Sample Size Calculation with GPower

Dr. Mark Williamson, Statistician
Biostatistics, Epidemiology, and Research Design Core
DaCCoTA

Purpose

- This Module was created to provide instruction and examples on sample size calculations for a variety of statistical tests on behalf of BERDC
- The software used is GPower, the premiere free software for sample size calculation that can be used in Mac or Windows

https://dl2.macupdate.com/images/icons256/24037.png

Background

- The Biostatistics, Epidemiology, and Research Design Core (BERDC) is a component of the DaCCoTA program
- Dakota Cancer Collaborative on Translational Activity has as its goal to bring together researchers and clinicians with diverse experience from across the region to develop unique and innovative means of combating cancer in North and South Dakota
- If you use this Module for research, please reference the DaCCoTA project

DaCCoTA
DAKOTA CANCER COLLABORATIVE ON TRANSLATIONAL ACTIVITY

The Why of Sample Size Calculation

- In designing an experiment, a key question is:

  How many animals/subjects do I need for my experiment?

- Too small of a sample size can under-detect the effect of interest in your experiment
- Too large of a sample size may lead to unnecessary wasting of resources and animals
- Like Goldilocks, we want our sample size to be 'just right'
- **The answer: *Sample Size Calculation***
- **Goal:** We strive to have enough samples to reasonably detect an effect if it really is there without wasting limited resources on too many samples.

https://upload.wikimedia.org/wikipedia/commons/thumb/e/ef/The_Three_Bears_-_Project_Gutenberg_eText_17034.jpg/1200px-The_Three_Bears_-_Project_Gutenberg_eText_17034.jpg

Key Bits of Sample Size Calculation

Effect size: magnitude of the effect under the alternative hypothesis

- The larger the effect size, the easier it is to detect an effect and require fewer samples

Power: probability of correctly rejecting the null hypothesis if it is false

- AKA, probability of detecting a true difference when it exists
- Power = **1-β**, where **β** is the probability of a Type II error (false negative)
- The higher the power, the more likely it is to detect an effect if it is present and the more samples needed
- Standard setting for power is 0.80

Significance level (α): probability of falsely rejecting the null hypothesis even though it is true

- AKA, probability of a Type I error (false positive)
- The lower the significance level, the more likely it is to avoid a false positive and the more samples needed
- Standard setting for **α** is 0.05

- Given those three bits, and other information based on the specific design, you can calculate sample size for most statistical tests

https://images-na.ssl-images-amazon.com/images/I/61YIBfLPPuL._SX355_.jpg

Effect Size in detail

- While ***Power*** and ***Significance level*** are usually set irrespective of the data, the effect size is a property of the sample data
- It is essentially a function of the difference between the means of the null and alternative hypotheses over the variation (standard deviation) in the data

<u>How to estimate Effect Size:</u>

A. Use background information in the form of preliminary/trial data to get means and variation, then calculate effect size directly

B. Use background information in the form of similar studies to get means and variation, then calculate effect size directly

C. With no prior information, make an estimated guess on the effect size expected, or use an effect size that corresponds to the size of the effect

   - Broad effect sizes categories are small, medium, and large
   - Different statistical tests will have different values of effect size for each category

$$Effect\ Size \approx \frac{Mean_1 - Mean_2}{Std.deviation}$$

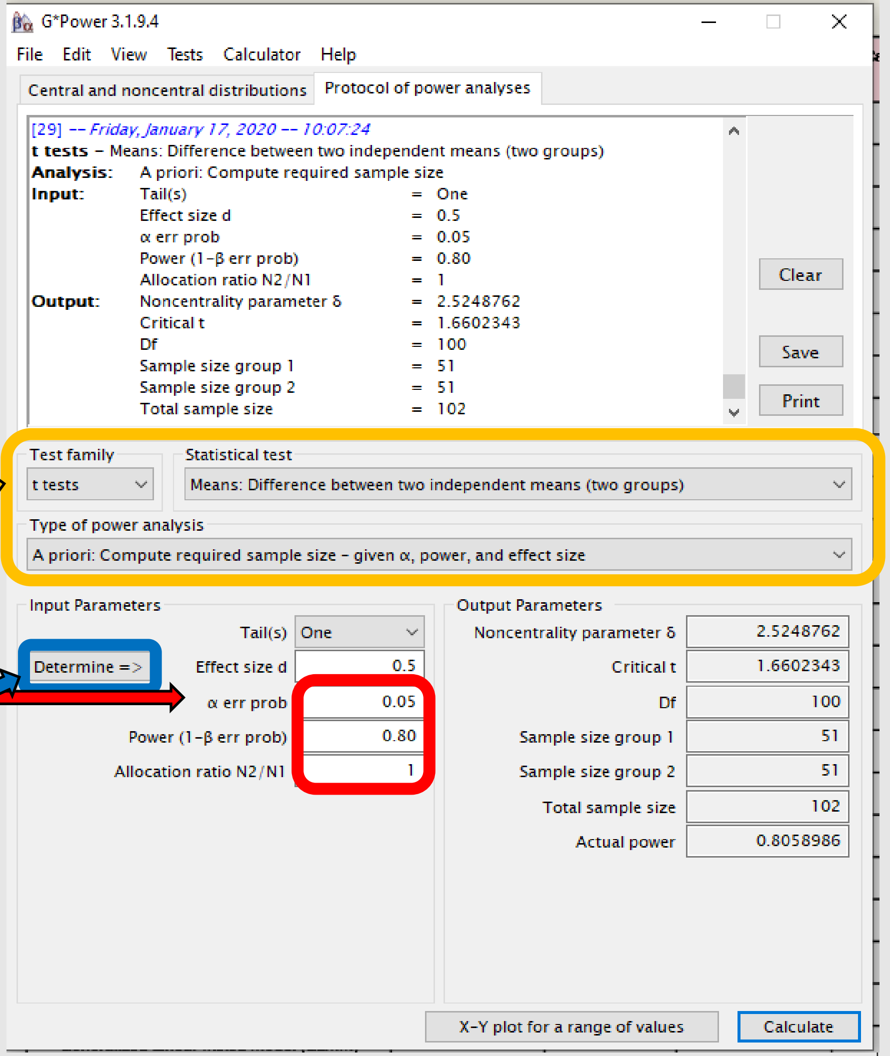
Statistical Rules of the Game

Here are a few pieces of terminology to refresh yourself with before embarking on calculating sample size:

- **Null Hypothesis (H0):** default or 'boring' state; your statistical test is run to either Reject or Fail to Reject the Null
- **Alternative Hypothesis (H1)**: alternative state; usually what your experiment is interested in retaining over the Null
- **One-Tailed Test:** looking for a deviation from the H0 in only one direction (ex: ***Is variable X larger than 0?***)
- **Two-tailed Test:** looking for a deviation from the H0 in either direction (ex: ***Is variable Y different from 0?***)
- **Parametric data:** approximately fits a normal distribution; needed for many statistical tests
- **Non-parametric data:** does not fit a normal distribution; alternative and less powerful tests available
- **Paired (dependent) data:** categories are related to one another (often result of before/after situations)
- **Un-paired (independent) data:** categories are not related to one another
- **Dependent Variable:** Depends on other variables; the variable the experimenter cares about; also known as the Y or response variable
- **Independent Variable:** Does not depend on other variables; usually set by the experimenter; also known as the X or predictor variable

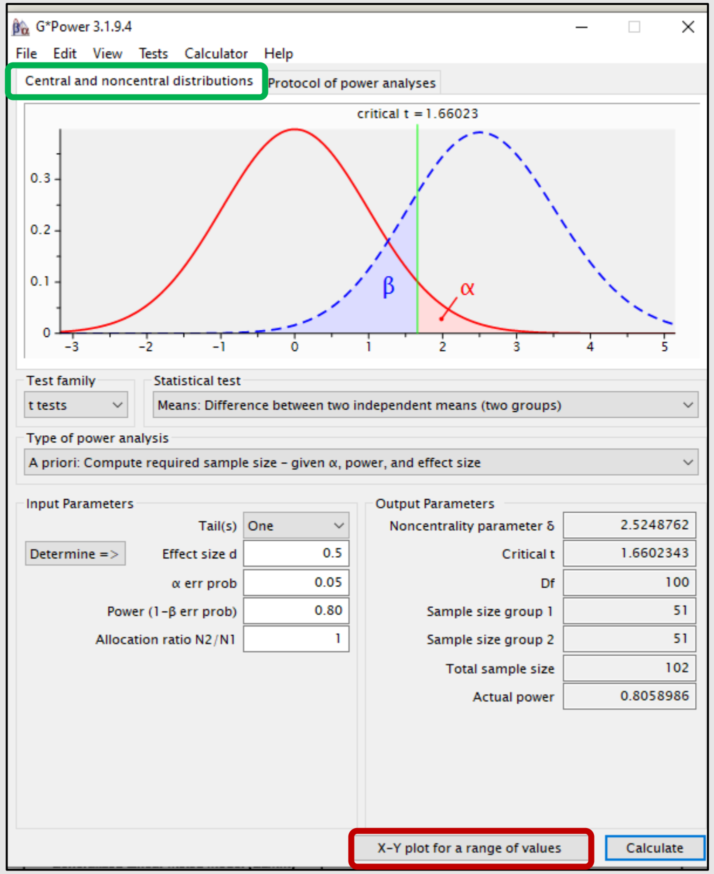
Using GPower: Basics

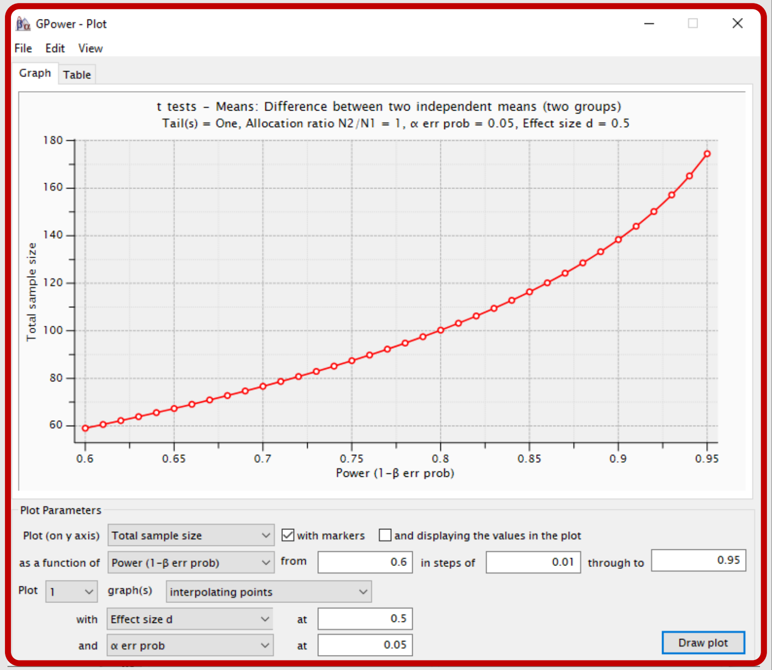
- Download for Mac or PC
  - http://www.psychologie.hhu.de/arbeitsgruppen/allgemeine-psychologie-und-arbeitspsychologie/gpower.html
- Three basic steps:
  - Select appropriate test:
  - Input parameters
  - Determine effect size (can use background info or guess)
- For situations when you have some idea of parameters such as mean, standard deviation, etc., I will refer to this as having **Background Information**
- If not, I will refer to this as **Naïve**



Using GPower: Graphics

- **Central and noncentral distributions**
  - Shows the distribution of the null hypothesis (red) and the alternative (blue)
  - Also has the critical values
- **X-Y plot for a range of values**
  - Can generate plots of one of the parameters α, effect size, power and sample size, depending on a range of values of the remaining parameters.





Taxonomy of Designs Covered

- **1 Numerical**
  - Parametric: **One mean T-test**
  - Non-parametric: **One mean Wilcoxon Test**
- **1 Numerical + 1 Categorical**
  - Categorical groups=2:
    - Independent (non-paired):
      - Parametric: **Two means T-test**
      - Non-parametric: **Mann-Whitney Test**
    - Dependent (paired):
      - Parametric: **Paired T-test**
      - Non-Parametric: **Paired Wilcoxon** Test
  - Categorical groups>2:
    - Independent (non-paired):
      - Parametric: **One-way ANOVA**
      - Non-Parametric: **Kruskal Wallace Test**
    - Dependent (paired):
      - Parametric: **Repeated Measures ANOVA**
      - Non-Parametric: **Friedman Test**

- **1 Numerical + 2^+ Categorical**
  - Single Category of Interest: **Multi-Way ANOVA**
    Blocked ANOVA
    Nested ANOVA
    Split-Plot ANOVA
  - Multiple Categories of Interest: **Multi-Way ANOVA**
- **1 Categorical: Proportion Test**
- **1 Categorical + 1 Categorical**
  - Independent (non-paired): **Fisher's Exact Test**
  - Dependent (paired): **McNamar's Test**
- **1 Categorical + 1^+ Categorical**
  - Categorical groups≥2: **Goodness-of-Fit Test**
- **1 Numerical + 1 Numerical**
  - Parametric: **Simple Linear Regression**
  - Non-parametric: **Spearman Rank-order Regression**
- **1 Numerical + 2^+ Numerical**
  - Parametric: **Multiple Linear Regression**
  - Non-Parametric: **Logistic** and **Poisson Regression**
- **1 Numerical + 1^+ Numerical + 1^+ Categorical:**
  - Only 1 Category of Interest: **ANCOVA**
  - *Multiple Categories of Interest: {**GLMM**}*

#	Name of Test	Numeric. Var(s)	Cat. Var(s)	Cat. Var Group #	Cat Var. # of Interest	Parametric	Paired
1	One Mean T-test	1	0	0	0	Yes	N/A
2	One Mean Wilcoxon Test	1	0	0	0	No	N/A
3	Two Means T-test	1	1	2	1	Yes	No
4	Mann-Whitney Test	1	1	2	1	No	No
5	Paired T-test	1	1	2	1	Yes	Yes
6	Paired Wilcoxon Test	1	1	2	1	No	Yes
7	One-way ANOVA	1	1	>2	1	Yes	No
8	Kruskal Wallace Test	1	1	>2	1	No	No
9	Repeated Measures ANOVA	1	1	>2	1	Yes	Yes
10	Friedman Test	1	1	>2	1	No	Yes
11	Multi-way ANOVA (1 Category of interest)	1	≥2	≥2	1	Yes	No
12	Multi-way ANOVA (>1 Category of interest)	1	≥2	≥2	>1	Yes	No
13	Proportions Test	0	1	2	1	N/A	N/A
14	Fisher's Exact Test	0	2	2	2	N/A	No
15	McNemar's Test	0	2	2	2	N/A	Yes
16	Goodness-of-Fit Test	0	≥1	≥2	1	N/A	No
17	Simple Linear Regression	2	0	N/A	N/A	Yes	N/A
18	Multiple Linear Regression	>2	0	N/A	N/A	Yes	N/A
19	Pearson's Correlation	2	1	N/A	N/A	Yes	No
20	Non-Parametric Regression (Logistic)	≥2	0	N/A	N/A	No	N/A
21	Non-Parametric Regression (Poisson)	≥2	0	N/A	N/A	No	N/A
22	ANCOVA	>1	≥1	>1	≥1	Yes	N/A

Format for each test

Overview

Example

{Parameter Calculations}

Practice

Answers

One Mean T-Test: Overview

Description: this tests if a sample mean is any different from a set value for a normally distributed variable.

Numeric. Var(s)	Cat. Var(s)	Cat. Var Group #	Cat Var. # of Interest	Parametric	Paired
1	0	0	0	Yes	N/A

Example:

- Is the average body temperature of college students any different from 98.6°F?
- H_0=98.6°F, H_1≠98.6°F

GPower:

- Select **t tests** from Test family
- Select **Means: difference from constant (one sample case)** from Statistical test
- Select **A priori** from Type of power analysis
- Background Info:
  a) Select **One** or **Two** from the Tail(s), depending on type
  b) Enter 0.05 in **α err prob** box – or a specific **α** you want for your study
  c) Enter 0.80 in **Power (1-β err prob**) box - or a specific power you want for your study
  d) Hit Determine =>
  e) Enter in the Mean H0, Mean HI, and SD, then hit Calculate and transfer to main window (this will calculate effect size and add it to the Input Parameters)
  f) Hit Calculate on the main window
  g) Find **Total sample size** in the Output Parameters
- Naïve:
  a) Run a-c as above
  b) Enter Effect size guess in the **Effect size d** box (small=0.2, medium=0.5, large=0.8)
  c) Hit Calculate on the main window
  d) Find **Total sample size** in the Output Parameters

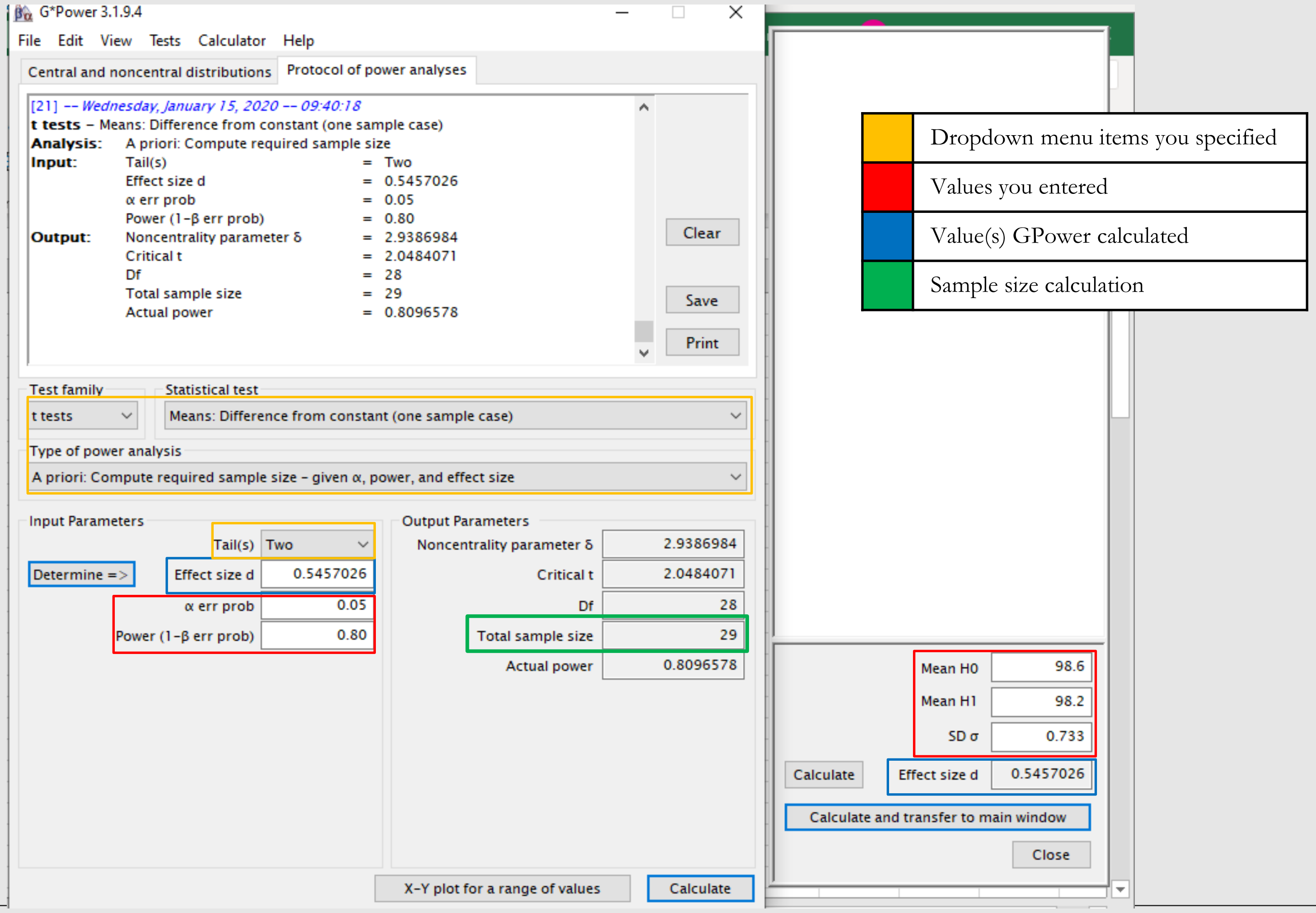
One Mean T-Test: Example

Is the average body temperature of college students any different from 98.6°F?

- H_0=98.6°F, H_1≠98.6°F
- From a trial study, you found the mean temperature to be **98.2°** with a standard deviation of **0.733**.
- Selected Two-tailed, because we were asking if temp differed, not whether it was simply lower or higher

Results:

- Total number of samples needed is 29.



One Mean T-Test: Practice

Calculate the sample size for the following scenarios (with α=0.05, and power=0.80):

1. You are interested in determining if the average income of college freshman is less than $20,000. You collect trial data and find that the mean income was $14,500 (SD=6000).
2. You are interested in determining if the average sleep time change in a year for college freshman is different from zero. You collect the following data of sleep change (in hours).

Sleep Change	-0.55	0.16	2.6	0.65	-0.23	0.21	-4.3	2	-1.7	1.9

3. You are interested in determining if the average weight change in a year for college freshman is greater than zero.

One Mean T-Test: Answers

Calculate the sample size for the following scenarios (with α=0.05, and power=0.80):

1. You are interested in determining if the average income of college freshman is less than $20,000. You collect trial data and find that the mean income was $14,500 (SD=6000).
   - **Found an effect size of 0.91, then used a one-tailed test to get a total sample size of 9**
2. You are interested in determining if the average sleep time change in a year for college freshman is different from zero. You collect the following data of sleep change (in hours).

