

TÕENÄOSUSTEORIA JA MATEMAATILINE STATISTIKA

Tõenäosusteooria uurib massiliselt toimuvates juhuslikes nähtustes esinevaid üldisi seaduspärasusi sõltumatult nende nähtuste konkreetsest olemusest.

Statistika on teadus, mis käsitleb katse- või vaatlusandmete kogumise, klassifitseerimise ja **oluliste** karakteristikute eraldamise meetodeid.

SÜNDMUSE TÕENÄOSUS JA SELLE LEIDMINE

1. ELEMENTAARSÜNDMUSTE RUUM. JUHUSLIKUD SÜNDMUSED

Tõenäosusteoorias mõistetakse katse all teatud tingimuste kompleksi realiseerumist, mille tulemusena võivad toimuda mingid sündmused.

Definitsioon 1.1: Juhuslik katse on igasugune tegevus, mille tulemus ei ole antud tingimustes üheselt määratud.

Näiteks märgi tabamine tulistamisel, mille puhul katseks on tulistamine ja sündmuseks märgi tabamine.

Iga katse korral võib leida mingi üksteist välistavate sündmuste (vt. käesolevat paragrahvi edasi) hulga nii, et katse tulemusena tingimata toimub üks neist sündmustest. Niisugust hulka nimetatakse antud katsega seotud elementaarsündmuste ruumiks (tähis Ω) ja ruumi kuuluvaid sündmusi elementaarsündmusteks (tähis ω). Elementaarsündmusi ω nimetatakse ruumi Ω *elementideks*.

Toome täpsema elementaarsündmuste definitsiooni

Definitsioon 1.2: Juhusliku katse võimalikke tulemusi nimetatakse elementaarsündmusteks.

Elementaarsündmuse kohta võib ka öelda, et ta on niisugune sündmus, mis osasündmusteks enam ei lahutu. Elementaarsündmuses sõltuvad sellest, mida katse tulemusena „kirja pannakse“. Seega võib sama reaalse katse korral vaadelda erinevaid elementaarteisenduste ruume.

Näide 1.1: Mündiviise. Tavaliselt vaadeldakse $\Omega = \{vapp, kiri\}$. Sõltuvalt katse sooritamise viisist, eesmärgist ja kasutatavast mündist võib mõnikord olla vajalik ka vaadelda olukorda, kus võimalikeks elementaarsündmusteks on $\Omega' = \{vapp, kiri, serva peal\}$ või $\Omega'' = \{lapiti, serva peal\}$. Viimasel juhul huvitab meid ainult see, kas münt jääb serva peale seisma või mitte.

Näide 1.2: Katse seisneb täringu viskamises. Katse tulemuseks saame teatud silmade arvu. Sel juhul koosneb elementaarsündmuste ruum kuuest elemendist (elementaarsündmusest): ω_1 – täringul tuleb üks silm, ω_2 – kaks silma, ω_3 – kolm silma, ω_4 – neli silma, ω_5 – viis silma, ω_6 – kuus silma. Selles näites on elementaarsündmuste ruum *lõplik hulk*.

Näide 1.3: Märki tulistatakse üksiklaskudega kuni esimese tabamiseni. Võimalikud on järgmised elementaarsündmused: ω_1 – tabamine esimesel tulistamisel, ω_2 – möödalaskmine esimesel ja tabamine teisel tulistamisel, ω_3 – möödalaskmine esimesel ja teisel tulistamisel ning tabamine kolmandal tulistamisel jne. Siin koosneb elementaarsündmuste ruum lõpmatust arvust elementidest, neid kõiki võib nummerdada, s.t. elementaarsündmuste ruum on selles näites *loenduv hulk*.

Näide 1.4: Juhuslikult visatakse kuuli maha joonistatud ringi. Eeldame, et kuul sattub igal viskel ringi. Ringi igale punktile M seame vastavusse katsetulemuse $\omega(M)$ – kuuli puutepunkti ringiga. Katsetulemuste $\omega(M)$ hulk moodustab elementaarsündmuste ruumi, mis sisaldab mitteloenduva hulga elementaarsündmusi ω . Niisugust ruumi nimetatakse *pidevaks*.

Näide 1.5: Kardipakk peale kaartide segamist. Olgu pakis on K kaarti. Kaardipakki võib seega vaadelda kaartide hulga $\{1, \dots, K\}$ järjestusena. Seega iga elementaarsündmus on hulga $\{1, \dots, K\}$ permutatsioon, elementaarsündmuste ruumi võimsus on siis $K!$.

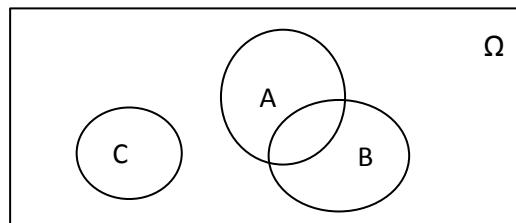
Olgu Ω mingi katsega seotud elementaarsündmuste ω ruum. Ruumi Ω elementaarsündmuste igale alamhulgale vastab nn. juhuslik sündmus, mille toimumine katse tulemusena seisneb sellesse alamhulka kuuluva mingi elementaarsündmuse ω toimumises. Juhuslikke sündmusi tähistatakse ladina tähestiku suurtähtedega.

Seega võib tõenäosusteoorias igat katset kirjeldada elementaarsündmuste hulgaga Ω ja selle hulga alamhulkade mingi klassiga (klass on hulk, mille elementideks on hulgad) $\mathcal{U}$. Neid alamhulki nimetataksegi juhuslikeks sündmusteks.

Elementaarsündmusi, mis kuuluvad juhuslikku sündmusele A , nimetatakse sündmuse A toimumise suhtes **soodsateks sündmusteks**. See tähendab, et katse tulemusena juhuslik sündmus A toimub, kui leiab aset mingi sündmus vaadeldava sündmuse suhtes soodsate sündmuste hulgast ja ei toimu vastupidisel juhul.

