Indira Gandhi
National Open University
School of Sciences

MST-004
STATISTICAL INFERENCE

Block

3

TESTING OF HYPOTHESIS

Curriculum and Course Design Committee

Prof. K. R. Srivathasan
Pro-Vice Chancellor
IGNOU, New Delhi

Prof. Parvin Sinclair
Pro-Vice Chancellor
IGNOU, New Delhi

Prof. Geeta Kaicker
Director, School of Sciences
IGNOU, New Delhi

Prof. Jagdish Prasad
Department of Statistics
University of Rajasthan, Jaipur

Prof. R. M. Pandey
Department of Bio-Statistics
All India Institute of Medical Sciences
New Delhi

Prof. Rahul Roy
Math. and Stat. Unit
Indian Statistical Institute, New Delhi

Dr. Diwakar Shukla
Department of Mathematics and Statistics
Dr. Hari Singh Gaur University, Sagar

Prof. Rakesh Srivastava
Department of Statistics
M.S. University of Baroda, Vadodara

Prof. G. N. Singh
Department of Applied Mathematics
I.S.M., Dhanbad

Dr. Gulshan Lal Taneja
Department of Mathematics
M.D. University, Rohtak

Faculty members of School of Sciences, IGNOU

Statistics
Dr. Neha Garg
Dr. Nitin Gupta
Mr. Rajesh Kaliraman
Dr. Manish Trivedi

Mathematics
Dr. Deepika
Prof. Poornima Mital
Prof. Sujatha Varma
Dr. S. Venkataraman

Block Preparation Team

Dr. Ramkishan **(Editor)**
Department of Statistics
D. A. V. (PG) College
C.C. S. University, Merrut

Dr. Parmod Kumar **(Language Editor)**
School of Humanities, IGNOU

Dr. Diwakar Shukla **(Units 1 - 4)**
Department of Mathematics and Statistics
Dr. Hari Singh Gaur University, Sagar

Mr. Prabhat Kumar Sangal **(Units 1 - 4)**
School of Sciences, IGNOU

Course Coordinator: Mr. Prabhat Kumar Sangal
Programme Coordinator: Dr. Manish Trivedi

Block Production

Mr. Sunil Kumar, AR (P),School of Sciences, IGNOU
CRC prepared by Mr. Prabhat Kumar Sangal, School of Sciences, IGNOU

Acknowledgement: I gratefully acknowledge my colleague Mr. Rajesh Kaliraman, Statistics Discipline, School of Sciences for their great support.

July, 2013

ISBN-978-81-266-

Further information on the Indira Gandhi National Open University may be obtained from University's Office at Maidan Garhi, New Delhi-110068 or visit University's website http://www.ignou.ac.in

Printed and published on behalf of the Indira Gandhi National Open University, New Delhi by the Director, School of Sciences.

Printed at:

TESTING OF HYPOTHESIS

In previous block of this course, we have discussed one part of statistical inference, that is, estimation and we have learnt how we estimate the unknown population parameter(s) by using **point estimation** and **interval estimation**. In this block we will focus on the second part of statistical inference which is known as **testing of hypothesis**.

In this block, you will study the basic concept of testing of hypothesis and different kinds of well-known tests. At the end of this block you will be aware of the idea, procedure and applications of the hypothesis testing. This block contains four units.

Unit 9: Concepts of Testing of Hypothesis

This unit explains the fundamental aspect relating to testing of hypothesis with examples. It describes basic concepts and methodologies as hypothesis, critical region, types of errors, level of significance, general procedure of testing a hypothesis, concept of p-value, relation between confidence interval and testing of hypothesis.

Unit 10: Large Sample Tests

This unit explores the procedure of testing the hypothesis based on one sample and two samples when sample size is large and taken from normal or non-normal population(s). In this unit, you will learn about Z-test for testing of hypothesis about mean, difference of two means, proportion, difference of two proportions, variance and equality of two variances.

Unit 11: Small Sample Tests

This unit is devoted to test a hypothesis based on one sample and two samples when sample sizes are small. In this unit, you will learn about t-test for testing of hypothesis about mean and difference of two means. This unit also explores the idea for testing the hypothesis for equality of two means when samples are dependent and testing of hypothesis about population correlation coefficient.

Unit 12: Chi-square and F Tests

The last unit of this block describes the chi-square test for population variance and F-test for equality of two population variances.

Notations and Symbols

$X_1, X_2, \ldots, X_n$	:	Random sample of size n
H_0	:	Null hypothesis
H_1 or H_A	:	Alternative hypothesis
ω	:	Critical (rejection) region
$\overline{\omega}$	:	Non-rejection region
α	:	Size of critical region or type-I error or level of significance
β	:	Type-II error
$1-\beta$	:	Power of the test
z_α	:	Critical value of Z-test at α level of significance
$t_{(\nu), \alpha}$	:	Critical value of t-test with ν degrees of freedom at α level of significance
$\chi^2_{(\nu), \alpha}$	:	Critical value of χ^2-test with ν degrees of freedom at α level of significance
$F_{(\nu_1, \nu_2), \alpha}$	:	Critical value of F-test with (ν_1, ν_2) degrees of freedom at α level of significance

UNIT 9 CONCEPTS OF TESTING OF HYPOTHESIS

Structure

9.1 INTRODUCTION

In previous block of this course, we have discussed one part of statistical inference, that is, estimation and we have learnt how we estimate the unknown population parameter(s) by using **point estimation** and **interval estimation**. In this block, we will focus on the second part of statistical inference which is known as **testing of hypothesis**.

In our day-to-day life, we see different commercials advertisements in television, newspapers, magazines, etc. such as

(i) The refrigerator of certain brand saves up to 20% electric bill,

(ii) The motorcycle of certain brand gives 60 km/liter mileage,

(iii) A detergent of certain brand produces the cleanest wash,

(iv) Ninety nine out of hundred dentists recommend brand A toothpaste for their patients to save the teeth against cavity, etc.

Now, the question may arise in our mind "can such types of claims be verified statistically?" Fortunately, in many cases the answer is "yes".

The technique of testing such type of claims or statements or assumptions is known as testing of hypothesis. The truth or falsity of a claim or statement is never known unless we examine the entire population. But practically it is not possible in mostly situations so we take a random sample from the population under study and use the information contained in this sample to take the decision whether a claim is true or false.

This unit is divided into 11 sections. Section 9.1 is introductory in nature. In Section 9.2, we defined the hypothesis. The concept and role of critical region in testing of hypothesis is described in Section 9.3. In Section 9.4, we explored the types of errors in testing of hypothesis whereas level of significance is explored in Section 9.5. In Section 9.6, we explored the types of tests in testing of hypothesis. The general procedure of testing a hypothesis is discussed in Section 9.7. In Section 9.8, the concept of p-value in decision making about the null hypothesis is discussed whereas the relation between confidence interval and testing of hypothesis is discussed in Section 9.9. Unit ends by providing summary of what we have discussed in this unit in Section 9.10 and solution of exercises in Section 9.11.

Objectives

After reading this unit, you should be able to:

- define a hypothesis;
- formulate the null and alternative hypotheses;
- explain what we mean by type-I and type-II errors;
- explore the concept of critical region and level of significance;
- define one-tailed and two-tailed tests;
- describe the general procedure of testing a hypothesis;
- concept of p-value; and
- test a hypothesis by using confidence interval.

Before coming to the procedure of testing of hypothesis, we will discuss the basis terms used in this procedure one by one in subsequent sections.

9.2 HYPOTHESIS

As we have discussed in previous section that in our day-to-day life, we see different commercials advertisements in television, newspapers, magazines, etc. and if someone may be interested to test such type of claims or statement then we come across the problem of testing of hypothesis. For example,

(i) a customer of motorcycle wants to test whether the claim of motorcycle of certain brand gives the average mileage 60 km/liter is true or false,

(ii) the businessman of banana wants to test whether the average weight of a banana of Kerala is more than 200 gm,

(iii) a doctor wants to test whether new medicine is really more effective for controlling blood pressure than old medicine,

(iv) an economist wants to test whether the variability in incomes differ in two populations,

(v) a psychologist wants to test whether the proportion of literates between two groups of people is same, etc.

In all the cases discussed above, the decision maker is interested in making inference about the population parameter(s). However, he/she is not interested in estimating the value of parameter(s) but he/she is interested in testing a claim or statement or assumption about the value of population parameter(s). Such claim or statement is postulated in terms of hypothesis.

In statistics, a hypothesis is a statement or a claim or an assumption about the value of a population parameter (e.g., mean, median, variance, proportion, etc.).

Similarly, in case of two or more populations a hypothesis is comparative statement or a claim or an assumption about the values of population parameters. (e.g., means of two populations are equal, variance of one population is greater than other, etc.). The plural of hypothesis is hypotheses.

In hypothesis testing problems first of all we should being identifying the claim or statement or assumption or hypothesis to be tested and write it in the words. Once the claim has been identified then we write it in symbolical form if possible. As in the above examples,

(i) Customer of motorcycle may write the claim or postulate the hypothesis "the motorcycle of certain brand gives the average mileage 60 km/liter." Here, we are concerning the **average** mileage of the motorcycle so let μ represents the average mileage then our hypothesis becomes $\mu = 60$ km / liter.

(ii) Similarly, the businessman of banana may write the statement or postulate the hypothesis "the average weight of a banana of Kerala is greater than 200 gm." So our hypothesis becomes $\mu > 200$ gm.

(iii) Doctor may write the claim or postulate the hypothesis " the new medicine is really more effective for controlling blood pressure than old medicine." Here, we are concerning the **average** effect of the medicines so let μ_1 and μ_2 represent the average effect of new and old medicines respectively on controlling blood pressure then our hypothesis becomes $\mu_1 > \mu_2$.

(iv) Economist may write the statement or postulate the hypothesis " the variability in incomes differ in two populations." Here, we are concerning the **variability** in income so let σ_1^2 and σ_2^2 represent the variability in incomes in two populations respectively then our hypothesis becomes $\sigma_1^2 \neq \sigma_2^2$.

(v) Psychologist may write the statement or postulate the hypothesis "the proportion of literates between two groups of people is same." Here, we are concerning the **proportion** of literates so let P_1 and P_2 represent the proportions of literates of two groups of people respectively then our hypothesis becomes $P_1 = P_2$ or $P_1 - P_2 = 0$.

The hypothesis is classified according to its nature and usage as we will discuss in subsequent subsections.

9.2.1 Simple and Composite Hypotheses

In general sense, if a hypothesis specifies only one value or exact value of the population parameter then it is known as simple hypothesis. And if a hypothesis specifies not just one value but a range of values that the population parameter may assume is called a composite hypothesis.

A hypothesis which completely specifies parameter(s) of a theoretical population (probability distribution) is called a simple hypothesis otherwise called composite hypothesis.

As in the above examples, the hypothesis postulated in (i) $\mu = 60$ km/liter is simple hypothesis because it gives a single value of parameter ($\mu = 60$) whereas the hypothesis postulated in (ii) $\mu > 200$ gm is composite hypothesis because it does not specify the exact average value of weight of a banana. It may be 260, 350, 400 gm or any other.

Similarly, (iii) $\mu_1 > \mu_2$ or $\mu_1 - \mu_2 > 0$ and (iv) $\sigma_1^2 \neq \sigma_2^2$ or $\sigma_1^2 - \sigma_2^2 \neq 0$ are not simple hypotheses because they specify more than one value as $\mu_1 - \mu_2 = 4$, $\mu_1 - \mu_2 = 7$, $\sigma_1^2 - \sigma_2^2 = 2, \sigma_1^2 - \sigma_2^2 = -5$, etc. and (v) $P_1 = P_2$ or $P_1 - P_2 = 0$ is simple hypothesis because it gives a single value of parameter as $P_1 - P_2 = 0$.

9.2.2 Null and Alternative Hypotheses

As we have discussed in last page that in hypothesis testing problems first of all we identify the claim or statement to be tested and write it in symbolical form. After that we write the complement or opposite of the claim or statement in symbolical form. In our example of motorcycle, the claim is $\mu = 60$ km/liter then its complement is $\mu \neq 60$ km/liter. In (ii) the claim is $\mu > 200$ gm then its complement is $\mu \leq 200$ gm. If the claim is $\mu < 200$ gm then its complement is $\mu \geq 200$ gm. The claim and its complement are formed in such a way that they cover all possibility of the value of population parameter.

We state the null and alternative hypotheses in such a way that they cover all possibility of the value of population parameter.

Once the claim and its compliment have been established then we decide of these two which is the null hypothesis and which is the alternative hypothesis. The thump rule is that the statement containing equality is the null hypothesis. That is, the hypothesis which contains symbols $=$ or $\leq$ or $\geq$ is taken as null hypothesis and the hypothesis which does not contain equality i.e. contains $\neq$ or $<$ or $>$ is taken as alternative hypothesis. The null hypothesis is denoted by H_0 and alternative hypothesis is denoted by H_1 or H_A.

In our example of motorcycle, the claim is $\mu = 60$ km/liter and its complement is $\mu \neq 60$ km/liter. Since claim $\mu = 60$ km/liter contains equality sign so we take it as a null hypothesis and complement $\mu \neq 60$ km/liter as an alternative hypothesis, that is,

$$H_0: \mu = 60 \text{ km/liter and } H_1: \mu \neq 60 \text{ km/liter}$$

In our second example of banana, the claim is $\mu > 200$ gm and its complement is $\mu \leq 200$ gm. Since complement $\mu \leq 200$ gm contains equality sign so we take complement as a null hypothesis and claim $\mu > 200$ gm as an alternative hypothesis, that is,

$$H_0: \mu \leq 200 \text{ gm and } H_1: \mu > 200 \text{ gm}$$

Formally these hypotheses are defined as

The hypothesis which we wish to test is called as the null hypothesis.

According to Prof. R.A. Fisher,

“A null hypothesis is a hypothesis which is tested for possible rejection under the assumption that it is true.”

The hypothesis which complements to the null hypothesis is called alternative hypothesis.

Note 1: Some authors use equality sign in null hypothesis instead of $\geq$ and $\leq$ signs.

The alternative hypothesis has two types:

(i) Two-sided (tailed) alternative hypothesis

(ii) One-sided (tailed) alternative hypothesis

If the alternative hypothesis gives the alternate of null hypothesis in both directions (less than and greater than) of the value of parameter specified in null hypothesis then it is known as two-sided alternative hypothesis and if it gives an alternate only in one direction(less than or greater than) only then it is known as one-sided alternative hypothesis. For example, if our alternative hypothesis is $H_1: \theta \neq 60$ then it is a two-sided alternative hypothesis because its means that the value of parameter θ is greater than or less than 60. Similarly, if $H_1: \theta > 60$ then it is a right-sided alternative hypothesis because its means that the value of parameter θ is greater than 60 and if $H_1: \theta < 60$ then it is a left-sided alternative hypothesis because its means that the value of parameter θ is less than 60.

In testing procedure, we assume that the null hypothesis is true until there is sufficient evidence to prove that it is false. Generally, the hypothesis is tested with the help of a sample so evidence in testing of hypothesis comes from a sample. If there is enough sample evidence to suggest that the null hypothesis is false then we reject the null hypothesis and support the alternative hypothesis. If the sample fails to provide us sufficient evidence against the null hypothesis we are not saying that the null hypothesis is true because here, we take the decision on the basis of a random sample which is a small part of the population. To say that null hypothesis is true we must study all observations of the population under study. For example, if someone wants to test that the person of India has two hands then to prove that this is true we must check all the persons of India whereas to prove that it is false we require a person he / she has one hand or no hand. So we can only say that there is not enough evidence against the null hypothesis.

Note 2: When we assume that null hypothesis is true then we are actually assuming that the population parameter is equal to the value in the claim. In our example of motorcycle, we assume that $\mu = 60$ km/liter whether the null hypothesis is $\mu = 60$ km/liter or $\mu \leq 60$ km/liter or $\mu \geq 60$ km/liter.

Now, you can try the following exercises.

E1) A company manufactures car tyres. Company claims that the average life of its tyres is 50000 miles. To test the claim of the company, formulate the null and alternative hypotheses.

E2) Write the null and alternative hypotheses in case (iii), (iv) and (v) of our example given in Section 9.2.

E3) A businessman of orange formulates different hypotheses about the average weight of the orange which are given below:

(i) $H_0: \mu = 100$ (ii) $H_1: \mu > 100$ (iii) $H_0: \mu \leq 100$ (iv) $H_1: \mu \neq 100$

(v) $H_1: \mu > 150$ (vi) $H_0: \mu = 130$ (vii) $H_1: \mu \neq \mu_0$

Categorize the above cases into simple and composite hypotheses.

After describing the hypothesis and its types our next point in the testing of hypothesis is critical region which will be described in next section.

9.3 CRITICAL REGION

As we have discussed in Section 9.1 that generally, null hypothesis is tested by the sample data. Suppose $X_1, X_2, \ldots, X_n$ be a random sample drawn from a population having unknown population parameter θ. The collection of all possible values of $X_1, X_2, \ldots, X_n$ is a set called sample space(S) and a particular value of $X_1, X_2, \ldots, X_n$ represents a point in that space.

In order to test a hypothesis, the entire sample space is partitioned into two disjoint sub-spaces, say, ω and $S - \omega = \overline{\omega}$. If calculated value of the test statistic lies in ω, then we reject the null hypothesis and if it lies in $\overline{\omega}$, then we do not reject the null hypothesis. The region ω is called a **"rejection region or critical region"** and the region $\overline{\omega}$ is called a **"non-rejection region"**. Therefore, we can say that

A statistic is a function of sample observations (not including parameter). Basically, a test statistic is a statistic which is used in decision making about the null hypothesis.

"A region in the sample space in which if the calculated value of the test statistic lies, we reject the null hypothesis then it is called critical region or rejection region."

This can be better understood with the help of an example. Suppose, 100 students appear in total 10 papers two of each in English, Physics, Chemistry, Mathematics and Computer Science of a Programme. Suppose the scores in these papers are denoted by $X_1, X_2, \ldots, X_{10}$ and maximum marks =100 for each paper. For obtaining the distinction award in this Programme, a student needs to have total score equal to or more than 750 which is a rule.

Suppose, we select one student randomly out of 100 students and we want to test that the selected student is a distinction award holder. So we can take the null and alternative hypotheses as

H_0: Selected student is a distinction award holder

H_1: Selected student is not a distinction award holder

For taking the decision about the selected student, we define a statistic $T_{10} = \sum_{i=1}^{10} X_i$ as the sum of the scores in all the 10 papers of the student. The range of T_{10} is $0 \le T_{10} \le 1000$. Now, we divide the whole space (0-1000) into two regions as no-distinction awarded region (less than 750) and distinction awarded region (greater than or equal to 750) as shown in Fig. 9.1. Here, 750 is the critical value which separates the no-distinction and distinction awarded regions.

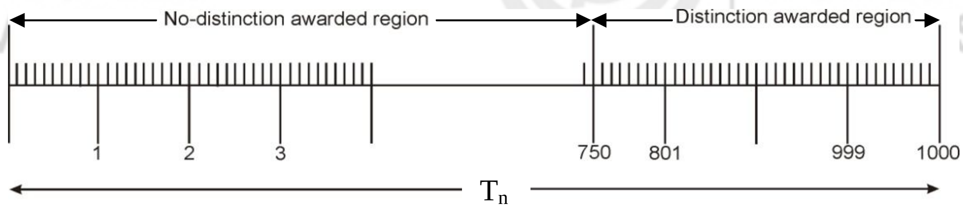

Fig. 9.1: Non-rejection and critical regions for distinction award

On the basis of scores in all the papers of the selected student, we calculate the value of the statistic $T_{10} = \sum_{i=1}^{10} X_i$. And calculated value may fall in distinction award region or not, depending upon the observed value of test statistic.

For making a decision to reject or do not reject H_0, we use test statistic $T_{10} = \sum_{i=1}^{10} X_i$ (sum of scores of 10 papers). If calculated value of test statistic T_{10} lies in no-distinction awarded region (critical region), that is, $T_{10} < 750$ then we reject H_0 and if calculated value of test statistic T_{10} lies in distinction awarded region (non-rejection region), that is, $T_{10} \ge 750$ then we do not reject H_0. It is a basic structure of the procedure of testing of hypothesis which needs two regions like:

(i) Region of rejection of null hypothesis H_0

(ii) Region of non-rejection of null hypothesis H_0

The point of discussion in this test procedure is "**how to fix the cut off value 750"**? What is the justification for this value? The distinction award region may be like $T_{10} \ge 800$ or at $T_{10} \ge 850$ or at $T_{10} \ge 900$. So, there must be a scientific justification for the cut-off value 750. In a statistical test procedure it is obtained by using the probability distribution of the test statistic.

The region of rejection is called critical region. It has a pre-fixed area generally denoted by α, corresponding to a cut-off value in a probability distribution of test statistic.

The rejection (critical) region lies in one-tail or two-tails on the probability curve of sampling distribution of the test statistic its depends upon the alternative hypothesis. Therefore, three cases arise:

Case I: If the alternative hypothesis is right-sided such as $H_1: \theta > \theta_0$ or $H_1: \theta_1 > \theta_2$ then the entire critical or rejection region of size α lies on right tail of the probability curve of sampling distribution of the test statistic as shown in Fig. 9.2.

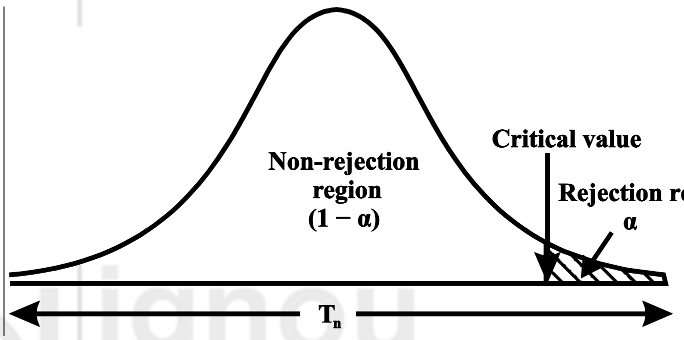

Fig. 9.2

Critical value is a value or values that separate the region of rejection from the non-rejection region.

Case II: If the alternative hypothesis is left-sided such as $H_1: \theta < \theta_0$ or $H_1: \theta_1 < \theta_2$ then the entire critical or rejection region of size α lies on left tail of the probability curve of sampling distribution of the test statistic as shown in Fig. 9.3.

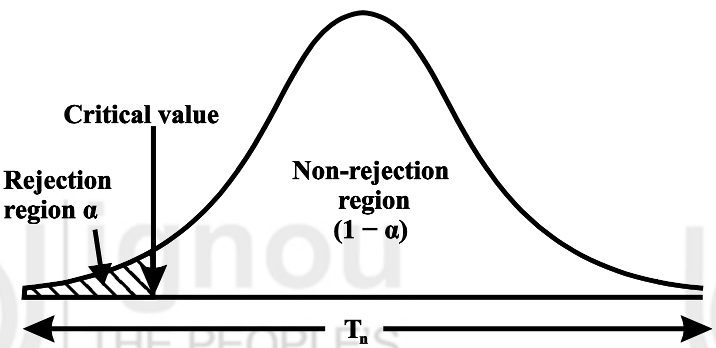

Fig. 9.3

Case III: If the alternative hypothesis is two sided such as $H_1: \theta \neq \theta_0$ or $H_1: \theta_1 \neq \theta_2$ then critical or rejection regions of size $\alpha/2$ lies on both tails of the probability curve of sampling distribution of the test statistic as shown in Fig. 9.4.

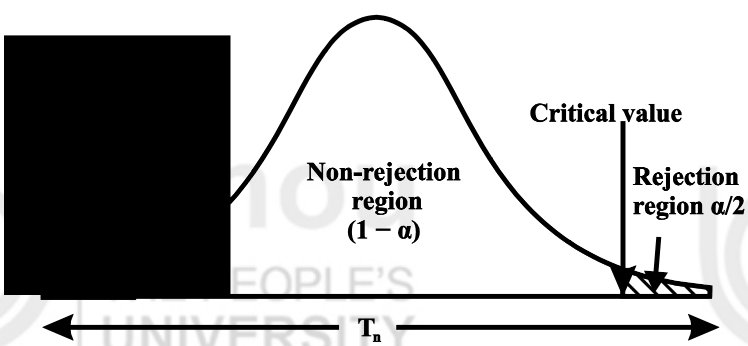

Fig. 9.4

Now, you can try the following exercise.

E4) If $H_0: \theta = 60$ and $H_1: \theta \neq 60$ then critical region lies in one-tail or two-tails.

9.4 TYPE-I AND TYPE-II ERRORS

In Section 9.3, we have discussed a rule that if the value of test statistic falls in rejection (critical) region then we reject the null hypothesis and if it falls in the non-rejection region then we do not reject the null hypothesis. A test statistic is calculated on the basis of observed sample observations. But a sample is a small part of the population about which decision is to be taken. A random sample may or may not be a good representative of the population.

A faulty sample misleads the inference (or conclusion) relating to the null hypothesis. For example, an engineer infers that a packet of screws is sub-standard when actually it is not. It is an error caused due to poor or inappropriate (faulty) sample. Similarly, a packet of screws may infer good when actually it is sub-standard. So we can commit two kinds of errors while testing a hypothesis which are summarised in Table 9.1 which is given below:

Table 9.1: Type of Errors

Decision	**H_0 True**	**H_1 True**
Reject H_0	Type-I Error	Correct Decision
Do not reject H_0	Correct Decision	Type-II Error

Let us take a situation where a patient suffering from high fever reaches to a doctor. And suppose the doctor formulates the null and alternative hypotheses as

H_0: The patient is a malaria patient

H_1: The patient is not a malaria patient

Then following cases arise:

Case I: Suppose that the hypothesis H_0 is really true, that is, patient actually a malaria patient and after observation, pathological and clinical examination, the doctor rejects H_0, that is, he / she declares him / her a non-malaria-patient. It is not a correct decision and he / she commits an error in decision known as type-I error.

Case II: Suppose that the hypothesis H_0 is actually false, that is, patient actually a non-malaria patient and after observation, the doctor rejects H_0, that is, he / she declares him / her a non-malaria-patient. It is a correct decision.

Case III: Suppose that the hypothesis H_0 is really true, that is, patient actually a malaria patient and after observation, the doctor does not reject H_0, that is, he / she declares him / her a malaria-patient. It is a correct decision.

Case IV: Suppose that the hypothesis H_0 is actually false, that is, patient actually a non-malaria patient and after observation, the doctor does not reject H_0, that is, he / she declares him / her a malaria-patient. It is not a correct decision and he / she commits an error in decision known as type-II error.

Thus, we formally define type-I and type-II errors as below:

Type-I Error:

The decision relating to rejection of null hypothesis H_0 when it is true is called type-I error. The probability of committing the type-I error is called size of test, denoted by α and is given by

$$\alpha = P\ [\text{Reject } H_0 \text{ when } H_0 \text{ is true}] = P\ [\text{Reject } H_0 / H_0 \text{ is true}]$$

We reject the null hypothesis if random sample / test statistic falls in rejection region, therefore,

$$\alpha = P[\, X \in \omega / H_0 \,]$$

where $X = (X_1, X_2, \ldots, X_n)$ is a random sample and ω is the rejection region and

$$1 - \alpha = 1 - P[\text{Reject } H_0 / H_0 \text{ is true}]$$

$$= P[\text{Do not reject } H_0 / H_0 \text{ is true}] = P[\text{Correct decision}]$$

The $(1-\alpha)$ is the probability of correct decision and it correlates to the concept of $100(1-\alpha)\%$ confidence interval used in estimation.

Type-II Error:

The decision relating to non-rejection of null hypothesis H_0 when it is false (i.e. H_1 is true) is called type-II error. The probability of committing type-II error is generally denoted by β and is given by

$$\beta = P[\text{Do not reject } H_0 \text{ when } H_0 \text{ is false}]$$

$$= P[\text{Do not reject } H_0 \text{ when } H_1 \text{ is true}]$$

$$= P[\text{Do not reject } H_0 / H_1 \text{ is true}]$$

$$= P[\, X \in \overline{\omega} / H_1 \,] \text{ where, } \overline{\omega} \text{ is the non-rejection region.}$$

and

$$1 - \beta = 1 - P[\text{Do not reject } H_0 / H_1 \text{ is true}]$$

$$= P[\text{Reject } H_0 / H_1 \text{ is true}] = P[\text{Correct decision}]$$

The $(1-\beta)$ is the probability of correct decision and also known as **"power of the test"**. Since it indicates the ability or power of the test to recognize correctly that the null hypothesis is false, therefore, we wish a test that yields a large power.

We say that a statistical test is ideal if it minimizes the probability of both types of errors and maximizes the probability of correct decision. But for fix sample size, α and β are so interrelated that the decrement in one results into the increment in other. So minimization of both probabilities of type-I and type-II errors simultaneously for fixed sample size is not possible without increasing sample size. Also both types of errors will be at zero level (i.e. no error in decision) if size of the sample is equal to the population size. But it involves huge cost if population size is large. And it is not possible in all situations such as testing of blood.

Depending on the problem in hand, we have to choose the type of error which has to minimize. For this, we have to look at a situation, suppose there is a decision making problem and there is a rule that if we make type-I error, we lose10 rupees and if we make type-II error we lose 1000 rupees. In this case, we try to eliminate the type-II error, since it is more expensive.

In another situation, suppose the Delhi police arrests a person whom they suspect is a murderer. Now, policemen have to test hypothesis:

H_0: Arrested person is innocent (not murderer)

H_1: Arrested person is a murderer

The type-I error is

$$\alpha = P\,[\text{Reject } H_0 \text{ when it is true}]$$

That is, suspected person who is actually an innocent will be sent to jail when H_0 rejects, although H_0 being a true.

The type-II error is

$$\beta = P\text{ [Do not reject } H_0 \text{ when } H_1 \text{ is true]}$$

That is, when arrested person truly a murderer but released by the police. Now, we see that in this case type-I error is more serious than type-II error because a murderer may be arrested / punished later on but sending jail to an innocent person is serious.

Consider another situation, suppose we want to test the null hypothesis $H_0 : p = 0.5$ against $H_1 : p \neq 0.5$ on the basis of tossing a coin once, where p is the probability of getting a head in a single toss (trial). And we reject the null hypothesis if a head appears and do not reject otherwise. The type-I error, that is, the probability of Reject H_0 when it is true can be calculated easily(as shown in Example 1) but the computation of type-II error is not possible because there are infinitely many alternatives for p such as p = 0.6, p = 0.1, etc.

Generally, strong control on α is necessary. It should be kept as low as possible. In test procedure, we prefix it at very low level like $\alpha = 0.05$ (5%) or 0.01 (1%) .

Now, it is time to do some examples relating to α and β.

Example 1: It is desired to test a hypothesis $H_0 : p = p_0 = 1/2$ against the alternative hypothesis $H_1 : p = p_1 = 1/4$ on the basis of tossing a coin once, where p is the probability of "getting a head" in a single toss (trial) and agreeing to reject H_0 if a head appears and accept H_0 otherwise. Find the value of α and β.

Solution: In such type of problems, first of all we search for critical region.

Here, we have critical region ω = {head}

Therefore, the probability of type-I error can be obtained as

$$\alpha = P[\text{Reject } H_0 \text{ when } H_0 \text{ is true}]$$

$$= P[X \in \omega / H_0] = P[\text{Head appears} / H_0]$$

$$= P\left[\text{Head appears}\right]_{p=\frac{1}{2}} = \frac{1}{2} \quad \left[\begin{array}{l}\because H_0 \text{ is true so we take value} \\ \text{of parameter p given in } H_0\end{array}\right]$$

Also,

$$\beta = P[\text{Do not reject } H_0 \text{ when } H_1 \text{ is true}]$$

$$= P\left[X \notin \omega / H_1\right] = P\ [\text{Tail appears} / H_1]$$

$$= P\left[\text{Tail appears}\right]_{p=\frac{1}{4}} \qquad \left[\begin{array}{l}\because H_1 \text{ is true so we take value} \\ \text{of parameter p given in } H_1\end{array}\right]$$

$$= 1 - P\left[\text{Head appears}\right]_{p=\frac{1}{4}} = 1 - \frac{1}{4} = \frac{3}{4}$$

Example 2: For testing H_0: $\theta = 1$ against H_1: $\theta = 2$, the pdf of the variable is given by

$$f(x, \theta) = \begin{cases} \frac{1}{\theta}; & 0 \leq x \leq \theta \\ 0; & \text{elsewhere} \end{cases}$$

Obtain type-I and type-II errors when critical region is $X \geq 0.4$. Also obtain power function of the test.

