Grandària Mostral

Quan és prou gran?

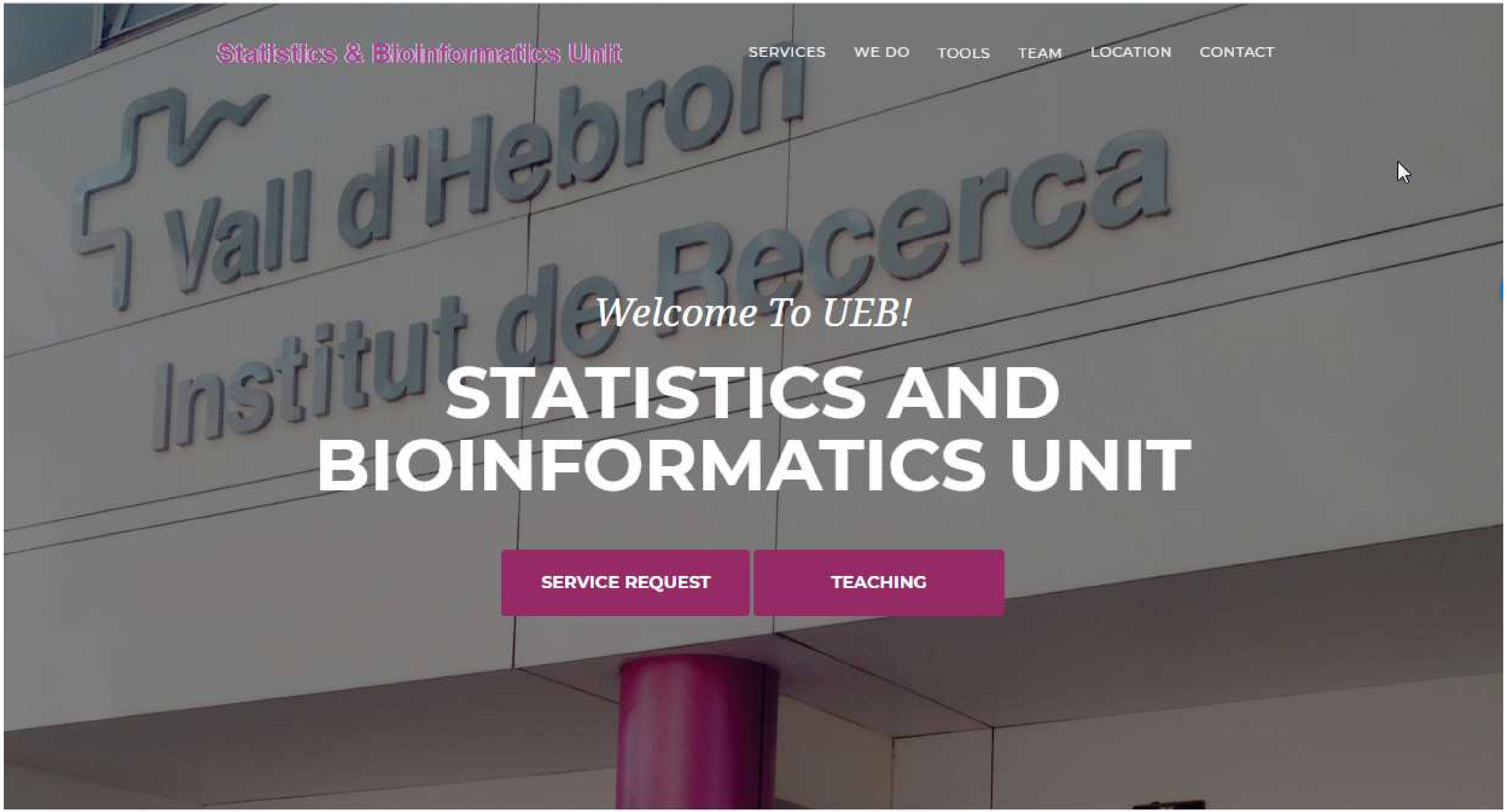
PINDOLES ESTADÍSTIQUES UEB-VHIR

Unitat d'Estadística I Bioinformàtica

28 de Gener de 2022

Vall d'Hebron | Vall d'Hebron Research | UNITAT D'ESTADÍSTICA I BIOINFORMÀTICA

ueb.vhir.org



Vall d'Hebron | Vall d'Hebron Research | UNITAT D'ESTADÍSTICA I BIOINFORMÀTICA

Pye et al. BMC Res Notes (2016) 9:90
DOI 10.1186/s13104-016-1893-x

BMC Research Notes

TECHNICAL NOTE Open Access

When is enough, enough? Understanding and solving your sample size problems in health services research

Victoria Pye*, Natalie Taylor, Robyn Clay-Williams and Jeffrey Braithwaite

The usual situation when a researcher ask a statistician about sample size for a research study

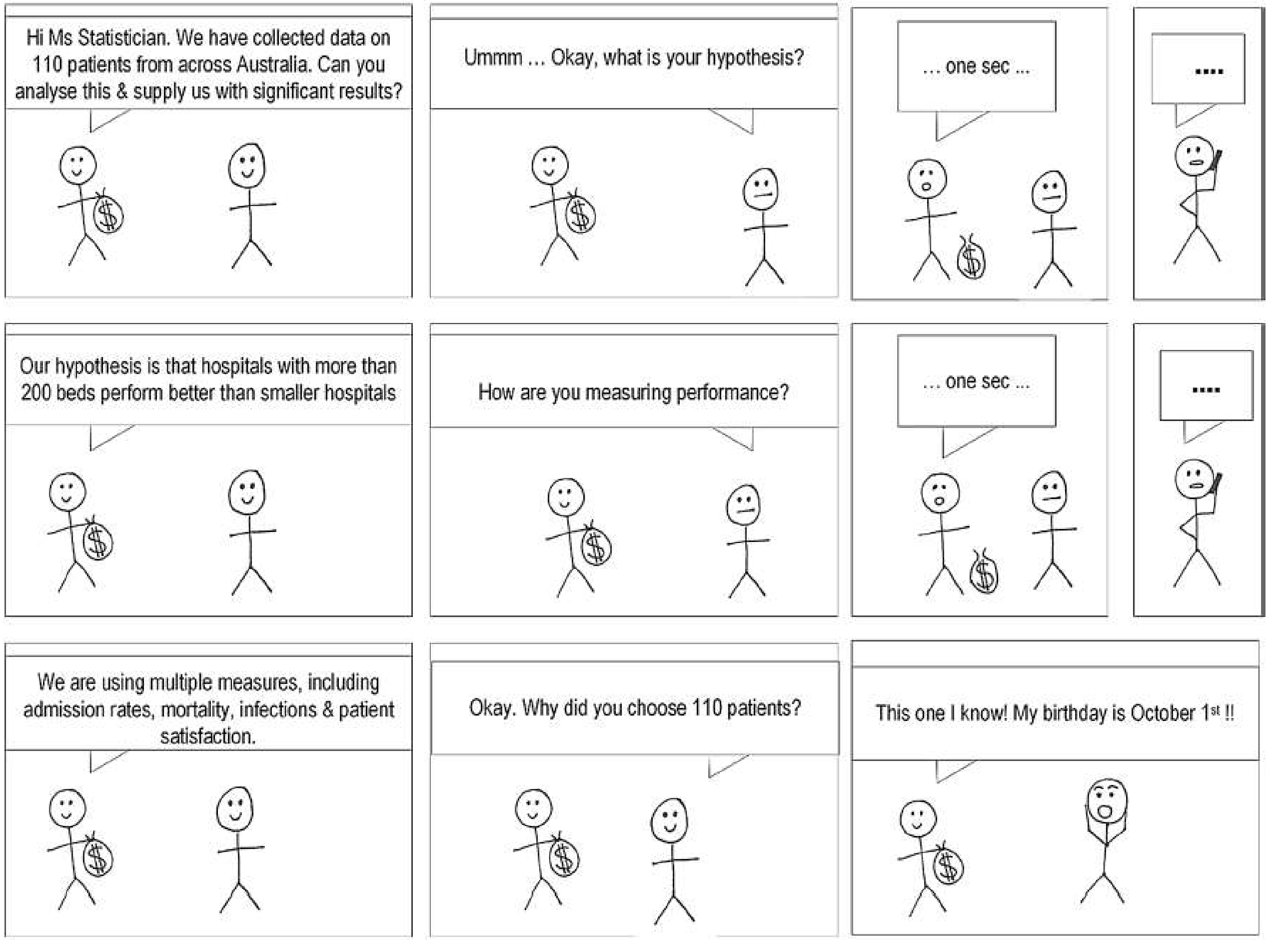
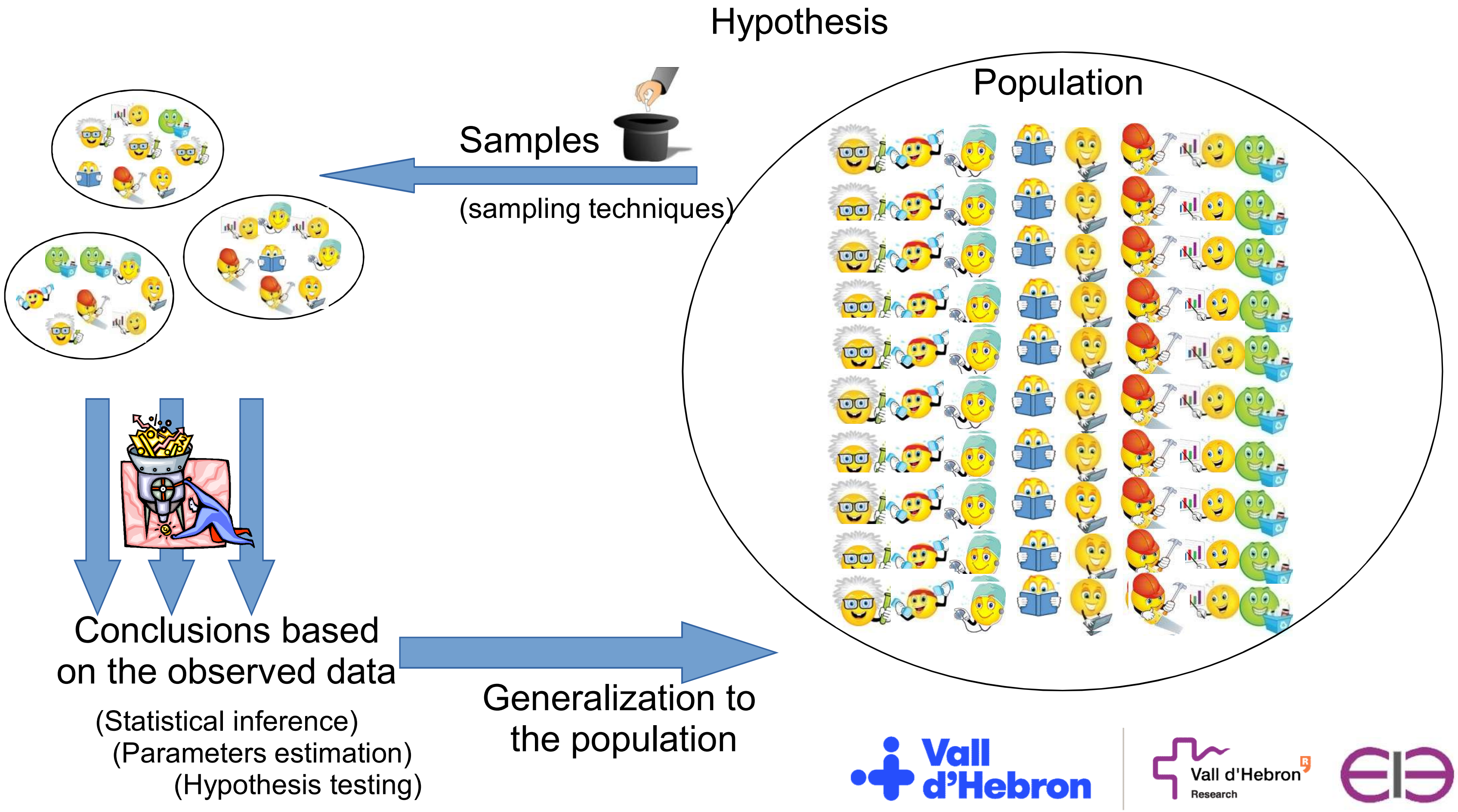


