Lab 05: Power Calculations

Step by Step guidance

Dr. Gordon Wright

g.wright@gold.ac.uk

October 31, 2023

Overview of Lab 05

Power

In the following slides AND IN THE HANDOUT, you can find step by step guidance on using G*Power to calculate your Power - and therefore the number of participants you would be required to test.

Remember, it’s not life and death if you don’t meet sample size this year, but it would be next year.

You will have potential access to the RPS next year.

You will be asked to take part in all available Mini-Dissertation studies this year in weeks 11-15

I got the, I got the power!

The formal definition of a p value is the probability of observing a result at least as extreme as the one observed, assuming the null hypothesis is true (e.g. Cohen, 1994).

> Cohen, J (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Lawrence Erlbaum Associates. I wouldn’t bother.

What?

This means, and assuming we have a crystal ball to know the ‘truth’ of whether there is an effect of our manipulation..

- a small p value indicates the results are surprising if the null hypothesis <u>IS</u> true
- a large p value indicates the results are not very surprising if the null <u>IS</u> true.

G*Power

You can download G*Power on this page. Under the heading “download” click on the appropriate version for whether you have a Windows or Mac computer.

https://www.psychologie.hhu.de/arbeitsgruppen/allgemeine-psychologie-und-arbeitspsychologie/gpower

2x2 factorial ANOVA Power analysis

In a 2x2 design, there are three effects to compare:

- the main effect of IV1,
- the main effect of IV2,
- and the interaction between IV1 and IV2.

Look after the little effect first

There are three comparisons to make here and it is unlikely you would expect the same effect size for all three comparisons, meaning you must ensure all three effects are sufficiently powered by a single experimental design. This just means that your sample size will need to be sufficient to detect the <u>weakest</u> effect size of the three.

You might expect the effect size (or difference) for IV1 to be a lot larger than the effect size (or difference) for IV2, for example.

This means you would have to ensure the smallest effect is sufficiently powered, in the above example, the IV2 effect size would drive the sample size calculation.

if the smallest effect is covered, by virtue of the fact power exists along a curve - then the larger effects would have sufficient power!

G*Power guide

If you are running a design with 2 between participant IVs, use Slide 1

If you are running a mixed design, and want to explore the between Main effect - slide 2

If you are running a mixed design, and want to explore the within Main effect - slide 3

If you are running a mixed design, and want to explore the interaction effect - slide 4

