Chapter 10

If there is an opportunity to make a mistake, sooner or later the mistake will be made.
—Edmond C. Berekely

Hypothesis Testing

LEARNING OBJECTIVES

After studying this chapter, you should be able to

- explain why hypothesis testing is important.
- know how to establish null and alternative hypotheses about a population parameter.
- develop hypothesis testing methodology for accepting or rejecting null hypothesis.
- use the test statistic z, t and F to test the validly of a claim or assertion about the true value of any population parameter.
- understand Type I and Type II errors and its implications in making a decision.
- compute and interpret p-values.
- interpret the confidence level, the significance level and the power of a test.

10.1 INTRODUCTION

Three techniques of inferential statistics namely, (i) a point estimate, (ii) confidence interval that is likely to contain the true parameter value, and (iii) degree of confidence associated with a parameter value which lies within an interval values, as discussed in Chapter 9, helps decision makers in determining an interval estimate of a population parameters value with the help of sample statistic along with certain level of confidence of the interval containing those values.

Another method of estimating the true value of population parameters is to test the validity of the claim (assertion, belief, assumption or statement), also called *hypothesis,* made about this true value using sample statistics.

10.2 HYPOTHESIS AND HYPOTHESIS TESTING

A *statistical hypothesis* (assertion, statement, belief or assumption) about an unknown population parameter value is tested and analysed based on sample data. On the basis of such a test, the hypothesized value of the population parameter is either accepted or rejected. *The process that enables a decision maker to test the validity (or significance) of his claim by analysing the difference between the value of sample statistic and the corresponding hypothesized population parameter value is called hypothesis testing*. Few examples of hypothesis are

- A judge assumes that a person charged with a crime is innocent and subject this assumption (hypothesis) to verification by reviewing the evidence and hearing testimony before reaching to a verdict.
- A pharmaceutical company claims the efficacy of a medicine against a disease that 95 per cent of all persons suffering from the said disease get cured.
- An investment company claims that the average return across all its investments is 20 per cent, and so on.

10.2.1 Formats of Hypothesis

As stated earlier, a hypothesis is a statement to be tested about the true value of population parameters using sample statistics. To examine whether any significant difference exists or not between true value of population parameters and sample statistics, a hypothesis can be stated in the form of *If…, then* statement. Consider, for example, the following statements:

- If inflation rate has decreased, then wholesale price index will also decrease.
- If employees are healthy, then they may take sick leave less frequently.

If terms such as 'positive', 'negative', 'more than' and 'less than' are used to make a statement, then such a hypothesis is called *directional hypothesis and* indicates the direction of the relationship between two or more populations under the study with respect to a parameter value as illustrated below:

- Side effects of particular medicine were experienced by less than 20 per cent of people.
- Greater stress causes lower job satisfaction to employees in any organization.

The *non-directional hypothesis* indicates the relationship but does not indicate the direction of relationship. In other words, there may be a significant relationship between two populations with respect to a parameter, even than nothing can be said whether the relationship would be positive or negative. Similarly, even if two populations differ with respect to a parameter, nothing can be said whether parameter value of any population will be more or less. The following examples illustrate non-directional hypotheses.

- The relationship between age and sick leaves.
- The difference between average pulse rates of men and women.

10.3 RATIONALE FOR HYPOTHESIS TESTING

Hypothesis Testing: The process of testing a statement or belief about a population parameter by the use of information collected from a sample(s).

If a hypothesis claim or assumption is made about the specific value of population parameter, then it is expected that the corresponding sample statistic is close to the hypothesized parameter value. It is possible only when hypothesized parameter value is reasonably close to true value of parameter and the sample statistic turns out to be the good estimator of the parameter value. This approach to test a hypothesis is called a *test statistic*.

Since sample statistics are random variables therefore their sampling distributions show the tendency of variation. Consequently, sample statistic value is not expected to be exactly equal to the hypothesized parameter value. The difference, if any, is due to chance and/or sampling error. But if the value of the sample statistic differs significantly from the hypothesized parameter value, then hypothesized parameter value may not be correct and increases the doubt about the correctness of the hypothesis.

In statistical analysis, difference between the value of sample statistic and hypothesized parameter is specified in terms of the probability whether the particular level of difference is significant or not. The probability that a particular level of difference exists by chance can be calculated from the known sampling distribution of the test statistic.

The probability with which a decision maker concludes that observed difference between the value of the test statistic and hypothesized parameter value cannot be due to chance is called the ***level of significance*** of the test.

10.4 GENERAL PROCEDURE FOR HYPOTHESIS TESTING

As mentioned before to test the validity of the hypothesis about the population parameter, a sample is drawn from the population and analysed. The results of the analysis are used to decide whether the claim is true or not. The general procedure for testing a hypothesis is summarized below:

Step 1: State the Null Hypothesis and Alternative Hypothesis

The ***null hypothesis***, H_0, represents the claim or statement made about the value or range of values of the population parameters. The capital letter H stands for hypothesis and the subscript '*zero*' implies '*no difference*' between sample statistic and the parameter value. Thus, hypothesis testing requires that the null hypothesis be considered *true* (*status quo* or *no difference*) until it is proved false on the basis of results derived from the sample data. The null hypothesis is expressed using mathematical sign (≤, =, ≥) to make a claim about specific value of the population parameter as follows:

Null Hypothesis: The hypothesis which is initially assumed to be true, although it may in fact be either true or false based on the sample data.

$$H_0 : \mu\ (\leq, =, \geq)\ \mu_0$$

where μ is population mean and μ_0 represents a hypothesized value of μ. Only one sign out of ≤ = and ≥ will appear at a time when stating the null hypothesis

An ***alternative hypothesis***, H_1, is the counter claim or statement made against the value or range of values of the population parameter. Thus, an alternative hypothesis represents the claim that specific value or range of values of the population parameters is not equal to the value claimed in the null hypothesis and is written as

Alternative Hypothesis: The hypothesis concluded to be true if the null hypothesis is rejected.

$$H_1 : \mu \neq \mu_0$$

Consequently, $H_1 : \mu < \mu_0$ or $H_1 : \mu > \mu_0$

Each of the following statements is an example of a null hypothesis and alternative hypothesis

• $H_0 : \mu = \mu_0$;	$H_1 : \mu \neq \mu_0$
• $H_0 : \mu \leq \mu_0$;	$H_1 : \mu > \mu_0$
• $H_0 : \mu \geq \mu_0$;	$H_1 : \mu < \mu_0$

Few examples of directional and non-directional hypothesis are as follows:

Directional Hypothesis

H_0 : The average pulse rates of men and women are same.

H_1 : Men have lower average pulse rates than women.

Non-directional Hypothesis

H_0 : The verbal abilities of men and women are same.

H_1 : The verbal abilities of men and women are different.

Step 2: State the Level of Significance, α (alpha)

The level of significance, usually denoted by α (alpha), is specified before the samples are drawn, so that the results obtained should not influence the choice of the decision maker. The level of significance is specified in terms of the likelihood (probability) of rejecting a null hypothesis when it is true, i.e. it is *the risk that a decision maker takes of rejecting the null hypothesis when it is really true*. Usually, $\alpha = 0.05$ is considered for consumer research projects, $\alpha = 0.01$, for quality assurance and $\alpha = 0.10$ for political polling.

Step 3: Establish Critical or Rejection Region

The area under the sampling distribution curve of the test statistic is divided into two mutually exclusive regions (areas) as shown in Fig. 10.1. These regions are called the *acceptance region* and the *rejection* (or *critical*) *region*.

The acceptance region represents the *range of values* of the sample statistic spread around the *null hypothesized population parameter*. If values of the sample statistic fall within the range (limits) of acceptance region, the null hypothesis is accepted, otherwise it is rejected.

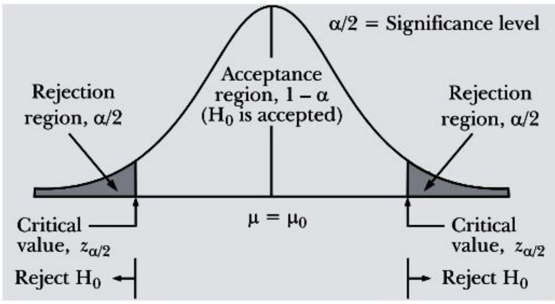


Figure 10.1
Areas of Acceptance and Rejection of H_0 (Two-tailed Test)

The **rejection region** represents the values of the sample statistic that fall outside the limits of the acceptance region. Consequently, null hypothesis is rejected. The size of the rejection region is related to the level of precision to make decisions about a population parameter.

Rejection Region: The range of values that will lead to the rejection of a null hypothesis.

The value of the sample statistic that separates the regions of acceptance and rejection is called **critical value**, i.e. the *level of significance α is used as the cut-off point which separates the area of acceptance from the area of rejection.*

Critical Value: A table value with which a test statistic is compared to determine whether a null hypothesis should be rejected or not.

Decision Rules

- If probability (H_0 is true) $\leq \alpha$, then reject H_0.
- If probability (H_0 is true) $> \alpha$, then accept H_0.

Step 4: Select the Suitable Test of Significance or Test Statistic

The tests of significance or test statistic are classified into two categories: ***parametric and non-parametric tests***. The data for parametric tests are derived from interval and ratio measurements. Whereas data for non-parametric tests are derived from nominal and ordinal measurements. Certain assumptions for parametric tests of significance are as follows:

(i) Each element (or member) of the population should have equal chance of being selected every time a sample is drawn from the population.
(ii) The samples should be drawn from a normally distributed population.
(iii) Normally distributed populations should have equal variances.

Non-parametric tests of significance do not specify normally distributed populations or equality of variances.

Selection of Tests of Significance

The following three factors are considered for choosing a particular test of significance:

(i) Number of samples: one, two or *k* samples?
(ii) Samples are independent or not?
(iii) Measurement scale chosen: nominal, ordinal, interval or ratio?

Besides above information, investigator should also know (i) sample size, (ii) number of samples and (iii) whether data have been weighted.

The value of test statistic is calculated from the distribution of sample statistic by using the following formula:

$$\text{Test statistic} = \frac{\text{Value of sample statistic} - \text{Value of hypothesized population parameter}}{\text{Standard error of the sample statistic}}$$

The choice of the nature of probability distribution of a sample statistic is guided by the sample size n and the value of population standard deviation σ as shown in Table 10.1 and Fig. 10.2.

Table 10.1: Choice of Probability Distribution

Sample Size n	*Population Standard Deviation σ*	
	Known	*Unknown*
• $n > 30$	Normal distribution	Normal distribution
• $n \leq 30$, population being assumed normal	Normal distribution	t-distribution

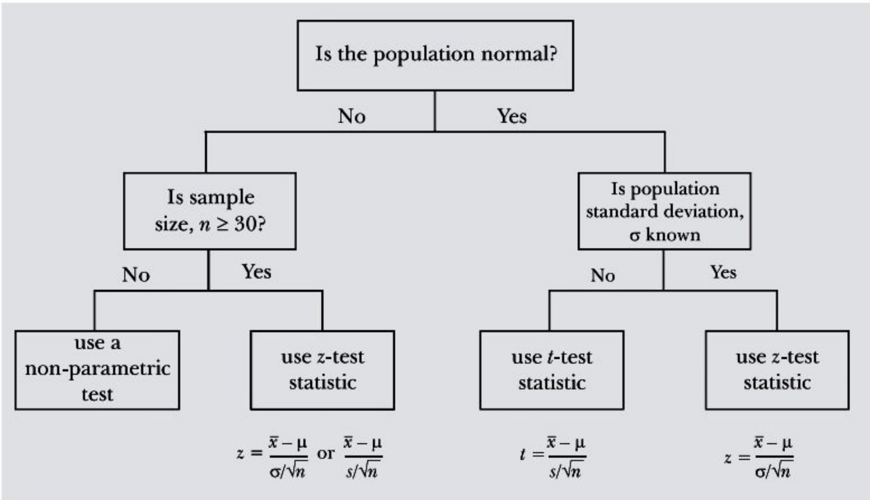


Figure 10.2
Choice of the Test Statistic

Step 5: Formulate a Decision Rule to Accept Null Hypothesis

Compare the calculated value of the test statistic with its critical value (also called *standard table value* of test statistic). The decision rules for null hypothesis are as follows:

- If the test statistic value falls within the area of acceptance, then accept the null, hypothesis, H_0
- Otherwise reject, H_0.

In other words, if the calculated absolute value of a test statistic is less than or equal to its critical (or table) value, then accept the null hypothesis otherwise reject it.

10.5 DIRECTION OF THE HYPOTHESIS TEST

The location of rejection region (or area) under the sampling distribution curve determines the direction of the hypothesis test, i.e. either lower tailed or upper tailed of the sampling distribution. It indicates the range of sample statistic values that would lead to a rejection of the null hypothesis. Figures 10.1, 10.3(a) and 10.3(b) illustrate the acceptance region and rejection region for a null hypothesized population mean, μ, by three different ways of formulating the null hypothesis.

(i) Null hypothesis and alternative hypothesis stated as

$$H_0 : \mu = \mu_0 \quad \text{and} \quad H_1 : \mu \neq \mu_0$$

imply that the sample statistic values which are either significantly smaller than or greater than the null hypothesized population mean, μ_0, value will lead to rejection of the null hypothesis. Hence, it is necessary to keep the rejection region at 'both tails' of the sampling distribution of the sample statistic. This type of test is called *two-tailed test* or *non-directional test* as shown in Fig. 10.1. If the significance level for the test is α per cent, then rejection region equal to $\alpha/2$ per cent is kept in each tail of the sampling distribution.

(ii) Null hypothesis and alternative hypothesis stated as

$$H_0 : \mu \leq \mu_0 \text{ and } H_1 : \mu > \mu_0 \text{ (Right-tailed test)}$$

or

$$H_0 : \mu \geq \mu_0 \text{ and } H_1 : \mu < \mu_0 \text{ (Left-tailed test)}$$

One-tailed Test: The test of a null hypothesis which can only be rejected when the sample test statistic value is in one extreme end of the sampling distribution.

imply that the value of sample statistic is either 'higher than ' or 'lower than ' than the hypothesized population mean, μ_0, value. This leads to the rejection of null hypothesis for significant difference from the specified value μ_0 in one direction (or tail) of the sampling distribution. Thus, the entire rejection region corresponding to the level of significance, α per cent, lies only in one tail of the sampling distribution of the sample statistic as shown in Figs 10.3(a) and (b). This type of test is called **one-tailed test** *or directional test.*

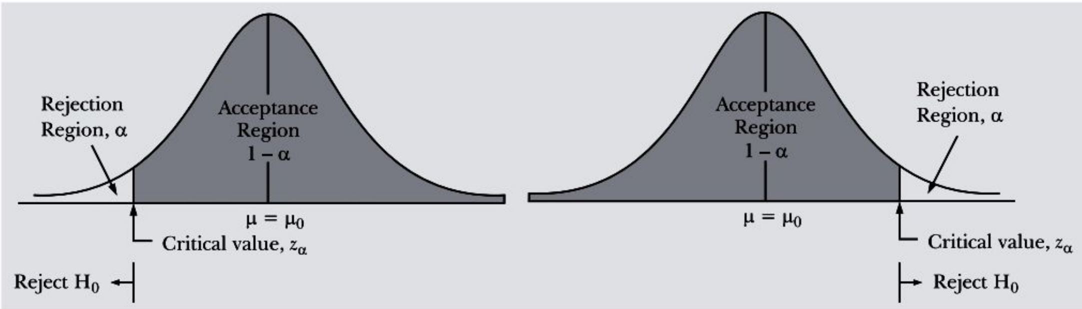


Fig. 10.3 (a) $H_0: \mu \geq \mu_0; H_1: \mu < \mu_0$, Left-tailed Test **Fig. 10.3 (b)** $H_0: \mu \leq \mu_0; H_1: \mu > \mu_0$, Right-tailed Test

10.6 ERRORS IN HYPOTHESIS TESTING

Type I Error: The probability of rejecting a true null hypothesis.

Since decision to reject or accept a null hypothesis is based on sample data, there is a possibility of an incorrect decision or error. A decision maker may commit two types of errors while testing a null hypothesis as shown in Table 10.2.

Table 10.2 Errors in Hypothesis Testing

Decision	*State of Nature*	
	H_0 is True	*H_0 is False*
Accept H_0	Correct decision with confidence $(1-\alpha)$	Type II error (β)
Reject H_0	Type I error (α)	Correct decision $(1 - \beta)$

Level of Significance: The probability of rejecting a true null hypothesis due to sampling error.

Type I Error: This is the probability of rejecting the null hypothesis when it is true and an alternative hypothesis is wrong. The probability of making a **Type I error** is denoted by the symbol α. It is represented by the area under the sampling distribution curve over the region of rejection.

The probability of making a Type I error is referred to as the **level of significance.** The probability of this error is decided by the decision maker before the hypothesis test is performed and based on his tolerance in terms of risk of rejecting the true null hypothesis. The risk of making Type I error usually depends on the cost and/or goodwill loss. The complement $(1 - \alpha)$ of the probability of Type I error measures the probability of not rejecting a true null hypothesis and is referred to as ***confidence level***.

Type II Error: This is the *probability of accepting the null hypothesis when it is false* and an alternative hypothesis is true. The probability of making a **Type II error** is denoted by the symbol β.

The probability of making a Type II error varies with the actual values of the population parameter being tested when null hypothesis H_0 is false. The probability of committing a Type II error depends on five factors: (i) the actual value of the population parameter being tested, (ii) the level of significance selected, (iii) type of test (one or two tailed test) used to evaluate the null hypothesis, (iv) the sample standard deviation (also called standard error), and (v) the size of sample.

A summary of certain critical values at various significance levels for test statistic z is given in Table 10.3.

Table 10.3 Summary of Certain Critical Values for Sample Statistic z

Rejection Region	*Level of Significance, α per cent*			
	$\alpha = 0.10$	$\alpha = 0.05$	0.01	$\alpha = 0.005$
One-tailed region	± 1.28	± 1.645	± 2.33	± 2.58
Two-tailed region	± 1.645	± 1.96	± 2.58	± 2.81

Type II Error: The probability of accepting a false null hypothesis.

10.6.1 Power of a Statistical Test

Another way of evaluating the goodness of a statistical test is to look at the complement of Type II error, which is stated as

$$1 - \beta = P(\text{reject } H_0 \text{ when } H_1 \text{ is true})$$

The complement $1 - \beta$ of β (probability of Type-II error) referred to as **power of a statistical test** measures the probability of rejecting H_0 when it is true.

Illustration Suppose null and alternative hypotheses are stated as

$$H_0 : \mu = 80 \text{ and } H_1 : \mu = 80$$

For all possible values of the population mean, μ, the probability of committing Type II error is required to be calculated.

Suppose a sample of size $n = 50$ is drawn from a population to compute the probability of committing Type II error for a specific value of the population mean, μ. Let sample mean so obtained be $\bar{x} = 71$ with a standard deviation, $s = 21$. For significance level, $\alpha = 0.05$ and a two-tailed test, the table value of $z_{0.05} = \pm 1.96$. But calculated value from sample data is

$$z = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}} = \frac{71 - 80}{21/\sqrt{50}} = -3.03$$

Since absolute value, $z_{cal} = -3.03$, is more than its critical $z_{0.05} = \pm 1.96$, the null hypothesis H_0 is rejected. The rejection of null hypothesis leads to either make a correct decision or commit a Type II error. If the population mean is $\mu = 75$ instead of $\mu = 80$, then probability of committing Type II error is determined by computing a critical region for the sample mean, $\bar{x}_c$. This value is used as the cutoff point between the area of acceptance and the area of rejection. If any other sample mean is less than (or greater than for right-tail rejection region), $\bar{x}_c$, then null hypothesis is rejected. Solving for the critical value of mean gives

Power of a Test: The ability (probability) of a test to reject the null hypothesis when it is false.

$$z_c = \frac{\bar{x}_c - \mu}{\sigma_{\bar{x}}} \quad \text{or} \quad \pm 1.96 = \frac{\bar{x}_c - 80}{21/\sqrt{50}}$$

$$\bar{x}_c = 80 \pm 5.82 \quad \text{or} \quad 74.18 \text{ to } 85.82$$

If critical value falls in the range $\bar{x}_c = 74.18$ to 85.82, then probability of accepting the false null hypothesis $H_0 : \mu = 80$ at$\mu = 75$ is calculated as follows:

$$z_1 = \frac{74.18 - 75}{21/\sqrt{50}} = -0.276$$

The corresponding area under normal curve for $z_1 = -0.276$ is 0.1064.

$$z_2 = \frac{85.82 - 75}{21\sqrt{50}} = 3.643$$

The corresponding area under normal curve for $z_2 = 3.643$ is 0.4995

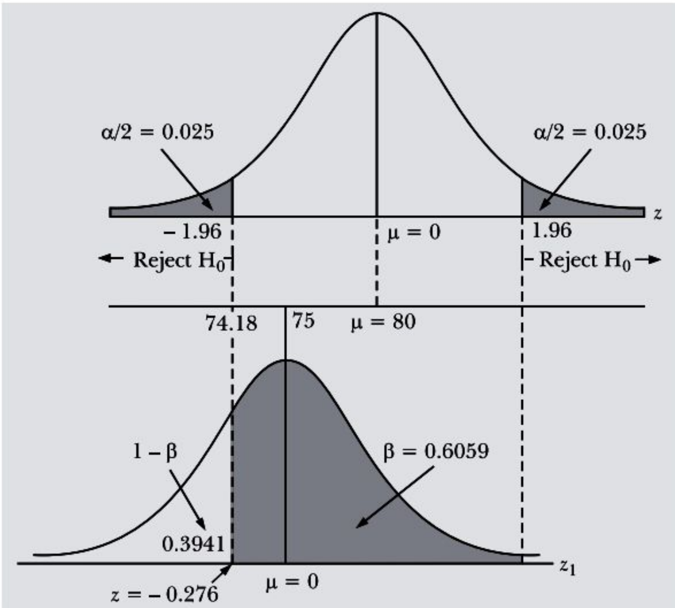
Hence, probability of committing Type II error (β) falls in the region:

$$\beta = P(74.18 < \bar{x}_c < 85.82) = 0.1064 + 0.4995 = 0.6059$$

The probability 0.6059 of committing Type II error (β) is the area to the right of $\bar{x}_c = 74.18$ in the sampling distribution. So, the power of the test is $1 - \beta = 1 - 0.6059 = 0.3941$ as shown in Fig. 10.4(b).

Figure 10.4 (a) Sampling Distribution with H_0: $\mu = 80$

Figure 10.4 (b) Sampling Distribution with H_0: $\mu = 75$



To keep probability of committing error of either Type I or Type II low depends on the cost of committing such an error. However, if cost is high in both the cases, then to keep both Type I or Type II errors low, it is preferred to have large sample size and a low level of significance, α, value.

