Sample size in Animal Studies

Santi Pérez-Hoyos

Statistics and Bioinformatics Unit (UEB)

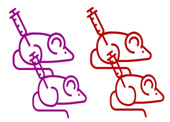
Topics

- Introduction
- Statistical concepts in the calculation of sample size
  - Effect size
  - Standard deviation or variability
  - Power
  - Level of significance
- Sample size calculation for dichotomous response variables in 2 groups
- Sample size calculation for continuous response variables in 2 groups
- Software for sample size calculation: Granmo, G Power
- Practical examples in different experimental designs

Sample size Calculation In Animal Studies

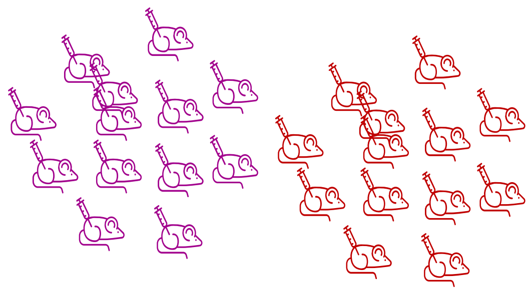
- Experiments using laboratory animals should be well designed, efficently executed, correctly analysed, clearly presented and correctly interpreted
- Sample size calculated obtained from similar studies or "our experience" are not justified
- Sample size from resource equation can be used in exploratory analysis but it is not good option for testing of hypothesis.
- Power analysis used to calculate sample size is the most scientific method to calculate sample size.

Too few animals



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

Too many animals



- Waste of resources
- Give a harmful treatment to animals
- Identify irrelevant treatment effects “significative”
- Potentially Not ethical

Billion dollar Question

How many animals do we need for our Study

UNITAT
D'ESTADÍSTICA I
BIOINFORMÀTICA

Before calculating sample size

- Have a Clear Research Question
- Identify the outcomes or end-points
- Know how variables are going to be measured (frequency, proportions, ordinal, means)
- How many subjects you can afford
- Have a clear idea of the expected results

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

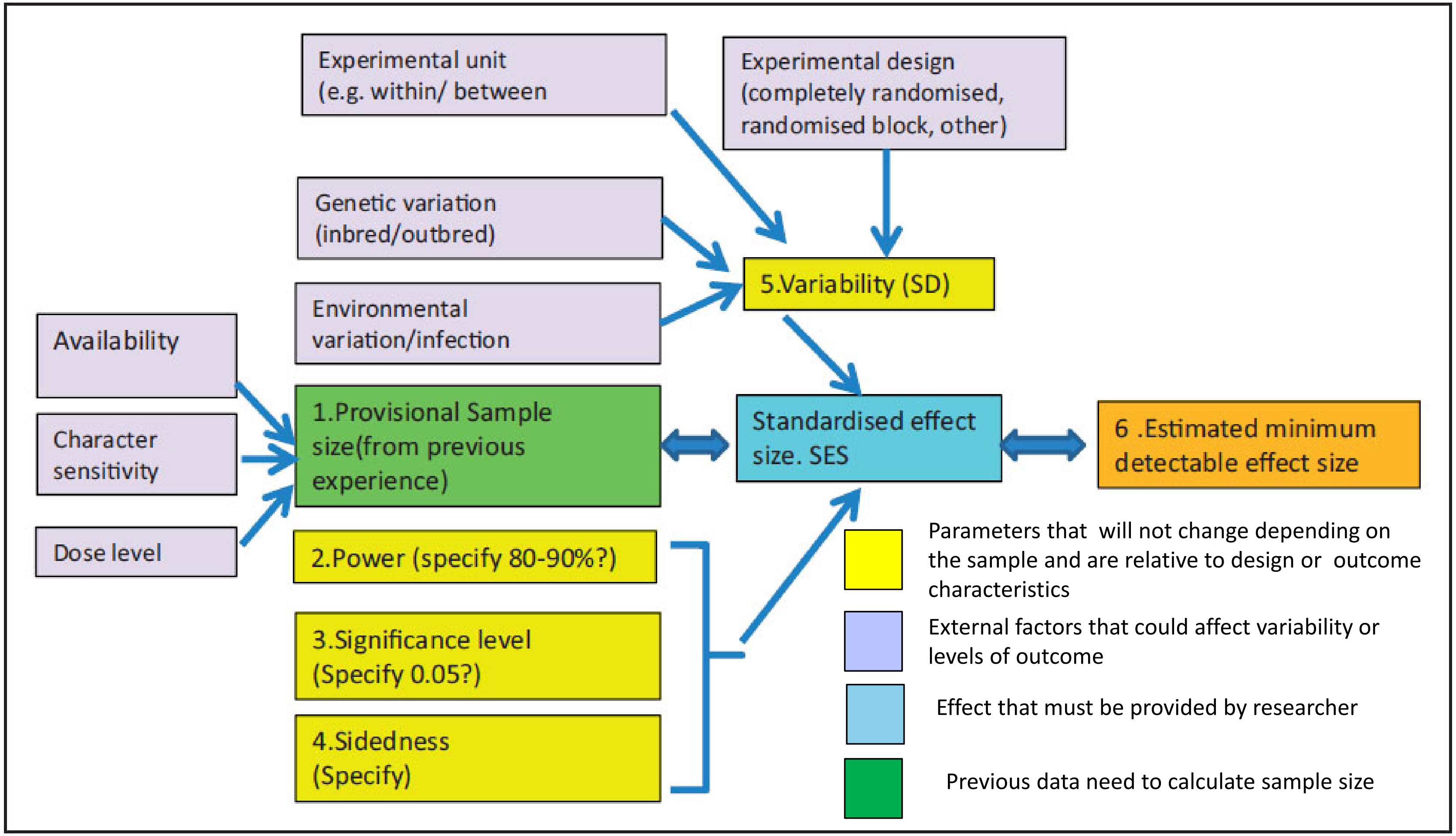
Preliminary concepts to be used

- Effect size
- Standard error of an estimator
- Confidence interval
- Statistical test
- Type I error
- Power of a test (1-β)

Sample size based on Hypothesis testing

Effect Size δ	Minimum detectable difference betweeen the two groups to compare
Type I Error α	Probability of rejecting the Null Hypothesis when its true (5%)
Power 1- β(Type II Error)	Probability of rejecting the Null Hypothesis when its false (80%,90%)
Variability of the data σ^2	Variance of the data

Factors de calculate sample size



UNITAT D'ESTADÍSTICA I BIOINFORMÀTICA

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

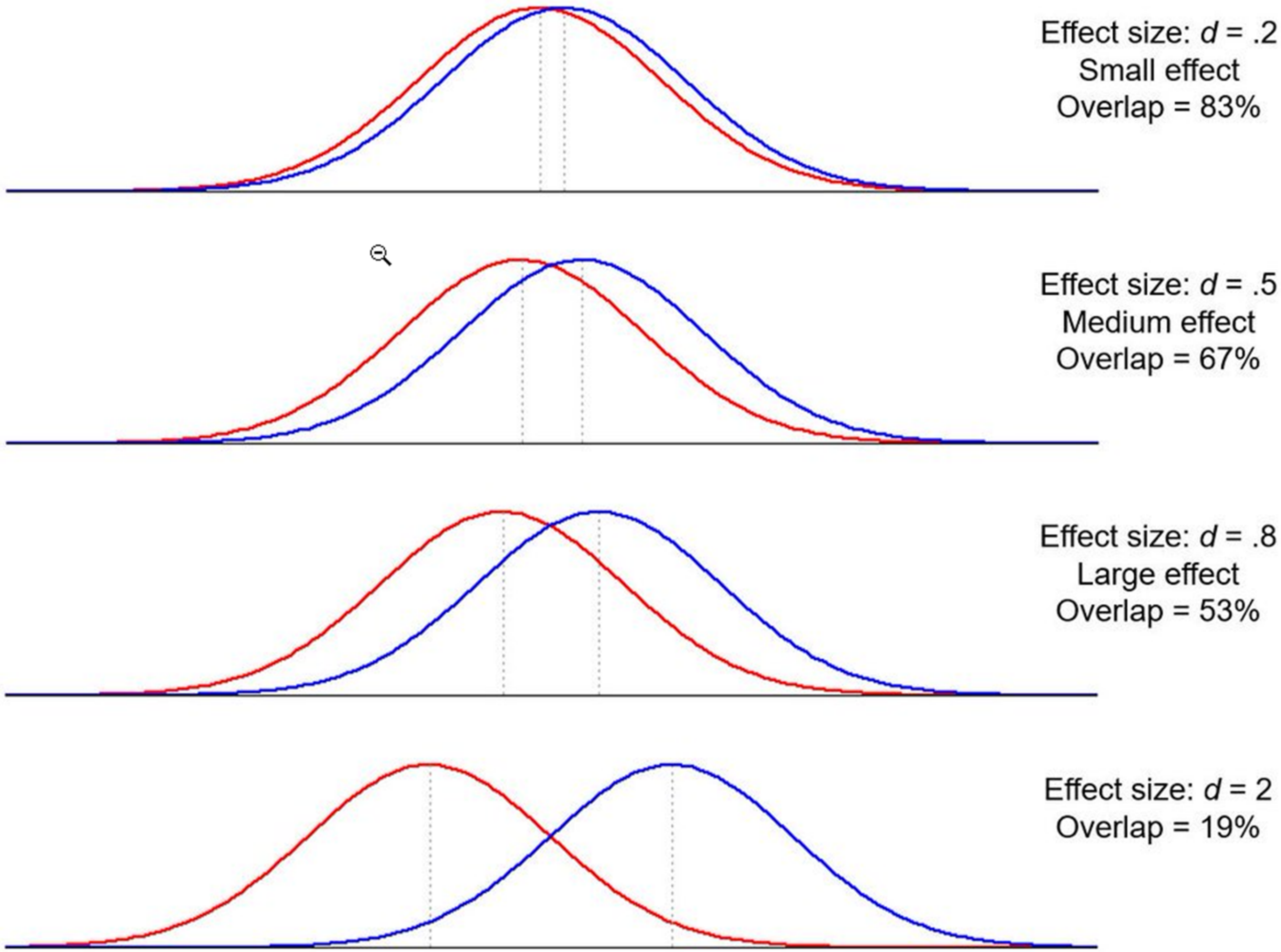
- Cohen d $$d = \frac{\bar{x}_1 - \bar{x}_2}{s}.$$
- Cohen w $$w = \sqrt{\sum_{i=1}^{m} \frac{(p_{1i} - p_{0i})^2}{p_{0i}}}$$
- Cohen f^2 $$f^2 = \frac{R^2}{1 - R^2}$$
- Eta-squared (η^2) $$\eta^2 = \frac{SS_{\text{Treatment}}}{SS_{\text{Total}}}.$$

