BLOCK II:
TESTING OF HYPOTHESIS

Unit-1

Testing of Hypothesis-an Introduction

Unit Structure:

1.1 Introduction

In many situations, we have to make decisions consulting only the sample observations. That is, inferences regarding popution characteristics are deduced on the basis of sample survey or sample information. But, inferences drawn such way are not free from the risk of errors due to sampling. Since a sample is only a part of the population, it may not be able to truly represent the entire population. In such situations, the estimated value of the populaton characteristic calculated from the sample may differ from the true value of that population characteristic. This difference is generally termed as sampling error. The problem of decision making arises here when one has to make decision on the basis of sample results even after knowing about sampling error. The modern theory of probability plays a vital role in this kind of decision making and the branch of Statistics that helps us in arriving at the criterion for decisions is known as testing of hypothesis. The theory of testing of hypothesis, initiated by J. Neyman and E.S. Pearson, involves the employment of different Statistical tools to arrive at decisions in certain situations where there is an element of uncertainty on the basis of a fixed size sample.

1.2 Objectives

After going through this unit-

- understand the concept of testing of hypothesis
- define testing of hypothesis
- explain the steps involved in the process of hypothesis testing

1.3 Concept and background of Testing of Hypothesis :

The conventional research approach does not comprise all the units of the entire population concerned. Only a group of units or individuals (called sample) are selected to represent the population. Different symbols are employed to indicate the sample characteristics (called statistic) and the population characteristics (called parameters). For sample values the Roman alphabets like $\bar{x}$, p, s etc are used, whereas the Greek letters like μ, σ, π etc are used for population values. The difference between the sample value and the population value is assumed to be due to the sampling error. The volume of sampling error plays a vital role in estimating the population parameters. To estimate the average variability of the population value for all the possible sample we must consider only random samples. In other words, if random or probability sampling technique is used, we can estimate the sampling error.

In practical situation, it is seen that the distribution of sample statistic always follows the normal distribution. i. e. if we draw random sample of size n to get the sample mean $\bar{x}$ as an estimate of the population mean μ. The value of sample mean will vary from sample to sample. Out of values of different values of the mean, most of them would lie near the true value of the population mean μ. This shows that the distribution of sample mean approximately follows the normal distribution.

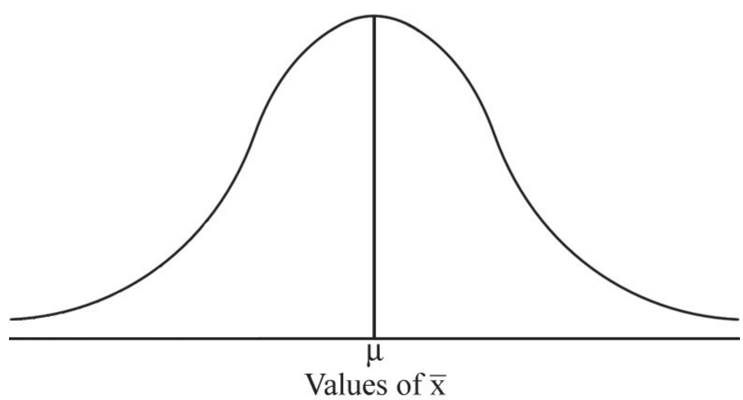
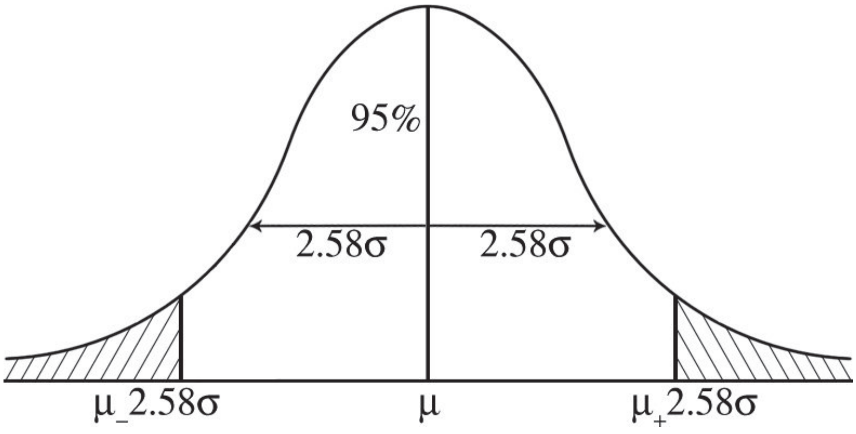


Fig : Distribution of Sample Means

To find the normal distribution uniquely for a specific distribution, we must know the value of avarage and the value of variation. Generally, arithmetic mean and standard deviation are taken as the parameters of the normal distribution. But, the distribution of sample means will not exactly follow a normal distribution. The closeness of the approximation mostly depends on the size of the sample. It is found that, the larger the sample size, the closer will be the approximation. If the size of the samples are 30 or more, the distribution of sample statistic approximately follows normal distribution. Considering the area property of normal distribution, we can say that for known value of μ and σ, $\mu \pm 1.96\sigma$ will contain 95% observations of the distribution. Similarly, $\mu \pm 2.58\sigma$ will contain 99% observations of the distribution.



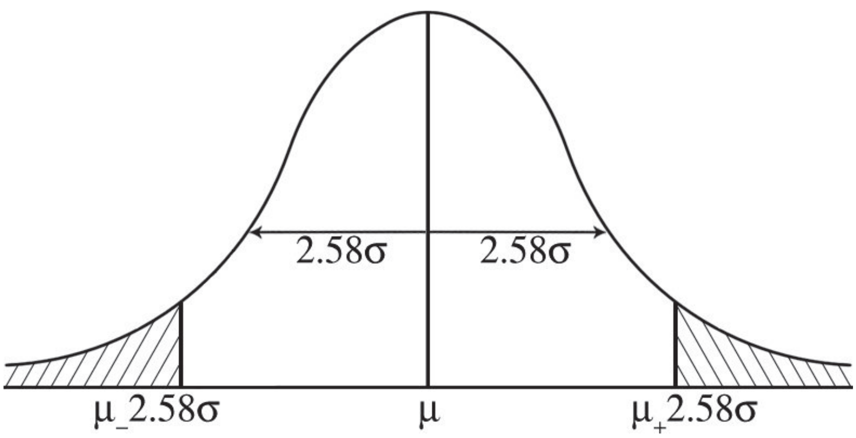


Fig : 95% and 99% intervals of the normal distribution measured about the population mean.

It should be noted that, the distribution of sample means we are considering is not based on original observations and hence we require the standard deviation of the sample means as our measure of variability. The standard deviation obtained from sample means is called the standard error of the mean and it is abbreviated as S.E. ($\overline{x}$). Similarly, we can find standard error of sample deviation, standard error of sample correlation coefficient, etc. But in order to calculate standard error of sample mean we need atleast two sample means which is true for most of the sample statistic. The standard error of sample means can be estimated indirectly by relating it to the standard deviation of the population (σ) and the size of the sample (n). The value of S.E. ($\overline{x}$) is obtained as $\frac{\sigma}{\sqrt{n}}$. Further, the value of σ is substituted with other estimated values derived from sample observations. In case of small samples (i.e. $n<30$) one must refer to some other distributions.

Statistical Decisions :

In most cases we are to take decisions about population on the basis of samples information. Such decisions are termed as statistical decisions. Great statistician J. Negman and E.S. Pearson initiated the concept of decision making by introducing the theory of Testing of Hypothesis.

The testing of hypothesis is a technique by which we test the validity of a given statement about a population. Generally, this is done based on a random sample drawn from the parent population. In other words, it is a rule or procedure for deciding whether to accept or reject the hypothesis within an optimum risk.

It is a branch of statistics which is primarily based on the concept of probability theory. Here sample data are used to draw conclusions about a population parameter. The process of testing of hypothesis is started by making a tentative assumption regarding the concerned parameter.

Steps involved in the Process of Testing of Hypothesis

The logical steps involved in the process of Testing of Hypothesis are as follows :

1. Making a Formal Statement : The first step in the process of Testing of Hypothesis is to set up a formal statement regarding the characteristic of the concerned population. The statement should be so made that it relates to the nature of the research problem. Such formal statements are also called Null Hypothesis. A complementary statement (called Alternative Hypothesis) should also be set up against the Null Hypothesis.

Then we collect sample information to find sample statistics and compare this value with the hypothetical value of the parameter. Suppose, we have made an assumption regarding population parameter. Now, to test the validity of this hypothetical value of the parameter we have to use sample data and find the value of sample statistics. Further, we calculate the difference between the hypothetical parametric value and the sample value. If the difference is minimum the probability of validity of the hypothetical value of the parameter will be higher.

In the procedure of hypothesis testing, we assume two hypothesis at a time at the beginning so that if one hypothesis is rejected, the other one will be accepted. Both the hypothesis are complementary to each other. Generally they are referred to as Null hypothesis and Alternative hypothesis. The null and alternative hypotheses are symbolised as H_0 and H_a or H_1 respectively.

2. Selection of the level of significance : The null hypothesis are tested for a specific value of level of significance. The level of significance is affected by number of factors like the size of the samples, difference between the sample means, the variability of measurements within samples, etc. In most cases, the level of significance is adopted either as 5% or 1% level.

3. Deciding the Sampling Distribution : After selection of the level of significance, we are to determine the appropriate sampling distribution for testing of hypothesis. The procedure of selecting the correct distribution are similar to those which are used in the context of estimation.

4. Random Sampling and computing the value of the Test Statistic : A sample should be drawn by using Random Sampling technique from the population under study. Further, required sample values are obtained from the drawn sample. The Test Statistic is calculated on the basis of these sample values.

5. Finding the Sampling Distribution : After calculating the value of the test statistic, we find the sampling distribution of it under the null hypothesis (i.e. considering the null hypothesis true).

6. Finding the Acceptance Region : After determining the sampling distribution of the test statistic, we find the acceptance region for it considering the type of the test-one tailed test or two tailed test.

7. Drawing of Conclusion : If the value of the calculated test statistic falls within the acceptance region (determined for a certain level of significance) we consider that the calculated value is insignificant and the test has provided no evidences against the null hypothesis.

