UNIT 15 TESTS OF HYPOTHESIS–I

STRUCTURE

15.0 OBJECTIVES

After studying this unit, you should be able to:

- estimate population characteristics (parameters) on the basis of a sample,
- get familiar with the criteria of a good estimator,
- differentiate between a point estimator and an interval estimator,
- comprehend the concept of statistical hypothesis,
- perform tests of significance of population mean and population proportion, and
- make decisions on the basis of testing hypothesis.

15.1 INTRODUCTION

Let us suppose that we have taken a random sample from a population with a view to knowing its characteristics, also known as its parameters. We are then confronted with the problem of drawing inferences about the population on the basis of the known sample drawn from it. We may look at two different scenarios. In the first case, the population is completely unknown and we would like to throw some light on its parameters with the help of a random sample drawn from the population. Thus, if μ denotes the population mean, then we intend to make a guess about it on the basis of a random sample. This is known as estimation. For example, one may be interested to know the average income of people living in the city of Delhi or the average life in burning hours of a fluorescent tube light produced by 'Indian Electrical' or proportion of people suffering from T.B. in city 'B' or the percentage of smokers in town 'C' and so on.

A somewhat different situation may arise when some information about a parameter is either known or specified and we would like to verify whether that information holds good for the sample drawn from the population as well.

This is known as problem of testing of hypothesis. In the previous examples, we may be interested of in testing whether the average income in the city of Delhi is, say, Rs. 2,000 per month. In the second example, we may like to verify whether the claims made by Indian Electrical, that their fluorescent lamps would last 5,000 hours, is justified. Some social workers may believe that 20% of the population in city B suffers from T.B. We would like to make our comment after a test of hypothesis. In the last example, some human activists, concerned about the hazards of passive smoking, assert that 30% of the people staying in town C are smokers. We may share their opinion once we have satisfied ourselves after peforming a statistical test of hypothesis.

It may be noted that testing of hypothesis plays a vital role in decision-making. In the first example, the statistician may be concerned about whether to bracket Delhi with the top metropoliton cities depending on the average income based on his/her recommendations. If on the basis of a statistical test, it is found that the claim made by the manufacturer of India Electrical is justified, then the sales of his lamps would increase. In the third example if there is evidence, again on the basis of testing hypothesis, that the social worker is right about his statement, suitable steps may be undertaken to improve the living conditions of the marginalized section in the city so that the percentage of people suffering from T.B. is reduced. Some strict legislation banning smoking or reducing smoking to a desirable level may be enacted on the basis of a hypothesis tested in the last example.

15.2 POINT ESTIMATION AND STANDARD ERRORS

Estimation is an integral part of our daily lives. In order to construct a new house or renovate an old house or flat, we demand an estimate of the cost involved. A student estimates his/her chance of success before appearing for an expensive competitive examination.

Now we shall consider estimation from the viewpoint of a statistician. As we discussed in Unit 4: sampling, is the means to find the true value of the parameter which can be correctly obtained only through census study. In many cases it is not practicable due to various constraints. Therefore, the alternative approach is to select some items as a sample from the population and collect the data and analyse the data, then estimate the chracteristics of the population. This is called **estimation**. Point estimate is one type of estimate. It is a single number which is used as an estimate of the unknown population parameter. Let us assume that we have taken a random sample of n observations, x_1, x_2, $x_3 \ldots x_n$, from a population characterized by a parameter θ (read theta). This symbol θ is used to denote a parameter that could be mean, mode or some measure of variation etc. Thus θ may be the mean (μ)of a normal distribution or the probability of success (P) of a binomial distribution with parameters 'n' and 'p' and so on. In theory of estimation, we try to find a statistic (i.e., a function of sample observations) T which estimates the unknown parameter θ. Thus the sample mean $(\bar{x}) = \Sigma x_i / n$, x_1, x_2, $x_3 \ldots x_n$ being the income per month of 'n' persons selected at random from the city of Delhi, may be considered to be the estimate of the average income per month (μ) of the people of Delhi.

This is denoted by $\hat{\mu} = \bar{x}$ i.e., the estimate of μ is $\bar{x}$.

To be more precise, $\bar{x}$ is known as a point estimator of μ as we try to estimate the population mean (μ) by a single value, namely, the sample mean. On the basis of a random sample of incomes from Delhi, if it is found that the sample mean is Rs. 2,000/-, then one may conclude that the estimate of average income per month of the people living in that city is Rs. 2,000/-.

As opposed to a point estimate, one may think of an interval estimate that is supposed to contain the average income of the people of Delhi per month. This would be discussed in Section 15.3.

At this juncture, we must make a distinction between the two terms Estimator and estimate. 'T' is defined to be an estimator of a parameter θ, if T esimtates θ. Thus T is a statistic and its value may differ from one sample to another sample. In other words, T may be considered as a random variable. The probability distribution of T is known as sampling distribution of T. As already discussed, the sample mean $\bar{x}$ is an estimator of population mean μ. The value of the estimator, as obtained on the basis of a given sample, is known as its estimate. Thus $\bar{x}$ is an estimator of μ, the average income of Delhi, and the value of $\bar{x}$ i.e., Rs. 2,000/-, as obtained from the sample, is the estimate of μ.

Selection of the best estimator: Our next endeavour would be to discuss different criteria for selecting the best estimator.

Unbiasedness and Minimum Variance: A statistic T is defined to be unbiased for a parameter θ if expectation of T is θ, i.e., E(T) = θ. On the other hand if E(T) = θ + a (θ), then the difference a (θ) = E(T) – θ is known as bias. The bias is known to be positive if a (θ) > 0 and negative if a (θ) < 0. Our first priority would be to select an unbiased estimator of θ. However, there may be many unbiased estimators of θ. If x_1, x_2 …, x_n denote n sample observations from a population with an unknown parameter θ, then any of the n observations or any linear function of them would be an unbiased estimator of θ.

In order to choose the best estimator among these estimators along with "unbiasedness", we introduce a second criterion, known as, minimum variance. A statistic T is defined to be a minimum variance unbiased estimator (MVUE) of θ if T is unbiased for θ and T has minimum variance among all the unbiased estimators of θ. We may note that sample mean ($\bar{x}$) is an MVUE for μ.

We know that $\bar{x} = \frac{\Sigma x_i}{n}$ …(15.1)

$$\therefore E(\bar{x}) = E\left[\frac{(\Sigma x_i)}{n}\right]$$

$$= \frac{1}{n}.[\Sigma E(x_i)]$$

$$= \frac{1}{n}[\Sigma \mu]$$ [x_1, x_2, …x_n are taken from population having as μ population mean]

$$= \frac{1}{n}.\, n\mu$$

$$E(\bar{x}) = \mu$$

Hence $(\bar{x})$ is an unbiased estimator of μ.

Further, variance of $(\bar{x})$ is given by :

$$v(\bar{x}) = v(\Sigma x_i / n) \text{ (where v denotes variance)}$$

$$= \frac{1}{n^2}.[\Sigma v(x_i)] \quad \text{Since } x_i^1\text{s are independent}$$

$$= \frac{1}{n^2}\Sigma\sigma^2 = \text{[where } \sigma^2 \text{ is population variance]}$$

$$= \frac{1}{n^2}.[n\sigma^2] = \frac{\sigma^2}{n} \qquad \text{……(15.2)}$$

It can be proved that v $(\bar{x})$ has the minimum variance among all the unbiased estimators of μ.

Consistency: If T is an estimator of θ, then it is obvious that T should be in the neighbourhood of θ. T is known to be consistent for θ, if the difference between T and θ can be made as small as we please by increasing the sample size n sufficiently.

We can further add that T would be a consistent estimator of θ if

i) E (T) → θ and

ii) V (T) → 0 for a very large n i.e., as n → ∝

For example, sample mean $(\bar{x})$ is a consistent estimator of μ as E $(\bar{x}) = \mu$

And V $(\bar{x}) = \frac{\sigma^2}{n} \rightarrow 0 \text{ as } n \rightarrow \propto$.

It may be noted that if T is a consistent estimator of θ, then any function of T is also a consistent estimator of θ.

Efficiency: A statistic T is called as an efficient estimator of θ if it has the minimum standard error among all the estimators of θ for a fixed sample size n. Both the sample mean and sample median are consistent estimators for μ. But standard error (a term, to be defined and explained in this section) of sample mean is less than that of sample median. Hence sample median is only a consistent estimator of μ, whereas sample mean is both consistent and efficient estimator of μ.

Sufficiency: A statistic T is known to be a sufficient estimator of θ if T contains sufficient information about θ so that we do not have to look for any other estimator of θ. Sample mean $(\bar{x})$ is a sufficient estimator of μ.

Now let us consider the following point estimates that are commonly used.

A) **Estimating Population Mean:** It is obvious that sample mean is the best estimator of population mean μ. It is an MVUE. It is both consistent and efficient estimator for μ. Further more, $\bar{x}$ is a sufficient estimator for μ. Thus we estimate the average income of the people of Delhi by the sample mean or the average life of bulbs, manufactured by Indian Electricals, by the corresponding sample mean.

B) **Estimating Population Proportion:** If a discrete random variable x follows binomial distribution with parameters n and P, then we have

$$\mu = E(x) = nP$$

$$\sigma^2 = v(x) = nP\,(1-p)$$

[n denoting the number of trials and P denoting the probability of a success].

Hence, it follows that :

$$E(p) = E\left(\frac{x_i}{n}\right) = \frac{nP}{n} = P \qquad \ldots(15.3)$$

and V (p)= V (x_i/n) $= \dfrac{V(x_i)}{n^2}$

$$= \frac{nP\,(1-P)}{n^2}$$

$$= \frac{P\,(1-P)}{n} \qquad \ldots(15.4)$$

Thus if we take a random sample of size 'n' from a population where the proportion of population possessing a certain characteristic is 'P' and the sample contains x units possessing that characteristic, then an estimate of population proportion (P) is given by:

$$\hat{P} = \frac{x}{n} \qquad \ldots(15.5)$$

In other words, the estimate of the population proportion is given by the corresponding sample estimate i.e., $\hat{P} = p$ ...(15.6)

From (15.3) E {p} = P

So p is an unbiased estimator of P. It can be shown that p has the minimum variance among all the unbiased estimators of p. In other words, p is an MVUE of P.

$$\text{As } v\,(p) = \frac{P(1-P)}{n} \to 0 \text{ as } n \to \infty$$

it follows from Eq. (15.4) that p is a consistent estimator of P. We can further establish that p is an efficient as well as a sufficient estimator of P. Thus we advocate the use of sample proportion to estimate the population proportion as p which satisfies all the desirable properties of an estimator.

In order to estimate the proportion of people suffering from T.B. in city B, if we find the number of people suffering from T.B. is 'x' in a random sample of size 'n', taken from city B, then sample estimate p = x/n would provide the estimate of the proportion of people in that city suffering from T.B. Similarly, the percentage of smokers as found from a random sample of people of town C would provide the estimate of the percentage of smokers in town C.

C) **Estimation of Population Variance and Standard Error:** Standard error of a statistic T, to be denoted by S.E. (T), may be defined as the standard deviation of T as obtained from the sampling distribution of T. In order to compute the standard error of sample mean, it may be noted that from Eq.

(15.2) $S.E.\,(\bar{x}) = \dfrac{\sigma}{\sqrt{n}}$ for simple random sampling with replacement (SRSWR).

$S.E.\,(\bar{x}) = \dfrac{\sigma}{\sqrt{n}}\sqrt{\dfrac{N-\eta}{N-1}}$ for simple random sampling without replacement (SRSWOR)].

where σ is the population standard deviation (S.D.), n is sample size, N is population size and the factor $\sqrt{\dfrac{N-\eta}{N-1}}$ is known as finite population corrector (f.p.c.) or finite population multiplier (f.p.m.) which may be ignored for a large population.

In order to find S.E., it is necessary to estimate σ^2 or σ in case it is unknown. If $x_1, x_2, \ldots, x_n$ denote n sample observations drawn from a population with mean μ and variance σ^2, then the sample variance:

$$S^2 = \frac{\Sigma(x_i - \bar{x})^2}{n} \qquad \ldots(15.7)$$

may be considered to be an estimator of σ^2

Since $E(x) = \mu$ and $V(\not{z}) = E(x-\mu)^2 = \sigma^2 \quad \ldots(15.8)$

We have

$$ns^2 = \Sigma(x_i - \bar{x})^2 \qquad \ldots(\text{from } 15\text{-}8)$$

$$= \Sigma[(x_i - \mu) - (\bar{x} - \mu)]^2$$

$$= \Sigma[(x_i - \mu)^2 - 2(x_i - \mu)(\bar{x} - \mu) + (\bar{x} - \mu)^2]$$

$$= \Sigma(x_i - \mu)^2 - 2(\bar{x} - \mu)\ \Sigma(x_i - \mu) + n(\bar{x} - \mu)^2$$

$$= \Sigma(x_i - \mu)^2 - 2(\bar{x} - \mu).n(\bar{x} - \mu) + n(\bar{x} - \mu)^2$$

$$[\text{since } \Sigma(x_i - \mu) = \Sigma x_i - \Sigma\mu$$

$$= n\bar{x} - n\mu = n(\bar{x} - \mu)]$$

$$= \Sigma(x_i - \mu)^2 - 2n(\bar{x} - \mu)^2 + n(\bar{x} - \mu)^2$$

$$= \Sigma(x_i - \mu)^2 + n(\bar{x} - \mu)^2 \qquad \ldots(15.9)$$

As x_i is the ith sample observation from a population with μ as mean and σ^2 as variance, it follows that :

$$E(x_i - \mu)^2 = \sigma^2$$

And $$E(\bar{x} - \mu)^2 = v(\bar{x}) = \frac{\sigma^2}{n} \qquad \ldots(15.10)$$

From (15-9), $E(ns^2) = \Sigma E(x_i - \mu)^2 - n.E(\bar{x} - u)^2$

$$= \Sigma\sigma^2 - n.\frac{\sigma^2}{n} = n\sigma^2 - \sigma^2 = (n-1)\sigma^2$$

$$\therefore E(S^2) = \frac{n-1}{n}\sigma^2 \quad \neq \sigma^2 \qquad \ldots(15.11)$$

Hence S^2, the sample variance, is a biased estimator of σ^2.

$$\text{As } E(S^2) = \frac{n-1}{n}\sigma^2$$

$$\therefore E\left(\frac{n}{n-1}s^2\right) = \sigma^2 \qquad \ldots(15.12)$$

thus $\frac{n}{n-1}s^2 = \frac{(\Sigma x_i - \bar{x})^2}{n-1} = (s^{|})^2$ is an unbiased estinator of σ^2 $\ldots(15.13)$

so, we use $(s^{|})^2 = \frac{\Sigma (x_i - \bar{x})^2}{n-1}$ as an estimator of σ^2 and

$S^{|} = \sqrt{\frac{\Sigma (x_i - \bar{x})^2}{n-1}}$ as an estimator of σ

An estimate of S.E. $(\bar{x})$ is given by:

$$\hat{\text{S.E.}}\,(\bar{x}) = \frac{S^{|}}{\sqrt{n}} \quad \text{for SRSWR}$$

$$= \frac{S^{|}}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}} \quad \text{for SRSWOR} \qquad \ldots\ldots(15.14)$$

From (15.4), it follows that $v(p) = \frac{P\,(1-P)}{n}$

$$\text{S.E.}(p) = \frac{\sqrt{P(1-P)}}{n} \quad \text{for SRSWR}$$

$$= \sqrt{\frac{P\,(1-P)}{n}} . \sqrt{\frac{N-n}{N-1}} \quad \text{for SRSWOR} \qquad \ldots\ldots(15.15)$$

An estimate of standard error of sample proportion is given by:

$$\hat{\text{S.E.}}\,(p) = \sqrt{\frac{p\,(1-p)}{n}} \quad \text{for SRSWR}$$

$$= \sqrt{\frac{p\,(1-p)}{n}} . \sqrt{\frac{N-n}{n}} \quad \text{for SRSWOR} \qquad \ldots\ldots(15.16)$$

Let us consider the following illustrations to estimate variance from sample and also estimate the standard error.