Solution: Here, we have critical (rejection) and non-rejection regions as

$$\omega = \{X : X \geq 0.4\} \text{ and } \overline{\omega} = \{X : X < 0.4\}$$

We have to test the null hypothesis

$$H_0 : \theta = 1 \text{ against } H_1 : \theta = 2$$

The size of type-I error is given by

$$\alpha = P[X \in \omega / H_0] = P[X \geq 0.4 / \theta = 1]$$

$$= \left[\int_{0.4}^{\theta} f(x, \theta)\, dx\right]_{\theta=1} \qquad \left[\because P[X \geq a] = \int_{a}^{\infty} f(x, \theta)\, dx\right] \qquad \ldots (1)$$

Now, by using $f(x, \theta) = \frac{1}{\theta}; \quad 0 \leq x \leq \theta$, we get from equation (1)

$$\alpha = \left[\int_{0.4}^{\theta} \frac{1}{\theta} dx\right]_{\theta=1} = \int_{0.4}^{1} dx = (x)_{0.4}^{1} = 1 - 0.4 = 0.6$$

Similarly, the size of type-II error is given by

$$\beta = P[X \in \overline{\omega} / H_1] = P[X < 0.4 / \theta = 2]$$

$$\beta = \left[\int_{0}^{0.4} \frac{1}{\theta} dx\right]_{\theta=2} = \int_{0}^{0.4} \frac{1}{2} dx = \frac{1}{2}(x)_{0}^{0.4} = \frac{1}{2}(0.4 - 0) = 0.2$$

The power function of the test = 1 - β = 1 - 0.2 = 0.8.

Now, you can try the following exercise.

E5) An urn contains either 4 white and 2 black balls or 2 white and 4 black balls. Two balls are to be drawn from the urn. If less than two white balls are obtained, it will be decided that this urn contains 2 white and 4 black balls. Calculate the values of α and β.

9.5 LEVEL OF SIGNIFICANCE

So far in this unit, we have discussed the hypothesis, types of hypothesis, critical region and types of errors. In this section, we shall discuss very useful concept **"level of significance"**, which play an important role in decision making while testing a hypothesis.

The probability of type-I error is known as level of significance of a test. It is also called the size of the test or size of critical region, denoted by α. Generally, it is pre-fixed as 5% or 1% level (α = 0.05 or 0.01). As we have discussed in Section 9.3 that if calculated value of the test statistic lies in rejection(critical) region, then we reject the null hypothesis and if it lies in non-rejection region, then we do not reject the null hypothesis. Also we note that when H_0 is rejected then automatically the alternative hypothesis H_1 is accepted. Now, one point of our discussion is that how to decide critical value(s) or cut-off value(s) for a known test statistic.

If distribution of test statistic could be expressed into some well-known distributions like Z, χ^2, t, F etc. then our problem will be solved and using the probability distribution of test statistic, we can find the cut-off value(s) that provides us critical area equal to 5% (or 1%).

Another viewpoint about the level of significance relates to the trueness of the conclusion. If H_0 do not reject at level, say, $\alpha = 0.05$ (5% level) then a person will be confident that "concluding statement about H_0" is true with 95% assurance. But even then it may false with 5% chance. There is no cent-percent assurance about the trueness of statement made for H_0.

As an example, if among 100 scientists, each draws a random sample and use the same test statistic to test the same hypothesis H_0 conducting same experiment, then 95 of them will reach to the same conclusion about H_0. But still 5 of them may differ (i.e. against the earlier conclusion).

Similar argument can be made for, say, $\alpha = 0.01$ (=1%). It is like when H_0 is rejected at $\alpha = 0.01$by a scientist , then out of 100 similar researchers who work together at the same time for the same problem, but with different random samples, 99 of them would reach to the same conclusion however, one may differ.

Now, you can try the following exercise.

E6) If probability of type-I error is 0.05 then what is the level of significance?

9.6 ONE-TAILED AND TWO-TAILED TESTS

We have seen in Section 9.3 that rejection (critical) region lies at one-tail or two-tails on the probability curve of sampling distribution of the test statistic its depend upon the form of alternative hypothesis. Similarly, the test of testing the null hypothesis also depends on the alternative hypothesis.

A test of testing the null hypothesis is said to be two-tailed test if the alternative hypothesis is two-tailed whereas if the alternative hypothesis is one-tailed then a test of testing the null hypothesis is said to be one-tailed test.

For example, if our null and alternative hypothesis are

$$H_0 : \theta = \theta_0 \text{ and } H_1 : \theta \neq \theta_0$$

then the test for testing the null hypothesis is two-tailed test because the alternative hypothesis is two-tailed that means, the parameter θ can take value greater than θ_0 or less than θ_0.

If the null and alternative hypotheses are

$$H_0 : \theta \leq \theta_0 \text{ and } H_1 : \theta > \theta_0$$

then the test for testing the null hypothesis is right-tailed test because the alternative hypothesis is right-tailed.

Similarly, if the null and alternative hypotheses are

$$H_0 : \theta \geq \theta_0 \text{ and } H_1 : \theta < \theta_0$$

then the test for testing the null hypothesis is left-tailed test because the alternative hypothesis is left-tailed.

The above discussion can be summarised in Table 9.2.

Table 9.2: Null and Alternative Hypotheses and Corresponding One-tailed and Two-tailed Tests

Null Hypothesis	Alterative Hypothesis	Types of Critical Region / Test
$H_0 : \theta = \theta_0$	$H_1 : \theta \neq \theta_0$	Two-tailed test having critical regions under both tails.
$H_0 : \theta \leq \theta_0$	$H_1 : \theta > \theta_0$	Right-tailed test having critical region under right tail only.
$H_0 : \theta \geq \theta_0$	$H_1 : \theta < \theta_0$	Left- tailed test having critical region under left tail only.

Let us do one example based on type of tests.

Example 3: A company has replaced its original technology of producing electric bulbs by CFL technology. The company manager wants to compare the average life of bulbs manufactured by original technology and new technology CFL. Write appropriate null and alternate hypotheses. Also say about one tailed and two tailed tests.

Solution: Suppose the average lives of original and CFL technology bulbs are denoted by μ_1 and μ_2 respectively.

If company manager is interested just to know whether any significant difference exists in average-life time of two types of bulbs then null and alternative hypotheses will be:

$H_0: \mu_1 = \mu_2$ [average lives of two types of bulbs are same]

$H_1: \mu_1 \neq \mu_2$ [average lives of two types of bulbs are different]

Since alternative hypothesis is two-tailed therefore corresponding test will be two-tailed.

If company manager is interested just to know whether average life of CFL is greater than original technology bulbs then our null and alternative hypotheses will be

$H_0: \mu_1 \geq \mu_2$

$$H_1: \mu_1 < \mu_2 \quad \begin{bmatrix} \text{average life of CFL technology bulbs} \\ \text{is greater than average life of original technology} \end{bmatrix}$$

Since alternative hypothesis is left-tailed therefore corresponding test will be left-tailed test.

Now, you can try the following exercises.

E7) If we have null and alternative hypotheses as

$H_0: \theta = \theta_0$ and $H_1: \theta \neq \theta_0$

then corresponding test will be

(i) left-tailed test (ii) right-tailed test (iii) both-tailed test

Write the correct option.

E8) The test whether one-tailed or two-tailed depends on

(i) Null hypothesis (H_0) (ii) Alternative hypothesis (H_1)

(iii) Neither H_0 nor H_1 (iv) Both H_0 and H_1

9.7 GENERAL PROCEDURE OF TESTING A HYPOTHESIS

Testing of hypothesis is a huge demanded statistical tool by many discipline and professionals. It is a step by step procedure as you will see in next three units through a large number of examples. The aim of this section is just give you flavor of that sequence which involves following steps:

Step I: First of all, we have to setup null hypothesis H_0 and alternative hypothesis H_1. Suppose, we want to test the hypothetical / claimed /

assumed value θ_0 of parameter θ. So we can take the null and alternative hypotheses as

$$H_0 : \theta = \theta_0 \text{ and } H_1 : \theta \neq \theta_0 \qquad [\text{for two-tailed test}]$$

$$\text{or} \quad \left.\begin{array}{l} H_0 : \theta \leq \theta_0 \text{ and } H_1 : \theta > \theta_0 \\ H_0 : \theta \geq \theta_0 \text{ and } H_1 : \theta < \theta_0 \end{array}\right\} \qquad [\text{for one-tailed test}]$$

In case of comparing same parameter of two populations of interest, say, θ_1 and θ_2, then our null and alternative hypotheses would be

$$H_0 : \theta_1 = \theta_2 \text{ and } H_1 : \theta_1 \neq \theta_2 \qquad [\text{for two-tailed test}]$$

$$\text{or} \quad \left.\begin{array}{l} H_0 : \theta_1 \leq \theta_2 \text{ and } H_1 : \theta_1 > \theta_2 \\ H_0 : \theta_1 \geq \theta_2 \text{ and } H_1 : \theta_1 < \theta_2 \end{array}\right\} \qquad [\text{for one-tailed test}]$$

Step II: After setting the null and alternative hypotheses, we establish a criteria for rejection or non-rejection of null hypothesis, that is, decide the level of significance (α), at which we want to test our hypothesis. Generally, it is taken as 5% or 1% (α = 0.05 or 0.01).

Step III: The third step is to choose an appropriate test statistic under H_0 for testing the null hypothesis as given below:

$$\text{Test statistic} = \frac{\text{Statistic} - \text{Value of the parameter under } H_0}{\text{Standard error of statistic}}$$

After that, specify the sampling distribution of the test statistic preferably in the standard form like Z (standard normal), χ^2, t, F or any other well-known in literature.

Step IV: Calculate the value of the test statistic described in Step III on the basis of observed sample observations.

Step V: Obtain the critical (or cut-off) value(s) in the sampling distribution of the test statistic and construct rejection (critical) region of size α. Generally, critical values for various levels of significance are putted in the form of a table for various standard sampling distributions of test statistic such as Z-table, χ^2-table, t-table, etc.

Step VI: After that, compare the calculated value of test statistic obtained from Step IV, with the critical value(s) obtained in Step V and locates the position of the calculated test statistic, that is, it lies in rejection region or non-rejection region.

Step VII: In testing of hypothesis ultimately we have to reach at a conclusion. It is done as explained below:

(i) If calculated value of test statistic lies in rejection region at α level of significance then we reject null hypothesis. It means that the sample data provide us sufficient evidence against the null hypothesis and there is a significant difference between hypothesized value and observed value of the parameter.

(ii) If calculated value of test statistic lies in non-rejection region at α level of significance then we do not reject null hypothesis. Its means that the sample data fails to provide us sufficient evidence against the null hypothesis and the difference between hypothesized value and observed value of the parameter due to fluctuation of sample.

Note 3: Nowadays the decision about the null hypothesis is taken with the help of p-value. The concept of p-value is very important, because computer packages and statistical software such as SPSS, SAS, STATA, MINITAB, EXCEL, etc. all provide p-value. So, Section 9.8 is devoted to explain the concept of p-value.

Now, with the help of an example we explain the above procedure.

Example 4: Suppose, it is found that average weight of a potato was 50 gm and standard deviation was 5.1 gm nearly 5 years ago. We want to test that due to advancement in agricultural technology, the average weight of a potato has been increased. To test this, a random sample of 50 potatoes is taken and calculate the sample mean $(\bar{X})$ as 52gm. Describe the procedure to carry out this test.

Solution: Here, we are given that

Specified value of population mean = μ_0 = 50 gm,

Population standard deviation = σ = 5.1 gm,

Sample size = n = 50,

Sample mean = $\bar{X}$ = 52 gm

To carry out the above test, we have to follow up the following steps:

Step I: First of all, we setup null and alternative hypotheses. Here, we want to test that the average weight of potato is increased. So our claim is "average weight of potato has increased" i.e. $\mu > 50$ and its complement is $\mu \leq 50$. Since complement contains equality sign so we can take the complement as the null hypothesis and claim as the alternative hypothesis, that is,

$$H_0: \mu \leq 50 \text{ gm and } H_1: \mu > 50 \text{ gm [Here, } \theta = \mu]$$

Since the alternative hypothesis is right-tailed, so our test is right-tailed.

Step II: After setting the null and alternative hypotheses, we fix level of significance α. Suppose, α = 0.01 (= 1 % level).

Step III: Define a test statistic to test the null hypothesis as

$$\text{Test staistic} = \frac{\text{Statistic} - \text{Value of the parameter under } H_0}{\text{Standard error of statistic}}$$

$$T = \frac{\bar{X} - 50}{\sigma / \sqrt{n}}$$

Since sample size is large (n = 50 > 30) so by central limit theorem the sampling distribution of test statistic approximately follows standard normal distribution (as explained in Unit 1 of this course), i.e. $T \sim N(0,1)$

Step IV: Calculate the value of test statistic on the basis of sample observations as

$$T = \frac{52 - 50}{5.1/\sqrt{50}} = \frac{2}{0.72} = 2.78$$

Step V: Now, we find the critical value. The critical value or cut-off value for standard normal distribution is given in **Table I (Z-table)** in the Appendix at the end of Block 1 of this course. So from this table, the critical value for right-tailed test at α = 0.01 is z_α = 2.33.

Step IV: Now, to take the decision about the null hypothesis, we compare the calculated value of test statistic with the critical value.

Since calculated value of test statistic (= 2.78) is greater than critical value (= 2.33), that means calculated value of test statistic lies in rejection region at 1% level of significance as shown in Fig. 9.5. So we reject null hypothesis and support the alternative hypothesis. Since alternative hypothesis is our claim, so we support the claim.

Thus, we conclude that sample does not provide us sufficient evidence against the claim so we may assume that the average weight of potato has increased.

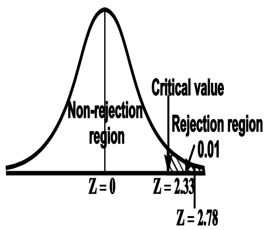

Fig. 9.5

Now, you can try the following exercise.

E9) What is the first step in testing of hypothesis?

9.8 CONCEPT OF p-VALUE

In Note 3 of Section 9.6, we promised that p-value will be discussed in Section 9.8. So, it is the time to keep our promise. Nowadays use of p-value is becoming more and more popular because of the following two reasons:

- most of the statistical software provides p-value rather than critical value.
- p-value provides more information compared to critical value as far as rejection or do not rejection of H_0 .

The first point listed above needs no explanation. But second point lies in the heart of p-value and needs to explain more clearly. Moving in this direction, we note that in scientific applications one is not only interested simply in rejecting or not rejecting the null hypothesis but he/she is also interested to assess how strong the data has the evidence to reject H_0. For example, as we have seen in Example 4 based on general procedure of testing a hypothesis where we tested the null hypothesis

$$H_0: \theta \leq 50 \text{ gm against } H_1: \theta > 50 \text{ gm}$$

To test the null hypothesis, we calculated the value of test statistic as 2.78 and the critical value (z_α) at $\alpha = 0.01$ was $z_\alpha = 2.33$.

Since calculated value of test statistic (= 2.78) is greater than critical (tabulated) value (= 2.33), therefore, we reject the null hypothesis at 1% level of significance.

Now, if we reject the null hypothesis at this level (1%) surely we have to reject it at higher level because at $\alpha = 0.05$, $z_\alpha = 1.645$ and at $\alpha = 0.10$, $z_\alpha = 1.28$. However, the calculated value of test statistic is much higher than 1.645 and 1.28, therefore, the question arises "could the null hypothesis also be rejected at values of α smaller than 0.01?" The answer is "yes" and we can compute the smallest level of significance (α) at which a null hypothesis can be rejected. This smallest level of significance (α) is known as "p-value".

The p-value is the smallest value of level of significance(α) at which a null hypothesis can be rejected using the obtained value of the test statistic and can be defined as:

The p-value is the probability of obtaining a test statistic equal to or more extreme (in the direction of sporting H_1) than the actual value obtained when null hypothesis is true.

The p-value also depends on the type of the test. If test is one-tailed then the p-value is defined as:

For right-tailed test:

p-value = P[Test Statistic (T) $\geq$ observed value of the test statistic]

For left-tailed test:

p-value = P[Test Statistic (T) $\leq$ observed value of the test statistic]

If test is two-tailed then the p-value is defined as:

For two-tailed test:

$$\text{p-value} = 2P\left[T \geq \left|\text{observed value of the test statistic}\right|\right]$$

Procedure of taking the decision about the null hypothesis on the basis of p-value:

To take the decision about the null hypothesis based on p-value, the p-value is compared with level of significance (α) and if p-value is equal or less than α then we reject the null hypothesis and if the p-value is greater than α we do not reject the null hypothesis.

The p-value for various tests can be obtained with the help of the tables given in the Appendix of the Block 1 of this course. But unless we are dealing with the standard normal distribution, the exact p-value is not obtained with the tables as mentioned above. But if we test our hypothesis with the help of computer packages or softwares such as SPSS, SAS, MINITAB, STATA, EXCEL, etc. then these types of computer packages or softwares present the p-value as part of the output for each hypothesis testing procedure. Therefore, in this block we will also describe the procedure to take the decision about the null hypothesis on the basis of critical value as well as p-value concepts.

9.9 RELATION BETWEEN CONFIDENCE INTERVAL AND TESTING OF HYPOTHESIS

In Units 7 and 8 of this course, we have learned about confidence intervals which were used to estimate the unknown population parameters with certain confidence. In Section 9.7, we have discussed the general procedure of testing a hypothesis which has been used in making decision about the specified/ assumed/ hypothetical values of population parameters. Both confidence interval and hypothesis testing have been used for different purposes but have been based on the same set of concepts. Therefore, there is an extremely close relationship between confidence interval and hypothesis testing.

In confidence interval, if we construct $(1-\alpha)$ 100% confidence interval for an unknown parameter then this interval contained all probable values for the parameter being estimated and relatively improbable values are not contained by this interval.

So this concept can also be used in hypothesis testing. For this, we contract an appropriate $(1-\alpha)$ 100% confidence interval for the parameter specified by the null hypothesis as we have discussed in Units 7 and 8 of this course and if the value of the parameter specified by the null hypothesis lies in this confident interval then we do not reject the null hypothesis and if this specified value does not lie in this confidence interval then we reject the null hypothesis.

Therefore, three cases may arise:

Case I: If we want to test the null hypothesis H_0: $\theta = \theta_0$ against the alternative hypothesis H_1: $\theta \neq \theta_0$ at 5% or 1% level of significance then we construct two-sided $(1-\alpha)100\% = 95\%$ or 99% confidence interval for the parameter θ. And we have 95% or 99% (as may be the case) confidence that this interval will include the parameter value θ_0. If the value of the parameter specified by the null hypothesis i.e. θ_0 lies in this confidence interval then we do not reject the null hypothesis otherwise we reject the null hypothesis.

Case II: If we want to test the null hypothesis H_0: $\theta \leq \theta_0$ against the alternative hypothesis H_1: $\theta > \theta_0$ then we construct the lower one-sided confidence bound for parameter θ. If the value of the parameter specified by the null hypothesis i.e. θ_0 is greater than or equal to this lower bound then we do not reject the null hypothesis otherwise we reject the null hypothesis.

Case III: If we want to test the null hypothesis H_0: $\theta \geq \theta_0$ against the alternative hypothesis H_1: $\theta < \theta_0$ then we construct the upper one-sided confidence bound for parameter θ. If the value of the parameter specified by the null hypothesis i.e. θ_0 is less than or equal to this upper bound then we do not reject the null hypothesis otherwise we reject the null hypothesis.

For example, referring back to Example 4 of this unit, here we want to test the null hypothesis

$$H_0 : \mu \leq 50\,\text{gm against } H_1 : \mu > 50\,\text{gm}$$

This was tested with the help of test statistic

$$T = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \qquad [\text{Here}, \theta = \mu]$$

and we reject the null hypothesis at $\alpha = 0.01$.

This problem could also have been solved by obtaining confidence interval estimate of population mean which is described in Section 7.4 of Unit 7.

Here, we are given that

$$n = 50,\ \bar{X} = 50 \text{ and } \sigma = 5.1$$

Since alternative hypothesis is right-tailed, therefore, we construct lower one-sided confidence bound for population mean.

Since population variance is known so lower one-sided $(1-\alpha)$ 100 % confidence bound for population mean when population variance is known is given by

$$\bar{X} - z_\alpha \frac{\sigma}{\sqrt{n}}$$

Since we test our null hypothesis at $\alpha = 0.01$ therefore, we contract 99% lower confidence bound and for $\alpha = 0.01$, we have $z_\alpha = z_{0.01} = 2.33$.

Thus, lower one-sided 99% confidence bound for average weight of potatoes is

$$52 - 2.33 \frac{5.1}{\sqrt{50}} = 52 - 1.68 = 50.32$$

Since the value of the parameter specified by the null hypothesis i.e. $\mu = 50$ is less than lower bound for average weight of potato so we reject the null hypothesis.

Thus, we can use three approaches (critical value, p-value and confidence interval) for taking decision about null hypothesis.

With this we end this unit. Let us summarise what we have discussed in this unit.

9.10 SUMMARY

In this unit, we have covered the following points:

1. Statistical hypothesis, null hypothesis, alternative hypothesis, simple & composite hypotheses.
2. Type-I and Type-II errors.
3. Critical region.
4. One-tailed and two-tailed tests.
5. General procedure of testing a hypothesis.
6. Level of significance.
7. Concept of p-value.
8. Relation between confidence interval and testing of hypothesis.

9.11 SOLUTIONS / ANSWERS

E1) Here, we wish to test the claim of the company that the average life of its car tyres is 50000 miles so

Claim: $\mu = 50000$ miles and complement: $\mu \neq 50000$ miles

Since claim contains equality sign so we take claim as the null hypothesis and complement as the alternative hypothesis i.e.

$H_0: \mu = 50000$ miles

$H_1: \mu \neq 50000$ miles

E2) (iii) Here, doctor wants to test whether new medicine is really more effective for controlling blood pressure than old medicine so

Claim: $\mu_1 > \mu_2$ and complement: $\mu_1 \leq \mu_2$

Since complement contains equality sign so we take complement as the null hypothesis and claim as the alternative hypothesis i.e.

$H_0: \mu_1 \leq \mu_2$

$H_1: \mu_1 > \mu_2$

(iv) Here, economist wants to test whether the variability in incomes differ in two populations so

Claim: $\sigma_1^2 \neq \sigma_2^2$ and complement: $\sigma_1^2 = \sigma_2^2$

Since complement contains equality sign so we take complement as the null hypothesis and claim as the alternative hypothesis i.e.

$H_0: \sigma_1^2 = \sigma_2^2$

$H_1: \sigma_1^2 \neq \sigma_2^2$

(v) Here, psychologist wants to test whether the proportion of literates between two groups of people is same so

Claim: $P_1 = P_2$ and complement: $P_1 \neq P_2$

Since claim contains equality sign so we take claim as the null hypothesis and complement as the alternative hypothesis i.e.

H_0: $P_1 = P_2$

H_1: $P_1 \neq P_2$

E3) Here, (i) and (vi) represent the simple hypotheses because these hypotheses tell us the exact values of parameter average weight of orange μ as $\mu = 100$ and $\mu = 130$.

The rest (ii), (iii), (iv), (v) and (vii) represent the composite hypotheses because these hypotheses do not tell us the exact values of parameter μ.

E4) Since alternative hypothesis H_1: $\theta \neq 60$ is two tailed so critical region lies in two-tails.

E5) Let A and B denote the number of white balls and black balls in the urn respectively. Further, let X be the number of white balls drawn among the two balls from the urn then we can take the null and alternative hypotheses as

H_0: A = 4 & B = 2 and H_1: A = 2 & B = 4

The critical region is given by

$$w = \{X : X < 2\} = \{X : X = 0, 1\}$$

Thus,

$$\alpha = P\,[\text{Reject } H_0 \text{ when } H_0 \text{ is true}]$$

$$= P[X \in w / H_0] = P[X = 0 / H_0] + P[X = 1 / H_0]$$

$$= \frac{{}^4C_0\ {}^2C_0}{{}^6C_2} + \frac{{}^4C_1\ {}^2C_1}{{}^6C_2} = \frac{1 \times 1}{15} + \frac{4 \times 2}{15} = \frac{1}{15} + \frac{8}{15}$$

$$\alpha = \frac{9}{15} = \frac{3}{5}$$

Similarly,

$$\beta = P\,[\text{Do not reject } H_0 \text{ when } H_1 \text{ is true}]$$

$$= P[X \notin w / H_1] = P[X = 2 / H_1] = \frac{{}^2C_2\ {}^4C_0}{{}^6C_2} = \frac{1 \times 1}{15} = \frac{1}{15}$$

E6) Since level of significance is the probability of type-I error so in this case level of significance is 0.05 or 5%.

E7) Here, the alternative hypothesis is two-tailed therefore, the test will be two-tailed test.

E8) Whether the test of testing a hypothesis is one-tailed or two-tailed depends on the alternative hypothesis. So correct option is (ii).

E9) First step in testing of hypothesis is to setup null and alternative hypotheses.

UNIT 10 LARGE SAMPLE TESTS

Structure

10.1 INTRODUCTION

In previous unit, we have defined basic terms used in testing of hypothesis. After providing you necessary material required for any test, we can move towards discussing particular tests one by one. But before doing that let us tell you the strategy we are adopting here.

First we categories the tests under two heads:

- Large sample tests
- Small sample tests

After that, their unit wise distribution is done. In this unit, we will discuss large sample tests whereas in Units 11 and 12 we will discuss small sample tests. The tests which are described in these units are known as **"parametric tests"**.

Sometimes in our studies in the fields of economics, psychology, medical, etc. we take a sample of objects / units / participants / patients, etc. such as 70, 500, 1000, 10,000, etc. This situation comes under the category of large samples.

As a thumb rule, a sample of size n is treated as a large sample only if it contains more than 30 units (or observations, $n > 30$). And we know that, for large sample ($n > 30$), one statistical fact is that almost all sampling distributions of the statistic(s) are closely approximated by the normal distribution. Therefore, the test statistic, which is a function of sample observations based on $n > 30$, could be assumed follow the normal distribution approximately (or exactly).

But story does not end here. There are some other issues which need to be taken care off. Some of these issues have been highlighted by making different cases in each test as you will see when go through Sections 10.3 to 10.8 of this unit.

This unit is divided into ten sections. Section 10.1 is introductory in nature. General procedure of testing of hypothesis for large samples is described in

Section 10.2. In Section 10.3, testing of hypothesis for population mean is discussed whereas in Section 10.4, testing of hypothesis for difference of two population means with examples is described. Similarly, in Sections 10.5 and 10.6, testing of hypothesis for population proportion and difference of two population proportions are explained respectively. Testing of hypothesis for population variance and two population variances are described in Sections 10.7 and 10.8 respectively. Unit ends by providing summary of what we have discussed in this unit in Section 10.9 and solution of exercises in Section 10.10.

Objectives

After studying this unit, you should be able to:

- judge for a given situation whether we should go for large sample test or not;
- Applying the Z-test for testing the hypothesis about the population mean and difference of two population means;
- Applying the Z-test for testing the hypothesis about the population proportion and difference of two population proportions; and
- Applying the Z-test for testing the hypothesis about the population variance and two population variances.

10.2 PROCEDURE OF TESTING OF HYPOTHESIS FOR LARGE SAMPLES

As we have described in previous section that for large sample size (n > 30), one statistical fact is that almost all sampling distributions of the statistic(s) are closely approximated by the normal distribution. Therefore, when sample size is large one can apply the normal distribution based test procedures to test the hypothesis.

In previous unit, we have given the procedure of testing of hypothesis in general. Let us now discuss the procedure of testing a hypothesis for large samples in particular.

Suppose $X_1, X_2, \ldots, X_n$ is a random sample of size n (> 30) selected from a population having unknown parameter θ and we want to test the hypothesis about the hypothetical / claimed / assumed value θ_0 of parameter θ. For this, a test procedure is required. We discuss it step by step as follows:

Step I: First of all, we have to setup null hypothesis H_0 and alternative hypothesis H_1. Here, we want to test the hypothetical / claimed / assumed value θ_0 of parameter θ. So we can take the null and alternative hypotheses as

$$H_0 : \theta = \theta_0 \text{ and } H_1 : \theta \neq \theta_0 \qquad [\text{for two-tailed test}]$$

$$\text{or} \quad \left.\begin{array}{l} H_0 : \theta \leq \theta_0 \text{ and } H_1 : \theta > \theta_0 \\ H_0 : \theta \geq \theta_0 \text{ and } H_1 : \theta < \theta_0 \end{array}\right\} \qquad [\text{for one-tailed test}]$$

In case of comparing same parameter of two populations of interest, say, θ_1 and θ_2, then our null and alternative hypotheses would be

$$H_0 : \theta_1 = \theta_2 \text{ and } H_1 : \theta_1 \neq \theta_2 \qquad [\text{for two-tailed test}]$$

$$\text{or} \quad \left.\begin{array}{l} H_0: \theta_1 \le \theta_2 \text{ and } H_1: \theta_1 > \theta_2 \\ H_0: \theta_1 \ge \theta_2 \text{ and } H_1: \theta_1 < \theta_2 \end{array}\right\} \quad [\text{for one-tailed test}]$$

Step II: After setting the null and alternative hypotheses, we have to choose level of significance. Generally, it is taken as 5% or 1% ($\alpha = 0.05$ or 0.01). And accordingly rejection and non-rejection regions will be decided.

Step III: Third step is to determine an appropriate test statistic, say, Z in case of large samples. Suppose T_n is the sample statistic such as sample mean, sample proportion, sample variance, etc. for the parameter θ then for testing the null hypothesis, test statistic is given by

$$Z = \frac{T_n - E(T_n)}{SE(T_n)} = \frac{T_n - E(T_n)}{\sqrt{Var(T_n)}} \quad \left[\begin{array}{l} \because \text{ we know that SE of a statistic is} \\ \text{the SD of the sampling distribution} \\ \text{of that statistic} \\ \therefore SE(T_n) = SD(T_n) = \sqrt{Var(T_n)} \end{array}\right]$$

where, $E(T_n)$ is the expectation (or mean) of T_n and $Var(T_n)$ is variance of T_n.

Step IV: As already mentioned for large samples, statistical fact is that almost all sampling distributions of the statistic(s) are closely approximated by the normal distribution as the parent population is normal or non-normal. So, the test statistic Z will assumed to be approximately normally distributed with mean 0 and variance 1 as

$$Z = \frac{T_n - E(T_n)}{\sqrt{Var(T_n)}} \sim N(0,1)$$

By putting the values of T_n, $E(T_n)$ and $Var(T_n)$ in above formula we calculate the value of test statistic Z. Let z be the calculated value of test statistic Z.

Step V: After that, we obtain the critical (cut-off or tabulated) value(s) in the sampling distribution of the test statistic Z corresponding to α assumed in Step II. These critical values are given in **Table-I (Z-table)** at the Appendix of Block 1 of this course corresponding to different level of significance (α). For convenient some useful critical values at $\alpha = 0.01, 0.05$ for Z-test are given in **Table 10.1** in this section. After that, we construct rejection (critical) region of size α in the probability curve of the sampling distribution of test statistic Z.

Step VI: Take the decision about the null hypothesis based on the calculated and critical values of test statistic obtained in Step IV and Step V. Since critical value depends upon the nature of the test that it is one-tailed test or two-tailed test so following cases arise:

In case of one-tailed test:

Case I: When $H_0: \theta \le \theta_0$ and $H_1: \theta > \theta_0$ (right-tailed test)

In this case, the rejection (critical) region falls under the right tail of the probability curve of the sampling distribution of test statistic Z. Suppose z_α is the critical value at α level of significance so entire region greater than or equal to z_α is the rejection region and less than z_α is the non-rejection region as shown in Fig. 10.1.

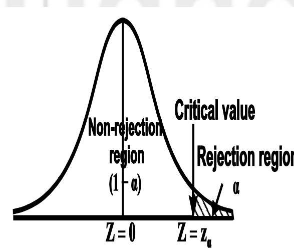

Fig. 10.1

If z (calculated value) $\geq z_\alpha$ (tabulated value), that means the calculated value of test statistic Z lies in the rejection region, then we reject the null hypothesis H_0 at α level of significance. Therefore, we conclude that sample data provides us sufficient evidence against the null hypothesis and there is a significant difference between hypothesized or specified value and observed value of the parameter.

If $z < z_\alpha$, that means the calculated value of test statistic Z lies in non-rejection region, then we do not reject the null hypothesis H_0 at α level of significance. Therefore, we conclude that the sample data fails to provide us sufficient evidence against the null hypothesis and the difference between hypothesized value and observed value of the parameter due to fluctuation of sample.

so the population parameter θ may be θ_0.

Case II: When $H_0 : \theta \geq \theta_0$ and $H_1 : \theta < \theta_0$ (left-tailed test)

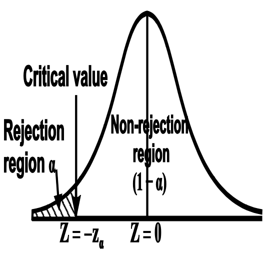

Fig. 10.2

In this case, the rejection (critical) region falls under the left tail of the probability curve of the sampling distribution of test statistic Z. Suppose $-z_\alpha$ is the critical value at α level of significance then entire region less than or equal to $-z_\alpha$ is the rejection region and greater than $-z_\alpha$ is the non-rejection region as shown in Fig. 10.2.

If $z \leq -z_\alpha$, that means the calculated value of test statistic Z lies in the rejection region, then we reject the null hypothesis H_0 at α level of significance.

If $z > -z_\alpha$, that means the calculated value of test statistic Z lies in the non-rejection region, then we do not reject the null hypothesis H_0 at α level of significance.

