POWER ANALYSIS

Bioinformatics Awareness Days @ TIGEM
July 10th, 2023

Eugenio Del Prete, M. Eng., Ph.D.
BIOINFORMATICS CORE
e.delprete@tigem.it

Bioinformatics Core: Tasks

- **STATISTICAL DATA ANALYSIS**
  Experimental Design, Hypothesis Testing, Power Analysis Differential Expression Analysis, Cluster Analysis, Time Series Data Analysis, Survival Analysis, Correlation Analysis
- **OMICS**
  Microarray Analysis, Gene Networks, Pathway Analysis, TFBS Identification, Gene Annotation, Integration, Protein Analysis, Drug Networks
- **NEXT GENERATION SEQUENCING**
  Whole Exome, Targeted Gene, RNA, miRNA, ChIP, Visualization, Interpretation
- **DATABASE AND SOFTWARE**
  DB Creation, DB Maintenance, Web Sites Creation, Web Service Support
- **BIOINFORMATICS AND (BIO)STATISTICS TRAINING**

Bioinformatics Core: People

DIEGO DI BERNARDO

https://www.tigem.it/research/facilities/core-facilities/bioinformatics

https://bioinformatics.tigem.it/

DIEGO CARRELLA

ROSSELLA DE CEGLI

XAVIER BUJANDA CUNDIN

EUGENIO DEL PRETE

Bioinformatics Core: Something about Me

- **TLC ENGINEER** @ UNIVERSITY OF ROME ‘SAPIENZA’
  MAIN TOPICS: Signal Processing, Remote Sensing, Bioinformatics
  THESIS: miRNA Analysis, Genomic Data Mining, Consensus Analysis, PSSM Creation
- **BIOINFORMATICS RESEARCH FELLOW** @ INSTITUTE OF FOOD SCIENCES (CNR)
  Protein Prediction and Classification, Protein Analysis, Proteomic Mass Spectra Analysis, Sequence Alignment and Phylogenetic Tree, Docking
- **PHD IN APPLIED BIOLOGY** @ UNIVERSITY OF BASILICATA
  Celiac Disease and Comorbities, Microarray Data Analysis, Ontologies, Gene Set Enrichment Analysis, Semantic Similarity, Proteomic Mass Spectra Analysis
- **BIOINFORMATICS RESEARCH FELLOW** @ INSTITUTE OF APPLIED MATHEMATICS (CNR)
  Proteomic Mass Spectra Analysis, Metabolomic (Lipidomic) Data Analysis, Web Tools Developer, Hypothesis Tests, Omics Data Integration
- **BIOSTATISTICIAN AND DATA SCIENTIST** @ TIGEM

Outline

DEFINITIONS
- Sample Size
- Replicates
- Power
- Significance Level
- Effect Size
- Power Analysis

G*POWER
- Tool
- Graphical User Interface

EXAMPLES
- One Way ANOVA
- Repeated Measure ANOVA
- Multi Way ANOVA

CONCLUSION
- Take Home Message
- Final Remarks

Sample Size

Question

- **How many samples do I need for my experiment?**

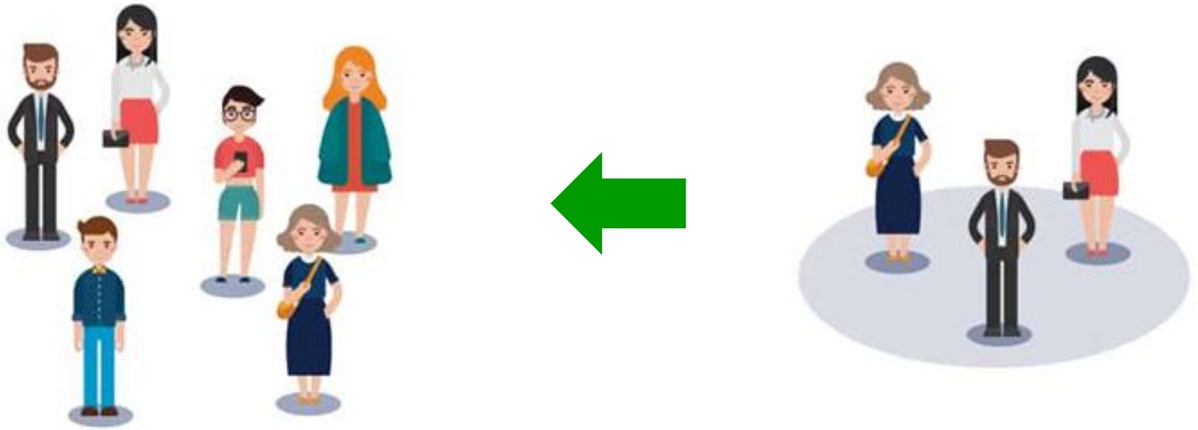
Problems

- **Impossibility** to collect data from an entire population of interest
- **Too small**: under-detection of effect of interest
- **Too large**: unnecessary wasting of resources and/or animals

Goal

- Enough samples to **reasonably detect an effect** (if it really exists) **without wasting limited resources** on too many samples

POPULATION SAMPLE



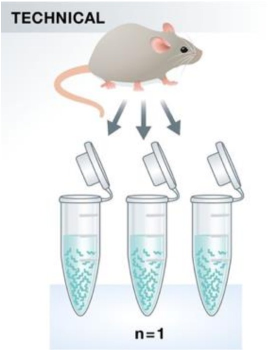
REALITY ≈ MODEL

a) PROBABILITY OF ERROR
b) MARGIN OF ERROR

Replicates

Technical Replicates

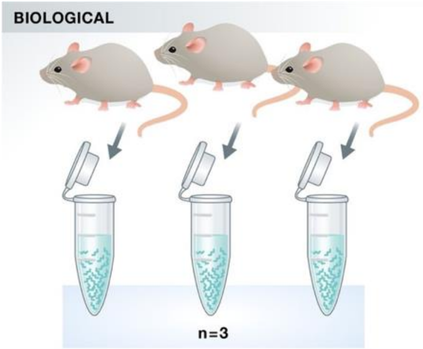
- Use the same biological sample to repeat the technical or experimental steps in order to accurately measure technical variation and remove it during analysis



WITHIN VARIABILITY

Biological Replicates

- Use different biological samples of the same condition to measure the biological variation between samples



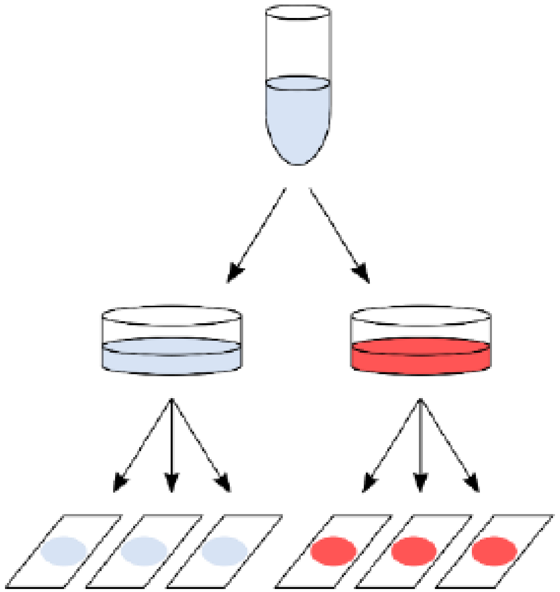
BETWEEN VARIABILITY

WHY THREE (3) IS THE ‘MAGIC’ NUMBER?

