Tartu Ülikool
Keskkonnafüüsika instituut

JAOTUSFUNKTSIOONID JA MÕÕTEMÄÄRAMATUSED

I VIHIK

LOENGUKONSPEKT

Rein Rõõm

TARTU 2005

Käesolev loengukonspekt *JAOTUSFUNKTSIOONID JA MÕÕTEMÄÄRAMATUSED* on mõeldud kasutamiseks eeskätt füüsika-keemiateaduskonna esimese kursuse üliõpilastele, kes läbivad füüsika, keemia ja materjaliteaduse bakalaureuseõppe kavu. Käesolev kursus eelneb *Füüsikaliste mõõtmiste aluste* kursusele ja on selle eeldusaineks .

Kursuse *JAOTUSFUNKTSIOONID JA MÕÕTEMÄÄRAMATUSED* sisuks on tutvumine tõenäosusteooria ja matemaatilise statistika alustega mahus, mis on minimaalselt tarvilik selleks, et mõista looduses ja ühiskonnas aset leidvaid juhuslikke nähtusi ja protsesse, õppida neid täppisteaduslike meetoditega analüüsima ning välja selgitama neis peituvaid seaduspärasusi. See on kursuse esimene, laiem ülesanne ja eesmärk. Teine, kitsam rakenduseesmärk on tutvumine füüsikaliste mõõtmiste alustega. Mõõtmisoskus on vajalik teadusuuringutes, tehnilises tegevuses, meditsiinis, majanduses, spordis, koduses majapidamises. Mõõtmisoskus kuulub mistahes tehnoloogilise eriala alustesse. Keskse tähtsusega igasugustel mõõtmistel, füüsikalistel mõõtmistel sealhulgas, on mõõtmistulemuste statistiline töötlus ja mõtestamine tõenäosuslikes kategooriates. Suure osa füüsikaliste mõõtmiste alustest moodustabki mõõtmistulemuste töötlemine. Kokkuvõtvalt, kui kursuse esimene pool on sissejuhatus tõenäosusteooriasse ja matemaatilisse statistikasse, siis teine pool tegeleb omandatud matemaatilise aparaadi rakendamisega füüsikalistel mõõtmistel.

Sissejuhatus

Looduses ja ühiskonnas – sealhulgas teaduslikus uurimistöös, mitmesugustel tehnikaaladel ning tööstuslikus masstootmises - on sageli tegemist korduvalt esinevate nähtustega, mis toimuvad või kulgevad iga kord veidi erinevalt. Niisuguseid nähtusi nimetatakse juhuslikeks.

Näiteid juhuslikest nähtustest:

- Tänaval vastutulevate inimeste pikkus (vanus, kehakaal, ...);
- Märklaua tabamine märkilaskmisel püssi või vibuga;
- Radioaktiivse elemendi aatomi eluiga;
- Tuule suund ja suurus;
- Viking-loto loosimise kuus võidunumbrit;
- Masstootmises konveierilt tulevate seadmete või detailide parameetrid (näiteks detaili pikkus);
- Ühe ja sama eseme pikkus (kaal, elektrijuhtivus, ...) erinevatel (korduvatel) mõõtmistel;
- ....

Üldiselt saab kõiki nähtusi ja sündmusi liigitada kindlateks, võimatuteks ja juhuslikeks.

Nähtust nimetatakse **kindlaks**, kui ta seatud tingimustel alati toimub. Nähtust nimetatakse **võimatuks**, kui see seatud tingimustel ei toimu iialgi.

- Näide.

  Nähtus, mis seisneb selles, et käestpillatud ese kukub maha (ja ei lähe selle asemel hoopis lendu), on kindel sündmus Maal, kuid ei ole seda mitte ümber Maa tiirlevas orbitaaljaamas. Samal ajal sündmus „käestpillatud pliiats jääb õhku hõljuma" on Maal võimatu.

Nähtust nimetatakse **juhuslikuks**, kui see seatud tingimustel võib toimuda või mitte toimuda. Võime öelda ka nii, et seatud tingimustel juhuslik nähtus kulgeb korduvatel katsetel iga kord erinevalt.

- Näide.

  Täringu viskamisel silmade 1, 2, 3, 4, 5 või 6 saamine on juhuslik nähtus (eeldusel, et täring on korrapärane ja valmistatud homogeensest materjalist).

Paljudel nähtustel on harilikult korduv iseloom, mis on korduvalt realiseeritavate põhjuste (tingimuste) tagajärg. Niisuguste korduvate nähtuste kohta võib kujuneda teatud "objektiivne" ettekujutus nende juhuslikkuse või võimalikkuse kohta. Osutub, **korduvaid juhuslikke nähtusi iseloomustavad seaduspärasused, mis ei ole juhuslikud**. Selliseid seaduspärasusi nimetatakse **tõenäosuslikeks.**

- Näide.

  Tööpingil ühesugustel tingimustel toodetud detailide mõõtmed (pikkus, laius, diameeter) on üldiselt erinevad, kuid nende mõõtmete arvväärtused kõiguvad väikeses vahemikus, erinedes küllalt vähe mingist keskmisest. Kõikumistel on küll juhuslik iseloom, kuid suurtes partiides on erinevate partiide detailide mõõtmete aritmeetilised keskmised ligikaudu võrdsed. Enamgi – mõõtmete hälbed aritmeetilistest keskmistest erinevates partiides on peaaegu ühesuguse sagedusega.

Juhuslikes suurustes või nähtustes ilmnevaid mittejuhuslikke seaduspärasusi uurib, põhjendab ja selgitab **tõenäosusteooria**. Juhuslike suuruste tõenäosuslike seaduspärasuste eksperimentaalse uurimisega ja põhjendamisega tegeleb **matemaatiline statistika.** Kui tõenäosusteooria tegeleb peaasjalikult juhuslike sündmuste ja juhuslike suuruste analüüsiga, keskendudes tõenäosustele, jaotusfunktsioonidele ja juhulike suuruste arvkarakteristikute uurimisele, siis matemaatilise statistika eesmärgiks on korduvatel mõõtmistel või vaatlustel saadud andmete tõenäosuslik--statistiliste seaduspärasuste väljaselgitamise ja põhjendamise. Selleks kasutab matemaatiline statistika nn. *suurte arvude seadusi* – seaduspärasusi, mis ilmnevad väga suure arvu ühesugustes tingimustes toimuvate sõltumatute katsete keskmistamisel. Suurte arvude seaduste kehtivust tõestatakse omakorda tõenäosusteooria meetodite ja vahenditega.

Tõenäosusteooria kui rakendusmatemaatika haru tekkis 15. ja 16. sajandil ja arenes sõltuvalt praktika vajadustest. Tõenäosusteooria oma arengu algstaadiumis oli seotud majanduslike ülesannetega (näiteks elukindlustus, mis on seotud niisuguste nähtustega nagu haigestumus, suremus) ja hasartmängudega. Tõenäosusteooria arengut stimuleeris praktiliste ja teaduslike probleemide hulk, mis omakorda põhinesid vaatlusandmetel ja teoreetilistel üldistustel.

Tänapäeval on tõenäosuse mõiste, tõenäosuslik-teoreetiline meetod, levinud mitte ainult majanduses, tööstuses ja tehnikaaladel, vaid ka rida täppisteadusi rajaneb tõenäosuslik-teoreetilistel alustel. Näitena olgu nimetatud termodünaamika (statistiline füüsika) ja kogu mikrofüüsika alustades kvantmehaanikaga.

Käesoleva kursusel on kaks eesmärki.

Esimene, laiem ülesanne on anda ülevaade tõenäosusteooria ja matemaatilise statistika alustest mahus, mis on tarvilik selleks, et mõista looduses ja ühiskonnas aset leidvaid juhuslikke nähtusi ja protsesse „üleüldse", õppida neid täppisteaduslikult analüüsima ning välja selgitama juhuslikes nähtustes ja sündmustes peituvaid seaduspärasusi.

Teine, kitsam ja rakenduslikum eesmärk on tutvumine füüsikaliste mõõtmiste tõenäosus-teoreetiliste alustega. Mõõtmistega tuleb tegemist teha kogu füüsikapraktikumi ulatuses, samuti keemia praktikumides. Suure osa praktikumide töömahust moodustab mõõtmistulemuste statistiline töötlus, mille eesmärgiks on võimalikult täpsete mõõtmistulemuste saamine ja mõõtmise kui juhuslikkust sisaldava protsessi usaldusväärsuse hindamine (mõõtemääramatuste analüüs). Nagu pikaajaline kogemus näitab, on see osa üliõpilastele üks kõige raskemini omandatavaid mõõtmispraktikumide alalõike. Seetõttu on käesolev kursus füüsikaliste mõõtmiste aluste omandamiseks vajalik ja oluline.

1. JUHUSLIKUD SÜNDMUSED

1.1. Sündmuste liigitamine

Iga teadusharu põhineb teatud arvul põhimõistetel, mille abil antakse teooria loogiline ülesehitus, s.t määratakse keerukamad mõisted. Tõenäosusteoorias on esimesteks niisugusteks mõisteteks katse (eksperiment) ja sündmus. Neid ei defineerita, vaid selgitatakse näidetega. Üldiselt märgime, et **sündmuseks** võime nimetada kõigepealt igat loodusnähtust. Harilikult nii nimetame nähtust, mis mingil põhjusel on meie huvisfääris ja püüame seda uurida ja tunnetada. Isoleeritud nähtusi looduses ei eksisteeri. Mingi nähtuse tekkimiseks peavad olema täidetud teatud tingimused.

Täpsustame nüüd katse ja sündmuse mõisteid. **Katse** all mõistame teatud tingimuste kompleksi realiseerumist, mille tulemusena võivad toimuda mingid sündmused. Tingimused võivad olla loodud kunstlikult (tahtlikult), või nad eksisteerivad sõltumatult eksperimentaatorist. Võime öelda ka, et **katse** on niisugune olukord või seisund, mille tulemusena võivad toimuda mingid sündmused. **Sündmuse** all mõistame igat fakti, mis katse tulemusena võib toimuda või mitte. Rõhutame, et sõltuvalt sündmuse sisust kasutame nimetuse *katse* asemel veel nimetusi **eksperiment** või **vaatlus**. Niisiis katse tulemusena toimuvad teatavad (aimatavad) sündmused. Katse on määratletud, kui on teada katse tingimused ja oodatavate sündmuste hulk. Katse võib olla korraldatud kindla eesmärgi suunitlusega, planeerimisega, aga ka passiivse vaatlusena, lihtsalt andmete registreerimisega. Rõhutame veel, et katse ei tarvitse olla korraldatud inimese poolt; katse võib kulgeda sõltumatult inimesest. Siinjuures on inimene ainult vaatleja osas või toimuva registreerija.

Katsed võib jaotada kahte klassi. Esimeses klassis on katsetulemused ettearvatavad loodusteaduste seaduste alusel. Niisuguseid katseid nimetatakse **determineeritud katseteks**. Teise klassi arvame sündmused, mille korral samadel katsetingimustel võivad esineda üksteist välistavad sündmused. Niisuguste katsete uurimine moodustabki tõenäosusteooria aine, ja katseid nimetatakse **juhuslikeks** ehk **stohhastilisteks** ehk **tõenäosuslikeks katseteks**. Edaspidi kasutame lühemat nimetust – katse.

Tähistame sündmusi ladina tähestiku suurtähtedega, varustades neid vajaduse korral indeksitega: $A, B, \ldots\ldots A_1, B_1$ jne.

Juhuslikuks sündmuseks nimetatakse sündmust, mis katse tulemusena võib toimuda või mitte, kusjuures varem ei ole teada, kas sündmus toimub. Juhusliku sündmuse toimumine on küll võimalik, kuid selle reaalne toimumine sõltub mitmetest omavahel seotud või sõltumatutest põhjustest. Teadusharu, mis käsitleb juhuslikke sündmusi, nimetatakse sageli **stohhastikaks.**

Vaatleme näiteid juhuslike sündmuste kohta.

Näide 1. Lihtsamateks juhuslikeks sündmusteks on täringu viskamisel silmade tulek, kaardi võtmine kaardipakist, mündi viskamine, võit loteriil, inimese eluiga jne.

Näide 2. Valime Tartu elanikest juhuslikult suvalise esindaja. Juhuslikuks sündmuseks on meesterahva/naisterahva saamine, aga samuti vanema/noorema kui 35 aastat inimese saamine.

Näide 3. Koosnegu toodang n seadmest. Seadme kvaliteedi kontrollimine olgu aga niisugune, et selle tulemusel seade hävib. Et mitte hävitada kogu toodangut, valitakse m seadet $(m < n)$ ja kontrollitakse neid. Katse tulemuseks saame defektsete seadmete arvu. Siit võib juba teha järeldusi toodangu kohta.

Näide 4. Vaatleme Browni liikumist – vedelikus heljuvate väikeste aineosakeste kaootilist liikumist. Osake muudab oma asendit juhuslikult, põrkudes kokku vedeliku molekulidega. Osakese asend fikseeritud ajamomendil t on aga raskesti prognoositav ja seega juhuslik sündmus. Kui aga aeg t muutub pidevalt, jõuame üldisema mõiste – juhusliku protsessi juurde.

Vaadeldud näidetest tuleneb, et tõenäosusteoorias sündmus on väga üldine ja universaalne mõiste, mille alla mahuvad kõik meid igapäevaelus ümbritsevad nähtused. Tõenäosusteoorias on aga eriline koht kahel sündmusel, mida nimetame edaspidi *kindel sündmus* ja *võimatu sündmus*. **Kindlaks sündmuseks** nimetatakse sündmust, mis katse tulemusena teatud tingimustel alati toimub ja tähistatakse tähega Ω. Näiteks kindlateks sündmusteks on:

visatud kivi kukub alati maa peale tagasi;

iga elusolend hukub temperatuuril 1000 °C.

Võimatu sündmus on sündmus, mis vaadeldava katse tulemusena kunagi ei toimu. Tähistame võimatu sündmuse sümboliga Φ. Näiteks kui võrgus voolu ei ole, siis võimatuks sündmuseks on – elektripirn põleb.

Kui sündmuse A toimumisest järeldub sündmuse B toimumine, siis tähistame

$$A \subset B$$

või

$$A \Rightarrow B.$$

Sel korral nimetatakse sündmust A ka sündmuse B **alamsündmuseks** või **osasündmuseks** ja sündmust B sündmuse A **ülemsündmuseks**. Näiteks, kui sündmus A tähendab, et täringul tuleb 6 silma ja B – täringul tuleb paarisarv silmi, siis ilmselt $A \subset B$.

Ilmselt kehtivad seosed

$$A \subset \Omega, \ \ \Phi \subset A.$$

Esimese seose korral võime öelda ka nii – kuna sündmus Ω toimub alati, siis toimub ta ka juhul, kui on toimunud sündmus A, teise seose korral – kuna sündmus Φ ei toimu kunagi, siis võib öelda, et juhul, kui võimatu sündmus toimub, toimub mis tahes sündmus.

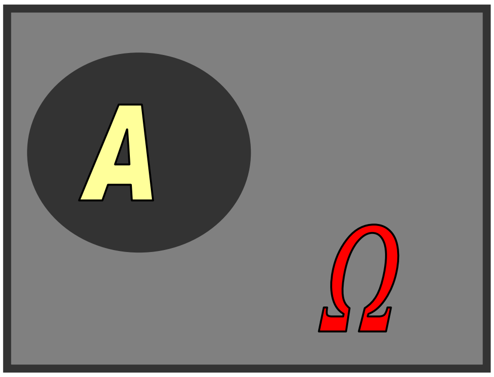
Sündmusi A ja B nimetatakse **võrdseteks** või **samasteks** või **identseteks** (**ekvivalentseteks**), kui kehtivad seosed

$$A \subset B, B \subset A$$

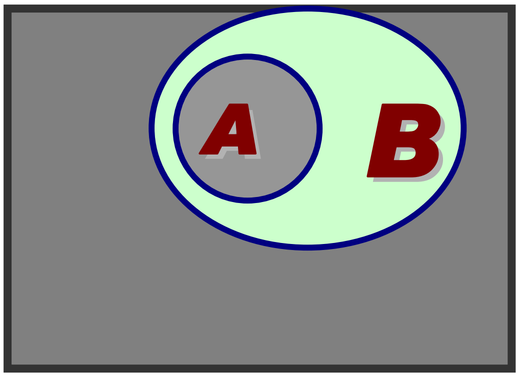
ja sel korral kirjutatakse $A = B$.

Venni diagrammid

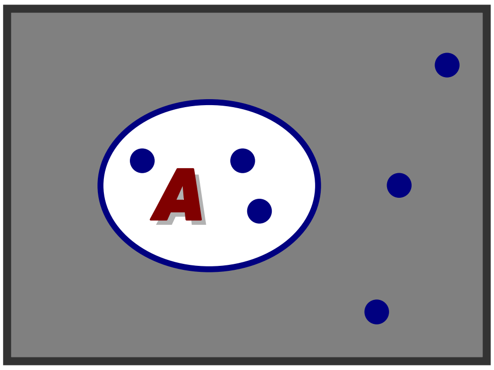
Graafiliselt kujutame edaspidi sündmusi punktihulkadena Venni[1] diagrammi näol. Sündmus A olgu ekvivalentne juhusliku punkti sattumisega piirkonda A (ühe-, kahe- või kolmemõõtmelises ruumis), mis on piirkonna Ω alamhulk. Vaatleme siin kahemõõtmelisi piirkondi. Sel korral kindlat sündmust kujutame ristkülikuna. Kui on tarvis eraldi eristada või rõhutada mingit sündmust, siis viirutame vastava piirkonna. Kindel sündmus Ω ja juhuslik sündmus A on kujutatud joonisel 1.1. Seos $A \subset B$ on kujutatud joonisel 1.2. Sündmuste graafilisel kujutamisel võivad neile vastavad hulgad olla lõplikud või lõpmatud (kontiinuumi võimsusega). Lõplike hulkade korral koosneb piirkond joonisel lõplikust arvust diskreetsetest punktidest (joonis 1.3); võimatu sündmus sel juhul on kujutatud joonisel 1.4.



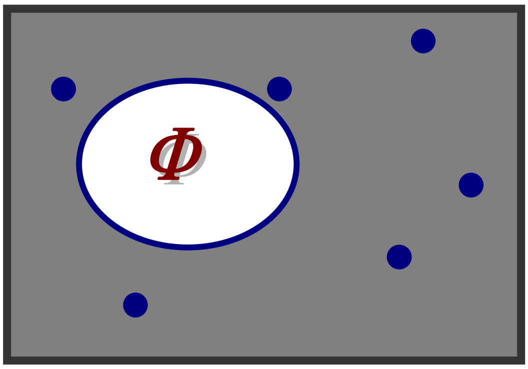
Joonis 1.1



Joonis 1.2



Joonis 1.3



Joonis 1.4

1.2. Tehted sündmustega

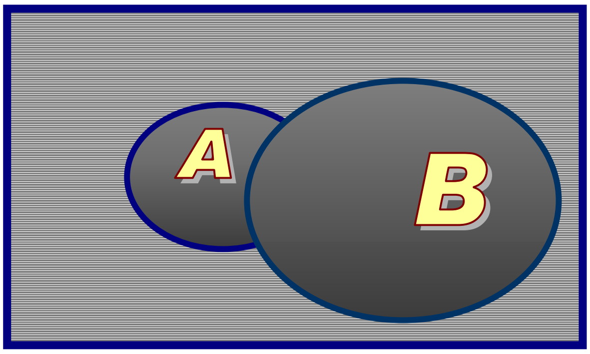
Sündmuste A ja B **summaks** nimetatakse sündmust $A \cup B$ (alternatiivtähis $A+B$), mis seisneb vähemalt ühe sündmuse toimumises sündmustest A ja B. Summa on kujutatud graafiliselt joonisel 2.1 varjutatud alana.

Ilmselt kehtivad seosed

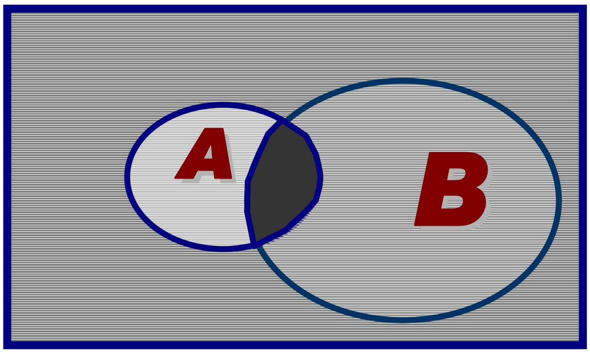
[1] *John Venn (1834–1923), inglise matemaatik ja loogik*

$$A \bigcup A = A, \quad \Omega \bigcup A = \Omega, \quad A \bigcup \Phi = A.$$

Sündmuste A ja B **korrutiseks** nimetatakse sündmust $A \bigcap B$ $(A \cdot B)$, mis seisneb sündmuste A ja B toimumises. Korrutis on kujutatud joonisel 2.2.



Joonis 2.1



Joonis 2.2

Kehtivad seosed

$$A \bigcap A = A, \;\; \Omega \bigcap A = A, \;\; A \bigcap \Phi = \Phi .$$

Definitsioonidest järeldub, et sündmuste summa ja korrutis on kommutatiivsed, s.o

$$A \bigcup B = B \bigcup A,$$
$$A \bigcap B = B \bigcap A.$$

Sündmuste A_k $(k = 1, 2, ..., n)$ **summaks** nimetatakse sündmust

$$\bigcup_{k=1}^{n} A_k \quad , \qquad \left(\text{Alternatiivne tähistus:} \qquad \sum_{k=1}^{n} A_k \right) \quad ,$$

mis seisneb vähemalt ühe sündmuse toimumises nendest sündmustest.

Sündmuste A_k $(k = 1, 2, ..., n)$ **korrutiseks** nimetatakse sündmust

$$\bigcap_{k=1}^{n} A_k \qquad \left(\text{Alternatiivne tähistus:} \qquad \prod_{k=1}^{n} A_k \right) \quad ,$$

mis seisneb kõigi nende sündmuste toimumises.

