**Power Analysis for Correlation and Regression Models**

**G\*Power – Regression Model**

[G\*Power](https://www.psychologie.hhu.de/arbeitsgruppen/allgemeine-psychologie-und-arbeitspsychologie/gpower.html) is available free, for PC and for Macs, and is designed for the regression model (Y is random but the predictors are fixed). The R2 program (discussed below) is designed for correlation analysis (all variables are random). Under most circumstances you will get the similar results from R2 and G\*Power. For example, suppose I ask how much power I would have for a large effect, alpha = .05, *n* = 5, one predictor.

**Using G\*Power, correlation, point biserial**

You expect to have only 5 scores to test the correlation between two variables. As usual, you will use the .05 alpha. One of the variables is group membership and the other is a normally distributed variable. In other words, you are doing a regression analysis, not a correlation analysis. You boot up G\*Power:

A screenshot of a cell phone

Description automatically generated

The “Test family” is set to “t tests.” That works, *t* can be used to the test the linear association between two variables. “Statistical test” is set to “Correlation.” That works, you will be testing the correlation between two variables. “Type of power analysis” is set to “A priori.” That would work if you wanted to know how many subjects you would need to obtain a desired amount of power, but that is not the situation at hand. You expect to have only 5 subjects and want to know how much power that would give you. So you change “Type of power analysis” to “Post hoc.” “Tails” is set to “One,” that is, directional hypotheses, but these are almost never employed. Change it to “Two.” “Effect size” is set to “0.3,” but you expect a large effect. Change it to “0.5.” Alpha at .05 is fine. “Total sample size” is set to 100, but you expect only 5, so change it to 5.

A screenshot of a cell phone

Description automatically generated

Now just click “Calculate.”

A screenshot of a cell phone

Description automatically generated

Even with a large effect you will have only 15% power. You need a larger sample size.

Click on the “Protocol of power analysis” to get the settings and results in a more concise format:

**t tests -** Correlation: Point biserial model

**Analysis:** Post hoc: Compute achieved power

**Input:** Tail(s) = Two

Effect size |r| = 0.5

α err prob = 0.05

Total sample size = 5

**Output:** Noncentrality parameter δ = 1.290994

Critical t = 3.182446

Df = 3

Power (1-β err prob) = 0.151938

* Warning: The default setting for “Tails” is “One,” which is appropriate for directional hypotheses, which rarely are tested. My most common mistake in G\*Power is to forget to set it to Two Tails.
* The expected value of *t* moves away from 0, in the negative or the positive direction, as the value of rho moves away from zero, in either direction. Accordingly, we use a two-tailed test for nondirectional hypotheses. This is not true if we elect to use *F* as the test statistic. The expected value of *F* only increases as the expected value of rho-squared increases, whether that results from the value of *r* moving away from 0 in the negative direction or away from zero in the positive direction. Accordingly, with *F*, we use a one-tailed test for nondirectional hypotheses.

**Equivalently, using G\*Power, Multiple regression, omnibus *R2***

**F tests -** Multiple Regression: Omnibus (R² deviation from zero)

**Analysis:** Post hoc: Compute achieved power

**Input:** Effect size f² = 0.3333333

α err prob = 0.05

Total sample size = 5

Number of predictors = 1

**Output:** Noncentrality parameter λ = 1.666667

Critical F = 10.127964

Numerator df = 1

Denominator df = 3

Power (1-β err prob) = 0.151938

Here G\*Power uses Cohen’s *f2* effect size statistic, which is *R2* / (1-*R2*). For a rho of .5, that is .25/.75 = .333333333. The *F* test here is appropriately one-tailed with nondirectional hypotheses, and G\*Power appropriately returns exactly the same value of power as it did with the two-tailed *t* test. Do note both analysis here have been for regression analysis.

**Correlation, Bivariate Normal Model, Exact Test Family**

One can test the null hypothesis that rho is a particular value. That particular value need not be zero. To get the power for such a test, select Test Family = Exact. Here I try this routine but with the null value being zero and the true value .5.