Iga elementaarsündmus on samuti juhuslik sündmus. Kogu ruum Ω on sündmus, mis katse tulemusena tingimata toimub. Niisugust sündmust nimetatakse **kindlaks sündmuseks**.

Juhuslikke sündmusi kujutatakse geomeetriselt punktihulkadena piirkonnas Ω . Kindlale sündmusele vastab kogu piirkond Ω .



Definitsioon 1.3: Kaht või enamat juhuslikku sündmust nimetatakse võrdvõimalikeks, kui tingimused nende sündmuste toimumiseks on ühesugused.

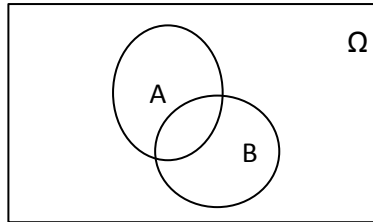
Definitsioon 1.4: Kaht sündmust nimetatakse võrdseteks (samaväärseteks) ja kirjutatakse

$$A = B,$$

kui need sündmused koosnevad ruumi Ω kuuluvatest samadest elementaarsündmustest.

Näiteks mingi seade koosneb n jadaühendusega blokist. Mingi bloki riknemine tähendab kogu seadme riknemist. Sündmus A – vähemalt ühe bloki riknemine, ja sündmus B – seadme riknemine, on võrdsed.

Definitsioon 1.5: Kaht sündmust A ja B nimetatakse mittevälstavateks sündmusteks, kui ühe sündmuse toimumine ei välista teise sündmuse toimumist. Ehk teisiti, sündmusi A ja B nimetatakse mittevälstavateks, kui leiduvad elementaarsündmused, mis kuuluvad üheaegselt nii sündmusele A kui ka sündmusele B .

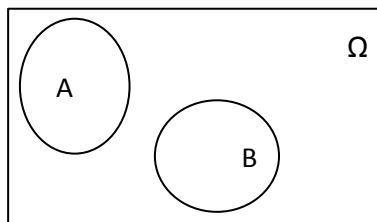


Näiteks kahe täringu viskamisel olgu sündmus A ühel täringul kolme silma esiletulek ja sündmus B teisel täringul kolme silma esiletulek. Sel juhul on sündmused A ja B mittevälstavavad.

Lause 1.1: Sündmused moodustavad mittevälstavate sündmuste süsteemi, kui vähemalt kaks sündmust sellest süsteemist on mittevälstavavad.

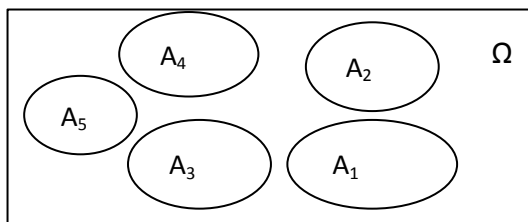
Näiteks tulistatakse kolm korda märki. Olgu sündmus A_1 tabamine esimesel tulistamisel, A_2 tabamine teisel tulistamisel, A_3 tabamine kolmandal tulistamisel. Sündmused A_1, A_2, A_3 moodustavad mittevälstavate sündmuste süsteemi.

Definitsioon 1.6: Kaht sündmust A ja B nimetatakse teineteist välistavateks, kui ühe sündmuse toimumine välistab teise sündmuse toimumise. Ehk teisiti, sündmusi A ja B nimetatakse teineteist välistavateks, kui ei leidu ühtegi elementaarsündmust, mis kuuluks üheaegselt nii sündmusele A kui ka B .



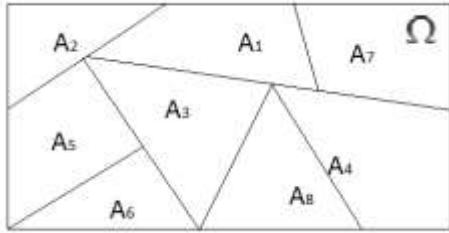
Näiteks saabus kauplusesse partii ühesuguse tegumoe ja suurusega, kuid erinevat värvi kingi. Sündmus A – juhuslikult võetud karbis on mustad kingad, sündmus B – karbis on valged kingad. Sündmused A ja B on teineteist välistavad sündmused.

Definitsioon 1.7: Sündmuste süsteemi $A_1, A_2, \dots, A_n$ nimetatakse üksteist välistavate sündmuste süsteemiks, kui iga kaks süsteemi kuuluvat sündmust on teineteist välistavad sündmused.



Definitsioon 1.8: Üksteist välistavate sündmuste süsteemi $A_1, A_2, \dots, A_n$ nimetatakse täissüsteemiks (täielikuks süsteemiks), kui katse tulemusel toimub tingimata üks ja ainult üks nendest sündmustest.

Ilmselt peab ruumi Ω iga elementaarsündmus ω kuuluma ainult ühte täissüsteemi sündmustest $A_1, A_2, \dots, A_n$.

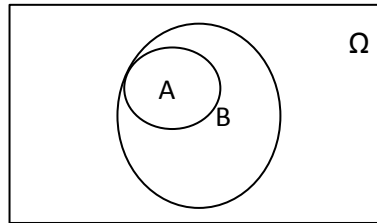


Näiteks tulistatakse kolm korda märki. Olgu sündmus A – möödalask, B – üks tabamine, C – kaks tabamist, D – kolm tabamist. Sündmused A, B, C, D moodustavad täissüsteemi.

Antud katsega seotud kõigi elementaarsündmuste hulk moodustab alati sündmuste täissüsteemi.

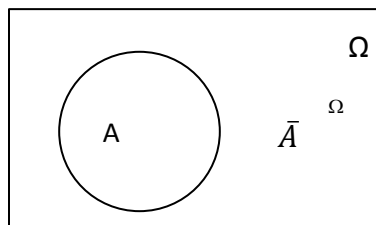
Definitsioon 1.9: Sündmust B nimetatakse sündmuse A järelduks, kui sündmuse A toimumisest tuleneb sündmuse B toimumine.