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

Effect size δ

TABLE 3. Thresholds for interpreting the effect size

Test	Relevant effect size	Effect Size (ES)		
		Small	Medium	Large
t-test for means	Cohen's d	0.2	0.5	0.8
Chi-Square	Cohen's ω	0.1	0.3	0.5
r x c frequency tables	Cramer's V or Phi	0.1	0.3	0.5
Correlation studies	***r***	0.2	0.5	0.8
2 x 2 table case control	Odd Ratio (OR)	1.5	2	3
2 x 2 table cohort studies	Risk Ratio (RR)	2	3	4
One-way an(c)ova (regression)	Cohen's f	0.1	0.25	0.4
ANOVA (for large sample) ANOVA (for small size)	Eta Square η^2 Omega square Ω^2	0.01	0.06	0.14
Friedman test	Average spearman Rho	0.1	0.3	0.5
Multiple regression	η^2	0.02	0.13	0.26
Coefficient of determination	r^2	0.04	0.25	0.64
Number needed to treat	NNT	1 / Initial risk		

Serdar CC, Cihan M, Yücel D, Serdar MA. Sample size, power and effect size revisited: simplified and practical approaches in pre-clinical, clinical and laboratory studies. Biochem Med (Zagreb). 2021 Feb 15;31(1):010502. doi: 10.11613/BM.2021.010502. Epub 2020 Dec 15. PMID: 33380887; PMCID: PMC7745163

How to proceed to calculate sample size

- Plan the experiment
- From previous studies obtain estimates of the Mean and Standard deviation
- Decide the precision of the Confidence interval
- Calculate de Effect size according to the design and comparison measures
- Select the type I error (alpha =0.05) and the Power (at least 0.80)
- Proceed to calculate the sample size using formulas or specific software

Sample size based on Hypothesis testing

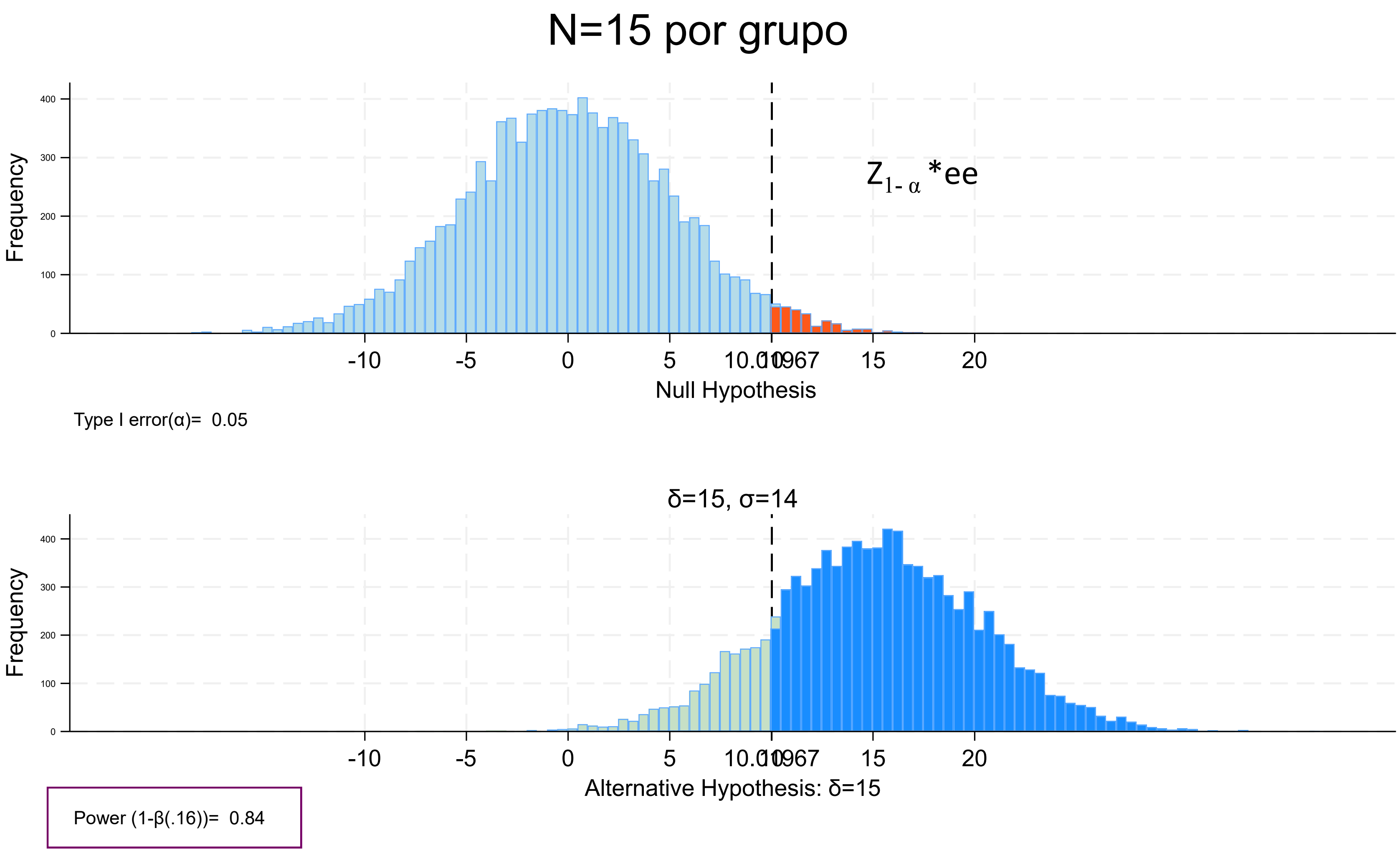
Imagine a follow up Study

- Two groups with N=30(15 by group)
- H_0: No differences $\delta=0$
- Standard deviation $\sigma=14$

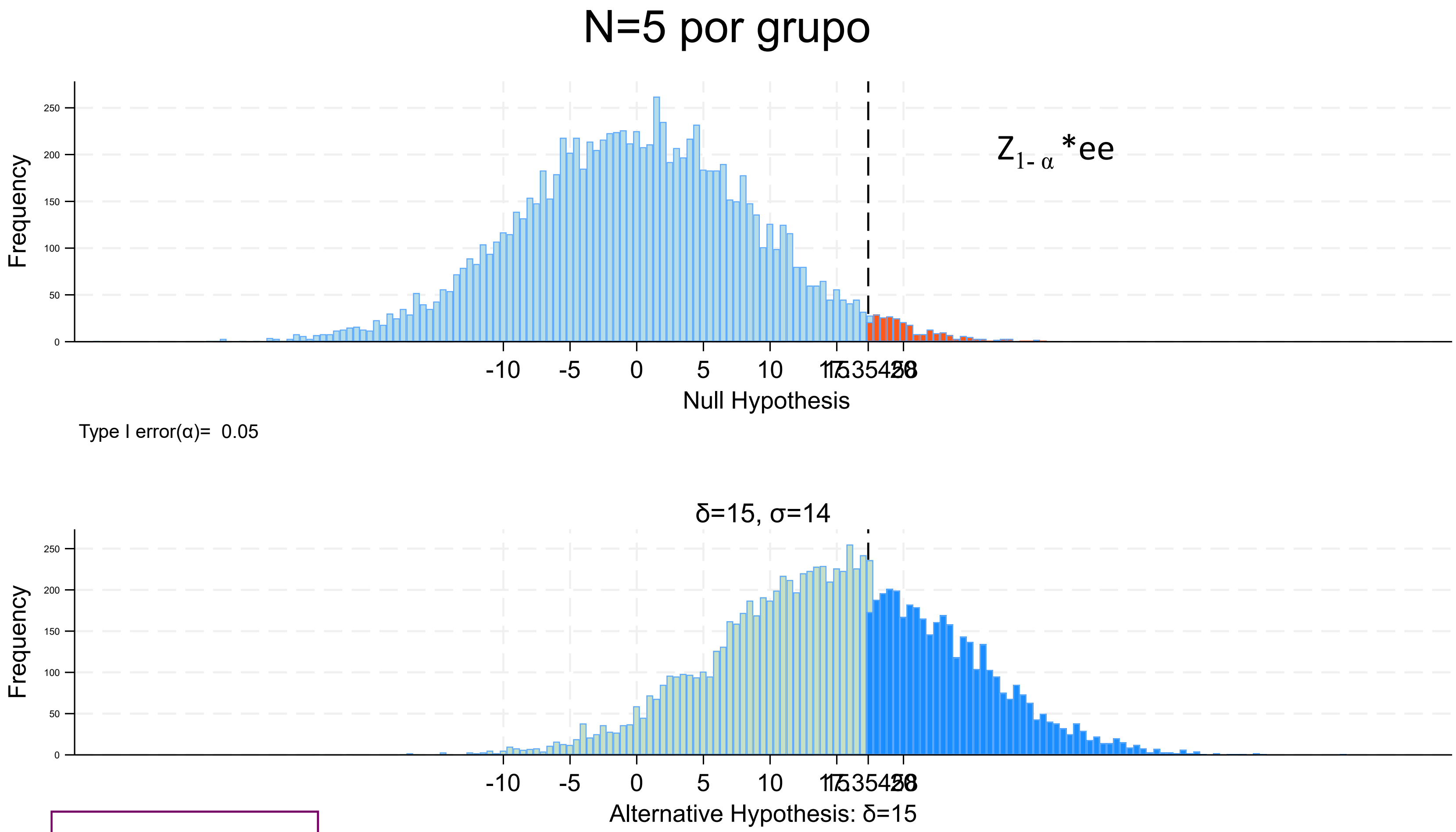
If H_0: is not true we observe

Difference $\delta=15$ (Effect size)

