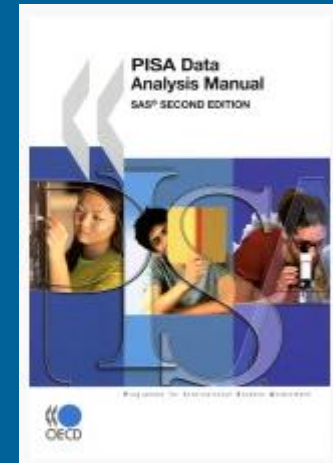


Computation of Standard Errors for Multistage Samples

Guide to the PISA Data Analysis Manual

OECD
BETTER POLICIES FOR BETTER LIVES

What is a Standard Error (SE)

- In PISA, as well as in IEA studies, results are based on a sample
  - Published statistics are therefore estimates
    - Estimates of the means, of the standard deviations, of the regression coefficients …
  - The uncertainty due to the sampling process has to be quantified
    - Standard Errors, Confidence Intervals, P Value

Table 2.3*a*

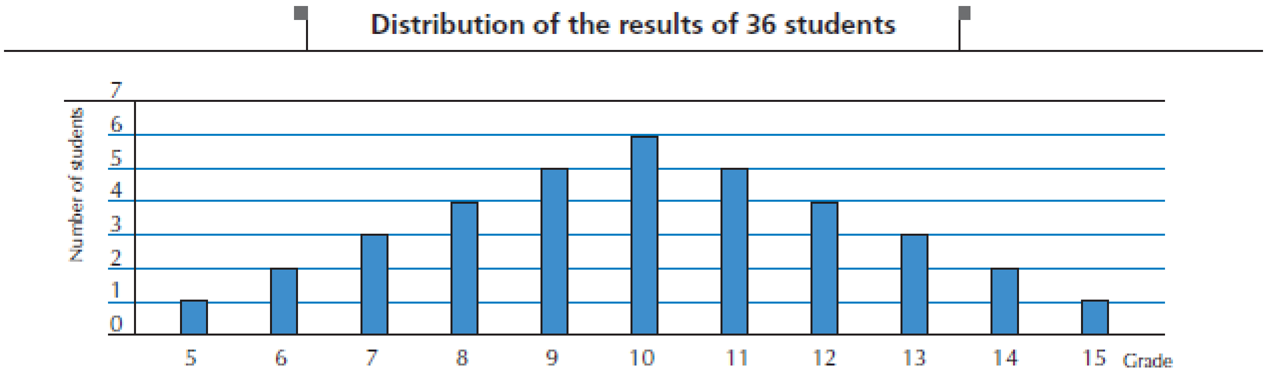
Variation in student performance on the combined reading literacy scale

		Mean		Standard deviation		Percentiles											
						5th		10th		25th		75th		90th		95th	
		Mean score	S.E.	S.D.	S.E.	Score	S.E.	Score	S.E.	Score	S.E.	Score	S.E.	Score	S.E.	Score	S.E.
OECD COUNTRIES	Australia	528	(3.5)	102	(1.6)	354	(4.8)	394	(4.4)	458	(4.4)	602	(4.6)	656	(4.2)	685	(4.5)
	Austria	507	(2.4)	93	(1.6)	341	(5.4)	383	(4.2)	447	(2.8)	573	(3.0)	621	(3.2)	648	(3.7)
	Belgium	507	(3.6)	107	(2.4)	308	(10.3)	354	(8.9)	437	(6.6)	587	(2.3)	634	(2.5)	659	(2.4)
	Canada	534	(1.6)	95	(1.1)	371	(3.8)	410	(2.4)	472	(2.0)	600	(1.5)	652	(1.9)	681	(2.7)
	Czech Republic	492	(2.4)	96	(1.9)	320	(7.9)	368	(4.9)	433	(2.8)	557	(2.9)	610	(3.2)	638	(3.6)
	Denmark	497	(2.4)	98	(1.8)	326	(6.2)	367	(5.0)	434	(3.3)	566	(2.7)	617	(2.9)	645	(3.6)

OECD (2001). *Knowledge and Skills for Life: First Results from PISA 2000*. Paris: OECD.

What is a Standard Error (SE)

- Let us imagine a teacher willing to implement the mastery learning approach, as conceptualized by B.S. Bloom.
- Need to assess students after each lesson
- With 36 students and 5 lessons per day…



What is a Standard Error (SE)

- Description of the population distribution

$$\mu = \frac{1}{N}\sum_{i=1}^{N} x_i = \frac{1}{36}(5+6+6+....+14+14+15) = 10$$

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2 = \frac{1}{36}\sum_{i=1}^{36}(x_i - 10)^2 = \frac{210}{36} = 5.833$$

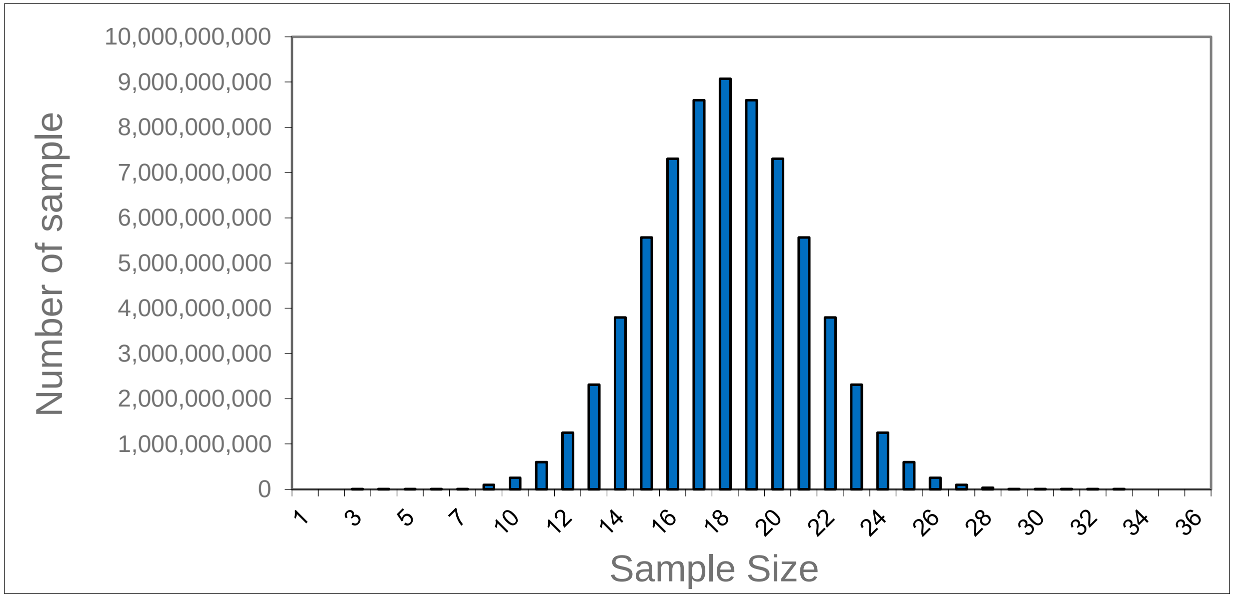
- The teacher decides to randomly draw 2 student's tests for deciding if a remediation is needed
- How many samples of 2 students from a population of 36 students?

