- **i.** Definition of probability
- **ii.** Conditional probability
- **iii.** Independent an dependent events
- **iv.** Mutually Exclusive events
- **v.** Probability distribution
- **vi.** Mathematical Expectations
- **vii.** Relation between population, sample, mean and variance
- **viii.** Combinatorial analysis
- **ix.** Combinations
- **x.** Stirling Approximation to n!
- **xi.** Relation of probability to point set theory
- **xii.** Euler or Venn diagram and probability

---

1. Sampling theory
2. Random Sample and random numbers
3. Sampling with and without replacement
4. Sampling distribution of mean
5. Sampling distribution of proportion
6. Sample distribution of difference and sums
7. Standard errors.

1. Sampling theory

For the purpose of determining population characteristics, instead of enumerating entire population, the individual in the sample, only observed then the sample characteristics are utilised to approximately determine or estimate the population.

For example: for examining the sample of a particular stuff, we arrive at a decision of purchasing or rejecting the stuff. The error involved in such approximation is called SAMPLING ERROR and is inherent and unavoidable in any or every sampling scheme. But sampling result in considerable gain, especially I time and cost, not only in respect of making observations of characteristics but also in the subsequent handling of data.

In **statistics**, quality assurance, and survey methodology, sampling is the selection of a subset of individuals from within a statistical population to estimate characteristics of the whole population.

In business, medical, social and psychological sciences etc., research, sampling theory is widely used for gathering information about a population. The sampling process comprises several stages:

- Defining the population of concern
- Specifying the sampling frame (set of items or events possible to measure)
- Specifying a sampling method for selecting the items or events from the sampling frame
- Determining the appropriate sample size
- Implementing the sampling plan
- Sampling and data collecting
- Data which can be selected

Every sampling system is used to obtain some estimates having certain properties of the population under study.

Some of the commonly known an frequently used type of sampling are:

i. Purposive sampling
ii. Random sampling
iii. Stratified sampling
iv. Systematic sampling

In case of Random Sampling the sample units are selected at random and in which each unit of population has an equal chance of being included in it.

Suppose we take sample of size n from a finite population size of N, then there are $^{N}C_{n}$ possible samples. A sampling techniques in which each of the $^{N}C_{n}$ samples has an equal chance of being selected is known as SAMPLING and the sample obtained by this technique is termed as RANDOM SAMPLE.

Fairly good random sample obtained by the Trippet's number tables or by throwing a dice, draw of a lottery etc.

(Trippet's random number tables consist of 10400 four-digited numbers, giving in all 10400× 4 i.e. 41600 digits, taken from the British census report. These table have proved to be fairly random in character. Any page of table is selected at random and the numbers in any row or column or diagonal selected at random may be taken to constitute the sample.

2. **Sampling with and without replacement**

   Sampling where each member of the population may be chosen more than once it is called SAMPLING WITH REPLACEMENT, while if each member cannot be chosen more than once it is called SAMPLING WITHOUT REPLACMENT.

3. **Sampling distribution of mean**

   The probability distribution of any statistic is known as SAMPLING DISTRIBUTION.

   The sample mean and sample variance are the most common statistic that are computed for sample.

   Hey both have sampling distribution that have general properties regardless of the probability distributions of the parent population.

   SAMPLING DISTRIBUTION OF MEAN:

   If all possible samples of size n are drawn without replacement from a finite population of size N_p,

   Let we denote mean and standard deviation of sampling distribution by $\bar{x}$ and $\sigma_{\bar{x}}$ and population mean and standard deviation by μ *and* σ then

   $$\bar{x} = \mu \text{ and } \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{N}} = \sqrt{\frac{Np-N}{Np-1}}$$

- For large value of N (N > 30) the sampling distribution of mean approximately a normal distribution with mean $\bar{x}$ and standard deviation (s.d.) by $\sigma_{\bar{x}}$ irrespective of the population

- This result for infinite population is a special case of CENTRLA LIMIT THEOREM of advanced probability theory, which shows that accuracy of the approximation improved as N gets large.

