REVIEW ARTICLE

SAMPLE SIZE DETERMINATION (FOR DESCRIPTIVE STUDIES)

*Simarjeet Kaur

M.Sc. Community Health Nursing

ARTICLE INFO	ABSTRACT
	Sample size determination is the essential step of research methodology. It is an act of choosing the number of observer or replicates to include in statistical sample. In some situations, increase in precision for larger sample size is minimal or even nonexistent. Sample sizes are judged based on the quality of resulting estimates. Sample size determination in descriptive studies is different than experimental studies. Sample size determination helps in increasing the quality of evidence based research.

INTRODUCTION

Quantitative researcher needs to pay careful attention to the number of subjects required to test the hypothesis adequately. A sample is the representation of given population and it is worth if it actually represents the population.

GENERAL INSTRUCTIONS

- **First,** the need for considering sample size will be **reviewed.**
- **Second,** the study **design parameters** affecting sample size will be identified.
- **Third**, **formulae** for calculating appropriate sample sizes for some common study designs will be defined.
- Sample size should be estimated **early in the design** phase of the study.

Definition

Sample size determination is the act of choosing the number of observations or replicates to include in a statistical *sample*.

Or

Sample size determination is the mathematical estimation of the number of subjects/units to be included in a study.

Corresponding author: **Simarjeet Kaur,**
M.Sc. Community Health Nursing

Need

Optimum sample size determination is required for the following reasons:

1. To allow for appropriate analysis.
2. To provide the desired level of accuracy
3. To allow validity of significance test etc.

How large a sample do i need?

If the sample is too small:

- Even a well conducted study may fail to answer its research question
- It may fail to detect important effect or associations
- It may associate this effect or association imprecisely

If the sample size is too large:

- The study will be difficult and costly.
- Time constraint
- Available cases e.g. rare disease.
- Loss of accuracy.

Steps to use sample size formulae

1. 1st Formulate a research question
2. 2nd Select appropriate study design, statistical significance.

3. 3rd use the appropriate formula to calculate the sample size

Sample size in quantitative studies

1. The larger the sample, the more representative
2. The larger the sample, the smaller the sampling error.
3. Descriptive studies and correlational studies require large samples.
4. Quasi–experimental and experimental studies use smaller samples than descriptive and co -relational studies

Sample size in quantitative studies (Non-Experimental/ Descriptive Studies)

Level of confidence

- Probability of value of parameters will fall within specified range, closely connected with the level of significance for statistical tests.
- For example, we can be '95% confident' that the true mean value lies somewhere within a valid 95% confidence level, corresponds to significance testing at the 5% level ($P < 0.05$) of significance.
- Likewise, we can be '99% confident' that the true mean value lies somewhere within a valid 99% confidence level, corresponds to significance testing at the 1% level ($P < 0.01$) of significance

Precision

A measure of **how close an estimate** is to the **true value** of a population parameter. Or it can be thought of as the **amount of fluctuation from the population parameter** that we can expect by chance alone in sample estimates

Degree of Precision

- This is presented in the form of a *confidence interval (Range of values within which confidence lies)*. For example, a survey of a sample of patients indicates that 35 per cent smoke.
- We can accept that the figure for the wider population lies between 25 and 45 per cent, (allowing a margin for random error (MRE) of 10% either way)

Smaller the allowed margin for random error, the larger the sample must be (See in Table)

Margin for random error	Sample size
+ or – 10%	88
+ or – 5%	350
+ or – 3%	971
+ or – 2%	2188
+ or – 1%	8750

*Two way of calculation in non-experimental studies

Estimate mean value   Estimate proportion/%
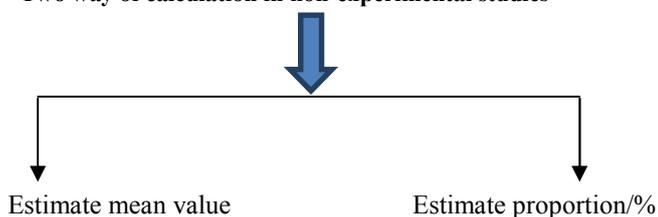
Example 1: How large must a sample be to estimate the mean value of the population?

