Unit-IV
Sampling theory

Model No 4.1: **Introduction to population and sample**
Model No 4.2: **Standard error**
Model No 4.3: **Sampling distribution of means and variances**
Model No 4.4: **Central limit theorem**
Model No 4.5: **Problems on series**
Model No 4.6: **Estimations & Point and interval estimations**
Model No 4.7: **Unbiased Estimations**
Model No 4.8: **Maximum error of estimate**

Model No 4.1: Introduction to population and sample

It is not easy to collect all the information about population and also it is not possible to study the characteristics of the entire population (finite or infinite) due to time factor, cost factor and other constraints. Thus we need sample. Sample is a finite subset of statistical individuals in a population and the number of individuals in a sample is called the sample size. Sampling is quite often used in our day-to-day practical life.

For example in a shop we assess the quality of rice, wheat or any other commodity by taking a handful of it from the bag and then to decide to purchase it or not.

Population

The population is a complete set of all possible observations of the type which is to be investigated. Total number of students studying in a school or college, total number of books in a library, total number of houses in a village or town is some examples of population.

Sometimes it is possible and practical to examine every person or item in the population we wish to describe. We call this a complete enumeration, or census. We use sampling when it is not possible to measure every item in the population. Statisticians use the word population to refer not only to people but to all items that have been chosen for study.

Finite population and infinite population:

A population is said to be finite if it consists of finite number of units. Number of workers in a factory, production of articles in a particular day for a company is examples of finite population. The total number of units in a population is called population size (N). A population is said to be infinite if it has infinite number of units. For example the number of stars in the sky, the number of people seeing the Television programmes etc.,

Sample

Statisticians use the word sample to describe a portion chosen from the population. A finite subset of statistical individuals defined in a population is called a sample. The number of units in a sample is called the sample size (n).

Types of sampling:

i) **Purposive sampling:** Purposive sampling is one in which sample units are selected with a definite purpose in view.
Ex: Suppose you want to collect feedback from students on the pedagogical methods in their school.

ii) **Random sampling:** Random sample is the one in which each unit of population has an equal chance of being included in it. And the sample obtained by this sampling is termed as random sample.
Ex: 25 students were selected in WIPRO from VVIT out of a hat from 3000 students who are studying in VVIT.

iii) **Simple sampling:** Simple sampling in which each unit of the population has an equal chance of being included in the sample and this probability is independent of the previous drawings.

Note:

1. Simple sampling may be regarded as random sampling but a random sampling is not necessarily a simple sampling.
2. For a finite population, random sampling with replacement is a simple sampling while random sampling without replacement is not a simple sampling.
3. For an infinite population, any random sampling is simple.

Example for Simple sampling: 25 students were selected in WIPRO from VVIT out of a hat from 3000 students who are studying in VVIT. In this case, the population is all 3000 students, and the sample is random because each STUDENT has an equal chance of being chosen.

iv) Stratified random sampling:

A method of sampling that involves dividing a population into smaller groups–called **strata**. The groups or strata are organized based on the shared characteristics or attributes of the members in the group. The process of classifying the population into groups is called **stratification**.

Examples for Stratified random sampling One might divide a sample of adults into subgroups by age, like 18–29, 30–39, 40–49, 50–59, and 60 and above.

Large Sample & Small Sample:

If the sample size
$n \geq 30$ i.e, referred as Large Sample
$n < 30$ i.e, referred as Small Sample

Sampling is done in 2 ways:

i) With replacement (infinite)
ii) Without replacement (finite)

Parameters and statistics:

We can describe samples and populations by using measures such as the mean, median, mode and standard deviation. When these terms describe the characteristics of a population, they are called parameters. When they describe the characteristics of a sample, they are called statistics. A parameter is a characteristic of a population and a statistic is a characteristic of a sample. Since samples are subsets of population statistics provide estimates of the parameters. That is, when the parameters are unknown, they are estimated from the values of the statistics.

	Parameters (Population)	Statistics (Sample)
Mean	μ	$\bar{x}$
Proportion	P (Capital)	p (small)
Variance	σ^2	s^2
Standard deviation	σ	s

Model No 4.2: Standard Error

The standard deviation of the sampling distribution of a statistic is known as its standard error. It is abbreviated as S.E.

For example, the standard deviation of the sampling distribution of the mean x known as the standard error of the mean.

S. No.	Standard error (S. E.)	With replacement Infinite Population	Without replacement Finite Population
1	Standard error of sample mean ($\bar{x}$)	$\frac{\sigma}{\sqrt{n}}$	$\frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$
2	Standard error of sample Proportion (p)	$\sqrt{\frac{PQ}{n}}$ Where $Q=1-P$	$\sqrt{\frac{PQ}{n}}\sqrt{\frac{N-n}{N-1}}$ Where $Q=1-P$
3	Standard error of sample Standard deviation (s)	$\frac{\sigma}{\sqrt{2n}}$	
4	Standard error of the difference of two sample means $\bar{x}_1$ and $\bar{x}_2$	$\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}$	
5	Standard error of the difference of two sample proportions p_1 and p_2	$\sqrt{\frac{P_1Q_1}{n_1}+\frac{P_2Q_2}{n_2}}$	
6	Standard error of the difference of two standard deviations s_1 and s_2	$\sqrt{\frac{\sigma_1^2}{2n_1}+\frac{\sigma_2^2}{2n_2}}$	