Hulgateoreetilises sõnastuses tähendab see, et hulk A on hulga B alamhulk: $A \subset B$.



Näiteks tähendagu katse tulemus kahe täringu viskamisel saadavat silmade arvu. Olgu sündmus B – silmade arv on paarisarv, sündmus A – silmade arv jagub neljaga. Sündmus B on sündmuse A järelmus.

Lause 1.2: Igale sündmusele A vastab selle vastandsündmus $\bar{A}$, mille toimumine tähendab sündmuse A mittetoimumist. Sündmus $\bar{A}$ koosneb ruumi Ω kõigist nendest elementaarsündmustest, mis ei kuulu sündmuse A koosseisu.



Näiteks tulistatakse märki kolm korda. Sündmus A – vähemalt üks tabamine, ja sündmus $\bar{A}$ – kõik kolm möödalaskmist, on vastandsündmused.

Vastandsündmuste A ja $\bar{A}$ paar on kahest sündmusest koosneva täissüsteemi lihtsaim juht.

Definitsioon 1.10: Kindla sündmuse Ω vastandsündmust $\bar{\Omega}$ nimetatakse võimatuks sündmuseks. Ehk teisiti, sündmust nimetatakse võimatuks, kui ta ei saa katse tulemusel toimuda. Võimatu sündmus ei sisalda ühtegi elementaarsündmust. Sellele vastab nn. tühi hulk $\emptyset$, mis ei sisalda ühtegi elementi.

2. TEHTED SÜNDMUSTEGA. SÜNDMUSTE HULGAALGEBRA

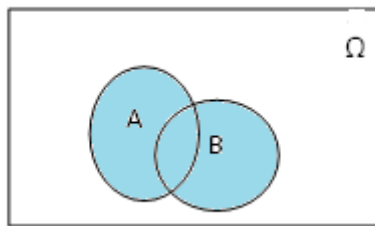
Sündmused on hulgad, seega saab neile rakendada ka hulgateoreetilisi tehteid: ühend, ühisosa, täiend jne. Järgnevad definitsioonid annavad neile tehetele tõenäosusteoreetilise interpretatsiooni.

Definitsioon 2.1: Sündmuste A ja B summaks nimetatakse sündmust C , mille puhul toimub kas sündmus A või sündmus B või mõlemad ("A **või** B **või** mõlemad").

Sündmuste summat tähistatakse

$$C = A + B \text{ või } C = A \cup B.$$

Sündmus $A + B$ saadakse seega sündmustesse A ja B kuuluvate elementaarsündmuste ühendamisel.



Märkus: Üldiselt sündmuste $A_1, A_2, \dots, A_n$ summaks nimetatakse sündmust C , mis seisneb vähemalt ühe sündmuse toimumises nendest sündmustest. Summat tähistatakse

$$C = \sum_{k=1}^n A_k \text{ või } C = \bigcup_{k=1}^n A_k.$$

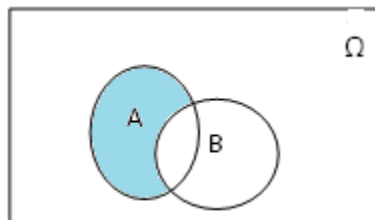
Lause 2.1: Kui sündmused $A_1, A_2, \dots, A_n$ moodustavad täissüsteemi, siis nende sündmuste summa on kindel sündmus, s.o.

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

Definitsioon 2.2: Sündmuste A ja B **vaheks** nimetatakse sündmust, kui toimub sündmus A ja ei toimu sündmus B .

Sündmuste vahet tähistatakse

$$C = A - B \text{ või } C = A \setminus B.$$



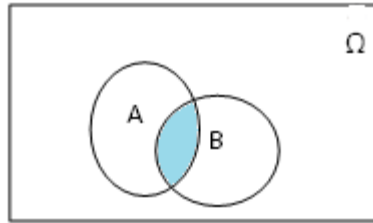
Näitena vaatleme kahe täringu viskamist. Olgu sündmus A – kahe täringuga saadavate silmade summa on paarisarv, sündmus B – kummalgi täringul on paarisarv silmi. Nüüd tähendab sündmus $A - B$, et kummalgi täringul on paaritu arv silmi.

Definitsioon 2.3: Sündmuste A ja B korrutiseks nimetatakse sündmust C , mille puhul toimub nii sündmus A kui ka sündmus B ("A ja B").

Sündmuste korrutist tähistatakse

$$C = AB \text{ või } C = A \cap B.$$

Seega koosneb sündmus AB nii sündmusele A kui ka sündmusele B kuuluvatest elementaarsündmustest.



Märkus: Üldiselt sündmuste $A_1, A_2, \dots, A_n$ korrutiseks nimetatakse sündmust C , mis seisneb kõigi nende sündmuste toimumises. Korrutist tähistatakse

$$C = \prod_{k=1}^n A_k \text{ või } C = \bigcap_{k=1}^n A_k.$$

Lause 2.2: Kui sündmused $A_1, A_2, \dots, A_n$ moodustavad täissüsteemi, siis iga suvalise sündmuse B puhul

$$B = \sum_{k=1}^n BA_k.$$

Suvaliste sündmuste A ja B puhul kehtivad valemid

$$\overline{A+B} = \bar{A}\bar{B}, \quad \overline{AB} = \bar{A} + \bar{B},$$

mida nimetatakse duaalsusseosteks ja mis väljendavad nn. duaalsusprintsipi, mille järgi vastandsündmusele üle minnes asendub sündmuste summa nende korrutisega ja vastupidi.

