Dealing with Numbers

"Surviving the post p-values era"

Estibaliz Gómez de Mariscal

esgomezm@pa.uc3m.es

I. Errors & Uncertainty

II. Statistical measures to characterize uncertainty

III. Statistical sampling

IV. Sample size

Surviving the post p-values era:

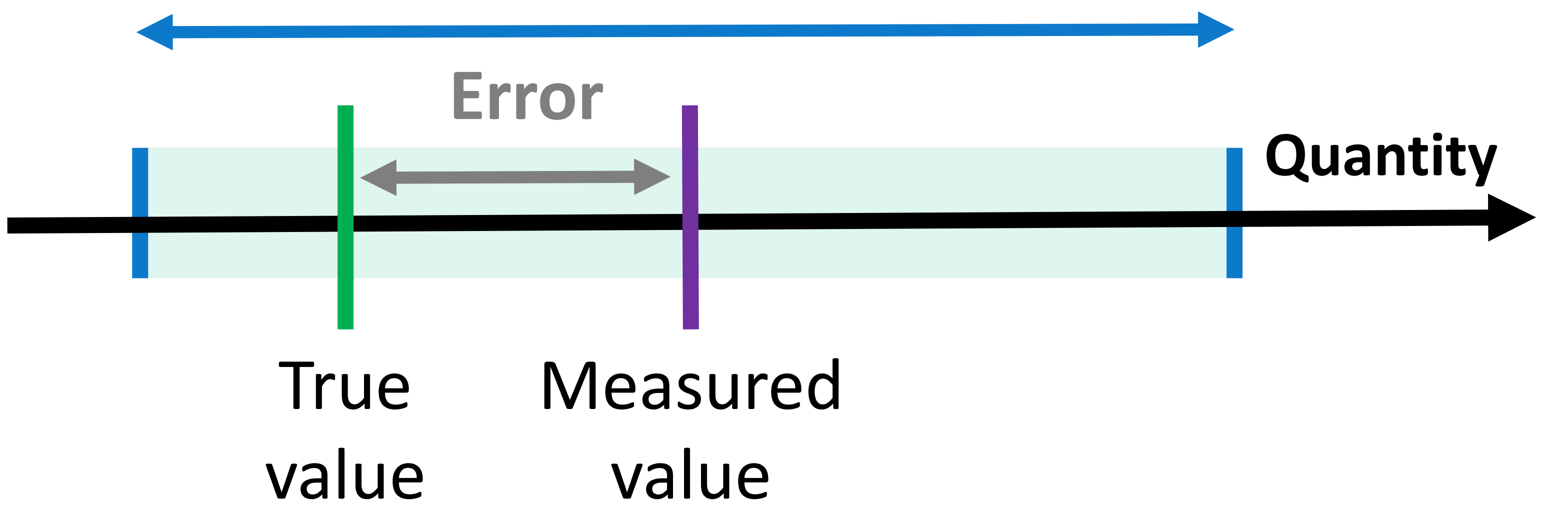
- **Use of the p-value as a size-dependent function: model and applications**

Measurement is a process of experimentally obtaining the value of a quantity.

Optimal measurement procedure to be as close as possible to the true value

Uncertainty range

Measured value $\pm$ U



https://sisu.ut.ee/measurement/uncertainty

ERROR

most times cannot be used in practice

UNCERTAINTY

Interval around the measured value. It provides the <u>quality of a measurement</u>

- Resolution
- Estimation method
- Calibration
- Repeatability uncertainty (human factor)
- Reproducibility
- Data acquisition (measurement error)
- Biological systems
- Data analysis

Final result report

Sources of variance / errors / uncertainty / bias / noise

- Resolution
- Calibration
- Data acquisition (measurement error)

Controlled measures
e_c

- Biological systems
- Reproducibility
- Repeatability uncertainty (human factor)

Uncontrolled & subjective
e_u

- Data analysis
- Estimation method

Calculated using statistical methods
e

Final result
$\pm e_c \pm e_u \pm e$

Statistical measures to characterize uncertainty

Standard deviation (σ)

- Variability or dispersion of the data w.r.t. the mean.
- $s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n}}$; ($s^2$: sample variance, biased)
- $S = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}}$; ($S^2$: quasi-variance non-biased)

Standard error of the mean (SEM)

- Standard deviation of the mean: how far the sample mean of the data is likely to be from the true population mean
- It can be used to validate the accuracy of a sample.
- $s.e.m_{\bar{x}} = \frac{S}{\sqrt{n}}$: standard error

Confidence Intervals (CI at 95%)

- "A range of values it can be 95% confident that it contains the true mean value"
- If the variance of the population (σ^2) is known:

$$CI = \bar{x} \pm z_{(1-\alpha/2)} \frac{\sigma}{\sqrt{n}}$$

- If the σ is unknown but n is large:

$$CI = \bar{x} \pm z_{(1-\alpha/2)} \frac{S}{\sqrt{n}}$$

- If the σ is unknown but n is small:

$$CI = \bar{x} \pm t_{(n-1, 1-\alpha/2)} \frac{S}{\sqrt{n}}$$

Statistical measures to characterize uncertainty

Standard deviation (σ)