Fig. 1 A statistician's dilemma

Vall d'Hebron
Vall d'Hebron Research
UNITAT D'ESTADÍSTICA I BIOINFORMÀTICA

The objective of statistical inference



Vall d'Hebron | Vall d'Hebron Research | UNITAT D'ESTADÍSTICA I BIOINFORMÀTICA

Billion dollar Question

How many subjects do we need for our Study

Before calculating sample size

- Have a Clear Research Question
- Selected the design of the study (Clinical trial, Observational,..)
- Identify the outcomes or end-points
- Know how variables are going to be measured(proportions, ordinal, continous)
- How many subjects you can afford
- Have a clear idea of the expected results

Too few subjects

- You will not able to answer the question
- Potentially Not ethical

Too many subjects

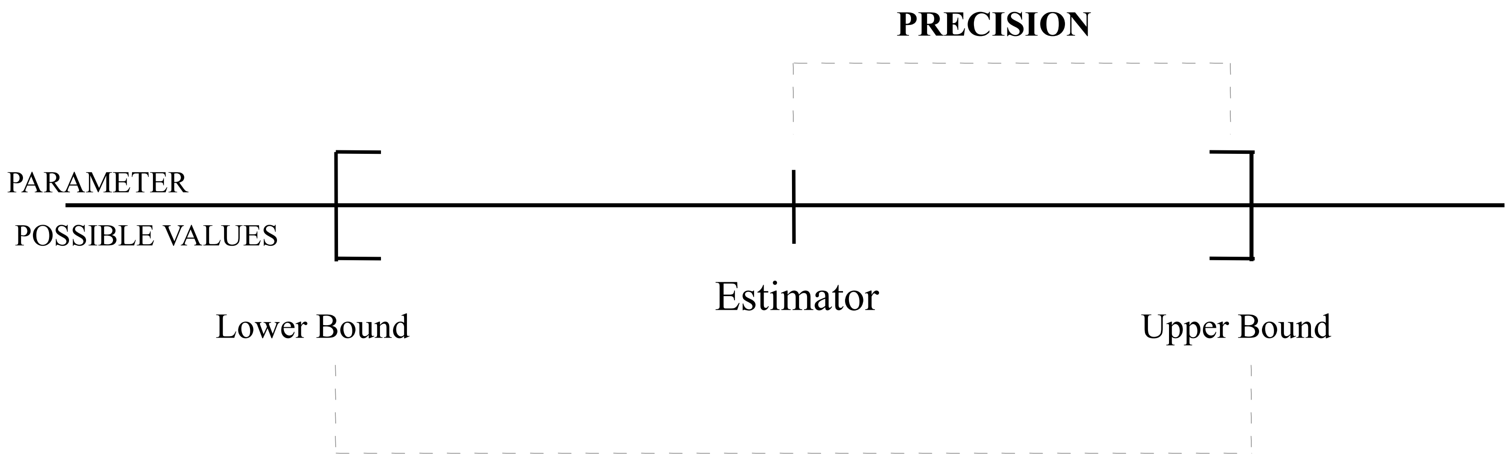
- Waste of resources
- Give a harmful treatment to patients
- Identify irrelevant treatment effects "significative"
- Potentially Not ethical

Vall d'Hebron | Vall d'Hebron Research | UNITAT D'ESTADÍSTICA I BIOINFORMÀTICA

Before discussing sample size calculations there are several things to keep in mind

- Type of calculations depend on study goal.
  - Estimation
  - Testing
- Preliminary concepts to be used
  - Standard error of an estimator
  - Confidence interval
  - Hypothesis testing
  - Type I error
  - Power of a test (1-β)

Sample size based on Confidence Intervals



Values in which we are confident that real population parameter is inside
With a prefixed confidence level (Usually 95%)

$$\text{Estimator} \pm \text{Coef.}_{1-\alpha/2} \times \text{Standard Error}$$

Vall d'Hebron | Vall d'Hebron Research | UNITAT D'ESTADÍSTICA I BIOINFORMÀTICA

Standard error of the mean

A measure of how variable is the sample mean when computed in all samples of size n

Standard deviation of the distribution of sample mean

$$\text{standard error} = \frac{\sigma}{\sqrt{\text{n}}} \cong \frac{s}{\sqrt{\text{n}}}$$

Standard error of a proportion

The standard error of a proportion is computed similarly to the SEM.

Instead of the standard deviation it uses the population proportion in the formula

If p is not known the estimated proportion is used

$$\text{standard error} = \sqrt{\frac{p \cdot (1-p)}{\text{n}}} \cong \sqrt{\frac{\hat{p} \cdot (1-\hat{p})}{\text{n}}}$$

Formulas for confidence intervals

Data normally distributed

Population variance known
(unrealistic assumption)

$$\bar{X}_n - z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}} \le \mu \le \bar{X}_n + z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}$$

Population variance unkown, estimated by sample variance

$$\bar{X}_n - t_{1-\alpha/2}\frac{s}{\sqrt{n}} \le \mu \le \bar{X}_n + t_{1-\alpha/2}\frac{s}{\sqrt{n}}$$

Data: Counts of presence or absence of an event

Sample must be "big enough"

$$\hat{p} \pm z_{1-\alpha/2}\sqrt{\frac{\hat{p}\hat{q}}{n}}; \quad n \ge 30, n\hat{p} \ge 5, n\hat{q} \ge 5$$

$Z_{1-\alpha/2}$ are quantiles of standard Normal N(0,1) distribution

$1-\alpha$	0,90	0,95	0,99
$Z_{1-\alpha/2}$	1,64	1,96	2,58

UNITAT D'ESTADÍSTICA I BIOINFORMÀTICA

Sample SIZE for mean

$$Precision = z_{1-\alpha/2} * ee = z_{1-\alpha/2} * \frac{\sigma}{\sqrt{n}}$$

$$n = \frac{z^2_{1-\alpha/2}\sigma^2}{precision^2}$$

- If interval range is 10 (precision =10/2=5)
- Confidence level is 95%
- Standard deviation is 20

$$n = \frac{1.96^2 20^2}{5^2} = 62$$


Calculadora de Grandària Mostral GRANMO
Versió 7.12 Abril 2012
Català
Castellano
English
Mitjanes : Estimació Poblacional
Nivell de confiança:
0.95
0.90
Altre
Població de referència (Intro => S'assumeix una població infinita):
99999999
Estimació de la desviació estàndard:
20
Precisió de l'estimació pel nivell de confiança sel·leccionat:
5
Proporció estimada de reposicions necessaries:
0
calcula
Neteja resultats
Neteja tot
Selecciona tot
Imprimir
20/01/2022 14:19:13 Estimació Poblacional (Mitjanes)
Una mostra aleatòria de 62 individus és suficient per estimar, amb una confiança del 95% i una precisió de +/- 5 unitats, la mitjana poblacional d'uns valors que es preveu que tinguin una desviació estàndard al voltant de 20 unitats. En percentatge de reposicions necessària s'ha previst que serà del 0%.
Proporcions
Mitjanes
Dos mitjanes independents
Mitjanes aparellades (repetides en un grup)
Observada respecte d'una de referència
Mitjanes aparellades (repetides en dos grups)
Estimació Poblacional
Anàlisi de la variança
Potència d'un contrast
Altres

Sample size for proportion

$$Precision = z_{1-\alpha/2} * ee = 1.96 * \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

$$n = \frac{1.96^2\, \hat{p}(1-\hat{p})}{precision^2} =$$

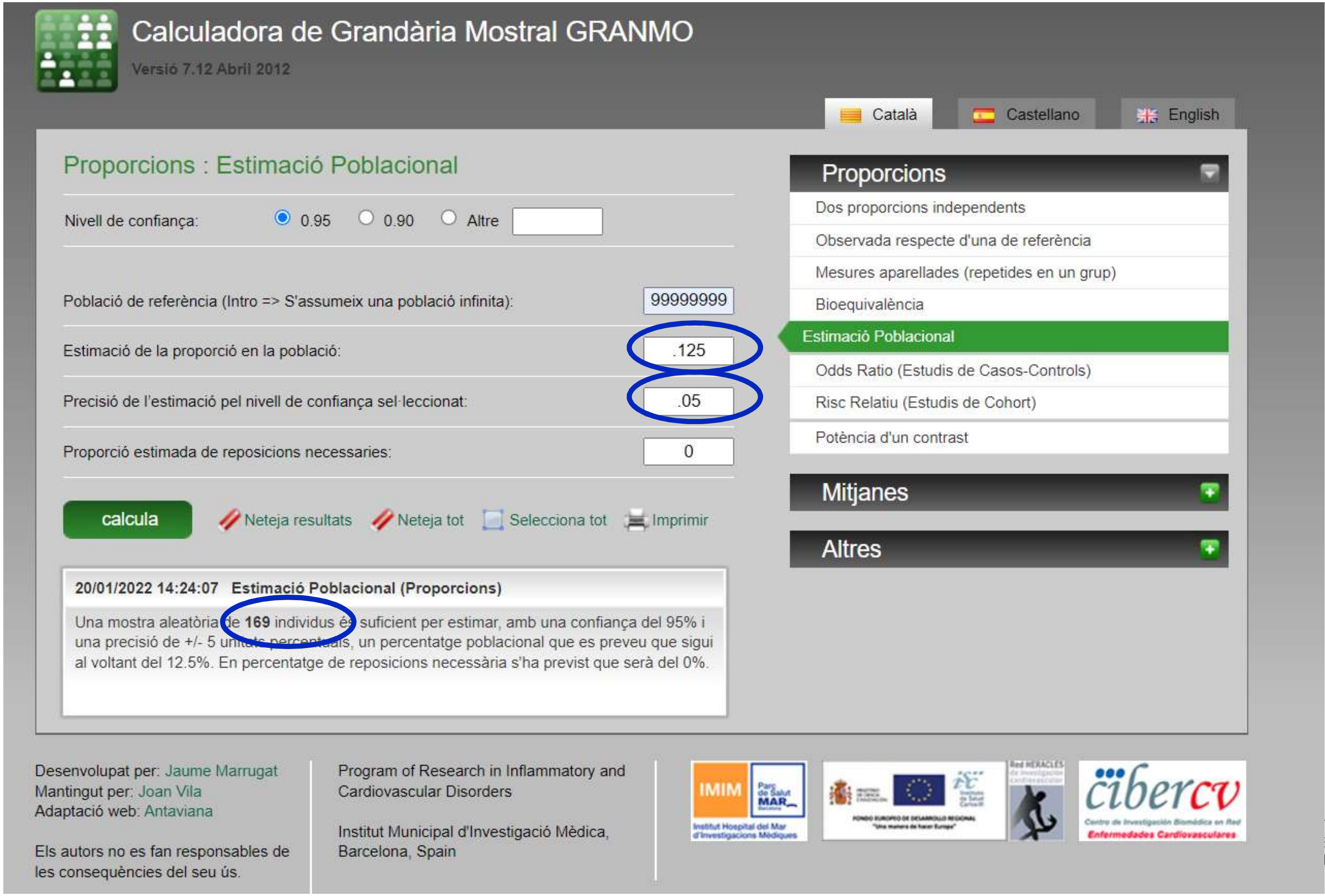
- Assume precision is 5% (Interval = p±.05)
- Confidence level is 95%
- If it is known that p is around 12.5%

$$n = \frac{1.96^2 .125(1-.125)}{.05^2} = 168$$

- If p is unknown maximum sample size will be if p=50%

$$n = \frac{1.96^2 .50(1-.50)}{.05^2} = 384$$

Vall d'Hebron | Vall d'Hebron Research | UNITAT D'ESTADÍSTICA I BIOINFORMÀTICA



TAT
STADÍSTICA I
INFORMÀTICA

Sample size based on Hypothesis testing

Type I Error α

Probability of rejecting the Null Hypothesis when its true (5%)

Power $1-\beta$ (Type II Error)

Probability of rejecting the Null Hypothesis when its false (80%,90%)