Focus on: Experimental Design and Sample Size

Design 1

- Single vial, two culture dishes
- Same time for cellular growth
- Cells counting for the glass slides
- **Which is the sample size?**



ANSWER: N = 1

There is not independence between the three glass slides within the two conditions

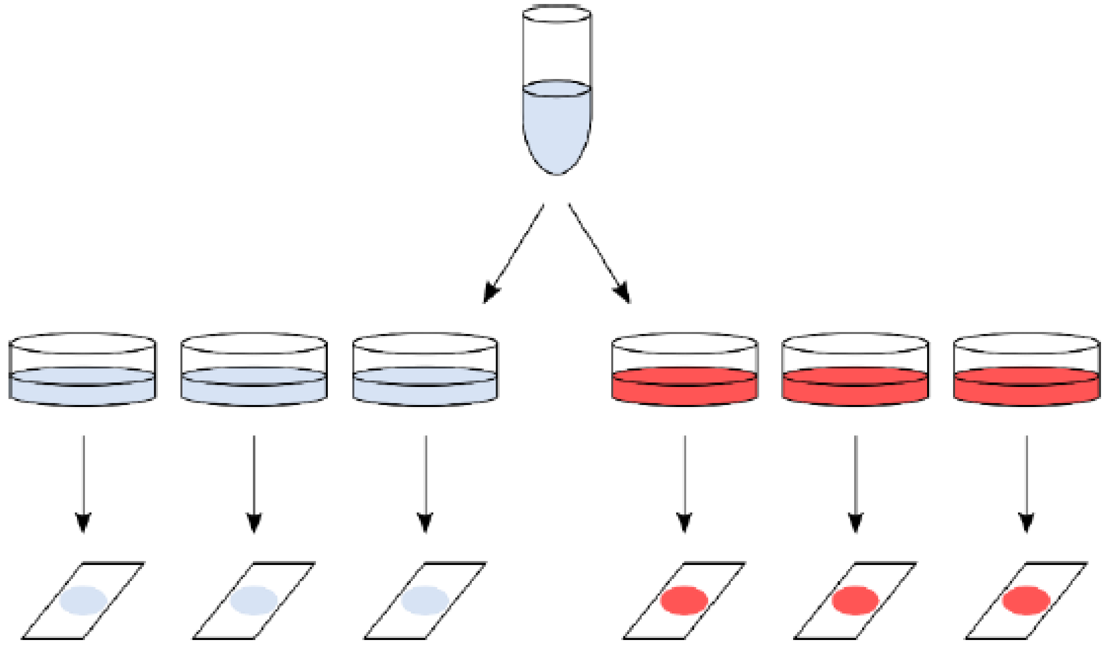
AND

The variability among the glass slides is only due to pipetting error.

Focus on: Experimental Design and Sample Size

Design 2

- Single vial, six culture dishes
- Same time for cellular growth
- Cells counting for the glass slides
- **Which is the sample size?**



ANSWER: N = 1

The glass slides are processed in the same day, same medium, same incubator, and same time.

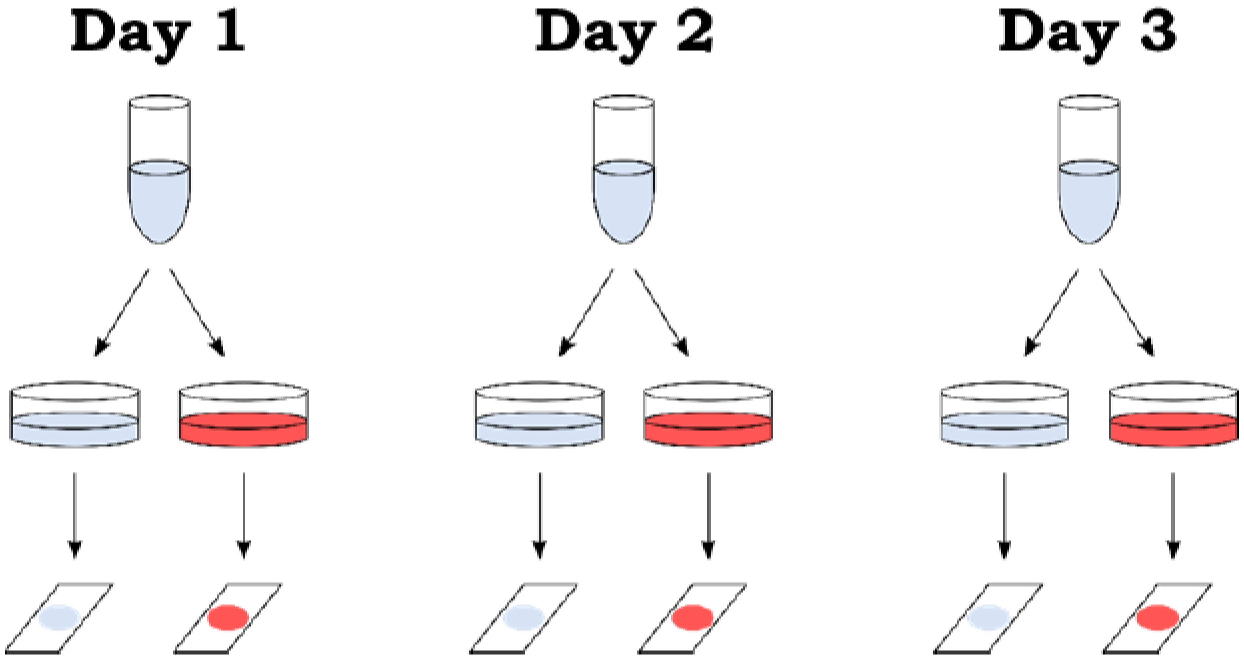
AND

Different plastic does not provide independence.

