11-1-2011

JMASM31: MANOVA Procedure for Power Calculations (SPSS)

Alan Taylor

Macquarie University, alan.taylor@mq.edu.au

Recommended Citation

JMASM Algorithms and Code
JMASM31: MANOVA Procedure for Power Calculations (SPSS)

Alan Taylor
Macquarie University
Sydney, Australia

D'Amico, Neilands & Zambarano (2001) showed how the SPSS MANOVA procedure can be used to conduct power calculations for research designs. This article demonstrates a simple way of entering data required for power calculations into SPSS and provides examples that supplement those given by D'Amico, Neilands & Zambarano.

Key words: Power, sample size, MANOVA, SPSS.

Introduction

Most researchers acknowledge the importance of conducting power calculations prior to embarking on research projects to ensure that there is a good chance that effects regarded as theoretically or practically important will be determined statistically significant. Despite this acknowledgement, many studies are under-powered. Over the years, a number of writers have called on researchers to consider the power of their research designs and have attempted to facilitate power calculations. Cohen (1992), for example, followed up his book on power (1988) by providing a primer aimed at facilitating power calculation and sample size estimation in the behavioral sciences. Others have provided software for the same purpose; *Gpower*, for example, is a free program which allows the calculation of power and sample size in a variety of designs intended to "address the weaknesses of existing power analysis tools" (Erdfelder, Faul & Buchner, 1992, p. 1). In the same spirit, D'Amico, Neilands & Zambarano (2001), noted that the MANOVA procedure in SPSS (now called IBM SPSS Statistics, but referred to as

SPSS in the remainder of this article) can be used to calculate power in repeated measures and multivariate designs, which may not be readily accommodated by dedicated power programs. Their suggestion was particularly valuable because SPSS is a package that is accessible to many researchers.

D'Amico, Neilands & Zambarano's (2001) method takes advantage of the fact that the SPSS MANOVA procedure can both write and read datasets that are in a matrix form. This format allows a user to easily experiment with different numbers of cases, means, standard deviations and correlations between dependent and independent variables. In their examples, the authors read the data into SPSS in a matrix format and then read the matrix into MANOVA for the power calculation. This article suggests a minor extension to D'Amico, Neilands & Zambarano's procedure based on MANOVA's ability to write, as well as read, data in matrix format, a suggestion which facilitates power calculations with this method. Further examples of the use of the procedure to illustrate its usefulness in a number of design contexts are also provided; these examples supplement those given by D'Amico, Neilands & Zambarano (2001), which included ANCOVA, MANOVA and a repeated measures ANOVA.

A Summary of the Procedure

1. Set up a dummy dataset in SPSS that is similar to the one for which power estimates

Alan Taylor is a Senior Lecturer in the Department of Psychology. He teaches statistics and data analysis and the use of statistical software. Email him at: alan.taylor@mq.edu.au.

are to be obtained. The means and standard deviations can be drawn from previous research or may be guesswork. It may be easiest to use standardized measures so that effects (e.g., differences between group means) can be specified in terms of standard deviations.

2. Run the MANOVA procedure in order to save the data in matrix form.

3. Alter the values (the number of cases, differences between means, correlations) and run the MANOVA procedure to conduct the power calculations. Continue altering the relevant values (usually the number of cases) and re-running the MANOVA analysis to observe what values are necessary to obtain an acceptable level of power.

Detailed Examples

Example 1: Pre-Post Design with Two Groups

Consider a pre/post design with two groups, treatment and control. The dependent variable is anxiety, measured on a 10-point scale. The goal is to determine whether anxiety decreases more for the treatment group than for the control group. In other words, if there is a significant interaction between group and time. From previous research or from guesswork it is hypothesized that the mean score for each group (to which participants are randomly assigned from a waiting list of people who have come to an anxiety clinic) at pre-test will be about 6 and that the control group score will decline somewhat without treatment to around 5.5, but that the effect of treatment will be strong so the post-test mean for the treatment group will be approximately 4. The standard deviation of this anxiety measure is known to be about 1.5. With such changes – 2 versus 0.5 – a researcher would want to have good chance of finding the interaction significant at alpha 0.05. A dummy dataset, such as that shown in Figure 1, is first created in SPSS; the variable names were entered in the Variable View and the numbers were entered in the Data View.

