SESUG Paper 2023-year

How to use SAS POWER procedure to plan a study

Xiaofeng “Steven” Liu University of South Carolina

ABSTRACT

SAS PROC POWER procedure can be used to choose the required sample size for hypothesis testing and confidence interval, the latter of which is less known to practitioners. The purpose of the paper is to review PROC POWER procedure and introduce sample size determination for building a precise confidence interval that is characterized by a sufficiently narrow interval width. Practitioners can use PROC POWER procedure to achieve adequate statistical power and/or high probability of obtaining a precise interval estimate while planning a scientific study. Examples are provided to illustrate sample size planning for interval precision.

INTRODUCTION

Sample size can be chosen to achieve a desired width in a confidence interval. The probability of obtaining a narrow width given the confidence interval including the population parameter is defined as the power of the confidence interval, the concept of which is not familiar to many practitioners. This paper shows how to use the SAS PROC POWER procedure to select an appropriate sample size and achieve sufficient power for a confidence interval. Examples are provided to illustrate sample size choice for confidence intervals in an independent *t* test.

A confidence interval can be constructed by extending a certain number of standard errors around a point estimate. The point estimate of a population parameter is the middle point of the confidence interval. The margin of error on either side of the point estimate provides positive and negative allowances. The margin of error is half the width of a confidence interval, and it is indicative of the precision of the confidence interval.

POWER OF CONFIDENCE INTERVAL

Beal (1989) provided the seminal idea for determining sample sizes for confidence intervals. He defined the probability of obtaining a narrow width, conditional on the confidence interval including the population parameter. This probability is denoted by $P[w \leq U|I]$, where the half width of the confidence interval is *w*; the upper bound of the half width is *U*. The inequality $w \leq U$ suggests that the halfwidth is less than or equal to a specified upper bound U. The event *I* indicates that the confidence interval contains the population parameter (e.g., the difference between two population means). The probability of the confidence interval including the population parameter is equivalent to the confidence level or $P[I] = 1 - \alpha$. The significance level α is traditionally set at .05. The probability $P[w \leq U|I]$ is a conditional probability, and the conditional probability is related to its unconditional counterpart, that is, the probability of obtaining a narrow width of a confidence interval $P[w \leq U]$. The relationship between the conditional and unconditional probabilities can be shown in the following

$$P[w \leq U] = P[w \leq U|I]P[I] + P[w \leq U|\bar{I}]P[\bar{I}].$$

The symbol ($\bar{I}$) means an opposite event of *I*, that is, the confidence interval does not contain the population parameter. The probability $P[\bar{I}]$ therefore means the likelihood that the confidence interval fails to include the parameter value. The conditional probability of achieving a certain width conditional on the confidence interval not including the population parameter is $P[w \leq U|\bar{I}]$. The conditional probability $P[w \leq U|I]$ is of more interest than the conditional probability $P[w \leq U|\bar{I}]$. A narrow confidence interval is only desired when it includes the population parameter. The conditional probability $P[w \leq U|I]$ is defined as the power of the confidence interval, and it is the default output in the SAS PROC POWER procedure

(SAS, 2004). Although there are some differences in value between the conditional probability $P[w \leq U|I]$ and the unconditional probability $P[w \leq U]$, the difference is usually very small. The two probabilities are mathematically about the same, but they represent different philosophies. Both Beal's paper and the SAS procedure seem to reflect the same view. Therefore, we will try to compute the conditional probability $P[w \leq U|I]$ in the example, using SAS.

EXAMPLE

Sample size can be chosen to obtain a narrow confidence interval on the two-mean difference in an independent *t* test, which is often used to compare the outcome performance between a treatment group and a control group. A sufficient sample size can also be determined to obtain a narrow confidence interval on the two-mean difference in an independent *t* test, which is often used to compare the outcome performance between a treatment group and a control group. If there is any performance differential between the treatment and control groups, one can make a causal inference about the treatment effect. A confidence interval of the mean difference can be computed to measure the magnitude of the treatment effect. For planning purposes, we assume a balanced design that has equal numbers of participants in the treatment and control conditions (i.e., $n_1 = n_2 = n$). The margin of error or the half width of the confidence interval is w.

