VALIK-UURINGUD I

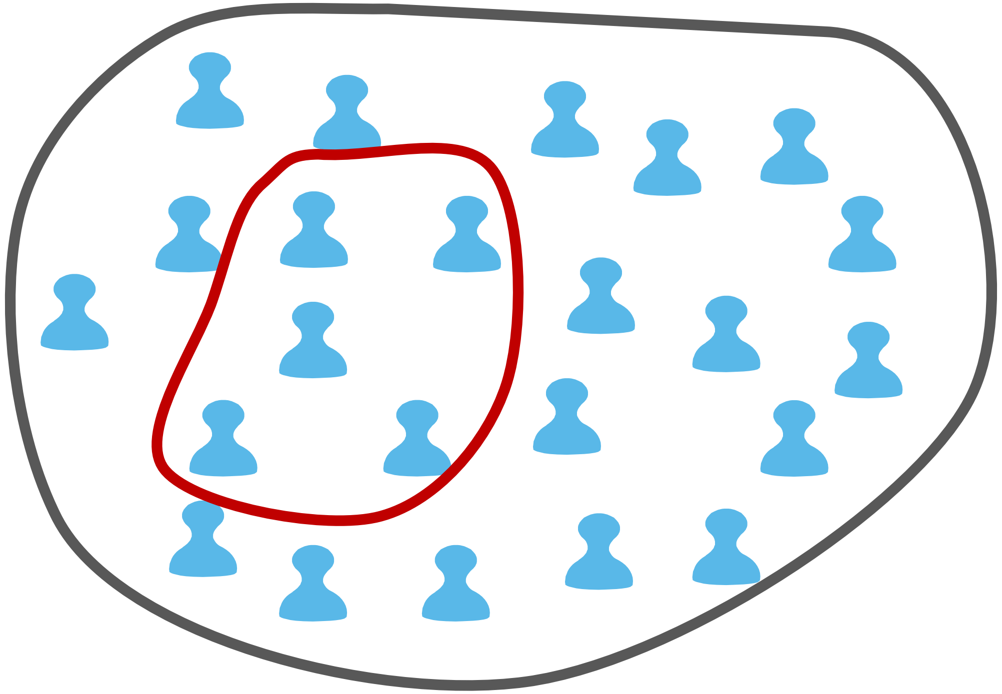
LOENGU TEEMAD

- Kogum, valim ja valikumeetodid
- Hinnangud
  - Punkthinnang
  - Vahemikhinnang
- Keskväärtuse usalduspiirid
  - Keskväärtuse, dispersiooni ja standardhälbe punkthinnangud
  - Valimi keskmise valimjaotus
  - Keskväärtuse usalduspiirid suure valimi korral

VALIM JA KOGUM

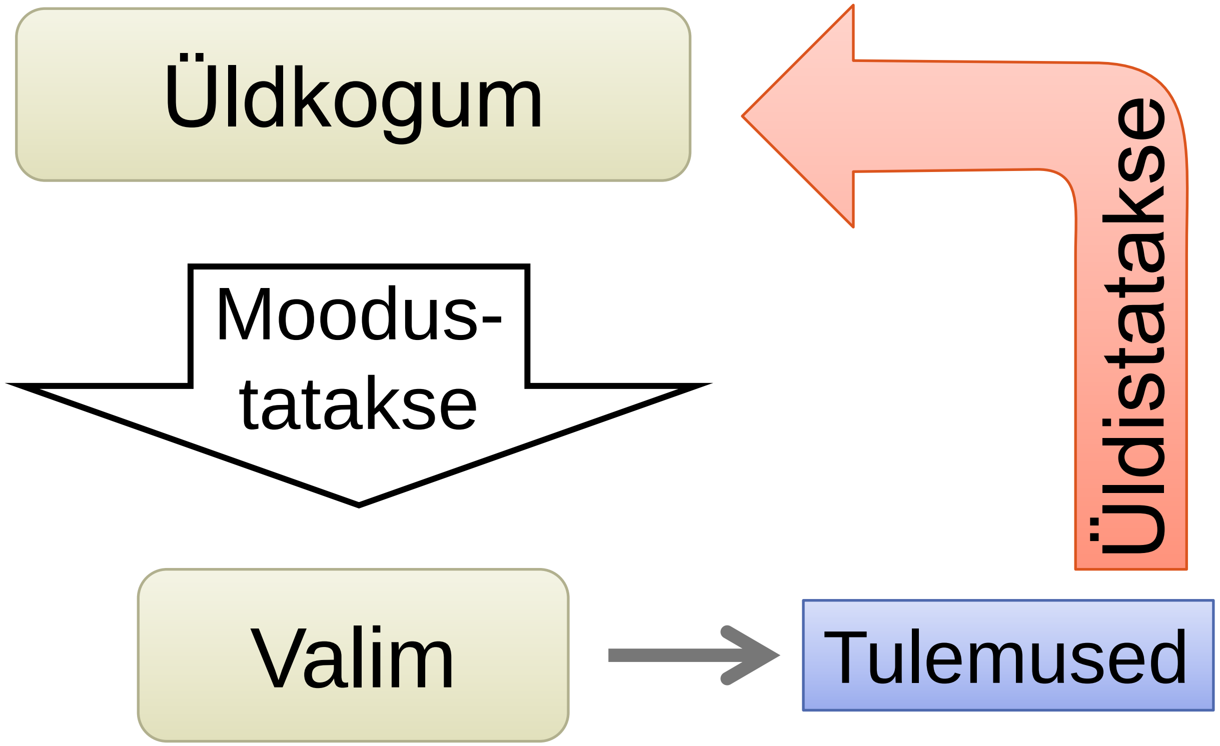
Statistiline uuring võib olla:

- **kõikne** – uuritakse läbi terve üldkogum;
- **valikuline** – uuritakse läbi üldkogumit esindav osa: valim.



VALIKUURINGU EESMÄRK

Valikuuringu eesmärgiks on valimi põhjal järelduste tegemine üldkogumi kohta.



VALIKUURINGUTE KASUTAMISE PÕHJUSED

- Väiksem maksumus.
- Suurem kiirus.
- Suurem paindlikkus.
- Laiem rakendatavus:
  - spetsiaalne aparatuur;
  - spetsiaalselt ettevalmistatud töötajad.
- Suurem täpsus andmete kogumisel:
  - suurema kvalifikatsiooniga tööjõud;
  - võimalik paremini kontrollida töötlemisvigu.
- Mõnikord võib objekti testimine rikkuda objekti.

Video: The Original Mattress Factory Test center

NÄIDE: UURINGU KULUD

- Kõikne statistika: rahvaloendus (2011)

Ca 65% e-loendusel, ülejäänute juurde küsitlejad.

Värvati 2000 inimest.

Planeeritud kulud 18,9 mln eurot.

- Valikuuring TEAN ja OSKAN (2011)

Valim: 13 000 inimest.

Intervjueerijate arv 120.

Planeeritud kulud 1,6 mln eurot.

LOEND

Objektide võtmine valimisse toimub loendi abil.

Loend (freim) on vahend pääsemiseks üldkogumi objektide juurde.

Näited:

- elanike register;
- äriregister;
- klientide andmebaas.

Näiteks AS Emor. Kodudes toimuvad küsitlused, loendiks on aadressibaasis olevad aadressid.

VASTAMISMÄÄR

Vastamismäär on vastanute osakaal valimis.

1607 uuringu põhjal keskmine vastamismäär

- isikuküsitluste korral 52,7%
- organisatsioonide korral 35,7%

Eesti ettevõtete aastastatistika üldkogum, valim ja vastanud

Aasta IV kvartal	Kogum	Valim	Vastanud	Valimi osatähtsus kogumis, %	Vastamis-määr, %
2016	88 241	10 361	7166	11,7	69,2
2020	108 131	10 993	7166	10,1	69,7
2024	130 742	11 224	8042	8,6	71,7

Allikas: Eesti Statistikaamet, tabel EM026: Ettevõtete kvartalistatistika kvaliteedinäitajad

VALIKUMEETODID

Tõenäosuslikud valikumeetodid

- Iga objekti korral on teada selle valimisse kaasamise tõenäosus.
- Võimalik hinnata valimi põhjal tehtud hinnangute **usaldusvahemikku** ja teame **tõenäosust**, et meid huvitava parameetri väärtus kogumis langeb sellesse vahemikku.

Empiirilise valiku korral ei ole objektide valimisse sattumise tõenäosused teada.

JUHUSLIKU VALIKU MEETODID

Tõmbeviisi valik

- Ühe juhusliku katse tulemus otsustab, millised üldkogumi objektid valimisse võetakse.
- Kasutatakse näiteks juhuarvude generaatorit.
- Valimi maht ette antud.

Loeteluviisi valik

- Korduvad juhuslikud katsed: üldkogumi iga objekti korral sooritatakse katse, mille tulemusel otsustatakse, kas see objekt võetakse valimisse või mitte.
- Valimi maht pole täpselt ette teada.