If you are running a design with 2 between participant IVs, use Slide 5

G*Power guide - Key info

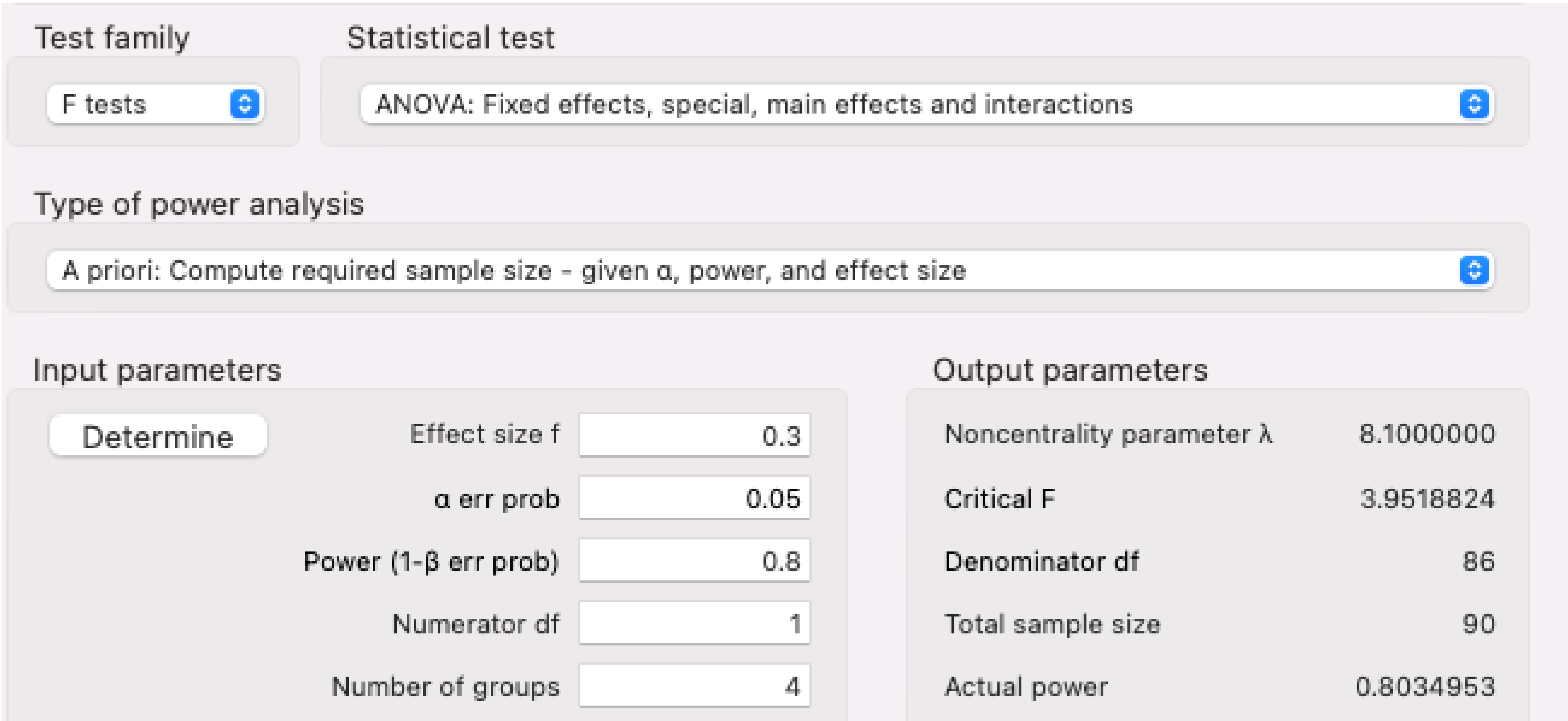
The following slides give you the specifics of what you need to include depending on the effect you are exploring. The only thing you need is the effect size of interest.

This is usually gleaned from your target paper (or other literature).

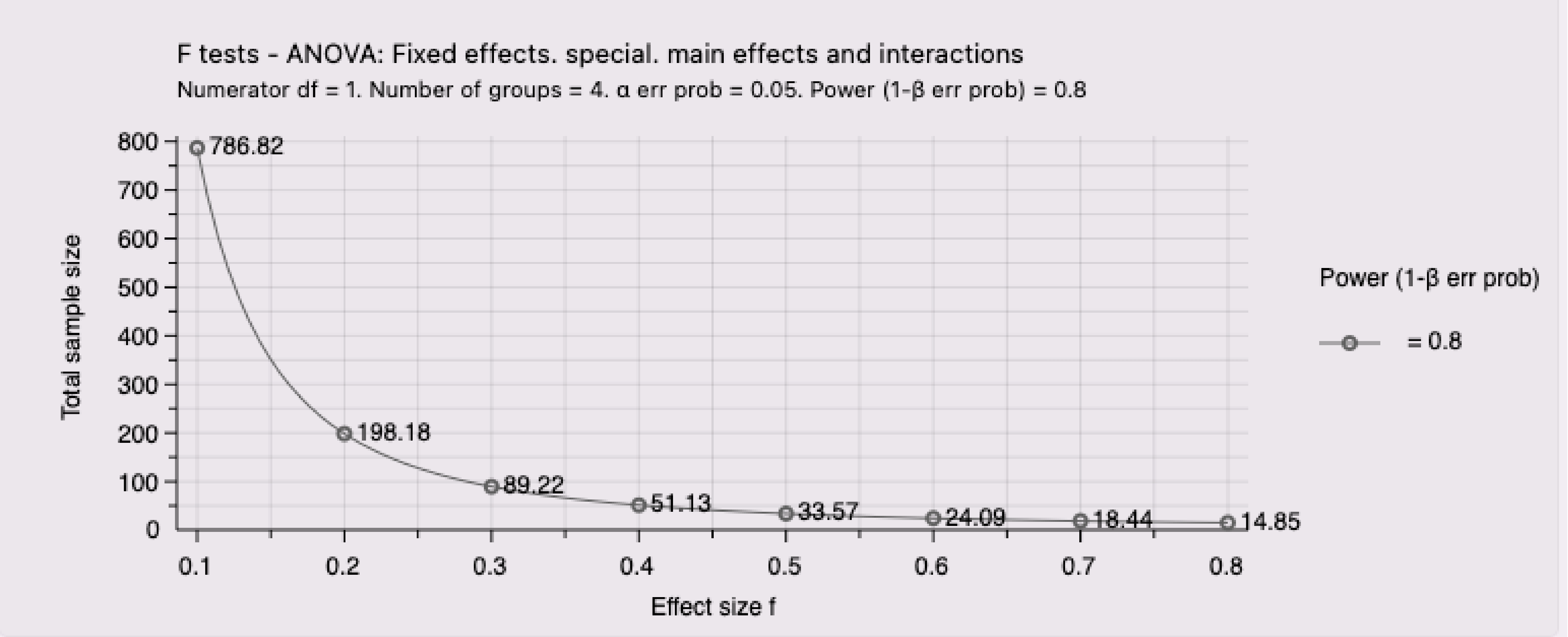
You will need to examine each of your effect sizes in your groups and cover the weakest with the the sample size, so the effect that requires the largest sample size will give the group sample size.