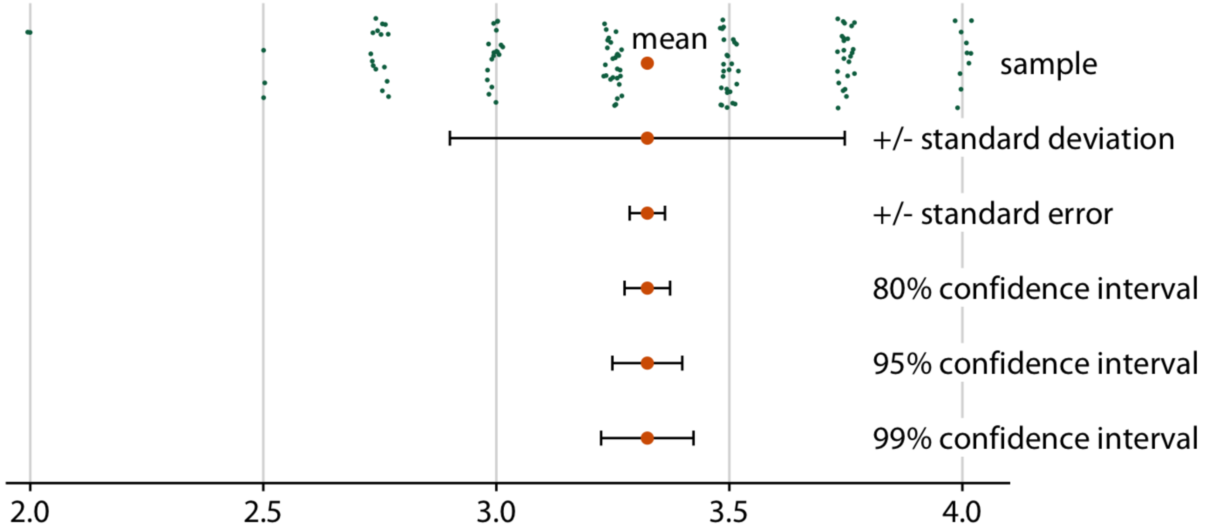
- Variability or dispersion of the data w.r.t. the mean.

Confidence Intervals (CI at 95%)

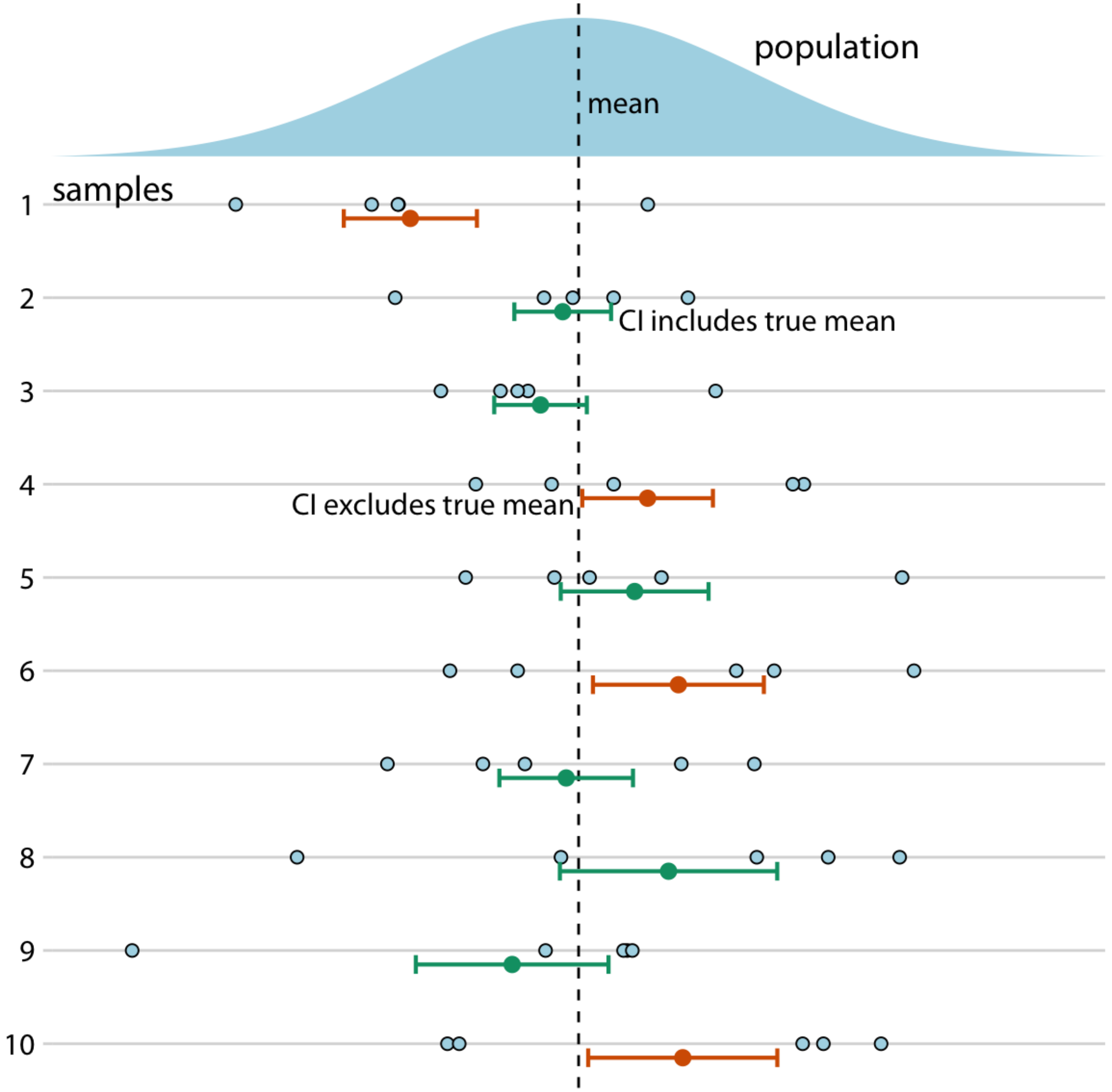
- A range of values it can be 95% confident that it contains the true mean value. 

Standard error of the mean (SEM)

- Standard deviation of the mean: how far the sample mean of the data is likely to be from the true population mean
- It can be used to validate the accuracy of a sample.