Restricted Use - À usage restreint

What is a Standard Error (SE)

- Number of possible sample of size **n** from a population of size **N**

$$\binom{n}{N} = C_N^n = \frac{N!}{(N-n)!n!}$$



- If the thickness of a coin is 1 mm, then 1 billion of coins on the edge corresponds to 1000 km

Restricted Use - À usage restreint

What is a Standard Error (SE)

Description of the 630 possible samples of 2 students selected from 36 students, according to their mean

Sample mean	Results of the two sampled students	Number of combinations of the two results	Number of samples
5.5	5 and 6	2	2
6.0	6 and 6 5 and 7	1 3	4
6.5	5 and 8 6 and 7	4 6	10
7.0	7 and 7 5 and 9 6 and 8	3 5 8	16
7.5	5 and 10 6 and 9 7 and 8	6 10 12	28
8.0	8 and 8 5 and 11 6 and 10 7 and 9	6 5 12 15	38
8.5	5 and 12 6 and 11 7 and 10 8 and 9	4 10 18 20	52
9.0	9 and 9 5 and 13 6 and 12 7 and 11 8 and 10	10 3 8 15 24	60
9.5	5 and 14 6 and 13 7 and 12 8 and 11 9 and 10	2 6 12 20 30	70
10.0	10 and 10 5 and 15 6 and 14 7 and 13 8 and 12 9 and 11	15 1 4 9 16 25	70

What is a Standard Error (SE)

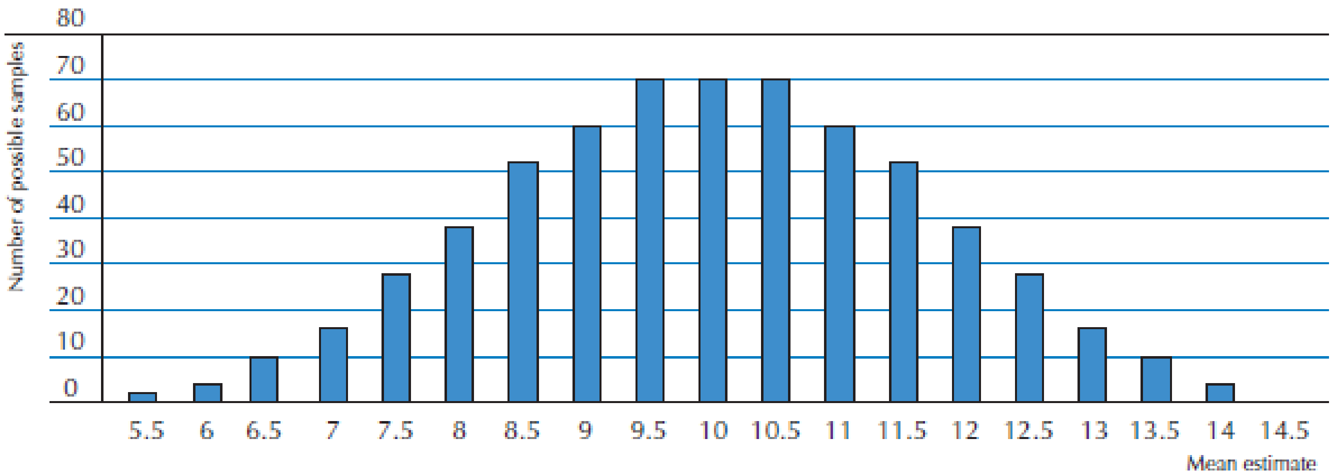
10.5	6 and 15 7 and 14 8 and 13 9 and 12 10 and 11	2 6 12 20 30	70
11.0	7 and 15 8 and 14 9 and 13 10 and 12 11 and 11	3 8 15 24 10	60
11.5	8 and 15 9 and 14 10 and 13 11 and 12	4 10 18 20	52
12.0	9 and 15 10 and 14 11 and 13 12 and 12	5 12 15 6	38
12.5	10 and 15 11 and 14 12 and 13	6 10 12	28
13.0	11 and 15 12 and 14 13 and 13	5 8 2	16
13.5	12 and 15 13 and 14	4 6	10
14.0	13 and 15 14 and 14	3 1	4
14.5	14 and 15	2	2
			630

Restricted Use - À usage restreint

What is a Standard Error (SE)

- Graphical representation of the population mean estimate for all possible samples

Sampling variance distribution of the mean



What is a Standard Error (SE)

- The distribution of sampling variance on the previous slide has:
  - a mean of 10

$$\mu_{(\hat{\mu})} = \frac{\left[(2*5.5)+(4*6)+......(4*14)+(2*14.5)\right]}{630} = 10$$

  - a Standard Deviation (STD) of 1.7

$$\sigma^2_{(\hat{\mu})} = \frac{\left[(5.5-10)^2+(5.5-10)^2+(6-10)^2+.......(14.5-10)^2\right]}{630}$$

$$\sigma_{(\hat{\mu})} = \sqrt{\frac{1785}{630}} = 1.68$$

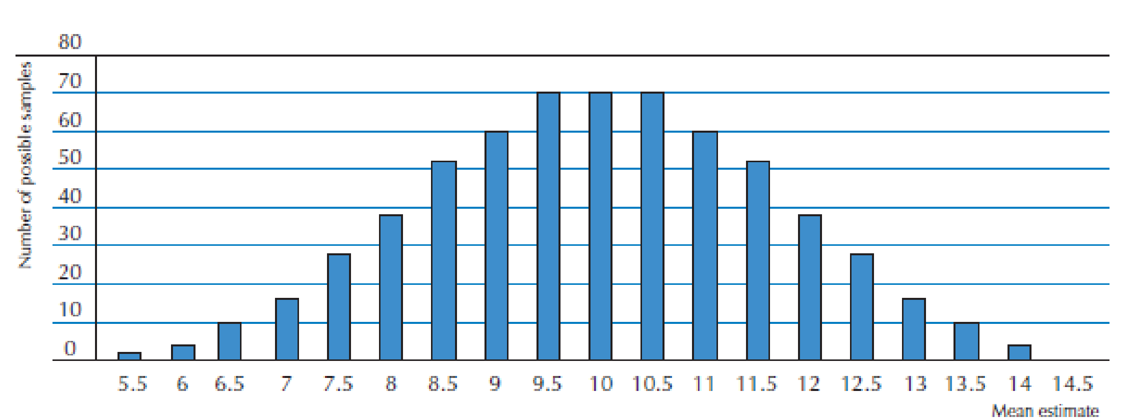
  - The STD of a sampling distribution is denoted Standard Error (SE)

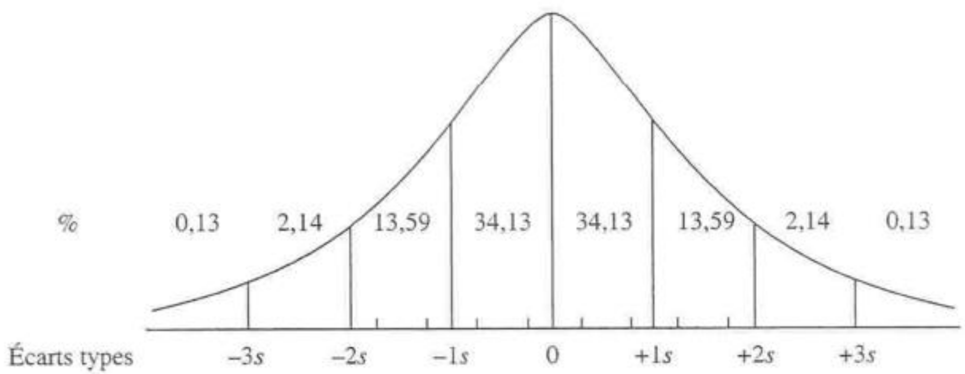
Restricted Use - À usage restreint

What is a Standard Error (SE)

- The sampling distribution on the mean looks like a normal distribution

Sampling variance distribution of the mean





Restricted Use - À usage restreint

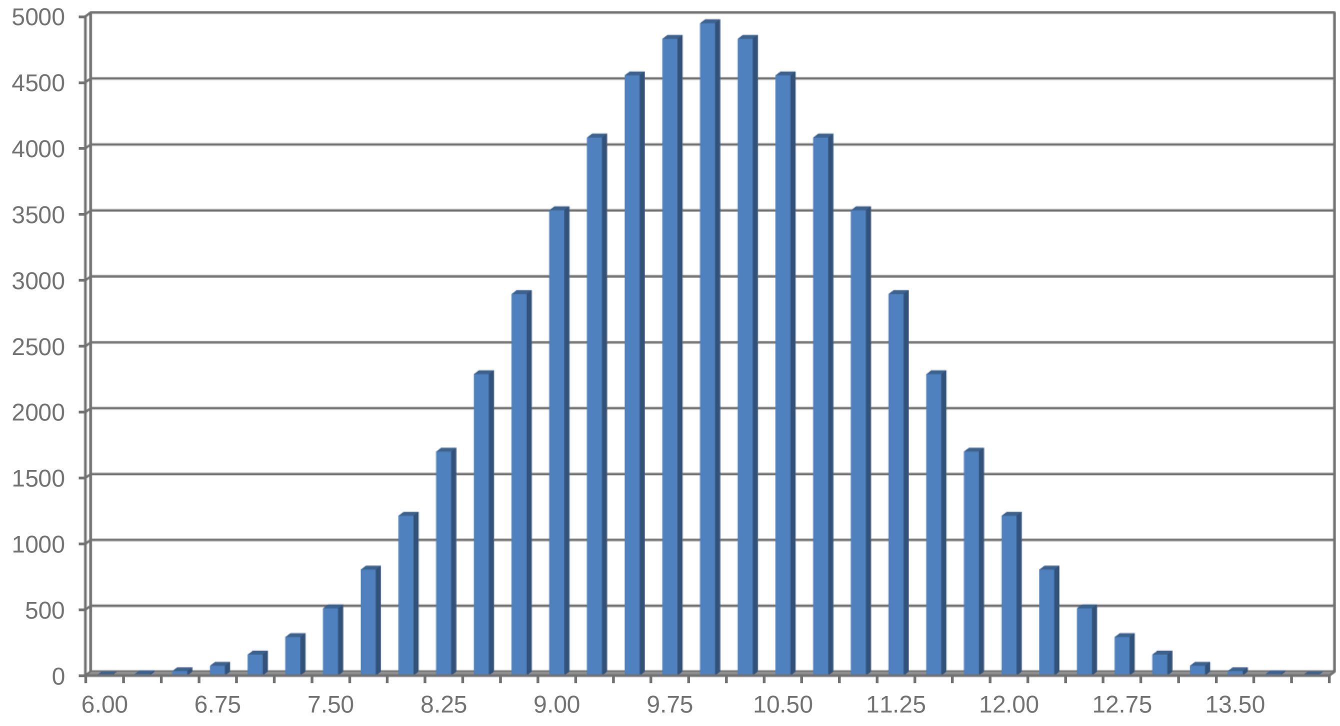
What is a Standard Error (SE)

