Chapter 16

TESTS OF HYPOTHESIS

16.1. Introduction

In most of the situations, it is very difficult to study the whole population. The value of population parameter is usually unknown, and one objective of sampling is to estimate its value. The choice of appropriate statistics depends on which population parameter is of interest to statistician. Any inference drawn about the population is based on sample statistic. Now the question arises, whether the sample statistic is a representative value of respective population parameter or whether there is any significant difference between the parameter and statistic to some extent. This matter can be ensured by the test of hypothesis. For example, a reputed toothpaste producer claims that average weight of its big size toothpaste is 140 gm. The consumers association of Bangladesh (CAB) can verify this claim by the following steps:

i) Collecting a random sample of toothpastes
ii) Determining the average weight of toothpastes and
iii) Performing a test of hypothesis about the mean weight.

> **Hypothesis Testing** The process that enables a decision maker to draw an inference about population characteristics by analyzing the difference between the value obtained from sample and the hypothesized value of parameter is called hypothesis testing.

The tests of hypothesis involve all of the above-mentioned three steps. The second step is simply the determination of sample statistic corresponding to the parameter, which is of interest to the researchers.

5a

16.2 Concepts of Hypothesis Testing

The process of hypothesis testing can be compared with criminal jury trial. In a jury trial it is assumed that the criminal is innocent, and the jury will decide that a person is guilty only if there is very strong evidence against the presumption of innocence. This criminal jury trial process for choosing between guilt and innocence possesses

i) Rigorous procedures for presenting and evaluating evidence;
ii) A judge to enforce the rules;
iii) A decision process that assumes innocence unless there is evidence to prove guilt beyond a reasonable doubt.

It is to be noted here that this process will fail to convict a number of people who are, in fact, guilty. But if a person's innocence is rejected and the person is found guilty, we have a strong belief that the person is guilty. In testing the hypothesis, the sample collected from population acts as evidence of trial.

However, the following concepts are related to this type of decision.

16.2.1 Hypothesis and Statistical hypothesis Every moment, we use different types of statements regarding different events of our life. Statistics is not concerned with all types of statements. Statistics usually deals with statements, which have some relation with uncertainty. Thus, any statement about any aspect of a phenomenon is considered as hypothesis. Tomorrow will be a sunny day, s(he) is the president of some association, the firm is not running well, etc. are examples of the statements about tomorrow's weather, chief of an association, state of the firm, respectively. Hence these are simply hypothesis. However, in attempting to take decision regarding some characteristics of population on the basis of sample, it is necessary to make some assumption regarding parameters of the population, such assumptions which may or may not be true, are called statistical hypothesis. Hence, statistical hypothesis is a claim or statement (belief or assumption) about unknown feature (distribution or parameter) of a population. For example, average monthly sale of a store is Taka 10,000, average hourly production of a machine is 200 units, proportion of defective products produced by a certain machine is less than other and so on.

Hypothesis Any statement about any phenomenon is termed as hypothesis.

Statistical Hypothesis Statistical hypothesis is a statement about population characteristic that can be tested on the basis of sample data.

A pious person will go to the heaven is not a statistical hypothesis since it cannot be proved with the statistical data. It does not mean that it has no value. It is a faith.

In statistical tests of significance, two mutually exclusive hypotheses are to be used; these are null hypothesis and alternative hypothesis.

16.2.2 Null hypothesis The approach of statistical hypothesis testing starts with a statement complement to the original claim. The hypothesis about the parameter of a population such as mean μ, the variance σ^2, or the proportion π which is formulated for sole purpose of rejecting or nullifying it, is called null hypothesis. Hence, null hypothesis is a statement about no difference between the parameter and statistic. Null hypothesis is denoted by H_0. For example, if we want to decide whether a given coin is not fair, we formulate the null hypothesis that the coin is fair, i.e. $H_0 : \pi = 0.5$.

Null Hypothesis The hypothesis that is formulated for its possible rejection using sample data is called null hypothesis. **Null hypothesis is denoted by H_0.**

16.2.3 Alternative hypothesis The alternative hypothesis is a logical opposite statement of null hypothesis. If null hypothesis is rejected (or actually it is false), then some alternative form of parameter should be true. Thus, any hypothesis that differs from a given null hypothesis is called an alternative hypothesis. Alternative hypothesis is denoted by H_A or H_1. For example, if the null hypothesis about population mean is $H_0 : \mu_0 = 55$, an alternative might be

$$H_A : \mu_1 \neq 55, \text{ or } \mu_1 < 55 \text{ or } \mu_1 > 55.$$

Alternative Hypothesis The hypothesis, which is true if the null hypothesis is false is called alternative hypothesis. Alternative hypothesis indicates the type of test (left, right, or two-tail). It is denoted by: H_A or H_1.

Once the test has been carried out, the final conclusion is always given in terms of the null hypothesis. We either "Reject H_0 in favour of H_1" or "Do not reject H_0"; we never conclude "Reject H_1", or even "Accept H_1". However, if we conclude, "Do not reject H_0", this does not necessarily mean that the null hypothesis is true, it only suggests that there is no sufficient evidence against H_0 or in favour of H_1. Rejecting the null hypothesis then, suggests that the alternative hypothesis may be true.

Clear and precise formulation of the null and alternative hypothesis is the first and foremost step in the test of significance. Without formulation of this hypothesis one cannot proceed to perform any test of significance. Once a null hypothesis is formulated, it is required to formulate an alternative to this that is to be considered if there is not enough evidence to accept null hypothesis. Hence, role of null and alternative hypothesis in a test of significance lies in taking decision against or in favor of the postulated presumption regarding population parameter. The decision will be taken in favor of that hypothesis which will be more evident from the test for a given level of significance. Thus the null and alternative hypothesis in a test of significance enables us to take decision whether the postulated presumption is justified or not. Moreover, all hypothesis testing are done under the assumption that the null hypothesis is true. The decision in a test of significance is based on null hypothesis and the type of test (left, right, or two-tail) is based on the alternative hypothesis.

16.2.4 Simple hypothesis A hypothesis is said to be a simple hypothesis if it completely specifies the distribution of the population from which the sample has been considered. In this case, the information about all the parameters of population distribution is known. For example, if a coin is tossed 50 times (that means n of binomial distribution is 50) to determine if the coin is fair one, the null hypothesis to be formulated is $H_0 : \pi = 0.50$, which is a simple hypothesis because it specifies the population distribution completely. Again for testing $H_0 : \mu = \mu_0$ of a normal distribution if the population variance σ^2 is known, it is a simple hypothesis.

Simple Hypothesis The hypothesis, which completely specifies all the parameters of the related population, is called simple hypothesis.

16.2.5 Composite hypothesis On the other hand, if a hypothesis does not specify the population distribution completely, it is called a composite hypothesis. In the above coin tossing example, if $n = 50$ is not specified, $H_0 : \pi = 0.50$, would be a composite hypothesis, because, there is a number of distributions all with $\pi = 0.50$. In this case, n is called a nuisance parameter. Similarly, even if we know $n = 50$ but the hypothesis to be tested is defined as $H_0 : \pi \neq 0.50$ or $H_0 : \pi > 0.50$, it would be composite hypothesis, because value of π is not specified by a single value. Again for testing $H_0 : \mu = \mu_0$ of a normal distribution if the information about σ^2 is not given, it is a composite hypothesis. In this case, the parameter σ^2 is known as nuisance parameter.

Composite hypothesis The hypothesis, which does not completely specify the parameters, is called a composite hypothesis.

16.2.6 Errors in decision-making In any decision-making, if the decision is not correct, the decision maker may commit error in two mutually exclusive ways, termed as type I error and type II error.

Table 16.1 Errors in decision-making

Decision	State of Nature	
	H_0 is True	H_0 is False
Reject H_0	Type I Error	Correct decision
Fail to reject H_0	Correct decision	Type II Error

5b

Type I error The error of rejecting the null hypothesis when it is in fact true is called type 1 error. That means, type I error occurs when null hypothesis is wrongly rejected. A type I error is also known as first kind of error.

Type II error The error of accepting the null hypothesis when it is false is called type II error. Type II error occurs when null hypothesis is not rejected wrongly. A type II error is also known as second kind of error.

A type I error is often considered to be more serious, and therefore more important to avoid than a type II error. The hypothesis test procedure is therefore adjusted so that there is a guaranteed 'low' probability of rejecting the null hypothesis wrongly; this probability is never 0. While the exact probability of a type II error is generally unknown.

If we do not reject the null hypothesis, it may still be false (a type II error) as the sample may not be representative enough to identify the falseness of the null hypothesis (especially if the truth is very close to hypothesis).

For any given set of data, type I and type II errors are inversely related; the smaller the risk of one, the higher the risk of the other.

Although it is desirable to keep the both types of errors at minimum level, but unfortunately in practice it is not possible. Hence the probability of type I error is kept fixed (to be considered at the beginning of testing procedure) and then try to minimize the probability of type II error.

16.2.7 Level of significance In testing a given hypothesis, the maximum probability with which we would be willing to take risk of rejecting a hypothesis when it should be accepted, is called the level of significance of the test. This probability is denoted by α, generally specified before any sample is drawn so that the results obtained will not influence the choice of the decision maker. Since *Type I is the more serious error* (usually) that is the one we concentrate on. We usually pick to be very small such as 0.05, 0.01or in some cases 0.001. It is to be noted here that alpha (α) is not a Type I error, alpha is the *probability of committing a Type I error* and beta (β) is the *probability of committing a Type II error.*

5b

Level of significance The probability of committing a type I error is called the level of significance. In other words, it is the total area under critical region. Symbolically, $\alpha = P$ (reject $H_0 | H_0$ is true). It is also known as size of a test.

$1-\alpha = P$ (accept $H_0 | H_0$ is true) is called the confidence coefficient.

Definition Power of a test The complement of the probability of type II error is called the power of a test. That means, the probability of rejecting a false null hypothesis is the power of a test. In other words, the probability of correct decision is power of a test. Symbolically, Power of a test $1-\beta = 1-P$ (accept $H_0 | H_0$ is false) $= P'$(reject $H_0 | H_0$ is false).

- **Interpretation of level of significance** Generally a significance level of 0.05 or 0.10 is considered, although other values are also used. Thus if the level of significance is 0.05, it will mean that there are about 5 samples out of 100 that would direct to reject the hypothesis when it should be actually accepted. So $(1-0.05)=0.95$ is the probability of accepting null hypothesis when it is true, i.e. there is 95% confidence in taking the right decision. In such case it is said that the hypothesis has been rejected at 5% level of significance, which again means that the probability of wrong decision is 0.05.

5e

16.2.8 One tailed and two tailed test A one-tailed test is a test, which is concerned about possible deviation of the value of the parameter in only one direction from the specified value defined in the null hypothesis, while a two-tailed test is a test, which is concerned about the possible deviation of the parametric value in both directions. These are also called a one-sided alternative or two-sided alternative. In this case, the parameter can take any value other than the value specified by null hypothesis. For example, $H_0 : \mu = 0$, against $H_1 : \mu \neq 0$ is a two tailed test, while, $H_0 : \mu = 0$, against $H_1 : \mu > 0$, or $H_0 : \mu = 0$, against $H_1 : \mu < 0$ is a one tailed tests.

- **A left – tailed test** When the rejection region is in the left tail of the distribution of the test statistic, the test is called a left-tailed test. If the null hypothesis is $H_0 : \mu_0 = 0$, then the alternative hypothesis will be $H_1 : \mu < 0$.

- **A right – tailed test** When the rejection region is in the right tail of the distribution of the test statistic, the test is called a right-tailed test. If the null hypothesis is $H_0 : \mu_0 = 0$, then the alternative hypothesis will be $H_1 : \mu_0 > 0$.

Alternative hypotheses defined in the above two tests are called one – sided alternatives.

- **A Two- tailed test** When the rejection region is equally divided in the left and right tails of the distribution of the test statistic, the test is called a two-tailed test. The alternative hypothesis defined in this test is called two –sided alternatives. If the null hypothesis is $H_0 : \mu_0 = 0$, then the two – sided alternative hypothesis is defined by $H_1 : H_1 : \mu \neq 0$.

In the test of significance, the decision whether the use of a one-tailed or a two-tailed test is appropriate depends on the objective of the test, i.e. it is chosen on the basis of the direction of claimed deviation of the parameter.

Suppose that a manufacturer of ballpoint pen, whose machine produces on average 1000 pens per hour, is planning to purchase a new machine. The authority will not buy a new machine unless it is definitely proved as superior. For this he would test the hypothesis that the new machine is no better than the existing machine or similar to the machine now available in the market against the alternative that the new machine is superior with respect to the hourly average production of existing machine. In other word, it is required to test the null hypothesis $H_0 : \mu = 1000$ against the alternative $H_1 : \mu > 1000$ and buy the new machine only if the null hypothesis is rejected. Such an alternative test will result in a one tailed test with the critical region in the right tail.

Similarly, consider a drug, on an average five doses of which is enough in order to get cured a certain disease. A company introduces a new drug and claims that in average only two doses of this new drug will need to get cure of the same disease. In order to verify the companies claim one has to formulate $H_0 : \mu = 5$ against the alternative $H_1 : \mu < 5$. The company's claim will be true if the null hypothesis is rejected. Such an alternative test will result in a one tailed test with the critical region in the left tail.

Again, suppose one wishes to test the inequality of income of two populations. In this case, the deviation of the income of one population to other may happen in any of the two sides, i.e. income of one population may be either more or less than other population. In such case, it is required to consider a two-tailed test.

16.2.9 Test Statistic The decision about test of a hypothesis or the acceptance or rejection of a hypothesis is based on the statistical evidence from sample data. In testing procedure, it is desirable to select an appropriate statistic to be computed from sample data depending on the assumptions or available information or nature of population from which the sample has been drawn.

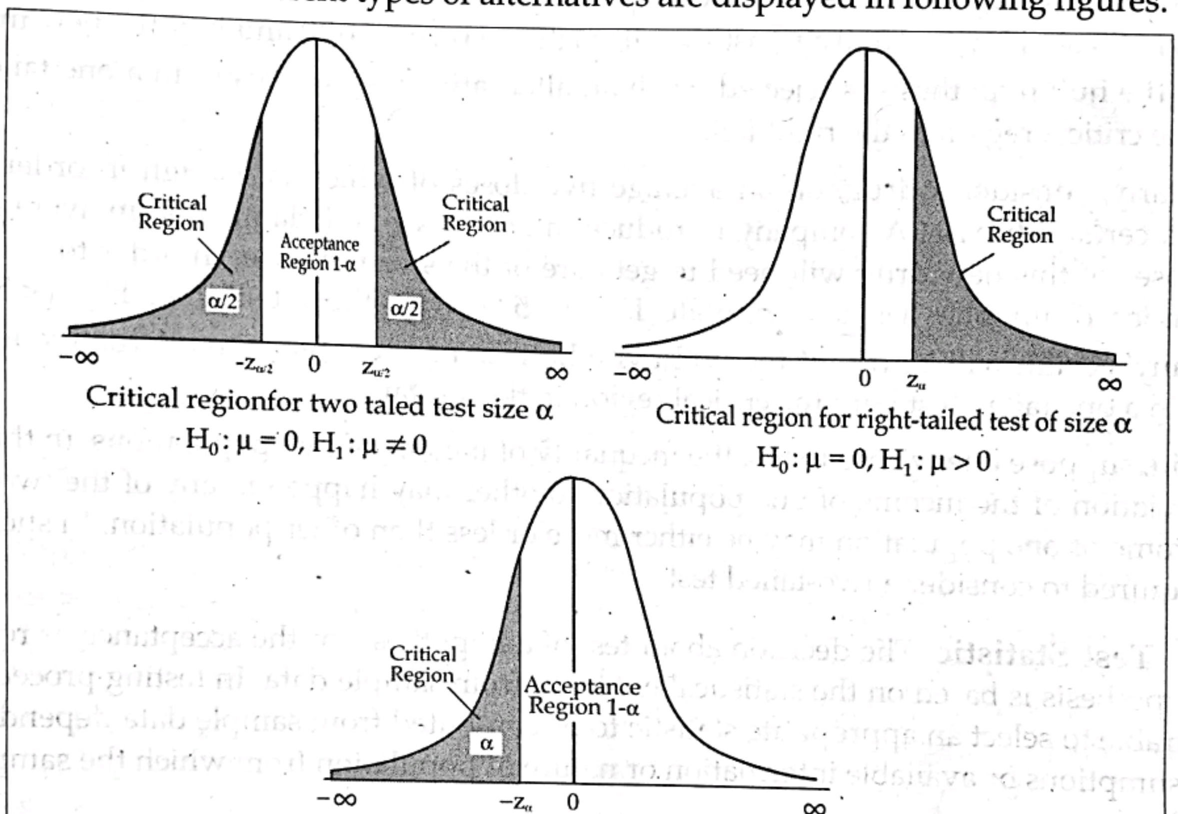
Test statistic The statistic, which is used to provide evidence about the rejection or acceptance of null hypothesis, is called test statistic. The decision about the rejection or acceptance of a null hypothesis is taken comparing the observed (calculated) and theoretical (tabulated) value of test statistic.

16.2.10 Critical Region and Acceptance region Sample space of experiment, which corresponds to the area under the sampling distribution curve of the test statistic, is divided into two mutually exclusive regions, such as acceptance region and rejection or critical region. If the value of the test statistic falls within the rejection region, the null hypothesis is rejected, and if the value of the test statistic falls within the acceptance region, we fail to reject it. In case of one-tailed test the critical region, the area of which is exactly equal to the level of significance, lies entirely on one extreme end of the curve depending on whether the test is right or left tailed, while in case of two-tailed test area of rejection is divided into two regions lie at both ends of the curve, area of each of these regions is usually exactly half of level of significance.

Critical region The set of possible values of the test statistic, which provides evidence to contradict with null hypothesis and lead to the rejection of null hypothesis is called critical region. The set of values of the test statistic that support the alternative hypothesis and lead to rejecting the null hypothesis is called the rejection region.

Acceptance region The set of values of the test statistic, which provides evidence to agree with the null hypothesis and lead to the acceptance of null hypothesis is called acceptance region.

Critical regions for different types of alternatives are displayed in following figures.



Critical regionfor two taled test size α
$H_0: \mu = 0, H_1: \mu \neq 0$

Critical region for right-tailed test of size α
$H_0: \mu = 0, H_1: \mu > 0$

Critical region for left-tailed test of size α
$H_0: \mu = 0, H_1: \mu < 0$

Fig. 16.1 *Areas of acceptance and rejection regions for one tailed and two-tailed test.*

16.2.11 Critical Value

16.2.11 Critical Value A critical value is measured in the same units of measurement as the test statistics and identifies the value of the test statistic that would lead to the rejection or acceptance of null hypothesis at the specified level of significance. This is also called the theoretical value of statistic. The critical values are determined independently of the sample statistics from the sampling distribution under null hypothesis from the table. The critical value (s) for a hypothesis test is a threshold to which the value of the test statistic in a sample is compared to determine whether or not the null hypothesis is to be rejected. The critical value for any hypothesis test depends on (i) the significance level at which the test is carried out, and (ii) the type of test (one-sided or two-sided). There are two critical values for a two-tailed test, while one for a one-tailed test. The critical values of popular test statistics for different level of significance are available in Statistical tables.

5f

Critical value The value of the sample statistic that separates acceptance region and rejection region is called critical value. This is the value of the test statistic with which observed value is compared and decision regarding acceptance or rejection of null hypothesis is taken.

Some useful critical values of Z for both one-tailed and two-tailed tests at various levels of significance (negative value of Z stands for left-tailed test and positive value for right-tailed test) are presented in following table.

Table 16.2 Critical values of Z statistic for different level of significance

Level of significance α	0.10	0.05	0.025	0.01	0.005	0.002
Critical values of Z for right-tailed test	1.28	1.645	1.96	2.33	2.58	2.88
Critical values of Z for left-tailed test	−1.28	−1.645	−1.96	−2.33	−2.58	−2.88
Critical values of Z for two-tailed test	± 1.645	±1.96	± 2.33	± 2.58	± 2.81	± 3.08

16.2.12 p-value The decision to reject or accept the null hypothesis is taken by comparing the observed value of test statistic with the critical value, which is obtained on the basis of level of significance. Since the critical value of test statistic is different at different levels of significance, so the decision about the acceptance or rejection may be different at different levels of significance. For example, suppose for a right-tailed test the value of Z-statistic observed from the sample is 2.03 and at the 0.05 level of significance the critical value of Z is 1.645, here 2.03 lies in the critical region, so the null hypothesis may be rejected at 5% level of significance. However, we cannot reject H_0 at the 0.01 level because the test statistic is less than the critical value $z = 2.33$. That means the null hypothesis may be rejected or may not be rejected at varied level of significance. To avoid this type of confusion in decision making about rejection of a null hypothesis it is desirable to consider the smallest level of significance at which it is rejected. This smallest level of significance is termed as p-value. By doing this, the actual risk of committing Type I error can be established. We see from the table of standard normal distribution that the critical value of Z at 2.12% level of significance is 2.03, which is the smallest value at which the hypothesis may be rejected, so here the p value is 0.0212.

5f

* **p-value** The p-value or observed significance level of a statistical test is the smallest value of the level of significance at which the null hypothesis can be rejected. It is the actual risk of committing a type I error, if H_0 is rejected based on the observed value of the test statistic. The p-value measures the strength of the evidence against H_0.