Previous measures come from experimental data → BIAS

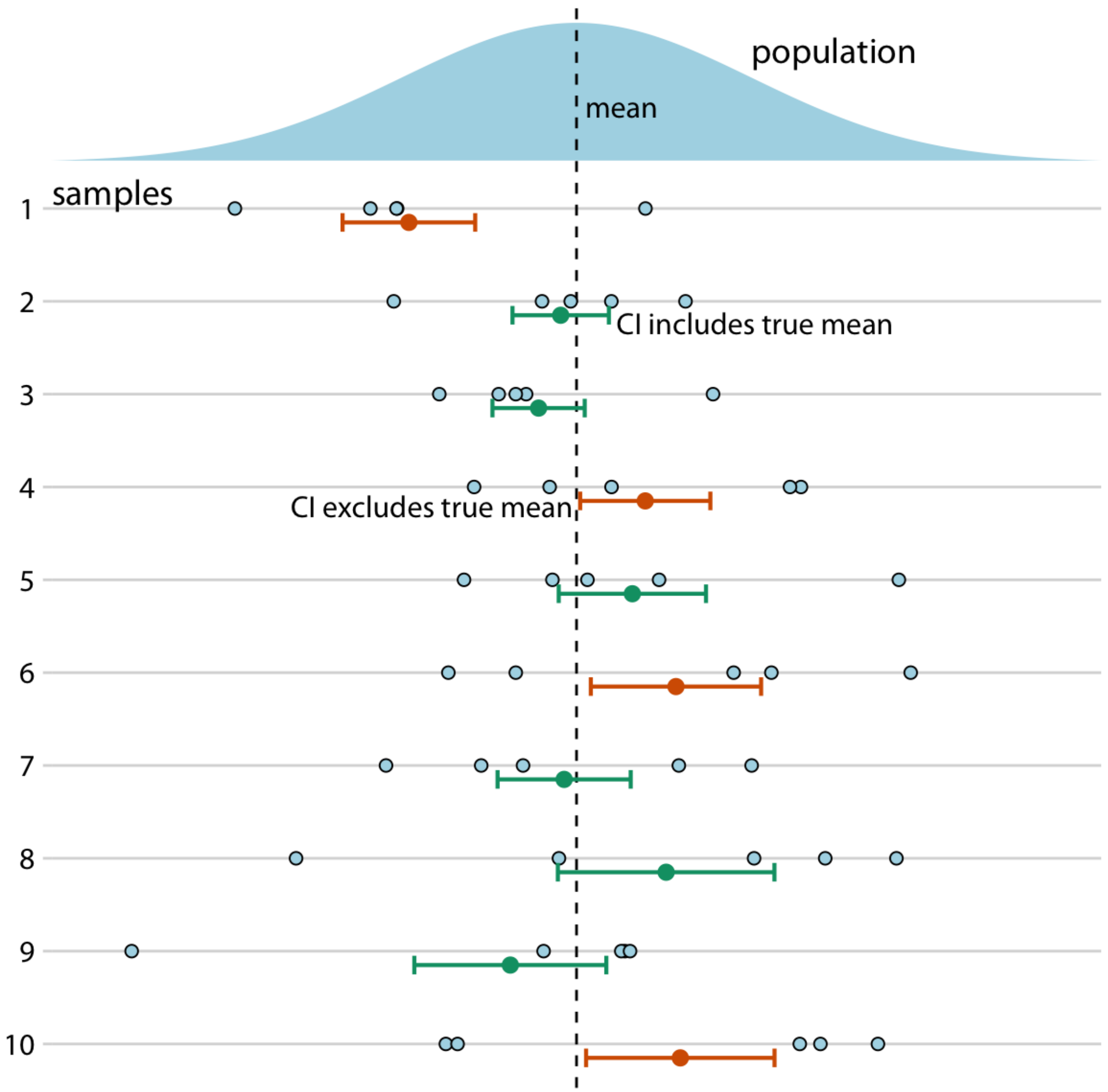


True meaning of a Confidence Interval

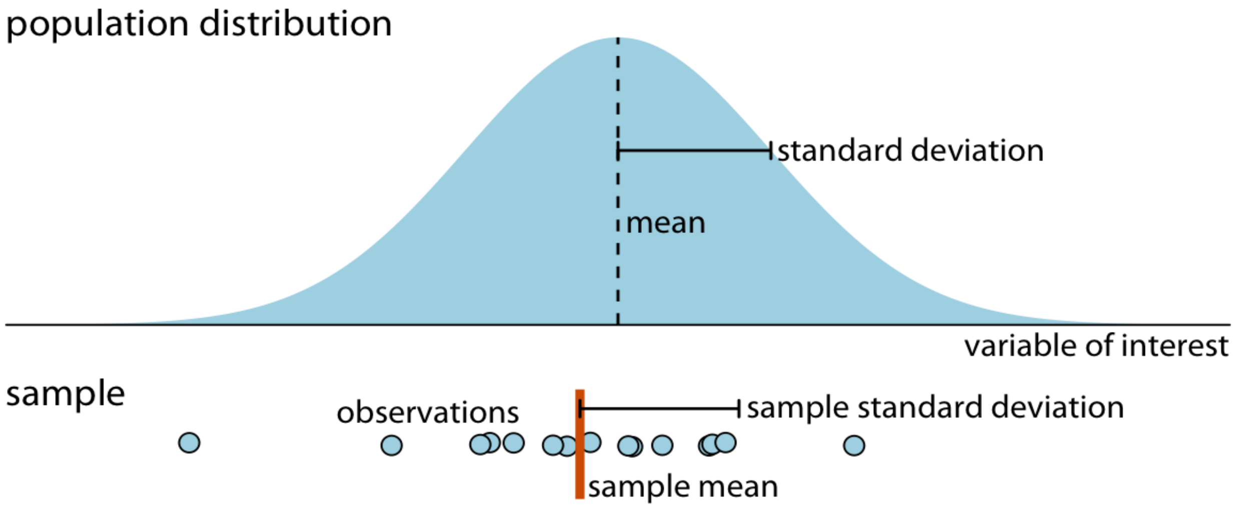
For each 100 intervals, we belief that at least in 95 of them we will find the expected value.