If you cannot find an effect, you may use a 0.3 f effect size, but this is to be used only when no alternative is available.

1 - Fully Between Designs (any effect of interest)



Power plot


F tests - ANOVA: Fixed effects. special. main effects and interactions
Numerator df = 1. Number of groups = 4. α err prob = 0.05. Power (1-β err prob) = 0.8
Total sample size
800
700
600
500
400
300
200
100
0
786.82
198.18
89.22
51.13
33.57
24.09
18.44
14.85
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
Effect size f
Power (1-β err prob)
= 0.8

2 - Mixed design (looking at a between main effect)

Test family: F tests

Statistical test: ANOVA: Repeated measures, between factors

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

Input parameters

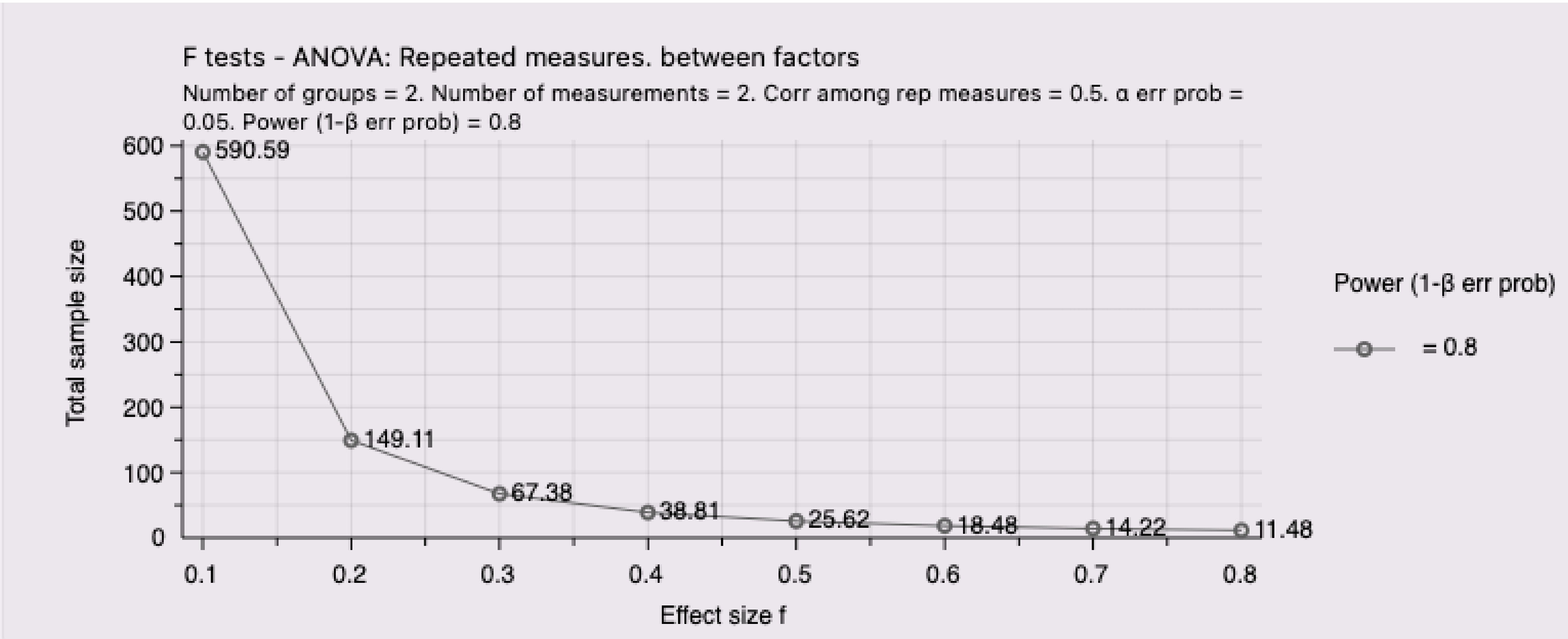
Determine

Effect size f	0.3
α err prob	0.05
Power (1-β err prob)	0.8
Number of groups	2
Number of measurements	2
Corr among rep measures	0.5

Output parameters

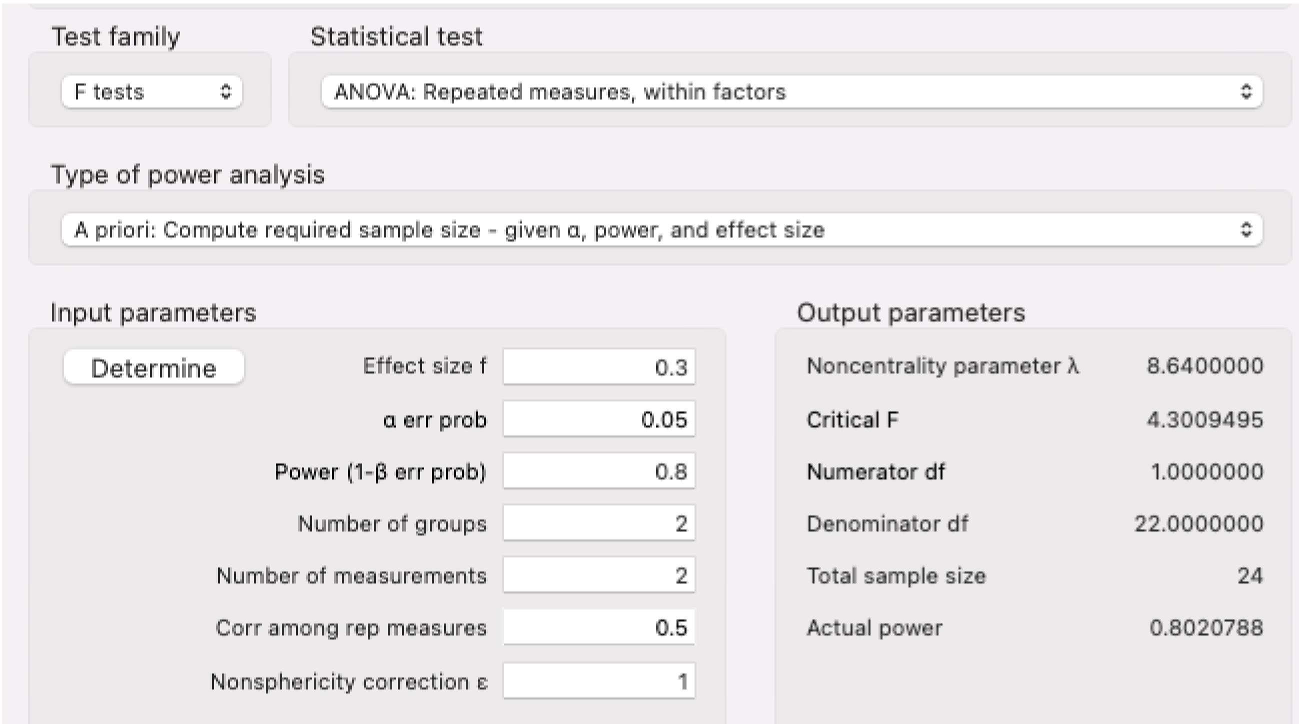
Noncentrality parameter λ	8.1600000
Critical F	3.9862695
Numerator df	1.0000000
Denominator df	66.0000000
Total sample size	68
Actual power	0.8036734

graph



F tests - ANOVA: Repeated measures. between factors
Number of groups = 2. Number of measurements = 2. Corr among rep measures = 0.5. α err prob = 0.05. Power (1-β err prob) = 0.8
Total sample size
600
500
400
300
200
100
0
590.59
149.11
67.38
38.81
25.62
18.48
14.22
11.48
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
Effect size f
Power (1-β err prob)
= 0.8