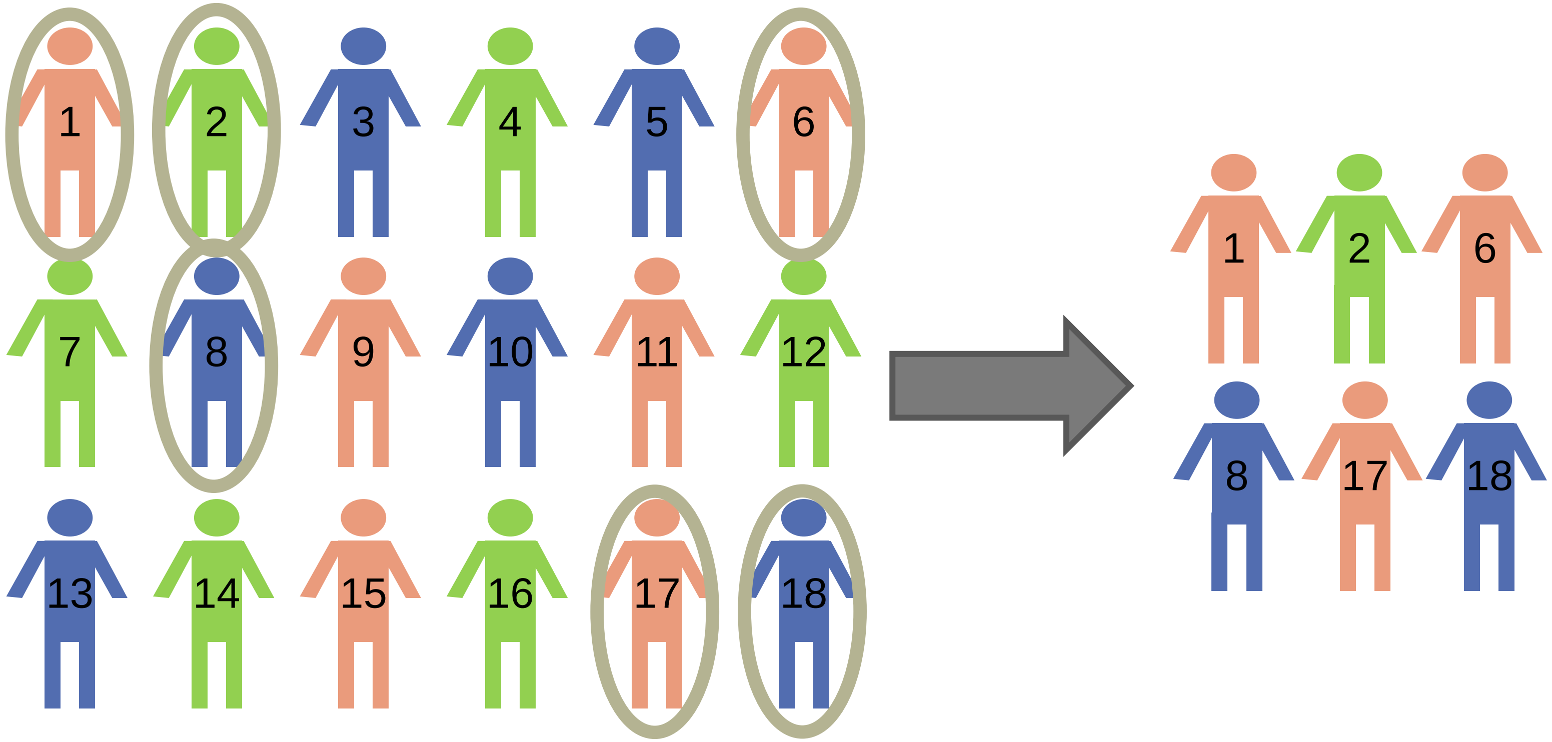
Demo: valikuviisid

TÕENÄOSUSLIKUD VALIKUMEETODID

- Lihtne juhuvalik
- Süstemaatiline valik
- Kihtvalik
- Klastervalik
- Kaheastmeline valik

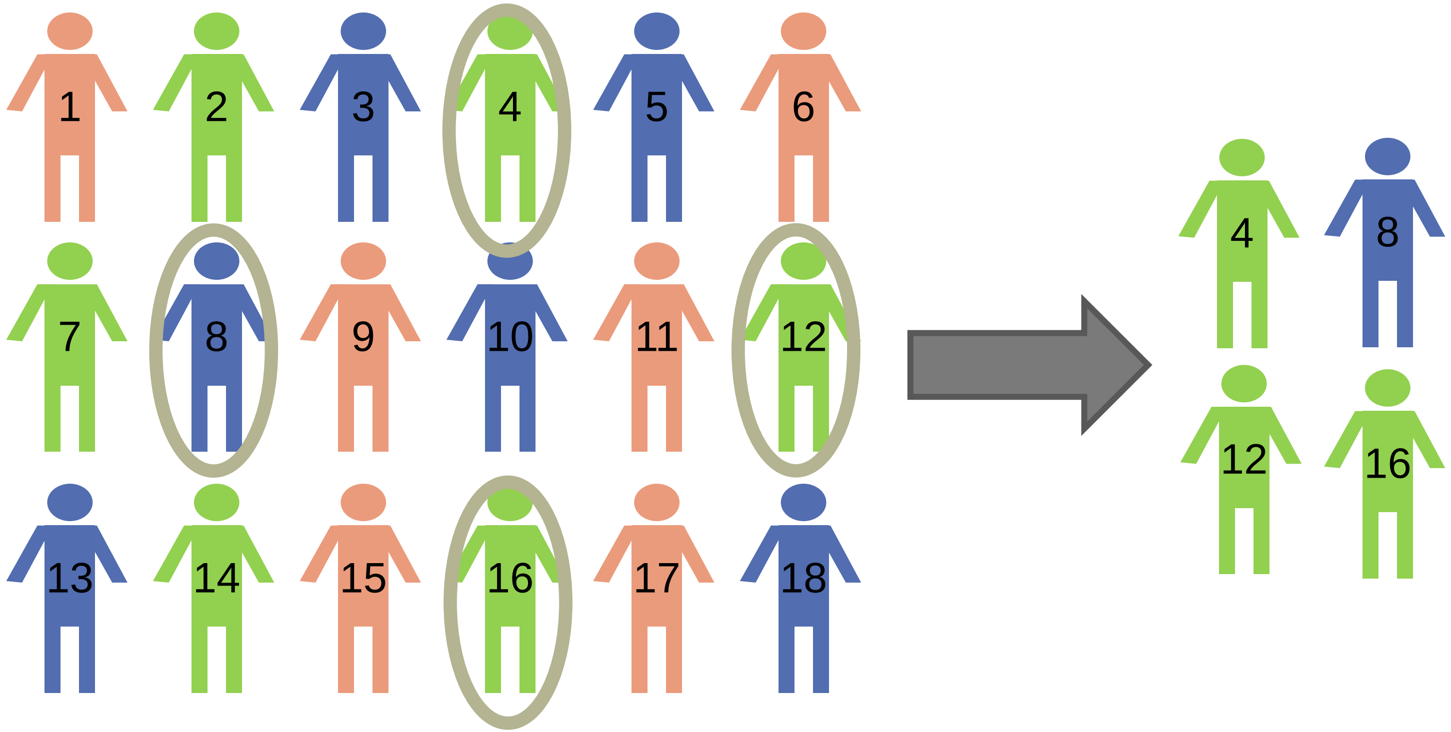
LIHTNE JUHUVALIK

- Kõigil objektidel ühesugune valimisse sattumise tõenäosus.
- Ei taga piisavat representatiivsust.



SÜSTEMAATILINE VALIK

- Objektide valik loendist toimub fikseeritud sammuga, mis määratakse juhuslikult.
- Ei taga piisavat representatiivsust.



NÄIDE: SÜSTEMAATILINE VALIK

Uuring „Vabatahtlikus tegevuses osalemine Eestis 2013“.

Poliitikauuringute keskus Praxis.

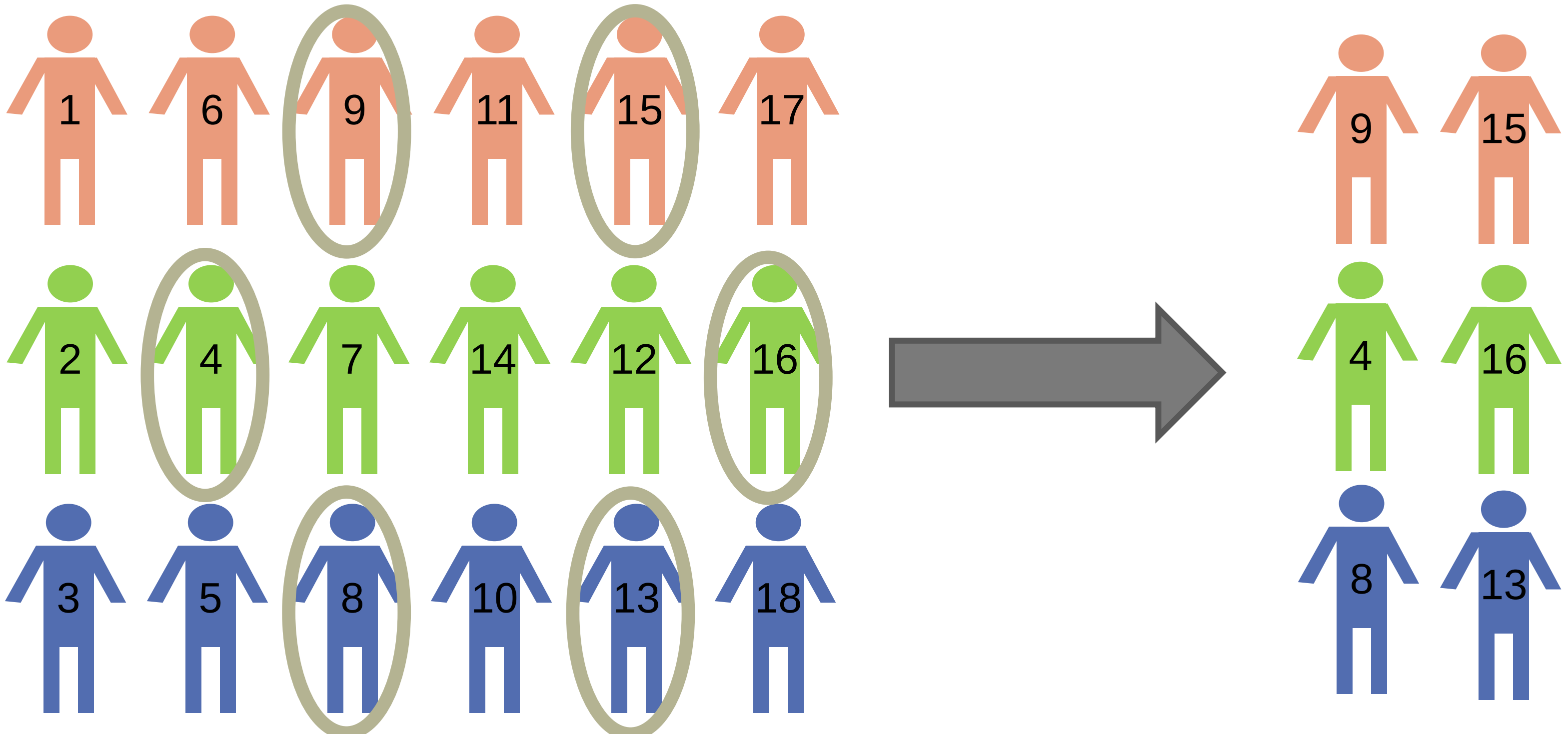
Valimi suurus 1000 inimest.

Aadressi valikul rakendati lähte-aadressi meetodit, mille puhul antakse igale küsitlejale ette juhuslikult valitud aadress esimese intervjuu läbiviimiseks.

Edasi liiguti kindla sammu alusel – iga kolmas korter või iga teine eramaja, et tagada valikusse sattunud elupaikade juhuslikkus.

KIHTVALIK

- Üldkogum jaotatakse tausttunnuste järgi kihtideks.
- Igas kihis rakendatakse mingit tõenäosuslikku valikumeetodit.



NÄIDE: KIHTVALIK

Eesti ühiskonna lõimumismonitooring 2023. Kantar Emor, Kultuuriministeerium.

Kihtideks jaotamine taustunnuste alusel:

Taustunnus	Kihtide arv
Sugu	2
Vanus	4
Rahvus	2
Asulatüüp	3
Haridustase	3

Aluseks vastavad proportsioonid Eesti elanikkond 15-aastase ja vanemad.

See tagab valimi esinduslikkuse.