https://serialmentor.com/dataviz/visualizing-uncertainty.html

Previous measures come from experimental data → BIAS



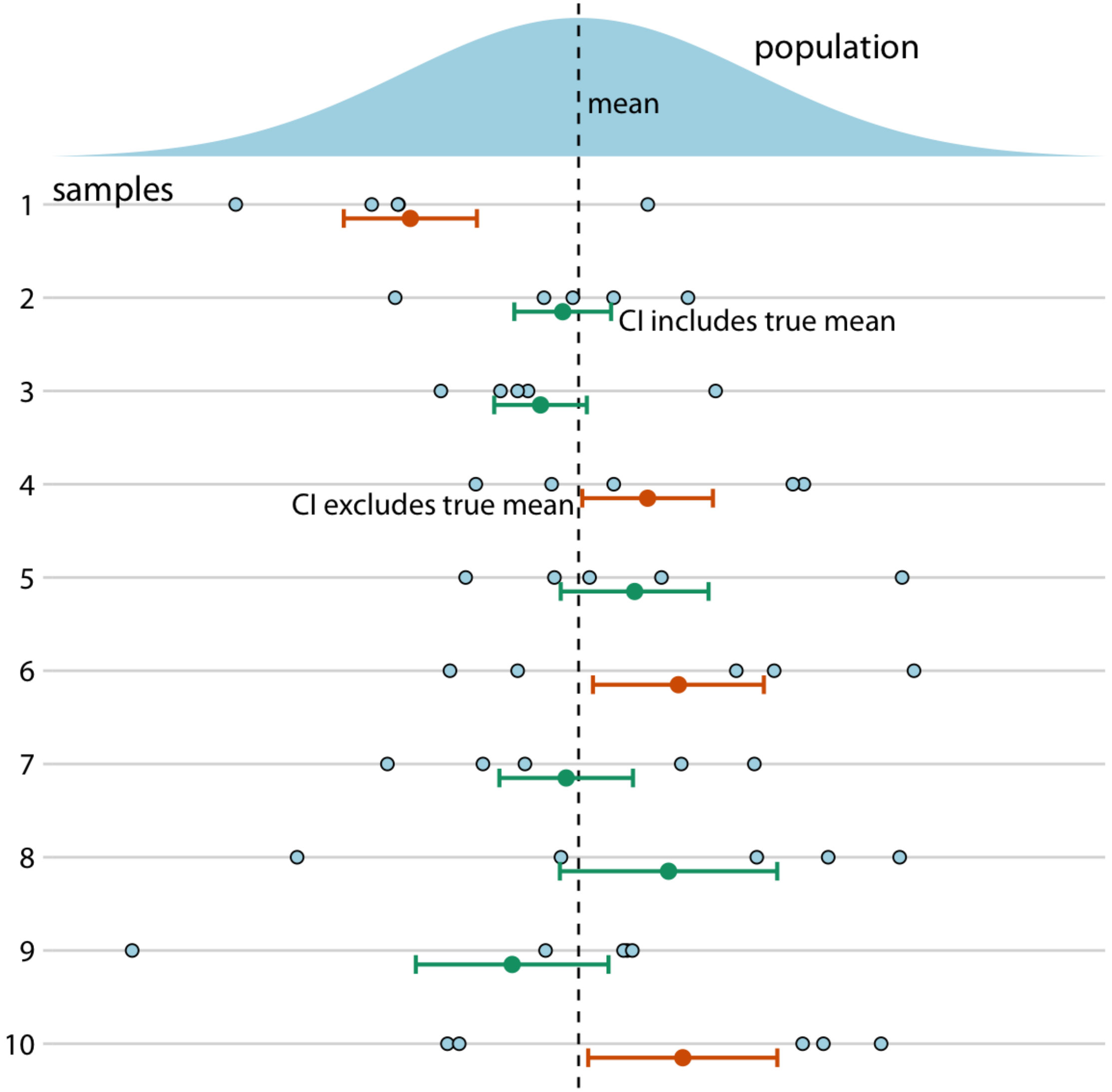
Statistical sampling



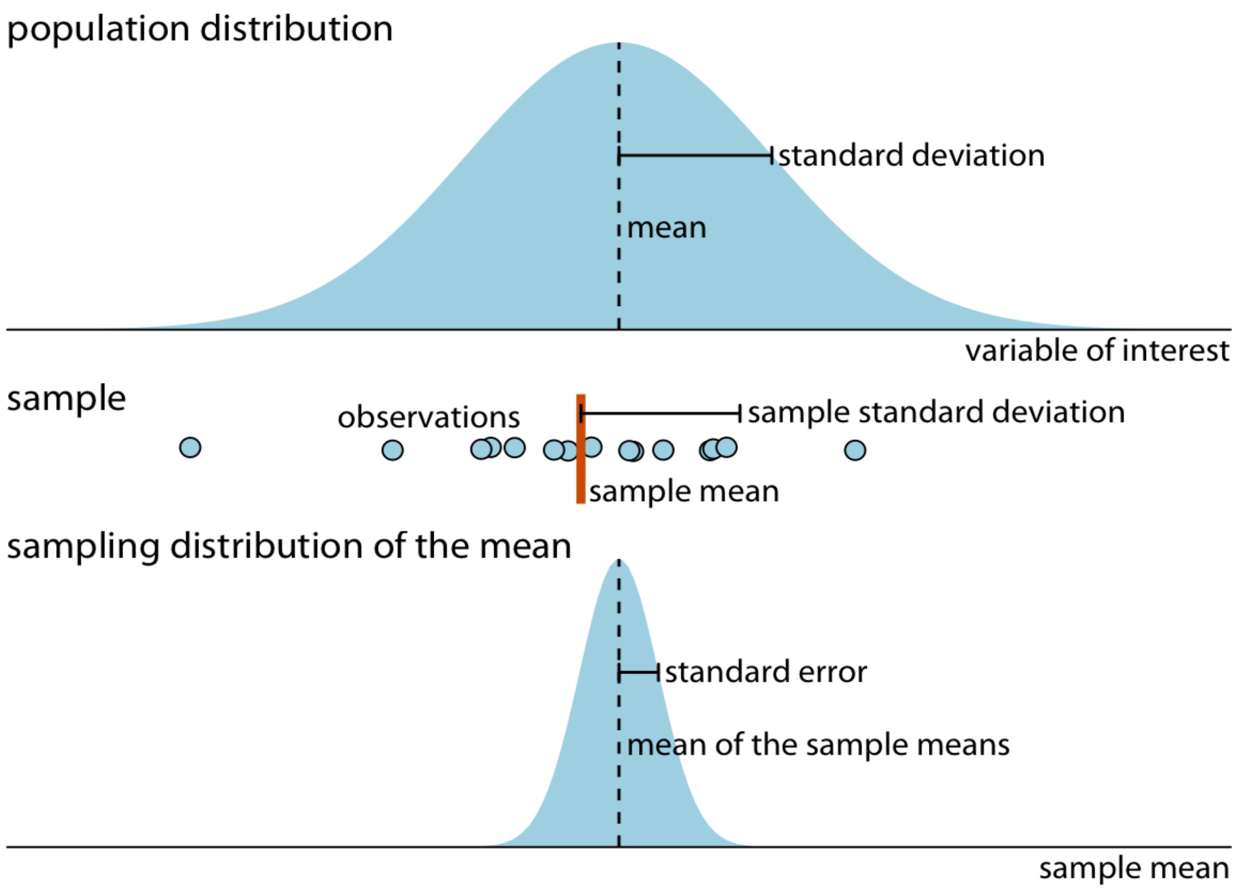
1. Any finite sample of that variable will have a different sample mean and standard deviation.
2. Repeat the sampling and calculate a mean each time.

