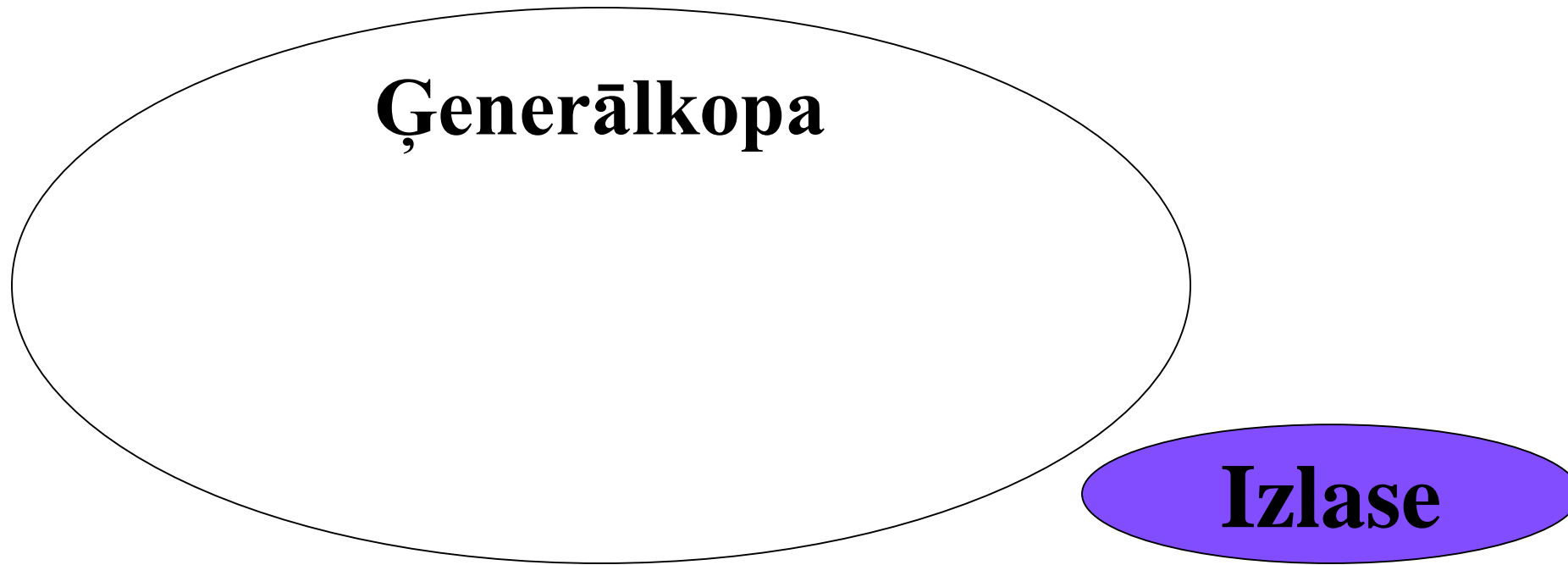


**Mērījumu rezultātu un kļūdu
aprēķini kompleksām
klasterveida izlasēm
starptautisko salīdzinošo
izglītības pētījumu datu
analīzē**

Prof. Andrejs Geske



Ģenerālkopa ir visu mūs interesējošo objektu kopa

Izlase ir kādas apakškopas izraudzīšana no ģenerālkopas tālākiem pētījumiem

+
•
。


IZLASES KĻŪDAS

+
•
。

Izlases kļūdas

Kļūdu, kura rodas, izlases raksturotājiem atšķiroties no ģenerālkopas raksturotājiem, sauc par izlases kļūdu

Ja ģenerālās kopas vidējais ir μ un izlases vidējais ir \bar{x} , tad izlases kļūda ir $\bar{x} - \mu$



Nosakot kādu raksturotāju, piemēram, vidējo vērtību, varam izvēlēties dažādas izlases. Katrā izlasē tiks iegūta sava vidējā vērtība. Šo vidējo vērtību sadalījums nosaka izlases kļūdu.

Izlašu vidējo vērtību standartnovirze

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

kur

σ – izlašu vidējo standartnovirze

\bar{x} – visu izlašu aritmētiskais vidējais

n – izlašu skaits

x_i – i -tās izlases vidējais

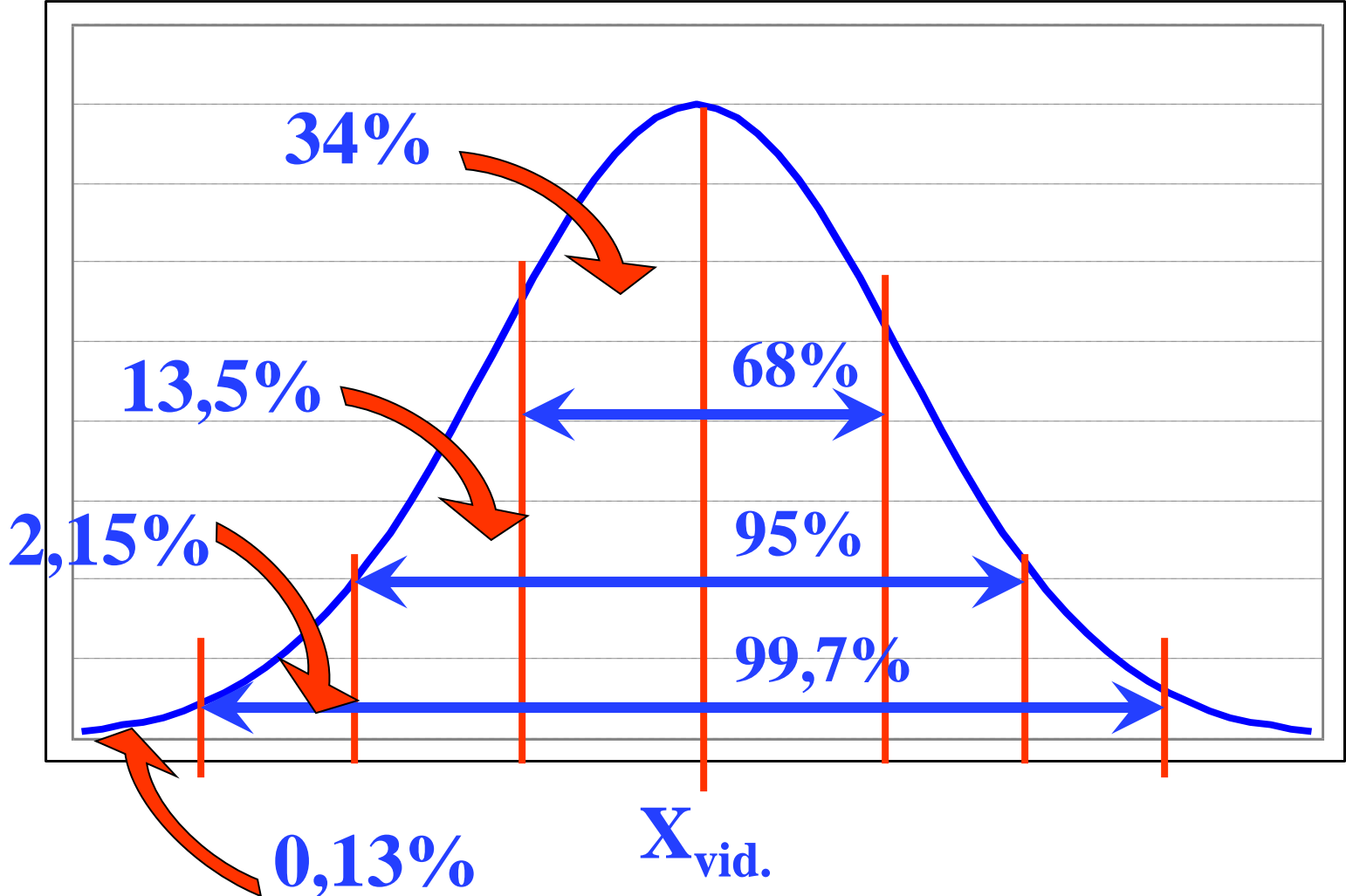
Izlasses standartklūda

**Izlašu vidējo vērtību
standartnovirzi sauc par
izlasses standartklūdu**

Standartklūdas lietojums

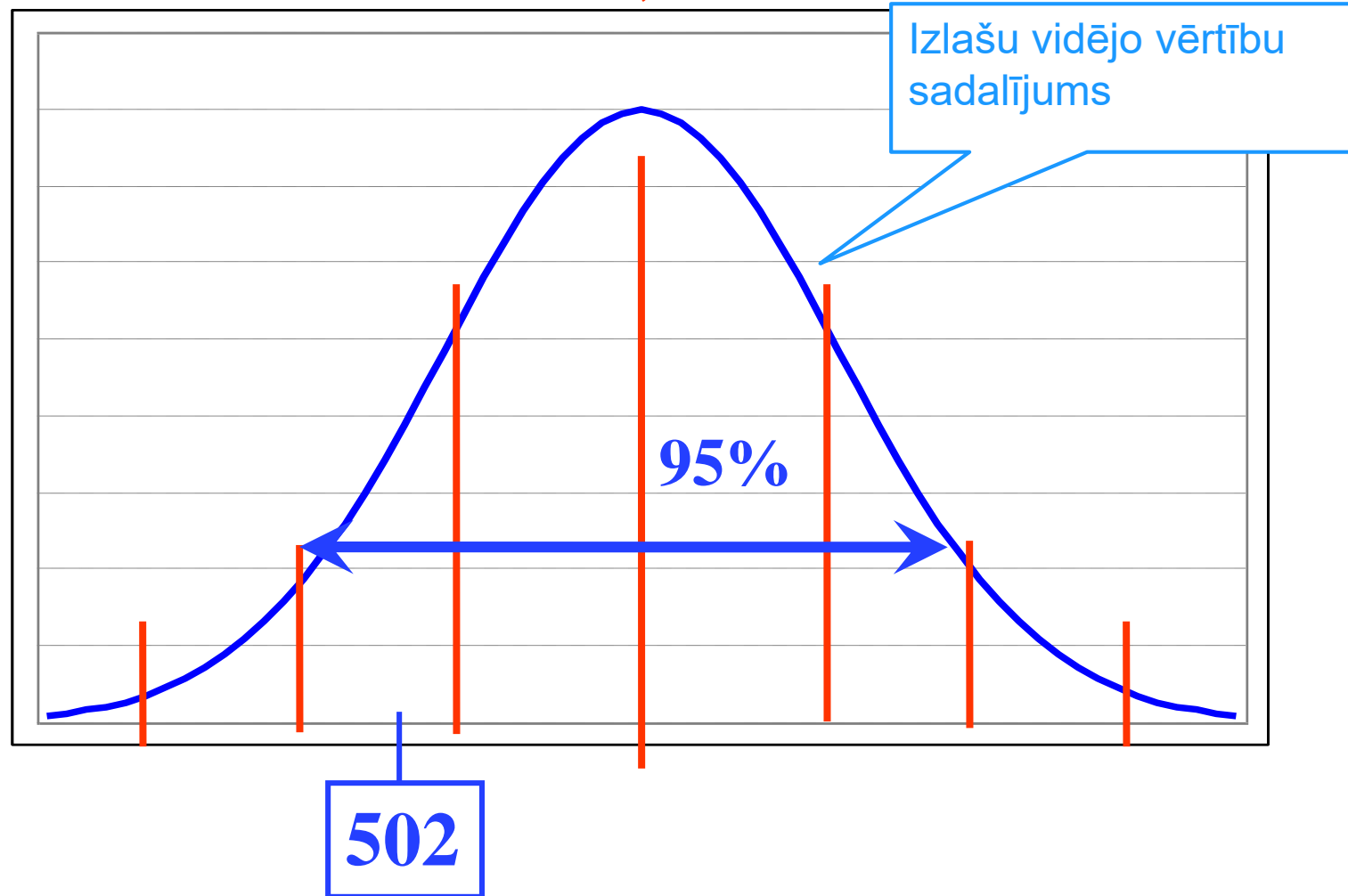
- Lai noteiktu atšķirības starp grupām
- Lai noteiktu ticamības intervālus, piemēram 95% ticamībai $\pm 1,96 SK$

