Chapter 4

Sampling

Chapter Index

Learning Objectives

- Discuss the concept of sampling

- Explain the errors in measurement and sampling errors

- Describe the non-sampling errors

- Explain the different types of sampling methods

Concept of Sampling

Different Definitions of Sampling

Sampling is the process in which a sufficient number of elements are selected from a population for conducting a study. The characteristics and properties of the selected sample are generalised for the whole population.

Sampling is the selection of certain percentage of a group of items according to a predetermined plan.

Sampling method is the process or method of drawing a definite number of individuals, cases or observations from a particular universe, selecting part of a total group for investigation.

The term sample should be reserved for a set of units or portion of an aggregate of material which has been selected in the belief that it will be a representative of the whole aggregate.

Sampling is a small piece of the population obtained by a probability process that mirrors with known precision the various patterns and subclasses of the population.

Census Versus Sampling Survey

Census refers to the procedure of systematically acquiring and recording information related to all items in a population or universe.

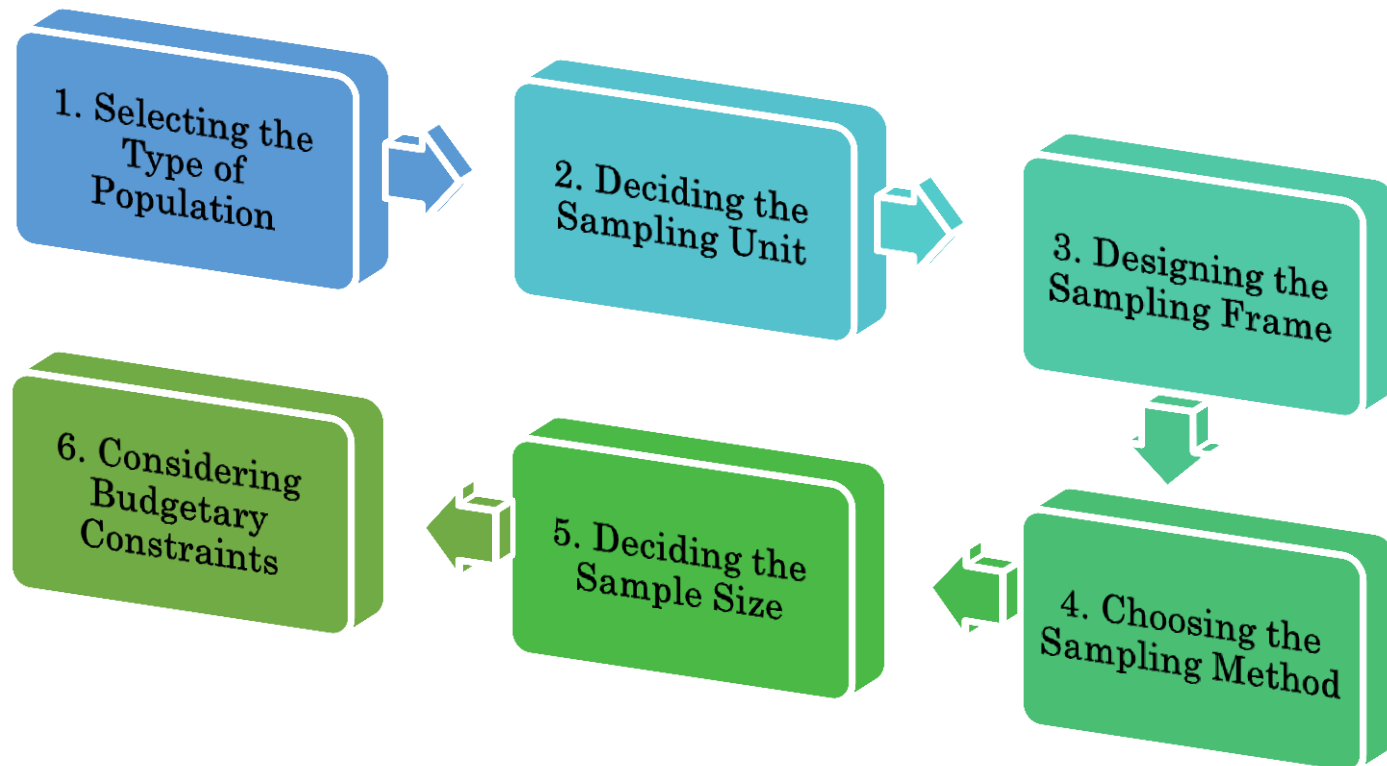Sample survey refers to the study of some selected elements from the population.