Sleep Change	-0.55	0.16	-2.6	0.65	-0.23	0.21	-4.3	2	-1.7	1.9

   - **Mean H0=0, Mean H1=-0.446 with SD=1.96; found an effect size of 0.228 then used a two-tailed test to get a total sample size of 154**
3. You are interested in determining if the average weight change in a year for college freshman is greater than zero.
   - **Guessed a large effect size (0.8), then used a one-tailed test to get a total sample size of 12**

One Mean Wilcoxon: Overview

Description: this tests if a sample mean is any different from a set value for a non-normally distributed variable

Numeric. Var(s)	Cat. Var(s)	Cat. Var Group #	Cat Var. # of Interest	Parametric	Paired
1	0	0	0	No	N/A

Example:

- Is the average number of children in Grand Forks families greater than 1?
- H_0=1 child, H_1>1 child

GPower:

  - Select **t tests** from Test family
  - Select **Means: Wilcoxon signed-rank test (one sample case)** from Statistical test
  - Select **A priori** from Type of power analysis
- Background Info:
  a) Select **One** or **Two** from the Tail(s), depending on type
  b) Select Parent Distribution (**Laplace, Logistic,** or **min ARE**) depending on variable (min ARE is good default if you don't know for sure)
  c) Enter 0.05 in **α err prob** box – or a specific **α** you want for your study
  d) Enter 0.80 in **Power (1-β err prob**) box - or a specific power you want for your study
  e) Hit Determine =>
  f) Enter in the Mean H0, Mean HI, and SD, then hit Calculate and transfer to main window (this will calculate effect size and add it to the Input Parameters)
  g) Hit Calculate on the main window
  h) Find **Total sample size** in the Output Parameters
- Naïve:
  a) Run a-d as above
  b) Enter Effect size guess in the **Effect size d** box
  c) Hit Calculate on the main window
  d) Find **Total sample size** in the Output Parameters

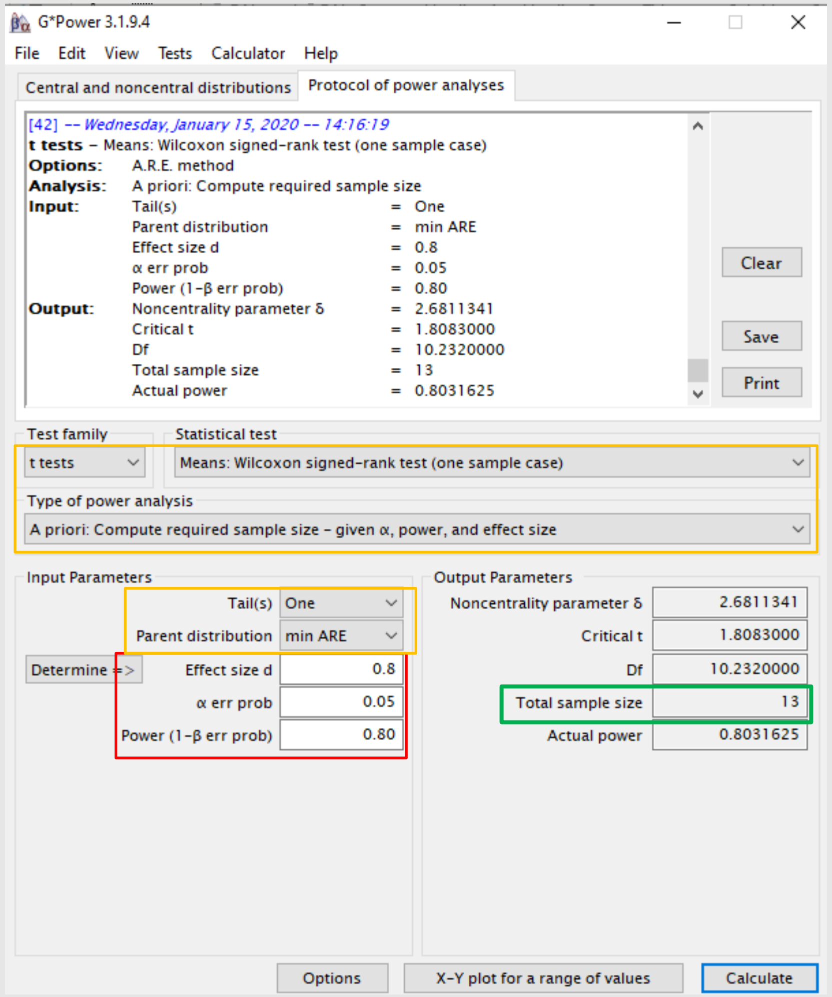
One Mean Wilcoxon: Example

Is the average number of children in Grand Forks families greater than 1?

- H_0=1 child, H_1>1 child
- You have no background information and you don't think it is normally distributed
- Default distribution: for non-normal, use **min ARE** – weakest, but least assumptions
- Try looking at a **large** effect
- Selected One-tailed, because we only cared if the number is greater than the null (1 child)

Results:

- Total number of samples needed is 13.



Yellow	Dropdown menu items you specified
Red	Values you entered
Blue	Value(s) GPower calculated
Green	Sample size calculation

One Mean Wilcoxon: Practice

Calculate the sample size for the following scenarios (with α=0.05, and power=0.80):

1. You are interested in determining if the average number of insurance claims for Grand Forks families different from 3. You collect trial data and find that the claim number was 2.8 (SD=0.53, with a Laplace distribution).
2. You are interested in determining if the average number of pets in Grand Forks families is greater than 1. You collect the following trial data for pet number.

Pet Number	1	1	1	3	2	1	0	0	0	4

3. You are interested in determining if the average number of yearly trips to the emergency room for Grand Forks families is different from 5?

One Mean Wilcoxon: Answers

Calculate the sample size for the following scenarios (with α=0.05, and power=0.80):

1. You are interested in determining if the average number of insurance claims for Grand Forks families different from 3. You collect trial data and find that the claim number was 2.8 (SD=0.53, with a Laplace distribution).
   - **Found an effect size of 0.38, then used a two-tailed test with Laplace distribution to get a total sample size of 39**
2. You are interested in determining if the average number of pets in Grand Forks families is greater than 1. You collect the following trial data for pet number.

Pet Number	1	1	1	3	2	1	0	0	0	4

   - **Mean H0=1, Mean H1=1.3 with SD=1.34;**
   - **Found an effect size of 0.224**
   - **Because I didn't know the distribution, went conservatively with min ARE, then used a one-tailed test to get a total sample size of 145**
3. You are interested in determining if the average number of yearly trips to the emergency room for Grand Forks families is different from 5?
   - **Guessed a medium effect (0.5) and because I didn't know the distribution went conservatively with min ARE, then used a two-tailed test to get a total sample size of 39**

Two Means T-Test: Overview

Description: this tests if a mean from one group is different from the mean of another group for a normally distributed variable. AKA, testing to see if the difference in means is different from zero.

Numeric. Var(s)	Cat. Var(s)	Cat. Var Group #	Cat Var. # of Interest	Parametric	Paired
1	**1**	**2**	**1**	**Yes**	**No**

Example:

- Is the average body temperature higher in men than in women?
- H_0=0°F, H_1>0°F

GPower:

- Select **t tests** from Test family
- Select **Means: Difference between two independent means (two groups)** from Statistical test
- Select **A priori** from Type of power analysis

- Background Info:
  a) Select **One** or **Two** from the Tail(s), depending on type
  b) Enter 0.05 in **α err prob** box – or a specific **α** you want for your study
  c) Enter 0.80 in **Power (1-β err prob**) box - or a specific power you want for your study
  d) Enter 1.00 in the **Allocation ratio N2/N1** box – unless your group sizes are not equal, then enter the ratio
  e) Hit Determine =>
  f) Enter in the Mean and SD for each group, then hit Calculate and transfer to main window (this will calculate effect size and add it to the Input Parameters)
  g) Hit Calculate on the main window
  h) Find **Total sample size** in the Output Parameters
- Naïve:
  a) Run a-d as above
  b) Enter Effect size guess in the **Effect size d** box
  c) Hit Calculate on the main window
  d) Find **Total sample size** in the Output Parameters

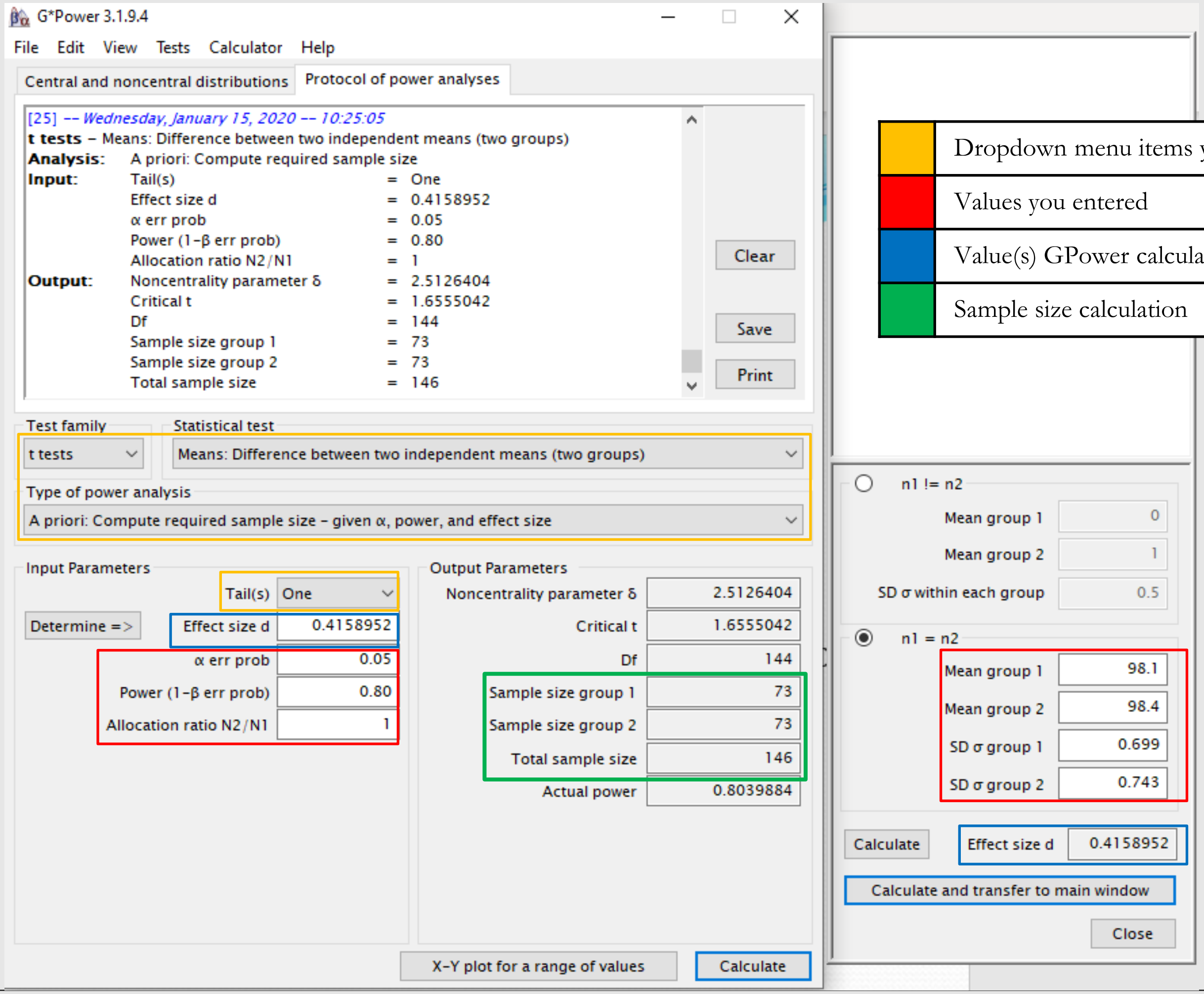
Two Means T-test : Example

Is the average body temperature higher in men than in women?

- From a trial study, you found the mean temperature to be **98.1°** for men with a standard deviation (SD) of **0.699** and **98.4 °** for women with a SD of **0.743.**
- Selected one-tailed, because we only cared to test if men's temp was higher than women's, not lower
- Group 1 is men, group 2 is women

Results:

- Need a sample size of 73 per gender, for a total of 146.



- Dropdown menu items you specified
- Values you entered
- Value(s) GPower calculated
- Sample size calculation

Two Means T-Test: Practice

Calculate the sample size for the following scenarios (with α=0.05, and power=0.80):

1. You are interested in determining if the average daily caloric intake different between men and women. You collected trial data and found the average caloric intake for 10 males to be 2350.2 (SD=258), 5 while females had intake of 1872.4 (SD=420).
2. You are interested in determining if the average protein level in blood different between men and women. You collected the following trial data on protein level (grams/deciliter).

Male Protein	1.8	5.8	7.1	4.6	5.5	2.4	8.3	1.2
Female Protein	9.5	2.6	3.7	4.7	6.4	8.4	3.1	1.4

3. You are interested in determining if the average glucose level in blood is lower in men than women.

Two Means T-Test: Answers

Calculate the sample size for the following scenarios (with α=0.05, and power=0.80):

1. You are interested in determining if the average daily caloric intake different between men and women. You collected trial data and found the average caloric intake for 10 males to be 2350.2 (SD=258), while 5 females had intake of 1872.4 (SD=420).
   - **Found an effect size of 1.37, then used a two-tailed test with Allocation Ratio of 0.5 (5 women/10 men) to get a total sample size of 22 (15 men, 7 women)**

2. You are interested in determining if the average protein level in blood different between men and women. You collected the following trial data on protein level (grams/deciliter).

Male Protein	1.8	5.8	7.1	4.6	5.5	2.4	8.3	1.2
Female Protein	9.5	2.6	3.7	4.7	6.4	8.4	3.1	1.4

   - **Mean group 1 (men)=4.5875 with SD=2.575, mean group 2 (women)=4.975 with SD=2.875**
   - **Found an effect size of 0.142, then used a two-tailed test with Allocation Ratio of 1.0 (equal group sizes) to get a total sample size of 1560 (780 each for men and women)**

1. You are interested in determining if the average glucose level in blood is lower in men than women in a predominately female college.
   - **Guessed a small effect (0.2), then used a two-tailed test with Allocation Ratio of 3.0 (3 women: 1 man) to get a total sample size of 826 (207 men, 619 women)**

Mann-Whitney Test: Overview

Description: this tests if a mean from one group is different from the mean of another group for a non-normally distributed variable. AKA, testing to see if the difference in means is different from zero.

Numeric. Var(s)	Cat. Var(s)	Cat. Var Group #	Cat Var. # of Interest	Parametric	Paired
1	1	2	1	No	No

Example:

- Does the average number of snacks per day for individuals on a diet differ between young and old persons?
- H_0=0 difference in snack number,
  $H_1 \neq 0$ difference in snack number

GPower:

- Select **t tests** from Test family
- Select **Means: Wilcoxon-Mann-Whitney test (two groups)** from Statistical test
- Select **A priori** from Type of power analysis
- Background Info:
  a) Select **One** or **Two** from the Tail(s), depending on type
  b) Select Parent Distribution (**Laplace, Logistic,** or **min ARE**) depending on variable (min ARE is good default if you don't know for sure)
  c) Enter 0.05 in **α err prob** box – or a specific **α** you want for your study
  d) Enter 0.80 in **Power (1-β err prob**) box - or a specific power you want for your study
  e) Enter 1.00 in the **Allocation ratio N2/N1** box – unless your group sizes are not equal, then enter the ratio
  f) Hit Determine =>
  g) Enter in the Mean and SD for each group, then hit Calculate and transfer to main window (this will calculate effect size and add it to the Input Parameters)
  h) Hit Calculate on the main window
  i) Find **Total sample size** in the Output Parameters
- Naïve:
  a) Run a-e as above
  b) Enter Effect size guess in the **Effect size d** box
  c) Hit Calculate on the main window
  d) Find **Total sample size** in the Output Parameters

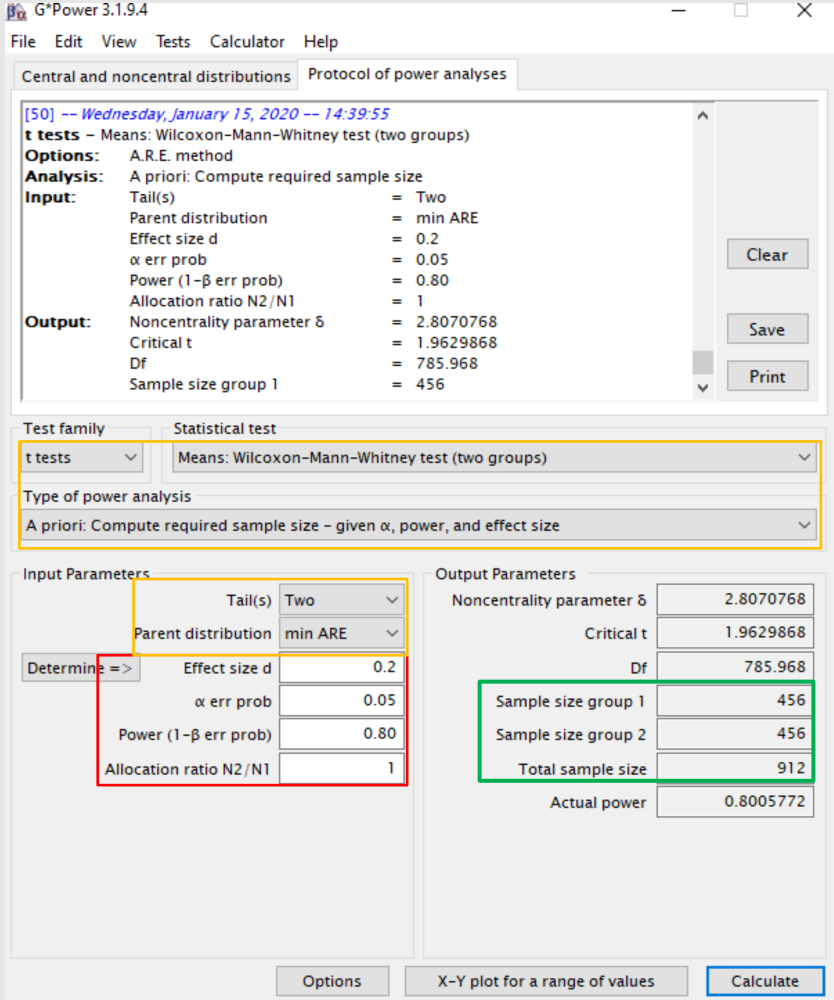
Mann-Whitney Test: Example

Does the average number of snacks per day for individuals on a diet differ between young and old persons?

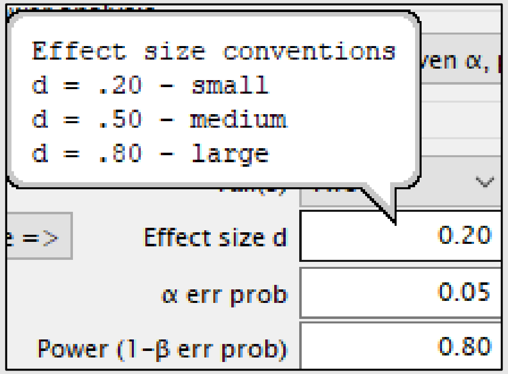
- H_0=0 difference in snack number, $H_1 \neq 0$ difference in snack number
- You have no background information and you think it might be a **small** effect (0.20)
- Selected two-tailed, because we only care if there is a difference in the two groups
  Hovering over the **Effect size d** box will show a bubble of size conventions

Results:

Need a sample size of 456 for each age for a total of 912.



Yellow	Dropdown menu items you specified
Red	Values you entered
Blue	Value(s) GPower calculated
Green	Sample size calculation



Mann-Whitney Test: Practice

Calculate the sample size for the following scenarios (with α=0.05, and power=0.80):

1. You are interested in determining if the average number of walks per week for individuals on a diet is lower in younger people than older. You collected trial data and found mean walk number for young dieters to be 1.4 (SD=0.18) and 2.5 for older (SD=0.47).
2. You are interested in determining if the number of meals per day for individuals on a diet is higher in younger people than older. You collected trial data on meals per day.

Young meals	1	2	2	3	3	3	3	4
Older meals	1	1	1	2	2	2	3	3

3. You are interested in determining if the average number of weight loss websites visited per day for individuals on a diet is different in younger people than older.

Mann-Whitney Test: Answers

Calculate the sample size for the following scenarios (with α=0.05, and power=0.80):

1. You are interested in determining if the average number of walks per week for individuals on a diet is lower in younger people than older. You collected trial data and found mean walk number for young dieters to be 1.4 (SD=0.18) and 2.5 for older (SD=0.47).
   - **Found an effect size of 3.09, then used a one-tailed test with a min ARE distribution and Allocation Ratio of 1.0 to get a total sample size of 6 (3 each for young and old)**
2. You are interested in determining if the number of meals per day for individuals on a diet is higher in younger people than older. You collected trial data on meals per day

Young meals	1	2	2	3	3	3	3	4
Older meals	1	1	1	2	2	2	3	3

   - **Mean group 1 (older)=1.875 with SD=0.834, mean group 2 (younger)=2.625 with SD=0.916**
   - **Found an effect size of 0.53, then used a one-tailed test with a Logistic distribution and Allocation Ratio of 1.0 to get a total sample size of 82 (41 each for young and old)**
3. You are interested in determining if the average number of weight loss websites visited per day for individuals on a diet is different in younger people than older.
   - **Guessed a medium effect (0.5) and Logistic Regression, then used a two-tailed test and Allocation Ratio of 1.0 to get a total sample size of 118 (59 each for young and old)**

Paired T-test: Overview

Description: this tests if a mean from one group is different from the mean of another group, where the groups are dependent (not independent) for a normally distributed variable. Pairing can be leaves on same branch, siblings, the same individual before and after a trial, etc.