- Let us count the number of samples with a mean included between
  - [(10-1.96SE);(10+1.96SE)]
  - [(10-3.30);(10+3.30)]
  - [6.70;13.30]
  - There are: 6+28+38+52+60+70+70+70+60+52 +38+28+16=598 samples, thus 94.9 % of all possible samples
- With a population *N*(10, 5, 83), 95% of all possible samples of size 2 will have a population mean estimate included between 6.70 and 13.30

Restricted Use - À usage restreint

What is a Standard Error (SE)

- Sampling distribution of the mean estimates of all possible samples of size 4



Restricted Use - À usage restreint

What is a Standard Error (SE)

- The distribution of sampling variance on the previous slide has:
  - a mean of 10

$$\mu_{(\hat{\mu})} = \frac{[(3*6)+(10*6.25)+......(10*13.75)+(3*14)]}{58905} = 10$$

  - a Standard Deviation of 1.7

$$\sigma^2_{(\hat{\mu})} = \frac{[(6-10)^2+(6-10)^2+(6-10)^2+.......(14-10)^2]}{58905} = 1.335$$

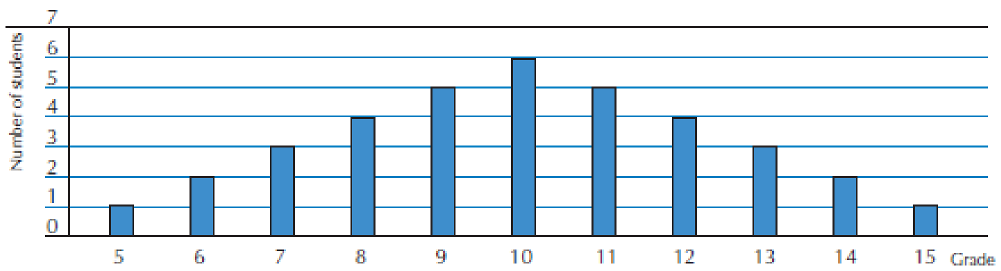
$$\sigma_{(\hat{\mu})} = 1.155$$

Restricted Use - À usage restreint

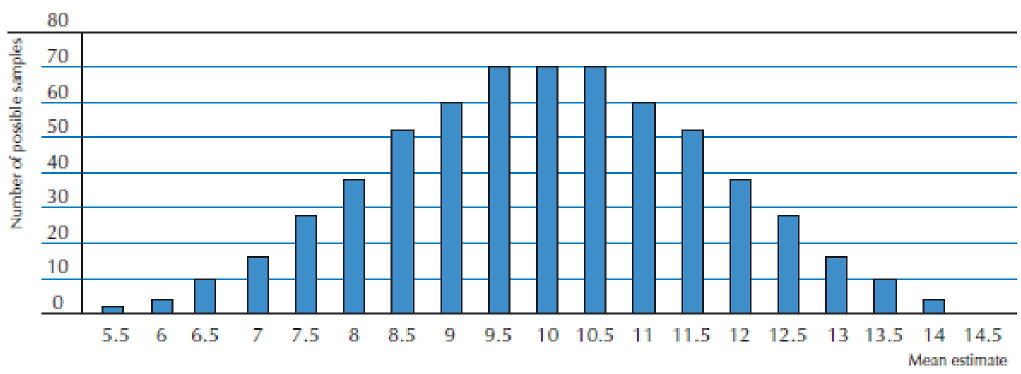
What is a Standard Error (SE)

- Distribution of the scores versus distribution of the mean estimates

Distribution of the results of 36 students



Sampling variance distribution of the mean



What is a Standard Error (SE)

- The sampling variance of the mean is inversely proportional to the sample size:
  - If two students are sampled, then the smallest possible mean is 5.5 and the highest possible mean is 14.5
  - If four students are sampled, it ranges from 6 to 14
  - If 10 students are sampled, it ranges from 7 to 13
- The sampling variance is proportional to the variance:
  - If the score are reported on 20, with a sample of size 2, then it ranges from 5,5 to 14,5
  - If the score are reported on 40 (multiplied by 2) then it ranges from 11 to 29

Restricted Use - À usage restreint

What is a Standard Error (SE)

	Individual 1	Individual 2	Individual 3	Individual 4	Mean estimates
Sample 1	X_{11}	X_{21}	X_{31}	X_{41}	$\hat{u}_1$
Sample 2	X_{12}	X_{22}	X_{32}	X_{42}	$\hat{u}_2$
Sample 3	X_{13}	X_{23}	X_{33}	X_{43}	$\hat{u}_3$
Sample 4	X_{14}	X_{24}	X_{34}	X_{44}	$\hat{u}_4$
Sample 5	X_{15}	X_{25}	X_{35}	X_{45}	$\hat{u}_5$
Sample 6	X_{16}	X_{26}	X_{36}	X_{46}	$\hat{u}_6$
Sample 7	X_{17}	X_{27}	X_{37}	X_{47}	$\hat{u}_7$
Sample 8	X_{18}	X_{28}	X_{38}	X_{48}	$\hat{u}_8$
Sample 9	X_{19}	X_{29}	X_{39}	X_{49}	$\hat{u}_9$
……					
Sample X	X_{1x}	X_{2x}	X_{3x}	X_{4x}	$\hat{u}_x$

$$\sigma^2_{(\hat{\mu})} = \sigma^2_{(\frac{1}{n}\sum_{i=1}^{n} X_i)}$$

Restricted Use - À usage restreint

What is a Standard Error (SE)

$$\sigma^2_{(\hat{u})} = \sigma^2_{(\frac{1}{n}\sum_{i=1}^{n} X_i)}$$

$$\sigma^2_{(a.X)} = a^2.\sigma^2_{(X)}$$

$$\sigma^2_{(\hat{u})} = \frac{1}{n^2}\sigma^2_{(\sum_{i=1}^{n} X_i)}$$

$$\sigma^2_{(A+B)} = \sigma^2_{(A)} + \sigma^2_{(B)} + 2\,\mathrm{cov}(A,B)$$

$$\sigma^2_{(\hat{u})} = \frac{1}{n^2}\sum_{i=1}^{n}\sigma^2_{(X_i)} + \sum_{i=1}^{n-1}\sum_{j=i+1}^{n} 2\,\mathrm{cov}(X_i, X_j)$$

$$\sigma^2_{(\hat{u})} = \frac{1}{n^2} n\sigma^2 = \frac{\sigma^2}{n}$$

$$SE = \sigma_{(\hat{u})} = \sqrt{\sigma^2_{(\hat{u})}} = \sqrt{\frac{\sigma^2}{n}}$$

Restricted Use - À usage restreint

What is a Standard Error (SE)

- As we don't know the variance in the population, the SE for a mean of θ as obtained from a sample is calculated as

$$\hat{\sigma}^2_{(\hat{\mu})} = \frac{\hat{\sigma}^2}{n} \qquad \hat{\sigma}_{(\mu)} = \frac{\hat{\sigma}_{\hat{\mu}}}{\sqrt{n}}$$

