ISBN: 978-81-971650-3-0

9. Sampling Fundamentals

Suranjana Das

MSc. Student, Department of Botany,
Gauhati University, Guwahati.

Abstract:

Collection of relevant information on the subject of study is the one of the most important steps in research methodology. In other words, selection of the appropriate data and collecting it for the purpose of testing the hypotheses or to make a generalisation of the study constitute an important part of the research.

The process of sample selection demands thorough understanding and knowledge on the concept of population, sample and various sampling techniques. This article focusses on a brief discussion about the concept of sampling, its importance in research field and the fundamentals associated with the sampling process.

Furthermore, there exist a number of different types of sampling methods about which the researcher should have a clear idea for its proper application during the research. In the regards, this article also presents a brief about the different sampling methods.

Keywords:

Sampling Method, Research Methodology, Sampling fundamentals, Probability Sampling, Non-probability Sampling.

9.1 Introduction:

Samples are the part of every research from which inferences could be drawn. The accuracy of a research is heavily influenced by the process of sampling. Sampling involves the strategic selection of a sample from a large group of individuals or population for the purpose of conducting research.

The sample could be a biological specimen or any individual chosen to derive statistical inferences and to predict the characteristics of the entire population. Sampling offers a pragmatic and practical approach to identify the features of the entire population, which otherwise it would be extremely expensive and time-consuming to survey the entire population for a research study.

Various sampling strategies are available to define and collect samples that will form the basis of a research study. Best sampling strategy for research is mainly determined by the purpose of study, available resources, time constrains and research hypothesis. This article provides a brief description about the various aspects of sampling procedure.

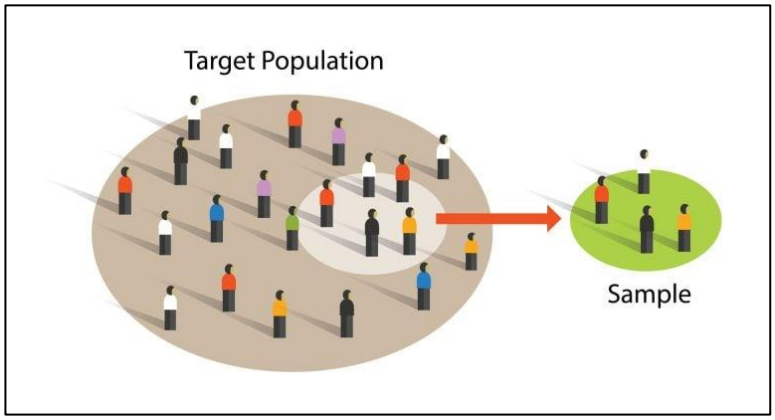

Figure 9.1: Selection of sample from the target population (adapted from: Saul Mcleod 2023)

9.2 Need of Sampling in Research:

- Sampling is cost-effective as well as time-effective. Sampling helps to reduce expenditure, permit measurement of greater scope and produces results at a relatively faster rate [13].
- A sample is the representative of a population which helps the researcher to obtain a generalised finding about the characteristics of the entire population.
- Sampling helps the researcher to choose a sample size appropriate for the study, which enables to obtain more accurate measurements as well as improves the precision of the findings [13].
- Sampling improves the feasibility of a research by allowing the researcher to conduct the studies that would have been impossible if the entire population has to be studied.
- Although there is always a possibility of the occurrence of biases during a study, but use of sampling in research can help to reduce the bias to certain extent by allowing the researcher to choose the participants intentionally.
- Sampling allows the researcher to draw conclusions about the population based on the characteristics of the sample, which makes sampling an essential component for statistical analysis.

9.3 Sampling Fundamentals in Research Methodology:

A. Universe / Population:

There consists of a particular collection of units which are required during the time of a particular study. Each unit can be considered as an entity to make observations based on certain well-defined procedure. The entire collection of such units which is the focus of a particular enquiry represents Universe or Population. The term ‘Universe’ in terms of statistics refers to the total of the entities or units in the field of inquiry. Whereas ‘Population’ refers to the total no. of the entities about which one desires to gather information i.e., population is the target group to be studied. However, a population can be

finite or infinite. Finite population consists of a fixed number of elements whereas in infinite population, the number of elements is infinite. Finite population helps to enumerate the results in totality, whereas in infinite population, it is theoretically impossible to observe all the elements. The population of a city, the number of workers in a factory are the examples of finite population whereas the number of stars in sky or the number of listeners of a broadcasting programme represents the examples of an infinite population.