Special Case : Suppose hypotheses are defined as

$$H_0 : \mu = 80 \text{ and } H_1 : \mu < 80$$

Given $n = 50$, $s = 21$ and $\bar{x} = 71$. For $\alpha = 0.05$ and left-tailed test, the table value $z_{0.05} = -1.645$. The calculated value of z from sample data is

$$z = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}} = \frac{71 - 80}{21/\sqrt{50}} = -3.03$$

For $z_{0.05} = -1.645$, the critical value of the sample mean, $\bar{x}_c$, at population mean, $\mu = 80$ is $\bar{x}_c = 75.115$.

The null hypothesis, H_0, will be rejected for any value less than critical value of mean, $\bar{x}_c = 75.115$, as shown in Fig. 10.5 (a). The distribution of values when the alternative population mean value $\mu = 78$ is true is shown in Fig 10.5(b). The null hypothesis, H_0, will not be rejected when sample mean falls in the acceptance region, $\bar{x}_c \geq 75.151$.

Thus, new critical value of z for the area to the right of $\bar{x}_c \geq 75.151$ is as follows:

$$z_1 = \frac{\bar{x}_c - \mu}{\sigma_{\bar{x}}} = \frac{75.115 - 78}{21/\sqrt{50}} = -0.971$$

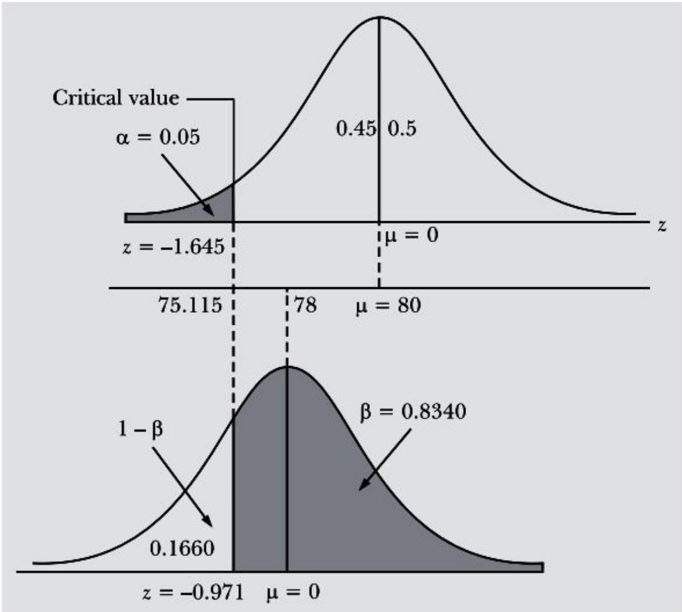


Figure 10.5 (a) Sampling Distribution with $H_0 : \mu = 80$

Figure 10.5 (b) Sampling Distribution with $H_0 : \mu = 78$

The corresponding area under normal curve for $z_1 = -0.971$ is 0.3340.ve.
Hence, probability of committing Type II error (β) falls in the region:

$$\beta = P(75.115 < \bar{x}_c < 80) = 0.3340 + 0.5000 = 0.8340$$

The probability 0.8340 of committing Type II error (β) is the area to the right of $\bar{x}_c = 75.115$ in the sampling distribution. So, the power of the test is $1 - \beta = 1 - 0.8340 = 0.1660$ as shown in Fig. 10.4(b).

Conceptual Questions 10A

1. Discuss the difference in purpose between the estimation of parameters and the testing of statistical hypothesis.
2. Describe the various steps involved in testing of hypothesis. What is the role of standard error in testing of hypothesis? *[Delhi Univ., M.Com, 2005]*
3. What do you understand by null hypothesis and level of significance? Explain with the help of one example. *[HP Univ., MBA, 2003]*
4. What is a test statistic? How is it used in hypothesis testing?
5. Define the term 'level of significance'. How is it related to the probability of committing a Type I error?
   (a) Explain the general steps needed to carry out a test of any hypothesis.
   (b) Explain clearly the procedure of testing hypothesis. Also point out the assumptions in hypothesis testing in large samples.
6. This is always a trade-off between Type I and Type II errors. Discuss. *[Delhi Univ., M.Com, 2006]*
7. When should a one-tailed alternative hypothesis be used? Under what circumstances is each type of test used?
8. Explain the general procedure for determining a critical value needed to perform a test of a hypothesis.
9. What is meant by the terms hypothesis and a test of a hypothesis?
10. Define the standard error of a statistic. How it is helpful in testing of hypothesis and decision-making?
11. Define the terms 'decision rule' and 'critical value'. What is the relationship between these terms?
12. (a) How is power related to the probability of making a Type II error?
    (b) What is the power of a hypothesis test? Why is it important?
    (c) How can the power of a hypothesis test be increased without increasing the sample size?
13. Write short notes on the following:
    (a) Acceptance and rejection regions
    (b) Type I and Type II errors
    (c) Null and alternative hypotheses
    (d) One-tailed and two-tailed tests
14. When planning a hypothesis test, what should be done if probabilities of both Type I and Type II are to be small?

10.7 HYPOTHESIS TESTING FOR POPULATION PARAMETERS WITH LARGE SAMPLES

Hypothesis testing for population parameters with large samples (n>30) is based on the assumption that the population from which the samples are drawn has a normal distribution. Consequently, sampling distribution of mean, $\bar{x}$, is also normal. Since sample size is large therefore even if the population does not have a normal distribution, the sampling distribution of mean $\bar{x}$ is assumed to be normal due to the central limit theorem.

10.7.1 Hypothesis Testing for Single Population Mean

Two-tailed Test

Let μ_0 be the hypothesized value of the population mean, μ, to be tested. For this, the null and alternative hypotheses for two-tailed test are defined as

$$H_0 : \mu = \mu_0 \quad \text{or} \quad \mu - \mu_0 = 0$$

and

$$H_1 : \mu \neq \mu_0$$

Population standard deviation (σ) is known : Since sample size is large therefore sampling distribution of mean, $\bar{x}$, is assumed to be normal due to the central limit theorem. The z-test statistic is given by

$$\text{Test-statistic, } z = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}} = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

In this formula, the numerator, $\bar{x} - \mu$, measures the difference (in an absolute sense) between the observed sample mean, $\bar{x}$, and the hypothesized mean, μ. The denominator represents the ***standard error of the mean***, so z-test statistic represents how many standard errors (deviations) the sample mean, $\bar{x}$, is from the actual population mean, μ.

Population standard deviation (σ) is unknown : In this case, the sample standard deviation, s, is used to estimate population standard deviation, σ. The z-test statistic is given by

$$z = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

The rejection area is represented by two-tails of the sampling distribution of mean, $\bar{x}$, so that half of the level of significance, $\alpha/2$, falls in each tail of the distribution. Hence, $z_{a/2}$ represents the standardized normal variable corresponding to $\alpha/2$ in both the tails of the distribution as shown in Fig 10.1.

Decision Rules

- If calculated value, z_{cal} is more than its absolute critical value, i.e., $z_{cal} \geq |z_{\alpha/2}|$, then reject null hypothesis, H_0.
- Otherwise accept, H_0.

where $z_{\alpha/2}$ is the table value (also called critical value) of z at a chosen level of significance α.

Left-tailed Test

Large sample (n > 30) hypothesis testing about a population mean for a left-tailed test is of the form

$$H_0 : \mu \geq \mu_0 \quad \text{and} \quad H_1 : \mu < \mu_0$$

$$\text{Test statistic, } z = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}} = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

Decision Rules

- If $z_{cal} \le -z_\alpha$, (critical table value), then reject null hypothesis, H_0.
- Otherwise accept H_0.

Right-tailed Test

Large sample ($n > 30$) hypothesis testing about a population mean for a right-tailed test is of the form

$$H_0 : \mu \le \mu_0 \quad \text{and} \quad H_1 : \mu > \mu_0$$

$$\text{Test statistic, } z = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}} = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

Decision Rules

- If $z_{cal} \ge z_\alpha$, (critical table value), then reject the null hypothesis, H_0.
- Otherwise accept H_0.

10.7.2 Relationship Between Interval Estimation and Hypothesis Testing

Consider following statements of null and alternative hypothesis

- $H_0 : \mu = \mu_0$ and $H_1 : \mu \ne \mu_0$ (Two-tailed test)
- $H_0 : \mu \le \mu_0$ and $H_1 : \mu > \mu_0$ (Right-tailed test)
- $H_0 : \mu \ge \mu_0$ and $H_1 : \mu < \mu_0$ (Left-tailed test)

The following are the confidence intervals in all above three cases where hypothesized value μ_0 of population mean, μ, is likely to fall. Accordingly, the decision to accept or reject the null hypothesis will be taken.

Two-tailed Test

Two critical values CV_1 and CV_2 one for each tail of the sampling distribution is computed as follows:

Two-tailed Test: The test of a null hypothesis which can be rejected when the sample statistic is in either extreme end of the sampling distribution.

(a) Known σ

Normal population :	Any sample size, n
Any population :	Large sample size n
$CV_1 = \mu_0 - z_{\alpha/2}\,\sigma_{\bar{x}}$ $CV_2 = \mu_0 + z_{\alpha/2}\,\sigma_{\bar{x}}$	

(b) Unknown σ

Any population :	Large sample size, n
$CV_1 = \mu_0 - z_{\alpha/2}\,s_{\bar{x}}$ $CV_2 = \mu_0 + z_{\alpha/2}\,s_{\bar{x}}$	

Decision Rules

- Reject H_0 when $\bar{x} \le CV_1$ or $\bar{x} \ge CV_2$.
- Accept H_0 when $CV_1 < \bar{x} < CV_2$.

Left-tailed Test

The critical value for left tail of the sampling distribution is computed as follows:

(a) Known σ

Normal population :	Any sample size, n
Any population :	Large sample size, n
$CV = \mu_0 - z_\alpha\,\sigma_{\bar{x}}$	

(b) Unknown σ

Any population :	Large sample size, n
$CV = \mu_0 - z_\alpha\,s_{\bar{x}}$	

Decision Rules

- Reject H_0 when $\bar{x} \le CV$.
- Accept H_0 when $\bar{x} > CV$.

Right-tailed Test

The critical value for right tail of the sampling distribution is computed as follows:

(a) Known σ		(b) Unknown σ	
Normal population :	Any sample size, n	Any population :	Large sample size, n
Any population :	Large sample size, n		
$CV = \mu_0 + z_\alpha \sigma_{\bar{x}}$		$CV = \mu_0 + z_\alpha s_{\bar{x}}$	

Decision Rules

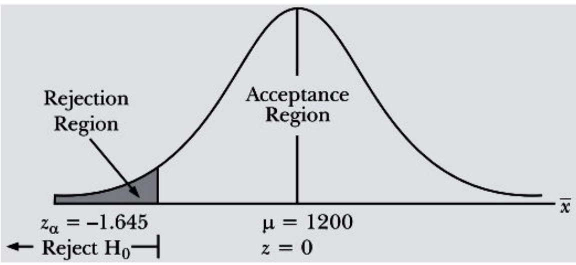
- Reject H_0 when $\bar{x} \ge CV$.
- Accept H_0 when $\bar{x} < CV$.

Example 10.1: Individual filing of income tax returns prior to 30 June had an average refund of ₹1200. Consider the population of 'last minute' filers who file their returns during the last week of June. For a random sample of 400 individuals who filed a return between 25 and 30 June, the sample mean refund was ₹1054 with standard deviation of ₹1600. Using 5 per cent level of significance, test the belief that the individuals who wait until the last week of June to file their returns to get a higher refund than early filers.

Solution: Let null hypothesis H_0 : Individuals who wait until the last week of June to file their returns get a higher return than the early filers, i.e.

$$H_0 : \mu \ge 1200 \quad \text{and} \quad H_1 : \mu < 1200$$

Figure 10.6



Given $n = 400, s = 1600, \bar{x} = 1054, \alpha = 5$ per cent. Using z-test statistic, we have

$$z = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{1054 - 1200}{1600/\sqrt{400}} = -\frac{146}{80} = -1.825$$

Since the calculated value, $z_{cal} = -1.825$ is less than its critical value $z_\alpha = -1.645$ at $\alpha = 0.05$ level of significance, the null hypothesis, H_0, is rejected as shown in Fig. 10.6. Hence, we conclude that individuals who wait until the last week of June are likely to receive a refund of less than ₹1200.

Alternative Approach: $CV = \mu_0 - z_\alpha \sigma_{\bar{x}} = 1200 - 1.645 \times (1600/\sqrt{400})$

$$= 1200 - 131.6 = 1068.4$$

Since $\bar{x}$ (= 1054) <CV(= 1068.4), the null hypothesis H_0 is rejected

Example 10.2: A packaging device is set to fill detergent powder packets with a mean weight of 5 kg with a standard deviation of 0.21 kg. The weight of packets can be assumed to be normally distributed. The weight of packets is known to increase over a period of time due to machine fault, which is not tolerable. A random sample of 100 packets is taken and weighed. This sample has a mean weight of 5.03 kg. Can we conclude that the mean weight produced by the machine has increased? Use a 5 per cent level of significance.

Solution: Let null hypothesis H_0 : Mean weight of packets produced by the machine has increased, i.e.,

$$H_0 : \mu \geq 5 \text{ and } H_1 : \mu < 5$$

Given $n=100$, $\bar{x}=5.03$ kg, $\sigma=0.21$ kg and $\alpha=5$ per cent. Using z-test statistic, we have

$$z = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}} = \frac{5.03-5}{0.21\sqrt{100}} = \frac{0.03}{0.021} = 1.428$$

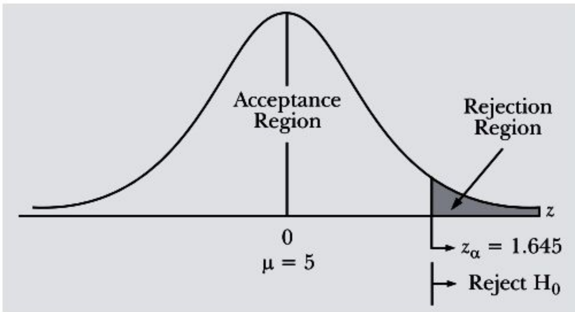


Figure 10.7

Since calculated value, $z_{cal} = 1.428$ is less than its critical value $z_\alpha = 1.645$ at $\alpha = 0.05$, the null hypothesis, H_0 is accepted. Hence, we conclude that mean weight produced by the machine has increased more than 5 kg.

Alternative Approach: $\text{CV} = \mu_0 + z_\alpha = 5 + 1.645 \times (0.21/\sqrt{100})$

$= 5 + 0.034 = 5.034$

Since $(\bar{x} = 5.03) < \text{CV}(= 5.034)$, H_0 is accepted.

Example 10.3: The mean life time of a sample of 400 fluorescent light bulbs produced by a company is found to be 1600 hours with a standard deviation of 150 hours. Test the hypothesis that the mean life time of the bulbs produced in general is higher than the mean life of 1570 hours at $\alpha = 0.01$ level of significance.

Solution: Let null hypothesis H_0 : Mean life time of bulbs is higher than 1570 hours, i.e.

$$H_0 : \mu \geq 1570 \text{ and } H_1 : \mu < 1570$$

Given $n = 400$, $\bar{x} = 1600$ hours, $s = 150$ hr and $\alpha = 0.01$. Using z-test statistic, we have

$$z = \frac{\bar{x}-\mu}{\sigma_{\bar{x}}} = \frac{\bar{x}-\mu}{s/\sqrt{n}} = \frac{1600-1570}{150/\sqrt{400}} = \frac{30}{7.5} = 4$$

Since the calculated value, $z_{cal} = 4$, is more than its critical value $z_\alpha = \pm 2.33$, the H_0 is rejected. Hence, we conclude that the mean lifetime of bulbs produced by the company is not higher than 1570 hours.

Alternatively Approach: $\text{CV} = \mu_0 + z_\alpha s_{\bar{x}} = 1570 + 2.33 \times (150/\sqrt{400})$

$= 1570 + 17.475 = 1587.475$

Since $\bar{x}(= 1600) > \text{CV}(= 1587.47)$, the null hypothesis H_0 is rejected.

Example 10.4: A continuous manufacturing process of steel rods is said to be in 'state of control' and produces acceptable rods if the mean diameter of all rods produced is 2 inches. Although the process standard deviation exhibits stability over time with standard deviation, $\sigma = 0.01$ inch. The process mean may vary due to operator error or problems of process adjustment. Periodically, random samples of 100 rods are selected to determine whether the process is producing acceptable rods. If result of a test indicates that the process is out of control, it is stopped and the source of trouble is identified. Otherwise, it is allowed to continue operating. A random sample of 100 rods is selected resulting in a mean of 2.1 inches. Test the hypothesis to determine whether the process is be continued.

Solution: Since rods that are either too narrow or too wide are unacceptable, the low values and high values of the sample mean lead to the rejection of the null hypothesis. Consider the null hypothesis H_0 that the process may be allowed to continue when diameter is 2 inches. Consequently, rejection region is on both tails of the sampling distribution. The null and alternative hypotheses are stated as follows:

$$H_0 : \mu = 2 \text{ inches (continue process)}$$

$$H_1 : \mu \neq 2 \text{ inches (stop the process)}$$

Given $n = 100$, $\bar{x} = 2.1$, $\sigma = 0.01$, $\alpha = 0.01$. Using the z-test statistic, we have

$$z = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}} = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{2.1 - 2}{0.01/\sqrt{100}} = \frac{0.1}{0.001} = 100$$

Since $z_{cal}=100$ value is more than its critical value $z_{\alpha/2} = 2.58$ at $\alpha = 0.01$, the null hypothesis, H_0 is rejected. Thus, stop the process in order to identify the source of trouble.

Alternative Approach:

$$CV_1 = \mu_0 - z_{\alpha/2}\,\sigma_{\bar{x}} = \mu_0 - z_{\alpha/2}(\sigma/\sqrt{n})$$

$$= 2 - 2.58 \times (0.01/\sqrt{100}) = 2 - 0.003 = 1.997$$

$$CV_2 = \mu_0 + z_{\alpha/2}\,\sigma_{\bar{x}} = 2 + 2.58 \times (0.01/\sqrt{100})$$

$$= 2 + 0.003 = 2.003$$

Since $\bar{x}(= 2.1) \geq CV_2(= 2.003)$, the null hypothesis is rejected.

Example 10.5: An ambulance service claims that it takes on the average 8.9 minutes to reach its destination in emergency calls. To check on this claim, the agency which licenses ambulance services has then timed on 50 emergency calls, getting a mean of 9.3 minutes with a standard deviation of 1.8 minutes. Does this indicate that the average time claimed is too low at the one per cent significance level?

***p*-value:** The probability of getting the sample statistic or a more extreme value, when null hypothesis is true.

Solution: Let Null hypothesis H_0= Average time claimed and observed are same. The null and alternative hypotheses are stated as follows:

$$H_0 : \mu = 8.9 \quad \text{and} \quad H_1 : \mu \neq 8.9$$

Given $n = 50$, $\bar{x} = 9.3$, and $s = 1.8$. Using z-test statistic, we get

$$z = \frac{\bar{x} - \mu}{s_{\bar{x}}} = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{9.3 - 8.9}{1.8/\sqrt{50}} = \frac{0.4}{0.254} = 1.574$$

Since $z_{cal} = 1.574$ in less than its critical value $z_{\alpha/2} = \pm 2.58$ at $\alpha = 0.01$, the null hypothesis is accepted. Hence, we conclude that there is no difference between the average time observed and claimed.

10.7.3 p-value Approach to Test Hypothesis of Single Population Mean

The ***p*-value** is another approach to test null hypothesis about population mean based on a large sample. This approach is also referred to as ***observed significance level approach*** (smallest significance level, α, for which null hypothesis can be rejected). The p-value measures the strength of evidence against null hypothesis That is, p-value indicates the actual risk of committing Type I error when the true null hypothesis is rejected based on the observed value of the test statistic. In other words, *p-value is the probability of observing a sample statistic that is more than the value of test statistic given that the null hypothesis H_0 is true*. The advantage of this approach is that the p-value can be compared directly to the level of significance α.

- The smaller the p-value, the more evidence there is against null hypothesis, H_0.
- A small p-value indicates the value of the test statistic is unusual given the assumption that null hypothesis, H_0 is true

Decision Rules

(a) For a left-tailed test, the p-value is the area to the left of the calculated value of the test statistic. For example, if $z_{cal} = -1.76$, then the area to the left of it is 0.5000 – 0.4608 = 0.0392 or the p-value is 3.92 per cent.

(b) For right-tailed test, the p-value is the area to the right of the calculated value of the test statistic. For example, if $z_{cal} = +2.00$, then the area to the right of it is 0.5000 –0.4772 = 0.0228 or the p value is 2.28 per cent.

Decision rules for left-tailed test and right-tailed test are as under.

- Reject H_0 if p-value $\leq \alpha$
- Accept H_0 if p-value $> \alpha$

(c) For a two-tailed test, the p-value is twice the tail area. If the calculated value of the test statistic falls in the left tail (or right tail), then the area to the left (or right) of the calculated value is multiplied by 2.

- If z is in the upper tail ($z > 0$), find the area under the standard normal curve to the right of z.
- If z is in the lower tail ($z < 0$), find the area under the standard normal curve to the left of z.
- Double the tail area obtained in step 2 to obtain the p-value
- The rejection rule: reject H_0 if the p-value $\leq \alpha$.

Example 10.6: 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. It was found that the mean petrol consumption for 50 cars was 10 km per litre with a standard deviation of 3.5 km per litre. Test the claim that in the new car petrol consumption is 9.5 km per litre on the average at 5 per cent level of significance. *[HP Univ., MBA, 2001]*

Solution: Let Null hypothesis, H_0= Difference is not significant between the company's claim and the sample average value, i.e.,

$$H_0 : \mu = 9.5 \text{ km/litre} \quad \text{and} \quad H_1 : \mu \neq 9.5 \text{ km/litre}$$

Given $\bar{x} = 10, n = 50, s = 3.5$, and $z_{\alpha/2} = 1.96$ at $\alpha = 0.05$ level of significance. Using the z-test statistic, we get

$$z = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}} = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{10 - 9.5}{3.5/\sqrt{50}} = 1.010$$

Since $z_{cal} = 1.010$ is less than its critical value $z_{\alpha/2} = 1.96$ at $\alpha = 0.05$ level of significance, the null hypothesis is accepted. Hence, we can conclude that the new car's petrol consumption is 9.5 km/litre.

The p-value Approach

For $z_{cal} = 1.010$, the cumulative probability is 0.3437. The p-value = 0.5000 – 0.3437 = 0.1563. The p-value is the area to the right as well as left of the calculated value of z-test statistic (for two-tailed test). Since $z_{cal} = 1.010$, then the area to its right is 0.5000 – 0.3437 = 0.1563.