If the calculated value falls outside the acceptance region (i.e. in the rejection region) then it will be significant and then we reject our null hypothesis at the level of significance adopted.

Check Your Progress

1. What do you mean by hypothesis testing?
2. Define level of significance
3. What are the errors in testing of hyposthesis ?

Flowchart for Testing of Hypothesis

State null as well as alternative hypothesis

↓

Specify the level of significance of the α value

↓

Decide the correct sampling distribution

↓

Draw a Random Sample and work out an appropriate value from sample data

↓

Obtain the value of the Test Statistic using sample values

↓

Find the sampling distribution of the Test Statistic

↓

Determine the Acceptance Region for given value of α

↓

Is the calculated value of the Test Statistic within the Acceptance Region

↓	↓
If Yes	If No
↓	↓
Accept the nul hypothesis	Reject the null hypothesis

1.4 Summing Up:

In testing of hypothesis, we test an assumption regarding a population paramenter. For testing such an assumption sample data are used and decisions are made about population on the basis of sample information. The theory of testing of hypothesis is based on the concept of level of significance and the test statistic. Finally decision rules are followed for interpretation.

1.5 Model Questions:

1. Describe the concept of testing of hypothesis.
2. Explain the procedure of testing of hypothesis briefly.
3. How do you set up a suitable significance level? Explain.
4. What are the decision rules to be followed during testing of hypothesis.
5. Distinguish between one tailed and two tailed test.

1.6 References and Suggested Readings

1. Hogg, Tanis,Rao; Probability and statistical inference; Pearson
2. Bhuyan K.C; Probability Distribution Theory and Statistical Inference; New Cemtral Book Agency(P) Ltd.
3. Elhance D.N, Elhance Veena, Agarwar B.M,;Fundamentals of Statistics; Kitab Mahal
4. Gupta S.C.,Kapoor V.K.; Fundamentals of Mathematical Statistics ; Sultan Chand & Sons

******

Unit-2
Types of Hypothesis

Unit Structure

2.1 Introduction

When a researcher observes some known facts and takes up a problem for analysis, he has to start with some assumptions regarding the population under study. Such assumption are often called hypothesis. The researcher has to proceed further on the basis of this hypothis and the facts that are already known. Formulation of an appropriate hypothesis is so crucial for any kind of research. Without having a valid hypothesis, investigations can't be carried out in right direction. Generally, such hypotheses are tested to evaluate possible intepretation about the population characteristics. Since the hypotheses reflect a generalised proposition regarding the population, we must take adaquate care in framing the hypothesis scientifically.

In this unit we will discuss different kinds of hypotheses with examples.

2.2 Objectives

After going through this unit, you will be able to-

- know the different types of hypothesis
- understand the different types of hypothesis with their definitions

2.3 Types of Hypothesis

There are several types of hypotheses used for research activities. They are discussed below–

2.3.1 Statistical Hypothesis

A hypothesis is a preconceived idea or assumption or statement about the nature of a population or about the value of its parameters. Such a hypothesis (which may or may not be true) about a population parameter that is testable on the basis of the evidence from a random sample is called a statistical hypothesis. The procedure by which we test the validity of a given statistical hypothesis is termed as Testing of Hypothesis or Tests of Hypothesis. Following are different types of hypothesis :

2.3.1.1 Null Hypothesis

The hypothesis to be tested is termed as Null Hypothesis and it is denoted by the symbol H_0. Some times null hypothesis is also called simply hypothesis. Null hypothesis asserts that there is no (significant) difference between the statistic and the population parameter under consideration. This hypothesis also considers that whatever observed difference is there is merely due to fluctuations of sampling from the same population. The no difference (or null) attitude before drawing a sample is the basis of null hypothesis. Great statistician R.A. Fisher defines this hypothesis as "Null Hypothesis is the hypothesis which is tested for possible rejection under the assumption that it is true." If we want to decide, for example, if a coin is unbiased or not, we formulate the null hypothesis that the coin is unbiased, i.e. H_0:p=0.5, where p is the probability of getting a head in tossing a coin. Similarly, if we want to decide whether one procedure is better than the other, we consider the hypothesis that there is no difference between the procedures and any observed difference is merely due to fluctuations in sampling from the same population. Again, if we are to test if the mean of a particular population (i.e. μ) has a specified value or not then the null hypothesis H_0 will be that the population mean μ has a specified value μ_0, i.e., the null hypothesis H_0 will be stated as $H_0: \mu = \mu_0$, i.e. there is no difference between the population mean μ and the specified value μ_0. In an another context if we are to test whether a particular medicine is effective in curing a particular disease then our null hypothesis will be that the medicine is not effective which means that taking the medicine and not taking the medicine makes no difference.

2.3.1.2 Alternative Hypothesis

Any hypothesis other than null hypothesis is known as Alternative Hypothesis and it is denoted by the symbol H_1. For example, if the null hypothesis is $H_o : \mu = \mu_o$, the possible alternative hypothesis may be as follows :

$H_1 : \mu \neq \mu_0$

or, $H_1 : \mu > \mu_o$

or, $H_1 : \mu < \mu_0$

In any decision procedure we are to formulate both the null hypothesis (Ho) and alternative hypothesis (H_1) and they are such that when one is true, the other one is false and coversely.

Check Your Progress

1. What do you mean by Statistical hypothesis ?
2. How are the null hypothesis constructed?
3. Distinguish between null and alternative hypothesis.

2.3.2 Simple and Composite Hypothesis :

Any statistical hypothesis that completely specifies the distribution of the concerned population is called a simple hypothesis. Otherwise if the hypothesis doesnot completely specify the concerned distribution, then it is known as composite hypothesis. For example, for a normal distribution N (μ, σ^2) with σ^2 known, the two hypothesis Ho : $\mu = 50$, against the alternative $H_1 : \mu = 20$ are simple hypotheses.

But, usually the alternative hypothesis H_1 is a composite one which may be stated as–

$H_1 : \mu \neq 50$

or $H_1 : \mu < 50$

or $H_1 : \mu > 50$

depending upon the situation and each of these does not specifically state the value of the population mean μ. Such hypotheses are termed as composite hypothesis as they are not specifying the value of the population mean μ.

Let us consider an another example. Suppose sampling is done from a normal distribution with mean μ and variance σ^2, where the value of σ^2 is unknown. In this case,

Ho : $\mu=0$ and $\mu=5$

is a simple hypothesis as it completely specifies the concerned distribution.

But, each of the following hypotheses is a composite hypothesis.

1. $H_0 : \mu=0$
2. $H_0 : \sigma=\sigma_0$
3. $H : \mu \neq \mu_0$ and $\sigma=\sigma_0$
4. $H : \mu_1 \leq \mu \leq \mu_2$ and $\sigma=\sigma_0$

All the above ideas can be put in a mathematical formalism as follows :

Let the distribution of the population be given by the p.d.f. f(x; θ_1, θ_2, ..., θ_k) with k unknown parametes θ_1, θ_2, ..., θ_k. The set of all values of the parameters constitute a subset θ of R^k, called the parametric space. Every element of θ is a k-vector called a parametric point.

A statistical hypothesis H is then any assumption of the form H : $\theta \in \Omega \subset \theta$.

If Ω is singleton, then H is called a simple hypothesis. If Ω is not a singleton, then H is called a composite hypothesis.

2.4 Summing Up

There can be several types of hypothesis. But in the context of testing of hypothesis we use only the Statistical Hypothesis. The two hypothesis in a statistical test are normally termed as - Null hypothesis and Alternative hypothesis. Both the hypothesis are very useful tool in testing the significance of difference between the statistics and the population parameter under consideration. After testing the null hypothesis, we make decision whether to accept or reject the null hypothesis.

2.5 Model Questions

1. What is Statistical hypothesis? Describe briefly
2. State the similarities and differences between null and alternative hypothesis.

3. How to construct the null and alternative hypothesis. Describe with example

4. Distinguish between simple and composite hypothesis

5. Define Statistical Hypothesis. State the different types of statistical hypothesis.

2.6 References and Suggested Reading

1. Hogg, Tanis,Rao; Probability and statistical inference; Pearson

2. Bhuyan K.C; Probability Distribution Theory and Statistical Inference; New Cemtral Book Agency(P) Ltd.

3. Elhance D.N, Elhance Veena, Agarwar B.M,;Fundamentals of Statistics; Kitab Mahal

4. Gupta S.C.,Kapoor V.K.; Fundamentals of Mathematical Statistics ; Sultan Chand & Sons

******

Unit-3

Errors in Testing of Hypothesis

Unit Structure:

3.1 Introduction

The decision of acceptance or rejection of a null hypothesis regarding a population parameter is made on the basis of sample drawn from the parent population. That is why an element of risk of taking wrong decisions, i.e. an element of uncertainly is always involved in making such decisions. Such probable errors are, therefore, categorised and tried to keep minimum during the process of testing of hypothesis. The two distinct types of errors are Type I error and Type II error. In deciding whether to accept or reject a null hypothesis, both the errors play a vital role. In this unit we shall discuss the definitions of these errors and related terms used in testing of hypothesis.

3.2 Objectives

After going through this unit, you will be able to-

- know the definitions of different types of errors,
- understand the concept of level of significance and critical region,
- explain the distinction between parametric and non parametric test.

3.3 Type I error and Type II error :

When we reject a true null hypothesis, the error is called as the type I error. The probability of committing such an error is denoted by the symbol α. Thus,

$$\alpha = \Pr\text{ (Type I error)}$$
$$= \Pr\text{ (Rejecting Ho when Ho is true)}$$

Again, when we accept a false null hypothesis, the error is called as the type II error. The probability of committing this error is denoted by the symbol β. Thus,

$$\beta = \Pr\text{ (Type II error)}$$
$$= \Pr\text{ (Accepting Ho when Ho is false)}$$

In Statistical Quality Control (SQC) terminology Type I error is known as producer's risk whereas Type II error is known as concumer's risk.

Executive will take the decision if it is worthwhile to have a larger proportion of Type I error than that of Type II error. It would depend on the relative losses incurred by comitting Type I and Type II errors respectively. However, in general it is more risky to accept a false hypothesis (or commit Type II error) than to reject a true hypothesis (or commit Type I error).