Normālais sadalījums



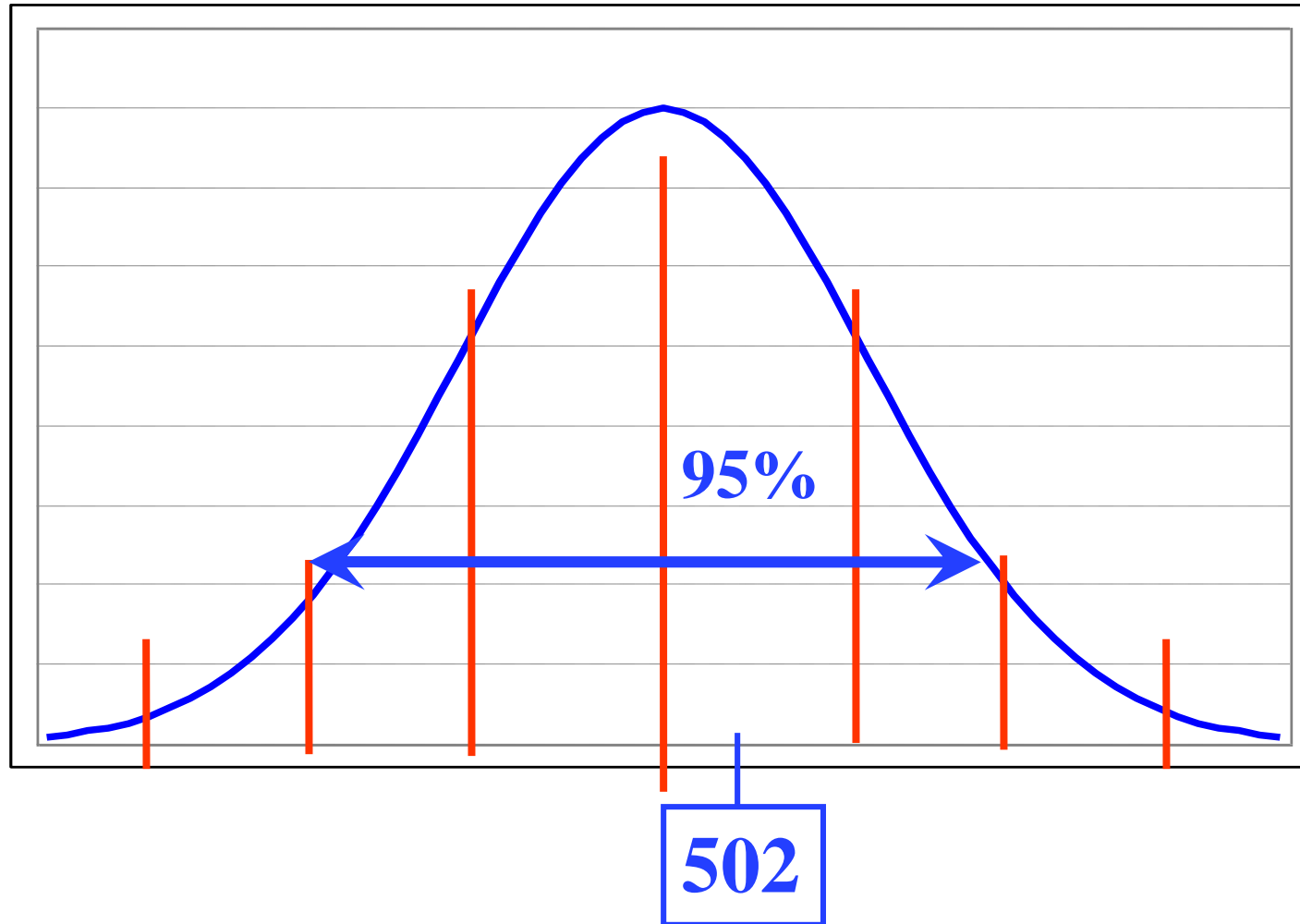
Standartklūdas interpretācija

Mērījums: vidējā vērtība **502**, standartklūda **3,0**
Ar **95%** ticamību ģenerālkopas vidējā vērtība būs starp **496** un **508**. Statistiskā kļūda **6,0** pie **95%** ticamības.



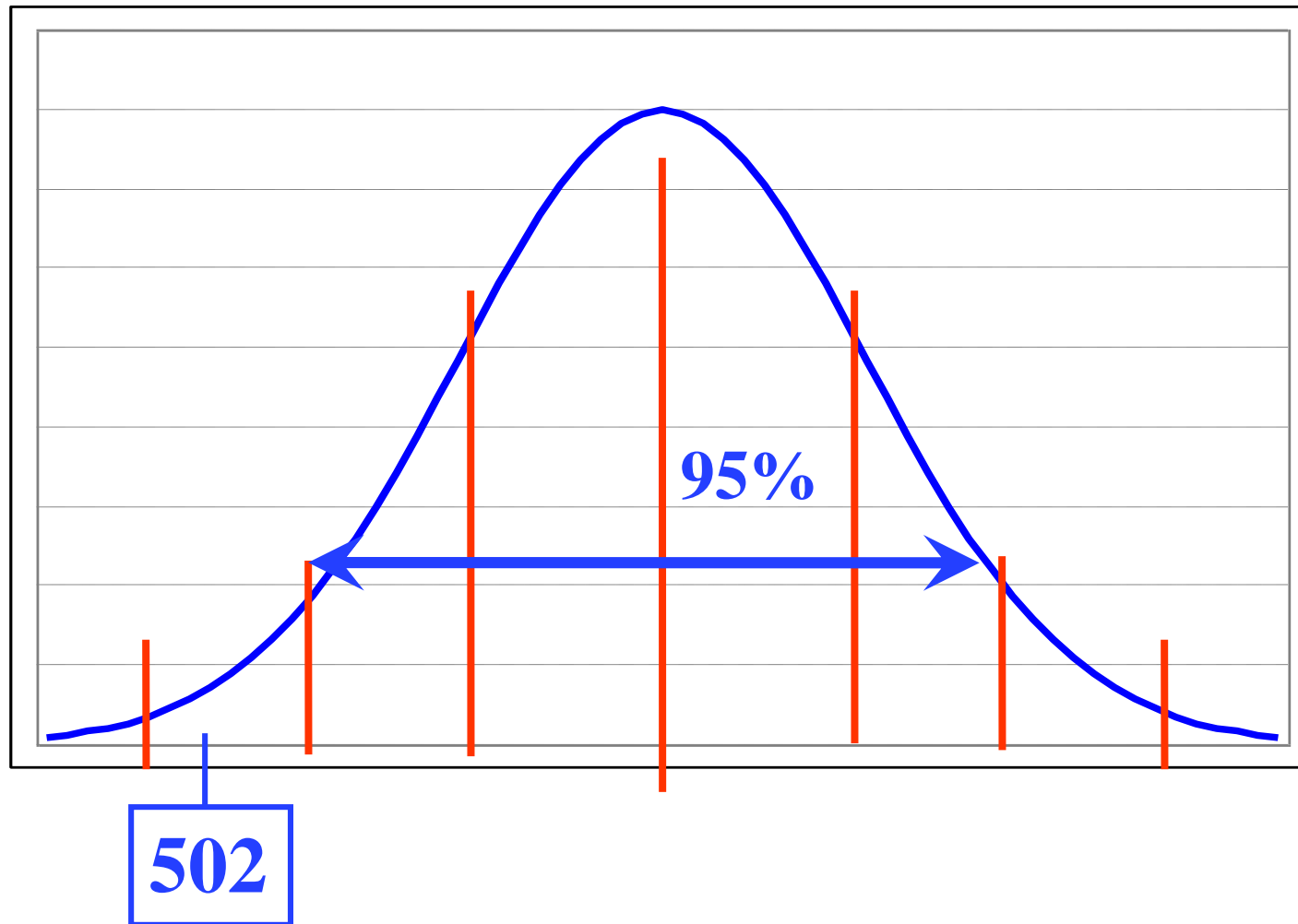
Standartklūdas interpretācija

Mērījums: vidējā vērtība **502**, standartklūda **3,0**
Ar **95%** ticamību ģenerālkopas vidējā vērtība būs starp **496** un **508**. Statistiskā kļūda **6,0** pie **95%** ticamības.



Standartklūdas interpretācija

Mērījums: vidējā vērtība **502**, standartklūda **3,0**
Ar **95%** ticamību ģenerālkopas vidējā vērtība būs starp **496** un **508**. Statistiskā kļūda **6,0** pie **95%** ticamības.



Izlašu vidējo standartnovirze, ja veikta gadījumi izlase

$$\text{SK} \approx \frac{\sigma}{\sqrt{n}}$$

σ

Mērījumu standartnovirze

n

Mērījumu skaits

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

PĒTĪJUMA
DALĪBNIĒKU IZLASES
VEIDOŠANA



Izlases veidošanas metožu klasifikācija

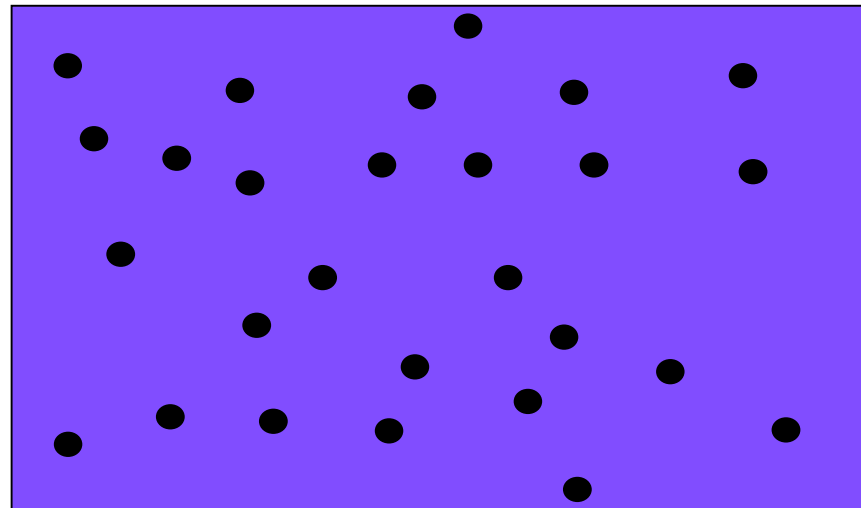
- Nevarbūtīgās izlases metodes
- Varbūtīgās izlases metodes

