Research Article

Effect size for saw tooth power function in binomial trials

Serpil Aktaş

Department of Statistics, Hacettepe University, Beytepe, Ankara, TURKEY.
Email: spxl@hacettepe.edu.tr

Abstract Experiments should be designed to include an adequate number of units to address the study. With aninadequate sample size, we may not get valid conclusions. This paper aims to show the power analysis for a single sample from binomial data and to emphasize the saw tooth shape of power plots.The behaviour of power by sample size with desired power for specific alternative hypotheses for one sample proportion is illustrated on the graphics for various cases.
Keywords: sample size estimation, statistical power, binomial distribution, one sample case.

***Address for Correspondence:**
Dr. Serpil Aktaş, Department of Statistics, Hacettepe University, Beytepe, Ankara, TURKEY.
Email: spxl@hacettepe.edu.tr
Received Date: 02/05/2015 Accepted Date: 17/05/2015

Access this article online	
Quick Response Code:	Website: www.statperson.com
	DOI: 20 May 2015

INTRODUCTION

In the designing procedure of an experiment, what sample size is needed to detect a significant difference between the population and sample proportions is an issue for the researchers who want to estimate the minimum samplesize to assure a power to detect the difference. Particularly in the clinical studies it is important to minimize the number of individuals if it is a very time consuming or expensive experiment (McDonald, 2009). In statistical analyses, the power curve is generally expected to be monotonically increasing, namely, the more samples we draw, the more power we get (Cohen,1988). However, such as one sample proportion tests, the power function might be non-monotonic. Unlike the other tests, a relative increase in sample size can result in a decrease in power due to the actual significance level remains behind the target level. This "sawtoothed" similitude was discussed in Chernick and Liu (2002). The authors studied the sawtooth nature of the power function via nQuery Advisor for those whomay askhow large the sample size would be for the given power.We are also interested inhow to find sample sizes with desired power for specific alternativehypotheses for proportions. Power and Sample Size tools help the researchers to balance their limited resources.We illustrate the sawtooth behavior of the power plot on graphical illustrations in case of a single binomial proportion.

POWER AND SAMPLE SIZE ESTIMATION FOR ONE SAMPLE PROPORTION

Binomial distribution is very common in probability and statistics, because it may appropriate to summarize a group of independent observations by the number of observations in the group that represent one of two outcomes which are commonly labeled "success" and "failure". Even though the sampling distribution for proportions follows a binomial distribution, the normal approximation is used for the derivations. A one-sample proportion test involves testing the null

How to site this article: Serpil Aktaş. Effect size for saw tooth power function in binomial trials. *International Journal of Statistika and Mathemtika* May to July 2015; 13(1): 24-31. http://www.statperson.com (accessed 24 May 2015).

hypothesis H_0: p = p_0 versus the two-sided alternative hypothesis H_A: p ≠ p_0, the upper one-sided alternative H_A: p > p_0, or the lower one-sided alternative H_A: p < p_0. We want to test the hypothesis H_0:p=p_0 against the alternative H_A:p≠p_0 Thus, the minimum sample sizefor a two-sided interval follows,

$$n \geq \frac{\left(z_{1-\alpha/2}+z_{1-\beta}\right)^2 p(1-p)}{(p-p_0)^2} \qquad (1)$$

and the power,

$$1-\beta = \Phi\left(\frac{p-p_0}{\sqrt{\frac{p(1-p)}{n}}} - z_{1-\frac{\alpha}{2}}\right) + \Phi\left(-\frac{p-p_0}{\sqrt{\frac{p(1-p)}{n}}} - z_{1-\frac{\alpha}{2}}\right). \qquad (2)$$

Where, n is the sample size; p_0 is the null proportion or the value of the proportion under the null hypothesis; Φ is the standard normal distribution; α is Type I error ; β is Type II error, meaning 1- β is power. In order to do a power analysis, we need to specify an effect size which is the size of the difference between the null hypothesis and the alternative hypothesis that we expect to detect (Box and Hunter, 1978; Spiegel, 1992). The binomial test is based on the binomial distribution and the exact sampling distribution. Magnitude of an effect of interest or clinically meaningful difference, often expressed asan effect size. Specifying the effect size has an impact on estimating the sample size. The interpretation of “clinically meaningful” is usually determined by the researcher. For example, in clinical trials, if there is no prior knowledge about the performance of the considered clinical procedure, then a standardized effect size between 0.25 and 0.5 may be considered clinically meaningful (Chow et.al, 2008).