The distinction between these two types of errors can be made clear by an example. Let us assume that the difference between the two population mean in actually zero. If our test of significance when applied to the sample means leads us to believe that the difference in population means is significant, we commit a Type I error. On the other hand, suppose there is true difference between the two population means. Now if our test of significance leads to the conclusion "not significant", we commit a Type II error. We thus find ourselves in the situation which may be described by the following table :

	Accept H_0	Reject H_0
H_0 is true	Correct Decision	Type I error
H_0 is false	Type II error	Correct Decision

While testing hypothesis the aim is to reduce both the types of error, i.e. Type I and Type II error. But due to fixed sample size, it is not possible to control both the errors simultaneously. There is a trade-off between these two types errors : the probability of committing one type of error can only be reduced if we are willing to increase the probability of committing the other type of error. In order to get a low β, we will have to put up with a high α. To deal with this trade-off in business situations, managers decide the appropriate level of significance by examining the costs or penalties attached to both types of errors.

In practical situation, it is more dangerous to accept a false hypothesis (Type II error) than to reject a correct one (Type I error). That is why, we keep the probability of committing Type I error at a certain level. This level of type I evolves the concept of level of significance.

3.4 Level of Significance

The maximum probability with which a true null hypothesis (Ho) is rejected (i.e. committing a type I error) in the test procedure is termed as the Level of Significance. Generally, it is denoted by the symbol of type I error, i.e. by α. Thus,

Level of Significance = Pr (Type I error)

= Pr (Rejecting Ho when Ho is true)

= α

Usually the value of α is taken as either 5% (i.e. α=0.05) or 1% (i.e. α=0.01).

α = 5%=0.05 means that in 5% of the total number of samples, each of the same fixed size, that can be drawn from a population we are likely to reject a correct Ho. In other words, when we consider level of significance i.e. α=5%, there are 5 cases in 100 that we would reject the correct null hypothesis. This implies that we are 95% confident that we have made the right decision in rejecting the null hypothesis and accepting the alternative hypothesis. Similarly, we can interpret other levels of significance.

The concept of level of significance also uses the theory of area property of Normal distribution. As we know that for all random samples of size greater than or equal to 30 that can be drawn from a population with mean μ_0, in case of 95% samples the range $\overline{x} \pm 1.96 \, s/\sqrt{n}$ will contain the population mean μ_0 where s is the sample estimate of the population standard deviation σ. Therefore, for any random sample, we can be 95%

sure that the range $\bar{x} \pm 1.96\ {}^{s}\!/\!_{\sqrt{n}}$ contains the population mean μ_0 where $\bar{x}$ is the mean of the random sample. In case of the remaining 5% samples only, the range $\bar{x} \pm 1.96\ {}^{s}\!/\!_{\sqrt{n}}$ does not contain the population mean μ_0.

$\bar{x} \pm 1.96\ {}^{s}\!/\!_{\sqrt{n}}$ is called the 95% confidence interval. Hence if for a sample drawn randomly from the population, the range $\bar{x} \pm 1.96\ {}^{s}\!/\!_{\sqrt{n}}$ does not contain the population mean μ_0. Hence, we reject our null hypothesis H_0 which states that population mean $\mu = \mu_0$. But in case of 5% of the samples drawn from a population with mean μ_0, the ranges $\bar{x} \pm 1.96\ {}^{s}\!/\!_{\sqrt{n}}$ does not contain μ_0.

As such we can be only 95% confident that our decision of rejecting the null hypothesis is correct whenever $\bar{x} \pm 1.96\ {}^{s}\!/\!_{\sqrt{n}}$ does not contain μ_0 for a sample drawn randomly from the population. Although the difference between $\bar{x}$ and μ_0 may be due to fluctuations of sampling we have rejected the null hypothesis. Thus Type I error that is induced in our decision is 5% which we call the level of significance.

Let us suppose that under the hypothesis considered the sample statistic t approximately follows normal distribution with mean =E(t) and standard deviation $= \sigma_t$ =S.E.(t). Then the variable Z defined as follows

$$Z = \frac{t - E(t)}{\sigma_t}$$

is called the standard normal variate which again follows normal distribution with mean=0 and standard deviation=1. The variable Z is termed as standard normal test statistic or the Z-score.

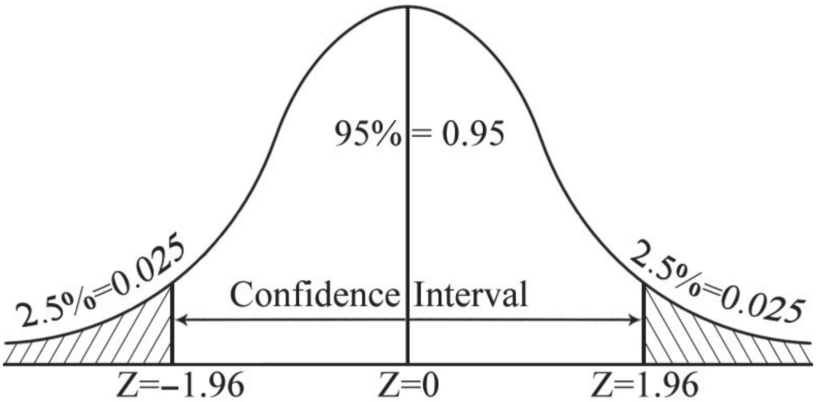


Fig : Two-tailed test

In the above figure we can see that if the test statistic of a sample statistic lies between –1.96 and 1.96 then we are 95% confident that the null hypothesis is true. This happens because the area under the normal curve between Z=–1.96 and Z=1.96 is 0.95 which is 95% of the total area under the Z curve. It should be noted that total area under Z curve is unity or I which is 100%.

However, the probability of being Z scores of a statistic outside the range –1.96 and 1.96 would be 0.05 if the given hypothesis were true. We would then say that this Z score differed significantly from what be expected under the hypothis, and we would be inclined to reject the hypothesis. The total shaded area 0.025+0.025=0.05 is the level of significance of the test which represents the probability of being wrong in rejecting the hypothesis, i.e., the probability of committing Type I error. Therefore, we say that the hypothesis is rejected at 0.05 (i.e. 5%) level of significance or that the Z-score of the given sample statistic is significant at 0.05 or 5% level of significance.

3.5 Critical Region

The set of values of the test statistic that lead to the rejection of the hypothesis is called the critical region or the rejection region or the region of significance. On the other hand, the set of values which lead to the acceptance of the hypothesis is termed as the acceptance region. Generally, the critical region corresponds to the predetermined value of the level of significance α and the acceptance region corresponds to $1-\alpha$.

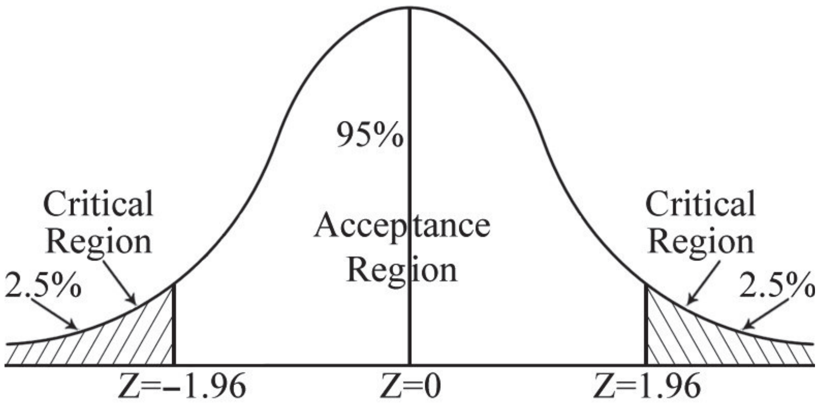


Fig : Two-tailed test

In other words, the set of Z scores of a sample statistic outside the range –1.96 to 1.96 constitutes what is called the critical region or the rejection region of the hypothesis. (as shown in fig.) In fact, the critical region is the region under the standard normal curve beyond the range –1.96 to 1.96. Again, the Z scores falling inside the range –1.96 to 1.96 represents the area under the standard normal curve (or Z curve) between the two ordinates at $Z = \pm 1.96$ and this is called the Acceptance Region or the region of non significance. Similarly, we can define critical region at any other level of significance, say, 1%, 10% etc.

Check Your Progress

1. What are the different types of errors in testing of hypothesis ?
2. What are producer's risk and consumer's risk ?
3. Define Critical Region .

3.6 One-Tailed and Two-Tailed Test

In the procedure of testing hypothesis, the test will be a one-tailed or two-tailed depends entirely on the nature of the alternative hypothesis H_1. Suppose, the null hypothesis to be tested is, Ho : $\mu = \mu_o$. In this case, if the alternative hypothesis adopted is $H_1 : \mu > \mu_o$ or $H_1 : \mu < \mu_o$, then the test is called a one-tailed test. More specifically, for $H_1 : \mu > \mu_o$ the test is called a right-tailed test. Otherwise, for $H_1: \mu < \mu_o$ it will be termed as a left-tailed test. In this case, the critical region for the interval $(-\alpha, -1.645)$ is shown as the shaded area in the following figure :

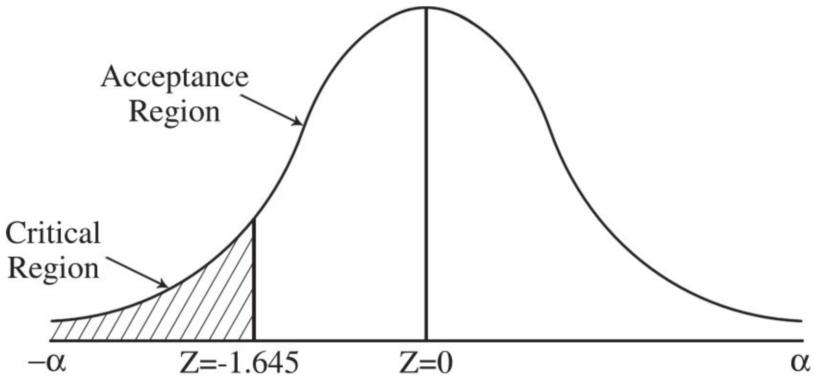


Fig : Left-tailed test

Again, for the null hypothesis Ho: $\mu = \mu_o$, if the alternative hypothesis is H_1: $\mu \neq \mu_o$, then the test will be called a two-tailed test, Here, for α =0.05 and for the standard normal variate Z.

