SAMPLE AND SAMPLING DESIGN

CENSUS AND SAMPLE SURVEY All items in any field of inquiry constitute a 'Universe' or 'Population.' A complete enumeration of all items in the 'population' is known as a census inquiry. It can be presumed that in such an inquiry, when all items are covered, no element of chance is left and highest accuracy is obtained. But in practice this may not be true. Even the slightest element of bias in such an inquiry will get larger and larger as the number of observation increases. Moreover, there is no way of checking the element of bias or its extent except through a resurvey or use of sample checks. Besides, this type of inquiry involves a great deal of time, money and energy. Therefore, when the field of inquiry is large, this method becomes difficult to adopt because of the resources involved. At times, this method is practically beyond the reach of ordinary researchers. Perhaps, government is the only institution which can get the complete enumeration carried out. Even the government adopts this in very rare cases such as population census conducted once in a decade. Further, many a time it is not possible to examine every item in the population, and sometimes it is possible to obtain sufficiently accurate results by studying only a part of total population. In such cases there is no utility of census surveys. However, it needs to be emphasised that when the universe is a small one, it is no use resorting to a sample survey. When field studies are undertaken in practical life, considerations of time and cost almost invariably lead to a selection of respondents i.e., selection of only a few items. The respondents selected should be as representative of the total population as possible in order to produce a miniature cross-section. The selected respondents constitute what is technically called a 'sample' and the selection process is called 'sampling technique.' The

survey so conducted is known as 'sample survey'. Algebraically, let the population size be N and if a part of size n (which is < N) of this population is selected according to some rule for studying some characteristic of the population, the group consisting of these n units is known as 'sample'. Researcher must prepare a sample design for his study i.e., he must plan how a sample should be selected and of what size such a sample would be. 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 the procedure the researcher would adopt in selecting items for the sample.

Basic Terms:

Population: The entire group under study.

Sample: A subset of the population selected for the study.

Sampling: The process of selecting a representative group from a larger population.

Advantages and Limitations of Sampling:

Advantages:

1.Cost-effective and time-efficient.

2.Easier to manage and analyze.

3.Feasible for large populations.

Advantages of sampling over census It is difficult in most of the cases to take up census enumeration for lack of time money, trained personnel and other constraints.

1. Less Time: There is considerable saving in time and labour since we study only a portion of the population in a sample survey. The sampling results can be obtained more rapidly and the data can be analyzed much faster since relatively fewer data have to be collected and processed**.**

2. Less Cost: Sampling certainly results in reducing the cost of survey in terms of money and man- hours. Although the amount of expenses involved in collecting information is generally greater per unit in a sample than in a complete enumeration, the total cost of sample survey is expected to be much smaller than that of a complete census. Since in most of the cases funds are limited in research study, sampling helps in reducing the costs of data collection.

3. Greater Accuracy: The results of a sample survey are usually much more accurate and therefore, more reliable than those obtained from a compete census. The errors due to factors such as training of field workers, measuring and recording observations, location of units, bias due to interviewers, incompleteness due to non-response etc. will certainly be larger and therefore more serious in a census than in sample survey.

4. Greater Scope: Sample survey has usually greater scope than census survey. Some inquires may require highly trained personnel or specialized equipment for collection of data, thus making a census practically impossible, or even inconceivable. In a sample survey we may have greater coverage both in respect of the information collected and in respect of the geographical, demographic or other boundaries taken into account.

5. Census is impossible: There are many situations in which complete population survey is impossible and one has to resort to sample survey only. Some of such situations are:

(a) if the population is infinite or too large to be observed.

(b) if the population is hypothetical like the population of all throws that may be made with a coin.

(c) if the nature of experimentation is destructive to observe the desired information, like observing the life-length of electricity bulb, or sound of crackers etc.

6. Administrative Convenience: The organization and administration of sample surveys are easy as one can choose the units of the sample in such a way that their investigations are administratively convenient. For example, if we wish to choose five districts of Uttar Pradesh, we choose two near by districts from western U.P. and three adjacent districts from eastern U.P., instead of choosing five randomly from all over the U.P.

Limitations:

1. Risk of sampling bias.

2. Generalization limitations to the entire population.

3. Possible underrepresentation or overrepresentation of certain groups.

(a) Sample survey is not suitable if higher order accuracy is required.

(b) If the items of the sample are not selected without any bias, the conclusions may not be correct.

(c) The investigator's personal bias regarding the choice of units and drawing of sample may lead to false conclusions.

(d) Sample investigation method is not suitable if the information is required about each individual of the population.