Variability of the data σ^2

Variance of the data

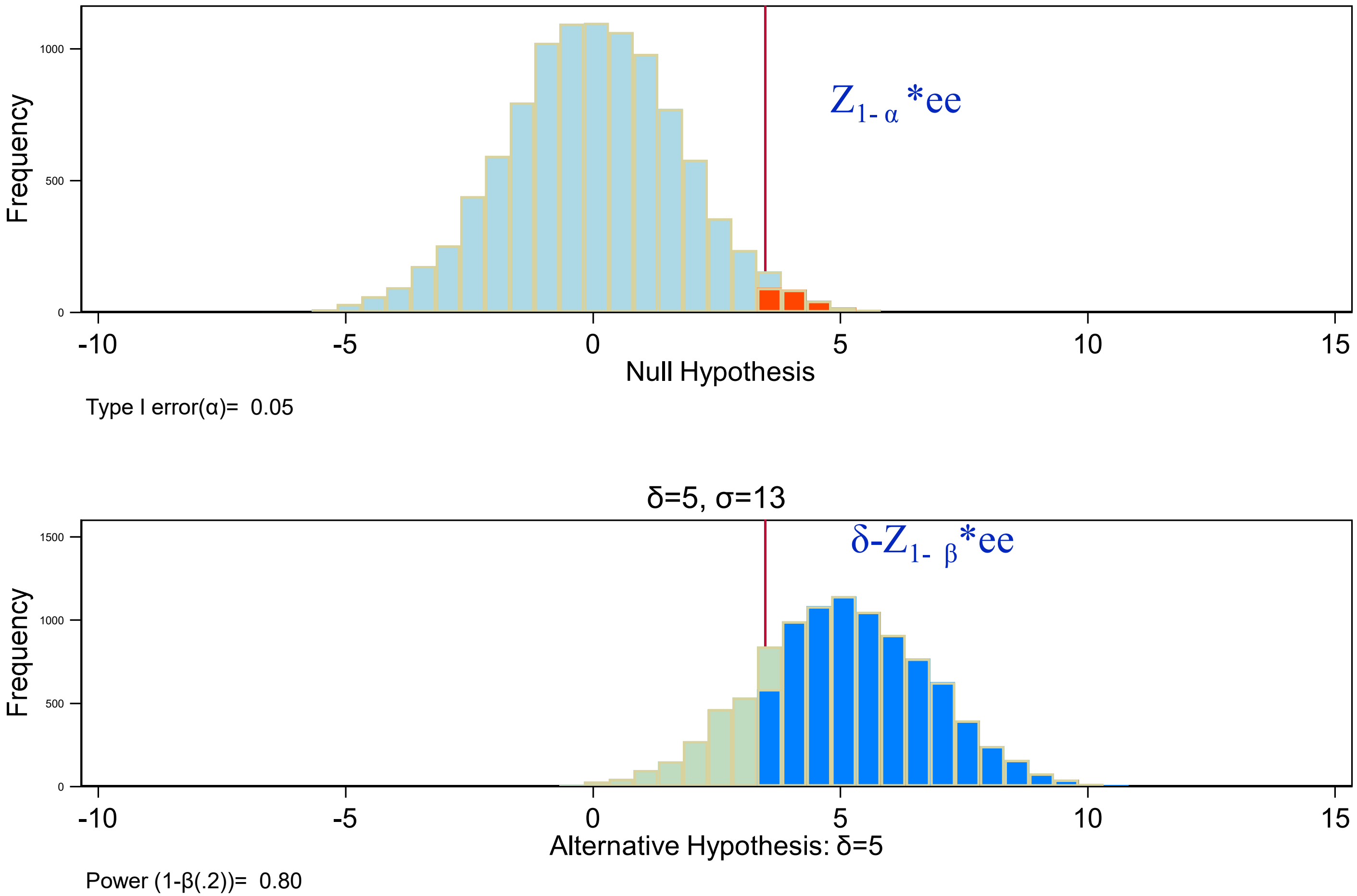
Effect Size δ

Minimum detectable difference betweeen the two groups to compare

Vall d'Hebron | Vall d'Hebron Research | UNITAT D'ESTADÍSTICA I BIOINFORMÀTICA

Sample size based on Hypothesis testing

N=107 por grupo

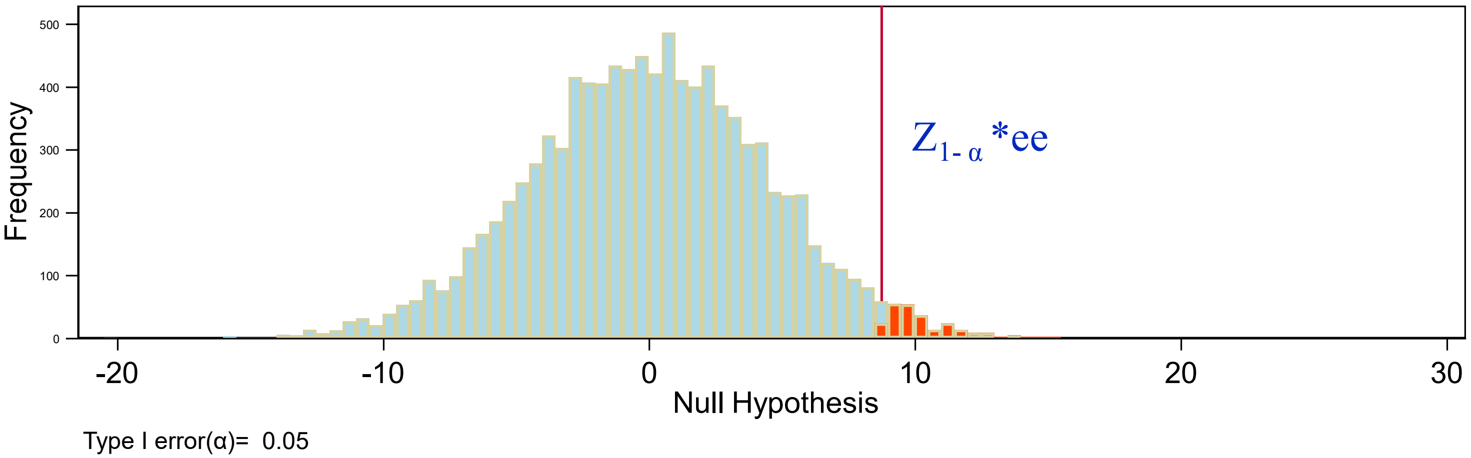


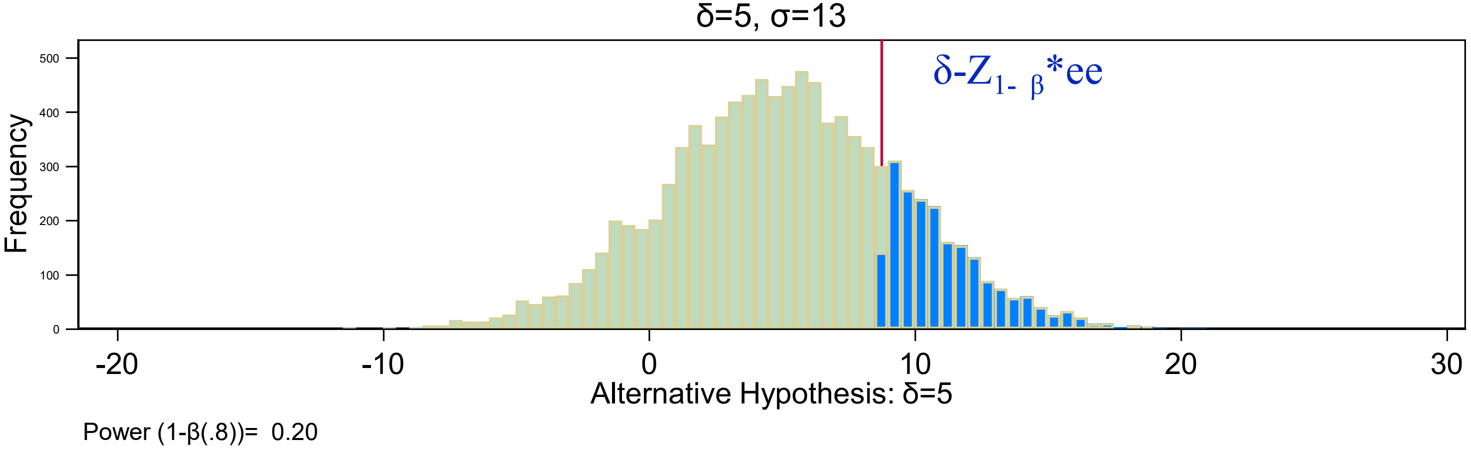
bron

UNITAT
D'ESTADÍSTICA I
BIOINFORMÀTICA

Sample size based on Hypothesis testing

N=17 por grupo

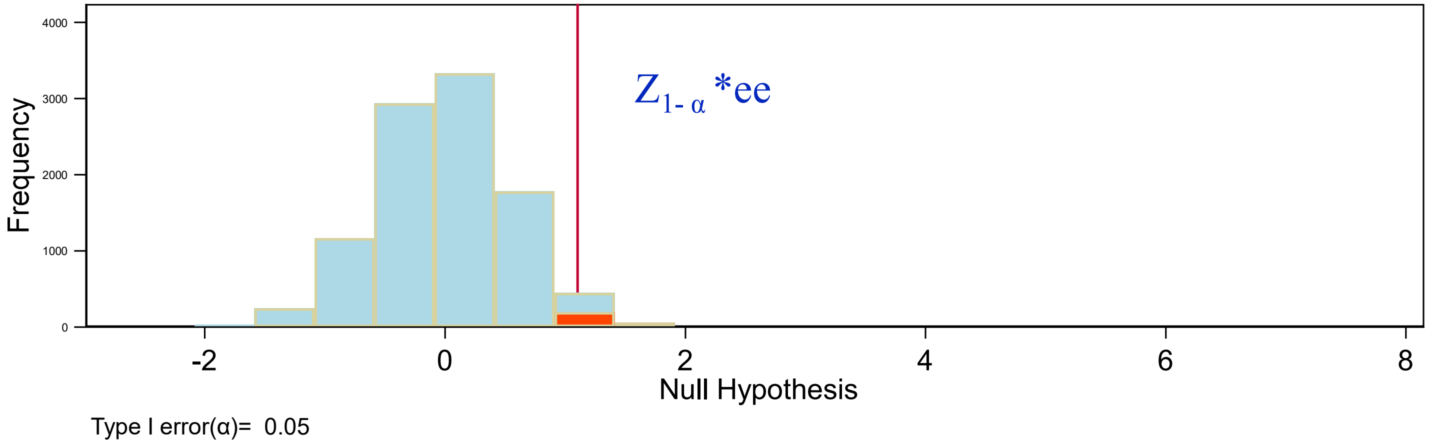




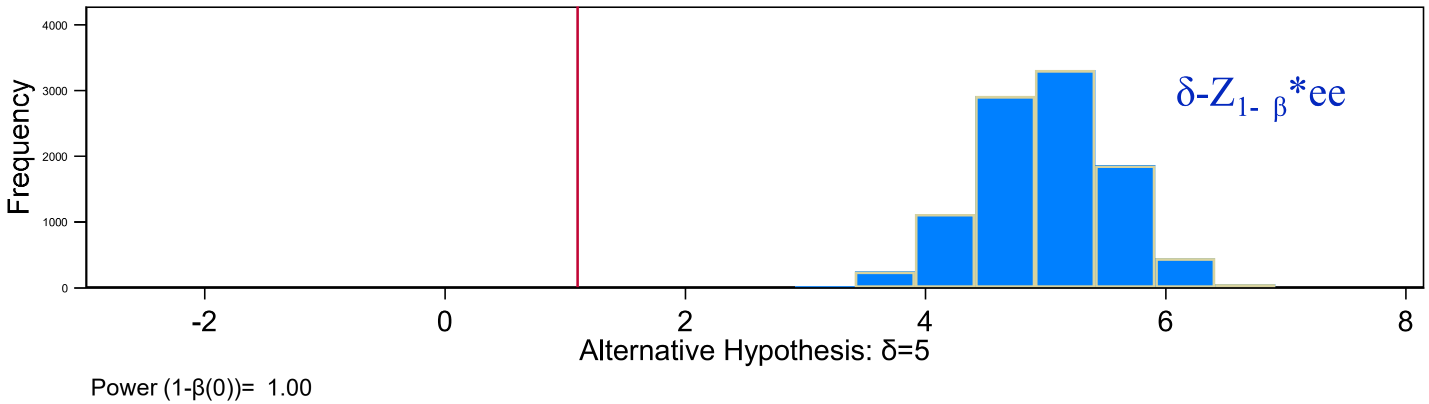
UNITAT D'ESTADÍSTICA I BIOINFORMÀTICA

Sample size based on Hypothesis testing

N=1070 por grupo



δ=5, σ=13



UNITAT D'ESTADÍSTICA I BIOINFORMÀTICA

Sample size based on Hypothesis testing

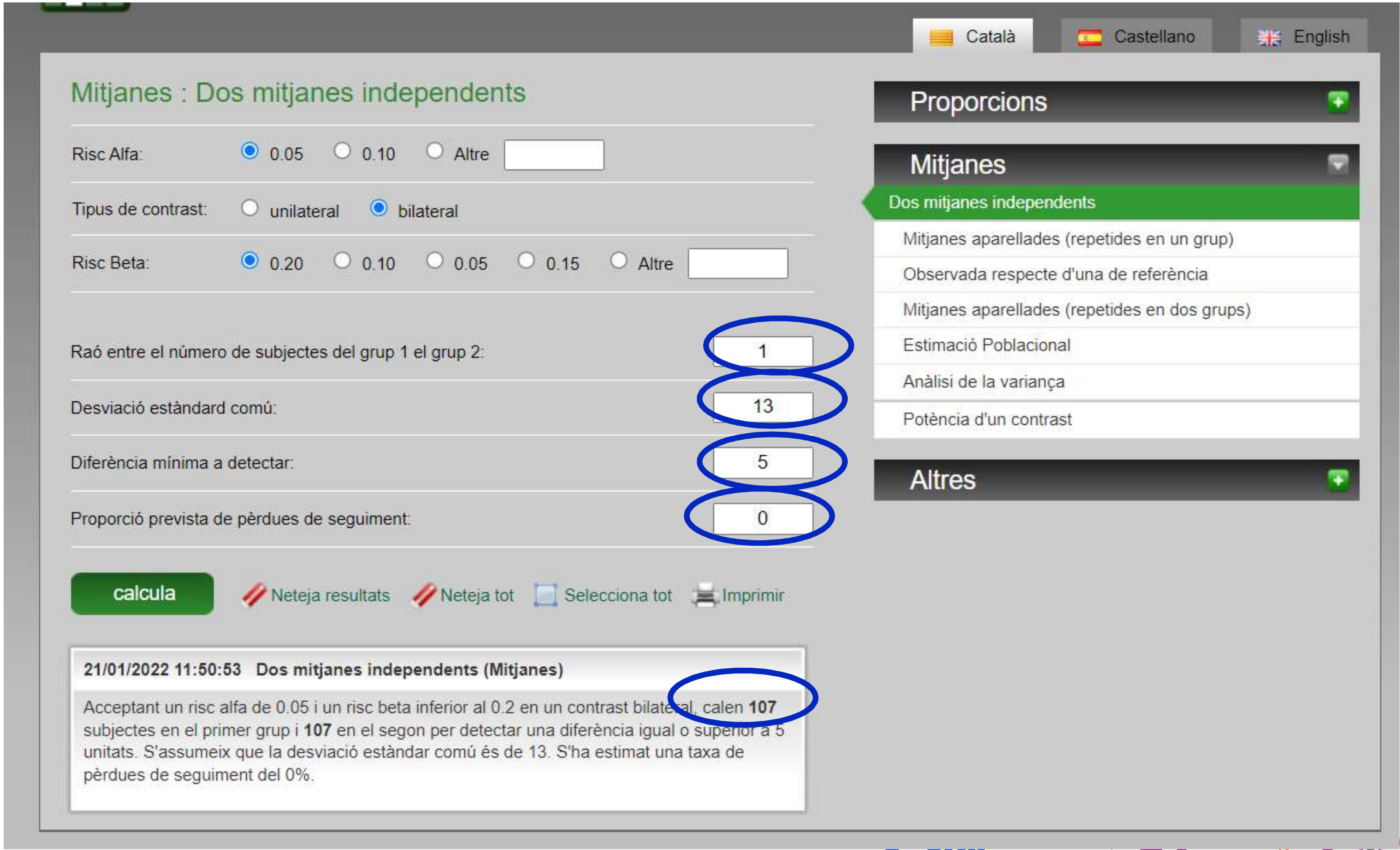
We want

$$Z_{1-\alpha} * ee = \delta - Z_{1-\beta} * ee$$

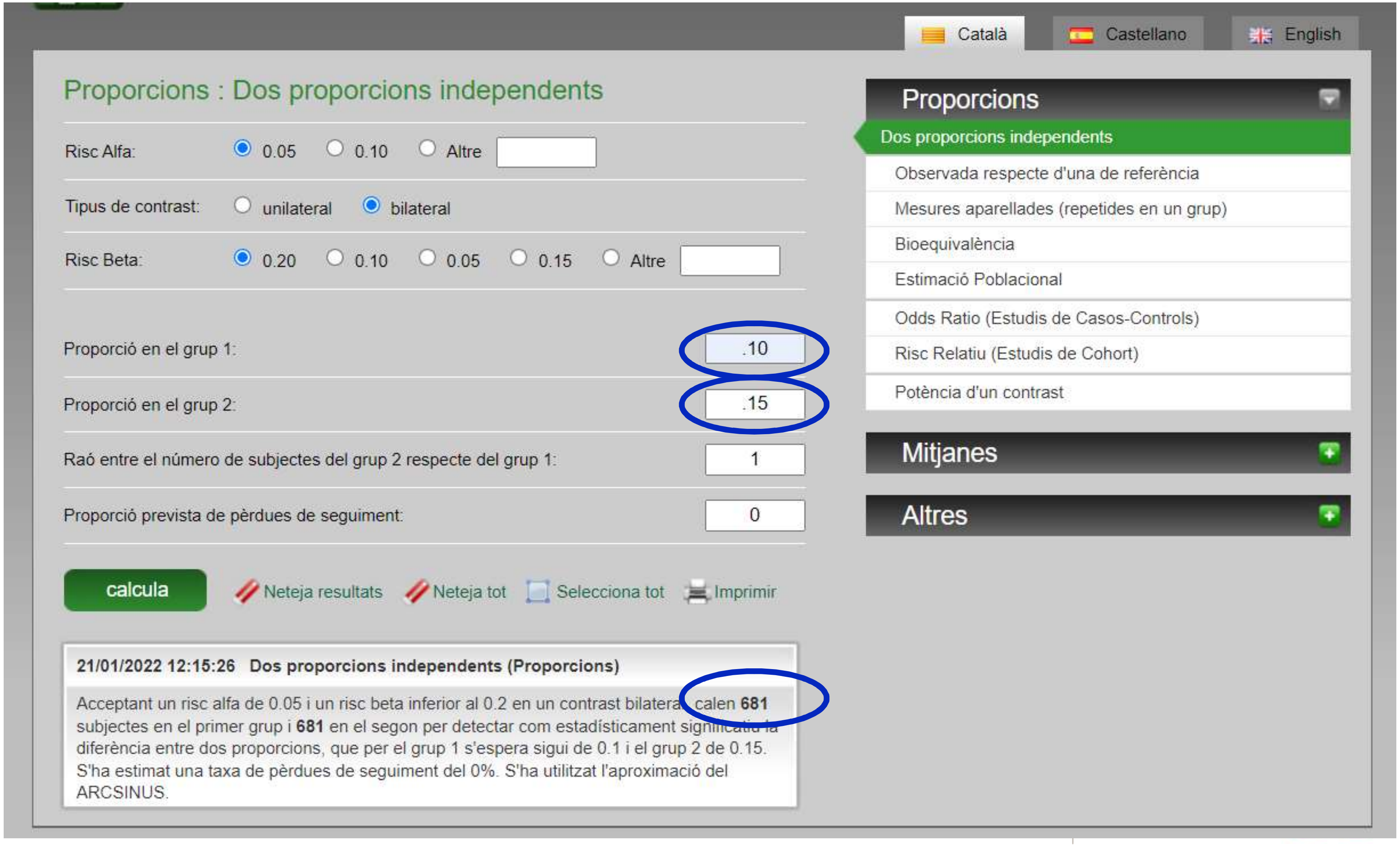
We know α, β and $ee = \sigma/\sqrt{n}$

$$\boxed{n = \frac{2\sigma^2 (z_{1-\alpha} + z_{1-\beta})^2}{\delta^2}}$$

Sample size for mean differences



Sample size for differences in proportions



UNITAT
D'ESTADÍSTICA I
BIOINFORMÀTICA

Some formulas to calculate sample size

Problema	Datos necesarios	Tamaño muestral
Estimación de una proporción	p=Proporción esperada o 0.50 d= precisión(mitad amplitud I.C.) e= error porcentual sobre verdadero parámetro $z_{1-\alpha/2}$= percentil N(0,1) para significación α	$n = \frac{z^2_{1-\alpha/2}\, p(1-p)}{d^2}$ $n = \frac{z^2_{1-\alpha/2}(1-p)}{e^2}$
Estimación de una media	σ^2=varianza esperada d= precisión(mitad amplitud I.C.) $z_{1-\alpha/2}$= percentil N(0,1) para significación α	$n = \frac{z^2_{1-\alpha/2}\sigma^2)}{d^2}$
Diferencia de proporciones	p_1=proporción grupo 1 p_2= proporción grupo 2 D=p_1-p_2 $z_{1-\alpha/2}$= percentil N(0,1) para significación α $z_{1-\beta}$= percentil N(0,1) para poder β p= $(p_1+p_2)/2$	$n = \frac{\left[z_{1-\alpha/2}\sqrt{2p(1-p)} + z_{1-\beta}\sqrt{p_1(1-p_1)+p_2(1-p_2)}\right]^2}{D^2}$ $n = \frac{\left[z_{1-\alpha/2} + z_{1-\beta}\right]^2}{2(\arcsin\sqrt{p_2} - \arcsin\sqrt{p_1})^2}$ para enfermedades de ocurrencia rara
Diferencia de medias	σ_1^2= varianza grupo 1 σ_2^2= varianza grupo 2 $z_{1-\alpha/2}$= percentil N(0,1) para significación α $z_{1-\beta}$= percentil N(0,1) para poder β D= diferencia de medias	$n = \frac{\left[z_{1-\alpha/2} + z_{1-\beta}\right]^2(\sigma_1^2 + \sigma_2^2)}{D^2}$