**Exact -** Correlation: Bivariate normal model

**Options:** exact distribution

**Analysis:** Post hoc: Compute achieved power

**Input:** Tail(s) = Two

Correlation ρ H1 = 0.5

α err prob = 0.05

Total sample size = 5

Correlation ρ H0 = 0

**Output:** Lower critical r = -0.8783394

Upper critical r = 0.8783394

Power (1-β err prob) = 0.1311324

Notice that the estimated value of power is not what it was when I used the Point Biserial Model or the Multiple Regression Model. Why? A point biserial *r* comes from a regression model, since the dichotomous variable is fixed rather than random. Here G\*Power identifies the model as “Bivariate Normal,” which implies a correlation model. Maybe that is the difference. [Eric Heidel](https://www.scalelive.com/) stated that the [Point Biserial Model](https://www.scalelive.com/sample-size-for-point-biserial.html) should be used only if the analysis is a regression analysis (predicting a normally distributed random variable from a fixed variable), with the [Bivariate Normal Model](https://www.scalelive.com/sample-size-for-pearsons-r.html) being used for a correlation analysis (predicting one random variable from another, both normally distributed)

Now for a null with no zero. The correlation between grades and IQ is typically found to be about .5. I have a population for which I expect the correlation not to be .5. I ask how many cases I’ll need for 80% power if the true rho is .8?

**Exact -** Correlation: Bivariate normal model

**Options:** exact distribution

**Analysis:** A priori: Compute required sample size

**Input:** Tail(s) = Two

Correlation ρ H1 = 0.7

α err prob = 0.05

Power (1-β err prob) = 0.80

Correlation ρ H0 = .5

**Output:** Lower critical r = 0.3177643

Upper critical r = 0.6503523

Total sample size = 80

Actual power = 0.8009210

What if the actual correlation is .3?

**Exact -** Correlation: Bivariate normal model

**Options:** exact distribution

**Analysis:** A priori: Compute required sample size

**Input:** Tail(s) = Two

Correlation ρ H1 = 0.3

α err prob = 0.05

Power (1-β err prob) = 0.80

Correlation ρ H0 = .5

**Output:** Lower critical r = 0.3653659

Upper critical r = 0.6164209

Total sample size = 139

Actual power = 0.8012822

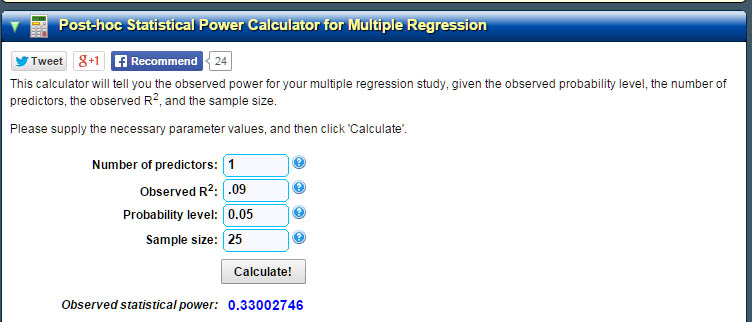
Wow, the power analysis is affected not only by the magnitude of the difference between the true rho and the null rho but the direction of that difference.

Other Alternatives

There are several online calculators for computing power for linear correlation/regression analyses, but my experiences has been that many of them produce incorrect results. I have found one that seems to work correctly, producing results very close to those produced by G\*Power. It can be found here: <http://www.danielsoper.com/statcalc3/calc.aspx?id=9> .

To use this calculator, just enter the number of predictor variables (for bivariate regression, this will be 1), assumed value of *r2* – note that is *r-squared* -- , the alpha level, the sample size, and then click Calculate.

For example, suppose that you have 25 pairs of scores and want to determine what your chances are getting significant results if the correlation between the two variables is, in the population, medium. Cohen defined a medium rho as .3. This calculator requires rho-squared, so you will enter the value .09.



Now for the G\*Power solution.

**t tests -** Correlation: Point biserial model

**Analysis:** Post hoc: Compute achieved power

**Input:** Tail(s) = Two

Effect size |ρ| = 0.3

α err prob = 0.05

Total sample size = 25

**Output:** Noncentrality parameter δ = 1.5724273

Critical t = 2.0686576

Df = 23

Power (1-β err prob) = 0.3255412

Close enough for me.

The power table in Howell can also be used to estimate the power for a correlation/regression design. This is explained in my document [Bivariate Linear Correlation](http://core.ecu.edu/psyc/wuenschk/docs01/Corr2101.pdf).

Power analysis for *r* is exceptionally simple:, assuming that *df* are large enough for *t* to be approximately normal. Cohen’s benchmarks for effect sizes for *r* are: **.10** is small but not necessarily trivial, **.30** is medium, and **.50** is large (Cohen, J. A Power Primer, *Psychological Bulletin*, 1992, *112*, 155-159). For the example here, . Using the Power Table in our textbook, power would be .29 for δ = 1.4 and .32 for δ = 1.5. 1.47 is seven tenths of the way from 1.4 to 1.5, so our power is estimated to be .29 + .7(.32-.29) = .311.