Numeric. Var(s)	Cat. Var(s)	Cat. Var Group #	Cat Var. # of Interest	Parametric	Paired
1	**1**	**2**	**1**	**Yes**	**Yes**

Example:

- Is heart rate higher in patients after a run compared to before a run?
- H_0=0 bpm, H_1>0 bpm

GPower:

  - Select **t tests** from Test family
  - Select **Means: Difference between two dependent means (matched)** from Statistical test
  - Select **A priori** from Type of power analysis
- Background Info:
  a) Select **One** or **Two** from the Tail(s), depending on type
  b) Enter 0.05 in **α err prob** box – or a specific α you want for your study
  c) Enter 0.80 in **Power (1-β err prob)** box - or a specific power you want for your study
  d) Hit Determine =>
  e) Enter in the Mean and SD for each group, as well as the **Correlation between groups**, then hit Calculate and transfer to main window (this will calculate effect size and add it to the Input Parameters)
  f) Hit Calculate on the main window
  g) Find **Total sample size** in the Output Parameters
- Naïve:
  a) Run a-c as above
  b) Enter Effect size guess in the **Effect size d** box
  c) Hit Calculate on the main window
  d) Find **Total sample size** in the Output Parameters

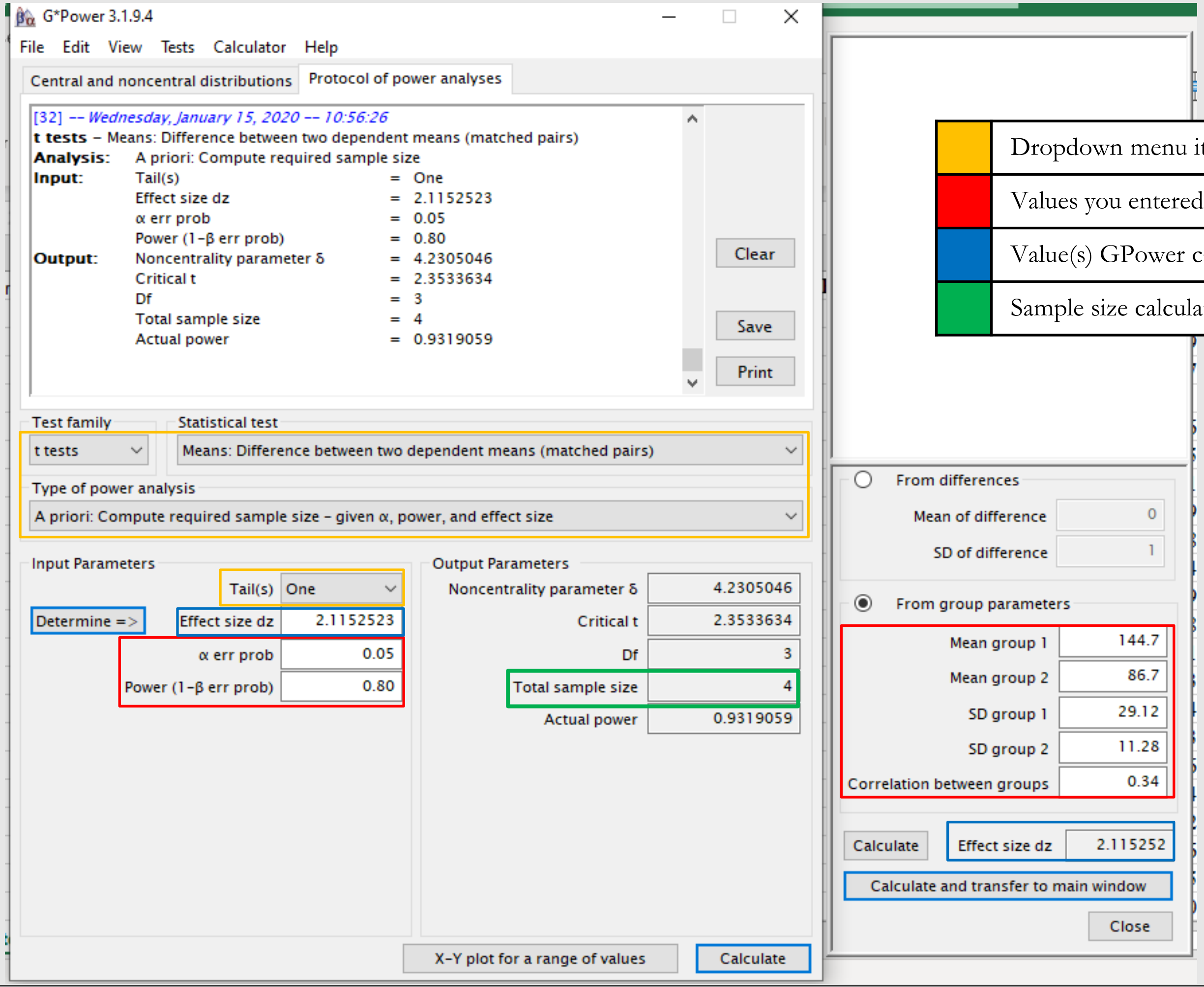
Paired T-test: Example

Is heart rate higher in patients after a run compared to before a run?

- H_0=0 bpm, H_1>0 bpm
- From a trial study, you found the mean beats per minute (bpm) before the run to be **86.7** with SD=**11.28** and **144.7** with SD=**29.12** after the run. The correlation between measurements before and after is **0.34**.
- Selected One-tailed, because we only cared if bpm was higher after a run
- Group 1 is after the run, while group 2 is before the run
- If you only have the mean difference, can use that in the **From differences** option
- If you don't have correlation data, can guess it (0.5 is good standard if you have no idea)

Results:

- Need a sample size of 4 individuals (each with a before and after run measurement).
- Because the difference in means was so large, the effect size was huge, so it does not take many samples to test this question



- Yellow: Dropdown menu items you specified
- Red: Values you entered
- Blue: Value(s) GPower calculated
- Green: Sample size calculation

Paired T-Test: Practice

Calculate the sample size for the following scenarios (with α=0.05, and power=0.80):

1. You are interested in determining if heart rate is higher in patients after a doctor's visit compared to before a visit. You collected the following trial data and found mean heart rate before and after a visit.

BPM before	126	88	53.1	98.5	88.3	82.5	105	41.9
BPM after	138.6	110.1	58.44	110.2	89.61	98.6	115.3	64.3

2. You are interested in determining if metabolic rate in patients after surgery is different from before surgery. You collected trial data and found a mean difference of 0.73 (SD=2.9).

3. You are interested in determining if glucose levels in patients after surgery are lower compared to before surgery.

Paired T-Test : Answers

Calculate the sample size for the following scenarios (with α=0.05, and power=0.80):

1. You are interested in determining if heart rate is higher in patients after a doctor's visit compared to before a visit. You collected the following trial data and found mean heart rate before and after a visit.

BPM before	126	88	53.1	98.5	88.3	82.5	105	41.9
BPM after	138.6	110.1	58.44	110.2	89.61	98.6	115.3	64.3

   - **Mean 1 (before)=85.4 with SD=27.2; Mean 2 (after)=98.1 with SD=26.8; correlation between groups=0.96**
   - **Found an effect size of 1.66, then used a one-tailed test to get a total sample size of 4 pairs**

2. You are interested in determining if metabolic rate in patients after surgery is different from before surgery. You collected trial data and found a mean difference of 0.73 (SD=2.9).
   - **Found an effect size of 0.25, then used a two-tailed test to get a total sample size of 126**

3. You are interested in determining if glucose levels in patients after surgery are lower compared to before surgery.
   - **Guessed a small effect (0.20), then used a one-tail test to get a total sample size of 156**

Paired Wilcoxon: Overview

Description: this tests if a mean from one group is different from the mean of another group, where the groups are dependent (not independent) for a non-normally distributed variable.

Numeric. Var(s)	Cat. Var(s)	Cat. Var Group #	Cat Var. # of Interest	Parametric	Paired
1	**1**	**2**	**1**	**No**	**Yes**

Example:

- Are genome methylation patterns different between identical twins?
- H_0=0% methylation, H_1≠0% methylation

GPower:

- Select **t tests** from Test family
- Select **Means: Wilcoxon signed-rank test (matched pairs)** from Statistical test
- Select **A priori** from Type of power analysis
- Background Info:
  a) Select **One** or **Two** from the Tail(s), depending on type
  b) Select Parent Distribution (**Laplace, Logistic,** or **min ARE**) depending on variable (min ARE is good default if you don't know for sure)
  c) Enter 0.05 in **α err prob** box – or a specific **α** you want for your study
  d) Enter 0.80 in **Power (1-β err prob**) box - or a specific power you want for your study
  e) Hit Determine =>
  f) Enter in the Mean and SD for each group, as well as the **Correlation between groups**, then hit Calculate and transfer to main window (this will calculate effect size and add it to the Input Parameters)
  g) Hit Calculate on the main window
  h) Find **Total sample size** in the Output Parameters
- Naïve:
  a) Run a-d as above
  b) Enter Effect size guess in the **Effect size d** box
  c) Hit Calculate on the main window
  d) Find **Total sample size** in the Output Parameters

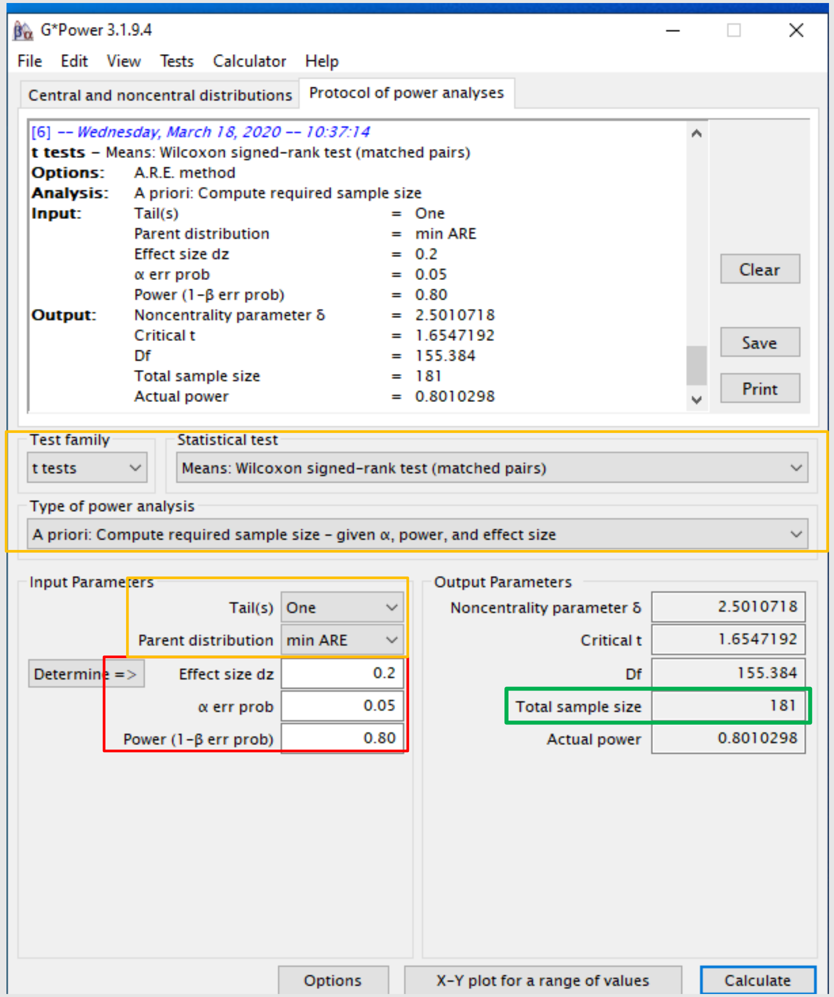
Paired Wilcoxon: Example

Are genome methylation patterns different between identical twins?

- H_0=0% methylation, H_1>0% methylation
- You have no background information on this, so assume there is going to be a **small** effect (0.2).
- Selected one-tailed, because can't have negative methylation

Results:

- Need a sample size of 181 pairs of twins (362 individuals total).



Yellow	Dropdown menu items you specified
Red	Values you entered
Blue	Value(s) GPower calculated
Green	Sample size calculation

Paired Wilcoxon: Practice

Calculate the sample size for the following scenarios (with α=0.05, and power=0.80):

1. You are interested in determining if genome methylation patterns are different between fraternal twins. You collected trial data and found mean methylation levels in twin A at 43.2 (SD=20.9) and at 65.7 (SD=28.1) for twin B.
2. You are interested in determining if genome methylation patterns are higher in the first fraternal twin born compared to the second. You collected the following trial data on methylation level difference (in percentage).

Methy. Diff (%)	5.96	5.63	1.25	1.17	3.59	1.64	1.6	1.4

3. You are interested in determining if the cancer risk rate is lower in the first identical twin born compared to the second.

Paired Wilcoxon: Answers

Calculate the sample size for the following scenarios (with α=0.05, and power=0.80):

1. You are interested in determining if genome methylation patterns are different between fraternal twins. You collected trial data and found mean methylation levels in twin A at 43.2 (SD=20.9) and at 65.7 (SD=28.1) for twin B.
    - **With a default correlation of 0.5, found an effect size of 0.89, then used a two-tailed test with a min ARE distribution to get a total sample size of 14 pairs**
2. You are interested in determining if genome methylation patterns are higher in the first fraternal twin born compared to the second. You collected the following trial data on methylation level difference (in percentage).

Methy. Diff (%)	5.96	5.63	1.25	1.17	3.59	1.64	1.6	1.4

- **Mean of difference=2.505 with SD=1.87**
- **Found an effect size of 1.33, then used a one-tailed test with a min ARE distribution to get a total sample size of 6 pairs**

3. You are interested in determining if the cancer risk rate is lower in the first identical twin born compared to the second.
    - **Guessed a small effect (0.2) and because I didn't know the distribution went conservatively with min ARE, then used a one-tailed test to get a total sample size of 181 pairs**

One-way ANOVA: Overview

Description: this tests if at least one mean is different among groups, where the groups are larger than two, for a normally distributed variable. ANOVA is the extension of the Two Means T-test for more than two groups.

Numeric. Var(s)	Cat. Var(s)	Cat. Var Group #	Cat Var. # of Interest	Parametric	Paired
1	**1**	**>2**	**1**	**Yes**	**No**

Example:

- Is there a difference in new car interest rates across 6 different cities?
- $H_0 = 0$, $H_1 \neq 0$

GPower:

- Select **F tests** from Test family
- Select **ANOVA: Fixed effects, omnibus, one-way** from Statistical test
- Select **A priori** from Type of power analysis

- Background Info:
  a) Enter 0.05 in **α err prob** box – or a specific α you want for your study
  b) Enter 0.80 in **Power (1-β err prob**) box - or a specific power you want for your study
  c) Enter the number of groups
  d) Hit Determine =>
  e) Enter in the **Mean** and **Size** for each group, as well as the SD within each group, then hit Calculate and transfer to main window (this will calculate effect size and add it to the Input Parameters)
  f) Hit Calculate on the main window
  g) Find **Total sample size** in the Output Parameters
- Naïve:
  a) Run a-c as above
  b) Enter Effect size guess in the **Effect size d** box
  c) Hit Calculate on the main window
  d) Find **Total sample size** in the Output Parameters

One-way ANOVA: Example

Is there a difference in new car interest rates across 6 different cities?

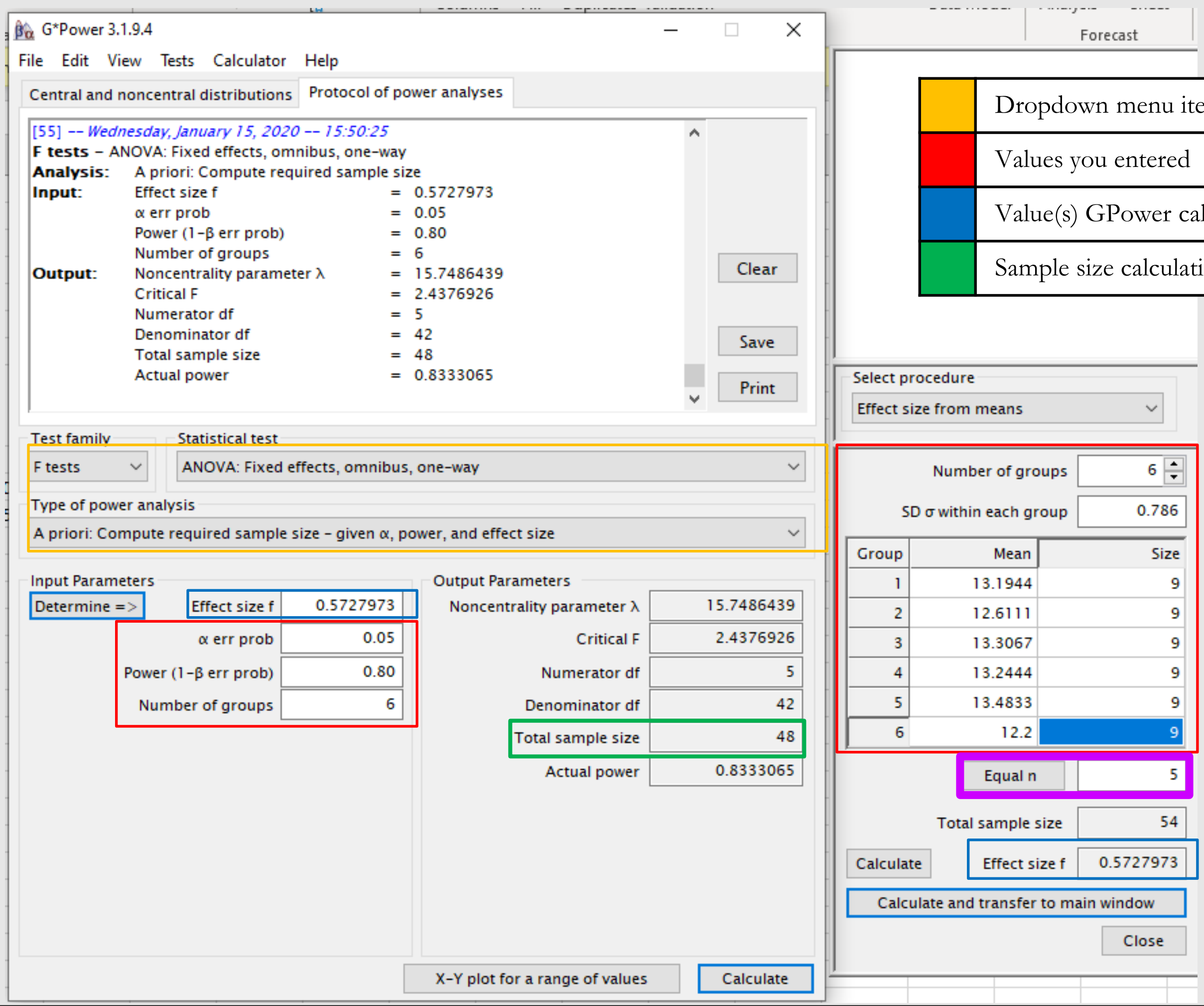
- $H_0=0\%$, $H_1 \neq 0\%$
- From a trial study, you found the following information for means

City	1	2	3	4	5	6
mean	13.19444	12.61111	13.30667	13.24444	13.48333	12.2

- With a SD of 0.786 and group sizes of 9
- No Tails in ANOVA
- If groups are equal size, can enter the size in the Purple box, then click **Equal n**

Results:

- A total number of 48 samples are needed (8 per group).



(yellow)	Dropdown menu items you specified
(red)	Values you entered
(blue)	Value(s) GPower calculated
(green)	Sample size calculation

One-way ANOVA: Practice

Calculate the sample size for the following scenarios (with α=0.05, and power=0.80):

1. You are interested in determining there is a difference in weight lost between 4 different surgery options. You collect the following trial data of weight lost in pounds (shown on right)
2. You are interested in determining if there is a difference in white blood cell counts between 5 different medication regimes.

Option 1	Option 2	Option 3	Option 4
6.3	9.9	5.1	1.0
2.8	4.1	2.9	2.8
7.8	3.9	3.6	4.8
7.9	6.3	5.7	3.9
4.9	6.9	4.5	1.6

One-way ANOVA: Answers

Calculate the sample size for the following scenarios (with α=0.05, and power=0.80):

1. You are interested in determining there is a difference in weight lost between 4 different surgery options. You collect the following trial data of weight lost in pounds (shown on right)
   - **4 groups with group size of 5; means of 5.94, 6.22, 4.36, and 2.82 with SD=2.23**
   - **Found an effect size of 0.61 for a total sample size of 36**
2. You are interested in determining if there is a difference in white blood cell counts between 5 different medication regimes.
   - **5 groups; guessed a medium effect size (0.25) for a total sample size of 200**

Option 1	Option 2	Option 3	Option 4
6.3	9.9	5.1	1.0
2.8	4.1	2.9	2.8
7.8	3.9	3.6	4.8
7.9	6.3	5.7	3.9
4.9	6.9	4.5	1.6

Kruskal Wallace Test: Overview

Description: this tests if at least one mean is different among groups, where the groups are larger than two for a non-normally distributed variable. There really isn't a standard way of calculating sample size in GPower, but you can use a rule of thumb:

1. Run Parametric Test
2. Add 15% to total sample size

(https://www.graphpad.com/guides/prism/7/statistics/index.htm?stat_sample_size_for_nonparametric_.htm)

Numeric. Var(s)	Cat. Var(s)	Cat. Var Group #	Cat Var. # of Interest	Parametric	Paired
1	**1**	**>2**	**1**	**No**	**No**

Example:

- Is there a difference in draft rank across 3 different months?
- $H_0=0$, $H_1\neq0$

GPower:

- Select **F tests** from Test family
- Select **ANOVA: Fixed effects, omnibus, one-way** from Statistical test
- Select **A priori** from Type of power analysis
- Background Info:
  a) Enter 0.05 in **α err prob** box – or a specific α you want for your study
  b) Enter 0.80 in **Power (1-β err prob)** box - or a specific power you want for your study
  c) Enter the number of groups
  d) Hit Determine =>
  e) Enter in the **Mean** and **Size** for each group, as well as the SD within each group, then hit Calculate and transfer to main window (this will calculate effect size and add it to the Input Parameters)
  f) Hit Calculate on the main window
  g) Find **Total sample size** in the Output Parameters
  h) Add 15% to size (Total + Total*0.15)
- Naïve:
  a) Run a-c as above
  b) Enter Effect size guess in the **Effect size d** box
  c) Hit Calculate on the main window
  d) Find **Total sample size** in the Output Parameters
  e) Add 15% to size (Total + Total*0.15)

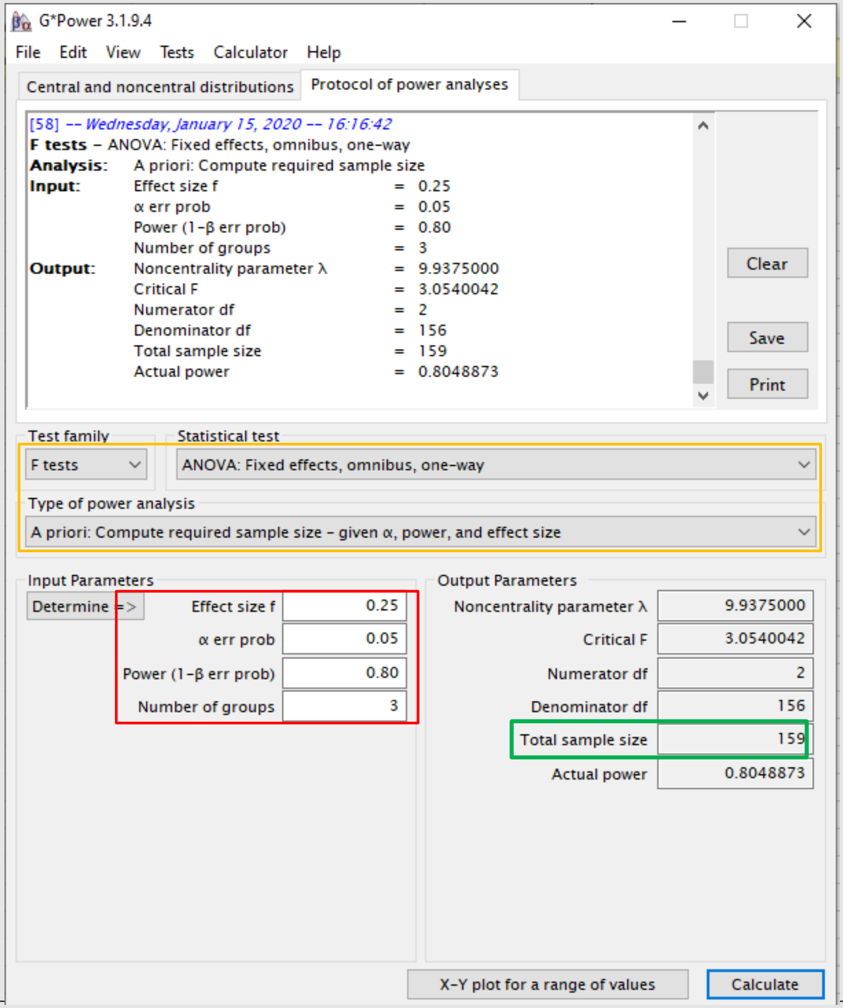
Kruskal Wallace Test: Example

Is there a difference in draft rank across 3 different months?

