

JAOTUSSEADUSED II

LOENGUTEEMAD

- Poissoni jaotus
- Eksponentjaotus
- Normaaljaotus

POISSONI JAOTUS

BINOOMJAOTUS

Kas minutis astub poodi ostja või mitte? Kaheväärtuseline tunnus: astub sisse/ei astu sisse. Olgu üksiksündmuse tõenäosus näiteks $p=0,8$.

Kui suur on tõenäosus, et tunnis on m minutit, kus siseneb üks ostja?
Katsete arv $n=60$. Binoomjaotusest

$$P(X = m) = C_{60}^m 0,8^m 0,2^{60-m}$$

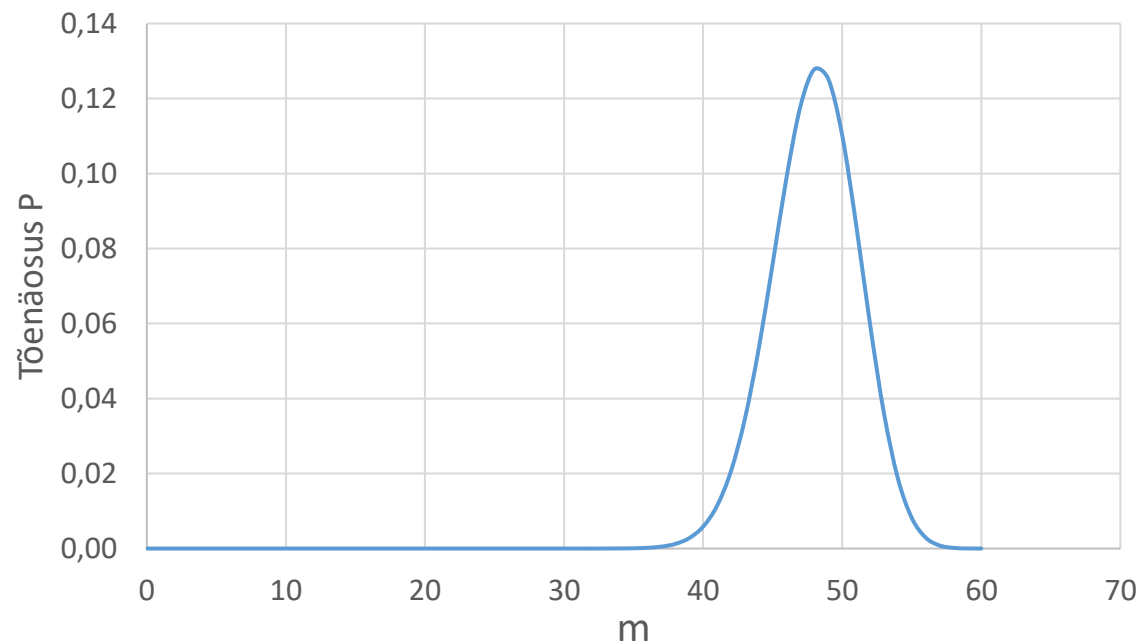
$$P(X = 0) = 1,15 \cdot 10^{-42}$$

...

$$P(X = 48) = 0,128$$

...

$$P(X = 60) = 1,53 \cdot 10^{-6}$$



BINOOMJAOTUS -> POISSONI JAOTUS

Kas minutis astub poodi ostja või mitte? Kaheväärtuseline tunnus: : astub sisse/ei astu sisse.
Olgu üksiksündmuse tõenäosus näiteks $p=0,8$.

Aga kui mõnes minutis siseneb **mitu** ostjat?

Siis pole kaheväärtuseline tunnus, binoomjaotust kasutada ei saa.

Lahendus: vaatleme sekundeid. Kas sekundis astub poodi ostja või mitte?

Kaheväärtuseline tunnus, binoomjaotus: $X \sim B(n, p)$

Tunnis sekundeid $n = 60 \cdot 60 = 3600$

Tõenäosus, et sekundis $p = \frac{0,8}{60} \approx 0,0133$

$$X \sim B(3600, 0,0133)$$

Läheme üle piirile,

kus katsete arv $n \rightarrow \infty$

positiivse sündmuse tõenäosus $p \rightarrow 0$

$$C_n^m p^m (1-p)^{n-m} \xrightarrow[n \rightarrow \infty]{p \rightarrow 0} e^{-\lambda} \frac{\lambda^m}{m!}$$



Binoomjaotus $\rightarrow$ Poissoni jaotus

POISSONI VALEM

Kui λ on keskmine haruldaste sündmuste arv intervallis, siis m sündmuse toimumise tõenäosus on leitav Poissoni valemi abil:

$$P(X = m) = e^{-\lambda} \frac{\lambda^m}{m!}$$

$e=2,718\dots$ on Euleri arv

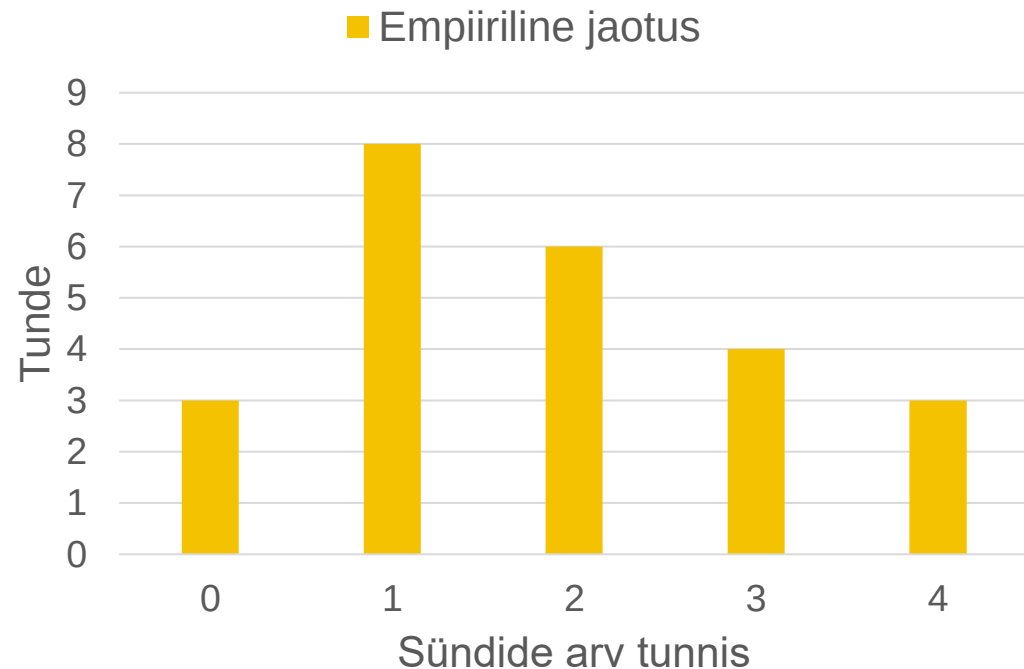
NÄIDE: BEEBIBUUM JÕULUDE AJAL AUSTRALIAS, I

18. detsembril 1997 sündis Austraalias Mater Mother's Hospitalis 24 tunni jooksul 44 last.

Allikas: The Sunday Mail, 21. dets 1997

Empiiriline jaotus

Sündide arv tunnis	Tunde
0	3
1	8
2	6
3	4
4	3
	24



NÄIDE: BEEBIBUUM JÕULUDE AJAL AUSTRALIAS, II

18. detsembril 1997 sündis Austraalias Mater Mother's Hospitalis 24 tunni jooksul 44 last.

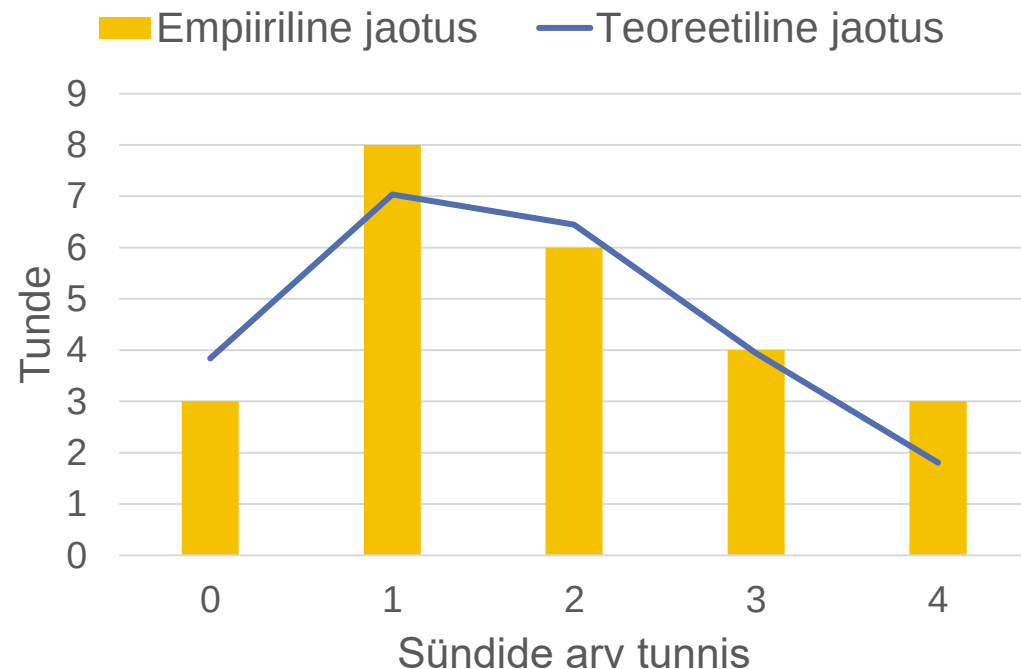
Keskmine sündide arv tunnis $\lambda = \frac{44}{24} \approx 1,833$ on Poissoni jaotuse keskväärtus.

Tõenäosus, et tunnis m sündi: $P(X = m) = e^{-\lambda} \frac{\lambda^m}{m!} = e^{-1,833} \frac{1,833^m}{m!}$

Sündide arv tunnis m	Tõenäosus P	Tunde
0	0,160	3,8
1	0,293	7,0
2	0,269	6,4
3	0,164	3,9
4	0,075	1,8

Teoreetiline jaotus

Tunde, kus $X=m$
 $24 \cdot P(X = m)$

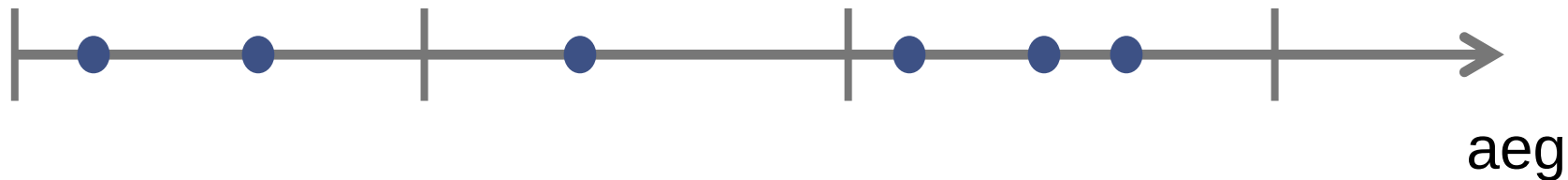


POISSONI JAOTUS

Poissoni jaotus kirjeldab diskreetset juhuslikku suurust, mille väärtuseks on haruldaste ($p < 0,1$) sündmuste arv m teatud intervallis.

Sündmused toimuvad

- juhuslikult ja üksteisest sõltumatult;
- harva;
- kindlas intervallis (ajaline või ka ruumiline intervall).