Duaalsusseosed võib üldistada suvalise lõpliku arvu sündmuste $A_1, A_2, \dots, A_n$ juhule:

$$\overline{\sum_{k=1}^n A_k} = \prod_{k=1}^n \bar{A}_k, \quad \overline{\prod_{k=1}^n A_k} = \sum_{k=1}^n \bar{A}_k.$$

Kui ruum Ω on pidev, siis on selle ruumi kõigi võimalike alamhulkade klass tõkestamata. Siis piisab praktikas vaadelda ruumi Ω alamhulkade kitsamat klassi U , mis rahuldab järgmisi tingimusi:

- 1) $\Omega \in U$ (Ω on kindel sündmus),
- 2) $\emptyset \in U$ ($\emptyset$ on võimatu sündmus),
- 3) kui $A \in U$, siis ka $\bar{A} = \Omega \setminus A \in U$,
- 4) kui $A \in U, B \in U$, siis ka $A \cup B \in U, A \cap B \in U$.

Definitsioon 2.4: Alamhulkade klassi U , mis rahuldab tingimusi 1) – 4), nimetatakse hulgaalgebraks. Juhul, kui Ω on lõplik, on sündmuste hulk U samane Ω kõigi alamhulkade klassiga.

Definitsioon 2.5: Mittetühja sündmuste hulka U , mis rahuldab tingimusi:

- 1) kui $A \in U$, siis ka $\bar{A} \in U$,
- 2) kui $A \in U, B \in U$, siis ka $A \cup B \in U$

nimetatakse sündmuste hulgaalgebraks.

Tihti tehakse juhuslik katse, mille käigus mõõdetakse mingit suurust X . Elementaarsündmusteks on sündmused $X = x$, kus x on mingi fikseeritud väärtus. Siin on loomulik samastada elementaarsündmuste ruumi Ω sirge

punkti hulga. Kui on teada, et suurus X võib omandada väärtusi ainult mingitest hulkadest, siis neid hulki tulebki vaadelda elementaarsündmuste ruumina Ω . Mõõtmisprotsessis on loomulik jälgida niisuguseid juhuslikke sündmusi $a \leq X < b$, kus $a < b$ (a ja b on suvalised reaalarvud, ka lõpmatud). Kirjeldatud poolloikude kõikvõimalikke lõplikke summasid võib vaadelda antud katsega seotud sündmuste hulgaalgebrana.

3. SÜNDMUSE SAGEDUS JA SELLE OMADUSED

Olgu tegemist katsete seeriaga, mis koosneb n katsest. Igal katsel võib sündmus A toimuda või mitte.

Definitsioon 3.1: Sündmuse A sageduseks antud katsete seerias nimetatakse sündmuse A esinemiste arvu ja kõigi katsete arvu suhet.

Sündmuse A sagedust tähistatakse $P^*(A)$. Seega

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

kus m on sündmuse A esinemiste arv ja n on kõigi katsete arv.

Sageduse omadused

- 1) $0 \leq P^*(A) \leq 1$ (sündmus A võib n katse puhul toimuda 0 kuni n korda, s.t. $0 \leq m \leq n$).
- 2) $P^*(\Omega) = 1$ (kindla sündmuse sagedus on 1).
- 3) $P^*(\emptyset) = 0$ (võimatu sündmuse sagedus on null).

Kui sündmused A ja B on teineteist välistavad, siis

$$4) P^*(A+B) = P^*(A) + P^*(B).$$

Tõestus: Olgu $P^*(A) = \frac{m}{n}$ ja $P^*(B) = \frac{k}{n}$. Kuna A ja B on teineteist välistavad sündmused, siis puuduvad katsed, mille korral sündmused A ja B oleks toimunud samaaegselt. Seega saab sündmuste summa definitsiooni põhjal öelda, et sündmus $A+B$ toimus $m+k$ korda ja järelikult

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

Definitsioon 3.2: Sündmuse A sagedust tingimusel, et sündmus B toimus, nimetatakse sündmuse A tinglikuks sageduseks $P^*(A|B)$.

Sündmuse A tinglik sagedus avaldub kujul

$$P^*(A|B) = \frac{P^*(AB)}{P^*(B)}. \quad (3.1)$$

Tõestus. Olgu $P^*(B) = \frac{m}{n}$ ja $P^*(AB) = \frac{k}{n}$. Kuna sündmus B toimus m katses ja omakorda nendest m katsest k katses toimus sündmus B koos sündmusega A , siis $P^*(A|B) = \frac{k}{m}$. Asetades eelnevad sageduste väärtused

seosesse (1), saame samasuse $\frac{k}{m} = \frac{\frac{k}{n}}{\frac{m}{n}}$.

4. SÜNDMUSE TÕENÄOSUS

Sündmuse sagedus on määratav vaid pärast katseid. Kuna eri katseseeriade puhul saadakse erinevaid tulemusi, siis ei ole sagedus sündmuse kõige paremaks iseloomustajaks. Kuid võib tähele panna, et pikas katseseerias sagedus stabiliseerub, kõikudes mingi konstantse väärtuse p ümber. Nii jõuame tõenäosuse statistilise definitsioonini:

Definitsioon 4.1: Juhusliku sündmuse tõenäosuseks nimetatakse konstanti, mille ümber grupeeruvad selle sündmuse sagedused katsete arvu suurenedes.

Eelnev definitsioon iseloomustab küll tõenäosuse mõistet, kuid ei anna võtteid tõenäosuse arvutamiseks. Seda võimaldab teha tõenäosuse klassikaline definitsioon (e. Laplace'i mudel):

Definitsioon 4.2: Sündmuse A tõenäosuseks nimetatakse suurust

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

kus n on kõikide võrdvõimalike elementaarsündmuste arv ja m on sündmuses A sisalduvate elementaarsündmuste arv (m on n -ö. soodsate elementaarsündmuste arv).

Tuletame meelde, et elementaarsündmus on niisugune sündmus, mis osasündmusteks enam ei lahutu. Elementaarsündmused on võrdvõimalikud, kui sisulistel kaalutlustel ei saa lugeda ühe elementaarsündmuse toimumise võimalikkust erinevaks teise elementaarsündmuse toimumise võimalikkusest. Klassikaline definitsioon võimaldab tõenäosust arvutada ilma katseid tegemata. Selle definitsiooni puuduseks on elementaarsündmuste võrdvõimalikkuse nõue.

