International Journal of Advances in Engineering and Management (IJAEM)
Volume 3, Issue 1 Jan-Feb 2021, pp: 147-149 www.ijaem.net ISSN: 2395-5252

Basics of Sample Size Determination

Himangini Singh,
University Institute of Management, RDVV Jabalpur

Submitted: 10-01-2021 Revised: 23-01-2021 Accepted: 26-01-2021

ABSTRACT: Applied statistics research plays pivotal role in diverse problems of social sciences, agricultural sciences, health sciences, and business research. The technique of sampling and determination of sample size have crucial role in survey-based research problems in applied statistics. If the sample size is inappropriate it may lead to erroneous outcome and similarly if excessive sample size is drawn from the population it can influence the conclusion and quality of the research. The present paper gives an overview of the basic terms which emphasizes on several parameters for selecting appropriate sample size.

Keywords: Sampling, Sample Size, Level of precision, Confidence level, Degree of variability

I. INTRODUCTION:

Sampling approach plays a vital role in marketing research. The question about the sample size invariably is raised while conducting any research or survey. The most frequently asked questions are: How to calculate the sample size? The level of statistical accuracy? Is the size important? How important is the range of variability?

There are various factors which affect these questions like the objective of the study, population size (known or unknown), benefits of selecting a good sample size, disadvantage of choosing a bad sample and how much error is permissible. Too small sample may prevent the findings from being extrapolated, whereas too large sample may amplify the detection of differences, emphasizing statistical differences that are not relevant.[1] The determination of sample size cannot be done without considering the statistical and non-statistical features in research setting. Resources, budget, time, ethical standards falls into the category of the non statistical factors and precision level, confidence level and the degree of variability are the statistical factors (Miaoulis and Michener, 1976).

Population to be sampled

The word population is used to denote the aggregate from which the sample is chosen. The population to be sampled should represent the target population. But in many studies the sample chosen from the population is deviated towards the convenience of the researcher rather than the actual source of information. The analysis and the results are drawn from the chosen sample and are applicable for the target population. Though there might be some differences between the chosen sample and target population.

The three criteria's of sample size determination should reflect the purpose of the study and the population to be sampled is described below:

The level of precision:

The specification of the level of precision is a crucial step. The level of precision is also known as sampling error is the range in which the true value of the population is estimated to be. This range is often expressed in percentage points, (e.g., ±5 percent). If the sample size increases, the level of standard error declines and it becomes more accurate in terms of statistical precision. The outcome of any study or research always has the probability of some uncertainty due to the definite portion chosen from the population to be measured. It refers to the acceptable errors in the estimate. The probability of the uncertainty can be reduced by taking large sample size.

The confidence level

A **confidence level** refers to the percentage of all possible samples which can represent the inclusion of the population parameter within the confidence interval (specified range). The 95% **confidence level** means you can be 95% certain that the outcome will meet the standard expectations 95% of the time. It means if you repeat the survey you, the level of certainty to obtain the same result is 95%. This confidence level is common and widely used by the researchers. A 100% confidence level is imaginary in statistics unless the researcher surveys the whole population and exempts all types of bias. The

[1] Altman DG. Practical Statistics for Medical Research. London, UK: Chapman & Hall; 1991.

confidence level when stated as a proportion is known as confidence coefficient. For example, the confidence level of 95% when transformed into the confidence coefficient would be .95.

The formula derived for the confidence interval for µ as

$$\bar{X} \pm Z\frac{\sigma}{\sqrt{n}}$$

- The formula to determine sample size where the mean of continuous outcome variable in a single population is supposed to be estimated by:

$$n = \left(\frac{Z\sigma}{E}\right)^2$$

- The formula to determine sample size where the mean is to be calculated for single sample but for dichotomous outcome is:

$$n = p\left(1-p\right)\left(\frac{Z}{E}\right)^2$$

- The formula to determine sample size where the difference in mean between two independent population is to be calculated is:

$$n_i = 2\left(\frac{Z\sigma}{ES}\right)^2$$

- The formula to determine sample size where the difference in mean of a continuous outcome which depends on matched data is:

$$n = \left(\frac{Z\sigma_d}{E}\right)^2$$

- The formula to determine the sample size where the difference in proportions between two independent is:

$$n_i = \left\{p_1\left(1-p_1\right)+p_2\left(1-p_2\right)\right\}\left(\frac{Z}{E}\right)^2$$

Z is the value from the standard normal distribution
σ is the standard deviation of the outcome variable
E is the desired margin of error.
p is the proportion of successes in the population.
σ_d is the standard deviation of the difference scores.

Degree of variability

The sampling variability defines the range of estimate variation between samples. There are two basic measures of variability: the variance (σ^2) and standard deviation (σ). **Sampling variability refers to the distribution of attributes statistically in the population and never a population** Changes in variability fluctuates with the increase and decrease ratio in the sample size. Increasing or decreasing sample sizes leads to changes in the variability of samples. For example, a sample size of 10 respondents is taken from the population of 50000 and if the 100 respondents are chosen as the sample size from the same population of 50000. It is very likely to conclude a very different outcome than a sample size of 10. There is no such sample size which can lead to the accurate estimate for any statistical measures of variability. The requirement of the larger and smaller sample size depends on the heterogeneity and homogeneity of the population. The researcher has to select large sample size if the population is heterogeneous and small size if the population is more homogeneous(less variability).

II. CONCLUSION:

Sample size determination is a very important aspect of any research. A researcher could study this as a tool for understanding the basics of determining the sample size. The sample size also is often increased by 30% to balance the nonresponsive rates and other barriers to achieve the specified level of confidence and the degree of precision. The three basic mentioned factors affect the validity and reliability of the conclusion. The optimization of sample size plays an important role so the important the number of experimental units should not be very small as to avoid the interpretation of the existing effect and not very large as it leads to the wastage of time, money and other resources.

To ensure the external validity or generalizability of research findings to the target population, the researcher must satisfactorily answer the question of whether the results of the survey would have been the same if a 100% response rate had been achieved (Richardson, 2000). Controlling for nonresponse error begins with designing and implementing research, following generally acceptable protocols and procedures (Dillman, 2000). Appropriate sampling protocols and procedures should be used to maximize participation in a study. Once participation has been maximized, the researcher will have obtained a high enough Journal of Agricultural Education 50 Volume 42, Issue 4, 2001 Lindner, Murphy, Briers Handling Nonresponse… response rate to conclude that nonresponse is not a threat to external validity or obtained a response rate that warrants additional procedures for ensuring that nonresponse is not a threat to external validity.

REFERENCES:

[1]. Lindner, Jimmy & Murphy, Tim & Briers, Gary. (2001). Handling Nonresponse in Social Science Research. Journal of Agricultural Education. 42. 10.5032/jae.2001.04043.
[2]. Cochran, W. G. 1963. Sampling Techniques, 2nd Ed., New York: John Wiley and Sons, Inc.
[3]. Miaoulis, George, and R. D. Michener. 1976. An Introduction to Sampling. Dubuque, Iowa: Kendall/Hunt Publishing Company.
[4]. Israel, Glenn D. 1992. Determining Sample Size. IFAS Extension, University of Florida. PEOD-6.
[5]. Lisa Sullivan,_Power and Sample Size Determination. PhD Professor of Biosatistics Boston Univeristy School of Public Health

International Journal of Advances in Engineering and Management

ISSN: 2395-5252

IJAEM

Volume: 03 Issue: 01 DOI: 10.35629/5252

www.ijaem.net

Email id: ijaem.paper@gmail.com