Focus on: Experimental Design and Sample Size

Design 3

- Three vials, one sample, six culture dishes
- Three time for cellular growth
- Cells counting for the glass slides
- **Which is the sample size?**



ANSWER: N = 3

Three different times introduce the independence, even if the variability is still more technical than biological

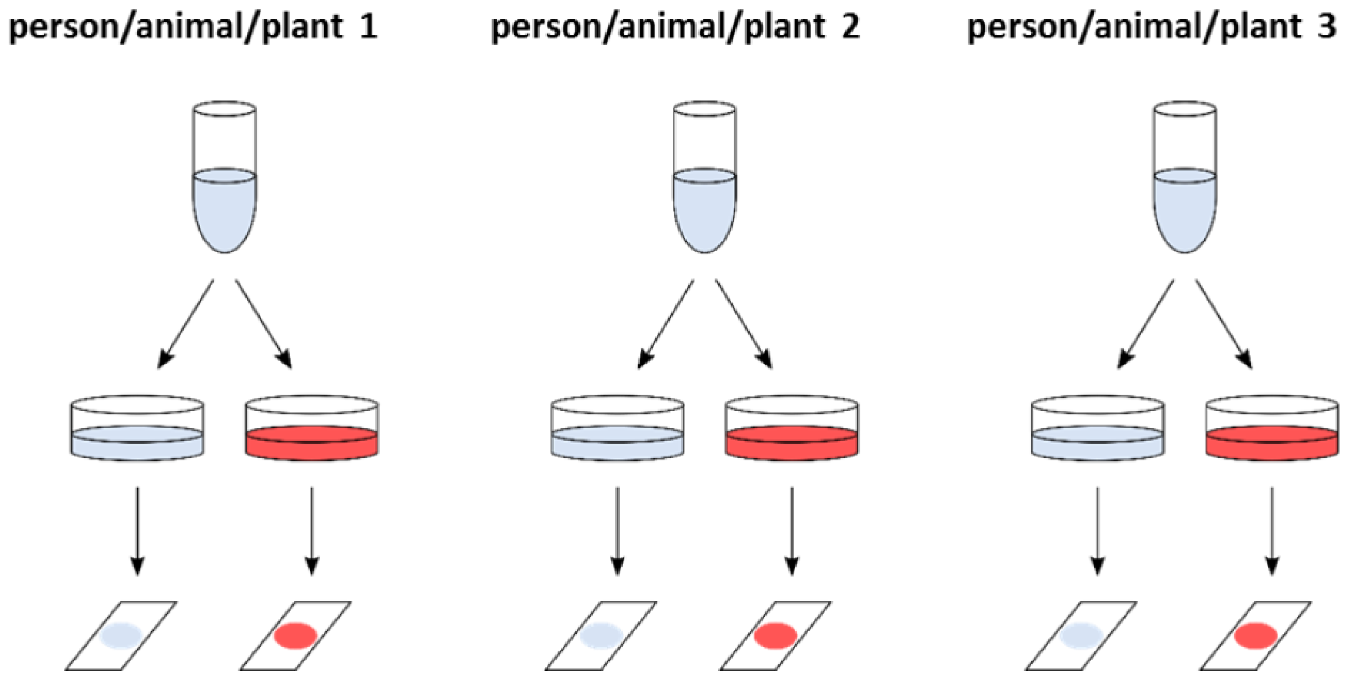
AND

The glass slides for each time are paired

Focus on: Experimental Design and Sample Size

Design 4

- Three vials, three samples, six culture dishes
- Same time for cellular growth
- Cells counting for the glass slides
- **Which is the sample size?**



<u>ANSWER: N = 3</u>

Three independent samples provides biological replicates

AND

The glass slides for each sample are paired

AND

Pay attention to batch effects (e.g., litter)

Power and Significance Level

- H0: LEMUR (NOT SLOTH)

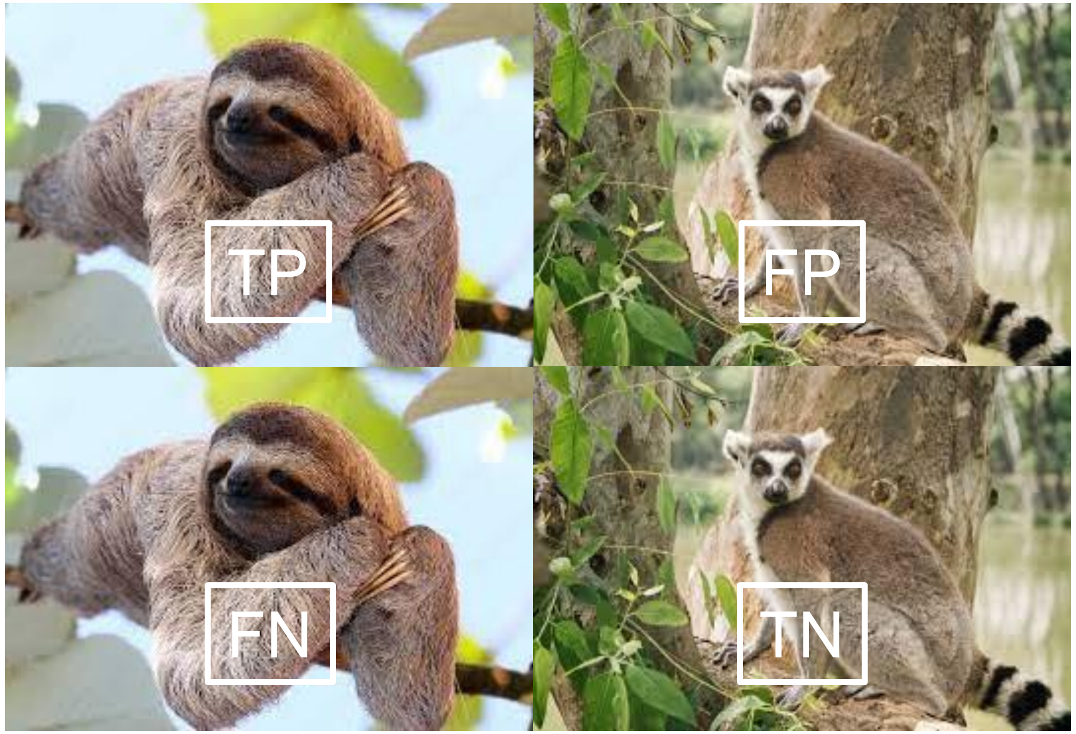
REALITY

H0 FALSE | H0 TRUE

STUDY FINDINGS

REJECT H0

ACCEPT H0



TRUE POSITIVE (TP) - POWER (1-β)
Probability to REJECT H0 when H0 is FALSE

FALSE POSITIVE (FP) - TYPE I ERROR, α
Probability to REJECT H0 when H0 is TRUE

FALSE NEGATIVE (FN) - TYPE II ERROR, β
Probability to ACCEPT H0 when H0 is FALSE

TRUE NEGATIVE (TN)
Probability to ACCEPT H0 when H0 is TRUE

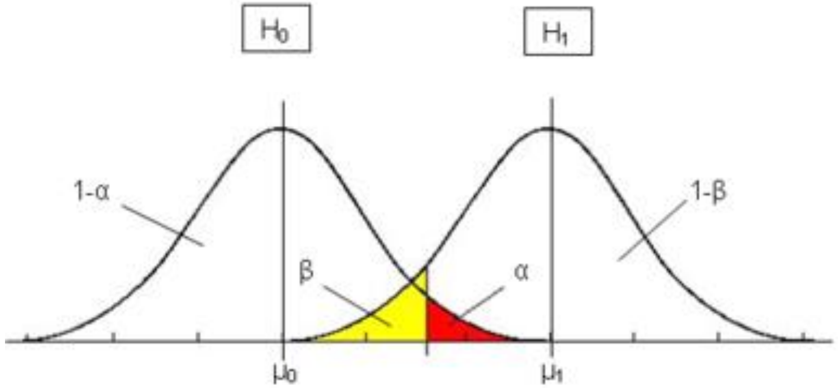
Power and Significance Level

Power (1-β)