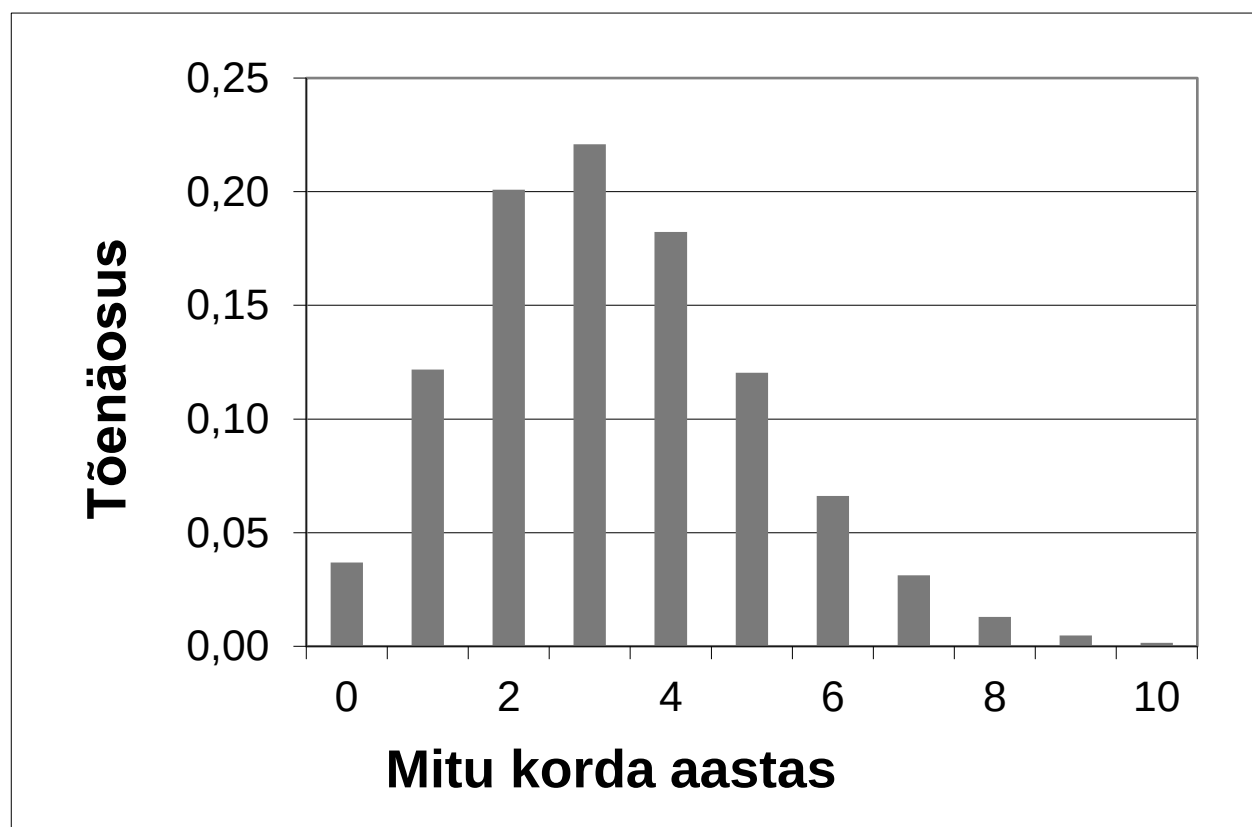


Demo: Poissoni protsess

Demo: Poissoni jaotus

NÄIDE: PEREARSTI KÜLASTAMINE, 1

Praxise uuring perearsti praksiste kohta Tallinnas ja Harjumaal, 2008.
Viimase 12 kuu jooksul külastasid uuritavad perearsti keskmiselt 3,3 korda.



Teoreetiline jaotus: Poissoni jaotus.

Jaotuse keskväärtus
empiirilisest jaotusest.

Poissoni jaotus
keskväärtusega 3,3.

Tõenäosuste arvutamisel
kasutame vastavat
Poissoni jaotust.

NÄIDE: PEREARSTI KÜLASTAMINE, 2

Konkreetsse perearsti nimistus on 1500 patsienti.

1. Mitu patsienti ei külasta perearsti aasta jooksul **kordagi**?

Keskväärtus $\lambda = 3,3$ külastust aastas

Tõenäosus, et külastuste arv aastas $m = 0$ $P(X = 0) = e^{-\lambda} \frac{\lambda^m}{m!} = e^{-3,3} \frac{3,3^0}{0!} \approx 0,037$

n katse korral $m=0$ 3,7% katsetest.

Mitu patsienti 1500-st ei külasta kordagi? $m = Pn = 0,037 \cdot 1500 \approx 55$

NÄIDE: PEREARSTI KÜLASTAMINE, 3

Konkreetsse perearsti nimistus on 1500 patsienti.

2. Mitu patsienti külastab perearsti aasta jooksul **rohkem kui 3 korda**?

$$P(X > 3) = 1 - P(X \leq 3) = 1 - (P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3))$$

Tõenäosus, et külastuste arv on kuni 3

$$P(X \leq 3) = e^{-3,3} \left(\frac{3,3^0}{0!} + \frac{3,3^1}{1!} + \frac{3,3^2}{2!} + \frac{3,3^3}{3!} \right) \approx 0,58$$

Tõenäosus, et külastuste arv on suurem kui 3

$$P(X > 3) = 1 - 0,58 = 0,42$$

Mitu patsienti 1500-st

$$0,42 \cdot 1500 \approx 629$$

NÄIDE: PEREARSTI KÜLASTAMINE, 4

Konkreetsse perearsti nimistus on 1500 patsienti.

3. Mitu patsienti külastab perearsti kord **kahe kuu** jooksul?

Keskväärtus $\lambda = 3,3$ külastust aastas

Keskväärtus **kahe kuu** kohta $\lambda = \frac{3,3}{6} = 0,55$ külastust

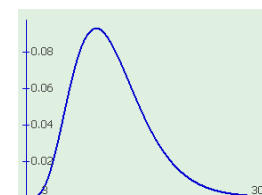
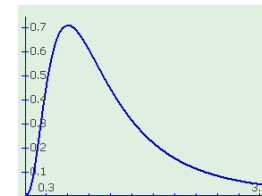
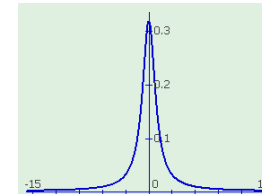
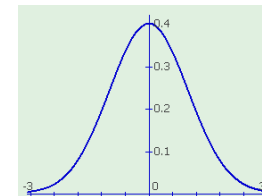
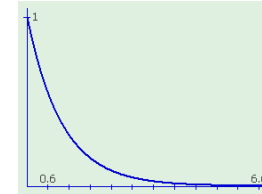
Muudame perioodi:
uus keskväärtus.

Tõenäosus, et külastuste arv on 1. $P(X = 1) = e^{-0,55} \frac{0,55^1}{1!} \approx 0,368$

Mitu patsienti 1500-st $0,368 \cdot 1500 \approx 551$

PIDEVAD JAOTUSED

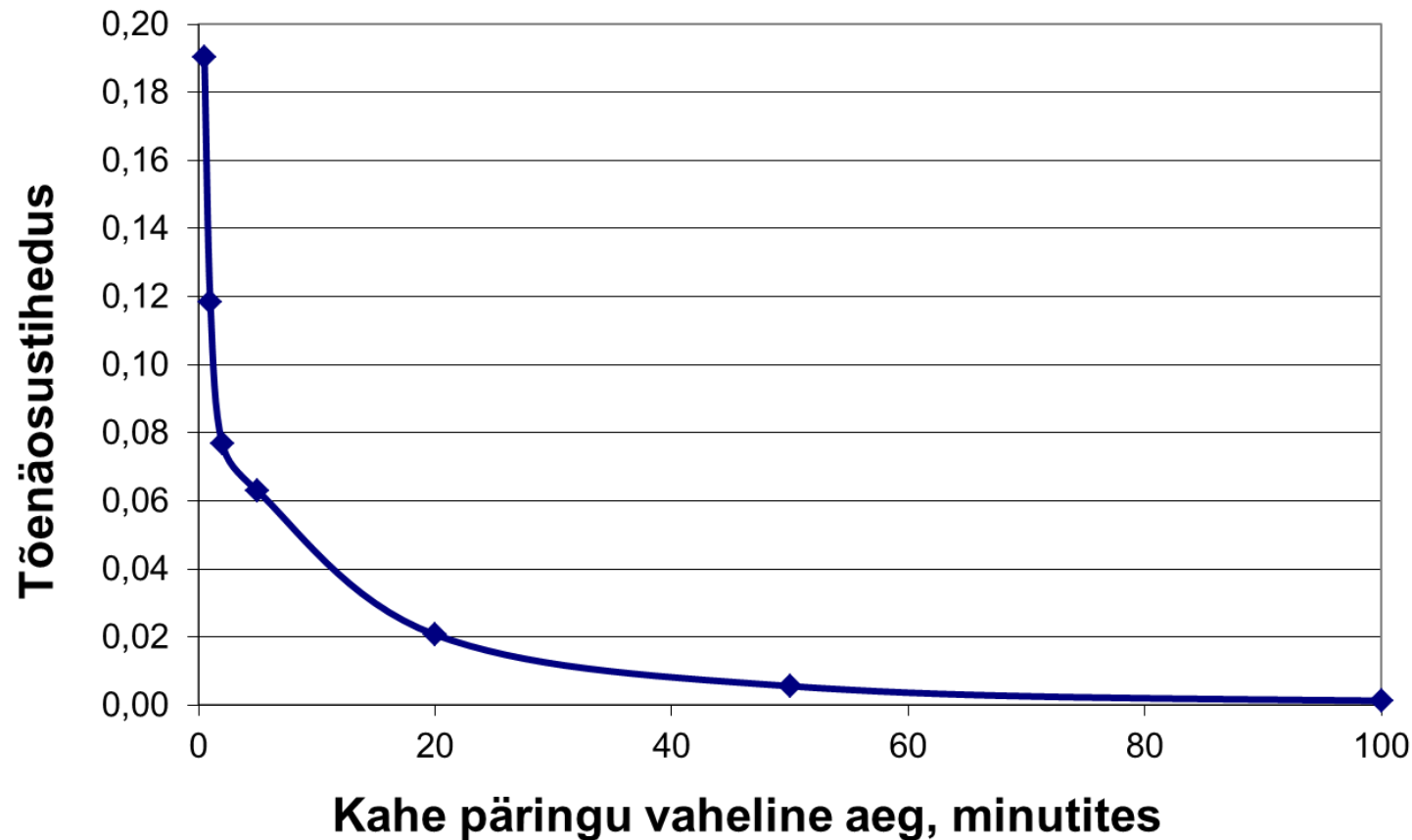
- Eksponentjaotus
 - Normaaljaotus
 - t -jaotus ehk Studenti jaotus
 - F -jaotus ehk Fisheri jaotus
 - χ^2 jaotus (hii-ruut)
-



EKSPONENTJAOTUS

NÄIDE: PÄRINGUTE VAHELINE AEG

Päringud Audentese Ülikooli õppeinfosüsteemist 28.01.2003-25.02.2003.
1183 päringut, registreeritud kahe päringu vaheline aeg.



EKSPONENTJAOTUS

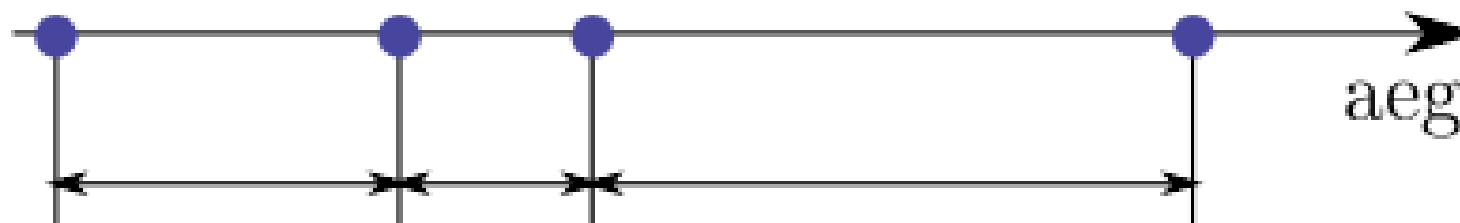
EkspONENTJAOTUSELE allub ajavahemik kahe sõltumatult toimuva samaliigilise järjestikuse sündmuse vahel.

