

## Teaching - Estimation of Minimum Sample Size and The Impact of Effect Size and Altering The Type - I & II Errors on it, in Clinical Research

K. R. Sundaram, J. Amrutha

### ABSTRACT

In any research study, one of the most important questions asked for at the time of planning the study is 'what should be the sample size in my study'. Estimation of minimum sample size depends upon- design of the study, whether it is an estimation problem or a hypothesis testing problem, the type of variable(s) to be studied, type-I & II errors acceptable and rough estimates of the parameters to be studied. In this presentation estimation of sample size for a clinical trial, comparing the efficacy of two different drugs, in case of categorical variables and the effect of altering the type I & II errors and the effect size have been discussed with examples. The results showed that the sample size increases when the type-I or / and II errors are decreased. Symmetric nature of sample size on either side with respect to the effect size has also been explained.

### BACKGROUND

In any research study, one of the most important questions asked for at the time of planning the study is 'what should be the sample size in my study'. Many students / researchers think that the figure on sample size can be got just by asking for it. It depends upon many factors-design of the study, whether it is an estimation problem or a hypothesis testing problem, the type of variable(s) to be studied (measurable or categorical), type - I & II errors acceptable and rough estimates of the parameters to be studied. Too large a sample size and too small a sample size are not recommended for many reasons. Too large a sample size, apart from resulting in unnecessary increase in the budget, may make a small difference in the parameter values between the groups statistically significant which may not be clinically significant. Similarly too small a sample size may not make a clinically important difference in the parameter values between the groups statistically significant. The pertinent question, then, is what should be the ideal / optimum / minimum sample size for any given

study based on the information required for estimation of sample size.

For any research study there could be budget and time constraints. We would like to have our estimates for a pre-fixed Confidence Coefficient (1 - type I error) and allowable error which are determined based on the time & budget at the disposal of the investigator. Similarly we would like to test the statistical significance of the research hypothesis for a pre-fixed Confidence Coefficient and Power (1 - type II error) of the test which are determined based on the time & budget at the disposal of the investigator. Also whether the test is one-tailed or two-tailed affects the sample size. Besides these factors the effect size (obtained from the previous studies or from a pilot study) which is the difference in the proportions of positive effect between the test group and the comparative control (placebo) group also determines the minimum sample size. Demonstration of the determination of minimum sample size for the above mentioned parameters based on hypothetical data has been done in this presentation.

### OBJECTIVES

To study the effect of increasing the Confidence coefficient and Power of the test in the estimation of minimum sample size.

To study the effect of testing one - sided hypothesis compared to two-sided hypothesis on the sample size.

To study the effect of the magnitude of the differences in the proportions of the variable, say, positive response between two populations, say, test group and control (placebo) group on the sample size.

### KNOWLEDGE REQUIRED

Concept of Standard Error, Type I & II errors, Confidence Coefficient & Power of the test Normal curve and One and Two -sided (tailed) tests (Bland, 2000; Alvan R F, 2001; Amritage, et. al., 2002; Sundaram, et. al., 2009).It is assumed that all the participants/readers have the necessary knowledge on the above indicated statistical concepts.

### METHODS

*Formula: (define  $n, Z, p, q, p_1$  &  $p_2$ )*

$$n = \frac{2 \times (Z_{1-\alpha/2} + Z_{1-\beta})^2 \times p \times q}{(p_1 - p_2)^2}$$

(Two - sided test)

where  $p = \frac{(p_1 + p_2)}{2}$ .

**TABLE ON Z - VALUES**

Significance level ( ) and Power (1- ) values for sample size calculation:

		Two - sided test	One - sided test
		$(Z_{1-\alpha/2})$	$(Z_{1-\alpha})$
Significance Level	0.01	2.576	2.326
	0.05	1.960	1.645
	0.10	1.645	1.282

		$Z_{1-}$
Power	0.80	0.842
	0.90	1.282
	0.95	1.645
	0.99	2.326

**RESULTS**

Table 1 gives the minimum sample size estimated for increasing Confidence Coefficients (95%, 99% & 99.5%) and Power (80%, 90% & 95%) for the effect size -0.05 ( $p_1 = 0.05$  &  $p_2 = 0.10$ ) and ( $p_1 = 0.90$  &  $p_2 = 0.95$ ), the two extreme values of  $p_1$  and  $p_2$ .

