Optimal Sampling through Power Analysis

An Automated Approach

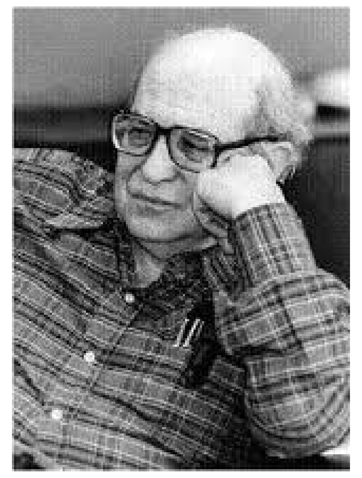
Micah A. Thornton
Southern Methodist University

Contact Information:
Department of CSE
Southern Methodist University
6425 Boaz Lane Dallas TX 75205

Phone: +1 (214) 564 9637
Email: mathornton@smu.edu

SMU. Darwin Deason Institute for Cyber Security
LYLE SCHOOL OF ENGINEERING

Introduction



Since the inception of statistical meta-analysis, power analysis has taken to the stage as a means of retrieving optimal sample sizes for specified statistical analysis. To this ends it uses a set of three interdependent values to solve for a fourth unknown, the four values are known as: Sample Size, Effect Size, Significance Level, and Power Level. The methodology proposed in this poster takes advantage of power analysis when selecting the optimal amount of data needed to perform specific statistical analysis, in an automated fashion.

Background

Recall that there are four possible outcomes of any particular hypothesis test:

Table 1: Test Outcomes

Decision	null is true	null is false
Reject null	Type 1 Error	Correct Decision
Accept null	Correct Decision	Type 2 Error

Also recall the notation for the two different types of error as well as power from the table above.

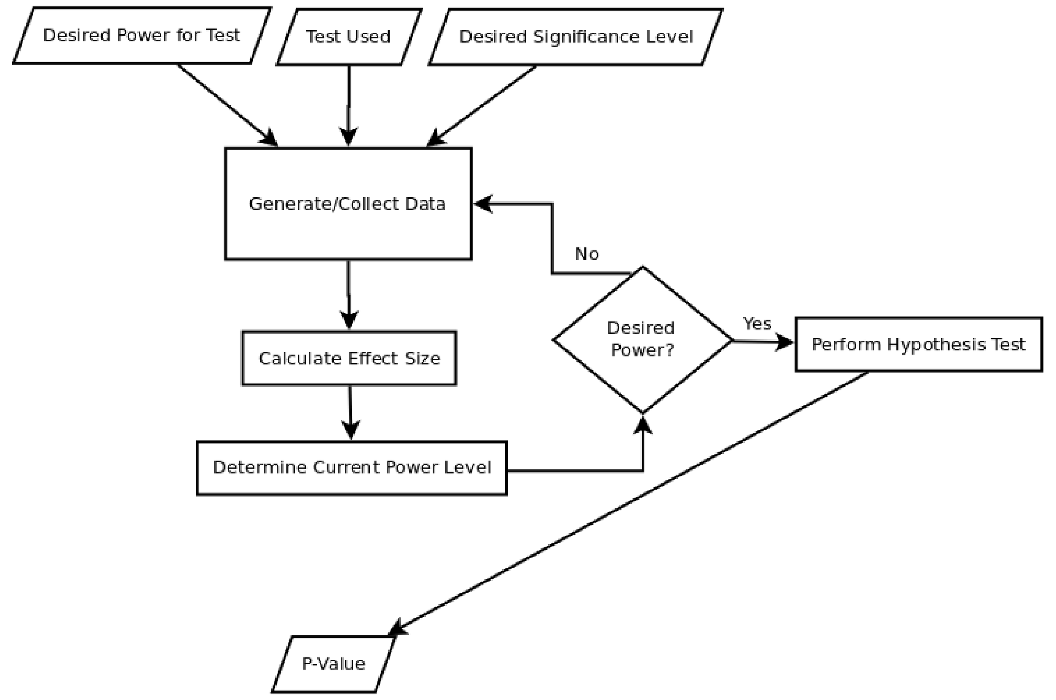
$$P(\text{Type I error}) = \alpha$$
$$P(\text{Type II error}) = \beta$$
$$P(H_0 \text{ is rejected } | H_0 \text{ is false}) = \gamma$$

In the above definitions γ is also known as the Power of a hypothesis test and α is known as the significance level of a hypothesis test. A slightly more elusive concept is the notion of effect size. The effect size of a particular test was defined as Cohen (above) as "the degree to which the phenomenon is present in the population, or the degree to which the null hypothesis is false." Cohen went on to provide several measures of effect size which are useful for different hypothesis tests, which are discussed in further detail later.