Varbūtīgās izlases

Katram ģenerālās kopas elementam ir lielāka par nulli varbūtība iekļūt izlasē

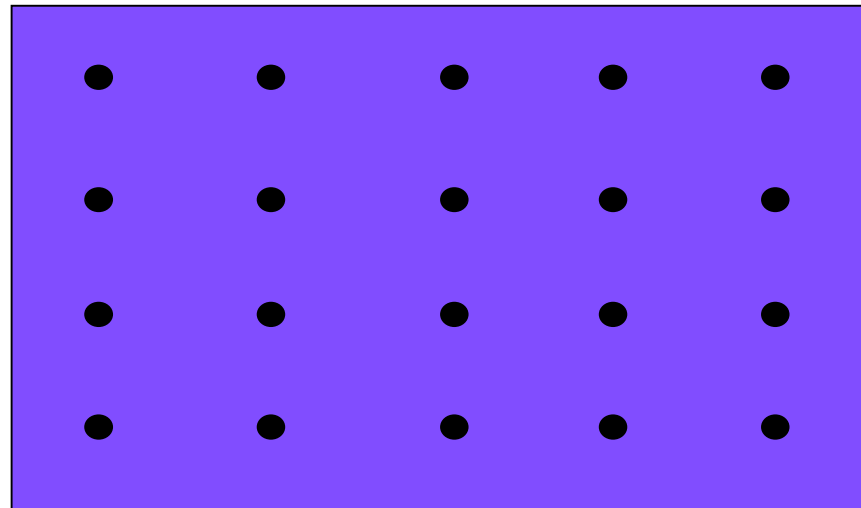
Vienkāršā gadījumizlase

Vienkāršajā gadījumizlasē katram ģenerālās kopas elementam ir vienāda iespēja iekļūt izlasē



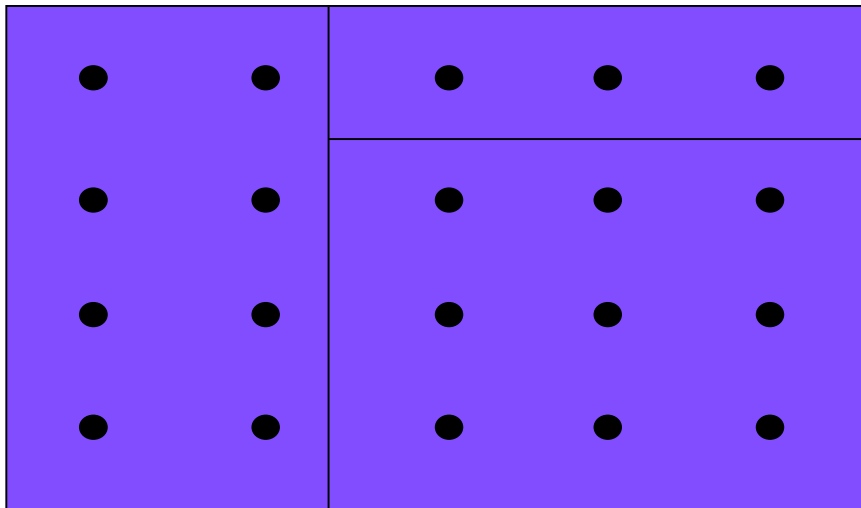
Sistemātiskā izlase

Sistemātiskajā izlases metodē no visu elementu saraksta izvēlās elementus, kuri atrodas noteiktā attālumā viens no otra

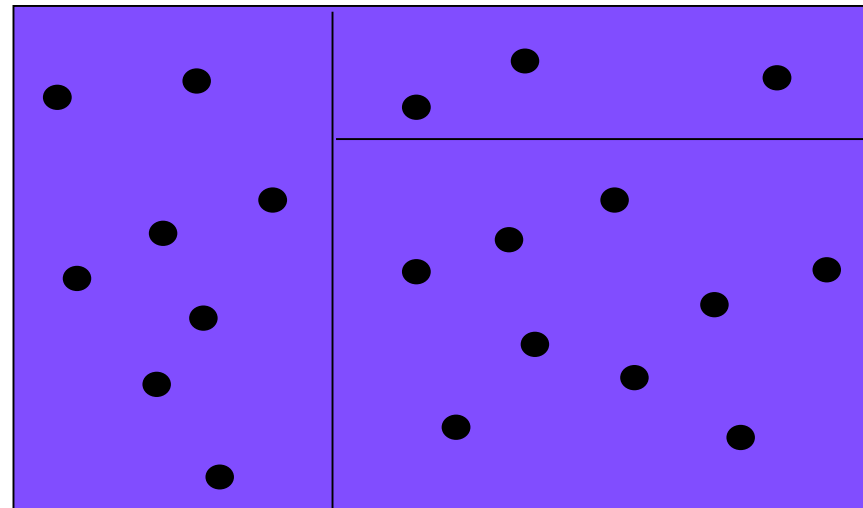


Stratificētā izlase

Veicot stratificēto izlasi ģenerālās kopas elementi vispirms tiek sakārtoti slāņos pēc kādas noteiktas pazīmes, tad katra slāņa iekšpusē tiek veikta vienkārša gadījumizlase vai sistemātiskā izlase



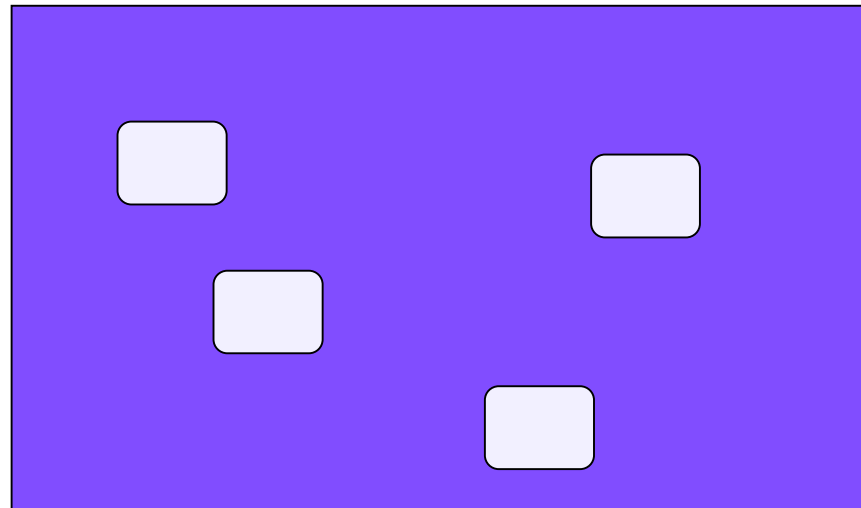
Otrais solis -
sistemātiskā izlase



Otrais solis -
vienkārša gadījumizlase

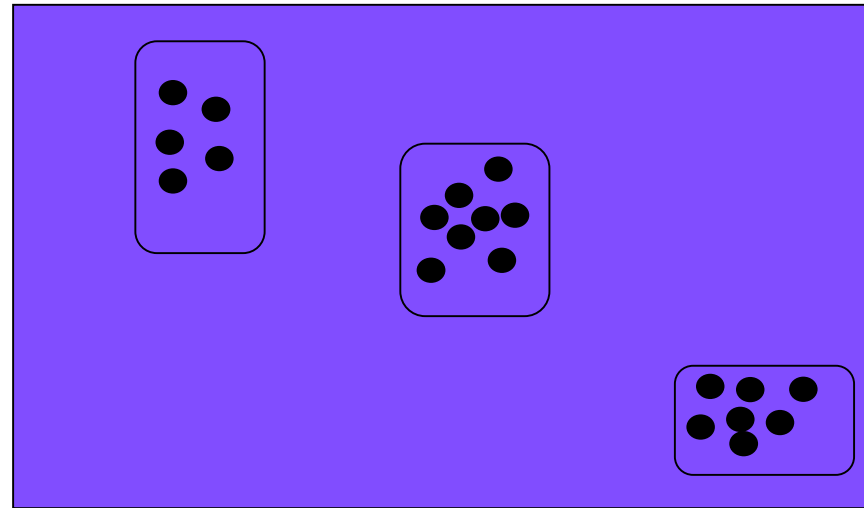
Klasteru izlase

Izlasē izvēlās kādas grupas, kurās esošos elementus pilnībā iekļauj pētījumā.



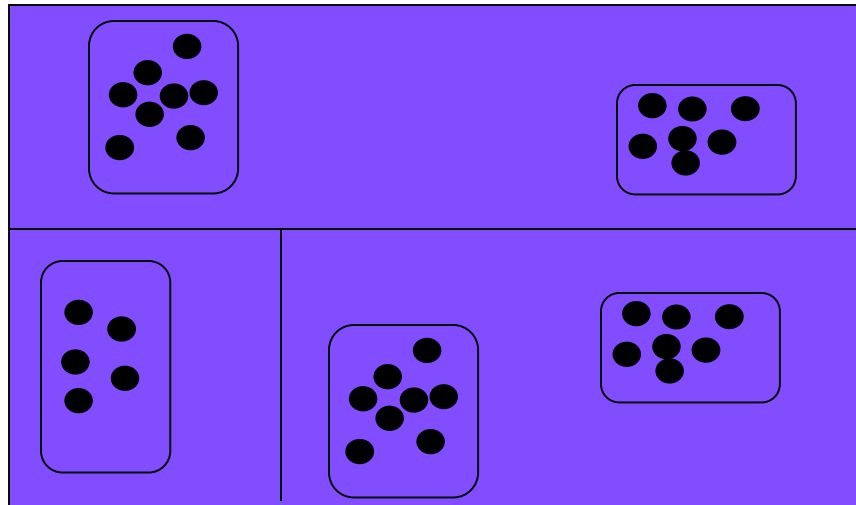
Vairākpakāpju izlase

Izlasē izvēlās kādas noteiktas grupas, nākamo izlases soli veic izvēloties elementus no šīm grupām