Märgime, et kasutatakse ka lühendatud tähistusi

$$\bigcup_{k} A_k , \qquad \bigcap_{k} A_k .$$

Kehtivad sündmuste liitmise ja korrutamise assotsiatiivsuse seadus

$$(A \cup B) \cup C = A \cup (B \cup C),$$
$$(A \cap B) \cap C = A \cap (B \cap C)$$

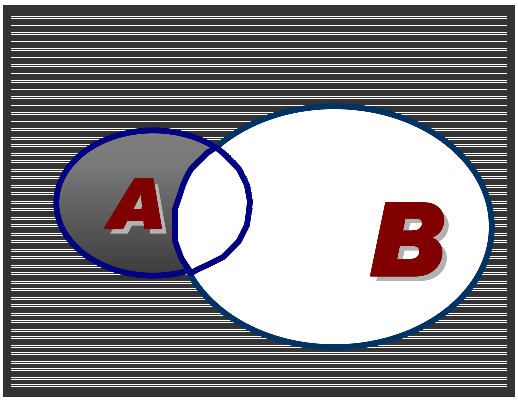
ja distributiivsuse seadus

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

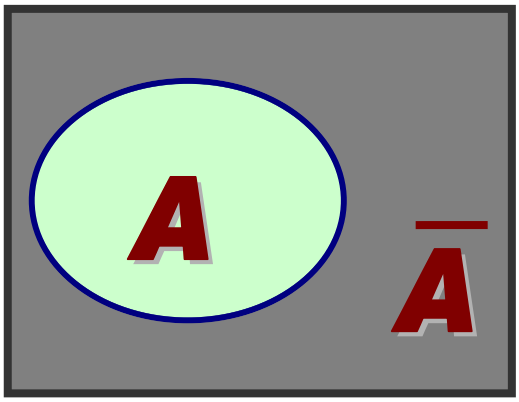
Sündmuste A ja B **vaheks** nimetatakse sündmust

$$A \setminus B,$$

mis seisneb selles, et toimub sündmus A, aga ei toimu sündmust B. Vahe on kujutatud joonisel 2.3.



Joonis 2.3



Joonis 2.4

Kehtivad seosed

$$A \setminus \Omega = \Phi, \;\; A \setminus \Phi = A, \;\; \Phi \setminus A = \Phi.$$

Vahet $\Omega \setminus A$ nimetatakse sündmuse A **vastandsündmuseks** ja tähistatakse

$$\overline{A} = \Omega \setminus A,$$

mis on kujutatud joonisel 2.4. Definitsioonist järeldub, et kui toimub sündmus A, siis ei toimu sündmust $\overline{A}$.

Kehtivad seosed

$$\overline{\Omega} = \Phi, \;\; \overline{\Phi} = \Omega,$$

$$(\overline{\overline{A}}) = A, \;\; A \cup \overline{A} = \Omega, \;\; A \cap \overline{A} = \Phi.$$

Olgu märgitud, et sündmuste vahe võib defineerida ka võrdusega

$$A \setminus B = A \cap \overline{B}.$$

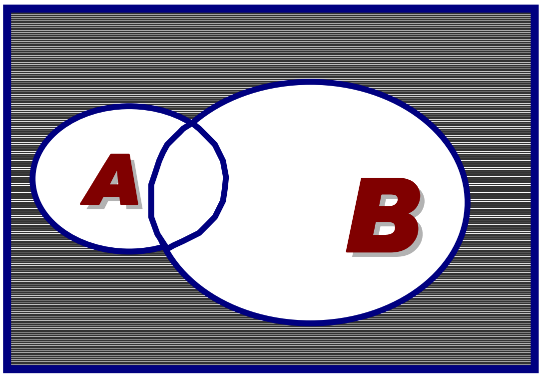
Mitmed tehted avalduvad teiste tehete kaudu, näiteks

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

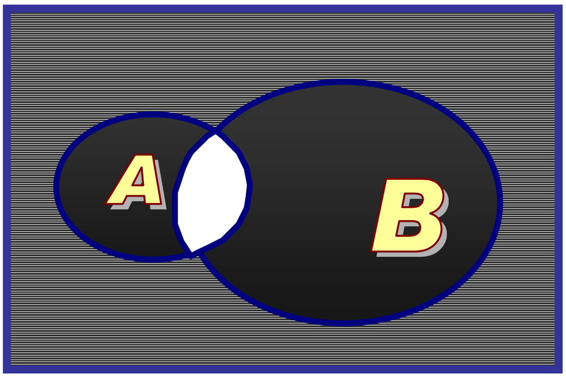
Kasutades vaadeldud tehteid, saab keerukamaid sündmusi (liitsündmusi) avaldada antud sündmuste (lihtsündmuste) kaudu. Ülesannete lahendamisel on sageli otstarbekas lihtsustada valemeid, kus mingid sündmused avalduvad teiste sündmuste kaudu. Seda võib teha vastavalt summa või korrutise kommutatiivsuse, assotsiatiivsuse või distributiivsuse omaduste põhjal. Kasulikud on siinjuures järgmised **duaalsusseosed** ehk **Morgani seadused**:

$$\overline{A\bigcup B}=\overline{A}\bigcap\overline{B},\ \ \overline{A\bigcap B}=\overline{A}\bigcup\overline{B},$$

mis formaalselt tähendavad, et korrutamise ja summa tehtemärgid tuleb ära vahetada, kui sündmused asendada vastandsündmustega. Esimene seadus tähendab, et kumbki sündmustest A ja B ei toimu, teine – ülimalt üks sündmustest A ja B toimub (toimub sündmus A või B). Need sündmused on kujutatud joonistel 2.5 ja 2.6.



Joonis 2.5. Seose $\overline{A\bigcup B}=\overline{A}\bigcap\overline{B}$ tõestus



Joonis 2.6. Seose $\overline{A\bigcap B}=\overline{A}\bigcup\overline{B}$ tõestus

Duaalsusseosed võib üldistada sündmuste hulga jaoks:

$$\overline{\bigcup_k A_k}=\bigcap_k\overline{A}_k,\ \overline{\bigcap_k A_k}=\bigcup_k\overline{A}_k.$$

Märkus: Nagu näha, on tehted sündmustega sarnased hulkade tehetega. See on seletatav ja ka arusaadav selles mõttes, et iga sündmus on seotud teatud hulga võimalike juhtudega. Sündmus toimub, kui juht kuulub teatud juhtude hulka, ega toimu, kui juht ei kuulu sellesse hulka.

Näide 1. Olgu sündmus A – vähemalt üks võit maleturniiril esimeses kahes partiis, A_1 – võit esimeses partiis ja A_2 – võit teises partiis. Nende sündmuste vahel ilmselt kehtib seos

$$A=A_1\bigcup A_2.$$

Võib arutleda aga ka järgmiselt: võib võita esimese partii ja kaotada teise $(A_1\bigcap\overline{A}_2)$ või võita mõlemad partiid $(A_1\bigcap A_2)$ või kaotada esimese ja võita teise $(\overline{A}_1\bigcap A_2)$. Seega

$$A=(A_1\bigcap\overline{A}_2)\bigcup(A_1\bigcap A_2)\bigcup(\overline{A}_1\bigcap A_2).$$

Võib leida ka kolmanda variandi sündmuse A kirjeldamiseks. Paneme tähele, et sündmus A on sündmuse $\overline{A}_1\bigcap\overline{A}_2$ – kaotatakse mõlemas partiis – vastandsündmus. Järelikult

$$A=\overline{\overline{A}_1\bigcap\overline{A}_2}.$$

Näide 2. Kirjutada liitsündmusena järgmised sündmused: 1) toimub ainult üks sündmustest A ja B, 2) ülimalt üks (A või B) sündmustest A ja B, 3) kumbki ei toimu.

Esimene sündmus on

$$(A \cap \overline{B}) \cup (\overline{A} \cap B),$$

teine $\overline{A \cap B}$, mis Morgani teise seaduse põhjal avaldub kujul

$$\overline{A \cap B} = \overline{A} \cup \overline{B}.$$

Kolmas sündmus on $\overline{A \cup B}$, mis Morgani esimese seaduse järgi on

$$\overline{A \cup B} = \overline{A} \cap \overline{B}.$$

Näide 3. Olgu

$$A = (B \cup C) \cap (B \cup \overline{C}) \cap (\overline{B} \cup C).$$

Lihtsustada sündmuse avaldist.

Et suvaliste sündmuste B ja C korral

$$B \cup (B \cap C) = B,$$

siis

$$(B \cup C) \cap (B \cup \overline{C}) = (B \cap B) \cup (B \cap \overline{C}) \cup (C \cap B) \cup (B \cap \overline{C}) =$$
$$= B \cup (C \cap B) \cup (B \cap \overline{C}) \cup \Phi = B$$

ja

$$A = B \cap (\overline{B} \cup C) = (B \cap \overline{B}) \cup (B \cap C) = B \cap C.$$

Näide 4. Analoogiliselt saame

$$\bigcap_{k=1}^{n} (B \cup C_k) = B \cup \bigcap_{k=1}^{n} C_k.$$

1.3. Elementaarsündmuste ruum

Juhuslikke sündmusi A ja B nimetatakse **teineteist välistavateks**, kui ühe sündmuse toimumine välistab teise sündmuse toimumise, s.t (joonis 3.1)

$$A \cap B = \Phi.$$

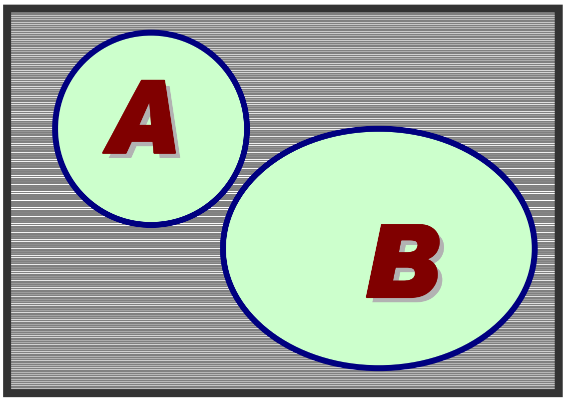
Kaht sündmust A ja B nimetatakse **mittevälistavateks**, kui ühe sündmuse toimumine ei välista teise sündmuse toimumist (joonis 3.2).

Vaatleme sündmuste lõplikku hulka

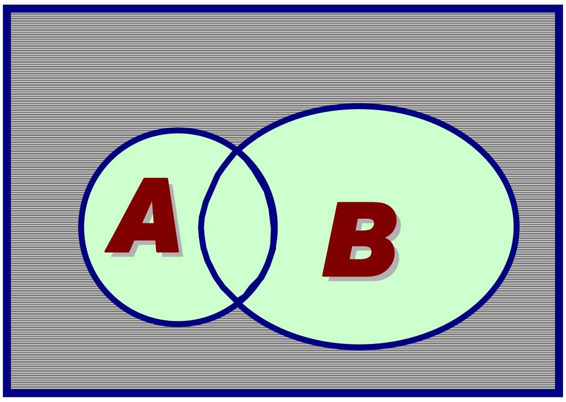
$$\mathbf{A} = \{A_1,\ A_2, ..., A_n\},$$

kus kõik sündmused on erinevad. Hulka $\mathcal{A}$ nimetatakse **üksteist välistavate sündmuste hulgaks,** kui mistahes kaks sellesse hulka kuuluvat sündmust on teineteist välistavad sündmused, s.o (joonis 3.3)

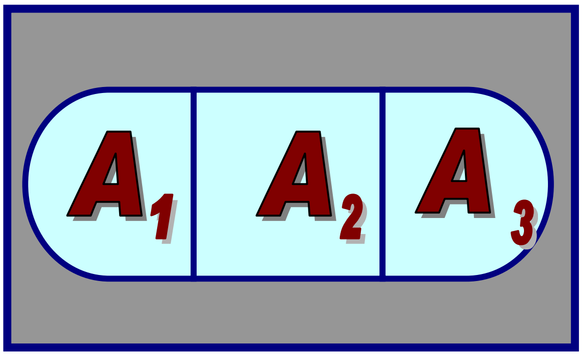
$$A_i \cap A_k = \Phi \quad (i \neq k).$$



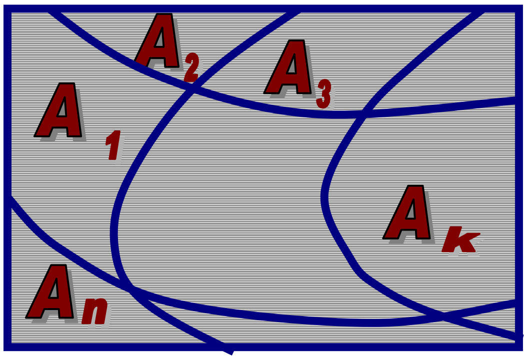
Joonis 3.1



Joonis 3.2



Joonis 3.3



Joonis 3.4

Üksteist välistavate sündmuste hulka $\mathcal{A}$ nimetatakse **sündmuste täissüsteemiks**, kui kehtib võrdus (joonis 3.4)

$$\bigcup_{k=1}^{n} A_k = \Omega .$$

Näiteks täissüsteemiks on hulk $\{A,\ \overline{A}\}$.

Sündmusi nimetatakse **võrdvõimalikeks,** kui ükski neist sündmustest ei ole rohkem eelistatud kui teised. Võrdvõimalikkuse mõiste on üks põhimõisteid ja raskesti kontrollitav. Sageli tuvastatakse võrdvõimalikkust sümmeetria abil (münt, täring, mängukaardid, rulett).

Näide 1. Kui täring on valmistatud homogeensest materjalist ja tal on korrapärase hulktahuka kuju, siis täringu viskamisel ühe või teise silmade arvu tulek on võrdvõimalikud sündmused.

Sündmusi nimetatakse **ainuvõimalikeks**, kui katse tulemusena ühe ja ainult ühe sündmuse toimumine on kindel sündmus.

Ainuvõimalike ja võrdvõimalike sündmuste täissüsteemi nimetatakse **elementaarsündmuste ruumiks**. Sündmusi, mis moodustavad elementaarsündmuste ruumi, nimetatakse **elementaarsündmusteks** ja tähistatakse $E_k (k = 1,2,...,n)$.

Näide 2. Täringu viskamisel sündmused $E_k (k = 1,2,...,6)$ – täringul tuleb k silma, on elementaarsündmused. Sündmuse A – täringul tuleb paarisarv silmi, võime kirjutada kujul

$$A = E_2 \bigcup E_4 \bigcup E_6 .$$

Järeldus. Vaadeldud mõistete põhjal võime öelda, et *sündmus on määratud elementaarsündmuste ruumi mingi alamhulgaga*.

Rakendustes on sageli ülesandeid, kus võimalike tulemuste arv on lõpmatu. Seetõttu üldistame elementaarsündmuse mõistet. Sündmust E nimetatakse **elementaarsündmuseks**, kui ei eksisteeri sündmusi

$$E_1 \neq E, \; E_2 \neq E$$

nii, et kehtiks võrdus

$$E = E_1 \bigcup E_2 .$$

Teisisõnu – sündmust ei saa lahutada "lihtsamate" teineteist välistavate võimatust sündmusest erinevate sündmuste summaks, s.o elementaarsündmus ei sisalda teisi alamsündmusi peale võimatu sündmuse ja iseenda.

Lõpmatut hulka $\{E_1, E_2, ...\}$ nimetatakse **elementaarsündmuste ruumiks**, kui iga kaks sündmust E_k ja E_m on elementaarsündmused, sündmused on võrdvõimalikud ja ainuvõimalikud ning

$$\bigcup_{k=1}^{\infty} E_k = \Omega .$$

Näide 3. Edaspidi on väga oluline niisugune katse, mille korral mõõdetakse mingit suurust X. Elementaarsündmusteks on siin sündmused kujul $X = x$, kus x on mingi fikseeritud väärtus. Seega on loomulik samastada elementaarsündmuste hulk mingi punktihulgaga sirgel. Kui on eelnevalt teada, et suurus X võib omandada väärtusi mingist punktihulgast H, siis seda hulka tulebki vaadelda elementaarsündmuste ruumina. See hulk on üldjuhul aga kontiinumi võimsusega hulk.

Näide 4. Katse all vaatleme ringikujulise märklaua tabamist. Ringi keskpunkt olgu koordinaadisüsteemi alguspunktis ja raadius olgu r. Elementaarsündmusena vaatleme suvalise punkti (x, y) tabamist. Elementaarsündmuste ruum Ω on siin kogu tasand $\mathbf{R}^2$. Märklaua tabamine (sündmus A) tähendab seda, et juhuslik punkt sattub ringi $\{x^2 + y^2 \leq r^2\}$. Ilmselt $A \subset \Omega$.

Märkus 2. Samas katses võib elementaarsündmusi esitada erinevalt. Näiteks märklaua korral juhusliku punkti võib esitada ristkoordinaatides või polaarkoordinaatides punktipaaridena vastavalt (x, y) või (ρ, φ).

Kokkuvõttes oleme jõudnud järgmise mudelini. Vaadeldava katsega seome mingi elementaarsündmuste ruumi Ω nii, et katse tulemusel toimub üks ja ainult üks elementaarsündmus ω, mida interpreteerime hulgana, mis sisaldab ainult ühe elemendi (üheelemendiline hulk). Kindel sündmus on seega kõigi elementaarsündmuste hulk Ω ja võimatu sündmus tühi hulk Φ.

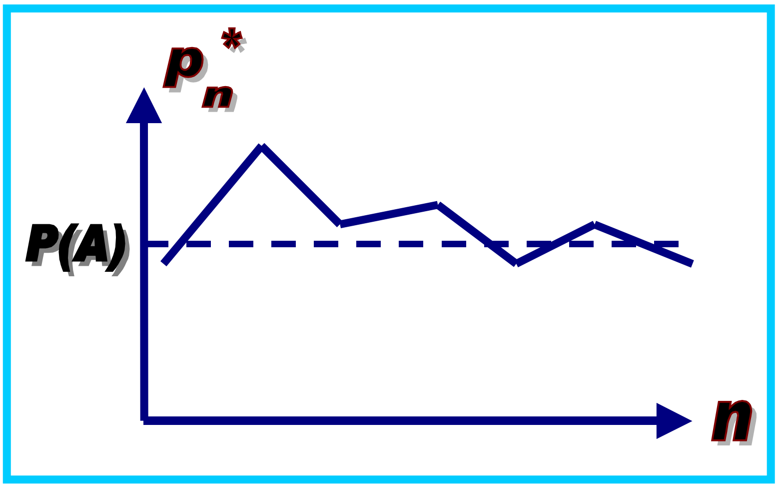
1.4. Sündmuse sagedus

Vaatleme katset, mille tulemusena võib toimuda sündmus A. Eeldame, et katset võib korrata, kusjuures katsetingimused ei muutu ja igal katsel ei sõltu katsetulemus teiste katsete tulemustest. Nimetame niisuguseid katseid **sõltumatuteks katseteks** sündmuse A suhtes. Olgu selles katseseerias katsete arv n ja $m(A)$ – selles seerias sündmuse A toimumiste (esinemiste) arv. Vaadeldavas katseseerias nimetatakse **sündmuse** A **sageduseks** suurust

$$P^*(A) \equiv p_n^* = \frac{m(A)}{n}.$$

Kui katsete arv n ei ole suur, siis sündmuse sagedus on oluliselt juhusliku iseloomuga. Katsete suure arvu n korral on aga sagedused p_n^* erinevates katseseeriates, nagu näitab pikaajaline praktika, ligikaudu võrdsed. S.t sagedusel on tendents **stabiliseeruda** (on **stabiilsuse omadus**). Siin peab märkima, et tõenäosusteooria uurib ainult niisuguseid sündmusi, millel on stabiilsuse omadus. Teisisõnu, eksisteerib niisugune konstant $P(A)$, mida nimetatakse **sündmuse** A **tõenäosuseks**, mille ümber sagedused võnguvad, s.o erinevad vähe sellest konstandist $P(A)$ (joonis 4.1). Lühidalt, kehtib ligikaudne võrdus

$$P(A) \approx \frac{m(A)}{n}.$$



Joonis 4.1

Öeldust järeldub, et vaadeldava sündmusega on võimalik siduda konstantset suurust $P(A)$, mille ümbruses on kõik sagedused ja mis iseloomustab sündmuse ja katsetingimuste vahelist objektiivset seost. Rõhutame, et sageduse leidmiseks ei piisa katsetingimuste teadmisest, vaid on tarvis veel statistilisi andmeid (katsed on vaja korraldada). Niisiis sagedus on sündmuse eksperimentaalne (katseline) karakteristik.

Sageduse põhiomadused

1. Iga sündmuse A korral kehtivad võrratused

$$0 \le P^*(A) \le 1,$$

s.o sündmuse sagedus on mittenegatiivne reaalarv nulli ja ühe vahel.

Tõepoolest, ilmselt kehtivad võrratused

$$0 \le m(A) \le n.$$

Jagades need võrratused suurusega n, saame

$$0 \le \frac{m(A)}{n} \le 1$$

ehk

$$0 \le P^*(A) \le 1.$$

2. Kindla sündmuse sagedus on võrdne ühega, s.t

$$P^*(\Omega) = 1.$$

Sageduse mõiste järgi

$$\frac{m(\Omega)}{n} = \frac{n}{n} = 1,$$

mida oligi vaja tõestada.