Since it is the two-tailed test, p-value becomes 2(0.1563)=0.3126. Since 0.3126 > α =0.05, null hypothesis H_0 is accepted.

Example 10.7: A sample of 100 tyres is taken from a lot. The mean life of these tyres is found to be 39,350 km. with a standard deviation of 3260 km. Could the sample come from a population with mean life of 40,000 km? Establish 99% confidence limits within which the mean life of tyres is expected to lie. *[Delhi Univ., B.A. (Eco. H), 2001]*

Solution: (a) Let us take the null hypothesis that there is no significant difference between the sample mean and the hypothetical population mean. Calculating standard error of means as follows:

$$\text{Standard error, } \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{3260}{\sqrt{100}} = 326 \cdot$$

Since, $z = \dfrac{\bar{x} - \mu}{\sigma_{\bar{x}}} = \dfrac{40{,}000 - 39{,}350}{326} = \dfrac{650}{326} = 1.994$

is less than the table value, $z = 2.58$ at 1 per cent level of significance, therefore the null hypothesis is accepted. Hence difference (if any) could have arisen due to fluctuation of sampling.

(b) 99% confidence limits: $\bar{x} \pm 2.58\,\sigma_{\bar{x}} = 39{,}350 \pm 2.58\,(326)$

$= 39{,}350 \pm 841.08$

Hence the mean life, $\bar{x}$ of tyres is expected to lie between 38,509 and 40,191.

Example 10.8: The lifetime of electric bulbs for a random sample of 10 from a large consignment gave the following data :

Item :	1	2	3	4	5	6	7	8	9	10
Life in '000 hours :	4.2	4.6	3.9	4.1	5.3	3.8	3.9	4.3	4.4	5.6

Can we accept the hypothesis that there is no significant difference in the sample mean and the hypothetical population mean, $\mu = 4$.

Solution: Let us take the hypothesis that there is no significant difference in the sample mean and the hypothetical population mean. Calculations for sample mean $\bar{x}$ and standard deviation, s are shown in Table 10.4. Applying the t-test,

$$t = \frac{\bar{x} - \mu}{s\sqrt{n}} = \frac{4.4 - 4}{0.589/\sqrt{10}} = \frac{0.4}{0.186} = 2.150\,.$$

Table 10.4 Calculation of $\bar{x}$ and s

x	$(x - \bar{x}) = (x - 4.4)$	$(x - 4.4)^2$
4.2	–0.2	0.04
4.6	0.2	0.04
3.9	–0.5	0.25
4.1	–0.3	0.09
5.2	0.8	0.64
3.8	–0.6	0.36
3.9	–0.5	0.25
4.3	–0.1	0.01
4.4	0	0
5.6	1.2	1.44
44		3.12

where $\bar{x} = \dfrac{1}{n}\Sigma x = \dfrac{44}{10} = 4.4$, and $s = \sqrt{\dfrac{\Sigma(x - \bar{x})^2}{n - 1}} = \sqrt{\dfrac{3.12}{9}} = 0.589$;

Since the calculated value t_{cal} is less than its table value, $t_{0.05} = 2.262$ at $df = n - 1 = 10 - 1 = 9$, therefore hypothesis is accepted. Hence, there is no significant difference in sample mean life and population mean life of bulbs.

Example 10.9: 10 workers are selected at random from a large number of workers in a factory. The number of items produced by them on a certain day are found to be : 51, 52, 53, 55, 56, 57, 58, 59, 60.

In the light these data, would it be appropriate to suggest the mean of the number of items produced in the population is 58? *[M.D. Univ., M.Com., 2006]*

Solution: The calculations for sample mean, $\bar{x}$ and standard deviation, s are shown in Table 10.5. Let us take the hyothesis that there is no significant difference in the sample mean and hypothetical population mean.

Table 10.5 Calculation of $\bar{x}$ and s

x	$(x - \bar{x}) = (x - 56)$	$(x - 56)^2$
51	–5	25
52	–4	16
53	–3	9
55	–1	1
(56) ← *A*	0	0
57	1	1
58	2	4
59	3	9
59	3	9
60	4	16
560		90

$$\bar{x} = \frac{\Sigma x}{n} = \frac{560}{10} = 56 \text{, and } s = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n-1}} = \sqrt{\frac{90}{9}} = 3.162$$

Applying t-test,

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{56 - 58}{3.162/3.162} = 2$$

Since calculated value, $t_{cal} = 2$ is less than its table value $t_\alpha = 2.262$ at $\alpha = 0.05$ and $df = 9$, the hypothesis is accepted. Hence, the mean of the numbers of item produced in the population is 58.

10.7.4 Hypothesis Testing for Difference Between Two Population Means

If there are two independent populations, each having its mean and standard deviation as

Population	*Mean*	*Standard Deviation*
1	μ_1	σ_1
2	μ_2	σ_2

then hypothesis testing procedure explained in the previous section is modified to test whether there is any significant difference between the means of these populations.

The Procedure

Let two independent random samples of large size n_1 and n_2 be drawn from the first and second population, respectively. Let the sample means so calculated be $\bar{x}_1$ and $\bar{x}_2$. The z-test statistic used to determine the difference between the population means $(\mu_1 - \mu_2)$ is based on the difference between the sample means $(\bar{x}_1 - \bar{x}_2)$ because sampling distribution of two mean values has the property $E(\bar{x}_1 - \bar{x}_2) = (\mu_1 - \mu_2)$. This test statistic will follow the standard normal distribution for a large sample due to the central limit theorem.

The z-test statistic is

$$\text{Test statistic, } z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sigma_{\bar{x}_1 - \bar{x}_2}} = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

where $\sigma_{\bar{x}_1 - \bar{x}_2}$ is standard error of the statistic $(\bar{x}_1 - \bar{x}_2)$; $\bar{x}_1 - \bar{x}_2$ is difference between two sample means, i.e., sample statistic; and $\mu_1 - \mu_2$ is difference between population means, i.e., hypothesized population parameter.

If $\sigma_1^2 = \sigma_2^2$, the above formula algebraically reduces to

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

If the standard deviations σ_1 and σ_2 of each of the populations are *not known*, then the standard error of sampling distribution of the sample statistic $\bar{x}_1 - \bar{x}_2$ may estimated by substituting the sample standard deviations s_1 and s_2 as estimates of the population standard deviations. Under this condition, the standard error of $\bar{x}_1 - \bar{x}_2$ is estimated as

$$s_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

The standard error of the *difference between standard deviation of sampling distribution* is given by

$$\sigma_{\sigma_1 - \sigma_2} = \sqrt{\frac{\sigma_1^2}{2n_1} + \frac{\sigma_2^2}{2n_2}}$$

The null and alternative hypothesis are stated as

Null hypothesis : $H_0 : \mu_1 - \mu_2 = d_0$
Alternative hypothesis :

One-tailed Test	*Two-tailed Test*
$H_1 : (\mu_1 - \mu_2) > d_0$	$H_1 : (\mu_1 - \mu_2) \neq d_0$
$H_1 : (\mu_1 - \mu_2) < d_0$	

where d_0 is some specified difference that is desired to be tested. If there is no difference between μ_1 and μ_2, i.e. $\mu_1 = \mu_2$, then $d_0 = 0$.

Decision Rule: Reject H_0 at a specified level of significance α when

One-tailed test	*Two-tailed test*
• $z_{cal} > z_\alpha$ or $z < -z_a$	• $z_{cal} > z_{\alpha/2}$ or $z_{cal} < -z_{\alpha/2}$
• When p-value $< \alpha$	

Example 10.10: A firm believes that the tyres produced by process A on an average last longer than tyres produced by process B. To test this belief, random samples of tyres produced by the two processes were tested and the results are:

Process	*Sample Size*	*Average Lifetime (in km)*	*Standard Deviation (in km)*
A	50	22,400	1000
B	50	21,800	1000

Is there evidence at a 5 per cent level of significance that the firm is correct in its belief?

Solution: Let us take the null hypothesis that there is no significant difference in the average life of tyres produced by processes A and B, i.e.,

$$H_0 : \mu_1 = \mu_2 \quad \text{or} \quad \mu_1 - \mu_2 = 0 \quad \text{and} \quad H_1 : \mu_1 \neq \mu_2$$

Given, $\bar{x}_1 = 22{,}400$ km, $\bar{x}_2 = 21{,}800$ km, $\sigma_1 = \sigma_2 = 1000$ km, and $n_1 = n_2 = 50$. Using the z-test statistic, we get

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sigma_{\bar{x}_1 - \bar{x}_2}} = \frac{\bar{x}_1 - \bar{x}_2}{\sigma_{\bar{x}_1 - \bar{x}_2}} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

$$= \frac{22{,}400 - 21{,}800}{\sqrt{\frac{(1000)^2}{50} + \frac{(1000)^2}{50}}} = \frac{600}{\sqrt{20{,}000 + 20{,}000}} = \frac{600}{200} = 3$$

Since the calculated value, z_{cal} =3, is more than its critical value, $z_{\alpha/2}$ = ±1.645 at α = 0.05 level of significance, the null hypothesis, H_0, is rejected. Hence, we can conclude that the tyres produced by process A last longer than those produced by process B.

The p-value Approach:

$$p\text{-value} = P(z > 3.00) + P(z < -3.00) = 2\,P(z > 3.00)$$
$$= 2(0.5000 - 0.4987) = 0.0026$$

Since p-value of 0.026 is less than specified significance level α = 0.05, null hypothesis, H_0 is rejected.

Example 10.11: An experiment was conducted to compare the mean time (in days) required to recover from a common cold for people given daily dose of 4 mg of vitamin C versus those who were not given a vitamin supplement. Suppose that 35 people were randomly selected for each treatment category and the mean recovery times and standard deviations for the two groups were as follows:

	Vitamin C	*No Vitamin Supplement*
Sample size	35	35
Sample mean	5.8	6.9
Sample standard deviation	1.2	2.9

Test the hypothesis that the use of vitamin C reduces the mean time required to recover from a common cold and its complications at the level of significance, α = 0.05.

Solution: Let us take the null hypothesis that the use of vitamin *C* reduces the mean time required to recover from the common cold and its complications, that is

$$H_0 : (\mu_1 - \mu_2) = 0 \quad \text{and} \quad H_1 : (\mu_1 - \mu_2) \neq 0$$

Given, n_1 = 35, $\bar{x}_1$ = 5.8, s_1 = 1.2 and n_2 = 35, $\bar{x}_2$ = 6.9, s_2 = 2.9, and level of significance, α =0.05. Using z-test statistic, we get

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sigma_{\bar{x}_1 - \bar{x}_2}} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$$

$$= \frac{5.8 - 6.9}{\sqrt{\dfrac{(1.2)^2}{35} + \dfrac{(2.9)^2}{35}}} = \frac{-1.1}{\sqrt{0.041 + 0.240}} = -\frac{1.1}{0.530} = -2.605$$

Since calculated value, z_{cal}= – 2.605 is more than its critical value z_α = –1.645 at significance level α = 0.05, the null hypothesis, H_0 is rejected. Hence, we can conclude that the use of vitamin C does not reduce the mean time required to recover from the common cold and its complications.

Example 10.12: You are working as a purchase manager for a company. The following information has been supplied to you by two manufactures of electric bulbs:

	Comapny A	*Company B*
Mean life (in hours)	1300	1248
Standard deviation (in hours)	82	93
Sample size	100	100

Which brand of bulbs are you going to purchase if you desire to take a risk ot 5%?

[*Madurai Univ., M.Com., 2006; Delhi Univ., MBA, 2004, 2006; Madras Univ., B.Com., 2006*]

Solution: Let us take the hypothesis that there is no significant difference in the mean life of bulbs produced by both the companies.

Given, $\bar{x}_A$ = 1300 , σ_1 = 82, n_1 = 100 and $\bar{x}_B$ = 1248, σ_2 = 93, n_2 = 100; level of significance, α = 0.05. Applying z-test statistic, we get:

$$z = \frac{\bar{x}_A - \bar{x}_B}{\sigma_{\bar{x}_1 - \bar{x}_2}} = \frac{1300 - 1248}{12.398} = 4.19$$

where $$s_{(\bar{x}_1 - \bar{x}_2)} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} = \sqrt{\frac{(82)^2}{100} + \frac{(93)^2}{100}} = \sqrt{\frac{6724}{100} + \frac{8649}{100}} = 12.398$$

Since calculated value, $z_{cal} = 4.19$ is more than its critical value $z_\alpha = 1.96$ at significance value, $\alpha = 0.05$, therefore the null hypothesis is rejected. Hence, there is a significant difference in the mean life of the two makes of bulbs.

Example 10.13: An Educational Testing Service conduced a study to investigate difference between the scores of female and male students on the Mathematics Aptitude Test. The study identified a random sample of 562 female and 852 male students who had achieved the same high score on the mathematics portion of the test. That is, the female and male students viewed as having similar high ability in mathematics. The verbal scores for the two samples are given below:

	Female	*Male*
Sample mean	547	525
Sample stundard deviation	83	78

Do the data support the conclusion that given populations of female and male students with similar high ability in mathematics, the female students will have a significantly high verbal ability? Test at $\alpha = 0.05$ significance level. What is your conclusion?

[*Delhi Univ., MBA, 2003*]

Solution: Let us take the null hypothesis that the female students have high level verbal ability, that is,

$$H_0 : (\mu_1 - \mu_2) \geq 0 \quad \text{and} \quad H_1 : (\mu_1 - \mu_2) < 0$$

Given for female students: $n_1 = 562$, $\bar{x}_1 = 547$, $s_1 = 83$; for male students: $n_2 = 852$, $\bar{x}_2 = 525$, $s_2 = 78$, and $\alpha = 0.05$.

Substituting these values into the z-test statistic, we get

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} = \frac{547 - 525}{\sqrt{\frac{(83)^2}{562} + \frac{(78)^2}{852}}}$$

$$= \frac{22}{\sqrt{12.258 + 7.140}} = \frac{22}{\sqrt{19.398}} = \frac{22}{4.404} = 4.995$$

For one-tailed test at $\alpha = 0.05$ significance level, the critical value of z-test statistic is $z_\alpha = \pm 1.645$. Since $z_{cal} = 4.995$ is more than the critical value $z_\alpha = 1.645$, null hypothesis, H_0 is rejected. Hence, we conclude that there is no evidence to say that difference between verbal ability of female and male students is significant.

Example 10.14: In a sample of size 1000, the mean is 17.5 and the standard deviation is 2.5. In another sample of size 800, the mean is 18 and the standard deviation is 2.7. Assuming that the samples are independent, discuss whether the two samples could have come from population which has the same standard deviation. [*Saurashtra Univ., B.Com, 2006*]

Solution: Let us take the null hypothesis that two samples have come from a population which have the same standard deviation, that is,

$$H_0 : \sigma_1 = \sigma_2 \quad \text{and} \quad H_1 : \sigma_1 \neq \sigma_2.$$

Given, $\sigma_1 = 2.5$, $n_1 = 1000$ and $\sigma_2 = 2.7$, $n_2 = 800$. Applying the z-test statistic, we get

$$z = \frac{\sigma_1 - \sigma_2}{\sigma_{\sigma_1 - \sigma_2}} = \frac{2.7 - 2.5}{0.0876} = \frac{0.2}{0.0876} = 2.283$$

$$\text{where} \quad \sigma_{\sigma_1 - \sigma_2} = \sqrt{\frac{\sigma_1^2}{2n_1} + \frac{\sigma_2^2}{2n_1}} = \sqrt{\frac{(2.5)^2}{2000} + \frac{(2.7)^2}{1600}} = \sqrt{\frac{6.25}{2000} + \frac{7.29}{1600}} = 0.0876$$

Since the $z_{cal} = 2.283$ is more than its critical value $z_\alpha = 1.96$ at $\alpha = 0.05$ level of significance, the null hypothesis, H_0 is rejected. Hence, we conclude that the two samples have not come from a population which has the same standard deviation.

Example 10.15: The mean production of wheat from a sample of 100 fields is 200 tons per acre with a standard deviation of 10 tons. Another sample of 150 fields gives the mean at 220 tons per acre with a standard deviation of 12 tons. Assuming that the standard deviation of the universe is 11 tons, find at 1 per cent level of significance, whether the two results are consistent. *[Punjab Univ., M.Com., Mangalore MBA, 2004]*

Solution: Let us take the null hypothesis that the two results are consistent, that is

$$H_0: \sigma_1 = \sigma_2 \quad \text{and} \quad H_1: \sigma_1 \neq \sigma_2.$$

Given $\sigma_1 = \sigma_2 = 11, n_1 = 100, n_2 = 150$. Applying the z-test statistic, we have

$$z = \frac{\sigma_1 - \sigma_2}{\sigma_{\sigma_1 - \sigma_2}} = \frac{10 - 12}{1.004} = -\frac{2}{1.004} = -1.992$$

$$\text{where} \quad \sigma_{\sigma_1 - \sigma_2} = \sqrt{\frac{\sigma^2}{2}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} = \sqrt{\frac{(11)^2}{2}\left(\frac{1}{100} + \frac{1}{150}\right)} = 1.004$$

Since $z_{cal} = -1.992$ is more than its critical value $z_\alpha = -2.58$ at $\alpha = 0.01$ level of significance, the null hypothesis is accepted. Hence, we conclude that the two results are likely to be consistent.

Example 10.16: A man buys 50 electric bulbs of 'Philips' and 50 electric bulbs of 'HMT'. He finds that 'Philips' bulbs have an average life of 1500 hours with a standard deviation of 60 hours and 'HMT' bulbs have an average life of 1512 hours with a standard deviation of 80 hours. Is there a significant difference in the mean life of the two makes of bulbs? *[Kumaon Univ., MBA, 2006]*

Solution: Let us take the hypothesis that there is no significant difference in the mean life of the two makes of bulbs.

Given, $n_1 = 50$, $\bar{x}_1 = 1500$ hrs, $\sigma_1 = 60$ hrs, and $n_2 = 50$, $\bar{x}_2 = 1512$ hrs, $\sigma_2 = 80$ hrs; level of significance, $\alpha = 0.05$. Applying z-test statistic, we get

$$z = \frac{\bar{x}_p - \bar{x}_h}{\sigma_{\bar{x}_1 - \bar{x}_2}} = \frac{1512 - 1500}{14.14} = 0.849$$

$$\text{where} \quad \sigma_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} = \sqrt{\frac{(60)^2}{50} + \frac{(80)^2}{50}} = \sqrt{\frac{3600 + 6400}{50}} = \sqrt{200} = 14.14$$

Since calculated value, $z_{cal} = 0.849$ is less than its critical value $z_\alpha = 2.58$ at $\alpha = 0.01$ level of significance therefore the null hypothesis, H_0 is accepted. Hence, the difference in the mean life of two makes is not significant.

Self-practice Problems 10A

10.1 The mean breaking strength of the cables supplied by a manufacturer is 1800 with a standard deviation of 100. By a new technique in the manufacturing process it is claimed that the breaking strength of the cables has increased. In order to test this claim a sample of 50 cables is tested. It is found that the mean breaking strength is 1850. Can we support the claim at one per cent level of significance?

10.2 A sample of 100 households in a village was taken and the average income was found to be ₹628 per month with a standard deviation of ₹60 per month. Find the standard error of mean and determine 99 per cent confidence limits within which the incomes of all the people in this village are expected to lie. Also test the claim that the average income was ₹640 per month.

10.3 A random sample of boots worn by 40 combat soldiers in a desert region showed an average life of 1.08 years with a standard deviation of 0.05. Under the standard conditions, the boots are known to have an average life of 1.28 years. Is there reason to assert at a level of significance of 0.05 that use in the desert causes the mean life of such boots to decrease?

10.4 An ambulance service claims that it takes, on an average, 8.9 minutes to reach its destination in emergency calls. To check on this claim, the agency which licenses ambulance services had them timed on 50 emergency calls, getting a mean of 9.3 minutes with a standard deviation of 1.8 minutes. At the level of significance of 0.05, does this constitute evidence that the figure claimed is too low?

10.5 A sample of 100 tyres is taken from a lot. The mean life of the tyres is found to be 39,350 km with a standard deviation of 3260 km. Could the sample come from a population with mean life of 40,000 km? Establish 99 per cent confidence limits within which the mean life of the tyres is expected to lie.
[Delhi Univ., BA(H) Eco., 2004]

10.6 A simple sample of the heights of 6400 Englishmen has a mean of 67.85 inches and a standard deviation of 2.56 inches, while a simple sample of heights of 1600 Austrians has a mean of 68.55 inches and a standard deviation of 2.52 inches. Do the data indicate that the Austrians are on the average taller than the Englishmen? Give reasons for your answer.

10.7 Consider the following hypothesis $H_0 : \mu = 15$ and $H_1 : \mu \neq 15$. A sample of 50 provided a sample mean of 14.2 and standard deviation of 5. Compute the p-value, and conclude about H_0 at the level of significance 0.02.

10.8 A product is manufactured in two ways. A pilot test on 64 items from each method indicates that the products of method A have a sample mean tensile strength of 106 tons and a standard deviation of 12 tons, whereas in method B the corresponding values of mean and standard deviation are 100 tons and 10 tons, respectively. Greater tensile strength in the product is preferable. Use an appropriate large sample test of 5 per cent level of significance to test whether or not method 1 is better for processing the product. State clearly the null hypothesis.
[Delhi Univ., MBA, 2007]

10.9 Two types of new cars produced in India are tested for petrol mileage. One group consisting of 36 cars averaged 14 km per litre. While the other group consisting of 72 cars averaged 12.5 km per litre.
(a) What test-statistic is appropriate if $\sigma_1^2 = 1.5$ and $\sigma_2^2 = 2.0$?
(b) Test whether there exist a significant difference in the petrol consumption of these two types of cars. (use $\alpha = 0.01$) *[Roorkee Univ., MBA, 2004]*

Hints and Answers

10.1 Let $H_0 : \mu = 1800$ and $H_1 : \mu \neq 1800$ (Two-tailed test)
Given $\bar{x} = 1850$, $n = 50$, $\sigma = 100$, $z_\alpha = \pm 2.58$ at $\alpha = 0.01$ level of significance

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{1850 - 1800}{100/\sqrt{50}} = 3.54$$

Since z_{cal} (=3.54) $> z_\alpha$ (= 2.58), reject H_0. The breaking strength of the cables of 1800 does not support the claim.

10.2 Given $n = 100$, $\sigma = 50$; $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{50}{\sqrt{100}} = 5$.

Confidence interval at 99 per cent is $\bar{x} \pm z_\alpha\, \sigma_{\bar{x}} = 628 \pm 2.58\,(5) = 628 \pm 12.9$; $615.1 \le \mu \le 640.9$

Since hypothesized population mean $\mu = 640$ lies in the interval, H_0 is accepted.