Census	Sample Survey
Used when it is required to gather information regarding each element or member of the population.	Used when a population contains numerous elements and it is not feasible to conduct a census.
Used when the researcher has enough time and other resources including money to conduct complete enumeration.	Used when the researcher does not have enough time and/or resources for conducting a census.

Developing Sample Design/Sampling Process

Sample design refers to a methodological plan to obtain a sample from a population.

The steps involved in developing a sample design are:

1. Selecting the Type of Population → 2. Deciding the Sampling Unit → 3. Designing the Sampling Frame → 4. Choosing the Sampling Method → 5. Deciding the Sample Size → 6. Considering Budgetary Constraints

Selecting the type of population

The researcher, while developing a sample design, must choose the type of population according to his/her research study which can be finite or infinite.

Deciding the sampling unit

A sampling unit is one of the units into which an aggregate is divided for the purpose of sampling, each unit is regarded as individual and indivisible when selection is made.

Designing the sampling frame

A sampling frame refers to the list of all in a population. If it is not available it should be prepared by the researcher. It should be comprehensive, authentic and appropriate.

Choosing the sampling method

A sampling method is a procedure of selecting sampling units from a population.

Deciding the sample size

A sample size refers to the number of items in a sample. Before selecting a sample size, the researcher must consider the population variance for the variable of interest.

Considering budgetary constraints

These refer to finance-related constraints. While developing a sample design, the researcher must consider those financial aspects that greatly affect not only the size of sample but also the type of sample.

Characteristics of a Good Sampling Design

A sample design must be feasible in the context of available funds

A sample design should result in such type of sample whose results can be applied to the whole population

A sample design should not be based on personal bias of the researcher

Determining Sample Size

There is a relation between the sample size and the accuracy of the research. The larger the sample size, the more accurate the research findings will be.

Before determining the sample size, it is important to have information regarding the following aspects:

Population size: In order to determine the sample size, the population size must be known.

Confidence Interval (margin of error): CI determines the allowable difference between the population mean and the sample mean.

Confidence level: It refers to the percentage confidence that the sample mean falls within the given confidence interval.

Standard deviation: It refers to the standard deviation that a researcher may expect in his responses for the variable of interest.

How to calculate the sample size?

- First, find out the Z-score or the constant value that represents the confidence level corresponding to a Z-score which could be found out using Z-score table.
- Next, insert the values of Z-score, confidence interval and standard deviation in the formula to calculate the sample size:

$$\text{Required sample size} = \frac{(Z_{Score})^2 \sigma^2}{(Margin\ of\ Error)^2}$$

- Z-score for 99% confidence level is 2.576.
- If a researcher choses to conduct the research with a confidence level of 99%, confidence interval of ±3% with a Standard Deviation (SD) of 0.5; then, the required sample size will be: $= \frac{(2,567)^2 (0.5)^2}{(0.03)^2} = 1843.055$
- Therefore, a sample containing 1844 subjects is required.

Example Calculate Sample size for 99% CI

Errors in Measurement and Sampling Errors

Some important errors in measurement are:

Errors due to the interviewer

These errors occur because of the biased attitude of the interviewer.

Errors due to respondents

These errors occur due to the reluctance of respondents to respond to questions.

Errors due to ineffectiveness of the instrument

These errors occur because of the ineffective measuring instrument, such as a questionnaire.

Errors due to the situation

These errors occur because of situational factors.

