Paper S1-06 2013

Power Trip: A Road Map of PROC POWER

Melissa Plets, Grand Valley State University, Allendale, MI
Julie Strominger, Grand Valley State University, Allendale, MI

ABSTRACT

Sample size and power analyses are extremely important components to consider when designing, planning and recruiting for prospective clinical research projects. Unfortunately, these are often considered to be scary and complicated calculations for research clinicians. PROC POWER offers a systematic solution to finding the balance between efficiency and conclusive results, while remaining simple enough for non-statisticians to use.

More power is universally considered to be advantageous, even outside of the domain of statistics. It is especially important in statistics, as more power results in a higher probability that the null hypothesis will be rejected when it is false. Additionally, sample size is directly related to power. In general, a larger sample size will result in more accuracy, precision, and higher power. An important aspect of study design is maximizing power while remaining within the bounds of financial feasibility. This is where PROC POWER comes in, providing methods of determining sample sizes and power calculations. Throughout this paper, we will construct a road map of the POWER procedure to assist both statisticians and non-statisticians with the implementation of this POWERful SAS tool.

INTRODUCTION

Research is what drives the medical world to the development of new methods, treatments and cures. Clinical studies are an essential part of the research process and are constantly utilized in the field of healthcare. When designing a clinical research study, there are four main components to consider:

- The *significance level*, denoted alpha, is the probability of rejecting the null hypothesis when it is true. In other words, it is the probability that significant results will be found by chance, not due to a true significant difference. This is also referred to as the Type I error rate. The most common choice for the significance level is $\alpha = 0.05$.

- *Power* is derived by calculating (1- β), where β represents the Type II error rate. Type II error is defined as the probability of not rejecting the null hypothesis when it is false. Power is therefore the probability of rejecting the null hypothesis when it is false. Put simply, it is the probability of finding significance when there is a significant difference to be found. The most common choice is power = 0.80.

- *Effect size* tells us the minimum number of standard deviations of difference between the interest groups which can be detected by the test. It can be precisely calculated by dividing the difference in means (between 2 groups of interest) by their pooled standard deviation. However, if these values are unknown, as a general rule of thumb an effect size of 0.20 is considered a “small” effect, 0.50 is seen as a “medium” effect and 0.80 is thought of as a “large” effect. This component is inversely related to sample size, where a larger sample size is required to detect a smaller effect.

- *Sample size* is the total number of subjects that are to participate in the study. Additionally, if the study aims to compare two groups, a sample size calculation should also include the number of subjects in each group. It is important to conduct a minimum sample size calculation prior to the start of a study to avoid unnecessary expenses and wasted resources. Sample size is the key to the ideal balance between efficiency and conclusive results.

Not only are these four components correlated with one another, but any three of these elements can be used to calculate the fourth. The goals of any great clinical research study are to minimize alpha and effect size, and maximize power while remaining within the bounds of a feasible minimal sample size. Instead of using lengthy, complicated mathematic formulas to calculate one of these components, the POWER procedure in SAS 9.3 provides a much more user-friendly solution.

STEP 1: IDENTIFY TYPE OF ANALYSIS

When approaching a research study, knowing the aims and goals of the study are of utmost importance. Understanding the purpose of the study is imperative to accurately analyze the data which, in turn, is necessary to know prior to determining the power or sample size for a study. Questions will need to be asked of the principal investigator/research team if the overall goals of the study are unclear. It is absolutely vital to understand what kind of

statistical tests will be performed post-collection in order to determine power or sample size. Are there multiple groups? Are all groups being compared to one another or is there one control and a treatment group/multiple treatment groups? Will there be multiple comparisons?

In conclusion, what the research question is, what type of data will be collected (quantitative, qualitative, both), and the calculation the researcher wants (e.g. sample size, power, sensitivity) need to be known prior to beginning any kind of coding/analysis. After that is complete, a statistical plan should be created that outlines in detail which statistical tests will be performed following data collection. Using the statistical plan, sample sizes or power calculations - depending on what the researcher wants to know - should be calculated for all statistical tests that will be performed. The sample size should be chosen based on the largest sample size given and power for each statistical test will be given.