Description of Approach

The flow chart below provides a description of the procedure that is proposed in this poster. It is important to note that this procedure cannot, and should not be applied in all cases, however there are some instances when this procedure can benefit the experimenter.

Figure 1: Flow Diagram of Proposed Procedure



For example, this procedure may be useful at labs, which have sensors that automatically sample a particular subject many times a second, such as fluid flow or pressure in pipes.

When is this Approach Useful

1. Sampling Quickly and Accumulating Data is not a Problem.
2. Adding Effect size Calculations to processing is possible.
3. A desired power level and significance level is known.

Effect Size Measures

As mentioned above, Jacob Cohen, and several of his contemporaries created several different forms of effect size measurements, for various different tests. Some commonly used effect size measures for typical hypothesis tests are below:

Student's T-test:

$$d = \frac{\bar{x}_1 - \bar{x}_2}{s}$$

F-Test:

$$f^2 = \frac{R^2}{1-R^2}$$

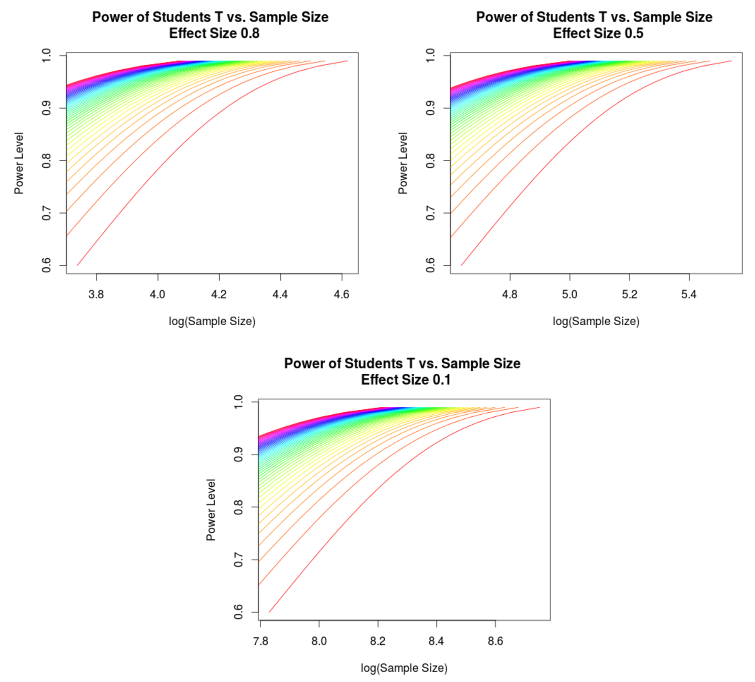
Chi-Square Test:

$$w = \sqrt{\sum_{i=1}^{n} \frac{(p_{0i} - p_{1i})^2}{p_{0i}}}$$

Results

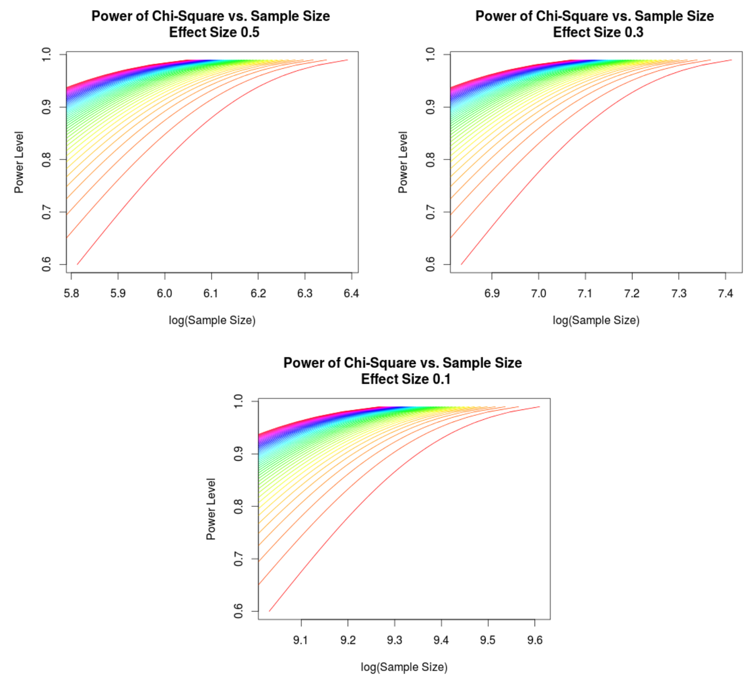
As mentioned in the abstract several plots were produced showing the various tradeoffs in sample size, effect size, power level, and significance level. Here are three such plots produced for our use of the Student's T-test.