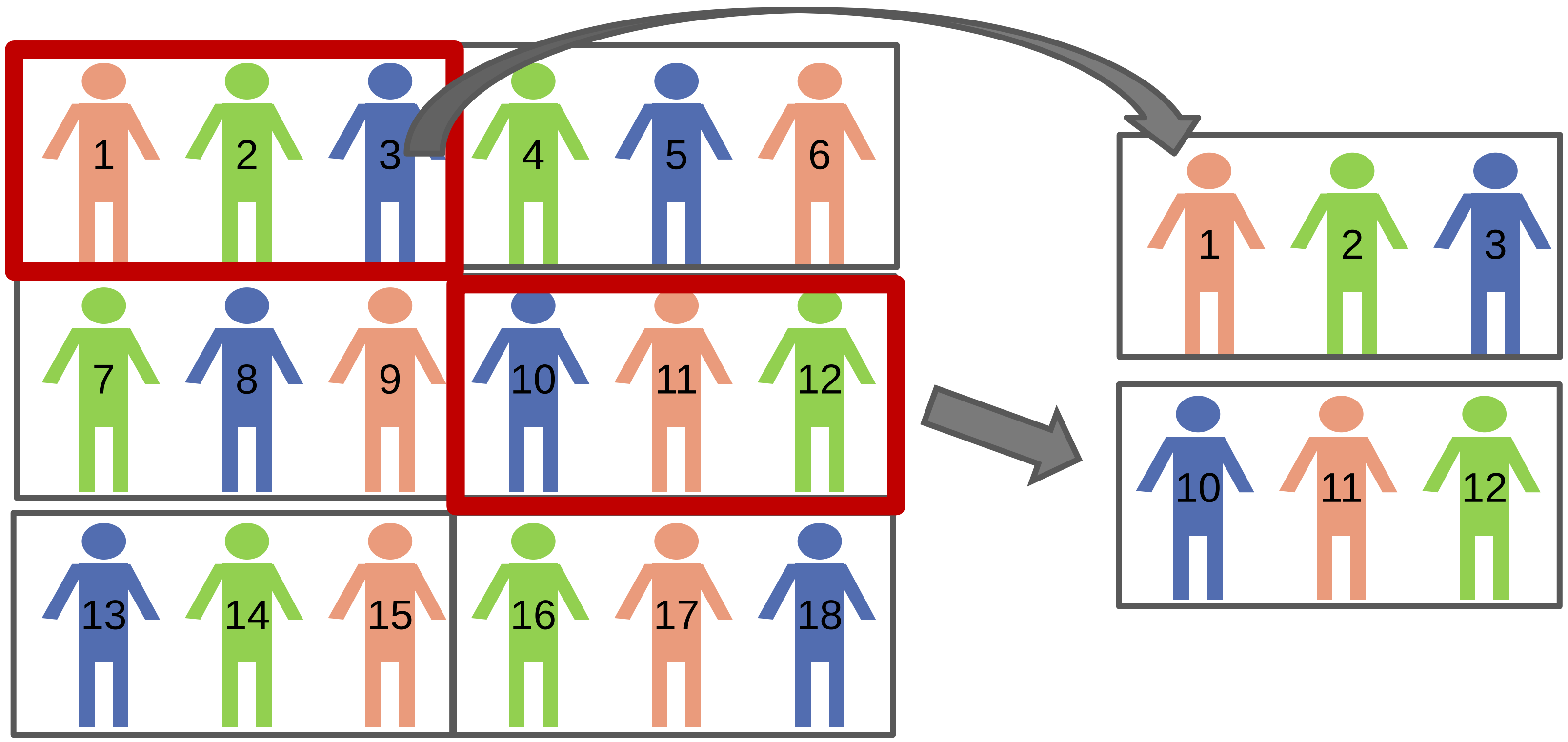
16. Püsielanikkonna uuringu metoodika

16.1 PLANEERITUD VALIM

ESA 1.1.22 15+ elanikkond			Kokku valim
SUGU	mees	47%	**656**
	naine	53%	**744**
VANUS	15-29a	17%	**244**
	30-44a	26%	**368**
	45-59a	24%	**330**
	60+a	33%	**458**
RAHVUS	eestlane	60%	**837**
	muu	40%	**563**
ASULATÜÜP	Tallinn	35%	**494**
	muu linn	34%	**473**
	maa-asula	31%	**433**
HARIDUS*	I tase	20%	**279**
	II tase	43%	**601**
	III tase	37%	**520**
KOKKU		100%	**1400**

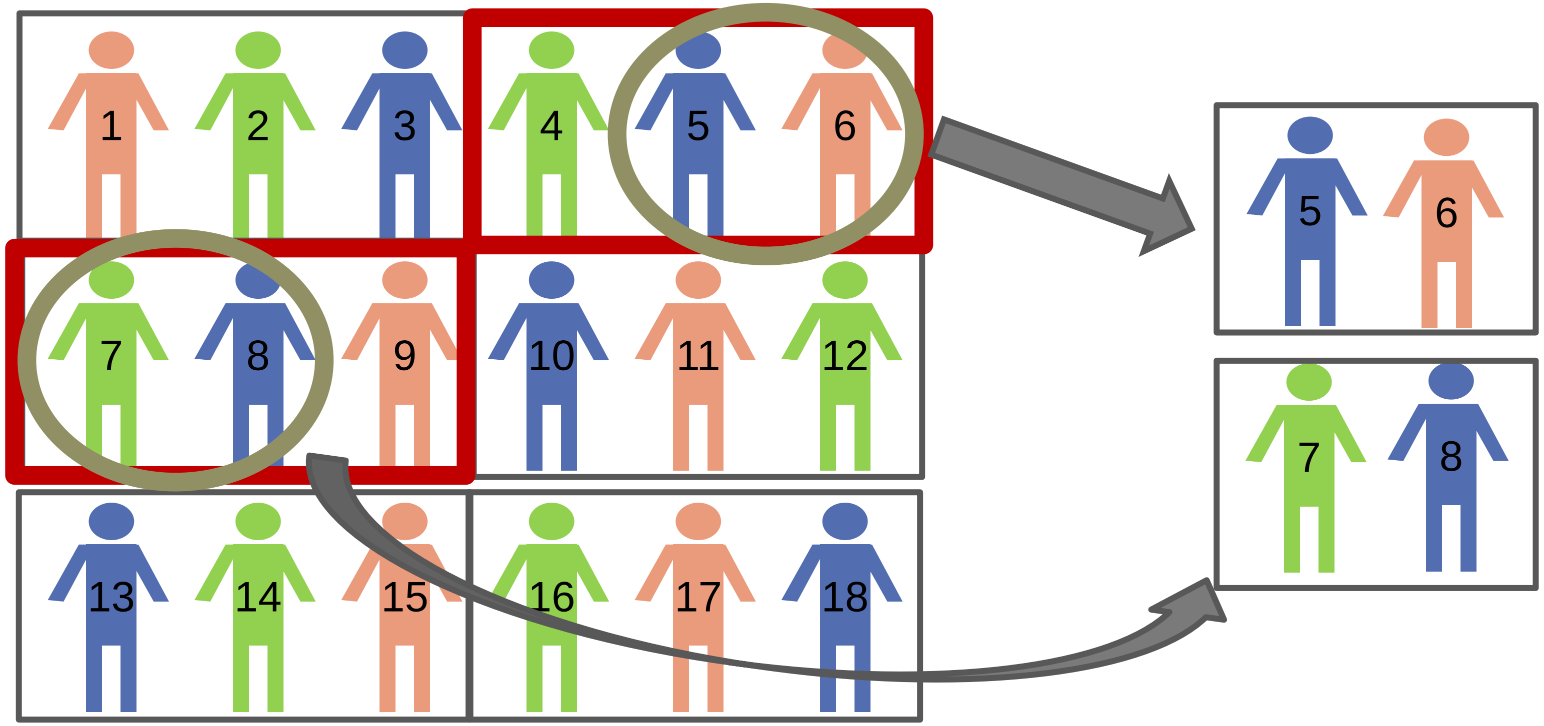
KLASTERVALIK

- Üldkogum koosneb objektigruppidest ehk klastritest (üldhariduskoolid, vallad).
- Toimub juhuslik klastrite valik ning iga klastri parameetrid leitakse selle klastri kõigi objektide põhjal.



KAHEASTMELINE VALIK

- Esimesel astmel klastrite juhuslik valik.
- Teisel astmel klastritest objektide juhuslik valik.



NÄIDE: NOORSOOUURING KISS1999

Uuriti noorte seksuaalset küpsemist.

Loend: eesti ja vene õppekeelega üldhariduskoolide 9. klasside nimekirjad

Kihtvalim: eesti ja vene koolid, tausttunnus emakeel

Kaheastmeline valik:

1. Klastervalik õpilaste arvu järgi: väikesed, keskmised, suured klassid.

2. Juhuvalik klastrites.

EMPIIRILINE VALIK I

Empiirilise valiku korral ei ole objektide valimisse sattumise tõenäosused teada.

- Kvootide meetod

Ette antakse valimi struktuur, st antakse ette kvoodid, kui palju objekte tuleb vaadeldavatest tunnusrühmadest valida. Kvoodid määratakse vastavalt üldkogumi struktuurile.

- Tasakaalustatud valik

Analoogne kvootide meetodiga, kuid üldkogumi ja valimi võrdlemiseks ei kasutata mitte tausttunnuse väärtuse sagedusi vaid näiteks keskmisi.

EMPIIRILINE VALIK II

- Ekspertvalik ehk subjektiivne valik

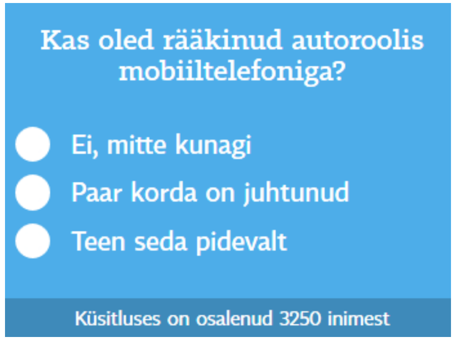
Näiteks nimetab ekspert kümme tänavat, mis on tema arvates uuritava linna jaoks tüüpilised ning andmed korjatakse nendelt tänavatelt.

- Sobivusvalim

Valitakse välja objektid, mis mingil põhjusel on sobivad. Näiteks küsitletakse inimesi, keda isiklikult tuntakse.

- Spontaanne valim

Kui inimesed ise vabatahtlikult otsustavad küsitluses osaleda.



www.postimees.ee

KUI SUUR PEAB OLEMA VALIM?

Valimi suurus sõltub tunnuse varieerumisest üldkogumis.

Näiteks soovime teada, milline on keskmine palk maakonnas A, kus elab 10 000 inimest.
Oletame, et kõik saavad seal ühesugust palka. Mitut inimest peame küsitlema, ehk milline on valimi maht?

- Kui tunnus ei varieeru, siis valimi maht n=1.
- Suurema varieerumise korral suurem valim.
- Täpsemalt vt õpikust ptk „Valimi mahu planeerimine“

NÄITED ERINEVATEST VALIMITEST

Uuring	Kogum	Kogumi maht	Valimi maht	Valimi osakaal kogumist
Erakondade reiting (Emor)	18 a ja vanemad	1068 tuh	1000	0,094%
Tean ja oskan (2011)	16-65 a	860 tuh	13 000	1,5%
Lõimumismonitooring 2023	15 a ja vanemad püsielanikud	1 114 tuh	1400	0,13%
Sotsiaaluuring (Eesti Statistikaamet)	Leibkonnad	560 tuh	8200	1,5%
Ettevõtted (Eesti Statistikaamet, 2024)	Eesti ettevõtted	130 tuh	11 000	8,6%

VALIMI ESINDUSLIKKUS

Esinduslik ehk representatiivne valim:

- piisavalt suur (suurus sõltub varieerumisest);
- kõigil üldkogumi objektidel võimalus sattuda valimisse;
- kajastab üldkogumi seesmist struktuuri, taustunnuste jaotumist:
  - meeste ja naiste osakaal valimis sama, mis kogumis;
  - vanuseline struktuur valimis sama, mis kogumis;
  - …..