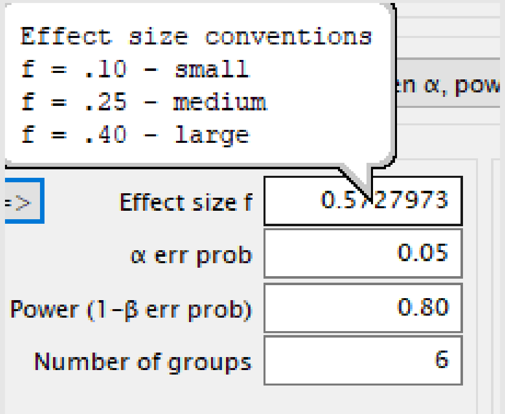
- H_0=0, H_1≠0
- You don't have background info, so you guess that there is a **medium** effect size (0.25)

Results:

- A total number of 159 samples are needed (53 per group) for the parametric (ANOVA)
- For the non-parametric:
  - 159 + 159*0.15 = 182.85
  - Round up
  - Total of 183 samples (61 per group) are needed.
- Notice that non-parametric is weaker



(yellow)	Dropdown menu items you specified
(red)	Values you entered
(blue)	Value(s) GPower calculated
(green)	Sample size calculation



Kruskal Wallace Test: Practice

Calculate the sample size for the following scenarios (with α=0.05, and power=0.80):

1. You are interested in determining there is a difference in hours worked across 3 different groups (faculty, staff, and hourly workers). You collect the following trial data of weekly hours (shown on right).
2. You are interested in determining there is a difference in assistant professor salaries across 25 different departments.

Faculty	Staff	Hourly
42	46	29
45	45	42
46	37	33
55	42	50
42	40	23

Kruskal Wallace Test: Answers

Calculate the sample size for the following scenarios (with α=0.05, and power=0.80):

1. You are interested in determining there is a difference in hours worked across 3 different groups (faculty, staff, and hourly workers). You collect the following trial data of weekly hours (shown on right).
   - **3 groups with group size of 5; means of 46, 42 and 35.4 with SD=8.07**
   - **Found an effect size of 0.54 for a parametric sample size of 39**
   - **39*1.15=44.85 -> round up to total sample size of 45 (15 per group)**
2. You are interested in determining there is a difference in assistant professor salaries across 25 different departments.
   - **25 groups; guessed a small effect size (0.10) for a parametric sample size of 2275**
   - **2275*1.15=2616.26 -> round up to total sample size of 2625 (105 per group)**

Faculty	Staff	Hourly
42	46	29
45	45	42
46	37	33
55	42	50
42	40	23

Repeated Measures ANOVA: Overview

Description: this tests if at least one mean is different among groups, where the groups are repeated measures (more than two) for a normally distributed variable. Repeated Measures ANOVA is the extension of the Paired T-test for more than two groups.

Numeric. Var(s)	Cat. Var(s)	Cat. Var Group #	Cat Var. # of Interest	Parametric	Paired
1	1	>2	1	Yes	Yes

Example:

- Is there a difference in blood pressure at 1, 2, 3, and 4 months post-treatment?
- H_0=0 bpm, $H_1 \neq 0$ bpm

GPower:

- Select **F tests** from Test family
- Select **ANOVA: Repeated measures, within factors** from Statistical test
- Select **A priori** from Type of power analysis
- Background Info or Naïve:
  a) Enter 0.05 in **α err prob** box – or a specific α you want for your study
  b) Enter 0.80 in **Power (1-β err prob)** box - or a specific power you want for your study
  c) Enter the number of **groups** and **measurements** (see next page)
  d) Enter the **Corr among rep measures** and the **Nonsphericity correction ε** (see next page)
  e) Hit Determine =>
  f) Enter calculated **Partial** η^2 (eta squared) (see next page)
     - for ***Background,*** enter calculated eta squared and select **as in SPSS** for the **Effect size specification** in the **Options** bar
     - for ***Naïve***, enter guess eta squared (based on convention) and keep the Effect size specification on **as in Gpower 3.0** and will also have to enter a guess for the **Corr among rep measures** parameter
  g) Hit **Calculate and transfer to main window**
  h) Hit Calculate on the main window
  i) Find **Total sample size** in the Output Parameters

Repeated Measures ANOVA: parameters and how to calculate them

- **Number of groups**: how many different groups are being subjected to the repeated measurements
  - In the simple case, it is one (college students)
  - In more complex designs, it may be more than one (college freshman, sophomores, juniors, and seniors)
- **Number of measurements**: how many repeats of a measurement
  - Ex. number of times blood pressure is measured
- **Corr among rep measures**: Correlation
  - No easy way to get a single correlation value
  - Can average all the values from a correlation table
  - Or default to 0.5 unless you have reason to believe it higher or lower
- **Nonsphericity correction ε**: The assumption of sphericity,
  - Can be estimated with background info
  - Otherwise, if the data is assumed to be spherical, enter 1
  - Spherical: variances of the differences between all possible pairs of the within subjects variable should be equivalent
- **Partial η2 (eta squared):** proportion of the total variance in a dependent variable that is associated with the membership of different groups defined by an independent variable
  - The larger the eta, the larger the effect size
  - Can be estimated with background information (Equation is $\eta 2 = SS_{eff} / (SS_{eff} + SS_{err})$ where SS_{eff} is the sum of squares between groups and the SS_{err} is the sum of squares within groups
  - Otherwise, enter guessed value; convention is **small=0.02, medium=0.06, large=0.14**

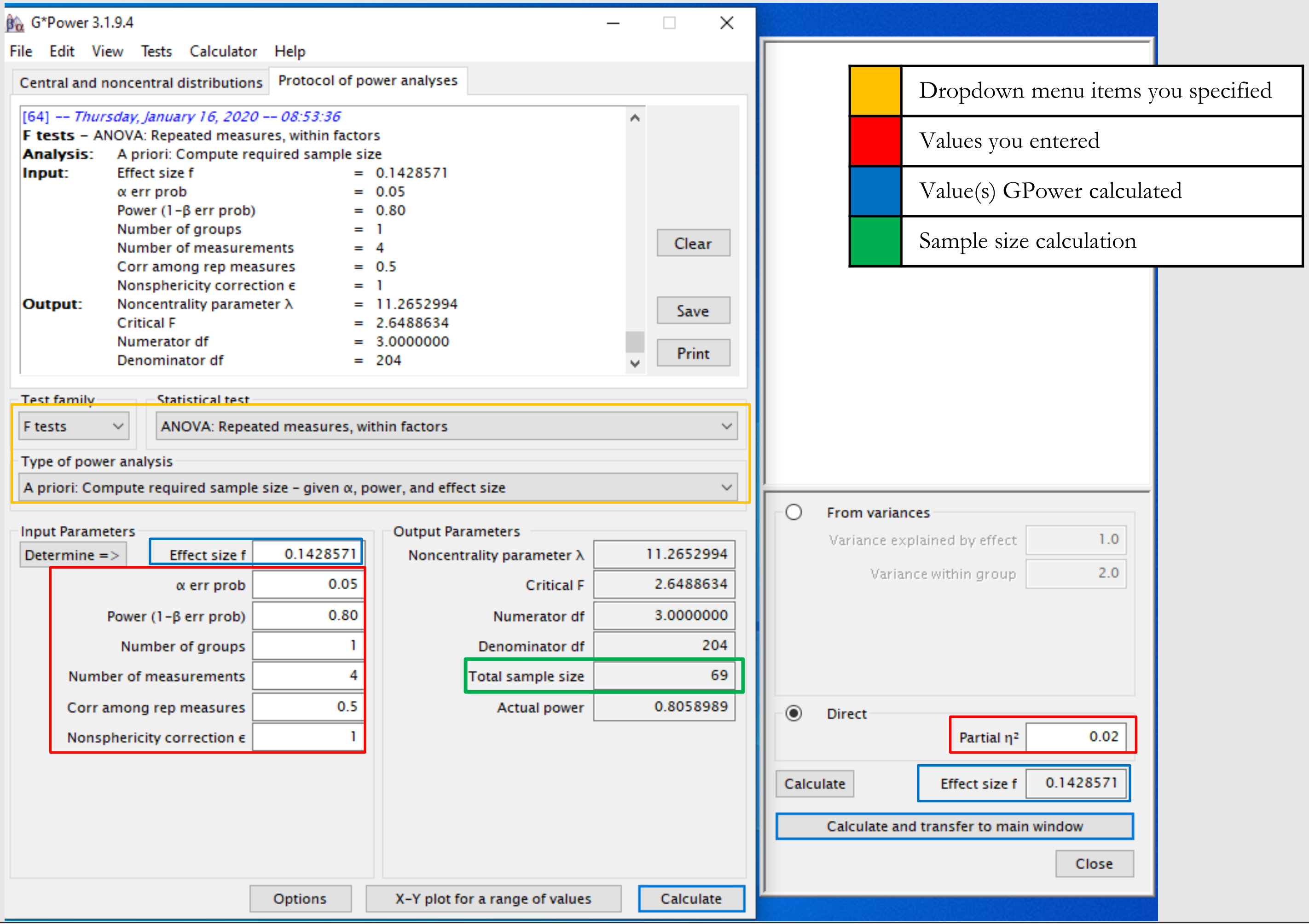
Repeated Measures ANOVA: Example

Is there a difference in blood pressure at 1, 2, 3, and 4 months post-treatment?

- $H_0=0$, $H_1 \neq 0$
- 1 group, 4 measurements
- Have no background info
- Assume 0.5 correlation and Nonsphericity correction of 1.0
- Assume **small** partial eta-squared (0.02)

Results:

- A total number of 69 samples are needed (each getting 4 measurements).



Repeated Measures ANOVA: Practice

Calculate the sample size for the following scenarios (with α=0.05, and power=0.80):

1. You are interested in determining if there is a difference in blood serum levels at 6, 12, 18, and 24 months post-treatment. You collect the following trial data of blood serum in mg/dL (shown on right).
2. You are interested in determining if there is a difference in antibody levels at 1, 2, and 3 months post-treatment.
   - **Info**: no background info, but expect nonsphericity correction of **1**, correlation of **0.5**, and **medium** eta squared (remember to select ***as in GPower 3.0*** in options)

6 months	12 months	18 months	24 months
38	38	46	52
13	44	15	29
32	35	53	60
35	48	51	44
21	27	29	36

Repeated Measures ANOVA: Answers

Calculate the sample size for the following scenarios (with α=0.05, and power=0.80):

1. You are interested in determining if there is a difference in blood serum levels at 6, 12, 18, and 24 months post-treatment? You collect the following trial data of blood serum in mg/dL (shown on right).
   - **1 group, 4 measurements**
   - **Ran Repeated measures in SPSS with data to get $\eta 2 = SS_{eff} / (SS_{eff} + SS_{err})$**
   - **$\eta 2$ = 19531.2 / (19531.2 + 1789.5) = 0.92 (selected *as in SPSS* in options)**
   - **Sphericity was non-significant (p=0.712), so nonsphericity correction is 1.0**
   - **Got effect size of 3.39 for a total sample size of 3**
2. You are interested in determining if there is a difference in antibody levels at 1, 2, and 3 months post-treatment?
   - **1 group, 3 measurements**
   - **Guessed a medium eta squared (0.06), selected *as in GPower 3.0* in options**
   - **Set correlation to default of 0.5 and nonsphericity correction to 1.0**
   - **Got effect size of 0.25 for a total sample size of 27**

6 months	12 months	18 months	24 months
38	38	46	52
13	44	15	29
32	35	53	60
35	48	51	44
21	27	29	36

Friedman Test: Overview

Description: this tests if at least one mean is different among groups, where the groups are repeated measures (more than two) for a non-normally distributed variable. The Friedman test is the extension of the Two Means Wilcoxon test for more than two groups.

Numeric. Var(s)	Cat. Var(s)	Cat. Var Group #	Cat Var. # of Interest	Parametric	Paired
1	**1**	**>2**	**1**	**No**	**Yes**

Example:

- Is there a difference in taste preference across three different desserts for a group of students?
- $H_0=0, \quad H_1\neq0$

GPower:

- Select **F tests** from Test family
- Select **ANOVA: Repeated measures, within factors** from Statistical test
- Select **A priori** from Type of power analysis
- Background Info or Naïve:

a) Enter 0.05 in **α err prob** box – or a specific **α** you want for your study
b) Enter 0.80 in **Power (1-β err prob)** box - or a specific power you want for your study
c) Enter the number of **groups** and **measurements**
d) Enter the **Nonsphericity correction ε**
e) Hit Determine =>
f) Enter **Partial η^2** (eta squared)
   - for ***Background,*** enter calculated eta squared and select **as in SPSS** for the **Effect size specification** in the **Options** bar
   - for ***Naïve***, enter guess eta squared (based on convention) and keep the Effect size specification on **as in Gpower 3.0** and will also have to enter a guess for the **Corr among rep measures** parameter
g) Hit **Calculate and transfer to main window**
h) Hit Calculate on the main window
i) Find **Total sample size** in the Output Parameters
j) Add 15% to size (Total + Total*0.15)

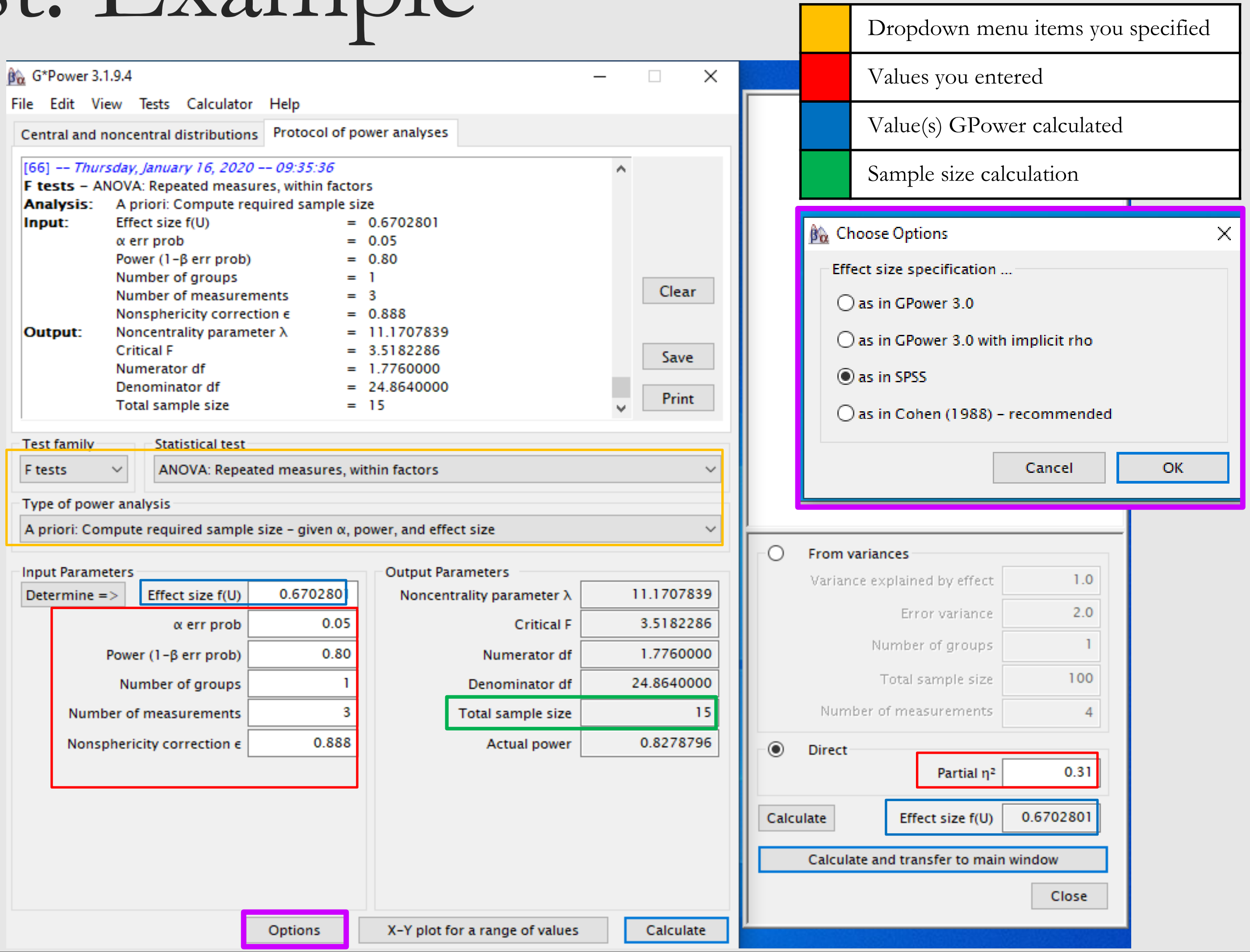
Friedman Test: Example

Is there a difference in taste preference across three different desserts for a group of students?

- $H_0=0, \quad H_1 \neq 0$
- 1 group, 3 measurements
- From a trial study, you found a sphericity of 0.888
- The eta-squared was
  - $SS_{eff} \,/\, (SS_{eff} + SS_{err})$
  - $18 \,/\, (18 + 40) = 0.31$
- Make sure to select the correction **Option**

Results:

- A total number of 15 samples are needed for parametric
- Non-parametric:
  - 15 + 15*0.15 =17.25 (round up) -> 18
  - A total of 18 samples are needed (each getting 3 measurements).



Friedman Test: Practice

Calculate the sample size for the following scenarios (with α=0.05, and power=0.80):

1. You are interested in determining if there is a difference in taste preference across four different main courses for a group of students. You collect the following trial data of food preferences(shown on right).
2. You are interested in determining if there is a difference in movie preference across five different genres for a group of students.

steak	lobster	burgers	noodles
9	1	6	6
8	8	9	4
9	8	10	7
5	9	5	2
8	2	7	3

Friedman Test: Answers

Calculate the sample size for the following scenarios (with α=0.05, and power=0.80):

1. You are interested in determining if there is a difference in taste preference across four different main courses for a group of students. You collect the following trial data of food preferences(shown on right).
   - 1 group, 4 measurements
   - Ran Repeated measures in SPSS with data to get $\eta 2 = SS_{eff} / (SS_{eff} + SS_{err})$
   - $\eta 2$ = 793.8 / (793.8 + 36.7) = 0.95 (selected *as in SPSS* in options)
   - Sphericity was non-significant (p=0.165), so nonsphericity correction is 1.0
   - Got effect size of 4.35 for a parametric sample size of 3
   - 3*1.15=3.45 -> round up to total sample size of 4
2. You are interested in determining if there is a difference in movie preference across five different genres for a group of students.
   - 1 group, 5 measurements
   - Guessed a large eta squared (0.14), selected *as in GPower 3.0* in options
   - Set correlation to default of 0.5 and nonsphericity correction to 1.0
   - Got effect size of 0.40 for a parametric sample size of 9
   - 9*1.15=10.35 -> round up to total sample size of 11

steak	lobster	burgers	noodles
9	1	6	6
8	8	9	4
9	8	10	7
5	9	5	2
8	2	7	3

Multi-Way ANOVA (1 Category of interest): Overview

Description: this test is an extension of ANOVA, where there is more than one category, but only one category is of interest. The other category/categories are things that need to be controlled for (blocking/nesting/random effects/etc.).