Note that there are only two observations per case for this dummy dataset and that the required means are obtained by

Figure 1: Dummy Data Entered into an SPSS Dataset

having one case in each group one unit lower than the mean, and the other case one unit higher than the mean. The following MANOVA commands are now run:

    manova pre post by group(0,1)/
    wsfactor=time(2)/
    matrix=out(*)/
    design.

The MANOVA procedure can only be run using syntax. The matrix subcommand asks for the data to be saved in matrix format (see Figure 2).

The data are now in a form that allows the various values to be altered to simulate the data that might be obtained. In the present case, the goal is to determine if having 10 cases per group provides enough power, thus, the N in the top row is increased to 20 and the Ns for each group to 10. It is also necessary to change the standard deviation to 1.5 at each time point and to reduce the correlation between pre and post scores to a more realistic value, such as 0.5. The altered dataset is shown in Figure 3.

The following new set of MANOVA commands are run to obtain the power values:

    manova pre post by group(0,1)/
    wsfactor=time(2)/
    matrix=in(*)/
    power=f(.05) exact/
    design.

This syntax reads the matrix dataset and requests a power analysis with an alpha of 0.05. The relevant section of the MANOVA output (see Table 1), indicates that the power for the interaction is too low to be acceptable.

Figure 2: Example 1 Dummy Data in Figure 1 Shown in Matrix Format

	ROWTYPE_	group	VARNAME_	pre	post
1	N	.		4.0000000	4.0000000
2	MEAN	.00		6.0000000	5.5000000
3	N	.00		2.0000000	2.0000000
4	MEAN	1.00		6.0000000	4.0000000
5	N	1.00		2.0000000	2.0000000
6	STDDEV	.		1.4142136	1.4142136
7	CORR	.	pre	1.0000000	1.0000000
8	CORR	.	post	1.0000000	1.0000000

Figure 3: Matrix Form of the Example 1 Dummy Dataset with Altered Ns and Standard Deviations

	rowtype_	group	varname_	pre	post
1	N	.		20.0000000	20.0000000
2	MEAN	.00		6.0000000	5.5000000
3	N	.00		10.0000000	10.0000000
4	MEAN	1.00		6.0000000	4.0000000
5	N	1.00		10.0000000	10.0000000
6	STDDEV	.		1.5000000	1.5000000
7	CORR	.	PRE	1.0000000	.5000000
8	CORR	.	POST	.5000000	1.0000000

Table 1: SPSS MANOVA Output Showing the Results of the Power Calculation
for Example 1 Dummy Dataset

```
Observed Power at the .0500 Level

Source of Variation    Noncentrality    Power

TIME                       13.889        0.941
GROUP BY TIME               5.000        0.562
```

The number of subjects can be increased and the analysis repeated until an acceptable value is obtained. In this case, if the group size is doubled the power for detecting the interaction is 0.869, which is much more acceptable. In fact, 15 per group, which gives a power of 0.753, may be considered sufficient. Note that, in repeated measures analyses, changes in the correlation between measures may have a dramatic effect on power. For example, if the correlation between the pre- and post-test measures in this example was a still-realistic 0.7 as opposed to 0.5, the power with only 10 cases per group is 0.779 rather than 0.562.

If no information is available regarding the values to expect, the standard deviation can be set to 1 (as for a standard score), and differences between the means specified in terms of the standard deviation. Consider a simple example in which two groups are to be compared. With a standard deviation of 1, if one group had a mean of zero and the other a mean of 0.5 in the dummy dataset this would represent a moderate effect size in terms of Cohen's (1992) classification.