  Olulised on need taustunnused, mis mõjutavad uuritavaid tunnuseid.

NÄIDE: VALIMISED USA-s 1936

Alates aastast 1916 oli ajakiri **Literary Digest** edukalt prognoosinud USA presidendivalimiste võitjaid.

Aastal 1936 kasutati esmakordselt telefoniküsitlust. Valimi maht ca 2,4 mln.

Prognoos valimi põhjal: Alf Landon.

Tegelikult sai presidendiks Franklin Roosevelt.

Gallupi valimi maht oli 50 tuhat, prognoos Roosevelt oli õige.

VALIKUNIHE

Valikunihe tekib siis, kui valimi moodustamisel on teatud tüüpi objektidel suurem võimalus valimisse sattuda kui teistel.

Põhjused võivad olla erinevad.

Näiteks valimis ainult need, keda antud probleem huvitab:

- küsitlused veebilehtedel;
- küsitlused telesaadetes (helistage numbril …, kui vastate "ei" ja numbril …, kui vastate "jaa").

Valimis need, kes külastavad seda veebilehte, kes vaatavad seda saadet.

NÄIDE: KÜSITLUS TALLINNAS 2012

Kas toetate tasuta ühistransporti Tallinnas?

Kogum 416 000 elanikku.

Küsitlus 19.-25. märts 2012 viidi läbi kaubanduskeskustes, linnaosade küsitluspunktides.

Valim	68 059	16%
„Jah“	51 242	75%

Veebilehel Petitsioon.ee 17. märts kuni 15. mai 2012.

Valim	374
„Jah“	7,2%

Mõlemas esines valikunihe.

Küsitluspunkti läks see, kes oli huvitatud.

Veebilehele läks see, kes oli selle idee vastu.

NÄIDE: VÕIMALIK VALIKUNIHE MÕJU-UURINGUS

Kas EASi toetuse saamine mõjutas ettevõtete müügitulu?

Valimis toetust saanud ettevõtted.

Tulemus: jah müügitulu suurenes, toetusest oli kasu.

AGA vaadeldud aastatel võis müügitulu suureneda kõikidel ettevõtetel (hea majandusaasta).

EAS: Ettevõtluse Arendamise Sihtasutus

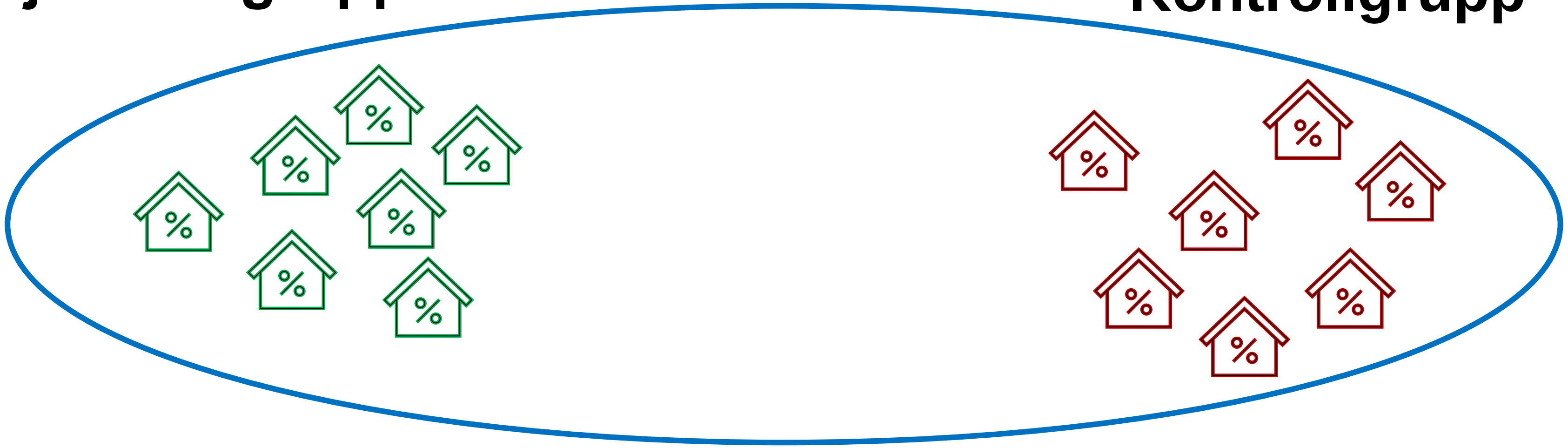
NÄIDE: VALIKUNIHKEST HOIDUMINE MÕJU-UURINGUS

Kas EASi toetuse saamine mõjutas ettevõtete müügitulu?

Toetust saanud ettevõtted.
Mõjutatud grupp

Valimis

Ettevõtted, kes ei saanud toetust (juhuvalim äriregistrist).
Kontrollgrupp



Kui mõjutatud grupil suurenes müügitulu oluliselt rohkem kui kontrollgrupil, siis oli toetusest kasu.

Vt näiteks Hartsenko, J. & Sauga, A. (2013) The role of financial support in SME and economic development in Estonia. *Business and Economic Horizons*, Vol 9, 10-22.

VALIMI SUURUS VS ESINDUSLIKKUS

Valimi suurusest on tähtsam valimi **esinduslikkus.**

Valesti koostatud suur valim annab halvema tulemuse kui õigesti koostatud väike valim.

HINNANGUD

HINNANGUD

Uuringu eesmärk: teha järeldusi üldkogumi kohta.

Näiteks: Eesti elanike sissetulekute aritmeetiline keskmine on …… eurot.

Kõikne statistika: saame selle aritmeetilise keskmise leida.

Valikuuring: saame seda aritmeetilist keskmist **hinnata.**

Valimi põhjal leitakse üldkogumi parameetrite **hinnangud.**

Kasutatakse spetsiaalseid hindamisreegleid.

PUNKTHINNANG

Demo: punkthinnang

Punkthinnang: parameetri hindamise tulemuseks on üks arv.

Hinnang leitakse valimi põhjal.
Valim on juhuvalim.

$\Downarrow$

Punkthinnang on juhuslik suurus.

HINNANGU NIHE

Hinnangu nihe on hinnangute keskväärtuse $E[\hat{a}]$ ja hinnatava parameetri tegeliku väärtuse a vahe:

$$b = E[\hat{a}] - a$$

Näiteks 5 valimit

Valim	Keskmine
1	1216
2	902
3	964
4	996
5	904

Hinnangute keskväärtus $E[\hat{a}] = 996,4$

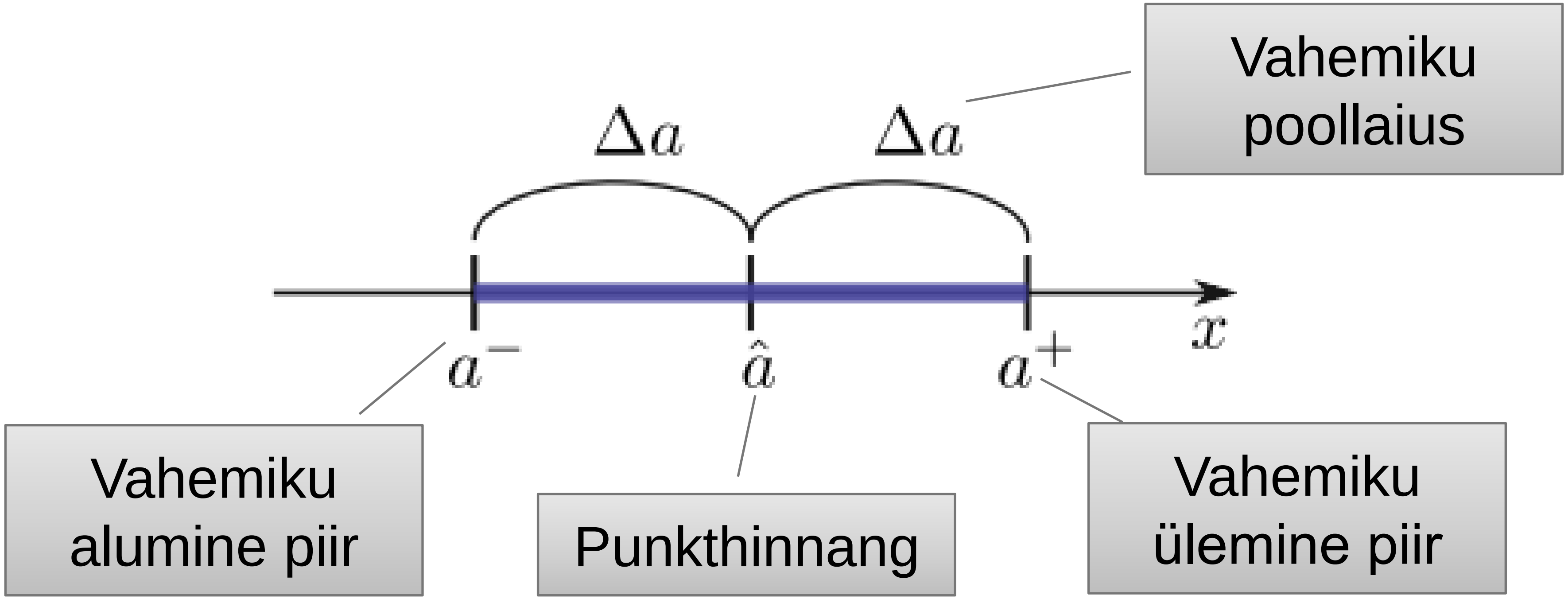
Tegelik väärtus $a = 942,8$

Hinnangu nihe $b = 53,6$

Demo: hinnangu nihe

VAHEMIKHINNANG

Vahemikhinnang on valimi põhjal määratud vahemik, mis katab parameetri tegeliku väärtuse etteantud (küllalt suure) tõenäosusega.



NÄIDE: PUNKTHINNANG JA VAHEMIKHINNANG

Riigikogu valimised olid 5. märts 2023.
Uuringufirma Norstat küsitles iganädalaselt inimeste erakondlikku eelistust.
Valimi maht 1000. Maksimaalne viga usaldatavusega 0,95 on 3,1%.

Tabelis on viimase valimi (27.02-03.03 2023) põhjal saadud hinnangud ja tegelikud valimistulemused.