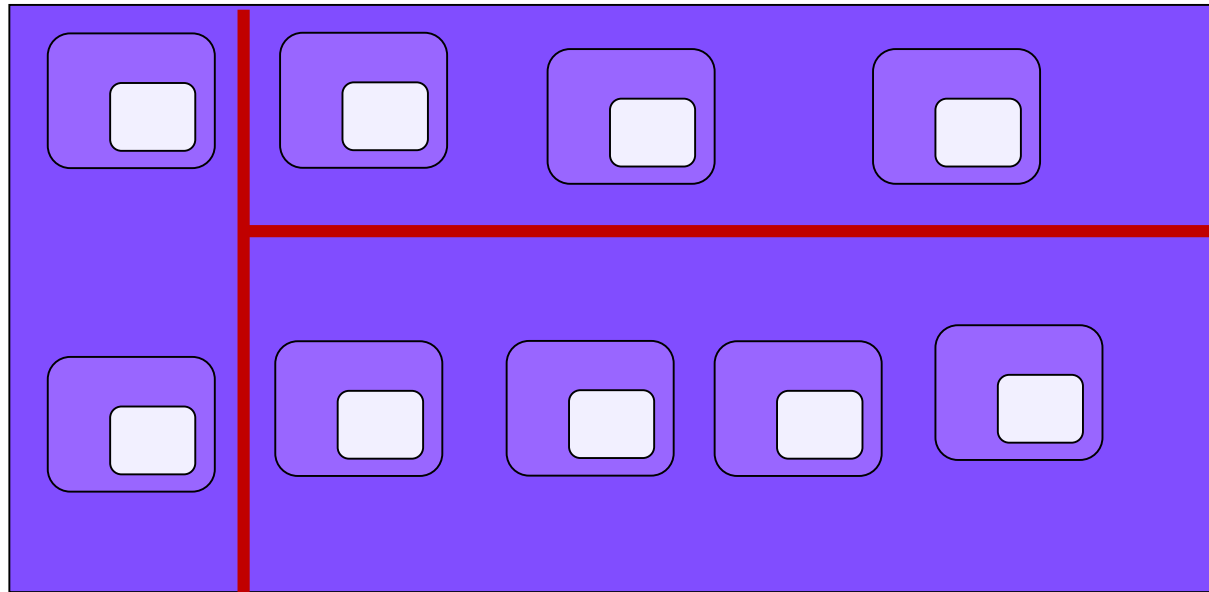
Vairākpakāpju stratificētā izlase

Veicot stratificēto izlasi ģenerālās kopas elementi vispirms tiek sakārtoti slāņos pēc kādas noteiktas pazīmes, tad katra slāņa iekšpusē izvēlās kādas noteiktas grupas, nākamo izlases soli veic izvēloties elementus no šīm grupām



Startificētā klaseru izlase

- Slāņi – Skolas –
 - Klases (IEA PIRLS, TIMSS, u.c) vai
 - Skolēni (OECD PISA) vai
 - Skolotāji (OECD TALIS)



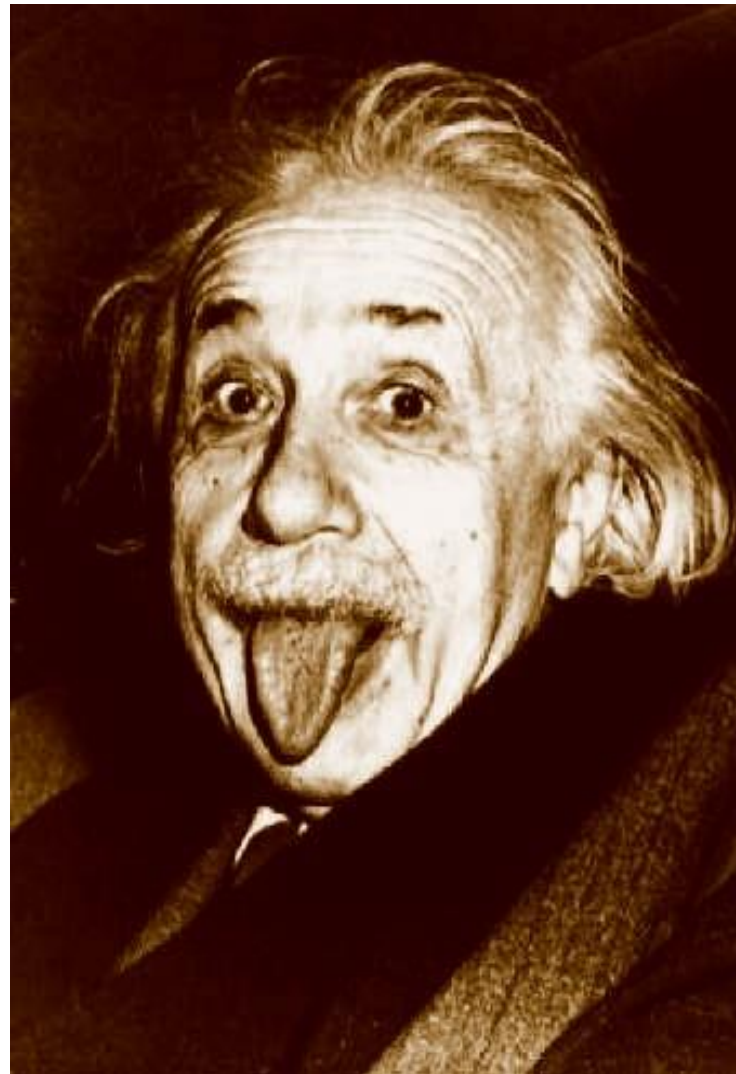
OECD PISA 2022 slāņi

	Stratification Variable MS	Explicit or Implicit	Number of Levels	Level Names
MS Strata Var name 1	Urbanisation	Explicit	4	Riga, City, Town, Rural
MS Strata Var name 2	School type	Implicit	5	Basic, secondary, gymnasium, vocational, part time

(1) Explicit Strata	(2) Implicit Strata	(3) Schools	(4) Students
Latvian - Gymnasium	ALL - Town-Rural	14	806
	ALL - City	11	999
	ALL - Riga	5	541
Latvian - Gymnasium	Total	30	2,346
Latvian - Basic	ALL - Town-Rural	175	2,525
	ALL - City	11	340
	ALL - Riga	16	626
Latvian - Basic	Total	202	3,491
Latvian - Secondary	ALL - Town-Rural	113	4,281
	ALL - City	19	1,312
	ALL - Riga	32	2,194
Latvian - Secondary	Total	164	7,787
Russian - ALL	Gymnasium - Riga	1	171
	Basic - Town-Rural	5	60
	Basic - City	6	250
	Basic - Riga	7	228
	Secondary - Town-Rural	8	261
	Secondary - City	19	1,114
	Secondary - Riga	39	2,777
Russian - ALL	Total	85	4,861
Latvian & Russian - ALL	Basic - Town-Rural	3	138
	Basic - City	1	51
	Basic - Riga	1	42
	Secondary - Town-Rural	11	639
	Secondary - City	5	261
	Secondary - Riga	7	209
Latvian & Russian - ALL	Total	28	1,340
Latvia	Total:	509	19,825

