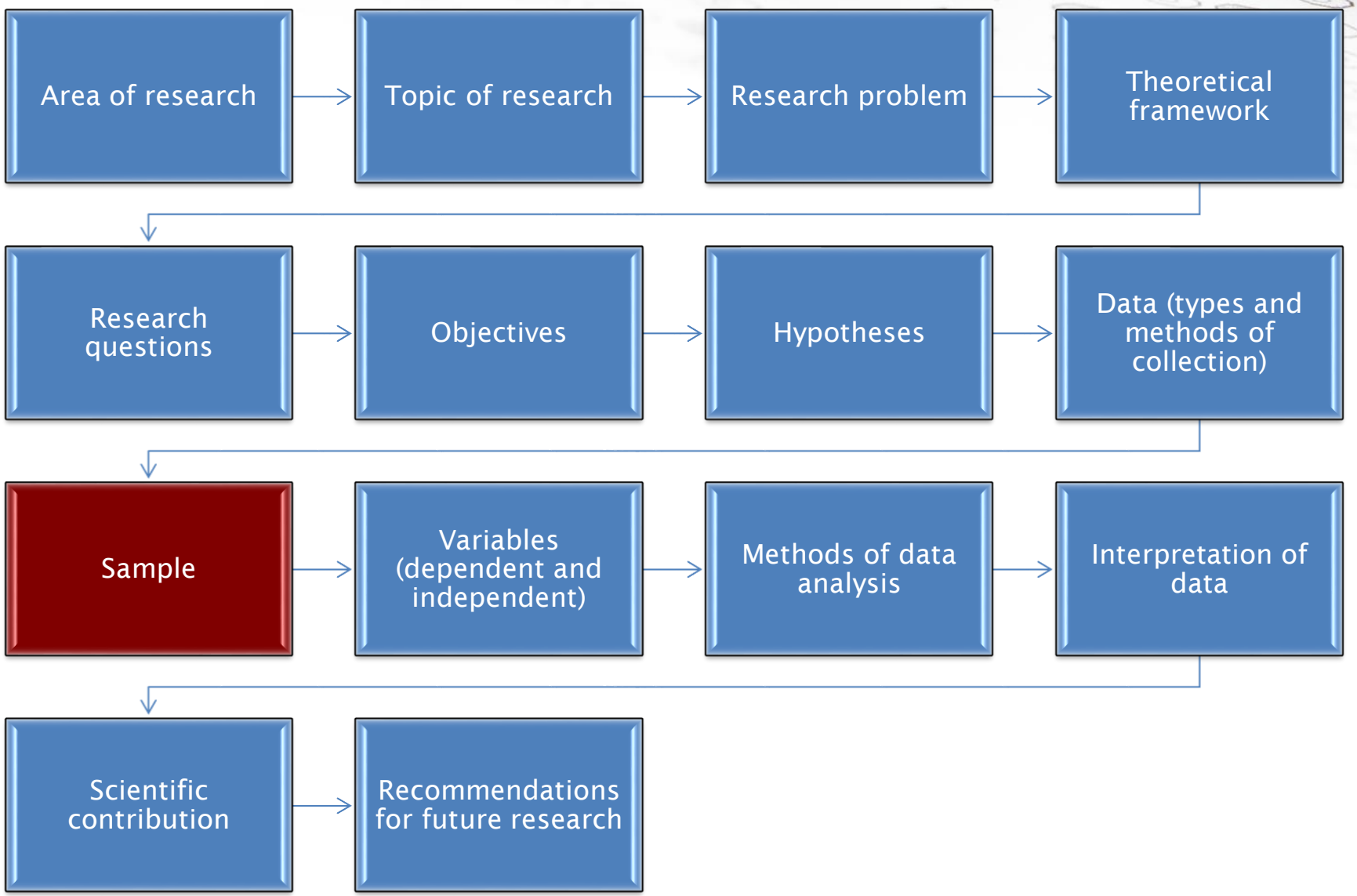
STATISTICAL RESEARCH METHODS

International inter-university postgraduate interdisciplinary doctoral study ENTREPRENEURSHIP AND INNOVATIVENESS

Jasna Horvat, Ph.D.

Josipa Mijoč, Ph.D.

RESEARCH PROCESS

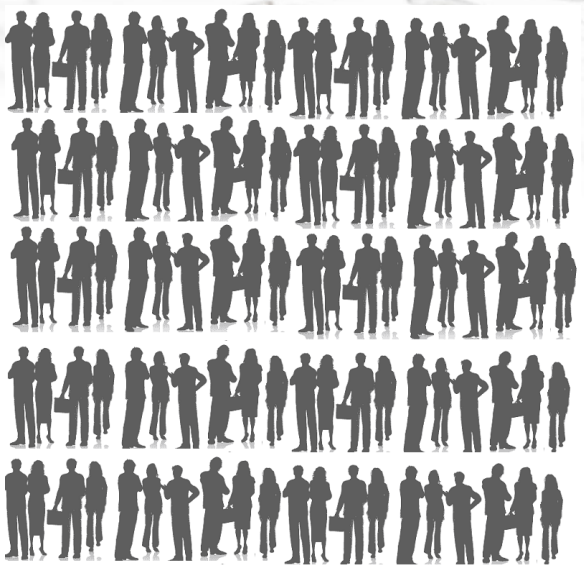


Sample and population

- POPULATION
  - Defined group of research subjects that are being sampled
  - All individuals of interest
  - Parameter
    - Characteristic of population
  - Population size = N
- SAMPLE
  - Subset from whole population
  - calculate a statistic
  - Sample size = n

POPULATION

SAMPLE



n = 100

N = 10 000

THE BIG PICTURE OF STATISTICS

Theory

↓

Research question / Hypothesis to test

↓

Design Research Study

↓

Collect Data

(measurements, observations)

USING STATISTICS!

Depends on our goal:

Describe characteristics
organize, summarize, condense data

DESCRIPTIVE STATISTICS

Test hypothesis, Make conclusions,
interpret data, understand relations

INFERENTIAL STATISTICS

Statistical procedures can be divided into two major categories

- descriptive statistics
- inferential statistics

DESCRIPTIVE STATISTICS

- descriptive statistics describe the statistical data
- using numerical and graphical methods to the collected data presented in an understandable and clear manner

TEXT

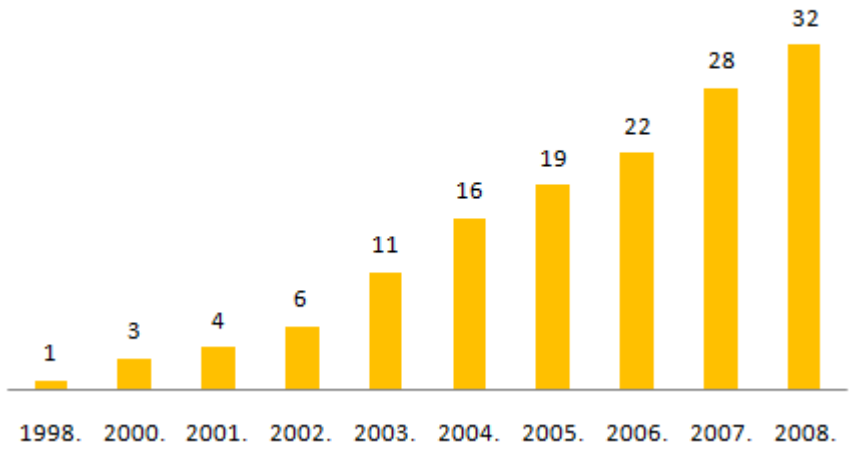
According to HZZ in late October 344,444 people were registered as unemployed, which means that the number of unemployed in 20 days more than 5500. In October 2013, 25,933 people were left out from the unemployment registration, of which 14.776 found job. The unemployment register during the same month, joined 46,594 people, which is three percent more than October 2012 .

TABLE

Ulasci u evidenciju: Prostorna jedinica - županija, Godina - Mjesec

Godina / Prostorna jedinica - županija	▶ 2004	▶ 2005	▶ 2006	▶ 2007	▶ 2008	▶ 2009
ZAGREBAČKA	11.345	11.101	11.031	8.941	8.774	14.207
KRAPINSKO-ZAGORSKA	5.552	5.353	5.049	4.008	4.049	6.669
SISAČKO-MOSLAVAČKA	11.568	11.224	10.925	9.772	10.296	13.200
KARLOVAČKA	8.048	7.404	6.834	6.221	6.158	8.503
VARAŽDINSKA	9.798	8.934	8.372	7.159	6.678	10.689
KOPRIVNIČKO-KRIŽEVAČKA	6.242	5.490	5.541	5.276	4.899	6.933
BJELOVARSKO-BILOGORSKA	9.325	8.847	8.985	8.412	7.923	9.882
PRIMORSKO-GORANSKA	15.229	14.487	15.207	12.971	12.767	17.428
LIČKO-SENJSKA	3.421	2.625	2.877	2.201	2.222	2.703

Number of unemployed



GRAPH

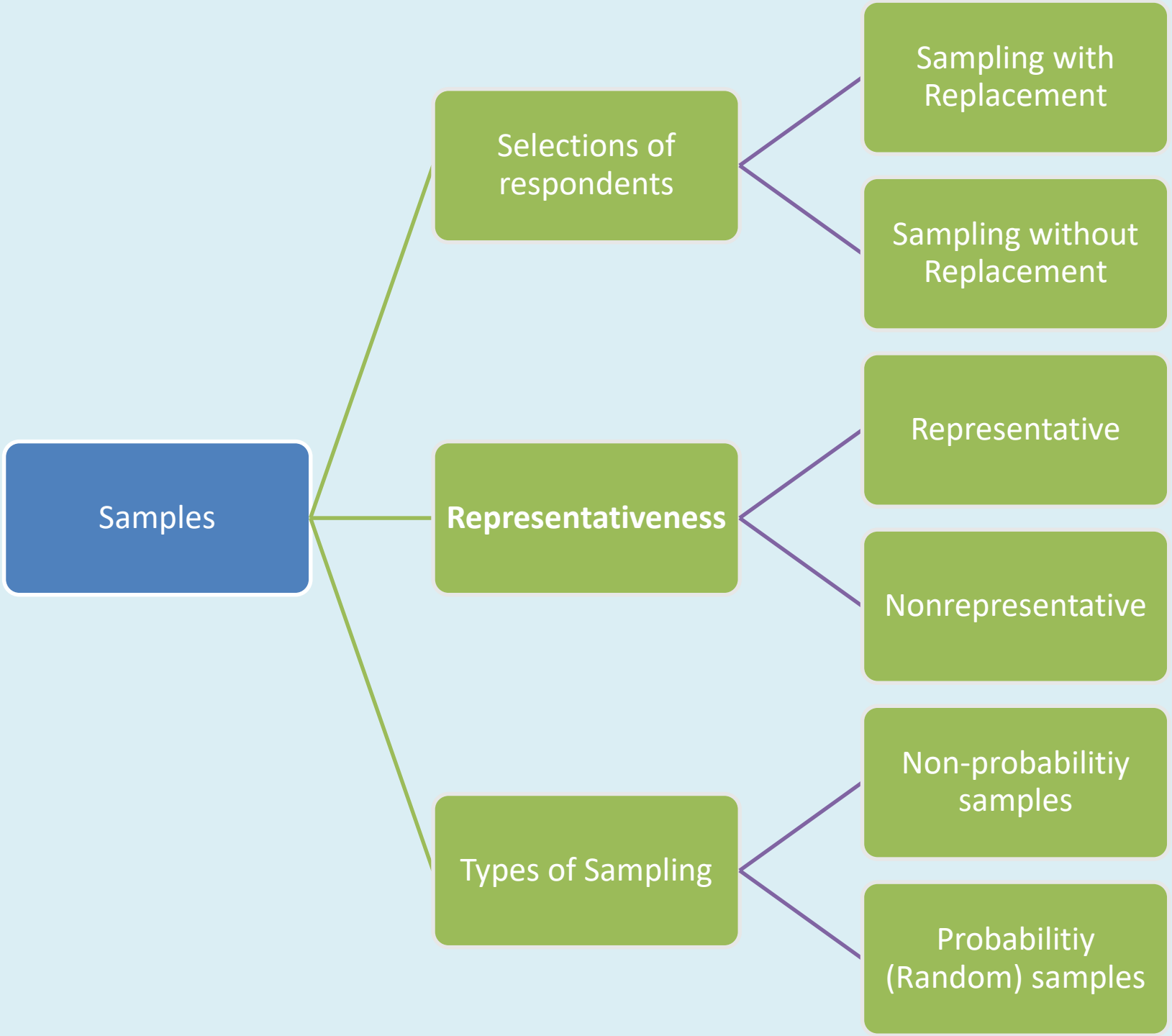
INFERENTIAL STATISTICS

- Inferential statistics make conclusions about the population based on the sample using
  - estimation
  - hypothesis testing
  - determining the relationship between variables
  - predictions about population

- Descriptive statistics
  - How many women are employed in management positions in Croatian companies?
  - How many hours of overtime during one month reach employees of the company?
- Inferential statistics
  - Is there a connection between gender and the decisions on starting your own business?
  - What factors affect the younger age groups (young people) when deciding on the selection of career times?

Descriptive statistics

- SPSS ☺
- Tomorrow ☺

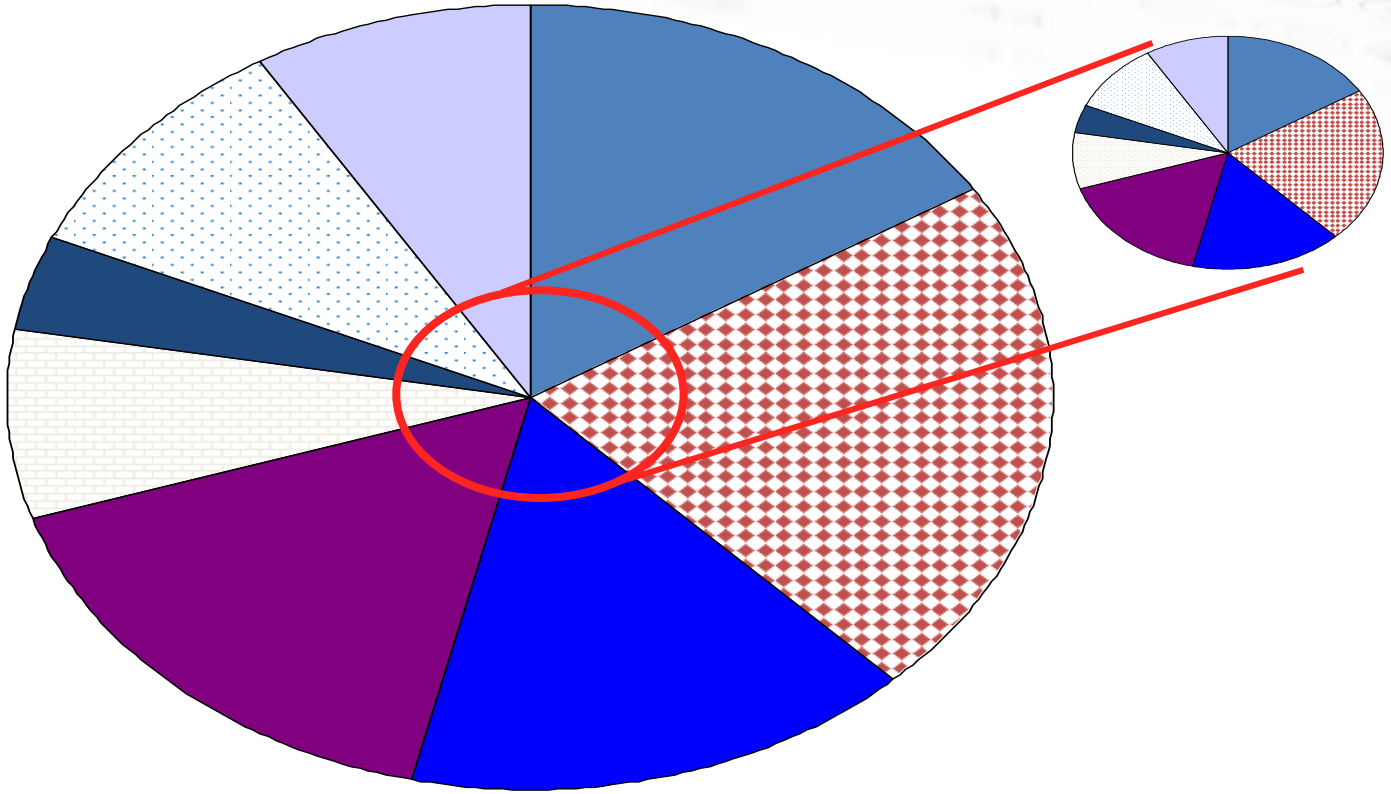

Samples
Selections of respondents
Sampling with Replacement
Sampling without Replacement
Representativeness
Representative
Nonrepresentative
Types of Sampling
Non-probabilitiy samples
Probabilitiy (Random) samples

Sampling

- the process of selecting units (e.g., people, cases, items or data) from a population of interest so that by studying the sample we may fairly generalize our results back to the population from which they were chosen (Trochim, 2006).
- A precise of the target population is essential and usually done in terms of:
  - Elements
    - person or object which data is sought and about which inferences are to be made
      - (e.g., people, cases, items or data)
  - Sampling units
    - target population element available for selection during the sampling process
  - Sampling frame
    - a representation of the elements of the target population