Suppose we have considered the level of significance α =0.05. Then for a standard normal variate Z, under right tailed test we have, $P[Z \leq 1.645] = 1 - 0.05 = 0.95$. In this case we reject Ho if the computed value of Z from a sample lies outside the interval $(1.645, \alpha)$ and accept it otherwise. The region for the interval $(1.645, \alpha)$ is known as the critical region and is shown in the following figure :

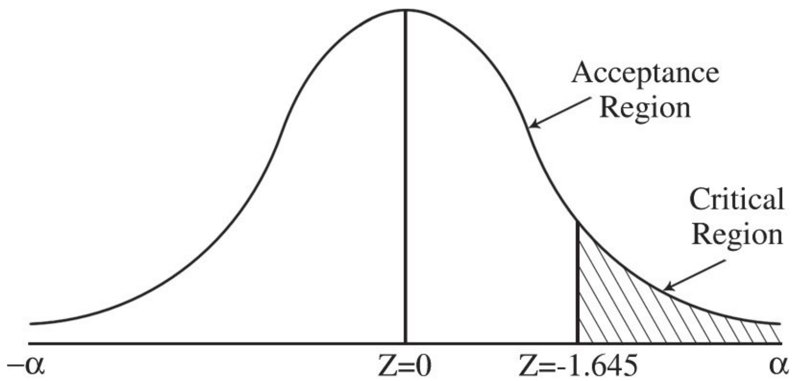


Fig : Right-tailed test

Similarly, for a left-tailed test we have, $P[-1.645 \leq Z) = 1 - 0.05 = 0.95$ we have.

$P[-1.96 \leq Z < 1.96] = 1 - 0.05 = 0.95$. For such a test we reject Ho if the computed value of Z from a sample lies outside the interval (–1.96, 1.96). This the critical region or rejection rejion will lie both side of the tail as shown in the following figure :

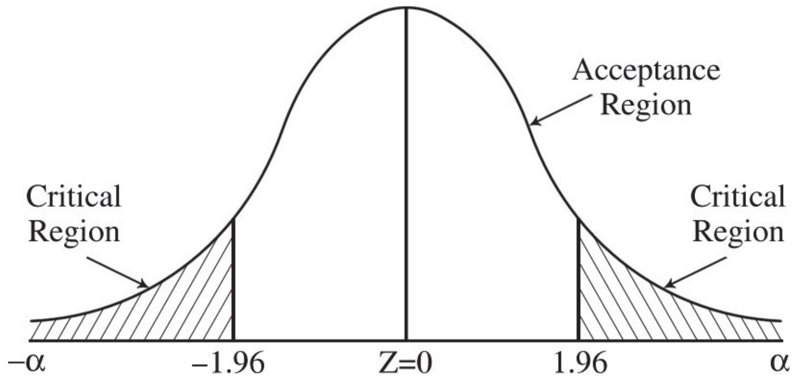


Fig : Two-tailed test

Critical values of Z for various levels of significance are summarized in the following table :

Level of Significance	1%	5%
Critical value for two tailed test	$\lvert Z_\alpha \rvert = 2.58$	$\lvert Z_\alpha \rvert = 1.96$
Critical value for one tailed test	$\lvert Z_\alpha \rvert = 2.33$	$\lvert Z_\alpha \rvert = 1.645$

3.7 Test of Hypothesis

There are different types of tests of hypotheses for the purpose of testing the hypotheses. The tests of hypotheses can be classified into two categories. They are :

1. Parametric tests or standard tests of hypotheses
2. Non-Parametric tests of distribution-free test of hypotheses

Parametric Test

Any statistical test that makes assumption about the parameters (defining properties) of the parent population distribution(s) from which we draw samples is usually called parametric test. The typical assumptions made are:

- Normality : Data have drawn from a normal distribution
- Homogeneity of variances : Data drawn from different populations have the same variance.
- Linearity : Data have a linear relationship
- Independence : Data are independent each other

Almost all of the most commonly used statistical tests are based on the above assumptions. The following are some important parametric tests :

- Z-test or Large Sample test
- t-test
- χ^2 –test
- F-test

Non-Parametric Test

In contrast to parametric tests, non-parametric tests do not require any assumptions about the parameters or about the nature of population. In other words, a statistical test used in the case of non-metric independent variabables, is called non-parametric test. When an investigator has no idea regarding the population parameter and the data concerned are strongly non-normal, such tests are adopted to test statistical hypothesis. Moreover, most of the non-parametric tests are applicable to data measured in an ordinal or nominal scale. Following are some of the important non parametric tests:

- The Runs test for Randomness
- The Median test for Randomness
- Wilcoxon Signed Rank test
- The Matched-Pairs Sign test
- Wilcoxon Matched-Pairs Signed Rank-sum test
- Mann-whitney Wilcoxon test

3.10 Model Questions

1. What are the two types of errors associated with testing of hypothesis. Explain them with examples
2. Define and discuss the Level of Significance.
3. Distinguish between one tailed and two tailed test.
4. What do you mean by Level of Significance? Explain
5. Clarify the underlying concept of Critical Region and Acceptance Region.

3.8 Summing Up

In testing of hypothesis, it is aimed to reduce both the types of error,i.e. Type-I and Type -II. But, it is not possible to reduce both the errors together due to fixed samples size. The reduction of one type of error leads to an increase in the other types of error. As the Type -I error is considered to be more dangerous ,we keep the probability of committing Type-I error at a certain level , which is called the level of significance. Generally , the level of significance is denoted by α . Further, critical values of the statistic for one -tailed or two- tailed test for a certain level of significance are obtained for any decision procedure. In this process of obtaining the critical values , the concept of Critical Region and Acceptance Region arises.

3.9 References and Suggested Reading

1. Hogg, Tanis,Rao; Probability and statistical inference; Pearson
2. Bhuyan K.C; Probability Distribution Theory and Statistical Inference; New Cemtral Book Agency(P) Ltd.
3. Elhance D.N, Elhance Veena, Agarwar B.M,;Fundamentals of Statistics; Kitab Mahal
4. Gupta S.C.,Kapoor V.K.; Fundamentals of Mathematical Statistics ; Sultan Chand & Sons
5. Bhardwaj R.S.; Business Statistics ; Excel Books
6. Choudhury L, Sarma R, Deka M, Gogoi S.J.; An Introduction to Statistics; L. Choudhury.

******

Unit-4
Parametric Test

Unit Structure:

4.1 Introductions

Parametric tests are usually based on certain properties of the parent population from which samples are to be drawn. The basic assumptions are–

- Normality : Under this assumption we assume that the data drawn from a normal distribution.
- Homoscedasticity : In case of Parametric tests involving more than one population we assume that each population has equal variance.
- Independence : Here we assume that data in each group are randomly and independently drawn from the population.

The following are some important parametric tests :

- Large Sample Test or Z-test
- Student's t-test or t-test
- Snedecore's F-test or F-test
- Chi-square test

All these tests are based on the assumption of normality, i.e. the source of data is considered to be normally distributed. In some cases, the population may not follow normal distribution, yet the tests are applicable on account of the fact that most of the samples drawn and their sampling distributions closely approach to normal distribution.

4.2 Objectives

After going through this unit, you will be able to-

- understand the concept of parametric test
- explain the definitions of different parametric test
- describe the technique of solving practical problems using different parametric test

4.3 Large Sample Test or Z test

This test is based on the normal probability distribution and is used for large samples i.e. the samples that are greater than equl to 30. The assumptions made in such test are

1. The sampling distribution of the sample statistic follows normal distribution, and
2. the sample values are sufficiently close to the population value and hence can be used in its place for calculating the estimate of the standard error.

If t is a statistic calculated from any sample, then the test statistic under this test is given as :

$$Z = \frac{t - E(t)}{S.E.(t)} \sim N(0,1)$$

Some of different types of Z-test are discussed below :

Z-test for a specified mean

Suppose a sample is drawn from a normal population. To test the null hypothesis, Ho: $\mu \neq \mu_0$ (specified) against the alternative H_1 $\mu \neq \mu_0$ (two-tailed test) the test statistic used is

$$Z_{cal} = \left| \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} \right| \sim N(0,1)$$

The decision rule would be :

Reject Ho at 5% (say) level of significance if $Z_{cal} > Z_{tab}$ i.e. 1.96. Otherwise, there is no evidence against Ho at this level significance.

The tabulated value or critical value of Z i.e. Z_{tab} may be different according to the nature of the test (two-tailed test or one-tailed test) and the level of significance considered during the test procedure.

Example : The mean life of 100 electric bulbs produced by a company is found to be 1570 hours with a standard deviation of 120 hours. If μ is the mean life time of all the bulbs produced by the company, test whether μ=1600 hours or not. Test your hypothesis at 5% level of significance.

Solution : Here the null hypothesis to be tested is

Ho: μ =1600 hours

against the alternative

$H_1: \mu \neq 1600$ hours

We are given,

n= sample size = 100

$\bar{x}$ = sample mean = 1570 hours

s= sample s.d.=120 hours

Now, under the null hypothesis, the test statistic is

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

$$= \left| \frac{1570 - 1600}{120 / \sqrt{100}} \right|$$

, considering sample s.d. as the estimated value of population s.d. δ

$$= \frac{30}{120/10} = \frac{30}{12} = 2.5$$

The table value of Z at 5% level of significance is 1.96.

Conclusion : Since the calculated Z>1.96, the tabulated value of Z, we may reject the null hypothesis at 5% level of significance and conclude that probably the population mean is different from 1600 hours.

Z-test for specified proportion

Let us consider a random sample of size $n(n \geq 30)$ out of which x number of observations possessing a certain attribute.

Now, $p = \frac{x}{n}$ is the sample proportion of observations possessing the attribute.

If we want to test the hypothesis that the population proportion P has a specified value P_o, the null hypothesis to be tested would be

H_o: P= P_o

against the alternative

$H_1 : P \neq P_o$

The test statistic under this null hypothesis is given as

$$Z = \left| \frac{P - E(p)}{S.E.(p)} \right| = \left| \frac{p - P}{\sqrt{\frac{PQ}{n}}} \right| \sim N(0,1)$$

Where Q=1–P

Comparing the calculated value of Z with the tabulated value we make the decisions for different types test and for different level of significance as in the previous case.

