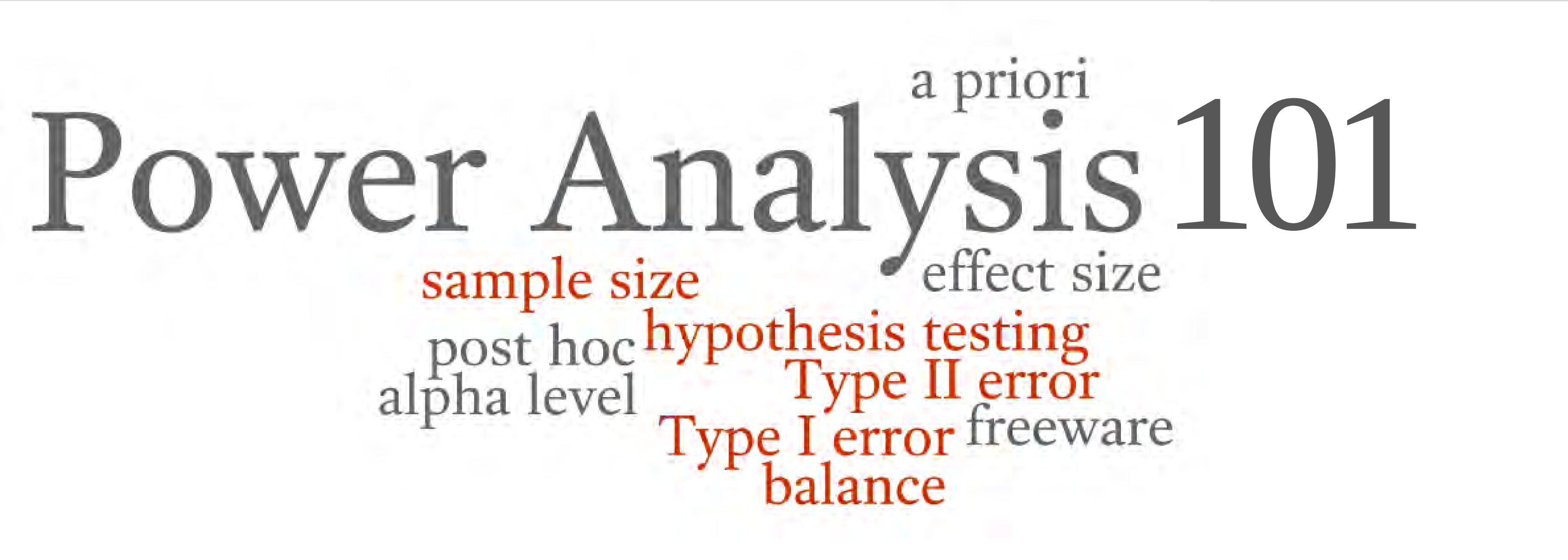


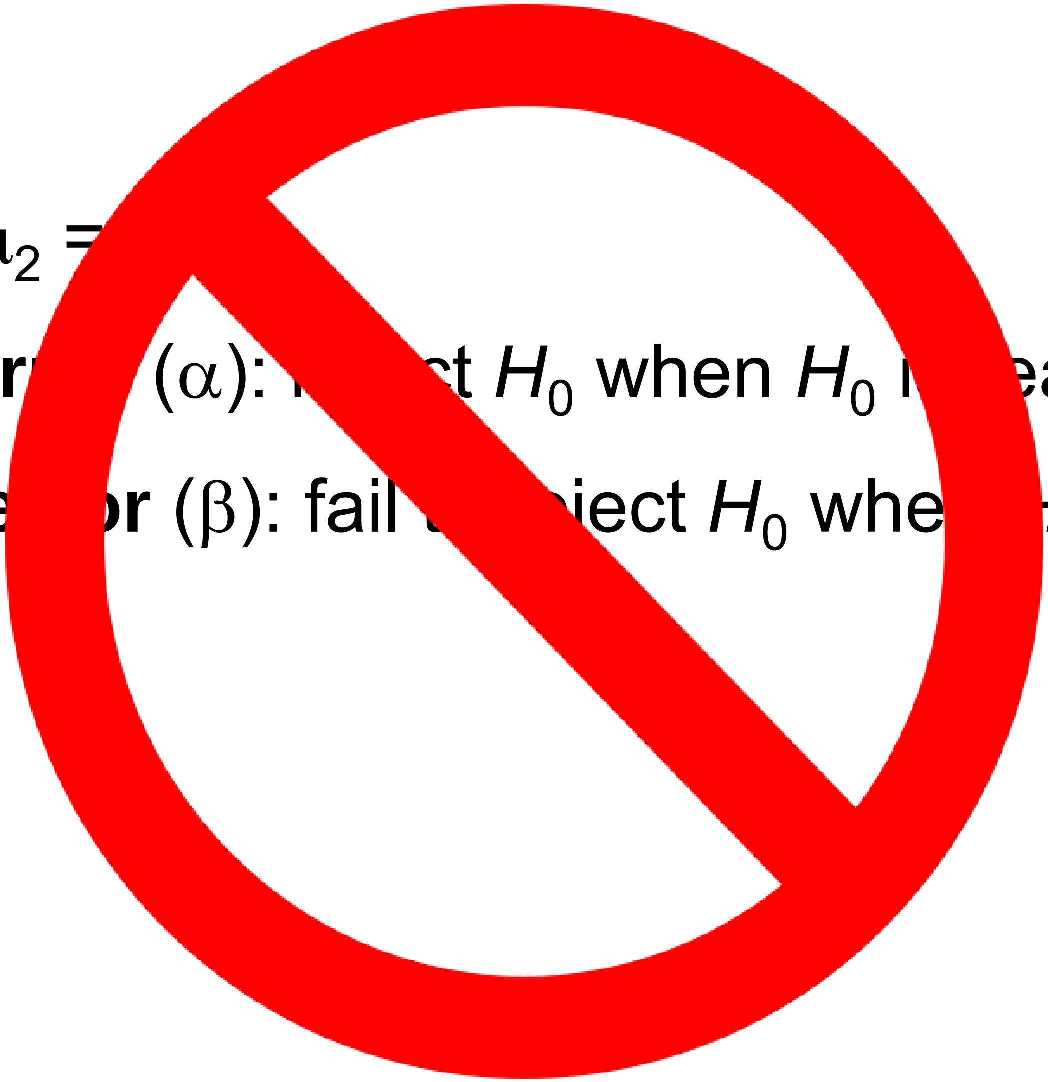
Richard G. Lomax
Susie Mauck
Robert Nichols

Outline

- Introduction to Power
- Importance of Power
- Determinants of Power
- Types of Power
- Power Analysis Software

The Ohio State University

What is Power?



H_0: $\mu_1 - \mu_2 =$

Type I er (α): r ct H_0 when H_0 i eally true.

Type II e or (β): fail t ject H_0 whe H_0 is really false.

TYPE I ERROR

Finding a difference when a difference isn't there.

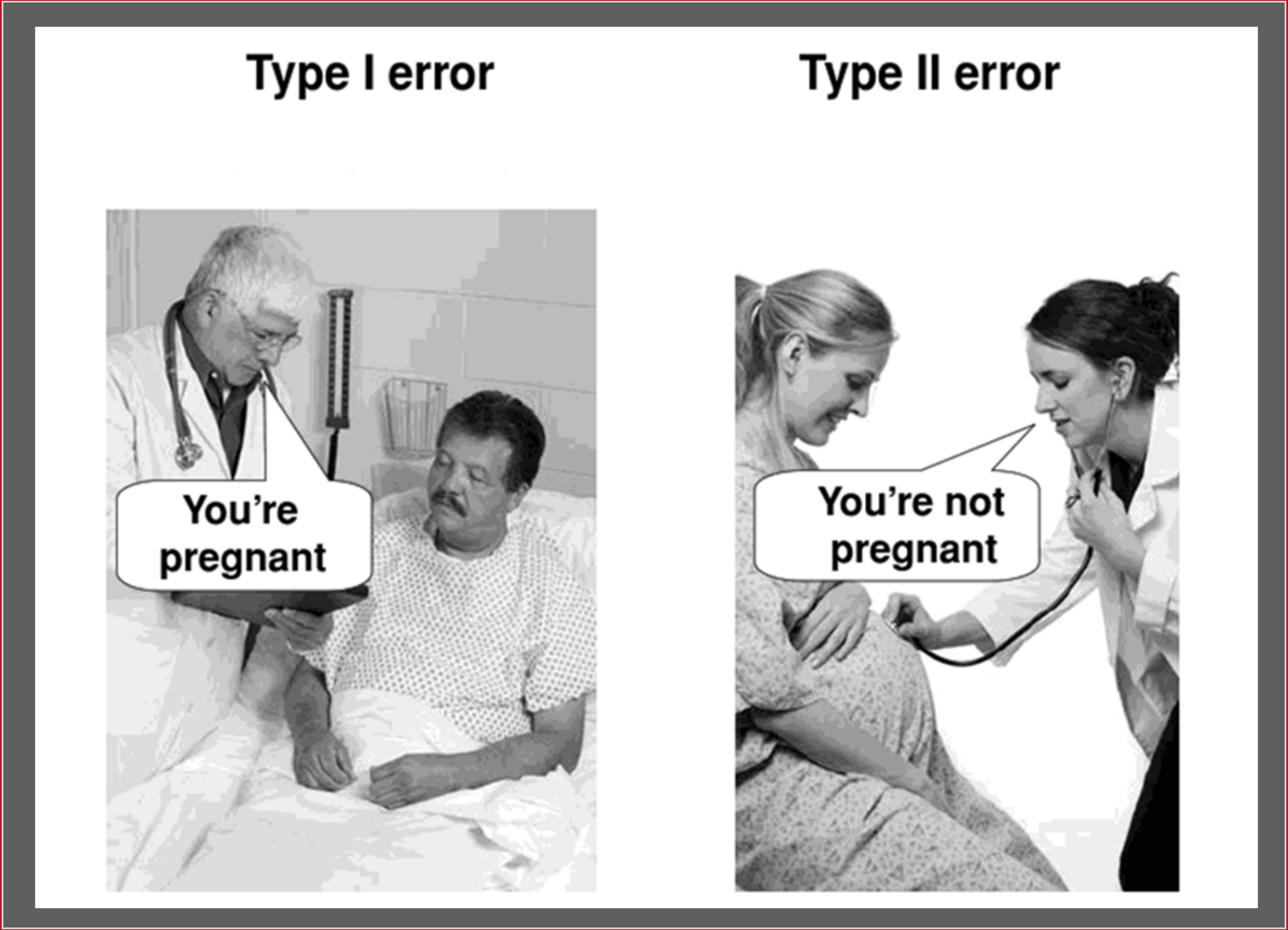
TYPE II ERROR

Not finding a difference when a difference is there.

Type I error: Thinking they are different when they are not.

Type II error: Thinking they are not different when they are.

Introduction to Power

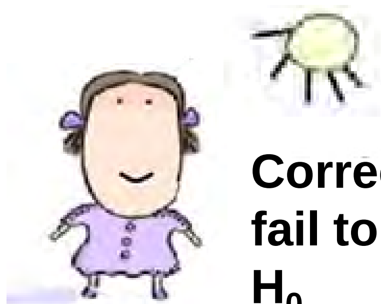
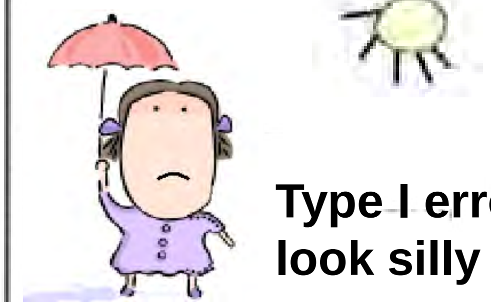


From The essential guide to effect sizes: Statistical power, meta-analysis, and the interpretation of research results by Paul D. Ellis, 2010.