A small p-value indicates that the observed value of the test statistic lies far away from the hypothesized value. This presents strong evidence that H_0 is false and should be rejected. A large p-value indicates that the observed test statistic is not far from the hypothesized value and does not support rejection of H_0.

If the p-value is less than a pre-assigned significance level α, then the null hypothesis can be rejected, and we can report that the results are statistically significant at level α.

Many researchers classify p-values as follows.

i) If the p-value is less than 0.01, H_0 is rejected. The results are highly significant.

ii) If the p-value is between 0.01 and 0.05, H_0 is rejected. The results are statistically significant.

iii) If the p-value is between 0.05 and 0.10, H_0 is usually not rejected. The results are only tending toward statistically significant.

iv) If the p-value is greater than 0.10, H_0 is not rejected. The results are not statistically significant.

In particular, for testing the null hypothesis $H_0 : \mu = \mu_0$, against $H_1 : \mu > \mu_0$, p-value is given by

$$\text{p-value} = P\left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \geq z_{obs.} \,\middle|\, H_0 : \mu = \mu_0\right)$$

where, $z_{obs.}$ is the observed value of the test statistic associated with the smallest significance level at which the null hypothesis can be rejected.

It is to be noted here that all hypothesis testing are done under the assumption that the null hypothesis is true. It is also important to understand that the rejection of null hypothesis is to conclude that it is false, while fail to reject it does not necessarily mean that it is true. We fail to reject null hypothesis since we have no sufficient evidence to believe otherwise.

16.2.13 Assumption The suppositions regarding population and/or sample, which are needed to take decision about the distribution of test statistic, are known as assumption. For example, the test statistic Z follows N(0, 1) under the assumption that the sample observations are drawn independently from normal population with known variance or the sample size is large. On the other hand, the test statistic t follows Student's t distribution under the assumption that the sample observations are independently drawn from a normal population with unknown variance and the sample size is small.

16.3 Survey of Important Test Statistics

The important test statistics are

1) Z-test or normal test   2) t-test   3) χ^2-test   4) F-test

A brief survey of the important parametric tést statistics is provided below. Applications of these test statistics have been described in section 16.5.

6C

16.3.1 Z-test or Normal test In a normal tests we find U whose expected value E(U) is specified by the null hypothesis. The standard error $\sigma(U)$ of U is either known or estimated from a large sample. Then a statistic

$$Z = \frac{U - E(U)}{\sigma(U)}$$

is taken as a normal variable with mean 0 and standard deviation 1. It is symbolically expressed as $Z \sim N(0, 1)$. If the distribution of U is normal and σ(U) is known, then Z is exactly normally distributed. Frequently, however, the distribution of U is approximately normal or σ(U) is estimated from a sample or perhaps both. When samples are large this approximation is usually quite satisfactory. That is why; normal test is often regarded as a large sample test.

Large sample ($n \geq 30$)

In particular, when U is sample mean $\overline{X}$ then

$$Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1).$$

The normal curve is symmetrical about mean; hence, the critical value at a particular level of significance at right side is the same as that of left side with a negative sign. Normal tests can be used in case of one-tailed as well as two-tailed tests.

16.3.2 t-tests (Student's t-test) In case of normal tests it is assumed that population variance is either known or estimated from large sample (usually $n > 30$), but very often we have to deal with small samples where population variance is unknown. In this situation the test statistic t is to be used instead of Z-statistic. The distribution of t contains a parameter ν (nu) known as degrees of freedom (d.f.). This is a positive integer and always less than n, the size of the sample. The relationship between ν and n depends on how σ(U) calculated. The normal test can be regarded as a special case of t-test when ν is large. So, a t-test is also called a small sample test. ($n < 30$)

This statistic t is generally known as Student's t is defined in similar algebraic form of Z-statistic except that standard error is estimated from small sample. Thus the algebraic form of t is:

$$t = \frac{U - E(U)}{\text{Estimated } \sigma(U)} \quad \text{with} \quad \nu = n - 1$$

In particular, for sample mean $\overline{X}$, $t = \dfrac{\overline{X} - \mu}{s/\sqrt{n}}$

where $s^2 = \dfrac{1}{n-1}\sum(x - \overline{x})^2$ with $\nu = n - 1$.

Like the normal distribution, t distribution is also symmetrical about mean; hence, the critical value at a particular level of significance with certain degrees of freedom at right side is the same as that of left side with a negative sign. Like the normal test, the t-tests can also be one-tailed or two-tailed.

Another form of t-test is also used for testing the difference between two means in case of dependent or correlated or repeated samples (x, y). This is known as paired t-test, defined as

$$t = \frac{d - \bar{d}}{s(d)} \quad \text{with } \nu = n-1, \text{ where, } d = x - y.$$

6e

16.3.3 χ^2-tests χ^2-tests are used mainly for testing hypothesis that specify the nature of one or more distributions as a whole. Thus a hypothesis may define the mathematical form of a distribution or assert the two or more distributions are identical or two attributes are independent, etc. The elements common to the test statistics used for testing the above hypothesis is that each involves the comparison of an observed set of frequencies with a corresponding set of expected set of frequencies under null hypothesis. If O_i ($i = 1, 2, \ldots, k$) denotes the observed frequency, and E_i denotes the corresponding expected frequency, then the test statistic χ^2 is defined as

$$\chi^2_{(\nu)} = \sum \frac{(O_i - E_i)^2}{E_i}$$

where ν denotes the degrees of freedom, the only parameter of the theoretical χ^2 distribution.

χ^2 is used for testing varying types of hypotheses and the value or definition of ν varies with type of hypothesis under consideration. For example,

i) χ^2- test for specific variance here: $\nu = n-1$;
ii) χ^2- test for goodness of fit, here: $\nu = k - r$;
iii) χ^2- test for independence of attributes, here: $\nu = (r-1)(c-1)$.

Non parametric

Here n is the sample size, k is the number of cells or values of a variable, r is the number of restrictions imposed on the set of frequencies while calculating the expected frequencies, r is the number of rows and c is the number of columns. It is to be mentioned here that the first one is the parametric test while the next two are non-parametric. Tests of goodness of fit are beyond the scope of this book.

χ^2 is used for testing the hypothesis that normal population from which the sample is available has a specified variance. This is an exact and one can consider a one tailed as well as a two-tailed test. In this case χ^2 is defined as

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2} \quad \text{with } n-1 \text{ degrees of freedom.}$$

However, for last two cases, χ^2 tests are approximate in the sense that the test statistic has an approximate χ^2 distribution under null hypothesis and these are one-tailed tests.

16.3.4 F-tests. R.A. Fisher originally devised this test and Snedecor called it F-test in honour of Fisher. Suppose s_1^2 and s_2^2 denote the sample variances of the same variance σ^2 of a normal population computed from two independent samples of sizes n_1 and n_2

respectively with $v_1 = n_1 - 1$ and $v_2 = n_2 - 1$ degrees of freedom. Then the statistic F is defined as

$$F = \frac{s_1^2}{s_2^2} \text{ with } v_1 \text{ and } v_2 \text{ degrees of freedom, where, } s_1^2 > s_2^2.$$

The statistic F is so defined that s_1^2 in the numerator is expected to be larger than denominator s_2^2 when the null hypothesis is not true. Hence, only the upper tail of F-distribution is usually used as the critical region.

16.4 Steps in Hypothesis Testing

5d

As it is clear from above discussion that in order to test the validity of claim or assumption about the population parameter, at first it is necessary to draw a sample from the respective population and then analyzed. Some assumptions regarding the population distribution may also be necessary for suitability of tests. The results of the analysis are used to decide whether the claim is true or not. Thus, the general procedure for hypothesis testing consists of following basic steps.

i) **State the null hypothesis and alternative hypothesis** Clear and precise formulation of the null and alternative hypothesis is the first and foremost step in the test of significance. Without formulation of this hypothesis, one cannot proceed to perform any test of significance. That means, at first it is required to state the assumed value of population parameter which is to be tested as null hypothesis. A justified alternative hypothesis along with null hypothesis is also to be established depending on the statement of problem. Care should be taken regarding the alternative whether it would be one-tailed or two-tailed. For example, suppose we want to test the hypothesis that the monthly average price of a commodity is Taka 100 per kg. In this case, the following null and alternative hypothesis are to be considered $H_0 : \mu = 100$, against the alternative, $H_1 : \mu \neq 100$

ii) **Specify the level of significance (α) prior to sampling** At the second step of test of hypothesis, it is required to specify the level of significance. Because, the risk of taking wrong decision depends on the nature of study, so maximum risk should be determined before drawing sample from the population. It is at the discretion of investigator to select its value. Although usually $\alpha = 0.05$ is considered, but value of α may vary depending on the sensitivity of the study. For example, 5% risk might be more for taking decision about the effectiveness of a drug, in that case, 1% or 0.1% level of significance may be considered.

iii) **Select the suitable test statistic** Depending on the formulated hypothesis, assumption made about the population distribution, sample size, at this stage the appropriate test statistic is to be selected.

iv) **Establish the critical region** At this stage the critical region is to be established on the basis of above steps. That means, the critical region is selected depending on alternative hypothesis whether it is one-tailed or two-tailed, the level of significance and the selected test statistic. For example, for the two tailed alternative $H_1 : \mu \neq 100$ at 5%

level of significance, suppose normal test statistic Z is selected for testing the hypothesis, then the critical region is the lower and upper 2.5% area of a standard normal distribution which are $Z<-1.96$ and $Z>1.96$, hence the critical values are ± 1.96 (the values of z at certain level of significance can be obtained from the table 'Area under the normal curve' available in statistical tables or in the appendix of almost all books on statistics).

v) **Collect sample and Compute the value of the test statistic** After selecting the critical region, a sample of predetermined size n is collected. Then the value of selected test statistic is calculated from sample. In this case, it is assumed that the null hypothesis is true.

vi) **Compare observed and critical values** The value of the test statistic computed in earlier step is compared with the critical value or values. The computed value is checked whether it falls within or beyond the critical region.

vii) **Make the decision** The decision about the acceptance or rejection of hypothesis is taken on the basis of critical value. It is either "reject the null hypothesis" or "fail to reject the null hypothesis or not reject the null hypothesis". If the observed value of test statistic lies within the critical region, then "reject the null hypothesis", on the other hand if the observed value of test statistic falls beyond the critical region, then decision is taken as 'fail to reject null hypothesis'.

For the convenience of the users, a set of decision rules for normal test for $\alpha = 1\%$, 5% and 10% are provided below.

Table 16.3 Decision rule for one-tailed and two-tailed test using Z-statistic

Alternative Hypothesis	Decision Rule		
	$\alpha = 0.01$ Reject H_0 if	$\alpha = 0.05$ Reject H_0 if	$\alpha = 0.10$ Reject H_0 if
$\mu \neq \mu_0$	$Z > 2.58$ or $Z < -2.58$	$Z > 1.96$ or $Z < -1.96$	$Z > 1.645$ or $Z < -1.645$
$\mu > \mu_0$	$Z > 2.53$	$Z > 1.645$	$Z > 1.28$
$\mu < \mu_0$	$Z < -2.53$	$Z < -1.645$	$Z < -1.28$

viii) **Draw conclusion** This is a statement, which indicates the level of evidence (sufficient or insufficient) at given level of significance, and/or, whether the original claim is rejected or accepted. If decision is taken in favour of alternative hypothesis, then it is concluded that there is sufficient evidence to reject null hypothesis or accept the original claim.

16.5 Applications of Test Statistics

5C

In this and the following sections we will discuss some specific applications of the test statistics viz. Z, t, χ^2 in testing various types of hypotheses related to business and management.

The applications are classified as follows.

- ***Applications of Z-statistic***
  i) Test of a single population mean
  ii) Test of equality of two population means
  iii) Test of a single population proportion
  iv) Test for difference between two population proportions
  v) Test of a specified correlation co-efficient
  vi) Test of equality of two-population correlation co-efficient

mean, proportion, correlation coefficient

6c

- ***Applications of t-statistic (small sample test)***
  i) Test of a single population mean
  ii) Test of difference between two population means
  iii) Test of significance of a correlation co-efficient with zero value
  iv) Test of a population of regression co-efficient with zero or specified value
  v) Test of difference between two population regression co-efficient

- ***Applications of χ^2 -statistic***
  i) Test of a population variance with specific value
  ii) Test of equality of several variances
  iii) Test of equality of several correlation co-efficient
  iv) Test of equality of several population proportions
  v) Tests of independence of attributes
  vi) Test of goodness of fit

variance

6e

- ***Applications of F -statistic***
  i) Test of significance of difference between two population variances
  ii) Test of significance of several population means
  iii) Test of significance of two or more regression co-efficient

The above-mentioned applications of the test statistics are discussed below.

6a

16.6 Hypothesis Testing for Single Population Mean

Although it is difficult to draw a clear-cut line of demarcation between large and small samples, it is generally agreed that if the size of sample exceeds 29, then it may be regarded as a large sample. The test of significance used for large samples are different from that of small samples for the reasons that the assumptions we make in case of large samples do not hold for small samples. The following assumptions are to be made for Z-test.

- ***For small sample*** The sample is randomly selected from a normally distributed population with known variance, and

- ***For large sample*** The sample is randomly selected from a normally distributed population with unknown variance.

Some practical examples are cited below.

1 Population normal and variance known for any sample size (small and large)
Suppose $X_1, X_2, \ldots, X_n$ be a random sample of size n drawn independently from a normal population with mean μ and variance σ^2. In this case,

$$X \sim N(\mu, \sigma^2), \text{ then } \overline{X} \sim N(\mu, \sigma^2/n).$$

The following null and alternative hypothesizes may be considered for testing the population mean

i) $H_0 : \mu = \mu_0$ against $H_1 : \mu \neq \mu_0$ (for a two tailed alternative);
ii) $H_0 : \mu = \mu_0$ against $H_1 : \mu > \mu_0$ (for a right tailed alternative);
iii) $H_0 : \mu = \mu_0$ against $H_1 : \mu < \mu_0$ (for a left tailed alternative).

The test statistic for testing the null hypothesis H_0 for all the alternatives is

$$z = \frac{\overline{x} - \mu_0}{\sigma/\sqrt{n}}.$$

- **Decision rule (i)** Suppose the level of significance is α. We find $z_{\alpha/2}$ from the standard normal integral table by using $P(|Z| > z_{\alpha/2}) = \alpha/2$. We reject the null hypothesis if the absolute observed value of z is greater than $z_{\alpha/2}$.

(ii) For the second case, we find the z_α from the standard normal integral table by using $P(Z > z_\alpha) = \alpha$. We reject the null hypothesis if the observed z value is greater than z_α.

(iii) For the third case, we find the $-z_\alpha$ from the standard normal integral table by using $P(Z < -z_\alpha) = \alpha$. We reject the null hypothesis if the observed z value is less than $-z_\alpha$.

The above three cases can be shown in the following table.

Table 16.4 Decision rule for a single mean test using Z statistic

Case No.	Type of test	Decision rule Reject H_0, if
1	Two-tailed test $H_1 : \mu \neq \mu_0$	$\lvert Z \rvert > z_{\alpha/2}$
2	Right-tailed test $H_1 : \mu > \mu_0$	$Z > z_\alpha$
3	Left-tailed test $H_1 : \mu < \mu_0$	$Z < -z_\alpha$

Example 16.6.1 The managing director of a firm claims that his firm produces 110 items on average daily. A random sample of 15 days gives the following data set.

110, 118, 130, 140, 142, 146, 112, 100, 95, 98, 96, 122, 123, 124, 130.

It is known that the number of items produced by the firm follows normal distribution with variance 300.

Can we conclude at 5% level of significance that the average daily production of items of that firm is

a) 110 items b) More than 110 items c) Less then 110 items?

d) Compute p-value for each case.

Solution (a) Steps involved in testing the hypothesis will be followed in this case:

(i) First we have to formulate null and alternative hypothesis. It is a two-tailed test. Since if the average number of items produced by the firm is more or less than 110 to some extent, then the claim of the managing director will be proved as false. In that case the claim would be rejected.

So, the null hypothesis and alternative hypothesis can be formulated as follows.

Null hypothesis: $H_0 : \mu = 110$ Alternative hypothesis: $H_a : \mu \neq 110$.

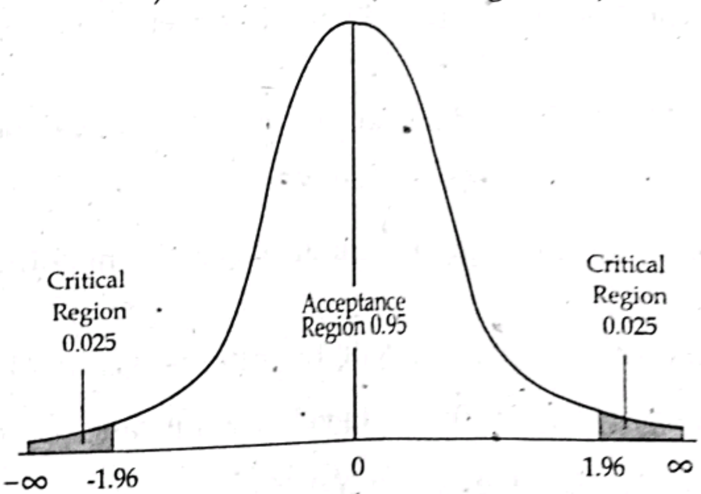
(ii) Level of significance is: $\alpha = 0.05$.

(iii) Here, the sample is taken from a normal population with known variance. The appropriate test statistic is

$$Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,\ 1).$$

where $\overline{X}$ = The sample mean
μ = Population mean (to be tested)
σ = Standard deviation of the population
$\sigma/\sqrt{n}$ = Standard error of the sample mean $\overline{X}$

(iv) Here $\alpha = 0.05$ and it is a two tailed-test, the critical region will be on both sides of curve of Z in such a way that the critical region will comprise 2.5% or 0.025 area at the right end and 2.5% at the left end. From the table of area of standard normal distribution, we see that these values of Z are ± 1.96, that means the critical regions are $Z < -1.96$ (at the left end) and $Z > 1.96$ (at the right end). Here $|z_{0.025}| = 1.96$.



(v) Under the null hypothesis, the value of Z is: $z = \dfrac{\overline{x} - \mu_0}{\sigma/\sqrt{n}}$.

(vi) Here, $\overline{x} = \dfrac{\Sigma X}{n} = \dfrac{1786}{15} = 119.07$, $\mu_0 = 110$,

$\sigma^2 = 300$, $\sigma = 17.32$ and $n = 15$. Substituting these values in the formula of Z, we have

$$z = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} = \frac{119.07-110}{17.32/\sqrt{15}} = 2.03.$$

(vii) It is found that the observed value of Z is 2.03, which is greater than the right tail critical value 1.96, hence it falls in the upper critical region.

(viii) Since the observed value of the test statistic falls in the critical region, so we fail to accept the null hypothesis at 5% level of significance.

- **Conclusion** Hence we cannot accept the claim of the managing director at 5% level of significance.

- **p-value** From the table of the standard normal distribution we find that $P(Z>2.03)=0.0212$ since it is a two tailed test, the p-value is $0.0212\times 2=0.0424$ that means the smallest level of significance at which the hypothesis may be rejected is approximately 4. 34%.

(b) Null hypothesis is the same as (a) Null hypothesis $H_0 : \mu = 110$

Alternative hypothesis $H_a : \mu > 110$ Level of significance is: $\alpha = 0.05$.

The apropriate test statistic is the same as (a). That is

$$Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,\ 1).$$

Here $\alpha = 0.05$ and it is one-sided right tailed-test, the critical value $z_{0.05}$ will be found in such a way that $P[Z > z_{0.05}] = 0.05$. It is found from the standard normal distribution that $z_{0.05} = 1.645$.

The calculated value of Z under the null hypothesis is 2.03 which is greater than 1.645. That is observed value of Z lies in the rejection region. Hence we have no reason to accept the null hypothesis.

- **p-value** From the table of the standard normal distribution we find that

$$p = P(Z > 2.03) = 0.0212$$

That means the smallest level of significance at which the hypothesis may be rejected is approximately 2.12%.

(c) Null hypothesis is the same as (a) Null hypothesis $H_0 : \mu = 110$

Alternative hypothesis $H_a : \mu < 110$ Level of significance is: $\alpha = 0.05$.

The appropriate test statistic is the same as (a). That is

$$Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,\ 1)$$

Here $\alpha = 0.05$ and it is one-sided left tailed-test, the critical value $z_{0.05}$ will be found in such a way that $P[Z < -z_{0.05}] = 0.05$. It is found from the standard normal distribution that for left tail, $z_{0.05} = -1.645$.

The calculated value of Z under the null hypothesis is 2.03 which is greater than 1.645. That is observed value of Z lies in the acceptance region. Hence we have no reason to reject the null hypothesis at 5% level of significance under the alternative hypothesis that the average production of the firm is less than 110.

- **p-value** From the table of the standard normal distribution we find that

$$p = P(Z > 2.03) = 0.0212$$

This is a left tailed test, so the p-value is

$$p = P(Z > -2.03) = 1 - 0.0212 = 0.9788$$

That means the smallest level of significance at which the hypothesis may be rejected is approximately 97.88%.

Note In practice, if we are in a position to reject a null hypothesis, we compute p-value to find the exact level of significance. The p-value found in this case is quite unjustified.

