hulgal T määratud kindla funktsiooni $m_x(t)$, mida me järgnevalt nimetame juhusliku funktsiooni $X(t)$ *keskväärtuseks* ehk *keskväärtusfunktsiooniks.* Lühidalt on eelnev arutelu kirja pandav kujul

$$m_x(t)\overset{\text{def.}}{=}\mathrm{E}X(t). \tag{4.2.1}$$

Olgu

$$X^{\circ}(t) \overset{\text{def.}}{=} X(t) - m_x(t). \tag{4.2.2}$$

Funktsioon $X^{\circ}(t)$ on *tsentreeritud* juhuslik funktsioon. Keskväärtuse $m_x(t)$ saame esitada ühemõõtmelise jaotustiheduse $f_1(x;t)$ abil

$$m_x(t) = \int_{-\infty}^{+\infty} x f_1(x;t)\, dx. \tag{4.2.3}$$

Analoogilist mõttekäiku kasutades defineerime juhusliku funktsiooni $X(t)$ *dispersiooni* ehk *dispersioonifunktsiooni*

$$D_x(t) \overset{\text{def.}}{=} E\left(X^{\circ}(t)\right)^2, \tag{4.2.4}$$

mis on kindel funktsioon, kusjuures ühemõõtmelise jaotustiheduse abil avaldub ta kujul

$$D_x(t) = \int_{-\infty}^{+\infty} \left(x - m_x(t)\right)^2 f_1(x;t)\, dx. \tag{4.2.5}$$

Defineerime juhusliku funktsiooni $X(t)$ *standardhälbe*:

$$\sigma_x(t) \overset{\text{def.}}{=} \sqrt{D_x(t)}. \tag{4.2.6}$$

Keskväärtus $m_x(t)$ on kindel funktsioon, mille väärtus on igal argumendi väärtusel t juhusliku funktsiooni $X(t)$ võimalike realisatsioonide väärtuste keskmine. Juhusliku funktsiooni $X(t)$ standardhälve $\sigma_x(t)$ iseloomustab juhusliku funktsiooni kõrvalekallet keskväärtusest $m_x(t)$, st võimalike realisatsioonide pere hajuvust keskväärtuse $m_x(t)$ suhtes.

Juhusliku funktsiooni $X(t)$ *kovariatsiooni* ehk *kovariatsioonifunktsiooni* $K_x(t_1, t_2)$ defineerime seosega

$$K_x(t_1, t_2) \overset{\text{def.}}{=} \mathrm{E}\left(\left(X^{\circ}(t_1)\right)\left(X^{\circ}(t_2)\right)\right), \tag{4.2.7}$$

st kovariatsiooni defineerime argumendi väärtustel t_1 ja t_2 sooritatud juhusliku funktsiooni lõigetest $X(t_1)$ ja $X(t_2)$ moodustatud vektori $(X(t_1), X(t_2))$ kovariatsioonina. Juhusliku vektori $(X(t_1), X(t_2))$ korral kehtib võrdus

$$\mathrm{E}\left(\left(X^{\circ}(t_1)\right)\left(X^{\circ}(t_2)\right)\right) = \mathrm{E}\left(\left(X^{\circ}(t_2)\right)\left(X^{\circ}(t_1)\right)\right).$$

Seega järeldub seosest (4.2.7) $K_x(t_2, t_1) = K_x(t_1, t_2)$, st kovariatsioon $K_x(t_1, t_2)$ on sümmeetriline oma muutujate suhtes. Avaldame kovariatsiooni $K_x(t_1, t_2)$ kahemõõtmelise jaotustiheduse abil:

$$K_x(t_1, t_2) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \left(x_1 - m_x(t_1)\right)\left(x_2 - m_x(t_2)\right) f_2(x_1, x_2; t_1, t_2)\, dx_1 dx_2. \tag{4.2.8}$$

Defineerime juhusliku funktsiooni $X(t)$ *korrelatsiooni* ehk *korrelatsioonifunktsiooni*

$$R_x(t_1, t_2) \overset{\text{def.}}{=} \frac{K_x(t_1, t_2)}{\sigma_x(t_1)\sigma_x(t_2)}, \tag{4.2.9}$$

mis rahuldab seost $|R_x(t_1,t_2)| \leq 1$. Juhusliku funktsiooni $X(t)$ korrelatsioon $R_x(t_1,t_2)$ iseloomustab seost funktsiooni lõigete $X(t_1)$ ja $X(t_2)$ vahel.

Leiame

$$K_x(t,t) = \mathrm{E}\left(\left(X^\circ(t)\right)\left(X^\circ(t)\right)\right) = \mathrm{E}\left(X^\circ(t)\right)^2 = D_x(t),$$

s.o

$$K_x(t,t) = D_x(t). \tag{4.2.10}$$

Seoseni (4.2.10) jõuame ka valemist (4.2.8), lähtudes valikust $t_1 = t_2 = t$:

$$\begin{aligned}
K_x(t_1,t_2) &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \left(x_1 - m_x(t_1)\right)\left(x_2 - m_x(t_2)\right) f_2\left(x_1,x_2;t_1,t_2\right) dx_1 dx_2 = \\
&= \left[\begin{array}{c} t_1 = t_2 = t \;\Rightarrow X\left(t_1\right) = X\left(t_2\right) = X\left(t\right) \Rightarrow \\ \Rightarrow\; f_2\left(x_1,x_2;t,t\right) = f_1\left(x_1;t\right)\cdot 1 \cdot \delta\left(x_2 - x_1\right) \end{array}\right] = \\
&= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \left(x_1 - m_x(t)\right)\left(x_2 - m_x(t)\right) f_1\left(x_1;t\right)\delta\left(x_2 - x_1\right) dx_1 dx_2 = \\
&= \int_{-\infty}^{+\infty} \left(x_1 - m_x(t)\right) f_1\left(x_1;t\right) dx_1 \int_{-\infty}^{+\infty} \left(x_2 - m_x(t)\right)\delta\left(x_2 - x_1\right) dx_2 = \\
&= \int_{-\infty}^{+\infty} \left(x_1 - m_x(t)\right)^2 f_1\left(x_1;t\right) dx_1 = D_x\left(t\right).
\end{aligned}$$

Näide 1. Olgu $X(t) = t\,U + b$, kus b on kindel suurus ja $U \sim N\left(a,\sigma\right)$. Leiame funktsiooni $X(t)$ keskväärtuse $m_x(t)$, tsentreeritud juhusliku funktsiooni $X^\circ(t)$, kovariatsiooni $K_x(t_1,t_2)$, dispersiooni $D_x(t)$ ja standardhälbe $\sigma_x(t)$ ning korrelatsiooni $R_x(t_1,t_2)$.

Saame

$$\begin{aligned}
m_x(t) &= \mathrm{E}\left(tU + b\right) = t\,\mathrm{E}U + \mathrm{E}b = at + b, \\
X^\circ(t) &= X(t) - m_x(t) = \left(tU + b\right) - \left(at + b\right) = \\
&= tU - at = t\left(U - a\right) = tU^\circ, \\
K_x(t_1,t_2) &= \mathrm{E}\left(\left(X^\circ(t_1)\right)\left(X^\circ(t_2)\right)\right) = \mathrm{E}\left(\left(t_1U^\circ\right)\left(t_2U^\circ\right)\right) = \\
&= t_1t_2E\left(\left(U^\circ\right)^2\right) = t_1t_2\mathrm{D}U = \sigma^2 t_1t_2, \\
D_x(t) &= K_x(t,t) = \sigma^2t^2,\ \sigma_x(t) = \sqrt{t^2\sigma^2} = \sigma\left|t\right|, \\
R_x(t_1,t_2) &= \frac{K_x(t_1,t_2)}{\sigma_x(t_1)\sigma_x(t_2)} = \frac{\sigma^2t_1t_2}{\sigma\left|t_1\right|\sigma\left|t_2\right|} = \operatorname{sign}\left(t_1t_2\right). \qquad \diamondsuit
\end{aligned}$$

Näide 2. Olgu $Z(t) = Xe^t + Ye^{-t}$, kus X ja Y on juhuslikud suurused ning $\mathrm{E}X = 1$, $\mathrm{E}Y = -1$ ja $\begin{pmatrix} 9 & -3 \\ & 4 \end{pmatrix}$ on vektori (X,Y) kovariatsioonimaatriks.

Leiame keskväärtuse $m_z(t)$, tsentreeritud juhusliku funktsiooni $Z^\circ(t)$, kovariatsiooni $K_z(t_1,t_2)$, dispersiooni $D_z(t)$ ja standardhälbe $\sigma_z(t)$ ning korrelatsiooni $R_z(t_1,t_2)$.

Saame

$$m_z(t) = \mathrm{E}\,(Xe^t + Ye^{-t}) = e^t\,\mathrm{E}X + e^{-t}\mathrm{E}Y = e^t - e^{-t},$$

$$Z^\circ(t) = Z(t) - m_z(t) = (Xe^t + Ye^{-t}) - (e^t - e^{-t}) =$$

$$= (Xe^t - e^t) + (Ye^{-t} - e^{-t}) = X^\circ e^t + Y^\circ e^{-t},$$

$$K_z(t_1,t_2) = \mathrm{E}\,((Z^\circ(t_1))\,(Z^\circ(t_2))) =$$

$$= \mathrm{E}\,((X^\circ e^{t_1} + Y^\circ e^{-t_1})\,(X^\circ e^{t_2} + Y^\circ e^{-t_2})) =$$

$$= \mathrm{E}\left((X^\circ)^2\, e^{t_1+t_2} + Y^\circ X^\circ e^{-t_1+t_2} + X^\circ Y^\circ e^{t_1-t_2} + (Y^\circ)^2\, e^{-t_1-t_2}\right) =$$

$$= e^{t_1+t_2}\mathrm{E}\,(X^\circ)^2 + e^{-t_1+t_2}\mathrm{E}\,(Y^\circ X^\circ) + e^{t_1-t_2}\mathrm{E}\,(X^\circ Y^\circ) + e^{-t_1-t_2}\mathrm{E}\,(Y^\circ)^2 =$$

$$= e^{t_1+t_2}\mathrm{D}X + e^{-t_1+t_2}\mathrm{cov}\,(Y,X) + e^{t_1-t_2}\mathrm{cov}\,(X,Y) + e^{-t_1-t_2}\mathrm{D}Y =$$

$$= 9e^{t_1+t_2} - 3\,(e^{-t_1+t_2} + e^{t_1-t_2}) + 4e^{-t_1-t_2},$$

$$D_z(t) = K_z(t,t) = 9e^{2t} - 6 + 4e^{-2t},\ \sigma_z(t) = \sqrt{9e^{2t} - 6 + 4e^{-2t}},$$

$$R_z(t_1,t_2) = \frac{K_z(t_1,t_2)}{\sigma_z(t_1)\sigma_z(t_2)} = \frac{9e^{t_1+t_2} - 3\,(e^{-t_1+t_2} + e^{t_1-t_2}) + 4e^{-t_1-t_2}}{\sqrt{(9e^{2t_1} - 6 + 4e^{-2t_1})\,(9e^{2t_2} - 6 + 4e^{-2t_2})}}. \qquad \diamondsuit$$

Defineerime kahe juhusliku funktsiooni $X(t)$ ja $Y(t)$ *vastastikuse kovariatsiooni* ehk *vastastikuse kovariatsioonifunktsiooni,* kui

$$K_{xy}(t_1,t_2) \stackrel{\text{def.}}{=} \mathrm{E}\,((X^\circ(t_1))\,(Y^\circ(t_2))) \tag{4.2.11}$$

ja *vastastikuse korrelatsiooni* ehk *vastastikuse korrelatsioonifunktsiooni*

$$R_{xy}(t_1,t_2) \stackrel{\text{def.}}{=} \frac{K_{xy}(t_1,t_2)}{\sigma_x(t_1)\sigma_y(t_2)}. \tag{4.2.12}$$

Juhuslikke funktsioone $X(t)$ ja $Y(t)$ nimetatakse *mittekorreleeruvateks*, kui $K_{xy}(t_1,t_2) \equiv 0$, ja *korreleeruvateks* vastandjuhul.

Kompleksete väärtustega juhuslikku funktsiooni $Z(t)$ võime esitada kujul $Z(t) = X(t) + iY(t)$. Olgu $\overline{Z(t)} \stackrel{\text{def.}}{=} X(t) - iY(t)$,

$$m_z(t) \stackrel{\text{def.}}{=} m_x(t) + im_y(t) \tag{4.2.13}$$

ja

$$Z^\circ(t) \stackrel{\text{def.}}{=} Z(t) - m_z(t) = X(t) + iY(t) - (m_x(t) + im_y(t)) = X^\circ(t) + iY^\circ(t)$$

ning

$$K_z(t_1,t_2) \stackrel{\text{def.}}{=} \mathrm{E}\left((Z^\circ(t_1))\,\overline{(Z^\circ(t_2))}\right) =$$

$$= \mathrm{E}\left((X^\circ(t_1)+iY^\circ(t_1))\,(X^\circ(t_2)-iY^\circ(t_2))\right) =$$

$$= \mathrm{E}\left(X^\circ(t_1)X^\circ(t_2)\right) - i\,\mathrm{E}\left(X^\circ(t_1)Y^\circ(t_2)\right) +$$

$$+i\,\mathrm{E}\left(Y^\circ(t_1)X^\circ(t_2)\right) + \mathrm{E}\left(Y^\circ(t_1)Y^\circ(t_2)\right) =$$

$$= K_x(t_1,t_2) - iK_{xy}(t_1,t_2) + iK_{yx}(t_1,t_2) + K_y(t_1,t_2),$$

st

$$K_z(t_1,t_2) = K_x(t_1,t_2) + K_y(t_1,t_2) + i\left(K_{yx}(t_1,t_2) - K_{xy}(t_1,t_2)\right). \qquad (4.2.14)$$

Kui

$$D_z(t) \stackrel{\text{def.}}{=} \mathrm{E}\left((Z^\circ(t))\,\overline{(Z^\circ(t))}\right) = K_z(t,t) =$$

$$= K_x(t,t) - iK_{xy}(t,t) + iK_{yx}(t,t) + K_y(t,t) =$$

$$= D_x(t) + D_y(t),$$

siis

$$D_z(t) = D_x(t) + D_y(t). \qquad (4.2.15)$$

Lause 1. Komplekssete väärtustega juhusliku funktsiooni $Z(t) = X(t) + iY(t)$ korral kehtivad seosed (4.2.14) ja (4.2.15).

Näide 3. Olgu $Z(t) = X\cos t + i\,Y\sin t$, kus X ja Y on juhuslikud suurused ning $\mathrm{E}X = 1$, $\mathrm{E}Y = -1$ ja

$$\begin{pmatrix} 16 & -6 \\ & 9 \end{pmatrix}$$

on vektori (X,Y) kovariatsioonimaatriks. Leiame keskväärtuse $m_z(t)$, tsentreeritud juhusliku funktsiooni $Z^\circ(t)$, kovariatsiooni $K_z(t_1,t_2)$, dispersiooni $D_z(t)$ ja standardhälbe $\sigma_z(t)$ ning korrelatsiooni $R_z(t_1,t_2)$.

Saame

$$\begin{aligned} m_z(t) &= \mathrm{E}\left(X\cos t + i\,Y\sin t\right) = \cos t\ \mathrm{E}X + i\,\sin t\,\mathrm{E}Y = \cos t\, - i\,\sin t, \\ Z^\circ(t) &= Z(t) - m_z(t) = (X\cos t + i\,Y\sin t) - (\cos t\, - i\,\sin t) = \\ &= (X\cos t - \cos t\,) + i\,(Y\sin t + \sin t) = X^\circ\cos t + i\,Y^\circ\sin t, \end{aligned}$$

$$K_z(t_1,t_2) = \mathrm{E}\left(Z^\circ(t_1)\overline{Z^\circ(t_2)}\right) =$$

$$= \mathrm{E}\left(\left(X^\circ \cos t_1 + i\,Y^\circ \sin t_1\right)\left(X^\circ \cos t_2 - i\,Y^\circ \sin t_2\right)\right) =$$

$$= \cos t_1 \cos t_2\, \mathrm{E}\left(X^\circ\right)^2 - i\,\mathrm{E}\left(X^\circ Y^\circ\right) \cos t_1 \sin t_2 +$$

$$+ i\, \sin t_1 \cos t_2\, \mathrm{E}\left(Y^\circ X^\circ\right) + \mathrm{E}\left(Y^\circ\right)^2 \sin t_1 \sin t_2 =$$

$$= \cos t_1 \cos t_2 \mathrm{D}X - i\, \cos t_1 \sin t_2 \operatorname{cov}(X,Y) +$$

$$+ i\, \sin t_1 \cos t_2 \operatorname{cov}(Y,X) + \sin t_1 \sin t_2 \mathrm{D}Y =$$

$$= 16 \cos t_1 \cos t_2 + 6i\left(\cos t_1 \sin t_2 - \sin t_1 \cos t_2\right) + 9 \sin t_1 \sin t_2 =$$

$$= 7 \cos t_1 \cos t_2 + 6i \sin\left(t_2 - t_1\right) + 9 \cos\left(t_2 - t_1\right),$$

$$D_z(t) = 16 \cos^2 t + 9 \sin^2 t = 9 + 7\cos^2 t, \quad \sigma_z(t) = \sqrt{9 + 7\cos^2 t},$$

$$R_z(t_1,t_2) = \frac{16 \cos t_1 \cos t_2 + 6i\left(\cos t_1 \sin t_2 - \sin t_1 \cos t_2\right) + 9 \sin t_1 \sin t_2}{\sqrt{9 + 7\cos^2 t_1}\sqrt{9 + 7\cos^2 t_2}}. \quad \diamondsuit$$

4.3 Tehted juhuslike funktsioonidega

4.3.1 Juhuslike funktsioonide liitmine

Olgu antud kaks juhuslikku funktsiooni $X(t)$ $(t \in T)$ ja $Y(t)$ $(t \in T)$. Kui $Z(t) = X(t) + Y(t)$, siis

$$m_z(t) = \mathrm{E}\left(X(t) + Y(t)\right) = \mathrm{E}X(t) + \mathrm{E}Y(t) = m_x(t) + m_y(t),$$
$$Z^\circ(t) = Z(t) - m_z(t) = X(t) + Y(t) - \left(m_x(t) + m_y(t)\right) = X^\circ(t) + Y^\circ(t)$$

ja

$$K_z(t_1,t_2) = \mathrm{E}\left(\left(Z^\circ(t_1)\right)\left(Z^\circ(t_2)\right)\right) =$$

$$= \mathrm{E}\left(\left(X^\circ(t_1) + Y^\circ(t_1)\right)\left(X^\circ(t_2) + Y^\circ(t_2)\right)\right) =$$

$$= \mathrm{E}\left(X^\circ(t_1) X^\circ(t_2)\right) + \mathrm{E}\left(X^\circ(t_1) Y^\circ(t_2)\right) +$$

$$+ \mathrm{E}\left(X^\circ(t_2) Y^\circ(t_2)\right) + \mathrm{E}\left(Y^\circ(t_1) Y^\circ(t_2)\right) =$$

$$= K_x\left(t_1,t_2\right) + K_y\left(t_1,t_2\right) + K_{xy}\left(t_1,t_2\right) + K_{xy}\left(t_2,t_1\right)$$

ning

$$D_z(t) = K_z(t,t) = D_x\left(t\right) + D_y\left(t\right) + 2K_{xy}\left(t,t\right),$$
$$\sigma_z(t) = \sqrt{D_z(t)} = \sqrt{D_x\left(t\right) + D_y\left(t\right) + 2K_{xy}\left(t,t\right)}.$$

Sõnastame tõestatu.

Lause 1. Kui $Z(t) = X(t) + Y(t)$, siis

$$\begin{aligned} m_z(t) &= m_x(t) + m_y(t),\ D_z(t) = D_x(t) + D_y(t) + 2K_{xy}(t,t), \\ K_z(t_1,t_2) &= K_x(t_1,t_2) + K_y(t_1,t_2) + K_{xy}(t_1,t_2) + K_{xy}(t_2,t_1), \\ \sigma_z(t) &= \sqrt{D_x(t) + D_y(t) + 2K_{xy}(t,t)},\ R_z(t_1,t_2) = \frac{K_z(t_1,t_2)}{\sigma_z(t_1)\sigma_z(t_2)}. \end{aligned}$$

Näide 1. Leiame kahe juhusliku funktsiooni $X(t)$ ja $Y(t)$ summa $Z(t)$ keskväärtuse $m_z(t)$, kovariatsiooni $K_z(t_1,t_2)$, dispersiooni $D_z(t)$, standardhälbe ja korrelatsiooni $R_z(t_1,t_2)$, kui

$$m_x(t) = t,\ K_x(t_1,t_2) = t_1t_2,\ m_y(t) = -t,\ K_y(t_1,t_2) = \exp(\alpha(t_1+t_2))$$

ning

$$K_{xy}(t_1,t_2) = \exp(\beta(t_1-t_2)).$$

Lause 1 alusel saame

$$\begin{aligned} m_z(t) &= m_x(t) + m_y(t) = t - t = 0, \\ K_z(t_1,t_2) &= K_x(t_1,t_2) + K_y(t_1,t_2) + K_{xy}(t_1,t_2) + K_{xy}(t_2,t_1) = \\ &= t_1t_2 + \exp(\alpha(t_1+t_2)) + \exp(\beta(t_1-t_2)) + \exp(\beta(t_2-t_1)), \\ D_z(t) &= D_x(t) + D_y(t) + 2K_{xy}(t,t) = \\ &= t^2 + \exp(2\alpha t) + 2, \\ \sigma_z(t) &= \sqrt{D_z(t)} = \sqrt{t^2 + \exp(2\alpha t) + 2}, \\ R_z(t_1,t_2) &= \frac{K_z(t_1,t_2)}{\sigma_z(t_1)\sigma_z(t_2)} = \\ &= \frac{t_1t_2 + \exp(\alpha(t_1+t_2)) + \exp(\beta(t_1-t_2)) + \exp(\beta(t_2-t_1))}{\sqrt{t_1^2 + \exp(2\alpha t_1) + 2}\sqrt{t_2^2 + \exp(2\alpha t_2) + 2}}. \end{aligned}$$

$\diamondsuit$

4.3.2 Juhusliku funktsiooni korrutamine kindla funktsiooniga

Olgu antud kindel funktsioon $h(t)$ ja juhuslik funktsioon $X(t)$, kusjuures on teada $m_x(t)$ ja $K_x(t_1,t_2)$. Leiame funktsiooni $Y(t) = h(t)X(t)$ põhilised karakteristikud. Saame

$$\begin{aligned} m_y(t) &= \mathrm{E}Y(t) = \mathrm{E}(h(t)X(t)) = h(t)(\mathrm{E}X(t)) = h(t)m_x(t), \\ Y^\circ(t) &= Y(t) - m_y(t) = h(t)X(t) - h(t)m_x(t) = \\ &= h(t)(X(t) - m_x(t)) = h(t)X^\circ(t), \end{aligned}$$

$$\begin{aligned}
K_y(t_1,t_2) &= \mathrm{E}\left(\left(Y^\circ(t_1)\right)\left(Y^\circ(t_2)\right)\right) = \\
&= \mathrm{E}\left(h(t_1)X^\circ(t_1)h(t_2)X^\circ(t_2)\right) = \\
&= h(t_1)h(t_2)\mathrm{E}\left(X^\circ(t_1)X^\circ(t_2)\right) = h(t_1)h(t_2)K_x(t_1,t_2), \\
D_y(t) &= K_y(t,t) = h^2(t)K_x(t,t) = h^2(t)D_x(t), \\
\sigma_y(t) &= \sqrt{K_y(t,t)} = \sqrt{h^2(t)D_x(t)} = |h(t)|\,\sigma_x(t), \\
R_y(t_1,t_2) &= \frac{K_y(t_1,t_2)}{\sigma_y(t_1)\sigma_y(t_2)} = \frac{h(t_1)h(t_2)K_x(t_1,t_2)}{|h(t_1)|\,\sigma_x(t_1)\,|h(t_2)|\,\sigma_x(t_2)} = \\
&= (\operatorname{sign} h(t_1))\,(\operatorname{sign} h(t_2))\,R_x(t_1,t_2).
\end{aligned}$$

Lause 1. Kui $Y(t) = h(t)X(t)$, $h(t)$ on kindel funktsioon ja $X(t)$ juhuslik funktsioon, siis

$$\begin{aligned}
&m_y(t) = h(t)m_x(t),\ K_y(t_1,t_2) = h(t_1)h(t_2)K_x(t_1,t_2), \\
&D_y(t) = h^2(t)D_x(t),\ R_y(t_1,t_2) = (\operatorname{sign} h(t_1))\,(\operatorname{sign} h(t_2))\,R_x(t_1,t_2).
\end{aligned}$$

Näide 1. Leiame juhusliku funktsiooni $Y(t) = X(t)\cos t$ korral $m_y(t)$, $K_y\left(t_1,t_2\right)$, $D_y(t)$, $\sigma_y(t)$ ja $R_y(t_1,t_2)$, kui

$$m_x(t) = t,\ K_x\left(t_1,t_2\right) = \exp\left(-\left|t_2-t_1\right|\right).$$

Et

$$\begin{aligned}
D_x(t) &= K_x\left(t,t\right) = \exp\left(-\left|t-t\right|\right) = 1,\ \sigma_x(t) = 1, \\
R_x(t_1,t_2) &= \frac{K_x(t_1,t_2)}{\sigma_x(t_1)\sigma_x(t_2)} = \frac{\exp\left(-\left|t_2-t_1\right|\right)}{1\cdot 1} = \exp\left(-\left|t_2-t_1\right|\right),
\end{aligned}$$

siis Lause 1 põhjal saame

$$\begin{aligned}
m_y(t) &= h(t)m_x(t) = t\cos t, \\
K_y(t_1,t_2) &= h(t_1)h(t_2)K_x(t_1,t_2) = \\
&= \cos t_1 \cos t_2 \exp\left(-\left|t_2-t_1\right|\right), \\
D_y(t) &= h^2(t)D_x(t) = \cos^2 t,\ \sigma_y(t) = \left|\cos t\right|, \\
R_y(t_1,t_2) &= (\operatorname{sign} h(t_1))\,(\operatorname{sign} h(t_2))\,R_x(t_1,t_2) = \\
&= (\operatorname{sign}\,\cos t_1)\,(\operatorname{sign}\,\cos t_2)\exp\left(-\left|t_2-t_1\right|\right). \qquad \diamondsuit
\end{aligned}$$

4.3.3 Juhusliku funktsiooni integraal

Olgu antud juhuslik funktsioon $X(t)\ \ (t\in T)$, mille kõik realisatsioonid on integreeruvad lõigul $[0;t]\subset T$. Olgu see lõik $[0;t]$ jaotatud punktidega τ_i $(i=0;1;2;\ldots;n)$ n osalõiguks $[\tau_{i-1},\tau_i]$ $(i=1;2;\ldots;n)$, kusjuures $0=\tau_0<\tau_1<\tau_2<\ldots<\tau_n=t$. Olgu $\varsigma_i\in[\tau_{i-1},\tau_i]$ ja $\Delta\tau_i=\tau_i-\tau_{i-1}$. Suurust

$$Y\left(t\right) = \int_0^t X\left(\tau\right)d\tau \overset{\text{def.}}{=} \lim_{\max\Delta\tau_i\to 0}\sum_{i=1}^n X\left(\varsigma_i\right)\Delta\tau_i$$

nimetatakse *juhusliku funktsiooni* $X(t)$ *integraaliks* lõigul $[0;t]$. Suurus $Y(t)$ on juhuslik funktsioon, sest igal argumendi t väärtusel hulgast T on $Y(t)$ väärtus juhuslik.

Leiame juhusliku funktsiooni $Y(t)$ keskväärtuse

$$\begin{aligned} m_y(t) &= \mathrm{E}Y(t) = \mathrm{E}\int_0^t X(\tau)\,d\tau = \mathrm{E}\left(\lim_{\max\Delta\tau_i\to 0}\sum_{i=1}^n X(\varsigma_i)\,\Delta\tau_i\right) = \\ &= \left[\begin{array}{c}\text{eeldusel, et piirväärtuse ja keskväärtuse}\\ \text{leidmise järjekord on muudetav}\end{array}\right] = \\ &= \lim_{\max\Delta\tau_i\to 0}\mathrm{E}\left(\sum_{i=1}^n X(\varsigma_i)\,\Delta\tau_i\right) = \lim_{\max\Delta\tau_i\to 0}\sum_{i=1}^n \left(\mathrm{E}X(\varsigma_i)\right)\Delta\tau_i = \\ &= \lim_{\max\Delta\tau_i\to 0}\sum_{i=1}^n m_x(\varsigma_i)\,\Delta\tau_i = \int_0^t m_x(\tau)\,d\tau. \end{aligned}$$

Et

$$\begin{aligned} Y^\circ(t) = Y(t) - m_y(t) &= \int_0^t X(\tau)\,d\tau - \int_0^t m_x(\tau)\,d\tau = \\ &= \int_0^t \left(X(\tau) - m_x(\tau)\right)d\tau = \int_0^t X^\circ(\tau)\,d\tau, \end{aligned}$$

siis saame leida juhusliku funktsiooni $Y(t)$ kovariatsiooni

$$K_y(t_1,t_2) = \mathrm{E}\left(\left(Y^\circ(t_1)\right)\left(Y^\circ(t_2)\right)\right) =$$

$$\begin{aligned} &= \mathrm{E}\left(\left(\int_0^{t_1} X^\circ(\tau_1)\,d\tau_1\right)\left(\int_0^{t_2} X^\circ(\tau_2)\,d\tau_2\right)\right) = \\ &= \mathrm{E}\left(\int_0^{t_1}\int_0^{t_2} X^\circ(\tau_1)\,X^\circ(\tau_2)\,d\tau_1 d\tau_2\right) = \\ &= \left(\int_0^{t_1}\int_0^{t_2} \mathrm{E}\left(X^\circ(\tau_1)\,X^\circ(\tau_2)\right)d\tau_1 d\tau_2\right) = \int_0^{t_1}\int_0^{t_2} K_x(\tau_1,\tau_2)d\tau_1 d\tau_2, \end{aligned}$$

millest omakorda saame dispersiooni ja standardhälbe

$$\begin{aligned} D_y(t) &= K_y(t,t) = \int_0^t\int_0^t K_x(\tau_1,\tau_2)d\tau_1 d\tau_2, \\ \sigma_y(t) &= \sqrt{D_y(t)} = \sqrt{\int_0^t\int_0^t K_x(\tau_1,\tau_2)d\tau_1 d\tau_2} \end{aligned}$$

ja korrelatsiooni

$$R_y(t_1,t_2) = \frac{K_y(t_1,t_2)}{\sigma_y(t_1)\sigma_y(t_2)} =$$

$$= \frac{\int_0^{t_1}\int_0^{t_2} K_x(\tau_1,\tau_2)d\tau_1 d\tau_2}{\sqrt{\int_0^{t_1}\int_0^{t_1} K_y(\tau_1,\tau_2)d\tau_1 d\tau_2}\sqrt{\int_0^{t_2}\int_0^{t_2} K_y(\tau_1,\tau_2)d\tau_1 d\tau_2}}.$$

Lause 1. Kui $Y(t) = \int_0^t X(\tau)\,d\tau$, siis

$$m_y(t) = \int_0^t m_x(\tau)\,d\tau,\ K_y(t_1, t_2) = \int_0^{t_1}\int_0^{t_2} K_x(\tau_1, \tau_2) d\tau_1 d\tau_2,$$

$$D_y(t) = \int_0^t\int_0^t K_x(\tau_1, \tau_2) d\tau_1 d\tau_2,\ \sigma_y(t) = \sqrt{\int_0^t\int_0^t K_x(\tau_1, \tau_2) d\tau_1 d\tau_2},$$

$$R_y(t_1, t_2) = \frac{\int_0^{t_1}\int_0^{t_2} K_x(\tau_1, \tau_2) d\tau_1 d\tau_2}{\sqrt{\int_0^{t_1}\int_0^{t_1} K_y(\tau_1, \tau_2) d\tau_1 d\tau_2}\sqrt{\int_0^{t_2}\int_0^{t_2} K_y(\tau_1, \tau_2) d\tau_1 d\tau_2}}.$$

Näide 1. Olgu juhusliku funktsiooni $X(t)$ korral teada, et $m_x(t) = 0$ ja $K_x(t_1, t_2) = 1/\left(1 + (t_2 - t_1)^2\right)$. Leiame juhusliku funktsiooni

$$Y(t) = \int_0^t X(\tau)\,d\tau$$

keskväärtuse ja kovariatsiooni, dispersiooni ja standardhälbe.

Lause 1 põhjal saame

$$m_y(t) = \int_0^t m_x(\tau)\,d\tau = \int_0^t 0 d\tau = 0,$$

$$K_y(t_1, t_2) = \int_0^{t_1}\int_0^{t_2} K_x(\tau_1, \tau_2) d\tau_1 d\tau_2 = \int_0^{t_1}\int_0^{t_2} \frac{1}{1 + (\tau_2 - \tau_1)^2} d\tau_1 d\tau_2 =$$

$$= \int_0^{t_1} \arctan(\tau_2 - \tau_1)\Big|_0^{t_2} d\tau_1 =$$

$$= \int_0^{t_1} \left(\arctan(t_2 - \tau_1) + \arctan\tau_1\right) d\tau_1 =$$

$$= t_1 \arctan t_1 + t_2 \arctan t_2 - (t_1 - t_2)\arctan(t_1 - t_2) - \frac{1}{2}\ln\frac{\left(1 + t_1^2\right)\left(1 + t_2^2\right)}{1 + (t_1 - t_2)^2},$$

$$D_y(t) = K_y(t, t) = 2t\arctan t - \ln\left(1 + t^2\right),$$

$$\sigma_y(t) = \sqrt{2t\arctan t - \ln(1 + t^2)}. \qquad \diamondsuit$$

4.3.4 Juhusliku funktsiooni diferentseerimine

Olgu antud juhuslik funktsioon $X(t)\ (t \in T)$, mille kõik realisatsioonid on diferentseeruvad hulgal T. Suurust

$$Y(t) = \frac{dX(t)}{dt} \overset{\text{def.}}{=} \lim_{\Delta t \to 0} \frac{X(t + \Delta t) - X(t)}{\Delta t}$$

nimetame *juhusliku funktsiooni* $X(t)$ *tuletiseks* hulgal T. Leiame juhusliku funktsiooni $X(t)$ tuletise $Y(t)$ keskväärtuse

$$m_y(t) = \mathrm{E}\left(\lim_{\Delta t\to 0}\frac{X(t+\Delta t)-X(t)}{\Delta t}\right) = \left[\begin{array}{c}\text{kui piirväärtuse ja keskväärtuse}\\ \text{leidmise järjekord on muudetav}\end{array}\right]$$

$$= \lim_{\Delta t\to 0}\mathrm{E}\left(\frac{X(t+\Delta t)-X(t)}{\Delta t}\right) = \lim_{\Delta t\to 0}\frac{\mathrm{E}X(t+\Delta t)-\mathrm{E}X(t)}{\Delta t} =$$

$$= \lim_{\Delta t\to 0}\frac{m_x(t+\Delta t)-m_x(t)}{\Delta t} = \frac{d\,m_x(t)}{dt}.$$

Seega

$$\begin{aligned} Y^\circ(t) &= Y(t) - m_y(t) = \frac{dX(t)}{dt} - \frac{d\,m_x(t)}{dt} = \frac{d\,(X(t)-m_x(t))}{dt} = \frac{d\,X^\circ(t)}{dt} \Rightarrow \\ &\Rightarrow K_y(t_1,t_2) = \mathrm{E}\left((Y^\circ(t_1))\,(Y^\circ(t_2))\right) = \mathrm{E}\left(\frac{\partial\,X^\circ(t_1)}{\partial t_1}\frac{\partial\,X^\circ(t_2)}{\partial t_2}\right) = \\ &= \mathrm{E}\left(\frac{\partial^2}{\partial t_1\partial t_2}\left(X^\circ(t_1)X^\circ(t_2)\right)\right) = \frac{\partial^2}{\partial t_1\partial t_2}\mathrm{E}\left(X^\circ(t_1)X^\circ(t_2)\right) = \\ &= \frac{\partial^2}{\partial t_1\partial t_2}K_x(t_1,t_2), \end{aligned}$$

millest saame omakorda dispersiooni, standardhälbe ja korrelatsiooni

$$D_y(t) = K_y(t,t),\ \sigma_y(t) = \sqrt{D_y(t)},\ R_y(t_1,t_2) = \frac{K_y(t_1,t_2)}{\sigma_y(t_1)\sigma_y(t_2)}.$$

Lause 1. Kui $Y(t) = \dfrac{dX(t)}{dt}$, siis

$$m_y(t) = \frac{d\,m_x(t)}{dt},\ K_y(t_1,t_2) = \frac{\partial^2 K_x(t_1,t_2)}{\partial t_1\partial t_2},\ D_y(t) = K_y(t,t),$$

$$\sigma_y(t) = \sqrt{D_y(t)},\ R_y(t_1,t_2) = \frac{\dfrac{\partial^2 K_x(t_1,t_2)}{\partial t_1\partial t_2}}{\sqrt{\left.\dfrac{\partial^2 K_x(t_1,t_2)}{\partial t_1\partial t_2}\right|_{t_2=t_1}}\sqrt{\left.\dfrac{\partial^2 K_x(t_1,t_2)}{\partial t_1\partial t_2}\right|_{t_1=t_2}}}.$$

Näide 1. Olgu juhusliku funktsiooni $X(t)$ korral $m_x(t) = 1$ ja $K_x(t_1,t_2) = \exp\left(\alpha\,(t_1+t_2)\right)$. Leiame juhusliku funktsiooni $Z(t) = t\dfrac{dX(t)}{dt} + 2$ keskväärtuse, kovariatsiooni, dispersiooni, standardhälbe ja korrelatsiooni.