Vahemikhinnang = punkthinnang ± maksimaalne viga

Erakond	Punkt-hinnang, %	Valimis-tulemused, %
Eesti Reformierakond	32,5	31,2
Eesti Konservatiivne Rahvaerakond	16,2	16,1
Eesti Keskerakond	15,4	15,3
Eesti 200	14,5	13,3
Sotsiaaldemokraatlik Erakond	7,8	9,3

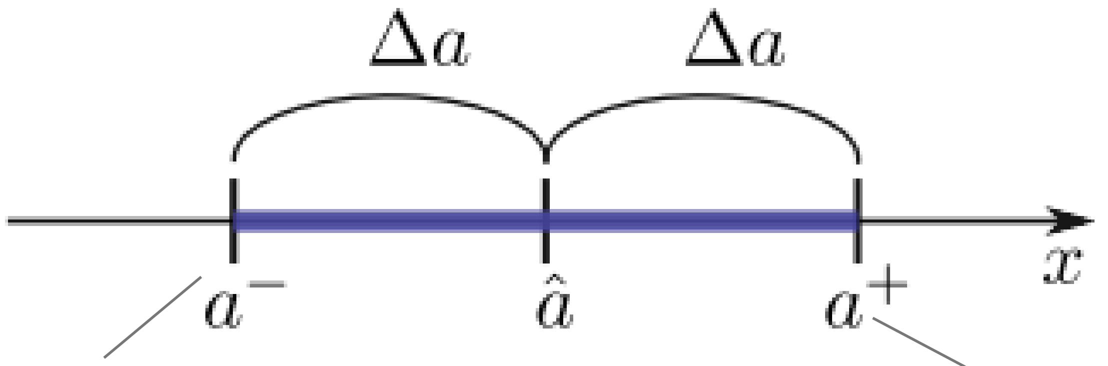
Kõikide erakondade korral langesid valimistulemused küsitluse vahemikhinnangu piiridesse.

Allikas: Erakondade reitingud, https://reitingud.ee/ Vaadatud 12.03.2023

USALDUSVAHEMIK

Parameetri *a* **usaldusvahemikuks** usaldatavusega *β* nimetatakse vahemikku, mis katab parameetri *a* väärtuse tõenäosusega *β*:

$$\beta = P(|\hat{a} - a| < \Delta a)$$



Alumine usalduspiir

Ülemine usalduspiir

KUIDAS LEIDA USALDUSVAHEMIKKU?

Sõltub, millise parameetri usaldusvahemikku soovime leida

- keskväärtuse,
- osakaalu,
- mediaani.

Sõltub valimi suurusest.

ÜLDKOGUMI KESKVÄÄRTUSE USALDUSVAHEMIKU LEIDMINE, SUUR VALIM

ÜLDKOGUMI KESKVÄÄRTUSE PUNKTHINNANG

Üldkogumi keskväärtuse μ punkthinnanguks on valimi aritmeetiline keskmine:

$$\boxed{\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i}$$

Summeerimine üle valimisse kuuluvate objektide.

n valimi maht

ÜLDKOGUMI DISPERSIOONI PUNKTHINNANG

Üldkogumi dispersiooni σ^2 punkthinnang

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

Summeerimine üle valimisse kuuluvate objektide.

See on **valimi dispersioon.**

MIKS NIMETAJAS *n* -1?

Demo: dispersiooni hinnangud

Kui kasutame valemit $$S^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n}$$

saame nihkega hinnangu. Hinnangu keskväärtus

$$E\left[S^2\right] = \frac{n-1}{n}\sigma^2$$

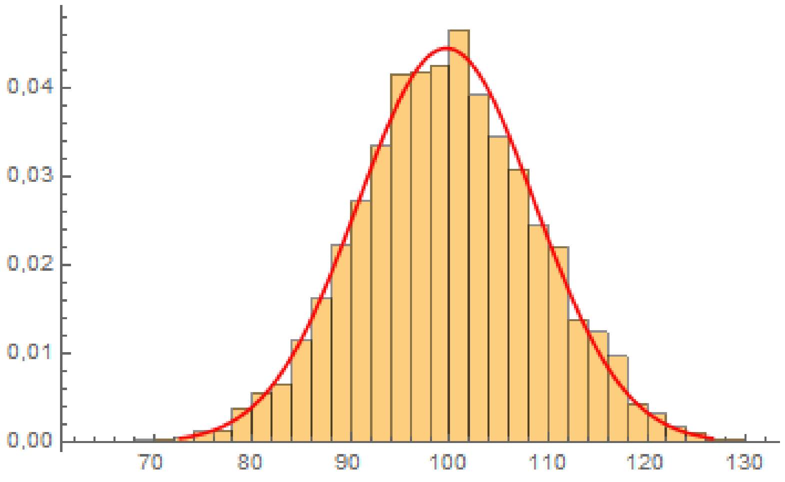
Tõestus õpikus lisa A.7

Nihketa hinnangu jaoks teisendus

$$\frac{n}{n-1}\cdot S^2 = \frac{n}{n-1}\cdot\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}$$

VALIMI KESKMISE VALIMJAOTUS

Võtame ühest ja samast kogumist palju erinevaid valimeid ja leiame iga valimi keskmise.



Eeldused

- lihtne juhuvalik;
- suured valimid.

Valimite keskmised alluvad normaaljaotusele.

Ei sõltu sellest, millisele jaotusele allub vaadeldav tunnus kogumis.

Demo: keskmiste valimjaotus

TSENTRAALNE PIIRTEOREEM

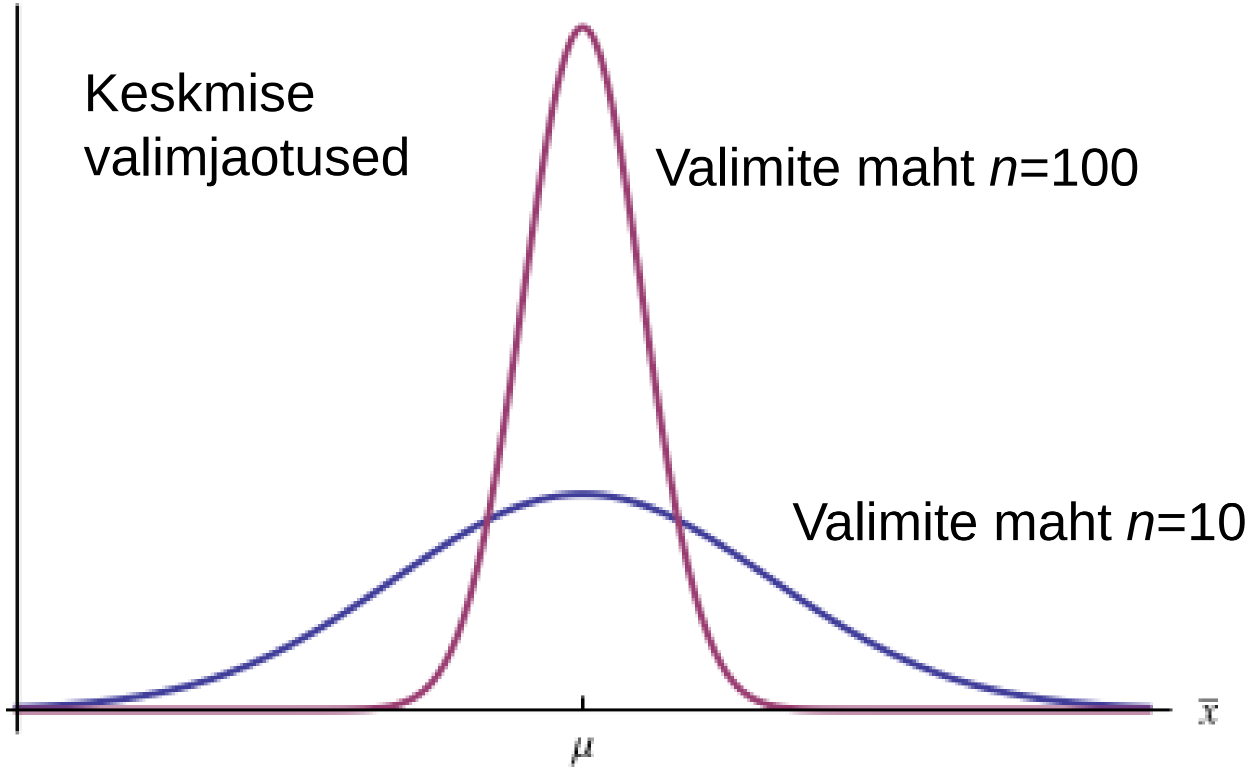
Küllalt suure valimi mahu n korral alluvad valimite keskmised $\bar{x}$ normaaljaotusele keskväärtusega μ ja standardhälbega $\frac{\sigma}{\sqrt{n}}$, kus σ on kogumi standardhälve.

$$\bar{x} \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$$

VALIMJAOTUSE STANDARDHÄLVE

σ iseloomustab üksikväärtuste hajumist.

$\dfrac{\sigma}{\sqrt{n}}$ iseloomustab valimite (maht *n*) keskmiste hajuvust, see on valimi keskmiste valimjaotuse standardhälve



VALIMITE KESKMISTE HAJUVUS

Normaaljaotuse korral jääb vahemikku

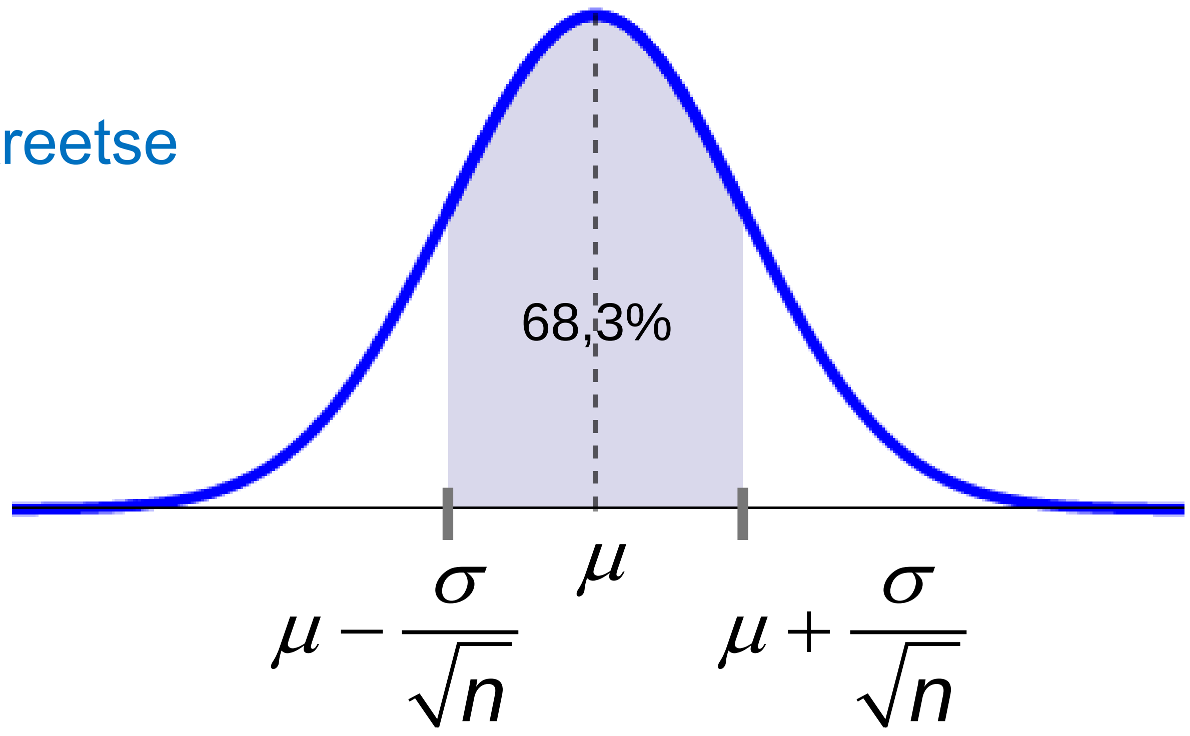
- keskväärtus ± 1 standardhälve 68,3%
- keskväärtus ± 2 standardhälvet 95,4%
- keskväärtus ± 3 standardhälvet 99,7%

kõikidest väärtustest.