2 Population normal, variance unknown and sample size is large (n>29)

5/n

Example 16.6.2 Manager of a fertilizer factory claims that the average daily production of his factory follows normal distribution with mean production 880 kg. A random sample of 50 days shows that average production is 871 kg with standard deviation 21 kg. Test the significance of the claim of the manager at 5% level of significance. Also find p-value.

Solution Here null and alternative hypotheses are

$$H_0: \mu = 880 \text{ and } H_1: \mu \neq 880$$

Here population is normal but variance is unknown and the sample size is large. The sample standard deviation can be taken as a good estimate of the population standard to estimate the standard error of the sample mean $\overline{X}$. That is $\frac{\sigma}{\sqrt{n}}$ can be replaced by $\frac{s}{\sqrt{n}}$ for defining Z. The appropriate test statistic is $Z = \frac{\overline{X} - \mu}{s/\sqrt{n}}$, which is approximately a standard normal variable. The value of the test statistic Z under the null hypothesis is

$$z = \frac{\overline{x} - \mu_0}{s/\sqrt{n}}$$

Here s^2 is the estimate of population variance σ^2, defined as

$$s^2 = \frac{1}{n-1}\Sigma(x - \overline{x})^2.$$

It is a two-tailed test, so the critical region at 5% level of significance is

$$|Z|>1.96.$$

We have, $\bar{x}=871$, $s=21$, $n=50$, so the computed value of Z under null hypothesis is

$$z=\frac{\bar{x}-\mu}{s/\sqrt{n}}=\frac{871-880}{21/50}=-3.03.$$

- **Decision** Since the observed value of Z lies in the critical region, so we fail to accept null hypothesis. That means, the manager's claim is not justified.
- **p-value** From the standard normal integral table, we find that $P(Z>3.03)=0.0005$ and $P(Z<-3.03)=0.0005$ so the value of p is $0.0005\times 2=0.001$ Since the p-value is far less than 0.01, the value of Z is highly significant.

5h
M
3 Population is not normal, variance known and sample size is large

Example 16.6.3 The producer of a company claims that the selling price of his product is very standard and it is TK. 1500 per unit with standard deviation TK. 45. There is some doubt of CAB (Consumers' Association of Bangladesh) regarding this price. They want to verify this price using statistical testing procedure. A random sample of the sailing prices of 100 products of this company from different areas were collected. The average price per unit was found Tk. 1510.

Can the CAB conclude at 5% level of significance that the average price of the product is standard? Also calculate p-value and 95% confidence interval for population mean.

Solution We have to test the hypothesis that the sailing price of the product is TK. 1500. So, the null hypothesis and alternative hypotheses are

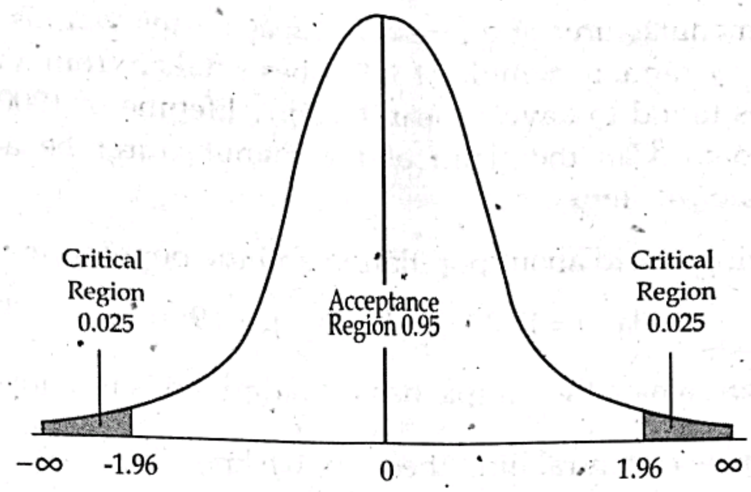
Null hypothesis $H_0:\mu=1500$ Alternative hypothesis $H_a:\mu\neq 1500$

Level of significance: $\alpha=0.05$.

Here, we have to test the significance of a population mean with known population variance. But nothing is said about the form of distribution, but mean and variance exists. Since the sample size is large $(n=100)$, according the central limit theorem, the sampling distribution of the mean is approximately normally distributed with mean μ and standard error $\frac{\sigma}{\sqrt{n}}$, so the appropriate test statistic is:

$$Z=\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}\sim N(0,\ 1).$$

Here $\alpha=0.05$ (given), and it is a two tailed-test, the critical region will be on both sides of curve of z in such a way that the critical region will comprise 2.5% or 0.025 area at the right end and 2.5% at the left end. From the table of area of normal curve, we see that these values of Z are ± 1.96, that means the critical regions are $Z<-1.96$ (at the left end) and $Z>1.96$ (at the right end).



Under the null hypothesis Z is given by : $z = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$.

Here, $\bar{x} = 1510$, $\mu_0 = 1500$, $\sigma = 45$ and $n = 100$, then;

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{1510 - 1500}{45/\sqrt{100}} = 2.22.$$

It is found that the observed value of Z is 2.22. It is greater than the critical value 1.96; hence it falls in the upper critical region. So we fail to accept the null hypothesis at 5% level of significance.

- **Conclusion** The claim of the producer is not right, that means, the price of the products as claimed by the producer is not standard.

- **p-value** From the table of area under the standard normal distribution we find that the value of $P(Z > 2.22) = 0.0132$ since it is a two tailed test, the p-value is 0.0132 × 2 = 0.0264, that means the smallest level of significance at which the hypothesis is rejected is 2.64%.

- **Confidence interval** 95% confidence interval for μ is given by

$$\bar{x} \pm z_{\alpha/2} \frac{\sigma}{\sqrt{n}}, \quad \text{here } z_{\alpha/2} = z_{0.025} = 1.96,$$

So, the required CI is: $1510 \pm 1.96 \frac{45}{\sqrt{100}} = (1501, 1519)$.

Thus, we may be 95% confident that the true mean price will be between TK. 1501 and TK. 1519.

4 Population is not normal, variance is unknown and the sample size is large

Si

Example 16.6.4 A manufacturer of fluorescent tubes claims that his tubes have a lifetime of 1950 burning hours. A random sample of 100 tubes is taken from a day's output and tested for burning life. It is found to have a mean burning lifetime of 1900 hours with a standard deviation of 150 hours. Can the claim of the manufacturer be accepted at 5% level of significance? Also find p-value.

Solution Here nothing is said about population and the population variance is not known.

$$H_0 : \mu = 1900 \quad \text{and} \quad H_1 : \mu \neq 1900$$

The population from which the sample has been derived is not normal, but the sample size is large. So by the virtue of central limit theorem, we know that: $Z = \dfrac{\overline{X} - \mu}{s/\sqrt{n}} \sim N(0,\ 1)$.

Thus the appropriate test statistic is: $Z = \dfrac{\overline{X} - \mu}{s/\sqrt{n}} \sim N(0,\ 1)$.

Here s^2 is the estimate of population variance σ^2 and is defined as

$$s^2 = \frac{1}{n-1}\Sigma(x - \overline{x})^2.$$

It is a two-tailed test, so the critical region at 5% level of significance is

$$|Z| > 1.96.$$

Given, $\overline{x} = 1900$, $s = 150$, $n = 100$, so the computed value of Z under null hypothesis is.

$$z = \frac{\overline{x} - \mu_0}{s/\sqrt{n}} = \frac{1900 - 1950}{150/100} = -3.33.$$

- **Decision.** Since the observed value of Z lies in the critical region, so we fail to accept null hypothesis. That means, the manufacturer's claim is not accepted at 5% level of significance.
- **p-value** From the standard normal integral table, we find that $P(Z < -3.33) = 0.0004$ so the value of p is $0.0004 \times 2 = 0.0008$ (approx). Since the p-value is far less than 0.01, the value of z is highly significant.

CD

5 Population normal, variance unknown and sample size is small $(n < 30)$

Suppose $X_1, X_2, \ldots, X_n$ be a random sample of size n drawn independently from a normal population with mean μ and unknown variance σ^2.

The following null and alternative hypothesizes may be considered.

i) $H_0 : \mu = \mu_0$ against $H_1 : \mu \neq \mu_0$ (for a two tailed alternative)

ii) $H_0 : \mu = \mu_0$ against $H_1 : \mu > \mu_0$ (for a right tailed alternative)

iii) $H_0 : \mu = \mu_0$ against $H_1 : \mu < \mu_0$ (for a left tailed alternative)

The test statistic for testing the null hypothesis H_0 for all the alternatives is

$$t = \frac{\bar{x}-\mu_0}{s/\sqrt{n}}, \quad \text{where} \quad s^2 = \frac{1}{n-1}\Sigma(x-\bar{x})^2.$$

Here t follows Student's t-distribution with $(n-1)$ degrees of freedom.

The decision rules at 100α percent level of significance for hypothesis testing using t-statistic are as follows:

Table 16.5 Decision rule for t-test

Case No.	Type of test	Decision rule Reject H_0, if
1	Two-tailed test: $H_1: \mu \neq \mu_0$	$\lvert t \rvert > t_{\alpha/2;(n-1)}$
2	Right-tailed test: $H_1: \mu > \mu_0$	$t > t_{\alpha;(n-1)}$
3	Left-tailed test: $H_1: \mu < \mu_0$	$t < -t_{\alpha;(n-1)}$

5i

Example 16.6.5 A wholesaler knows that on average sales in its store is 20% higher in December than in November. For the current year, a random sample of six stores was taken. Their percentage of sales increased in December was found to be 19.2, 18.4, 19.8, 20.2, 20.4, 19.0. Assuming that the sample has been drawn from a normal population with mean μ and unknown variance σ^2.

a) test the null hypothesis at 10% level of significance whether the true mean percentage sales increase is 20%, against the two sided alternative.

b) do you think that the true mean percentage sales increase is more than 20% at 10% level of significance.

c) do you think that the true mean percentage sales increase is less than 20% at 10% level of significance.

Solution Here the population variance is unknown and the sample size is small. The estimated standard error of sample mean $\bar{x}$ is given by

$$\hat{se}(\bar{x}) = s/\sqrt{n} \quad \text{where,} \quad s^2 = \frac{\Sigma(x-\bar{x})^2}{n-1}.$$

(a) We want to test the null hypothesis

$H_0: \mu = \mu_0 = 20$ against the alternative $H_1: \mu = \mu_1 \neq 20.$

Since the sample size is small and population variance is unknown, the value of the test statistic under null hypothesis is

$$t = \frac{\bar{x}-\mu_0}{s/\sqrt{n}}$$

which is distributed as Student's t with $(n-1)$ degrees of freedom.

It is a two tailed test, so the decision rule is

Reject H_0 in favor of H_1 if $t = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}} > t_{\alpha/2;(n-1)}$ or, $t = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}} < -t_{\alpha/2;(n-1)}$

The critical values of t at 10% level of significance with $n-1=5$ df are $\pm t_{n-1;\alpha/2} = \pm t_{5;05} = \pm 2.015$ (From table of t-distribution).

Here, $n=6$, $\bar{x}=19.5$, $s^2=0.588$, and $s/\sqrt{n} = 0.24$.

Then, we have, $t = \dfrac{19.5-20}{0.24} = -1.08$.

Since the observed value $t=-1.08$ lies between -2.015 and 2.015, hence we fail to reject the null hypothesis at 10% level of significance.

That means, the true mean sales is 20% higher in December than in November.

(b) In this case, we have to perform a one-tailed test, given by

$$H_0 : \mu = \mu_0 = 20 \text{ against the alternative } H_1 : \mu = \mu_1 > 20.$$

The decision rule is to reject H_0 in favor of H_1 if $t = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}} > t_{\alpha;n-1}$.

The computed value of t is the same as before i,e, $t=-1.08$, the critical value of t at 10% level of significance is: $t_{0.10;5} = 1.476$

Since the observed value of $t=-1.08$ which is less than the critical value, we fail to reject the null hypothesis at 10% level of significance, which means the average sales increased by more than 20 percent is not evident from the given data.

(c) In this case, we have to perform a one-tailed test, given by

$$H_0 : \mu = \mu_0 = 20 \text{ against the alternative } H_1 : \mu = \mu_1 < 20.$$

The decision rule is to reject H_0 in favor of H_1 if $t = \dfrac{\bar{x} - \mu_0}{s/\sqrt{n}} < -t_{\alpha;n-1}$.

The computed value of t is the same as before i.e. $t=-1.08$ the critical value of t at 10% level of significance is: $t_{0.10;5} = -1.476$

Since the observed value of $t=-1.08$ lies beyond the critical region, we fail to reject the null hypothesis at 10% level of significance, which means the average sales increased by less than 20 percent is not evident from the given data.

51 **Example 16.6.6 (Large sample with unknown variance when parent population is not normal)** A fertilizer factory manager claims that its average daily production is 912 kg. A random sample of 50 days shows that average production is 903 kg with standard deviation 21 kg. Test the significance of the claim of the manager at 5% level of significance.

Solution Here, $H_0 : \mu = 912$ and $H_1 : \mu \neq 91.2$.

The population from which the sample has been derived is not normal, but the sample size is large. So by the virtue of central limit theorem, we know that $\frac{\overline{X}-\mu}{s/\sqrt{n}} \sim N(0,\ 1)$.

Thus under the null hypothesis, the appropriate test statistic is

$$z = \frac{\overline{X}-\mu_0}{s/\sqrt{n}}$$

where s^2 is the estimate of population variance σ^2, defined as

$$s^2 = \frac{\Sigma(x-\overline{x})^2}{n-1}.$$

It is a two tailed test, so the critical region at 5% level of significance is

$$|Z| > 1.96.$$

Given, $\overline{x}=871,\ s=21,\ n=50$, so the computed value of Z under null hypothesis is

$$z = \frac{\overline{x}-\mu}{s/\sqrt{n}} = \frac{903-912}{21/\sqrt{50}} = -3.03.$$

- **Decision** Since the observed value of Z lies in the critical region, so we fail to accept null hypothesis. That means, the manager's claim is not justified.

Example 16.6.7 (Two-tailed test for small sample with known variance) The yields of wheat from a random sample of six test plots are as 1.40, 1.80, 1.30, 1.90, 1.60 and 2.20 tons per acre, test whether the information supports the claim that the average yield for this kind of wheat is 1.5 tons/acre with standard deviation 0.43 tons/acre. Also find the p-value.

Solution The null and alternative hypotheses for this test are

Null hypothesis $H_0: \mu = 1.5$ Alternative hypothesis $H_A: \mu \neq 1.5$

Here, the sample size is small, the population variance is known and the sample is taken from normal population, so the appropriate test statistic is

$$Z = \frac{\overline{X}-\mu}{\sigma/\sqrt{n}}.$$

Let the level of significance $\alpha = 0.05$ (if α is not given, it is considered as 0.05).

Thus the critical values of Z are ± 1.96 that means the critical regions are

$$Z < -1.96 \quad \text{and} \quad Z > 1.96.$$

The average yield as calculated for the observations is 1.7 tons/acre, so the value of Z under null hypothesis is

$$z = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{1.7 - 1.5}{\frac{0.43}{\sqrt{6}}} = 1.14.$$

Since the observed value of Z is less than 1.96, that means the observed value lies in the acceptance region, so we fail to reject null hypothesis. Hence, the information supports that the average yield is 1.5 tons/acre.

- **p-value** From the table of standard normal distribution, we find that $P(Z>1.14)=0.1271$ and $P(Z<-1.14)=0.1271$ so, the p-value is $0.1271\times2=0.2542$, that means the test will be significance at 25.42% level of significance.

Example 16.6.8 (Right tailed test for small sample size with known variance) A stenographer claims that she can take dictation at the rate of more than 100 words per minute with a standard deviation of 15 words. Can we reject the claim on the basis of 10 trials in which she demonstrates a mean of 102 words per minute? Use 5% level of significance.

Solution If the stenographer can take dictation of even at the rate of 100 words per minute, her claim can not be accepted. So the null hypothesis and alternative hypothesis to be considered are

Null hypothesis $H_0: \mu = 100$ Alternative hypothesis $H_1: \mu > 100$

Since the population variance is known, the appropriate test statistic is

$$Z = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}}.$$

The critical region at 5% level of significance is: $Z>1.645$.

Here, $\bar{x}=102$, $\sigma=15$, $n=10$, so the value of Z is

$$Z = \frac{102-100}{15/\sqrt{10}} = 0.42.$$

The computed value of Z does not exceed the critical value 1.645, so we fail to reject null hypothesis. So stenographer's claim is not correct.

6n

Example 16.6.9 (Left tailed test for small sample size with known variance) An automobile manufacturer company claims that a new model car achieves on an average 31.5 miles per gallon in highway driving. The distribution is known to be normal with standard deviation 2.4 miles per gallon. A random sample of sixteen automobiles provided an average of 30.6 miles per gallon in highway trials. Test the claim of company at the 5% level of significance against the population mean is less than 31.5 miles per gallon.

Solution It is a left-tailed test, because, if the average coverage of distance is even equal to 31.5 miles, the company's claim will not be correct. Thus, the null and alternative hypotheses are Null hypothesis $H_0 : \mu = 31.5$ Alternative hypothesis $H_A: \mu < 31.5$

Since, the population variance is known, the appropriate test statistic is

$$Z = \frac{\overline{X}-\mu}{\sigma/\sqrt{n}}.$$

The critical region at 5% level of significance is: $Z<-1.645$.

Here, $\bar{x}=30.6,\ \sigma=2.4, n=16$, so the value of Z is

$$Z=\frac{30.6-31.5}{2.4/\sqrt{16}}=-1.50.$$

6n The computed value of Z does not fall in the critical region, so we fail to the reject the null hypothesis that means, the average achievement of car is not less than 31.5 miles per gallon.

Example 16.6.10 **(Two-tailed test for large sample with known variance)** A large manufacturer of stereo components is concerned about the efficiency of many new employees hired during the last six months. The efficiency rating of all employees has been reasonably stable with mean rating 200 and a standard deviation of 20. A random sample of 75 new employees has been selected and their average efficiency rating is found as 197.5. Test the null hypothesis at 1% level of significance that the mean efficiency rating is still 200. Find 99% confidence interval for population mean.

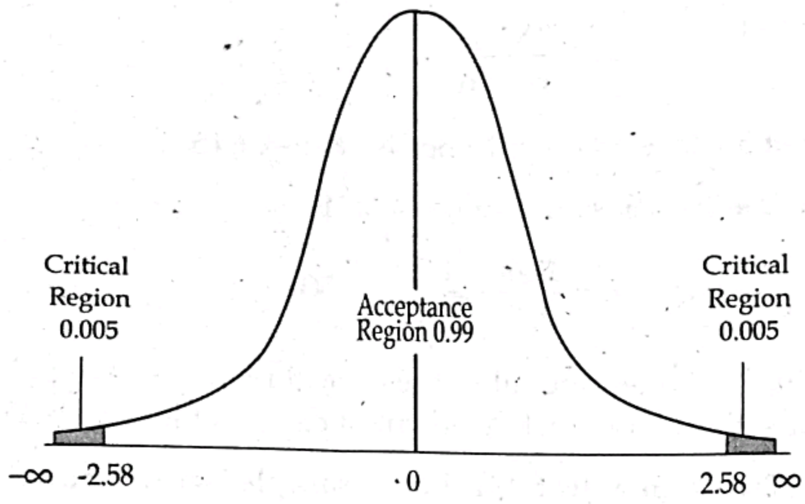
Solution It is a two-tailed test because if the mean efficiency rating obtained from sample is too high or too low in comparison with population mean rating, the null hypothesis would be rejected, so the null and alternative hypotheses to be considered are

Null hypothesis $H_0: \mu = 200$ Alternative hypothesis $H_1: \mu \neq 200$

We have to test the significance of a population mean with known population variance $\sigma^2=20^2$, and the sample size is also large, so the sampling distribution of the mean is normally distributed with standard error $\frac{\sigma}{\sqrt{n}}$, the appropriate test statistic for the selected hypothesis is

$$Z = \frac{\overline{X}-\mu}{\sigma/\sqrt{n}}.$$

Here the level of significance is $\alpha=0.01$ (as given in the problem).



It is a two- tailed test, the critical region will be on both ends of curve of Z that will comprise 0.5%or 0.005 area at the right end and 0.005 area at the left end. From the table of normal curve, we see that this critical values of Z are ± 2.58, that means the critical regions are $Z < -2.58$ and $Z > 2.58$.

Now, we have to calculate: $z = \dfrac{\overline{x} - \mu_0}{\sigma / \sqrt{n}}$.

Here, $\overline{x} = 197.5$, $\mu_0 = 200$, $\sigma = 20$ and $n = 75$, we have

$$z = \frac{\overline{x} - \mu_0}{s/\sqrt{n}} = \frac{197.5 - 200}{20/\sqrt{75}} = -1.08.$$

The observed value of Z is greater than the lower critical value –2.58. Since the observed value of test statistic does not fall within the critical region, so we fail to reject the null hypothesis at 1% level of significance. That means the mean efficiency rating is still 200.

- **Confidence interval** 99% confidence limits are given by

$$\overline{x} \pm z_{0.005} \frac{\sigma}{\sqrt{n}} = 197.5 \pm 2.58 \frac{20}{\sqrt{75}} = (191, 203),$$

that means, we are 99% confident that true average efficiency rating of all employees will lie between 191 and 203.

