The Stata Journal

The *Stata Journal* publishes reviewed papers together with shorter notes or comments, regular columns, book reviews, and other material of interest to Stata users. Examples of the types of papers include 1) expository papers that link the use of Stata commands or programs to associated principles, such as those that will serve as tutorials for users first encountering a new field of statistics or a major new technique; 2) papers that go "beyond the Stata manual" in explaining key features or uses of Stata that are of interest to intermediate or advanced users of Stata; 3) papers that discuss new commands or Stata programs of interest either to a wide spectrum of users (e.g., in data management or graphics) or to some large segment of Stata users (e.g., in survey statistics, survival analysis, panel analysis, or limited dependent variable modeling); 4) papers analyzing the statistical properties of new or existing estimators and tests in Stata; 5) papers that could be of interest or usefulness to researchers, especially in fields that are of practical importance but are not often included in texts or other journals, such as the use of Stata in managing datasets, especially large datasets, with advice from hard-won experience; and 6) papers of interest to those who teach, including Stata with topics such as extended examples of techniques and interpretation of results, simulations of statistical concepts, and overviews of subject areas.

The *Stata Journal* is indexed and abstracted by *CompuMath Citation Index*, *Current Contents/Social and Behavioral Sciences*, *RePEc: Research Papers in Economics*, *Science Citation Index Expanded* (also known as *SciSearch*), *Scopus*, and *Social Sciences Citation Index*.

For more information on the *Stata Journal*, including information for authors, see the webpage

http://www.stata-journal.com

Subscriptions are available from StataCorp, 4905 Lakeway Drive, College Station, Texas 77845, telephone 979-696-4600 or 800-STATA-PC, fax 979-696-4601, or online at

http://www.stata.com/bookstore/sj.html

Subscription rates listed below include both a printed and an electronic copy unless otherwise mentioned.

U.S. and Canada		Elsewhere	
Printed & electronic		**Printed & electronic**	
1-year subscription	$ 98	1-year subscription	$138
2-year subscription	$165	2-year subscription	$245
3-year subscription	$225	3-year subscription	$345
1-year student subscription	$ 75	1-year student subscription	$ 99
1-year institutional subscription	$245	1-year institutional subscription	$285
2-year institutional subscription	$445	2-year institutional subscription	$525
3-year institutional subscription	$645	3-year institutional subscription	$765
Electronic only		**Electronic only**	
1-year subscription	$ 75	1-year subscription	$ 75
2-year subscription	$125	2-year subscription	$125
3-year subscription	$165	3-year subscription	$165
1-year student subscription	$ 45	1-year student subscription	$ 45

Back issues of the *Stata Journal* may be ordered online at

http://www.stata.com/bookstore/sjj.html

Individual articles three or more years old may be accessed online without charge. More recent articles may be ordered online.

http://www.stata-journal.com/archives.html

The *Stata Journal* is published quarterly by the Stata Press, College Station, Texas, USA.

Address changes should be sent to the *Stata Journal*, StataCorp, 4905 Lakeway Drive, College Station, TX 77845, USA, or emailed to sj@stata.com.

The *Stata Journal* (ISSN 1536-867X) is a publication of Stata Press. Stata, STATA, Stata Press, Mata, MATA, and NetCourse are registered trademarks of StataCorp LP.

The Stata Journal (2014)
14, Number 2, pp. 389–397

Power analyses for detecting effects for multiple coefficients in regression

Christopher L. Aberson
Department of Psychology
Humboldt State University
Arcata, CA
chris.aberson@humboldt.edu

Abstract. I present applications of Stata to conduct power analysis for multiple regression. Simple power analyses are provided through the Stata command `power` and the user-written command `powerreg`. However, tools for power analysis for multiple regression coefficients are limited. I present strategies for detecting significant regression coefficients in a single study (for example, power estimates for two or more coefficients). I also address the power for detecting multiple effects (that is, the probability that all coefficients are statistically significant) in the same sample through the Monte Carlo approaches afforded by the `simulate` command.