Some formulas to calculate sample size

Problema	Datos necesarios	Tamaño muestral
Estimación de una OR	OR= Odds ratio que se estima p_2= Proporción de expuestos en los controles p_1= Proporción de expuestos en los casos e=amplitud relativa del C.I. $z_{1-\alpha/2}$= percentil N(0,1) para significación α	$p_1 = \frac{ORp_2}{ORp_2 + (1-p_2)}$ $n = z^2_{1-\alpha/2} \frac{1/p_1(1-p_1) + 1/p_2(1-p_2)}{\ln(1-e)^2}$
Estimación de un RR	RR= Riesgo relativo que se estima p_2= Proporción de casos en los no expuestos p_1= Proporción de casos en los expuestos e=amplitud relativa del C.I. $z_{1-\alpha/2}$= percentil N(0,1) para significación α $z_{1-\beta}$= percentil N(0,1) para poder β	$p_1 = RRp_2$ $n = z^2_{1-\alpha/2} \frac{(1-p_1)/p_1 + (1-p_2)/p_2}{\ln(1-e)^2}$
Contraste OR>1	OR= Odds ratio que se estima p_2= Proporción de expuestos en los controles p_1= Proporción de expuestos en los casos $z_{1-\alpha/2}$= percentil N(0,1) para significación α $z_{1-\beta}$= percentil N(0,1) para poder β	$p_1 = \frac{ORp_2}{ORp_2 + (1-p_2)}$ $n = \frac{\left[z_{1-\alpha/2}\sqrt{2p_2(1-p_2)} + z_{1-\beta}\sqrt{p_1(1-p_1)+p_2(1-p_2)}\right]^2}{(p_1-p_2)^2}$
Contraste RR >1	RR= Riesgo relativo que se estima p_2= Proporción de casos en los no expuestos p_1= Proporción de casos en los expuestos $z_{1-\alpha/2}$= percentil N(0,1) para significación α $z_{1-\beta}$= percentil N(0,1) para poder β	$p_1 = RRp_2$ $n = \frac{\left[z_{1-\alpha/2}\sqrt{2p_2(1-p_2)} + z_{1-\beta}\sqrt{p_1(1-p_1)+p_2(1-p_2)}\right]^2}{(p_1-p_2)^2}$

Vall d'Hebron Research

UNITAT D'ESTADÍSTICA I BIOINFORMÀTICA

Some tips

- Sample size goes up
  - for smaller α
  - For higher β
  - For smaller δ
  - For higher σ
  - For p closer to 50%
- Sample size is higher for proportions than means
- Sample size must be calculated a priori. Is not sensible to calculate power after
- SD can be calculated from 95% CI
- Upper-Lower limit of a CI is about 4 Standard Error and $SE=s/\sqrt{n}$
- Some % of survivors can be obtained from Kaplan-Meier survival curves and can be used for calculations
- Sample size is not an exact science and must be the product of calculations and reality

Effect size δ

Is the number measuring the strength of the relation between two (o more) group of comparison. It can be calculated from the data.