Leiame samm-sammult. Kui $Y(t) = \dfrac{dX(t)}{dt}$, siis Lause 1 põhjal

$$\begin{aligned}
m_y(t) &= \frac{d\,m_x(t)}{dt} = \frac{d\,1}{dt} = 0, \\
K_y(t_1,t_2) &= \frac{\partial^2}{\partial t_1 \partial t_2} K_x(t_1,t_2) = \frac{\partial^2}{\partial t_1 \partial t_2} \exp\left(\alpha\left(t_1+t_2\right)\right) = \alpha^2 e^{\alpha(t_1+t_2)}, \\
D_y(t) &= K_y(t,t) = \alpha^2 e^{2\alpha t},\ \sigma_y(t) = \sqrt{D_y(t)} = \sqrt{\alpha^2 e^{2\alpha t}} = |\alpha|\, e^{\alpha t}, \\
R_y(t_1,t_2) &= \frac{K_y(t_1,t_2)}{\sigma_y(t_1)\sigma_y(t_2)} = \frac{\alpha^2 e^{\alpha(t_1+t_2)}}{|\alpha|\, e^{\alpha t_1}\, |\alpha|\, e^{\alpha t_2}} = 1.
\end{aligned}$$

Kui $U(t) = tY(t)$, siis Lause 4.3.2.1 põhjal

$$m_u(t) = t \cdot m_y(t) = 0,\ K_u(t_1,t_2) = t_1 t_2 K_y(t_1,t_2) = \alpha^2 t_1 t_2 e^{\alpha(t_1+t_2)}.$$

Et $Z(t) = U(t) + 2$, siis Lause 4.3.1.1 põhjal

$$\begin{aligned}
m_z(t) &= m_u(t) + 2 = 2, \\
K_z(t_1,t_2) &= K_u\left(t_1,t_2\right) + 0 + 0 + 0 = \alpha^2 t_1 t_2 e^{\alpha(t_1+t_2)}, \\
D_z(t) &= K_z(t,t) = \alpha^2 t^2 e^{2\alpha t},\ \sigma_z(t) = \sqrt{D_z(t)} = \sqrt{\alpha^2 t^2 e^{2\alpha t}} = |\alpha t|\, e^{\alpha t}, \\
R_z(t_1,t_2) &= \frac{K_z(t_1,t_2)}{\sigma_z(t_1)\sigma_z(t_2)} = \frac{\alpha^2 t_1 t_2 e^{\alpha(t_1+t_2)}}{|\alpha t_1|\, e^{\alpha t_1}\, |\alpha t_2|\, e^{\alpha t_2}} = \operatorname{sign}\left(t_1 t_2\right). \qquad \diamondsuit
\end{aligned}$$

Kui $U(t) = tY(t)$, siis Lause 4.3.2.1 põhjal

$$m_u(t) = t \cdot m_y(t) = 0,\ K_u(t_1,t_2) = t_1 t_2 K_y(t_1,t_2) = \alpha^2 t_1 t_2 e^{\alpha(t_1+t_2)}.$$

Et $Z(t) = U(t) + 2$, siis Lause 4.3.1.1 põhjal

$$\begin{aligned}
m_z(t) &= m_u(t) + 2 = 2, \\
K_z(t_1,t_2) &= K_u\left(t_1,t_2\right) + 0 + 0 + 0 = \alpha^2 t_1 t_2 e^{\alpha(t_1+t_2)}, \\
D_z(t) &= K_z(t,t) = \alpha^2 t^2 e^{2\alpha t},\ \sigma_z(t) = \sqrt{D_z(t)} = \sqrt{\alpha^2 t^2 e^{2\alpha t}} = |\alpha t|\, e^{\alpha t}, \\
R_z(t_1,t_2) &= \frac{K_z(t_1,t_2)}{\sigma_z(t_1)\sigma_z(t_2)} = \frac{\alpha^2 t_1 t_2 e^{\alpha(t_1+t_2)}}{|\alpha t_1|\, e^{\alpha t_1}\, |\alpha t_2|\, e^{\alpha t_2}} = \operatorname{sign}\left(t_1 t_2\right). \qquad \diamondsuit
\end{aligned}$$

4.4 Juhusliku funktsiooni kanooniline arendus

Järgnevalt vaatleme keerukama juhusliku funktsiooni esitamist lihtsamate juhuslike funktsioonide abil.

Definitsioon 1. Kui juhuslik funktsioon $X(t)$ on esitatud kujul

$$X(t) = m_x(t) + \sum\nolimits_{j=1}^{n} U_j \varphi_j(t), \tag{4.4.1}$$

kus $\varphi_j(t)$ $(j = 1; \ldots; n)$ on kindlad funktsioonid ja suurused U_j $(j = 1; \ldots; n)$ on tsentreeritud mittekorreleeruvad juhuslikud suurused ning $m_x(t)$ on juhusliku funktsiooni $X(t)$ keskväärtusfunktsioon, siis nimetatakse seda kuju juhusliku funktsiooni $X(t)$ *kanooniliseks arenduseks.*

Uurime, milline on seosega (4.4.1) antud juhusliku funktsiooni $X(t)$ kovariatsioon ja dispersioon. Kuna tsentreeritud suuruste U_j korral $U_j^\circ = U_j$ ja $X^\circ(t) = \sum_{j=1}^n U_j^\circ \varphi_j(t)$, siis saame

$$\begin{aligned}
K_x(t_1,t_2) &= \mathrm{E}\left(X^\circ(t_1)X^\circ(t_2)\right) = \\
&= \mathrm{E}\left(\left(\sum_{i=1}^n U_i^\circ \varphi_i(t_1)\right)\left(\sum_{j=1}^n U_j^\circ \varphi_j(t_2)\right)\right) = \\
&= \sum_{i=1}^n \sum_{j=1}^n \varphi_i(t_1)\varphi_j(t_2)\mathrm{E}\left(U_i^\circ U_j^\circ\right) = \\
&= \sum_{i=1}^n \sum_{j=1}^n \varphi_i(t_1)\varphi_j(t_2)\,\mathrm{cov}\,(U_i, U_j) = \\
&= \left[\begin{array}{c} U_i \text{ ja } U_j \text{ on mitte-} \\ \text{korreleeruvad} \end{array} \Rightarrow \mathrm{cov}\,(U_i,U_j) = \delta_{i,j}\,\mathrm{D}U_j\right] = \\
&= \sum_{j=1}^n \varphi_j(t_1)\varphi_j(t_2)\,\mathrm{D}U_j
\end{aligned}$$

ja $D_x(t) = K_x(t,t) = \sum_{j=1}^n \varphi_j^2(t)\,\mathrm{D}U_j$.

Lause 1. Kui (4.4.1) on funktsiooni $X(t)$ kanooniline arendus, siis

$$K_x(t_1,t_2) = \sum_{j=1}^n \varphi_j(t_1)\varphi_j(t_2)\,\mathrm{D}U_j \tag{4.4.2}$$

ja

$$D_x(t) = \sum_{j=1}^n \varphi_j^2(t)\,\mathrm{D}U_j. \tag{4.4.3}$$

Näide 1. Olgu juhuslikud suurused U_j $(j = 1;\ldots;n)$ tsentreeritud ja mittekorreleeruvad, kusjuures $\mathrm{D}U_j = 2^j$. Leiame juhusliku funktsiooni

$$X(t) = \sin t + \sum_{j=1}^n U_j \cos jt$$

kovariatsiooni ja dispersiooni.

Valemite (4.4.2) ja (4.4.3) abil saame

$$K_x(t_1,t_2) = \sum_{j=1}^n 2^j \cos(jt_1)\cos(jt_2),\ D_x(t) = \sum_{j=1}^n 2^j \cos^2(jt)\,. \qquad \diamondsuit$$

4.5 Statsionaarsed juhuslikud funktsioonid

Definitsioon 1. Juhuslikku funktsiooni $X(t)$ $(t \in T)$ nimetatakse *statsionaarseks kitsas mõttes,* kui kõik selle juhusliku funktsiooni mitmedimensionaalsed jaotustihedused on invariantsed muutuja t suvalise võimaliku nihke suhtes, st iga $n \in \mathbf{N}$ ja iga sellise Δ, kus $t_i + \Delta \in T$ $(i = 1;\ldots;n)$, korral

$$f_n(x_1,x_2,\ldots,x_n;t_1+\Delta,t_2+\Delta,\ldots,t_n+\Delta) =$$

$$= f_n(x_1,x_2,\ldots,x_n;t_1,t_2,\ldots,t_n).$$

Kui kitsas mõttes statsionaarse juhusliku funktsiooni $X(t)$ argumendiks on aeg, siis kõneldakse *kitsas mõttes statsionaarsest juhuslikust protsessist.*

Kui juhuslik funktsioon $X(t)$ on kitsas mõttes statsionaarne, siis

$$\begin{aligned} m_x(t+\Delta) &= \int_{-\infty}^{+\infty} x\, f_1(x;t+\Delta)\, dx = \begin{bmatrix} f_1(x;t+\Delta) = \\ = f_1(x;t) \end{bmatrix} = \\ &= \int_{-\infty}^{+\infty} x\, f_1(x;t)\, dx = m_x(t), \\ D_x(t+\Delta) &= \int_{-\infty}^{+\infty} (x - m_x(t+\Delta))^2\ f_1(x;t+\Delta)\, dx = \\ &= \int_{-\infty}^{+\infty} (x - m_x(t))^2\ f_1(x;t)\, dx = D_x(t), \end{aligned}$$

st

$$m_x(t+\Delta) = m_x(t), \quad D_x(t+\Delta) = D_x(t). \tag{4.5.1}$$

Kitsas mõttes statsionaarse juhusliku funktsiooni $X(t)$ korral

$$K_x(t_1+\Delta, t_2+\Delta) =$$

$$\int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} (x_1 - m_x(t_1+\Delta))\,(x_2 - m_x(t_2+\Delta))\, f_2(x_1,x_2;t_1+\Delta,t_2+\Delta) dx_1 dx_2 =$$

$$= [m_x(t_i+\Delta) = m_x(t_i),\ f_2(x_1,x_2;t_1+\Delta,t_2+\Delta) = f_2(x_1,x_2;t_1,t_2)] =$$

$$= \int\limits_{-\infty}^{+\infty}\int\limits_{-\infty}^{+\infty} (x_1 - m_x(t_1))\,(x_2 - m_x(t_2))\, f_2(x_1,x_2;t_1,t_2) dx_1 dx_2 = K_x(t_1,t_2),$$

st

$$K_x(t_1+\Delta, t_2+\Delta) = K_x(t_1,t_2). \tag{4.5.2}$$

Lause 1. Kitsas mõttes statsionaarse juhusliku funktsiooni $X(t)$ keskväärtus $m_x(t)$ ja dispersioon $D_x(t)$ on konstantsed funktsioonid ning kovariatsioon $K_x(t_1,t_2)$ sõltub vaid argumentide t_1 ja t_2 vahest $t_2 - t_1$, st

$$K_x(t_1,t_2) = k_x(t_2 - t_1) \tag{4.5.3}$$

ning

$$D_x(t) = k_x(0), \tag{4.5.4}$$

kus $k_x(\tau) \overset{\text{def.}}{=} K_x(0,\tau) \ \ (\tau = t_2 - t_1)\,.$

Tõestus. Seosest (4.5.1) järeldub, et funktsioonid $m_x(t)$ ja $D_x(t)$ on konstantsed. Kuna seos (4.5.2) peab kehtima suvalise (võimaliku) Δ korral, siis kehtib ta ka $\Delta = -t_1$ korral. Seega

$$K_x(t_1 - t_1, t_2 - t_1) = K_x(0, t_2 - t_1) = k_x(t_2 - t_1) = k_x(\tau). \qquad \square$$

Definitsioon 2. Juhuslikku funktsiooni $X(t)$ $(t \in T)$ nimetatakse *statsionaarseks (laias mõttes),* kui selle juhusliku funktsiooni keskväärtus on konstantne funktsioon ja kovariatsioon $K_x(t_1, t_2)$ sõltub vaid argumentide vahest $t_2 - t_1$.

Lause 2. Iga kitsas mõttes statsionaarne juhuslik funktsioon on (laias mõttes) statsionaarne, kuid mitte vastupidi.

Lause 3. Statsionaarse juhusliku funktsiooni $X(t)$ kovariatsioon $k_x(\tau)$ on paarisfunktsioon, kusjuures

$$|k_x(\tau)| \leq k_x(0) = D_x(t).$$

Tõestus. Väite esimesele osale sobib järgmine arutelu

$$X(t_1)X(t_2) = X(t_2)X(t_1) \;\Rightarrow\; X^\circ(t_1)X^\circ(t_2) = X^\circ(t_2)X^\circ(t_1) \;\Rightarrow$$

$$\Rightarrow\; \mathrm{E}\,(X^\circ(t_1)X^\circ(t_2)) = \mathrm{E}\,(X^\circ(t_2)X^\circ(t_1)) \;\Leftrightarrow$$

$$\Leftrightarrow\; K_x(t_1, t_2) = K_x(t_2, t_1) \;\overset{(4.5.3)}{\Leftrightarrow}$$

$$\Leftrightarrow\; k_x(t_2 - t_1) = k_x(t_1 - t_2) \;\Leftrightarrow\; k_x(-\tau) = k_x(\tau)$$

ja teisele osale sobib arutelu

$$|R_x\,(t_1, t_2)| \leq 1 \;\Leftrightarrow\; |\mathrm{cov}\,(X(t_1), X(t_2))| \leq \sigma_x(t_1) \cdot \sigma_x(t_2) \;\Leftrightarrow$$

$$\Leftrightarrow\; |K_x(t_1, t_2)| \leq \sqrt{D_x\,(t_1)\,D_x\,(t_2)} \;\Rightarrow$$

$$\Rightarrow\; [X(t) \text{ on statsionaarne } \;\Rightarrow\; D_x\,(t) \text{ on konstantne}] \;\Rightarrow$$

$$\Rightarrow\; |K_x(t_1, t_2)| \leq \sqrt{D_x\,(t)\,D_x\,(t)} = D_x\,(t) \;\overset{(4.5.3)}{\Rightarrow}$$

$$\Rightarrow\; |k_x(t_2 - t_1)| \leq D_x\,(t) \overset{(4.5.4)}{=} k_x(0) \;\Rightarrow\; |k_x(\tau)| \leq k_x\,(0) = D_x\,(t)\,. \qquad \square$$

Näide 1. Olgu

$$X(t) = t + U\cos\omega t + V\sin\omega t,$$

kusjuures U ja V on tsentreeritud mittekorreleeruvad juhuslikud suurused, $\mathrm{D}U = \mathrm{D}V$ ja ω on kindel suurus. Uurime juhuslike funktsioonide $X(t)$ ja $X^\circ(t)$ statsionaarsust.

Juhuslike suuruste U ja V tsentreeritusest järeldub

$$U^\circ = U, \quad V^\circ = V.$$

Kuna

$$\begin{aligned} m_x(t) = \mathrm{E}X(t) &= \mathrm{E}\,(t + U\cos\omega t + V\sin\omega t) = \\ &= \mathrm{E}\,t + (\cos\omega t)\;\mathrm{E}U + (\sin\omega t\,)\,\mathrm{E}V = t, \end{aligned}$$

siis $m_x(t) \neq const$ ja Definitsiooni 2 põhjal $X(t)$ ei ole statsionaarne. Kuna

$$\begin{aligned} X^{\circ}(t) &= X(t) - m_x(t) = t + U\cos\omega t + V\sin\omega t - t = \\ &= U\cos\omega t + V\sin\omega t, \end{aligned}$$

siis

$$m_{X^{\circ}}(t) = \mathrm{E}X^{\circ}(t) = \mathrm{E}\left(U\cos\omega t + V\sin\omega t\right) = 0.$$

Seega funktsiooni $X^{\circ}(t)$ keskväärtus on konstantne. Et juhuslikud suurused U ja V on tsentreeritud ja mittekorreleeruvad, siis $\mathrm{E}\left(U^{\circ}V^{\circ}\right) = 0$ ja

$$K_{X^{\circ}}(t_1, t_2) = \mathrm{E}\left(\left(X^{\circ}(t_1) - m_{x^{\circ}}(t_1)\right)\left(X^{\circ}(t_2) - m_{x^{\circ}}(t_2)\right)\right) =$$

$$= \mathrm{E}\left(\left(U^{\circ}\cos\omega t_1 + V^{\circ}\sin\omega t_1\right)\left(U^{\circ}\cos\omega t_2 + V^{\circ}\sin\omega t_2\right)\right) =$$

$$= \left(\cos\omega t_1\cos\omega t_2\right)\mathrm{E}\left(U^{\circ}\right)^2 + \left(\cos\omega t_1\sin\omega t_2\right)\mathrm{E}\left(U^{\circ}V^{\circ}\right) +$$

$$+ \left(\sin\omega t_1\cos\omega t_2\right)\mathrm{E}\left(V^{\circ}U^{\circ}\right) + \left(\sin\omega t_1\sin\omega t_2\right)\mathrm{E}\left(V^{\circ}\right)^2 =$$

$$= \left(\cos\omega t_1\cos\omega t_2\right)\mathrm{D}U + \left(\sin\omega t_1\sin\omega t_2\right)\mathrm{D}V =$$

$$= \left(\cos\omega t_1\cos\omega t_2 + \sin\omega t_1\sin\omega t_2\right)\mathrm{D}U = \mathrm{D}U\cos\left(\omega\left(t_2 - t_1\right)\right)$$

ning $k_{x^{\circ}}(\tau) = \mathrm{D}U\cos(\omega\tau)$. Seega on juhuslik funktsioon $X^{\circ}(t)$ statsionaarne (laias mõttes). $\diamondsuit$

Näide 2. Olgu $X(t) = \sin(t + \Phi)$, kus Φ on juhuslik suurus, mis allub ühtlasele jaotusele vahemikus $(0; 2\pi)$. Kas $X(t)$ on statsionaarne?

Kuna

$$\begin{aligned} m_x(t) &= \mathrm{E}\sin(t + \Phi) = \int_0^{2\pi} \frac{1}{2\pi}\sin(t + \varphi)\,d\varphi = -\frac{1}{2\pi}\cos(t + \varphi)\Big|_0^{2\pi} = \\ &= -\frac{1}{2\pi}\cos(t + 2\pi) + \frac{1}{2\pi}\cos(t) = 0 \end{aligned}$$

ja

$$K_x(t_1, t_2) = \mathrm{E}\left(X^{\circ}(t_1)X^{\circ}(t_2)\right) = \mathrm{E}\left(\sin(t_1 + \Phi)\sin(t_2 + \Phi)\right) =$$

$$= \mathrm{E}\left(\left(\sin t_1\cos\Phi + \cos t_1\sin\Phi\right)\left(\sin t_2\cos\Phi + \cos t_2\sin\Phi\right)\right) =$$

$$= \sin t_1\sin t_2\mathrm{E}\left(\cos^2\Phi\right) + \sin t_1\cos t_2\mathrm{E}\left(\cos\Phi\sin\Phi\right) +$$

$$+ \cos t_1\sin t_2\mathrm{E}\left(\cos\Phi\sin\Phi\right) + \cos t_1\cos t_2\mathrm{E}\left(\sin^2\Phi\right) = 0.5\cos\left(t_2 - t_1\right),$$

sest

$$\mathrm{E}\left(\cos^2\Phi\right) = \int_0^{2\pi} \frac{\cos^2\varphi}{2\pi} d\varphi = \int_0^{2\pi} \frac{1+\cos 2\varphi}{4\pi} d\varphi = \frac{1}{2},$$
$$\mathrm{E}\left(\cos\Phi\sin\Phi\right) = \int_0^{2\pi} \frac{\cos\varphi\sin\varphi}{2\pi} d\varphi = \int_0^{2\pi} \frac{\sin 2\varphi}{4\pi} d\varphi = 0,$$
$$\mathrm{E}\left(\sin^2\Phi\right) = \int_0^{2\pi} \frac{\sin^2\varphi}{2\pi} d\varphi = \int_0^{2\pi} \frac{1-\cos 2\varphi}{4\pi} d\varphi = \frac{1}{2}.$$

Seega on $m_x(t)$ konstantne ja $K_x(t_1, t_2)$ sõltub vaid argumentide vahest $t_2 - t_1$, st $X(t)$ on statsionaarne. $\diamondsuit$

Näidake, et statsionaarse juhusliku funktsiooni korrelatsioon $r_x(\tau)$ avaldub kujul

$$r_x(\tau) = k_x(\tau)/k_x(0).$$

Definitsioon 3. Öeldakse, et juhuslikud funktsioonid $X(t)$ ja $Y(t)$ $(t \in T)$ on *statsionaarselt seotud,* kui nende juhuslike funktsioonide vastastikune kovariatsioon $K_{xy}(t_1, t_2)$ sõltub vaid argumentide t_2 ja t_1 vahest $t_2 - t_1$.

Olgu statsionaarselt seotud funktsioonide $X(t)$ ja $Y(t)$ korral

$$k_{xy}(t_2 - t_1) \overset{\text{def.}}{=} K_{xy}(t_1, t_2). \tag{4.5.5}$$

Näide 3. Olgu U, V, W tsentreeritud juhuslikud suurused ja

$$X(t) = U\sin t + V\cos t, \quad Y(t) = V\sin t + W\cos t$$

ning

$$\begin{pmatrix} 16 & 4 & -1 \\ & 1 & 4 \\ & & 25 \end{pmatrix}$$

juhusliku vektori (U, V, W) kovariatsioonimaatriks. Kas $X(t)$ ja $Y(t)$ on statsionaarselt seotud?

Vektori (U, V, W) kovariatsioonimaatriksi põhjal

$$\mathrm{D}V = 1, \ \operatorname{cov}(U, V) = 4, \ \operatorname{cov}(U, W) = -1, \ \operatorname{cov}(V, W) = 4.$$

Kuna

$$K_{xy}(t_1, t_2) = \mathrm{E}\left(X^\circ(t_1)Y^\circ(t_2)\right) = [U = U^\circ,\ V = V^\circ,\ W = W^\circ] =$$

$$= \mathrm{E}\left(\left(U^\circ\sin t_1 + V^\circ\cos t_1\right)\left(V^\circ\sin t_2 + W^\circ\cos t_2\right)\right) =$$

$$= \sin t_1 \sin t_2 \mathrm{E}\left(U^\circ V^\circ\right) + \sin t_1 \cos t_2 \mathrm{E}\left(U^\circ W^\circ\right) +$$

$$+ \cos t_1 \sin t_2 \mathrm{E}\left(V^\circ V^\circ\right) + \cos t_1 \cos t_2 \mathrm{E}\left(V^\circ W^\circ\right) =$$

$$= \sin t_1 \sin t_2 \mathrm{cov}\,(U, V) + \sin t_1 \cos t_2 \mathrm{cov}\,(U, W) +$$

$$+ \cos t_1 \sin t_2 \mathrm{D}V + \cos t_1 \cos t_2 \mathrm{cov}\,(V, W) =$$

$$= 4 \sin t_1 \sin t_2 - \sin t_1 \cos t_2 + \cos t_1 \sin t_2 + 4 \cos t_1 \cos t_2 =$$

$$= 4 \cos (t_2 - t_1) + \sin (t_2 - t_1)\,,$$

siis funktsioonid $X(t)$ ja $Y(t)$ on statsionaarselt seotud ja seose (4.5.5) põhjal $k_{xy}(\tau) = 4 \cos \tau + \sin \tau.$ $\diamondsuit$

4.6 Lõplikus vahemikus statsionaarse funktsiooni spektraalarendus

Olgu juhuslik funktsioon $X(t)$ $(T = (-l, l)\,,\ l \neq \infty)$ statsionaarne. Et

$$t_1, t_2 \in (-l, l) \ \Rightarrow\ \tau = t_2 - t_1 \in (-2l, 2l)\,,$$

siis selle statsionaarse juhusliku funktsiooni kovariatsioon $k_x\,(\tau)$ on määratud vahemikus $(-2l, 2l)$ ja on selles vahemikus Lause 4.5.2 põhjal paarisfunktsioon. Kui eeldada, et $k_x\,(\tau)$ on arendatav vahemikus $(-2l, 2l)$ Fourier' ritta $4l$-perioodilise trigonomeetrilise süsteemi järgi, siis on see rida koosinusrida

$$k_x\,(\tau) \sim \sum_{j=0}^{\infty} D_j \cos \frac{j\pi\tau}{2l},$$

kus

$$D_0 = \frac{1}{2l} \int_0^{2l} k_x\,(\tau)\, d\tau,\ \ D_j = \frac{1}{l} \int_0^{2l} k_x\,(\tau) \cos \frac{j\pi\tau}{2l}\, d\tau \quad (j = 1; 2; 3; \ldots)\,.$$

Kui tähistada $\omega_j = (j\pi)\,/\,(2l)$ ja eeldada, et saadud Fourier' rida koondub vahemikus $(-2l, 2l)$ funktsiooniks $k_x\,(\tau)\,,$ siis

$$k_x\,(\tau) = \sum_{j=0}^{\infty} D_j \cos (\omega_j \tau)\,, \tag{4.6.1}$$

kus

$$D_0 = \frac{1}{2l} \int_0^{2l} k_x\,(\tau)\, d\tau,\ \ D_j = \frac{1}{l} \int_0^{2l} k_x\,(\tau) \cos (\omega_j \tau)\ d\tau \quad (j = 1; 2; 3; \ldots)\,. \tag{4.6.2}$$

Järgnevalt piirdume juhuga $D_j \geq 0$ $\ (j \in \mathbf{N}_0)\,.$ Seostest (4.5.4) ja (4.6.1) järeldub, et vahemikus $(-l, l)$ statsionaarse juhusliku funktsiooni $X(t)$ korral

$$D_x(t) = \sum_{j=0}^{\infty} D_j$$

ja

$$K_x(t_1,t_2)=k_x\left(t_2-t_1\right)=\sum_{j=0}^{\infty}D_j\cos\left(\omega_j\left(t_2-t_1\right)\right)=$$
$$=\sum_{j=0}^{\infty}D_j\cos\left(\omega_j t_2-\omega_j t_1\right)=$$
$$=\sum_{j=0}^{\infty}D_j\left(\cos\left(\omega_j t_2\right)\cos\left(\omega_j t_1\right)+\sin\left(\omega_j t_2\right)\sin\left(\omega_j t_1\right)\right).$$

Selline kovariatsioon on statsionaarsel juhuslikul funktsioonil

$$X(t)=m_x(t)+\sum_{j=0}^{\infty}\left(U_j\cos\left(\omega_j t\right)+V_j\sin\left(\omega_j t\right)\right), \tag{4.6.3}$$

kus U_j ja V_j on tsentreeritud mittekorreleeruvad juhuslikud suurused ning $\mathrm{D}U_j=\mathrm{D}V_j=D_j$. Kontrollige! Viimasest tingimusest selgub, miks funktsiooni $k_x\left(\tau\right)$ arenduses Fourier' ritta (4.6.1) on kasutatud kordajate tähistust D_j. Vahemikus $(-l;l)$ statsionaarse juhusliku funktsiooni $X(t)$ esitust kujul (4.6.3) nimetatakse funktsiooni $X(t)$ *spektraalarenduseks vahemikus* $(-l;l)$. Sõnastame saadud tulemuse.

Lause 1. Kui on teada vahemikus $(-l,l)$ statsionaarse juhusliku funktsiooni kovariatsioon $k_x(\tau)$, mille Fourier' arendus

$$k_x\left(\tau\right)=\sum_{j=0}^{\infty}D_j\cos\left(\omega_j\tau\right)$$

koondub funktsiooniks $k_x\left(\tau\right)$ vahemikus $\left(-2l,2l\right)$, kusjuures $\omega_j=\left(j\pi\right)/\left(2l\right)$ ja $D_j\geq 0$, siis avaldub $X(t)$ kujul (4.6.3). Seejuures on U_j ning V_j tsentreeritud mittekorreleeruvad juhuslikud suurused ning kordajad $D_j=\mathrm{D}V_j=\mathrm{D}U_j$ $(j=0;1;2;\ldots)$ on leitavad valemite (4.6.2) abil.

Näide 1. Leiame vahemikus $(-1;1)$ statsionaarse juhusliku funktsiooni $X(t)$ spektraalarenduse, kui on antud funktsiooni $X(t)$ kovariatsioon

$$k_x(\tau)=\begin{cases}\left(|\tau|-2\right)^2, & \text{kui } |\tau|<2,\\ 0, & \text{kui } |\tau|\geq 2.\end{cases}$$

Selle spektraalarenduse leidmiseks kasutame Lauset 1. Et antud näite korral $l=1$, siis $\omega_j=\left(j\pi\right)/2$ ja valemite (4.6.2) abil saame

$$D_0=\frac{1}{2}\int_0^2\left(|\tau|-2\right)^2d\tau=\frac{1}{2}\int_0^2\left(\tau-2\right)^2d\tau=4/3>0,$$

$$D_j = \frac{1}{1} \int_0^2 (|\tau| - 2)^2 \cos(j\pi\tau/2)\ d\tau = \int_0^2 (\tau - 2)^2 \cos(j\pi\tau/2)\ d\tau =$$

$$= \left[\begin{array}{cc} u = (\tau - 2)^2 & du = 2(\tau - 2)\, d\tau \\ dv = \cos(j\pi\tau/2)\ d\tau & v = \dfrac{2}{j\pi} \sin(j\pi\tau/2) \end{array} \right] =$$

$$= (\tau - 2)^2 \frac{2}{j\pi} \sin(j\pi\tau/2) \bigg|_0^2 - \int_0^2 \frac{4(\tau - 2)}{j\pi} \sin(j\pi\tau/2)\, d\tau =$$

$$= \left[\begin{array}{cc} u = \dfrac{4(\tau - 2)}{j\pi} & du = \dfrac{4}{j\pi} d\tau \\ dv = -\sin(j\pi\tau/2)\ d\tau & v = \dfrac{2}{j\pi} \cos(j\pi\tau/2) \end{array} \right] =$$

$$= \frac{8(\tau - 2)}{j^2\pi^2} \cos(j\pi\tau/2) \bigg|_0^2 - \int_0^2 \frac{8}{j^2\pi^2} \cos(j\pi\tau/2)\, d\tau =$$

$$= \frac{16}{j^2\pi^2} > 0 \quad (j = 1; 2; 3; \ldots).$$

Seega

$$X(t) = m_x(t) + \sum_{j=0}^{\infty} \left(U_j \cos \frac{j\pi t}{2} + V_j \sin \frac{j\pi t}{2} \right),$$

kus U_j ja V_j on tsentreeritud mittekorreleeruvad juhuslikud suurused ning $\mathrm{D}_0 = 4/3$, $\mathrm{D}V_j = \mathrm{D}U_j = 16/\left(j^2\pi^2\right)$ $(j = 1; 2; 3; \ldots)$. ♢

4.7 Lõpmatus vahemikus statsionaarse juhusliku funktsiooni spektraalarendus

Definitsioon 1. Kindlat funktsiooni

$$s_x(\omega) \overset{\text{def.}}{=} \frac{1}{2\pi} \int_{-\infty}^{\infty} k_x(\tau) \exp(-i\omega\tau)\, d\tau \tag{4.7.1}$$

nimetatakse statsionaarse juhusliku funktsiooni $X(t)$, mille kovariatsioon on $k_x(\tau)$, *spektraaltiheduseks.*

Et statsionaarse juhusliku funktsiooni $X(t)$ spektraaltihedus $s_x(\omega)$ on defineeritud kui kovariatsiooni $k_x(\tau)$ Fourier' pöördteisendus, siis kovariatsioon $k_x(\tau)$ on spektraaltiheduse $s_x(\omega)$ Fourier' teisendus, st

$$k_x(\tau) = \int_{-\infty}^{\infty} s_x(\omega) \exp(i\omega\tau)\, d\omega. \tag{4.7.2}$$

Näide 1. Olgu

$$k_x(\tau) = \begin{cases} 1-\tau^2, \text{ kui } |\tau| \le 1, \\ 0, \text{ kui } |\tau| > 1 \end{cases}$$

statsionaarse juhusliku funktsiooni $X(t)$ kovariatsioon. Leiame spektraaltiheduse $s_x(\omega)$.

Valemi (4.7.1) abil saame

$$\begin{aligned} s_x(\omega) &= \frac{1}{2\pi}\int_{-1}^{1}(1-\tau^2)\exp(-i\omega\tau)\,d\tau = \\ &= \frac{1}{2\pi}\int_{-1}^{1}(1-\tau^2)(\cos(\omega\tau) - i\sin(\omega\tau))\,d\tau = \\ &= \frac{1}{\pi}\int_{0}^{1}(1-\tau^2)\cos(\omega\tau)\,d\tau = \\ &= \left[\begin{array}{c} u = 1-\tau^2,\ du = -2\tau d\tau \\ dv = \cos(\omega\tau)\,d\tau, v = \dfrac{\sin(\omega\tau)}{\omega} \end{array}\right] = \\ &= \frac{1}{\pi}\left(\left(1-\tau^2\right)\frac{\sin(\omega\tau)}{\omega}\Big|_0^1 + \int_0^1 \frac{\sin(\omega\tau)}{\omega}2\tau d\tau\right) = \\ &= \left[\begin{array}{c} u = 2\tau,\ du = 2d\tau \\ dv = \dfrac{\sin(\omega\tau)}{\omega}d\tau, v = -\dfrac{\cos(\omega\tau)}{\omega^2} \end{array}\right] = \\ &= \frac{1}{\pi}\left(-\frac{\cos(\omega\tau)}{\omega^2}2\tau\Big|_0^1 + \int_0^1 \frac{2\cos(\omega\tau)}{\omega^2}d\tau\right) = \\ &= \frac{1}{\pi}\left(-\frac{2\cos\omega}{\omega^2} + \frac{2\sin(\omega\tau)}{\omega^3}\Big|_0^1\right) = \\ &= \frac{2}{\pi\omega^3}(\sin\omega - \omega\cos\omega). \qquad \diamondsuit \end{aligned}$$

Saab näidata, et koosinusteisenduse abil on $s_x(\omega)$ ja $k_x(\tau)$ seotud järgnevalt

$$s_x(\omega) = \frac{1}{\pi}\int_0^\infty k_x(\tau)\cos(\omega\tau)\,d\tau, \tag{4.7.3}$$

$$k_x(\tau) = 2\int_0^\infty s_x(\omega)\cos(\omega\tau)\,d\omega. \tag{4.7.4}$$

Definitsioon 2. Kindlat funktsiooni

$$s_{xy}(\omega) \overset{\text{def.}}{=} \frac{1}{2\pi}\int_{-\infty}^{\infty} k_{xy}(\tau)\exp(-i\omega\tau)\,d\tau \tag{4.7.5}$$

nimetatakse statsionaarsete ja statsionaarselt seotud juhuslike funktsioonide $X(t)$ ja $Y(t)$, mille vastastikune kovariatsioon on $k_{xy}(\tau)$, *vastastikuseks spektraaltiheduseks.*

Seega

$$k_{xy}(\tau) = \int_{-\infty}^{\infty} s_{xy}(\tau) \exp(i\omega\tau)\, d\omega. \tag{4.7.6}$$

Formaalselt kehtivad δ-funktsiooni $\delta(\tau)$ ja selle Fourier' teisendi $\widehat{\delta}(\omega)$ korral seosed

$$\widehat{\delta}(\omega) = \int_{-\infty}^{\infty} \delta(\tau) \exp(i\omega\tau)\, d\tau = \exp(i\omega 0) = 1 \quad (\omega \in \mathbf{R})$$

ja

$$\delta(\tau) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \widehat{\delta}(\omega) \cdot \exp(-i\omega\tau)\, d\omega = \frac{1}{2\pi}\int_{-\infty}^{\infty} \exp(-i\omega\tau)\, d\omega =$$
$$= \left[\begin{array}{c} \text{teostame muutujate vahetused} \\ \omega = -\rho \text{ ja siis } \rho = \omega \end{array} \right] = \frac{1}{2\pi}\int_{-\infty}^{\infty} \exp(i\omega\tau)\, d\omega.$$

Seega

$$\int_{-\infty}^{\infty} \exp(i\omega\tau)\, d\omega = 2\pi\delta(\tau). \tag{4.7.7}$$

Definitsioon 3. Statsionaarset juhuslikku funktsiooni $X(t)$, mille spektraaltihedus $s_x(\omega)$ on konstantne, nimetatakse *statsionaarseks valgeks müraks.*

Statsionaarse valge müra $X(t)$ korral, st juhul

$$s_x(\omega) = s_0 = const,$$

saame valemitest (4.7.2) ja (4.7.7), et

$$k_x(\tau) = \int_{-\infty}^{\infty} s_x(\omega) \exp(i\omega\tau)\, d\omega = s_0 \int_{-\infty}^{\infty} \exp(i\omega\tau)\, d\omega = 2\pi s_0 \delta(\tau).$$

Lause 1. Spektraaltihedusega s_0 statsionaarse valge müra $X(t)$ korral kehtib seos

$$k_x(\tau) = 2\pi s_0 \delta(\tau). \tag{4.7.8}$$

Üldistame statsionaarse valge müra mõistet.