NUMERICAL EXAMPLES

Power and sample size calculations are a function of the specificalternative hypothesis of interest, in addition to some parameters. We use G*Power (Faul *et al.*. 2007) and SAS v9.2 to calculate power across for overall sample. Effect size depends upon the type of comparison under study. Effect size of 0.05 represents a "small" effect, around 0.15 a "medium" effect and 0.25 , a "large". The alpha level used in determining the sample size in most of the researches is 0.01, 0.05 or 0.10. We fixedα = 0.01 and 0.05 as thenominal significance level.The power of a hypothesis test is the probability of rejectingthe null hypothesis when the alternative hypothesis is true. Considering the possible scienerios given in Table 1, we draw plots of power versus sample size tosee how the sample size and power act each other. Power curves are produced for each combination of values.

Table 1: Parameters for calculating the power and sample size

Case	Constant Proportion	α Error Probability	Effect size
1	0.10	0.01	0.10, 0.15, 0.20, 0.25, 0.30
2	0.10	0.05	0.10, 0.15, 0.20, 0.25, 0.30
3	0.20	0.01	0.10, 0.15, 0.20, 0.25, 0.30
4	0.20	0.05	0.10, 0.15, 0.20, 0.25, 0.30
5	0.30	0.01	0.10, 0.15, 0.20, 0.25, 0.30
6	0.30	0.05	0.10, 0.15, 0.20, 0.25, 0.30
7	0.40	0.01	0.10, 0.15, 0.20, 0.25, 0.30
8	0.40	0.05	0.10, 0.15, 0.20, 0.25, 0.30
9	0.50	0.01	0.10, 0.15, 0.20, 0.25, 0.30
10	0.50	0.05	0.10, 0.15, 0.20, 0.25, 0.30
11	0.60	0.01	0.10, 0.15, 0.20, 0.25, 0.30
12	0.60	0.05	0.10, 0.15, 0.20, 0.25, 0.30

Instead of power curves, we can plot estimated sample sizes for a range of power values to get the approxiamate sample size for powers of interest or vice versa. In most cases, statisticians will probably be satisfied with the graphs rather than producing tables. Therefore we skipped the tables here. In the following graphs, we see the power plots for a particular test for testing H_0: p =p_0, against H_A: p > p_0. As it can be seen from the graphics that more data included in an analysis do not ensure the higher power. Therefore the power curve shapes like sawtooth eg. in Figure 13.