- **Probability** to reject a **false** null hypothesis (H0) when the alternative hypothesis (H1) is **true**
- **Probability** of detecting a **true difference** (there is a difference to detect)
- Direct relation with the **Type II Error**

Significance Level (α)

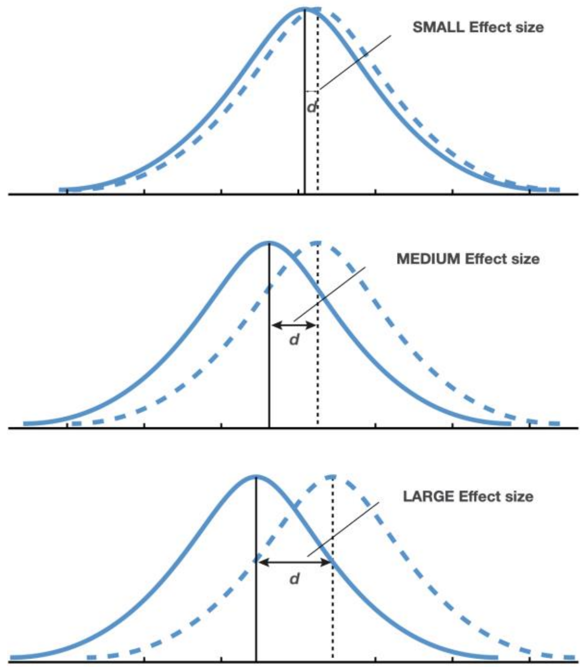
- **Probability** to reject a **false** null hypothesis (H0) when the null hypothesis (H0) is **true**
- **Probability** of detecting a **false difference** (there is not a difference to detect)
- It is the **Type I Error**



Effect Size

Considerations

- **Property** of the sample data (unlike power and significance level)
- To be **estimated**
- **Different statistical** tests have **different values for each category**



Estimation

- Background information from **preliminary data** or **trial data** or similar study
- **Convention** (forced or not)

Formula (Generic)

$$Effect\ Size \approx \frac{Mean\ of\ H1 - Mean\ of\ H0}{Standard\ Deviation}$$

Test	Relevant effect size	Effect Size Threshold		
		Small	Medium	Large
t-test for means	d	0.2	0.5	0.8
F-test for ANOVA	f	0.1	0.25	0.4
t-test for correlation	r	0.1	0.3	0.5
Chi-square	w	0.1	0.3	0.5
2 proportions	h	0.2	0.5	0.8

Power Analysis

Hypothesis
- Biological question
- **One Tail** or **Two Tails** test

One Tail: $\begin{cases} H_0: \mu_0 > \mu_1 \text{ or } H_0: \mu_0 < \mu_1 \\ H_1: \mu_0 < \mu_1 \text{ or } H_1: \mu_0 > \mu_1 \end{cases}$

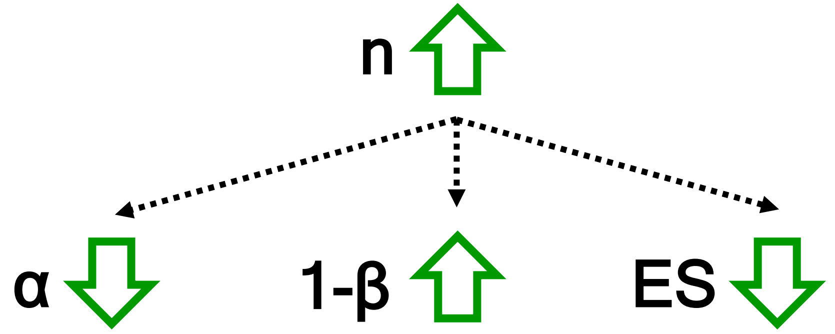
Two Tails: $\begin{cases} H_0: \mu_0 = \mu_1 \\ H_1: \mu_0 \neq \mu_1 \end{cases}$

Experimental Design
- **Variables** (Treatment, Timepoint, Dose, Sex, …)

Statistical Test
- **Comparison** of interest
- **Type of variables**

Parameters
- **Power (1-β)**
- **Significance Level (α)**
- **Effect Size (ES)**
- **Sample Size (n)**

WHICH IS THE LINK WITH n AMONG THE PARAMETERS?



Focus on: Type of Power Analysis

A Priori Power Analysis

- Determine **how many samples** you will need to have a good chance of detecting an effect of a specified size with the desired amount of power

⇩

GIVEN α, 1-β, ES
THEN COMPUTE n

A priori analyses are performed as **part of the research planning process**. They allow you to determine the sample size you need in order to reach a desired level of power.

A Posteriori Power Analysis

- Allow to find **how much power** you would have if you had a specified and sufficient number of samples when you had already conducted a research

⇩

GIVEN α, n, ES
THEN COMPUTE 1-β

Post hoc analyses are performed after your study has been conducted and can be used to assist in **explaining any potential non-significant results**.

G*Power

Heinrich Heine Universität Düsseldorf
Cognitive and Industrial Psychology

Availability

- **WINDOWS:** 3.1.9.7 (March 2020)
- **MAC**: 3.1.9.6 (February 2020)

Statement

- **Free** software for everyone
- **Downloadable** from the webpage
- **Screenshot** without asking permissions