Representative Sample

- Representative sample is one that:
  - represent the key characteristics of total population
    - so you can *generalize* to population
  - thumbnail picture of selected population
    - contains the essential characteristics of the entire population



Types of Sampling

Non-probability samples	Convenience
	Quota
	Judgment
	Snow-ball
Probability (Random) samples	Simple Random Sampling
	Stratified Random Sampling
	Cluster Sampling
	Systematic Sampling

Non-probability samples

- The process of selecting a sample from a population **without using** (statistical) probability theory.
- Not every element of the population has the opportunity for selection in the sample
  - the researcher CANNOT estimate the error caused by not collecting data from all elements/members of the population

Non-probability samples

- No sampling frame
- Not reliable
- Non-random selection
- More likely to produce a biased sample
- Restricts generalization
- Questionable representativeness

1. Convenience sample

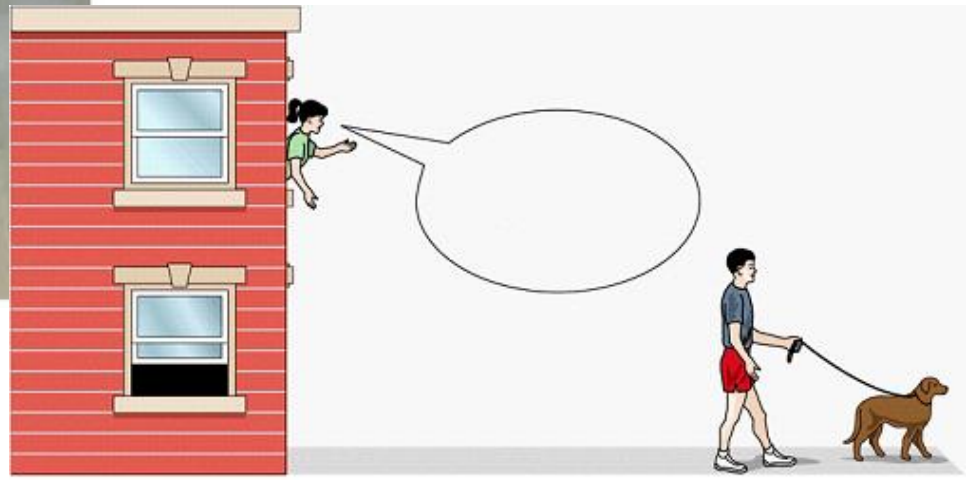
- Selection strategy
  - Select cases based on respondents availability for the study
    - individuals who are easiest to reach
    - „Man on the street”
  - Selecting easily accessible respondents with no randomization
  - Available or accessible clients
  - It is done at the “convenience” of the researcher

1. Convenience sample

- Purpose:
  - Saves time, money and effort; but at the expense of information and credibility.
- Problem: *No* evidence for representativeness


© Shibuya246.com


Petrikovac
KAWA SAKI
3P



2. Quota samples

- Selection strategy
  - Select a sample that yields the **same proportions as the population proportions** on easily identified variables
    - e.g. population has 40% women and 60% men, you want your sample to meet that quota
- Represent major characteristics of population by sampling a proportional amount of each

2. Quota samples

- Representative Sample?
  - representative only for the characteristics that are the basis for distinguishing groups of units
    - E.g.
      - Gender
      - Education

2. Quota samples

- Purpose
  - Taking a set number of cases from each subgroup to raise analytic confidence and representativeness
- Problem:
  - How do you pick the characteristics?
  - How do you know their proportion in population?

3. Judgment sample (Expert sampling)

- Selection strategy
  - rely on the judgment of the researcher
  - according to an experienced researcher's belief that they will meet the requirements of study panel of experts make a judgment about the representativeness of your sample
- Advantage
  - expert judgment supports the sampling

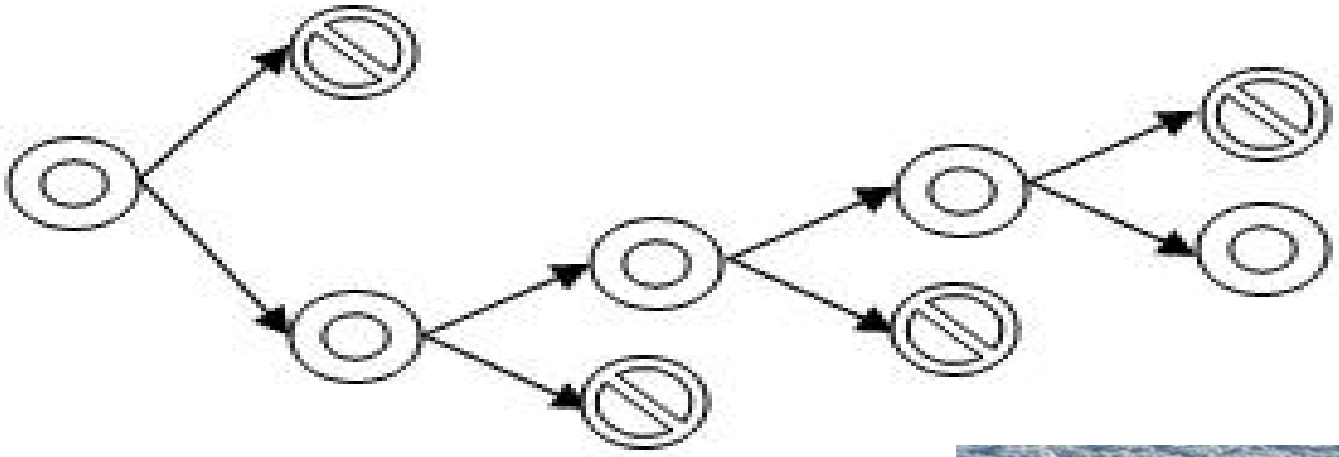
3. Judgment sample

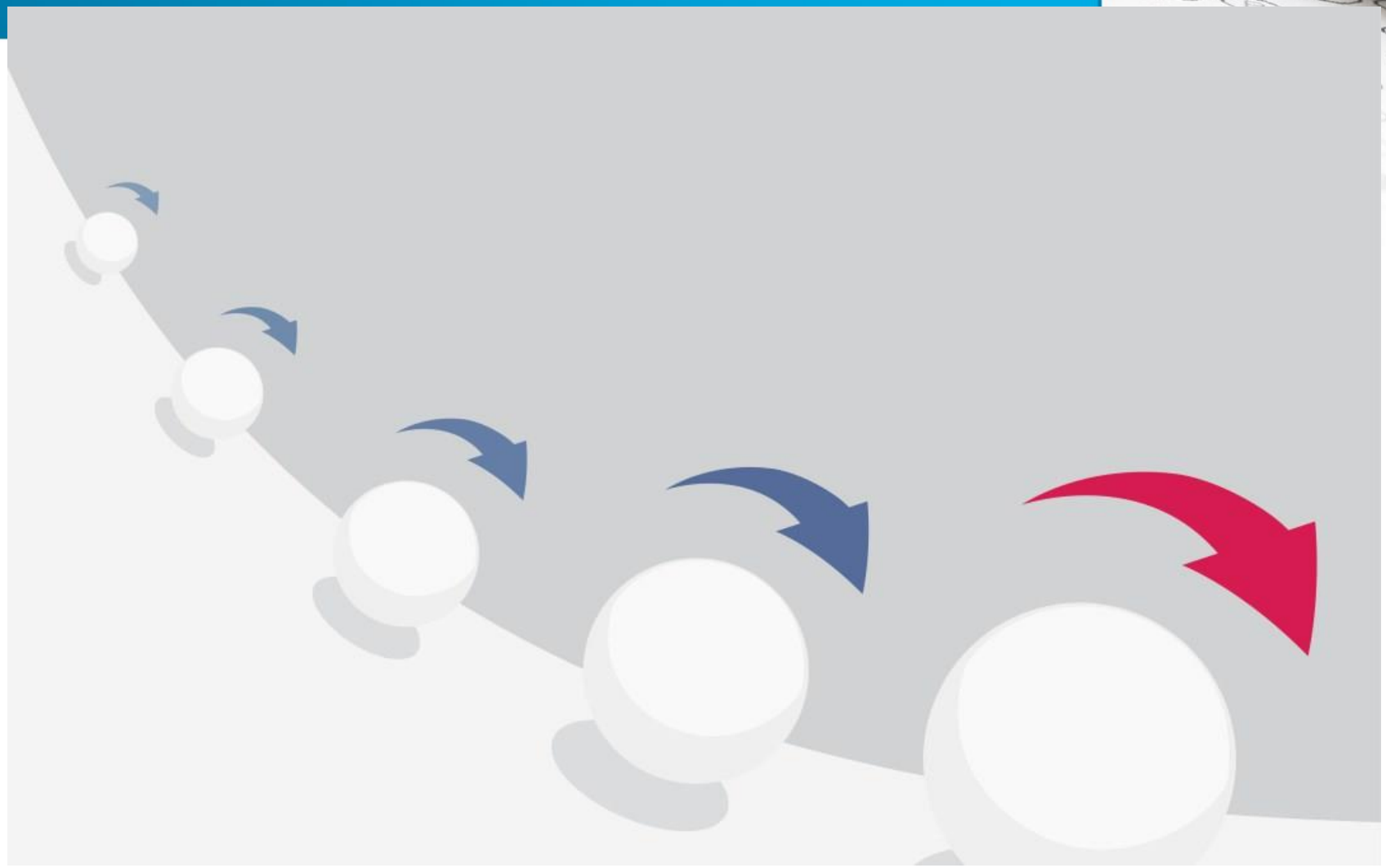
- Problem
  - great deal of sampling error since the researcher's judgment may be wrong
    - the "experts" may be wrong
- Example:
  - Specific People
  - Specific cases/organizations
  - Specific events
  - Specific pieces of data

3. Judgment sample

4. Snow-ball sample

- target population is rare and unique and compiling a complete list of sampling units is a nearly impossible task
  - identify hard-to-reach populations
- Selection strategy
  - One person recommends another, who recommends another, who recommends another, etc.
- rare groups of people tend to form their own unique social circles.





Types of Sampling

Non-probability samples	Convenience
	Quota
	Judgment
	Snow-ball
Probability (Random) samples	Simple Random Sampling
	Stratified Random Sampling
	Cluster Sampling
	Systematic Sampling

Probability samples

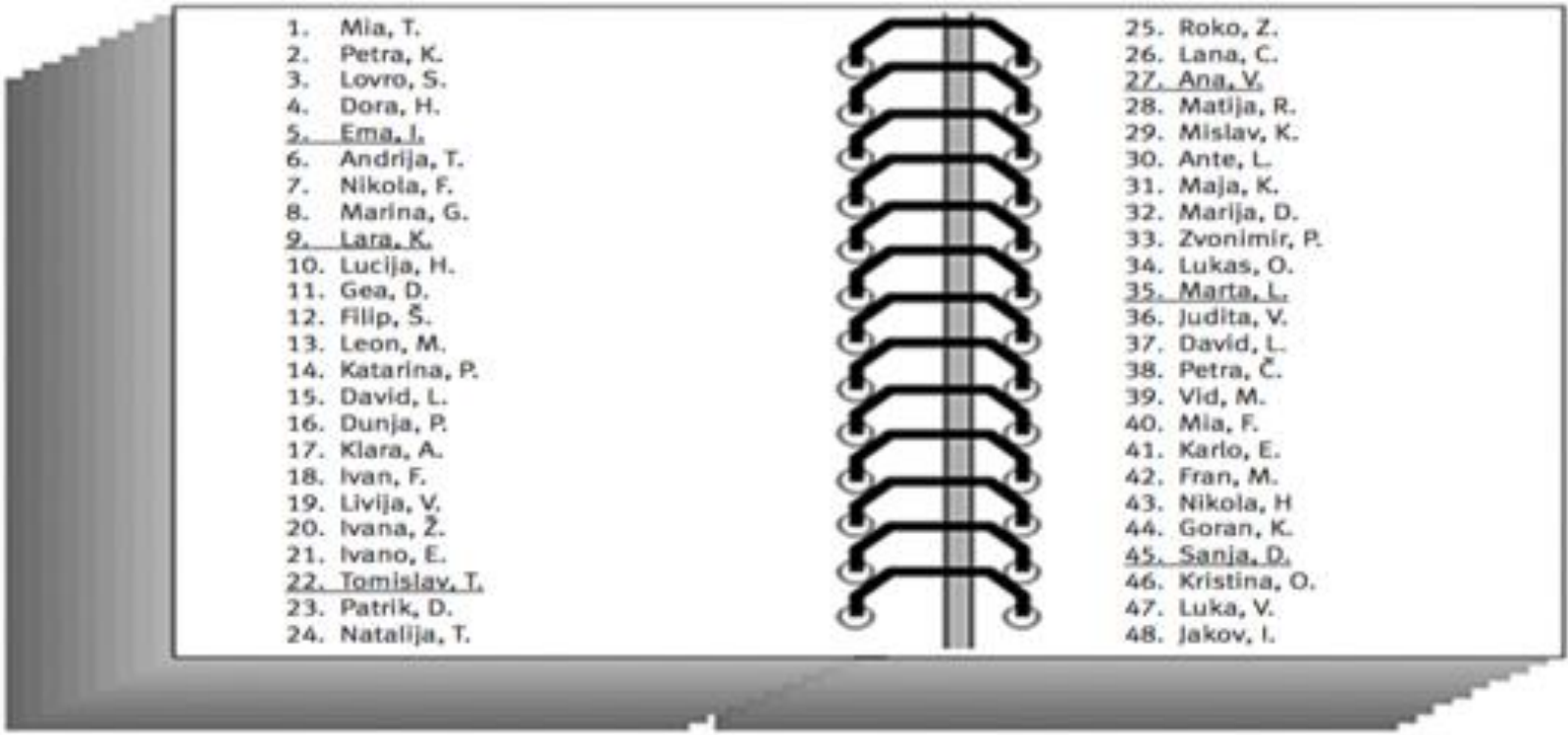
- Everyone in the population has equal opportunity for selection as a subject
- Increases sample's representativeness of the population
- Decreases sampling error and sampling bias
- Provide *unbiased* selection of units in the sample

1. Simple Random Sampling

- All participants have equal chance of being selected
  - Roll dice, flip coin, draw from hat, random number
- Random sampling is the only way to ensure that your sample is truly representative of the target population
- Random selection reduce bias
- Ideal sampling
  - but sometimes not possible

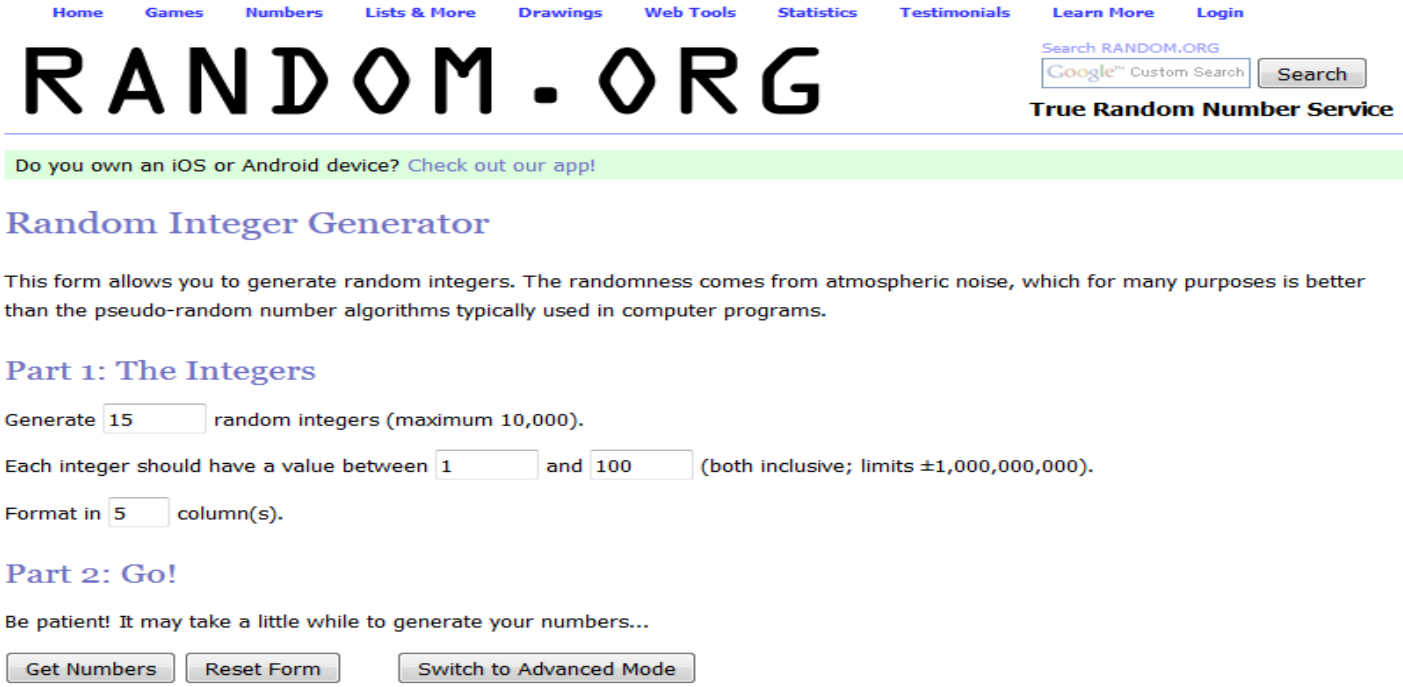
List of buyers Sample frame

Population (N=48)
Sample (n=6)



Example: random selection of n numbers from the list of N numbers

- N = 100
- n = 15



RANDOM.ORG

True Random Number Service

Random Integer Generator

Here are your random numbers:

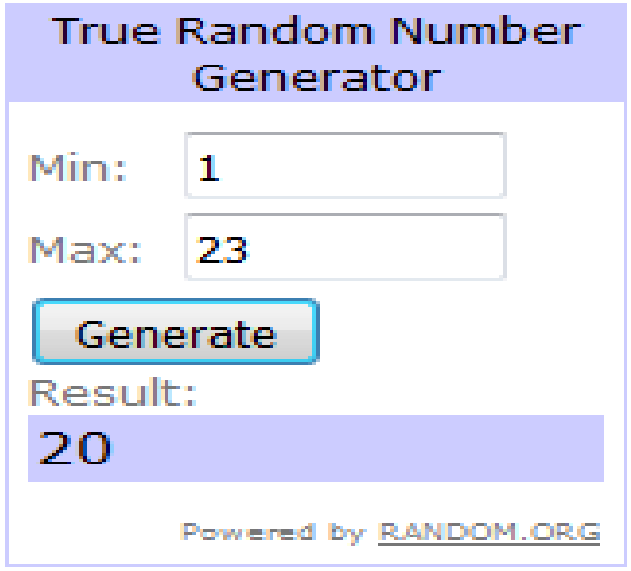
11	53	6	25	11
88	94	55	80	13
86	7	37	8	50

Timestamp: 2016-01-28 14:07:37 UTC

Again! Go Back

Note: The numbers are generated left to right, i.e., across columns.

