EMPÜ Matemaatika Instituut

E. ARUVEE, U. ENGSTRAND, U. OLSSON

# Matemaatiline statistika

Põhimõisted

Bioloogilistele erialadele

EPMÜ Matemaatika Instituut

E. ARUVEE, U. ENGSTRAND, U. OLSSON

# Matemaatiline statistika

Põhimõisted

Bioloogilistele erialadele

TARTU 2000

# Matemaatiline statistika

Põhimõisted

Bioloogilistele erialadele

Toimetaja N. Veske

Käsikirja ettevalmistamist toetas Rootsi Instituut
(grant 2080/1999 (380/98))



ISBN

# Eessõna

Koostöös Rootsi Põllumajandusülikooli kolleegidega Ulf Olssoniga ja Ulla Engstrandiga ning Rootsi Instituudi grandi toetusel valmis käesolev õpik. Seda tingis vajadus hoogustada matemaatilise statistika õpetamist Eesti Põllumajandusülikoolis. Käesoleva materjali valikul on tuginetud Rootsi Põllumajandusülikooli loengute materjalidele ja pikaajalistele kogemustele matemaatilise statistika õpetamisel. Arvutite laialdase kasutamisega koguneb mitmesuguseid andmestikke kiiresti, kuid sageli ei osata nednega midagi peale hakata. Seepärast on hariduse vajalikuks osaks oskus töötada arvandmestikuga, mida võimaldavad matemaatilise statistika meetodid.

Raamatus käsitletakse matemaatilise statistika põhimõisteid. Näidatakse, kuidas leida statistikute väärtusi käsitsiarvutamisel ja programmipakette kasutades. Teksti on illustreeritud statistikapakettide Minitab ja SAS väljatrükkidega. Minitabi kasutamisel võlub tema lihtsus, kuid SAS-i puhul professionaalsus. Selle raamatu loogilise jätkuna ilmub peatselt raamat "Andmeanalüüs" mis koos käesoleva raamatuga moodustavad põllumajandusülikoolis õpetatava matemaatilise statistika kursuse.

Tänan kolleeg Jüri Vardjat käsikirja läbilugemise ja tehtud kasulike märkuste eest. Raamatu eestikeelse käsikirja valmimise eest kuulub eriline tänu toimetaja Nora Veskele, kelle tähelepanelik töö oli raamatu valmimisel hindamatuks abiks.

Tartus, juuni 2000. a Eve Aruvee

# Peatükk 1

# Statistiline vaatlus

## 1.1 Statistika ja bioloogia

Füüsika põhjal näib elu väga lihtne. Me formuleerime füüsikalise seaduse, nagu näiteks $E = mC^2$, ning seejuures võime kindlad olla, et antud seadus kehtib kõikjal; vähemalt järgmise suure teadusliku avastuseni. Füüsika on laialdase ulatusega deterministlik teadus. See tähendab, et paljusid füüsikalisi seadusi võime väljendada valemiga

$$y = f(x),$$

kus $y$ on tundmatu suurus, mida soovime leida, $f$ on mingi funktsioon ja $x$ on funktsiooni argument.

Loodusseadused bioloogias, kaasa arvatud põllumajanduslikes teadustes, ei ole nii lihtsad. Sageli toob bioloogiline nähtus esile muutuse, senitundmatud seadused, mida soovime hinnata. Kui käsitleme kahte ala ühesuguse meetodiga, võime saada erinevad tulemused. Kahel ühevanusel konnal võivad olla erinevad kaalud. Kaks lehma, kes tarvitavad sama tüüpi võrdse koguse sööta, võivad anda erineva hulga piima. Järelikult: *varieeruvus* (muutlikkus) on bioloogilise andmestiku tüüpiline tunnus.

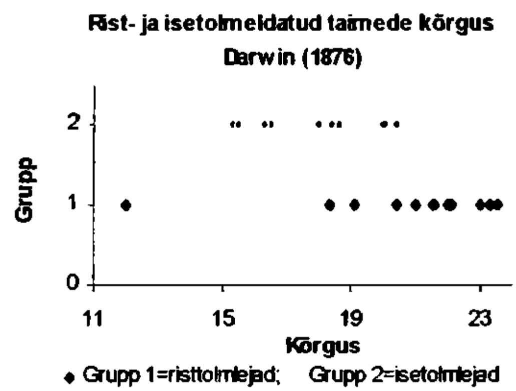
Varieeruvus, mida me ei oska selgitada, põhjustab probleeme, kui püüame koostada järeldusi bioloogilise andmestiku kohta. Üheks meetodiks bioloogilise varieeruvuse käsitlemisel on vaadelda seda juhuslikkusena. Seega paljud loodusseadused bioloogias on tõenäosuslikud. See tähendab, et me usaldame tõenäosuslikke väiteid.

Näiteks mingi bioloogilise nähtuse mudeli võime kirjutada kujul

$$y = f(x) + e,$$

kus $y$ on mingi muutuja uuritav väärtus, $f$ on mingi funktsioon, mis sümboliseerib oodatavat või keskmist muutuja $y$ väärtust mingis üldkogumis, $x$ on muutuja, mis mõjutab $y$-it ja $e$ on mudeli viga. Mudeli vea ehk *jäägi* tõttu on sellist tüüpi mudelid tõenäosuslikud. Hiljem näeme, et statistilisi meetodeid ja mudeleid saab vaadelda selle üldise valemi erijuhuna.

Näide. Aastal 1876 ilmus Charles Darwinilt raamat "Taimeriigi rist- ja isetolmlemise mõjud" Darwin püüdis selles raamatus näidata, et taimed, mida on paljundatud risttolmeldamise teel, on elujõulisemad kui taimed, mida on paljundatud isetolmeldamise teel. Ta tegi katse, kus ühe liigi 15 taime kasvatati ühesugustes tingimustes. Mõne aja pärast mõõtis ta taimede kõrgused. Tulemused võib graafiliselt kujutada järgmiselt:




Graafikult näeme, et grupi 1 taimed on keskmiselt pikemad kui taimed, mis kuuluvad gruppi 2, aga tulemus ei ole päris tõepärane. Siin on osaline kattumine gruppide vahel ja grupis 1 oli ka väga väikesi taimi. On raske hinnata, kas Darwin'il oli õigus, lähtudes ainult graafikust. Küsimus seisneb: missugusele üldisele järeldusele võime jõuda selle andmestiku põhjal?

Üheks võimaluseks jõuda sellist tüüpi andmestiku põhjal teaduslikule järeldusele on formuleerida statistiline mudel. Darwin'i juhul võib mudel olla midagi taolist:

$$\text{Taime pikkus} = \begin{array}{c}\text{(kõikide taimede}\\ \text{keskmine pikkus)}\end{array} + \begin{array}{c}\text{(muutus, kui taime}\\ \text{on risttolmeldatud)}\end{array} + \begin{array}{c}\text{(juhuslik}\\ \text{variatsioon)}\end{array}$$

See on tõenäosuslik mudel, mis on analoogiline eelpool esitatule $y = f(x) + e$. Enamikel juhtudel püstitame väite mudeli süstemaatilise osa $f(x)$ kohta. Selle näite korral soovime leida vastust küsimusele: "Kas risttolmeldatud taimed on keskmiselt teistest pikemad"?

## 1.2 Üldkogumid ja valimid

Teaduslikud järeldused peavad olema mingis mõttes üldised. Ühe katse tulemusi peab olema võimalik seostada mingi üldise teooriaga. Katset või uuringut planeerides peame määratlema uurimisobjekti, mille kohta soovime järeldusi teha.

Statistikas nimetatakse uurimisobjekti üldkogumiks, kuid bioloogilistes uurimisvaldkondades kasutatakse üldkogumi sünonüümina ka populatsiooni mõistet. Soovime formuleerida isendite populatsiooni kohta üldist teooriat. Darwin'i näites tahame saada järeldust ainult ühe liigiga piirduva taimede populatsiooni kohta. Üldkogum võib eksisteerida mingis reaalses tähenduses, näiteks "Kõik Eestis elavad piimalehmad seisuga 1. jaanuar 1999"; "Eesti üliõpilaste hulk 1999/2000. õppeaastal" Teisel juhul on ta kujutletav. Näiteks soovis Darwin arvatavasti teha järeldusi taimede kohta üldse, mitte ainult antud liigi isendite kohta. On keeruline koostada head üldkogumi definitsiooni, aga teeme proovi.

**Definitsioon 1** *Üldkogum on mingit kindlat liiki objektide hulk.*

Katsetes kasutame harva kogu üldkogumit. Üldiselt teeme mõõtmisi üldkogumist võetud valimi elementide kohta. Ning kasutame seda järelduste tegemiseks üldkogumi kohta.

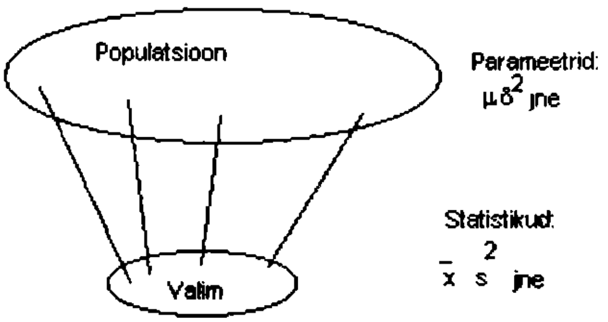
**Definitsioon 2** *Valim on üldkogumi alamhulk.*

Igas katses peab olema mõningane sarnasus valimi ja üldkogumi vahel. Valim peab *esindama üldkogumit* (olema *representatiivne*). Kuidas me selle saavutame?

Statistiline vastus on: valim tuleb moodustada juhusliku valiku teel, üldkogumi igal elemendil peab olema ühesugune võimalus sattuda valimisse. Selle eelis on, et uurija ei saa luua valimit süstemaatilisel teel, mis üldkogumist oluliselt erineks. Sageli teostatakse juhuslik valik juhuslike arvude generaatori või tabeli abil. Niisiis, juhusliku valiku tõttu võime kasutada käsitleme meetodeid, milles valim on moodustatud juhuslikult. Seega valimi all mõistame juhuslikku valikuga moodustatud valimit mingist üldkogumist.

## 1.3 Statistiline otsustus

Suur osa sellest kursusest on pühendatud üldiste järelduste tegemise probleemidele, mis põhinevad andmestiku lõplikel valimitel. Näiteks soovime järeldust teha mingi üldkogumi keskväärtuse kohta. Moodustame üldkogumist valimi ja arvutame tema keskväärtuse. Tulemust kasutame järelduste tegemiseks üldkogumi keskväärtuse kohta mingi kindla täpsusega. Meetodid, mida kasutame, põhinevad tõenäosusteoorial. Antud olukorda saab kujutada järgnevalt:




Kursus on üldjoontes jagatud kolmeks osaks.

Esimeses osas uurime kogutud andmete kirjeldavaid meetodeid. Andmete paremaks mõistmiseks on otstarbekas arvutada andmete numbrilisi statistikuid või esitada andmed diagrammil.

Teine osa on rohkem teoreetiline. See on tõenäosusest ja statistilistest jaotustest. Kui kursuse esimene osa aitab meil kirjeldada valimit, siis teine esitab üldkogumi kirjeldamise meetodeid.

Kõige mahukamas, kursuse kolmandas osas (ilmub käesoleva raamatu teises osas), käsitleme meetodeid järelduste tegemiseks valimi alusel üldkogumi kohta ja neid nimetatakse *statistilisteks otsustusteks.*

# Peatükk 2

# Kirjeldava statistika põhimõisted

## 2.1 Andmete kogumine

### 2.1.1 Sissejuhatus

Bioloogilise andmestiku kogumise tulemuseks võib olla suur hulk andmeid. Nende interpreteerimiseks on vaja leida mingi sobiv mudel, aga sageli on seda raske valida, kuna võimalusi on palju. Kirjeldava statistika eesmärk on korrastada andmed mitmel viisil selliselt, et tähtsamad suunad oleksid selgitatud. See ei ole formaalne probleem; sageli on see "detektiivi töö", kus me pöörame ja painutame andmeid erineval viisil seni, kuni märkame mudelit ja seni, kuni oleme võimelised selgitama mudelit teistele.

Kirjeldav statistika sisaldab endas ka edastamist. Uurija, kes on töötanud eksperimendiga mingi aja, võib sageli märgata andmestikus seda, mida teine ei oska näha. Illustreerides leidu graafikul või tabelis, võime veenda teisi, et järeldused on õiged. Niisiis, kirjeldav statistika on andmeanalüüsi tähtis osa nii uurijale kui ka uuringu tulemuse kasutajale.

### 2.1.2 Tunnused

Tunnus on mingi näitaja, mida võime kirjeldada või mõõta andmestikku kuuluval objektil. Näiteks kui objektideks on inimesed, võime igat neist kirjeldada vanuse, soo, kaalu, õdede - vendade arvu, aastase sissetuleku, juuste värvuse, matemaatika hinde või elukutse põhjal. Sel juhul on tunnused:

vanus, sugu, kaal, õdede - vendade arv, aastane sissetulek, juuste värvus, matemaatika hinne või elukutse.

### 2.1.3 Kvalitatiivsed ja kvantitatiivsed tunnused

Kirjeldava statistika meetodeid (samuti ka teisi statistika meetodeid) kasutatakse sõltuvalt sellest, kas tunnus on kvalitatiivne või kvantitatiivne. Kvantitatiivsete (või arvuliste) tunnuste väärtused on reaalarvud. Kvantitatiivse tunnuse näiteks on vanus, kaal ja aastane toodang.

Kvantitatiivne tunnus võib olla pidev või diskreetne. Pidev tunnus peab omama mingit väärtust mingist vahemikust. Näiteks inimeste kaal võib olla mõõdetud (vähemalt põhimõtteliselt) määramatu täpsusega. Järelikult – kaal ja vanus on pideva kvantitatiivse tunnuse näited.

Diskreetne tunnus saab omandada vaid kindlaid erinevaid väärtusi, tunnuse väärtused määratakse loendamise teel. Näiteks õdede-vendade arvuks saavad olla vaid väärtused 0, 1, 2, 3, ..., st nad on vaid täisarvud.

Kvalitatiivsed tunnused e mittearvulised tunnused jaotatakse omakorda järjestatud -, binaarsed - ja nominaaltunnusteks. Kõige levinum on järjestus- (järjestatud) tunnus. Seda iseloomustab väärtuste hulga sisuline järjestatus. Näiteks matemaatika hinne, hinnang etendusele. Binaarse tunnuse näiteks on sugu. Sool võib olla vaid kaks väärtust (enamike liikide juures): naine ja mees. Juuste värvus on nominaaltunnuse näide. Selle korral ei leidu üht ühist sisuliselt põhjendatud skaalat, mille alusel iga kaht väärtust saaks võrrelda. Sageli kasutatakse andmetöötlusprotsessis mingeid koode "soo" või "juuste värvuse" jaoks. Näiteks võime kasutada koodina arvu 0 väärtuse "naine" ja arvu 1 väärtuse "mees" jaoks, juuste värvuse koodid võivad olla järgnevad: 1 - blondid, 2 - pruunijuukselised, 3 - mustajuukselised jne. Nominaaltunnuse korral võime ka kasutada koode, kuid ei tohi neid töödelda arvulistena. Kodeerimine ei tohi muuta tunnuste fundamentaalset olemust: sugu ja juuste värvus on vaid kvalitatiivsed tunnused ja neid peab analüüsima ka selliselt.

### 2.1.4 Skaala tüübid

Erinevate tunnuste mõõtmiseks võib kasutada erinevaid skaalasid.

| Skaala tüüp | Väärtused on järjestatud | Väärtused on võrdvahemikulised | Skaalal on 0-punkt |
|---|---|---|---|
| Nominaalne | | | |
| Ordinaaalne | X | | |
| Intervall | X | X | |
| Proportsionaalne | X | X | X |

Nominaalne skaala ainult nimetab väärtusi, mitte ei järjesta neid. Näiteks sugu.

Ordinaalne skaala järjestab kategooriaid. Näiteks hinne matemaatikas.

Intervallskaala on võrdvahemikuline, kuid tal ei ole alguspunkti. Näiteks temperatuur Celsiuse järgi.

Proportsionaalsel skaalal on 0-punkt, kuid tunnusel see omadus puudub. Näiteks kaal.

Meetodid, mida kasutatakse statistilise andmestiku (võivad olla esitatud tabelite või graafikute kujul) jaoks, sõltuvad sellest, missugust tüüpi tunnustega me töötame ja missugust tüüpi skaalas on tehtud tunnuse mõõtmised.

## 2.2 Numbriliste andmete ülevaade

### 2.2.1 Aritmeetiline keskmine

**Definitsioon 3** *Valimis vaatluste (aritmeetiline) keskmine arvutatakse*

$$\bar{x} = \frac{\sum x_i}{n}$$

Selles valemis tuleb kõigi vaatluste väärtuste summa jagada vaatluste arvuga.

Näide 1.

Järgnev andmestik on rottide, kelle vanus on 1 päev ja rohkem, hapniku hingamise hulk ($VO_2$, ml/t)[1]

1,57 1,60 1,81 1,60 1,88 1,79 1,75 1,85

1,92 1,68 1,75 1,73 1,72 1,68 1,76 1,78

[1] Andmed on võetud: Claes Lilja A note on the metabolic rate in the young rat. Swedish J. agric. Res. 17:103-104, 1987

Kõikide vaatluste väärtuste summa
$\sum x_i = 1,57 + 1,60 + \quad + 1,76 + 1,78 = 27,87$
Valimi, 16 roti, keskmine on $\bar{x} = \frac{\sum x_i}{n} = \frac{27,87}{16} = 1,741875$.
Keskmist võib ümardada vajaliku täpsuseni. Tulemuse võib esitadada $\bar{x} = 1,74$ või $\bar{x} = 1,742$.

Märkus: tulemust ei tohi ümardada, kui soovime seda kasutada edaspidistes arvutustes. Keskmist on mõttekas arvutada ainult numbrilistele tunnustele (vähemalt) intervallskaalas. Mõnikord arvutatakse ka aritmeetilist keskmist ordinaalses skaalas mõõdetud andmete korral, mis on teataval määral sobimatu. Näiteks õpilaste hinded koolis.

### 2.2.2 Mediaan

Valimi mediaani tähistatakse sageli Me või $\tilde{x}$ Mediaan on väärtus, millest valimis pooled vaatlused on väiksemad ja pooled on suuremad. Mediaani leidmiseks on vaja moodustada valimi elementidest järjestatud jada, mida nimetatakse variatsioonreaks. Seejärel leiame mediaani asukoha.

**Definitsioon 4** *Mediaan on variatsioonrea vaatlus positsiooniga $\frac{n+1}{2}$ Kui positsioon ei ole täisarv, siis mediaan on kahe keskmise vaatluse väärtuste poolsumma.*

Näite 1 andmetest on moodustatud variatsioonrida, st valimi elemendid on sorteeritud kasvavasse järjekorda; iga arvu positsioon on antud tema all.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 1,57 | 1,60 | 1,60 | 1,68 | 1,68 | 1,72 | 1,73 | 1,75 |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| 1,75 | 1,76 | 1,78 | 1,79 | 1,81 | 1,85 | 1,88 | 1,92 |
| 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |

Vaatlusandmeid on $n$=16 (paarisarv), mediaani asukoht on $\frac{16+1}{2} = 8,5$. Mediaan on kahe keskmise vaatluse väärtuse keskmine, st 8. ja 9. variatsioonrea elemendi väärtuste keskmine. Need vaatlused on 1,75 ja 1,75. Seega $Me = \tilde{x} = 1,75$. Mediaani võib leida numbriliste tunnuste (intervalli või proportsionaalse skaala) korral ning ordinaalses skaalas tunnustele.

### 2.2.3 Kohendatud keskmine (Trimmed mean)

Aritmeetiline keskmine on kõikide vaatlusväärtuste keskmine. Ta on tundlik ekstreemsete vaatlusväärtuste suhtes. Näiteks aritmeetiline keskmine arvudest 1, 2, 3, 4 ja 5 on 3. Arimeetiline keskmine arvudest 1, 2, 3, 4 ja 50 on

12. Järelikult võib aritmeetilne keskmine ühe vaatluse tõttu palju muutuda. Mediaan ei ole tundlik mõnede ekstreemsete vaatlusväärtuste suhtes: mediaan on 3 mõlema eelpool vaadeldud näite korral.

Teiste sõnadega võib öelda, et keskmine $\bar{x}$ põhineb kõigil $n$ vaatlusel, aga mediaan $\tilde{x}$ baseerub peamiselt kahel vaatlusel. Kompromissina kasutatakse järgmist protseduuri: näiteks kustutatakse 5% suurimatest ja 5% väikseimatest vaatlusväärtustest ning arvutatakse aritmeetiline keskmine ülejäänud vaatluste väärtustest.

Vaatleme näites 1 esitatud andmestikku, kustutame variatsioonreast suurima (1,92) ja vähima (1,57) elemendi. Aritmeetiline keskmine ülejäänud elementidest on 1,7414.

Märkus. Kuigi kohendatud keskmine võib intuitiivselt meeldida, eriti kui andmestik sisaldab erilisi väärtusi (erindeid), ei kasutata teda statistiliste järelduste tegemisel kuigi sageli.

### 2.2.4 Dispersioon

Dispersioon on valimis mõõdetav kui vaatlusväärtuste hajuvus keskväärtuse suhtes. Kui vaatlused on kõik lähedaste väärtustega, siis on dispersioon väike, aga kui vaatluste väärtused on väga erinevad, siis on dispersioon suur.

**Definitsioon 5** *Valimi dispersioon on defineeritud valemiga*

$$s^2 = \frac{\sum (x_i - \bar{x})^2}{n-1}$$

See valem on käsitsi arvutamisel veidi tülikas, kuna on vaja lahutada iga vaatluse väärtusest keskväärtus $\bar{x}$ ning seejärel tulemus ruutu tõsta. Valem, mis annab sama tulemuse, aga mida on käsitsi kergem kasutada, on järgmine: $s^2 = \frac{\sum x_i^2 - \frac{(\sum x_i)^2}{n}}{n-1}$

Näite 1 andmete põhjal võib dispersiooni arvutada järgmiselt.
Kõigi vaatlusväärtuste summa on
$\sum x_i = 1,57 + 1,60 + \quad + 1,76 + 1,78 = 27,87$

Kõigi vaatlusväärtuste ruutude summa on
$\sum x_i^2 = 1,57^2 + 1,60^2 + \quad + 1,76^2 + 1,78^2 = 2,4649 + 2,5600 + \quad + 3,0976 + 3,1684 = 48,6955$.

Meenutame, et valimimaht $n = 16$, dispersioon on
$s^2 = \frac{\sum x_i^2 - \frac{(\sum x_i)^2}{n}}{n-1} = \frac{48,6955 - \frac{27,87^2}{16}}{15} = 0,009963$.
Tulemuse võime ümardada $s^2 = 0,01$.

Märkus. Taskuarvutid ja tabelarvutusprogrammid nagu Excel omavad dispersiooni jaoks kahte dispersiooni standardfunktsiooni, mis sageli põhjustab veidi segadust. Näiteks Excelis on üks funktsioon VAR(), mis on paketi abitekstis kirjeldatud järgmiselt: "Arvutab dispersiooni, mis põhineb valimil", ja teine funktsioon VARP() on kirjeldatud järgmiselt: "Arvutab dispersiooni, mis põhineb üldkogumil" Nendevaheline erinevus seisneb selles, et üldkogumi valem kasutab $n$-i murru nimetajas $(n-1)$ asemel. Käesolevas tekstis eeldame, et dispersiooni arvutamiseks kasutame eelpool toodud valimi valemit.

### 2.2.5 Standardhälve

**Definitsioon 6** *Valimi standardhälve on ruutjuur valimi dispersioonist:*

$$s = \sqrt{s^2}$$

Näite 1 andmestiku standardhälve on
$s = \sqrt{s^2} = \sqrt{0,009963} = 0,099814$. Tulemuse võime ümardadada
$s = 0,1$.

Märkus. Sageli esitatav küsimus on: "Miks vajame mõlemat - dispersiooni ja standardhälvet"? Vastus on: standardhälvet on lihtsam interpreteerida, kuna ta on mõõdetud samades ühikutes kui andmed. Kui andmed on cm-tes, siis standardhälve on ka cm-tes. Järelikult, kui on vaja kirjeldada muutust andmehulgas, võime kasutada pigem $s$-i kui $s^2$ Kuna aga paljud statistilised arvutused põhinevad dispersioonil, tuleb sagedamini kasutada $s^2$ kui $s$-i.

### 2.2.6 Haare

Haare on defineeritud kui suurima ja vähima vaatlusväärtuse erinevus, ta on andmestikus hõlpsasti arvutatav hajuvuskarakteristik.

Näite 1 andmetes on suurim vaatluse väärtus 1,92 ja vähim väärtus on 1,57 Järelikult haare on $1,92 - 1,57 = 0,35$.

Märkus. Andmete korral, kui valimi suurus ei ole nii suur, kasutatakse haaret, et anda esialgne hinnang standardhälbele. Sageli on see järgmine $s \approx \frac{haare}{4}$ Seda lihtsat "rusikareeglit" võib kasutada standarhälbe arvutuste kontrollimiseks. Selle reegli kohaselt näite 1 andmete jaoks on $s \approx \frac{0,35}{4} = 0,0875$, mis ei erine väga palju eelpool arvutatud õigest väärtusest $0,1$.

### 2.2.7 Variatsioonikordaja (CV)

Mõnel juhul vajame varieeruvusele andmestikus väljundit mingis ühikus, mis ei tohi sõltuda kasutatavate mõõtmiste skaalast. Sel juhul arvutame variatsioonikordaja, mis on tunnuse standardhälbe ja keskväärtuse suhe;

$$CV = \frac{s}{\bar{x}}$$

Näite 1 andmete korral $CV = \frac{0{,}099814}{1{,}7419} = 0,057303$. $CV$ väljendatakse sageli protsendina. Nende andmete korral on standardhälve umbes 5,7% keskväärtusest, st $CV = 5,7\%$.

### 2.2.8 Keskväärtuse standardviga (SEM)

Keskväärtuse standardviga (SEM) ei ole andmestikus varieeruvuse hulga mõõt. Pigem on ta üldkogumi keskväärtuse $\bar{x}$ kui hinnangu täpsuse mõõt. Kui palju eeldame keskväärtuse $\bar{x}$ varieerumist?

Keskväärtuse standardviga sõltub:

1) dispersioonist,

2) valimi suurusest,

3) valimi moodustamise viisist.

Juhusliku valimi keskväärtuse standardvea võime arvutada järgmiselt

$$SEM = s_{\bar{x}} = \frac{s}{\sqrt{n}}.$$

Näite 1 andmete korral $SEM = s_{\bar{x}} = \frac{s}{\sqrt{n}} = \frac{0{,}099814}{\sqrt{16}} = 0,02495 \approx 0,025.$

### 2.2.9 Kvartiilid ja protsentiilid

Nagu mediaan jagab andmed kahte ossa, võime defineerida esimest kvartiili $q_1$ kui väärtust, millest 25% valimi vaatlusi on sattunud ettepoole. Mediaan on teine kvartiil. Kolmas kvartiil $q_3$ on väärtus, millest 75% valimi vaatlusi on sattunud ettepoole. Ka teisi protsentiile võime kasutada; võime rääkida 20-st protsentiilist (või 95-st protsentiilist) kui väärtusest, millest 20% (95%) valimi vaatlusi on sattunud ettepoole.

Interpoleerimise juures on sageli vaja arvutada valimist kvartiile ja protsentiile.

Näite 1 andmete korral $q_3$ asukoht on $\frac{3(n+1)}{4} = \frac{51}{4} = 12,75$, st kolmanda kvartiili asukoht on vaatlus numbriga 12,75. Vaatlusel numbriga 12 on väärtus 1,79 ja vaatlusel numbriga 13 on väärtus 1,81. Interpoleerime nende kahe vaatluse vahel ja saame kolmanda kvartiili järgmiselt:
$1,79 + 0,75(1,81 - 1,79) = 1,805$.

Esimese kvartiili $q_1$ asukoht arvutatakse $\frac{n+1}{4} = \frac{17}{4} = 4,25$. Vaatlusel numbriga 4 on väärtus 1,68 ja vaatlusel numbriga 5 on samuti väärtus 1,68, seega $q_1 = 1,68$.

### 2.2.10 Kvartiilide vahe

Kvartiilide vahe on andmestikus hajuvuskarakteristik. See on alternatiivne standardhälbele. Kvartiilide vahe arvutatakse

$$v = q_3 - q_1.$$

Näite 1 andmete korral $v = 1,805 - 1,68 = 0,125$.

### 2.2.11 Arvuti programmipaketid kirjeldava statistika jaoks

Arvuti programmipaketid nagu Minitab ja SAS arvutavad eelnevas kirjeldatud suurusi. Näite 1 andmete korral annab Minitab järgmise väljatrüki:

```
Descriptive Statistics
Variable  N       Mean    Median  Tr Mean  StDev   SE Mean
HAPNIK    16      1.7419  1.7500  1.7414   0.0998  0.0250

Variable  Min     Max     Q1      Q3
HAPNIK    1.5700  1 9200  1.6800  1.8050
```

Märkus. TrMean on keskmine, kui andmestikust on 5% väikseimaid ja 5% suurimaid vaatlusi välja jäetud.

SAS paketi protseduur MEANS annab järgmise väljatrüki (kui kvartiilid või teised statistikud on vajalikud, siis sellised statistikud tuleb eraldi arvutatada).

*Analysis Variable : HAPNIK*

| N | Mean | Std Dev | Minimum | Maximum |
|---|---|---|---|---|
| 16 | 1.7418750 | 0.0998144 | 1.5700000 | 1.9200000 |

Võime kasutada SAS protseduuri UNIVARIATE, mis annab väljatrükis suurema hulga statistikute väärtusi:

Univariate procedure

Variable=HAPNIK

Moments

| | | | |
|---|---|---|---|
| N | 16 | Sum Wgts | 16 |
| Mean | 1.741875 | Sum | 27.87 |
| Std Dev | 0.099814 | Variance | 0.009963 |
| Skewness | -0.09452 | Kurtosis | -0.4475 |
| USS | 48.6955 | CSS | 0.149444 |
| CV | 5.73086 | Std Mean | 0.024954 |
| T:Mean=0 | 69.80455 | Pr>\|T\| | 0.0001 |
| Num ^= 0 | 16 | Num > 0 | 16 |
| M(Sign) | 8 | Pr>=\|M\| | 0.0001 |
| Sgn Rank | 68 | Pr>=\|S\| | 0.0001 |

Quantiles (Def=5)

| | | | |
|---|---|---|---|
| 100% Max | 1.92 | 99% | 1.92 |
| 75% Q3 | 1.8 | 95% | 1.92 |
| 50% Med | 1.75 | 90% | 1.88 |
| 25% Q1 | 1.68 | 10% | 1.6 |
| 0% Min | 1.57 | 5% | 1.57 |
| | | 1% | 1.57 |
| Range | 0.35 | | |
| Q3-Q1 | 0.12 | | |
| Mode | 1.6 | | |

Extremes

| Lowest | Obs | Highest | Obs |
|---|---|---|---|
| 1.57 ( | 1) | 1.79 ( | 6) |
| 1.6 ( | 4) | 1.81 ( | 3) |
| 1.6 ( | 2) | 1.85 ( | 8) |
| 1.68 ( | 14) | 1.88 ( | 5) |
| 1.68 ( | 10) | 1.92 ( | 9) |

Märkus. Selle väljatrüki mõned lühendid vajavad selgitust. Skewness ja Kurtosis on jaotuse kuju mõõdud; USS ("Uncorrected sum of squares") on $\sum x^2$; CSS ("corrected sum of squares) on $\sum (x-\bar{x})^2$; CV on variatsioonikoefitsent; Std Mean on keskväärtuse standardviga. Med on mediaan ning Q1 ja Q3 on kvartiilid. Ülejäänud lühendid on kas arusaadavad või praegu mitteolulised.

## 2.3 Numbrilise andmestiku graafiline esitlus

### 2.3.1 Tüvi-leht-diagramm

Tüvi-leht-diagramm on kiire ülevaade numbrilise andmestiku hulgast. Teda on sobiv kasutada, kui tunnuse väärtuste hulk pole väga suur. Illustreerime selle ideed näite 1 andmete põhjal. Tüvi-leht-diagrammi koostamiseks on esiteks vaja valida “tüve“ suurus. Käesolevas näites kasutame “tüvena“ täisosa ning esimest kümnendkohta. Kirjutame üksteise alla kasvavasse järjekorda kõikvõimalikud “tüved“ Iga väärtuse ülejäänud kümnendkohad kirjutame kõrvale vertikaalsesse tulpa:

| "Tüvi" | "Leht" | | "Tüvi" | "Leht" |
|---|---|---|---|---|
| 1,5 | 7 | Teeme “ilusa“ tabeli, | 1,5 | 7 |
| 1,6 | 0088 | võime igas “tüves“ | 1,6 | 0088 |
| 1,7 | 9553268 | sorteerida vaatlused | 1,7 | 2355689 |
| 1,8 | 185 | suuruse järgi. Ühtlasi | 1,8 | 158 |
| 1,9 | 2 | kergendab see | 1,9 | 2 |
| | | kvartiilide arvutamist jne. | | |

Tüvi-leht-diagrammid on kasutatavad, kui soovime saada kiiret ülevaadet numbrilisest andmestikust. Need annavad üheaegselt andmestikust tabeli ja graafiku.

### 2.3.2 Karpdiagramm

Näite 1 andmete põhjal karpdiagrammi koostamiseks on vaja eelnevalt arvutada järgmised statistikud:

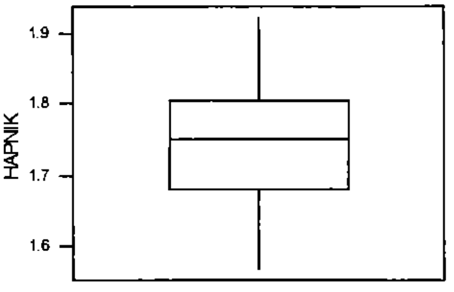
esimene kvartiil $q_1 = 1,64$ ,

teine kvartiil $q_2$, mis on ka mediaan $Me = q_2 = 1,75$,

kolmas kvartiil $q_3 = 1,805$,

kvartiilide vahe $\mathrm{v} = q_3 - q_1 = 1,805 - 1,64 = 0,165$.

Karpdiagrammi kujutamiseks joonistame kõigepealt vertikaalse skaala. Horisontaalsed jooned tõmbame kvartiilide $q_1$, $q_2$=Me ja $q_3$ kohalt. Ühendame need jooned karbiks. Karbist väljapoole joonistame "vurrud", need on jooned, mis ühendavad karbiga vaatluse vähimat ja suurimat väärtust. Järgnev karpdiagramm on koostatud näite 1 andmetele, selle koostamiseks on kasutatud programmipaketti Minitab.




Märkus. Karpdiagrammi võib ka joonistada horisontaalselt.

Karpdiagrammi "vurrud" on pikendus vähima ja suurima vaatluseni, nad on piiratud seestpoolt alumise ja ülemise piiriga. Alumist ja ülemist piiri analüütilise andmestiku žargoonis kutsutakse "seesmiseks aiaks" Need on defineeritud järgmiselt:

alumine piir: $q_1 - 1,5$v,

ülemine piir: $q_3 + 1,5$v.

Vaatlused, mis on "seesmisest aiast" väljaspool, joonistatakse eraldi ja võivad tulla arvesse kui "leebed eemalasujad" Mõned programmid joonistavad diagrammile veel teisigi punkte. Vaatlusi, mis asuvad väljaspool, kutsutakse "väliseks aiaks", joonistatakse diagrammile mingi muu sümboliga ja neid vaadeldakse kui "ekstreemsed eemalasujad" Alumine ja ülemine "väline aed" defineeritakse vastavalt $q_1 - 3$v ja $q_3 + 3$v. Karpdiagrammi vastav näide koos eemalasujatega on toodud edaspidi selles peatükis.

### 2.3.3 Sagedustabel

Sagedustabeliks nimetatakse tabelit, milles on ära näidatud tunnuse erinevad väärtused ja nende väärtuste esinemiste arv (sagedus) uuritavas andmestikus. Kui tunnusel on palju erinevaid väärtusi, on sobiv koostada andmestikust ülevaate saamiseks sagedustabel (jaotustabel), kus andmed on jaotatud klassideks, ja lugeda, mitu vaatlust satub vastavasse klassi. Mõned üldised juhised, mida vaja arvestada:

- Valida klassid selliselt, et vaatlused oleksid üheselt klassidesse jaotatud.
- Enamikel juhtudel on soovitav kasutada 5-10 klassi.

- Praktikas sobivad klassipiirideks tihti nn ümmargused, näit nulli või viiega lõppevad arvud.

- Vältida avatud klasse.

- Kui võimalik, valida klassid võrdse laiusega.

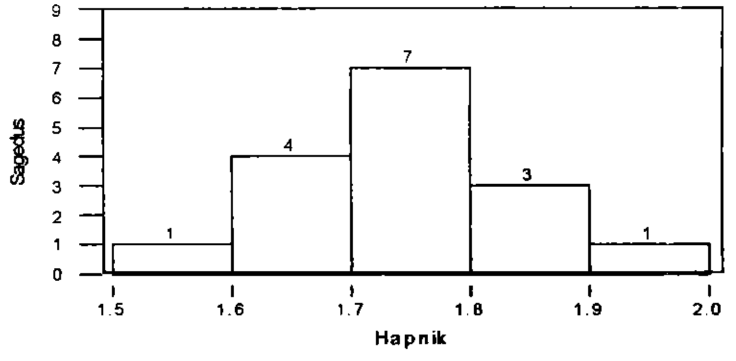
Sagedustabel näite 1 andmete jaoks, kus andmed on jaotatud viieks võrdse laiusega klassiks, on järgmine:

| Klassipiirid | Keskpunkt | Sagedus |
|---|---|---|
| [1,50 −1,60) | 1,55 | 1 |
| [1,60 −1,70) | 1,65 | 4 |
| [1,70 −1,80) | 1,75 | 7 |
| [1,80 −1,90) | 1,85 | 3 |
| [1,90 −2,00) | 1,95 | 1 |

### 2.3.4 Histogrammid

Histogramm ehk tulpdiagramm on graafik, millel kujutatakse andmestiku sagedustabelit. Iga klass on esitatud tulbana. Tulba kõrgus on võrdeline sagedusega, st on võrdne vaatluste arvuga selles klassis. Märkusena olgu öeldud, et tulba kõrgus on võrdeline sagedusega ainult siis, kui klassid on võrdse laiusega.