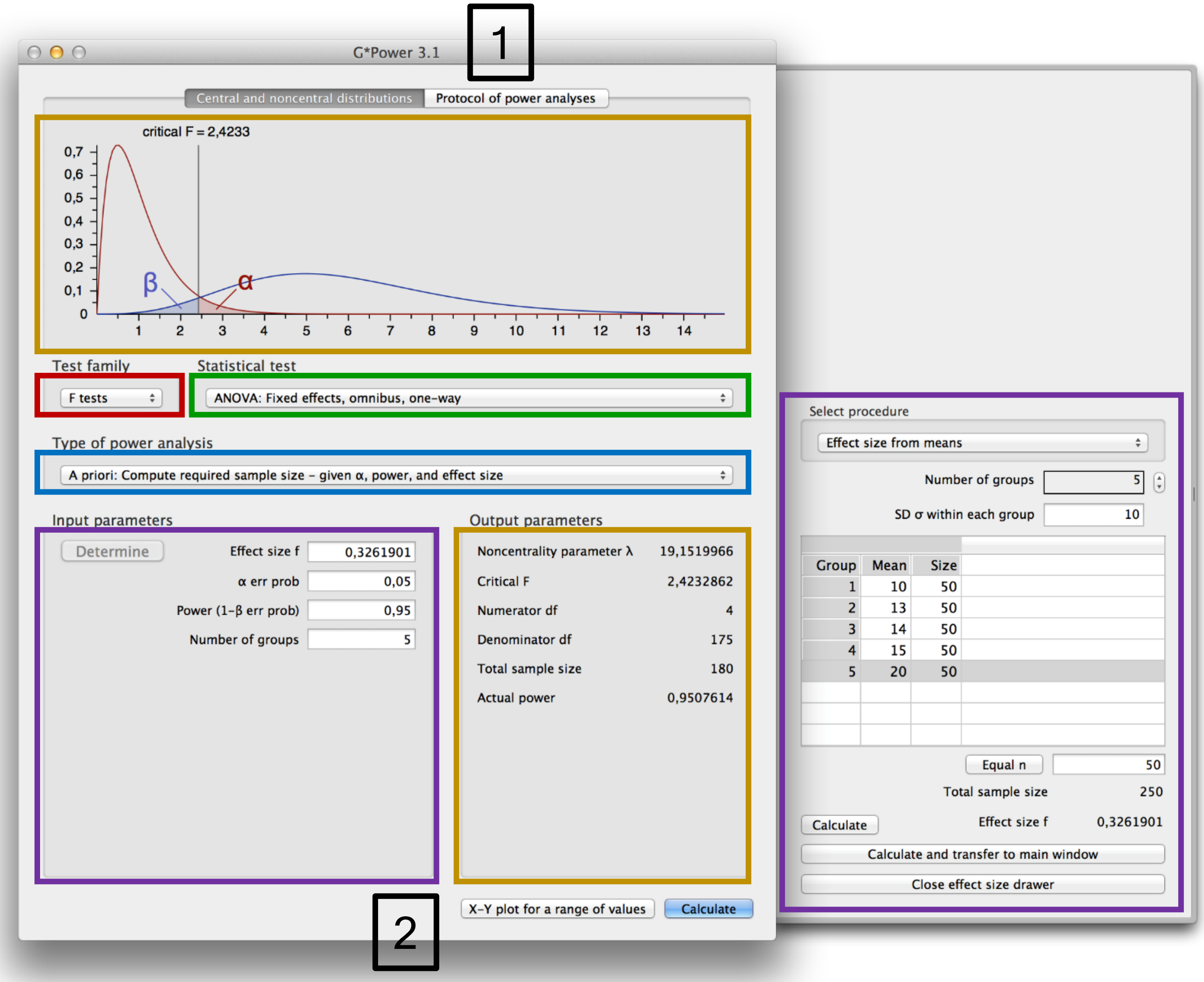
Analysis Type

- Selection for the **parameter of interest**
- Selection for the **test of interest**

Analysis Type	Calculated Value
A priori	Required sample size
Compromise	Implied power and alpha values
Criterion	Required alpha value
Post hoc	Achieved power
Sensitivity	Required effect size

EXACT: 8 specific + 1 general
F-TEST: 16 specific + 1 general
T TEST: 11 specific + 1 general
χ^2 TEST: 2 specific + 1 general
Z TEST: 7 specific + 1 general

G*Power GUI



a) Type of power analysis

b) Test family

c) Statistical test

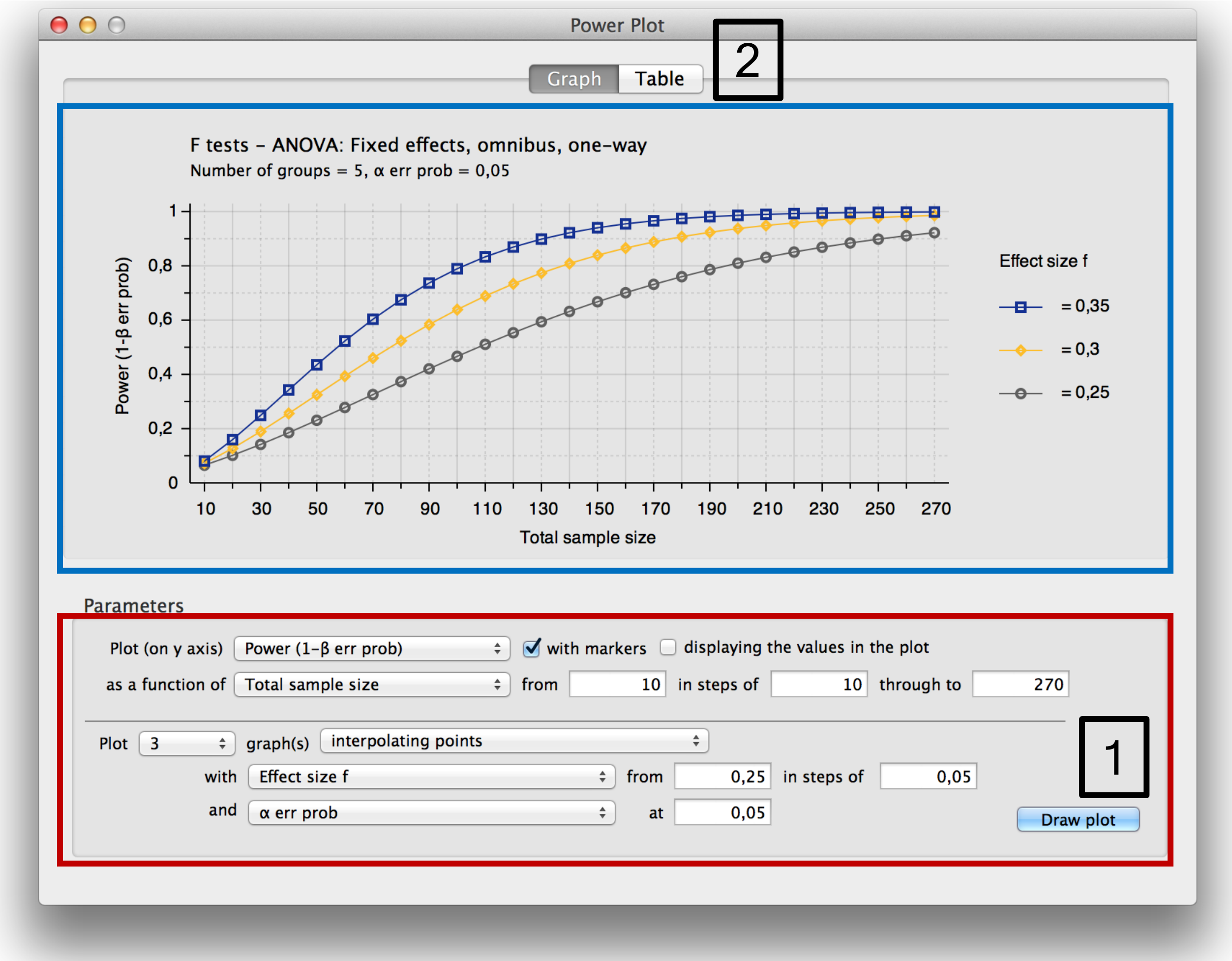
d) Input parameters (with values)

e) Output parameters with density plots

If you click on:

