

Unit III

SAMPLING: DEFINITION

Sampling is the process of obtaining information regarding the entire research population or aggregate or totality by examining only a part of it. It is a picture in miniature to represent the larger whole, to enable the researcher to make a judgment or inference about the whole with the presumption that the sample data will provide valid and reliable conclusions.

Census and Sample method

Under the census or complete enumeration survey method, data are collected for each and every unit (person, household, field, shop, factory, etc.) as the case may be of the population which is the complete set of items which are of interest in any particular situation. For example, if the average wage of workers working in sugar industry in India is to be calculated, then wage figures would be obtained from each and every worker working in the sugar industry and by dividing the total wages which all these workers receive by the number of workers working in sugar industry, we would get the figure of average wage.

However, despite the advantages the census method is not very popularly used in practice. The effort, money and time required for carrying out complete enumeration will generally be very large and in many cases cost may be so prohibitive that the very idea of collecting information may have to be dropped. This is more true of underdeveloped countries where resources constitute a big constraint. Also if the population is infinite or the evaluation process destroys the population unit, the method cannot be adopted.

Sampling is simply the process of learning about the population on the basis of a sample drawn from it. Thus, in the sampling technique instead of every unit of the universe is studied and the conclusions are drawn on that basis for the entire universe. A sample is a subset of population units.

(B) Advantages and Limitations of Sampling:

Merits

1. It saves time, because fewer items are collected and processed. When the results are urgently required, this method is very helpful.
2. It reduces cost only a few and selected items are studied in sampling. So there is reduction in cost of money and reduction in terms of man-hours. It is advantageous to underdeveloped countries.
3. More reliable results can be obtained because (a) there are fewer chances of sampling statistical errors. If there is sampling error, it is possible to estimate and control the results. (b) Highly expert and trained persons can be employed for scientific processing and analysing of relatively limited data, and they can use their high technical knowledge and get more accurate and reliable results.
4. Sampling provides more detailed information: As it saves time, money and energy, we can collect more detailed information in a sample survey.

5. Sampling method is sometimes the only method possible. If the population under study is infinite, sampling method is the only method to be used. For example, to test the breaking strength of bricks manufactured in a factory, under census method, all the bricks would be broken in the process of testing. There would be no bricks left for use. Thus, census method is impracticable. Also if the population under investigation is infinite, sampling is the only possible solution.
6. Administrative convenience: The organisation and administration of sample survey are easy.
7. More scientific: the method has full justification for the expenditure involved.
8. The degree of accuracy obtainable in this method is higher than that in the census method.

Shortcomings

1. Illusory conclusion: If a sample enquiry is not carefully planned and executed, the conclusions may be inaccurate and misleading.
2. Sample not representative: To make a representative sample is taken from the universe; the result is applicable to the whole population. If the sample is not representative of the universe, the result may be false and misleading.
3. Lack of experts: As there is lack of experts to plan and conduct a sample survey, its execution and analysis, the results of the sample survey are not satisfactory and trustworthy.
4. Sometimes the sampling plan may be complicated and requires more money, labour, time than a census method.
5. There is organisational problem in sample investigation.
6. Personal bias: There may be personal biases and prejudices with regard to the choice of technique and drawing of sampling units.
7. Choice of sample sizes: If the size of the sample is not appropriate then it may lead to untrue characteristics of the population.
8. Conditions of complete coverage: If the information is required for each and every item of the universe, then a complete enumeration survey is better.

Types of Sampling

From the practical point of view the sample designs are basically of two types:

- (1) Probability Sampling.
- (2) Non-probability sampling

Probability Sampling:

It provides a scientific technique of drawing samples from population in accordance with certain laws or chance in which each unit in the universe has some definite pre-assigned probability of being selected in the sample. That is why it is called random sampling.

1. Random Sampling (Probability sampling)

A random sample is one where each item in the universe has an equal chance of known opportunity of being selected. According to Harper, "A random sample is a sample selected in such a way that every item in the population has an equal chance of being included."