Illustration 1

A sample of 32 fluorescent lights taken from Indian Electricals was tested for the lives of the lights in burning hours. The data are presented below:

Table 15.1: The Lives in Hours of 32 Lights

Sl. No.	Life (Hours)	Sl. No.	Life (Hours)	Sl. No.	Life (Hours)
1	4895	12	4992	23	4987
2	4907	13	4997	24	5021
3	5013	14	5003	25	5009
4	4996	15	4985	26	5016
5	5015	16	5015	27	5019
6	4899	17	5317	28	4903
7	4723	18	4990	29	4925
8	4968	19	4989	30	4972
9	5023	20	4923	31	5009
10	5021	21	4946	32	4998
11	5015	22	5024		

Solution: We are interested in estimating the average life of fluorescent lights manufactured by Indian Electricals. As discussed in this section, the estimate of the population mean (μ) is given by the corresponding sample mean. Then

$\hat{\mu} = \bar{x}$. If we are further interested in estimating the standard error of $\bar{x}$, then we are to compute

$$\hat{S.E.}(\bar{x}) = \frac{s^{|}}{\sqrt{n}}$$

$$\text{where, } s^{|} = \sqrt{\frac{\Sigma(x_i - \bar{x})^2}{\eta - 1}} = \sqrt{\frac{\Sigma x_i^{\,2} - n\bar{x}^2}{n-1}}$$

$$\text{and } \bar{x} = \frac{\Sigma x_i}{n}, \; n = \text{Sample size}$$

We ignore f.p.c. as the population of lights is very large.

Table 15.2: Computation of sample mean and sample S.D.

Life in Hours x_i	$u_i = x_i - 5000$	u_i^2
4895	–105	11025
4907	–93	8649
5013	13	169
4996	–4	16
5015	15	225
4899	–101	10201
4723	–277	76729
4968	–32	1024
5023	23	529
5021	21	441
5015	15	225
4992	–8	64
4997	–3	9
5003	3	9
4985	–15	225
5015	15	225
5317	317	100489
4990	–10	100
4984	–16	256
4923	–77	5929
4946	–54	2916
5024	24	576
4987	–13	169
5021	21	441
5009	9	81
5016	16	256

5019	19	361
4903	–97	9409
4925	–75	5625
4972	–28	784
5009	9	81
4998	–2	4
Total	**–490**	**237242**

From the above Table, $\Sigma u_i = -490, \quad \Sigma u_i^2 = 237242$

$$\therefore \bar{u} = \frac{\Sigma u_i}{32} = \frac{-490}{32} = -15.3125$$

As $u_i = x_i - 5000$

$$\therefore \bar{u} = \bar{x} - 5000$$

Or, $\overline{X} = 5000 + \bar{u} = 4984.6875$ (approximately) 4985

$$(s^|)^2 = (s^|{}_x)^2 = (s^|)^2 = \frac{\Sigma u_i^2 - n\bar{u}^2}{n-1} = \frac{23742 - 7503.1248}{31} = 7410.9316$$

$$\therefore s^| = 86.0868$$

hence $\text{S.E.}(\bar{x}) = \frac{s^|}{\sqrt{n}} = \frac{86.0868}{\sqrt{32}} = 15.2183$

so the estimate of the average life of lights as manufactured by Indian Electricals is 4985 hours. Estimate of the population variance in 7410.9316 $(\text{hours})^2$ and the standard error is 15.2183 hours.

Illustration 2

A sample of 350 people from city C contained 70 smokers. Find an estimate of the proportion of smokers in the city. Also find an estimate of the standard error of the proportion of smokers in the sample.

Solution: In this case x = no. of smokers in the sample = 70, n = 350.

Thus we have $p = \frac{x}{n} = \frac{70}{350} = 0.2$

Hence the estimate of the proportions of smokers in the city is 0.2 or 20%.

Further

$$\hat{\text{S.E.}}\,(p) = \sqrt{\frac{p\,(1-p)}{n}} = \sqrt{\frac{0.2\,(1-0.2)}{350}} = 0.0214$$

$\therefore$ The estimate of the standard error of the proportion of smokers in the sample is 0.0214.

Self Assessment Exercise A

1) State, with reasons, whether the following statements are true or false.

   a) Both the statistic and parameter are functions of sample observations.

   b) Any type of sampling would lead to the same inference about the population.

   c) Statistical inference is a statistical process to know about a population from the knowledge of a sample drawn from it.

   d) Any type of estimator can be used for estimating a parameter.

   e) In most cases, decision-making depends on estimation.

   f) There may be more than one estimator for a parameter.

   g) Assumption of normality is a must for point estimation.

   h) Every consistent estimator is necessarily an efficient estimator.

   i) A consistent estimator approaches the parameter with an increase in sample size.

   j) Point estimator is used as an estimate of the unknown population parameter.

2) Differentiate between estimator and estimate.

   ....

   ....

3) In choosing between sample mean and sample median – which one would you prefer?

   ....

   ....

4) The monthly earnings of 20 families, obtained from a random sample from a village in West Bengal are given below:

Sl. No.	Monthly earnings (Rs.)	Sl. No.	Monthly earnings (Rs.)
1	1023	11	1012
2	976	12	998
3	898	13	1015
4	1012	14	989
5	980	15	923
6	963	16	767
7	1023	17	897
8	946	18	1013
9	1007	19	947
10	977	20	958

   Find an estimate of the average monthly earnings of the village. Also obtain an estimate of the S.E. of the sample estimate.

   ....

....

....

....

....

5) In a sample of 900 people, 429 people are found to be consumers of tea. Estimate the proportion of consumers of tea in the population. Also find the corresponding standard error.

....

....

....

6) Obtain an unbiased estimate of population mean and population variance on the basis of the following sample observations:

50, 46, 52, 53, 45, 43, 46, 48, 51

....

....

....

15.3 INTERVAL ESTIMATION

This is another type of estimation. As opposed to estimating a parameter by a single value i.e., point estimation discussed in the previous section, we may think of an interval or a range of values that is supposed to contain the parameter. An interval estimate would always be specified by two values i.e., the lower value and the upper value, within which the parameter lies. This is known as **Interval Estimation**. Thus interval estimation may be defined as estimating an interval to which the unknown parameter θ may belong, in all likelihood.

Regarding the estimation of the average income of the people of Delhi city, one may argue that it would be better to provide an interval which is likely to contain the population mean. Thus, instead of saying the estimate of the average income of Delhi is Rs. 2,000/-, we may suggest that, in all probability, the estimate of the average income of Delhi would be from Rs. 1,900/- to Rs. 2,100/-. In the second example of estimating the average life of lights produced by Indian Electricals where the estimate came out to be 4985 hours, the point estimation may be a bone of contention between the producer and the potential buyer. The buyer may think that the average life is rather less than 4985 hours. An interval estimation of the life of lights might satisfy both the parties. Figure 15.1 shows some intervals for θ on the basis of different samples of the same size from a population characterized by a parameter θ. A few intervals do not contain θ.

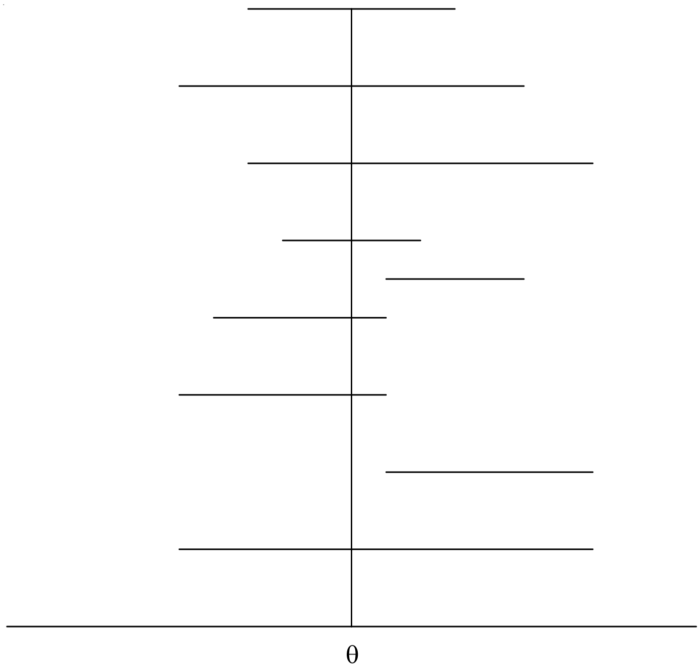

Fig.15.1: Confidence Intervals to θ

15.4 CONFIDENCE LIMITS, CONFIDENCE INTERVAL AND CONFIDENCE CO-EFFICIENT

Let us assume that we have taken a random sample of size 'n' from a population characterized by a parameter θ. Let us further suppose that based on these sample observations, it is possible to find two statistics t_1 and t_2 such that:

$P\ (t_1 < \theta) = \alpha_1$

And $P\ (t_2 > \theta) = \alpha_2$

Where α_1 and α_2 are two small positive numbers. Combining these two conditions, we may write:

$P\ (t_1 \leq \theta \leq t_2) = 1 – \alpha$ …(15.17)

Where $\alpha = \alpha_1 + \alpha_2$

Equation (15.17) could be interpreted as the probability that θ lies between t_1 and t_2 is (1–α), whatever may be the value of θ, satisfying (15.17). The interval $[t_1, t_2]$, t_1 being less than t_2, that contains the parameter θ is known Confidence Interval to θ, t_1 being known as Lower Confidence limit and t_2 as Upper Confidence Limits. (1–α) is known to be Confidence Co-efficient corresponding to the confidence interval $[t_1, t_2]$.

One may like to know why the term 'confidence' comes into the picture. If we choose α_1 and α_2 such a way that α = 0.01, then the probability that θ would belong to the random interval $[t_1, t_2]$ is 0.99. In other words, one feels 99% confident that $[t_1, t_2]$ would contain the unknown parameter θ. Similarly if we select α = 0.05, then $P\ [t_1 \leq \theta \leq t_2] = 0.95$, thereby implying that we are 95% confident that θ lies between t_1 and t_2. (15.17) suggests that as α decreases, (1–α) increases and the probability that the confidence interval $[t_1, t_2]$ would include the parameter θ also increases. Hence our endeavour would be to reduce 'α' and thereby increase the confidence co-efficient (1–α).

Referring to the estimation of the average life of lights (θ), if we observe that θ lies between 4935 hours and 5035 hours with probability 0.98, then it would imply that if repeated samples of a fixed size (say n = 32) are taken from the population of lights, as manufactured by Indian Electricals, then in 98 per cent of cases, the interval [4935 hours, 5035 hours] would contain θ, the average life of lights in the population while in 2 per cent of cases, the interval would not contain θ. In this case, the confidence interval for θ is [4935 hours, 5035 hours]. Lower Confidence Limit of θ is 4935 hours, Upper Confidence Limit of θ is 5035 hours, and the Confidence Co-efficient is 98 per cent.

Selection of Confidence Interval

Our next task would be to select the basis for estimating confidence interval. Let us assume that we have taken a random sample of size ‘n’ from a normal population characterized by the two parameters μ and σ, the population mean and standard deviation respectively. Thus, in the case of estimating a Confidence Interval for average income of people dwelling in Delhi city, we assume that the distribution of income is normal and we have taken a random sample from the city. In another example concerning average life of fluorescent lights as produced by Indian Electricals, we assume that the life of a fluorescent light is normally distributed and we have taken a random sample from the population of fluorescent lights manufactured by Indian Electricals.

Figure 15.2 shows percentage of area under Normal Curve. It can be shown that if a random sample of size ‘n’ is drawn from a normal population with mean ‘μ’ and variance σ^2, then $(\bar{x})$, the sample mean also follows normal distribution with ‘μ’ as mean and σ^2/n as variance. Further as we have observed in Section 15.2.

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

From the properties of normal distribution, it follows that the interval :

$[\mu - \text{S.E.}\,(\bar{x}),\ \mu + \text{S.E.}\,(\bar{x})]$ covers 68.27% area.

The interval $[\mu - 2\,\text{S.E.}\,(\bar{x}), \mu + 2\,\text{S.E.}(\bar{x})]$ covers 95.45% area and the interval

$[\mu - 3\,\text{S.E.}\,(\bar{x}),\ \mu + 3\,\text{S.E.}\,(\bar{x})]$ covers 99.73% area. Figure 15.2 depicts this information.