In case of two-tailed test: When $H_0 : \theta = \theta_0$ and $H_1 : \theta \neq \theta_0$

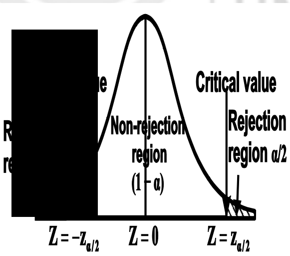

Fig. 10.3

In this case, the rejection region falls under both tails of the probability curve of sampling distribution of the test statistic Z. Half the area (α) i.e. $\alpha/2$ will lies under left tail and other half under the right tail. Suppose $-z_{\alpha/2}$ and $z_{\alpha/2}$ are the two critical values at the left-tailed and right-tailed respectively. Therefore, entire region less than or equal to $-z_{\alpha/2}$ and greater than or equal to $z_{\alpha/2}$ are the rejection regions and between $-z_{\alpha/2}$ and $z_{\alpha/2}$ is the non-rejection region as shown in Fig. 10.3.

If $z \geq z_{\alpha/2}$ or $z \leq -z_{\alpha/2}$, that means the calculated value of test statistic Z lies in the rejection region, then we reject the null hypothesis H_0 at α level of significance.

If $-z_{\alpha/2} < z < z_{\alpha/2}$, that means the calculated value of test statistic Z lies in the non-rejection region, then we do not reject the null hypothesis H_0 at α level of significance.

Table 10.1 shows some commonly used critical (cut-off or tabulated) values for one-tailed test and two-tailed test at different level of significance (α) for Z-test.

Table 10.1: Critical Values for Z-test

Level of Significance (α)	Two-Tailed Test	One-Tailed Test	
		Right-Tailed Test	Left- Tailed Test
$\alpha = 0.05$ (= 5%)	$\pm z_{\alpha/2} = \pm 1.96$	$z_\alpha = 1.645$	$-z_\alpha = -1.645$
$\alpha = 0.01$ (= 1%)	$\pm z_{\alpha/2} = \pm 2.58$	$z_\alpha = 2.33$	$-z_\alpha = -2.33$

Note 1: As we have discussed in Step IV of this procedure that when sample size is large then test statistic follows the normal distribution as the parent population is normal or non-normal so we do not require any assumption of the form of the parent population for large sample size but when sample size is small (n < 30) then for applying parametric test we must require the assumption that the population is normal as we shall in Units 11 and 12. If this assumption is not fulfilled then we apply the non-parametric tests which will be discussed in Block 4 of this course.

Decision making procedure about the null hypothesis using the concept of p-value:

To take the decision about the null hypothesis on the basis of p-value, the p-value is compared with given level of significance (α) and if p-value is less than or equal to α then we reject the null hypothesis and if the p-value is greater than α we do not reject the null hypothesis.

Since test statistic Z follows approximately normal distribution with mean 0 and variance unity, i.e. standard normal distribution and we also know that standard normal distribution is symmetrical about Z = 0 line therefore, if z represents the calculated value of Z then p-value can be calculated as follows:

For one-tailed test:

For $H_1: \theta > \theta_0$ (right-tailed test)

$$\text{p-value} = P[Z \geq z]$$

For $H_1: \theta < \theta_0$ (left-tailed test)

$$\text{p-value} = P[Z \leq z]$$

For two-tailed test:

For $H_1: \theta \neq \theta_0$

$$\text{p-value} = 2P\left[Z \geq |z|\right]$$

These p-values for Z-test can be obtained with the help of **Table-I (Z-table)** given in the Appendix at the end of Block 1 of this course (which gives the probability $[0 \leq Z \leq z]$ for different value of z) as discussed in Unit 14 of MST-003.

For example, if test is right-tailed and calculated value of test statistic Z is 1.23 then

$$\text{p-value} = P[Z \geq z] = P[Z \geq 1.23] = 0.5 - P[0 < Z < 1.23]$$

$$= 0.5 - 0.3907 \begin{bmatrix}\text{From Z-table given in Appendix} \\ \text{of Block 1of this course.}\end{bmatrix}$$

$$= 0.1093$$

Now, you can try the following exercises.

E1) If an investigator observed that the calculated value of test statistic lies in non-rejection region then he/she will

(i) reject the null hypothesis

(ii) accept the null hypothesis

(iii) not reject the null hypothesis

Write the correct option.

E2) If we have null and alternative hypotheses as

$H_0: \theta = \theta_0$ and $H_1: \theta \neq \theta_0$

then the rejection (critical) region lies in

(i) left tail

(ii) right tail

(iii) both tails

Write the correct option.

E3) If test is two-tailed and calculated value of test statistic Z is 2.42 then calculate the p-value for the Z-test.

10.3 TESTING OF HYPOTHESIS FOR POPULATION MEAN USING Z-TEST

In previous section, we have discussed the general procedure for Z-test. Now we are discussing the Z-test for testing the hypothesis or claim about the population mean when sample size is large. Let population under study has mean μ and variance σ^2, where μ is unknown and σ^2 may be known or unknown. We will consider both cases under this heading. For testing a hypothesis about population mean we draw a random sample $X_1, X_2,\ldots,X_n$ of size $n > 30$ from this population. As we know that for drawing the inference about the population mean we generally use sample mean and for test statistic, we require the mean and standard error of sampling distribution of the statistic (mean). Here, we are considering large sample so we know by central limit theorem that sample mean is asymptotically normally distributed with mean μ and variance σ^2/n whether parent population is **normal or non-normal**. That is, if $\bar{X}$ is the sample mean of the random sample then

$$E\left(\bar{X}\right) = \mu,\ \ \mathrm{Var}\left(\bar{X}\right) = \frac{\sigma^2}{n} \qquad \ldots (1)$$

But we know that standard error = $\sqrt{\text{Variance}}$

$$\therefore \qquad SE\left(\bar{X}\right) = \sqrt{\mathrm{Var}\left(\bar{X}\right)} = \frac{\sigma}{\sqrt{n}} \qquad \ldots (2)$$

Now, follow the same procedure as we have discussed in previous section, that is, first of all we have to setup null and alternative hypotheses. Since here we want to test the hypothesis about the population mean so we can take the null and alternative hypotheses as

$$H_0 : \mu = \mu_0 \text{ and } H_1 : \mu \neq \mu_0 \; [\text{for two-tailed test}] \begin{bmatrix} \text{Here, } \theta = \mu \text{ and } \theta_0 = \mu_0 \\ \text{if we compare it with} \\ \text{general procedure.} \end{bmatrix}$$

$$\text{or} \quad \left. \begin{array}{l} H_0 : \mu \leq \mu_0 \text{ and } H_1 : \mu > \mu_0 \\ H_0 : \mu \geq \mu_0 \text{ and } H_1 : \mu < \mu_0 \end{array} \right\} [\text{for one-tailed test}]$$

For testing the null hypothesis, the test statistic Z is given by

$$Z = \frac{\bar{X} - E(\bar{X})}{SE(\bar{X})}$$

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \qquad \begin{bmatrix} \text{Using equations (1) and (2) and} \\ \text{under } H_0 \text{ we assume that } \mu = \mu_0. \end{bmatrix}$$

When we assume that the null hypothesis is true then we are actually assuming that the population parameter is equal to the value in the null hypothesis. For example, we assume that $\mu = 60$ whether the null hypothesis is $\mu = 60$ or $\mu \leq 60$ or $\mu \geq 60$.

The sampling distribution of the test statistic depends upon σ^2 that it is known or unknown. Therefore, two cases arise:

Case I: When σ^2 is known

In this case, the test statistic follows the normal distribution with mean 0 and variance unity when the sample size is the large as the population under study is normal or non-normal. If the sample size is small then test statistic Z follows the normal distribution only when population under study is normal. Thus,

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \sim N(0, 1)$$

Case II: When σ^2 is unknown

In this case, we estimate σ^2 by the value of sample variance (S^2) where,

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2$$

Then become test statistic $\frac{\bar{X} - \mu_0}{S/\sqrt{n}}$ follows the t-distribution with (n−1) df as the sample size is large or small provided the population under study follows normal as we have discussed in Unit 2 of this course. But when population under study is not normal and sample size is large then this test statistic approximately follows normal distribution with mean 0 and variance unity, that is,

$$Z = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \sim N(0,1)$$

After that, we calculate the value of test statistic as may be the case (σ^2 is known or unknown) and compare it with the critical value given in **Table 10.1** at prefixed level of significance α. Take the decision about the null hypothesis as described in the previous section.

From above discussion of testing of hypothesis about population mean, we note following point:

(i) When σ^2 is known then we apply the Z-test as the population under study is normal or non-normal for the large sample. But when sample size is small then we apply the Z-test only when population under study is normal.

(ii) When σ^2 is unknown then we apply the t-test only when the population under study is normal as sample size is large or small. But when the assumption of normality is not fulfilled and sample size is large then we can apply the Z-test.

(iii) When sample is small and σ^2 is known or unknown and the form of the population is not known then we apply the non-parametric test as we will be discussed in Block 4 of this course.

Following examples will help you to understand the procedure more clearly.

Example 1: A light bulb company claims that the 100-watt light bulb it sells has an average life of 1200 hours with a standard deviation of 100 hours. For testing the claim 50 new bulbs were selected randomly and allowed to burn out. The average lifetime of these bulbs was found to be 1180 hours. Is the company's claim is true at 5% level of significance?

Solution: Here, we are given that

Specified value of population mean = μ_0 = 1200 hours,

Population standard deviation = σ = 100 hours,

Sample size = n = 50

Sample mean = $\bar{X}$ = 1180 hours.

In this example, the population parameter being tested is population mean i.e. average life of a bulb (μ) and we want to test the company's claim that average life of a bulb is 1200 hours. So our claim is μ = 1200 and its complement is $\mu \neq 1200$. Since claim contains the equality sign so we can take the claim as the null hypothesis and complement as the alternative hypothesis. So

$$H_0 : \mu = \mu_0 = 1200 \quad [\text{average life of a bulb is 1200 hours}]$$

$$H_1 : \mu \neq 1200 \quad [\text{average life of a bulb is not 1200 hours}]$$

Also the alternative hypothesis is two-tailed so the test is two-tailed test.

Here, we want to test the hypothesis regarding mean when population SD (variance) is known and sample size n = 50(> 30) is large. So we will go for Z-test.

Thus, for testing the null hypothesis the test statistic is given by

$$Z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}}$$

$$= \frac{1180 - 1200}{100 / \sqrt{50}} = \frac{-20}{14.14} = -1.41$$

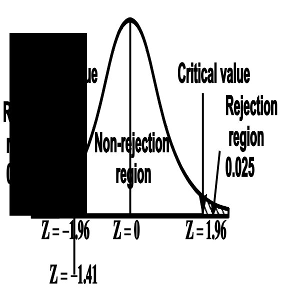

Fig. 10.4

The critical (tabulated) values for two-tailed test at 5% level of significance are $\pm z_{\alpha/2} = \pm z_{0.025} = \pm 1.96$.

Since calculated value of test statistic Z (= –1.41) is greater than critical value (= − 1.96) and less than the critical value (= 1.96), that means it lies in non-rejection region as shown in Fig. 10.4, so we do not reject the null hypothesis. Since the null hypothesis is the claim so we support the claim at 5% level of significance.

Decision according to p-value:

The test is two-tailed, therefore,

p-value = $2P\left[Z \ge |z|\right] = 2P\left[Z \ge 1.41\right]$

$$= 2\left[0.5 - P\left[0 \le Z \le 1.41\right]\right] = 2\left(0.5 - 0.4207\right) = 0.1586$$

Since p-value (= 0.1586) is greater than α (= 0.05) so we do not reject the null hypothesis at 5% level of significance.

Decision according to confidence interval:

Here, test is two-tailed, therefore, we contract two-sided confidence interval for population mean.

Since population standard deviation is known, therefore, we can use $(1-\alpha)$ 100 % confidence interval for population mean when population variance is known which is given by

$$\left[\bar{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}},\ \bar{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right]$$

Here, α = 0.05, we have $z_{\alpha/2} = z_{0.025} = 1.96.$

Thus, 95% confidence interval for average life of a bulb is given by

$$\left[1180 - 1.96\frac{100}{\sqrt{50}}, 1180 + 1.96\frac{100}{\sqrt{50}}\right]$$

or $\quad \left[1180 - 27.71, 1180 + 27.71\right]$

or $\quad \left[1152.29,\ 1207.71\right]$

Since 95% confidence interval for average life of a bulb contains the value of the parameter specified by the null hypothesis, that is, $\mu = \mu_0 = 1200$ so we do not reject the null hypothesis.

Thus, we conclude that sample does not provide us sufficient evidence against the claim so we may assume that the company's claim that the average life of a bulb is 1200 hours is true.

Note 2: Here, we note that the decisions about null hypothesis based on three approaches (critical value or classical, p-value and confidence interval) are same. The learners are advised to make the decision about the claim or statement by using only one of the three approaches in the examination. Here, we used all these approaches only to give you an idea how they can be used in a given problem. Those learners who will opt biostatistics specialisation will see and realize the importance of confidence interval approach in Unit 16 of MSTE-004.

Example 2: A manufacturer of ball point pens claims that a certain pen manufactured by him has a mean writing-life at least 460 A-4 size pages. A purchasing agent selects a sample of 100 pens and put them on the test. The mean writing-life of the sample found 453 A-4 size pages with standard deviation 25 A-4 size pages. Should the purchasing agent reject the manufacturer's claim at 1% level of significance?

Solution: Here, we are given that

Specified value of population mean = μ_0 = 460, Sample size = n = 100,

Sample mean = $\overline{X}$ = 453,

Sample standard deviation = S = 25

Here, we want to test the manufacturer's claim that the mean writing-life (μ) of pen is at least 460 A-4 size pages. So our claim is $\mu \geq 460$ and its complement is $\mu < 460$. Since claim contains the equality sign so we can take the claim as the null hypothesis and the complement as the alternative hypothesis. So

$$H_0 : \mu \geq \mu_0 = 460 \text{ and } H_1 : \mu < 460$$

Also the alternative hypothesis is left-tailed so the test is left-tailed test.

Here, we want to test the hypothesis regarding population mean when population SD is unknown. So we should used t-test for if writing-life of pen follows normal distribution. But it is not the case. Since sample size is n = 100 (n > 30) large so we go for Z-test. The test statistic of Z-test is given by

$$Z = \frac{\overline{X} - \mu_0}{S/\sqrt{n}}$$

$$= \frac{453 - 460}{25/\sqrt{100}} = \frac{-7}{2.5} = -2.8$$

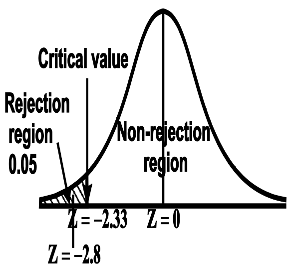

Fig. 10.5

We have from Table 10.1 the critical (tabulated) value for left-tailed Z test at 1% level of significance is $z_\alpha = -2.33$.

Since calculated value of test statistic Z (= –2.8) is less than the critical value (= −2.33), that means calculated value of test statistic Z lies in rejection region as shown in Fig. 10.5, so we reject the null hypothesis. Since the null hypothesis is the claim so we reject the manufacturer's claim at 1% level of significance.

Decision according to p-value:

The test is left-tailed, therefore,

$$\text{p-value} = P[Z \leq z] = P[Z \leq -2.8] = P[Z \geq 2.8] \quad \begin{bmatrix} \because Z \text{ is symetrical} \\ \text{about } Z = 0 \text{ line} \end{bmatrix}$$

$$= 0.5 - P[0 \leq Z \leq 2.8] = 0.5 - 0.4974 = 0.0026$$

Since p-value (= 0.0026) is less than α (= 0.01) so we reject the null hypothesis at 1% level of significance.

Therefore, we conclude that the sample provide us sufficient evidence against the claim so the purchasing agent rejects the manufacturer's claim at 1% level of significance.

Now, you can try the following exercises.

E4) A sample of 900 bolts has a mean length 3.4 cm. Is the sample regarded to be taken from a large population of bolts with mean length 3.25 cm and standard deviation 2.61 cm at 5% level of significance?

E5) A big company uses thousands of CFL lights every year. The brand that the company has been using in the past has average life of 1200 hours. A new brand is offered to the company at a price lower than they are paying for the old brand. Consequently, a sample of 100 CFL light of new brand is tested which yields an average life of 1220 hours with standard deviation 90 hours. Should the company accept the new brand at 5% level of significance?

10.4 TESTING OF HYPOTHESIS FOR DIFFERENCE OF TWO POPULATION MEANS USING Z-TEST

In previous section, we have learnt about the testing of hypothesis about the population mean. But there are so many situations where we want to test the hypothesis about difference of two population means or two population means. For example, two manufacturing companies of bulbs are produced same type of bulbs and one may be interested to test that one is better than the other, an investigator may want to test the equality of the average incomes of the peoples living in two cities, etc. Therefore, we require an appropriate test for testing the hypothesis about the difference of two population means.

Let there be two populations, say, population-I and population-II under study. Also let μ_1, μ_2 and σ_1^2, σ_2^2 denote the means and variances of population-I and population-II respectively where both μ_1 and μ_2 are unknown but σ_1^2 and σ_2^2 may be known or unknown. We will consider all possible cases here. For testing the hypothesis about the difference of two population means, we draw a random sample of large size n_1 from population-I and a random sample of large size n_2 from population-II. Let $\overline{X}$ and $\overline{Y}$ be the means of the samples selected from population-I and II respectively.

These two populations may or may not be normal but according to the central limit theorem, the sampling distribution of difference of two large sample means asymptotically normally distributed with mean $(\mu_1 - \mu_2)$ and variance $(\sigma_1^2/n_1 + \sigma_2^2/n_2)$ as described in Unit 2 of this course.

Thus,

$$E\left(\overline{X}-\overline{Y}\right) = E\left(\overline{X}\right) - E\left(\overline{Y}\right) = \mu_1 - \mu_2 \qquad \dots (3)$$

and

$$\text{Var}\left(\overline{X}-\overline{Y}\right) = \text{Var}\left(\overline{X}\right) + \text{Var}\left(\overline{Y}\right) = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$$

But we know that standard error = $\sqrt{\text{Variance}}$

$$\therefore \quad SE\left(\overline{X}-\overline{Y}\right) = \sqrt{\text{Var}\left(\overline{X}-\overline{Y}\right)} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \qquad \dots (4)$$

Now, follow the same procedure as we have discussed in Section 10.2, that is, first of all we have to setup null and alternative hypotheses. Here, we want to test the hypothesis about the difference of two population means so we can take the null hypothesis as

$$H_0 : \mu_1 = \mu_2 \text{ (no difference in means)} \quad \left[\begin{array}{l}\text{Here, } \theta_1 = \mu_1 \text{ and } \theta_2 = \mu_2 \\ \text{if we compare it with} \\ \text{general procedure.}\end{array}\right]$$

or $\quad H_0: \mu_1 - \mu_2 = 0$ (difference in two means is 0)

and the alternative hypothesis as

$$H_1: \mu_1 \neq \mu_2 \qquad \text{[for two-tailed test]}$$

$$\left.\begin{array}{l} H_0: \mu_1 \leq \mu_2 \text{ and } H_1: \mu_1 > \mu_2 \\ \text{or} \\ H_0: \mu_1 \geq \mu_2 \text{ and } H_1: \mu_1 < \mu_2 \end{array}\right\} \qquad \text{[for one-tailed test]}$$

For testing the null hypothesis, the test statistic Z is given by

$$Z = \frac{\left(\bar{X} - \bar{Y}\right) - E\left(\bar{X} - \bar{Y}\right)}{SE\left(\bar{X} - \bar{Y}\right)}$$

or
$$Z = \frac{\left(\bar{X} - \bar{Y}\right) - \left(\mu_1 - \mu_2\right)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \quad \text{[using equations (3) and (4)]}$$

Since under null hypothesis we assume that $\mu_1 = \mu_2$, therefore, we have

$$Z = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

Now, the sampling distribution of the test statistic depends upon σ_1^2 and σ_2^2 that both are known or unknown. Therefore, four cases arise:

Case I: When $\sigma_1^2 \,\&\, \sigma_2^2$ are known and $\sigma_1^2 = \sigma_2^2 = \sigma^2$

In this case, the test statistic follows normal distribution with mean 0 and variance unity when the sample sizes are large as both the populations under study are normal or non-normal. But when sample sizes are small then test statistic Z follows normal distribution only when populations under study are normal, that is,

$$Z = \frac{\bar{X} - \bar{Y}}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim N\left(0, 1\right)$$

Case II: When $\sigma_1^2 \,\&\, \sigma_2^2$ are known and $\sigma_1^2 \neq \sigma_2^2$

In this case, the test statistic also follows the normal distribution as described in case I, that is,

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

Case III: When $\sigma_1^2 \,\&\, \sigma_2^2$ are unknown and $\sigma_1^2 = \sigma_2^2 = \sigma^2$

In this case, σ^2 is estimated by value of pooled sample variance S_p^2 where,

$$S_p^2 = \frac{1}{n_1 + n_2 - 2}\left[n_1 S_1^2 + n_2 S_2^2\right]$$

and

$$S_1^2 = \frac{1}{(n_1 - 1)}\sum_{i=1}^{n_1}\left(X_i - \bar{X}\right)^2 \text{ and } S_2^2 = \frac{1}{(n_2 - 1)}\sum_{i=1}^{n_2}\left(Y_i - \bar{Y}\right)^2$$

and test statistic follows t-distribution with ($n_1 + n_2 - 2$) degrees of freedom as the sample sizes are large or small provided populations under study follow normal distribution as described in Unit 2 of this course. But when the populations are under study are not normal and sample sizes n_1 and $n_2$are large (> 30) then by central limit theorem, test statistic approximately normally distributed with mean 0 and variance unity, that is,

$$Z = \frac{\bar{X} - \bar{Y}}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim N(0,1)$$

Case IV: When $\sigma_1^2 \& \sigma_2^2$ are unknown and $\sigma_1^2 \neq \sigma_2^2$

In this case, $\sigma_1^2 \& \sigma_2^2$ are estimated by the values of the sample variances $S_1^2 \& S_2^2$ respectively and the exact distribution of test statistic is difficult to derive. But when sample sizes n_1 and n_2 are large (> 30) then central limit theorem, the test statistic approximately normally distributed with mean 0 and variance unity, that is,

$$Z = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \sim N(0, 1)$$

After that, we calculate the value of test statistic and compare it with the critical value given in **Table 10.1** at prefixed level of significance α. Take the decision about the null hypothesis as described in Section10.2.

From above discussion of testing of hypothesis about population mean, we note following point:

(i) When $\sigma_1^2 \& \sigma_2^2$ are known then we apply the Z-test as both the population under study are normal or non-normal for the large sample. But when sample sizes are small then we apply the Z-test only when populations under study are normal.

(ii) When $\sigma_1^2 \& \sigma_2^2$ are unknown then we apply the t-test only when the populations under study are normal as sample sizes are large or small. But when the assumption of normality is not fulfilled and sample sizes are large then we can apply the Z-test.

(iii) When samples are small and $\sigma_1^2 \& \sigma_2^2$ are known or unknown and the form of the population is not known then we apply the non-parametric test as we will be discussed in Block 4 of this course.

Let us do some examples based on above test.

Example 3: In two samples of women from Punjab and Tamilnadu, the mean height of 1000 and 2000 women are 67.6 and 68.0 inches respectively. If population standard deviation of Punjab and Tamilnadu are same and equal to 5.5 inches then, can the mean heights of Punjab and Tamilnadu women be regarded as same at 1% level of significance?

Solution: We are given

$$n_1 = 1000,\ n_2 = 2000,\ \bar{X} = 67.6,\ \bar{Y} = 68.0 \text{ and } \sigma_1 = \sigma_2 = \sigma = 5.5$$

Here, we wish to test that the mean height of Punjab and Tamilnadu women is same. If μ_1 and μ_2 denote the mean heights of Punjab and Tamilnadu women respectively then our claim is $\mu_1 = \mu_2$ and its complement is $\mu_1 \neq \mu_2$. Since the claim contains the equality sign so we can take the claim as the null hypothesis and complement as the alternative hypothesis. Thus,

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

Since the alternative hypothesis is two-tailed so the test is two-tailed test.

Here, we want to test the hypothesis regarding two population means. The standard deviations of both populations are known and sample sizes are large, so we should go for Z-test.

So, for testing the null hypothesis, the test statistic Z is given by

$$Z = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

$$= \frac{67.6 - 68.0}{\sqrt{\left(\frac{(5.5)^2}{1000} + \frac{(5.5)^2}{2000}\right)}} = \frac{-0.4}{5.5\sqrt{\left(\frac{1}{1000} + \frac{1}{2000}\right)}}$$

$$= \frac{-0.4}{5.5 \times 0.0387} = -1.88$$

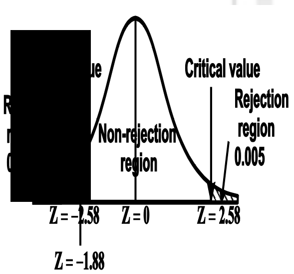

Fig. 10.6

The critical (tabulated) values for two-tailed test at 1% level of significance are $\pm z_{\alpha/2} = \pm z_{0.005} = \pm 2.58$.

Since calculated value of Z ($= -1.88$) is greater than the critical value ($= -2.58$) and less than the critical value ($= 2.58$), that means it lies in non-rejection region as shown in Fig. 10.6, so we do not reject the null hypothesis i.e. we fail to reject the claim.

Decision according to p-value:

The test is two-tailed, therefore,

$$\text{p-value} = 2P\left[Z \geq |z|\right] = 2P\left[Z \geq 1.88\right]$$

$$= 2\left[0.5 - P\left[0 \leq Z \leq 1.88\right]\right] = 2\left(0.5 - 0.4699\right) = 0.0602$$

Since p-value ($= 0.0602$) is greater than $\alpha (= 0.01)$ so we do not reject the null hypothesis at1% level of significance.

Thus, we conclude that the samples do not provide us sufficient evidence against the claim so we may assume that the average height of women of Punjab and Tamilnadu is same.

Example 4: A university conducts both face to face and distance mode classes for a particular course indented both to be identical. A sample of 50 students of face to face mode yields examination results mean and SD respectively as:

$$\bar{X} = 80.4, \quad S_1 = 12.8$$

and other sample of 100 distance-mode students yields mean and SD of their examination results in the same course respectively as:

$$\bar{Y} = 74.3, \quad S_2 = 20.5$$

Are both educational methods statistically equal at 5% level?

Solution: Here, we are given that

$$n_1 = 50, \quad \bar{X} = 80.4, \quad S_1 = 12.8;$$

$$n_2 = 100, \quad \bar{Y} = 74.3, \quad S_2 = 20.5$$

We wish to test that both educational methods are statistically equal. If μ_1 and μ_2 denote the average marks of face to face and distance mode students respectively then our claim is $\mu_1 = \mu_2$ and its complement is $\mu_1 \neq \mu_2$. Since the claim contains the equality sign so we can take the claim as the null hypothesis and complement as the alternative hypothesis. Thus,

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

Since the alternative hypothesis is two-tailed so the test is two-tailed test.

We want to test the null hypothesis regarding two population means when standard deviations of both populations are unknown. So we should go for t-test if population of difference is known to be normal. But it is not the case. Since sample sizes are large (n_1, and $n_2 > 30$) so we go for Z-test.

For testing the null hypothesis, the test statistic Z is given by

$$Z = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

$$= \frac{80.4 - 74.3}{\sqrt{\frac{(12.8)^2}{50} + \frac{(20.5)^2}{100}}} = \frac{6.1}{\sqrt{3.28 + 4.20}} = 2.23$$

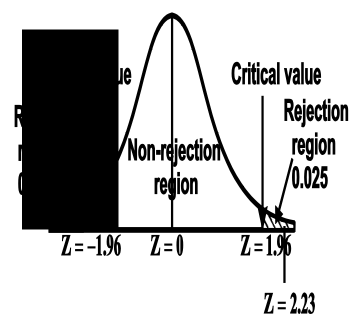

Fig. 10.7

The critical (tabulated) values for two-tailed test at 5% level of significance are $\pm z_{\alpha/2} = \pm z_{0.025} = \pm 1.96$.

Since calculated value of Z (= 2.23) is greater than the critical values (= ±1.96), that means it lies in rejection region as shown in Fig. 10.7, so we reject the null hypothesis i.e. we reject the claim at 5% level of significance.

Decision according to p-value:

The test is two-tailed, therefore,

$$\text{p-value} = 2P\left[Z \geq |z|\right] = 2P\left[Z \geq 2.23\right]$$

$$= 2\left[0.5 - P\left[0 \leq Z \leq 2.23\right]\right] = 2\left(0.5 - 0.4871\right) = 0.0258$$

Since p-value (= 0.0258) is less than $\alpha\,(= 0.05)$ so we reject the null hypothesis at 5% level of significance.

Thus, we conclude that samples provide us sufficient evidence against the claim so both methods of education, i.e. face-to-face and distance-mode, are not statistically equal.

Now, you can try the following exercises.

E6) Two brands of electric bulbs are quoted at the same price. A buyer was tested a random sample of 200 bulbs of each brand and found the following information:

	Mean Life (hrs.)	SD(hrs.)
Brand A	1300	41
Brand B	1280	46

Is there significant difference in the mean duration of their lives of two brands of electric bulbs at 1% level of significance?

E7) Two research laboratories have identically produced drugs that provide relief to BP patients. The first drug was tested on a group of 50 BP patients and produced an average 8.3 hours of relief with a standard deviation of 1.2 hours. The second drug was tested on 100 patients, producing an average of 8.0 hours of relief with a standard deviation of 1.5 hours. Does the first drug provide a significant longer period of relief at a significant level of 5%?

10.5 TESTING OF HYPOTHESIS FOR POPULATION PROPORTION USING Z-TEST

In Section 10.3, we have discussed the procedure of testing of hypothesis for population mean when sample size is large. But in many real world situations, in business and other areas where collected data are in form of counts or the collected data are classified into two categories or groups according to an attribute or a characteristic. For example, the peoples living in a colony may be classified into two groups (male and female) with respect to the characteristic sex, the patients in a hospital may be classified into two groups as cancer and non-cancer patients, the lot of articles may be classified as defective and non-defective, etc. Here, collected data are available in dichotomous or binary outcomes which is a special case of nominal scale and the data categorized into two mutually exclusive and exhaustive classes generally known as success and failure out comes. For example, the characteristic sex can be measured as success if male and failure if female or vice versa. So in such situations, proportion is suitable measure to apply.

In such situations, we require a test for testing a hypothesis about population proportion.

For this purpose, let $X_1, X_2, ..., X_n$ be a random sample of size n taken from a population with population proportion P. Also let X denotes the number of observations or elements possess a certain attribute (number of successes) out of n observations of the sample then sample proportion p can be defined as

$$p = \frac{X}{n} \le 1$$

As we have seen in Section 2.4 of the Unit 2 of this course that mean and variance of the sampling distribution of sample proportion are

$$E(p) = P \text{ and } Var(p) = \frac{PQ}{n}$$

where, Q = 1– P.

Now, two cases arise:

Case I: When sample size is not sufficiently large i.e. either of the conditions np > 5 or nq > 5 does not meet, then we use exact binomial test. But exact binomial test is beyond the scope of this course.

Case II: When sample size is sufficiently large, such that np > 5 and nq > 5 then by central limit theorem, the sampling distribution of sample proportion p is approximately normally distributed with mean and variance as

$$E(p) = P \text{ and } Var(p) = \frac{PQ}{n} \qquad \dots (5)$$

But we know that standard error = $\sqrt{\text{Variance}}$

$$\therefore \qquad SE(p) = \sqrt{\frac{PQ}{n}} \qquad \dots (6)$$

Now, follow the same procedure as we have discussed in Section 10.2, first of all we setup null and alternative hypotheses. Since here we want to test the hypothesis about specified value P_0 of the population proportion so we can take the null and alternative hypotheses as

$$H_0 : P = P_0 \text{ and } H_1 : P \neq P_0 \left[\text{for two-tailed test}\right] \begin{bmatrix} \text{Here, } \theta = P \text{ and } \theta_0 = P_0 \\ \text{if we compare it with} \\ \text{general procedure.} \end{bmatrix}$$

or

$$\left.\begin{matrix} H_0 : P \leq P_0 \text{ and } H_1 : P > P_0 \\ H_0 : P \geq P_0 \text{ and } H_1 : P < P_0 \end{matrix}\right\} \quad \left[\text{for one-tailed test}\right]$$

For testing the null hypothesis, the test statistic Z is given by

$$Z = \frac{p - E(p)}{SE(p)}$$

$$Z = \frac{p - P_0}{\sqrt{\frac{P_0 Q_0}{n}}} \sim N(0, 1) \text{ under } H_0 \left[\text{using equations (5) and (6)}\right]$$

After that, we calculate the value of test statistic and compare it with the critical value(s) given in **Table 10.1** at prefixed level of significance α. Take the decision about the null hypothesis as described in Section 10.2.