Sample size based on Hypothesis testing



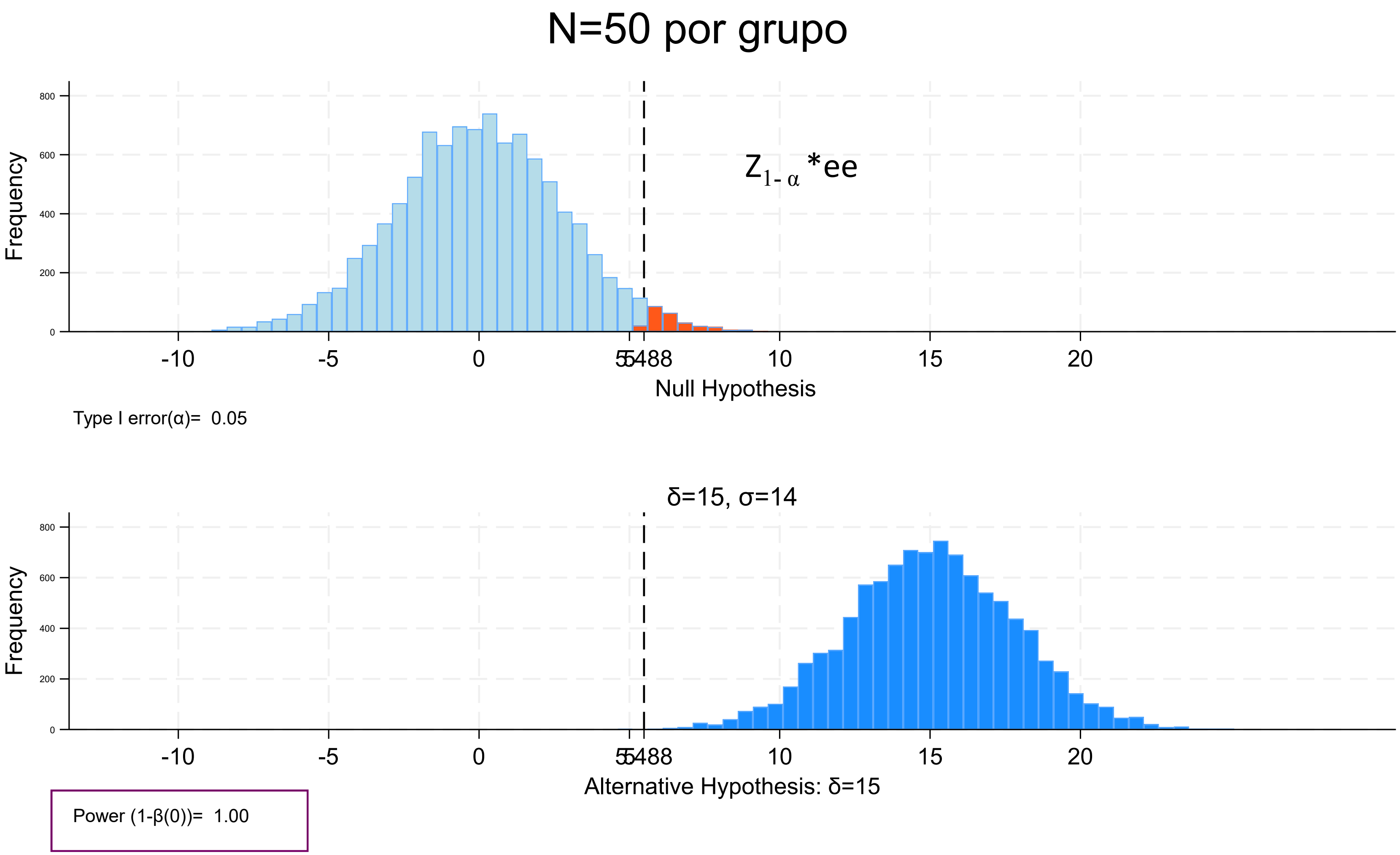
Sample size based on Hypothesis testing



Power (1-β(.6))= 0.40

UNITAT D'ESTADÍSTICA I BIOINFORMÀTICA

Sample size based on Hypothesis testing



UNITAT
D'ESTADÍSTICA I
BIOINFORMÀTICA

Sample size based for mean differences

We want

$$Z_{1-\alpha/2} * ee = \delta - Z_{1-\beta} * ee$$

We know α,β and $ee=\sigma/\sqrt{n}$ if σ is the same for both groups

$$n = \frac{2\sigma^2 \left(Z_{1-\frac{\alpha}{2}} + Z_{1-\beta}\right)^2}{\delta^2}$$

Sample size for mean differences

https://apisal.es/Investigacion/Recursos/granmo.html

UNITAT D'ESTADÍSTICA I BIOINFORMÀTICA

Sample size based for mean proportions

We want

$$Z_{1-\alpha/2} * ee = \delta - Z_{1-\beta} * ee$$

We know α,β and $ee = p(1-p)/\sqrt{n}$ i $\delta = p_2 - p_1$

$$n = \frac{\left(Z_{1-\frac{\alpha}{2}} + Z_{1-\beta}\right)^2 \left(p_1(1-p_1) + p_2(1-p_2)\right)}{\delta^2}$$

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

Sample size for differences in proportions

https://apisal.es/Investigacion/Recursos/granmo.html

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	$n = \dfrac{z_{1-\alpha/2}^2\, p(1-p)}{d^2}$ $n = \dfrac{z_{1-\alpha/2}^2 (1-p)}{e^2}$
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	$n = \dfrac{z_{1-\alpha/2}^2 \sigma^2)}{d^2}$
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 $p= (p_1+p_2)/2$	$n = \dfrac{\left[z_{1-\alpha/2}\sqrt{2p(1-p)} + z_{1-\beta}\sqrt{p_1(1-p_1)+p_2(1-p_2)}\right]^2}{D^2}$ $n = \dfrac{\left[z_{1-\alpha/2} + z_{1-\beta}\right]^2}{2(\arcsin\sqrt{p_2} - \arcsin\sqrt{p_1})^2}$ 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	$n = \dfrac{\left[z_{1-\alpha/2} + z_{1-\beta}\right]^2 (\sigma_1^2 + \sigma_2^2)}{D^2}$

Some formulas to calculate sample size

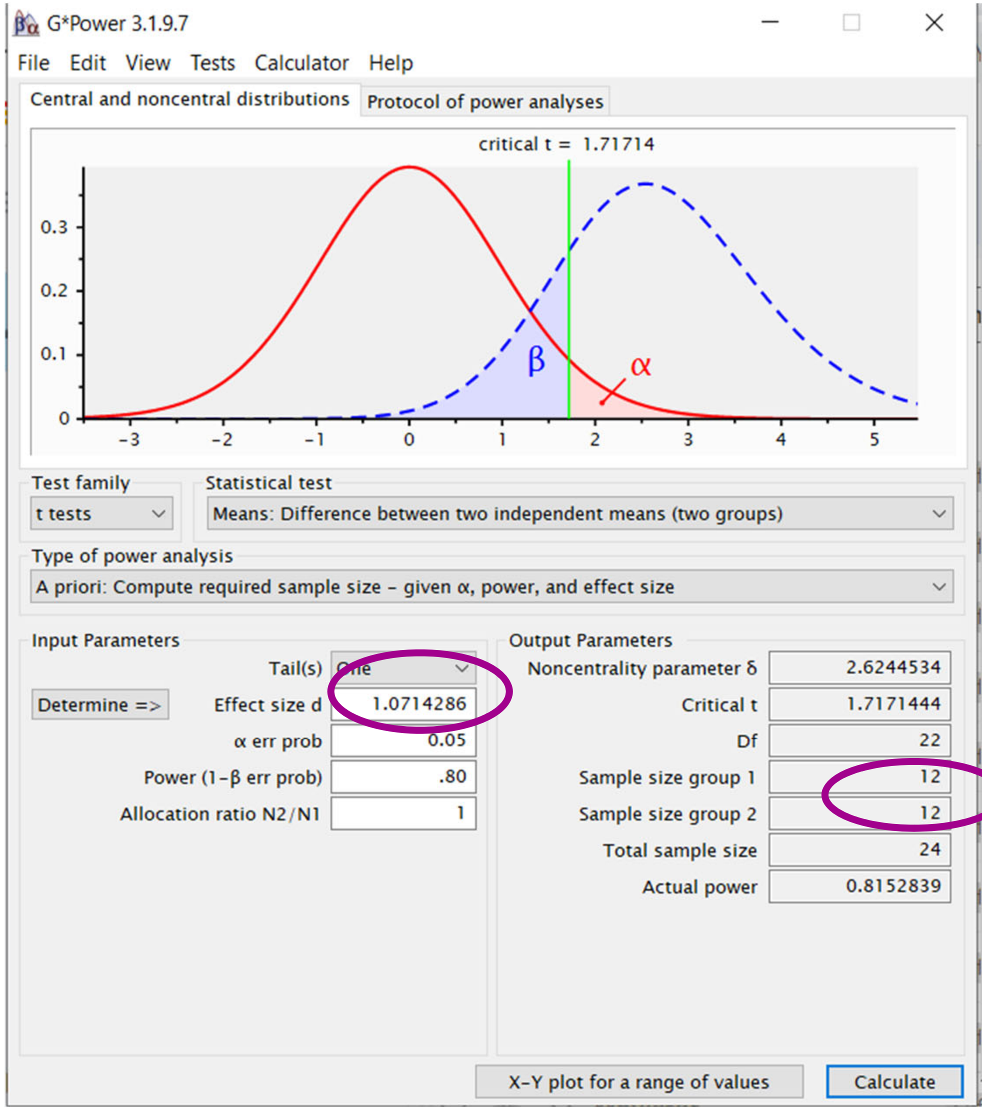
Problema	Datos necesarios	Tamaño muestral
Estimación de una OR	OR= Odds ratio que se estima p_2= Proporción de expuestos en los controles p_1= Proporción de expuestos en los casos e=amplitud relativa del C.I. $z_{1-\alpha/2}$= percentil N(0,1) para significación α	$p_1 = \frac{ORp_2}{ORp_2 + (1-p_2)}$ $n = z^2_{1-\alpha/2} \frac{1/p_1(1-p_1) + 1/p_2(1-p_2)}{\ln(1-e)^2}$
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	$p_1 = RRp_2$ $n = \frac{\left[z_{1-\alpha/2}\sqrt{2p_2(1-p_2)} + z_{1-\beta}\sqrt{p_1(1-p_1)+p_2(1-p_2)}\right]^2}{(p_1-p_2)^2}$