3. Võimatu sündmuse sagedus on null, s.t

$$P^*(\Phi) = 0.$$

Et $m(\Phi) = 0$, siis

$$P^*(\Phi) = \frac{0}{n} = 0.$$

4. Sageduste liitmisreegel. Kahe teineteist välistava sündmuse A ja B korral kehtib võrdus

$$P^*(A \cup B) = P^*(A) + P^*(B).$$

Et

$$P^*(A) = \frac{m(A)}{n}, \ P^*(B) = \frac{m(B)}{n}$$

ja

$$m(A \bigcup B) = m(A) + m(B),$$

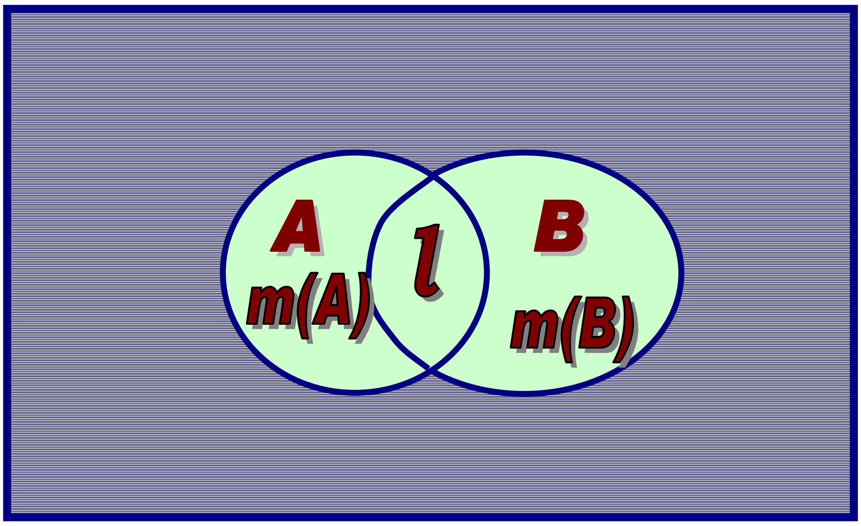
siis

$$P^*(A \bigcup B) = \frac{m(A) + m(B)}{n} = \frac{m(A)}{n} + \frac{m(B)}{n} = P^*(A) + P^*(B).$$

5. Sageduste korrutamisreegel. Kahe juhusliku sündmuse A ja B korral kehtivad võrdused

$$P^*(A \bigcap B) = P^*(A) P^*(B \mid A),$$
$$P^*(A \bigcap B) = P^*(B) P^*(A \mid B),$$

kus $P^*(A \mid B)$ on sündmuse A tinglik sagedus sündmuse B suhtes. ***Ühe sündmuse A sagedust, mis on leitud tingimusel, et teine sündmus B toimus, nimetatakse <u>sündmuse A tinglikuks sageduseks sündmuse B suhtes</u> ja tähistatakse $P^*(A \mid B)$.***



Joonis 4.2

Olgu (joonis 4.2)

$$P^*(A) = \frac{m(A)}{n}, \quad P^*(B) = \frac{m(B)}{n}, \quad P^*(A \bigcap B) = \frac{l}{n}.$$

Tingliku sageduse mõiste järgi

$$P^*(B \mid A) = \frac{l}{m(A)}, \quad P^*(A \mid B) = \frac{l}{m(B)},$$

järelikult

$$P^*(A \bigcap B) = \frac{l}{n} = \frac{m(A)}{n} \cdot \frac{l}{m(A)} = P^*(A) P^*(B \mid A).$$

Analoogiliselt

$$P^*(A \cap B) = \frac{m(B)}{n} \cdot \frac{l}{m(B)} = P^*(B)P^*(A|B).$$

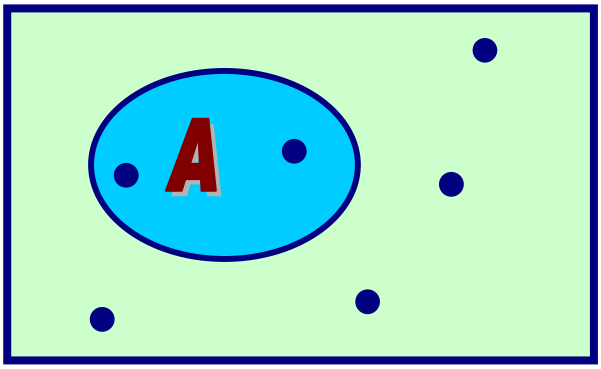
1.5. Sündmuse tõenäosus

Sündmuse tõenäosus on tõenäosusteooria olulisemaid mõisteid. Tõenäosus on sündmuse objektiivse võimalikkuse arvmõõt. See iseloomustab objektiivset põhjuslikkust teatud tingimuste ja sündmuste vahel.

Olgu $\boldsymbol{E} = \{E_1, E_2, ..., E_n\}$ lõplik elementaarsündmuste ruum. Sündmus A avaldugu m elementaarsündmuse kaudu, s.o

$$A = \bigcup_{k=1}^{m} E_{i_k} \quad (i_k \in \{1, 2, ..., n\}).$$

Viimane tähendab seda, et sündmus A toimub, kui toimub üks elementaarsündmustest E_{i_k} (joonis 5.1).



Joonis 5.1

Rõhutame, et tõenäosusteoorias vaadeldakse ainult niisuguseid sündmusi, mis avalduvad elementaarsündmuste summana.

Sündmuse A **tõenäosuseks** nimetatakse arvu m ja elementaarsündmuste arvu n jagatist ja tähistatakse [2]

$$P(A) \equiv p(A) \equiv p = \frac{m}{n}. \tag{5.1}$$

See on **tõenäosuse** nn **klassikaline definitsioon**.

---

[2] tähistus tuleneb nii ladinakeelsest sõnast *probābilitās* kui ka ingliskeelsest *probability*

Sageli sõnastatakse see definitsioon teistsuguses terminoloogias. Moodustagu sündmused $E_k\,(k = 1,\ 2,\ ...,\ n)$ sündmuste täissüsteemi. Üksikuid sündmusi E_k nimetatakse **juhtudeks**. Juhtu nimetatakse **soodsaks** sündmuse toimumiseks, kui selle juhu realiseerumisel toimub sündmus A. Sündmuse A tõenäosuseks nimetatakse selle sündmuse kõigi soodsate juhtude arvu m ja kõigi juhtude arvu n suhet, s.t

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

Klassikalisest definitsioonist (5.1) järeldub (nii nagu sündmuse sageduse korral), et kehtivad võrratused

$$0 \le P(A) \le 1,$$

kusjuures

$$P(\Omega) = 1,\ \ P(\Phi) = 0.$$

Tõepoolest, kui $A = \Omega$, siis $m = n$ ja

$$P(\Omega) = \frac{m}{n} = 1.$$

Kui $A = \Phi$, siis

$$P(\Phi) = \frac{0}{n} = 0.$$

Suvalise sündmuse A korral $0 < m < n$ ehk

$$0 < \frac{m}{n} < 1,$$

s.o

$$0 < P(A) < 1.$$

Kokkuvõttes

$$0 \le P(A) \le 1.$$

Seega on võetud sündmuse toimumise võimalikkuse mõõtühikuks kindla sündmuse tõenäosus. Mida lähemal arvule 1 on sündmuse tõenäosus, seda võimalikum on sündmuse toimumine.

Vaatleme näiteid tõenäosuse klassikalise definitsiooni rakenduse kohta.

Näide 1. Leida tõenäosus, et täringu viskamisel silmade arv on: 1) kuus, 2) paarisarv. Olgu sündmus A – täringul on 6 silma, B – täringul on paarisarv silmi. Seega

$$P(A) = \frac{1}{6},\ \ P(B) = \frac{3}{6} = \frac{1}{2}.$$

Näide 2. Valides telefoninumbrit oli abonent unustanud kaks viimast numbrit ja teades, et need on erinevad, valis neid seetõttu juhuslikult. Leida tõenäosus, et valitakse õige telefoninumber.

Elementaarsündmuste arv (kõigi juhtude arv) on

$$n = V_{10}^{2} = 10 \cdot 9 = 90,$$

soodsate juhtude arv on 1. Järelikult

$$p = \frac{1}{90} = 0{,}011(1).$$

Kui aga kaks viimast numbrit võivad olla ühesugused, siis

$$n = W_{10}^{2} = 10^{2} = 100$$

ja tõenäosus

$$p = \frac{1}{100} = 0{,}01.$$

Tuletame meelde: <u>Variatsioonid</u>.

Variatsioonid *n elemendist m kaupa V_n^m on arv, mis näitab, mitmel erineval viisil saab n elemendist valida m erinevat elementi, kusjuures elementide järjestus valimis on oluline. Variatsioonide arvutusvalem on*

$$V_n^m = n(n-1)(n-1)\cdot \ldots\ldots\ldots \cdot (n-m+1).$$

Variatsioonid kordumistega *n elemendist m kaupa W_n^m on arv, mis näitab, mitmel erineval viisil saab n elemendist valida m elementi, kusjuures elementide järjestus valimis on oluline, kuid (erinevalt variatsioonidest) üht ja sama elementi võib valida korduvalt (kuni m korda). Kordumistega variatsioonide arvutusvalem on*

$$W_n^m = n^m.$$

Näide 3. Seitsmele kaardile on kirjutatud tähed

$$a, i, l, l, n, n, t.$$

Leida tõenäosus, et juhuslikult järjestatud kaartidel tuleb sõna "tallinn".

Kõigi juhtude arv on $7!$. Soodsate juhtude arv on $2!\cdot 2!$, sest *l*-isid ja *n*-isid võib omavahel vahetada. Seega

$$p = \frac{2!\cdot 2!}{7!} = \frac{1}{1260} \approx 0{,}000794 \quad .$$

Tuletame meelde: <u>Permutatsioonid</u>.

Permutatsioonid n *elemendist tähistatakse* P_n *(mitte segi ajada tõenäosusega!) või* $n!$. *Permutatsioonid on arv, mis näitab, mitmel eri viisil saab järjestada* n *elementi. Permutatsioonide arvutusvalem on*

$$P_n \equiv n! = 1 \cdot 2 \cdot 3 \cdot .... \cdot (n-1) \cdot n.$$

Tuletame meelde: <u>Kombinatsioonid.</u>

Kombinatsioonid n *elemendist* m *kaupa* $C_n^m \equiv \binom{n}{m}$ *on arv, mis näitab, mitmel eri viisil võib valida* n *elemendist* m *elementi, kusjuures **elementide järjestus valimis ei ole oluline.** Kombinatsioonide arvutusvalem on*

$$C_n^m \equiv \binom{n}{m} = \frac{V_n^m}{P_m} = \frac{n(n-1) \cdot .... \cdot (n-m+1)}{m!} = \frac{n!}{m! \cdot (n-m)!}.$$

Näide 4. Urnis on 8 valget ja 4 musta kuuli. Juhuslikult valitakse 2 kuuli. Leida tõenäosus, et mõlemad kuulid on valged.

Kõikide juhtude arv

$$n = C_{12}^2 = \binom{12}{2},$$

soodsate juhtude arv

$$m = C_8^2 = \binom{8}{2}$$

ja

$$p = \frac{8!\ 2!\ 10!}{2!\ 6!\ 12!} = \frac{7 \cdot 8}{11 \cdot 12} = \frac{14}{33} = 0{,}424(24) \quad .$$

Märkus. Ka näited 1–3 võime sõnastada nii, et on tegemist kuuli võtmisega urnist. Üldiselt kõiki ülesandeid, kus eeldatakse sündmuste võrdvõimalikkust, võib taandada nn **urniskeemile**.

Näide 5. *Eesti Loto "5 32-st"*. Leida tõenäosus, et ühel täidetud piletil on märgitud õigesti: 1) 3 numbrit, 2) 4 numbrit, 3) kõik 5 numbrit.

Ilmselt saame ülesande taandada urniskeemile, s.t urnile, milles on 5 valget (võidunumbrit) ja 27 musta kuuli (ei ole võidunumbrid). Kõikide erinevate juhtude arv võtta urnist 5 kuuli on

$$n = \binom{32}{5} = 201376,$$

soodsate juhtude arvud on vastavalt

$$m_3 = \binom{5}{3}\binom{27}{2} = 3510,\ \ m_4 = \binom{5}{4}\binom{27}{1} = 135,\ m_5 = 1$$

ja seega on vastavad tõenäosused

$$p_3 = \frac{3510}{201376} \approx 0{,}01743, \;\; p_4 = \frac{135}{201376} \approx 0{,}00067,$$

$$p_5 = \frac{1}{201376} \approx 0{,}000005 \quad .$$

(ehk 1,7 korral sajast, 6,7 korral kümnest tuhandest ja 5 korral miljonist katsest).

Näide 6. (Analoogiline eelnevaga). Viking Lotol on 48 numbrit (kuuli) ja võidud on 3, 4, 5 ja 6 õiget numbrit (6 on ‚Jackpot'). Vastavad võidutõenäosused tulevad

$$p_3 = 0{,}02313 \approx 1/43, p_4 = 0{,}00116 \approx 1/860, p_5 = 0{,}000021 \approx 1/50000,$$

$$p_6 = 1/(12{,}3 \cdot 10^6).$$

Näide 7. Toodangupartii koosneb n seadmest, kusjuures n_1 on defektsed. Juhuslikult valitakse kontrolliks $m(m < n)$ seade. Leida tõenäosus, et m seadme hulgas on täpselt $k\,(k < m)$ defektset.

Kõikide juhtude arv on ilmselt erinevate kombinatsioonide arv n elemendist m kaupa

$$\binom{n}{m}.$$

Soodsate juhtude arv aga

$$\binom{n_1}{k}\binom{n-n_1}{m-k},$$

sest k seadet peavad olema defektsed ja seega $m-k$ mittedefektsed; mittedefektseid on aga üldse $n-n_1$. Järelikult nõutud tõenäosus on

$$p = \frac{\binom{n_1}{k}\binom{n-n_1}{m-k}}{\binom{n}{m}}.$$

Märkus. Sageli on tegemist sündmustega, millede tõenäosus on väga väike (lähedane nullile). Siit ei järeldu, et üksikul katsel sündmus ei toimu. Pikaajaline praktika aga näitab, et vähe tõenäoline sündmus üksikul katsel enamikel juhtudel ei toimu. Näiteks võit loteriil! Siit tuleneb nn väikese tõenäosusega **sündmuse praktilise mittetoimumise printsiip**: *kui sündmuse tõenäosus on väga väike, siis praktiliselt üksikul katsel seda sündmust ei toimu*. Loomulikult tekib küsimus – kui väike peab olema sündmuse tõenäosus, et lugeda see sündmus võimatuks üksikul katsel? Ühest vastust siin anda ei saa. Erineva sisuga sündmuste korral on vastused erinevad ega ole seega seotud matemaatilise teooriaga. Kõik sõltub katse tulemuse tähtsusest ja tagajärgedest. Mida ohtlikum on võimalik katsetulemus, seda lähedasem nullile peab olema vastava sündmuse tõenäosus, et lugeda seda praktiliselt võimatuks. Näiteks on lubamatu, et kosmoselaeva langevarju mitteavanemise tõenäosus on 0.01. Samal ajal aga tõenäosus 0.01, et näiteks Tallinn-Tartu kiirrong saabub hilinemisega, on väga hea näitaja. Praktiliselt saabub rong õigel ajal, hilinedes keskmiselt ühel korral 3.3 kuu jooksul.

Kasutame edaspidi ka nimetust **praktiliselt võimatu sündmus**, kui

$$p(A) \approx 0.$$

Analoogiliselt nimetatakse sündmust **praktiliselt kindlaks sündmuseks**, kui

$$p(A) \approx 1.$$

Praktiliselt kindlatel ja võimatutel sündmustel on eriline koht tõenäosusteoorias – neile on rajatud kogu igapäevane tunnetuslik väärtus. Näiteks lennates lennukiga või sõites laevaga (autoga), ei arvesta me võimaliku õnnetusega, kuigi niisuguse sündmuse tõenäosus erineb nullist (kuid on väga väike).

Juhime veel tähelepanu sündmuse praktilise mittetoimumise printsiibi sõnastuses väljendile *üksikul katsel*. Asi on selles, et korrates katseid, suurendame tõenäosust, et sündmus toimuks kas või üks kord selles katseseerias. Vaatleme loteriid, mis koosneb 1 miljonist piletist, kusjuures ainult üks on võidupilet. Ühe pileti omanikul on võidu tõenäosus 0.000001. Niisiis võit on praktiliselt võimatu sündmus. Olgu müüdud kõik piletid. Keegi ostjatest võitis! Seega toimub praktiliselt võimatu sündmus. Mille arvel? Selle arvel, et katset (pileti ostmist) korrati väga palju kordi.

1.6. Geomeetriline tõenäosus

Tõenäosuse klassikaline definitsioon nõuab, et elementaarsündmuste ruum on lõplikumõõtmeline. Rakendustes esinevad aga sageli ülesanded, kus elementaarsündmuste ruum on lõpmatumõõtmeline. Mõnikord saab niisugustel juhtudel kasutada tõenäosuse leidmiseks niisuguseid meetodeid, kus oluline osa on ikkagi sündmuste võrdvõimalikkuse mõistel. Kasutada saab seda ülesannetes, mis taanduvad juhusliku punkti sattumisele (viskamisele) lõiku, lõplikku tasandilisse või ruumilisse piirkonda. Siit pärinebki meetodi nimetus – *geomeetriline tõenäosus.* Vaatleme tasandilist juhtu. Olgu D ja D_1 tasandilised piirkonnad nii, et $D_1 \subset D$ ja nende vastavad pindalad on $A(D)$ ja $A(D_1)$ (joonis 6.1). Vaatleme punkti juhuslikku sattumist piirkonda D ja leiame tõenäosuse, et punkt sattub piirkonda D_1. Siinjuures eeldame, et juhuslik punkt võib sattuda piirkonna D mis tahes punkti ja tõenäosus sattuda piirkonda D mingisse osapiirkonda on võrdeline selle osapiirkonna pindalaga ning ei sõltu osapiirkonna asendist ega kujust. Teisiti öeldes, me seame elementaarsündmuste ruumile Ω vastavaks piirkonna D ja juhuslikule sündmusele B vastavaks piirkonna D_1. Tõenäosus, et juhuslik punkt sattub piirkonda D_1 (sündmus B), arvestades eeldust, on seega

$$P(B) = cA(D_1).$$

Kordaja c määramiseks oletame, et piirkond D_1 ühtib piirkonnaga D. Siis on sündmus B samaväärne sündmusega Ω (s.t $B = \Omega$). Järelikult

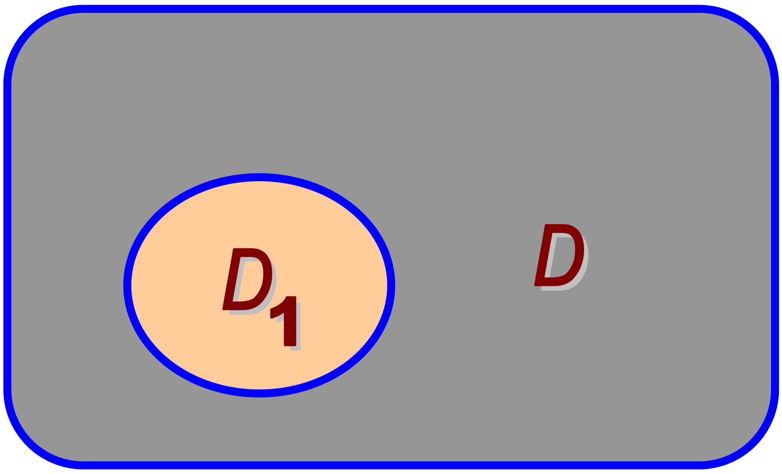
$$P(\Omega) = cA(D) = 1,$$

millest

$$c = \frac{1}{A(D)}.$$

Seega oleme saanud sündmuse B tõenäosuse, mida nimetatakse **geomeetriliseks tõenäosuseks**

$$P(B) = \frac{A(D_1)}{A(D)}. \tag{6.1}$$



Joonis 6.1

Teisisõnu, tõenäosus $P(B)$ on võrdne piirkonna D_1 (punktide soodsate asendite hulk) ja kogu piirkonna D (punktide kõikvõimalike asendite hulk) pindalade suhtega. (Tegelikult on valem (6.1) intuitiivselt kergesti adutav.)

Üldjuhul, kui punkti juhusliku sattumise võimalikkus mingisse mitmemõõtmelisse piirkonda (erijuhul sirgel, tasandil või ruumis) ei sõltu selle piirkonna asendist ja raja kujust, vaid sõltub üksnes piirkonna mõõdust A (vastavalt pikkusest, pindalast või ruumalast), siis on mingisse piirkonda D_1 juhusliku punkti sattumise tõenäosus võrdne selle piirkonna mõõduga. Nii et valem (6.1) kehib suvalise dimensiooniga geomeetrias selle parandusega, et A on piirkonna „ruumala" vastavas ruumis.

Näide 1. Leida tõenäosus, et ruutvõrrandi

$$x^2 + 2ax + b = 0$$

lahendid on reaalsed, kui kordajate a ja b väärtused on võrdvõimalikud ja rahuldavad tingimusi

$$|a| \le 1, \ |b| \le 1.$$

Millise tõenäosusega on võrrandil positiivsed lahendid?

Võrrandi lahendid on

$$x = -a \pm \sqrt{a^2 - b}.$$

Need lahendid on reaalsed, kui

$$a^2 - b \geq 0. \tag{6.2}$$

Kõigi elementaarsündmuste (a, b) hulk D moodustab ruudu $-1 \leq a \leq 1, \ -1 \leq b \leq 1$ joonisel 6.2. Soodsate elementaarsündmuste hulk D_1, mis on määratud võrratusega (6.3), on joonisel 6.2 toonitud halliks. Ilmselt

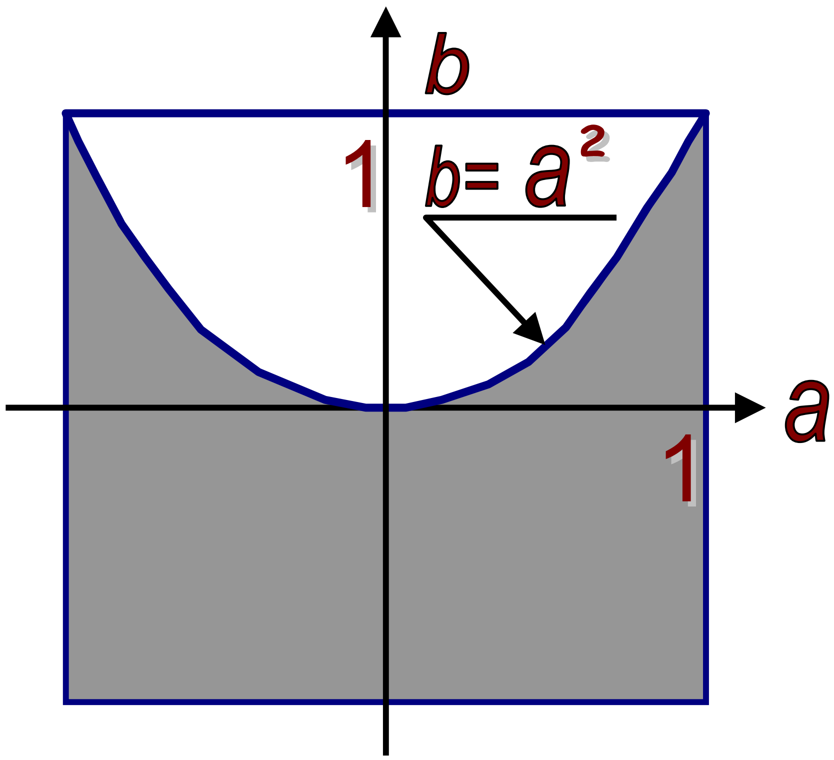
$$A(D) = 4$$

ja

$$A(D_1) = 2 + 2\int_0^1 a^2 da = 2 + \frac{2}{3} = \frac{8}{3}$$

ning tõenäosus, et lahendid on reaalsed, avaldub

$$p_1 = \frac{8}{3 \cdot 4} = \frac{2}{3} = 0{,}6(6).$$



Joonis 6.2

Ruutvõrrandil on positiivsed lahendid, kui $a < 0$ ja $b > 0$. Nüüd on

$$A(D_1) = \int_{-1}^{0} a^2 da = \frac{a^3}{3}\bigg|_{-1}^{0} = \frac{1}{3}$$

ja

$$p_2 = \frac{1}{12} = 0{,}083(3).$$

Näide 2. Kaks tuttavat otsustavad kohtuda teatud kohas ajavahemikul $[T, T+t]$. Saabuja ootab $t/2$ min. Kui teist tuttavat selle aja jooksul ei tulnud või oli enne lahkunud, siis kohtumist ei toimu. Leida tõenäosus, et kohtumine toimub.

