SAMPLING AND STATISTICAL INFERENCE

SAMPLING THEORY

SOME IMPORTANT TERMS ASSOCIATED WITH SAMPLING

Population and Sample

Population or Universe. Population in statistics means the whole of the information which comes under the purview of statistical investigation. It is the totality of all the observations of a statistical experiment or enquiry. In other words, an aggregate of objects; animate or inanimate under study is the population. It is also known as the universe. For example, the population of the heights of students in a school, or the population of the sum of points obtained in throwing three dice.

Finite or Infinite Population. A population may be finite or infinite according as the number of observations or items in it are finite or infinite. The population of weights of students of class XII in a government school is an example of a finite population. The population of pressure at different points in the atmosphere is an example of an infinite population.

Sample. A part of the population selected for study is called a sample. In other words, the selection of a group of individuals or items from a population in such a way that this group represents the population, is called a sample. For example, a housewife tests a small quantity of rice to see that it has been cooked or not. This
A small quantity of rice is a sample and represents the entire quantity of rice cooked. A sample is a selected portion of the population. A sample drawn from a population provides valuable information about its parent population. It gives a fairly accurate result and a reliable picture of the total observations under investigation. It is always used to measure and estimate the corresponding characteristics of its parent population. When the sample drawn is perfectly representative, it is identical with its parent population almost in every respect except that it is smaller than the population.

Size of a Sample. The number of individuals or items included in a finite sample is called the size of the sample.

### Parameter and Statistic

There are various statistical measures in statistics such as mean, median, mode, standard deviation, coefficient of variation, variance etc. These statistical measures can be computed both from population (or universe) data and sample data.

**Parameter:** Any statistical measure computed from population data is known as parameter.

**Statistic:** Any statistical measure computed from sample data is known as statistic.

Thus a parameter is a statistical measure which relates to the population and is based on population data, whereas a statistic is a statistical measure which relates to the sample and is based on sample data. Thus a population mean, population median, population variance, population coefficient of variation etc., are all parameters. Statistic computed from a sample such as sample mean, sample variance etc., which are drawn from the parent population plays an important role in (i) The theory of Estimation and, (ii) Testing of hypothesis.

The usual notation used for parameter (in the case of population data) and Statistic (in the case of sample data) are given below:

Formally, a parameter is any function of population values, while a statistic is a function of sample values. Very often the values of various parameters are unknown and these are estimated by the corresponding statistic. For example, sample mean \( \bar{X} \) (or sample standard deviation \( S \)) is used as an estimator of population mean (or population standard deviation).

**Notations**

<table>
<thead>
<tr>
<th>Statistical Measure</th>
<th>Population</th>
<th>Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>( \mu )</td>
<td>( \bar{X} )</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>( \sigma )</td>
<td>( S )</td>
</tr>
<tr>
<td>Proportion</td>
<td>( P )</td>
<td>( p )</td>
</tr>
<tr>
<td>Size</td>
<td>( N )</td>
<td>( n )</td>
</tr>
</tbody>
</table>
SAMPLING

It is the procedure or process of selecting a sample from a population. A sampling can also be defined as the process of drawing a sample from a population and of compiling a suitable statistic from such a sample in order to estimate the parameter of the parent population and to test the significance of the statistic computed from such sample.

SAMPLING THEORY

The study of relationship existing between a population and the samples drawn from the population is called Sampling. Sampling theory is based on sampling. It deals with statistical inferences drawn from sampling results. Statistical inference made on the basis of sampling results are of the following three types:

(i) **Statistical Estimation.** It helps in estimating an unknown population parameter (such as population mean, median, mode, standard deviation, kurtosis etc.) on the basis of suitable statistic (such as sample mean, mode, median, variance etc.) computed from the sample drawn from such parent population.

(ii) **Tests of Significance.** Sampling theory helps in testing of significance about the population characteristics on the basis of suitable statistic computed from a sample drawn from such parent population. In other words, it helps in determining whether observed differences between two samples are actually due to chance or whether they are really significant. Such questions help us in deciding whether one production is better than the other. The test of significance plays an important role in the decision theory.

(iii) **Statistical Inference.** Statistical inference means drawing conclusions about some matters on the basis of certain statistical results. These statistical results are obtained by drawing samples from the population and then by computing suitable statistic from these samples so as to make statistical inferences. These statistical inferences enable us to draw statistical conclusions about some measures of a population on the basis of such statistic.

TYPES OF SAMPLING

A sample can be selected from a population in various ways. Different situations call for different methods of sampling. There are two methods of Sampling:

1. Random Sampling or Probability Sampling Method
2. Non-Random Sampling or Non-Probability Sampling Method.
**Random Sampling or Probability Sampling**

**Random Sampling.** Random or Probability sampling is the scientific technique of drawing samples from the population according to some laws of chance in which each unit in the universe or population has some definite pre-assigned probability of being selected in the sample. It is of two types.

(a) **Simple Random Sampling (SRS).**

It is the method of selection of a sample in such a way that each and every member of population or universe has an equal chance or probability of being included in the sample. Random sampling can be carried out in two ways.

(i) **Lottery Method.** It is the simplest, most common and important method of obtaining a random sample. Under this method, all the members of the population or universe are serially numbered on small slips of a paper. They are put in a drum and thoroughly mixed by vibrating the drum. After mixing, the numbered slips are drawn out of the drum one by one according to the size of the sample. The number of slips so drawn constitute a random sample.

(ii) **Random Number Method.** In this method, sampling is conducted on the basis of random numbers which are available from the random number tables. The various random number tables available are:

(a) Trippet's Random Number Series;

(b) Fisher's and Yales Random Number Series;

(c) Kendall and Badington Random Number Series;

(d) Rand Corporation Random Number Series;

One major dis-advantage of random sampling is that all the members of the population must be known and be serially numbered. It will entail a lot of difficulties in case the population is of large size and will be impossible in case the population is of infinite size.

(b) **Restricted Random Sampling.**

It is of three types

(i) Stratified Sampling

(ii) Systematic Sampling

(iii) Multi-stage Sampling
Stratified Sampling. In stratified random sampling, the population is divided into strata (groups) before the sample is drawn. Strata are so designed that they do not overlap. Elementary units from each stratum is drawn at random and the units so drawn constitute a sample. Stratified sampling is suitable in those cases where the population is heterogeneous but there is a homogeneity within each of the groups or strata.

Advantages

(i) It is a representative sample of the heterogeneous population.

(ii) It lessens the possibility of bias of one sidedness.

Disadvantages

(i) It may be difficult to divide the population into heterogeneous groups.

(ii) There may be over-lapping of different strata of the population which will provide an unrepresentative sample.

Systematic Sampling. In this method every elementary unit of the population is arranged in order and the sample units are distributed at equal and regular intervals. In other words, a sample of suitable size is obtained (from the orderly arranged population) by taking every unit say tenth unit of the population. One of the first units in this ordered arrangement is chosen at random and the sample is computed by selecting every tenth unit (say) from the rest of the lot. If the first unit selected is 4, then the other units constituting the sample will be 14, 24, 34, 44, and so on.

Advantages. It is most suitable where the population units are serially numbered or serially arranged.

Disadvantages. It may not provide a desirable result due to large variation in the items selected.

Multi-stage Sampling. In this sampling method, sample of elementary units is selected in stages. Firstly a sample of cluster is selected and from among them a sample of elementary units is selected. It is suitable in those cases where population size is very big and it contains a large number of units.

Non-Random Sampling or Non-Probability Sampling Method

A sample of elementary units that is being selected on the basis of personal judgement is called a non-probability sampling. It is of five types.
(i) Purposive Sampling; (ii) Cluster Sampling;
(iii) Quota Sampling; (iv) Convenience Sampling;
(v) Sequential Sampling.

**Purposive Sampling.** Purposive sampling is the method of sampling by which a sample is drawn from a population based entirely on the personal judgement of the investigator. It is also known as **Judgement Sampling or Deliberate Sampling.** A randomness finds no place in it and so the sample drawn under this method cannot be subjected to mathematical concepts used in computing sampling error.

**Cluster Sampling.** Cluster Sampling involves arranging elementary items in a population into heterogeneous subgroups that are representative of the overall population. One such group constitutes a sample for study.

**Quota Sampling.** In quota sampling method, quotas are fixed according to the basic parameters of the population determined earlier and each field investigator is assigned with quotas of number of elementary units to be interviewed.

**Convenience Sampling.** In convenience sampling, a sample is obtained by selecting convenient population elements from the population.

**Sequential Sampling.** In sequential sampling a number of sample lots are drawn one after another from the population depending on the results of the earlier samples drawn from the same population. Sequential sampling is very useful in Statistical Quality Control. If the first sample is acceptable, then no further sample is drawn. On the other hand if the initial lot is completely unacceptable, it is rejected straightway. But if the initial lot is of doubtful and marginal character falling in the area lying between the acceptance and rejection limits, a second sample is drawn and if need be a third sample of bigger size may be drawn in order to arrive at a decision on the final acceptance or rejection of the lot. Such sampling can be based on any of the random or non-random method of selection.

**ADVANTAGES OF RANDOM (OR PROBABILITY) SAMPLING**

1. Random sampling is objective and unbiased. As a result, it is defensible before the superiors or even before the court of law.

2. The size of sample depends on demonstrable statistical method and therefore, it has a justification for the expenditure involved.
3. Statistical measures, i.e., parameters based on the population can be estimated and evaluated by sample statistic in terms of certain degree of precision required.

4. It provides a more accurate method of drawing conclusions about the characteristics of the population as parameters.

5. It is used to draw the statistical inferences.

6. The samples may be combined and evaluated, even though accomplished by different individuals.

7. The results obtained can be assessed in terms of probability, and the sample is accepted or rejected on a consideration of the extent to which it can be considered representative.

**ERRORS IN SAMPLE SURVEY**

A sample is a part of the whole population. A sample drawn from the population depends on chance and as such all the characteristics of the population may not be present in the sample drawn from the same population. Any statistical measure, say, mean of the sample, may not be equal to the corresponding statistical measure (mean) of the population from which the sample has been drawn. Thus there can be discrepancies in the statistical measure of population, i.e., parameter and the statistical measures of sample drawn from the same population, i.e., statistic. These discrepancies are known as Errors in Sampling. Errors in sampling are of two types.

(i) Sampling Errors

(ii) Non-sampling Errors or Bias.

**Sampling Errors**

Sampling Errors is inherent in the method of sampling. Sampling depends on chance and due to the existence of chance in sampling, the sampling errors occur. Errors in sampling arise primarily due to the following reasons:

1. **Faulty selection of the sample.** This may be due to selection of defective sampling techniques which may introduce the element of bias, e.g., purposive or judgement sampling, in which investigator deliberately selects a non-representative sample.

