CHAPTER -2

2. Sampling and sampling techniques

Introduction

Given a variable X, if we arrange its values in ascending order and assign probability to each of the values or if we present X_i in a form of relative frequency distribution the result is called *Sampling Distribution of X*.

Definitions:

1. *Parameter*: Characteristic or measure obtained from a population.
2. *Statistic*: Characteristic or measure obtained from a sample.
3. *Sampling*: The process or method of sample selection from the population.
4. *Sampling unit*: the ultimate unit to be sampled or elements of the population to be sampled.

Examples:

- If some body studies Scio-economic status of the households, households are the sampling unit.
- If one studies performance of freshman students in some college, the student is the sampling unit.

5. *Sampling frame*: is the list of all elements in a population.

Examples:

- List of households.
- List of students in the registrar office.

6. *Errors in sample survey*:

There are two types of errors

a) *Sampling error*:
- It is the discrepancy between the population value and sample value.
- May arise due to inappropriate sampling techniques applied

b) *Non sampling errors*: are errors due to procedure bias such as:
- Due to incorrect responses
- Measurement
- Errors at different stages in processing the data.

The Need for Sampling

- Reduced cost
- Greater speed
- Greater accuracy
- Greater scope
- More detailed information can be obtained.

There are two types of sampling.

1. Random Sampling or probability sampling.

It is a method of sampling in which all elements in the population have a pre-assigned non-zero probability to be included in to the sample.

Examples:

- Simple random sampling
- Stratified random sampling

- Cluster sampling
- Systematic sampling

1. Simple Random Sampling*:*

- It is a method of selecting items from a population such that every possible sample of specific size has an equal chance of being selected. In this case, sampling may be with or without replacement. Or
- All elements in the population have the same pre-assigned non-zero probability to be included in to the sample.
- Simple random sampling can be done either using the lottery method or table of random numbers.

2. **Stratified Random Sampling***:*

- The population will be divided in to non-overlapping but exhaustive groups called strata.
- Simple random samples will be chosen from each stratum.
- Elements in the same strata should be more or less homogeneous while different in different strata.
- It is applied if the population is heterogeneous.
- Some of the criteria for dividing a population into strata are: Sex (male, female); Age (under 18, 18 to 28, and 29 to 39); Occupation (blue-collar, professional, and others).

3. Cluster Sampling*:*

- The population is divided in to non-overlapping groups called clusters.
- A simple random sample of groups or cluster of elements is chosen and all the sampling units in the selected clusters will be surveyed.
- Clusters are formed in a way that elements with in a cluster are heterogeneous, i.e. observations in each cluster should be more or less dissimilar.
- Cluster sampling is useful when it is difficult or costly to generate a simple random sample. For example, to estimate the average annual household income in a large city we use cluster sampling, because to use simple random sampling we need a complete list of households in the city from which to sample. To use stratified random sampling, we would again need the list of households. A less expensive way is to let each block within the city represent a cluster. A sample of clusters could then be randomly selected, and every household within these clusters could be interviewed to find the average annual household income.

4. Systematic Sampling*:*

- A complete list of all elements with in the population (sampling frame) is required.
- The procedure starts in determining the first element to be included in the sample.
- Then the technique is to take the k^{th} item from the sampling frame.

Let

$$N = population\ size,\ n = sample\ size,\ k = \frac{N}{n} = sampling\ \text{int}\ erval.$$

- Chose any number between 1 and k. Suppose it is $j\,(1 \le j \le k)$.
- The j^{th} unit is selected at first and then $(j+k)^{th}, (j+2k)^{th}, ...etc$ until the required sample size is reached.

2. Non Random Sampling or non-probability sampling.

- It is a sampling technique in which the choice of individuals for a sample depends on the basis of convenience, personal choice or interest.

Examples:

- Judgment sampling.
- Convenience sampling
- Quota Sampling.

1. Judgment Sampling

- In this case, the person taking the sample has direct or indirect control over which items are selected for the sample.

2. Convenience Sampling

- In this method, the decision maker selects a sample from the population in a manner that is relatively easy and convenient.

3. Quota Sampling

- In this method, the decision maker requires the sample to contain a certain number of items with a given characteristic. Many political polls are, in part, quota sampling.

Note:

let $N = population\ size,\ n = sample\ size.$

1. Suppose simple random sampling is used N^n
   - We have N^n possible samples if sampling is with replacement.
   - We have $\binom{N}{n}$ possible samples if sampling is with out replacement.
2. After this on wards we consider that samples are drawn from a given population using simple random sampling.