Keywords: st0342, powermr3, powersim3, powersim3_simulator, statistical power, corr2data, simulate, Monte Carlo

1 Introduction

Statistical power analysis became prominent with Jacob Cohen's (1988) seminal work on the topic. Since that time, extensive literature and several software packages focused on power (for example, PASS, nQuery, Sample Power, and G*Power) have emerged. In recognition of the important role of power, grant applications often require or encourage statistical power analysis as part of the design process (Maddock and Rossi 2001). Despite these advances, surveys across fields such as abnormal psychology (for example, Sedlmeier and Gigerenzer [1989]), management (Cashen and Geiger 2004; Mone, Mueller, and Mauland 1996), rehabilitation counseling (Kosciulek and Szymanski 1993), psychiatry (Brown and Hale 1992), adult education (West 1985), and consulting, clinical, and social psychology (Rossi 1990) all suggest that low power is common in published work. Focusing on multiple regression, I address barriers to accurate power analysis that likely contribute to the persistence of underpowered research and provide tools for conducting more accurate power analyses.

One primary issue for power analysis in multiple regression is the presence of several different types of hypotheses. The most common hypothesis tests focus on the squared multiple correlation (R^2) and the significance of regression coefficients. Power analyses for R^2, in terms of models and change, are handled well by the `powerreg` command (Ender and Chen 2011). Power analysis for regression coefficients is more complicated. Existing freeware sources (for example, PiFace and G*Power) either require complex input information (for example, a variance inflation factor) or only provide tests for a single coefficient's impact when added to an existing model. These are useful approaches for dealing with power for a single predictor. However, researchers often want to design studies that detect significance for multiple predictors in the same model. Using Stata's `corr2data` and `regress` commands with a handful of additional calculations, I provide the `powermr3` command for addressing power for multiple coefficients in the same sample.

Another source of low power for designs with multiple predictors is a lack of attention to power for detecting a set of outcomes. For example, a researcher conducting a multiple regression analysis with three predictors is often interested in detecting significant regression coefficients for all three predictors. Power analyses of studies with multiple predictors yield an estimate on power for each predictor but not on the power to detect all of them in the same study. I term the power to detect all effects in a study "Power(All)". Power to detect multiple effects differs from power for individual effects: in most research situations, the former is considerably lower than the latter. The lack of attention to Power(All) is a likely source of underpowered research in the behavioral sciences (Maxwell 2004). I detail how to address this form of power using Stata's `simulate` command. All examples use programs designed for 3 predictor variables; however, versions of the program are available for 2 through 10 predictors.

2 Power for individual effects

In multiple linear regression designs, there are several significance tests of interest. Two of the most prominent are tests of the fit of the overall model (R^2) and tests of the individual regression coefficients (often termed slopes or b). The power for detecting these effects is a function of sample size, type I error rate (also known as alpha), the strength of each predictor's correlation with the dependent variable (DV), and the correlation between predictors (multicollinearity).

The most common conceptualization of statistical power reflects the power to detect any specific effect in a study. For example, in a three-predictor multiple regression analysis, power might be discussed in terms of the power to detect a significant R^2 or the power to detect effects for the first, second, or third predictor. Many sources on power analysis (for example, Cohen [1988]; Faul et al. [2009]; Lenth [2006]; and Ender and Chen [2011]) provide estimation procedures for R^2 or a single coefficient. More recent work develops approaches for establishing power for multiple coefficients in SPSS (Aberson 2010). The first part of this article presents a general framework for conducting power analysis for multiple regression in Stata.

The basic process for power analysis detailed in this article provides estimates of correlations between all the variables involved in the proposed study. From this information, I generate a dataset of a specified sample size with these parameters. From those data, I calculate noncentrality parameters (λ) for each test. Using these values, in conjunction with the `nFtail()` function, provides the power statistic.