- Mean difference of a quantitative measure between two populations
- Difference of proportions of a dichotomous measure in two populations
- Correlation coefficient between two quantitative variatons
- Regression coeficient from a multivariable regression model
- Risk difference, Relative Risk, Odds Ratio, Hazard Ratio

It is preferred that the effect size measurement is standardized

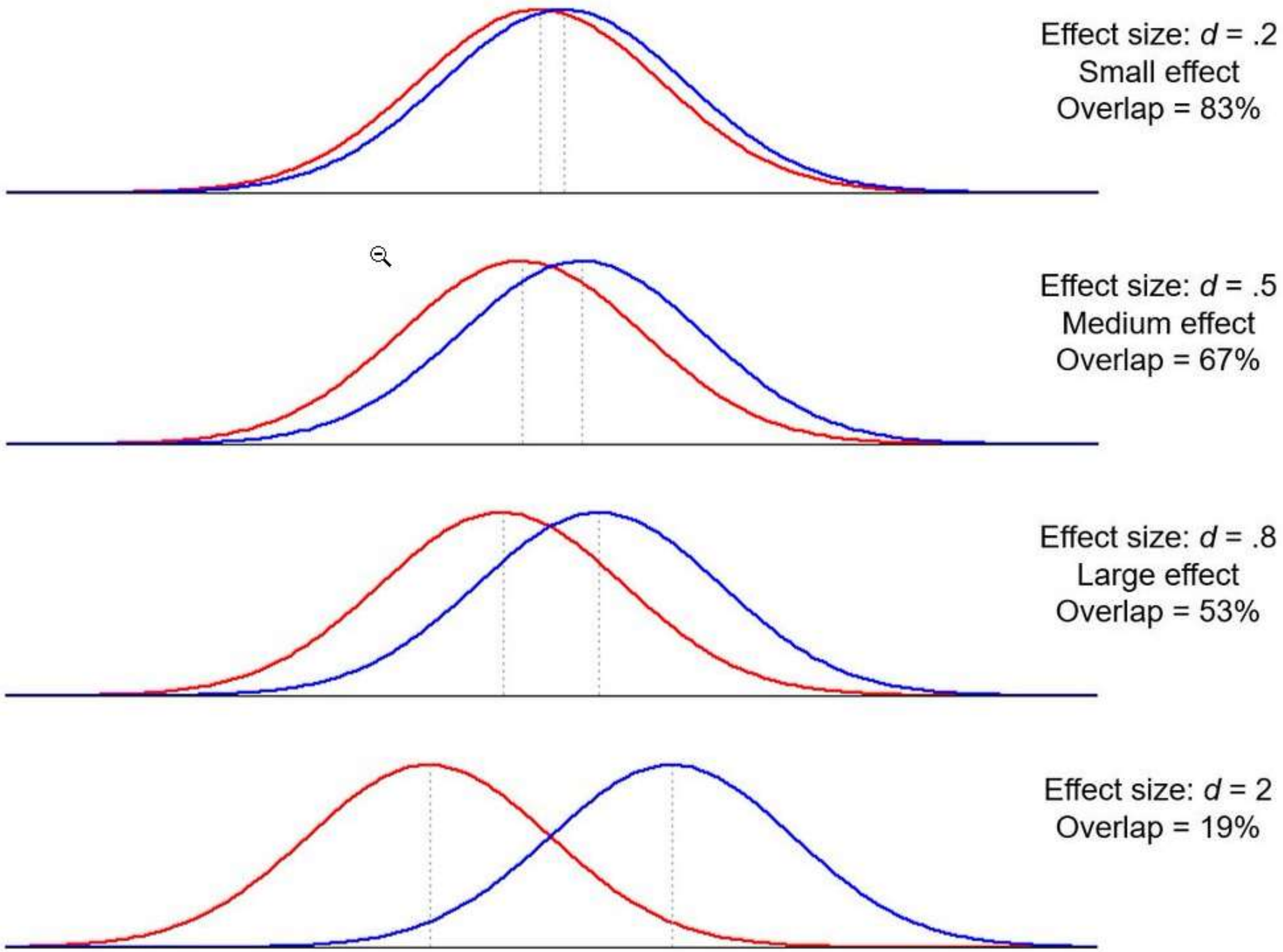
- Cohen d $\quad d = \frac{\bar{x}_1 - \bar{x}_2}{s}.$
- Cohen w $\quad w = \sqrt{\sum_{i=1}^{m} \frac{(p_{1i} - p_{0i})^2}{p_{0i}}}$
- Cohen f^2 $\quad f^2 = \frac{R^2}{1 - R^2}$
- Eta-squared (η^2) $\quad \eta^2 = \frac{SS_{\text{Treatment}}}{SS_{\text{Total}}}.$

Table 1 Values of Effect Sizes and Their Interpretation

Kind of Effect Size	*Small*	*Medium*	*Large*
r	.10	.30	.50
d	0.20	0.50	0.80
η^2_p	.01	.06	.14
f^2	.02	.15	.35

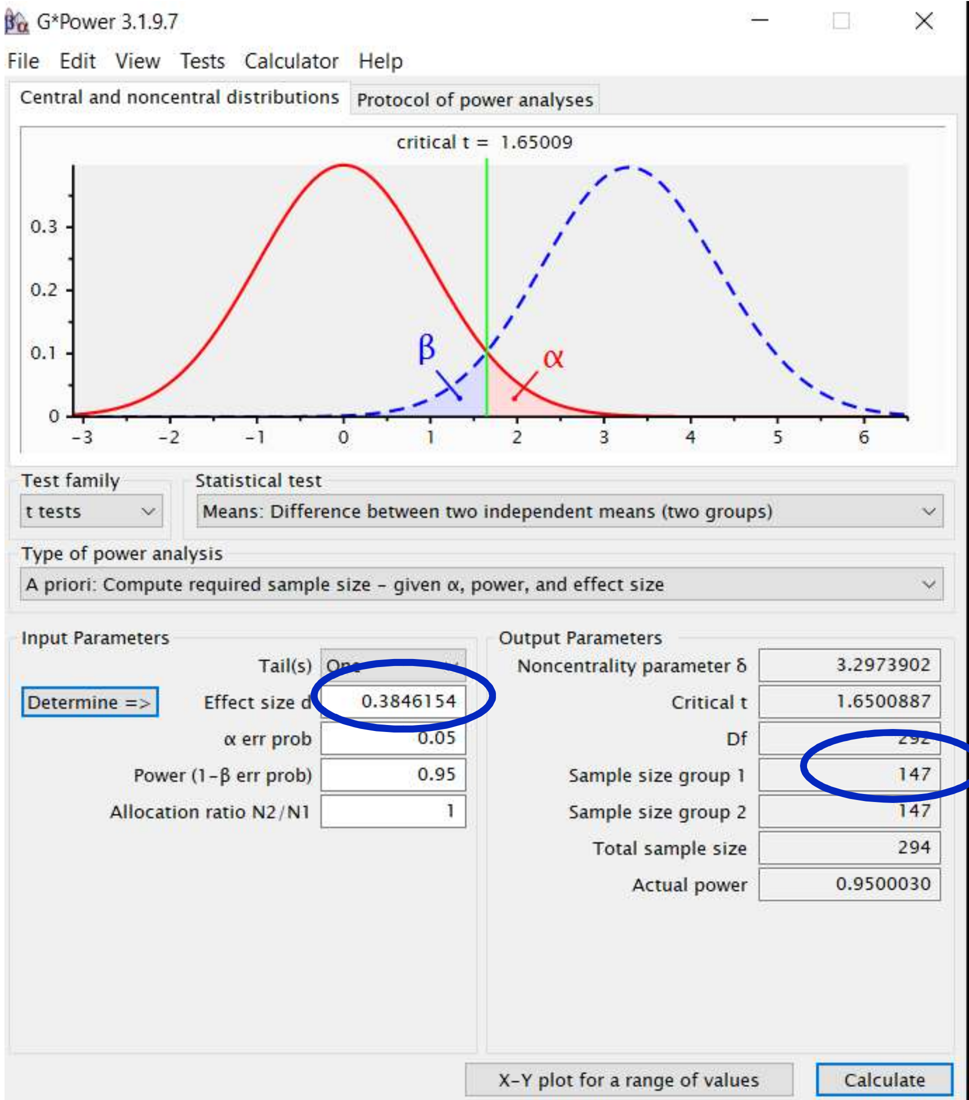
Source: Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*, 155–159. doi:10.1037/0033-2909.112.1.155