https://serialmentor.com/dataviz/visualizing-uncertainty.html

Previous measures come from experimental data → BIAS



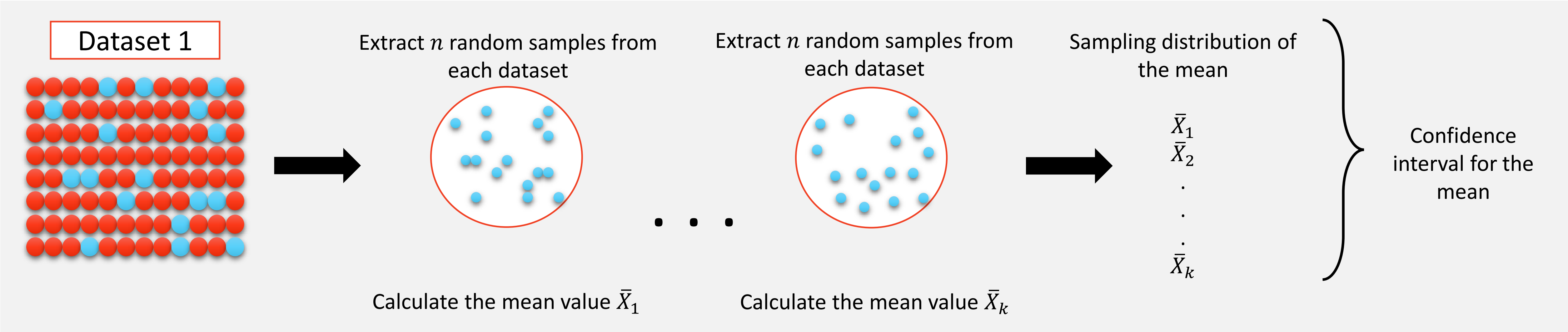
Statistical sampling



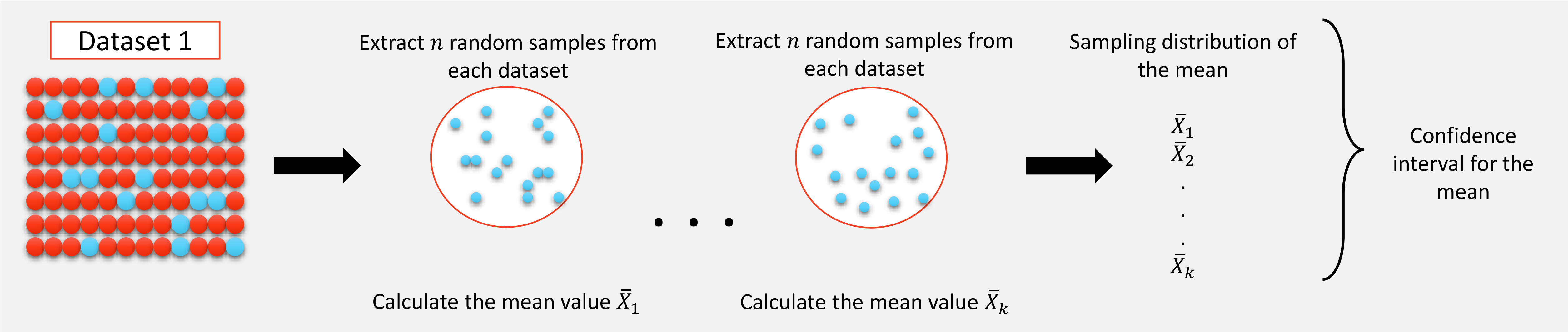
The resulting means would be distributed according to the sampling distribution of the mean.

https://serialmentor.com/dataviz/visualizing-uncertainty.html

Computation of (confidence) intervals of the sample mean



Computation of (confidence) intervals of the sample mean



The following conditions must be satisfied:

I. All the samples must be **randomly chosen**

II. Confidence intervals are designed for **normal distributions**.

- ➢ The mean follows a normal conditions (central limit theorem).

III. Sampling has to satisfy the **independence** condition.

- Sampling with replacement.
- Case without replacement: Questionnaires to people who is leaving a store, you cannot ask them to go back again.
  - ➢ Recommendation: n value smaller than 10% of the sample → Calculate first n.