6h

Example 16.6.11 Suppose in Example 16.6.3, instead of checking if the price is different from the standard price, CAB decided to verify whether the specified price of the product is more than the standard price.

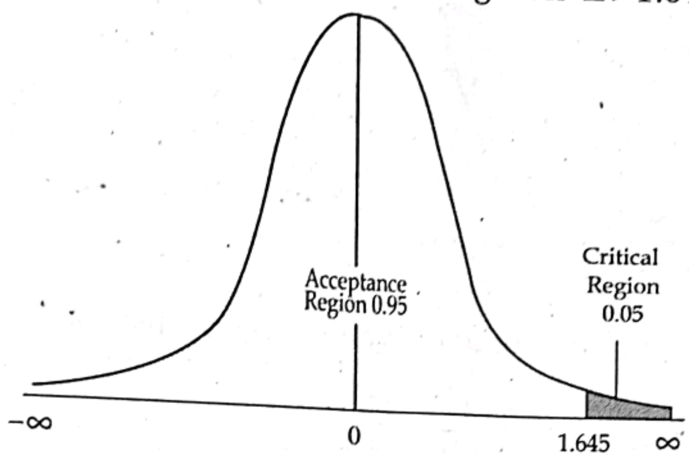
Solution It is obviously a one-tailed test, because, if average price obtained from the sample is less than or equal to the standard price, the claim of the producer will be established. Hence we have to consider a composite hypothesis defined as

Null hypothesis $H_0 : \mu \leq 1500$ Alternative hypothesis $H_1 : \mu > 1500$

The test statistic for testing the null hypothesis is: $Z = \dfrac{\bar{x}-\mu}{\sigma/\sqrt{n}}$.

Here $\alpha = 0.05$ (given).

It is a right- tailed test; the critical region will be on right side of curve of Z that will comprise 5% or 0.05 area at the right end. From the table of area of normal curve, we see that this critical value of Z is 1.645 that means the critical region is $Z > 1.645$ (at the right end).



We have found $z = 2.22$ (from Example 16.6.3).

Hence, the observed value of $Z = 2.22$ which is greater than the critical value 1.645, hence, the observed value of test statistic falls in the critical region, we fail to accept the null hypothesis at 5% level of significance.

The claim of the producer is not right, that means, the average price of the products of the producer is more than the standard price.

- **p-value** The p-value is given by $P(Z > 2.22) = 0.0132$ that means the smallest level of significance at which the hypothesis is rejected is 0.0132 or 1.3%.

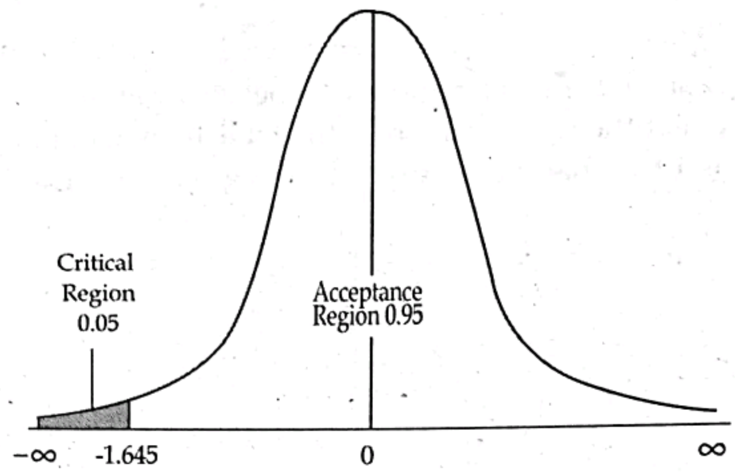
Example 16.6.12 The average petrol consumption of existing auto engines is 10.5 km. per liter. An auto company decided to introduce a new six cylinder car whose mean petrol consumption is claimed to be lower than that of the existing auto engine. In order to verify company's claim, a sample of 50 new cars was randomly selected and it was found that the mean petrol consumption was 10 km. per liter with a standard deviation of 3.5 km. per liter. Test whether the claim of the company is acceptable at 5% level of significance.

Solution Here we have to consider the following hypotheses.

$$H_0 : \mu = 10.5 \text{ and}$$
$$H_1 : \mu < 10.5.$$

Although the population variance is not known, since sample size is large, the appropriate test statistic is $Z = \dfrac{\bar{X}-\mu}{s/\sqrt{n}}$ where s^2 is sample variance and n is the sample size.

Since it is a left tailed test, the critical region lies in the left end of the curve given by $Z<-1.645$.



Given, $\overline{x}=10$ km., $s=3.5$, $n=50$, so the value of test statistic Z under H_0 is given by

$$z = \frac{\overline{x}-\mu_0}{s/\sqrt{n}} = \frac{10-10.5}{3.5/\sqrt{50}} = -1.01.$$

This observed value of Z does not lie in the critical region, so we fail to reject the null hypothesis. That means the claim of the company is not acceptable.

6i **Example 16.6.13** Suppose the daily number of items produced by a firm for randomly selected 15 days is as follows

110, 118, 130, 140, 142, 146, 112, 100, 95, 98, 96, 122, 123, 124, 130.

Can we conclude at 5% level of significance that the average daily production of items of that firm is 110?

Solution It is a two-tailed test because if the average hourly number of items produced by the company is more or less than 110, then the statement that the average daily production of items will be proved as false, then the null hypothesis that the average daily production of items is 110 would be rejected.

So, we the null hypothesis and alternative hypothesis are as follows

Null hypothesis $H_0: \mu = 110$ Alternative hypothesis $H_a: \mu \neq 110$

Level of significance as given in the problem is $\alpha=0.05$, since the variance is unknown and sample size is small, the appropriate test statistic is

$$t = \frac{\overline{X}-\mu_0}{s/\sqrt{n}}$$ which is distributed as Student's t with $n-1$ df.

Here, $n=11$, so the $df=15-1=14$.

Here $\alpha=0.05$, and it a two tailed-test, the critical region will be on both sides of t-distribution, in such a way that the critical region will comprise 2.5% or 0.025 area at the right end and 2.5% at the left end. From the table of t-distribution, we find for $df=14$, $\alpha=0.025$, the values of t are ± 2.145, that means the critical regions are $t<-2.145$ (at the left end) and $t>2.145$ (at the right end).

Here, $\bar{x}=119.07$, $\mu=\mu_0=110$.

$$s^2=\frac{\Sigma(X-\bar{X})^2}{n-1}=\frac{1}{n-1}\left[\Sigma X^2-\frac{(\Sigma X)^2}{n}\right]=\frac{1}{14}\left[216682-\frac{(1786)^2}{15}\right]=\frac{1}{14}\left[216682-\frac{(1786)^2}{15}\right]$$

$=287.78$, so, $s=16.96$ and $n=15$.

So, we have, $t=\dfrac{\bar{x}-\mu_0}{s/\sqrt{n}}=\dfrac{119.07-110}{16.96/\sqrt{15}}=2.07.$

It is found that the observed value 2.07 is less than the critical value 2.145; hence it does not fall in the critical region, so we fail to reject the null hypothesis at 5% level of significance.

- **Conclusion** We can conclude that the average daily production of the items of the given firm can be accepted as 110.

6) **Example 16.6.14 (Small sample with unknown variance)** A gas station repair shop claims that the average time it takes to do a lubrication job and oil change is maximum 30 minutes. The consumer protection department wants to test the claim. A sample of six cars was sent to the station for oil change and lubrication. The job took an average of 34 minutes with a standard deviation of 4 minutes. Assuming that the population is normal, do you think that the job took an average of time more than 30 minutes? (use $\alpha=0.05$)

Solution It is obviously a one-tailed test, because, the claim will not be rejected if the average time taken on a car for the oil change and lubrication job is considerably less than 30 minutes. So, in this case we have to consider a composite hypothesis given by (although in practice we use simple hypothesis)

$$H_0: \mu=\mu_0 \le 30 \text{ against the alternative } H_1: \mu=\mu_1 > 30.$$

Since the sample size is small and population variance is unknown, t statistic will be used.

Under the null hypothesis the value of t is

$$t=\frac{\bar{x}-\mu_0}{s/\sqrt{n}} \text{ which is distributed as Student's } t \text{ with } n-1 \text{ df.}$$

Here, $n=6$, $\bar{x}=34$, $\mu_0=30$, $s=4$ and $\hat{\sigma}(\bar{x})=s/\sqrt{n}=1.63$.

Therefore, we get: $t=\dfrac{\bar{x}-\mu_0}{s/\sqrt{n}}=\dfrac{34-30}{1.63}=2.45$

The critical value of t at $\alpha=0.05$ and $df=n-1=5$, for a right tailed test is given by $t_{5;0.05}=2.02$, since the computed value of $t=2.45$ is higher than the critical value of $t=2.02$, we reject the null hypothesis, that means the claim of the shop is not considered to be correct.

- **p-value** In this case p-value is given by $P(t_5 > 2.45) = 0.03$ (from the table of critical value of t distribution), hence p-value is 0.03 or 3%.

Example 16.6.15 A process that produces bottles of shampoo, when operating correctly, produces bottles whose contents weigh, on average, 20 ounces. A random sample of nine bottles from a single production run yielded the following weights

21.4, 19.7, 19.7, 20.6, 20.8, 20.1, 19.7, 20.3, 20.9.

Assuming that the population distribution is normal, test the hypothesis that the process is operating correctly at 5% level of significance. Also calculate 95% confidence interval for population mean.

Solution It is a two-tailed test, because, the claim will be rejected if the average weight deviates from 20 ounces in any direction. So, the hypothesis are given by

$$H_0 : \mu = \mu_0 = 20 \text{ against the alternative } H_1 : \mu \neq 20$$

Since the sample size is small and population variance is unknown, t statistic will be used. Under the null hypothesis the value of t is

$$t = \frac{\overline{x} - \mu_0}{s/\sqrt{n}} \text{ which is distributed as Student's t with } n-1 \text{ df.}$$

Here, $n = 9$, $\overline{x} = 20.36$, $\mu_0 = 20$, $s = 0.61$ and $\hat{\sigma}(\overline{x}) = s/\sqrt{n} = 0.203$

Hence, $t = \dfrac{\overline{x} - \mu_0}{s/\sqrt{n}} = \dfrac{20.36 - 20}{0.203} = 1.77.$

The critical values of t at $\alpha = 0.05$ with $n-1 = 8$ df for a two- tailed test are given by

$$\pm t_{8;0.025} = \pm 2.316.$$

Since the computed value of t is 1.77 which does not fall in the critical region, so we fail to reject the null hypothesis, that means the process is operating correctly.

95% confidence interval for population mean μ is given by

$$\overline{x} \pm t_{n-1;\alpha/2} \frac{s}{\sqrt{n}} = 20.36 \pm 2.316 \times \frac{0.61}{\sqrt{9}}.$$

So, the lower and upper confidence limits are 19.89 and 20.83 respectively.

16.7 Test of Hypothesis Concerning Two Population Means

Suppose $X_{11}, X_{12}, \ldots\ldots, X_{1n_1}$ be a random sample of size n_1 drawn from normal population with mean μ_1 and variance σ_1^2, and $X_{21}, X_{22}, \ldots\ldots, X_{2n_2}$ be another sample of size n_2 drawn from normal population with mean μ_2 and variance σ_2^2. Suppose, the observed sample means are $\overline{X}_1$ and $\overline{X}_2$. In the earlier chapter, it is mentioned that the distribution of the difference between two sample means follows normal distribution when variances are known for all possible sample sizes, that means

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,\ 1).$$

The possible null and alternative hypotheses considered for testing the significance of difference between two population means are

i) $H_0: \mu_1 = \mu_2$ against $H_a: \mu_1 \neq \mu_2$ (for a two tailed alternative)

ii) $H_0: \mu_1 = \mu_2$ against $H_a: \mu_1 > \mu_2$ (for a right tailed alternative)

iii) $H_0: \mu_1 = \mu_2$ against $H_a: \mu_1 < \mu_2$ (for a left tailed alternative)

Again, the following alternative situations may arise in testing the above mentioned null and alternative hypotheses

i) Independent samples with known population variances, sample sizes are large or small.

ii) Independent samples with unknown population variances, sample sizes are large.

iii) Independent populations for small sample sizes (≤ 29) with unknown but equal variances.

iv) Independent populations for small sample sizes (≤ 29) with unknown and unequal variances.

v) Correlated sample or matched sample (the sample obtained from a bi-variate normal population or paired observations).

The test statistic to be used for testing the simple hypothesis $H_0: \mu_1 - \mu_2 = 0$ against a one-tailed or two-tailed alternative is given by

Under situation (i): $Z = \dfrac{(X_1 - X_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0, 1)$.

Under situation (ii): $Z = \dfrac{(X_1 - X_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \sim N(0, 1)$.

where $s_1^2 = \dfrac{\Sigma(x_1 - \bar{x}_1)^2}{n_1 - 1}$ and $s_2^2 = \dfrac{\Sigma(x_2 - \bar{x}_2)^2}{n_2 - 1}$.

Under situation (iii): $t = \dfrac{(X_1 - X_2) - (\mu_1 - \mu_2)}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$

which is distributed as Student's t with $n_1 + n_2 - 2$ degrees of freedom, where, s^2 is pooled-estimate of variance, given by

$$s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

The value of $(\mu_1 - \mu_2) = 0$ for all Z and t defined above under null hypothesis $H_0 : \mu_1 - \mu_2 = 0$, but if any other value is specified (say, $\mu_1 - \mu_2 = \theta_0$) by null hypothesis for the difference between means, that means if

$$H_0 : \mu_1 - \mu_2 = \theta_0$$

then the value of $(\mu_1 - \mu_2)$ will be θ_0 instead of zero.

For testing the hypothesis regarding difference between two means under situation (i) to (iii) it is desirable to test the equality of two population variance to check if the assumptions of equal variances are valid or not. If the variances are found not to be equal, then, Student's t statistic can not be applied. Hence, under situation (iv), under null hypothesis, the test statistic is given by:

$$t' = \frac{(\overline{X}_1 - \overline{X}_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}.$$

Here, t' is not a Student's t statistic. The critical values of t' at $100\alpha\%$ level of significance are computed using the formula:

$$t'_\alpha = \frac{\frac{s_1^2 t_1}{n_1} + \frac{s_2^2 t_2}{n_2}}{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

where t_1 and t_2 are Student's with $(n_1 - 1)$ and $(n_2 - 1)$ df respectively at $100\alpha\%$ level of significance.

Again, under situation (v), let us consider a random sample of n matched pairs of observations (x_i, y_i) from a bi-variate normal population, then the test of the hypothesis $H_0 : \mu_1 - \mu_2 = 0$ requires to compute a statistic d defined as $d = x_i - y_i$, and the testing procedure is the same as testing the significance of a single mean of the observations obtained from the difference of two variables assuming that small sample size and unknown population variance. The statistic used for testing this type of hypothesis is called paired-t test, defined as

$$t = \frac{\overline{d}}{se(\overline{d})} \sim t_{n-1} \quad \text{under situation (v)}$$

which is distributed as t with $n-1$ df.

where, $\bar{d} = \frac{\sum d}{n}$, $s_d^2 = \frac{\sum (d-\bar{d})^2}{n-1}$ and $se(\bar{d}) = \frac{s_d}{\sqrt{n}}$.

- **Decision rule** The decision rules in all of the above cases are the same as that of testing the significance of a single mean using Z or t statistic.

Example 16.7.1 It is wanted to investigate if male and female typists earn comparable wages. The sample data for daily wages of male and female provide with the following information.

Table 16.6 Sample mean and variance of male and female typists

	Male	Female
Sample size	60	60
Mean wage	Taka 158.50	Taka 141.60
SD (Population)	Taka 18.20	Taka 20.60

Test whether the mean wages of male typists is more than female typists at 5% and 1% level of significance.

Solution Let the wages of male and female are normally and independently distributed with means μ_1 and μ_2, and known variances σ_1^2 and σ_2^2 respectively. It is a one-tailed test, so we consider the following hypothesis

$$H_0 : \mu_1 = \mu_2 \text{ against } H_1 : \mu_1 > \mu_2 .$$

As the sample sizes are large, under null hypothesis, the value of test statistic is given by

$$z = \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}.$$

For $\alpha = 0.05$, the critical region is $Z > 1.645$, and for $\alpha = 0.01$, the critical region is $Z > 2.33$.

Here, $\bar{x}_1 = 158.50$, $\bar{x}_2 = 141.6$, $\sigma_1 = 18.20$, $\sigma_2 = 20.60$, $n_1 = n_2 = 60$.

Thus, the computed value of Z is: $z = \frac{(158.50 - 141.6)}{\sqrt{\frac{(18.20)^2}{60} + \frac{(20.60)^2}{60}}} = 4.76.$

- **Conclusion** The computed value of Z is greater than critical values at both the level of significance, hence in the given city, male typists have on the average higher earnings than their female counterpart at 1% and 5% levels of significance. Here, the value z is highly significant.

Example 16.7.2 (Large sample sizes with unknown population variances) A potential buyer of electric bulbs bought 100 bulbs each of two famous brands A and B. Upon testing both these sample, he found that brand A had a mean life of 1500 hours with standard deviation of 50 hours whereas brand B had an average life of 1530 hours with standard deviation of 60 hours. Can it be concluded at 5% level of significance that the bulbs of two brands differ significantly in quality?

Solution We assume that the parent population of these two lifetimes are independently distributed with means μ_1 and μ_2 and unknown variances σ_1^2 and σ_2^2. We also assume that there is no significant difference in the quality of both brands so that brand A is as good as brand B in terms of operating hours. Hence, we have to test

$$H_0 : \mu_1 = \mu_2 \text{ against } H_1 : \mu_1 \neq \mu_2 .$$

Since sample sizes are large, so under the null hypothesis, the value of test statistic is

$$z = \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \sim N(0, 1) .$$

Again, $\alpha = 0.05$ and since it is a two tailed test, the critical region is

$$|Z| > 1.96.$$

Here, $\bar{x}_1 = 1500$, $\bar{x}_2 = 1530$, $s_1 = 50$ and $s_2 = 60$, $n_1 = n_2 = 100$.

Now, $\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = \sqrt{\frac{(50)^2}{100} + \frac{(60)^2}{100}} = 7.81.$

Thus, $z = \frac{(1500 - 1530)}{7.81} = -3.841$

Since the observed value of Z is greater than the critical value, the null hypothesis may be rejected at 5% level of significance.

- **p-value** From the table of Z, we have $P(Z < -3.00) = P(Z > 3.00) = 0.00$, hence the null hypothesis may be rejected even at 0% level of significance, so, $p = 0.00$. It is said that the value z is highly significant since the p-value is 0.00.

- **Confidence Interval** 95% confidence limits for the difference of two population means $(\mu_1 - \mu_2)$ are given by

$$(\bar{x}_1 - \bar{x}_2) \pm 1.96\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = (-30) \pm 7.81 = (22.19,\ 37.81)$$

Example 16.7.3 (Large sample sizes with unknown population variances) A professor taught two sections of an introductory marketing course using very different styles. In the first section approach was extremely formal and rigid, while in the second section an independent, more relaxed and informal attitude was adopted. At the end of the course, a common final examination was administered. In the first section the seventy two students obtained a mean score of 71.03 and the sample standard deviation was 22.91. In the second section there were sixty four students, with mean score 80.92 and standard deviation 23.11.

Assume that these two groups of students can be regarded as independent random samples from the populations of all students who might be exposed to these teaching methods. Test at 5% level of significance (i) whether the performance of these two methods are the same, (ii) whether second method is better than first one.

Solution We assume that the parent populations are independently distributed with means μ_1 and μ_2 and unknown variances σ_1^2 and σ_2^2.

(i) We have to test that there is no significant difference in two method, so that first method is as good as second method as far as the teaching system is concerned. Hence, we have to test

$$H_0 : \mu_1 = \mu_2 \text{ against } H_1 : \mu_1 \neq \mu_2 .$$

Since sample sizes are large, so under null hypothesis the value of test statistic is

$$z = \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \sim N(0,\ 1) .$$

Again, $\alpha = 0.05$ and since it is a two tailed test, the critical region is

$$|Z| > 1.96.$$

Here, $\bar{x}_1 = 71.03$, $\bar{x}_2 = 80.92$, $s_1 = 22.91$ and $s_2 = 23.11$, $n_1 = 72$, $n_2 = 64$.

Now, $$\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = \sqrt{\frac{(22.91)^2}{72} + \frac{(23.11)^2}{64}} = 3.95.$$

Thus, $$z = \frac{(71.03 - 80.92)}{3.95} = -2.50.$$

Since the computed absolute value of Z is greater than the critical value, the null hypothesis may be rejected at 5% level of significance. The sample supports that there is significant difference between average performances of the two methods.

- **p-value** From the table of Z, we have $P(Z < -2.50) = P(Z > 2.50) = 0.007$, hence the null hypothesis may be rejected even at 1.4% (since $0.007 \times 2 = 0.014$) level of significance, so $p = 0.014$.

- **Confidence Interval** 95% confidence limits for the difference of two population means $(\mu_1 - \mu_2)$ are given by

$$(\bar{x}_1 - \bar{x}_2)) \pm 1.96\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} = (-9.89) \pm 7.81 = (-17.70,\ 2.08).$$

(ii) Since the second method would be proved to be better than the first method, if the average score obtained by first method is significantly smaller than that of second method, so, in order to test whether second method is better than the first method, we have to consider a one-tailed test defined as

$$H_0 : \mu_1 = \mu_2 \text{ against } H_1 : \mu_1 < \mu_2 .$$

Here, since the sample sizes are large, the test statistic is also Z as defined in (i), but the critical region will cover the 5% area only in the left tail, thus the critical region is $Z < -1.645$.