Histogrammil peavad tulbad asuma kõrvuti, ei tohi jätta tühja ruumi kahe tulba vahele, näidates, et kõik väärtused antud vahemikus on võimalikud. (Mõned programmid teevad histogrammid sellest erinevalt, aga näiteks Excel ja Minitab joonistavad tulbad sel teel.) Eelnevate andmete põhjal on histogramm järgmine:




### 2.3.5 Histogrammid kui klassid ei ole võrdse laiusega

Eelpool soovitasime, et histogrammi tulba kõrgus peab olema võrdeline sagedusega. Püüame selgitada, miks see on nii ja kuidas joonistada histogrammi, kui klassid ei ole võrdsed.

Näide. Oletame, et soovime illustreerida 30 farmi pindala jaotusi järgmise tabeli järgi:

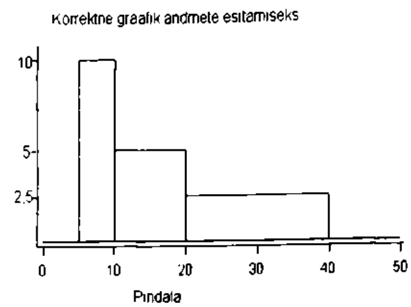
| Pindala | Farmide arv |
|---|---|
| 5 – 10 ha | 10 |
| 10 – 20 ha | 10 |
| 20 – 40 ha | 10 |

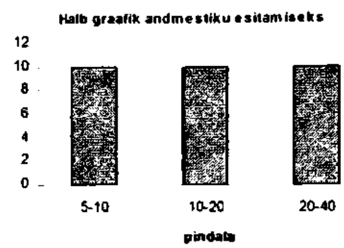
Igas klassis on ühesugune arv farme, aga klassid pole pindalalt võrdsed. Kui jagame andmed võrdse laiusega klassidesse, eeldame farmide võrdset jaotust kõigis klassides, saame:

| Pindala | Farmide arv |
|---|---|
| 5 – 10 ha | 10,0 |
| 10 – 15 ha | 5,0 |
| 15 – 20 ha | 5,0 |
| 20 – 25 ha | 2,5 |
| 25 – 30 ha | 2,5 |
| 30 – 35 ha | 2,5 |
| 35 – 40 ha | 2,5 |

Järelikult, järgnev vasakpoolne graafik, mis on koostatud esimese tabeli põhjal, näitab jaotuse moonutatud pilti, aga graafik paremal pool on parem.

Seal on kõik tulbad võrdse pindalaga ning esile on toodud fakt, et farmide arvukus on võrdne kõikides klassides.

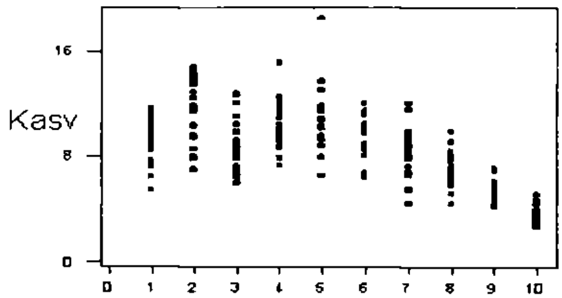

Parempoolne alumine graafik on pigem halb. Klassid on kujutatud võrdse laiusega, ehkki nendel laius ei ole võrdne.

## 2.3.6 Hajuvusdiagramm

Tüvi-leht-diagrammid, histogrammid ja karpdiagrammid iseloomustavad enamasti ainult ühte tunnust. Vahel on huvitav uurida mitme numbrilise tunnuse vahelisi seoseid. Kujutame graafikul ühe tunnuse väärtusi horisontaalsel teljel ja mõne teise tunnuse väärtusi vertikaalsel teljel, kus iga vaatlus tähendab punkti graafikul: moodustuvat graafikut nimetatakse hajuvusdiagrammiks ehk korrelatsiooniväljaks.

Näide. Järgneva graafiku andmed on võetud uurimusest[2], kus mõõdeti seemikute juurekasvu. Seemikuid väetati erinevate 2-(2,4-diklorofenoksüü)-propioonhappe (Dichloroprop) annustega. Annused on graafikul logaritmitud. Graafikul näeme suurt hulka varieeruvust, aga näeme ka süstemaatilist mudelit: keskmine juure kasv väheneb annuse kasvades.

[2]Ulla Didon and Ulf Olsson (1997): A bioassay method for detection of Dichlorprop and MCPA in water. Swedish J. Agr. Res., 27: 121-128.

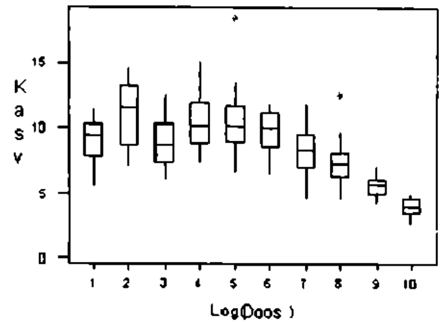



### 2.3.7 Varieeruvuse esitamine graafikul

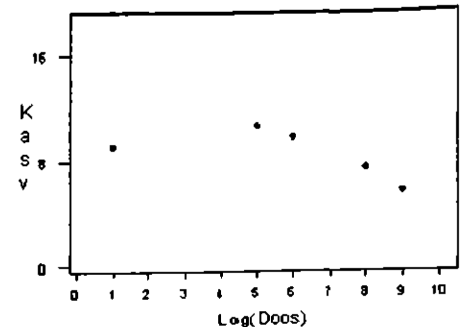
Mõnikord suure andmestiku korral võib vajalik informatsioon graafikul kaduma minna. Näiteks joonisel on liiga palju punkte ja mudel on raskesti nähtav Sel juhul on kasulik mingil viisil summeerida vaatlused ja kujutada graafikul ainult summad. Mõned alternatiivid:

- kujutada graafikul ainult keskväärtusi,
- koostada karpdiagramm iga x-i väärtuse jaoks,
- joonistada iga x-i väärtuse juurde keskväärtus ± standardhälve,
- joonistada iga x-i väärtuse juurde keskväärtus ± keskväärtuse standardviga.

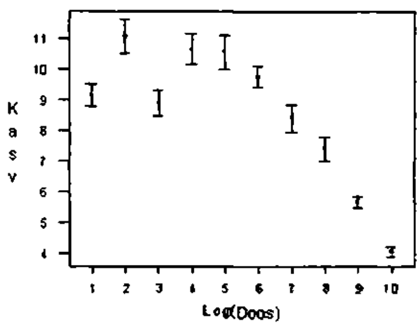
Järgnevatel graafikutel on esitatud sama andmestiku põhjal need neli alternatiivi. Segaduste vältimiseks on joonisel kindlasti vaja märkida missugust alternatiivi kasutame. Meie näite korral kasutame vastavalt karpdiagrammi, ainult keskväärtust, Mean±SEM ja Mean±$s$.



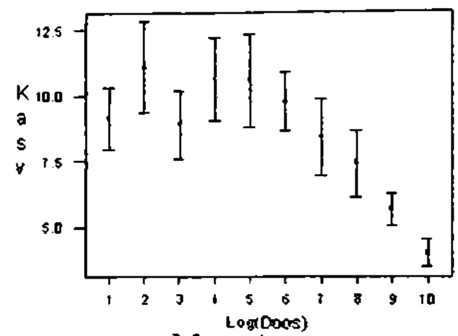

Karpdiagramm




Ainult keskmised




Mean±SEM




Mean±s

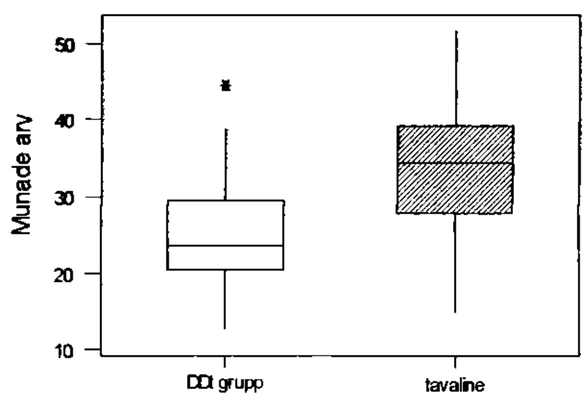
### 2.3.8 Teised bivariantsed graafikud

Mõnikord on kasulik kombineerida erinevaid diagrammide tüüpe. Üheks näiteks võiks olla erinevate gruppide karpdiagrammid (näiteks erinevate töötluste) esitamine ühel ja samal graafikul. Sellel on siis kvalitatiivne muutuja horisontaalteljel ja kvantitatiivne muutuja on vertikaalteljel.

Näide. Uuriti kahte gruppi puuviljakärbseid. Üks grupp valiti selline, mis on vastupidav DDT suhtes, teine grupp oli normaalne selekteerimata grupp. Vaatlustena fikseeriti iga emase poolt munetud keskmine munade arv päevas ning mõõtmisi teostati kahe nädala jooksul.

| Grupp 1 | | Grupp 2 | |
|---|---|---|---|
| 12.8 | 27.3 | 35.4 | 22.6 |
| 21.6 | 22.4 | 27.4 | 40.4 |
| 14.8 | 27.5 | 19.3 | 34.4 |
| 23.1 | 20.3 | 41.8 | 30.4 |
| 34.6 | 38.7 | 20.3 | 14.9 |
| 19.7 | 26.4 | 37.6 | 51.8 |
| 22.6 | 23.7 | 36.9 | 33.8 |
| 29.6 | 26.1 | 37.3 | 37.9 |
| 16.4 | 29.5 | 28.2 | 29.5 |
| 20.3 | 38.6 | 23.4 | 42.4 |
| 29.3 | 44.4 | 33.7 | 36.6 |
| 14.9 | 23.2 | 29.2 | 47.4 |
| | 23.6 | 41.7 | |

Missugune on kahe grupi vahelise viljakuse erinevus? Vastuse küsimusele võime välja lugeda järgnevalt jooniselt:




Olgugi et kaks gruppi osaliselt katavad sama ala, näeme, et grupil 1 on madalam viljakus kui grupil 2. Hiljem vaatleme formaalseid viise, kuidas seda järeldust hinnata.

## 2.4 Kvalitatiivse andmestiku kokkuvõte

### 2.4.1 Sagedustabel

Lihtsaim võimalus ülevaate saamiseks diskreetsest andmestikust (numbrilise või ka nominaalse skaala andmestikust) on koostada sagedustabel. Selle paremaks illustreerimiseks kasutame näidet. Märgiti üles erinevat liiki puude arvukus Lansing Wood'i, Michigan[3]metsas. Tulemus on esitatud tabelis, mis ongi sagedustabel:

| Liik | Puude arv |
|---|---|
| Punane tamm | 346 |
| Valge tamm | 448 |
| Must tamm | 135 |
| Hikkoripuu | 703 |
| Vaher | 514 |
| Teised liigid | 105 |
| Kokku | 2251 |

### 2.4.2 Risttabel

**Risttabel on tabel, kus ühikud jaotatakse kasutatavate kahe (või rohkema) kriteeriumi järgi. Risttabelit kasutatakse tunnustevaheliste sõltuvuste leidmiseks.**

Näide.

Uuriti, kas leidub sõltuvust norskamise ja südamehaiguste vahel[4] Andmeid koguti 2484 patsiendilt, kellel oli mingi südamehaigus ning nendelt küsiti, kui suurel määral nad öösiti norskavad. Saadud andmestik on järgmine:

| Kas on südamega probleem? | Norskavad harva/mitte kunagi | Norskavad sageli/alati | Kokku |
|---|---|---|---|
| Ja | 59 | 51 | 110 |
| Ei | 1958 | 416 | 2374 |
| Kokku | 2017 | 467 | 2484 |

[3]Anmed on võetud: P. G. N. Digby and R. A. Kempton (1987): Multivariate analysis of ecological communities. London, Chapman & Hall.

[4]Andmed on võetud: P G. Norton and E. V Dunn (1985): Snoring as a risk factor for disease. British Medical Journal, pp. 630-632.

Mõnikord on risttabelit raske lugeda, seepärast on kasulik esitada arvud protsentidena. Selleks on vaja läbida kolm järgnevat sammu:

1. Arvutada kokkuvõtted protsentidena (antud juhul protsent 2484-st).

2. Arvutada reakokkuvõtted protsentidena. 51 patsienti südamehaigustega moodustab $\frac{51}{467}$ $100 = 10,9\%$ kõigist, kes norskavad sageli/alati.

3. Arvutada veerukokkuvõtted protsentidena. Kõigist, kel on südamehaigusi, norskasid sageli või alati $\frac{51}{110}$ $100 = 46,4\%$.

Kumba protsenti arvutada, sõltub sellest, mida on vaja näidata. Käesolevas näites südamehaigustega patsientidest norskasid 46% sageli/alati. Samal ajal kui 18% kontrollgrupist norskasid. See viib mõttele, et on mingi seos norskamise ja südamehaiguste vahel. Hiljem uurime seda, kas see on "reaalne" või võib seda selgitada juhuslikkusena.

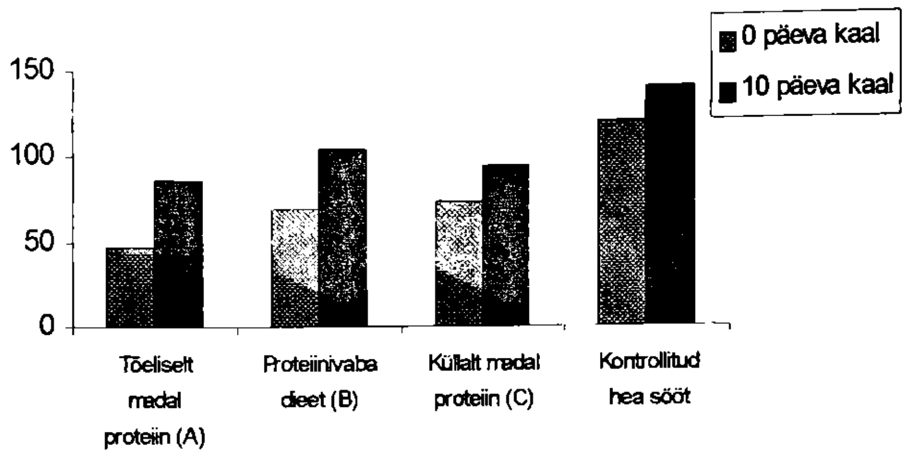
### 2.4.3 Tulpdiagrammid

Järgnevas tabelis on esitatud andmed hiirte kehakaalu kohta, keda söödeti erinevate toiduratsioonidega[5] Andmed on kogutud eksperimendi esimesel päeval (0. päev) ja peale kümnendat päeva (10. päev).

| Grupp | 0. päeva kaal (g) | 10. päeva kaal (g) |
|---|---|---|
| Tõeliselt madal proteiin (A) | 46 | 86 |
| Proteiinivaba dieet (B) | 70 | 105 |
| Küllalt madal proteiin (C) | 73 | 95 |
| Kontrollitud hea sööt | 120 | 140 |

Tulpdiagramm, mis kirjeldab hiirte juurdekasvu, on järgmine:

[5] Andmed on võetud: R.O. Lawal: The effect of a single oral dose of polyphenols obtained from the outercoat of the fruit of Treculia africana in protein-deficient rats. Food Chemistry 44 (1992) 321-323.



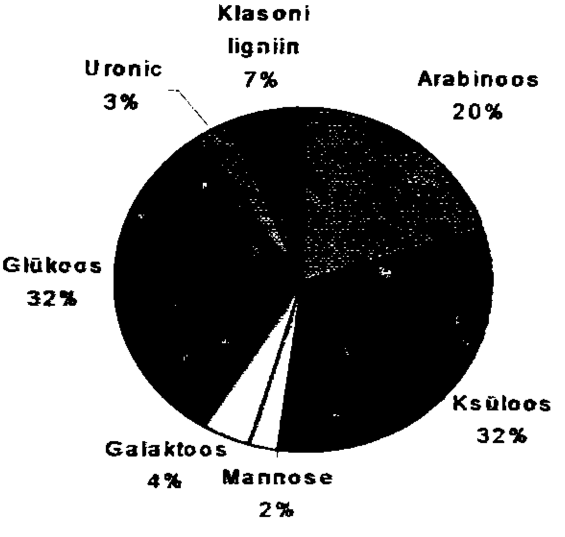

### 2.4.4 Sektordiagramm (Pie graafik)

Kevadise nisu kiudainete sisaldus on toodud järgmises tabelis[6]:

| Dieedi kiud | Keskmine (%) |
|---|---|
| Arabinosi sete | 2,30 |
| Xylosi sete | 3,60 |
| Mannosi sete | 0,28 |
| Galactosi sete | 0,49 |
| Glucosi sete | 3,60 |
| Uronic happe sete | 0,35 |
| Klason ligniin | 0,74 |

Nende andmete kujutamiseks sektordiarammil arvutame erinevate komponentide protsendid ja transleerime need nurgakraadidesse.

[6]Andmed on võetud: Per Åman: The variation in chemical composition of Swedish wheats. Swedish J.agric.Res.18:27-30, 1988.


Klasoni
ligniin
7%
Uronic
3%
Arabinoos
20%
Glükoos
32%
Ksüloos
32%
Galaktoos
4%
Mannose
2%

## 2.5 Ülesanded

**2.1**

20 taimest koosneva valimi kasvud (cm-tes) mõõdeti järgmiselt:

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 12 | 16 | 24 | 13 | 8 | 15 | 22 | 18 | 14 | 10 |
| 18 | 26 | 17 | 12 | 11 | 17 | 21 | 18 | 16 | 5 |

A. Kujuta andmestik sagedustabelina, moodustades selleks eelnevalt sobivad klassid.

B. Kasuta saadud sagedustabelit ja joonista selle põhjal histogramm.

C. Arvuta keskmine, dispersioon, standardhälve ja mediaan.

**2.2**

Emaseid rotte (28-84 päeva vanad) söödeti kindla toiduga. Nad kaaluti ja saadi järgmised andmed:

134 146 104 119 124 101 107 83 113 129 97 123 92 118

A. Arvuta keskväärtus.

B. Arvuta dispersioon.

C. Kirjelda andmed, kasutades stem-and-leaf diagrammmi.

D. Arvuta mediaan.

E. Koosta kastdiagramm.

**2.3**

Uurija kasvatas kasvuhoone ühel peenral 15 tomatitaime. Peale 21. päeva mõõtis ta taimede varrepikkused (cm), mis olid järgmised:

| | | | | |
|---|---|---|---|---|
| 12,4 | 10,9 | 11,8 | 14,1 | 12,6 |
| 12,2 | 12,2 | 13,5 | 12,7 | 11,9 |
| 13,4 | 12,1 | 12,0 | 13,2 | 13,1 |

A. Konstrueeri nende andmete põhjal tüvi-leht-diagramm.

B. Arvuta keskväärtus.

C. Arvuta dispersioon.

D. Arvuta mediaan.

E. Koosta karpdiagramm.

**2.4**

Katses sooviti uurida, kuidas mõjub valgus sinepitaimede juurekasvule. 20 sinepitaime istutati pottidesse. 10 juhuslikult valitud taime pagutati pimedasse, ülejäänud 10 taime jäeti kasvama päevavalgusesse. Mõne aja möödudes mõõdeti iga taime juure pikkust ja saadi järgmised tulemused:

| Taimed päevavalgusest | Taimed pimedast |
|---|---|
| 21 | 22 |
| 39 | 16 |
| 31 | 20 |
| 13 | 14 |
| 52 | 32 |
| 39 | 28 |
| 55 | 36 |
| 50 | 41 |
| 29 | 17 |
| 17 | 22 |

Küsimused:

A. Arvuta keskväärtus mõlemas grupis.

B. Arvuta dispersioon mõlemas grupis.

C. Koosta karpdiagramm nii, et mõlemad grupid oleksid samal diagrammil.

D. Tee tüvi-leht-diagrammid kahele grupile, kasutades ühesugust "tüve" mõlemale grupile.

# Peatükk 3

# Tõenäosus

## 3.1 Mõned kontseptsioonid ja märkused

Tõenäosusteooria sisaldab arvutusi ja tõenäosuslikke otsuseid selle kohta, mis võib juhtuda. See on väärtuslik bioloogidele, näiteks geneetikutele. Selle abil formuleeritakse ka metoodilised alused järelduste tegemiseks andmete kohta. Tõenäosusteooria põhimõisteks on sündmus, mis defineeritakse katse abil.

**Definitsioon 7** *Juhuslikuks katseks nimetatakse katset, mille tulemus ei ole katsetingimustega üheselt määratud. Katse on suvaline arv korratav, tal on lõplik arv võrdvõimalikke tulemusi.*

Juhusliku katse kirjeldamisel ei lähtuta üksikust katsetulemusest, vaid sageli ilmneb katse paljukordsel kordamisel mingi seaduspärasus. Katse korraldamise tingimused tuleb enne katse sooritamist fikseerida.

Näide. Kui me viskame täringut, siis katsetulemuste hulk on {1, 2, 3, 4, 5, 6}. Resultaat "6" on tulemus.

**Definitsioon 8** *Sündmuseks nimetame iga katsetulemuste hulka.*

Ütleme, et sündmus A toimub, kui katse tulemusena esineb mõni selle võimalikest sündmuse katsetulemustest. Sündmus on määratud parajasti siis, kui on fikseeritud katse ja määratletud tema katsetulemuste hulk. Seega ka iga üksik katsetulemus võib olla sündmus.

Näide. Võime defineerida sündmuse A kui "täringu kõik paarisarvulised arvud". Sel juhul on sündmus $A = \{2, 4, 6\}$.

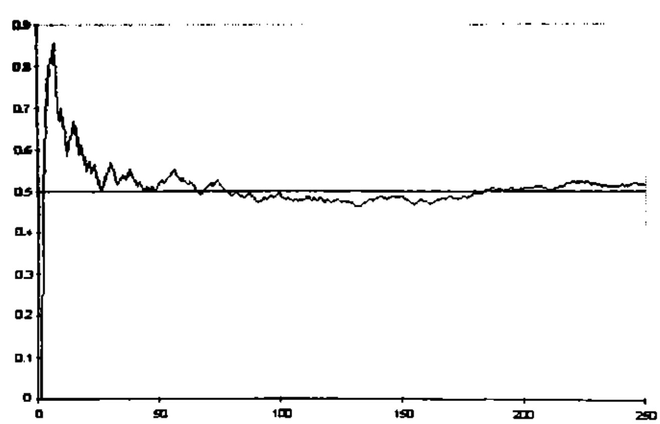
**Definitsioon 9** *Täiend mingile sündmusele A sisaldab kõiki tulemusi, kus A ei esine. Sageli tähistatakse seda* $\overline{A}$.

Näide. Kui $A = \{2, 4, 6\}$, siis esitatud näites täiend on $\overline{A} = \{1, 3, 5\}$

## 3.2 Tõenäosuse üldmõisted

### 3.2.1 Tõenäosus kui suhteline sagedus

Näide. Oletame, et viskame münti. Märgime tulemused järgnevalt: mündi esikülje saamise korral 1 ning tagakülje korral 0. Esimesel viskel saime mündi tagakülje. Peale esimest viset on mündi esikülgede suhteline sagedus 0. Teise viske tulemus on mündi esikülg. Kui me tähistame mündi esikülje ilmumise arvu $f_n$ ja kõikide visete arvu $n$, siis saame $f_n/n = 0,5$. Mündi esikülje ilmumise suhteline sagedus on 0,5. Võime jätkata mündi viskamist. Peale 250 viset saame näiteks:



Näeme, et suhteline sagedus (proportsioon) stabiliseerub arvu 0,5 ümbruses. Võime oletada, et tõenäosus selleks, et viskel tuleb mündi esikülg, on 0,5. Tõenäosuse ebatäpne definitsioon võiks olla $P = \lim_{n\to\infty} \frac{f_n}{n}$ See on Bernoulli poolt tõestatud suurte arvude seadus, mis väidab suhteliste sageduste jada koondumist tõenäosuseks.

### 3.2.2 "Soodsad" ja "võimalikud" tulemused

Tõenäosuse teine kontseptsioon on järgmine: oletame, et mingis katses saadi $n$ võimalikku tulemust. Nendest $f$ on soodsad. Missugune on tõenäosus, et üks soodne tulemus esineb?

Vastus on: kui erinevad tulemused on võrdtõenäosed ehk võrdvõimalikud, siis soodsate tulemuste saamise tõenäosus

$$P = \frac{f}{n}.$$

Näide. Meil on 100 loteriipiletit, millest igal piletil on ühesugune võimalus võiduks. Tõmbame nendest 5 piletit. Tõenäosus selleks, et võidame, on $\frac{5}{100} = 0,05$.

### 3.2.3 Aksiomaatiline tõenäosus

Matemaatikud eelistavad sageli tõenäosuse defineerimiseks kasutada kolme aksioomi.

1. Mingi sündmuse $A$ tõenäosus $P(A)$ on arv, mis asub vahemikus

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

2. Mingi liitsündmuse tõenäosus on iga üksiku sündmuse tõenäosuste summa.

Näide. Tõenäosus, et täringuviskel tuleb paarisarv, on

$$P(2) + P(4) + P(6).$$

3. Iga sündmuse esinemistõenäosus on 1.

Näide. Tõenäosus, et täring näitab ühte numbritest (1, 2, 3, 4, 5, või 6), on 1[1]

## 3.3 Mõned tõenäosuse reeglid

### 3.3.1 Liitmise reegel

Kui $A$ ja $B$ on kaks sündmust, siis tõenäosus selleks, et toimub üks sündmustest $A$ või $B$ või mõlemad ($A$ ja $B$) on $P(A) + P(B) - P(A \text{ ja } B)$.

Märkus. Loogiline *ja* tähistatakse sageli märgiga $\cap$ ja loogiline *või* tähistatakse $\cup$. Liitmise reegli võib kirjutada järgmiselt

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

[1] Jätame tähelepanemata võimaluse, et viskel võib täring kaduma minna.

Näide. Teatav geen päritakse vastavalt järgmistele tõenäosustele:

| | Mees | Naine | Kokku |
|---|---|---|---|
| Omab geeni | 0,167 | 0,083 | 0,250 |
| Ei oma geeni | 0,333 | 0,417 | 0,750 |
| Kokku | 0,500 | 0,500 | 1,000 |

Mis on tõenäolisem, kas sündiv indiviid on mees või omab geeni või mõlemad - on mees ja omab geeni?

Vastus. Tõenäosus $P(A)$, et sündiv indiviid on mees, on 0,5. Tõenäosus $P(B)$, et see indiviid omab teatud geeni, on 0,25. Tõenäosus $P(A \text{ ja } B) = P(A \cap B)$, et indiviid on mees ja omab teatavat geeni, on 0,167 Seega kahe sündmuse koos esinemise tõenäosus on $0,500 + 0,250 - 0,167 = 0,583$.

### 3.3.2 Sõltumatud sündmused

Kui ühe sündmuse toimumune ei mõjuta teise sündmuse toimumise tõenäosust, siis need sündmused on sõltumatud. Järelikult, kaks sündmust $A$ ja $B$ on sõltumatud, kui kehtib võrdus

$$P(A \cap B) = P(A) \quad P(B).$$

Näide. Eelneva näite põhjal küsime, kas sündmused, geeni olemasolu ja sugu, on sõltumatud?

Vastus. Kaks sündmust on sõltumatud, kui $P(\text{mees ja omab geeni}) = P(\text{mees}) \quad P(\text{omab geeni}) = 0,5 \quad 0,25 = 0,125$. Tabelist saame, et $P(\text{mees ja omab geeni}) = 0,167$ Seega 0,167 $\neq$ 0,125. Järelikult, geeni olemasolu ja sugu ei ole sõltumatud.

### 3.3.3 Sõltumatute sündmuste korrutamise reegel

Kui $A$ ja $B$ on sõltumatud sündmused, siis nende sündmuste korrutise tõenäosus võrdub tegursündmuste tõenäosuste korrutisega.

$$P(A \cap B) = P(A) \quad P(B).$$

Näide. Teatavat liiki seemne idanevus on 90%, st kõikidest seemnetest 90% hakkavad idanema. Külvati kaks seemet. Missugune on tõenäosus, et mõlemad seemned idanevad?

Lahendus. $P(\text{mõlemad seemned idanevad}) = P(\text{esimene seeme idaneb})$ $P(\text{teine seeme idaneb}) = 0,9 \quad 0,9 = 0,81$.

Näide. Sama liigi seemneid külvati 10 seemet. Kui suur on tõenäosus, et kõik seemned hakkavad idanema?

Lahendus. Kasutame korrutamise reeglit korduvalt, selle põhjal $P$(kõik seemned idanevad) = $P$(esimene seeme idaneb) $P$(teine seeme idaneb) ... $P$(kümnes seeme idaneb) = 0,9 0,9 0,9 0,9 0,9 0,9 0,9 0,9 0,9 0,9 $= 0,9^{10}$ = 0,3487

### 3.3.4 Tinglik tõenäosus

Tõenäosus selleks, et sündmus $A$ toimub tingimusel, et sündmus $B$ on toimunud, tähistatakse $P(A|B)$.

Märkus. Sõltumatute sündmuste korral $P(A|B) = P(A)$. Tõenäosus ei tohi sõltuda $B$-st.

### 3.3.5 Bayes'i teoreem

Oletame, et on toimunud sündmus $B$ ning me teame selle sündmuse toimumise tõenäosust $P(B)$. Nüüd toimus sündmus $A$, mille tulemusena võime esialgsele sündmusele arvutada täpsustatud tõenäosuse Bayes'i valemiga

$$P(A|B) = \frac{P(A \cap B)}{P(B)}.$$

Näide. Eelpool vaadeldud pärilikkuse näites kindlad indiviidid on mehed. Missugune on tõenäosus, et need indiviidid omavad teatud geeni?

Lahendus. Tabelist näeme, et tõenäosus $P$(mees ja omab geeni) $= 0,167$. Tabelist saame ka $P$(mees) $= 0,5$. Järelikult tõenäosus, et indiviid omab teatavat geeni, juhul kui ta on mees, on 0,167/0,5 = 0,333. (Üks kolmandik kõikidest meestest omab teatavat geeni.)

### 3.3.6 Korrutamise reegel

Kui $A$ ja $B$ on sündmused (ei pea olema sõltumatud), siis tõenäosus, et $A$ ja $B$ toimuvad, arvutatakse

$$P(A \cap B) = P(A) \quad P(B|A).$$

Näide. Pärilikkuse näites on tõenäosus selleks, et indiviid on mees ja omab teatavat geeni, 0,5 0,333 = 0,167 (mida just arvutasime). Seda näeme ka ilma igasuguste arvutusteta otse tabelist.

Märkus. Bayes'i teoreem ja korrutamise reegel (lihtsas formuleeringus) on kaks erinevat teed sama asja kirjeldamiseks.

## 3.4 Ülesanded

### 3.1

Pärilikkuse näites omab indiviid teatud geeni. Missugune on tõenäosus, et see indiviid on naine?

### 3.2

Missugune on seemnete näites tõenäosus selleks, et 10 seemne hulgast vähemalt üks seeme hakkab idanema?

### 3.3

Ajalehes "Aftonbladet" 20.10.1994. võetud artikkel "'Tervel" naisel oli rinnavähk" Vastavalt artiklile said mõned naised, keda uuriti ultraheliga, vale diagnoosi. 4% naistest, kellel diagnoositi rinnavähk, olid tegelikult terved ja 2% naistel, kes olid terveks tunnistatud, oli rinnavähk.

Eeldame, et 3%-l vaadeldud naistest diagnoositi rinnavähk. Kasuta eelpool toodud informatsiooni ning vasta järgmistele küsimustele:

A. Missugune on tõenäosus, et naisel, kes diagnoositi kui "haige", tegelikult on ka rinnavähk?

B. Missugune on tõenäosus, et juhuslikult valitud naisel on rinnavähk?

C. Eeldame, et naisel on rinnavähk. Missugune on tõenäosus, et seda ei ole avastatud?

# Peatükk 4

# Tõenäosuslikud jaotused

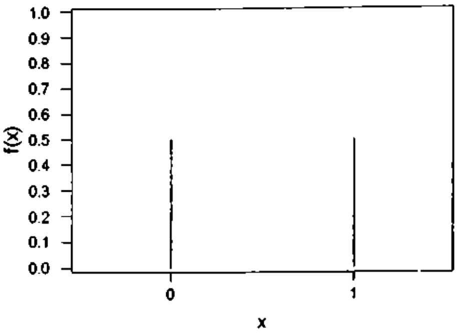
## 4.1 Sissejuhatus

Juhuslik suurus on suurus, mille väärtused on määratud juhusega. Seda võib kirjeldada tõenäosusfunktsiooni või tõenäosusjaotusega. Matemaatiliselt on tõenäosusjaotus funktsioon, mis iseloomustab juhusliku suuruse tõenäosuslikku käitumist. Mõned nendest jaotustest on osutunud sobivateks mudeliteks reaalsete protsesside või nähtuste kirjeldamisel. Järelikult võib tõenäosusfunktsiooni vaadelda kui matemaatilist mudelit üldkogumile, millest on võetud valim.

## 4.2 Diskreetsed juhuslikud suurused

Diskreetset juhuslikku suurust kirjeldatakse tema tõenäosusfunktsiooni f(x) abil. Igale juhusliku suuruse väärtusele seab funktsioon f(x) vastu tema esinemistõenäosuse. Tõenäosusfunktsiooni võib esitada tabelina või valemina, mis näitab juhusliku suuruse ja tema tõenäosuse vahelist seost. Juhusliku suuruse kõikide tõenäosuste summa on 1. Kui tähistame tõenäosusfunktsiooni f(x), siis peab kehtima $\sum f(x_i) = 1$.

Näide Bernoulli jaotuse kohta, kui $p = 0,5$. Vaatleme lihtsat katset mündi viskamise kohta. Kui münt maandub esikülg peal, registreerime selle numbriga 1 ja kui maandub tagapool peal, siis märgime numbriga 0. Järelikult $x = 1$ ja $x = 0$ on kaks võimalikku tulemust. Kui münt on "lennus", on mõlemad tulemused võrdtõenäolised. Tõenäosusjaotuse võime esitada järgmise tabelina ja kujutada graafikul:

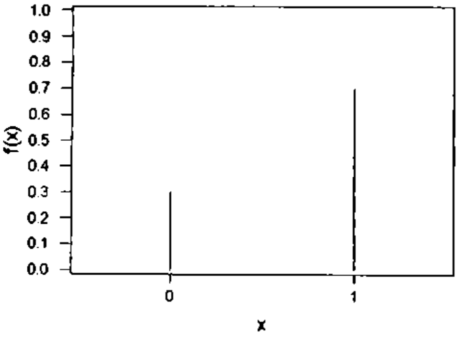



| $x$ | $f(x)$ |
|---|---|
| 0 | 0,5 |
| 1 | 0,5 |

$f(x)$ on kasutusel *tõenäosusfunktsioonina*. See näitab, et juhuslik suurus $X$ omandab väärtuse x, mille tõenäosus on: $f(x) = P(X = x)$. Eelnevas lihtsas näites $f(x) = 0,5$, kui $x = 0$ ja $f(x) = 0,5$, kui $x = 1$ ning kõigil teistel juhtudel $f(x) = 0$.

Märkus. Tavaliselt kasutatakse $X$ juhusliku suuruse tähistamiseks ja juhusliku suuruse väärtuse tähistamiseks tähte $x$. Tähistus $P(X = x)$ tähendab: "Tõenäosus selleks, et juhuslik suurus $X$ omandab väärtuse $x$" Seda tähistust kasutame juhtudel, kui on vaja vahet teha suuruse ja väärtuse vahel. Kui kontekstist on selge, siis kasutame väikest tähte.