2. **Substitution.** Sometimes an investigator while collecting the information from a particular sampling unit, included in the random selection substitutes a convenient member of the population and this may lead to some bias as the characteristic possessed by the substituted unit may be different from those possessed by the original unit included in sampling.
3. Faulty demarcation of sampling units.

4. Variability of the population. Sampling error may also depend on the variability or heterogeneity of the population from which the samples are drawn.

Non-Sampling Errors or Bias

Non-sampling errors or Bias automatically creep in due to human factors which always varies from one investigator to another. Bias may arise in the following different ways:

(i) due-to negligence and carelessness on the part of investigator;
(ii) due to faulty planning of sampling;
(iii) due to the faulty selection of sample units;
(iv) due to incomplete investigation and sample survey;
(v) due to framing of a wrong questionnaire;
(vi) due to negligence and non-response on the part of the respondents;
(vii) due to substitution of a selected unit by another;
(viii) due to error in compilation;
(ix) due to applying wrong statistical measure.

SAMPLING DISTRIBUTION OF A STATISTIC

Sampling distribution of a statistic is the frequency distribution which is formed with various values of a statistic computed from different samples of the same size drawn from the same population. We can draw a large number of samples of same size from a population of fixed size, each sample containing different population members. Any statistic (statistical measure of sample) like mean, median, variance, standard deviation etc. may be computed for each of these samples. As a result a series of various values of that statistic may be obtained. These various values can be arranged into a frequency distribution table, which is known as the sampling distribution of the statistic.

Sampling may be done with replacement or without replacement. Sampling with replacement means that the same unit of the population may be included in each sample more than once. Sampling without replacement means that the same unit of population may not be included in each sample more than once. In the case of sampling with replacement, the total number of possible samples each of size n drawn out of population of size N is $N^n$. But if the sampling is
without replacement the total number of possible samples will be \( C(N, n) m \) (say). For each of these samples a value of statistic, say sample mean \( \bar{X} \), is computed. As the samples are formed with different sample units so the value of each of the sample means will be different. The sample mean can be regarded as a random variable \( \bar{X} \) and each sample mean then constitute as the observed value of this new random variable \( \bar{X} \). Let these values be: \( \bar{X}_1, \bar{X}_2, \bar{X}_3, ..., \bar{X}_m \). These mean values \( \bar{X}_1, \bar{X}_2, \bar{X}_3, ..., \bar{X}_m \) can be used to form a frequency distribution. Then this frequency distribution of the statistic \( \bar{X} \) is known as the sampling distribution of sample mean.

Similarly, the sampling distribution of standard deviation or coefficient of variation or variance may be constructed with the various values of standard deviation or coefficient of variation or variance respectively.

Let \( t_1, t_2, t_3, ..., t_m \), be the \( m \) values of a statistic for \( m \) possible samples. Then the statistic \( t \) can be regarded as a random variable which can take any one of the values \( t_1, t_2, t_3, ..., t_m \). This set of \( t \) values constitutes the sampling distribution of the statistic.

We can compute the mean, variance and other statistical measures of the sampling distribution of the statistic \( t \). For example,

\[
\text{Mean : } E(t) = \frac{1}{m} \sum t = \bar{t} \text{ (say)} \quad \text{.....(1)}
\]

\[
\text{Variance : } \text{Var}(t) = \frac{1}{m} \sum \left[ t - E(t) \right]^2 = \frac{1}{m} \sum (t - \bar{t})^2.
\]

If the population size is infinitely large or sampling is done with replacement, then the total number of possible samples of the same size which may be drawn from the population cannot be determined. In such a case, a large number of repeated random samples from the population of fixed size can be drawn and the values of statistic for these samples may be computed. These values of the statistic can be used to form a frequency distribution. This frequency distribution of the statistic is known as the sampling distribution of the statistic. The main characteristic of the sampling distribution of a statistic is that it approaches normal distribution even when the population distribution is not normal provided the sample size is sufficiently large (greater than 30). Another important feature of the sampling distribution of statistic is that the mean and the standard deviation of the sampling distribution of sample mean bear a definite relation to the corresponding parameters, i.e., mean and standard deviation of parent population. These characteristics of the sampling distribution help us:

(i) To estimate the unknown population parameter from the known statistic.
(ii) To set the confidence limits of the parameter within which the parameter values are expected to lie.

(iii) To test a hypothesis and to draw a statistical inference from it.

**SAMPLING DISTRIBUTION OF SAMPLE MEAN**

Theorem 1. If \( X_1, X_2, X_3, ..., X_n \) are \( n \) random samples drawn from a finite population of size \( N \) with mean \( \mu \) and variance \( \sigma^2 \), and \( \bar{X} \) is the mean of samples, then.

\[
(i) \quad \mu = \mathbb{E} \bar{X} \\

(a) \quad \text{Var} \, \bar{X} = \frac{N-n}{N-1} \frac{\sigma^2}{n}, \text{ if the sampling is without replacement} \\

(b) \quad \text{Var} \, \bar{X} = \frac{\sigma^2}{N}, \text{ if the sampling with replacement or the population is infinite, i.e., or if } N \to \infty
\]

**STANDARD ERROR OF A STATISTIC**

The statistical measure of standard deviation may be computed both from the observations of the population and also from the observations of a sample. Also, we know that the standard deviation is a measure of the average amount of the variability of all the observations of variable from their mean. When the average amount of the variability of the observations of a population is computed, it is called the standard deviation. But when the average amount of the variability of the observations of a sampling distribution of a statistic is computed, it is known as Standard Error. Thus the standard deviation computed from the observations of a sampling distribution of a statistic is called the standard error of the statistic. In other words, the standard deviation used to measure the variability of the values of a statistic from sample to sample is called the standard error of the statistic. Thus, the standard deviation and standard error have the same meaning and same connection but are used in different cases and different circumstances. Both of them are used to measure the variability of observations.

Standard error is used to measure the variability of the values of a statistic computed from the samples of the same size drawn from the population, whereas standard deviation is used to measure the variability of the observations of the population itself.

The following two standard errors are frequently used in statistics.
1. **Standard Error of Sample Mean.** It is the standard deviation of the sampling distribution of sample means. It is denoted by $\sigma_{\bar{x}}$ and is given by $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$, where $\sigma$ is the standard deviation of population and n is the sample size.

2. **Standard Error of Proportion.** It is computed from the proportions of all possible samples of the same size drawn from a population. It is denoted by $s_p$, and is given by

$$
S_p = \sqrt{\frac{P(1-P)}{n}}, \text{ if } P \text{ is known},
$$

where $P = \text{population proportion}, n = \text{sample size}$.

In other words, the standard deviation of the sampling distribution of a statistic ‘it’ is known as standard Error of T and is denoted by S.E. (t)

\[ S.E.(t) = \sqrt{\text{var}(t)} = \sqrt{\frac{1}{n} \sum (t - \bar{t})^2}; \]

where $n = \text{size of the sample}$.

The standard errors of the sampling distributions of some of the well-known statistic, where n is the sample size, $a$ is the population standard deviation, $P$ is the population proportion and $Q = 1 - P$; $n_1$ and $n_2$ represent the sizes of two samples respectively, is given below.

**REMARKS:**

1. The standard deviation of a statistic is termed as standard error. The standard error of $\bar{x}$ to be written as S.E.($\bar{x}$) and is equal to $\frac{\sigma}{\sqrt{n}}$, when sampling is with replacement.

and it is equal to $\frac{\sigma}{\sqrt{n}}, \sqrt{\frac{N-n}{N-1}}$, when sampling is without replacement.

2. S. F. ($\bar{x}$) is inversely proportional to the sample size. Thus larger the sample size, smaller is S.E. ($\bar{x}$) and consequently more efficient is $\bar{x}$ as an estimator of $\mu$.

3. The term $\sqrt{\frac{N-n}{N-1}}$ is termed as finite population correction (fpc,). We note that fpc tends to become closer and closer to unity as population size becomes larger and larger.

4. As a general rule, fpc may be taken to be equal to unity when sample size is less than 5% of population size, i.e., $n < 0.05N$. 

354  
**STATISTICS (CPT)**


**SAMPLING DISTRIBUTION OF PROPORTION**

Let us suppose that the population under study is classified only into two mutually exclusive and exhaustive classes, according as the given unit possesses or does not possess the given attribute, say, smoking, drinking, honesty, beauty etc., under consideration.

Let us consider a population consisting of \( N \) units and let the number of units possessing the attribute under study be ‘\( a \)’, (say).

\[ P = \frac{\text{Proportion of units (in the population) possessing the given attribute}}{a / N} \]

\[ \Rightarrow Q = 1 - P \quad \text{Proportion of population units which do not possess the given attribute}. \]

Let us take a **simple random sample without replacement** (srswor) of size \( n \) from this population. If \( X \) is the number of units possessing the given attribute in the sample, then we define:

\[ p = \text{proportion of sampled units possessing the given attribute} \Rightarrow p = \frac{X}{n}. \]

\[ \therefore q = 1 - p = \text{proportion of units (in the sample) which do not possess the given attribute}. \]

We shall now given some important results in simple random sampling without replacement from a finite population in the form of theorems.

**Theorem 2.** Sample proportion ‘\( p \)’ is an unbiased estimate of the population proportion \( P \). In other words \( E(p) = P \).

Thus the sample proportion ‘\( p \)’ can be used as a ‘point estimate’ for population proportion \( P \).

**Theorem 3.** (i) In case of simple random sample without replacement

\[ \text{Var}(p) \text{ in srswor} = \frac{N-n}{N-1} \cdot \frac{PQ}{n} \]

\[ \Rightarrow \text{S.E.}(p) \text{ in srswor} = \sqrt{\frac{N-n}{N-1}} \cdot \sqrt{\frac{PQ}{n}} \]

(ii) In case of simple random sample with replacement (srswoi) or in case of sampling from an infinite population, i.e., \( N \to \infty \) we have.

\[ \text{Var} (p) \text{ in srswr} = \frac{PQ}{n} ; \]

\[ \text{S.E.}(p) \text{ in srswr} = \sqrt{\frac{PQ}{n}} \]
Remark. In practice, the population proportion $P$ and consequently $Q = 1 - P$ are unknown. In that case, an unbiased estimate of $\text{Var}(P)$ is provided by $\text{Var}(p)$. It is defined as follows:

I. In srswr,

$$\text{Var}(p) = E[\text{Var}(p)] = \frac{(N - n) pq}{N(n-1)} \Rightarrow S.E.(p) = \sqrt{\frac{N - n}{N(n-1)}} \frac{pq}{n}$$

II. In srswr or in sampling from an infinite population,

$$\text{Var}(p) = E[\text{Var}(p)] = \frac{pq}{n} \Rightarrow S.E.(p) = \sqrt{\frac{pq}{n}}$$

BASIC STATISTICAL LAWS

Sample survey is the study of the unknown population on the basis of a proper sample drawn from it. The four popular and well defined statistic laws which indicate the utility of larger size of the sample for the purpose of reducing sampling errors are:

1. Law of Statistical Regularity

2. Law of Inertia of Large Numbers

3. Principle of Optimization

4. Principle of Validity

I. Law of Statistical Regularity:

It states that a reasonably larger number of items selected at random from a large group of items, will on the average, represent the characteristics of the group. In the words of the statistician W. I. King, “The law of statistical regularity lays down that a moderately large number of items chosen at random from a large group, are almost sure (on the average) to possess the characteristics of the large group “.