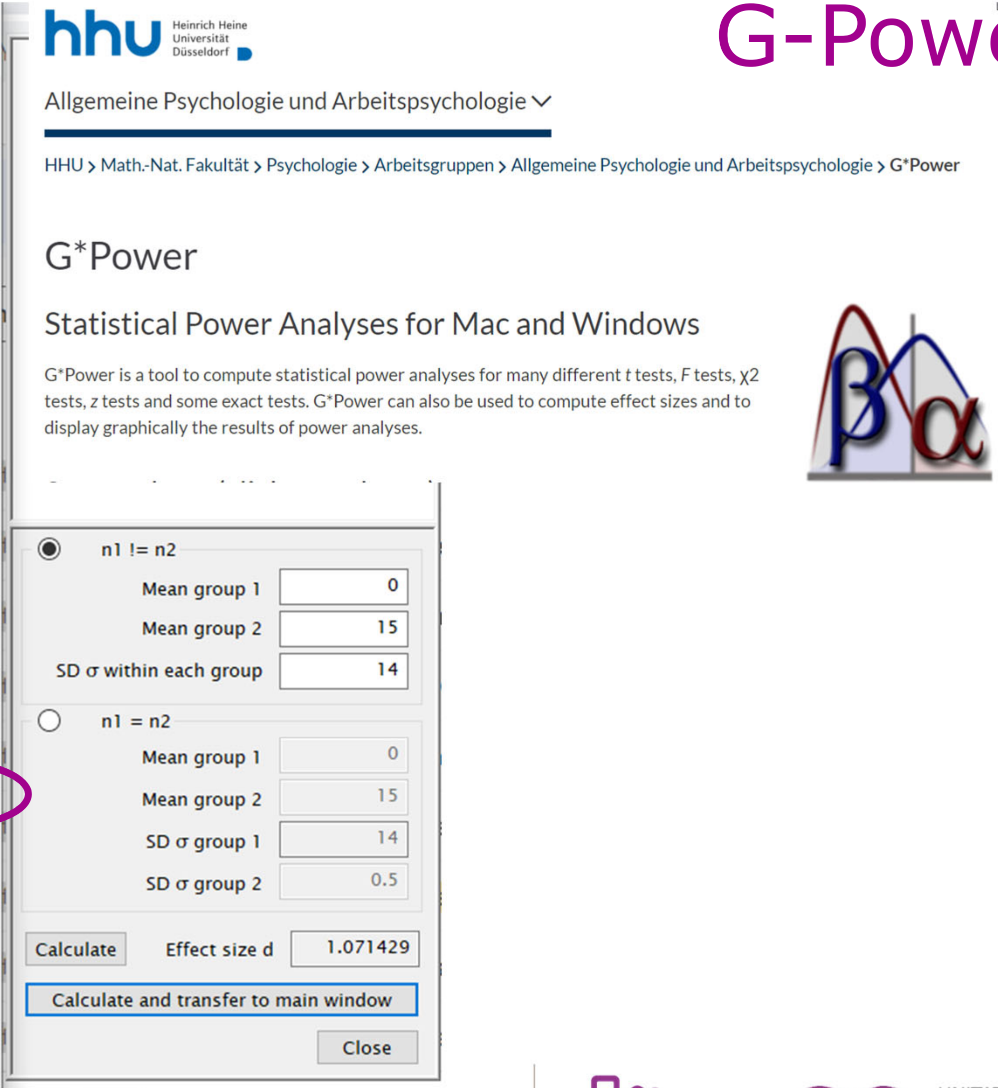
Some tips

- 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

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

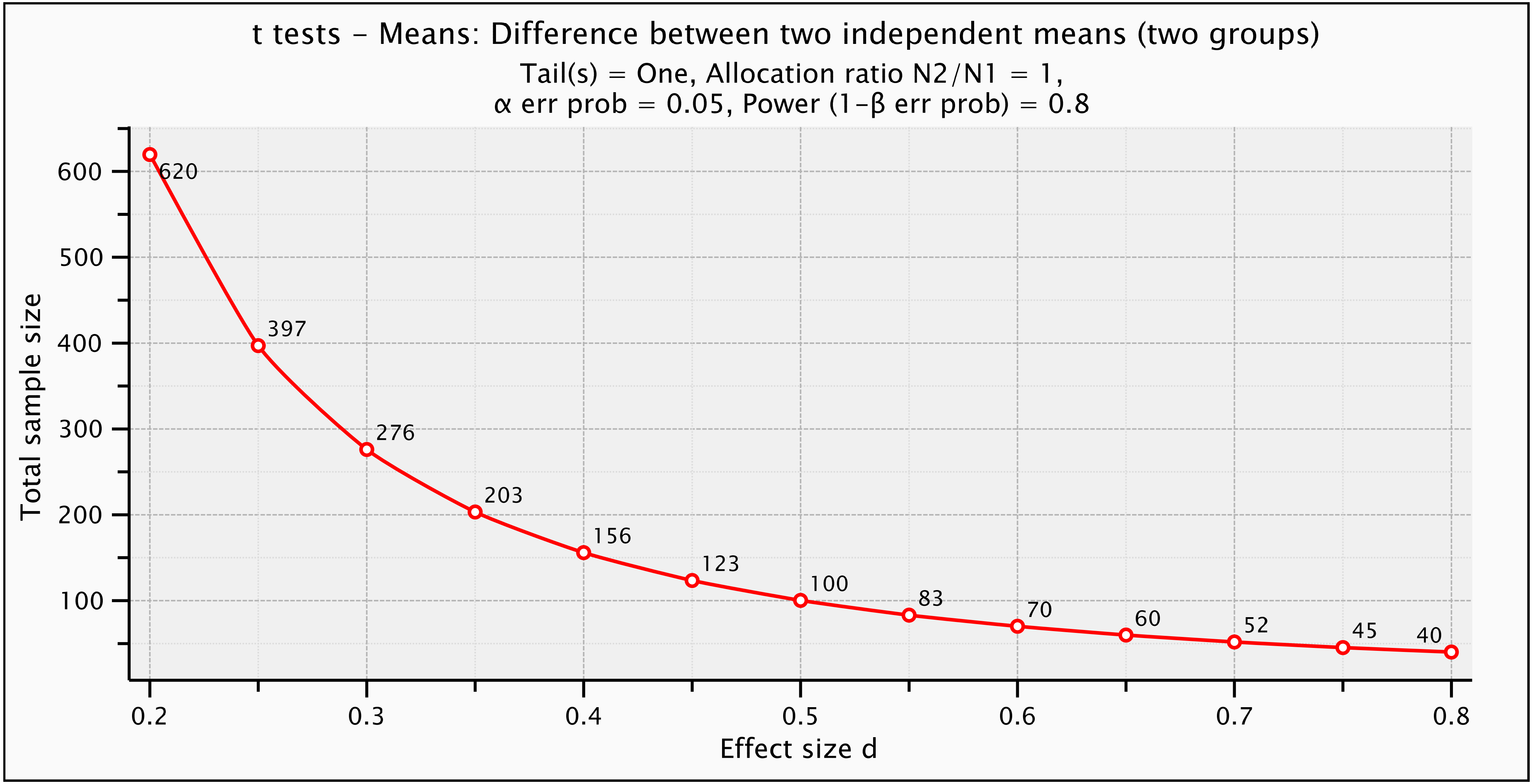
G-Power



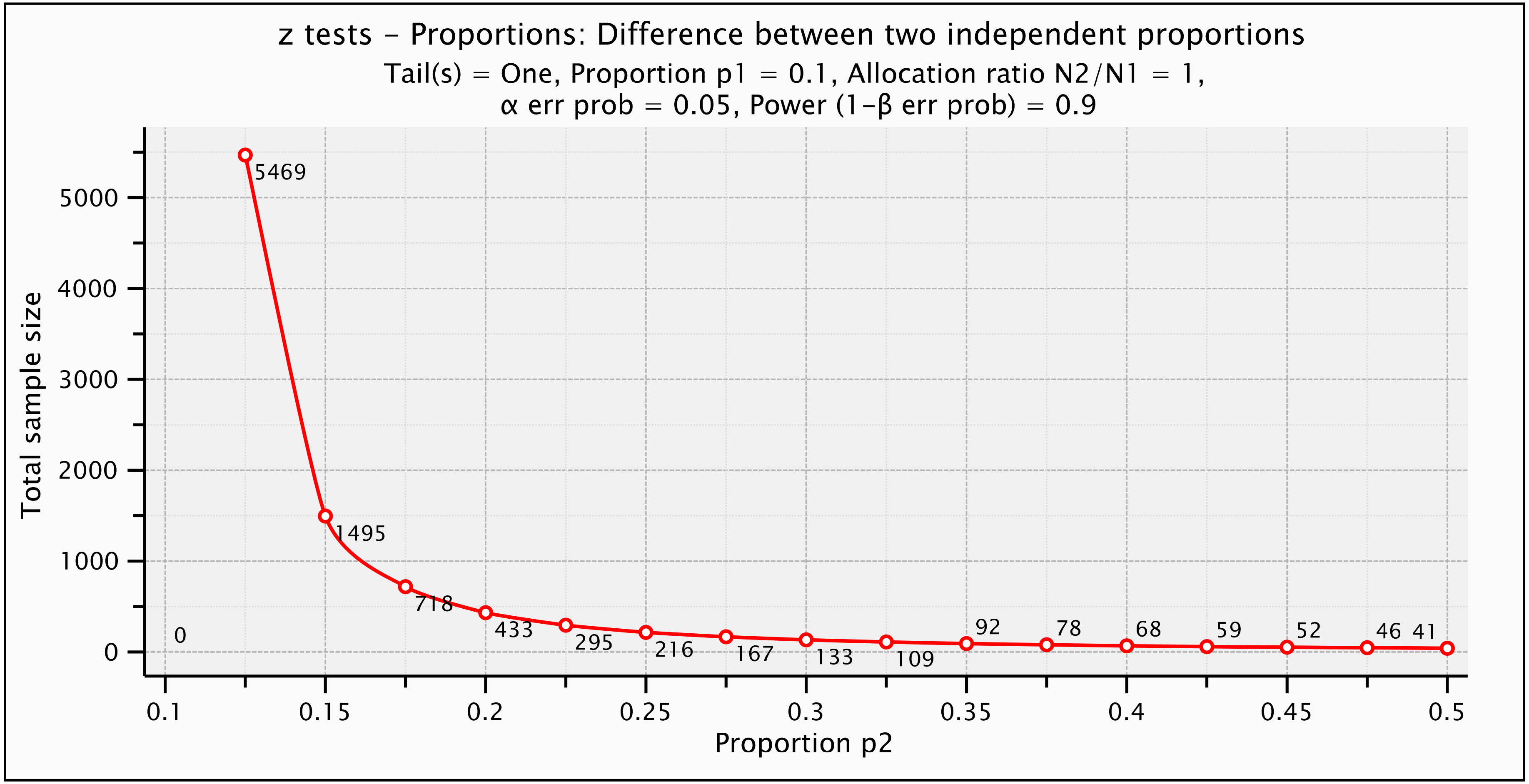


Vall d'Hebron Research

UNITAT D'ESTADÍSTICA I BIOINFORMÀTICA



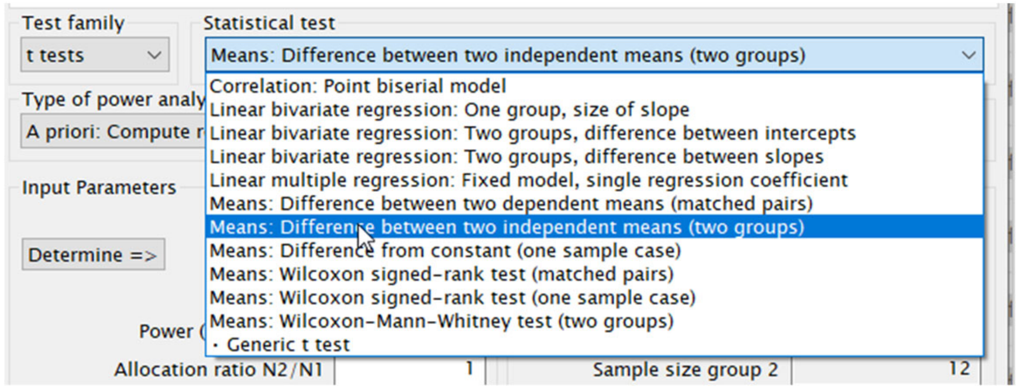