ICCS 2022

Izlases piemērs, skatot foto

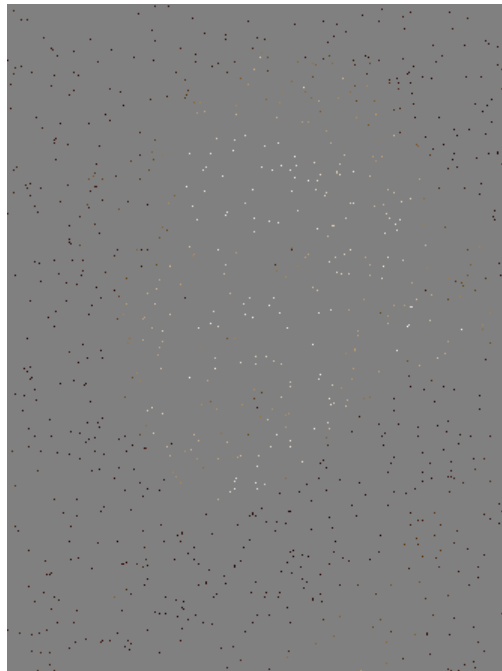


Nejaušā izlase

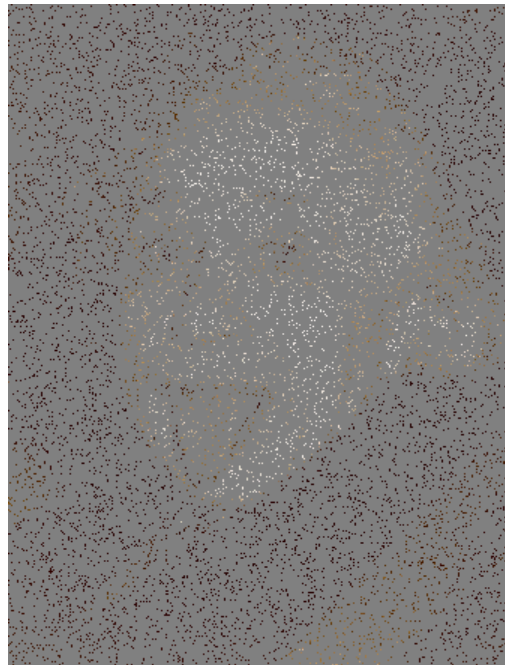
Izlases lielums

=

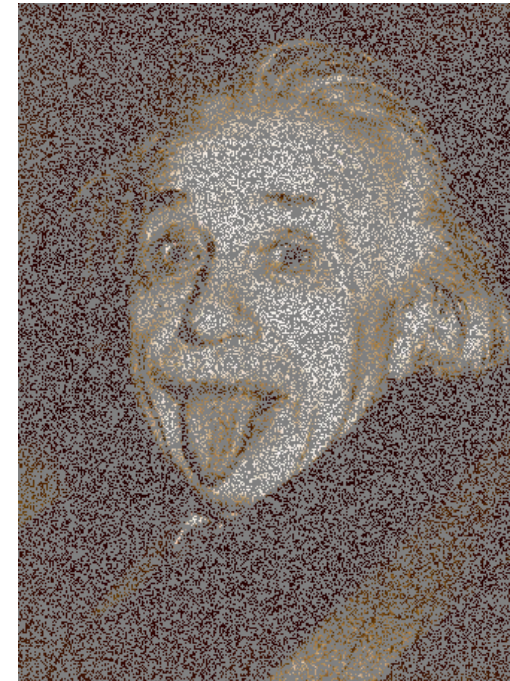
1,000



10,000

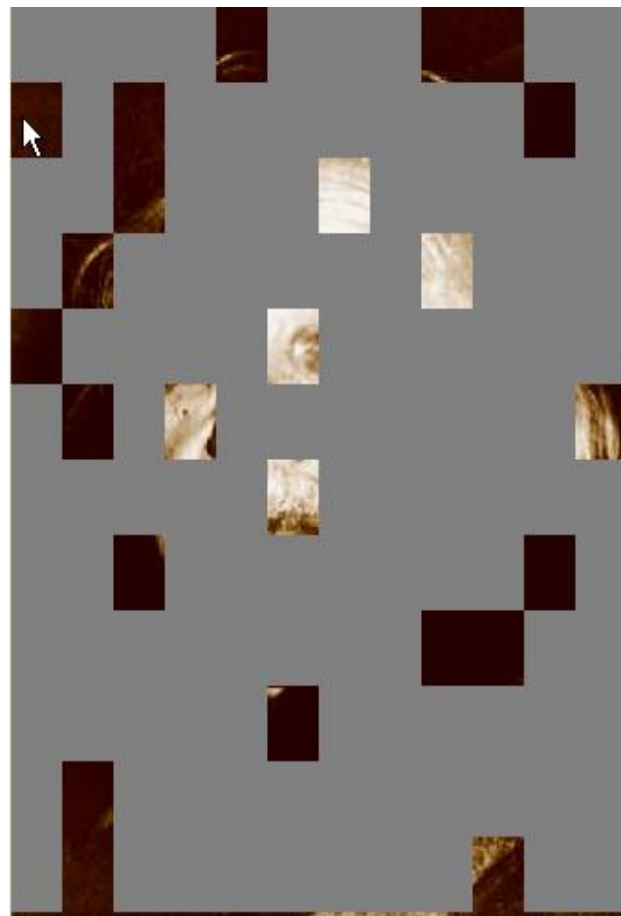
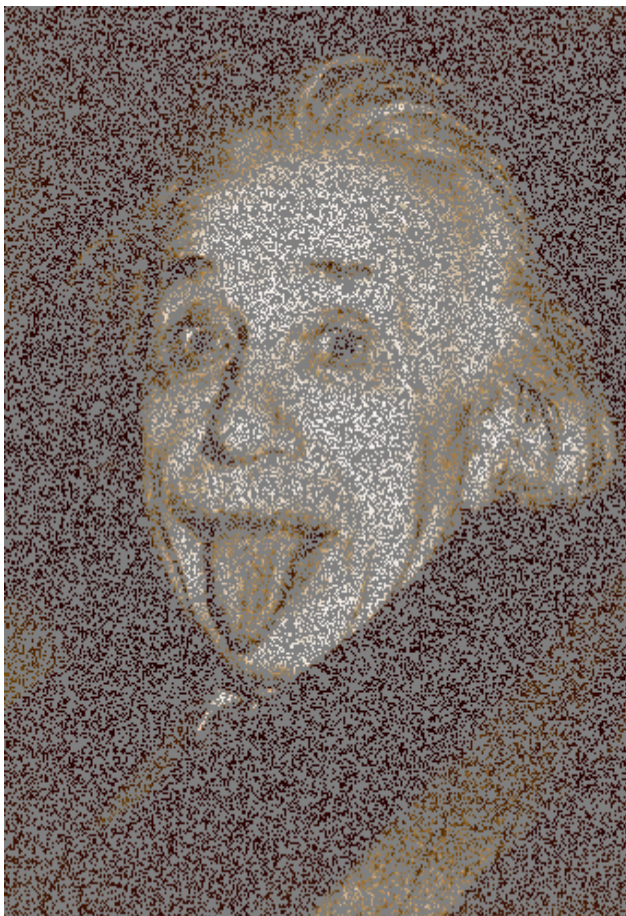


50,000



Klasetru izlase

Izlases
lielums=
50 000



Klasterveida izlasei izlases kļūdas ir lielākas nekā neajušai izlase. Klasterī elementi ir līdzīgāki viens otram nekā ģenerālkopā

+
•
◦

SVARU JĒDZIENS

+
•
◦

Svari



Ģenerālā kopa –
2 skolās 800 skolēni



Izlase

Katra skolēna svars = 5

Katra skolēna svars = 5

$$120 \times 5 + 40 \times 5 = 800$$



Ģenerālā kopa –
2 skolās 800 skolēni



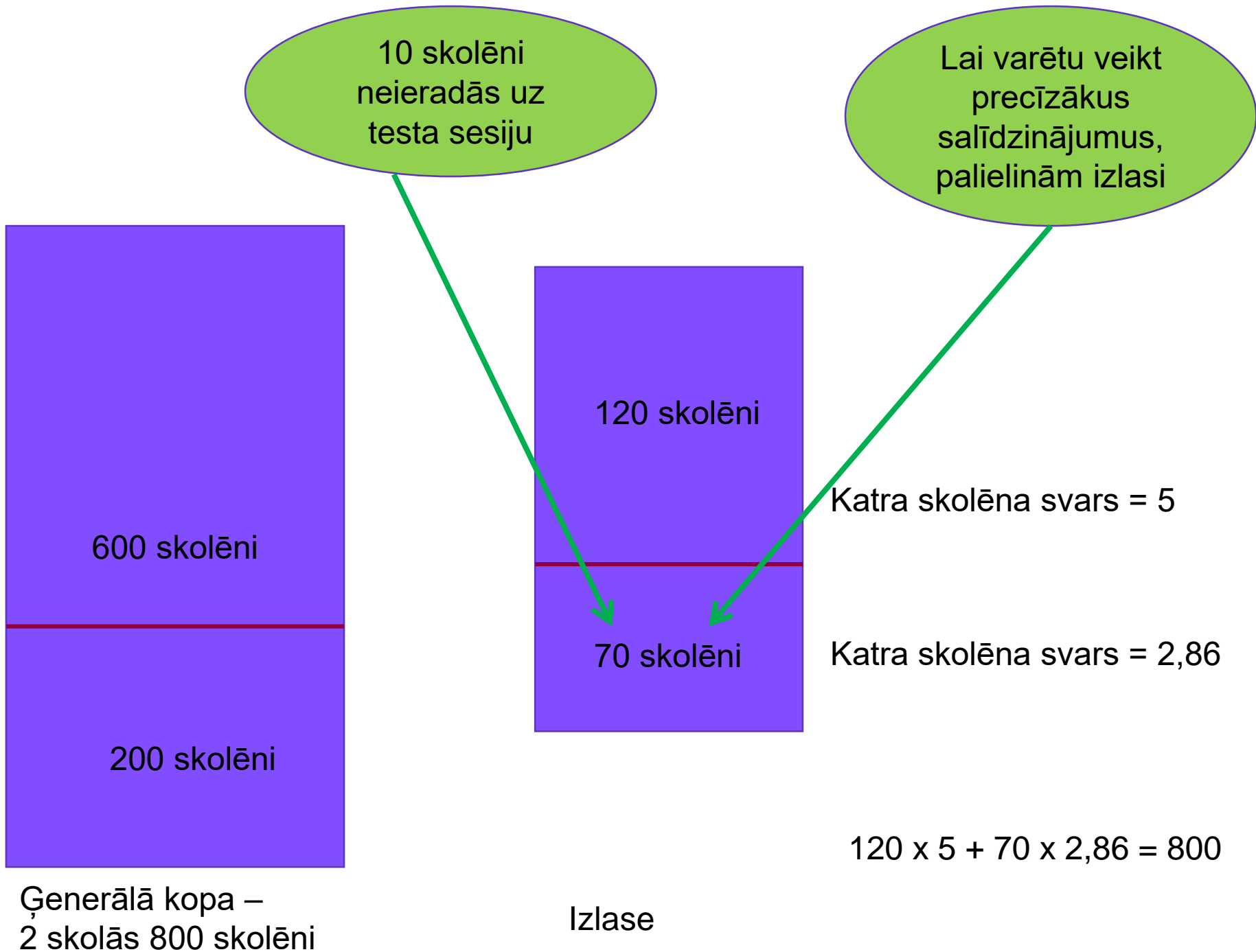
Izlase

Lai varētu veikt
precīzākus
salīdzinājumus,
palielinām izlasi

Katra skolēna svars = 5

Katra skolēna svars = 2,5

$$120 \times 5 + 80 \times 2,5 = 800$$



Piemērs svaru nozīmei

	1. slānis	2. slānis	Kopā
Kopa	1 000	10 000	11 000
Izlase	100	100	200
Izlases svars	10	100	
Nesvērts vidējais	500	600	550

Nesvērtais vidējais

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

Acīmredzami nepareizs risinājums!

Piemērs svaru nozīmei

	1. slānis	2. slānis	Kopā
Kopa	1 000	10 000	11 000
Izlase	100	100	200
Izlases svars	10	1000	
Nesvērts vidējais	500	600	550
Svērts vidējais	500	600	591

Svērtais vidējais

$$= \frac{\sum wgt \cdot x}{\sum wgt}$$

Svaru atšķirības nosaka

- Varbūtības tikt iekļautam izlasē dažādām skolām un skolēniem nav vienādas
 - Dažādos slāņos ir dažādas varbūtības iekļaut skolu
 - Skolām ir dažādi lielumi (PPS -- Probability proportional to size sample)
 - Arī ļoti lielā skolā izvēlas noteiktu skaitu klašu / skolēnu (piem. PISA – max 35 skolēni)
- Atšķirīgi nepiedalīšanās parametri -
 - Dažādos ārējos un iekšējos slāņos (explicit and implicit strata)
 - Dažādās skolās
 - Dažādās klasēs

Piemērs – IEA PIRLS 2016 – Latvija – 150 skolu 218 klases

Lielas skolas,
augsta dalība

Klases ID	Svars	Dalīb skolēni
401	1.00	10
201	1.66	28
202	1.71	21
303	1.80	15
301	1.92	25
1401	1.94	32
...
28601	14.82	5
29001	14.82	5
29201	14.82	4
29601	14.82	2
27201	16.47	9
28801	17.29	6

Mazas skolas,
zema dalība



ATPAKAĻ PIE
IZLASES KĻŪDU
NOVĒRTĒŠANAS

Izlases kļūdu novērtēšana

- **Apmēram pirms pusgadsimta tika izstrādātas metodes, kuras ļauj novērtēt izlases kļūdas kompleksās izlasēs :**
- **Atkārtotas aizstāšanas metodes (Repeated replication methods)**
 - **Jackknife Repeated Replication (JRR)**
 - **Balanced Repeated Replication (BRR)**
 - **u.c.**
- **Kopīga iezīme – daudzkārtīgi tiek izveidotas izlases apakšizlases**

Jackknife Repeated Replication

- Tiek izvaidoti apakšizlases manipulējot ar svariem – t.i. tiek lietoti vēl papildus svari
- IEA PIRLS, ICCS, TIMSS u.c. pētījumos lieto “JK2”-metodi:
- 2 maksimāli līdzīgas skolas tiek apvienotas vienā zonā, kopumā tiek izveidotas 75 zonas. Tiek izveidots jauns mainīgais JKZONE
- Katrā zonā vienai skolai tiek piekārtots indikators JKREP. Indikatora vērtība 0 norāda, ka, veidojot apakšizlasi, skola tiks izslēgata. Indikatora vērtība 1 norāda, ka, veidojot apakšizlasi, skolas svari tiks dubultoti.