Olgu tuttavate saabumise ajamomendid x ja y. Kohtumine toimub, kui

$$|x - y| \le \frac{t}{2}.$$

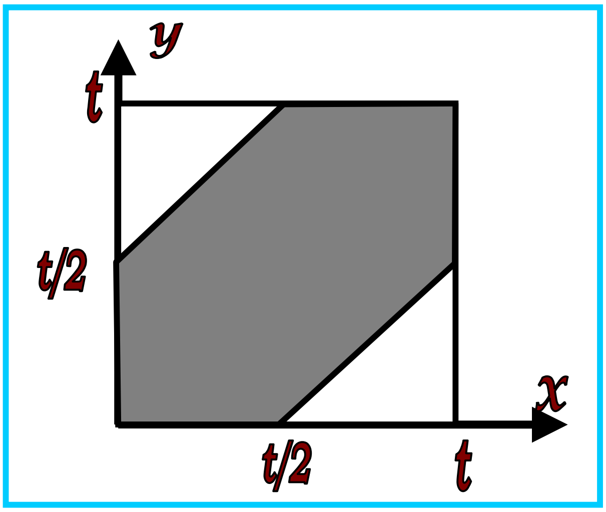
Arvutuste ja joonise lihtsustamiseks võime ilmselt võtta $T = 0$. Elementaarsündmuste $(x,\ y)$ hulk D on ruut $[0,\ t] \otimes [0,\ t]$. Kohtumistele vastav elementaarsündmuste hulk D_1 (rahuldab võrratust (6.4)) on joonisel 6.3 tähistatud halli värviga. Nüüd

$$A(D) = t^2,$$

$$A(D_1) = t^2 - 2\frac{1}{2}\left(\frac{t}{2}\right)^2 = t^2 - \frac{t^2}{4} = \frac{3}{4}t^2$$

ja järelikult

$$p = \frac{3t^2}{4 \cdot t^2} = 0{,}75.$$



Joonis 6.3

Näide 4. Kuubis $R: [-a < x < a], [-a < y < a], [-a < z < a]$ tekitatakse n juhuslikku punkti (X, Y, Z). Leida tõenäosus, et lähima genereeritud punkti kaugus kuubi keskpunktist on suurem kui $r\,(r < a)$.

Olgu lähima punkti kaugus r, s.t keras K, mille keskpunkt langeb kokku kuubi keskpunktiga, ei ole ühtegi genereeritud punkti (joonis 6.5). Ilmselt tõenäosus, et üks juhuslikult genereeritud punkt satub sellesse kerasse, on võrdeline kera ruumalaga, s.o tõenäosus ***punkti saamiseks*** sellesse kerasse on

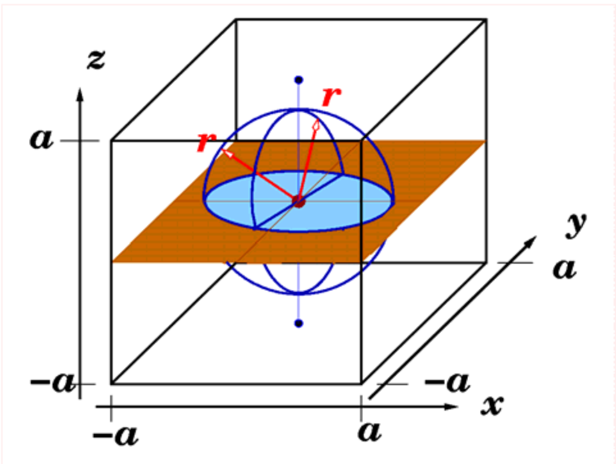
$$p = \frac{V(K)}{V(R)} = \frac{\frac{4}{3}\pi r^3}{8a^3} = \frac{\pi}{6}\left(\frac{r}{a}\right)^3.$$

Vastavalt tõenäosus, et see punkt ***ei satu kerasse*** K on selle sündmuse vastandsündmus tõenäosusega

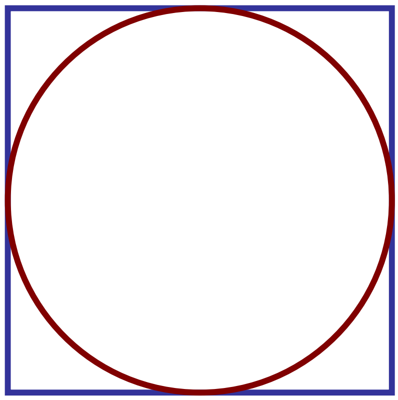
$$p' = 1 - p = 1 - \frac{\pi}{6}\left(\frac{r}{a}\right)^3.$$

Järelikult avaldub tõenäosus, et kõik n genereeritud punkti paiknevad väljaspool kera K, valemiga

$$P = (p')^n = \left(1 - \frac{\pi}{6}\left(\frac{r}{a}\right)^3\right)^n.$$



Joonis 6.4



Joonis 6.5

Näide 4. Kui suur on tõenäosus, et maapinnal paiknevasse ühikruutu sattuv vihmapiisk satub ühtlasi ka sellesse ruutu paigutatud puutujaringjoone sisse (joonis 6.5)?

Vastav tõenäosus on ühikulise diameetriga ringi pindala $S_O = \pi / 4$ suhe ühikruudu pindalasse $S_R = 1$, seega

$$P = S_O / S_R = \pi / 4 \approx 0{,}785.$$

DEMO: π väärtuse arvutamine vihmasajus.

1.7. Märkusi tõenäosuse mõiste kohta

1. Kui klassikaline tõenäosus ei ole rakendatav, siis võib määrata sündmuse sageduse. Selleks on aga vaja korraldada katseseeria. Kui katseseerias on n katset ja sündmus toimub m_n korda, siis sündmuse sagedus on

$$p_n^* = \frac{m_n}{n}.$$

Vastandades tõenäosuse ja sageduse mõisted, näeme, et sündmuse tõenäosuse määramiseks ei pea korraldama katseseeriat (eksperimenti), sageduse leidmiseks on see aga eeltingimus. Teisisõnu – tõenäosus leitakse enne katset (*a priori*), sagedus – peale katset (*a posteriori*).

2. Uurimused ja praktika on näidanud (mida juba punktis 1.4 märkisime), et küllalt suure katsete arvu korral on sündmuste sagedusel p_n^* **stabiilsuse omadus,** mis seisneb selles, et erinevates katseseeriates sündmuse sagedus muutub vähe (seda vähem, mida suurem on katsete arv) ja kõigub mingi konstandi ümber, mida nimetatakse sündmuse tõenäosuseks.

Öeldust tuleneb, et kui eksperimentaalselt on määratud sagedus, siis võib selle võtta sündmuse tõenäosuse ligikaudseks väärtuseks.

3. Kui on olemas piirväärtus

$$\lim_{n\to\infty} \frac{m_n}{n} = p^*,$$

siis nimetatakse seda **sündmuse statistiliseks tõenäosuseks** (R.E. Mises, 1918).

Märgime aga, et põhimõtteliselt ei ole võimalik kindlaks määrata Misese mõttes defineeritud statistilist tõenäosust, kuna ei saa korraldada lõpmatut katseseeriat. Samal ajal võib aga iga $\varepsilon > 0$ korral määrata katsete arvu $n(\varepsilon)$ nii, et sündmuse sagedus p_n^* erineb piirväärtusest (statistilisest tõenäosusest) $p*$ absoluutväärtuse poolest vähem kui ε, s.o

$$\left| p_n^* - p* \right| < \varepsilon.$$

Praktikas (ja ka käesolevas konspektis) kasutatakse statistilist tõenäosust harilikult selles mõttes, et selle väärtuseks võetakse sagedus p_n^* fikseeritud n korral või viimasest ümardamisega saadud väärtus.

4. Paljudel juhtudel on teoreetiliste kaalutlustega võimatu leida sündmuse tõenäosust. Nii näiteks ei osata teoreetiliselt leida tõenäosust, et sündiv laps on näiteks poiss või leida tõenäosust, et inimene sureb teatud haigusesse.

Näide. Võrreldes andmeid tüdrukute ja poiste sündide kohta kogu maailmas, on saadud sagedused

$$p_t^* = 0.486, \;\; p_n^* = 0{,}514 \quad ,$$

mis on tuntud demograafias kui *tüdrukute ja poiste sündide tõenäosused*. Asjaolu, et poisse sünnib rohkem kui tüdrukuid, märkis esmakordselt J. Arbuthnot[3] 1710. aastal, uurides 82 aasta jooksul (1629–1710) toimunud sünde Londonis. Aastal 1812 leidis P.S. Laplace, et poiste ja tüdrukute sündide suhe on 22:21, mis langeb kokku A. Humboldti andmetega Lõuna-Ameerika troopiliste alade kohta. See annab sageduse $p_p^* = 0{,}5116$.

Esitame veel tabeli aastatel 1990–1998 Eestis sündinud poisslaste sageduste kohta.

1990	1991	1992	1993	1994	1995	1996	1997	1998
0,5152	0,5141	0,5130	0,5199	0,5100	0,5141	0,5137	0,5188	0,5165

Kokkuvõtteks võib öelda, et sündmuse tõenäosuse võib määrata eksperimentaalselt. Selleks on vaja korraldada katseseeria ja leida sündmuse sagedus. Seega võib väita, et (mõnedel kitsendustel) on igal sündmusel olemas tõenäosus.

[3] *John Arbuthnot (1667–1735), šoti arst ja matemaatik*

2. JUHUSLIKUD SUURUSED

2.1. Juhusliku suuruse mõiste

Teame, et loodusnähtused on nii kvalitatiivselt kui ka kvantitatiivselt määratletud. Nähtuse ja katse kvalitatiivne tulemus kujutab endast sündmust. Täppisteaduste eesmärgiks on saada kvantitatiivseid karakteristikuid. Uurijat huvitab mitte ainult teatud sündmuse toimumine, vaid ka mõnede arvsuuruste väärtused, mis on seotud vaadeldava sündmusega. Sageli tehaksegi mingi sündmuse toimumine kindlaks sellega, et mingi füüsikaline suurus omandas katse (uurimuse) käigus teatud väärtuse. Niisugust, juhusliku sündmusega seotud suurust, mis katse tulemusena omandab mingi, varem mitteteadaoleva väärtuse, nimetatakse *juhuslikuks suuruseks*. Niisiis juhuslik suurus on seotud sündmusega ja juhusliku suuruse mõiste üldistab juhusliku sündmuse mõistet. Täpsemini, **juhuslikuks suuruseks** nimetatakse suurust, mis katse tulemusel omandab ühe või teise, varem mitteteadaoleva väärtuse mingist väärtuste hulgast. Viimast hulka nimetatakse juhusliku suuruse **võimalike väärtuste hulgaks**.

Juhuslikuks suuruseks on näiteks:

täringu viskamisel saadud silmade arv,

telefonikeskjaama saabuvate kutsungite arv,

spordivõistlustel kaugushüppe võitja tulemus,

elektripirni põlemiskestus,

gaasimolekuli kiirus,

välisõhu temperatuur (mingil kindlal ajahetkel) Tartus Raekoja Platsil.

Siin kahel esimesel suurusel on naturaalarvulised, kolmel järgneval – reaalarvulised positiivsed võimalikud väärtused, viimasel – nii positiivsed kui negatiivsed (kui mõõdame Celsiuse kraadides).

Tähistame juhuslikke suurusi ladina tähestiku suurtähtedega $X, Y, \ldots;$ nende võimalikke väärtusi aga vastavate väiketähtedega $x, y, \ldots$. Kui katse tulemusel juhuslik suurus X omandab väärtuse $X = x$, siis öeldakse ka, et $\boldsymbol{x}$ **on juhusliku suuruse** $\boldsymbol{X}$ **realisatsioon.** Üldiselt nimetatakse **juhuslikuks suuruseks** suurust X, kui iga reaalväärtuselise x korral vahemikus

$$-\infty < x < \infty$$

(seda tähistatakse ka nii: $x \in \mathbf{R}$, s.t „x kuulub reaalteljele“)

on määratud või antud tõenäosus leida juhuslik suurus punktist x vasemal $P(X < x)$.

Liigitame juhuslikke suurusi diskreetseteks ja pidevateks suurusteks (võimalikud on ka segatüüpi juhuslikud suurused, aga neid selles kursuses ei vaadelda).

Juhuslikku suurust nimetatakse **diskreetseks**, kui selle võimalikeks väärtusteks on üksikud, diskreetsed või isoleeritud väärtused. Võimalike väärtuste hulk on siin lõplik või loenduv hulk.

Pideva juhusliku suuruse all mõistame niisugust suurust, mille võimalike väärtuste hulk on reaaltelg **R** või selle mingi alamhulk, näiteks lõik $[a,b]$. Pideva juhusliku suuruse mõistet täpsustame hiljem.

Juhuslikke suurusi iseloomustatakse jaotusseadustega. Öeldakse, et **jaotusseadus** ehk **jaotus** ehk **tõenäosusjaotus** on antud, kui on määratud: **1)** juhusliku suuruse võimalike väärtuste hulk, **2)** eeskiri, mis võimaldab leida tõenäosust, et juhusliku suuruse väärtus on mingis piirkonnas. Jaotusseaduse enamlevinumateks ja praktilisteks vormideks on jaotustabel, jaotustihedus ja integraalne jaotusfunktsioon.

Kui juhuslikul suurusel on antud jaotusseadus, siis öeldakse, et juhuslik suurus on antud jaotusega või juhuslik suurus allub antud jaotusele.

2.2. Diskreetse juhusliku suuruse jaotusseadus

Vaatleme diskreetset juhuslikku suurust X võimalike väärtustega $x_1, x_2, ..., x_n$, mida lühidalt tähistame edaspidi $X : x_1, x_2, ..., x_n$. See, et juhuslik suurus X omandab ühe neist väärtustest, näiteks $X = x_2$, on tüüpiline juhuslik sündmus mingi tõenäosusega. Tähistame tõenäosused

$$P(X = x_k) = p_k \quad (k = 1, 2, ...n),$$

mida nimetame diskreetse juhusliku suuruse X **tõenäosusfunktsiooniks**. Ilmselt kehtib

$$\sum_{k=1}^{n} p_k = 1, \tag{2.1}$$

kuna sündmused $X = x_k$ on ainuvõimalikud, üksteist välistavad ja moodustavad täissüsteemi. Võrduse (2.1) paremal poolel olev arv 1 on seega jaotatud juhusliku suuruse võimalike väärtuste vahel. Siit tuleneb ka nimetus *jaotus*.

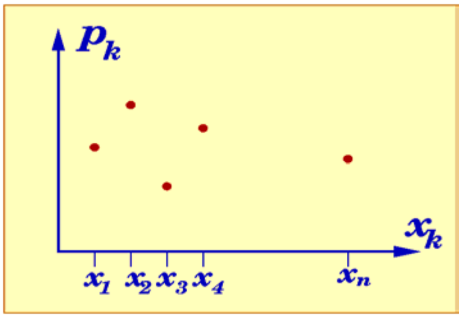
Diskreetse juhusliku suuruse lihtsaimaks jaotusseaduse esitamisvormiks on tabel, kus on loetletud võimalikud väärtused ja neile vastavad tõenäosused

X	x_1	x_2	...	x_n
P	p_1	p_2	...	p_n

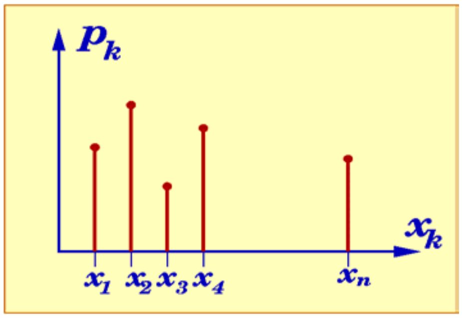
Niisugust tabelit nimetatakse juhusliku suuruse **jaotustabeliks.**

Praktikas sobib kasutada jaotustabelit, kui võimalike väärtuste arv ei ole suur. Mõnedel juhtudel saab jaotusseaduse anda ka valemi kujul

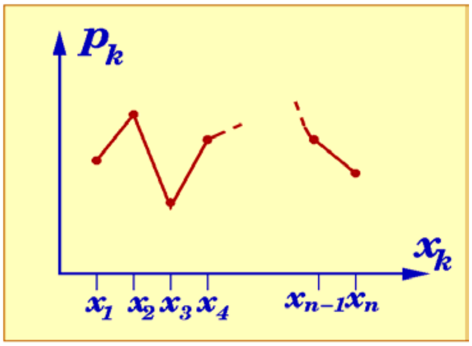
$$P(X = x_k) = g(x_k).$$



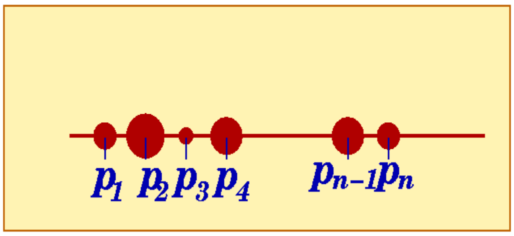
Joonis 2.1



Joonis 2.2



Joonis 2.3



Joonis 2.4

Jaotustabeli võib esitada ka graafiliselt, märkides juhusliku suuruse võimalikud väärtused abstsissteljel ja tõenäosused ordinaatteljel. Seega saame joonisel üksikud isoleeritud punktid (joonis 2.1). Selle paremaks visualiseerimiseks võib toimida mitmeti. Ühendades punktid (x_k, p_k) nende projektsioonidega $(x_k, 0)$ abstsissteljel sirglõikude abil, saame nn **jaotusspektri** (joonis 2.2). Kui ühendame punktid (x_k, p_k) omavahel sirglõikudega (see on ainult näitlikustamiseks), saame nn **jaotuspolügooni** (joonis 2.3). Kasutatakse veel mehaanilist interpretatsiooni (joonis 2.4). Siin on punktid x_k kui materiaalsed punktid massidega p_k, kusjuures masside summa on võrdne ühega .

Näide 1. Teelõigul on 4 valgusfoori. Igaüks neist fooridest lubab liiklusvahendi liikumise tõenäosusega p. Leida liiklusvahendi poolt peatuseta (kuni esimese peatuseni) läbitud fooride arvu jaotustabel.

Olgu juhuslik suurus X liiklusvahendi poolt peatuseta läbitud fooride arv. Ilmselt X võimalikud väärtused on 0 (ei läbi ühtki foori), 1, 2, 3 ja 4. Tähistame tõenäosuse, et liiklusvahend ei läbi konkreetset foori $q = 1 - p$. Siis tõenäosused avalduvad

Tõenäosus, et ei läbita ühtki foori:

$$P(X = 0) = q.$$

Tõenäosus, et läbitakse esimene foor, kuid peatutakse teise ees:

$$P(X = 1) = p\ q\,.$$

Tõenäosus, et läbitakse esimene ja teine foor, kuid peatutakse kolmanda ees:

$$P(X=2)=p^2q.$$

Tõenäosus, et läbitakse esimene, teine ja kolmas foor, kuid peatutakse neljanda ees:

$$P(X=3)=p^3q.$$

Tõenäosus, et peatuseta läbitakse kõik neli foori:

$$P(X=4)=p^4.$$

Niisiis on jaotustabel kujul

X	0	1	2	3	4
P	q	pq	p^2q	p^3q	p^4

Lihtne on kontrollida, et kehtib võrdus (2.1):

$$\sum_{k=0}^{4} p_k = q(1+p+p^2+p^3)+p^4=(1-p)(1+p+p^2+p^3)+p^4=1.$$

Konkreetsel erijuhul $p=0.5$ saame

X	0	1	2	3	4
P	0,5	0,25	0,125	0,0625	0,0625

Näide 2. **Binoomjaotus** ehk **Bernoulli jaotus** (A. Moivre, 1711; Jakob Bernoulli, 1713). Vaatleme n sõltumatut katset, kui igal katsel toimub sündmus A konstantse tõenäosusega p. Olgu diskreetne juhuslik suurus X – sündmuse A toimumiste arv selles katseseerias. Leida juhusliku suuruse X jaotusseadus.

Ilmselt $X: 0, 1, 2, ..., n$.

Sündmuse A toimumise tõenäosus üksikkatsel on p, mittetoimumise tõenäosus aga on $q=1-p$. Kui teeme n katset ja saame mingis konkreetses järjestuses sündmuse A toimumise k -l korral, ning järelikult, vastandsündmuse $\overline{A}$ toimumise $n-k$ -l korral:

$$A,\ldots\overline{A},\overline{A},\ldots\ldots A,\overline{A},\overline{A},A,A,\overline{A},\overline{A},\ldots\ldots\ldots\overline{A},A,\overline{A},\ldots\ldots,\overline{A},A,A,\overline{A},\ldots\ldots\ldots\overline{A}.A,\overline{A},\ldots\ldots\ldots A,\overline{A},\overline{A},\ldots\ldots\ldots$$

konkreetne katseseeria, kokku n katset,

siis on sellise tulemuse saamise tõenäosus p^kq^{n-k}. Kokku on erinevaid võimalusi saada erinevas järjestuses k soodsat katset võrdne kombinatsioonide arvuga n liikmest k kaupa, s.o C_n^k -ga. Seega

$$P(X=k)=P_n(k)\equiv\binom{n}{k}p^k q^{n-k}.$$

Seega oleme jaotusseaduse esitanud tõenäosusfunktsioonina valemi abil. Loomulikult võime selle esitada ka jaotustabelina

X	0	1 …		k	…	n
P	q^n	npq^{n-1}	...	$\binom{n}{k}p^k q^{n-k}$	...	p^n

Niisugust jaotusseadust nimetatakse **binoomjaotuseks** ehk **Bernoulli jaotuseks**. Nimetus binoomjaotus tuleneb sellest, et Newtoni binoomvalemis

$$(p+q)^n=\sum_{k=0}^{n}\binom{n}{k}p^k q^{n-k}$$

üldliige ongi meid huvitav tõenäosus $P_n(k)$. Viimasest võrdusest järeldub veel, et kehtib võrdus (2.1), kuna $p+q=1$.

Näide 3. **Poissoni jaotus.**

Vaatleme veel kord katseseeriat, milles on n katset ja sündmuse A tõenäosus igal katsel on p. Tõenäosus, et katseseerias sündmus A toimub k korda, avaldub Bernoulli valemiga. Kui n on küllalt suur ja p väga väike, siis valem praktiliseks arvutamiseks ei sobi, kuna on tegemist väga suurte ja väga väikeste arvude korrutamisega. Leiame sobivama valemi.