1) **Protocol of power analysis**: Table with I/O parameters
2) **X-Y plot for a range of values:** Plot for parameters

G*Power GUI



a) Parameters and attributes

b) Power Plot

If you click on:

1) **Draw plot**: Power Plot
2) **Table**: X/Y values in tabular format

One Way ANOVA

Description

- Check if **at least one mean** is different among groups
- **Normally distributed (numeric) variables**
- Extension of unpaired t-test for more than two groups

Conditions

- **Numeric variables**: one — MEASUREMENTS
- **Categorical variables**: one — E.G. TREATMENT, TIMEPOINT, DOSE
- **Number of group per categorical variable**: more than two
- **Details**: parametric, unpaired — NORMAL DISTRIBUTION AND DIFFERENT SAMPLES

Parameters

- Significance level: 0.05
- Power: 0.80

One Way ANOVA

Example

- Five different medication regimes on patients after surgery
- Differences in white cell counts
- H0: no difference among medication regimes; H1: difference among medication regimes

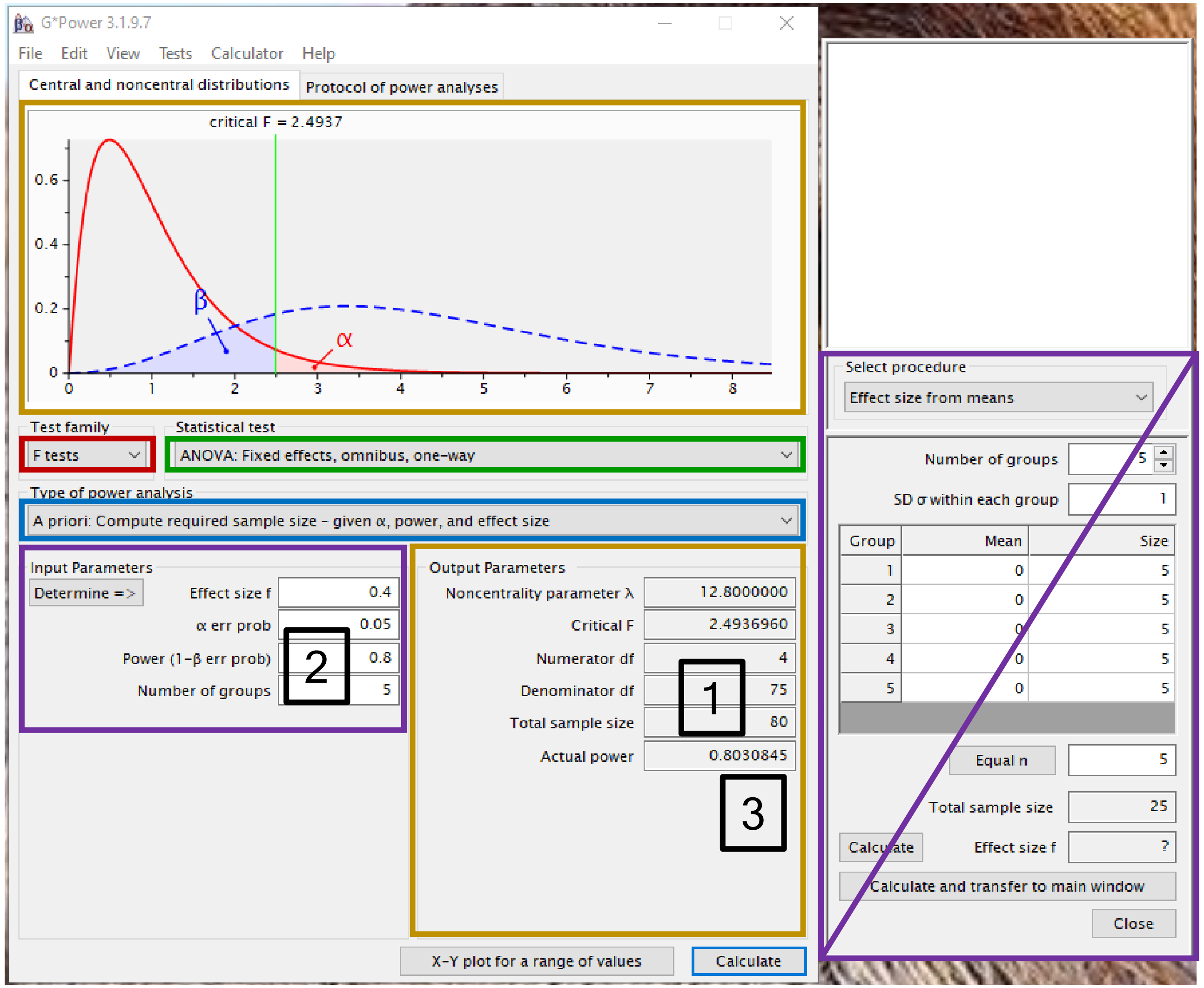
What we know

- **Number of groups**: five
- **Number of measurements**: one per sample
- **Significance level**: 0.05
- **Power**: 0.80

What we guess

- **Effect Size f**: 0.4 (large)

One Way ANOVA



a) A priori: compute required sample size

b) F tests

c) ANOVA: Fixed effects, omnibus, one-way

d) Input parameters (without values)

e) Output parameters with density plots

RESULTS:

1) **Total sample size** (Tss): 80
2) **Sample size per group**: Tss / n_g = 16
3) **Actual power:** 0.801 (check)

Repeated Measure ANOVA

Description

- Check if **at least one mean** is different among groups
- **Normally distributed (numeric) variables**
- Extension of paired t-test for more than two groups

Conditions

- **Numeric variables**: one — MEASUREMENTS
- **Categorical variables**: one — E.G. TREATMENT, TIMEPOINT, DOSE
- **Number of group per categorical variable**: more than two
- **Details**: parametric, paired — NORMAL DISTRIBUTION AND SAME SAMPLES

Parameters

- Significance level: 0.05
- Power: 0.80

Repeated Measure ANOVA

Example

- One treatment on patients
- Four measurements of blood pressure after treatment (1, 2, 3, 4 months)
- H0: no difference during time; H1: difference during time