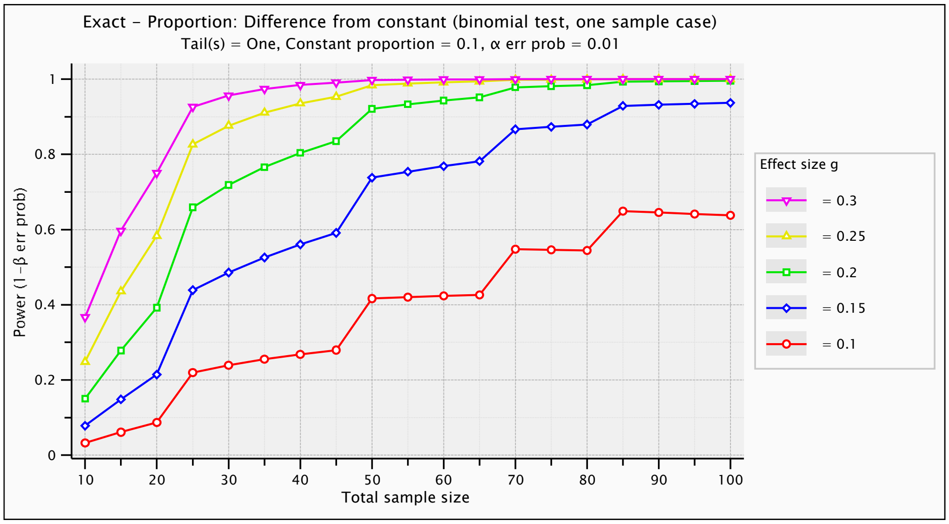

Figure 1: Power plot for P_0=0.1 ;α=0.01

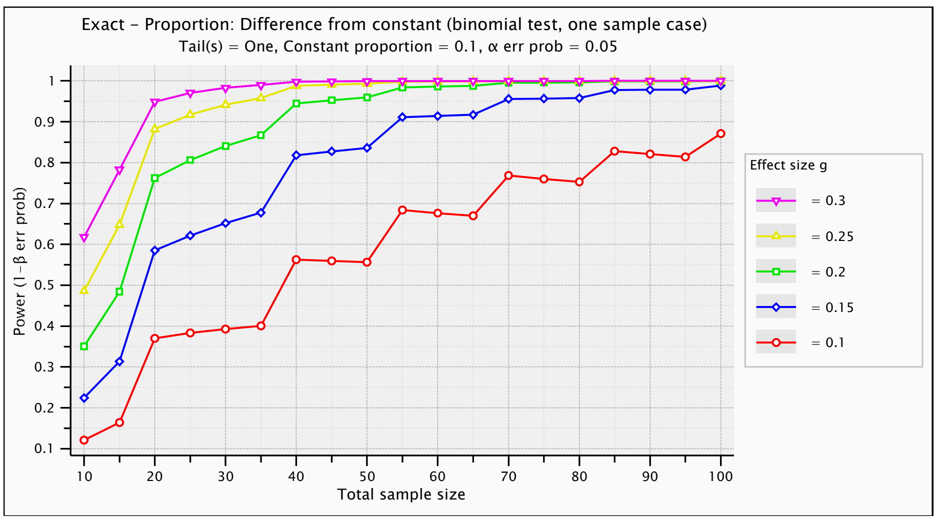

Figure 2: Power plot for P_0=0.1 ;α=0.05

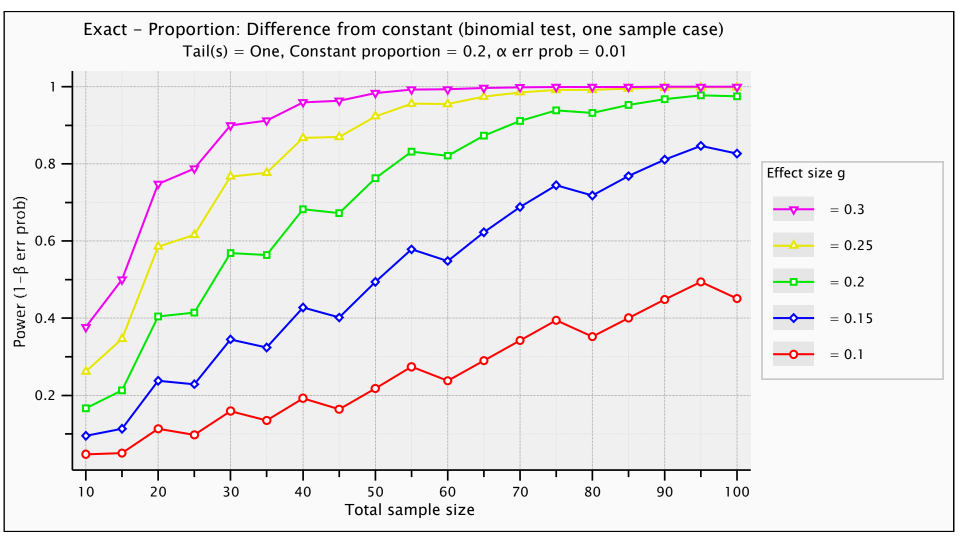

Figure 3: Power plot for P_0=0.2; α=0.01

From Figure 3, we have to draw at least 95 subjects are needed to detect the difference with 50% power using a 1% level for one sided test.