- similarly, the SE for a percentage P is calculated as

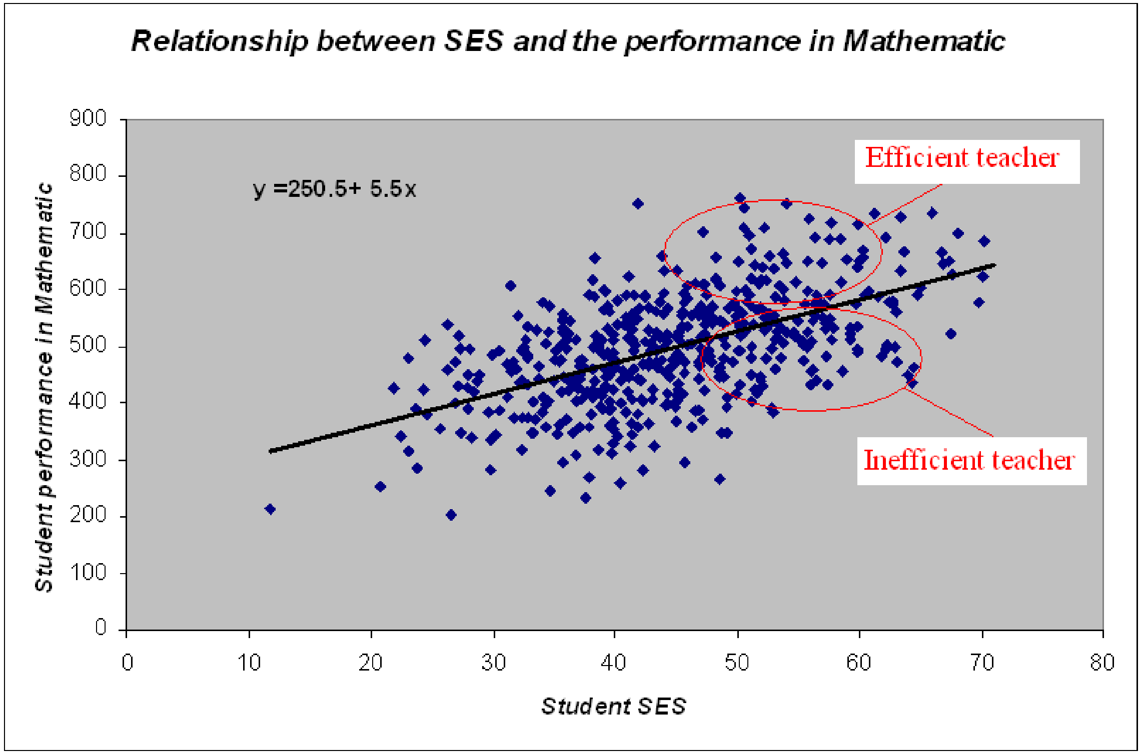
$$SE_{\hat{P}} = \sqrt{\frac{PQ}{n}}$$

What is a Standard Error (SE)

- Linear regression assumptions:
  - Homoscedasticity
    - the variance of the error terms is constant for each value of x
  - Linearity
    - the relationship between each X and Y is linear
  - Error Terms are normally distributed
  - Independence of Error Terms
    - successive residuals are not correlated
    - If not, the SE of regression coefficients is biased

Restricted Use - À usage restreint

What is a Standard Error (SE)



- With a multistage sample design, errors will be correlated

Standard Errors for multistage samples

- Multistage samples are usually implemented in International Surveys in Education:
  - schools (PSU=primary sampling units)
  - classes
  - students
- If schools/classes / students are considered as infinite populations and if units are selected according to a SRS procedures, then:

$$\sigma^2_{(\hat{\mu})} = \frac{\sigma^2_{sch}}{n_{sch}} + \frac{\sigma^2_{cla|sch}}{n_{sch}.n_{Cla|sch}} + \frac{\sigma^2_{stu|cla|sch}}{n_{sch}.n_{Cla|sch}.n_{stu|cla|sch}}$$

Restricted Use - À usage restreint

Standard Errors for multistage samples

- PISA:
  - 2 stage samples : schools and then students

$$\sigma^2_{(\hat{\mu})} = \frac{\sigma^2_{sch}}{n_{sch}} + \frac{(\sigma^2_{cla|sch} + \sigma^2_{stu|cla})}{n_{sch} n_{stu/sch}} \longrightarrow \rho = \frac{\sigma^2_B}{\sigma^2_B + \sigma^2_W}$$

- IEA:
  - 2 stage samples : schools and then 1 class per selected school

$$\sigma^2_{(\hat{\mu})} = \frac{(\sigma^2_{sch} + \sigma^2_{cla|sch})}{n_{sch}.1} + \frac{\sigma^2_{stu|cla}}{n_{sch}.n_{stu/cla}} \longrightarrow \rho = \frac{\sigma^2_B}{\sigma^2_B + \sigma^2_W}$$

Restricted Use - À usage restreint

Standard Error for multistage samples

- Three fictitious examples in PISA

$$\sigma^2_{(\hat{\mu})} = \frac{\sigma^2_{sch}}{150} + \frac{\sigma^2_{stud|sch}}{(150).(35)} = \frac{1000}{150} + \frac{9000}{5250} = 6.66 + 1.71 = 8.38$$

$$\sigma^2_{(\hat{\mu})} = \frac{\sigma^2_{sch}}{150} + \frac{\sigma^2_{stud|sch}}{(150).(35)} = \frac{3000}{150} + \frac{7000}{5250} = 20 + 1.33 = 21.33$$

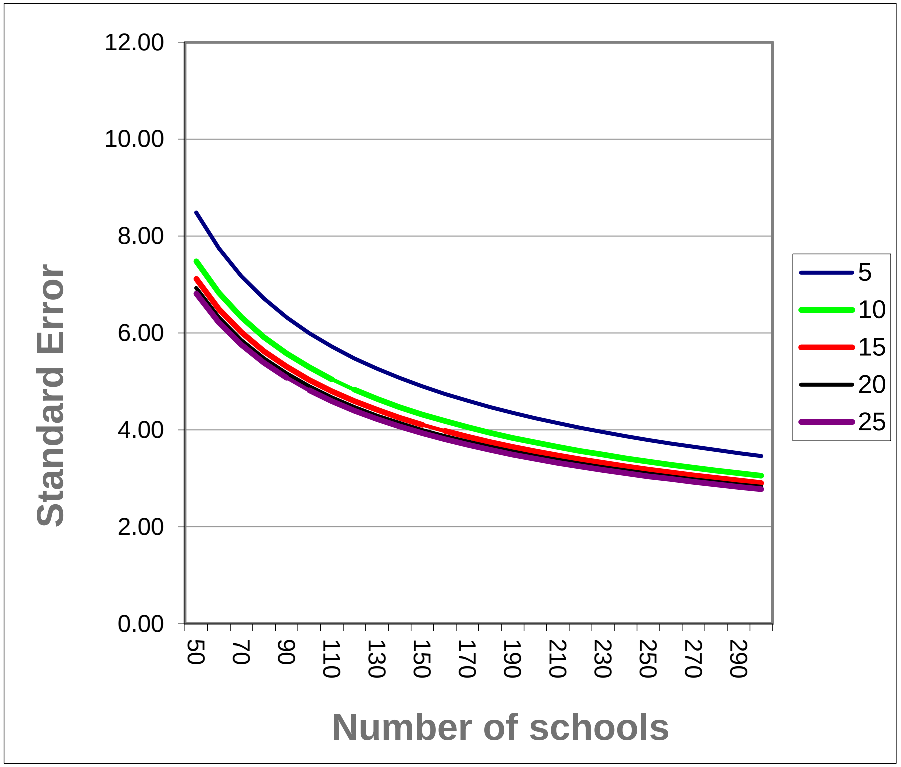
$$\sigma^2_{(\hat{\mu})} = \frac{\sigma^2_{sch}}{150} + \frac{\sigma^2_{stud|sch}}{(150).(35)} = \frac{6000}{150} + \frac{4000}{5250} = 40 + 0.76 = 40.76$$