This law explains that if a reasonable large sample is selected at random without bias (i.e., probability sampling), it is almost certain that on an average, the sample so chosen, shall have the same characteristics as those of the parent population from where the units constituting the sample have been drawn. It is on the basis of this theory that the law of statistical theory tells us that a random selection is very likely to give a representative sample.

II. Law of Inertia of Large Numbers:

It states that “large groups or aggregate of data show high degree of stability because there is a greater possibility that the extremes on one side are compensated by the extremes on the other side.”
The law of inertia of large number is a corollary to the law of statistical regularity. It emphasizes the fact that large numbers are relatively more stable and more reliable than small ones. In a large number, it is unlikely that the data would move in only one direction. Thus, the greater the size of the sample, the greater will be the compensation or tendency to neutralise one another and consequently more stable would be the result. For example, the birth rate, death rate etc. may vary from place to place in India but in India as a whole country, they will be found somewhat stable over a number of years.

The simplest method of increasing the accuracy of a sample is to increase its size. The larger the size of the sample, the more reliable is the result. The other things remaining unchanged, the sampling is inversely proportional to the root of the number of items in the sample.

The Law of Statistical Regularity and the Law of the Inertia of Large Numbers have great importance in the theory of sampling as the sampling error is reduced to minimum if these laws are correctly applied.

III. Principle of Optimization

The principle of optimization ensures that an optimum level of efficiency at a minimum cost or the maximum efficiency at a given level of cost can be achieved with the selection of an appropriate sampling design.

IV. Principle of Validity

The principle of validity states that a sampling design is valid only if it is possible to obtain valid estimates and valid tests about population parameters. Only a probability sampling ensures this validity.

UTILITY OF STANDARD ERROR OF STATISTIC

The standard error of sample mean and the standard error of proportion are used in sampling in order to obtain the following facts:

1. Standard error is used to set up the confidence limits within which the population parameter may lie.
2. Standard error is used to test the hypothesis and to draw a statistical conclusion from it.
3. Standard error is used to measure the variability of the values of a statistic from its mean.
4. It is generally used for large samples and gives us the idea about the average amount of error which actually occurs in estimating the values of a parameter on the basis of a statistic.

**ESTIMATION**

It is possible to draw valid conclusion about the population parameters from sampling distribution. We have a sample from a population involving unknown parameter, such as the sample mean. The problem is to construct a sample quantity that will serve to estimate the unknown parameter, viz, population mean \( \mu \). Such a sample quantity is called the estimator and the actual numerical value obtained by evaluating an estimator in a given instance is the estimate. For example, the sample mean \( \bar{X} \) is an estimator of the population mean \( \mu \). If for a specific sample, the sample mean is 7.85, then 7.85 is our estimator for the population mean \( \mu \).

It is important to note that an estimator must be a statistic and it must depend only on the sample and not on the parameter to be estimated. An estimator is a statistic which for all practical purposes, can be used in place of unknown parameter of the population.

Estimators are bound to differ from the true value of the population parameters. But the tolerable divergence between the estimated and the true value of the population parameter may be specified beforehand.

**CHARACTERISTICS OF A GOOD ESTIMATOR**

A good estimator is one which is as close to the true value of the parameter as possible. A good estimator must possess the following characteristics:

(i) Unbiasedness;
(ii) Consistency;
(iii) Efficiency;
(iv) Sufficiency.

**Unbiasedness.** A statistic \( t = t(x_1, x_2, x_3, \ldots, x_n) \) is said to be an unbiased estimate of the corresponding population parameter \( \theta \), if \( E(t) = \theta \) \( \Rightarrow \) the mean value of the sampling distribution of the statistic \( t \) is equal to the parameter of the population.

**Consistency.** A statistic \( t = t_n = t(x_1, x_2, x_3, \ldots, x_n) \) based on the sample size \( n \) is said to be a consistent estimator of the parameter \( \theta \) if \( t_n \rightarrow \theta \) as \( n \rightarrow \infty \)

Symbolically,

\[
\lim_{n \to \infty} P(t_n \to \theta) = 1
\]

**Efficiency.** If \( t_1 \) and \( t_2 \) are two consistent estimators of a parameter \( \theta \) such that \( \text{Var}(t_1) < \text{Var}(t_2) \) for all samples of size \( n \), then \( t_1 \) is said to be more efficient than \( t_2 \). In other words, an estimator with lesser variability is said to be more efficient and consequently more reliable than the other.
**Sufficiency.** A statistic \( t = t(x_1, x_2, x_3, ..., x_n) \) is said to be a sufficient estimator of parameter \( \theta \) if it contains all the information in the sample regarding the parameter. In other words, if \( t = t(x_1, x_2, x_3, ..., x_n) \) is a statistic based on a random sample \( (x_1, x_2, x_3, ..., x_n) \) of size \( n \) from a population with probability density function P.D.F = \( P(x, q) \), then it is a sufficient estimator of \( q \), if the conditional probability \( P\left[\frac{x_1 \cap x_2 \cap x_3 \cap x_4 \cap x_n}{t = k}\right] \) does not depend on \( q \), i.e., \( P(x, q) = P(x) \).

**TYPES OF ESTIMATION**

Estimation is divided into two groups:

(i) **Point Estimation**

(ii) **Interval Estimation**

**Point Estimation**

In point estimation a single statistic is used to provide an estimate of the population parameter. In other words, the estimate of a population parameter given by a single number is called the point estimation of the parameter. In point estimation, we find a statistic which may be used for to replace an unknown parameter of the population for all practical purposes.

**Interval Estimation**

There are situations where the point estimation is not desirable and we are interested in finding such limits within which with a known probability or to a known degree of reliability, the value of the population parameter is expected to lie. Such a process of estimation is called the interval estimation. In other words, an interval estimation is the range of values used in making estimation of a population parameter. Thus the interval estimation of the population parameter is the estimation of the population parameter by an interval around a point. The interval estimation of a population parameter \( \theta \) is the estimation of the parameter \( q \) with the help of the interval \([t - s, t + s]\), where \( t \) is the sample statistic, i.e., \( t - s \leq q \leq t + s \).

Both point and interval estimations are the methods of drawing inductive references and therefore, involve an element of uncertainty. But the interval estimation has an advantage over the point estimation as it provides a measure of degree of uncertainty in terms of probability attached to the interval. In the interval estimation, the estimate for the parameter lies between two limits. The two limits within which the estimate for the parameter lies are known as Confidence Limits or Fiducial Limits and the interval bounded by these two limits as Confidence Interval or Fiducial Interval. The confidence interval depends upon the confidence that is required to set up. The probability that we associate with an interval is called the confidence...
level. It shows how confidently we can say that the interval estimate will include the population parameter. The higher the probability, the more is the confidence. Although any confidence level can be considered, but the most commonly used confidence levels are 90%, 95%, 98% and 99%. In the light of above discussion, the interval estimation is defined as:

\[ P \left[ C_1 < q < C_2 \right] = 1 - \alpha \], where \( \alpha \) is the level of significance. \( [C_1, C_2] \) is the interval within which the unknown parameter \( q \) is expected to lie. It is known as Confidence Interval or Fiducial Interval. \( C_1, C_2 \) are respectively known as lower limit and upper limit of the confidence interval \( [C_1, C_2] \) and \( 1 - \alpha \) is called the confidence coefficient, depending upon the desired precision of the estimate (For example \( \alpha = 0.01 \) gives 99% confidence limits).

If ‘t’ is the sample statistic used to estimate the corresponding population parameter \( q \), then \( (1 - \alpha) \% \) confidence limits for \( q = t \pm S.E.(t) \times t \alpha \), where S.E. (t) is the standard error of \( t \), and \( t \alpha \) the significant or critical value at significance level \( \alpha \). Also \( t + S.E.(t) \times t \alpha \) is the upper limit and \( t - S.E.(t) \times t \alpha \) is the lower limit of confidence interval

\[ [t - S.E. (t) \times t \alpha , t + S.E. (t) \times t \alpha ] \]

**CONFIDENCE INTERVAL OF THE MEAN**

Let \( \mu \) be the population mean and \( \bar{x} \) be the sample mean of the sampling distributions of means. It is also assumed that the sample mean is normal if the sample is large. Then the interval estimate of population mean \( \mu \) by the sample mean \( \bar{x} \) of the sampling distribution of means is given by the following rule:

**WORKING RULE**

**Step I.** Compute \( \bar{x} \) or take \( \bar{x} \).

**Step II.** Select the confidence level and the corresponding to that specific level of confidence, find, from the table, the critical value \( Z \alpha \) or \( t \alpha \).

**Step III.** Compute S.E. (\( \bar{x} \)) with the help of following results:

**Case I.** When \( \sigma \), the standard deviation of the normal population is known.

\[ S.E. (\bar{x}) = \frac{\sigma}{\sqrt{n}}, \quad \text{where} \ n \ is \ the \ sample \ size. \]

**Case II.** When \( \sigma \), the standard deviation of population is not known.

\[ S.E. (\bar{x}) = \frac{s}{\sqrt{n-1}}, \]

where \( n = \) Sample size and \( n \) is large.
s = standard deviation of the sampling distribution of sample.