(e) Sample survey is a specialized technique and every body cannot use it. Its use requires specialized knowledge and trained personnel.

Sampling Process:

STEPS IN SAMPLE DESIGN while developing a sampling design, the researcher must pay attention to the following points:

(i) **Type of universe**: The first step in developing any sample design is to clearly define the set of objects, technically called the Universe, to be studied. The universe can be finite or infinite. In finite universe the number of items is certain, but in case of an infinite universe the number of items is infinite, i.e., we cannot have any idea about the total number of items. The population of a city, the number of workers in a factory and the like are examples of finite universes, whereas the number of stars in the sky, listeners of a specific radio programme, throwing of a dice etc. are examples of infinite universes.

(ii) **Sampling unit:** A decision has to be taken concerning a sampling unit before selecting sample. Sampling unit may be a geographical one such as state, district, village, etc., or a construction unit such as house, flat, etc., or it may be a social unit such as family, club, school, etc., or it may be an individual. The researcher will have to decide one or more of such units that he has to select for his study.

(iii) **Source list**: It is also known as 'sampling frame' from which sample is to be drawn. It contains the names of all items of a universe (in case of finite universe only). If source list is not available, researcher has to prepare it. Such a list should be comprehensive, correct, reliable and appropriate. It is extremely important for the source list to be as representative of the population as possible.

(iv) **Size of sample**: This refers to the number of items to be selected from the universe to constitute a sample. This a major problem before a researcher. The size of sample should neither be excessively large, nor too small. It should be optimum. An optimum sample is one which fulfills the requirements of efficiency, representativeness, reliability and flexibility. While deciding the size of sample, researcher must determine the desired precision as also an acceptable confidence level for the estimate. The size of population variance needs to be considered as in

case of larger variance usually a bigger sample is needed. The size of population must be kept in view for this also limits the sample size. The parameters of interest in a research study must be kept in view, while deciding the size of the sample. Costs too dictate the size of sample that we can draw. As such, budgetary constraint must invariably be taken into consideration when we decide the sample size.

(v) **Parameters of interest**: In determining the sample design, one must consider the question of the specific population parameters which are of interest. For instance, we may be interested in estimating the proportion of persons with some characteristic in the population, or we may be interested in knowing some average or the other measure concerning the population.

(vi) **Budgetary constraint**: Cost considerations, from practical point of view, have a major impact upon decisions relating to not only the size of the sample but also to the type of sample. This fact can even lead to the use of a non-probability sample.

(vii) **Sampling procedure**: Finally, the researcher must decide the type of sample he will use i.e., he must decide about the technique to be used in selecting the items for the sample. In fact, this technique or procedure stands for the sample design itself. There are several sample designs (explained in the pages that follow) out of which the researcher must choose one for his study. Obviously, he must select that design which, for a given sample size and for a given cost, has a smaller sampling error.

Define Population: Clearly identify the target group.

Select Sampling Frame: The list from which the sample will be drawn.

Types of Sampling -There are a number of ways of drawing a sample from a population depending on the nature of data and type of enquiry. These sampling techniques can broadly be classified into two headings:

(a) Purposive sampling

(b) Random sampling

(**a) Purposive Sampling** This type of Sampling is used with a definite purpose in view. In this sampling method the investigator uses his discretion in the matter of selecting the items that are to be included in the sample. In this kind of selection, the sample entirely depends upon the judgment of the investigator and no formula or principle is followed. A wiser investigator will include those units in the sample, which he thinks are most representative of the population characteristic under study. For example, to calculate the per capita income of a city a wise investigator will choose only a few rich families, more of middle class people and mostly poor people by his own judgment, Non –Sampling Errors

1. Faulty definitions of objectives

2. Response- bias

3. Non- response bias

4. Errors due to interviewers

5. Errors in measurement and publication to make the sample true representative of the population of the city. This can also be misused if he selects a sample consisting of only rich people in order to show a high per capita income. This is the biggest limitation of this sampling method. This sampling technique is also known as Subjective sampling or Judgment sampling or Non-probability sampling. In general, this method is not recommended due to element of subjectiveness on the part of investigator. However, if the researcher is an expert and experienced in the art of sampling, he can carefully apply this technique and then, purposive sampling would provide reliable results.