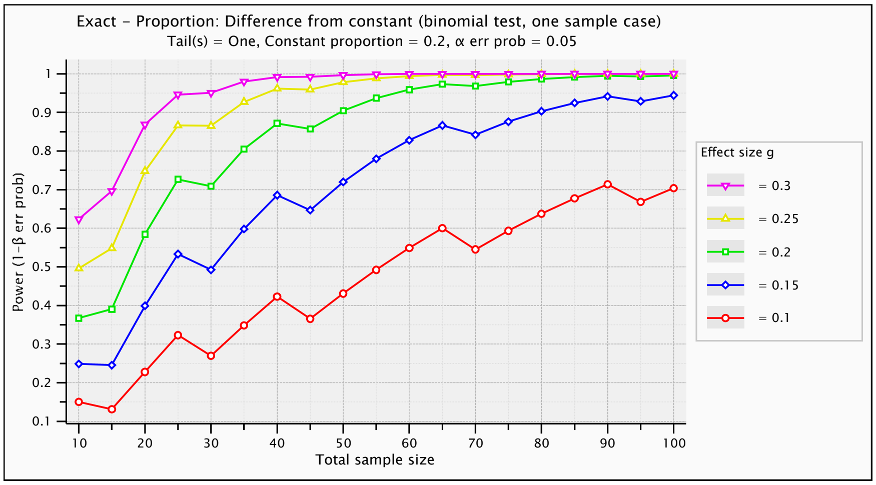

Figure 4: Power plot for P_0=0.2; α=0.05

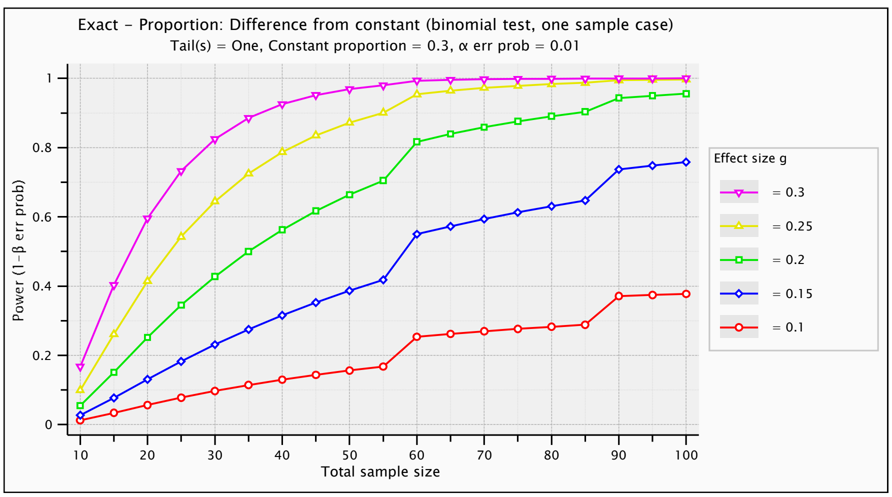

Figure 5: Power plot for P_0=0.3 ;α=0.01

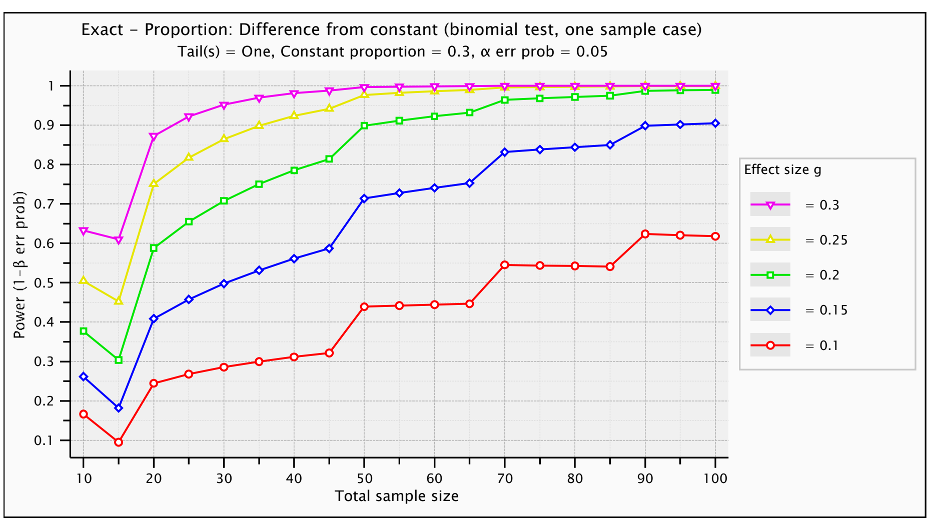

Figure 6: Power plot for P_0=0.3 ;α=0.05