**Step IV.** Construct a confidence level as follows:

**Case I.** When $\sigma$ is known and population is normal or any population with large n.

Confidence Interval = $[\bar{x} - \text{S.E.}(\bar{x}) \times Z_{\alpha}, \bar{x} + \text{S.E.}(\bar{x}) \times Z_{\alpha}]$

or $[\bar{x} - \frac{\sigma}{\sqrt{n}} \times Z_{\alpha}, \bar{x} + \frac{\sigma}{\sqrt{n}} \times Z_{\alpha}]$ 

$\therefore$ S.E. $(\bar{x}) = \frac{\sigma}{\sqrt{n}}$

**Case II.** When $\sigma$ is unknow and n is large.

In this case, the Confidence Interval = $[\bar{x} - \text{S.E.}(\bar{x}) \times Z_{\alpha}, \bar{x} + \text{S.E.}(\bar{x}) \times Z_{\alpha}]$

$= [\bar{x} - \frac{s}{\sqrt{n-1}} \times Z_{\alpha}, \bar{x} + \frac{s}{\sqrt{n-1}} \times Z_{\alpha}]$

or $[\bar{x} - \frac{s}{\sqrt{n-1}} \times Z_{\alpha}, \bar{x} + \frac{s}{\sqrt{n-1}} \times Z_{\alpha}]$ 

$\therefore$ S.E. $(\bar{x}) = \frac{s}{\sqrt{n-1}}$

**Step V.** Select the value of $Z_{\alpha}$ from the following table

<table>
<thead>
<tr>
<th>Confidence level</th>
<th>Value of confidence coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(1 - \alpha)$</td>
<td>$Z_{\alpha}$</td>
</tr>
<tr>
<td>90%</td>
<td>1.64</td>
</tr>
<tr>
<td>95%</td>
<td>1.96</td>
</tr>
<tr>
<td>98%</td>
<td>2.33</td>
</tr>
<tr>
<td>99%</td>
<td>2.58</td>
</tr>
</tbody>
</table>

Without any reference to the confidence level 3.00

**Note:** Where no reference to the confidence level is given, then always take $Z_{\alpha} = 3$.

**INTERVAL ESTIMATE OF THE PROPORTION**

Population proportion. The population proportion $P$ is the ratio of the number of elements possessing a characteristic to the total number of elements in the population, i.e.,
Population proportion: \( P = \frac{\text{Number of elements possessing the characteristic}}{\text{Total number of elements in the Population}} \)

If we multiply the population by 100, then we get the percentage and we may make use of percentage for the proportion and vice-versa.

Sample Proportion. The sample proportion \( p \) is the ratio of the number of elements possessing a characteristic to the total number of elements ‘\( n \)’ in the sample.

\[
\text{Sample Proportion: } p = \frac{\text{Number of elements possessing the characteristic}}{\text{Total number of elements in the sample (n)}}
\]

It is important to note that the mean of sampling distribution of \( p \) equals the population proportion, i.e., \( E(p) = P \).

**Standard Error (or Standard Deviation) of Sample Proportion**

Case I. When the population size is infinitely large or the sample is drawn with replacement.

(i) Standard Error of Sample Proportion is

\[
S.E. (p) = \sqrt{\frac{PQ}{n}} \text{ if } P \text{ is known, where } Q = 1 - P, n = \text{the sample size.}
\]

(ii) Standard Error of Sample Proportion is

\[
S.E. (p) = \sqrt{\frac{pq}{n}} \text{ if } P \text{ is not known,}
\]

\( n \) = sample size and \( p \) = sample proportion where \( q = 1 - p \).

It is an estimated value.

Case II. When the population size is finite and the sample is drawn without replacement.

(i) Standard Error of Sample Proportion

\[
S.E.(p) = \sqrt{\frac{PQ}{n}} \times \sqrt{\frac{N - n}{N - 1}} \text{, if } P \text{ is known,}
\]

where \( Q = 1 - P \), \( N \) = Size of population, \( n \) = Sample size.

(ii) Standard Error of Sample Proportion

\[
S.E.(p) = \sqrt{\frac{pq}{n}} \times \sqrt{\frac{N - n}{N - 1}} \text{ if } P \text{ is not known,}
\]

\( N \) = size of population, \( n \) = size of sample ; \( p \) = sample proportion, and \( q = 1 - p \).

**Confidence Interval Estimate**

Confidence Interval Estimate of the population proportion \( P \) is given by the following results:
Case I. When \( P \) is known. In this case, the confidence interval of the sample proportion is given by
\[
(P - Z_\alpha \times S.E. (p), P + Z_\alpha \times S.E. (p))
\]
where \( Z_\alpha \) is the confidence coefficient, and S.E. (p) is the standard error of p.

Case II. When \( P \) is not known. In this the confidence interval of the sample proportion is given by
\[
[p - S.E. (p) \times Z_\alpha, p + S.E. (p) \times Z_\alpha]
\]
where \( Z_\alpha \) = confidence coefficient; Standard Error of p = S.E. (p).

**DETERMINATION OF SAMPLE SIZE**

The determination of sample size for estimating a mean or proportion is a crucial question. By selecting a sample size lower than the correct size may affect reliability and a higher one will mean more cost and time. The determination of the size of a sample is the most important factor for the purposes of estimation of the value of the population parameters. We have the following formula for it.

**Sample Size for Estimation a Mean**

In order to determine the sample size for estimating a population mean, the following factors must be known:

(i) The desired confidence level.

(ii) The permissible sampling error \( E = \bar{x} - \mu \).

(iii) The standard deviation \( \sigma \).

After having known the above mentioned three factors, the size of sample mean \( n \) is given by
\[
n = \left( \frac{\sigma Z}{E} \right)^2
\]

**Sample Size for Estimating a Proportion**

In this case, we must know the following three factors:

(i) the desired confidence level.

(ii) the permissible sampling error \( E = \) difference between the estimate from the sample \( p \) and the parameter \( P \) to be estimated = \( P - p \).

(iii) the estimated true proportion of success.

The sample size \( n \) is given by
\[
n = \frac{Z^2 pq}{E^2} \text{ where } q = 1 - p.
\]

***************
1. Sampling errors are:
   (a) caused by inaccurate measurement  (b) the result of the chance selection of the sampling units.
   (c) of no great concern.  (d) larger for a census than for a sample.

2. Non-sampling errors are:
   (a) caused by inaccurate measurement.  (b) the result of the chance selection of the sampling units.
   (c) of no great concern.  (d) always larger for a census than for a sample.

3. If \( \mu_x \) is the population mean, and \( \sigma^2 \) is the population variance, then the mean and variance of a sample mean are equal to
   (a) \( \mu_x \) and \( \sigma^2 \)
   (b) \( \mu_x/n \) and \( \sigma^2/n \)
   (c) \( \mu_x/n \) and \( \sigma^2/n^2 \)
   (d) \( \mu_x \) and \( \sigma^2/n \).

4. A sample consists of
   (a) all units of the population.  (b) 50 per cent units of the population.
   (c) 5 per cent units of the population.  (d) any fraction of the population.

5. Probability of selection varies at each subsequent draw in
   (a) sampling without replacement.  (b) sampling with replacement.
   (c) both (A) and (B).  (d) neither (A) nor (B).

6. The number of possible samples of size \( n \) from a population of \( N \) units with replacement is
   (a) \( n^2 \)  (b) \( N^2 \)
   (c) \( \infty \)  (d) none of these.

7. A function of variates for estimating a parameter is called
   (a) an estimate  (b) an estimator
   (c) a frame  (d) a statistic.

8. Let the standard error of an estimator \( T \) under srswor is more than the standard error of \( T \) under stratified randomly sampling. Then \( T \) under stratified sampling as compared to \( T \) under srswor is:
   (a) more reliable  (b) equally reliable.
   (c) less reliable.  (d) not comparable.
9. Which of the following basis distinguishes cluster sampling from stratified sampling?
(a) A sample is always drawn from each (b) Clusters are preferably heterogeneous whereas stratum whereas no sample of elementary units is drawn from clusters.
(c) Small size clusters are better whereas there is no such restriction for stratum size.
(d) all of these.

10. Non-response in surveys mean
(a) non-return of questionnaire by the respondents.
(b) non-availability of respondents.
(c) refusal to give information by the respondents.
(d) all of these.

11. Choose the pair of symbols that best completes this sentence: ______ is a parameter, whereas ______ is a static.
(a) N, μ (b) N,n
(c) σ, s (d) all of these.

12. Regarding the number of strata, which statement is true?
(a) More the number of strata, poorer it is.
(b) Lesser the number of strata, better it is.
(c) More the number of strata, better it is.
(d) Not more than ten items should be there in a stratum.

13. The magnitude of the standard error of an estimate is an index of its
(a) accuracy (b) precision
(c) efficiency (d) all of these.

14. An estimate based on a fixed set of values of a sample always possess
(a) a single value (b) any value
(c) a value equal to one (d) all of these.

15. If the sample values are 1, 3, 5, 7, 9, the standard error of sample mean is:
(a) S.E. = √2 (b) S.E. = 1/√2
(c) S.E. = 2.0 (d) S.E. = 1/2.

16. What precaution(s) make(s) cluster sampling more efficient?
(a) Choosing clusters having largest within variation.
(b) By taking clusters of small size.
(c) Choosing clusters having least variation clusters.
(d) All of these.
17. A population is divided into sub population and it has been found that all items within a sub population are alike. Which of the following sampling procedures would you adopt?
(a) Cluster sampling  (b) Simple random sampling.
(c) Systematic sampling.  (d) Stratified sampling.

18. In which of the following situations would \( \sigma_x = \frac{\sigma}{\sqrt{n}} \) be the correct formula to use for computing \( \sigma_x \)?
(a) Sampling is from a finite population with replacement.
(b) Sampling is from an infinite population.
(c) Sampling is from a finite population without replacement.
(d) (A) and (B) only.

19. If population variance of an infinite population is \( \sigma^2 \) and a sample of \( n \) items is selected from this population, the standard error of sample mean is equal to
(a) \( \sigma / n \)  (b) \( \sigma^2 / n \)
(c) \( \sigma / \sqrt{n} \)  (d) \( \sigma \)

20. A population is perfectly homogeneous in respect of a characteristic. What size of sample would you prefer?
(a) A large sample  (b) A small sample
(c) A single item  (d) No item.

21. A systematic sample does not yield good results if:
(a) variation in units is periodic.  (b) units at regular intervals are corrected.
(c) both (A) and (B).  (d) none of (A) and (B).

22. Double sampling is also known as
(a) two stage sampling.  (b) two phase sampling.
(c) two directional sampling.  (d) all of these.

23. The number of possible samples of size \( n \) out of \( N \) population units without replacement is
(a) \( ^nC_n \)  (b) \( (N)_n \)
(c) \( n^2 \)  (d) \( n! \).

24. Probability of any one sample of size \( n \) being drawn out of \( N \) units is:
(a) \( 1 / N \)  (b) \( n / N \)
(c) \( 1 / n! \)  (d) \( 1/(\binom{N}{n}) \)

25. The discrepancy between estimates and population parameters is known as
(a) human error.  (b) enumeration error.
(c) sampling error.  (d) formula error.
26. A selection procedure of a sample having no involvement of probability is known as:
   (a) judgement sampling.  (b) purposive sampling.
   (c) subjective sampling.  (d) all of these.

27. Sampling is inevitable in the situation(s).
   (a) when the population is infinite.  (b) blood test of a person.
   (c) testing of life of dry battery cells.  (d) all of these.

