1. Introduction

Welcome to the series of E-learning modules on Introduction to Sampling Techniques. In this module, we are going to cover the concept of sampling design, various techniques of survey, need for sampling and its objectives, process of sampling design and various types of sampling techniques.

By the end of this session, you will be able to:

- Explain the sample design
- Explain the various methods of survey and the implications of the sample design
- Explain the need, objectives and principles of sampling
- Explain the process of sampling design
- Explain the essentials of a good sampling design
- Explain the various types of sampling techniques

The science of statistics may be broadly studied under the following two headings:

- Descriptive
- Inductive

In Descriptive Statistics, we have been considering the statistical treatment of the available data, which was available to the statistician in the secondary form. For example, Data on sales, production etc. of a company are available in company's yearbook, data on population is available in Government survey given in published form, and data on student's progress is available in school or college records and so on.

For a problem under study, the statistician has to carry out the collection of data, which is referred to as primary data. This collection of primary data may be done either by census method or by sampling method.

The inductive statistics is also known as statistical inference. This may be termed as the logic of drawing statistically valid conclusions about the population based on examining the part of the population, which is drawn from the population in scientific manner.

In modern decision-making process, the different fields of human activity - including ordinary actions of our daily life, most of our decisions and attitudes depend very much upon the inspection or examination of only a few objects or items out of the total lot. This process of studying only a sample data and then generalizing the results to the population involves an element of risk. The risk is of making wrong decisions unless the samples are selected with utmost care.

2. Area of Application

In the present era of technological civilization, the government of a country requires detailed information about the requirements of the people, the production in the country, the resources available etc. to frame its policies in all fields. Such data is also very important for industrial organization to plan their production, expansion of activities, sales organization etc. The collection of data is a major activity in the life of a nation.

With large population and multiplication of activities in almost all fields the collection of information from each unit of a population say from each farm, each household is simply out of question. The financial outlay and time required for such a program will be very disproportionate to the utility of the final result when available. The only course open is to collect data from a small part of the population and obtain information about the population based on this data.

We can collect information from the universe either by enumerating each and every unit of the universe called as census survey or by enumerating a representative portion of the population known as Sample Survey. Hence, in a census survey all the units are taken into account whereas in a sample survey, only few selected items are studied. Based on such studies, attempts are made to draw generalization, which may be applied to the whole data.

For example, let us find out the average monthly expenditure of the two thousand students studying in the hostels of a University. If we hold a census investigation, we shall have to study the monthly expenditure of each one of these two thousand students. However, if we hold a sample investigation we shall select say two hundred students out of two thousand students and then study their expenditure.

The study of these two hundred units is technically called as a “sample”. We can draw conclusions, which will hold good for the expenditure of all the two thousand students (technically called a “universe” or a “population”). The sample is considered as a representative of the universe if the sample has been properly selected and if the size is all right. Whatever holds good for the sample will also hold good for the population.

If the scope of the enquiry is very wide, the census enquiry will not only be very expensive but highly cumbersome also. In such cases, sample investigation is very suitable. However, it needs to be emphasized that when the population is a small one, it is of no use resorting to a sample survey. When field studies are undertaken in practical life, consideration of time and cost almost invariably leads to a selection of respondents that is selection of only a few items.

The respondents selected must be a representative of the total population as possible in order to produce a miniature cross section. The selected respondents constitute a “sample” and the selection process is called “sampling technique”. The survey conducted is known as “sample survey”.

Let us discuss the implications of a sample design.
A sample design is a definite plan for obtaining a sample from a given population. It refers to the technique or a procedure an investigator would adopt in selecting the item for the sample

that is, size of the sample. A sample design is determined before data are collected. There are many sample designs or techniques and the investigator can select the appropriate one. Some designs are relatively more precise and easy to apply than others are. One must select a sample design, which should be reliable and appropriate for their study.

Let us identify the need for sampling.
Sampling is used for various of reasons such as:

- Sampling can save time and money
- Sampling may enable measurements that are more accurate for a sample study, which is generally conducted by trained and experienced investigators
- Sampling remains the only way when the population consists of infinitely many members
- Sampling remains the only choice when the test involves the destruction of the item under study
- Sampling usually enables to estimate sampling errors

Let us discuss the characteristics of a good sampling design.