**Multiple Regression/Correlation**

Consider the research published in the article: Patel, S., Long, T. E., McCammon, S. L., & Wuensch, K. L. (1995). [Personality and emotional correlates of self-reported antigay behaviors](http://core.ecu.edu/psyc/wuenschk/Articles/JInterpersonalViolence95.htm). *Journal of Interpersonal Violence, 10,* 354-366. We had data from 80 respondents. We wished to predict self-reported antigay behaviors from five predictor variables. Suppose we wanted to know how much power we would have if the population *ρ2* was .13 (a medium sized effect according to Cohen). All of the variables are random rather than fixed, so a multiple correlation analysis will be conducted.

**Exact -** Linear multiple regression: Random model

**Options:** Exact distribution

**Analysis:** Post hoc: Compute achieved power

**Input:** Tail(s) = One

H1 ρ² = .13

H0 ρ² = 0

α err prob = 0.05

Total sample size = 80

Number of predictors = 5

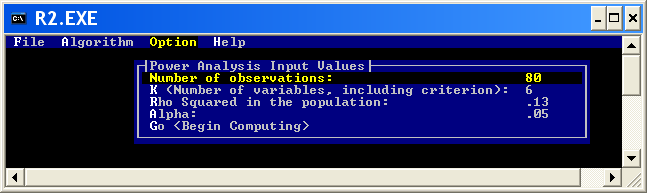
**Output:** Lower critical R² = 0.1364360

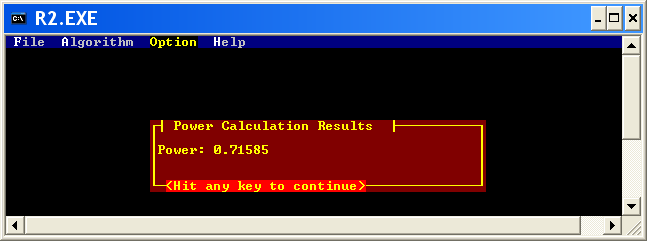
Upper critical R² = 0.1364360

Power (1-β err prob) = 0.7158714

Notice that I specified a one-tailed test. That is because ρ2 cannot be less than zero. The null hypothesis is ρ2 = 0 and the alternative hypothesis is ρ2 > 0.

Many years ago Steiger and Fouladi created a DOS program to do power analysis for the random model (it also will construct confidence intervals for ρ2). It produces results nearly identical to those of G\*Power.





Suppose that all of the predictor variables were fixed rather than random (most psychologists would call that analysis an ANOVA rather than a multiple regression). I used “Determine” to convert ρ2 to f2 and then calculated power to be 73%. Power for the random effects model will never be greater than for the fixed effects model, and the difference will be small.

Graphical user interface, text, application

Description automatically generated

When would one ever use a two-tailed test with a multiple correlation? Only when the null is a nonzero value. For example, suppose that the correlation relating freshmen year grade point averages to SAT and high school grades was usually about .5, but you had reason to suspect that for one particular group of students it would be different value. Your null is ρ2 = .5 and your alternative is ρ2 ≠ .5. How much power would you have if the actual ρ2 were .3 and you had 50 subjects?

**Exact -** Linear multiple regression: Random model

**Options:** Exact distribution

**Analysis:** Post hoc: Compute achieved power

**Input:** Tail(s) = Two

H1 ρ² = .3

H0 ρ² = .5

α err prob = 0.05

Total sample size = 50

Number of predictors = 2

**Output:** Lower critical R² = 0.3031232

Upper critical R² = 0.6899616

Power (1-β err prob) = 0.4422557

You would reject the null if the sample *R2* fell outside of the range .303 to .690.

What if you had directional hypotheses, null ρ2 ≥ .5 and alternative ρ2 < .5, and the true ρ2 = .3.

**Options:** Exact distribution

**Analysis:** Post hoc: Compute achieved power

**Input:** Tail(s) = One

H1 ρ² = .3

H0 ρ² = .5

α err prob = 0.05

Total sample size = 50

Number of predictors = 2

**Output:** Lower critical R² = 0.3385471

Upper critical R² = 0.3385471

Power (1-β err prob) = 0.5692433

Notice that you would have more power than with the nondirectional hypotheses. You would reject the null if the sample *R2* were less than .3385.

A person wearing glasses

Description automatically generated

[Return to Wuensch’s Statistics Lesson Page](http://core.ecu.edu/psyc/wuenschk/StatsLessons.htm)

[Karl L. Wuensch](http://core.ecu.edu/psyc/wuenschk/klw.htm), Dept. of Psychology, East Carolina University, October, 2020.