B. Sampling frame:

A sampling frame represents the comprehensive and accurate list of individuals with attributes of the target population from where the samples can be drawn. In other words, it is the basis for defining, identifying, selecting and representing all the members who will the potential representative of a particular study. The selection of sampling frame depends on the research objectives, characteristics of the target population and the availability of the relevant data. It becomes an essential element in any research endeavour, because it ensures that the samples selected from the population represents the actual members being targeted in the research. It also helps to eliminate sample bias or any random group unrepresentative of the target research.

C. Sampling Design:

Sampling design is a definite criterion for obtaining a sample from the population of interest. It refers to the procedures for selecting items for the sample during research. There are various sample designs from which a researcher can choose. Selection of a sample design is an important task in the part of a researcher. Some designs are relatively more precise and easier in application than others. A researcher should always choose a sample design which is easily applicable, reliable and appropriate for the research purpose. Sampling design should be determined before any data has been collected.

D. Statistics and Parameters:

Statistics and parameters are basically the numbers that are used to summarise the properties of a sample or population. Parameters represent the numbers that define the properties of entire population whereas statistics represents the numbers that describe the population of a sample. The main objective of a sample analysis is to obtain the estimate of a parameter from a statistic. The mean value of a sample statistic in a sample distribution is assumed to be an estimate of the unknown population parameter [12]. There are a number of attributes that exist to evaluate, which can generate a number of different parameters and statistics. For e.g.: when any measurement is done on a continuous scale, then various summary values can be calculated for statistics and parameters like mean, median, standard deviation and correlation. And when the measurement is categorical, the parameter or statistics will often be a proportion.

E. Sampling Error:

Sampling errors are the random deviation of the selected sample from the original characteristics, behaviours, qualities or figures of the entire population. The most common

cause of occurrence of such an error is due to a biased sampling procedure or randomly by chance. The magnitude of a sampling error can be affected by a number of factors including sample size, nature of the sample and the population's variability. The more homogenous the sample, lower is the sampling error. Sampling error is inversely related to the size of the sample i.e., as the sample size approaches the size of a population, the standard error approaches zero. Moreover, variation in population can also affect the occurrence of sampling error. Increase in the variation in a population can cause increase in the variability of a sample and make it unrepresentative of the overall population. Sampling errors can never be completely eliminated. However, with the application of different sampling theories, the effect of sampling error can be minimised [11].

F. Precision:

Precision of a sample can be defined as the variation among all the samples that are used to measure the population parameters. Precision is analogous to sample variance with respect to its reflection in the width of confidence interval of a sample. A number of random measurement errors as well as the variability that exist within a population can have an effect on the precision of a sample. Precision is an important factor to consider when designing the inventory or monitoring program because sample variance is the foundation of subsequent statistical analysis. Therefore, sampling precision can be used as an expression of the confidence in the collected sample although it can be quantified from the collected data **[**1].

G. Confidence Level & Significance Level:

Confidence level depicts the surety of an inference. It is expressed in terms of percentage and represents the number of times that the actual value will fall within the stated precision limits. For e.g. If we assume a confidence level of 95%, we mean that there are 95 chances in 100 that the sample results represent the true condition of the population within a specified precision range against 5 chances in 100 that it does not.

However, the level of significance is the fixed probability of wrong rejection of the null hypothesis, when in fact it is true. Hence, confidence level indicates the likelihood of the result to fall within the precision range and significance level indicates the likelihood that the result will fall outside the precision range. It can thus be inferred that if the confidence level is 95%, then the significance level will be (100-95) i.e., 5% [12]. Also, it is an important fact to note that the area of the normal curve within the precision limit for the specified confidence level constitute the acceptance region while the area of the curve outside these limits in either direction constitutes the rejected region.