Numeric. Var(s)	Cat. Var(s)	Cat. Var Group #	Cat Var. # of Interest	Parametric	Paired
1	≥2	≥2	1	Yes	No

Example:

- Is there difference in treatment (Drug A, B, and C) from a series of four different hospital sections (Block 1, 2, 3, and 4)?
- $H_0=0$, $H_1 \neq 0$

GPower:

- Select **F tests** from Test family
- Select **ANOVA: Fixed effects, special, main effects and interactions** from Statistical test
- Select **A priori** from Type of power analysis

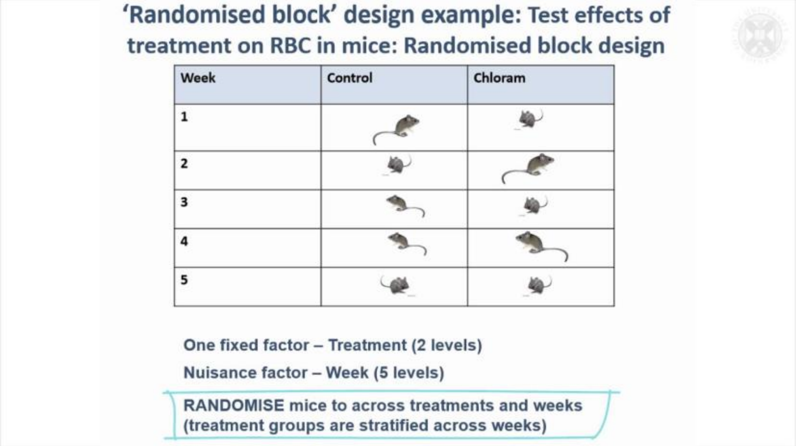
- Background Info:
  a) Enter 0.05 in **α err prob** box – or a specific α you want for your study
  b) Enter 0.80 in **Power (1-β err prob)** box - or a specific power you want for your study
  c) Enter the **number of groups** and **Numerator df** (df for the category of interest only)
  d) Hit Determine =>
  e) Enter calculated **Partial** η^2 (eta squared)
  f) Hit **Calculate and transfer to main window**
  g) Hit Calculate on the main window
  h) Find **Total sample size** in the Output Parameters
- Naïve:
  a) Follow steps a-c as above
  b) Estimate **effect size f**
  c) Hit Calculate on the main window
  d) Find **Total sample size** in the Output Parameters

Multi-Way ANOVA: Parameters and how to calculate them

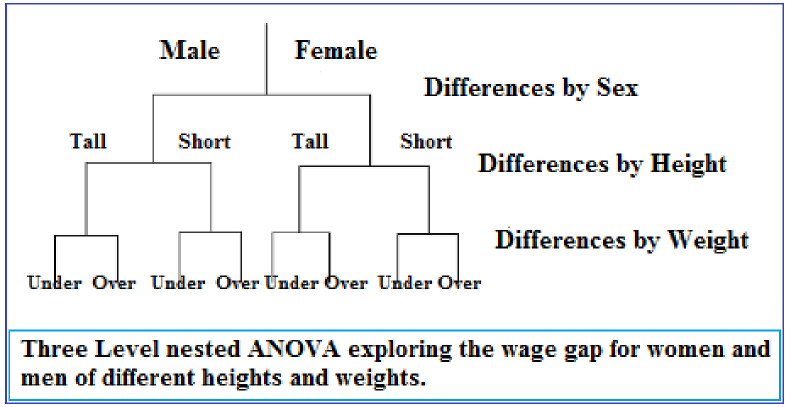
- **Numerator df** = degrees of freedom for the effect you want to test
  - Degrees of freedom found by taken the number of categories for the predictor variable or interaction and subtracting one
  - For interactions between variables, the degrees of freedom need to be multiplied together
  - For example, in a design with three seating locations and two genders:
    - For the predictor variable seating location, enter 3 (locations) – 1 = **2 df**
    - For the predictor variable gender, enter 2 (genders) – 1 = **1 df**
    - For the interaction, enter (3 – 1) * (2 – 1) = **2 df**
  - Since we only want to look at the effect of a single category: just use that category's DF
    - If in the sample above, we only care about seating (we included gender to control for the difference b/t genders) = **2df**
- **Number of groups** = found by multiplying the number of levels in both predictor variables. In this example there are 2 genders and 3 seating locations, so 2 x 3 = 6

Multi-Way ANOVA: Categories to control for

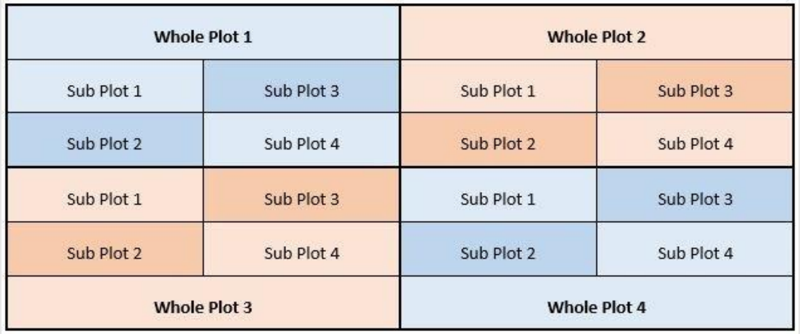
- When you are running a multi-way ANOVA for only 1 category of interest, that means you only care about one category (say drug type), but there are other factors that might interfere with your ability to detect that category because those categories have variation among the groups
- Examples of interfering factors: gender, location, age, etc.
- Below are common classes of effects that are controlled for in ANOVA
  - **Blocking**: samples come from different blocks (locations) –ex. Nutrient treatment types across different fields
  - **Nesting**: samples are located within higher order variables –ex. Leaves are nested within branch
  - **Split-Plot**: More complicated version of blocking with two levels of experimental units – ex. Pesticide treatment in greenhouse trays across whole and subplots



https://i.ytimg.com/vi/gKkMOaANugI/maxresdefault.jpg



https://www.statisticshowto.datasciencecentral.com/wp-content/uploads/2015/07/nested-anova.png



https://blog.minitab.com/hubfs/Imported_Blog_Media/split_plot_3.jpg

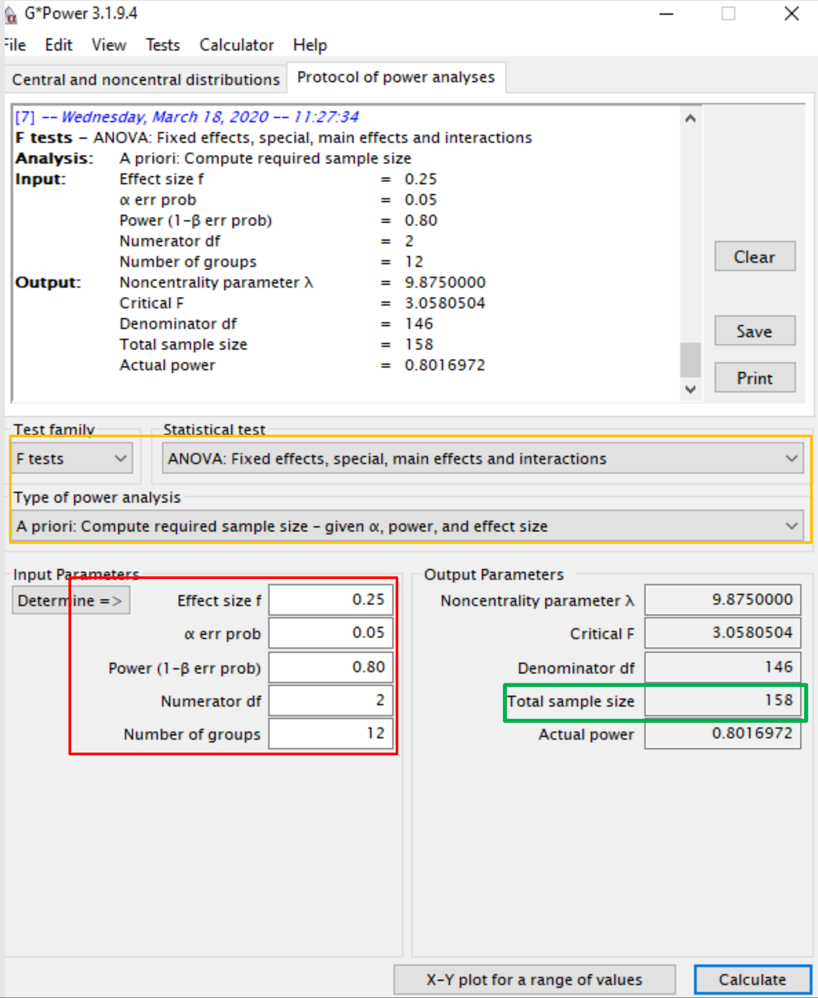
Multi-Way ANOVA (1 Category of interest): Example

Is there difference in treatment (Drug A, B, and C) from a series of four different hospital sections (Block 1, 2, 3, and 4)?

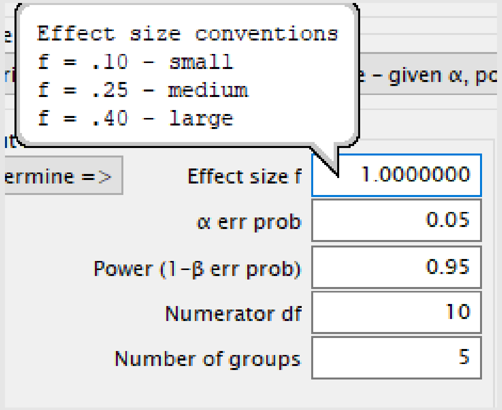
- H_0=0, H_1≠0
- Category of interest: Treatment
- Want to control for the Sections (Blocking)
- No background information
- Assume **medium** effect size
- Numerator df (Treatment) = 3-1=2
- Number of groups (Treatment * Sections) = 3*4 =12

Results:

- A total number of 158 samples are needed, so 13.17 per group, rounding up to 14 per group (168 total).



(yellow)	Dropdown menu items you specified
(red)	Values you entered
(blue)	Value(s) GPower calculated
(green)	Sample size calculation



Multi-Way ANOVA (>1 Category of interest): Overview

Description: this test is an extension of ANOVA, where there is more than one category, and each category is of interest. If there is two categories, it is 2-way ANOVA; three categories, 3-way ANOVA, etc.

Numeric. Var(s)	Cat. Var(s)	Cat. Var Group #	Cat Var. # of Interest	Parametric	Paired
1	≥2	≥2	>1	Yes	No

Example:

- Is there difference in treatment (Drug A, B, and C) across age (child, adult, elder) and cancer stage (I, II, III, IV, V)?
- $H_0=0$, $H_1 \neq 0$

GPower:

- Select **F tests** from Test family
- Select **ANOVA: Fixed effects, specia, main effects and interactions** from Statistical test
- Select **A priori** from Type of power analysis

- Background Info:
  a) Enter 0.05 in **α err prob** box – or a specific **α** you want for your study
  b) Enter 0.80 in **Power (1-β err prob**) box - or a specific power you want for your study
  c) Enter the **number of groups** and **Numerator df** (df for all categories of interest)
  d) Hit Determine =>
  e) Enter calculated **Partial** η^2 (eta squared)
  f) Hit **Calculate and transfer to main window**
  g) Hit Calculate on the main window
  h) Find **Total sample size** in the Output Parameters
- Naïve:
  a) Follow steps a-c as above
  b) Estimate **effect size f**
  c) Hit Calculate on the main window
  d) Find **Total sample size** in the Output Parameters

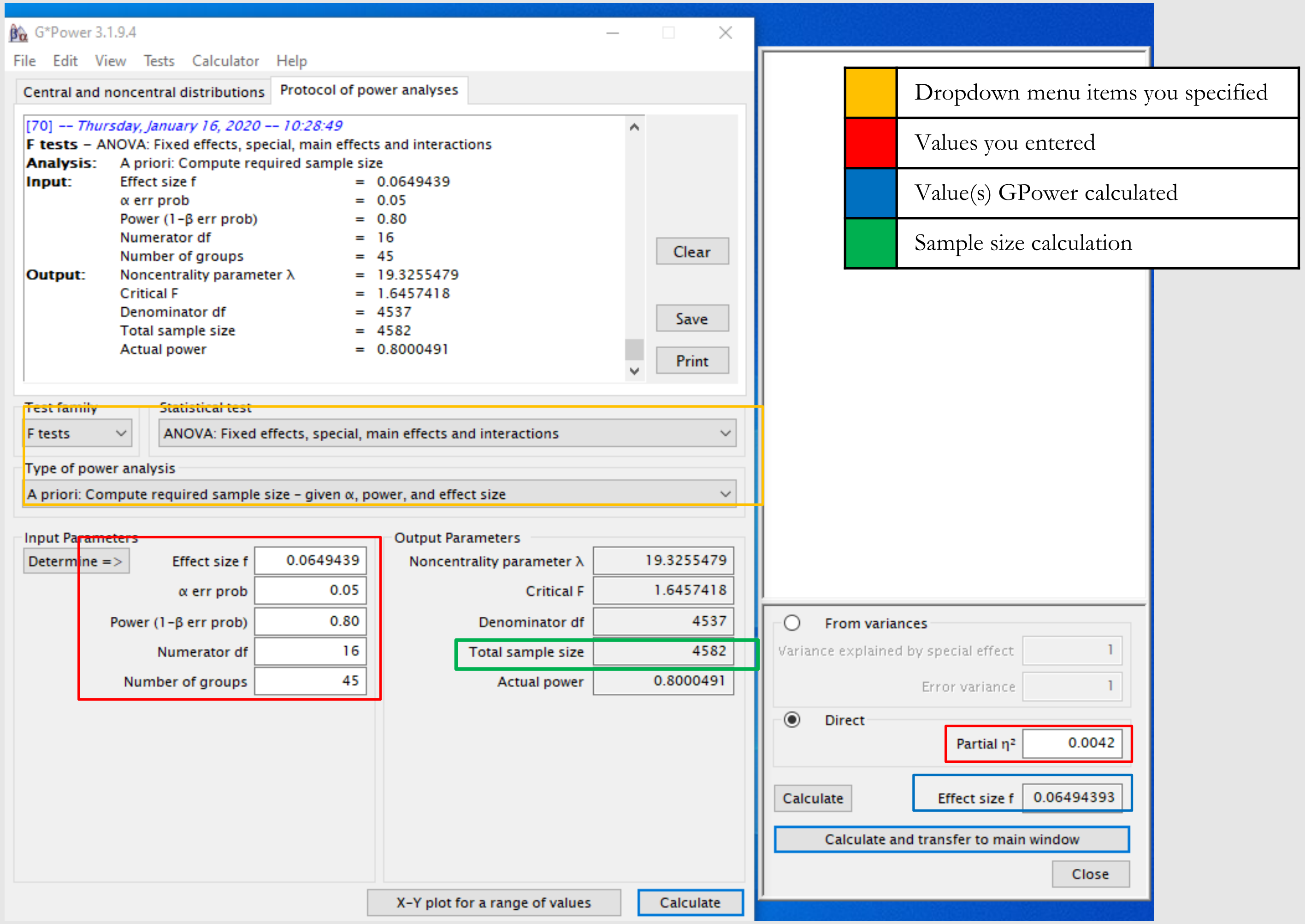
Multi-Way ANOVA (>1 Category of interest): Example

Is there difference in treatment (Drug A, B, and C) across age (child, adult, elder) and cancer stage (I, II, III, IV, V)?

- $H_0=0$, $H_1 \neq 0$
- Categories of interest: Treatment, Age, and Caner Stage
- Based on trial study, the **partial eta-squared=0.0042**
- Numerator df = Treat DF * Age DF * Stage DF = (3-1)*(3-1)*(5-1)=2*2*4=16
- Number of groups = Treat*Age*Stage = 3*3*5=45

Results:

- A total number of 4582 samples are needed, so 101.82 per group, rounding up to 102 (4590 total).



Multi-Way ANOVA: Non-Parametric

- There really isn't a non-parametric option like in the one-way or repeated measures ANOVA
- Two options:
  - Turn to a more sophisticated model for your statistical test, like a Generalized Linear Mixed Model (explained later)
  - Run a parametric test, then add the extra 15% at the end like we've done before for simpler non-parametric ANOVA paralogs

Non-Parametric Reanalysis				
Test	**Question**	**Parametric Sample size**	**Calculations**	**Non-parametric Sample size**
Multi-Way, 1 category of interest	Is there difference in treatment (Drug A, B, and C) from a series of four different hospital sections (Block 1, 2, 3, and 4)?	**251**	251+251*0.15=288.65 288.65/12(groups)=24.05 -> 25 25*12(groups)=300	**300**
Multi-Way, 1 category of interest	Is there difference in treatment (Drug A, B, and C) across age (child, adult, elder) and cancer stage (I, II, III, IV, V)?	**4582**	4582+4582*0.15=5269.3 5269.3/45(groups)=117.09 ->118 118*45(groups) = 5310	**5310**

Multi-Way ANOVA: Practice

Calculate the sample size for the following scenarios (with α=0.05, and power=0.80):

1. You are interested in determining if there is a difference in treatment (Drug A, B, and C), while controlling for age (child, adult, elder). You collect the following trial data for treatment (shown on right).
2. You are interested in determining if there is a difference in treatment (Drug A, B, and C) across age (child, adult, elder) and cancer stage (I, II, III, IV, V).

Drug A			Drug B			Drug C		
child	adult	elder	child	adult	elder	child	adult	elder
-6.4	8.7	-3.1	1.3	-6.0	6.8	-2.0	-4.3	-1.2
-8.2	-6.3	-6.5	3.6	1.3	2.4	1.5	1.3	1.1
7.9	-1	-1.5	3.9	-1.9	1.3	2.48	-8.2	-9.7

Multi-Way ANOVA: Answers

Calculate the sample size for the following scenarios (with α=0.05, and power=0.80):

1. You are interested in determining if there is a difference in treatment (Drug A, B, and C), while controlling for age (child, adult, elder). You collect the following trial data for treatment (shown on right).
   - **Only care about treatment so numerator df is 3-1=2**
   - **Number of groups is 3*3=9**
   - **Ran Univariate ANOVA in SPSS with data to get $\eta 2 = SS_{eff} / (SS_{eff} + SS_{err})$**
   - **$\eta 2 = 68.8 / (68.8 + 96.3) = 0.416$**
   - **Got effect size of 0.84 for a total sample size of 19**
2. You are interested in determining if there is a difference in treatment (Drug A, B, and C) across age (child, adult, elder) and cancer stage (I, II, III, IV, V).
   - **Care about treatment, age, and cancer stage**
   - **Numerator df = (3-1)*(3-1)*(5-1)=2*2*4=16**
   - **Number of groups is 3*3*5=45**
   - **Guessed a medium effect size (0.25) for a total sample size of 323**

Drug A			Drug B			Drug C		
child	adult	elder	child	adult	elder	child	adult	elder
-6.4	8.7	-3.1	1.3	-6.0	6.8	-2.0	-4.3	-1.2
-8.2	-6.3	-6.5	3.6	1.3	2.4	1.5	1.3	1.1
7.9	-1	-1.5	3.9	-1.9	1.3	2.48	-8.2	-9.7

Proportion Test: Overview

Description: this tests when you only have a single categorical value with only two groups, and you want to know if the proportions of certain values differ from some constant proportion. (binomial, AKA=Yes/No, 1/0, etc.)

Numeric. Var(s)	Cat. Var(s)	Cat. Var Group #	Cat Var. # of Interest	Parametric	Paired
0	1	2	1	N/A	N/A

Example:

- Is there a significance difference in cancer prevalence of middle-aged women who have a sister with breast cancer compared to the general population prevalence (proportion=0.02, or 2%)?
- $H_0=0$, $H_1 \neq 0$

GPower:

- Select **Exact** from Test family
- Select **Proportion: Difference from constant (binomial test, one sample case)** from Statistical test
- Select **A priori** from Type of power analysis

- Background Info or Naïve:
  a) Select **One** or **Two** Tail(s) as appropriate to your question
  b) Enter 0.05 in **α err prob** box – or a specific **α** you want for your study
  c) Enter 0.80 in **Power (1-β err prob**) box - or a specific power you want for your study
  d) Enter the **constant proportion** (whatever constant proportion you are measuring your proportion of interest against)
  e) Hit Determine =>
  f) Select **Difference P2-P1** from the **Calc P2 from** option (doesn't really matter, all give you about the same effect size)
  g) Put the **constant proportion** in the **P1 box**
  h) Put the **alternative proportion** in the **P2 box:**
     a) For ***Background***: enter the value from your background information
     b) For ***Naïve***: guess the value and enter it in
        - 0.2=small, 0. 5=medium, and 0.8 large effect sizes
  i) Hit **Sync Values**, then **Calculate and transfer to main window**
  j) Hit Calculate on the main window
  k) Find **Total sample size** in the Output Parameters

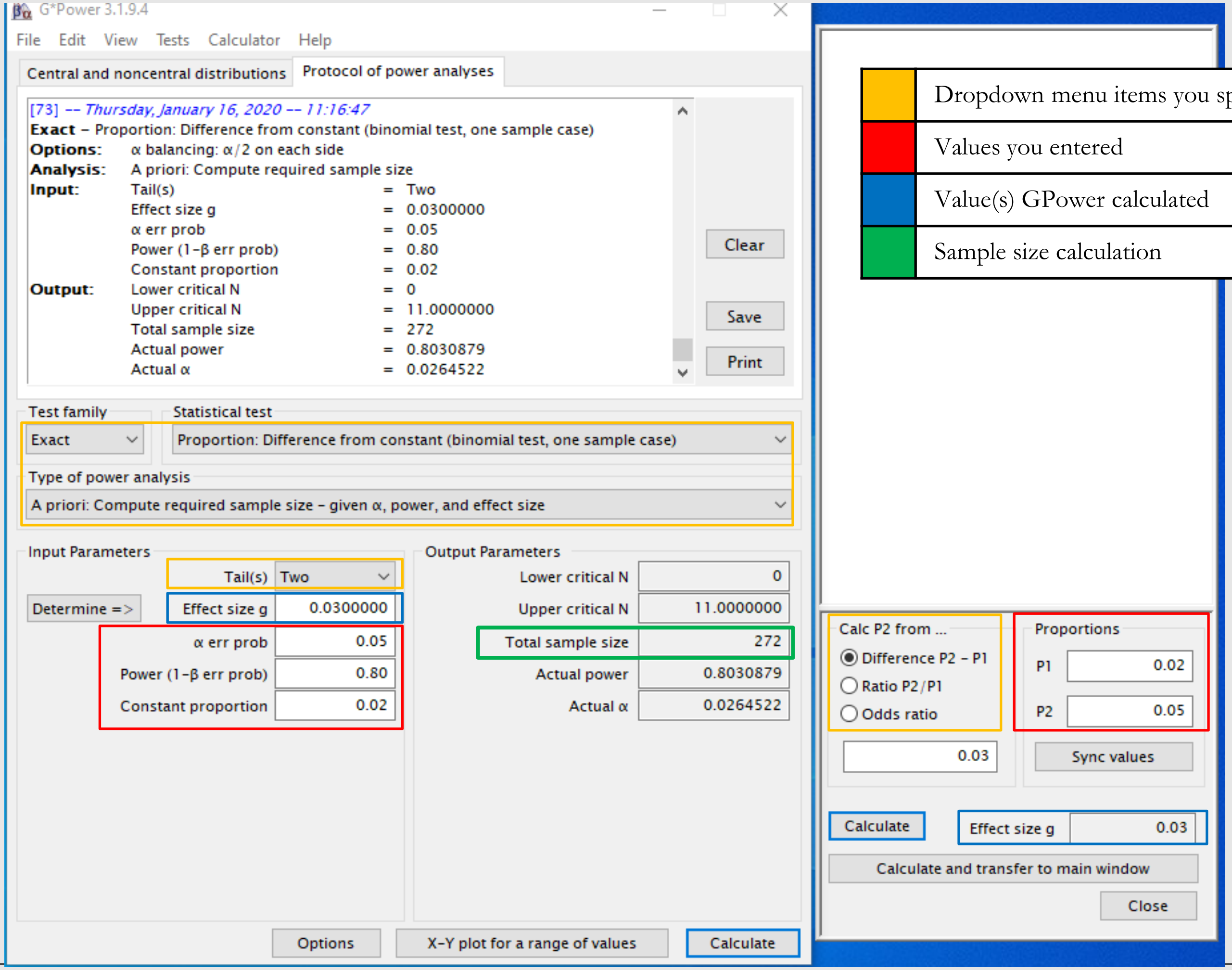
Proportion Test: Example

Is there a significance difference in cancer prevalence of middle-aged women who have a sister with breast cancer compared to the general population prevalence (proportion=0.02, or 2%)?

- $H_0=0$, $H_1 \neq 0$
- Based on background information, we found the prevalence (proportion) of women w/ sister who has breast cancer to get breast cancer was 0.05.
- Want to use two-tailed, because only care if there is a difference, not directionality

Results:

- A total number of 272 samples are needed.



Yellow	Dropdown menu items you specified
Red	Values you entered
Blue	Value(s) GPower calculated
Green	Sample size calculation

Proportion Test: Practice

Calculate the sample size for the following scenarios (with α=0.05, and power=0.80):

1. You are interested in determining if the male incidence rate proportion of cancer in North Dakota is higher than the US average (prop=0.00490). You find trial data cancer prevalence of 0.00495.
2. You are interested in determining if the female incidence rate proportion of cancer in North Dakota is lower than the US average (prop=0.00420).