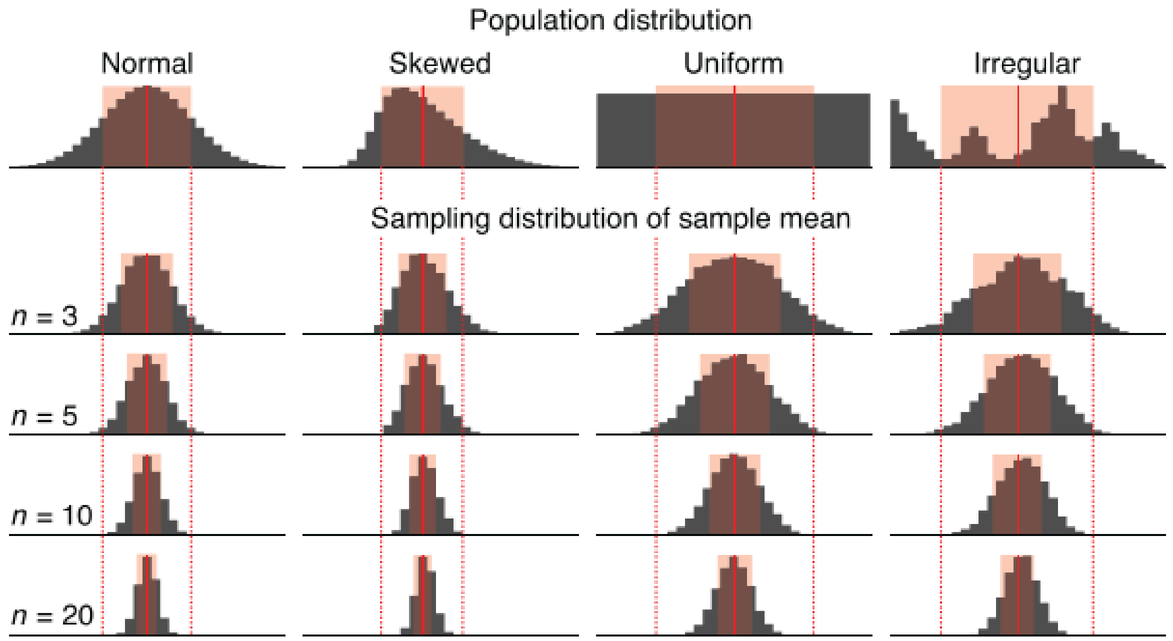
How robust are these measures?

What is the main problem with estimation using experimental data?

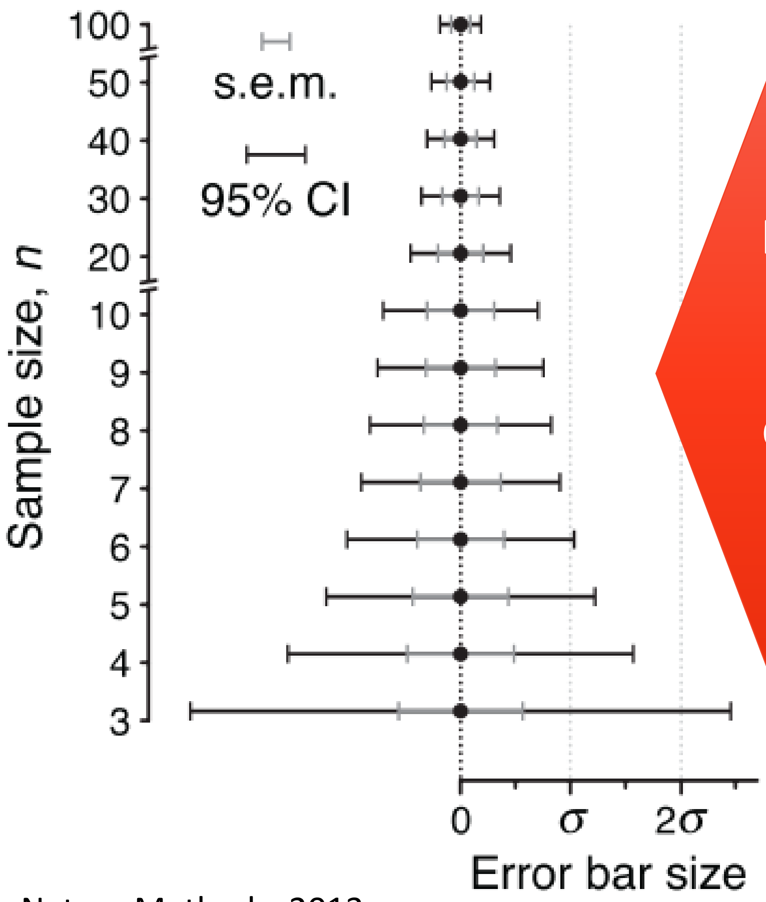
Previous measures come from experimental data → BIAS

- Pooled mean (mean of the sample means): $\mu_X = \frac{\sum_i N_{X_i} \mu_{X_i}}{\sum_i N_{X_i}}$
- Pooled standard deviation: $\sigma_X = \frac{\sum_i (N_{X_i} - 1) \sigma_{X_i}^2}{\sum_i (N_{X_i} - 1)}$

Sampling distributions of sample means for 10,000 samples for indicated sample sizes



Relationship between s.e.m. and CI at 95% error bars when increasing the sample size (n)



Intervals of the error bars decrease when *n* increases

Martin Krzywinski, Naomi Altman, Points of Significance: Importance of being uncertain, Nature Methods, 2013
Martin Krzywinski, Naomi Altman, Points of Significance: Error bars, Nature Methods, 2013

Sample size (n)