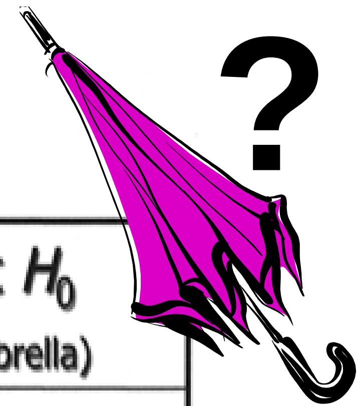
Null = no rain

Decision

State of Nature

	Fail to Reject H_0 (No umbrella)	Reject H_0 (Take umbrella)
H_0 is true (No rain)	 **Correctly fail to reject H_0**	 **Type I error – look silly**
H_0 is false (Rains)	 **Type II error – get wet**	 **Correctly reject H_0 POWER!**

Introduction to Power



Decision

State of Nature (Reality)	Fail to Reject H_0 (don't carry umbrella)	Reject H_0 (carry umbrella)
H_0 is TRUE (no rain)	Correct Decision (no umbrella, no rain; thus OK) $(1 - \alpha)$	Type I Error (bring umbrella, no rain; thus look silly) (α)
H_0 is FALSE (rains)	Type II Error (no umbrella, rains; thus get wet) (β)	Correct Decision (bring umbrella, rains; thus OK) $(1 - \beta)$ = Power

Null = no rain

You need to balance the two types of error…

TYPE I ERROR

TYPE II ERROR

Decision Making

Null hypothesis – hurricane won't hit NYC
Alternative hypothesis – hurricane hits NYC

Mayor Bloomberg could risk making a Type I error:

- Concludes Hurricane Sandy is coming.
- Orders evacuation.

Decision Making

Null hypothesis – hurricane won't hit NYC
Alternative hypothesis – hurricane hits NYC

Mayor Bloomberg could risk making a Type II error:

- Concludes Hurricane Sandy is not coming.
- Does nothing.

Which is worse?   It depends…

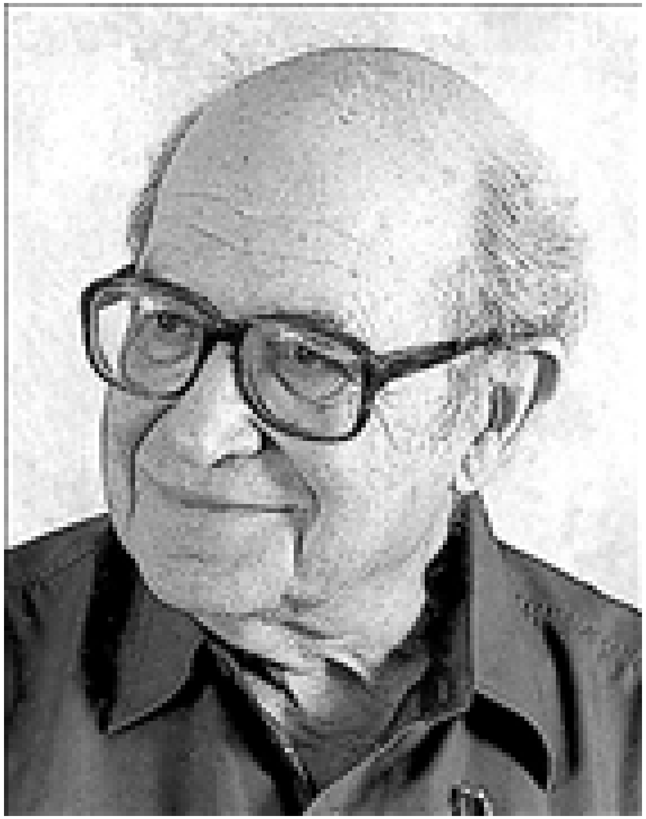
What is the appropriate level of power?

$\alpha = .05$
Power = .8

α = .05
Power = .8

They're really more like guidelines…

The five-eighty convention…



Jacob Cohen, 1923 – 1998
http://archpsyc.jamanetwork.com

Power (1 – β) :

Power of .8 → comfortable accepting a 20% chance you will miss something that is really there

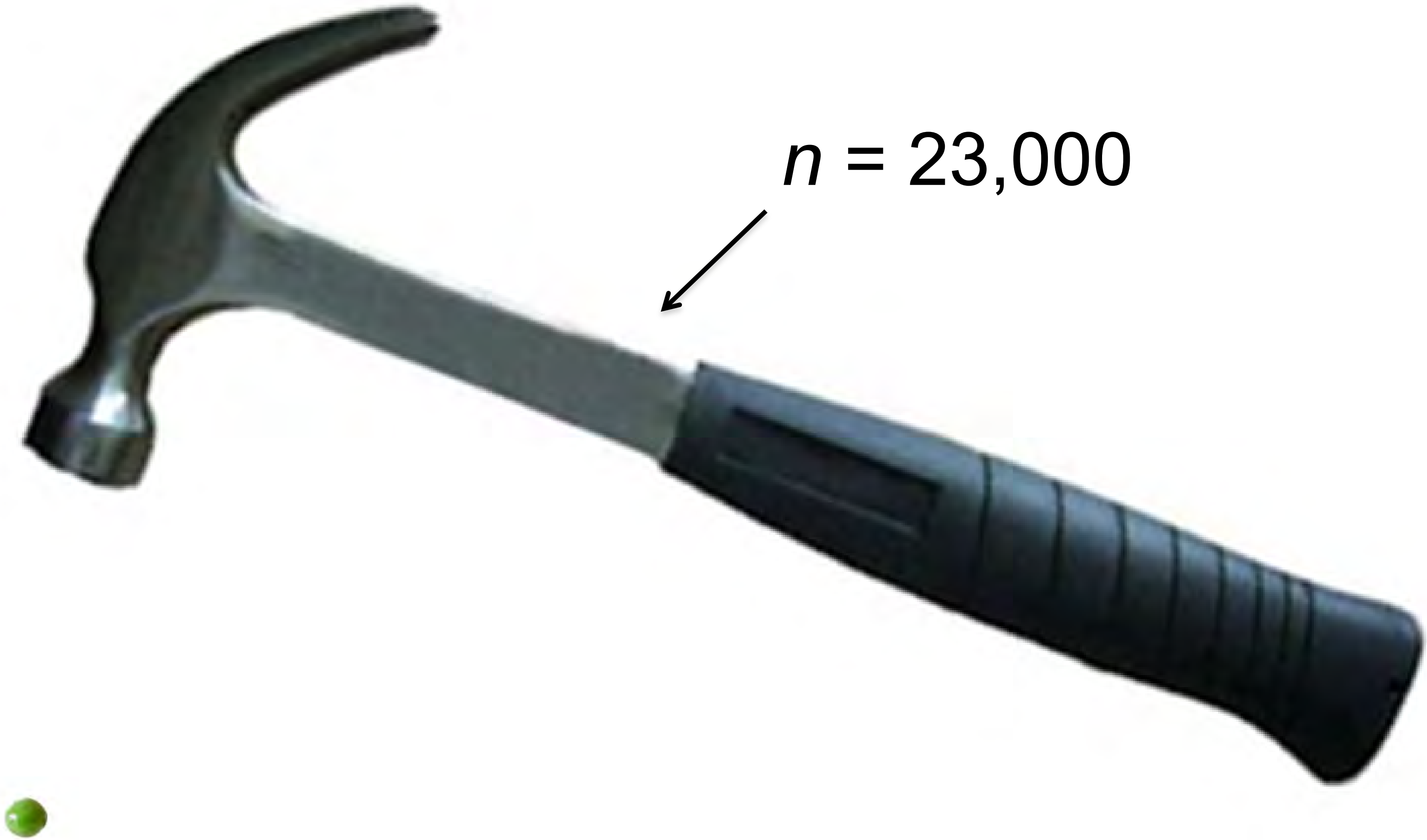
Why is Power Important?

Importance of Power

a priori power analysis can protect against wasting resources.

http://commonsense4worldequality.wordpress.com/2011/12/12/wasting-resources-is-wasting-time-to-come-up-with-real-solutions/

Too much power…



Too little power…

$n = 8$

Importance of Power

a priori power analysis is often required for dissertations, papers, & grant proposals…

http://english.marion.ohiostate.edu/EPs2012/EPs2012-images/grant_money.gif

"Sorry, you guys didn't do a power analysis, so no grant money for you…"

a priori power analysis is an indication of how well you have planned…

Poorly planned studies rarely yield meaningful results…

Determinants of Power

Power is largely determined by…

1. Level of Significance (α)
2. Sample size (*n*)
3. Effect size (e.g., Cohen's *d*, *r*, ω, η^2, partial η^2)

Power can be increased by…

1. Increasing α

$\alpha = .10$

Rejected

Power can be increased by

2. Increasing sample size

n ⟶ ***n***

3. Power can be increased by increasing effect size, but effect size is not always within the researcher's control.

http://teachers.bergencatholic.org/faculty/hornerj/literature3h/frankenstein%20movie%20shots/frankensteinscreenshots.htm

However…

Your research design can improve your effect size and therefore increase power.

For example:

- Add a covariate
- Use the most reliable measures available
- Include multiple sites

How do I know what the effect size is?

Effect size (e.g., Cohen's *d*, *r*, ω, η^2, partial η^2)

- Prior research
- Estimate a range
- Pilot study

Determinants of Power

Table 2.1 *Cohen's effect size benchmarks*

Test	Relevant effect size	Effect size classes: Small	Medium	Large
Comparison of independent means	d, Δ, Hedges' g	.20	.50	.80
Comparison of two correlations	q	.10	.30	.50
Difference between proportions	Cohen's g	.05	.15	.25
Correlation	r	.10	.30	.50
	r^2	.01	.09	.25
Crosstabulation	w, φ, V, C	.10	.30	.50
ANOVA	f	.10	.25	.40
	η^2	.01	.06	.14
Multiple regression	R^2	.02	.13	.26
	f^2	.02	.15	.35

Notes: The rationale for most of these benchmarks can be found in Cohen (1988) at the following pages: Cohen's d (p. 40), q (p. 115), Cohen's g (pp. 147–149), r and r^2 (pp. 79–80), Cohen's w (pp. 224–227), f and η^2 (pp. 285–287), R^2 and f^2 (pp. 413–414).

From The essential guide to effect sizes: Statistical power, meta-analysis, and the interpretation of research results by Paul D. Ellis, 2010.

The Ohio State University

Two Types of Power

A power analysis can either be done …

- <u>before</u> data are collected
    - *a priori* or prospective power analysis

Or… a power analysis can be done …

- <u>after</u> data are collected
  - *post hoc* or retrospective power analysis

The Abuse of Power: The Pervasive Fallacy of Power Calculations for Data Analysis

John M. HOENIG and Dennis M. HEISEY

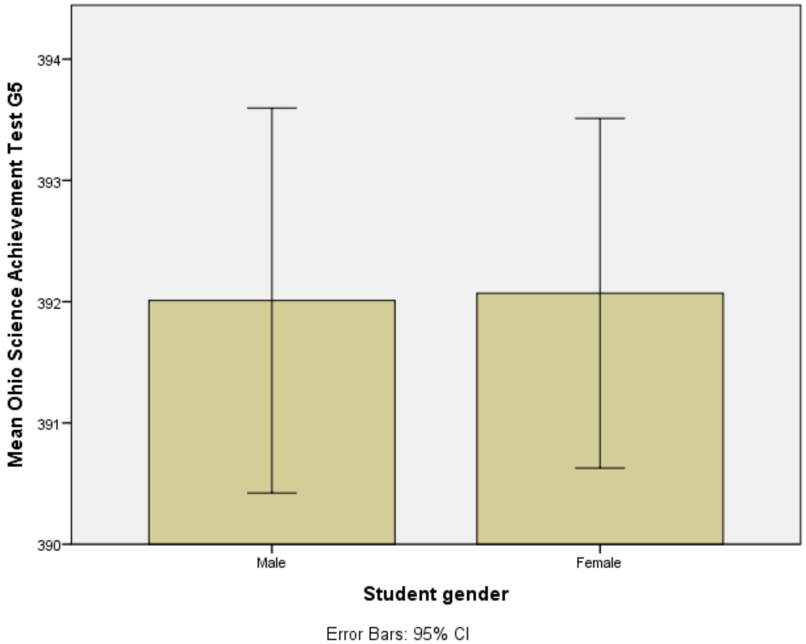
Post Hoc Power Analysis:
An Idea Whose Time Has Passed?

Marc Levine, Ph.D., and Mary H. H. Ensom, Pharm.D., FASHP, FCCP

The Pitfalls of Post Hoc Power

n = 2,600 students
Is there a gender difference on the 5th grade science test?