Proportion Test: Answers

Calculate the sample size for the following scenarios (with α=0.05, and power=0.80):

1. You are interested in determining if the male incidence rate (proportion per 100,000) of cancer in North Dakota is higher than the US average (prop=0.00490). You find trial data cancer prevalence of 0.00495.
   - **The difference in P2-P1 is 0.00005 (constant proportion is 0.0049)**
   - **Got effect size of 0.00005 and used a one-tailed test for a total sample size of 12,113,157**
2. You are interested in determining if the female incidence rate proportion of cancer in North Dakota is lower than the US average (prop=0.00420).
   - **Guessed a very small effect size (0.0001) and entered a constant proportion of 0.00420**
   - **Used a one-tailed test for a total sample size of 2,490,591**

Fisher's Exact Test: Overview

Description: this test when you only have 2 categorical variables and you want to know if the proportions are different between groups, where the groups are not related. Essentially testing between an observed proportion and an expected one. Also called independent Chi-Squared test.

Numeric. Var(s)	Cat. Var(s)	Cat. Var Group #	Cat Var. # of Interest	Parametric	Paired
0	2	2	2	N/A	No

Example:

- Is the expected proportion of students passing a stats course taught by psychology teachers of 0.85 different from the observed proportion of students passing the same stats class taught by mathematics teachers of 0.95
- $H_0=0$, $H_1 \neq 0$

GPower:

- Select **Exact** from Test family
- Select **Proportion: Inequality, two independent groups (Fisher's exact test)** from Statistical test
- Select **A priori** from Type of power analysis
- Background Info or Naïve:
  - a) Select **One** or **Two** Tail(s) as appropriate to your question
  - b) Enter values for **Proportion p1 and p2**
    - a) P1 is the 'expected value'
    - b) P2 is the 'observed value' – if you don't have background info, you will have to guess on this
  - c) Enter 0.05 in **α err prob** box – or a specific **α** you want for your study
  - d) Enter 0.80 in **Power (1-β err prob**) box - or a specific power you want for your study
  - e) Enter 1.0 in **Allocation ration N2/N1** – unless the group sizes are different
  - f) Hit Calculate on the main window
  - g) Find **Total sample size** in the Output Parameters

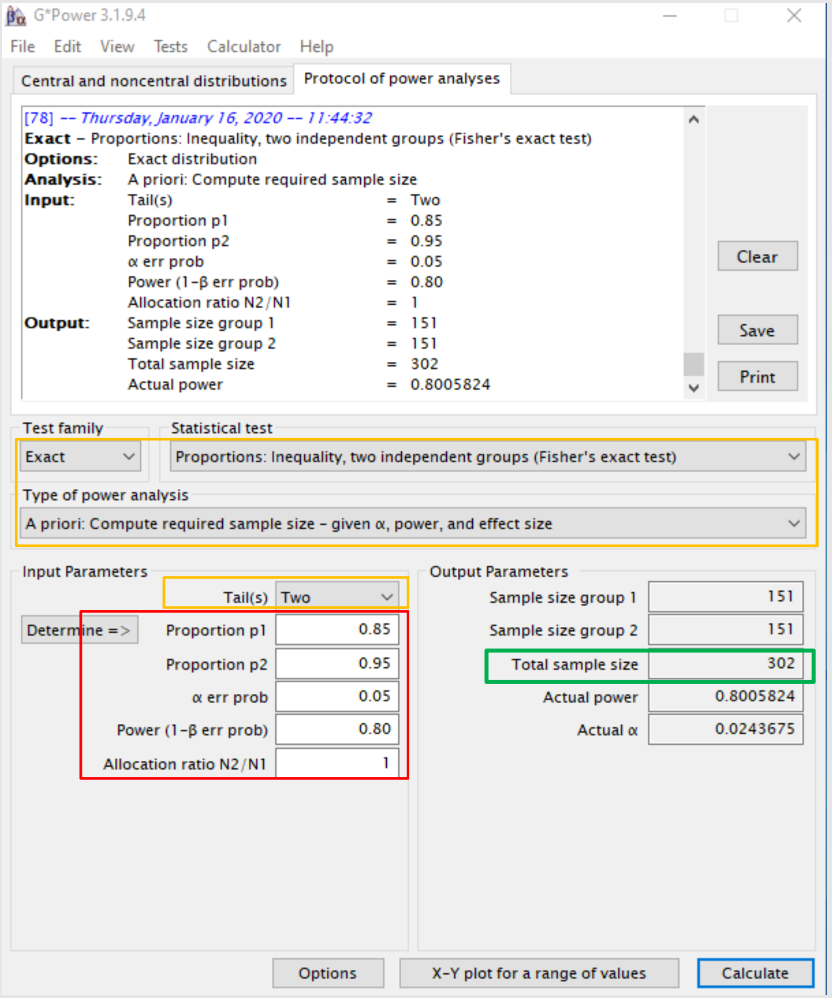
Fisher's Exact Test: Example

Is the expected proportion of students passing a stats course taught by psychology teachers of 0.85 different from the observed proportion of students passing the same stats class taught by mathematics teachers of 0.95?

- $H_0=0$, $H_1 \neq 0$
- P1 is the expected value of 0.85
- P2 is the observed value of 0.95
- Want to use two-tailed, because only care if there is a difference, not directionality

Results:

- A total number of 302 samples are needed.



Yellow	Dropdown menu items you specified
Red	Values you entered
Blue	Value(s) GPower calculated
Green	Sample size calculation

Fisher's Exact Test: Practice

Calculate the sample size for the following scenarios (with α=0.05, and power=0.80):

1. You are interested in determining if the expected proportion (P1) of students passing a stats course taught by psychology teachers is different than the observed proportion (P2) of students passing the same stats class taught by biology teachers. You collected the following data of passed tests. You also know that twice as many students take the psychology class than the biology one.

Psychology	Yes	Yes	Yes	No	No	Yes	Yes	Yes	Yes	No
Biology	No	No	Yes	Yes	Yes	No	Yes	No	Yes	Yes

1. You are interested in determining of the expected proportion (P1) of female students who selected YES on a question was higher than the observed proportion (P2) of male students who selected YES. The observed proportion of males who selected yes was 0.75.

Fisher's Exact Test: Practice

Calculate the sample size for the following scenarios (with α=0.05, and power=0.80):

1. You are interested in determining if the expected proportion (P1) of students passing a stats course taught by psychology teachers is <u>different</u> than the observed proportion (P2) of students passing the same stats class taught by biology teachers. You collected the following data of passed tests. You also know that twice as many students take the psychology class than the biology one.

Psychology	Yes	Yes	Yes	No	No	Yes	Yes	Yes	Yes	No
Biology	No	No	Yes	Yes	Yes	No	Yes	No	Yes	Yes

   - **Psychology proportion is 7/10=0.70; biology proportion is 6/10=0.60; allocation ratio is 0.5**
   - **For a two-tailed test, the total sample size is 812 (541 for psychology class, 271 for biology class)**

2. You are interested in determining of the expected proportion (P1) of female students who selected YES on a question was <u>higher</u> than the observed proportion (P2) of male students who selected YES. The observed proportion of males who selected yes was 0.75.
   - **Don't have any info on the female students, but will guess that it is 0.10 higher (P1=0.85)**
   - **For a one-tailed test, the total sample size is 430 (215 each male and female)**

McNamar’s Test: Overview

Description: this test when you only have 2 categorical variables and you want to know if the proportions are different between groups, where the groups are related (paired). Still testing between an observed proportion and an expected one. Also called dependent Chi-Squared test.

Numeric. Var(s)	Cat. Var(s)	Cat. Var Group #	Cat Var. # of Interest	Parametric	Paired
0	2	2	2	N/A	Yes

Example:

- Is there a difference in innate immune response (good/poor) for patients taking Vitamin D or not (yes/no)?
- $H_0=0$, $H_1 \neq 0$

GPower:

- Select **Exact** from Test family
- Select **Proportion: Inequality, two dependent groups (McNemar)** from Statistical test
- Select **A priori** from Type of power analysis

- Background Info or Naïve:
  - a) Select **One** or **Two** Tail(s) as appropriate to your question
  - b) Enter value for **Odds ratio**: (see next page)
  - c) Enter 0.05 in **α err prob** box – or a specific **α** you want for your study
  - d) Enter 0.80 in **Power (1-β err prob**) box - or a specific power you want for your study
  - e) Enter value for **Prop discordant pairs:** (see next page)
  - f) Hit **Calculate** on main window
  - g) Find **Total sample size** in the Output Parameters

McNamar’s Test: Proportion table

- To calculate the **Odds Ratio** and **Prop discordant pairs,** it is important to understand what the data looks like
- To the right is trial data from our example: the number are proportions for each category pair
- Note: the proportions must all sum to 1.0
- Discordant pairs are the proportions for the groups that don’t match up
  - Good Immune Response for Vitamin D and Poor response with No Vitamin D
  - Poor Immune Response for Vitamin D and Good response with No Vitamin D
  - The stronger the two categories are related to each other, the smaller the discordant pair proportions should be
- The **Odds Ratio** is the quotient of the discordant pairs
  - 0.02/0.13= 0.1538 (Note: The opposite quotient 0.13 / 0.02 = 6.5 produces the same result in GPower)
- The **Proportion of discordant pairs** is the sum of the discordant pairs
  - 0.02+0.13=0.15

		Vitamin D	
		good	poor
No Vitamin D	good	0.70	0.02
	poor	0.13	0.15

discordant pairs

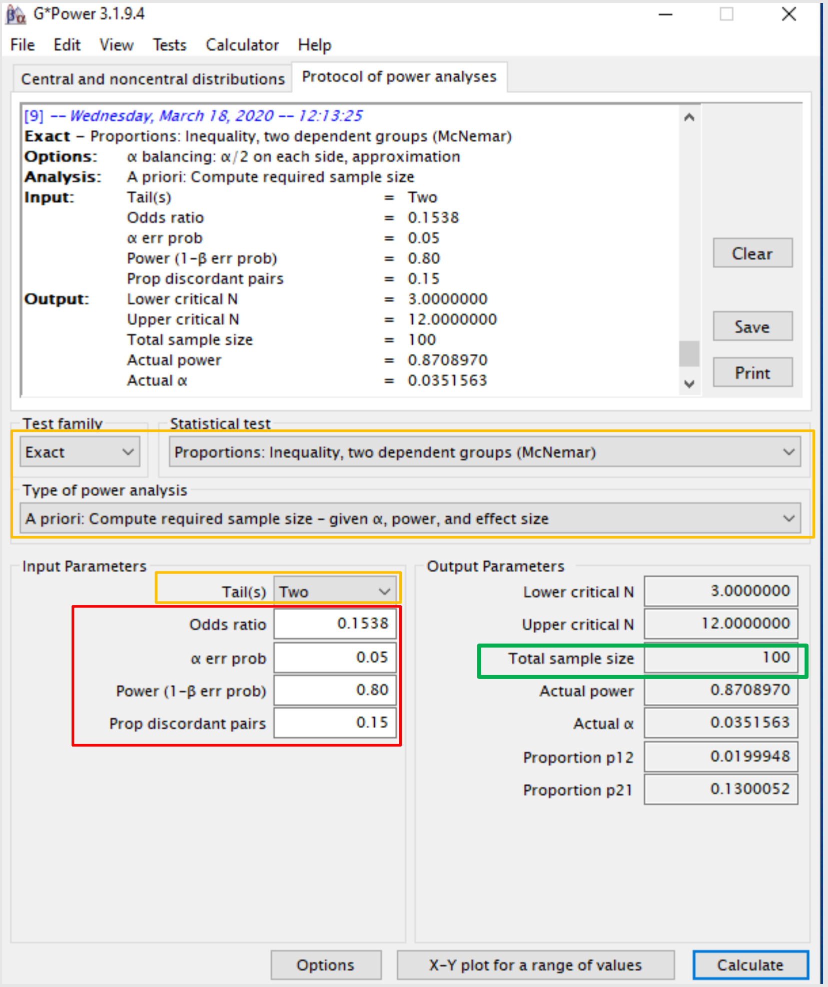
McNamar's Test: Example

Is there a difference in innate immune response (good/poor) for patients taking Vitamin D or not (yes/no)?

- H_0=0, $H_1 \neq 0$
- Based on background information from the previous page,
  - Odds Ratio=0.02/0.13=**0.1538**
  - Prop discordant pairs=0.02+0.13=**0.15**
- Want to use two-tailed, because only care if there is a difference, not directionality

Results:

- A total number of 100 samples are needed.



(yellow)	Dropdown menu items you specified
(red)	Values you entered
(blue)	Value(s) GPower calculated
(green)	Sample size calculation

McNamar's Test: Practice

Calculate the sample size for the following scenarios (with α=0.05, and power=0.80):

1. You are interested in determining if there is a difference in overnight stay (yes/no) for patients having surgery with anesthesia or not (yes/no). You collected the following trial information.

		Anesthesia yes	Anesthesia no
No Anesthesia	yes	0.30	0.10
	no	0.08	0.52

1. You are interested in determining if there is a difference in snacking (yes/no) for students in the afternoon whether they had school or not (yes/no). You collected the following trial information.

Student	1	2	3	4	5	6	7	8	9	10
Snack	Yes	No	Yes	Yes	No	Yes	No	Yes	No	Yes
School	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No

McNamar's Test: Answers

Calculate the sample size for the following scenarios (with α=0.05, and power=0.80):

1. You are interested in determining if there is a difference in overnight stay (yes/no) for patients having surgery with anesthesia or not (yes/no). You collected the following trial information.

		Anesthesia yes	Anesthesia no
No Anesthesia	yes	0.30	0.10
	no	0.08	0.52

- **Odds ratio is 0.08/0.10=0.80; proportion of discordant pairs is 0.08+0.10=0.18**
- **For a two-tailed test, total sample size is 3562**

2. You are interested in determining if there is a difference in snacking (yes/no) for students in the afternoon whether they had school or not (yes/no). You collected the following trial information.

Student	1	2	3	4	5	6	7	8	9	10
Snack	Yes	No	Yes	Yes	No	Yes	No	Yes	No	Yes
School	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No

		School yes	School no
No School	yes	4/10	3/10
	no	2/10	1/10

- **Odds ratio is 0.2/0.3=0.667; proportion of discordant pairs is 0.2+0.3=0.5**
- **For a two-tailed test, total sample size is 398**

Goodness-of-Fit Test: Overview

Description: Extension of Chi-squared tests, which asks if table of observed values are any different from a table of expected ones. Also generically called Chi-Squared.

Numeric. Var(s)	Cat. Var(s)	Cat. Var Group #	Cat Var. # of Interest	Parametric	Paired
0	**≥1**	**≥2**	**1**	**N/A**	**No**

Example:

- Does the observed proportions of phenotypes from a genetics experiment different from the expected 9:3:3:1?
- $H_0=0$, $H_1 \neq 0$

GPower:

- Select **X^2** from Test family
- Select **Goodness-of-fit tests: Contingency tables** from Statistical test
- Select **A priori** from Type of power analysis

- Background Info:
  - a) Enter 0.05 in **α err prob** box – or a specific **α** you want for your study
  - b) Enter 0.80 in **Power (1-β err prob**) box - or a specific power you want for your study
  - c) Enter X in **DF** (degrees of freedom), which = **number of categories -1**
  - d) Hit Determine =>
  - e) Add the number of cells that equals the number of proportions
  - f) Enter the expected proportions in the p(H0) column
  - g) Enter in the observed proportions in the p(H1) column
    - If you have counts rather than proportions, you can enter them in and then click on **Normalize p(H0)** and **Normalize p(H1)**
  - h) Hit **Calculate and transfer to main window**
  - i) Hit Calculate on the main window
  - j) Find **Total sample size** in the Output Parameters
- Naïve:
  - a) Run steps a-c above
  - b) Estimate an **effect size w**
  - c) Hit Calculate on the main window
  - d) Find **Total sample size** in the Output Parameters

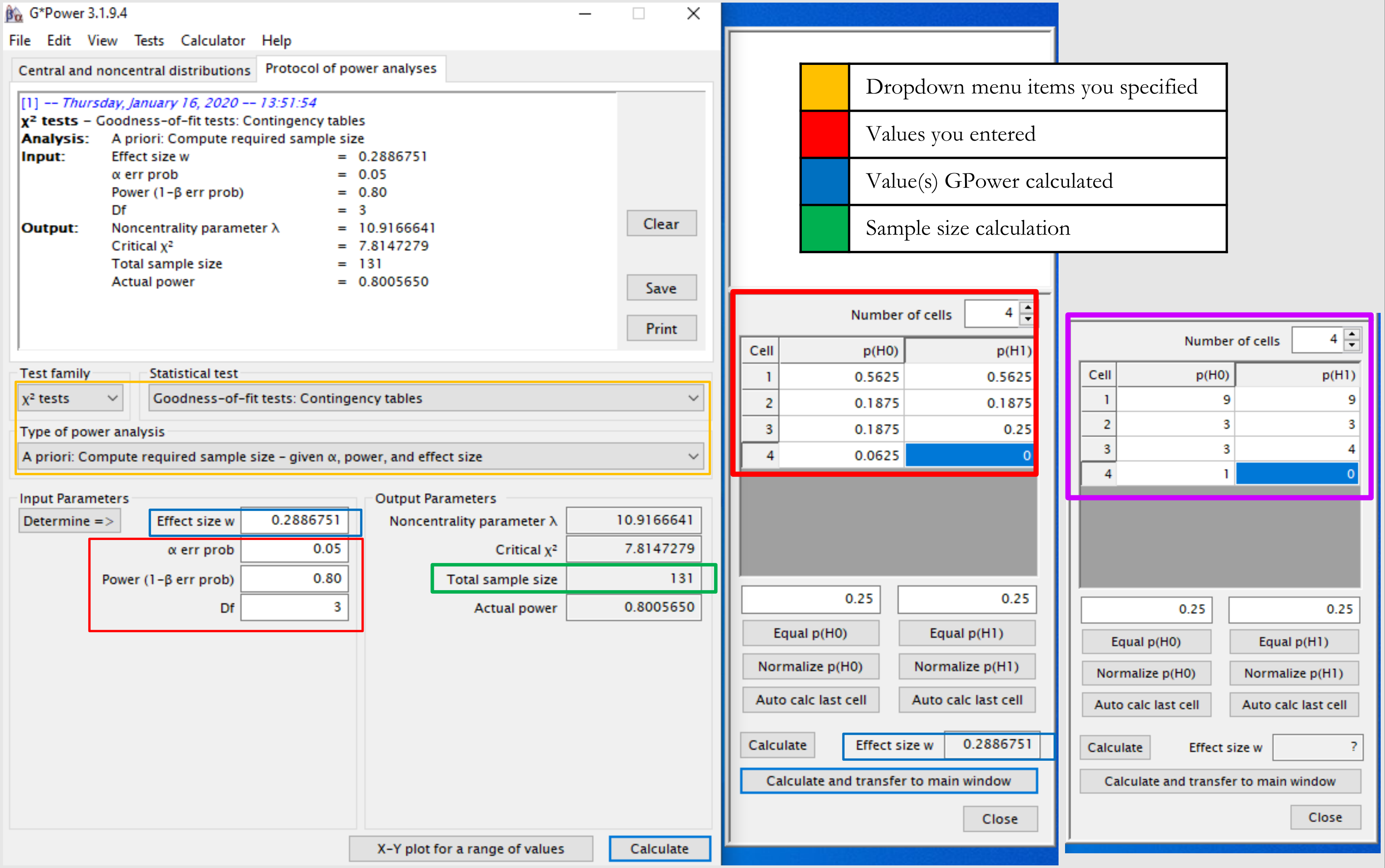
Goodness-of-Fit Test: Example

Does the observed proportions of phenotypes from a genetics experiment different from the expected 9:3:3:1?

- $H_0=0$, $H_1\neq0$
- The observed ratios we'd like to test are 9:3:4:0
- DF=4-1=3
- Notice the change between the **counts** and the **normalized proportions**

Results:

- A total number of 131 samples are needed.



Goodness-of-Fit Test: Practice

Calculate the sample size for the following scenarios (with α=0.05, and power=0.80):

1. You are interested in determining if the ethnic ratios in a company differ by gender. You collect the following trial data.

Gender	White	Black	Am. Indian	Asian
Male	**0.60**	**0.25**	**0.01**	**0.14**
Female	**0.65**	**0.21**	**0.11**	**0.03**

2. You are **interested** in determining if the proportions of student by year (Freshman, Sophomore, Junior, Senior) is any different from 1:1:1:1. You collect the following trial data.

Student	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Grade	**Frs**	**Frs**	**Frs**	**Frs**	**Frs**	**Frs**	**Frs**	**Soph**	**Soph**	**Soph**	**Soph**	**Soph**	**Jun**	**Jun**	**Jun**	**Jun**	**Jun**	**Sen**	**Sen**	**Sen**

Goodness-of-Fit Test: Answers

Calculate the sample size for the following scenarios (with α=0.05, and power=0.80):

1. You are interested in determining if the ethnic ratios in a company differ by gender. You collect the following trial data.

Gender	White	Black	Am. Indian	Asian
Male	0.60	0.25	0.01	0.14
Female	0.65	0.21	0.11	0.03

   - **Got an effect size of 1.04 and degrees of freedom (4-1) of 3; total sample size is 10**

2. You are **interested** in determining if the proportions of student by year (Freshman, Sophomore, Junior, Senior) is any different from 1:1:1:1. You collect the following trial data.

Student	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Grade	Frs	Frs	Frs	Frs	Frs	Frs	Frs	Soph	Soph	Soph	Soph	Soph	Jun	Jun	Jun	Jun	Jun	Sen	Sen	Sen

   - **Proportions were: Freshman=7/20=0.35, Sophomore=5/20=0.25, Junior=5/20=0.25, Senior=3/20=0.15**
   - **Got an effect size of 0.28 and degrees of freedom (4-1) of 3; total sample size is 137**

Simple Linear Regression: Overview

<u>Description</u>: this test determines if there is a significant relationship between two normally distributed numerical variables. The predictor variable is used to try to predict the response variable.