H. Sampling Distribution:

Sampling distribution is the frequency distribution of a sample statistic obtained from a set of samples. It is a type of probability distribution based on a large number of samples of size 'n' from a given population. The sampling distribution tends to have more values clustered around the mean with fewer values scattered around the mean. With infinitely large number of samples, this distribution will approach a normal distribution [12].

The shape of a normal sampling distribution is a bell-shaped curve with a single peak and two tails extending symmetrically in either direction. The centre of the sampling distribution of sample mean is the true population mean, μ. Sampling distribution is closely associated with the Central Limit Theorem. The Central Limit Theorem helps in constructing the sampling distribution of a mean.

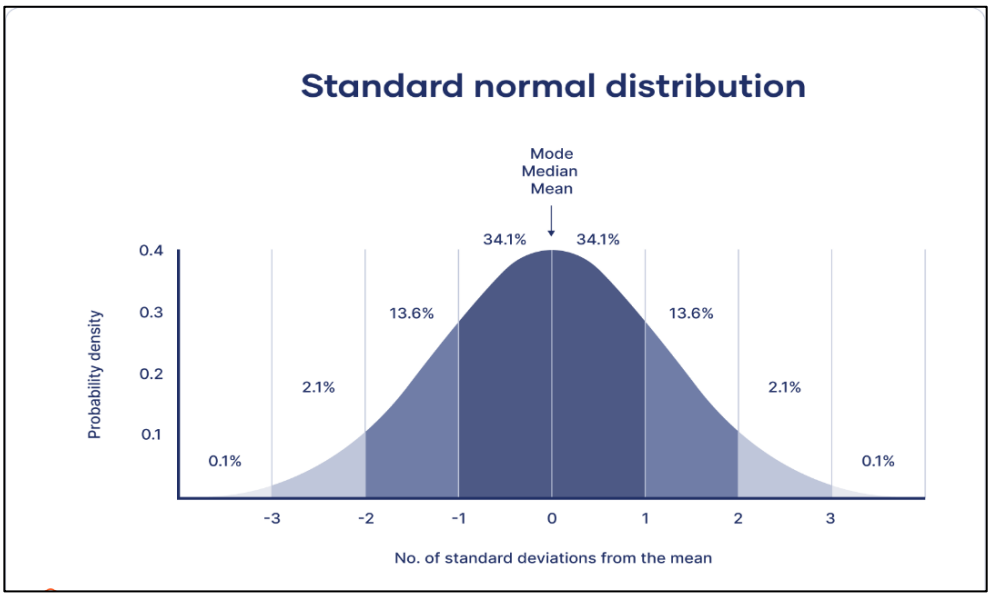

Figure 9.2: Normal Sampling Distribution curve
(Adapted from: Pritha Bhandari, 2020)

9.4 Stages of Sampling:

A. Identify the Target Population:

The first step in sampling process is to define the target of interest. The target population, in general includes all the different units from which the samples could be drawn. Depending on the subject and purpose of study, units of target population may include individual people, people of specific characteristics (like women of age above sixty), households or communities.

B. Establishing a Sample Frame:

The next step is to design a sample frame. A sample frame is the list of all the units from which a sample could be selected. The level of difficulties in designing a sample frame depends on the purpose and situation of a study. Sometimes it can be easier to develop a sample frame covering the entire target population, if the unit of analysis is organisation supported, campaigns run or events organised. At times it may be difficult when it is based on individuals or households living in a geographical region. In such cases, approximation is a must. Whenever possible, it is necessary to include the marginalised or vulnerable groups in the sample frame.

C. Selection of a Sampling Methodology:

The Sampling methods provides an outline about the way of sample selection. Sampling methods are mostly influenced by the objectives of the research, financial resources, time constrains and the nature of the study to be conducted. All sampling methods can be grouped under two categories- Probability sampling & Non- Probability Sampling, which will be discussed next in this article.

D. Determination of Sample size:

The selection of an appropriate sample size actually depends on the purpose of data collection, the type of analysis to be conducted and the questions that needs to be answered through the sampling [8]. Size of the sample has an important role in the sampling process. There are various ways of classifying the techniques used in determining the sample size. In non-probability sampling, sample size determination is based on the importance of decision, number of variables and completion rates. Formulas are used to determine the sample size in case of probability sampling.