A. Simple random sampling:

It is a technique in which sample is so drawn that each and every unit in the population has an equal and independent chance of being included in the sample. Several methods have been adopted for random selection of the sample. They are:

(i) **Lottery Method:** This is the most popular and simplest method. In this method, all the items of the universe are numbered on separate slips of paper of same size, shape and colour. They are folded and mixed up in a drum or container. A blindfold selection is made. The required number of slips are selected for the desired sample size. The selection of items thus depends on chance. For example, if we want to select 5 students, out of 50 students, then we must write the names of all the 50 students on slips of the same size and mix them; then we make a blindfold selection of 5 students. This method is also called unrestricted random sampling, because units are selected from the unrestricted random sampling, because units are selected from the population without any restriction. This method is mostly used in lottery draws. If the universe is infinite, this method is inapplicable. There is a lot of possibility of personal prejudice if the size and shape of the slips are not identical.

(ii) **Table of Random Numbers:** As the lottery method cannot be used, when the population is infinite, the alternative method is that of using the table of random numbers.

There are several standard tables of random numbers. But the credit for this technique goes to Prof. L.H.C. Tippett (1927). The random number table (taken from the British Census Report) consists of 10,400 four-figured numbers giving in all $10,400 \times 4 = 41,600$. There are various other random numbers. They are Fisher and Yates (1938) comprising of 15,000 digits arranged in twos, Kendall and B.B. Smith (1939) consisting of 1,00,000 digits grouped in 25,000 sets of 4 digit random numbers, Rand Corporation (1955) consisting of 2,00,000 random numbers of 5 digits each, etc.

Merits

1. Scientific method: there is less chance for personal bias.
2. More representative: when the size of the sample increases, it is representative of the population, as the Law of Inertia of large Numbers and the Law of Statistical Regularity beings to operate.
3. Sampling error can be measured.
4. Theory of probability is inapplicable, if a sample is random.
5. this method is economical as it saves time, money and labour

Demerits

1. This requires a complete list of the population but such up-to-date lists are not available in many enquiries.
2. If the size of the sample is small, then it will not be a representative of the population.
3. When the distribution between items is very large, this method cannot be used.

B. Restricted Random Sampling:

(i) **Stratified sampling:** When the population is heterogeneous or of different segments of strata with respect to the variable or characteristic under study, then it is stratified. First the population is divided into a number of sub-groups or strata. Each stratum is homogeneous. A sample is drawn from each stratum at random.

There are two types of stratified random sampling. They are proportional and non-proportional. In the proportional sampling, equal and proportionate representation is given to sub-groups or strata. If the number of items is large in the population, the same will have a higher size and vice versa.

In disproportionate or non-proportionate sample, equal representation is given to all the sub-strata regardless of their existence in the population.

Merits

1. It is more representative.
2. It ensures greater accuracy.
3. It is easy to administer as the universe is sub-divided.
4. Greater geographical concentration reduces the time and expenses.
5. When the original population is badly skewed, this method is an appropriate one.
6. For non-homogeneous population, it may yield more reliable results.

Demerits

1. To divide the population into homogeneous strata, it requires more money, time and statistical experience which is a difficult one.
2. If proper stratification is not done, the sample will have an effect of bias. If different strata of population overlap, such a sample will not be a representative one.

(ii) Systematic sampling: It is also known as quasi-random sampling. A systematic sample is selected at random sampling. When a complete list of the population is available, this method is used. We arrange the items in numerical, alphabetical, geographical or any other order. If we want to select a sample of 10 students from 100 students, under this method Kth item is picked up from the sample frame and K is the sample interval.

K = Sampling interval

N = Size of universe

n = Sample size

in the above example $k=10$. 10 is the sampling interval. Every 10th student will be taken as sample, i.e., 10th 20th, 30th, and so on.

Merits

This is simple and convenient. the time and work is reduced much. If we take care, the result will be a satisfactory one. It can also be used in infinite population.

Demerits

It may not represent the whole population. There is the element of personal bias of investigators.

(iii) Cluster sampling or multistage sampling: It is also called as sampling stages. It refers to a sampling procedure, which is carried out in several stages. The whole population is divided into sampling units, and these units are again divided into sub-units. This process will continue when we reach a lease number.

For example, we want to take 5000 students from Madhya Pradesh. We want to take 5000 students from Madhya Pradesh. We must take universities at the first stage, then the number of colleges at the second stage, selection of students from the colleges at the third stage etc.