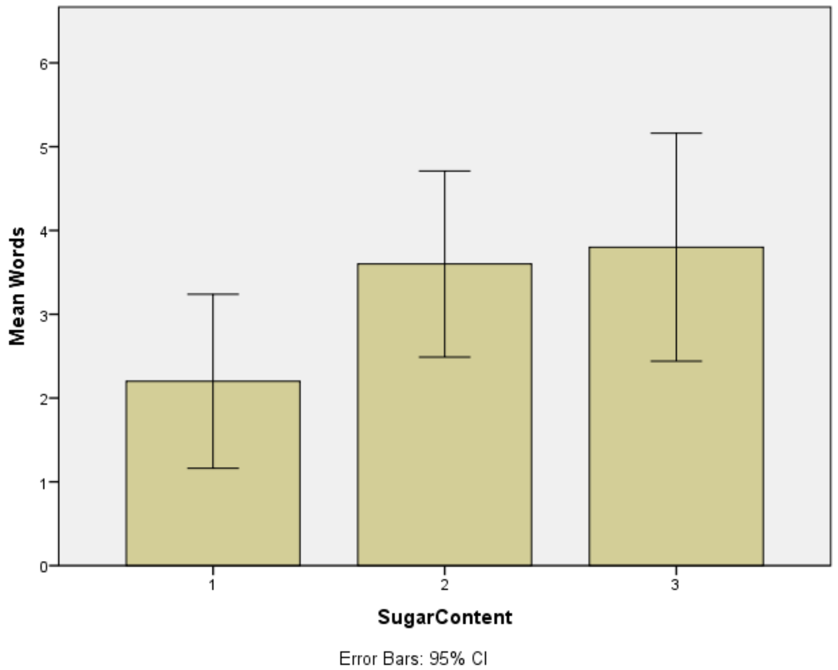
Results:
p = .658
Effect size partial η^2 = .01
Observed power = .07



The Pitfalls of Post Hoc Power

n = 15 subjects
5 in each group
Does the amount of sugar ingested affect memory?

Results:
$p = .055$
Effect size partial $\eta^2 = .384$
Observed power = .57



The Ohio State University

Old School Power

THE OHIO STATE UNIVERSITY

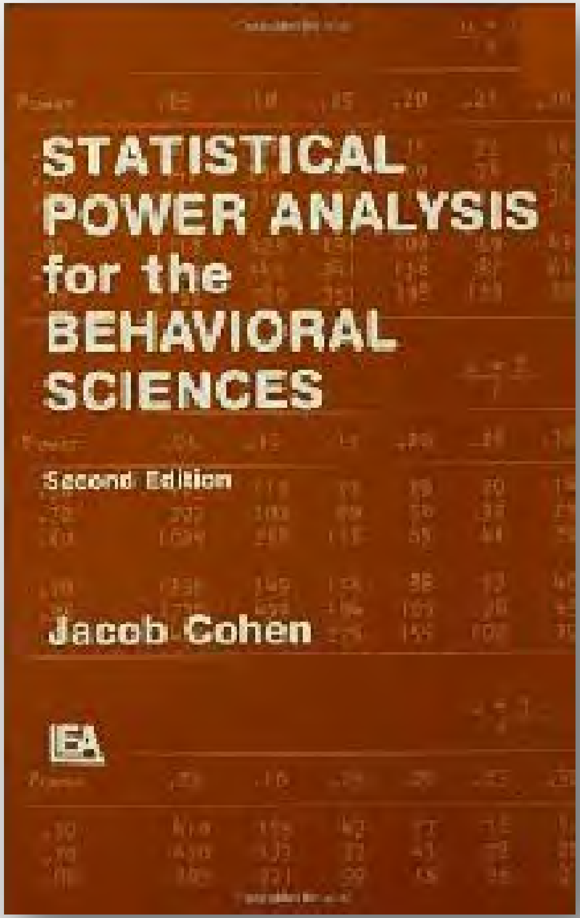


Table 8.3.21

Power of F test at a = .05, u = 15

		f											
n	F_c	.05	.10	.15	.20	.25	.30	.35	.40	.50	.60	.70	.80
2	2.352	05	05	06	07	08	10	12	14	20	28	39	51
3	1.992	05	06	07	09	11	15	19	25	39	57	74	87
4	1.880	05	06	08	11	15	20	28	37	58	78	92	98
5	1.826	05	07	09	13	19	27	38	50	74	91	98	*
6	1.794	05	07	10	15	23	34	47	61	85	96	*	
7	1.772	06	07	11	18	28	41	56	71	92	99		
8	1.757	06	08	12	21	33	48	65	79	96	*		
9	1.745	06	08	14	24	38	55	72	85	98			
10	1.736	06	09	15	27	43	61	78	90	99			
11	1.729	06	09	17	30	47	67	83	93	*			
12	1.724	06	10	18	33	52	72	87	96				
13	1.719	06	10	20	36	57	77	90	97				
14	1.715	06	11	21	39	61	81	93	98				
15	1.711	06	11	23	42	65	84	95	99				
16	1.708	06	12	25	45	69	87	96	99				
17	1.706	07	12	26	48	72	90	97	*				
18	1.704	07	13	28	51	76	92	98					
19	1.702	07	14	30	54	78	93	99					
20	1.700	07	14	31	57	81	95	99					
21	1.698	07	15	33	60	84	96	99					
22	1.696	07	16	35	63	86	97	*					
23	1.695	07	16	37	65	88	97						
24	1.694	07	17	39	68	89	98						
25	1.693	07	17	40	70	91	98						
26	1.692	07	18	42	72	92	99						
27	1.691	08	19	44	74	93	99						
28	1.690	08	20	46	75	94	99						
29	1.689	08	20	47	78	95	*						
30	1.688	08	21	49	80	96							
31	1.687	08	22	51	82	97							
32	1.687	08	22	52	83	97							
33	1.686	08	23	54	84	98							
34	1.686	08	24	56	86	98							
35	1.685	09	25	57	87	98							
36	1.684	09	25	59	88	99							
37	1.684	09	26	60	89	99							
38	1.683	09	27	62	90	99							
39	1.683	09	28	63	91	99							

The Ohio State University

Power Software

Power Analysis Software

Software Version and Pricing

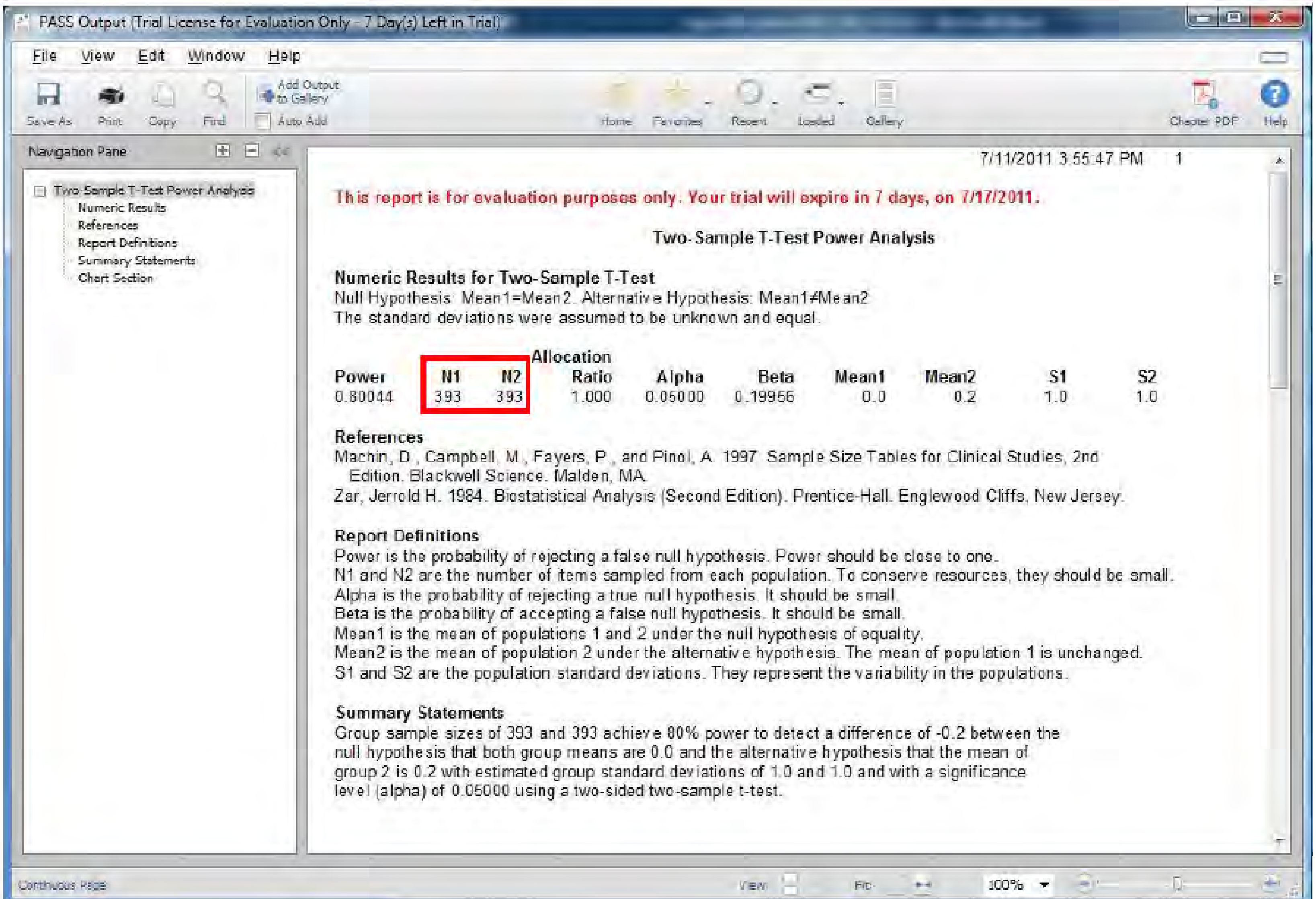
Software	*Current Version*	*Pricing*	*A Priori*	*Post Hoc*	*Graphics*
G*Power	3.1.9	Free	✓	✓	✓
PASS	13	$795	✓	✓	✓
SAS	9.3	Free for OSU students. Contact OCIO for Faculty/Staff licenses	✓		✓
Stata	13	Gradplan prices vary by University and STATA module	✓	✓	
SPSS/Statistics	21	Free for OSU students. Contact OCIO for Faculty/Staff licenses		✓	
SPSS/SamplePower	3.0.1	$720/year	✓		✓
Optimal Design	3.01	Free	✓		✓

Power Analysis Software

Software Version and Pricing

Software	*Current Version*	*Pricing*	*Website*
G*Power	3.1.9	Free	http://www.gpower.hhu.de/en.html
PASS	13	$795	http://www.ncss.com/software/pass/
SAS	9.3	Free for OSU students. Contact OCIO for Faculty/Staff licenses	
Stata	13	Gradplan prices vary by University and STATA module	http://www.stata.com/order/new/edu/gradplans/campus-gradplan/
SPSS/Statistics	21	Free for OSU students. Contact OCIO for Faculty/Staff licenses	
SPSS/SamplePower	3.0.1	$720/year	http://www-03.ibm.com/software/products/en/spss-samplepower
Optimal Design	3.01	Free	http://sitemaker.umich.edu/group-based/optimal_design_software

Screenshots of PASS


PASS Output (Trial License for Evaluation Only - 7 Day(s) Left in Trial)

File View Edit Window Help

Save As Print Copy Find Add Output to Gallery Auto Add Home Favorites Recent Loaded Gallery Chapter PDF Help

Navigation Pane

Two-Sample T-Test Power Analysis
Numeric Results
References
Report Definitions
Summary Statements
Chart Section

7/11/2011 3:55:47 PM 1

This report is for evaluation purposes only. Your trial will expire in 7 days, on 7/17/2011.

Two-Sample T-Test Power Analysis

Numeric Results for Two-Sample T-Test
Null Hypothesis: Mean1=Mean2. Alternative Hypothesis: Mean1≠Mean2
The standard deviations were assumed to be unknown and equal.

Power	N1	N2	Allocation Ratio	Alpha	Beta	Mean1	Mean2	S1	S2
0.80044	393	393	1.000	0.05000	0.19956	0.0	0.2	1.0	1.0

References
Machin, D., Campbell, M., Fayers, P., and Pinol, A. 1997. Sample Size Tables for Clinical Studies, 2nd Edition. Blackwell Science. Malden, MA.
Zar, Jerrold H. 1984. Biostatistical Analysis (Second Edition). Prentice-Hall. Englewood Cliffs, New Jersey.

Report Definitions
Power is the probability of rejecting a false null hypothesis. Power should be close to one.
N1 and N2 are the number of items sampled from each population. To conserve resources, they should be small.
Alpha is the probability of rejecting a true null hypothesis. It should be small.
Beta is the probability of accepting a false null hypothesis. It should be small.
Mean1 is the mean of populations 1 and 2 under the null hypothesis of equality.
Mean2 is the mean of population 2 under the alternative hypothesis. The mean of population 1 is unchanged.
S1 and S2 are the population standard deviations. They represent the variability in the populations.