- Sample design must result in a truly representative sample
- Sample design must be such which results in small sampling errors
- Sample design must be viable in the context of funds available for the research study
- Sample design must be so that a systematic bias can be controlled in a better way
- Sample should be such that the results of the sample study can be applied for the population with a reasonable level of confidence

3. Principal Steps

Let us list the principal steps of Sampling Design Process.
The sampling design process may be grouped under the following heads:

- Define the population to be sampled with the objectives
- Determine of the frame and the sampling units
- Determine the sampling procedure
- Determine the appropriate sample size
- Execute sampling design

Define the target population to be sampled with objectives of the survey
First, define in clear words about the objectives of the survey. Sometimes the sponsoring agents will also not know about the objectives of the survey like what it wants and how it is going to make use of the results. However, they should take utmost care so that their objectives should go along with the available resource, interims of money, manpower and the time limit required for the completion of the survey. Next decide whom do you want to survey? What are the characteristics of those who have the information?

The main step in developing any sampling design is to clearly define the set of objects from which the sample is chosen technically called target Population. For example, in sampling of farms, clear rules must be framed regarding the size, shape etc. of the farm giving importance for the borderline cases. This will help the investigator to decide in the field without any hesitation whether to include or not to include a given farm in the population.

Determine the Sampling frame:
Frame can be defined as some list or map or other acceptable material that covers the population decided upon and which serves as a guide for the population to be covered. For example:

- Students who eat Samosas
- Young people at random in the street
- Professional Association's members list
- Students' enrolment list
- College mailing list

Let us list the problems with list or frame.
Since frame is the one, which determines the structure of the sample survey it has become one of the major practical problems. A frame, which has been already prepared for some other purpose, has to be scrutinized and should be checked to see that it is free from all sorts of defects. The various sorts of defects are like unknown amount of duplication, ineligibles, omissions etc. and should be brought up-to-date before using them. A good experience helps in constructing a good frame.

Selecting the sampling Design:
The most important part is the selection of the samples. A sample study would give dependable conclusions only if the sample is the true representative of the universe. There are three methods by which the samples can be selected and they are:

- Probability (Random) sampling
- Non-probability (Deliberate) sampling
- Mixed sampling

Some of the probability sampling techniques in which each unit has an equal chance of being included in the sample are:

- Simple random sampling
- Systematic sampling
- Stratified sampling
- Cluster sampling

Non-probability (non-random) sampling is one in which each unit has an unequal chance of being included in the sample.
Convenience sampling
Judgement sampling
Snowball sampling
Quota sampling

Probability sampling:
In random sampling, selection of the sample is done in such a manner that the chance of selection of each unit of the universe is the same. In other words, the selection of the units depends entirely on chance.

Random selection of the samples has many advantages over the non-random selection.
The most important merit of this technique is the theory of probability. It is possible to have an idea about the errors of estimation and assign limits within which the true value of the measure of universe must invariably lie.
Another point in favour of this technique is that the selection is not affected by the prejudice or bias of the investigator.

4. Limitations

Limitations of Probability Sampling:
It should always be kept in mind that in many cases it is difficult to say that the selection has been purely random and is more expensive than non-probability samples. This random sampling technique is purely based on the theory of probability.

Non-probability (Non -Random) sampling:
In Non-probability or Non–Random or Deliberate sampling the investigator himself chooses few such units from the population. His selection is based on his own ideas about the representativeness of the sampled units. Hence, this is a subjective procedure in which, the probability of selection for some population units is zero or unknown before drawing the sample.

The advantages of non-probability sampling are:

- Cheaper and faster than Probability Sampling
- Reasonably representative, if collected in a thorough manner

Non-probability technique of selection has many drawbacks.
The first and foremost is that the bias or prejudice of the investigator has enough scope to work and influence the selection.
Another defect of this technique is that it is not possible to have any idea about the degree of accuracy achieved in any statistical investigation conducted by this method.