Replicated Weights Using JRR Sampling Information

School	Student	JKZone	JKRep	Score	Weight	RWgt1	RWgt2	RWgt3	RWgt4	RWgt5
01	0101	1	1	96	6	12	6	6	6	6
01	0102	1	1	74	7	14	7	7	7	7
02	0201	1	0	56	9	0	9	9	9	9
02	0202	1	0	62	2	0	2	2	2	2
03	0301	2	0	34	2	2	0	2	2	2
03	0302	2	0	53	7	7	0	7	7	7
04	0401	2	1	96	5	5	10	5	5	5
04	0402	2	1	57	2	2	4	2	2	2
05	0501	3	0	78	2	2	2	0	2	2
05	0502	3	0	87	2	2	2	0	2	2
06	0601	3	1	71	8	8	8	16	8	8
06	0602	3	1	47	7	7	7	14	7	7
07	0701	4	0	56	1	1	1	1	0	1
07	0702	4	0	25	2	2	2	2	0	2
08	0801	4	1	82	2	2	2	2	4	2
08	0802	4	1	68	2	2	2	2	4	2
09	0901	5	1	16	10	10	10	10	10	20
09	0902	5	1	80	8	8	8	8	8	16
10	1001	5	0	20	5	5	5	5	5	0
10	1002	5	0	35	9	9	9	9	9	0
Unweighted Mean:		59.7	Sum of Weights:		100	96	109	99	102	
Weighted Mean:		57.2	Weighted Means:		60.7	60.0	56.6	58.6	58.8	

- Kopumā tiek aprēķinātas 75 izlašu vērtības
- No tām tiek aprēķināta izlases standartklūda

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

kur

σ – izlašu vidējo standartnovirze

$n = 75$

Balanced Repeated Replication

- Tāpat klasteri tiek sadalīti zonās
- Pēc īpašiem algoritmiem tiek izveidoti 80 mainīgie, kuri ietver papildus svarus $W_FSTURWT01 \dots W_FSTURWT80$
- Tiek aprēķinātas vērtības 80 apakšizlasēm
- Tiek atrasta šo izlašu vērtību standartnovirze, kura arī ir standartkļūda

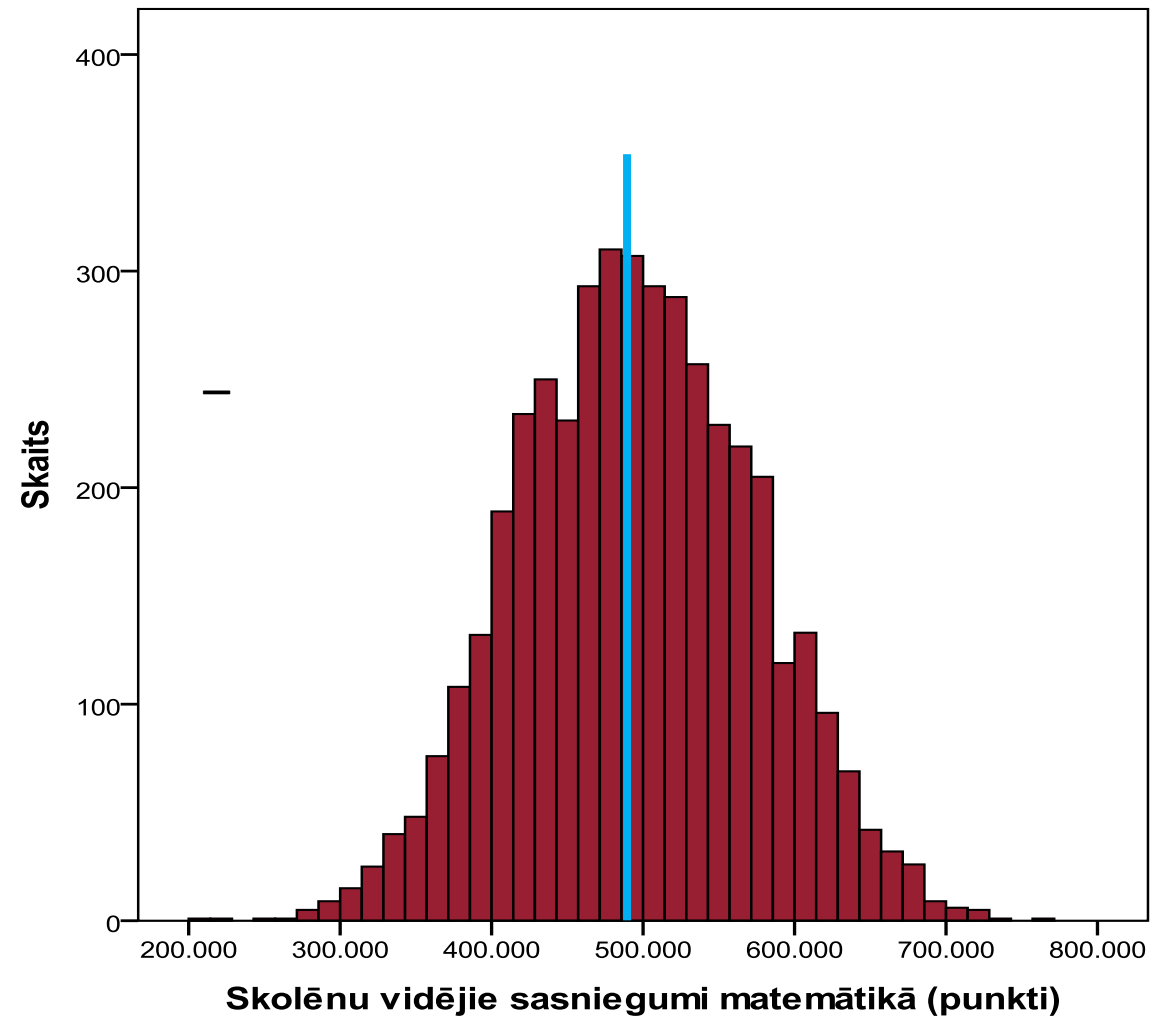
Piemērs no OECD PISA 2015 datu failā

W_FSTUWT	W_FSTURW T1	W_FSTURW T2	W_FSTURW T3	W_FSTURW T4	W_FSTURW T5	W_FSTURW T6	W_FSTURW T7
2.42888	3.85299	3.53558	1.13673	1.26531	3.85299	3.53558	1.13673
2.42888	3.85299	3.53558	1.13673	1.26531	3.85299	3.53558	1.13673
2.42888	3.85299	3.53558	1.13673	1.26531	3.85299	3.53558	1.13673
2.42888	3.85299	3.53558	1.13673	1.26531	3.85299	3.53558	1.13673
2.42888	3.85299	3.53558	1.13673	1.26531	3.85299	3.53558	1.13673
2.42888	3.85299	3.53558	1.13673	1.26531	3.85299	3.53558	1.13673
2.42888	3.85299	3.53558	1.13673	1.26531	3.85299	3.53558	1.13673
2.42888	3.85299	3.53558	1.13673	1.26531	3.85299	3.53558	1.13673
2.42888	3.85299	3.53558	1.13673	1.26531	3.85299	3.53558	1.13673
2.08329	3.15774	3.15774	1.02300	1.02300	3.15774	3.15774	1.02300
2.08329	3.15774	3.15774	1.02300	1.02300	3.15774	3.15774	1.02300

OECD PISA 2012 Latvija

Vidējais Latvija = 491
SN Latvija = 82
SN OECD = 98

Ja būtu gadījuma
izlase un $n=4000$, tad
 $SK = 82 / 63 = 1,3$
tomēr
 $SK = 2,8$



SASNIEGUMU
NOVĒRTĒŠANA UN
MĒRĪJUMU KĻŪDA



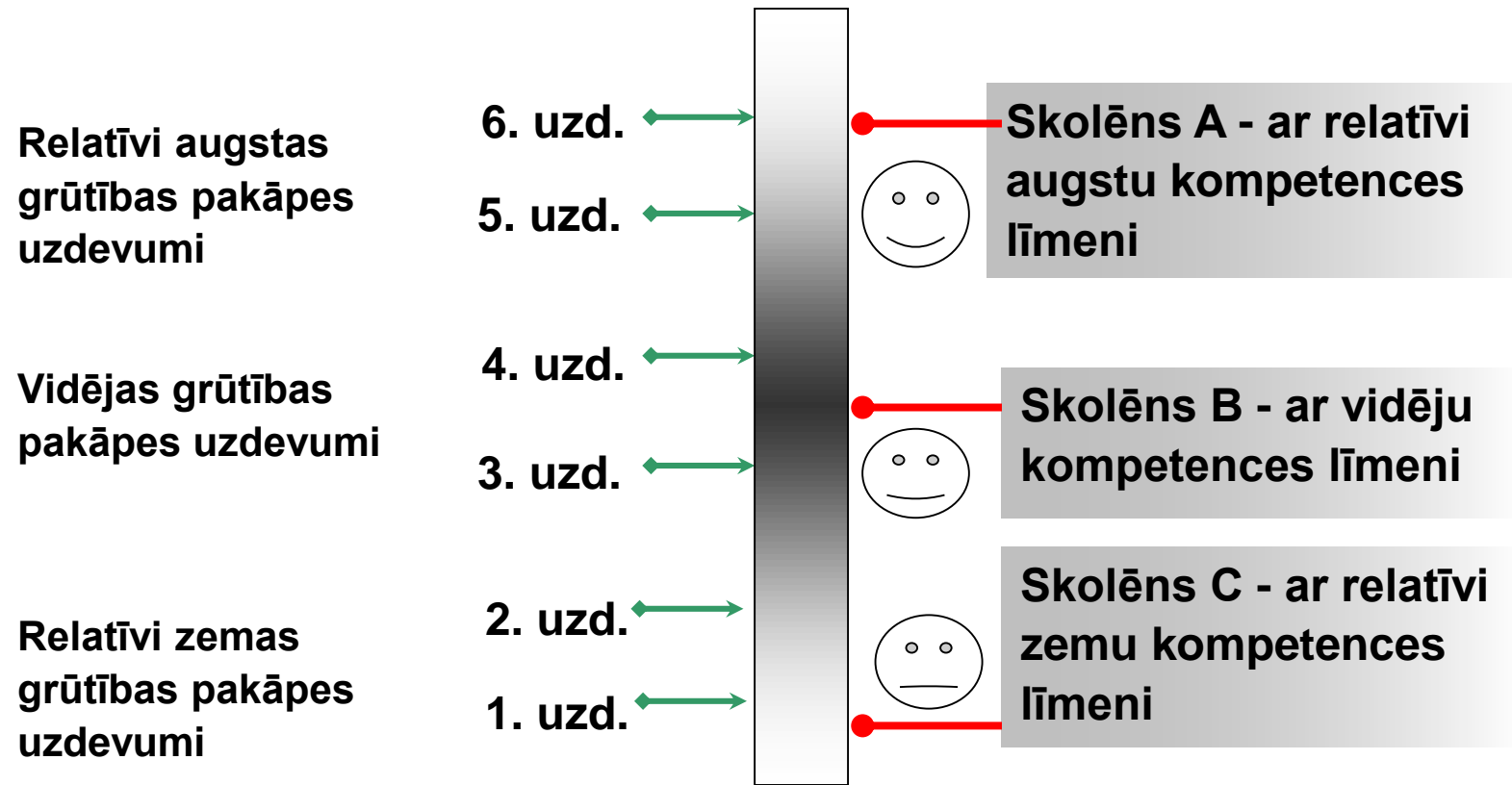
Sasniegumu skalas veidošana – *Jautājuma atbildes teorija* *Item Response Theory – IRT*

Vienota skala uzdevumu grūtībai un skolēnu kompetencei

Uzdevumu
grūtība

Uzdevumu un sasniegumu
skala (punktos)

Skolēnu
kompetence



Galvenās priekšrocības

- Vienā skalā var apvienot atšķirīgu formu un grūtību uzdevumus
- Vienotā skalā var iegūt tādu respondentu vērtējumu, kuri veikuši atšķirīgus uzdevumus
- Piemērota adaptīviem un daļēji adaptīviem testiem