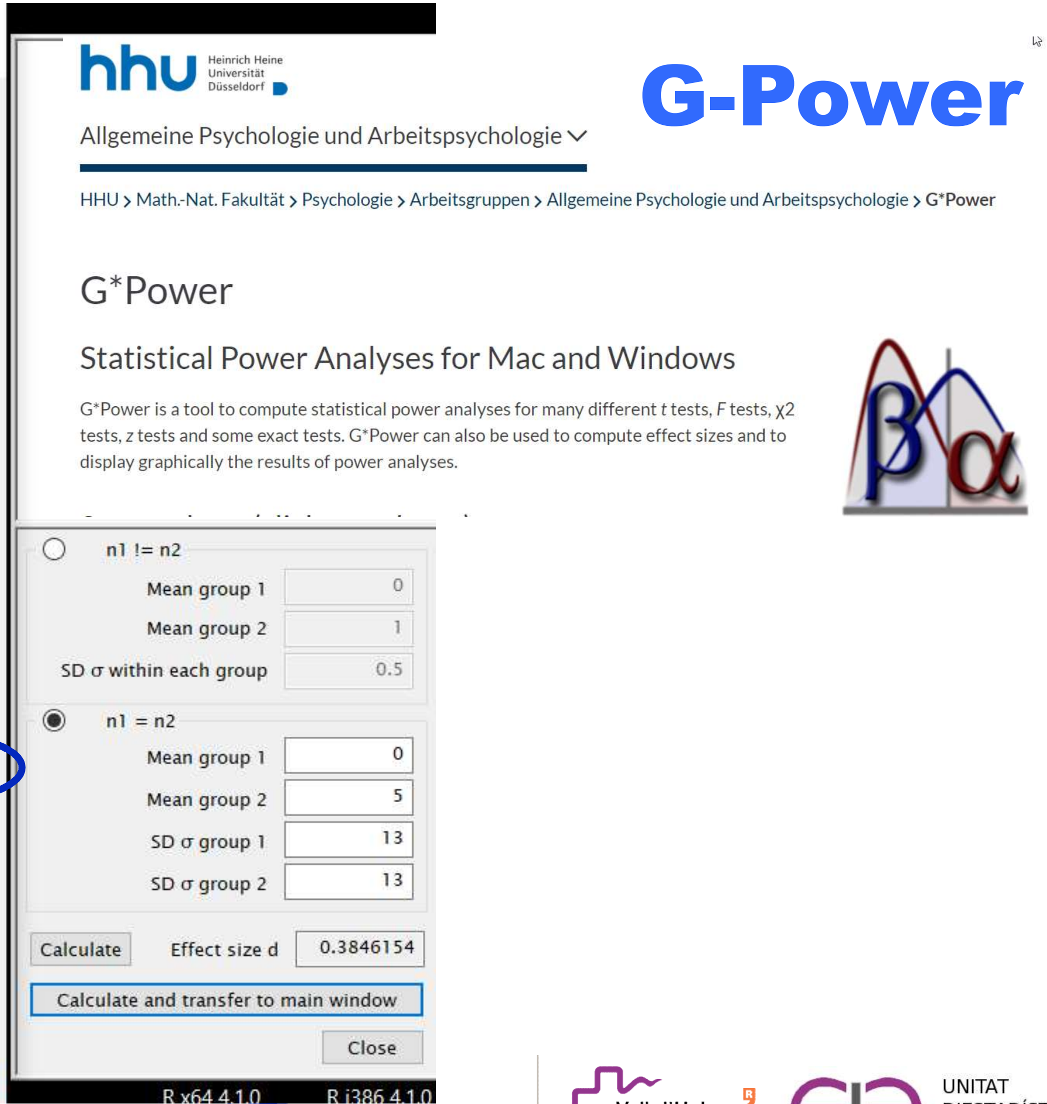
Understanding Effect Sizes



UNITAT
D'ESTADÍSTICA I
BIOINFORMÀTICA

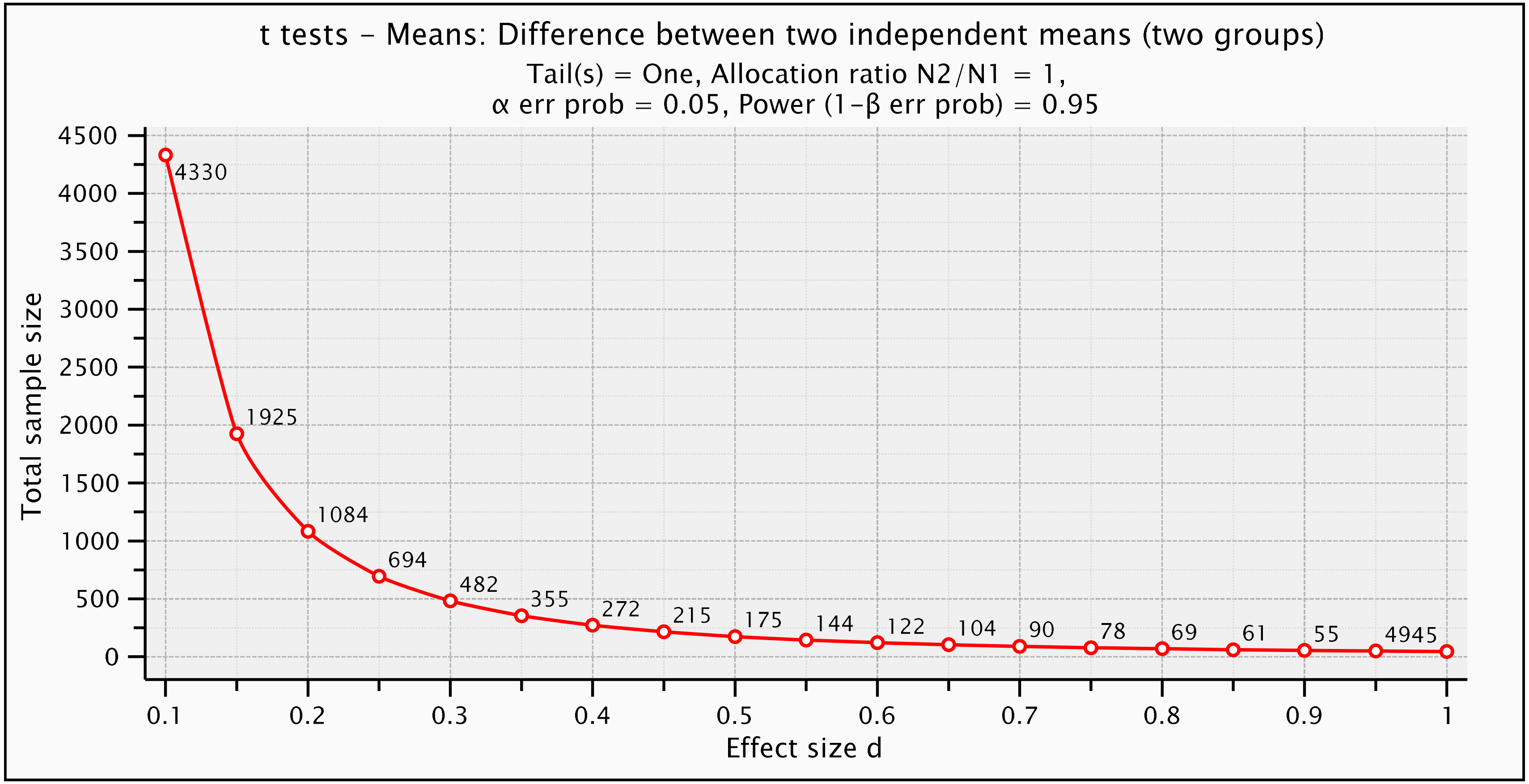
G-Power



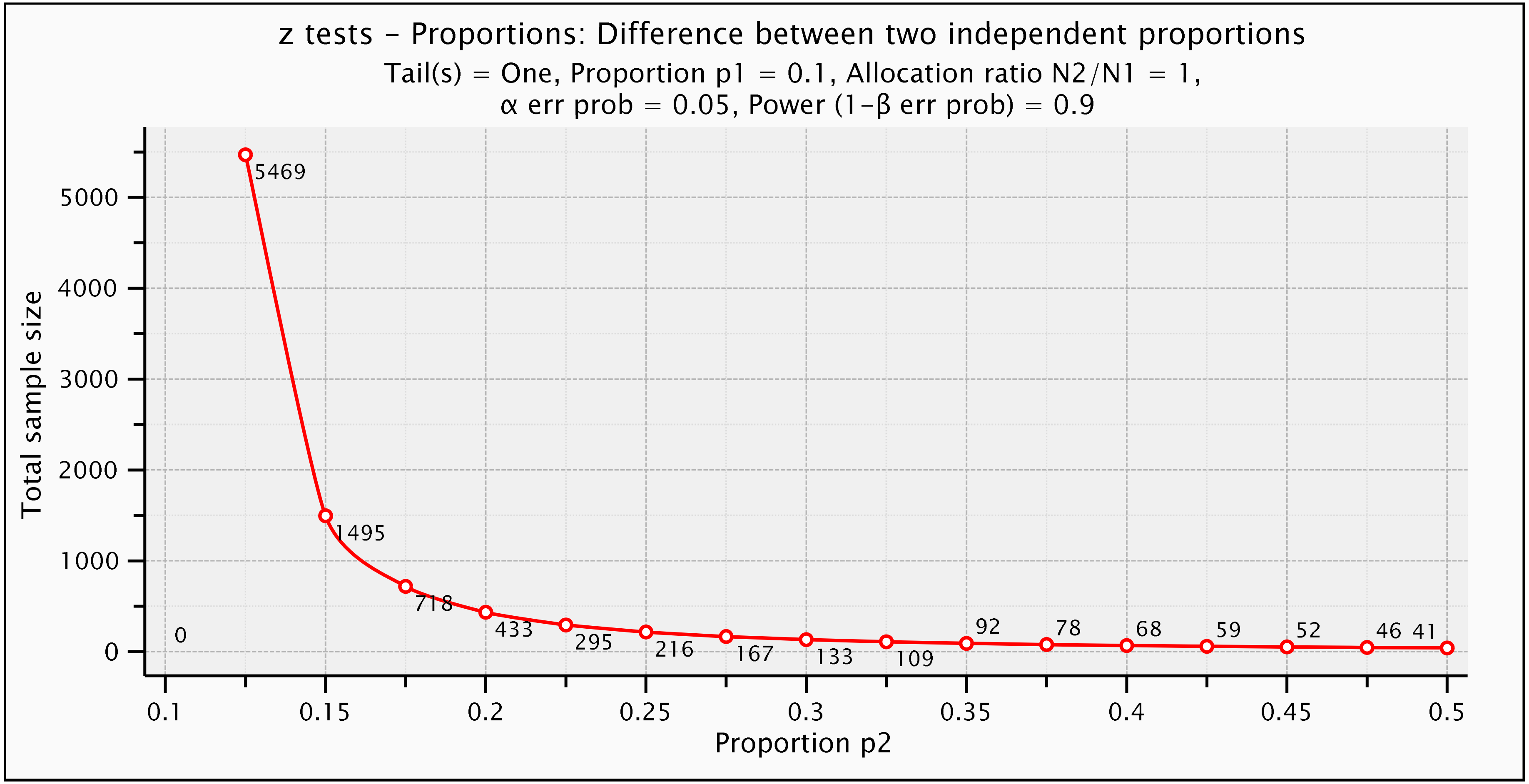


Vall d'Hebron Research

UNITAT D'ESTADÍSTICA I BIOINFORMÀTICA

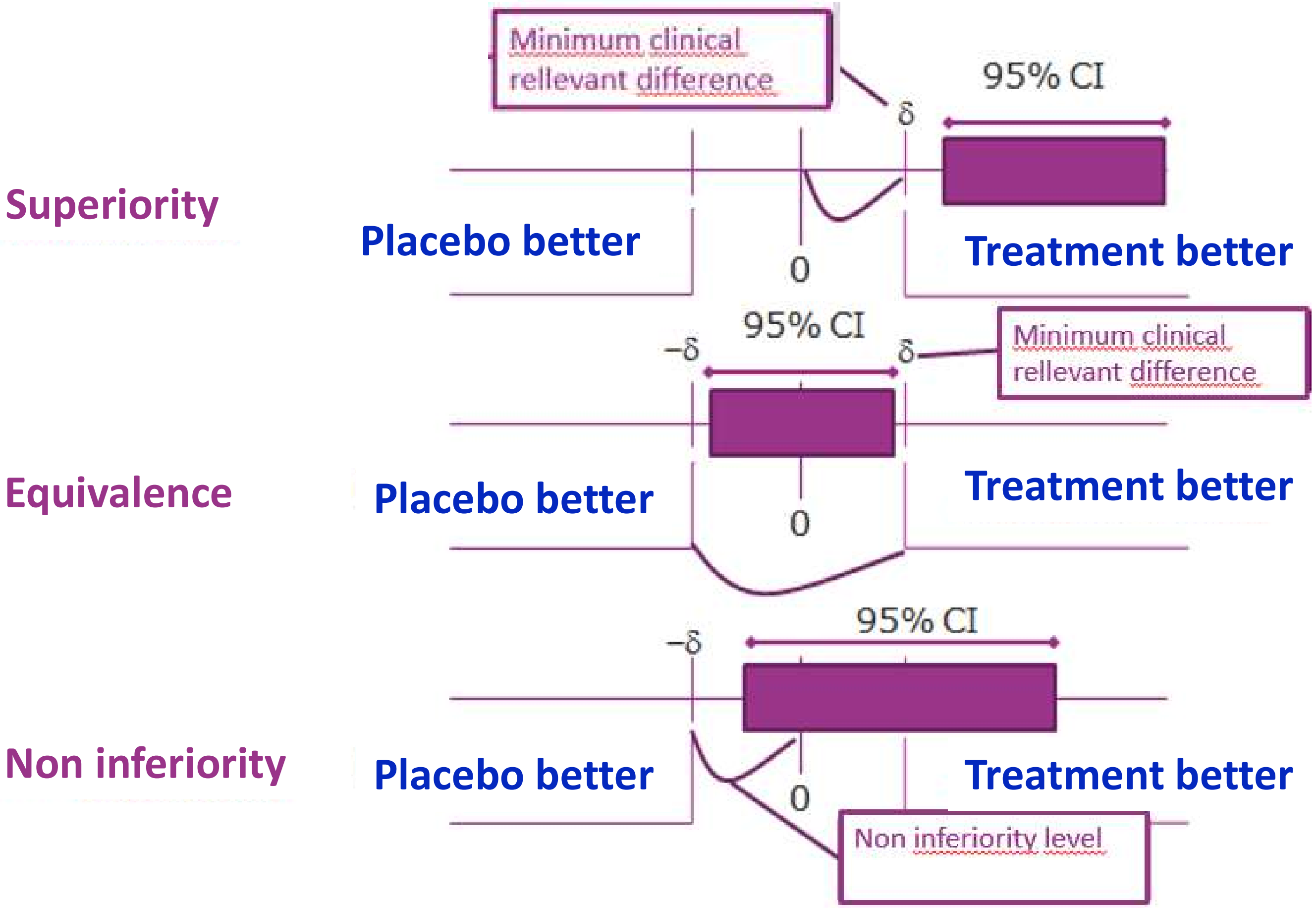


UNITAT
D'ESTADÍSTICA I
BIOINFORMÀTICA



z tests – Proportions: Difference between two independent proportions
Tail(s) = One, Proportion p1 = 0.1, Allocation ratio N2/N1 = 1,
α err prob = 0.05, Power (1-β err prob) = 0.9
Total sample size
Proportion p2
5469
1495
718
433
295
216
167
133
109
92
78
68
59
52
46
41
0
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.45
0.5
0
1000
2000
3000
4000
5000

Equivalence and Non inferiority Trials



UNITAT D'ESTADÍSTICA I BIOINFORMÀTICA

51 Journal of Thoracic Disease, Vol 1, No 1, December 2009

Practical Biostatistics

How to Calculate Sample Size in Randomized Controlled Trial?