Confidence Coefficient (%)	Power (%)	Proportions $p_1 = 0.05$ & $p_2 = 0.10$		Proportions $p_1 = 0.90$ & $p_2 = 0.95$	
		2 - sided	1 - sided	2 - sided	1 - sided
95	80	435	342	435	342
	90	581	474	581	474
	95	718	598	718	598
99	80	648	557	648	557
	90	825	722	825	722
	95	987	874	987	874
99.5	80	946	856	946	856
	90	1157	1058	1157	1058
	95	1348	1240	1348	1240

Table 1: Sample size for different Confidence Coefficients & Power values for the two extreme combination of Proportions

It could be seen from Table 1 and Figures 1 (a) and 1 (b) that minimum sample size increases when the Confidence Coefficient & Power increase. Also, sample size required, for a given Confidence Coefficient & Power, is less for a one-sided test compared to two-sided test.

The results also showed that sample size is the same for the effect size ( $p_1 - p_2$ ) at the lower & upper extremes.

Table 2 gives the minimum sample size estimate for 95% confidence coefficient and 80% power for different effect sizes for different combinations from lower proportions to upper proportions (Table 2 (a) for effect size 0.05, Table 2 (b) for 0.10, Table 2 (c) for 0.20, Table 2 (d) for 0.30, Table 2 (e) for 0.40 and Table 2 (f) for effect size 0.50).



Fig.1(a)

Sample size for different Confidence Coefficients & Power values for the two extreme combination of Proportions for two-sided tests

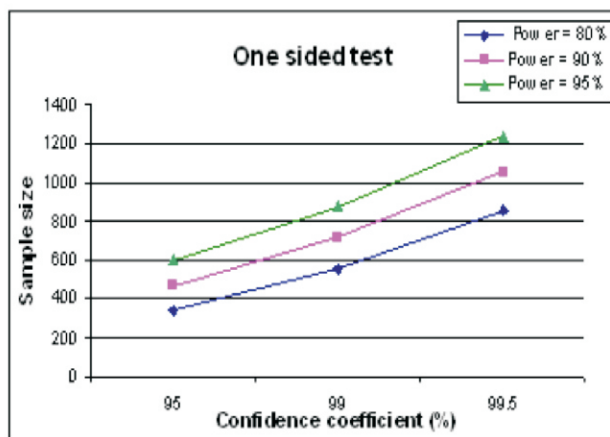


Fig.1(b)

Sample size for different Confidence Coefficients & Power values for the two extreme combination of Proportions for one-sided tests

The results showed that for all the effect sizes, the sample size was lower at the two extremes of the proportion of the two groups and was increasing for the corresponding increase in  $p_1$  and  $p_2$  from the lower side and for the corresponding decrease from the upper side resulting symmetry of sample size on either sides (upper and lower), showing symmetric nature of the sample size for the same effect size with different combinations.

**Table 2 (a): Sample size for Effect size = 0.05 for different combination of Proportions with 95% Confidence Coefficient & 80% Power**

Proportions	$p_1$	0.0	0.1	0.1	0.2	0.25	0.30	0.35	0.40	0.45	0.50
	$p_2$	0.1	0.1	0.2	0.25	0.30	0.35	0.40	0.45	0.50	0.55
Sample size		435	686	906	1094	1251	1377	1471	1534	1566	1566

0.55	0.60	0.65	0.70	0.75	0.80	0.85	0.90
0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95
1534	1471	1377	1251	1094	906	686	435