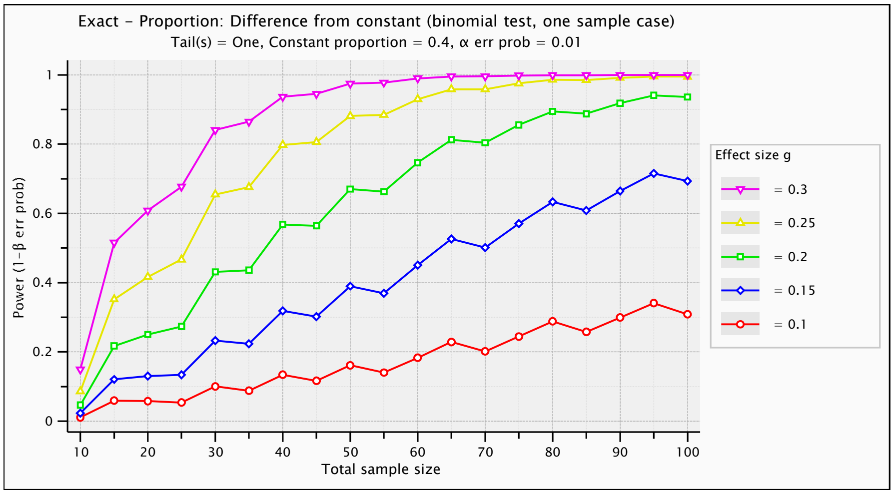

Figure 7: Power plot for P_0=0.4; α=0.01

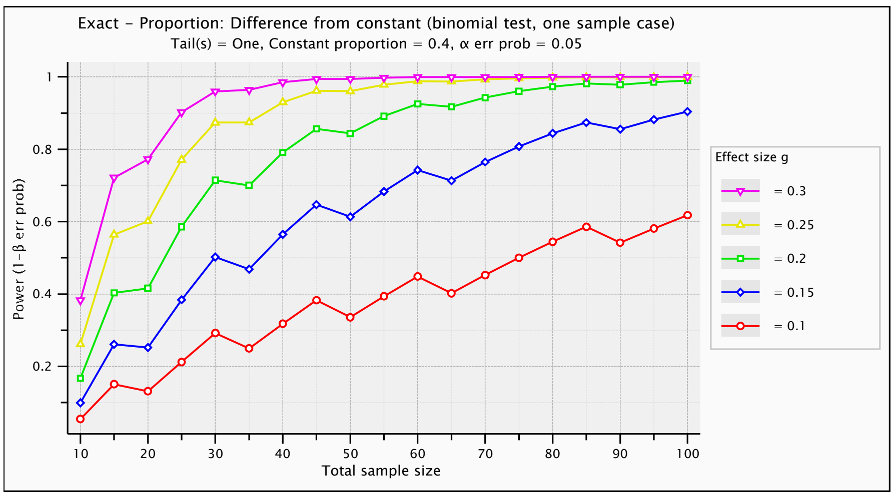

Figure 8: Power plot for P_0=0.4; α=0.05

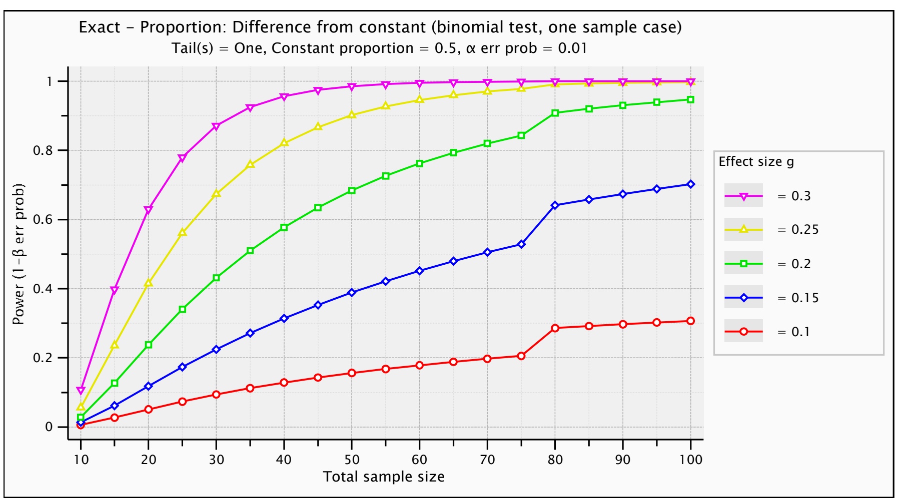

Figure 9: Power plot for P_0=0.5 ;α=0.01