Example: random selection of 1 number from the list of N numbers



RANDOM.ORG

Lottery Quick Pick

This form allows you to quick pick lottery tickets. The randomness comes from atmospheric noise, which for many purposes is better than the pseudo-random number algorithms typically used in computer programs.

Pick 2 ticket(s) for the lottery in Croatia
called Eurojackpot
which uses 5 numbers, the highest of which is 50
and 2 numbers, the highest of which is 10

Your chance of matching all numbers with this combination is 2 in 95,344,200.

Pick Tickets | Reset Form | Save Settings | Restore Settings | Clear Settings

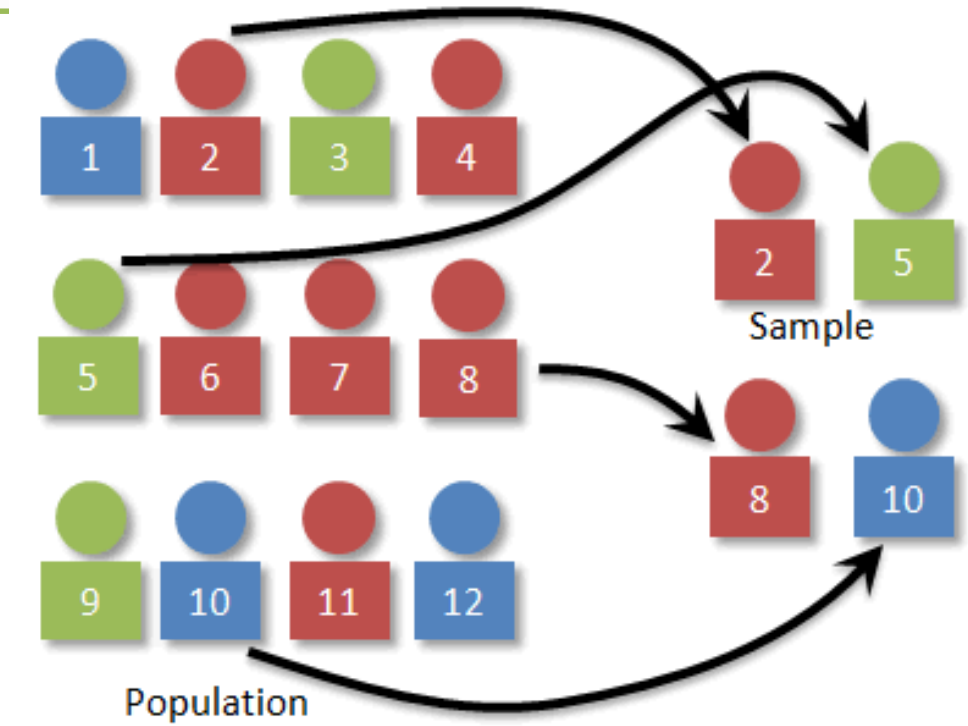
(settings can be saved in your browser)

If your lottery is not listed, pick any country, then select 'Other Lottery' and fill in the other fields. You
you want to know more about your chances, you can use Gerry Quinn's Lotto Odds Calculator (
RANDOM.ORG's lottery quick pick will help you win.

1. Simple Random Sampling

- Required list of the entire population
- Researchers use
  - table of random numbers
  - random digit dialing
  - other random selection methods

1. Simple Random Sampling

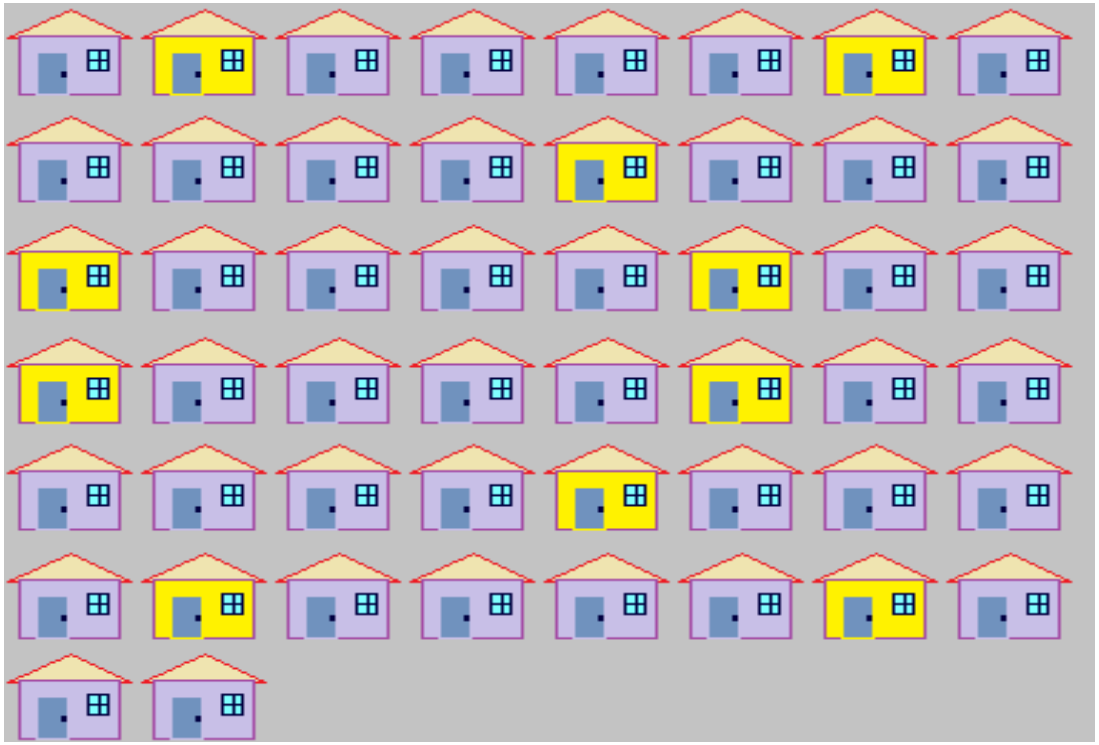


2. Systematic Sampling

- Method that requires selecting samples based on a **system of intervals** in a numbered population
- Selection strategy
  - random starting point
    - (1) selects a subject at random from the first k names in the sampling frame
  - then picking each ith
    - (2) selects every ith element listed after that one

2. Systematic Sampling

e.g. List of household



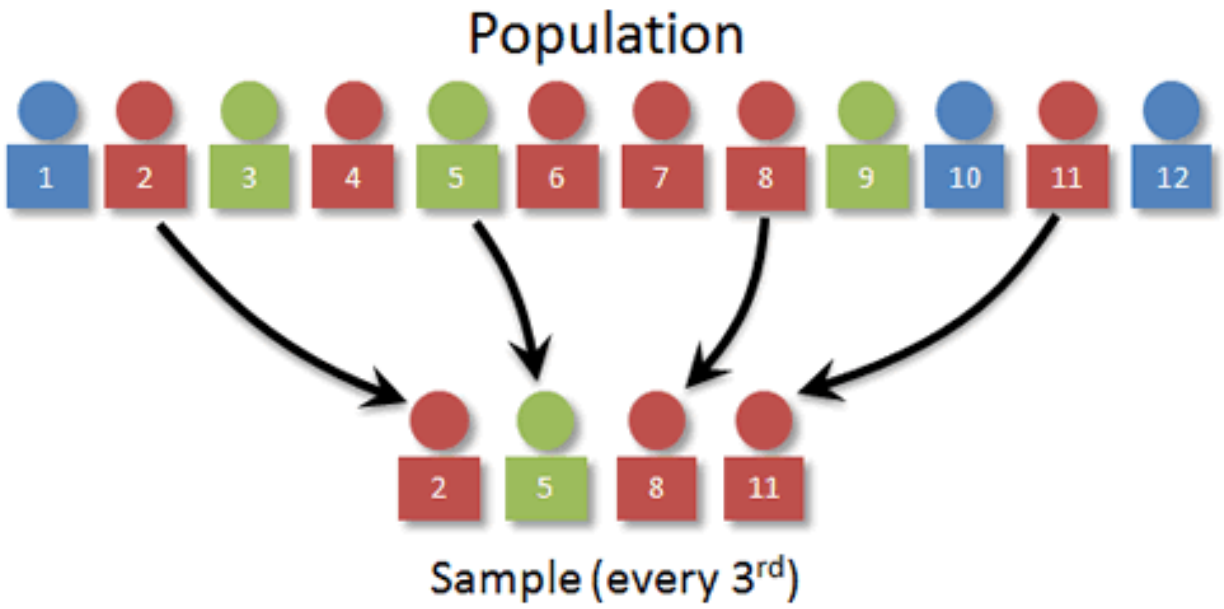
2. Systematic Sampling

The number k is called the skip number.

- Population size is N, sample size is n, k = N/n
- N= 50, sample size n = 10, 50/10 = 5
- Every fifth participant

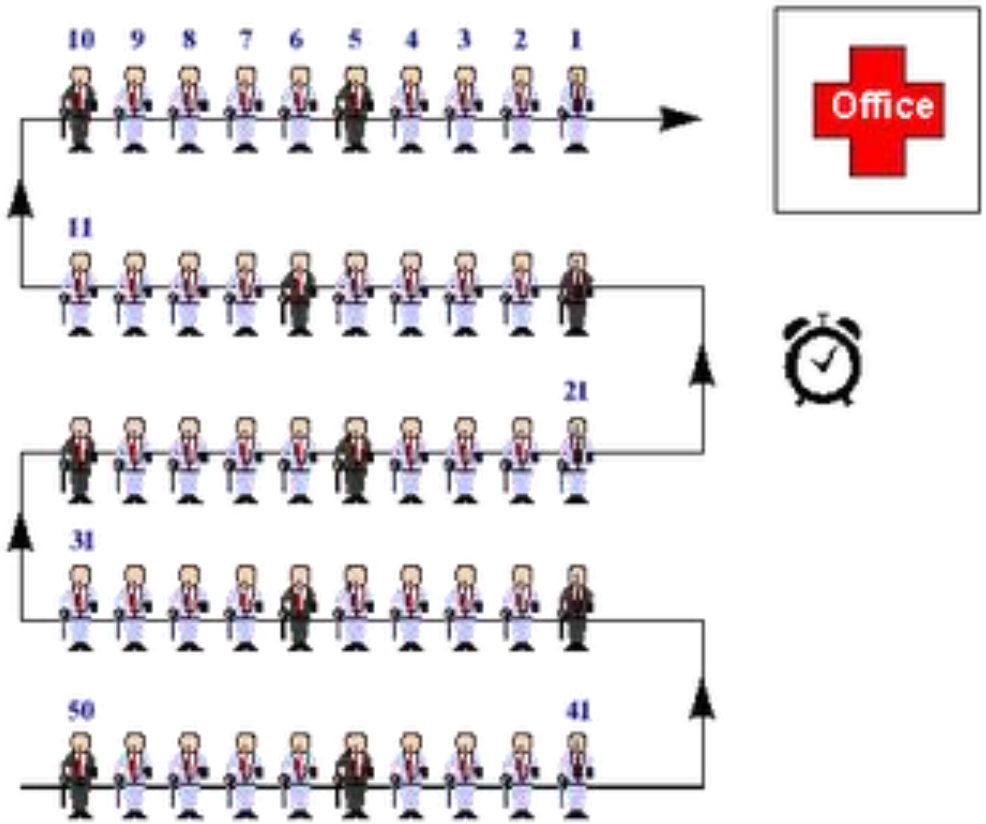
- Systematic random sample elements can be obtained via various means such as customer list, membership list, taxpayer, roll and so on.

2. Systematic Sampling



2. Systematic Sampling

- Every **k**th member (for example every 10th person) is selected from a list of all population members.



2. Systematic Sampling

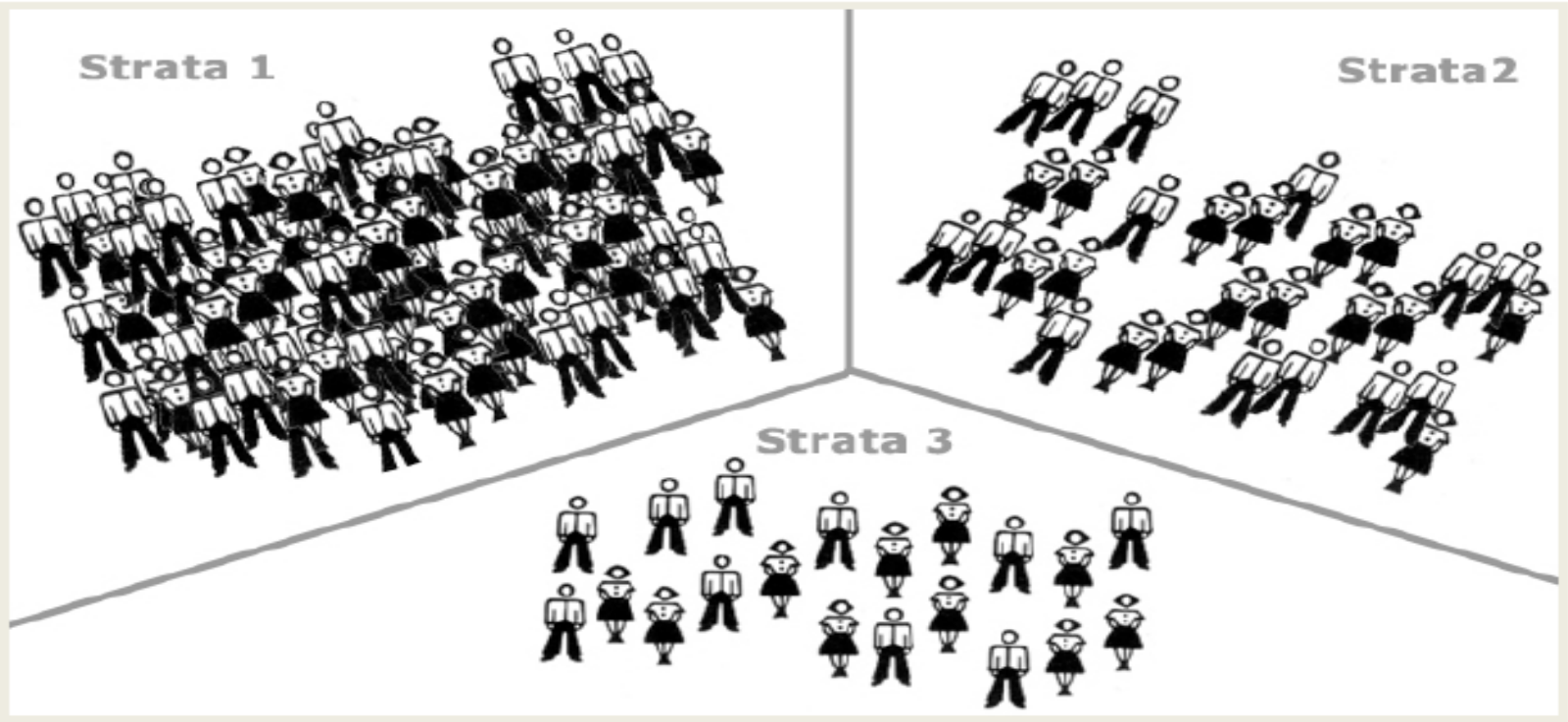
Yearbook

List of students in one class

3. Stratified Random Sampling

- Selection strategy
  - Two-steps procedure
    - First step the population is divided into mutually exclusive and collectively exhaustive sub-populations, which are called strata
      - population is divided into two or more groups called strata, according to some criterion, example age and subsamples are randomly selected from each strata
      - grouping elements that share certain characteristics
    - Second step – randomly chosen population elements

3. Stratified Random Sampling



3. Stratified Random Sampling

- Used when there is considerable diversity among the population elements.
- The major aim of it is to reduce cost without losing in precision:
  - proportionate stratified sampling
  - disproportionate stratified sampling