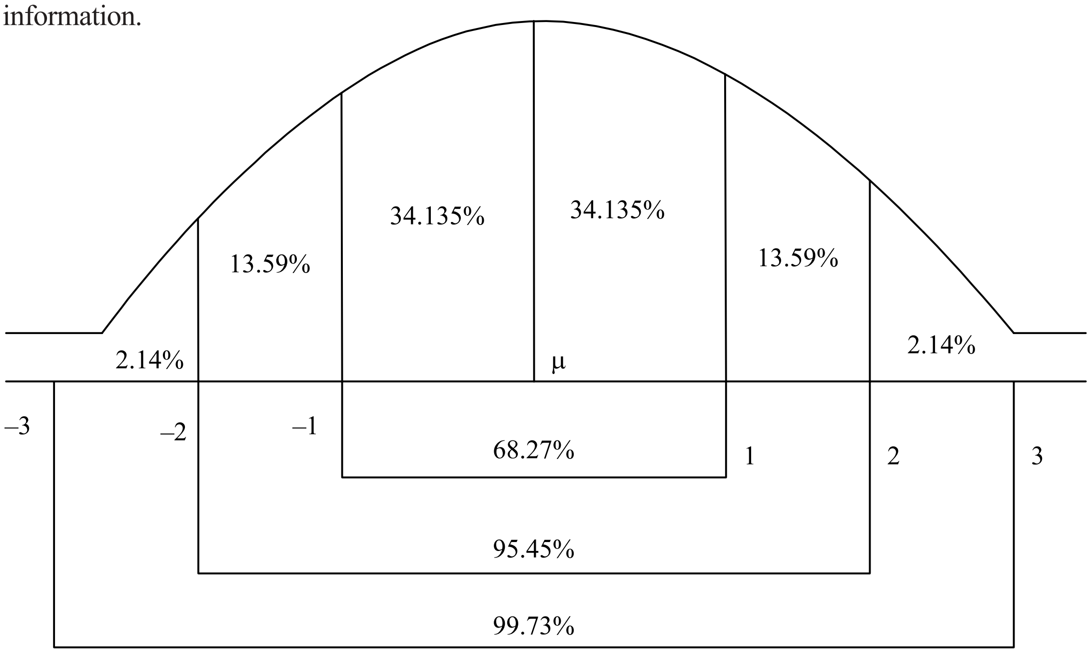

Fig. 15.2: Percentages of Area under a Normal Curve

Now let us consider a situation where the assumption of normality may not hold. If the sample size is large enough, then the sample mean $\bar{x}$ follows approximately, i.e., asymptotically normal distribution with mean as μ and standard error as $\sigma/\sqrt{n}$, μ and σ being the mean and S.D. of the population under consideration. In case σ is unknown, we can replace it by the corresponding sample standard deviation. One may ask the question as to how large 'n' should be. It is rather difficult to specify an exact value of 'n' so that the distribution of $\bar{x}$ would be asymptotically normal. Larger the value of 'n', the better. However for practical purposes, if 'n' exceeds 30, then we may assume that $\bar{x}$ is asymptotically normal.
Our next question may be what would be the confidence interval for μ.

Will it be $\mu \pm \text{S.E.}(\bar{x})$, or $\mu \pm 2\,\text{S.E.}(\bar{x})$, or $\mu \pm 3\,\text{S.E.}(\bar{x})$, or some other interval?

Suppose that the Confidence Interval to μ is given by :

$\mu \pm u\,\text{S.E.}(\bar{x})$ and we are to determine u such that :

$$P[\bar{x} - u \times \text{S.E.}(\bar{x}) \le \mu \le \bar{x} + u \times \text{S.E.}(\bar{x})] = 1 - \alpha \qquad \ldots\ldots(15.18)$$

Or, $$P[-u \le \frac{\bar{x} - \mu}{\text{S.E.}(\bar{x})} \le u] = 1 - \alpha$$

Or, $P[-u \le Z \le u] = 1 - \alpha$ [where $Z = \dfrac{\bar{x} - \mu}{\text{S.E.}(\bar{x})}$ is a standard normal variable]

Or. $\underline{\bar{\phi}}$ (u) – $\underline{\bar{\phi}}$ (– u) = 1–α [where $\underline{\bar{\phi}}$ (K) = P(Z ≤ K), area under standard normal curve from – ∝ to K].

Or, $\underline{\bar{\phi}}(u) - [1 - \underline{\bar{\phi}}(u) = 1 - \alpha$

Or. $2\underline{\bar{\phi}}(u) = 2 - \alpha$

Or. $$\underline{\bar{\phi}}(u) = 1 - (\alpha/2) \qquad \ldots(15.19)$$

Putting α = 0.10 in (15.19), we get

$$\underline{\bar{\phi}}(u) = 1 - 0.05 = 0.95$$

or, $\underline{\bar{\phi}}(u) = \underline{\bar{\phi}}(1.645)$

or. u = 1.645

Thus 100 (1–α) % or 100 (1–0.1)% or 90% confidence interval to population mean μ is :

Given by $$\left[\bar{X} - 1.645\frac{\sigma}{\sqrt{n}},\ \bar{X} + 1.645\frac{\sigma}{\sqrt{n}}\right]$$

Putting α =0.05, 0.02 and 0.01 respectively in (15.19) and proceeding in a similar manner, we get 95% Confidence Interval to $\mu = \left[\bar{x} - 1.96\frac{\sigma}{\sqrt{n}}, \bar{x} + \frac{1.96\sigma}{\sqrt{n}}\right]$ …(15.20)

98% Confidence Interval to $\mu = \left[\bar{x} - 2.33\frac{\sigma}{\sqrt{n}}, \bar{x} + 2.33\frac{\sigma}{\sqrt{n}}\right]$ …(15.21) and 99% Confidence Interval to $\mu = \left[\bar{x} - 2.58\frac{\sigma}{\sqrt{n}}, \bar{x} + 2.58\frac{\sigma}{\sqrt{n}}\right]$ ...(15.22)

Theoretically we may take any Confidence interval by choosing 'u' accordingly. However in a majority of cases, we prefer 95% or 99% Confidence Interval. These are shown in Figure 15.3 and Figure 15.4 below.

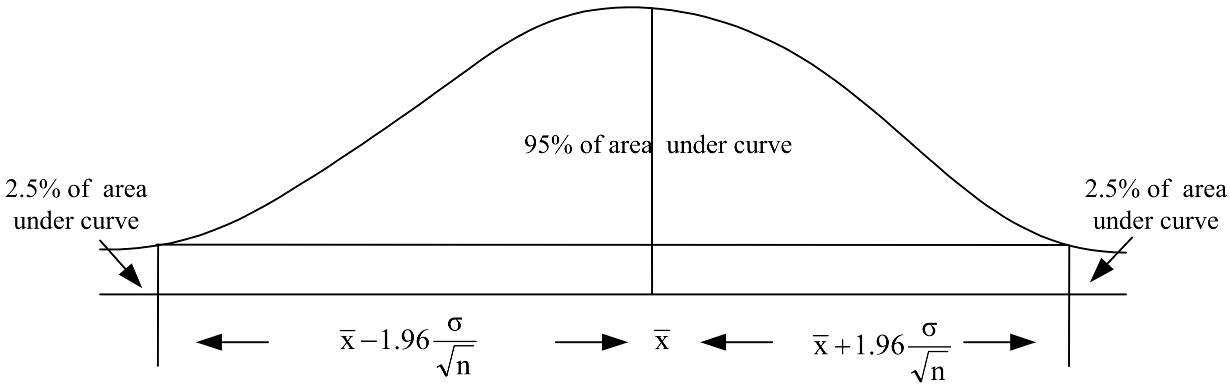

Fig. 15.3: 95% Confidence Interval for Population Mean

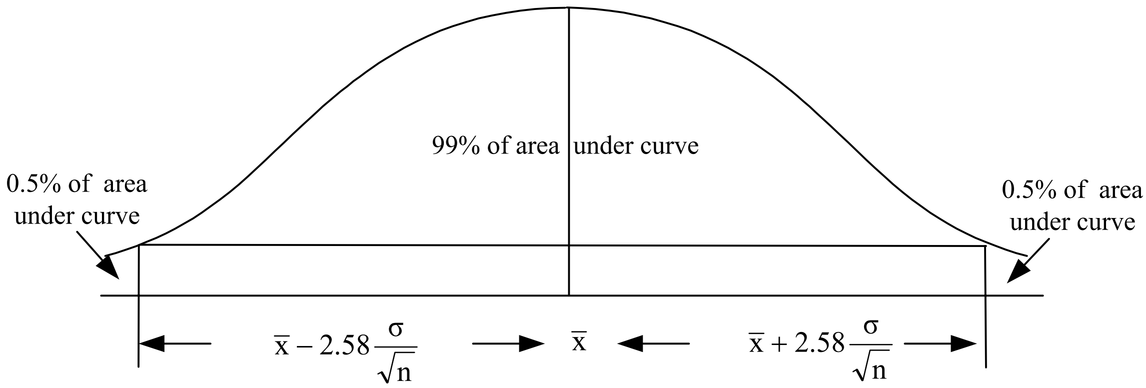

Fig. 15.4: 99% Confidence Interval for Population Mean

Next we consider Interval Estimation in the following cases:

Interval Estimation of Population Mean

As suggested in this section under assumption of normality, 95% confidence interval to μ, the population mean, is given by

$$\left[\bar{x} - 1.96\frac{\sigma}{\sqrt{n}}, \quad \bar{x} + 1.96\frac{\sigma}{\sqrt{n}}\right]$$

If the assumption of normality does not hold but 'n' is greater than 30, the above 95% confidence interval still may be used for estimating population mean. In case σ is unknown, it may be replaced by the corresponding unbiased estimate of σ, namely $S^{|}$, so long as 'n' exceeds 30. However, we may face a difficult situation in case σ is unknown and 'n' does not exceed 30. This problem has been discussed in the next unit (Unit-16). Similarly, 99% confidence interval to μ is given by :

$$\left[\bar{x} - 2.58\frac{\sigma}{\sqrt{n}}, \quad \bar{x} + 2.58\frac{\sigma}{\sqrt{n}}\right]$$

In case σ is unknown. The 99% confidence interval to μ is :

$$\left[\bar{x} - 2.58\frac{S^{|}}{\sqrt{n}}, \quad \bar{x} + 2.58\frac{S^{|}}{\sqrt{n}}\right] \qquad \text{...(15.23)}$$

in case σ is unknown and n > 30.

Interval Estimation of Unknown Population Proportion

It can be assumed that when n is large and neither 'p' nor (1–p) is small (one may specify np ≥ 5 and n (1–p) ≥ 5), then the sample proportion p is asymptotically normal with mean as P and $S.E.(p) = \sqrt{\frac{P(1-P)}{n}}$, P being the unknown population proportion in which we are interested. The estimate of S.E. (p) is given by :

$$\widehat{S.E.}(\hat{p}) = \sqrt{\frac{p(1-p)}{n}}$$

Hence, 95% confidence interval to p is given by :

$$\left[p - 1.96\sqrt{\frac{p(1-p)}{n}},\quad p + 1.96\sqrt{\frac{p(1-p)}{n}}\right] \qquad \ldots(15.24)$$

and 99% confidence interval to P is :

$$\left[p - 2.58\sqrt{\frac{p(1-p)}{n}},\quad p + 2.58\sqrt{\frac{p(1-p)}{n}}\right] \qquad \ldots(15.25)$$

Let us consider the following illustrations to understand the procedure for interval estimation.

Illustration 3

In a random sample of 1,000 families from the city of Delhi, it was found that the average of income as obtained from the sample is Rs. 2,000/-, it is further known that population S.D. is Rs. 258. Find 95% as well as 99% confidence interval to population mean.

Solution: Let x denote income of the people of Delhi city. If μ denotes average income of people dwelling in Delhi, then 95% confidence interval to μ is:

$$\left[\bar{x} - 1.96\frac{\sigma}{\sqrt{n}},\quad \bar{x} + 1.96\frac{\sigma}{\sqrt{n}}\right]$$

and 99% confidence interval to μ is :

$$\left[\bar{x} - 2.58\frac{\sigma}{\sqrt{n}},\ \bar{x} - 2.58\frac{\sigma}{\sqrt{n}}\right]$$

Where $\bar{x}$ = Sample mean; n = Sample size, and σ = Population standard deviation.

In our case,

$\bar{x}$ = Rs. 2000, n = 1000, σ = Rs. 258

$$\therefore \bar{x} - 1.96\frac{\sigma}{\sqrt{n}} = Rs.\,2000 - 1.96 \times \frac{258}{\sqrt{1000}} = Rs.\ 1984.01$$

$$\bar{x} + 1.96\frac{\sigma}{\sqrt{n}} = Rs.\,2000 + 1.96 \times \frac{258}{\sqrt{1000}} = Rs.\ 2015.99$$

$$\bar{x} - 2.58\frac{\sigma}{\sqrt{n}} = \text{Rs. } 2000 - 2.58 \times \frac{258}{\sqrt{1000}} = \text{Rs. } 1979$$

and
$$\bar{x} + 2.58\frac{\sigma}{\sqrt{n}} = \text{Rs. } 2000 + 2.58 \times \frac{258}{\sqrt{1000}} = \text{Rs. } 2021$$

Hence we have

95% confidence interval to average income for the people of Delhi = [Rs. 1984.01 to Rs. 2015.99] and 99% confidence interval to average income for the people of Delhi = [Rs. 1979 to Rs. 2021].

Illustration 4

Calculate the 95% and 99% confidence limits to the average life of fluorescent lights produced by Indian Electricals.

Solution: Since σ, the population standard deviation is unknown and n = 32 (> 30), we replace σ by $S^{|}$, the sample S.D. with divisor as (n–1) in our previous example and get 95% confidence interval to μ is:

$$\left[\bar{x} - 1.96\,\frac{S^{|}}{\sqrt{n}}, \quad \bar{x} + 1.96\,\frac{S^{|}}{\sqrt{n}}\right]$$

Similarly, 99% confidence interval for μ = $\left[\bar{x} - 2.58\frac{S^{|}}{\sqrt{n}}, \quad \bar{x} + 2.58\frac{S^{|}}{\sqrt{n}}\right]$

Where, $\bar{x}$ = Sample mean = 4985 hours, n = Sample size = 32; and

$S^{|}$ = Sample S.D. with (n–1) division = 86.0868 hours (as computed earlier).

$$\therefore \quad \bar{x} - 1.96\frac{S^{|}}{\sqrt{n}} = 4985 - 1.96 \times \frac{86.0868}{\sqrt{32}} = 4955.17 \text{ hours}$$

$$\bar{x} + 1.96\frac{S^{|}}{\sqrt{n}} = 4985 + 1.96 \times \frac{86.0868}{\sqrt{32}} = 5014.83 \text{ hours}$$

$$\bar{x} - 2.58\frac{S^{|}}{\sqrt{n}} = 4985 - 2.58 \times \frac{86.0868}{\sqrt{32}} = 4945.74 \text{ hours}$$

$$\bar{x} + 2.58\frac{S^{|}}{\sqrt{n}} = 4985 + 2.58\,\frac{\times 86.0868}{\sqrt{32}} = 5024.26 \text{ hours}$$

$\therefore$ 95% Confidence Interval to the average life of lights = [4945.17 hours, 5014.83 hours].