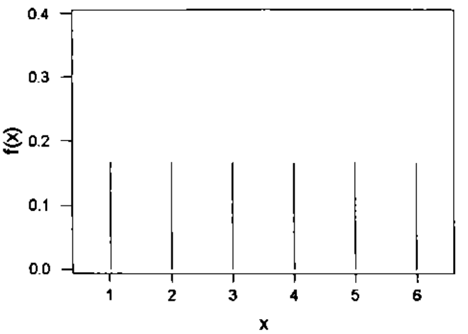
Näide *üldise Bernoulli jaotuse* kohta. Oletame, et tulemused $x = 0$ ja $x = 1$ ei ole võrdtõenäolised $P(x = 1) = p$ ja $P(x = 0) = 1 - p$. See on Bernoulli jaotus, mis sõltub $p$ väärtusest. Jaotuse *parameeter* on $p$.




| $x$ | $f(x)$ |
|---|---|
| 0 | $1 - p$ |
| 1 | $p$ |

Näide täisarvulise jaotuse kohta. Kui viskame täringut, siis võime saada 1, 2, 3, 4, 5 või 6 punkti. Kui täring veel veereb, on kõik tulemused võrdtõenäolised. Järelikult võime tõenäosusjaotuse kujutada järgmiselt:

| $x$ | $f(x)$ |
|---|---|
| 1 | 1/6 |
| 2 | 1/6 |
| 3 | 1/6 |
| 4 | 1/6 |
| 5 | 1/6 |
| 6 | 1/6 |



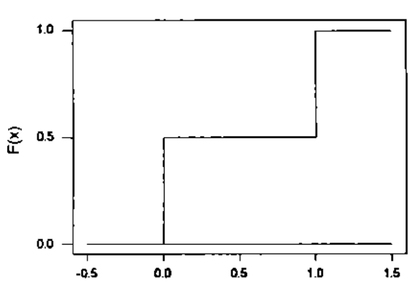

## 4.2.1 Diskreetse juhusliku suuruse jaotusfunktsioon

Jaotusfunktsioon $F(x)$ näitab iga väärtuse $x$ korral tõenäosust, et diskreetne juhuslik suurus $X$ omandab selle väärtuse või on sellest väiksem. Järelikult *jaotusfunktsioon* on

$$F(x) = P(X \leq x) = \sum_{x_i \leq x} f(x_i).$$

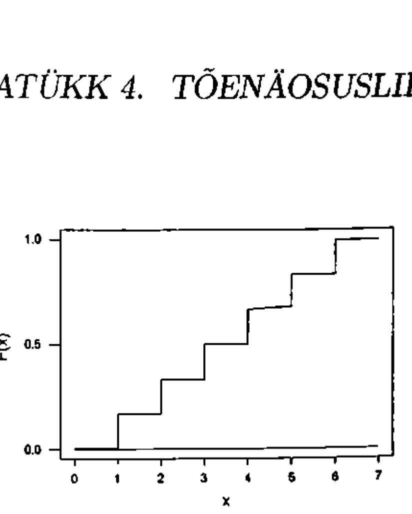
Näide. Järgnevas tabelis ja kõrvaloleval graafikul on esitatud Bernoulli jaotuse korral tõenäosusfunktsioon ja jaotusfunktsioon, kui parameeter on $p = 0,5$.

| $x$ | $f(x)$ | $F(x)$ |
|---|---|---|
| 0 | 0,5 | 0,5 |
| 1 | 0,5 | 1,0 |




Näide. Esitame juhusliku suuruse tõenäosusfunktsiooni ja jaotusfunktsiooni, kus numbrid 1 kuni 6 on võrdtõenäolised. Antud olukord vastab soodsatele täringu visetele.

| $x$ | $f(x)$ | $F(x)$ |
|---|---|---|
| 1 | 1/6 | 1/6 |
| 2 | 1/6 | 2/6 |
| 3 | 1/6 | 3/6 |
| 4 | 1/6 | 4/6 |
| 5 | 1/6 | 5/6 |
| 6 | 1/6 | 6/6 |




Statistiliste jaotuste paljud tabelid on jaotusfunktsioonide tabelid.

### 4.2.2 Diskreetse juhusliku suuruse keskväärtus

Diskreetse juhusliku suuruse keskväärtus on defineeritud kui

$$EX = \sum x \;\; f(x),$$

kus summeeritud on üle kõigi $x$ väärtuste.

Näide. Bernoulli jaotuse korral, kus sündmused $x = 0$ ja $x = 1$ on võrdtõenäolised, on keskväärtus $EX = \sum x \;\; f(x) = 0 \;\; 0,5 + 1 \;\; 0,5 = 0,5$.

Näide. Üldise Bernoulli jaotuse korral, kus $P(x = 1) = p$ ja $P(x = 0) = 1 - p$, on keskväärtus $EX = \sum x \;\; f(x) = 0 \;\; (1 - p) + 1 \;\; p = p$.

Näide. Kui heita täringut, siis keskväärtus on
$EX = 1 \;\; \frac{1}{6} + 2 \;\; \frac{1}{6} + 3 \;\; \frac{1}{6} + 4 \;\; \frac{1}{6} + 5 \;\; \frac{1}{6} + 6 \;\; \frac{1}{6} = 3,5$.

### 4.2.3 Diskreetse juhusliku suuruse dispersioon

Diskreetse juhusliku suuruse dispersioon defineeritakse valemiga

$$DX = \sum (x - EX)^2 \;\; f(x).$$

Näide. Bernoulli jaotuse korral, kus sündmused $x = 0$ ja $x = 1$ on võrdtõenäolised, arvutatakse dispersioon
$DX = (0 - 0,5)^2 \;\; 0,5 + (1 - 0,5)^2 \;\; 0,5 = 0,25$.

Näide. Üldise Bernoulli jaotuse korral
$DX = (0 - p)^2 \;\; (1 - p) + (1 - p)^2 \;\; p = p(1 - p)$.

Näide. Kui heita täringut, siis

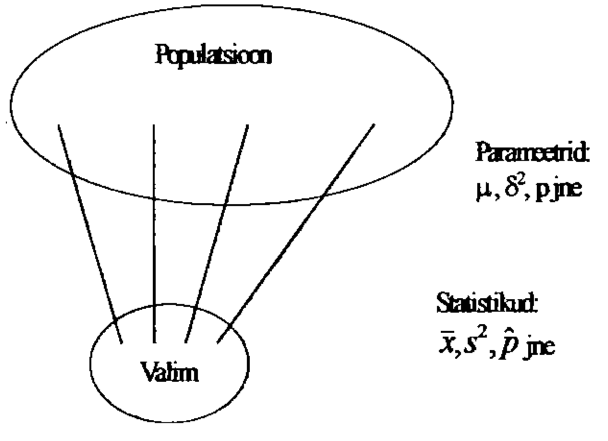
$DX = (1-3,5)^2 \ \frac{1}{6} + (2-3,5)^2 \ \frac{1}{6} + (3-3,5)^2 \ \frac{1}{6} + (4-3,5)^2 \ \frac{1}{6} +$
$+ (5-3,5)^2 \ \frac{1}{6} + (6-3,5)^2 \ \frac{1}{6} = \frac{35}{12} \approx 2,9167$

Märkus. Siin on risk suureks segaduseks. Varem me näitasime, kuidas arvutada keskväärtus ja dispersioon andmestiku valimi jaoks. Nüüd andsime keskväärtuse ja dispersiooni valemid juhusliku suuruse jaoks. Viimati mainitut võib interpreteerida kui mingi üldkogumi keskväärtust ja dispersiooni. Nagu märkisime, on juhusliku suuruse tähendus sageli kasutusel kui mingi üldkogumi mudel. Palun püüa neid eristada:

a) valimi statistikud: $\bar{x}$, $s^2$ jne,

b) üldkogumi (populatsiooni) parameetrid: EX $= \mu$, DX $= \sigma^2$ jne on määratud jaotusega juhusliku suuruse parameetrid. (Kreeka tähed on sageli kasutusel üldkogumi parameetrite tähistamisel.)

Ehk aitab joonis kergendada selle erinevuse arusaamisel:




## 4.2.4 Binoomjaotus

Näide. Seemnehunnikust võeti 5-st seemnest koosnev valim. Seemnetel oli 70% idanevus (st tõenäosus, et seeme hakkab idanema, on 0,7). Kui suur on tõenäosus, et kõik 5 seemet hakkavad idanema?

Lahendus. Kasutades tõenäosuste korrutamise reeglit, saame tõenäosuse $0,7 \ \ 0,7 \ \ 0,7 \ \ 0,7 \ \ 0,7 = 0,7^5 = 0,16807$

Näeme, et *binoomjaotus* määrab n katse hulgas "edukaks" osutunud katsete jaotuse. Tal on kaks parameetrit: n on katseseeria pikkus (täisarvuline) ja p on "edu" võimalikkus üksikkatsel (väärtus 0 ja 1 vahel).

Näide (järg). Missugune on tõenäosus, et 3 seemet 5-st idanevad?

Lahendus. Tähistame idanevat seemet numbriga 1 ja mitteidanevat seemet numbriga 0. Näiteks üks võimalik tulemus võib olla 1, 1, 1, 0, 0. Sellise järgnevuse tõenäosus on $0,7 \quad 0,7 \quad 0,7 \quad 0,3 \quad 0,3 = 0,7^3 \quad 0,3^2 = 0,03087$

Võime valida kuidas iganes kolm idanevat seemet 5 seemne hulgast. Mõned näited on 1, 1, 0, 1, 0; 0, 1, 0, 1, 1; 1, 0, 1, 1, 0 ja nii edasi. Järgnev arvutus näitab erinevate võimaluste arvu 3 idaneva seemne valimiseks 5 seemne seast

$\binom{5}{3} = \frac{5!}{3! \cdot (5-3)!} = \frac{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}{3 \cdot 2 \cdot 1 \cdot (2 \cdot 1)} = 10$. Nende hulgast igal on esinemistõenäosus 0,03087 Järelikult, tõenäosus 3 idaneva seemne valimiseks on $10 \cdot 0,03087 = 0,3087$

Märkus. Sümbol $x!$ tähendab, et korrutame $x$-ga võrdse ja sellest väiksemad positiivsed täisarvud omavahel. Järelikult $5! = 5 \quad 4 \quad 3 \quad 2 \quad 1 = 120$. $x!$ loetakse "$x$ faktoriaal"

Märkus. Sümbol $\binom{n}{x}$ on defineeritud järgmiselt: $\binom{n}{x} = \frac{n!}{x!(n-x)!}$.

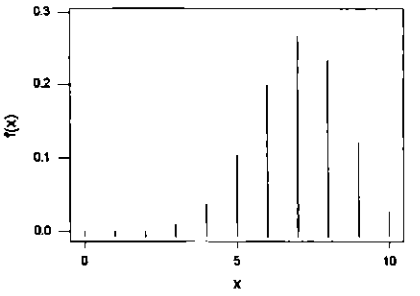
### 4.2.5 Üldine binoomjaotus

Oletame, et mingi katse võib anda kaks erinevat tulemust, mida märgime koodide 1 ja 0 (näiteks idaneb ja ei idane) abil. Koodiga 1 tähistatud tulemusel on tõenäosus $p$. Seega, igal katsel on üldine Bernoulli jaotus. Katset korrati $n$ korda. Missugune on tõenäosus, et 1-de arv on $x$?

Tõenäosus, et juhuslik suurus X omandab väärtuse x $(0 \leq x \leq n)$, on arvutatav järgmise valemiga:

$$P(X = x) = \frac{n!}{x!\,(n-x)!} p^x (1-p)^{n-x}$$

See on *binoomjaotuse tõenäosusfunktsioon* $f(x)$. On võimalt näidata, et binoomjaotuse keskväärtus ja dispersioon on vastavalt $EX = np$ ja $DX = np(1-p)$. Näitena esitame binoomjaotuse graafiku, kus $n = 10$ ja $p = 0,7$




Binoomjaotuse jaotusfunktsiooni väärtused on esitatud raamatu lisas "Tabelid" tabel 2. Illustreerime näite abil tabeli kasutamist.

Näide. Oletame, et meil on plaanis külvata 10 seemet. Valime need juhuslikult seemnehunnikust, seemnetel on 70% idanevus (st tõenäosus, et seeme hakkab idanema, on 0,7).

A. Missugune on tõenäosus, et rohkem kui 8 seemet 10-st idanevad?

B. Missugune on tõenäosus, et 8 seemet 10-st seemnest idanevad?

Lahendus. Väljavõte binoomjaotuse jaotusfunktsiooni tabelist, kus $n = 10$, on järgmine:

| n | x | p=0.1 | p=0.2 | p=0.25 | p=0.3 | p=0.4 | p=0.5 | p=0.6 | p=0.7 |
|---|---|---|---|---|---|---|---|---|---|
| 10 | 0 | 0.3487 | 0.1074 | 0.0563 | 0.0282 | 0.0060 | 0.0010 | 0.0001 | 0.0000 |
| | 1 | 0.7361 | 0.3758 | 0.2440 | 0.1493 | 0.0464 | 0.0107 | 0.0017 | 0.0001 |
| | 2 | 0.9298 | 0.6778 | 0.5256 | 0.3828 | 0.1673 | 0.0547 | 0.0123 | 0.0016 |
| | 3 | 0.9872 | 0.8791 | 0.7759 | 0.6496 | 0.3823 | 0.1719 | 0.0548 | 0.0106 |
| | 4 | 0.9984 | 0.9672 | 0.9219 | 0.8497 | 0.6331 | 0.3770 | 0.1662 | 0.0473 |
| | 5 | 0.9999 | 0.9936 | 0.9803 | 0.9527 | 0.8338 | 0.6230 | 0.3669 | 0.1503 |
| | 6 | 1.0000 | 0.9991 | 0.9965 | 0.9894 | 0.9452 | 0.8281 | 0.6177 | 0.3504 |
| | 7 | 1.0000 | 0.9999 | 0.9996 | 0.9984 | 0.9877 | 0.9453 | 0.8327 | 0.6172 |
| | 8 | 1.0000 | 1.0000 | 1.0000 | 0.9999 | 0.9983 | 0.9893 | 0.9536 | $\boxed{0.8507}$ |
| | 9 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 0.9999 | 0.9990 | 0.9940 | 0.9718 |
| | 10 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |

A. Meie näites $p = 0,7$ Järelikult peame kasutama tulpa, kus $p = 0,7$ Raami sees olev arv on 0,8507 See tähendab, et kui idanevate seemnete arv $(x)$ on 8 või väiksem, on tõenäosus: $P(x \leq 8) = 0,8507$

B. Tõenäosus, et täpselt $x$=8 seemet idanevad on,
$P(x = 8) = P(x \leq 8) - P(x \leq 7) = 0,8507 - 0,6172 = 0,2335.$

### 4.2.6 Poisson'i jaotus

Näide. Vaatleme kindlalt liiki väikesi krabisid. Nad püüti piki randa täiesti juhuslikult. Rand jagati 0, 5 0, 5 m suurusteks ruutudeks. Loeti üle, mitu krabi on igas ruudus. Keskmiselt oli neid $\mu = 2$. Kui suur on tõenäosus, et kindel ruut sisaldab 3 krabi?

Lahendus. Üks võimalus selle probleemi vaatlemisel on mõttes jagada pindala ikka väiksemateks ja väikesemateks ruutudeks. Lõpuks on võimalik, et igas ruudus on mitte rohkem kui üks krabi. Neid ruute võib olla väga suur arv $(n)$ ja krabide igasse ruutu sattumise tõenäosus $(p)$ võib olla väga väike. Selle probleemi lahendamiseks võime kasutada binoomjaotust. Matemaatiliselt on võimalik näidata, et kui bimoomjaotuse korral $n \to \infty$ ja $p \to 0$, siis $np$ on konstant. Tähistame $np = \lambda s$. Seda tulemusjaotust tuntakse kui *Poisson'i jaotust* ja kirjutatakse

$$f(x) = \frac{\lambda s^x e^{-\lambda s}}{x!}$$

Märkus. Miks tähistasime parameetri $\lambda s$-ga. Kas ei oleks olnud küllaldane kasutada ainult $\lambda$? Vastus sellele on: Poisson'i jaotus on seotud kasutavate ühikutega. Kui krabide arv m$^2$ kohta järgib Poisson'i jaotust parameetriga 3, siis krabide arv 0,5 m$^2$ kohta järgib Poisson'i jaotust parameetriga 1,5.

Näiteks kui $\lambda s = 2$, siis tõenäosus, et täpselt $X = 3$ krabi satub ühte ruutu, on $f(3) = \frac{2^3 e^{-2}}{3!} = 0,1804$.

Poisson'i jaotuse keskväärtus $EX = \lambda s$ ja dispersioon $DX = \lambda s$. Järelikult, on keskväärtus ja dispersioon võrdsed.

Poisson'i jaotuse jaotusfunktsiooni tabelist (tabel 3 raamatu lisas) on esitatud väljavõte. Esitame tabeli kasutamise näite.

Oletame, et krabide arv ruutmeetri kohta järgib Poisson'i jaotust keskväärtusega $\lambda s = 3$.

A. Missugune on tõenäosus, et kindel ruut sisaldab mitte rohkem kui 5 krabi?

B. Missugune on tõenäosus, et kindel ruut sisaldab 5 krabi?

Lahendus. Järgnev tabel on Poisson'i jaotuse jaotusfunktsiooni tabeli osa.

| x | λs=0.5 | λs=1 | λs=2 | λs=3 |
|---|---|---|---|---|
| 0 | 0.607 | 0.368 | 0.135 | 0.050 |
| 1 | 0.910 | 0.736 | 0.406 | 0.199 |
| 2 | 0.986 | 0.920 | 0.677 | 0.423 |
| 3 | 0.998 | 0.981 | 0.857 | 0.647 |
| 4 | 1.000 | 0.996 | 0.947 | 0.815 |
| 5 | | 0.999 | 0.983 | $\boxed{0.916}$ |
| 6 | | 1.000 | 0.995 | 0.966 |
| 7 | | | 0.999 | 0.988 |

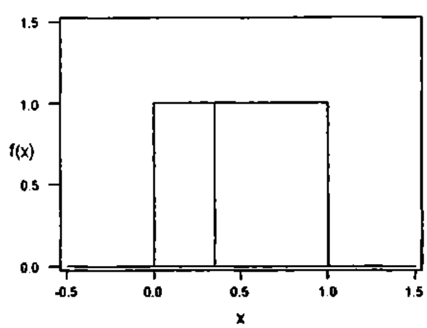
A. Raami sees olev number annab vastuse: $P(X \leq 5) = 0,916$.

B. Nõutav tõenäosus arvutatakse:
$P(X = 5) = P(X \leq 5) - P(X \leq 4) = 0,916 - 0,815 = 0,101$.

## 4.3 Pidevad juhuslikud suurused

### 4.3.1 Ristkülikjaotus

Mitmed taskuarvutid ja tabelarvutusprogrammid omavad juhuslike arvude funktsiooni, mida sageli tähistatakse RND(). Kasutades seda funktsiooni, leiab arvuti juhusliku arvu enamsti 0 ja 1 vahel. Selliste numbrite jaotust nimetatakse *ristkülikjaotuseks*, mille väärtused on: $f(x) = 1$, kui $0 \leq x \leq 1$, ning teistel juhtudel on funktsiooni väärtus 0. Jaotusealune pindala on võrdne 1. Graafik on väga lihtne.




Eelnevalt käsitlesime diskreetseid juhuslikke suurusi. Nüüd käsitletavad juhuslikud suurused on pidevad juhuslikud suurused. Pidev juhuslik suurus võib omandada väärtusi kogu arvteljelt. Tema defineerimiseks kasutatakse tihedusfunktsiooni f(x), mis on alati mittenegatiivne ja tema integraal üle kogu juhusliku suuruse määramispiirkonna on 1. Järelikult on järgnevate juhuslike suuruste korral tõenäosus esitatav pindalana. Näiteks ristkülikjaotuse korral tõenäosus, et $X$ on väiksem kui 0,3, on 0,3. See on eeloleval graafikul vasakpoolne pindala väärtuseni 0,3. See tähendab, et jaotusfunktsioonil $F(x)$ on väärtus $0,3$ kui $x = 0,3$.

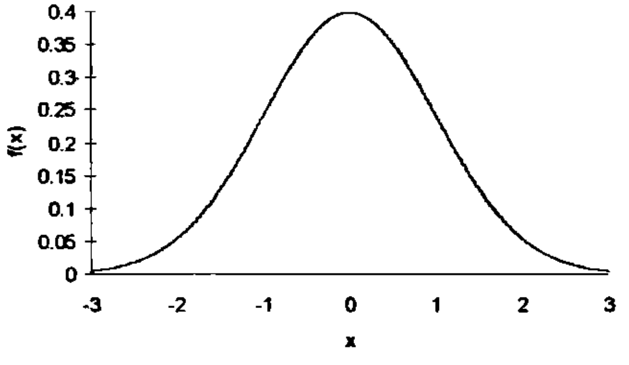
### 4.3.2 Normaaljaotus

Normaaljaotus on tähtis kahel põhjusel. Esiteks, järgivad paljud muutujad bioloogilise uurimuse juures normaaljaotust, vähemalt ligikaudselt. Näitena võib tuua inimese pikkuse ja kaalu. Teiseks, valimi keskväärtus ja teised statistikud, mis kehtivad valimis, püüavad järgida valimi mahu $n$ kasvades normaaljaotust, mida näitame hiljem.

Normaaljaotus on pidev ja tal on kena kelluka kuju. Teda kirjeldavad kaks parameetrit: keskväärtus $\mu$ ja dispersioon $\sigma^2$ Matemaatiline avaldis või tihedusfunktsioon *normaaljaotusele* on esitatav kui

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-0,5\left[\frac{(x-\mu)}{\sigma}\right]^2}$$

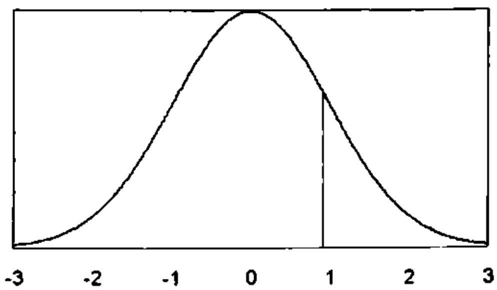
Tihedusfunktsioon on punkti $\mu$ suhtes sümmeetriline ja omab selles punktis kõige suuremat väärtust. Parameeter $\sigma$ määrab tunnuse väärtuse hajutatuse keskväärtuse suhtes ja tema väärtus peab olema positiivne, $\sigma > 0$. Normaaljaotust parameetritega $\mu = 0$ ja $\sigma^2 = 1$ nimetatakse standardseks normaaljaotuseks ning tema graafik on järgmine:




Normaaljaotusega suuruste tõenäosus on arvutatav kui pindala, kusjuures kogu jaotusfunktsiooni alune pindala on 1. Pindala saab leida, kui arvutada vastav integraal. Praktiliseks kasutamiseks on valem arvutuslikult keerukas. Seetõttu esitatakse jaotusfunktsiooni (integraali) väärtused tabelis (tabelid 4 ja 5 raamatu lisas). Normaaljaotuse jaotusfunktsiooni tabeleid on sobiv kasutada normaaljaotusega juhusliku suuruse abil defineeritud sündmuste tõenäosuste arvutamiseks.

Näide. Kui suur on tõenäosus, et normaaljaotusega suuruse $z$ väärtus on väiksem kui 1, kui tema keskväärtus on 0 ja dispersioon on 1 ?

Lahendus. Otsitav tõenäosus on näidatud graafikul väärtusest 1 vasakul pool:




Kasutame normaaljaotuse tabelit ja leiame soovitava tõenäosuse, mis on 0,8413.

Tabel annab igale $z$ positiivsele väärtusele tõenäosuse $F(z)$ omandades väiksema väärtuse kui tabuleeritud, $P(Z < z)$. Kui soovime arvutada tõenäosuse, mis omandab suurema väärtuse, kui on tabelis, kasutame seost. $P(Z > z) = 1 - P(Z < z) = 1 - F(z)$.

Näide. Missugune on tõenäosus, et $Z > 1$?

Lahendus. See tõenäosus kujutab eelneval graafikul väärtusest 1 paremale poole jäävat pindala. $P(Z > 1) = 1 - P(Z < 1) = 1 - 0,8413 = 0,1587$

Negatiivsete $z$ väärtuste korral kasutame fakti
$P(Z \leq -z) = P(Z \geq z) = 1 - P(Z \leq z)$.

Näide. Missugune on tõenäosus, et $Z \leq -1$?

Vastus. $P(Z \leq -1) = P(Z \geq 1) = 1 - P(Z \leq 1) = 1 - 0,8413 = 0,1587$

Märkus: Matemaatilistele puristidele meeldiks ülaltoodud avaldises näha range võrratuse märki, st $P(Z > z)$. Pideva muutuja jaoks pole tähtsust faktil, et tõenäosus, kus $Z$ omandab väärtuse $z$, on 0. Järelikult,
$P(Z \leq 1) = P(Z < 1)$ nii võime kasutada ükskõik kumba kahest.

Märkus. Üldine nõuanne: kui on vaja arvutada tõenäosust normaaljaotuse jaoks, skitseeri jaotuse graafik ja viiruta otsitav pindala. See kergendab otsutamist, kas vastus on reaalne.

Siiani oleme vaadelnud ainult ühte erilist normaaljaotust, mille keskväärtus $\mu = 0$ ja dispersioon $\sigma^2 = 1$. Ometi leidub erinevate keskväärtuste ja dispersioonidega lõpmatult palju teisi normaaljaotusi. Normaaljaotuse sümmeetrilisuse tõttu saame tabeleid lihtsa teisendusega kasutada ka suvalise normaaljaotusega juhusliku suuruse kaudu defineeritud sündmuste tõenäosuste arvutamiseks.

Kui $X$ järgib normaaljaotust keskväärtusega $\mu$ ja dispersiooniga $\sigma^2$, siis muutuja $Z = \frac{X-\mu}{\sigma}$ järgib normaaljaotust keskväärtusega 0 ja dispersiooniga 1. $Z$-i nimetame standardiseeritud normaaljaotuseks. Selle paremaks illustreerimiseks vaatleme näidet.

Oletame, et meeste pikkus üldkogumis (ligikaudu) järgib normaaljaotust $\mu = 175$ ja $\sigma = 10$. Missugune on tõenäosus selleks, et:

a) juhuslikult valitud mees on lühem kui 185 cm,

b) juhuslikult valitud mees on lühem kui 165 cm,

c) juhuslikult valitud mehe pikkus on 165 ja 185 cm vahel?

Lahendus.

a) $P(X < 185) = P\left(Z = \frac{X-\mu}{\sigma} < \frac{185-175}{10}\right) = P(Z < 1) = 0,8413$

b) $P(X < 165) = P\left(Z = \frac{X-\mu}{\sigma} < \frac{165-175}{10}\right) = P(Z < -1) = 0,1587$

c) $P(165 < X < 185) = P(-1 < Z < 1) = 0,8413 - 0,1587 = 0,6826.$

### 4.3.3 Üldkogumi keskväärtuse ja dispersiooni reegel

On kindlad reeglid juhusliku suuruse keskväärtuse (matemaatilise ootuse) ja dispersiooni kasutamiseks. Me ei tõesta seda reeglit, vaid kasutame peagi mõnda nendest. Need reeglid on kokkuvõtvalt esitatud järgmises tabelis:

| **Funktsioon** | **Keskväärtus** | **Dispersioon** | **Märkused** |
|---|---|---|---|
| $a + bX$ | $a + b\mu_x$ | $b^2\delta^2_x$ | |
| $a + bX + cY$ | $a + b\mu_x + c\mu_y$ | $b^2\delta^2_x + c^2\delta^2_y + 2bc \cdot Cov(X,Y)$ | |
| $a + bX + cY$ | $a + b\mu_x + c\mu_y$ | $b^2\delta^2_x + c^2\delta^2_y$ | Kui X ja Y on mittekorreleeritud, st $Cov(X,Y) = 0$ |
| $X_1 + X_2 + \ldots + X_n$ | $\mu_1 + \mu_2 + \ldots + \mu_n$ | $\delta^2_1 + \delta^2_2 + \ldots + \delta^2_n$ | Kui $X_1, X_2, \ldots, X_n$ on mittekorreleeritud, st $Cov(X_i, Y_j) = 0$ iga i,j korral |
| $X_1 + X_2 + \ldots + X_n$ | $\mu + \mu + \ldots + \mu = n\mu$ | $\delta^2 + \delta^2 + \ldots + \delta^2$ | Vaatlused on mittekorreleeritud võrdse keskväärtuse $\mu$ ja dispersiooni $\sigma^2$ korral |
| $\overline{X} = \frac{\sum X_i}{n}$ | $\mu$ | $\frac{\delta^2}{n}$ | Vaatlused on mittekorreleeritud võrdse keskväärtuse $\mu$ ja dispersiooni $\sigma^2$ korral |

### 4.3.4 Juhusliku suuruse valimkeskmise $\overline{X}$ jaotus

Oletame, et mingist üldkogumist (või jaotusest) on moodustatud n vaatlusega juhuslik valim, mille keskväärtus on $\mu$ ja dispersioon on $\sigma^2$ Valimi keskväärtuse võime leida tavalisel viisil $\bar{x} = \frac{\sum x_i}{n}$ Valim on saadud sõltuvalt juhusest, kuna vaatluste valik valimisse oli juhuslik. Kui kordame valikut, saame erineva keskväärtusega teise valimi. Järelikult sõltub valimkeskmine $\bar{X}$ juhusest: $\bar{X}$ on juhuslik suurus!

Siin on kontseptsioon, millest on raske aru saada. Teadlane, kes tegi mingi katse, sai näiteks keskväärtuse 10,0. Ta võib intuitsioonivastaselt vaadelda seda kui "juhuslikku" Seejuures väärtus 10,0 oli juba vaadeldud. Trikk on kujutada ette, mis juhtuks, kui katset korrata: saame terve hulga keskväärtusi. Kui laboratooriumi reeglid seda lubavad, võime katset korrata ja võime saada lõpmatult palju valimeid. Teooria põhjal võime ennustada valimkeskmise $\bar{X}$ käitumist katse kujutletavates kordustes.

Tuleme tagasi üldkogumi juurde, mille keskväärtus on $\mu$ ja dispersioon on $\sigma^2$ Mida võime ennustada üldkogumi kohta, kui vaatleme üldkogumist moodustatud valimit, mille suurus on $n$?

"Keskväärtuse ja dispersiooni reeglite" tabeli põhjal saame:

1. $E\left(\bar{X}\right) = \mu$ Korduvalt valitud valimkeskmine on üldkogumi keskväärtus.

2. $D\left(\bar{X}\right) = \frac{\sigma^2}{n}$ Valimkeskmise dispersioon, kui kasutame korduvat valikut, on üldkogumi dispersiooni ja valimi mahu suhe.

Märkus. Tähtis on seejuures eeldus, et vaatlused valimis peavad olema sõltumatud. See tingimus on täidetud, kui valim on moodustatud juhusliku valiku abil.

Märkus. $\sqrt{D\left(\bar{X}\right)} = \sqrt{\frac{\sigma^2}{n}} = \frac{\sigma}{\sqrt{n}}$ tähendab ka siin valimkeskmise standardviga. (Võrdle seda avaldist kirjeldava statistika peatükis olevaga.)

Kolmas keskväärtuse omadus on esitatud ilma tõestuseta ja seda tuntakse kui tsentraalse piirväärtuse teoreemi:

3. *Valimkeskmise* $\bar{X}$ *jaotus*, mis on saadud sama üldkogumi korduvate valimite põhjal, läheneb n-i kasvades normaaljaotusele. See avaldub paljudes mittemandunud valikutes, isegi kui $X$ ei ole normaaljaotusega.

Võtame kokku eelneva: kui meil on mingi üldkogumi valim mahuga $n$, siis võime vaadelda valimkeskmist $\bar{X}$ kui ühte vaatlust üldkogumist (jaotusest), millel on keskväärtus $\mu$ ja dispersioon $\frac{\sigma^2}{n}$ Kui valim on suur ja kui jaotus üldkogumis ei ole normaalne, võime lisaks eeldada, et $\bar{X}$ on vähemalt ligikaudu normaaljaotusega. Kui üldkogum on normaaljaotusega, siis on ka

$\bar{X}$ normaaljaotusega isegi juhul, kui n on väike.

Näide. Oletame, et meeste populatsioonis tunnus pikkus järgib (ligikaudu) normaaljaotust, mille $\mu = 175$ ja $\sigma = 10$. Missugune on tõenäosus selleks, et 4-st mehest koosnevas valimis keskmine pikkus on väiksem kui 180 cm?

Lahendus. Vaatleme seda valimit, mis järgib nornaaljaotust, mille keskväärtus on $E(\bar{X}) = \mu = 175$ ja dispersioon on $D(\bar{X}) = \frac{\sigma^2}{n} = \frac{100}{4} = 25$ , seega keskväärtuse standardviga on $\sqrt{25} = 5$. Nõutava tõenäosuse arvutamiseks kasutame $Z$ teisendust, mille põhjal:

$$P(\bar{X} < 180) = P\left(Z = \frac{\bar{X}-\mu}{\sqrt{\frac{\sigma^2}{n}}} < \frac{180-175}{\sqrt{\frac{100}{4}}}\right) = P(Z < 1) = 0,8413$$

Näide. (Üldkogum, mis ei järgi normaaljaotust.) Oletame, et suurte majapidamiste populatsioonis autode omamine on jaotunud järgnevalt:

| Autode arv $x$ | Suhe kõikidest majapidamistest $f(x)$ |
|---|---|
| 0 | 0,40 |
| 1 | 0,50 |
| 2 | 0,07 |
| 3 | 0,03 |

Kui suur on tõenäosus, et 200 majapidamise juhuslikus valimis valimkeskmine on väiksem kui 0,8?

Lahendus. Üldkogumi jaoks võime leida
$\mu = \sum x \quad f(x) = 0 \quad 0,40 + 1 \quad 0,50 + 2 \quad 0,07 + 3 \quad 0,03 = 0,73$ ja $\sigma^2 = \sum x^2 \quad f(x) - \mu^2 = 0 \quad 0,40 + 1 \quad 0,50 + 4 \quad 0,07 + 9 \quad 0,03 - 0,73^2 = 1,05 - 0,5329 = 0,5171$. Suure valimi korral (n=200) võime oodata $\bar{x}$, mis vähemalt ligikaudu järgib normaaljaotust, hoolimata faktist, et üldkogum oli defineeritud mittenormaalsena. Järelikult
$P(\bar{x} < 0,8) = P\left(z = \frac{\bar{x}-\mu}{\sqrt{\frac{\sigma^2}{n}}} < \frac{0,80-0,73}{\sqrt{\frac{0.5171}{200}}}\right) = P(z < 1,38) = 0,9162.$

## 4.4 Ülesanded

**4.1**

Väga suures puuviljade kaubalastis on kaoprotsendiks määratud 0,1%. Kui suur on tõenäosus, et juhuslikult moodustatud 9 puuviljast koosnev valim sisaldab:

A. ühte riknenud puuvilja,

B. vähemalt kolme riknenud puuvilja?

**4.2**

Kõrge amplituudiga lained järgivad sekundis Poissoni jaotust keskväärtusega $\lambda s = 2$ sekundis. Arvuta tõenäosus, et 5-sekundilise perioodi kestel:

A. ei esinenud selliseid laineid, st $P(X = 0)$,

B. $P(X \leq 16)$,

C. $P(X > 19)$?

**4.3**

Võib eeldada, et lehmade piimatoodang järgib normaaljaotust keskväärtusega $\mu = 24,5$ ja dispersiooniga $\sigma^2 = 30$.

A. Missugune on tõenäosus, et juhuslikult valitud lehma toodang on alla 18 kg?

B. Missugune on tõenäosus, et 10 juhuslikult valitud lehma keskmine toodang on väiksem kui 26 kg?

C. Missugune on tõenäosus, et 10 juhuslikult valitud lehma kogutoodang on rohkem kui 280 kg?

D. Leia ülemine ja alumine piir lehmade piimatoodangule, eeldades, et 95% kõikide lehmade piimatoodangust on nende piiride vahel. (Piirid peavad olema sümmeetrilised keskväärtuse ümber.)

E. Leia ülemine ja alumine piir 10 lehma keskmisele piimatoodangule, eeldades, et 95% kõikidest gruppidest, mis koosnevad 10 juhuslikult valitud lehmast, omavad keskväärtust nende piiride vahel.

**4.4**

Laboratooriumis uuriti alaniini aminotransferaasi (ALT) seerumit. On teada, et korduvate katsete seerias järgib üksiku proovi tulemus normaaljaotust keskväärtusega, mis on võrdne selle katse õige ALT kontsentratsiooniga, ja standardhälbega $\sigma$=4.

A. Oletame, et ALT kontsentratsioon rohkem kui 40 on vaadeldav kui "ebaharilikult kõrge" Kui õige ALT kontsentratsioon on 35, missugune on tõenäosus, et tema proov on "ebaharilikult kõrge"?

B. Oletame, et laboratoorium teeb iga prooviga kaks analüüsi. Missugune on tõenäosus, et nende kahe mõõtmise keskväärtus $\bar{X}$ on suurem kui 40, kui proovi õige ALT kontsentratsioon on 35?

C. Leia 95% vea usalduspiirid katse tulemustele, kui proovi õige ALT kontsentratsioon on 35, eeldades, et

i) tehti üks ülesmärkimine,

ii) tehti kahekordne ülesmärkimine.

**4.5**

Farmaatsiakompanii korraldas diagnostilise testi X, et uurida inimeste medikamentide kuritarvitamist. Test tehti kõigile, isegi kui inimene ei olnud vaadeldaval ajal ravimit kuritarvitanud. X kõrge väärtus näitab, et inimene on kuritarvitanud ravimeid. Test tehti inimestega, kelle kohta on teada, et nad kuritarvitavad ravimeid, ja kontrollgrupiga. Võib eeldada, et X järgib igas grupis normaaljaotust. Teadaolevad üldkogumi parameetrid on toodud järgnevas tabelis:

| | Kuritarvitajad | Kontroll |
|---|---|---|
| Keskväärtus $\mu$ | 5,0 | 2,5 |
| Standardhälve $\sigma$ | 1,0 | 0,5 |

Klassifitseerimise reegel on järgmine: inimene on ravimi kuritarvitaja, kui X-i väärtus on üle 3,5.

Diagnostilise testi tundlikkus on tõenäosus, et haige inimene on ka testi järgi haigeks diagnoositud.

Diagnostilise testi eripära on tõenäosus, et terve inimene on ka testi järgi terveks diagnoositud.

Küsimused:

A. Arvuta esitatud testi tundlikkus.

B. Arvuta esitatud testi eripära.

C. Võib arvata, et kontrollgrupis on inimesi, kes kuritarvitavad ravimeid, aga kellel õnnestus seda varjata. Need on inimesed, kellel X-i väärtus on tõenäoliselt kõrge. Eeldame, et sellised inimesed leitakse ja kustutatakse kontrollgrupist. Mis juhtub siis:

i) tundlikkusega,

ii) eripäraga?

**4.6**

Paljud meditsiinilised katsed baseeruvad niinimetatud "kahe sigma piiridel". Moodustame vere koostise kohta valimi erinevatest mõõtmistest, milles on andmed kaaliumi, kaltsiumi, kolesterooli jne kohta. "Kahe sigma piirid" on (keskväärtus)$\pm$ (2-kordne standardhälve). Üldiselt võib öelda, et 95% kõigist tervetest patsientidest omavad väärtusi nende piiride vahel.

"2 sigma" kolesterooli piirid naistel on:

| Vanus | 2 sigma piirid |
|---|---|
| Noorem kui 21 | 140 – 180 |
| 21 kuni 49 | 140 – 280 |
| 50 või üle | 180 – 280 |

Küsimused.