The observed value of Z is -2.50, as found in (i), falls in the critical region, so the null hypothesis may be rejected at 5% level of significance. That means, the sample supports that the second method is on an average better than the first method.

Example 16.7.4 (Large sample sizes with known variances) A firm believes that the tires produced by process I on an average last longer than tires produced by process II. To test this belief, random samples of tires produced by two processes were tested and the results are as.

Table 16.7 Mean and standard deviation of lifetimes of tires

	Process I	Process II
Sample size	50	50
Average lifetime (in km.)	22,400	21,800
Population Standard deviation (in km.)	1000	1000

Is there any evidence at 5% level of significance that the firm is correct in its belief?

Solution We have to consider the following null and alternative hypothesis

$$H_0 : \mu_1 = \mu_2 \text{ against } H_1 : \mu_1 > \mu_2 .$$

Since the sample sizes are large, the value of test statistic under null hypothesis is

$$z = \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} .$$

Since it is a one tailed test, the critical region is: $Z > 1.645$.

Given, $\bar{x}_1 = 22400$, $\bar{x}_2 = 21800$, $\sigma_1 = 1000$, $\sigma_2 = 1000$, $n_1 = n_2 = 50$.

So, the value of Z is: $z = \dfrac{(22400 - 21800)}{\sqrt{\frac{(1000)^2}{50} + \frac{(1000)^2}{50}}} = 3.00.$

Since the calculated value of Z is more than its critical value at 5% level, therefore null hypothesis may be rejected. Hence, we can conclude that the tires produced by process I has longer life than process II.

- **p-value** From the table of area under the standard normal probability distribution, we have $P(Z > 3.00) = 0.001$, so the smallest critical value for which null hypothesis may be rejected is 0.001, thus the p-value is 0.001.

Example 16.7.5 (Small sample sizes with unknown and equal population variance) Manager of a factory I claims that the average wage of its workers is higher than that of factory II. A firm conducted a survey on daily wages of workers of two factories to see if the claim of

Table 16.8 Mean and standard deviation of lifetimes of tires

	Factory I	Factory II
Sample size	16	11
Sample mean wage	Taka 290	Taka 250
Sample standard deviation	15	50

Solution Let us assume that the wages of corresponding population of two factories are independent and normally distributed with common unknown variance σ^2. Thus the null and alternative hypothesis are

$$H_0: \mu_1 = \mu_2 \quad \text{against} \quad H_1: \mu_1 > \mu_2 .$$

Since the sample sizes are small and variances are unknown but equal, the test statistic is

$$t = \frac{\bar{x}_1 - \bar{x}_2 - (\mu_1 - \mu_2)}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

which is distributed as Student's t with $n_1 + n_2 - 2$ degrees of freedom,

where, $s^2 = \dfrac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$ is an estimate of common variance σ^2.

Here, $n_1 = 16$, $n_2 = 11$, so the degrees of freedom is $n_1 + n_2 - 2 = 25$.

It is a one tailed test, thus from the table of t-distribution we have the critical values of t at 5% level of significance with 25 df is 1.708, that means, $t_{25;0.05} = 1.708$, in other words, the critical region is: $t > 1.708$

Given, $\bar{x}_1 = 290$, $\bar{x}_2 = 250$, $s_1 = 15$, $s_2 = 50$, $s^2 = 1135$ and $s = 33.69$

So, under the null hypothesis, the value of t is

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{290 - 250}{13.19} = 3.03.$$

- **Decision** The observed value of t is greater than the critical value, it lies in the critical region, so the null hypothesis may be rejected at 5% level of significance.
- **Conclusion** The claim of manager of factory-I is justified.

Example 16.7.6 (Small sample sizes with unknown and equal population variances) The residence of Dhaka city complains that traffic speeding fines given in their city are higher than the traffic speeding fines that are given in Chittagong city. The appropriate authority agreed to study the problem. To check if the complaints were reasonable, independent random samples of the amounts paid by the residents for speeding fines in each of two cities over the last three months were obtained and shown in following table.

Table 16.9 Amounts of traffic speeding fines

Dhaka city	100	125	135	128	140	142	128	137	156	142
Chittagong city	95	87	100	75	110	105	85	95		

Assuming an equal population variance, test

i) whether there is any significant difference in the mean cost of speeding in these two cities and find the 95% confidence interval.

ii) whether the mean speeding cost in Dhaka city is higher than Chittagong city at 1% level of significance.

Solution **(i)** Let X_1 be the speeding cost in Dhaka city and X_2 be the speeding cost in Chittagong city. Assuming the samples have been drawn independently from two normal populations with means μ_1 and μ_2 respectively with common variance σ^2. We have to test the following hypothesis

$$H_0: \mu_1 = \mu_2 \text{ against } H_1: \mu_1 \neq \mu_2.$$

It is a two tailed test, the sample sizes are small and variances are also unknown, so, the appropriate test statistic is given by

$$t = \frac{\bar{x}_1 - \bar{x}_2 - (\mu_1 - \mu_2)}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

which is distributed as Student's t with $n_1 + n_2 - 2$ degrees of freedom, where, $s^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1 + n_2 - 2}$ is an estimate of common variance σ^2.

Given, $n_1 = 10$, $n_2 = 8$, so the degrees of freedom is: $n_1 + n_2 - 2 = 16$.

From the table of t-distribution we have the critical values of t at 5% level of significance with 16 df are ±2.12, that means, $t_{16;0.025} = 2.12$, in other words, the critical region is

$$|t| > 2.12.$$

From the given observations, we have,

$$\bar{x}_1 = 133.30,\ \bar{x}_2 = 94.00,\ s_1 = 18.20,\ s_2 = 20.60,\ s^2 = 371.98$$

So, under null hypothesis, the calculated value of t is

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{133.30 - 94.00}{9.14} = 4.30.$$

- **Decision** The observed absolute value of t is greater than the critical value, it lies in the critical region, so the null hypothesis may be rejected at 5% level of significance.
- **Conclusion** There is significant difference between the average traffic fines in two cities.

Again, the 95% confidence interval for $(\mu_1-\mu_2)$ is given by

$$(\bar{x}_1-\bar{x}_2)-t_{16;0.025}\times s\sqrt{\frac{1}{n_1}+\frac{1}{n_2}} < (\mu_1-\mu_2) < (\bar{x}_1-\bar{x}_2)+t_{16;0.25}\times s\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}$$

or, $39.30 \pm 2.12\times 4.30 = 39.30 \pm 9.12 = (30.18,\ 48.42)$.

- **p-value** From the table of critical values of t-distribution, we find that $t_{16;0.005}=2.921$, that means $2\times 0.005=0.010=1\%$ can be considered as the smallest level of significance at which the null is still rejected, so the required p-value is 0.01.

(ii) In this case we have to perform a one-tailed test given by

$$H_0:\mu_1=\mu_2 \text{ against } H_1:\mu_1>\mu_2.$$

Here the critical value of t at 1% level of significance with 16 df is: $t_{16;0.01}=2.583$

The observed value of t is same as previous one, so $t=4.30$.

In this case too, the observed t falls in the critical region, so the null hypothesis may be rejected at 1% level of significance. We can conclude that on an average the traffic fines at Dhaka city is higher than that of Chittagong city.

- **p-value** In this case the p-value is 0.005 because, $P(t>4.30)=0.005$.

Example 16.7.7 (Matched observations or paired sample) A study was conducted by a pharmaceutical company to compare the difference in effectiveness of two particular drugs in cholesterol levels. The company used paired sample approach to control variation in reduction that might be due to factors other than the drug itself. Each member of a pair was matched by age, weight, lifestyle, and other pertinent factors. Drug X was given to one person randomly selected from each pair, and drug Y was given to the other individual in the pair. After a specific period of time each person's cholesterol level was measured again. Suppose a random sample of eight pairs of patients with known cholesterol problems is selected from the large populations of participants. The following table gives the number of points by which each person's cholesterol level was reduced.

Table 16.10 Reduction levels of cholesterol by drugs

Pair	1	2	3	4	5	6	7	8
Drug X	29	32	31	32	32	29	31	30
Drug Y	26	27	28	27	30	26	33	36

i) Test whether there is any significant difference between the mean reductions of cholesterol levels by two drugs at 1% level of significance.

ii) Find 99% confidence interval for the difference between the population means.

Solution (i) Let the observations have been selected from a bi-variate normal population. So, it is necessary to use paired t-test for testing the difference between the mean reduction levels.

Let us formulate the null and alternative hypothesis as

$$H_0 : \mu_x = \mu_y \text{ against } H_1 : \mu_x \neq \mu_y .$$

The appropriate test statistic is: $t = \dfrac{\bar{d}}{se(\bar{d})} \sim t_{n-1}$, which is distributed as t with $n-1$ df,

where, $\bar{d} = \dfrac{\Sigma d}{n}$, $s_d^2 = \dfrac{\Sigma(d-\bar{d})^2}{n-1}$ and $se(\bar{d}) = \dfrac{s_d}{\sqrt{n}}$.

Here, $n=8$, degrees of freedom is $n-1=7$.

It is two-tailed test, so at 1% level of significance the critical region is

$$|t| > t_{n-1;\alpha/2} = t_{7;0.005} = 3.499$$

Now, let us construct the following table for calculation of mean and standard deviation of d.

Table 16.11 Calculation of mean and standard deviation of d

Pair	1	2	3	4	5	6	7	8
Drug X	29	32	31	32	32	29	31	30
Drug Y	26	27	28	27	30	26	33	36
$d=X-Y$	3	5	3	5	2	3	–2	–6
$(d-\bar{d}))$	1.38	3.38	1.38	3.38	0.38	1.38	–3.63	–7.63
$(d-\bar{d})^2$	1.89	11.39	1.89	11.39	0.14	1.89	13.14	58.14

We have $\bar{d} = 1.625$, $s_d^2 = 14.27$, $s_d = 3.777$.

So, $t = \dfrac{\bar{d}}{se(\bar{d})} = \dfrac{\bar{d}}{s_d/\sqrt{n}} = \dfrac{1.625}{3.777/\sqrt{8}} = 1.22.$

- **Decision** The observed value of t is smaller than absolute value of critical value, that means it does no fall in the critical region, so we fail to reject null hypothesis.
- **Conclusion** There is no significant difference between the average reduction of cholesterol level by two drugs, that means, drug X and drug Y are equally effective.

(ii) The 99% confidence interval for the difference of population average reduction of cholesterol levels is given by

$$\bar{d} - t_{n-1;\alpha/2} \cdot se(\bar{d}) < \mu_x - \mu_y < \bar{d} + t_{n-1;\alpha/2} \cdot se(\bar{d}) = (-3.05,\ 6.30).$$

Example 16.7.8 (Matched observations or paired sample) Ten persons were appointed as probationary officers in an office. Their performance was noted by taking a test and the marks were recorded out of 100. After training for a 6-months period, another test was conducted. The marks obtained by the officers before and after training were as follows.

Table 16.12 Marks of employees before and after training

Employees	A	B	C	D	E	F	G	H	I	J
Before training	80	76	92	60	70	56	74	56	70	56
After training	84	70	96	80	70	52	84	72	72	50

Were the employees benefited by the training?

Solution It would be proved that the employees were not benefited by the training if there is no significant difference between the average score obtained by them before and after training. On the other hand, it would be proved to be benefited if the average score obtained by employees significantly improved after training.

So we have to conduct a one tailed test defined by the null and alternative hypothesis as

$$H_0 : \mu_a = \mu_b \text{ against } H_1 : \mu_a < \mu_b .$$

The appropriate test statistic is $t = \dfrac{\bar{d}}{se(\bar{d})} \sim t_{n-1}$, which is distributed as t with $n-1$ df,

where, $\bar{d} = \dfrac{\sum d}{n}$, $s_d^2 = \dfrac{\Sigma(d-\bar{d})^2}{n-1}$ and $se(\bar{d}) = \dfrac{s_d}{\sqrt{n}}$.

Here, $n=10$, degrees of freedom is: $n-1=9$.

It is one-tailed test, so at 5% level of significance the critical region is,

$$t < -t_{n-1;\alpha} = -t_{9;0.05} = -2.62.$$

Now, let us construct the following table for calculation of mean and standard deviation of d.

Table 16.13. Computation of mean and standard deviation of d

Employees	A	B	C	D	E	F	G	H	I	J
Before training (X)	80	76	92	60	70	56	74	56	70	56
After training (Y)	84	70	96	80	70	52	84	72	72	50
$d = X - Y$	–4	6	–4	–20	0	4	–10	–16	-2	6
$(d - \bar{d})$	0.00	10.00	0.00	–16.00	4.00	8.00	–6.00	-12.00	2.00	10.00
$(d - \bar{d})^2$	0.00	100.00	0.00	256.00	16.00	64.00	36.00	144.00	4.00	100.00

We have: $\bar{d} = -40/10 = -4$, $s_d^2 = 80$, $s_d = 8.944$.

So, $t = \dfrac{\bar{d}}{se(\bar{d})} = \dfrac{\bar{d}}{s_d/\sqrt{n}} = \dfrac{-4}{8.944/\sqrt{10}} = -1.414.$

- **Decision** The observed value of t does not fall in the critical region, so we fail to reject null hypothesis that means the null hypothesis holds true.

- **Conclusion** It can be concluded that the employees were not benefited by training.

6d

16.8 Test of Hypothesis Concerning Attributes

In case of attributes we can only find out the presence or absence of a certain qualitative characteristics. For example, in the study of attribute 'employment', a survey may be conducted and the people may be classified as employed and unemployed, in the study of the attribute 'effectiveness' of a drug, the patients may be classified as cured and not cured and in the study of attribute 'size', the items may be classified as small and large. The appearance of an attribute may be considered as success and its non-appearance as failure. Obviously, this type of two outcomes will follow binomial distribution. Thus, when the sample size is large, we can perform the following tests with the attributes having two categories

i) Test of a population proportion for a specified value;
ii) Test of the difference between two population proportions;
iii) Tests of independence of two attributes (having two or more categories of each attribute).

The test regarding independence of attributes is discussed in section 16.12.

6d

16.8.1 Test of hypothesis about a population proportion Suppose we have a sample of n observations from a population, a proportion π of population follow a particular attribute. Then, if the number of sample observations n is large, and the observed sample proportion is P. For testing the significance of this population proportion for a given value π_0, we consider the following null and alternative hypothesis

$$H_0 : \pi = \pi_0, \text{ against the alternative } H_1 : \pi \neq \pi_0 \text{ (two tailed test).}$$

We know, the sampling distribution of P is normal with mean π and variance $\sigma_P^2 = \frac{\pi(1-\pi)}{n}$ and under H_0 it is $\frac{\pi_0(1-\pi_0)}{n}$.

So, under the null hypothesis, the test statistic is

$$Z = \frac{P-\pi_0}{se(P)} = \frac{P-\pi_0}{\sqrt{\frac{\pi_0(1-\pi_0)}{n}}} \sim N(0,\ 1).$$

The decision rule is: Reject H_0 in favor of H_1 at $100\alpha\%$ level of significance if

$$z = \frac{P-\pi_0}{\sqrt{\frac{\pi_0(1-\pi_0)}{n}}} > z_{\alpha/2} \text{ or, } \frac{P-\pi_0}{\sqrt{\frac{\pi_0(1-\pi_0)}{n}}} < -z_{\alpha/2} \text{ for a two-tailed test,}$$

or, if $|Z| > z_{\alpha/2}$

Similarly, for a right tailed test, the decision rule is

Reject H_0 in favor of H_1 at $100\alpha\%$ level of significance if $\dfrac{P-\pi_0}{\sqrt{\dfrac{\pi_0(1-\pi_0)}{n}}} > z_\alpha$.

and for a left-tailed test, the decision rule is

Reject H_0 in favor of H_A at $100\alpha\%$ level of significance if $\dfrac{P-\pi_0}{\sqrt{\dfrac{\pi_0(1-\pi_0)}{n}}} < -z_\alpha$.

Remark P is an estimator or statistic and p is a estimate which is sample value of P.

Example 16.8.1 For the following questions carry out the test of significance of population proportions at 5% level of significance (where x represents the number of things of particular category).

i) $H_0: \pi = 0.25,\ H_1: \pi \neq 0.25,\ n = 100,\ x = 40;$
ii) $H_0: \pi = 0.40,\ H_1: \pi > 0.40,\ n = 200,\ x = 100;$
iii) $H_0: \pi = 0.30,\ H_1: \pi < 0.30,\ n = 400,\ x = 100.$

Solution The statistic to be used for testing the given hypothesis is given by

$$Z = \frac{P-\pi_0}{\sqrt{\dfrac{\pi_0(1-\pi_0)}{n}}}$$ where, P is the estimator of population proportion.

(i) This is a two tailed test, so the decision rule is

Reject H_0 in favour of alternative at 5% level if $|Z| > z_{0.025}$ or, $|Z| > 1.96$.

Here, $n = 100$, so, $p = \dfrac{40}{100} = 0.40$ and $\pi_0 = 0.25$, so the computed value of test statistic is

$$z = \frac{p-\pi_0}{\sqrt{\dfrac{\pi_0(1-\pi_0)}{n}}} = \frac{0.40-0.25}{\sqrt{\dfrac{0.25\times(1-0.25)}{100}}} = 3.46.$$

- **Decision** $|Z| > 1.96$, the computed value of Z falls in the critical region, so we fail to accept null hypothesis.

(ii) This is a right tailed test, so the decision rule is

Reject H_0 in favour of alternative at 5% level if $Z > z_{0.05} = 1.645$.

Here, $n = 200$, so, $P = \dfrac{100}{200} = 0.50$ and $\pi_0 = 0.40$, so the computed value of Z is

$$z = \frac{P - \pi_0}{\sqrt{\frac{\pi_0(1-\pi_0)}{n}}} = \frac{0.50 - 0.40}{\sqrt{\frac{0.40 \times (1-0.40)}{200}}} = 2.89.$$

- **Decision** The observed value of $Z > 1.645$, which falls in the critical region, so we fail to accept null hypothesis.

(iii) This is a left tailed test, so the decision rule is

Reject H_0 in favour of alternative at 5% level if $Z < -z_{0.05} = -1.645$.

Here, $n = 400$, so, $P = \frac{100}{400} = 0.25$ and $\pi_0 = 0.30$, so the computed value of Z is

$$z = \frac{P - \pi_0}{\sqrt{\frac{\pi_0(1-\pi_0)}{n}}} = \frac{0.25 - 0.30}{\sqrt{\frac{0.30 \times (1-0.30)}{400}}} = -2.18.$$

- **Decision** The o bserved value of $Z < -1.645$, which falls in the critical region, so we fail to accept null hypothesis.

Example 16.8.2 Forecasts of corporate earnings per share are made on a regular basis by many financial analysts. In a random sample of 600 forecasts, it was found that 382 of these forecasts exceeded the actual outcome for earnings. Test against a two tailed alternative the null hypothesis that the population proportion of forecasts that are higher than actual outcomes is 0.50 at 5% level of significance.

Solution Let π denotes the population proportion and P denotes the sample proportion of forecasts that are above the actual outcomes. We are interested to test the hypothesis

$H_0 : \pi = 0.50$, against the alternative $H_1 : \pi \neq 0.50$.

The appropriate test statistic is: $Z = \frac{P - \pi_0}{\sqrt{\frac{\pi_0(1-\pi_0)}{n}}}$.

The decision rule is: Reject H_0 in favour of alternative if $|Z| > z_{\alpha/2}$ or $|Z| > 1.96$.

Here, $n = 600$, $P = \frac{382}{600} = 0.637$ and $\pi_0 = 0.50$.

So, the computed value of test statistic is:

$$z = \frac{P - \pi_0}{\sqrt{\frac{\pi_0(1-\pi_0)}{n}}} = \frac{0.637 - 0.50}{\sqrt{\frac{0.50(1-0.50)}{600}}} = 6.71.$$

Since 6.71 is much bigger than 1.96, the null hypothesis is clearly rejected. That means, forecasts of corporate earnings that exceed the actual values are significantly different from 0.50.

Example 16.8.3 A manufacturer claims that at least 95% of the equipments which he supplied to a factory conformed to the specification. An examination of the sample of 200 pieces of equipment revealed that 18 were faulty. Test the claim of manufacturer at 5% level of significance.

Solution According to the statement of the problem, it is better if we consider a composite hypothesis such that at least 95% if the equipments supplied by the company conformed to the specification, that means:

$$H_0: \pi \geq \pi_0 = 0.95, \text{ against the alternative } H_1: \pi < 0.95 \text{ (one tailed test).}$$

The appropriate test statistic is: $Z = \dfrac{P - \pi_0}{\sqrt{\dfrac{\pi_0(1-\pi_0)}{n}}}$

$\alpha = 0.05$, the decision rule is : Reject H_0 in favour of alternative if $Z < -z_{\alpha/2}$ or $Z < -1.645$ (for this left-tailed test).

Here, $n = 200$, out of 200, 18 were found faulty, that means $(200-18) = 182$ equipments conform to the specification, so $P = \dfrac{182}{200} = 0.91$ and $\pi_0 = 0.95$.