3. Stratified Random Sampling

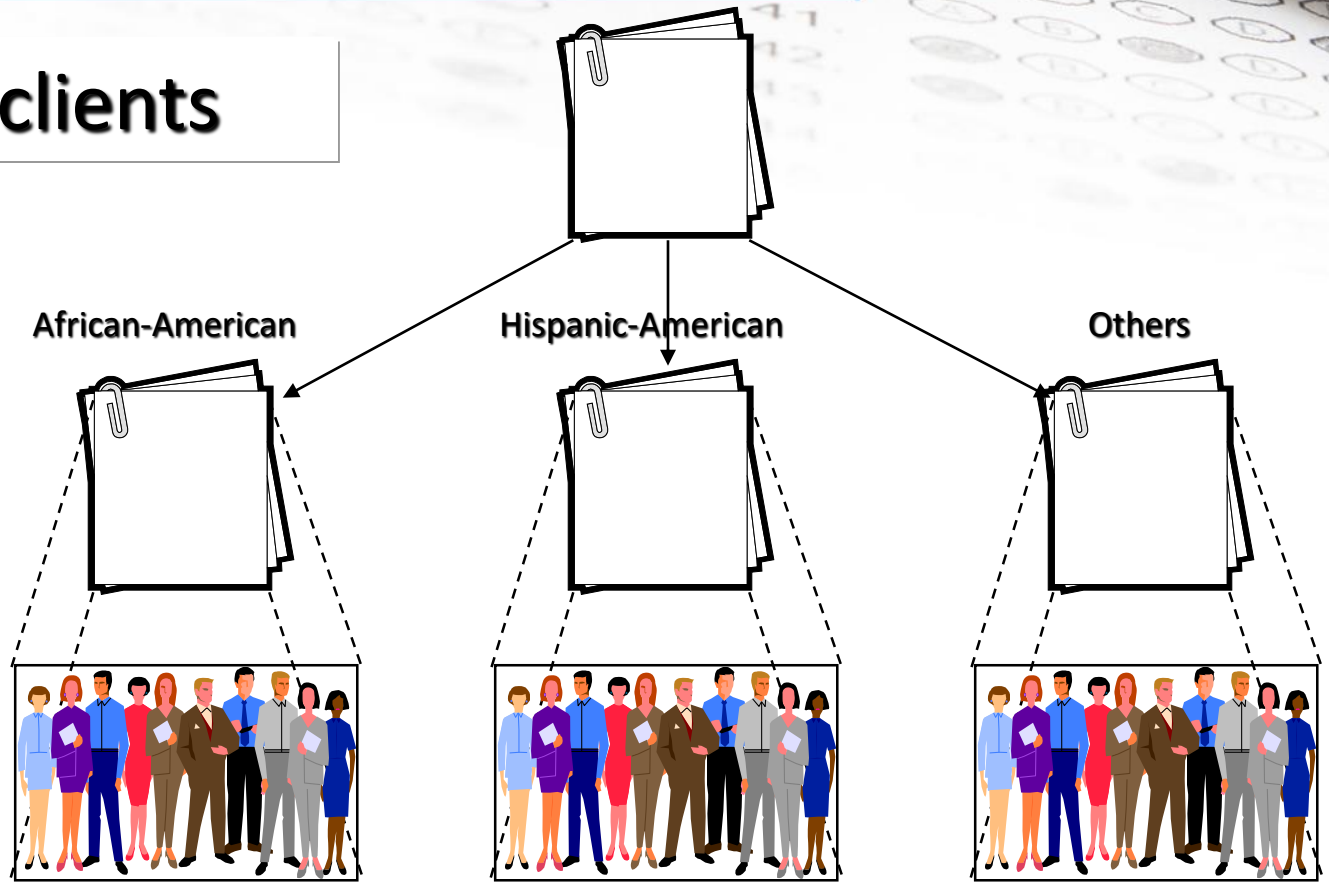
- Potential problem
  - assurance of representativeness
  - comparison between strata
  - understanding of each stratum as well as unique characteristics.
- **Variables** often used include: age, gender, geographic region, or other socio-demographic characteristics, religion, or maybe type of games used

3. Stratified Random Sampling

3. Stratified Random Sampling

List of clients

Strata



Random subsamples of n/N

4. Cluster Sampling

- Selection strategy
  - Two-steps procedure
  - (1)Divides the population into groups or clusters
    - The population is divided into subgroups (clusters) like families
    - A number of clusters are selected randomly to represent the total population
    - Population divided into clusters of homogeneous units, usually based on geographical contiguity (but element of each cluster is heterogeneous)
  - The population is divided into mutually exclusive and collectively exhaustive sub-populations
  - (2)A simple random sample is taken of the subgroups and then all members of the cluster selected are surveyed

4. Cluster Sampling

Klaster iz Slavonije

Klaster iz Cent. Hrv.

Klaster iz Zagreba

Klaster iz Istre

Klaster iz Kvarnera

4. Cluster Sampling

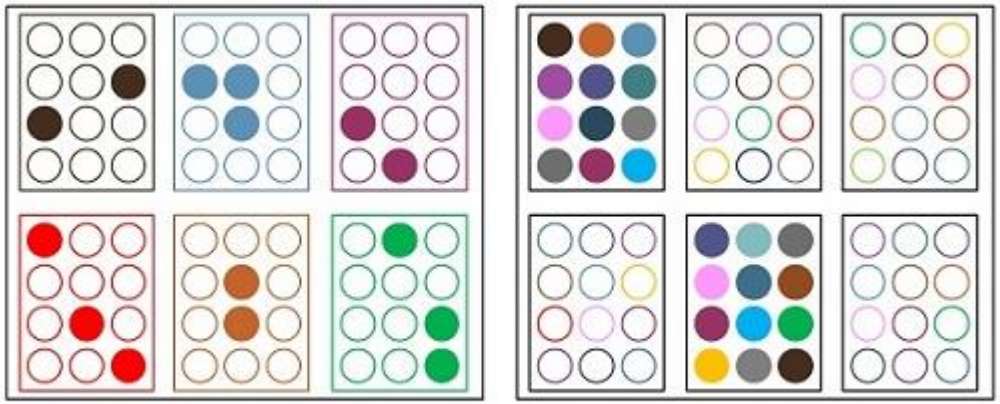
- Advantages :
  - Cuts down on the cost of preparing a sampling frame
  - Administratively useful, especially when you have a wide geographic area to cover
- Disadvantages:
  - Sampling error is higher for a simple random sample of the same size
  - Problem with representatives

Difference Between Strata and Clusters

- Although strata and clusters are both non-overlapping subsets of the population, they differ in several ways.

- With stratified sampling, the best survey results occur when elements within strata are
  - internally homogeneous

  However, with cluster sampling, the best results occur when elements within clusters are
  - internally heterogeneous



Stratified Sampling Vs Cluster Sampling

Example

Sample:

Quantitative research: University students of economics were selected for the sample since student sample is very common in testing of self-employment intentions. Students of the graduate study are the target group of this research since they are nearing completion of their studies and thus also before the decision on the choice of career. It is expected that graduate students will have enough time and energy for planning future business ventures (Audet, 2004). Wu and Wu (2008) argue that by understanding entrepreneurial intentions of university students it is possible to better predict their future real decision on launching a new business venture.

- The questionnaire was completed by a total of 453 students, 428 of which were used in the analysis.

Components of Sampling Designs

- Sampling (Random; Other probability sampling or nonprobability – such as convenience, purposive, or snowball)
- **Groups** – one group; two, or multiple group comparisons
- **Time intervals** – pre and post tests or multiple observations

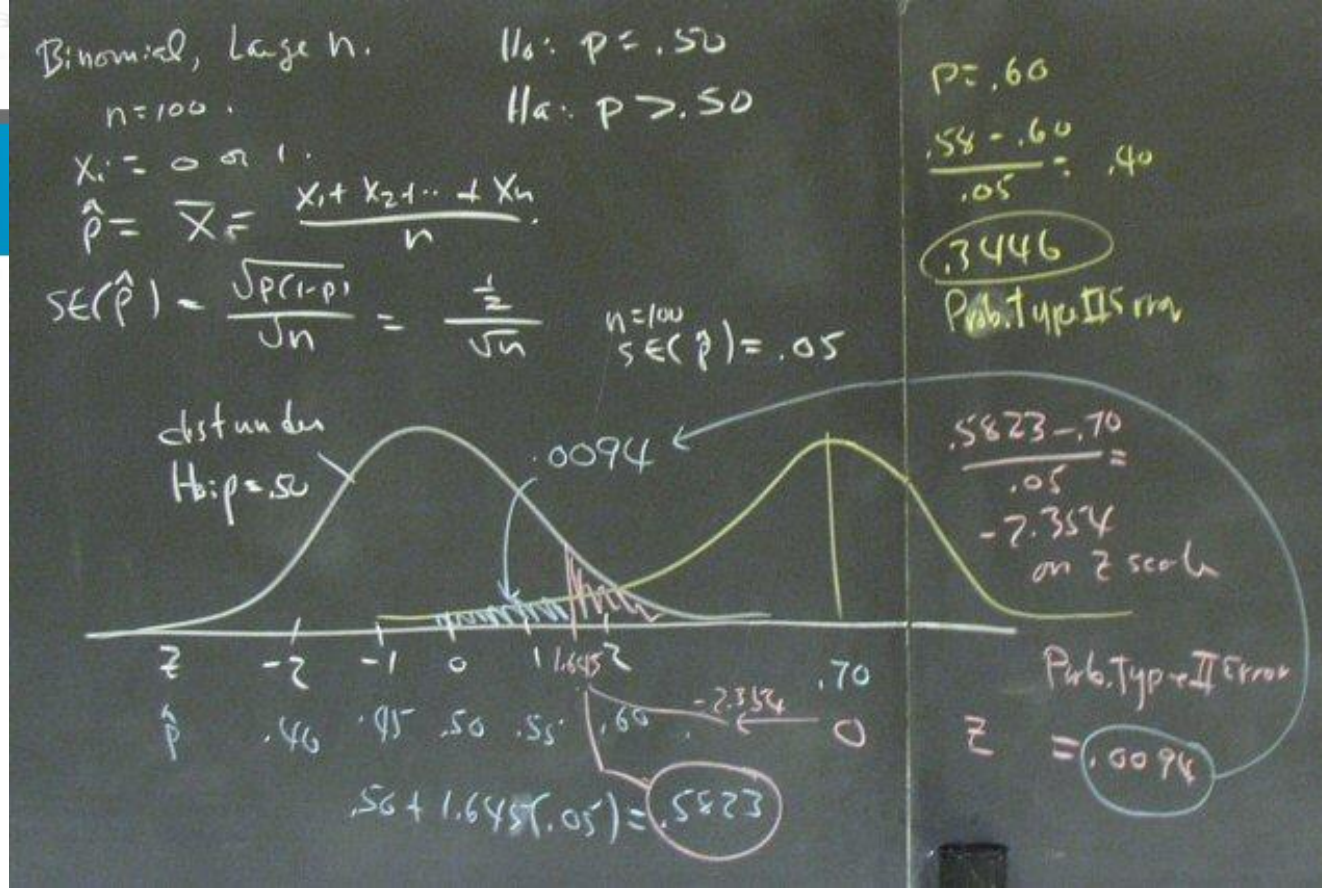

Σ
Sample
Size

What is an appropriate sample size?

- ????
- There is no universal answer! ☹
- It depends on:
  - Research objective
  - The nature of target population
  - Statistics methods
    - Number of variables

Why do we worry about Sample Size and Power?

- Sample size too big
  - too much power wastes money and resources on extra subjects without improving statistical results
- Sample size too small
  - having too little power to detect meaningful differences
    - exposure (treatment) discarded as not important when in fact it is useful
- Improving your research design



Inferential statistics

The intention of statistics

- The intention of the statistical analysis is to answer two questions:
  - Is there a significant relationship/difference/influence between the variables (do we reject the null hypothesis?)
  - If there is connection/difference/impact - how big it is?

Hypotheses

- The null hypothesis (H_0) is the claim of the population parameter
- H_0 - it is assumed that there is no statistically significant difference (change) between the actual value of the population parameter and claimed values
- The null hypothesis is the one that is being tested and the most common objective of testing its rejection

H_0 –the claim of the parameter

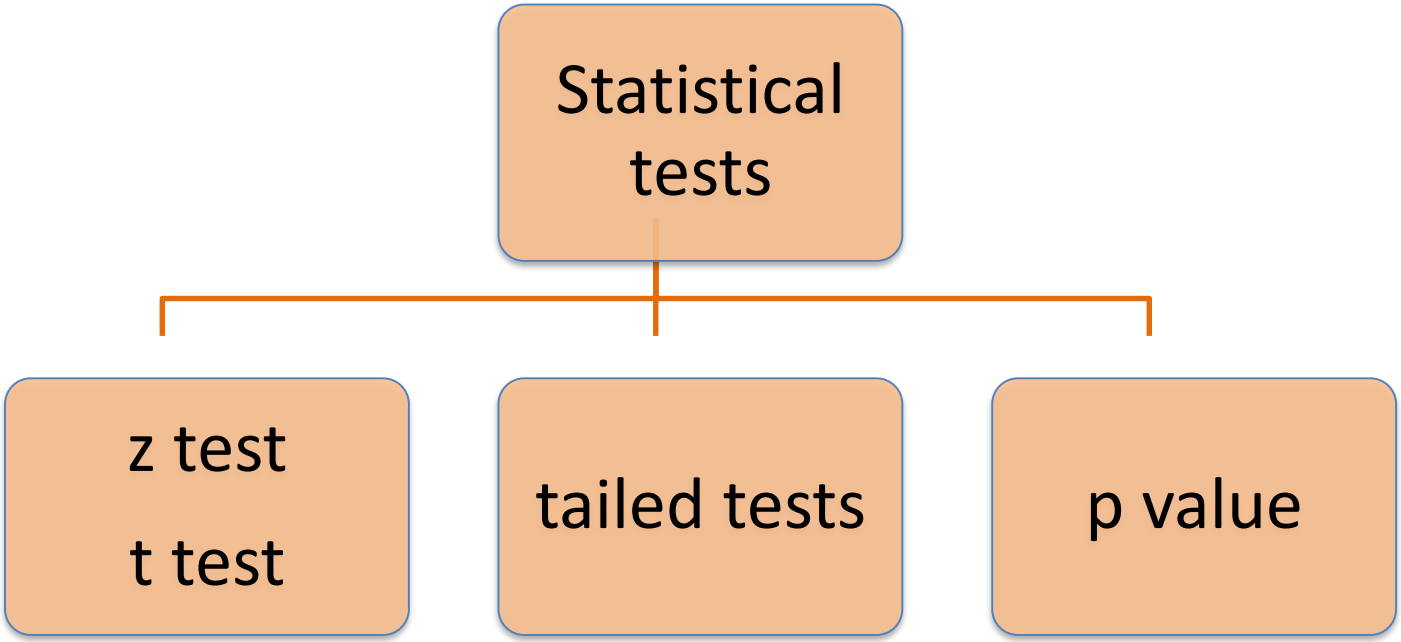
H_1 – alternative hypothesis (claim opposite of H_0)

The null hypothesis can be:

- **Reject H_0** **–**
- **Do not reject H_0** **+**

Statistical tests

Figure: Statistical test for testing null hypothesis



What is the goal of hypothesis testing?

- ???
- To reject H0

Criteria for rejecting a null hypothesis

- Level of Significance (Alpha Level)
  - Traditional levels are .05 or .01
- Region of distribution of sample means defined by alpha level is known as the “critical region”
- No hypothesis is ever “proven”; we just fail to reject null
- When the null is retained, alternatives are also retained.