5. ÜKSTEIST VÄLISTAVATE SÜNDMUSTE LIITMISTEOREEM

Tõenäosuse statistilise definitsiooni alusel ja lähtudes sündmuse sageduse omadustest, tuuakse sisse viis tõenäosusteooria aksioomi (neist viimane järgmises punktis). Niisuguse idee põhjal ülesehitatud tõenäosusteooria aksiomaatika tagab teeria kooskõla katsetega.

Aksioom 5.1: $0 \leq P(A) \leq 1$.

Aksioom 5.2: $P(\Omega) = 1$.

Aksioom 5.3: $P(\emptyset) = 0$.

Aksioom 5.4: Kui sündmused A ja B on **teineteist välistavad**, siis nende summa tõenäosus avaldub kujul

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

Viimase üldistuseks on **üksteist välistavate sündmuste liitmisteoreem:**

$$P(A_1 + A_2 + \dots + A_n) = P(A_1) + P(A_2) + \dots + P(A_n).$$

Tuletame meelde, et üksteist välistavad sündmused $A_1, A_2, \dots, A_n$ moodustavad **täieliku süsteemi**, kui katse tulemusena toimub neist tingimata üks ja ainult üks. Seega täieliku süsteemi korral $P(A_1 + A_2 + \dots + A_n) = 1$.

Järeldus 5.1: Kui üksteist välistavad sündmused $A_1, A_2, \dots, A_n$ moodustavad täieliku sündmuste süsteemi, siis vastavalt eelnevale valemile ning liitmisteoreemile

$$P(A_1) + P(A_2) + \dots + P(A_n) = 1.$$

Järeldus 5.2: Kuna vastandsündmused moodustavad täieliku süsteemi, siis on nende tõenäosuste summa võrdne ühega:

$$P(A) + P(\bar{A}) = 1 \quad \text{ja} \quad P(A) = 1 - P(\bar{A}).$$

6. KORRUTAMISTEOREEM

Defineerime tingliku tõenäosuse analoogiliselt sellele, kuidas oli defineeritud ühe sündmuse tinglik sagedus teise sündmuse toimumisel.

Definitsioon 6.1: Sündmuse A tõenäosust, mis leitakse tingimusel, et toimub teine sündmus B , nimetatakse sündmuse A tinglikuks tõenäosuseks sündmuse B suhtes ja tähistatakse $P(A|B)$.

Siinjuures eeldatakse, et $P(B) \neq 0$.

Toetudes punkti 3 valemile (3.1), tuuakse sisse tõenäosusteooria viies aksioom:

$$P(AB) = P(A) \cdot P(B|A) = P(B) \cdot P(A|B). \quad (6.1)$$

Viienda aksioomi üldistuseks n sündmusele on **korrutamisteoreem**:

$$P(A_1 A_2 \dots A_n) = P(A_1) \cdot P(A_2|A_1) \cdot \dots \cdot P(A_n|A_1 A_2 \dots A_{n-1}).$$

Definitsioon 6.2: Kui $P(A|B) = P(A)$ ja $P(B|A) = P(B)$ ehk sündmuse B toimumine ei muuda sündmuse A tõenäosust ja vastupidi, siis neid kaht sündmust nimetatakse teineteisest sõltumatuteks.

Sõltumatute sündmuste A ja B korral on viies aksioom esitatav kujul

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

Näide 6.1: Vaatlused on näidanud, et iga-aastaselt maratonijooksul läbib 90% osavõtjatest $\frac{3}{4}$ distantist ja 80% kogu distantsti. Leida tõenäosus, et $\frac{3}{4}$ distantist läbinud sportlane lõpetab jooksu.

Lahendus.

Sündmus A : sportlane on läbinud $\frac{3}{4}$ distantist; $P(A) = 0,9$.

Sündmus B : sportlane lõpetab jooksu.

Sündmus AB : sportlane on läbinud $\frac{3}{4}$ distantist ja lõpetab jooksu; $P(AB) = 0,8$.

Sündmus $B|A$: sportlane lõpetab jookse tingimusel, et ta on läbinud $\frac{3}{4}$ distantist.

Valemist (6.1) saab:

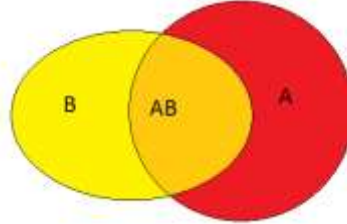
$$P(B|A) = \frac{P(AB)}{P(A)} = \frac{0,8}{0,9} \approx 0,89.$$

7. ÜKSTEIST MITTEVÄLISTAVATE SÜNDMUSTE SUMMA TÕENÄOSUS

Teoreem 7.1: Kahe teineteist mittevälistava sündmuse summa tõenäosuse kohta kehtib järgmine:

$$P(A+B) = P(A) + P(B) - P(AB). \quad (7.1)$$

Teoreemi geomeetrilist tähendust võib illustreerida järgmiselt:



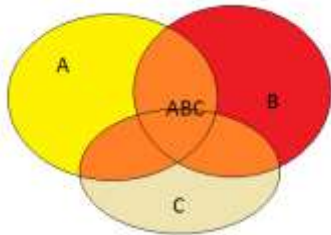
Joonisel olev oranž osa tuleb arvesse võtta üks kord. Tõlgendades hulkasid A ja B piltlikult märklaudadena, oleks ühisosa kohal muidu kahekordne märklaud.