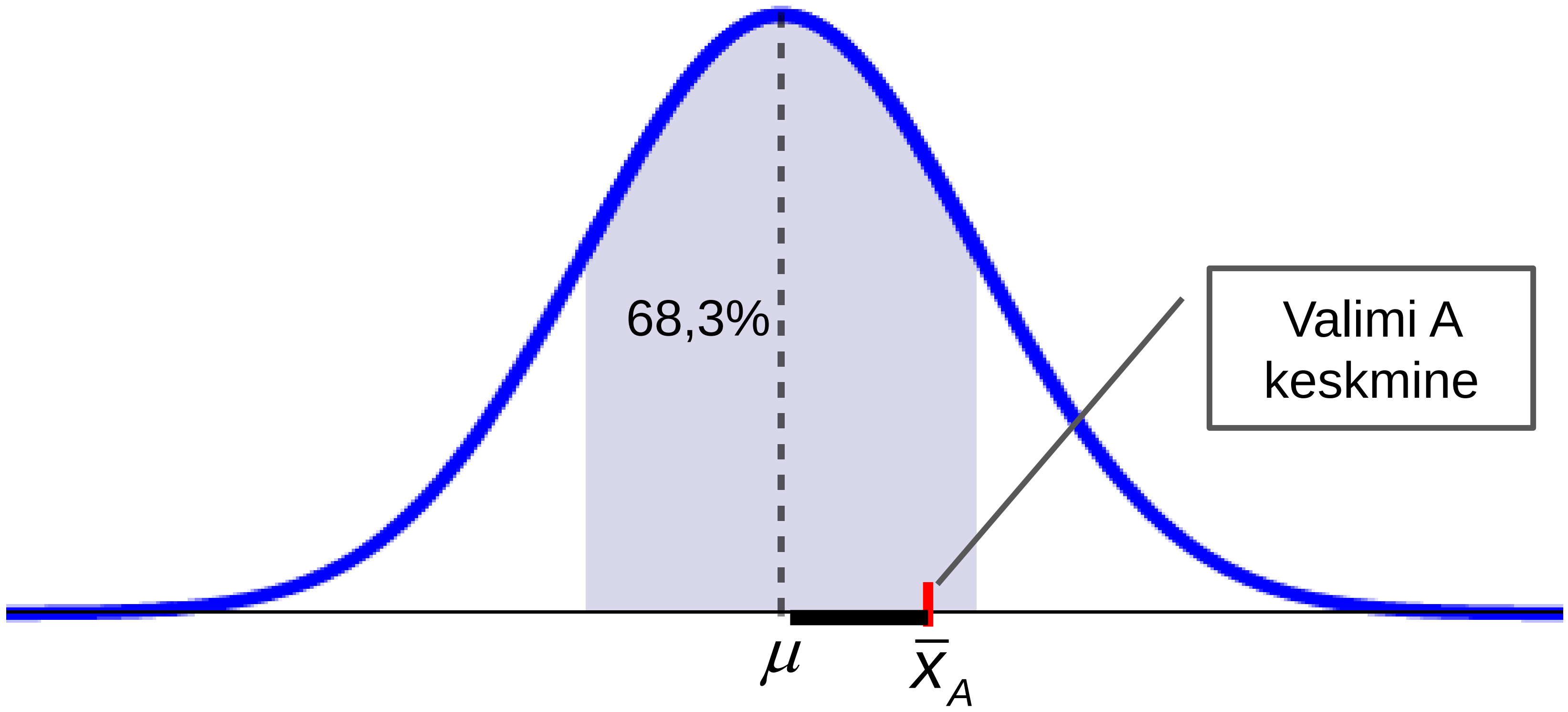
Tõenäosusega 68,3% jääb ühe konkreetse valimi keskmine vahemikku

$$\mu \pm \frac{\sigma}{\sqrt{n}}$$

Valimjaotuse standardhälve



ÜHE VALIMI KESKMINE

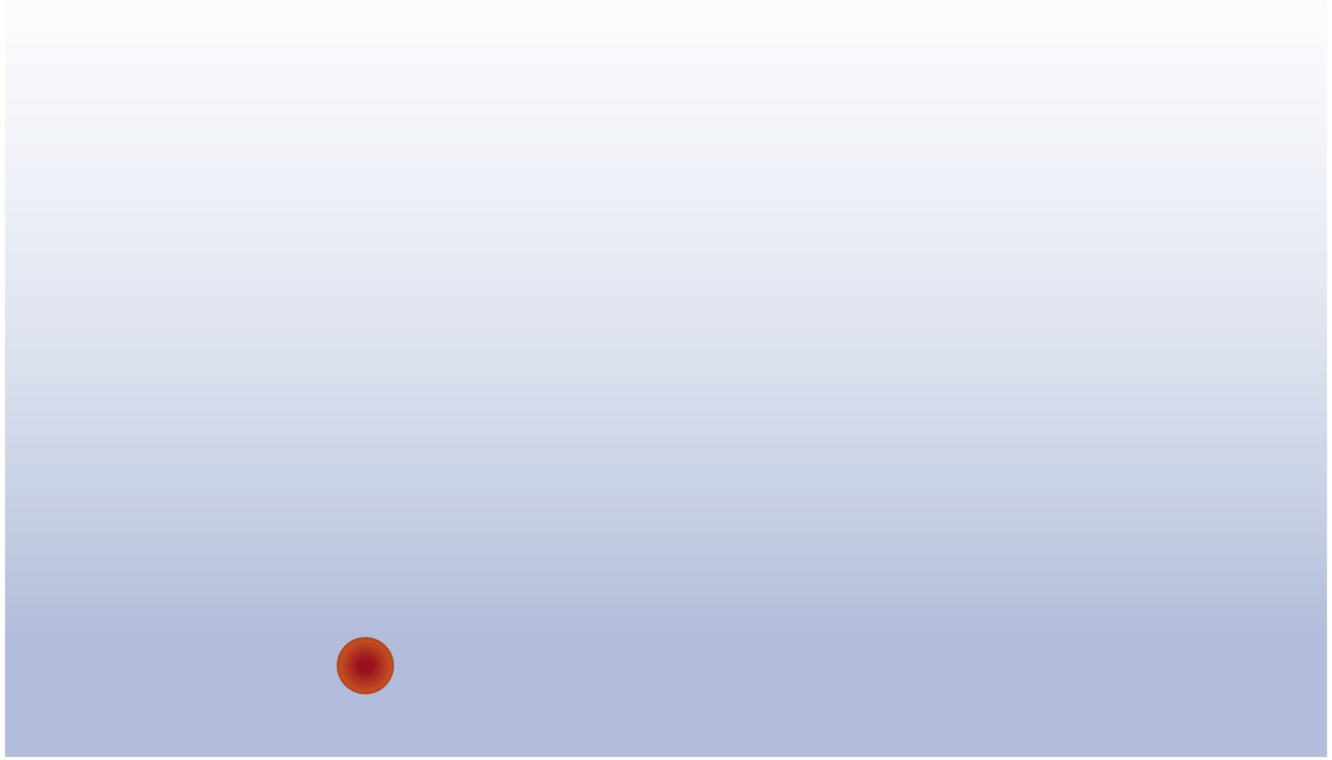


Kogumi keskväärtuse μ kaugus valimi keskmisest $\bar{x}_A - \mu$

Tõenäosusega 68,3% $\left|\bar{x}_A - \mu\right| < \frac{\sigma}{\sqrt{n}}$

Demo: usaldusvahemiku saamine

Kui me teame, kuidas maha kukkunud õunad hajuvad ümber puu,



siis leides ühe õuna, teame, kui kaugelt puud otsida.

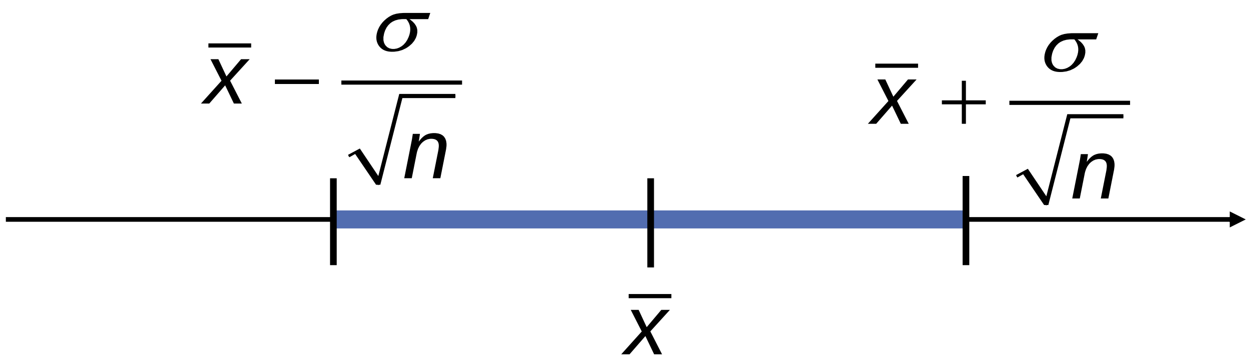
USALDUSVAHEMIK

Tõenäosusega 68,3% $$\bar{x} - \frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + \frac{\sigma}{\sqrt{n}}$$

Järelikult kogumi keskväärtuse usaldusvahemik

$$\boxed{\mu = \bar{x} \pm \frac{\sigma}{\sqrt{n}}}$$

usaldatavusega β=68,3%



TÕENÄOSUSKORDAJA

$$\mu = \bar{x} \pm \frac{\sigma}{\sqrt{n}}$$ usaldatavusega β=68,3%

$$\mu = \bar{x} \pm 2 \cdot \frac{\sigma}{\sqrt{n}}$$ usaldatavusega β=95,4%

$$\mu = \bar{x} \pm 3 \cdot \frac{\sigma}{\sqrt{n}}$$ usaldatavusega β=99,7%

$$\mu = \bar{x} \pm k_{\beta} \cdot \frac{\sigma}{\sqrt{n}}$$ k_{β} on **tõenäosuskordaja**