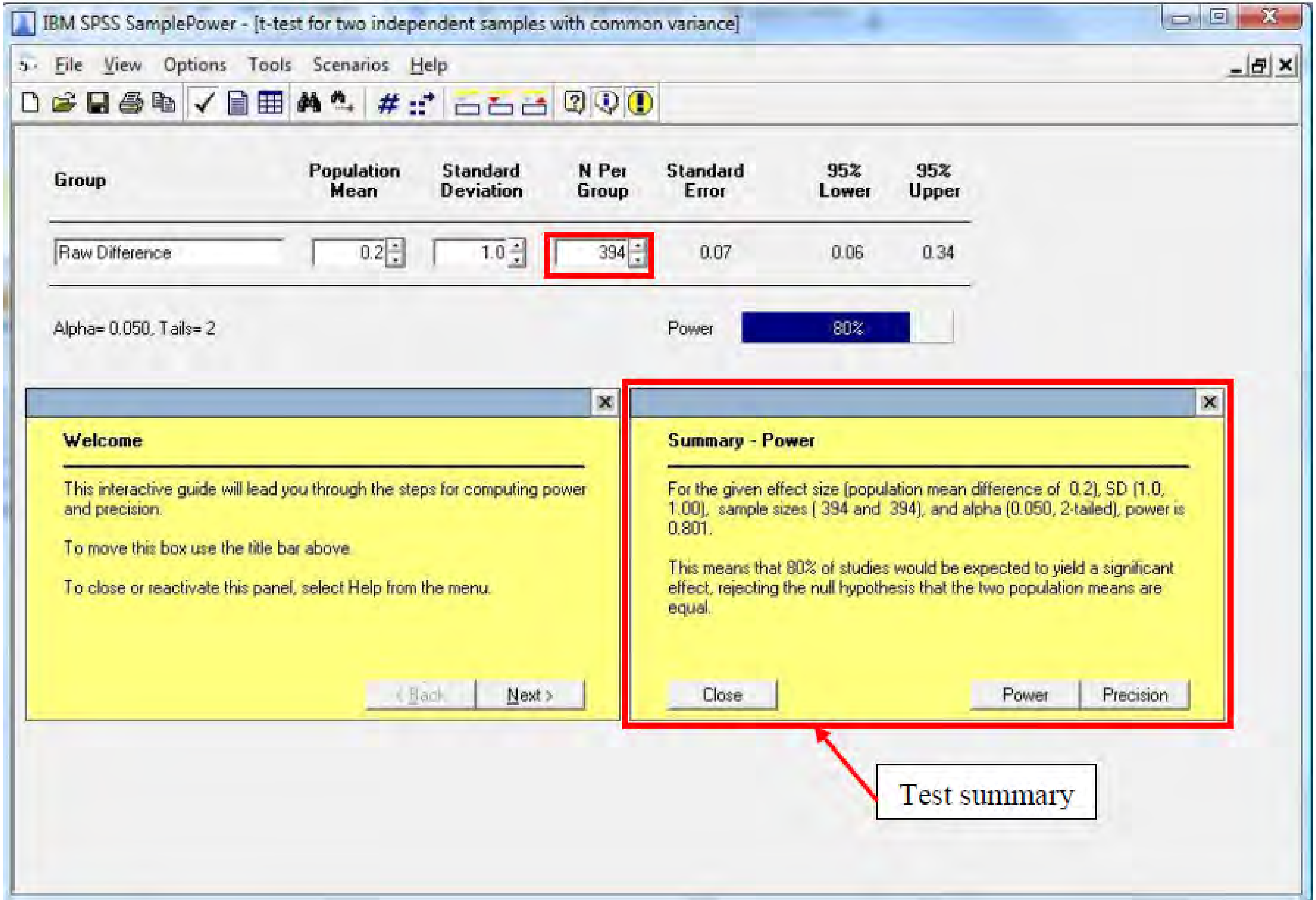
Summary Statements
Group sample sizes of 393 and 393 achieve 80% power to detect a difference of -0.2 between the null hypothesis that both group means are 0.0 and the alternative hypothesis that the mean of group 2 is 0.2 with estimated group standard deviations of 1.0 and 1.0 and with a significance level (alpha) of 0.05000 using a two-sided two-sample t-test.

Continuous Page View: Fit: 100%

Peng, C. Y. J., Long, H., & Abaci, S. (2012). Power Analysis Software for Educational Researchers. *The Journal of Experimental Education, 80*(2), 113-136.

Power Analysis Software

Screenshots of SPSS SamplePower



Peng, C. Y. J., Long, H., & Abaci, S. (2012). Power Analysis Software for Educational Researchers. *The Journal of Experimental Education, 80*(2), 113-136.

Screenshots of SAS

The SAS System

The POWER Procedure
Two-Sample t Test for Mean Difference

Fixed Scenario Elements	
Distribution	Normal
Method	Exact
Number of Sides	2
Group 1 Mean	0.2
Group 2 Mean	0
Standard Deviation	1
Nominal Power	0.8
Null Difference	0
Alpha	0.05

Computed N Per Group	
Actual Power	N Per Group
0.801	394

Peng, C. Y. J., Long, H., & Abaci, S. (2012). Power Analysis Software for Educational Researchers. *The Journal of Experimental Education, 80*(2), 113-136.

Screenshots of STATA

Stata 12 Syntax

Example 1: One-sample *t*-test (2-tailed)

Power = .80 *Hypothesiswise α = .05* *Effect Size = 0.2*

Syntax:

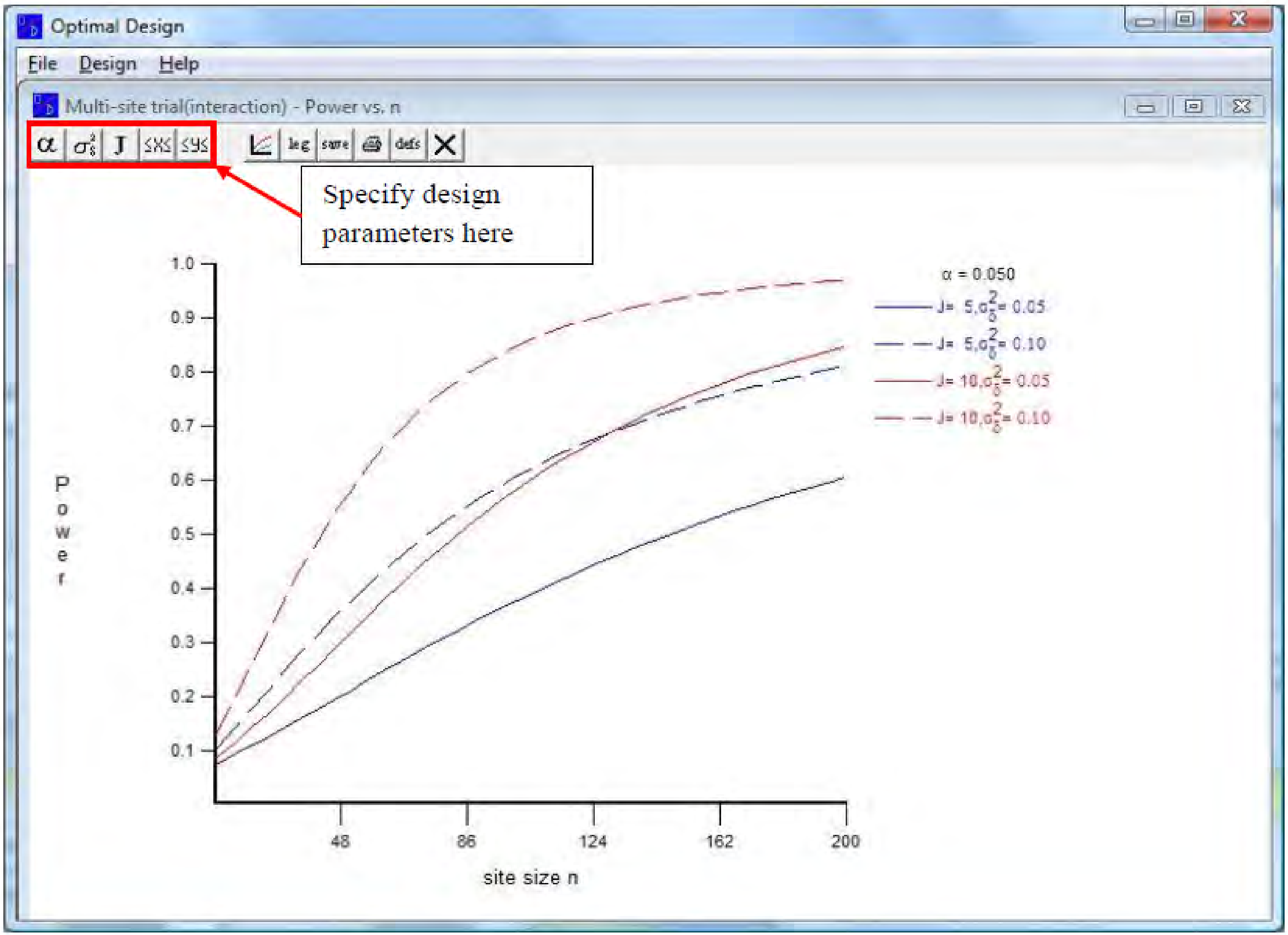
```
sampsi 0.2 0, sd(1) power(.8) onesamp
```

Output:

```
Estimated sample size for one-sample comparison of mean
 to hypothesized value

Test Ho: m =    .2, where m is the mean in the population

Assumptions:
        alpha          =  0.0500  (two-sided)
        power          =  0.8000
        alternative m  =  0
        sd             =  1

Estimated required sample size:
           n =     197
```

Peng, C. Y. J., Long, H., & Abaci, S. (2012). Power Analysis Software for Educational Researchers. *The Journal of Experimental Education*, *80*(2), 113-136.

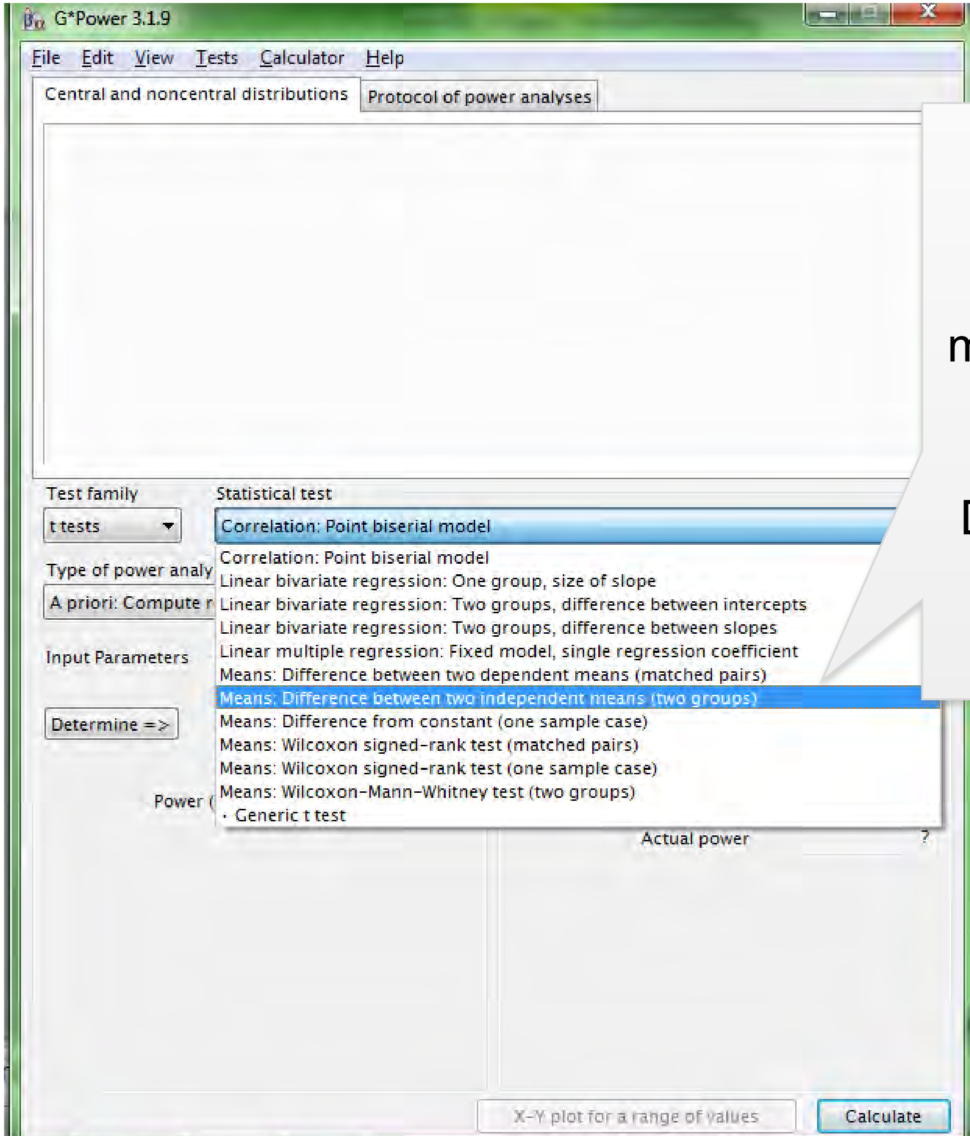
Screenshots of Optimal Design



Peng, C. Y. J., Long, H., & Abaci, S. (2012). Power Analysis Software for Educational Researchers. *The Journal of Experimental Education, 80*(2), 113-136.