Numeric. Var(s)	Cat. Var(s)	Cat. Var Group #	Cat Var. # of Interest	Parametric	Paired
2	**0**	**N/A**	**N/A**	**Yes**	**N/A**

<u>Example</u>:

- Can height predict weight in college males?
- $H_0=0$, $H_1\neq0$

<u>GPower</u>:

- Select **F tests** from Test family
- Select **Linear multiple regression: Fixed model, R2 deviation from zero** from Statistical test- simple linear regression is just the limit case of multiple regression where there is only 1 predictor variable
- Select **A priori** from Type of power analysis

- Background Info or Naïve:
  - a) Enter 0.05 in **α err prob** box – or a specific α you want for your study
  - b) Enter 0.80 in **Power (1-β err prob**) box - or a specific power you want for your study
  - c) Enter 1 in **Number of predictors**
  - d) Hit Determine =>
  - e) Selected the **From correlation coefficient** option and add **squared multiple correlation**
    - Add the correlation coefficient from background information after squaring it
  - f) Hit **Calculate and transfer to main window**
  - g) Hit Calculate on the main window
  - h) Find **Total sample size** in the Output Parameters
- Naïve:
  - a) Run steps a-c above
  - b) Estimate an **effect size f^2**
  - c) Hit Calculate on the main window
  - d) Find **Total sample size** in the Output Parameters

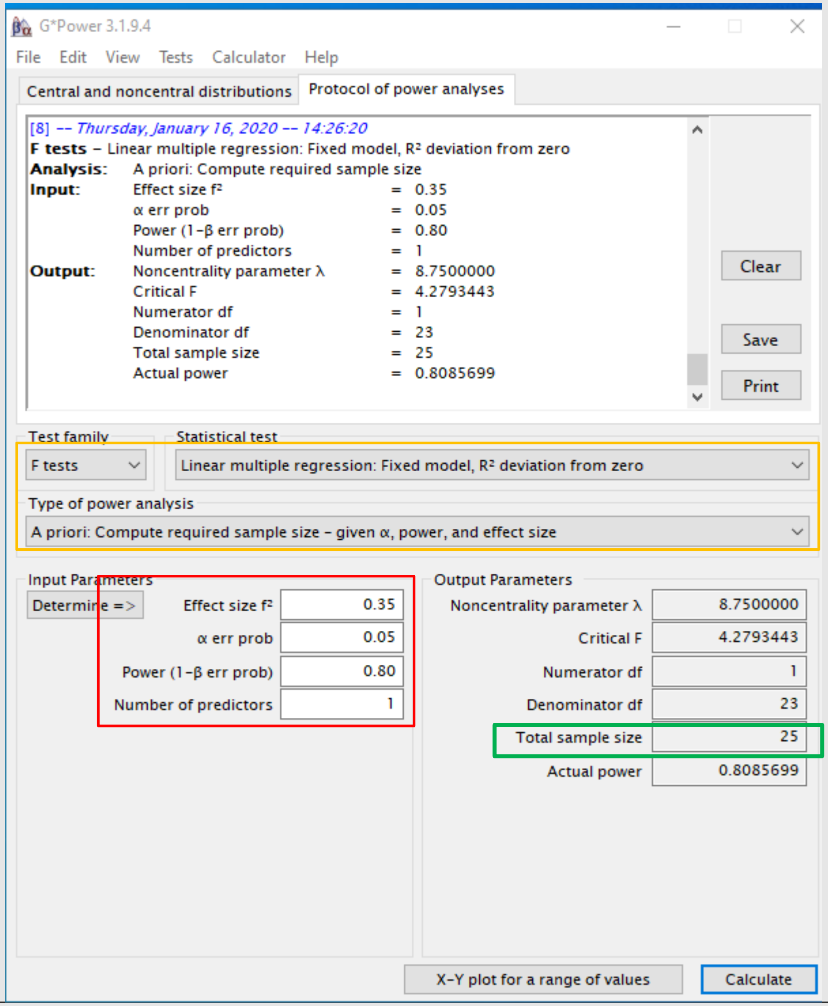
Simple Linear Regression: Example

Is there a relationship between height and weight in college males?

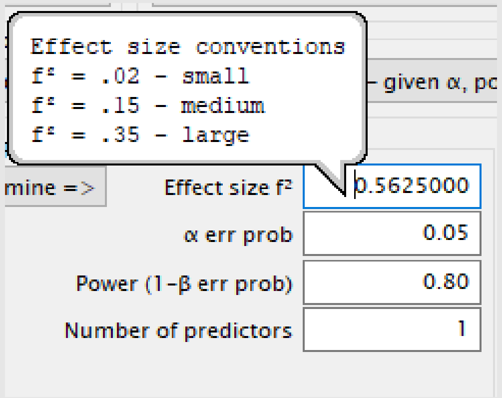
- H_0=0, H_1≠0
- No background, but expect the correlation to be strong between the two, so go with **large** effect size

Results:

- A total number of 25 samples are needed.



Yellow	Dropdown menu items you specified
Red	Values you entered
Blue	Value(s) GPower calculated
Green	Sample size calculation



Simple Linear Regression: Practice

Calculate the sample size for the following scenarios (with α=0.05, and power=0.80):

1. You are interested in determining if height (meters) in plants can predict yield (grams of berries). You collect the following trial data.

Yield	**46.8**	**48.7**	**48.4**	**53.7**	**56.7**
Height	**14.6**	**19.6**	**18.6**	**25.5**	**20.4**

2. You are interested in determining if the size of a city (in square miles) can predict the population of the city (in # of individuals).

Simple Linear Regression: Answers

Calculate the sample size for the following questions:

1. You are interested in determining if height (meters) in plants can predict yield (grams of berries). You collect the following trial data.

Yield	**46.8**	**48.7**	**48.4**	**53.7**	**56.7**
Height	**14.6**	**19.6**	**18.6**	**25.5**	**20.4**

  - **Running a regression found a squared multiple correlation coefficient (R-squared)=0.459**
  - **Got an effect size of 0.85; for 1 predictor, found total sample size of 12**

2. You are interested in determining if the size of a city (in square miles) can predict the population of the city (in # of individuals).

  - **Guessed a large effect size (0.35); for 1 predictors, found total sample size of 25**

Multiple Linear Regression: Overview

Description: The extension of simple linear regression. The first major change is there are more predictor variables. The second change is that interaction effects can be used. Finally, the results typically can't be plotted.

Numeric. Var(s)	Cat. Var(s)	Cat. Var Group #	Cat Var. # of Interest	Parametric	Paired
>2	0	N/A	N/A	Yes	N/A

Example:

- Can height, age, and time spent at the gym, predict weight in adult males?
- $H_0=0$, $H_1\neq0$

GPower:

- Select **F tests** from Test family
- Select **Linear multiple regression: Fixed model, R2 deviation from zero** from Statistical test
- Select **A priori** from Type of power analysis
- Background Info or Naïve:
  a) Enter 0.05 in **α err prob** box – or a specific **α** you want for your study
  b) Enter 0.80 in **Power (1-β err prob**) box - or a specific power you want for your study
  c) Enter the number of predictor variables in **Number of predictors**
  d) Hit Determine =>
  e) Enter correlation coefficient:
     a) From correlation coefficient: enter squared multiple correlation coefficient
     b) From predictor correlations: add number of predictors, then click **Specify matrices**, from there add the correlation coefficients from a correlation matrix of the outcome (response) with all the predictor variables, and finally **Accept values**
  f) Hit **Calculate and transfer to main window**
  g) Hit Calculate on the main window
  h) Find **Total sample size** in the Output Parameters
- Naïve:
  a) Run steps a-c above
  b) Estimate an **effect size f^2**
  c) Hit Calculate on the main window
  d) Find **Total sample size** in the Output Parameters

Multiple Linear Regression: Example

(yellow)	Dropdown menu items you specified
(red)	Values you entered
(blue)	Value(s) GPower calculated
(green)	Sample size calculation

Can height, age, and time spent at the gym, predict weight in adult males?

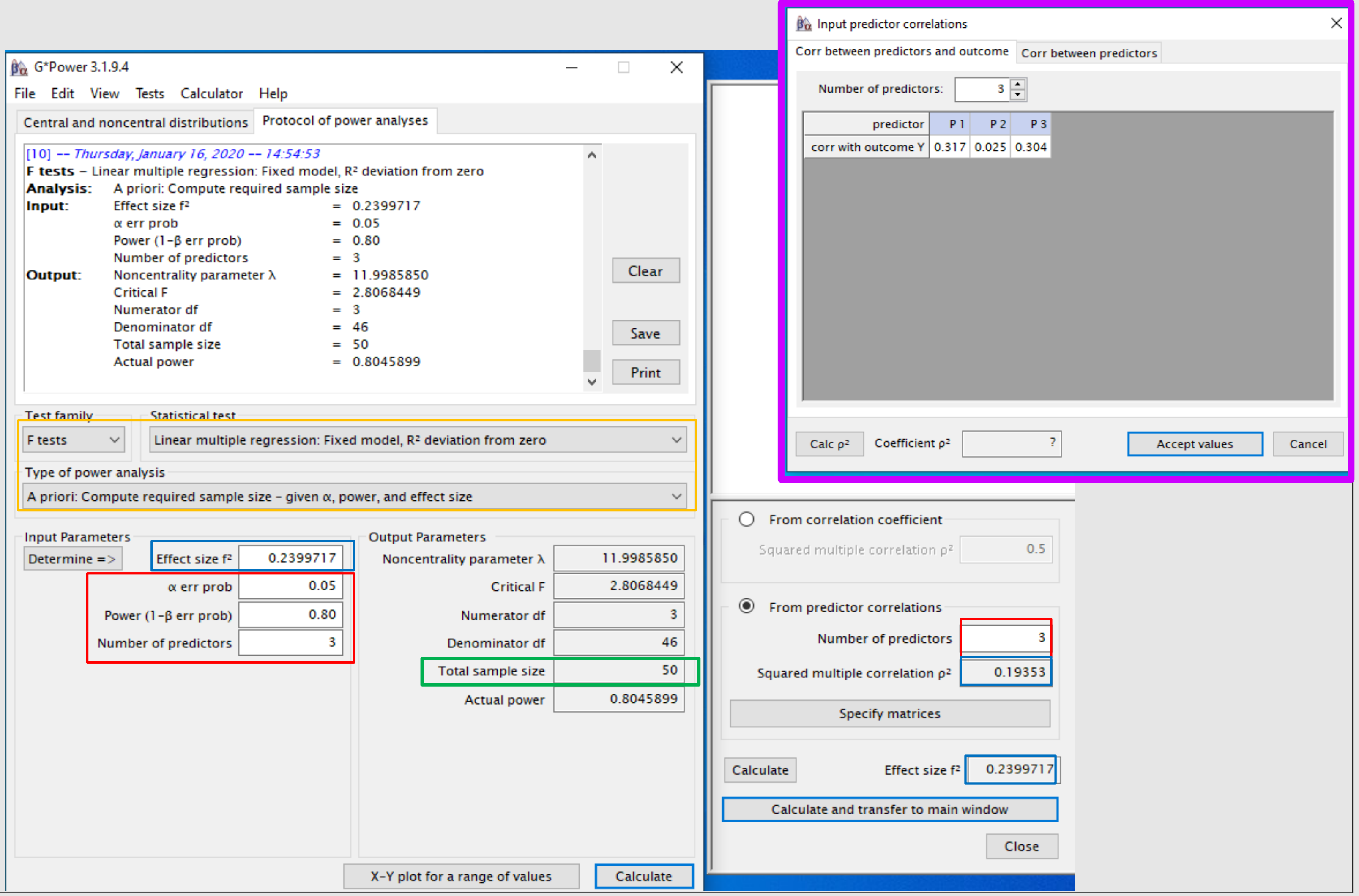
- $H_0=0$, $H_1 \neq 0$
- From trial data, we obtained the following correlation matrix

	Weight	Height	Age	GymTime
Weight	1	0.317	0.025	0.304
Height	**0.317**	1	-0.466	0.800
Age	**0.025**	-0.466	1	-0.485
GymTime	**0.304**	0.800	-0.485	1

- Only need about highlighted values for the matrix

Results:

- A total number of 50 samples are needed.



Multiple Linear Regression: Practice

Calculate the sample size for the following scenarios (with α=0.05, and power=0.80):

1. You are interested in determining if height (meters), weight (grams), and fertilizer added (grams) in plants can predict yield (grams of berries). You collect the following trial data.

Yield	**46.8**	**48.7**	**48.4**	**53.7**	**56.7**
Height	**14.6**	**19.6**	**18.6**	**25.5**	**20.4**
Weight	**95.3**	**99.5**	**94.1**	**110**	**103**
Fertilizer	**2.1**	**3.2**	**4.3**	**1.1**	**4.3**

2. You are interested in determining if the size of a city (in square miles), number of houses, number of apartments, and number of jobs can predict the population of the city (in # of individuals).

Multiple Linear Regression: Answers

Calculate the sample size for the following questions:

1. You are interested in determining if height (meters), weight (grams), and fertilizer added (grams) in plants can predict yield (grams of berries). You collect the following trial data.

Yield	46.8	48.7	48.4	53.7	56.7
Height	14.6	19.6	18.6	25.5	20.4
Weight	95.3	99.5	94.1	110	103
Fertilizer	2.1	3.2	4.3	1.1	4.3

   - **Running a regression found a squared multiple correlation coefficient (R-squared)=0.944**
   - **Got an effect size of 16.85; for 3 predictors, found total sample size of 6**

2. You are interested in determining if the size of a city (in square miles), number of houses, number of apartments, and number of jobs can predict the population of the city (in # of individuals).

   - **Guessed a large effect size (0.35); for 4 predictors, found total sample size of 40**

Pearson’s Correlation: Overview

Description: this test determines if there is a difference between two correlation values. The categorical variable is a grouping variable that is binomial.

Numeric. Var(s)	Cat. Var(s)	Cat. Var Group #	Cat Var. # of Interest	Parametric	Paired
2	1	N/A	N/A	Yes	No

Example:

- Is the correlation between hours studied and test score for group A statistically different than the correlation between hours studied and test score for group B?
- $H_0=0$, $H_1 \neq 0$

GPower:

  - Select **z tests** from Test family
  - Select **Correlations: Two Independent Person r’s** from Statistical test
  - Select **A priori** from Type of power analysis
- Background Info:
  a) Select **One** or **Two** Tail(s) as appropriate to your question
  b) Enter 0.05 in **α err prob** box – or a specific **α** you want for your study
  c) Enter 0.80 in **Power (1-β err prob**) box - or a specific power you want for your study
  d) Enter 1.0 in **Allocation ratio N2/N1** – unless the group sizes are different between the categories
  e) Hit Determine =>
  f) Add the correlation coefficients 1 and 2
  g) Hit **Calculate and transfer to main window**
  h) Hit Calculate on the main window
  i) Find **Total sample size** in the Output Parameters
- Naïve:
  a) Run steps a-d above
  b) Estimate an **effect size q**
  c) Hit Calculate on the main window
  d) Find **Total sample size** in the Output Parameters

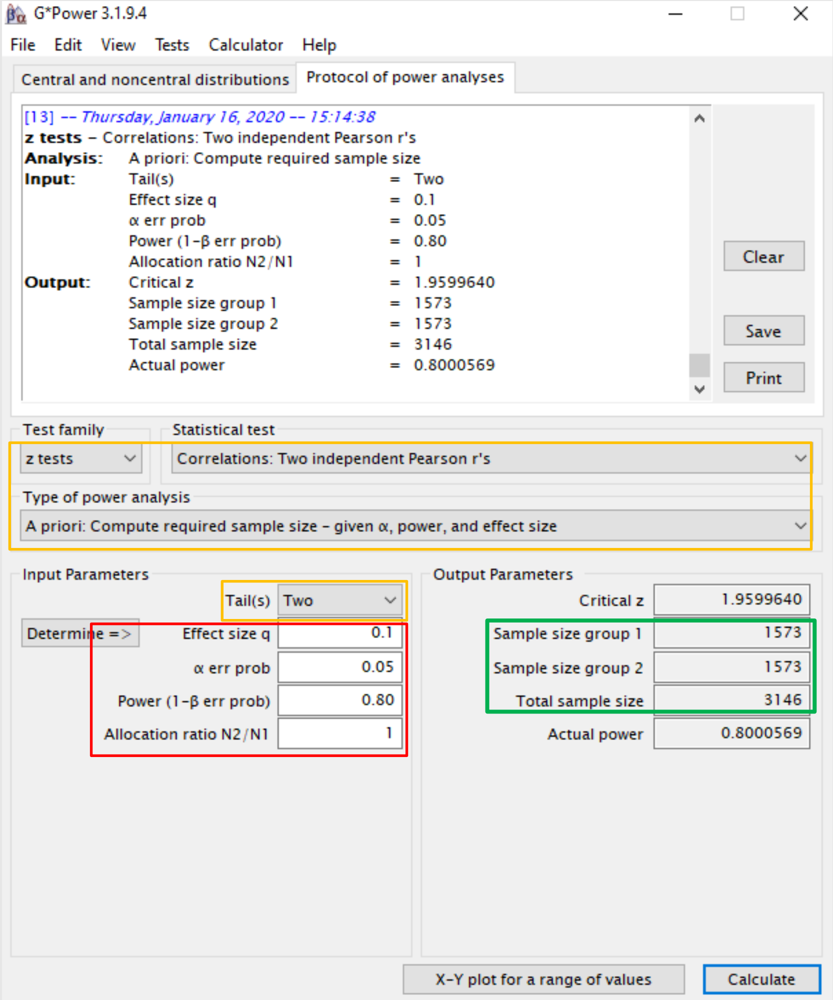
Pearson's Correlation: Example

Is the correlation between hours studied and test score for group A statistically different than the correlation between hours studied and test score for group B?

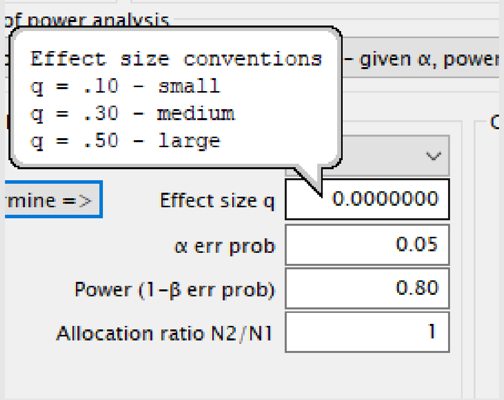
- H_0=0, H_1≠0
- Don't have data, but expect that the two groups will be quite similar so expect effect size to be **small**
- Two tails because we only care about difference, not direction

Results:

- A total number of 3146 samples are needed, 1573 for each of the two groups.



	Dropdown menu items you specified
	Values you entered
	Value(s) GPower calculated
	Sample size calculation



Pearson's Correlation: Practice

Calculate the sample size for the following scenarios (with α=0.05, and power=0.80):

1. You are interested in determining if the correlation between height and weight in men is **differen**t from the correlation between height and weight in women. You collect the following trial data.

Males	**Height**	178	166	172	186	182
	Weight	165	139	257	225	196
Females	**Height**	174	175	157	168	166
	Weight	187	182	149	132	143

2. You are interested in determining if, in lab mice, the correlation between longevity (in months) and average protein intake (grams) is **higher** from the correlation between longevity and average fat intake (grams).

Pearson's Correlation: Answers

Calculate the sample size for the following scenarios (with α=0.05, and power=0.80):

1. You are interested in determining if the correlation between height and weight in men is **different** from the correlation between height and weight in women. You collect the following trial data.

Males	**Height**	178	166	172	186	182
	Weight	165	139	257	225	196
Females	**Height**	174	175	157	168	166
	Weight	187	182	149	132	143

- **Correlation coefficient for men=0.37; coefficient for women=0.66**
- **Got effect size of -0.40, and with an allocation ratio of 1.0, total sample size is 198 (99 per group)**
- **NOTE: same result if effect size is 0.40, sign doesn't matter here**

2. You are interested in determining if, in lab mice, the correlation between longevity (in months) and average protein intake (grams) is **higher** from the correlation between longevity and average fat intake (grams).

- **Guess a medium effect (0.3), and with an allocation ratio of 1.0, total sample size is 282 (141 per group)**

Non-Parametric Regression (Logistic): Overview

Description:

Non-parametric regression is any regression where the numerical variables are not normally distributed. The two that can be calculate in GPower are Logistic and Poisson.

In **Logistic regression**, the response variable (Y) is binary (0/1). It tests whether a continuous predictor is a significant predictor of a binary outcome, with or without other covariates. (Can also run with a dichotomous predictor but will only show continuous here).

Numeric. Var(s)	Cat. Var(s)	Cat. Var Group #	Cat Var. # of Interest	Parametric	Paired
≥2	0	N/A	N/A	No	N/A

Example:

- Does body mass index (BMI) influences mortality (yes 1, no 0)?
- $H_0=0$, $H_1\neq 0$

GPower:

- Select **z tests** from Test family
- Select **Logistic Regression** from Statistical test
- Select **A priori** from Type of power analysis
- Select **Options** button and set the **Input effect size as .. Two probabilities**
- Background Info or Naïve:
  - a) Select **One** or **Two** Tail(s) as appropriate to your question
  - b) Enter 0.05 in **α err prob** box – or a specific **α** you want for your study
  - c) Enter 0.80 in **Power (1-β err prob**) box - or a specific power you want for your study
  - d) Enter value in **Pr(Y=1|Y=1) H0** box-(explained in next slide)
  - e) Enter value in **Pr(Y=1|Y=1) H1** box-(explained in next slide)
  - f) Enter value in **R^2 other X** box-(explained in next slide)
  - g) Select **X distribution** – will have to examine or predict distribution of predictor (X) variable-(explained in next slide)
  - h) Enter parameters in the **X parameter box**(es) – will depend on the distribution
  - i) Hit Calculate on the main window
  - j) Find **Total sample size** in the Output Parameters

Logistic Regression: parameters and how to calculate them

- **Pr(Y=1|X=1) H0 –** probability of Y variable (response) =1, when main predictor is one standard deviation above its mean
  - Based on background information, or make informed guess
- **Pr(Y=1|X=1) H1** - probability of Y variable (response) =1, when main predictor is at its mean
  - Based on background information, or make informed guess
- **R^2 other X** – Expected R-squared between main predictor variable and over covariates; amount of variability in main predictor that is accounted for by covariates
  - If there are no covariates (as in the simplest case of a single predictor), enter 0
  - Otherwise, calculate with background data by regressing main predictor onto data for all other covariates
  - Rule of thumb for naïve estimation: low association=0.04, moderate association=0.25, strong association=0.81
- **X distribution** – will have to examine or predict distribution of predictor (X) variable
  - Select normal unless you think the main predictor is distributed differently
- **X parameter box**(es) – will depend on the distribution
  - For normal, the μ (mu) is the z-score population mean of main predictor, while sigma (σ) is that predictor's z-score population Standard Deviation

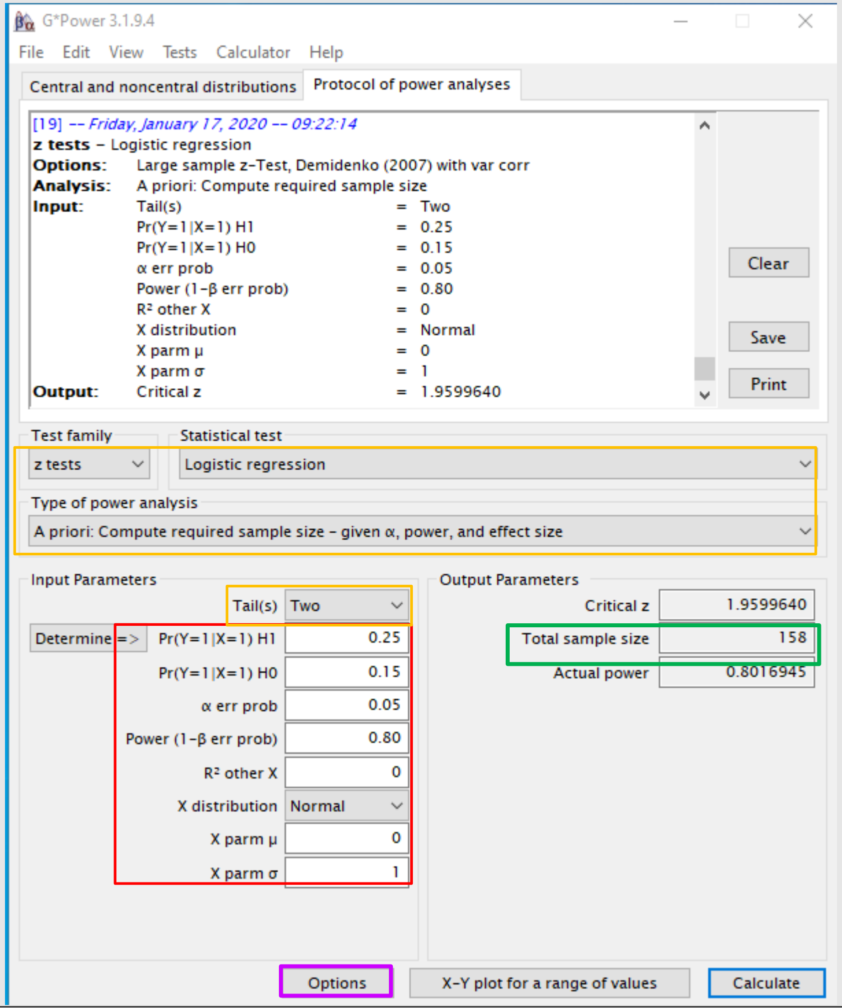
Non-Parametric Regression (Logistic): Example

Does body mass index (BMI) influences mortality (yes 1, no 0)?