UNITAT D'ESTADÍSTICA I BIOINFORMÀTICA



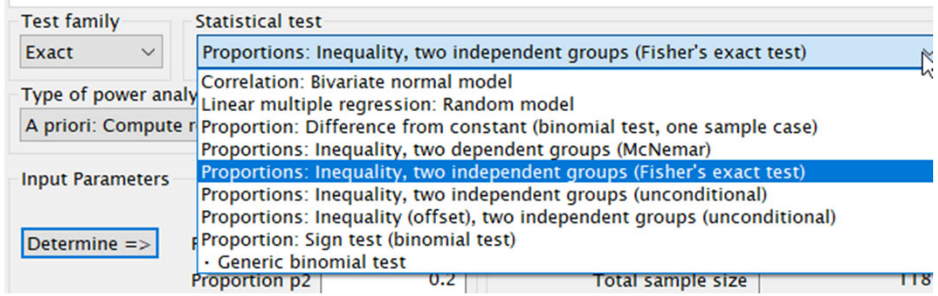
UNITAT D'ESTADÍSTICA I BIOINFORMÀTICA

Use G Power

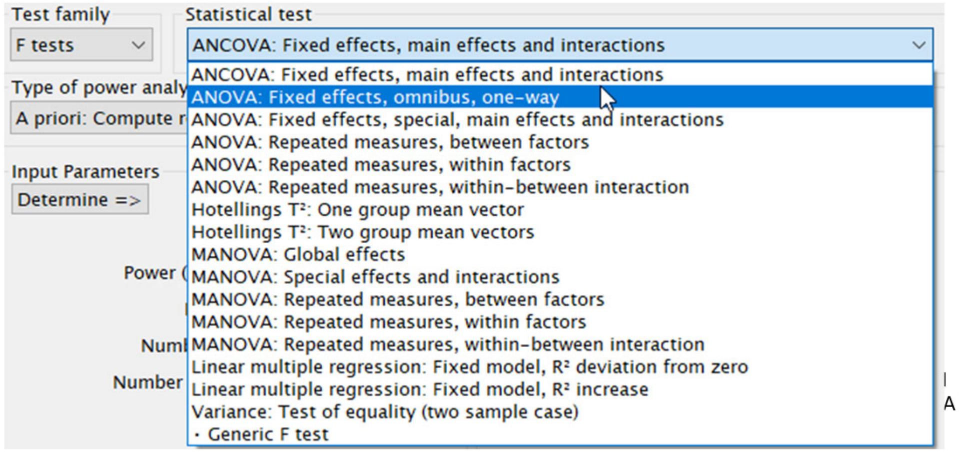
To compare 2 means (t-test)



To compare 2 proportions (χ2-test. Exact)

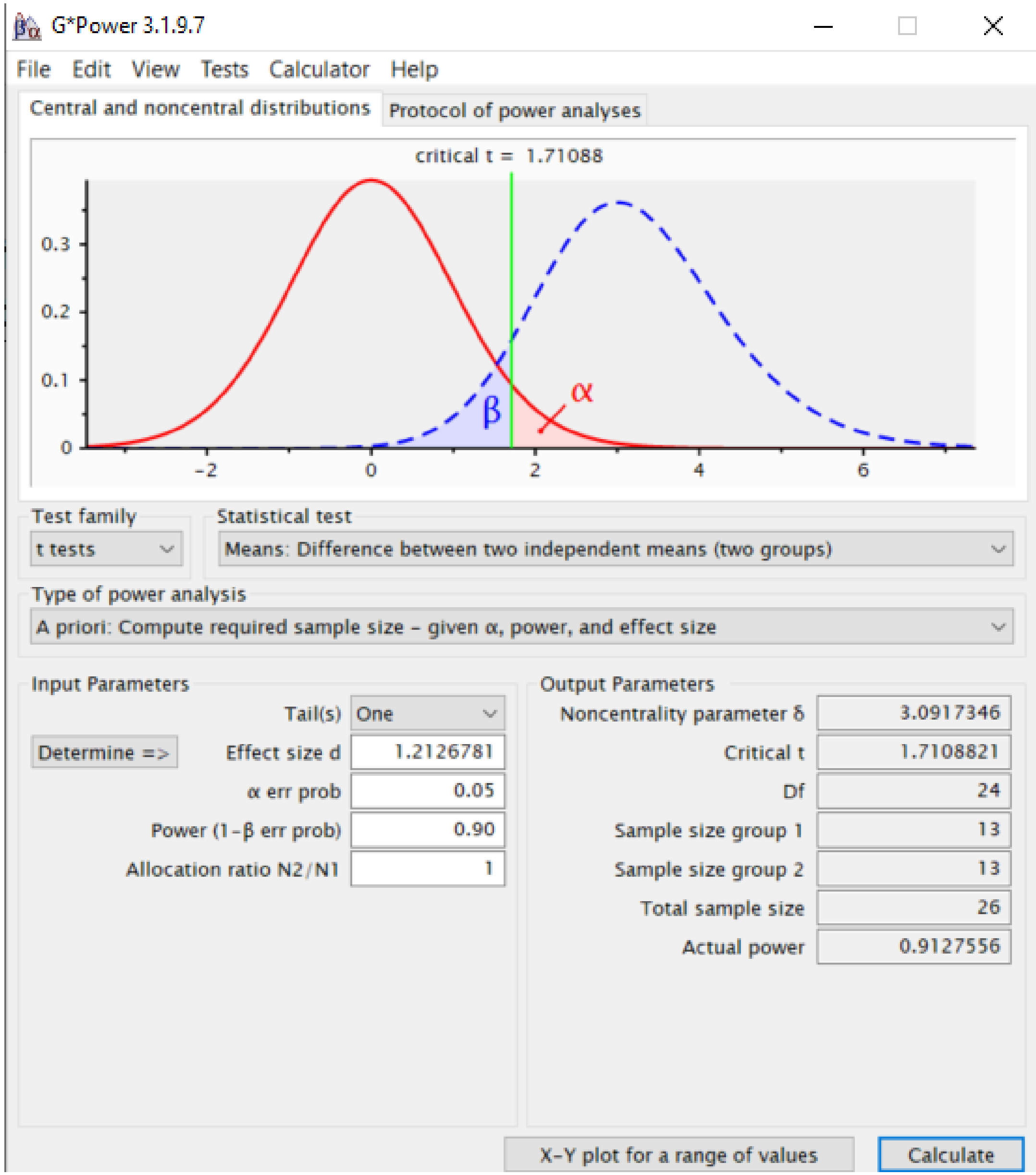


To compare 3 means (F-tests Anova)