Examples using G*Power

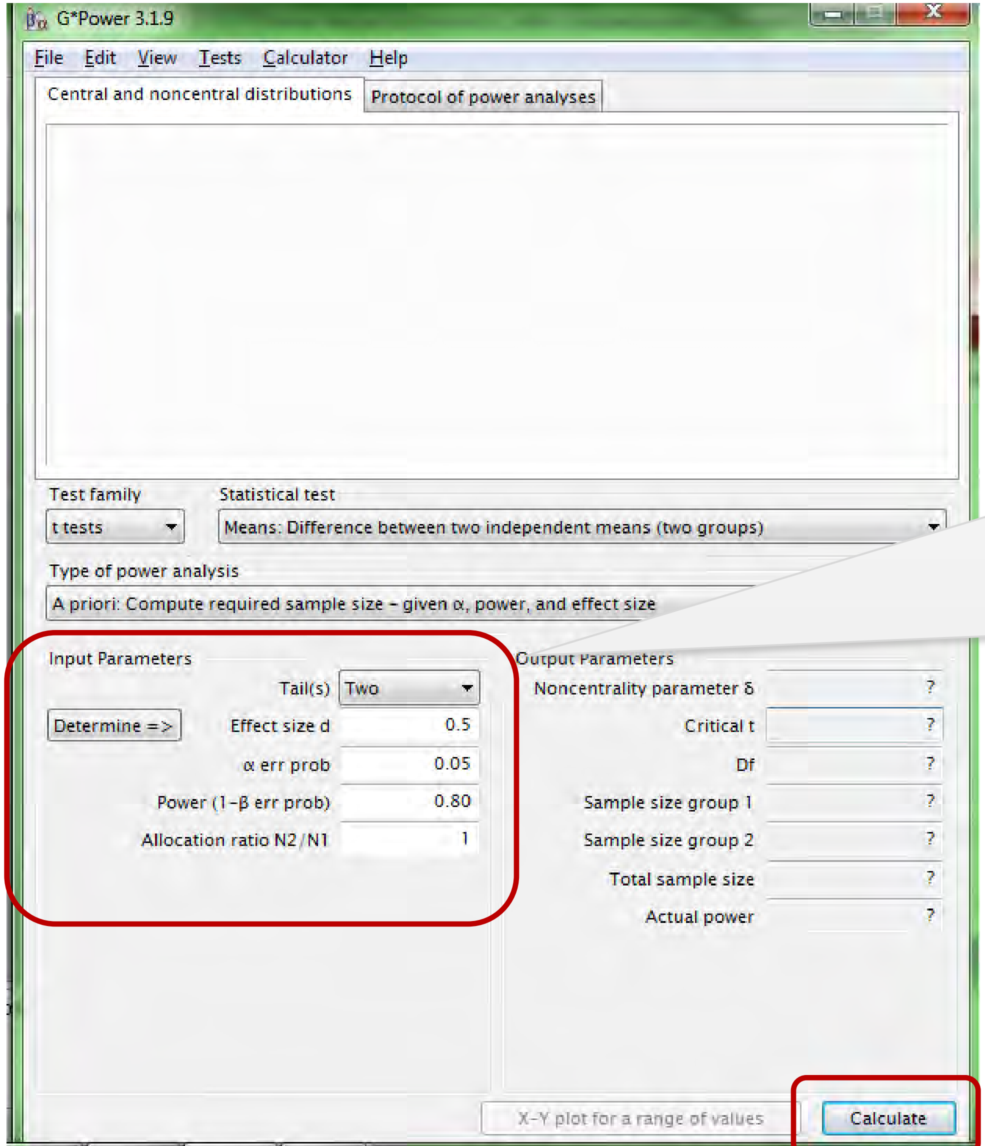
Power Analysis Software



Two-Sample t-test: A priori power

The default selection for “Statistical Test” is “Correlation: Point biserial model.” Use the arrow to toggle to the desired statistical test. For the two-sample t-test, we need “Means: Difference between two independent means (two groups).”

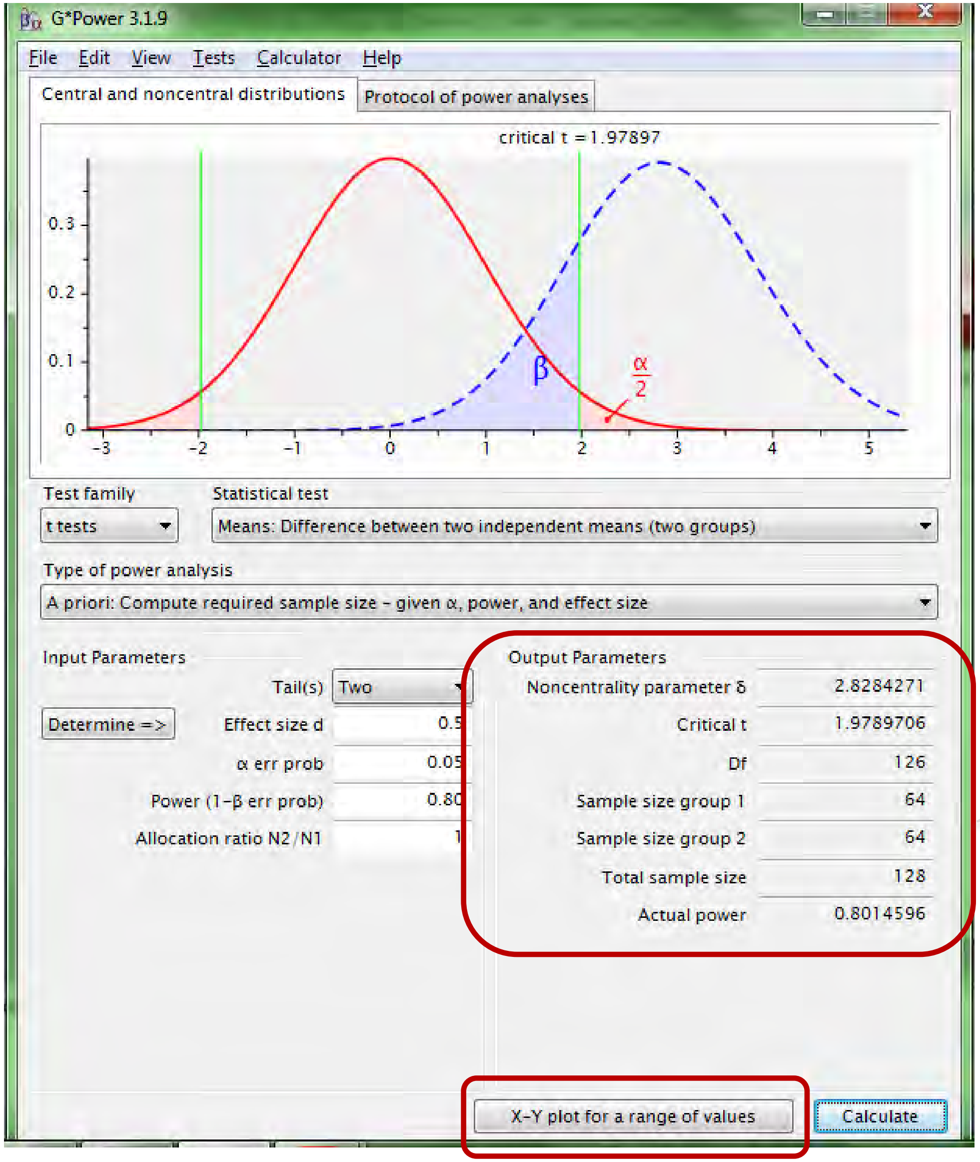
Two-Sample t-test: A priori power



The "Input Parameters" for computing *a priori* power must be specified:

1. One versus two tailed test;
2. Anticipated effect size *d*;
3. Alpha level;
4. Desired power;
5. Allocation ratio (of *n*'s).

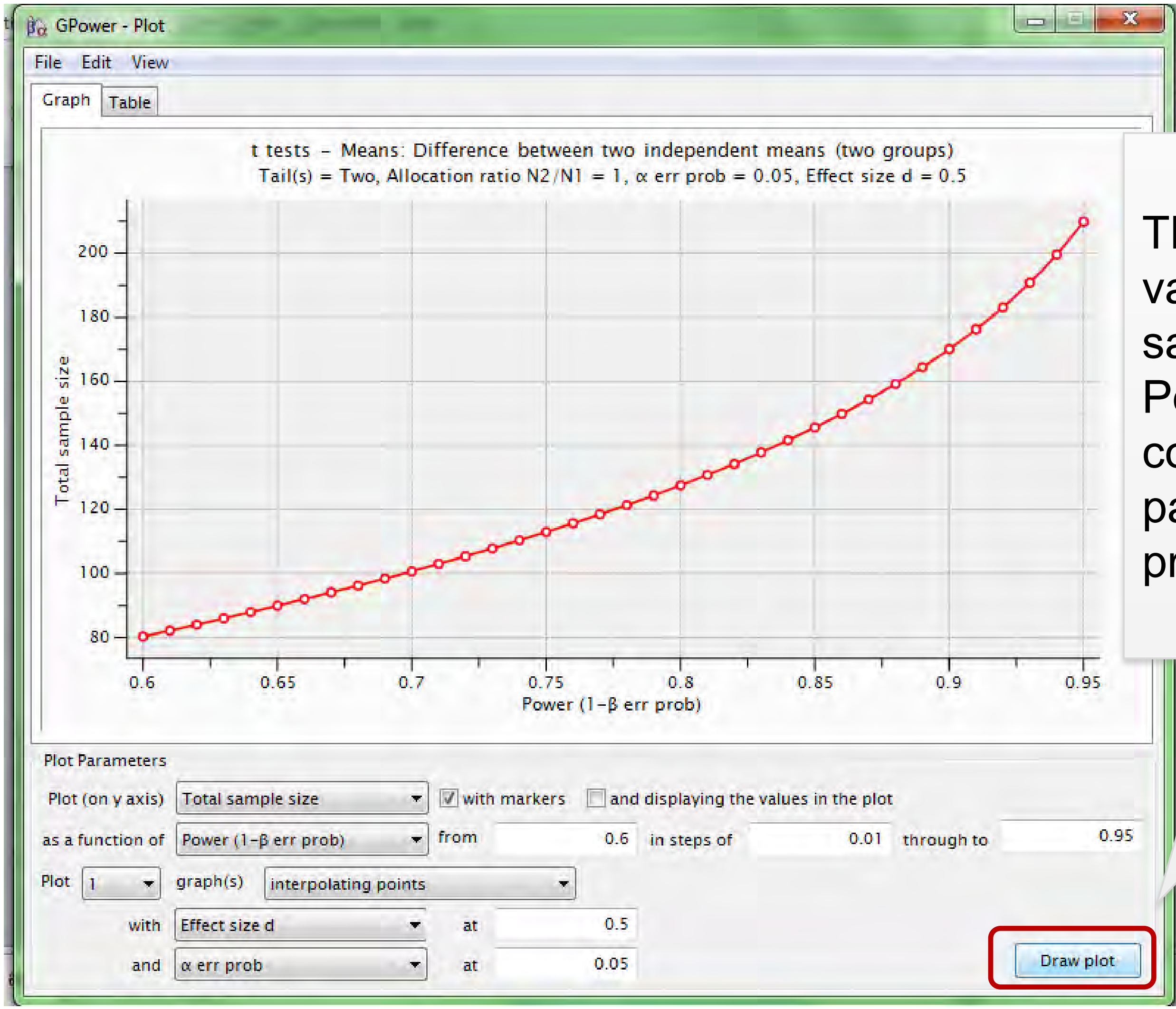
Two-Sample t-test: A priori power



The "Output Parameters" provide the relevant statistics given the input specified. Based on the parameters specified, we need a sample size of 128 for our two sample t-test.