Eeldame, et korrutis np on konstantne ja tähistame

$$\lambda=np,$$

täpsemini

$$\lambda=\lim_{\substack{n\to\infty\\ p\to 0}} np.$$

Teisendame Bernoulli valemit (asendades suuruse p eelmisest võrdusest)

$$P_n(k)=\frac{n(n-1)\ldots[n-(k-1)]}{k!}p^k(1-p)^{n-k}=$$

$$=\frac{n(n-1)\ldots[n-(k-1)]}{k!}\left(\frac{\lambda}{n}\right)^k\left(1-\frac{\lambda}{n}\right)^{n-k}.$$

Et n on väga suur, siis leiame piirväärtuse

$$\lim_{n\to\infty}P_n(k)=\frac{\lambda^k}{k!}\lim_{n\to\infty}\left[1\cdot\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\ldots\left(1-\frac{k-1}{n}\right)\left(1-\frac{\lambda}{n}\right)^{n-k}\right]=$$

$$=\frac{\lambda^k}{k!}\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^n\left(1-\frac{\lambda}{n}\right)^{-k}=\frac{\lambda^k}{k!}e^{-\lambda}.$$

Saadud jaotust

$$P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!} \qquad (k = 0, 1, ...) \tag{2.2}$$

nimetatakse **Poissoni jaotuseks** (A. de Moivre, 1718; S.D. Poisson, 1830, 1837). Jaotust nimetatakse samuti **väikeste arvude seaduseks** ehk **harva esinevate sündmuste seaduseks**, sest see kirjeldab harva esinevaid nähtusi.

Ka siin kehtib võrdus (2.1), sest

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

Poissoni jaotuse jaotustabel on

X	0	1	…	k	…
P	$e^{-\lambda}$	$e^{-\lambda}\lambda$	...	$\dfrac{e^{-\lambda}\lambda^k}{k}$	...

Binoomjaotust võib lähendada Poissoni jaotusega, kui $p < 0.1$.

2.3. Pideva juhusliku suuruse jaotustihedus

Pidevalt muutuva suuruse X korral on tõenäosus, et tuleb etteantud fikseeritud väärtus $X = x$, võrdne nulliga:

$$P(X = x) = 0.$$

Et seda demonstreerida, vaatleme näiteks lõiku $[0,1]$ *ehk* $0 \le x \le 1$.

Küsime – kui suur on tõenäosus, et sellest intervallist juhuslikult valitud arv on võrdne arvuga $x = 0.31813452$?

Küsimus on sisuliselt mõttetu, sest tõenäosuse arvutamise reegli järgi

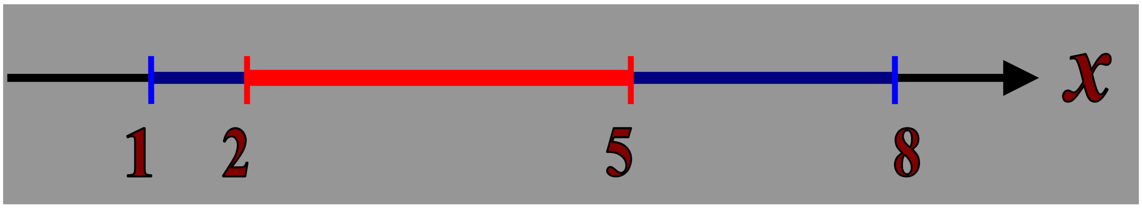
$$P = \frac{\text{soodsad} \quad \text{võimalused}}{\text{kõik võimalused}}$$

saame tulemuseks

$$P = \frac{1}{\infty} = 0 \ .$$

Seepärast on mõtet vaid tõenäosusel, et juhusliku suuruse realisatsioon asub mingis intervallis $x_1 < X < x_2$, või poollõigul $x_1 \le X < x_2$, või rahuldab tingimust $X < x$, kus x on ette antud.

Näide. Juhuslik suurus x võib omandada võrdse tõenäosusega mistahes väärtusi lõigust **[1, 8].** Kui suur on tõenäosus, et juhuslikult valitud x väärtus kuulub alamlõiku **[2, 5]** (joonis 3.1)



Joonis 3.1

Otsitav tõenäosus määratakse vastavate lõikude pikkuste suhtena

$$p = \frac{\text{soodsa piirkonna pikkus}}{\text{kogu piirkonna pikkus}} = \frac{5-2}{8-1} = \frac{3}{7} = 0{,}428\,.$$

Siin on ilmselt tegemist geomeetrilise tõenäosuse juhtumiga, kus geomeetriliseks kujundiks on sirglõik ning juhusliku sündmuse realiseerumise tõenäosus on võrdeline selle lõigu pikkusega.

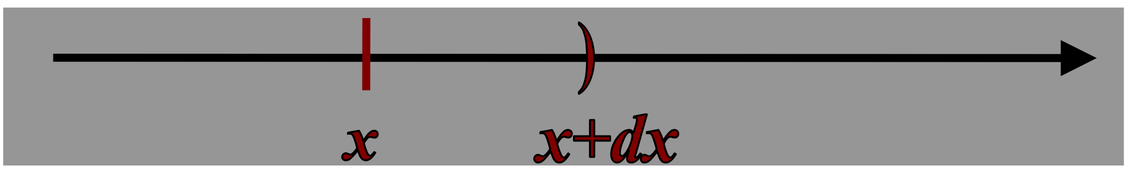
Toodud näites ei olnud soodsas piirkonnas $[2,5]$ eelistatumaid ega vähemeelistatumaid alasid. Praktilistes ülesannetes aga harilikult on.

Näiteid. Kujutagu X endast patsiendi kehatemperatuuri **°C,** siis praktiliselt kõik võimalikud x väärtused katab piirkond $[34,42]$. Selles piirkonnas on aga temperatuuride alampiirkond $[36,38]$ üldiselt hoopis tõenäosem kui $[40,42]$.

Või teine näide. Kujutagu x endast täiskasvanud mehe kehakaalu (korrektsem oleks öelda massi!) kilogrammides. Kogu lubatav x väärtuste piirkond on hinnanguliselt $[10,300]$. Samas on ilmne, et alampiirkond $[60,90]$ on palju tõenäosem kui sama lai piirkond $[160,190]$.

Juhusliku suuruse võimalike väärtuste erinevate piirkondade tõenäosust väljendab **juhusliku suuruse jaotustihedus.** Kasutatakse ka termineid **tõenäosustihedus, tihedusfunktsioon, jaotusfunktsioon.** *NB! Termin **jaotusfunktsioon** on levinud ja kasutusel füüsikute ja keemikute hulgas ning vastavate suundade erialakirjanduses. Matemaatikud tähistavad selle terminiga tavaliselt suurust, mida käesolevas kursuses (hiljem, allpool) tähistame terminiga **integraalne jaotusfunktsioon.** Seepärast peab olema termini **jaotusfunktsioon** kasutamisega ettevaatlik ja alati täpsustama, mida selle termini all konkreetselt silmas peetakse. Käesolevas kursuses on jaotusfunktsioonil alati juhusliku suuruse jaotustiheduse tähendus.*

Vaatleme reaalteljel (lõpmata) lühikest poolintervalli $[x, x+dx)$, mis sisaldab vasema otspunkti x, kuid ei sisalda parempoolset otspunkti $x+dx$ (vt joonis 3.2).



Joonis 3.2

Poolintervalli kasutamine on tarvilik, et üks ja sama punkt ei oleks korraga mitmes alamhulgas. Vaadeldes kaht kõrvutiseisvat poolintervalli $[x, x+dx)$ *ja* $[x+dx, x+dx+dx_1)$ *on kindel, et nende kokkupuutepunkt* $x+dx$ *kuulub parempoolsele intervallile, kuid ei kuulu vasempoolsele, ning samal ajal ei ole ka olukorda, kus see punkt ei kuuluks kummalegi alamhulgale.*

Tähistame tõenäosuse, et pideva muutkonnaga juhuslik suurus X satuks sellesse poolintervalli $[x, x+dx)$ sümboliga dP. Ilmselt tõenäosus dP on võrdeline elementaarlõigu pikkusega dx:

$$dP = f(x)dx. \tag{3.1}$$

Võrdetegurit $f(x)$ nimetatakse juhusliku suuruse jaotustiheduseks. Vastavalt toodud valemile võime kirjutada

$$f(x) = \frac{dP}{dx}, \tag{3.2}$$

s.o tõenäosustihedus on tõenäosus, et X satub punkti x lähiümbrusesse, arvutatuna ühikulise intervalli kohta.

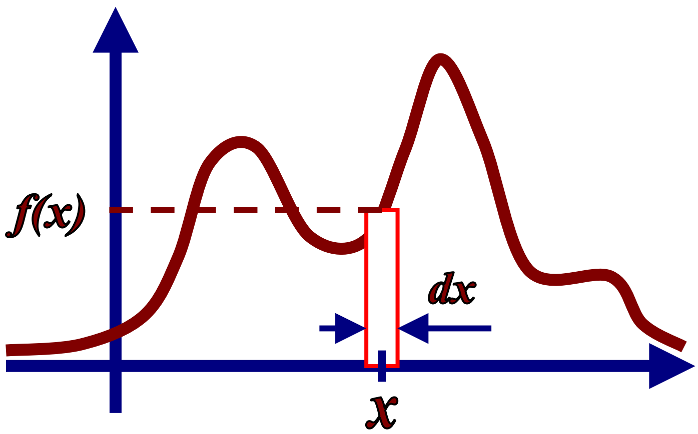
Tõenäosustiheduse olemust illustreerib võrdlus tavalise massitihedusega. Kui vaadelda reaaltelge muutuva läbimõõduga vardana, mille kogumass on üks, siis dP on poolintervalli $[x, x+dx)$ mass ja $f(x)$ varda massi(joon-)tihedus punktis x.

Tõenäosustiheduse põhiomadused.

1. Mittenegatiivsus:

$$f(x) \geq 0\ .$$

See omadus järeldub määrangust (3.1). Kuna loeme intervalli dx positiivseks ja kehtib $dP \geq 0$, siis järeldub meie väide. Jaotusfunktsiooni põhimõtteline kuju on toodud joonisel 3.3.



Joonis 3.3

2. Elementaartõenäosuse geomeetriline tõlgendus. Elementaartõenäosus dP on ristküliku pindala, mille aluse laius on dx ja kõrgus on $f(x)$ (joonis 3.3). Sel põhjusel nimetatakse elementaartõenäosust dP ka ***tõenäosuselemendiks.***

3. Tõenäosus lõplikul lõigul. Summeerides valemi (3.1) üle elementaarintervallide dx lõplikul poolintervallil $[a,b)$, s.o summeerides intervalli elementaartõenäosused dP, saame tõenäosuse, et juhuslik suurus X satub sellesse vahemikku:

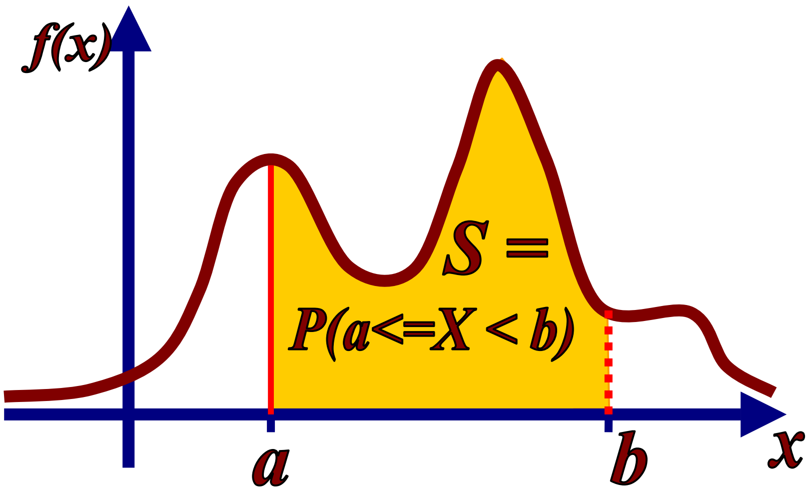
$$P(a \le X < b) = \int_a^b dP = \int_a^b f(x)dx. \tag{3.3}$$

Vastavalt määratud integraali definitsioonile on paremal oleva integraali väärtuseks kõvera $f(x)$ alla jääva, vasemalt ja paremalt vertikaaljoontega $x = a$ ja $x = b$ ning alt reaalteljega piiratud kujundi pindala S (joonis 3.4)

Definitsioon. Tõenäosust $p = P(a \le X < b)$ nimetame lõigu $[a,b)$ **usaldusnivooks.**

Usaldusnivoo sõltub konkreetsest jaotusest ja lõigu asukohast ja pikkusest.

Usaldusnivood väljendatakse sageli (näiteks füüsikalistel mõõtmistel) ka protsentides.



Joonis 3.4

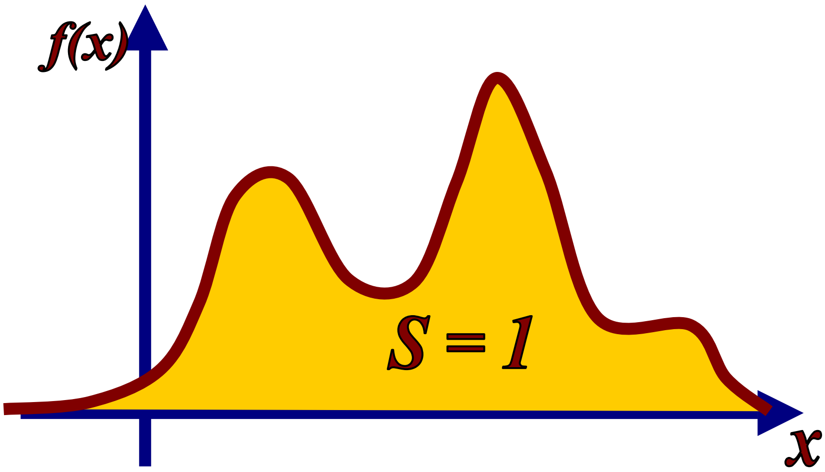
4. Nullile lähenemine argumendi kasvamisel. Selleks, et integraal (3.3) eksisteeriks kõikide a ja b väärtuste korral, peab kehtima

$$\lim_{x\to\pm\infty} f(x) = 0.$$

5. Normeerimistingimus:

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

See tingimus väljendab tõsiasja, et suvalise X väärtuse saamine peab olema tõene sündmus.



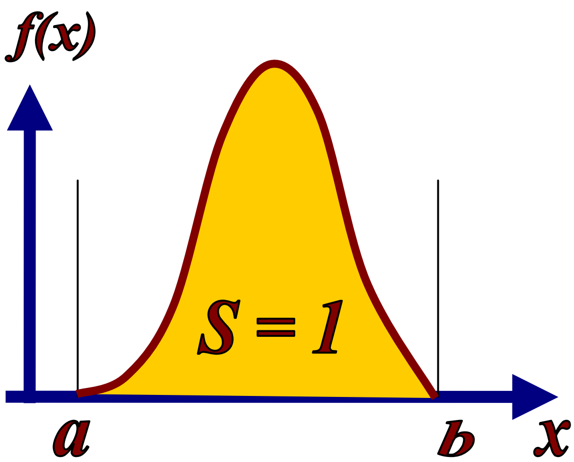
Joonis 3.5

Geomeetriliselt tähendab see, et jaotustiheduse ja x-telje poolt piiratud kujundi kogupindala on võrdne ühega, nagu näidatud joonisel 3.5.

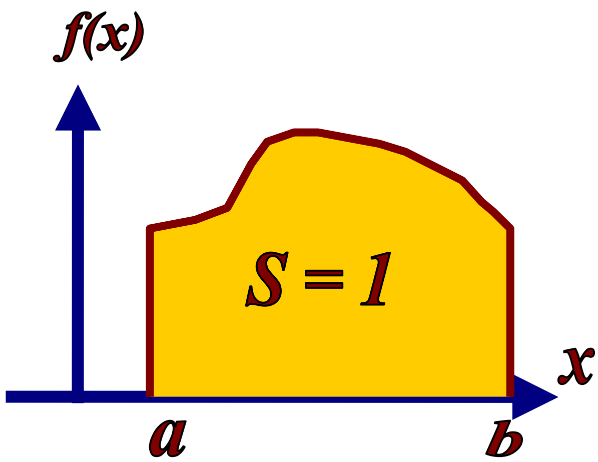
Märkus. Iga funktsioon, mis rahuldab mittenegatiivsuse tingimust 1 ja normeerimistingimust 5, on tõenäosustihedus (s.t võib olla mingi tõenäosustihedus).

6. Lõplikul lõigul lokaliseeritud tõenäosustihedus.

Võib osutuda, et tõenäosustihedus on null väljaspool mingit lõplikku intervalli $[a,b]$. Sel juhul on ka tõenäosus, et juhusliku suuruse realisatsioon satub väljapoole seda intervalli, võrdne nulliga. Tõenäosustihedus võib seejuures saada nulliks intervalli otstes (joonis 3.6a), aga see ei ole tingimata tarvilik ning ta võib olla ka nullist erinevate väärtustega otstes (joonis3.6b). Intervalli $[a,b]$ nimetatakse **jaotuse kandjaks.**



Joonis 3.6a



Joonis 3.6b

***Märkus**. Iga lõplikul lõigul lokaliseeritud jaotust $f(x)$ võib vaadelda antuna kogu reaalteljel, jätkates teda nulliga a-st vasemale ja b-st paremale. Uus jaotus*

$$f^*(x) = \begin{cases} 0, & x < a \\ f(x), & a \le x \le b \\ 0, & x > b \end{cases}$$

rahuldab kõiki jaotustihedusele esitatavaid nõudeid. Selles mõttes on kõik jaotused antud kogu reaalteljel. Rakendustes on muidugi suur erinevus, kas jaotus on tegelikult antud lõplikul lõigul või on suurus jaotunud nullist erineva tõenäosustihedusega kogu reaalteljel.

Lõpmatu muutumispiirkonnaga jaotust võib ka vaadelda kui joonisel 3.6a antud jaotuse piirjuhtu protsessis $a \to -\infty, b \to \infty$.

Samuti on mõttekad pooltelgedel $(-\infty, b]$ ja $[a, \quad \infty)$ lokaliseeritud jaotused.

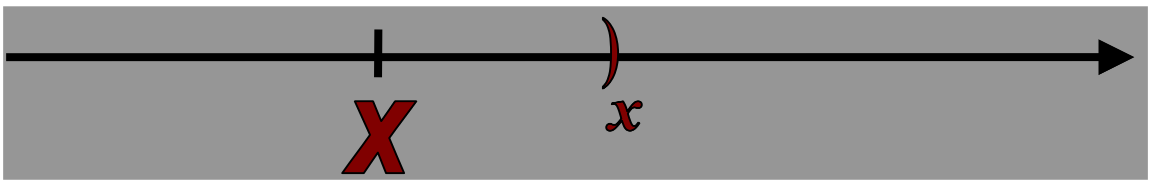
2.4. Integraalne jaotusfunktsioon

Integraalne jaotusfunktsioon $F(x)$ on kõrvuti jaotustihedusega teine enamlevinud vahend pideva juhusliku suuruse X kirjeldamiseks.

Juhusliku suuruse X **integraalseks jaotusfunktsiooniks** $F(x)$ nimetatakse funktsiooni, mis on võrdne tõenäosusega, et kehtib võrratus $X < x$ **:**

$$F(x) = P(X < x).$$

Seega, **funktsioon $F(x)$ on tõenäosus, et juhuslik punkt X asetseb arvsirgel punktist x vasakul, piirkonnas** $(-\infty, x)$ (joonis 4.1).



Joonis 4.1

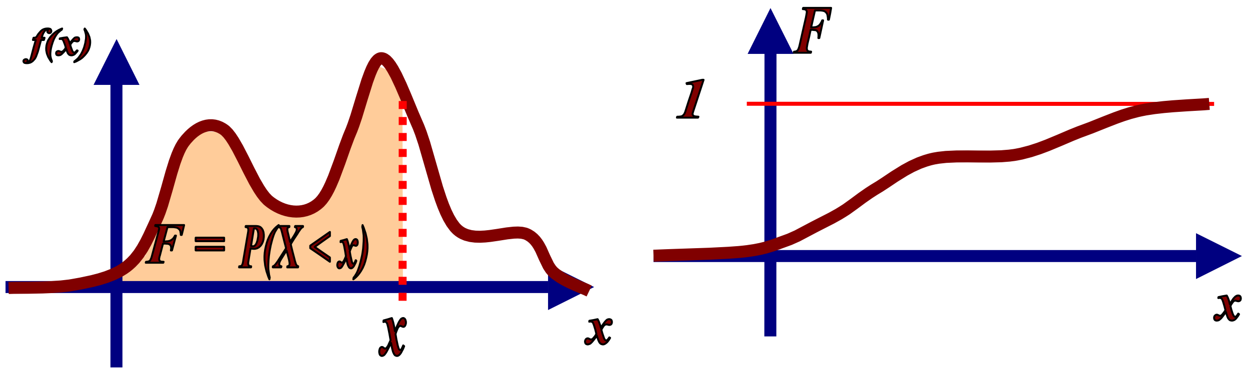
Ümarsulg x väärtuse kohal tähistab asjaolu, et tegemist on paremalt avatud intervalliga, mille otspunkt (käesoleval juhul x) ei kuulu selle intervalli koosseisu.

Võttes avaldises (3.3) vasakpoolse raja võrdseks miinus lõpmatusega ja parempoole raja b võrdseks x-ga, saame

$$P(-\infty < X < x) \quad \equiv \quad P(X < x) \quad = \quad \int_{-\infty}^{x} f(x')dx'.$$

Seega

$$F(x) = \int_{-\infty}^{x} f(y)dy. \tag{4.1}$$



Joonis 4.2

Tegemist on ülemise raja funktsiooniga ($F(x)$ argumendiks on integraali ülemine raja). Segaduse vältimiseks on seejuures integraalialune muutuja tähistatud y-ga.

Nagu näeme, on integraalne jaotusfunktsioon leitav, kui on teada jaotustihedus $f(x)$. Vastupidi, teades integraalset jaotusfunktsiooni, on võimalik arvutada jaotustihedus, nagu näitab allpool omaduste kirjelduses omadus 1.