99% Confidence Interval to the average life of lights = [4945.74 hours, 5024.26 hours].

Illustration 5

While interviewing 350 people in a city, the number of smokers was found to be 70. Obtain 99% lower confidence limit and the corresponding upper confidence limit to the proportion of smokers in the city.

Solution: As discussed in the previous section, 99% Lower Confidence Limit to P, the proportion of smokers in the city is given by:

$$p-2.58\sqrt{\frac{p\,(1-p)}{n}}$$

and 99% Upper Confidence Limt to P is:

$$p+2.58\sqrt{\frac{p\,(1-p)}{n}}$$

provided np ≥ 5 and np (1–p) ≥ 5.

In this case x = no. of smokers = 70

n = no. of people interviewed = 350

$$\therefore\ \ p=\frac{x}{n}=\frac{70}{350}=0.2$$

As np = 350 × 0.2 = 70 and n (1–p) = 350 × 0.8 = 280 are rather large, we can apply the formula for 99% Confidence Limit as mentioned already.

∴ 99% Lower Confidence Limit to P is :

$$0.2-1.96\times\sqrt{\frac{0.2\times(1-0.2)}{350}}=0.2-0.0214=0.1786$$

99% Upper Confidence Limit to P is :

$$0.2+1.96\times\sqrt{\frac{0.2\times(1-0.2)}{350}}=0.2+0.0214=0.2214$$

Hence 99% Lower Confidence Limit and 99% Upper Confidence Limit for the proportion of smokers in the city are 0.1786 and 0.2214 respectively.

Illustration 6

In a random sample of 19586 people from a town, 2358 people were found to be suffering from T.B. With 95% Confidence as well as 98% Confidence, find the limits between which the percentage of the population of the town suffering from T.B. lies.

Solution: Let x be the number of people suffering from T.B. in the sample and ‘n’ as the number of people who were examined. Then the proportion of people suffering from T.B. in the sample is given by:

$$p\ =\ \frac{x}{n}\ =\ \frac{2358}{19586}\ =\ 0.1204$$

As np = x = 2358 and n (1–p) = n–np = n–x

= 19586–2358 = 17228

are both very large numbers, we can apply the formula for finding Confidence Interval as mentioned in the previous section. Thus 95% Confidence Interval to

P, the proportion of the population of the town suffering from T.B., is given by :

$$\left[p - 1.96\sqrt{\frac{p(1-p)}{n}},\quad p + 1.96\sqrt{\frac{p(1-p)}{n}}\right]$$

$$= \left[0.1204 - 1.96\sqrt{\frac{0.1204 \times (1-0.1204)}{19586}},\ 0.1204 + 1.96\sqrt{\frac{0.1204 \times (1-0.1204)}{19586}}\right]$$

= [0.1204 – 0.0045, 0.1204 + 0.0045] = [0.1181, 0.1227]

In a similar way, 98% Confidence Interval to P is given by:

$$\left[p - 2.33\sqrt{\frac{p(1-p)}{n}},\quad p + 2.33\sqrt{\frac{p(1-p)}{n}}\right]$$

$$= \left[0.1204 - 2.33\sqrt{\frac{0.1204 \times (1-0.1204)}{19856}},\ 0.1204 + 2.33\sqrt{\frac{0.1204 \times (1-0.1204)}{19856}}\right]$$

= [0.1150, 0.1258]

Thereby, we can say with 95% confidence that the percentrage of population in the town suffering from T.B. lies between 11.81 and 12.27 and with 98% confidence that the percentage of population suffering from T.B. lies between 11.50 and 12.58.

Illustration 7

A famous shoe company produces 80,000 pairs of shoes daily. From a sample of 800 pairs, 3% are found to be of poor quality. Find the limits for the number of substandard pair of shoes that can be expected when the Confidence Level is 0.99.

Solution: Let p be the sample proportion of defective shoes as produced by the shoe company. In this case sample size (n) is 800 and population size (N) is 80,000. Since the population is very large, we do not apply finite population correction.

p = 3% = 0.03

$$\therefore \hat{\text{S.E.}}(\hat{p}) = \sqrt{\frac{p(1-p)}{n}} = \sqrt{\frac{0.03\,(1-0.03)}{800}} = 0.0060$$

Thus 99% Lower Confidence Limit to P, the proportion of defective shoes in the daily production of the shoe company is :

$$p - 2.58 \quad \hat{\text{S.E.}}(\hat{p})$$

= 0.03–2.58 × 0.006 = 0.01452

similarly 99% Upper Confidence Limit to p is :

$$P + 2.58\,\hat{\text{S}}\text{.E.}(\hat{P}) = 0.03 + 2.58 \times 0.006 = 0.04548$$

Hence, the Lower limit to the number of substandard i.e., defective pairs of shoes at 99% Level of Confidence = N × 0.01452.

= 80,000 × 0.01452 = 1161.6 = (approximately) 1162

The Upper Limit to the number of substandard, pairs of shoes at 99% Level of Confidence is

$$80{,}000 \times 0.04548 = 3638.4 = \text{(approximately) } 3638$$

Self Assessment Exercise B

1) State with reasons, whether the following statements are true or false.

   a) Confidence Interval provides a range of values that may not contain the parameter.

   b) Confidence Interval is a function of Confidence Co-efficient.

   c) 95% Confidence Interval for population mean is $\bar{x} \pm 1.96\, \text{S.E.}\, (\bar{x})$.

   d) While computing Confidence Interval for population mean, if the population S.D. is unknown, we can always replace it by the corresponding sample S.D.

   e) 99% Upper Confidence Limit for population proportion is $p + 1.96\sqrt{\frac{p\,(1-p)}{n}}$.

   f) Confidence co-efficient does not contain Lower Confidence Limit and Upper Confidence Limit.

   g) If np ≥ 5 and np (1–p) ≥ 5, one may apply the formula $p \pm z_\alpha \sqrt{\frac{p\,(1-p)}{n}}$ for computing Confidence Interval for population proportion.

   h) The interval $\mu \pm 3\, \text{S.E.}\, (\bar{x})$ covers 96% area of the normal curve.

2) Differentiate between Point Estimation and Interval Estimation.

   ....

   ....

3) Distinguish between Confidence Limit and Confidence Interval.

   ....

   ....

   ....

4. Out of 25,000 customer's ledger accounts, a sample of 800 accounts was taken to test the accuracy of posting and balancing and 50 mistakes were found. Assign limits within which the number of wrong postings can be expected with 99% confidence.

   ....

   ....

   ....

5. A sample of 20 items is drawn at random from a normal population comprising 200 items and having standard deviation as 10. If the sample mean is 40, obtain 95% Interval Estimate of the population mean.

   ....

   ....

   ....

6) A new variety of potato grown on 400 plots provided a mean yield of 980 quintals per acre with a S.D. of 15.34 quintals per acre. Find 99% Confidence Limits for the mean yield in the population.

....

....

....

15.5 TESTING HYPOTHESIS – INTRODUCTION

Referring to the problem of the status to be given to Delhi City, one of the criteria for determining the status would be the average income of the people of Delhi. Let us suppose that if 'μ', the average income of the people is Rs. 3,000 per month, then Delhi would belong to the group of top cities. In order to estimate 'μ', we take a random sample of people living in that city and compute $\overline{x}$, the sample mean. If $\overline{x}$ is in the neighbourhood of Rs. 3,000, then we have no hesitation in declaring the status of Delhi as one belonging to the top grade. But the most important question would be as to what difference between the sample mean and Rs. 3000 (population mean) can be accepted as the difference due to only sampling fluctuations.

In order to answer this question, let us familiarise ourselves with a few terms associated with the problem. A statement like 'The average income of the people belonging to the city of Delhi is Rs. 3,000 per month' is known as a **null hypothesis.** Thus, a null hypothesis may be described as an assumption or a statement regarding a parameter (population mean, 'μ', in this case) or about the form of a population. The term 'null' is used as we test the hypothesis on the assumption that there is no difference or, to be more precise, no significant difference between the value of a parameter and that of an estimator as obtained from a random sample taken from the population. A hypothesis may be simple or composite.

A **simple hypothesis** is one that specifies the population distribution completely. Thus testing $\mu = 3{,}000$ is a simple hypothesis if the population standard deviation (σ) is known.

A **composite hypothesis** is one that does not specify the population completely. Testing $\mu = 3{,}000$ when σ is unknown is a composite hypothesis as it does not specify the population completely. A null hypothesis is denoted by H_0. Thus we may write :

$$H_0 : \mu = 3{,}000$$

i.e., the null hypothesis is that the population mean is Rs. 3,000. Generally, we write

$$H_0 : \mu = \mu_0$$

i.e., the null hypothesis is that the population mean is μ, whereas μ_0 may be any value as specified in a given situation.

Obviously a null hypothesis (H_0) is to be tested against an appropriate alternative hypothesis (H_1). **Any hypothesis that contradicts a null hypothesis is known as an alternative hypothesis.** If the null hypothesis is rejected, the alternative hypothesis is accepted. Procedures enabling us to decide whether to accept or reject a hypothesis is known as test of hypothesis or test of significance or decision rule. Thus, the entire process of hypothesis testing is either to reject or accept H_0 only.

In the present problem, one may argue that since many people of Delhi city are living in the slums and even on the pavements, the average income should be less than Rs. 3000. So one alternative hypothesis may be :

$H_1 : \mu < 3{,}000$ i.e., the average income is less than Rs. 3,000 or symbolically:

or, $H_1 : \mu < \mu_0$ i.e., the population mean (μ) is less than μ_0.

Again one may feel that since there are many multistoried buildings and many new models of vehicles run through the streets of the city, the average income must be more than Rs. 3,000. So another alternative hypothesis may be :

$H_2 : \mu > 3000$ i.e., the average income is more than Rs. 3,000.

or, $H_2 : \mu > \mu_0$ i.e., the population mean is more than μ_o.

Lastly, another group of people may opine that the average income is significantly different from μ_0. So the third alternative could be :

$H : \mu \neq 3000$ i.e., the average income is anything but Rs. 3,000.

or, $H : \mu \neq \mu_0$ i.e., the population mean is not μ_0.

15.6 THEORY OF TESTING HYPOTHESIS — LEVEL OF SIGNIFICANCE, TYPE-I AND TYPE-II ERRORS AND POWER OF A TEST

In order to take a decision about acceptance or rejection of a null hypothesis, let us consider the theory involving testing of hypothesis. Suppose that we have a random sample of size ‘n’ taken from a population characterized by an unknown parameter ‘θ’. We denote the n sample observations by $x = (x_1, x_2, x_3, \ldots x_n)$ and we would like to test

$H_0 : \theta = \theta_0$ against

$H_1 : \theta = \theta_1$

If n = 2, then $x = (x_1, x_2)$ can be represented as a point in the 2-dimensional plane taking, say x_1, on the horizontal axis and x_2 on the vertical axis. In a similar way, it is possible to conceive $x = (x_1, x_2, x_3, \ldots x_n)$ as a point in the n-dimensional plane. Consider all the possible samples of a fixed size ‘n’, i.e., ${}^NC_n$ in case of SRSWOR and N^n in the case of SRSWR, N denoting the population size. Next we consider the sample space formed by all these points and let it be denoted by Ω. We divide Ω into two parts ω and $A = \Omega - \omega$, the boundary of ω is taken within A. We frame a simple rule which says that if the sample point x falls on ω, we reject H_0 and if x falls on A, we accept H_0. ω is known as the critical region or rejection region and A, as the acceptance region. At this juncture, let us make one point clear. Acceptance of H_0 does not mean that H_0 is always true. It just reflects the idea that on the basis of the given data, there is not enough evidence to support the validity of H_1. In a similar manner rejection of H_0 indicates the null hypothesis does not hold good in the light of the given sample observations.

Type-I and Type-II Errors

Now while testing H_0 we are liable to commit two types of errors. In the first case, it may be that H_0 is true but x falls on ω and as such, we reject H_0. This is known as type-I error or error of the first kind. Thus type-I error is committed in rejecting a null hypothesis which is, in fact, true. Secondly, it may be that H_0 is false but x falls on A and hence we accept H_0. This is known as type-II error or error of the second kind. So type-II error may be described as the error committed in accepting a null hypothesis which is, in fact, false. The two kinds of errors are shown in Table 15.3.

Table 15.3: Types of Errors in Testing Hypothesis

Real Situation	**Statistical decision based on sample**	
	H_0 Accepted	**H_0 Rejected**
H_0 True	Right decision	Type-I error
H_0 False	Type-II error	Right decision

It is obvious that we should take into account both types of errors and must try to reduce them.Since committing these two types of errors may be regarded as random events, we may modify our earlier statement and suggest that an appropriate test of hypothesis should aim at reducing the probabilities of both types of errors. Let 'α' (read as 'alpha') denote the probability of type-I error and 'β' (read as 'beta') the probability of type-II error. thus by definition, we have

α = The probability of the sample point falling on the critical region when H_0 is true i.e., the value of θ is $\theta_0 = P\ (x \in \omega \mid \theta_0)$ …(15.26)

and β = The probability of the sample point falling on the critical region when H_1 is true, i.e., the value of θ is θ_1

$$= P\ (x \in A \mid \theta_1) \qquad \ldots (15.27)$$

Surely, our objective would be to reduce both type-I and type-II errors. But since we have taken recourse to sampling, it is not possible to reduce both types of errors simultaneously for a fixed sample size. As we try to reduce 'α', β increases and a reduction in the value of β results in an increase in the value of 'α'. Thus, we fix α, the probability of type-I error to a given level (say, 5 per cent or 1 per cent) and subject to that fixation, we try to reduce β, probability of type-II error. 'α' is also known as size of the critical region. It is further known as level of significance as 'α' constitutes the basis for making the difference $(\theta - \theta_0)$ as significant. The selection of 'α' level of significance, depends on the experimenter.