28. An unordered sample of size \( n \) can occur in:
   (a) \( n \) ways  (b) \( n! \) ways
   (c) \( n^2 \) ways  (d) one way.

29. Suppose that a population with \( N = 144 \) has \( \mu = 24 \). What is the mean of the sampling distribution of the mean for samples of size 25?
   (a) 24  (b) 2
   (c) 4.8  (d) Cannot be determined from the information given.

30. In simple random sampling with replacement, the same sampling unit may be included in the sample:
   (a) only twice  (b) only once
   (c) more than once  (d) none of these.

31. If each and every unit of a population has equal chance of being included in the sample, it is known as
   (a) purposive sampling.  (b) restricted sampling.
   (c) subjective sampling.  (d) unrestricted sampling.

32. The errors emerging out of faulty planning of surveys are categorised as:
   (a) non-sampling errors.  (b) non-response errors.
   (c) absolute error.  (d) sampling errors.

33. Probability of drawing a unit at each selection remains same in
   (a) srswor.  (b) srswr.
   (c) both (A) and (B).  (d) none of (A) and (B).

34. The central limit theorem assures us that the sampling distribution of the mean:
   (a) is always normal.  (b) is always normal for large sample sizes.
   (c) approaches normality as sample size increases.  (d) appears normal only when \( N \) is greater than 1,000.
35. Simple random, sample can be drawn with the help of:
   (a) random number tables  (b) chit method
   (c) roulette wheel        (d) all of these.

36. Increase in reliability and accuracy of results from a sampling study with the increase in sample size is known as the principle of:
   (a) statistical regularity. (b) optimisation.
   (c) law of increasing returns. (d) inertia of large numbers.

37. A population consisting of all the items which are physically present is called
   (a) hypothetical population. (b) real population.
   (c) infinite population.    (d) none of these.

38. A population consisting of all real numbers is an example of:
   (a) an infinite population. (b) a finite population.
   (c) an imaginary population. (d) none of these.

39. The number of all possible samples of size two from a population of 4 units is
   (a) 2  (b) 6
   (c) 8  (d) 12.

40. Probability of including a specified unit in a sample of size n selected out of N units is
    (a) 1/ n  (b) 1/ N
    (c) n/ N  (d) N/ n.

41. To meet requirement of the principle of validity of sampling methods, one must adopt:
    (a) purpose sampling. (b) restricted sampling.
    (c) probability sampling. (d) none of these.

42. An estimator can possess:
    (a) a fixed value  (b) any value
    (c) both (A) and (B) (d) none of these.

43. The finite population multiplier does not have to be used when the sampling fraction is:
    (a) greater than 0.05. (b) greater than 0.50.
    (c) less than 0.05.    (d) none of these.

44. If the items are destroyed under investigation, we have to go for
    (a) complete enumeration  (b) sampling studies
    (c) both (A) and (B)    (d) none of these.

45. Systematic sampling means:
    (a) selection of n contiguous units.  (b) selection of n units situated at equal distances.
    (c) selection of n largest units.     (d) none of these.
46. The standard error of the mean for a sample size of two or more is:
   (a) always greater than the standard deviation of the population.
   (b) generally, greater than the standard deviation of the population.
   (c) usually, less than the standard deviation of the population.
   (d) none of these.

47. In which of the following situation(s) cluster sampling is appropriate?
   (a) When the units are situated for apart.
   (b) When sampling frame is not available.
   (c) When all the elementary units are not easily identifiable.
   (d) All of these.

48. The selected items of a sample resulted into same values pertaining to a character. The variance of the sample is:
   (a) 2
   (b) 0
   (c) 1
   (d) none of these.

49. A sample of 16 items from an infinite population having S.D. = 4, yielded total scores as 160. The standard error of sampling distribution of mean is:
   (a) 1
   (b) 20
   (c) 30
   (d) none of these.

50. A border patrol checkpoint which stops every passenger van is utilising
   (a) simple random sampling.
   (b) systematic sampling.
   (c) stratified sampling.
   (d) complete enumeration.

51. Selected units of a systematic sample are:
   (a) not easily locateable.
   (b) easily locateable.
   (c) not representing the whole population.
   (d) all of these.

52. Greatest drawback of systematic sampling is that:
   (a) one requires a large sample.
   (b) data are not easily accessible.
   (c) no single reliable formula for standard error of mean is available.
   (d) none of these.

53. In a normally distributed population, the sampling distribution of the mean
   (a) is normally distributed
   (b) has a mean equal to the population mean.
   (c) has standard deviation equal to the population standard deviation divided by the square root of the sample size.
   (d) all of these.
54. Under proportional allocation, the size of the sample from each stratum depends on:
   (a) size of the stratum.  (b) total sample size.
   (c) population size.  (d) all of these.

55. Which of the following statements is true?
   (a) All sampling procedures involve sampling with constant probability.
   (b) There exists sampling procedure in which the units are selected with varying probability.
   (c) Every selection procedure of a sample involves probability.
   (d) none of these.

56. The central limit theorem:
   (a) requires some knowledge of the frequency distribution.
   (b) permits us to use sample statistics to make inferences about population parameters.
   (c) relates the shape of a sampling distribution of the mean to the mean of the sample.
   (d) requires a sample to contain fewer than 30 observations.

57. Under equal allocation in stratified sampling, the sample from each stratum is:
   (a) proportional to stratum size.
   (b) of same size from each stratum.
   (c) in proportion to the per unit cost of survey of the stratum.
   (d) none of these.

58. The errors in a survey other than sampling errors are called
   (a) planning error  (b) formula errors
   (c) non-sampling error  (d) none of these.

59. Stratified sampling belongs to the category of:
   (a) subjective sampling.  (b) judgement sampling.
   (c) controlled sampling.  (d) non-random sampling

60. There are more chances of non-sampling errors than sampling errors in case of:
   (a) studies of large samples.  (b) complete enumeration.
   (c) inefficient investigators.  (d) all of these.

61. Stratified sampling comes under the category of:
   (a) unrestricted sampling.  (b) subjective sampling.
   (c) purposive sampling.  (d) restricted sampling.

62. Sampling error can be reduced by:
   (a) choosing a proper probability sampling,
   (b) selecting a sample of adequate size.
(c) using a suitable formula for estimation.
(d) all of these.

63. Which of the following statement is correct?
(a) Simple random sample is inferior than systematic sample.
(b) Systematic sample is superior than stratified random sample.
(c) Stratified random sample is better than systematic sample.
(d) None of these.

64. In what situation(s), a systematic sample is more preferred than others?
(a) When the items are in row.
(b) When the data are on cards.
(c) When the items situated at equal distances are uncorrelated.
(d) all of these.

65. Which of the following advantage of systematic sampling you approve?
(a) Economical
(b) Easy selection of sample
(c) Spread of sample over the whole population
(d) all of these.

66. Sampling is the process of obtaining a
(a) population
(b) sample
(c) frequency
(d) none of these.

67. Sampling can be described as a statistical procedure
(a) to infer about the unknown universe from a knowledge of any sample.
(b) to infer about the known universe from a knowledge of a sample drawn from it.
(c) to infer about the unknown.
(d) Both (A) and (B).

68. A parameter is a characteristic of
(a) Population
(b) Sample
(c) Both (A) and (B)
(d) none of these.

69. Statistical decision about an unknown universe is taken on the basis of
(a) sample observations.
(b) a sampling frame,
(c) sample survey.
(d) compute enumeration.

70. Statistical data may be collected by complete enumeration called
(a) census inquiry.
(b) sample inquiry.
(c) Both
(d) none
71. A sample survey is prone to:
   (a) Non-sampling errors.  (b) Sampling errors.
   (c) Either (A) or (B).  (d) Both (A) and (B).

72. A sample is a selected part of the
   (a) estimation  (b) population
   (c) both  (d) none of these.

73. Two basic Statistical laws concerning a population are
   (a) the law of statistical irregularity and the law of inertia of large numbers.
   (b) the law of statistical regularity and the law of inertia of large numbers.
   (c) the law of statistical regularity and the law of inertia of small numbers.
   (d) the law of statistical irregularity and the law of inertia of small numbers.

74. Sampling Fluctuations may be described as:
   (a) the variation in the values of a sample.
   (b) the variation in the values of a sample.
   (c) the differences in the values of a parameter.
   (d) the variation in the values of observations.

75. The Law of Statistical Regularity says that:
   (a) Sample drawn from the population under discussion possesses the characteristics of the population.
   (b) A large sample drawn at random from the population would possess the characteristics of the population.
   (c) A large sample drawn at random from the population would possess the characteristics of the population on an average.
   (d) An optimum level of efficiency can be attained at a minimum cost.

76. Which sampling is subjected to the discretion of the sampler?
   (a) Simple random sampling  (b) Systematic sampling
   (c) Purposive sampling  (d) Quota sampling.

77. Which sampling provides separate estimates for population means for different segments and also overall estimate?
   (a) Multistage sampling  (b) Stratified sampling
   (c) Systematic sampling  (d) Simple random sampling

78. The difference of the actual value and the expected value using a model is:
   (a) Error in statistics.  (a) Absolute error.
   (c) Percentage error.  (d) Relative error.
79. Which sampling is affected most if the sampling frame contains an undetected periodicity?
   (a) Simple random sampling. (b) Multistage sampling.
   (c) Stratified sampling. (d) Systematic sampling.

80. Sample mean is an example of:
   (a) parameter  (b) statistic
   (c) both  (d) none.

81. Large sample is that sample whose size is:
   (a) greater than 30. (b) greater than or equal to 30.
   (c) less than 20. (d) less than or equal to 30.

82. Population mean is an example of:
   (a) parameter  (b) statistic
   (c) both (A) and (B) (d) none.

83. The finite population multiplier is ignored when the sampling fraction is
   (a) greater than 0.05. (b) less than 0.6.
   (c) less than 0.05. (d) greater than 0.6.

84. The ways of selecting a sample are:
   (a) Random sampling (b) Multi-stage sampling
   (c) both (a) and (b) (d) none of these.

85. Random sampling implies
   (a) Probability sampling (b) Haphazard sampling
   (c) Systematic sampling (d) Sampling with the same probability for each unit.

86. Simple random sampling is
   (a) a probabilistic sampling (b) a mixed sampling.
   (c) a non-probabilistic sampling. (d) Both (B) and (C).

87. If random sampling with replacement is applied, then the mean of sample means will be ______ the population mean
   (a) greater than (b) less than :
   (c) exactly equal to (d) none of these.

88. Simple random sampling is very effective if
   (a) the population is not very large.
   (b) the population is not much heterogeneous.
89. The number of types of random sampling is:
   (a) 2  (b) 3  (c) 1  (d) 4.