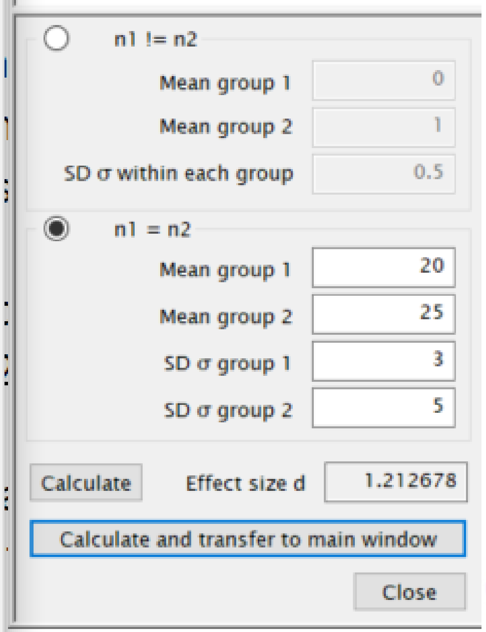


To compare two means

- Primary interest compare Group 1 vs Group 2
- Needed
  - Mean Group 1   (20)
  - Mean Group 2   (25)
  - Or mínimum difference wanted to detect
  - Standard deviation group1 /group2 3/5
  - Type 1 Error (usually 0.05)
  - Power (0.80 or 0.90)



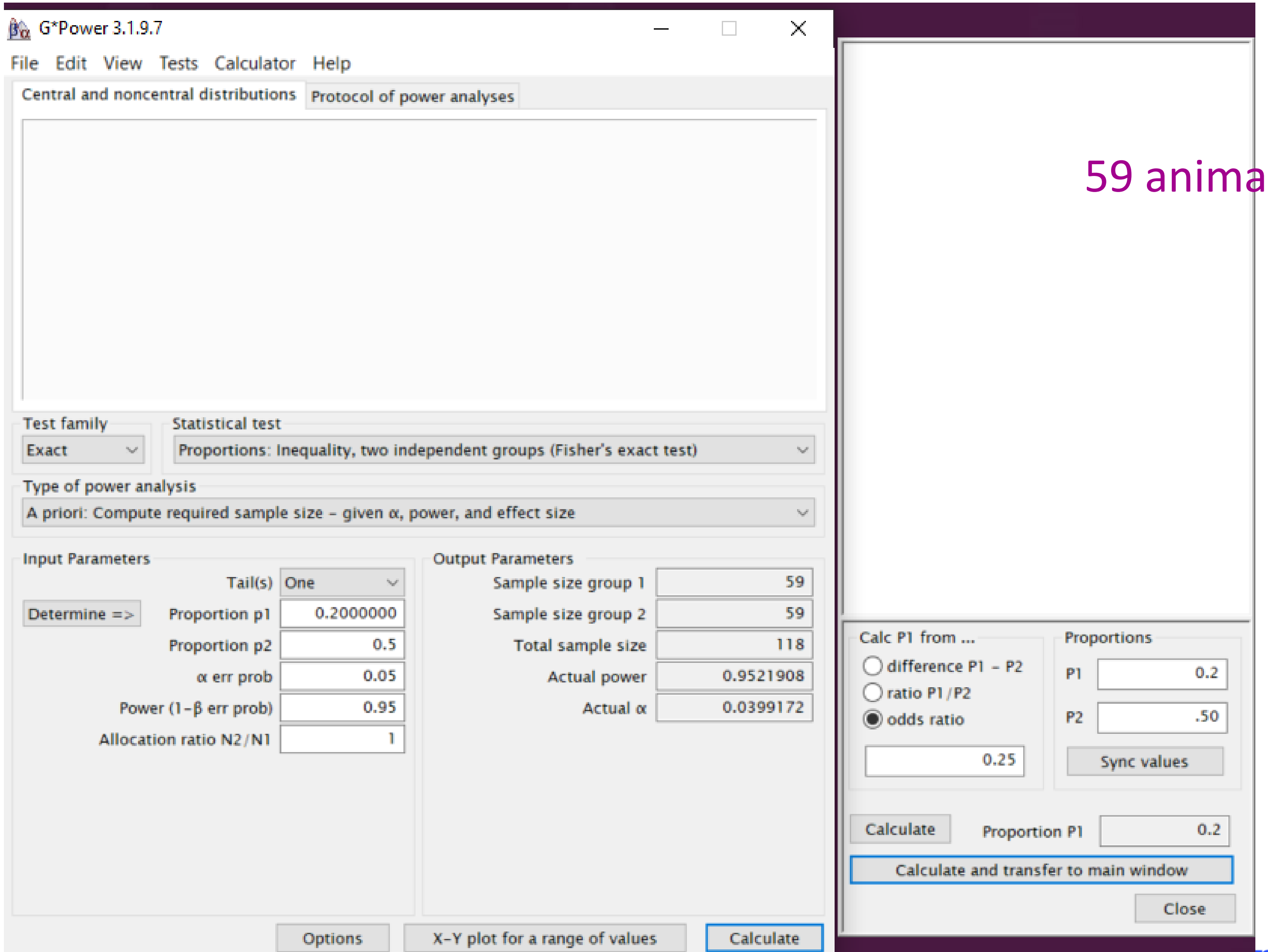
13 animals per group=26



UNITAT D'ESTADÍSTICA I BIOINFORMÀTICA

To compare two proportions

- Primary interest compare Group 1 vs Group 2
- Needed
  - Proportion Group 1 (20%)
  - Proportion Group 2 (50%)
  - Type 1 Error (usually 0.05)
  - Power (0.80 or 0.90)

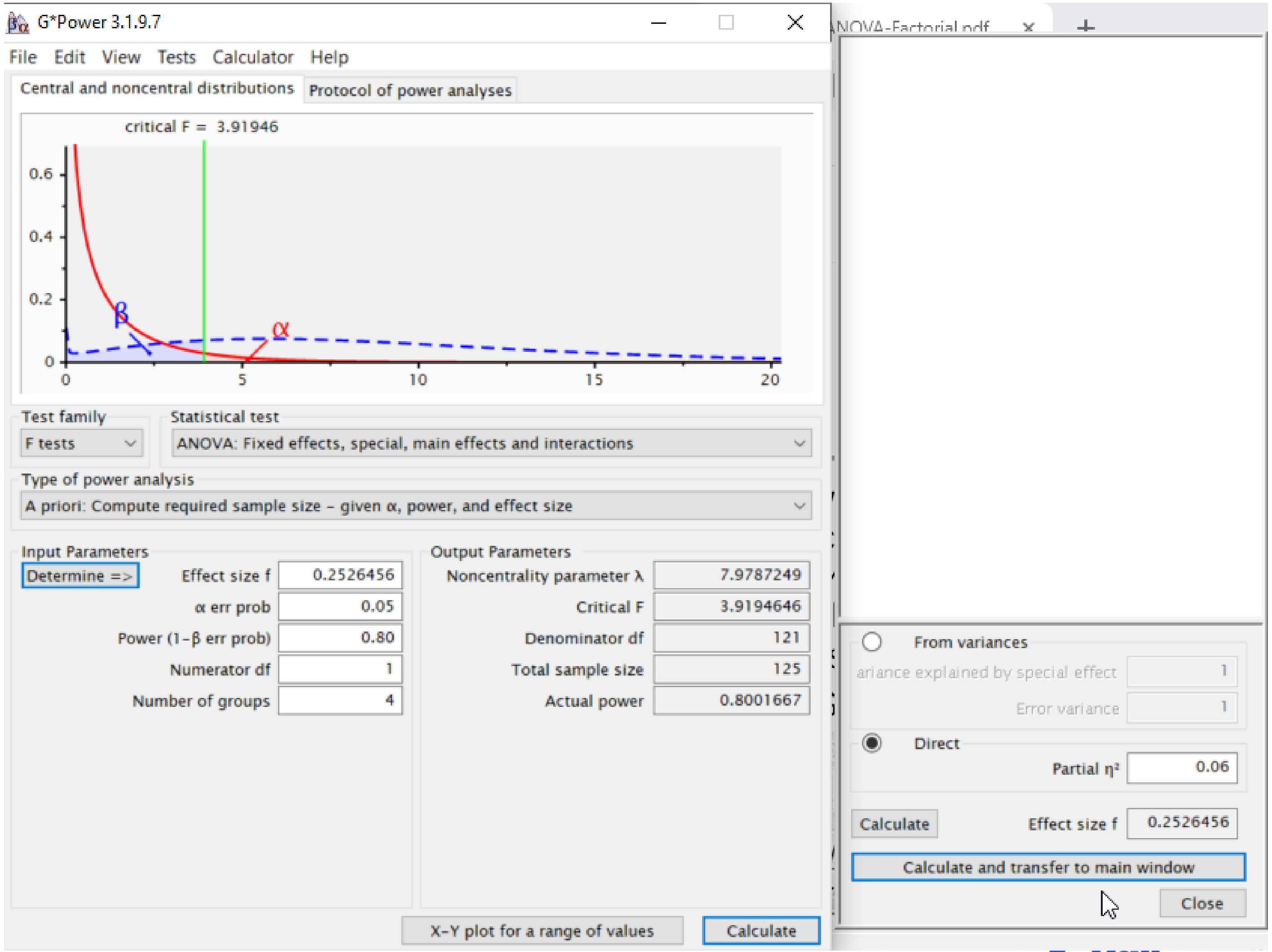


59 animals per group=118

To compare more than two means

- Option 1
  - Stablish number of comparisons (g1 vs g2 , g1 vs g3,...)
  - Calculate sample size for two means for each comparison with type 1 error= 0.05/n comparisons (Bonferroni)
  - Select the màximum number for group
- Option 2
  - Based on effect size η , number of groups of the factorial design.