Power of a Test: By definition, we have

$$\beta = P\ (x \in A \mid \theta = \theta_1) \qquad \text{[from 15.27]}$$

$$: \quad 1 - \beta = 1 - P\ (x \in A \mid \theta = \theta_1)$$

$$= P\ (x \in \omega \mid \theta = \theta_1) \qquad \text{[from 15.26]}$$

[Since θ_1 may fall either on ω or A,

therefore, $P\ (x \in \omega \mid \theta = \theta_1) + P\ (x \in A \mid \theta = \theta_1) = 1$

and we have $1 - P\ (x \in A \mid \theta = \theta_1) = P\ (x \in \omega \mid \theta = \theta_1)$]

Now P ($x \in \omega \mid \theta = \theta_1$) is the probability of rejecting H_0 when H_0 is false and the alternative hypothesis H_1 is true which should be the desirable property of an appropriate test. It is obvious that a low value of β would ensure a high value of (1–β). Hence we try to minimize β, the probability of type-II error, as the minimization of β ensures the maximization of (1–β). The expression (1–β) serves as an indicator of the validity of the test as a very high value of (1–β) indicates that the test is doing fine in its endeavour to reject a false hypothesis. Hence (1–β) is known as power of the test as it tells us how well the test under consideration is performing when the null hypothesis is not true. It is obvious that we should try to make our test as powerful as possible subject to a fixed value of α. One may regard power of a test as a function of θ. The function P (θ) = 1–β (θ) is known as the **power function of the test.** The curve obtained by plotting P (θ) against θ is known as **power curve**. Look at the following figure 15.5 which exhibits a **power curve**.

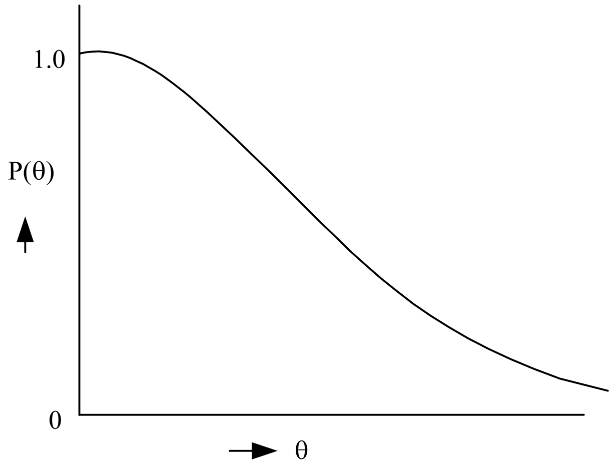

Fig. 15.5: Power Curve of a Test

15.7 TWO-TAILED AND ONE-TAILED TESTS

In order to test the null hypothesis $H_0 : \theta = \theta_0$ against a plausible alternative hypothesis, let us suppose that we find a statistic T which is a sufficient estimator of θ. We assume further that, based on a random sample taken from the population characterized by an unknown parameter θ, it is possible to find a function of T and θ and let u = u (T, θ) by such a function. T is known as test statistic for testing $H_0 : \theta = \theta_0$. Lastly let us assume that when $\theta = \theta_0$, $u_0 = u(T, \theta_0)$ i.e., u_0 is the value of 'u' under H_0 (i.e., assuming the null hypothesis to be true). Based on the sampling distribution of the test statistic u under H_0, it may be possible to find 4 values of u, namely, $u_{\alpha/2}$, $u_{(1-\alpha/2)}$, u_α and $u_{(1-\alpha)}$ for a fixed level of significance α, such that :

$$P\ (u_0 \geq u_{\alpha/2}) = \frac{\alpha}{2} \qquad \ldots (15.28)$$

$$P\ (u_0 \leq u_{(1-\alpha/2)}) = \frac{\alpha}{2} \qquad \ldots (15.29)$$

$$P\ (u_0 \geq u_\alpha) = \alpha \qquad \ldots (15.30)$$

$$P\ (u_0 \leq u_{(1-\alpha)}) = \alpha \qquad \ldots (15.31)$$

u_α may be described as the upper α-point of the distribution of u and $u_{(1-\alpha)}$ as the corresponding lower α-point.

Two-tailed test: Adding (15.28) and (15.29), we get :

$$P\ (u_0 \geq u_{\alpha/2}) + P\ (u_0 \leq u_{(1-\alpha/2)}) = \alpha \qquad \ldots (15.32)$$

i.e., the probability that u_0 would exceed $u_{\alpha/2}$ or u_0 is less than $u_{(1-\alpha/2)}$ is α.

In order to test $H_0 : \theta = \theta_0$ against $H_1 : \theta \neq \theta_0$, if we select a low value of α, say $\alpha = 0.01$, then (15.32) suggests that the probability u_0 is greater than $u_{a/2}$ or u_0 is less than $u_{(1-\alpha/2)}$ is 0.01 which is pretty low. So on the basis of a random sample drawn from the population, if it is found that u_0 is greater than $u_{a/2}$ or u_0 is less than $u_{(1-\alpha/2)}$, then we have rather strong evidence that H_0 is not true. Then we reject $H_0 : \theta = \theta_0$ and accept the alternative hypothesis $H_1 : \theta \neq \theta_0$. As shown in the following Figure 15.6, here the critical region lies on both tails of the probability distribution of u.

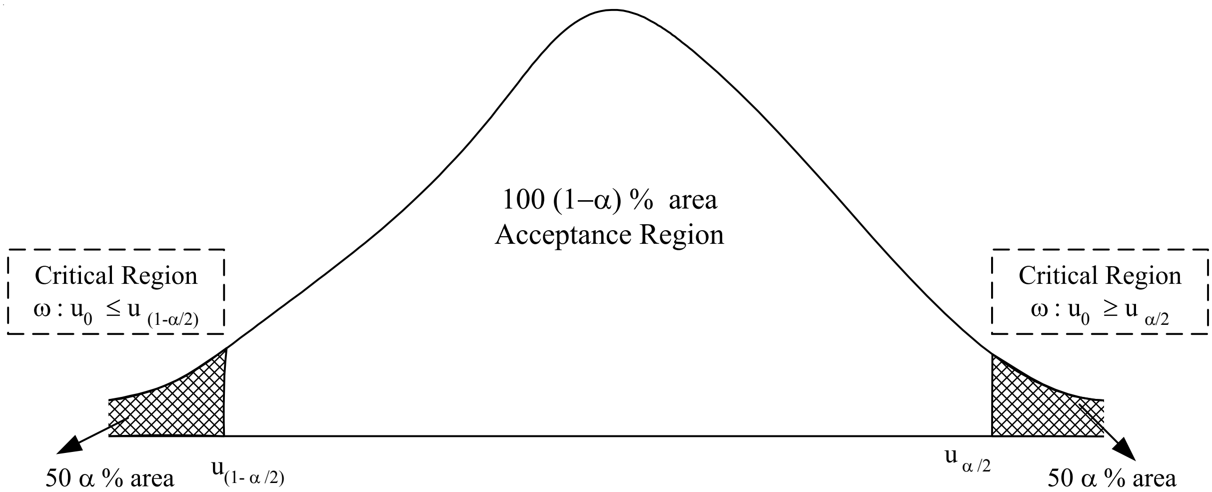

Fig. 15.6: Critical region of a two-tailed Test

If the sample point x falls on one of the two tails, we reject H_0 and accept $H_1 : \theta \neq \theta_0$. The statistical test for $H_0 : \theta = \theta_0$ against $H_1 : \theta \neq \theta_0$ is known as both-sided test or two-tailed test as the critical region, 'ω' lies on both sides of the probability curve, i.e., on the two tails of the curve. The critical region is $\omega : u_0 \geq u_{\alpha/2}$ and $\omega : u_0 \leq u_{(1-\alpha/2)}$. It is obvious that a two-tailed test is appropriate when there are reasons to believe that 'u' differs from θ_0 significantly on both the left side and the right side, i.e., the value of the test statistic 'u' as obtained from the sample is significantly either greater than or less than the hypothetical value.

For testing the null hypothesis $H_0 : \mu = 3000$, i.e., the average income of the people of Delhi city is Rs. 3000, one may think that the alternative hypothesis would be $H_1 : \mu \neq 3000$ i.e., the average income is not Rs. 3000 and as such, we may advocate the application of a two-tailed test. Similarly, for testing the null hypothesis that the average life of lights produced by Indian Electricals is 5,000 hours against the alternative hypothesis that the average life is not 5,000 hours, i.e., for testing $H_0 : \mu = 5{,}000$ against $H_1 : \mu \neq 5{,}000$, we may prescribe a two-tailed test. In the problem concerning the health of city B, we may be interested in testing whether 20% of the population of city B really suffers from T.B. i.e., testing $H_0 : P = 0.2$ against $H_1 : P \neq 0.2$ and again a two-tailed test is necessary and lastly regarding the harms of smoking, we may like to test $H_0 : P = 0.3$ against $H_1 : P \neq 0.3$.

Right-tailed Tests

We may think of testing a null hypothesis against another pair of alternatives. If we wish to test $H_0 : \theta = \theta_0$ against $H_1 : \theta > \theta_0$, then from (15.30) we have $P(u_0 \geq u_\alpha) = \alpha$. This suggests that a low value of α, say $\alpha = 0.01$, implies that the probability that u_0 exceeds u_α is 0.01. So the probability that u_0 exceeds u_α is rather small. Thus on the basis of a random sample drawn from this population if it is found that u_0 is greater than u_α, then we have enough evidence to suggest that H_0 is not true. Then we reject H_0 and accept H_1. This is exhibited in Figure 15.7 as shown below:

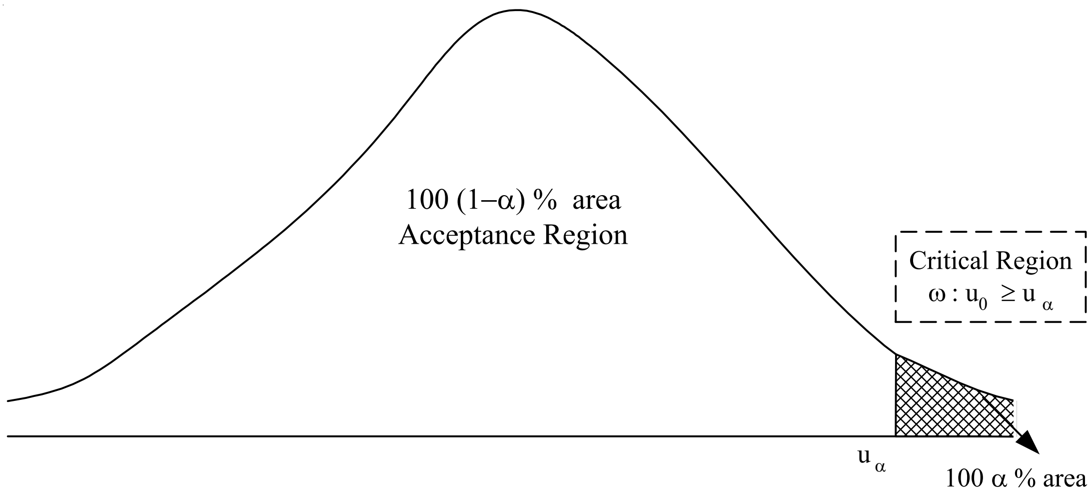

Fig. 15.7: Critical region of a right-tailed Test

As shown in figure 15.7, the critical region lies on the right tail of the curve. This is a one-sided test and as the critical region lies on the right tail of the curve, it is known as **right-tailed test or upper-tailed test**. We apply a right-tailed test when there is evidence to suggest that the value of the statistic u is significantly greater than the hypothetical value θ_0. In case of testing about the average income of the citizens of Delhi, if one has prior information to suggest that the average income of Delhi is more than Rs. 3,000, then we would like to test $H_0 : \mu = 3{,}000$ against $H_1 : \mu > 3{,}000$ and we select the right-tailed test. In a similar manner for testing the hypothesis that the average life of lights by Indian Electricals is more than 5,000 hours or for testing the hypothesis that more than 20 per cent suffer from T.B. in city B or for testing the hypothesis that the per cent of smokers in town C is more than 30, we apply the right-tailed test.

Left-tailed test

Lastly, we may be interested to test $H_0 : \theta = \theta_0$ against $H_2 : \theta < \theta_0$.From (15.31), we have $P(u_0 \leq u_{1-\alpha}) = \alpha$. Choosing $\alpha = 0.01$, this implies that the probability that u_0 would be less than u_α is 0.01, which is surely very low. So, if on the basis of a random sample taken from the population, it is found that u_0 is less than $u_{1-\alpha}$, then we have very serious doubts about the validity of H_0. In this case, we reject H_0 and accept $H_2 : \theta < \theta_0$. This is reflected in Figure 15.8 shown below.