4. **Sampling distribution of proportion**

Suppose that a population is infinite and that the probability of occurrence of an event (probability of success) is p, while the probability of failure is q=1-p, consider all possible sample of size N drawn from population and we obtained a sampling distribution of proportions whose mean $\boldsymbol{\bar{x} = p}$ & $\boldsymbol{\sigma_{\bar{x}}} = \sqrt{\frac{\mathbf{pq}}{\mathbf{N}}}$

5. **Sample distribution of difference and sums**

Suppose we are giving two populations. For each of sample size n_1 drawn from first population, let us compute a statistic s_1, this yield a sampling distribution for the statistics whose mean and standard deviations we denoted by μ_{s_1} and σ_{s_1} respectively. Similarly for other sample of size n_2, let us compute a statistic s_2, this yield a sampling distribution for the statistics whose mean and standard deviations we denoted by μ_{s_2} and σ_{s_2} respectively.

From all possible combinations of these samples from the two populations we can obtain a distribution of the difference $s_1 - s_2$, which is called the sampling distribution of differences of statistic.

The mean and standard deviation of these distribution is denoted respectively by

$\mu_{s1-s2} = \mu_{s_1} - \mu_{s_2}$ and $\sigma_{s1-s2} = \sqrt{\sigma_{s_1}{}^2 + \sigma_{s_2}{}^2}$

Provided that the sample chosen do not in any way depend on each other.
(if s_1 , s_2 are two sample mean denoted by $\boldsymbol{\overline{x_1}}$ ***and*** $\boldsymbol{\overline{x_2}}$ then

$\mu_{\overline{x_1}-\overline{x_2}} = \mu_{\overline{x_1}} - \mu_{\overline{x_2}}$ & $\sigma_{\overline{x_1}-\overline{x_2}} = \sqrt{\sigma_{\overline{x_1}}{}^2 + \sigma_{\overline{x_2}}{}^2} = \sqrt{\frac{\sigma_{\overline{x_1}}{}^2}{n1} + \frac{\sigma_{\overline{x_2}}{}^2}{n2}}$

Similarly we will write for proportion.

6. Standard errors.

The difference between the observed value of a statistic and the value of the parameter is known as the sampling error.

Sampling error also called estimation error is the amount of inaccuracy in estimating some value that is caused by only a portion of a population (i.e. *sample*) rather than the whole population. It is the difference between the statistic (value of sample, such as sample mean) and the corresponding parameter (value of population, such as population mean) is called the sampling error.

The standard deviation of the sampling distribution of a statistics is known as its standard Error, abbreviated as S.E.

The standard errors of some of the well-known statistics, for large samples, are given below, where n is the sample size, σ^2 the population variance (or $\boldsymbol{\sigma}$ be the standard deviation), and P the population proportion, and Q=1-P; n_1 *and* n_2 represent the sizes of two independent random samples respectively drawn from the given population.

Sr.No.	Statistic	Standard Error
1	Sample mean	$\frac{\sigma}{\sqrt{n}}$
2	Observed sample proportion	$\sqrt{\frac{PQ}{n}}$
3	Sample standard deviation	$\sqrt{\frac{\sigma^2}{2n}}$
4	Sample variance	$\sigma^2\sqrt{\frac{2}{n}}$
5	Difference of two sample means	$\sqrt{\frac{{\sigma_1}^2}{n_1}+\frac{{\sigma_2}^2}{n_2}}$
6	Difference of two sample deviation	$\sqrt{\frac{{\sigma_1}^2}{2n_1}+\frac{{\sigma_2}^2}{2n_2}}$
7	Difference of two sample proportion	$\sqrt{\frac{P_1Q_1}{n_1}+\frac{P_2Q_2}{n_2}}$
8	Sampling from finite population of size N	$\sqrt{\frac{(N-n)PQ}{(N-1)n}}$