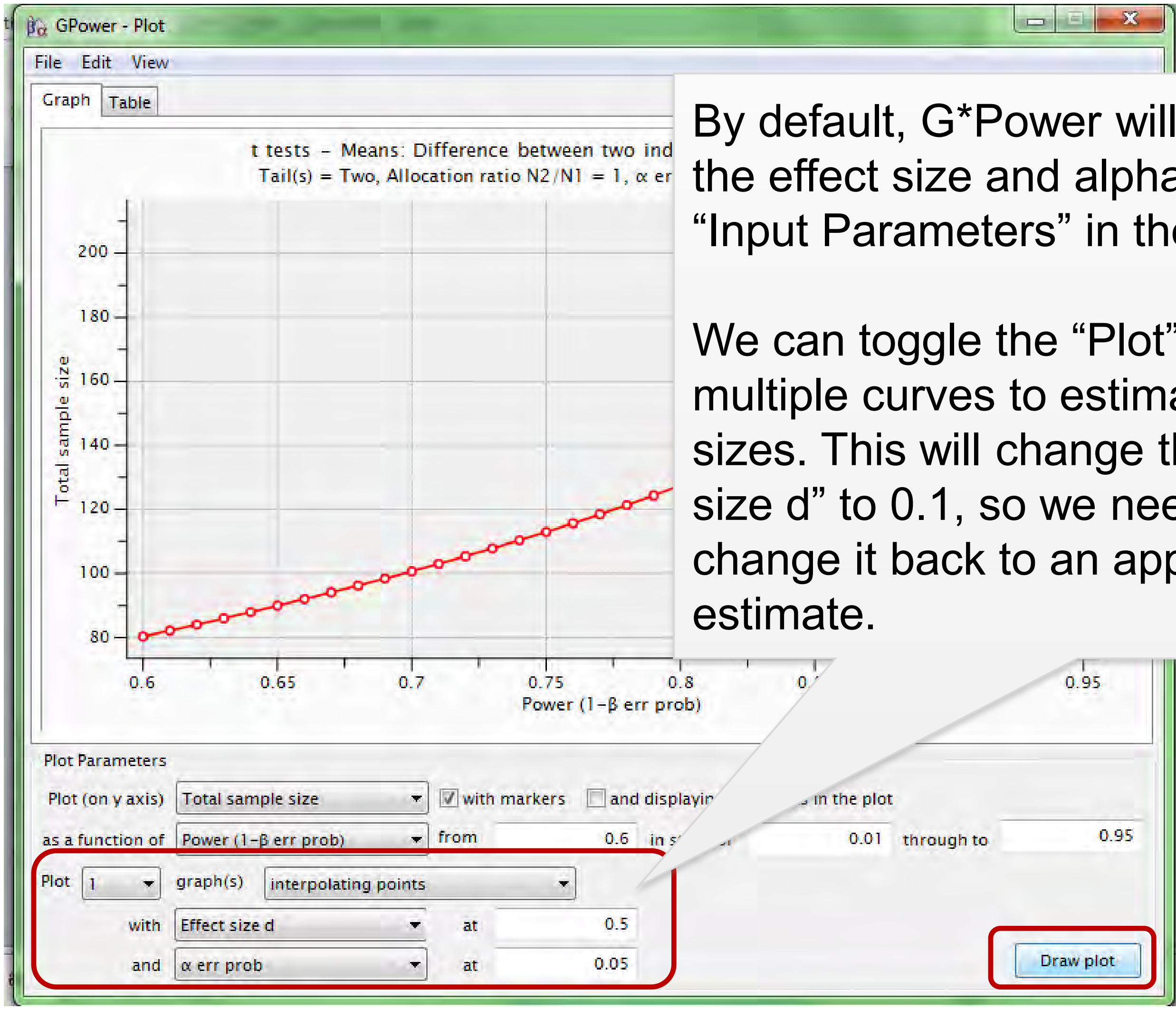
Power Analysis Software

Two-Sample t-test: A priori power



The “X-Y plot for a range of values” will plot the Total sample size as a function of Power. This plot can be copied from G*Power and pasted into your paper or presentation.

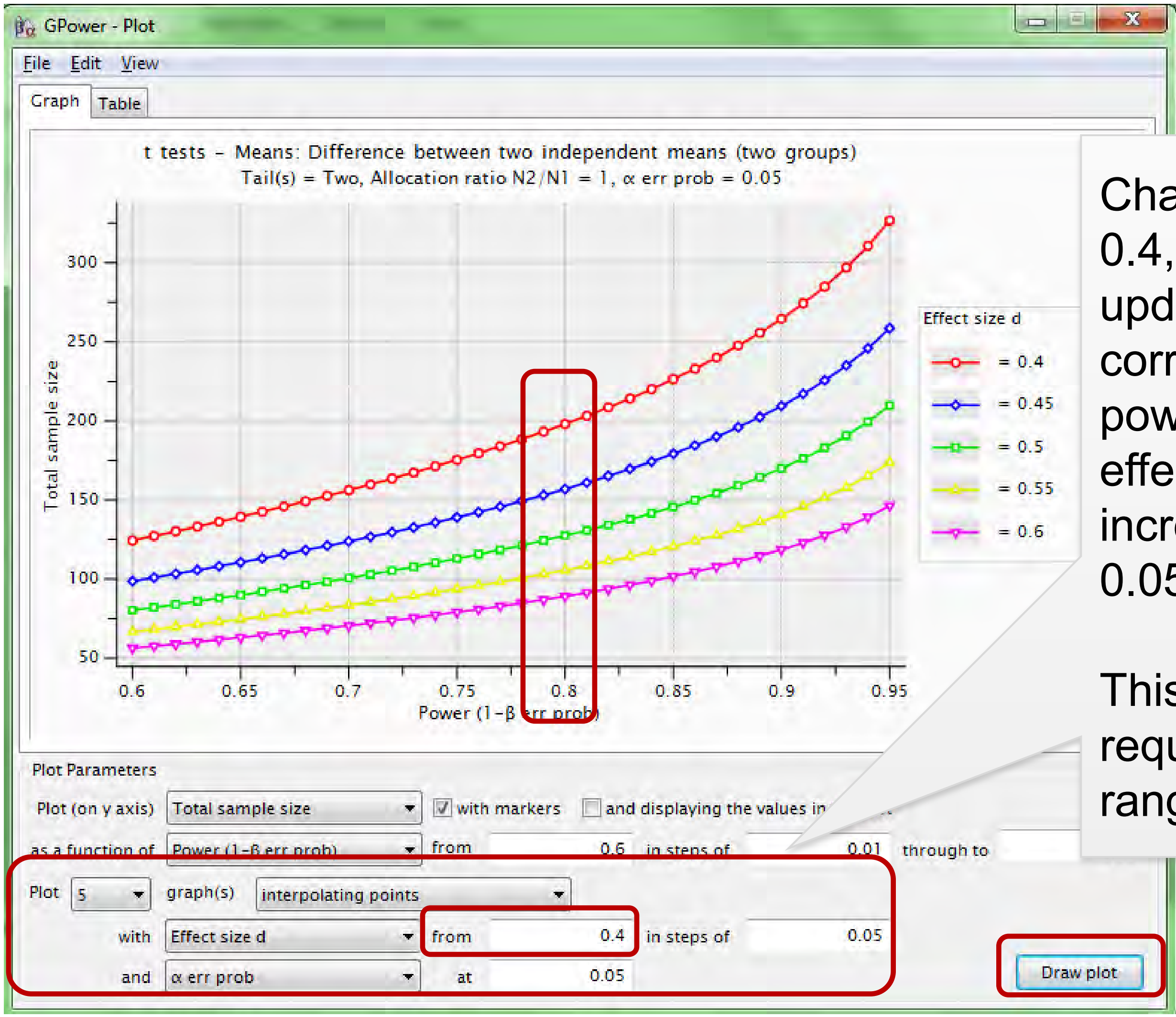
Two-Sample t-test: A priori power



By default, G*Power will plot 1 curve using the effect size and alpha level specified under "Input Parameters" in the main window.

We can toggle the "Plot" number to plot multiple curves to estimate a range of effect sizes. This will change the value of "effect size d" to 0.1, so we need to remember to change it back to an appropriate effect size estimate.

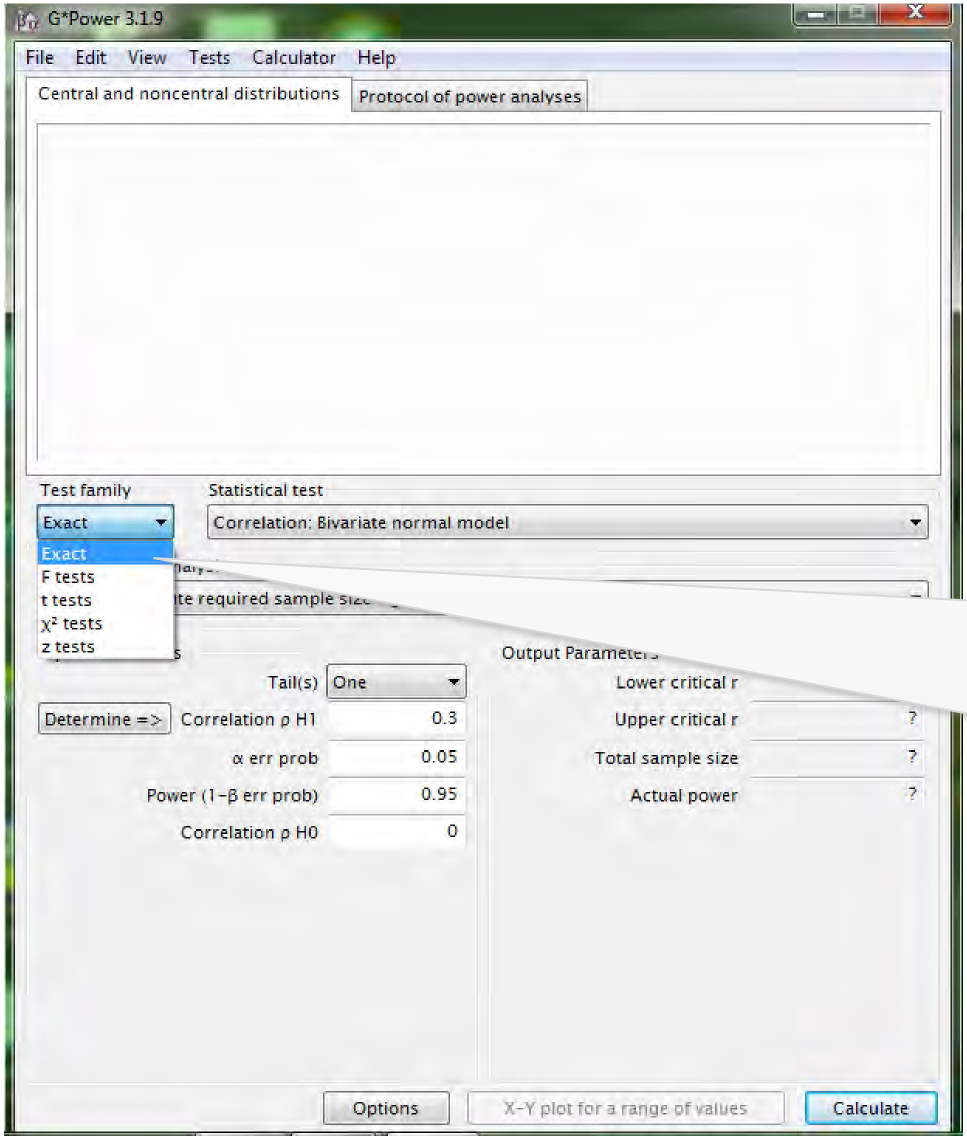
Two-Sample t-test: A priori power



Change "Plot" to 5, "from" to 0.4, and click "Draw plot" to update the graph with the correct values. This will plot 5 power curves, starting with an effect size of 0.4 and increasing in increments of 0.05.