What we know

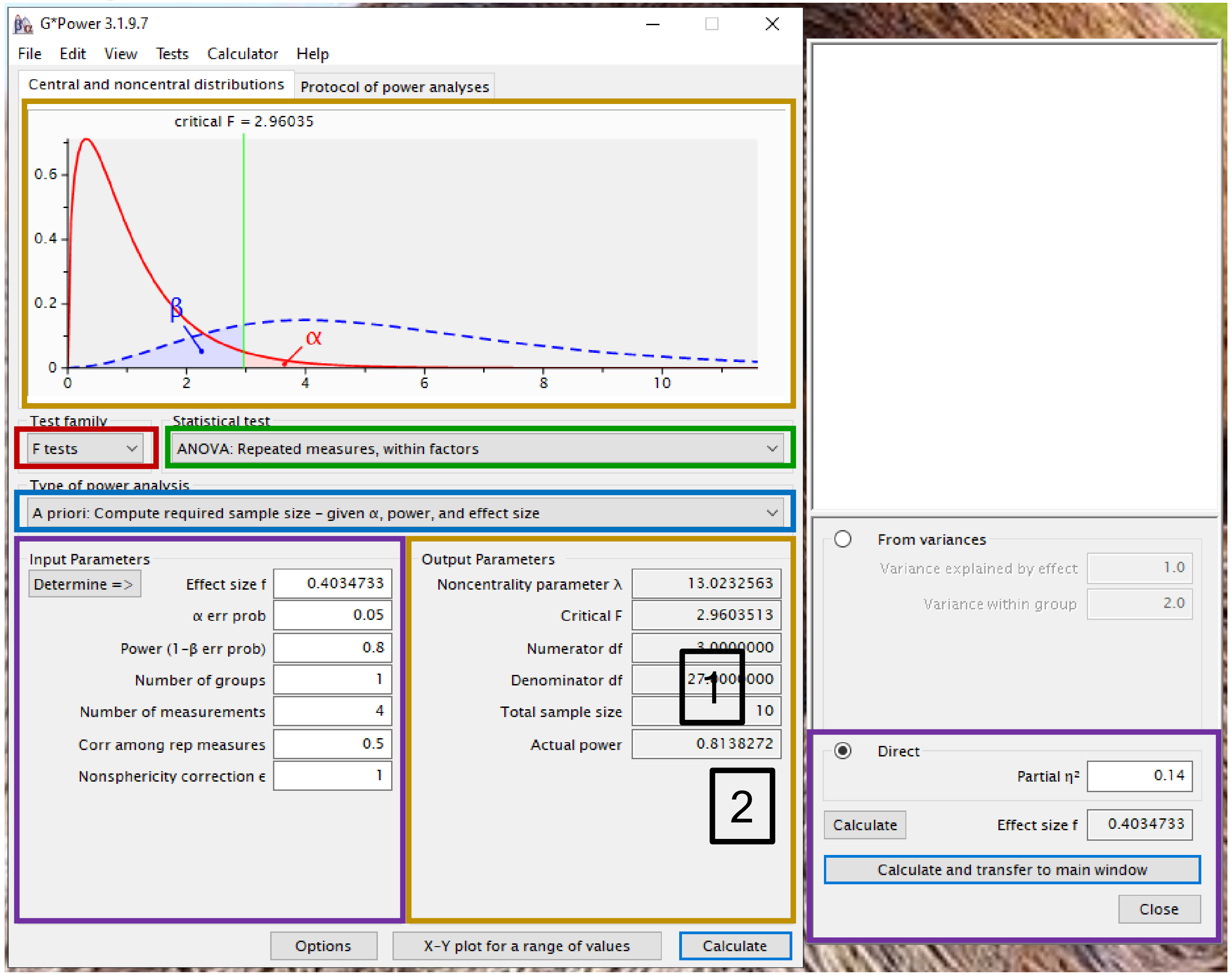
- **Number of groups**: one
- **Number of measurements**: four per sample
- **Significance level**: 0.05
- **Power**: 0.80

What we guess

- **Correlation**: 0.5
- **Non-sphericity correction**: 1
- **Partial** η^2: 0.14 (large)

CORRELATION: relationships among measurements
SPHERICITY: same differences of variances
PARTIAL η^2: ratio between variances (EFFECT SIZE)

Repeated Measure ANOVA



a) A priori: compute required sample size

b) F tests

c) ANOVA: Repeated measures, within factors

d) Input parameters (with Partial η^2)

e) Output parameters with density plots

RESULTS:

1) **Total sample size**: 10
2) **Actual power:** 0.814 (check)

Multi Way ANOVA

Description

- Define the number of **categories of interest** (one or more than one)
- **Normally distributed (numeric) variables**
- Extension of one way ANOVA for more than one category

Conditions

- **Numeric variables**: one — MEASUREMENTS
- **Categorical variables**: more than one — E.G. TREATMENT, TIMEPOINT, DOSE
- **Number of group per categorical variable**: more than one
- **Number of categories of interest**: one or more than one
- **Details**: parametric, unpaired — NORMAL DISTRIBUTION AND DIFFERENT SAMPLES

Parameters

- Significance level: 0.05
- Power: 0.80

Multi Way ANOVA

Example

- Different treatments on patients (a, b, c)
- Different ages (child, adult, elder) and different cancer stages (I, II, III, IV, V)
- H0: no difference in treatment across ages/stages;
  H1: difference in treatment across ages/stages