Power calculations require calculation of a noncentrality (λ). This value, along with the relevant comparison value on the central F distribution (that is, the critical value given degrees of freedom [d.f.] and type I error rate) and d.f., determines the likelihood that a population with the specified parameters produces a sample that allows for rejection of the null hypothesis. For tests of R^2, λ is a function of the proportion of explained variance and d.f. For the approaches detailed in this article, this value is equivalent to the F statistic produced as a significance test for R^2.

$$\lambda_{R^2} = \frac{R^2_{\text{Model}}}{1 - R^2_{\text{Model}}} \text{d.f.}_{\text{error}} \tag{1}$$

The noncentrality parameter for regression coefficients is based on the regression coefficient and its standard error. This is equivalent to the squared t statistic produced as a significance test for the regression coefficient.

$$\lambda_b = \left(\frac{b}{\text{se}}\right)^2 \tag{2}$$

Once these values are calculated, the Stata function `nFtail()` is used to calculate power. This command requires λ, the two relevant d.f. values, and the critical F value for rejecting the null hypothesis.

The example below uses the notation y for the DV and 1, 2, and 3 for a series of predictor variables. In this example, the researcher wants to determine a sample size that provides a power of 0.80, at minimum, for each of the three predictors.

Table 1. Correlations between DV (y) and predictors (1, 2, 3)

	y	1	2	3
1	0.30	1.00		
2	0.30	0.30		
3	0.30	0.40	0.30	

3 The powermr3 command

3.1 Description

`powermr3` calculates the power for the R^2 model and the regression coefficients for a range of sample sizes.

3.2 Syntax

`powermr3, ry1(#) ry2(#) ry3(#) r12(#) r13(#) r23(#) nstart(#) nend(#) by(#) [my(#) m1(#) m2(#) m3(#) sy(#) s1(#) s2(#) s3(#) alpha(#)]`

3.3 Options

`ry1(#)` specifies the correlation between the first predictor and the criterion. `ry1()` is required.

`ry2(#)` specifies the correlation between the second predictor and the criterion. `ry2()` is required.

`ry3(#)` specifies the correlation between the third predictor and the criterion. `ry3()` is required.

`r12(#)` specifies the correlation between the first predictor and the second predictor. `r12()` is required.

`r13(#)` specifies the correlation between the first predictor and the third predictor. `r13()` is required.

`r23(#)` specifies the correlation between the second predictor and the third predictor. `r23()` is required.

`nstart(#)` specifies the starting sample size. `nstart()` is required.

`nend(#)` specifies the ending sample size. `nend()` is required.

`by(#)` specifies the units between the first sample size and each subsequent sample size. `by()` is required.

`my(#)`, `m1(#)`, `m2(#)`, and `m3(#)` specify the user-defined means for the power calculation. The defaults are `my(0)`, `m1(0)`, `m2(0)`, and `m3(0)`.

`sy(#)`, `s1(#)`, `s2(#)`, and `s3(#)` specify the user-defined standard deviations for the power calculation. The defaults are `sy(1)`, `s1(1)`, `s2(1)`, and `s3(1)`.

`alpha(#)` specifies the user-defined type I error rate. The default is `alpha(.05)`.

3.4 Example

To run this code for the example above, the command line would look like this:

```
. powermr3, ry1(.3) ry2(.3) ry3(.30) r12(.3) r23(.3) r13(.4) nstart(200)
> nend(300) by(10)
```

This line does not provide values for means, standard deviations, or alpha, so they default to 0, 1, and 0.05, respectively.

The code produces the output below. Given our parameters, each of the coefficients (b1, b2, and b3) exceeds a power of 0.80 somewhere between 280 and 290. The second predictor produces more power than the others. The power for R^2 for the model is particularly strong regardless of sample size. For a finer-grained estimate, I ran the code again using nstart(280), nend(290), and by(1). This approach identified a sample size of 282 as the point where power exceeded 0.80 for all predictors.

```
Summary statistics: mean
  by categories of: n
```

n	power_b1	power_b2	power_b3	power_R2
200	.6517	.7982082	.6517	.999803
210	.6735311	.8176227	.6735312	.999891
220	.6942767	.8353996	.6942767	.9999402
230	.7139592	.8516429	.7139592	.9999674
240	.7326048	.8664554	.7326047	.9999824
250	.7502428	.8799374	.7502428	.9999905
260	.766905	.8921868	.766905	.9999949
270	.782625	.9032973	.782625	.9999973
280	.7974376	.9133585	.7974375	.9999986
290	.8113786	.9224555	.8113786	.9999992
300	.8244846	.9306687	.8244846	.9999996