Ajavahemik kahe järjestikuse

- kliendi saabumise vahel;
- telefonikõne vahel;
- veebilehelt tehtud päringu vahel.

Mõnikord ka mingi tegevuse aeg

- kliendi teenindamine;
- detaili eluiga.

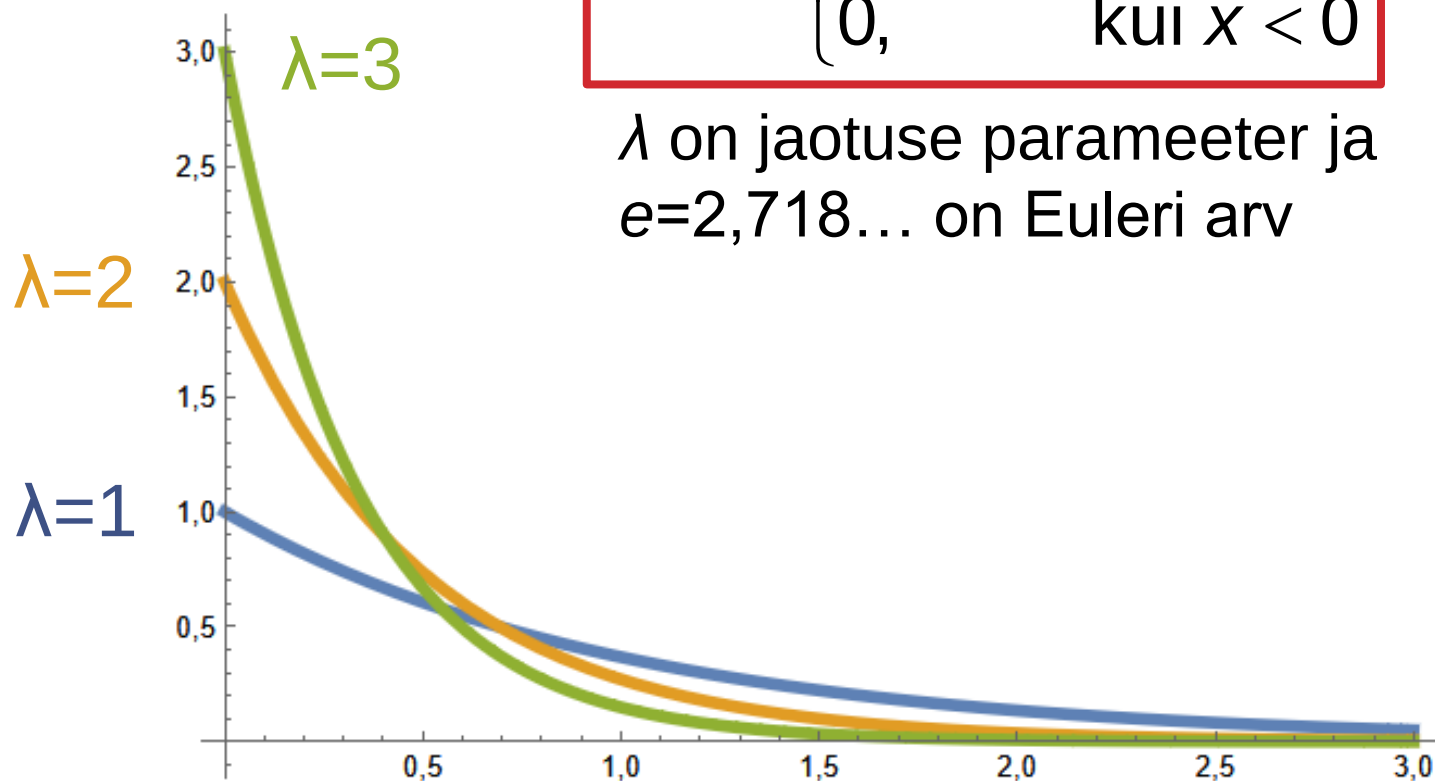


JAOTUSTIHEDUS

Eksponentjaotuse jaotustihedus

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

λ on jaotuse parameeter ja
 $e=2,718\dots$ on Euleri arv



KESKVÄÄRTUS

Pideva jaotuse keskväärtus üldjuhul

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

jaotustihedus

Eksponentjaotuse korral

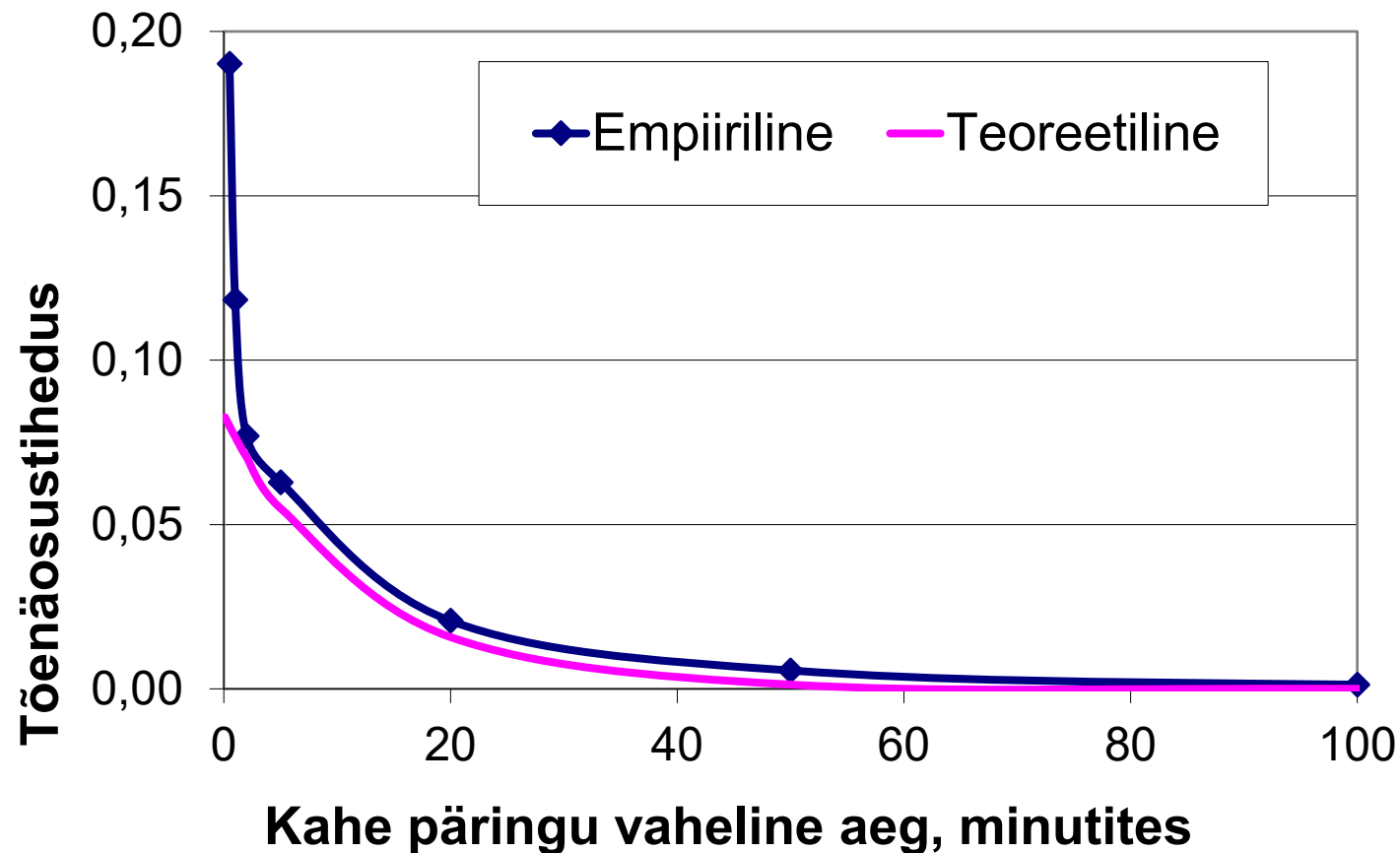
$$\mu = \int_0^{\infty} \lambda e^{-\lambda x} x dx = \frac{1}{\lambda}$$

Eksponentjaotuse keskväärtus

$$\mu = \frac{1}{\lambda}$$

NÄIDE: PÄRINGUTE VAHELINE AEG, 2

Päringud Audentese õppeinfosüsteemist.
Keskväärtus 12 min.



JAOTUSFUNKTSIOON

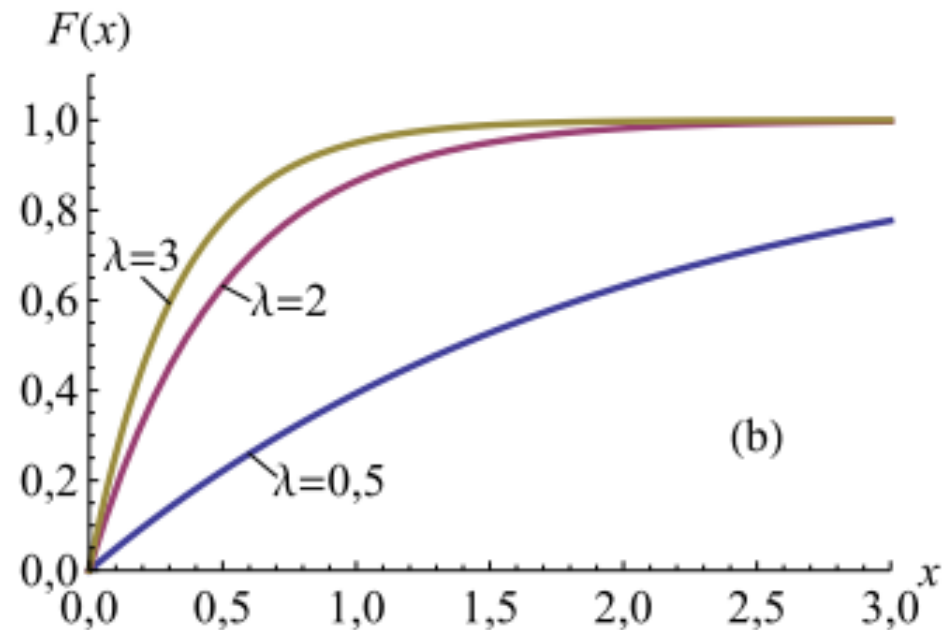
Pideva jaotuse
jaotusfunktsioon üldjuhul