The test(s) decided upon will then be entered into PROC POWER to determine the calculation the researcher wants (sample size, power, effect size). This can be done for a variety of statistical analyses which are listed in the first column of Table 1. The line of code following PROC POWER specifies which test will be used to determine the desired calculation. Options for this are listed in the second column of Table 1.

Statistical Analysis	PROC POWER Code
T-Tests	ONESAMPLEMEANS,PAIREDMEANS,TWOSAMPLEMEANS
Confidence Intervals for Means	ONESAMPLEMEANS,PAIREDMEANS,TWOSAMPLEMEANS
Exact and Approximate Tests	ONESAMPLEFREQ,PAIREDFREQ,TWOSAMPLEFREQ
Tests of proportions	ONESAMPLEFREQ,PAIREDFREQ,TWOSAMPLEFREQ
ANOVA (one-way)	ONEWAYANOVA
Wilcoxon Mann-Whitney Test	TWOSAMPLEWILCOXON
Correlation	ONECORR,TWOCORR
Logistic Regression	LOGISTIC
Multiple Linear Regression	MULTREG
Survival Curve Comparison	TWOSAMPLESURVIVAL

Table 1. Statistical Analyses and Corresponding PROC POWER Code

STEP 2: SET KNOWN COMPONENTS

After determining the statistical analysis to be performed and corresponding code, the next step is to enter all of the known component values and set the unknown component to missing, denoted by a period. The component which is set as “missing” is the one SAS will solve for.

```
POWER=.;
```

One unique attribute that PROC POWER provides is the option to enter multiple values for a single component. This is one of the advantages PROC POWER has over most of the other user-friendly, “point and click” power calculation software programs, such as GPower 3.1.5. This can be especially advantageous for non-statisticians who may need a better understanding of the relationships among the four main components. The example below illustrates how power changes as the sample size is increased.

```
POWER=.;
NPERGROUP= 2 5 10 20
```

Computed Power		
Index	N Per Group	Power
1	2	0.129
2	5	0.458
3	10	0.805
4	20	0.984

Figure 2. Computed Power for Different Group Sizes

We have found throughout our experience in a hospital affiliated research department that 2-sample t-tests are frequently used in medical research. Therefore, to illustrate the use of plotting in PROC POWER we will work through a 2-sample t-test example for the remainder of this paper.

STEP 3: ADDITIONAL OPTIONS (PLOTTING)

The plotting options available in PROC POWER can be tremendously helpful for both statisticians and non-statisticians when attempting to visualize relationships among components. This can come in handy when feasibility of a study may only allow for a certain number of subjects. The example below demonstrates how PROC POWER can be used to visualize how power changes as sample size per group increases for a 2-sample t-test.

A graph is produced by using the PLOT option, where one of the aforementioned components is assigned to either the X or Y axis. This can be specified by using X= or Y= in the PLOT statement. Furthermore, the YOPTS= option coupled with the REF= suboption creates horizontal reference lines. Here, we see reference lines highlighting power=0.8, power =0.9 and power =0.95.

```
PROC POWER;
     TWOSAMPLEMEANS TEST=DIFF
     NULLDIFF= 0
     MEANDIFF= 4
     STDDEV= 3
     POWER= 0.8 0.9 0.95
     ALPHA = 0.05
     NPERGROUP = .;
     PLOT Y=POWER YOPTS=(REF=0.8 0.9 0.95);
RUN;
```