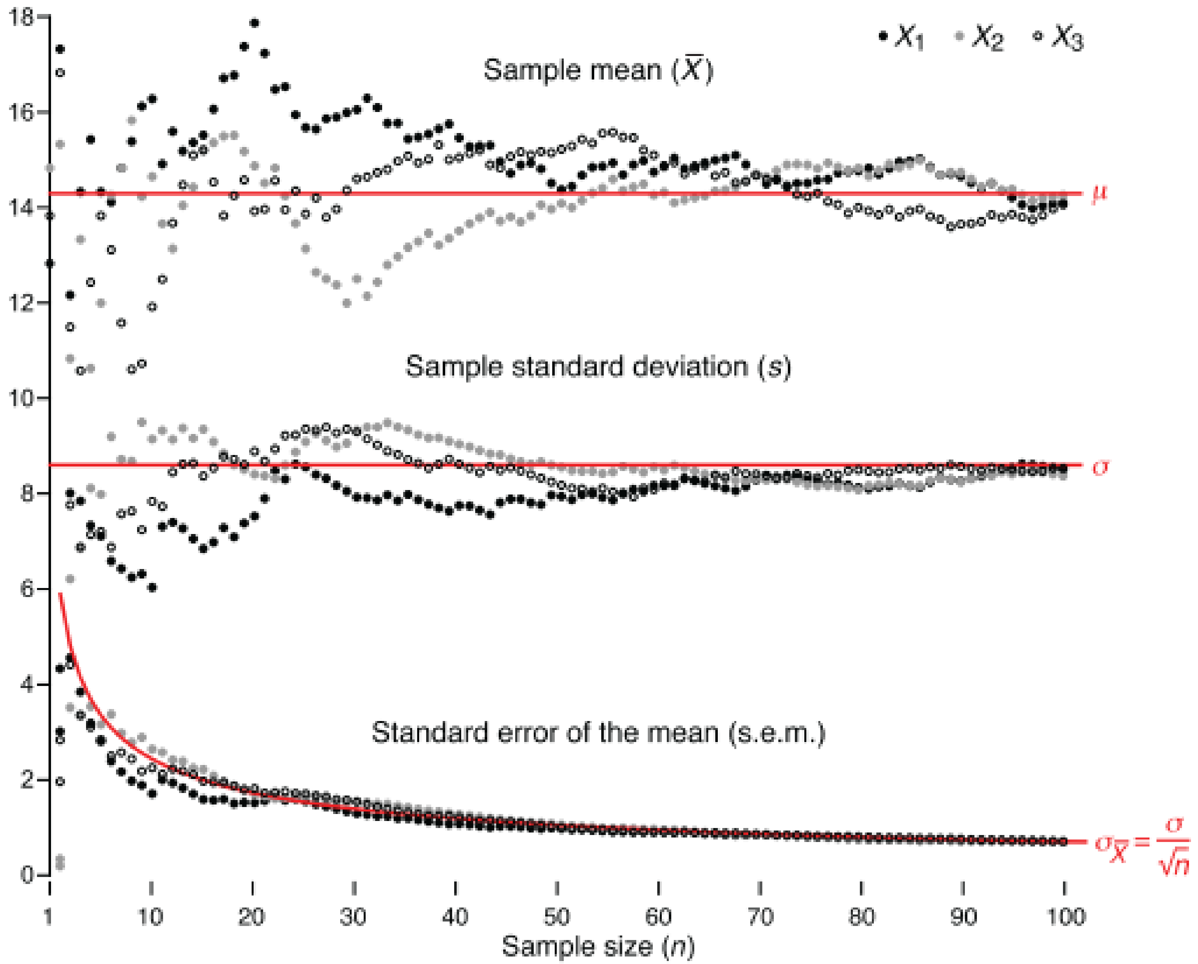


Figure 4 | The mean ($\bar{X}$), s.d. (s) and s.e.m. of three samples of increasing size drawn from the distribution in **Figure 2a**. As n is increased, $\bar{X}$ and s more closely approximate μ and σ. The s.e.m. ($s/\sqrt{n}$) is an estimate of $\sigma_{\bar{X}}$ and measures how well the sample mean approximates the population mean.

There is a certain value of n for which the population mean is well approximated.

Martin Krzywinski, Naomi Altman, Points of Significance: Importance of being uncertain, Nature Methods, 2013

Sample size (n)

To calculate n it is necessary to set parameters that depend on the type of study:

- Standard deviation of the sample (σ)
- Confidence level (1-α)
- Error range (e) / Length of the confidence interval (L)

If we are interested in the confidence interval of the sample mean (μ):

- If the standard deviation of the population (σ) is known: $$n = \frac{(2z_{(1-\alpha/2)}\sigma)^2}{L^2}.$$
- If σ is unknown but n is large: $$n = \frac{(2z_{(1-\alpha/2)}S)^2}{L^2}.$$
- If σ is unknown but n is small: $$n = \frac{(2t_{(n-1,1-\alpha/2)}S)^2}{L^2},$$

where $z_{(1-\alpha/2)}$ and $t_{(n-1,1-\alpha/2)}$ are the corresponding quantiles in the Normal and t-Student distributions respectively.

Martin Krzywinski, Naomi Altman, Points of Significance: Importance of being uncertain, Nature Methods, 2013

Sample size (n)

		Fórmula	R
One parameter Estimation	Mean	$n=\left(\frac{2\cdot z_{\alpha/2}\cdot\sigma}{A}\right)^2$	*sample.size.mean*
	Probabilities	$n=\left(\frac{z_{\alpha/2}}{A}\right)^2$	*sample.size.prop*
Means comparison	Independent	$n=\frac{2\cdot\sigma^2\cdot(z_{\alpha/2}+Z_\beta)^2}{\Delta^2}$	*TwoSampleMean.Equality*
	Paired	$N=\frac{2\cdot\sigma_I^2\cdot(z_{\alpha/2}+z_\beta)^2}{\Delta^2}$	-
	Análisis del cambio	$n=\frac{2\cdot\sigma_C^2\cdot(z_{\alpha/2}+z_\beta)^2}{\Delta^2}$	-
	Equivalent	$n=\frac{2\sigma^2(z_\alpha+z_{\beta/2})^2}{\Delta^2}$	*TwoSampleMean.Equivalence*
	Non-inferiority	$n=\frac{2\sigma^2(z_\alpha+z_\beta)^2}{\Delta^2}$	*TwoSampleMean.NIS*
	Precision	$n=\frac{8\cdot\sigma^2\cdot z_{\alpha/2}^2}{A^2}$	-
Probabilities comparison	Independent	$n=\left(\frac{z_{\alpha/2}\sqrt{2p(1-p)}}{p_A-p_B}+\frac{z_\beta\sqrt{p_A(1-p_A)+p_B(1-p_B)}}{p_A-p_B}\right)^2$	*TwoSampleProportion.Equality*
Times comparison	Instant recruitment	$\mathrm{N}=\frac{2E}{2-\pi_A-\pi_B}$	*ssizeCT.default (para HRR constantes)*
	Recruitment for a period	$N=\frac{2(z_{1-\alpha/2}+z_{1-\beta})^2(\Phi(\lambda_A)+\Phi(\lambda_B))}{(\lambda_A-\lambda_B)^2}$	*TwoSampleSurvival.Equality*