The following SAS code produces the required sample sizes to achieve .80 in the probability of having the half width narrower than the upper bound U, given that the confidence interval includes the true mean difference.

```
proc power;
   twosamplemeans ci=diff
      halfwidth = .1 to 1.0 by .1
      stddev =1
      probwidth =.80
      ntotal = .;
      plot x=effect min=.1 max=1.0;
run;
```

For simplicity of illustration, the standard deviation σ is set to unity, and the upper bound U is expressed as a multiple of the standard deviation, which ranges from .1 to 1.0 with an increment of .1 in the argument "halfwidth". The keyword "twosamplemeans ci=diff" indicates that the required sample size is for the confidence interval, based on the independent *t* test.

Table 1 shows the required total sample size for the power of confidence interval above .80, based on the desired precision. For instance, we can modify the SAS code to obtain the required total sample size 74 specifically for the desired half width of .5 standard deviation . This is tantamount to setting "halfwidth =" to .5 and "stddev =" to 1 in the SAS code:

```
proc power;
   twosamplemeans ci=diff
      halfwidth =.5
      stddev =1
      probwidth =.80
      ntotal = .;
run;
```

$w = U/\sigma$	Power of confidence interval	Total sample size $N=2n$
0.2	0.808	410
0.3	0.829	190
0.4	0.81	110
0.5	0.834	74
0.6	0.85	54
0.7	0.804	40
0.8	0.811	32

Table 1. Power of confidence interval and required sample size

CONCLUSION

There has been a movement to shift the emphasis from null hypothesis testing to confidence interval estimation. The critics of hypothesis testing argue that *p*-value does not suggest the size of the treatment effect and that laymen confuse statistical significance with practical importance. Some journal editorials (Wilkinson, 1999) have directly called for the use of confidence interval in lieu of hypothesis test. Although misconceptions about hypothesis testing are common, it will remain the main vehicle of research inquiry in social sciences. Nevertheless, a confidence interval accompanying hypothesis test shall be encouraged in research practice. The wide use of confidence intervals will not automatically improve statistical practice unless people pay attention to sample size and the precision of confidence interval. A wide confidence interval is not informative about the true size of the parameter because such a confidence interval shows a long range of plausible values the parameter can take on (Liu, 2013).

In conclusion, sample size planning is very important to both null hypothesis testing and confidence interval estimation. A sufficient sample size is necessary to assure a reasonable chance of rejecting the null hypothesis and confirming the research hypothesis in a significance test. Following a statistically significant result, a confidence interval can be constructed to measure the magnitude of the treatment effect. Sample size has great bearing on the width of the confidence interval as well. An inadequate sample size will likely return a large interval width, which is not informative about the size of the treatment effect. As high statistical power and estimate precision are essential to good statistical practice, sample size determination should become an integral part of any study planning. The power of confidence interval offers a simple and coherent way to determine sample size for statistical power and estimate precision in a randomized experiment, the goal of which is to confirm the existence of the treatment effect in a hypothesis test and then to measure the size of the treatment effect in a confidence interval. The primary goal of the randomized experiment can be linked to statistical power, and the secondary goal can be tied to the probability of obtaining a short interval. SAS PROC POWER procedure allows practitioners to calculate both statistical power and power of confidence interval, the latter of which is readily available with the right setup of a few statements.

REFERENCES

Beal, S. L. (1989). Sample size determination for confidence intervals on the population mean and on the difference between two population means. *Biometrics*, *45*, 969-977.

Liu, X. S. (2013). *Statistical power analysis for the social and behavioral sciences: Basic and advanced techniques*. New York, NY: Routledge.

SAS. (2004). SAS/STAT User guide, version 9.1. 2. Cary, NC: SAS Institute.

Wilkinson, L. (1999). Statistical methods in psychology journals: Guidelines and explanations. *American psychologist*, *54*(8), 594.

CONTACT INFORMATION

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

Xiaofeng “Steven” Liu
Department of Education and Development Sciences
University of South Carolina, Columbia
803-777-6084
E-mail: xliu@email.sc.edu