Another strategy that can be adopted when there is uncertainty about the magnitude of differences is to perform a series of analyses with various combinations of N and effect sizes (and correlations in a repeated measures design). If the power is reasonable over a range of approximate realistic combinations of values, then the research has a good chance of obtaining a significant result. If not, it may be considered that the research is not worth doing with the number of subjects available.

Example 2: Oneway ANOVA with Contrasts and Unknown Means and Standard Deviation

This example has three groups of subjects, a control group (group = 1) and two treatment groups (2 and 3). In this case expected means and standard deviations are unknown, so the standard deviation is set equal to one, the mean of the control group to zero and the means of groups 2 and 3 to 0.5 and 0.8 respectively. The difference between the means of groups 1 and 2 is therefore $(0.5 - 0) = 0.5$, and Cohen's d $= 0.5/1 = 0.5$, which Cohen (1992) terms a medium effect size. The difference between the

control group and group 3 is 0.8, and d = 0.8, a large effect size.

Again, a dummy dataset with the necessary structure is entered into SPSS using, in this case, arbitrary numbers for the dependent variable (see Figure 4).

Figure 4: Dummy Dataset for Example 2

The MANOVA commands to create the matrix version of the dataset (see Figure 5) are:

```
manova score by group(1,3)/
matrix=out(*)/
design.
```

The standard deviation of 0.7071 is replaced with 1 and the means are given the values noted above. It is assumed that 60 subjects can be recruited, 20 in each group. The revised matrix dataset is shown in Figure 6.

It is assumed that there is no interest in the overall ANOVA result, but rather in two *a priori* contrasts: group 2 versus group 1, and group 3 versus group 1. MANOVA can give the power for each of the contrasts with the commands:

```
manova score by group(1,3)/
contrast(group)=simple(1)/
matrix=in(*)/
power=f(.025) exact/
design=group(1) group(2).
```

The simple (1) option asks for the required contrasts and, in the design statement, group (1) represents the group 2 versus group 1 contrast and group (2) represents the group 3 versus

group 1 contrast. To hold the overall Type I error at 0.05, alpha is set at 0.025 for the two contrasts. The relevant output, (see Table 2) indicates that the power for the first contrast is very low, 0.24, but that for the second is 0.59. At this stage more cases could be added to determine how many more cases would be needed to achieve an acceptable level of power.

Figure 5: Matrix Form of the Data for Example 2

	rowtype_	group	varname_	score
1	N	.		6.0000000
2	MEAN	1.00		1.5000000
3	N	1.00		2.0000000
4	MEAN	2.00		3.5000000
5	N	2.00		2.0000000
6	MEAN	3.00		5.5000000
7	N	3.00		2.0000000
8	STDDEV	.		.7071000
9	CORR	.	SCORE	1.0000000
10				

Figure 6: Matrix Format of the Dataset for Example 2 with Amended Values of *N*, Means and Standard Deviations

	rowtype_	group	varname_	score
1	N	.		60.0000000
2	MEAN	1.00		0E-7
3	N	1.00		20.0000000
4	MEAN	2.00		.5000000
5	N	2.00		20.0000000
6	MEAN	3.00		.8000000
7	N	3.00		20.0000000
8	STDDEV	.		1.0000000
9	CORR	.	SCORE	1.0000000
10				

Table 3: MANOVA output showing the results of the power calculation for Example 2

```
         Observed Power at the  .0250 Level
                              Noncen-
      Source of Variation     trality    Power

      GROUP(1)                  2.500     .244
      GROUP(2)                  6.400     .592
```

Example 3: Correlation

Assume that a researcher seeks to assess the correlation between measures of anger and narcissism, which is expected to be very low. It is desirable to have a good chance (power at least 0.80) of obtaining a significant result if the correlation in the population is 0.30 or higher. The dummy dataset (see Figure 7) is created and then the following commands are used to produce a matrix version of the data (see Figure 8):

    manova anger with narciss/
    matrix=out(*)/
    design.