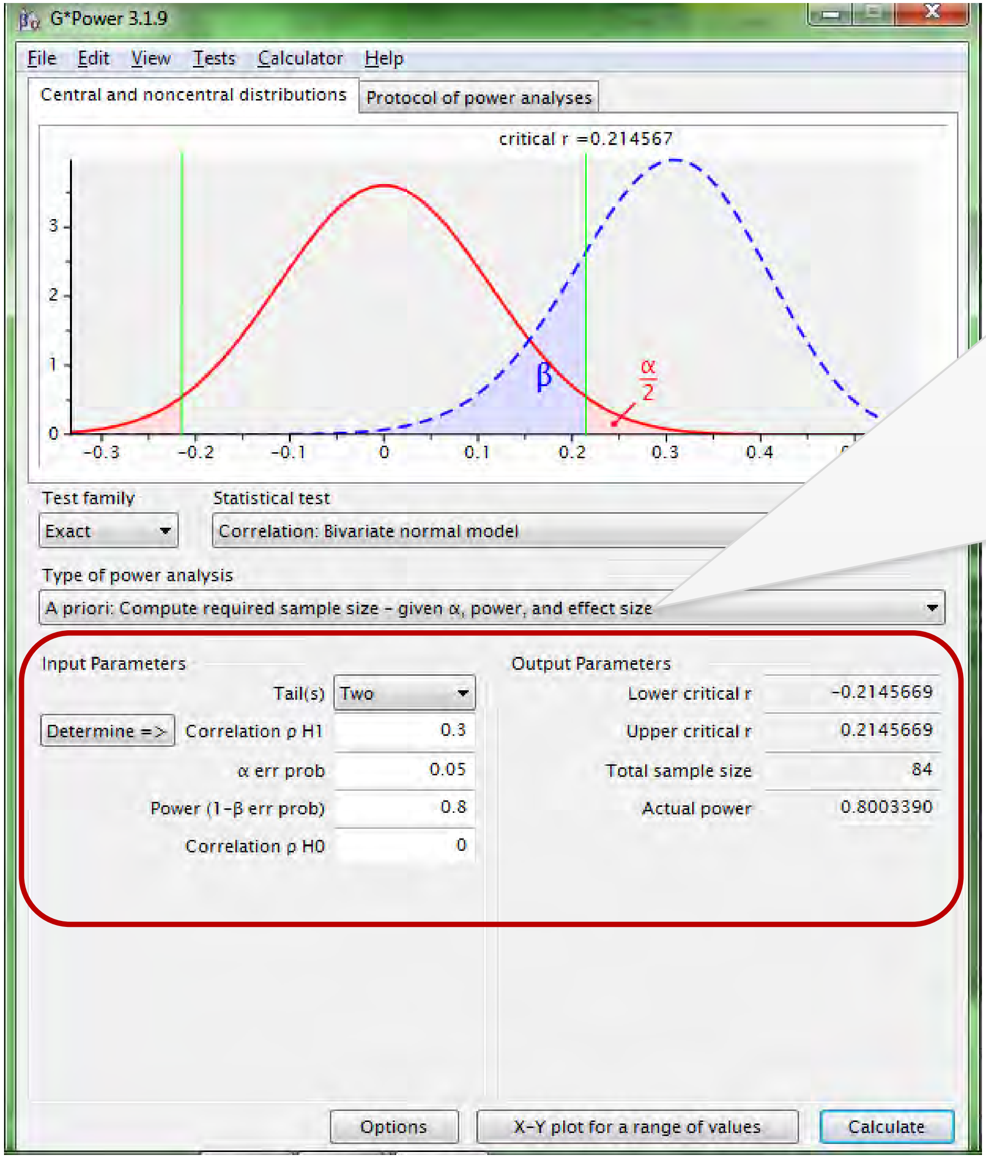
This allows us to estimate the required sample size for a range of effect sizes.

Correlation: A priori power



The default selection for "Test family" is "t tests." Use the arrow to toggle to "Exact." We'll use the default selection for "Statistical test," which is "Correlation: Bivariate normal model."

Correlation: A priori power

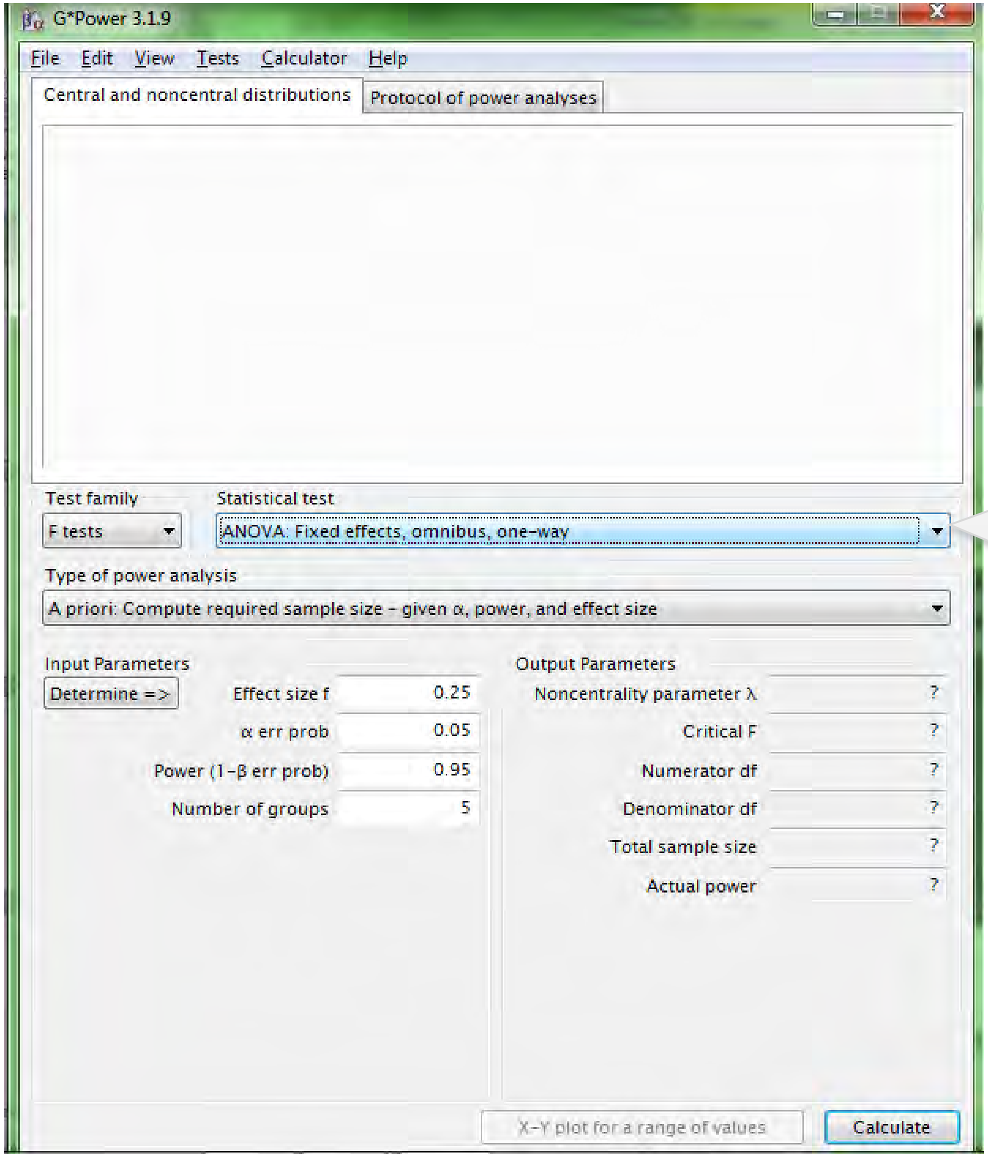


The “Input Parameters” for computing *a priori* power must be specified:

1. One or two tailed test
2. Anticipated correlation coefficient value
3. Alpha level
4. Desired power
5. Correlation coefficient under the null hypothesis (H0)

“Output Parameters” provides the relevant statistics given the input specified. Based on the parameters specified, we need a sample size of 84.

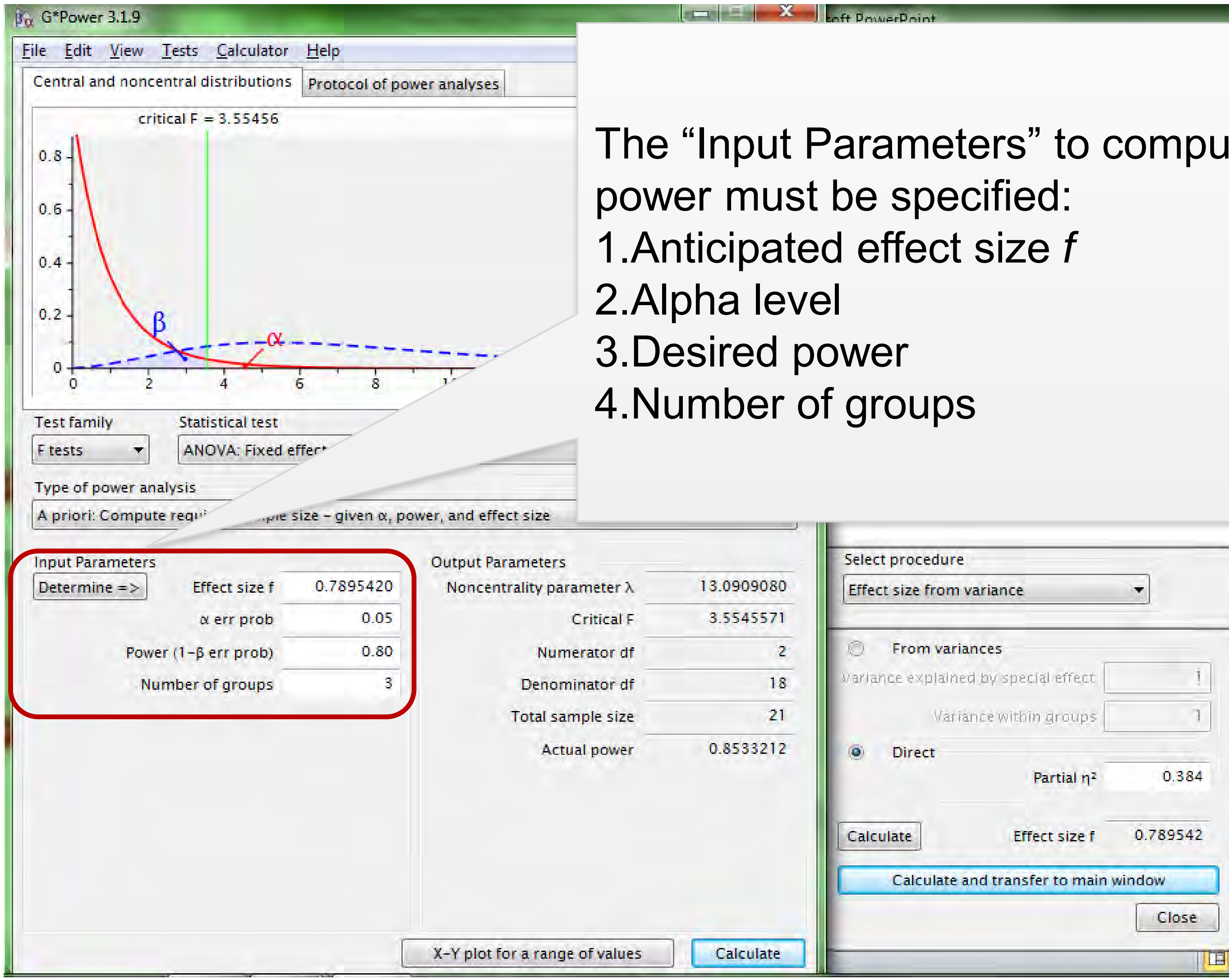
One-factor ANOVA: A priori power



Use the arrow to toggle the "Test family" to "F tests." Use the arrow to toggle "Statistical test" to select "ANOVA: Fixed effects, omnibus, one-way."