Figure 2: Power Level Plots for T-Test



Note that the effect sizes used in each of these correspond to Cohen's standard effect sizes for large, medium, and small effects respectively. Similarly, plots for the Chi-Square Goodness of fit test were produced.

Figure 3: Power Level Plots for Chi-Square Tests

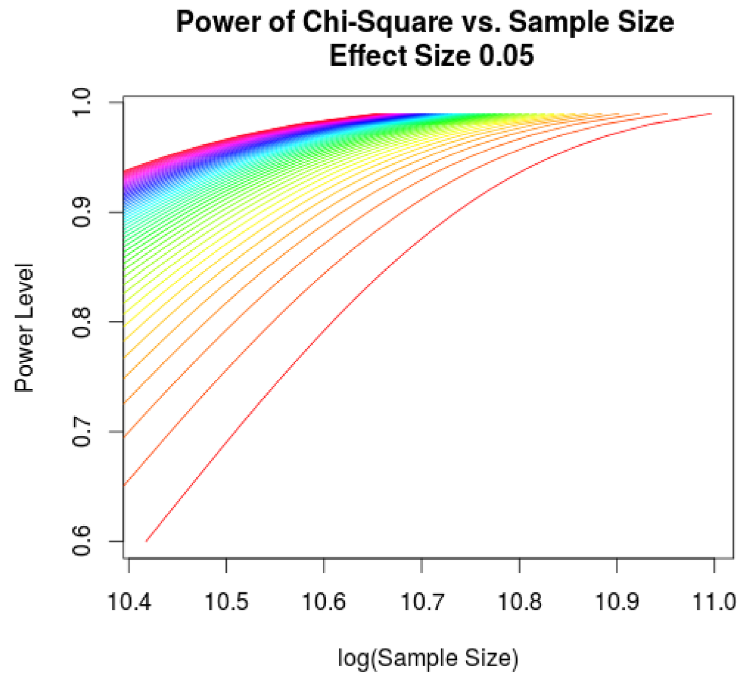


Note also that the x-axis which denotes sample size is on a log scale. These plots also represent a spectrum of different significance levels ranging from $\alpha = 0.001$ to $\alpha = 0.05$. The closer the line's color to red, the lower the significance level.

Hazard

NOTE: It is dangerous to estimate effect size as was done in these plots, as is shown in the flow diagram, effect size IS data dependent. As a final indicator of just how critical effect size is when performing power analysis consider the plot below:

Figure 4: Power Plot for Chi Square Test with Low effect size



Conclusions

- Power analysis is heavily dependent on the effect size measure. Effect size is a statistic (that is to say a function of the data). Therefore Power analysis is very data dependent.
- Power analysis should be done in conjunction with sampling, or after data is collected
- If effect size is estimated poorly, then Power analysis is not a useful way to glean optimal numbers of samples.
- If the theoretical effect size is known ahead of time, then power analysis can be used very effectively (no pun intended.)

Forthcoming Research

Time and memory complexity analysis for the inclusion of iteratively updating effect size measures. To this end we wish to minimize the impact on sampling speed of updating effect size measures, and wish to discover optimal methods for implementing the updating of effect sizes such that there prevalence is minimized in the sampling procedure, and their effect on timing is minimized.

Acknowledgements

The work presented in this poster could not have been done without the effort and results that Jacob Cohen proposed in his seminal text on Power Analysis: Statistical Power Analysis for the Behavioral Science (second edition) 1988, Routledge.