Possible decisions (errors)

- In the process of testing - four decisions

Table: Likelihood of correct and incorrect decisions for H_0

	H_0 correct	H_0 incorrect
reject H_0	Type I error **The probability of making error α** **Type I error = α**	Right decision **(1- β)** (Power of the test)
fail to reject H_0	Right decision **(1 – α)**	Type II error **The probability of making error** **Type II error = β**

α and β Levels

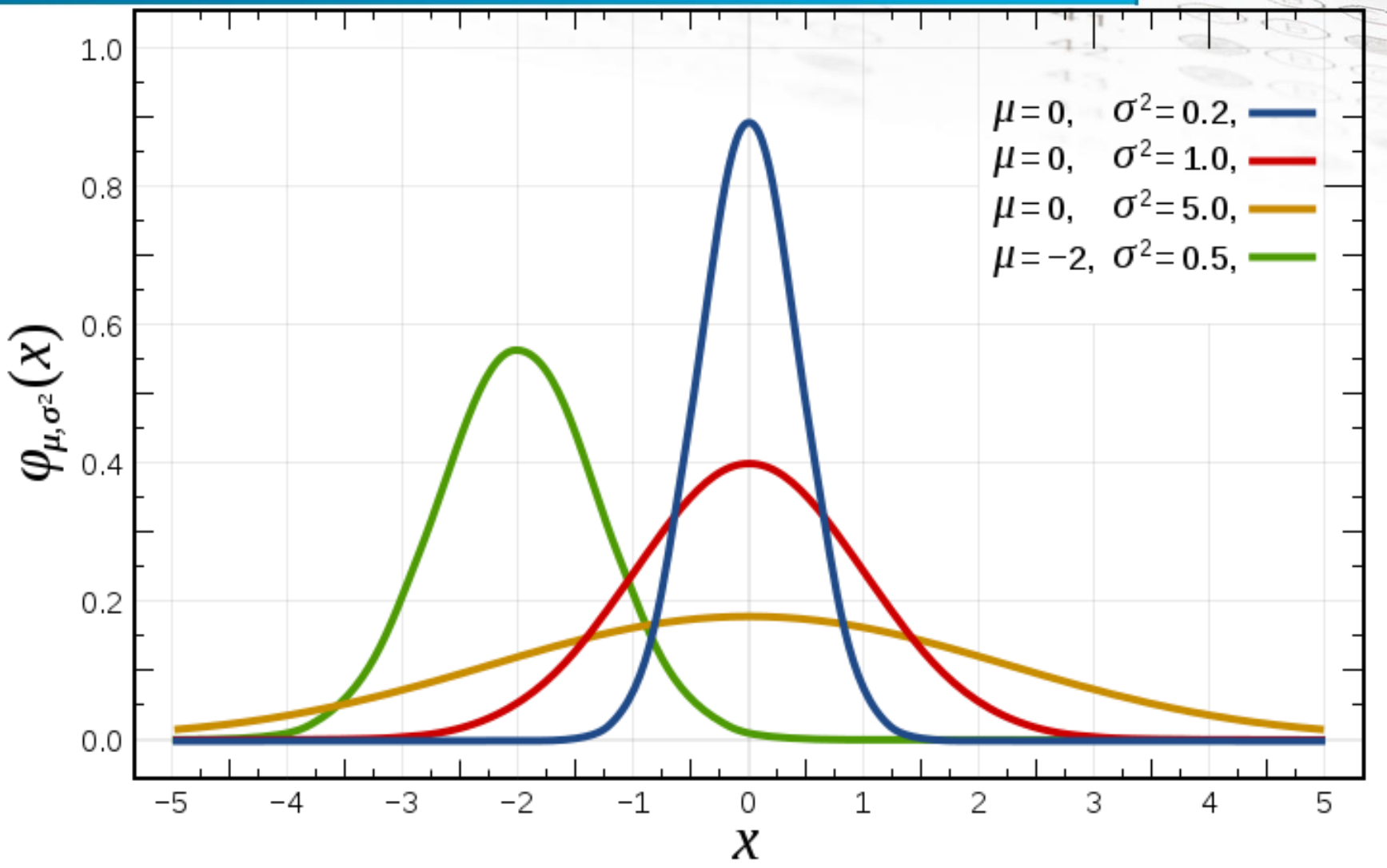
- Usually range from 0.01-.10 (α) and from 0.05-.20 (β)
- Convention α=0.05 and β=0.20
- Use low alpha's to avoid false positives
- Use low beta's to avoid false negatives
- Increased sample size will reduce type I and type II errors

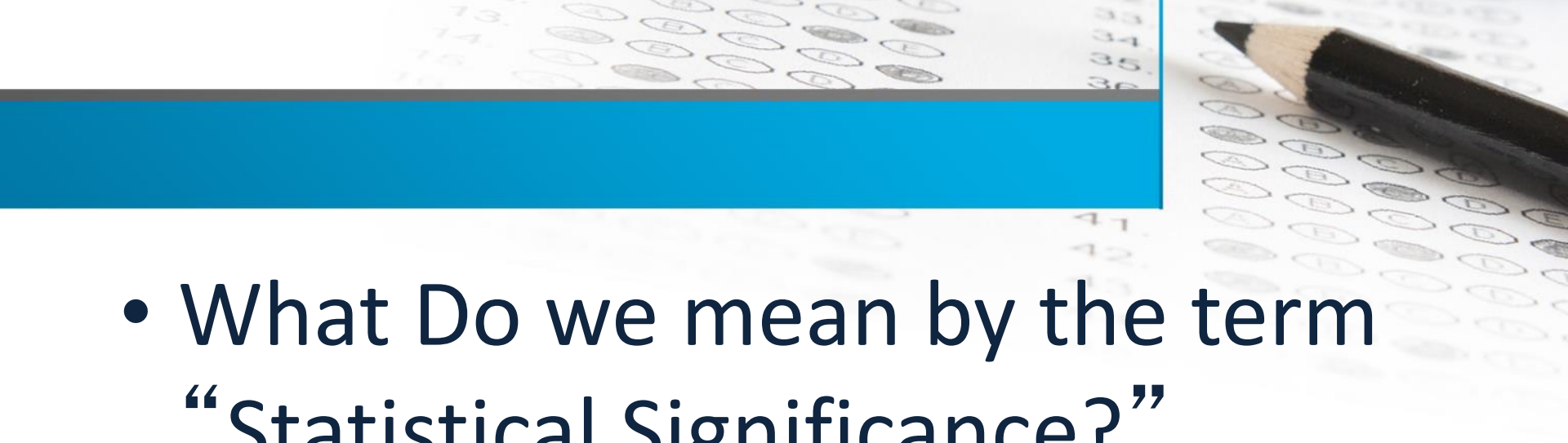
Statistical Power (1- β)

- How sensitive is a test to detecting real effects?
- A powerful test decreases the chances of making a Type II Error
- Ways of Increasing Power:
  - Increase sample size
  - Make alpha level less conservative
  - Use one-tailed versus a two-tailed test
- If β =.20 then power =.80

Normality of distribution

- The central limit theorem
  - Each distribution will be approximately normally distributed
    - If the sample is greater than 30 units
    - And the small sample size (*n*) if the sample selected from the population that depends on the normal distribution

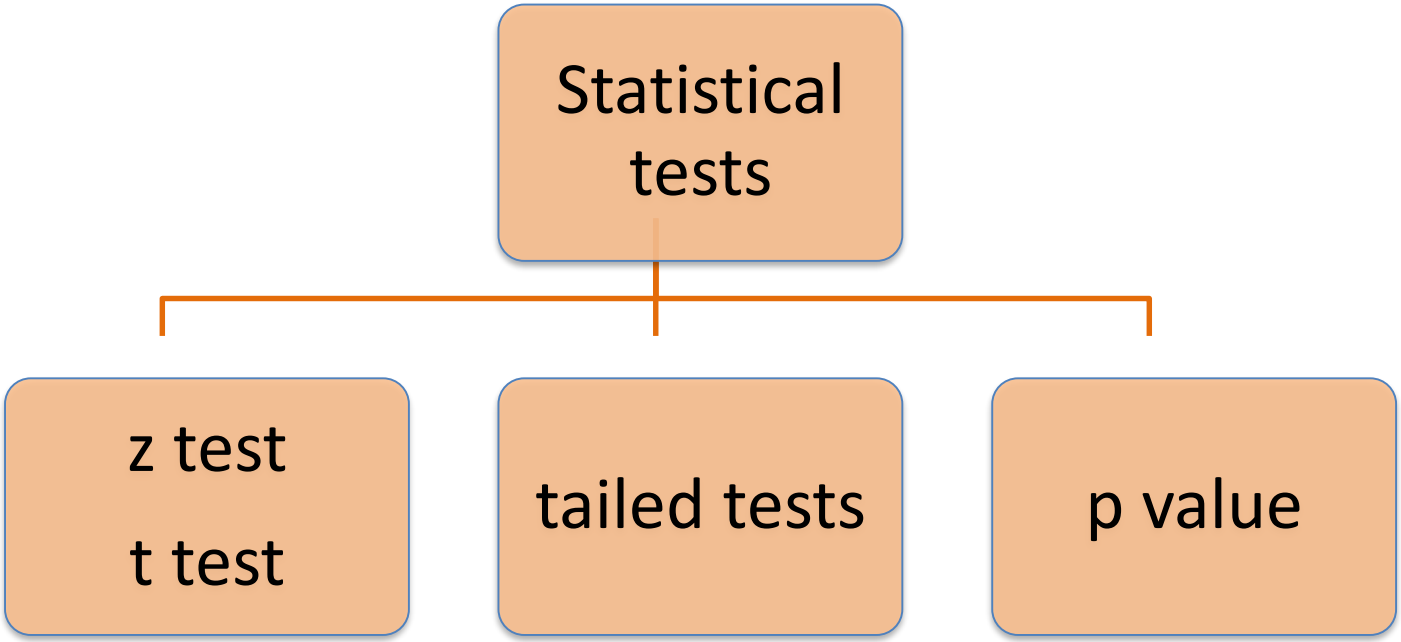

$\mu = 0, \quad \sigma^2 = 0.2,$
$\mu = 0, \quad \sigma^2 = 1.0,$
$\mu = 0, \quad \sigma^2 = 5.0,$
$\mu = -2, \quad \sigma^2 = 0.5,$
$\varphi_{\mu,\sigma^2}(x)$
x
1.0
0.8
0.6
0.4
0.2
0.0
−5
−4
−3
−2
−1
0
1
2
3
4
5



- What Do we mean by the term “Statistical Significance?”

Statistical tests

Figure: Statistical test for testing null hypothesis



P-values

- A measure of a Type I error (random error)
  - Our very important error ☺
- If p=0.05, there is just a 5% chance that an observed association in your sample is due to random error

Using the P-Value

- Reject H_o if $p-value \leq \alpha$

- Fail to Reject H_o if $p-value > \alpha$

- The smaller the *P*-value, the stronger the evidence against H_0.


THIS IS WHAT
P<.05
FEELS LIKE


p value
p<0.05 Reject H0
p>0.05 Don't reject H0
0.05 Magical number in statistics :)

P VALUE

$p < 0.05$ statistically significant difference

$p > 0.05$ no statistically significant difference

Sample size

Sample size: The number of elements in the obtained sample.

Factors that decide about the sample size:

1. The confidence you need in data.
2. The **margin of error** that you can tolerate.
3. The type of analysis that you are going to undertake.
4. The size of total population.

Sample size

- General rule - as large as possible to increase the representativeness of the sample

- Increased size decreases **sampling error**
- As the number of **variables** studied increases, the sample size also needs to increase in order to detect significant relationships or differences

Why do we worry about Sample Size and Power?

- Sample size too big
  - too much power wastes money and resources on extra subjects without improving statistical results
- Sample size too small
  - having too little power to detect meaningful differences
    - exposure (treatment) discarded as not important when in fact it is useful
- Improving your research design

- **Optimal / Ideal Sample SIZE!!!!**

- Sample size table

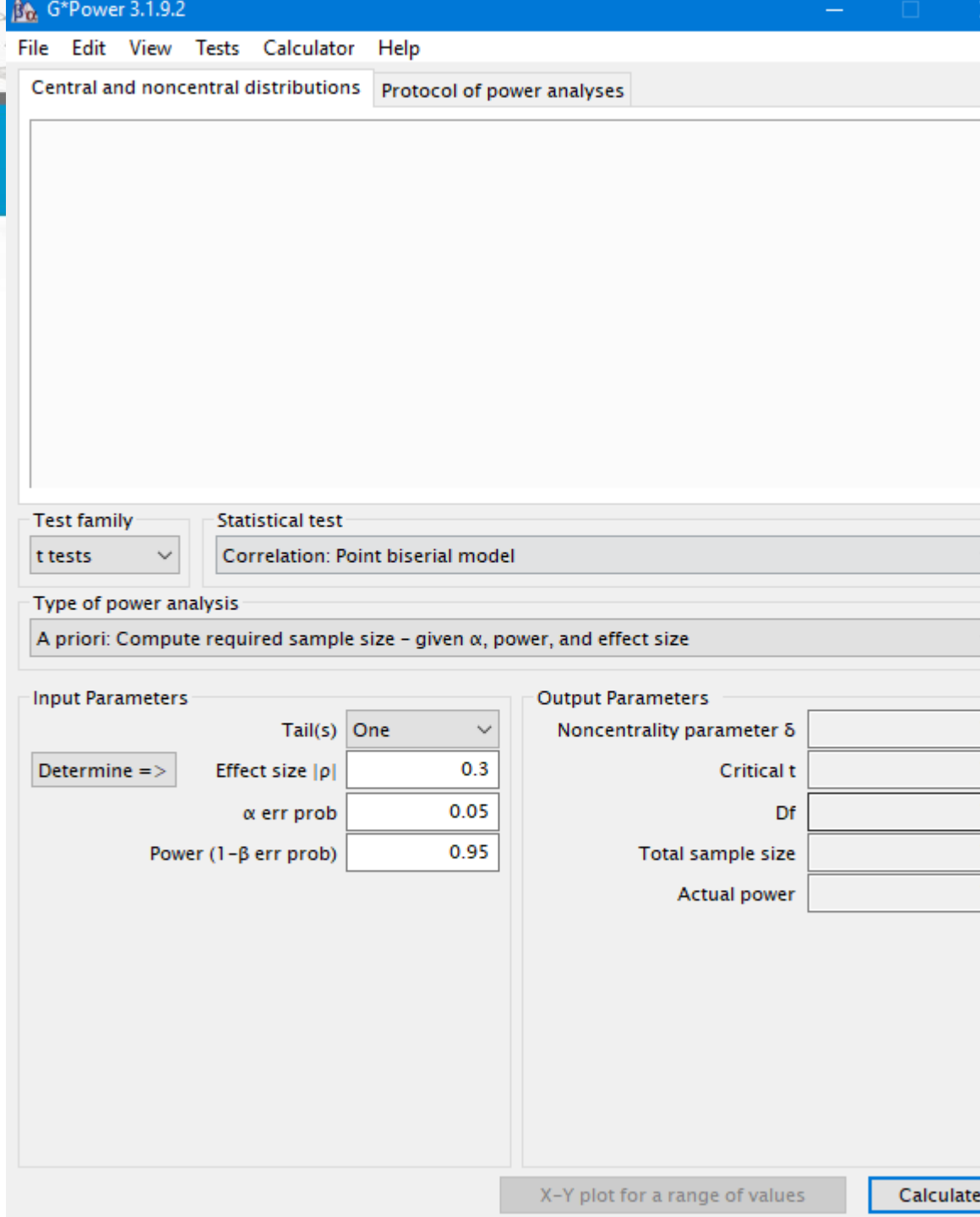
Sample Size and Power Software

- EpiInfo
  - Programs→Statcalc→Sample size and Power
  - User-friendly; easily accessible
- PASS, Power and Precision