In research study, two types of errors can occur

Sampling errors

Non-sampling errors
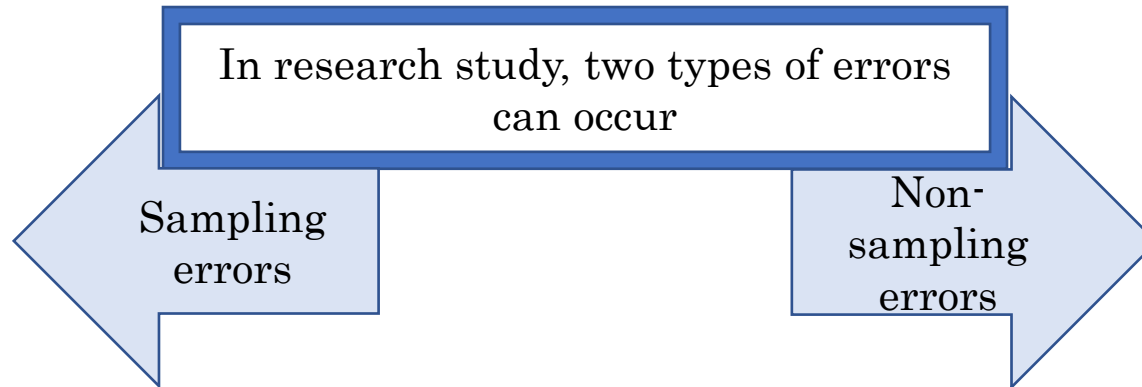
Sampling errors (random errors)
- These errors occur due to observing only a sample of an entire population for a study instead of studying the entire population.
- Any value computed from a sample is called sample statistic.
- The sample statistic may or may not be close to the value obtained for the whole population.
- A deviation in the values leads to a sampling error.
- A sampling error occurs because of selection of only a specific part of the population as the sample.

Some important methods that result in less sampling errors are:

Increasing sample size:
- Increasing the sample size will reduce sampling errors.
- If the sample size is equal to the complete population, the scope of sampling error is zero.

Choose the target population carefully:
- The target population from which the researcher draws a sample should be selected carefully.
- **Example:** If a research relates to the search or usage pattern of Indian Facebook users, the population would include all the Facebook users whose nationality is Indian irrespective of the fact whether or not they are located in India.

Some important methods that result in less sampling errors are:

Stratification:

- It refers to dividing the given population into homogeneous and non-overlapping units or sub-groups (known as stratum) to make the sample more representative.
- Grouping is done on the basis of one or more common attributes.
- Example: There is a population of 1000 people, out of the entire population, 300 belong to high-class, 400 belong to middle-class, & 300 belong to low-class. The sizes of these strata are denoted by S_1, S_2, and S_3.

Process of Stratification of the Given Sample of 1000 People

Stratum Number	Size of Stratum	Size of Sample from Each Stratum
1.	S_1 (High-class) = 300	$s_1 = (s \times s_1)/1000$ $= (100 \times 300)/1000 = 30$
2.	S_2 (Middle-class) = 400	$s_2 = (s \times s_2)/1000$ $s_2 = (100 \times 400)/1000 = 40$
3.	S_3 (Low – class) = 300	$s_3 = (s \times s_3)/1000$ $= (100 \times 300)/1000 = 30$
	$S_1 + S_2 + S_3 = 1000$ (Total Population)	$s_1 + s_2 + s_3 = 100$ (Desired Sample Size)

Non-Sampling Errors

Non-sampling errors occur as a result of factors other than sample selection. Reasons because of which these errors occur are:

Improper division of sampling units of a population:
- Example: An organisation wants to conduct a skill development programme for its employees who do not have the required skills to perform job duties.
- To do so, employee population of an organisation is divided as skilled population & unskilled population. However, this division may not be a precise division of the employee population.
- The reason is some people are involved in nontechnical jobs, some are involved in technical jobs and some people work on multiple projects simultaneously.
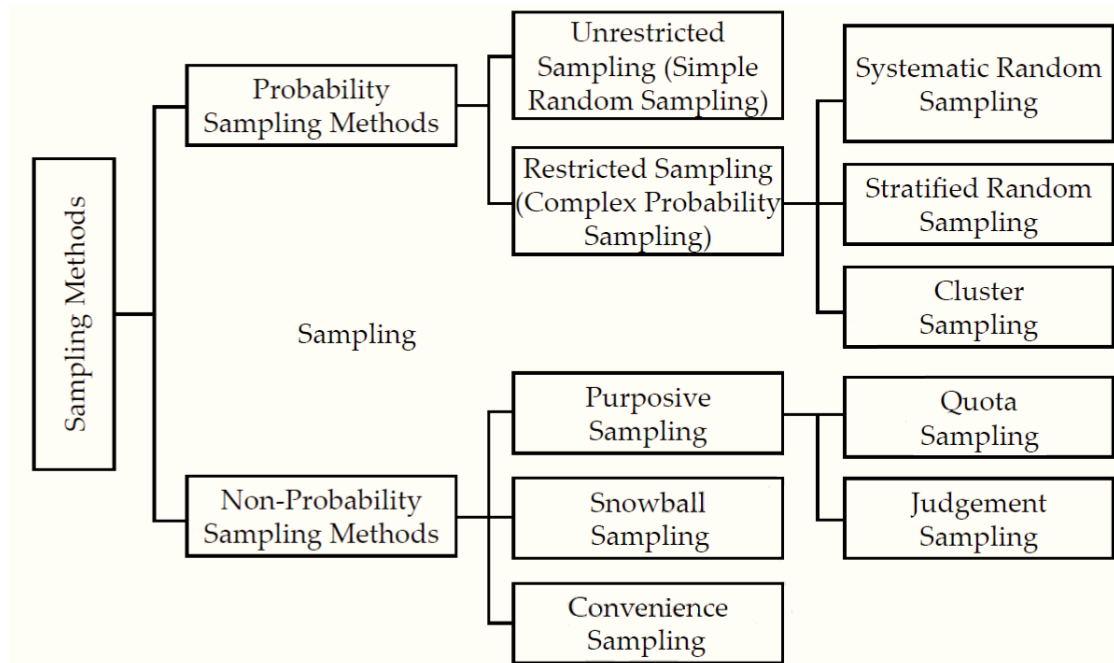- In addition, some people may be more qualified but they perform unskilled jobs.

Poor response of respondents:
- This makes it difficult for the researcher to derive accurate results.
- Usually, respondents show reluctance in supplying accurate information about their income, age and current level of education and skills.
- The incorrect information provided by respondents leads to erratic results, even if each element of the population is considered.

Methods of Sampling

Sampling methods are:

		Unrestricted Sampling (Simple Random Sampling)	Systematic Random Sampling
	Probability Sampling Methods	Restricted Sampling (Complex Probability Sampling)	Stratified Random Sampling
			Cluster Sampling
Sampling Methods	Sampling		
		Purposive Sampling	Quota Sampling
	Non-Probability Sampling Methods	Snowball Sampling	Judgement Sampling
		Convenience Sampling	

Some important factors that affect the choice of a sampling method are:

- Accuracy
- Supplementary information available about the given population
- Cost concerns
- Generalisation

Probability Sampling Methods

Probability sampling refers to the method of sampling in which the probability of selecting each item in the sample is known.
Random sampling refers to selecting a sample from the given population on a random basis.

Three types of probability sampling are:

1. Unrestricted Sampling (Simple Random Sampling)

In this type of probability sampling, each element in the population has equal chance of selection in the sample.

- **Example:** A researcher needs to select five elements out of the total population of 10.
- In this case, he/she knows that the probability of selection of each element is equal because elements would be selected on the basis of random sampling without any bias involved.

2. Restricted Sampling (Complex Sampling Method)

In this method, sampling is done on a restricted basis. Some of the common types of restricted sampling are:

Stratified random sampling

- It refers to a sampling method in which the sample is selected by dividing population into different subgroups or strata.