Nomenclature

$\boldsymbol{\sigma}^2$: total variance

$\boldsymbol{\sigma}_I^2$: intra-subject variance

$\boldsymbol{\sigma}_C^2$: variance of the changing variable

Δ: difference we want to detect

$\boldsymbol{\rho}$: correlation between both observations

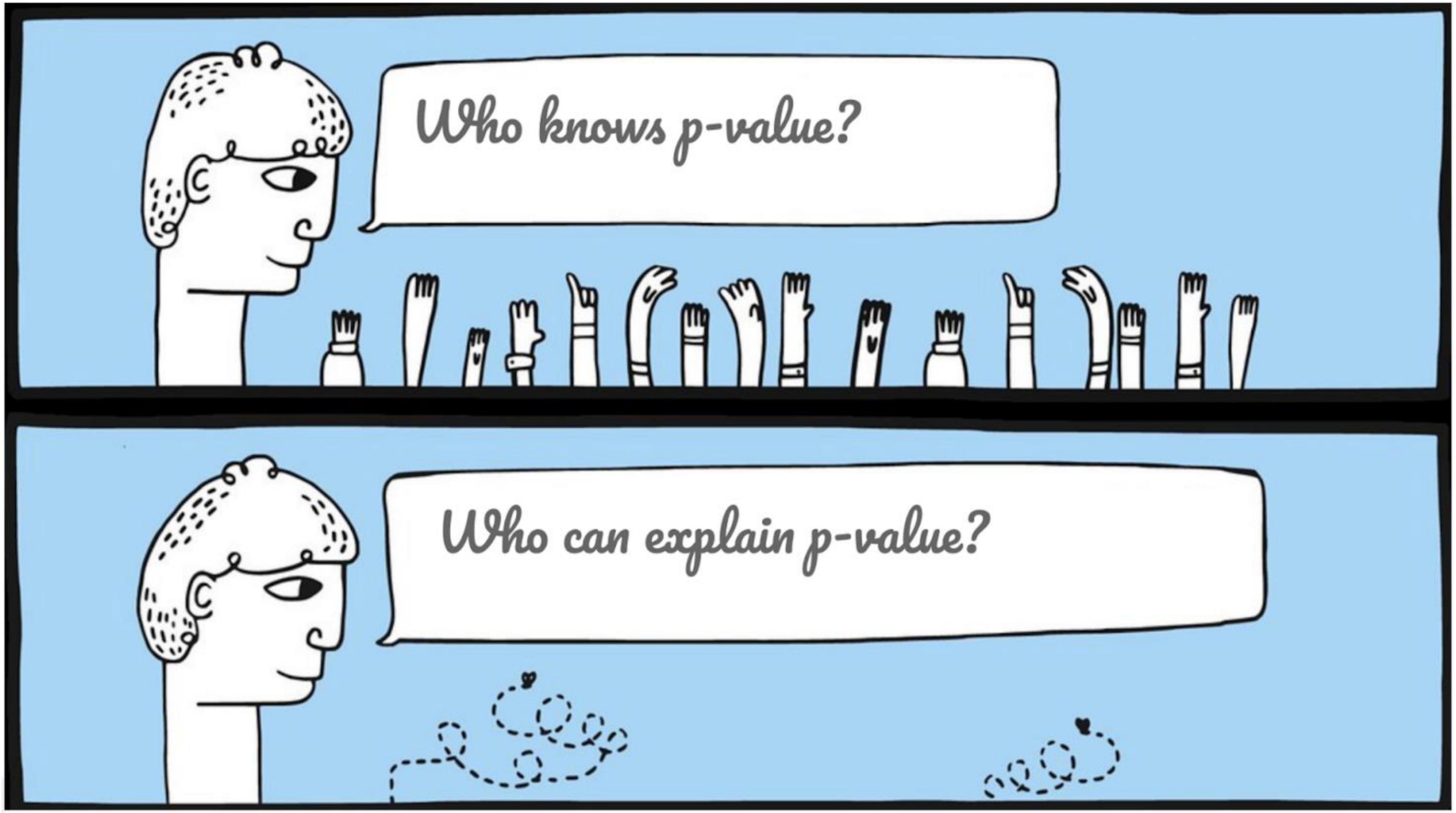
$\boldsymbol{A}$: amplitude of the interval

$\boldsymbol{E}$: number of events we need to observe

$\boldsymbol{\lambda}$: event appearance rate

$\boldsymbol{\pi}$: estimated proportion of cases for which the event will NOT be present during the study.

Surviving the post *p-values* era



Acknowledgments: This work was produced with the support of the Spanish Ministry of Economy and Competitiveness (TEC2015-73064-EXP, TEC2016–78052-RTC2017-6600-1, RTC-2017-6600-1), a 2017 Leonardo Grant for Researchers and Cultural Creators, BBVA Foundation, and NEUBIAS (Network of European Bioimage Analysists).

uc3m | Universidad Carlos III de Madrid

Instituto de Investigación Sanitaria Gregorio Marañón

neubias network of european bioimage analysts

cost EUROPEAN COOPERATION IN SCIENCE & TECHNOLOGY

GOBIERNO DE ESPAÑA

MINISTERIO DE ECONOMÍA, INDUSTRIA Y COMPETITIVIDAD

Fundación BBVA