90. Random sampling is called lottery sampling.
   (a) True  (b) False  
   (c) Both (A) and (B)  (d) none of these.

91. Stratified random sampling is: appropriate when the universe is not homogeneous.
   (a) True  (b) False  
   (c) Both (A) and (B)  (d) none of these.

92. Random numbers are also called Random sampling number.
   (a) True  (b) False  
   (c) Both (A) and (B)  (d) none of these.

93. Cluster sampling is ideal in case the data are widely scattered.
   (a) True  (b) False  
   (c) Both (A) and (B)  (d) none of these.

94. In stratified sampling, the sampling is subdivided into several parts, called
   (a) strata.  (b) strati.  
   (c) start.  (d) none of these.

95. Deliberate sampling is free from bias.
   (a) True  (b) False  
   (c) Both (A) and (B)  (d) none of these.

96. Purposive selection is resorted to in case of judgment sampling.
   (a) True  (b) False  
   (c) Both (A) and (B)  (d) none of these.

97. The ratio of the number of elements possessing a characteristic to the total number of elements
    in a sample is known as:
   (a) characteristic proportion  (b) sample proportion  
   (c) Both (A) and (B)  (d) none of these.

98. Finite population multiplier is:
   (a) square of \((N - 1) / (N - n)\)  (b) square root of \((N - n) / (N - 1)\)  
   (c) square root of \((N - 1) / (N - n)\)  (d) square of \((N - n) / (N - 1)\).
99. Standard error can be described as
   (a) the error committed in sampling.
   (b) the error committed in sample survey.
   (c) the error committed in estimating a parameter.
   (d) Standard deviation of a statistic.

100. A measure of precision obtained by sampling is given by :
   (a) Standard error  (b) Expectation
   (c) Sampling distribution  (d) Sampling fluctuation.

101. The standard deviation in the sampling deviation is called:
   (a) standard error.  (b) absolute error.
   (c) relative error.  (d) none of these.

102. As the sample size increases, standard error
   (a) increases  (b) decreases
   (c) remains constant  (d) decreases proportionately.

103. A population comprises 5 members. The number of all possible samples of size 2 that can be
     drawn from it with replacement is :
     (a) 100  (b) 25  (c) 125  (d) 25.

104. Standard error of mean may be defined as the standard deviation in the sampling distribution
     of :
     (a) mean  (b) median  (c) mode  (d) none of these.

105. If from a population with 25 members, a random sample without replacement of 2 members is
     taken, the number of all such samples is :
     (a) 300  (b) 725  (c) 150  (d) 540.

106. The sample proportion is taken as an estimate of the population proportion of defectives.
     (a) True  (b) False
     (c) both (A) and (B)  (d) none of these.

107. Standard deviation of a sampling distribution is itself the standard error.
     (a) True  (b) False
     (c) both (A) and (B)  (d) none of these.

108. The standard error of the mean for finite population is very close to the standard error of the
     mean for infinite population when the sampling fraction is
     (a) small  (b) moderate
     (c) large  (d) none of these.
109. Sampling error increases with an increase in the size of the sample.
   (a) True  (b) False
   (c) both (A) and (B)  (d) none of these.

110. Testing the assumption that an assumed population is located at a known level of significance is known as
   (a) confidence testing  (b) point estimation
   (c) interval estimation  (d) hypothesis testing.

111. The standard deviation of the _____ distribution is called standard error.
   (a) Normal  (b) Poisson
   (c) Binomial  (d) Sampling.

112. Under_______ method selection is often based on certain predetermined criteria.
   (a) Area sampling  (b) Block or Cluster sampling
   (c) Quota sampling  (d) Deliberate, purposive or judgment sampling.

113. Which would you prefer for   when “The universe is large”?
   (a) Full enumeration  (b) Sampling
   (c) both (A) and (B)  (d) none of these.

114. A ------------ is a complete or whole set or possible measurements data corresponding to the entire collection of units.
   (a) Sample  (b) Population
   (c) both (A) and (B)  (d) none of these.

115. The finite population correction factors should be used when the population is
   (a) infinite  (b) finite and large
   (c) finite and small  (d) none of these

116. A statistic is a ........ variable
   (a) compound  (b) simple
   (c) random  (d) none of these

117. Which would you prefer for .... “The Statistical inquiry is in depth”.
   (a) Full enumeration  (b) Sampling
   (c) both (A) and (B)  (d) none.

118. The primary object of sampling is to obtain ....... information about population with ....... effort.
   (a) maximum, minimum  (b) minimum, maximum
   (c) some,less  (d) none.
119. The measure of divergence is ........ as the size of the sample approaches that of the population.
   (a) more  (b) less
   (c) same  (d) none of these.

120. For ........ samples, the sample proportion is an unbiased estimate of the population proportion.
   (a) large  (b) small
   (c) moderate  (d) none of these.

121. Value of a ........ is different for different samples.
   (a) statistic  (b) skill
   (c) both (A) and (B)  (d) none of these.

122. Sampling error is ______ proportional to the square root of the number of items in the sample.
   (a) inversely  (b) directly
   (c) equally  (d) none.

123. Which would you prefer for _____ “Where testing destroys the quality of the product”.
   (a) Full enumeration  (b) Sampling
   (c) both (A) and (B)  (d) none of these.

124. The distribution of sample ........ is normally or approximately normally distributed about the
     population
   (a) median  (b) mode
   (c) mean  (d) none of these.

125. The standard error of the ........ is the standard deviation of sample means
   (a) mode  (b) median
   (c) mean  (d) none of these.

126. ----------- sampling is the most appropriate in cases when the population is more or less
     homogeneous with respect to the characteristic under study.
   (a) Stratified  (b) Multi - stage
   (c) Random  (d) none of these.

127. The mean of the sampling distribution of sample proportion is ----------- the population proportion.
   (a) greater than  (b) less than
   (c) equal to (d) none of these.

128. ----------- sampling is similar to cluster sampling.
   (a) Judgment  (b) Quota
   (c) Area  (d) none of these.
129. Which would you prefer when Previous experiences reveals a low rate of error.
   (a) Larger Sample  (b) Small sample
   (c) both (A) and (B)  (d) none of these.

130. A statistic is
   (a) a function of sample observations.  (b) a function of population units.
   (c) a characteristic of a population.  (d) a part of a population.

131. The estimate which is used in making estimation of a population parameter is:
   (a) point estimation  (b) interval estimation
   (c) both (A) and (B)  (d) none of these.

132. The sample standard deviation is
   (a) a biased estimator  (b) an unbiased estimator
   (c) a biased estimator for population S.D.  (d) a biased estimator for population variance.

133. The desired confidence level is required to determine sample size for
   (a) estimating a mean  (b) estimating a proportion
   (c) both (A) and (B)  (d) none of these.

134. For an unknown parameter, how many interval estimates exist?
   (a) only one  (b) two
   (c) three  (d) many

135. The criteria for an ideal estimator are
   (a) unbiasedness, consistency, efficiency and sufficiency
   (b) unbiasedness, expectation, sampling and estimation
   (c) estimation, consistency, sufficiency and efficiency
   (d) estimation, expectation, unbiasedness and sufficiency.

136. The confidence limits are the upper and lower limits of the:
   (a) point estimate  (b) interval estimate
   (c) confidence interval  (d) none of these.

137. The most commonly used confidence interval is
   (a) 95 per cent  (b) 90 per cent
   (c) 94 per cent  (d) 98 per cent.

138. When we have an idea or the error that might be involved, we use:
   (a) point estimate  (b) interval estimate
   (c) both (A) and (B)  (d) none of these.
139. Parameters are those constants which occur in
(a) samples. (b) probability density function.
(c) a formula (d) none of these.

140. An unbiased estimator
(a) has the smallest variance among all estimators.
(b) is always the best estimator.
(c) has an expected value equal to the true parameter value.
(d) always generates the true value of the parameter.

141. An estimator is considered to be the best if its distribution is
(a) continuous (b) discrete
(c) concentrated about the true (d) none of these.
parameter value.

142. An efficient estimator
(a) has a small variance.
(b) gets closer to the true parameter value as the sample size increases.
(c) has an expected value equal to the true parameter value.
(d) always generates the true value of the parameter.

143. When choosing an estimator of a population parameter, one should consider
(a) sufficiency (b) efficiency
(c) both (A) and (B) (d) none of these.

144. An estimator of a parametric function \( T(q) \) is said to be the best if it possesses:
(a) at least three properties of a good estimator.
(b) any two properties of a good estimator.
(c) all the properties of a good estimator.
(d) all the above.

145. A consistent estimator
(a) has the smallest variance among all estimators.
(b) gets closer to the true parameter value as the sample size increases.
(c) has an expected value equal to the true parameter value.
(d) is always sufficient.

146. Other thing equal, width of a confidence interval will be
(a) wider if the population variance is known.
(b) narrower if the population variance is unknown.
(c) wider if the population variance is unknown.
(d) the same width whether the population variance is known or not.
147. The type of estimates are
   (a) point estimate           (b) interval estimates
   (c) estimation of confidence region (d) all the above.

148. The larger the population variance the
   (a) narrower the width of the confidence interval, other things being equal.
   (b) wider the width of the confidence interval, other things being equal.
   (c) larger the mean.            (d) smaller the mean.

149. We can use the normal distribution to represent the sampling distribution of the population when
   (a) the sample size is more than 10.  (b) the sample size is less than 50.
   (c) the sample size is more than 5. (d) none of these.

150. If an estimator $T_n$ of population parameter $\theta$ converges in probability to $\theta$ as $n$ tends to infinity is said to be:
   (a) sufficient                   (b) efficient
   (c) consistent                  (d) unbiased.

151. Other things being equal, the width of a 90% confidence interval will be:
   (a) wider than a 95% confidence interval.
   (b) narrower than a 95% confidence interval.
   (c) may be wider or narrower depending on the population variance.
   (d) the same width as a 95% confidence interval.

152. If population proportion information is unknown, the standard error of the proportion can be estimated by the formula:
   (a) $\sqrt{npq}$           (b) $\sqrt{npq}$           (c) $\sqrt{pq/n}$       (d) none of these.

153. In determining the sample size for estimating a population mean, the number of factors must be known is:
   (a) 2            (b) 3            (c) 5            (d) 4

154. The standard deviation is required to determine sample size for
   (a) estimating a mean                  (b) estimating a proportion
   (c) both (A) and (B)                    (d) none of these.

155. The difference between sample S.D. and the estimate of population S.D. is negligible if the sample size is
   (a) small                  (b) moderate
   (c) sufficiently large    (d) none of these.
156. The difference between the estimate from the sample and the parameter to be estimated is
(a) sampling error.  
(b) permissible sampling error.
(c) confidence level.  
(d) none of these.

157. If the expected value of the estimator is the value of the parameter of estimation then a good estimator shall be
(a) biased  
(b) unbiased
(c) both (A) and (B)  
(d) none of these.