A. Leia iga vanusegrupi korral, missuguse keskväärtuse $\mu$ ja standardhälbe $\sigma$ korral on tabelis esitatud piirid määratud.

B. Missugune on tõenäosus, et indiviid on väljaspool kolme sigma piire?

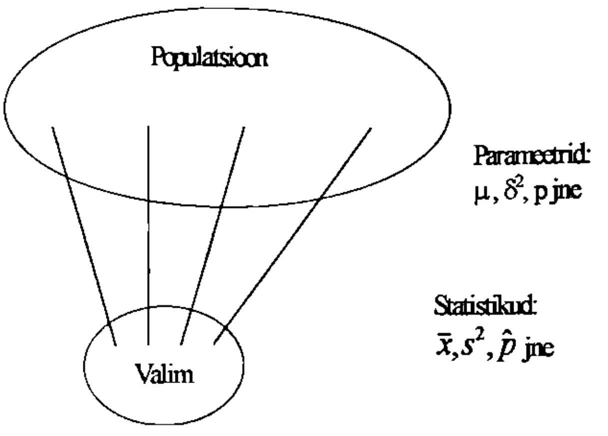
C. 18-aastase tüdruku kolesterooli tase on 225. Missugune on tõenäosus, et tüdruk on terve, kui ta kolesterooli tase on 225 või suurem?

# Peatükk 5

# Statistiline järeldus: hindamine

Enamike teaduslike andmete kogumise eesmärk on teha järeldusi, mida saab üldistada. Veterinaar või arst on tavaliselt huvitatud üksikust patsiendist. Statistiline järeldus tehakse mingi kogumi kohta: "Katses raviti tüsedaid patsiente ravimiga A, mille tulemusena vähenes nende kaal (keskmiselt) 3 kg. Teisi patsiente raviti ravimiga B ning nende keskmine kehakaal vähenes 1.7 kg" Arstid kasutavad sageli selliseid tulemusi, otsustades, missugust ravimit määrata. Nad võivad eeldada, et patsient on "sarnane" patsiendiga, keda kasutati katses, ning ravim A on tema jaoks katsetulemuse põhjal parim.

Paljude statistiliste otsuste kujunemine on järgmine. Eeldame, et eksisteerib üldkogum, mille jaoks ongi vaja tulemusi üldistada. Kui eeldame, et üldkogumis on mingil tunnusel mingi jaotus, siis selle jaotuse võib kindlaks määrata mingite jaotusparameetrite abil: tunnuse keskväärtuse $\mu$, tunnuse dispersiooni $\sigma^2$, tunnuse osakaalu $p$ jne järgi. Üldkogumi tunnuste parameetrid on üldiselt tundmatud ja andmete kogumise üheks eesmärgiks on hinnata neid parameetreid. Järelikult tuleb moodustada üldkogumist valim. Saame arvutada valimi statistikud, mis kirjeldavad valimit: keskväärtus $\bar{x}$; dispersioon $s^2$; osakaal $\hat{p}$ jne. Seda vaadet võib illustreerida järgmiselt:




Valimit kasutatakse selleks, et teha järeldusi üldkogumi kohta. Neid järeldusi nimetakse statistilisteks järeldusteks.

Näide. 1000 valija seast valiti juhuslikult 350 valijat. Nad ütlesid, et valivad sotsiaaldemokraate, kui valimised oleksid täna. Mida võib öelda sotsiaaldemokraatide poolt hääletajate proportsiooni (suhte) kohta kõigi valijate hulgast?

Näide. Instrumenti kasutati viis korda, et kindlaks määrata laktoosi taset lahuses. Iga kord saadi veidi erinev tulemus, kuna mõõtmisel esineb alati mingi viga. Võime arvutada viie mõõtmise keskväärtuse $\bar{x} = 3,22$ ja dispersiooni $s^2 = 0,047$ Missugune on "õige" laktoosi tase lahuses?

Näide. A sorti nisu kasvatati 10 põllul ja B sorti nisu 10 teisel põllul. A sorti nisu andis keskmiselt saaki 8,0 ja B sorti nisu andis saaki 7,4. Kas võime julgesti soovitada, et A sorti nisu on parem?

## 5.1 Märkus

Märkisime juba varem, et on kasulik kasutada erinevaid sümboleid üldkogumi parameetriteks ja valimi statistikuteks. Kokkuvõtvalt on järgnevas tabelis esitatud mõned sageli kasutatavad sümbolid.

| | Üldkogumi parameetrid | Valimi statistikud |
|---|---|---|
| Ühikute arv | $N$ | $n$ |
| Keskväärtus | $\mu$ | $\bar{x}$ |
| Dispersioon | $\sigma^2$ | $s^2$ |
| Standardhälve | $\sigma$ | $s$ |
| Proportsioon | $p$ või $p_X$ | $\hat{p}$ või $\hat{p}_X$ |

## 5.2 Punkthinnang

Üldkogumi parameetri hinnang on mingi valimi parameetri lähendfunktsioon. Saadud väärtust nimetakse hinnanguks. Hinnangut võib esitada üksikväärtusena, punkthinnanguna või usalduspiiridena.

Tavaliselt lisatakse hindamisel mingi märk vastavale tähisele, ^ ("katus") sümboli kohale. Keskväärtuse $\mu$ hinnangut tähistatakse $\hat{\mu}$ , dispersiooni $\sigma^2$ hinnangut tähistatakse $\hat{\sigma}^2$

Näide. $\bar{x}$ on (punkt-)hinnang $\mu$. Järelikult kehtib võrdus $\hat{\mu} = \bar{x}$ Kui arvutame valimi põhjal keskmse $\bar{x}$ =12,4, siis 12,4 on (punkt-)hinnang $\mu$.

Näide. $\hat{\sigma}^2$ on (punkt-)hinnang $\sigma^2$ Järelikult $\hat{\sigma}^2 = s^2$ Kui valimist saame $s^2$=3,5, siis 3,5 on (punkt-)hinnang $\sigma^2$

## 5.3 Hinnangu omadused

Soovime, et hinnangud oleksid mingis mõttes "lähedased" tegelikule parameetri väärtusele. Hinnangus sisaldub alati mingi süstemaatiline viga, st hinnang üle- või alahindab hinnatava parameetri õiget väärtust. Esitame mõned hinnangu omadused Need on: nihutamatus, efektiivsus ja ruutkeskmine viga.

### 5.3.1 Nihutamatus

Mingi parameetri $\theta$ hinnang $\hat{\theta}$ on nihketa, kui hinnang on keskmiselt õige, st $E\left(\hat{\theta}\right) = \theta$.

Näide. Võib näidata, et $E\left(\bar{X}\right) = \mu$. Järelikult valimi keskväärtus on üldkogumi keskväärtusele $\mu$ nihketa hinnanguks.

Näide. Võib näidata, et $E\left(s^2\right) = \sigma^2$ Järelikult valimi dispersioon on üldkogumi dispersioonile $\sigma^2$ nihketa hinnanguks.

Näide. Üldiselt, $E\left(s\right) \neq \sigma$ Järelikult valimi standardhälve ei ole üldkogumi standardhälbe nihketa hinnanguks.

Nihe hinnangule on defineeritud järgmiselt $Bias\left(\hat{\theta}\right) = E\left(\hat{\theta}\right) - \theta$.

### 5.3.2 Efektiivsus

Mõnikord hinnatakse ühte ja sama parameetrit erineval viisil. Üks võimalus hinnangute võrdlemiseks on võrrelda nende dispersioone. Võib öelda, et hinnang $\hat{\theta}_1$ on efektiivsem kui hinnang $\hat{\theta}_2$, kui $D\left(\hat{\theta}_1\right) < D\left(\hat{\theta}_2\right)$. Suhteline efektiivsus $\hat{\theta}_1$ ja $\hat{\theta}_2$ vahel defineeritakse $\frac{D(\hat{\theta}_2)}{D(\hat{\theta}_1)}$.

Näide. Tähistame valimi mediaani $Md$. Saab näidata, et normaaljaotuse korral $D(Md) = \frac{\pi}{2}\frac{\sigma^2}{n}$ Teame, et $D\left(\bar{X}\right) = \frac{\sigma^2}{n}$. Mediaani suhteline efektiivsus, kui võrrelda seda valimi keskväärtusega, on järelikult $\frac{D(Md)}{D(\bar{X})} = \frac{\pi}{2} \approx 1,57$ $\bar{X}$ on 57% efektiivsem kui mediaan. Teisti öeldes, sama täpsusega $\mu$ hindamiseks vajame 57% võrra suuremat valimit, kui kasutame $Md$ $\bar{x}$ asemel.

### 5.3.3 Ruutkeskmine viga

Mõnikord võime leida kaks erinevat hinnangut: nihutatud hinnangu, millel on väike dispersioon, ning mingi teise nihutamata hinnangu, millel on suur dispersioon. Missugust hinnangut valida? Hinnangu valimise kriteeriumiks defineerime hinnangu $\hat{\theta}$ ruutkeskmise vea järgmiselt: $MSE = E\left[\left(\hat{\theta} - \theta\right)^2\right]$ Valime hinnangu, millel on väiksem MSE. See näitab, et $MSE\left(\hat{\theta}\right) = D\left(\hat{\theta}\right) + \left[Bias\left(\hat{\theta}\right)\right]^2$

## 5.4 Vahemikhinnang

### 5.4.1 Keskväärtuse $\mu$ usalduspiirid, kui $\sigma^2$ on teada

Usaldusvahemik on avaldis üldkogumi parameetri kohta. Avaldis on esitatud kui vahemik: "Vahemik 2,5 kuni 5,9 katab keskväärtuse $\mu$ õige väärtuse" Väide on esitatud 95% usaldusnivool. Me võime 95% kindlad olla, et see väide on õige.

Sissejuhatav näide. Instrumenti kasutati viis korda, et määrata kindlaks laktoosi tase lahuses. Iga kord saadi veidi erinev tulemus, kuna mõõtmisel esineb alati mingi viga.. Mõõtmised olid järgmised:

3,4; 3,3; 3,1; 2,9; 3,4.

Tööjuhises on lubatud instrumendi mõõtmisviga dispersiooniga $\sigma^2 = 0,05$. Arvuta 95% usaldusvahemik laktoosi õigele tasemele lahuses.

Teame järgmist:

1) valimi keskväärtus on $\bar{x} = 3,22$ ,

2) valim sisaldab $n = 5$ vaatlust,

3) üldkogumi dispersiooni $\sigma^2 = 0,05$ ,

4) me ei tea $x$ jaotust, aga eeldame, et $x$ järgib normaaljaotust (mõõtmise vead järgivad sageli normaaljaotust),

5) ei tea, kas vaatlused on tehtud sõltumatult, aga eeldame, et on.

Nendele faktidele ja oletustele toetudes peab vaadeldav väärtus $\bar{X}$ näitama ühte vaatlust, mis järgib normaaljaotust keskväärtusega $\mu$ (mis on teadmata) ja dispersiooniga $\frac{\sigma^2}{n}$ (meie juhul on teada). Järelikult:

$\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$. Võime teisendada selle standardnormaaljaotuseks:

$Z = \frac{\bar{X}-E(\bar{X})}{\sqrt{Var(\bar{X})}} = \frac{\bar{X}-\mu}{\sqrt{\frac{\sigma^2}{n}}}$

Standardnormaaljaotus $Z$ näitab, et tõenäosus, omandada $z$-i väärtust vahemikus $-1,96$ ja $+1,96$,

on 0,95 (st 95%). Märkus. Selle väärtuse leiad raamatu lõpus tabelist 5 viimaselt realt veerust 0,975.

Järelikult võime öelda, et

$P\left(-1,96 < \frac{\bar{X}-\mu}{\sqrt{\frac{\sigma^2}{n}}} < 1,96\right) = 0,95$. Teisendame seda avaldist:

$P\left(\bar{X} - 1,96 \sqrt{\frac{\sigma^2}{n}} < \mu < \bar{X} + 1,96 \sqrt{\frac{\sigma^2}{n}}\right) = 0,95$. Kui asetame oma näite väärtused avaldisse, saame vahemiku $3,22 - 1,96 \cdot \sqrt{\frac{0,05}{5}} < \mu < 3,22 + 1,96 \cdot \sqrt{\frac{0,05}{5}}$, millest

$3,22 - 0,196 < \mu < 3,22 + 0,196$ ning lõpuks saame $(3,024 < \mu < 3,416)$.

Oleme 95% veendunud, et vahemik (3,024; 3,416) sisaldab $\mu$ õige väärtuse. Neid arve nimetatakse keskväärtuse $\mu$ 95% usalduspiirideks.

Järelikult, keskväärtuse $\mu$ 95% usalduspiirid arvutatakse järgmiselt:

$$\bar{x} \pm z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}.$$

Suurusele $z$ saame väärtuse $t$-jaotuse tabelist raamatu lõpus. 95% usalduspiiride jaoks kasutame veergu 0,975.

Usaldusvahemiku laius sõltub kolmest parameetrist:

1) suurem $\sigma$ väärtus annab laiema vahemiku,

2) suurem $n$ väärtus annab väiksema vahemiku (kui suurendame valimi mahtu, saame väiksema juhusliku varieeruvuse),

3) kui muudame usaldusnivood, muutub ka vahemik. Näiteks 99% usaldusvahemik on laiem kui 95% usaldusvahemik.

### 5.4.2 Keskväärtuse $\mu$ usalduspiirid, kui $\sigma^2$ ei ole teada

Kui üldkogumi dispersioon ei ole teada, siis ei saa enam arvutada nii nagu eelpool. $Z$-i avaldises ei tea me $\sigma$ väärtust. Sel juhul kasutame Dublini Guinnessi õllevabriku ärisaladust. Aastal 1908 avaldas üks Guinnessi töötaja W S. Gossett tõestuse, et teatava eelduse korral järgib statistik $\frac{\bar{X}-\mu}{\sqrt{\frac{s^2}{n}}}$ niinimetatud *t-jaotust*. Vabriku juhatus ei lubanud Gossett'il artiklit teaduslikus ajakirjas avaldada, nii avaldas ta oma tulemused pseudonüümi Student all. Sellepärast kutsutakse seda jaotust sageli kui Student'i $t$-jaotuseks.

Teeme $t$-jaotuse korral järgmised eeldused:

1. Vaatlused peavad olema vastastikku sõltumatud. Lihtsaim viis selle saamiseks on valida andmestikust lihtne juhuslik valim.

2. $X$-i jaotus peaks olema normaaljaotus. See eeldus sobib suuremale valimile.

Kui need eeldused on täidetud, siis järgib statistik $\frac{\bar{X}-\mu}{\sqrt{\frac{s^2}{n}}}$ $t$-jaotust. Jaotus on määratud parameetriga, mida nimetatakse vabadusastmete arvuks. Sel juhul märgime tõestuseta, et vabadusastmete arv on $(n-1)$. (See on $s^2$ avaldise nimetaja.) Osa tabelist 5 on $t$-jaotuse tabel.

Näide. Vaatleme eelnevas punktis esitatud näidet, kuid nüüd ei tea üldkogumi dispersiooni $\sigma^2$ Arvuta 95% usalduspiirid laktoosi õigele tasemele lahuses.

Näites olevate andmete põhjal võime arvutada keskväärtuse $\bar{x} = 3,22$ ja dispersiooni

$$s^2 = \frac{\sum x^2 - \frac{(\sum x)^2}{n}}{n-1} = \frac{52{,}03 - \frac{(16{,}1)^2}{5}}{4} = \frac{0{,}188}{4} = 0,047$$

Meil on $n = 5$ vaatlust, seega vabadusastmete arv on $n - 1 = 4$. Leiame $t$-jaotuse tabelist, et $t$ väärtusele tõenäosusega 95% vastab $t = 2,776$. Võime arvutada tõenäosuse

$P\left(-2,776 < \frac{\bar{X}-\mu}{\sqrt{\frac{s^2}{n}}} < 2,776\right) = 0,95$. Teisendame avaldist

$P\left(\bar{X} - 2,776 \sqrt{\frac{s^2}{n}} < \mu < \bar{X} + 2,776 \sqrt{\frac{s^2}{n}}\right) = 0,95$. Peale teadaolevate väärtuste asetamist avaldisse saame usaldusvahemiku

$3,22 - 2,776 \sqrt{\frac{0{,}047}{5}} < \mu < 3,22 + 2,776 \sqrt{\frac{0{,}047}{5}}$,

$3,22 - 0,269 < \mu < 3,22 + 0,269$ ; $(2,95 < \mu < 3,49)$.

Võime 95% kindlusega öelda, et vahemik $(2,95; 3,49)$ sisaldab $\mu$ õiget väärtust.

*Keskväärtuse $\mu$ usalduspiirid* (kui üldkogumi $\sigma$ on teadmata) arvutatakse valemist

$$\bar{x} \pm t\frac{s}{\sqrt{n}}.$$

$t$ väärtus saadakse $t$-jaotuse tabelist. 95% usalduspiiride jaoks kasutame veergu 0,975. $t$-jaotuse vabadusastmete arv on $n-1$.

Maksimaalne vabadusastmete arv ($\gamma$) $t$-jaotuse tabelis on 100. Suurema $\gamma$ väärtuse korral peame kasutama $z$ väärtust kui head lähendit. Nagu näeme $t$ väärtus 100 vabadusastmega (veerus 0,975) on võrdne 1,984 ning vastav $z$-i väärtus on 1,96. Püüdlikud tudengid kasutavad suurte valimite korral $\gamma$ väärtusena tabelis toodud väärtuste interpolatsiooni, samal ajal kui paljud tudengid kasutavad lihtsalt ligilähedast väärtust.

### 5.4.3 Kui suur peab valim olema?

Sissejuhatav näide. Kasutame punkti 5.4.1 näidet.

Soovime teha uusi mõõtmisi ja vajame valimit nii, et keskväärtuse $\mu$ 95% usalduspiirid oleksid kitsamad kui $\pm 0,10$. Kui suurt valimit vajame?

Eelduse kohaselt teame, et $\sigma^2 = 0,05$. Eespool on juba arvutatud nende andmete keskväärtuse $\mu$ 95% usalduspiirid: $3,22 \pm 1,96\sqrt{\frac{0,05}{5}}$; $3,22 \pm\ 0,196$.

Uue valimi jaoks peame jälgima, et $1,96\sqrt{\frac{\sigma^2}{n}} \leq 0,10$, millest

$1,96^2\sigma^2 \leq 0,10^2 \cdot n$ ; $n \geq \frac{1,96^2\sigma^2}{0,10^2}$ Kuna teame, et $\sigma^2 = 0,05$, siis järelikult $n \geq \frac{1,96^2 \cdot 0.05}{0,10^2}$ Ümardame ülespoole ja saame valimi mahuks $n = 20$.

### 5.4.4 Valimi mahu määramise üldine valem

Valime $n = \frac{z^2\sigma^2}{L^2}$, kus L on soovitav täpsus. (Järelikult L on usalduspiiride pool pikkust.) Ümardame n lähima täisarvuni ülespoole.

Valimi mahu määramisel on probleemiks, et peame teadma dispersiooni $\sigma^2$ Praktikas teame harva dispersiooni väärtust, eriti enne andmete kogumist. Järgnevad meetodid annavad vähemalt mingi hinnangu dispersioonile.

1. Kui üldkogumi dispersioon on teada, kasuta muidugi seda.

2. Kui üldkogumi dispersioon on teadmata, võib kasutada mõne varajasema samatüübilise andmestiku valimi dispersiooni.

3. Kui isegi see ei ole võimalik, järgi oma töös "kiiret ja musta" meetodit. Püüa (oletuse ja ettekujutuse abil) leida suurim x ($x_{max}$) ja vähim x ($x_{min}$) väärtus, mida nendest muutujatest võib võtta. Paljude korralike jaotuste korral võib see olla $(x_{max} - x_{min}) \approx 4s$. (Kui jaotus ei ole hea, võib kasutada selle asemel 5s.) See teeb võimalikuks oletada dispersiooni $\sigma$ suurust, mis on sageli valimi mahu määramisel küllaldane.

## 5.5 Ülesanded

### 5.1

McIntosh (ära aja segi arvuti margiga) sorti õunte hindade valim koguti juhuslikult erinevatest kauplustest. Hinnad olid (dollarites):

2,31 2,50 2,30 2,40 2,34 2,20 2,42
2,46 2,30 2,25

A. Arvuta keskväärtus ja dispersioon.

B. Arvuta keskväärtuse $\mu$ 95% usalduspiirid.

C. Mitu protsenti valimi vaatlustest on usalduspiiride vahel? Kommenteeri vastust.

### 5.2

Juhuslikult valitud ülikooli 10 meesõppejõu vere kolesterooli tase oli (grammi liitri kohta):

3,0 1,8 2,1 1,4 2,5 1,6 1,9 2,8 2,2 2,7

A. Arvuta keskväärtus ja dispersioon.

B. Arvuta ja interpreteeri keskväärtuse $\mu$ 95% usalduspiirid.

### 5.3

On teada, et 90% normaalkaaluliste kanade tapakaal on 1,9 ja 2,1 vahel. Tehti katse, kus kontrolliti uut toitumisskeemi. Arvuta soovitav valimimaht nii, et keskväärtuse $\mu$ 95% usalduspiirid ei oleks laiemad kui $\pm 20$ grammi (st $\pm 0,02$ kg). Eeldame, et standardhälve on sama suur kui normaalkaaluliste kanade korral.

### 5.4

Mingil kindlal maal on teada, et seal elab 35000 fertiilsuseas naist. Terviseprogrammi planeerimiseks on vaja informatsiooni. Sellel maal on vaja hinnata (0 - 6 aastaste) laste arvu. Moodustati juhuslik valim n=400 naisest ning küsiti nendelt, mitu 0 - 6 aasta vanuses last on neil. Tulemused on esitatud tabelis:

| Laste arv x | 0 | 1 | 2 | 3 | 4 | Kokku |
|---|---|---|---|---|---|---|
| Naiste arv f | 40 | 100 | 120 | 110 | 30 | 400 |

A. Arvuta keskväärtus ja dispersioon laste arv/ naiste valimis.

B. Arvuta keskväärtuse $\mu$ 95% usalduspiirid, keskmine laste arv ühe naise kohta.

C. Arvuta 95% usalduspiirid 0- kuni 6-aastaste laste üldarvule sellel maal.

D. Küsitlust tahetakse korrata ning sellel juhul keskväärtuse $\mu$ usalduspiirid ei tohi olla laiemad kui $\pm 0,05$. Mitu naist on nüüd valimisse vaja?

# Peatükk 6

# Hüpoteeside kontrollimine

## 6.1 Sissejuhatav näide

(R.A. Fisheri, kes on üheks suureks meheks statistikas, klassikaline näide).

"Daam ütleb, et ta võib tee maitse järgi määrata, kas tassi valati esimesena tee või piim. Testimine kinnitas seda. Valmistati ette viis teetassi. Mõnda valati esimesena tee ja mõnda tassi esimesena piim. Katsealune määras iga tassi kohta, mida sinna esimesena valati. Daam määras viis korda järjest õigesti.

Testi hüpotees on, et daam oskab vahet teha tee kahe valmistamisviisi vahel."

On ilmne, et kunagi ei saa kindlalt öelda, kas meil on õigus daami uskuda või mitte. Võime kehtestada mingisuguse "otsuse reegli" nii, et teeksime võimalikult vähem vigu.

Teaduslik lähenemine on alustada daami umbusaldamisega. Esimese hüpoteesina oletame, et daam ei tunne erinevust, mis tähendab, et edu tõenäosus iga tassi korral on 0,5. Seda nimetatakse nullhüpoteesiks:

$H_0 : p = 0,5.$

Tegelikult, kui daam tunneb erinevust kahte liiki tee vahel, siis nullhüpotees on vale. Võime olla suuremeelsed ja möönda, et daam tunneb erinevust, seepärast on tal edu tõenäosus suurem kui 0,5. Sellepärast kasutame alternatiivset ehk sisukat hüpoteesi:

$H_1 : p > 0,5.$

Teaduslik ülesanne on valida nende kahe hüpoteesi vahel. Selleks peame enne katse sooritamist kehtestama otsustusreeglid näiteks järgmiselt:

a) kui daam määrab õigesti vähem kui 5 tassi, siis usaldame $H_0$;

b) kui daam määrab õigesti 5 tassi 5-st, siis usaldame $H_1$.

Mõlemal juhul võib otsus olla vale. Puhta õnne korral võib daam viis korda järjest õigesti ära arvata, isegi kui $p = 0,5$, mis tähendab, et lükkame tagasi $H_0$, isegi kui see on õige. Sellist tüüpi vea tegemise tõenäosust tähistatakse sageli $\alpha$.

Ja isegi kui daam tunneb erinevust (näiteks, isegi kui tal on õigete otsuste tegemise tõenäosus 0,8), võib ta teha aegajalt vigu. Ta võib õigesti arvata näiteks 4 korral ja 1 korral eksida, mis tähendab otsustamise reegli järgi, et ei tohi usaldada $H_1$, ehkki ta on õige. Tähistame seda liiki vea tegemise tõenäosuse $\beta$-ga.

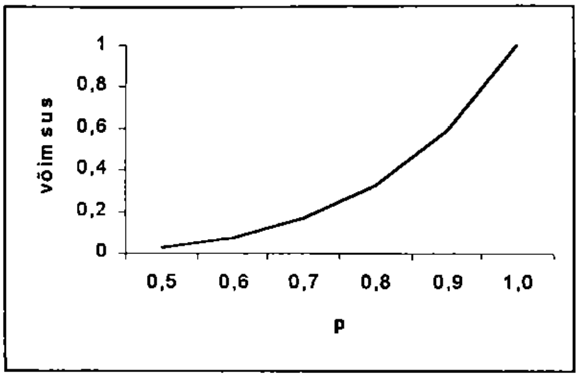
Katse võimalikud tulemused on:

| | Võetakse vastu $H_0$ | Võetakse vastu $H_1$ |
|---|---|---|
| **Kehtib $H_0$** | Korrektne otsus. Tõenäosus=1-$\alpha$ | "Esimest liiki viga" Tõenäosus = $\alpha$ |
| **Kehtib $H_1$** | "Teist liiki viga" Tõenäosus = $\beta$ | Korrektne otsus. Tõenäosus =1-$\beta$ |

Praktikas on kogetud, et esimest liiki viga on raskemate tagajärgedega kui teist liiki viga. Seetõttu valitakse otsuse langetamiseks kriitiline piir nii, et esimest liiki vea tegemise tõenäosus ei ületaks meie valitud tõket. Uurija poolt kehtestatud tõket nimetatakse olulisuse nivooks, mida tähistatakse $\alpha$-ga. Teist liiki vea tõenäosus on $\beta \leq 1 - \alpha$. Testi olulisuse nivoo väärtus määratakse, lähtudes konkreetsest ülesande sisust. Sagedamini kasutatakse väärtusi 0,1; 0,05 ja 0,01. Mida väiksem on olulisuse nivoo, seda rangem on kriteerium ning seda raskem on sisukat hüpoteesi tõestada.

Meie näites $\alpha = 0,5^5 = 0,031$.
Kui näiteks $p = 0,75$, siis $\beta = P(x < 5) = 1 - P(x = 5) = 1 - 0,75^5 = 0,763$.

Sageli nimetetakse $1 - \beta$ testi võimsuseks. Võimsus (samuti ka $\beta$) näitab, kui vale on $H_0$. On võimalik teha võimsuse graafik, niinimetatud võimsuse funktsioon, mis näitab, kuidas $(1 - \beta)$ sõltub $p$ õigest väärtusest:




## 6.2 Hüpoteesid üldkogmi keskväärtuse kohta

Sissejuhatav näide. Instrumenti kasutati viiel erineval korral, et määrata laktoosi taset lahuses. Saadi järgmised tulemused:

3,4 3,3 3,1 2,9 3,4.

Instrumendi tööjuhis lubab mõõtmisviga dispersiooniga $s^2 = 0,05$. Nüüd väidab ettevõtja, et lahus, mida parajasti analüüsiti, on standardlahus, millel on teada õige laktoosi sisaldus $\mu = 3,0$. Kui instrument näitab süstemaatilist kõrvalekallet sellest väärtusest, vajab seade kordaseadmist. Kas instrument vajab kordaseadmist?

Püstitame hüpoteesid:

$H_0 : \mu = 3,0$ (nullhüpotees) üldkogumi keskväärtus on 3,

$H_1 : \mu \neq 3,0$ (sisukas hüpotees).

Hüpoteeside kontrollimiseks kasutatakse statistikut, mis on valimi keskväärtuse $\bar{X}$ ja standardse väärtuse $\mu_0$ normeeritud vahe Z, kus $Z = \frac{\bar{X}-\mu_0}{\sqrt{\frac{\sigma^2}{n}}}$ Kui $H_0$ on õige, siis Z jaotus peab olema standardiseeritud normaaljaotus, st $Z \sim N(0,1)$. $\mu_0$ on $\mu$ väärtus vastavalt nullhüpoteesile. Kaks hüpoteesi on õiged, kui kehtivad järgmised eeldused:

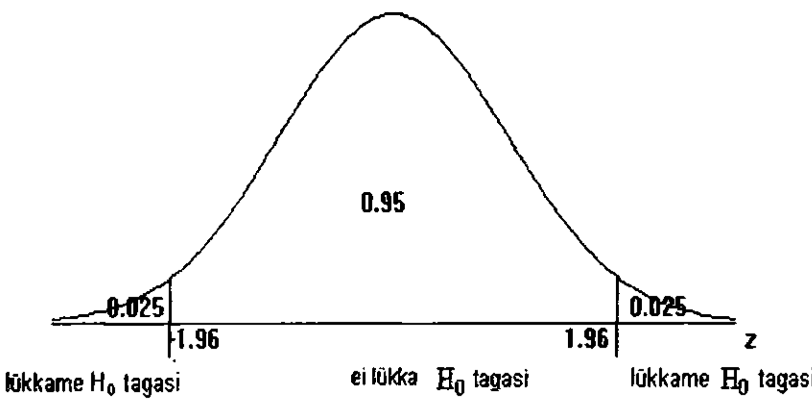
1. Jaotus peab olema normaaljaotus.

See kehtib, kui $x$ jaotus on normaaljaotus (sõltumata valimi mahust $n$) või $x$ jaotus läheneb normaaljaotusele ja n on mõistlikult suur (vastavalt tsentraalsele piirteoreemile).

2. Valimis peavad olema vaatlused vastastikku sõltumatud. See on tagatud, kui on moodustatud juhuslik valim.

Nende kahe eelduse põhjal võime püstitada järgmise otsustusreegli. Arvutame z väärtuse. Kui $z$ on arvuliselt suur, siis kummutame $H_0$. Vastasel juhul jääme $H_0$ juurde. Reegel põhineb väitel, et on vähetõenäoline saada $z$-i suurt väärtust, kui $H_0$ on õige. Seega kaotame usu $H_0$ suhtes, kui $z$ on arvuliselt suur.

Mida peaksime mõistma "suure" või "väikese" $z$-i väärtuse all? Näiteks võime valida $H_0$ kummutamiseks piirid, mis vastavad 5% tõenäosusele, seega olulisuse nivoo $\alpha$ on $0,05$.




Järelikult võime otsustuse reegli põhjal öelda:
kui $z < -1,96$, loeme $H_1$ tõestatuks;
kui $z > 1,96$, loeme samuti $H_1$tõestatuks;
kui $z$ on vahemikus $-1,96 \leqslant z \leqslant 1,96$ jääme $H_0$juurde.

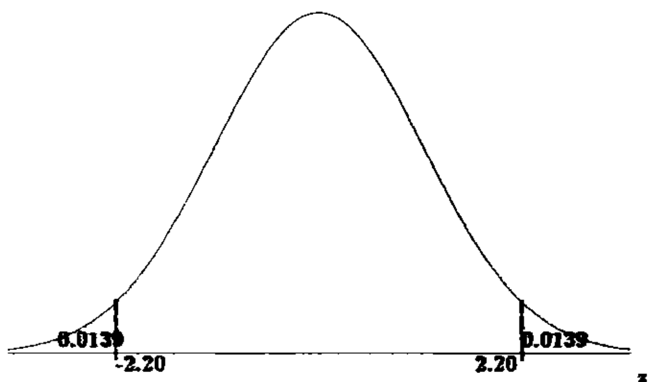
Kasutades seda otsustusreeglit, on risk teha "esimest liiki viga" $\alpha = 0,05$, st lükata tagasi $H_0$, kuigi $H_0$ on õige.

Meie andmete korral arvutame statistiku Z väärtuse $z = \frac{\bar{X}-\mu_0}{\sqrt{\frac{\sigma^2}{n}}} = \frac{3,22-3,0}{\sqrt{\frac{0,05}{5}}} = 2,20$

Kui $z$ >1,96, siis ütleme, et tulemus on oluline, ning kummutame $H_0$. Me ei usu enam, et $\mu = 3$. Sisukat hüpoteesi vastu võttes teeme esimest liiki vea, eksimise risk ei ole suurem kui 5%. Kuna sisuka hüpoteesi vastuvõtmisega kaasnev risk on meie poolt ette määratud, loetakse sisukas hüpotees tõestatuks.

## 6.3 Testi p (tõenäosus-)väärtus

Selle asemel, et kummutada $H_0$ määratud $z$ piiride korral, võime kasutades normaaljaotuse tabelit arvutada, kui tõenäoline on vaadeldav $z$ väärtus. Sealt leiame, et $P(z > 2,20) = 0,0139$ ja $P(z < -2,20) = 0,0139$. Kokkuvõttes oleme saanud tulemuse, kus tõenäosus on $0,0139 + 0,0139 = 0,0278$, kui $H_0$ on õige. Seda nimetatakse $p$-väärtuseks. "Tõenäosus saada tulemust valimis, kui $H_0$ on õige, on 0,0278"




Mõnel juhul on $p$ väärtust raskem arvutada kui anda ainult olulisuse nivoo. Aga arvutiprogrammid saavad $p$ vääruse arvutamisega ja esitamisega kergelt hakkama. Järgnev on programmipaketi Minitab analüüs meie näite andmete põhjal:

**Z-Test**

```
Test of mu = 3.000 vs mu not = 3.000
The assumed sigma = 0.224

Variable    N     Mean    StDev   SE Mean        Z    P-Value
X           5    3.220    0.217     0.100     2.20      0.028
```

### 6.3.1 "Tähe süsteem"

Teaduslikes publikatsioonides on vaja avaldada $p$ väärtust iga kord, kui olulisuse test on tehtud. Siis võib lugeja otsustada, missugune tulemus vastab tema olulisuse nivoole. "Test näitab, et keskväärtus ei ole 3,0 (p=0,028)" On kasulik meeles pidada järgnevat süsteemi:

Kui $p > 0,05$ , siiis tulemus on "mitte oluline"
Kui $0,01 < p \leqslant 0,05$ , siis tulemus on "oluline 5% tasemel", märgitakse *
Kui $0,001 < p \leqslant 0,01$ , siis tulemus on "oluline 1% tasemel", märgitakse **

Kui $p \leqslant 0,001$ , siis tulemus on "oluline 0,1% tasemel", märgitakse ***

Paljud toimetajad ja retsensendid tõrguvad tunnistamast väljendeid nagu: "oluline 9% tasemel", "peaaegu oluline" jne. 5% tase on muutunud standardiks, kui nimetame midagi "oluliseks"

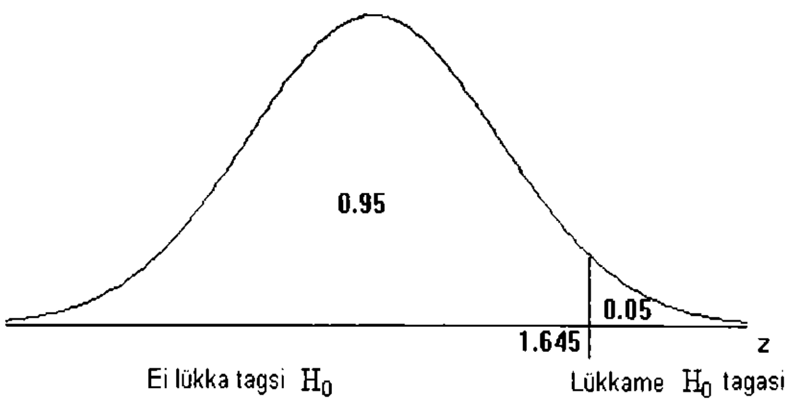
## 6.4 Ühepoolsed alternatiivsed hüpoteesid

Mõnel juhul oleme huvitatud ainult erilistes suundades erinevuse näitamisest. Näiteks võime demonstreerida, et uue toiduratsiooni kasutamine annab suuremat kaalu kasvu kui mingi varajasem või öelda, et kindel keskväärtus on suurem kui 3,0. Need on ühepoolsed hüpoteesid. Oletame, et meie näites võib olla $\mu$ suurem kui 3. Selle uurimiseks püstitame järgmised hüpoteesid:

$H_0$: $\mu = 3,0$ (samuti võime kirjutada $H_0$: $\mu \leqslant 3,0$; tähtis on, et $H_0$ sisaldaks võrdusmärki),

$H_1$: $\mu > 3,0$.

Sel korral peaksime $H_0$-i kummutama vaid juhul, kui z väärtus on suur:




Normaaljaotuse tabeli põhjal näeme, et $H_0$-i peaks tagasi lükkama olulisuse nivoo $\alpha = 0,05$ korral, kui $z > 1,645$. Arvutatud $z$-i väärtus on $z = 2,20$. Kuna see on suurem kui piir 1,645, on tulemus oluline ja kummutame $H_0$.

Alternatiivselt võime arvutada p väärtuse. See on $P(z > 2,20) = 0,0139$.

## 6.5 Hüpoteesi kontrollimine, kui $\sigma$ on teadmata

Sissejuhatav näide.

Instrumenti kasutati viiel erineval korral, et määrata laktoosi taset lahuses. Saadi järgmised tulemused:

3,4 3,3 3,1 2,9 3,4.

Ettevõtja väidab, et lahus, mida parajasti analüüsiti, on standardlahus, millel on teada laktoosi õige sisaldus $\mu = 3,0$. Kui instrument näitab süstemaatilist kõrvalekallet sellest väärtusest, on vaja seade korda seada. Vajab instrument seda?