Definitsioon 4. Juhuslikku funktsiooni $X(t)$, mille kovariatsioon on kujul

$$K_x(t_1, t_2) = \sqrt{h(t_1)\, h(t_2)}\,\delta(t_2 - t_1), \tag{4.7.9}$$

nimetatakse *valgeks müraks,* kusjuures funktsiooni $h(t)$ nimetatakse *valge müra intensiivsuseks.*

4.8 Juhuslikud jadad ja Markovi ahelad

Juhuslik jada $\{X_n\}$, kus X_n $(n \in \mathbf{N})$ on juhuslikud suurused, on juhusliku funktsiooni $X(t)$ $(t \in T)$ erijuht, kus hulk T on loenduv, $T = \{t_n\}$, ja $X_n \overset{\text{def.}}{=} X(t_n)$. Juhusliku jada $\{X_k\}$ keskväärtus $m_x(n) \overset{\text{def.}}{=} \mathrm{E}X_n$ ja kovariatsioon $K_x(n,m) \overset{\text{def.}}{=} \mathrm{E}(X_n^o X_m^o)$ on kindlad jadad. Kui jada $\{X_n\}$ liikmed on sõltumatud, siis $K(n,m) = \delta_{n,m}\,\mathrm{D}X_n$.

Definitsioon 1. Juhuslikku jada $\{X_n\}$ nimetatakse *lihtsaks Markovi ahelaks,* kui selle jada iga elemendi X_n $(n > 1)$ tinglik jaotus sellele elemendile eelnevate elementide suhtes sõltub vaid elemendist X_{n-1}.

Järeldus 1. Jada $\{X_n\}$ on lihtne Markovi ahel parajasti siis, kui elemendi X_n tinglik jaotustihedus rahuldab seost

$$f_n(x_n \mid x_1, \ldots, x_{n-1}) = f_n(x_n \mid x_{n-1}) \quad (n = 2; 3; \ldots), \tag{4.8.1}$$

kus $f_n(x_n \mid x_1, \ldots, x_{n-1})$ on elemendi X_n tinglik jaotustihedus eeldusel, et leidsid aset sündmused $X_k = x_k$ $(k = 1; \ldots; n-1)$.

Näide 1. Vaatleme ühesugustel tingimustel sooritatud sõltumatute katsete jada, mille igal katsel sündmus A toimub ülimalt üks kord. Olgu p sündmuse A toimumise tõenäosus ja X_n sündmuse A toimumiste arv n-ndal katsel. Leiame selle jada keskväärtuse ja kovariatsiooni. Kas see jada on lihtne Markovi ahel?

Saame $m_x(n) = \mathrm{E}X_n = p$. Kuna katsed jadas on sõltumatud, siis on sõltumatud ka juhuslikud suurused X_n ja X_m $(n \neq m)$. Seega $K_x(n,m) = \operatorname{cov}(X_n, X_m) = 0$ $(n \neq m)$. Iga sõltumatute elementidega jada on lihtne Markovi ahel. Tingimus (4.8.1) omandab kuju

$$f_n(x_n \mid x_1, \ldots, x_{n-1}) = f_n(x_n \mid x_{n-1}) = (1-p)\,\delta(x_n - 0) + p\delta(x_n - 1). \quad \diamondsuit$$

Näide 2 (vt [32]). Kui juhuslik jada $\{X_n\}$ allub normaaljaotusele (vt Definitsiooni 4.1.4) ja $K_x(n,m) = Dq^{|n-m|}$, siis on see jada lihtne Markovi ahel.

Definitsioon 2. Markovi ahelat $\{X_n\}$ nimetatakse *diskreetsete seisunditega ahelaks*, kui selle ahela iga elemendi X_n võimalike väärtuste hulk on lõplik või loenduv.

Definitsioon 3. Suurust

$$p_{i,j}(k) \overset{\text{def.}}{=} P(X_{k+1} = s_j \mid X_k = s_i), \tag{4.8.2}$$

kus s_j on diskreetsete seisunditega ahela elemendi X_{k+1} *võimalik seisund* (väärtus) ja s_i elemendi X_k võimalik seisund, nimetatakse diskreetsete seisunditega Markovi ahela $\{X_n\}$ *üleminekutõenäosuseks.*

Definitsioon 4. Diskreetsete seisunditega Markovi ahelat $\{X_n\}$ nimetatakse *homogeenseks,* kui üleminekutõenäosus $p_{i,j}(k)$ ei sõltu arvust k, vaid ainult arvudest i ja j.

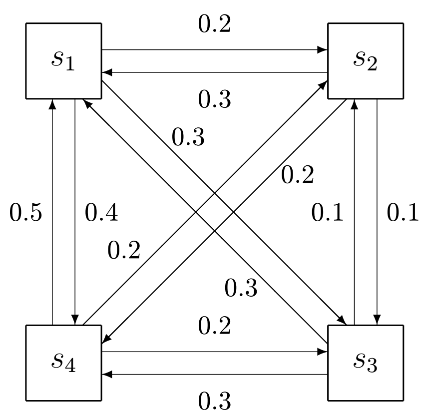
Seega sobib $p_{i,j}$ homogeense diskreetsete seisunditega Markovi ahela üleminekutõenäosuse tähistuseks.

Definitsioon 5. Maatriksit $(p_{i,\,j})$ nimetatakse homogeense diskreetsete seisunditega Markovi ahela *üleminekumaatriksiks.*

Veenduge, et homogeense diskreetsete seisunditega Markovi ahela üleminekumaatriksi korral $\sum_j p_{i,\,j} = 1$.

Näide 3. Olgu diskreetsete seisunditega homogeensel Markovi ahelal neli seisundit s_1, s_2, s_3 ja s_4, kusjuures üleminekutõenäosused ühest seisundist teise on antud skeemil ja $p_{i,i} = 1 - \sum_{j=1,\, j\neq i}^{4} p_{i,j}$ $(i = 1; 2; 3; 4)$. Leiame üleminekumaatriksi.

Skeemi põhjal koostame maatriksi



$$\begin{pmatrix} 0.1 & 0.2 & 0.3 & 0.4 \\ 0.3 & 0.4 & 0.1 & 0.2 \\ 0.3 & 0.1 & 0.3 & 0.3 \\ 0.5 & 0.2 & 0.2 & 0.1 \end{pmatrix}$$

♢

4.9 Ülesanded

1. Juhuslikku suurust V jaotustihedusega $f(v)$ vaadeldakse kui juhuslikku funktsiooni $V(t)$, st $V(t) = V$. Leidke funktsiooni $V(t)$ korral: 1) ühemõõtmeline jaotusfunktsioon $F_1(v;t)$ ja jaotustihedus $f_1(v;t)$; 2) keskväärtus $m_v(t)$ ja dispersioon $D_v(t)$; 3) kahedimensionaalne jaotusfunktsioon $F_2(v_1, v_2; t_1, t_2)$ ja kovariatsioon $K_v\,(t_1, t_2)$. V: $F_1(v;t) = \int_{-\infty}^{v} f(v)dv$, $f_1(v;t) = f(v)$, $m_v(t) = \mathrm{E}V$, $\mathrm{D}_v(t) = \mathrm{D}V$, $F_2(v_1, v_2; t_1, t_2) = \int_{-\infty}^{\min\{v_1,v_2\}} f(v)dv$, $K_v\,(t_1, t_2) = \mathrm{D}V$.

2. Olgu $X(t) = U\cos t \quad (t \in [0; \pi/3])$, kus juhuslik suurus U allub ühtlasele jaotusele lõigul $[a, b]$. Leidke $F_1(x;t)$, $f_1(x;t)$, $m_x(t)$, $K_x(t_1, t_2)$, $D_x(t)$, $\sigma_x(t)$ ja $R_x(t_1, t_2)$. V: $F_1(x;t) = \begin{cases} 0, \text{ kui } x < a\cos t, \\ \dfrac{x - a\cos t}{(b-a)\cos t}, \text{ kui } a\cos t \le x \le b\cos t, \\ 1, \text{ kui } x > b\cos t, \end{cases}$

$f_1(x;t) = \begin{cases} \dfrac{1}{(b-a)\cos t}, \text{ kui } a\cos t \le x \le b\cos t, \\ 0, \text{ kui } x < a\cos t \ \vee \ x > b\cos t, \end{cases}$ $\quad m_x(t) = \dfrac{(a+b)\cos t}{2}$,

$K_x(t_1, t_2) = \dfrac{(b-a)^2}{12}\cos t_1 \cos t_2$, $D_x(t) = \dfrac{(b-a)^2}{12}\cos^2 t$, $\sigma_x(t) = \dfrac{b-a}{2\sqrt{3}}\cos t$, $R_x(t_1, t_2) = 1$.

3. Olgu $X(t) = t^3 U + a$, kus a on kindel suurus ja $f(u) = (u/2)\left(\mathbf{1}(u) - \mathbf{1}(u-2)\right)$ on juhusliku suuruse U jaotustihedus ning $t > 0$. Leidke $F_1(x;t)$, $f_1(x;t)$, $m_x(t)$, $K_x(t_1,t_2)$, $D_x(t)$, $\sigma_x(t)$ ja $R_x(t_1,t_2)$.
V: $F_1(x;t) = \begin{cases} 0, \text{ kui } x < a, \\ (x-a)^2 / \left(4t^6\right), \text{ kui } a \le x \le a + 2t^3, \\ 1, \text{ kui } x > a + 2t^3, \end{cases} \quad m_x(t) = \dfrac{4t^3}{3} + a,$
$f_1(x;t) = \begin{cases} 0, \text{ kui } x < a \vee x > a + t^3, \\ (x-a) / \left(2t^6\right), \text{ kui } a \le x \le a + t^3, \end{cases} \quad K_x(t_1,t_2) = \dfrac{2}{9} t_1^3 t_2^3,$
$D_x(t) = \dfrac{2t^6}{9}, \ \sigma_x(t) = \dfrac{\sqrt{2} t^3}{3}, \ R_x(t_1,t_2) = 1.$

4. Olgu $X(t) = \exp(-|tU|)$, kus U on juhuslik suurus. Leidke selle juhusliku suuruse realisatsioon $x(t)$ juhul, kui U omandab katse käigus väärtuse $1/4$.
V: $x(t) = \exp(-|t|/4)$.

5. Olgu $X(t) = U \cos t$ ja $Y(t) = U \sin t$, kus U on juhuslik suurus. Leidke juhuslike funktsioonide $X(t)$ ja $Y(t)$ vastastikune korrelatsioon $R_{xy}(t_1,t_2)$.
V: $R_{xy}(t_1,t_2) = \operatorname{sign}(\cos t_1) \cdot \operatorname{sign}(\sin t_2)$.

6. On antud $X(t)$ kovariatsioon $K_x(t_1,t_2) = \exp(-|t_1 - t_2|)$. Leidke juhusliku funktsiooni $Y(t) = X(t) \sin\left(t^2\right) + \cos^2 t$ kovariatsioon $K_y(t_1,t_2)$.
V: $K_y(t_1,t_2) = \sin\left(t_1^2\right) \sin\left(t_2^2\right) \exp(-|t_1 - t_2|)$.

7. Olgu U juhuslik suurus. Leidke juhuslike funktsioonide $X(t) = (t+1)U$ ja $Y(t) = (t-1)U$ vastastikune korrelatsioon $R_{xy}(t_1,t_2)$.
V: $R_{xy}(t_1,t_2) = \operatorname{sign}(t_1 + 1) \cdot \operatorname{sign}(t_2 - 1)$.

8. Näidake, et funktsioonidel $X(t)$ ja $X^o(t) = X(t) - m_x(t)$ on sama kovariatsioon.

9. On antud $X(t)$ kovariatsioon $K_x(t_1,t_2) = |\cos t_1 \cos t_2| \exp\left(-(t_1 - t_2)^2\right)$.
Leidke $R_x(t_1,t_2)$. V: $R_x(t_1,t_2) = \operatorname{sign}(\cos t_1) \cdot \operatorname{sign}(\cos t_2) \exp\left(-(t_1 - t_2)^2\right)$.

10. Näidake, et funktsioonide paaril $X(t)$ ja $Y(t)$ on sama vastastikune kovariatsioon kui funktsioonide paaril $X^o(t)$ ja $Y^o(t)$.

11. On antud $X(t)$, $Y(t)$ ja $Z(t)$ kovariatsioonid ja nende vastastikused kovariatsioonid. Leidke $U(t) = X(t) + Y(t) + Z(t)$ kovariatsioon $K_u(t_1,t_2)$.
V: $K_u(t_1,t_2) = K_x(t_1,t_2) + K_y(t_1,t_2) + K_z(t_1,t_2) + K_{xy}(t_1,t_2) + K_{xz}(t_1,t_2) +$
$+ K_{yx}(t_1,t_2) + K_{yz}(t_1,t_2) + K_{zx}(t_1,t_2) + K_{zy}(t_1,t_2)$.

12. Näidake, et kahe mittekorreleeruva juhusliku funktsiooni $X(t)$ ja $Y(t)$ korrutise $Z(t) = X(t)Y(t)$ keskväärtus $m_z(t)$ võrdub tegurite keskväärtuste $m_x(t)$ ja $m_y(t)$ korrutisega $m_x(t) \cdot m_y(t)$.

13. Tõestage, et kahe tsentreeritud mittekorreleeruva juhusliku funktsiooni korrutise kovariatsioon on tegurite kovariatsioonide korrutis.

14. Tõestage, et kolme sõltumatu tsentreeritud juhusliku funktsiooni korrutise kovariatsioon on tegurite kovariatsioonide korrutis.

15. Olgu antud $X(t) = a + \sum_{k=1}^{n} V_k \exp(-\alpha_k t)$, kus V_k on tsentreeritud mittekorreleeruvad juhuslikud suurused dispersioonidega D_k ja a ning α_k on kindlad suurused. Leidke $m_x(t)$, $K_x(t_1, t_2)$ ja $D_x(t)$. Kas $X(t)$ on statsionaarne?
V: $m_x(t) = a$, $K_x(t_1, t_2) = \sum_{k=1}^{n} D_k \exp(-\alpha_k (t_1 + t_2))$,
$D_x(t) = \sum_{k=1}^{n} D_k \exp(-2\alpha_k t)$, $X(t)$ ei ole statsionaarne.
16. Olgu antud $X(t) = a + \sum_{k=1}^{3} V_k \exp(-kt)$, kus V_k on tsentreeritud juhuslikud suurused. Olgu a kindel suurus ja

$$\begin{pmatrix} 4 & -8 & -6 \\ & 16 & 12 \\ & & 9 \end{pmatrix}$$

vektori (V_1, V_2, V_3) kovariatsioonimaatriks. Leidke $m_x(t)$, $K_x(t_1, t_2)$ ja $D_x(t)$. Kas $X(t)$ on statsionaarne? V: $m_x(t) = a$, $K_x(t_1, t_2) = 4\exp(-t_1 - t_2) -$
$-8\exp(-t_1 - 2t_2) - 6\exp(-t_1 - 3t_2) - 8\exp(-2t_1 - t_2) + 16\exp(-2t_1 - 2t_2) +$
$+12\exp(-2t_1 - 3t_2) - 6\exp(-3t_1 - t_2) + 12\exp(-3t_1 - 2t_2) + 9\exp(-3t_1 - 3t_2)$
$D_x(t) = 4\exp(-2t) - 16\exp(-3t) + 4\exp(-4t) + 24\exp(-5t) + 9\exp(-6t)$, ei ole statsionaarne.
17. Olgu $X(t) = t + U\cos\omega t + V\sin\omega t$, kusjuures U ja V on tsentreeritud mittekorreleeruvad juhuslikud suurused dispersioonidega $\mathrm{D}U = \mathrm{D}V = 2$. Leidke $m_x(t)$, kovariatsioon $K_x(t_1, t_2)$ ja dispersioon $D_x(t)$. Kas $X(t)$ on statsionaarne?
V: $m_x(t) = t$, $K_x(t_1, t_2) = 2\cos(\omega(t_1 - t_2))$, $D_x(t) = 2$, $X(t)$ ei ole statsionaarne, küll aga $X^o(t)$ on statsionaarne.
18. Olgu $X(t) = t + U\cos\omega t + V\sin\omega t$, kusjuures U ja V on tsentreeritud juhuslikud suurused ning $\begin{pmatrix} 2 & -1 \\ & 2 \end{pmatrix}$ on vektori (U, V) kovariatsioonimaatriks. Leidke $m_x(t)$, $K_x(t_1, t_2)$ ja $D_x(t)$. Kas $X(t)$ on statsionaarne? V: $m_x(t) = t$, $K_x(t_1, t_2) = 2\cos(\omega(t_1 - t_2)) - \sin(\omega(t_1 + t_2))$, $D_x(t) = 2 - \sin(2\omega t)$, ei ole statsionaarne.
19. Olgu $X(t) = V_1\cos\omega t + V_2\sin\omega t$, $Y(t) = U_1\cos\omega t + U_2\sin\omega t$, kusjuures $\mathrm{E}V_k = \mathrm{E}U_k = 0$ $(k = 1; 2)$ ning $\mathrm{D}V_k = 1$, $\mathrm{D}U_k = 4$ $(k = 1; 2)$. Vektori (V_1, V_2, U_1, U_2) korrelatsioonimaatriks on

$$\begin{pmatrix} 1 & 0 & 0.5 & 0 \\ & 1 & 0 & -0.5 \\ & & 1 & 0 \\ & & & 1 \end{pmatrix}.$$

Leidke selle vektori kovariatsioonimaatriks ja $R_{xy}(t_1, t_2)$. Kas $X(t)$ ja $Y(t)$ on statsionaarselt seotud?
V: $R_{xy}(t_1, t_2) = 0.5\cos(\omega(t_1 + t_2))$, $X(t)$ ja $Y(t)$ ei ole statsionaarselt seotud,

$$\begin{pmatrix} 1 & 0 & 1 & 0 \\ & 1 & 0 & -1 \\ & & 4 & 0 \\ & & & 4 \end{pmatrix}.$$

20. Olgu $Z(t) = X(t) + jY(t)$, kus j on imaginaarühik ja
$X(t) = \sum_{k=1}^{3} (a_k + V_k)\exp(-\alpha_k t)$, $Y(t) = \sum_{k=1}^{3} (b_k + U_k)\exp(-\beta_k t)$.

Seejuures on a_k, α_k, b_k ja β_k $(k = 1; 2; 3)$ kindlad suurused ning U_k ja V_k tsentreeritud juhuslikud suurused. Olgu

$$\begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ & 2 & 0 & 0 & -1 & 0 \\ & & 3 & 0 & 0 & 3 \\ & & & 1 & 0 & 0 \\ & & & & 2 & 0 \\ & & & & & 3 \end{pmatrix}$$

vektori $(V_1, V_2, V_3, U_1, U_2, U_3)$ kovariatsioonimaatriks. Leidke $m_z(t)$ ja $K_z(t_1, t_2)$ ning $D_z(t)$. V: $m_z(t) = \sum_{k=1}^{3} a_k \exp(-\alpha_k t) + j \sum_{k=1}^{3} b_k \exp(-\beta_k t)$,
$K_z(t_1, t_2) = \exp(-\alpha_1 (t_1 + t_2)) + 2\exp(-\alpha_2 (t_1 + t_2)) + 3\exp(-\alpha_3 (t_1 + t_2)) -$
$-j(\exp(-\alpha_1 t_1 - \beta_1 t_2) - \exp(-\alpha_2 t_1 - \beta_2 t_2) + \exp(-\alpha_3 t_1 - \beta_3 t_2)) +$
$+j(\exp(-\beta_1 t_1 - \alpha_1 t_2) - \exp(-\beta_2 t_1 - \alpha_2 t_2) + \exp(-\beta_3 t_1 - \alpha_3 t_2)) +$
$+\exp(-\beta_1 (t_1 + t_2)) + 2\exp(-\beta_2 (t_1 + t_2)) + 3\exp(-\beta_3 (t_1 + t_2))$,
$D_z(t) = \exp(-2\alpha_1 t) + 2\exp(-2\alpha_2 t) + 3\exp(-2\alpha_3 t) + \exp(-2\beta_1 t) +$
$+2\exp(-2\beta_2 t) + 3\exp(-2\beta_3 t)$.

21. Olgu $Z(t) = X(t) + Y(t)$. Leidke $m_z(t)$, $D_z(t)$ ja $K_z(t_1, t_2)$, kui
$m_x(t) = \cos^2 t$, $K_x(t_1, t_2) = \left(1 + \sin^2(t_1 - t_2)\right) / \left(1 + (t_1 - t_2)^2\right)$,
$m_y(t) = \sin^2 t$, $K_y(t_1, t_2) = \left(\cos^2(t_1 - t_2)\right) / \left(1 + (t_1 - t_2)^2\right)$ ja
$K_{xy}(t_1, t_2) = (\sin(t_1 - t_2)\cos(t_1 - t_2)) / \left(1 + (t_1 - t_2)^2\right)$.
V: $m_z(t) = 1$, $K_z(t_1, t_2) = 2/\left(1 + (t_1 - t_2)^2\right)$ ja $D_z(t) = 2$.

22. Olgu $K_x(t_1, t_2) = e^{-(t_2 - t_1)^2}$ ja $Y(t) = X'(t)$. Leidke $K_y(t_1, t_2)$, $K_{xy}(t_1, t_2)$, $K_{yx}(t_1, t_2)$, $R_y(t_1, t_2)$, $R_{xy}(t_1, t_2)$ ja $R_{yx}(t_1, t_2)$.
V: $K_y(t_1, t_2) = \left[2 - 4(t_2 - t_1)^2\right] \exp\left(-(t_2 - t_1)^2\right)$, $D_y(t) = 2$, $\sigma_y(t) = \sqrt{2}$,
$R_y(t_1, t_2) = \left[1 - 2(t_2 - t_1)^2\right] \exp\left(-(t_2 - t_1)^2\right)$, $\quad K_{xy}(t_1, t_2) =$
$= 2(t_1 - t_2)\exp\left(-(t_2 - t_1)^2\right)$, $K_{yx}(t_1, t_2) = 2(t_2 - t_1)\exp\left(-(t_2 - t_1)^2\right)$,
$R_{xy}(t_1, t_2) = \sqrt{2}(t_1 - t_2)\exp\left(-(t_2 - t_1)^2\right)$, $R_{yx}(t_1, t_2) = -R_{xy}(t_1, t_2)$.

23. Olgu $Y(t) = X'(t)$ ja $Z(t) = Y(t) - X(t)$. Kuidas on seotud $K_z(t_1, t_2)$ ja $K_x(t_1, t_2)$?
V: $K_z(t_1, t_2) = \dfrac{\partial^2 K_x(t_1, t_2)}{\partial t_1 \partial t_2} - \dfrac{\partial K_x(t_1, t_2)}{\partial t_1} - \dfrac{\partial K_x(t_1, t_2)}{\partial t_2} + K_x(t_1, t_2)$.

24. Olgu $K_x(t_1, t_2) = \sin\omega t_1 \sin\omega t_2$. Leidke $Y(t) = \int_0^t X(\tau)d\tau$ korral $K_y(t_1, t_2)$, $D_y(t)$, $\sigma_y(t)$ ja $R_y(t_1, t_2)$. V: $K_y(t_1, t_2) = \dfrac{(1 - \cos\omega t_1)(1 - \cos\omega t_2)}{\omega^2}$,
$D_y(t) = \dfrac{(1 - \cos\omega t)^2}{\omega^2}$, $\sigma_y(t) = \dfrac{1 - \cos\omega t}{\omega}$, $R_y(t_1, t_2) = 1$.

25. Olgu $Y(t) = \exp(t)\int_0^t X(\tau)\,d\tau$, kusjuures $K_x(t_1,t_2) = \sin\omega t_1 \sin\omega t_2$. Leidke $K_y(t_1,t_2)$, $D_y(t)$ ja $K_{xy}(t_1,t_2)$.
V: $K_y(t_1,t_2) = \dfrac{(1-\cos\omega t_1)(1-\cos\omega t_2)}{\omega^2}\exp(t_1+t_2)$,
$D_y(t) = \dfrac{(1-\cos\omega t)^2}{\omega^2}\exp(2t)$, $K_{xy}(t_1,t_2) = \dfrac{(1-\cos\omega t_2)\sin\omega t_1}{\omega}\exp(t_2)$.

26. Olgu $Y(t) = \int_0^t X(\tau)d\tau$ ja $K_x(t_1,t_2) = t_1+t_2+t_1t_2$. Leidke $K_y(t_1,t_2)$, $D_y(t)$, $\sigma_y(t)$, $R_y(t_1,t_2)$, $K_{xy}(t_1,t_2)$ ja $R_{xy}(t_1,t_2)$.
V: $K_y(t_1,t_2) = t_1t_2(2t_1+2t_2+t_1t_2)/4$, $D_y(t) = t^2\left(4t+t^2\right)/4$,
$\sigma_y(t) = |t|\sqrt{4t+t^2}/2$, $R_y(t_1,t_2) = \operatorname{sign}(t_1)\operatorname{sign}(t_2)\dfrac{2t_1+2t_2+t_1t_2}{\sqrt{4t_1+t_1^2}\sqrt{4t_2+t_2^2}}$,
$K_{xy}(t_1,t_2) = \left(2t_1t_2+t_2^2+t_1t_2^2\right)/2$, $R_{xy}(t_1,t_2) = \dfrac{2t_1t_2+t_2^2+t_1t_2^2}{\sqrt{t_2^2(2t_1+t_1^2)(4t_2+t_2^2)}}$.

27. Olgu $Y(t) = (\sin t)\int_0^t X(\tau)d\tau$ ja $K_x(t_1,t_2) = \exp(t_1+t_2)$. Leidke $K_y(t_1,t_2)$, $D_y(t)$, $\sigma_y(t)$, $K_{xy}(t_1,t_2)$ ja $R_{xy}(t_1,t_2)$.
V: $K_y(t_1,t_2) = (1-\exp t_1)(1-\exp t_2)\sin t_1\sin t_2$, $D_y(t) = (1-\exp t)^2\sin^2 t$,
$\sigma_y(t) = |(1-\exp t)\sin t|$, $K_{xy}(t_1,t_2) = (\exp t_2 - 1)\exp t_1 \sin t_2$, $R_{xy}(t_1,t_2) =$
$= \operatorname{sign}(\sin t_2)\operatorname{sign}(\exp t_2 - 1)$.

28. Olgu $K_x(t_1,t_2) = t_1^2t_2^2$ ja $Y(t) = \int_0^t X(\tau)d\tau$. Leidke $K_y(t_1,t_2)$, $D_y(t)$, $K_{xy}(t_1,t_2)$ ja $K_{yx}(t_1,t_2)$. V: $K_y(t_1,t_2) = t_1^3t_2^3/9$, $D_y(t) = t^6/9$,
$K_{xy}(t_1,t_2) = t_1^2t_2^3/3$, $K_{yx}(t_1,t_2) = t_1^3t_2^2/3$.

29. Olgu $Y(t) = \varphi(t)X(t)+\psi(t)X'(t)$, kus $\varphi(t)$ ja $\psi(t)$ on kindlad funktsioonid. Milline on seos $K_y(t_1,t_2)$ ja $K_x(t_1,t_2)$ vahel?
V: $K_y(t_1,t_2) = \varphi(t_1)\varphi(t_2)K_x(t_1,t_2) + \varphi(t_1)\psi(t_2)\dfrac{\partial}{\partial t_2}K_x(t_1,t_2) +$
$+\psi(t_1)\varphi(t_2)\dfrac{\partial}{\partial t_1}K_x(t_1,t_2) + \psi(t_1)\psi(t_2)\dfrac{\partial^2}{\partial t_1\partial t_2}K_x(t_1,t_2)$.

30. Olgu $X(t) = \cos(t+\Phi)$, kus Φ on juhuslik suurus, mis allub lõigul $[0;2\pi]$ ühtlasele jaotusele. Kas $X(t)$ on statsionaarne? V: $X(t)$ on statsionaarne.

31. Kas $X(t) = a\sin(\omega t+\Phi)$ $(a,\omega>0)$ on statsionaarne juhuslik funktsioon, kui

$$f(\varphi) = \begin{cases} \cos\varphi, & \text{kui } \varphi\in(0;\pi/2), \\ 0, & \text{kui } \varphi\notin(0;\pi/2) \end{cases}$$

on juhusliku suuruse Φ jaotustihedus? V: $X(t)$ ei ole statsionaarne.

32. Kas $X(t) = a\cos(\omega t+\Phi)$ $(a,\omega>0)$ on statsionaarne juhuslik funktsioon, kui

$$f(\varphi) = \begin{cases} \sin\varphi, & \text{kui } \varphi\in(0;\pi/2), \\ 0, & \text{kui } \varphi\notin(0;\pi/2) \end{cases}$$

on juhusliku suuruse Φ jaotustihedus? V: $X(t)$ ei ole statsionaarne.

33. Olgu $X(t)$ statsionaarne juhuslik funktsioon, $k_x(\tau) = 4\exp\left(-\alpha^2\tau^2\right)$ ja $Y(t) = 3X'(t)+2$. Leidke $k_y(\tau)$, $D_y(\tau)$ ja $r_y(\tau)$. V: $D_y(\tau) = 72\alpha^2$,
$k_y(\tau) = 72\alpha^2\exp\left(-\alpha^2\tau^2\right)\left(1-2\alpha^2\tau^2\right)$, $r_y(\tau) = \exp\left(-\alpha^2\tau^2\right)\left(1-2\alpha^2\tau^2\right)$.

34. Olgu $X(t)$ statsionaarne. Tõestage, et $k_{x'}(\tau) = -k_x''(\tau)$.

35. Olgu $X(t)$ statsionaarne ja $k_x(\tau) = (\cos\tau)\exp(-\tau^2)$. Leidke $k_{x'}(\tau)$ ja $\max|k_{x'}(\tau)|$. V: $k_{x'}(\tau) = 3(\cos\tau)\exp(-\tau^2) - 4\tau\exp(-\tau^2)(\sin\tau + \tau\cos\tau)$, $\max|k_{x'}(\tau)| = 3$.

36. Olgu $X(t)$ statsionaarne. Näidake, et

$$k_{xx'}(\tau) = k_x'(\tau)\,,\ k_{x'x}(\tau) = -k_x'(\tau)\,,\ k_{xx''}(\tau) = k_x''(\tau)\,.$$

37. Olgu teada vahemikus $(-1;1)$ statsionaarse juhusliku funktsiooni $X(t)$ kovariatsioon $k_x(\tau) = 2 - |\tau|$ $(\tau \in (-2;2))$. Leidke $X(t)$ spektraalarendus.
V: $X(t) = m_x(t) + \sum_{j=0}^{\infty} U_{2j+1}\cos((2j+1)\pi t/2) + V_{2j+1}\sin((2j+1)\pi t/2)$,
kus $\mathrm{D}U_{2j+1} = \mathrm{D}V_{2j+1} = 8/\left(\pi^2(2j+1)^2\right)$.

38. Olgu $X(t)$ statsionaarne. Leidke selle funktsiooni spektraaltihedus $s_x(\omega)$, kui $k_x(\tau) = \exp(-|\tau|)$. V: $s_x(\omega) = 1/\left(\pi\left(1+\omega^2\right)\right)$.

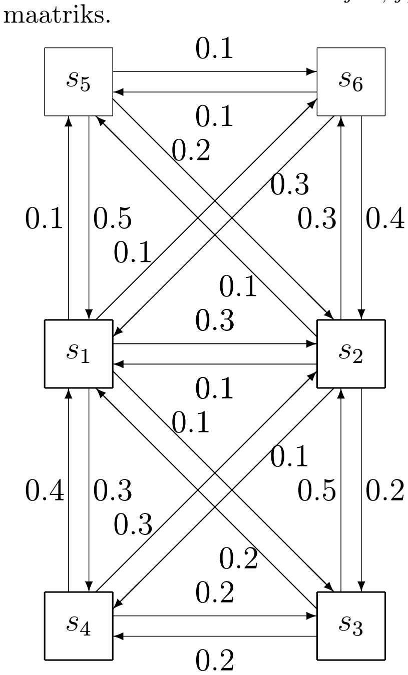
39. Leidke statsionaarse juhusliku funktsiooni $X(t)$ spektraaltihedus $s_x(\omega)$, kui

$$k_x(\tau) = \begin{cases} 1 - |\tau|\,, \text{ kui } |\tau| \le 1, \\ 0, \text{ kui } |\tau| > 1. \end{cases}$$

V: $s_x(\omega) = (1 - \cos\omega)/\left(\pi\omega^2\right)$.

40. Leidke statsionaarse juhusliku funktsiooni $X(t)$ dispersioon $D_x(t)$, kui $s_x(\omega) = 5/\left(\pi\left(\omega^2+1\right)\right)$ on selle funktsiooni spektraaltihedus. V: $D_x(t) = 5$.

41. Olgu diskreetsete seisunditega homogeensel Markovi ahelal kuus seisundit s_1, s_2, s_3, s_4, s_5 ja s_6, kusjuures üleminekutõenäosused ühest seisundist teise on antud skeemil ja $p_{i,i} = 1 - \sum_{j=1,\ j\neq i}^{6} p_{i,j}$ $(i = 1;2;3;4;5;6)$. Leidke üleminekumaatriks.



$V:$

$$\begin{pmatrix} 0.1 & 0.3 & 0.1 & 0.3 & 0.1 & 0.1 \\ 0.1 & 0.2 & 0.2 & 0.1 & 0.1 & 0.3 \\ 0.2 & 0.5 & 0.1 & 0.2 & 0 & 0 \\ 0.4 & 0.3 & 0.2 & 0.1 & 0 & 0 \\ 0.5 & 0.2 & 0 & 0 & 0.2 & 0.1 \\ 0.3 & 0.4 & 0 & 0 & 0.1 & 0.2 \end{pmatrix}$$

Peatükk 5

Matemaatiline statistika

5.1 Sissejuhatus. Põhimõisted

Oleme eelnevalt uurinud juhuslikke suurusi, juhuslikke vektoreid ja juhuslikke funktsioone, lähtudes nende teadaolevatest jaotusfunktsioonidest. Praktikas esinevate ülesannete korral me tavaliselt uuritava juhusliku suuruse (vektori, funktsiooni) jaotusfunktsiooni ei tea või ei tea seda täpselt. Tutvume järgnevas osas põgusalt, milliseid meetodeid sel juhul kasutada.

Statistikaks nimetatakse massnähtuste seaduspärasusi käsitlevat teadusharu. *Statistilisteks andmeteks* on katse, vaatluse, mõõtmise, küsitluse jms tulemusel saadud väärtused. *Matemaatiliseks statistikaks* nimetatakse matemaatika haru, mis tõenäosusteooriale tuginedes uurib statistiliste andmete põhjal järelduste tegemise meetodeid. Nimetame mõningad matemaatilise statistika suunad: *katseplaneerimine*, *statistiliste hüpoteeside kontrollimine, statistilised hinnangud, statistilised otsustused, mitmemõõtmeline statistiline analüüs, korrelatsioonanalüüs, komponentanalüüs, faktoranalüüs kanooniline analüüs, regressioonanalüüs, dispersioonanalüüs, kovariatsioonanalüüs, diskriminantanalüüs, klasteranalüüs, asümptootiliste meetodite teooria, juhuslike protsesside statistika* ja *mitteparameetriline statistika.*

Rakendusstatistikaks ehk *andmeanalüüsiks* nimetatakse mingis valdkonnas kogutud andmestiku töötlemist sisuliste järelduste saamiseks. Rakendusstatistika meetodid põhinevad matemaatilisel statistikal.

Üldkogumiks (*statistiliseks kogumiks, populatsiooniks*) nimetatakse objektide, mille kohta soovitakse teha statistilisi järeldusi, hulka. Üldkogumi objekte kirjeldatakse ühe või mitme *tunnusega*. Sellest tuleneb *ühe-* või *mitmemõõtmeline statistiline analüüs.* Need tunnused on juhuslikud suurused või juhuslikud vektorid. *Valim (väljavõte, väljavõtukogum, võend)* on üldkogumi objektide hulga lõplik alamhulk, mille põhjal tehakse järeldusi üldkogumi objektide tunnuse või tunnuste kohta. Eeldame järgnevalt, et

1) valim on juhuslik, st iga valimisse sattuv objekt on üldkogumist valitud juhuslikult,

2) objektid võetakse valimisse üksteisest sõltumatult,
3) iga üldkogumi objekt võib esineda valimis ühekordselt (kordusteta valim).