Test	Relevant effect size	Effect Size (ES)		
		Small	Medium	Large
ANOVA (for large sample)	Eta Square η^2	0.01	0.06	0.14
ANOVA (for small size)	Omega square Ω^2			



Treatment (2 levels)
Sex (2 levels)

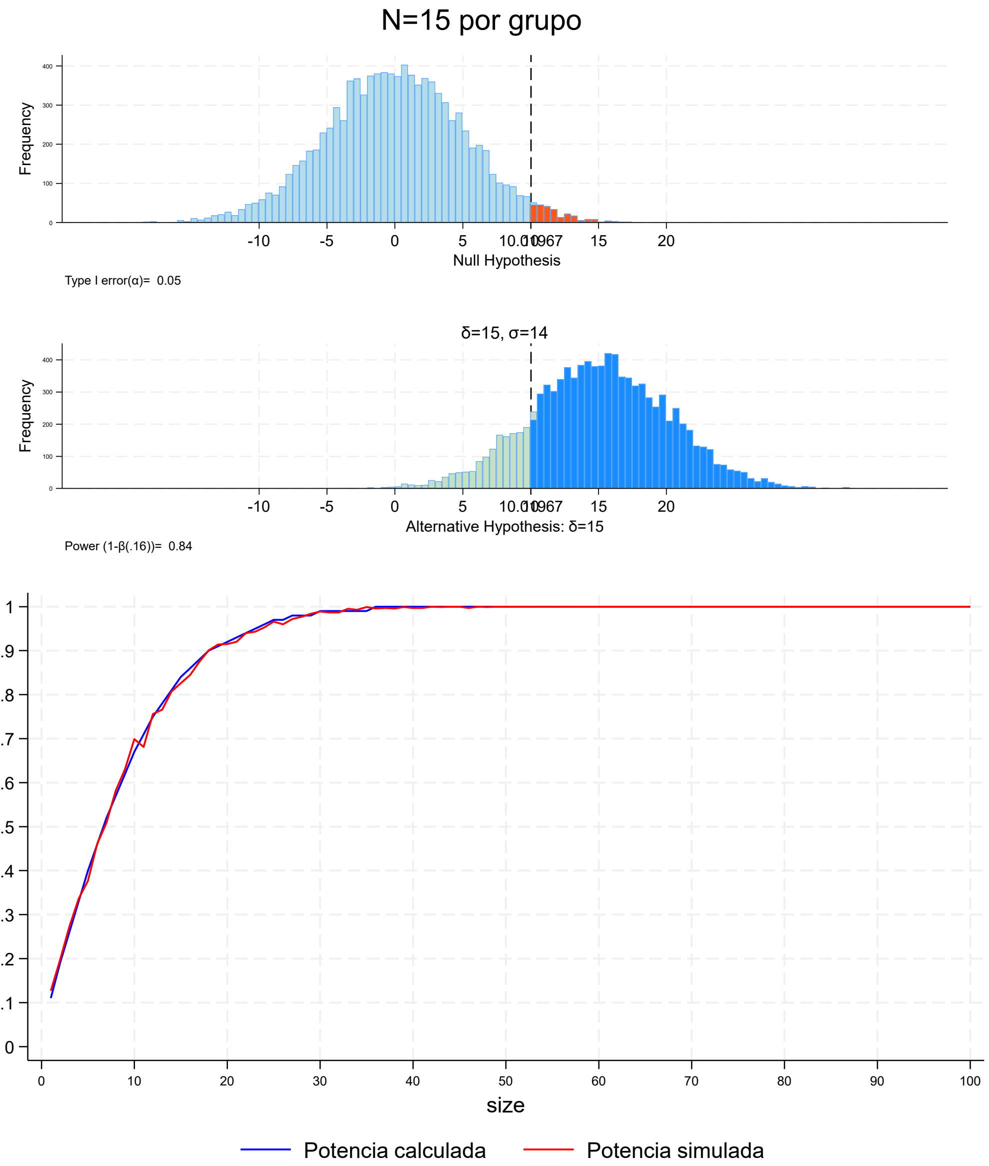
Df of treatment (2-1)

Number of groups= 4

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
- 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=15
   σ=14
2. Select series of sample size n=1,2,3 ...., 10,20...,100
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

Some considerations about sex and replications

- Any experiment can be represesented as follows:

Outcome = Treatment + Error

Degrees of freedom	(n-1)	(2-1)	(n-2)

- If sex is not considered and all animals are male, to answer the questions relative to effects on female another experiment with the same n should be done.

- But if a factorial design is performed double sample size is not needed

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

- When a factor is included is not need to double the sample size, only a slight increase should be done to account for the extra parameter.
- A factorial analysis should be done in that case

Outcome = Treatment + Sex + Treatment*Sex + Error

Degrees of freedom	(n-1)	(2-1)	(2-1)	(2-1)*(2-1)	(n-4)

	Treatment (n=10)	Control (n=10)
Female (n=10)	n=5	n=5
Male (n=10)	n=5	n=5

- Only **4** additional animals would be required

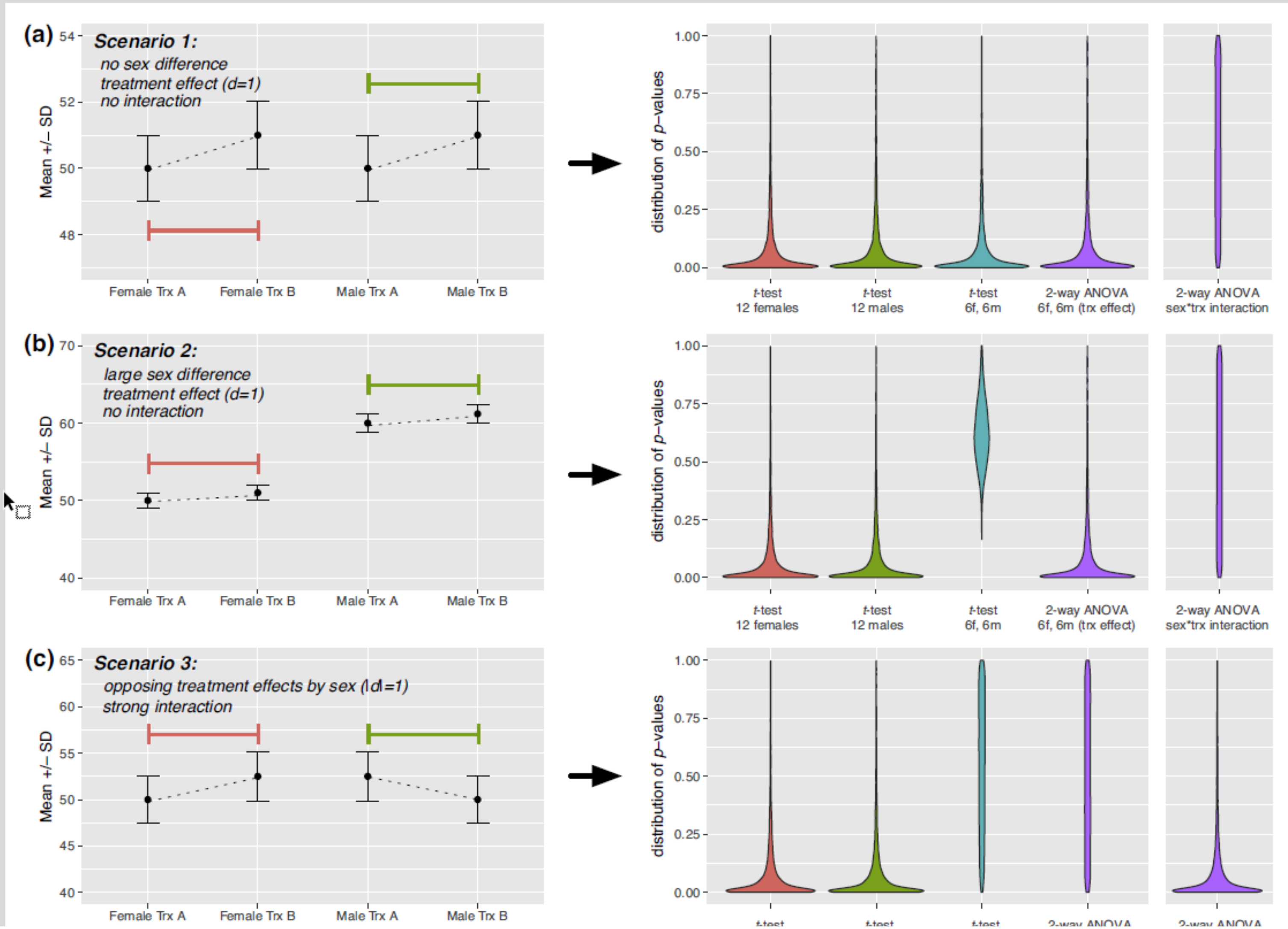
Factorial Designs

- Can test for sex differences using the same number of animals with a little loss of power
- Can answer three questions at once
  - Does the outcome differ between treated and untreated?
  - Does the outcome differ between males and females?
  - Does the treatment have the same effect on males and females?
- To test this effects a two-way anova should be used instead of t-test

	Treatment (n=10)	Control (n=10)
Female (n=10)	n=5	n=5
Male (n=10)	n=5	n=5

SCIENCE FACT OR SCIENCE FICTION: IF I INCLUDE FEMALE ANIMALS, DO I NEED TO DOUBLE MY SAMPLE SIZE? Issue 13\November 2018. Canadian institutes of Health Research. . Institut of Gender and Health. https://cihr-irsc.gc.ca/e/documents/igh_mythbuster_november_2018_en.pdf

Figure 2



Beery AK. Inclusion of females does not increase variability in rodent research studies. Curr Opin Behav Sci. 2018 Oct;23:143-149. doi: 10.1016/j.cobeha.2018.06.016. Epub 2018 Aug 2. PMID: 30560152; PMCID: PMC6294461.

