

Sampling

Introduction

Sampling is an essential part of any research investigation. Almost all research studies involve sampling. It is, therefore, essential that we understand the main concepts of sampling and are familiar with the sampling methods.

Research studies assume that the people selected for studies are representative of a large group about whom generalizations are to be made. We normally cannot survey everyone in the population; but through sampling techniques, we can be confident that only a small part of the total population can fairly represent the total population. Sampling, then, is a technique that saves the time and trouble of questioning 100 percent of the population.

What is Sample?

A sample is a collection of items or elements from a population or universe. Hence, a sample is only a portion or subset of the universe or population. It comprises some observations selected from the population. For instance, if 50 students are drawn from a population of 500 students of a college, these 50 students form the sample for the study.

Population or universe refers to the entire group of people, events, or things of interest that the researcher wishes to investigate. For example, if you are interested in investigating the smoking habits of employees in a chemical factory, then all employees in that factory will form the population.

Sometimes, the entire population will be sufficiently small, and you can include the entire population in your study. If the total items are studied, that is called a census study. However, it is not always possible to study every items or elements in a universe. Usually, the population is too large. Hence, a small, but carefully chosen sample can be used to represent the population. The sample thus selected reflects the characteristics of the population from which it is drawn. For thesis or project work to be undertaken by us, the study of the total population is neither possible nor necessary. Making a census study of the entire universes is not possible on account of limitations of time and money. Hence, sampling becomes inevitable.

Population may be finite or infinite. A finite population is one containing a fixed number of elements. An infinite population is one without limits of any kind and is therefore indeterminate.

Sample Design and Related Terminologies

Sampling design or strategy is the way in which you design your sample plan and select your samples from the population. In designing a sample, you must consider three things: sampling frame, selection of sampling items, and sample size. These terminologies can be explained as follows.

■ Sampling frame is the list identifying each unit in the study population. All the elements in a sampling population constitute its sampling frame. Thus, it may be all the students at university Campus, Kirtipur, all names in the telephone directory, or all persons having their bank accounts with the Nepal Bank Limited. After determining the sampling frame, the researcher will decide how sample will be selected.

■ *Sampling item or unit* is an element (person, institution, etc) of your study that becomes the basis for selecting your sample.

■ The *size of the sample* must be determined. What should be the sample size? Should fifty or eighty employees be interviewed? Though accuracy is greater with large samples, so are costs.

■ *Sample statistics* are the information obtained from the respondents selected for your study. Your sample statistics become the basis of estimating the prevalence of the characteristics in the study population.

■ *Population parameters* or population mean are the characteristics of the population estimated from the sample statistics. If you measure the entire population and calculate a value like a mean or average, this is called a population parameter.

Thus, the basic components of a sample design are: (a) choosing the sample units (who are to be surveyed), (b) choosing the sample size (how many to be surveyed), (c) choosing the sampling procedure (how to ensure that those who are to be interviewed are included in the sample), and (d) choosing the media (how to reach respondents in the sample? - through mail survey, personal interview, or telephone interview).

THE PRINCIPAL STEPS IN A SAMPLE SURVEY

The main steps involved in the planning and execution of a sample survey may be grouped somewhat arbitrarily under the following heads.

1. Objectives of the Survey.

The first step is to define in clear and concrete terms, the objectives of the survey. It is generally found that even the sponsoring agency is not quite clear in mind as to what it wants and how it is going to use the results. The sponsors of the survey should take care that these objectives are commensurate with the available resources in terms of money, manpower and the time limit required for the availability of the results of the survey.

2. Defining the Population to be sampled.

The population, i.e., the aggregate of objects (animate or in-animate) from which sample is chosen should be defined in clear and unambiguous terms. For example, in sampling of farms clear-cut rules must be framed to define a farm regarding shape, size., etc., keeping in mind the border-line cases so as to enable the investigator to decide in the field without much hesitation whether or not to include a given farm in the population.

3. The Frame and Sampling Units.

The population must be capable of division into what are called sampling units for purposes of sample selection. The sampling units must cover the entire population and they must be distinct, unambiguous and non-overlapping in the sense that every element of the population belongs to one and only one sampling unit. For example, in socio-economic survey for selecting people in a town, the sampling unit might be an individual person, a family, a household or a block in a locality. In order to cover the population decided upon, there should be some list, map or other acceptable material, called the frame, which serves as a guide to the population to be covered. Only good experience person helps to construct a good frame.

4. Data to be collected.

The data should be collected keeping in view the objectives of the survey. The tendency should not be to collect too many data some of which are never subsequently examined and analyzed. A practical method is to chalk out an outline of the tables that the survey should produce. This would help in eliminating the collection of irrelevant information and ensure that no essential data are omitted.