- If considered as a SRS or random assignment to schools

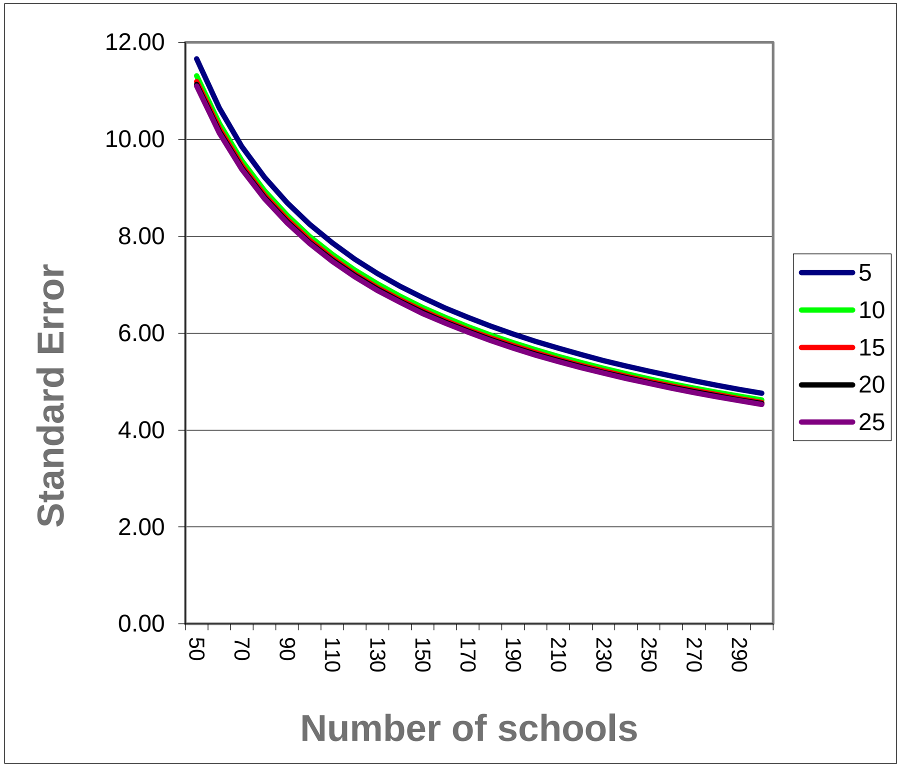
$$\sigma^2_{(\hat{\mu})} = \frac{\sigma^2}{n} = \frac{10000}{5250} = 1.90$$

Standard Error for multistage samples

$$\sigma^2 = 10000$$
$$\rho = 0.20$$



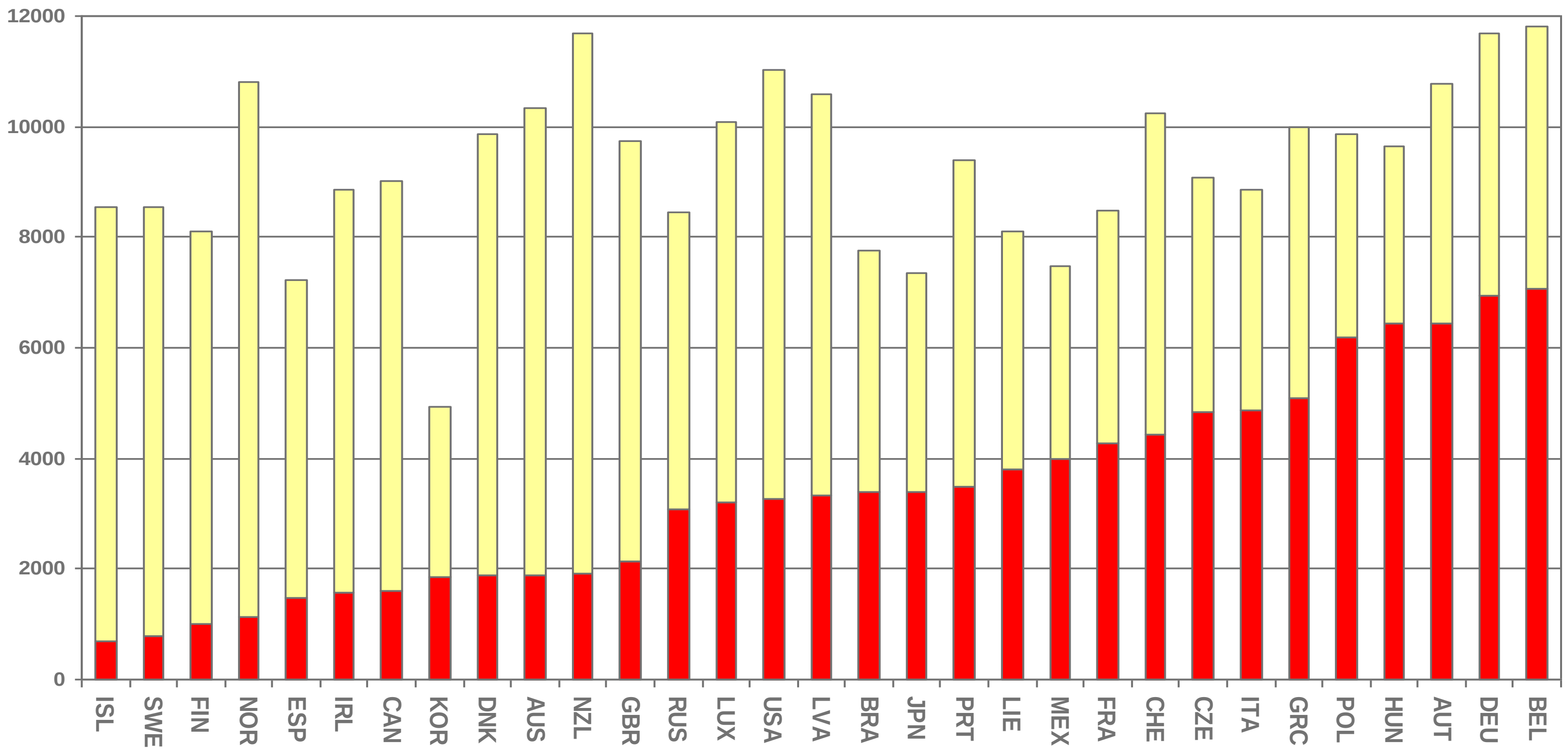
$$\sigma^2 = 10000$$
$$\rho = 0.60$$



Restricted Use - À usage restreint

Standard Error for multistage samples

Variance Decomposition for Reading Literacy in PISA 2000



Restricted Use - À usage restreint

Standard Error for multistage samples

- Impact of the stratification variables on sampling variance

$$\hat{\mu} = \frac{(N_1\hat{\mu}_1) + (N_2\hat{\mu}_2)}{N}$$

$$\sigma^2_{(\hat{\mu})} = \sigma^2\left[\frac{(N_1\mu_1) + (N_2\mu_2)}{N}\right]$$

N_1, N_2 considered as constant

$$= \frac{N_1^2\sigma^2_{(\mu_1)} + N_2^2\sigma^2_{(\mu_2)} + 2\operatorname{cov}(N_1\hat{\mu}_1, N_2\hat{\mu}_2)}{N^2}$$

Independent samples so COV=0

$$= \frac{N_1^2\sigma^2_{(\mu_1)} + N_2^2\sigma^2_{(\mu_2)}}{N^2}$$

Restricted Use - À usage restreint

Standard Error for multistage samples

Effect	*Sum of Squares*	*Degree of Freedom*	*Mean square*
Gender (50F+50G)	2500	L-1 (1)	2500
ERROR	7500	N-L (98)	76.53
TOTAL	10000	N-1 (99)	101.01

ERROR mean square → σ_W^2

TOTAL mean square → σ^2

$$\sigma_{(\hat{\mu})} = \sqrt{\frac{\sigma^2}{n}} = \sqrt{\frac{101.01}{100}} = 1.005$$

$$\sigma^2_{\left(\frac{\hat{\mu}_F + \hat{\mu}_M}{2}\right)} = \frac{\sigma^2_{(\hat{\mu}_F)} + \sigma^2_{(\hat{\mu}_M)}}{4} = \frac{\frac{76.53}{50} + \frac{76.53}{50}}{4} = 0.7653$$

$$\sigma_{(\hat{\mu})} = \sqrt{0,7653} = 0.875$$

Restricted Use - À usage restreint

Standard Error for multistage samples

School and within school variances of the student performance in reading, intraclass correlation with and without control of the explicit stratification variables (OECD, PISA 2000 database)

Country	School variance	Within school variance	School variance under control of stratification	Rho	Rho under control of stratification
AUT	6356	4243	624	0.60	0.13
BEL	7050	4724	3489	0.60	0.42
CHE	4517	5909	3119	0.43	0.35
CZE	4812	4203	604	0.53	0.13
DNK	1819	7970	1696	0.19	0.18
ESP	1477	5649	823	0.21	0.13
FIN	998	7096	869	0.12	0.11
FRA	4181	4219	910	0.50	0.18
GBR	2077	7637	1990	0.21	0.21
GRC	4995	4907	3619	0.50	0.42
HUN	6604	3230	4638	0.67	0.59
IRL	1589	7349	1495	0.18	0.17
ISL	652	7884	563	0.08	0.07
ITA	4719	4028	2031	0.54	0.34