Example : A coin is tossed 900 times and heads appear 490 time. Does the result test the hypothesis that the coin is unbiased? Test the hypothesis at 1% level of significance.

Solution :

Let P denotes the population proportion of heads.

Here the null hypothesis to be tested is,

H_o : the coin is unbiased, i.e. $P=\frac{1}{2}=0.5$

against the alternative.

H_1 : the coin is biased, i.e. $P \neq 0.5$

We are given,

n= sample size = 900

x= No. of heads appeared = 490

p= sample proportion of heads $= \frac{490}{900} = 0.54$

Now, considering the null hypothesis to be true (or under the null hypothesis), the test statistic is

$$|Z| = \left|\frac{p - E(p)}{S.E.\,of\,(p)}\right|$$

$$= \left|\frac{p - P}{\sqrt{\frac{P(1-P)}{n}}}\right| \sim N(0,1)$$

$$= \left|\frac{0.54 - 0.5}{\sqrt{\frac{0.5(1-0.5)}{900}}}\right| = \frac{0.04}{\sqrt{0.00278}}$$

=2.39

The table value of Z at 1% level of significance (for two tailed test) is 2.58 i.e. $Z_{0.01}$=2.58.

Conclusion : Since the calculated value of Z=2.39<2.58, the table value of Z, we may accept the null hypothesis at 1% level of significance and conclude that probably the coin is unbiased.

Z-test for difference of two Means :

Let a random sample of size n_1 is drawn from a population having population mean μ_1 and standard deviation σ_1. Again, let an another sample of size n_2 is drawn from a population having population mean μ_2 and same standard deviation σ_2. If $\overline{x_1}$ and $\overline{x_2}$ are the sample means of the two random samples respectively and we want to test the null hypothesis, Ho: $\mu_1 = \mu_2$, against the alternative H_1: $\mu_1 \neq \mu_2$, the test statistic under the null hypothesis is given as

$$Z = \left|\frac{\left(\overline{x_1} - \overline{x_2}\right) - E\left(\overline{x_1} - \overline{x_2}\right)}{S.E.\left(\overline{x_1} - \overline{x_2}\right)}\right| \sim N(0,1)$$

$$= \left| \frac{\overline{x_1} - \overline{x_2}}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \right| \text{ (after simplification)}$$

The decision rule is same as the previous cases of Z test.

Note : (i) If $\sigma_1^2 = \sigma_2^2 = \sigma^2$ (say) then

$$Z = \left| \frac{\overline{x_1} - \overline{x_2}}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \right| \sim N(0,1)$$

(ii) If σ_1^2 and σ_2^2 are unknown they can be replaced by the sampling variance s_1^2 and s_2^2 respectively (since n_1 and n_2 are large). In that case

$$Z = \left| \frac{\overline{x_1} - \overline{x_2}}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \right| \sim N(0,1)$$

Example : In order to compare the scooters of two well-known companies, say, A and B in respect of efficiency in petrol mileage, a random sample of 50 scooters has been selected from each company and following results are obtained :

Company A : Mean mileage = 32.5 mile/litre

S.D. = 4.5 mile/litre

Company B : Mean mileage = 34.8 mile/litre

S.D. = 5.6 mile/litre

Test whether there is any significant difference between the two brands in respect of mean mileage.

Solution : Let μ_1 and μ_2 denote the mean mileage (population mean) of all scooters of companies A and B respectively.

Here the null hypothesis to be tested is

Ho: $\mu_1 = \mu_2$

against the alternative

$H_1 = \mu_1 \neq \mu_2$

We are given,

Mean mileage of company A $= \overline{x}_1 = 32.5$

S.D. of company A $= s_1 = 4.5$

Mean mileage of company B $= \overline{x}_2 = 34.8$

S.D. of company B $= s_2 = 5.6$

Sample size of company A = sample size of Company B, i.e. $n_1 = n_2 = 50$

Under the null hypothesis Ho, the test statistic is

$$|Z| = \left|\frac{\overline{x}_1 - \overline{x}_2}{S.E.(\overline{x}_1 - \overline{x}_2)}\right| = \left|\frac{\overline{x}_1 - \overline{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}\right| \sim N(0,1)$$

$$= \left|\frac{\overline{x}_1 - \overline{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}\right|$$, considering sample S.D.s are estimated value of the population S.D.s

$$= \left|\frac{32.5 - 34.8}{\sqrt{\frac{(4.5)^2}{50} + \frac{(5.6)^2}{50}}}\right| = \frac{2.3}{1.016} = 2.26$$

The table value of Z at 5% level of significance is 1.96.

Conclusion : Since the calculated value of Z=2.26>1.96, the table value of Z, we may reject the null hypothesis at 5% level of significance and conclude that probably the mean mileage of the two brands of scooters are different.

Z-test for difference of two Proportions :

Let n_1 and n_2 be the number of individuals in two samples selected from populations I and II respectively. Let x_1 and x_2 be the number of individuals possessing a certain attribute in the samples drawn.

Let, P_1 and P_2 be the population proportion of individuals possessing the certain attribute in population I and II respectively.

Now, suppose we want to test the null hypothesis, $H_o: P_1 = P_2 = P$ (say) against the alternative $H_1: P_1 \neq P_2$.

Under the null hypothesis the test statistic is given as

$$Z = \left| \frac{p_1 - p_2}{S.E.(p_1 - p_2)} \right|, \text{ where } p_1 = \text{sample proportion of the first sample}$$

$$= \frac{x_1}{n_1}$$

p_2 = sample proportion of the second sample

$$= \frac{x_2}{n_2}$$

$$= \left| \frac{p_1 - p_2}{\sqrt{PQ\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \right| \sim N(0,1)$$

Where Q=1–P

The conclusion is done in usual manner of comparing the calculated and tabulated value of Z.

Note : If the value of P is not known, it is estimated by

$$\hat{P} = p = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2}$$

and $\hat{Q} = 1 - p = q$

In this case,

$$Z=\left|\frac{p_1-p_2}{\sqrt{pq\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}\right|\sim N(0,1)$$

Example : A company has the head office at Kolkata and a branch at Guwahati. The personnel director wanted to know if the workers at the two places would like the introduction of a new plan of work and a survey was conducted for the purpose. Out of a sample of 500 workers at Kolkata 62% favoured the new plan. At Guwahati out of a sample of 400 workers 41% were against the new plan. Is there a significant difference between the two group in their attitude towards the new plan at 5% level of significance?

Solution : Let P_1 and P_2 be the population proportion in Kolkata and Guwahati respectively who prefer the new plan.

Here the null hypothesis to be tested is,

$Ho{:}P_1{=}P_2$

against the alternative

$H_1{:}\ P_1\neq P_2$

We are given,

n_1= sample size of workers of Kolkata=500

n_2= sample size of workers of Guwahati= 400

p_1= sample proportion of workers of Kolkata who favours the new plan = 62%=0.62

p_2= sample propertion of workers of Guwahati who favours the new plan = (100–41%)

= 59%=0.59

$$\text{Now, } P=\frac{n_1p_1+n_2p_2}{n_1+n_1}$$

$$=\frac{500\times0.62+400\times0.59}{500+400}=\frac{546}{900}=0.607$$

$\therefore$ q = 1–p=1–0.6079=0.393

Now, under the null hypothesis Ho, the test statistic is

$$|z| = \left|\frac{p_1 - p_2}{S.E.(p_1 - p_2)}\right| = \left|\frac{p_1 - p_2}{\sqrt{pq\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}\right| \sim N(0,1)$$

$$= \left|\frac{0.62 - 0.59}{\sqrt{0.607 \times 0.393\left(\frac{1}{500} + \frac{1}{400}\right)}}\right|$$

$$= \frac{0.03}{0.0327} = 0.917$$

The table value of Z at 5% level of significance is 1.96.

Conclusion : Since the calculated value of Z=0.917<1.96, the table value of Z, we may accept the null hypothesis at 5% level of significance and conclude that probably there is no significant difference between the two groups of workers in their attitude towards the new plan.

Check Your Progress

1. What is a parametric test?
2. A random sample of 200 tins of coconut oil gave an average weight of 4.95kgs with a standard deviation of 0.21kg. Do you accept the hypothesis of net weight of 5 kgs per tin at 5% level of significance?
3. A random sample of 1000 workers from Tamil Nadu shows that their mean wages are Rs. 47.00 per day with a standard deviation of Rs. 28.00. A random sample of 1500 workers from Assam give a mean wage of Rs. 49.00 per day with a standard deviation of Rs. 40.00. Is there a significant difference between the mean daily wages of the workers in the two regions? Conduct your test at α =0.05.
4. What is standard normal test? State uses of this test.

Student t-test (small sample Test)

In case of small samples i.e. when n<30, we have to use the concept of a new distribution known as Students-t distribution. Here population is considered to follow normal distribution whose S.D. σ is not known. The distribution is basically based on the concept of degrees of freedom.

The degrees of freedom of a set of observations is the number of values which could be chosen independently with the specification of the system. For example, if a variable x assumes n different values and k different linear restrictions are imposed on the values of x, then the degrees of freedom, denoted by γ, will be γ=n–k.

In general, k is considered to be 1 in most cases due to the linear restriction. In that case, degrees of freedom γ=n–1.

It should be noted that the critical value (or table value) of the statistic t at a specified level of significance vary with the degrees of freedom of the distribution.

Application of t-distribution :

Student-t distribution has a large number of applications in statistics some of which are enumerated below :

(i) To test for a single population mean.

(ii) To test the difference between two population means.

(iii) To test the significance of observed correlation co-efficient.

(iv) To test the significance of observed regression co-efficient.

(v) To test the significance of observed partial correlation coefficient.

t-test for the single population mean :

Suppose a random sample of size n (n<3) is drawn from a Normal population with unknown mean μ and variance σ^2 and $\overline{x}$ is the sample mean. Now, to test the null hypothesis that the population mean μ has a specified mean μ_0 when population S.D. σ is unknown, i.e.

Ho: $\mu = \mu_0$

against the alternative

$H_1 = \mu \neq \mu_0$

the test statistic under the null hypothesis is given by,

$$t = \left| \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n-1}}} \right| \sim t_{0.0\alpha,(n-1)}$$

where $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} xi$

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})$$

Which is an unbiased estimate of the population variance σ^2

We reject Ho if calculated value of t greater than the tabulated value of t. Otherwise we accept the Ho at the specified level of significance.