TÕENÄOSUSKORDAJA LEIDMINE

Demo: usaldatavus

$k_\beta = 1,\ \beta = 68{,}3\%$
$k_\beta = 2,\ \beta = 95{,}4\%$
$k_\beta = 3,\ \beta = 99{,}7\%$

k_β täpselt, β ligikaudu

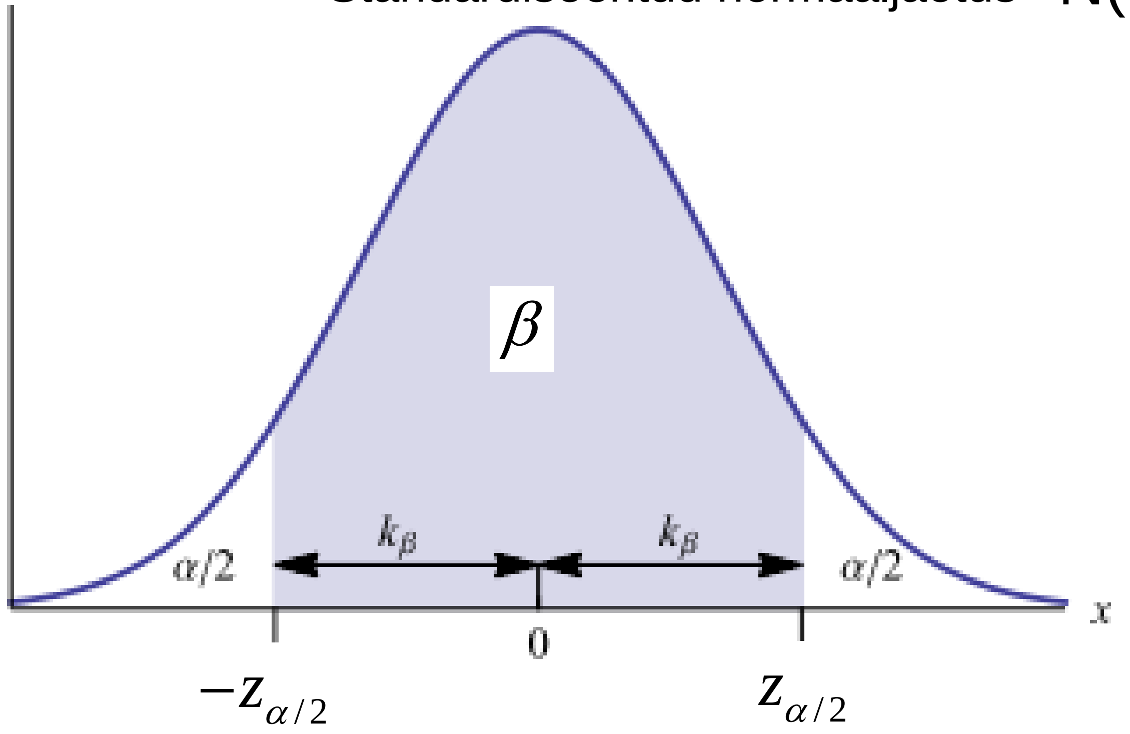
k_β on standardiseeritud normaaljaotuse täiendkvantiil

$$k_\beta = z_{\alpha/2}$$

$$\alpha = 1 - \beta$$

Aga kui võtta β täpselt?
Näiteks 90%?

Standardiseeritud normaaljaotus N(0, 1)



TÕENÄOSUSKORDAJA VÄÄRTUSED

k_β	β
• keskväärtus ± 1 standardhälve	68,3%
• keskväärtus ± 2 standardhälvet	95,4%
• keskväärtus ± 3 standardhälvet	99,7%

Täpsed väärtused ↑ Ligikaudsed väärtused ↑

Ligikaudsed väärtused ↓ Täpsed väärtused ↓

• keskväärtus ± 1,15 standardhälvet	75%
• keskväärtus ± 1,64 standardhälvet	90%
• keskväärtus ± 1,96 standardhälvet	95%

STANDARDVIGA

Praktikas pole meil üldkogumi standardhälbe tegelik väärtus σ teada ning kasutatakse selle hinnangut, valimi standardhälvet *s*.

$$\frac{\sigma}{\sqrt{n}} \quad \rightarrow \quad \boxed{se = \frac{s}{\sqrt{n}}}$$

standardviga (*standard error*)

kus valimi standardhälve

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

STANDARDHÄLVE JA STANDARDVIGA

Valimi standardhälve

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

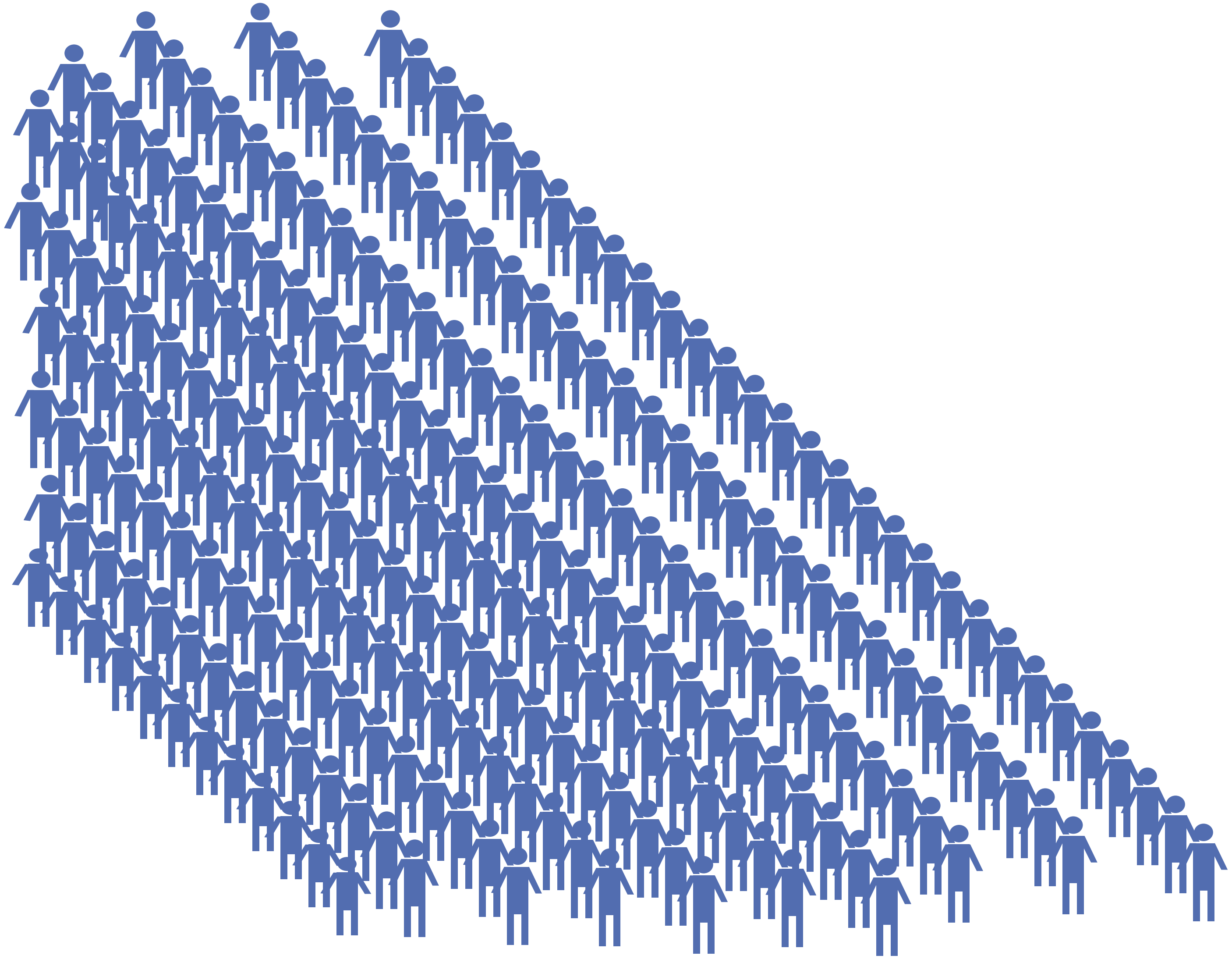
Kogumi standardhälbe hinnang.
Iseloomustab üksikute objektide hajumist.

Standardviga

$$se = \frac{s}{\sqrt{n}} = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n(n-1)}}$$

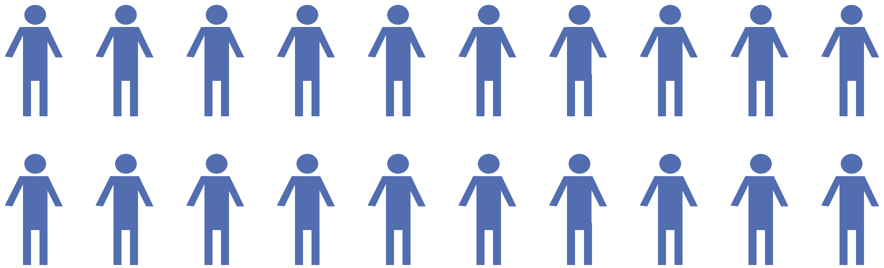
Keskväärtuse valimjaotuse standardhälbe hinnang.
Iseloomustab valimite keskmiste hajumist.

LÕPMATULT SUUR KOGUM



Valimi maht *n*=3
Kogumi maht enne *N*=200.
Valimi võtmisel väheneb
197-ni, so 1,5%.

LÕPLIK KOGUM



Valimi maht *n*=3.
Kogumi maht enne *N*=20.
Valimi võtmisel väheneb 17-ni,
so 15%.

LÕPLIK JA LÕPMATU KOGUM

Tsentraalne piirteoreem kehtib, kui kogumi maht $N \to \infty$.

Siis valimjaotuse standardhälbe hinnang $se = \frac{s}{\sqrt{n}}$

Kui kogumi maht N on lõplik ja väheneb valimi tegemisel

siis $$se = \frac{s}{\sqrt{n}}\sqrt{1 - \frac{n}{N}}$$

Kasutada siis, kui $\frac{n}{N} > 5\%$

KOGUMI KESKVÄÄRTUSE USALDUSPIIRID

Suure (n>30) valimi korral on üldkogumi keskväärtuse usalduspiirid usaldatavusega β

$$\boxed{\begin{aligned} &\bar{x} \pm \Delta x \\ &\Delta x = z_{\alpha/2} \frac{s}{\sqrt{n}} \end{aligned}}$$

$\bar{x}$ valimi keskmine

s valimi standardhälve

n valimi maht

$z_{\alpha/2}$ standardiseeritud normaaljaotuse täiendkvantiil, α =1-β

Demo: usalduspiirid

KOGUMI KESKVÄÄRTUSE USALDUSPIIRID LÕPLIKU KOGUMI KORRAL

Lõpliku kogumi mahu *N* korral

$$\bar{x} \pm \Delta x$$

$$\Delta x = z_{\alpha/2} \frac{s}{\sqrt{n}} \sqrt{1 - \frac{n}{N}}$$

$\bar{x}$ valimi keskmine

s valimi standardhälve

n valimi maht

$z_{\alpha/2}$ standardiseeritud normaaljaotuse täiendkvantiil, $\alpha = 1 - \beta$

MILLEST SÕLTUB USALDUSVAHEMIKU POOLLAIUS

Standardhälvet muuta ei saa.

$$z_{(1-\beta)/2}\frac{s}{\sqrt{n}}$$

Usaldatavust saame valida.