Merits

1. It introduces flexibility in the sampling method.
2. It is helpful in large -scale survey where the preparation of list is difficult, time-consuming or expensive.
3. It is valuable in underdeveloped countries, where no detailed and accurate framework is available.

Demerits

It is less accurate than other methods.

2. Non-random sampling method (Non Probability Sampling)

(a) Purposive or Judgment sampling: The investigator has the power to select or reject any item in an investigation. The choice of sample items depends on the judgments of the investigator. He has the vital role to play in collecting the information. For example, if a sample of 5 students are to be selected from a B.Com. Class of 50 students for analyzing the habit of picture-seeing, the investigator would select 5 students who, according to his opinion are the representative of the class.

Merits

1. It is a simple method.
2. It is used to obtain a more representative sample.
3. It is very helpful to make public policies, decisions, etc. The executives and public officials use this method for their urgent problem.

Demerits

1. Due to individual bias the sample may not be a representative one.
2. It is difficult to get correct sampling errors.
3. The estimates are not accurate.
4. Its results cannot be compared with other sampling studies.

(b) Quota sampling: This sampling is similar to stratified sampling. It is used in the U.S.A. for investigating public opinion and consumer research. To collect data, the universe is divided into quota according to some characteristics. Each enumerator is then told to interview a certain number of persons who are his quota. The selection of sample items depends on personal judgment. It is a stratified-cum-purposive sampling and thus has the advantages of both the methods. There is saving of time and money. If there are trained investigators, the sampling will give quite reliable results.

Personal prejudice and individual bias are there. It is not based on random sampling, and so sampling error cannot be estimated.

(c) Convenience or Chunk Sampling: Chunk is a convenient slice of a population which is commonly referred to as a sample. It is obtained by selecting convenient population units.

1. It is suitable when the universe is not clearly defined.
2. Sample unit is not clear.
3. Complete source list is not available.

A sample obtained from automobile registration, telephone directories, etc., is a convenience sample. The results of this sampling cannot be representative. They are unsatisfactory. They are biased. But they are used for pilot studies.

(d) Snow-ball Sampling: In sociology and statistics research, snow-ball sampling (or chain sampling, chain-referral sampling, referral sampling) is a non-probability sampling technique where existing study subjects recruit future subjects from among their acquaintances. Thus the sample group appears to grow like a rolling snowball. As the sample builds up, enough data is gathered to be useful for research. This sampling technique is often used in hidden populations which are difficult for researchers to access; example populations would be drug users or sex workers. As sample members are not selected from a sampling frame, snowball

samples are subject to numerous biases. For example, people who have many friends are more likely to be recruited into the sample.

What is snowball sampling? Snowball sampling uses a small pool of initial informants to nominate, through their social networks, other participants who meet the eligibility criteria and could potentially contribute to a specific study. The term "snowball sampling" reflects an analogy to a snowball increasing in size as it rolls downhill

Advantages

1. Locate hidden populations: It is possible for the surveyors to include people in the survey that they would not have known.
2. Locating people of a specific population: There is no lists or other obvious sources for locating members of the population of specific interest.

Disadvantages

1. Community Bias: The first participants will have strong impact on the sample. Snowball sampling is inexact, and can produce varied and inaccurate results. The method is heavily reliant on the skill of the individual conducting the actual sampling, and that individual's ability to vertically network and find an appropriate sample. To be successful requires previous contacts within the target areas, and the ability to keep the information flow going throughout the target group.
2. Not Random: Snowball sampling contradicts many of the assumptions supporting conventional notions of random selection and representativeness. However, Social systems are beyond researcher's ability to recruit randomly. Snowball sampling is inevitable in social systems.
3. Vague Overall Sampling Size: There is no way to know the total size of the overall population.
4. Wrong Anchoring: Another disadvantage of snowball sampling is the lack of definite knowledge as to whether or not the sample is an accurate reading of the target population. By targeting only a few select people, it is not always indicative of the actual trends within the result group. Identifying the appropriate person to conduct the sampling, as well as locating the correct targets is a time consuming process which renders the benefits only slightly outweighing the costs.