Püstitame oma hüpoteesid:

$H_0 : \mu = 3,0$,

$H_1 : \mu \neq 3,0$.

Kuna me ei tea populatsiooni dispersiooni $\sigma^2$, kasutame sama ideed nagu usalduspiiride korral. Kasutame tõsiasja, et $t = \frac{\bar{x}-\mu_0}{\sqrt{\frac{s^2}{n}}}$ järgib $t$-jaotust $(n-1)$ vabadusastmega, samadel eeldustel kui vastavate usaldusvahemike juhul. Kahepoolse testi korral näeme $t$-jaotuse tabelist ($\alpha = 0,05$), et peame $H_0$ tagasi lükkama, kui $t > 2,776$ või $t < -2,776$. Meie andmete korral $s^2 = 0,047$, nii saame $t = \frac{3,22-3,0}{\sqrt{\frac{0,047}{5}}} = 2,269$. Kuna vaadeldav $t$ väärtus on väiksem kui piir 2,776, siis jõuame järeldusele, et tulemus ei ole oluline ja jääme $H_0$ juurde.

Meie $t$-jaotuse tabel ei ole küllalt täiuslik, et leida $t$-jaotuse testi põhjal $p$ väärtusi. Analüüsi tegemiseks on mõistlik kasutada mõnda arvutiprogrammi, $p$ väärtus arvutatakse siis järgmiselt:

T-Test of the Mean

Test of mu = 3.0000 vs mu not = 3.0000

| Variable | N | Mean | StDev | SE Mean | T | P-Value |
|---|---|---|---|---|---|---|
| X | 5 | 3.2200 | 0.2168 | 0.0970 | 2.27 | 0.086 |

Kui alternatiivne hüpotees on ühepoolne, siis tulemus peab olema oluline ($p = 0,043$).

Märkused.

1. Isegi kui oleme suutnud näidata, et miski "on oluline", ei tähenda see,et oleme absoluutselt veendunud, et $H_0$ on vale. On risk ($\alpha$), et me eksime.

2. Isegi kui oleme suutnud näidata, et "ei ole oluline", ei tähenda see, et $H_0$ on õige. Isegi siin on risk ($\beta$), et järeldus on vale. $\beta$ võib olla üsna suur, eriti kui valim on väike.

| **Parameeter; hüpotees** | **Punkthinnang** | **Test statistik** | **Eeldused, kommentaarid** |
|---|---|---|---|
| $H_0 : \Theta = \Theta_0$ | $\hat{\Theta}$ | $\left\{\begin{matrix} t \\ v\tilde{o}i \\ z \end{matrix}\right\} = \dfrac{\hat{\Theta} - \Theta_0}{\sqrt{D(\hat{\Theta})}}$ | z kui populatsioonis on $D(\hat{\Theta})$ teada, mis on praktikas ebatavaline; z kasutame lähendina, kui $D(\hat{\Theta})$ on hinnatud valimist ja valim on suur; t kui $D(\hat{\Theta})$ on teadmata ja on hinnatud valimist. |

Testid põhinevad normaaljaotusel. (Vaata eraldi lehte erinevate võimaluste kohta raamatu lõpus.)

## 6.6 Ülesanne

### 6.1

Juhuslik valim koosneb 10-st ülikooli meesõppejõust. Nendel mõõdeti kolesterooli taset veres ja saadi järgmised tulemused (grammi liitri kohta):

3,0 1,8 2,1 1,4 2,5 1,6 1,9 2,8 2,2 2,7.

Üldises meeste populatsioonis kolesterooli keskmine tase on 2,0. Kontrolli hüpoteesi, et ülikooli meesõppejõudude populatsiooni kolesterooli keskväärtus on 2,0.

# Peatükk 7

# Osade suhe tervikusse - järeldused

## 7.1 Tõenäosuse $p$ usalduspiirid

Sissejuhatav näide. Hääletajatest on moodustatud juhuslik valim, kus n = 1000. Nendest 350 (35%) teatasid, et nad hääletaksid sotsiaaldemokraatide poolt, kui hääletus toimuks täna. Kui ebakindel on see väide, st missugune on piiri viga?

Parameeter, mida soovime hinnata, on $p$. See on sotsiaaldemokraatide poolt hääletajate suhe kõikide hääletajate hulgast. Tõenäosuse $p$ hinnanguks on selle suhteline sagedus $\hat{p} = \frac{350}{1000} = 0,35$. Olgu valimis hääletajate arv $x$. Elementide väärtused valimis on kas "0" või "1", sest nad on poolt või ei ole. Seega $x$ (ja seepärast samuti $\hat{p} = x/n$ ) on binoomjaotusega juhuslik suurus (vaata peatükk 4). Kuna meil on suur valim, siis binoomjaotusega juhusliku suuruse jaotus läheneb normaaljaotusele (piirteoreemi põhjal). Erinevates õpikutes on erinevad hinnangud, mida nimetame "suureks valimiks" Sageli on rusikareegel, et nii $np$ kui ka $n(1-p)$ peavad olema suuremad kui 10. Mõned õpikud ütlevad isegi, et nad peavad olema suuremad kui 5. Sama reeglit kasutatakse juhul, kui $p$ on teadmata. Tema asemel kasutame siis tõenäosuse $p$ hinnangut $\hat{p}$. Meie näites on $n\hat{p}$ =350 ja $n(1-\hat{p})$ =650, mis on kindlasti normaaljaotusele lähendamiseks küllaldane.

Kui lisame eelduse, et vaatlused on sõltumatud (mis järeldub juhuslikust valikust), võime kasutada käsitletud usalduspiiride leidmise üldist reeglit. See on:

| | | | | |
|---|---|---|---|---|
| (Parameeter) = | (Hinnang) | $\pm$ | t või z | Hinnangu standardviga |
| $\theta =$ | $\hat\theta$ | $\pm$ | t või z | $\sqrt{D(\hat\theta)}$ |

Parameeter on $p$ ja selle parameetri hinnang on $\hat p$. On vaja teada $\hat p$ dispersiooni. Võib näidata, et $D(\hat p) = \frac{p(1-p)}{n}$ Probleemiks on, et valem sisaldab tundmatut $p$ väärtust. Paneme lihtsalt $\hat p$ selle asemele ja saame dispersioonile mõistliku lähendi $D(\hat p) \approx \frac{\hat p(1-\hat p)}{n}$. Järelikult võime *tõenäosuse p usalduspiiride* valemi kirjutada järgmiselt:

$$\hat p \pm z\sqrt{\frac{\hat p(1-\hat p)}{n}}.$$

Eelnev valem on seotud tõsiasjaga, et meil on suur valim: siis võib $\hat p$ jaotust lähendada normaaljaotusele ning võime kasutada valemis $z$-i.

Sissejuhatava näite andmeid ja normaaljaotuse z väärtust $z = 1,96$ (95%-le vastav vahemik) kasutades saame $0,8 \pm 1,96\sqrt{\frac{0,8(1-0,8)}{100}}$, st 0,35 ± 0,0296. Saadud tulemuse põhjal võime 95 juhul sajast öelda, et tegelik tõenäosuse protsent asub vahemikus (0,32; 0,38). "Piiri viga" on umbes 3%, st kui kasutame valimi põhjal saadud hinnangut 35%, pole tehtud viga 95 juhul sajast suurem kui 3%.

### 7.1.1 Kui suur valim on vajalik tõenäosuse $p$ usalduspiiride leidmiseks?

Kui suurt valimit vajame sotsiaaldemokraatide poolt hääletajate suhte leidmiseks tõenäosusega $p$ 95%-lisete usalduspiiride jaoks eeldusel, et usaldusvahemik oleks lühem kui 0,02?

Eelpool arvutasime tõenäosuse $p$ 95% usalduspiirid valemiga $\hat p \pm z\sqrt{\frac{\hat p(1-\hat p)}{n}}$ Saime tulemuse $0,35 \pm 0,03$. Kui soovime leida piire, mis oleksid lühemad kui 0,02, siis eeldame, et tõenäosus $p$ on ligikaudu sama suur kui varem. Saame $1,96\sqrt{\frac{0,35(1-0,35)}{n}} \leq 0,05$. Leiame lahendi $n$ jaoks: $n \geq \frac{1,96^2 \cdot 0,35 \cdot (1-0,35)}{0,05^2} = 2184,91$ Järelikult antud tingimuste korral vajame valimisse vähemalt 2185 hääletajat.

Selles arvutuses kasutasime vana andmestiku väärtust $\hat p = 0,35$. Kerkib küsimus: mida teeme siis, kui meil ei ole ühtegi vastavat vana andmestikku. Tegelikult on võimalik arvutadada nõutav valimi suurus ka juhul, kui pole teada $p$ (või $\hat p$). Kehtib, et $\frac{p(1-p)}{n}$ suurim väärtus $p = 0,5$. Kasutades seda väärtust, saame kindlasti valimimahu, mis ei olegi nii väike. Kasutame seda meie näites, saame $n \geq \frac{1{,}96^2 \cdot 0{,}5 \cdot (1-0{,}5)}{0{,}05^2} = 2401$. Kui moodustame valimi üle 2400 hääletajaga, on vea piir väiksem kui 0,02 hoolimata hääletajate õigest suhtest.

## 7.2 Hüpoteesi $H_0$ $p = p_0$ kontrollimine

Sissejuhatav näide. Ostsime koti seemneid, mille idanevus oli 90%. Selle kontrollimiseks moodustasime juhusliku valimi $n = 200$ seemnest, millest $x = 166$ seemet idanes.

Soovime kontrollida hüpoteese:

$H_0$: $p = 0,90$ ,

$H_1$: $p \neq 0,90$.

Soovime teha järeldusi parameetri $p$ kohta, mis on tõenäosus selleks, et juhuslikult valitud seeme idaneb. Tõenäosuse $p$ parim hinnang on $\hat{p} = \frac{x}{n} = \frac{166}{200} = 0,83$. Kui vaatlused on sõltumatud, siis $x$ (ja samuti $\hat{p}$) järgib binoomjaotust, aga kuna meie "rusikareegel" on rahuldatud, siis võime jaotust normaaljaotusega lähendada. Sel juhul $n \ p_0 = 180$ ja $n \ (1 - p_0) = 20$. Mõlemad on suuremad kui 10, seega võime kasutada $Z = \frac{\hat{p}-p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$, mis on (ligikaudu) standardnormaaljaotus.

Märkused. Kasutame $p_0$ väärtust (nullhüpoteesist) murru nimetajas, sest usume $H_0$ kehtivust seni, kuni (kuidagi) tõestame, et ta on väär. Kui kasutada sama normaaljaotuse lähendit usalduspiiride leidmisel, siis kasutame seda hinnangu $\hat{p}$ asemel lugejas.

Kahepoolse testi korral, kui olulisuse nivoo on 5%, kummutame $H_0$, kui |z|>1,96. Kui me valime testi 1%-lise olulisuse nivooga, siis vastavad piirid on |z|>2,576. Meie näites saame $z = \frac{0{,}83-0{,}9}{\sqrt{\frac{0{,}9 \cdot 0{,}1}{200}}} = -3,30$. Saadud tulemus on oluline mõlemal juhul, 5% tasemel ja 1% tasemel. Võtame vastu sisuka hüpoteesi $H_1$. $p$ väärtus (kahepoolse testi puhul) on p=2·0,0005=0,001. Tulemus on koguni "***" tähtsusega.

## 7.3 $(p_A - p_B)$ usalduspiirid

Sissejuhatav näide. On tehtud katse, selgitamaks kahe erineva pestitsiidi mõju lutikatele. 200 lutikat töödeldi pestitsiidiga A ja 180 lutikat samal ajal pestitsiidiga B. Lutikatest, keda töödeldi A-ga, suri 140 ja teistest, keda töödeldi B-ga, suri 100. Tähistame lutikate surmatõenäosused, vastavalt $p_A$ ja $p_B$. Arvuta ja interpreteeri 95%-lised $(p_A - p_B)$ usalduspiirid.

Parameeter, mille kohta soovime järeldusi teha, on $p_A - p_B$. Mõistlik hinnang on $\hat{p}_A - \hat{p}_B$ Selleks, et saaks kasutada usalduspiiride standardset meetodit, peame teadma:

1. Missugune on $\hat{p}_A - \hat{p}_B$ jaotus?

Hinnangu avaldises mõlemad (st $\hat{p}_A$ ja $\hat{p}_B$ ) järgivad binoomjaotust. Eespool näitasime, et binoomjaotust võime lähendada normaaljaotusele, kui $n$ on suur, nagu see hetkel ongi. Niisiis, kui eeldame, et iga $\hat{p}_A$ ja $\hat{p}_B$ järgib normaaljaotust (vähemalt ligikaudu), siis ka $\hat{p}_A - \hat{p}_B$ jaotus järgib ligikaudu normaaljaotust. Normaaljaotusega muutujate summad ja vahed on samuti normaaljaotusega.

2. Missugune on $\hat{p}_A - \hat{p}_B$ dispersioon?

Kehtib $D(\hat{p}_A - \hat{p}_B) = D(\hat{p}_A) + D(\hat{p}_B)$, kui $\hat{p}_A$ ja $\hat{p}_B$ on sõltumatud. Järelikult, $D(\hat{p}_A - \hat{p}_B) = D(\hat{p}_A) + D(\hat{p}_B) = \frac{p_A(1-p_A)}{n_A} + \frac{p_B(1-p_B)}{n_B}$ Me ei tea $p_A$ ja $p_B$. Peame tegema teise lähendi: kasutama hinnanguid $\hat{p}_A$ ja $\hat{p}_B$ dispersiooni avaldises, siis saame dispersiooni hinnangu: $\hat{D}(\hat{p}_A - \hat{p}_B) = \frac{\hat{p}_A(1-\hat{p}_A)}{n_A} + \frac{\hat{p}_B(1-\hat{p}_B)}{n_B}$

*Usalduspiirid* võib nüüd kirjutada järgmiselt

$$\hat{p}_A - \hat{p}_B \pm z\sqrt{\frac{\hat{p}_A(1-\hat{p}_A)}{n_A} + \frac{\hat{p}_B(1-\hat{p}_B)}{n_B}}.$$

Kasutame sissejuhatava näite andmeid $\hat{p}_A = 140/200 = 0,7$ ja $\hat{p}_B = 100/180 = 0,56$. 99%-lised usalduspiirid võib arvutada valemist

$0,70 - 0,56 \pm 2,576\sqrt{\frac{0,70 \cdot 0,30}{200} + \frac{0,56 \cdot 0,44}{180}}$, millest saame $0,14 \pm 0,127$ Oleme 99 juhul sajast veendunud, et vahemik $(0,013; 0,267)$ sisaldab õiget $p_A - p_B$ väärtust. Kuna vahemik ei sisalda väärtust 0, siis see tähendab, et kahel töötlusel on lutikatele erinev mõju.

## 7.4 Hüpoteesi $H_0$: $p_A = p_B$ kontrollimine

Sissejuhatav näide. Tehti katse, kus võrreldi kahe erineva pestitsiidi mõju lutikatele. 200 lutikat töödeldi pestitsiidiga A ja 180 lutikat pestitsiidiga B. Lutikatest, keda töödeldi A-ga, suri 140 ja lutikatest, keda töödeldi B-ga, suri 100. Kontrollime hüpoteesi $H_0$: $p_A = p_B$.

Situatsioon on sarnane usalduspiiride leidmisega. Soovime teha järeldust parameetri $p_A - p_B$ kohta. Hinnangu $\hat{p}_A - \hat{p}_B$ jaotust võib lähendada normaaljaotusega, kui valim on suur. Dispersiooni avaldis hinnangule $\hat{p}_A - \hat{p}_B$ on: $D(\hat{p}_A - \hat{p}_B) = D(\hat{p}_A) + D(\hat{p}_B) = \frac{p_A(1-p_A)}{n_A} + \frac{p_B(1-p_B)}{n_B}$. Eelpool dispersiooni avaldises usalduspiiride arvutamisel lihtsalt asendasime $p_A$ ja $p_B$ vastavalt $\hat{p}_A$ ja $\hat{p}_B$. Selline lähenemine töötab üsna hästi testide jaoks, aga hüpoteeside kontrollimisel on võimalik leida veel parem hinnang.

Kui kontrollime hüpoteesi, siis kasutame teststatistikut eeldusel, et nullhüpotees on õige. Sel juhul võime eeldada, et $p_A = p_B = p$, kuni oleme veendunud, et see on vale. Kui $H_0$ on õige, siis võime leida parima $p$ hinnangu, mis saadakse kombineerides andmeid kahest valimist ühte hinnangusse: $\hat{p}_0 = \frac{n_A\hat{p}_A + n_B\hat{p}_B}{n_A + n_B}$ (see on surnud lutikate summaarne arv jagatud lutikate kogu arvuga). Asendame dispersiooni avaldises $\hat{p}_A$ ja $\hat{p}_B$ $\hat{p}$-ga, saame $\hat{D}(\hat{p}_A - \hat{p}_B) = \frac{\hat{p}_0(1-\hat{p}_0)}{n_A} + \frac{\hat{p}_0(1-\hat{p}_0)}{n_B} = \hat{p}_0(1-\hat{p}_0)\left(\frac{1}{n_A} + \frac{1}{n_B}\right)$. Hüpoteesi $H_0$ kontrollimiseks arvutame $z = \frac{\hat{p}_A - \hat{p}_B}{\sqrt{\hat{p}_0(1-\hat{p}_0)\left(\frac{1}{n_A} + \frac{1}{n_B}\right)}}$ ja lükkame $H_0$ tagasi valitud usaldusnivoo juures, kui $z$ on väljaspool kriitilise väärtuse piire.

Meie näite korral $\hat{p}_0 = \frac{140+100}{200+180} = 0,63$. Teststatistik on $z = \frac{0,70-0,56}{\sqrt{0,63(1-0,63)\left(\frac{1}{200} + \frac{1}{180}\right)}} = 2,91$. See on väljaspool 5%-lisi piire (1,96), väljaspool 1%-lisi piire (2,576), aga seespool 0,1%-lisi piire (3,291). $p$ väärtus on 0,0036. Tulemus on "**" tähtsusega, mis näitab kahe pestitsiidi efektiivsuse erinevust.

## 7.5 Ülesanded

**7.1**

Uusim saavutus akne ravimiseks on ravim, mida kutsutakse "retinoidhape" Testis uuriti selle ravimi mõju 20 patsiendil, kellel oli äge akne. 20-st testitud patsienst 14-l märgati ravimi kasutamisel paranemist.

Küsimused:

A. Arvuta 99%-lised tõenäosuse $p$ usalduspiirid, st ravimit tarbivate paranevate patsientide suhe üldkogumis. Samuti selgita oma arvutuste aluseks olevaid oletusi.

B. Arvatakse, et usalduspiirid ülesandes A on liiga laiad. Seetõttu planeeri uus katse. Selles peaks valimi maht $n$ olema nii suur, et 95%-lise tõenäosuse $p$ usalduspiiride laius oleks väiksem kui 0,1. Leia valimi maht $n$.

# Peatükk 8

# Kahe keskväärtuse võrdlemine

## 8.1 Sissejuhatus

Sagedane probleem bioloogilise andmestiku uurimisel on kahe keskväärtuse võrdlemine. Mõned näited: eksperimentaalne grupp koos kontrollgrupiga, ühe nisu sordi saagi võrdlemine mõne teise nisusordi saagiga, patsiendi vererõhu näitaja võrdlemine enne ja peale ravikuuri. Parameeter, mida niisuguste näidete korral soovime uurida, on $\mu_1 - \mu_2$. Kasutatavate meetodite valik sõltub seejuures sellest, mil viisil on andmed kogutud. Põhimõtteliselt võime eristada kahte olukorda:

**1. Paarikaupa andmed**

Paarikaupa andmed tähendavad seda, et andmestiku väärtused kuuluvad paarikaupa kokku ("plokid"). Mõned näited:

• Ühel ja samal indiviidil tehti üks mõõtmine (x) enne ravimist, teine mõõtmine (y) peale ravimist.

• Ühel ja samal isikul tehti mõõtmised vasakul käel (x) ja paremal käel (y).

• 20-st katseloomast moodustati 10 paari. Iga paari loomad olid ühe ema järglased.

• 10 erinevat põldu jagati kahte ossa. Igal põllul kasvatati liiki A ühel osal ja liiki B teisel osal.

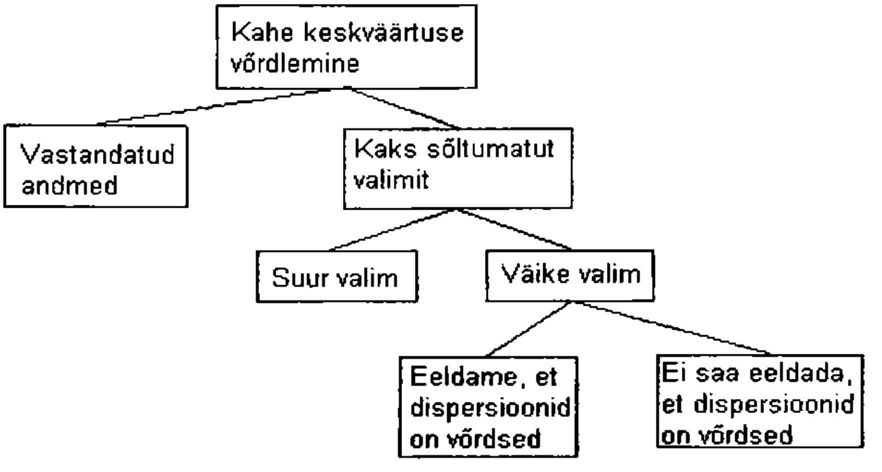
**2. Kaks sõltumatut gruppi**

Katseühikute rühmitamine on sel juhul tehtud täiesti juhuslikult. Mõned näited:

• Liiki A kasvatati 10-l põllul, liiki B 10-l teisel põllul.

• 10 katselooma on juhuslikult määratud sööma sööta A, 10 teist looma söövad sööta B.

Kui tegemist on kahe sõltumatu grupiga, siis võib andmeanalüüsi jagada alajaotusteks. Järgmisel skeemil esitame kokkuvõtvalt erinevad võimalused:




# 8.2 Usalduspiirid ja hüpoteesid $\mu_1 - \mu_2$ korral

## 8.2.1 Paarikaupa andmed

Sissejuhatav näide. 1. peatükis vaatlesime andmestikku, mis oli avaldatud Charles Darwini poolt 1876. aastal. Darwin püüdis näidata, et taimed, mida paljundati risttolmeldamise teel, olid kiirema kasvuga kui taimed, mida paljundati isetolmeldamise teel. Darwin tegi taimepaaridega katse, kus igast paarist üks oli isetolmeldatud ja teine oli risttolmeldatud. Kõiki taimepaare kasvatati koos ühesugustes tingimustes. Sageli on see hea katseplaan, sageli viitab plokk katseplaanile.

Taimepaaride kõrgused (tollides) on esitatud järgmises tabelis:

| Paari nr | $x_k$ risttolmeldatud | $x_s$ isetolmeldatud | $d = x_k - x_s$ |
|---|---|---|---|
| 1 | 23,5 | 17,4 | 6,1 |
| 2 | 12,0 | 20,4 | -8,4 |
| 3 | 21,0 | 20,0 | 1,0 |
| 4 | 22,0 | 20,0 | 2,0 |
| 5 | 19,1 | 18,4 | 0,7 |
| 6 | 21,5 | 18,6 | 2,9 |
| 7 | 22,1 | 18,6 | 3,5 |
| 8 | 20,4 | 15,3 | 5,1 |
| 9 | 18,3 | 16,5 | 1,8 |
| 10 | 21,6 | 18,0 | 3,6 |
| 11 | 23,3 | 16,3 | 7,0 |
| 12 | 21,0 | 18,0 | 3,0 |
| 13 | 22,1 | 12,8 | 9,3 |
| 14 | 23,0 | 15,5 | 7,5 |
| 15 | 12,0 | 18,0 | -6,0 |

Kas Darwini väide oli õige? Sellele küsimusele võime vastata, kui arvutame ja interpreteerime usalduspiirid kahe keskväärtuse erinevusele $(\mu_k - \mu_s)$ või kontrollime hüpoteesi $H_0$: $\mu_k - \mu_s = 0$. See on sama, kui teha järeldus iga paari $x_k$ ja $x_s$ vahelise erinevuse $d$ jaoks.

$d$ keskväärtus on $\bar{d} = \frac{39,1}{15} = 2,607$ ja dispersioon $s_d^2 = \frac{\sum d^2 - \frac{(\sum d)^2}{n}}{(n-1)} = \frac{412,9 - \frac{39,1^2}{15}}{14} = 22,21$.

Meie parameeter on $\mu_k - \mu_s = \mu_d$, kus $\mu_d$ tähendab üldkogumi paarikaupa väärtuste erinevuse $d$ keskväärtust. Kontrollime hüpoteesi, et kahel erinevalt paljundatud taimel on võrdne keskmine kõrgus. Võime kontrollida kahepoolset hüpoteesi:

$H_0$: $\mu_d = 0$,

$H_1$: $\mu_d \neq 0$;

või kontrollida ühepoolset hüpoteesi:

$H_1$: $\mu_d > 0$ (või $H_1$: $\mu_d < 0$).

Hüpoteesi valik sõltub küsimusest, aga mitte andmestikust. Antud juhul võib olla eelistatud ühepoolne hüpotees. Darwin soovis näidata, et risttolmeldatud taimed on kõrgemad. Hüpoteesi kontrollimiseks arvutame t-statistiku $t = \frac{\bar{d} - 0}{\sqrt{\frac{s_d^2}{n}}} = \frac{2,607}{\sqrt{\frac{22,21}{15}}} = 2,142$. Kahepoolse $t$-test korral on piirid $\pm 2,145$, kui olulisuse nivoo $\alpha = 0,05$ ja vabadusastmed on $(n-1) = 14$. Hetkel on arvutatud $t$ väärtus just nende piiride vahel. Vastav ühepoolne piir on $1,761$.

Järelikult hüpoteesi $H_1$ võib lugeda tõestatuks ka ühepoolse testi põhjal. Minitabi väljatrükist loeme välja sama järelduse.

**T-Test of the Mean**

Test of mu = 0.0000 vs mu > 0.0000

| Variable | N | Mean | StDev | SE Mean | T | P |
|---|---|---|---|---|---|---|
| D | 15 | 2.61 | 4.71 | 1.22 | 2.14 | 0.025 |

Järelikult võime toetada Darwini väidet, et risttolmeldatud taimed on keskmiselt kõrgemad kui isetolmeldatud taimed.

Võime arvutada ka $\mu_d$ usalduspiirid. Illustratsiooniks leiame kahepoolsed piirid. Arvutame usaldusvahemiku $2,607 \pm t_{0,975;14}\sqrt{\frac{22,21}{15}}$ ; $2,607 \pm 2,61$; $(0,00;\ 5,22)$. Sama tulemust näeme järgnevas Minitab'i väljatrükis:

**Confidence Intervals**

| Variable | N | Mean | StDev | SE Mean | 95.0 % CI |
|---|---|---|---|---|---|
| D | 15 | 2.61 | 4.71 | 1.22 | (-0.00, 5.22) |

Siin on usalduspiirid kahepoolsed, nad vastavad kahepoolsele testile. Näeme, et usaldusvahemik sisaldab väärtust 0, mille põhjal võime järeldada, et erinevus kahe grupi vahel ei ole oluline.

### 8.2.2 Sõltumatud valimid

Oletame, et soovime võrrelda kahe töötluse keskväärtusi. Katseplaan on randomiseeritud (vt ptk 9) st ei ole kasutatud andmete sobitamist, mis tähendab, et katse ühikud on juhuslikult määratud kahele töötlusele.

Soovime teha otsustusi parameetri $\mu_1 - \mu_2$ kohta. Kasutame valimi keskmiste erinevust $\bar{x}_1 - \bar{x}_2$ kui hinnangut $\mu_1 - \mu_2$ kohta. Kehtib $E(\bar{x}_1 - \bar{x}_2) = \mu_1 - \mu_2$. Järelikult meie hinnang on nihketa. Usalduspiiride leidmiseks ja hüpoteeside kontrollimiseks vajame dispersiooni hinnangut ning hinnangu jaotust.

Dispersiooni jaoks kehtib $D(\bar{x}_1 - \bar{x}_2) = 1^2 D(\bar{x}_1) + (-1)^2 D(\bar{x}_2) =$
(eeldades sõltumatust rühmade vahel)

$=\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$(eeldades sõltumatust rühmade sees).

Kui üldkogumid on mõlemas grupis normaaljaotusega, siis ka vastavad keskväärtused $\bar{x}_1$ ja $\bar{x}_2$ on normaaljaotusega, st meie hinnang $\bar{x}_1 - \bar{x}_2$ järgib normaaljaotust. Kui üldkogumid ei ole normaaljaotusega, siis suurte valimite korral hinnangu $\bar{x}_1 - \bar{x}_2$ jaotus on lähendatud normaaljaotusele. See järeldub tsentraalsest piirteoreemist. Missugust valimit peaksime vaatlema suure valimina, on otsustuse asi.

Kui dispersioon on teada ning kasutame normaaljaotust, võime arvutada $\mu_1 - \mu_2$ *usalduspiirid*

$$\underline{\mu_1 - \mu_2} = (\bar{x}_1 - \bar{x}_2) - z\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}},$$
$$\overline{\mu_1 - \mu_2} = (\bar{x}_1 - \bar{x}_2) + z\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

Samal viisil võime kontrollida hüpoteesi $H_0 : \mu_1 - \mu_2 = 0$, arvutades $z = \frac{(\bar{x}_1 - \bar{x}_2) - 0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$ Tegelikus elus me ei tea üldkogumi dispersiooni. Lahendame selle probleemi kahel viisil. Lihtsam lahendus on see, kui eeldame, et üldkogumi dispresioonid on võrdsed.

### 8.2.3 Võrdsed dispersioonid, väike valim

Näide. Võrreldi kahte erinevalt toidetud pullide rühma. 8 looma jaotati nii, et nendest 4 said sööta A ja 4 said sööta B. Märgiti üles kaalu juurdekasv:

| sööt 1 | sööt 2 |
|---|---|
| 330 | 290 |
| 360 | 320 |
| 400 | 340 |
| 350 | 370 |

Selle andmestiku põhjal saame arvutada hetkel vajalikud valimi statistikud:

sööt 1: $n_1 = 4$; $\bar{x}_1 = 360$; $s_1 = 29,44$,

sööt 2: $n_2 = 4$; $\bar{x}_2 = 330$; $s_2 = 33,67$

Meid huvitav parameeter on $\mu_1 - \mu_2$, mis näitab kaalu juurdekasvu keskmist erinevust. Selle parameetri hinnang on $\bar{x}_1 - \bar{x}_2$ ning hinnangu dispersioon on $D(\bar{x}_1 - \bar{x}_2) = D(\bar{x}_1) + D(\bar{x}_2) = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$

Paljudel juhtudel on sobiv eeldada, et dispersioonid on võrdsed ning ei sõltu töötlusest. Dispersioon võib kirjeldada näiteks loomulikku juhuslikku varieeruvust loomade vahel, mis ei ole mõjutatud sööda tüübist. On ka kindel tehniline põhjus võrdse dispersiooni eeldamiseks. Näeme, et siis on võimalik üsna lihtsal viisil leida dispersiooni hinnangu vabadusastmete arvu.

Kui eeldame võrdset dispersiooni, siis vajame dispersiooni hinnangut, mis on mõlemal rühmitaval tunnusel ühine. Dispersiooni tähistame $\sigma^2$ Eeldame, et $\sigma_1^2 = \sigma_2^2 = \sigma^2$ Saab näidata, et dispersiooni $\sigma^2$ hinnang on $\hat{\sigma}^2 = \frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}$, kus vabadusastmete arv on $(n_1 + n_2 - 2)$.

Kasutades neid põhimõtteid t-testi jaoks, kontrollime hüpoteese:

$H_0 : \mu_1 = \mu_2$ (st $H_0 : \mu_1 - \mu_2 = 0$),

$H_1 : \mu_1 - \mu_2 \neq 0$.

Kasutame teststatistikut $t = \frac{\bar{x}_1-\bar{x}_2}{\sqrt{\hat{\sigma}^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$, kus $\hat{\sigma}^2 = \frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}$.

Test statistikul t on $(n_1 + n_2 - 2)$ vabadusastet.

Meie näites saame dispersiooni hinnangu $\hat{\sigma}^2 = \frac{3\cdot 866,67+3\cdot 1133,33}{4+4-2} = \frac{6000}{6} = 1000$, mille põhjal arvutame t väärtuse: $t = \frac{360-330}{\sqrt{1000(\frac{1}{4}+\frac{1}{4})}} = 1,34$. Kriitiline väärtus 6 vabadusastme juures on 2,447 (olulisuse nivoo on 5%). Järelikult oleme tõestanud $H_1$. Ei saa öelda, et kaks erinevat sööta annavad keskmiselt ühesuguse kaalu juurdekasvu.

Minitabi väljatrükk selle kohta on järgmine:

**Two Sample T-Test and Confidence Interval**

```
Twosample T for Fodder A vs Fodder B
           N     Mean     StDev   SE Mean
Fodder A   4    360.0      29.4        15
Fodder B   4    330.0      33.7        17

95% C.I. for mu Fodder A - mu Fodder B: ( -25,  85)
T-Test mu Fodder A = mu Fodder B (vs not =): T= 1.34  P=0.23  DF=  6
Both use Pooled StDev = 31.6
```

Näeme, et väljatrükk sisaldab teststatistikut ja samuti ka $\mu_1 - \mu_2$ usalduspiire. Usaldusvahemik arvutatakse valemi $\bar{x}_1 - \bar{x}_2 \pm 2,447\sqrt{\hat{\sigma}^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}$ põhjal, kus (nagu ennegi) $\hat{\sigma}^2 = \frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}$ Arv 2,447 on 6 vabadusastmega kriitiline t väärtus 95%-liste usalduspiiride jaoks (tabeli põhjal).

### 8.2.4 Valim on väike ning dispersioonid ei ole võrdsed

Kui ei ole võimalik eeldada, et dispersioonid on kahes grupis võrdsed, võime kasutada lähendamist, mis on tuntud erinevate nimedega: Behrens-Fisher'i probleem, Satterwaite'i lähendamine ja teised.

Meid huvitav parameeter on ikka $\mu_1 - \mu_2$, millele kasutame hinnangut $\bar{x}_1 - \bar{x}_2$. Suure valimi korral on hinnangu dispersioon $\hat{D}(\bar{x}_1 - \bar{x}_2) = \frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}$ Selle arvutamiseks pole teada vabadusastmete arv. Saab näidata, et selle saab leida valemi $\gamma = \frac{\left[s_1^2/n_1 + s_2^2/n_2\right]^2}{\frac{\left[s_1^2/n_1\right]^2}{n_1 - 1} + \frac{\left[s_2^2/n_2\right]^2}{n_2 - 1}}$ põhjal. See on üsna komplitseeritud avaldis, aga arvutid saavad sellega lihtsalt hakkama. Kui eeldame, et valimite dispersioonid ei ole võrdsed, siis Minitab annab järgmise tulemuse:

**Two Sample T-Test and Confidence Interval**

```
Two sample T for Fodder A vs Fodder B
          N     Mean     StDev   SE Mean
Fodder A  4     360.0     29.4        15
Fodder B  4     330.0     33.7        17

95% CI for mu Fodder A - mu Fodder B: ( -27,  87)
T-Test mu Fodder A = mu Fodder B (vs not =): T= 1.34  P=0.24  DF=  5
```

Nagu näeme, saime praegu 5 vabadusastet 6 asemel. See on hind, et loobusime võrdsete dispersioonide eeldusest. Kuid järeldused on samad, mis eespool.

### 8.2.5 Valimid on suured

Suure valimi korral võime kasutada lähendamist, kui soovime kiiresti otsustada, kas erinevus kahe töötluse vahel on oluline või mitte.

Sissejuhatav näide. Võrreldakse kahe erinevat tõugu põrsaste keskmist sünnikaalu. Kummastki tõust on põrsaste arv juhuslikult valitud ning nende valimite kohta on antud arvutatavad statistikud:

tõug 1: $n_1 = 61$; $\bar{x}_1 = 1,33$; $s_1 = 0,25$;

tõug 2: $n_2 = 40$; $\bar{x}_2 = 1,58$; $s_2 = 0,19$.

Parameeter, mida soovime hinnata, on $\mu_1 - \mu_2$, mis väljendab keskmist erinevust sünnikaalus. Selle parameetri hinnanguna kasutame $\bar{x}_1 - \bar{x}_2$ Kuna meil on suured valimid, võime arvata, et $\bar{x}_1$ ja $\bar{x}_2$ on ligikaudu normaaljaotusega, isegi kui kaalud ise ei ole normaaljaotusega. Eeldamegi, et hinnang $\bar{x}_1 - \bar{x}_2$ järgib normaaljaotust.

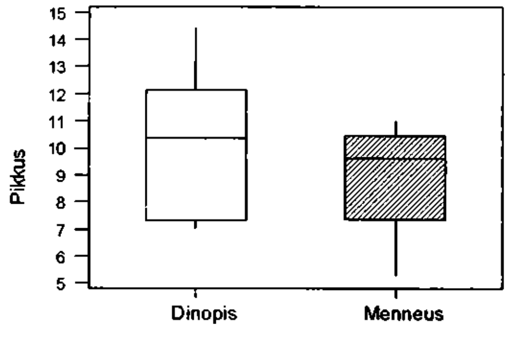
Esitame hinnangu dispersiooni $D(\bar{x}_1 - \bar{x}_2) = D(\bar{x}_1) + D(\bar{x}_2) = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$. Dispersiooni hinnang on $\hat{D}(\bar{x}_1 - \bar{x}_2) = \frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}$ Me ei tea, mitu vabadusastet on haaratud sellesse dispersiooni, aga kuna valim on suur, võime lihtsalt öelda "paljud" vabadusastmed ning kasutada $z$-jaotust. Kontrollime hüpoteese $H_0 : \mu_1 = \mu_2$, $H_1 : \mu_1 \neq \mu_2$.Arvutame $z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{1{,}33 - 1{,}58}{\sqrt{\frac{0{,}25^2}{61} + \frac{0{,}19^2}{40}}} = 5,69.$ See on oluline võrreldes 1,96 (5%); 2,576 (1%) või 3,291 (0,1%). $p$ väärtus on tegelikult null.