Example : A machine is designed to produce insulating washers for electrical devices of average thickness of 0.025cm. A random sample of 10 washers was found to have an average thickness of 0.024cm. with a S.D. of 0.002cm. Test the significance of deviation.

(Given, table value of t at 5% level of significance for 9 degrees of freedom is 2.262 i.e. $t_{0.05,9} = 2.262$)

Solution : Let μ denotes the average thickness of population of washers.

Here the null hypothesis to be tested is,

H_0: μ=0.025cm

against the alternative

H_1: $\mu \neq$ 0.025cm

We are given,

n=sample size=10

$\bar{x}$ = sample mean = 0.024cm

s=sample s.d.= 0.002cm

Now, under the null hypothesis the test statistic is

$$|t| = \left| \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n-1}}} \right| = \left| \frac{0.024 - 0.025}{\frac{0.002}{\sqrt{10-1}}} \right|$$

$$= \frac{0.003}{0.002} = \frac{3}{2} = 1.5$$

Degrees of freedom=n–1=10–1=9

The table value of t at 5% level of significance for 9 d.f. i.e. $t_{.05,9} = 2.262$

Conclusion : Since the calculated value of t=1.5<2.262, the table value of t, we may accept the null hypothesis at 5% level of significance and conclude that probably the deviation is not significant.

t-test for the difference of two population Means

Let $\overline{x}_1$ be the sample mean of a random sample of size n_1(<30) drawn from a population with mean μ_1 and vareance σ_1^2.

Let $\overline{x}_2$ be an another sample mean of a random sample of size n_2(<3) drawn from a population with mean μ_2 and variance σ_2^2.

If we want to test the null hypothesis that the two population means are equal i.e. Ho: $\mu_1 = \mu_2$

against the alternative

H_1: $\mu_1 \neq \mu_2$

then the test statistic (assuming that $\sigma_1^2 = \sigma_2^2 = \sigma^2$ (say)) under this H_0 is given by

$$t = \left| \frac{\overline{x}_1 - \overline{x}_2}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \right| \sim t_{0.0\alpha,(n_1+n_2-2)}$$

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

and s_1^2, s_2^2 are the sample variances of the first and second sample respectively

Here, s^2 is considered as an unbiased estimate of σ^2, i.e. $E(s^2)=\sigma^2$

The table value of t is obtained for (n_1+n_2-2) d.f. for α% of level of significance.

If the calculated value of t is greater than the tabulated one, we reject the null hypothesis at α% level of significance. Otherwise we accept the Ho.

Example : Two samples of 6 and 5 items respectively gave the following results :

Mean of the first sample=40

S.D. of the first sample=8

Mean of the second sample=50

S.D. of the second sample=10

Is the difference of the means significant? (Given $t_{0.05,9} = 2.26$)

Solution : Let μ_1 and μ_2 denote the population means of the first and second population respectively from which samples are drawn.

Here the null hypothesis to be tested is

Ho: $\mu_1 = \mu_2$

against the alternative

H_1: $\mu_1 \neq \mu_2$

We are given, $n_1 = 6, \bar{x} = 40, s_1 = 8$

$n_2 = 5, \bar{x}_2 = 50, s_2 = 10$

$$\text{Now, } s = \sqrt{\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2}}$$

$$= \sqrt{\frac{6 \times 8^2 + 5 \times 10^2}{6 + 5 - 2}} = \sqrt{98.22} = 9.91$$

Under the null hypothesis, the test statistic is

$$|t| = \left| \frac{\bar{x}_1 - \bar{x}_2}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \right| \text{ where } s = \sqrt{\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2}}$$

$$= \left| \frac{40-50}{9.91\sqrt{\frac{1}{6}+\frac{1}{5}}} \right|$$

$$= \frac{10}{6} = 1.67$$

The degrees of freedom=$n_1+n_2-2=5+6-2=9$

The table value of t at 5% level of significance for 9 d.t. is 2.26.

Conclusion : Since the calculated value of t=1.67<2.26, the table value of t, we may accept the null hypothesis at 5% level of significance and conclude that probably the population means are same.

Chi-Square (χ^2) test

Chi-Square test, also called pearson's chi-square test, is a statistical test applied to sets of categorical data to test whether there is any significant difference between two sample results. The test statistic of this test follows a chi-square distribution that tends to approach normality as the number of degrees of freedom increases.

Application of Chi-Square (χ^2) test

Some of the application of Chi-square test are enumerated below :

(i) to test the hypothetical value of the population variance

(ii) to test the goodness of fit,

(iii) to test the independence of attributes

Chi-Square (χ^2) test to test the hypothetical value of the population variance

Suppose a random sample of size n is drawn from a normal population. If we want to test the null hypothesis that population variance $\sigma^2 = \sigma_0{}^2$ (say)

i.e. $Ho: \sigma^2 = \sigma_0{}^2$ (say) against the alternative,

$H_1 : \sigma^2 \neq \sigma_0{}^2$

then the test statistic, under this Ho is given by,

$$\chi^2 = \sum_{i=1}^{n} \frac{(x_1 - \overline{x})^2}{\sigma_0{}^2} \sim \lambda^2_{0.0\alpha,(n-1)}$$

We reject Ho if calculated χ^2 is greater then tabulated one. Otherwise we accept H_0.

Chi-Square test for test the goodness of fit

The chi-square goodness of fit test is used to find out how the observed value of a given phenomena is significantly different from the expected value.

Let $0_1, 0_2, --- 0_n$ be a set of observed frequencies and $E_1, E_2, ----, E_n$ be the corresponding set of expected frequencies. Here the null hypothesis to be tested is,

H_0: there is no significant difference between observed and expected frequencies.

The alternative hypothesis would be.

H_1: there is significant difference between observed and expected frequencies.

Under the null hypothesis H_0, the test statistic is given as

$$\chi^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i} \sim \chi^2_{(n-1)}$$

To apply Chi-Square test, the following conditions should be satisfied :

(1) N, the total frequency should be reasonably large, say, greater than 50

(2) $\sum_{i=1}^{n} O_i = \sum_{i=1}^{n} E_i = N$

(3) No expected cell frequency should be less than 5. If any expected frequency is less than 5, it should be pooled with the adjacent frequency. The degrees of freedom lost due to pooling should be adjusted accordingly.

Note : The test statistic may be simplified as follows :

$$\chi^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$$

$$= \sum_{i=1}^{n} \frac{O_i^2 - 2O_iE_i + E_i^2}{E_i}$$

$$= \sum_{i=1}^{n} \frac{O_i^2}{E_i} - 2\sum_{i=1}^{n} O_i + \sum_{i=1}^{n} E_i$$

$$= \sum_{i=1}^{n} \frac{O_i^2}{E_i} - N \left[\because \sum_{i=1}^{n} O_i = \sum_{i=1}^{n} E_i = N \right]$$

Example : Following values of observed and expected frequencies obtained for a distribution.

x :	5	10	15	20	25	30	35	40	45	50
Observed freq :	2	2	6	13	15	23	16	13	6	4
Expected freq. :	1	3	6	12	17	20	17	13	7	4

Test if the fit is good.

Solution : Here the null hypothesis to be tested is,

H_0 : there is no significant difference between the observed and the expected frequency

against the alternative.

H_1 : there is significant difference between the observed and the expected frequency.

Let us now prepare the following table :

X	O_i	E_i	O_i^2/E_i
5	2	1	
10	2 } 10	3 } 10	10.00
15	6	6	
20	13	12	14.08
25	15	17	13.23
30	23	20	26.45
35	16	17	15.06
40	13	13	13.00
45	6	7	5.14
50	4	4	4.00

$$\sum O_i = 100 \qquad \sum E_i = 100 \qquad \sum \frac{O_i^2}{E_i} = 100.96$$

$$\therefore \chi^2 = \sum_{i=1}^{n} \frac{O_i^2}{E_i} - N$$

$$= 100.96 - 100 = 0.96$$

The table value of χ^2 at 5% level of significance for 10–2–2=6 d.f. is 12.592.

Conclusion : Since the calculated χ^2<12.592, the tabulated value we may accept the null hypothesis at 5% level of significance and conclude that probably there is no significant difference between observed and expected frequency.

Chi-Square test for independence of attributes :

Suppose the observations be classified according to two attributes A and B where A is divided into 'm' classes namely A_1, A_2, -----, A_m and. B is divided into 'n' classes namely B_1, B_2, ----, B_n. Let (A_iB_j) denote the number of persons possessing the attribute A_i and B_j. [i=1, 2, ----m; j=1, 2, ----n]. Now, the table showing saved frequencies in different categories having m-rows and n-columns is called a m × n contingency table.

A \ B	B_1	B_2	------	B_j	------	B_n	Total
A_1	(A_1B_1)	(A_1B_2)	------	(A_1B_j)	------	(A_1B_n)	(A_1)
A_2	(A_1B_1)	(A_2B_2)	------	(A_2B_j)	------	(A_2B_n)	(A_2)
⋮	⋮	⋮		⋮		⋮	⋮
A_i	(A_iB_1)	(A_iB_2)	------	(A_iB_j)	------	(A_iB_n)	(A_i)
⋮	⋮	⋮		⋮		⋮	⋮
A_m	(A_mB_1)	(A_mB_2)	------	(A_mB_j)	------	(A_mB_n)	(A_m)
Total	(B_1)	(B_2)	------	(B_j)	------	(B_n)	N

Here, (A_i) = the number of persons possessing the attribute A_i

$$= \sum_{j=1}^{n}(A_iB_j) \qquad (i=1, 2, -----, m)$$

(B_j) = the number of persons possessing the attribute B_j

$$= \sum_{i=1}^{n}(A_iB_j) \qquad (j=1, 2, ----, n)$$

$$\sum_{i=1}^{m}(A_i) = \sum_{j=1}^{n}(B_j) = N \text{, total frequency}$$

Now, we are to test whether the attributes A and B are independent or not. The null hypothesis to be tested would be

H_0 : the two attributes are independent against the alternative

H_1: the two attributes are not independent under the null hypothesis, the test statistic is given as

$$\chi^2 \quad = \sum \frac{(O-E)^2}{E} \sim \lambda^2_{(m-1)(n-1)}$$

$$= \sum \frac{O^2}{E} - N \text{ (simplified form)}$$

Where the expected frequency (E) of any cell is obtained as

$$E = \frac{\text{Row total} \times \text{column total}}{N}, \text{N=Grand total}$$

The table value of χ^2 is obtained for specified level of significance with d.f. $\gamma = (m-1) \times (n-1)$

If the calculated value of χ^2 is greater than the tobulated one, we reject the null hypothesis at the specified level. Otherwise we accept the H_0.