Let us do some examples of testing of hypothesis about population proportion.

Example 5: A machine produces a large number of items out of which 25% are found to be defective. To check this, company manager takes a random sample of 100 items and found 35 items defective. Is there an evidence of more deterioration of quality at 5% level of significance?

Solution: The company manager wants to check that his machine produces 25% defective items. Here, attribute under study is defectiveness. And we define our success and failure as getting a defective or non defective item.

Let P = Population proportion of defectives items = 0.25(= P_0)

p = Observed proportion of defectives items in the sample = 35/100 = 0.35

Here, we want to test that machine produces more defective items, that is, the proportion of defective items (P) greater than 0.25. So our claim is P > 0.25

and its complement is $P \leq 0.25$. Since complement contains the equality sign so we can take the complement as the null hypothesis and the claim as the alternative hypothesis. So

$$H_0 : P \leq P_0 = 0.25 \text{ and } H_1 : P > 0.25$$

Since the alternative hypothesis is right-tailed so the test is right-tailed test.

Before proceeding further, first we have to check whether the condition of normality meets or not.

$$\therefore \qquad np = 100 \times 0.35 = 35 > 5$$

$$nq = 100 \times (1 - 0.35) = 100 \times 0.65 = 65 > 5$$

We see that condition of normality meets, so we can go for Z-test.

So, for testing the null hypothesis, the test statistic Z is given by

$$Z = \frac{p - P_0}{\sqrt{\frac{P_0 Q_0}{n}}}$$

$$= \frac{0.35 - 0.25}{\sqrt{\frac{0.25 \times 0.75}{100}}} = \frac{0.10}{0.0433} = 2.31$$

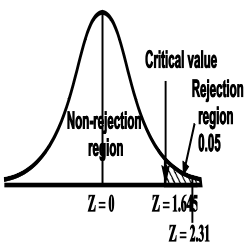

Fig. 10.8

Since test is right-tailed so the critical value at 5% level of significance is $z_\alpha = z_{0.05} = 1.645$.

Since calculated value of test statistic Z (= 2.31) is greater than the critical value (= 1.645), that means it lies in the rejection region as shown in Fig. 10.8. So we reject the null hypothesis and support the alternative hypothesis i.e. we support the claim at 5% level of significance.

Decision according to p-value:

The test is right-tailed, therefore,

$$\text{p-value} = P[Z \geq z] = P[Z \geq 2.31]$$

$$= 0.5 - P[0 < Z < 2.31] = 0.5 - 0.4896$$

$$= 0.0104$$

Since p-value (= 0.0104) is less than $\alpha (= 0.05)$ so we reject the null hypothesis at 5% level of significance.

Thus, we conclude that the sample fails to provide us sufficient evidence against the claim so we may assume that deterioration in quality exists at 5% level of significance.

Example 6: A die is thrown 9000 times and draw of 2 or 5 is observed 3100 times. Can we regard that die is unbiased at 5% level of significance.

Solution: Let getting a 2 or 5 be our success, and getting a number other than 2 or 5 be a failure then in usual notions, we have

$$n = 9000, X = \text{number of successes} = 3100, p = 3100/9000 = 0.3444$$

Here, we want to test that the die is unbiased and we know that if die is unbiased then proportion or probability of getting 2 or 5 is

P = Probability of getting a 2 or 5

= Probability of getting 2 + Probability of getting 5

$$=\frac{1}{6}+\frac{1}{6}=\frac{1}{3}=0.3333$$

So our claim is P = 0.3333 and its complement is $P \neq 0.3333$. Since the claim contains the equality sign so we can take the claim as the null hypothesis and complement as the alternative hypothesis. Thus,

$$H_0: P = P_0 = 0.3333 \text{ and } H_1: P \neq 0.3333$$

Since the alternative hypothesis is two-tailed so the test is two-tailed test.

Before proceeding further, first we have to check whether the condition of normality meets or not.

$$\therefore \quad np = 9000 \times 0.3444 = 3099.6 > 5$$

$$nq = 9000 \times (1 - 0.3444) = 9000 \times 0.6556 = 5900.4 > 5$$

We see that condition of normality meets, so we can go for Z-test.

So, for testing the null hypothesis, the test statistic Z is given by

$$Z = \frac{p - P_0}{\sqrt{\frac{P_0 Q_0}{n}}}$$

$$= \frac{0.3444 - 0.3333}{\sqrt{\frac{0.3333 \times 0.6667}{9000}}} = \frac{0.0111}{0.005} = 2.22$$

Since test is two-tailed so the critical values at 5% level of significance are $\pm z_{\alpha/2} = \pm z_{0.025} = \pm 1.96$.

Since calculated value of Z (= 2.22) is greater than the critical value (= 1.96), that means it lies in rejection region, so we reject the null hypothesis i.e. we reject our claim.

Decision according to p-value:

Since test is two-tailed so

$$\text{p-value} = 2P\left[Z \geq |z|\right] = 2P\left[Z \geq 2.22\right]$$

$$= 2\left[0.5 - P\left[0 < Z < 2.22\right]\right] = 2\left(0.5 - 0.4868\right) = 0.0264$$

Since p-value (= 0.0264) is less than $\alpha (= 0.05)$, so we reject the null hypothesis at 5% level of significance.

Thus, we conclude that the sample provides us sufficient evidence against the claim so die cannot be considered as unbiased.

Now, you can try the following exercises.

E8) In a sample of 100 MSc. Economics first year students of a University, it was seen that 54 students came from Science background and the rest are from other background. Can we assume that 50% of the students are from Science background in MSc. Economics first year students in the University at 1% level of significance?

E9) Out of 200 patients who are given a particular injection 180 survived. Test the hypothesis that the survival rate is more than 80% at 5% level of significance?

10.6 TESTING OF HYPOTHESIS FOR DIFFERENCE OF TWO POPULATION PROPORTIONS USING Z-TEST

In Section 10.5, we have discussed the testing of hypothesis about the population proportion. In some cases, we are interested to test the hypothesis about difference of two population proportions of an attributes in the two different populations or groups. For example, one may wish to test whether the proportions of alcohol drinkers in the two cities are same, one may wish to test proportion of literates in a group of people is greater than the proportion of literates in other group of people, etc. Therefore, we require the test for testing the hypothesis about the difference of two population proportions.

Let there be two populations, say, population-I and population-II under study. And also let we draw a random sample of size n_1 from population-I with population proportion P_1 and a random sample of size n_2 from population-II with population proportion P_2. If X_1 and X_2 are the number of observations /individuals / items / units possessing the given attribute in the sample of sizes n_1 and n_2 respectively then sample proportions can be defined as

$$p_1 = \frac{X_1}{n_1} \text{ and } p_2 = \frac{X_2}{n_2}$$

As we have seen in Section 2.5 of the Unit 2 of this course that mean and variance of the sampling distribution of difference of sample proportions are

$$E(p_1 - p_2) = P_1 - P_2$$

and variance

$$\text{Var}(p_1 - p_2) = \frac{P_1Q_1}{n_1} + \frac{P_2Q_2}{n_2}$$

where, $Q_1 = 1 - P_1$ and $Q_2 = 1 - P_2$.

Now, two cases arise:

Case I: When sample sizes are not sufficiently large i.e. any of the conditions $n_1p_1 > 5$ or $n_1q_1 > 5$ or $n_2p_2 > 5$ or $n_2q_2 > 5$ does not meet, then we use exact binomial test. But exact binomial test is beyond the scope of this course.

Case II: When sample sizes are sufficiently large, such that $n_1p_1 > 5$, $n_1q_1 > 5$, $n_2p_2 > 5$ and $n_2q_2 > 5$ then by central limit theorem, the sampling distribution of sample proportions p_1 and p_2 are approximately normally as

$$p_1 \sim N\left(P_1, \frac{P_1Q_1}{n_1}\right) \text{ and } p_2 \sim N\left(P_2, \frac{P_2Q_2}{n_2}\right)$$

Also, by the property of normal distribution described in Unit 13 of MST-003, the sampling distribution of the difference of sample proportions follows normal distribution with mean

$$E(p_1 - p_2) = E(p_1) - E(p_2) = P_1 - P_2 \qquad \ldots (7)$$

and variance

$$Var(p_1 - p_2) = Var(p_1) + Var(p_2) = \frac{P_1Q_1}{n_1} + \frac{P_2Q_2}{n_2}$$

That is,

$$p_1 - p_2 \sim N\left(P_1 - P_2, \frac{P_1Q_1}{n_1} + \frac{P_2Q_2}{n_2}\right)$$

Thus, standard error is given by

$$SE(p_1 - p_2) = \sqrt{Var(p_1 - p_2)} = \sqrt{\frac{P_1Q_1}{n_1} + \frac{P_2Q_2}{n_2}} \qquad \ldots (8)$$

Now, follow the same procedure as we have discussed in Section 10.2, first of all we have to setup null and alternative hypotheses. Here, we want to test the hypothesis about the difference of two population proportions so we can take the null hypothesis as

$$H_0 : P_1 = P_2 \text{ (no difference in proportions)} \left[\begin{array}{l}\text{Here, } \theta_1 = P_1 \text{ and} \\ \theta_2 = P_2 \text{ if we compare} \\ \text{it with general} \\ \text{procedure.}\end{array}\right]$$

or $\quad H_0 : P_1 - P_2 = 0$ (difference in two proportions is 0)

and the alternative hypothesis may be

$$H_1 : P_1 \neq P_2 \qquad [\text{for two-tailed test}]$$

$$\text{or} \quad \left.\begin{array}{l} H_0 : P_1 \leq P_2 \text{ and } H_1 : P_1 > P_2 \\ H_0 : P_1 \geq P_2 \text{ and } H_1 : P_1 < P_2 \end{array}\right\} \qquad [\text{for one-tailed test}]$$

For testing the null hypothesis, the test statistic Z is given by

$$Z = \frac{(p_1 - p_2) - E(p_1 - p_2)}{SE(p_1 - p_2)}$$

$$\text{or} \quad Z = \frac{(p_1 - p_2) - (P_1 - P_2)}{\sqrt{\frac{P_1Q_1}{n_1} + \frac{P_2Q_2}{n_2}}} \qquad [\text{using equations (7) and (8)}]$$

Since under null hypothesis we assume that $P_1 = P_2 = P$, therefore, we have

$$Z = \frac{p_1 - p_2}{\sqrt{PQ\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

where, Q = 1 - P.

Generally, P is unknown then it is estimated by the value of pooled proportion $\hat{P}$, where

$$\hat{P} = \frac{n_1p_1 + n_2p_2}{n_1 + n_2} = \frac{X_1 + X_2}{n_1 + n_2} \text{ and } \hat{Q} = 1 - \hat{P}$$

After that, we calculate the value of test statistic and compare it with the critical value(s) given in **Table 10.1** at prefixed level of significance α. Take the decision about the null hypothesis as described in Section 10.2.

Now, it is time for doing some examples for testing of hypothesis about the difference of two population proportions.

Example 7: In a random sample of 100 persons from town A, 60 are found to be high consumers of wheat. In another sample of 80 persons from town B, 40 are found to be high consumers of wheat. Do these data reveal a significant difference between the proportions of high wheat consumers in town A and town B (at $\alpha = 0.05$)?

Solution: Here, attribute under study is high consuming of wheat. And we define our success and failure as getting a person of high consumer of wheat and not high consumer of wheat respectively.

We are given that

n_1 = total number of persons in the sample of town A = 100

n_2 = total number of persons in the sample of town B = 80

X_1 = number of persons of high consumer of wheat in town A = 60

X_2 = number of persons of high consumer of wheat in town B = 40

The sample proportion of high wheat consumers in town A is

$$p_1 = \frac{X_1}{n_1} = \frac{60}{100} = 0.60$$

and the sample proportion of wheat consumers in town B is

$$p_2 = \frac{X_2}{n_2} = \frac{40}{80} = 0.50$$

Here, we want to test that the proportion of high consumers of wheat in two towns, say, P_1 and P_2, is not same. So our claim is $P_1 \neq P_2$ and its complement is $P_1 = P_2$. Since the complement contains the equality sign, so we can take the complement as the null hypothesis and the claim as the alternative hypothesis. Thus,

$$H_0 : P_1 = P_2 = P \text{ and } H_1 : P_1 \neq P_2$$

Since the alternative hypothesis is two-tailed so the test is two-tailed test.

Before proceeding further, first we have to check whether the condition of normality meets or not.

$\therefore \quad n_1p_1 = 100 \times 0.60 = 60 > 5,\ n_1q_1 = 100 \times 0.40 = 40 > 5$

$n_2p_2 = 80 \times 0.50 = 40 > 5,\ n_2q_2 = 80 \times 0.50 = 40 > 5$

We see that condition of normality meets, so we can go for Z-test.

The estimate of the combined proportion (P) of high wheat consumers in two towns is given by

$$\hat{P} = \frac{n_1p_1 + n_2p_2}{n_1 + n_2} = \frac{X_1 + X_2}{n_1 + n_2} = \frac{60 + 40}{100 + 80} = \frac{5}{9}$$

$$\hat{Q} = 1 - \hat{P} = 1 - \frac{5}{9} = \frac{4}{9}$$

For testing the null hypothesis, the test statistic Z is given by

$$Z = \frac{p_1 - p_2}{\sqrt{\hat{P}\hat{Q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$= \frac{0.60 - 0.50}{\sqrt{\frac{5}{9} \times \frac{4}{9}\left(\frac{1}{100} + \frac{1}{80}\right)}} = \frac{0.10}{0.0745} = 1.34$$

The critical values for two-tailed test at 5% level of significance are $\pm z_{\alpha/2}$ $= \pm z_{0.025} = \pm 1.96$.

Since calculated value of Z (=1.34) is less than the critical value (= 1.96) and greater than critical value (= −1.96), that means calculated value of Z lies in non-rejection region, so we do not reject the null hypothesis and reject the alternative hypothesis i.e. we reject the claim.

Decision according to p-value:

Since the test is two-tailed, therefore

$$\text{p-value} = 2\,P[Z \geq z] = 2P[Z \geq 1.34]$$

$$= 2\left[0.5 - P[0 \leq Z \leq 1.34]\right] = 2(0.5 - 0.4099) = 0.1802$$

Since p-value (= 0.1802) is greater than $\alpha(= 0.05)$ so we do not reject the null hypothesis at 5% level of significance.

Thus, we conclude that the samples provide us the sufficient evidence against the claim so we may assume that the proportion of high consumers of wheat in two towns A and B is same.

Example 8: A machine produced 60 defective articles in a batch of 400. After overhauling it produced 30 defective in a batch of 300. Has the machine improved due to overhauling? (Take $\alpha = 0.01$).

Solution: Here, the machine produced articles and attribute under study is defectiveness. And we define our success and failure as getting a defective or non defective article. Therefore, we are given that

X_1 = number of defective articles produced by the machine before overhauling = 60

X_2 = number of defective articles produced by the machine after overhauling = 30 and $n_1 = 400, \ n_2 = 300,$

Let p_1 = Observed proportion of defective articles in the sample before the overhauling

$$= \frac{X_1}{n_1} = \frac{60}{400} = 0.15$$

and p_2 = Observed proportion of defective articles in the sample after the overhauling

$$= \frac{X_2}{n_2} = \frac{30}{300} = 0.10$$

Here, we want to test that machine improved due to overhauling that means the proportion of defective articles is less after overhauling. If P_1 and P_2 denote the proportion defectives before and after the overhauling the machine so our claim is $P_1 > P_2$ and its complement $P_1 \le P_2$. Since the complement contains the equality sign so we can take the complement as the null hypothesis and claim as the alternative hypothesis. Thus,

$$H_0 : P_1 \le P_2 \text{ and } H_1 : P_1 > P_2$$

Since the alternative hypothesis is right-tailed so the test is right-tailed test.

Since P is unknown, so the pooled estimate of proportion is given by

$$\hat{P} = \frac{X_1 + X_2}{n_1 + n_2} = \frac{60+30}{400+300} = \frac{90}{700} = \frac{9}{70} \text{ and } \hat{Q} = 1 - \hat{P} = 1 - \frac{9}{70} = \frac{61}{70}.$$

Before proceeding further, first we have to check whether the condition of normality meets or not.

$$\therefore \quad n_1p_1 = 400 \times 0.15 = 60 > 5, \ n_1q_1 = 400 \times 0.85 = 340 > 5$$

$$n_2p_2 = 300 \times 0.10 = 30 > 5, \ n_2q_2 = 300 \times 0.90 = 270 > 5$$

We see that condition of normality meets, so we can go for Z-test.

For testing the null hypothesis, the statistic is given by

$$Z = \frac{p_1 - p_2}{\sqrt{\hat{P}\hat{Q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$= \frac{0.15 - 0.10}{\sqrt{\frac{9}{70} \times \frac{61}{70}\left(\frac{1}{400} + \frac{1}{300}\right)}} = \frac{0.05}{0.0256} = 1.95$$

The critical value for right-tailed test at 1% level of significance is $z_\alpha = z_{0.01} = 2.33$.

Since calculated value of Z (= 1.95) is less than the critical value (= 2.33) that means calculated value of Z lies in non-rejection region, so we do not reject the null hypothesis and reject the alternative hypothesis i.e. we reject the claim at 1% level of significance.

Decision according to p-value:

Since the test is right-tailed, therefore,

p-value = $P[Z \ge z] = P[Z \ge 1.95]$

$$= 0.5 - P[0 \le Z \le 1.95] = 0.5 - 0.4744 = 0.0256$$

Since p-value (= 0.0256) is greater than $\alpha (= 0.01)$ so we do not reject the null hypothesis at1% level of significance.

Thus, we conclude that the samples provide us sufficient evidence against the claim so the machine has not been improved after overhauling.

Now, you can try the following exercises.

E10) The proportions of literates between groups of people of two districts A and B are tested. Out of the 100 persons selected at random from each district, 50 from district A and 40 from district B are found literates. Test whether the proportion of literate persons in two districts A and B is same at 1% level of significance?

E11) In a large population 30% of a random sample of 1200 persons had blue-eyes and 20% of a random sample of 900 persons had the same blue-eyes in another population. Test the proportion of blue-eyes persons is same in two populations at 5% level of significance.

10.7 TESTING OF HYPOTHESIS FOR POPULATION VARIANCE USING Z-TEST

In Section 10.3, we have discussed testing of hypothesis for population mean but when analysing quantitative data, it is often important to draw conclusion about the average as well as the variability of a characteristic of under study. For example, a company manufactured the electric bulbs and the manager of the company would probably be interested in determining the average life of the bulbs and also determining whether or not the variability in the life of bulbs is within acceptable limits, the product controller of a milk company may be interested to know variance of the amount of fat in the whole milk processed by the company is no more than the specified level, etc. So we require a test for this purpose.

The procedure of testing a hypothesis for population variance or standard deviation is similar to the testing of population mean.

For testing a hypothesis about the population variance, we draw a random sample $X_1, X_2, ..., X_n$ of size $n > 30$ from the population with mean μ and variance σ^2 where, μ be known or unknown.

We know that by central limit theorem that sample variance is asymptotically normally distributed with mean σ^2 and variance $2\sigma^4/n$ whether parent population is **normal or non-normal**. That is, if S^2 is the sample variance of the random sample then

$$E(S^2) = \sigma^2 \text{ and } Var(S^2) = \frac{2\sigma^4}{n} \qquad \dots (9)$$

But we know that standard error = $\sqrt{\text{Variance}}$

$$\therefore \quad SE(S^2) = \sqrt{Var(S^2)} = \sqrt{\frac{2}{n}}\sigma^2 \qquad \dots (10)$$

The general procedure of this test is explained in the next page.

As we are doing so far in all tests, first Step in hypothesis testing problems is to setup null and alternative hypotheses. Here, we want to test the hypothesis specified value σ_0^2 of the population variance σ^2 so we can take our null and alternative hypotheses as

$$H_0 : \sigma^2 = \sigma_0^2 \text{ and } H_1 : \sigma^2 \neq \sigma_0^2 \qquad [\text{for two-tailed test}]$$

or

$$\left.\begin{array}{l} H_0 : \sigma^2 \leq \sigma_0^2 \text{ and } H_1 : \sigma^2 > \sigma_0^2 \\ H_0 : \sigma^2 \geq \sigma_0^2 \text{ and } H_1 : \sigma^2 < \sigma_0^2 \end{array}\right\} \qquad [\text{for one-tailed test}]$$

For testing the null hypothesis, the test statistic Z is given by

$$Z = \frac{S^2 - E(S^2)}{SE(S^2)} \sim N(0,1)$$

$$Z = \frac{S^2 - \sigma_0^2}{\sigma_0^2\sqrt{\frac{2}{n}}} \qquad \left[\begin{array}{l}\text{Using equations (9) and (10) and} \\ \text{under } H_0 : \sigma^2 = \sigma_0^2\end{array}\right]$$

After that, we calculate the value of test statistic and compare it with the critical value given in **Table 10.1** at prefixed level of significance α. Take the decision about the null hypothesis as described in Section 10.2.

Note 2: When population under study is normal then for testing the hypothesis about population variance or population standard deviation we use chi-square test which will be discussed in Unit 12 of this course. Whereas when the distribution of the population under study is not known and sample size is large then we apply Z-test as discussed above.

Now, it is time to do example based on above test.

Example 9: A random sample of size 65 screws is taken from a population of big box of screws and measured their length (in mm) which gives sample variance 9.0. Test that the two years old population variance 10.5 is still maintained at present at 5% level of significance.

Solution: We are given that

$$n = 65, \quad S^2 = 9.0, \quad \sigma_0^2 = 10.5$$

Here, we want to test that the two years old screw length population variance (σ^2) is still maintained at 10.5. So our claim is $\sigma^2 = 10.5$ and its complement is $\sigma^2 \neq 10.5$. Since the claim contains the equality sign so we can take the claim as the null hypothesis and complement as the alternative hypothesis. Thus,

$$H_0 : \sigma^2 = \sigma_0^2 = 10.5 \text{ and } H_1 : \sigma^2 \neq 10.5$$

Since the alternative hypothesis is two-tailed so the test is two-tailed test.

Here, the distribution of population under study is not known and sample size is large (n > 30) so we can go for Z-test.

For testing the null hypothesis, the test statistic Z is given by

$$Z = \frac{S^2 - \sigma_0^2}{\sigma_0^2\sqrt{\frac{2}{n}}} \sim N(0,1)$$

$$= \frac{9.0 - 10.5}{10.5\sqrt{\frac{2}{65}}} = \frac{-1.5}{10.5 \times 0.175} = \frac{-1.5}{1.84} = -0.81$$

The critical values for two-tailed test at 5% level of significance are $\pm z_{\alpha/2} = \pm z_{0.025} = \pm 1.96$.

Since calculated value of Z (= −0.81) is less than critical value (= 1.96) and greater than the critical value (= −1.96), that means it lies in non-rejection region, so we do not reject the null hypothesis i.e. we support the claim.

Thus, we conclude that sample fails to provide us sufficient evidence against the claim so we may assume that two years old screw length population variance is still maintained at 10.5mm.

Now, you can try the following exercise.

E12) A random sample of size 120 bulbs is taken from a lot which gives the standard deviation of the life of electric bulbs 7 hours. Test the standard deviation of the life of bulbs of the lot is 6 hours at 5% level of significance.

10.8 TESTING OF HYPOTHESIS FOR TWO POPULATION VARIANCES USING Z-TEST

In previous section, we have discussed testing of hypothesis about the population variance. But there are so many situations where we want to test the hypothesis about equality of two population variances or standard deviations. For example, an economist may want to test whether the variability in incomes differ in two populations, a quality controller may want to test whether the quality of the product is changing over time, etc.

Let there be two populations, say, population-I and population-II under study. Also let μ_1, μ_2 and σ_1^2, σ_2^2 denote the means and variances of population-I and population-II respectively where both σ_1^2 and σ_2^2 are unknown but μ_1 and μ_2 may be known or unknown. For testing the hypothesis about equality of two population variances or standard deviations, we draw a random sample of large size n_1 from population-I and a random sample of large size n_2 from population-II. Let S_1^2 and S_2^2 be the sample variances of the samples selected from population-I and population-II respectively.

These two populations **may or may not be normal** but according to the central limit theorem, the sampling distribution of difference of two large sample variances asymptotically normally distributed with mean $(\sigma_1^2 + \sigma_2^2)$ and variance $(2\sigma_1^4/n_1 + 2\sigma_2^4/n_2)$.

Thus,

$$E\left(S_1^2 - S_2^2\right) = E\left(S_1^2\right) - E\left(S_2^2\right) = \sigma_1^2 - \sigma_2^2 \qquad \ldots (11)$$

and

$$\text{Var}\left(S_1^2 - S_2^2\right) = \text{Var}\left(S_1^2\right) + \text{Var}\left(S_2^2\right) = \frac{2\sigma_1^4}{n_1} + \frac{2\sigma_2^4}{n_2}$$

But we know that standard error = $\sqrt{\text{Variance}}$

$$\therefore \quad SE\left(S_1^2 - S_2^2\right) = \sqrt{Var\left(S_1^2 - S_2^2\right)} = \sqrt{\frac{2\sigma_1^4}{n_1} + \frac{2\sigma_2^4}{n_2}} \quad \ldots (12)$$

Now, follow the same procedure as we have discussed in Section 10.2, that is, first of all we have to setup null and alternative hypothesis. Here, we want to test the hypothesis about the two population variances, so we can take our null and alternative hypotheses as

$$H_0 : \sigma_1^2 = \sigma_2^2 = \sigma^2 \text{ and } H_1 : \sigma_1^2 \neq \sigma_2^2 \qquad [\text{for two-tailed test}]$$

$$\text{or} \quad \left.\begin{array}{l} H_0 : \sigma_1^2 \leq \sigma_2^2 \text{ and } H_1 : \sigma_1^2 > \sigma_2^2 \\ H_0 : \sigma_1^2 \geq \sigma_2^2 \text{ and } H_1 : \sigma_1^2 < \sigma_2^2 \end{array}\right\} \qquad [\text{for one-tailed test}]$$

For testing the null hypothesis, the test statistic Z is given by

$$Z = \frac{\left(S_1^2 - S_2^2\right) - E\left(S_1^2 - S_2^2\right)}{SE\left(S_1^2 - S_2^2\right)} \sim N(0,1)$$

$$\text{or} \quad Z = \frac{\left(S_1^2 - S_2^2\right) - \left(\sigma_1^2 - \sigma_2^2\right)}{\sqrt{\frac{2\sigma_1^4}{n_1} + \frac{2\sigma_2^4}{n_2}}} \quad [\text{using equations (11) and (12)}]$$

Since under null hypothesis we assume that $\sigma_1^2 = \sigma_2^2 = \sigma^2$, therefore, we have

$$Z = \frac{S_1^2 - S_2^2}{\sigma^2 \sqrt{2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \sim N(0,1)$$

Generally, population variances σ_1^2 and σ_2^2 are unknown, so we estimate them by their corresponding sample variances S_1^2 and S_2^2 as

$$\hat{\sigma}_1^2 = S_1^2 \quad \text{and} \quad \hat{\sigma}_2^2 = S_2^2$$

Thus, the test statistic Z is given by

$$Z = \frac{S_1^2 - S_2^2}{\sqrt{\left(\frac{2S_1^4}{n_1} + \frac{2S_2^4}{n_2}\right)}} \sim N(0,1)$$

After that, we calculate the value of test statistic as may be the case and compare it with the critical value given in **Table 10.1** at prefixed level of significance α. Take the decision about the null hypothesis as described in Section 10.2.

Note 3: When populations under study are normal then for testing the hypothesis about equality of population variances we use F- test which will be discussed in Unit 12 of this course. Whereas when the form of the populations under study is not known and sample sizes are large then we apply Z-test as discussed above.

Now, it is time to do an example based on above test.

Example 10: A comparative study of variation in weights (in pound) of Army-soldiers and Navy- sailors was made. The sample variance of the weight of 120 soldiers was 60 pound^2 and the sample variance of the weight of 160 sailors was 70 pound^2. Test whether the soldiers and sailors have equal variation in their weights. Use 5% level of significance.

Solution: Given that

$$n_1 = 120,\ \ S_1^2 = 60,\ \ n_2 = 160\ ,\ \ S_2^2 = 70$$

We want to test that the Army-soldiers and Navy-sailors have equal variation in their weights. If σ_1^2 and σ_2^2 denote the variances in the weight of Army-soldiers and Navy-sailors so our claim is $\sigma_1^2 = \sigma_2^2$ and its complement is $\sigma_1^2 \neq \sigma_2^2$. Since the claim contains the equality sign so we can take the claim as the null hypothesis and complement as the alternative hypothesis. Thus,

$$H_0 : \sigma_1^2 = \sigma_2^2 \ \left[\begin{array}{l}\text{Army-soldiers and Navy-sailors} \\ \text{have equal variation in their weights}\end{array}\right]$$

and the alternative hypothesis as

$$H_1 : \sigma_1^2 \neq \sigma_2^2 \ \left[\begin{array}{l}\text{Army-soldiers and Navy-sailors} \\ \text{have different variation in their weights}\end{array}\right]$$

Here, the distributions of populations under study are not known and sample sizes are large $\left(n_1 = 120 > 30,\ n_2 = 160 > 30\right)$ so we can go for Z-test.

Since population variances are unknown so for testing the null hypothesis, the test statistic Z is given by

$$Z = \frac{S_1^2 - S_2^2}{\sqrt{\left(\frac{2S_1^4}{n_1} + \frac{2S_2^4}{n_2}\right)}}$$

$$= \frac{60 - 70}{\sqrt{\frac{2 \times (60)^2}{120} + \frac{2 \times (70)^2}{160}}}$$

$$= \frac{-10}{\sqrt{60.0 + 61.25}} \quad \frac{-10}{11.01} = -0.91$$

The critical values for two-tailed test at 5% level of significance are $\pm z_{\alpha/2} = \pm z_{0.025} = \pm 1.96$.

Since calculated value of Z (= −0.91) is less than critical value (= 1.96) and greater than the critical value (= −1.96), that means it lies in non-rejection region, so we do not reject the null hypothesis i.e. we support the claim.

Thus, we conclude that samples fail to provide us sufficient evidence against the claim so we may assume that the Army-soldiers and Navy-sailors have equal variation in their weights.

Now, you can try the following exercise.

E 13) Two sources of raw materials of bulbs are under consideration by a bulb manufacturing company. Both sources seem to have similar characteristics but the company is not sure about their respective uniformity. A sample of 52 lots from source A yields variance 25 and a sample of 40 lots from source B yields variance of 12. Test whether the variance of source A significantly differs to the variances of source B at $\alpha = 0.05$?

We now end this unit by giving a summary of what we have covered in it.

10.9 SUMMARY

In this unit we have covered the following points:

1. How to judge a given situation whether we should go for large sample test or not.
2. Applying the Z-test for testing the hypothesis about the population mean and difference of two population means.
3. Applying the Z-test for testing the hypothesis about the population proportion and difference of two population proportions.
4. Applying the Z-test for testing the hypothesis about the population variance and two population variances.

10.10 SOLUTIONS / ANSWERS

E1) Since we have a rule that if the observed value of test statistic lies in rejection region then we reject the null hypothesis and if calculated value of test statistic lies in non-rejection region then we do not reject the null hypothesis. Therefore in our case, we do not reject the null hypothesis. So (iii) is the correct answer. Remember always on the basis of the one sample we never accept the null hypothesis.

E2) Since the test is two-tailed therefore rejection region will lie under both tails.

E3) Since test is two-tailed, therefore,

$$\text{p-value} = 2P\left[Z \geq |z|\right] = 2P\left[Z \geq 2.42\right]$$

$$= 2\left[0.5 - P\left[0 \leq Z \leq 2.42\right]\right] = 2\left(0.5 - 0.4922\right) = 0.0156$$

E4) We are given that

$$n = 900,\ \bar{X} = 3.4 \text{ cm},\ \mu_0 = 3.25 \text{ cm and } \sigma = 2.61 \text{ cm}$$

Here, we wish to test that the sample comes from a large population of bolts with mean (μ) 3.25cm. So our claim is $\mu = 3.25$cm and its complement is $\mu \neq 3.25$cm. Since the claim contains the equality sign so we can take the claim as the null hypothesis and the complement as the alternative hypothesis. Thus,

$$H_0 : \mu = \mu_0 = 3.25 \text{ and } H_1 : \mu \neq 3.25$$

Since the alternative hypothesis is two-tailed, so the test is two-tailed test.

Here, we want to test the hypothesis regarding population mean when population SD is unknown, so we should use t-test if the population of bolts known to be normal. But it is not the case. Since the sample size is large (n > 30) so we can go for Z-test instead of t-test as an approximate. So test statistic is given by

$$Z = \frac{\overline{X} - \mu_0}{\sigma / \sqrt{n}}$$

$$= \frac{3.40 - 3.25}{2.61/\sqrt{900}} = \frac{0.15}{0.087} = 1.72$$

The critical (tabulated) values for two-tailed test at 5% level of significance are $\pm z_{\alpha/2} = \pm z_{0.025} = \pm 1.96$.