Restricted Use - À usage restreint

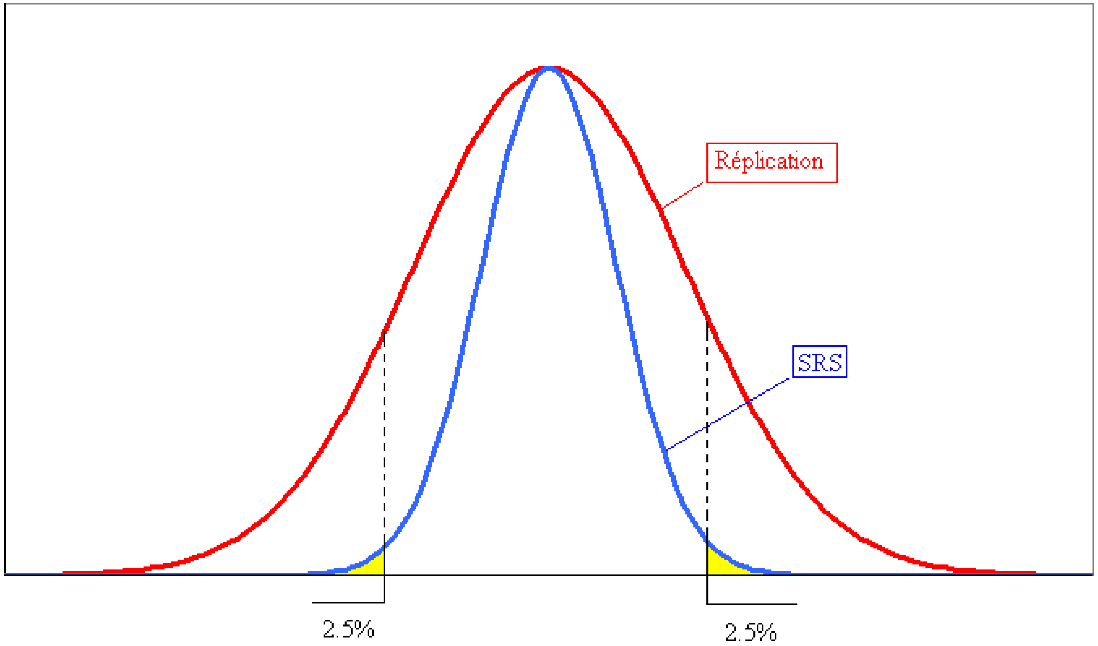
Standard Error for multistage samples

- Consequences of considering PISA samples as simple random samples
  - In most cases, underestimation of the sampling variance estimates
    - Non significant effect will be reported as significant
    - How can we measure the risk?
    - Computation of the Type I error

Restricted Use - À usage restreint

Standard Error for multistage samples

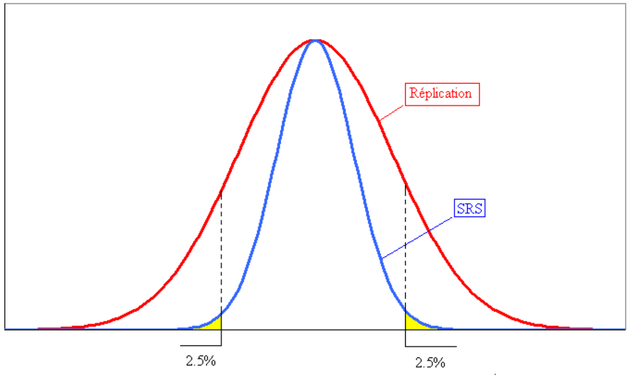
- Consequences : Type I error underestimation



Standard Error for multistage samples

- Consequences : Type I error underestimation

	Sampling Variance	Standard Error	Ratio	Ratio x Z score	Type I Error
Unbiased estimate	24	4.90			
Biased estimate	20	4.47	0.91	1.79	0.07
	16	4.00	0.82	1.60	0.11
	12	3.46	0.71	1.38	0.17
	8	2.83	0.58	1.13	0.26
	4	2.00	0.41	0.80	0.42



Standard Error for multistage samples

Sampling Design Effect

$$SDE = \frac{\sigma^2_{(\hat{\theta}_{reel})}}{\sigma^2_{(\hat{\theta}_{SRS})}}$$

Design effect and type I errors

Design effect (coefficient of increase)	Type I error	Design effect (coefficient of increase)	Type I error
1.5	0.11	11.0	0.55
2.0	0.17	11.5	0.56
2.5	0.22	12.0	0.57
3.0	0.26	12.5	0.58
3.5	0.29	13.0	0.59
4.0	0.33	13.5	0.59
4.5	0.36	14.0	0.60
5.0	0.38	14.5	0.61
5.5	0.40	15.0	0.61
6.0	0.42	15.5	0.62
6.5	0.44	16.0	0.62
7.0	0.46	16.5	0.63
7.5	0.47	17.0	0.63
8.0	0.49	17.5	0.64
8.5	0.50	18.0	0.64
9.0	0.51	18.5	0.65
9.5	0.52	19.0	0.65
10.0	0.54	19.5	0.66
10.5	0.55	20.0	0.66

Restricted Use - À usage restreint

Standard Error for multistage samples

- Sampling design effect in PISA 2000 Reading

Country	SDE	Type I	Country	SDE	Type I
Australia	5.90	0.42	Korea	5.89	0.42
Austria	3.10	0.27	Latvia	**10.16**	0.54
Belgium	7.31	0.47	Liechtenstein	0.48	0.00
Brazil	6.14	0.43	Luxembourg	0.73	0.02
Canada	9.79	0.53	Mexico	6.69	0.45
Czech Republic	3.18	0.27	Netherlands	3.52	0.30
Denmark	2.36	0.20	New Zealand	2.40	0.21
Finland	3.90	0.32	Norway	2.97	0.26
France	4.02	0.33	Poland	7.12	0.46
Germany	2.36	0.20	Portugal	9.72	0.53
Greece	**12.04**	0.57	Russian Federation	**13.53**	0.59
Hungary	8.64	0.50	Spain	6.18	0.43
Iceland	0.73	0.02	Sweden	2.32	0.20
Ireland	4.50	0.36	Switzerland	**10.52**	0.55
Italy	1.90	0.16	United Kingdom	5.97	0.42
Japan	**19.28**	0.66	United States	**17.29**	0.64

Restricted Use - À usage restreint

Standard Error for multistage samples

- Factors influencing the SE other than the sample size
  - School Variance: depending on the variable
    - Usually high for performance
    - Low for other variables
  - Efficiency of the stratification variables
    - A stratification variable can be efficient for some variables and not for others
  - Population parameter estimates

Restricted Use - À usage restreint

Standard Error for multistage samples

- A few examples (PISA2000, Belgium)
  - Intraclass correlation
    - Performance in reading: Rho=0.60 SDE=7.19
    - Social Background (HISEI): Rho= 0.24 SDE=3.45
    - Enjoyment for Reading: Rho=0.10 SDE=1.86
  - Statistics
    - Regression analyses: Reading = HISEI + GENDER
      - Intercept SDE= 5.50
      - HISEI SDE=3.78
      - GENDER SDE=3.91
    - Logistic regression Level (0/1 Reading below or above 500) =HISEI
      - Intercept SDE= 2.39
      - HISEI SDE=2.09

Restricted Use - À usage restreint

Standard Error for multistage samples

- Very few mathematical solutions for the estimation of the sampling variance for multistage samples
  - For mean estimates under the condition
    - Simple Random Sample (SRS) and stratified
    - Probability Proportional to Size (PPS) sample but with no stratification variables
  - No mathematical solutions for other statistics
  - **Use of replication methodologies for the estimation of sampling variance**
- For SRS:
  - Jackknife: *n* replications de *n-1* cases
  - Bootstrap : an infinite number of samples of *n* cases randomly drawn with replacement.