Integraalse jaotusfunktsiooni omadused

1. Jaotustihedus on integraalse jaotusfunktsiooni tuletis

$$f(x) = \frac{dF}{dx}. \tag{4.2}$$

Tõestuseks vaatame diferentsiaali

$$dF = F(x+dx) - F(x) = \int_{-\infty}^{x+dx} f(y)dy - \int_{-\infty}^{x} f(y)dy = $$

$$= \int_{x}^{x+dx} f(y)dy = f(x)dx$$

Et viimases integraalis integreerimislõik $[x, x+dx]$ on lõpmata lühike, siis võib lugeda integraalialuse funktsiooni konstandiks ning ta integraali alt välja tuua, peale mida järgnebki viimane võrdus. Seega oleme saanud

$$dF = f(x)dx,$$

millest järeldub (4.2).

Kommentaar. *Tihedusfunktsiooni ja integraalse jaotusfunktsiooni kasutamine on samaväärne – teades ühte neist, on alati võimalik üle minna teisele. Üldiselt on nii, et füüsikas ja keemias on levinum jaotustiheduse (jaotusfunktsiooni) kasutamine, kuna matemaatikud eelistavad integraalset jaotusfunktsiooni.*

2. Integraalse jaotusfunktsiooni tõkestatus. Integraalne jaotusfunktsioon $F(x)$ rahuldab võrratusi

$$0 \le F(x) \le 1.$$

Et jaotusfunktsioon on sündmuse $X < x$ tõenäosus ja tõenäosus on reaalarv nulli ja ühe vahel, siis siit järeldubki omadus. Geomeetriliselt tähendab see, et jaotusfunktsiooni graafik paikneb kahe paralleelse sirge $y = 0$ ja $y = 1$ vahel (vt parempoolne joonis 4.2).

3. Integraalse jaotusfunktsiooni monotoonsus. Integraalne jaotusfunktsioon $F(x)$ on mittekahanev funktsioon:

$$\text{Kui} \quad x_1 < x_2, \quad \text{siis} \quad F(x_1) \le F(x_2).$$

Kõige vahetumalt järeldub F monotoonsus tema tuletise f mittenegatiivsusest (vt eelmine punkt, omadus 1). Kui funktsiooni tuletis on null või suurem nullist, siis funktsioon ei saa kahaneda, sest tema tõusunurga tangens (võrdub tuletisega) on kõikjal mittenegatiivne.

2.5. *Ühtlane jaotus*

Ühtlane jaotus on alati antud lõplikul lõigul $[a,b]$. See on lihtsaim, kuid rakendustes üsna levinud jaotus. Ühtlase jaotuse üldkuju on

$$f(x) = C, \qquad a \le x \le b .$$

Seejuures konstandi C saame normeerimistingimusest. Arvutame

$$\int_a^b f(x)dx = \int_a^b Cdx = C\int_a^b dx = C(b-a).$$

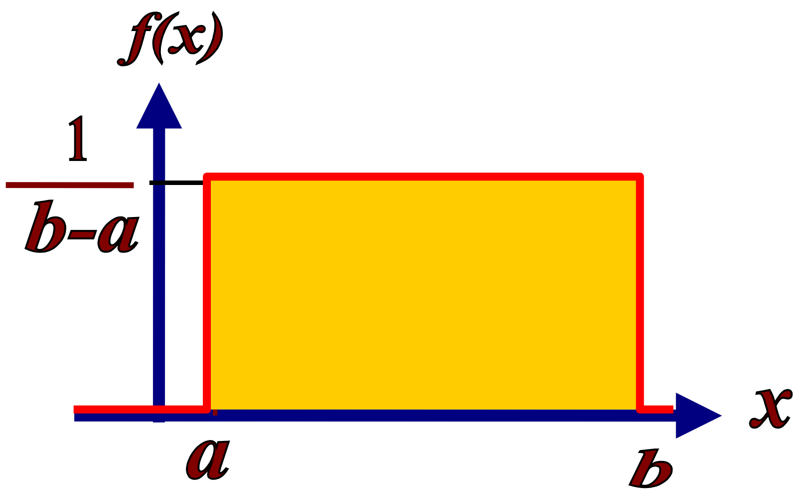
Et norm oleks üks, peab kehtima

$$C = \frac{1}{b-a}.$$

Seega ühtlase jaotuse lõplik kuju on

$$f(x) = \frac{1}{b-a}, \qquad a \le x \le b. \tag{5.1}$$

Ühtlase jaotuse graafik on toodud joonisel 5.1.



Joonis 5.1

Leiame ka ühtlase jaotuse integraalse jaotusfunktsiooni.

Kui $x < a$, siis $F(x) = 0$.

Kui $a < x < b$, siis valemi (4.1) põhjal

$$F(x) = \int_{-\infty}^{x} f(y)dy = \int_{a}^{x} \frac{dy}{b-a} = \frac{x-a}{b-a}.$$

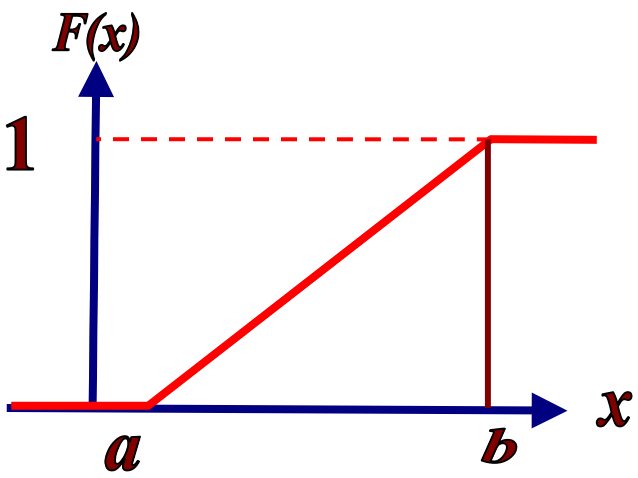
Kui $x > b$, siis

$$F(x) = \int_{a}^{b} \frac{dy}{b-a} + \int_{b}^{x} 0 \cdot dy = \frac{b-a}{b-a} = 1.$$

Seega

$$F(x) = \begin{cases} 0, & x \le a, \\ \dfrac{x-a}{b-a}, & a \le x \le b, \\ 1, & x \ge b. \end{cases} \tag{5.2}$$

Ühtlase jaotuse integraalse jaotusfunktsiooni graafik on joonisel 5.2.



Joonis 5.2

Leiame veel tõenäosuse, et juhusliku suuruse *X* väärtus on vahemikus $(\alpha,\beta)\subset[a,b]$. Valemi (3.3) põhjal

$$P(\alpha < X < \beta) = \int_{\alpha}^{\beta} \frac{1}{b-a}\,dx = \frac{\beta-\alpha}{b-a},$$

mille võime saada ka geomeetrilise tõenäosuse mõistet kasutades.

Ühtlase jaotuse näiteid.

1. Browni osakesed, mis alguses on vedelikku 'sisse pritsitud' ühte punkti, jaotuvad aja kasvades (s.o $t \to \infty$ korral) lõplikus anumas ühtlaselt. (Kui tahame olla korrektsed ja piirduda siin vaid ühemõõtmelise juhusliku suurusega, tuleks valida anumaks kapilaartoru, siis osakeste jaotus piki kapillaartoru on ühemõõtmelise juhusliku suuruse jaotustihedus.)

2. Mõõteriistast põhjustatud mõõtmisvead loetakse ühtlaselt jaotunuks mingis intervallis (nn B-tüüpi e. mittestatistilise päritoluga mõõtemääramatus, millest tuleb lähemalt juttu loengukursuse teises, füüsikalistele mõõtmistele pühendatud osas).

2.6. Eksponentjaotus

Eksponentjaotus on lihtsaim jaotus, mille kandjaks on kogu positiivne pooltelg $[0,\infty)$. Eksponentjaotuse integraalne jaotusfunktsioon on

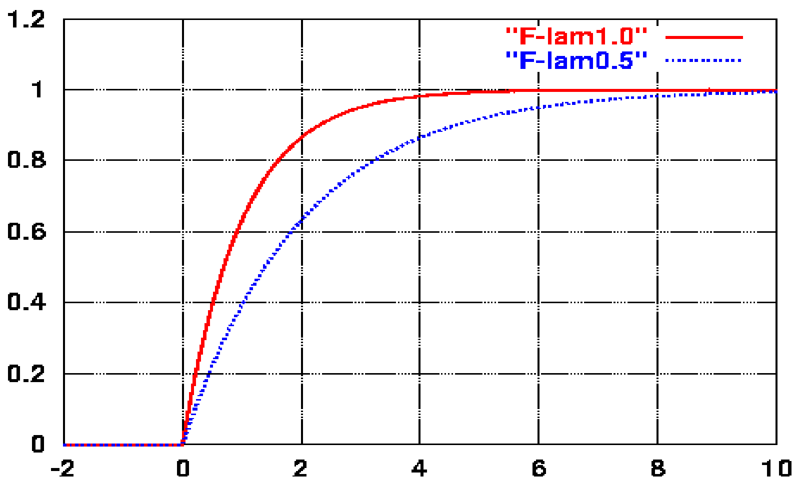
$$F(x) = \begin{cases} 0, & x < 0, \\ 1 - e^{-\lambda x}, & x \geq 0, \end{cases} \tag{6.1}$$

kus positiivne konstant λ on jaotuse parameeter. Integraalse jaotusfunktsiooni graafik on toodud joonisel 6.1a.

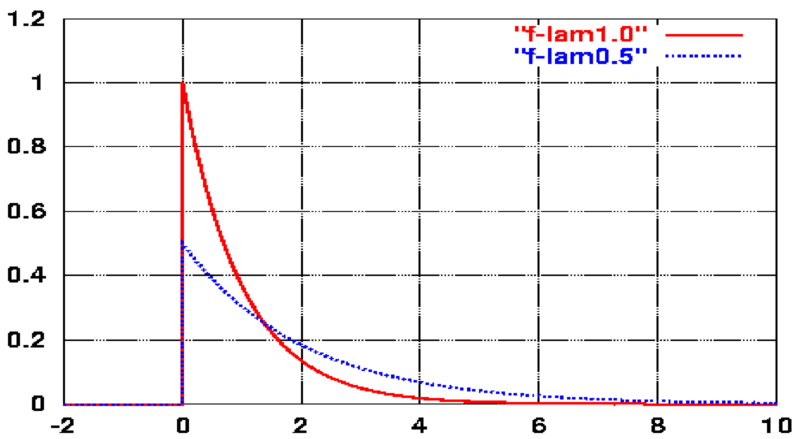
Võttes integraalsest jaotusfunktsioonist tuletise, saame jaotustiheduse jaoks avaldise

$$f(x) = \begin{cases} 0, & x < 0, \\ \lambda e^{-\lambda x}, & x \geq 0. \end{cases} \tag{6.2}$$

Seega, eksponentjaotuse tihedus on katkev punktis null, hüppe suurus on λ. Eksponentjaotuse graafik on toodud joonisel 6.1b.



Joonis 6.1a



Joonis 6.1b

Näide 1. *<u>Radioaktiivse isotoobi aatomi lõhustumise tõenäosus</u>* ajahetkel t ühikulise ajaintervalli jooksul on

$$f(t) = \frac{1}{t_0} \exp(-t / t_0), \tag{6.3}$$

kus parameeter t_0 on isotoobi karakteristlik eluiga (aeg, mille jooksul isotoobi aatomite arv kahaneb esialgsega võrreldes e (= naturaallogaritmi alus) korda. Vastavalt, integraalne jaotusfunktsioon

$$F(t) = 1 - \exp(-t/t_0) \tag{6.4}$$

annab tõenäosuse, et aatom on lõhustunud hetkeks t. Parameeter $\lambda = 1/t_0$ käesolevas näites. Tuumafüüsikas ja kiirguskaitses on kõrvuti t_0-ga kasutusel poolestusaeg $t_{1/2}$ – aeg, mille jooksul isotoobi aatomite arv kahaneb kaks korda:

$$F(t_{1/2}) = 1/2 \quad \text{ehk} \quad \exp(-t_{1/2}/t_0) = 1/2.$$

See jaotusseadus saadakse järeldusena lõhustumisseadusest, mille tuletuskäik on järgmine.

Olgu isotoobi aatomite arv antud konkreetses proovis $n(t)$ ja olgu nende muut aja dt jooksul dn. Ilmselt $dn \sim n$ ja $dn \sim dt$ (märk „~“ tähendab „võrdeline“). Seega $dn \sim n\, dt$. Tähistame võrdeteguri $-1/t_0$-ga:

$$dn = -(1/t_0)\, n\, dt.$$

Miinusmärk võtab arvesse, et osakeste arv kahaneb ja dn on negatiivne. Tegemist on diferentsiaalides antud võrrandiga, mis kirjeldab isotoobi aatomite arvu kahanemist ajas. Eraldades siin muutujad

$$\frac{dn}{n} = -\frac{dt}{t_0},$$

saame integreerimisel osakeste arvuks hetkel t:

$$n(t) = n_0 \exp(-t/t_0),$$

kus n_0 on isotoobi aatomite arv proovis alghetkel. Siit saame veel allesolevate aatomite suhteliseks hulgaks $n(t)/n_0 = \exp(-t/t_0)$. Kuna aatomite arv on väga suur ($n_0 \sim 10^{23}$, sest ühes gramm-moolis sisaldub Avogadro arv $N_A = 6.02 \cdot 10^{23}$ aatomit), siis on suhe $n(t)/n_0$ väga suure täpsusega seesama, mis tõenäosus, et konkreetne osake ei ole veel laguneda jõudnud. Järeldusena on $1 - n(t)/n_0$ võrdne tõenäosusega, et osake on hetkeks t juba laguneda jõudnud. Viimane ei ole aga midagi muud kui integraalne jaotusseadus (6.4).

Näide 2. Footoni neeldumise tõenäosus hägusas keskkonnas allub eksponentjaotusele. Tõenäosus, et footon läbib hägusas keskkonnas vahemaa x enne kui ta neeldub, on

$$P(X > x) = \exp(-x/l),$$

kus l on nn *footoni vaba tee pikkus* selles keskkonnas (l suurus sõltub keskkonna läbipaistvusest ja on seda väiksem, mida hägusem on keskkond). Vastavalt, tõenäosus, et footon on vahemaa x jooksul neeldunud, on antud integraalse jaotusfunktsiooniga (6.1), mis käesoleval juhul on

$$F(x) = 1 - \exp(-x/l).$$

2.7. Normaaljaotus

Normaaljaotus e. ***Gaussi jaotus*** on üks tähtsamaid jaotusseadusi juhuslikele suurustele, mis on jaotunud kogu reaalteljele ja võivad omandada väärtusi vahemikus $(-\infty, \infty)$.

Juhusliku suuruse jaotust nimetatakse **normaaljaotuseks** ehk **Gaussi jaotuseks**, kui jaotustihedus on

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-a)^2}{2\sigma^2}\right), \tag{7.1}$$

kus a on fikseeritud reaalarv (võib olla nii negatiivne kui positiivne, aga võib olla ka null), ja $\sigma > 0$ on fikseeritud positiivne reaalarv. Parameetreid a ja σ nimetatakse normaaljaotuse **keskväärtuseks** ja **ruuthälbeks**. Nende nimetuste päritolu saab selgemaks veidi hiljem, kui asume uurima jaotuste arvkarakteristikuid.

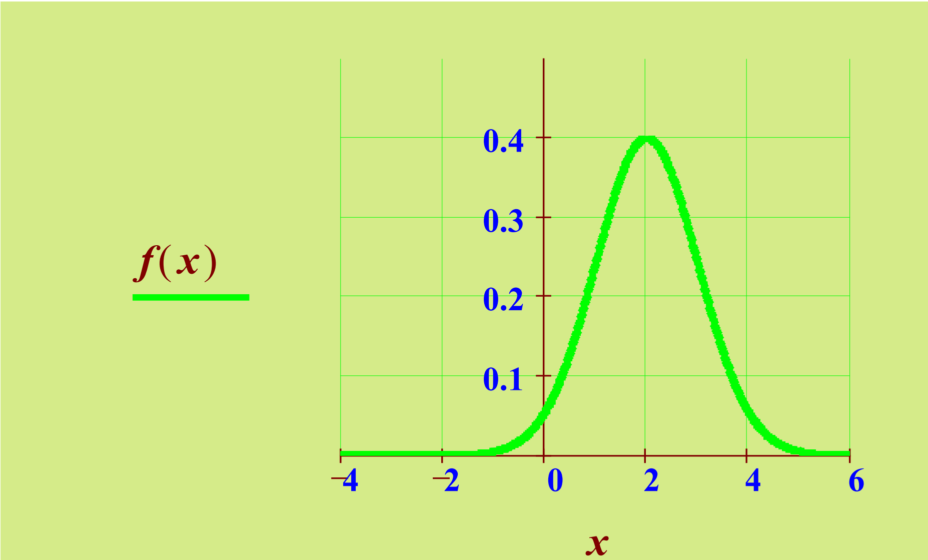
Jaotustiheduse määramispiirkond on kogu reaaltelg $\mathbf{R}$ (s.t argument x võib omandada väärtusi kogu reaalteljel). Kui $x \to \pm\infty$, siis $f(x) \to 0$. Seega on sirge $y = 0$ jaotustiheduse $f(x)$ asümptoot protsessis $x \to \pm\infty$.

Punktis $x = a$ on jaotustiheduse maksimum, kusjuures ta saavutab seal väärtuse

$$\max f(x) = f(a) = \frac{1}{\sigma\sqrt{2\pi}} \approx \frac{0.3989}{\sigma} \approx \frac{0.4}{\sigma}.$$

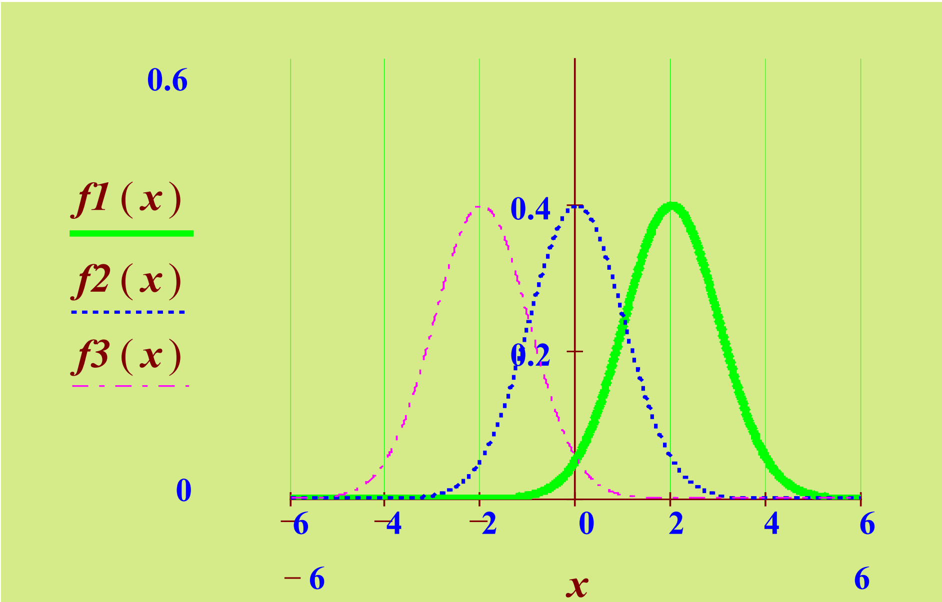
Jaotusjoon on sümmeetriline vertikaalsirge $x = a$ suhtes.

Normaaljaotuse tiheduse graafik on parameetrite $a = 2$, $\sigma = 1$ korral toodud joonisel 7.1.



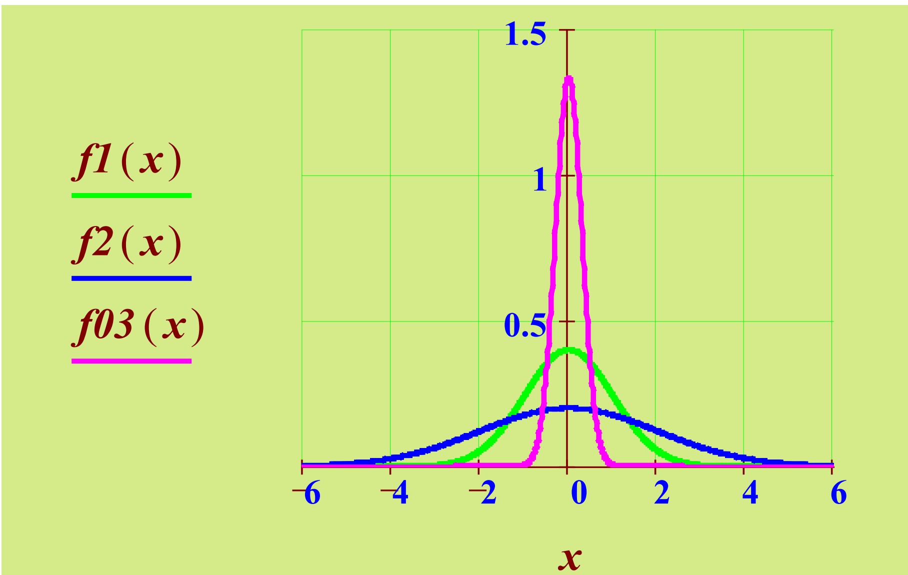
Joonis 7.1. Normaaljaotus $a = 2$, $\sigma = 1$ korral.

Parameetri a muutumisel joone asend muutub x-telje suhtes: a kasvades jaotus nihkub paremale (toimub kujundi x-telje sihiline paralleellüke). Mida suurem on a, seda paremal paikneb kõver (joonis 7.2).



Joonis 7.2. Erinevad normaaljaotused $\sigma = 1$ korral: $a = -2$ (kõver *f3(x)*), $a = 0$ (kõver *f2(x)*) ja $a = 2$ (kõver *f1(x)*).

Kui parameeter σ kasvab, siis kahanevad funktsiooni väärtused ja joon muutub lamedamaks – kõver surutakse kokku *y*-telje suunas. Kui σ kahaneb, siis muutub joon teravatipulisemaks – kõver venitatakse välja *y*-telje suunas (joonis 7.3).



Joonis 7.3. Normaaljaotus ($a = 0$) erinevate σ-de korral: *f1(x)* – $\sigma = 1$, *f2(x)* – $\sigma = 2$ ja *f03(x)* – $\sigma = 0{,}3$.