10.3 Let $H_0 : \mu = 1.28$ and $H_1 : \mu < 1.28$ (One-tailed test)
Given $n = 40$, $\bar{x} = 1.08$, $s = 0.05$, $z_\alpha = \pm 1.645$ at $\alpha = 0.05$ level of significance

$$z = \frac{\bar{x} - \mu}{s\sqrt{n}} = \frac{1.08 - 1.28}{0.05/\sqrt{40}} = -28.57$$

Since z_{cal} (= – 28.57) $< z_{\alpha/2}$ = – 1.64, H_0 is rejected. Mean life of the boots is less than 1.28 and affected by use in the desert.

10.4 Let $H_0 : \mu = 8.9$ and $H_1 : \mu \neq 8.9$ (Two-tail test)
Given $n = 50$, $\bar{x} = 9.3$, $s = 1.8$, $z_{\alpha/2} = \pm 1.96$ at $\alpha = 0.05$ level of significance

$$z = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{9.3 - 8.9}{1.8/\sqrt{50}} = 1.574$$

Since z_{cal} (= 1.574) $< z_{\alpha/2}$ (= 1.96), H_0 is accepted, that is, claim is valid.

10.5 Let $H_0 : \mu = 40{,}000$ and $H_1 : \mu \neq 40{,}000$ (Two-tail test)
Given $n = 100$, $\bar{x} = 39{,}350$, $s = 3{,}260$, and $z_{\alpha/2} = \pm 2.58$ at $\alpha = 0.01$ level of significance

$$z = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{39{,}350 - 40{,}000}{3260/\sqrt{100}} = -1.994$$

Since z_{cal} (= –1.994) $> z_{\alpha/2}$ (= – 2.58), H_0 is accepted. Thus the difference in the mean life of the tyres could be due to sampling error.

10.6 Let $H_0: \mu_1 = \mu_2$ and $H_1: \mu_1 > \mu_2$; μ_1 and μ_2 = mean height of Austrians and Englishmen, respectively.

Given Austrian: $n_1 = 1600,\ \bar{x}_1 = 68.55, s_1 = 2.52$ and

Englishmen; $n_2 = 6400,\ \bar{x}_2 = 67.85, s_2 = 2.56$

$$z = \frac{x_1 - x_2}{\sigma_{\bar{x}_1 - \bar{x}_2}} = \frac{x_1 - x_2}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$$

$$= \frac{68.55 - 67.85}{\sqrt{\dfrac{(2.52)^2}{1600} + \dfrac{(2.56)^2}{6400}}} = 9.9$$

Since $z_{cal} = 9.9 > z_\alpha (= 2.58)$ for right tail test, H_0 is rejected. Austrian's are on the average taller than the Englishmen.

10.7 Given $n = 50,\ \bar{x} = 14.2, s = 5$, and $\alpha = 0.02$

$$z = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{14.2 - 15}{5/\sqrt{50}} = -1.13$$

Table value of $z = 1.13$ is 0.3708. Thus p-value $= 2\,(0.5000 - 0.3708) = 0.2584$. Since p-value $> \alpha$, H_0 is accepted.

10.8 Let $H_0: \mu_1 = \mu_2$ and $H_1: \mu_1 > \mu_2$; μ_1 and μ_2 = mean life of items produced by Method A and B, respectively.

Given Method 1: $n_1 = 64,\ \bar{x}_1 = 106, s_1 = 12$;

Method 2: $n_2 = 64,\ \bar{x}_2 = 100, s_2 = 10$

$$z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}} = \frac{106 - 100}{\sqrt{\dfrac{(12)^2}{64} + \dfrac{(10)^2}{64}}} = 3.07$$

Since $z_{cal} (= 3.07) > z_\alpha (= 1.645)$ for a right-tailed test, H_0 is rejected. Method A is better than Method B.

10.9 Let $H_0: \mu_1 = \mu_2$ and $H_1: \mu_1 \neq \mu_2$; μ_1 and μ_2 = mean petrol mileage of two types of new cars, respectively

Given $n_1 = 36,\ \bar{x}_1 = 14,\ \sigma^2{}_1 = 1.5$ and $n_2 = 72$, $\bar{x}_2 = 12.5, \sigma_2^2 = 2.0$

$$z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} = \frac{14 - 12.5}{\sqrt{\dfrac{1.5}{36} + \dfrac{2}{72}}}$$

$$= \frac{1.5}{0.2623} = 5.703$$

Since $z_{cal} (= 5.703) > z_{\alpha/2} (= 2.58)$ at $\alpha = 0.01$ level of significance, H_0 is rejected. There is a significant difference in petrol consumption of the two types of new cars.

10.8 HYPOTHESIS TESTING FOR SINGLE POPULATION PROPORTION

Suppose instead of testing a hypothesis regarding a population mean, a population proportion (a fraction, ratio or percentage) p of values that indicates the part of the population or sample having a particular attribute of interest is considered. Then in the same way, a random sample of size n is selected to determine the proportion of elements having a particular attribute of interest (also called success) in it as follows:

$$\bar{p} = \frac{\text{Number of successes in the sample}}{\text{Sample size}} = \frac{x}{n}$$

The value of this sample statistic is compared with the population proportion, p, so as to arrive at a conclusion about the hypothesis.

The three forms of null hypothesis and alternative hypothesis about the hypothesized population proportion, $p_0 p_0$, are as follows:

Null hypothesis	*Alternative hypothesis*
• $H_0: p = p_0$	$H_1: p \neq p_0$ (Two-tailed test)
• $H_0: p \geq p_0$	$H_1: p < p_0$ (Left-tailed test)
• $H_0: p \leq p_0$	$H_1: p > p_0$ (Right-tailed test)

To validate a null hypothesis, it is assumed that the sampling distribution of a proportion, $\bar{p}$, follows a standardized normal distribution. Then, using the value of the sample proportion $\bar{p}$ and its standard deviation $\sigma_{\bar{p}}$, apply the z-test statistic as follows:

$$\text{Test statistic, } z = \frac{\bar{p} - p_0}{\sigma_{\bar{p}}} = \frac{\bar{p} - p_0}{\sqrt{\dfrac{p_0\,(1 - p_0)}{n}}}$$

The comparison of the z-test statistic value to its critical (table) value at a given level of significance enables us to test the null hypothesis about a population proportion based on the difference between the sample proportion $\bar{p}$ and the hypothesized population proportion, p_0.

Decision Rule: Reject H_0 when

One-tailed test	*Two-tailed test*
• $z_{cal} > z_\alpha$ or $z_{cal} < -z_\alpha$	• $z_{cal} > z_{\alpha/2}$ or $z_{cal} < -z_{\alpha/2}$
• p-value $< \alpha$	

10.8.1 Hypothesis Testing for Difference Between Two Population Proportions

Let two independent populations each having proportion and standard deviation of an attribute be as follows:

Population	*Proportion*	*Standard Deviation*
1	p_1	σ_{p1}
2	p_2	σ_{p2}

The procedure of testing null hypothesis for single population can be extended for testing null hypothesis about any difference between the proportions of two populations. The null hypothesis that there is no difference between two population proportions is stated as

$$H_0: p_1 = p_2 \text{ or } p_1 - p_2 = 0 \text{ and } H_1: p_1 \neq p_2$$

The sampling distribution of difference between sample proportions $\bar{p} - \bar{p}_2$ is based on the assumption that the difference between two population proportions, $p_1 - p_2$, is normally distributed. The standard deviation (or error) of sampling distribution of population proportions, $p_1 - p_2$, is given by

$$\sigma_{\bar{p}_1 - \bar{p}_2} = \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}\,;\; q_1 = 1 - p_1 \text{ and } q_2 = 1 - p_2$$

The difference $\bar{p}_1 - \bar{p}_2$ between sample proportions of two independent simple random samples is the point estimator of the difference between two population proportions, $p_1 - p_2$. Then expected value, $E(\bar{p}_1 - \bar{p}_2) = p_1 - p_2$. Thus, z-test statistic for the difference between two population proportions is stated as

$$z = \frac{(\bar{p}_1 - \bar{p}_2) - (p_1 - p_2)}{\sigma_{\bar{p}_1 - \bar{p}_2}} = \frac{\bar{p}_1 - \bar{p}_2}{\sigma_{\bar{p}_1 - \bar{p}_2}}$$

Most often, the standard error of difference between sample proportions is not known. Thus to test a null hypothesis that there is no difference between the population proportions, the two sample proportions $\bar{p}_1$ and $\bar{p}_2$ are combined to get one unbiased estimate of population proportion as follows:

$$\text{Pooled estimate } \bar{p} = \frac{n_1\bar{p}_1 + n_2\bar{p}_2}{n_1 + n_2}$$

The z-test statistic is then restated as

$$z = \frac{\bar{p}_1 - \bar{p}_2}{s_{\bar{p}_1 - \bar{p}_2}}\,;\; s_{\bar{p}_1 - \bar{p}_2} = \sqrt{\bar{p}(1-\bar{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

Example 10.17: An auditor claims that 10 per cent of customers' ledger accounts are carrying mistakes of posting and balancing. A random sample of 600 ledger accounts was taken to test the accuracy of posting and balancing and 45 mistakes were found. Are these sample results consistent with the claim of the auditor? Use 5 per cent level of significance.

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

$$H_0: p = 0.10 \text{ and } H_1: p \neq 0.10$$

Given $\bar{p} = 45/600 = 0.075$, $n = 600$ and $\alpha = 5$ per cent. Using the z-test statistic, we get

$$z = \frac{\bar{p} - p_0}{\sigma_{\bar{p}}} = \frac{0.075 - 0.10}{\sqrt{\frac{0.10 \times 0.90}{600}}} = -\frac{0.025}{0.0122} = -2.049$$

Since z_{cal} ($= -2.049$) is less than its critical value z_α ($= -1.96$) at $\alpha = 0.05$ level of significance, null hypothesis, H_0 is rejected. Hence, we conclude that the sample results are not consistent with the claim of the auditor.

Example 10.18: A company is considering two different television advertisements for promotion of a new product. Management believes that advertisement A is more effective than advertisement B. Two test market areas with identical consumer characteristics are selected: advertisement A was used in one area and advertisement B in another area. In a random sample of 60 customers who saw advertisement A, 18 had tried the product. In a random sample of 100 customers who saw advertisement B, 22 had tried the product. Does this indicate that advertisement A is more effective than advertisement B, if a 5 per cent level of significance is used? *[Delhi Univ., MFC 2004; MBA, 2006]*

Solution: Let us take the null hypothesis that both advertisements are equally effective, that is,

$$H_0: p_1 = p_2 \text{ and } H_1: p_1 > p_2$$

where p_1 and p_2 are proportion of customers who saw advertisement A and advertisement B, respectively.

Given $n_1 = 60$, $\bar{p}_1 = 18/60 = 0.30$; $n_2 = 100$, $\bar{p}_2 = 22/100 = 0.22$ and level of significance $\alpha = 0.05$. Using the z-test statistic, we get

$$z = \frac{(\bar{p}_1 - \bar{p}_2) - (p_1 - p_2)}{s_{\bar{p}_1 - \bar{p}_2}} = \frac{\bar{p}_1 - \bar{p}_2}{s_{\bar{p}_1 - \bar{p}_2}}; \; p_1 = p_2$$

$$= \frac{0.30 - 0.22}{0.0707} = \frac{0.08}{0.0707} = 1.131$$

where

$$s_{\bar{p}_1 - \bar{p}_2} = \sqrt{\bar{p}(1-\bar{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}; \; q = 1 - p$$

$$= \sqrt{0.25 \times 0.75\left(\frac{1}{60} + \frac{1}{100}\right)}$$

$$= \sqrt{0.1875\left(\frac{160}{600}\right)} = 0.0707$$

$$\bar{p} = \frac{n_1\bar{p}_1 + n_2\bar{p}_2}{n_1 + n_2} = \frac{60\,(18/60) + 100\,(22/100)}{60 + 100}$$

$$= \frac{18 + 22}{160} = \frac{40}{160} = 0.25$$

Since $z_{cal} = 1.131$ is less than its critical value $z_\alpha = 1.645$ at $\alpha = 0.05$ level of significance, the null hypothesis, H_0, is accepted. Hence, we conclude that there is no significant difference in the effectiveness of the two advertisements.

Example 10.19: In a simple, random sample of 600 men chosen from city A, 400 are found to be smokers. In another simple, random sample of 900 men chosen from city B, 450 are found to be smokers. Do the data indicate that there is a significant difference in the habit of smoking in two cities? *[Raj Univ., M.Com., 2004; Punjab Univ., M.Com., 2002]*

Solution: Let us take the null hypothesis that there is no significant difference in the habit of smoking in the two cities, that is,

$$H_0: p_1 = p_2 \text{ and } H_1: p_1 \neq p_2$$

where p_1 and p_2 are proportion of men found to be smokers in city A and city B, respectively

Given $n_1 = 600$, $\bar{p}_1 = 400/600 = 0.667$; $n_2 = 900$, $\bar{p}_2 = 450/900 = 0.50$ and level of significance $\alpha = 0.05$. Using the z-test statistic, we have

$$z = \frac{(\bar{p}_1 - \bar{p}_2) - (p_1 - p_2)}{s_{\bar{p}_1 - \bar{p}_2}} = \frac{\bar{p}_1 - \bar{p}_2}{s_{\bar{p}_1 - \bar{p}_2}}; \; p_1 = p_2$$

$$= \frac{0.667 - 0.500}{0.026} = \frac{0.167}{0.026} = 6.423$$

where
$$s_{\bar{p}_1 - \bar{p}_2} = \sqrt{\bar{p}(1-\bar{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}; \; q = 1 - p$$

$$= \sqrt{0.567 \times 0.433\left(\frac{1}{600} + \frac{1}{900}\right)}$$

$$= \sqrt{0.245\,(0.002)} = 0.026;$$

$$\bar{p} = \frac{n_1\bar{p}_1 + n_2\bar{p}_2}{n_1 + n_2} = \frac{600\,(400/600) + 900\,(450/900)}{600 + 900}$$

$$= \frac{400 + 450}{1500} = \frac{850}{1500} = 0.567$$

Since $z_{cal} = 6.423$ is greater than its critical value $z_{\alpha/2} = 2.58$, at $\alpha/2 = 0.05$ level of significance, the null hypothesis, H_0, is rejected. Hence, we conclude that there is a significant difference in the habit of smoking in two cities.

Example 10.20: In a random sample of 500 people belonging to urban area, 200 are found to be commuters of public transport. In another sample of 400 people belonging to rural area, 200 are found to be commuters of public transport. Discuss whether the data reveal a significant difference between urban area so far as the proportion of commuters of public transport is concerned. *[Madras Univ., B.Com., 2005]*

Solution: Let us take the hypothesis that there is no significant difference between rural and urban area so far as the proportion of commuters of public transport is concerned.

$$H_0 : p_1 = p_2 \text{ and } H_1 : p_1 \neq p_2$$

where p_1 and p_2 are the proportion of people who travel with public transport, respectively.

Given, $n_1 = 500$, $\bar{p}_1 = 200/500 = 0.4$ and $n_2 = 400$, $\bar{p}_2 = 200/400 = 0.5$; level of significance, $\alpha = 0.05$. Applying z-test statistic, we get

$$z = \frac{\bar{p}_1 - \bar{p}_2}{\sigma_{p_1 - p_2}} = \frac{0.4 - 0.5}{0.033} = -3.03$$

where, $\sigma_{\bar{p}_2 - \bar{p}_2} = \sqrt{pq\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} = \sqrt{0.444 \times 0.556\left(\frac{1}{500} + \frac{1}{400}\right)} = 0.033$

and $p = \frac{n_1\bar{p}_1 + n_2\bar{p}_2}{n_1 + n_2} = \frac{500 \times 0.4 + 400 \times 0.5}{500 + 400} = \frac{200 + 200}{500 + 400} = 0.444.$

Since $z_{cal} = -3.03$ is less than its critcial value $z_\alpha = -2.58$ at $\alpha = 0.01$ level of significance, therefore the null hypothesis, H_0 is rejected. Hence, there is significance difference in the rural urban areas so far as the proportion of commuters of public transport is concerned.

10.9 HYPOTHESIS TESTING FOR A BINOMIAL PROPORTION

If elements in a sample have mutually exclusive attributes such as good or bad then instead of counting *proportion of elements having same attribute (called success)* in a sample, the number of elements having desired attributes (also called *number of successes*)are counted in a sample. The z-test statistic for computing the difference between the number of successes in a sample and the hypothesized (expected) number of successes in the population is given by

$$z = \frac{\text{Sample estimate} - \text{Expected value}}{\text{Standard error of estimate}} = \frac{x - np}{\sqrt{npq}}$$

Since sampling distribution of the number of successes in the sample follows a binomial distribution with mean, $\mu = np$ and standard deviation $\sqrt{npq}$, then normal distribution provides a good approximation to the binomial distribution provided the sample size is large, i.e., both $np \geq 5$ and $n\,(1 - p) \geq 5$.

Example 10.21: In a hospital 480 female and 520 males babies were born in a week. Do these figures confirm the hypothesis that males and females are born in equal number?

[Madras Univ, M.Com., 2000]

Solution: Let us take the hypothesis that male and female babies are born in equal number, that is, $p = 1/2$, that is

$$H_0 : p_1 = p_2 = 0.5 \text{ and } H_1 : p_1 \neq p_2.$$

Given, $p_1 = 480, p_2 = 520$. Applying z-test statistic, we get

$$z = \frac{p_1 - p_2}{\sigma_{p_1 - p_2}} = \frac{480 - 520}{15.81} = -\frac{20}{15.81} = -1.265$$

where $\sigma_{p_1 - p_2} = \sqrt{npq} = \sqrt{1000 \times 0.5 \times 0.5} = 15.81$

Since $z_{cal} = -1.265$ is more than its critical value $z_\alpha = -1.96$ at $\alpha = 0.05$ level of significance, the null hypothes is H_0 is accepted. Hence, it can be concluded that the male and females babies are born in equal number.

Example 10.22: Suppose a production manager implements a newly developed sealing system for boxes. He takes a random sample of 200 boxes from the daily output and finds that 12 boxes need rework. He is interested to determine whether the new sealing system has increased defective packages below 10 per cent. Use 1 per cent level of significance.

Solution: Let us take the null hypothesis that the new sealing system has increased defective packages below 10 per cent, that is,

$$H_0: p \geq 0.10 \quad \text{and} \quad H_1: p < 0.10$$

Given $n = 200, p = 0.10$, and level of significance $\alpha = 0.01$. Using the z-test statistic, we get

$$z = \frac{x - np}{\sqrt{npq}} = \frac{12 - 200\,(0.10)}{\sqrt{200\,(0.10)\,(0.90)}} = \frac{12 - 20}{\sqrt{18}} = -\,1.885$$

Since z_{cal} ($= -1.885$) is more than its critical value $z_\alpha = -\,2.33$ at $\alpha = 0.01$ level of significance, the null hypothesis, H_0, is accepted. Hence, we conclude that the new sealing system has increased defective packages below 10 per cent.

Example 10.23: In 324 throws of a six-faced dice, odd numbers appeared 180 times. Would you say that the dice is fair at 5 per cent level of significance? *[MD Univ., M.Com., 2005]*

Solution: Let us take the hypothesis that the dice is fair, that is,

$$H_0: p = 162/324 = 0.5 \text{ and } H_1: p \neq 0.5$$

Given $n = 324, p = q = 0.5$ (i.e., 162 odd or even numbers out of 324 throws). Applying the z-test statistic, we get

$$z = \frac{x - np}{\sqrt{npq}} = \frac{180 - 162}{\sqrt{324 \times 0.5 \times 0.5}} = \frac{18}{9} = 2$$

Since $z_{cal} = 2$ is more than its critical value $z_{\alpha/2} = 1.96$ at $\alpha = 0.05$ significance level, the null hypothesis H_0 is rejected. Hence, we conclude that the dice is not fair.

Example 10.24: Of those women who are diagnosed to have early-stage breast cancer, one-third eventually die of the disease. Suppose an NGO launch a screening programme to provide for the early detection of breast cancer and to increase the survival rate of those diagnosed to have the disease. A random sample of 200 women was taken from among those who were periodically screened and who were diagnosed to have the disease. If 164 women in the sample of 200 survive the disease, can screening programme be considered effective? Test using $\alpha = 0.01$ level of significance and explain the conclusions from your test.

Solution: Let us take the null hypothesis that the screening programme was effective, that is,

$$H_0 : p = 1 - (1/3) = 2/3 \quad \text{and} \quad H_1 : p > 2/3$$

Given $n = 200, p = 2/3, q = 1/3$ and $\alpha = 0.05$. Applying the z-test statistic, we get

$$z = \frac{x - np}{\sqrt{npq}} = \frac{164 - 200 \times (2/3)}{\sqrt{200 \times (2/3)(1/3)}}$$

$$= \frac{164 - 133.34}{\sqrt{44.45}} = \frac{30.66}{6.66} = 4.60$$

Since $z_{cal} = 4.60$ is greater than its critical value $z_\alpha = 2.33$ at $\alpha = 0.01$ significance level, the null hypothesis, H_0, is rejected. Hence, we conclude that the screening programme was not effective.

Self-practice Problems 10B

10.10 A company manufacturing a certain type of breakfast cereal claims that 60 per cent of all housewives prefer that type to any other. A random sample of 300 housewives contains 165 who do prefer that type. At 5 per cent level of significance, test the claim of the company.

10.11 An auditor claims that 10 per cent of a company's invoices are incorrect. To test this claim a random sample of 200 invoices is checked and 24 are found to be incorrect. At 1 per cent significance level, test whether the auditor's claim is supported by the sample evidence.

10.12 A sales clerk in the department store claims that 60 per cent of the shoppers entering the store leave without making a purchase. A random sample of 50 shoppers showed that 35 of them left without buying anything. Are these sample results consistent with the claim of the sales clerk? Use a significance level of 0.05. *[Delhi Univ., MBA, 2000, 2003]*

10.13 A dice is thrown 49,152 times and of these 25,145 yielded either 4, 5 or 6. Is this consistent with the hypothesis that the dice must be unbiased?

10.14 A coin is tossed 100 times under identical conditions independently yielding 30 heads and 70 tails. Test at 1 per cent level of significance whether or not the coin is unbiased. State clearly the null hypothesis and the alternative hypothesis.