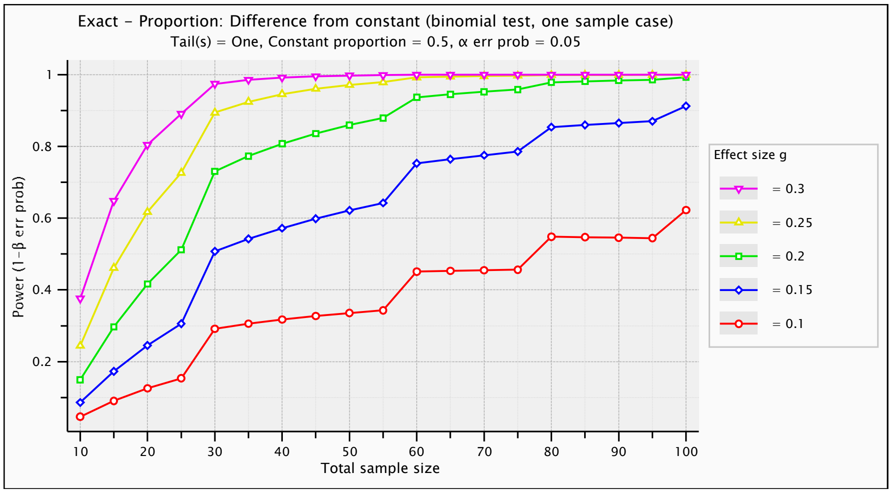

Figure 10: Power plot for P_0=0.5; α=0.05

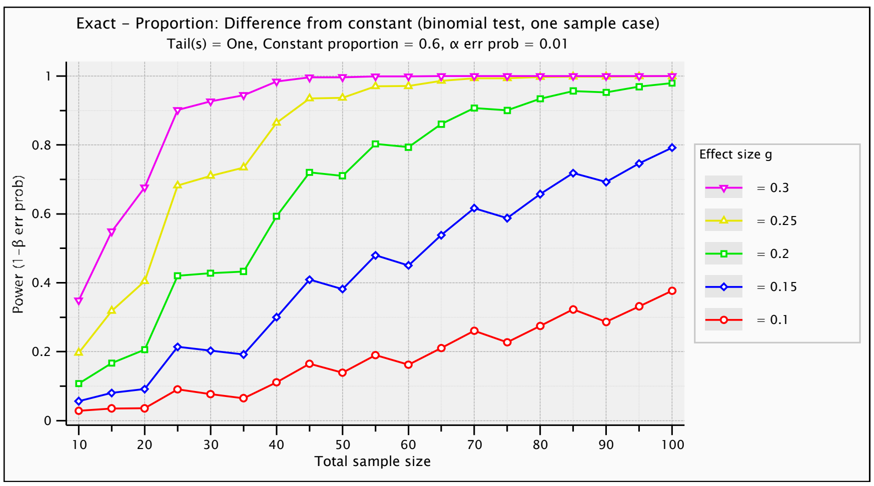

Figure 11: Power plot for P_0=0.6; α=0.01 As the output in Figure 11 shows, a sample size of 40 and effect size=0.3 would meet the highest power