158. Standard error is used to set confidence limits for population parameter and in tests of significance.
(a) True  
(b) False
(c) both (A) and (B)  
(d) none of these.

159. The interval bounded by upper and lower limits is known as
(a) estimate interval  
(b) confidence interval
(c) point interval  
(d) none of these.

160. A die was thrown 400 times and ‘six’ resulted 80 times then observed value of proportion is:
(a) 0.4  
(b) 0.2
(c) 5  
(d) none of these.

161. In a sample of 400 parts manufactured by a factory, the number of defective parts was found to be 30. The observed value is
(a) 7/ 60  
(b) 3/ 40
(c) 40/ 3  
(d) 60/ 7

162. If S.D. = 20 and sample size is 100, then standard error of mean is
(a) 2  
(b) 5
(c) 1/ 5  
(d) none of these.

163. The most commonly used confidence levels are
(a) greater than and equal to 90%.  
(b) less than 90%
(c) greater than 90%.  
(d) less than and equal to 90%.

164. Statistical hypothesis is an
(a) error  
(b) assumption
(c) both (A) and (B)  
(d) none of these.

165. The confidence interval for the population mean when the population variance is known is
(a) $P_t(x \leq Z)$  
(b) $\bar{x} \pm Z_{\alpha/2} \sigma / \sqrt{n}$
(c) $Pr(x \geq z)$  
(d) 1.
166. A sufficient statistic
(a) is consistent  (b) is unbiased
(c) uses all information a sample contains about the parameter to be estimated.
(d) is always efficient.

167. The width of \((1-\alpha)\) confidence interval for a population mean is
(a) \(2\sigma Z_{\alpha/2}\)
(b) \(\sigma Z_{\alpha/2}\)
(c) \((2\sigma/\sqrt{n})Z_{\alpha/2}\)
(d) \((\sigma/\sqrt{n})Z_{\alpha/2}\)

168. Bias of an estimator can be
(a) positive  (b) negative
(c) either positive or negative  (d) always zero

169. When constructing interval estimates for the population mean we should use the
(a) Z-distribution when the population variance is unknown and the t-distribution when the population variance is known.
(b) Z-distribution when the population variance is known and the t-distribution when the population variance is unknown and is estimated with a small sample.
(c) Z-distribution whether the population variance is known or not, if the sample size is large enough.
(d) None of these.

170. Suppose a polster survey 300 people to see if they favour prayer in school. Of the 300 people surveyed, 127 favour prayer in school. An estimator of the proportion of people who favour prayer in school would be:
(a) 127  (b) 300
(c) 127/300  (d) 127/300 [1 — (127 — 300) /300]

171. If the expected value of an estimator is not equal to its parametric function \(T(q)\) it is said to be a
(a) unbiased estimator  (b) biased estimator
(c) consistent estimator  (d) none of these.

172. The ....... that we associate with an interval estimate is called the confidence level.
(a) probability  (b) statistics
(c) both  (d) none.

173. A ......... estimate is a single number.
(a) point  (b) interval
(c) both (A) and (B)  (d) none of these.
174. The higher the probability the .......... is the confidence.
   (a) moderate (b) less
   (c) more (d) none of these.

175. The sample standard deviation may be a good estimate for population standard deviation in case of .......... samples.
   (a) small (b) moderately sized
   (c) large (d) none of these.

176. A range of values is :
   (a) a point estimate (b) an interval estimate
   (c) both (A) and (B) (d) none of these.

177. If we do not have any knowledge of population variance, then we have to estimate it from
   (a) frequency (b) sample data
   (c) distribution (d) none of these.

178. The procedures for determining the sample size for estimating a population proportion are similar to those of estimating a population mean. In this case we must know ____ factor.
   (a) 2 (b) 5
   (c) 4 (d) 3

179. In .......... estimation, the estimate is given by a single quantity
   (a) interval (b) point
   (c) both (d) none of these.

180. We use t-distributions when samples are drawn from the ____ population.
   (a) normal (b) binomial
   (c) Poisson (d) none of these.

181. For 2 sample values, we have _______ values, degree of freedom
   (a) 2 (b) 1
   (c) 3 (d) 4

182. The estimate of the parameter is stated as on interval with a specified degree of
   (a) confidence (b) interval
   (c) class (d) none of these.

183. We use t-distributions when the sample size is
   (a) large (b) small
   (c) moderate (d) none of these.
184. For 5 sample values, we have ........... degree of freedom.
   (a) 5  (b) 3  (c) 4  (d) none of these.

185. Assume that you take a sample and calculate \( \bar{x} \) as 100. You then calculate the upper limit of a 90 per cent confidence interval for \( \mu \) : its value is 112. What is the lower limit of this confidence interval?
   (a) 88  (b) 92  (c) 100  (d) cannot be determined from the given information.

186. An interval estimate is
   (a) a range of values used to estimate the population parameter.
   (b) a single value that is used to estimate the population parameter.
   (c) always unbiased.
   (d) always a sufficient statistic.

187. A simple random sample of size 16 is drawn from a population with 50 members. What is the S.E. of sample mean if the population variance is known to be 25 given that the sampling is done with replacement?
   (a) 1.25  (b) 6.25  (c) 1.04  (d) 1.56

188. If a random sample of 500 oranges produces 25 rotten oranges, then the estimate of S.E. of the proportion of rotten oranges in the sample is
   (a) 0.01  (b) 0.05  (c) 0.028  (d) 0.0593

189. If the population S.D. is known to be 5 for a population containing 80 units, then the standard error of sample mean for a sample of size 25 without replacement is
   (a) 5  (b) 0.20  (c) 1  (d) 0.834

190. A sample of size 3 is taken from a population of 10 members with replacement. If the sample observations are 1, 3 and 5, what is the estimate of the standard error of sample mean?
   (a) 1.96  (b) 1.02  (c) 1.15  (d) 2.28

191. A simple random sample of size 10 is drawn without replacement from a universe containing 85 units. If the mean and S.D. as obtained from the sample, are 90 and 4 respectively, what is the estimate of the standard error of sample mean?
   (a) 0.58  (b) 1.26  (c) 0.67  (d) 0.72

192. A random sample of the heights of 100 students from a large population of students having S.D. as 0.35m show an average height of 1.75m. What are the 95% confidence limits for the average height of all the students forming the population?
   (a) [1.68m, 1.82m]  (b) [1.58m, 1.90m]  (c) [1.58m, 1.92m]  (d) [1.5m, 2.0m]
193. A random sample of a group of people is taken and 120 were found to be in favour of liberalizing licensing regulations. If the proportion of people in the population found in favour of liberalization with 95% confidence lies between 0.683 and 0.817, then the number of people in the group is:
(a) 140 (b) 150 (c) 160 (d) 175

194. If it is known that the 95% L.C.L. and U.C.L. to population mean are 48.04 and 51.96 respectively, what is the value of the population standard S.D. when the sample size is 100?
(a) 8 (b) 10 (c) 12 (d) 12.50

195. A factory produces 60,000 pairs of shoes on a daily basis. From a sample of 600 pairs, 3 per cent were found to be of inferior quality. Estimate the number of pairs that can be reasonably expected to be spoiled in the daily production process at 95% level of confidence.
(a) [989, 2612] (b) [782, 2618] (c) [982, 2618] (d) none of these.

196. A random sample of 100 days shows an average daily sale of `1000 with a standard deviation of `250 in a particular shop. Assuming a normal distribution, find the limits which have a 95% chance of including the expected sales per day.
(a) [950.75, 1049.25] (b) [950.75, 1049.25] (c) [950.75, 1149.25] (d) none of these.

197. The incidence of a particular disease in an area is such that 20 per cent people of that area suffers from it. What size of sample should be taken so as to ensure that the error of estimation of the proportion should not be more than 5 per cent with 95 per cent confidence?
(a) 246 (b) 236 (c) 226 (d) 286.
## ANSWER KEYS

<table>
<thead>
<tr>
<th>Q.</th>
<th>A.</th>
<th>Q.</th>
<th>A.</th>
<th>Q.</th>
<th>A.</th>
<th>Q.</th>
<th>A.</th>
<th>Q.</th>
<th>A.</th>
<th>Q.</th>
<th>A.</th>
<th>Q.</th>
<th>A.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>31</td>
<td>61</td>
<td>91</td>
<td>121</td>
<td>151</td>
<td>181</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>32</td>
<td>62</td>
<td>92</td>
<td>122</td>
<td>152</td>
<td>182</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>33</td>
<td>63</td>
<td>93</td>
<td>123</td>
<td>153</td>
<td>183</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>34</td>
<td>64</td>
<td>94</td>
<td>124</td>
<td>154</td>
<td>184</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>35</td>
<td>65</td>
<td>95</td>
<td>125</td>
<td>155</td>
<td>185</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>36</td>
<td>66</td>
<td>96</td>
<td>126</td>
<td>156</td>
<td>186</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>37</td>
<td>67</td>
<td>97</td>
<td>127</td>
<td>157</td>
<td>187</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>38</td>
<td>68</td>
<td>98</td>
<td>128</td>
<td>158</td>
<td>188</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>39</td>
<td>69</td>
<td>99</td>
<td>129</td>
<td>159</td>
<td>189</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>40</td>
<td>70</td>
<td>100</td>
<td>130</td>
<td>160</td>
<td>190</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>41</td>
<td>71</td>
<td>101</td>
<td>131</td>
<td>161</td>
<td>191</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>42</td>
<td>72</td>
<td>102</td>
<td>132</td>
<td>162</td>
<td>192</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>43</td>
<td>73</td>
<td>103</td>
<td>133</td>
<td>163</td>
<td>193</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>44</td>
<td>74</td>
<td>104</td>
<td>134</td>
<td>164</td>
<td>194</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>45</td>
<td>75</td>
<td>105</td>
<td>135</td>
<td>165</td>
<td>195</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>46</td>
<td>76</td>
<td>106</td>
<td>136</td>
<td>166</td>
<td>196</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>47</td>
<td>77</td>
<td>107</td>
<td>137</td>
<td>167</td>
<td>197</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>48</td>
<td>78</td>
<td>108</td>
<td>138</td>
<td>168</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>49</td>
<td>79</td>
<td>109</td>
<td>139</td>
<td>169</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>50</td>
<td>80</td>
<td>110</td>
<td>140</td>
<td>170</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>51</td>
<td>81</td>
<td>111</td>
<td>141</td>
<td>171</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>52</td>
<td>82</td>
<td>112</td>
<td>142</td>
<td>172</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>53</td>
<td>83</td>
<td>113</td>
<td>143</td>
<td>173</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>54</td>
<td>84</td>
<td>114</td>
<td>144</td>
<td>174</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>55</td>
<td>85</td>
<td>115</td>
<td>145</td>
<td>175</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>56</td>
<td>86</td>
<td>116</td>
<td>146</td>
<td>176</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>57</td>
<td>87</td>
<td>117</td>
<td>147</td>
<td>177</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>58</td>
<td>88</td>
<td>118</td>
<td>148</td>
<td>178</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>59</td>
<td>89</td>
<td>119</td>
<td>149</td>
<td>179</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>60</td>
<td>90</td>
<td>120</td>
<td>150</td>
<td>180</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>211</td>
<td>212</td>
<td>213</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

388

STATISTICS (CPT)
1. The best method to collect data in case of natural calamity is
   (a) Personal interview.              (b) Telephone interview.
   (c) Mailed questionnaire method.    (d) Indirect interview.