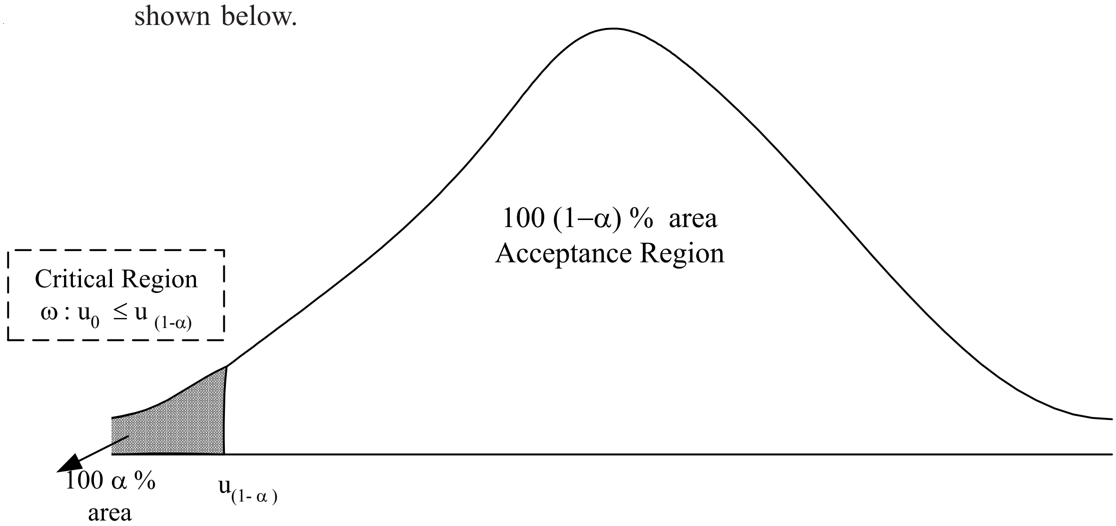

Fig. 15.8: Critical Region of a Left-tailed Test

The test for $H_0 : \theta = \theta_0$ against $H_2 : \theta < \theta_0$ is another one-sided test and as the critical region lies on the left tail of the curve, this is known as a **left-**

tailed test or a lower-tailed test. We apply a left-tailed test when there is enough indication to suggest that the value of the test statistic 'u' is significantly less than the hypothetical value. Then for determining the status of Delhi city, if somebody suggests with evidence that the average income is less than Rs. 3,000 and as such Delhi should not be regarded as a top grade city, then we are to test $H_0 : \mu = 3000$ against $H_1 : \mu < 3000$, which is a left-tailed test. We may further note that we apply left-tailed test when we would like to test the hypothesis that the average life of lights of Indian Electricals is less than 5,000 hours or less than 20 per cent are suffering from T.B. in city B or less than 30 per cent are smokers in town C.

15.8 STEPS TO FOLLOW FOR TESTING HYPOTHESIS

While testing hypothesis, one must go through the following steps.

1) Set up the null hypothesis $H_0 : \theta = \theta_0$ and one of the alternative hypothesis $H : \theta \neq \theta_0$ or $H_1 : \theta > \theta_0$ or $H_2 : \theta < \theta_0$ depending upon the problem. Selecting the proper alternative plays a significant role in decision making in connection with testing of hypothesis.

2) Choose the appropriate test statistic 'u' and sampling distribution of 'u' under H_0. In most cases 'u' follows a standard normal distribution under H_0 and hence Z-test can be recommended in such a case.

3) Select α, the level of significance of the test if it is not provided in the given problem. In most cases, we choose $\alpha = 0.05$ and $\alpha = 0.01$ which are known as 5% level of significance and 1% level of significance.

4) Define critical region ω, based on the alternative hypothesis. For testing $H_0 : \theta = \theta_0$ against both-sided alternative $H_1 : \theta \neq \theta_0$, the critical region is given by $\omega : u_0 \geq u_{\alpha/2}$ and $\omega : u_0 \leq u_{(1-\alpha/2)}$. Similarly, the critical region for the right-sided alternative is given by $\omega : u_0 \geq u_\alpha$ and the critical region for the left-sided is given by $\omega : u_0 \leq u_{1-\alpha}$.

5) Obtain the value of u_0 on the basis of the given sample observations.

6) Reject H_0 if u_0 falls on ω. Otherwise accept H_0.

7) Draw your own conclusion in very simple language which should be understood even by a layman.

15.9 TESTS OF SIGNIFICANCE FOR POPULATION MEAN – Z-TEST FOR VARIABLES

Let us assume that we have taken a random sample of size 'n' from a normal population with mean as μ and standard deviation as σ. Let the sample observations be denoted by $x_1, x_2, x_3, \ldots x_n$. While testing for the unknown population mean μ, we are to consider the following cases.

Case 1: When the standard deviation σ is known. We want to test $H_0 : \mu = \mu_0$ against one of the following alternative hypothesis.

$H : \mu \neq \mu_0$ or,

$H_1 : \mu > \mu_0$ or,

$H_2 : \mu < \mu_0$.

As we have discussed in Section 15.2, the best statistic for the parameter μ is $\bar{x}$. It has been proved in that Section, $E(\bar{x}) = \mu$.

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

As such the test statistic :

$$z = \frac{\bar{x} - E(\bar{x})}{S.E.(\bar{x})} = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

is a standard normal variable. Under H_0, i.e., assuming the null hypothesis to be true,

$z_0 = \dfrac{\sqrt{n}(x - \mu_0)}{\sigma}$ is a standard normal variable. As such, the test is known as standard normal variate test or standard normal deviate test or Z-test. In order to find the critical region for testing H_0 against H from (15.28) and (15.29), we find that :

$$P(u_0 \geq u_{(\alpha/2)}) = \frac{\alpha}{2}$$

and $$P(u_0 \leq u_{(1-\alpha/2)}) = \frac{\alpha}{2}$$

If we denote the standard normal variate by Z, and the upper α-point of the standard normal distribution by Z_α, and by $Z_{(1-\alpha/2)} = -Z_{\alpha/2}$, (as the standard normal distribution is symmetrical about 0), the lower α-point of the standard normal distribution, then the above two equations are reduced to :

$$P(Z_0 \geq Z_{\alpha/2}) = \frac{\alpha}{2} \qquad \text{……(15.33)}$$

and $$P(Z_0 \leq -Z_{\alpha/2}) = \frac{\alpha}{2} \qquad \text{……(15.34)}$$

From (15.33), we have:

$$1 - P(Z_0 < Z_{(\alpha/2)}) = \frac{\alpha}{2}$$

or $$1 - \underline{\bar{\phi}}(Z_{\alpha/2}) = \frac{\alpha}{2}$$

or $$\underline{\bar{\phi}}(Z_{\alpha/2}) = 1 - \frac{\alpha}{2}$$

choosing $\alpha = 0.05, \underline{\bar{\phi}}(Z_{0.025}) = 1 - 0.025 = 0.975$

or, $\underline{\bar{\phi}}(Z_{0.0025}) = \underline{\bar{\phi}}(1.96)$ [from Section 15.4]

Thus, $Z_{0.025} = 1.96$

Hence from (15.33) and (15.34), we have

$P(Z_0 \geq 1.96) = 0.025$

and $P(Z_0 \leq -1.96) = 0.025$

Combining these two equations, we get $P(|z_0| \ge 1.96) = 0.05$ ......[15.35]

Thus for testing H_1: $\mu \ne \mu_0$, the critical region is given by :

$$\omega : |z_o| \ge 1.96$$

When the level of significance is 5% and

$$Z_0 = \frac{\sqrt{n}(x - \mu_0)}{\sigma} \quad ......(15.36)$$

Proceeding in a similar manner, the critical region for the two-tailed test at 1% level of significance is given by :

$$\omega : |z_o| \ge 2.58 \quad ......(15.37)$$

Now if we decide to test H_0 against the alternative hypothesis $H_1 : \mu > \mu_0$, from (15.30), we have

$$P(Z_0 \ge Z_\alpha) = \alpha$$

or, $1 - P(Z_0 < Z_\alpha) = \alpha$

or, $1 - \underline{\overline{\phi}}(Z_\alpha) = \alpha$

or, $1 - \underline{\overline{\phi}}(Z_{\alpha/2}) = \dfrac{\alpha}{2}$ ...... (15.38)

Putting $\alpha = 0.05$ in (15.38), we get

$$\underline{\overline{\phi}}(Z_{0.05}) = 0.95 = \underline{\overline{\phi}}(1.645) \quad ...... \text{[from Section 15.4]}$$

Hence the critical region for this right-tailed test at 5% level of significance is :

$$\omega : Z_0 \ge 1.645$$

Similarly the critical region at 1% level of significance would be :

$$\omega : Z_0 \ge 2.35$$

Finally if we make up our minds to test H_0 against H_2 : $\mu < \mu_0$, then from (15.35), we get

$$P(Z_0 \le -Z_\alpha) = \alpha$$

or, $\underline{\overline{\phi}}(-Z_\alpha) = \alpha$

or, $1 - \underline{\overline{\phi}}(Z_\alpha) = \alpha$

or, $\underline{\overline{\phi}}(Z_\alpha) = 1 - \alpha$

Thus as before, putting $\alpha = 0.05$, we get $Z_\alpha = 1.645$

And as such the critical region for this left-tailed test at 5% level of significance is:

$$\omega : Z_0 \le -1.645$$

The critical region, when the level of significance is 1%, is :

$$\omega : Z_0 \le -2.35$$

Figures 15.9 and 15.10 describe critical regions at 5% and 1% level of significance for **Two-tailed Tests**. Right-tailed tests are shown in Figures 15.11 and 15.12 and left-tailed tests are exhibited in Figures 15.13 and 15.14.

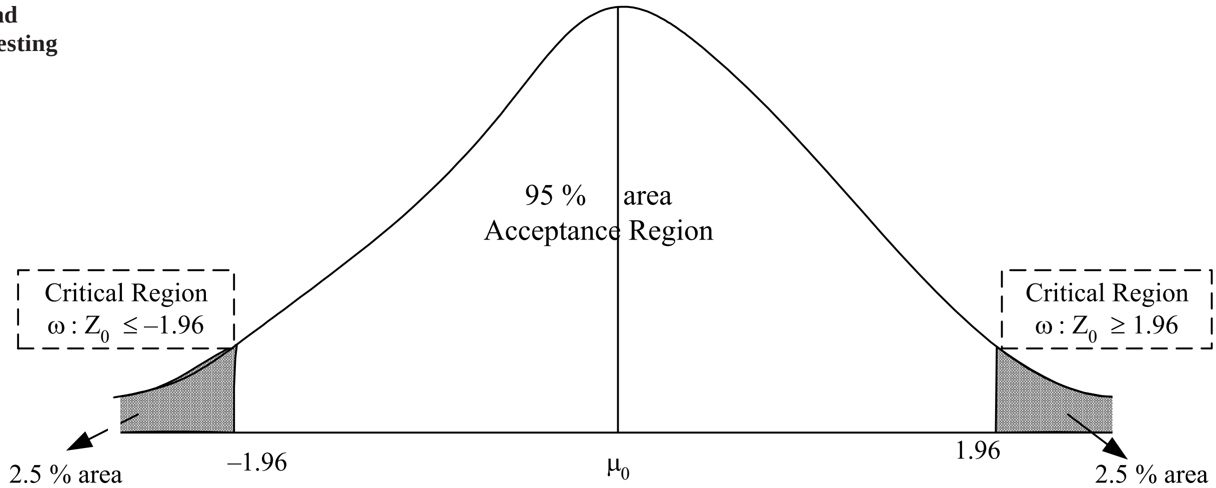

Figure 15.9: Two-tailed Critical Region for Testing Population Mean at 5% Level of Significance

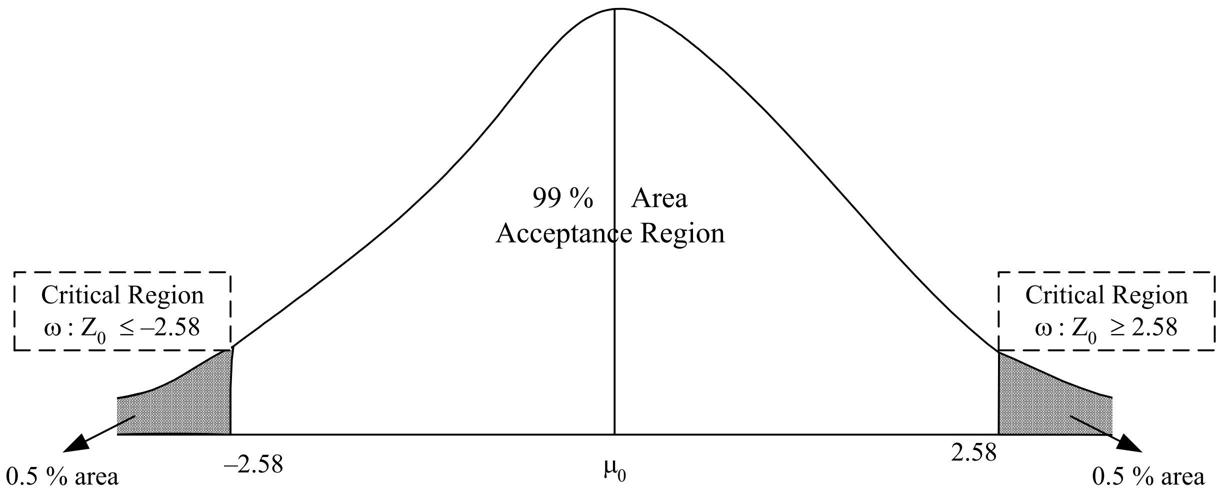

Figure 15.10: Two-tailed Critical Region for Testing Population Mean at 1% Level of Significance.