So, the computed value of test statistic is: $z = \dfrac{P - \pi_0}{\sqrt{\dfrac{\pi_0(1-\pi_0)}{n}}} = \dfrac{0.91 - 0.95}{\sqrt{\dfrac{0.95(1-0.95)}{200}}} = -2.67.$

Since the observed value of $z = -2.67$ is less than the critical value -1.645, it lies in the critical region, so the null hypothesis is clearly rejected. That means, the proportion of equipments conforming to the specification is greater than 95%.

Example 16.8.4 An auditor claims that 10 percent of the customers' ledger accounts of a bank are carrying mistakes of posting and balancing. A random sample of 600 accounts was taken to test the accuracy of posting and balancing, and 45 accounts were found to have mistakes. Are these sample results consistent with the claim of auditor? (use 5% level of significance).

Solution Let us take the null that the claim of the auditor is valid, that means

$$H_0: \pi \geq \pi_0 = 0.10, \text{ against the alternative } H_1: \pi_0 \neq 0.10$$

The appropriate test statistic is: $Z = \dfrac{P - \pi_0}{\sqrt{\dfrac{\pi_0(1-\pi_0)}{n}}}$.

$\alpha = 0.05$, the decision rule is: Reject H_0 in favour of alternative if

$$|Z| > z_{\alpha/2} \text{ or } |Z| > 1.96.$$

Here, n = 200, so, $P = \frac{45}{600} = 0.075$ and $\pi_0 = 0.10$.

So, the computed value of test statistic is: $z = \frac{p-\pi_0}{\sqrt{\frac{\pi_0(1-\pi_0)}{n}}} = \frac{0.075-0.10}{\sqrt{\frac{0.10(1-0.10)}{600}}} = -2.049.$

Since the observed value of $|z| = 2.049$ is greater than the critical value 1.96, it lies in the critical region, so the null hypothesis is rejected at 5% level of significance. That means, the claim of the auditor is not valid.

Example 16.8.5 Suppose machine produces 12% faulty items. A manufacturer of the same type of machine claims that there machine is better than this machine. In order to test the manufacturer's claim, a random sample of 300 items were checked and 30 items were found to be faulty. On the basis of the information, comment on the claim of manufacturer.

Solution The manufacturer's claim will be proved to be justified if it produces less proportion of defective items than the existing one. So, let us think that the existing machine is as better as the new machine, and consider the null hypothesis as:

$$H_0 : \pi = \pi_0 = 0.12, \text{ against the alternative } H_1 : \pi_0 < 0.12 \text{ (one tailed test).}$$

The appropriate test statistic is: $Z = \frac{P-\pi_0}{\sqrt{\frac{\pi_0(1-\pi_0)}{n}}}$.

Let $\alpha = 0.05$, the decision rule is: Reject H_0 in favour of alternative if

$$Z < -z_{\alpha/2} \text{ or } Z < -1.96.$$

Here, n = 300, so, $P = \frac{30}{300} = 0.10$ and $\pi_0 = 0.12$.

So, the computed value of test statistic is: $z = \frac{P-\pi_0}{\sqrt{\frac{\pi_0(1-\pi_0)}{n}}} = \frac{0.10-0.12}{\sqrt{\frac{0.12(1-0.12)}{300}}} = -1.066.$

Since the observed value of $z = -1.06$ does not fall in the critical region, so we fail to reject the null hypothesis at 5% level of significance. That means, the claim of the manufacturer is not justified and the new machine is as better as the existing one.

69 **16.8.2 Test of hypothesis about difference between two population proportions** Let P_1 and P_2 be the sample proportions obtained from large samples of sizes n_1 and n_2 from respective population having proportions π_1 and π_2. We are interested to test the hypothesis that there is no difference between the population proportions, that means,

$$H_0 : \pi_1 = \pi_2, \text{ against the alternative } H_1 : \pi_1 \neq \pi_2$$

The sampling distribution of difference between sample proportions $(P_1 - P_2)$ is normal with mean $E(P_1 - P_2) = \pi_1 - \pi_2$ and variance $\sigma_p^2 = \frac{\pi_1(1-\pi_1)}{n_1} + \frac{\pi_2(1-\pi_2)}{n_1}$.

Since the sample sizes are large, the test statistic is: $Z = \frac{(P_1 - P_2) - (\pi_1 - \pi_2)}{\sqrt{\frac{\pi_1(1-\pi_1)}{n_1} + \frac{\pi_2(1-\pi_2)}{n_2}}} \sim N(0, 1)$.

However, since all tests are undertaken under null hypothesis, so if the null hypothesis is true, then P_1 and P_2 are two independent unbiased estimators of the same population parameter $\pi_1 = \pi_2 = \pi$. Thus the best estimate of the common proportion π is the pooled proportions P for two samples. The pooled estimate of π is the weighted mean of two sample proportions, given by

$$P = \frac{n_1 P_1 + n_2 P_2}{n_1 + n_2}.$$

The test statistic Z then becomes

$$Z = \frac{(P_1 - P_2) - (\pi_1 - \pi_2)}{\sqrt{P(1-P)\left[\frac{1}{n_1} + \frac{1}{n_2}\right]}} = \frac{(P_1 - P_2)}{\sqrt{P(1-P)\left[\frac{1}{n_1} + \frac{1}{n_2}\right]}} \text{ under } H_0.$$

The decision rule is for a two tailed test, reject H_0. in favor of H_1 at $100\alpha\%$ level of significance if

$$\frac{(P_1 - P_2)}{\sqrt{P(1-P)\left[\frac{1}{n_1} + \frac{1}{n_2}\right]}} > z_{\alpha/2} \text{ or, } \frac{(P_1 - P_2)}{\sqrt{P(1-P)\left[\frac{1}{n_1} + \frac{1}{n_2}\right]}} < -z_{\alpha/2} \text{ or, if } |Z| > z_{\alpha/2}.$$

Similarly, for a right tailed test, the decision rule is : Reject H_0 in favor of H_1 at $100\alpha\%$ level of significance if $\frac{(P_1 - P_2)}{\sqrt{P(1-P)\left[\frac{1}{n_1} + \frac{1}{n_2}\right]}} > z_\alpha$.

And for a left-tailed test, the decision rule is : Reject H_0 in favor of H_1 at $100\alpha\%$ level of significance if

$$\frac{(P_1 - P_2)}{\sqrt{P(1-P)\left[\frac{1}{n_1} + \frac{1}{n_2}\right]}} < z_\alpha.$$

Example 16.8.6 A company is considering two different television advertisements for promotion of a new product. Manager believes that advertisement A is more effective than advertisement B. Two test market areas with virtually identical consumer characteristics are selected: advertisement A is used in one area and advertisement B is used in other area. In a random sample of 60 customers who saw the advertisement A, 18 had tried to buy the product, on the other hand, in a random sample of 100 consumers who saw advertisement B, 22 had tried to buy the product. Does this indicate that the advertisement A is more efficient than advertisement B, if level of significance is 5%?

Solution Let π_1 and π_2 be the population proportions of customers who had tried to buy the products after seeing the advertisement A and advertisement B respectively, then we consider the null hypothesis as both advertisements are equally effective, that means

$$H_0 : \pi_1 = \pi_2, \text{ against the alternative } H_1 : \pi_1 > \pi_2.$$

(one tailed test, because, advertisement A will be considered more effective if proportion of customers who had tried to buy in this case is more than that of advertisement B).

Under the null hypothesis, the appropriate test statistic is

$$Z = \frac{(P_1 - P_2)}{\sqrt{P(1-P)\left[\frac{1}{n_1} + \frac{1}{n_2}\right]}}$$

where, P is the pooled estimate of proportion.

Given, $\alpha = 0.05$, and it is a right-tailed test, so the decision rule is: Reject H_0 in favour of H_1 if $Z > z_{0.05}$ or $Z > 1.645$.

Here, $n_1 = 60$, proportions $p_1 = \frac{18}{60} = 0.30$ and $n_2 = 100$, $p_2 = \frac{22}{100} = 0.22$ and the pooled estimate of proportion is

$$p = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2} = \frac{60 \times 0.30 + 100 \times 0.22}{60 + 100} = 0.25.$$

Thus, $$z = \frac{0.30 - 0.22}{\sqrt{0.25(1-0.25)\left\{\frac{1}{60} + \frac{1}{100}\right\}}} = 1.131.$$

Since, observed value z is less than the critical value 1.645, we fail to reject the null hypothesis at 5% level of significance.

Hence, we can conclude that there is no significant difference in the effectiveness of the two advertisements.

Example 16.8.7 800 units from factory A are inspected and 12 are found to be defective, 500 units from factory B are inspected and 12 are found to be defective. Can it be concluded at 5% level of significance that production at factory A is better than factory B?

Solution Let π_1 and π_2 be the population proportions of defectives of factory A and factory B respectively, then we consider that null hypothesis that performance of both factories are the same, that means,

$$H_0 : \pi_1 = \pi_2, \text{ against the alternative } H_A : \pi < \pi_2,$$

(it is a one tailed test, because, factory A can be considered as better if the proportion of defectives found in factory A is less than the proportion of defectives found in factory B).

The appropriate test statistic is $Z = \dfrac{(P_1 - P_2)}{\sqrt{P(1-P)\left[\dfrac{1}{n_1} + \dfrac{1}{n_2}\right]}}$

where, P_1 and P_2 are the sample proportions, and P is the pooled estimate of proportion given by

$$P = \frac{n_1 P_1 + n_2 P_2}{n_1 + n_2}.$$

Given, $\alpha = 0.05$, and it is a left-tailed test, so the decision rule is: Reject H_0 in favour of H_1 if $Z < -z_{0.05}$ or, $Z < -1.645$

Here, $n_1 = 800$, $p_1 = \dfrac{12}{800} = 0.015$ and $n_2 = 500$, $p_2 = \dfrac{12}{500} = 0.024$ and the pooled estimate of proportion

$$p = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2} = \frac{800 \times 0.015 + 500 \times 0.024}{800 + 500} = 0.018.$$

Thus, $$z = \frac{0.015 - 0.024}{\sqrt{0.018(1-0.018)\left\{\dfrac{1}{800} + \dfrac{1}{100}\right\}}} = -1.184.$$

Since, observed values of Z is not less than the critical value -1.645, we fail to reject the null hypothesis, that means the null hypothesis holds good at 5% level of significance.

Hence, we cannot conclude that the production at factory A is better than B.

Example 16.8.8 In a random sample of 700 workers from a particular factory of Bangladesh 200 are found to be smokers. In another factory out of 1300 workers 400 were found to be smokers. Can you conclude that there is a significant difference between the two factories with regard to the smoking habit?

Solution Let π_1 and π_2 be the population proportions of smokers in two factories respectively. Let us take the hypothesis that there is no difference in smoking habits in the two factories, and then the null hypothesis is

$$H_0 : \pi_1 = \pi_2, \text{ against the alternative } H_1 : \pi_1 \neq \pi_2.$$

The appropriate test statistic is: $Z = \dfrac{(P_1 - P_2)}{\sqrt{P(1-P)\left[\dfrac{1}{n_1} + \dfrac{1}{n_2}\right]}}$

where, P_1 and P_2 are the sample proportions, and P is the pooled estimate of proportion given by $P = \dfrac{n_1P_1 + n_2P_2}{n_1 + n_2}$.

Let $\alpha = 0.05$, and it is a two-tailed test, so the decision rule is: Reject Ho in favour of H_1 if $|Z| > 1.96$.

Here, $n_1 = 700$, $p_1 = \dfrac{200}{700} = 0.2857$ and $n_2 = 500$, $p_2 = \dfrac{400}{1300} = 0.3077$ and

the pooled estimate of proportion: $p = \dfrac{n_1p_1 + n_2p_2}{n_1 + n_2} = \dfrac{200400}{700+1300} = 0.30$.

Thus, $z = \dfrac{0.2857 - 0.3077}{\sqrt{0.30(1-0.30)\left\{\dfrac{1}{700} + \dfrac{1}{1300}\right\}}} = -1.023$.

Since the observed absolute value of Z is less than 1.96, there is no evidence to doubt the hypothesis that means, workers of two factories do not differ significantly with respect to the smoking habit.

Example 16.8.9. In a simple random sample of 600 men taken from a big city in which 400 are found to be smokers. In another simple random sample of 900 men taken from another city in which 450 are found smokers. Do the data indicate that there is a significant difference in the habit of smoking in the two cities? N.U.BBA -2012, 2008

Solution. Let π_1 *and* π_2 be the population proportions of smokers habit in the two cities respectively. The null hypothesis is H_0: $\pi_1 = \pi_2$; $H_1: \pi_1 \neq \pi_2$

Let p_1 and p_2 are the sample proportions. Here $p_1 = \dfrac{400}{600} = 0.667$,

$p_2 = \dfrac{450}{900} = 0.50$ and

$p = \dfrac{400 + 450}{600 + 900} = \dfrac{850}{1500} = 0.567$.

The appropriate test statistic is: $Z = \dfrac{(P_1 - P_2)}{\sqrt{P(1-P)\left[\dfrac{1}{n_1} + \dfrac{1}{n_2}\right]}}$

Thus

$$z=\frac{0.667-0.50}{\sqrt{0.567(1-0.567)\left\{\frac{1}{600}+\frac{1}{900}\right\}}}=\frac{0.167}{\sqrt{0.567(0.433)(0.0028}}=\frac{0.167}{\sqrt{0.00069}}=\frac{0.167}{0.026}=6.42$$

Comment . The critical value of Z at 1% level of significane of two tailed test is 2.575 but the observed vale of Z is 6.42 which is greater than that . The null hypothesis rejected and the observed value is highly signicant. Hence there exists a significance difference in the smoking habits of the two cities.

16.9 Test of Hypothesis about Correlation Co-efficient

Suppose (x_1, y_1), (x_2, y_2), ..., (x_n, y_n) be n pairs of observations of a random sample drawn from a bi-variate normal population with correlation co-efficient ρ, then the sample correlation co-efficient between n pairs of observations is given by:

$$r=\frac{\sum(x-\bar{x})(y-\bar{y})}{\sqrt{\sum(x-\bar{x})^2\sum(y-\bar{y})^2}}=\frac{\sum xy-\frac{\sum x\sum y}{n}}{\sqrt{\left\{\sum x^2-\frac{(\sum x)^2}{n}\right\}\left\{\sum y^2-\frac{(\sum y)^2}{n}\right\}}}.$$

The significance of population correlation co-efficient from which the sample has been drawn may be tested under following two assumptions:

(i) $H_0:\rho=0$, when the population correlation co-efficient is zero (with one-tailed or two-tailed alternative),

(ii) $H_0:\rho=\rho_0$, when the population correlation co-efficient is equal to some specified value ρ_0 (with one-tailed or two-tailed alternative).

16.9.1 Testing the hypothesis when the population correlation co-efficient equals zero Here the null hypothesis is considered as there is no correlation in the population that means, the relationship between the variables is not linear. The test is undertaken considering the following null hypothesis $H_0:\rho=0$: (in this case usually a two-tailed test $H_0:\rho\neq 0$ is conducted, however, one-tailed test is also conducted when the situation deserves)

Let r be the sample correlation co-efficient (which is the best estimate of population correlation co-efficient ρ). It is to be noted here that, for a sample of size n, the variance of r is given by $\text{var}(r)=\sqrt{\frac{1-r^2}{n-2}}$, then the appropriate test statistic to be used for testing the above mentioned hypothesis is:

$$t=\frac{r}{\sqrt{\frac{1-r^2}{n-2}}}=\frac{r\sqrt{n-2}}{\sqrt{1-r^2}}\sim t_{n-2}$$

which follows t-distribution with $n-2$ degrees of freedom.

Thus, if the computed value of r is greater than the tabulated value of t at 100α% level of significance, then the null hypothesis is rejected that means the decision rule is

$$\text{Reject } H_0 \text{ if } |t| > t_{n-2;\alpha/2} \quad \text{or,} \quad \left|\frac{r\sqrt{n-2}}{\sqrt{1-r^2}}\right| > t_{n-2;\alpha/2}$$

which indicates that the sample data provide sufficient evidence that $\rho \neq 0$.

However, an approximate 'rule of thumb' for considering population correlation co-efficient ρ to be significantly different from zero is given by

$$|r| > \frac{2}{\sqrt{n}}.$$

16.9.2 Testing the hypothesis when the population correlation co-efficient equals some specified value ρ_0 Let us consider the following null and alternative hypothesis

$$H_0: \rho = \rho_0 \quad \text{against} \quad H_1: \rho \neq \rho_0.$$

The test statistic t used for testing $H_0: \rho = 0$ is appropriate when $\rho = 0$ (under null hypothesis), however, when $\rho \neq 0$, the test statistic is not appropriate. Thus, in testing the above mentioned hypothesis, at first we have to use Fisher's z transformation. Hence, r is to be transformed into z given by $z = \frac{1}{2}\log_e \frac{1+r}{1-r} = 1.1513\log_{10}\frac{1+r}{1-r}$ (since $\log_e x = 2.3026\log_{10} x$).

It has been found that Z is normally distributed with mean

$$z_0 = \frac{1}{2}\log_e \frac{1+\rho_0}{1-\rho_0} = 1.1513\log_{10}\frac{1+\rho_0}{1-\rho_0}$$

(under null hypothesis) and standard deviation $\frac{1}{\sqrt{n-3}}$.

Therefore, the test statistic for testing the null hypothesis $H_0: \rho = \rho_0$ is given by

$$Z = \frac{z - z_0}{\sqrt{\frac{1}{n-3}}} = (z - z_0)\sqrt{n-3}$$

which approximately follows N(0, 1). The approximation is reasonably good if sample size is large.

16.10 Test of Hypothesis about Regression Co-efficient

We know, for n pairs of observations $(x_1, y_1), (x_2, y_2), \dots\dots\dots, (x_n, y_n)$ drawn from a bi-variate normal population, the least squares estimate of parameter β, the regression co-efficient of a population regression model, $y = \alpha + \beta x + \varepsilon$, where $\varepsilon \sim NID(\mu, \sigma^2)$, is given by

$$b = \frac{\Sigma(x_i - \bar{x})(y_i - \bar{y})}{\Sigma(x_i - \bar{x})^2} = \frac{sp(x,y)}{ss(x)}$$

It has been found that b is an unbiased estimate of β, that means E(b)=β, and variance of b is given by $Var(b) = \dfrac{s^2}{\Sigma(x_i - \bar{x})^2}$, s^2 is unbiased estimate of σ^2.

$$s^2 = s^2 = \frac{\text{SS due to Error}}{n-2} = \frac{\Sigma(y_i - \hat{y})^2}{n-2} = \frac{\Sigma(y_i - a - bx)^2}{n-2},$$

s^2 can also be estimated using the relationship:

$$s^2 = \frac{\Sigma(y-\bar{y})^2 - b\Sigma(y-\bar{y})(x-\bar{x})}{n-2}$$

Then, $se(b) = \dfrac{s}{\sqrt{\Sigma(x_i - \bar{x})^2}} = \dfrac{s}{\sqrt{\Sigma x^2 - \dfrac{(\Sigma x)^2}{n}}}$

In order to test the significance of parameter β, the following null and alternative hypothesis is formulated:

$$H_0 : \beta = \beta_0 \text{ against } H_1 : \beta \neq \beta_0.$$

The test statistic for testing the hypothesis is given by: $t = \dfrac{b - \beta_0}{se(b)} \sim t_{n-2}$.

The decision rule is: Reject H_0 at $100\alpha\%$ level of significance, if $|t| > t_{n-2;\alpha/2}$

or, if $t = \dfrac{b-\beta_0}{se(b)} > t_{n-2;\alpha/2}$ or, $t = \dfrac{b-\beta_0}{se(b)} < -t_{n-2;\alpha/2}$.

Similarly, one-tailed test of regression parameter can be conducted, if necessary.

Note that very often testing the significance of regression co-efficient implies testing the null hypothesis of zero population regression co-efficient, in that case, it is required to formulate null hypothesis as $H_0 : \beta = 0$, that means, substitute zero in the place of β_0, all other steps are the same as testing the null hypothesis $H_0 : \beta = \beta_0$.

The decision rule for testing the significance of correlation co-efficient and regression co-efficient using t-test statistic are as follows.

Table 16.14 Decision rule for correlation and regression co-efficient test

Case No.	Types of test		Decision rule
	For correlation co-efficient	For regression co-efficient	Reject H_0, if
1	$H_1 : \rho \neq 0$	$H_1 : \beta \neq \beta_0$	$\|t\| > t_{\alpha/2;(n-2)}$
2	$H_1 : \rho > 0$	$H_1 : \beta > \beta_0$	$t > t_{\alpha;(n-2)}$
3	$H_1 : \rho < 0$	$H_1 : \beta < \beta_0$	$t < -t_{\alpha;(n-2)}$

• **Confidence interval (CI) for β** If the regression errors ε's are normally distributed and all the assumptions of linear regression model hold, then the $100(1-\alpha)\%$ CI for the population regression co-efficient β is given by

$$b - t_{n-2;\alpha/2}\,se(b) < \beta < b + t_{n-2;\alpha/2}\,se(b).$$

Example 16.10.1 In a study of relationship between expenditure (x) and sales volume (y), a sample of 10 firms yields the co-efficient of correlation $r=0.93$. Can it be concluded on the basis of information that x and y are linearly related? (use $\alpha=0.05$).

Solution We have to test the hypothesis $\rho=0$ against $H_0:\rho \neq 0$.