Baoliang Zhong, MD

Sample size goes up

Table 1 Hypothesis testing of different design

design	null hypothesis	alternative hypothesis	test statistics
non-inferiority	H0:T-S=-δ	Ha: T-S>-δ	Z=(d+δ)/sd
equivalence	H10:T-S=-δ	H1a: T-S>-δ	Z1=(d+δ)/sd
	H20:T-S=δ	H2a: T-S<δ	Z2=(δ -d)/sd
statistical superiority	H0:T-S=0	Ha: T-S>0	Z=d/sd
clinical superiority	H0:T-S=δ	Ha: T-S>δ	Z=(d-δ)/sd

T: new treatment; S: standard treatment; δ: clinically admissible margin of non–inferiority/equivalence/ superiority; d: the effectiveness difference between T and S; sd: the standard error of d; Z: Z obeys standard normal distribution.

Vall d'Hebron | Vall d'Hebron Research | UNITAT D'ESTADÍSTICA I BIOINFORMÀTICA

Schumi and Wittes *Trials* 2011, **12**:106
http://www.trialsjournal.com/content/12/1/106

TRIALS

REVIEW Open Access

Through the looking glass: understanding non-inferiority

Jennifer Schumi* and Janet T Wittes

Sample size goes up

Table 2 Approximate sample sizes required for non-inferiority comparison of proportions

True proportion in active control	Non-inferiority bound using 10% margin	Approximate sample size per group assuming 1:1 randomization to new treatment and control required under:		
		Equal effects	5% benefit	10% benefit
0.1	0.09	19,200	8,725	5,050
0.2	0.18	8,500	3,900	2,250
0.3	0.27	4,970	2,260	1,300
0.4	0.36	3,200	1,450	825
0.5	**0.45**	**2,100**	**1,000**	**550**
0.6	0.54	1,440	640	360
0.7	0.63	930	405	225

Sample sizes calculated using Pass 2008 methods for non-inferiority tests of two independent proportions, using the Z statistic with continuity correction and pooled variance, with a target power of 90% and α level of 0.025.

Vall d'Hebron | Vall d'Hebron Research | UNITAT D'ESTADÍSTICA I BIOINFORMÀTICA

Interim analysis

Innovative Trial Designs in GI Drug Development: Why Trials Succeed and Fail—A Brief Report From 2018 Digestive Disease Week

M. Scott Harris[1,2,4] and Colin W. Howden[3,]

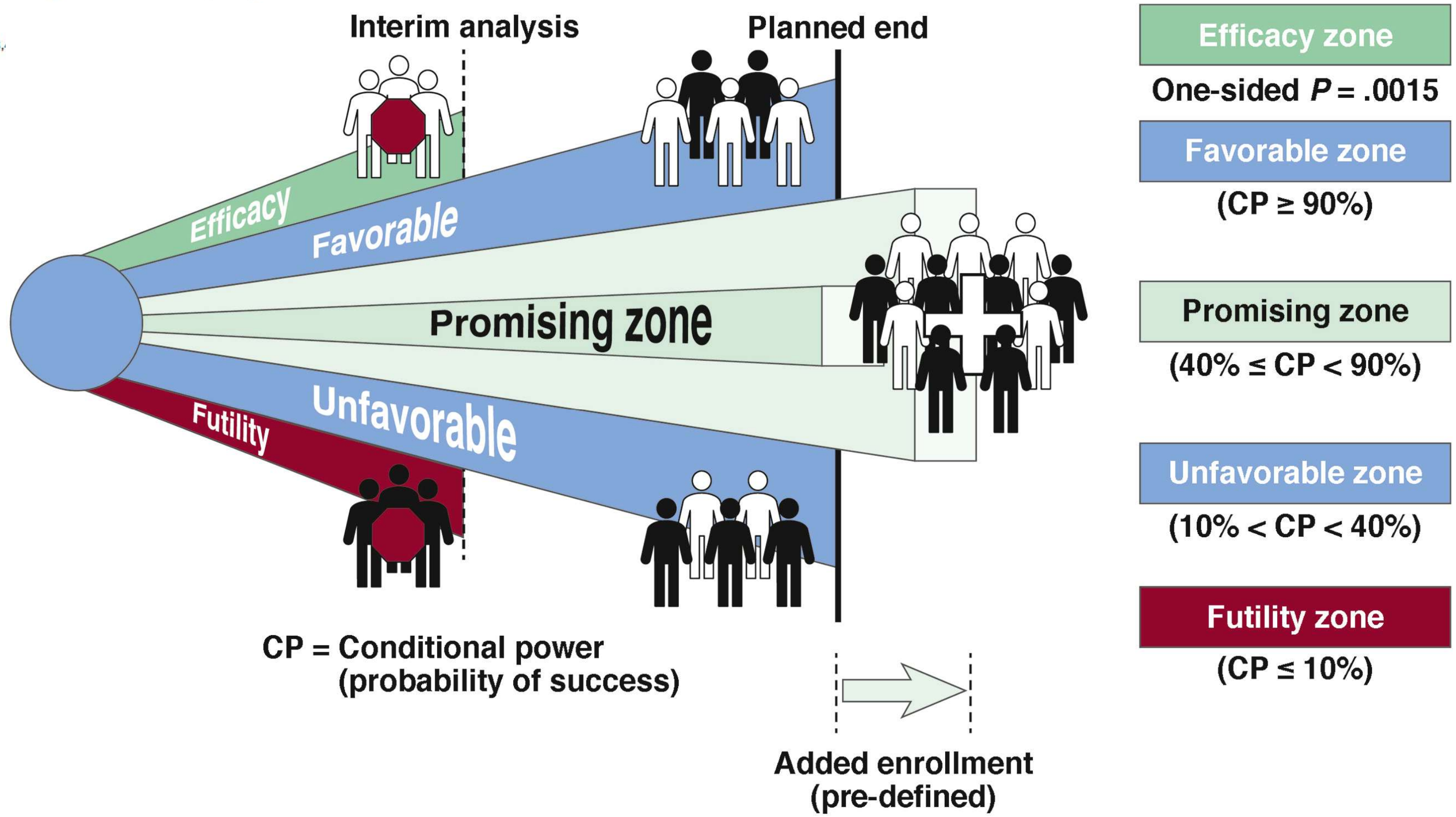
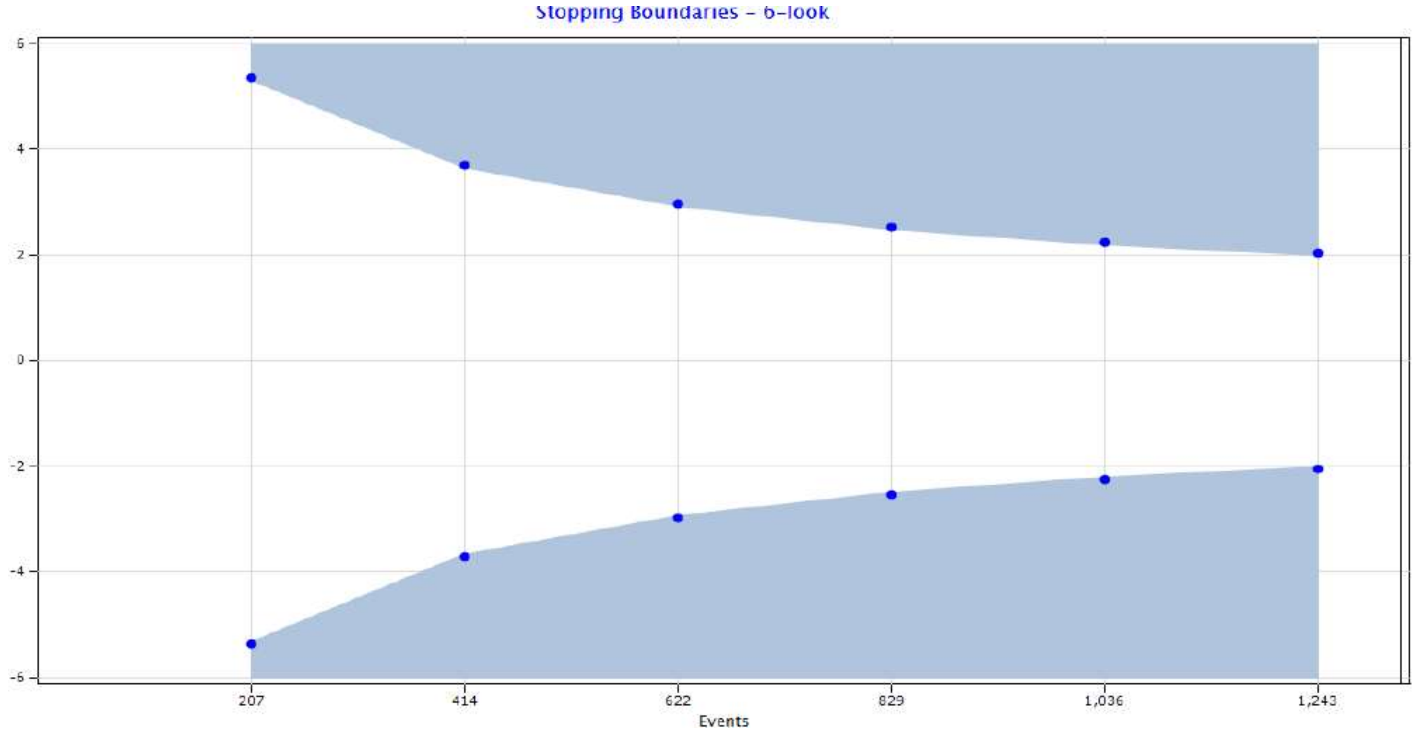


Figure 1In blinded sample size re-estimation, the study commences with an optimistic sample size, assuming a large treatment effect. Patients can be added at IA if the conditional power (CP) falls in the promising zone, the example showing CP ≥ 40% and ≤ 90%. A favorable (≥90%) or unfavorable (10% to <40%) CP would result in no change in enrollment. The trial could also be stopped at IA for success ($P \leq .0015$) or futility (CP of <10%). The starting sample size of the blinded sample size re-estimate is typically smaller than grouped sequential design because of the flexibility of adding patients at the interim analysis (IA). The study is conducted as a same phase trial, because unlike a phase II/III 2-stage design, no change in dose, treatment, subpopulation, or study phase is implemented at the IA. Modified with permission from Mehta CR.[5]