E. Specifying the Sampling Plan:

This step includes the specification and illustration of the decision regarding the implementation of the research processes.

F. Selecting the Sample:

This is the step in the sampling process, where the actual selection of the sampling element is carried out.

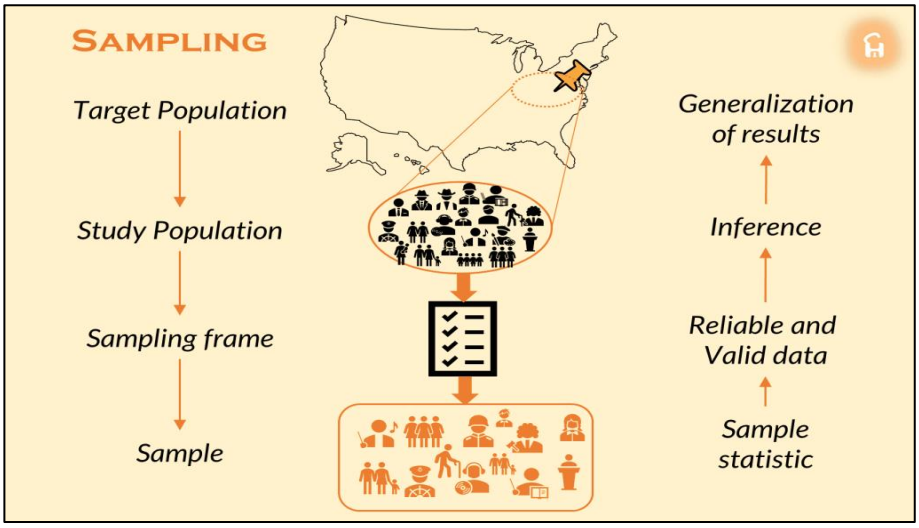

Figure 9.3: Different stages of Sampling
(Adapted from: Chitvan Trivedi, 2020)

9.5 Methods of Sampling:

In terms of research, a sample is a group of people, objects or items that are taken from a large population for measurement. To select an appropriate sample from such a large population for the purpose of research, sampling is done.

Sampling is an invaluable tool for the researcher, which enables them to collect the meaningful data as well as facilitates analysis to identify specific characteristics of the people. Based on the features of the population, purpose of study and the available resources, sampling can be broadly classified into two categories-

1. **Probability Sampling**
2. **Non- Probability Sampling**

9.5.1 Probability Sampling:

In this type of sampling, there is an equal chance of each member of the population of being selected in the sample [2,3]. Probability sampling is mostly done in quantitative research. Probability sampling is the most valid choice to obtain results of a sample that is representative of the whole population.

Probability sampling can be further classified into four categories:

A. Simple Random Sampling:

In simple random sampling, each member has an equal chance of being selected and each selection is independent of the others [7]. In this type of sampling, members of the population are selected randomly and the selection is entirely based on chance, for which it is also known as chance selection. This type of sampling is a great choice when the population is highly homogenous.

B. Stratified Sampling:

In this type of sampling, the population is divided into subgroups called as strata on the basis of similarities and then random samples are selected from each stratum in proportion to its size in the population [13]. The purpose of stratified sampling is to make a true representative sample and to address the issue of less homogenous population.

C. Systematic Sampling:

Systematic sampling is an advanced form of simple random sampling in which the units or elements are selected at regular intervals from an ordered list of the population.

The starting element in this type of sampling is chosen at random and then the rest of the elements are selected after a fixed interval. Therefore, it is considered to be more convenient than the simple random sampling [7].

D. Clustered Sampling:

In this type of sampling, various parts of the population are divided into clusters and members from each cluster are selected at random. Clustered sampling is distinct from stratified sampling in that it involves subgroups or clusters that possess characteristics similar to the whole sample.

All members can be included in the final sample when the cluster size is small and in case of larger cluster, individuals within each cluster may be sampled using simple random sampling technique [7].