Normaaljaotus (7.1) **on normeeritud**, s.t kehtib

$$\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty} \exp\left(-\frac{(x-a)^2}{2\sigma^2}\right)dx \quad = \quad 1. \tag{7.2}$$

Et selles veenduda, tuleb arvutada integraal

$$I = \int_{-\infty}^{\infty} \exp\left(-\frac{(x-a)^2}{2\sigma^2}\right)dx \quad = \int_{-\infty}^{\infty} \exp\left(-\frac{x^2}{2\sigma^2}\right)dx .$$

(Siin esimene integraal läheb üle teiseks muutujavahetusel $x \to x'$: $\quad x' = x - a$ *.)*

Tavaliselt seda integraali arvutatakse kompleksfunktsioonide teooria vahendeid kasutades ja tõenäosusteooria elementaarkursustes jäetakse valem (7.2) tõestamata. Tegelikult on I *leidmine üsna elementaarne järgmise kavaluse kasutamisel. Arvutame esmalt suuruse* I^2 *ja seejärel võtame tulemusest ruutjuure. Suuruse* I^2 *arvutamiseks tuleb leida kahekordne integraal*

$$I^2 = \int_{-\infty}^{\infty} \exp\left(-\frac{x^2}{2\sigma^2}\right)dx \int_{-\infty}^{\infty} \exp\left(-\frac{y^2}{2\sigma^2}\right)dy \quad =$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp\left(-\frac{x^2}{2\sigma^2}\right)\exp\left(-\frac{y^2}{2\sigma^2}\right)dxdy$$

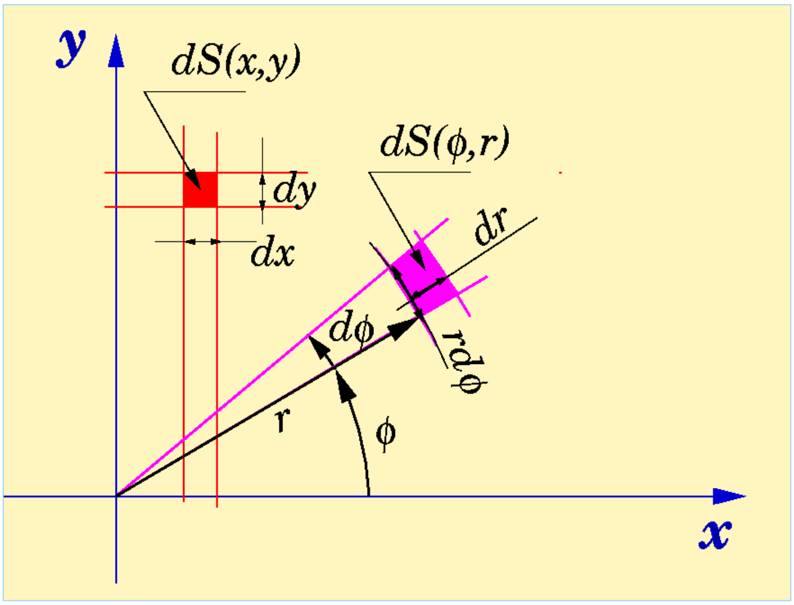
$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right)dxdy \quad = \iint_S \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right)dS(x,y) ,$$

kus $\quad dS(x,y) = dx \cdot dy$.

Läheme siin viimases pindintegraalis muutujatelt (ristkoordinaatidelt) x, y *üle polaarkoordinaatidele* r, φ *seostega*

$$x = r\cos(\varphi) , \quad y = r\sin(\varphi) .$$

Seejuures integreerimispiirkond S *, milleks on x,y-tasand, ja mis ristkoordinaatides esitub* $S = (-\infty < x < \infty , -\infty < y < \infty)$*, teiseneb piirkonnaks* $S = (0 < \varphi \le 2\pi , 0 \le r < \infty)$*, kusjuures pinnaelement on defineeritud kujul* $dS(\varphi, r) = r\, dr\, d\varphi$ *(vt joonis 7.4).*



Joonis 7.4

Et seejuures kehtib $x^2 + y^2 = r^2$*, on uutes muutujates pindintegraali arvutus õige lihtne*

$$I^2 = \iint_S \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right) dS(x,y) = \int_0^{2\pi} d\varphi \int_0^{\infty} dr\, r\, \exp\left(-\frac{r^2}{2\sigma^2}\right) =$$

$$= 2\pi \int_0^{\infty} d\,\frac{r^2}{2}\, \exp\left(-\frac{r^2}{2\sigma^2}\right)$$

$$= 2\pi\sigma^2 \int_0^{\infty} d\,\frac{r^2}{2\sigma^2} \exp\left(-\frac{r^2}{2\sigma^2}\right)$$

ja minnes siin integraali all üle muutujale $\varsigma = r^2/(2\sigma^2)$*, saame lõplikult*

$$I^2 = 2\pi\sigma^2 \int_0^{\infty} d\varsigma \exp(-\varsigma) = 2\pi\sigma^2.$$

Siit järeldub $I = \sigma\sqrt{2\pi}$ *. Normeering (7.2) on tõestatud.*

Kui $a = 0$ ja $\sigma = 1$, siis nimetatakse jaotust **standardiseeritud (standardseks) normaaljaotuseks**. Sel korral on jaotustihedus

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right). \tag{7.3}$$

Normaaljaotuse ja standardiseeritud normaaljaotuse vahel kehtib seos

$$f(x) = \frac{1}{\sigma}\varphi\left(\frac{x-a}{\sigma}\right).$$

Leiame tõenäosuse, et normaaljaotusega juhuslik suurus X omandab väärtuse vahemikust (α, β). Selleks kasutame valemit (3.3)

$$P(\alpha < X < \beta) = \int_{\alpha}^{\beta} f(x)dx = \frac{1}{\sigma\sqrt{2\pi}}\int_{\alpha}^{\beta}\exp\left(-\frac{(x-a)^2}{2\sigma^2}\right)dx.$$

Teisendame integraali muutujate vahetusega

$$t = \frac{x-a}{\sqrt{2}\,\sigma}, \qquad dx = \sqrt{2}\,\sigma\,dt,$$

siis

$$\begin{aligned} P(\alpha < X < \beta) &= \frac{1}{\sqrt{\pi}}\int_{(\alpha-a)/(\sqrt{2}\sigma)}^{(\beta-a)/(\sqrt{2}\sigma)}\exp\left(-t^2\right)dt = \\ &= \frac{1}{\sqrt{\pi}}\int_{0}^{(\beta-a)/(\sqrt{2}\sigma)}\exp\left(-t^2\right)dt - \frac{1}{\sqrt{\pi}}\int_{0}^{(\alpha-a)/(\sqrt{2}\sigma)}\exp\left(-t^2\right)dt \quad . \end{aligned} \tag{7.4}$$

Et viimased integraalid ei avaldu kvadratuurides (s.t, nad ei ole avaldatavad teadaolevate elementaarfunktsioonide kaudu), siis defineerime uue funktsiooni, mille kaudu avaldame nõutava tõenäosuse.

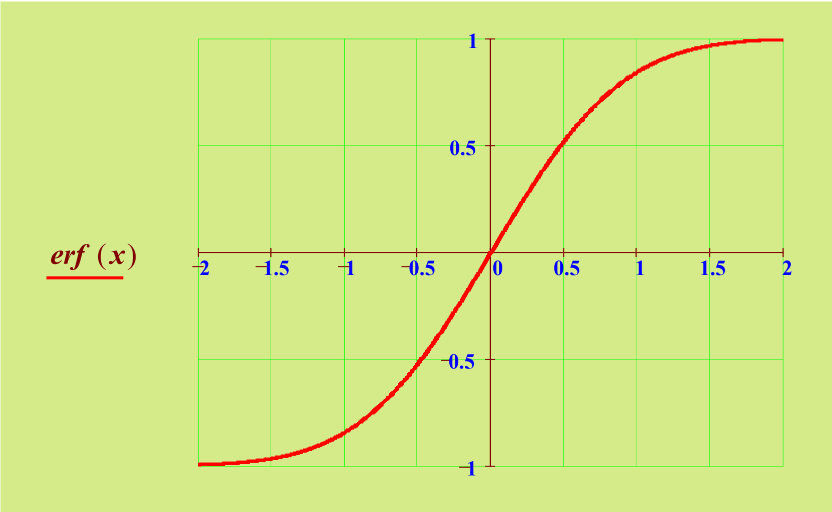
Veafunktsiooniks nimetatakse integraali

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}}\int_{0}^{x}\exp(-t^2)dt, \tag{7.5}$$

mida vaadatakse ülemise raja funktsioonina. Funktsiooni tähis **erf** tuleneb ingliskeelsest nimetusest **er**ror **f**unction. Funktsiooni graafik on esitatud joonisel 7.5. Veafunktsiooni tuleb käsitleda analoogiliselt elementaarfunktsioonidega (**sin, ln, exp,**...). Veafunktsiooni arvutamise võimalus on isegi mõnedel kallimatel taskuarvuteil. Tema arvutusalgoritm on kõikides enamlevinud matemaatilistes tarkvarapakettides (*Mathcad, Mathematica, Maple, ...*). Joonis 7.5 näiteks on valmistatud *Mathcad* keskkonnas.

Nõutava tõenäosuse (7.4) saab veafunktsiooni kaudu avaldada järgmiselt

$$P(\alpha < X < \beta) = \frac{1}{2}\left[\operatorname{erf}\left(\frac{\beta-a}{\sqrt{2}\,\sigma}\right) - \operatorname{erf}\left(\frac{\alpha-a}{\sqrt{2}\,\sigma}\right)\right]. \tag{7.6}$$



Joonis 7.5. Veafunktsioon.

Veafunktsiooni omadusi.

Mõlemad alljärgnevad omadused on näha graafikult joonisel 7.5. Tõestame need siin ka definitsioonist (7.5) lähtudes.

1. Veafunktsioon on paaritu funktsioon: kehtib

$$\mathrm{erf}(-x) = -\mathrm{erf}(x) \quad .$$

Tõepoolest, teisendades integraali (7.5) muutujate vahetusega

$$u = -t, \qquad du = -dt\,,$$

saame

$$\mathrm{erf}(x) = \frac{1}{\sqrt{\pi}} \int_0^x \exp\left(-u^2\right)(-du) = -\frac{1}{\sqrt{\pi}} \int_0^{-x} \exp\left(-u^2\right) du = -\mathrm{erf}(-x).$$

Seega, tõepoolest on tegu koordinaatide alguse suhtes paaritu funktsiooniga. Seetõttu tema väärtused tabuleeritakse või salvestatakse vajadusel vaid positiivsete argumendiväärtuste jaoks.

2. Kehtib võrdus

$$\lim_{x \to +\infty} \mathrm{erf}(x) = 1 \quad .$$

Kuna vastavalt valemile (7.6)

$$\frac{1}{2}\mathrm{erf}(\infty) - \frac{1}{2}\mathrm{erf}(-\infty) = 1$$

ja äsjatõestatud omaduse 1. põhjal $\mathrm{erf}(-\infty) = \quad -\mathrm{erf}(-\infty)$*, saame*

$$\frac{1}{2}\operatorname{erf}(\infty)-\frac{1}{2}\operatorname{erf}(-\infty)=\operatorname{erf}(\infty)=1 \quad .$$

3. Leiame tõenäosuse, et kehtiks võrratus

$$|X-a|<\varepsilon,$$

mis on samaväärne võrratustega

$$a-\varepsilon<X<a+\varepsilon.$$

Valemi (7.6) abil

$$P(|X-a|<\varepsilon)=P(a-\varepsilon<X<a+\varepsilon)=$$

$$=\frac{1}{2}\operatorname{erf}\left(\frac{\varepsilon}{\sqrt{2}\,\sigma}\right)-\frac{1}{2}\operatorname{erf}\left(-\frac{\varepsilon}{\sqrt{2}\,\sigma}\right)=\operatorname{erf}\left(\frac{\varepsilon}{\sqrt{2}\,\sigma}\right) \quad ,$$

s.t

$$P(|X-a|<\varepsilon)=\operatorname{erf}\left(\frac{\varepsilon}{\sqrt{2}\,\sigma}\right) \quad . \tag{7.7}$$

Lõpuks, **avaldame normaaljaotuse integraalse jaotusfunktsiooni veafunktsiooni kaudu**. Et

$$F(x)=\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{x}\exp\left(-\frac{(y-a)^2}{2\sigma^2}\right)dy,$$

siis muutujate vahetusega

$$u=\frac{y-a}{\sqrt{2}\,\sigma}, \quad dy=\sqrt{2}\,\sigma\,du$$

saame

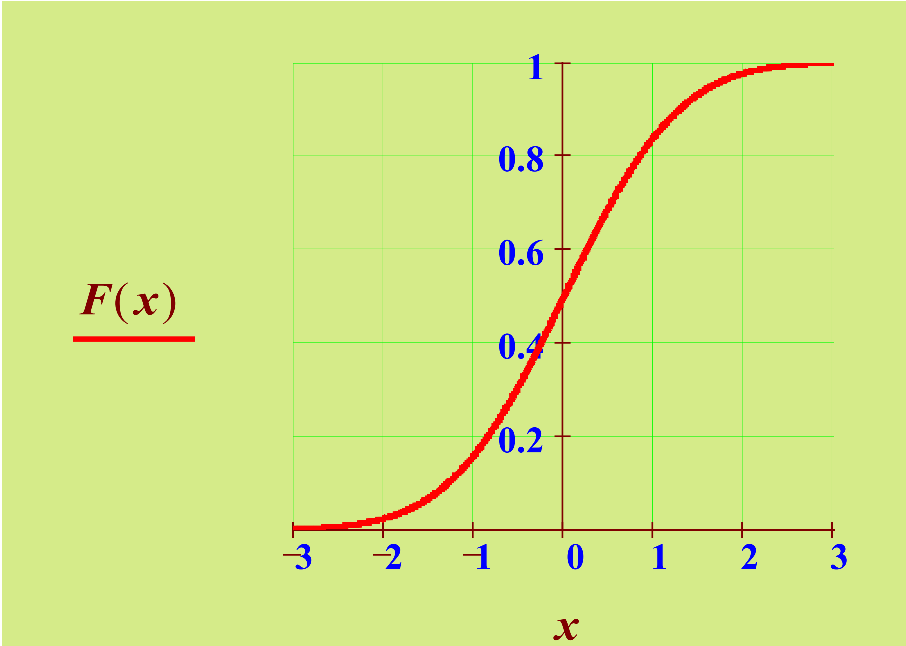
$$F(x)=\frac{1}{\sqrt{\pi}}\int_{-\infty}^{(x-a)/(\sigma\sqrt{2})}\exp\left(-u^2\right)du=\frac{1}{\sqrt{\pi}}\left(\int_{-\infty}^{0}+\int_{0}^{(x-a)/(\sigma\sqrt{2})}\right)\exp\left(-u^2\right)du.$$

Kuna

$$\frac{1}{\sqrt{\pi}}\int_{-\infty}^{0}\exp\left(-u^2\right)du=\frac{1}{\sqrt{\pi}}\int_{0}^{\infty}\exp\left(-u^2\right)du=\frac{1}{2}erf(\infty)=\frac{1}{2} \quad ,$$

Saame

$$F(x)=\frac{1}{2}\left[1+erf\left(\frac{x-a}{\sigma\sqrt{2}}\right)\right] \quad .$$



Joonis 7.6. Integraalne normaaljaotus.

,

Integraalsse jaotuse graafik on juhul $a = 0$, $\sigma = 1$ esitatud joonisel 7.6.

Normaaljaotuse näiteid

Näide 1. Gaasi molekulide kiirus V_x mingi konkreetse telje x sihis allub normaaljaotusele, mille keskväärtus a on null ja standardhälbe ruut $\sigma^2 \sim T$ (on võrdeline temperatuuriga). Lähemalt tehakse sellega tutvust statistilise füüsika kursuses.

Näide 2. „Süstime“ kapillaartorusse punkti $x = a$ Browni osakese. Igal järgneval ajahetkel on osakese tõenäosusjaotus piki kapillaartoru (tõenäosus leida osake punkti x ühikulises ümbruses) antud normaaljaotusega, mille keskväärtus on parajasti a ja mille standardhälbe ruut $\sigma^2 \sim \Delta t$, kus Δt on eksperimendi algusest möödunud ajavahemik. Tegemist on n.ö ajas laiali valguva normaaljaotusega (mis läbib järjestikku faasid 3, 2, 1 joonisel 7.3). Kui toru on piisavalt pikk, siis jaotustihedus läheneb lõpuks nullile. Küllalt lühikese kapillaari korral tekib suurte Δt-de piirkonnas kõrvalekaldumine normaalsest seadusest, kuna osake hakkab põrkumisel toru otsaseinaga sellelt peegelduma, mistõttu Browni osakese tõenäosusjaotus läheb lõpuks üle ühtlaseks jaotusseaduseks.

Näide 3. Normaaljaotusele allub nn A-tüüpi e. statistilise päritoluga mõõtemääramatus (mõõtmisviga), millest lähemalt tuleb juttu loengukursuse teises, füüsikalistele mõõtmistele pühendatud osas.

Näide 4. Kui mingi suuruse kujunemist põhjustab väga suure arvu tegurite koosmõju, mis kõik on juhuslikud, vastastikku sõltumatud ning ühtmoodi jaotunud, siis see suurus on jaotunud normaalselt. Eelnevad näited 1–3 on kõik just taolise tekkemehhanismiga. Konkreetse näitena võib tuua summa

$$s = \frac{1}{K}\sum_{i=1}^{K} X_i ,$$

kus kõik juhuslikud suurused X_i on ühtmoodi jaotunud (näiteks ühtlane jaotus lõigul $[-\varepsilon, \varepsilon]$), vastastikku sõltumatud suurused. Kui siin lasta K --l saada hästi suureks (teoreetilisel piiril $K \to \infty$), siis on s normaalselt jaotunud suurus (igakordne matemaatiline probleem on vaid tema parameetrite a ja σ leidmine sõltuvalt X_i parameetritest.

2.8. Deltajaotus

Kui punktis x_0 ainsat maksimumi omava jaotusfunktsiooni $f(x, x_0)$ laius läheneb nullile, nii et maksimumpunkti asend ei nihku, saame piiril lõpmata kitsa, punktis x_0 lõpmata kõrge jaotustiheduse, mida nimetatakse **deltajaotuseks, delta-funktsiooniks** ehk **Diraci funktsiooniks** (inglise füüsiku Paul Dirac'i auks, kes võttis selle funktsiooni esmakordselt kasutusele kvantmehaanikas) ja tähistatakse sümboliga $\delta(x - x_0)$.

Vaatame näiteks normaaljaotust (7.1):

$$\xi(x - x_0;\sigma) \equiv f(x) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{(x - x_0)^2}{2\sigma^2}\right),$$

kus keskväärtus a on tähistatud x_0-iga. Selle jaotuse abil saame määrata deltajaotuse kui piirväärtuse

$$\delta(x - x_0) \quad = \quad \lim_{\sigma\to 0} \xi(x - x_0;\sigma). \tag{8.1}$$

Analoogiliselt, olgu ühtlane jaotus (5.1) lokaliseeritud lõigul $[x_0 - \Delta/2, x_0 + \Delta/2]$ (s.t $a = x_0 - \Delta/2$, $b = x_0 + \Delta/2$). Tähistame selle ühtlase jaotuse

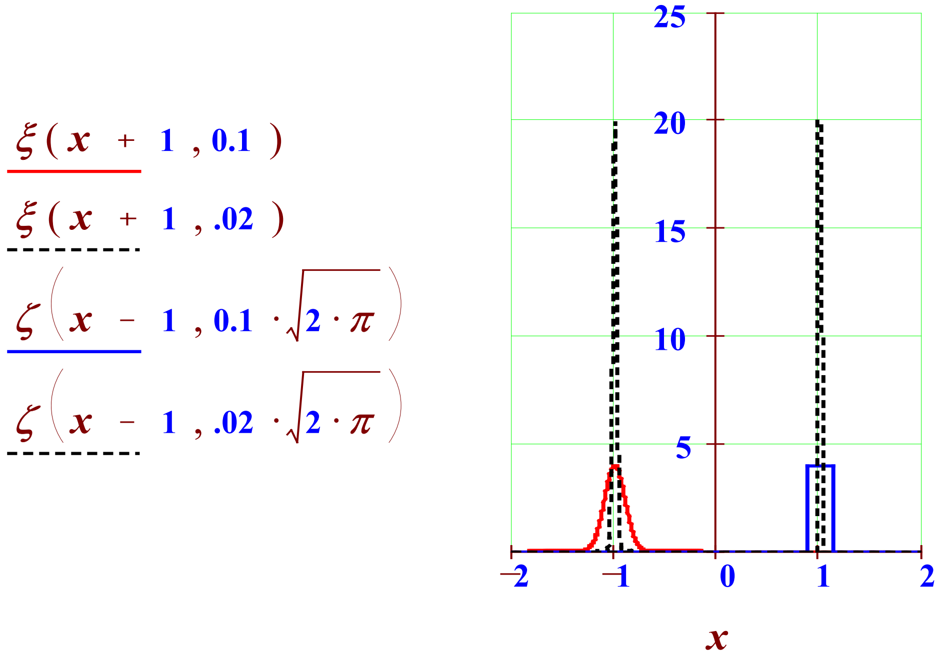
$$\varsigma(x - x_0;\Delta) = \begin{cases} 0, & |x - x_0| > \Delta/2 \\ \dfrac{1}{\Delta}, & |x - x_0| \le \Delta/2 \end{cases} .$$

Selle jaotuse abil saab delta-funktsiooni defineerida analoogiliselt eelnevaga kui piirväärtuse

$$\delta(x - x_0) \quad = \quad \lim_{\Delta\to 0} \xi(x - x_0;\Delta). \tag{8.2}$$

Joonis 8.1 demonstreerib, mil viisil kitsad jaotused ξ ja ς lähendavad delta-funktsiooni erinevate σ ja Δ väärtuste korral. ξ-jaotuse puhul on parameetri σ väärtusteks valitud $\sigma = 0.1$ ja $\sigma = 0.02$. ς-jaotuse parameetri Δ väärtused on võetud vastavalt võrdseks

$\Delta = \sigma\sqrt{2\pi}$, mis tagab ξ- ja ς-jaotuste võrdse kõrguse ja seega parema vastastikuse võrreldavuse. Lisaks on ξ-jaotuse tsenter x_0 paigutatud punkti $x_0 = -1$ ja ς-jaotuse tsenter vastavalt punkti $x_0 = 1$. On näha, et mida väiksem on σ (vastavalt Δ) väärtus, seda kitsam ja kõrgem on jaotustihedus.



Joonis 8.1.

Delta-funktsiooni omadused.