Vastandsündmusele üle minnes võib valemi (7.1) asemel kasutada järgmist valemit:

$$P(A+B) = 1 - P(\bar{A}\bar{B}).$$

Kolme üksteist mittevälistava sündmuse puhul

$$P(A+B+C) = P(A) + P(B) + P(C) - P(AB) - P(AC) - P(BC) + P(ABC).$$



Vastandsündmusele üleminek nõuab aga ilmselt väiksemaid arvutusi:

$$P(A+B+C) = 1 - P(\bar{A}\bar{B}\bar{C}).$$

8. TÄISTÕENÄOSUSE VALEM

Moodustugu üksteist välistavatest elementaarsündmustest $H_1, H_2, \dots, H_n$ täielik süsteem. Nimetame neid sündmusi hüpoteesideks. Toimugu meid huvitav sündmus A vaid ühega hüpoteesidest:

$$A = AH_1 + AH_2 + \dots + AH_n.$$

Selles seoses olevad liidetavad on üksteist välistavad sündmused, kuna hüpoteesid on üksteist välistavad. Seega võib kasutada üksteist välistavate sündmuste tõenäosuste liitmisteoreemi:

$$P(A) = \sum_{i=1}^n P(AH_i).$$

Korrutamisteoreemi abil

$$P(AH_i) = P(H_i) \cdot P(A|H_i)$$

ja

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

Viimane ongi **täistõenäosuse valem**.

9. BAYESI VALEM (HÜPOTEESI TINGLIK TÕENÄOSUS)

Tuletame valemi hüpoteesi H_i tingliku tõenäosuse $P(H_i|A)$ arvutamiseks.

Olgu antud hüpoteeside täielik süsteem $H_1, H_2, \dots, H_n$ ning olgu teada nende hüpoteeside tõenäosused $P(H_1), P(H_2), \dots, P(H_n)$. Tehakse katse, mille tulemuseks on mingi sündmus A , mille tinglikud tõenäosused $P(A|H_1), P(A|H_2), \dots, P(A|H_n)$ olgu teada. Siis avaldub hüpoteesi H_i tinglik tõenäosus järgneva (**Bayesi**) valemi kujul:

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

Tõestus. Vastavalt korrutamisteoreemile

$$P(AH_i) = P(A) \cdot P(H_i|A) = P(H_i) \cdot P(A|H_i),$$

millest

$$P(H_i|A) = \frac{P(H_i) \cdot P(A|H_i)}{P(A)}.$$

Avaldades $P(A)$ täistõenäosuse valemi abil, saabki Bayesi valemi.

Bayesi valemi kasutamist nõudvates ülesannetes on iseloomulik, et kõigepealt püstitatakse hüpoteesid. Seejärel tehakse katse, mille tulemus on teada. Sõltuvalt katse tulemusest hinnatakse hüpoteeside esialgsed tõenäosused ümber.

10. BERNOULLI VALEM

Püstitame järgmise ülesande. Leida tõenäosus, et n ühesuguse ja sõltumatu katse tulemusel toimuks sündmus A täpselt m korda. Olgu igal üksikkatsel sündmuse A toimumise tõenäosus konstantne: $P(A) = p$. Vastandsündmuse $\bar{A}$ toimumise tõenäosus on siis $P(\bar{A}) = 1 - p = q$. Tähistame otsitava tõenäosuse $P_{m,n}$. Tõenäosuse $P_{m,n}$ arvutamine oleks liitmis- ja korrutamisteoreemi abil võimalik, kuid katsete arvu suurenedes nõuaks see mahukaid arvutusi. Tuletame arvutusi lihtsustava valemi.

Meid huvitab tõenäosus, et sündmus A toimuks n katse puhul täpselt m korda, aga ülejäänud $n - m$ katse puhul toimuks tema vastandsündmus $\bar{A}$. Seejuures pole tähtis, millistel katsetel sündmus A toimub. Erinevate

võimaluste arvu sündmuse A toimumiseks täpselt m korda n -kordsel katsetamisel annab kombinatsioonide arv n elemendist m elemendi kaupa, s.t. C_n^m (selle tulemuseni jõudmiseks võib sündmustele anda järjekorranumbrid). Ühe niisuguse kombinatsiooni näiteks oleks sündmus B mille puhul sündmus A toimub esimese m katse korral ja vastandsündmus $\bar{A}$ järgmise $n-m$ katse korral:

$$B = \underbrace{A A \dots A}_m \underbrace{\bar{A} \bar{A} \dots \bar{A}}_{n-m}.$$

Kuna katsed on sõltumatud, siis kasutades korrutamisteoreemi sõltumatute sündmuste kohta, saame:

$$P(B) = \underbrace{P(A)P(A)\dots P(A)}_m \underbrace{P(\bar{A})P(\bar{A})\dots P(\bar{A})}_{n-m} = p^m q^{n-m}.$$

Kuna sündmusega B analoogsed sündmused on üksteist välistavad, siis võib kasutada üksteist välistavate sündmuste liitmisteoreemi. Tulemuseks saame **Bernoulli valemit**:

$$P_{m,n} = C_n^m p^m q^{n-m} = \frac{n!}{m!(n-m)!} p^m q^{n-m}.$$

Sündmused, mis seisnevad sündmuse A toimumises n katsel $0, 1, 2, \dots, n$ korda, on üksteist välistavad ning moodustavad sündmuste täieliku süsteemi. Seega üldiselt

$$\sum_{m=0}^n P_{m,n} = 1$$

ja

$$\sum_{m=0}^n C_n^m p^m q^{n-m} = q^n + C_n^1 p q^{n-1} + \dots + C_n^m p^m q^{n-m} + \dots + p^n = (q+p)^n = 1. \quad (10.1)$$

Eelnevas $q=1-p$. Kuna summa $\sum_{m=0}^n P_{m,n}$ on ühtlasi binoomi $(q+p)^n$ arenduseks, siis nimetatakse tõenäosuste jaotust kujul (10.1) **binoomjaotuseks**. Binoomjaotust käsitletakse lähemalt edaspidi.

Toetudes analoogiale binoomiga, defineerime tõenäosusi $P_{m,n}$ **genereeriva funktsiooni** $\varphi_n(x)$:

$$\varphi_n(x) = (q + px)^n = q^n x^0 + C_n^1 p q^{n-1} x^1 + \dots + C_n^m p^m q^{n-m} x^m + \dots + p^n x^n.$$

(10.2)

$P_{m,n}$ on genereerivas funktsioonis liikme x^m kordaja.