Objektide arvu valimis nimetatakse *valimi mahuks.* Valimi põhjal saadud *valimväärtuste* kirjapanekuks kasutatakse tavaliselt *statistilist rida,* mis on valimväärtuste esitus registreerimise järjekorras. Olgu X üldkogumi juhuslik suurus (objektide tunnus) meile tundmatu jaotusseadusega ja $\{x_1, x_2, \ldots, x_n\}$ selle tunnuse väärtused valimi objektide korral. Järgnevalt kasutame tinglikult nimetust valim ka statistilise rea $\{x_1, x_2, \ldots, x_n\}$ korral. Tänu valimi juhuslikkusele ja valimväärtuste sõltumatusele võime valimit $\{x_1, \ldots, x_n\}$ käsitleda kui juhuslikku vektorit $(x_1, \ldots, x_n)$, mille komponendid on sõltumatud ja alluvad kõik samale jaotusseadusele mis tunnus X. Seega oleks korrektne kasutada valimi jaoks tähistust $(X_1, \ldots, X_n)$ ja selle juhusliku vektori realisatsiooni jaoks tähistust $(x_1, \ldots, x_n)$. Tavaliselt piirdutakse siiski valimi tähistusega $\{x_1, \ldots, x_n\}$, jättes talle nii juhusliku vektori kui ka selle realisatsiooni osa. Viimane asjaolu esialgu raskendab jälgimist. *Variatsioonreaks* nimetatakse ühe tunnuse järgi järjestatud (ühemõõtmelist) valimit. Üldkogumi tunnuse X *empiiriliseks jaotuseks* nimetatakse tunnuse X iseloomustamiseks kasutatava diskreetse juhusliku suuruse X^*, kus $P(X^* = x_k) = 1/n \ \ (k = 1; \ldots; n)$, jaotust. Seega *empiirilist jaotusseadust* saab esitada tabeli

x_i	x_1	x_2	$\ldots$	x_n
$P(X^* = x_i)$	$1/n$	$1/n$	$\ldots$	$1/n$

Tabel 1

kujul. Kui valimis mõningad väärtused korduvad, siis on otstarbekas esitada variatsioonrida kujul

x_{i_j}	x_{i_1}	x_{i_2}	$\ldots$	x_{i_q}
n_{i_j}	n_{i_1}	n_{i_2}	$\ldots$	n_{i_q}

Tabel 2

kus x_{i_j} on väärtuste kasvamise (kahanemise) järgi järjestatud valimi erinevad väärtused ja *sagedus* n_{i_j} on väärtuse x_{i_j} esinemiste arv valimis, kusjuures $\sum_{j=1}^{q} n_{i_j} = n$. Sagedustabeli 2 põhjal skitseeritakse xn-tasandil *sageduste polügoon,* murdjoon, mis ühendab punkte (x_{i_k}, n_{i_k}) ja $\left(x_{i_{k+1}}, n_{i_{k+1}}\right)$ omavahel, kus $k = 1, \ldots, q-1$ ja $x_{i_1} < x_{i_2} < \ldots < x_{i_q}$. Korduvate väärtuste korral saame empiirilisele jaotusseadusele anda kuju

x_{i_j}	x_{i_1}	x_{i_2}	$\ldots$	x_{i_q}
$p^*_{i_j}$	$p^*_{i_1}$	$p^*_{i_2}$	$\ldots$	$p^*_{i_q}$

Tabel 3

kus $p^*_{i_j} = P\left(X^* = x_{i_j}\right) = \dfrac{n_{i_j}}{n}$ on väärtuse x_{i_j} *suhteline sagedus.* Tunnuse X *empiiriliseks karakteristikuks* nimetatakse tema empiirilise jaotuse põhjal leitud karakteristikut. Defineerime valimi $\{x_1, x_2, \ldots, x_n\}$ põhjal juhusliku suuruse X *empiirilise jaotusfunktsiooni*

$$F_n^*(x) \stackrel{def}{=} \frac{1}{n} \sum_{x_i < x} 1, \tag{5.1.1}$$

st empiirilise jaotusfunktsiooni $F_n^*(x)$ väärtuse arvutamiseks punktis x tuleb leida valimis tingimust $x_i < x$ rahuldavate väärtuste arv ja jagada see valimi mahuga. Et valim on juhuslik, siis on juhuslik ka väärtuste x_i arv, mis rahuldavad tingimust $x_i < x$ ja seega iga fikseeritud x korral on $F_n^*(x)$ juhuslik suurus. Järelikult on empiiriline jaotusfunktsioon $F_n^*(x)$ juhuslik funktsioon.

Kui valimi maht on suur, siis on otstarbekas variatsioonrida jaotada klassidesse, kusjuures klasside arvuks soovitatakse

valimi maht n	alla 50	50-100	100-250	üle 250
klasside arv m	5-7	6-10	7-12	10-20

Tabel 4

Mõningad autorid soovitavad võtta $m = [\sqrt{n}]$, kus $[\sqrt{n}]$ on arvu $\sqrt{n}$ täisosa. Järgmise sammuna koostatakse *klasside sagedustabel suhteliste* sageduste järgi

klass	$[a_0, a_1]$	$(a_1, a_2]$	$\cdots$	$(a_{m-1}, a_m]$
suhteline sagedus	p_1^{**}	p_2^{**}	$\cdots$	p_m^{**}

,

Tabel 5

milles on fikseeritud klassid ja väärtuste sagedus igas klassis, kusjuures $p_i^{**} = k_i/n$ ja k_i on i-ndasse klassi kuuluvate väärtuste arv. Kui klassid on ühesuguse ulatusega, siis tavaliselt valitakse

$$h = (x_n - x_1)/m, \ \ a_0 = x_1, \ \ a_i = a_0 + i\,h \quad (i = 1; \ldots ; m) \tag{5.1.2}$$

või

$$h = (x_n - x_1)/(m-1), \ \ a_0 = x_1 - h/2, \ \ a_i = a_0 + i\,h \quad (i = 1; \ldots ; m). \tag{5.1.3}$$

Histogrammiks (astmikdiagrammiks, sagedusjaotuse tulpdiagrammiks) nimetatakse sagedustabeli 5 graafilist kujutist, mil klasside sagedustele vastavad üksteise kõrval paiknevad tulbad, kusjuures tulba aluseks on klassi laius ja tulba kõrguseks klassi suhtelise sageduse ning klassi laiuse jagatis. Sellise valiku korral on tulba pindalaks klassi suhteline sagedus ja tulpade pindalade summa on üks. Milline seos on tulpdiagrammil ja üldkogumi pideva jaotusega uuritava juhusliku suuruse X jaotustihedusel $f(x)$? Osutub, et mõnikord on otstarbekas valimi $\{x_1, x_2, \ldots, x_n\}$ suhteliste sageduste Tabeli 5 omamisel tunnuse X empiirilise jaotusfunktsiooni $F_n^*(x)$ asemel kasutada funktsiooni

$$F_n^{\triangle}(x) \stackrel{\text{def.}}{=} \sum_{a_i < x} p_i^{**}, \tag{5.1.4}$$

st $F_n^{\triangle}(x)$ on sageduste p_i^{**}, mille korral $a_i < x$, summa.

Näide 1. Kolmekümne minuti jooksul fikseeriti iga minut kohvikusse sisenevate külastajate arv. Saadi valim mahuga 30 :

$$\{5, 4, 7, 4, 1, 1, 2, 5, 6, 2, 4, 7, 5, 3, 3, 6, 7, 6, 5, 5, 3, 2, 4, 7, 4, 6, 6, 6, 2, 11\}$$

Järjestame selle valimväärtused kasvamise järgi, koostame sagedustabeli, teeme sageduste polügooni, leiame empiirilise jaotusseaduse ja empiirilise jaotusfunktsiooni $F_{30}^*(x)$ ning selle graafiku. Jaotame variatsioonrea klassidesse ja koostame klasside sagedustabeli, suhteliste sageduste tabeli ning histogrammi. Leiame klassidele vastava empiirilise jaotusfunktsiooni $F_{30}^{\triangle}(x)$ ja skitseerime selle graafiku.

Saame variatsioonrea

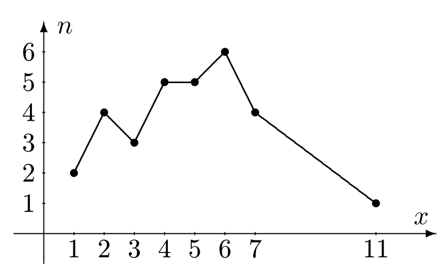
$$\{1, 1, 2, 2, 2, 2, 3, 3, 3, 4, 4, 4, 4, 4, 5, 5, 5, 5, 5, 6, 6, 6, 6, 6, 6, 7, 7, 7, 7, 11\}$$

Kuna selles valimis on võrdseid väärtusi, koostame sagedustabeli

x_{i_j}	1	2	3	4	5	6	7	11
n_{i_j}	2	4	3	5	5	6	4	1

Tabel 6

Skitseerime sageduste polügooni



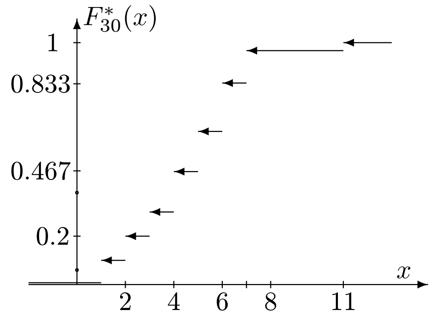
Esitame empiirilise jaotusseaduse ka kujul

x_{i_j}	1	2	3	4	5	6	7	11
$p_{i_j}^*$	2/30	4/30	3/30	5/30	5/30	6/30	4/30	1/30

Tabel 7

Leiame empiirilise jaotusfunktsiooni $F^*_{30}(x)$ ja skitseerime graafiku

$$F^*_{30}(x) = \begin{cases} 0, \ x \le 1, \\ 2/30, \ 1 < x \le 2, \\ 6/30, \ 2 < x \le 3, \\ 9/30, \ 3 < x \le 4, \\ 14/30, \ 4 < x \le 5, \\ 19/30, \ 5 < x \le 6, \\ 25/30, \ 6 < x \le 7, \\ 29/30, \ 7 < x \le 11, \\ 1, \ 11 < x \end{cases}$$



Jaotame variatsioonrea klassidesse. Tabeli 4 abil valime klasside arvu 6. Eeskirja (5.1.3) põhjal saame

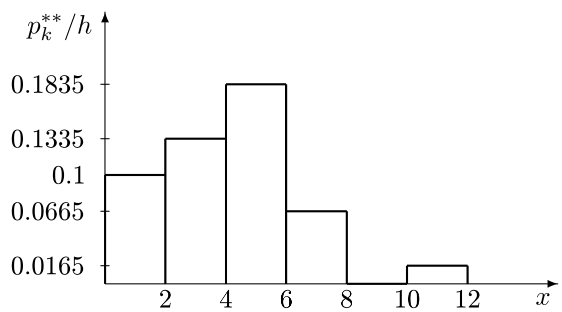
$$h = (11-1)\,/\,(6-1) = 2, \ \ a_0 = 1 - 2/2 = 0, \ \ a_i = 2i \quad (i = 1; \ldots ; 6)\,.$$

Koostame sageduste ja suhteliste sagedustabeli klasside korral tabeli

klass	$[0;2]$	$(2;4]$	$(4;6]$	$(6;8]$	$(8;10]$	$(10;12]$
k_i	6	8	11	4	0	1
$p_i^{**} = k_i/n$	6/30	8/30	11/30	4/30	0/30	1/30

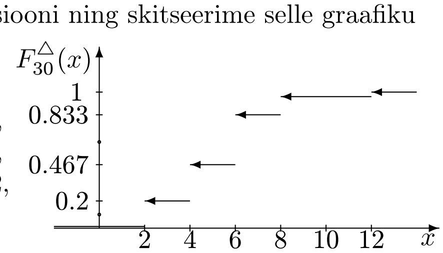
Tabel 8

Joonistame histogrammi



Leiame empiirilise jaotusfunktsiooni ning skitseerime selle graafiku

$$F^{\triangle}_{30}(x) = \begin{cases} 0, \ x \le 2, \\ 6/30, \ 2 < x \le 4, \\ 14/30, \ 4 < x \le 6, \\ 25/30, \ 6 < x \le 8, \\ 29/30, \ 8 < x \le 12, \\ 1, \ 12 < x \end{cases}$$



$\diamondsuit$

Lisaks eelnevalt esitatud valimi moodustamise viisile, mille abil saadakse nn *täiesti juhuslik valik,* kasutatakse veel *mehaanilist valikut, tüüpilist valikut* ja *seeriavalikut.* Mehaaniline valik toimub kindla intervalli järgi, näiteks valitakse iga sajas objekt. Tüüpilise valiku korral jaotatakse üldkogum mingi tunnuse põhjal osadeks ja igast osast valitakse proportsionaalselt vastav valimi osa. Seeriavaliku korral valitakse juhuslikult teatud üldkogumi osad, mille kõiki elemente seejärel uuritakse.

Tutvume järgmiste matemaatilise statistika ülesannetega:
1) jaotuse parameetrite määramine;
2) hüpoteeside kontroll;
3) katseandmete silumine vähimruutude meetodil.

5.2 Punkthinnangud

Olgu üldkogumi objektide tunnuse X kui juhusliku suuruse jaotus sõltuv m parameetrist $\alpha_k \quad (k = 1; \ldots ; m)$, st suuruse X jaotustihedus f on kujul $f(x, \alpha_1, \ldots, \alpha_m)$ ja kõik või osa neist parameetritest α_k on teadmata. Valimi $\{x_1, \ldots, x_n\}$ abil leitud parameetri α_k empiirilist väärtust $\alpha_k^* = \alpha_k^*(x_1, \ldots, x_n)$ nimetatakse parameetri α_k *punkthinnanguks* (*punktihinnanguks*). *Statistik* ehk *hinnangfunktsioon* $\alpha_k^*(x_1, \ldots, x_n)$ seab valimile $\{x_1, , \ldots, x_n\}$ vastavusse üldkogumi jaotust iseloomustava arvsuuruse α_k^*. Et valim on juhuslik, siis α_k^* on juhuslik suurus. Rõhutame, et meie poolt tehtud eeldustel on kõik valimväärtused x_i sõltumatud juhuslikud suurused ja on sama jaotusega, mis X. Kuna soovime, et punkthinnang α_k^* iseloomustaks parameetrit α_k võimalikult hästi, siis suvaline funktsioon $\alpha_k^*(x_1, \ldots, x_n)$ ei sobi punkthinnanguks. Punkthinnangu korral uuritakse kolme tingimuse täidetust.

1. Parameetri α_k punkthinnangut α_k^* nimetatakse *nihketa hinnanguks,* kui

$$\mathrm{E}\alpha_k^* = \alpha_k.$$

2. Parameetri α_k punkthinnangut α_k^* nimetatakse *mõjusaks hinnanguks,* kui

$$\forall \varepsilon > 0 : \quad \lim_{n \to \infty} P(|\alpha_k^* - \alpha_k| < \varepsilon) = 1.$$

3. Parameetri α_k punkthinnangut α_k^* nimetatakse *efektiivseks,* kui sellel on parameetri α_k teiste punkthinnangutega võrreldes vähim dispersioon.

Parameetri α_k punkthinnangut α_k^* nimetatakse *asümptootiliselt efektiivseks,* kui

$$\lim_{n \to \infty} \frac{\sigma^2_{\alpha_k^*}}{\sigma^2_{\alpha_k^{*e}}} = 1,$$

kus α_k^{*e} on parameetri α_k efektiivne punkthinnang.

Järgnevalt uurime põhiliselt üldkogumi juhusliku suuruse X keskväärtuse $\mathrm{E}X$ ja dispersiooni $\mathrm{D}X$ punkthinnanguid $(\mathrm{E}X)^*$ ja $(\mathrm{D}X)^*$, mis on vastavalt erijuhud suuruse X algmomendi ν_k ja keskmomendi μ_k punkthinnangutest ν_k^* ja μ_k^*. Samuti uurime vektori (X, Y) kovariatsiooni $\operatorname{cov}(X, Y)$ punkthinnanguid.

5.2.1 Algmomendi punkthinnang. Valimi keskmine

Defineerime üldkogumi juhusliku suuruse X algmomendi ν_k punkthinnangu valimi $\{x_1, x_2, \ldots, x_n\}$ korral

$$\nu_k^* \stackrel{\text{def.}}{=} c\sum_{i=1}^{n} x_i^k.$$

Valime kordaja c nii, et hinnang ν_k^* oleks nihketa. Saame

$$\mathrm{E}\nu_k^* = \mathrm{E}c\sum_{i=1}^{n} x_i^k = c\sum_{i=1}^{n} \mathrm{E}x_i^k = c\sum_{i=1}^{n} \mathrm{E}X^k = c \cdot n \cdot \nu_k.$$

Seega $c = 1/n \;\Rightarrow\; \mathrm{E}\nu_k^* = \nu_k$. Järelikult algmomendi ν_k otsitav punkthinnang

$$\nu_k^* = \frac{1}{n}\sum_{i=1}^{n} x_i^k \tag{5.2.1}$$

on nihketa. Erijuhul $k = 1$ saame valemist (5.2.1) keskväärtuse $\mathrm{E}X$ nihketa punkthinnangu

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i, \tag{5.2.2}$$

mida nimetatakse *valimi keskmiseks.* Seejuures

$$\mathrm{E}\,\overline{x} = \mathrm{E}X. \tag{5.2.3}$$

Leiame punkthinnangu ν_k^* dispersiooni

$$\begin{gathered}
\mathrm{D}\nu_k^* = \mathrm{D}\left(\nu_k^* - \mathrm{E}\nu_k^*\right)^2 = \mathrm{E}\left(\frac{1}{n}\sum_{i=1}^{n} x_i^k - \nu_k\right)^2 = \\
= \mathrm{E}\left(\frac{1}{n}\sum_{i=1}^{n} x_i^k - \frac{1}{n}\sum_{i=1}^{n}\nu_k\right)^2 = \mathrm{E}\left(\frac{1}{n}\sum_{i=1}^{n}\left(x_i^k - \nu_k\right)\right)^2 = \\
= \frac{1}{n^2}\mathrm{E}\sum_{i=1}^{n}\sum_{j=1}^{n}\left(x_i^k - \nu_k\right)\left(x_j^k - \nu_k\right) = \\
= \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\mathrm{E}\left(\left(x_i^k - \nu_k\right)\left(x_j^k - \nu_k\right)\right) = \\
= \left[\begin{array}{c} x_i \text{ sõltumatud } \Rightarrow\; x_i^k - \nu_k \text{ sõltumatud } \Rightarrow \\ \Rightarrow\; \mathrm{E}\left(\left(x_i^k - \nu_k\right)\left(x_j^k - \nu_k\right)\right) = \delta_{i,j}\mathrm{E}\left(\left(x_i^k - \nu_k\right)^2\right) \end{array}\right] = \\
= \frac{1}{n^2}\sum_{i=1}^{n}\mathrm{E}\left(x_i^k - \nu_k\right)^2 = \frac{1}{n^2}\sum_{i=1}^{n}\mathrm{E}\left(X^k - \nu_k\right)^2 = \frac{1}{n}\mathrm{E}\left(X^k - \nu_k\right)^2 = \frac{\mathrm{D}\left(X^k\right)}{n}.
\end{gathered}$$

Seega

$$\mathrm{D}\nu_k^* = \frac{1}{n}\mathrm{D}\left(X^k\right), \tag{5.2.4}$$

millest juhul $k = 1$ saame

$$\mathrm{D}\overline{x} = \frac{\mathrm{D}X}{n}. \tag{5.2.5}$$

Seosest (5.2.5) järeldub, et $n \to \infty \ \Rightarrow\ \mathrm{D}\overline{x} \to 0$. Seega punkthinnang $\overline{x}$ on mõjus. Ka hinnang ν_k^* on mõjus. Vastus küsimusele, kas hinnang ν_k^* on efektiivne või mitte, sõltub suuruse X jaotusest. Näiteks suuruse $X \sim N(a, \sigma)$ korral on hinnang $\overline{x}$ efektiivne. Sõnastame tõestatu.

Lause 1. Valemiga (5.2.1) määratud algmomendi ν_k punkthinnang ν_k^* on nihketa ja mõjus, kusjuures suuruse ν_k^* dispersioon $\mathrm{D}\nu_k^*$ on leitav valemi (5.2.4) abil.

Järeldus 1. Valemiga (5.2.2) määratud valimi keskmine $\overline{x}$ on suuruse X keskväärtuse $\mathrm{E}X$ nihketa mõjus punkthinnang, kusjuures kehtib seos (5.2.5).

Märkus 1. Kui valimis on korduvaid väärtusi ja variatsioonrida on esitatud Tabeli 5.1.2 abil, siis algmomendi ν_k valemiga (5.2.1) esitatud punkthinnang ν_k^* on esitatav kujul

$$\nu_k^* = \left(\sum\nolimits_{j=1}^{q} n_{i_j} x_{i_j}^k\right) / \left(\sum\nolimits_{j=1}^{q} n_{i_j}\right) \tag{5.2.1.6}$$

ja keskväärtuse $\mathrm{E}X = \nu_1$ valemiga (5.2.2) esitatud punkthinnang $\overline{x}$ kujul

$$\overline{x} = \left(\sum\nolimits_{j=1}^{q} n_{i_j} x_{i_j}\right) / \left(\sum\nolimits_{j=1}^{q} n_{i_j}\right). \tag{5.2.7}$$

Märkus 2. Kui valimi elemendid x_i on suured, kuid lähedased arvud, siis arvutuste lihtsustamiseks on otstarbekas lahutada valimi igast elemendist sama arv c, st vaadelda hulka $\{u_1, \ldots, u_n\}$, kus $u_i = x_i - c$, ja esitada (5.2.7) kujul

$$\overline{x} = c + \left(\sum\nolimits_{j=1}^{q} n_{i_j} u_{i_j}\right) / \left(\sum\nolimits_{j=1}^{q} n_{i_j}\right). \tag{5.2.8}$$

Näide 1. Leiame valemi (5.2.7) abil Näites 5.1.1 esitatud valimi keskmise:

$$\overline{x} = \frac{1}{30}(2\cdot 1 + 4\cdot 2 + 3\cdot 3 + 5\cdot 4 + 5\cdot 5 + 6\cdot 6 + 4\cdot 7 + 1\cdot 11) \approx 4.633. \quad \diamondsuit$$

Näide 2. Olgu variatsioonrida esitatud kujul

x_i	992	998	1007	1009
n_i	7	9	8	4

Leiame valemi (5.2.8) abil valimi keskmise.

Valime $c = 1000$. Miks? Koostame tabeli

u_i	-8	-2	7	9
n_i	7	9	8	4

Valemi (5.2.8) abil saame

$$\overline{x} = 1000 + \frac{1}{28}(7\cdot(-8) + 9\cdot(-2) + 8\cdot 7 + 4\cdot 9) \approx 1000.643. \quad \diamondsuit$$

Valimi keskmist on tihti otstarbekas leida rekursiivselt.

Lause 2. Valimi $\{x_1, x_2, \ldots, x_n\}$ keskmist saab leida rekursiivselt

$$\overline{x}_{j+1} = \overline{x}_j + \frac{x_{j+1} - \overline{x}_j}{j+1} \quad (j = 1; 2; \ldots; n-1)\,, \tag{5.2.9}$$

kus

$$\overline{x}_j \overset{\text{def.}}{=} \frac{1}{j}\sum_{i=1}^{j} x_i \quad (j = 1; 2; \ldots; n)\,,\ \overline{x} = \overline{x}_n. \tag{5.2.10}$$

Tõestage antud väide iseseisvalt. □

Millist lisainfot annab eeskirja (5.2.9) rakendamine?

5.2.2 Üldkogumi dispersiooni punkthinnangud

Uurime üldkogumi juhusliku suuruse X dispersiooni $\mathrm{D}X = \mu_2$ punkthinnangut valimi $\{x_1, x_2, \ldots, x_n\}$ korral. Eristame kaht juhtu.

I On teada üldkogumi juhusliku suuruse keskväärtus.

Otsime sel korral dispersiooni nihketa hinnangut kujul

$$\mu_2^* \overset{\text{def.}}{=} c\sum_{i=1}^{n} (x_i - \nu_1)^2\,. \tag{5.2.11}$$

Valime kordaja c nii, et hinnang μ_2^* oleks nihketa. Kuna valimisse võetakse elemendid üksteisest sõltumatult, siis

$$\begin{aligned}\mathrm{E}\mu_2^* &= \mathrm{E}c\sum_{i=1}^{n} (x_i - \nu_1)^2 = c\sum_{i=1}^{n} \mathrm{E}\,(x_i - \nu_1)^2 = \\ &= c\sum_{i=1}^{n} \mathrm{E}\,(X - \nu_1)^2 = c \cdot n \cdot \mathrm{D}X\end{aligned}$$

ja seosega (5.2.11) antud juhusliku suuruse X dispersiooni $\mathrm{D}X$ punkthinnang on nihketa, kui $c = 1/n$. Seega, teades keskväärtust $\mathrm{E}X = \nu_1$, on otstarbekas valida

$$\mu_2^* = \frac{1}{n}\sum_{i=1}^{n} (x_i - \mathrm{E}X)^2 \tag{5.2.12}$$

dispersiooni $\mathrm{D}X$ punkthinnanguks. Sel korral $\mathrm{E}\mu_2^* = \mathrm{D}X$. Leiame punkthinnangu (5.2.12) dispersiooni

$$\begin{aligned}\mathrm{D}\mu_2^* &= \mathrm{E}\,(\mu_2^* - \mathrm{E}\mu_2^*)^2 = \mathrm{E}\,(\mu_2^* - \mu_2)^2 = \mathrm{E}\left(\frac{1}{n}\sum_{i=1}^{n} (x_i - \nu_1)^2 - \mu_2\right)^2 = \\ &= \mathrm{E}\left(\frac{1}{n}\left(\sum_{i=1}^{n} (x_i - \nu_1)^2 - n\mu_2\right)\right)^2 = \frac{1}{n^2}\mathrm{E}\left(\sum_{i=1}^{n} (x_i - \nu_1)^2 - n\,\mu_2\right)^2 =\end{aligned}$$

$$= \frac{1}{n^2}\mathrm{E}\left(\left(\sum_{i=1}^{n}(x_i-\nu_1)^2 - n\,\mu_2\right)\left(\sum_{j=1}^{n}(x_j-\nu_1)^2 - n\,\mu_2\right)\right) =$$

$$= \frac{1}{n^2}\sum_{i=1}^{n}\mathrm{E}\,(x_i-\nu_1)^4 + \frac{1}{n^2}\sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n}\mathrm{E}\left((x_i-\nu_1)^2(x_j-\nu_1)^2\right) -$$

$$-\frac{1}{n^2}n\mu_2\left(\sum_{j=1}^{n}\mathrm{E}\,(x_j-\nu_1)^2 + \sum_{i=1}^{n}\mathrm{E}\,(x_i-\nu_1)^2\right) + \frac{1}{n^2}\mathrm{E}n^2\mu_2^2 =$$

$$= \frac{n\mu_4}{n^2} + \frac{n\,(n-1)\,\mu_2^2}{n^2} - \frac{2n^2\mu_2^2}{n^2} + \frac{n^2\mu_2^2}{n^2} = \frac{1}{n}\left(\mu_4-\mu_2^2\right).$$

Sõnastame tõestatu.

Lause 1. Kui on teada EX, siis seosega (5.2.12) määratud dispersiooni DX punkthinnang μ_2^* on nihketa. Kui eksisteerib μ_4, siis hinnang μ_2^* on ka mõjus. Seejuures $\mathrm{D}\mu_2^* = \frac{1}{n}\left(\mu_4-\mu_2^2\right)$.

Märkus 1. Kui valimis on korduvaid elemente ja variatsioonrida on esitatud Tabeli 5.1.2 abil, siis dispersiooni DX valemiga (5.2.12) esitatud punkthinnang μ_2^* on esitatav kujul

$$\mu_2^* = \frac{1}{n}\sum_{j=1}^{q} n_{i_j}\left(x_{i_j}-\mathrm{E}X\right)^2. \tag{5.2.13}$$

II Üldkogumi juhusliku suuruse keskväärtust ei ole teada.

Kui EX ei ole teada, siis otsime juhusliku suuruse X dispersiooni DX nihketa punkthinnangut kujul

$$s^2 \overset{\text{def.}}{=} c\sum_{i=1}^{n}(x_i-\overline{x})^2. \tag{5.2.14}$$

Valime kordaja c nii, et hinnang s^2 on nihketa. Kuna

$$\mathrm{E}s^2 = \mathrm{E}c\sum_{i=1}^{n}(x_i-\overline{x})^2 = \mathrm{E}c\sum_{i=1}^{n}\left(x_i-\frac{1}{n}\sum_{j=1}^{n}x_j\right)^2 =$$

$$= \mathrm{E}c\sum_{i=1}^{n}\left(\frac{1}{n}\sum_{j=1}^{n}(x_i-x_j)\right)^2 = \mathrm{E}\frac{c}{n^2}\sum_{i=1}^{n}\left(\sum_{j=1}^{n}((x_i-\nu_1)-(x_j-\nu_1))\right)^2 =$$

$$= \mathrm{E}\frac{c}{n^2}\sum_{i=1}^{n}\left(\sum_{j=1}^{n}((x_i-\nu_1)-(x_j-\nu_1))\right)\left(\sum_{k=1}^{n}((x_i-\nu_1)-(x_k-\nu_1))\right) =$$

$$= \frac{c}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\mathrm{E}\left(x_i-\nu_1\right)^2 - \frac{c}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\mathrm{E}\left(x_i-\nu_1\right)\left(x_k-\nu_1\right) -$$

$$- \frac{c}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\mathrm{E}\left(x_j-\nu_1\right)\left(x_i-\nu_1\right) + \frac{c}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\mathrm{E}\left(x_j-\nu_1\right)\left(x_k-\nu_1\right) =$$

$$= \frac{cn^3\mu_2}{n^2} - \frac{cn^2\mu_2}{n^2} - \frac{cn^2\mu_2}{n^2} + \frac{cn^2\mu_2}{n^2} = c\left(n-1\right)\mu_2,$$

siis $c = 1/\left(n-1\right)$ korral on punkthinnang s^2 nihketa. Seega, mitte teades üldkogumi juhusliku suuruse X keskväärtust EX, on

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i-\overline{x}\right)^2 \qquad (5.2.15)$$

dispersiooni DX nihketa punkthinnang. Suurust s^2 nimetatakse *valimi dispersiooniks.* Kuna

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i-\overline{x}\right)^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n}x_i^2 - 2\overline{x}\sum_{i=1}^{n}x_i + \sum_{i=1}^{n}\overline{x}^2\right) =$$

$$= \frac{1}{n-1}\left(n\overline{x^2} - 2n\overline{x}^2 + n\overline{x}^2\right) = \frac{n}{n-1}\left(\overline{x^2} - \overline{x}^2\right),$$

siis saame valimi dispersiooni leida ka valemi

$$s^2 = \frac{n}{n-1}\left(\overline{x^2} - \overline{x}^2\right) \qquad (5.2.16)$$

abil, kus $\overline{x^2} = (1/n)\sum_{i=1}^{n}x_i^2$.

Lause 2. Kui EX ei ole teada, siis s^2 on suuruse DX nihketa hinnang.

Saab näidata, et s^2 on teatud lisatingimustel ka mõjus.

Kui valimis on korduvaid elemente ja variatsioonrida on esitatud Tabeli 5.1.2 abil, siis s^2 on esitatav kujul

$$s^2 = \left(\sum\nolimits_{j=1}^{q} n_{i_j}\left(x_{i_j}-\overline{x}\right)^2\right) / \left(\sum\nolimits_{j=1}^{q} n_{i_j} - 1\right). \qquad (5.2.17)$$

Suurust s^2 võib leida rekursiivselt.

Lause 3. Kui EX ei ole teada, siis valimi dispersiooni võib leida rekursiivselt

$$s_{j+1}^2 = \left(1-\frac{1}{j}\right)s_j^2 + (j+1)\left(\overline{x}_{j+1}-\overline{x}_j\right)^2 \quad \left(j = 1;2;\ldots;n-1\right), \qquad (5.2.18)$$

kus suurus $\overline{x}_j$ on määratud valemiga (5.2.10) ja $s_1^2 = 0$ ning

$$s_j^2 = \frac{1}{j-1}\sum_{i=1}^{j}\left(x_i-\overline{x}\right)^2 \;\; \left(j = 2;3;\ldots;n\right), \;\; s_n^2 = s^2. \qquad (5.2.19)$$

Tõestage Laused 4 ja 5 iseseisvalt.

Lause 4. Kui

$$m_2 \overset{\text{def.}}{=} \frac{1}{n}\sum_{i=1}^{n}(x_i-\overline{x})^2, \tag{5.2.20}$$

siis kehtivad seosed

$$m_2 = \frac{n-1}{n}s^2, \quad s^2 = \frac{n}{n-1}m_2. \tag{5.2.21}$$

Suuremahulise valimi korral erinevad dispersiooni nihkega hinnang m_2 ja nihketa hinnang s^2 teineteisest vähe ja üldkogumi juhusliku suuruse X dispersiooni DX punkthinnanguna kasutatakse sageli suurust m_2.

Lause 5. Kui valim on antud Tabeli 5.1.2 kujul, siis

$$m_2 = \overline{x^2} - (\overline{x})^2 = \frac{1}{n}\sum_{j=1}^{q} n_{i_j}x_{i_j}^2 - \frac{1}{n^2}\left(\sum_{j=1}^{q} n_{i_j}x_{i_j}\right)^2 \tag{5.2.22}$$

ja

$$s^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n}x_i^2 - n\overline{x}^2\right) = \frac{1}{n-1}\sum_{j=1}^{q} n_{i_j}x_{i_j}^2 - \frac{1}{n(n-1)}\left(\sum_{j=1}^{q} n_{i_j}x_{i_j}\right)^2. \tag{5.2.23}$$

Lause 6 (vt [30])**.** Üldkogumi sümmeetrilise jaotuse korral on punkthinnangud $\overline{x}$ ja s^2 mittekorreleeruvad, st cov $\left(\overline{x},\, s^2\right) = 0$.

Näide 1. Leiame Näites 5.1.1 esitatud valimi põhjal suuruse X dispersiooni DX punkthinnangu s^2.

Et suuruse X keskväärtust EX ei ole antud, siis on tegemist juhuga II. Näites 5.2.1.1 leidsime selle valimi keskmise $\overline{x} \approx 4.633$. Dispersiooni D$X$ punkthinnangu s^2 saame valemi (5.2.17) abil

$$s^2 = \frac{1}{30-1}\left(2\cdot(1-4.633)^2 + 4\cdot(2-4.633)^2 +\right.$$

$$+3\cdot(3-4.633)^2 + 5\cdot(4-4.633)^2 + 5\cdot(5-4.633)^2 +$$

$$\left.+6\cdot(6-4.633)^2 + 4\cdot(7-4.633)^2 + 1\cdot(11-4.633)^2\right) \approx 4.79. \quad \diamondsuit$$

Näide 2. Valimi maht on 100 ja $m_2 = 4$ on dispersiooni nihkega hinnang. Leiame dispersiooni nihketa hinnangu s^2.

Seostest (5.2.21) teise abil saame

$$s^2 = \frac{n}{n-1}m_2 = \frac{100}{99}\cdot 4 \approx 4.04. \qquad \diamondsuit$$

5.2.3 Normaaljaotusele alluva juhusliku suuruse keskväärtuse ja dispersiooni punkthinnangud

Kui täiendavalt eeldada, et üldkogumi juhuslik suurus allub normaaljaotusele, siis on võimalik tõestada mitu huvitavat väidet.

Lause 1. Kui $X \sim N(a, \sigma)$, siis:
1° $\overline{x} \sim N(a, \sigma/\sqrt{n})$, $\sqrt{n}(\overline{x} - a)/\sigma \sim N(0;1)$;
2° $\overline{x}$ ja s^2 on sõltumatud juhuslikud suurused;
3° $(n-1)s^2/\sigma^2$ on juhuslik suurus, millel on χ^2-jaotus vabadusastmete arvuga $n-1$.

Tõestus. Esitame väidete 1° ja 3° tõestused. Väite 2° tõestuse leiate monograafiast [30]. Kuna

$$g_{\overline{x}}(\omega) = \mathrm{E}e^{i\omega\overline{x}} = \mathrm{E}e^{i(\omega/n)\sum_{k=1}^{n} x_k} = \mathrm{E}\left(e^{i(\omega/n)x_1} e^{i(\omega/n)x_2} \cdots e^{i(\omega/n)x_n}\right) =$$

$$= [\text{sõltumatud } x_k\text{-d}] = \left(\mathrm{E}e^{i(\omega/n)x_1}\right)\left(\mathrm{E}e^{i(\omega/n)x_2}\right)\cdots\left(\mathrm{E}e^{i(\omega/n)x_n}\right) =$$
$$= \left(\mathrm{E}e^{i(\omega/n)X}\right)^n = \left[X \sim N(a,\sigma) \Leftrightarrow g_X(\omega) = e^{ia\omega - \sigma^2\omega^2/2}\right] =$$
$$= \left(e^{ia(\omega/n) - \sigma^2(\omega/n)^2/2}\right)^n = e^{ia\omega - \left(\sigma/\sqrt{n}\right)^2 \omega^2/2} \Leftrightarrow \overline{x} \sim N\left(a, \sigma/\sqrt{n}\right) \Leftrightarrow$$
$$\Leftrightarrow \sqrt{n}(\overline{x} - a)/\sigma \sim N(0;1),$$

siis väide 1° on tõene. Esitame võrduste ahela

$$\sum_{i=1}^{n}(x_i - a)^2 = \sum_{i=1}^{n}((x_i - \overline{x}) + (\overline{x} - a))^2 =$$
$$= \sum_{i=1}^{n}(x_i - \overline{x})^2 + 2\sum_{i=1}^{n}(x_i - \overline{x})(\overline{x} - a) + \sum_{i=1}^{n}(\overline{x} - a)^2 =$$
$$= \sum_{i=1}^{n}(x_i - \overline{x})^2 + 2(\overline{x} - a)\sum_{i=1}^{n}(x_i - \overline{x}) + n(\overline{x} - a)^2 =$$
$$= \left[\sum_{i=1}^{n}(x_i - \overline{x}) = \sum_{i=1}^{n} x_i - n\,\overline{x} = 0\right] = \sum_{i=1}^{n}(x_i - \overline{x})^2 + n(\overline{x} - a)^2.$$

Seega

$$\sum_{i=1}^{n}(x_i - a)^2 = \sum_{i=1}^{n}(x_i - \overline{x})^2 + n(\overline{x} - a)^2 \Leftrightarrow$$
$$\sum_{i=1}^{n}\left(\frac{x_i - a}{\sigma}\right)^2 = \frac{\sum_{i=1}^{n}(x_i - \overline{x})^2}{\sigma^2} + \left(\frac{\sqrt{n}(\overline{x} - a)}{\sigma}\right)^2 \Leftrightarrow$$
$$\sum_{i=1}^{n}\left(\frac{x_i - a}{\sigma}\right)^2 = \frac{(n-1)s^2}{\sigma^2} + \left(\frac{\sqrt{n}(\overline{x} - a)}{\sigma}\right)^2. \qquad (5.2.24)$$

Valimi elemendid x_i on sõltumatud. Seega on sõltumatud ka $(x_i - a)/\sigma \sim N(0;1)$ ja seose (5.2.24) vasakul poolel olev suurus allub χ^2-jaotusele vabadusastmete arvuga n. Kuna väite 2° põhjal on $\overline{x}$ ja s^2 sõltumatud, siis on sõltumatud ka seose (5.2.24) paremal poolel olevad liidetavad. Kuna $\sqrt{n}\,(\overline{x}-a)/\sigma \sim N(0;1)$, siis $\left(\sqrt{n}\,(\overline{x}-a)/\sigma\right)^2$ allub χ^2-jaotusele vabadusastmete arvuga 1. Seega suurus $(n-1)\,s^2/\sigma^2$ on χ^2-jaotusega, mille vabadusastmete arv on $n-1$. □

Näide 1. Uuriti tudengi lõunatamiseks kuluvat aega. See on juhuslik suurus, mis allub normaaljaotusele standardhälbega 3 minutit. Kaastudeng mõõtis 15 päeval selle tudengi lõunatamiseks kuluvat aega ja leidis selle dispersiooni. Milline on tõenäosus, et ta saab valimi dispersiooni suurema kui 12?