4 Detecting multiple effects

The previous section focused on power for individual effects. Another way to conceptualize statistical power is in terms of the power to detect multiple effects within the same analysis. In the previous example, this would refer to the probability of detecting significant effects for b1, b2, and b3 all in the same sample. I term this "Power(All)".

Let's take a simple example. In a study designed with power_b1 = 0.50, power_b2 = 0.50, and power_b3 = 0.50, you might be tempted to think Power(All) is 0.50 as well. This is not the case. As a thought exercise, consider flipping a coin three times. The probability of the coin coming up heads is 0.50 on each flip. This is analogous to power_b1, power_b2, and power_b3 with each set at 0.50. The probability that the coin comes up heads on all three flips is far less than 0.50. In this case, the probability of obtaining heads on all three flips would be 0.13. This is analogous to Power(All) or how likely any one study is to reject all three null hypotheses. For models with uncorrelated predictors and equal predictor-DV correlations, calculation of Power(All) involves a simple application of the binomial expansion approach.

Most data addressed via regression analysis do not meet these criteria. First, predictor variables are often correlated. The size of these correlations impacts Power(All). Second, the power for each predictor is not always the same. The approaches detailed below use Stata's simulate command to simulate multiple regression analyses. The output provides empirical estimates of power.

5 The powersim3 command

5.1 Description

The `powersim3` command estimates Power(All) using a Monte Carlo approach and the example in the previous section. The major difference is that instead of creating a dataset with the specified parameters and basing power calculations on those data, we repeatedly draw random samples from a population with those parameters. Each sample is analyzed with multiple regression, and the results of these analyses are saved to a data file. A few additional calculations derive the power estimates. Again the example focuses on 3 predictors, but programs are available for 2 through 10 predictors.

5.2 Syntax

`powersim3, ry1(#) ry2(#) ry3(#) r12(#) r13(#) r23(#) n(#)` [`my(#) m1(#) m2(#) m3(#) sy(#) s1(#) s2(#) s3(#) alpha(#) reps(#)`]

5.3 Options

`ry1(#)` specifies the correlation between the first predictor and the criterion. `ry1()` is required.

`ry2(#)` specifies the correlation between the second predictor and the criterion. `ry2()` is required.

`ry3(#)` specifies the correlation between the third predictor and the criterion. `ry3()` is required.

`r12(#)` specifies the correlation between the first predictor and the second predictor. `r12()` is required.

`r13(#)` specifies the correlation between the first predictor and the third predictor. `r13()` is required.

`r23(#)` specifies the correlation between the second predictor and the third predictor. `r23()` is required.

`n(#)` specifies the sample size. `n()` is required.

`my(#)`, `m1(#)`, `m2(#)`, and `m3(#)` specify the user-defined means for the power calculation. The defaults are `my(0)`, `m1(0)`, `m2(0)`, and `m3(0)`.

`sy(#)`, `s1(#)`, `s2(#)`, and `s3(#)` specify the user-defined standard deviations for the power calculation. The defaults are `sy(1)`, `s1(1)`, `s2(1)`, and `s3(1)`.

`alpha(#)` specifies the user-defined type I error rate. The default is `alpha(.05)`.

`reps(#)` specifies the number of replications. The default is `reps(100)`. This value is useful for the initial analyses to produce a quick estimate of power (for example, is a sample of 300 close to what is needed?). For the final analyses, 10,000 or more replications are recommended.

5.4 Example

The output that follows is produced by the code

```
. powersim3, ry1(.3) ry2(.3) ry3(.30) r12(.3) r23(.3) r13(.4) n(282)
> reps(10000)
```

The user-provided values reflect the values from the example demonstrating the `powermr3` command. I recommend 10,000 replications to ensure there are enough values to address power.