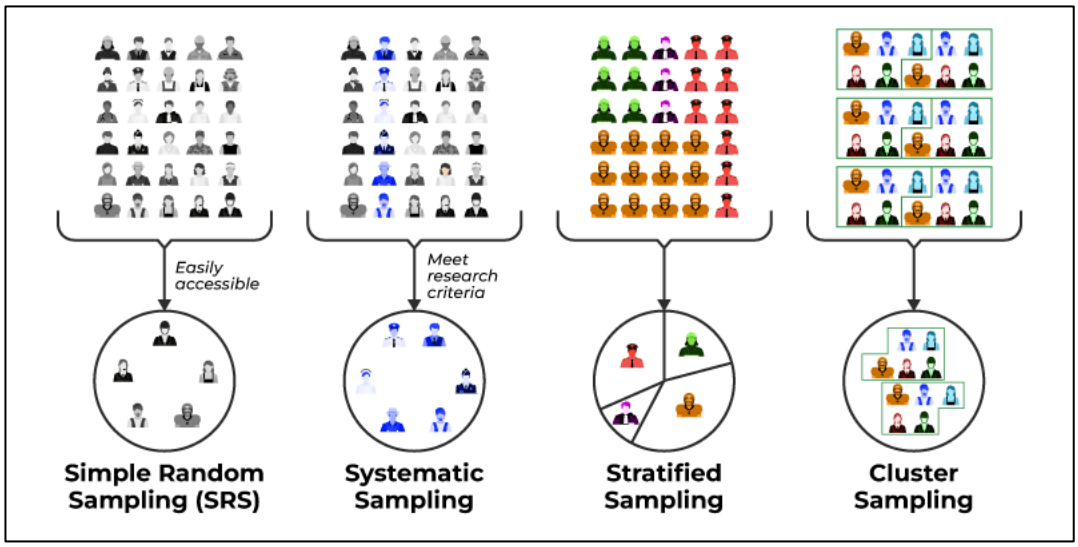

Figure 9.4: Types of Probability Sampling
(Adapted from: Chitvan Trivedi 2020)

9.5.2 Non-Probability Sampling:

Non-Probability sampling is based on the collection of data through a non-random selection based on a predefined criteria [3,4].

It is an easier and more economical method of sampling but at times it introduces sampling bias, resulting in weaker inferences compared to probability sampling technique in research. Non-probability sampling can be further classified into five categories-

A. Convenience Sampling:

Selecting the members of a sample directly from the population on the basis of proximity and accessibility to the researcher is known as convenience sampling. This is a simple, inexpensive and practical method of sampling based on the available data without any additional requirements [6]. However, there are high chances of sampling and selection bias in this type of sampling due to inappropriate population representation.

B. Consecutive Sampling:

Here, the researchers select the participants based on their availability and desire to participate in the study as they become available. This strategy helps in sequential recruitment of the individuals who fulfil the researcher's requirements.

C. Quota Sampling:

This kind of sampling allows the researcher to select the individuals non-randomly based on some specific characteristics which are representative of the population. On the basic of these characteristics, predetermined quota for specific subgroups are set and then the sample units are selected from these subgroups until the set quota is reached.

D. Purposive Sampling:

In purposive sampling the researcher selects a sample relevant to the purpose of study. This method of sampling is suitable for small size population and is commonly applicable in qualitative research, mainly to understand a particular phenomenon. [7]

E. Snowball Sampling:

This type of sampling is adopted when identification of the members of a population is challenging. Collection of samples in snowball sampling is based on a chain referral process, where one respondent aids in identifying the others [6]. These respondents share common features representative of the target population. This method is commonly applicable in qualitative research, during the study of a phenomenon related to a stigmatised or hidden populations.