- G*power
  - Free
  - Calculating sample size
  - Calculating statistical power

G*power3

- Some examples

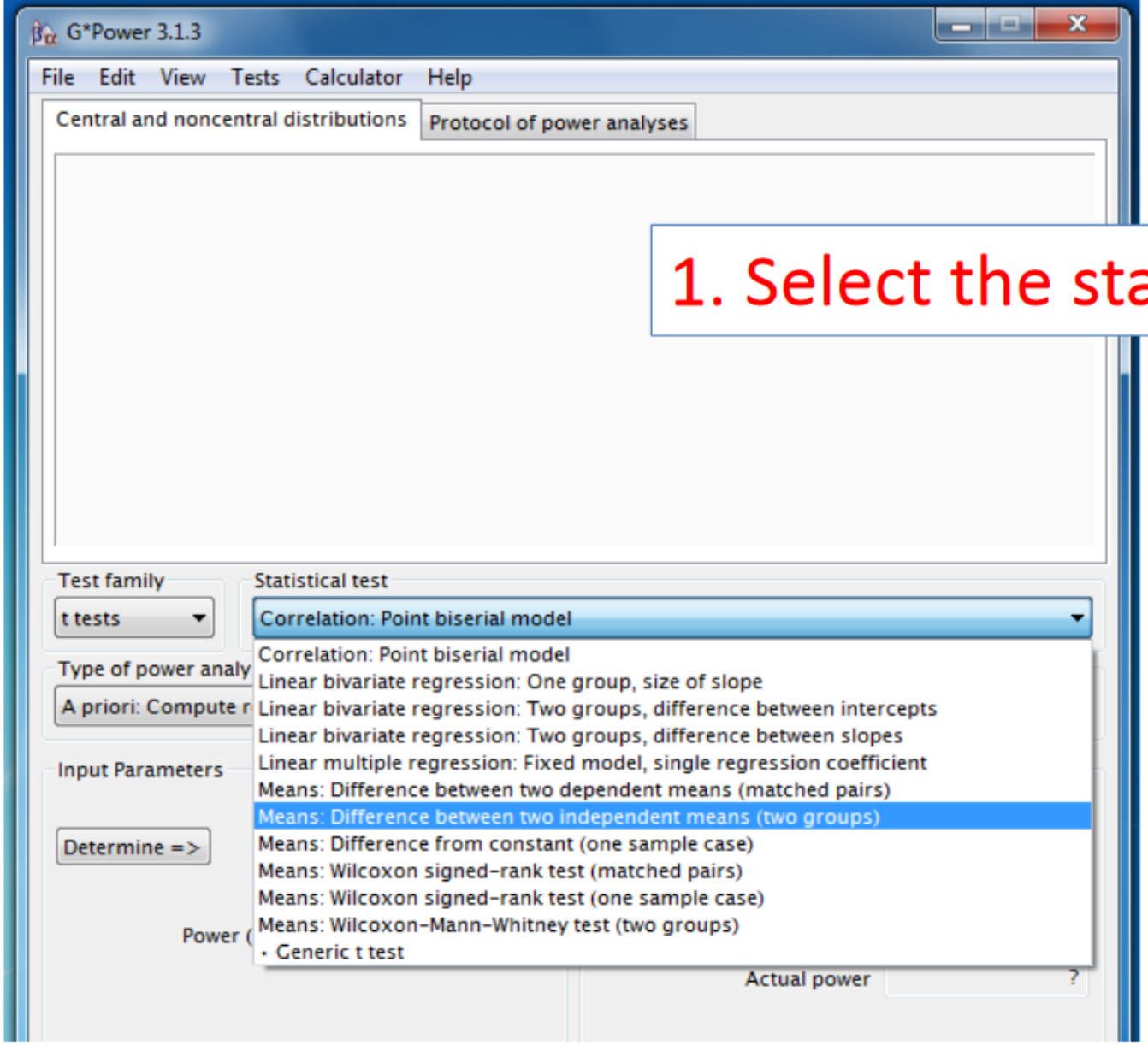


Before calculating

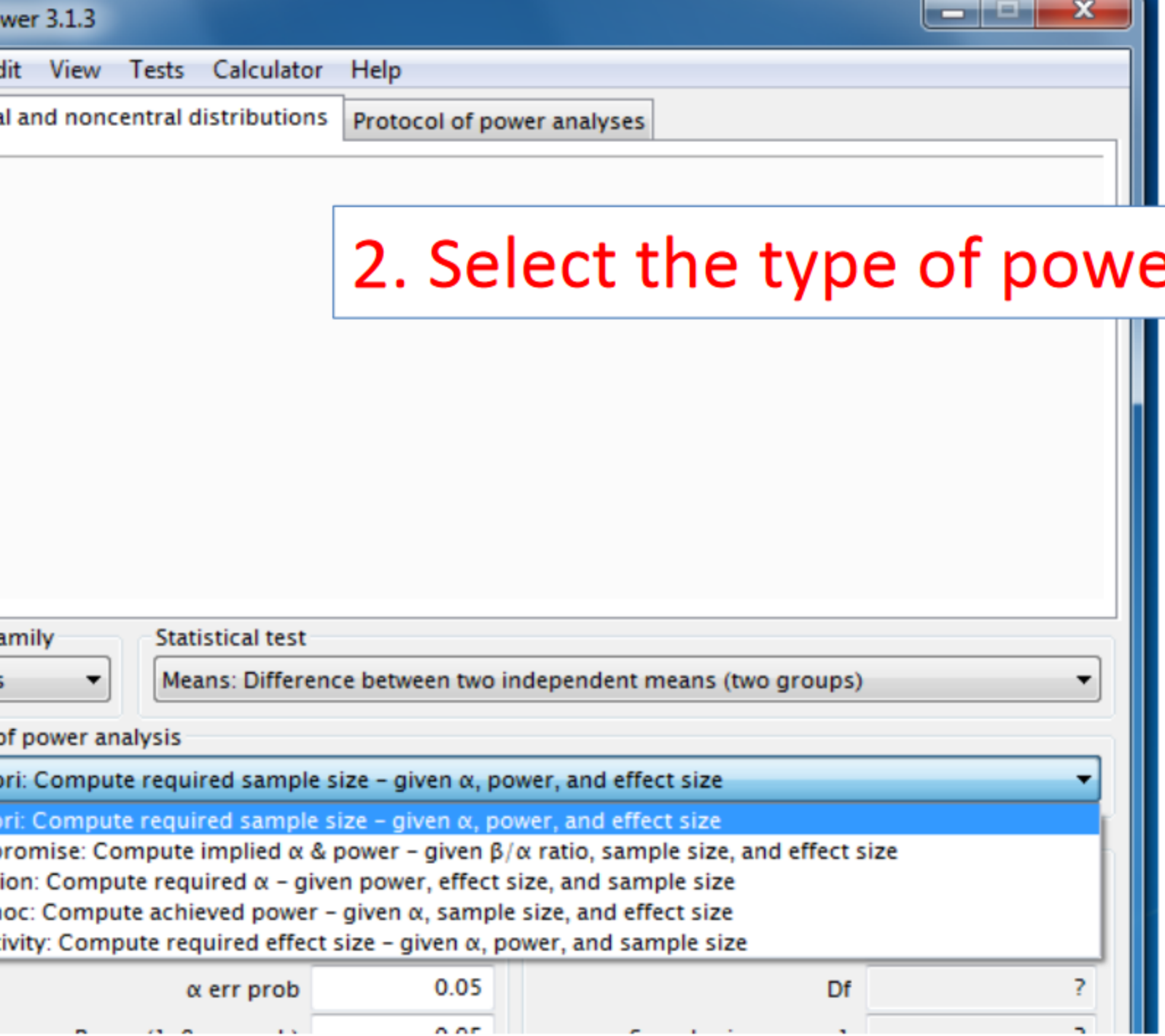
- Understand which statistical method you plan to use
- Do you know the number of predictive variables?

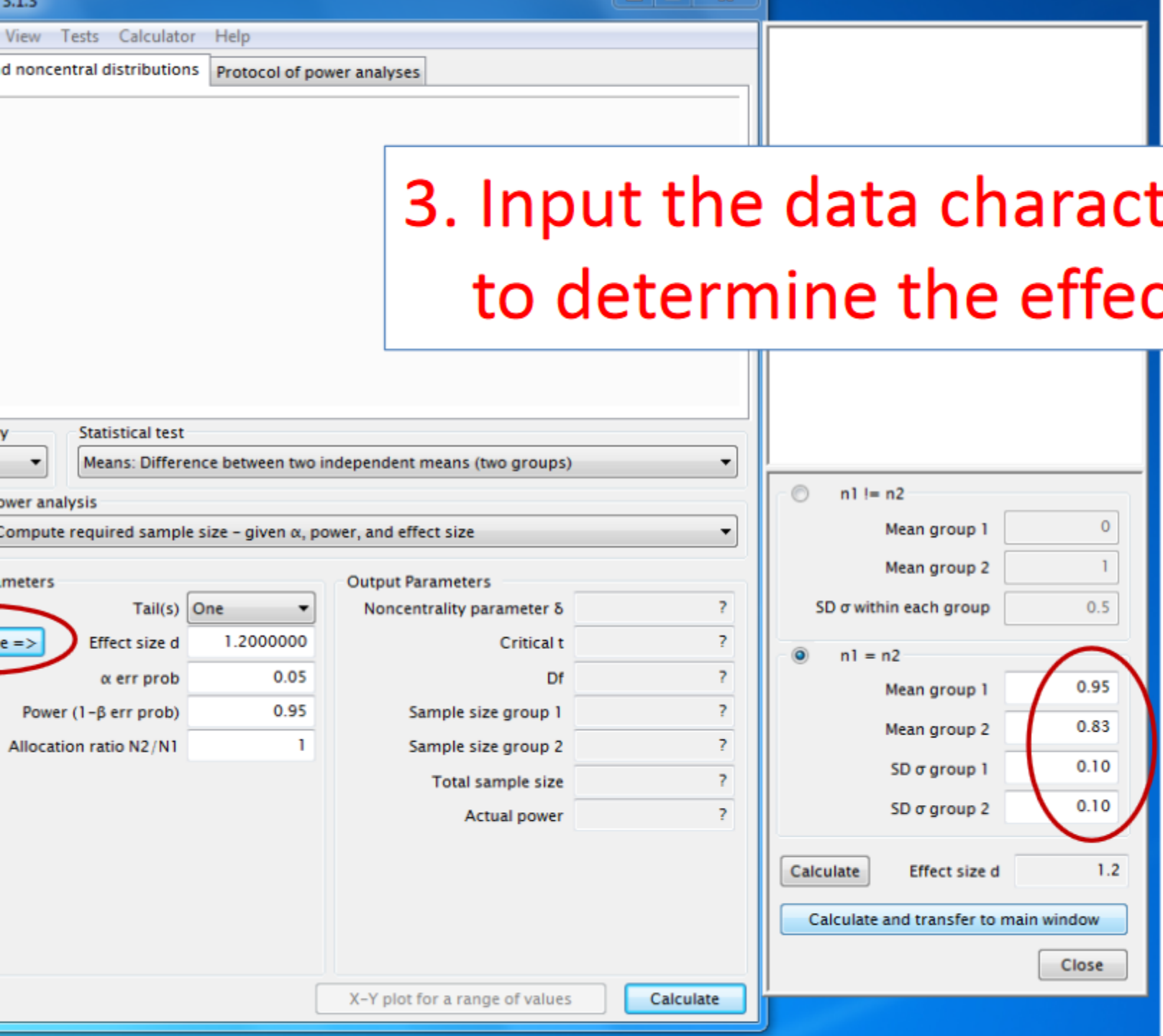
G* Power

- Correlations & regressions (univariate, multiple variate, logistic)
- Means (one, two, many groups, un/paired, non-parametric)
- Proportions (one, two groups, un/paired)
- Variances (one, two groups)

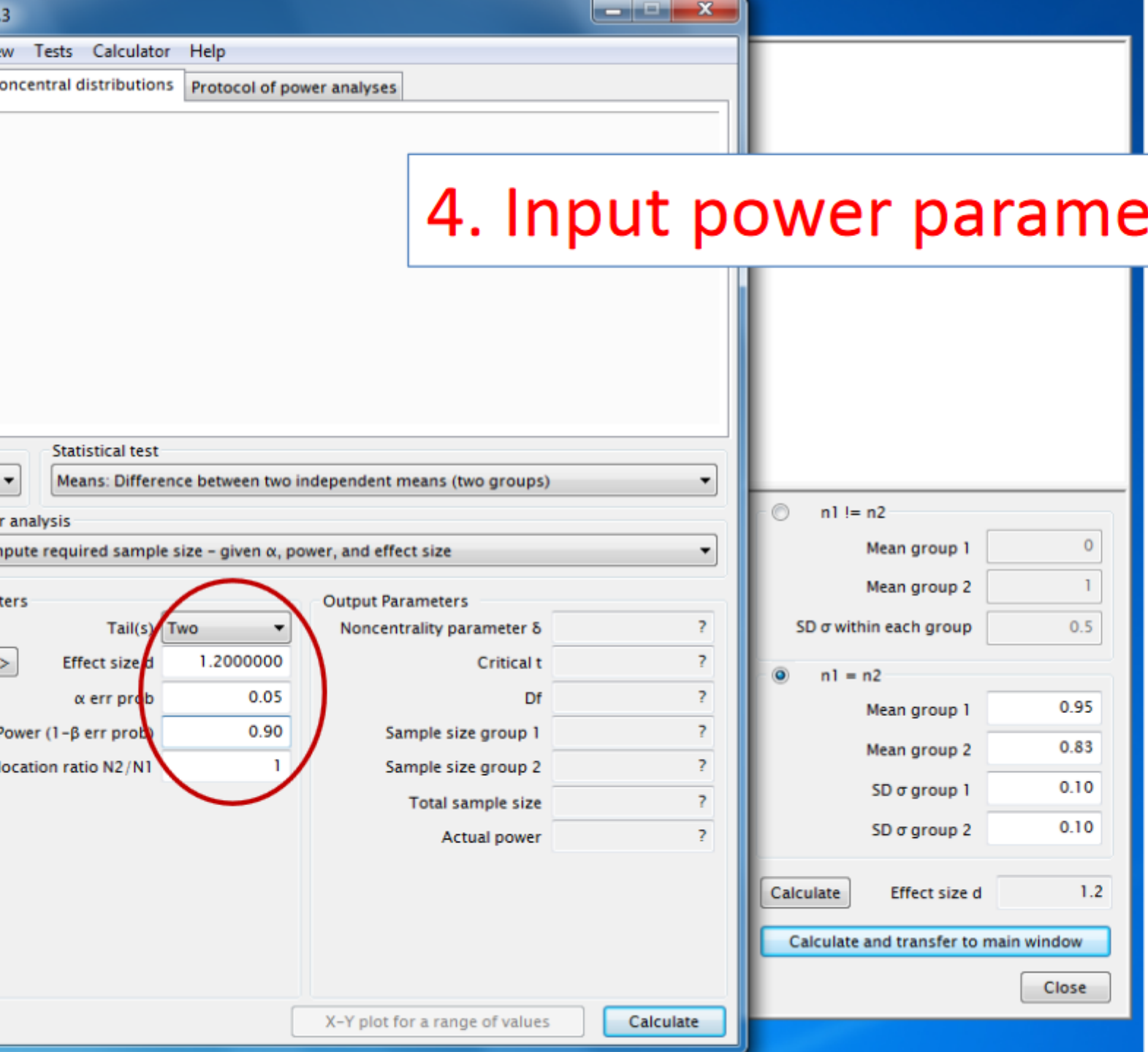


1. Select the statistical test

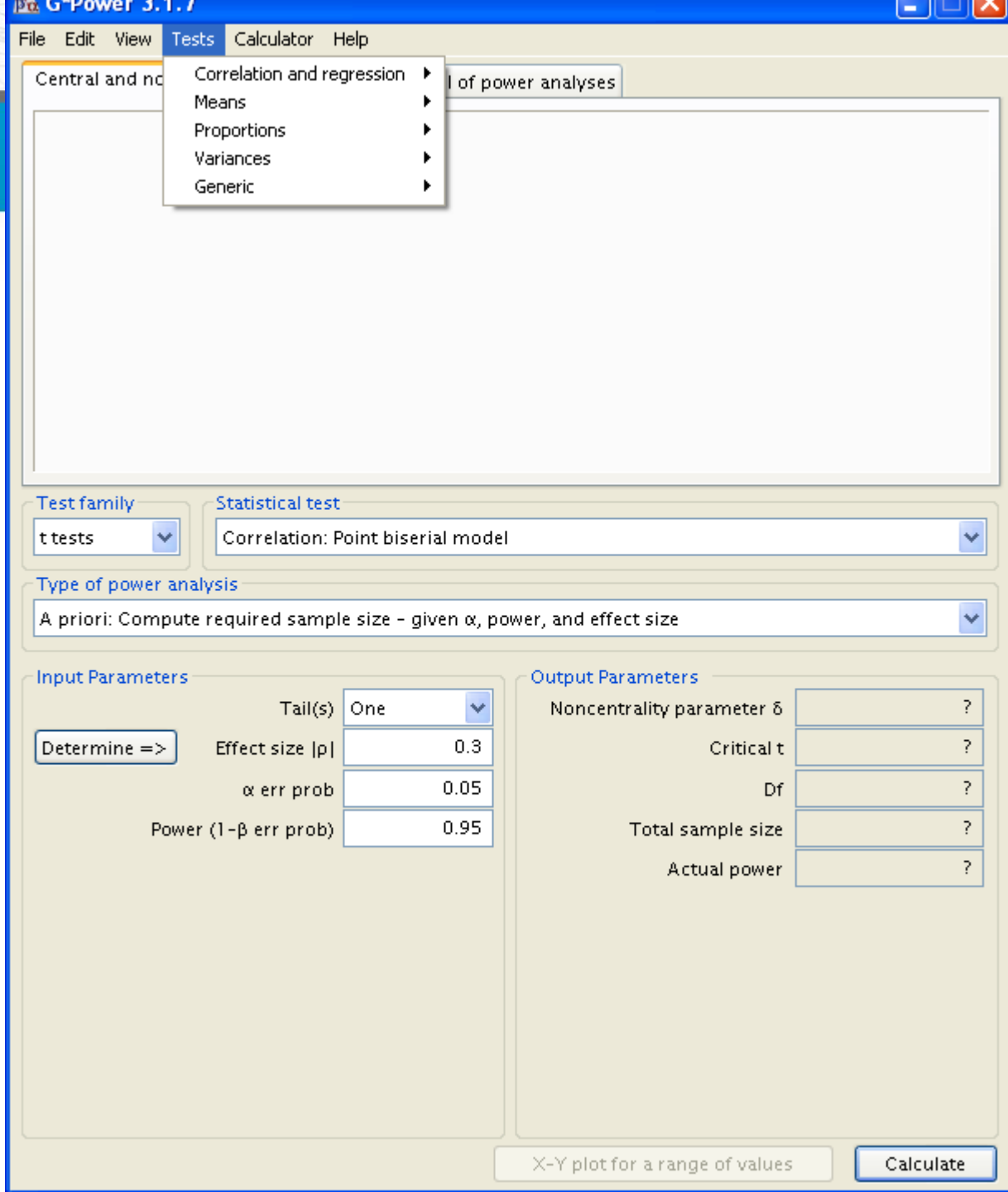

wer 3.1.3
dit View Tests Calculator Help
al and noncentral distributions
Protocol of power analyses
2. Select the type of power analys
amily
Statistical test
Means: Difference between two independent means (two groups)
of power analysis
ori: Compute required sample size – given α, power, and effect size
ori: Compute required sample size – given α, power, and effect size
promise: Compute implied α & power – given β/α ratio, sample size, and effect size
ion: Compute required α – given power, effect size, and sample size
oc: Compute achieved power – given α, sample size, and effect size
tivity: Compute required effect size – given α, power, and sample size
α err prob
0.05
Df
?