Restricted Use - À usage restreint

Replication methods for SRS

- *Jackknife* for SRS

$$\hat{\sigma}^2_{jack} = \frac{n-1}{n}\sum_{i=1}^{n}(\hat{\theta}_{(i)} - \hat{\bar{\theta}}_{(i)})^2$$

$$\hat{\sigma}^2_{jack} = \frac{n-1}{n}\sum_{i=1}^{n}(\hat{\theta}_{(i)} - \hat{\theta})^2$$

Student	1	2	3	4	5	6	7	8	9	10	Mean
Value	10	11	12	13	14	15	16	17	18	19	14.50
Replication 1	0	1	1	1	1	1	1	1	1	1	15.00
Replication 2	1	0	1	1	1	1	1	1	1	1	14.88
Replication 3	1	1	0	1	1	1	1	1	1	1	14.77
Replication 4	1	1	1	0	1	1	1	1	1	1	14.66
Replication 5	1	1	1	1	0	1	1	1	1	1	14.55
Replication 6	1	1	1	1	1	0	1	1	1	1	14.44
Replication 7	1	1	1	1	1	1	0	1	1	1	14.33
Replication 8	1	1	1	1	1	1	1	0	1	1	14.22
Replication 9	1	1	1	1	1	1	1	1	0	1	14.11
Replication 10	1	1	1	1	1	1	1	1	1	0	14.00

Restricted Use - À usage restreint

Replication methods for SRS

- *Jackknife* for SRS
  - Estimation of the SE by replication

$$\hat{\sigma}^2_{jack} = \frac{n-1}{n}\sum_{i=1}^{n}(\hat{\theta}_{(i)} - \hat{\theta})^2$$

$$\sigma^2_{(\hat{\mu})} = \frac{9}{10}\left[(15.00 - 14.50)^2 + (14.88 - 14.50)^2 + ... + (15.11 - 14.50)^2 + (14.00 - 14.50)^2\right]$$

$$\sigma^2_{(\hat{\mu})} = \frac{9}{10}(1.018519) = 0.9167$$

  - Estimation by using the mathematical formula

$$\sigma^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \hat{\mu}\right)^2 = \frac{1}{9}\left[(10-14.5)^2 + (11-14.5)^2 + ... + (18-14.5)^2 + (19-14.5)^2\right] = 9.17$$

$$\sigma^2_{(\hat{\mu})} = \frac{\sigma^2}{n} = \frac{9.17}{10} = 0.917$$

Replication methods for SRS

$$\hat{\mu}_{(i)} - \hat{\mu} = \frac{\left[\left(\sum_{i=1}^{n} x_i\right) - x_i\right]}{(n-1)} - \frac{\left[\sum_{i=1}^{n} x_i\right]}{n} = \left(\frac{-x_i}{(n-1)} + \frac{\left[\sum_{i=1}^{n} x_i\right]}{(n-1)}\right) - \frac{\left[\sum_{i=1}^{n} x_i\right]}{n} = -\frac{x_i}{n-1} + \left[\sum_{i=1}^{n} x_i\right]\left[\frac{1}{n-1} - \frac{1}{n}\right]$$

$$= -\frac{x_i}{n-1} + \left[\sum_{i=1}^{n} x_i\right]\left[\frac{1}{(n-1)}\left(1 - \frac{(n-1)}{n}\right)\right] = -\frac{1}{(n-1)}\left[x_i - \left(\sum_{i=1}^{n} x_i\right)\left(\frac{n}{n} - \frac{(n-1)}{n}\right)\right]$$

$$= -\frac{1}{(n-1)}\left[(x_i - \hat{\mu})(n-(n-1))\right] = -\frac{1}{(n-1)}\left[(x_i - \hat{\mu})(n-n+1)\right] = -\frac{1}{(n-1)}(x_i - \hat{\mu})$$

$$\Rightarrow \left(\hat{\mu}_{(i)} - \hat{\mu}\right)^2 = \frac{1}{(n-1)^2}(x_i - \hat{\mu})^2$$

$$\Rightarrow \sum_{i=1}^{n}\left(\hat{\mu}_{(i)} - \hat{\mu}\right)^2 = \frac{1}{(n-1)^2}\sum_{i=1}^{n}(x_i - \hat{\mu})^2 = \frac{1}{(n-1)}\frac{\sum_{i=1}^{n}(x_i - \hat{\mu})^2}{(n-1)} = \frac{1}{(n-1)}\hat{\sigma}^2$$

$$\Rightarrow \sigma_{jack}^2 = \frac{(n-1)}{n}\sum_{i=1}^{n}\left(\hat{\mu}_{(i)} - \hat{\mu}\right)^2 = \frac{(n-1)}{n}\frac{1}{(n-1)}\hat{\sigma}^2 = \frac{\hat{\sigma}^2}{n}$$

Restricted Use - À usage restreint

Replication methods for SRS

- *Bootstrap* for SRS

$$\hat{\sigma}^2_{boot} = \frac{1}{G-1}\sum_{i=1}^{n}(\hat{\theta}_{(i)} - \bar{\hat{\theta}}_{(i)})^2$$

Student	1	2	3	4	5	6	7	8	9	10	Mean
Value	10	11	12	13	14	15	16	17	18	19	14.50
Structure 1	1	1	1	1	1	1	1	1	1	1	14.50
Structure 2	2	1	1	1	1	1	1	1	1	0	From 13.7 to 15.4
Structure 3	2	2	1	1	1	1	1	1	0	0	From 12.9 to 16.1
Structure	5	1	1	1	1	1	0	0	0	0	
Structure	6	1	1	1	1	0	0	0	0	0	
Structure	7	1	1	1	0	0	0	0	0	0	
Structure	8	1	1	0	0	0	0	0	0	0	
Structure	9	1	0	0	0	0	0	0	0	0	
Structure	10	0	0	0	0	0	0	0	0	0	From 10 to 19

Restricted Use - À usage restreint

Replication methods for multistage sample

- *Jackknife* for unstratified Multistage Sample

Replicate	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10
School 1	0.00	1.11	1.11	1.11	1.11	1.11	1.11	1.11	1.11	1.11
School 2	1.11	0.00	1.11	1.11	1.11	1.11	1.11	1.11	1.11	1.11
School 3	1.11	1.11	0.00	1.11	1.11	1.11	1.11	1.11	1.11	1.11
School 4	1.11	1.11	1.11	0.00	1.11	1.11	1.11	1.11	1.11	1.11
School 5	1.11	1.11	1.11	1.11	0.00	1.11	1.11	1.11	1.11	1.11
School 6	1.11	1.11	1.11	1.11	1.11	0.00	1.11	1.11	1.11	1.11
School 7	1.11	1.11	1.11	1.11	1.11	1.11	0.00	1.11	1.11	1.11
School 8	1.11	1.11	1.11	1.11	1.11	1.11	1.11	0.00	1.11	1.11
School 9	1.11	1.11	1.11	1.11	1.11	1.11	1.11	1.11	0.00	1.11
School 10	1.11	1.11	1.11	1.11	1.11	1.11	1.11	1.11	1.11	0.00

- Each replicate= contribution of a school

Restricted Use - À usage restreint

Replication methods for SRS

- *Jackknife* for stratified Multistage Sample

Pseudo-stratum	School		R1	R2	R3	R4	R5
1	1		2	1	1	1	1
1	2		0	1	1	1	1
2	3		1	0	1	1	1
2	4		1	2	1	1	1
3	5		1	1	2	1	1
3	6		1	1	0	1	1
4	7		1	1	1	0	1
4	8		1	1	1	2	1
5	9		1	1	1	1	2
5	10		1	1	1	1	0

- Each replicate= contribution of a pseudo stratum

Restricted Use - À usage restreint

Replication methods for multistage sample

- *Balanced Replicated Replication*

Pseudo-stratum	School		R1	R2	R3	R4	R5	R6	R7	R8
1	1		2	2	2	2	2	2	2	2
1	2		0	0	0	0	0	0	0	0
2	3		2	0	2	0	2	0	2	0
2	4		0	2	0	2	0	2	0	2
3	5		2	2	0	0	2	2	0	0
3	6		0	0	2	2	0	0	2	2
4	7		2	0	0	2	2	0	0	2
4	8		0	2	2	0	0	2	2	0
5	9		2	2	2	2	0	0	0	0
5	10		0	0	0	0	2	2	2	2

- Each replicate= an estimate of the sampling variance

Restricted Use - À usage restreint

Replication methods for multistage sample

- How to form the pseudo-strata, i.e. how to pair schools?

ID	Size	From	To	SAMPLED
1	15	1	15	1
2	20	26	35	0
3	25	36	60	0
4	30	61	90	0
5	35	91	125	1
6	40	126	165	0
7	45	166	210	0
8	50	211	260	1
9	60	261	320	0
10	80	321	400	1
Total	400			

– Within explicit strata, with a systematic sampling procedure, schools are sequentially selected.