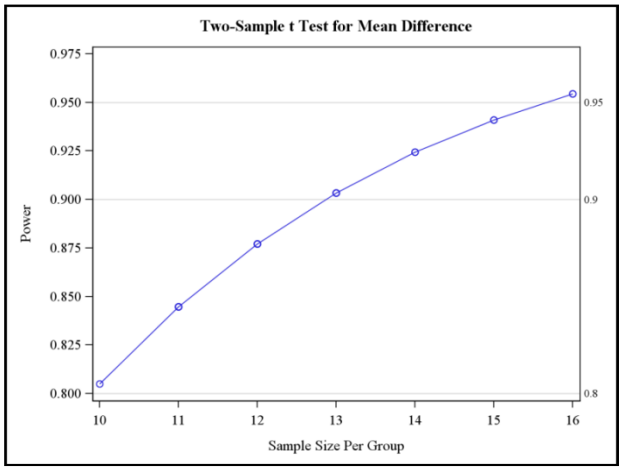

Figure 2. Computed Sample Size for Two-Sample T-Test Mean Difference

Additionally, these plotting options can be used to illustrate relationships among multiple values of multiple components.

The X= EFFECT option in the PLOT statement puts the mean difference values on the x-axis. The VARY option together with the COLOR BY POWER suboption will create a separate line for each specified power value and label each line a different color. In this example we are able to visualize how power and sample size change as the magnitude of the mean difference between the two groups changes.

```
PROC POWER PLOTONLY;
       TWOSAMPLEMEANS TEST=DIFF
       NULLDIFF= 0
       MEANDIFF= 2 3 4 5
       STDDEV= 3
       POWER= 0.8 0.9 0.95
       ALPHA = 0.05
       NPERGROUP = .;
       PLOT X=EFFECT  VARY(COLOR BY POWER);
RUN;
```

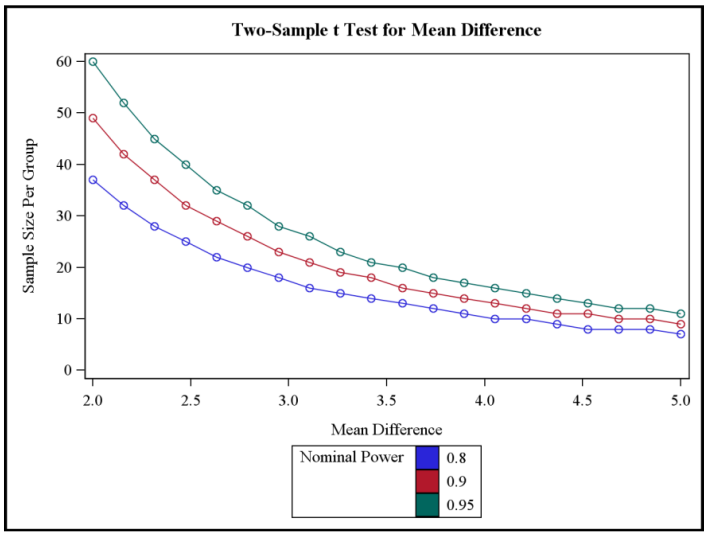

Figure 3. Computed Sample Sizes for Different Power Levels – Two Sample T-Test Mean Difference

Other labeling suboptions within the VARY option in the PLOT statement include labeling by SYMBOL or LINESTYLE. Also, power is not the only component which can be used within the VARY option. Any component can be used here as long as it is numerically defined earlier in the PROC POWER procedure.
To output only the graph, specify PLOTONLY in the PROC POWER statement.

CONCLUSION

In a research setting being able to determine the power or sample size needed for an analysis is priceless. This paper, combined with user-friendly PROC POWER, allows even the most novice researcher to compute sample sizes and determine power for different analyses.

ACKNOWLEDGMENTS

We would like to acknowledge Dr. Robert Downer, Dr. Gerald Shoultz, and Erica Goodrich for their help and guidance.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the authors at:

Name: Melissa Plets
Enterprise: Grand Valley State University
E-mail: pletsm@mail.gvsu.edu

Name: Julie Strominger
Enterprise: Grand Valley State University
E-mail: strominj@mail.gvsu.edu