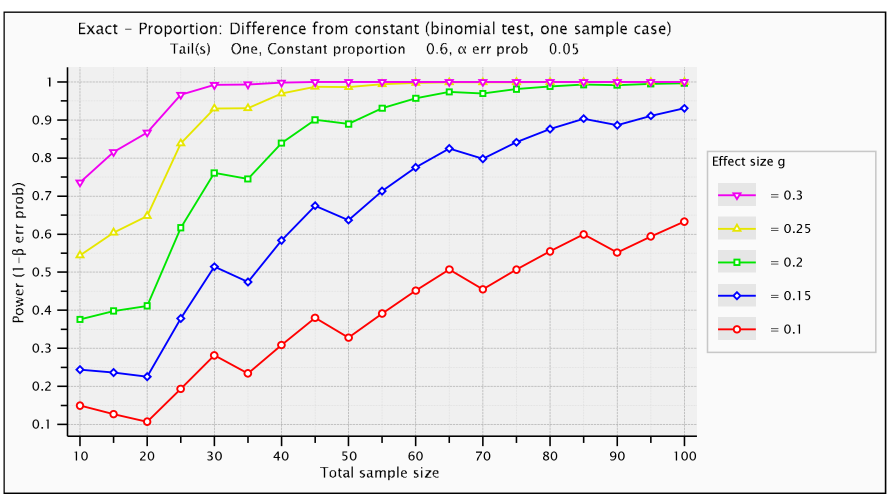

Figure 12: Power plot for P_0=0.4 ;α=0.05

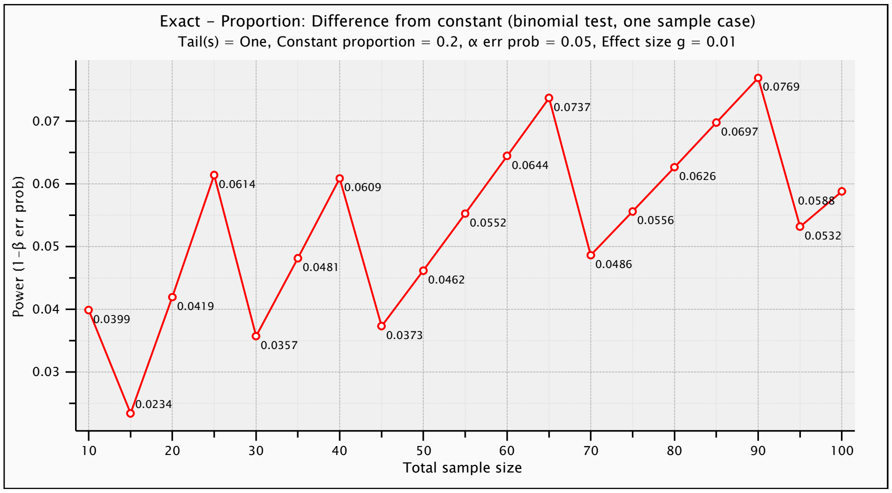

Figure 13: Power plot for P_0=0.2 ;α=0.05 ; Effect size=0.01

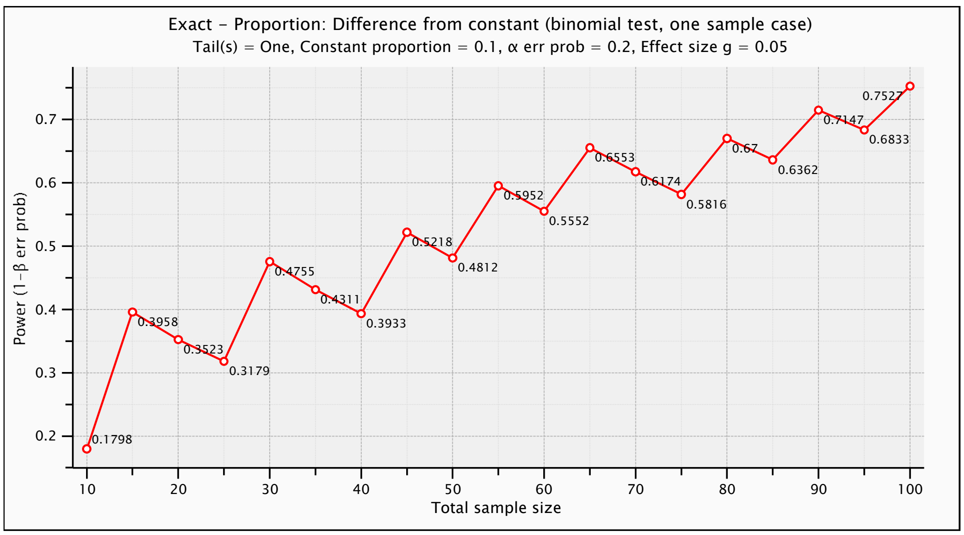

Figure 14: Power plot for P_0=0.1; α=0.2; Effect size=0.05

For instance, from the output in Figure 14, while n=90, we achive the power 0.7147 but for n=95, the power is 0.6833. Thesepower plots confirm the saw tooth shape for each combination. Sometimes fewer experiments give higher power and vice versa. Increasing trend sometimes is followed by a downward jump and then an upward trend again. A small increase in sample size results in a decrease in power.

CONCLUSION

In recent years, the exact methods have been replaced by normal approximations. Using the exact method, we pointed out the saw tooth nature of the power function for a single binomial proportion. Unlike a monotonic increase in the rate of rejection of H_0 as this null hypothesis is false, binomial test for one proportion does not exhibit this desirable property. The shape is like saw tooth. A study with large enough sample size will not have high enough power to detect the differences. The statistical expectation that power typically increases with increasing sample size might not hold in Binomial experiments. Experiments should be designed to include an adequate number of subjects to address the study. Studies that have either less number of subjects or larger number of subjects could be wasteful in terms of resources. Therefore, determining the appropriate sample size is the most important phase of a study. Performing sample size and power computations is often quite important.Power analysis can optimize the limited resources.