Lause 1 väite 3° põhjal on juhuslik suurus

$$Q = (n-1)\,s^2/\sigma^2 = 14s^2/9$$

χ^2-jaotusega vabadusastmete arvuga $n-1=14$. Seega

$$\begin{aligned} P\left(s^2 > 12\right) &= P\left(14s^2/9 > 14\cdot 12/9\right) = P\,(Q > 18.667) = \\ &= 1 - P\,(Q \le 18.667) = \\ &= 1 - \mathrm{ChiSquareDist}\,(18.667; 14) \approx \\ &\approx 1 - 0.8219 = 0.1781. \qquad \diamondsuit \end{aligned}$$

Lause 2. Kui X allub normaaljaotusele keskväärtusega a, siis $\sqrt{n}\,(\overline{x}-a)/s$ on Studenti jaotusega, mille vabadusastmete arv on $n-1$.

Näide 2. Suurus X on normaaljaotusega $N(a,\sigma)$. Saadakse valim mahuga 25, mille korral $\overline{x}=10$ ja $s^2=9$. Leiame sündmuse $a\in[9;12]$ tõenäosuse.

Kuna Lause 2 põhjal on suurus $T=\sqrt{n}\,(\overline{x}-a)/s = 5\,(10-a)/3$ Studenti jaotusega vabadusastmete arvuga $n-1=24$, siis

$$\begin{aligned} P\,(9\le a\le 12) &= P\,(-9\ge -a\ge -12) = P\,(10-9\ge 10-a\ge 10-12) = \\ &= P\,(1\ge 10-a\ge -2) = P\,(5/3\ge T\ge -10/3) = \\ &= \mathrm{TDist}\,(5/3;24) - \mathrm{TDist}\,(-10/3;24) \approx 0.9443. \qquad \diamondsuit \end{aligned}$$

5.2.4 Juhusliku vektori arvkarakteristikute punkthinnangud

Uurime üldkogumi juhusliku vektori (X,Y) komponentide kovariatsioonimomendi ja korrelatsioonikordaja punkthinnanguid. Olgu vektori (X,Y) korral sooritatud n sõltumatut katset, mis andsid valimi

$$\{(x_1;y_1),(x_2;y_2),\ \ldots\ ,(x_n;y_n)\}.$$

Vaatame kahte juhtu.

I Vektori (X,Y) **juht**, kui EX ja EY on teada. Näidake, et sellel juhul

$$\mathrm{K}^{***}_{x,\,y} = \frac{1}{n}\sum_{i=1}^{n}(x_i-\mathrm{E}X)\,(y_i-\mathrm{E}Y)$$

on kovariatsiooni $\mathrm{cov}\,(X,Y)=\mathrm{K}_{x,\,y}$ nihketa punkthinnang.

II Vektori (X, Y) juht, kui EX ja EY ei ole teada.

Valemite (5.2.2) ja (5.2.15) abil saame leida parameetrite $\mathrm{E}X = \nu_x$, $\mathrm{E}Y = \nu_y$, DX ja DY punkthinnangud $\overline{x}$, $\overline{y}$, s_x^2 ja s_y^2. Kovariatsiooni $\operatorname{cov}(X, Y) = \mathrm{K}_{x,\,y}$ nihketa punkthinnangut $\mathrm{K}^*_{x,\,y}$ otsime kujul

$$\mathrm{K}^*_{x,\,y} = c \sum_{i=1}^{n} (x_i - \overline{x})(y_i - \overline{y})\,. \tag{5.2.25}$$

Leiame

$$\mathrm{E}\,\mathrm{K}^*_{x,\,y} = \mathrm{E}\, c \sum_{i=1}^{n} (x_i - \overline{x})(y_i - \overline{y}) =$$

$$= c \sum_{i=1}^{n} \mathrm{E}\left[\left(x_i - \frac{1}{n}\sum_{j=1}^{n} x_j\right)\left(y_i - \frac{1}{n}\sum_{j=1}^{n} (y_j)\right)\right] =$$

$$= \frac{c}{n^2}\sum_{i=1}^{n} \mathrm{E}\left[\left(\sum_{j=1}^{n} ((x_i - \nu_x) - (x_j - \nu_x))\right)\left(\sum_{p=1}^{n} ((y_i - \nu_y) - (y_p - \nu_y))\right)\right] =$$

$$= \frac{c}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{p=1}^{n} \mathrm{E}\left[((x_i - \nu_x) - (x_j - \nu_x))((y_i - \nu_y) - (y_p - \nu_y))\right] =$$

$$= \frac{c}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{p=1}^{n} \mathrm{E}\left[(x_i - \nu_x)(y_i - \nu_y)\right] - \frac{c}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{p=1}^{n} \mathrm{E}\left[(x_i - \nu_x)(y_p - \nu_y)\right] -$$

$$- \frac{c}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{p=1}^{n} \mathrm{E}\left[(x_j - \nu_x)(y_i - \nu_y)\right] + \frac{c}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{p=1}^{n} \mathrm{E}\,(x_j - \nu_x)(y_p - \nu_y) =$$

$$= \left[\mathrm{E}\left[(x_i - \nu_x)(y_i - \nu_y)\right] = \mathrm{K}_{x,\,y},\ \mathrm{E}\left[(x_j - \nu_x)(y_i - \nu_y)\right] = \delta_{i,\,j}\mathrm{K}_{x,\,y}\right] =$$

$$= \frac{c}{n^2}\left(n^3\mathrm{K}_{x,\,y} - n^2\mathrm{K}_{x,\,y}\right) = c\,(n-1)\,\mathrm{K}_{x,\,y}.$$

Seega on hinnang (5.2.25) nihketa, kui $c = 1/(n-1)$, st

$$\mathrm{K}^*_{x,\,y} = \frac{1}{n-1}\sum_{i=1}^{n} (x_i - \overline{x})(y_i - \overline{y})\,. \tag{5.2.26}$$

Lause 1. Suurus $\mathrm{K}^*_{x,\,y}$ on vektori (X, Y) kovariatsioonimomendi $\operatorname{cov}(X, Y)$ nihketa punkthinnang.

Hinnangu (5.2.26) asemel kasutatakse tihti suuruse $\operatorname{cov}(X, Y)$ nihkega punkthinnangut

$$\mathrm{K}^{**}_{x,\,y} = \frac{1}{n}\sum_{i=1}^{n} (x_i - \overline{x})(y_i - \overline{y})\,, \tag{5.2.27}$$

kusjuures

$$\mathrm{K}^*_{x,\,y} = \frac{n}{n-1}\mathrm{K}^{**}_{x,\,y}\,. \tag{5.2.28}$$

Lause 2. Kehtib seos

$$\mathrm{K}^{**}_{x,\,y} = \overline{xy} - \overline{x}\,\overline{y}, \tag{5.2.29}$$

kus

$$\overline{xy} = \frac{1}{n}\sum_{i=1}^{n} x_i y_i. \tag{5.2.30}$$

Tõestus. Leiame

$$\mathrm{K}^{**}_{x,\,y} = \frac{1}{n}\sum\nolimits_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y}) =$$

$$= \frac{1}{n}\sum\nolimits_{i=1}^{n} x_i y_i - \overline{y}\frac{1}{n}\sum\nolimits_{i=1}^{n} x_i - \overline{x}\frac{1}{n}\sum\nolimits_{i=1}^{n} y_i + \overline{xy}\frac{1}{n}\sum\nolimits_{i=1}^{n} 1 =$$

$$= \overline{xy} - \overline{y}\,\overline{x} - \overline{x}\,\overline{y} + \overline{x}\,\overline{y} = \overline{xy} - \overline{x}\,\overline{y}. \qquad \diamondsuit$$

Definitsioon 3. Kindlat suurust

$$\mathrm{r}^{*}_{xy} \stackrel{\text{def.}}{=} \frac{\mathrm{K}^{*}_{x,\,y}}{s_x\ s_y} \tag{5.2.31}$$

nimetatakse juhuslike suuruste X ja Y *korrelatsioonikordaja* r_{xy} *punkthinnanguks.*

Näide 1. Olgu vektori (X, Y) korral sooritatud 5 sõltumatut katset, mis andsid valimi

$$\{(1;0)\,,\,(2;6)\,,\,(-1,4)\,;\,(1;1)\,,\,(2;7)\}\,.$$

Leiame suurused $\mathrm{K}^{*}_{x,\,y}$ ja r^{*}_{xy} ning valimi korrelatsioonimaatriksi.

Valemi (5.2.2) abil saame leida suurused $\overline{x}$ ja $\overline{y}$:

$$\overline{x} = \frac{1}{5}\sum\nolimits_{i=1}^{5} x_i = \frac{1}{5}(1+2-1+1+2) = 1,$$

$$\overline{y} = \frac{1}{5}\sum\nolimits_{i=1}^{5} y_i = \frac{1}{5}(0+6+4+1+7) = \frac{18}{5}.$$

Valemi (5.2.16) abil saame s_x^2 ja s_y^2 :

$$s_x^2 = \frac{5}{5-1}\left(\frac{1}{5}\sum\nolimits_{i=1}^{5} x_i^2 - \overline{x}^2\right) =$$

$$= \frac{1}{4}\left(1^2 + 2^2 + (-1)^2 + 1^2 + 2^2 - 5\cdot 1\right) = \frac{3}{2},$$

$$s_y^2 = \frac{1}{5-1}\left(\sum\nolimits_{i=1}^{5} y_i^2 - 5\cdot\overline{y}^2\right) =$$

$$= \frac{1}{4}\left(0^2 + 6^2 + 4^2 + 1^2 + 7^2 - 5\cdot\left(\frac{18}{5}\right)^2\right) = \frac{93}{10}.$$

Valemi (5.2.30) põhjal leiame

$$\overline{xy} = \frac{1}{5}\sum\nolimits_{i=1}^{5} x_i y_i = \frac{1}{5}\left(1\cdot 0 + 2\cdot 6 - 1\cdot 4 + 1\cdot 1 + 2\cdot 7\right) = \frac{23}{5}.$$

Valemitest (5.2.29), (5.2.28) ja (5.2.31) saame

$$\mathrm{K}^{**}_{x,\,y} = \overline{xy} - \overline{x}\,\overline{y} = \frac{23}{5} - 1\cdot\frac{18}{5} = 1, \quad \mathrm{K}^{*}_{x,\,y} = \frac{5}{5-1}\mathrm{K}^{**}_{x,\,y} = \frac{5}{4}$$

ja

$$\mathrm{r}^*_{xy} = \frac{\mathrm{K}^*_{x,\,y}}{s_x\ s_y} = \frac{5/4}{\sqrt{(3/2)\cdot(93/10)}} = \frac{5}{186}\sqrt{155} \approx 0.335.$$

Järelikult on valimi korrelatsioonimaatriks

$$\begin{pmatrix} 1 & 0.335 \\ & 1 \end{pmatrix}. \qquad \diamondsuit$$

III Vektori $(X_1, X_2, \ldots, X_m)$ **juht,** kui $\mathrm{E}X_i$ ei ole teada.

Lause 3. Olgu vektori $(X_1, X_2, \ldots, X_m)$ korral sooritatud n sõltumatut katset, mis andsid valimi

i	x_1	x_2	$\cdots$	x_m
1	x_{11}	x_{21}	$\cdots$	x_{m1}
2	x_{12}	x_{22}	$\cdots$	x_{m2}
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
n	x_{1n}	x_{2n}	$\cdots$	x_{mn}

kus x_{ji} on j-nda komponendi väärtus i-ndal katsel. Kui

$$\overline{x}_j \overset{\text{def.}}{=} \frac{1}{n}\sum_{i=1}^{n} x_{ji}, \quad s_j^2 \overset{\text{def.}}{=} \frac{1}{n-1}\sum_{i=1}^{n}\left(x_{ji} - \overline{x}_j\right)^2, \tag{5.2.32}$$

siis

$$\mathrm{K}^*_{x_j,\,x_m} \overset{\text{def.}}{=} \frac{1}{n-1}\sum_{i=1}^{n}\left(x_{ji} - \overline{x}_j\right)\left(x_{mi} - \overline{x}_m\right), \quad \mathrm{r}^*_{x_j,x_m} = \frac{\mathrm{K}^*_{x_j,\,x_m}}{s_{x_j}\ s_{x_m}}. \tag{5.2.33}$$

Näide 2. Olgu vektori (X_1, X_2, X_3) korral sooritatud 5 sõltumatut katset, mis andsid valimi

i	x_1	x_2	x_3
1	2.8	3.5	3.7
2	2.0	3.2	3.0
3	3.1	2.5	2.7
4	2.8	2.9	2.6
5	3.0	2.3	3.1

Leiame valimi kovaritsioonimaatriksi ja korrelatsioonimaatriksi.

Valemite (5.2.32) ja (5.2.33) abil saame $\overline{x}_1 = 2.74$, $\overline{x}_2 = 2.88$, $\overline{x}_3 = 3.02$, $s_1^2 \approx 0.19$, $s_2^2 \approx 0.24$, $s_3^2 \approx 0.19$,

$$\left(\mathrm{K}^*_{x_k,\, x_m}\right) \approx \begin{pmatrix} 0.19 & -0.12 & -0.02 \\ & 0.24 & 0.12 \\ & & 0.19 \end{pmatrix},$$

$$\left(\mathrm{r}^*_{x_k, x_m}\right) \approx \begin{pmatrix} 1.0 & -0.57 & -0.09 \\ & 1.0 & 0.57 \\ & & 1.0 \end{pmatrix}. \qquad \diamondsuit$$

5.2.5 Suurima tõepära meetod

I Suurus X on pidev

Kui suuruse X jaotus sõltub m parameetrist $\alpha_1, \ldots, \alpha_m$, siis suuruse X jaotustihedusel on kuju $f(x; \alpha_1, \ldots, \alpha_m)$. Kuna valimi elemendid on sõltumatud, siis

$$f(x_1, \ldots, x_n; \alpha_1, \ldots, \alpha_m) = \prod_{i=1}^n f(x_i; \alpha_1, \ldots, \alpha_m)$$

on ühelt poolt valimi kui juhusliku vektori jaotustihedus. Teiselt poolt on funktsioon $\prod_{i=1}^n f(x_i; \alpha_1, \ldots, \alpha_m)$ konkreetse valimi korral m muutuja funktsioon. Parameetrite $\alpha_1, \ldots, \alpha_m$ hinnangud $\alpha_1^*, \ldots, \alpha_m^*$ valitakse selliselt, et

$$\max_{(\alpha_1, \alpha_2, \ldots, \alpha_m) \in \mathbf{R}^m} f(x_1, x_2, \ldots, x_n; \alpha_1, \alpha_2, \ldots, \alpha_m) =$$

$$= f(x_1, x_2, \ldots, x_n; \alpha_1^*, \alpha_2^*, \ldots, \alpha_m^*).$$

Kui parameetrite $\alpha_1, \alpha_2, \ldots, \alpha_m$ hinnangud $\alpha_1^*, \alpha_2^*, \ldots, \alpha_m^*$ määrata selliselt, siis kõneldakse, et need hinnangud on saadud *suurima tõepära meetodil.* Kui funktsioon $f(x_1, \ldots, x_n; \alpha_1, \ldots, \alpha_m)$ saavutab maksimaalse väärtuse, siis maksimaalse väärtuse saavutab ka $\ln f(x_1, \ldots, x_n; \alpha_1, \ldots, \alpha_m)$, st on vaja leida m muutuja funktsiooni

$$\ln f(x_1, \ldots, x_n; \alpha_1, \ldots, \alpha_m) = \ln \Pi_{i=1}^n f(x_i; \alpha_1, \ldots, \alpha_m) =$$

$$= \sum_{i=1}^n \ln f(x_i; \alpha_1, \ldots, \alpha_m)$$

statsionaarsed punktid. Seega määratakse parameetrite $\alpha_1, \ldots, \alpha_m$ punkthinnangud $\alpha_1^*, \ldots, \alpha_m^*$ võrrandisüsteemist

$$\frac{\partial}{\partial \alpha_k^*} \sum_{i=1}^n \ln f(x_i; \alpha_1^*, \ldots, \alpha_m^*) = 0 \quad (k = 1; 2; \ldots; m). \tag{5.2.34}$$

Lause 1 (suurima tõepära meetod)**.** Kui $f(x; \alpha_1, \ldots, \alpha_m)$ on pideva juhusliku suuruse X jaotustihedus, siis parameetrite α_i $(i = 1; \ldots; m)$ punkthinnangud α_i^* leitakse võrrandisüsteemist (5.2.34).

Näide 1. Kui $X \sim N(a, \sigma)$, siis

$$f(x_1, \ldots, x_n; a, \sigma) = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n \exp\left(-\frac{\sum_{i=1}^n (x_i - a)^2}{2\sigma^2}\right)$$

ja

$$\ln f(x_1, \ldots, x_n; a, \sigma) = -n \ln\left(\sigma\sqrt{2\pi}\right) - \frac{\sum_{i=1}^n (x_i - a)^2}{2\sigma^2}.$$

Olgu a^* ja σ^* parameetrite a ja σ punkthinnangud. Võrrandisüsteemile (5.2.34) saame kuju

$$\left\{\begin{array}{l} \dfrac{\partial}{\partial a^*}\left(-n\ln\left(\sigma^*\sqrt{2\pi}\right) - \dfrac{\sum_{i=1}^n (x_i - a^*)^2}{2(\sigma^*)^2}\right) = 0 \\ \dfrac{\partial}{\partial \sigma^*}\left(-n\ln\left(\sigma^*\sqrt{2\pi}\right) - \dfrac{\sum_{i=1}^n (x_i - a^*)^2}{2(\sigma^*)^2}\right) = 0 \end{array}\right. \Rightarrow$$

$$\Rightarrow \left\{\begin{array}{l} \dfrac{\sum_{i=1}^n 2(x_i - a^*)}{2(\sigma^*)^2} = 0 \\ -\dfrac{n}{\sigma} + \dfrac{\sum_{i=1}^n (x_i - a^*)^2}{(\sigma^*)^3} = 0 \end{array}\right. \Leftrightarrow \left\{\begin{array}{l} \sum_{i=1}^n x_i = n\, a^* \\ \sum_{i=1}^n (x_i - a^*)^2 = n(\sigma^*)^2 \end{array}\right. \Rightarrow$$

$$\Rightarrow a^* = \frac{1}{n}\sum_{i=1}^n x_i = \overline{x} \wedge (\sigma^*)^2 = \frac{1}{n}\sum_{i=1}^n (x_i - a^*)^2 = m_2. \qquad \diamondsuit$$

II Suurus X on diskreetne

Olgu $P(X = x) = h(x; \alpha_1, \ldots, \alpha_m)$, kus $\alpha_1, \ldots, \alpha_m$ on selle diskreetse jaotuse parameetrid, mille punkthinnanguid $\alpha_1^*, \ldots, \alpha_m^*$ me valimi $\{x_1, \ldots, x_n\}$ põhjal leiame. Seega $P(X_i = x_i) = h(x_i; \alpha_1, \ldots, \alpha_m)$, kus X_i on X i-nda elemendi võtmisel, ja

$$\begin{aligned} P((X_1, \ldots, X_n) = (x_1, \ldots, x_n)) &= \Pi_{i=1}^n P(X_i = x_i) = \\ &= \Pi_{i=1}^n h(x_i, \alpha_1, \ldots, \alpha_m). \end{aligned}$$

Kasutades suurima tõepära meetodit diskreetse jaotuse korral leitakse parameetrite $\alpha_1, \ldots, \alpha_m$ punkthinnangud $\alpha_1^*, \ldots, \alpha_m^*$ selliselt, et

$$\max_{(\alpha_1, \alpha_2, \ldots, \alpha_m) \in \mathbf{R}^m} \Pi_{i=1}^n h(x_i, \alpha_1, \ldots, \alpha_m) =$$

$$= \Pi_{i=1}^n h(x_i, \alpha_1^*, \ldots, \alpha_m^*).$$

Kui funktsioon $\Pi_{i=1}^n h(x_i, \alpha_1^*, \ldots, \alpha_m^*)$ saavutab maksimaalse väärtuse, siis saavutab maksimaalse väärtuse ka

$$\ln \Pi_{i=1}^n h(x_i, \alpha_1^*, \ldots, \alpha_m^*) = \sum_{i=1}^n \ln h(x_i, \alpha_1^*, \ldots, \alpha_m^*).$$

Võrrandisüsteemist

$$\frac{\partial}{\partial \alpha_k^*} \sum_{i=1}^{n} \ln\, h\left(x_i, \alpha_1^*, \ldots, \alpha_m^*\right) = 0 \quad (k = 1; 2; \ldots; m) \tag{5.2.35}$$

leiame hinnangute vektori (vektorid) $\left(\alpha_1^*, \ldots, \alpha_m^*\right)$.

Lause 2. Kui X on diskreetne juhuslik suurus ja

$$P\left(X = x\right) = h\left(x; \alpha_1, \ldots, \alpha_m\right),$$

siis parameetrite α_i punkthinnangud α_i^* $(i = 1; \ldots; m)$ leitakse võrrandisüsteemist (5.2.35).

Näide 2. Olgu suuruse X võimalikud väärtused 1 ja 0 ning

$$P\left(X = 1\right) = p, \quad P\left(X = 0\right) = 1 - p = q.$$

Leiame valimi $\{x_1, \ldots, x_n\}$ põhjal parameetri p punkthinnangu p^*.

Kasutame Lauset 2. Antud jaotusel on üks parameeter, $\alpha_1 = p$. Seejuures $h\left(x, p\right) = p^x \left(1 - p\right)^{1-x}$. Tingimuse (5.2.35) abil saame

$$\frac{d}{d\,p^*} \sum_{i=1}^{n} \ln\left(\left(p^*\right)^{x_i}\left(1 - p^*\right)^{1-x_i}\right) = 0 \;\Leftrightarrow$$

$$\Leftrightarrow\; \frac{d}{d\,p^*}\left(\sum_{i=1}^{n} x_i \ln p^* + \left(1 - x_i\right)\ln\left(1 - p^*\right)\right) = 0 \;\Rightarrow$$

$$\Rightarrow\; \sum_{i=1}^{n}\left(\frac{x_i}{p^*} - \frac{1 - x_i}{1 - p^*}\right) = 0 \Leftrightarrow \frac{1}{p^*}\sum_{i=1}^{n} x_i - \frac{1}{1 - p^*}\sum_{i=1}^{n}\left(1 - x_i\right) = 0 \Leftrightarrow$$

$$\Leftrightarrow\; \left(\frac{1}{p^*} + \frac{1}{1 - p^*}\right)\sum_{i=1}^{n} x_i = \frac{n}{1 - p^*} \;\Leftrightarrow\; p^* = \frac{1}{n}\sum_{i=1}^{n} x_i. \qquad \diamondsuit$$

Näide 3. Olgu X Poissoni jaotusele alluv suurus. Seega on suuruse X võimalikud väärtused mittenegatiivsed täisarvud. Kasutame Lauset 2 parameetri λ hinnangu λ^* leidmiseks.

Antud juhul on jaotusel üks parameeter, $\alpha_1 = \lambda$. Seejuures

$$h\left(x, \lambda\right) = \left(\lambda^x / \left(x!\right)\right)\exp\left(-\lambda\right).$$

Tingimuse (5.2.35) abil saame

$$\frac{d}{d\lambda^*}\sum_{i=1}^{n} \ln\left(\exp\left(-\lambda^*\right)\frac{\left(\lambda^*\right)^{x_i}}{x_i!}\right) = 0 \;\Leftrightarrow$$

$$\Leftrightarrow\; \frac{d}{d\lambda^*}\sum_{i=1}^{n}\left(-\lambda^* + x_i \ln \lambda^* - \ln x_i!\right) = 0 \;\Rightarrow$$

$$\Rightarrow\; \textstyle\sum_{i=1}^{n}\left(-1 + \dfrac{x_i}{\lambda^*}\right) = 0 \Leftrightarrow \lambda^* = \dfrac{1}{n}\textstyle\sum_{i=1}^{n} x_i = \overline{x}. \qquad \diamondsuit$$

Nii süsteemi (5.2.34) kui ka (5.2.35) nimetatakse *tõepära võrrandite süsteemiks.* Kui $k = 1$, siis kõneldakse *tõepäravõrrandist.* Kehtib järgmine väide.

Lause 3. Kui eksisteerib parameetri α efektiivne punkthinnang α^*, siis tõepäravõrrand on üheselt lahenduv ja lahendiks on see efektiivne hinnang.

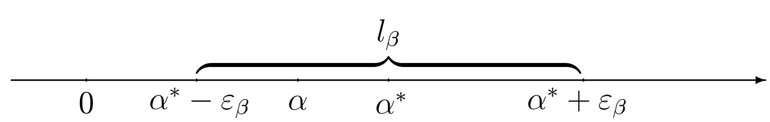
Märkus 1. Lause 3 ei anna mingit infot, kui parameetril α ei ole efektiivset punkthinnangut.

5.3 Vahemikhinnangud

Olgu α juhusliku suuruse X jaotuse parameeter ja $\alpha^* = \alpha^*(x_1, \ldots, x_n)$ selle parameetri α hinnang. Olgu $\varepsilon > 0$. Suurus α^* on juhuslik suurus. Seega on juhuslikud ka vahemiku $(\alpha^* - \varepsilon,\, \alpha^* + \varepsilon)$ otspunktid. Vaatleme seost

$$P(a \in (\alpha^* - \varepsilon,\, \alpha^* + \varepsilon)) = \beta \tag{5.3.1}$$

ehk lühidalt $P(\alpha \in l_\beta) = \beta$, kus $l_\beta = (\alpha^* - \varepsilon,\, \alpha^* + \varepsilon)$. Tõlgendame seost $\alpha \in (\alpha^* - \varepsilon,\, \alpha^* + \varepsilon)$ kui sündmust, et juhuslike otspunktidega vahemik l_β katab punkti α. Seoses (5.3.1) esinevat arvu β nimetatakse *usaldusnivooks* ja vahemikku l_β *usaldusvahemikuks* ehk *usalduspiirkonnaks* ning vahemiku l_β otspunkte $\alpha^* - \varepsilon$ ja $\alpha^* + \varepsilon$ *usalduspiirideks.* Avaldis (5.3.1) seob suurusi β ja ε. Saame suuruse ε suuruse β funktsioonina $\varepsilon = \varepsilon(\beta) \equiv \varepsilon_\beta$.



Seega on seos (5.3.1) esitatav kujul

$$P(\alpha \in (\alpha^* - \varepsilon_\beta,\, \alpha^* + \varepsilon_\beta)) = \beta. \tag{5.3.2}$$

Kui

$$P(\alpha \le \alpha^* - \varepsilon_\beta) = P(\alpha \ge \alpha^* + \varepsilon_\beta) = (1 - \beta)/2,$$

siis l_β usalduspiire nimetatakse *sümmeetrilisteks usalduspiirideks* ja usalduspiirkonda *sümmeetriliseks usaldusvahemikuks.* Kui kasutada *ühepoolseid usaldusvahemikke*, siis *vasakpoolne usaldusvahemik* $(-\infty, \alpha_v)$ usaldusnivooga β määratakse seosest $P(\alpha < \alpha_v) = \beta$ ja *parempoolne usaldusvahemik* $(\alpha_p, +\infty)$ seosest $P(\alpha > \alpha_p) = \beta$. Harilikult valitakse $\beta = 0.90, \ldots, 0.95$. Kasutatakse *ligikaudseid usaldusvahemikke* ja *täpseid usaldusvahemikke.* Täpsete usaldusvahemike leidmisel on vajalik lisainfo X jaotuse kohta.

5.3.1 Üldkogumi keskväärtuse usaldusvahemik

Valimi keskmine $\overline{x}$ on defineeritud seosega (5.2.2). Suuruse $\overline{x}$ keskväärtuseks on $\mathrm{E}X$ ja $\mathrm{D}\overline{x} = \mathrm{D}X/n = \sigma^2/n$. Suurus $\overline{x}$ on tsentraalse piirteoreemi põhjal asümptootiliselt normaalne. Seega allub $\overline{x}$ valimi piisavalt suure mahu korral ligikaudu normaaljaotusele parameetritega $\mathrm{E}X$ ja $\sqrt{\mathrm{D}\overline{x}} = \sigma/\sqrt{n}$. Vaatleme seoste ahelat

$$P(\mathrm{E}X \in (\overline{x}-\varepsilon_\beta, \overline{x}+\varepsilon_\beta)) = \beta \ \Leftrightarrow$$

$$\Leftrightarrow\ P(\mathrm{E}X-\varepsilon_\beta < \overline{x} < \mathrm{E}X+\varepsilon_\beta)) = \beta \ \Leftrightarrow$$

$$\Leftrightarrow\ \Phi\left(\frac{\mathrm{E}X+\varepsilon_\beta-\mathrm{E}X}{\sigma/\sqrt{n}}\right) - \Phi\left(\frac{\mathrm{E}X-\varepsilon_\beta-\mathrm{E}X}{\sigma/\sqrt{n}}\right) \approx \beta \ \Leftrightarrow$$

$$\Leftrightarrow\ \Phi\left(\frac{\varepsilon_\beta\sqrt{n}}{\sigma}\right) - \Phi\left(\frac{-\varepsilon_\beta\sqrt{n}}{\sigma}\right) \approx \beta \ \Leftrightarrow 2\Phi\left(\frac{\varepsilon_\beta\sqrt{n}}{\sigma}\right) \approx \beta \ \Leftrightarrow$$

$$\Leftrightarrow\ \Phi\left(\frac{\varepsilon_\beta\sqrt{n}}{\sigma}\right) \approx \frac{\beta}{2} \ \Leftrightarrow \frac{\varepsilon_\beta\sqrt{n}}{\sigma} \approx \Phi^{-1}\left(\frac{\beta}{2}\right) \ \Leftrightarrow\ \varepsilon_\beta \approx \frac{\sigma}{\sqrt{n}}\Phi^{-1}\left(\frac{\beta}{2}\right).$$

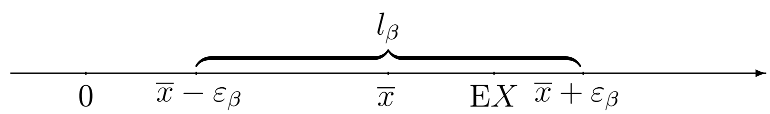
Et $\sigma \approx s$, siis

$$\varepsilon_\beta \approx \frac{s}{\sqrt{n}}\Phi^{-1}\left(\frac{\beta}{2}\right) \tag{5.3.1.1}$$

ja

$$l_\beta \approx \left(\overline{x} - \frac{s}{\sqrt{n}}\Phi^{-1}\left(\frac{\beta}{2}\right), \overline{x} + \frac{s}{\sqrt{n}}\Phi^{-1}\left(\frac{\beta}{2}\right)\right) \tag{5.3.4}$$

on usaldusnivoole β vastav ligikaudne usaldusvahemik



ning

$$P\left(\mathrm{E}X \in \left(\overline{x} - \frac{s}{\sqrt{n}}\Phi^{-1}\left(\frac{\beta}{2}\right), \overline{x} + \frac{s}{\sqrt{n}}\Phi^{-1}\left(\frac{\beta}{2}\right)\right)\right) \approx \beta. \tag{5.3.5}$$

Lause 1. Valimi piisavalt suure mahu korral on keskväärtuse $\mathrm{E}X$ ligikaudne sümmeeetriline usaldusvahemik leitav valemi (5.3.4) abil.

Näide 1. Olgu valimi $\{x_1, x_2, \ldots, x_{25}\}$ põhjal leitud, et $\overline{x} = 3$ ja $s = 1$. Leiame usaldusnivoole 0.92 vastava keskväärtuse sümmeetrilise usaldusvahemiku $l_{0.92}$.

Et valimi maht n on piisavalt suur, siis võime rakendada Lauset 1. Valemi (5.3.3) abil leiame

$$\varepsilon_{0.92} \approx \frac{1}{\sqrt{25}}\Phi^{-1}\left(\frac{0.92}{2}\right) = \frac{1}{5}\Phi^{-1}(0.46) \approx 0.2 \cdot 1.75 = 0.35.$$

Seega on $l_{0.92} \approx (2.65; 3.35)$ Lause 1 põhjal keskväärtuse $\mathrm{E}X$ sümmeetriline usaldusvahemik. $\diamondsuit$

5.3.2 Normaaljaotusele alluva üldkogumi keskväärtuse usalduspiirkond

Kui X on normaaljaotusega, siis juhuslik suurus

$$T_{n-1} = \sqrt{n}\left(\overline{x} - \mathrm{E}X\right)/s \tag{5.3.6}$$

on Lause 5.2.3.2 põhjal Studenti jaotusega, mille vabadusastmete arv on $n-1$. Kuna

$$P\left(|\overline{x} - \mathrm{E}X| < \varepsilon_\beta\right) = \beta \;\Leftrightarrow\; P\left(\sqrt{n}\frac{|\overline{x} - \mathrm{E}X|}{s} < \frac{\sqrt{n}\varepsilon_\beta}{s}\right) = \beta \;\Leftrightarrow$$

$$\Leftrightarrow\; P\left(|T_{n-1}| < \frac{\sqrt{n}\varepsilon_\beta}{s}\right) = \beta \;\Leftrightarrow\; \int_{-\varepsilon_\beta\sqrt{n}/s}^{\varepsilon_\beta\sqrt{n}/s} f_{T_{n-1}}(t)dt = \beta \;\overset{\varepsilon_\beta\sqrt{n}/s = t_\beta}{\Leftrightarrow}$$

$$\Leftrightarrow\; 2\int_0^{t_\beta} f_{T_{n-1}}(t)dt = \beta \;\overset{SWP30}{\Rightarrow}\; t_\beta = \mathrm{TInv}\left(0.5 + \beta/2; n-1\right),$$

siis

$$\varepsilon_\beta = st_\beta/\sqrt{n} \tag{5.3.7}$$

ja

$$l_\beta = \left(\overline{x} - st_\beta/\sqrt{n}, \, \overline{x} + st_\beta/\sqrt{n}\right) \tag{5.3.8}$$

on usaldusnivoole β vastav usaldusvahemik, st kehtib seos

$$P\left(\overline{x} - st_\beta/\sqrt{n} < \mathrm{E}X < \overline{x} + st_\beta/\sqrt{n}\right) = \beta. \tag{5.3.9}$$

Märkus 1. Studenti jaotuse tabeleid on kahte tüüpi

$$2\int_0^{t_\beta} f_{T_n}(t)dt = \beta \;\Rightarrow\; (\beta, n) \mapsto t_\beta$$

või

$$2\int_{t_\alpha}^{+\infty} f_{T_n}(t)dt = \alpha \;\Rightarrow\; (\alpha, n) \mapsto t_\alpha,$$

kus $\beta = 1 - \alpha$.

Lause 1. Kui X on normaaljaotusega, siis keskväärtuse $\mathrm{E}X$ sümmeetriline usaldusvahemik on leitav valemi (5.3.8) abil.

Näide 1. Olgu saadud valim $\{-2.5; 3.4; -2.0; 1.0; 2.1\}$. Leiame $\mathrm{E}X$ sümmeetrilise usaldusvahemiku usaldusnivool $\beta = 0.9$.