Power Analysis Software

One-factor ANOVA: A priori power

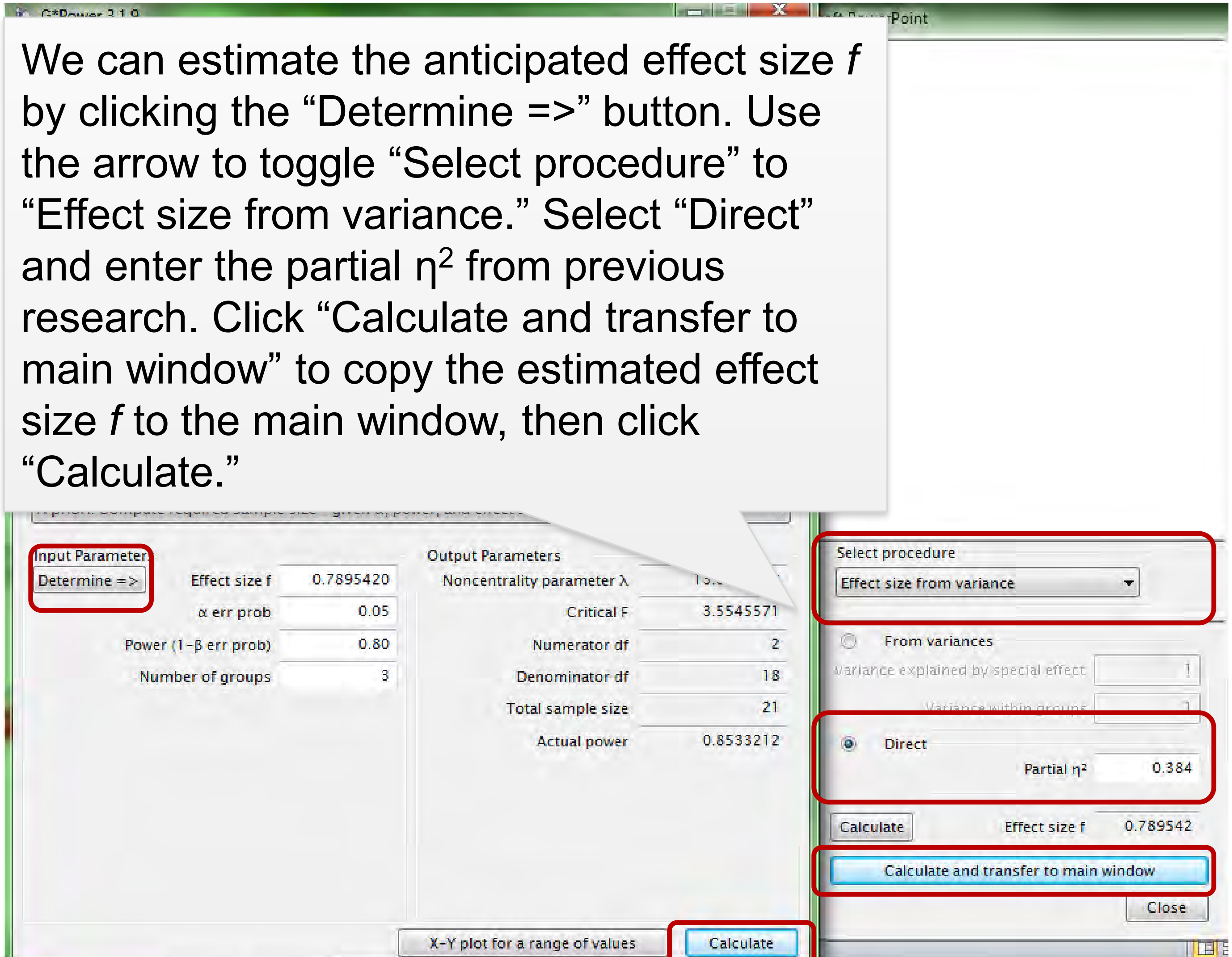


The "Input Parameters" to compute *a priori* power must be specified:

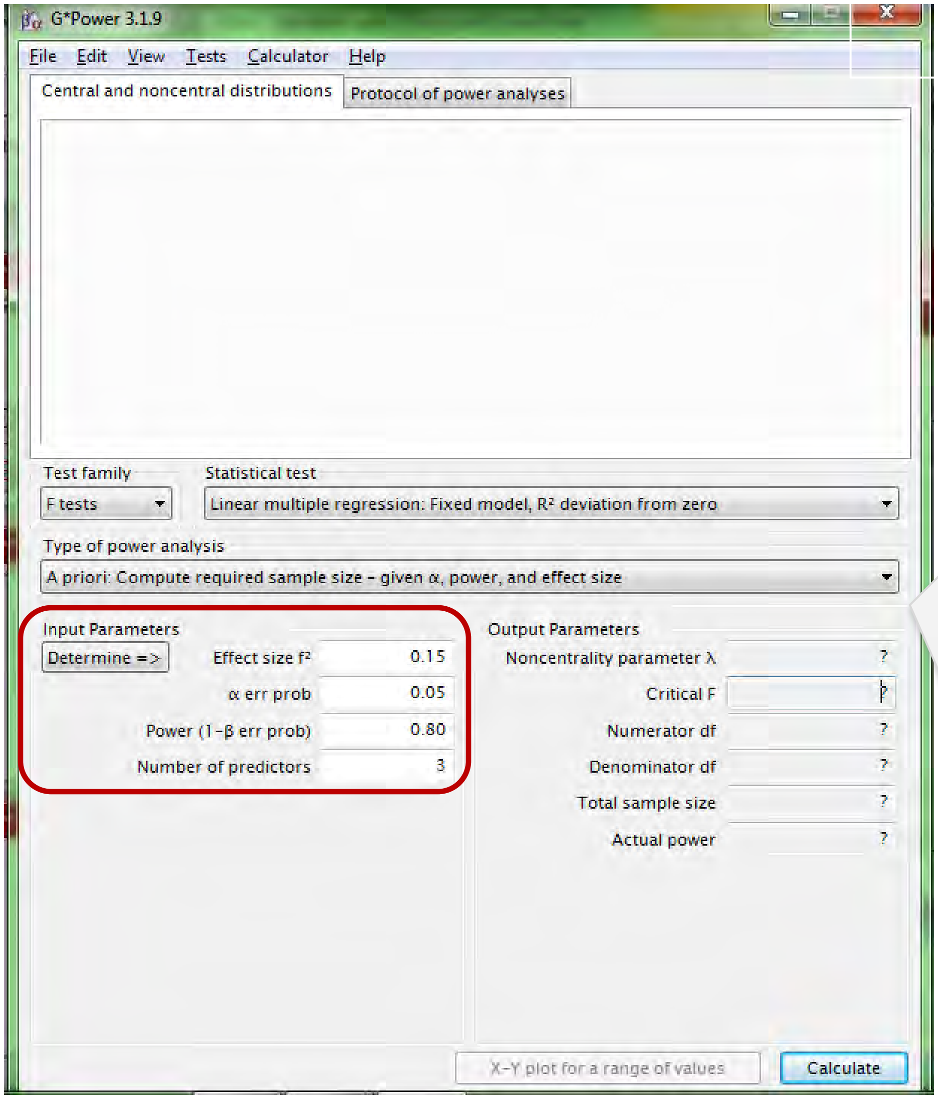
1. Anticipated effect size *f*
2. Alpha level
3. Desired power
4. Number of groups

One-factor ANOVA: A priori power

We can estimate the anticipated effect size *f* by clicking the “Determine =>” button. Use the arrow to toggle “Select procedure” to “Effect size from variance.” Select “Direct” and enter the partial η^2 from previous research. Click “Calculate and transfer to main window” to copy the estimated effect size *f* to the main window, then click “Calculate.”



Multiple Linear Regression: A priori power



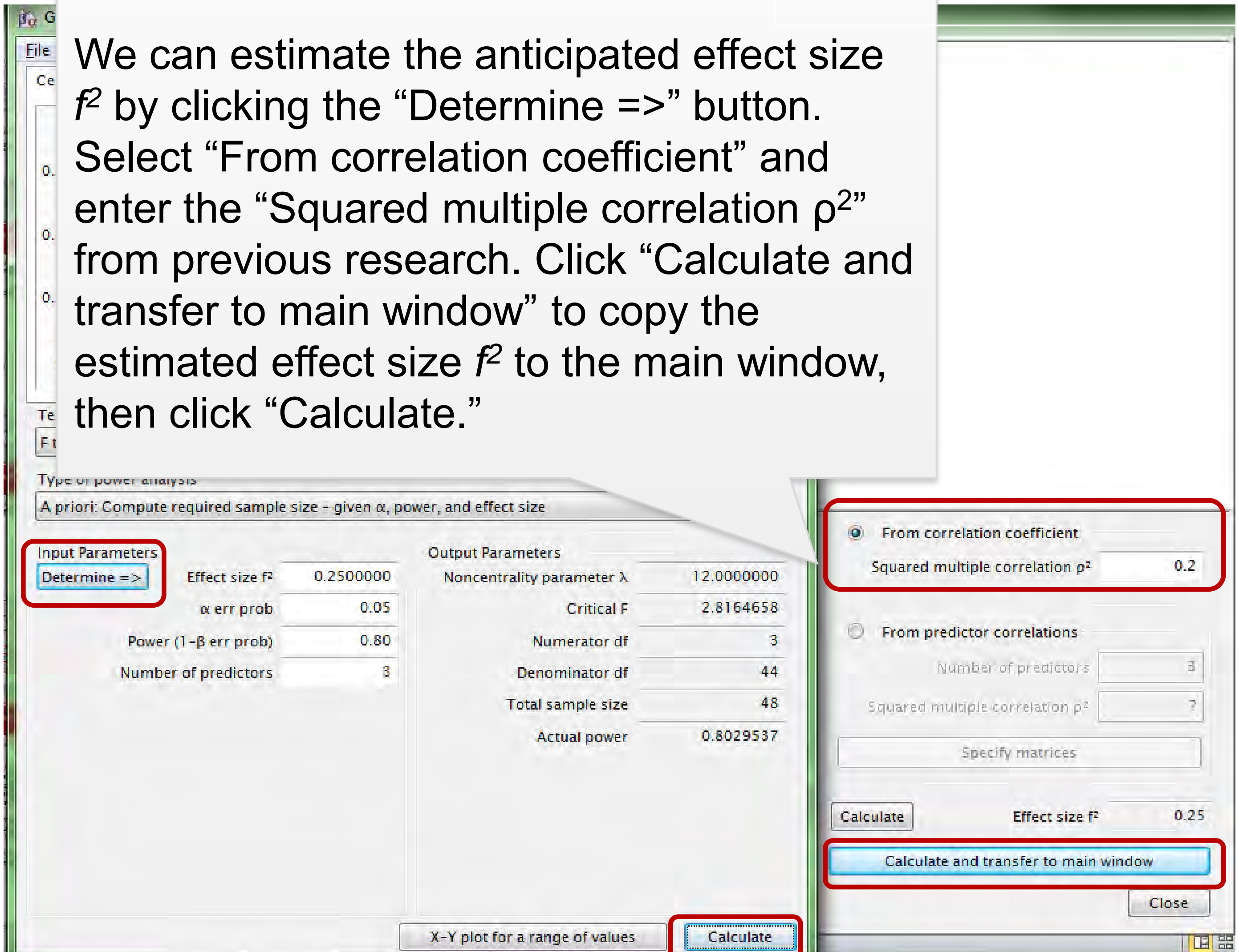
Use the arrow to toggle the “Test family” to “F tests.” Use the arrow to toggle “Statistical test” to select “Linear multiple regression: Fixed model, R^2 deviation from zero.”

The “Input Parameters” to compute *a priori* power must be specified:

1. Anticipated Effect size f^2
2. Alpha level
3. Desired power
4. Number of predictors in your model.

Multiple Linear Regression: A priori power

We can estimate the anticipated effect size f^2 by clicking the "Determine =>" button. Select "From correlation coefficient" and enter the "Squared multiple correlation ρ^2" from previous research. Click "Calculate and transfer to main window" to copy the estimated effect size f^2 to the main window, then click "Calculate."



G*Power 3 Website and Free Download

http://www.gpower.hhu.de/en.html

References

Abraham, W. T., & Russell, D. W. (2008). Statistical power analysis in psychological research. *Social and Personality Psychology Compass*, *2*(1), 283-301.

Cohen, J. 1988. *Statistical power analysis for the behavioral sciences* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Ellis, P. D. (2010). *The essential guide to effect sizes: Statistical power, meta-analysis, and the interpretation of research results*. Cambridge University Press.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior research methods*, *39*(2), 175-191.

Hoenig, J. M., & Heisey, D. M. (2001). The abuse of power. *The American Statistician*, *55*(1).

Levine, M., & Ensom, M. H. (2001). Post hoc power analysis: an idea whose time has passed? *Pharmacotherapy: The Journal of Human Pharmacology and Drug Therapy*, *21*(4), 405-409.

Murphy, K. R. & Myors, B. (2004). *Statistical power analysis: A simple and general model for traditional and modern hypothesis tests* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

O'Keefe, D. J. (2007). Brief report: post hoc power, observed power, a priori power, retrospective power, prospective power, achieved power: sorting out appropriate uses of statistical power analyses. *Communication Methods and Measures*, *1*(4), 291-299.

Peng, C. Y. J., Long, H., & Abaci, S. (2012). Power analysis software for educational researchers. *The Journal of Experimental Education*, *80*(2), 113-136.

Thomas, L., & Krebs, C. J. (1997). A review of statistical power analysis software. *Bulletin of the Ecological Society of America*, *78*(2), 126-138.

The Essential Guide to Effect Sizes: Statistical Power, Meta-analysis, and Interpretation of Research Results

by Paul D. Ellis is available as an e-book through the OSU library.

5

NO IMAGE AVAILABLE

The essential guide to effect sizes [electronic resource] : statistical power, meta-analysis, and the interpretation of research results / Paul D. Ell

Ellis, Paul D., 1969-

Cambridge ; New York : Cambridge University Press, 2010

Book

LOCATION	CALL NO.	YEAR	STATUS	NOTE
Web E-books	Q180.55.S7 E45 2010eb		USE ONLINE	