- *Suppose we wish to measure the number of times that the average patient with asthma consults her/his general practitioner for treatment?*

The formula to calculate the sample size for a mean estimate is:
$N = (SD/SE)^2$

Where N = the required sample size,
SD = the standard deviation, and
SE = the standard error of the mean

The standard deviation could be estimated either by looking at some previous study or by carrying out a pilot study.

- Suppose that previous data showed that the standard deviation of the number of visits made in a year was 20.

- First, the SE (standard error) is calculated by deciding upon the accuracy level which you require.
- If you want a 95% confidence level, then divide the maximum acceptable MRE (margin for random error) by 1.96 to calculate the SE.
- If instead you want a 99% confidence level, then divide the maximum acceptable MRE by 2.56 to calculate the SE.
- b) the standard error is 5 divided by 1.96 = 2.55

The formula as follows:
$N = (SD/SE)^2$
$(20/2.55)2 = (7.84)2$
$= 61.4 = 61$ (rounded to nearest patient)

Example 2: How large must a sample be to estimate a proportion / percentage?

- You want to conduct a survey of the proportion of men over 65 who have cardiac symptoms
- Your significance level is 95%
- Your acceptable margin for random error is plus or minus 2 per cent
- From previous studies work you estimate that the proportion is about 20 per cent

a) Calculate the SE = ...MRE/1.96...............
b) Using the formula for the sample sizes for a proportion, calculate:
$N = P(100\%-P)/(SE)^2$

First, the SE can be calculated by dividing the confidence interval by 1.96
$SE = 2/1.96 = 1.02$
We then calculate:
$N = P(100\%-P)/(SE)^2$
With P = 20% and SE = 1.02,
We have $N = 20(100-20)/(1.02)2$
$20 \times 80/1.04 = 1539$

- In case no literature for proportion, it is safer to assume the worst case scenario and assume that the proportion is likely to be 50%.
- Other things being equal, this will allow for the largest possible sample size - and in most circumstances it is

preferable to have a slight overestimate of the number of people needed, rather than an underestimate

***Modified Simple Formula**

For survey type- prevalence study

where p is prevalence
q is the 1 – p
d is the precision

(Assuming 95% confidence level)

$$ n = \frac{4\,p\,q}{d^2} $$

n is the sample size required for the study

****Key points to remember**

- Increasing the sample size increases the sensitivity of the study to detect a difference between the groups being compared
- To estimate the sample size for a descriptive study, in order to estimate a mean or a proportion, it is necessary to specify the maximum acceptable margin for random error.

***Software for calculating Sample Size**

PS - Power & Sample Size Calculation
by Vanderbilt University, USA

http://www.calculator.net/sample-size-calculator.html
http://www.raosoft.com/samplesize
Http://www.kck.usm.my/ppsg/stats_resources.htm

Conclusion

The purpose of sample size formulae 'is not to only give an samples number…but rather to subject the study design to scrutiny, including an assessment of the validity and reliability of data collection, and to give an estimate to distinguish whether tens, hundreds, or thousands of participants are required.

REFERENCES

Kothari, C.R. 1985. Research Methodology. 3rd edi. Delhi: New age international publishers; p.162-170

Macfarlane SB. 1997. Conducting a Descriptive Survey: Choosing a Sampling Strategy. *Trop Doct*., 27(1): 14-21.

Naing, L., T. Winn, B.N. Rusli1. 2006. Practical Issues in Calculating the Sample Size for Prevalence Studies. *Archives of Orofacial Sciences*, 1: 9-14

Nihssamplingsamplesizecalcutaion.pdfavailablefrom:https://www.researchgate.net/file.PostFileLoader.html

Yamane, Taro. 1967. *Statistics: An Introductory Analysis*, 2[nd]Ed. New York: Harper and Row.

****