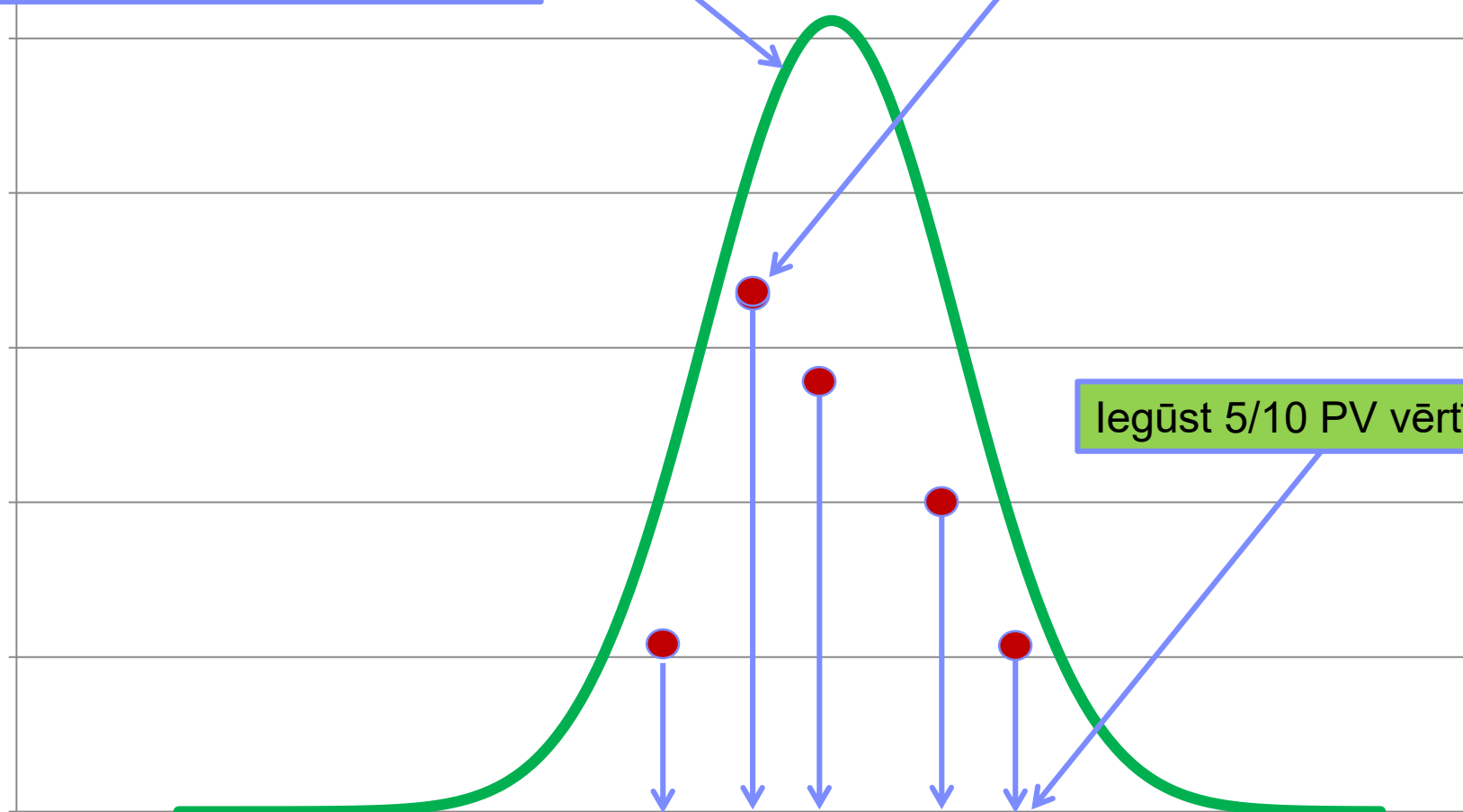
Ticamās vērtības (*Plausible Values – PV*)

- Ietver nejaušas izkliedes komponentes
- Katram skolēnam tiek piešķirtas piecas vai desmit ticamās vērtības (PV1, ..., PV10)
- Veicot datu apstrādi, katra operācija jāatkārto piecas/desmit reizes, beigās aprēķinot vidējo vērtību
- Aprēķinot ticamās vērtības, tiek ņemtas vērā ne tikai skolēnu atbildes uz testa jautājumiem, bet arī atbildes uz aptaujas jautājumiem
- Nav piemērotas individuālam respondenta vērtējumam

PV skaidrojums

Sastāda katram respondentam, ņemot vērā sasniegumus IRT modelī un aptaujas jautājumus

Zem līknes nejauši ģenerē 5/10 punktus

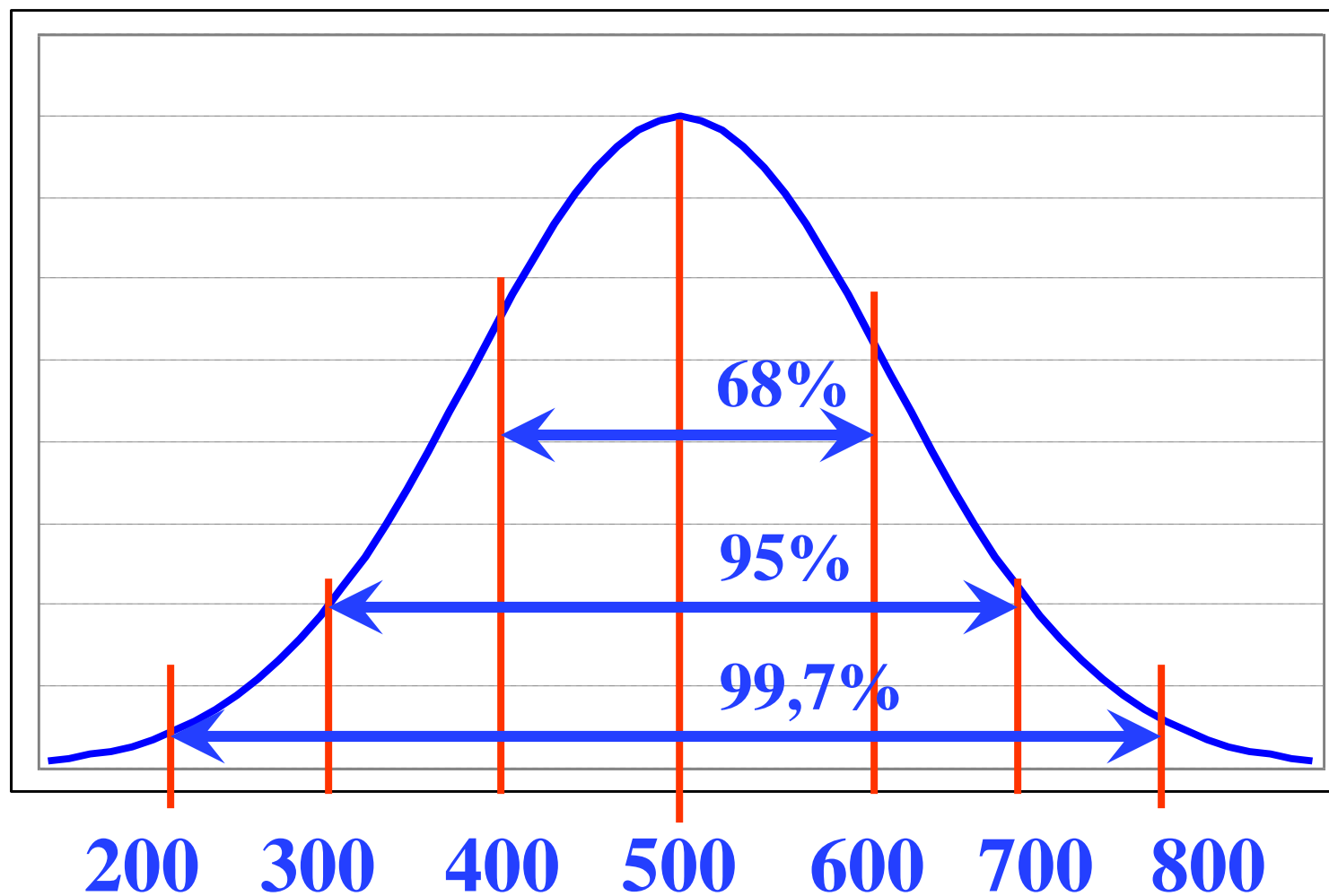


legūst 5/10 PV vērtības

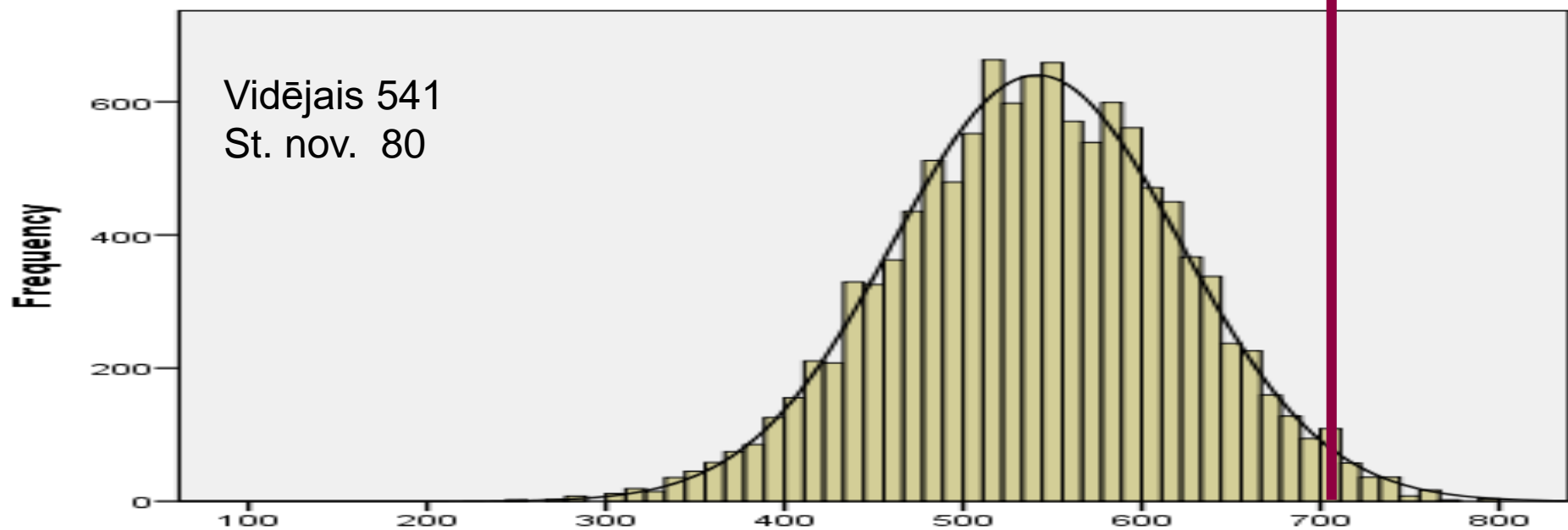
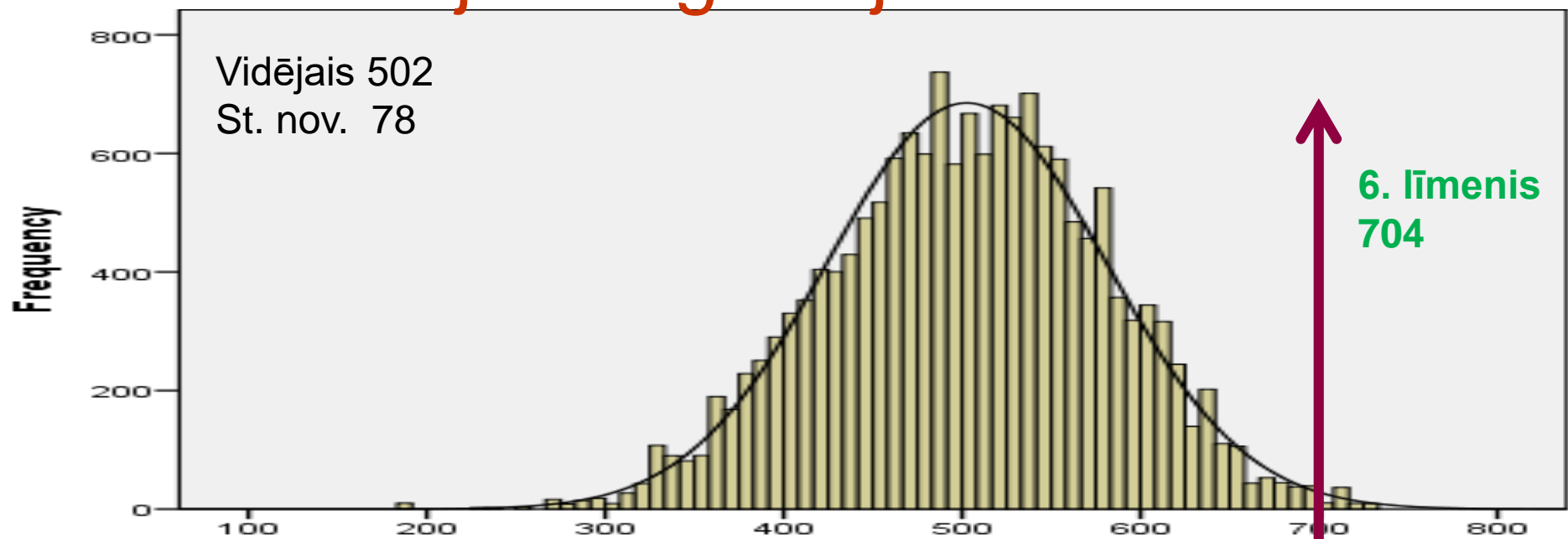
Kopējā standartklūda