The correlation is replaced with 0.30, and 30 subjects are used at first (the values of the means and standard deviations are immaterial in this case). The revised version of the dataset is shown in Figure 9 and relevant output is shown in Table 4. The value of 0.362 is unacceptably low; thus, further experimentation was conducted to show that 85 subjects are needed to achieve a power of 0.80.

Figure 7: Dummy Dataset for Example 3

	anger	narciss
1	1.00	2.00
2	2.00	1.00
3	3.00	4.00
4	4.00	3.00
5	5.00	5.00
6		

Figure 8: Matrix Form of the Data for Example 3

	rowtype_	varname_	anger	narciss
1	N		5.0000000	5.0000000
2	MEAN		3.0000000	3.0000000
3	STDDEV		1.5811388	1.5811388
4	CORR	ANGER	1.0000000	.8000000
5	CORR	NARCISS	.8000000	.8000000
6				

Figure 9: Amended Matrix Dataset for Example 3

	rowtype_	varname_	anger	narciss
1	N		30.0000000	30.0000000
2	MEAN		3.0000000	3.0000000
3	STDDEV		1.5811388	1.5811388
4	CORR	ANGER	1.0000000	.3000000
5	CORR	NARCISS	.3000000	1.0000000

Table 4: MANOVA output showing the results of the power calculation for Example 3

```
            Observed Power at the  .0500 Level
                                    Noncen-
            Source of Variation     trality    Power

            Regression               2.769     .362
```

Example 4: Using an Existing Dataset

It is often not sensible to calculate the power for an existing dataset (if the effects are significant, the power will be viewed as adequate; if the effects are not significant, the power may be considered too low). However, it can be sensible to ask, for an effect which was not significant: How many more cases would be needed to have a good chance of finding a significant effect if the population characteristics are the same as those of my sample?

This example uses a partly synthetic dataset called *glmdemo.sav* (available for download from: http://www.psy.mq.edu.au/ psystat/download.htm). Suppose a multivariate analysis with three variables, *test1* to *test3*, as the dependent variables and *group* (with four categories) as the grouping variable has been conducted. The results for the 99 cases in the dataset are shown in Table 5.

The MANOVA commands to produce the matrix version of the data are:

```
manova test1 to test3 by group(1,4)/
matrix=out(*)
```

The matrix version of the dataset is shown in Figure 10.

The following commands can be used to calculate the observed power:

```
manova test1 to test3 by group(1,4)/
matrix=in(*)/
power=f(.05) exact.
```

The power for the Wilks' Lambda statistic is 0.62; now, determine what improvement would result if five cases were added to each group for an addition of 20 more subjects overall. The subject numbers could be added manually, but the following commands will complete the task and will make it easier to add more in the future:

```
do if (rowtype_ eq "N" and
sysmis(group)).
compute test1=test1 + 20.
compute test2=test2 + 20.
compute test3=test3 + 20.
else if (rowtype_ eq "N" and
~sysmis(group)).
compute test1=test1 + 5.
compute test2=test2 + 5.
compute test3=test3 + 5.
end if.
execute.
```

When the MANOVA commands are run again, the resulting power is 0.72. If the above syntax is used with the addition of another five subjects to each group – 139 subjects overall – the power is found to be exactly 0.80.