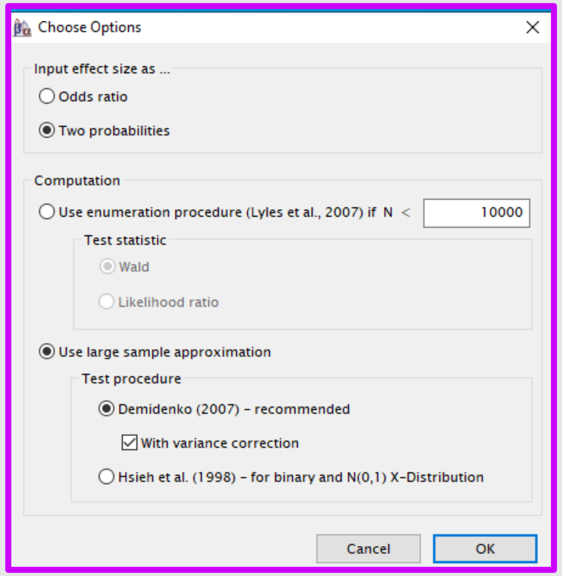
- $H_0=0, \quad H_1 \neq 0$
- Based on background data, expect Pr(Y=1|X=1) H0 to be low at 0.15 and Pr(Y=1|X=1) H1 to be higher, at 0.25
- Since BMI is only predictor, the R^2 other X is 0
- Expect BMI to be normally distributed (mean=0, SD=1)
- Two tails because we only care about difference, not direction
- Make sure to Choose Options

Results:

- A total number of 158 samples are needed.



(yellow)	Dropdown menu items you specified
(red)	Values you entered
(blue)	Value(s) GPower calculated
(green)	Sample size calculation



Non-Parametric Regression (Poisson): Overview

Description:

In **Poisson regression**, the response variable (Y) is a rate. It tests whether a continuous predictor variable influences the rate of events over a set period, with or without covariates (can also do dichotomous predictor but will only show continuous here).

Note that Poisson regression assumes independence of observations which is that the occurrence or absence of a previous event should not influence whether another event will occur. Subjects can have multiple events, as long as they are independent.

Numeric. Var(s)	Cat. Var(s)	Cat. Var Group #	Cat Var. # of Interest	Parametric	Paired
≥2	0	N/A	N/A	No	N/A

Example:

- Does a change in drug dose decrease the rate of adverse affects?
- H_0=0, H_1≠0

GPower:

- Select **z tests** from Test family
- Select **Poisson Regression** from Statistical test
- Select **A priori** from Type of power analysis
- Background Info or Naïve:

a) Select **One** or **Two** Tail(s) as appropriate to your question

b) Enter value in **Exp(β1)** box – (explained in next slide)

c) Enter 0.05 in **α err prob** box – or a specific α you want for your study

d) Enter 0.80 in **Power (1-β err prob**) box - or a specific power you want for your study

e) Enter value in **Base rate exp(β)** box-(explained in next slide)

f) Enter value in **Mean exposure** box –(explained in next slide)

g) Enter value in **R^2 other X** box-same as logistic; enter 0 if no covariates

h) Select **X distribution** – will have to examine or predict distribution of predictor (X) variable

i) Enter parameters in the **X parameter box**(es) – will depend on the distribution

j) Hit Calculate on the main window

k) Find **Total sample size** in the Output Parameters

Poisson Only:

Poisson Regression: parameters and how to calculate them

- **Exp(β1)** –change in response rate with a 1-unit increase in the predictor variable
  - If we expect response rate to go **UP** by 25% per 1-unit predictor increase, the Exp(β1) would be 1.25
  - If we expect response rate to go **DOWN** by 30% per 1-unit predictor increase, the Exp(β1) would be 0.70
- **Base rate exp(β)** –The response rate we expect if there is no intervention
  - Need to know your unit of exposure, what rate is being counted-could be time, distance, sample size, volume, etc.
  - For that unit of exposure, the **base rate** is the number of events expected per length of exposure
  - Ex.: 13 events/30 days= base rate of 0.43; 1 event/20 miles= base rate of 0.05; 5 events/10 volunteers= base rate of 0.5
- **Mean exposure** – length of exposure (how long you want the study to last in terms of your unit of exposure)
- **R^2 other X** – Expected R-squared between main predictor variable and over covariates; amount of variability in main predictor that is accounted for by covariates
  - If there are no covariates (as in the simplest case of a single predictor), enter 0
  - Otherwise, calculate with background data by regressing main predictor onto data for all other covariates
  - Rule of thumb for naïve estimation: low association=0.04, moderate association=0.25, strong association=0.81
- **X distribution** – will have to examine or predict distribution of predictor (X) variable
  - Select normal unless you think the main predictor is distributed differently
- **X parameter box**(es) – will depend on the distribution
  - For normal, the μ (mu) is the z-score population mean of main predictor, while sigma (σ) is that predictor's z-score population Standard Deviation

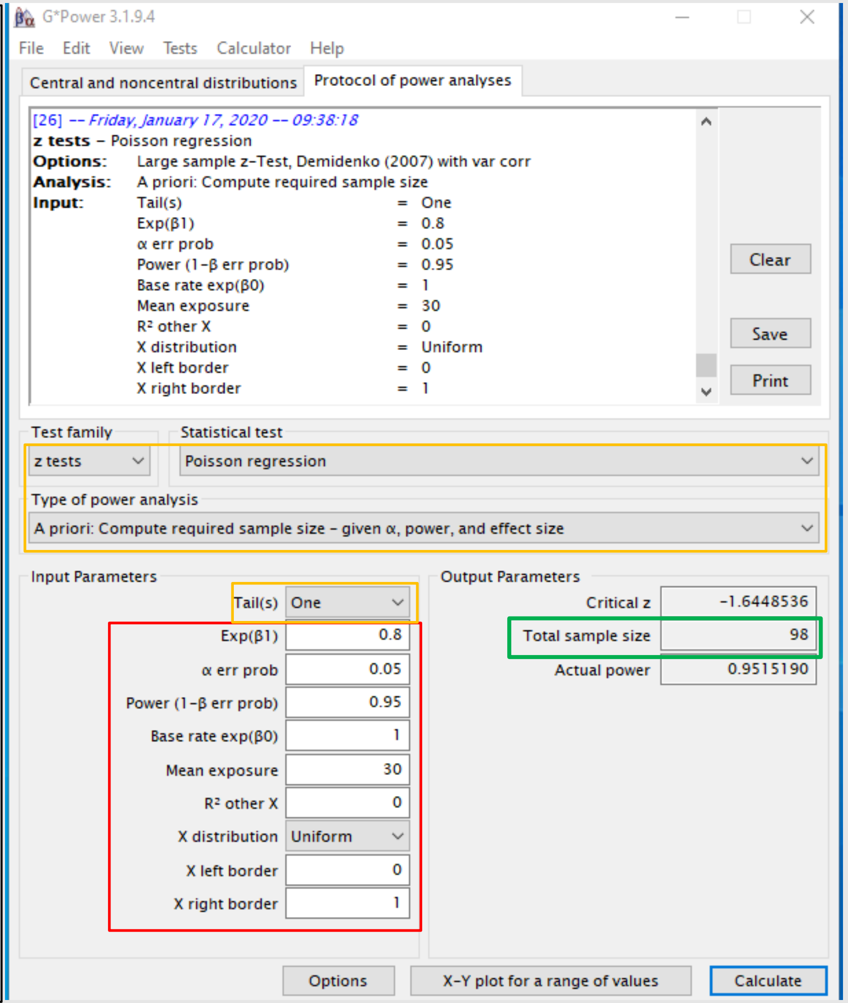
Non-Parametric Regression (Poisson): Example

Does a change in drug dose decrease the rate of adverse affects?

- H_0=0, H_1≠0
- No background information
- Expect a moderate decrease in rate with increase in drug dose, so will set Exp(β1) to **0.80**
- Our unit of exposure will be days and evens will be adverse effects, so let's set the Base rate exp(β0) to be:
  - **1 adverse effect/1 day = 1**
- We want to try this for a month, so let's have Mean exposure be 30 days
- Since Drug dose is only predictor, the R^2 other X is 0
- Not sure what rate of adverse effects' distribution is, so to be safe, went with uniform
- One tail, because we are specifically interested in a decrease in rate

Results:

- A total number of 98 samples are needed.



(yellow)	Dropdown menu items you specified
(red)	Values you entered
(blue)	Value(s) GPower calculated
(green)	Sample size calculation

Non-Parametric Regression: Practice

Calculate the sample size for the following scenarios (with α=0.05, and power=0.80):

1. You are interested in determining if body temperature influences sleep disorder prevalence (yes 1, no 0). You collect the following trial data.

Temperature	98.6	98.5	99.0	97.5	98.8	98.2	98.5	98.4	98.1
Sleep Disorder?	No	No	Yes	No	Yes	No	No	Yes	No

2. You are interested in determining if the rate of lung cancer incidence changes with a drug treatment.

Non-Parametric Regression: Answers

Calculate the sample size for the following scenarios (with α=0.05, and power=0.80):

1. You are interested in determining if body temperature influences sleep disorder prevalence (yes 1, no 0). You collect the following trial data.

Temperature	98.6	98.5	99.0	97.5	98.8	98.2	98.5	98.4	98.1
Sleep Disorder?	No	No	Yes	No	Yes	No	No	Yes	No

   - Want logistic regression with two-tails; temperature mean is 98.4 with SD=0.436 -> (97.964 --98.836)
   - Pr(Y=1|X=1) H0 = 0.33 (as only one had sleep disorder at ranges outside one SD); Pr(Y=1|X=1) H1= 0.67
   - No other Xs, so R-squared is 0
   - Temperature is normally distributed, but set the mean to 0 and SD=1 (want it standardized rather than actual)
   - Got a total sample size of 195

2. You are interested in determining if the rate of lung cancer incidence changes with an environmental factor over a 3 years.
   - Want poisson regression with two-tails
   - Expect an Exp(β1) of 1.02 (up by 2% per every unit increase)
   - Researching rate of lung cancer in men per year found rate of 57.8 (per 100,000)
   - Mean exposure is 3, as we want to try this for 3 years; No other Xs, so R-squared is 0
   - Try normal distribution of X with 0,1
   - Got a total sample size of 116

ANCOVA: Overview

Description: Blend of ANOVA and regression.

Tests whether the means of a dependent variable (DV) are equal across levels of a categorical independent variable (IV) often called a treatment, while statistically controlling for the effects of other continuous variables that are not of primary interest, known as covariates.

Numeric. Var(s)	Cat. Var(s)	Cat. Var Group #	Cat Var. # of Interest	Parametric	Paired
>1	≥1	>1	≥1	Yes	N/A

Example:

- Is there a difference in recovery time across two conditions and three diagnostic groups with baseline-scores as the co-variate?
- $H_0=0$, $H_1\neq0$

GPower:

- Select **F tests** from Test family
- Select **ANCOVA: Fixed effects, main effects and interactions** from Statistical test
- Select **A priori** from Type of power analysis
- Background Info:
  a) Enter 0.05 in **α err prob** box – or a specific α you want for your study
  b) Enter 0.80 in **Power (1-β err prob**) box - or a specific power you want for your study
  c) Fill out boxes for **Numerator df** and **Number of groups-**same as Multi-way ANOVA
  d) Fill out box for **Number of covariates**-depends on number you are including in design
  e) Hit Determine =>
  f) Select Direct
  g) Add **partial eta-squared**-same as ANOVA
  h) Hit **Calculate and transfer to main window**
  i) Hit Calculate on the main window
  j) Find **Total sample size** in the Output Parameters
- Naïve:
  a) Run steps a-c above
  b) Estimate an **effect size f**
  c) Hit Calculate on the main window
  d) Find **Total sample size** in the Output Parameters

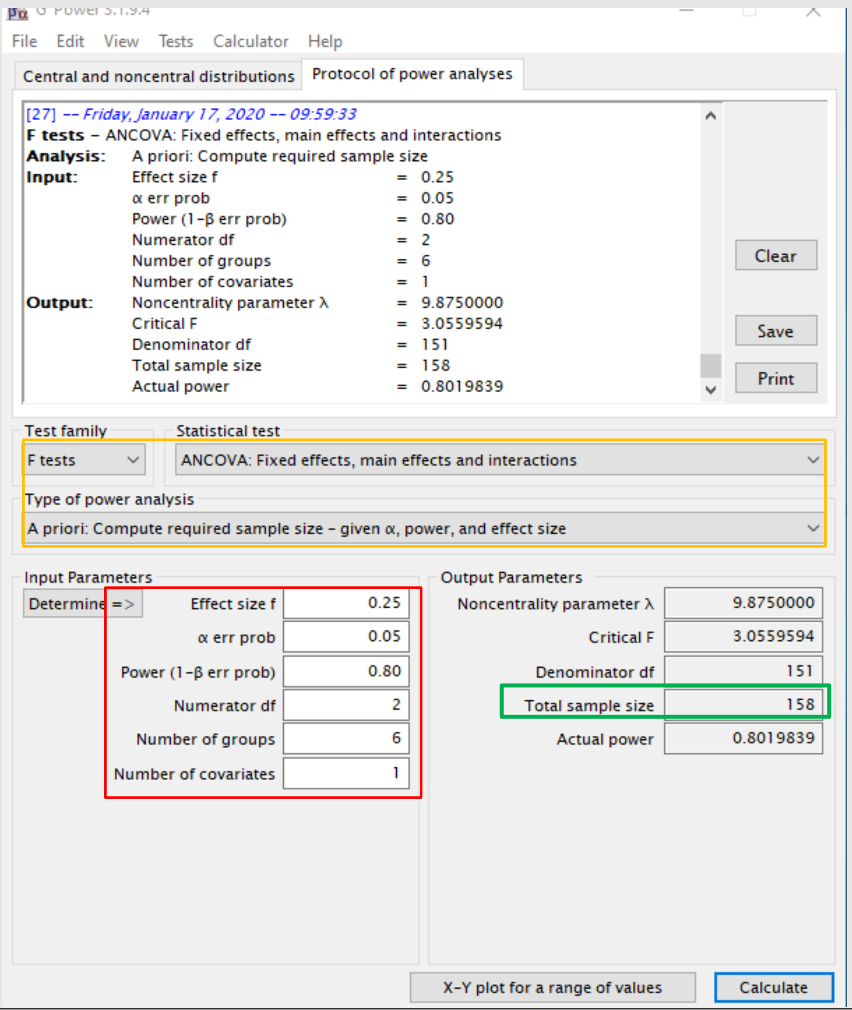
ANCOVA: Example

Is there a difference in recovery time across two conditions and three diagnostic groups with baseline-scores as the co-variate?

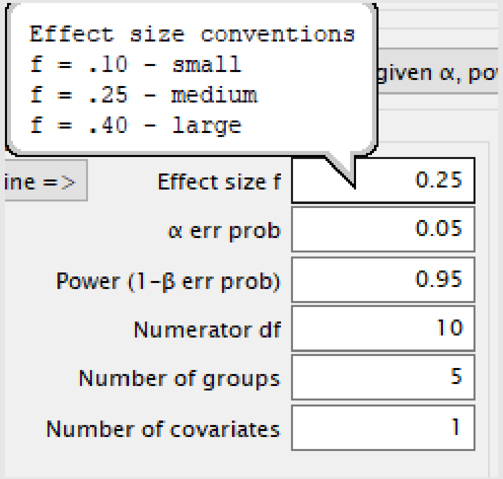
- $H_0=0$, $H_1 \neq 0$
- No background information
- Expect a **medium effect**
- Numerator df = (2-1)*(3-2)=2
- Number of groups=2*3=6
- Only one covariate

Results:

- A total of 158 samples are needed.



- Dropdown menu items you specified
- Values you entered
- Value(s) GPower calculated
- Sample size calculation



ANCOVA: Practice

Calculate the sample size for the following scenarios (with α=0.05, and power=0.80):

1. You are interested in determining if there is a difference in weight loss due to treatment (Drug X, Y, and Z), while controlling for body weight. You collect the following trial data for treatment (shown on right).

2. You are interested in determining if there is a difference in weight loss due to a diet style (1, 2, 3, or 4), while controlling for starting body weight and body height.

Drug X		Drug Y		Drug Z	
Starting weight	Weight lost	Starting weight	Weight lost	Starting weight	Weight lost
120	**1.9**	**120**	**5.7**	**239**	**1.1**
144	**6.6**	**236**	**9**	**332**	**4.6**
270	**5.9**	**204**	**4.4**	**268**	**12.6**
169	**9.6**	**127**	**7.5**	**207**	**12.5**
189	**3.9**	**134**	**3.7**	**183**	**18.5**

ANCOVA: Answers

Calculate the sample size for the following scenarios (with α=0.05, and power=0.80):

1. You are interested in determining if there is a difference in weight loss due to treatment (Drug X, Y, and Z), while controlling for body weight. You collect the following trial data for treatment (shown on right).
   - **Numerator df =3-1=2**
   - **3 groups (3 treatments) and 1 covariate (weight)**
   - **Ran general linear model in SPSS with data to get $\eta 2 = SS_{eff} / (SS_{eff} + SS_{err})$**
   - **$\eta 2 = 64 / (64 + 235.3) = 0.213$**
   - **Got effect size of 0.52 for a total sample size of 39**
2. You are interested in determining if there is a difference in weight loss due to a diet style (1, 2, 3, or 4), while controlling for starting body weight and body height.
   - **Numerator df =4-1=3**
   - **4 groups (4 diet styles) and 2 covariates (weight, and height)**
   - **Guessed medium effect size (0.25) for a total sample size of 179**

Drug X		Drug Y		Drug Z	
Starting weight	Weight lost	Starting weight	Weight lost	Starting weight	Weight lost
120	1.9	120	5.7	239	1.1
144	6.6	236	9	332	4.6
270	5.9	204	4.4	268	12.6
169	9.6	127	7.5	207	12.5
189	3.9	134	3.7	183	18.5

Generalized Linear Mixed Models, a postlude

Description: The most expansive model for statistical analysis. Can used fixed effects, random effects, categorical predictor variables, numerical predictor variables, and a variety of distributions (Not just normal). Often abbreviated as GLMMs.

Numeric. Var(s)	Cat. Var(s)	Cat. Var Group #	Cat Var. # of Interest	Parametric	Paired
≥1	≥0	≥0	≥0	Y/N	N/A

Example:

- Can height, weight, rate of driving speed, gender, age, ethnicity, and number of weekly drinks predict the probability of car crashes?
- $H_0=0$, $H_1\neq0$

- **Unfortunately, GLMMS are beyond the scope of GPower**
- **Stay tuned for approaches in other software programs such as R, SAS, and Excel**

Common Clinical Study Designs

Name	Description	Typical Statistical Method
Randomized Control Trial	A controlled clinical trial that randomly (by chance) assigns participants to two or more groups. There are various methods to randomize study participants to their groups	**ANOVA**
Double Blind Method	A type of randomized controlled clinical trial/study in which neither medical staff/physician nor the patient knows which of several possible treatments/therapies the patient is receiving.	**ANOVA**
Cohort Study	A clinical research study in which people who presently have a certain condition or receive a particular treatment are followed over time and compared with another group of people who are not affected by the condition.	**Regression**
Case Control	Case-control studies begin with the outcomes and do not follow people over time. Researchers choose people with a particular result (the cases) and interview the groups or check their records to ascertain what different experiences they had. They compare the odds of having an experience with the outcome to the odds of having an experience without the outcome.	**Regression**
Cross-sectional Study	The observation of a defined population at a single point in time or time interval. Exposure and outcome are determined simultaneously	**Regression**

Info from https://research.library.gsu.edu/c.php?g=115595&p=755213 and https://hsl.lib.umn.edu/biomed/help/understanding-research-study-designs

Acknowledgements

- The DaCCoTA is supported by the National Institute of General Medical Sciences of the National Institutes of Health under Award Number U54GM128729.

- For the labs that use the Biostatistics, Epidemiology, and Research Design Core in any way, including this Module, please acknowledge us for publications. **"Research reported in this publication was supported by DaCCoTA (the National Institute of General Medical Sciences of the National Institutes of Health under Award Number U54GM128729).**

DaCCoTA
DAKOTA CANCER COLLABORATIVE ON TRANSLATIONAL ACTIVITY