Six look sequential desing

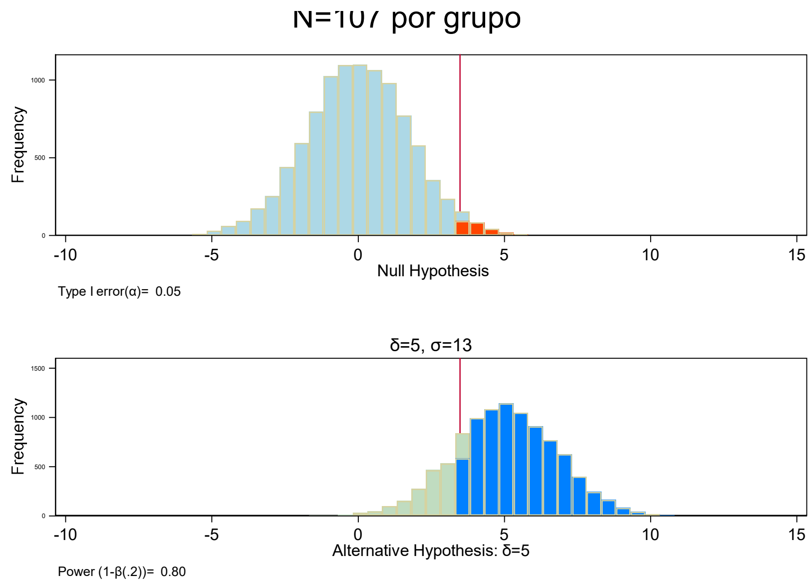


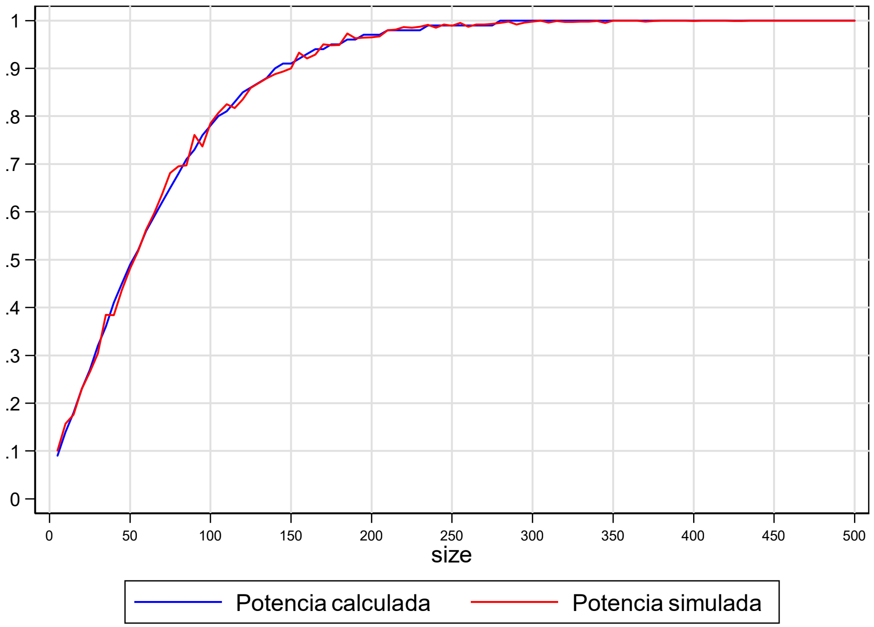
UNITAT D'ESTADÍSTICA I BIOINFORMÀTICA

The present and the future of sample size: SIMULATION

- Simulation: Computing intensive approach that mimics the data generating process to evalute properties of a test statistic or confidence interval (Power)
- Model the population parameters of the study(Survival function, Regression models, etc. Lost of follow up , Atrition rate, Changes in treatments, etc.
- Sampling strategy and simulate sampling
- Analytics methods and analysis model :
- Performance of the model (power and precision).
- Calculated Distribution of the effects measurements
- Select sample size

Vall d'Hebron | Vall d'Hebron Research | UNITAT D'ESTADÍSTICA I BIOINFORMÀTICA





Example Simulation mean difference

1. Select Parameters
   μ_0-μ_0=0
   μ_1-μ_0=δ=5
   σ=13
2. Select series of sample size n=10,20...,490,500
3. Calculate ee=Standard error=$\sigma\sqrt{\frac{1}{n}+\frac{1}{n}}$
4. Sampling 1000 from Normal(δ,ee)
5. Calculate Critical value=cv=$z_{1-\alpha/2}\sigma\sqrt{\frac{1}{n}+\frac{1}{n}}$
6. Calculate power as % of observations over cv
7. Plot sample size against power

Vall d'Hebron
Vall d'Hebron Research
UNITAT D'ESTADÍSTICA I BIOINFORMÀTICA

AMERICAN JOURNAL OF INDUSTRIAL MEDICINE 44:532–539 (2003)

Power Calculations for Survival Analyses Via Monte Carlo Estimation

David B. Richardson, PhD*

Example Simulation Clinical Trial

- Period follow up 20 years
- Two groups(x) equally distributed(1:1) (sampling from binomial p=0.50
- Generate I cohorts of size N
- Effect size ϕ were hazard ratio=exp(ϕ)
- Baseline risk exponential exp(δ)
- Disease status= log(h)= δ+ ϕ x
- Calculate proportional hazard
- Power= proportion fof pvalues less than .05 for a big number of repetitions.

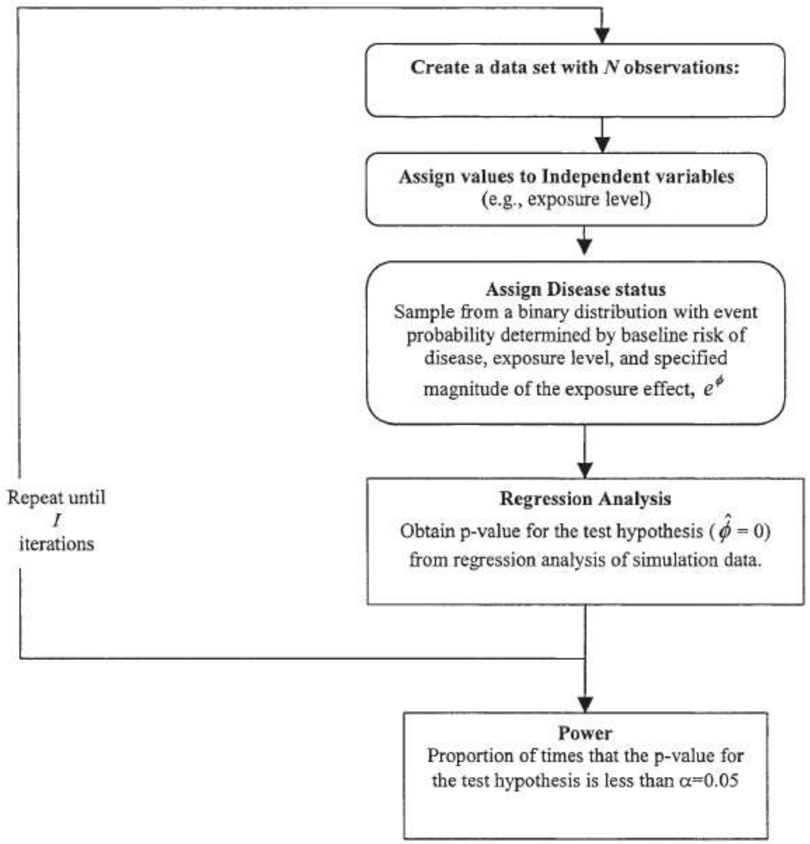


FIGURE 1. Flowchart of the process followed for Monte Carlo estimation of statistical power.

Vall d'Hebron | Vall d'Hebron Research | UNITAT D'ESTADÍSTICA I BIOINFORMÀTICA

AMERICAN JOURNAL OF INDUSTRIAL MEDICINE 44:532–539 (2003)

Power Calculations for Survival Analyses Via Monte Carlo Estimation

David B. Richardson, PhD*

Parameter	Example 1
Type I error probability (α)	0.05
Number of iterations (I)	1,000
Number of persons in study cohort (N)	2,500, 5,000, 10,000
Exposure effect (e^{δ})	1.01, 1.05, 1.10, 1.25
Length of follow-up (in years)	20
Exposure (x)	B (1,0.5)
Age at entry (in years)	n.a.
Length of employment (in years)	n.a.
Baseline annual mortality rate (e^{δ}) (e^{δ_0} in Example 2—4)	0.0084

TABLE II. Estimates of Statistical Power by Total Study Size and Exposure Effect Size for a Cohort Study Comparing Survival in Exposed Versus Unexposed Groups of Equal Size (Example 1)

Study size	Relative risk (hazard ratio)			
	1.01	**1.05**	**1.10**	**1.25**
2,500	0.04[a]	0.07[a]	0.15[a]	0.62[a]
	0.03[b]	0.07[b]	0.16[b]	0.63[b]
5,000	0.05[a]	0.10[a]	0.25[a]	0.91[a]
	0.03[b]	0.10[b]	0.27[b]	0.90[b]
10,000	0.05[a]	0.16[a]	0.50[a]	1.00[a]
	0.04[b]	0.16[b]	0.48[b]	1.00[b]

[a]Statistical power estimated by Monte Carlo approach under conditions specified in Table I using the program in Appendix 1. (Values shown include estimates of power derived via Monte Carlo simulation.)
[b]Statistical power calculated using the SSP internet based computer program for power and sample size calculations under conditions specified in Table I. (Values shown include estimates of power derived via the method described by Lachin and Foulkes.)

UNITAT D'ESTADÍSTICA I BIOINFORMÀTICA

Pye et al. BMC Res Notes (2016) 9:90
DOI 10.1186/s13104-016-1893-x
BMC Research Notes

TECHNICAL NOTE Open Access

When is enough, enough? Understanding and solving your sample size problems in health services research

Victoria Pye*, Natalie Taylor, Robyn Clay-Williams and Jeffrey Braithwaite

Next time you want to calculate a sample size do not forget to know the Distribution of the Outcome, the Effect Size, and the possible Confounding variables

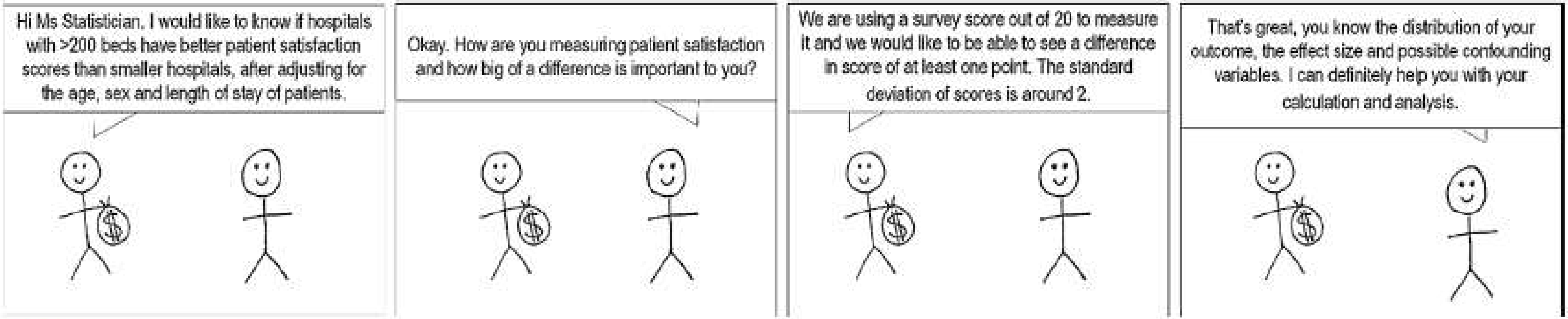


Fig. 4 A statistician's dream

Vall d'Hebron | Vall d'Hebron Research | UNITAT D'ESTADÍSTICA I BIOINFORMÀTICA

Summary

- Important step in designing in study
- Although for simple situation is easy to calculate sample size, when having several outcomes and covariates performing a sample size can be complicated
- Because sample size errors and costs related to performing a RCT is recomended to take caution and ask a for statistical advice
- Simulation is the actual option to perform sample size before starting a RCT

Gràcies \ Gracias \Thank you

Vall d'Hebron

Vall d'Hebron Research