To veido Izlases standartklūda un Mērījumu (PV) standartklūda

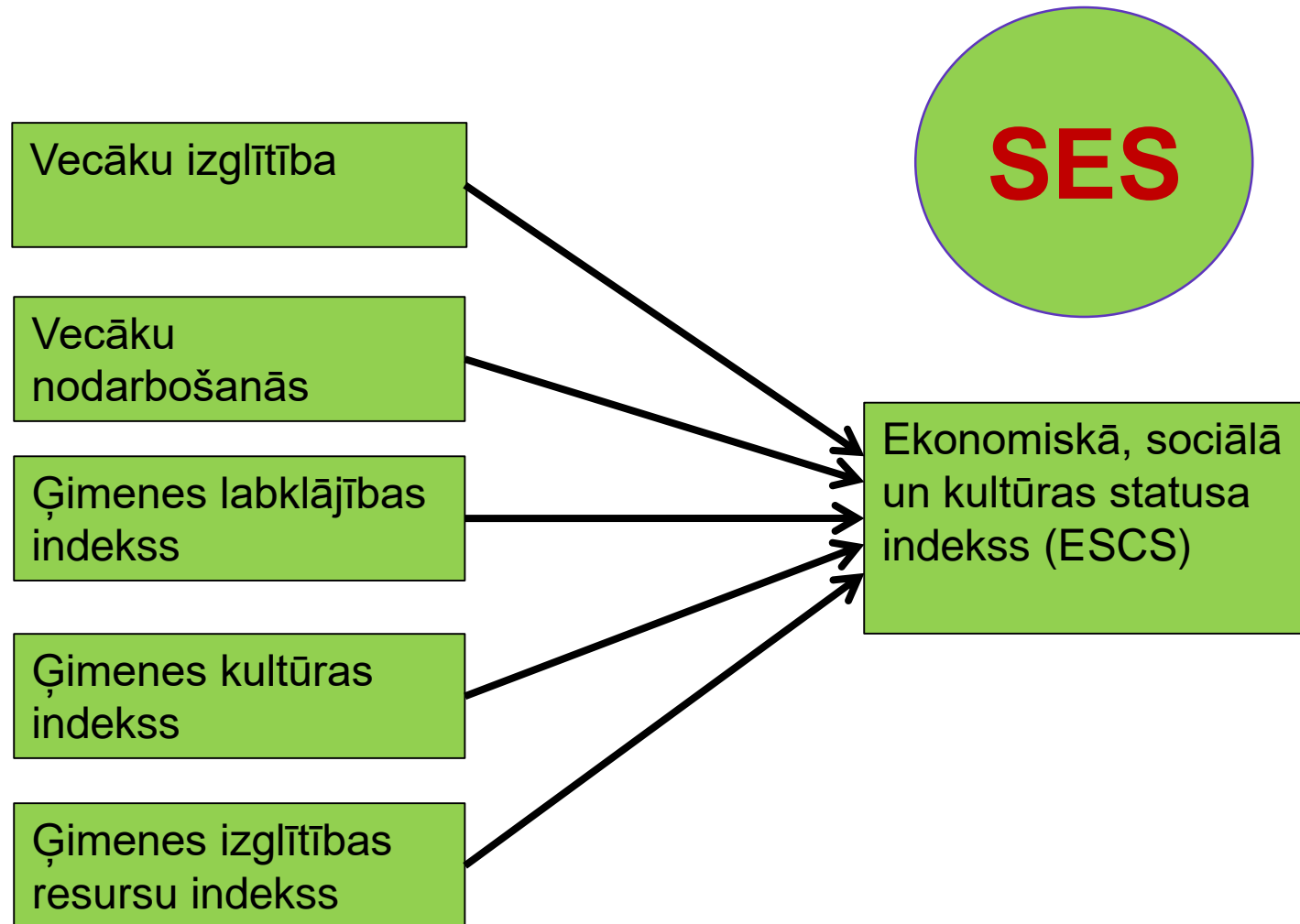
OECD skolēnu sadalījums pa sasniegumiem



Latvija un Igaunija DabZ 2012



Datu analīzes gaitā tiek veidoti dažādi indikatori



KĀ TO PRAKTISKI
APRĒĶINĀT?



- Vienkāršā veidā SPSS neatbalsta šādus aprēķinus
- Jāveido speciāla SPSS programma (skripts)
- 90tie gadi – bija pieejami atsevišķi skripti. Nepārskatāmi, jāveic tekstos izmaiņas
- 2000šie – mēģinājumi veidot SPSS iekļāvumus (Add-ons). Nesaderīgi ar katru nākošo SPSS versiju
- Tagad – labi izveidots IEA IDB Analyzer
- Ir arī jauna R pakete – RREPEST



Please select Statistical Software and choose an Action:

- SPSS
- SAS
- R

Convert Files from SPSS to R

Access the Merge Module

Access the Analysis Module

View the Help Manual

Access the Sample Files

Exit

IEA IDB Analyzer

- Atbalsta visus IEA un OECD pētījumus
- Atbalsta arī dažus citus lielus starptautiskus izglītības pētījumus
- Nav lietojams neatkarīgi veidotu pētījumu izlases kļūdu aprēķiniem
- Darbības princips – tiek ģenerēts atbilstošs SPSS skripts

1 Analysis File: E:\Dokumenti\LIGUMDAR\Pirls2016\Dati_V30\2_Data\SPSS\ASGLVAR4.sav

Select

2

Analysis Type:

PIRLS (Using Student Weights)

Statistic Type:

Percentages and Means

Plausible Value Option:

Use PVs

Number of Decimals:

2

Show Graphs

Yes



3 Select Variables:

Name	Description
ASRLIT01-05	1ST TO 5TH PLAUSIBLE VALUE: LITERARY PURPOSE PV1
ASRINF01-05	1ST TO 5TH PLAUSIBLE VALUE: INFORMATIONAL PURPOS PV1
ASRIIE01-05	1ST TO 5TH PLAUSIBLE VALUE: INTERPRETING PROCESS PV1
ASRRSI01-05	1ST TO 5TH PLAUSIBLE VALUE: STRAIGHTFORWARD PROC PV1

4 Grouping Variables: Exclude Missing From Analysis

Name	Description
IDCNTY	Country ID - Numeric Code
ASBG01	GEN\SEX OF STUDENT

5 Plausible Values:



Name	Description
ASRREA01-05	1ST TO 5TH PLAUSIBLE VALUE: OVERALL READING PV1

6 Weight Variable:



Name	Description
TOTWGT	TOTAL STUDENT WEIGHT

4 Output Files: E:\Dokumenti\LIGUMDAR\Pirls2016\Dati_V30\2_Data\SPSS\aaa.*

Modify

Return to Main Menu

Help

5

Start SPSS

Izveidotais SPSS skripts

```
1 * Script created using the IEA IDB Analyzer (Version 4.0.20).
2 * Created on 15-Apr-18 at 11:20 AM.
3 * Press Ctrl+A followed by Ctrl+R to submit this analysis.
4
5 include file = "C:\Users\A\AppData\Roaming\IEA\IDBAnalyzerV4\bin\Data\Templates\SPSS_Macros\JB_PV.ieasps".
6
7 JB_PV infile="E:\Dokumenti\LIGUMDAR\Pirls2016\Dati_V30\2_Data\SPSS\ASGLVAR4.sav"
8   cvar=IDCNTRY ASBG01 /
9   rootpv=ASRREA0 /
10  tailpv=/
11  npv=5/
12  wgt=TOTWGT/
13  nrwgt=150 /
14  rwgt=/
15  jkz=JKZONE/
16  jkr=JKREP/
17  jk2type=FULL/
18  nomiss=Y/
19  method=JRR/
20  kfac=0/
21  shrtcut=N/
22  viewcod=N/
23  ndec=2/
24  clean = Y/
25  strctry = N/
26  intavg = Y/
27  graphs=Y/
28  selcrit = /
29  selvar = /
30  outdir="E:\Dokumenti\LIGUMDAR\Pirls2016\Dati_V30\2_Data\SPSS/"
31  outfile="aaa".
```

Report

□

Average for PVMATH by CNTRYID ST004D01T

Country Identifier	Student (Standardized) Gender	N of Cases	Sum of W_FSTUWT	Sum of W_FSTUWT (s.e.)	Percent Percent	Percent (s.e.)	PVMATH (Mean)	PVMATH (s.e.)	Std.Dev	Std.Dev. (s.e.)	pctmiss
Estonia	Female	3120	6467	153,64	48,46	,50	506,69	2,47	82,01	1,21	,00
	Male	3272	6878	153,00	51,54	,50	513,01	2,19	87,51	1,40	,00
Finland	Female	4995	28891	665,23	49,01	,41	486,72	2,08	84,23	1,18	,00
	Male	5244	30063	612,46	50,99	,41	481,66	2,29	93,80	1,09	,00
Latvia	Female	2736	8505	137,86	50,53	,59	478,46	2,35	76,31	1,61	,00
	Male	2637	8328	111,63	49,47	,59	487,96	2,31	83,61	1,40	,00
Table Average	Female	.	.	.	49,33	,29	490,62	1,33	80,85	,78	.
	Male	.	.	.	50,67	,29	494,21	1,31	88,31	,75	.

□

**PALDIES PAR
UZMANĪBU!**