The test statistic for testing the hypothesis is given by

$$t=\frac{r\sqrt{n-2}}{\sqrt{1-r^2}} \sim t_{n-2}.$$

Here, $n=10, r=0.93, df=n-2=8.$

It is a two-tailed test and $\alpha=0.05$, so from the table of t-distribution, we find the critical value of t at 2.5% level of significance with 8 df is 2.306.

Now the value of test statistic is: $t=\dfrac{0.93\sqrt{10-2}}{\sqrt{1-(0.93)^2}}=7.03.$

Since the computed value of t is much greater than the tabulated value of t, the null hypothesis may be rejected at 5% level of significance. Hence, it may be concluded that x and y are linearly related.

Example 16.10.2 Suppose, in a study of demand and supply of a commodity for 20 months, it has be found that $r=0.884$. Test whether this correlation is significantly different from a hypothesized value 0.92 at 5% level of significance.

Solution Here we have to test the following null and alternative hypothesis

$$H_0:\rho = 0.92 \text{ against } H_1:\rho \neq 0.92.$$

The test statistic for testing the hypothesis is given by

$$Z=(z-z_0)\sqrt{n-3}$$

where, $z=\frac{1}{2}\log_e\frac{1+r}{1-r}=1.1513\times\log_{10}\frac{1+r}{1-r}$ and $z_0=1.1513\times\log_{10}\frac{1+\rho_0}{1-\rho_0}$.

This is a two-tailed test, so the critical region of Z at 5% level of significance is: $|Z|>1.96$.

Here, $r=0.884$, $\rho_0=0.92$ and $n=20$,

$$z=1.1513\times\log_{10}\frac{1+r}{1-r}=1.1513\times\log_{10}\frac{1+0.884}{1-0.884}=1.3938.$$

and $z_0 = 1.1513 \times \log_{10}\frac{1+\rho_0}{1-\rho_0} = 1.1513 \times \log_{10}\frac{1+0.92}{1-0.92} = 1.5890.$

Thus, $z = (z - z_0)\sqrt{n-3} = -0.80.$

Since this value of Z does not fall in the critical region, we fail to reject null hypothesis at 5% level of significance, that means, the population correlation co-efficient is not significantly different from the hypothesized value.

Example 16.10.3 A research team was attempting to determine if political risk in countries is related to inflation for these countries. For this purpose a survey of political risk analysis was conducted with the mean political risk score for each of 49 countries. The political risk score is scaled such that the higher the score, the greater the political risk. The sample correlation co-efficient between political risk and inflation for these countries were found as 0.43. Test whether there is a positive linear relationship between the political risk and inflation at 0.5% level of significance.

Solution We want to test: $H_0:\rho=0$ against $H_1:\rho>0$.

We have, $n=45$, $r=0.43$, so $df=49-2=47$.

The test is based on the statistic: $t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} \sim t_{n-2}$.

This is a right tailed test, so the critical region is given by: $t > t_{0.005;47}$, or, $t > 2.704$

The computed value of test statistic is: $t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} = \frac{0.43\sqrt{49-2}}{\sqrt{1-(0.43)^2}} = 3.265.$

Since the observed value of t lies in critical region, we can reject null hypothesis at 0.5% level of significance.

We have strong evidence of a positive linear relationship between inflation and political risk of the countries.

Example 16.10.4 A regression of retail sales on disposable income provides with the following results: $n=22$, $b=0.3815$, $se(b)=0.0253$. Test the significance of regression co-efficient at 1% level of significance. Also, compute 99% confidence interval for regression parameter.

Solution We have to test the hypothesis $H_0:\beta=\beta_0=0$ against $H_1:\beta\neq 0$. Assume that the regression errors ε's are normally distributed and all the assumptions of linear regression model hold, then the test statistic for testing the above mentioned hypothesis is given by

$$t = \frac{b}{se(b)} \text{ which distributed as Student's t with } n-2\,\text{df}.$$

From the given information, we have the computed value of t as

$$t = \frac{b}{se(b)} = (0.3815 - 0)/0.0253 = 15.08.$$

Now, from table of t-distribution, we have for $(n-2)=20\,\text{df}$, $t_{20,\,0.005}=2.845$, hence, the null hypothesis is very clearly rejected, that means, the regression coefficient is highly significant.

Again, the 99% CI is given by $b \pm t_{n-2,\alpha/2} \times se(b) = 0.3815 \pm 2.845 \times 0.0253 = (0.31, 0.45)$

Example 16.10.5 A Courier Express authority recorded the following data regarding distance (x) and the time usage (y) for hand delivery of 15 packets and obtained the following results

$$\Sigma x=243,\ \Sigma y=503,\ \Sigma x^2=3999,\ \Sigma y^2=17181,\ \Sigma xy=8229.$$

i) Compute the correlation between distance and time usage and test whether it is significantly different from zero against an alternative that there is a positive linear relationship between these two variables at 1% level of significance.
ii) Compute the regression co-efficient and test its significance.
iii) Also obtain 99% confidence interval for regression parameter.

Solution (i) We know the correlation co-efficient is given by

$$r=\frac{\Sigma xy-\frac{\Sigma x \Sigma y}{n}}{\sqrt{\left\{\Sigma x^2-\frac{(\Sigma x)^2}{n}\right\}\left\{\Sigma y^2-\frac{(\Sigma y)^2}{n}\right\}}}.$$

Substituting the values from given information, we have

$$\Sigma xy-\frac{\Sigma x \Sigma y}{n}=8229-(243\times 503)/15=80.4$$

$$\Sigma x^2-\frac{(\Sigma x)^2}{n}=3999-(243)^2/15=62.4$$

$$\Sigma y^2-\frac{(\Sigma y)^2}{n}=17181-(503)^2/15=313.73.$$

$$\text{Thus, } r=\frac{\Sigma xy-\frac{\Sigma x \Sigma y}{n}}{\sqrt{\left\{\Sigma x^2-\frac{(\Sigma x)^2}{n}\right\}\left\{\Sigma y^2-\frac{(\Sigma y)^2}{n}\right\}}}=\frac{80.4}{\sqrt{62.4\times 313.73}}=0.57.$$

In order to perform the required test, we have to formulate the following null and alternative hypotheses

$$H_0:\rho=0 \quad \text{against} \quad H_1:\rho>0.$$

The appropriate test statistic is: $t=\dfrac{r\sqrt{n-2}}{\sqrt{1-r^2}}$.

Here the degrees of freedom is: $n-2=15-2=13$.

This is a right tailed test, so the critical region is given by: $t>t_{0.01;13}$ or $t>2.650$ (From table of t-distribution).

The computed value of test statistic is: $t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} = \frac{0.57\sqrt{15-2}}{\sqrt{1-(0.57)^2}} = 2.50.$

Thus, the observed value of t does not fall in the critical region, so we fail to reject null hypothesis at 1% level of significance.

Hence, it can be concluded that the population correlation co-efficient is not significantly different from zero, that means, there is no linear relationship between the variables.

(ii) We know, for the regression line $y = \alpha + \beta x + \varepsilon$, the least squares estimate of β and α are given by

$$b = \frac{\Sigma xy - \frac{\Sigma x \Sigma y}{n}}{\Sigma x^2 - \frac{(\Sigma x)^2}{n}} \quad \text{and} \quad a = \bar{y} - b\bar{x}.$$

Using the values of sum of the products and sum of squares of x from the calculation of correlation co-efficient, we have: $b = \frac{80.4}{62.4} = 1.29$ and

$$a = \bar{y} - b\bar{x} = \frac{503}{15} - 1.29 \times \frac{243}{15} = 35.53 - 1.29 \times 16.20 = 13.64.$$

The fitted regression line is: $y = 13.64 + 1.29x$

In order to test the significance of regression co-efficient, we have to consider the following null and alternative hypothesis

$$H_0 : \beta = 0 \text{ against a two-tailed alternative } H_1 : \beta \neq 0,$$

The appropriate test statistic is: $t = \frac{b - \beta_0}{se(b)} \sim t_{n-2}.$

Here, degrees of freedom is also 13 and since it is a two-tailed test, at 1% level of significance the critical region is given by: $|t| > t_{0.005;13}$ or $|t| > 3.012$

$$\text{Here, } s^2 = \frac{\Sigma(y-\bar{y})^2 - b\Sigma(y-\bar{y})(x-\bar{x})}{n-2} = \frac{313.73 - 1.29 \times 80.4}{13} = 16.15 \text{ and } s = 4.02.$$

$$se(b) = \frac{s}{\sqrt{\Sigma(x-\bar{x})^2}} = \frac{4.02}{\sqrt{624}} = 0.5089.$$

Thus, under null hypothesis, the value of the test statistic is

$$t = \frac{b}{se(b)} = \frac{1.29}{0.5089} = 2.54.$$

The computed value of t does not fall in the critical region, so we fail to reject null hypothesis at 1% level of significance.

Thus, we may conclude that population regression co-efficient is not significantly different from zero.

(iii) 99% confidence interval for β is given by

$$b - t_{n-2,\alpha/2}\,se(b) < \beta < b + t_{n-2,\alpha/2}\,se(b)$$

$$= 1.29 - 3.012 \times 0.6103 < \beta < 1.29 + 3.012 \times 0.6103 = (-0.548,\ 3.128).$$

Example 16.10.6 The fitted regression model for profit (y) on costs (x) for 11 items produced by a company has been found as (the variables are measured in hundred Taka)

$$y = 1922.40 + 1.2865x$$
$$se = (274.9) \qquad (0.0253$$

(where the values in the parenthesis represent the standard error of respective statistic)

The company claims that for every thousand of increase of the cost of their products, they incur an average profit of Taka 1.2 thousand. In the light of the claim of company, test the significance of population regression co-efficient, against a two-tailed alternative at 5% level of significance, find p value of the test and 95% confidence interval for population regression co-efficient.

Solution For testing the significance of regression co-efficient, we have to formulate the null and alternative hypothesis as follows.

$H_0 : \beta = 1.2$ against a two-tailed alternative $H_1 : \beta \neq 1.2$.

The appropriate test statistic is: $t = \dfrac{b - \beta_0}{se(b)} \sim t_{n-2}$.

Here, $n = 11$, so degrees of freedom is $11 - 2 = 9$ and since it is a two-tailed test, at 5% level of significance the critical region is: $|t| > t_{0.025;9}$ or, $|t| > 2.262$

From the output of results we have: $b = 1.2865$, $se(b) = 0.0253$, $\beta_0 = 1.2$.

So, under H_0 the value of: $t = \dfrac{b - \beta_0}{se(b)} = \dfrac{1.2865 - 1.2}{0.0253} = 3.42.$

The computed value of t falls in the critical region, so we fail to accept null hypothesis at 5% level of significance. That means, the regression co-efficient is significant.

Thus, we may conclude that population regression co-efficient is significantly different from 1.2.

- **p-value** From the table of area under t-distribution, we have, for degrees of freedom 9, $P(t < -3.250) = 0.005$, so the $P(t > 3.250) = 0.005$, so the value of $p = 2 \times 0.005 = 0.01$.

Again, 95% confidence interval for β is given by

$$b - t_{n-2,\alpha/2} \times se(b) < \beta < b + t_{n-2,\alpha/2} \times se(b)$$

$$= 1.2865 - 2.262 \times 0.0253 < \beta < 1.2865 + 2.262 \times 0.0253 = (1.2292,\ 1.3437).$$

Example 16.10.7 Suppose the fitted linear regression model for production (y) of certain commodity on electricity consumption (x) for 20 years has been found as

$$y = 42.34 + 0.55x$$
$$se = (2.88) \quad (0.0197.$$

Test the significance of regression co-efficient at 5% level of significance, find 95% CI for β.

Solution For testing the significance of regression co-efficient, let us formulate the null and alternative hypothesis as follows

$$H_0 : \beta = 0 \text{ against a two-tailed alternative } H_1 : \beta \neq 0.$$

The appropriate test is by: $t = \dfrac{b - \beta_0}{se(b)} \sim t_{n-2}$.

Here, $n = 20$, so degrees of freedom is $20 - 2 = 18$ and since it is a two-tailed test, at 5% level of significance the critical region is given by: $|t| > t_{0.025;18}$ or $|t| > 2.101$

From the output of results we have

$$b = 0.55, \ se(b) = 0.0197, \ \beta_0 = 0.$$

So under null hypothesis, the value of

$$t = \frac{b}{se(b)} = \frac{0.55}{0.0197} = 27.92.$$

The computed value of t falls in the critical region and far away from the critical value, so we fail to accept null hypothesis at 5% level of significance. That means, the regression co-efficient is very significant.

95% confidence interval for β is given by

$$b - t_{n-2,\alpha/2} \times se(b) < \beta < b + t_{n-2,\alpha/2} \times se(b)$$

$$= 0.55 - 2.101 \times 0.0197 < \beta < 0.55 + 2.101 \times 0.0197 = (0.5086, \ 0.5914).$$

Example 16.10.8 Twelve secretaries, already working for different periods, at an office were asked to take a special three-day intensive course to improve their keyboard skill. At the beginning and again at the end of the course, they were given a particular two-page letter to type and the improved flawless number of words typed is recorded. The recorded data are shown in the following table.

Table 16.15 Experience and improvement in typing speed

Secretary	Number of Years of Experience	Improvement (words per minute)
A	2	9
B	6	11
C	3	8
D	8	12

E	10	14
F	5	9
G	10	14
H	11	13
I	12	14
J	9	10
K	8	9
L	10	10

i) Compute the product moment correlation co-efficient and test its significant against a positive alternative;

ii) Fit a regression line of improvement on experience and test the significance of regression co-efficient;

iii) Find 95% confidence interval for population regression co-efficient.

Solution (i) We know, the product moment correlation co-efficient is given by

$$r=\frac{\Sigma xy-\frac{\Sigma x\Sigma y}{n}}{\sqrt{\left\{\Sigma x^2-\frac{(\Sigma x)^2}{n}\right\}\left\{\Sigma y^2-\frac{(\Sigma y)^2}{n}\right\}}}.$$

The statistics needed for the calculation of correlation co-efficient are shown in following table.

Table 15.16 Calculating statistics for correlation and regression co-efficient

Secretary	x	y	x^2	y^2	xy
A	2	9	4	81	18
B	6	11	36	121	66
C	3	8	9	64	24
D	8	12	64	144	96
E	10	14	100	196	140
F	5	9	25	81	45
G	10	14	100	196	140
H	11	13	121	169	143
I	12	14	144	196	168
J	9	10	81	100	90
K	8	9	64	81	72
L	10	10	100	100	100
Total	$\Sigma x = 94$	$\Sigma y = 133$	$\Sigma x^2 = 848$	$\Sigma y^2 = 1529$	$\Sigma xy = 1102$

We have:

$$r=\frac{\Sigma xy-\frac{\Sigma x\Sigma y}{n}}{\sqrt{\left\{\Sigma x^2-\frac{(\Sigma x)^2}{n}\right\}\left\{\Sigma y^2-\frac{(\Sigma y)^2}{n}\right\}}}=\frac{1102-94\times 133/12}{\sqrt{\left\{848-\frac{(94)^2}{12}\right\}\left\{1529-\frac{(133)^2}{12}\right\}}}=0.89.$$

For testing the significance of correlation co-efficient, we formulate the null and alternative hypothesis

$$H_0:\rho=0 \text{ against } H_1:\rho>0.$$

The appropriate test statistic is: $t=\dfrac{r\sqrt{n-2}}{\sqrt{1-r^2}}$.

Here the degrees of freedom of t is: $n-2=12-2=10$.

This is a right- tailed test, so at 5% level of significance, the critical region is given by

$$t>t_{0.05;10} \text{ or, } t>1.812 \text{ (From table of t-distribution).}$$

Under null hypothesis, the value of test statistic is

$$t=\frac{r\sqrt{n-2}}{\sqrt{1-r^2}}=\frac{0.89\sqrt{12-2}}{\sqrt{1-(0.89)^2}}=6.25.$$

Since, the observed value of t falls in the critical region, so we fail to accept null hypothesis at 5% level of significance. That means, there is a significant positive correlation between experience and improvement of typing speed.

(ii) The simple linear regression line to be fitted is: $y=a+bx$.

Where, $b=\dfrac{\sum xy-\dfrac{\sum x\sum y}{n}}{\left\{\sum x^2-\dfrac{(\sum x)^2}{n}\right\}}$ and $a=\bar{y}-b\bar{x}$.

Thus, $b=\dfrac{\sum xy-\dfrac{\sum x\sum y}{n}}{\left\{\sum x^2-\dfrac{(\sum x)^2}{n}\right\}}=\dfrac{1102-94\times133/12}{848-(94)^2/12}=0.54$ approx.

and $a = a=\bar{y}-b\bar{x}=133/12-0.54\times94/12=6.85$ approx.

Substituting the computed values of a and b leads to the following fitted regression equation: $\hat{y}=6.85+0.54x$.

For testing the significance of regression co-efficient, let us consider $H_0:\beta=0$ against a two-tailed alternative $H_1:\beta\neq0$.

The appropriate test statistic is: $t=\dfrac{b-\beta_0}{se(b)}\sim t_{n-2}$.

Here, $s^2=\dfrac{\sum(y-\bar{y})^2-b\sum(y-\bar{y})(x-\bar{x})}{n-2}=2.2427$ and $s=1.4976$

And $se(b) = \frac{s}{\sqrt{\Sigma(x-\bar{x})^2}} = \frac{1.4976}{\sqrt{60.3}} = 0.1928.$

Thus, $t = \frac{b-\beta_0}{se(b)} = \frac{0.54-0}{0.1928} = 2.280.$

Here, at 5% level of significance with 10 degrees of freedom, the critical values are $|t| > t_{0.025;10}$ or, $|t| > 2.228$ (from table). The computed value of t falls in the critical region, so we fail to accept null hypothesis at 5 % level of significance.

Thus, we may conclude that population regression co-efficient is significantly different from zero.

(iii) 95% confidence interval for β is given by

$$b - t_{n-2,\alpha/2} \times se(b) < \beta < b + t_{n-2,\alpha/2} \times se(b)$$

$$= 0.54 - 2.228 \times 0.1928 < \beta < 0.54 - 2.228 \times 0.1928 = (0.1104, 0.9695).$$

16.11 Test of Significance of Single Variance (χ^2 Parametric Test)

Let $X_1, X_2, \ldots\ldots, X_n$ be a random sample of size n from normal population with mean μ and variance σ^2, where μ is unknown. We want to test the hypothesis that the population variance is σ_0^2, a specified value of σ^2, against possible three types of alternatives, that means,

i) $H_0 : \sigma^2 = \sigma_0^2$ against $H_A : \sigma^2 \neq \sigma_0^2$ (against a two-tailed alternative);

ii) $H_0 : \sigma^2 = \sigma_0^2$ against $H_A : \sigma^2 > \sigma_0^2$ (against a right-tailed alternative);

iii) $H_0 : \sigma^2 = \sigma_0^2$ against $H_A : \sigma^2 < \sigma_0^2$ (against a left -tailed alternative).

Since μ is unknown, it is estimated by its unbiased estimate $\bar{X} = \frac{\Sigma X_i}{n}$ and variance σ^2 estimated by its unbiased estimate

$$s^2 = \frac{\Sigma(X_i - \bar{X})^2}{n-1} = \frac{1}{n-1}\left[\Sigma X_i^2 - \frac{(\Sigma X_i)^2}{n}\right].$$

Then, under null hypothesis, the test statistic is: $\chi^2 = \frac{(n-1)s^2}{\sigma_0^2}$

- **Decision** In case of two tailed alternative, reject H_0 at α level of significance

$$\text{if } \chi^2_{cal} \leq \chi^2_{1-\alpha/2;n-1} \text{ or, } \chi^2_{cal} \geq \chi^2_{\alpha/2;n-1} .$$

In case of a right-tailed alternative, reject H_0 at α level of significance

$$\text{if } \chi^2_{cal} \geq \chi^2_{\alpha;n-1}$$

and in case of a left-tailed alternative, reject H_0 at α level of significance

$$\text{if } \chi^2_{cal} \le \chi^2_{1-\alpha;n-1}.$$

Example 16.11.1 The following are the weights (in gram) of a randomly selected sample of 11 apples in a shop.

70, 85, 92, 90, 95, 79, 80, 85, 90, 85, 95.

The weight of apples follows normal distribution with mean μ and variance σ^2. Can we conclude that the population variance of apples of the shop is more than 50 gm²?

Solution The null and alternative hypothesis for this test is

$$H_0 : \sigma^2 = 50 \text{ against } H_A : \sigma^2 > 50.$$

Under null hypothesis, the test statistic is

$$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2} \text{ which is distributed as } \chi^2 \text{ with } n-1 \text{ df.}$$

It is a right tailed test, so, the critical region is given by: $\chi^2 \ge \chi^2_{\alpha;n-1}$.

$$\text{Here, } \bar{X} = \frac{\Sigma X_i}{n} = \frac{946}{11} = 86 \text{ and}$$

$$s^2 = \frac{\Sigma(X_i - \bar{X})^2}{n-1} = \frac{1}{n-1}\left[\Sigma X_i^2 - \frac{(\Sigma X_i)^2}{n}\right]$$

$$= \frac{1}{10}\left[81930 - \frac{(946)^2}{11}\right] = 57.4.$$

So, the computed value of χ^2 is: $\chi^2 = \dfrac{(n-1)s^2}{\sigma_0^2} = \dfrac{10 \times 57.4}{50} = 11.48.$

Here, at 5% level of significance, the critical value is $\chi^2_{0.05,10} = 18.307$ which is more than the computed value, so we fail to reject H_0, that means, the variance of the weights of the apples is not more than 50 gm².