5. The Questionnaire or Schedule. Having decided about the type of the data to be collected, the next important part of the sample survey is the construction of the questionnaire (to be filled in by the respondent) or schedule of enquiry (to be completed by the interviewer) which requires skill, special technique as well as familiarity with the subject-matter under study. The questions should be clear, brief, non-offending, courteous in tone, unambiguous and to the point so that not much scope of guessing is left on the part of the respondent or interviewer. Suitable and detailed instructions for filling up the questionnaire or schedule should also be prepared.

6. Method of collecting information

The two methods commonly employed for collecting data for human populations are : (i) Interview Method. In this method, the investigator goes from house to house and interviews the individuals personally. He asks the questions one by one and fills up the schedule on the basis of the information supplied by the individuals. (ii) Mailed Questionnaire Method. In this method, the questionnaire is mailed to the individuals who are required to fill it up and returns it duly completed. Whether the data should be collected by interview method or mail questionnaire method or by physical observation has to be decided keeping in view the costs involved and the accuracy aimed at. Although mail surveys are less costly, there is scope for considerable non-response. Moreover mail method is practicable only among the educated people who are really interested in the particular survey being conducted. On the other hand, interview method costs more and there are interviewer errors also but without investigators the data collected may be worthless.

7. Non-respondents.

Quite often (due to practical difficulties), the data cannot be collected for all the sampled units. For example, the selected respondent may not be available at his place when the investigator goes there or he may fail or even refuse to give certain information when contacted. This incompleteness, called non-response, obviously tends to change the results. Such cases of non-response should be handled with caution in order to draw unbiased and valid conclusions. Procedures will have to be devised to deal with those who do not furnish information. The reasons for non-response should be recorded by the investigator.

8. Selection of Proper Sampling Design.

The size of the sample (n), the procedure of selection and the estimation of the population parameters along with their margins of uncertainty are some of the important statistical problems that should receive the most careful attention. A number of designs (plans) for the selection of a sample are available and a judicious selection will guarantee good and reliable estimates. For each sampling plan, rough estimates of sample size n can be obtained for a desired degree of precision. The relative costs and time involved should also be considered before making a final selection of the sampling plan.

9. Organization of Field Work. It is absolutely essential that the personnel should be thoroughly trained in locating the sample units, recording the measurements, the methods of collection of required data before starting the field work. The success of a survey to a great extent depends upon the reliable field work. It is very necessary to make provisions for adequate supervisory staff for inspection after field work. From practical point of view a small pretest, (i.e., trying out the questionnaire and field methods on a small scale) has been found to be immensely useful. It always helps to decide upon effective method of asking questions and results in the improvement of the questionnaire. Moreover, it might disclose certain problems and troubles that will otherwise be quite serious on a large-scale survey such as "the cost and the time may far exceed the available money and stipulated period."

10 Summary and Analysis of the Data. The analysis of the data may be broadly classified into the following heads:

a) Scrutiny and editing of the data: An initial quality check should be carried out by the supervisory staffs while the investigators are in the field. Accordingly, the schedule should be thoroughly scrutinized to examine the plausibility and consistency of the data obtained.

(b) Tabulation of data: Before carrying out the tabulation of the data, we must decide about the procedure for tabulation of the data which are incomplete due to non-response to certain items in the questionnaire and where certain questions are deleted in editing process. The method of tabulation, viz., hand tabulation or machine tabulation, will depend upon the quantity of the data. For large-scale survey, machine tabulation will obviously be much quicker and economical. For

a large-scale sample survey, the use of code numbers for qualitative variables is essential for machine tabulation. With simple questionnaire, the answers can sometimes be pre-coded, i.e., entered in a manner in which they can be conveniently or routinely transferred to mechanical equipment such as personal computers, etc. Finally, the tables that lead to the estimates are prepared.

(c) Statistical analysis. After the data has been properly scrutinized, edited and tabulated, a very careful statistical analysis is to be made. Different methods of estimation may be available for the same data. Appropriate formulae should then be used to provide final estimates of the required information. Efforts should be made to keep the procedure free from errors.

(d) Reporting and conclusions. Finally, a report incorporating detailed statement of the different stages of the survey should be prepared. In the presentation of the results, it is good practice to report the technical aspect of the design, viz., the types of the estimators used along with the amount of error to be expected in the most important estimate.

11. Information gained for Future Surveys. Any completed survey is helpful in providing a note of caution and taking lessons from it for designing future surveys. The information gained from any completed sample in the form of the data regarding the means, standard deviations and the nature of the variability of the principal measurements together with the cost involved in obtaining the data serves as a potential guide for improved together sampling.