Arvutame, $\overline{x} = 0.4$ ja $s^2 = 6.6$ ning $n - 1 = 4$. SWP abil leiame $t_{0.9} = \mathrm{TInv}(0.95; 4) = 2.13$. Valemi (5.3.7) abil saame $\varepsilon_{0.9} = 2.13\sqrt{6.6}/\sqrt{4} \approx 2.736$. Lause 1 põhjal on $l_{0.9} = (-2.336; 3.136)$ keskväärtuse $\mathrm{E}X$ usaldusvahemik, mis vastab usaldusnivoole 0.9. $\diamondsuit$

5.3.3 Normaaljaotusele alluva üldkogumi dispersiooni usalduspiirkond

Kui X on normaaljaotusega, siis suurus $Y_{n-1} = (n-1)\,s^2/\mathrm{D}X$ on χ^2-jaotusega juhuslik suurus, mille vabadusastmete arv on $n-1$. Lähtume seosest

$$P\,(v_1 < Y_{n-1} < v_2) = \beta,$$

kus

$$P\,(Y_{n-1} < v_1) = \frac{1-\beta}{2} \Rightarrow v_1 = \mathrm{ChiSquareInv}(\frac{1-\beta}{2}, n-1) \tag{5.3.10}$$

ja

$$P\,(Y_{n-1} < v_2) = \frac{1+\beta}{2} \Rightarrow v_2 = \mathrm{ChiSquareInv}(\frac{1+\beta}{2}, n-1). \tag{5.3.11}$$

Kuna

$$P\,(v_1 < Y_{n-1} < v_2) = \beta \;\Leftrightarrow\; P\left(\frac{1}{v_2} < \frac{1}{Y_{n-1}} < \frac{1}{v_1}\right) = \beta \;\Leftrightarrow$$

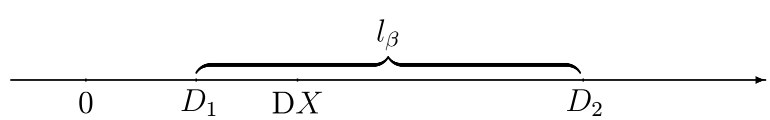
$$\Leftrightarrow P\left(\frac{(n-1)\,s^2}{v_2} < \frac{(n-1)\,s^2}{Y_{n-1}} < \frac{(n-1)\,s^2}{v_1}\right) = \beta \Leftrightarrow$$

$$\Leftrightarrow P\left(\frac{(n-1)\,s^2}{v_2} < \mathrm{D}X < \frac{(n-1)\,s^2}{v_1}\right) = \beta,$$

siis

$$l_\beta = \left(\frac{(n-1)\,s^2}{v_2}, \frac{(n-1)\,s^2}{v_1}\right) \tag{5.3.12}$$

on dispersiooni DX sümmeetriline usaldusvahemik usaldusnivool β. Seega



$$P(D_1 < \mathrm{D}X < D_2) = \beta,$$

kus

$$D_1 = (n-1)\,s^2/v_2,\; D_2 = (n-1)\,s^2/v_1. \tag{5.3.13}$$

Lause 1. Kui X on normaaljaotusega, siis dispersiooni DX sümmeetriline usaldusvahemik on leitav valemi (5.3.12) abil.

Näide 3. Valimi $\{-2.5; 3.4; -2.0; 1.0; 2.1\}$ korral $s^2 = 6.6$ ja $n-1 = 4$. Leiame DX usaldusvahemiku usaldusnivool $\beta = 0.9$.

Seoste (5.3.10) ja (5.3.13) abil leiame

$$v_1 = 0.711,\ D_2 = \frac{4 \cdot 6.6}{0.711} \approx 37.13$$

ning seoste (5.3.11) ja (5.3.13) abil

$$v_2 = 9.49,\ D_1 = \frac{4 \cdot 6.6}{9.49} \approx 2.78.$$

Lause 1 põhjal on $l_{0.9} \approx (2.78; 37.13)$ dispersiooni DX sümmeetriline usaldusvahemik usaldusnivool 0.9. $\diamondsuit$

5.4 Hüpoteeside statistiline kontrollimine

Järgneva aluseks on *praktilise kindluse printsiip.* Nimelt, kui katse sooritamisel on sündmuse A toimumise tõenäosus väga väike, siis katse ühekordsel toimumisel võib olla kindel, et sündmus A ei toimu. Praktilise kindluse printsiip on aluseks inimese igapäevasel tegevusel. Näiteks lennukisse minnes me ei arvesta, et lennuk teeb avarii.

Definitsioon 1. Iga oletust tundmatu jaotusseaduse kuju või parameetrite kohta nimetatakse *(statistiliseks) hüpoteesiks.*

Definitsioon 2. Iga hüpoteesi, mis määrab jaotusseaduse täielikult, nimetatakse *lihthüpoteesiks.* Hüpoteesi, mis ei ole lihthüpotees, nimetatakse *liithüpoteesiks.*

Kontrollitavat hüpoteesi nimetatakse tavaliselt *nullhüpoteesiks* ja tähistatakse H_0. Kõrvuti nullhüpoteesiga vaadeldakse *konkureerivat* ehk *alternatiivset* hüpoteesi H_1, st H_0 ja H_1 on teineteist välistavad.

Hüpoteesi H_0 kontrollimiseks kasutatakse valimi $\{x_1, x_2, \ldots, x_n\}$ põhjal spetsiaalselt koostatud statistikut $\theta_n^*(x_1, x_2, \ldots, x_n)$, mille kui juhusliku suuruse täpne või ligikaudne jaotus on teada. Statistiku θ_n^* kõigi võimalike väärtuste hulk Δ jaotatakse kaheks mittelõikuvaks osahulgaks: *kriitiliseks hulgaks* Δ_1 (hüpoteesi H_0 tagasilükkamise piirkond) ja *lubatud hulgaks* Δ_0 (hüpoteesi H_0 vastuvõtmise piirkond). Valimi jaotuse põhjal määratakse Δ_1 selliselt, et kui hüpotees H_0 on õige, siis $P(\theta_n^* \in \Delta_1) = \alpha$, kus α on etteantud väike arv. Kui α on "piisavalt väike," siis sündmust $\theta_n^* \in \Delta_1$ võib praktilise kindluse printsiibi põhjal lugeda praktiliselt võimatuks ja sündmuse $\theta_n^* \in \Delta_1$ toimumisel hüpotees H_0 lükatakse tagasi. Kui $\theta_n^* \in \Delta_0$, siis hüpotees H_0 võetakse vastu, täpsemini, seda ei lükata tagasi. Reeglit, mille järgi hüpotees lükatakse tagasi või võetakse vastu, nimetatakse *statistiliseks kriteeriumiks.*

Lihtsamad kriitilised hulgad Δ_1 on: *parempoolne kriitiline hulk* $(\theta_{kr}, +\infty)$, *vasakpoolne kriitiline hulk* $(-\infty, \theta_{kr})$, *kahepoolne kriitiline hulk* $(-\infty, \theta_{krv}) \cup (\theta_{krp}, +\infty)$, kusjuures

$$P(\theta_n^* \in (-\infty, \theta_{krv})) = P(\theta_n^* \in (\theta_{krp}, \infty)),$$

ja sümmeetriline kriitiline hulk

$$(-\infty, -\theta_{kr}) \cup (\theta_{kr}, +\infty).$$

Seega parempoolse kriitilise hulga korral

$$\begin{aligned} \theta_n^* > \theta_{kr} \;&\Rightarrow\; \text{hüpotees } H_0 \text{ lükatakse tagasi,} \\ \theta_n^* < \theta_{kr} \;&\Rightarrow\; \text{hüpotees } H_0 \text{ võetakse vastu} \end{aligned} \tag{5.4.1}$$

ja vasakpoolse kriitilise hulga korral

$$\begin{aligned} \theta_n^* < \theta_{kr} \;&\Rightarrow\; \text{hüpotees } H_0 \text{ lükatakse tagasi,} \\ \theta_n^* > \theta_{kr} \;&\Rightarrow\; \text{hüpotees } H_0 \text{ võetakse vastu.} \end{aligned} \tag{5.4.2}$$

Kahepoolse kriitilise hulga korral

$$\begin{aligned} (\theta_n^* < \theta_{krv}) \,\vee\, (\theta_n^* > \theta_{krp}) \Rightarrow \text{ hüpotees } H_0 \text{ lükatakse tagasi,} \\ \theta_{krv} < \theta_n^* < \theta_{krp} \;\Rightarrow\; \text{hüpotees } H_0 \text{ võetakse vastu} \end{aligned} \tag{5.4.3}$$

ja sümmeetrilise kriitilise hulga korral

$$\begin{aligned} |\theta_n^*| > \theta_{kr} \;&\Rightarrow\; \text{hüpotees } H_0 \text{ lükatakse tagasi,} \\ |\theta_n^*| < \theta_{kr} \;&\Rightarrow\; \text{hüpotees } H_0 \text{ võetakse vastu.} \end{aligned} \tag{5.4.4}$$

Statistilise kriteeriumi korral on neli võimalikku juhtu

hüpotees H_0	võetakse vastu	lükatakse tagasi
õige	õige otsustus	esimest liiki viga
vale	teist liiki viga	õige otsustus

Definitsioon 3. Esimest liiki vea lubatavuse tõenäosust α nimetatakse k*riteeriumi olulisuse nivooks.*

Olgu teist liiki vea lubatavuse tõenäosus β.

Definitsioon 4. Teist liiki vea mittelubatavuse tõenäosust $1-\beta$ nimetatakse *kriteeriumi võimsuseks.*

Kasutades statistilise kontrolli terminoloogiat pooljuhtide tootmisel, on H_0 hüpotees, et partii vastab nõuetele, ja α on tootja risk, et ta kannab maha kogu partii, kui pistelise kontrolli näitajad ei vasta lubatule. Samas on β tarbija risk võtta positiivse pistelise kontrolli tulemuste põhjal vastu nõuetele mittevastav partii. Kahe erineva vea võimalus eristab hüpoteeside kontrolli parameetrite vahemikhinnanguist. Tõenäosused α ja β on üheselt määratud kriitilise piirkonna valikuga. Et

$$\alpha = P\left((\theta_n^* \in \Delta_1)\,(H_0 \text{ on õige})\right), \tag{5.4.5}$$

siis suuruse α vähendamiseks tuleb muuta hulka Δ_1 väiksemaks, st suurendada hulka Δ_0. Aga suuruse

$$\beta = P\left((H_0 \text{ on vale})\,(\theta_n^* \in \Delta_0)\right) \tag{5.4.6}$$

vähendamine nõuab vastupidi hulga Δ_0 vähendamist. Seega ei saa suurusi α ja β samaaegselt vähendada, st valimi mahtu suurendamata võib kui tahes väikeseks teha vaid ühe neist.

Millistest printsiipidest lähtudes valida kriitilist piirkonda Δ_1? Parima tulemuse saamiseks tuleb etteantud olulisuse nivoo α korral valida kriitiline piirkond selliselt, et kriteeriumi võimsus $1-\beta$ oleks maksimaalne.

5.4.1 Kahe jaotuse keskväärtuste võrdsuse kontrollimine

Olgu teada DX ja DY ning sõltumatud valimid

$$\{x_1, x_2, \ldots, x_n\}\,,\ \{y_1, y_2, \ldots, y_m\}\,.$$

Kontrollime keskväärtuste EX ja EY võrdsust, st meil on vaja kontrollida hüpoteesi

$$H_0 : \mathrm{E}X = \mathrm{E}Y$$

olulisuse nivool α, kasutades sümmeetrilist kriitilist hulka. Leiame antud valimite korral $\overline{x}$ ja $\overline{y}$. Tsentraalse piirteoreemi põhjal on $\overline{x}$ ja $\overline{y}$ asümptootiliselt normaalsed. Lause 5.2.1.1 abil saame

$$\mathrm{E}\overline{x} = \mathrm{E}X,\ \mathrm{E}\overline{y} = \mathrm{E}Y,\ \sigma_{\overline{x}} = \sigma_X \,/\sqrt{n},\ \sigma_{\overline{y}} = \sigma_Y \,/\sqrt{m}.$$

Seega ligikaudu

$$\overline{x} \sim N\left(\mathrm{E}X, \sigma_X \,/\sqrt{n}\right), \quad \overline{y} \sim N\left(\mathrm{E}Y, \sigma_Y \,/\sqrt{m}\right).$$

Kui H_0 on õige, siis ligikaudu (miks?)

$$\overline{x} - \overline{y} \sim N\left(0, \sqrt{\sigma_X^2 \,/n + \sigma_Y^2 \,/m}\right) \tag{5.4.7}$$

ja statistik

$$\theta^* = \frac{\overline{x} - \overline{y}}{\sqrt{\sigma_X^2 \,/n + \sigma_Y^2 \,/m}} \sim N\,(0;1)\,. \tag{5.4.8}$$

Arvutame valimite põhjal suuruse θ^*. Kriitilise piirkonna valik sõltub konkureerivast hüpoteesist H_1.

1. Kui $H_1 : \mathrm{E}X \neq \mathrm{E}Y$, siis kasutame sümmeetrilist kriitilist hulka, kusjuures suuruse θ_{kr} määrame tingimusest

$$\Phi(\theta_{kr}) = 0.5 - \alpha/2. \tag{5.4.9}$$

Seejuures $\Phi(x) = \left(1/\sqrt{2\pi}\right)\int_0^x e^{-t^2/2}dt$. Kui $|\theta^*| > \theta_{kr}$, siis lükkame hüpoteesi H_0 tagasi. Kui $|\theta^*| < \theta_{kr}$, siis loeme hüpoteesi H_0 õigeks. Täpsemini, loeme hüpoteesi katseandmetega kooskõlas olevaks.

2. Kui $H_1 : \mathrm{E}X > \mathrm{E}Y$, siis kasutame parempoolset kriitilist hulka, kusjuures suuruse θ_{kr} määrame tingimusest

$$\Phi(\theta_{kr}) = 0.5 - \alpha. \tag{5.4.10}$$

Kui $\theta^* > \theta_{kr}$, siis lükkame H_0 tagasi. Kui $\theta^* < \theta_{kr}$, siis loeme H_0 õigeks.

3. Kui $H_1 : \mathrm{E}X < \mathrm{E}Y$, siis kasutame vasakpoolset kriitilist hulka, kusjuures suuruse θ_{kr} määrame tingimusest (5.4.10). Kui $\theta^* < -\,\theta_{kr}$, siis lükkame H_0 tagasi. Kui $\theta^* > -\theta_{kr}$, siis loeme H_0 õigeks.

Näide 1. Olgu $\mathrm{D}X = 80$ ja $\mathrm{D}Y = 70$. Olgu sõltumatute valimite

$$\{x_1, x_2, \ldots, x_{50}\}, \ \{y_1, y_2, \ldots, y_{40}\}$$

põhjal leitud, et $\overline{x} = 60$ ja $\overline{y} = 67$. Kontrollime keskväärtuste $\mathrm{E}X$ ja $\mathrm{E}Y$ võrdsuse hüpoteesi

$$H_0 : \mathrm{E}X = \mathrm{E}Y$$

õigsust olulisuse nivool $\alpha = 0.02$.

Leiame (5.4.8) abil

$$\theta^* = \frac{\overline{x} - \overline{y}}{\sqrt{\sigma_X^2/n + \sigma_Y^2/m}} \approx -3.82.$$

Kriitilise hulga valik sõltub konkureeriva hüpoteesi H_1 valikust.

1. Kui $H_1 : \mathrm{E}X \neq \mathrm{E}Y$, siis tingimuse (5.4.9) põhjal

$$\Phi(\theta_{kr}) = 0.49 \Leftrightarrow \theta_{kr} = 2.33.$$

Et $|\theta^*| > \theta_{kr}$, siis lükkame hüpoteesi H_0 tagasi ehk loeme hüpoteesi H_1 kehtivaks.

2. Kui $H_1 : \mathrm{E}X > \mathrm{E}Y$, siis tingimuse (5.4.10) põhjal

$$\Phi(\theta_{kr}) = 0.48 \Leftrightarrow \theta_{kr} \approx 2.054.$$

Et $\theta^* < \theta_{kr}$, siis selle alternatiivse hüpoteesi korral on hüpotees H_0 õige.

3. Kui $H_1 : \mathrm{E}X < \mathrm{E}Y$, siis tingimuse (5.4.10) põhjal $\theta_{kr} \approx 2.054$. Et $\theta^* < -\,\theta_{kr}$, siis selle alternatiivse hüpoteesi korral lükkame hüpoteesi H_0 tagasi.

Kuidas mõtestada saadud tulemusi? $\diamondsuit$

5.4.2 Binoomjaotuse parameetrite võrdlemine

1° Olgu $P(A_1) = p_1$, $P(A_2) = p_2$, α olulisuse nivoo ja $H_0 : p_1 = p_2$. Suuruse p_1 määramiseks sooritatakse n_1 sõltumatut katset, kusjuures sündmus A_1 toimub m_1 katse korral. Suuruse p_2 määramiseks sooritatakse n_2 sõltumatut katset, kusjuures sündmus A_2 toimub m_2 katse korral. Saame sagedused $\omega_1 = \dfrac{m_1}{n_1}$ ja $\omega_2 = \dfrac{m_2}{n_2}$, mis piisavalt suurte n_1 ja n_2 korral juhuslike suurustena alluvad Lause 2.11.1 põhjal ligikaudu vastavalt normaaljaotustele $N\left(p_1, \sqrt{p_1(1-p_1)/n_1}\right)$ ja $N\left(p_2, \sqrt{p_2(1-p_2)/n_2}\right)$. Kui H_0 on tõene, st $p_1 = p_2 = p$, siis $\omega_1 - \omega_2$ allub ligikaudu normaaljaotusele

$$N\left(0, \sqrt{\widehat{p}\,(1-\widehat{p})\,(1/n_1 + 1/n_2)}\right)$$

(miks?) ja statistik

$$\theta^* = \frac{\omega_1 - \omega_2}{\sigma_{\omega_1 - \omega_2}} = (\omega_1 - \omega_2) / \sqrt{\widehat{p}\,(1 - \widehat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} \tag{5.4.11}$$

allub ligikaudu normaaljaotusele $N\,(0;\,1)$, kusjuures $\widehat{p} = \dfrac{m_1 + m_2}{n_1 + n_2}$. Kriitilise piirkonna valik sõltub konkureerivast hüpoteesist H_1.

1. Kui $H_1:\ p_1 \neq p_2$, siis kasutame sümmeetrilist kriitilist hulka, kusjuures suuruse θ_{kr} määrame tingimusest (5.4.9). Kui $|\theta^*| > \theta_{kr}$, siis lükkame hüpoteesi H_0 tagasi. Kui $|\theta^*| < \theta_{kr}$, siis loeme hüpoteesi õigeks.

2. Kui $H_1:\ p_1 > p_2$, siis kasutame parempoolset kriitilist hulka, kusjuures suuruse θ_{kr} määrame tingimusest (5.4.10). Kui $\theta^* > \theta_{kr}$, siis lükkame H_0 tagasi. Kui $\theta^* < \theta_{kr}$, siis loeme H_0 õigeks.

3. Kui $H_1:\ p_1 < p_2$, siis kasutame vasakpoolset kriitilist hulka, kusjuures suuruse θ_{kr} määrame tingimusest (5.4.10). Kui $\theta^* < -\,\theta_{kr}$, siis lükkame H_0 tagasi. Kui $\theta^* > -\theta_{kr}$, siis loeme H_0 õigeks.

Näide 1. Kontrolltöös kasutati kaht varianti. Esimest varianti kirjutas 105 üliõpilast, kellest 65 said läbi. Teist kirjutas 140 üliõpilast, kellest 69 said läbi. Kontrollime olulisuse nivool 0.04 hüpoteesi H_0 : variandid on sama raskusega.

Leiame

$$\omega_1 = \frac{65}{105},\ \omega_2 = \frac{69}{140},\ \widehat{p} = \frac{65 + 69}{105 + 140} = \frac{134}{245}$$

ja

$$\theta^* \overset{(5.4.11)}{=} \frac{\dfrac{65}{105} - \dfrac{69}{140}}{\sqrt{\dfrac{134}{245} \cdot \dfrac{111}{245}\left(\dfrac{1}{105} + \dfrac{1}{140}\right)}} \approx 1.964.$$

Kriitilise piirkonna valik sõltub konkureerivast hüpoteesist H_1.

1. Kui $H_1 : p_1 \neq p_2$, siis tingimuse (5.4.9) põhjal

$$\Phi(\theta_{kr}) = 0.48 \Leftrightarrow \theta_{kr} = 2.05.$$

Et $|\theta^*| < \theta_{kr}$, siis selle konkureeriva hüpoteesi korral on hüpotees H_0 õige.

2. Kui $H_1 : p_1 > p_2$, siis tingimuse (5.4.10) põhjal

$$\Phi(\theta_{kr}) = 0.46 \Leftrightarrow \theta_{kr} \approx 1.75.$$

Et $\theta^* > \theta_{kr}$, siis lükkame hüpoteesi H_0 tagasi.

3. Kui $H_1 : p_1 < p_2$, siis tingimuse (5.4.10) põhjal $\theta_{kr} \approx 1.75$. Et $\theta^* > -\,\theta_{kr}$, siis selle konkureeriva hüpoteesi korral on hüpotees H_0 õige. ♢

2° Kui on tegemist l sündmusega A_i $(i = 1; \ldots; l)$ ja $P(A_i) = p_i$ ning $H_0:\ p_1 = p_2 = \ldots = p_l$, siis moodustatakse statistik

$$q^* = \frac{1}{\widehat{p}(1 - \widehat{p})} \sum_{i=1}^{l} n_i\,(\omega_i - \widehat{p})^2\,, \tag{5.4.12}$$

mis on piisavalt suurte mahtude n_i korral ligikaudu χ^2-jaotusega, mille vabadusastmete arv on $l-1$. Seejuures

$$\omega_i = \frac{m_i}{n_i} \;\; (i = 1; \ldots ; l)\,, \quad \widehat{p} = \frac{\sum_{i=1}^{l} m_i}{\sum_{i=1}^{l} n_i}. \tag{5.4.13}$$

Olgu $H_1 = \overline{H}_0$. Olulisuse nivool α kasutatakse parempoolset kriitilist hulka. Arvuti või Lisa 3 abil leitakse

$$(\alpha,\, l-1) \mapsto q_{\alpha,\, l-1}.$$

Hüpotees lükatakse tagasi, kui $q^* > q_{\alpha,\, l-1}$.

Näide 2. Kontrolltööl kasutati nelja varianti. Esimest varianti kirjutas 105 üliõpilast, kellest 61 said läbi. Teise variandi korral olid need arvud 140 ja 69, kolmanda variandi korral 130 ja 72 ning neljanda variandi korral 90 ja 71. Kas kriteeriumi olulisuse nivool 0.02 on variandid sama raskusega?

Olgu $H_0 : p_1 = p_2 = p_3 = p_4$ ja $H_1 = \overline{H}_0$, kus $\overline{H}_0$ on hüpoteesi H_0 vastandhüpotees. Et

$$\begin{aligned} \omega_1 &= \frac{61}{105},\ \omega_2 = \frac{69}{140},\ \omega_3 = \frac{72}{130},\ \omega_4 = \frac{71}{90} \;\; (i = 1; \ldots ; l)\,, \\ \widehat{p} &= \frac{61+69+72+71}{105+140+130+90} = \frac{91}{155} \end{aligned}$$

ja

$$\begin{aligned} n_1\,(\omega_1 - \widehat{p})^2 &= 105\left(\frac{61}{105} - \frac{91}{155}\right)^2 \approx 0.004, \\ n_2\,(\omega_2 - \widehat{p})^2 &= 140\left(\frac{69}{140} - \frac{91}{155}\right)^2 \approx 1.243\,, \\ n_3\,(\omega_3 - \widehat{p})^2 &= 130\left(\frac{72}{130} - \frac{91}{155}\right)^2 \approx 0.144, \\ n_4\,(\omega_4 - \widehat{p})^2 &= 90\left(\frac{71}{90} - \frac{91}{155}\right)^2 \approx 3.665, \end{aligned}$$

siis

$$q^* \overset{(5.4.12)}{=} \frac{1}{\widehat{p}(1-\widehat{p})} \sum_{i=1}^{4} n_i\,(\omega_i - \widehat{p})^2 \approx 20.86.$$

Kuna

$$(0.02;\ 4-1) \mapsto q_{0.02;\ 3} \overset{\text{Lisa 3}}{\approx} 9.84,$$

siis $q^* > q_{0.02;\ 3}$. Hüpotees H_0 lükatakse tagasi. $\diamondsuit$

5.4.3 Normaaljaotuste dispersioonide võrdlemine

1° Olgu X_1 ja X_2 kaks normaaljaotusega suurust dispersioonidega σ_1^2 ja σ_2^2. Võtame valimid vastavalt mahtudega n_1 ja n_2. Hüpoteesi $H_0 : \sigma_1^2 = \sigma_2^2$ kontrollimine taandub valimite dispersioonide s_1^2 ja s_2^2 võrdlemisele. Olgu $s_1^2 \geq s_2^2$. Kui $\sigma_1^2 = \sigma_2^2 = \sigma^2$, siis Lause 5.2.3.1 põhjal on juhuslikud suurused $\dfrac{(n_1-1)\,s_1^2}{\sigma^2}$ ja $\dfrac{(n_2-1)\,s_2^2}{\sigma^2}$ χ^2-jaotusega vastavalt vabadusastmete arvudega $n_1 - 1$ ja $n_2 - 1$. Järelikult (vt Definitsiooni 3.10.1) on esitatud tingimustel statistik

$$z^* = \frac{\dfrac{(n_1-1)\,s_1^2}{\sigma^2} \cdot \dfrac{1}{n_1-1}}{\dfrac{(n_2-1)\,s_2^2}{\sigma^2} \cdot \dfrac{1}{n_2-1}} = \frac{s_1^2}{s_2^2} \tag{5.4.14}$$

Fisheri jaotusega vabadusastmete arvudega $\widehat{n}_1 = n_1 - 1$ ja $\widehat{n}_2 = n_2 - 1$. Kriitilise piirkonna valik sõltub konkureeriva hüpoteesi valikust.

1. Kui $H_1 : \mathrm{D}X_1 > \mathrm{D}X_2$, siis kasutatakse parempoolset kriitilist piirkonda. Paketi SWP kasutamisel leitakse

$$\mathrm{FInv}(1-\alpha, \widehat{n}_1, \widehat{n}_2) = z_{kr} \tag{5.4.15}$$

või Lisa 5 abil

$$(\alpha, \widehat{n}_1, \widehat{n}_2) \mapsto z_{\alpha,\,\widehat{n}_1,\,\widehat{n}_2} = z_{kr}. \tag{5.4.16}$$

Kui $z^* > z_{kr}$, siis hüpotees H_0 lükatakse tagasi. Kui $z^* < z_{kr}$, siis on hüpotees H_0 selle alternatiivse H_1 korral kooskõlas valimitega.

2. Kui $H_1 : \mathrm{D}X_1 \neq \mathrm{D}X_2$, siis kasutatakse kahepoolset kriitilist piirkonda. Lisa 5 abil leitakse muutuja z väärtused z_{krv} ja z_{krp}, kusjuures

$$(\alpha/2, \widehat{n}_1, \widehat{n}_2) \mapsto z_{\alpha/2,\,\widehat{n}_1,\,\widehat{n}_2} = z_{krp} \tag{5.4.17}$$

ja

$$(1-\alpha/2, \widehat{n}_1, \widehat{n}_2) \mapsto z_{1-\alpha/2,\,\widehat{n}_1,\,\widehat{n}_2} = \frac{1}{z_{\alpha/2,\,\widehat{n}_2,\,\widehat{n}_1}} = z_{krv}. \tag{5.4.18}$$

Kui $z^* < z_{krv} \vee z^* > z_{krp}$, siis hüpotees H_0 lükatakse tagasi. Kui $z_{krv} < z^* < z_{krp}$, siis on hüpotees H_0 selle alternatiivse H_1 korral kooskõlas valimitega. Paketi SWP kasutamisel leiame

$$z_{krv} = \mathrm{FInv}(\alpha/2, \widehat{n}_1, \widehat{n}_2), \quad z_{krp} = \mathrm{FInv}(1-\alpha/2, \widehat{n}_1, \widehat{n}_2). \tag{5.4.19}$$

Näide 1. Olgu X_1 ja X_2 kaks normaaljaotusega juhuslikku suurust dispersioonidega σ_1^2 ja σ_2^2. On võetud valimid vastavalt mahtudega $n_1 = 30$ ja $n_2 = 20$. Nende põhjal on leitud $s_1^2 = 12$ ja $s_2^2 = 8$. Kontrollime olulisuse nivool $\alpha = 0.02$ hüpoteesi $H_0 : \sigma_1^2 = \sigma_2^2$.

Leiame

$$z^* \overset{(5.4.14)}{=} \frac{s_1^2}{s_2^2} = 1.5.$$

1. Kui $H_1 : \mathrm{D}X_1 > \mathrm{D}X_2$, siis kasutatakse parempoolset kriitilist piirkonda. Paketi SWP30 kasutamisel leitakse

$$z_{kr} = \mathrm{FInv}(1 - 0.02, 29, 19) \approx 2.51.$$

Kuna $z^* < z_{kr}$, siis on hüpotees H_0 selle konkureeriva H_1 korral kooskõlas valimitega.

2. Kui $H_1 : \mathrm{D}X_1 \neq \mathrm{D}X_2$, siis kasutatakse kahepoolset kriitilist piirkonda. Lisa 5 abil leitakse muutuja z väärtused z_{krv} ja z_{krp}, kusjuures

$$(0.01; 29; 19) \overset{(5.4.17)}{\mapsto} z_{0.01;29;19} \approx 2.90 \approx z_{krp}$$

ja

$$(0.99; 29; 19) \overset{(5.4.18)}{\mapsto} z_{0.99;\,29;19} = \frac{1}{z_{0.01;19;29}} \approx \frac{1}{2.59} \approx 0.39 \approx z_{krv}.$$

Suuruste z_{krv} ja z_{krp} täpsemad väärtused saadakse (5.4.19) abil

$$z_{krv} = \mathrm{FInv}(0.01, 29, 19) \approx 0.38,\ z_{krp} = \mathrm{FInv}(0.99, 29, 19) \approx 2.86.$$

Et $z_{krv} < z^* < z_{krp}$, siis on hüpotees H_0 selle konkureeriva H_1 korral kooskõlas valimitega. $\diamondsuit$

2° Olgu uurimisel l normaaljaotusega suurust X_k vastavalt dispersioonidega σ_k^2 $(k = 1; 2; \ldots; l)$ ja võetud sõltumatud valimid mahtudega $n_1, \ldots, n_l$. Püstitame hüpoteesi $H_0 : \sigma_k^2 = \sigma^2 \ (k = 1; \ldots; l)$. Konkureerivaks valime hüpoteesi $H_1 = \overline{H}_0$. Hüpoteesi H_0 kontrollimiseks kasutame *Bartletti testi*, mis on rakendatav, kui $n_k \geq 3 \ (k = 1; \ldots; l)$. Selle testi korral kasutatakse statistikut

$$q^* = \frac{\sum_{k=1}^{l} (n_k - 1) \ln \left(\overline{s}^2 / s_k^2\right)}{1 + \dfrac{1}{3\,(l-1)} \left(\sum_{k=1}^{l} \dfrac{1}{n_k - 1} - \dfrac{1}{n_1 + \ldots + n_l - l}\right)}, \tag{5.4.20}$$

millel on χ^2-jaotus vabadusastmete arvuga $l - 1$. Seejuures on s_k^2 suuruse X_k valimi dispersioon ja

$$\overline{s}^2 = \frac{1}{n_1 + \ldots + n_l - l} \sum_{k=1}^{l} n_k s_k^2. \tag{5.4.21}$$

Hüpotees H_0 lükatakse tagasi, kui $q^* > q_{\alpha,\,l-1}$, kusjuures $(\alpha,\, l-1) \mapsto q_{\alpha,\,l-1}$ leitakse arvuti või χ^2-jaotuse tabeli abil (vt Lisa 3).

Näide 2. Olgu uurimisel neli normaaljaotusega suurust X_k dispersioonidega σ_k^2 $(k = 1; 2; 3; 4)$. On võetud valimid mahtudega vastavalt $n_1 = 10$, $n_2 = 16$, $n_3 = 12$, $n_4 = 20$. Eelnevalt on leitud $s_1^2 = 8$, $s_2^2 = 11$, $s_3^2 = 9$, $s_4^2 = 17$. Kontrollime Bartletti testi abil olulisuse nivool $\alpha = 0.02$ hüpoteesi

$$H_0 : \sigma_k^2 = \sigma^2 \, (k = 1; 2; 3; 4) \, .$$

Olgu $H_1 = \overline{H}_0$. Kuna

$$\overline{s}^2 \overset{(5.4.21)}{\approx} 13.037, \; q^* \overset{(5.4.20)}{\approx} 5.7824$$

ja Lisa 3 abil

$$(0.02; 3) \mapsto q_{0.02;\,3} = 9.84$$

või SWP3 abil

$$\text{ChiSquareInv}(0.98, 3) = 9.8374,$$

siis $q^* \approx 5.7824 < q_{0.02;\,3} \approx 9.8374$ ja H_0 vastab valimitele. $\diamondsuit$

5.4.4 Otsustused jaotusseaduste kohta

Kontrollime hüpoteesi H_0 : juhusliku suuruse X jaotusfunktsiooniks on just sellise kujuga $F(x)$. Olgu $H_1 = \overline{H}_0$

1° Esitame *Pearsoni* χ^2*-kriteeriumi* valimi jaotuse ja teoreetilise jaotuse erinevuse kohta. Olgu n valimi maht ja m variatsioonrea klasside arv. Olgu $n = n_1 + \ldots + n_m$, kus n_k näitab, mitu elementi on k-ndas klassis. Seega $w_k = \dfrac{n_k}{n}$. Olgu k-ndasse klassi kuulumise tõenäosus teoreetilise jaotuse põhjal p_k. Moodustame Pearsoni karakteristiku

$$q^* = \sum_{k=1}^{m} \frac{n}{p_k} \left(w_k - p_k\right)^2 = \sum_{k=1}^{m} \frac{\left(n_k - np_k\right)^2}{np_k}, \tag{5.4.22}$$

millel on χ^2-jaotus vabadusastmete arvuga $k = m - r - 1$, kus m on empiirilise jaotuse variatsioonrea klasside arv ja r on teoreetilise jaotuse parameetrite arv.

Pearsoni χ^2-kriteeriumi rakendamisel tuleb:
1) leida valemi (5.4.22) abil suurus q^*;
2) leida antud olulisuse nivoo α korral tabelist või arvuti abil kriitiline väärtus $q_{\alpha,k}$;
3) võrrelda suurusi q^* ja $q_{\alpha,k}$, kusjuures juhul $q^* \leq q_{\alpha,k}$ hüpotees H_0 ei ole vastuolus valimiga ja juhul $q^* > q_{\alpha,k}$ H_0 on vastuolus valimiga.

Näide 1. Kasutame Näites 5.1.1 saadud variatsioonrea klassideks jaotust

klassid	$(0; 2]$	$(2; 4]$	$(4; 6]$	$(6; 8]$	$(8; 10]$	$(10; 12]$
elementide arv	6	8	11	4	0	1
sagedus (suhteline)	6/30	8/30	11/30	4/30	0/30	1/30

Olgu olulisuse nivoo $\alpha = 0.05$. Näidetes 5.2.1.1 ja 5.2.2.1 leidsime, et $\overline{x} = 4.633$ ja $s^2 \approx 4.79$, st $s \approx 2.19$. Kasutame teoreetilise jaotusena $N\,(4.63; 2.19)$. Kas hüpotees

$$H_0 : \; X \sim N(4.63; 2.19)$$

on kooskõlas Näites 5.1.1 esitatud valimiga?

Leiame selle normaaljaotuse korral tõenäosused:

$$\begin{aligned}
p_1 &= P(-\infty \le X \le 2) = \Phi\left(\frac{2-4.63}{2.19}\right) - \Phi\left(\frac{-\infty-4.63}{2.19}\right) \approx \\
&= \Phi(-1.20) - \Phi(-\infty) = \Phi(+\infty) - \Phi(1.20) \approx 0.115 \\
p_2 &= P(2 \le X \le 4) = \Phi(-0.288) - \Phi(-1.20) \approx 0.272; \\
p_3 &= P(4 \le X \le 6) = \Phi(0.626) - \Phi(-0.288) \approx 0.347; \\
p_4 &= P(6 \le X \le 8) = \Phi(1.539) - \Phi(0.626) \approx 0.204 \\
p_5 &= P(8 \le X \le 10) = \Phi(2.452) - \Phi(1.539) \approx 0.055; \\
p_6 &= P(10 \le X \le +\infty) = \Phi(+\infty) - \Phi(2.452) \approx 0.007.
\end{aligned}$$

Valemi (5.4.22) abil leiame

$$q^* = \frac{(6-30\cdot 0.115)^2}{30\cdot 0.115} + \frac{(8-30\cdot 0.272)^2}{30\cdot 0.272} + \frac{(11-30\cdot 0.347)^2}{30\cdot 0.347} +$$

$$+\frac{(4-30\cdot 0.204)^2}{30\cdot 0.204} + \frac{(0-30\cdot 0.055)^2}{30\cdot 0.055} + \frac{(1-30\cdot 0.007)^2}{30\cdot 0.007} \approx 7.28.$$

Kuna Lisas 3 esitatud tabeli põhjal $q_{0.05;3} = 7.82$, siis $q^* \le q_{0.05;3}$. Seega on olulisuse nivool $\alpha = 0.05$ hüpotees $H_0 :\ X \sim N(4.67; 2.29)$ Pearsoni χ^2-kriteeriumi põhjal valimiga kooskõlas. $\diamondsuit$

2° *Kolmogorovi kriteerium* valimi jaotuse ja teoreetilise jaotuse erinevuse jaoks. Lähtume valimi põhjal koostatud sagedustabelist

klassid	$(a_0, a_1]$	$(a_1, a_2]$	$\cdots$	$(a_{m-1}, a_m]$
sagedus (suhteline)	p_1	p_2	$\cdots$	p_m

milles on fikseeritud klassid ja elementide suhteline sagedus igas klassis. Koostame tabeli

x	a_0	a_1	a_2	$\cdots$	a_{m-1}	a_m
$F_n(x)$	0	p_1	p_1+p_2	$\cdots$	$p_1+\ldots+p_{m-1}$	1
$F(x)$	$F(a_0)$	$F(a_1)$	$F(a_2)$	$\cdots$	$F(a_{m-1})$	$F(a_m)$

kus $F(a_k)$ on leitud teoreetilise jaotuse jaotusfunktsiooni $F(x)$ abil. Leiame *Kolmogorovi statistiku*

$$D = \max_{0\le k\le m} |F_n(a_k) - F(a_k)|. \tag{5.4.23}$$

On tõestatud, et iga pideva juhusliku suuruse X korral

$$P\left(D\sqrt{n} \ge \lambda\right) \overset{n\to\infty}{\to} P(\lambda) \overset{\text{def.}}{=} 1 - \sum_{k=-\infty}^{\infty} (-1)^k e^{-2k^2\lambda^2} \quad (\lambda > 0). \tag{5.4.24}$$

Etteantud olulisuse nivool α tuleb kriitiline väärtus λ_α leida võrrandist

$$P(\lambda_\alpha) = \alpha. \tag{5.4.25}$$

Esitame funktsiooni

$$P^{-1}: \ \alpha \mapsto \lambda_\alpha$$

lühikese tabeli:

α	0.4	0.3	0.2	0.1	0.05	0.025	0.01	0.005	0.001	0.0005
λ_α	0.89	0.97	1.07	1.22	1.36	1.48	1.63	1.73	1.95	2.03

Kolmogorovi kriteeriumi rakendamisel tuleb:
1) koostada punktides a_k $(k = 0, 1, \ldots; m)$ sagedustabeli põhjal funktsioonide $F_n(x)$ ja $F(x)$ väärtuste tabel;
2) leida D ja $\lambda = D\sqrt{n}$;
3) leida λ_α;
4) kontrollida nullhüpoteesi H_0 vastavust valimile antud olulisuse nivoo α korral, st kui $\lambda \le \lambda_\alpha$, siis võetakse hüpotees H_0 vastu ja $\lambda > \lambda_\alpha$ korral lükatakse H_0 tagasi.