Table 5: MANOVA results of a multivariate analysis of the *glmdemo* dataset used in Example 4

```
EFFECT .. GROUP
 Multivariate Tests of Significance (S = 3, M = -1/2, N = 45 1/2)


 Test Name          Value     Approx. F  Hypoth. DF   Error DF    Sig. of F

 Pillais            .14202     1.57358      9.00       285.00       .123
 Hotellings         .15739     1.60308      9.00       275.00       .114
 Wilks              .86122     1.59333      9.00       226.49       .118
 Roys               .11361


 - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
 EFFECT .. GROUP (Cont.)
 Univariate F-tests with (3,95) D. F.


 Variable  Hypoth. SS  Error SS  Hypoth. MS   Error MS    F       Sig. of F

 TEST1       2.04393   59.70197    .68131      .62844    1.08412    .360
 TEST2       5.62424   57.69468   1.87475      .60731    3.08695    .031
 TEST3       2.08582   69.51679    .69527      .73176     .95014    .420
```

Example 5: Multiple Regression

This analysis regresses a dependent variable *y* on *x1* and *x2* (numeric variables), and *x3* (dichotomous variable). A dummy dataset is created as usual using arbitrary numbers, but with one exception: with the dichotomous variable, *x3*, it is a good idea to insert zeroes and ones (e.g., female = 0, male = 1) in the same proportion as they would be expected to occur in the sample. Because the mean and standard deviation of a proportion are linked, this will help avoid the need to change one or both after creating the matrix dataset. In this example, a 50/50 distribution is assumed. The initial dataset is shown in Figure 11.

The MANOVA commands used to produce the matrix version of the data are:

```
manova y with x1 x2 x3/
matrix=out(*).
```

Figure 12 shows the initial matrix version of the dataset.

Assume, based on past research or theory, that *x1* is moderately (0.5) correlated with *y*, but *x2* and *x3* are only weakly correlated with *y* (both 0.3). Furthermore assume that *x1* and *x2* are highly correlated (0.6), but neither is correlated with *x3* (0.1) and that 50 subjects can be obtained for the research. The amended dataset is shown in Figure 13. The MANOVA commands below provide the output shown in Table 6:

```
manova y with x1 x2 x3/
matrix=in(*)/
power=f(.05).
```

There is ample power (0.97) for the overall regression, but the power for each independent variable is also of interest. The results show the power for an alpha of 0.05; a different alpha, e.g., 0.05/3 = 0.0167, may be used for the predictors. The MANOVA command is run again, with power = f(0.0167), to obtain the results shown in Table 7 (only part of the output is shown).

The power for *x1* remains adequate, but a bigger sample would be needed to achieve acceptable values for *x2* and *x3*. A sample of

100 gives a value of 0.70 for *x3*; *x2* is a lost cause due to its correlation with *x1*.

Conclusion

As D'Amico, Neilands, & Zambarano (2001) noted, the SPSS MANOVA procedure provides a way of conducting power and sample size calculations for multivariate and repeated measures designs that may be impossible or difficult with dedicated power and sample size software. This article illustrated a simple method of creating the matrix dataset, which is at the core of procedure described by D'Amico, Neilands, & Zambarano, and provided additional examples of the method to supplement their work.

Comprehensive power and sample size software, such as NCSS PASS (Hintz, 2012), is available and programs such as *Mplus* (Muthén & Muthén, 1998-2010) offer sophisticated simulation facilities for models that are outside the scope of those that can be handled by the methods described herein. The variety of designs for which power can be calculated using the SPSS MANOVA procedure, together with the ubiquity of the package, make it a valuable contribution to facilitating the routine calculation of power during research design.

References

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*, 2nd *Ed.* Hillsdale, N.J.: L. Erlbaum Associates.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112, 155-159.

D'Amico, E., Neilands, T., & Zambarano, R. (2001). Power designs for multivariate and repeated measures designs: A flexible approach using the SPSS MANOVA procedure. *Behavior Research Methods, Instruments, & Computers*, *33*, 479-484.

Erdfelder, E., Faul, F., & Buchner, A. (1996). GPOWER: A general power analysis program. *Behavior Research Methods, Instruments and Computers*, *28*, 1-11.

Hintz, J. (2011). PASS 11 (computer software). NCSS: Kayesville, Utah.