Mixed Sampling:
In mixed sampling, there is a mixture of random sampling and non-random sampling. The universe is first divided into same groups based on non-random sampling and from each subdivision, certain items are selected in accordance with random sampling.

Determine the Appropriate sample size:
The data is to be collected by keeping in mind the objectives of the survey. One should not collect too many data, which are never examined and analyzed. This refers to the number of items to be selected from the universe to constitute a sample, which is one of the major problems in sample design. The size of the sample should be neither excessively large nor too small. It should be optimum. An optimum sample is the one, which fulfils the requirements of efficiency, representativeness, reliability and flexibility. The size and the parameters of the population should also be kept in mind while deciding the size of the sample. In practical situation, one can chalk out an outline of the table, which would help us in eliminating the collection of irrelevant information.

Execute Sampling design:
The procedure for selection of the sample and estimation of population parameters along with the margin of error are some of the important statistical problems that should receive most of our attention. Finally, the investigator should decide the type of the sample he/she will use that is he must decide about the technique to be used in selecting the items for the sample.
In fact, this procedure or technique stands for sample design itself. There are several sample designs out of which an investigator has to select judiciously which will guarantee reliable

estimates. Obviously he must select the design which for a given sample size and cost has a smaller sampling error.

Based on the above discussion, it can be concluded that the various steps in the process of sampling are interwoven.

5. Essentials of Sampling

Let us list the Essentials of Sampling.
For the sampling technique to be useful, it is necessary that samples are drawn using one of the sampling techniques, which possess few essentials. They are

- Adequacy
- Independent selection
- Homogeneity
- Representative

The theory of sampling is based on the following important principles.
1. Principle of Statistical regularity:

This principle is related to the mathematical theory of probability. According to King, "The law of statistical regularity lays down that a moderately large number of items chosen at random from a large group are almost sure on the average to possess the characteristics of the large group". This principle focuses the importance of selecting a sample at random so that each and every unit in the population has an equal chance of being selected in the sample. In other words, selection must not be made by deliberate exercise of one's direction. A sample selected in this manner would be representative of the population.

An immediate derivation from the principle of Statistical regularity is the principle of inertia of large numbers which states that "other things being equal, as the sample size increases the results tend to be more reliable and accurate". This is because when we deal with large numbers the variations in the aggregate result is likely to be insignificant (Because variations in the component parts tend to balance each other).

2. Principle of Validity:
Sample design should enable us to obtain valid tests and estimates about the parameters of the population. The sample obtained by the method of probability sampling satisfies this principle.

3. Principle of Optimization:
This principle impresses upon obtaining optimum results in terms of efficiency and cost of the design with the resources at our disposal. The reciprocal of the sampling variance of an estimate provides a measure of efficiency while a measure of the cost of the design is provided by the total expenses incurred in terms of money and manpower. The principle of optimization aims at

- Achieving the given level of efficiency at a minimum cost
- Obtaining maximum possible efficiency with the given level of cost

We can generalize the results of the sample only to the population from which we sampled and we cannot cross the limits of the target population.

- School students not college students - That is if we have obtained the results for the information collected from the school students only, we cannot generalize this result to the college students.

- College students not Indian population - younger, poorer, etc.
- Indians not people from other countries - less traditional, more affluent, etc.

Although the theory of sampling has been developed recently, the practice of using sampling methods is very old. Sampling in Statistics is as common and important as salt in food.
The origin of Statistics is as old as our civilization. From time immemorial, homemaker's see a sample of cooked rice to ensure all the rice grains in the pot are cooked or not. Also we take a handful grains to ensure all of it in the bag is good.

Similarly, a businessperson places an order for the material by just examining the small sample; doctor takes few drops of blood to check for any disease or deficiency. Therefore, in every field sampling is used consciously or unconsciously.

Hence, one has to select a sample design, which is reliable and appropriate for their study.

Here's a summary of our learning in this session, where we have discussed:

- The concept of sample design
- The various types of survey and its implications
- The need for sampling and its objectives
- The essentials of a good sampling design
- The sample design process
- The various techniques of sampling
- The Principles of sampling
- The caution for generalization