3 - Mixed design (looking at within main effect)



4 - Mixed design (looking at interaction effect)

Test family: F tests

Statistical test: ANOVA: Repeated measures, within-between interaction

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

Input parameters

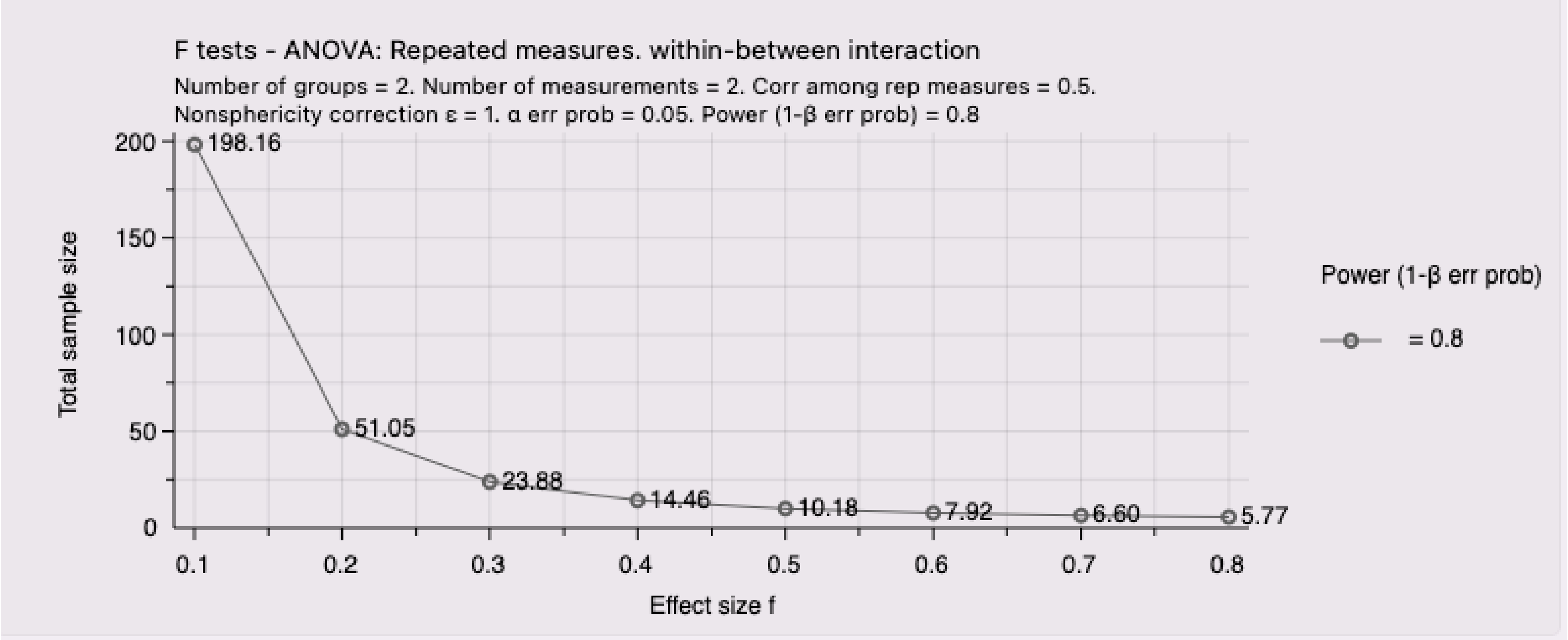
Determine

Effect size f	0.3
α err prob	0.05
Power (1-β err prob)	0.8
Number of groups	2
Number of measurements	2
Corr among rep measures	0.5
Nonsphericity correction ε	1