Since calculated value of test statistic Z (= 1.72) is less than the critical value (= 1.96) and greater than critical value (= −1.96), that means it lies in non-rejection region, so we do not reject the null hypothesis i.e. we support the claim at 5% level of significance.

Thus, we conclude that sample does not provide us sufficient evidence against the claim so we may assume that the sample comes from the population of bolts with mean 3.25cm.

E5) Here, we given that

$$\mu_0 = 1200,\ n = 100,\ \overline{X} = 1220,\ S = 90$$

Here, the company may accept the new CFL light when average life of CFL light is greater than 1200 hours. So the company wants to test that the new brand CFL light has an average life greater than 1200 hours. So our claim is $\mu > 1200$ and its complement is $\mu \leq 1200$. Since complement contains the equality sign so we can take the complement as the null hypothesis and the claim as the alternative hypothesis. Thus,

$$H_0 : \mu \leq \mu_0 = 1200 \text{ and } H_1 : \mu > 1200$$

Since the alternative hypothesis is right-tailed so the test is right-tailed test.

Here, we want to test the hypothesis regarding population mean when population SD is unknown, so we should use t-test if the distribution of life of bulbs known to be normal. But it is not the case. Since the sample size is large (n > 30) so we can go for Z-test instead of t-test. Therefore, test statistic is given by

$$Z = \frac{\overline{X} - \mu_0}{S / \sqrt{n}}$$

$$= \frac{1220 - 1200}{90/\sqrt{100}} = \frac{20}{9} = 2.22$$

The critical values for right-tailed test at 5% level of significance is $z_{\alpha} = z_{0.05} = 1.645$.

Since calculated value of test statistic Z (= 2.22) is greater than critical value (= 1.645), that means it lies in rejection region so we reject the null hypothesis and support the alternative hypothesis i.e. we support our claim at 5% level of significance.

Thus, we conclude that sample does not provide us sufficient evidence against the claim so we may assume that the company accepts the new brand of bulbs.

E6) Given that

$$n_1 = 200, \quad \bar{X} = 1300, \quad S_1 = 41;$$

$$n_2 = 200, \quad \bar{Y} = 1280, \quad S_2 = 46$$

Here, we want to test that there is significant difference in the mean duration of their lives of two brands of electric bulbs. If μ_1 and μ_2 denote the mean lives of two brands of electric bulbs respectively then our claim is $\mu_1 \neq \mu_2$ and its complement is $\mu_1 = \mu_2$. Since the complement contains the equality sign so we can take the complement as the null hypothesis and the claim as the alternative hypothesis. Thus,

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

Since the alternative hypothesis is two-tailed so the test is two-tailed test.

We want to test the null hypothesis regarding equality of two population means. The standard deviations of both populations are unknown so we should go for t-test if population of difference is known to be normal. But it is not the case. Since sample sizes are large (n_1, and $n_2 > 30$) so we go for Z-test.

So for testing the null hypothesis, the test statistic Z is given by

$$Z = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

$$= \frac{1300 - 1280}{\sqrt{\frac{(41)^2}{200} + \frac{(46)^2}{200}}} = \frac{20}{\sqrt{8.41 + 10.58}} = \frac{20}{4.36} = 4.59$$

The critical (tabulated) values for two-tailed test at 1% level of significance are $\pm z_{\alpha/2} = \pm z_{0.005} = \pm 2.58$.

Since calculated value of test statistic Z (= 4.59) is greater than the critical values (= ± 2.58), that means it lies in rejection region, so we reject the null hypothesis an support the alternative hypothesis i.e. support the claim at 1% level of significance.

Thus, we conclude that samples do not provide us sufficient evidence against the claim so there is significant difference in the mean duration of their lives of two brands of electric bulbs.

E7) Given that

$$n_1 = 50, \quad \bar{X} = 8.3, \quad S_1 = 1.2;$$

$$n_2 = 100, \quad \bar{Y} = 8.0, \quad S_2 = 1.5$$

Here, we want to test that the first drug provides a significant longer period of relief than the other. If μ_1 and μ_2 denote the mean relief time due to first and second drugs respectively then our claim is $\mu_1 > \mu_2$ and its complement is $\mu_1 \le \mu_2$. Since complement contains the equality sign so we can take the complement as the null hypothesis and the claim as the alternative hypothesis. Thus,

$$H_0 : \mu_1 \le \mu_2 \text{ and } H_1 : \mu_1 > \mu_2$$

Since the alternative hypothesis is right-tailed so the test is right-tailed test.

We want to test the null hypothesis regarding equality of two population means. The standard deviations of both populations are unknown. So we should go for t-test if population of difference is known to be normal. But it is not the case. Since sample sizes are large (n_1, and $n_2 > 30$) so we go for Z-test.

So for testing the null hypothesis, the test statistic Z is given by

$$Z = \frac{\overline{X} - \overline{Y}}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

$$= \frac{8.3 - 8.0}{\sqrt{\frac{(1.2)^2}{50} + \frac{(1.5)^2}{100}}} = \frac{0.3}{\sqrt{0.0288 + 0.0255}}$$

$$= \frac{0.3}{0.2265} = 1.32$$

The critical (tabulated) value for right-tailed test at 5% level of significance is $z_\alpha = z_{0.05} = 1.645$.

Since calculated value of test statistic Z (= 1.32) is less than the critical value (=1.645), that means it lies in non-rejection region, so we do not reject the null hypothesis and reject the alternative hypothesis i.e. we reject the claim at 5% level of significance.

Thus, we conclude that samples provide us sufficient evidence against the claim so the first drug does not has longer period of relief than the other.

E8) Here, students are classified as Science background and other. We define our success and failure as getting student of Science background and other respectively. We are given that

n = Total number of students is the sample = 100

X = Number of students from Science background = 54

p = Sample proportion of students from Science background

$$= \frac{54}{100} = 0.54$$

We want to test whether 50% of the students are from Science background in MSc. If P denotes the proportion of first year Science background students in the University. So our claim is $P = P_0 = 0.5$ and its complement is $P \ne 0.5$. Since the claim contains the equality sign so we can take the claim as the null hypothesis and complement as the alternative hypothesis. Thus,

$H_0 = P = P_0 = 0.5$ (= 50%)

H_1: $P \neq 0.5$ [Science background differs to 50%]

Since the alternative hypothesis is two-tailed so the test is two-tailed test.

Before proceeding further, first we have to check whether the condition of normality meets or not.

$\therefore \quad np = 100 \times 0.54 = 54 > 5$

$nq = 100 \times (1 - 0.54) = 100 \times 0.46 = 46 > 5$

We see that condition of normality meets, so we can go for Z-test.

For testing the null hypothesis, the test statistic Z is given by

$$Z = \frac{p - P_0}{\sqrt{\frac{P_0 Q_0}{n}}} = \frac{0.54 - 0.50}{\sqrt{0.5 \times 0.5 / 100}} \qquad [\because Q_0 = 1 - P_0]$$

$$= \frac{0.04}{0.05} = 0.80$$

The critical (tabulated) values for two-tailed test at 1% level of significance are $\pm z_{\alpha/2} = \pm z_{0.005} = \pm 2.58$.

Since calculated value of test statistic Z (= 0.80) is less than the critical value (= 2.58) and greater than critical value (= −2.58), that means it lies in non-rejection region, so we do not reject the null hypothesis i.e. we support the claim at 1% level of significance.

Thus, we conclude that sample fails to provide us sufficient evidence against the claim so we may assume that 50% of the first year students in MSc. Economics in the University are from Science background.

E9) We define our success and failure as a patient is survived and not survived. Here, we are given that

n = 200

X = Number of survived patients who are given a particular injection

= 180

p = Sample proportion of survived patients who are given a particular injection

$$= \frac{X}{n} = \frac{180}{200} = 0.9$$

$$P_0 = 80\% = \frac{80}{100} = 0.80 \Rightarrow Q_0 = 1 - P_0 = \frac{80}{100} = 0.20$$

Here, we want to test that the survival rate of the patients is more than 80%. If P denotes the proportion of survival patients then our claim is P > 0.80 and its complement is P ≤ 0.80. Since complement contains the equality sign so we can take the complement as the null hypothesis and the claim as the alternative hypothesis. Thus,

$$H_0 : P \leq P_0 = 0.80 \text{ and } H_1 : P > 0.80$$

Since the alternative hypothesis is right-tailed so the test is right-tailed test.

Before proceeding further, first we have to check whether the condition of normality meets or not.

$$\therefore \quad np = 200 \times 0.9 = 180 > 5$$

$$nq = 200 \times (1 - 0.9) = 200 \times 0.1 = 20 > 5$$

We see that condition of normality meets, so we can go for Z-test.

For testing the null hypothesis, the test statistic Z is given by

$$Z = \frac{p - P_0}{\sqrt{\frac{P_0 Q_0}{n}}}$$

$$= \frac{0.9 - 0.8}{\sqrt{\frac{(0.8)(0.2)}{200}}} = \frac{0.1}{0.0283} = 3.53$$

The critical (tabulated) value for right-tailed test at 5% level of significance is $z_{\alpha} = z_{0.05} = 1.645$.

Since calculated value of test statistic Z (= 3.53) is greater than the critical value (=1.645), that means it lies in rejection region, so we reject the null hypothesis and support the alternative hypothesis i.e. support our claim at 5% level of significance.

Thus, we conclude that sample fails to provide us sufficient evidence against the claim so we may assume that the survival rate is greater than 80% in the population using that injection.

E10) Let X_1 and X_2 stand for number of literates in districts A and B respectively. Therefore, we are given that

$$n_1 = 100, X_1 = 50 \Rightarrow p_1 = \frac{X_1}{n_1} = \frac{50}{100} = 0.50$$

$$n_2 = 100, X_2 = 40 \Rightarrow p_2 = \frac{X_2}{n_2} = \frac{40}{100} = 0.40$$

Here, we want to test whether the proportion of literate persons in two districts A and B is same. If P_1 and P_2 denote the proportions of literate persons in two districts A and B respectively then our claim is $P_1 = P_2$ and its complement $P_1 \neq P_2$. Since the claim contains the equality sign so we can take the claim as the null hypothesis and complement as the alternative hypothesis. Thus,

$$H_0 : P_1 = P_2 = P \text{ and } H_1 : P_1 \neq P_2$$

Since the alternative hypothesis is two-tailed so the test is two tailed test.

The estimate of the combined proportion (P) of literates in districts A and B is given by

$$\hat{P} = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2} = \frac{X_1 + X_2}{n_1 + n_2} = \frac{50 + 40}{100 + 100} = 0.45$$

$$\hat{Q} = 1 - \hat{P} = 1 - 0.45 = 0.55$$

Before proceeding further, first we have to check whether the condition of normality meets or not.

$$\therefore \quad n_1p_1 = 100 \times 0.50 = 50 > 5,\ n_1q_1 = 100 \times 0.50 = 50 > 5$$

$$n_2p_2 = 100 \times 0.40 = 40 > 5,\ n_2q_2 = 100 \times 0.60 = 60 > 5$$

We see that condition of normality meets, so we can go for Z-test.

For testing the null hypothesis, the test statistic Z is given by

$$Z = \frac{p_1 - p_2}{\sqrt{\hat{P}\hat{Q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$= \frac{0.50 - 0.40}{\sqrt{0.45 \times 0.55\left(\frac{1}{100} + \frac{1}{100}\right)}} = \frac{0.10}{0.0704} = 1.42$$

The critical (tabulated) values for two-tailed test at 1% level of significance are $\pm z_{\alpha/2} = \pm z_{0.005} = \pm 2.58$.

Since calculated value of test statistic Z (= 1.42) is less than the critical value (= 2.58) and greater than critical value (= −2.58), that means it lies in non-rejection region, so we do not reject the null hypothesis i.e. we support our claim at 1% level of significance.

Thus, we conclude that samples fail to provide us sufficient evidence against the claim so we may assume that the proportion of literates in districts A and B is equal.

E11) Here, we are given that

$$n_1 = 1200, \quad p_1 = 30\% = 0.30$$

$$n_2 = 900, \quad p_2 = 20\% = 0.20$$

Here, we want to test that the proportion of blue-eye persons in both the population is same. If P_1 and P_2 denote the proportions of blue-eye persons in two populations respectively then our claim is $P_1 = P_2$ and its complement $P_1 \neq P_2$. Since the claim contains the equality sign so we can take the claim as the null hypothesis and complement as the alternative hypothesis. Thus,

$$H_0 : P_1 = P_2 = P \text{ and } H_1 : P_1 \neq P_2$$

Since the alternative hypothesis is two-tailed so the test is two tailed test.

The estimate of the combined proportion (P) of literates in districts A and B is given by

$$\hat{P} = \frac{n_1p_1 + n_2p_2}{n_1 + n_2} = \frac{1200 \times 0.30 + 900 \times 0.20}{1200 + 900} = 0.257$$

$$\hat{Q} = 1 - \hat{P} = 1 - 0.257 = 0.743$$

Before proceeding further, first we have to check whether the condition of normality meets or not.

$$\therefore \qquad n_1p_1 = 1200 \times 0.30 = 360 > 5,\ n_1q_1 = 1200 \times 0.70 = 840 > 5$$

$$n_2p_2 = 900 \times 0.20 = 180 > 5,\ n_2q_2 = 900 \times 0.80 = 720 > 5$$

We see that condition of normality meets, so we can go for Z-test.

For testing the null hypothesis, the test statistic Z is given by

$$Z = \frac{p_1 - p_2}{\sqrt{\hat{P}\hat{Q}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$= \frac{0.30 - 0.20}{\sqrt{0.257 \times 0.743\left(\frac{1}{1200} + \frac{1}{900}\right)}} = \frac{0.10}{0.019} = 5.26$$

The critical (tabulated) values for two-tailed test at 5% level of significance are $\pm z_{\alpha/2} = \pm z_{0.005} = \pm 1.96$.

Since calculated value of test statistic Z (= 5.26) is greater than critical values (= ±1.96), that means it lies in rejection region, so we reject the null hypothesis i.e. we reject our claim at 5% level of significance.

Thus, we conclude that samples provide us sufficient evidence against the claim so the proportion of blue-eyed person in both the population is not same.

E12) Here, we are given that

$$n = 120, \quad S = 7, \quad \sigma_0 = 6$$

Here, we want to test that standard deviation (σ) of the life of bulbs of the lot is 6 hours. So our claim is $\sigma = 6$ and its complement is $\sigma \neq 6$. Since the claim contains the equality sign so we can take the claim as the null hypothesis and complement as the alternative hypothesis. Thus,

$$H_0 : \sigma = \sigma_0 = 6 \text{ and } H_1 : \sigma \neq 6$$

Since the alternative hypothesis is two-tailed so the test is two-tailed test.

Here, the distribution of population under study is not known and sample size is large (n > 30) so we can go for Z-test.

For testing the null hypothesis, the test statistic Z is given by

$$Z = \frac{S^2 - \sigma_0^2}{\sigma_0^2\sqrt{\frac{2}{n}}} \sim N(0,1)$$

$$= \frac{(7)^2 - (6)^2}{(6)^2\sqrt{\frac{2}{120}}} = \frac{13}{36 \times 0.129} = \frac{13}{4.64} = 2.80$$

The critical values for two-tailed test at 5% level of significance are $\pm z_{\alpha/2} = \pm z_{0.025} = \pm 1.96$.

Since calculated value of Z (= 2.8) is greater than critical values (= ±1.96), that means it lies in rejection region, so we reject the null hypothesis i.e. we reject our claim at 5% level of significance.

Thus, we conclude that sample provides us sufficient evidence against the claim so standard deviation of the life of bulbs of the lot is not 6.0 hours.

E13) Here, we are given that

$$n_1 = 52, \quad S_1^2 = 25$$

$$n_2 = 40, \quad S_2^2 = 12$$

Here, we want to test that variance of source A significantly differs to the variances of source B. If σ_1^2 and σ_2^2 denote the variances in the raw materials of sources A and B respectively so our claim is $\sigma_1^2 \neq \sigma_2^2$ and its complement is $\sigma_1^2 = \sigma_2^2$. Since complement contains the equality sign so we can take the complement as the null hypothesis and the claim as the alternative hypothesis. Thus,

$$H_0 : \sigma_1^2 = \sigma_2^2 \text{ and } H_1 : \sigma_1^2 \neq \sigma_2^2$$

Since the alternative hypothesis is two-tailed so the test is two-tailed test.

Here, the distributions of populations under study are not known and sample sizes are large $(n_1 = 52 > 30,\ n_2 = 40 > 30)$ so we can go for Z-test.

Since population variances are unknown so for testing the null hypothesis, the test statistic Z is given by

$$Z = \frac{S_1^2 - S_2^2}{\sqrt{\left(\frac{2S_1^4}{n_1} + \frac{2S_2^4}{n_2}\right)}}$$

$$= \frac{25-12}{\sqrt{\frac{2(25)^2}{52} + \frac{2(12)^2}{40}}} \quad = \frac{13}{5.5} = 2.36$$

The critical values for two-tailed test at 5% level of significance are $\pm z_{\alpha/2} = \pm z_{0.025} = \pm 1.96$.

Since calculated value of Z (= 2.36) is greater than critical values (= ±1.96), that means it lies in rejection region, so we reject the null hypothesis and support the alternative hypothesis i.e. we support our claim at 5% level of significance.

Thus, we conclude that samples fail to provide us sufficient evidence against the claim so variance of source A significantly differs to the variance of source B.

UNIT 11 SMALL SAMPLE TESTS

Structure

11.1 INTRODUCTION

In previous unit, we have discussed the testing of hypothesis for large samples in details. Recall that throughout the unit, we were making an assumption that "if sample size is sufficiently large then test statistic follows approximately standard normal distribution". Also recall two points highlighted in this course, i.e.

- Cost of our study increases as sample size increases.
- Sometime nature of the units in the population under study is such that they destroyed under investigation.

If there are limited recourses in terms of money then first point listed above force us not to go for large sample size when items /units under study are very costly such as airplane, computer, etc. Second point listed above give an alarm for not to go for large sample if population units are destroyed under investigation.

So, we need an alternative technique which is used to test the hypothesis based on small sample(s). Small sample tests do this job for us. But in return they demand one basic assumption that population under study should be normal as you will see when you go through the unit. t, χ^2 and F -tests are some commonly used small sample tests.

In this unit, we will discuss t-test in details which is based on the t-distribution described in Unit 3 of this course. And χ^2 and F-tests will be discussed in next unit which are based on χ^2 and F-distributions described in Unit 3 and Unit 4 of this course respectively.

This unit is divided into eight sections. Section 11.1 is described the need of small sample tests. The general procedure of t-test for testing a hypothesis is described in Section 11.2. In Section 11.3, we discuss testing of hypothesis for population mean using t-test. Testing of hypothesis for difference of two population means when samples are independent is described in Section 11.4 whereas in Section 11.5, the paired t-test for difference of two population means when samples are dependent(paired) is discussed. In Section 11.6 testing of hypothesis for population correlation coefficient is explained. Unit ends by providing summary of what we have discussed in this unit in Section 11.7 and solution of exercises in Section 11.8.

Before moving further a humble suggestion to you that please revise what you have learned in previous two units. The concepts discussed there will help you a lot to better understand the concepts discussed in this unit.

Objectives

After studying this unit, you should be able to:

- realize the importance of small sample tests;
- know the procedure of t-test for testing a hypothesis;
- describe testing of hypothesis for population mean for using t-test;
- explain the testing of hypothesis for difference of two population means when samples are independent using t-test;
- describe the procedure for paired t-test for testing of hypothesis for difference of two population means when samples are dependent or paired; and
- explain the testing of hypothesis for population correlation coefficient using t-test.

11.2 GENERAL PROCEDURE OF t-TEST FOR TESTING A HYPOTHESIS

The general procedure of t-test for testing a hypothesis is similar as Z-test already explained in Unit 10. Let us give you similar details here.

For this purpose, let $X_1, X_2, \ldots, X_n$ be a random sample of **small size n (< 30)** selected from a **normal population** (recall the demand of small sample tests pointed out in previous Section 11.1) having parameter of interest, say, θ which is actually unknown but its hypothetical value, say, θ_0 estimated from some previous study or some other way is to be tested. t-test involves following steps for testing this hypothetical value:

Step I: First of all, we setup null and alternative hypotheses. Here, we want to test the hypothetical value θ_0 of parameter θ so we can take the null and alternative hypotheses as

$$H_0 : \theta = \theta_0 \text{ and } H_1 : \theta \neq \theta_0 \qquad [\text{for two-tailed test}]$$

$$\text{or} \quad \left.\begin{array}{l} H_0 : \theta \leq \theta_0 \text{ and } H_1 : \theta > \theta_0 \\ H_0 : \theta \geq \theta_0 \text{ and } H_1 : \theta < \theta_0 \end{array}\right\} \qquad [\text{for one-tailed test}]$$

In case of comparing same parameter of two populations of interest, say, θ_1, and θ_2 then our null and alternative hypotheses would be

$$H_0 : \theta_1 = \theta_2 \text{ and } H_1 : \theta_1 \neq \theta_2 \qquad [\text{for two-tailed test}]$$

$$\text{or} \quad \left.\begin{array}{l} H_0 : \theta_1 \leq \theta_2 \text{ and } H_1 : \theta_1 > \theta_2 \\ H_0 : \theta_1 \geq \theta_2 \text{ and } H_1 : \theta_1 < \theta_2 \end{array}\right\} \qquad [\text{for one-tailed test}]$$

Step II: After setting the null and alternative hypotheses our next step is to decide a criteria for rejection or non-rejection of null hypothesis i.e.

decide the level of significance α, at which we want to test our null hypothesis. We generally take α = 5 % or 1%.

Step III: The third step is to determine an appropriate test statistic, say, t for testing the null hypothesis. Suppose T_n is the sample statistic (may be sample mean, sample correlation coefficient, etc. depending upon θ) for the parameter θ then test-statistic t is given by

$$t = \frac{T_n - E(T_n)}{SE(T_n)}$$

Step IV: As we know, t-test is based on t-distribution and t-distribution is described with the help of its degrees of freedom, therefore, test statistic t follows t-distribution with specified degrees of freedom as the case may be.

By putting the values of T_n, $E(T_n)$ and $SE(T_n)$ in above formula, we calculate the value of test statistic t. Let t_{cal} be the calculated value of test statistic t after putting these values.

Step V: After that, we obtain the critical (cut-off or tabulated) value(s) in the sampling distribution of the test statistic t corresponding to α assumed in Step II. The critical values for t-test are given in **Table-II (t-table)** of the Appendix at the end of Block 1 of this course corresponding to different level of significance (α). After that, we construct rejection (critical) region of size α in the probability curve of the sampling distribution of test statistic t.

Step VI: Take the decision about the null hypothesis based on calculated and critical value(s) of test statistic obtained in Step IV and Step V respectively. Since critical value depends upon the nature of the test that it is one-tailed test or two-tailed test so following cases arise:

In case of one-tailed test:

Case I: When $H_0 : \theta \le \theta_0$ and $H_1 : \theta > \theta_0$ (right-tailed test)

In this case, the rejection (critical) region falls under the right tail of the probability curve of the sampling distribution of test statistic t. Suppose $t_{(\nu),\alpha}$ is the critical value at α level of significance then entire region greater than or equal to $t_{(\nu),\alpha}$ is the rejection region and less than $t_{(\nu),\alpha}$ is the non-rejection region as shown in Fig. 11.1.

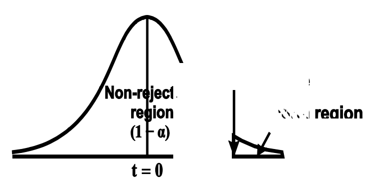

Fig. 11.1

If $t_{cal} \ge t_{(\nu,\alpha)}$, that means calculated value of test statistic t lies in the rejection (critical) region, then we reject the null hypothesis H_0 at α level of significance. Therefore, we conclude that sample data provides us sufficient evidence against the null hypothesis and there is a significant difference between hypothesized value and observed value of the parameter.

If $t_{cal} < t_{(\nu,\alpha)}$, that means calculated value of test statistic t lies in non-rejection region, then we do not reject the null hypothesis H_0 at α level of significance. Therefore, we conclude that the sample data fails to provide us sufficient evidence against the null hypothesis and the difference between hypothesized value and observed value of the parameter due to fluctuation of sample.

Case II: When $H_0 : \theta \geq \theta_0$ and $H_1 : \theta < \theta_0$ (left-tailed test)

In this case, the rejection (critical) region falls under the left tail of the probability curve of the sampling distribution of test statistic t. Suppose $-t_{(\nu),\alpha}$ is the critical value at α level of significance then entire region less than or equal to $-t_{(\nu),\alpha}$ is the rejection region and greater than $-t_{(\nu),\alpha}$ is the non-rejection region as shown in Fig. 11.2.

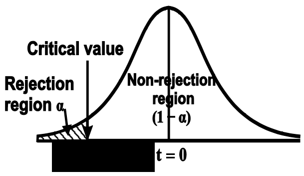

Fig. 11.2

If $t_{cal} \leq -t_{(\nu),\alpha}$, that means calculated value of test statistic t lies in the rejection (critical) region, then we reject the null hypothesis H_0 at α level of significance.

If $t_{cal} > -t_{(\nu),\alpha}$, that means calculated value of test statistic t lies in the non-rejection region, then we do not reject the null hypothesis H_0 at α level of significance.

In case of two-tailed test:

That is, when $H_0 : \theta = \theta_0$ and $H_1 : \theta \neq \theta_0$

In this case, the rejection region falls under both tails of the probability curve of sampling distribution of the test statistic t. Half the area (α) i.e. α/2 will lies under left tail and other half under the right tail. Suppose $-t_{(\nu),\alpha/2}$ and $t_{(\nu),\alpha/2}$ are the two critical values at the left- tailed and right-tailed respectively. Therefore, entire region less than or equal to $-t_{(\nu),\alpha/2}$ and greater than or equal to $t_{(\nu),\alpha/2}$ are the rejection regions and between $-t_{(\nu),\alpha/2}$ and $t_{(\nu),\alpha/2}$ is the non-rejection region as shown in Fig. 11.3.

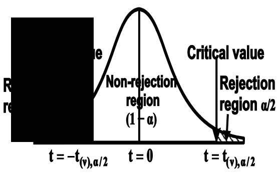

Fig. 11.3

If $t_{cal} \geq t_{(\nu),\alpha/2}$, or $t_{cal} \leq -t_{(\nu),\alpha/2}$, that means calculated value of test statistic t lies in the rejection(critical) region, then we reject the null hypothesis H_0 at α level of significance.

And if $-t_{(\nu),\alpha/2} < t_{cal} < t_{(\nu),\alpha/2}$, that means calculated value of test statistic t lies in the non-rejection region, then we do not reject the null hypothesis H_0 at α level of significance.

Procedure of taking the decision about the null hypothesis on the basis of p-value:

To take the decision about the null hypothesis on the basis of p-value, the p-value is compared with given level of significance (α). And if p-value is less than or equal to α then we reject the null hypothesis and if p-value is greater than α then we do not reject the null hypothesis at α level of significance.

Since the distribution of test statistic t follows t-distribution with ν df and we also know that t-distribution is symmetrical about t = 0 line therefore, if t_{cal} represents calculated value of test statistic t then p-value can be defined as:

For one-tailed test:

For H_1: $\theta > \theta_0$ (right-tailed test)

$$\text{p-value} = P[t \geq t_{cal}]$$

For H_1: $\theta < \theta_0$ (left-tailed test)

$$\text{p-value} = P\left[t \leq t_{cal}\right]$$

For two-tailed test: For H_1: $\theta \neq \theta_0$

$$\text{p-value} = 2P\left[t \ge |t_{cal}|\right]$$

These p-values for t-test can be obtained with the help of **Table-II (t-table)** given in the Appendix at the end of Block 1 of this course. But this table gives the t-values corresponding to the standard values of α such as 0.10, 0.05, 0.025, 0.01 and 0.005 only, therefore, the exact p-values are not obtained with the help of this table and we can approximate the p-value for this test.

For example, if test is right-tailed and calculated (observed) value of test statistic t is 2.94 with 9 df then p-value is obtained as:

Since calculated value of test statistic t is based on the 9 df therefore, we use row for 9 df in the t-table and move across this row to find the values in which calculated t-value falls. Since calculated t-value falls between 2.821 and 3.250, which are corresponding to the values of one-tailed area α = 0.01 and 0.005 respectively, therefore, p-value will lie between 0.005 and 0.01, that is,

$$0.005 < \text{p-value} < 0.01$$

If in the above example, the test is two-tailed then the two values 0.01 and 0.005 would be doubled for p-value, that is,

$$0.005 \times 2 = 0.01 < \text{p-value} < 0.02 = 2 \times 0.01$$

Note 1: With the help of computer packages and softwares such as SPSS, SAS, MINITAB, EXCEL, etc. we can find the exact p-values for t-test.

Now, you can try the following exercise.

E1) If test is two-tailed and calculated value of test statistic t is 2.42 with 15 df then find the p-value for t-test.

11.3 TESTING OF HYPOTHESIS FOR POPULATION MEAN USING t-TEST

In Section 10.3 of the previous unit, we have discussed Z-test for testing the hypothesis about population mean when population variance σ^2 is known and unknown.

Recall form these, we have already pointed out that one basic difference between Z-test and t-test is that Z-test is used when population SD is known whether sample size is large or small and t-test is used when population SD is unknown whether sample size is small or large. But in case of large sample size Z-test is an appropriate of t-test as we did in previous unit. But in practice standard deviation of population is not known and sample size is small so in this situation, we use t-test provided population under study is normal.

Assumptions

Virtually every test has some assumptions which must be met prior to the application of the test. This t-test needs following assumptions to work:

(i) The characteristic under study follows normal distribution. In other words, populations from which random sample is drawn should be normal with respect to the characteristic of interest.

(ii) Sample observations are random and independent.

(iii) Population variance σ^2 is unknown.

For describing this test, let $X_1, X_2, \ldots, X_n$ be a random sample of **small size n (< 30)** selected from a **normal population** with mean μ and unknown variance σ^2.

Now, follow the same procedure as we have discussed in previous section, that is, first of all we setup the null and alternative hypotheses. Here, we want to test the claim about the specified value μ_0 of population mean μ so we can take the null and alternative hypotheses as

$$H_0 : \mu = \mu_0 \text{ and } H_1 : \mu \neq \mu_0 \; [\text{for two-tailed test}] \; \left[\begin{array}{l}\text{Here, } \theta = \mu \text{ and } \theta_0 = \mu_0 \\ \text{if we compare it with} \\ \text{general procedure.}\end{array}\right]$$

or
$$\left.\begin{array}{l} H_0 : \mu \leq \mu_0 \text{ and } H_1 : \mu > \mu_0 \\ H_0 : \mu \geq \mu_0 \text{ and } H_1 : \mu < \mu_0 \end{array}\right\} [\text{for one-tailed test}]$$

For testing the null hypothesis, the test statistic t is given by

$$t = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \sim t_{(n-1)} \quad \text{under } H_0$$

where, $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ is the sample mean and $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2$ is the sample variance.

For computational simplicity, we may use the following formulae for $\bar{X}, S^2$:

$$\bar{X} = a + \frac{1}{n}\sum d \text{ and } S^2 = \frac{1}{n-1}\left[\sum d^2 - \frac{\left(\sum d\right)^2}{n}\right]$$

where, d = (X − a), 'a' being the assumed arbitrary value.

Here, the test statistic t follows t-distribution with (n − 1) degrees of freedom as we discussed in Unit 3 of this course.

After substituting values of $\bar{X}$, S and n , we get calculated value of test statistic t. Then we look for critical (or cut-off or tabulated) value(s) of test statistic t from the t-table. On comparing calculated value and critical value(s), we take the decision about the null hypothesis as discussed in previous section.

Let us do some examples of testing of hypothesis about population mean using t-test.

Example 1: A manufacturer claims that a special type of projector bulb has an average life 160 hours. To check this claim an investigator takes a sample of 20 such bulbs, puts on the test, and obtains an average life 167 hours with standard deviation 16 hours. Assuming that the life time of such bulbs follows normal distribution, does the investigator accept the manufacturer's claim at 5% level of significance?

Solution: Here, we are given that

$$\mu_0 = 160, \quad n = 20, \quad \bar{X} = 167 \quad \text{and} \quad S = 16$$

Here, we want to test the manufacturer claims that a special type of projector bulb has an average life (μ) 160 hours. So claim is $\mu = 160$ and its complement is $\mu \neq 160$. Since the claim contains the equality sign so we can take the claim as the null hypothesis and complement as the alternative hypothesis. Thus,

$$H_0 : \mu = \mu_0 = 160 \text{ and } H_1 : \mu \neq 160$$

Since the alternative hypothesis is two-tailed so the test is two-tailed test.

Here, we want to test the hypothesis regarding population mean when population SD is unknown. Also sample size is small n = 20(n < 30) and population under study is normal, so we can go for t-test for testing the hypothesis about population mean.

For testing the null hypothesis, the test statistic t is given by

$$t = \frac{\overline{X} - \mu_0}{S/\sqrt{n}}$$

$$= \frac{167 - 160}{16/\sqrt{20}} = \frac{7}{3.58} = 1.96$$

The critical value of the test statistic t for various df and different level of significance α are given in **Table II** of the Appendix at the end of the Block 1of this course.