Example : Two sample poles votes for two candidates A and B for a public office are taken, one from among residents of urban areas, and the other from residents of rural areas. The results are given below. Examine whether the nature of the area is related to voting preference in this election.

Votes for / Area	A	B	Total
Rural	620	380	1000
Urban	550	450	1000
Total	1170	830	2000

(Given, $\chi^2_{0.05,1} = 3.84$)

Solution : Here the null hypothesis to be tested is

Ho : there is no association between the nature of area and voting preference against the alternative

H_1: there is association between the nature of area and voting preference.

Now, we prepare the following table :

Observed frequency O	Expected frequency E	$(O-E)^2$	$\frac{(O-E)^2}{E}$
620	585	1225	$\frac{1225}{585} = 2.10$
550	585	1225	$\frac{1225}{585} = 2.10$
380	415	1225	$\frac{1225}{415} = 2.95$
$\sum O = 2000$	$\sum E = 2000$		$\chi^2 = \sum \frac{(O-E)^2}{E} = 10.10$

$$\left[\begin{array}{l}\text{Formula to obtain expected frequency:}\\ \text{Expected frequency} = \dfrac{\text{Row total} \times \text{column total}}{\text{Grand total}}\end{array}\right]$$

Under the null hypothesis, the test statistic is

$$\chi^2 = \sum \frac{(O-E)^2}{E} = 10.10$$

Degrees of freedom= $(r-1) \times (c-1) = (2-1) \times (2-1) = 1$

The table value of χ^2 for 1 d.f. at 5% level of significance, i.e. $\chi^2_{0.05,1} = 3.84$

Conclusion : Since the calculated value of $\chi^2 > 3.84$, the table value of χ^2, we may reject the null hypothesis at 5% level of significance and conclude that probably the nature of area and voting preference are associated.

Example :

A sample survey conducted in a region regarding educational qualification of 100 men and women reveals the following information :

Sex \ Education	Middle school	High School	College	Total
Male	10	15	25	50
Female	25	10	15	50
Total	35	25	40	100

Test if the educational qualification depend on sex.

Solution : Here the null hypothesis to be tested is,

H_0 : educational qualification is independent of sex against the alternative

H_1 : educational qualification is dependent of sex

Now we prepare the following table :

Observed frequency O	Expected frequency E	$\frac{O^2}{E}$
10	$\frac{50\times 35}{100}=17.5$	$\frac{10^2}{17.5}$
25	$\frac{50\times 35}{100}=17.5$	$\frac{25^2}{17.5}$
15	$\frac{50\times 25}{100}=12.5$	$\frac{15^2}{12.5}$
10	$\frac{50\times 25}{100}=12.5$	$\frac{10^2}{12.5}$
25	$\frac{50\times 40}{100}=20$	$\frac{25^2}{12.5}$
15	$\frac{50\times 40}{100}=20$	$\frac{15^2}{20}$
$\sum O=100$	$\sum E=100$	$\sum\frac{O^2}{E}=109.93$

$$\therefore \chi^2 = \sum \frac{O^2}{E} - N, \quad \text{N=total frequency}$$

$= 109.93–100$

$= 9.93$

Degrees of freedom=(3–1)(2–1)=2

The tabulated value of χ^2 at 5% level of significance for 2 d.f. is 5.99.

Conclusion : Since the calculated value of χ^2>5.99, the tabulated value, we may reject the null hypothesis at 5% level of significance and conclude that probably educational qualification depend on sex.

F-test (F-Distribution)

The F-distribution is a continuous probability distribution. An F-variate is defined by the ratio of two independent chi-square variables divided by their respective degrees of freedom. If λ_1^2 and λ_2^2 are two independent chi-square variates with n_1 and n_2 degrees of freedom respectively, then F is defined as

$$F = \frac{\chi_1^2 / n_1}{\chi_2^2 / n_2}$$

Which follows F-distribution with (n_1, n_2) degrees of freedom.

Application of F-test

Some of the application of F-test are enumerated below :

(1) to test the equality of two population variances

(2) to test the equality of several population means

(3) to test the linearity of regression

(4) to test the significance of observed correlation ratio

(5) to test the significance of observed multiple correlation coefficient

F-test for the equality of two population variances

Let a random sample of size n_1 is drawn from a normal population with variance σ_1^2. Let an another sample of size n_2 is drawn from a normal population with variance σ_2^2. To test whether the two population variances σ_1^2 and σ_2^2 are equal, we construct the null hypothesis as

$H_0 : \sigma_1^2 = \sigma_2^2 = \sigma^2$ (say)

against the alternative

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

Under this Ho, the test statistic is given as

$$F = \frac{S_1^2}{S_2^2}; \left(S_1^2 > S_2^2\right) \sim F(n_1 - 1, n_2 - 1)$$

Here S_1^2 and S_2^2 are unbiased estimates of the common population variance δ^2 and are given by

$$S_1^2 = \frac{1}{n_1 - 1} \sum_{i=1}^{n_1} \left(x_i - \bar{x}\right)^2$$

and $$S_2^2 = \frac{1}{n_2 - 1} \sum_{j=1}^{n_2} \left(y_j - \bar{y}\right)^2$$

If the calculated F is greater than the tabulated value, we reject the null hypothesis at the specified level of significance. Otherwise we accept the null hypothesis.

Note : The greater of the two mean squares S_1^2 and S_2^2 is taken in the numerator of the F-statistic. This implies that if $S_2^2 > S_1^2$, the F-statistic would be

$$F = \frac{S_2^2}{S_1^2} \sim F_{(n_2 - 1, n_1 - 1)}$$

Example : Following results are obtained from two independent samples :

	Size	Mean	Sum of squares of deviation from mean
I = Sample	9	68	36
II= Sample	10	69	42

Test if the population variances are equal.

Solution : Here, the null hypothesis to be tested is

H_0 : the population variances are equal i.e.

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

against the alternative

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

We are given,

$$n_1=9, \quad \bar{x}=68, \sum_{i=1}^{n_1}(x_i-\bar{x})^2 = 36$$

$$n_2=10, \quad \bar{y}=69, \sum_{j=1}^{n_2}(y_j-\bar{y})^2 = 42$$

$$\therefore S_1^2 = \frac{1}{n_1-1}\sum_{i=1}^{n_1}(x_i-\bar{x})^2 = \frac{36}{8} = 4.50$$

$$S_2^2 = \frac{1}{n_2-1}\sum_{j=1}^{n_2}(y_j-\bar{y})^2$$

$$= \frac{42}{9} = 4.66$$

Under the Ho, the test statistic is given by

$$F = \frac{S_2^2}{S_1^2} \quad \left[\because S_2^2 > S_1^2\right]$$

$$= \frac{4.66}{4.50} = 1.03$$

Here the degrees of freedom is $(n_2–1, n_1–1)$

= (10–1=9, 9–1=8)= (9, 8)

The table value of F at 5% level of significance for (9, 8) d.f. is 3.4.

Conclusion : Since the calculated value of $F<3.4$, the tabulated value of F, we may accept the null hypothesis at 5% level of significance and conclude that probably the population variances are same.

Example : Two sources of raw materials are under consideration by a company. Both sources sum to have similar characteristics, but the company is not sure about their respective uniformity. A sample of ten lots from source A yeilds a variance of 225, and a sample of eleven lots from source B yeilds a variance of 200. Is it likely that the variance of source A is significantly greater than the variance of source B?

(Consider α=0.01)

Solution : Here, the null hypothesis to be tested is

H_0 : the variance of sources A and B are equal i.e. $\sigma_1{}^2 = \sigma_2{}^2$

against the alternative

$H_1 : \sigma_1{}^2 > \sigma_2{}^2$

We are given,

$$n_1 = 10, \ S_1{}^2 = 225$$

$$n_2 = 11, \ S_2{}^2 = 200$$

$$\therefore \quad S_1^2 = \frac{n_1}{n_1 - 1} S_1{}^2 = \frac{10}{10-1} \times 225 = 250$$

$$\text{and } S_2^2 = \frac{n_2}{n_2 - 1} S_2{}^2 = \frac{11}{11-1} \times 200 = 220$$

Under Ho, the test statistic is

$$F = \frac{S_1^2}{S_2^2} \qquad \left[\because S_1^2 > S_2^2\right]$$

$$= \frac{250}{220} = 1.14$$

Here, the degrees of freedom is $(n_1–1, n_2–1)$

$= (10–1, 11–1)=(9, 10)$

The table value of F at 1% level of significance for (9, 10) d.f. is 4.94.

Conclusion : Since the calculated value of F<4.94, the tabulated value of F, we may accept the null hypothesis at 1% level of significance and conclude that probably the two population variances are equal.

4.4 Summing Up

The tests based on the assumption that the samples are drawn from a normally distributed population and population parameters are partially known or atleast estimable are termed as Parametric test. Further in parametric tests the most commonly used parameters are mean and standard deviation. Although it is difficult to draw a clear cut line between large and small samples, the statisticians have agreed upon that a sample is to be considered as large if its size is greater than or equal to 30, i.e. if sample size ≥ 30. The tests of significance used for dealing with problems relating to large samples are different from the ones used for small samples. This is due to the reason that the assumptions made in case of large samples do not hold good for small samples.

4.5 Model Questions

1. Is it likely that a sample of 300 items whose mean is 60.0 is a random sample from a large population whose mean is 16.1 and standard deviation 5.2?

2. A college conducts both day and night classes. A sample of 100 'day student' yields examination results as under :

   $\overline{x_1} = 72.4, \quad s_1 = 14.8$

   A sample of 200 'night student' yields examination results as under :

   $\overline{x_1} = 72.4, \quad s_1 = 17.9$

   Are the two means statistically equal at 1% level?