- Example: A human resource manager wants to assess the need of training for employees working in his/her organisation. Here, the total population of that organisation constitutes the total number of employees. Since the different levels of employees have different needs for training, the manager divides the population into exclusive subgroups, such as clerical workers, computer administrators, supervisors, lower-level managers & middle-level managers.

Two ways to select the subgroup in Stratified random sampling

Proportionate stratified sampling

- In this method, elements from all the subgroups of population are selected corresponding to their exact proportion in the population.

- Example of Proportionate Stratified Sampling

Employees	Number of Elements in the Population	Number of Elements Selected for the Sample by Proportionate Sampling
Clerical workers	500	100
Computer administrators	400	80
Supervisors	200	40
Lower-level managers	150	30
Middle-level managers	40	8
Top-level managers	10	2
	1300	260

Disproportionate stratified sampling

- In this method, there is no proportionate allocation of elements for the sample.

- Example of Disproportionate Sampling

Employees	Number of Elements in the Population	Number of Elements Selected for the Sample by Disproportionate Sampling
Clerical workers	500	70
Computer administrators	400	60
Supervisors	200	50
Lower-level managers	150	50
Middle-level managers	40	24
Top-level managers	10	6
	1300	260

Systematic sampling

- In systematic sampling, a sample is selected from the sampling frame at regular intervals.
- Steps:
  - ✓ In this type of sampling, the elements in the sampling frame are numbered consecutively.
  - ✓ After this, a random number is chosen.
  - ✓ Thereafter, the sampling fraction is calculated as the ratio of actual sample size and the total population.
  - ✓ Starting with the randomly selected number, samples are drawn using a frequency (inverse of the sampling fraction).

Cluster sampling

- Steps
  - ✓ In this type of probability sampling, the entire universe is divided into groups or clusters.
  - ✓ After that, these clusters are selected on the basis of random sampling.
  - ✓ All elements of the selected clusters should be included in the sample leaving out all the elements of the non-selected clusters.
- In cluster sampling, due care should be taken to make clusters as heterogeneous as possible.
- Cluster sampling is usually done on the basis of geographical areas due to which more data can be collected using face-to-face methods.

Non-Probability Sampling Methods

In non-probability (non-random) sampling, samples are collected in a manner according to which all individuals in the population do not get equal chances of being selected.

Important types of nonprobability sampling are:

- In this type of non-probability sampling, the researcher selects those elements or subjects from the target population that are easily accessible to him/her.
- Example: in a college, volunteers are required to organise a tree plantation camp. The strength of the college is 2000 and the number of volunteers required is 50. In this case, the easiest way to select volunteers is their accessibility. The researcher can select those students as volunteers who are easily accessible to him/her.

Convenience Sampling

Purposive Sampling

- It is a non-probability sampling method wherein the sample is chosen purposively on the basis of certain characteristics of a population and the objective of the study.
- **Example**: A researcher wants to gather opinion of working mothers about conditions at their workplaces. In this case, the researcher would contact only those women who are mothers and working. The females not falling in this category would not be surveyed.

Quota Sampling

The given population is first divided into mutually nonoverlapping subgroups. Thereafter, a sample is formed by selecting members from each subgroup according to the proportion of each subgroup in the total population; this is called proportionate quota sampling.

The two subtypes of purposive sampling

Judgement Sampling

The elements or units from the population are selected on the recommendation of experts in the field of research work that is being carried out.

The experts are asked to select the units that should be included in the sample so that the sample is truly representative of the population.

Snowball Sampling

Snowball Sampling (Chain sampling or Referral sampling)

- In the snowball technique, the researcher first picks up one or more subjects (to be included in sample) and then he/she asks them to recommend or refer subjects who conform to the criteria for being included in the sample.
- This process of referral is repeated with the new subjects till the required number of subjects in the sample is fulfilled.
- Example: For a research study, a researcher needs environmental photographers. The researcher finds only four photographers; then he/she asks these four people to provide information of more environmental photographers. On the basis of their information, contacts are increased and sufficient data is collected.

Thank You