Näide 2. Kontrollime Kolmogorovi kriteeriumi abil olulisuse nivool $\alpha = 0.05$ Näite 1 variatsioonrea klassideks jaotuse korral hüpoteesi

$$H_0: \ X \sim N(4.63; 2.72).$$

1. Koostame tabeli

x	0	2	4	6	8	10	12
$F_n(x)$	0	0.200	0.467	0.833	0.967	0.967	1
$F(x)$	0.044	0.167	0.408	0.693	0.892	0.976	0.997

2. Leiame suurused

$$D = |0.833 - 0.693| = 0.140, \ \lambda = D\sqrt{n} = 0.140\sqrt{30} \approx 0.77.$$

3. Funktsiooni P^{-1} tabeli põhjal $\alpha = 0.05 \mapsto 1.36 = \lambda_\alpha$.
4. Kuna $\lambda \le \lambda_\alpha$, siis olulisuse nivool $\alpha = 0.05$ on hüpotees H_0 Kolmogorovi kriteeriumi põhjal kooskõlas valimiga. $\diamondsuit$

5.5 Vähimruutude meetod ja regressioonijooned

Uurime järgnevalt empiirilise sõltuvuse esitamist analüütilise avaldisega.

1° Olgu antud juhusliku vektori (X, Y) valim $(x_i, y_i) \quad (i = 1, \ldots, n)$. Valime funktsioonide perest

$$\left\{ f(x) = \sum_{k=1}^{m} c_k \varphi_k(x) \right\},$$

kus kordajad $c_k \in \mathbf{R}$ on konstandid ja $\varphi_k(x) \quad (k = 1; \ldots; m)$ on etteantud sõltumatud funktsioonid, funktsiooni $f(x)$, mis *vähimruutude* mõttes lähendab kõige paremini antud valimit. Funktsioonide $\varphi_k(x)$ valik sõltub infost üldkogumi kohta. Tavaliselt kasutatakse lähendamist kas algebraliste või trigonomeetriliste polünoomidega. Kasutatakse ka lähendamist splainidega ja lainekestega. Arvutame $f(x_i)$ $(i = 1, \ldots, n)$, vahed $y_i - f(x_i)$ ja nende ruudud

$$(y_i - f(x_i))^2 \quad (i = 1, \ldots, n).$$

Moodustame suuruse

$$g(c_1, \ldots, c_m) = \sum_{i=1}^{n} \left(y_i - \sum_{k=1}^{m} c_k \varphi_k(x_i) \right)^2. \tag{5.5.1}$$

Määrame kordajad c_k $(k = 1, \ldots, m)$ selliselt, et $g(c_1, \ldots, c_m)$ oleks minimaalne. Lokaalse ekstreemumi tarviliku tingimuse põhjal

$$\frac{\partial g(c_1, \ldots, c_m)}{\partial c_j} = 0 \quad (j = 1, \ldots, m), \tag{5.5.2}$$

st

$$2 \sum_{i=1}^{n} \left(y_i - \sum_{k=1}^{m} c_k \varphi_k(x_i) \right) (-\varphi_j(x_i)) = 0 \quad (j = 1, \ldots, m)$$

ehk

$$\sum_{i=1}^{n} \left(y_i - \sum_{k=1}^{m} c_k \varphi_k(x_i) \right) \varphi_j(x_i) = 0 \quad (j = 1, \ldots, m)$$

või

$$\sum_{k=1}^{m} \left(\sum_{i=1}^{n} \varphi_k(x_i) \varphi_j(x_i) \right) c_k = \sum_{i=1}^{n} y_i \varphi_j(x_i) \quad (j = 1, \ldots, m). \tag{5.5.3}$$

Lahendades süsteemi (5.5.3), saame kordajad c_k ja valimile vastava funktsiooni $f(x)$ konkreetse kuju.

Kui $m = 2$ ja $\varphi_1(x) = 1$ ja $\varphi_2(x) = x$, siis saame funktsiooni $f(x) = c_1 + c_2 x$ kordajate c_1 ja c_2 leidmiseks süsteemi (5.5.3):

$$\left\{ \begin{array}{rcrcr} c_1 \sum\limits_{i=1}^{n} 1 & + & c_2 \sum\limits_{i=1}^{n} x_i & = & \sum\limits_{i=1}^{n} y_i \\ c_1 \sum\limits_{i=1}^{n} x_i & + & c_2 \sum\limits_{i=1}^{n} x_i^2 & = & \sum\limits_{i=1}^{n} x_i y_i \end{array} \right. \Leftrightarrow$$

$$\Leftrightarrow \left\{ \begin{array}{ccccc} c_1 & + & c_2 \frac{1}{n} \sum\limits_{i=1}^{n} x_i & = & \frac{1}{n} \sum\limits_{i=1}^{n} y_i \\ c_1 \frac{1}{n} \sum\limits_{i=1}^{n} x_i & + & c_2 \frac{1}{n} \sum\limits_{i=1}^{n} x_i^2 & = & \frac{1}{n} \sum\limits_{i=1}^{n} x_i y_i \end{array} \right. \Leftrightarrow$$

$$\Leftrightarrow \left\{ \begin{array}{ccccc} c_1 & + & c_2 \overline{x} & = & \overline{y} \\ c_1 \overline{x} & + & c_2 \overline{x^2} & = & \overline{xy} \end{array} \right. , \qquad (5.5.4)$$

kus

$$\overline{x} = \frac{1}{n} \sum_{i=1}^{n} x_i, \quad \overline{y} = \frac{1}{n} \sum_{i=1}^{n} y_i\,, \ \overline{x^2} = \frac{1}{n} \sum_{i=1}^{n} x_i^2, \ \overline{xy} = \frac{1}{n} \sum_{i=1}^{n} x_i y_i. \qquad (5.5.5)$$

Süsteemi (5.5.4) esimesest võrrandist järeldub, et punkt $(\overline{x}, \overline{y})$ asetseb sirgel $y = c_1 + c_2 x$. Leiame kordajad c_1 ja c_2, lahendades süsteemi (5.5.4). Saadud sirget nimetatakse suuruse Y *regressioonisirgeks* suuruse X suhtes.

Kui $m = 3$, $\varphi_1(x) = 1$, $\varphi_2(x) = x$ ja $\varphi_3(x) = x^2$, siis (5.5.3) abil saame funktsiooni $f(x) = c_1 + c_2 x + c_3 x^2$ kordajate c_1, c_2 ja c_3 määramiseks süsteemi:

$$\left\{ \begin{array}{ccccccc} c_1 \sum\limits_{i=1}^{n} 1 & + & c_2 \sum\limits_{i=1}^{n} x_i & + & c_3 \sum\limits_{i=1}^{n} x_i^2 & = & \sum\limits_{i=1}^{n} y_i \\ c_1 \sum\limits_{i=1}^{n} x_i & + & c_2 \sum\limits_{i=1}^{n} x_i^2 & + & c_3 \sum\limits_{i=1}^{n} x_i^3 & = & \sum\limits_{i=1}^{n} x_i y_i \\ c_1 \sum\limits_{i=1}^{n} x_i^2 & + & c_2 \sum\limits_{i=1}^{n} x_i^3 & + & c_3 \sum\limits_{i=1}^{n} x_i^4 & = & \sum\limits_{i=1}^{n} x_i^2 y_i \end{array} \right. \Leftrightarrow$$

$$\left\{ \begin{array}{ccccccc} c_1 & + & c_2 \overline{x} & + & c_3 \overline{x^2} & = & \overline{y} \\ c_1 \overline{x} & + & c_2 \overline{x^2} & + & c_3 \overline{x^3} & = & \overline{xy} \\ c_1 \overline{x^2} & + & c_2 \overline{x^3} & + & c_3 \overline{x^4} & = & \overline{x^2 y}, \end{array} \right. \qquad (5.5.6)$$

kus

$$\overline{x} = \frac{1}{n} \sum_{i=1}^{n} x_i, \ \overline{x^2} = \frac{1}{n} \sum_{i=1}^{n} x_i^2, \ \overline{x^3} = \frac{1}{n} \sum_{i=1}^{n} x_i^3, \ \overline{x^4} = \frac{1}{n} \sum_{i=1}^{n} x_i^4, \qquad (5.5.7)$$

$$\overline{y} = \frac{1}{n} \sum_{i=1}^{n} y_i\,, \ \overline{xy} = \frac{1}{n} \sum_{i=1}^{n} x_i y_i, \quad \overline{x^2 y} = \frac{1}{n} \sum_{i=1}^{n} x_i^2 y_i. \qquad (5.5.8)$$

Lause 1. Suuruse Y regressioonisirge X suhtes $y = c_1 + c_2 x$ parameetrid c_1 ja c_2 on leitavad süsteemist (5.5.4), mille kordajad on arvutatavad valemite (5.5.5) abil. *Regressiooniparabooli* $y = c_1 + c_2 x + c_3 x^2$ parameetrid c_1, c_2 ja c_3 on leitavad süsteemist (5.5.6), mille kordajad on arvutatavad valemite (5.5.7) ja (5.5.8) abil.

Näide 1. Olgu saadud valim $\{(x_i, y_i)\} \quad (i = 1, \ldots, 10)$

x_i	6.21	3.76	4.16	3.22	10.6	5.07	4.60	2.71	3.57	2.02
y_i	10.3	5.26	2.65	3.12	8.68	3.38	1.94	4.54	5.02	1.81

Leiame suuruse Y regressioonisirge $y = c_1 + c_2 x$ ja regressiooniparabooli $y = c_1 + c_2 x + c_3 x^2$ suuruse X suhtes. Teeme joonise.

Valemite (5.5.7) ja (5.5.8) põhjal leiame $\overline{x} = 4.592$, $\overline{x^2} = 26.377$, $\overline{x^3} = 189.03$, $\overline{x^4} = 1606.1$, $\overline{y} = 4.67$, $\overline{xy} = 25.676$, $\overline{x^2y} = 175.77$. Süsteemile (5.5.4) saame kuju

$$\begin{cases} c_1 + 4.59c_2 = 4.67 \\ 4.59c_1 + 26.38c_2 = 25.68, \end{cases}$$

millest $c_1 \approx 1.00$ ja $c_2 \approx 0.80$. Seega on regressioonisirgeks

$$y = 1 + 0.8x.$$

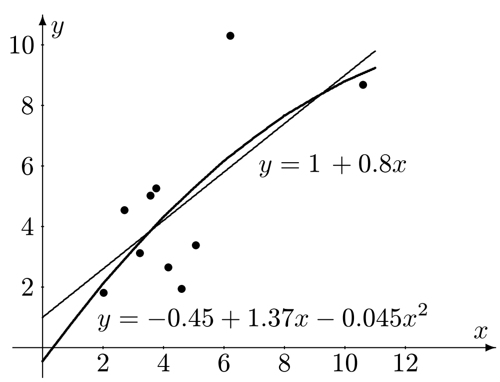
Süsteemile (5.5.6) saame kuju

$$\begin{cases} c_1 + 4.59c_2 + 26.38c_3 = 4.67 \\ 4.59c_1 + 26.38c_2 + 189.03c_3 = 25.68 \\ 26.38c_1 + 189.03c_2 + 1606.1c_3 = 175.77, \end{cases}$$

millest $c_1 \approx -0.45$ $c_2 \approx 1.37$ ja $c_3 \approx -0.045$. Seega

$$y = -0.45 + 1.37x - 0.045x^2$$

on regressiooniparabool. Teeme joonise, millele kanname *korrelatsioonivälja* (punktide (x_i, y_i) parv tasandil), regressiooniparabooli jämeda ja regressioonisirge peene joonega



2° Olgu saadud juhusliku vektori (X, Y) valim mahuga n. Valimi elementideks on paarid (x_i, y_j) $(i = 1, \ldots, m;\ j = 1, \ldots, l)$, kusjuures n_{ij} on paari sagedus, st selle paari esinemiste arv valimis, ja $x_{i+1} > x_i$ ning $y_{j+1} > y_j$. Selline olukord tekib, kui suuruste X ja Y variatsioonread on jaotatud klassidesse ja x_i ning y_j on klasside keskpunktid. Kui tähistada

$$n_j = \sum\nolimits_{i=1}^{m} n_{ij},\ \widehat{n}_i = \sum\nolimits_{j=1}^{l} n_{ij}, \tag{5.5.9}$$

siis

$$n = \sum\nolimits_{i=1}^{m} \widehat{n}_i = \sum\nolimits_{j=1}^{l} n_j = \sum\nolimits_{i=1}^{m} \sum\nolimits_{j=1}^{l} n_{ij}.$$

Iga x_i $(i = 1, \ldots, m)$ korral leitakse

$$\overline{y}_i = \left(\sum\nolimits_{j=1}^{l} y_j n_{ij} \right) / \widehat{n}_i. \tag{5.5.10}$$

Saame skitseerida murdjoone tippudega $(x_i, \overline{y}_i)$ $(i = 1, \ldots, m)$. Seda murdjoont nimetatakse suuruse Y *empiiriliseks regressioonijooneks empiiriline!regressioonijoon* suuruse X suhtes. Analoogiliselt leitakse iga y_j $(j = 1, \ldots, l)$ korral

$$\overline{x}_j = \left(\textstyle\sum_{i=1}^m x_i n_{ij}\right)/n_j. \tag{5.5.11}$$

Sellisel viisil saadakse suuruse X empiiriline regressioonijoon suuruse Y suhtes $(\overline{x}_j, y_j)$ $(j = 1, \ldots, l)$. Empiirilise regressioonijoone kuju järgi võib aimata, kas suuruste X ja Y vahel eksisteerib lineaarne sõltuvus või mitte. Suuruse Y lineaarset regressioonijoont X suhtes saab nagu jaotises 1° otsida kujul

$$y_x = c_1 + c_2 x. \tag{5.5.12}$$

Seejuures määratakse parameetrid c_1 ja c_2 vähimruutude meetodil, minimeerides suuruse

$$g(c_1, c_2) = \textstyle\sum_{i=1}^m (c_1 + c_2 x_i - \overline{y}_i)^2 \widehat{n}_i.$$

Seega tuleb lahendada süsteem

$$\begin{cases} \dfrac{\partial g(c_1, c_2)}{\partial c_1} = 0 \\ \dfrac{\partial g(c_1, c_2)}{\partial c_2} = 0 \end{cases} \Leftrightarrow \begin{cases} 2\sum\limits_{i=1}^m (c_1 + c_2 x_i - \overline{y}_i)\widehat{n}_i = 0 \\ 2\sum\limits_{i=1}^m (c_1 + c_2 x_i - \overline{y}_i)\widehat{n}_i x_i = 0 \end{cases}$$

ehk

$$\begin{cases} c_1 \sum\limits_{i=1}^m \widehat{n}_i + c_2 \sum\limits_{i=1}^m \widehat{n}_i x_i = \sum\limits_{i=1}^m \widehat{n}_i \overline{y}_i \\ c_1 \sum\limits_{i=1}^m \widehat{n}_i x_i + c_2 \sum\limits_{i=1}^m \widehat{n}_i x_i^2 = \sum\limits_{i=1}^m \widehat{n}_i x_i \overline{y}_i. \end{cases} \tag{5.5.0.1}$$

Kuna

$$\sum_{i=1}^m \widehat{n}_i \overline{y}_i \overset{(5.5.10)}{=} \sum_{i=1}^m \widehat{n}_i \left(\textstyle\sum_{j=1}^l y_j n_{ij}\right)/\widehat{n}_i = \textstyle\sum_{j=1}^l y_j \sum\limits_{i=1}^m n_{ij} \overset{(5.5.9)}{=} \textstyle\sum_{j=1}^l n_j y_j,$$

$$\sum_{i=1}^m \widehat{n}_i x_i \overline{y}_i \overset{(5.5.10)}{=} \sum_{i=1}^m \widehat{n}_i x_i \left(\textstyle\sum_{j=1}^l y_j n_{ij}\right)/\widehat{n}_i = \sum_{i=1}^m \textstyle\sum_{j=1}^l n_{ij} x_i y_j,$$

siis saadakse (5.5.13) kujul

$$\begin{cases} c_1 & + & c_2\overline{x} & = & \overline{y} \\ c_1\overline{x} & + & c_2\overline{x^2} & = & \overline{xy}, \end{cases} \tag{5.5.0.2}$$

kus

$$\overline{x} = \frac{1}{n}\sum_{i=1}^m \widehat{n}_i x_i,\ \overline{x^2} = \frac{1}{n}\sum_{i=1}^m \widehat{n}_i x_i^2,\ \overline{y} = \frac{1}{n}\textstyle\sum_{j=1}^l n_j y_j,\ \overline{xy} = \dfrac{1}{n}\sum\limits_{i=1}^m \textstyle\sum_{j=1}^l n_{ij} x_i y_j. \tag{5.5.15}$$

Analoogiliselt saab leida suuruse X lineaarse regressioonijoone Y suhtes

$$x_y = d_1 + d_2 y, \tag{5.5.0.3}$$

lahendades süsteemi

$$\begin{cases} d_1 & + & d_2\overline{y} & = & \overline{x} \\ d_1\overline{y} & + & d_2\overline{y^2} & = & \overline{xy}, \end{cases} \tag{5.5.0.4}$$

kus

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{m}\widehat{n}_i x_i,\ \overline{y} = \frac{1}{n}\sum\nolimits_{j=1}^{l} n_j y_j,\ \overline{y^2} = \frac{1}{n}\sum\nolimits_{j=1}^{l} n_j y_j^2,\ \overline{xy} = \frac{1}{n}\sum_{i=1}^{m}\sum\nolimits_{j=1}^{l} n_{ij} x_i y_j. \tag{5.5.0.5}$$

Sõnastame tuletatu.

Lause 2. Kujul (5.5.12) on otsitava suuruse Y lineaarse regressioonijoone X suhtes parameetrid c_1 ja c_2 leitavad süsteemist (5.5.14), mille kordajad on arvutatavad valemite (5.5.15) abil. Kujul (5.5.16) on otsitava suuruse X lineaarse regressioonijoone Y suhtes parameetrid d_1 ja d_2 leitavad süsteemist (5.5.17), mille kordajad on arvutatavad valemite (5.5.18) abil.

Näide 2. Olgu saadud juhusliku vektori (X, Y) valim mahuga $n = 150$. Leiame suuruse Y empiirilise regressioonijoone $y = a(x)$ suuruse X suhtes ja suuruse X empiirilise regressioonijoone $x = b(y)$ suuruse Y suhtes. Leiame suuruse Y lineaarse regressioonijoone X suhtes ja suuruse X lineaarse regressioonijoone Y suhtes. Teeme joonise. Seejuures on valimi elementideks paarid (x_i, y_j) $(i = 1, \dots, 7;\ j = 1, \dots, 5)$, kusjuures paaride sagedused n_{ij} on esitatud järgnevas tabelis, st $(x_i, y_j) \mapsto n_{ij}$

Y	X							n_j
	25	30	35	40	45	50	55	
13	3	2	7	1	0	1	0	14
23	5	13	4	12	4	0	0	38
33	2	4	13	9	15	6	4	53
43	0	3	5	10	6	9	2	35
53	0	0	0	2	2	5	1	10
$\widehat{n}_i$	10	22	29	34	27	21	7	$n = 150$

Valemi (5.5.10) abil leiame funktsiooni $x_i \to \overline{y}_i$ $(i = 1; \dots; 7)$ tabeli

$\overline{y}_i$	22.00	26.64	28.52	33.00	35.22	41.10	38.714
x_i	25	30	35	40	45	50	55

Valemi (5.5.11) abil leiame funktsiooni $y_j \to \overline{x}_j$ $(j = 1; \dots; 5)$ tabeli

$\overline{x}_j$	33.57	34.61	41.13	42.71	47.50
y_j	13	23	33	43	53

Valemite (5.5.15) abil leiame

$$\overline{x} = 39.\,57,\ \overline{y} = 32.\,27,\ \overline{x^2} = 1628.8,\ \overline{xy} = 1318.\,4.$$

Seega saame süsteemi (5.5.14) kujul

$$\begin{cases} c_1 & + & 39.\,57c_2 & = & 32.\,27 \\ 39.\,57c_1 & + & 1628.\,8c_2 & = & 1318.\,4, \end{cases}$$

millest $c_1 \approx 6.\,225$ ja $c_2 \approx 0.\,658$ ning

$$y_x \approx 6.\,225 + 0.\,658\,x.$$

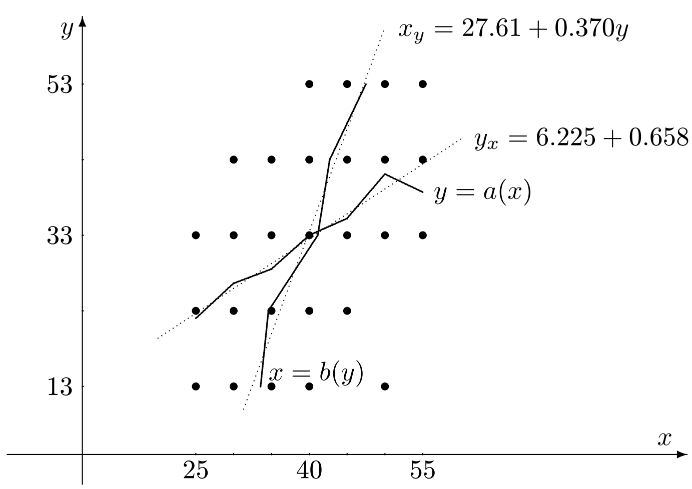
Valemi (5.5.18) abil leiame täiendavalt, et $\overline{y^2} = 1153.\,3$. Saame süsteemile (5.5.17) kuju

$$\begin{cases} d_1 & + & 32.\,27d_2 & = & 39.\,57 \\ 32.\,27d_1 & + & 1153.\,3d_2 & = & 1318.\,4, \end{cases}$$

millest $d_1 \approx 27.\,61$ ja $d_2 \approx 0.37$ ning

$$x_y \approx 27.\,61 + 0.\,37y.$$

Teeme joonise



$\diamondsuit$

5.6 Ülesanded

1. On antud valim

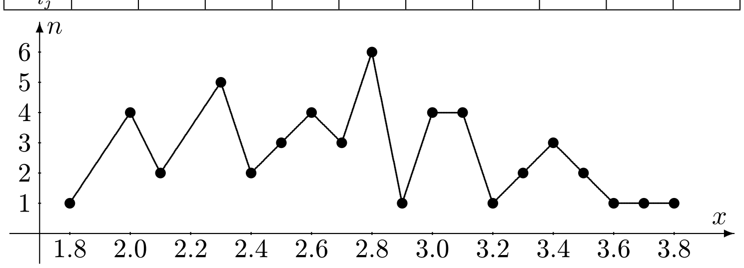
$$\{2.8,\ 3.5,\ 3.7,\ 2.3,\ 2.0,\ 3.8,\ 2.3,\ 2.6,\ 2.8,\ 2.3,\ 2.0,\ 3.2,\ 3.0,\ 2.1,\ 3.3,\ 3.5,\ 3.6,\ 3.0,\ 2.3,\ 3.4,\ 3.1,\ 2.5,\ 2.7,\ 2.5,\ 2.6,\ 1.8,\ 2.5,\ 3.4,\ 2.7,\ 2.0,\ 2.8,\ 2.9,\ 2.6,\ 2.1,\ 3.1,\ 2.8,\ 2.4,\ 2.0,\ 3.1,\ 3.0,\ 3.0,\ 2.3,\ 3.1,\ 3.4,\ 3.3,\ 2.8,\ 2.6,\ 2.7,\ 2.8,\ 2.4\}.$$

Järjestage see valim väärtuste kasvamise järgi, koostage sageduste tabel, skitseerige sageduste polügoon, leidke empiiriline jaotusseadus ja empiiriline jaotusfunktsioon $F_{50}^*(x)$ ja selle graafik. Jaotage variatsioonrida klassidesse ja koostage klasside sagedustabel, suhteliste sageduste tabel ning histogramm. Leidke klassidele vastav empiiriline jaotusfunktsioon $F_{50}^{\triangle}(x)$ ja skitseerige selle graafik.

V:

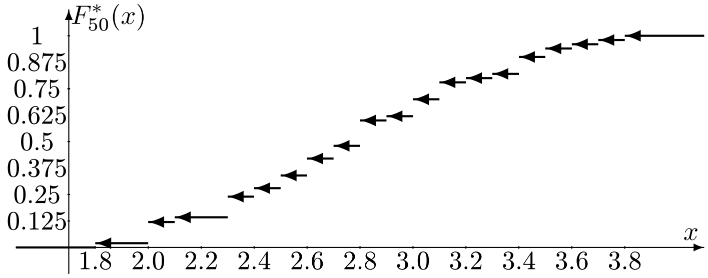
$$\{1.8,\ 2.0,\ 2.0,\ 2.0,\ 2.0,\ 2.1,\ 2.1,\ 2.3,\ 2.3,\ 2.3,\ 2.3,\ 2.3,\ 2.4,\ 2.4,\ 2.5,\ 2.5,\ 2.5,\ 2.6,\ 2.6,\ 2.6,\ 2.6,\ 2.7,\ 2.7,\ 2.7,\ 2.8,\ 2.8,\ 2.8,\ 2.8,\ 2.8,\ 2.8,\ 2.9,\ 3.0,\ 3.0,\ 3.0,\ 3.0,\ 3.1,\ 3.1,\ 3.1,\ 3.1,\ 3.2,\ 3.3,\ 3.3,\ 3.4,\ 3.4,\ 3.4,\ 3.5,\ 3.5,\ 3.6,\ 3.7,\ 3.8\},$$

x_{i_j}	1.8	2.0	2.1	2.3	2.4	2.5	2.6	2.7	2.8	2.9
$n_{i_j}^*$	1	4	2	5	2	3	4	3	6	1
x_{i_j}	3.0	3.1	3.2	3.3	3.4	3.5	3.6	3.7	3.8	
$n_{i_j}^*$	4	4	1	2	3	2	1	1	1	

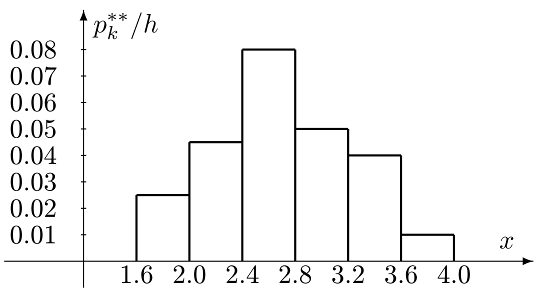


x_{i_j}	1.8	2.0	2.1	2.3	2.4	2.5	2.6	2.7	2.8	2.9
$p_{i_j}^*$	0.02	0.08	0.04	0.1	0.04	0.06	0.08	0.06	0.12	0.02
x_{i_j}	3.0	3.1	3.2	3.3	3.4	3.5	3.6	3.7	3.8	
$p_{i_j}^*$	0.08	0.08	0.02	0.04	0.06	0.04	0.02	0.02	0.02	

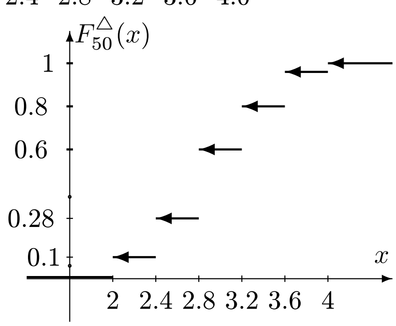
x	$(-\infty; 1.8]$	$(1.8; 2.0]$	$(2.0; 2.1]$	$(2.1; 2.3]$	$(2.3; 2.4]$
$F_n^*(x)$	0	0.02	0.12	0.14	0.24
x	$(2.4; 2.5]$	$(2.5; 2.6]$	$(2.6; 2.7]$	$(2.7; 2.8]$	$(2.8; 2.9]$
$F_n^*(x)$	0.28	0.34	0.42	0.48	0.60
x	$(2.9; 3.0]$	$(3.0; 3.1]$	$(3.1; 3.2]$	$(3.2; 3.3]$	$(3.3; 3.4]$
$F_n^*(x)$	0.62	0.70	0.78	0.80	0.82
x	$(3.4; 3.5]$	$(3.5; 3.6]$	$(3.6; 3.7]$	$(3.7; 3.8]$	$(3.8; +\infty)$
$F_n^*(x)$	0.90	0.94	0.96	0.98	1



klass	$(1.6;2]$	$(2;2.4]$	$(2.4;2.8]$	$(2.8;3.2]$	$(3.2;3.6]$	$(3.6;4]$
k_i	5	9	16	10	8	2
$p_i^{**} = \dfrac{k_i}{n}$	0.1	0.18	0.32	0.2	0.16	0.04
p_i^{**}/h	0.025	0.045	0.08	0.05	0.04	0.01



$$F_{50}^{\triangle}(x) = \begin{cases} 0, \ x \le 2.0, \\ 0.1, \ 2.0 < x \le 2.4, \\ 0.28, \ 2.4 < x \le 2.8, \\ 0.6, \ 2.8 < x \le 3.2, \\ 0.8, \ 3.2 < x \le 3.6, \\ 0.96, \ 3.6 < x \le 4.0, \\ 1, \ 4.0 < x \end{cases}$$



$\diamondsuit$

2. Leidke valimi

$$\{4.32,\ 3.62,\ 0.91,\ 5.62,\ 0.18,\ 2.96,\ 9.91,\ 10.42,\ 6.72,\ 2.14,$$
$$5.26,\ 2.65,\ 3.13,\ 8.69,\ 3.39,\ 1.42,\ 1.95,\ 4.57,\ 5.03,\ 1.82\}$$

keskmine, dispersioon ja standardhälve.
V: $\overline{x} = 4.24$, $s^2 \approx 8.29$, $s \approx 2.88$.

3. Olgu variatsioonrida esitatud kujul

x_i	1	3	4	6	7
n_i	2	3	6	7	2

Leidke valimi keskmine, dispersioon ja standardhälve.
V: $\overline{x} = 4.55$, $s^2 \approx 3.21$, $s \approx 1.79$.

4. Kasutades rekursiivseid valemeid (5.2.1.1) ja (5.2.2.7), leidke Näites 5.1.1 esitatud valimi keskmine ja dispersioon.
5. Kondensaatori tööiga on juhuslik suurus, mis allub normaaljaotusele $N(a;8)$. Fikseeriti 30 kondensaatori tööiga. Saadi valim. Leidke sündmuse $s^2 \in [60;70]$ tõenäosus. V: $P\left(s^2 \in [60;70]\right) \approx 0.23$.
6. Üldkogumi juhuslik suurus X allub normaaljaotusele $N(a,\sigma)$. Saadakse valim mahuga 20, mille korral $\overline{x} = 5$ ja $s^2 = 16$. Leiame sündmuse $a < 4$ tõenäosuse. V: $P(a < 4) \approx 0.1447$.
7. Olgu vektori (X, Y) korral sooritatud 8 sõltumatut katset, mis andsid valimi

$$\{(2;4), (1;3), (-1;2); (1;1), (-2;3), (0;1), (2;3), (1;2)\}.$$

Leidke valimi kovariatsioonimaatriks ja korrelatsioonimaatriks. V:

$$\begin{pmatrix} 2.0 & 0.357 \\ & 1.125 \end{pmatrix}, \quad \begin{pmatrix} 1.0 & 0.24 \\ & 1.0 \end{pmatrix}.$$

8. Olgu vektori (X_1, X_2, X_3, X_4) korral sooritatud 8 sõltumatut katset, mis andsid valimi

i	1	2	3	4	5	6	7	8
x_1	1	2	0	-1	3	2	2	1
x_2	1	4	1	0	7	3	3	0
x_3	1	6	-1	-2	24	7	6	2
x_4	0	15	2	3	25	15	14	2

Leidke valimi kovariatsioonimaatriks ja korrelatsioonimaatriks. V:

$$\left(\mathrm{K}^*_{x_k,\,x_m}\right) = \begin{pmatrix} 1.64 & 2.61 & 8.75 & 9.57 \\ & 5.70 & 18.41 & 20.50 \\ & & 67.98 & 66.64 \\ & & & 80.86 \end{pmatrix},$$

$$\left(\mathrm{r}^*_{x_k,x_m}\right) = \begin{pmatrix} 1.0 & 0.85 & 0.83 & 0.83 \\ & 1.0 & 0.94 & 0.96 \\ & & 1.0 & 0.90 \\ & & & 1.0 \end{pmatrix}.$$

9. Juhuslik suurus X, valimiga $\{x_1, \ldots, x_n\}$, allub eksponentjaotusele tihedusega $f(x) = \lambda \exp(-\lambda x)\,\mathbf{1}(x)\ \ (\lambda > 0)$. Leidke λ^* suurima tõepära meetodil. V: $\lambda^* = 1/\overline{x}$.
10. Juhuslik suurus X, valimiga $\{x_1, \ldots, x_n\}$, allub geomeetrilisele jaotusele parameetriga p, st $P(X = k) = (1-p)^{k-1} p\ \ (k \in \mathbf{N})$. Leidke p^* suurima tõepära meetodil. V: $p^* = 1/\overline{x}$.
11. Juhuslik suurus X, valimiga $\{x_1, \ldots, x_n\}$, allub *Kaptayni jaotusele* tihedusega

$$f(x) = \frac{g'(x)}{\sigma\sqrt{2\pi}} \exp\left(-\left(g(x) - a\right)^2 / \left(2\sigma^2\right)\right),$$

kus $g(x)$ on diferentseeruv funktsioon, mille korral $g'(x) > 0$. Leidke a^* suurima tõepära meetodil, kui σ on teada. V: $a^* = \frac{1}{n}\sum_{k=1}^{n} g(x_k)$.

12. Juhuslik suurus X, valimiga $\{x_1, \ldots, x_n\}$, allub Kaptayni jaotusele. Leidke σ^* suurima tõepära meetodil, kui a on teada. V: $\sigma^* = \sqrt{\frac{1}{n}\sum_{k=1}^{n} (g(x_k) - a)^2}$.

13. Juhusliku suuruse X valimi $\{x_1, \ldots, x_{50}\}$ korral $\overline{x} = 8.1$ ja $s = 2.3$. Leidke usaldusnivoole 0.95 vastav keskväärtuse EX usaldusvahemik.
V: $P(\mathrm{E}X \in (7.46; 8.74)) \approx 0.95$.

14. Normaaljaotusele alluva juhusliku suuruse X valimi $\{x_1, \ldots, x_{50}\}$ korral on $\overline{x} = 8.1$ ja $s = 2.3$. Leidke usaldusnivoole 0.95 vastav keskväärtuse EX usaldusvahemik. Võrrelge ülesannete 13 ja 14 vastuseid sisuliselt.
V: $P(\mathrm{E}X \in (7.45; 8.75)) \approx 0.95$.

15. Olgu antud normaaljaotusele alluva juhusliku suuruse X valimi $\{x_1, \ldots, x_{20}\}$ dispersioon $s^2 = 16$. Leidke dispersiooni DX usalduspiirkond usaldusnivoo $\beta = 0.98$ korral. V: $P(\mathrm{D}X \in (8.40; 39.84)) \approx 0.98$.

16. Olgu D$X = 9$, D$Y = 4$. On leitud $\overline{x} = 6$ ja $\overline{y} = 7.5$ sõltumatute valimite $\{x_1, \ldots, x_{30}\}$ ning $\{y_1, \ldots, y_{20}\}$ korral. Kontrollige hüpoteesi $H_0 : \mathrm{E}X = \mathrm{E}Y$ õigsust olulisuse nivool 0.05, kui: 1) $H_1 : \mathrm{E}X \neq \mathrm{E}Y$; 2) $H_1 : \mathrm{E}X > \mathrm{E}Y$; 3) $H_1 : \mathrm{E}X < \mathrm{E}Y$. V: 1) hüpotees H_0 lükatakse tagasi $H_1 : \mathrm{E}X \neq \mathrm{E}Y$ korral $(|\theta^*| \approx |-2.12| > \theta_{kr} \approx 1.96)$; 2) hüpoteesi H_0 ei lükata tagasi $H_1 : \mathrm{E}X > \mathrm{E}Y$ korral $(\theta^* \approx -2.12 < \theta_{kr} \approx 1.65)$; 3) hüpotees H_0 lükatakse tagasi $H_1 : \mathrm{E}X < \mathrm{E}Y$ korral $(\theta^* \approx -2.12 < -\theta_{kr} \approx -1.65)$.