$$F(a) = P(X \leq a) = \int_{-\infty}^a f(x) dx$$

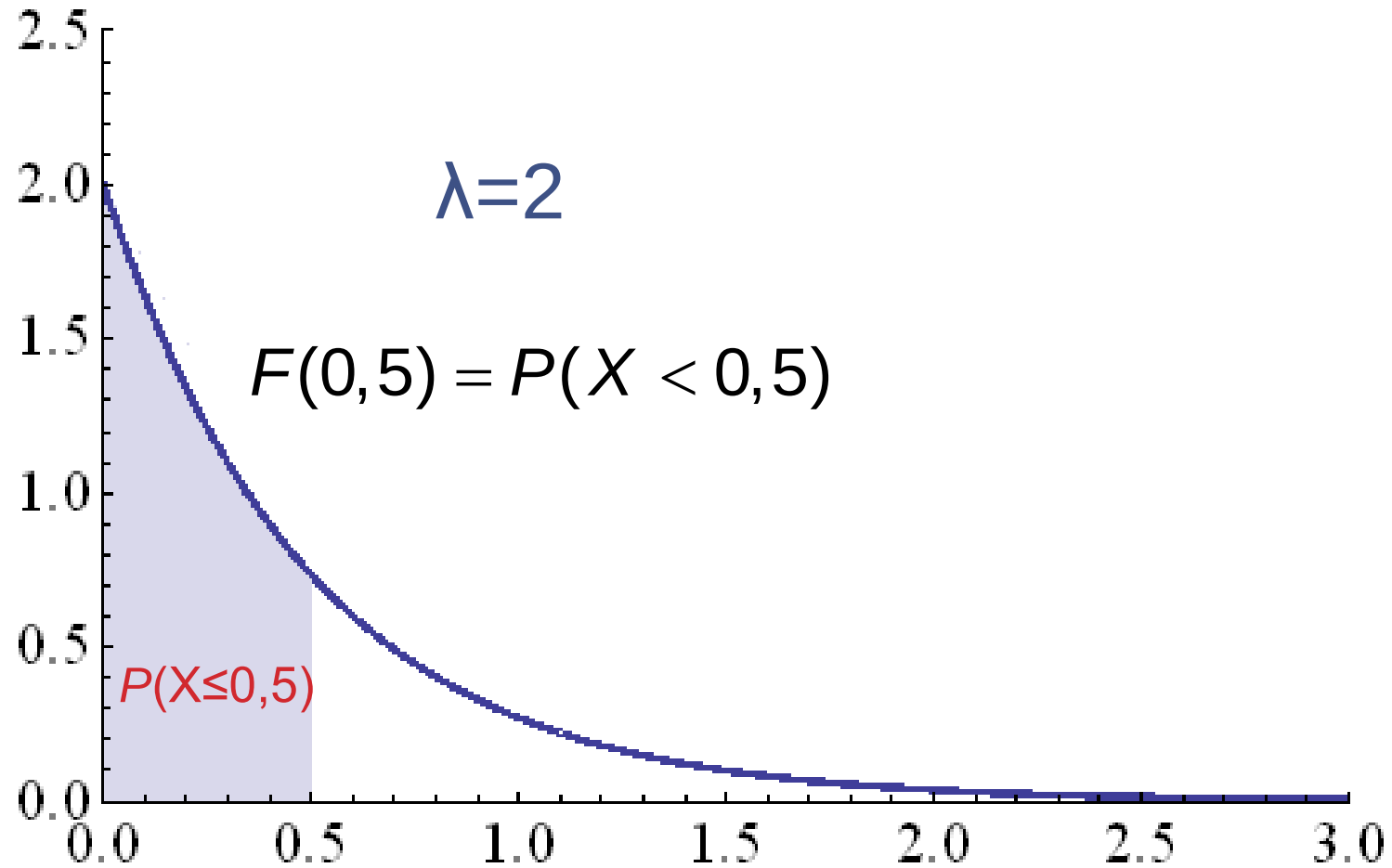
Eksponentjaotuse
jaotusfunktsioon

$$F(a) = \int_0^a \lambda e^{-\lambda x} dx$$

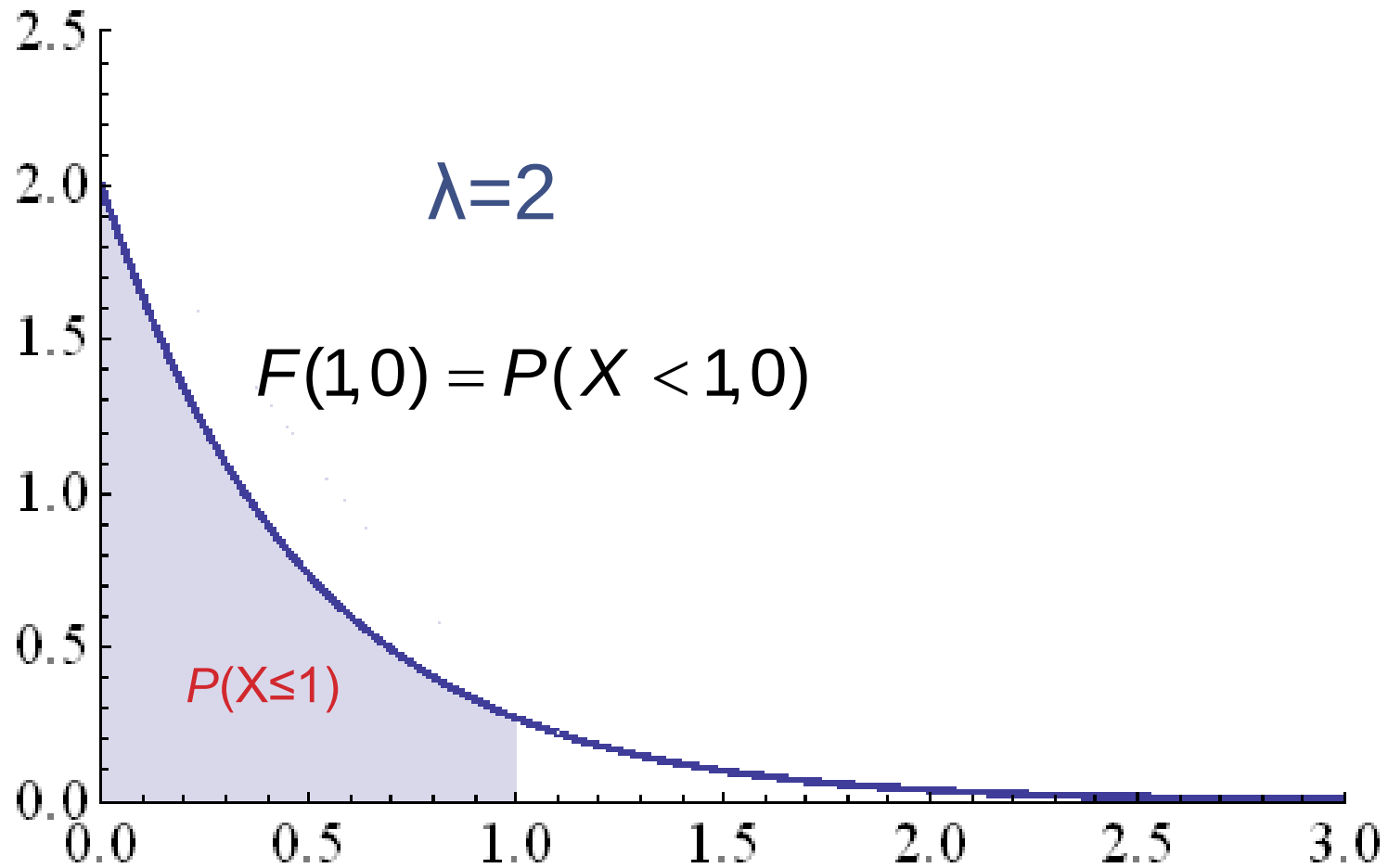
$$F(x) = \begin{cases} 1 - e^{-\lambda x}, & \text{kui } x \geq 0 \\ 0, & \text{kui } x < 0 \end{cases}$$



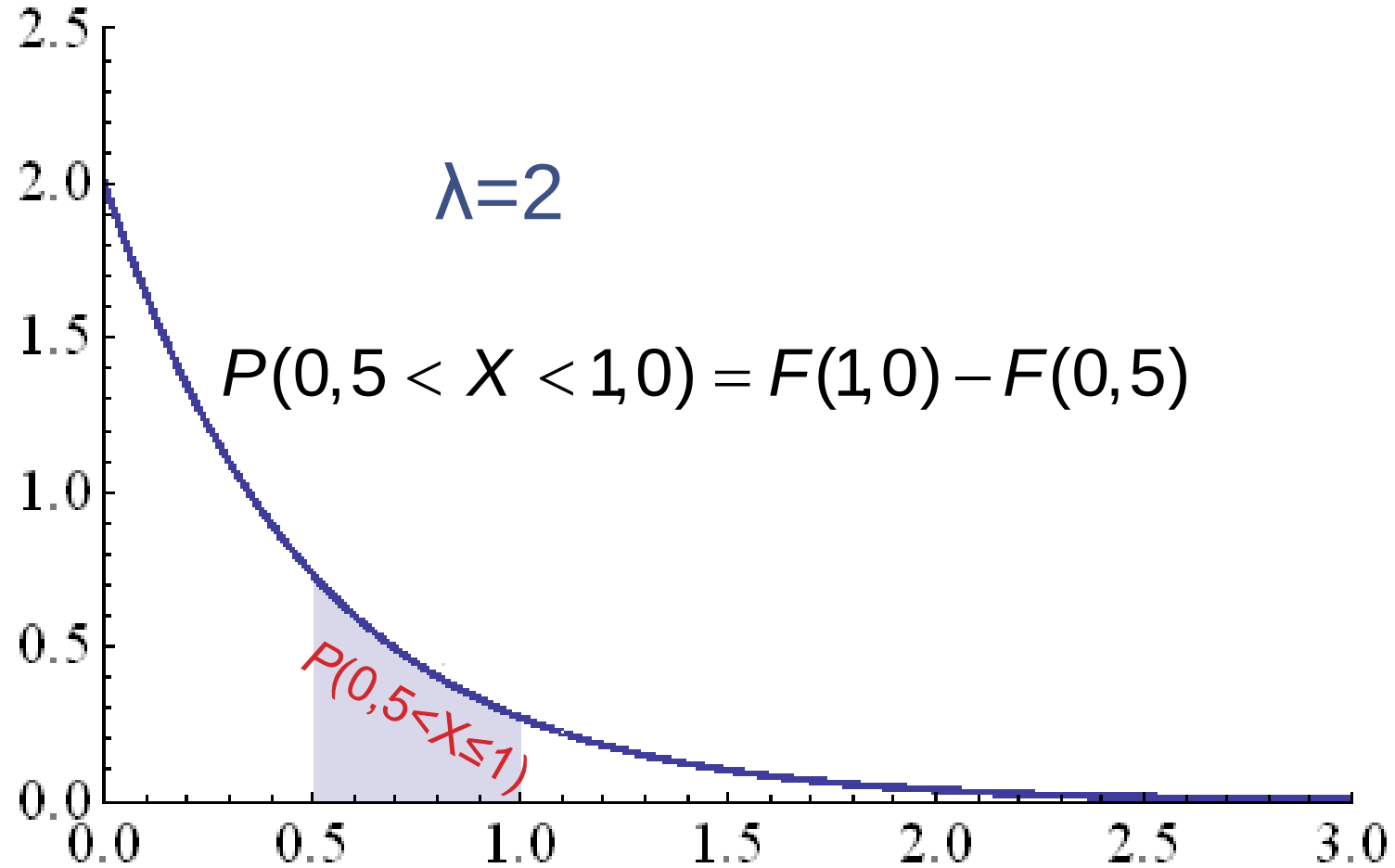
TÕENÄOSUS EKSPONENTJAOTUSE KORRAL, I



TÕENÄOSUS EKSPONENTJAOTUSE KORRAL, II



TÕENÄOSUS EKSPONENTJAOOTUSE KORRAL, III



NÄIDE: HÄDAABINUMBER 112, OOTEAEG



2024. aastal helistati hädaabiliinile 989 711 korral. Keskmine ooteaeg helistajale oli 5,6 sekundit *).