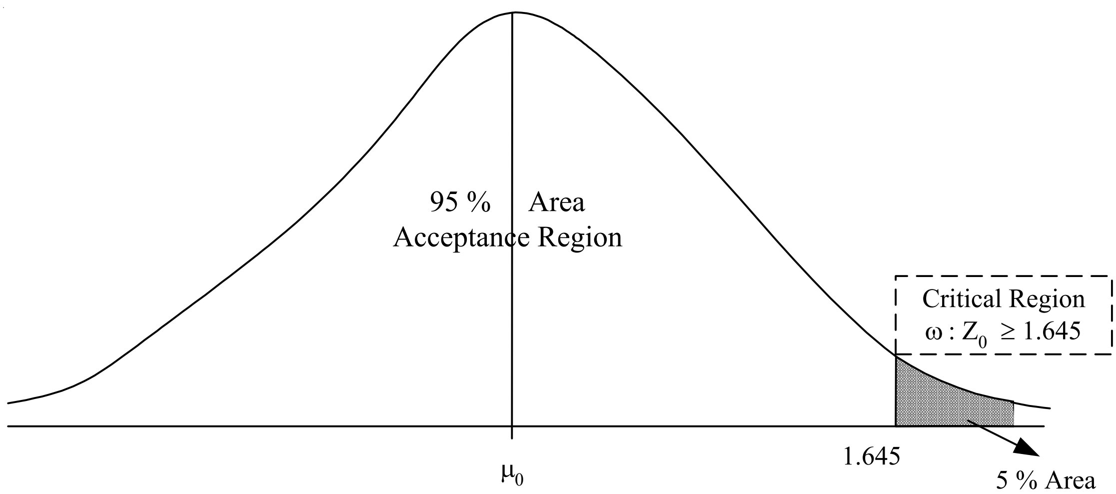

Figure 15.11: Right-tailed Critical Region for Testing Population Mean at 5% Level of Significance

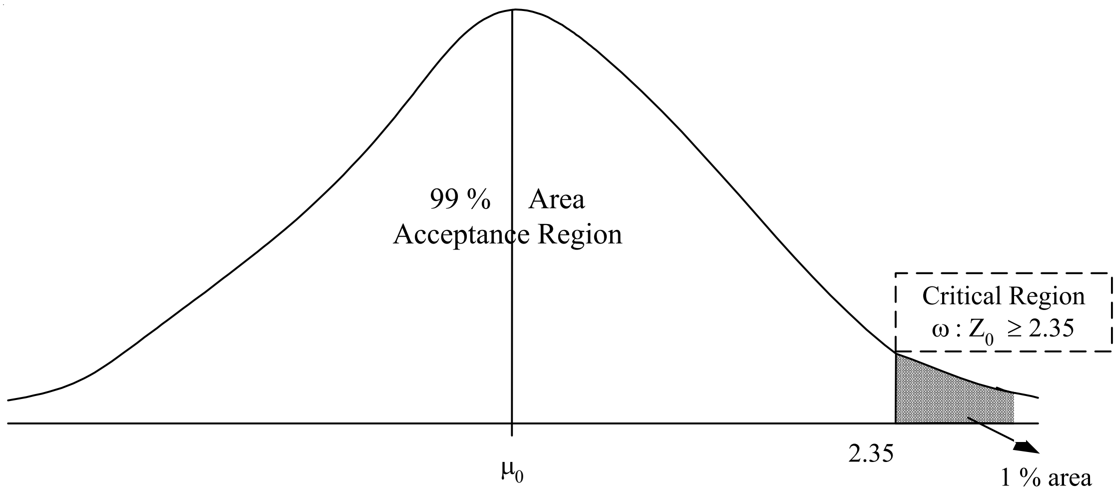

Figure 15.12: Right-tailed Critical Region for Testing Population Mean at 1% Level of Significance

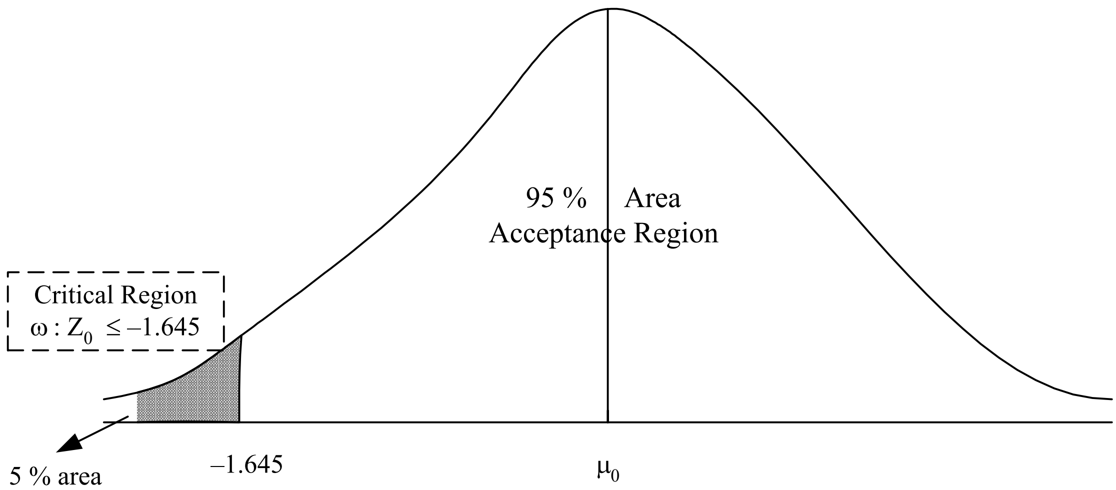

Figure 15.13: Left-tailed Critical Region for Testing Population Mean at 5% Level of Significance

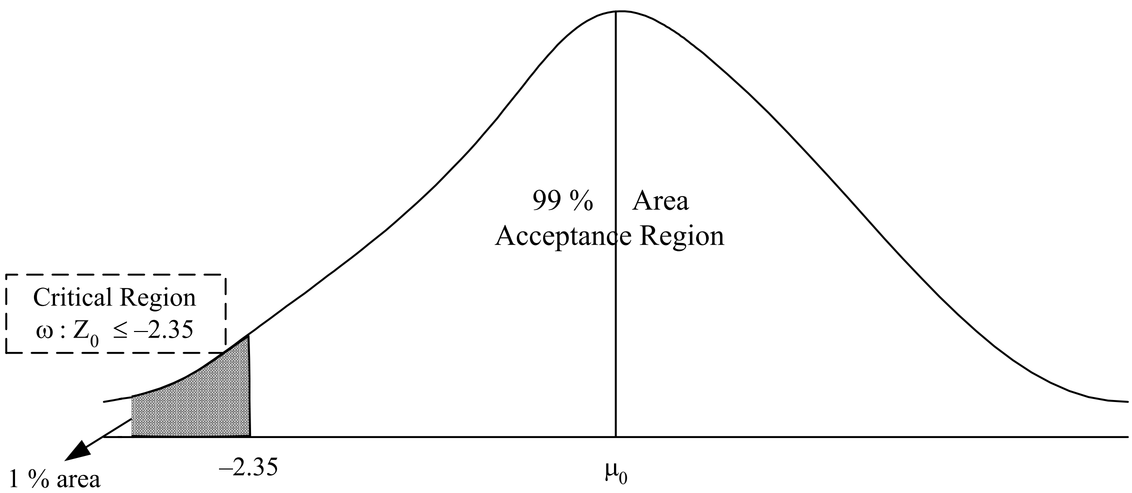

Figure 15.14: Left-tailed Critical Region for Testing Population Mean at 1% Level of Significance

Case II: When the population standard deviation is unknown.

In order to test for population mean, we replace σ by its unbiased estimator.

$$S' = \sqrt{\frac{\Sigma(X_i - \bar{x}^2)}{n-1}}$$

in the test statistic used in Case-I, provided we have a sufficiently large sample (as discussed earlier n should exceed 30). Thus we consider

$$Z_0 = \frac{\sqrt{n}(\bar{x} - \mu_0)}{s^{|}}$$

Z_0 is a standard normal variable. As before for testing $H_0 : \mu = \mu_0$ against both-sided alternative $H : \mu \neq \mu_0$, the critical region at 5% level of significance would be given by :

$$\omega \,|\, Z_0 \,| \geq 1.96$$

Also the critical region at 1% level of significance would be

$$\omega : | Z_0 | \geq 2.58$$

Further the critical region at 5% level of significance for the right-sided alternative $H_1 : \mu > \mu_0$ would be :

$$\omega : | Z_0 | \geq 1.645$$

and $\omega : Z_0 \geq 2.33$ when the level of significance is 1%.

Lastly the critical region for the left-sided alternative $H_2 : \mu > \mu_0$ would be provided by :

$$\omega : Z_0 \leq -1.645$$

and $\omega : Z_0 \leq -2.33$ when $\alpha = 0.05$ and 0.01 respectively.

15.10 TESTS OF SIGNIFICANCE FOR POPULATION PROPORTION – Z-TEST FOR ATTRIBUTES

We consider now the problem of testing $H_0 : P = P_0$, i.e., testing whether the proportion of units in the population possessing a certain characteristic is P_0, i.e., a specified value.

For example, if we want to test whether a fresh coin just coming out from a mint is unbiased, then we are to test $H_0 : P = 0.5$. Similarly, the problem of testing whether 20% population of city B is suffering from T.B amounts to testing $H_o : P = 0.2$ or testing whether 30% population of a town are smokers is equivalent to testing $H_0 : P = 0.3$.

As discussed earlier, the number of units in the population having a certain characteristic follows Binomial Distribution with parameters 'n' and P. If 'n' is large such that both nP and n(1–P) are not less than 5, then we can approximate a Binomial Distribution by a Normal Distribution with mean as μ = nP and variance as σ^2 = nP (1–P).

Hence, it follows that the sample proportion (p) = $\frac{x}{n}$ follows normal distribution

with mean as P_0 and S.D. as $\sqrt{\frac{P_0\,(1-P_0)}{n}}$ under H_0.

$$\text{Thus } Z_0 = \frac{p - P_0}{\sqrt{\dfrac{P_0(1-P_0)}{n}}} = \frac{\sqrt{n}(p - P_0)}{\sqrt{P_0(1-P_0)}}$$

is a standard normal variate and as such we can apply Z-test for attributes.

Hence, as discussed earlier, the critical region for testing $H_0 : P = P_0$ against two-sided alternative $H : P \neq P_0$ would be given by :

$\omega : |Z_0| \geq 1.96$ when the level of significance is 5% and by

$\omega : |Z_0| \geq 2.58$ at 1% level of significance.

The critical regions for the right-sided alternative $H_1 : P > P_0$ at 5% level of significance and 1% level of significance would be:

$\omega : Z_0 \geq 1.645$ and

$\omega : Z_0 \geq 2.33$ respectively.

Lastly when it comes to testing H_0 against the left-sided alternative $H_2 : P < P_0$,

We have the critical regions as $\omega : Z_0 \leq -1.645$ when $\alpha = 0.05$

and $\omega : Z_0 \leq -2.33$ when $\alpha = 0.01$

Let us consider the following illustrations to understand the application of this concept.

Illustration 8

The mean breaking strength of the cables supplied by a manufacturer is 1,900 units with a standard deviation of 110 units. By a new technique in the manufacturing process, the manufacturer claims, the breaking strength of the cables supplied by him has increased. In order to test his claim, a sample of 50 cables is tested. It is found that the mean breaking strength, as obtained from the sample, is 1926. Can you support the claim both at 5% and 1% levels of significance?

Solution: Let the mean breaking strength of the cables be denoted by $\bar{x}$. Since the sample size (n) is 50 which is more than 30, we can apply Z-test. Then we are to test,

$H_0 : \mu = 1900$ i.e., the mean breaking strength of the cables is 1900 units

Against $H_1 : \mu > 1900$; i.e., the mean breaking strength has increased

we use $Z_0 = \dfrac{\sqrt{n}\,(\bar{x} - 1900)}{\sigma}$

The critical region for this right-sided alternative is given by :

$\omega : Z_0 \geq 1.645$ at 5% level of significance and

$\omega : Z_0 \geq 2.33$ at 1% level of significance

As per given data, $n = 50$, $\bar{x} = 1926$, $\sigma = 110$

$$Z_0 = \frac{\sqrt{50}(1926 - 1900)}{110} = 1.671$$

Thus, we reject H_0 at 5% level of significance but accept the null hypothesis at 1% level of significance. On the basis of the given data, we thus conclude that the manufacturer's claim is justifiable at 5% level of significance but at 1% level of significance, we infer that the manufacturer has been unable to produce cables with a higher breaking strength.

Illustration 9

A random sample of 500 flower stems has an average length of 11 cm. Can this be regarded as a sample from a large population with mean as 10.8 cm and standard deviation as 2.38 cm?

Solution: Let the length of the stem be denoted by x. Assume that μ denotes the mean of stems in the population. The sample size 500 being very large, we apply Z-test for testing $H_0 : \mu = 10.8$, i.e., the population mean is 10.8 cm. against $H : \mu \neq 108$, i.e., the population mean is not 10.8.

As such we consider, as test statistic :

$$Z_0 = \frac{\sqrt{n}(\bar{x} - 10.8)}{\sigma}$$

and choosing the level of significance as 5%, we note that the critical region is :

$$\omega : |Z_0| \geq 1.96$$

as per given data,

$$m = 500, \quad \bar{x} = 11 \text{ cm}, \quad \sigma = 2.38 \text{ cm}$$

$$\therefore \quad Z_0 = \frac{\sqrt{500}\ (11 - 10.8)}{2.38} = 1.879$$

Thus we accept H_0. We conclude that on the basis of the given data, the sample can be regarded as taken from a large population with mean as 10.8 cm and standard deviation as 2.38 cm.

Illustration 10

A manufacturer of batteries asserts that the batteries made by him have a mean life of 650 hours with a standard deviation of 12.83 hours. Ten batteries were tested and the length of life of the batteries was recorded in hours as follows:

623, 648, 672, 685, 692, 650, 649, 666, 638, 629

Examine whether the manufacturer was right in his assertion.

Solution: We assume that x, the length of battery-life is normally distributed with mean as 650 hours and standard deviation as 12.83 hours. We are interested in testing $H_0 : \mu = 650$ i.e., the average life is less than 650 hours against $H_1 : \mu < 650$, i.e., the average life is less than 650 hours.

We consider $\quad Z_0 = \dfrac{\sqrt{n}\ (\bar{x} - 650)}{\sigma}$

and recall that the critical region at 1% level of significance (selecting α = 0.01) for this left-tailed test is given by

$$\omega : Z_0 < -2.33$$

since n = 10, σ = 12.83 hours, and

$$\bar{x} = \frac{623 + 648 + 672 + 685 + 692 + 650 + 649 + 666 + 638 + 629}{10} = 655.2 \text{ hours}$$

$$\therefore Z_0 = \frac{\sqrt{10}\,(655.2 - 650)}{12.83} = 1.282$$

As this does not fall in the critical region, H_0 is accepted. Thus on the basis of the given sample, we conclude that the manufactuer's assertion was right.

Illustration 11

The heights of 12 students taken at random from St. Nicholas College, which has 1,000 students and a standard deviation of height as 10 inches, are recorded in inches as 65, 67, 63, 69, 71, 70, 65, 68, 63, 72, 61 and 66. Do the data support the hypothesis that the mean height of all the students in that college is 68.2 inches?

Solution: Letting x stand for height of the students of St. Nicholas College, we would like to test

H_0 : μ = 68.2 i.e., the mean height is 68.2 inches against H : $\mu \neq 68.2$, i.e. the mean height is not 68.2 inches.

The critical region for this two-tailed test is :

$$\omega : |Z_0| \geq 1.96 \qquad \text{when } \alpha = 0.05$$

$$\omega : |Z_0| \geq 2.58 \qquad \text{when } \alpha = 0.01$$

where $Z_0 = \dfrac{\bar{x} - 68.2}{\text{S.E.}(\bar{x})}$

In this case,

$$\bar{x} = \frac{65 + 67 + 63 + 69 + 71 + 70 + 65 + 68 + 63 + 72 + 61 + 66}{12} = 66.67 \text{ inches}$$

n = sample size = 12; N = population size = 1000; σ = population S.D. = 10 inches

$$\text{S.E.}(\bar{x}) = \frac{\sigma}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}$$

$$= \frac{10}{\sqrt{12}}\sqrt{\frac{1000-12}{1000-1}} = 2.9027 \text{ inches}$$

$$\therefore Z_0 = \frac{66.67 - 68.2}{2.9027} = 0.527$$

Looking at ω, we accept H_0 at both 5% and 1% levels of significance. So on the basis of the given data, we comment that the mean height of the students of St. Nicholas is 68.2 inches.