17. Võrreldi kahe korvpalluri vabavisete tabavust meistrivõistlustel. Esimene neist tabas kuuekümne kaheksal viskel üheksakümne kolmest, teine neljakümne kolmel seitsmekümne kuuest. Kontrollige hüpoteesi $H_0 : p_1 = p_2$, kus p_1 ja p_2 on vastavalt esimese ja teise korvpalluri tabamise tõenäosus igal vabaviskel, olulisuse nivool 0.04, kui $H_1 : p_1 \neq p_2$. V: hüpotees H_0 lükatakse tagasi $(|\theta^*| \approx |2.25| > \theta_{kr} \approx 2.05)$.

18. Viis jahilaskurit sooritasid võistlustel igaüks sada lasku. Nende tabamuste arvud olid vastavalt 98, 85, 93, 89 ja 95. Uurige olulisuse nivool 0.025 hüpoteesi $H_0 : p_1 = p_2 = p_3 = p_4 = p_5$, kus p_i on i-nda laskuri tabamise tõenäosus igal lasul. V: hüpotees H_0 lükatakse tagasi $(q^* \approx 14.13 > q_{0.025;4} \approx 11.14)$.

19. Olgu kahe normaaljaotusele alluva juhusliku suuruse X ja Y valimid vastavalt $\{x_1, \ldots, x_{30}\}$ ja $\{y_1, \ldots, y_{40}\}$, kusjuures $s_x^2 = 6$ ja $s_y^2 = 8$. Kontrollige hüpoteesi $H_0 : \mathrm{D}X = \mathrm{D}Y$ olulisuse nivool 0.1, kusjuures $H_1 : \mathrm{E}X \neq \mathrm{E}Y$. V: hüpoteesi H_0 tagasi ei lükata $(z_{krv} \approx 0.57 < z^* \approx 1.33 < z_{krp} \approx 1.81)$.

20. Olgu viie normaaljaotusele alluva juhusliku suuruse X_k $(k = 1; 2; 3; 4; 5)$ valimi, iga neist mahuga sada, põhjal leitud $s_{x_1}^2 = 5$, $s_{x_2}^2 = 6$, $s_{x_3}^2 = 7$, $s_{x_4}^2 = 8$ ja $s_{x_5}^2 = 9$. Kontrollige hüpoteesi $H_0 : \mathrm{D}X_1 = \mathrm{D}X_2 = \mathrm{D}X_3 = \mathrm{D}X_4 = \mathrm{D}X_5$ olulisuse nivool 0.05, kui H_1 : kõik DX_k-d ei ole võrdsed. V: hüpotees H_0 lükatakse tagasi $(q^* \approx 15.39 > q_{0.05;4} \approx 9.49)$.

21.Olgu antud suuruse X variatsioonrea klassideks jaotus

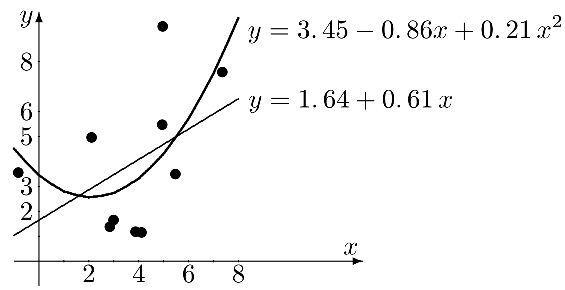
klass	(12; 14]	(14; 16]	(16; 18]	(18; 20]	(20; 22]	(22; 24]
keskpunkt	13	15	17	19	21	23
sagedus	7	12	22	29	21	9

Kontrollige olulisuse nivool 0.05 nii Pearsoni kui ka Kolmogorovi kriteeriumi abil hüpoteesi $H_0 : X$ allub normaaljaotusele. Kontrollime hüpoteesi: juhusliku suuruse X jaotusfunktsiooniks on just sellise kujuga $F(x)$. V: Pearsoni kriteeriumi järgi ei ole H_0 vastuolus valimiga $(q^* = 1.52 < q_{0.05;3} = 7.81)$, Kolmogorovi kriteeriumi järgi ei ole H_0 vastuolus valimiga $(D\sqrt{n} = 0.256 < 1.36 = \lambda_{0.05})$.

22. Katsetulemused on vormistatud tabeli kujul

k	1	2	3	4	5	6	7	8	9	10
x_k	7.35	3.87	5.47	2.11	2.84	4.96	−0.17	4.11	2.99	4.94
y_k	7.58	1.18	3.49	4.96	1.38	9.41	3.55	1.15	1.65	5.47

Leidke punkthinnangud $\overline{x}$, $\overline{y}$, s_x^2, s_y^2 k_{xy}^* ja kovariatsioonimaatriks ning korrelatsioonimaatriks. Leidke usaldusnivoole $\beta = 0.7$ vastavad keskväärtuste $\mathrm{E}X$ ja $\mathrm{E}Y$ usalduspiirkonnad. Leidke Y regressioonisirge ja regressiooniparabool X suhtes. Kandke joonisele korrelatsiooniväli, regressioonisirge peene ja regressiooniparabool jämeda joonega. V: $\overline{x} \approx 3.85$, $\overline{y} \approx 3.98$, $s_x^2 \approx 4.26$,
$s_y^2 = 8.23$, $k_{xy}^* = 2.59$, $\begin{pmatrix} 4.26 & 2.59 \\ & 8.23 \end{pmatrix}$, $\begin{pmatrix} 1.0 & 0.44 \\ & 1.0 \end{pmatrix}$,
$P(\mathrm{E}X \in (3.17; 4.52)) = P(\mathrm{E}Y \in (3.04; 4.92)) \approx 0.7$,
$y = 1.64 + 0.61x$, $y = 3.45 - 0.86x + 0.21x^2$,

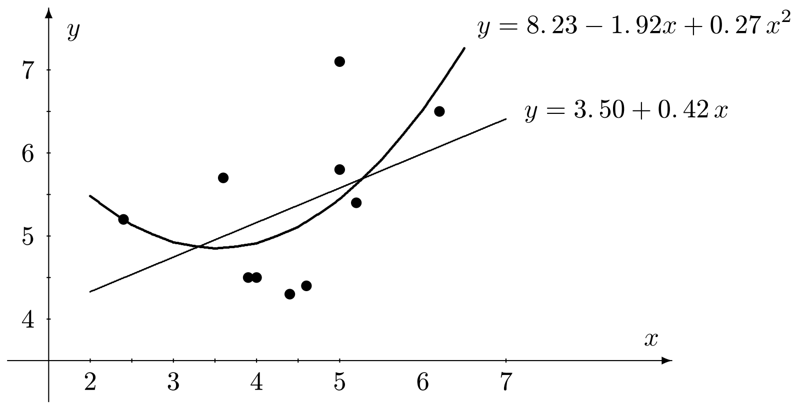


23. On antud juhusliku vektori (X, Y) lähtevalim

x_i	6.2	4.4	5.2	3.6	3.9	5.0,	2.4	4.6	4.0	5.0
y_i	6.5	4.3	5.4	5.7	4.5	7.1	5.2	4.4	4.5	5.8

Leidke: arvkarakteristikute $\mathrm{E}X$, $\mathrm{E}Y$, $\mathrm{D}X$, $\mathrm{D}Y$, σ_x, σ_y, $\mathrm{K}_{x,y}$, $r_{x,y}$ nihketa hinnangud $\overline{x}$, $\overline{y}$, s_x^2, s_x^2, s_x, s_y, $\mathrm{K}_{x,y}^*$, $r_{x,y}^*$; regressioonisirge $y = ax + b$ ja regressiooniparabool $y = ax^2 + bx + c$ (vähimruutude mõttes); keskväärtuse $\mathrm{E}X$ usaldusvahemik l_β usaldusnivool $\beta = 0.9$; keskväärtuse $\mathrm{E}X$ usaldusvahemik usaldusnivool $\beta = 0.9$, kui on teada lisaks, et X allub normaaljaotusele; $\mathrm{D}X$ usaldusvahemik usaldusnivool $\beta = 0.95$, kui lisaks on teada, et X allub normaaljaotusele. Skitseerige korrelatsiooniväli ja regressioonisirge ning regressiooniparabool. Kontrollige: hüpoteesi $H_0 : EX = EY$, võttes $\sigma_x \approx s_x$ ja $\sigma_y \approx s_y$, olulisuse nivool 0.05, kusjuures $H_1 : EX \neq EY$; hüpoteesi $H_0 : \mathrm{D}X = \mathrm{D}Y$ olulisuse nivool 0.1, kusjuures X ja Y on normaaljaotusega ja $H_1 : \mathrm{D}X \neq \mathrm{D}Y$.
V: $\overline{x} = 4.43$, $\overline{y} = 5.34$, $s_x^2 \approx 1.08$, $s_y^2 \approx 0.91$, $s_x \approx 1.04$, $s_y \approx 0.95$,
$\overline{xy} = 24.06$, $\mathrm{K}_{x,\,y}^* \approx 0.45$, $\mathrm{r}_{xy}^* \approx 0.45$;

$y = 3.50 + 0.42\,x$, $y = 8.23 - 1.92x + 0.27\,x^2$; $l_{0.9} \approx (3.89; 4.97)$,
$l_{0.9} \approx (3.83; 5.03)$, $l_{0.95} \approx (0.51; 3.59)$;



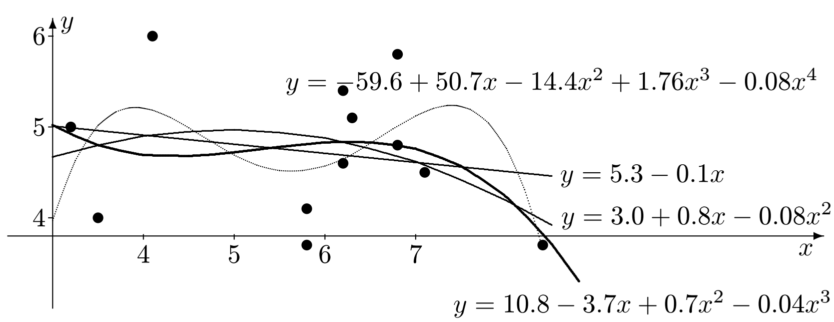
$|\theta^*| \approx |-2.04| > \theta_{kr} \approx 1.65 \Rightarrow$ hüpotees H_0 lükatakse tagasi,
$z_{krv} \approx 0.32 \leq z^* \approx 1.18 \leq z_{krp} \approx 3.18 \Rightarrow$ hüpotees H_0 on õige.

24. Katsetulemused on vormistatud tabeli kujul

x_i	6.2	6.2	3.2	6.8	4.1	5.8	8.4	7.1	6.8	5.8	3.5	6.3
y_i	5.4	4.6	5.0	5.8	6.0	3.7	3.7	4.5	4.8	4.1	4.0	5.1

Leidke esitatud valimi kovariatsioonimaatriks ja korrelatsioonimaatriks. Leidke vähimruutude meetodil regressioonijooned $y = ax + b$, $y = ax^2 + bx + c$, $y = ax^3 + bx^2 + cx + d$ ja $y = ax^4 + bx^3 + cx^2 + dx + e$. Skitseerige korrelatsiooniväli ja regressioonijooned.

V: $\begin{pmatrix} 2.36 & -0.24 \\ & 0.59 \end{pmatrix}$, $\begin{pmatrix} 1 & -0.20 \\ & 1 \end{pmatrix}$,



5.7 Lisad

Lisa 1. Kombinatoorika

Lihtsad ühendid

Mitu elementi on põhihulgas?	n	n	n
Mitmeelemendilisi alamhulki moodustatakse?	k	k	n
Kas elementide järjestus põhihulgas on oluline?	ei	jaa	jaa
Ülimalt mitu korda võib ühte elementi alamhulgas võtta?	1	1	1
Ühendite nimetused	Kombinatsioonid n elemendist k kaupa	Variatsioonid n elemendist k kaupa	Permutatsioonid n elemendist
Mitu erinevat ühendit saab moodustada?	C_n^k	V_n^k	P_n

$$C_n^k \equiv \binom{n}{k} = \frac{n!}{k!\,(n-k)!} = \frac{n\,(n-1)\cdots(n-k+1)}{k!},$$

$$V_n^k \equiv A_n^k = \frac{n!}{(n-k)!} = n\,(n-1)\cdots(n-k+1)\,,$$

$$P_n = V_n^n = n! = 1\cdot 2\cdots n,$$

$$0! = 1, \quad C_n^0 = V_n^0 = 1, \quad C_n^{k-1} + C_n^k = C_{n+1}^k,$$

$$\sum_{k=0}^{n} C_n^k = 2^n, \quad \sum_{k=0}^{n} (-1)^k\, C_n^k = 0, \quad \sum_{k=1}^{n} k\, C_n^k = 2^{n-1} n,$$

$$\sum_{k=0}^{n} \left(C_n^k\right)^2 = C_{2n}^n, \quad \sum_{k=1}^{n} (-1)^{k+1}\, k\, C_n^k = 0, \quad \sum_{k=0}^{m-n} C_{n+k}^k = C_{m+1}^{n+1},$$

$$\sum_{k=0}^{n} C_n^k C_m^k = C_{m+n}^n \;\; (m \ge n)\,, \quad \Gamma(x) \overset{\text{def.}}{=} \int_0^{+\infty} e^{-t} t^{x-1} dt \;\; (x > 0)\,,$$

$$\Gamma(x+1) = x\,\Gamma(x)\,, \quad \Gamma(1) = 1, \quad \Gamma(n+1) = n! \;\; (n \in \mathbf{N})\,.$$

Kordumistega ühendid

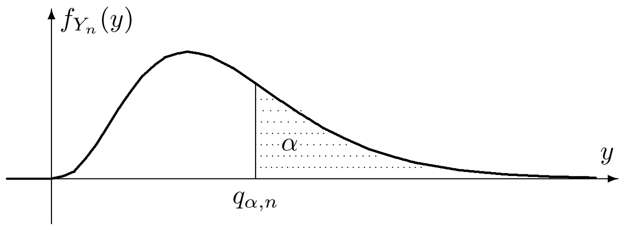
Mitu elementi on põhihulgas?	n erinevat elementi, igat neist suvaline arv eksemplare	n erinevat elementi, igat neist suvaline arv eksemplare	n elementi, nende seas on vaid k erinevat, i-ndat k_i eksemplari, $\sum_{i=1}^{n} k_i = n$
Mitmeelemendilisi alamhulki moodustatakse?	k	k	n
Kas elementide järjestus põhihulgas on oluline?	ei	jaa	eristatavuse piires jaa
Ülimalt mitu korda võib ühte elementi alamhulgas võtta?	k	k	$k_1, \ldots, k_n$
Ühendite nimetused	Kordumistega kombinatsioonid n elemendist k kaupa	Kordumistega variatsioonid n elemendist k kaupa	Kordumistega permutatsioonid n elemendist
Mitu erinevat ühendit saab moodustada?	Γ_n^k	W_n^k	$P_n(k_1, \ldots, k_n)$

$$\Gamma_n^k = C_{n+k-1}^k, \quad W_n^k = n^k, \quad P_k(k_1, \ldots, k_n) = \frac{n!}{k_1! \cdots k_n!}.$$

Lisa 2. Funktsiooni $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-t^2/2} dt$ tabel

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0.00	0.0000	0.58	0.2190	1.16	0.3770	1.74	0.4591
0.02	0.0080	0.60	0.2257	1.18	0.3810	1.76	0.4608
0.04	0.0160	0.62	0.2324	1.20	0.3849	1.78	0.4625
0.06	0.0239	0.64	0.2389	1.22	0.3888	1.80	0.4641
0.08	0.0319	0.66	0.2454	1.24	0.3925	1.82	0.4656
0.10	0.0398	0.68	0.2517	1.26	0.3962	1.84	0.4671
0.12	0.0478	0.70	0.2580	1.28	0.3997	1.86	0.4686
0.14	0.0557	0.72	0.2642	1.30	0.4032	1.88	0.4699
0.16	0.0636	0.74	0.2703	1.32	0.4066	1.90	0.4713
0.18	0.714	0.76	0.2764	1.34	0.4099	1.92	0.4726
0.20	0.0793	0.78	0.2823	1.36	0.4131	1.94	0.4738
0.22	0.0871	0.80	0.2881	1.38	0.4162	1.96	0.4750
0.24	0.0948	0.82	0.2939	1.40	0.4192	1.98	0.4761
0.26	0.1026	0.84	0.2995	1.42	0.4222	2.00	0.4772
0.28	0.1103	0.86	0.3051	1.44	0.4251	2.10	0.4821
0.30	0.1179	0.88	0.3106	1.46	0.4279	2.20	0.4861
0.32	0.1255	0.90	0.3159	1.48	0.4306	2.30	0.4893
0.34	0.1331	0.92	0.3212	1.50	0.4332	2.40	0.4918
0.36	0.1406	0.94	0.3264	1.52	0.4357	2.50	0.4938
0.38	0.1480	0.96	0.3315	1.54	0.4382	2.60	0.4953
0.40	0.1554	0.98	0.3365	1.56	0.4406	2.70	0.4965
0.42	0.1628	1.00	0.3413	1.58	0.4429	2.80	0.4974
0.44	0.1700	1.02	0.3461	1.60	0.4452	2.90	0.4981
0.46	0.1772	1.04	0.3508	1.62	0.4474	3.00	0.4986
0.48	0.1844	1.06	0.3554	1.64	0.4495	3.10	0.4990
0.50	0.1915	1.08	0.3599	1.66	0.4515	3.20	0.4993
0.52	0.1985	1.10	0.3643	1.68	0.4535	3.30	0.4995
0.54	0.2054	1.12	0.3686	1.70	0.4554	3.40	0.4997
0.56	0.2123	1.14	0.3729	1.72	0.4573	3.50	0.4998

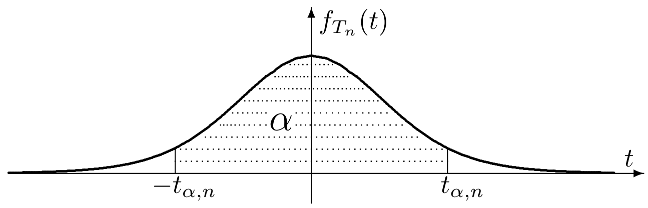
Lisa 3. χ^2-jaotus. Funktsiooni $(\alpha, n) \mapsto q_{\alpha,n}$, kus $\int_{q_{\alpha,n}}^{+\infty} f_{Y_n}(y)dy = \alpha$, tabel



$n \backslash \alpha$	0.99	0.98	0.95	0.90	0.80	0.70	0.50	0.30	0.20	0.10	0.02	0.01
1	0.00	0.00	0.00	0.02	0.06	0.15	0.45	1.07	1.64	2.71	5.41	6.64
2	0.02	0.04	0.10	0.21	0.45	0.71	1.39	2.41	3.22	4.60	7.82	9.21
3	0.11	0.18	0.35	0.58	1.00	1.42	2.37	3.66	4.64	6.25	9.84	11.3
4	0.30	0.43	0.71	1.06	1.65	2.20	3.36	4.88	5.99	7.78	11.7	13.3
5	0.55	0.75	1.14	1.61	2.34	3.00	4.35	6.06	7.29	9.24	13.4	15.1
6	0.87	1.13	1.63	2.20	3.07	3.83	5.35	7.23	8.56	10.6	15.0	16.8
7	1.24	1.56	2.17	2.83	3.82	4.67	6.35	8.38	9.80	12.0	16.6	18.5
8	1.65	2.03	2.73	3.49	4.59	5.53	7.34	9.52	11.0	13.4	18.2	20.1
9	2.09	2.53	3.32	4.17	5.38	6.39	8.34	10.7	12.2	14.7	19.7	21.7
10	2.56	3.06	3.94	4.86	6.18	7.27	9.34	11.8	13.4	16.0	21.2	23.2
11	3.05	3.61	4.58	5.58	6.99	8.15	10.3	12.9	14.6	17.3	22.6	24.7
12	3.57	4.18	5.23	6.30	7.81	9.03	11.3	14.0	15.8	18.5	24.1	26.2
13	4.11	4.76	5.89	7.04	8.63	9.93	12.3	15.1	17.0	19.8	25.5	27.7
14	4.66	5.37	6.57	7.79	9.47	10.8	13.3	16.2	18.1	21.1	26.9	29.1
15	5.23	5.98	7.26	8.55	10.3	11.7	14.3	17.3	19.3	22.3	28.3	30.6
16	5.81	6.61	7.96	9.31	11.1	12.6	15.3	18.4	20.5	23.5	29.6	32.0
17	6.41	7.26	8.67	10.1	12.0	13.5	16.3	19.5	21.6	24.8	31.0	33.4
18	7.02	7.91	9.39	10.9	12.9	14.4	17.3	20.6	22.8	26.0	32.3	34.8
19	7.63	8.57	10.1	11.6	13.7	15.3	18.3	21.7	23.9	27.2	33.7	36.2
20	8.26	9.24	10.8	12.4	14.6	16.3	19.3	22.8	25.0	28.4	35.0	37.6
21	8.9	9.92	11.6	13.2	15.4	17.2	20.3	23.9	26.2	29.6	36.3	38.9
22	9.54	10.6	12.3	14.0	16.3	18.1	21.3	24.9	27.3	30.8	37.7	40.3
23	10.2	11.3	13.1	14.8	17.2	19.0	22.3	26.0	28.4	32.0	39.0	41.6
24	10.9	12.0	13.8	15.7	18.1	19.9	23.3	27.1	29.6	33.2	40.3	43.0
25	11.5	12.7	14.6	16.5	18.9	20.9	24.3	28.2	30.7	34.4	41.7	44.3
26	12.2	13.4	15.4	17.3	19.8	21.8	25.3	29.2	31.8	35.6	42.9	45.6
27	12.9	14.1	16.1	18.1	20.7	22.7	26.3	30.3	32.9	36.7	44.1	47.0
28	13.6	14.8	16.9	18.9	21.6	23.6	27.3	31.4	34.0	37.9	45.4	48.3
29	14.3	15.6	17.7	19.8	22.5	24.6	28.3	32.5	35.1	39.1	46.7	49.6

Lisa 4. Studenti jaotus. Funktsiooni $(\alpha, n) \mapsto t_{\alpha,n}$, kus

$$\int_{-t_{\alpha,n}}^{t_{\alpha,n}} f_{T_n}(t)dt = 2\int_0^{t_{\alpha,n}} f_{T_n}(t)dt = \alpha,$$

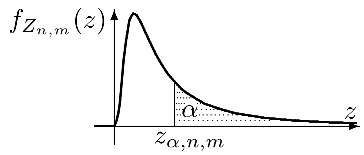


tabel

$n\backslash\alpha$	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.95	0.98	0.99
1	0.16	0.32	0.51	0.73	1.00	1.38	1.96	3.08	6.31	12.7	31.8	63.7
2	0.14	0.29	0.44	0.62	0.82	1.06	1.34	1.89	2.92	4.30	6.96	9.92
3	0.14	0.28	0.42	0.58	0.76	0.98	1.25	1.64	2.35	3.18	4.54	5.84
4	0.13	0.27	0.41	0.57	0.74	0.94	1.19	1.53	2.13	2.78	3.75	4.60
5	0.13	0.27	0.41	0.56	0.73	0.92	1.16	1.48	2.01	2.57	3.36	4.03
6	0.13	0.26	0.40	0.55	0.72	0.91	1.13	1.44	1.94	2.45	3.14	3.71
7	0.13	0.26	0.40	0.55	0.71	0.90	1.12	1.41	1.89	2.36	3.00	3.50
8	0.13	0.26	0.40	0.54	0.70	0.89	1.11	1.40	1.86	2.31	2.90	3.35
9	0.13	0.26	0.40	0.54	0.70	0.88	1.10	1.38	1.83	2.26	2.82	3.25
10	0.13	0.26	0.40	0.54	0.70	0.88	1.09	1.37	1.81	2.23	2.76	3.17
11	0.13	0.26	0.40	0.54	0.70	0.88	1.09	1.36	1.80	2.20	2.72	3.11
12	0.13	0.26	0.39	0.54	0.69	0.87	1.08	1.36	1.78	2.18	2.68	3.05
13	0.13	0.26	0.39	0.54	0.69	0.87	1.08	1.35	1.77	2.16	2.65	3.01
14	0.13	0.26	0.39	0.54	0.69	0.87	1.08	1.34	1.76	2.14	2.62	2.98
15	0.13	0.26	0.39	0.54	0.69	0.87	1.07	1.34	1.75	2.13	2.60	2.95
16	0.13	0.26	0.39	0.53	0.69	0.86	1.07	1.34	1.75	2.12	2.58	2.92
17	0.13	0.26	0.39	0.53	0.69	0.86	1.07	1.33	1.74	2.11	2.57	2.90
18	0.13	0.26	0.39	0.53	0.69	0.86	1.07	1.33	1.73	2.10	2.55	2.88
19	0.13	0.26	0.39	0.53	0.69	0.86	1.07	1.33	1.73	2.09	2.54	2.86
20	0.13	0.26	0.39	0.53	0.69	0.86	1.06	1.32	1.72	2.09	2.53	2.84
25	0.13	0.26	0.39	0.53	0.68	0.86	1.06	1.32	1.71	2.06	2.48	2.79
30	0.13	0.26	0.39	0.53	0.68	0.85	1.05	1.31	1.70	2.04	2.46	2.75
40	0.13	0.25	0.39	0.53	0.68	0.85	1.05	1.30	1.68	2.02	2.42	2.70
60	0.13	0.25	0.39	0.53	0.68	0.85	1.05	1.30	1.67	2.00	2.39	2.66
120	0.13	0.25	0.39	0.53	0.68	0.84	1.04	1.29	1.66	1.98	2.36	2.62
∞	0.13	0.25	0.38	0.52	0.67	0.84	1.04	1.28	1.64	1.96	2.33	2.58

Lisa 5. Fisheri jaotus. Funktsiooni $(\alpha, n, m) \mapsto z_{\alpha,n,m}$, kus

$$P\left(Z_{n,m} > z_{\alpha,n,m}\right) = \\ = \int_{z_{\alpha,n,m}}^{+\infty} f_{Z_{n,m}}(z)dz = \alpha,$$



tabel $\alpha = 0.05$ korral

$m\backslash n$	1	2	3	4	6	8	10	12	15	20	30	∞
1	161	200	216	225	234	239	242	244	246	248	250	254
2	18.5	19.0	19.2	19.2	19.3	19.4	19.4	19.4	19.4	19.4	19.5	19.5
3	10.1	9.55	9.28	9.12	8.94	8.85	8.79	8.74	8.70	8.66	8.62	8.53
4	7.71	6.94	6.59	6.39	6.16	6.04	5.96	5.91	5.86	5.80	5.75	5.63
5	6.61	5.79	5.41	5.19	4.95	4.82	4.74	4.68	4.62	4.56	4.50	4.36
6	5.99	5.14	4.76	4.53	4.28	4.15	4.06	4.00	3.94	3.87	3.81	3.67
7	5.59	4.74	4.35	4.12	3.87	3.73	3.64	3.57	3.51	3.44	3.38	3.23
8	5.32	4.46	4.07	3.84	3.58	3.44	3.35	3.28	3.22	3.15	3.08	2.93
9	5.12	4.26	3.86	3.63	3.37	3.23	3.14	3.07	3.01	2.94	2.86	2.71
10	4.96	4.10	3.71	3.48	3.22	3.07	2.98	2.91	2.85	2.77	2.70	2.54
11	4.84	3.98	3.59	3.36	3.09	2.95	2.85	2.79	2.72	2.65	2.57	2.40
12	4.75	3.89	3.49	3.26	3.00	2.85	2.75	2.69	2.62	2.54	2.47	2.30
13	4.67	3.81	3.41	3.18	2.92	2.77	2.67	2.60	2.53	2.46	2.38	2.21
14	4.60	3.74	3.34	3.11	2.85	2.70	2.60	2.53	2.46	2.39	2.31	2.13
15	4.54	3.68	3.29	3.06	2.79	2.64	2.54	2.48	2.40	2.33	2.25	2.07
16	4.49	3.63	3.24	3.01	2.74	2.59	2.49	2.42	2.35	2.28	2.19	2.01
17	4.45	3.59	3.20	2.96	2.70	2.55	2.45	2.38	2.31	2.23	2.15	1.96
18	4.41	3.55	3.16	2.93	2.66	2.51	2.41	2.34	2.27	2.19	2.11	1.92
19	4.38	3.52	3.13	2.90	2.63	2.48	2.38	2.31	2.23	2.16	2.07	1.88
20	4.35	3.49	3.10	2.87	2.60	2.45	2.35	2.28	2.20	2.12	2.04	1.84
21	4.32	3.47	3.07	2.84	2.57	2.42	2.32	2.25	2.18	2.10	2.01	1.81
22	4.30	3.44	3.05	2.82	2.55	2.40	2.30	2.23	2.15	2.07	1.98	1.78
23	4.28	3.42	3.03	2.80	2.53	2.37	2.27	2.20	2.13	2.05	1.96	1.76
24	4.26	3.40	3.01	2.78	2.51	2.36	2.25	2.18	2.11	2.03	1.94	1.73
25	4.24	3.39	2.99	2.76	2.49	2.34	2.24	2.16	2.09	2.01	1.92	1.71
30	4.17	3.32	2.92	2.69	2.42	2.27	2.16	2.09	2.01	1.93	1.84	1.62
40	4.08	3.23	2.84	2.61	2.34	2.18	2.08	2.00	1.92	1.84	1.74	1.51
60	4.00	3.15	2.76	2.53	2.25	2.10	1.99	1.92	1.84	1.75	1.65	1.39
120	3.92	3.07	2.68	2.45	2.17	2.02	1.91	1.83	1.75	1.66	1.55	1.25
∞	3.84	3.00	2.60	2.37	2.10	1.94	1.83	1.75	1.67	1.57	1.46	1.00

ja $\alpha = 0.01$ korral

$m\backslash n$	1	2	3	4	6	8	10	12	15	20	∞
1	4052	5000	5403	5625	5859	5982	6056	6106	6157	6209	6366
2	98.50	99.00	99.17	99.25	99.33	99.37	99.40	99.42	99.43	99.45	99.50
3	34.12	30.82	29.46	28.71	27.91	27.49	27.23	27.05	26.87	26.69	26.13
4	21.20	18.00	16.69	15.98	15.21	14.80	14.55	14.37	14.20	14.02	13.46
5	16.26	13.27	12.06	11.39	10.67	10.29	10.05	9.89	9.72	9.55	9.02
6	13.75	10.92	9.78	9.15	8.47	8.10	7.87	7.72	7.56	7.40	6.88
7	12.25	9.55	8.45	7.85	7.19	6.84	6.62	6.47	6.31	6.16	5.65
8	11.26	8.65	7.59	7.01	6.37	6.03	5.81	5.67	5.52	5.36	4.86
9	10.56	8.02	6.99	6.42	5.80	5.47	5.26	5.11	4.96	4.81	4.31
10	10.04	7.56	6.55	5.99	5.39	5.06	4.85	4.71	4.56	4.41	3.91
11	9.65	7.21	6.22	5.67	5.07	4.74	4.54	4.40	4.25	4.10	3.60
12	9.33	6.93	5.95	5.41	4.82	4.50	4.30	4.16	4.01	3.86	3.36
13	9.07	6.70	5.74	5.21	4.62	4.30	4.10	3.96	3.82	3.66	3.17
14	8.86	6.51	5.56	5.04	4.46	4.14	3.94	3.80	3.66	3.51	3.00
15	8.68	6.36	5.42	4.89	4.32	4.00	3.80	3.67	3.52	3.37	2.87
16	8.53	6.23	5.29	4.77	4.20	3.89	3.69	3.55	3.41	3.26	2.75
17	8.40	6.11	5.18	4.67	4.10	3.79	3.59	3.46	3.31	3.16	2.65
18	8.29	6.01	5.09	4.58	4.01	3.71	3.51	3.37	3.23	3.08	2.57
19	8.18	5.93	5.01	4.50	3.94	3.63	3.43	3.30	3.15	3.00	2.49
20	8.10	5.85	4.94	4.43	3.87	3.56	3.37	3.23	3.09	2.94	2.42
21	8.02	5.78	4.87	4.37	3.81	3.51	3.31	3.17	3.03	2.88	2.36
22	7.95	5.72	4.82	4.31	3.76	3.45	3.26	3.12	2.98	2.83	2.31
23	7.88	5.66	4.76	4.26	3.71	3.41	3.21	3.07	2.93	2.78	2.26
24	7.82	5.61	4.72	4.22	3.67	3.36	3.17	3.03	2.89	2.74	2.21
25	7.77	5.57	4.68	4.18	3.63	3.32	3.13	2.99	2.85	2.70	2.17
26	7.72	5.53	4.64	4.14	3.59	3.29	3.09	2.96	2.81	2.66	2.13
27	7.68	5.49	4.60	4.11	3.56	3.26	3.06	2.93	2.78	2.63	2.10
28	7.64	5.45	4.57	4.07	3.53	3.23	3.03	2.90	2.75	2.60	2.06
29	7.60	5.42	4.54	4.04	3.50	3.20	3.00	2.87	2.73	2.57	2.03
30	7.56	5.39	4.51	4.02	3.47	3.17	2.98	2.84	2.70	2.55	2.01
40	7.31	5.18	4.31	3.83	3.29	299	2.80	2.66	2.52	2.37	1.80
60	7.08	4.98	4.13	3.65	3.12	2.82	2.63	2.50	2.35	2.20	1.60
120	6.85	4.79	3.95	3.48	2.96	2.66	2.47	2.34	2.19	2.03	1.38
∞	6.63	4.61	3.78	3.32	2.80	2.51	2.32	2.18	2.04	1.88	1.00

Kirjandus

[1] Dougherty, E. R. Probability and statistics for the engineering, computing, and physical sciences. New York, Prentice-Hall, 1990.

[2] Everitt, Brian S., Der, G. A handbook of statistical analyses using SAS. CRC Press, 2001.

[3] Everitt, Brian S., Everitt, B. S. A handbook of statistical analyses using S-Plus. London, Chapman and Hall, 2001.

[4] Grimmet, G., Stirzaker, D. One thousand exercises in Probability. Oxford University Press, 2001.

[5] Gurski, J. I. Tõenäosusteooria ja matemaatilise statistika elemendid. Tallinn, Valgus, 1986.

[6] Hald, A. Statistical theory with engineering applications.

[7] Helstrom, C. W. Probability and stohhastic processes for engineers.

[8] Jõgi, A. Tõenäosusteooria I. Tallinn, TTÜ Kirjastus, 2000.

[9] Jõgi, A. Tõenäosusteooria II. Tallinn, TTÜ Kirjastus, 2000.

[10] Kiviste, A. Matemaatiline statistika MS Excel keskkonnas. Tallinn, GT Tarkvara, 1999.

[11] Käerdi, H. Statistika ja tõenäosusteooria alused. Tallinn, Sisekaitseakadeemia, 1999.

[12] Lõhmus, A., Petersen, I., Roos, H. Kõrgema matemaatika ülesannete kogu. Tallinn, Valgus, 1989.

[13] Milton, J. S., Jesse, C. A. Probability and statistics in engineering and coputing sciences. New York, McGraw-Hill,

[14] Peyton, Z., Peibles, J. R. Probability, random variables, and random signal principles.

[15] Rabe-Hesketh, S., Everitt, B. S. Handbook of statistical analyses Using Stata. CRC Press, 2000.

[16] Rice, J. A. Mathematical statistics and data analysis. Belmont, Duxbury Press, 1995.

[17] Sheskin, D. J. Handbook of parametric and nonparametric statistical procedures. CRC Press, 2000.

[18] Terrell, G. R. Mathematical statistics. Springer, 1999.

[19] Tiit, E.-M., Möls, M. Rakendusstatistika algkursus. Tartu, 1997.

[20] Tiit, E.-M., Parring, A., Möls, T. Tõenäosusteooria ja matemaatiline statistika. Tallinn, Valgus, 1977.

[21] Tõenäosusteooria ja matemaatiline statistika. Kõrgema matemaatika teatmik IV. Tallinn, TPI, 1979.

[22] Venables, W. N., Smith, D. M. An introduction to S-PLUS. Insightful Corporation,

[23] Venables, W. N., Smith, D. M. An introduction to R. www.r-project.org.

[24] Walpole, R. E., Myers, R. H., Myers, S. L., Ye, K. Probability and statistics for engineers and scientists. Prentice Hall, 2002.

[25] Ventsel, J. S. Teorija verojatnostei, Moskva, Võshaja shkola, 1998.

[26] Ventsel, J. S., Ovcharov, L. A. Teorija verojatnostei. Zadachi i uprazhnenija. Moskva, Nauka, 1969.

[27] Gmurman, V. J. Rukovodstvo k resheniju zadach po teori verojatnostei i matematicheskoi statistike. Moskva, Võshaja shkola, 1999.

[28] Gmurman, V. J. Teorija verojatnostei i matematicheskaja statistika. Moskva, Võshaja shkola, 1972.

[29] Koroljuk, V. S., Portenko, N. I., Skorohod, A. V., Turbin, A. F. Spravochnik po teori verojatnostei i matematicheskoi statistike. Moskva, Nauka, 1985.

[30] Cramér, H. Matematicheskije metodõ statistiki. Moskva, Mir, 1975.

[31] Kremer, N. Sh. Teorija verojatnostei i matematicheskaja statistika. Moskva, Juniti, 2000.

[32] Prohhorov, J V., ... Verojatnost i matematicheskaja statistika. Entsiklopedija. Moskva, Boshaja rossiskaja entsiklopedija, 1999.

[33] Pugachov, V, S. Vvedenije v teoriju verojatnostei. Moskva, Nauka, 1968.

[34] Pugachov, V, S. Teorija verojatnostei i matematicheskaja statistika. Moskva, Fizmatlit, 2002.

[35] Stojanov, I. Kontraprimerõ v teori verojatnostei. Moskva, Faktorial, 1999.

Indeks