1. Kui suure tõenäosusega oli hädaabinumbrile helistamisel ooteaeg kuni 2 sekundit?

$$P(X < 2) = F(2)$$

$$F(2) = 1 - e^{-\lambda \cdot 2} = 1 - e^{-\frac{1}{\mu} \cdot 2} = 1 - e^{-\frac{1}{5,6} \cdot 2} \approx 0,3 \quad \text{sest} \quad \lambda = \frac{1}{\mu} = \frac{1}{5,6}$$

2. Kui suure tõenäosusega oli hädaabinumbrile helistamisel ooteaeg kauem kui 30 sekundit?

$$P(X > 30) = 1 - P(X < 30) = 1 - F(30)$$

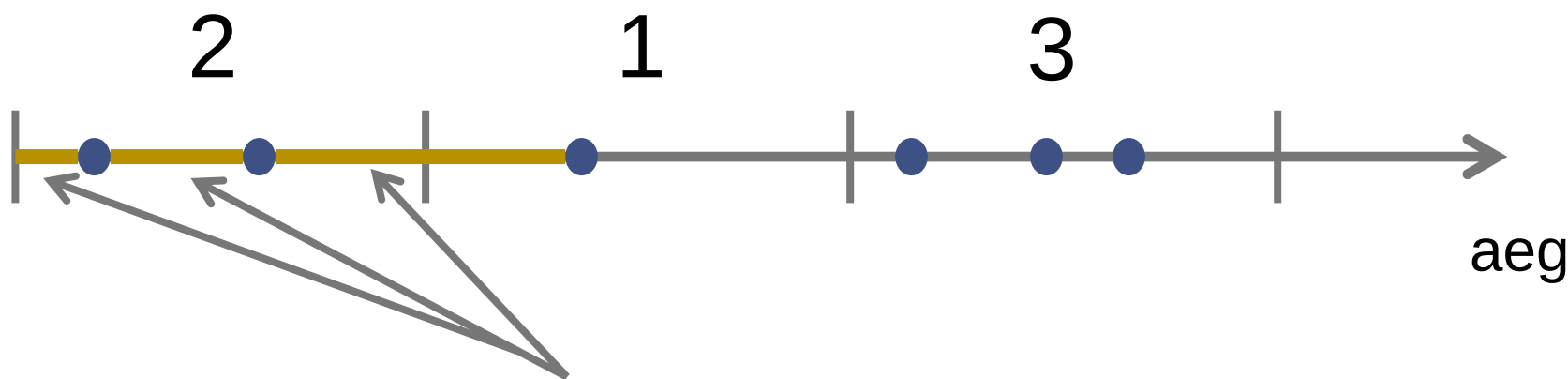
$$F(30) = 1 - e^{-\frac{1}{5,6} \cdot 30} \approx 0,995 \quad \text{Tõenäosus, et kuni 30 sekundit.}$$

$$P(X > 30) = 1 - 0,995 = 0,005 \quad \text{Kauem kui 30 s: viiel juhul tuhandest.}$$

*) Allikas: Häirekeskus. Statistika ja uuringud. <https://www.112.ee/et/juhend/statistika>

SEOS POISSONI JA EKSPONENTJAOTUSE VAHEL

Sündmuste arv ajaintervallis allub **Poissoni jaotusele**.



Sündmuste vaheline aeg allub **eksponentjaotusele**.

POISSONI JA EKSPONENTJAOTUSE KASUTAMINE

- Järjekorrateoorias (massteeninduse teooria)
- Logistikas
- Kaubavarude juhtimisel
- Infotehnoloogias

jm

Näide: järjekorra arvutused

NORMAALJAOTUS

NÄITED NORMAALJAOTUSEST

Kalade suurus.

Katkend filmist "Aegruum. Koodid"
(ETV2)



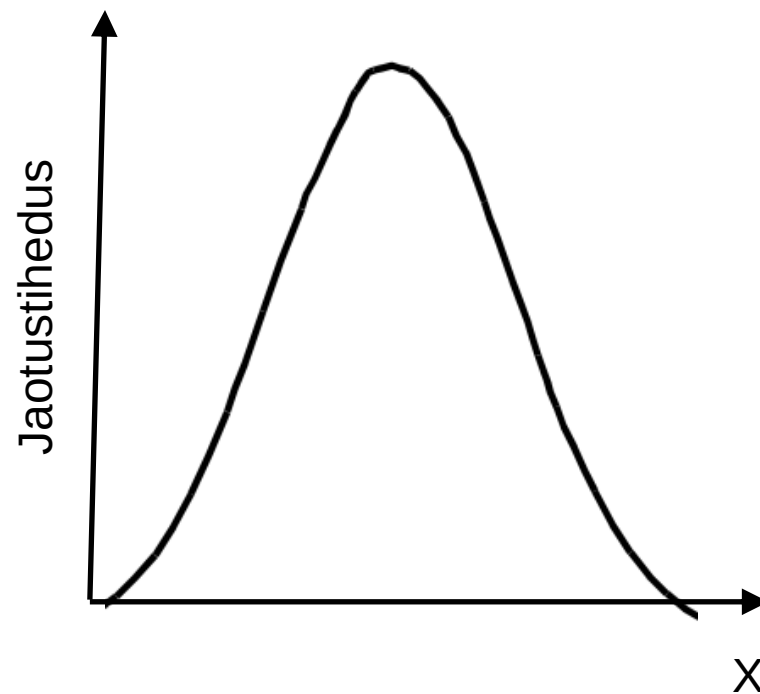
Animatsioon Galton Board

<https://www.mathsisfun.com/data/quincunx.html>

NORMAALJAOTUSELE ALLUV JUHUSLIK SUURUS

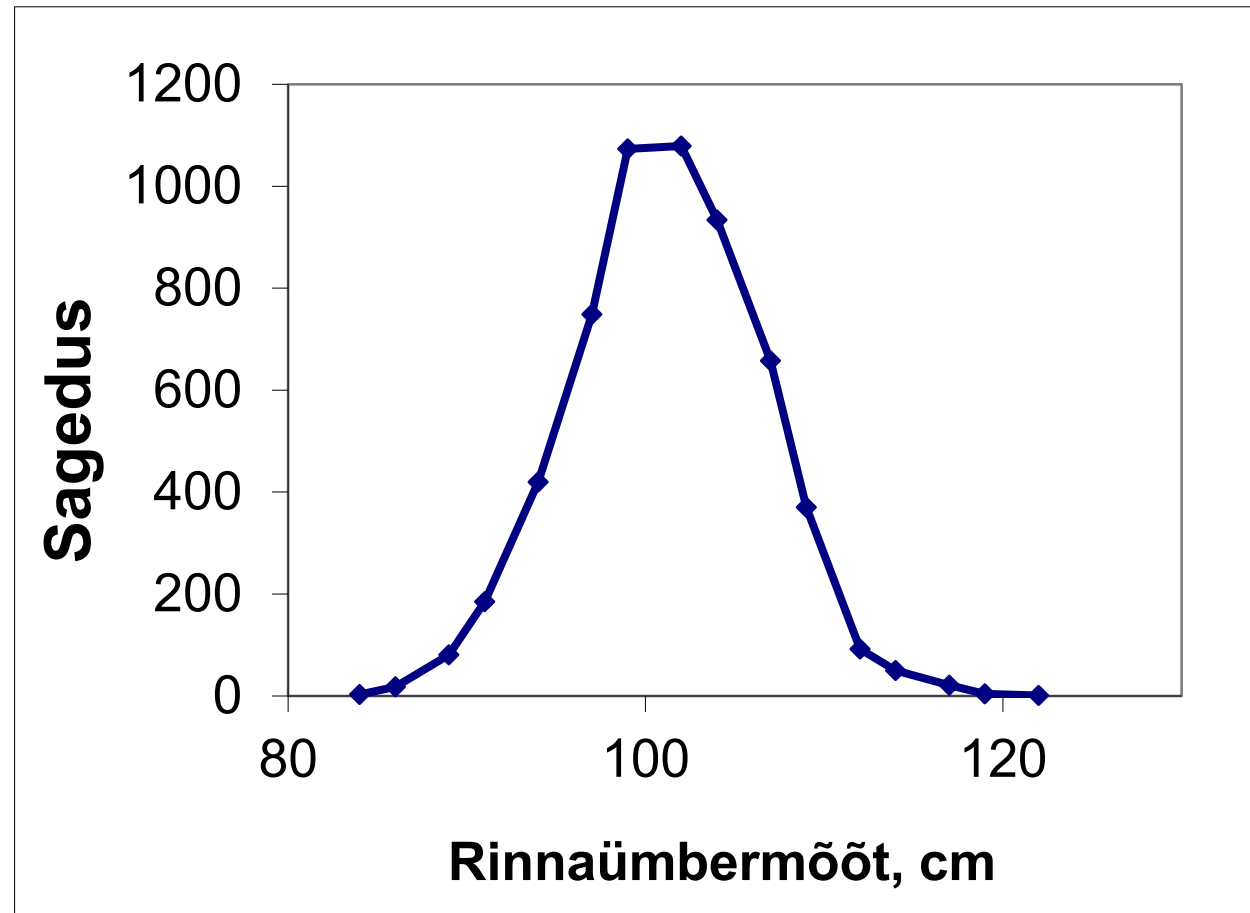
Juhuslik suurus allub normaaljaotusele, kui see

- on mõjutatud paljude faktorite poolt;
- iga üksikfaktori mõju on väike;
- puudub domineeriv faktor.



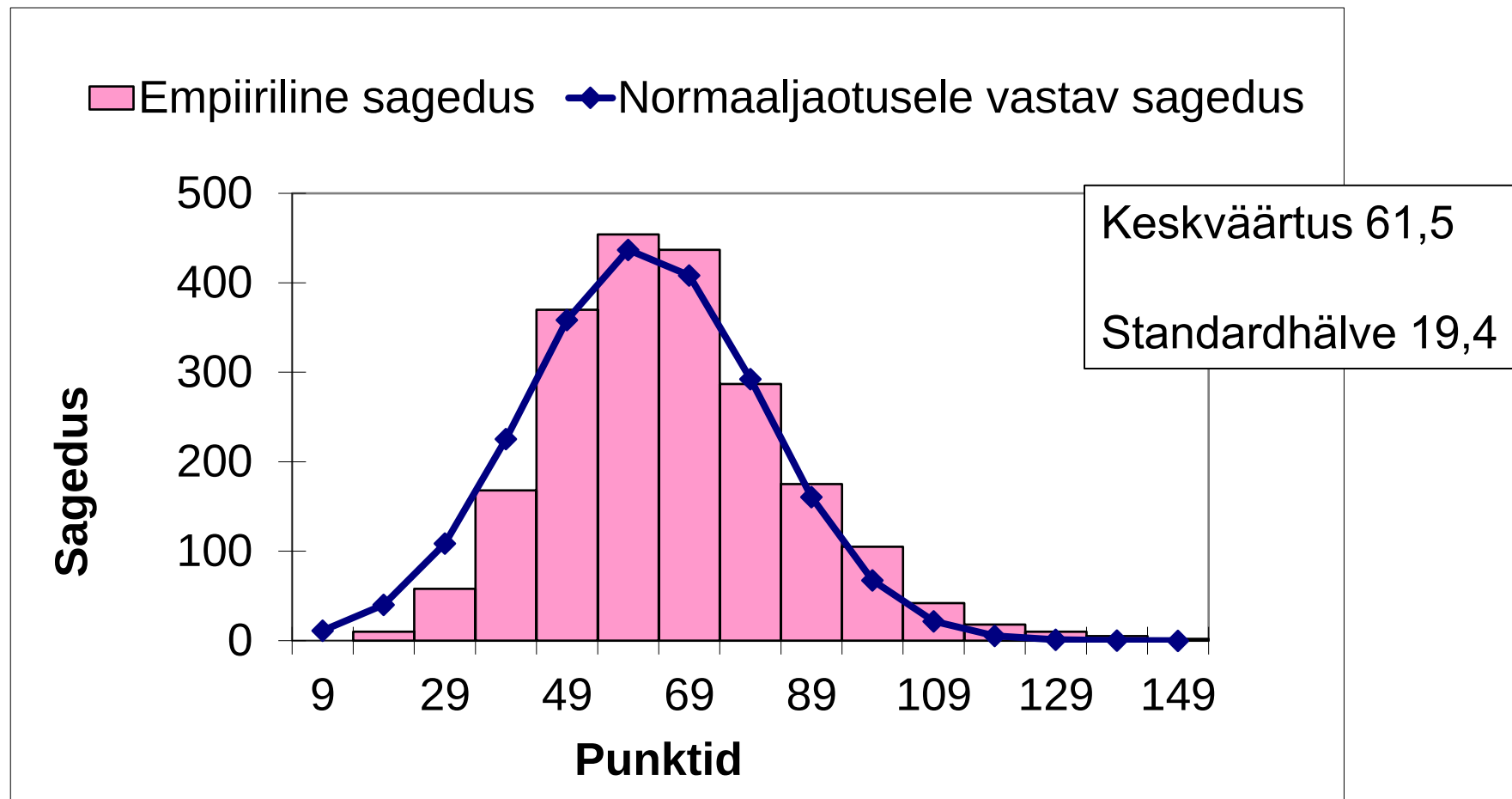
NÄIDE: NORMAALJAOTUS

Šoti sõdurite rinnaümbermõõt. Andmed pärinevad 19. saj algusest.