(b) Random Sampling In order to eliminate the possibility of human prejudice interfering in the selection of a representative sample, the method of random selection has been devised. A random sampling is that in which every item of the population has an equal chance of being selected in the sample. The selection is entirely objective. There are various ways in which a random sample may be drawn. The following are commonly used:

1. Simple random sampling

2. Stratified random sampling 3. Systematic sampling

4. Cluster sampling

5. Area sampling

6. Sub- sampling or Multistage sampling

7. Double sampling & Multiphase sampling

Choose Sampling Method: Random, stratified, cluster, etc.

1. Simple random sampling

Simple random sampling is a method of sample selection in which every item of the population has an equal and independent opportunity of being selected in the sample. The selection does not have any personal bias of the investigator. Random selection should not be confused with haphazard selection. There is nothing haphazard about such selection. When we speak of a simple random sample, we use the word random interchangeably with probability not with haphazardness. Random sampling has sometimes been referred to as representative or proportional sampling. If the sample is chosen at random and if the number of cases in it is sufficiently large, it will represent all the groups in the population in approximately correct proportion. Simple random sampling may be with or without replacement, according as a unit selected is replaced or not replaced back into the population before next draw. There are two methods of selecting a simple random sample: (i) Lottery method (ii) Table of random numbers (i) Lottery method: Under this method all the items of the universe are represented on cards and a blindfold selection is made of these cards. The selection of n (sample size) cards may be made in n draws, one by one. The one by one selection may be with or without replacement. In order to get accurate results, it is necessary that the cards should be similar in size, shape, and thickness, and in all other respects. It is one of the most reliable methods of drawing a random sample. This will, however, not be practicable if the parent population is large. (ii) Tables of random numbers: In the recent past, use of random number tables has been frequent for drawing of a simple random sample. A table of random digits is simply a table of digits, which have been generated by a random process. These numbers have been put to all possible tests and have been found to be truly random. With the help of these numbers, the work of selecting random samples has become very simple. What one has to do is to take any table of random numbers and start using the table from any position either horizontally or vertically. But once having started, not a single number should be left out and the order should also not be disturbed.

2. Stratified random sampling

If the population is heterogeneous in nature, that is widely apart in terms of characteristics under study, a simple random sample will not be a true representative of the population. In such cases, the entire heterogeneous population is divided into a

number of homogeneous groups called strata or sub-populations. Each such stratum or sub-population is homogeneous with in itself. These sub-populations are non-overlapping and together they comprise the whole of the population. Then units are sampled at random from each of these strata. Generally, the number of units selected from each stratum is proportional to the number of units in that stratum in the population. The sample, which is the set of all the sampling units drawn from each stratum, is called a stratified random sample and the technique of drawing this sample is termed as stratified random sampling. For example, suppose a college has 10,000 students out of which 5000 are in Arts faculty, 3000 in Commerce and rest 2000 are in Science. If a random sample of size 1000 students is to be drawn for the purpose of calculating their average percentage marks, it is advisable to use stratified random sampling. In this case, a sample to be chosen from Arts faculty will have its size = (5000 / 10000) x 1000 = 500. Similarly, the sub- samples of sizes 300 and 200 will be drawn from Commerce and Science faculties, respectively.

3. Systematic sampling Systematic sampling is sometimes known as mixed sampling. It has the feature of randomness in it along with a fixed procedure of selection of units. This is a convenient method when complete list of sampling units is assumed to be readily available or can be prepared. Such a list is known as a sampling frame. This sampling scheme consists of selecting only the first unit at random and the rest are then automatically selected according to some predetermined pattern. Let us suppose that the population size is N and a sample of size n is to be drawn. Then, we calculate the sampling interval k = N / n. Now, a unit from 1 to k is selected. Then, every kth unit in the sampling frame is selected thereafter. For example, in a class of 100 students with roll numbers 1 to 100, suppose a sample of size 10 is to be drawn. In this case k = N / n = 10. First we choose a student from roll numbers 1 to 10. Suppose it comes out be 7. Then, after 7 every 10th student will be selected in the sample. The sample will then, consist of students with roll numbers 7,17,27,37,47,57,67,77,87,97.

4. Cluster sampling In a random sampling scheme, the population consists of distinct and identifiable units called sampling units. The smallest unit into which the population can be divided is called an element of the population. A group of such elements is known as a cluster. When the sampling unit is a cluster the procedure is called cluster sampling. It is frequently used in large scale studies, as it is comparatively a less expensive and convenient sample design. In this sample design a large population is divided into smaller groups called clusters. These clusters are non-overlapping and exhaustive. Then out of all clusters a few are chosen by simple random sampling and all the units (elements) in a chosen cluster are observed. Clusters are usually formed of neighbouring elements and therefore tend to have similar characteristics. As a simple rule, the number of elements in a cluster should be small and the number of clusters should be large. The number of clusters and their formation depends on the research objectives and the resources available for research.