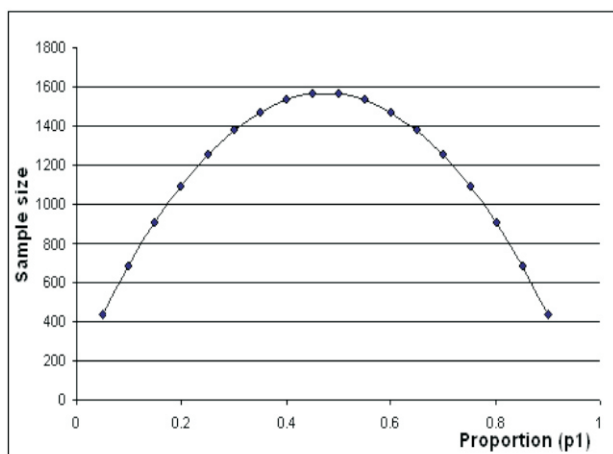


Fig. 2

**Sample size for Effect size = 0.05 for different combination of Proportions with 95% Confidence Coefficients & 80% Power**

**Table 2 (b): Sample size for Effect size = 0.10 for different combination of Proportions with 95% Confidence Coefficients & 80% Power**

Proportions	$p_1$	0.0	0.1	0.2	0.35	0.45	0.55	0.65	0.75	0.85
	$p_2$	0.1	0.2	0.3	0.45	0.55	0.65	0.75	0.85	0.95
Sample size		140	250	329	376	391	376	329	250	140

**Table 2 (c): Sample size for Effect size = 0.20 for different combination of Proportions with 95% Confidence Coefficients & 80% Power**

Proportions	$p_1$	0.0	0.1	0.2	0.35	0.45	0.55	0.65	0.75
	$p_2$	0.2	0.3	0.4	0.55	0.65	0.75	0.85	0.95
Sample size		49	72	88	96	96	88	72	49

**Table 2 (d): Sample size for Effect size = 0.30 for different combination of Proportions with 95% Confidence Coefficients & 80% Power**

Proportions	$p_1$	0.0	0.1	0.2	0.35	0.45	0.55	0.65
	$p_2$	0.3	0.4	0.5	0.65	0.75	0.85	0.95
Sample size		27	35	41	42	41	35	27

**Table 2 (e): Sample size for Effect size = 0.40 for different combination of Proportions with 95% Confidence Coefficients & 80% Power**

Proportions	$p_1$	0.0	0.1	0.2	0.35	0.45	0.55
	$p_2$	0.4	0.5	0.6	0.75	0.85	0.95
Sample size		17	21	23	23	21	17

**Table 2 (f): Sample size for Effect size = 0.50 for different combination of Proportions with 95% Confidence Coefficients & 80% Power**

Proportions	$p_1$	0.0	0.1	0.2	0.35	0.45
	$p_2$	0.5	0.6	0.7	0.85	0.95
Sample size		12	14	14	14	12

## DISCUSSION AND CONCLUSION

The aim of this paper is mainly for educating the students of Statistics and the research workers, the method of estimation of minimum sample size for different values of confidence coefficients and power of the test and for different effect sizes. The students and researchers will get clear idea of the impact of type I and type II errors and the direction of the test depending upon the effect size. Minimum sample size can be determined and finalized depending upon the time and budget available for the study by altering type I and type II errors and also depending upon the direction of the test (one-sided or two-sided).

**REFERENCES**

1. Bland, M. An introduction to Medical Statistics. Oxford University Press, 2000.
  2. Alvan, R. F. Principles of medical statistics, CRC press, 2001.
  3. Armitage P, Berry G, & Matthews, J. N. S. Statistical Methods in Medical Research. 4<sup>th</sup> Ed., Blackwell science, 2002.
  4. Sundaram, K. R., Dwivedi, S. N., & Sreenivas V. Medical Statistics -Principles & Methods. BI Publications, New Delhi, 2009.
-