1. Delta-funktsioon $\delta(x - x_0)$ on tõenäosustihedus mittejuhuslikule suurusele x, mis võib omandada vaid fikseeritud väärtuse x_0, s.t

$$P(X = x_0) = 1 \quad .$$

Järeldus. Diskreetse juhusliku suuruse (vt punkt 2.2) $X : x_1, x_2, \ldots, x_n$, mille puhul

$$P(X = x_k) = p_k \quad (k = 1, 2, \ldots n), \text{ kusjuures } \sum_{k=1}^{n} p_k = 1,$$

tõenäosustihedus on esitatav delta-funktsioonide kaalutud summana, kus kaaludeks on diskreetse suuruse tõenäosused p_k:

$$f(x) \quad = \quad \sum_{k=1}^{n} p_k \delta(x - x_k) \quad .$$

2. Delta-funktsioon on normeeritud ühele:

$$\int_{-\infty}^{\infty} \delta(x - x_0) dx \quad = \quad 1 \qquad \text{iga } x_0 \text{ korral.}$$

See omadus tuleneb nende jaotusfunktsioonide normeeritusest, mille piiriks delta-funktsioon on.

3. Integraal delta-funktsioonist suvalisel lõplikul lõigul $[a,b]$:

$$\int_a^b \delta(x - x_0) dx \quad = \begin{cases} 1 & \text{kui} \quad a < x_0 < b, \\ \frac{1}{2} & \text{kui} \quad \mathrm{x}_0 = a \text{ või } b, \\ 0 \quad \text{kui} \quad x_0 < a \text{ või } x_0 > b. \end{cases} \tag{8.3}$$

Seejuures a ja b võivad paikneda teineteisele kuitahes lähedal, kuid alati peab intervalli pikkus erinema nullist, s.t peab kehtima $a < b$. Tõestada on seda omadust lihtne, kasutades jällegi nende lõpliku laiusega jaotustiheduste abi, mille piirväärtusena delta-jaotus on defineeritud. Võttes konkreetselt näiteks funktsiooni $\xi(x - x_0; \sigma)$, kehtib piirväärtus

$$\lim_{\sigma\to 0} \int_a^b \xi(x - x_0;\sigma) dx \quad = \begin{cases} 1 & \text{kui} \quad a < x_0 < b, \\ \frac{1}{2} & \text{kui} \quad \mathrm{x}_0 = a \text{ või } b, \\ 0 \quad \text{kui} \quad x_0 < a \text{ või } x_0 > b, \end{cases}$$

mis sisuliselt tähendabki omadust (8.3).

4. Integraal delta- funktsiooni ja sileda, aeglaselt muutuva funktsiooni $g(x)$ korrutisest. Kehtib:

$$\int_{-\infty}^{\infty} \delta(x - x_0)\, g(x)\, dx \quad = \quad g(x_0) \ . \tag{8.4}$$

Selle omaduse näitamiseks tuleb punkti x_0 ümbruses eraldada välja kuitahes väike, kuid lõplik piirkond $[x_0 - \varepsilon, x_0 + \varepsilon]$, $\quad \varepsilon > 0$. Väljaspool seda piirkonda on delta-funktsioon $\delta(x - x_0)$ võrdne nulliga, selle piirkonna sees aga on $g(x)$ kui aeglaselt muutuv funktsioon ligikaudu konstantne, seetõttu kehtib

$$\int_{-\infty}^{\infty} \delta(x - x_0)\, g(x)\, dx \quad = \int_{x_0-\varepsilon}^{x_0+\varepsilon} \delta(x - x_0)\, g(x) dx \approx g(x_0) \int_{x_0-\varepsilon}^{x_0+\varepsilon} \delta(x - x_0) \ \ dx = g(x_0)$$

.

Omaduse (8.4) kohta öeldakse: Delta-funktsioon määrab integraaloperaatori, mis seab suvalisele integreeruvale funktsioonile vastavusse tema väärtuse üksikus fikseeritud punktis.

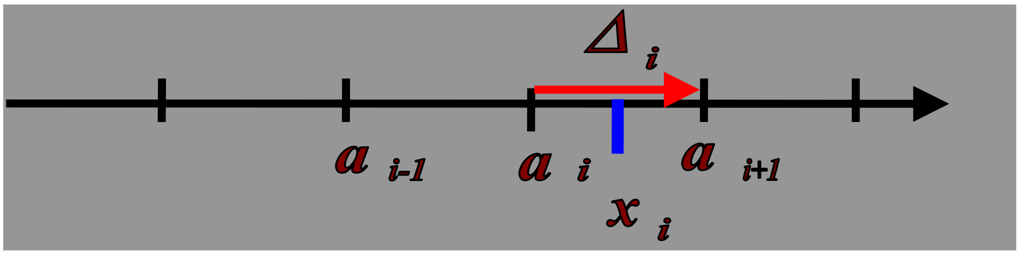
2.9. Tõenäosustiheduse hindamine eksperimendist

Paljudel juhtudel (võib öelda isegi et masendavalt määravas enamuses juhtudest) ei ole pideva juhusliku suuruse X jaotusseadus (s.o jaotusfunktsioon) teada aprioorsetest kaalutlustest ning see on vajalik leida katseliselt. Selleks tuleb läbi viia suur hulk N sõltumatuid juhusliku suuruse mõõtmisi, mille tulemusel saame juhusliku suuruse realisatsioonide jada

$$x_1,\ x_2, \ldots . x_{i-1}, x_i, x_{i+1, \ldots \ldots}, \ x_N \,. \tag{9.1}$$

Matemaatilises statistikas nimetatakse X kõikvõimalike väärtuste üldhulka ***üldkogumiks*** *(ingl. keeles* ***general ensemble****)*, mõõdetud lõpliku pikkusega väärtuste jada (9.1) aga nimetatakse ***valimiks.***

Pideva juhusliku suurusega eksperimenteerimisel on spetsiifiline, et kuigi suurus ise on pidev, on tema valim alati diskreetne kogum, ja selles mõttes analoogiline diskreetse juhusliku suuruse mõõtmisel saadud valimiga. Fundamentaalne erinevus on aga selles, et **pideva juhusliku suuruse üheski lõplikus valimis ei ole kaht võrdset suurust.** Teiste sõnadega, $P(x_i = x_j, \quad i \neq j) = 0$ alati. Seepärast peab pideva suuruse puhul jaotama X muutumispiirkonna väikesteks lõikudeks $\Delta_j = \lfloor a_j, a_{j+1}) = \{ a_j \leq X < a_{j+1} \}$ keskpunktide



Joonis 9.1

koordinaatidega $x_j = \dfrac{a_j + a_{j+1}}{2}$, nagu näidatud joonisel 9.1.

Iga poolintervalli Δ_j korral saame anda soodsate katsetulemuste arvu n_j, et katsetulemusel saadud juhusliku suuruse väärtus asub selles poolintervallis (joonis 9.2). Samuti saame arvutada eksperimentaalse sagedusjaotuse

$$p_j = \frac{n_j}{N}. \tag{9.2}$$

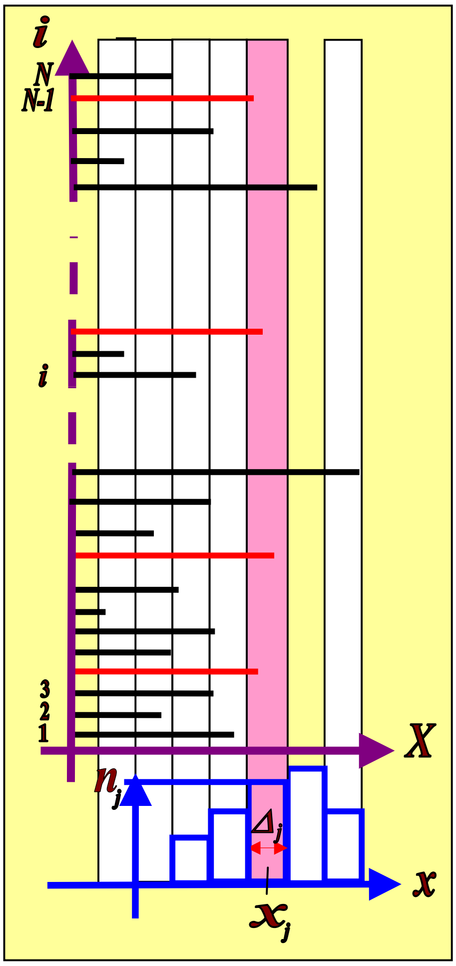
Kuna ilmselt alati kehtib

$$\sum_j n_j \quad = \quad N,$$

(grupeeritud katsetulemuste arvude summa peab võrduma katsetulemuste koguarvuga), siis

$$\sum_j p_j \quad = \quad 1.$$

Tulemused on esitatavad nn jaotustabelina (tabel 9.1).



Joonis 9.2

Pideva joonega on kujutatud histogrammile vastav ***polügoon*** – murdjoon, mis ühendab histogrammi tulpade tippude keskpunkte.

Me saame sisse tuua approksimeeriva jaotustiheduse $f_j^* = f^*(x_j)$, mille defineerime nii:

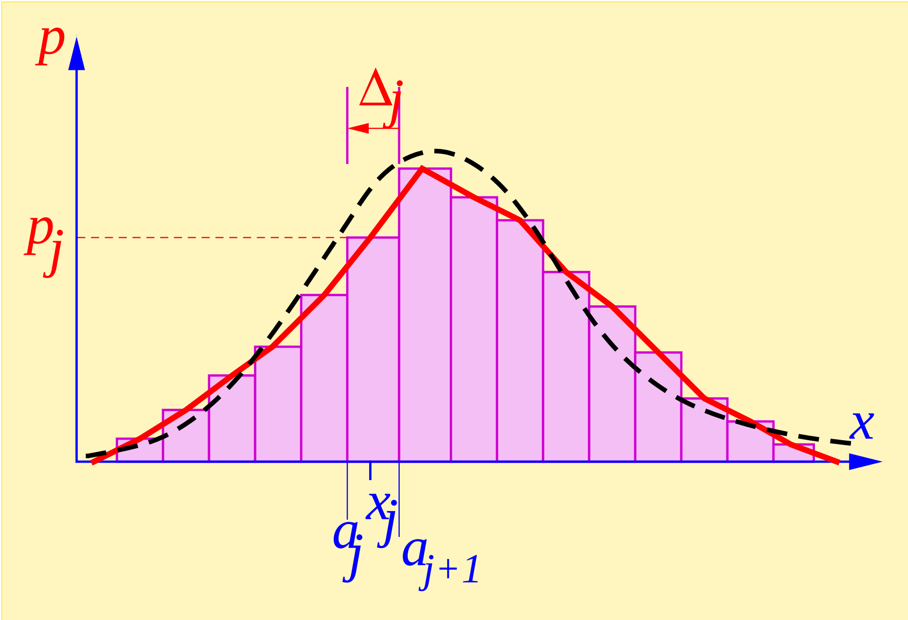
$$f^*(x_j) = \frac{p_j}{\Delta_j} = \frac{n_j}{N\Delta_j}.$$

Nüüd saame jaotustabelit täiendada f_j^* väärtuste reaga (viimane rida tabelis 9.1).

Tabel 9.1

j	1	2	...	j	...	M
x_j	x_1	x_2	...	x_j	...	x_M
Δ	Δ_1	Δ_2	...	Δ_j	...	Δ_M
n	n_1	n_2	...	n_j	...	n_M
p	p_1	p_2	...	p_j	...	p_M
f^*	f_1^*	f_2^*	...	f_j^*	...	f_M^*

Samuti saab tulemused esitada sagedusjaotuse histogrammina nagu kujutatud joonisel 9.3.



Joonis 9.3

Teeme hüpoteesi, et katsete arvu N lähenemisel lõpmatusse ja samaaegsel intervallide Δ_j pikkuste lähenemisel nullile eksisteerib piirväärtus

$$f(x) = \lim_{N\to\infty,\Delta_j\to 0,x_j\to x} f^*(x_j) = \lim_{N\to\infty,\Delta_j\to 0} \frac{n_j}{N\Delta_j}.$$

Seda ***piirväärtust nimetame juhusliku suuruse*** X ***tõenäosustiheduseks vaadeldavas punktis*** x. Joonisel 9.3 illustreerib katkendlik joon hüpoteetilist jaotustihedust, millele polügoon peaks/võiks läheneda piiril.

Piirile üleminek on võimalik vaid teoreetiliselt, praktikas tuleb alati piirduda lõpliku intervallide arvuga ja lõpliku katsete koguhulgaga. On selge, et mida suurem on katsete koguhulk ja mida detailsemalt on jaotatud arvtelg alalõikudeks, seda paremini approksimeerib f^* tõelist jaotustihedust f. Kuna siin on tegemist üheaegselt kahekordse piirprotsessiga – ühelt poolt peab intervallide arv lähenema lõpmatusele (kusjuures igaühe pikkus läheneb nullile), teisalt peab igasse alaintervalli Δ_j sattuvate sündmuste arv n_j samuti lähenema lõpmatusele – siis on tõenäosustiheduse eksperimentaalne hindamine väga suurt katsete arvu vajav, aeganõudev ja kallis ettevõtmine.

Praktikas ei ole võimalik kunagi sellist piirüleminekut teha, sest pole võimalik sooritada lõpmatut arvu katseid. Tekib küsimus, kuidas hinnata vajalikku minimaalset katsete arvu ja kuidas optimaalselt valida intervalle Δ_j ja nende arvu M. Vastused sõltuvad üldiselt vaadeldava juhusliku suuruse iseloomust. Põhjendatud vastuse neile küsimustele saab anda matemaatilise statistika vahenditega. Anname siinkohal lihtsa reegli Δ_j ja M valikuks. Olgu sooritatud katsete arv N ning olgu vähim ja suurim mõõdetud X realisatsioon vastavalt $x_{\min}$ ja $x_{\max}$. Valime intervallide arvuks

$$M \approx \sqrt{N} \qquad \text{(täisarvuks ümmardamise täpsusega),}$$

võtame kogu jaotuse paiknevaks lõigul $[x_{\min}, x_{\max}]$ ja kõik intervallid Δ_j võrdseks:

$$\Delta_j = \frac{x_{\max} - x_{\min}}{M}.$$

Sel juhul on lootust, et igasse intervalli sattuvate mõõtmiste hulk on $n_j \sim N/M \approx M$ (loe: on sama suurusjärku kui intervallide arv), mis tagab teatava optimaalsuse: kui oleks liiga palju intervalle, siis n_j oleks väikesed ja nende fluktuatsioonid ühest rakust teise oleksid suured. Valides aga intervallide arvu tunduvalt alla siintoodu, saaksime küll igal intervallil usaldusväärse sagedushinnangu, kuid jaotus (histogramm) ise oleks liiga jäme ja väheinformatiivne.

Näide 1. Olgu meil tehtud 100 mõõtmist, mis paiknevad mingil lõigul $[a,b] = [x_{\min}, x_{\max}]$, siis optimaalne intervallide arv oleks 10 ja igasse intervalli satuks suurusjärgult 10 üksikmõõtmist. Olgu nüüd aga sama juhusliku suurusega tehtud 100 korda rohkem mõõtmisi, s.o olgu mõõtmiste üldarv 10 000, nii et optimaalne intervallide arv oleks siin 100. Võime oletada, et paiknemine ja lõigu ulatus sellest ei muutu palju ja on lähedane $[a,b]$--le. Sel juhul tuleb siin üksikintervalli Δ_j pikkus ligikaudu 10 korda lühem kui eelmisel juhul, samal ajal kui mõõtmiste arv igal üksikintervallil n_j on suurenenud umbes kümme korda.

Näide 2. Silindri pikkuse jaotusfunktsiooni hindamine.

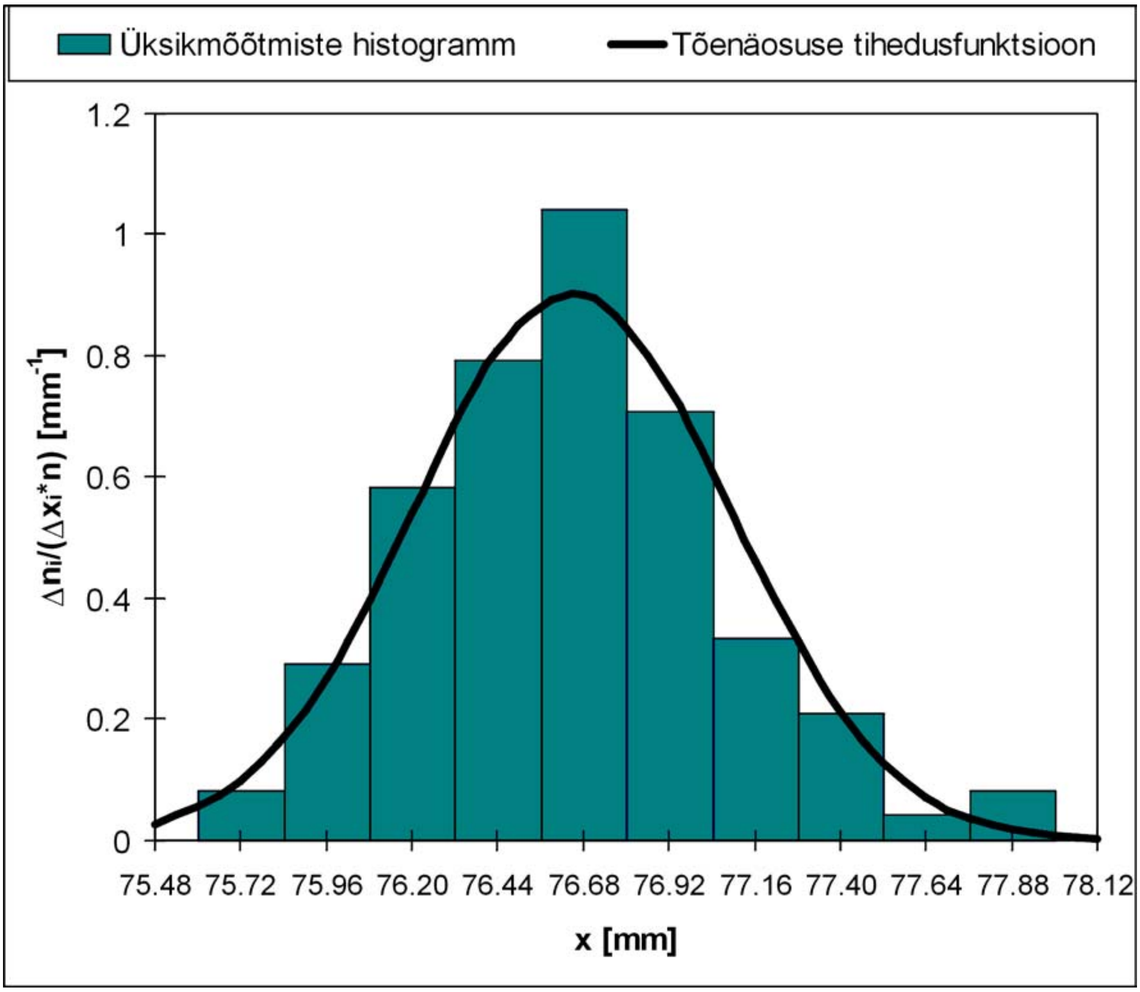
Vaatleme metallsilindri mõõtmist. Metallsilindri läbimõõt olgu 50 mm, aga see number 50 ei ole mõõtmiste seisukohast oluline. Oluline on, et silindri otsad on risti moodustajaga (lõigatud või treitud) ja me mõõdame silindri pikkust mitmest erinevast kohast.

Erinevatel mõõtmistel saame erinevad silindri pikkused. Oletame, et mõõtsime metallsilindri pikkust n =100 korda. Tulemuseks saime 100 lugemit, millest vähim oli $x_{\min}$ = 75,6 ja suurim $x_{\max}$ = 78,0. Mõõtmistulemuste diapasooni $[x_{\min}, x_{\max}]$ jagame $M = \sqrt{100} = 10$ võrdseks vahemikuks $\Delta_i = \Delta = 0{,}24\ mm$, seejärel loendame, mitu korda mõõtmistulemus igasse vahemikku sattus.

Tabel 9.3. Jaotustabel histogrammi joonistamiseks

j	1	2	3	4	5	6	7	8	9	10
x_j [mm]	75,72	75,96	76,20	76,44	76,68	76,92	77,16	77,40	77,64	77,88
n	2	7	14	19	25	17	8	5	1	2
f^* [1/mm]	0,083	0,29	0,58	0,79	1,04	0,71	0,33	0,21	0.042	0,083

Tulemused on esitatud **tabelis 9.3**. Tabeli alusel joonistame histogrammi (joonis 9.4). Pideva joonega on joonisel näidatud histogrammi kõige täpsemalt approksimeeriv normaaljaotus. Histogramm kannab olulist infot, näidates et ***mõõtmistulemused on ligikaudu normaalselt jaotunud.*** Siin kohtume esmakordselt jaotusfunktsiooni dimensiooniga. Seni vaatlesime juhuslikke suurusi just nagu dimensioonitutena (tegelikult ei pööranud sellele tähelepanu). Tegelikult on füüsikas ja keemias enamik mõõdetavaid suurusi **dimensiooniga suurused**. ***Jaotustiheduse dimensioon on*** alati ***juhusliku suuruse dimensiooni pöördväärtus***: Kui X dimensioon on $\dim X$, siis $f(x)$ dimensioon on $1/(\dim X)$. See järeldub asjaolust, et tõenäosus juhusliku suuruse sattumiseks elementaarintervalli dx: $dP = f(x)dx$ on dimensioonitu suurus. Lähemalt peatume sellel küsimusel veel füüsikaliste mõõtmiste peatükis. Käesoleval juhul on mõõdetava suuruse dimensioon pikkus väljendatuna *millimeetrites* ja vastavalt on tõenäosustiheduse dimensioon pöördpikkus *pöördmillimeetrites*.



Joonis 9.4. Silindri pikkuste jaotuse histogramm (tulbad) ja seda approksimeeriv jaotusfunktsioon.

Probleem mõtlemiseks: Kuidas valida intervall $[x_{\min}, x_{\max}]$, kui kõik mõõtmistulemused paiknevad lõigul $[a,b]$, välja arvatud üks mõõtmine, mis paikneb kaugel väljaspool seda lõiku?

3. ÕPPEKIRJANDUST TÕENÄOSUSTEOORIA KOHTA

1. Tiit E., 1968: Tõenäosusteooria I. Loengukonspekt. Tartu, TRÜ Rotaprint, 319 lk.
2. Tiit E., A. Parring, T. Möls, 1977: Tõenäosusteooria ja matemaatiline statistika [õpik TRÜ matemaatikateaduskonnas]. Tallinn, Valgus, 470 lk.
3. Gurski J, 1986: Tõenäosusteooria ja matemaatilise statistika elemendid. Tallinn, Valgus, 311 lk.
4. Jõgi A, 2000: Tõenäosusteooria, 1. osa. TTÜ Kirjastus, Tallinn, lk 1–198.
5. Jõgi A, 2000: Tõenäosusteooria, 2. osa. TTÜ Kirjastus, Tallinn, lk 199–416.