10.15 Before an increase in excise duty on tea, 400 people out of a sample of 500 persons were found to be tea drinkers. After an increase in the duty, 400 persons were known to be tea drinkers in a sample of 600 people. Do you think that there has been a significant decrease in the consumption of tea after the increase in the excise duty? *[Delhi Univ., M.Com., 2000; MBA, 2002]*

10.16 In a random sample of 1000 persons from UP 510 were found to be consumers of cigarettes. In another sample of 800 persons from Rajasthan, 480 were found to be consumers of cigarettes. Do the data reveal a significant difference between UP and Rajasthan so far as the proportion of consumers of cigarettes in concerned? *[MC Univ., M.Com, 2002]*

10.17 A machine puts out 10 defective units in a sample of 200 units. After the machine is overhauled it puts out 4 defective units in a sample of 100 units. Has the machine been improved? *[Madras Univ., M.Com., 2003]*

10.18 500 units from a factory are inspected and 12 are found to be defective, 800 units from another factory are inspected and 12 are found to be defective. Can it be concluded that at 5 per cent level of significance production at the second factory is better than in first factory? *[Kurukshetra Univ., MBA, 2004; Delhi Univ., MBA, 2006]*

10.19 A wholesaler of eggs claims that only 4 per cent of the eggs supplied by him are defective. A random sample of 600 eggs contained 36 defectives. Test the claim of the wholesaler. *[IGNOU, 2002]*

Hints and Answers

10.10 Let H_0: p = 60 per cent and H_1: p < 60 per cent

Given, sample proportion, $\bar{p} = 165/300 = 0.55$; $n = 300$ and $z_\alpha = 1.645$ at $\alpha = 5$ per cent

$$z = \frac{\bar{p} - p_0}{\sqrt{\frac{pq}{n}}} = \frac{0.55 - 0.60}{\sqrt{\frac{0.60 \times 0.40}{300}}} = -1.77$$

Since z_{cal} (= –1.77) is less than its critical value $z_\alpha = -1.645$, the H_0 is rejected. Percentage preferring the breakfast cereal is lower than 60 per cent.

10.11 Let H_0: p = 10 per cent and H_1: $p \neq$ 10 per cent

Given sample proportion, $\bar{p} = 24/200 = 0.12$; $n = 200$ and $z_{\alpha/2} = 2.58$ at $\alpha = 1$ per cent

$$z = \frac{\bar{p} - p_0}{\sqrt{\frac{pq}{n}}} = \frac{0.12 - 0.10}{\sqrt{\frac{0.10 \times 0.90}{200}}} = 0.943$$

Since z_{cal} (= 0.943) is less than its critical value $z_{\alpha/2}$ = 2.58, the H_0 is accepted. Thus, the percentage of incorrect invoices is consistent with the auditor's claim of 10 per cent.

10.12 Let H_0: p = 60 per cent and H_1: $p \neq$ 60 per cent

Given sample proportion, $\bar{p}$ = 35/60 = 0.70;

n = 50 and $z_{\alpha/2}$ = 1.96 at α = 5 per cent

$$z = \frac{\bar{p} - p_0}{\sqrt{\frac{pq}{n}}} = \frac{0.70 - 0.60}{\sqrt{\frac{0.60 \times 0.40}{50}}} = 1.44$$

Since z_{cal} (= 1.44) is less than its critical value $z_{\alpha/2}$ = 1.96, the H_0 is accepted. Claim of the sales clerk is valid.

10.13 Let H_0 : p = 50 per cent and H_1: $p \neq$ 50 per cent

Given sample proportion of success

p = 25, 145/49, 152 = 0.512; n = 49, 152

$$z = \frac{\bar{p} - p_0}{\sqrt{\frac{pq}{n}}} = \frac{0.512 - 0.50}{\sqrt{\frac{0.50 \times 0.50}{49{,}152}}} = \frac{0.012}{0.002} = 6.0$$

Since z_{cal} (= 6.0) is more than its critical value $z_{\alpha/2}$ = 2.58 at α = 0.01, the H_0 is rejected, dice is biased.

10.14 Let H_0: p = 50 per cent and H_1: $p \neq$ 50 per cent

Given n = 100, sample proportion of success, $\bar{p}$ = 30/100 = 0.30 and $z_{\alpha/2}$ = 2.58 at α = 0.01

$$z = \frac{\bar{p} - p_0}{\sqrt{\frac{pq}{n}}} = \frac{0.30 - 0.50}{\sqrt{\frac{0.50 \times 0.50}{100}}} = -\frac{0.20}{0.05} = -4$$

Since z_{cal} (= – 4) is less than its critical value $z_{\alpha/2}$ = – 2.58, the H_0 is rejected.

10.15 Let H_0: p = 400/500 = 0.80 and H_1: p < 0.80,

Given n_1 = 500, n_2 = 600, $\bar{p}_1$ = 400/500 = 0.80, $\bar{p}_2$ = 400/600 = 0.667 and z_α = 2.33 at α = 0.01

$$\bar{p} = \frac{n_1\bar{p}_1 + n_2\bar{p}_2}{n_1 + n_2} = \frac{400 + 400}{500 + 600} = 0.727;$$

$$q = 1 - 0.727 = 0.273$$

$$s_{\bar{p}_1 - \bar{p}_2} = \sqrt{\bar{p}\,\bar{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

$$= \sqrt{0.727 \times 0.273\left(\frac{1}{500} + \frac{1}{600}\right)} = 0.027$$

$$z = \frac{\bar{p}_1 - \bar{p}_2}{s_{\bar{p}_1 - \bar{p}_2}} = \frac{0.80 - 0.667}{0.027} = 4.93$$

Since z_{cal} (= 4.93) is more than its critical value z_α = 2.33, the H_0 is rejected. Decrease in the consumption of tea after the increase in the excise duty is significant.

10.16 Let H_0: $p_1 = p_2$ and H_1: $p_1 \neq p_2$

Given UP: n_1 = 1000, $\bar{p}_1$ = 510/1000 = 0.51;

Rajasthan: n_2 = 800, $\bar{p}_2$ = 480/800 = 0.60

$$\bar{p} = \frac{n_1\bar{p}_1 + n_2\bar{p}_2}{n_1 + n_2} = \frac{510 + 480}{1000 + 800} = 0.55;$$

$$q = 1 - 0.55 = 0.45$$

$$s_{\bar{p}_1 - \bar{p}_2} = \sqrt{\bar{p}\bar{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

$$= \sqrt{0.55 \times 0.45\left(\frac{1}{1000} + \frac{1}{800}\right)} = 0.024.$$

$$z = \frac{\bar{p}_1 - \bar{p}_2}{s_{\bar{p}_1 - \bar{p}_2}} = \frac{0.51 - 0.60}{0.024} = -3.75$$

Since z_{cal} (= –3.75) is less than its critical value $z_{\alpha/2}$= –2.58, the H_0 is rejected. The proportion of consumers of cigarettes in the two states is significant.

10.17 Let H_0 : $p_1 \leq p_2$ and H_1: $p_1 > p_2$

Given before overhaul: n_1 = 200, $\bar{p}_1$ = 10/200 = 0.05;

after overhaul: n_2 = 100, $\bar{p}_2$ = 4/100 = 0.04

$$\bar{p} = \frac{n_1\bar{p}_1 + n_2\bar{p}_2}{n_1 + n_2} = \frac{10 + 4}{200 + 100} = 0.047;$$

$$q = 1 - p = 0.953$$

$$s_{\bar{p}_1 - \bar{p}_2} = \sqrt{0.047 \times 0.953\left(\frac{1}{200} + \frac{1}{100}\right)} = 0.026;$$

$$z = \frac{\bar{p}_1 - \bar{p}_2}{s_{\bar{p}_1 - \bar{p}_2}} = \frac{0.05 - 0.04}{0.026} = 0.385$$

Since z_{cal} (= 0.385) is less than its critical value

z_α = 1.645 at α = 0.05, the H_0 is accepted.

10.18 Let H_0: $p_1 \leq p_2$ and H_1: $p_1 > p_2$

Given n_1 = 500, $\bar{p}_1$ = 12/500 = 0.024, n_2 = 800, $\bar{p}_2$ = 12/800 = 0.015

$$\bar{p} = \frac{n_1\bar{p}_1 + n_2\bar{p}_2}{n_1 + n_2} = \frac{12 + 12}{500 + 800} = 0.018;$$

$$q = 1 - p = 0.982$$

$$s_{\bar{p}_1 - \bar{p}_2} = \sqrt{\bar{p}\,\bar{q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

$$= \sqrt{0.018 \times 0.912\left(\frac{1}{500} + \frac{1}{800}\right)} = 0.0076$$

$$z = \frac{\bar{p}_1 - \bar{p}_2}{s_{\bar{p}_1 - \bar{p}_2}} = \frac{0.024 - 0.015}{0.0076} = 1.184$$

Since $z_{cal} = 1.184$ is less than its critical value $z_\alpha = 1.645$ at $\alpha = 0.05$, the H_0 is accepted. Production in second factory is better than in the first factory.

10.19 Let H_0: $p = 4$ per cent and H_1: $p \neq 4$ per cent

Given $n = 600$, $\bar{p} = 36/600 = 0.06$

Confidence limits: $\bar{p} \pm z_\alpha \sqrt{\dfrac{pq}{n}}$

$$= 0.06 \pm 1.96 \sqrt{(0.04 \times 0.96)/600}$$

$$= 0.06 \pm 0.016;\ 0.44 \leq p \leq 0.076$$

Since probability 0.04 of the claim does not fall into the confidence limit, the claim is rejected.

10.10 HYPOTHESIS TESTING FOR POPULATION MEAN WITH SMALL SAMPLES

When sample size is small (less than 30), the sampling distribution of a sample statistic, such as mean, $\bar{x}$ and proportion, $\bar{p}$, is not normal. Consequently, in such cases while testing a null hypothesis it is assumed that the samples have to be drawn from a normally or approximately normally distributed population. However, the critical values of sample statistic $\bar{x}$ or $\bar{p}$ depend on whether or not the population standard deviation σ is known. If value of the population standard deviation, σ, is not known, then its value is estimated by computing the standard deviation of sample, s, and the standard error of the mean is calculated by using the formula, $\sigma_{\bar{x}} = s/\sqrt{n}$. But the resulting sampling distribution may not be normal even if sampling is done from a normally distributed population. In all such cases, the sampling distribution referred as ***Student's t-distribution*** is used to test a null hypothesis.

10.10.1 Properties of Student's *t*-Distribution

If small samples of size are drawn from normal population with mean μ and a sample statistic of interest for each sample is computed, then probability density function of the t-distribution with degrees of freedom ν (a Greek letter *nu*) is given by

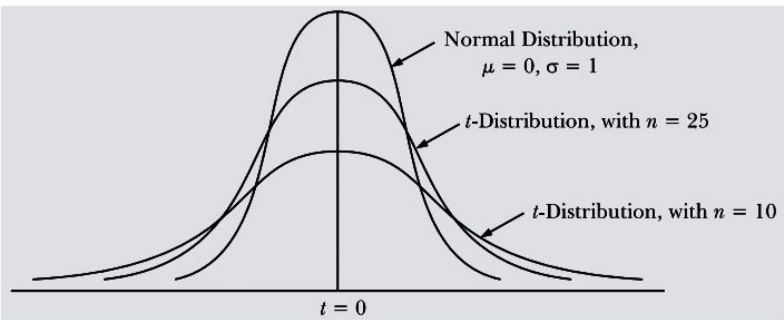
$$y = \frac{y_0}{\left(1 + \dfrac{t^2}{n-1}\right)^{\frac{n}{2}}} = \frac{y_0}{\left(1 + \dfrac{t^2}{\nu}\right)^{\frac{(\nu+1)}{2}}};\ -\infty \leq t \leq \infty$$

where y_0 is a constant whose value depends on sample size, n, such that the total area under the curve is unity and degree of freedom, $\nu = n - 1$.

The t-distribution is symmetrical about the line, $t = 0$ same as normal distribution.

The shape of the t-distribution depends on the sample size, n. The variability of t decreases with the increase of sample size, n. This implies that for different degrees of freedom, the shape of the t-distribution also different as shown in Fig. 10.8. When sample size, n, is infinitely large, the distribution of t and z random variables becomes identical.

Figure 10.8
Comparison of *t*-Distribution with Standard Normal Distribution



The t-distribution is less peaked than normal distribution at the center of the curve and longer in its tails.

The t-distribution has greater dispersion than standard normal distribution. This is because the t-statistic involves two random variables $\bar{x}$ and s, whereas z-statistic involves only the sample mean, $\bar{x}$. The variance of t-distribution is defined only when degree of freedom, $\nu \geq 3$ and is given by variance $(t) = \nu/(\nu - 2)$.

The degree of freedom refers to the number of independent squared deviations in the variance of sampling distribution, s^2 that are available for estimating population variance, σ^2.

10.10.2 Hypothesis Testing for Single Population Mean

The test-statistic to compute difference between the sample mean, $\bar{x}$, and population mean, μ, is given by

$$t = \frac{\bar{x} - \mu}{s_{\bar{x}}} = \frac{\bar{x} - \mu}{s/\sqrt{n}}; \; s = \sqrt{\frac{\Sigma(x - \bar{x})^2}{n-1}}$$

where s is an unbiased estimation of unknown population standard deviation, σ. This test-statistic has t-distribution with $n - 1$ degrees of freedom.

Confidence Interval

The confidence interval estimate of the population mean, μ, when unknown population standard deviation, σ, is estimated by sample standard deviation, s, is given by

- Two-tailed test : $\bar{x} \pm t_{\alpha}\frac{s}{\sqrt{n}}$; α = level of significance.
- One-tailed test : $\bar{x} \pm t_{\alpha/2}\frac{s}{\sqrt{n}}$; α = level of significance.

where t-test statistic value is based on a t-distribution with $n - 1$ degrees of freedom and $1 - \alpha$ is the confidence coefficient.

- Null hypothesis, $H_0: \mu = \mu_0$
- Alternative hypothesis

One-tailed test	Two-tailed test
$H_1: \mu > \mu_0$ or $\mu < \mu_0$	$H_1: \mu \neq \mu_0$

Decision rule: Reject null hypothesis, H_0, at the given degrees of freedom n–1 and level of significance when

One-tailed test	*Two-tailed test*
• $t_{cal} > t_{\alpha}$ or $t_{cal} < -t_{\alpha}$	$t_{cal} > t_{\alpha/2}$ or $t_{cal} < -t_{\alpha/2}$
• Reject H_0 when p-value $< \alpha$	

Example 10.25: The average breaking strength of steel rods is specified to be 18.5 thousand kg. To test the breaking strength a sample of 14 rods was taken. The mean and standard deviation so obtained were 17.85 thousand kg and 1.955 thousand kg, respectively. Is there any significant deviation in the breaking strength of the rods.

Solution: Let us take the null hypothesis that there is no significant deviation in the breaking strength of the rods, that is,

$$H_0: \mu = 18.5 \quad \text{and} \quad H_1: \mu \neq 18.5$$

Given $n = 14$, $\bar{x} = 17.85$, $s = 1.955$, $df = n - 1 = 13$, and $\alpha = 0.05$ level of significance. The critical value of t at $df = 13$ and $\alpha/2 = 0.025$ is $t_{\alpha/2} = 2.16$.

Using the z-test statistic, we get

$$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}} = \frac{17.85 - 18.5}{\frac{1.955}{\sqrt{14}}} = -\frac{0.65}{0.522} = -1.24$$

Since t_{cal} (= – 1.24) value is more than its critical value $t_{\alpha/2}$ = – 2.16 at level of significance, $\alpha/2$ = 0.025 and df = 13, the null hypothesis H_0 is accepted. Hence, we conclude that there is no significant deviation of sample mean from the population mean.

Example 10.26: An automobile tyre manufacturer claims that the average life of a particular grade of tyre is more than 20,000 km when used under normal conditions. A random sample of 16 tyres was tested and a mean and standard deviation of 22,000 km and 5000 km, respectively were computed. Assuming the life of the tyres in km to be approximately normally distributed, decide whether the manufacturer's claim is valid.

Solution: Let us take the null hypothesis that the manufacturer's claim is valid, that is,

$$H_0 : \mu \geq 20{,}000 \quad \text{and} \quad H_1 : \mu < 20{,}000$$

Given n = 16, $\bar{x}$ = 22,000, s = 5000, df = 15 and α = 0.05 level of significance. The critical value of t at df = 15 and α = 0.05 is t_α = 1.753.

Using the z-test statistic, we get

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{22{,}000 - 20{,}000}{5000/\sqrt{16}} = \frac{2000}{1250} = 1.60$$

Since t_{cal} (= 1.60) value is less than its critical value, t_α = 1.753, atα = 0.05 level of significance and df = 15, the null hypothesis H_0 is accepted. Hence, we conclude that the manufacturer's claim is valid.

Example 10.27: A fertilizer mixing machine is set to mix 12 kg of nitrate in every 100 kg of fertilizer. Ten bags of 100 kg each are examined. The percentage of nitrate so found is 11, 14, 13, 12, 13, 12, 13, 14, 11 and 12. Is there any reason to believe that the machine is defective?

Solution: Let us take the null hypothesis that the machine mixes 12 kg of nitrate in every 100 kg of fertilizer, and is not defective, that is,

$$H_0 : \mu = 12 \quad \text{and} \quad H_1 : \mu \neq 12$$

Given n = 10, df = 9, and α = 0.05 level of significance. Critical value $t_{\alpha/2}$ = 2.262 at df = 9 and $\alpha/2$ = 0.025.

Suppose the weight of nitrate in bags is normally distributed and its standard deviation is unknown. The calculations for sample mean, $\bar{x}$, and standard deviation, s, are shown in Table 10.6.

Table 10.6 Calculations of Sample Mean and Standard Deviation *s*

Variable x	*Deviation,* $d = x - 12$	d^2
11	–1	1
14	2	4
13	1	1
12	0	0
13	1	1
12	0	0
13	1	1
14	2	4
11	–1	1
12	0	0
125	5	13

$$\bar{x} = \frac{\sum x}{n} = \frac{125}{10} = 12.5$$

and
$$s = \sqrt{\frac{\sum(x-\bar{x})^2}{n-1}} = \sqrt{\frac{\sum d^2}{n-1} - \frac{\left(\sum d\right)^2}{n\,(n-1)}}$$

$$= \sqrt{\frac{13}{9} - \frac{(5)^2}{10\,(9)}} = 1.08$$

Using the z-test statistic, we have

$$t = \frac{\bar{x}-\mu}{\frac{s}{\sqrt{n}}} = \frac{12.5-12}{\frac{1.08}{\sqrt{10}}} = \frac{0.50}{0.341} = 1.466$$

Since t_{cal} (= 1.466) value is less than its critical value $t_{\alpha/2} = 2.262$ at $\alpha/2 = 0.025$ and $df = 9$, the null hypothesis H_0 is accepted. Hence, we conclude that the manufacturer's claim is valid, i.e., the machine is not defective.

Example 10.28: A random sample of size 16 has the sample mean 53. The sum of the squares of deviation taken from the mean value is 150. Can this sample be considered as taken from the population with mean, 56? Obtain 95 per cent and 99 per cent confidence limits of the sample mean.

Solution: Let us take the null hypothesis that the population mean is 56, that is,

$$H_0: \mu = 56 \quad \text{and} \quad H_1: \mu \neq 56$$

Given $n = 16$, $df = n - 1 = 15$, $\bar{x} = 53$; $s = \sqrt{\frac{\sum(x-\bar{x})^2}{(n-1)}} = \sqrt{\frac{150}{15}} = 3.162$

- 95 per cent confidence limit

$$\bar{x} \pm t_{0.05}\frac{s}{\sqrt{n}} = 53 \pm 2.13\,\frac{3.162}{\sqrt{16}} = 53 \pm 2.13\,(0.790) = 53 \pm 1.683$$

- 99 per cent confidence limit

$$\bar{x} \pm t_{0.01}\frac{s}{\sqrt{n}} = 53 \pm 2.95\,\frac{3.162}{\sqrt{16}} = 53 \pm 2.33$$

Example 10.29: A certain stimulus administered to each of 12 patients resulted in the following increases of blood pressures: 5, 2, 8, –1, 3, 0, 6, –2, 1, 5, 0, 4. Can it be concluded that the simulus will be, in general, accompanied by an increase in blood pressure?

Solution: Let us take the hypothesis that simulus accompanied an increase in blood pressure.

Let the assumed mean, A be 2. Calculations for mean, $\bar{x}$, and standard deviation, s, of the sample are shown in Table 10.7:

Table 10.7 Calculations for Mean and Standard Deviation

Blood Pressure Variation	*Number of Patients (f)*	$d = x - A$	fd	d^2	fd^2
–2	1	–4	–4	16	16
–1	1	–3	–3	9	9
0	2	–2	–4	4	8
1	1	–1	–1	1	1
(2) ← A	1	0	0	0	0
3	1	1	1	1	1
4	1	2	2	4	4
5	2	3	6	9	18
6	1	4	4	16	16
8	1	6	6	36	36
	12		7		109

$$\bar{x} = A + \frac{1}{n}\sum fd = 2 + \frac{7}{12} = 2 + 0.58 = 2.58$$

Applying t-test, we get

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n-1}} = \frac{2.58 - 0}{2.94/\sqrt{(12-1)}} = \frac{2.58}{0.942} = 2.738.$$

where $s = \sqrt{\left[\frac{\sum fd^2}{n} - \left(\frac{\sum fd}{n}\right)^2\right]} = \sqrt{\left[\frac{109}{12} - \left(\frac{7}{12}\right)^2\right]}$

$$= \sqrt{9.083 - (0.667)^2} = \sqrt{9.083 - 0.339} = \sqrt{8.744} = 2.957.$$

Since $t_{cal} = 2.957$ is more than its critical value $t_\alpha = 2.20$ at $\alpha = 0.05$ level of significance and $df = 11$, the hypothesis is rejected. Hence, stimulus will in general not be accompanied by an increase in blood pressure.

Example 10.30: The lifetime of electric bulbs for a random sample of 10 from a large consignment gave the following data :

Sample :	1	2	3	4	5	6	7	8	9	10
Life in '000 hours :	4.2	4.6	3.9	4.1	5.3	3.8	3.9	4.3	4.4	56

Can we accept the hypothesis that there is no significant difference in the sample mean and the hypothetical population mean.

Solution: Let us take the hypothesis that there is no significant difference in the sample mean lifetime of electric bulbs and the hypothetical population mean lifetime. Applying the t-test statistic, we get

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{4.4 - 4}{0.589/\sqrt{10}} = \frac{0.4}{0.186} = 2.150.$$

Table 10.8 Calculation of $\bar{x}$ and s

Lifetime of Bulbs (x)	$(x - \bar{x})$	$(x - \bar{x})^2$
4.2	−0.2	0.04
4.6	0.2	0.04
3.9	−0.5	0.25
4.1	−0.3	0.09
5.2	0.8	0.64
3.8	−0.6	0.36
3.9	−0.5	0.25
4.3	−0.1	0.01
4.4	0	0
5.6	1.2	1.44
44		3.12

where $\bar{x} = \frac{1}{n}\sum x_i = \frac{44}{10} = 4.4$ and $s = \sqrt{\frac{\sum(x - \bar{x})^2}{n-1}} = \sqrt{\frac{3.12}{9}} = 0.589$

Since, $t_{cal} = 2.150$ is less than its critical value, $t_\alpha = 2.262$ at $\alpha = 0.05$ level of significance and $df = 9$, the hypothesis is accepted. Hence, there is no significant difference in sample lifetime and population lifetime of bulbs.