95%-lised usalduspiirid sünnikaalude erinevuse jaoks on $\bar{x}_1 - \bar{x}_2 \pm 1,96\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$, millest saame $-0,25 \pm 0,086$.

### 8.2.6 Kahe üldkogumi keskväärtuse võrdlemine Minitab'iga

Kahe üldkogumi keskväärtuse võrdlemiseks vaatleme järgmist näidet. Katses olid kaks ämblikuliiki: Dinopis ja Menneus. Sooviti võrrelda, kumma liigi ämblikud püüavad keskmiselt suuremat saaki. Kummastki liigist valiti 10 ämblikku ja mõõdeti nende poolt püütud saagi suurus (mm-tes). Tulemused on esitatud jägnevas tabelis. Andmestikust esimese mulje saamiseks joonistame mõlema liigi jaoks karpdiagrammi.

| Dinopis | Menneus |
|---|---|
| 12,9 | 10,2 |
| 10,2 | 6,9 |
| 7,4 | 10,9 |
| 7,0 | 11,0 |
| 10,5 | 10,1 |
| 11,9 | 5,3 |
| 7,1 | 5,3 |
| 9,9 | 7,5 |
| 14,4 | 10,3 |
| 11,3 | 9,28,8 |




Kasutame Minitabi ja teeme kaks analüüsi. Esimesel korral eeldame, et üldkogumite dispersioonid on võrdsed.

### Two Sample T-Test and Confidence Interval

```
Twosample T for Dinopis vs Menneus
           N      Mean      StDev   SE Mean
Dinopis   10     10.26       2.51      0.79
Menneus   10      9.02       1.90      0.60

95% C.I. for mu Dinopis - mu Menneus: ( -0.85,  3.33)
T-Test mu Dinopis = mu Menneus (vs not =): T= 1.25  P=0.23  DF=  18
Both use Pooled StDev = 2.23
```

Teisel korral ei eeldata dispersioonide võrdsust.

### Two Sample T-Test and Confidence Interval

```
Twosample T for Dinopis vs Menneus
           N      Mean      StDev   SE Mean
Dinopis   10     10.26       2.51      0.79
Menneus   10      9.02       1.90      0.60

95% C.I  for mu Dinopis - mu Menneus: ( -0.87,  3.35)
T-Test mu Dinopis = mu Menneus (vs not =): T= 1.25  P=0.23  DF=  16
```

Mõlemad analüüsid on kooskõlas: ükskõik missuguse eelduse teeme, võime järeldada, et vaadeldes keskmist saagi suurust ($p = 0{,}23$), ei ole erinevus ämblike liikide vahel oluline. Järelikult meil ei ole mingit põhjust öelda, et kaks erinevat liiki ämblikku püüavad erineva suurusega saaki.

## 8.3 Ülesanded

### 8.1

Mõõdeti kahe erinevat tõugu lehmade keskmist piimatoodangut:

| Tüüp | Keskmine | St.hälve | n |
|---|---|---|---|
| A | 18,8 | 2,4 | 106 |
| B | 17,2 | 5,2 | 54 |

Kontrolli hüpoteesi, kas kahe erinevat tõugu lehmade keskmine piimatoodang on võrdne. Missuguseid eeldusi on vaja teha analüüsi käigus?

### 8.2

Katses kasvatati kahte erinevat sorti otra. Tulemused (vakk/aakri kohta) on esitatud järgnevas tabelis:

| Liik | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Trebi | 41,2 | 19,3 | 45,5 | 63,9 | 63,8 | 44,2 | 42,5 | 53,0 |
| Svanota | 39,4 | 30,8 | 44,5 | 51,5 | 41,1 | 26,5 | 35,7 | |

A. Kontrolli hüpoteesi, et kahel erineval odrasordil on keskmine saagikus ühesugune.

B. Arvuta 95%-lised usalduspiirid kahe sordi keskmise saagikuse erinevusele.

### 8.3

Kui tõsine mõju lapse arengule enne sündi on ema alkoholi tarvitamine?[1] Selleks, et uurida seda küsimust, leiti kuus naist kes olid olnud kroonilised alkohoolikud raseduse ajal. Nende lastele tehti 7-aastaselt IQ test järgmiste tulemustega:

78 102 64 84 100 70

A. Arvuta keskväärtus ja dispersioon IQ testi tulemuste kohta.

B. 46 naisest koosneva kontrollgrupi, kes ei olnud alkohoolikud, laste IQ testi tulemused on: $\bar{X} = 99$ ja $\sum(X_i - \bar{X})^2 = 11520$. Kontrolli hüpoteesi, et kahel naiste grupil on laste IQ testi tulemused keskmiselt võrdsed.

C. Kas võib eelnevat näidet pidada katseks? Selgita miks/miks mitte. Samuti arutle, kas ülesande B testi lõppjäreldused sõltuvad katsest tulenevatest andmetest.

### 8.4

Järgnev andmestik on uurimuse tulemus, mis selgitab, kas malaaria esineb võrdselt kahes ökoloogiliselt erinevas piirkonnas A ja B. Vere valimid on võetud juhuslikult valitud indiviididelt iga piirkonna 7-st erinevast kohast. Iga koha valimi protsendid näitavad, mil määral malaaria kindlaks määrati. Tulemused olid:

[1] Võetud Jones et al, Lancet, June 1974

| Koht | Piirkond A | Piirkond B |
|---|---|---|
| 1 | 5 | 13 |
| 2 | 12 | 18 |
| 3 | 13 | 11 |
| 4 | 2 | 17 |
| 5 | 4 | 14 |
| 6 | 3 | 11 |
| 7 | 8 | 21 |

A. Kontrolli, kas malaaria esineb võrdselt kahes piirkonnas.

B. Arvuta 95%-lised usalduspiirid kahe piirkonna erinevusele.

C. Missuguseid eeldusi on vaja analüüsiks?

**8.5**

Troopilisel maal kasutatakse pühvleid või härgi suhkru väljapressimise tööriista juures. Loomad töötavad 90 minutit. Sel ajal kõnnivad nad ringjooneliselt, kinnitatuna hoova külge. 10 pühvlit ja 8 härga osalesid katses, mille käigus mõõdeti nende südame tööd enne ja pärast tööperioodi. Tulemused olid järgmised:

| Pühvel | | Härg | |
|---|---|---|---|
| Südamelöögid enne tööd | Südamelöögid pärast tööd | Südamelöögid enne tööd | Südamelöögid pärast tööd |
| 50 | 52 | 65 | 68 |
| 50 | 58 | 65 | 72 |
| 56 | 55 | 59 | 65 |
| 54 | 59 | 65 | 70 |
| 43 | 46 | 63 | 71 |
| 48 | 55 | 53 | 63 |
| 53 | 52 | 56 | 64 |
| 46 | 48 | 61 | 68 |
| 51 | 55 | | |
| 47 | 50 | | |

A. Kontrolli hüpoteesi, et südame töö muutus on kahel loomaliigil ühesugune.

B. Missuguseid eeldusi on vaja analüüsiks?

**8.6**

Claes Lilja[2] artiklis olid andmed, kus oli mõõdetud jaapani vuti mune kahelt erinevalt liinilt. Üks liin ("P-liin") oli valitud vuti kasvatamise ajal kindlast generatsioonist nii, et iga generatsiooni kõrgeima kehakaaluga nelja nädala vanustel indiviididel lubati paljuneda. Teine liin ("C-liin") oli tavaline, selekteerimata

[2]Scand J agric res 17:149-151, 1987

grupp. Järgnevas tabelis on esitatud liinide kohta mõned muutujad (esimesena P-liin, siis C-liin):

| Muutuja | keskmine | SD | n | keskmine | SD | n |
|---|---|---|---|---|---|---|
| Värskete munade kaal | 14,1 | 1,88 | 30 | 10,2 | 0,61 | 29 |
| Kuiva koore kaal | 0,816 | 0,0931 | 30 | 0,699 | 0,0774 | 29 |
| Ca-sisaldus koores | 0,259 | 0,0035 | 10 | 0,256 | 0,0056 | 10 |

Tabelis on esitatud keskväärtuse, standardhälbe (SD) ja vaatluste arvu (n) väärtused kolmele muutujale: värskete munade kaal; kuiva koore kaal; Ca-sisaldus koores.

Küsimused:

A. Arvuta P-liini muutuja "värskete munade kaal" keskväärtuse standardviga (SEM).

B. Kas esineb olulisi erinevusi muutuja "kuiva koore kaal" P-liini ja C-liini keskväärtuse vahel?

C. Kontrolli, kas on mõistlik eeldada, et P-liinil ja C-liinil on võrdne dispersioon muutujale "värskete munade kaal"?

# AINEREGISTER



| Termin (eesti keel) | Inglise keel | Rootsi keel | lk |
|---|---|---|---|
| alternatiive e sisukas hüpotees | alternative hypothesis | mothypotesen | 65 |
| aritmeetriline keskmine | aritmetic mean | aritmetiskt medelvärde | 7 |
| Bayes'i teoreem | Bayes' theorem | Bayes sats | 33 |
| binoomjaotus | binomial distribution | bionomialfördelning | 41 |
| diskreetne juhuslik suurus | discrete random variable | diskret slumpvariabel | 37, 40 |
| dispersioon | variance | varians | 9 |
| efektiivsus | efficiency | effektivitet | 57 |
| esimesst liiki viga | error of the first kind | fel av första slaget | 66 |
| esindatav | representative | representativ | 3 |
| hajuvusdiagramm | scatter plot | punktdiagram | 18 |
| hindamine | estimation | estimation | 55 |
| hinnang | estimate | estimat | 57 |
| hinnang | estimator | estimator | 57 |
| histogramm, tulpdiagramm | histogram, bar chart | histogram | 16 |
| hüpoteesi kontrollimine | testing the hypothesis | hypotesprövning | 65 |
| intervallskaala | interval scale | intervallskala | 7 |
| jaotusfunktsioon | distribution function | Fördelningsfunktion | 39 |
| juhuslik katse | random experiment | slumpmässig experiment | 29 |
| juhuslik muutuja | random variable | slumpvariabel | 37 |
| juhuslik valik | random sampling | slumpmässig sampling | 3 |
| juhusliku muutuja dispersioon | variance of random variable | varians för en slumpmässig variabel | 40 |
| juhusliku muutuja keskväärtus | mean value of random variable | medelvärde (förväntat värde) för en slumpvariabel | 40 |

| | | | |
|---|---|---|---|
| kahe keskväärtuse võrdlemine | compare two mean values | jämföra två medelvärden | 81 |
| kaks sõltumatut gruppi | two independet groups | två oberoende grupper | 81 |
| karpdiagramm | box plot | box plot, lådogram | 14 |
| keskruudu viga | mean square error | felkvadratsumma | 58 |
| keskväärtus | mean | medelvärde | 7 |
| keskväärtuse standardviga | standard error of the mean | medelvärdets medelfel | 11 |
| kohendatud keskmine | trimmed mean | trimmat medelvärde | 8 |
| korrutamise reegel | multiplication rule | multiplikationssatsen | 32, 33 |
| kvalitatiivne muutuja | qualitative variable | kvalitativ variabel | 6 |
| kvanitatiivne muutuja | quantitative variable | kvantitativ variabel | 6 |
| kvartiil | quartile | kvartil | 11 |
| kvartiilide vahe | inter-quartile range | kvartilavstånd | 12 |
| liitmise reegel | addition rule | additionssatsen | 31 |
| mediaan | median | median | 8 |
| Minitab | Minitab | Minitab | 12 |
| mudel | model | modell | 1 |
| muutuja | variable | Variabel | 7 |
| nihketa | unbiased | väntvärderiktig | 57 |
| nihutamatus | unbiasedness | väntvärderiktighet | 57 |
| nominaalne skaala | nominal scale | nominalskala | 7 |
| normaaljaotus | normal distribution | normalfördelning | 46 |
| nullhüpotees | null hypothesis | nollhypotesen | 65 |
| oluline | significant | signifikans | 66 |
| ootusväärtus | expected value | förväntat värde | 40 |
| ordinaalne skaala | ordinal scale | ordinalskala | 7 |
| otsustused osakaalu kohta | inference on proportions | inferens om proportioner | 75 |
| paarikaupa võrdlemine | pairwise comparsion | Parvisa jämförelser | 81 |
| pidev muutuja | continuous variable | kontinuerlig variabel | 6 |
| Poissoni jaotus | Poisson distribution | Poissonfördelning | 44 |

| | | | |
|---|---|---|---|
| proportsionaalne skaala | ratio scale | kvotskala | 7 |
| protsentiil | percentile | percentil | 11 |
| punkthinnang | point estimation | punktestimation | 57 |
| ringdiagramm | pie chart | tårtdiagram | 24 |
| ristkülikjaotus | rectangular distribution | rektangulärfördelning | 46 |
| risttabel | cross table | korstabell | 21 |
| sagedustabel | frequency table | frekvenstabell | 15, 21 |
| segatud andmed | matched data | matchade data | 81 |
| SEM | SEM | SEM | 11 |
| sisemised aiad | inner fences | övre och undre gränsen | 15 |
| sõltumatu grupp | independent group | obereoende grupper | 81 |
| sõltumatu sündmus | independent events | oberoende händelser | 32 |
| standardhälve | standard deviation | standardavvikelse | 10 |
| standardiseeritud normaaljaotus | standarized normal distribution | standardiserad normalfördelning | 49 |
| statistiline otsustus | statistical inference | statistisk inferens | 4, 55 |
| Student'i t-test | Student's t | Student's t | 60 |
| suhteline sagedus | relative frequency | relativ frekvens | 30 |
| sündmus | event | händelse | 29 |
| täiend | complement | komplement | 30 |
| teist liiki viga | error of the second kind | fel av andra slaget | 66 |
| tinglik tõenäosus | conditional probability | betingad sannolikhet | 33 |
| t-jaotus | t distribution | t-fördelning | 60, 71 |
| tõenäosus | probability | sannolikhet | 30 |
| tõenäosuse usalduspiirid | confidence interval for p | konfidensintervall för p | 75 |
| tõenäosusfunktsioon | probability function | sannolikhetsfunktion | 37, 38 |
| tõenäosusjaotus | probability distribution | sannolikhetsfördelning | 38 |
| tõenäosuslik | probabilistic | probabilistisk | 1 |

| | | | |
|---|---|---|---|
| tõenäosusväärtus | p value | p-värde | 69 |
| tõenäosusväärtus | prob-value | prob-värde, p-värde | 30 |
| tsentraalne piirteoreem | central limit theorem | Centrala gränsvärdessatsen | 50 |
| t-test | t test | t-test | 71 |
| tulemus, tagajärg | outcome | utfall | 29 |
| tulpdiagramm | bar chart | stolpdiagram | 23 |
| tüvi- leht diagramm | stem-and-leaf displays | stem-and-leaf-diagram (stam-och-löv...) | 14 |
| ühepoolne hüpotees | single sided hypothesis | enkelsidig hypotes | 70 |
| üldkogum, populatsioon | population | population | 3 |
| üldkogumi parameetrid | population parameters | populationsparametrar | 41 |
| usalduspiirid | confidence interval | konfidensintervall | 58 |
| vabadusaste | degrees of freedom | frihetsgrader | 60 |
| vahemik | range | | 10 |
| vahemikhinnang | interval estimation | intervallestimation | 58 |
| valim | sample | samspel | 3 |
| valimi statistikud | sample statistics | stickprovsstorheter | 41 |
| valimkeskmine | distribution of $\overline{X}$ | fördelning för $\overline{X}$ | 50 |
| variatsioonikoefitsent | coefficient of variation | variationskoefficient | 11 |
| varieeruvus (muutlikkus) | variation | variation | 1 |
| võimsus | power | power, styrka | 66 |

## TABELID

1. Juhuslike arvude tabel ........ 98
2. Binoomjaotuse jaotusfunktsiooni tabel ........ 99
3. Poissoni jaotuse jaotusfunktsioon ........ 103
4. Normaaljaotuse tabel ........ 104
5. t-jaotusfunktsiooni tabel ........ 105
6. $\chi^2$-jaotuse tabel ........ 106
6. $\chi^2$-jaotuse tabel (järg) ........ 107
7. F-jaotuse tabel (1 – p = 0,90) ........ 108
8. F-jaotuse tabel (1 – p = 0,95) ........ 109
9. F-jaotuse tabel (1 – p = 0,99) ........ 110
10. F-jaotuse tabel (1 – p = 0,999) ........ 111

## 1. Juhuslike arvude tabel

Juhuslikud arvud on valitud tabelisse nii, et igal arvul on ühesugune tõenäosus olla valitud.

| Rida \ Veera | 1 2 3 4 5 | 6 7 8 9 10 | 11 12 13 14 15 | 16 17 18 19 20 | 21 22 23 24 25 | 26 27 28 29 30 |
|---|---|---|---|---|---|---|
| 1 | 1 8 7 8 6 | 7 4 5 9 1 | 3 2 0 1 9 | 6 7 6 6 1 | 3 0 0 7 0 | 8 9 9 1 2 |
| 2 | 9 6 0 2 0 | 7 3 9 8 3 | 5 5 0 0 9 | 0 4 4 0 0 | 0 8 2 1 1 | 4 3 8 8 5 |
| 3 | 4 6 4 3 2 | 6 4 4 5 4 | 6 0 7 4 8 | 1 7 0 0 5 | 4 1 6 3 1 | 4 9 4 2 5 |
| 4 | 7 2 3 2 0 | 6 4 6 2 8 | 6 2 2 7 5 | 4 5 6 0 8 | 0 3 6 4 9 | 6 2 8 9 4 |
| 5 | 1 0 0 8 6 | 2 5 4 5 1 | 1 4 9 8 5 | 5 3 7 0 4 | 8 7 9 1 1 | 0 9 7 3 5 |
| 6 | 1 8 2 7 6 | 3 6 8 0 8 | 5 3 4 0 4 | 8 9 4 3 9 | 1 8 9 8 0 | 6 8 7 7 1 |
| 7 | 6 7 8 8 3 | 5 2 9 6 7 | 4 5 8 7 5 | 9 7 9 4 9 | 7 4 3 8 9 | 0 9 8 7 4 |
| 8 | 9 5 6 5 5 | 1 5 8 5 2 | 1 6 1 2 0 | 1 8 3 5 7 | 3 1 4 9 4 | 5 4 2 9 4 |
| 9 | 8 0 0 1 6 | 2 5 4 2 5 | 9 0 2 9 4 | 7 3 8 7 8 | 7 0 6 6 5 | 5 7 0 1 8 |
| 10 | 4 4 4 8 8 | 6 2 7 5 8 | 2 5 9 0 3 | 7 9 4 1 9 | 4 0 0 0 9 | 9 0 9 7 5 |
| 11 | 8 7 8 2 1 | 6 3 5 7 2 | 3 2 2 3 2 | 3 7 0 8 1 | 1 0 7 7 0 | 2 8 2 5 8 |
| 12 | 0 4 8 5 4 | 5 2 8 2 3 | 9 5 1 8 3 | 6 4 8 7 4 | 1 3 7 6 4 | 1 6 1 1 3 |
| 13 | 1 3 4 8 5 | 0 7 7 6 6 | 0 3 3 1 5 | 0 8 8 7 7 | 3 5 9 7 1 | 3 1 4 0 2 |
| 14 | 3 2 9 5 6 | 2 8 8 7 4 | 3 1 7 5 2 | 4 8 0 6 3 | 7 5 3 0 8 | 7 1 8 2 8 |
| 15 | 7 1 1 7 4 | 5 9 7 3 5 | 7 2 2 3 6 | 5 3 6 6 3 | 1 0 3 9 5 | 9 0 1 7 9 |
| 16 | 2 2 0 4 2 | 4 7 4 0 7 | 8 8 8 1 5 | 5 3 6 2 7 | 9 1 0 0 4 | 0 9 1 1 8 |
| 17 | 4 2 6 5 8 | 2 8 7 0 3 | 8 0 3 1 0 | 7 9 4 6 6 | 2 8 7 3 2 | 5 0 4 1 2 |
| 18 | 4 8 4 1 2 | 4 6 0 7 2 | 6 8 6 7 6 | 3 1 3 9 5 | 7 1 5 9 2 | 2 3 0 2 8 |
| 19 | 7 6 2 0 4 | 0 9 1 8 3 | 8 3 5 7 8 | 9 0 4 1 8 | 5 6 2 7 7 | 8 2 7 2 4 |
| 20 | 3 7 3 8 3 | 2 3 9 0 9 | 6 9 6 3 5 | 0 1 0 5 1 | 1 0 4 1 9 | 7 8 0 2 5 |
| 21 | 0 9 9 0 9 | 2 4 0 0 6 | 5 7 2 2 5 | 1 0 0 7 1 | 9 8 4 5 8 | 0 1 3 4 6 |
| 22 | 7 9 4 7 2 | 2 3 7 4 4 | 0 2 9 7 5 | 2 5 7 6 6 | 6 5 2 5 3 | 2 9 5 1 6 |
| 23 | 9 6 2 9 8 | 3 7 2 2 8 | 5 6 7 2 1 | 5 4 7 6 8 | 7 1 3 5 1 | 3 5 2 8 5 |
| 24 | 9 9 4 4 0 | 7 6 4 4 3 | 1 7 6 6 8 | 1 8 3 1 1 | 1 2 2 9 3 | 9 3 2 4 2 |
| 25 | 5 3 7 2 3 | 6 8 0 1 6 | 6 3 1 1 5 | 2 5 9 6 1 | 9 8 4 2 8 | 2 5 9 4 5 |
| 26 | 5 8 2 0 3 | 8 7 1 8 4 | 7 0 4 6 5 | 7 4 3 4 3 | 6 8 5 9 2 | 5 2 5 2 6 |
| 27 | 0 7 1 2 2 | 9 3 3 6 0 | 7 1 6 2 0 | 5 8 8 0 6 | 9 6 4 8 9 | 6 5 2 2 1 |
| 28 | 6 6 7 8 4 | 2 0 1 5 5 | 8 0 7 5 2 | 3 2 4 9 7 | 0 1 7 8 4 | 9 9 8 6 5 |
| 29 | 1 7 5 3 2 | 4 1 1 2 0 | 9 5 0 8 0 | 0 0 7 5 1 | 7 8 6 6 4 | 9 4 2 9 1 |
| 30 | 4 0 9 9 9 | 0 1 4 2 1 | 0 0 2 5 4 | 4 1 4 2 1 | 2 3 1 1 7 | 4 9 8 8 1 |
| 31 | 7 6 9 9 2 | 8 0 9 9 8 | 2 0 3 3 9 | 4 9 6 4 7 | 4 1 6 0 1 | 3 3 1 2 5 |
| 32 | 9 6 5 3 1 | 3 4 0 2 8 | 6 5 1 0 2 | 0 6 7 7 7 | 2 2 3 5 7 | 7 8 2 8 0 |
| 33 | 3 4 6 8 9 | 8 2 2 7 3 | 4 4 2 3 2 | 6 3 5 3 3 | 4 0 8 6 5 | 3 8 8 9 5 |
| 34 | 3 1 9 0 3 | 9 2 5 1 8 | 9 0 3 9 1 | 6 0 7 9 7 | 8 8 9 6 5 | 7 1 0 9 2 |
| 35 | 5 3 1 0 4 | 0 0 3 1 0 | 0 8 4 6 9 | 7 5 3 9 1 | 5 2 7 1 1 | 2 6 4 2 8 |
| 36 | 8 8 6 1 6 | 1 3 9 3 9 | 0 8 3 1 4 | 5 0 1 1 2 | 8 5 4 3 0 | 0 9 8 4 9 |
| 37 | 6 9 1 0 3 | 9 5 9 7 7 | 6 6 0 1 6 | 4 8 1 2 9 | 5 8 9 3 3 | 3 3 5 6 1 |
| 38 | 9 0 7 5 1 | 8 9 1 4 8 | 8 6 7 0 5 | 0 0 9 1 7 | 9 2 1 6 8 | 0 9 7 2 4 |
| 39 | 0 7 0 1 8 | 7 7 6 9 9 | 6 8 7 8 3 | 1 8 5 0 9 | 3 3 1 2 3 | 2 9 5 6 4 |
| 40 | 5 3 6 3 0 | 9 0 9 0 9 | 2 3 7 6 2 | 6 9 0 8 6 | 0 5 9 0 9 | 3 3 6 2 1 |
| 41 | 0 6 9 2 4 | 3 9 7 5 2 | 4 0 7 1 2 | 9 2 3 3 8 | 1 9 5 5 7 | 8 0 7 0 0 |
| 42 | 6 5 3 3 8 | 7 4 3 3 1 | 3 2 9 6 6 | 9 2 7 0 0 | 2 1 0 4 2 | 0 9 8 8 9 |
| 43 | 8 8 3 6 3 | 1 4 9 4 0 | 0 8 5 9 3 | 5 6 2 9 3 | 4 3 0 4 4 | 8 4 2 9 5 |
| 44 | 0 9 0 5 6 | 4 0 2 8 1 | 5 1 2 7 6 | 5 2 7 5 4 | 9 3 7 8 8 | 1 5 6 7 0 |
| 45 | 1 3 5 0 8 | 5 6 3 1 4 | 2 9 7 5 0 | 1 9 6 1 4 | 9 2 8 3 1 | 2 0 2 6 3 |

## 2. Binoomjaotuse jaotusfunktsiooni tabel

Tabelis on antud n ja p erinevate väärtuste tõenäosused nii, et muutuja väärtus on väiksem või võrdne x-ga.

| n | x | p=0.1 | p=0.2 | p=0.2 | p=0.3 | p=0.4 | p=0.5 | p=0.6 | p=0.7 | p=0.7 | p=0.8 | p=0.9 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 5 | 0 | 0.590 | 0.327 | 0.237 | 0.168 | 0.077 | 0.031 | 0.010 | 0.002 | 0.001 | 0.000 | 0.000 |
| | 1 | 0.918 | 0.737 | 0.632 | 0.528 | 0.337 | 0.187 | 0.087 | 0.030 | 0.015 | 0.006 | 0.000 |
| | 2 | 0.991 | 0.942 | 0.896 | 0.836 | 0.682 | 0.500 | 0.317 | 0.163 | 0.103 | 0.057 | 0.008 |
| | 3 | 0.999 | 0.993 | 0.984 | 0.969 | 0.913 | 0.812 | 0.663 | 0.471 | 0.367 | 0.262 | 0.081 |
| | 4 | 1.000 | 0.999 | 0.999 | 0.997 | 0.989 | 0.968 | 0.922 | 0.831 | 0.762 | 0.672 | 0.409 |
| | 5 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| 6 | 0 | 0.531 | 0.262 | 0.178 | 0.117 | 0.046 | 0.015 | 0.004 | 0.000 | 0.000 | 0.000 | 0.000 |
| | 1 | 0.885 | 0.655 | 0.533 | 0.420 | 0.233 | 0.109 | 0.041 | 0.010 | 0.004 | 0.001 | 0.000 |
| | 2 | 0.984 | 0.901 | 0.830 | 0.744 | 0.544 | 0.343 | 0.179 | 0.070 | 0.037 | 0.017 | 0.001 |
| | 3 | 0.998 | 0.983 | 0.962 | 0.929 | 0.820 | 0.656 | 0.455 | 0.255 | 0.169 | 0.098 | 0.015 |
| | 4 | 0.999 | 0.998 | 0.995 | 0.989 | 0.959 | 0.890 | 0.766 | 0.579 | 0.466 | 0.344 | 0.114 |
| | 5 | 1.000 | 0.999 | 0.999 | 0.999 | 0.995 | 0.984 | 0.953 | 0.882 | 0.822 | 0.737 | 0.468 |
| | 6 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| 7 | 0 | 0.478 | 0.209 | 0.133 | 0.082 | 0.028 | 0.007 | 0.001 | 0.000 | 0.000 | 0.000 | 0.000 |
| | 1 | 0.850 | 0.576 | 0.444 | 0.329 | 0.158 | 0.062 | 0.018 | 0.003 | 0.001 | 0.000 | 0.000 |
| | 2 | 0.974 | 0.852 | 0.756 | 0.647 | 0.419 | 0.226 | 0.096 | 0.028 | 0.012 | 0.004 | 0.000 |
| | 3 | 0.997 | 0.966 | 0.929 | 0.874 | 0.710 | 0.500 | 0.289 | 0.126 | 0.070 | 0.033 | 0.002 |
| | 4 | 0.999 | 0.995 | 0.987 | 0.971 | 0.903 | 0.773 | 0.580 | 0.352 | 0.243 | 0.148 | 0.025 |
| | 5 | 1.000 | 0.999 | 0.998 | 0.996 | 0.981 | 0.937 | 0.841 | 0.670 | 0.555 | 0.423 | 0.149 |
| | 6 | 1.000 | 1.000 | 0.999 | 0.999 | 0.998 | 0.992 | 0.972 | 0.917 | 0.866 | 0.790 | 0.521 |
| | 7 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| 8 | 0 | 0.430 | 0.167 | 0.100 | 0.057 | 0.016 | 0.003 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| | 1 | 0.813 | 0.503 | 0.367 | 0.255 | 0.106 | 0.035 | 0.008 | 0.001 | 0.000 | 0.000 | 0.000 |
| | 2 | 0.961 | 0.796 | 0.678 | 0.551 | 0.315 | 0.144 | 0.049 | 0.011 | 0.004 | 0.001 | 0.000 |
| | 3 | 0.995 | 0.943 | 0.886 | 0.805 | 0.594 | 0.363 | 0.173 | 0.058 | 0.027 | 0.010 | 0.000 |
| | 4 | 0.999 | 0.989 | 0.972 | 0.942 | 0.826 | 0.636 | 0.405 | 0.194 | 0.113 | 0.056 | 0.005 |
| | 5 | 1.000 | 0.998 | 0.995 | 0.988 | 0.950 | 0.855 | 0.684 | 0.448 | 0.321 | 0.203 | 0.038 |
| | 6 | 1.000 | 0.999 | 0.999 | 0.998 | 0.991 | 0.964 | 0.893 | 0.744 | 0.632 | 0.496 | 0.186 |
| | 7 | 1.000 | 1.000 | 1.000 | 0.999 | 0.999 | 0.996 | 0.983 | 0.942 | 0.899 | 0.832 | 0.569 |
| | 8 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| 9 | 0 | 0.387 | 0.134 | 0.075 | 0.040 | 0.010 | 0.002 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| | 1 | 0.774 | 0.436 | 0.300 | 0.196 | 0.070 | 0.019 | 0.003 | 0.000 | 0.000 | 0.000 | 0.000 |
| | 2 | 0.947 | 0.738 | 0.600 | 0.462 | 0.231 | 0.089 | 0.025 | 0.004 | 0.001 | 0.000 | 0.000 |
| | 3 | 0.991 | 0.914 | 0.834 | 0.729 | 0.482 | 0.253 | 0.099 | 0.025 | 0.010 | 0.003 | 0.000 |
| | 4 | 0.999 | 0.980 | 0.951 | 0.901 | 0.733 | 0.500 | 0.266 | 0.098 | 0.048 | 0.019 | 0.000 |
| | 5 | 0.999 | 0.996 | 0.990 | 0.974 | 0.900 | 0.746 | 0.517 | 0.270 | 0.165 | 0.085 | 0.008 |
| | 6 | 1.000 | 0.999 | 0.998 | 0.995 | 0.975 | 0.910 | 0.768 | 0.537 | 0.399 | 0.261 | 0.053 |
| | 7 | 1.000 | 1.000 | 0.999 | 0.999 | 0.996 | 0.980 | 0.929 | 0.804 | 0.699 | 0.563 | 0.225 |
| | 8 | 1.000 | 1.000 | 1.000 | 1.000 | 0.999 | 0.998 | 0.989 | 0.959 | 0.924 | 0.865 | 0.612 |
| | 9 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| 10 | 0 | 0.348 | 0.107 | 0.056 | 0.028 | 0.006 | 0.001 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| | 1 | 0.736 | 0.375 | 0.244 | 0.149 | 0.046 | 0.010 | 0.001 | 0.000 | 0.000 | 0.000 | 0.000 |
| | 2 | 0.929 | 0.677 | 0.525 | 0.382 | 0.167 | 0.054 | 0.012 | 0.001 | 0.000 | 0.000 | 0.000 |
| | 3 | 0.987 | 0.879 | 0.775 | 0.649 | 0.382 | 0.171 | 0.054 | 0.010 | 0.003 | 0.000 | 0.000 |
| | 4 | 0.998 | 0.967 | 0.921 | 0.849 | 0.633 | 0.377 | 0.166 | 0.047 | 0.019 | 0.006 | 0.000 |
| | 5 | 0.999 | 0.993 | 0.980 | 0.952 | 0.833 | 0.623 | 0.366 | 0.150 | 0.078 | 0.032 | 0.001 |
| | 6 | 1.000 | 0.999 | 0.996 | 0.989 | 0.945 | 0.828 | 0.617 | 0.350 | 0.224 | 0.120 | 0.012 |
| | 7 | 1.000 | 0.999 | 0.999 | 0.998 | 0.987 | 0.945 | 0.832 | 0.617 | 0.474 | 0.322 | 0.070 |
| | 8 | 1.000 | 1.000 | 1.000 | 0.999 | 0.998 | 0.989 | 0.953 | 0.850 | 0.756 | 0.624 | 0.263 |
| | 9 | 1.000 | 1.000 | 1.000 | 1.000 | 0.999 | 0.999 | 0.994 | 0.971 | 0.943 | 0.892 | 0.651 |
| | 10 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |

| n | x | p=0.1 | p=0.2 | p=0.25 | p=0.3 | p=0.4 | p=0.5 | p=0.6 | p=0.7 | p=0.75 | p=0.8 | p=0.9 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 11 | 0 | 0.3138 | 0.0859 | 0.0422 | 0.0198 | 0.0036 | 0.0005 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| | 1 | 0.6974 | 0.3221 | 0.1971 | 0.1130 | 0.0302 | 0.0059 | 0.0007 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| | 2 | 0.9104 | 0.6174 | 0.4552 | 0.3127 | 0.1189 | 0.0327 | 0.0059 | 0.0006 | 0.0001 | 0.0000 | 0.0000 |
| | 3 | 0.9815 | 0.8389 | 0.7133 | 0.5696 | 0.2963 | 0.1133 | 0.0293 | 0.0043 | 0.0012 | 0.0002 | 0.0000 |
| | 4 | 0.9972 | 0.9496 | 0.8854 | 0.7897 | 0.5328 | 0.2744 | 0.0994 | 0.0216 | 0.0076 | 0.0020 | 0.0000 |
| | 5 | 0.9997 | 0.9883 | 0.9657 | 0.9218 | 0.7535 | 0.5000 | 0.2465 | 0.0782 | 0.0343 | 0.0117 | 0.0003 |
| | 6 | 1.0000 | 0.9980 | 0.9924 | 0.9784 | 0.9006 | 0.7256 | 0.4672 | 0.2103 | 0.1146 | 0.0504 | 0.0028 |
| | 7 | 1.0000 | 0.9998 | 0.9988 | 0.9957 | 0.9707 | 0.8867 | 0.7037 | 0.4304 | 0.2867 | 0.1611 | 0.0185 |
| | 8 | 1.0000 | 1.0000 | 0.9999 | 0.9994 | 0.9941 | 0.9673 | 0.8811 | 0.6873 | 0.5448 | 0.3826 | 0.0896 |
| | 9 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 0.9993 | 0.9941 | 0.9698 | 0.8870 | 0.8029 | 0.6779 | 0.3026 |
| | 10 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 0.9995 | 0.9964 | 0.9802 | 0.9578 | 0.9141 | 0.6862 |
| | 11 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 12 | 0 | 0.2824 | 0.0687 | 0.0317 | 0.0138 | 0.0022 | 0.0002 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| | 1 | 0.6590 | 0.2749 | 0.1584 | 0.0850 | 0.0196 | 0.0032 | 0.0003 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| | 2 | 0.8891 | 0.5583 | 0.3907 | 0.2528 | 0.0834 | 0.0193 | 0.0028 | 0.0002 | 0.0000 | 0.0000 | 0.0000 |
| | 3 | 0.9744 | 0.7946 | 0.6488 | 0.4925 | 0.2253 | 0.0730 | 0.0153 | 0.0017 | 0.0004 | 0.0001 | 0.0000 |
| | 4 | 0.9957 | 0.9274 | 0.8424 | 0.7237 | 0.4382 | 0.1938 | 0.0573 | 0.0095 | 0.0028 | 0.0006 | 0.0000 |
| | 5 | 0.9995 | 0.9806 | 0.9456 | 0.8822 | 0.6652 | 0.3872 | 0.1582 | 0.0386 | 0.0143 | 0.0039 | 0.0001 |
| | 6 | 0.9999 | 0.9961 | 0.9857 | 0.9614 | 0.8418 | 0.6128 | 0.3348 | 0.1178 | 0.0544 | 0.0194 | 0.0005 |
| | 7 | 1.0000 | 0.9994 | 0.9972 | 0.9905 | 0.9427 | 0.8062 | 0.5618 | 0.2763 | 0.1576 | 0.0726 | 0.0043 |
| | 8 | 1.0000 | 0.9999 | 0.9996 | 0.9983 | 0.9847 | 0.9270 | 0.7747 | 0.5075 | 0.3512 | 0.2054 | 0.0256 |
| | 9 | 1.0000 | 1.0000 | 1.0000 | 0.9998 | 0.9972 | 0.9807 | 0.9166 | 0.7472 | 0.6093 | 0.4417 | 0.1109 |
| | 10 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 0.9997 | 0.9968 | 0.9804 | 0.9150 | 0.8416 | 0.7251 | 0.3410 |
| | 11 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 0.9998 | 0.9978 | 0.9862 | 0.9683 | 0.9313 | 0.7176 |
| | 12 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 13 | 0 | 0.2542 | 0.0550 | 0.0238 | 0.0097 | 0.0013 | 0.0001 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| | 1 | 0.6213 | 0.2336 | 0.1267 | 0.0637 | 0.0126 | 0.0017 | 0.0001 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| | 2 | 0.8661 | 0.5017 | 0.3326 | 0.2025 | 0.0579 | 0.0112 | 0.0013 | 0.0001 | 0.0000 | 0.0000 | 0.0000 |
| | 3 | 0.9658 | 0.7473 | 0.5843 | 0.4206 | 0.1686 | 0.0461 | 0.0078 | 0.0007 | 0.0001 | 0.0000 | 0.0000 |
| | 4 | 0.9935 | 0.9009 | 0.7940 | 0.6543 | 0.3530 | 0.1334 | 0.0321 | 0.0040 | 0.0010 | 0.0002 | 0.0000 |
| | 5 | 0.9991 | 0.9700 | 0.9198 | 0.8346 | 0.5744 | 0.2905 | 0.0977 | 0.0182 | 0.0056 | 0.0012 | 0.0000 |
| | 6 | 0.9999 | 0.9930 | 0.9757 | 0.9376 | 0.7712 | 0.5000 | 0.2288 | 0.0624 | 0.0243 | 0.0070 | 0.0001 |
| | 7 | 1.0000 | 0.9988 | 0.9944 | 0.9818 | 0.9023 | 0.7095 | 0.4256 | 0.1654 | 0.0802 | 0.0300 | 0.0009 |
| | 8 | 1.0000 | 0.9998 | 0.9990 | 0.9960 | 0.9679 | 0.8666 | 0.6470 | 0.3457 | 0.2060 | 0.0991 | 0.0065 |
| | 9 | 1.0000 | 1.0000 | 0.9999 | 0.9993 | 0.9922 | 0.9539 | 0.8314 | 0.5794 | 0.4157 | 0.2527 | 0.0342 |
| | 10 | 1.0000 | 1.0000 | 1.0000 | 0.9999 | 0.9987 | 0.9888 | 0.9421 | 0.7975 | 0.6674 | 0.4983 | 0.1339 |
| | 11 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 0.9999 | 0.9983 | 0.9874 | 0.9363 | 0.8733 | 0.7664 | 0.3787 |
| | 12 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 0.9999 | 0.9987 | 0.9903 | 0.9762 | 0.9450 | 0.7458 |
| | 13 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 14 | 0 | 0.2288 | 0.0440 | 0.0178 | 0.0068 | 0.0008 | 0.0001 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| | 1 | 0.5846 | 0.1979 | 0.1010 | 0.0475 | 0.0081 | 0.0009 | 0.0001 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| | 2 | 0.8416 | 0.4481 | 0.2811 | 0.1608 | 0.0398 | 0.0065 | 0.0006 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| | 3 | 0.9559 | 0.6982 | 0.5213 | 0.3552 | 0.1243 | 0.0287 | 0.0039 | 0.0002 | 0.0000 | 0.0000 | 0.0000 |
| | 4 | 0.9908 | 0.8702 | 0.7415 | 0.5842 | 0.2793 | 0.0898 | 0.0175 | 0.0017 | 0.0003 | 0.0000 | 0.0000 |
| | 5 | 0.9985 | 0.9561 | 0.8883 | 0.7805 | 0.4859 | 0.2120 | 0.0583 | 0.0083 | 0.0022 | 0.0004 | 0.0000 |
| | 6 | 0.9998 | 0.9884 | 0.9617 | 0.9067 | 0.6925 | 0.3953 | 0.1501 | 0.0315 | 0.0103 | 0.0024 | 0.0000 |
| | 7 | 1.0000 | 0.9976 | 0.9897 | 0.9685 | 0.8499 | 0.6047 | 0.3075 | 0.0933 | 0.0383 | 0.0116 | 0.0002 |
| | 8 | 1.0000 | 0.9996 | 0.9978 | 0.9917 | 0.9417 | 0.7880 | 0.5141 | 0.2195 | 0.1117 | 0.0439 | 0.0015 |
| | 9 | 1.0000 | 1.0000 | 0.9997 | 0.9983 | 0.9825 | 0.9102 | 0.7207 | 0.4158 | 0.2585 | 0.1298 | 0.0092 |
| | 10 | 1.0000 | 1.0000 | 1.0000 | 0.9998 | 0.9961 | 0.9713 | 0.8757 | 0.6448 | 0.4787 | 0.3018 | 0.0441 |
| | 11 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 0.9994 | 0.9935 | 0.9602 | 0.8392 | 0.7189 | 0.5519 | 0.1584 |
| | 12 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 0.9999 | 0.9991 | 0.9919 | 0.9525 | 0.8990 | 0.8021 | 0.4154 |
| | 13 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 0.9999 | 0.9992 | 0.9932 | 0.9822 | 0.9560 | 0.7712 |
| | 14 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |

| n | x | p=0.1 | p=0.2 | p=0.25 | p=0.3 | p=0.4 | p=0.5 | p=0.6 | p=0.7 | p=0.75 | p=0.8 | p=0.9 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 15 | 0 | 0.2059 | 0.0352 | 0.0134 | 0.0047 | 0.0005 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| | 1 | 0.5490 | 0.1671 | 0.0802 | 0.0353 | 0.0052 | 0.0005 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| | 2 | 0.8159 | 0.3980 | 0.2361 | 0.1268 | 0.0271 | 0.0037 | 0.0003 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| | 3 | 0.9444 | 0.6482 | 0.4613 | 0.2969 | 0.0905 | 0.0176 | 0.0019 | 0.0001 | 0.0000 | 0.0000 | 0.0000 |
| | 4 | 0.9873 | 0.8358 | 0.6865 | 0.5155 | 0.2173 | 0.0592 | 0.0093 | 0.0007 | 0.0001 | 0.0000 | 0.0000 |
| | 5 | 0.9978 | 0.9389 | 0.8516 | 0.7216 | 0.4032 | 0.1509 | 0.0338 | 0.0037 | 0.0008 | 0.0001 | 0.0000 |
| | 6 | 0.9997 | 0.9819 | 0.9434 | 0.8689 | 0.6098 | 0.3036 | 0.0950 | 0.0152 | 0.0042 | 0.0008 | 0.0000 |
| | 7 | 1.0000 | 0.9958 | 0.9827 | 0.9500 | 0.7869 | 0.5000 | 0.2131 | 0.0500 | 0.0173 | 0.0042 | 0.0000 |
| | 8 | 1.0000 | 0.9992 | 0.9958 | 0.9848 | 0.9050 | 0.6964 | 0.3902 | 0.1311 | 0.0566 | 0.0181 | 0.0003 |
| | 9 | 1.0000 | 0.9999 | 0.9992 | 0.9963 | 0.9662 | 0.8491 | 0.5968 | 0.2784 | 0.1484 | 0.0611 | 0.0022 |
| | 10 | 1.0000 | 1.0000 | 0.9999 | 0.9993 | 0.9907 | 0.9408 | 0.7827 | 0.4845 | 0.3135 | 0.1642 | 0.0127 |
| | 11 | 1.0000 | 1.0000 | 1.0000 | 0.9999 | 0.9981 | 0.9824 | 0.9095 | 0.7031 | 0.5387 | 0.3518 | 0.0556 |
| | 12 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 0.9997 | 0.9963 | 0.9729 | 0.8732 | 0.7639 | 0.6020 | 0.1841 |
| | 13 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 0.9995 | 0.9948 | 0.9647 | 0.9198 | 0.8329 | 0.4510 |
| | 14 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 0.9995 | 0.9953 | 0.9866 | 0.9648 | 0.7941 |
| | 15 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 16 | 0 | 0.1853 | 0.0281 | 0.0100 | 0.0033 | 0.0003 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| | 1 | 0.5147 | 0.1407 | 0.0635 | 0.0261 | 0.0033 | 0.0003 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| | 2 | 0.7892 | 0.3518 | 0.1971 | 0.0994 | 0.0183 | 0.0021 | 0.0001 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| | 3 | 0.9316 | 0.5981 | 0.4050 | 0.2459 | 0.0651 | 0.0106 | 0.0009 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| | 4 | 0.9830 | 0.7982 | 0.6302 | 0.4499 | 0.1666 | 0.0384 | 0.0049 | 0.0003 | 0.0000 | 0.0000 | 0.0000 |
| | 5 | 0.9967 | 0.9183 | 0.8103 | 0.6598 | 0.3288 | 0.1051 | 0.0191 | 0.0016 | 0.0003 | 0.0000 | 0.0000 |
| | 6 | 0.9995 | 0.9733 | 0.9204 | 0.8247 | 0.5272 | 0.2272 | 0.0583 | 0.0071 | 0.0016 | 0.0002 | 0.0000 |
| | 7 | 0.9999 | 0.9930 | 0.9729 | 0.9256 | 0.7161 | 0.4018 | 0.1423 | 0.0257 | 0.0075 | 0.0015 | 0.0000 |
| | 8 | 1.0000 | 0.9985 | 0.9925 | 0.9743 | 0.8577 | 0.5982 | 0.2839 | 0.0744 | 0.0271 | 0.0070 | 0.0001 |
| | 9 | 1.0000 | 0.9998 | 0.9984 | 0.9929 | 0.9417 | 0.7728 | 0.4728 | 0.1753 | 0.0796 | 0.0267 | 0.0005 |
| | 10 | 1.0000 | 1.0000 | 0.9997 | 0.9984 | 0.9809 | 0.8949 | 0.6712 | 0.3402 | 0.1897 | 0.0817 | 0.0033 |
| | 11 | 1.0000 | 1.0000 | 1.0000 | 0.9997 | 0.9951 | 0.9616 | 0.8334 | 0.5501 | 0.3698 | 0.2018 | 0.0170 |
| | 12 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 0.9991 | 0.9894 | 0.9349 | 0.7541 | 0.5950 | 0.4019 | 0.0684 |
| | 13 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 0.9999 | 0.9979 | 0.9817 | 0.9006 | 0.8029 | 0.6482 | 0.2108 |
| | 14 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 0.9997 | 0.9967 | 0.9739 | 0.9365 | 0.8593 | 0.4853 |
| | 15 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 0.9997 | 0.9967 | 0.9900 | 0.9719 | 0.8147 |
| | 16 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 17 | 0 | 0.1668 | 0.0225 | 0.0075 | 0.0023 | 0.0002 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| | 1 | 0.4818 | 0.1182 | 0.0501 | 0.0193 | 0.0021 | 0.0001 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| | 2 | 0.7618 | 0.3096 | 0.1637 | 0.0774 | 0.0123 | 0.0012 | 0.0001 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| | 3 | 0.9174 | 0.5489 | 0.3530 | 0.2019 | 0.0464 | 0.0064 | 0.0005 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| | 4 | 0.9779 | 0.7582 | 0.5739 | 0.3887 | 0.1260 | 0.0245 | 0.0025 | 0.0001 | 0.0000 | 0.0000 | 0.0000 |
| | 5 | 0.9953 | 0.8943 | 0.7653 | 0.5968 | 0.2639 | 0.0717 | 0.0106 | 0.0007 | 0.0001 | 0.0000 | 0.0000 |
| | 6 | 0.9992 | 0.9623 | 0.8929 | 0.7752 | 0.4478 | 0.1662 | 0.0348 | 0.0032 | 0.0006 | 0.0001 | 0.0000 |
| | 7 | 0.9999 | 0.9891 | 0.9598 | 0.8954 | 0.6405 | 0.3145 | 0.0919 | 0.0127 | 0.0031 | 0.0005 | 0.0000 |
| | 8 | 1.0000 | 0.9974 | 0.9876 | 0.9597 | 0.8011 | 0.5000 | 0.1989 | 0.0403 | 0.0124 | 0.0026 | 0.0000 |
| | 9 | 1.0000 | 0.9995 | 0.9969 | 0.9873 | 0.9081 | 0.6855 | 0.3595 | 0.1046 | 0.0402 | 0.0109 | 0.0001 |
| | 10 | 1.0000 | 0.9999 | 0.9994 | 0.9968 | 0.9652 | 0.8338 | 0.5522 | 0.2248 | 0.1071 | 0.0377 | 0.0008 |
| | 11 | 1.0000 | 1.0000 | 0.9999 | 0.9993 | 0.9894 | 0.9283 | 0.7361 | 0.4032 | 0.2347 | 0.1057 | 0.0047 |
| | 12 | 1.0000 | 1.0000 | 1.0000 | 0.9999 | 0.9975 | 0.9755 | 0.8740 | 0.6113 | 0.4261 | 0.2418 | 0.0221 |
| | 13 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 0.9995 | 0.9936 | 0.9536 | 0.7981 | 0.6470 | 0.4511 | 0.0826 |
| | 14 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 0.9999 | 0.9988 | 0.9877 | 0.9226 | 0.8363 | 0.6904 | 0.2382 |
| | 15 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 0.9999 | 0.9979 | 0.9807 | 0.9499 | 0.8818 | 0.5182 |
| | 16 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 0.9998 | 0.9977 | 0.9925 | 0.9775 | 0.8332 |
| | 17 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |

| n | x | p=0.1 | p=0.2 | p=0.25 | p=0.3 | p=0.4 | p=0.5 | p=0.6 | p=0.7 | p=0.75 | p=0.8 | p=0.9 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 18 | 0 | 0.1501 | 0.0180 | 0.0056 | 0.0016 | 0.0001 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| | 1 | 0.4503 | 0.0991 | 0.0395 | 0.0142 | 0.0013 | 0.0001 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| | 2 | 0.7338 | 0.2713 | 0.1353 | 0.0600 | 0.0082 | 0.0007 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| | 3 | 0.9018 | 0.5010 | 0.3057 | 0.1646 | 0.0328 | 0.0038 | 0.0002 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| | 4 | 0.9718 | 0.7164 | 0.5187 | 0.3327 | 0.0942 | 0.0154 | 0.0013 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| | 5 | 0.9936 | 0.8671 | 0.7175 | 0.5344 | 0.2088 | 0.0481 | 0.0058 | 0.0003 | 0.0000 | 0.0000 | 0.0000 |
| | 6 | 0.9988 | 0.9487 | 0.8610 | 0.7217 | 0.3743 | 0.1189 | 0.0203 | 0.0014 | 0.0002 | 0.0000 | 0.0000 |
| | 7 | 0.9998 | 0.9837 | 0.9431 | 0.8593 | 0.5634 | 0.2403 | 0.0576 | 0.0061 | 0.0012 | 0.0002 | 0.0000 |
| | 8 | 1.0000 | 0.9957 | 0.9807 | 0.9404 | 0.7368 | 0.4073 | 0.1347 | 0.0210 | 0.0054 | 0.0009 | 0.0000 |
| | 9 | 1.0000 | 0.9991 | 0.9946 | 0.9790 | 0.8653 | 0.5927 | 0.2632 | 0.0596 | 0.0193 | 0.0043 | 0.0000 |
| | 10 | 1.0000 | 0.9998 | 0.9988 | 0.9939 | 0.9424 | 0.7597 | 0.4366 | 0.1407 | 0.0569 | 0.0163 | 0.0002 |
| | 11 | 1.0000 | 1.0000 | 0.9998 | 0.9986 | 0.9797 | 0.8811 | 0.6257 | 0.2783 | 0.1390 | 0.0513 | 0.0012 |
| | 12 | 1.0000 | 1.0000 | 1.0000 | 0.9997 | 0.9942 | 0.9519 | 0.7912 | 0.4656 | 0.2825 | 0.1329 | 0.0064 |
| | 13 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 0.9987 | 0.9846 | 0.9058 | 0.6673 | 0.4813 | 0.2836 | 0.0282 |
| | 14 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 0.9998 | 0.9962 | 0.9672 | 0.8354 | 0.6943 | 0.4990 | 0.0982 |
| | 15 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 0.9993 | 0.9918 | 0.9400 | 0.8647 | 0.7287 | 0.2662 |
| | 16 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 0.9999 | 0.9987 | 0.9858 | 0.9605 | 0.9009 | 0.5497 |
| | 17 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 0.9999 | 0.9984 | 0.9944 | 0.9820 | 0.8499 |
| | 18 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 19 | 0 | 0.1351 | 0.0144 | 0.0042 | 0.0011 | 0.0001 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| | 1 | 0.4203 | 0.0829 | 0.0310 | 0.0104 | 0.0008 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| | 2 | 0.7054 | 0.2369 | 0.1113 | 0.0462 | 0.0055 | 0.0004 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| | 3 | 0.8850 | 0.4551 | 0.2631 | 0.1332 | 0.0230 | 0.0022 | 0.0001 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| | 4 | 0.9648 | 0.6733 | 0.4654 | 0.2822 | 0.0696 | 0.0096 | 0.0006 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| | 5 | 0.9914 | 0.8369 | 0.6678 | 0.4739 | 0.1629 | 0.0318 | 0.0031 | 0.0001 | 0.0000 | 0.0000 | 0.0000 |
| | 6 | 0.9983 | 0.9324 | 0.8251 | 0.6655 | 0.3081 | 0.0835 | 0.0116 | 0.0006 | 0.0001 | 0.0000 | 0.0000 |
| | 7 | 0.9997 | 0.9767 | 0.9225 | 0.8180 | 0.4878 | 0.1796 | 0.0352 | 0.0028 | 0.0005 | 0.0000 | 0.0000 |
| | 8 | 1.0000 | 0.9933 | 0.9713 | 0.9161 | 0.6675 | 0.3238 | 0.0885 | 0.0105 | 0.0023 | 0.0003 | 0.0000 |
| | 9 | 1.0000 | 0.9984 | 0.9911 | 0.9674 | 0.8139 | 0.5000 | 0.1861 | 0.0326 | 0.0089 | 0.0016 | 0.0000 |
| | 10 | 1.0000 | 0.9997 | 0.9977 | 0.9895 | 0.9115 | 0.6762 | 0.3325 | 0.0839 | 0.0287 | 0.0067 | 0.0000 |
| | 11 | 1.0000 | 1.0000 | 0.9995 | 0.9972 | 0.9648 | 0.8204 | 0.5122 | 0.1820 | 0.0775 | 0.0233 | 0.0003 |
| | 12 | 1.0000 | 1.0000 | 0.9999 | 0.9994 | 0.9884 | 0.9165 | 0.6919 | 0.3345 | 0.1749 | 0.0676 | 0.0017 |
| | 13 | 1.0000 | 1.0000 | 1.0000 | 0.9999 | 0.9969 | 0.9682 | 0.8371 | 0.5261 | 0.3322 | 0.1631 | 0.0086 |
| | 14 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 0.9994 | 0.9904 | 0.9304 | 0.7178 | 0.5346 | 0.3267 | 0.0352 |
| | 15 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 0.9999 | 0.9978 | 0.9770 | 0.8668 | 0.7369 | 0.5449 | 0.1150 |
| | 16 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 0.9996 | 0.9945 | 0.9538 | 0.8887 | 0.7631 | 0.2946 |
| | 17 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 0.9992 | 0.9896 | 0.9690 | 0.9171 | 0.5797 |
| | 18 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 0.9999 | 0.9989 | 0.9958 | 0.9856 | 0.8649 |
| | 19 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 20 | 0 | 0.1216 | 0.0115 | 0.0032 | 0.0008 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| | 1 | 0.3917 | 0.0692 | 0.0243 | 0.0076 | 0.0005 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| | 2 | 0.6769 | 0.2061 | 0.0913 | 0.0355 | 0.0036 | 0.0002 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| | 3 | 0.8670 | 0.4114 | 0.2252 | 0.1071 | 0.0160 | 0.0013 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| | 4 | 0.9568 | 0.6296 | 0.4148 | 0.2375 | 0.0510 | 0.0059 | 0.0003 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| | 5 | 0.9887 | 0.8042 | 0.6172 | 0.4164 | 0.1256 | 0.0207 | 0.0016 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| | 6 | 0.9976 | 0.9133 | 0.7858 | 0.6080 | 0.2500 | 0.0577 | 0.0065 | 0.0003 | 0.0000 | 0.0000 | 0.0000 |
| | 7 | 0.9996 | 0.9679 | 0.8982 | 0.7723 | 0.4159 | 0.1316 | 0.0210 | 0.0013 | 0.0002 | 0.0000 | 0.0000 |
| | 8 | 0.9999 | 0.9900 | 0.9591 | 0.8867 | 0.5956 | 0.2517 | 0.0565 | 0.0051 | 0.0009 | 0.0001 | 0.0000 |
| | 9 | 1.0000 | 0.9974 | 0.9861 | 0.9520 | 0.7553 | 0.4119 | 0.1275 | 0.0171 | 0.0039 | 0.0006 | 0.0000 |
| | 10 | 1.0000 | 0.9994 | 0.9961 | 0.9829 | 0.8725 | 0.5881 | 0.2447 | 0.0480 | 0.0139 | 0.0026 | 0.0000 |
| | 11 | 1.0000 | 0.9999 | 0.9991 | 0.9949 | 0.9435 | 0.7483 | 0.4044 | 0.1133 | 0.0409 | 0.0100 | 0.0001 |
| | 12 | 1.0000 | 1.0000 | 0.9998 | 0.9987 | 0.9790 | 0.8684 | 0.5841 | 0.2277 | 0.1018 | 0.0321 | 0.0004 |
| | 13 | 1.0000 | 1.0000 | 1.0000 | 0.9997 | 0.9935 | 0.9423 | 0.7500 | 0.3920 | 0.2142 | 0.0867 | 0.0024 |
| | 14 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 0.9984 | 0.9793 | 0.8744 | 0.5836 | 0.3828 | 0.1958 | 0.0113 |
| | 15 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 0.9997 | 0.9941 | 0.9490 | 0.7625 | 0.5852 | 0.3704 | 0.0432 |
| | 16 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 0.9987 | 0.9840 | 0.8929 | 0.7748 | 0.5886 | 0.1330 |
| | 17 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 0.9998 | 0.9964 | 0.9645 | 0.9087 | 0.7939 | 0.3231 |
| | 18 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 0.9995 | 0.9924 | 0.9757 | 0.9308 | 0.6083 |
| | 19 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 0.9992 | 0.9968 | 0.9885 | 0.8784 |
| | 20 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |

## 3. Poissoni jaotuse jaotusfunktsioon

Tabelis on antud erinevate väärtuste $\lambda s$ tõenäosused nii, et muutuja väärtus, keskväärtusega $\lambda s$, on väiksem või võrdne x-ga.

| x | λs=0.5 | λs=1 | λs=2 | λs=3 | λs=4 | λs=5 | λs=6 | λs=7 | λs=8 | λs=9 | λs=10 | λs=11 | λs=12 | λs=13 | λs=14 | λs=15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 0.607 | 0.368 | 0.135 | 0.050 | 0.018 | 0.007 | 0.002 | 0.001 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| 1 | 0.910 | 0.736 | 0.406 | 0.199 | 0.092 | 0.040 | 0.017 | 0.007 | 0.003 | 0.001 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| 2 | 0.986 | 0.920 | 0.677 | 0.423 | 0.238 | 0.125 | 0.062 | 0.030 | 0.014 | 0.006 | 0.003 | 0.001 | 0.001 | 0.000 | 0.000 | 0.000 |
| 3 | 0.998 | 0.981 | 0.857 | 0.647 | 0.433 | 0.265 | 0.151 | 0.082 | 0.042 | 0.021 | 0.010 | 0.005 | 0.002 | 0.001 | 0.000 | 0.000 |
| 4 | 1.000 | 0.996 | 0.947 | 0.815 | 0.629 | 0.440 | 0.285 | 0.173 | 0.100 | 0.055 | 0.029 | 0.015 | 0.008 | 0.004 | 0.002 | 0.001 |
| 5 | | 0.999 | 0.983 | 0.916 | 0.785 | 0.616 | 0.446 | 0.301 | 0.191 | 0.116 | 0.067 | 0.038 | 0.020 | 0.011 | 0.006 | 0.003 |
| 6 | | 1.000 | 0.995 | 0.966 | 0.889 | 0.762 | 0.606 | 0.450 | 0.313 | 0.207 | 0.130 | 0.079 | 0.046 | 0.026 | 0.014 | 0.008 |
| 7 | | | 0.999 | 0.988 | 0.949 | 0.867 | 0.744 | 0.599 | 0.453 | 0.324 | 0.220 | 0.143 | 0.090 | 0.054 | 0.032 | 0.018 |
| 8 | | | 1.000 | 0.996 | 0.979 | 0.932 | 0.847 | 0.729 | 0.593 | 0.456 | 0.333 | 0.232 | 0.155 | 0.100 | 0.062 | 0.037 |
| 9 | | | | 0.999 | 0.992 | 0.968 | 0.916 | 0.830 | 0.717 | 0.587 | 0.458 | 0.341 | 0.242 | 0.166 | 0.109 | 0.070 |
| 10 | | | | 1.000 | 0.997 | 0.986 | 0.957 | 0.901 | 0.816 | 0.706 | 0.583 | 0.460 | 0.347 | 0.252 | 0.176 | 0.118 |
| 11 | | | | | 0.999 | 0.995 | 0.980 | 0.947 | 0.888 | 0.803 | 0.697 | 0.579 | 0.462 | 0.353 | 0.260 | 0.185 |
| 12 | | | | | 1.000 | 0.998 | 0.991 | 0.973 | 0.936 | 0.876 | 0.792 | 0.689 | 0.576 | 0.463 | 0.358 | 0.268 |
| 13 | | | | | | 0.999 | 0.996 | 0.987 | 0.966 | 0.926 | 0.864 | 0.781 | 0.682 | 0.573 | 0.464 | 0.363 |
| 14 | | | | | | 1.000 | 0.999 | 0.994 | 0.983 | 0.959 | 0.917 | 0.854 | 0.772 | 0.675 | 0.570 | 0.466 |
| 15 | | | | | | | 0.999 | 0.998 | 0.992 | 0.978 | 0.951 | 0.907 | 0.844 | 0.764 | 0.669 | 0.568 |
| 16 | | | | | | | 1.000 | 0.999 | 0.996 | 0.989 | 0.973 | 0.944 | 0.899 | 0.835 | 0.756 | 0.664 |
| 17 | | | | | | | | 1.000 | 0.998 | 0.995 | 0.986 | 0.968 | 0.937 | 0.890 | 0.827 | 0.749 |
| 18 | | | | | | | | | 0.999 | 0.998 | 0.993 | 0.982 | 0.963 | 0.930 | 0.883 | 0.819 |
| 19 | | | | | | | | | 1.000 | 0.999 | 0.997 | 0.991 | 0.979 | 0.957 | 0.923 | 0.875 |
| 20 | | | | | | | | | | 1.000 | 0.998 | 0.995 | 0.988 | 0.975 | 0.952 | 0.917 |
| 21 | | | | | | | | | | | 0.999 | 0.998 | 0.994 | 0.986 | 0.971 | 0.947 |
| 22 | | | | | | | | | | | 1.000 | 0.999 | 0.997 | 0.992 | 0.983 | 0.967 |
| 23 | | | | | | | | | | | | 1.000 | 0.999 | 0.996 | 0.991 | 0.981 |
| 24 | | | | | | | | | | | | | 0.999 | 0.998 | 0.995 | 0.989 |
| 25 | | | | | | | | | | | | | 1.000 | 0.999 | 0.997 | 0.994 |
| 26 | | | | | | | | | | | | | | 1.000 | 0.999 | 0.997 |
| 27 | | | | | | | | | | | | | | | 0.999 | 0.998 |
| 28 | | | | | | | | | | | | | | | 1.000 | 0.999 |
| 29 | | | | | | | | | | | | | | | | 1.000 |
| 30 | | | | | | | | | | | | | | | | |

# 4. Normaaljaotuse tabel

| z | Viimane kümnendkoht 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|---|
| 0.0 | 0.5000 | 0.5040 | 0.5080 | 0.5120 | 0.5160 | 0.5199 | 0.5239 | 0.5279 | 0.5319 | 0.5359 |
| 0.1 | 0.5398 | 0.5438 | 0.5478 | 0.5517 | 0.5557 | 0.5596 | 0.5636 | 0.5675 | 0.5714 | 0.5753 |
| 0.2 | 0.5793 | 0.5832 | 0.5871 | 0.5910 | 0.5948 | 0.5987 | 0.6026 | 0.6064 | 0.6103 | 0.6141 |
| 0.3 | 0.6179 | 0.6217 | 0.6255 | 0.6293 | 0.6331 | 0.6368 | 0.6406 | 0.6443 | 0.6480 | 0.6517 |
| 0.4 | 0.6554 | 0.6591 | 0.6628 | 0.6664 | 0.6700 | 0.6736 | 0.6772 | 0.6808 | 0.6844 | 0.6879 |
| 0.5 | 0.6915 | 0.6950 | 0.6985 | 0.7019 | 0.7054 | 0.7088 | 0.7123 | 0.7157 | 0.7190 | 0.7224 |
| 0.6 | 0.7257 | 0.7291 | 0.7324 | 0.7357 | 0.7389 | 0.7422 | 0.7454 | 0.7486 | 0.7517 | 0.7549 |
| 0.7 | 0.7580 | 0.7611 | 0.7642 | 0.7673 | 0.7704 | 0.7734 | 0.7764 | 0.7794 | 0.7823 | 0.7852 |
| 0.8 | 0.7881 | 0.7910 | 0.7939 | 0.7967 | 0.7995 | 0.8023 | 0.8051 | 0.8078 | 0.8106 | 0.8133 |
| 0.9 | 0.8159 | 0.8186 | 0.8212 | 0.8238 | 0.8264 | 0.8289 | 0.8315 | 0.8340 | 0.8365 | 0.8389 |
| 1.0 | 0.8413 | 0.8438 | 0.8461 | 0.8485 | 0.8508 | 0.8531 | 0.8554 | 0.8577 | 0.8599 | 0.8621 |
| 1.1 | 0.8643 | 0.8665 | 0.8686 | 0.8708 | 0.8729 | 0.8749 | 0.8770 | 0.8790 | 0.8810 | 0.8830 |
| 1.2 | 0.8849 | 0.8869 | 0.8888 | 0.8907 | 0.8925 | 0.8944 | 0.8962 | 0.8980 | 0.8997 | 0.9015 |
| 1.3 | 0.9032 | 0.9049 | 0.9066 | 0.9082 | 0.9099 | 0.9115 | 0.9131 | 0.9147 | 0.9162 | 0.9177 |
| 1.4 | 0.9192 | 0.9207 | 0.9222 | 0.9236 | 0.9251 | 0.9265 | 0.9279 | 0.9292 | 0.9306 | 0.9319 |
| 1.5 | 0.9332 | 0.9345 | 0.9357 | 0.9370 | 0.9382 | 0.9394 | 0.9406 | 0.9418 | 0.9429 | 0.9441 |
| 1.6 | 0.9452 | 0.9463 | 0.9474 | 0.9484 | 0.9495 | 0.9505 | 0.9515 | 0.9525 | 0.9535 | 0.9545 |
| 1.7 | 0.9554 | 0.9564 | 0.9573 | 0.9582 | 0.9591 | 0.9599 | 0.9608 | 0.9616 | 0.9625 | 0.9633 |
| 1.8 | 0.9641 | 0.9649 | 0.9656 | 0.9664 | 0.9671 | 0.9678 | 0.9686 | 0.9693 | 0.9699 | 0.9706 |
| 1.9 | 0.9713 | 0.9719 | 0.9726 | 0.9732 | 0.9738 | 0.9744 | 0.9750 | 0.9756 | 0.9761 | 0.9767 |
| 2.0 | 0.9772 | 0.9778 | 0.9783 | 0.9788 | 0.9793 | 0.9798 | 0.9803 | 0.9808 | 0.9812 | 0.9817 |
| 2.1 | 0.9821 | 0.9826 | 0.9830 | 0.9834 | 0.9838 | 0.9842 | 0.9846 | 0.9850 | 0.9854 | 0.9857 |
| 2.2 | 0.9861 | 0.9864 | 0.9868 | 0.9871 | 0.9875 | 0.9878 | 0.9881 | 0.9884 | 0.9887 | 0.9890 |
| 2.3 | 0.9893 | 0.9896 | 0.9898 | 0.9901 | 0.9904 | 0.9906 | 0.9909 | 0.9911 | 0.9913 | 0.9916 |
| 2.4 | 0.9918 | 0.9920 | 0.9922 | 0.9925 | 0.9927 | 0.9929 | 0.9931 | 0.9932 | 0.9934 | 0.9936 |
| 2.5 | 0.9938 | 0.9940 | 0.9941 | 0.9943 | 0.9945 | 0.9946 | 0.9948 | 0.9949 | 0.9951 | 0.9952 |
| 2.6 | 0.9953 | 0.9955 | 0.9956 | 0.9957 | 0.9959 | 0.9960 | 0.9961 | 0.9962 | 0.9963 | 0.9964 |
| 2.7 | 0.9965 | 0.9966 | 0.9967 | 0.9968 | 0.9969 | 0.9970 | 0.9971 | 0.9972 | 0.9973 | 0.9974 |
| 2.8 | 0.9974 | 0.9975 | 0.9976 | 0.9977 | 0.9977 | 0.9978 | 0.9979 | 0.9979 | 0.9980 | 0.9981 |
| 2.9 | 0.9981 | 0.9982 | 0.9982 | 0.9983 | 0.9984 | 0.9984 | 0.9985 | 0.9985 | 0.9986 | 0.9986 |
| 3.0 | 0.9987 | 0.9987 | 0.9987 | 0.9988 | 0.9988 | 0.9989 | 0.9989 | 0.9989 | 0.9990 | 0.9990 |
| 3.1 | 0.9990 | 0.9991 | 0.9991 | 0.9991 | 0.9992 | 0.9992 | 0.9992 | 0.9992 | 0.9993 | 0.9993 |
| 3.2 | 0.9993 | 0.9993 | 0.9994 | 0.9994 | 0.9994 | 0.9994 | 0.9994 | 0.9995 | 0.9995 | 0.9995 |
| 3.3 | 0.9995 | 0.9995 | 0.9995 | 0.9996 | 0.9996 | 0.9996 | 0.9996 | 0.9996 | 0.9996 | 0.9997 |
| 3.4 | 0.9997 | 0.9997 | 0.9997 | 0.9997 | 0.9997 | 0.9997 | 0.9997 | 0.9997 | 0.9997 | 0.9998 |
| 3.5 | 0.9998 | 0.9998 | 0.9998 | 0.9998 | 0.9998 | 0.9998 | 0.9998 | 0.9998 | 0.9998 | 0.9998 |
| 3.6 | 0.9998 | 0.9998 | 0.9999 | 0.9999 | 0.9999 | 0.9999 | 0.9999 | 0.9999 | 0.9999 | 0.9999 |
| 3.7 | 0.9999 | 0.9999 | 0.9999 | 0.9999 | 0.9999 | 0.9999 | 0.9999 | 0.9999 | 0.9999 | 0.9999 |
| 3.8 | 0.9999 | 0.9999 | 0.9999 | 0.9999 | 0.9999 | 0.9999 | 0.9999 | 0.9999 | 0.9999 | 0.9999 |
| 3.9 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 4.0 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |

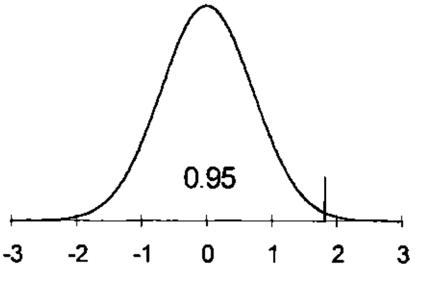
Tabel annab normaaljaotuse jaotusfunktsiooni $F(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{\frac{t^2}{2}} dt$ väärtused vahemikus z = 0,00(0,01)4,09. Negatiivsete väärtuste korral kasuta teisendust F(z) =1 – F(-z).