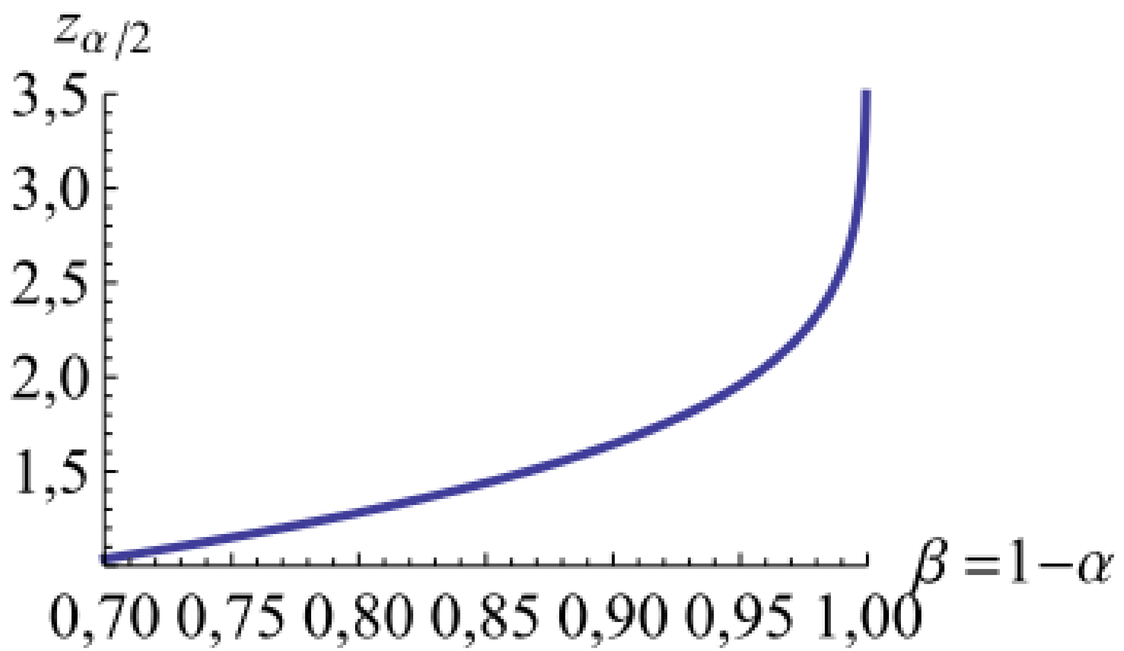
Valimi mahtu saame valida enne uuringut.

USALDATAVUSE VALIK

Kõige sagedamini kasutatakse usaldatavust **0,95.**
Mõnikord ka 0,90 või 0,99.

β	0,9	0,95	0,99
$z_{\alpha/2}$	1,64	1,96	2,59

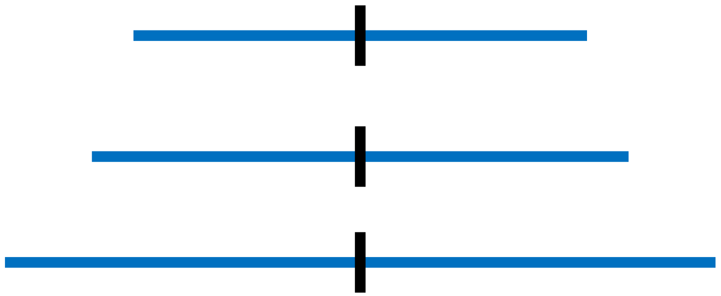
Tõenäosuskordaja sõltuvus usaldatavusest



USALDATAVUS JA USALDUSVAHEMIKU LAIUS

Ühe ja sama valimi korral:
suurem usaldatavus = laiem usaldusvahemik (suurem määramatus)

Usaldatavus β	Kordaja
0,90	1,64
0,95	1,96
0,99	2,58



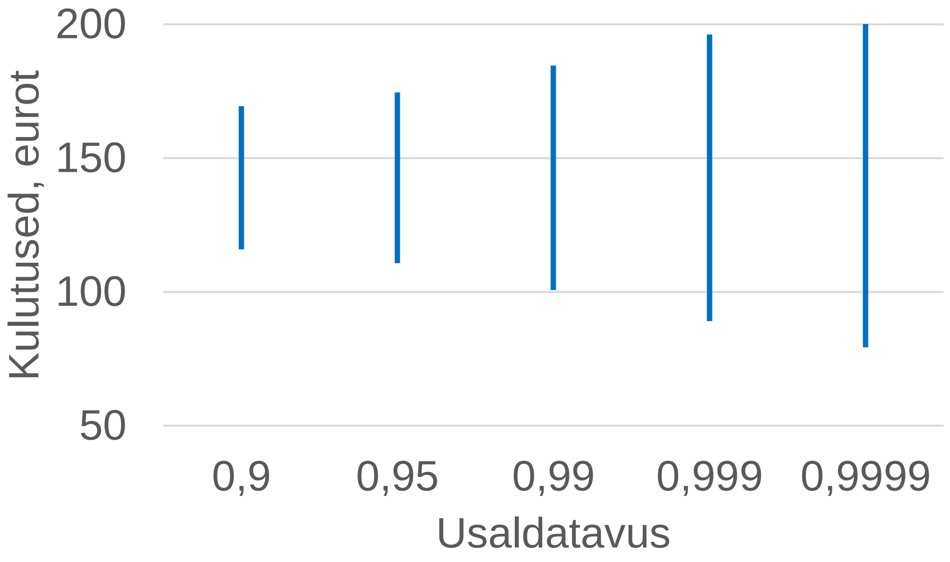
USALDATAVUSE SUURENDAMINE?

Leibkonna eelarve uuring Eestis 2002.
Valimi maht 886.

Kui suured olid kulutused riietele ja jalanõudele ühe leibkonnaliikme kohta aastas?

Keskmine 143 eurot aastas, standardhälve 485 eurot.

Usalduspiiride sõltuvus usaldatavusest



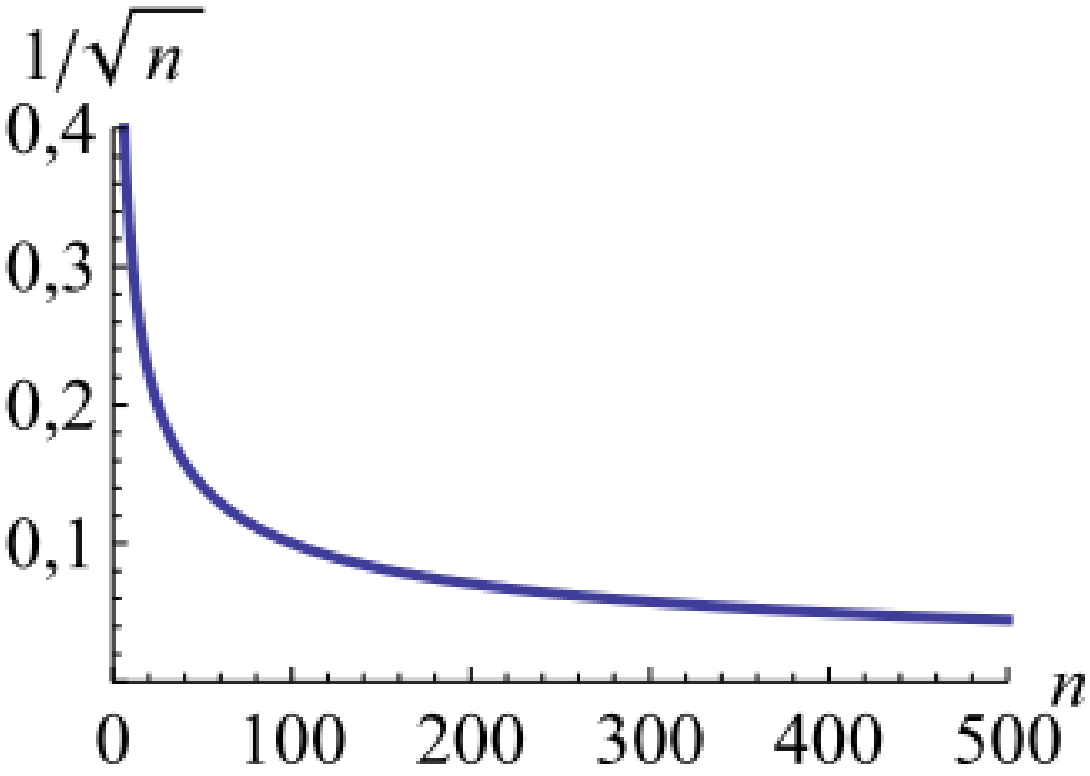
Suurem usaldatavus = suurem määramatus

sama valimi mahu ja standardhälbe korral.

POOLLAIUSE SÕLTUVUS VALIMI MAHUST

$$\Delta x = z_{\alpha/2} \frac{s}{\sqrt{n}}$$

Pöördvõrdeline ruutjuurega valimi mahust.



Kui soovime usaldusvahemiku poollaiust 10 korda vähendada, mitu korda tuleb suurendada valimi mahtu?

100 korda

NÄIDE: KULUD TOIDULE

Leibkonna eelarve uuring 2012.

Valimi maht $n = 9080$

Valimi keskmine $\bar{x} = 915,30$ € leibkonnaliikme kohta aastas

Valimi standardhälve $s = 541,72$ €

Usaldatavuseks võtame $\beta = 0,95$ Vastav tõenäosuskordaja $z_{0,025} = 1,96$

Usaldusvahemiku poollaiuse arvutus $$\Delta x = z_{\alpha/2} \frac{s}{\sqrt{n}} = 1,96 \cdot \frac{541,72}{\sqrt{9080}} \approx 11,1$$

Usaldatavusega 0,95 oli keskmine kulu toiduainetele ja mittealkohoolsetele jookidele $915,3 \pm 11,1$ eurot aastas pereliikme kohta.

Teine võimalus usaldusvahemiku esitamiseks: (904,2; 926,4) eurot aastas.

USALDUSPIIRIDE ESITUSVIISID

Esitusviis 1

$$\bar{x} \pm \Delta x$$

Esitatud on valimi keskmine ja usaldusvahemiku poollaius.

Millistesse piiridesse keskväärtus jääb: vaja arvutada.

915,3 ± 11,1 usaldatavusega 0,95

$915{,}3 \pm 11{,}1_{\,0,95}$

Esitusviis 2

alumine piir $\bar{x} - \Delta x$

ülemine piir $\bar{x} + \Delta x$

Esitatud on alumine ja ülemine usalduspiir.

Keskväärtuse punkthinnang ja usaldusvahemiku laius vaja arvutada.

Vahemik (904,2; 926,4) usaldatavusega 0,95.

NB! Alati tuleb märkida usaldatavus!

JÄRGMISEL LOENGUL JÄTKAME VALIKVAATLUSTEGA

- Keskväärtuse usaldusvahemik väikese valimi korral
- Mediaani usaldusvahemik
- Osakaalu usaldusvahemik