Example 10.31: A random sample of 27 pair of observations from a normal population gives a correlation of 0.42. Is it likely that the variables in the popoulation are uncorrelated?

[*Delhi Univ., M.Com., 2003*]

Solution: Let us take the hypothesis that there is no significant difference in the sample correlation and correlation in the population. Applying t-test statistic, we get

$$t = \frac{r}{\sqrt{(1-r^2)/(n-2)}} = \frac{0.42}{\sqrt{\{1-(0.42)^2\}/(27-2)}} = \frac{0.42}{0.1815} = 2.31$$

where, $r = 0.42$ and $n = 27$

Since calculated value, $t_{cal} = 2.31$ is more than its table value $t_\alpha = 1.708$ at $\alpha = 0.05$ level of significance and $df = 27 - 2 = 25$, the null hypothesis is rejected. Hence, it is likely that the variables in the population are not correlated.

Example 10.32: Is a correlation coefficient of 0.5 significant which is obtained from a random sample of 11 pairs of values from a normal population? *[Madras Univ., B.Com., 2005]*

Solution: Let us take the null hypothesis that the given correlation coefficient is not sufficient. Applying t-test statistic, we get:

$$t = \frac{r}{\sqrt{(1-r^2)/(n-2)}} = \frac{0.50}{\sqrt{\{1-(0.5)^2\}/(11-2)}} = \frac{0.50}{0.288} = 1.736$$

where $r = 0.5$, $n = 11$

Since calculated value $t_{cal} = 1.732$ is less than its critical table value $t_\alpha = 2.26$ at $\alpha = 0.05$ level of significance and $df = 11 - 2 = 9$, the null hypothesis accepted. Hence, the given correaltion coefficient is not significant.

Example 10.33: A random sample of size 16 has 53 as mean. The sum of squares of the deviations taken from mean is 135. Can this sample be regarded as taken from the population having 56 as mean? Obtain 95% and 99% confidence limits of the mean of the population.

[Annamalai Univ., M.Com., 2003; CAFC, 2006]

Solution: Let us take the hypothesis that there is no signigicant difference between the sample mean and hypothetical population mean.

Applying t-test statistic, we get

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{53-56}{3/\sqrt{16}} = -\frac{3}{0.75} = -4$$

where $$s = \sqrt{\frac{\sum(x-\bar{x})}{n-1}} = \sqrt{\frac{135}{15}} = 3$$

Given, $\bar{x} = 53$, $\mu = 56$, $n = 16$, $\sum(x - \bar{x})^2 = 135$, $df = 16 - 1 = 15$

Since the calculated value of $t_{cal} = |-4| = 4$ is more than its table value, $t_\alpha = 2.13$ at $\alpha = 0.05$ level of significance and $df = 15$, the hypothesis is rejected. Hence, the sample has not come from a population having 56 as mean.

At 95% confidence limit of the population mean, we have

$$\bar{x} \pm \frac{s}{\sqrt{n}} t_{0.05} = 53 \pm \frac{3}{\sqrt{16}} \times 2.13 = 53 \pm 1.6, \text{ i.e., } 51.4 \le \bar{x} \le 54.6.$$

At 99% confidence limits of the population mean

$$\bar{x} \pm \frac{s}{\sqrt{n}} t_{0.05} = 53 \pm \frac{3}{\sqrt{16}} \times 2.95 = 53 \pm \frac{3}{4} \times 2.95 = 53 \pm 2.212,$$

i.e., $50.788 \le \bar{x} \le 55.212$.

10.10.3 Hypothesis Testing for Difference of Two Population Means (Independent Samples)

For comparing the mean values of two normally distributed populations, independent random samples of sizes n_1 and n_2 are drawn from the two populations. If μ_1 and μ_2 are the mean values of two populations, then aim is to estimate the value of the difference $\mu_1 - \mu_2$ between mean values of the two populations.

Since sample mean values $\bar{x}_1$ and $\bar{x}_2$ are the best point estimators to draw inferences about population mean values μ_1 and μ_2, respectively, therefore the difference between the sample means of the two independent simple random samples, $\bar{x}_1 - \bar{x}_2$, is the best point estimator of the difference, $\mu_1 - \mu_2$.

The sampling distribution of sample means values $\bar{x}_1$ and $\bar{x}_2$ has the following properties:

- Expected value: $E(\bar{x}_1 - \bar{x}_2) = E(\bar{x}_1) - E(\bar{x}_2) = \mu_1 - \mu_2$
- Variance: $Var(\bar{x}_1 - \bar{x}_2) = Var.(\bar{x}_1) + Var.(\bar{x}_2) = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$

If the population standard deviations, σ_1 and σ_2, are known, then the large sample interval estimation can also be used for the small sample case. But, if population standard deviations, σ_1 and σ_2, are unknown, then their values are estimated by the sample standard deviations s_1 and s_2, respectively. The t-distribution is used to develop a small sample interval estimate for $\mu_1 - \mu_2$.

Case 1: Population Variances Are Unknown But Equal

If population variances, σ_1^2 and σ_2^2, are unknown but equal, i.e., both populations have exactly the same shape and $\sigma_1^2 = \sigma_2^2 = \sigma^2$, then standard error of estimation (also called standard deviation) for sampling distribution of difference in two sample means $\bar{x}_1 - \bar{x}_2$ can be written as

$$\sigma_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} = \sqrt{\sigma^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} = \sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

In such a case, values of $\sigma_1{}^2$ and $\sigma_2{}^2$ are need to be estimated separately and therefore data of two samples can be combined to get a pooled, single estimate of σ^2. If sample variance, s^2, is used to estimate value of the population variance, σ^2, then its pooled variance estimator is given by

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

This single variance estimator, s^2, is referred as ***weighted average*** of the values of s_1^2 and s_2^2, where weights are represented by degrees of freedom $n_1 - 1$ and $n_2 - 1$, respectively. Thus, the point estimate of $\sigma_{\bar{x}_1 - \bar{x}_2}$ when $\sigma_1^2 = \sigma_2^2 = \sigma^2$ is given by

$$s_{\bar{x}_1 - \bar{x}_2} = \sqrt{s^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} = s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

Since $s_1 = \sqrt{\frac{\sum(x_1 - \bar{x}_1)^2}{(n_1 - 1)}}$ and $s_2 = \sqrt{\frac{\sum(x_2 - \bar{x}_2)^2}{(n_2 - 1)}}$, therefore the pooled variance, s^2, can also be calculated as

$$s^2 = \frac{\sum(x_1 - \bar{x}_1)^2 + \sum(x_2 - \bar{x}_2)^2}{n_1 + n_2 - 2}$$

Following the same reason as discussed earlier section, the t-test statistic is defined as

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

The sampling distribution of t-statistic is approximated by the t-distribution with $n_1 + n_2 - 2$ degrees of freedom.

Null hypothesis: $H_0: \mu_1 - \mu_2 = d_0$

Alternative hypothesis:

One-tailed Test	Two-tailed Test
$H_1: (\mu_1 - \mu_2) > d_0$ or $(\mu_1 - \mu_2) < d_0$	$H_1: \mu_1 - \mu_2 \neq d_0$

Decision Rule: Reject null hypothesis, H_0 at $df = n_1 + n_2 - 2$ and specified level of significance α when

One-tailed test	*Two-tailed test*
$t_{cal} > t_\alpha$ or $t_{cal} < -t_\alpha$	$t > t_{\alpha/2}$ or $t_{cal} < -t_{\alpha/2}$

Confidence Interval

The confidence interval estimate of the difference between populations means μ_1 and μ_2 for small samples with unknown σ_1 and σ_2 estimated by sample standard deviation, s_1 and s_2, is given by

$$(\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2}\, s_{\bar{x}_1 - \bar{x}_2}$$

where $t_{\alpha/2}$ is the critical value of *t-statistic* and its value depends on the t-distribution with $n_1 + n_2 - 2$ degrees of freedom and confidence coefficient, $1 - \alpha$.

Case 2: Population Variances Are Unknown and Unequal

When two population variances, σ_1^2 and σ_2^2 are not equal, then standard error, $\sigma_{\bar{x}_1 - \bar{x}_2}$ of sampling distribution of difference, $\bar{x}_1 - \bar{x}_2$, of two mean values can be estimated by sample variances, s_1^2 and s_2^2. Thus, an estimate of standard error of $\bar{x}_1 - \bar{x}_2$ is given by

$$s_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

The sampling distribution of t-test statistic

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{s_{\bar{x}_1 - \bar{x}_2}} = \frac{\bar{x}_1 - x_2}{s_{\bar{x}_1 - \bar{x}_2}}$$

is approximated by t-distribution with degrees of freedom given by

$$\text{Degrees of freedom } (df) = \frac{[s_1^2/n_1 + s_2^2/n_2]^2}{\dfrac{(s_1^2/n_1)^2}{n_1 - 1} + \dfrac{(s_2^2/n_2)^2}{n_2 - 1}}$$

Example 10.34: In a test given to two groups of students, the marks obtained out of 50 are as follows:

First group	:	18	20	36	50	49	36	34	49	41
Second group	:	29	28	26	35	30	44	46		

Examine the significance of the difference between the arithmetic mean of the marks secured by the students of two groups. [*Madras Univ., M.Com., 2003; MD Univ., M.Com., 2004*]

Solution: Let us take the null hypothesis that there is no significant difference in arithmetic mean of the marks secured by students of the two groups of students, that is,

$$H_0\colon \mu_1 - \mu_2 = 0 \quad \text{or} \quad \mu_1 = \mu_2 \quad \text{and} \quad H_1\colon \mu_1 \neq \mu_2$$

Since sample size in both the cases is small and sample variances are not known, apply t-test statistic to test the null hypothesis

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s}\sqrt{\frac{n_1 n_2}{n_1 + n_2}}$$

Calculations of sample means, $\bar{x}_1, \bar{x}_2$, and pooled sample standard deviations are shown in Table 10.9.

Table 10.9 Calculation for $\bar{x}_1$, $\bar{x}_2$ and s

First Group x_1	$x_1 - \bar{x}_1 = x_1 - 37$	$(x_1 - \bar{x}_1)^2$	*Second Group* x_2	$x_2 - \bar{x}_2 = x_2 - 34$	$(x_2 - \bar{x}_2)^2$
18	−19	361	29	−5	25
20	−17	389	28	−6	36
36	−1	1	26	−8	64
50	13	169	35	1	1
49	12	144	30	−4	16
36	−1	1	44	10	100
34	−3	9	46	12	144
49	12	144			
41	4	16			
333	0	1,234	238	0	386

$$\bar{x}_1 = \frac{\sum x_1}{n_1} = \frac{333}{9} = 37 \text{ and } \bar{x}_2 = \frac{\sum x_2}{n_2} = \frac{238}{7} = 34$$

$$s = \sqrt{\frac{\sum (x_1 - \bar{x}_1)^2 + \sum (x_2 - \bar{x}_2)^2}{n_1 + n_2 - 2}} = \sqrt{\frac{1234 + 386}{9 + 7 - 2}} = 10.76$$

Substituting values in the t-test statistic, we get

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s}\sqrt{\frac{n_1 n_2}{n_1 + n_2}} = \frac{37 - 34}{10.76}\sqrt{\frac{9 \times 7}{9 + 7}} = \frac{3}{10.46} \times 1.984 = 0.551$$

Degrees of freedom, $df = n_1 + n_2 - 2 = 9 + 7 - 2 = 14$.

Since at $\alpha = 0.05$ and $df = 14$, the calculated value t_{cal} $(= 0.551)$ is less than its critical value $t_{\alpha/2} = 2.14$, the null hypothesis H_0 is accepted. Hence, we conclude that there is no significant difference in arithmetic mean of the marks secured by students of the two groups of students.

Example 10.35: The mean life of a sample of 10 electric light bulbs was found to be 1456 hours with standard deviation of 423 hours. A second sample of 17 bulbs chosen from a different batch showed a mean life of 1280 hours with standard deviation of 398 hours. Is there a significant difference between the mean life of electric bulbs chosen from two samples. [*Delhi Univ., M.Com., 2003*]

Solution: Let us take the null hypothesis that there is no significant difference between the mean life of electric bulbs of two batches, that is,

$$H_0: \mu_1 = \mu_2 \quad \text{and} \quad H_1: \mu_1 \neq \mu_2$$

Given Batch 1: $n_1 = 10$, $\bar{x}_1 = 1456$, $s_1 = 423$; Batch 2: $n_2 = 17$, $\bar{x}_2 = 1280$, $s_2 = 398$ and $\alpha = 0.05$. Thus,

$$\text{Pooled standard deviation, } s = \sqrt{\frac{(n_1 - 1)\, s_1^2 + (n_2 - 1)\, s_2^2}{n_1 + n_2 - 2}} = \sqrt{\frac{9\,(423)^2 + 16\,(398)^2}{10 + 17 - 2}}$$

$$= \sqrt{\frac{16{,}10{,}361 + 25{,}34{,}464}{25}} = \sqrt{1{,}65{,}793} = 407.18$$

Applying the t- test statistic, we have

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s}\sqrt{\frac{n_1 n_2}{n_1 + n_2}} = \frac{1456 - 1280}{407.18}\sqrt{\frac{10 \times 17}{10 + 17}}$$

$$= \frac{176}{407.18} \times 2.51 = 1.085$$

Since the calculated value, $t_{cal} = 1.085$ is less than its critical value $t_{\alpha/2} = 2.06$ at $df = 25$ and $\alpha = 0.05$ level of significance, the null hypothesis is accepted. Hence, we conclude that the mean life of electric bulbs of two batches does not differ significantly.

Example 10.36: The manager of a courier service believes that packets delivered at the end of the month are heavier than those delivered early in the month. As an experiment, he weighed a random sample of 20 packets at the beginning of the month. He found that the mean weight was 5.25 kg with a standard deviation of 1.20 kg. Ten packets randomly selected at the end of the month had a mean weight of 4.96 kg and a standard deviation of 1.15 kg. At the 0.05 significance level, can it be concluded that the packets delivered at the end of the month weigh more?

Solution: Let us take the null hypothesis that the mean weight of packets delivered at the end of the month is more than the mean weight of packets delivered at the beginning of the month, that is

$$H_0 : \mu_E \geq \mu_B \quad \text{and} \quad H_1 : \mu_E < \mu_B$$

Given beginning: $n_1 = 20$, $\bar{x}_1 = 5.25$, $s_1 = 1.20$; and: $n_2 = 10$, $\bar{x}_2 = 4.96$, $s_2 = 1.15$. Thus,

$$\text{Pooled standard deviation, } s = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}} = \sqrt{\frac{19\times(5.25)^2 + 9(4.96)^2}{20 + 10 - 2}}$$

$$= \sqrt{\frac{19 \times 27.56 + 9 \times 24.60}{28}} = \sqrt{26.60} = 5.16$$

Applying the t-test statistics, we have

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s}\sqrt{\frac{n_1 n_2}{n_1 + n_2}} = \frac{5.25 - 4.96}{5.16}\sqrt{\frac{20\times 10}{20 + 10}}$$

$$= \frac{0.29}{5.16}\sqrt{\frac{200}{30}} = 0.056\times 2.58 = 0.145$$

Since at $\alpha = 0.01$ and $df = 28$, the calculated value t_{cal} (= 0.145) is less than it's critical value $z_\alpha = 1.701$, the null hypothesis is accepted. Hence, packets delivered at the end of the month weigh more on an average.

Example 10.37: Show how you would use Student's t-test and Fisher's z-test to decide whether the two sets of observations: 17, 27, 18, 25, 27, 29, 27, 23, 17 and 16, 16, 20, 16, 20, 17, 15, 21 indicate samples from the same universe.

Solution: Let x_1 and x_2 be variables of two sets of observations, respectivel. Calculations for mean and standard deviation of the two series are:

Table 10.10 Calculations for Mean and Standard

Set I, x_1	*$x_1 - A = x_1 - 23$*	*$(x_1 - A)^2$*	*Set II, x_2*	*$x_2 - B = x_2 - 16$*	*$(x_2 - B)^2$*
17	–6	36	16	0	0
27	4	16	16	0	0
18	–5	25	20	4	16
25	2	4	(16) ← B	0	0
27	4	16	20	4	16
29	6	36	17	1	1
27	4	16	15	–1	1
(23) ← A	0	0	21	5	25
17	–6	36			
210	3	185	141	13	59

$$\bar{x}_1 = \frac{1}{n_1}\Sigma x_1 = \frac{210}{9} = 23.333, \quad \text{and} \quad \bar{x}_2 = \frac{1}{n_2}\Sigma x_2 = \frac{141}{8} = 17.625.$$

Let A = 23 and B = 16 be the assumed mean of set I and set II series, respectively. Then

$$s_1^2 = \frac{\Sigma(x_1 - A)^2}{n_1 - 1} - \frac{\Sigma(x_1 - A)^2}{n_1(n_1 - 1)} = \frac{185}{8} - \frac{(3)^2}{9(8)} = \frac{185}{8} - \frac{1}{8} = \frac{184}{8} = 23$$

and $$s_2^2 = \frac{\Sigma(x_2 - B)^2}{n_2 - 1} - \frac{\Sigma(x_2 - B)^2}{n_2(n_2 - 1)} = \frac{59}{7} - \frac{(13)^2}{8(7)} = \frac{59}{7} - \frac{169}{56} = \frac{303}{56} = 5.411$$

Applying z-test, we get

$$z = \frac{1}{2}\log_e \frac{s_1^2}{s_2^2} = 1.1513 \log_{10} \frac{s_1^2}{s_2^2} = 1.1513 \log\left(\frac{23}{5.41}\right)$$

$$= 1.1513\,[\log 23 - \log 5.41] = 1.1513\,[1.3617 - 0.7333] = 0.724$$

At $df_1 = 8$ and $df_2 = 7$, $z_{0.05} = 0.657$ and $z_{0.01} = 0.961$.

Since calculated value of $z = 0.724$ is more than its table value $z_{0.05} = 0.657$ and less than $z_{0.01} = 0.961$, the variance ratio is significant at $\alpha = 0.05$ level of significance and not significant at $\alpha = 0.01$ level of significance. But at $\alpha = 0.01$ level of significance, the two population variances are the same.

Apply t-test to measure the significance of the difference between the population mean values. Let us take the null hyothesis H0 : $\mu_1 = \mu_2$. Then

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

$$= \frac{(23.333 - 17.625)\sqrt{72}}{(3.846)\sqrt{17}} = \frac{(5.708)(8.485)}{(3.846)(4.123)} = \frac{48.432}{15.857} = 3.054,$$

where $$s^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} = \frac{(9-1)(23) + (8-1)(5.411)}{9 + 8 - 2} = \frac{184 + 37.877}{15}$$

$$= \frac{221.877}{(15)} = 14.729$$

or $$s = \sqrt{14.792} = 3.846$$

$df = n_1 + n_2 - 2 = 9 + 8 - 2 = 15$, and $t_{0.05} = 2.131$ and $t_{0.01} = 2.947$.

Since calculated value $t_{cal} = 3.054$ is more than both of its table values $t_{0.05} = 0.657$ and $t_{0.01} = 0.961$, the null hypothesis is rejected. Hence, it may be concluded that the difference between the population means is significant, i.e., the two samples do not belong to the same population.

Example 10.38: Strength test carried out of samples of two yarns spun to the same count gave the following results:

		Sample size	*Sample Mean*	*Sample Variance*
Yarn A	:	4	52	42
Yarn B	:	9	42	56

The strengths are expressed in pounds. Is the difference in mean strengths significant of the sources from which the samples are drawn? *[Delhi Univ., MBA, 2007]*

Solution: Let us take the hypothesis that the difference in the mean strength of the two yarns is not significant. Applying t-test of difference of means,

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s}\sqrt{\frac{n_1 n_2}{n_1 + n_2}} = \frac{52 - 42}{7.22}\sqrt{\frac{4 \times 9}{4 + 9}} = \frac{10}{7.22} \times 1.664 = 2.3$$

where $s = \sqrt{\dfrac{(n_1 - 1)\, s_1^2 + (n_2 - 1)\, s_2^2}{n_1 + n_2 - 2}} = \sqrt{\dfrac{(4-1)\,42 + (9-1)\,56}{4+9-2}} = 7.22$

$\bar{x}_1 = 52,\ \bar{x}_2 = 42,\ s = 7.22,\ n_1 = 4,\ n_2 = 9$, and $df = 4 + 9 - 2 = 11$.

Since, the calculated value $t_{cal} = 2.3$ is more than its table value, $t_\alpha = 1.769$ at $\alpha = 0.05$ level of significance and $df = 11$, the hypthesis is rejected. Hence, the difference in the mean strength of the two yarns '*A*' and '*B*' is significant.

10.10.4 Hypothesis Testing for Difference of Two Population Means (Dependent Samples)

If two samples of same size are paired so that each observation in one sample is associated with any particular observation in the second sample, then 'difference' between each pair of data is first calculated and these differences are treated as a single set of data in order to consider whether there has been any significant change or whether the differences could have occurred by chance.

The matched sampling plan often leads to a smaller sampling error than the independent sampling plan because in matched samples variation as a source of sampling error is eliminated.

Let μ_d be the mean of the difference between two population mean values. The mean of the difference, μ_d, is compared to some hypothesized value using the *t*-test statistic for a single sample. The *t*-test statistic is used because the standard deviation of the population of differences is unknown, and therefore the statistical inference about μ_d based on the average of the sample differences $\bar{d}$ would involve the *t*-distribution rather than the standard normal distribution. The *t*-test statistic also called *paired t-test statistic* is written as

$$t = \frac{\bar{d} - \mu_d}{s_d / \sqrt{n}}$$

where n is number of paired observations; $df = n - 1$ is degrees of freedom; n is the number of pairs of differences and s_d is sample standard deviation of the distribution of the difference between the paired observations

$$s_d = \sqrt{\frac{\sum (d - \bar{d})^2}{n-1}} = \sqrt{\frac{\sum d^2}{n-1} - \frac{\left(\sum d\right)^2}{n\,(n-1)}}$$

The null and alternative hypotheses are stated as

H_0: $\mu_d = 0$ or c (Any hypothesized Value)

H_1: $\mu_d > 0$ or ($\mu_d < 0$) (One-tailed Test)

$\mu_d \neq 0$ (Two-tailed Test)

Decision Rule: If the calculated value, t_{cal}, is less than its critical value, t_d at a specified level of significance and known degrees of freedom, then null hypothesis H_0 is accepted. Otherwise H_0 is rejected.