The critical (tabulated) values of test statistic for two-tailed test corresponding (n-1) = 19 df at 5% level of significance are $\pm t_{(n-1),\alpha/2} = \pm t_{(19),0.025} = \pm 2.093$.

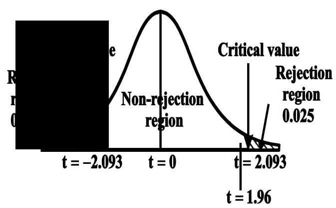

Fig. 11.4

Since calculated value of test statistic t (= 1.96) is greater than the critical value (= − 2.093) and is less than critical value (= 2.093), that means calculated value of test statistic lies in non-rejection region as shown in Fig. 11.4. So we do not reject the null hypothesis i.e. we support the manufacture's claim at 5% level of significance.

Decision according to p-value:

Since calculated value of test statistic t is based on 19 df therefore, we use row for 19 df in the t-table and move across this row to find the values in which calculated t-value falls. Since calculated t-value falls between 1.729 and 2.093 corresponding to one-tailed area α = 0.05 and 0.025 respectively therefore p-value lies between 0.025 and 0.05, that is,

$$0.025 < \text{p-value} < 0.05$$

Since test is two-tailed so

$$2 \times 0.025 = 0.05 < \text{p-value} < 0.10 = 0.05 \times 2$$

Since p-value is greater than $\alpha (= 0.05)$ so we do not reject the null hypothesis at 5% level of significance.

Thus, we conclude that sample fails to provide us sufficient evidence against the null hypothesis so we may assume that the manufacture's claim is true so the investigator may accept the manufacturer's claim at 5% level of significance.

Example 2: The mean share price of companies of Pharma sector is Rs.70. The share prices of all companies were changed time to time. After a month, a sample of 10 Pharma companies was taken and their share prices were noted as below:

70, 76, 75, 69, 70, 72, 68, 65, 75, 72

Assuming that the distribution of share prices follows normal distribution, test whether mean share price is still the same at 1% level of significance?

Solution: Here, we wish to test that the mean share price (μ) of companies of Pharma sector is still Rs.70 besides all changes. So our claim is $\mu = 70$ and its complement is $\mu \neq 70$. Since the claim contains the equality sign so we can take the claim as the null hypothesis and complement as the alternative hypothesis. Thus,

$$H_0 : \mu = \mu_0 = 70 \quad [\text{mean share price of companies is still Rs. 70}]$$

$$H_1 : \mu \neq \mu_0 = 70 \quad [\text{mean share price of companies is not still Rs. 70}]$$

Since the alternative hypothesis is two-tailed so the test is two-tailed test.

Here, we want to test the hypothesis regarding population mean when population SD is unknown. Also sample size is small n = 10(n < 30) and population under study is normal, so we can go for t-test for testing the hypothesis about population mean.

For testing the null hypothesis, the test statistic t is given by

$$t = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \sim t_{(n-1)} \qquad \ldots (1)$$

Calculation for $\bar{X}$ and S:

S. No.	**Sample value (X)**	**Deviation d = (X - a), a = 70**	**d^2**
1	70	0	0
2	76	6	36
3	75	5	25
4	69	-1	1
5	70	0	0
6	72	2	4
7	68	-2	4
8	65	-5	25
9	75	5	25
10	72	2	4
Total		12	124

The assumed value of a is 70.

From the above calculation, we have

$$\bar{X} = a + \frac{1}{n}\sum d = 70 + \frac{1}{10} \times 12 = 71.2$$

$$S^2 = \frac{1}{n-1}\left[\sum d^2 - \frac{\left(\sum d\right)^2}{n}\right]$$

$$= \frac{1}{10-1}\left[124 - \frac{(12)^2}{10}\right] = \frac{1}{9}\left[124 - \frac{144}{10}\right] = 12.18$$

$$\Rightarrow S = \sqrt{12.18} = 3.49$$

Putting the values in equation (1), we have

$$t = \frac{71.2 - 70}{3.49/\sqrt{10}} = \frac{1.2}{1.10} = 1.09$$

The critical (tabulated) values of test statistic for two-tailed test corresponding (n-1) = 9 df at 1% level of significance are $\pm t_{(n-1),\alpha/2} = \pm t_{(9),0.005} = \pm 3.250$.

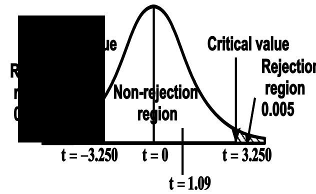

Fig. 11.5

Since calculated value of test statistic t (= 1.09) is less than the critical value (= 3.250) and greater than the critical value (= –3.250), that means calculated value of t lies in non-rejection region as shown in Fig. 11.5. So we do not reject the null hypothesis i.e. we support the claim at 1% level of significance.

Decision according to p-value:

Since calculated value of test statistic t is based on 9 df therefore, we use row for 9 df in the t-table and move across this row to find the values in which calculated t-value falls. Since all values in this row are greater than calculated t-value 1.09 and the smallest value is 1.383 corresponding to one-tailed area $\alpha = 0.10$ therefore p-value is greater than 0.10, that is,

$$\text{p-value} > 0.10$$

Since test is two-tailed so

$$\text{p-value} > 2 \times 0.10 = 0.20$$

Since p-value (= 0.20) is greater than $\alpha(= 0.01)$ so we do not reject the null hypothesis at 1% level of significance.

Thus, we conclude that the sample fails to provide us sufficient evidence against the claim so may assume that the mean share price is still Rs. 70.

Now, you can try the following exercises.

E2) A tyre manufacturer claims that the average life of a particular category of his tyre is 18000 km when used under normal driving conditions. A random sample of 16 tyres was tested. The mean and SD of life of the tyres in the sample were 20000 km and 6000 km respectively. Assuming that the life of the tyres is normally distributed, test the claim of the manufacturer at 1% level of significance using appropriate test.

E3) It is known that the average weight of cadets of a centre follows normal distribution. Weights of 10 randomly selected cadets from the same centre are as given below:

48, 50, 62, 75, 80, 60, 70, 56, 52, 77

Can we say that average weight of all cadets of the centre from which the above sample was taken is equal to 60 kg at 5% level of significance?

11.4 TESTING OF HYPOTHESIS FOR DIFFERENCE OF TWO POPULATION MEANS USING t-TEST

In Section 10.4 of the previous unit, we have discussed Z-test for testing the hypothesis about difference of two population means under different possibility of population variances σ_1^2 and σ_2^2. Recall from there, we have pointed out that one basic difference between Z-test and t-test is that, Z-test is used when standard deviations of both populations are known and t-test is used when standard deviations of both populations are unknown. But in practice standard deviations of both populations are not known, so in real life problems t-test is more suitable compared to Z-test.

Assumptions

This test works under following assumptions:

(i) The characteristic under study follows normal distribution in both the populations. In other words, both populations from which random samples are drawn should be normal with respect to the characteristic of interest.

(ii) Samples and their observations both are independent to each other.

(iii) Population variances σ_1^2 and σ_2^2 are both unknown but equal.

For describing this test, let there be two normal populations $N(\mu_1, \sigma_1^2)$ and $N(\mu_2, \sigma_2^2)$ under study. And we have to draw two independent random samples, say, $X_1, X_2, ..., X_{n_1}$ and $Y_1, Y_2, ..., Y_{n_2}$ of sizes n_1 and n_2 from these normal populations $N(\mu_1, \sigma_1^2)$ and $N(\mu_2, \sigma_2^2)$ respectively. Let $\bar{X}$ and $\bar{Y}$ be the means of first and second sample respectively. Further, suppose the variances of both the populations are unknown but are equal, i.e., $\sigma_1^2 = \sigma_2^2 = \sigma^2$ (say). In this case, σ^2 is estimated by value of pooled sample variance S_p^2 where,

$$S_p^2 = \frac{1}{n_1 + n_2 - 2}\left[(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2\right]$$

and

$$S_1^2 = \frac{1}{(n_1 - 1)}\sum_{i=1}^{n_1}\left(X_i - \bar{X}\right)^2, \quad S_2^2 = \frac{1}{(n_2 - 1)}\sum_{i=1}^{n_2}\left(Y_i - \bar{Y}\right)^2$$

This can also be written as

$$S_p^2 = \frac{1}{n_1 + n_2 - 2}\left[\sum_{i=1}^{n_1}\left(X_i - \bar{X}\right)^2 + \sum_{i=1}^{n_2}\left(Y_i - \bar{Y}\right)^2\right]$$

For computational simplicity, use the following formulae for $\bar{X}, \bar{Y}$ and S_p^2:

$$\bar{X} = a + \frac{1}{n_1}\sum d_1, \; \bar{Y} = b + \frac{1}{n_2}\sum d_2 \text{ and}$$

$$S_p^2 = \frac{1}{n_1 + n_2 - 2}\left[\left\{\sum d_1^2 - \frac{\left(\sum d_1\right)^2}{n_1}\right\} + \left\{\sum d_2^2 - \frac{\left(\sum d_2\right)^2}{n_2}\right\}\right]$$

where, $d_1 = (X - a)$ and $d_2 = (Y - b)$, 'a' and 'b' are the assumed arbitrary values.

Now, follow the same procedure as we have discussed in Section 11.2, that is, first of all we have to setup null and alternative hypotheses. Here, we want to test the hypothesis about the difference of two population means so we can take the null hypothesis as

$$H_0 : \mu_1 = \mu_2 \text{ (no difference in means)} \begin{bmatrix} \text{Here, } \theta_1 = \mu_1 \text{ and } \theta_2 = \mu_2 \\ \text{if we compare it with} \\ \text{general procedure.} \end{bmatrix}$$

or $\quad H_0 : \mu_1 - \mu_2 = 0$ (difference in two means is 0) and the alternative hypothesis as

$$H_1: \mu_1 \neq \mu_2 \qquad [\text{for two-tailed test}]$$

$$\left.\begin{array}{l} H_0: \mu_1 \leq \mu_2 \text{ and } H_1: \mu_1 > \mu_2 \\ H_0: \mu_1 \geq \mu_2 \text{ and } H_1: \mu_1 < \mu_2 \end{array}\right\} [\text{for one-tailed test}]$$

or

For testing the null hypothesis, the test statistic t is given by

$$t = \frac{\bar{X} - \bar{Y}}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t_{(n_1+n_2-2)} \quad \text{under } H_0$$

After substituting values of $\bar{X}$, $\bar{Y}$, S_p, n_1 and n_2, we get calculated value of test statistic t. Then we look for critical (or cut-off or tabulated) value(s) of test statistic t from the **t-table**. On comparing calculated value and critical value(s), we take the decision about the null hypothesis as discussed in Section 11.2.

Let us do some examples to become more user friendly with the test explained above.

Example 3: In a random sample of 10 pigs fed by diet A, the gain in weights (in pounds) in a certain period were

12, 8, 14, 16, 13, 12, 8, 14, 10, 9

In another random sample of 10 pigs fed by diet B, the gain in weights (in pounds) in the same period were

14, 13, 12, 15, 16, 14, 18, 17, 21, 15

Assuming that gain in the weights due to both foods follows normal distributions with equal variances, test whether diets A and B differ significantly regarding their effect on increase in weight at 5% level of significance.

Solution: Here, we can test that diets A and B differ significantly regarding their effect on increase in weight of pigs. If μ_1 and μ_2 denote the average gain in weights due to diet A and diet B respectively then our claim is $\mu_1 \neq \mu_2$ and its complement is $\mu_1 = \mu_2$. Since complement contains the equality sign so we can take the complement as the null hypothesis and the claim as the alternative hypothesis. Thus,

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

Since the alternative hypothesis is two-tailed so the test is two-tailed test.

Since it is given that the increase in the weight due to both foods follows normal distributions and population variances are equal and unknown. And other assumptions of t-test for testing a hypothesis about difference of two population means also meet. So we can go for this test.

For testing the null hypothesis, the test statistic t is given by

$$t = \frac{\bar{X} - \bar{Y}}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t_{(n_1+n_2-2)} \quad \text{under } H_0 \qquad \ldots (2)$$

Calculation for $\bar{X}$, $\bar{Y}$ and S_p :

Diet A			Diet B		
X	**$d_1 = (X-a)$ a =12**	**d_1^2**	**Y**	**$d_2 = (Y-b)$ b = 16**	**d_2^2**
12	0	0	14	−2	4
8	−4	16	13	−3	9
14	2	4	12	−4	16
16	4	16	15	−1	1
13	1	1	16	0	0
12	0	0	14	−2	4
8	−4	16	18	2	4
14	2	4	17	1	1
10	−2	4	21	5	25
9	−3	9	15	−1	1
Total	−4	70		−5	65

Here, a =12, b = 16 are assumed values.

From above calculations, we have

$$\bar{X} = a + \frac{1}{n_1}\sum d_1 = 12 + \frac{(-4)}{10} = 11.6,$$

$$\bar{Y} = b + \frac{1}{n_2}\sum d_2 = 16 + \frac{(-5)}{10} = 15.5$$

$$S_p^2 = \frac{1}{n_1 + n_2 - 2}\left[\left\{\sum d_1^2 - \frac{\left(\sum d_1\right)^2}{n_1}\right\} + \left\{\sum d_2^2 - \frac{\left(\sum d_2\right)^2}{n_2}\right\}\right]$$

$$= \frac{1}{10+10-2}\left[\left\{70 - \frac{(-4)^2}{10}\right\} + \left\{65 - \frac{(-5)^2}{10}\right\}\right]$$

$$= \frac{1}{18}(68.4 + 62.5) = 7.27$$

$$\Rightarrow S_p = \sqrt{7.27} = 2.70$$

Putting the values in equation (2), we have

$$t = \frac{11.6 - 15.5}{2.70\sqrt{\frac{1}{10} + \frac{1}{10}}}$$

$$= \frac{-3.90}{2.70 \times 0.45} = \frac{-3.90}{1.215} = -3.21$$

The critical values of test statistic t for two-tailed test corresponding $(n_1 + n_2 - 2) = 18$ df at 5% level of significance are $\pm t_{(n_1+n_2-2),\alpha/2} = \pm t_{(18),0.025} = \pm 2.101$.

Since calculated value of test statistic t (= −3.21) is less than critical values (± 2.101) that means calculated value of test statistic t lies in rejection region, so we reject the null hypothesis and support the alternative hypothesis i.e. support our claim at 5% level of significance.

Thus, we conclude that samples do not provide us sufficient evidence against the claim so diets A and B differ significantly in terms of gain in weights of pigs.

Example 4: The means of two random samples of sizes 10 and 8 drawn from two normal populations are 210.40 and 208.92 respectively. The sum of squares of the deviations from their means is 26.94 and 24.50 respectively. Assuming that the populations are normal with equal variances, can samples be considered to have been drawn from normal populations having equal mean.

Solution: In usual notations, we are given that

$$n_1 = 10,\ \ n_2 = 8,\ \ \bar{X} = 210.40,\ \ \bar{Y} = 208.92,$$

$$\sum\left(X - \bar{X}\right)^2 = 26.94, \quad \sum\left(Y - \bar{Y}\right)^2 = 24.50$$

Therefore,

$$S_p^2 = \frac{1}{n_1 + n_2 - 2}\left[\sum\left(X - \bar{X}\right)^2 + \sum\left(Y - \bar{Y}\right)^2\right]$$

$$= \frac{1}{10+8-2}\left[26.94 + 24.50\right] = \frac{1}{16} \times 51.44 = 3.215$$

$$\Rightarrow S_p = \sqrt{3.215} = 1.79$$

We wish to test that both the samples are drawn from normal populations having the same means. If μ_1 and μ_2 denote the means of both normal populations respectively then our claim is $\mu_1 = \mu_2$ and its complement is $\mu_1 \neq \mu_2$. Since the claim contains the equality sign so we can take the claim as the null hypothesis and complement as the alternative hypothesis. Thus,

$H_0 : \mu_1 = \mu_2$ [mean of both populations is equal]

$H_1 : \mu_1 \neq \mu_2$ [mean of both populations is not equal]

Since the alternative hypothesis is two-tailed so the test is two-tailed test.

Since it is given that two populations are normal with equal and unknown variances and other assumptions of t-test for testing a hypothesis about difference of two population means also meet. So we can go for this test.

For testing the null hypothesis, the test statistic t is given by

$$t = \frac{\bar{X} - \bar{Y}}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t_{(n_1+n_2-2)} \quad \text{under } H_0$$

$$= \frac{210.40 - 208.92}{1.79\sqrt{\frac{1}{10} + \frac{1}{8}}} = \frac{1.48}{1.79 \times 0.47} = \frac{1.48}{0.84} = 1.76$$

The critical values of test statistic t for two-tailed test corresponding $(n_1 + n_2 - 2) = 16$ df at 5% level of significance are $\pm t_{(n_1+n_2-2),\alpha/2} = \pm t_{(16),0.025} = \pm 2.12$.

Since calculated value of test statistic t (= 1.76) is less than the critical value (= 2.12) and greater than the critical value (= –2.12), that means calculated value of test statistic t lies in non-rejection region so we do not reject the null hypothesis i.e. we support the claim.

Thus, we conclude that samples fail to provide us sufficient evidence against the claim so we may assume that both samples are taken from normal populations having equal means.

Now, you can try the following exercises.

E4) Two different types of drugs A and B were tried on some patients for increasing their weights. Six persons were given drug A and other 7 persons were given drug B. The gain in weights (in ponds) is given below:

Drug A	5	8	7	10	9	6	−
Drug B	9	10	15	12	14	8	12

Assuming that increment in the weights due to both drugs follows normal distributions with equal variances, do the both drugs differ significantly with regard to their mean weights increment at 5% level of significance?

E5) To test the effect of fertilizer on wheat production, 26 plots of land with equal areas were chosen. Half of these plots were treated with fertilizer and the other half were untreated. Other conditions were the same. The mean yield of wheat on the untreated plots was 4.6 quintals with a standard deviation of 0.5 quintals, while the mean yield of the treated plots was 5.0 quintals with standard deviations of 0.3 quintals. Assuming that yields of wheat with and without fertilizer follow normal distributions with equal variances, can we conclude that there is significant improvement in wheat production due to effect of fertilizer at 1% level of significance?

11.5 PAIRED t-TEST

In the previous section, we have discussed t-test for equality of two population means in case of independent samples. However, there are so many situations where two samples are not independent and observations are recorded on the same individuals or items. Generally, such types of observations are recorded to assess the effectiveness of a particular training, diet, treatment, medicine, etc. In such situations, the observations are recorded **“before and after”** the insertion of training, treatment, etc. as the case may be. For example, if we wish to test a new diet on, say, 15 individuals then the weight of the individuals recorded before diet and after the diet will form two different samples in which observations will be paired as per each individual. Similarly, in the test of blood-sugar in human body, fasting sugar level before meal and sugar level after meal, both are recorded for a patient as paired observations, etc. The parametric test designed for this type of situation is known as paired t-test.

Now, come to the working principle of this test. This test first of all converts the two populations into a single population by taking the difference of paired observations.

Now, instead of two populations, we are left with one population, the population of differences. And the problem of testing equality of two population mean reduces to test the hypothesis that mean of the population of differences is equal to zero.

Assumptions

This test works under following assumptions:

(i) The population of differences follows normal distribution.

(ii) Samples are not independent.

(iii) Size of both the samples is equal.

(iv) Population variances are unknown but not necessarily equal.

Let $(X_1, Y_1), (X_2, Y_2), \ldots,(X_n, Y_n)$ be a paired random sample of size n and the difference between paired observations X_i & Y_i be denoted by D_i, that is,

$$D_i = X_i - Y_i \quad \text{for all } i = 1, 2, \ldots, n$$

Hence, we can assume that $D_1, D_2, \ldots, D_n$ be a random sample from normal population of differences with mean μ_D and unknown variance σ_D^2. This is same as the case of testing of hypothesis for population mean when population variance is unknown which is described in Section 11.3 of this unit.

Here, we want to test that there is an effect of a diet, training, treatment, medicine, etc. So we can take the null hypothesis as

$$H_0: \mu_1 = \mu_2 \text{ or } H_0: \mu_D = \mu_1 - \mu_2 = 0$$

and the alternative hypothesis

$$H_1: \mu_1 \neq \mu_2 \text{ or } H_1: \mu_D \neq 0 \qquad [\text{for two-tailed test}]$$

or

$$\left.\begin{array}{l} H_0: \mu_1 \leq \mu_2 \text{ and } H_1: \mu_1 > \mu_2 \\ H_0: \mu_1 \geq \mu_2 \text{ and } H_1: \mu_1 < \mu_2 \end{array}\right\} \quad [\text{for one-tailed test}]$$

For testing the null hypothesis, the test statistic t is given by

$$t = \frac{\bar{D}}{S_D/\sqrt{n}} \sim t_{(n-1)} \qquad \text{under } H_0$$

where, $\bar{D} = \frac{1}{n}\sum_{i=1}^{n} D_i$ and $S_D^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(D_i - \bar{D}\right)^2 = \frac{1}{n-1}\left[\sum_{i=1}^{n} D^2 - \frac{\left(\sum_{i=1}^{n} D\right)^2}{n}\right]$

After substituting values of $\bar{D}$, S_D and n we get calculated value of test statistic t. Then we look for critical (or cut-off or tabulated) value(s) of test statistic t from the **t-table**. On comparing calculated value and critical value(s), we take the decision about the null hypothesis as discussed in Section 11.2.

Let us do some examples to become more user friendly with paired t-test.

Example 5: A group of 12 children was tested to find out how many digits they would repeat from memory after hearing them once. They were given practice session for this test. Next week they were retested. The results obtained were as follows:

Child Number	1	2	3	4	5	6	7	8	9	10	11	12
Recall Before	6	4	5	7	6	4	3	7	8	4	6	5
Recall After	6	6	4	7	6	5	5	9	9	7	8	7

Assuming that the memories of the children before and after the practice session follow normal distributions, is the memory practice session improve the performance of children?

Solution: Here, we want to test that memory practice session improve the performance of children. If μ_1 and μ_2 denote the mean digit repetition before and after the practice so our claim is $\mu_1 < \mu_2$ and its complement is $\mu_1 \geq \mu_2$. Since complement contains the equality sign so we can take the complement as the null hypothesis and the claim as the alternative hypothesis. Thus,

$$H_0 : \mu_1 \geq \mu_2 \text{ and } H_1 : \mu_1 < \mu_2$$

Since the alternative hypothesis is left-tailed so the test is left-tailed test.

It is a situation of before and after. Also, it is given that the memories of the children before and after the practice session follow normal distributions. So, population of differences will also be normal. Also all the assumptions of paired t-test meet so we can go for paired t-test.

For testing the null hypothesis, the test statistic t is given by

$$t = \frac{\bar{D}}{S_D / \sqrt{n}} \sim t_{(n-1)} \text{ under } H_0 \qquad \ldots (3)$$

where, $\bar{D}$ and S_D are mean and standard deviation of the population of differences.

Calculation for $\bar{D}$ and S_D:

Child Number	Digit recall		D = (X−Y)	D^2
	Before (X)	After (Y)		
1	6	6	0	0
2	4	6	−2	4
3	5	4	1	1
4	7	7	0	0
5	6	6	0	0
6	4	5	−1	1
7	3	5	−2	4
8	7	9	−2	4
9	8	9	−1	1
10	4	7	−3	9
11	6	8	−2	4
12	5	7	−2	4
			$\sum D = -14$	$\sum D^2 = 32$

From above calculations, we have

$$\bar{D} = \frac{1}{n}\sum D = \frac{1}{12}(-14) = -1.17$$

$$S_D^2 = \frac{1}{n-1}\left\{\sum D^2 - \frac{\left(\sum D\right)^2}{n}\right\}$$

$$= \frac{1}{11}\left[32 - \frac{(-14)^2}{12}\right] = \frac{1}{11} \times 15.67 = 1.42$$

$$\Rightarrow \quad S_D = \sqrt{1.42} = 1.19$$

Substituting these values in equation (3), we have

$$t = \frac{-1.17}{1.19/\sqrt{12}} = \frac{-1.17}{0.34} = -3.44$$

The critical value of test statistic t for left-tailed test corresponding (n-1) = 11 df at 5% level of significance is $-t_{(n-1),\alpha} = -t_{(11),0.05} = -1.796$.

Since calculated value of test statistic t (= −3.44) is less than the critical value (=−1.796), that means calculated value of t lies in rejection region, so we reject the null hypothesis and support the alternative hypothesis i.e. support the claim at 5% level of significance.

Thus, we conclude that samples fail to provide us sufficient evidence against the claim so we may assume that memory practice session improves the performance of children.

Example 6: Ten students were given a test in Statistics and after one month's coaching they were again given a test of the similar nature and the increase in their marks in the second test over the first are shown below:

Roll No.	1	2	3	4	5	6	7	8	9	10
Increase in Marks	6	−2	8	−4	10	2	5	−4	6	0

Assuming that increment in marks follows normal distribution. Do the data indicate that students have gained knowledge from the coaching at 1% level of significance?

Solution: Here, we want to test that students have gained knowledge from the coaching. If μ_D denotes the average increment in the marks due to one month's coaching then our claim is $\mu_D < 0$ but here we are given increment $\mathbf{D_i = (Y_i - X_i)}$ instead of $\mathbf{D_i = (X_i - Y_i)}$ so we take our claim is $\mu_D > 0$ and its complement is $\mu_D \le 0$. Since complement contains the equality sign so we can take the complement as the null hypothesis and the claim as the alternative hypothesis. Thus,

$$H_0 : \mu_D \le 0 \text{ and } H_1 : \mu_D > 0$$

Since the alternative hypothesis is right-tailed so the test is right-tailed test.

It is given that increment in marks after one month coaching follows normal distribution and population variance is unknown. Also participants are same in both situations before and after the coaching. And all the assumption of paired t-test meet so we can go for paired t-test.

For testing H_0, the test statistic is given by

$$t = \frac{\bar{D}}{S_D/\sqrt{n}} \sim t_{(n-1)} \text{ under } H_0 \qquad \ldots (4)$$

Calculation for $\bar{D}$ and S_D:

Roll No.	1	2	3	4	5	6	7	8	9	10	Total
D	6	−2	8	−4	10	2	5	−4	6	0	$\sum D = 27$
D^2	36	4	64	16	100	4	25	16	36	0	$\sum D^2 = 301$

From above calculations, we have

$$\bar{D} = \frac{1}{n}\sum D = \frac{1}{10} \times 27 = 2.7$$

$$S_D^2 = \frac{1}{n-1}\left\{\sum D^2 - \frac{\left(\sum D\right)^2}{n}\right\}$$

$$= \frac{1}{10-1}\left[301 - \frac{(27)^2}{10}\right] = \frac{1}{9} \times 228.1 = 25.34$$

$$\Rightarrow \quad S_D = \sqrt{25.34} = 5.03$$

Substituting these values in equation (4), we have

$$t = \frac{2.7}{5.03/\sqrt{10}} = \frac{2.7}{1.59} = 1.70$$

The critical value of test statistic t for right-tailed test corresponding (n - 1) = 9 df at 1% level of significance is $t_{(n-1),\,\alpha} = t_{(9),\,0.01} = 2.821$.

Since calculated value of test statistic t (= 1.70) is less than the critical value (= 2.821), that means calculated value of test statistic t lies in non-rejection region, so we do not reject null hypothesis and reject the alternative hypothesis i.e. we reject our claim at 1% level of significance.

Thus, we conclude that sample provides us sufficient evidence against the claim so students are not gained knowledge from the coaching.

Now, you can try the following exercises.

E6) To verify whether the programme "Post Graduate Diploma in Applied Statistics (PGDAST)" improved performance of the graduate students in Statistics, a similar test was given to10 participants both before and after the programme. The original marks out of 100 (before course) recorded in an alphabetical order of the participants are 42, 46, 50, 36, 44, 60, 62, 43, 70 and 53. After the course the marks in the same order are 45, 46, 60, 42, 60, 72, 63, 43, 80 and 65. Assuming that marks of the students before and after the course follow normal distribution. Test whether the programme PGDAST has improved the performance of the graduate students in Statistics at 5% level of significance?

E7) A drug is given to 8 patients and the increments in their blood pressure are recorded to be 4, 0, 7, −2, 0, −3, 2, 0. Assume that increment in their blood pressure follows normal distribution. Is it reasonable to believe that the drug has no effect on the change of blood pressure at 5% level of significance?

11.6 TESTING OF HYPOTHESIS FOR POPULATION CORRELATION COEFFICIENT USING t-TEST

In Unit 6 of MST-002, we have discussed the concept of correlation. Where, we studied that if two variables are related in such a way that change in the value of one variable affects the value of another variable then the variables are said to be correlated or there is a correlation between these two variables. Correlation can be positive, which means the variables move together in the same direction, or negative, which means they move in opposite directions. And correlation coefficient is used to measure the intensity or degree of linear relationship between two variables. The value of correlation coefficient varies between −1 and +1, where −1 representing a perfect negative correlation, 0 representing no correlation, and +1 representing a perfect positive correlation.

Sometime, the sample data indicate for non-zero correlation but in population they are uncorrelated ($\rho = 0$).

For example, price of tomato in Delhi (X) and in London (Y) are not correlated in population ($\rho = 0$). But paired sample data of 20 days of prices of tomato at both places may show correlation coefficient (r) $\neq 0$. In general, in sample data $r \neq 0$ does not ensure in population $\rho \neq 0$ holds.

In this section, we will know how we test the hypothesis that population correlation coefficient is zero.

Assumptions

This test works under following assumptions:

(i) The characteristic under study follows normal distribution in both the populations. In other words, both populations from which random samples are drawn should be normal with respect to the characteristic of interest.

(ii) Samples observations are random.

Let us consider a random sample $(X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n)$ of size n taken from a bivariate normal population. Let ρ and r be the correlation coefficients of population and sample data respectively.

Here, we wish to test the hypothesis about population correlation coefficient (ρ), that is, linear correlation between two variables X and Y in the population, so we can take the null hypothesis as

$$H_0 : \rho = 0 \text{ and } H_1 : \rho \neq 0 \ [\text{for two-tailed test}] \quad \left[\begin{array}{l}\text{Here, } \theta = \rho \text{ and } \theta_0 = 0 \text{ if} \\ \text{we compare it with general} \\ \text{procedure given in Section 11.2.}\end{array}\right]$$

or
$$\left.\begin{array}{l} H_0 : \rho \leq 0 \text{ and } H_1 : \rho > 0 \\ H_0 : \rho \geq 0 \text{ and } H_1 : \rho < 0 \end{array}\right\} \ [\text{for one-tailed test}]$$

For testing the null hypothesis, the test statistic t is given by

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

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

After substituting values of r and n, we find out calculated value of test statistic t. Then we look for critical (or cut-off or tabulated) value(s) of test statistic t from the **t-table**. On comparing calculated value and critical value(s), we take the decision about the null hypothesis as discussed in Section 11.2.

Let us do some examples of testing of hypothesis that population correlation coefficient is zero.

Example 7: A random sample of 18 pairs of observations from a normal population gave a correlation coefficient of 0.7. Test whether the population correlation coefficient is zero at 5% level of significance.

Solution: Given that

$n = 18, r = 0.7$

Here, we wish to test that population correlation coefficient (ρ) is zero so our claim is $\rho = 0$ and its complement is $\rho \neq 0$. Since the claim contains the equality sign so we can take the claim as the null hypothesis and complement as the alternative hypothesis. Thus,

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

Since the alternative hypothesis is two-tailed so the test is two-tailed test.

Here, we want to test the hypothesis regarding population correlation coefficient is zero and the populations under study follow normal distributions, so we can go for t-test.

For testing the null hypothesis, the test statistic t is given by

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

$$= \frac{0.7\sqrt{18-2}}{\sqrt{1-(0.7)^2}} = \frac{0.7 \times 4}{\sqrt{0.51}} = \frac{2.8}{0.71} = 3.94$$

The critical value of test statistic t for two-tailed test corresponding (n-2) = 16 df at 5% level of significance are $\pm t_{(n-2),\alpha/2} = \pm t_{(16),0.025} = \pm 2.120$.

Since calculated value of test statistic t (= 3.94) is greater than the critical values (= ± 2.120), that means calculated value of test statistic t lies in rejection region, so we reject the null hypothesis. i.e. we reject the claim at 5% level of significance.

Thus, we conclude that sample provides us sufficient evidence against the claim so there exists a relationship between two variables.

Example 8: A random sample of 15 married couples was taken from a population consisting of married couples between the ages of 30 and 40. The correlation coefficient between the IQs of husbands and wives was found to be 0.68. Assuming that the IQs of husbands and wives follow normal distributions then test that IQs of husbands and wives in the population are positively correlated at1% level of significance.

Solution: Given that

$$n = 15, r = 0.68$$

Here, we wish to test that IQs of husbands and wives in the population are positively correlated. If ρ denote the correlation coefficient between IQs of husbands and wives in the population then the claim is $\rho > 0$ and its complement is $\rho \leq 0$. Since complement contains the equality sign so we can take the complement as the null hypothesis and the claim as the alternative hypothesis. Thus,

$$H_0 : \rho \leq 0 \text{ and } H_1 : \rho > 0$$

Since the alternative hypothesis is right-tailed so the test is right-tailed test.