Output parameters

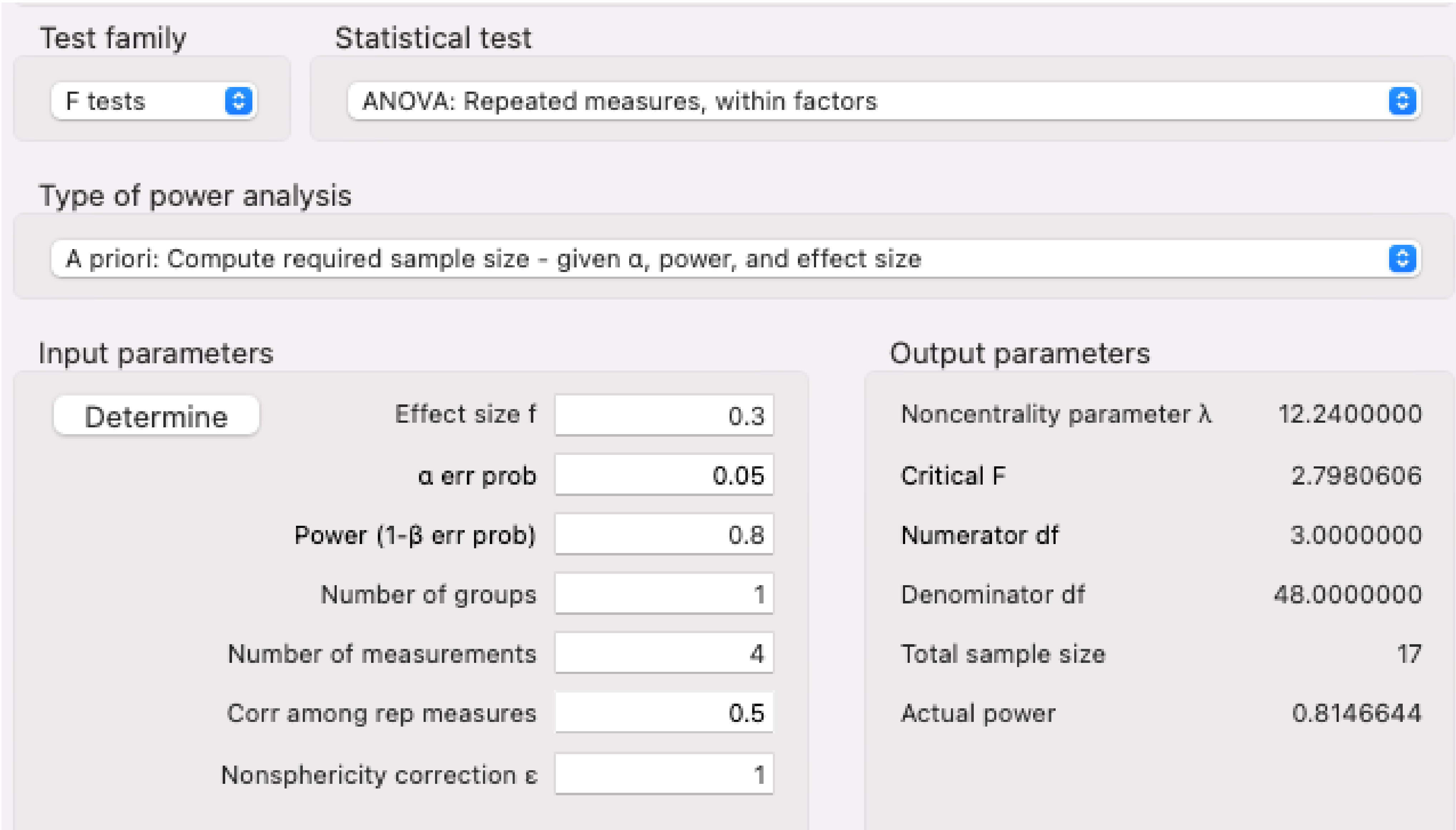
Noncentrality parameter λ	8.6400000
Critical F	4.3009495
Numerator df	1.0000000
Denominator df	22.0000000
Total sample size	24
Actual power	0.8020788