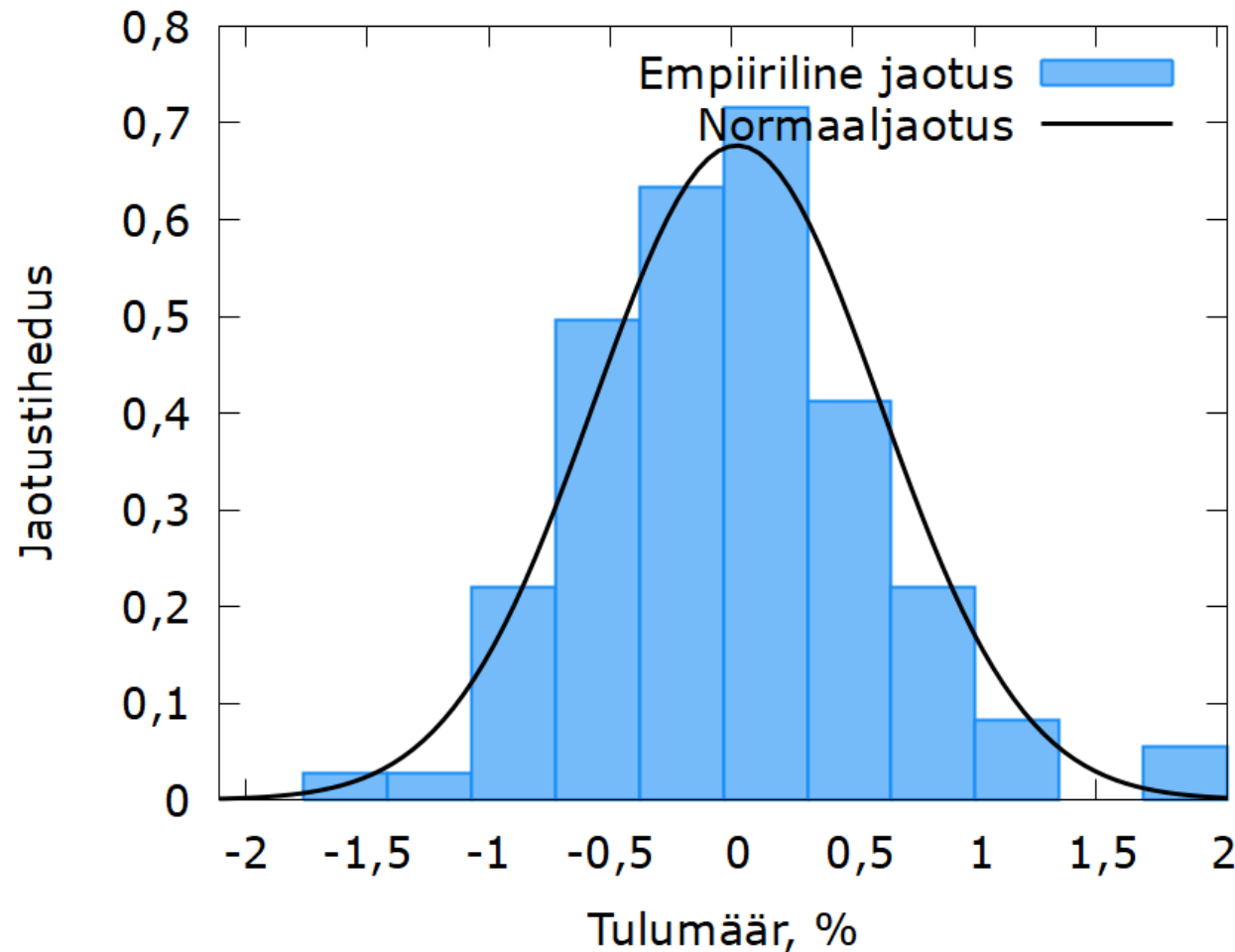
NÄIDE: NORMAALJAOTUS

Vabariikliku matemaatikaolümpiaadi KÄNGURU 2002 tulemused.
Osavõtjaid 2141.



NÄIDE: NORMAALJAOTUS

Merko Ehituse aktsia tulumäär 4. juuli kuni 29. dets 2023.



Keskväärtus 0,023%

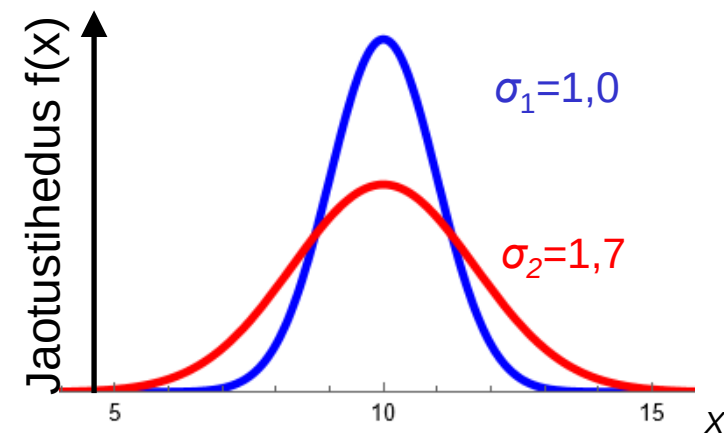
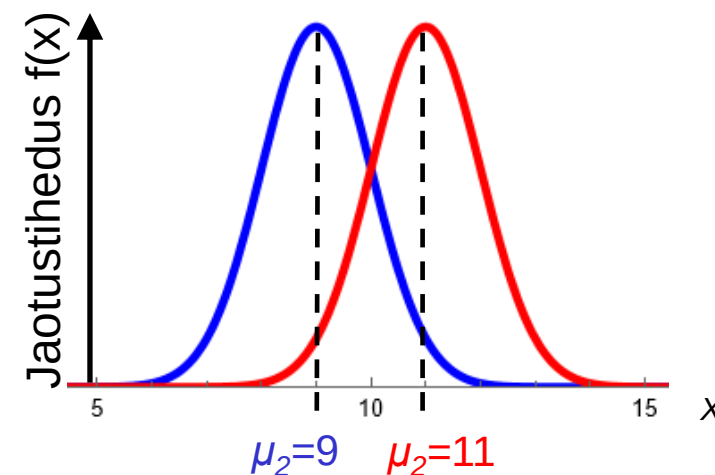
Standardhälve 0,59%

NORMAALJAOTUSE JAOTUSTIHEDUSE KÕVER

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Määratud kahe parameetriga:

- keskväärtus μ määrab ära jaotuskõvera asukoha;
- standardhälbest σ sõltub, kui lai on jaotuskõver.

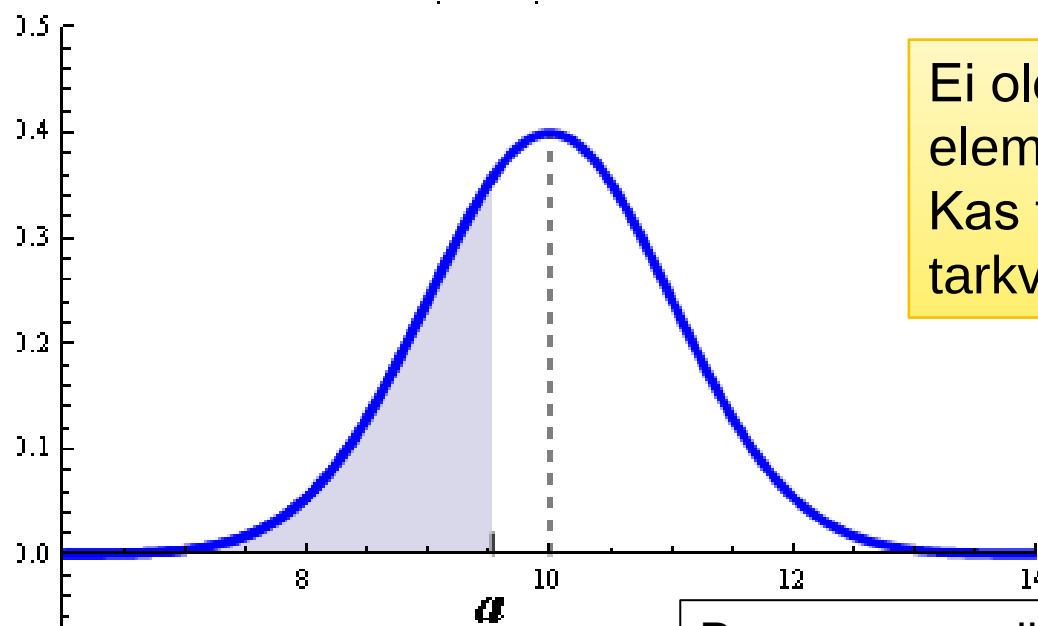


Demo: normaaljaotuse sõltuvus keskväärtusest ja standardhälbest

NORMAALJAOTUSE JAOTUSFUNKTSIOON

Tõenäosus, et normaaljaotusele alluva juhusliku suuruse X väärtus on väiksem kui mingi reaalarv a

$$F(a) = P(X < a) = \int_{-\infty}^a f(x) dx = \int_{-\infty}^a \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx$$



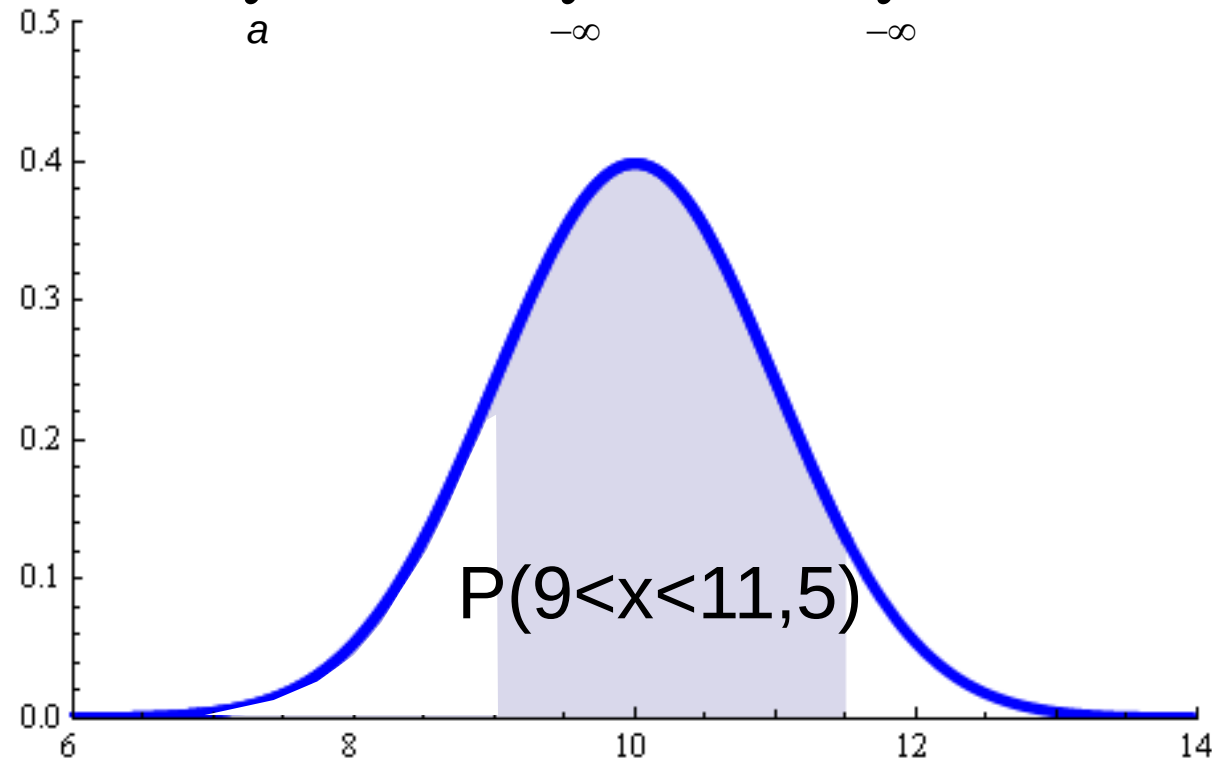
Ei ole võimalik leida
elementaarfunktsioonide abil.
Kas tabelitest või arvutis
tarkvara abil.