Illustration 12

A coin is tossed 950 times and heads appear 500 times. Does the result support the hypothesis that the coin is unbiased? Select $\alpha = 0.01$.

Solution: As explained in this section, we would denote by P the probability of getting a head. So testing the hypothesis that the coin is unbiased amounts to testing H_0 : P = 0.05 against H : P ≠ 0.5, i.e., the coin is biased.

Since n = 950; $nP_0 = 950 \times 0.5 = 475$; and $nP_0\ (1 – P_0) = 237.5$, we can apply Z-test for proportion. Thus we compute :

$$Z_0 = \frac{\sqrt{n}(p-0.5)}{\sqrt{0.5\,(1-0.5)}}$$ and note that the critical region at 1% level of significance for this two-tailed test is :

$$\omega : |Z_0| \geq 2.58$$

As $$p = \frac{x}{n} = \frac{500}{950} = 0.5263$$

$$\therefore Z_0 = \frac{\sqrt{950} \times (0.5263 - 0.5)}{0.5} = 1.176$$

So we accept H_0. On the basis of the given data, we conclude that the coin is unbiased.

Illustration 13

In a sample of 800 parts manufactured by a company, number of defective parts was found to be 60. The company, however, claims that only 7% of their product is defective. Apply an appropriate test to verify whether the manufacturer's claim is tenable.

Solution: Let 'p' be the sample proportion of defectives and P, the proportion of defective parts in the whole manufacturing process. Then we are to test

H_0 : P = 0.07, i.e., the proportion of defective parts in the process is 7% as claimed by the manufacturer against H_1 : P > 0.07, i.e., the proportion of defective parts is more than 7%.

We consider Z-test as $nP_0 = 800 \times 0.07 = 56$ as well as $nP_0\ (1 – p_0) = 800 \times 0.07 \times 0.93 = 52.09$ are quite large.

If we select $\alpha = 0.05$, then the critical region for this right-tailed test is :

$$\omega : Z_0 \geq 1.645$$

We have, as given, $$p = \frac{x}{n} = \frac{60}{800} = 0.075$$

$$Z_0 = \frac{\sqrt{n}\ (p-0.07)}{\sqrt{0.07 \times 0.93}} = \frac{\sqrt{800}\ (0.075-0.07)}{\sqrt{0.07 \times 0.93}}$$

= 0.5543, we ignore f.p.c as the population size is unknown.

Thus, Z_0 falls on the acceptance region and we accept the null hypothesis. We conclude that on the basis of the given information, the manufacturer's claim is valid.

Illustration 14

A family-planning activist claims that more than 33 per cent of the families in her town have more than one child. A random sample of 160 families from the town reveals that 50 families have more than one child. what is your inference ? Select $\alpha = 0.01$.

Solution: If 'P' denotes the proportion of families in the town having more than one child, then we want to test $H_0 : P = 0.33$ against $H_1 : P > 0.33$.

We consider $Z_0 = \frac{\sqrt{n}(p - 0.33)}{\sqrt{0.33(1 - 0.33)}}$ as test statistic and note that at 1% level of significance the critical region is $\omega : Z_0 \geq 2.35$.

Here, $p = \frac{50}{160} = 0.3125$, $n = 160$

$$\therefore Z_0 = \frac{\sqrt{160}\,(0.3125 - 0.33)}{\sqrt{0.33(1 - 0.33)}} = -0.4708$$

Thus H_0 is accepted and the claim of the activist is justifiable at 1% level of significance on the basis of the given sample.

Self Assessment Exercise C

1) Examine whether the following statements are true or false:

   a) A statistical hypothesis is an assumption about some parameter.

   b) A reduction of type-I error results in an increase in type-II error.

   c) Power of a test is a function of type-I error.

   d) Type-II error is committed when we reject a true null hypothesis.

   e) Probability of type-I error is also known as the level of significance of the test.

   f) The critical region for the two-tailed test for population mean at 5% level of significance is $\omega : |Z_0| \geq 2.58$.

   g) Z-test for population proportion is an exact test.

   h) When the sample size is very large, any test can be approximated by a Z-test.

2) Distinguish between TYPE-I and TYPE-II errors.

   ....

   ....

   ....

3) Differentiate between one-tailed tests and two-tailed tests.

   ....

   ....

   ....

4) A sample of 5 units is taken from a normal population having variance as 4 squared units. the sample observations are 23, 32, 35, 28 and 30. Do the data suggest that the population mean is 30 units? Test at 5% level of significance.

....

....

....

....

5) A producer making electronic components claims that not more than 2% of his components are defective. A sample of 300 components resulted in 16 defectives. Would you support his view ?

....

....

....

....

6) The numbers of male and female births in a hospital during a month were found to be 1980 and 1870 respectively. Do the data confirm to the hypothesis that the sexes are born in the same ratio?

....

....

....

....

15.11 LET US SUM UP

Statistical inference is a method to throw some light on the unknown population with the help of a sample drawn from it. There are two types of estimates with respect to estimating a parameter. They are : a) point estimates; b) interval estimates. We estimate a parameter with the help of a single value known as Point Estimate or a pair of values, known as Interval Estimate.

In a somewhat different situation, some information about some characteristic(s) of the population may be known and we would like to examine whether that information holds good for the sample as well. This is known as Test of hypothesis or test of Significance or Decision rule. While testing a hypothesis, one is likely to commit two types of Errors. Type-I error is committed in rejecting a true null hypothesis and Type-II error occurs when a false null hypothesis is accepted. A good test aims at reducing 'ρ', the probability of Type-II error, keeping α, the probability of type-I error at a fixed level. α is also known as size of the test or the level of significance. The test procedure comprises in finding the value of the test statistic assuming the null hypothesis to be true and comparing this value to the critical value.

We have concluded our discussion by conducting tests for population mean and population proportion under different types of alternative hypothesis.

15.12 KEY WORDS AND SYMBOLS USED

Alternative Hypothesis: A hypothesis contradicting a null hypothesis. It is denoted by H or H_1 or H_2.

Consistency: T is a consistent Estimator of θ if $E(T) \rightarrow \theta$ and $V(T) \rightarrow o$ for large n.

Critical Region or Rejection Region: The set of values of the test statistic leading to the rejection of H_0. It is a part of the sample space and is denoted by ω, if the sample point falls on ω, we reject H_0.

Efficiency: T is an efficient Estimator of θ if T has the minimum standard error among all the estimators of θ for a fixed sample size.

Interval Estimation: Estimation of a parameter θ by a pair of values, say, t_1 and t_2, $t_1 < t_2$. t_1 is known as Lower confidence Limit and t_2 as Upper Confidence Limit. The probability that $[t_1, t_2]$ contains θ is known as confidence co-efficient and denoted by $(1-\alpha)$.

95% confidence Interval to μ

$$= \left[\bar{x} - 1.96\frac{\sigma}{\sqrt{n}},\quad \bar{x} + 1.96\frac{\sigma}{\sqrt{n}}\right]$$

99% confidence Interval to μ

$$= \left[\bar{x} - 2.58\frac{\sigma}{\sqrt{n}},\quad \bar{x} + 2.58\frac{\sigma}{\sqrt{n}}\right]$$

where $\bar{x}$ = sample mean; σ = population S.D.; and n = sample size.

when σ is unknown, it can be replaced by

$$s^{|} = \sqrt{\frac{\Sigma(X_i - \bar{x})^2}{n-1}}$$ provided 'n' exceeds 30.

Level of Significance: This is the probability of type-I error and is denoted by α. Usually α is taken as .01 or 0.05 and accordingly we have 1% or 5% level of significance.

Null Hypothesis: An assumption or statement regarding the parameter or the form of a population distribution. A null hypothesis is denoted by H_0.

Point Estimation: Estimation of an unknown parameter θ by a statistic T with the help of a single value obtained from a random sample.

Power of a Test: Probability of rejecting a null hypothesis when it is false. This is given by $P(\theta) = 1-\beta\ (\theta)$ = 1-Probability of Type-II error.

Test statistic: A function of sample observations whose value, as computed from a random sample, determines the acceptance or rejection of the null hypothesis.

Type-I Error: Error committed in rejecting a true H_0.

Type-II Error: Error committed in accepting a false H_0.

Sufficiency: T is a sufficient estimator of θ if it contains all the information about θ.

Standard Error (S.E.): S.E of a statistic T is the standard deviation of T as obtained from its sampling distribution.

Unbiasedness and minimum variance: A statistic T is an unbiased estimator of θ if the expectation of T is θ. T is an MVUE (Minimum variance unbiased estimator) for θ if T has the minimum variance among all the unbiased estimators of θ.

Z-test for population mean: For testing H_0: $\mu = \mu_0$, test statistic is given by

$Z_0 = \sqrt{n}\,\frac{(\bar{x} - \mu_0)}{\sigma}$ If σ is unknown and n > 30, we replace σ by s| in the expression for Z_0.

Z-test for population proportion: For testing H_0 : P = P_0 we consider

$Z_0 = \frac{\sqrt{n}(p - P_0)}{\sqrt{P_0(1 - P_0)}}$ provided n is large,

where, p = sample proportion

Under the assumption that the null hypothesis is true, Z_0 follows standard normal distribution. At 5% level of significance, the critical region for the two-tailed test is given by

$\omega : |Z_o| \geq 1.96$

The critical region for the right-tailed test is

$\omega : Z_o \geq 1.645$

and the critical region for the left-tailed test is

$\omega : Z_o \leq -1.645$

Similarly when the level of significance is 1%, the critical region for the two-tailed test is

$\omega : |Z_o| \geq 2.58$

For the right-tailed test, the critical region is

$\omega : Z_o \geq 2.35$

and that for the left-tailed test is

$\omega : Z_o \leq -2.35$

15.13 ANSWERS TO SELF ASSESSMENT EXERCISES

A) 1. a) No, b) No. c) Yes, d) No e) Yes f) Yes, g) Yes, h) No, i) Yes j) Yes

4. $\bar{x}$ = Rs. 966.20, S| = Rs. 60.98

5. p = 0.4767, S.E. $(\hat{p})$ = 0.0166

6. $\hat{\mu} = 48.2222$; $\hat{\sigma} = 3.4564$

B) 1. a) No, b) Yes, c) Yes, d) Yes, e) No, f) No, g) Yes, h) No.

4. 1010 to 2115

5. [35.8318, 44.1682]

6. [978.02 quintals, 981.98 quintals]

C) 1. a) Yes, b) Yes, c) No, (d) No, e) Yes, f) No, g) Yes,

h) Yes.

4. Yes, $Z_0 = -0.447$

5. No, $Z_0 = 4.12$

6. Yes, $Z_0 = 1.774$

15.14 TERMINAL QUESTIONS/EXERCISES

1) Distinguish between Estimation and testing of hypothesis.

2) Explain the procedure for testing a statistical hypothesis.

3) Discuss the role of normal distribution in interval estimation and also in testing hypothesis.

4) What is an MVUE ? Examine whether a sample mean is an MVUE.

5) Discuss how far the sample proportion satisfies the desirable properties of a good estimator.

6) How do you proceed to set confidence limits to population mean ?

7) Describe how you could set confidence limits to population proportion on the basis of a large sample.

8) Explain how you would test for population mean.

9) Describe the different steps for testing the significance of population proportion.

10) 15 Life Insurance Policies in a sample of 250 taken out of 60,000 were found to be insured for less than Rs. 7500. How many policies can be reasonably expected to be insured for less than Rs. 7500 in the whole lot at 99% confidence level.

(Ans: 1278 to 5922)

11) A sample of 250 measurements of breaking strength of cotton threads provided a mean of 235 gm and a S.D of 32 gm. Find 95% confidence limits to the mean breaking strength.

(Ans: 231.033 gms, 238.967 gms)

12) A manufacturer of ball-point pens claims that a certain type of pen produced by him has a mean writing life of 550 pages with a S.D. of 35 pages. A purchaser selects 20 such pens and the mean life is found to be 539 pages. At 5% level of significance should the purchaser reject the manufacturer's claim ?

(Ans: Yes, $Z_0 = -2.30$)

13) In a sample of 550 guavas from a large consignment, 50 guavas are found to be rotten. Estimate the percentage of defective guavas and assign limits within which 95% of the rotten guavas would lie.

[Ans: (i) 9.09%; (ii) 0.0668 to 0.1150]

14) A die is thrown 59215 times out of which six appears 9500 times. Would you consider the die to be unbiased ?

(Ans: No, $Z_0 = -4.113$)

15) A sample of 50 items is taken from a normal population with mean as 5 and

standard deviation as 3. The sample mean comes out to be 4.38. Can the sample be regarded as a truly random sample?

(Ans: No, Z = –1.532)

16) A random sample of 600 apples was taken from a large consignment of 10,000 apples and 70 of them were found to be rotten. Show that the number of rotten apples in the consignment with 95% confidence may be expected to be from 910 to 1,424.

17) The mean life of 500 bulbs, as obtained in a random sample manufactured by a company, was found to be 900 hours with a standard deviation of 300 hours. Test the hypothesis that the mean life is less than 900 hours. Select $\alpha = 0.05$ and 0.01.

(Ans: Yes, $Z_0 = -3.7268$

Note: These questions/exercises will help you to understand the unit better. Try to write answers for them. But do not submit your answers to the university for assessment. These are for your practice only.

15.15 FURTHER READING

The following text books may be used for more indepth study on the topics dealt within this unit.

Levin and Rubin, 1996, *Statistics for Managers,* Printice Hall of India Pvt. Ltd., New Delhi.

Hooda, R.P., 2000, *Statistics for Business and Economics*, Macmillan India Ltd., New Delhi.

Gupta, S.P., *Statistical Methods*, 1999, Sultan Chand & Sons: New Delhi.

Gupta, C.B., and Vijay Gupta, 1998, *An Introduction to Statistical Methods*, Vikas Publishing House Pvt. Ltd., New Delhi.