Here, we want to test the hypothesis regarding population correlation coefficient is zero and the populations under study follow normal distributions, so we can go for t-test.

For testing the null hypothesis, the test statistic t is given by

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

$$= \frac{0.68\sqrt{15-2}}{\sqrt{1-(0.68)^2}} = \frac{0.68 \times 3.61}{0.73} = 3.36$$

The critical value of test statistic t for right-tailed test corresponding (n-2) = 13 df at 1% level of significance is $t_{(n-2),\alpha} = t_{(13),0.01} = 2.650$.

Since calculated value of test statistic t (= 3.36) is greater than the critical value (= 2.650), that means calculated value of test statistic t lies in rejection region, so we reject the null hypothesis and support the alternative hypothesis i.e. we support our claim at 1% level of significance.

Thus, we conclude that sample fail to provide us sufficient evidence against the claim so we may assume that the correlation between IQs of husbands and wives in the population is positive.

In the same way, you can try the following exercise.

E8) Twenty families were selected randomly from a colony to determine that correlation exists between family income and the amount of money spent per family member on food each month. The sample correlation coefficient was computed as r = 0.40. Assuming that the family income and the amount of money spent per family member on food each month follow normal distributions then test that there is a positive linear relationship between the family income and the amounts of money spent per family member on food each month in colony at 1% level of significance.

We now end this unit by giving a summary of what we have covered in it.

11.7 SUMMARY

In this unit, we have discussed the following points:

1. Need of small sample tests.
2. Procedure of testing a hypothesis for t-test.
3. Testing of hypothesis for population mean using t-test.
4. Testing of hypothesis for difference of two population means when samples are independent using t-test.
5. The procedure of paired t-test for testing of hypothesis for difference of two population means when samples are dependent or paired.
6. Testing of hypothesis for population correlation coefficient using t-test.

11.8 SOLUTIONS /ANSWERS

E1) Since calculated value of test statistic t is based on 15 df therefore, we use row for 15 df in the **t-table** and move across this row to find the values in which calculated t-value lies. Since calculated t-value falls between 2.131 and 2.602, which are corresponding to the values of one-tailed area α = 0.025 and 0.01 respectively, therefore p-value will lie between 0.01 and 0.025, that is, $0.01 < \text{p-value} < 0.025$

Since test is two-tailed, therefore, the values are doubled, so

$$0.02 = 2 \times 0.01 < \text{p-value} < 2 \times 0.025 = 0.05$$

E2) Here, we are given that

$$n = 16,\ \mu_0 = 18000,\ \bar{X} = 20000,\ S = 6000$$

Here, we want to test that manufacturer's claim is true that the average life (μ) of tyres is 18000 km. So claim is $\mu = 18000$ and its complement is $\mu \neq 18000$. Since the claim contains the equality sign so we can take the claim as the null hypothesis and complement as the alternative hypothesis. Thus,

$H_0: \mu = \mu_0 = 18000$ [average life of tyres is 18000 km]

$H_1: \mu \neq 18000$ [average life of tyres is not 18000 km]

Here, population SD is unknown and population under study is given to be normal. So we can go for t-test.

For testing the null hypothesis, the test statistic t is given by

$$t = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} = \frac{20000 - 18000}{6000/\sqrt{16}} = \frac{2000}{1500} = 1.33$$

The critical value of test statistic t for two-tailed test corresponding (n-1) = 15 df at 1% level of significance are $\pm\, t_{(15),0.005} = \pm\, 2.947$.

Since calculated value of test statistic t (= 1.33) is less than the critical (tabulated) value (= 2.947) and greater that critical value (= − 2.947), that means calculated value of test statistic lies in non-rejection region, so we do not reject the null hypothesis i.e. we support the manufacture's claim at 1% level of significance.

Thus, we conclude that sample fails to provide sufficient evidence against the claim so we may assume that manufacturer's claim is true.

E3) Here, we want to test that the average weight (μ) of all cadets of the centre is 60 kg. So our claim is $\mu = 60$ and its complement is $\mu \neq 60$. Since the claim contains the equality sign so we can take the claim as the null hypothesis and complement as the alternative hypothesis. Thus,

$H_0: \mu = \mu_0 = 60$ [average weight of all cadets is 60 kg]

$H_1: \mu \neq 60$ [average weight of all cadets is not 60 kg]

Since the alternative hypothesis is two-tailed so the test is two-tailed test.

Here, population SD is unknown and population under study is given to be normal. So we can go for t-test.

For testing the null hypothesis, the test statistic t is given by

$$t = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \sim t_{(n-1)} \text{ under } H_0 \qquad \ldots (5)$$

Before moving further, first we have to calculate value of $\bar{X}$ and S.

Calculation for $\bar{X}$ and S:

Sample value (X)	$(x - \bar{x})$	$(x - \bar{x})^2$
48	−15	225
50	−13	169
62	−1	1
75	12	144
80	17	289
60	−3	9
70	7	49

56	−7	49
52	−11	121
77	14	196
$\sum X = 630$		$\sum\left(X-\bar{X}\right)^2 = 1252$

From the above calculation, we have

$$\bar{X} = \frac{1}{n}\sum X = \frac{1}{10}\times 630 = 63$$

$$S^2 = \frac{1}{n-1}\sum\left(X-\bar{X}\right)^2$$

$$= \frac{1}{10-1}\times 1252 = 139.11$$

$$\Rightarrow S = \sqrt{139.11} = 11.79$$

Putting the values in equation (5), we have

$$t = \frac{63-60}{11.79/\sqrt{10}}$$

$$= \frac{3}{3.73} = 0.80$$

The critical values of test statistic t for two-tailed test corresponding (n - 1) = 9 df at 5% level of significance are $\pm\, t_{(9),0.025} = \pm\, 2.262$.

Since calculated value of test statistic t (= 0.80) is less than the critical value (= 2.262) and greater than the critical value (= −2.262), that means calculated value of test statistic t lies in non-rejection region so we do not reject H_0 i.e. we support the claim at 5% level of significance.

Thus, we conclude that sample fails to provide sufficient evidence against the claim so we may assume that the average weight of all the cadets of given centre is 60 kg.

E4) Here, we want to test that there is no difference between drugs A and B with regard to their mean weight increment. If μ_1 and μ_2 denote the mean weight increment due to drug A and drug B respectively then our claim is $\mu_1 = \mu_2$ and its complement is $\mu_1 \neq \mu_2$. Since the claim contains the equality sign so we can take the claim as the null hypothesis and complement as the alternative hypothesis. Thus,

$$H_0 : \mu_1 = \mu_2 \quad [\text{effect of both drugs is same}]$$

$$H_1 : \mu_1 \neq \mu_2 \quad [\text{effect of both drugs is not same}]$$

Since the alternative hypothesis is two-tailed so the test is two-tailed test.

Since it is given that increments in the weight due to both drugs follow normal distributions with equal and unknown variances and other assumptions of t-test for testing a hypothesis about difference of two population means also meet. So we can go for this test.

For testing the null hypothesis, the test statistic t is given by

$$t = \frac{\bar{X} - \bar{Y}}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t_{(n_1+n_2-2)} \quad \text{under } H_0 \qquad \dots (6)$$

Assume, a = 8, b = 12 and use short-cut method to find $\bar{X}$, $\bar{Y}$ and S_p.

Calculation for $\bar{X}$, $\bar{Y}$ and S_p :

Drug A			**Drug B**		
X	**d_1 = (X-a) a = 8**	**d_1^2**	**Y**	**d_2 = (Y-b) b = 12**	**d_2^2**
5	–3	9	9	–3	9
8	0	0	10	–2	4
7	–1	1	15	3	9
10	2	4	12	0	0
9	1	1	14	2	4
6	–2	4	8	–4	16
			12	0	0
	$\sum d_1 = -3$	$\sum d_1^2 = 19$		$\sum d_2 = -4$	$\sum d_2^2 = 42$

From above calculation, we have

$$\bar{X} = a + \frac{1}{n_1}\sum d_1 = 8 + \frac{1}{6}(-3) = 7.5,$$

$$\bar{Y} = b + \frac{1}{n_2}\sum d_2 = 12 + \frac{1}{7}(-4) = 11.43$$

$$S_p^2 = \frac{1}{n_1 + n_2 - 2}\left[\left\{\sum d_1^2 - \frac{\left(\sum d_1\right)^2}{n_1}\right\} + \left\{\sum d_2^2 - \frac{\left(\sum d_2\right)^2}{n_2}\right\}\right]$$

$$= \frac{1}{6+7-2}\left[\left\{19 - \frac{(-3)^2}{6}\right\} + \left\{42 - \frac{(-4)^2}{7}\right\}\right]$$

$$= \frac{1}{11}(17.5 + 39.71) = 5.20$$

$$\Rightarrow S_p = \sqrt{5.20} = 2.28$$

Putting these values in equation (6), we have

$$t = \frac{7.5 - 11.43}{2.28\sqrt{\frac{1}{6} + \frac{1}{7}}} = \frac{-3.93}{2.28 \times 0.56} = \frac{-3.93}{1.28} = -3.07$$

The critical values of test statistic t for two-tailed test corresponding $(n_1 + n_2 - 2)$ = 11 df at 5% level of significance are $\pm\, t_{(11),0.025} = \pm\, 2.201$.

Since calculated value of test statistic t (= −3.07) is less than the critical values (= ± 2.201) that means calculated value of test statistic t lies in rejection region, so we reject the null hypothesis i.e. we reject the claim at 5% level of significance.

Thus, we conclude that samples provide us sufficient evidence against the claim so drugs A and B differ significantly. Any one of them is better than other.

E5) Here, we are given that

$$n_1 = 13, \quad \bar{X} = 4.6, \quad S_1 = 0.5,$$

$$n_2 = 13, \quad \bar{Y} = 5.0, \quad S_2 = 0.3$$

Therefore, the pooled variance S_p^2 can be calculated as

$$S_p^2 = \frac{1}{n_1 + n_2 - 2}\left[(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2\right]$$

$$= \frac{1}{13 + 13 - 2}\left[12 \times (0.5)^2 + 12 \times (0.3)^2\right]$$

$$= \frac{1}{24}(3.00 + 1.08) = 0.17$$

$$\Rightarrow S_p = \sqrt{0.17} = 0.41$$

We want to test that the there is significant improvement in wheat production due to fertilizer. If μ_1 and μ_2 denote the mean wheat productions without and with the fertilizer respectively then our claim is $\mu_1 < \mu_2$ and its complement is $\mu_1 \geq \mu_2$. Since complement contains the equality sign so we can take the complement as the null hypothesis and the claim as the alternative hypothesis. Thus,

$$H_0 : \mu_1 \geq \mu_2 \text{ and } H_1 : \mu_1 < \mu_2$$

Since the alternative hypothesis is left-tailed so the test is left-tailed test.

Since it is given that yield of wheat with and without fertilizer follow normal distributions with equal and unknown variances and other assumptions of t-test for testing a hypothesis about difference of two population means also meet. So we can go for this test.

For testing the null hypothesis, the test statistic t is given by

$$t = \frac{\bar{X} - \bar{Y}}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t_{(n_1+n_2-2)} \quad \text{under } H_0$$

$$= \frac{4.6 - 5.0}{0.41\sqrt{\frac{1}{13} + \frac{1}{13}}} = \frac{-0.4}{0.41 \times 0.39} = \frac{-0.4}{0.16} = -2.5$$

The critical values of test statistic t for left-tailed test corresponding $(n_1 + n_2 - 2) = 24$ df at 1% level of significance is $-t_{(24), 0.01} = -2.492$.

Since calculated value of test statistic t (= −2.5) is less than the critical value (= −2.492), that means calculated value of test statistic t lies in rejection region, so we reject the null hypothesis and support the alternative hypothesis i.e. we support our claim at 1% level of significance.

Thus, we conclude that samples fail to provide us sufficient evidence against the claim so we may assume that there is significant improvement in wheat production due to fertilizer.

E6) Here, we want to test whether the programme PGDAST has improved the performance of the graduate students in Statistics. If μ_1 and μ_2 denote the average marks before and after the programmed so our claim is $\mu_1 < \mu_2$ and its complement is $\mu_1 \geq \mu_2$. Since complement contains the equality sign so we can take the complement as the null hypothesis and the claim as the alternative hypothesis. Thus,

$$H_0 : \mu_1 \geq \mu_2 \text{ and } H_1 : \mu_1 < \mu_2$$

Since the alternative hypothesis is left-tailed so the test is left-tailed test.

It is a situation of before and after. Also, the marks of the students before and after the programme PGDAST follow normal distributions. So, population of differences will also be normal. Also all the assumptions of paired t-test meet. So we can go for paired t-test.

For testing the null hypothesis, the test statistic t is given by

$$t = \frac{\bar{D}}{S_D / \sqrt{n}} \sim t_{(n-1)} \text{ under } H_0 \qquad \ldots (7)$$

Calculation for $\bar{D}$ and S_D :

Participant	**Marks**		**D = (X−Y)**	**D²**
	Before (X)	**After (Y)**		
1	42	45	−3	9
2	46	46	0	0
3	50	60	−10	100
4	36	42	−6	36
5	44	60	−16	256
6	60	72	−12	144
7	62	63	−1	1
8	43	43	0	0
9	70	80	−10	100
10	53	65	−12	144
			$\sum D = -70$	$\sum D^2 = 790$

From above calculation, we have

$$\bar{D} = \frac{1}{n}\sum D = \frac{1}{10}(-70) = -7$$

$$S_D^2 = \frac{1}{n-1}\left[\sum D^2 - \frac{\left(\sum D\right)^2}{n}\right]$$

$$= \frac{1}{9}\left[790 - \frac{(-70)^2}{10}\right] = \frac{1}{9} \times 300 = 33.33$$

$$\Rightarrow S_D = \sqrt{33.33} = 5.77$$

Putting the values in equation (7), we have

$$t = \frac{-7.0}{5.77/\sqrt{10}} = -3.83$$

The critical value of test statistic t for left-tailed test corresponding $(n-1) = 9$ df at 5% level of significance is $-t_{(9),\,0.05} = -1.833$.

Since calculated value of test statistic t (= −3.83) is less than the critical (tabulated) value (= −1.833), that means calculated value of test statistic t lies in rejection region, so we reject the null hypothesis and support the alternative hypothesis i.e. we support our claim at 5% level of significance.

Thus, we conclude that samples fail to provide us sufficient evidence against the claim so we may assume that the participants have significant improvement after the programme "Post Graduate Diploma in Applied Statistics (PGDAST)".

E7) Here, we want to test that the drug has no effect on change in blood pressure. If μ_D denotes the average increment in the blood pressure before drug then our claim is $\mu_D = 0$ and its complement is $\mu_D \neq 0$. Since the claim contains the equality sign so we can take the claim as the null hypothesis and complement as the alternative hypothesis. Thus,

$$H_0 : \mu_D = \mu_1 - \mu_2 = 0 \; [\text{the drug has no effect}]$$

$$H_1 : \mu_D \neq 0 \qquad [\text{the drug has an effect}]$$

It is given that increment in the blood pressure follows normal distribution. Also patients are same in both situations before and after the drug. And all the assumption of paired t-test meet. So we can go for paired t-test.

For testing H_0, the test statistic is given by

$$t = \frac{\bar{D}}{S_D/\sqrt{n}} \sim t_{(n-1)} \text{ under } H_0 \qquad \ldots (8)$$

Calculation for $\bar{D}$ and S_D :

Patient Number	1	2	3	4	5	6	7	8	Total
D	4	0	7	−2	0	−3	2	0	$\sum D = 8$
$(\mathbf{D}-\bar{\mathbf{D}})$	3	−1	6	−3	−1	−4	1	−1	
$(\mathbf{D}-\bar{\mathbf{D}})^2$	9	1	36	9	1	16	1	1	$\sum (D-\bar{D})^2 = 74$

From above calculation, we have

$$\bar{D} = \frac{1}{n}\sum D = \frac{1}{8} \times 8 = 1$$

$$S_D^2 = \frac{1}{n-1}\sum (D-\bar{D})^2 = \frac{1}{7} \times 74 = 10.57$$

$$\Rightarrow S_D = \sqrt{10.57} = 3.25$$

Putting the values in test statistic, we have

$$t = \frac{1}{3.25/\sqrt{8}} = \frac{1}{1.15} = 0.87$$

The critical value of test statistic t for two-tailed test corresponding (n-1) = 7 df at 5% level of significance are $\pm t_{(7),\,0.025} = \pm 2.365$.

Since calculated value of test statistic t (= 0.87) is less than the critical value (= 2.365) and greater than the critical value (=−2.365) that means calculated value of test statistic t lies in non-rejection region, so we do not reject the null hypothesis i.e. we support the claim at 5% level of significance.

Thus, we conclude that samples fail to provide us sufficient evidence against the claim so we may assume that the drug has no effect on the change of blood pressure of patients.

E8) We are given that

$$n = 20,\ r = 0.40$$

and we wish to test that there is a positive linear relationship between the family income and the amounts of money spent per family member on food each month in colony. If ρ denote the correlation coefficient between the family income and the amounts of money spent per family member then the claim is $\rho > 0$ and its complement is $\rho \le 0$. Since complement contains the equality sign so we can take the complement as the null hypothesis and the claim as the alternative hypothesis. Thus,

$$H_0 : \rho \le 0 \text{ and } H_0 : \rho > 0$$

Since the alternative hypothesis is right-tailed so the test is right-tailed test.

For testing the null hypothesis, the test statistic t is given by

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

$$= \frac{0.40\sqrt{20-2}}{\sqrt{1-(0.40)^2}} = \frac{0.40 \times 4.24}{0.92} = 1.84$$

The critical value of test statistic t for right-tailed test corresponding (n-2) = 18 df at 1% level of significance is $t_{(n-2),\alpha} = t_{(18),0.01} = 2.552$.

Since calculated value of test statistic t (= 1.84) is less than the critical value (= 2.552), that means calculated value of test statistic t lies in non-rejection region, so we do not reject the null hypothesis and reject the alternative hypothesis i.e. we reject our claim at 1% level of significance.

Thus, we conclude that sample provide us sufficient evidence against the claim so there is no positive linear correlation between the family income and the amounts of money spent per family member on food each month in colony.

UNIT 12 CHI-SQUARE AND F-TESTS

Structure

12.1 INTRODUCTION

Recall from the previous unit when we test the hypothesis about the difference of means of two populations, t-test needs an assumption of equality of variances of two populations under study. Other than this there are situations where we want to test the hypothesis about the variances of two populations. For example, an economist may want to test whether the variability in incomes differ in two population. In such situations, we use F-test when the populations under study follow the normal distributions.

Similarly, there are many other situations where we need to test the hypothesis about the hypothetical or specified value of the variance of the population under study. For example, the manager of the electric bulbs company would probably be interested whether or not the variability in the life of bulbs is within acceptable limits, the product controller of a milk company may be interested in the variance of the amount of fat in the whole milk processed by the company is no more than the specified level. In such situations, we use χ^2-test when the population under study follows the normal distribution.

This unit is divided into five sections. Section 12.1 is described the need of χ^2 and F-tests. χ^2-test for testing the hypothesis about the population variance is discussed in Section 12.2. And F-test for equality of variances of two populations is discussed in Section 12.3. Unit ends by providing summary of what we have discussed in this unit in Section 12.4 and solution of exercises in Section 12.5.

Objectives

After studying this unit, you should be able to:

- describe the testing of hypothesis for population variance; and
- explain the testing of hypothesis for two population variances.

12.2 TESTING OF HYPOTHESIS FOR POPULATION VARIANCE USING χ^2-TEST

In Section 11.3 of previous unit, we have discussed testing of hypothesis for population mean when the characteristic under study follows the normal distribution but when analysing quantitative data, it is often important to draw conclusion about the average as well as the variability of a characteristic under study. For example, if a company manufactured the electric bulbs then the manager of the company would probably be interested in determining the average life of the bulbs and also determining whether or not the variability in the life of bulbs is within acceptable limits, the product controller of a milk company may be interested in the variance of the amount of fat in the whole milk processed by the company is no more than the specified level, etc.

The procedure of testing a hypothesis for population variance or standard deviation is similar to the testing of population mean. The basic difference is that here we use chi-square test instead of t-test because here the sampling distribution of test statistic follows the chi-square distribution.

Assumptions

This test works under the following assumptions:

(i) The characteristic under study follows normal distribution. In other words, populations from which random sample is drawn should be normal with respect to the characteristic of interest.

(ii) Sample observations are random and independent.

Let $X_1, X_2, ..., X_n$ be a random sample of size n drawn from a normal population with mean μ and variance σ^2, where μ and σ^2 are unknown.

The general procedure of this test is explained below in detail:

As we are doing so far in all tests, first step in hypothesis testing problems is to setup null and alternative hypotheses. Here, we want to test the claim about the hypothesized or specified value σ_0^2 of population variance σ^2 so we can take our null and alternative hypotheses as

$$H_0: \sigma^2 = \sigma_0^2 \text{ and } H_1: \sigma^2 \neq \sigma_0^2 \qquad [\text{for two-tailed test}]$$

or

$$\left.\begin{array}{l} H_0: \sigma^2 \leq \sigma_0^2 \text{ and } H_1: \sigma^2 > \sigma_0^2 \\ H_0: \sigma^2 \geq \sigma_0^2 \text{ and } H_1: \sigma^2 < \sigma_0^2 \end{array}\right\} \qquad [\text{for one-tailed test}]$$

For testing the null hypothesis, the test statistic χ^2 is given by

$$\chi^2 = \frac{\sum\left(X_i - \bar{X}\right)^2}{\sigma_0^2} = \frac{(n-1)S^2}{\sigma_0^2} \sim \chi^2_{(n-1)} \text{ under } H_0 \qquad \text{... (1)}$$

where, $S^2 = \frac{1}{n-1}\sum\left(X - \bar{X}\right)^2$

Here, the test statistic χ^2 follows chi square distribution with $(n-1)$ degrees of freedom as we have discussed in Unit 3 of this course.

After substituting values of n, S and σ_0^2, we get calculated value of test statistic. Let χ^2_{cal} be the calculated value of test statistic χ^2.

Obtain the critical value(s) or cut-off value(s) in the sampling distribution of the test statistic χ^2 and construct rejection (critical) region of size α. The critical value of the test statistic χ^2 for various df and different level of significance α are given in **Table III** of the Appendix given at the end of the Block 1of this course.

After doing all the calculation discussed above, we have to take the decision about rejection or non-rejection of the null hypothesis. The procedure of taking the decision about the null hypothesis is explained is the next page:

For one-tailed test:

Case I: When $H_0: \sigma^2 \le \sigma_0^2$ and $H_1: \sigma^2 > \sigma_0^2$ (right-tailed test)

In this case, the rejection (critical) region falls under the right tail of the probability curve of the sampling distribution of test statistic χ^2. Suppose $\chi^2_{(\nu),\alpha}$ is the critical value at α level of significance where, $\nu = n - 1$, so entire region greater than or equal to $\chi^2_{(\nu),\alpha}$ is the rejection region and less than $\chi^2_{(\nu),\alpha}$ is the non-rejection region as shown in Fig. 12.1.

If $\chi^2_{cal} \ge \chi^2_{(\nu),\alpha}$, that means calculated value of test statistic lies in the rejection (critical) region, then we reject the null hypothesis H_0 at α level of significance. Therefore, we conclude that sample data provides us sufficient evidence against the null hypothesis and there is a significant difference between hypothesized value and observed value of the population variance σ^2.

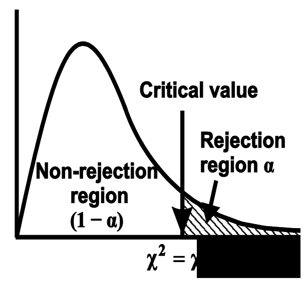

Fig. 12.1

If $\chi^2_{cal} < \chi^2_{(\nu),\alpha}$, that means calculated value of test statistic lies in non-rejection region, then we do not reject the null hypothesis H_0 at α level of significance. Therefore, we conclude that the sample data fails to provide us sufficient evidence against the null hypothesis and the difference between hypothesized value and observed value of the population variance σ^2 due to fluctuation of sample.

Case II: When $H_0: \sigma^2 \ge \sigma_0^2$ and $H_1: \sigma^2 < \sigma_0^2$ (left-tailed test)

In this case, the rejection (critical) region falls under the left tail of the probability curve of the sampling distribution of test statistic χ^2. Suppose $\chi^2_{(\nu),(1-\alpha)}$ is the critical value at α level of significance then entire region less than or equal to $\chi^2_{(\nu),(1-\alpha)}$ is the rejection (critical) region and greater than $\chi^2_{(\nu),(1-\alpha)}$ is the non-rejection region as shown in Fig. 12.2.

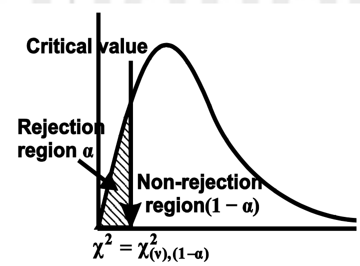

Fig. 12.2

If $\chi^2_{cal} \le \chi^2_{(\nu),(1-\alpha)}$, that means calculated value of test statistic lies in rejection (critical) region, then we reject the null hypothesis H_0 at α level of significance.

For $\chi^2_{cal} > \chi^2_{(\nu),(1-\alpha)}$, that means calculated value of test statistic lies in the non-rejection region, then we do not reject the null hypothesis H_0 at α level of significance.

For two-tailed test:

When $H_0: \sigma^2 = \sigma_0^2$ and $H_1: \sigma^2 \ne \sigma_0^2$

In this case, the rejection region falls under both tails of the probability curve of sampling distribution of the test statistic χ^2 and half the area (α) i.e. $\alpha/2$ of rejection (critical) region lies at left tail and other half on the right tail. Suppose $\chi^2_{(\nu),(1-\alpha/2)}$ and $\chi^2_{(\nu),\alpha/2}$ are the two critical values at the left-tailed and right-tailed respectively on pre-fixed α-level of significance. Therefore, entire region less than or equal to $\chi^2_{(\nu),(1-\alpha/2)}$ and greater than or equal to $\chi^2_{(\nu),\alpha/2}$ are the rejection

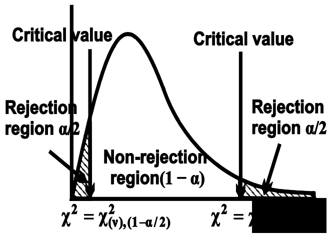

Fig. 12.3

(critical) regions and between $\chi^2_{(\nu),(1-\alpha/2)}$ and $\chi^2_{(\nu),\alpha/2}$ is the non-rejection region as shown in Fig. 12.3.

If $\chi^2_{cal} \ge \chi^2_{(\nu),\alpha/2}$ or $\chi^2_{cal} \le \chi^2_{(\nu),(1-\alpha/2)}$, that means calculated value of test statistic χ^2 lies in rejection (critical) region, then we reject the null hypothesis H_0 at α level of significance.

If $\chi^2_{(\nu),(1-\alpha/2)} < \chi^2_{cal} < \chi^2_{(\nu),\alpha/2}$, that means calculated value of test statistic χ^2 lies in non-rejection region, then we do not reject the null hypothesis H_0 at α level of significance.

Procedure of taking the decision about the null hypothesis on the basis of p-value:

You have done it in so many test, it has become a routine thing for you. You know that to take the decision about the null hypothesis on the basis of p-value, the p-value is compared with level of significance (α) and if p-value is less than or equal to α then we reject the null hypothesis and if the p-value is greater than α we do not reject the null hypothesis.

For χ^2-test, p-value is defined as:

For one-tailed test:

For $H_1: \sigma^2 > \sigma_0^2$ (right-tailed test)

$$\text{p-value} = P\left[\chi^2 \ge \chi^2_{cal}\right]$$

For $H_1: \sigma^2 < \sigma_0^2$ (left-tailed test)

$$\text{p-value} = P\left[\chi^2 \le \chi^2_{cal}\right]$$

For two-tailed test: $H_1: \sigma^2 \ne \sigma_0^2$

For two-tailed test the p-value is approximated as

$$\text{p-value} = 2P\left[\chi^2 \ge \chi^2_{cal}\right]$$

The p-value for χ^2-test can be obtained with the help of the **Table-III (χ^2-table)** given in the Appendix at the end of Block 1 of this course. Similar to t-test, this table gives the χ^2 values corresponding to the standard values of α such as 0.995, 0.99, 0.10, 0.05, 0.025, 0.01, etc only. therefore, the exact p-value is not obtained with the help of this table and we can approximate the p-value for this test.

For example, if test is right-tailed and calculated (observed) value of test statistic χ^2 is 25.10 with 12 df then p-value is calculated as:

Since test statistic is based on the 12 df therefore, we use row for 12 df in the χ^2-table and move across this row to find the values in which calculated χ^2-value falls. Since calculated χ^2-value falls between 23.24 and 26.22, corresponding to one-tailed area $\alpha = 0.025$ and 0.01 respectively, therefore p-value lies between 0.01 and 0.025, that is,

$$0.01 < \text{p-value} < 0.025$$

If in the above example the test is two-tailed then the two values 0.01 and 0.005 would be doubled for p-value, that is,

$$2 \times 0.01 = 0.02 < \text{p-value} < 0.05 = 2 \times 0.025$$

Note 1: With the help of computer packages and softwares such as SPSS, SAS, MINITAB, EXCEL, etc. we can find the exact p-value for χ^2-test.

Let us do some examples to become more user friendly with the test explained above.

Example 1: The variance of a certain dimension article produced by a machine is 7.2 over a long period. A random sample of 20 articles gave a variance 8. Is it justifiable to conclude that variability has increased at 5% level of significance assuming that the measurement of dimension article is normally distributed?

Solution: Here, we are given that

Sample size = n = 20

Sample variance = S^2 = 8

Specified value of population variance under test = $\sigma_0^2 = 7.2$

Here, we want to test that variability of dimension article produced by a machine has increased. Since variability is measured in terms of variance (σ^2) so our claim is $\sigma^2 > 7.2$ and its complement is $\sigma^2 \le 7.2$. Since complement contains the equality sign so we can take the complement as the null hypothesis and the claim as the alternative hypothesis. Thus,

$$H_0 : \sigma^2 \le \sigma_0^2 = 7.2$$

$$H_1 : \sigma^2 > 7.2 \qquad [\text{variability of dimension article has increased}]$$

Since the alternative hypothesis is right-tailed so the test is right-tailed test.

Here, we want to test the hypothesis about the population variance and sample size is small n = 20(< 30). Also we are given that measurement of dimension article follows normal distribution so we can go for χ^2 test for population variance.

So, test statistic is given by

$$\chi^2 = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)} \text{ under } H_0$$

$$= \frac{19 \times 8}{7.2} = 21.11$$

The critical (tabulated) value of test statistic χ^2 for right-tailed test corresponding (n-1) = 19 df at 5% level of significance is $\chi^2_{(n-1),\alpha} = \chi^2_{(19),0.05} = 30.14$.

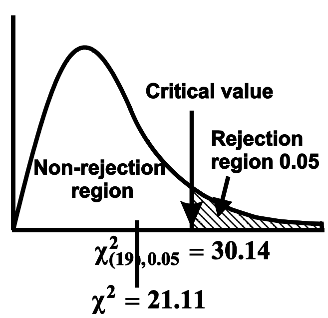

Fig. 12.4

Since calculated value of test statistic (= 21.11) is less than the critical (tabulated) value (= 30.14), that means calculated value of test statistic lies in non-rejection region as shown in Fig. 12.4, so we do not reject the null hypothesis and reject the alternative hypothesis i.e. we reject our claim at 5% level of significance.

Decision according to p-value:

Since test statistic is based on 19 df therefore, we use row for 19 df in the χ^2-table and move across this row to find the values in which calculated χ^2-value falls. Since calculated χ^2-value falls between 11.65 and 27.20, corresponding to one-tailed area α = 0.90 and 0.10 respectively therefore p-value lies between 0.10 and 0.90, that is,

$$0.10 < \text{p-value} < 0.90$$

Since p-value is greater than $\alpha(=0.05)$ so we do not reject the null hypothesis at 5% level of significance.

Thus, we conclude that sample provides us sufficient evidence against the claim so we may assume that the variability of dimension article produced by a machine is not increased.

Example 2: The12 measurements of the same object on an instrument are given below:

1.6, 1.5, 1.3, 1.5, 1.7, 1.6, 1.5, 1.4, 1.6, 1.3, 1.5, 1.5

If the measurement of the instrument follows normal distribution then carry out the test at 1% level of significance that variance in the measurement of the instrument is less than 0.016.