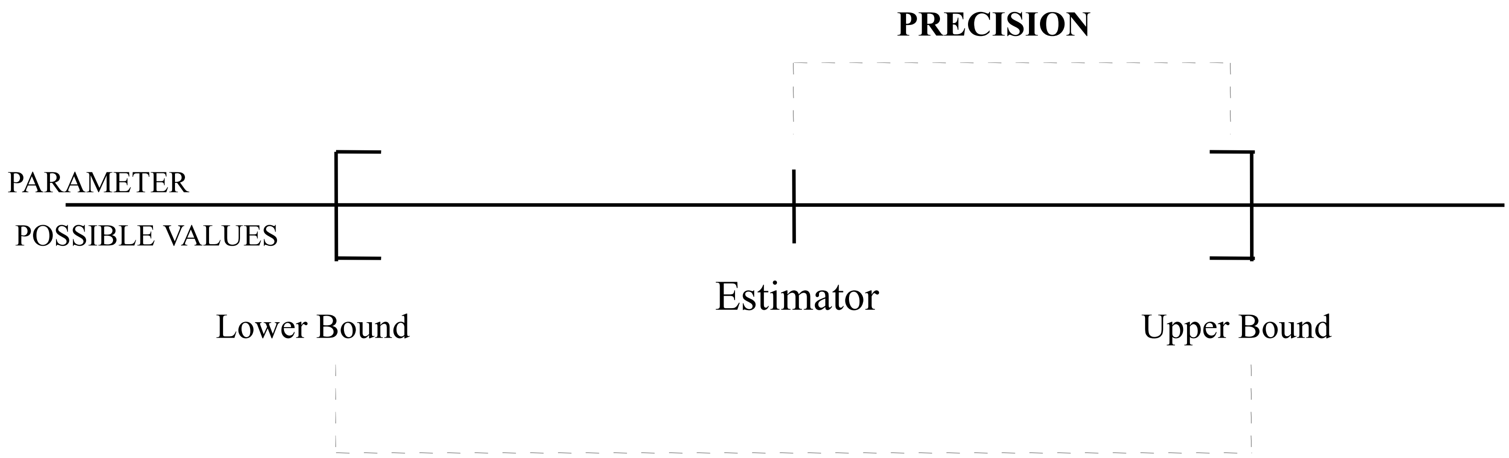
Summary

- Important step in designing a study
- Although for simple situations 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 an experiment , is recommended to take caution and ask for statistical advice
- Incorporating sex or other factors using a factorial design do not increase sample size significantively

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

Thank you / Gràcies/Gracias

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%)

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

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

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

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)

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

  - Population variance unkown, estimated by sample variance

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

- Data: Counts of presence or absence of an event
  - Sample must be “big enough”

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

- $Z_{1-\alpha/2}$ 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

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

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

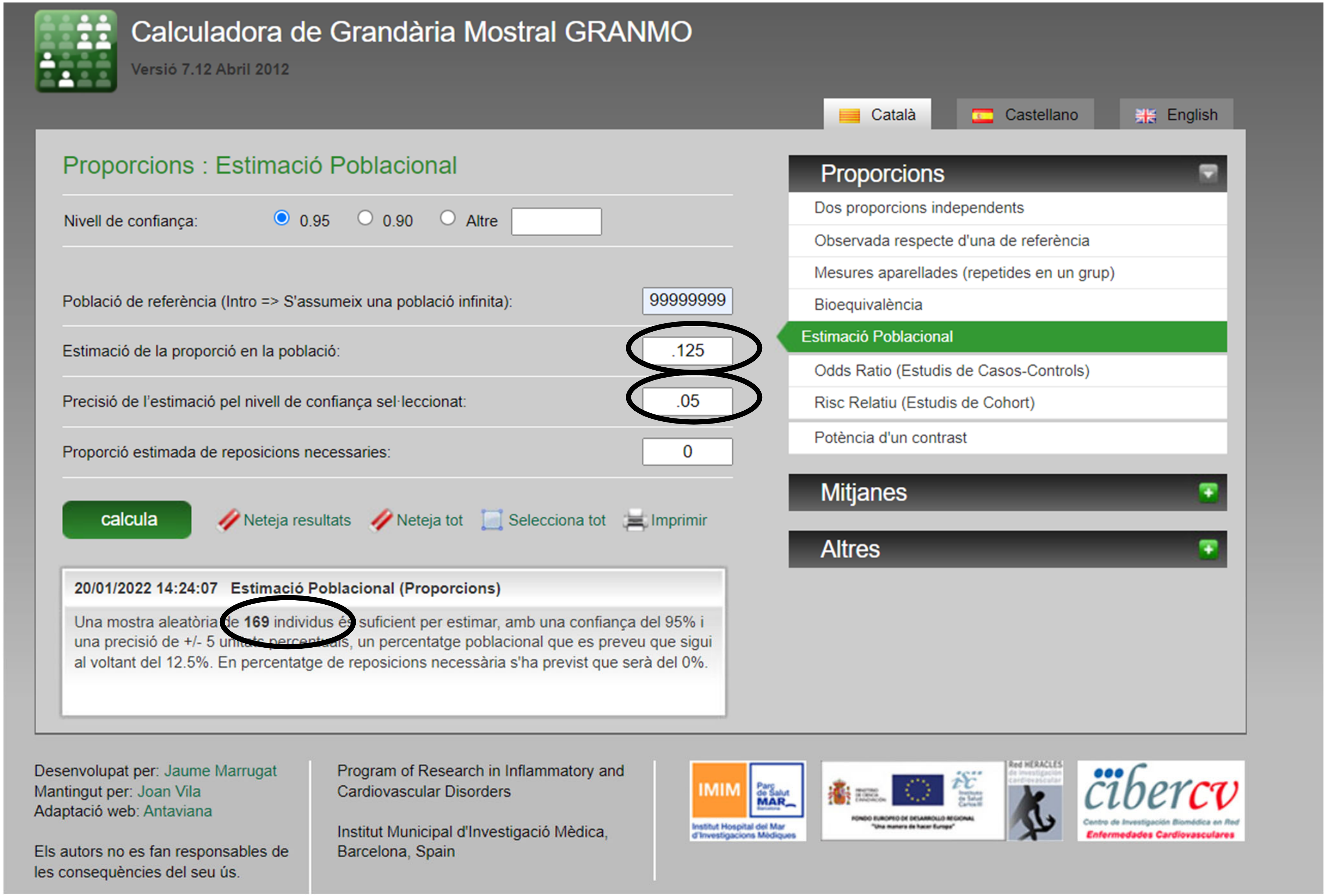
$$n = \frac{1.96^2 \hat{p}(1-\hat{p})}{precision^2} =$$

- 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%

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


Calculadora de Grandària Mostral GRANMO
Versió 7.12 Abril 2012
Català
Castellano
English
Proporcions : Estimació Poblacional
Nivell de confiança: 0.95 0.90 Altre
Població de referència (Intro => S'assumeix una població infinita): 99999999
Estimació de la proporció en la població: .125
Precisió de l'estimació pel nivell de confiança sel·leccionat: .05
Proporció estimada de reposicions necessaries: 0
calcula
Neteja resultats
Neteja tot
Selecciona tot
Imprimir
20/01/2022 14:24:07 Estimació Poblacional (Proporcions)
Una mostra aleatòria de 169 individus és suficient per estimar, amb una confiança del 95% i una precisió de +/- 5 unitats percentuals, un percentatge poblacional que es preveu que sigui al voltant del 12.5%. En percentatge de reposicions necessària s'ha previst que serà del 0%.
Proporcions
Dos proporcions independents
Observada respecte d'una de referència
Mesures aparellades (repetides en un grup)
Bioequivalència
Estimació Poblacional
Odds Ratio (Estudis de Casos-Controls)
Risc Relatiu (Estudis de Cohort)
Potència d'un contrast
Mitjanes
Altres
Desenvolupat per: Jaume Marrugat
Mantingut per: Joan Vila
Adaptació web: Antaviana
Els autors no es fan responsables de les conseqüències del seu ús.
Program of Research in Inflammatory and Cardiovascular Disorders
Institut Municipal d'Investigació Mèdica, Barcelona, Spain
IMIM
Institut Hospital del Mar d'Investigacions Mèdiques
Parc de Salut MAR
FONDO EUROPEO DE DESARROLLO REGIONAL
"Una manera de hacer Europa"
cibercv
Centro de Investigación Biomédica en Red
Enfermedades Cardiovasculares