Example 16.11.2 The daily duration of telephone calls received by the enquiry department of a small industry for a randomly selected 11 days over a quarter are as follows.

160, 172, 121, 144, 100, 108, 175, 200, 105, 95, 102.

The manager of industry says that the population variance of the daily duration of calls over the quarter is 1500. The authority thinks that it is overestimated. How would you comment on the variance?

Solution The null and alternative hypothesis for this test are

$$H_0 : \sigma^2 = 1500 \text{ against } H_A : \sigma^2 < 1500$$

Under null hypothesis, the test statistic is

$$\chi^2 = \frac{(n-1)s^2}{\sigma_0^{\,2}} \text{ which is distributed as } \chi^2 \text{ with } n-1 \text{ df} = 10 \text{ df.}$$

It is right tailed test, so, the critical region is given by $\chi^2 \le \chi^2_{1-\alpha,n-1}$ where at 5% level of significance, $\chi^2_{0.95,10} = 3.94$.

Here, $\Sigma X_i = 1482$ and

$$s^2 = \frac{\Sigma(X_i - \bar{X})^2}{n-1} = \frac{1}{n-1}\left[\Sigma X_i^{\,2} - \frac{(\Sigma X_i)^2}{n}\right]$$

$$= \frac{1}{10}\left[213340 - \frac{(1482)^2}{11}\right] = 1363.82.$$

So, the computed value of χ^2 is: $\chi^2 = \frac{(n-1)s^2}{\sigma_0^{\,2}} = \frac{10 \times 136382}{1500} = 9.09.$

Here, at 5% level of significance, the critical value is given by $\chi^2_{0.05,\,10} = 3.94$ which is less than the computed value, so we reject H_0, that means, the variance of daily duration of telephone calls is less than 1500. Hence, the authority's thought is beyond doubt.

16.12 Test of Hypothesis about Independence of Two Attributes (χ^2 Non parametric Test)

. Let a sample of size n be drawn from the population of interest and the observed frequencies are cross-classified according to the categories of two attributes. Let us consider two attributes A and B where A is assumed to have r categories and B is assumed to have c categories. The cross-classification can be conveniently displayed by means of a table called $r \times c$ contingency table (read r by c contingency table). A general contingency table of order $r \times c$ is shown in Table 16.17. The frequencies O_{ij} in the cells are termed as observed frequencies and the totals of the frequencies in each row (say, R_i, $i = 1, 2, \ldots, r$) and each column (say, C_j, $j = 1, 2, \ldots, c$) are termed as marginal frequencies.

> **Contingency Table** Data for attributes arranged in two-directional tabular form for the test of independence is called a contingency table. The order of the table is determined by the number of categories of two attributes.

Table 16.17 $r \times c$ Contingency table

		Attribute B								Total
		B_1	B_2	....	....	B_j	....	....	B_c	
Attribute A	A_1	O_{11}	O_{12}	....	....	O_{1j}	....	....	O_{1c}	R_1
	A_2	O_{21}	O_{22}	....	.....	O_{2j}	....	....	O_{2c}	R_2
	.	...	...	....	....	....	....	....	....	.

.	...	...	....	....	....	....	....	....	.
A_i	O_{i1}	O_{i2}	....	....	O_{ij}	....	....	O_{ic}	R_i
.	...	....	....	....	....	....	....	....	.
.	...	....	....	....	....	....	....	....	.
A_r	O_{r1}	O_{r2}	....	....	O_{rj}	....	....	O_{rc}	R_r
Total	C_1	C_2	....	....	C_j	....	....	C_c	N

For testing the independence of two attributes A and B, cross classified in different categories, the null hypothesis is considered as

$$H_0: \text{Two attributes } A \text{ and } B \text{ are independent}$$

against the alternative H_1 : Two attributes A and B are associated or dependent.

Under null hypothesis, the test statistic for testing the hypothesis is given by

$$\chi^2 = \sum \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \sim \chi^2_{(r-1)(c-1)}$$

where, O_{ij} and E_{ij} are the observed and expected frequencies corresponding to the (i, j)th cell of contingency table. Expected cell frequencies are computed according to the multiplicative rule of probability. If two events are independent, the probability of their joint occurrence is equal to the product of their individual probabilities. Applying this rule to a contingency table, it is equivalent to say that, if two criteria of classification are independent, a joint probability is equal to the product of the two corresponding marginal probabilities. Thus, the expected cell frequencies are computed by the formula

$$E_{ij} = \frac{R_i}{N} \times \frac{C_j}{N} \times N = \frac{R_i \times C_j}{N}.$$

Note that total expected frequencies should be equal to the total observed frequency, so while computing the expected frequencies it is required to adjust the expected frequencies for $rc-(r-1)\times(c-1)$ cells, that is why for this test the degree of freedom is $(r-1)\times(c-1)$ that means for $(r-1)\times(c-1)$ cells expected frequencies are calculated independently.

It is one-tailed test (right-tailed) test. So, if the calculated value of χ^2 is less than the table value of χ^2 at a specific level of significance with $(r-1)\times(c-1)$ degrees of freedom, the null hypothesis holds true, that means, the two attributes are independent. If calculated value of χ^2 is greater than the table value, the null hypothesis is rejected, that means the two attributes are associated or dependent.

Observed Frequency The frequencies obtained by observation. These are the sample frequencies.

Expected Frequency The frequency corresponding to a particular cell obtained by dividing the product of row total and column total passing through that cell by the total frequencies.

Example 16.12.1 A tobacco company claims that there is no relationship between smoking and lung ailments. To investigate the claims a random sample of 300 males in the age group 40-50 are given medical test. The observed sample results are shown below.

Table 16.18 Number of males according to smoking and lung ailment

	Found lung ailment	No lung ailment	Total
Smokers	75	105	180
Non-smokers	25	95	120
Total	100	200	300

On the basis of the information, can it be concluded that smoking and lung ailments are independent?

Solution Let us consider the hypothesis that smoking and lung ailments are independent, that means

H_0 : Smoking and lung ailments are independent;

or, there is no effect of smoking on lung ailments.

Against, H_1 : They are not independent, or smoking causes lung ailments.

The test statistic for testing the hypothesis is: $\chi^2 = \sum \frac{(O_{ij} - E_{ij})^2}{E} \sim \chi^2_{(r-1)(c-1)}$.

It is a 2×2 table, so the degree of freedom is 1, we have to calculate expected frequency for a cell independently, and others are to be obtained by adjustment so that the total marginal frequencies remain the same.

Again, let $\alpha = 0.05$, then the critical value of χ^2 with 1 df is 3.84 (which is also sometimes written as: $\chi^2_{0.05;1} = 3.84$).

Expected frequency for the cell corresponding to first row and first column (E_{11}) is computed as

$$E_{11} = \frac{R_1 \times C_1}{N} = \frac{180 \times 100}{300} = 60 .$$

So the expected frequencies for the remaining cells are

$$E_{12} = R_1 - E_{11} = 180 - 60 = 120, \quad E_{21} = C_1 - E_{11} = 100 - 60 = 40,$$

and $E_{22} = C_2 - E_{12} = 200 - 120 = 80$.

E_{22} can also be computed using R_2 as: $E_{22} = R_2 - E_{21} = 120 - 40 = 80$.

Arranging the observed and expected frequencies in the following table, we can easily compute the necessary columns for χ^2.

Table 16.19 Computation of χ^2

O	E	$(O-E)^2$	$(O-E)^2/E$

75	60	225	3.75
105	120	225	1.875
25	40	225	5.625
95	80	225	2.8125
Total			14.0625

Here the observed value of chi-squares is much more than the critical values, therefore, the null hypothesis is rejected at 5% level of significance. That means, it is evident that smoking has significant effect on lung ailments.

Example 16.12.2 A certain drug is claimed to be effective in curing cold. In an experiment on 500 persons suffering from cold, half of them were given the drug and half of them were given the sugar pills. The reaction to the treatment on patients is recorded as in the following table.

Table 16.20 Reaction to the treatment

	Helped	Harmed	No effect	Total
Drug	150	30	70	250
Sugar pills	130	40	80	250
Total	280	70	150	500

On the basis of the information can it be concluded that there is a significance difference in the effect of the drug and sugar pills.

Solution Let us take the null hypothesis that there is no difference in the drug and sugar pills as far as their effect on curing cold is concerned.

The test statistic is: $\chi^2 = \sum \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \sim \chi^2_{(r-1)(c-1)}$

Since it is a 2×3 table, the degrees of freedom would be $(2-1)\times(3-1)=2$, that means, we will have to calculate only two expected frequencies independently and other four can be calculated or determined by adjustment.

Expected frequencies are computed as follows.

$$E_{11} = \frac{R_1 \times C_1}{N} = \frac{250 \times 280}{500} = 140, \quad E_{12} = \frac{R_1 \times C_2}{N} = \frac{250 \times 70}{500} = 35.$$

Thus the table for the computation of expected frequencies is shown below.

Table 16.21 Computation of expected frequencies

	Helped	Harmed	No effect	Total
Drug	140	35	75	250
Sugar pills	140	35	75	250
Total	280	70	150	500

Arranging the observed and expected frequencies in the following table, we can easily compute the necessary columns for χ^2.

Table 16.22 Computation of χ^2

O	E	$(O-E)^2$	$(O-E)^2/E$
150	140	100	0.714
130	140	100	0.714
30	35	25	0.714
40	35	25	0.714
70	75	25	0.333
80	75	25	0.333
	Total		3.522

The calculated value of χ^2 is: 3.522.

The critical value of χ^2 for 2 df at 5% level of significance is 5.99 which is larger than the calculated value. Therefore we fail to reject the null hypothesis that means we can conclude that the drug and sugar pills do not make any significant difference in curing cold.

Example 16.12.3 In a survey on the daily production of a garments factory, the manager finds that 315 of the garments are of category A, 101 are of category B, 108 are of category C and 32 are of category D. According to the manager's long experience, he expects that the numbers of different categories of garments be in proportion 9 : 3 : 3 : 1. From the given information, can we say at 5% level of significance that the manager is correct?

Solution Here, the null hypothesis is

H_0: There is no significant difference between the observed and expected number of garments of different categories.

The appropriate test statistic for testing the hypothesis is

$$\chi^2 = \sum \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \sim \chi^2_{(r-1)(c-1)}.$$

The expected frequencies in this case are to be calculated as proportional to the expected ratio. The total number of garments observed $=315+101+108+32=556$ and the sum of ratios $=9+3+3+1=16$.

So, the expected number of garments different categories are as follows.

$$E_A = \frac{556 \times 9}{16} = 312.75, \quad E_B = E_C = \frac{556 \times 3}{16} = 104.25,$$

and, $E_D = 556-(E_A - E_B - E_C) = 34.75$

Computation of the value of χ^2 is shown in following table.

Table 16.23. Calculation of χ^2

Category	Observed number O	Expected number E	$(O-E)^2$	$(O-E)^2/E$

A	315	312.75	5.062	0.016
B	101	104.25	10.562	0.101
C	108	104.25	14.062	0.135
D	32	34.75	7.562	0.218
Total				0.470

The computed value of χ^2 is: 0.470.

Here, the degrees of freedom is $4-1=3$, and the critical value of χ^2 at 5% level of significance with 3 degrees freedom is 7.82 (from table).

Hence, the we fail to reject null hypothesis at 5% level of significance, hence, it can be concluded that there is no doubt about the manager's expectation.

16.13 Power of a Test

We know, the power of a test is the probability of rejecting null hypothesis when it is false. Let us consider a sample of size n from normal population with mean μ and known variance σ^2 and suppose we are interested to test the hypothesis

$H_0: \mu = \mu_0$, against the alternative $H_1: \mu > \mu_0$ at α level of significance.

We know, the test statistic to be used for testing the null hypothesis is given by

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} \sim N(0, 1)$$

and the decision rule is that Reject

$$H_0: \text{ if } \quad z > z_\alpha \quad \text{or} \quad z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} > z_\alpha \quad \text{or,} \quad \bar{x} > \bar{x}_c = \mu_0 + z_\alpha \sigma/\sqrt{n}.$$

Here, $\bar{x}_c = \mu_0 + z_\alpha \sigma/\sqrt{n}$ determine the values of sample mean that leads in rejecting the null hypothesis, which otherwise proves that the null hypothesis is false.

- **Steps in determination of power of a population mean test** The following steps are involved in computation of power of a test.

(i) Determine the value of: $\bar{x}_c = \mu_0 + z_\alpha \sigma/\sqrt{n}$.

(ii) Find the probability that the sample mean will be in the acceptance region for given false null hypothesis, which will give the probability of Type II error, β. At this step, a value of $\mu = \mu^*$ is considered such that $\mu^* > \mu_0$ (for a right tailed test) and β is computed as

$$\beta = P(\bar{x} \le \bar{x}_c \mid \mu = \mu^*) = P\left(z < \frac{\bar{x}_c - \mu^*}{\sigma/\sqrt{n}}\right).$$

(iii) Compute power of the test as : Power $= 1 - \beta$.

Power for the tests of other parameters can also be derived using the steps same as population mean.

The value of β and power of the test will be different for different values of μ^*. If powers of a test are plotted against different values of μ^* and a smooth curve is drawn, the curve which will be obtained is known as power curve.

Example 16.13.1 (Power for mean test) Suppose a ball bearing company claims that the average weight of its product is 5 ounces. The population distribution of weights is assumed to be normally distributed with standard deviation 0.1 ounce. An interested person drew a random sample of 16 observations to test whether the average weight is more than 5 ounces at 5% level of significance. Find power of this test.

Solution Here, $H_0: \mu = 5$, against the alternative $H_1: \mu > 5$.

The value of the test statistic for testing the hypothesis is given by

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}.$$

At $\alpha = 0.05$, $z_\alpha = 1.645$ hence the decision rule is

$$\text{Reject } H_0 \text{ if } \frac{\bar{x} - 5}{0.1/\sqrt{16}} > 1.645 \text{ or, } \bar{x} > 5 + 1.645 \times 0.1 \times 4 = 5.041 = \bar{x}_c.$$

Now if the sample mean is less than or equal to 5.041, then according to the rule, we will fail to reject null hypothesis.

Suppose, we want to determine the probability that the null hypothesis will not be rejected if the true mean is greater than 5.041 ounces, say 5.05 ounces. Thus the null hypothesis is wrong and the alternative is correct.

At this step, we want to determine the probability that we will fail to reject null hypothesis if $\mu = 5.05$, which is the probability of an incorrect decision and we will obtain β, as

$$\beta = P(X \le \bar{x}_c \mid \mu = 5.05) = P\left(Z \le \frac{5.041 - 5.05}{0.1/\sqrt{16}}\right)$$

$$= P(Z \le -0.36) = 1 - 0.6406 = 0.3594.$$

This means, when the population mean is 5.05, the value of β is 0.3594.

Finally, power is computed using the relationship

$$\text{Power} = 1 - \beta = 1 - 0.3594 = 0.6406$$

In the same way different values of β and power can be generated.

Example 16.13.2 (Power for proportion test) Suppose the authority of a big firm claims that they follow the gender equality convention of UN in the employment of workers in their firm. A random sample of 600 workers was obtained from the firm, and found that number of female workers was 382. Using a significance level of $\alpha = 0.05$, test the claim of authority and find the power of test.

Solution The null hypothesis and alternative hypothesis are respectively

$$H_0 : \pi = \pi_0 = 0.50,$$

Against the alternative $H_1 : \pi \neq 0.50 = \pi_1$ (where π_1 is any value of π other than π_0)

The appropriate test statistic for testing the hypothesis is given by

$$Z = \frac{P - \pi_0}{\sqrt{\frac{\pi_0(1-\pi_0)}{n}}}$$

where P is the estimated value of proportion from sample.

The decision rule is

$$\text{Reject } H_0 \text{ if } \quad \frac{P - \pi_0}{\sqrt{\frac{\pi_0(1-\pi_0)}{n}}} > z_{\alpha/2} \quad \text{or,} \quad \frac{P - \pi_0}{\sqrt{\frac{\pi_0(1-\pi_0)}{n}}} < -z_{\alpha/2}.$$

Here, $\alpha = 0.05$, $n = 600$, substituting the values, we have

$$\frac{P - 0.50}{\sqrt{\frac{0.50(1-0.50)}{600}}} > 1.96 \quad \text{or,} \quad \frac{P - 0.50}{\sqrt{\frac{0.50(1-0.50)}{600}}} < -1.96.$$

Hence, the decision rules becomes

$$\text{reject } H_0, \text{ if } \; P > 0.50 + 1.96 \times \sqrt{\frac{0.50(1-0.50)}{600}} = 0.50 + 0.04 = 0.54;$$

$$\text{or, } \; P < 0.50 - 1.96 \times \sqrt{\frac{0.50(1-0.50)}{600}} = 0.50 - 0.04 = 0.46.$$

The observed value of π is $P = 382/600 = 0.637$, which is greater than 0.54, the upper critical value, so we fail to accept null hypothesis. That means, the authority's claim is not justified.

Since the test proves that the null hypothesis is not true, let us suppose that the true value is $\pi_1 = 0.55$, thus for the probability of Type II error, β, we have to find the probability that the sample proportion is between 0.46 and 0.54, if the population proportion is 0.55. This probability is given by

$$\beta = P(0.46 \leq P \leq 0.54 \mid P = 0.55)$$

$$= P\left[\frac{0.46-\pi_1}{\sqrt{\frac{\pi_1(1-\pi_1)}{n}}} \le \frac{P-\pi_1}{\sqrt{\frac{\frac{\pi_1(1-\pi_1)}{n}}{n}}} \le \frac{0.0.54-\pi_1}{\sqrt{\frac{\pi_1(1-\pi_1)}{n}}}\right]$$

$$= P\left[\frac{0.46-0.55}{\sqrt{\frac{0.55(1-0.55)}{600}}} \le Z \le \frac{0.54-0.55}{\sqrt{\frac{0.55(1-0.55)}{600}}}\right]$$

$$= P(-1.43 \le Z \le -0.49) = 0.3121$$

Hence, $\text{Power} = 1-\beta = 1-0.3121 = 0.6879$

This probability β or power of the test can be calculated for any proportion P not equal to 0.50.

Grpup-A

16.14 Short questions and their answers

1. What is test of hypothesis?

   Ans. The process that enables a decision maker to draw an inference about population characteristic by analyzing the difference between the value obtained from sample and the hypothesized value of parameter is called test of hypothesis.

2. What is hypothesis?

   Ans. Any statement about any phenomenon is called hypothesis.

3. What is statistical hypothesis?

   Ans. A statistical hypothesis is a statement about population characteristic which can be tested on the basis of the sample data.

4. What is a null hypothesis?

   Ans. According to R.A. Fisher, a null hypothesis is the hypothesis which is tested for possible rejection under the assumption that it is true/Or A claim about a population parameter that is assumed to be true until proven otherwise.

5. What is an alternative hypothesis?

   Ans. A claim about a population parameter that will be true if the null hypothesis is false.

6. What is simple hypothesis?

   Ans. The hypothesis which completely specifies all the parameters of the related population is called simple hypothesis.

7. What is composite hypothesis?

   Ans. The hypothesis which does not completely specify all the parameters is called composite hypothesis.

8. What is type one error or first kind of error?

Ans. An error that occurswhen a true hypothesis is rejected is called type one erroe or first kind of error.

9. What is type two error or second kind of error?

Ans.An error that occurs than a false hypothesis is accepted is called type two error or second kind of error.

10. What is a level of significance or size of a test? What is α ?

Ans. The probability of first kind of error is called level of significance or size of a test. It is denoted by α . Actually, it is the probability of rejecting the null hypothesis when it is true.

11. What is β?

Ans. The probability of type two error is denoted by β.

12. What is the power of a test?

Ans. It is the probability of rejecting the null hypothesis when is false. It is denoted by $1-\beta$.

13. What is critical value or critical point?

Ans. One or two values that divide the whole region under the sampling distribution of the sample statistic into rejection and norejection regions is called critical value or values.

14. What is one tailed test?

Ans. A test in which there is only one rejection region either in the left tail or in the right tail of the distribution curve is called one tailed test.

15. What is two tailed test?

Ans. A test in which there are two rejection regions in each tail of the distribution curve is called two tailed test.

16. What is left tailed test?

Ans. A test in which the rejection region lies in the left tail of the distribution curve is called left tailed test.

17. What is right tailed test?

Ans. A test in which the rejection region lies in the right tail of the distribution curve is called right tailed test.

18. What is p-value? p value means probability value.

Ans. The smallest significance level at which a null hypothesis can be rejected is called p-value. If the p-value is less 0.05 but greater than 0.01, then the hypothesis is rejected at 5% level of signifacnce. In this case the calculated value of the test statistic is significant. But if the calculated value of the test statistic is less then 0.01, then the hypothesis is rejected at 1% level of significance and the calculated value of the test statistic is highly significant.

19. What is test statistic?

Ans. The statistic which is used to test the statistical hypothesis is called test statistic. The impotant parametric test statistics are i) Z-test or normal test, ii) t-test, iii) F-test, and iv) Chi-Square test.

Grpup-B & C