2. If from a population with 25 members, a random sample without replacement of 2 members is taken, the number of all such samples is
   (a) 300               (b) 625               (c) 50               (d) 600

3. Which sampling is subjected to the discretion of the sampler?
   (a) Systematic sampling           (b) Simple random sampling
   (c) Purposive sampling            (d) Quota sampling

4. Standard error can be described as
   (a) The error committed in sampling.
   (b) The error committed in sample survey.
   (c) The error committed in estimating a parameter.
   (d) Standard deviation of a statistic.

5. A population comprises 5 members. The number of all possible samples of size 3 that can be drawn from it with replacement is
   (a) 100               (b) 15               (c) 125               (d) 25

6. The sampling distribution is
   (a) The distribution of sample observations.
   (b) The distribution of random samples.
   (c) The distribution of a parameter.
   (d) The probability distribution of a statistic.

7. For 5 sample values, we have ______ degree of freedom.
   (a) 5               (b) 3               (c) 4               (d) 6

8. A sample survey is prone to
   (a) Sampling error.                (b) Non-sampling error.
   (c) Either (a) or (b).             (d) Both (a) and (b).
9. The data are known to be ______ if the data, as being already collected, are used by a different person or agency.
   (a) Primary       (b) Secondary
   (c) Specialized   (d) Subsidiary

10. The amount of non responses is likely to be maximum in ______ method of collecting data.
    (a) Telephone interview method
    (b) Personal interview method
    (c) Mailed questionnaire method
    (d) Observation method

11. The _____ the size of the sample more reliable is the result.
    (a) Medium
    (b) Smaller
    (c) Larger
    (d) None of these

12. The difference between sample S.D. and the estimate of population S.D. is negligible if the sample size is
    (a) Small.
    (b) Moderate.
    (c) Sufficiently large.
    (d) None of these.

13. A measure of precision obtained by sampling is given by
    (a) Standard error
    (b) Sampling fluctuation
    (c) Sampling distribution
    (d) Expectation

14. A range of value is
    (a) A point estimate
    (b) An interval estimate
    (c) Both
    (d) None of these

15. A random sample was taken to estimate the mean annual income of 100 families and the mean and standard deviation were found to be Rs. 6300 and Rs. 9.5 respectively find standard error for 95% confidence level.
    (a) 2.25
    (b) 1.50
    (c) 2.15
    (d) 1.862

16. For a (mxn) classification of bivariate data, the maximum number of conditional distributions is
    (a) q
    (b) p+q
    (c) pq
    (d) p

17. As the sample size decreases, standard error
    (a) Increases
    (b) Decreases
    (c) Remains constant
    (d) Increases proportionately
18. Standard deviation of sampling distribution is itself the standard error
   (a) True  (b) False  (c) Both  (d) None of these

19. If a sample of 500 eggs produces 25 rotten eggs arranges, then the estimates of SE of the proportion of rotten eggs in the sample is
   (a) 0.01  (b) 0.05  (c) 0.028  (d) 0.0593

20. A population comprises 3 numbers 2, 6, 4. Find all possible no. of samples of size two with replacement.
   (a) 27  (b) 6  (c) 9  (d) None of these

21. A random sample of 100 article taken from a large batch of articles contains 15 defective articles. What is the estimates of the proportion of defective articles in the entire batch.
   (a) 0.15  (b) 0.020  (c) 0.212  (d) None of these

22. A ______ estimate is a single number
   (a) Point  (b) Interval  (c) Both  (d) None of these

23. If sample mean is 20, population standard deviation is 3 and sample size is 64, find the interval estimate of the mean at confidence integral of 95%.
   (a) [19.265, 20.735]  (b) [19.801, 17.735]  (c) [20.735, 25.834]  (d) None of these

24. A company estimates the mean life of a drug under typical weather conditions. A simple random sample of 81 bottles yields the following information Sample mean = 23 months Population variance = 6.25 (months)$^2$
   The interval estimate with a confidence level of 90% is ____________
   (a) [22.543, 23.457]  (b) [22.6421, 23.5481]  (c) [22.451, 22.523]  (d) None of these

25. If x and y are two independent variables such that $x \sim B (n_1, P)$ and $y \sim B (n_2, p)$ then the parameter of $Z = x + y$ is
   (a) $(n_1+n_2), P$  (b) $(n_1-n_2), P$  (c) $(n_1+n_2), 2P$  (d) None of these

26. Stratified random sampling is used for ........ Population.
   (a) Homogeneous  (b) Non-homogeneous  (c) Either (a) or (b)  (d) None of these
27. Random sampling is also called lottery sampling.
   (a) False  (b) True
   (c) Both  (d) None of these

28. A simple random sample of size 36 is drawn from a finite population consisting of 101 units.
   If the population Standard Deviation is 12.6, find the Standard Error of sample mean when the
   sample is drawn with replacement.
   (a) 2.1  (b) 1.69  (c) 2.23  (d) None of these

29. A simple random sample of size 36 is drawn from a finite population consisting of 101 units.
   If the population Standard Deviation is 12.6, find the Standard Error of sample mean when the
   sample is drawn without replacement.
   (a) 2.1  (b) 1.69  (c) 2.45  (d) None of these

30. A random sample of size 9 is drawn without replacement from a finite population consisting of
    25 units. If the number of defective units in the population be 5, find the Standard Error of the
    sample proportion of defectives.
    (a) 0.1288  (b) 0.1088  (c) 0.0588  (d) None of these

31. A population consists of 4 numbers. Find the number of sample of size 2 for with replacement
    condition.
    (a) 16  (b) 6  (c) 10  (d) None of these

32. A population consists of 4 numbers. Find the number of sample of size two for without
    replacement condition.
    (a) 16  (b) 6  (c) 10  (d) None of these

33. The no. of factors must be known is __________ in determining the sample size for estimating
    a population mean.
    (a) 2  (b) 5  (c) 4  (d) 3

34. A sample of 100 gave a mean of 7.4 kg and a standard deviation of 1.2 kg. The standard error
    of mean will be
    (a) 0.12  (b) 0.001  (c) 0.0001  (d) 1

35. A sample of 100 gave a mean of 7.4 kg and a standard deviation of 1.2 kg. Find 95% confidence
    limits for population mean.
    (a) 7.164 and 7.635  (b) 5.164 and 5.635
    (c) 4.001 and 5.001  (d) None of these
36. A random sample of the heights of 500 oranges was taken from a large consignment. 65 were found to be defective. Find Standard Error of the proportion of defectives.

(a) 0.015  (b) 0.15  (c) 0.017  (d) None of these

37. Sample mean is a

(a) Parameter  (b) Statistic

(c) Both  (d) None of these

38. Deliberate sampling is a

(a) Random sampling  (b) Non - random sampling

(c) Both (a) and (b)  (d) None of these

39. A population consists units a, b, c, d, e, f. The total number of all possible samples of size four without replacement are

(a) 10  (b) 12

(c) 15  (d) None of these

40. A random sample of the heights of 100 students from a large population of students in a College having Standard Deviation of 0.75 ft. has an average height of 5.6 ft., Find 95% confidence limits for the average height of all the students of the College. (For 95%, Z = 1.96)

(a) 5.453, 5.747  (b) 16.453, 7.747

(c) 6.485, 7.647  (d) None of these

41. A sample of size 64 was drawn from a population consisting of 128 units. The sample mean of the measurements on a certain characteristic was found to be 28. If the population Standard Deviation is 4 then find the 96% confidence limit for the population mean? (For 96%, Z = 2.05)

(a) 30.72, 27.32  (b) 27.272, 28.728

(c) 30.272, 32.728  (d) None of these

42. In a large consignment of oranges a random sample of 500 oranges revealed that 65 oranges were bad. Find the 99.73% Confidence limit of bad ones? (For 99.73%, Z = 3)

(a) 8.5%, 17.5%  (b) 0.85%, 0.175%

(c) 8%, 12%  (d) None of these
43. Find the sample size such that the probability of the sample means differing from the population mean by not more than \( \frac{1}{10} \) th of the Standard Deviation is 0.95.

(a) 300  
(b) 384  
(c) 395  
(d) None of these

44. The 95% confidence limit for the sample mean \( \bar{x} \) is \( \mu \pm 1.96 \left( \frac{\sigma}{\sqrt{n}} \right) \).

This is

(a) True  
(b) False  
(c) Either (a) or (b)  
(d) None of these
# ANSWER KEYS

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>11</th>
<th>C</th>
<th>21</th>
<th>A</th>
<th>31</th>
<th>A</th>
<th>41</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>A</td>
<td>12</td>
<td>C</td>
<td>22</td>
<td>A</td>
<td>32</td>
<td>B</td>
<td>42</td>
<td>A</td>
</tr>
<tr>
<td>3</td>
<td>C</td>
<td>13</td>
<td>A</td>
<td>23</td>
<td>A</td>
<td>33</td>
<td>D</td>
<td>43</td>
<td>B</td>
</tr>
<tr>
<td>4</td>
<td>D</td>
<td>14</td>
<td>B</td>
<td>24</td>
<td>A</td>
<td>34</td>
<td>A</td>
<td>44</td>
<td>B</td>
</tr>
<tr>
<td>5</td>
<td>C</td>
<td>15</td>
<td>D</td>
<td>25</td>
<td>A</td>
<td>35</td>
<td>A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>D</td>
<td>16</td>
<td>B</td>
<td>26</td>
<td>B</td>
<td>36</td>
<td>A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>C</td>
<td>17</td>
<td>A</td>
<td>27</td>
<td>B</td>
<td>37</td>
<td>B</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>D</td>
<td>18</td>
<td>A</td>
<td>28</td>
<td>A</td>
<td>38</td>
<td>B</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>B</td>
<td>19</td>
<td>A</td>
<td>29</td>
<td>B</td>
<td>39</td>
<td>C</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>C</td>
<td>20</td>
<td></td>
<td>30</td>
<td>B</td>
<td>40</td>
<td>A</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* * *