View Tests Calculator Help
d noncentral distributions
Protocol of power analyses
3. Input the data characteristics to determine the effect size
Statistical test
Means: Difference between two independent means (two groups)
ower analysis
Compute required sample size – given α, power, and effect size
meters
Tail(s) One
e =>
Effect size d 1.2000000
α err prob 0.05
Power (1–β err prob) 0.95
Allocation ratio N2/N1 1
Output Parameters
Noncentrality parameter δ ?
Critical t ?
Df ?
Sample size group 1 ?
Sample size group 2 ?
Total sample size ?
Actual power ?
X-Y plot for a range of values
Calculate
n1 != n2
Mean group 1 0
Mean group 2 1
SD σ within each group 0.5
n1 = n2
Mean group 1 0.95
Mean group 2 0.83
SD σ group 1 0.10
SD σ group 2 0.10
Calculate
Effect size d 1.2
Calculate and transfer to main window
Close

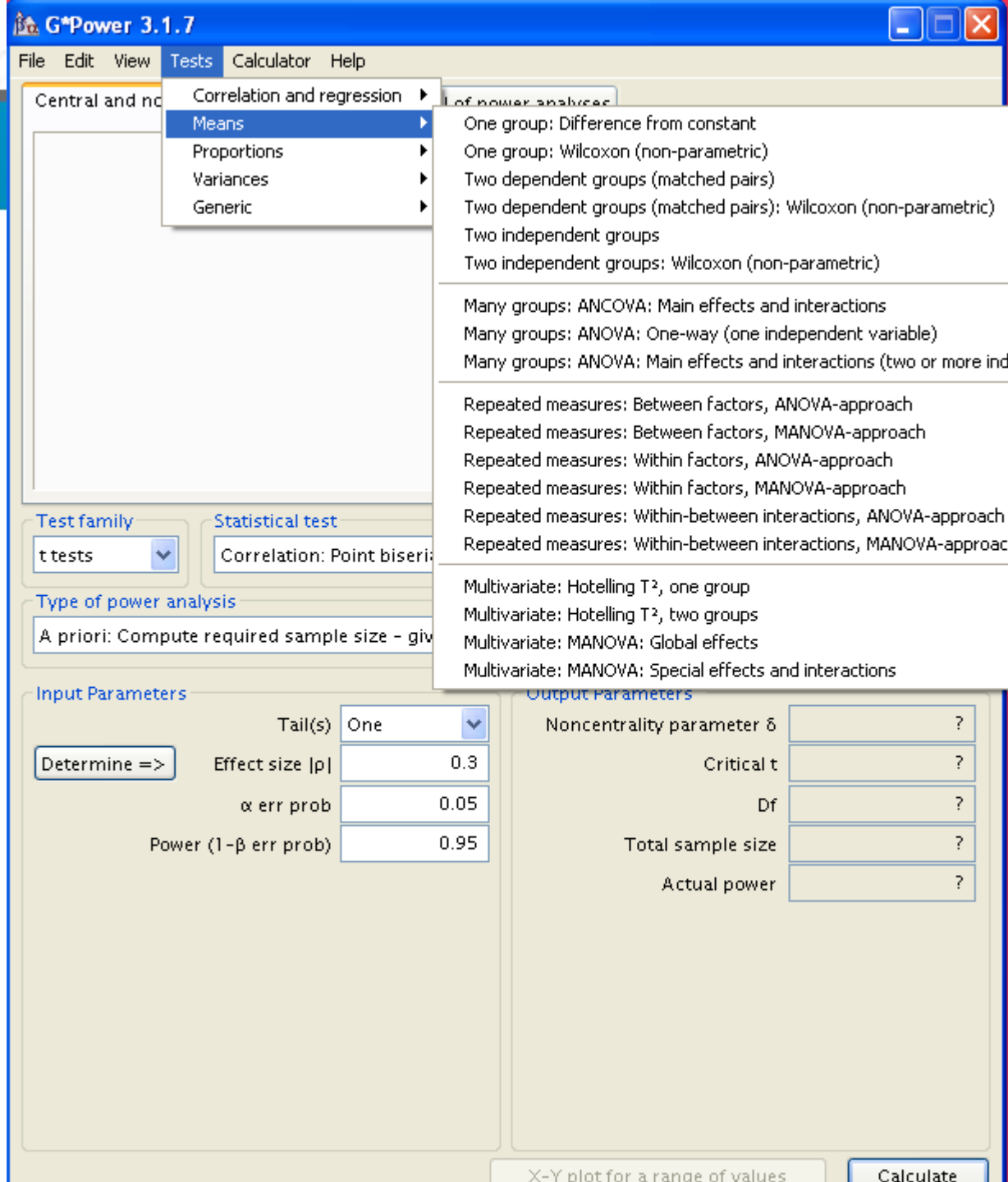
4. Input power parameters



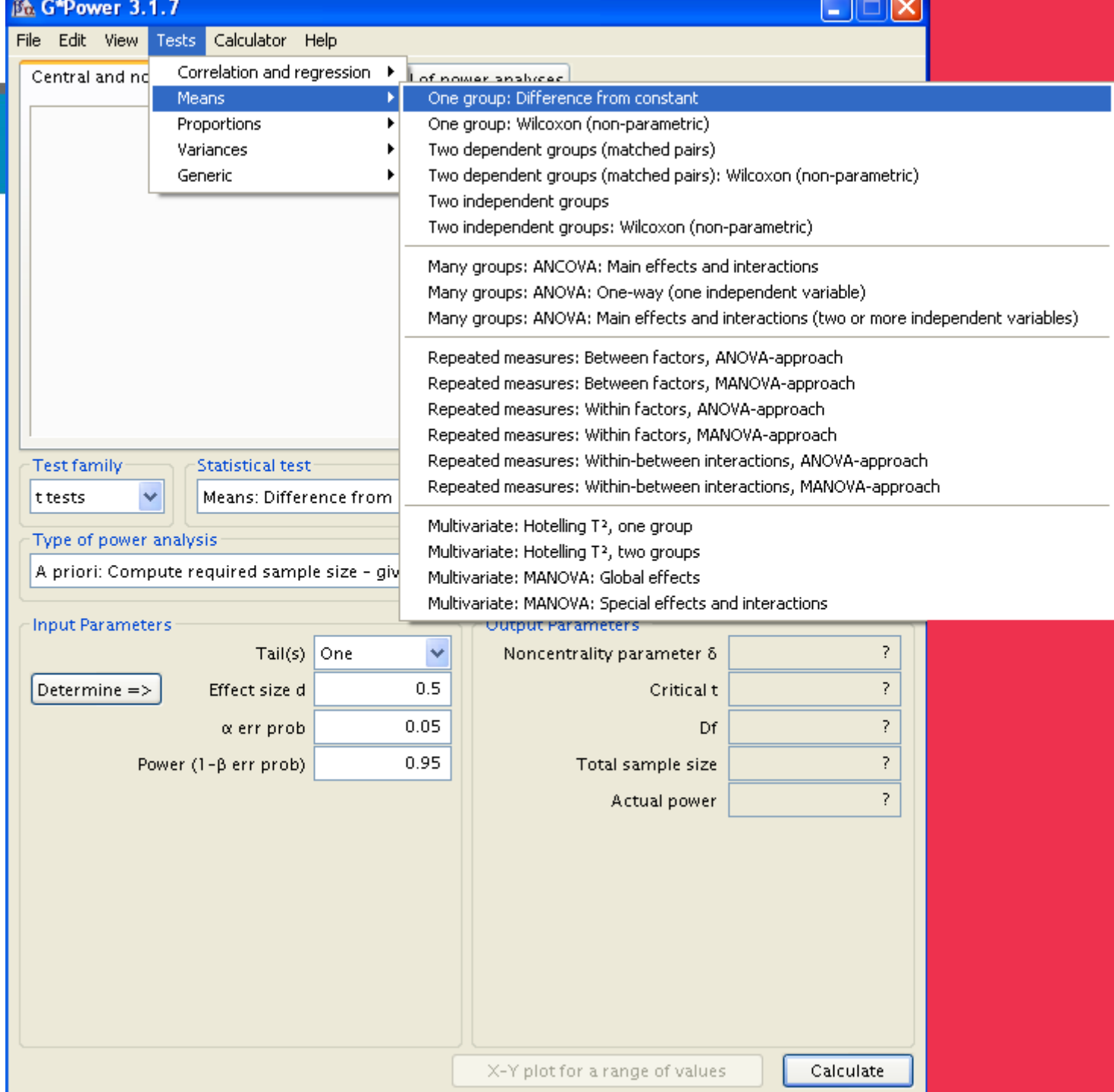
t-test

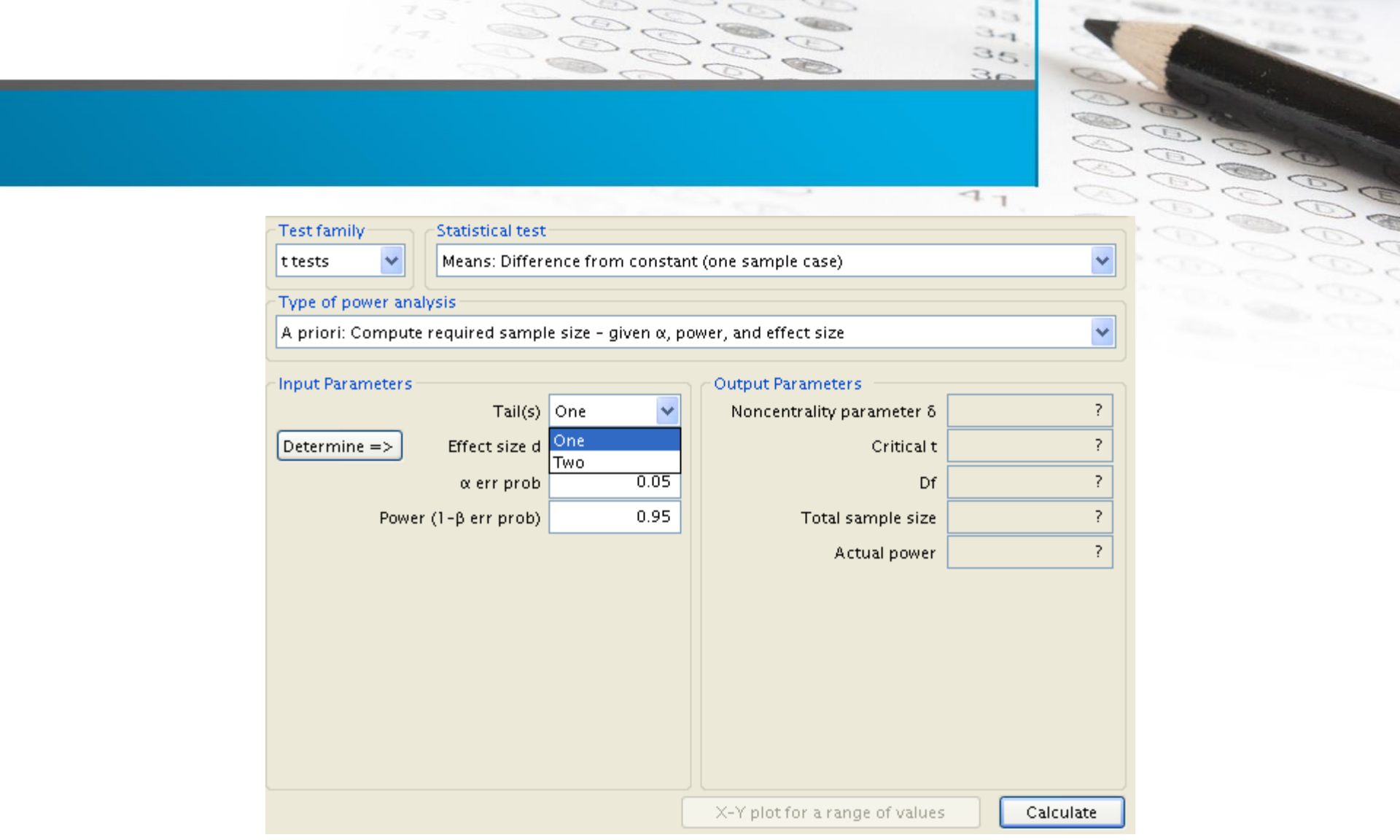

G*Power 3.1.7
File
Edit
View
Tests
Calculator
Help
Correlation and regression
Means
Proportions
Variances
Generic
of power analyses
Test family
t tests
Statistical test
Correlation: Point biserial model
Type of power analysis
A priori: Compute required sample size – given α, power, and effect size
Input Parameters
Tail(s)
One
Determine =>
Effect size |ρ|
0.3
α err prob
0.05
Power (1-β err prob)
0.95
Output Parameters
Noncentrality parameter δ
?
Critical t
?
Df
?
Total sample size
?
Actual power
?
X-Y plot for a range of values
Calculate

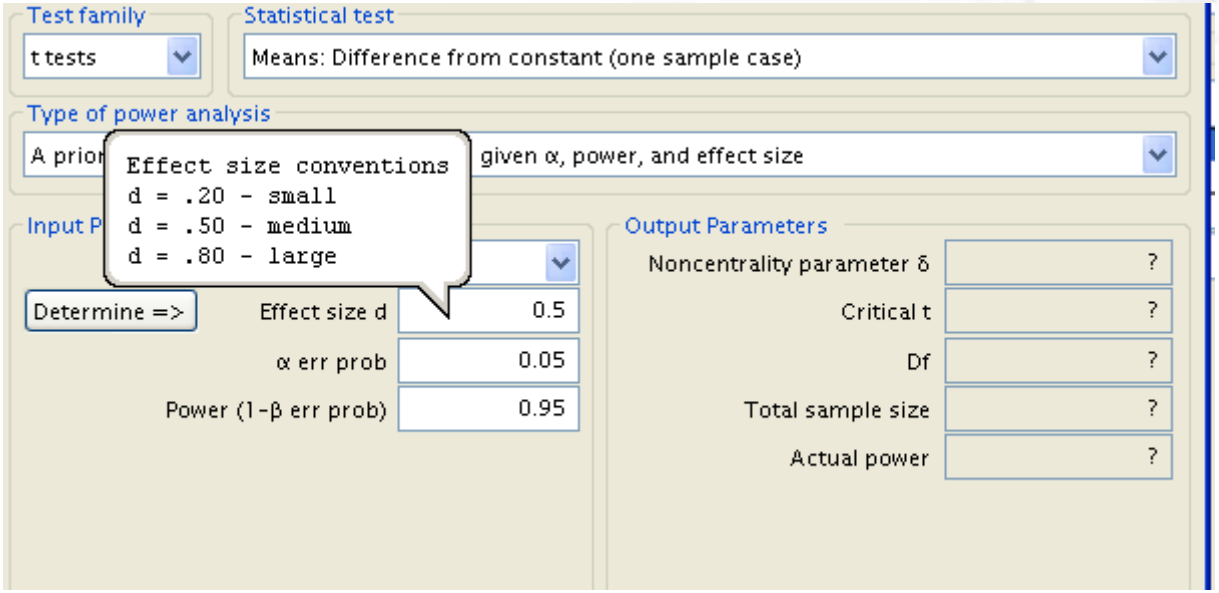
t-test

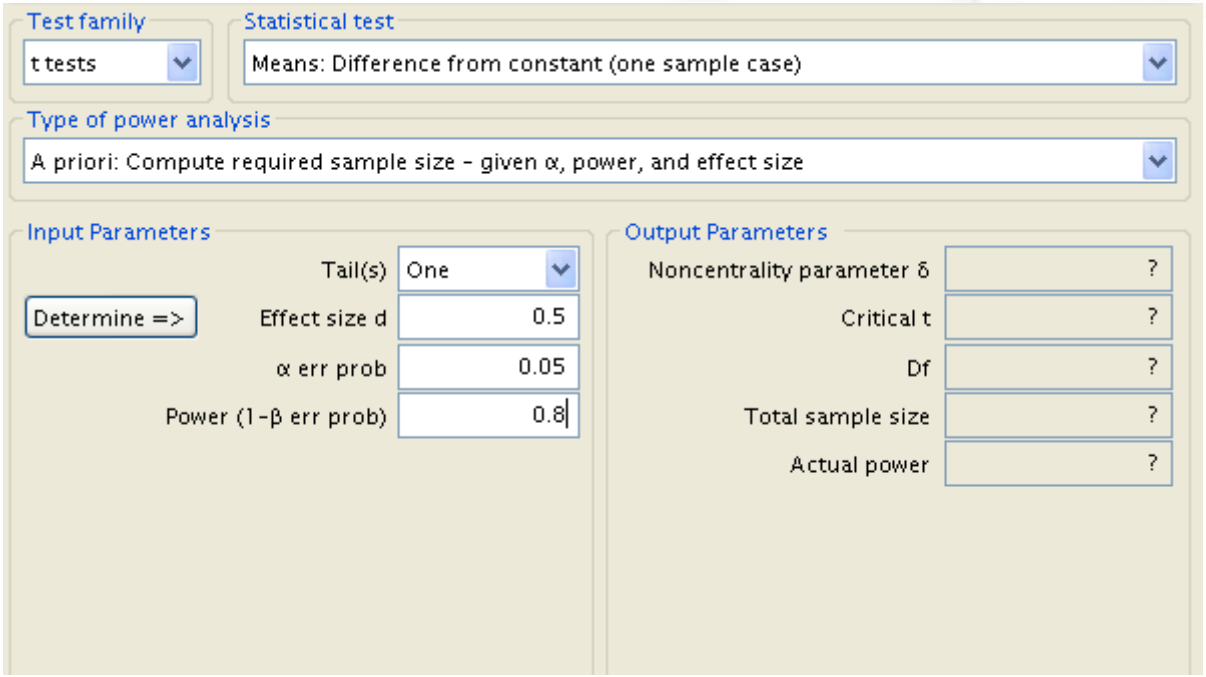

G*Power 3.1.7
File Edit View Tests Calculator Help
Correlation and regression
Means
Proportions
Variances
Generic
One group: Difference from constant
One group: Wilcoxon (non-parametric)
Two dependent groups (matched pairs)
Two dependent groups (matched pairs): Wilcoxon (non-parametric)
Two independent groups
Two independent groups: Wilcoxon (non-parametric)
Many groups: ANCOVA: Main effects and interactions
Many groups: ANOVA: One-way (one independent variable)
Many groups: ANOVA: Main effects and interactions (two or more independent variable
Repeated measures: Between factors, ANOVA-approach
Repeated measures: Between factors, MANOVA-approach
Repeated measures: Within factors, ANOVA-approach
Repeated measures: Within factors, MANOVA-approach
Repeated measures: Within-between interactions, ANOVA-approach
Repeated measures: Within-between interactions, MANOVA-approach
Multivariate: Hotelling T², one group
Multivariate: Hotelling T², two groups
Multivariate: MANOVA: Global effects
Multivariate: MANOVA: Special effects and interactions
Test family
t tests
Statistical test
Type of power analysis
A priori: Compute required sample size – giv
Input Parameters
Tail(s) One
Determine =>
Effect size |ρ| 0.3
α err prob 0.05
Power (1-β err prob) 0.95
Output Parameters
Noncentrality parameter δ ?
Critical t ?
Df ?
Total sample size ?
Actual power ?
X-Y plot for a range of values
Calculate

t-test


G*Power 3.1.7
File Edit View Tests Calculator Help
Correlation and regression
Means
Proportions
Variances
Generic
One group: Difference from constant
One group: Wilcoxon (non-parametric)
Two dependent groups (matched pairs)
Two dependent groups (matched pairs): Wilcoxon (non-parametric)
Two independent groups
Two independent groups: Wilcoxon (non-parametric)
Many groups: ANCOVA: Main effects and interactions
Many groups: ANOVA: One-way (one independent variable)
Many groups: ANOVA: Main effects and interactions (two or more independent variables)
Repeated measures: Between factors, ANOVA-approach
Repeated measures: Between factors, MANOVA-approach
Repeated measures: Within factors, ANOVA-approach
Repeated measures: Within factors, MANOVA-approach
Repeated measures: Within-between interactions, ANOVA-approach
Repeated measures: Within-between interactions, MANOVA-approach
Multivariate: Hotelling T², one group
Multivariate: Hotelling T², two groups
Multivariate: MANOVA: Global effects
Multivariate: MANOVA: Special effects and interactions
Test family
t tests
Statistical test
Means: Difference from
Type of power analysis
A priori: Compute required sample size - giv
Input Parameters
Tail(s)
One
Determine =>
Effect size d
0.5
α err prob
0.05
Power (1-β err prob)
0.95
Output Parameters
Noncentrality parameter δ
Critical t
Df
Total sample size
Actual power
X-Y plot for a range of values
Calculate