## 5. t-jaotusfunktsiooni tabel

| | F(t) | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| d.f. | 0.75 | 0.90 | 0.95 | 0.975 | 0.99 | 0.995 | 0.9975 | 0.999 | 0.9995 |
| 1 | 1.000 | 3.078 | 6.314 | 12.706 | 31.821 | 63.657 | 127.321 | 318.309 | 636.619 |
| 2 | 0.816 | 1.886 | 2.920 | 4.303 | 6.965 | 9.925 | 14.089 | 22.327 | 31.599 |
| 3 | 0.765 | 1.638 | 2.353 | 3.182 | 4.541 | 5.841 | 7.453 | 10.215 | 12.924 |
| 4 | 0.741 | 1.533 | 2.132 | 2.776 | 3.747 | 4.604 | 5.598 | 7.173 | 8.610 |
| 5 | 0.727 | 1.476 | 2.015 | 2.571 | 3.365 | 4.032 | 4.773 | 5.893 | 6.869 |
| 6 | 0.718 | 1.440 | 1.943 | 2.447 | 3.143 | 3.707 | 4.317 | 5.208 | 5.959 |
| 7 | 0.711 | 1.415 | 1.895 | 2.365 | 2.998 | 3.499 | 4.029 | 4.785 | 5.408 |
| 8 | 0.706 | 1.397 | 1.860 | 2.306 | 2.896 | 3.355 | 3.833 | 4.501 | 5.041 |
| 9 | 0.703 | 1.383 | 1.833 | 2.262 | 2.821 | 3.250 | 3.690 | 4.297 | 4.781 |
| 10 | 0.700 | 1.372 | 1.812 | 2.228 | 2.764 | 3.169 | 3.581 | 4.144 | 4.587 |
| 11 | 0.697 | 1.363 | 1.796 | 2.201 | 2.718 | 3.106 | 3.497 | 4.025 | 4.437 |
| 12 | 0.695 | 1.356 | 1.782 | 2.179 | 2.681 | 3.055 | 3.428 | 3.930 | 4.318 |
| 13 | 0.694 | 1.350 | 1.771 | 2.160 | 2.650 | 3.012 | 3.372 | 3.852 | 4.221 |
| 14 | 0.692 | 1.345 | 1.761 | 2.145 | 2.624 | 2.977 | 3.326 | 3.787 | 4.140 |
| 15 | 0.691 | 1.341 | 1.753 | 2.131 | 2.602 | 2.947 | 3.286 | 3.733 | 4.073 |
| 16 | 0.690 | 1.337 | 1.746 | 2.120 | 2.583 | 2.921 | 3.252 | 3.686 | 4.015 |
| 17 | 0.689 | 1.333 | 1.740 | 2.110 | 2.567 | 2.898 | 3.222 | 3.646 | 3.965 |
| 18 | 0.688 | 1.330 | 1.734 | 2.101 | 2.552 | 2.878 | 3.197 | 3.610 | 3.922 |
| 19 | 0.688 | 1.328 | 1.729 | 2.093 | 2.539 | 2.861 | 3.174 | 3.579 | 3.883 |
| 20 | 0.687 | 1.325 | 1.725 | 2.086 | 2.528 | 2.845 | 3.153 | 3.552 | 3.850 |
| 21 | 0.686 | 1.323 | 1.721 | 2.080 | 2.518 | 2.831 | 3.135 | 3.527 | 3.819 |
| 22 | 0.686 | 1.321 | 1.717 | 2.074 | 2.508 | 2.819 | 3.119 | 3.505 | 3.792 |
| 23 | 0.685 | 1.319 | 1.714 | 2.069 | 2.500 | 2.807 | 3.104 | 3.485 | 3.768 |
| 24 | 0.685 | 1.318 | 1.711 | 2.064 | 2.492 | 2.797 | 3.091 | 3.467 | 3.745 |
| 25 | 0.684 | 1.316 | 1.708 | 2.060 | 2.485 | 2.787 | 3.078 | 3.450 | 3.725 |
| 26 | 0.684 | 1.315 | 1.706 | 2.056 | 2.479 | 2.779 | 3.067 | 3.435 | 3.707 |
| 27 | 0.684 | 1.314 | 1.703 | 2.052 | 2.473 | 2.771 | 3.057 | 3.421 | 3.690 |
| 28 | 0.683 | 1.313 | 1.701 | 2.048 | 2.467 | 2.763 | 3.047 | 3.408 | 3.674 |
| 29 | 0.683 | 1.311 | 1.699 | 2.045 | 2.462 | 2.756 | 3.038 | 3.396 | 3.659 |
| 30 | 0.683 | 1.310 | 1.697 | 2.042 | 2.457 | 2.750 | 3.030 | 3.385 | 3.646 |
| 40 | 0.681 | 1.303 | 1.684 | 2.021 | 2.423 | 2.704 | 2.971 | 3.307 | 3.551 |
| 50 | 0.679 | 1.299 | 1.676 | 2.009 | 2.403 | 2.678 | 2.937 | 3.261 | 3.496 |
| 60 | 0.679 | 1.296 | 1.671 | 2.000 | 2.390 | 2.660 | 2.915 | 3.232 | 3.460 |
| 80 | 0.678 | 1.292 | 1.664 | 1.990 | 2.374 | 2.639 | 2.887 | 3.195 | 3.416 |
| 100 | 0.677 | 1.290 | 1.660 | 1.984 | 2.364 | 2.626 | 2.871 | 3.174 | 3.390 |
| ∞ | 0.674 | 1.282 | 1.645 | 1.960 | 2.326 | 2.576 | 2.807 | 3.092 | 3.291 |

**Näide:** $t$-jaotuse 10 vabadusastmega (d.f.) suuruse jaoks $F(1.812)=P(t\leq 1.812)=0.95$.

Jaotus on sümmeetriline, st $t$-jaotuse 10 vabadusastmega (d.f.) suuruse jaoks kehtib $F(-1.812)=P(t\leq -1.812)=0.05$.



## 6. $\chi^2$-jaotuse tabel

| | 1-p | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| d.f. | 0.0005 | 0.001 | 0.0025 | 0.005 | 0.01 | 0.025 | 0.05 | 0.1 | 0.5 |
| 1 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.001 | 0.004 | 0.016 | 0.455 |
| 2 | 0.001 | 0.002 | 0.005 | 0.010 | 0.020 | 0.051 | 0.103 | 0.211 | 1.386 |
| 3 | 0.015 | 0.024 | 0.045 | 0.072 | 0.115 | 0.216 | 0.352 | 0.584 | 2.366 |
| 4 | 0.064 | 0.091 | 0.145 | 0.207 | 0.297 | 0.484 | 0.711 | 1.064 | 3.357 |
| 5 | 0.158 | 0.210 | 0.307 | 0.412 | 0.554 | 0.831 | 1.145 | 1.610 | 4.351 |
| 6 | 0.299 | 0.381 | 0.527 | 0.676 | 0.872 | 1.237 | 1.635 | 2.204 | 5.348 |
| 7 | 0.485 | 0.598 | 0.794 | 0.989 | 1.239 | 1.690 | 2.167 | 2.833 | 6.346 |
| 8 | 0.710 | 0.857 | 1.104 | 1.344 | 1.646 | 2.180 | 2.733 | 3.490 | 7.344 |
| 9 | 0.972 | 1.152 | 1.450 | 1.735 | 2.088 | 2.700 | 3.325 | 4.168 | 8.343 |
| 10 | 1.265 | 1.479 | 1.827 | 2.156 | 2.558 | 3.247 | 3.940 | 4.865 | 9.342 |
| 11 | 1.587 | 1.834 | 2.232 | 2.603 | 3.053 | 3.816 | 4.575 | 5.578 | 10.341 |
| 12 | 1.934 | 2.214 | 2.661 | 3.074 | 3.571 | 4.404 | 5.226 | 6.304 | 11.340 |
| 13 | 2.305 | 2.617 | 3.112 | 3.565 | 4.107 | 5.009 | 5.892 | 7.042 | 12.340 |
| 14 | 2.697 | 3.041 | 3.582 | 4.075 | 4.660 | 5.629 | 6.571 | 7.790 | 13.339 |
| 15 | 3.108 | 3.483 | 4.070 | 4.601 | 5.229 | 6.262 | 7.261 | 8.547 | 14.339 |
| 16 | 3.536 | 3.942 | 4.573 | 5.142 | 5.812 | 6.908 | 7.962 | 9.312 | 15.338 |
| 17 | 3.980 | 4.416 | 5.092 | 5.697 | 6.408 | 7.564 | 8.672 | 10.085 | 16.338 |
| 18 | 4.439 | 4.905 | 5.623 | 6.265 | 7.015 | 8.231 | 9.390 | 10.865 | 17.338 |
| 19 | 4.912 | 5.407 | 6.167 | 6.844 | 7.633 | 8.907 | 10.117 | 11.651 | 18.338 |
| 20 | 5.398 | 5.921 | 6.723 | 7.434 | 8.260 | 9.591 | 10.851 | 12.443 | 19.337 |
| 21 | 5.896 | 6.447 | 7.289 | 8.034 | 8.897 | 10.283 | 11.591 | 13.240 | 20.337 |
| 22 | 6.404 | 6.983 | 7.865 | 8.643 | 9.542 | 10.982 | 12.338 | 14.041 | 21.337 |
| 23 | 6.924 | 7.529 | 8.450 | 9.260 | 10.196 | 11.689 | 13.091 | 14.848 | 22.337 |
| 24 | 7.453 | 8.085 | 9.044 | 9.886 | 10.856 | 12.401 | 13.848 | 15.659 | 23.337 |
| 25 | 7.991 | 8.649 | 9.646 | 10.520 | 11.524 | 13.120 | 14.611 | 16.473 | 24.337 |
| 26 | 8.538 | 9.222 | 10.256 | 11.160 | 12.198 | 13.844 | 15.379 | 17.292 | 25.336 |
| 27 | 9.093 | 9.803 | 10.873 | 11.808 | 12.879 | 14.573 | 16.151 | 18.114 | 26.336 |
| 28 | 9.656 | 10.391 | 11.497 | 12.461 | 13.565 | 15.308 | 16.928 | 18.939 | 27.336 |
| 29 | 10.227 | 10.986 | 12.128 | 13.121 | 14.256 | 16.047 | 17.708 | 19.768 | 28.336 |
| 30 | 10.804 | 11.588 | 12.765 | 13.787 | 14.953 | 16.791 | 18.493 | 20.599 | 29.336 |
| 40 | 16.906 | 17.916 | 19.417 | 20.707 | 22.164 | 24.433 | 26.509 | 29.051 | 39.335 |
| 50 | 23.461 | 24.674 | 26.464 | 27.991 | 29.707 | 32.357 | 34.764 | 37.689 | 49.335 |
| 60 | 30.340 | 31.738 | 33.791 | 35.534 | 37.485 | 40.482 | 43.188 | 46.459 | 59.335 |
| 80 | 44.791 | 46.520 | 49.043 | 51.172 | 53.540 | 57.153 | 60.391 | 64.278 | 79.334 |
| 100 | 59.896 | 61.918 | 64.857 | 67.328 | 70.065 | 74.222 | 77.929 | 82.358 | 99.334 |

# 6. $\chi^2$-jaotuse tabel (järg)

| | 1-p | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| d.f. | 0.75 | 0.9 | 0.95 | 0.975 | 0.99 | 0.995 | 0.9975 | 0.999 | 0.9995 |
| 1 | 1.323 | 2.706 | 3.841 | 5.024 | 6.635 | 7.879 | 9.141 | 10.828 | 12.116 |
| 2 | 2.773 | 4.605 | 5.991 | 7.378 | 9.210 | 10.597 | 11.983 | 13.816 | 15.202 |
| 3 | 4.108 | 6.251 | 7.815 | 9.348 | 11.345 | 12.838 | 14.320 | 16.266 | 17.730 |
| 4 | 5.385 | 7.779 | 9.488 | 11.143 | 13.277 | 14.860 | 16.424 | 18.467 | 19.997 |
| 5 | 6.626 | 9.236 | 11.070 | 12.833 | 15.086 | 16.750 | 18.386 | 20.515 | 22.105 |
| 6 | 7.841 | 10.645 | 12.592 | 14.449 | 16.812 | 18.548 | 20.249 | 22.458 | 24.103 |
| 7 | 9.037 | 12.017 | 14.067 | 16.013 | 18.475 | 20.278 | 22.040 | 24.322 | 26.018 |
| 8 | 10.219 | 13.362 | 15.507 | 17.535 | 20.090 | 21.955 | 23.774 | 26.124 | 27.868 |
| 9 | 11.389 | 14.684 | 16.919 | 19.023 | 21.666 | 23.589 | 25.462 | 27.877 | 29.666 |
| 10 | 12.549 | 15.987 | 18.307 | 20.483 | 23.209 | 25.188 | 27.112 | 29.588 | 31.420 |
| 11 | 13.701 | 17.275 | 19.675 | 21.920 | 24.725 | 26.757 | 28.729 | 31.264 | 33.137 |
| 12 | 14.845 | 18.549 | 21.026 | 23.337 | 26.217 | 28.300 | 30.318 | 32.909 | 34.821 |
| 13 | 15.984 | 19.812 | 22.362 | 24.736 | 27.688 | 29.819 | 31.883 | 34.528 | 36.478 |
| 14 | 17.117 | 21.064 | 23.685 | 26.119 | 29.141 | 31.319 | 33.426 | 36.123 | 38.109 |
| 15 | 18.245 | 22.307 | 24.996 | 27.488 | 30.578 | 32.801 | 34.950 | 37.697 | 39.719 |
| 16 | 19.369 | 23.542 | 26.296 | 28.845 | 32.000 | 34.267 | 36.456 | 39.252 | 41.308 |
| 17 | 20.489 | 24.769 | 27.587 | 30.191 | 33.409 | 35.718 | 37.946 | 40.790 | 42.879 |
| 18 | 21.605 | 25.989 | 28.869 | 31.526 | 34.805 | 37.156 | 39.422 | 42.312 | 44.434 |
| 19 | 22.718 | 27.204 | 30.144 | 32.852 | 36.191 | 38.582 | 40.885 | 43.820 | 45.973 |
| 20 | 23.828 | 28.412 | 31.410 | 34.170 | 37.566 | 39.997 | 42.336 | 45.315 | 47.498 |
| 21 | 24.935 | 29.615 | 32.671 | 35.479 | 38.932 | 41.401 | 43.775 | 46.797 | 49.011 |
| 22 | 26.039 | 30.813 | 33.924 | 36.781 | 40.289 | 42.796 | 45.204 | 48.268 | 50.511 |
| 23 | 27.141 | 32.007 | 35.172 | 38.076 | 41.638 | 44.181 | 46.623 | 49.728 | 52.000 |
| 24 | 28.241 | 33.196 | 36.415 | 39.364 | 42.980 | 45.559 | 48.034 | 51.179 | 53.479 |
| 25 | 29.339 | 34.382 | 37.652 | 40.646 | 44.314 | 46.928 | 49.435 | 52.620 | 54.947 |
| 26 | 30.435 | 35.563 | 38.885 | 41.923 | 45.642 | 48.290 | 50.829 | 54.052 | 56.407 |
| 27 | 31.528 | 36.741 | 40.113 | 43.195 | 46.963 | 49.645 | 52.215 | 55.476 | 57.858 |
| 28 | 32.620 | 37.916 | 41.337 | 44.461 | 48.278 | 50.993 | 53.594 | 56.892 | 59.300 |
| 29 | 33.711 | 39.087 | 42.557 | 45.722 | 49.588 | 52.336 | 54.967 | 58.301 | 60.735 |
| 30 | 34.800 | 40.256 | 43.773 | 46.979 | 50.892 | 53.672 | 56.332 | 59.703 | 62.162 |
| 40 | 45.616 | 51.805 | 55.758 | 59.342 | 63.691 | 66.766 | 69.699 | 73.402 | 76.095 |
| 50 | 56.334 | 63.167 | 67.505 | 71.420 | 76.154 | 79.490 | 82.664 | 86.661 | 89.561 |
| 60 | 66.981 | 74.397 | 79.082 | 83.298 | 88.379 | 91.952 | 95.344 | 99.607 | 102.695 |
| 80 | 88.130 | 96.578 | 101.879 | 106.629 | 112.329 | 116.321 | 120.102 | 124.839 | 128.261 |
| 100 | 109.141 | 118.498 | 124.342 | 129.561 | 135.807 | 140.169 | 144.293 | 149.449 | 153.167 |

## 7 F-jaotuse tabel (1 – p = 0,90)

| $\nu_2$ \ $\nu_1$ | 1 | 2 | 3 | 4 | 5 | 6 | 8 | 10 | 12 | 16 | 20 | 30 | 40 | 60 | 120 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 39.863 | 49.500 | 53.593 | 55.833 | 57.240 | 58.204 | 59.439 | 60.195 | 60.705 | 61.350 | 61.740 | 62.265 | 62.529 | 62.794 | 63.061 |
| 2 | 8.526 | 9.000 | 9.162 | 9.243 | 9.293 | 9.326 | 9.367 | 9.392 | 9.408 | 9.429 | 9.441 | 9.458 | 9.466 | 9.475 | 9.483 |
| 3 | 5.538 | 5.462 | 5.391 | 5.343 | 5.309 | 5.285 | 5.252 | 5.230 | 5.216 | 5.196 | 5.184 | 5.168 | 5.160 | 5.151 | 5.143 |
| 4 | 4.545 | 4.325 | 4.191 | 4.107 | 4.051 | 4.010 | 3.955 | 3.920 | 3.896 | 3.864 | 3.844 | 3.817 | 3.804 | 3.790 | 3.775 |
| 5 | 4.060 | 3.780 | 3.619 | 3.520 | 3.453 | 3.405 | 3.339 | 3.297 | 3.268 | 3.230 | 3.207 | 3.174 | 3.157 | 3.140 | 3.123 |
| 6 | 3.776 | 3.463 | 3.289 | 3.181 | 3.108 | 3.055 | 2.983 | 2.937 | 2.905 | 2.863 | 2.836 | 2.800 | 2.781 | 2.762 | 2.742 |
| 7 | 3.589 | 3.257 | 3.074 | 2.961 | 2.883 | 2.827 | 2.752 | 2.703 | 2.668 | 2.623 | 2.595 | 2.555 | 2.535 | 2.514 | 2.493 |
| 8 | 3.458 | 3.113 | 2.924 | 2.806 | 2.726 | 2.668 | 2.589 | 2.538 | 2.502 | 2.455 | 2.425 | 2.383 | 2.361 | 2.339 | 2.316 |
| 9 | 3.360 | 3.006 | 2.813 | 2.693 | 2.611 | 2.551 | 2.469 | 2.416 | 2.379 | 2.329 | 2.298 | 2.255 | 2.232 | 2.208 | 2.184 |
| 10 | 3.285 | 2.924 | 2.728 | 2.605 | 2.522 | 2.461 | 2.377 | 2.323 | 2.284 | 2.233 | 2.201 | 2.155 | 2.132 | 2.107 | 2.082 |
| 12 | 3.177 | 2.807 | 2.606 | 2.480 | 2.394 | 2.331 | 2.245 | 2.188 | 2.147 | 2.094 | 2.060 | 2.011 | 1.986 | 1.960 | 1.932 |
| 14 | 3.102 | 2.726 | 2.522 | 2.395 | 2.307 | 2.243 | 2.154 | 2.095 | 2.054 | 1.998 | 1.962 | 1.912 | 1.885 | 1.857 | 1.828 |
| 16 | 3.048 | 2.668 | 2.462 | 2.333 | 2.244 | 2.178 | 2.088 | 2.028 | 1.985 | 1.928 | 1.891 | 1.839 | 1.811 | 1.782 | 1.751 |
| 18 | 3.007 | 2.624 | 2.416 | 2.286 | 2.196 | 2.130 | 2.038 | 1.977 | 1.933 | 1.875 | 1.837 | 1.783 | 1.754 | 1.723 | 1.691 |
| 20 | 2.975 | 2.589 | 2.380 | 2.249 | 2.158 | 2.091 | 1.999 | 1.937 | 1.892 | 1.833 | 1.794 | 1.738 | 1.708 | 1.677 | 1.643 |
| 24 | 2.927 | 2.538 | 2.327 | 2.195 | 2.103 | 2.035 | 1.941 | 1.877 | 1.832 | 1.770 | 1.730 | 1.672 | 1.641 | 1.607 | 1.571 |
| 28 | 2.894 | 2.503 | 2.291 | 2.157 | 2.064 | 1.996 | 1.900 | 1.836 | 1.790 | 1.726 | 1.685 | 1.625 | 1.592 | 1.558 | 1.520 |
| 32 | 2.869 | 2.477 | 2.263 | 2.129 | 2.036 | 1.967 | 1.870 | 1.805 | 1.758 | 1.694 | 1.652 | 1.590 | 1.556 | 1.520 | 1.481 |
| 36 | 2.850 | 2.456 | 2.243 | 2.108 | 2.014 | 1.945 | 1.847 | 1.781 | 1.734 | 1.669 | 1.626 | 1.563 | 1.528 | 1.491 | 1.450 |
| 40 | 2.835 | 2.440 | 2.226 | 2.091 | 1.997 | 1.927 | 1.829 | 1.763 | 1.715 | 1.649 | 1.605 | 1.541 | 1.506 | 1.467 | 1.425 |
| 50 | 2.809 | 2.412 | 2.197 | 2.061 | 1.966 | 1.895 | 1.796 | 1.729 | 1.680 | 1.613 | 1.568 | 1.502 | 1.465 | 1.424 | 1.379 |
| 60 | 2.791 | 2.393 | 2.177 | 2.041 | 1.946 | 1.875 | 1.775 | 1.707 | 1.657 | 1.589 | 1.543 | 1.476 | 1.437 | 1.395 | 1.348 |
| 70 | 2.779 | 2.380 | 2.164 | 2.027 | 1.931 | 1.860 | 1.760 | 1.691 | 1.641 | 1.572 | 1.526 | 1.457 | 1.418 | 1.374 | 1.325 |
| 80 | 2.769 | 2.370 | 2.154 | 2.016 | 1.921 | 1.849 | 1.748 | 1.680 | 1.629 | 1.559 | 1.513 | 1.443 | 1.403 | 1.358 | 1.307 |
| 100 | 2.756 | 2.356 | 2.139 | 2.002 | 1.906 | 1.834 | 1.732 | 1.663 | 1.612 | 1.542 | 1.494 | 1.423 | 1.382 | 1.336 | 1.282 |

## 8. F-jaotuse tabel (1 – p = 0,95)

| | $v_1$ | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $v_2$ | 1 | 2 | 3 | 4 | 5 | 6 | 8 | 10 | 12 | 16 | 20 | 30 | 40 | 60 | 120 |
| 1 | 161.45 | 199.50 | 215.71 | 224.58 | 230.16 | 233.99 | 238.88 | 241.88 | 243.91 | 246.46 | 248.01 | 250.10 | 251.14 | 252.20 | 253.25 |
| 2 | 18.513 | 19.000 | 19.164 | 19.247 | 19.296 | 19.330 | 19.371 | 19.396 | 19.413 | 19.433 | 19.446 | 19.462 | 19.471 | 19.479 | 19.487 |
| 3 | 10.128 | 9.552 | 9.277 | 9.117 | 9.013 | 8.941 | 8.845 | 8.786 | 8.745 | 8.692 | 8.660 | 8.617 | 8.594 | 8.572 | 8.549 |
| 4 | 7.709 | 6.944 | 6.591 | 6.388 | 6.256 | 6.163 | 6.041 | 5.964 | 5.912 | 5.844 | 5.803 | 5.746 | 5.717 | 5.688 | 5.658 |
| 5 | 6.608 | 5.786 | 5.409 | 5.192 | 5.050 | 4.950 | 4.818 | 4.735 | 4.678 | 4.604 | 4.558 | 4.496 | 4.464 | 4.431 | 4.398 |
| 6 | 5.987 | 5.143 | 4.757 | 4.534 | 4.387 | 4.284 | 4.147 | 4.060 | 4.000 | 3.922 | 3.874 | 3.808 | 3.774 | 3.740 | 3.705 |
| 7 | 5.591 | 4.737 | 4.347 | 4.120 | 3.972 | 3.866 | 3.726 | 3.637 | 3.575 | 3.494 | 3.445 | 3.376 | 3.340 | 3.304 | 3.267 |
| 8 | 5.318 | 4.459 | 4.066 | 3.838 | 3.687 | 3.581 | 3.438 | 3.347 | 3.284 | 3.202 | 3.150 | 3.079 | 3.043 | 3.005 | 2.967 |
| 9 | 5.117 | 4.256 | 3.863 | 3.633 | 3.482 | 3.374 | 3.230 | 3.137 | 3.073 | 2.989 | 2.936 | 2.864 | 2.826 | 2.787 | 2.748 |
| 10 | 4.965 | 4.103 | 3.708 | 3.478 | 3.326 | 3.217 | 3.072 | 2.978 | 2.913 | 2.828 | 2.774 | 2.700 | 2.661 | 2.621 | 2.580 |
| 12 | 4.747 | 3.885 | 3.490 | 3.259 | 3.106 | 2.996 | 2.849 | 2.753 | 2.687 | 2.599 | 2.544 | 2.466 | 2.426 | 2.384 | 2.341 |
| 14 | 4.600 | 3.739 | 3.344 | 3.112 | 2.958 | 2.848 | 2.699 | 2.602 | 2.534 | 2.445 | 2.388 | 2.308 | 2.266 | 2.223 | 2.178 |
| 16 | 4.494 | 3.634 | 3.239 | 3.007 | 2.852 | 2.741 | 2.591 | 2.494 | 2.425 | 2.333 | 2.276 | 2.194 | 2.151 | 2.106 | 2.059 |
| 18 | 4.414 | 3.555 | 3.160 | 2.928 | 2.773 | 2.661 | 2.510 | 2.412 | 2.342 | 2.250 | 2.191 | 2.107 | 2.063 | 2.017 | 1.968 |
| 20 | 4.351 | 3.493 | 3.098 | 2.866 | 2.711 | 2.599 | 2.447 | 2.348 | 2.278 | 2.184 | 2.124 | 2.039 | 1.994 | 1.946 | 1.896 |
| 24 | 4.260 | 3.403 | 3.009 | 2.776 | 2.621 | 2.508 | 2.355 | 2.255 | 2.183 | 2.088 | 2.027 | 1.939 | 1.892 | 1.842 | 1.790 |
| 28 | 4.196 | 3.340 | 2.947 | 2.714 | 2.558 | 2.445 | 2.291 | 2.190 | 2.118 | 2.021 | 1.959 | 1.869 | 1.820 | 1.769 | 1.714 |
| 32 | 4.149 | 3.295 | 2.901 | 2.668 | 2.512 | 2.399 | 2.244 | 2.142 | 2.070 | 1.972 | 1.908 | 1.817 | 1.767 | 1.714 | 1.657 |
| 36 | 4.113 | 3.259 | 2.866 | 2.634 | 2.477 | 2.364 | 2.209 | 2.106 | 2.033 | 1.934 | 1.870 | 1.776 | 1.726 | 1.671 | 1.612 |
| 40 | 4.085 | 3.232 | 2.839 | 2.606 | 2.449 | 2.336 | 2.180 | 2.077 | 2.003 | 1.904 | 1.839 | 1.744 | 1.693 | 1.637 | 1.577 |
| 50 | 4.034 | 3.183 | 2.790 | 2.557 | 2.400 | 2.286 | 2.130 | 2.026 | 1.952 | 1.850 | 1.784 | 1.687 | 1.634 | 1.576 | 1.511 |
| 60 | 4.001 | 3.150 | 2.758 | 2.525 | 2.368 | 2.254 | 2.097 | 1.993 | 1.917 | 1.815 | 1.748 | 1.649 | 1.594 | 1.534 | 1.467 |
| 70 | 3.978 | 3.128 | 2.736 | 2.503 | 2.346 | 2.231 | 2.074 | 1.969 | 1.893 | 1.790 | 1.722 | 1.622 | 1.566 | 1.505 | 1.435 |
| 80 | 3.960 | 3.111 | 2.719 | 2.486 | 2.329 | 2.214 | 2.056 | 1.951 | 1.875 | 1.772 | 1.703 | 1.602 | 1.545 | 1.482 | 1.411 |
| 100 | 3.936 | 3.087 | 2.696 | 2.463 | 2.305 | 2.191 | 2.032 | 1.927 | 1.850 | 1.746 | 1.676 | 1.573 | 1.515 | 1.450 | 1.376 |

## 9. F-jaotuse tabel (1 – p = 0,99)

| $v_2$ \ $v_1$ | 1 | 2 | 3 | 4 | 5 | 6 | 8 | 10 | 12 | 16 | 20 | 30 | 40 | 60 | 120 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 4052.2 | 4999.5 | 5403.4 | 5624.6 | 5763.6 | 5859.0 | 5981.1 | 6055.8 | 6106.3 | 6170.1 | 6208.7 | 6260.6 | 6286.8 | 6313.0 | 6339.4 |
| 2 | 98.503 | 99.000 | 99.166 | 99.249 | 99.299 | 99.333 | 99.374 | 99.399 | 99.416 | 99.437 | 99.449 | 99.466 | 99.474 | 99.482 | 99.491 |
| 3 | 34.116 | 30.817 | 29.457 | 28.710 | 28.237 | 27.911 | 27.489 | 27.229 | 27.052 | 26.827 | 26.690 | 26.505 | 26.411 | 26.316 | 26.221 |
| 4 | 21.198 | 18.000 | 16.694 | 15.977 | 15.522 | 15.207 | 14.799 | 14.546 | 14.374 | 14.154 | 14.020 | 13.838 | 13.745 | 13.652 | 13.558 |
| 5 | 16.258 | 13.274 | 12.060 | 11.392 | 10.967 | 10.672 | 10.289 | 10.051 | 9.888 | 9.680 | 9.553 | 9.379 | 9.291 | 9.202 | 9.112 |
| 6 | 13.745 | 10.925 | 9.780 | 9.148 | 8.746 | 8.466 | 8.102 | 7.874 | 7.718 | 7.519 | 7.396 | 7.229 | 7.143 | 7.057 | 6.969 |
| 7 | 12.246 | 9.547 | 8.451 | 7.847 | 7.460 | 7 191 | 6.840 | 6.620 | 6.469 | 6.275 | 6.155 | 5.992 | 5.908 | 5.824 | 5.737 |
| 8 | 11.259 | 8.649 | 7.591 | 7.006 | 6.632 | 6.371 | 6.029 | 5.814 | 5.667 | 5.477 | 5.359 | 5.198 | 5.116 | 5.032 | 4.946 |
| 9 | 10.561 | 8.022 | 6.992 | 6.422 | 6.057 | 5.802 | 5.467 | 5.257 | 5.111 | 4.924 | 4.808 | 4.649 | 4.567 | 4.483 | 4.398 |
| 10 | 10.044 | 7.559 | 6.552 | 5.994 | 5.636 | 5.386 | 5.057 | 4.849 | 4.706 | 4.520 | 4.405 | 4.247 | 4.165 | 4.082 | 3.996 |
| 12 | 9.330 | 6.927 | 5.953 | 5.412 | 5.064 | 4.821 | 4.499 | 4.296 | 4.155 | 3.972 | 3.858 | 3.701 | 3.619 | 3.535 | 3.449 |
| 14 | 8.862 | 6.515 | 5.564 | 5.035 | 4.695 | 4.456 | 4.140 | 3.939 | 3.800 | 3.619 | 3.505 | 3.348 | 3.266 | 3.181 | 3.094 |
| 16 | 8.531 | 6.226 | 5.292 | 4.773 | 4.437 | 4.202 | 3.890 | 3.691 | 3.553 | 3.372 | 3.259 | 3.101 | 3.018 | 2.933 | 2.845 |
| 18 | 8.285 | 6.013 | 5.092 | 4.579 | 4.248 | 4.015 | 3.705 | 3.508 | 3.371 | 3.190 | 3.077 | 2.919 | 2.835 | 2.749 | 2.660 |
| 20 | 8.096 | 5.849 | 4.938 | 4.431 | 4.103 | 3.871 | 3.564 | 3.368 | 3.231 | 3.051 | 2.938 | 2.778 | 2.695 | 2.608 | 2.517 |
| 24 | 7.823 | 5.614 | 4.718 | 4.218 | 3.895 | 3.667 | 3.363 | 3.168 | 3.032 | 2.852 | 2.738 | 2.577 | 2.492 | 2.403 | 2.310 |
| 28 | 7.636 | 5.453 | 4.568 | 4.074 | 3.754 | 3.528 | 3.226 | 3.032 | 2.896 | 2.716 | 2.602 | 2.440 | 2.354 | 2.263 | 2.167 |
| 32 | 7.499 | 5.336 | 4.459 | 3.969 | 3.652 | 3.427 | 3.127 | 2.934 | 2.798 | 2.618 | 2.503 | 2.340 | 2.252 | 2.160 | 2.062 |
| 36 | 7.396 | 5.248 | 4.377 | 3.890 | 3.574 | 3.351 | 3.052 | 2.859 | 2.723 | 2.543 | 2.428 | 2.263 | 2.175 | 2.082 | 1.981 |
| 40 | 7.314 | 5.179 | 4.313 | 3.828 | 3.514 | 3.291 | 2.993 | 2.801 | 2.665 | 2.484 | 2.369 | 2.203 | 2.114 | 2.019 | 1.917 |
| 50 | 7 171 | 5.057 | 4.199 | 3.720 | 3.408 | 3.186 | 2.890 | 2.698 | 2.562 | 2.382 | 2.265 | 2.098 | 2.007 | 1.909 | 1.803 |
| 60 | 7.077 | 4.977 | 4.126 | 3.649 | 3.339 | 3.119 | 2.823 | 2.632 | 2.496 | 2.315 | 2.198 | 2.028 | 1.936 | 1.836 | 1.726 |
| 70 | 7.011 | 4.922 | 4.074 | 3.600 | 3.291 | 3.071 | 2.777 | 2.585 | 2.450 | 2.268 | 2.150 | 1.980 | 1.886 | 1.785 | 1.672 |
| 80 | 6.963 | 4.881 | 4.036 | 3.563 | 3.255 | 3.036 | 2.742 | 2.551 | 2.415 | 2.233 | 2.115 | 1.944 | 1.849 | 1.746 | 1.630 |
| 100 | 6.895 | 4.824 | 3.984 | 3.513 | 3.206 | 2.988 | 2.694 | 2.503 | 2.368 | 2.185 | 2.067 | 1.893 | 1.797 | 1.692 | 1.572 |

## 10. F-jaotuse tabel (1 – p = 0,999)

| $v_2$ \ $v_1$ | 1 | 2 | 3 | 4 | 5 | 6 | 8 | 10 | 12 | 16 | 20 | 30 | 40 | 60 | 120 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 405284 | 499999 | 540379 | 562500 | 576405 | 585937 | 598144 | 605621 | 610668 | 617045 | 620908 | 626099 | 628712 | 631337 | 633972 |
| 2 | 998.50 | 999.00 | 999.17 | 999.25 | 999.30 | 999.33 | 999.37 | 999.40 | 999.42 | 999.44 | 999.45 | 999.47 | 999.47 | 999.48 | 999.49 |
| 3 | 167.03 | 148.50 | 141.11 | 137.10 | 134.58 | 132.85 | 130.62 | 129.25 | 128.32 | 127.14 | 126.42 | 125.45 | 124.96 | 124.47 | 123.97 |
| 4 | 74.137 | 61.246 | 56.177 | 53.436 | 51.712 | 50.525 | 48.996 | 48.053 | 47.412 | 46.597 | 46.100 | 45.429 | 45.089 | 44.746 | 44.400 |
| 5 | 47.181 | 37.122 | 33.202 | 31.085 | 29.752 | 28.834 | 27.649 | 26.917 | 26.418 | 25.783 | 25.395 | 24.869 | 24.602 | 24.333 | 24.060 |
| 6 | 35.507 | 27.000 | 23.703 | 21.924 | 20.803 | 20.030 | 19.030 | 18.411 | 17.989 | 17.450 | 17.120 | 16.672 | 16.445 | 16.214 | 15.981 |
| 7 | 29.245 | 21.689 | 18.772 | 17.198 | 16.206 | 15.521 | 14.634 | 14.083 | 13.707 | 13.226 | 12.932 | 12.530 | 12.326 | 12.119 | 11.909 |
| 8 | 25.415 | 18.494 | 15.829 | 14.392 | 13.485 | 12.858 | 12.046 | 11.540 | 11.194 | 10.752 | 10.480 | 10.109 | 9.919 | 9.727 | 9.532 |
| 9 | 22.857 | 16.387 | 13.902 | 12.560 | 11.714 | 11.128 | 10.368 | 9.894 | 9.570 | 9.154 | 8.898 | 8.548 | 8.369 | 8.187 | 8.001 |
| 10 | 21.040 | 14.905 | 12.553 | 11.283 | 10.481 | 9.926 | 9.204 | 8.754 | 8.445 | 8.048 | 7.804 | 7.469 | 7.297 | 7.122 | 6.944 |
| 12 | 18.643 | 12.974 | 10.804 | 9.633 | 8.892 | 8.379 | 7.710 | 7.292 | 7.005 | 6.634 | 6.405 | 6.090 | 5.928 | 5.762 | 5.593 |
| 14 | 17.143 | 11.779 | 9.729 | 8.622 | 7.922 | 7.436 | 6.802 | 6.404 | 6.130 | 5.776 | 5.557 | 5.254 | 5.098 | 4.938 | 4.773 |
| 16 | 16.120 | 10.971 | 9.006 | 7.944 | 7.272 | 6.805 | 6.195 | 5.812 | 5.547 | 5.205 | 4.992 | 4.697 | 4.545 | 4.388 | 4.226 |
| 18 | 15.379 | 10.390 | 8.487 | 7.459 | 6.808 | 6.355 | 5.763 | 5.390 | 5.132 | 4.798 | 4.590 | 4.301 | 4.151 | 3.996 | 3.836 |
| 20 | 14.819 | 9.953 | 8.098 | 7.096 | 6.461 | 6.019 | 5.440 | 5.075 | 4.823 | 4.495 | 4.290 | 4.005 | 3.856 | 3.703 | 3.544 |
| 24 | 14.028 | 9.339 | 7.554 | 6.589 | 5.977 | 5.550 | 4.991 | 4.638 | 4.393 | 4.074 | 3.873 | 3.593 | 3.447 | 3.295 | 3.136 |
| 28 | 13.498 | 8.931 | 7.193 | 6.253 | 5.656 | 5.241 | 4.695 | 4.349 | 4.109 | 3.795 | 3.598 | 3.321 | 3.176 | 3.024 | 2.864 |
| 32 | 13.117 | 8.639 | 6.936 | 6.014 | 5.429 | 5.021 | 4.485 | 4.145 | 3.908 | 3.598 | 3.403 | 3.128 | 2.983 | 2.831 | 2.670 |
| 36 | 12.832 | 8.420 | 6.744 | 5.836 | 5.260 | 4.857 | 4.328 | 3.992 | 3.758 | 3.451 | 3.258 | 2.984 | 2.839 | 2.686 | 2.524 |
| 40 | 12.609 | 8.251 | 6.595 | 5.698 | 5.128 | 4.731 | 4.207 | 3.874 | 3.642 | 3.338 | 3.145 | 2.872 | 2.727 | 2.574 | 2.410 |
| 50 | 12.222 | 7.956 | 6.336 | 5.459 | 4.901 | 4.512 | 3.998 | 3.671 | 3.443 | 3.142 | 2.951 | 2.679 | 2.533 | 2.378 | 2.211 |
| 60 | 11.973 | 7.768 | 6.171 | 5.307 | 4.757 | 4.372 | 3.865 | 3.541 | 3.315 | 3.017 | 2.827 | 2.555 | 2.409 | 2.252 | 2.082 |
| 70 | 11.799 | 7.637 | 6.057 | 5.201 | 4.656 | 4.275 | 3.773 | 3.452 | 3.227 | 2.930 | 2.741 | 2.469 | 2.322 | 2.164 | 1.991 |
| 80 | 11.671 | 7.540 | 5.972 | 5.123 | 4.582 | 4.204 | 3.705 | 3.386 | 3.162 | 2.867 | 2.677 | 2.406 | 2.258 | 2.099 | 1.924 |
| 100 | 11.495 | 7.408 | 5.857 | 5.017 | 4.482 | 4.107 | 3.612 | 3.296 | 3.074 | 2.780 | 2.591 | 2.319 | 2.170 | 2.009 | 1.829 |