The power for each effect closely matches the estimates in the first section (`power_b1` and `power_b3` round to 0.80, and `power_b2` rounds to 0.91) with a sample of 282. More interesting are the power estimates for detecting all effects (they reject all three null hypothesis) in the same sample. Based on simulation results, only about 55% of our samples would detect significant effects for all three effects.

The power for each coefficient generated by the simulation is as follows:

(*output omitted*)

stats	power_b1	power_b2	power_b3
mean	.8004	.913	.8001

The power for detecting effects for zero, one, two, or three (all) effects in the same sample is as follows:

Rejected	Freq.	Percent	Cum.
0	2	0.02	0.02
1	331	3.31	3.33
2	4,197	41.97	45.30
3	5,470	54.70	100.00
Total	10,000	100.00	

To find a sample size that yields Power(All) = 0.80, I ran the test with gradually increasing sample sizes until settling on 400. This represents a sample that is 40% larger than the initial sample.

6 Conclusions

Power analysis for multiple regression designs where the researcher is interested in detecting multiple effects for the same sample produces unique issues for power analysis. The approaches here provide tools for addressing power for multiple effects. Attention to Power(All) is particularly important for designing studies where the goal is to detect multiple effects.

7 Acknowledgments

The author thanks members of Statalist and Stack Overflow, particularly Maarten Buis, for comments that helped to develop more elegant and functional code. Chuck Huber of StataCorp also provided valuable assistance and updates to code.

8 References

Aberson, C. L. 2010. *Applied Power Analysis for the Behavioral Sciences*. New York: Routledge.

Brown, J., and M. S. Hale. 1992. The power of statistical studies in consultation-liaison psychiatry. *Psychosomatics* 33: 437–443.

Cashen, L. H., and S. W. Geiger. 2004. Statistical power and the testing of null hypotheses: A review of contemporary management research and recommendations for future studies. *Organizational Research Methods* 7: 151–167.

Cohen, J. 1988. *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum.

Ender, P. B., and X. Chen. 2011. powerreg: Power analysis for linear regression models. UCLA: Statistical Consulting Group.
http://www.ats.ucla.edu/stat/stata/ado/analysis/.

Faul, F., E. Erdfelder, A. Buchner, and A.-G. Lang. 2009. Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods* 41: 1149–1160.

Kosciulek, J. F., and E. M. Szymanski. 1993. Statistical power analysis of rehabilitation counseling research. *Rehabilitation Counseling Bulletin* 36: 212–219.

Lenth, R. V. 2006. Java applets for power and sample size.
http://homepage.stat.uiowa.edu/~rlenth/Power/.

Maddock, J. E., and J. S. Rossi. 2001. Statistical power of articles published in three health psychology-related journals. *Health Psychology* 20: 76–78.

Maxwell, S. E. 2004. The persistence of underpowered studies in psychological research: Causes, consequences, and remedies. *Psychological Methods* 9: 147–163.

Mone, M. A., G. C. Mueller, and W. Mauland. 1996. The perceptions and usage of statistical power in applied psychology and management research. *Personnel Psychology* 49: 103–120.

Rossi, J. S. 1990. Statistical power of psychological research: What have we gained in 20 years? *Journal of Consulting and Clinical Psychology* 58: 646–656.

Sedlmeier, P., and G. Gigerenzer. 1989. Do studies of statistical power have an effect on the power of studies? *Psychological Bulletin* 105: 309–316.

West, R. F. 1985. A power analytic investigation of research in adult education: 1970–1982. *Adult Education Quarterly* 35: 131–141.

About the author

Chris Aberson is currently a professor of psychology at Humboldt State University. He earned his PhD at the Claremont Graduate University in 1999. His research interests in social psychology include prejudice, racism, and attitudes toward affirmative action. He serves as associate editor for Group Processes and Intergroup Relations. His quantitative interests focus on statistical power. His book, *Applied Power Analysis for the Behavioral Sciences*, was published in 2010.