Test family
t tests
Statistical test
Means: Difference from constant (one sample case)
Type of power analysis
A priori: Compute required sample size – given α, power, and effect size
Input Parameters
Tail(s)
One
Two
Determine =>
Effect size d
α err prob
0.05
Power (1–β err prob)
0.95
Output Parameters
Noncentrality parameter δ
?
Critical t
?
Df
?
Total sample size
?
Actual power
?
X-Y plot for a range of values
Calculate

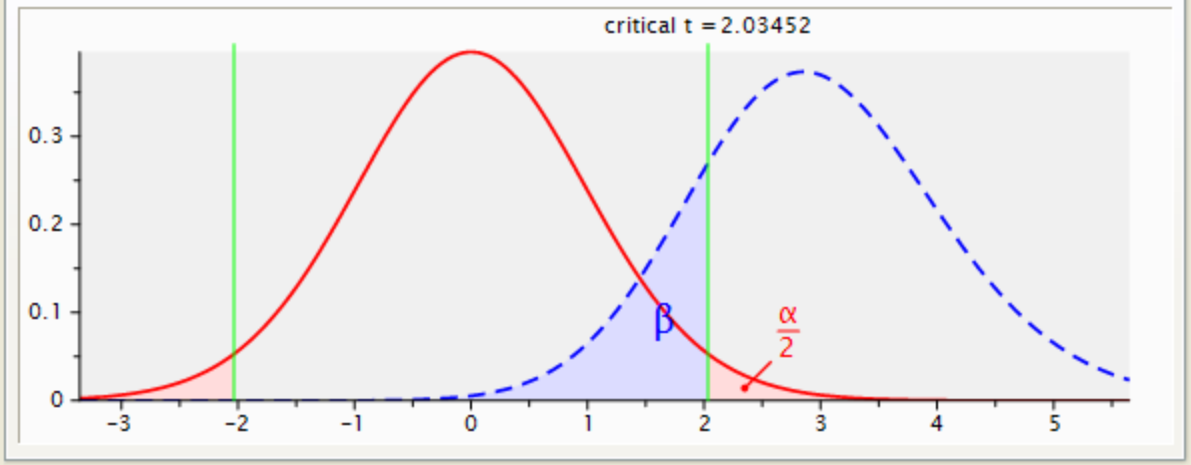

Test family
t tests
Statistical test
Means: Difference from constant (one sample case)
Type of power analysis
A prior
given α, power, and effect size
Effect size conventions
d = .20 - small
d = .50 - medium
d = .80 - large
Input P
Output Parameters
Determine =>
Effect size d
0.5
α err prob
0.05
Power (1-β err prob)
0.95
Noncentrality parameter δ
?
Critical t
?
Df
?
Total sample size
?
Actual power
?


Test family
t tests
Statistical test
Means: Difference from constant (one sample case)
Type of power analysis
A priori: Compute required sample size – given α, power, and effect size
Input Parameters
Tail(s)
One
Determine =>
Effect size d
0.5
α err prob
0.05
Power (1-β err prob)
0.8
Output Parameters
Noncentrality parameter δ
?
Critical t
?
Df
?
Total sample size
?
Actual power
?

G*Power 3.1.7

File Edit View Tests Calculator Help

Central and noncentral distributions | Protocol of power analyses



Test family: t tests

Statistical test: Means: Difference from constant (one sample case)

Type of power analysis: A priori: Compute required sample size – given α, power, and effect size

Input Parameters

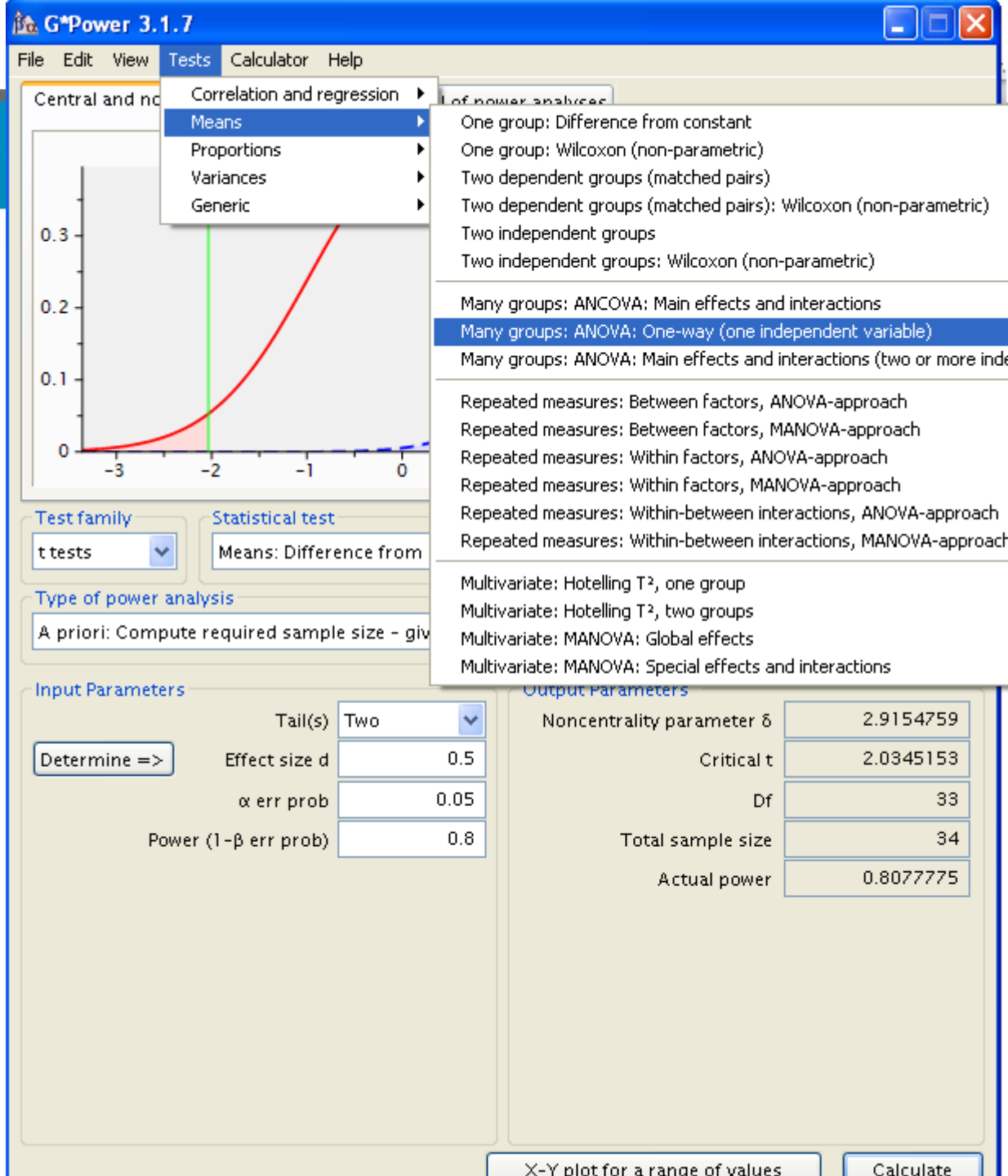
Tail(s)	Two
Determine => Effect size d	0.5
α err prob	0.05
Power (1-β err prob)	0.8

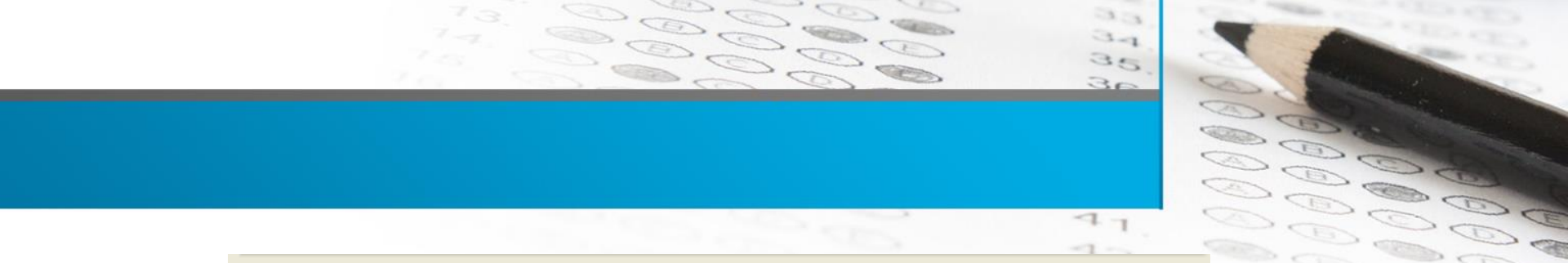
Output Parameters

Noncentrality parameter δ	2.9154759
Critical t	2.0345153
Df	33
Total sample size	34
Actual power	0.8077775

X-Y plot for a range of values | Calculate

ANOVA


G*Power 3.1.7
File Edit View Tests Calculator Help
Correlation and regression
Means
Proportions
Variances
Generic
One group: Difference from constant
One group: Wilcoxon (non-parametric)
Two dependent groups (matched pairs)
Two dependent groups (matched pairs): Wilcoxon (non-parametric)
Two independent groups
Two independent groups: Wilcoxon (non-parametric)
Many groups: ANCOVA: Main effects and interactions
Many groups: ANOVA: One-way (one independent variable)
Many groups: ANOVA: Main effects and interactions (two or more independent variables)
Repeated measures: Between factors, ANOVA-approach
Repeated measures: Between factors, MANOVA-approach
Repeated measures: Within factors, ANOVA-approach
Repeated measures: Within factors, MANOVA-approach
Repeated measures: Within-between interactions, ANOVA-approach
Repeated measures: Within-between interactions, MANOVA-approach
Multivariate: Hotelling T², one group
Multivariate: Hotelling T², two groups
Multivariate: MANOVA: Global effects
Multivariate: MANOVA: Special effects and interactions
Test family
t tests
Statistical test
Means: Difference from
Type of power analysis
A priori: Compute required sample size – giv
Input Parameters
Tail(s) Two
Determine =>
Effect size d 0.5
α err prob 0.05
Power (1-β err prob) 0.8
Output Parameters
Noncentrality parameter δ 2.9154759
Critical t 2.0345153
Df 33
Total sample size 34
Actual power 0.8077775
X-Y plot for a range of values
Calculate



Test family: F tests

Statistical test: ANOVA: Fixed effects, omnibus, one-way

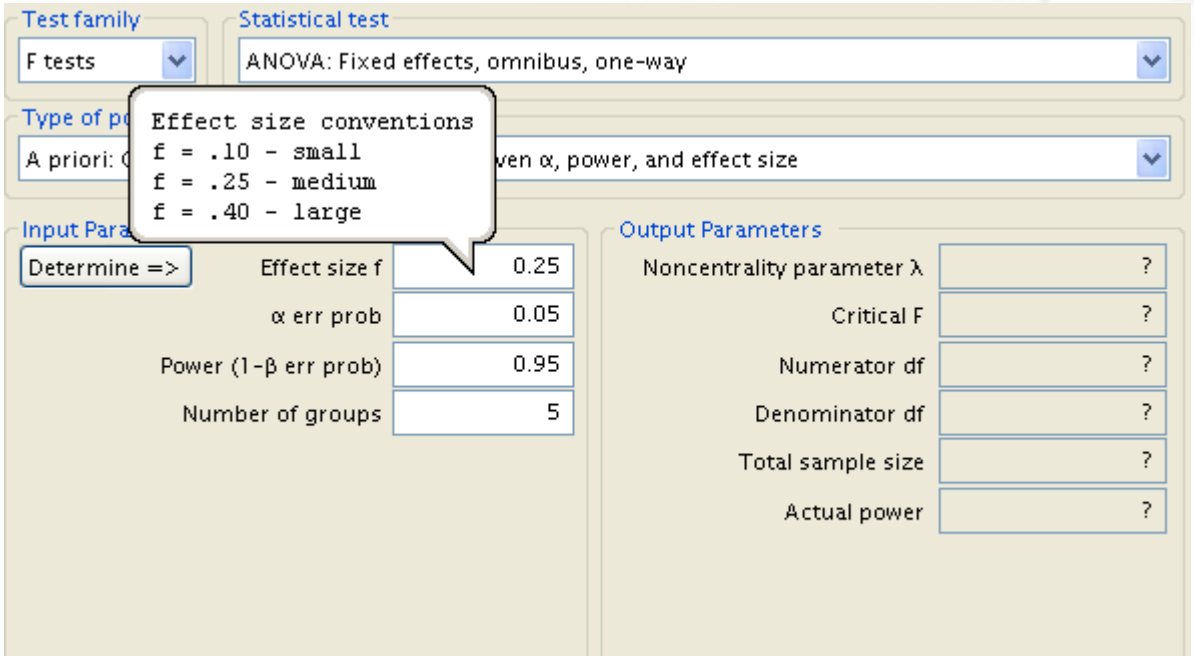
Type of power analysis: A priori: Compute required sample size – given α, power, and effect size

Input Parameters

Determine => Effect size f	0.25
α err prob	0.05
Power (1-β err prob)	0.95
Number of groups	5

Output Parameters

Noncentrality parameter λ	?
Critical F	?
Numerator df	?
Denominator df	?
Total sample size	?
Actual power	?


Test family
F tests
Statistical test
ANOVA: Fixed effects, omnibus, one-way
Type of po
A priori: C
ven α, power, and effect size
Effect size conventions
f = .10 - small
f = .25 - medium
f = .40 - large
Input Para
Determine =>
Effect size f
0.25
α err prob
0.05
Power (1-β err prob)
0.95
Number of groups
5
Output Parameters
Noncentrality parameter λ
?
Critical F
?
Numerator df
?
Denominator df
?
Total sample size
?
Actual power
?

Measuring Effect Size

- Statistical significance alone does not imply a substantial effect; just one larger than chance
- Cohen's *d* is the most common technique for assessing effect size
- Cohen's *d* = Difference between the means divided by the population standard deviation.
- *d > .8 means a large effect!*

Cohen's recommendations

Effect	d	r
Small	≥0.2	≥0.1
Medium	≥0.5	≥0.3
Large	≥0.8	≥0.5

Sample

- Qualitative research: The sample that was covered by in-depth interviews consisted of five participants (group of students that have pursued self-employment 10 years ago under mentoring guidance of a university professor from the Faculty of Economics in Osijek – development team of the CATI project).

Sample

- Quantitative research: University students of economics were selected for the sample since student sample is very common in testing of self-employment intentions. Students of the graduate study are the target group of this research since they are nearing completion of their studies and thus also before the decision on the choice of career. It is expected that graduate students will have enough time and energy for planning future business ventures (Audet, 2004). Wu and Wu (2008) argue that by understanding entrepreneurial intentions of university students it is possible to better predict their future real decision on launching a new business venture.
- The questionnaire was completed by a total of 453 students, 428 of which were used in the analysis.