Confidence Interval

The confidence interval estimate of the difference between two population means is given by

$$\bar{d} \pm t_{\alpha/2} \frac{s_d}{\sqrt{n}}$$

where $t_{\alpha/2}$ = critical value of *t*-test statistic at $n - 1$ degrees of freedom and α level of significance.

If the claimed value of null hypothesis H_0 lies within the confidence interval, then H_0 is accepted, otherwise rejected.

Example 10.39: The HR manager wishes to see if there has been any change in the ability of trainees after a specific training programme. The trainees take an aptitude test before the start of the programme and an equivalent one after they have completed it. The scores recorded are given below. Has any change taken place at 5 per cent significance level?

Trainee	:	A	B	C	D	E	F	G	H	I
Score before training	:	75	70	46	68	68	43	55	68	77
Score after training	:	70	77	57	60	79	64	55	77	76

Solution: Let us take the null hypothesis that there is no change that has taken place after the training, that is,

$$H_0: \mu_d = 0 \quad \text{and} \quad H_1: \mu_d \neq 0$$

Calculations for 'changes' are computed as shown in Table 10.11.

Table 10.11 Calculations of 'Changes'

Trainee	*Before Training*	*After Training*	*Difference in Scores, d*	d^2
A	75	70	5	25
B	70	77	7	49
C	46	57	–11	121
D	68	60	8	64
E	68	79	–11	121
F	43	64	–21	441
G	55	55	0	0
H	68	77	–9	81
I	77	76	1	1
			–45	903

$$\bar{d} = \frac{1}{n}\sum d = \frac{-45}{9} = -5 \text{ and}$$

$$s_d = \sqrt{\frac{\sum d^2}{n-1} - \frac{\left(\sum d\right)^2}{n(n-1)}} = \sqrt{\frac{903}{8} - \frac{(-45)^2}{9 \times 8}}$$

$$= \sqrt{112.87 - 28.13} = 9.21$$

Applying the t-test statistic, we have

$$t = \frac{\bar{d} - \mu_d}{s_d/\sqrt{n}} = \frac{-5-0}{9.21/\sqrt{9}} = -\frac{5}{3.07} = -1.63$$

Since the calculated value $t_{cal} = -1.63$ is more than its critical value, $t_{\alpha/2} = -2.31$, at $df = 8$ and $\alpha/2 = 0.025$ the null hypothesis is accepted. Hence, we conclude that there is no change in the ability of trainees after the training.

Example 10.40: Twelve students were given intensive coaching and 5 tests were conducted in a month. The scores of tests 1st and 5th are given below.

No. of students	:	1	2	3	4	5	6	7	8	9	10	11	12
Marks in 1st test	:	50	42	51	26	35	42	60	41	70	55	62	38
Marks in 5th test	:	62	40	61	35	30	52	68	51	84	63	72	50

Do the data indicate any improvement in the scores obtained in the first and fifth tests.

[*Punjab Univ., M.Com., 2007*]

Solution: Let us take the hypothesis that there is no improvement in the scores obtained in the first and fifth tests, that is,

$$H_0: \mu_d = 0 \quad \text{and} \quad H_1: \mu_d \neq 0$$

Calculations for 'changes' are as shown in Table 10.12:

Table 10.12 Calculations of 'Changes'

No. of Students	*Marks in 1st Test*	*Marks In 5th Test*	*Difference in Marks d*	*d^2*
1	50	62	12	144
2	42	40	−2	4
3	51	61	10	100
4	26	35	9	81
5	35	30	−5	25
6	42	52	10	100
7	60	68	8	64
8	41	51	10	100
9	70	84	14	196
10	55	63	8	64
11	62	72	10	100
12	38	50	12	144
			96	1122

$$\bar{d} = \frac{1}{n}\sum d = \frac{96}{12} = 8 \text{ and}$$

$$s_d = \sqrt{\frac{\sum d^2}{n-1} - \frac{\left(\sum d\right)^2}{n(n-1)}} = \sqrt{\frac{1122}{11} - \frac{(96)^2}{12 \times 11}} = 5.673$$

Applying the t-test statistic, we have

$$t = \frac{\bar{d} - \mu_d}{s_d / \sqrt{n}} = \frac{8}{5.673/\sqrt{12}} = 4.885$$

Since the calculated value $t_{cal} = 4.885$ is more than its critical value, $t_{\alpha/2} = 2.20$ at $df = 11$ and $\alpha/2 = 0.025$, the null hypothesis is rejected. Hence, we conclude that there is an improvement in the scores obtained in two tests.

Example 10.41: To test the desirability of a certain modification in typist's desks, 9 typists were given two tests of almost same nature, one on the desk in use and the other on the new type. The following difference in the number of words typed per minute was recorded:

Typists	:	A	B	C	D	E	F	G	H	I
Increase in number of words	:	2	4	0	3	−1	4	−3	2	5

Do the data indicate that the modification in desk increases typing speed?

Solution: Let us take the hypothesis that there is no change in typing speed with the modification in the typing desk, that is,

$$H_0: \mu_d = 0 \text{ and } H_1: \mu_d > 0$$

The calculations for 'changes' are shown in Table 10.13.

Table 10.13 Calculations of 'Changes'

Typist	*Increase in Number of Words d*	d^2
A	2	4
B	4	16
C	0	0
D	3	9
E	−1	1
F	4	16
G	−3	9
H	2	4
I	5	25
	16	84

$$\bar{d} = \frac{1}{n}\sum d = \frac{16}{9} = 1.778$$

and
$$s_d = \sqrt{\frac{\sum d^2}{n-1} - \frac{\left(\sum d\right)^2}{n\,(n-1)}} = \sqrt{\frac{84}{8} - \frac{(16)^2}{9 \times 8}} = 2.635$$

Applying the t-test statistic, we have

$$t = \frac{\bar{d} - \mu_d}{s_d / \sqrt{n}} = \frac{1.778}{2.625/\sqrt{9}} = 2.025$$

Since the calculated value $t_{cal} = 2.025$ is less than its critical value $t_\alpha = 2.306$ at $df = 8$ and $\alpha = 0.05$ significance level, the null hypothesis is accepted. Hence, we conclude that there is no change in typing speed with the modification in the desk.

Example 10.42: Ten pairs of maize plants were grown in parallel boxes and one member of each pair was treated by receiving a small electric current. The difference in heights between the treated and untreated in m.m. was:

(Treated) – (untreated) : 6.0, 1.3, 10.2, 23.9, 3.1, 6.8, –1.5, –14.7, –3.3 and 11.1.

Test whether the small electric current affected the growth of maize seedlings.

Solution: Let us take the null hypothesis that the electric current does not affect the growth, that is, $H_0 : \mu = 0$ and $H_1 = \mu \neq 0$

Calculations for s and $\bar{d}$ are shown in Table 10.14:

Table 10.14 Calculations for s and $\bar{d}$

Difference in Heights, d_i	$d_i - \bar{d} = d_i - 4.29$	$(d_i - \bar{d})^2$
6.0	1.71	2.924
1.3	−2.99	8.940
10.2	5.91	34.928
23.9	19.61	384.552
3.1	−1.19	1.416
6.8	2.51	6.300
−1.5	−5.79	33.524
−14.7	−18.99	360.620
−3.3	−7.50	57.608
11.1	6.81	46.376
42.9		937.188

$$\bar{d} = \frac{1}{n}\sum d_i = \frac{42.9}{10} = 4.29$$

$$s^2 = \frac{1}{n-1}\sum (d_i - \bar{d})^2 = \frac{937.188}{10-1} = \frac{937.188}{9} = 104.132$$

or

$$s = \sqrt{(104.132)} = 10.205 \text{ approx.}$$

Applying t-test statistic, we get

$$t = \frac{\bar{d}}{s/\sqrt{n}} = \frac{\bar{d}\sqrt{n}}{s} = \frac{4.29\sqrt{(10)}}{10.205} = \frac{4.29\,(3.162)}{10.205} = 1.329\,.$$

Since $t_{cal} = 1.329$ is less than its critical value $t_\alpha = 2.16$ at $\alpha = 0.05$ level of significance and $df = 9$, the hypothesis is accepted. Thus, sample does not provide satisfactory evidence that the electric current affects the growth of maize seedlings.

Self-practice Problems 10C

10.20 Ten oil tins are taken at random from an automatic filling machine. The mean weight of the tins is 15.8 kg and the standard deviation is 0.50 kg. Does the sample mean differ significantly from the intended weight of 16 kg? *[Delhi Univ., MBA, 2006]*

10.21 Nine items of a sample had the following values: 45, 47, 50, 52, 48 47, 49, 53 and 50. The mean is 49 and the sum of the square of the deviation from mean is 52. Can this sample be regarded as taken from the population having 47 as mean? Also obtain 95 per cent and 99 per cent confidence limits of the population mean. *[Delhi Univ., MBA, 2004]*

10.22 A random sample of size 16 has 53 as mean. The sum of the squares of the deviations taken from mean is 135. Can this sample be regarded as taken from the population having 56 as mean? Obtain 95 per cent and 99 per cent confidence limits of the mean of the population. *[Madras Univ., M.Com., 2006]*

10.23 A drug manufacturer has installed a machine which automatically fills 5 gm of drug in each phial. A random sample of fills was taken and it was found to contain 5.02 gm on an average in a phial. The standard deviation of the sample was 0.002 gm. Test at 5 per cent level of significance if the adjustment in the machine is in order. *[Delhi Univ., MBA, 2005]*

10.24 Two salesmen A and B are working in a certain district. From a sample survey conducted by the Head Office, the following results were obtained. State whether there is any significant difference in the average sales between the two salesmen.

		Salesman	
		A	*B*
No. of samples	:	20	18
Average sales (₹ in thousand)	:	170	205
Standard deviation (₹ in thousand)	:	20	25

[Delhi Univ., MBA, 2000; Kumaon Univ., MBA, 2060]

10.25 The means of two random samples of sizes 9 and 7 are 196.42 and 198.82 respectively. The sum of the squares of the deviations from the mean is 26.94 and 18.73 respectively. Can the sample be considered to have been drawn from the same normal population? *[Delhi Univ., M.Com., 2004]*

10.26 A random sample of 12 families in one city showed an average monthly food expenditure of ₹1380 with a standard deviation of ₹100 and a random sample of 15 families in another city showed an average monthly food expenditure of ₹1320 with a standard deviation of ₹120. Test whether the difference between the two means is significant at $\alpha = 0.01$ level of significance of $\alpha = 0.01$. *[AIMA Diploma in Mgt., 2003; Delhi Univ., MBA, 2000]*

10.27 You are given the following data about the life of two brands of bulbs:

	Mean life	*Standard deviation*	*Sample size*
Brand A	2000 hrs	250 hrs	12
Brand B	2230 hrs	300 hrs	15

Do you think there is a significant difference in the quality of the two brands of bulbs? *[Delhi Univ., MBA, 2004]*

10.28 An IQ test was administered to 5 persons before and after they were trained. The results are given below:

Candidate		*I*	*II*	*III*	*IV*	*V*
IQ before training	:	110	120	123	132	125
IQ after training	:	120	118	.25	136	121

Test whether there is any change in IQ level after the training programme. *[Delhi Univ., M.Com, 2005]*

10.29 Eleven sales executive trainees are assigned selling jobs right after their recruitment. After a fortnight they are withdrawn from their field duties and given a month's training for executive sales. Sales executed by them in thousands of rupees before and after the training, in the same period are listed below:

Sales Before Training	*Sales After Training*
23	24
20	19
19	21
21	18
18	20
20	22
18	20
17	20
23	23
16	20
19	27

Do these data indicate that the training has contributed to their performance?

[*Delhi Univ., M.Com., 2006*]

Hints and Answers

10.20 Let H_0: $\mu = 16$ and H_1: $\mu \neq 16$

Given $n = 10$, $\bar{x} = 15.8, s = 0.50.$,

Applying t-test statistic,

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{15.8 - 16}{0.50/\sqrt{10}} = -1.25$$

Since $t_{cal} = -1.25 >$ critical value $t_{\alpha/2} = -2.262$, at $df = 9$ and $\alpha/2 = 0.025$, the null hypothesis is accepted.

10.21 Let H_0: $\mu = 27$ and H_1: $\mu \neq 47$

Given $\bar{x} = 49$, $\sum(x - \bar{x})^2 = 52$, $n = 9$, and

$$s = \sqrt{\frac{\Sigma(x-\bar{x})^2}{(n-1)}} = \sqrt{\frac{52}{8}} = 2.55.$$

Applying t-test statistic,

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{49 - 47}{2.55/\sqrt{9}} = 2.35$$

Since $t_{cal} = 2.35 >$ critical value $t_{\alpha/2} = 2.31$ at $\alpha/2 = 0.025$, $df = 8$, the null hypothesis is rejected.

10.22 Let H_0: $\mu = 56$ and H_1: $\mu \neq 46$

Given: $n = 16$, $\bar{x} = 53$ and $\sum(x - \bar{x})^2 = 135$. Thus

$$s = \sqrt{\frac{\Sigma(x-\bar{x})^2}{(n-1)}} = \sqrt{\frac{135}{15}} = 3$$

Applying t-test statistic, $t = \dfrac{\bar{x} - \mu}{s/\sqrt{n}} = \dfrac{53 - 56}{3/\sqrt{16}} = -4$

Since $t_{cal} = -4 <$ critical value $t_{\alpha/2} = -2.13$ at $\alpha/2 = 0.025$, $df = 15$, the null hypothesis is rejected.

10.23 Let H_0: $\mu = 5$ and H_1: $\mu \neq 5$

Given $n = 10$, $\bar{x} = 5.02$ and $s = 0.002$.

Applying t-test statistic,

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{5.02 - 5}{0.002/\sqrt{10}} = 33.33$$

Since $t_{cal} = 33.33 >$ critical value $t_{\alpha/2} = 1.833$ at $\alpha/2 = 0.025$ and $df = 9$, the null hypothesis is rejected.

10.24 Let H_0: $\mu_1 = \mu_2$ and H_1: $\mu_1 \neq \mu_2$

Given $n_1 = 20$, $s_1 = 20$, $= 170$; $n_2 = 18$, $s_2 = 25$, $\bar{x}_2 = 205$,

Applying t-test statistic,

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s/\sqrt{n}} \sqrt{\frac{n_1 n_2}{n_1 + n_2}}$$

$$= \frac{170 - 205}{22.5} \sqrt{\frac{20 \times 18}{20 + 18}}$$

$$= \frac{-35}{22.5} \sqrt{\frac{360}{38}} = -4.8$$

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

$$= \sqrt{\frac{19\,(20)^2 + 17\,(25)^2}{20 + 18 - 2}}$$

$$= \sqrt{\frac{18{,}225}{36}} = 22.5$$

Since $t_{cal} = -4.8 <$ critical value $t_{\alpha/2} = -1.9$ at $\alpha/2 = 0.025$ and $df = n_1 + n_2 - 2 = 36$, the null hypothesis is rejected.

10.25 Let H_0: $\mu_1 = \mu_2$ and H_1: $\mu_1 \neq \mu_2$

Given $n_1 = 9$, $\bar{x}_1 = 196.42$, $\sum(x_1 - \bar{x}_1)^2 = 26.94$ and $n_2 = 7$, $\bar{x}_2 = 198.82$ and $\sum(x_2 - \bar{x}_2)^2 = 18.73$

Applying t-test statistic,

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s} \sqrt{\frac{n_1 n_2}{n_1 + n_2}} = \frac{52 - 42}{7.224} \sqrt{\frac{4 \times 9}{4 + 9}}$$

$$= -\frac{2.40}{1.81} \sqrt{\frac{63}{16}} = -2.63$$

$$= \sqrt{\frac{\sum(x_1-\bar{x}_1)^2+(x_2-\bar{x}_2)^2}{n_1+n_2-2}}$$

$$= \sqrt{\frac{26.94+18.73}{9+7-2}} = 1.81$$

Since $t_{cal} = -2.63 <$ critical value $t_{\alpha/2} = -2.145$ at $\alpha/2 = 0.025$ and $df = 14$, the null hypothesis is rejected.

10.26 Let H_0: $\mu_1 = \mu_2$ and H_1: $\mu_1 \neq \mu_2$

Given $n_1 = 12$, $s_1 = 100$, $\bar{x}_1 = 1380$ and $n_2 = 15$, $s_2 = 120$, $\bar{x}_2 = 1320$.

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

$$= \sqrt{\frac{11\,(100)^2+14\,(120)^2}{12+15-2}} = 111.64$$

Applying t-test statistic,

$$t = \frac{\bar{x}_1-\bar{x}_2}{s}\sqrt{\frac{n_1 n_2}{n_1+n_2}}$$

$$= \frac{1380-1320}{111.64}\sqrt{\frac{12\times 15}{12+15}}$$

$$= \frac{60}{111.64}\sqrt{\frac{180}{27}} = 1.39$$

Since $t_{cal} = 1.39 <$ critical value $t_{\alpha/2} = 2.485$ at $\alpha/2 = 0.025$ and $df = n_1 + n_2 - 2 = 25$, the null hypothesis is accepted.

10.27 Let H_0: $\mu_1 = \mu_2$ and H_1: $\mu_1 \neq \mu_2$ (Two-tailed test)

Given $n_1 = 12$, $\bar{x}_1 = 2000$, $s_1 = 250$ and $n_2 = 15$, $\bar{x}_2 = 2230$, $s_2 = 300$.

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

$$= \sqrt{\frac{11\,(250)^2+14\,(300)^2}{12+15-2}} = 279.11$$

Applying t-test statistic,

$$t = \frac{\bar{x}_1-\bar{x}_2}{s}\sqrt{\frac{n_1 n_2}{n_1+n_2}} = \frac{2000-2230}{279.11}\sqrt{\frac{12\times 15}{12+15}}$$

$$= -\frac{230}{279.11}\sqrt{\frac{180}{27}} = -2.126$$

Since $t_{cal} = -2.126 <$ critical value $t_{\alpha/2} = -1.708$ at $\alpha/2 = 0.025$ and $df = n_1 + n_2 - 2 = 25$, the null hypothesis is rejected.

10.28 Let H_0: No change in IQ level after training programme

IQ Level		d	d^2
Before	*After*		
110	120	10	100
120	118	−2	4
123	125	2	4
132	136	4	16
125	121	−4	16
		10	140

$$\bar{d} = \frac{1}{n}\sum d = \frac{10}{5} = 2$$

$$s = \sqrt{\frac{\sum d^2}{n-1}-\frac{(\sum d)^2}{n\,(n-1)}} = \sqrt{\frac{140}{4}-\frac{(10)^2}{5\times 4}}$$

$$= 5.477$$

Applying t-test statistic,

$$t = \frac{\bar{d}}{s/\sqrt{n}} = \frac{2}{5.477/\sqrt{5}} = 0.817$$

Since $t_{cal} = 0.814 <$ critical value $t_{\alpha/2} = 4.6$ at $\alpha/2 = 0.025$ and $df = 4$, the null hypothesis is accepted.

10.29 Let H_0: Training did not improve the performance of the sales executives

Difference in Sales, d	d^2
1	1
−1	1
2	4
−3	9
2	4
2	4
2	4
3	9
0	0
4	16
8	64
20	116

$$\bar{d} = \frac{1}{n}\sum d = \frac{20}{11} = 1.82$$

$$s = \sqrt{\frac{\sum d^2}{n-1}-\frac{(\sum d)^2}{n\,(n-1)}} = \sqrt{\frac{116}{10}-\frac{(20)^2}{11\times 10}} = 2.82$$

Applying t-test statistic,

$$t = \frac{\bar{d}}{s/\sqrt{n}} = \frac{1.82}{2.82/\sqrt{11}} = 2.14$$

Since $t_{cal} = 2.14 <$ critical value $t_{\alpha/2} = 2.23$ at $\alpha/2 = 0.025$ and $df = 10$, the null hypothesis is accepted.

10.11 HYPOTHESIS TESTING BASED ON F-DISTRIBUTION

Often population variances are required to be compared in certain statistical applications such as (i) product quality resulting from two different production processes; (ii) temperatures for two heating devices; (iii) assembly times for two assembly methods, and (iv) rate of return from investment in two types of stocks, and so on.

When independent random samples of size n_1 and n_2 are drawn from two normally distributed populations, the ratio

F-test: A hypothesis test for comparing the variance of two independent populations with the help of variances of two small samples.

$$F = \frac{s_1^2/\sigma_1^2}{s_2^2/\sigma_2^2}$$

follows F-distribution with degrees of freedom $df_1 = n_1 - 1$ and $df_2 = n_2 - 1$, where s_1^2 and s_2^2 are two sample variances and are given by

$$s_1^2 = \frac{\sum(x_1 - \bar{x}_1)^2}{n_1 - 1} \text{ and } s_2^2 = \frac{\sum(x_2 - \bar{x}_2)^2}{n_2 - 1}$$

If two normally distributed populations have equal variances, i.e. $\sigma_1^2 = \sigma_2^2$, then the ratio

$$F = \frac{s_1^2}{s_2^2}; \; s_1 > s_2$$

follows a probability distribution known as F-distribution with $n_1 - 1$ degrees of freedom for numerator and $n_2 - 1$ degrees of freedom for denominator. For computational purposes, a larger sample variance is placed in the numerator so that ratio is always equal to or more than one.

Assumptions

1. Independent random samples are drawn from each of two normally distributed populations.
2. The amount of variability in the two populations is same and can be measured by a common variance σ^2, i.e., ${\sigma^2}_1 = {\sigma^2}_2 = \sigma^2$.

The F-distribution, also called ***variance ratio distribution***, is not symmetric and the shape of any F-distribution depends on the degrees of freedom of the numerator and denominator. A typical graph of F-distribution is shown in Fig. 10.9 for equal degrees of freedom for both numerator and denominator.

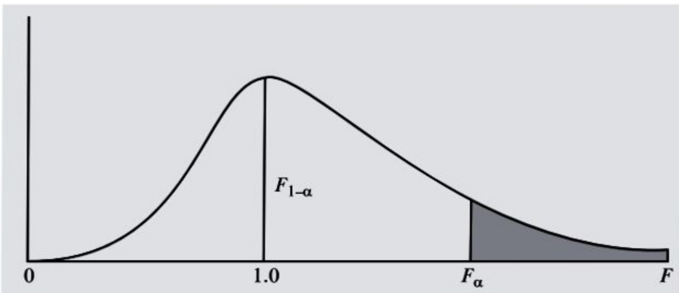


Figure 10.9
F-distribution for *n* Degrees of Freedom

10.11.1 Properties of F-distribution

1. The total area under F-distribution curve is unity. The value of F-test statistic denoted by F_α at a particular level of significance, α, indicates area (or probability) of α to the right of F_α value.
2. The F-distribution is positively skewed with a range 0 to ∞ and the degree of skewness decreases with the increase in degrees of freedom ν_1 for numerator and ν_2 for denominator. For $\nu_2 \geq 30$, F-distribution is approximately normal.