graph

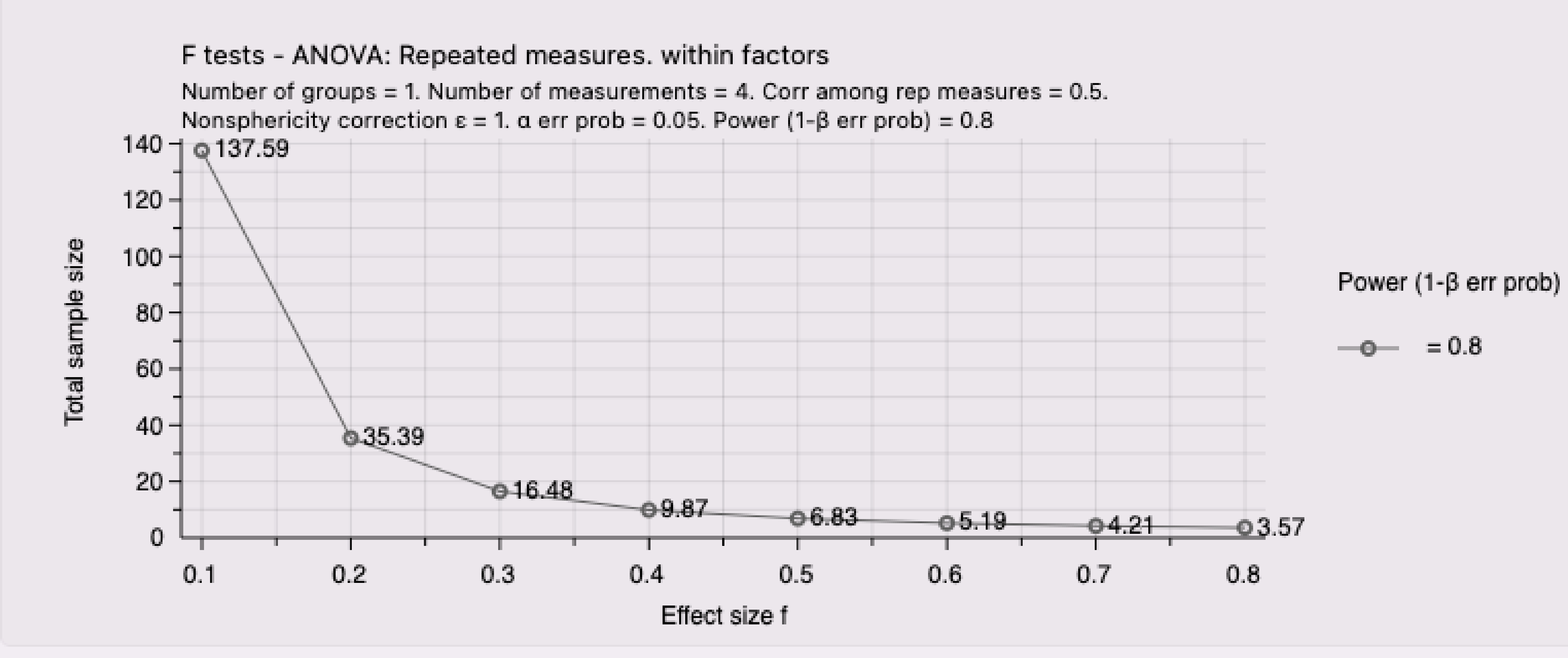


F tests - ANOVA: Repeated measures. within-between interaction
Number of groups = 2. Number of measurements = 2. Corr among rep measures = 0.5.
Nonsphericity correction ε = 1. α err prob = 0.05. Power (1-β err prob) = 0.8
Total sample size
200
150
100
50
0
198.16
51.05
23.88
14.46
10.18
7.92
6.60
5.77
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
Effect size f
Power (1-β err prob)
= 0.8

5 - Repeated Measures ANOVA (any effect)



Graph


F tests - ANOVA: Repeated measures. within factors
Number of groups = 1. Number of measurements = 4. Corr among rep measures = 0.5.
Nonsphericity correction ε = 1. α err prob = 0.05. Power (1-β err prob) = 0.8
Total sample size
140
120
100
80
60
40
20
0
137.59
35.39
16.48
9.87
6.83
5.19
4.21
3.57
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
Effect size f
Power (1-β err prob)
= 0.8

Writing up

> “In order to detect an effect size of Cohen’s d = 0.xx with 80% power (alpha = .05, two-tailed), G*Power analysis determines a sample size of XX participants in a *insert type* 2x2 ANOVA”. The smallest effect size of interest was set to d= 0.xx based on the meta-analysis by Dougal et al. (2023).”