Demo: normaaljaotus ja tõenäosus

VAHEMIKKU JÄÄMISE TÕENÄOSUS

Vahemik (a, b)

$$P(a < x < b) = \int_a^b f(x)dx = \int_{-\infty}^b f(x)dx - \int_{-\infty}^a f(x)dx = F(b) - F(a)$$



NÄIDE: KÕRGEMASSE KVALITEEDIKLASSI KUULUVATE TOODETE OSAKAAL

Toote läbimõõt peab olema 5 mm.

Toode kuulub kõrgemasse klassi, kui tegelikud mõõtmed ei erine nominaalmõõtmest rohkem kui 10%.

$$0,1 \cdot 5 = 0,5 \text{ mm}$$

Kõrgem klass, kui $d = 4,5$ kuni $5,5$ mm

Toodete mõõtmed alluvad normaaljaotusele, standardhälve on 0,8 mm.

Kui suur on oodatav kõrgemasse klassi kuuluvate toodete arv päevas, kui päevas toodetakse 40 tk seda toodet?

Leiame tõenäosuse, et mõõtmed on 4,5 ja 5,5 vahel

$$F(5,5) - F(4,5) = 0,734 - 0,266 = 0,468$$

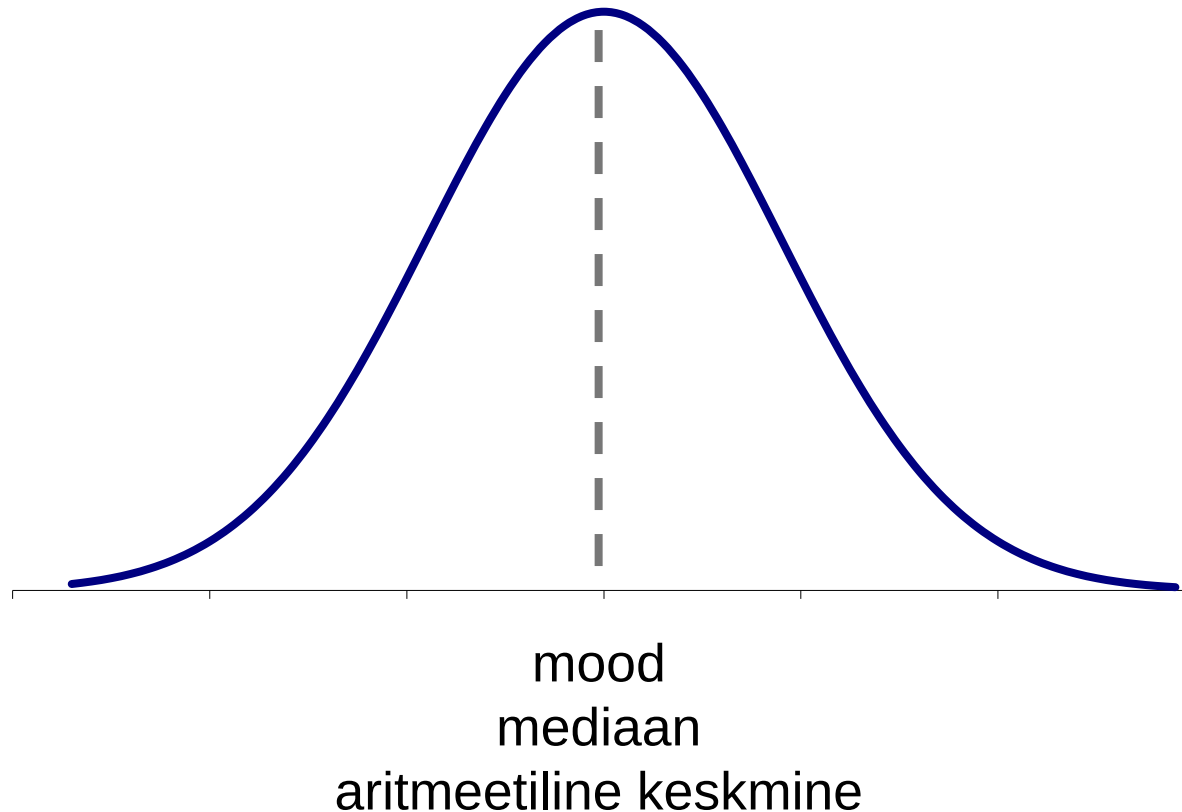
Normaaljaotusest

$$40 \cdot 0,468 \approx 19$$

40 tootest kuulub kõrgemasse kvaliteediklassi ligikaudu 19 toodet.

NORMAALJAOTUSE KESKMISED

Normaaljaotuse korral
aritmeetiline keskmine = mood = mediaan



NÄIDE: NORMAALJAOTUSELE ALLUV NÕUDLUS

Kingapaari hind hooajal 50 eurot. Peale hooaega allahinnatud, hind 25 eurot.

Omahind 35 eurot.

Iga hooajal müüdud kingapaar annab lisatulu $MR = 15$ eurot.

Iga müümata jäänud kingapaar: lisakulu $MC = 10$ eurot.

Müük hooaja jooksul: keskväärtus 400 paari, standardhälve 142 paari, allub normaaljaotusele.

Kui palju kingi tellida?

Tellimuskogus Q , nõudlus D , lisatulu MR , lisakulu MC

Demo: tellimine

Arvutuskäik:

$$E[MC] = E[MR]$$

$$P(D < Q)MC = P(D > Q)MR$$

$$P(D > Q) = 1 - P(D < Q)$$

$$P(D < Q)MC = MR - P(D < Q)MR$$

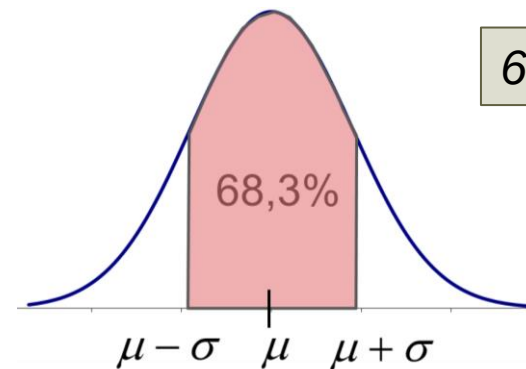
$$P(D < Q)(MC + MR) = MR \quad P(D < Q) = F(Q) = \frac{MR}{MR + MC} = \frac{15}{15 + 10} = 0,6$$

Normaaljaotusest $N(400, 142)$, kui $F(Q) = 0,6$, siis $Q = 436$ paari.

STANDARDHÄLBE KORDSED VAHEMIKUD

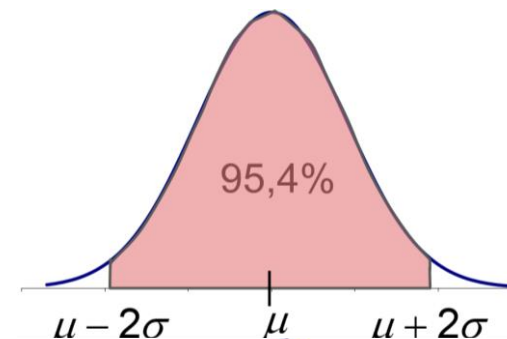
Normaaljaotuse korral jääb

vahemikku $(\mu - \sigma; \mu + \sigma)$ 68,3%

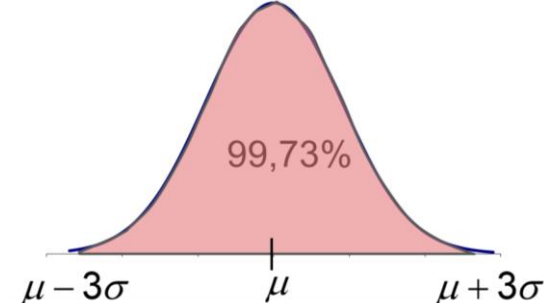


68-95-99.7% Rule

vahemikku $(\mu - 2\sigma; \mu + 2\sigma)$ 95,4%



vahemikku $(\mu - 3\sigma; \mu + 3\sigma)$ 99,7%



kõikidest väärtustest.

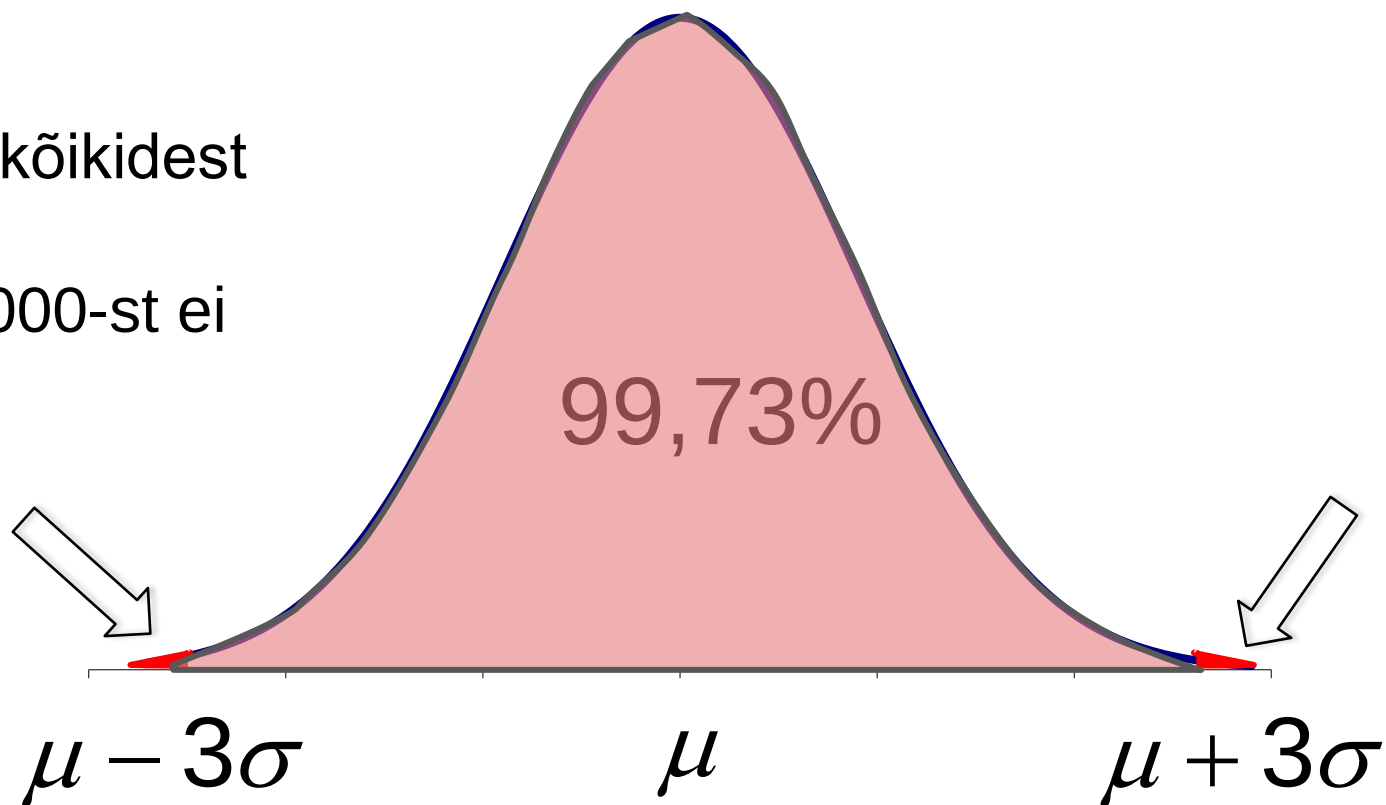
Demo: standardhälbe kordne vahemik normaaljaotuse korral

VAHEMIKUST VÄLJAS

Kui suur osa jääb välja vahemikust $\mu \pm 3\sigma$?