– Pairs are formed according to the sequence
  - School 1 with School 5
  - School 8 with School 10

Replication methods for multistage sample

- How to form the pseudo-strata, i.e. how to pair schools?

IEA TIMSS / PIRLS procedure		
ID	Participation	Pseudo-Stratum
14	1	1
21	1	1
35	1	2
56	0	
78	1	2
99	1	3
103	0	
115	1	3
126	1	4
137	1	4

PISA procedure		
ID	Participation	Pseudo-Stratum
14	1	1
21	1	1
35	1	2
56	0	2
78	1	3
99	1	3
103	0	4
115	1	4
126	1	5
137	1	5

Restricted Use - À usage restreint

Replication methods for multistage sample

- Balanced Replicated Replication
  - With L pseudo-strata, there are 2^L possible combinations
  - If 4 strata, then 16 combinations
  - Same efficiency with an Hadamard Matrix of Rank 4

	Stratum 1	Stratum 2	Stratum 3	Stratum 4
1	1	1	1	1
2	1	1	1	2
3	1	1	2	1
4	1	2	1	1
5	2	1	1	1
6	1	1	2	2
7	1	2	1	2
8	2	1	1	2
9	1	2	2	1
10	2	1	2	1
11	2	2	1	1
12	1	2	2	2
13	2	1	2	2
14	2	2	1	2
15	2	2	2	1
16	2	2	2	2

Restricted Use - À usage restreint

Replication methods for multistage sample

- *Hadamard* Matrix

Combination				
1	1	1	1	1
2	1	-1	1	-1
3	1	1	-1	-1
4	1	-1	-1	1

- Each row is orthogonal to all other rows, i.e. the sum of the products is equal to 0.
- Selection of school according to this matrix

Restricted Use - À usage restreint

Replication methods for multistage sample

$$H_{2n} = \left[\frac{H_n}{H_n} \frac{H_n}{-H_n} \right]$$

$$H_2 = \begin{bmatrix} +1 & +1 \\ +1 & -1 \end{bmatrix}$$

$$H_4 = \begin{bmatrix} +1 & +1 & +1 & +1 \\ +1 & -1 & +1 & -1 \\ +1 & +1 & -1 & -1 \\ +1 & -1 & -1 & +1 \end{bmatrix}$$

$$H_8 = \begin{vmatrix} +1 & +1 & +1 & +1 & +1 & +1 & +1 & +1 \\ +1 & -1 & +1 & -1 & +1 & -1 & +1 & -1 \\ +1 & +1 & -1 & -1 & +1 & +1 & -1 & -1 \\ +1 & -1 & -1 & +1 & +1 & -1 & -1 & +1 \\ +1 & +1 & +1 & +1 & -1 & -1 & -1 & -1 \\ +1 & -1 & +1 & -1 & -1 & +1 & -1 & +1 \\ +1 & +1 & -1 & -1 & -1 & -1 & +1 & +1 \\ +1 & -1 & -1 & +1 & -1 & +1 & +1 & -1 \end{vmatrix}$$

Restricted Use - À usage restreint

Replication methods for multistage sample

- *Fays method*

Pseudo-stratum	School		R1	R2	R3	R4	R5	R6	R7	R8
1	1		1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5
1	2		0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5
2	3		1.5	0.5	1.5	0.5	1.5	0.5	1.5	0.5
2	4		0.5	1.5	0.5	1.5	0.5	1.5	0.5	1.5
3	5		1.5	1.5	0.5	0.5	1.5	1.5	0.5	0.5
3	6		0.5	0.5	1.5	1.5	0.5	0.5	1.5	1.5
4	7		1.5	0.5	0.5	1.5	1.5	0.5	0.5	1.5
4	8		0.5	1.5	1.5	0.5	0.5	1.5	1.5	0.5
5	9		1.5	1.5	1.5	1.5	0.5	0.5	0.5	0.5
5	10		0.5	0.5	0.5	0.5	1.5	1.5	1.5	1.5

Restricted Use - À usage restreint

Replication methods for multistage sample

General formula

$$\sigma^2_{(\hat{\theta})} = c \sum_{i=1}^{G} (\hat{\theta}_{(i)} - \hat{\theta})^2$$

	c
BRR	$\frac{1}{G}$
Fay	$\frac{1}{G(1-k)^2}$
JK1	$\frac{G-1}{G}$
JK2	1

- BRR / Fay : each replicate is an estimate of the sampling variance
    - C = average
    - Same number of replicate for each country
- JK2 : each replicate corresponds of the pseudo-stratum to the sampling variance estimate
    - C = sum
    - Possibility of different number of replicates

Restricted Use - À usage restreint

Replication methods for multistage sample

- In the case of infinite populations, the sampling variance of the mean estimate consists of 2 components in the case of a PISA sampling design:
  - Between school variance
  - Within school variance

$$\sigma^2_{(\hat{\mu})} = \frac{\sigma^2_{sch}}{n_{sch}} + \frac{(\sigma^2_{cla|sch} + \sigma^2_{stu|cla})}{n_{sch} n_{stu/sch}}$$

- Replication methods, by removing entire schools, only integrate the uncertainty due to the selection of schools, not due to the selection of students within schools

$$\sigma^2_{(\hat{\theta})} = c \sum_{i=1}^{G} (\hat{\theta}_{(i)} - \hat{\theta})^2$$

Replication methods for multistage sample

- Let us image an educational system with no school variance and with infinite populations of students within schools

Effect	*Sum of Square*	*Degree of Freedom*	*Mean Square*
Schools	$SS_B = \sum_{j=1}^{K}\sum_{i=1}^{n_j}(\bar{y}_j - \bar{y})^2$	k-1	$MS_B = \frac{SS_B}{k-1}$
Residual	$SS_W = \sum_{j=1}^{K}\sum_{i=1}^{n_j}(x_i - \bar{y}_j)^2$	n-k	$MS_W = \frac{SS_W}{n-k} = \sigma_W^2$
Total	$SS_T = \sum_{i=1}^{n}(x_i - \bar{y})^2$	n-1	$MS_T = \frac{SS_T}{n-1} = \sigma^2$

$$\sigma_B^2 = \frac{MS_B - MS_w}{n_W}$$

$$\sigma^2_{(\hat{\mu})} = \frac{\sigma_B^2}{n_B} + \frac{\sigma_W^2}{n_B n_W} = \frac{\frac{MS_B - MS_W}{n_W}}{n_B} + \frac{MS_W}{n_B n_W} = \frac{MS_B - MS_W}{n_B n_W} + \frac{MS_W}{n_B n_W} = \frac{MS_B}{n_B}$$

Replication methods for multistage sample

- An illustration of this mathematical equality:
  - Estimation of the SE by JK1 replication method

School	Full	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10
1	100		100	100	100	100	100	100	100	100	100
2	110	110		110	110	110	110	110	110	110	110
3	120	120	120		120	120	120	120	120	120	120
4	130	130	130	130		130	130	130	130	130	130
5	140	140	140	140	140		140	140	140	140	140
6	150	150	150	150	150	150		150	150	150	150
7	160	160	160	160	160	160	160		160	160	160
8	170	170	170	170	170	170	170	170		170	170
9	180	180	180	180	180	180	180	180	180		180
10	190	190	190	190	190	190	190	190	190	190	
Mean	145.0	150.0	148.9	147.8	146.7	145.6	144.4	143.3	142.2	141.1	140.00
		25.00	15.12	7.72	2.78	0.31	0.31	2.78	7.72	15.12	25.00

$$\sigma^2_{(\hat{\mu})} = \frac{G-1}{G}\sum_{i=1}^{G}(\hat{\mu}_{(i)} - \hat{\mu})^2 = \frac{9}{10}\left(25 + 15.12 + \ldots + 25\right) = \frac{9}{10}101.85 = 91.67$$

$$SE = \sigma_{(\hat{\mu})} = \sqrt{\sigma^2_{(\hat{\mu})}} = \sqrt{91.67} = 9.57$$

Restricted Use - À usage restreint

Replication methods for multistage sample

- An illustration of this mathematical equality:
  - Estimation of the SE by the formula

School	*School mean*	$(\hat{\mu}_i - \hat{\mu})^2$
1	100	2025
2	110	1225
3	120	625
4	130	225
5	140	25
6	150	25
7	160	225
8	170	625
9	180	1225
10	190	2025
Mean	*145*	
	SS	*8250*
	MS	*916.666667*

$$\sigma^2_{(\hat{\mu})} = \frac{CM_B}{n_B} = \frac{916.7}{10} = 91.67$$

Restricted Use - À usage restreint