3. The sample variances s_1^2 and s_2^2 are the unbiased estimates of population variance. Since $s_1 > s_2$, the range of F-distribution curve is from 0 to ∞.
4. If the ratio s_1^2/s_2^2 is nearly equal to 1, then it indicates that σ_1^2 and σ_2^2 are unequal. On the other hand, a very large or very small value for s_1^2/s_2^2 indicates difference in the population variances.
5. The F-distribution credited to Sir Ronald Fisher is written as

$$z = \frac{1}{2}\log_e F = \frac{1}{2}\log_e \frac{s_1^2}{s_2^2} = \frac{1}{2}\log_e\left(\frac{s_1}{s_2}\right)^2$$

$$= \log_e\left(\frac{s_1}{s_2}\right) = \log_{10}\left(\frac{s_1}{s_2}\right)\log_e 10 = 2.3026\log_{10}\left(\frac{s_1}{s_2}\right)$$

The probability density function of F-distribution is given by

$$f(\text{F}) \text{ or } y = y_0 \frac{e^{\nu_1 z}}{(\nu_1 e^{2Z} + \nu_2)^{-(\nu_1+\nu_2)/2}}, -\infty < z < \infty$$

where y_0 is a constant and depends on the degrees of freedom, $v_1 = n_1 - 1$ and $v_2 = n_2 - 1$, such that the area under the curve is unity.

6. The mean and variances of the F-distribution are

$$\text{Mean } \mu = \frac{\nu_2}{\nu_2 - 2}, v_2 > 2$$

and
$$\text{Variance } \sigma^2 = \frac{2\nu_2^2(\nu_1 + \nu_2 - 2)}{\nu_1(\nu_2 - \nu_1)^2(\nu_2 - 4)}, \nu_2 > 4$$

This implies that F-distribution has no mean for $\nu_2 \le 2$ and no variance for $\nu_2 \le 4$.

7. The reciprocal property

$$F_{1-\alpha\,(v2,\,v1)} = \frac{1}{F_{\alpha\,(\nu_2,\,\nu_1)}}$$

of F-distribution helps to identify corresponding lower (left) tail F-values from the given upper (right) tail F- values. For example, if $\alpha = 0.05$, then for $v_1 = 24$ and $v_2 = 15$, the value of $F_{0.95(15,\,24)} = 2.11$. Thus $F_{0.95(24,\,15)} = 1/2.11 = 0.47$.

8. The variance of F-distribution with 1 and n degrees of freedom is same as t-distribution with n degrees of freedom.

10.11.2 Comparing Two Population Variances

The null hypothesis is stated as follows without taking into consideration the value of the ratio s_1^2/s_2^2:

Null Hypothesis	*Alternative Hypothesis*	
$H_0 : \sigma_1^2 = \sigma_2^2$	$H_1 : \sigma_1^2 > \sigma_2^2$ or $\sigma_1^2 < \sigma_2^2$	(One-tailed Test)
$H_0 : \sigma_1^2 = \sigma_2^2$	$H_1 : \sigma_1^2 \neq \sigma_2^2$	(Two-tailed Test)

The F-test statistic $= s_1^2/s_2^2$, where s_1^2 and s_2^2 are sample variances is based on random samples of size n_1 and n_2 drawn from two populations 1 and 2, respectively.

Decision Rules: If calculated value of F-test statistic is less than its critical value $F_{\alpha(v1,\,v2)}$ with $df_1 = n_1 - 1$ (numerator), and $df_2 = n_2 - 1$(denominator) for one-tailed test, then accept null hypothesis, H_0.

The population with larger variance is considered as population 1 because any population with larger sample variance always places the ratio s_1^2/s_2^2 in the right-tail direction, so that area of rejection of H_0 falls in the right (upper) tail of the F-distribution curve.

Confidence Interval

An interval estimate for all possible values of population variances ratio, σ_1^2/σ_2^2, is given by

$$\frac{s_1^2/s_2^2}{F_{(1-\alpha)}} \le \frac{\sigma_1^2}{\sigma_2^2} \le \frac{s_1^2/s_2^2}{F_\alpha}$$

where critical value of F is based on a F-distribution with degrees of freedom $(n_1 - 1)$, $(n_2 - 1)$ and $(1 - \alpha)$ confidence coefficient.

Example 10.43: A research was conducted to understand whether women have a greater variation in attitude on political issues than men. Two independent samples of 31 men and 41 women were used for the study. The sample variances so calculated were 120 for women and 80 for men. Test whether the difference in attitude toward political issues is significant at 5 per cent level of significance.

Solution: Let us take the hypothesis that the difference in attitude toward political issues is significant, that is,

$$H_0\colon \sigma_w^2 = \sigma_m^2 \text{ and } H_1\colon \sigma_w^2 > \sigma_m$$

Applying F-test statistic, we get $F = \frac{s_1^2}{s_2^2} = \frac{120}{80} = 1.50$

Since variance for women is in the numerator, F-distribution with $df_1 = 41 - 1 = 40$ (numerator) and $df_2 = 31 - 1 = 30$ (denominator) is used to conduct one-tailed test.

Since calculated value of F(= 1.50) is less than its critical value of $F_{\alpha = 0.05} = 1.79$ at $df_1 = 40$ and $df_2 = 30$, the null hypothesis is accepted. Hence, we conclude that women have a greater variation in attitude on political issues than men.

Example 10.44: The following figures relate to the number of units of an item produced per shift by two workers A and B for a number of days

A:	19	22	24	27	24	18	20	19	25		
B:	26	37	40	35	30	30	40	26	30	35	45

Can it be inferred that worker A is more stable compared to worker B? Answer using the F-test at 5 per cent level of significance.

Solution: Let us take the hypothesis that the two workers are equally stable, that is,

$$H_0\colon \sigma_A^2 = \sigma_B^2 \text{ and } H_1\colon \sigma_A^2 > \sigma_B^2$$

The calculations for population variances σ_A^2 and σ_B^2 of the number of units produced by workers A and B, respectively, are shown in Table 10.15.

Table 10.15 Calculation of σ_A^2 and σ_B^2

Worker A x_1	$x_1 - \bar{x}_1 = x_1 - 22$	$(x_1 - 22)^2$	*Worker B* x_2	$x_2 - \bar{x}_2 = x_2 - 34$	$(x_2 - \bar{x}_2)^2$
19	−3	9	26	−8	64
22	0	0	37	3	9
24	2	4	40	6	36
27	5	25	35	1	1
24	2	4	30	−4	16
18	−4	16	30	−4	16
20	−2	4	40	6	36
19	−3	9	26	−8	64
25	3	9	30	−4	16
			35	1	1
			45	11	121
198	0	80	374	0	380

$$\bar{x}_1 = \frac{\sum x_1}{n_1} = \frac{198}{9} = 22; \ \bar{x}_2 = \frac{\sum x_2}{n_2} = \frac{374}{11} = 34$$

$$s_A^2 = \frac{\sum (x_1 - \bar{x}_1)^2}{n_1 - 1} = \frac{80}{9-1} = 10;$$

$$s_B^2 = \frac{\sum (x_2 - \bar{x}_2)^2}{n_2 - 1} = \frac{380}{11-1} = 38$$

Applying F-test statistic, we have

$$F = \frac{s_B^2}{s_A^2} = \frac{38}{10} = 3.8 \text{ (because } s_B^2 > s_A^2)$$

Since calculated value of F is more than its critical value $F_{0.05(10,\ 8)} = 3.35$ at $\alpha = 5$ per cent level of significance and degrees of freedom $df_A = 8$, $df_B = 10$, the null hypothesis is rejected. Hence we conclude that worker A is more stable than worker B, because $s_A^2 < s_B^2$.

Example 10.45: It is known that the mean diameters of rivets produced by two firms, A and B, are practically the same, but the standard deviations may differ. For 22 rivets produced by firm A, the standard deviation is 2.9 m while for 16 rivets manufactured by firm B, the standard deviations 3.8 m.

Compute the statistic you would require to test whether the products of firm A have the same variability as those of firm B.

Solution: Let us denote the variates of the two series by x_1 and x_2.

For the firm *A*, $\sigma_{x1} = 2.9, n_1 = 22$

For the firm *B*, $\sigma_{x2} = 3.8, n_2 = 16$

Thus, $s_1^2 = \frac{n_1}{n_1 - 1}\sigma_{x1}^2 = \frac{22}{22-1}(2.9)^2 = \frac{185.02}{21} = 8.81$

$$s_2^2 = \frac{n_2}{n_2 - 1}\sigma_{x2}^2 = \frac{16}{16-1}(3.8)^2 = \frac{231.04}{15} = 15.403$$

$$F = \frac{s_2^2}{s_1^2} = \frac{15.403}{8.81} = 1.75 \text{ approx, because } s_2^2 > s_1^2$$

$df, v_1 = n_1 - 1 = 22 - 1 = 21$ and $df, v_2 = n_2 - 1 = 16 - 1 = 15$.

At 15 degree of freedom in the numerator and 21 degree of freedom in the denominator, the table value of $F_{0.05} = 2.18$ and $F_{0.01} = 3.04$. Since the calculated value of $F = 1.75$ is less than both of $F_{0.05}$ and $F_{0.01}$, result is significant both at 5% and 1% level.

Hence it may be concluded that the products of firm *A* are of better quality than those of firm *B*.

Self-practice Problems 10D

10.30 The mean diameter of a steel pipe produced by two processes, A and B, is practically the same but the standard deviations may differ. For a sample of 22 pipes produced by A, the standard deviation is 2.9 m, while for a sample of 16 pipes produced by B, the standard deviation is 3.8 m. Test whether the pipes produced by process A have the same variability as those of process B.

10.31 Tests for breaking strength, were carried out on two lots of 5 and 9 steel wires, respectively. The variance of one lot was 230 and that of the other was 492. Is there a significant difference in their variability?

10.32 Two random samples drawn from normal population are:

Sample 1	*Sample 2*
20	27
16	33
26	42
27	35

23	32
22	34
18	38
24	28
25	41
19	43
	30
	37

Obtain estimates of the variances of the population and test whether the two populations have the same variance.

10.33 In a sample of 8 observations, the sum of the squared deviations of items from the mean was 94.50. In another sample of 10 observations the value was found to be 101.70. Test whether the difference is significant at 5 per cent level of significance (at 5 per cent level of significance critical value of F for $v_1 = 3$ and $v_2 = 9$ degrees of freedom is 3.29 and for $v_1 = 8$ and $v_2 = 10$ degrees of freedom, its value is 3.07).

10.34 Most individuals are aware of the fact that the average annual repair costs for an automobile depends on the age of the automobile. A researcher is interested in finding out whether the variance of the annual repair costs also increases with the age of the automobile. A sample of 25 automobiles that are 4 years old showed a sample variance for annual repair cost of ₹850 and a sample of 25 automobiles that are 2 years old showed a sample variance for annual repair costs of ₹300. Test the hypothesis that the variance in annual repair costs is more for the older automobiles, for a 0.01 level of significance.

10.35 The standard deviation in the 12-month earnings per share for 10 companies in the software industry was 4.27 and the standard deviation in the 12-month earning per share for 7 companies in the telecom industry was 2.27. Conduct a test for equal variance at $\alpha = 0.05$. What is you conclusion about the variability in earning per share for two industries.

Hints and Answers

10.30 Let H_0 : There is no difference in the variability of diameters produced by process A and B, i.e.

$H_0 : \sigma_A^2 = \sigma_B^2$ and $H_1 : \sigma_A^2 \neq \sigma_B^2$

Given $\sigma_A = 2.9, n_1 = 22, df_A = 21; \sigma_B = 3.8,$
$n_2 = 16, df_B = 21.$

$$s_A^2 = \frac{n_1}{n_1 - 1}\sigma_A^2 = \frac{22}{22-1}(2.9)^2 = \frac{22}{21}(8.41) = 8.81$$

$$s_2^2 = \frac{n_2}{n_2 - 1}\sigma_B^2 = \frac{16}{16-1}(3.8)^2 = \frac{16}{15}(14.44) = 15.40$$

$$F = \frac{s_2^2}{s_1^2} = \frac{15.40}{8.81} = 1.75$$

Since the calculated value $f = 1.75$ is less than its critical value $F_{0.05\,(15,\,21)} = 2.18$, the null hypothesis is accepted.

10.31 Let H_0: No significant variability in the breaking strength of wires

Given $n_1 = 5, \sigma_1{}^2 = 230, df_1 = 4; n_2 = 9, \sigma_2{}^2 = 492,$ $df_2 = 8$

$$F = \frac{\sigma_2^2}{\sigma_1^2} = \frac{492}{230} = 2.139$$

Since calculated value F = 2.139 is less its critical value $F_{0.05(8,4)} = 6.04$ the null hypothesis is accepted.

10.32 Let H_0 : Two populations have the same variance, i.e.

H_0: $\sigma_1^2 = \sigma_2^2$ and H_1: $\sigma_1^2 \neq \sigma_2^2$.

Sample 1: $\bar{x}_1 = \frac{\sum x_1}{10} = 22;$

$$s_1^2 = \frac{\sum (x_1 - \bar{x}_1)^2}{n_1 - 1} = \frac{120}{9} = 13.33, df_1 = 9$$

Sample 2: $x_2 = \frac{\sum x_1}{12} = 35;$

$$s_2^2 = \frac{\sum (x_2 - \bar{x}_2)^2}{n_2 - 1} = \frac{314}{11} = 28.54, df_2 = 11$$

$$F = \frac{s_2^2}{s_1^2} = \frac{28.54}{13.33} = 2.14$$

Since calculated value F = 2.14 is less than its critical value $F_{0.05\,(11,\,9)} = 4.63$, the null hypothesis is accepted.

10.33 Let H_0: The difference is not significant

Sample 1: $n_1 = 8, \sum(x_1 - \bar{x}_1)^2 = 94.50, v_1 = 7$

Sample 2: $n_1 = 10, \sum(x_2 - \bar{x}_2)^2 = 101.70, v_2 = 9$

$$\therefore \quad s_1^2 = \frac{\sum (x_1 - \bar{x}_1)^2}{n_1 - 1} = \frac{94.50}{7} = 13.5;$$

$$s_2^2 = \frac{\sum (x_2 - \bar{x}_2)^2}{n_2 - 1} = \frac{101.70}{9} = 11.3$$

$$F = \frac{s_1^2}{s_2^2} = \frac{13.5}{11.3} = 1.195$$

Since the calculated value F = 1.195 is less than its critical value $F_{0.05\,(7,\,9)} = 3.29$, the null hypothesis is accepted.

10.34 Let H_0 : No significant difference in the variance of repair cost, $H_0: \sigma_1^2 = \sigma_2^2$ and $H_1: \sigma_1^2 > \sigma_2^2$

s_1^2 = ₹850; s_2^2 = ₹300

$n_1 = 25, df_1 = 24; n_2 = 25, df_2 = 24$

$$F = \frac{s_1^2}{s_2^2} = \frac{850}{300} = 2.833$$

Since the calculated value F = 2.833 is more than its critical value $F_{0.01(24, 24)} = 2.66$, the null hypothesis is rejected.

10.35 Let H_0 : No significant difference of variability in earning per share for two industries,

$H_0: \sigma_1^2 = \sigma_2^2$ and $H_1: \sigma_1^2 \neq \sigma_2^2$

Software industry:

$s_1^2 = (4.27)^2 = 18.23,$

$n_1 = 10, df_1 = 9$

Telecom industry:

$s_2^2 = (2.27)^2 = 5.15,$

$n_2 = 7, df_2 = 6$

$$\therefore \quad F = \frac{s_1^2}{s_2^2} = \frac{18.23}{5.15} = 3.54$$

Since the calculated value F = 3.54 is less than its critical value $F_{0.05\ (9,\ 6)} = 4.099$, the null hypothesis is accepted.

Formulae Used

1. Hypothesis testing for population mean with large sample (n> 30)

(a) Test statistic about a population mean μ

- σ assumed known, $z = \dfrac{|\bar{x} - \mu|}{\sigma/\sqrt{n}}$
- σ is estimated by s, $z = \dfrac{\bar{x} - \mu}{s/\sqrt{n}}$

(b) Test statistic for the difference between means of two populations

- Standard deviation of $\bar{x}_1 - \bar{x}_2$ when σ_1 and σ_2 are known

$$\sigma_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

$$\text{Test statistic } z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sigma_{\bar{x}_1 - \bar{x}_2}}$$

- Standard deviation of $\bar{x}_1 - \bar{x}_2$ when $\sigma_1^2 = \sigma_2^2$

$$\sigma_{\bar{x}_1 - \bar{x}_2} = \sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

- Point estimator of $\sigma_{\bar{x}_1 - \bar{x}_2}$

$$s_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

- Interval estimation for single population mean

$\bar{x} \pm z_{\alpha/2}\,\sigma_{\bar{x}}$; σ is known

$\bar{x} \pm z_{\alpha/2}\, s_{\bar{x}}$; σ is unknown

- Interval estimation for the difference of means of two populations

$(\bar{x}_1 - \bar{x}_2) \pm z_{\alpha/2}\,\sigma_{\bar{x}_1 - \bar{x}_2}$; σ_1 and σ_2 are known

$(\bar{x}_1 - \bar{x}_2) \pm z_{\alpha/2}\, s_{\bar{x}_1 - \bar{x}_2}$; σ_1 and σ_2 are unknown

2. Hypothesis testing for population proportion for large sample (n> 30)

(a) Test statistic for population proportion p

$$z = \frac{\bar{p} - p}{\sigma_{\bar{p}}}; \quad \sigma_{\bar{p}} = \sqrt{\frac{p\,(1 - p)}{n}}$$

(b) Test statistic for the difference between the proportions of two populations

- Standard deviation of $\bar{p}_1 - \bar{p}_2$

$$\sigma_{\bar{p}_1 - \bar{p}_2} = \sqrt{\frac{p_1(1 - p_1)}{n_1} + \frac{p_2(1 - p_2)}{n_2}}$$

- Point estimator of

$$s_{\bar{p}_1 - \bar{p}_2} = \sqrt{\frac{\bar{p}_1(1 - \bar{p}_1)}{n_1} + \frac{\bar{p}_2(1 - \bar{p}_2)}{n_2}}$$

- Interval estimation of the difference between the proportions of two populations

$$(\bar{p}_1 - \bar{p}_2) \pm z_{\alpha/2}\, s_{\bar{p}_1 - \bar{p}_2}$$

where all n_1p_1, $n_1(1 - p_1)$, n_2p_2 and $n_2(1 - p_2)$ are more than or equal to 5

- Test statistic for hypothesis testing about the difference between proportions of two populations

$$z = \frac{(\bar{p}_1 - \bar{p}_2) - (p_1 - p_2)}{\sigma_{\bar{p}_1 - \bar{p}_2}}$$

- Pooled estimator of the population proportion

$$\bar{p} = \frac{n_1\bar{p}_1 + n_2\bar{p}_2}{n_1 + n_2}$$

- Point estimator of $\sigma_{\bar{p}_1 - \bar{p}_2}$

$$s_{\bar{p}_1 - \bar{p}_2} = \sqrt{\bar{p}(1 - \bar{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

3. Hypothesis testing for population mean with small sample ($n \le 30$)
   - Test statistic when s is estimated by s

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

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

$$= \sqrt{\frac{\sum (x_1 - \bar{x}_1)^2 + (x_2 - \bar{x}_2)^2}{n_1 + n_2 - 2}}$$

   - Test statistic for difference between the means of two population proportions

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{s_{\bar{x}_1 - \bar{x}_2}}$$

where $s_{\bar{x}_1 - \bar{x}_2}$ is the point estimator of $\sigma_{\bar{p}_1 - \bar{p}_2}$

when $\sigma_1^2 = \sigma_2^2 = \sigma^2$ and

$$s_{\bar{x}_1 - \bar{x}_2} = \sqrt{s^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} = s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

   - Interval estimation of the difference between means of two populations $(\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2}\, s_{\bar{x}_1 - \bar{x}_2}$

4. Hypothesis testing for matched samples (small sample case)

   Test statistic for matched samples

$$t = \frac{\bar{d} - \mu_d}{s_d/\sqrt{n}};\; s_d = \sqrt{\frac{\sum (d - \bar{d})^2}{n - 1}};\; \bar{d} = \frac{\sum d}{n}$$

5. Hypothesis testing for two population variances

$$F = s_1^2/s_2^2;\; s_1^2 > s_2^2$$

Chapter Concepts Quiz

True or False

1. [T] [F] A tentative assumption about a population parameter is called the null hypothesis.
2. [T] [F] As a general guideline, a research hypothesis should be stated as the alternative hypothesis.
3. [T] [F] The equality part of the expression (either ≥, ≤, or =) always appears in the null hypothesis.
4. [T] [F] Type I error is the probability of accepting null hypothesis when it is true.
5. [T] [F] Type II error is the probability of accepting null hypothesis when it is true.
6. [T] [F] The probability of making a Type I error is referred to as the level of significance.
7. [T] [F] The estimated standard deviation of sampling distribution of a statistic is called standard error.
8. [T] [F] Type I error is more harmful than Type II error.
9. [T] [F] For a given level of significance, we can reduce β by increasing the sample size.
10. [T] [F] If the cost of Type I error is large, a small level of significance should be specified.
11. [T] [F] For a given sample size n, an attempt to reduce the level of significance results in an increase in β.
12. [T] [F] For a given level of significance, change in sample size changes the critical value.
13. [T] [F] The t-test statistic is used when $n \le 30$ and the population standard deviation is known.
14. [T] [F] The value of the test statistic that defines the rejection region is called critical region for the test.

Multiple Choice Questions

15. Sampling distribution will be approximately normal if sample size is
    (a) large (b) sufficiently large
    (c) small (d) none of these
16. Critical region is a region of
    (a) rejection (b) acceptance
    (c) indecision (d) none of these
17. The probability of Type II error is
    (a) α (b) β
    (c) 1 – α (d) 1 – β
18. The term 1 – β is called
    (a) level of the test (b) power of the test
    (c) size of the test (d) none of these
19. The test statistic to test $\mu_1 = \mu_2$ for normal population is
    (a) F-test (b) z-test
    (c) t-test (d) none of these
20. The usual notation the standard error of the sampling distribution is
    (a) $\sigma/\sqrt{n}$ (b) σ/n
    (c) σ^2 (d) none of these
21. If Type I and Type II errors are fixed, then the power of a test increases with
    (a) the increase of sample size
    (b) not related to sample size