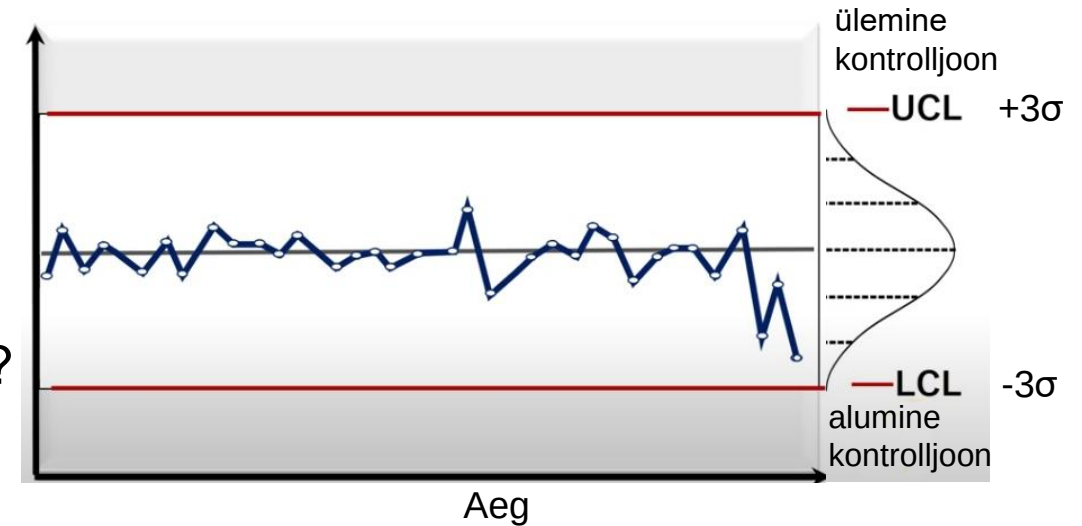
Välja jääb 0,27% kõikidest väärtustest.

Ligikaudu kolm 1000-st ei vasta nõuetele.

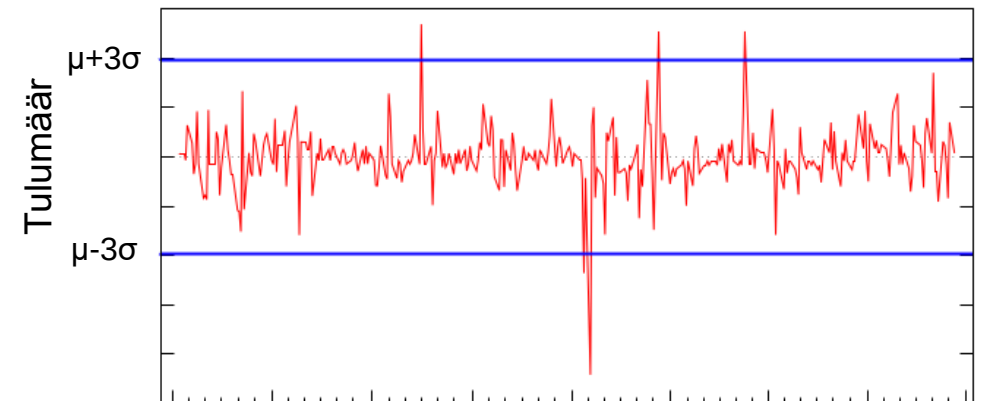


3 SIGMA REEGEL

- Kvaliteedikontroll, protsessiohje
 - Tootmises: toodete parameetrid varieeruvad juhuslikel põhjustel => alluvad normaaljaotusele.
 - Kõikumise lubatud piirid lähtuvad 3 sigma reeglist.
 - Kas toodete parameetrid jäävad lubatud piiridesse? Kui jah, siis on protsess kontrolli all.



- Finantsturgude analüüs
 - Efektiivse turu hüpotees: tulumäärade jaotus allub normaaljaotusele.
 - Tulumäär erineb keskmisest rohkem kui 3 sigma võrra => ebaharilik sündmus.



NÄIDE: PATAREIDE MÕÕTMETE VARIEERUVUS

Duracelli AAA (LR03) patarei MX2400 tehniline kirjeldus:
patarei pikkus on vahemikus 43,5 kuni 44,5 mm.

$$l = 44 \pm 0,5 \text{ mm}$$

Kasutatakse 3σ suurust poollaiust.

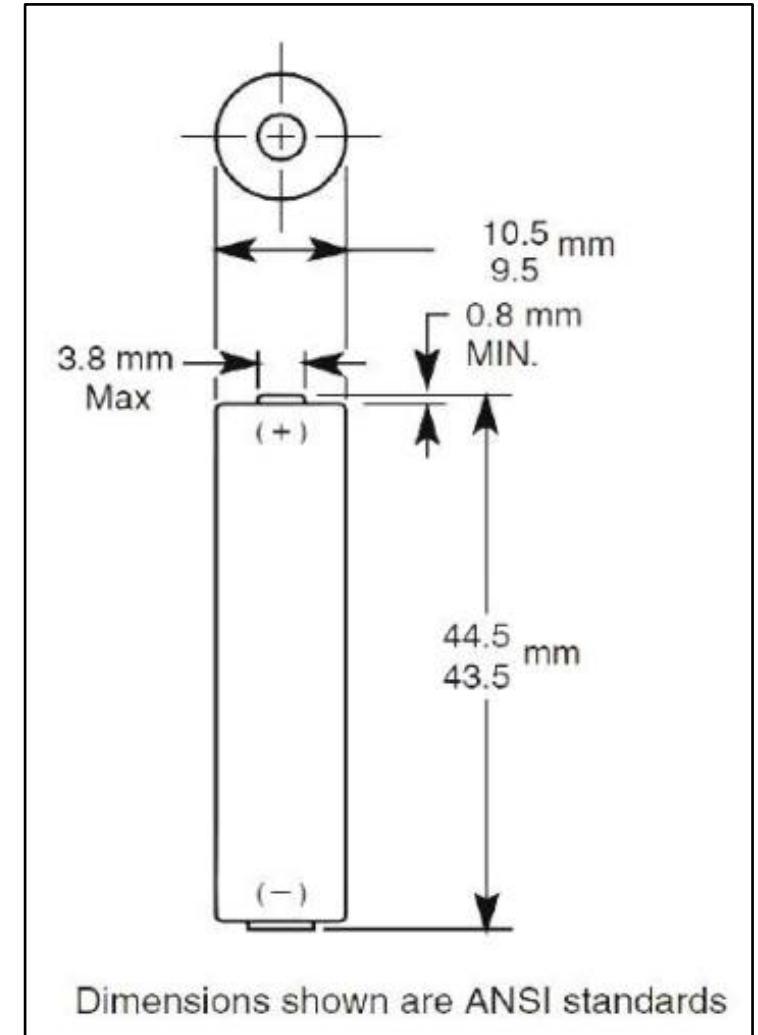
Kui garanteerida, et 99,7% patareidest jääb sellesse vahemikku, siis milline peaks olema standardhälve?

$$99,7\% \rightarrow 3\sigma$$

$$3\sigma = 0,5 \text{ mm}$$

Standardhälve

$$\sigma = \frac{0,5}{3} \approx 0,167 \text{ mm}$$



<https://www.duracell.com/wp-content/uploads/2020/02/AAA-Duracell-Ultra.pdf>

KVALITEEDIJUHTIMISE PRINTSIIP: SIX SIGMA

Võeti kasutusele Motorola poolt 1970. aastal.

Tootmises: varieerumine võib olla piirides $\pm 6\sigma$.

Vahemikust väljalangemise tõenäosus 0,00034%.

Defektimäär 3 miljoni kohta.

Kasutatakse ka mujal kui tootmises: teenindus, müük.

Nt maksimaalselt 3 kaebust miljoni teenindusjuhu kohta.

Six Sigma on

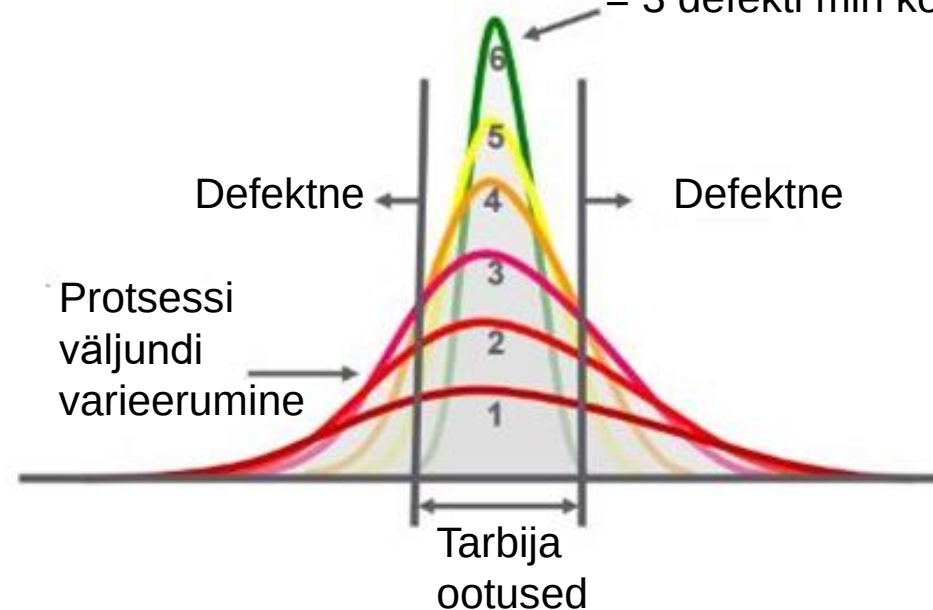
- juhtimisfilosoofia;
- projektijuhtimise ja probleemilahendamise metoodika;
- protsessi mõõtmise metoodika;
- ärikultuur.

Läbivaks jooneks on erinevate statistiliste tööriistade rakendamine:

- ärialaste probleemide tõlkimine statistilisteks probleemideks;
- nende lahendamine;
- tulemuste tõlkimine tagasi äritegevuses kasutatavaks lahenduseks.

6 σ

Protsessi väljund 6σ
= 3 defekti mln kohta



KVALITEEDIJUHTIMISE PRINTSIIP: SIX SIGMA

6σ

Mõned ettevõtted, kes on täitnud Six Sigma sertifikaadi.



TOSHIBA