Muthén, L. K., & Muthén, B. O. (1998-2010). *Mplus user's guide*, *6th Ed*. Los Angeles, CA: Muthén & Muthén.

Figure 10: Matrix Version of the *glmdemo* Dataset Used in Example 4

	ROWTYPE_	group	VARNAME_	test1	test2	test3
1	N	.		99.0000000	99.0000000	99.0000000
2	MEAN	1		4.4244974	4.5246307	4.2067255
3	N	1		29.0000000	29.0000000	29.0000000
4	MEAN	2		4.6060099	4.8869344	4.2621851
5	N	2		24.0000000	24.0000000	24.0000000
6	MEAN	3		4.7934063	5.1225354	4.2714072
7	N	3		31.0000000	31.0000000	31.0000000
8	MEAN	4		4.6287784	4.9799732	4.6443188
9	N	4		15.0000000	15.0000000	15.0000000
10	STDDEV	.		.7927432	.7793025	.8554272
11	CORR	.	test1	1.0000000	.0415152	.1233156
12	CORR	.	test2	.0415152	1.0000000	.2428773
13	CORR	.	test3	.1233156	.2428773	1.0000000

Figure 11: Dummy Dataset Used in Example 5

	y	x1	x2	x3
1	1.00	2.00	4.00	.00
2	2.00	1.00	1.00	.00
3	3.00	4.00	2.00	.00
4	4.00	3.00	5.00	1.00
5	5.00	6.00	3.00	1.00
6	6.00	5.00	6.00	1.00

Figure 12: Initial Matrix Form of the Dataset Used in Example 5

	rowtype_	varname_	y	x1	x2	x3
1	N		6.0000000	6.0000000	6.0000000	6.0000000
2	MEAN		3.5000000	3.5000000	3.5000000	.5000000
3	STDDEV		1.8708287	1.8708287	1.8708287	.5477226
4	CORR	Y	1.0000000	.8285714	.5428571	.8783101
5	CORR	X1	.8285714	1.0000000	.3714286	.6831301
6	CORR	X2	.5428571	.3714286	1.0000000	.6831301
7	CORR	X3	.8783101	.6831301	.6831301	1.0000000

Figure 13: Amended Version of the Matrix Dataset used in Example 5

	rowtype_	varname_	y	x1	x2	x3
1	N		50.0000000	50.0000000	50.0000000	50.0000000
2	MEAN		3.5000000	3.5000000	3.5000000	.5000000
3	STDDEV		1.8708287	1.8708287	1.8708287	.5477226
4	CORR	Y	1.0000000	.5000000	.3000000	.3000000
5	CORR	X1	.5000000	1.0000000	.6000000	.1000000
6	CORR	X2	.3000000	.6000000	1.0000000	.1000000
7	CORR	X3	.3000000	.1000000	.1000000	1.0000000

Table 6: MANOVA output for Example 5

```
Observed Power at the .0500 Level
                            Noncen-
Source of Variation         trality    Power

Regression                  20.986     .970


- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
Regression analysis for WITHIN CELLS error term
--- Individual Univariate .9500 confidence intervals
--- two-tailed observed power taken at .0500 level
Dependent variable .. Y

COVARIATE              B        Beta    Std. Err.    t-Value    Sig. of t

X1               .48418     .00000       .153        3.166        .003
X2              -.01582    -.11076       .153        -.103        .918
X3               .86472    6.05305       .420        2.059        .045

COVARIATE   Lower -95%  CL- Upper     Noncent.      Power

X1               .176       .792      10.025        .873
X2              -.324       .292        .011        .051
X3               .019      1.710       4.240        .522
```

Table 7: Output Obtained for the Amended Matrix Dataset Used in Example 5

```
two-tailed observed power taken at .0167 level

COVARIATE    Lower -95%  CL- Upper      Noncent.        Power
X1                .176       .792       10.025          .750
X2               -.324       .292         .011          .017
X3                .019      1.710        4.240          .345
```