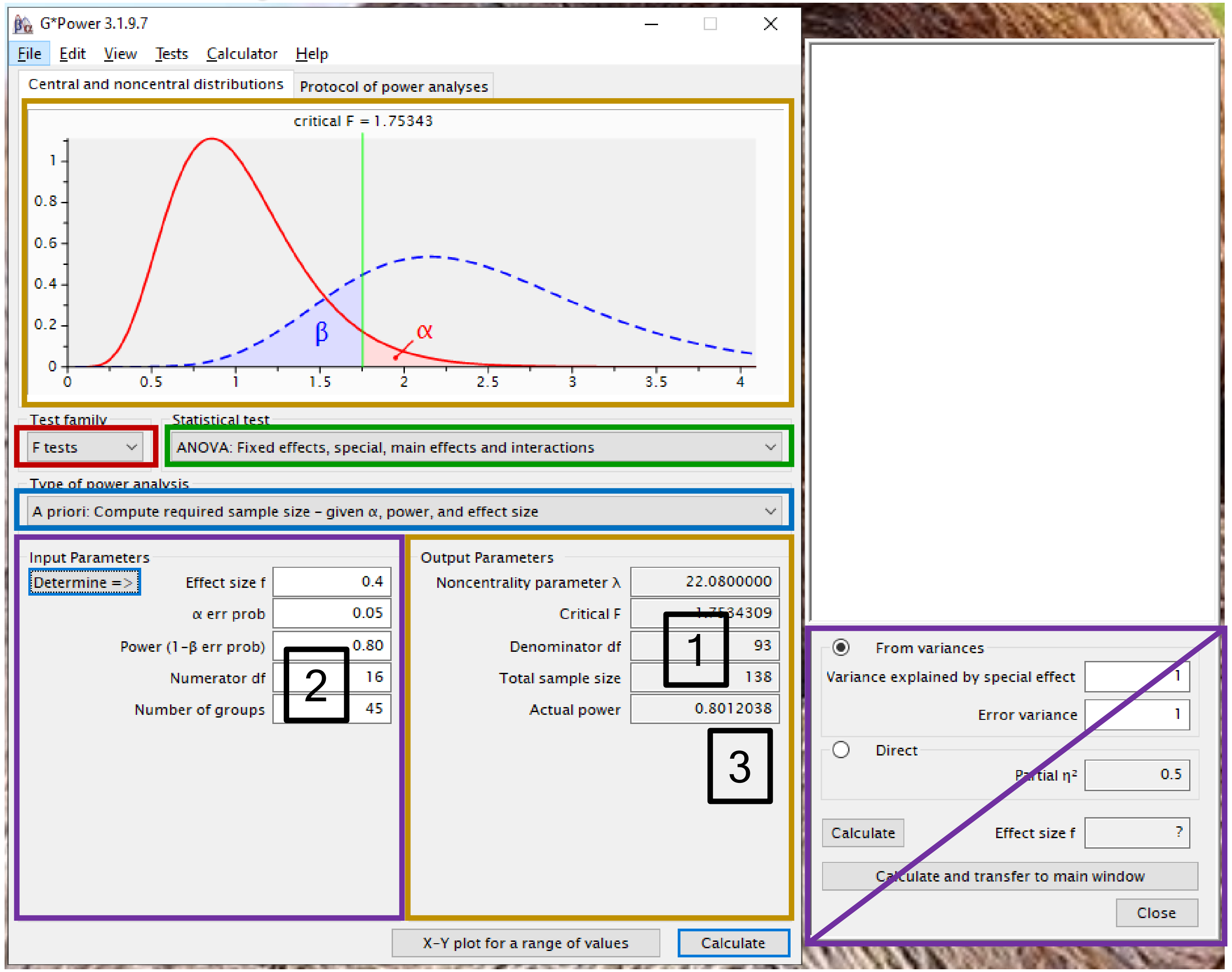
What we know

- **Number of groups**: three treatments (T), three ages (A), five stages (S), $n_g = TAS = 45$
- **Number of degrees of freedom**: $df = (T-1)(A-1)(S-1) = 16$
- **Significance level**: 0.05
- **Power**: 0.80

What we guess

- **Effect Size f**: 0.4 (large)

Multi Way ANOVA



a) A priori: compute required sample size

b) F tests

c) ANOVA: Fixed effects, special, main effects and interactions

d) Input parameters (without values)

e) Output parameters with density plots

RESULTS:

1) **Total sample size** (Tss): 138
2) **Sample size per group**: $Tss/n_g \approx 4$
3) **Actual power:** 0.801 (check)

Focus on: Non-Parametric Tests

Not normally distributed (numeric) variables

- Peculiar distribution and/or low number of samples
- Information obtained from a preliminary/previous/similar experiment
- No relaxation of the hypothesis

Rule of Thumb

- Compute the **parametric test and add 15%** to total sample size
- **Formula**: $Tss_{NP} = Tss + 0.15Tss = 1.15Tss$

Previous Examples

- **One Way ANOVA:** 1.15 x 80 = 92
- **Repeated Measure ANOVA**: 1.15 x 10 ≈ 12
- **Multi Way ANOVA**: 1.15 x 138 ≈ 159

GENERALIZED LINEAR MIXED MODELS

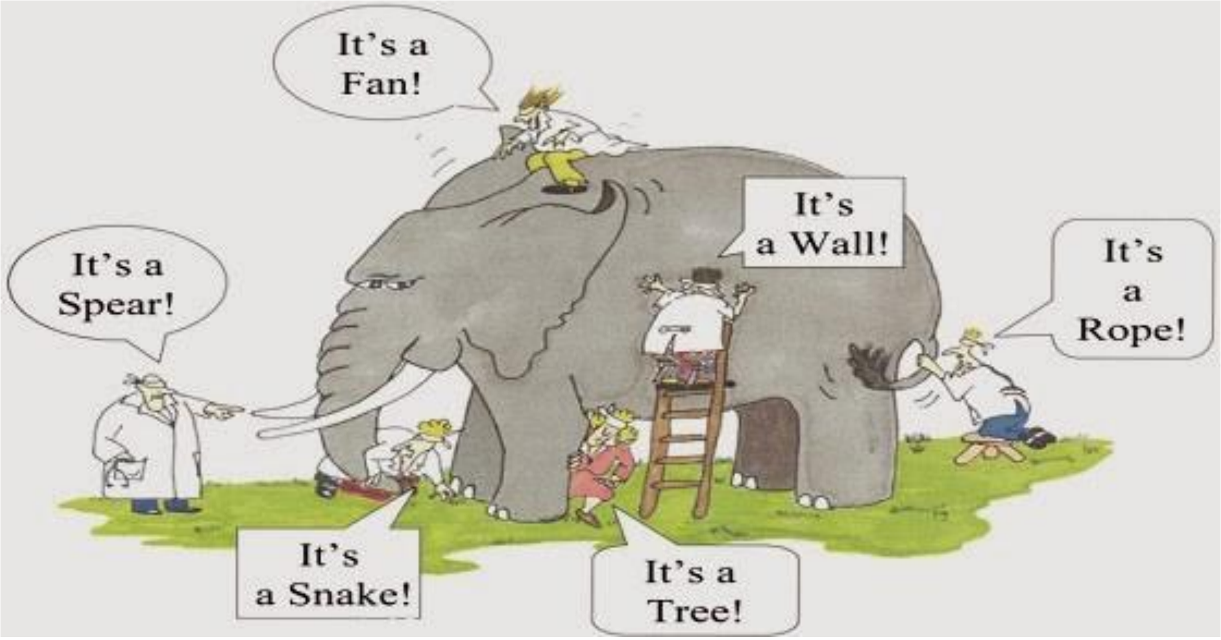
Take Home Message

- **Sample size** is the most important parameter to detect a possible effect among different conditions
- **Sample size**, **significance level**, **power**, and **effect size** are strictly connected each other (considering the distribution of variables)
- **A priori power analysis** provides the minimum number of the samples to detect a possible effect, **with fixed values for the parameters**
- **Different types** of a priori power analysis exist, based on the **parameter of interest** and on the **statistical test** to perform
- **G*Power** is a free, downloadable, and user-friendly tools to perform a priori and a posteriori power analysis

Final Remarks

To consult the statistician after an experiment is finished is often merely to ask him to conduct a postmortem examination. He can perhaps say what the experiment died of.

Sir R. A. Fisher
First Session of the Indian Statistical Conference, Calcutta, 1938



Eugenio Del Prete, M.Eng., Ph.D.
Biostatistician and Data Scientist
Telethon Institute of Genetics and Medicine (TIGEM)
Pozzuoli (NA), Italy
e-mail: e.delprete@tigem.it

References

[1] Faul, F. **Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses,** Behavior Research Methods (2009).
[2] Uttley, J. **Power Analysis, Sample Size, and Assessment of Statistical Assumptions - Improving the Evidential Value of Lighting Research**, LEUKOS (2019).
[3] Kang, H. **Sample size determination and power analysis using G*Power software**, Journal of Educational Evaluation of Health Professions (2021).

[h1] *https://www.bioinformatics.babraham.ac.uk/training/*
[h2] *https://med.und.edu/research/daccota/berdc-resources.html*
[h3] *https://www.psychologie.hhu.de/arbeitsgruppen/allgemeine-psychologie-und-arbeitspsychologie/gpower*