Solution: Here, we are given

Sample size = n = 12

Specified value of population variance under test $\sigma_0^2 = 0.016$

Here, we want to test that the variance (σ^2) in the measurements of the instrument is less than 0.016. So our claim is $\sigma^2 < 0.016$ and its complement is $\sigma^2 \geq 0.016$. Since complement contains the equality sign so we can take complement as null hypothesis and claim as the alternative hypothesis. Thus,

$$H_0 : \sigma^2 \geq \sigma_0^2 = 0.016 \text{ and } H_1 : \sigma^2 < 0.016$$

Since the alternative hypothesis is left-tailed so the test is left-tailed test.

Here, we want to test the hypothesis about the population variance and sample size is small n = 12(< 30). Also we are given that the measurement of the instrument follows normal distribution so we can go for χ^2-test for population variance.

For testing the null hypothesis, the test statistic χ^2 is given by

$$\chi^2 = \frac{\sum\left(X_i - \bar{X}\right)^2}{\sigma_0^2} \sim \chi^2_{(n-1)} \text{ under } H_0 \qquad \text{... (2)}$$

Calculation for $\sum\left(X_i - \bar{X}\right)^2$:

X	$(X-\bar{X})$	$(X-\bar{X})^2$
1.6	0.1	0.01
1.5	0	0
1.3	-0.2	0.04
1.5	0	0
1.7	0.2	0.04
1.6	0.1	0.01
1.5	0	0
1.4	-0.1	0.01
1.6	0.1	0.01
1.3	-0.2	0.04
1.5	0	0
1.5	0	0
$\sum X = 18$	0	$\sum(X-\bar{X})^2 = 0.16$

From above calculation, we have

$$\bar{X} = \frac{1}{n}\sum X = \frac{1}{12} \times 18 = 1.5$$

Putting the values of $\sum\left(X - \bar{X}\right)^2$ and σ_0^2 in equation (2), we have

$$\chi^2 = \frac{\sum\left(X_i - \bar{X}\right)^2}{\sigma_0^2} = \frac{0.16}{0.016} = 10$$

The critical value of test statistic χ^2 for left-tailed test corresponding (n - 1) = 11 df at 5% level of significance is $\chi^2_{(n-1),(1-\alpha)} = \chi^2_{(11),0.95} = 4.57$.

Since calculated value of test statistic (= 10) is greater than the critical value (= 4.57), that means calculated value of test statistic lies in non-rejection region as shown in Fig. 12.5 so we do not reject the null hypothesis and reject the alternative hypothesis i.e. we reject the claim at 5% level of significance.

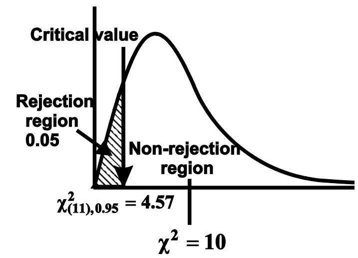

Fig. 12.5

Thus, we conclude that sample provide sufficient evidence against the claim so the variance in the measurement of the instrument is not less than 0.016.

In the same way, you can try the following exercises.

E1) An ambulance agency claims that the standard deviation in the length of serving times is less than 15 minutes. Investigator suspects that this claim is wrong and takes a random sample of 20 serving times which has a standard deviation of 17 minutes. Assume that the service time of the ambulance follows normal distribution. Test at α = 0.01, is there enough evidence to reject the agency's claim?

E2) A cigarette manufacturer claims that the variance of nicotine content of its cigarettes is 0.62. Nicotine content is measured in milligrams and is normally distributed. A sample of 25 cigarettes has a variance of 0.65. Test the manufacturer's claim at 5% level of significance.

12.5 TESTING OF HYPOTHESIS FOR TWO POPULATION VARIANCES USING F-TEST

In Section 12.1 we have already mentioned that before applying t-test for difference of two population means, one of the requirements is to check the equality of variances of two populations. This assumption can be checked with the help of F-test for two population variances. This F-test is also important in a number of contexts. For example, an economist may want to test whether the variability in incomes differ in two populations, a quality controller may want to test whether the quality of the product is changing over time, etc.

Some author uses this test as the name "homoscedastic test". Because two populations with equal variances are called homoscedastic. This word is derived from two wards, homo means "the same" and scedastic means "variability".

Assumptions

The assumptions for F-test for testing the variances of two populations are:

1. The populations from which the samples are drawn must be normally distributed.
2. The samples must be independent of each other.

Now, we come to the general procedure of this test.

Let $X_1, X_2, ..., X_{n_1}$ be a random sample of size n_1 from a normal population with mean μ_1 and variance σ_1^2. Similarly, $Y_1, Y_2, ..., Y_{n_2}$ be a random sample of size n_2 from another normal population with mean μ_2 and variance σ_2^2. Here, we want to test the hypothesis about the two population variances so we can take our alternative null and hypotheses as

$$H_0 : \sigma_1^2 = \sigma_2^2 = \sigma^2 \text{ and } H_1 : \sigma_1^2 \neq \sigma_2^2 \qquad [\text{for two-tailed test}]$$

or

$$\left.\begin{array}{l} H_0 : \sigma_1^2 \leq \sigma_2^2 \text{ and } H_1 : \sigma_1^2 > \sigma_2^2 \\ H_0 : \sigma_1^2 \geq \sigma_2^2 \text{ and } H_1 : \sigma_1^2 < \sigma_2^2 \end{array}\right\} \qquad [\text{for one-tailed test}]$$

For testing the null hypothesis, the test statistic F is given by

$$F = \frac{S_1^2}{S_2^2} \sim F_{(n_1-1, n_2-1)} \quad \text{under } H_0 \qquad \ldots (3)$$

where, $S_1^2 = \frac{1}{n_1-1}\sum\left(X-\bar{X}\right)^2$ and $S_2^2 = \frac{1}{n_2-1}\sum\left(Y-\bar{Y}\right)^2$.

For computational simplicity, we can also write

$$S_1^2 = \frac{1}{n_1-1}\left(\sum X^2 - \frac{\left(\sum X\right)^2}{n_1}\right) \text{ and } S_2^2 = \frac{1}{n_2-1}\left(\sum Y^2 - \frac{\left(\sum Y\right)^2}{n_2}\right)$$

The test statistic F follows F-distribution with $\nu_1 = (n_1 - 1)$ and $\nu_2 = (n_2 - 1)$ degrees of freedom as discussed in Unit 4 of this course.

After substituting the values of S_1^2 and S_2^2, we get calculated value of test statistic. Let F_{cal} be calculated value of test statistic F.

Obtain the critical value(s) or cut-off value(s) in the sampling distribution of the test statistic F and construct rejection (critical) region of size α. The critical values of the test statistic F for various df and different level of significance α are given in **Table IV (F-table)** of the Appendix at the end of Block 1of this course.

After doing all calculations discussed above, we have to take the decision about rejection or non rejection of the null hypothesis. This is explained below:

In case of one-tailed test:

Case I: When $H_0 : \sigma_1^2 \leq \sigma_2^2$ and $H_1 : \sigma_1^2 > \sigma_2^2$ (right-tailed test)

In this case, the rejection (critical) region falls at the right side of the probability curve of the sampling distribution of test statistic F. Suppose $F_{(\nu_1,\nu_2),\alpha}$ is the critical value of test statistic F with ($\nu_1 = n_1 - 1$, $\nu_2 = n_2 - 1$) df at α level of significance so entire region greater than or equal to $F_{(\nu_1,\nu_2),\alpha}$ is the rejection (critical) region and less than $F_{(\nu_1,\nu_2),\alpha}$ is the non-rejection region as shown in Fig. 12.6.

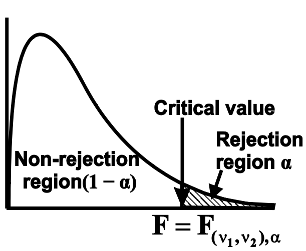

Fig. 12.6

If $F_{cal} \geq F_{(\nu_1,\nu_2),\alpha}$, that means calculated value of test statistic lies in rejection (critical) region, then we reject the null hypothesis H_0 at α level of significance. Therefore, we conclude that samples data provide us sufficient evidence against the null hypothesis and there is a significant difference between population variances.

If $F_{cal} < F_{(\nu_1,\nu_2),\alpha}$, that means calculated value of test statistic lies in non-rejection region, then we do not reject the null hypothesis H_0 at α level of significance. Therefore, we conclude that the samples data fail to provide us sufficient evidence against the null hypothesis and the difference between population variances due to fluctuation of sample.

Case II: When $H_0: \sigma_1^2 \geq \sigma_2^2$ and $H_1: \sigma_1^2 < \sigma_2^2$ (left-tailed test)

In this case, the rejection (critical) region falls at the left side of the probability curve of the sampling distribution of test statistic F. Suppose $F_{(\nu_1,\nu_2),(1-\alpha)}$ is the critical value at α level of significance then entire region less than or equal to $F_{(\nu_1,\nu_2),(1-\alpha)}$ is the rejection(critical) region and greater than $F_{(\nu_1,\nu_2),(1-\alpha)}$ is the non-rejection region as shown in Fig. 12.7.

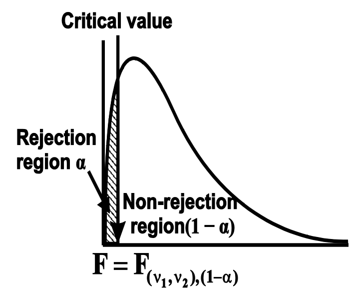

Fig. 12.7

If $F_{cal} \leq F_{(\nu_1,\nu_2),(1-\alpha)}$, that means calculated value of test statistic lies in rejection (critical) region, then we reject the null hypothesis H_0 at α level of significance.

If $F_{cal} > F_{(\nu_1,\nu_2),(1-\alpha)}$, that means calculated value of test statistic lies in non-rejection region, then we do not reject the null hypothesis H_0 at α level of significance.

Note 2: F-table mentioned above gives only the right tailed critical values with different degrees of freedom at different level of significance. The left tailed critical value of F-test can always be obtained by the formula given below (as described in Unit 4 of this course):

$$F_{(\nu_1,\nu_2),(1-\alpha)} = \frac{1}{F_{(\nu_2,\nu_1),\alpha}}$$

In case of two-tailed test:

When $H_0: \sigma_1^2 = \sigma_2^2$ and $H_1: \sigma_1^2 \neq \sigma_2^2$

In this case, the rejection (critical) region falls at both sides of the probability curve of the sampling distribution of test statistic F and half the area(α) i.e. α/2 of rejection (critical) region lies at left tail and other half on the right tail.

Suppose $F_{(\nu_1,\nu_2),(1-\alpha/2)}$ and $F_{(\nu_1,\nu_2),\alpha/2}$ are the two critical values at the left-tailed and right-tailed respectively on pre-fixed α level of significance. Therefore, entire region less than or equal to $F_{(\nu_1,\nu_2),(1-\alpha/2)}$ and greater than or equal to $F_{(\nu_1,\nu_2),\alpha/2}$ are the rejection (critical) regions and between $F_{(\nu_1,\nu_2),(1-\alpha/2)}$ and $F_{(\nu_1,\nu_2),\alpha/2}$ is the non-rejection region as shown in Fig. 12.8.

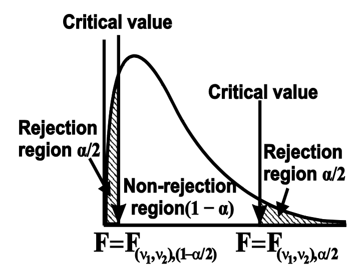

Fig. 12.8

If $F_{cal} \geq F_{(\nu_1,\nu_2),\alpha/2}$ or $F_{cal} \leq F_{(\nu_1,\nu_2),(1-\alpha/2)}$, that means calculated value of test statistic lies in rejection(critical) region, then we reject the null hypothesis H_0 at α level of significance.

If $F_{(\nu_1,\nu_2),(1-\alpha/2)} < F_{cal} < F_{(\nu_1,\nu_2),\alpha/2}$, that means calculated value of test statistic F lies in non-rejection region, then we do not reject the null hypothesis H_0 at α level of significance.

Procedure of taking the decision about the null hypothesis on the basis of p-value:

To take the decision about the null hypothesis on the basis of p-value, the p-value is compared with level of significance (α) and if p-value is less than or equal to α then we reject the null hypothesis and if the p-value is greater than α we do not reject the null hypothesis.

For F-test, p-value can be defined as:

For one-tailed test:

For $H_1: \sigma_1^2 > \sigma_2^2$ (right-tailed test)

$$\text{p-value} = P\left[F \geq F_{cal}\right]$$

For $H_1: \sigma_1^2 < \sigma_2^2$ (left-tailed test)

$$\text{p-value} = P\left[F \leq F_{cal}\right]$$

For two-tailed test: $H_1: \sigma_1^2 \neq \sigma_2^2$

For two-tailed test the p-value is approximated as

$$\text{p-value} = 2P\left[F \geq F_{cal}\right]$$

The p-value for F-test can be obtained with the help of **Table-IV (F-table)** given in the Appendix at the end of Block 1 of this course. Similar to t-test or χ^2-test, this table gives F values corresponding to the standard values of α such as 0.10, 0.05, 0.025 and 0.01 only. Therefore with the help of F-table, we cannot find the exact p-value. So we can approximate p-value for this test.

For example, if test is right-tailed and calculated (observed) value of test statistic F is 2.65 with 24 degrees of freedom of numerator and 14 degrees of freedom of the denominator then we can find the p-value as:

Since test statistic is based on (24, 14) df therefore, we move across all the values of tabulated F corresponding to (24, 14) df at α such as 0.10, 0.05, 0.025 and 0.01 and find the values in which calculated F-value falls. Since we have F tabulated with (24, 14) df at α = 0.10, 0.05, 0.025 and 0.01 as

α =	0.10	0.05	0.025	0.01
$F_{(24,\,14),\,\alpha}$	1.94	2.35	2.79	3.43

Since calculated F-value (= 2.65) falls between 2.35 and 2.79, corresponding to one-tailed area α = 0.05 and 0.025 respectively therefore p-value lies between 0.0025 and 0.05, that is,

$$0.025 < \text{p-value} < 0.05$$

If in the above example the test is two-tailed then the two values 0.025 and 0.05 would be doubled for p-value, that is,

$$0.05 < \text{p-value} < 0.10$$

Note 3: With the help of computer packages and software such as SPSS, SAS, MINITAB, EXCEL, etc. we can find the exact p-value for F-test.

Let us do some examples based on this test.

Example 3: The following data relate to the number of items produced in a shift by two workers A and B for some days:

A	26	37	40	35	30	30	40	26	30	35	45
B	19	22	24	27	24	18	20	19	25		

Assuming that the parent populations are normal, can it be inferred that B is more stable (or consistent) worker compared to A?

Solution: Here, we want to test that worker B is more stable than worker A. As we know that stability of data is related to variance of the data. Smaller value of the variance implies data that it is more stable. Therefore, to compare stability of two workers, it is enough to compare their variances. If σ_1^2 and σ_2^2 denote the variances of worker A and worker B respectively then our claim is $\sigma_1^2 > \sigma_2^2$ and its complement is $\sigma_1^2 \le \sigma_2^2$. Since complement contains the equality sign so we can take the complement as the null hypothesis and the claim as the alternative hypothesis. Thus,

$$H_0 : \sigma_1^2 \le \sigma_2^2$$

$$H_1 : \sigma_1^2 > \sigma_2^2 \left[\text{worker B is more stable than worker A}\right]$$

Since the alternative hypothesis is right-tailed so the test is right-tailed test.

Here, we want to test the hypothesis about two population variances and sample sizes n_1 = 11(< 30) and n_2 = 9 (< 30) are small. Also populations under study are normal and both samples are independent so we can go for F-test for two population variances.

For testing the null hypothesis, test statistic is given by

$$F = \frac{S_1^2}{S_2^2} \sim F_{(n_1-1,\, n_2-1)} \qquad \ldots (4)$$

where, $S_1^2 = \frac{1}{n_1 - 1}\sum_{i=1}^{n_1}\left(X_i - \bar{X}\right)^2$ and $S_2^2 = \frac{1}{n_2 - 1}\sum_{i=1}^{n_2}\left(Y_i - \bar{Y}\right)^2$

Calculation for S_1^2 and S_2^2:

Items Produced by A (Variable X)	$(X - \bar{X}) = (X - 34)$	$(X - \bar{X})^2$	Items Produced by B (Variable Y)	$(Y - \bar{Y}) = (Y - 22)$	$(Y - \bar{Y})^2$
26	−8	64	19	−3	9
37	3	9	22	0	0
40	6	36	24	2	4
35	1	1	27	5	25
30	−4	16	24	2	4
30	−4	16	18	−4	16
40	6	36	20	−2	4
26	−8	64	19	−3	9
30	−4	16	25	3	9
35	1	1			
45	11	121			
Total = 374	0	380	198	0	80

Therefore, we have

$$\bar{X} = \frac{1}{n_1}\sum X = \frac{1}{11}\times 374 = 34$$

and

$$\bar{Y} = \frac{1}{n_2}\sum Y = \frac{1}{9}\times 198 = 22$$

Thus,

$$S_1^2 = \frac{1}{n_1 - 1}\sum\left(X - \bar{X}\right)^2 = \frac{1}{10}\times 380 = 38$$

$$S_2^2 = \frac{1}{n_2 - 1}\sum\left(Y - \bar{Y}\right)^2 = \frac{1}{8}\times 80 = 10$$

Putting the value of S_1^2 and S_2^2 in equation (4), we have

$$F = \frac{38}{10} = 3.8$$

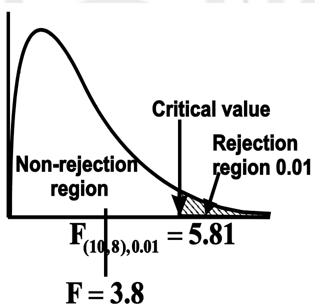

Fig. 12.9

The critical (tabulated) value of test statistic F for right-tailed test corresponding $(n_1 - 1, n_2 - 1) = (10, 8)$ df at 1% level of significance is $F_{(n_1-1, n_2-1),\alpha} = F_{(10,8),0.01} = 5.81$.

Since calculated value of test statistic (= 3.8) is less than the critical value (= 5.81), that means calculated value of test statistic lies in rejection region as shown in Fig. 12.9, so we do not reject the null hypothesis and reject the alternative hypothesis i.e. we reject the claim at 1% level of significance.

Thus, we conclude that samples provide us sufficient evidence against the claim so worker B is not more stable (or consistent) worker compared to A.

Example 4: Two random samples drawn from two normal populations gave the following results:

Sample	Size	Mean	Sum of Squares of Deviation from the Mean
Sample I	9	59	26
Sample II	11	60	32

Test whether both samples are from the same normal populations?

Solution: Since we have to test whether both the samples are from same normal population, therefore, we will test two hypotheses separately:

(i) Two population means are equal, i.e. $H_0 : \mu_1 = \mu_2$

(ii) Two population variances are equal, i.e. $H_0 : \sigma_1^2 = \sigma_2^2$

Since sample sizes are small and populations under study are normal so two means will be tested using t-test whereas two variances will be tested using F-test. But t-test is based on the prior assumption that both population variances are same, therefore, first we apply F-test and later the t- test (when F- test accepts equality hypothesis).

Given that

$$n_1 = 9, \quad \bar{X} = 59, \quad \sum\left(X - \bar{X}\right)^2 = 26$$

$$n_2 = 11, \quad \bar{Y} = 60, \quad \sum\left(Y - \bar{Y}\right)^2 = 32$$

Therefore,

$$S_1^2 = \frac{1}{n_1 - 1}\sum\left(X - \bar{X}\right)^2 = \frac{1}{9-1} \times 26 = 3.25$$

$$S_2^2 = \frac{1}{n_2 - 1}\sum\left(Y - \bar{Y}\right)^2 = \frac{1}{11-1} \times 36 = 3.60$$

First we want to test that the variances of both normal populations are equal so our claim is $\sigma_1^2 = \sigma_2^2$ and its complement is $\sigma_1^2 \neq \sigma_2^2$. Thus, we can take the null and alternative hypotheses as

$$H_0 : \sigma_1^2 = \sigma_2^2 \text{ and } H_1 : \sigma_1^2 \neq \sigma_2^2$$

Since the alternative hypothesis is two-tailed so the test is two-tailed test.

For testing this, the test statistic F is given by

$$F = \frac{S_1^2}{S_2^2} \sim F_{(n_1-1, n_2-1)}$$

$$= \frac{3.20}{3.65} = 0.88$$

The critical (tabulated) value of test statistic F for two-tailed test corresponding $(n_1 - 1, n_2 - 1) = (8, 11)$ df at 5% level of significance are $F_{(n_1-1, n_2-1), \alpha/2} =$

$$F_{(8,11), 0.025} = 3.66 \text{ and } F_{(n_1-1, n_2-1), (1-\alpha/2)} = \frac{1}{F_{(n_2-1, n_1-1), \alpha/2} =} = \frac{1}{F_{(11,8), 0.025}} = \frac{1}{4.24} = 0.24.$$

Since calculated value of test statistic (= 0.88) is less than the critical value (= 3.66) and greater than the critical value (= 0.24), that means calculated value of test statistic F lies in non-rejection region so we do not reject the null hypothesis i.e. we support the claim.

Thus, we conclude that both samples may be taken from normal populations having equal variances.

Now, we test that the means of two normal populations are equal so our claim is $\mu_1 = \mu_2$ and its complement is $\mu_1 \neq \mu_2$. Thus, we can take the null and alternative hypotheses as

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

The test statistic is given by

$$t = \frac{\bar{X} - \bar{Y}}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \qquad \ldots (5)$$

where, $S_p^2 = \frac{1}{n_1 + n_2 - 2}\left[\sum\left(X - \bar{X}\right)^2 + \sum\left(Y - \bar{Y}\right)^2\right]$

$$= \frac{1}{9 + 11 - 2}\left(26 + 32\right) = \frac{1}{18} \times 58 = 3.22$$

$$S_p = \sqrt{3.22} = 1.79$$

Putting the values of $\bar{X}, \bar{Y}, S_p, n_1$ and n_2 in equation (5), we have,

$$t = \frac{59 - 60}{1.79\sqrt{\frac{1}{9} + \frac{1}{11}}} = \frac{-1}{1.79 \times 0.45} = \frac{-1}{0.81} = -1.23$$

The critical values of test statistic t for ($n_1 + n_2$ -2) = 18 df at 5% level of significance for two-tailed test are ± $t_{(18),0.025}$ = ± 2.101.

Since calculated value of test statistic t (= −1.23) is less than the critical value (= 2.101) and greater than the critical value (= –2.101), that means calculated value of test statistic t lies in non-rejection region so we do not reject the null hypothesis i.e. we support the claim.

Thus, we conclude that both samples may be taken from normal populations having equal means.

Hence, overall we conclude that both samples may come from the same normal populations.

Now, you can try the following exercises.

E3) Two sources of raw materials are under consideration by a bulb manufacturing company. Both sources seem to have similar characteristics but the company is not sure about their respective uniformity. A sample of 12 lots from source A yields a variance of 125 and a sample of 10 lots from source B yields a variance of 112. Is it likely that the variance of source A significantly differs to the variance of source B at significance level α = 0.01?

E4) A laptop computer maker uses battery packs of two brands, A and B. While both brands have the same average battery life between charges (LBC), the computer maker seems to receive more complaints about shorter LBC than expected for battery packs of brand A. The computer maker suspects that this could be caused by higher variance in LBC for brand A. To check that, ten new battery packs from each brand are selected, installed on the same models of laptops, and the laptops are allowed to run until the battery packs are completely discharged. The following are the observed LBCs in hours:

Brand A	3.2	3.7	3.1	3.3	2.5	2.2	3.2	3.1	3.2	4.3
Brand B	3.4	3.6	3.0	3.2	3.2	3.2	3.0	3.1	3.2	3.2

Assuming that the LBCs of both brands follows normal distribution, test the LBCs of brand A have a larger variance that those of brand B at 5% level of significance.

We now end this unit by giving a summary of what we have covered in it.

12.4 SUMMARY

In this unit, we have discussed the following points:

1. Testing of hypothesis for population variance using χ^2-test.
2. Testing of hypothesis for two population variances using F-test.

12.5 SOLUTIONS / ANSWERS

E1) Here, we are given that

$$\sigma_0 = 15, \quad n = 20, \quad S = 17$$

Here, we want to test the agency's claim that the standard deviation (σ) of the length of serving times is less than 15 minutes. So our claim is $\sigma < 15$ and its complement is $\sigma \geq 15$. Since complement contains the equality sign so we can take complement as null hypothesis and claim as the alternative hypothesis. Thus,

$$H_0 : \sigma \geq \sigma_0 = 15$$

$$H_1 : \sigma < 15 \quad \begin{bmatrix} \text{SD of the length of serving} \\ \text{times is less than 15 minutes} \end{bmatrix}$$

Since the alternative hypothesis is left-tailed so the test is left-tailed test.

Here, we want to test the hypothesis about the population standard deviation and sample size is small n = 20(< 30). Also we are given that the service time of the ambulance follows normal so we can go for χ^2 test for population variance.

The test statistic is given by

$$\chi^2 = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)}$$

$$= \frac{19 \times (17.0)^2}{(15.0)^2} = 24.40$$

The critical value of test statistic χ^2 for left-tailed test corresponding (n - 1) = 19 df at 1% level significance is $\chi^2_{(n-1),(1-\alpha)} = \chi^2_{(19),0.99} = 7.63$.

Since calculated value of test statistic (= 24.40) is greater than the critical value (= 7.63), that means calculated value of test statistic lies in non-rejection region so we do not reject the null hypothesis and reject the alternative hypothesis i.e. we reject our claim at 1% level of significance.

Thus, we conclude that sample provides us sufficient evidence against the claim so agency's claim that the standard deviation (σ) of the length of serving times is less than 15 minutes is not true.

E2) Here, we are given that

$$\sigma_0^2 = 0.62, \quad n = 25, \quad S^2 = 0.65$$

Here, we want to test the cigarette manufacturer's claims that the variance (σ^2) of nicotine content of its cigarettes is 0.62 milligrams. So claim is $\sigma^2 = 0.62$ and its complement is $\sigma \neq 0.62$. Since claim contains the equality sign so we can take claim as null hypothesis and complement as the alternative hypothesis. Thus,

$$H_0 : \sigma^2 = \sigma_0^2 = 0.62 \text{ and } H_1 : \sigma^2 \neq 0.62$$

Since the alternative hypothesis is two-tailed so the test is two-tailed test.

Here, we want to test the hypothesis about the population variance and sample size is small n = 25(< 30). Also we are given that the nicotine content of its cigarettes follows normal distribution so we can go for χ^2 test for population variance.

The test statistic is given by

$$\chi^2 = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{(n-1)}$$

$$= \frac{24 \times 0.65}{0.62} = 25.16$$

The critical (tabulated) values of test statistic χ^2 for two-tailed test corresponding (n - 1) = 24 df at 5% level of significance are

$\chi^2_{(n-1),\alpha/2} = \chi^2_{(24),0.025} = 39.36$ and $\chi^2_{(n-1),(1-\alpha/2)} = \chi^2_{(24),0.975}$ =12.40.

Since calculated value of test statistic (= 25.16) is less than the critical value (= 39.36) and greater than the critical value (= 12.40), that means calculated value of test statistic lies in non-rejection region, so we do not reject the null hypothesis i.e. we support the claim at 5% level of significance.

Thus, we conclude that sample fails to provide sufficient evidence against the claim so we may assume that manufacturer's claim that the variance of the nicotine content of the cigarettes is 0.62 milligram is true.

E3) Here, we are given that

$$n_1 = 12, \qquad S_1^2 = 125, \qquad n_2 = 10, \qquad S_2^2 = 112$$

Here, we want to test that variance of source A significantly differs to the variances of source B. If σ_1^2 and σ_2^2 denote the variances in the raw materials of sources A and B respectively so our claim is $\sigma_1^2 \neq \sigma_2^2$ and its complement is $\sigma_1^2 = \sigma_2^2$. Since complement contains the equality sign so we can take the complement as the null hypothesis and the claim as the alternative hypothesis. Thus,

$$H_0 : \sigma_1^2 = \sigma_2^2 \text{ and } H_1 : \sigma_1^2 \neq \sigma_2^2$$

Since the alternative hypothesis is two-tailed so the test is two-tailed test.

Here, we want to test the hypothesis about two population variances and sample sizes n_1 = 12(< 30) and n_2 = 10 (< 30) are small. Also populations under study are normal and both samples are independent so we can go for F-test for two population variances.

For testing this, the test statistic is given by

$$F = \frac{S_1^2}{S_2^2} \sim F_{(n_1-1, n_2-1)}$$

$$= \frac{125}{112} = 1.11$$

The critical (tabulated) value of test statistic F for two-tailed test corresponding (n_1 - 1, n_2 - 1) = (11, 9) df at 5% level of significance are

$F_{(n_1-1, n_2-1),\alpha/2} = F_{(11,9),0.025} = 3.91$ and

$$F_{(n_1-1, n_2-1),(1-\alpha/2)} = \frac{1}{F_{(n_2-1, n_1-1),\alpha/2}} = \frac{1}{F_{(9, 11), 0.025}} = \frac{1}{3.59} = 0.28.$$

Since calculated value of test statistic (= 1.11) is less than the critical value (= 3.91) and greater than the critical value (= 0.28), that means calculated value of test statistic lies in non-rejection region, so we do not reject the null hypothesis and reject the alternative hypothesis i.e. we reject the claim at 5% level of significance.

Thus, we conclude that samples provide us sufficient evidence against the claim so we may assume that the variances of source A and B is differ.

E4) Here, we want to test that the LBCs of brand A have a larger variance than those of brand B. If σ_1^2 and σ_2^2 denote the variances in the LBCs of brands A and B respectively so our claim is $\sigma_1^2 > \sigma_2^2$ and its complement is $\sigma_1^2 \le \sigma_2^2$. Since complement contains the equality sign so we can take the complement as the null hypothesis and the claim as the alternative hypothesis. Thus,

$$H_0 : \sigma_1^2 \le \sigma_2^2 \text{ and } H_1 : \sigma_1^2 > \sigma_2^2$$

Since the alternative hypothesis is right-tailed so the test is right-tailed test.

Here, we want to test the hypothesis about two population variances and sample sizes n_1 = 10(< 30) and n_2 = 10 (< 30) are small. Also populations under study are normal and both samples are independent so we can go for F-test for two population variances.

For testing the null hypothesis, test statistic is given by

$$F = \frac{S_1^2}{S_2^2} \sim F_{(n_1-1,\, n_2-1)} \qquad \ldots (6)$$

where, $S_1^2 = \dfrac{1}{n_1-1}\sum\left(X-\bar{X}\right)^2 = \dfrac{1}{n_1-1}\left[\sum X^2 - \dfrac{\left(\sum X\right)^2}{n_1}\right]$ and

$$S_2^2 = \frac{1}{n_2-1}\sum\left(Y-\bar{Y}\right)^2 = \frac{1}{n_2-1}\left[\sum Y^2 - \frac{\left(\sum Y\right)^2}{n_2}\right].$$

Calculation for S_1^2 and S_2^2 :

LBCs of Brand A (X)	X^2	LBCs of Brand A (Y)	Y^2
3.7	13.69	3.6	12.96
3.2	10.24	3.2	10.24
3.3	10.89	3.2	10.24
3.1	9.61	3.0	9.00
2.5	6.25	3.0	9.00
2.2	4.84	3.2	10.24
3.1	9.61	3.2	10.24
3.2	10.24	3.1	9.61
4.3	18.49	3.2	10.24
3.2	10.24	3.1	9.61
Total = 3.18	104.1	31.8	101.38

From the calculation, we have

$$\bar{X} = \frac{1}{n_1}\sum X = \frac{1}{10}\times 31.8 = 3.18,$$

$$\bar{Y} = \frac{1}{n_2}\sum Y = \frac{1}{10}\times 31.8 = 3.18$$

Thus,

$$S_1^2 = \frac{1}{n_1 - 1}\left[\sum X^2 - \frac{\left(\sum X\right)^2}{n_1}\right] = \frac{1}{9}\left[104.10 - \frac{(31.8)^2}{10}\right]$$

$$= \frac{1}{9}\times 2.98 = 0.33$$

$$S_2^2 = \frac{1}{n_2 - 1}\left[\sum Y^2 - \frac{\left(\sum Y\right)^2}{n_2}\right] = \frac{1}{9}\left[101.38 - \frac{(31.8)^2}{10}\right]$$

$$= \frac{1}{9}\times 0.26 = 0.03$$

Putting the values of S_1^2 and S_2^2 in equation (6), we have

$$F = \frac{0.33}{0.03} = 1.1$$

The critical (tabulated) value of test statistic F for right-tailed test corresponding $(n_1 - 1, n_2 - 1) = (9, 9)$ df at 1% level of significance is $F_{(n_1-1, n_2-1),\alpha} = F_{(9,9),0.01} = 5.35$.

Since calculated value of test statistic (= 1.1) is less than the critical value (= 5.35), that means calculated value of test statistic lies in non-rejection region so we do not reject the null hypothesis and reject the alternative hypothesis i.e. we reject the claim at 1% level of significance.

Thus, we conclude that samples provide us sufficient evidence against the claim so variance in LBCs of brand A is not greater than variance in LBCs of brand B.