Kui tehakse **erinevates tingimustes** (mitte ühesugustes, nagu eespool) n sõltumatut katset, kusjuures sündmuse A toimumise tõenäosus i -ndal katsel on p_i , siis tõenäosus $P_{m,n}$ (et sündmus A toimuks n katsel täpselt m korda) on leitav genereeriva funktsiooni $\psi_n(x)$ abil:

$$\psi_n(x) = (q_1 + p_1 x)(q_2 + p_2 x) \dots (q_n + p_n x), \quad (10.3)$$

kus $q_i = 1 - p_i$. $P_{m,n}$ on liikme x^m kordaja.

Näide 10.1: Leida märgi tabamiste arvu tõenäosuste jaotus viie sõltumatu lasu korral, kui märgi tabamise tõenäosus igal üksikul lasul on $0,6$.

Lahendus: Ülesande tigimuste järgi $n=5$, $p=0,6$ ja $q=0,4$. Järelikult genereeriv funktsioon (2) on

$$\varphi_5(x) = (0,4 + 0,6x)^5 = 0,01024 + 0,0768x + 0,2304x^2 + 0,3456x^3 + 0,2592x^4 + 0,07776x^5.$$

Otsitavad tõenäosused on astme x^m vastavad kordajad:

$$P_{0,5} = 0,01024; \quad P_{1,5} = 0,0768; \quad P_{2,5} = 0,2304;$$

$$P_{3,5} = 0,3456; \quad P_{4,5} = 0,2592; \quad P_{5,5} = 0,07776.$$

Kontroll:

$$\sum_{m=0}^5 P_{m,5} = 0,01024 + 0,0768 + 0,2304 + 0,3456 + 0,2592 + 0,07776 = 1.$$

Lahendus on seega õige. Lahendusest ilmneb, et tõenäoseim on viiel lasul tabada kolm korda.

Näide 10.2: Üksteise järel saadetakse välja neli raadiosignaali. Iga signaali vastuvõtt on sõltumatu ülejäänud signaalide vastuvõtust. Üksikute signaalide vastuvõtu tõenäosused on 0,2 ; 0,3 ; 0,4 ; 0,5. Leida kolme signaali vastuvõtu tõenäosus.

Lahendus: Kasutades valemit (3), saame:

$$\begin{aligned} \psi_4(x) &= (0,8 + 0,2x)(0,7 + 0,3x)(0,6 + 0,4x)(0,5 + 0,5x) = \\ &= 0,168 + 0,394x + 0,32x^2 + 0,106x^3 + 0,012x^4. \end{aligned}$$

Otsitav $P_{3,4}$ on x^3 kordajaks: $P_{3,4} = 0,106$.

11. TÕENÄOSEIM SÜNDMUSE TOIMUMISTE ARV

Sündmuse A toimumise tõenäoseimaks arvuks n sõltumatu katse korral on m_0 , mille puhul tõenäosus $P_{m_0, n}$ ei ole väiksem ühestki teisest katsete arvu tõenäosusest $P_{m, n}$. Eelneva punkti teises näites on $\max P_{m_0, n} = P_{1,4} = 0,394$. Seega $m_0 = 1$, mis tähendab, et kõige tõenäosem on ühe signaali vastuvõtt.

Tõenäoseimat sündmuse toimumiste arvu on aga võimalik lihtsamalt leida. Ilmselt

$$P_{m_0, n} \geq P_{m_0+1, n} \quad \text{ja} \quad P_{m_0, n} \geq P_{m_0-1, n}.$$

Neist võrratustest esimene annab vastavalt Bernoulli valemile

$$\frac{n!}{m_0!(n-m_0)!} p^{m_0} q^{n-m_0} \geq \frac{n!}{(m_0+1)!(n-m_0-1)!} p^{m_0+1} q^{n-m_0-1}$$

ehk

$$\frac{q}{n-m_0} \geq \frac{p}{m_0+1},$$

millest

$$m_0 q + q \geq np - m_0 p$$

ja

$$m_0(q+p) \geq np - q$$

ning lõpuks

$$m_0 \geq np - q.$$

Seos $P_{m_0, n} \geq P_{m_0-1, n}$ annab analoogsete teisenduste tulemusena

$$m_0 \leq np + p$$

ja seega kokkuvõttes

$$np - q \leq m_0 \leq np + p.$$

12. GEOMEETRILINE TÕENÄOSUS

Olgu igale elementaarsündmusele seatud vastavaks punkt tasandil, kusjuures kõigi elementaarsündmuste hulga vastab tasandiline piirkond pindalaga S . Igale sündmusele A vastab teatav punktihulk selles tasandilises piirkonnas ja sündmust A võib vaadelda kui juhusliku punkti sattumist vastavasse osapiirkonda, mille pindala olgu S_1 . Kui sisulistel kaalutlustel võib lugeda, et juhusliku punkti sattumise tõenäosus mingisse osapiirkonda on võrdeline selle osapiirkonna pindalaga, siis

$$P(A) = \frac{S_1}{S}.$$

Seda tõenäosust nimetatakse sündmuse A **geomeetriliseks tõenäosuseks**.

Geomeetrilist tõenäosust kasutatakse ülesannetes, mida võib taandada ülesandele punkti juhusliku sattumise kohta ruumi lõplikku piirkonda. Paneme veel tähele: et punkti pikkus, pindala või ruumala on 0, saame, et geomeetrilise tõenäosuse korral on iga elementaarsündmuse tõenäosus 0.

Eelnevas piirduti kahemõõtmelise juhuga. Ühemõõtmeline ja kolmemõõtmeline juht erinevad ainult selle poolest, et pindala asemel on seal vastavalt kas pikkus või ruumala.

Näide 12.1: Egiptuse liinibuss saabub peatusesse juhuslikult ajavahemikus 8:00 kuni 20.00. Turist hakkab bussi ootama kell 12:00 ning ootab maksimaalselt poolteist tundi. Leida tõenäosus, et ta saab bussi peale. Geomeetrilise tõenäosuse korral ei ole oluline see, mis kell turist bussi ootama hakkas vaid see, kaua ta vastu peab. Antud juhul poolteist tundi. Otsitav tõenäosus on siis

Lahendus: $\frac{90}{12 \cdot 60} = \frac{1}{8}.$

Näide 12.2: Kaks isikut leppisid kokku kohtuda teatud kohas ajavahemiku T jooksul. Leida tõenäosus, et kohtumine toimub, kui isikute saabumishetked on sõltumatud ja esimesena saabuja ei oota kauem kui t .