5. **Area sampling** If the population or universe is represented by a geographical area and its segments are made to form clusters, the cluster sampling is called Area sampling. For example, to study some characteristics nationwide, India can be divided into small geographical regions, called clusters. These sub areas may be states, cities, districts, blocks or villages depending on the research objective.

6. Sub- sampling or Multistage sampling In cluster sampling, the whole population is divided into N clusters and then n clusters are chosen randomly. Then, all the elements in the selected cluster are enumerated. Instead of enumerating all the element in a cluster if we survey only a sample of units in each selected cluster it is known as sub-sampling or two stage sampling. In such sample design, since clusters are formed at first stage, they are known as first stage units or primary sampling units and the elements within a cluster are called second stage units. This procedure can be generalized to three or more stages and will be known as multistage sampling. For example, in a marketing survey districts can be used as first stage units, colonies can be called second stage units, and then household can be termed as the third stage units.

7. Double sampling & Multi-phase sampling There are situations in which it is useful to collect information on some auxiliary variable apart from observing the study variable. For example, if we are interested to know the performance of MBA students in research methodology course, the marks obtained in that paper will be the study variable. A related variable will be the background of the student in graduation classes like arts, science, and commerce. Therefore proportion of students belonging to arts, science, commerce and others will be observed in the first phase and depending on these proportions the second and main phase of the survey will be conducted i.e. obtaining their marks in research methodology paper. It is relatively cheaper and faster to collect data on the auxiliary variable. Therefore if such data are not readily available one conducts a large-scale fast survey to obtain information on auxiliary variable in the first phase of the sample design. Then, after this the data on study variable are collected by doing the main survey in the second phase. In the first phase of the survey only a part of the resources are used and most of the part of resources are spent on the main survey. This kind of sample design is known as double sampling or two-phase sampling. When the survey is conducted in three or more phases, it is known as multi-phase sampling.

Determine Sample Size: Based on statistical considerations.

Execute Sampling: Actually selecting the sample.

CHARACTERISTICS OF A GOOD SAMPLE DESIGN From what has been stated above, we can list down the characteristics of a good sample design as under:

(a) Sample design must result in a truly representative sample.

(b) Sample design must be such which results in a small sampling error.

(c) Sample design must be viable in the context of funds available for the research study.

(d) Sample design must be such so that systematic bias can be controlled in a better way.

(e) Sample should be such that the results of the sample study can be applied, in general, for the universe with a reasonable level of confidence.

Types of Sampling:

Random Sampling: Each member of the population has an equal chance of being selected.

Stratified Sampling: Dividing the population into subgroups and then randomly sampling from each subgroup.

Cluster Sampling: Dividing the population into clusters and randomly selecting entire clusters.

Systematic Sampling: Selecting every nth individual from the population.

Types of Sample Designs:

Cross-Sectional Design: Data collected at a single point in time.

Longitudinal Design: Data collected over an extended period.

Testing of Hypothesis:

Hypothesis: A testable statement or assumption about a population parameter.

Null Hypothesis (H0): The hypothesis to be tested and usually states no effect or no difference.

Alternative Hypothesis (H1): The hypothesis that contradicts the null hypothesis.

Determining the Sample Size:

Power Analysis: Estimating the sample size needed to detect a significant effect.

Margin of Error: The acceptable range of deviation from the true value.

Sampling Distribution of the Mean:

Central Limit Theorem: The distribution of sample means approaches a normal distribution as the sample size increases, regardless of the shape of the population distribution.

Scaling Techniques:

Concept of Attitude:

Attitude: A psychological tendency that reflects an individual's feelings, beliefs, or behavioral intentions towards a particular object, person, or situation.

Difficulty of Attitude Measurement:

1. Subjectivity and variability in interpreting responses.

2.Influence of social desirability bias.

Types of Scale:

1.Nominal Scale: Categorizes data without any order or ranking.

2.Ordinal Scale: Orders data, but the intervals between scale points are not equal.

3.Interval Scale: Orders data with equal intervals between scale points, but lacks a true zero point.

4.Ratio Scale: Has equal intervals between scale points and a true zero point.

Criteria for Good Test:

Reliability:

Description: Consistency of the measurement instrument.

Statistics: Intra-class correlation coefficient or Cronbach's alpha assess reliability.

Validity:

Description: The extent to which the test measures what it claims to measure.

Statistics: Correlation coefficients with established measures can assess validity.