REFERENCES

1. Box, G.E.P., W.G.Hunter , J.S.Hunter, 1978. Statistics forExperimenters: An Introduction to Design, Data Analysis, and ModelBuilding, Wiley.
2. Chernick M.R, C.Y.Liu, 2002. The saw-toothed behavior of power versus sample size and software solutions: single binomial proportion using exact methods. The American Statistician, 56:149–155.
3. Chow, S.C., J. Shao, and H. Wang, 2008. Sample Size Calculations in Clinical Research. 2nd ed. New York: Marcel Dekker.
4. Cohen, J. 1988. Statistical Power Analysis for the Behavioral Sciences. 2nd ed. Hillsdale, NJ: Erlbaum.
5. Faul, F., E. Erdfelder, A.G. Lang, A.Buchner, 2007. G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behavior Research Methods, 39: 175-191.
6. Fleiss, J. L., B. Levin, and M. C. Paik, 2003. Statistical Methods for Rates and Proportions. 3rd ed. New York: Wiley.
7. McDonald, J.H. 2009. Handbook of Biological Statistics, 2nd ed. Baltimore: Sparky House Publishing.
8. Nietert, P.J.,and M. J. Dooley, 2011.The Power of the Sign Test Given Uncertainty in the Proportion of Tied ObservationsContemp Clin Trials. 2011 January ; 32(1): 147–150
9. SAS Institute Inc. 1999.SAS/STAT User's Guide, Version 8, Cary, NC: SAS Institute Inc.
10. Spiegel, M. R, 1992.Theory and Problems of Probability and Statistics. New York: McGraw-Hill.

Source of Support: None Declared
Conflict of Interest: None Declared