Lahendus: Tähistame ühe isiku saabumishetke tähega x ja teise saabumishetke tähega y .

Elementaarsündmuste

ruum on nüüd $\Omega = [0, T] \times [0, T]$. Et kohtumine toimuks, on tarvis, et $|x - y| \leq t$.

Seega meid huvitav sündmus on

$$A = \{(x, y) \in [0, T] \times [0, T]: |x - y| \leq t\}.$$

Geomeetriline tõenäosus on seega

$$P(A) = \frac{S_1(A)}{S(\Omega)} = \frac{T^2 - (T - t)^2}{T^2}.$$

13. SÕLTUMATUD SÜNDMUSED

Väga sageli on intuiitselt selge, et ühe katse tulemus on täiesti sõltumatu teise katse tulemusest. Sellisel juhul on ka iga esimese katse tulemusega määratud sündmus sõltumatu teise katse tulemusest sõltuva sündmusega: sündmused on sõltumatud. Sõltumatute sündmuste A ja B korral ei mõjuta A toimumine või mittetoimumine kuidagi B toimumist. Siit järeldub, et A toimumine ei muuda B toimumise tõenäosust, st $P(B|A) = P(B)$. Tingliku tõenäosuse definitsioonist saame, toodud võrdus on ekvivalentne järgmise võrdusega:

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

Definitsioon 13.1: Sündmusi A ja B nimetatakse sõltumatuteks, kui kehtib $P(AB) = P(A)P(B)$.

Märkused:

1. Erinevalt tingliku tõenäosuse definitsioonist võib nii A kui ka B tõenäosus olla 0.
2. Kui A ja B on sõltumatud, siis on sõltumatud ka sündmused
 - a. $\bar{A}$ ja B
 - b. $\bar{B}$ ja A
 - c. $\bar{A}$ ja $\bar{B}$
3. Sündmus, mille tõenäosus on 0 või 1 on sõltumatu igast teisest sündmustest, kaasaarvatud iseendast.
4. Kui A ja B on sõltumatud ja $P(A) > 0$ ning $P(B) > 0$, siis $P(A|B) = P(A)$ ja $P(B|A) = P(B)$. Üaltoodust järeldub, et sellisel juhul ka $P(A|\bar{B}) = P(A)$, $P(\bar{A}|B) = P(\bar{A})$ ja $P(\bar{A}|\bar{B}) = P(\bar{A})$.

Sageli informatsioon ühe sündmuse toimumise kohta suurendab või vähendab teise sündmuse toimumise tõenäosust. Sel juhul on tegemist sõltuvate ehk korreleeritud sündmustega.

Definitsioon 13.2: Sündmusi A ja B nimetatakse positiivselt korreleerituteks, kui $P(A|B) > P(A)$ ning negatiivselt korreleerituteks, kui $P(A|B) < P(A)$.

Enam kui kahe sündmuse sõltumatus defineeritakse järgmiselt.

Definitsioon 13.3: Sündmusi $A_1, \dots, A_n$ nimetatakse sõltumatuteks, kui iga arvu $k \in \{2, 3, \dots, n\}$ ja iga võrratusi $1 \leq i_1 < i_2 < \dots < i_k \leq n$ rahuldava indeksite komplekti korral kehtib võrdus

$$P(A_{i_1} \cdot A_{i_2} \cdot \dots \cdot A_{i_k}) = P(A_{i_1})P(A_{i_2}) \dots P(A_{i_k}). \quad (13.1)$$

Definitsioonist tuleneb, et kolm sündmust A, B ja C on täielikult sõltumatud, kui kehtivad võrdused

$$P(AB) = P(A)P(B), \quad P(AC) = P(A)P(C), \quad P(BC) = P(B)P(C) \\ P(ABC) = P(A)P(B)P(C).$$

Seega $A_1, \dots, A_n$ on sõltumatud, kui kõik paarid A_i, A_j ($i < j$) on sõltumatud, kõik kolmikud A_i, A_j, A_k ($i < j < k$) on sõltumatud, kõik nelikud on sõltumatud jne.

Definitsioon 13.4: Sündmusi $A_1, \dots, A_n$ nimetatakse paarikaupa sõltumatuteks, kui kõik paarid A_i, A_j ($1 \leq i < j \leq n$) on sõltumatud.

Kahe sündmuse korral langevad sõltumatus ja paarikaupa sõltumatus kokku. Üldiselt see nii pole: sõltumatud sündmused on ka paarikaupa sõltumatud, kuid paarikaupa sõltumatud sündmused ei pruugi sõltumatud olla.

Märgime veel, et lõpmatut hulka sündmusi $\{A_\alpha\}$ (võib olla mitteloenduv) nimetatakse sõltumatuteks, kui iga lõplik alamhulk sündmusi on sõltumatud.

Definitsioon 13.5: Olgu B mingi positiivse tõenäosusega sündmus, st $P(B) > 0$. Sündmusi $A_1, \dots, A_n$ nimetatakse tinglikult sõltumatuks tingimusel et B on toimunud, kui iga arvu $k \in \{2, 3, \dots, n\}$ ja iga võrratusi $1 \leq i_1 < i_2 < \dots < i_k \leq n$ rahuldava indeksite komplekti korral kehtib võrdus

$$P(A_{i_1} \cdot A_{i_2} \cdot \dots \cdot A_{i_k} | B) = P(A_{i_1} | B)P(A_{i_2} | B) \dots P(A_{i_k} | B). \quad (13.2)$$

Seega tinglik sõltumatus pole midagi muud kui sõltumatus tõenäosusmõõdu $Q(A) := P(A|B)$ korral.