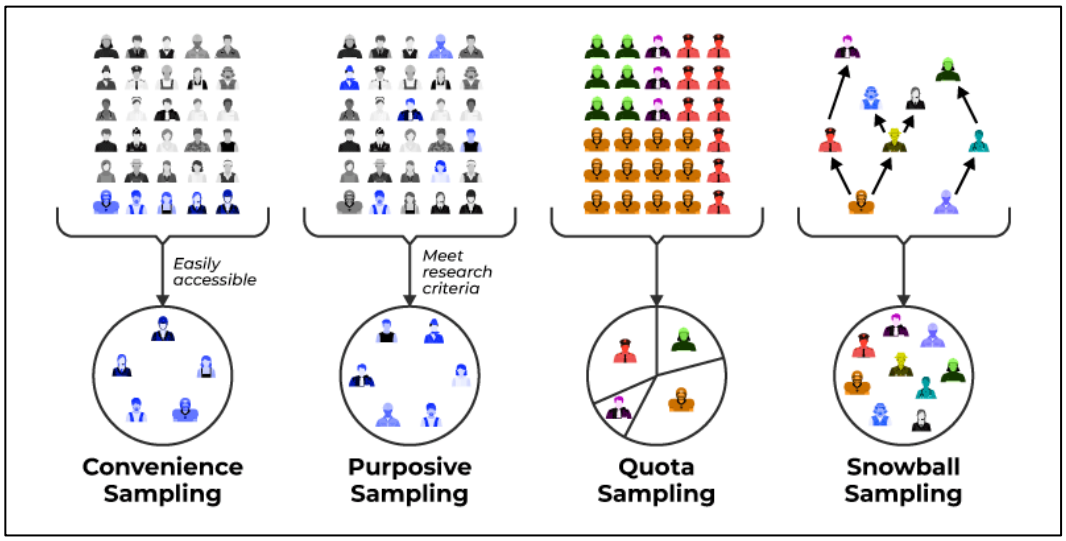

Figure 9.2: Types of Non-Probability Sampling
(Adapted from: Chitvan Trivedi, 2020)

9.6 Conclusion:

Sample is the representative of the population. The basic motive behind sampling is to analyse the features of the unit in the sample, deduce results from the study and to obtain a generalised conclusion about the population from where the sample was collected. Sampling can sometimes be an easier process from choosing a few community groups to visit during a field trip. However, at some other time it can be a complicated process and requires careful planning and expertise in statistics. In either case it is important to have a thorough understanding about the application of the analyses which will be used to obtain the results, before a sample is developed. It is also important in the part of a researcher to have a knowledge about the various strategies used in the selection of a sample from a large population. To choose the best sampling strategy, researchers must have a clear and detailed understanding about the study goal, characteristics of the target population and the resources at their disposal. Researcher can make use of sound sampling strategies to increase the validity of their findings and thereby advancing knowledge in a variety of domains [9].

9.6 References:

1. Cook, C.W., and J. Stubbendieck. (eds). 1986. Range research: Basic problems and techniques. Society for Range Management, Denver, CO. pp 215-220.
2. J. Wretman, Reflections on Probability vs non-probability sampling, official stat. Honour Daniel Thorburn. (2010) 29-35.
3. A. Shorten, C. Moorley, Selecting the Sample, Evid. Based. Nurs. 17 (2014) eb-2014-101747- doi: 10.1136/eb-2014-101747.
4. F. Gravetter, L. Forzano, Selecting Research Participants, Res. Methods Behav. Sci. (2012) 125-139.
5. P. Sampling, P. Guidelines, M.S. Choices, T. Oaks, CHAPTER 5 Choosing the Type of Sampling (2012) 125-174.
6. P. Bharadwaj. Types of Sampling in research. J Pract. Cardiovasc Sci 2019; 5: 157-63.
7. R. Bisht. What are Sampling Methods? Techniques, Types and Examples. R. Discovery (2023).
8. R. Popplewell (2013). A Rough Guide to Sampling. INTRAC, Oxford, 2013.
9. Makwana, P. Engineer, A. Dabhi, H. Chudasama. Sampling Methods in Research:
10. A Review. International Journal of Trend in Scientific Research and Development, ISSN:2456-6470, Vol-7, Issue 3 (2023). Pp.762-768.
11. M. Elfil, A. Negida. Sampling Methods in Clinical Research: an educational review, Emergency. 2017; 5
12. P. Sedgwick. What is Sampling error? BMJ 2012; 344: e4285 doi:10.1136/bmj. e4285 (2012).
13. A. Bhattacherjee. Social Science Research: Principles, Methods and Practices (Revised edition) (2019).
14. Z. Halukuz. M.A. Edu/ Research Methodology/ Sampling 1 (2021).