3. A dies is thrown 49152 times and of these 25145 yielded either 4 or 5 or 6. Is this consistent with the hypothesis that the die must be unbiased?

4. Mention the various assumptions of Student's t-test.

5. Explain clearly the concept of degrees of freedom.

6. A random sample of size 7 from a normal population gave a mean of 977.51 and standard deviation of 4.42. Find the 95% confidence interval for population mean.

7. A salesman is expected to effect an average sales of Rs. 3500. A sample test revealed that a particular salesman had made the following sales : Rs. 3700, Rs. 3400, Rs. 2500, Rs. 3000 and Rs. 2000. Using 0.05 level of significance, conclude whether his sale is below expectation or not.

8. Write down the conditions for the application of χ^2 test.

9. Three hundred digits were chosen at random from a set of tables. The frequencies of the digits were as follows :

Digit	:	0	1	2	3	4	5	6	7	8	9
Frequency	:	28	29	33	31	26	35	32	30	31	25

Using the χ^2 test assess the hypothesis that the digits were distributed in equal number in the tables. [Table value of $\chi^2_{0.05,9} = 16.92$]

10. A sample of 20 observations gave a standard deviation of 3.72. Is this compatible with the hypothesis that the sample is drawn from a normal population with variance 4.35? [Table value of $\chi^2_{0.05,19} = 30.144$]

11. Two samples were drawn from two normal populations and their values are :

A :	66	67	75	76	82	84	88	90	92		
B :	64	66	74	78	82	85	87	92	93	95	97

Test whether the populations have the same variances at 5% level of significance.

12. The following data relate to a random sample of government employees in the year 1995 in two states of the Indian union :

	State I	State II
Sample size	16	25
Mean monthly income of the sample employee (in Rs.)	440	460
Sample variance	40	42

Carry out a test of hypothesis that the variance of the two populations are equal.

4.6 Reference & Suggested Readings

1. Sheskin David J., Handbook of Parametric and Nonparametric Statistical Procedures, Fifth Edition, CRC Press.
2. Alvo Mayer, Yu Philip L. H., A Parametric Approach to Nonparametric Statistics, Springer
3. Pfanzagl Johann, Parametric Statistical Theory, De Gruyter Textbook.
4. Hogg, Tanis,Rao; Probability and statistical inference; Pearson
5. Bhuyan K.C; Probability Distribution Theory and Statistical Inference; New Cemtral Book Agency(P) Ltd.
6. Elhance D.N, Elhance Veena, Agarwar B.M,;Fundamentals of Statistics; Kitab Mahal
7. Gupta S.C.,Kapoor V.K.; Fundamentals of Mathematical Statistics ; Sultan Chand & Sons
8. Bhardwaj R.S.; Business Statistics ; Excel Books
9. Choudhury L, Sarma R, Deka M, Gogoi S.J.; An Introduction to Statistics; L. Choudhury.

******

Unit-5
Non-Parametric Tests

Unit Structure:

5.1 Introduction

The tests based on the assumption that the samples are drawn from a normally distributed population and the population parameters are partially known or at least estimable have been discussed in the previous unit of Parametric Tests. But there are some situations where such assumptions are not tenable. To over come this limitation we have to adopt an another technique of hypothesis testing called Non-Parametric method.

In this method no assumptions about the parameters or about the nature of the population require. That is why, sometimes this method is also called distribution free method. Non-Parametric methods of estimation and tests often depend on ordered sample and also order statistics.

A random sample (x_1, x_2, ...x_n) is said to be an ordered sample if $x_1<x_2<.......<x_n$. Any statistic defined on the basis of the order of a sample is called an order statistic. e.e. median, quartile, semi-inter quartile range etc.

5.2 Objectives

After going through this unit, you will be able to-

- know the definition of Non-Parametric test
- understand the different types of Non-Parametric test
- describe the technique of solving practical problems using different Non-Parametric tests.

5.3 Non-Parametric Tests of Hypothesis

As defined earlier, non-parametric tests donot depend on the distribution of the sampled population. That is why, such tests are also called 'distribution-free tests'. Also, non-parametric methods focus on the location of the probability distribution of the parent population, rather than on specific parameters of the population. Before discussing the different non-parametric tests, let us discuss the advantages and disadvantages of non parametric tests.

Advantages

1. Non-parametric tests require less restrictive assumptions as compared to parametric test.
2. These tests often require very few arithmetic computations.
3. There is no alternative to use a non parametric test if the data are available in ordinal or nominal scale.
4. Non-parametric tests are useful with small samples.

Disadvantages

1. Parametric tests are more powerful than non parametric tests. This means that there is a greater risk of accepting a false hypothesis.
2. Non parametric methods deal with test of hypothesis only. No non parametric method of estimation is available.

The Paired-Sample Sign Test

This test is used to test for consistent differences between pairs of observations. The test has very important application in problems involving paired data such as data relating to the responses of mother and daughter towards ideal family size, weight of patients before and after treatment, etc. In such problems, each pair of sample values can be replaced with a + (plus) sign if the first value is greater than the second, a – (minus) sign if the first value is smaller than the second or be discarded if the two values are same.

Let, P be the proportion of plus signs, i.e.

$$P = \frac{\text{Number of plus signs}}{\text{Total number of pairs}}$$

Let, P be thi proportion of plus sign in the population.

Then the null hypothesis to be tested is

Ho : P=0.5

If the difference is due to chance effects the probability of a + sign for any particular pair is $\frac{1}{2}$, as is the probability of a – sign. If S is the number of times the less frequent sign occurs, then S has the binomial distribution with P=$\frac{1}{2}$.

The table value for a two-tailed test at 5% level of significance can be conveniently found by the expression

$$K = \frac{(n-1)}{2} - (0.98)\sqrt{n}$$

Where n is the sample size minus number of tied pairs.

Ho is rejected if $S \le K$ for the sign test

In case of large samples (generally considered n>25) then the binomial distribution can be approximate with normal distribution with

$$Z = \frac{k - np}{\sqrt{npq}} \sim N(0,1)$$

Where K= the number of most frequently occuring signs

q=1–p

We compare the calculated value of Z with the table value of Z and draw conclusion as before. The approximation becomes better when a correction for continuity is employed and then

$$Z = \frac{(k \pm 0.5) - 0.5n}{0.5\sqrt{n}} \sim N(0,1)$$

Here, k+0.5 is used when $k < \frac{n}{2}$ and k–0.5 is used when $k > \frac{n}{2}$.

Conclusions are drawn in the similar manner.

Example : A group of 30 people have started physical exercise to reduce their body weight. Their body weights (in kg) are recorded before (x) and after (y) the physical exercise. The data are given below :

x : 56, 58, 62, 57, 56, 60, 64, 66, 67, 62, 64,

y : 55, 58, 60, 56, 56, 58, 62, 60, 63, 60, 63, x : 67, 68, 55, 70, 62, 61, 64, 60, 58, 57, 64,

y : 65, 66, 58, 70, 63, 62, 62, 61, 55, 56, 63,

x : 52, 56, 58, 57, 60, 62, 64, 67

y : 54, 57, 55, 60, 58, 61, 63, 66

Is there any change in body–weights due to physical exercise?

Solution : Here, the null hypothesis to be tested is,

Ho : there is no change in body-weights due to physical exercise. i.e. P=0.5

against the alternative,

$H_1 : P \neq 0.5$

Here, $D_i = x_i - y_i$

$= + 0 + + 0 - + + + + + + + + - 0 - -$

$+ - + + - - + - + + + +$

$\therefore$ n = 30–3=27

and $k = 19 > \frac{n}{2} = \frac{27}{2} = 13.5$

$\therefore$ Under the null hypothesis, the test statistic is

$$Z = \frac{(k \pm 0.5) - 0.5n}{0.5\sqrt{n}} \sim N(0,1)$$

$$= \frac{(k - 0.5) - 0.5n}{0.5\sqrt{n}}, \because k > \frac{n}{2}$$

$$= \frac{(19 - 0.5) - 0.5 \times 27}{0.5\sqrt{27}}$$

$$= \frac{18.5 - 13.5}{2.598}$$

$= 1.92$

The table value of Z at 5% level of significance is 1.96

Conclusion : Since calculated value of Z<1.96, the tabulated value, we may accept the null hypothesis at 5% level of significance and conclude that probably there is no change in body-weight due to physical exercise.

Wilcoxon signed Rank-Sum test

The Wilcoxon signed rank sum test is an another kind of non-parametric test and can be used to test the null hypothesis that the median of distribution is equal to some value. Moreover, the test can be used in lieu of a one-sample t-test or a paired t-test.

We perform the following steps to carry-out the test.

Case I : Paired Data

1. State the null hypothesis– in this case
   Ho : the median difference, M=O
2. Calculate each paired difference,
   $d_i=x_i-y_i$, where (x_i, y_i) are the pairs of observations. (i=1, 2, ..., n)
3. Rank these differences in ascending order ignoring the signs (i.e. assign rank 1 to the smallest |di|, rank 2 to the next, etc)
4. The cases of tied ranks are assigned ranks by the average method.
5. Label each rank with its sign, according to the sign of di.
6. Calculate T_+, the sum of the ranks of the positive d_is, and T_-, the sum of the ranks of the negative d_is. (As a check the total, $T_+ + T_-$, should be equal to $\frac{n(n+1)}{2}$, where n is the number of pairs of observations in the sample)

Case II : Single of Observations

1. State the null hypothesis – the median value is equal to some value M, i.e.
   Ho : Median = M
2. Calculate the difference between each observation and the specified value of median M, i.e. $d_i=x_i-M$
3. Rank these differences in ascending order ignoring the signs (i.e. assign rank 1 to the smallest |di|, rank 2 to the next, etc)
4. The cases of tied ranks are assigned ranks by the average method.
5. Label each rank with its sign, according to the sign of di.
6. Calculate T_+, the sum of the ranks of the positive d_is, and T_-, the sum of the ranks of the negative d_is. (As a check the total, $T_+ + T_-$, should be equal to $\frac{n(n+1)}{2}$, where n is the number of pairs of observations in the sample)

   Under the null hypothesis, the test statistic, to be denoted by T, is obtained as follows :

7. Choose T=min (T_-, T_+)
8. Use tables of critical values for the Wilcoxon signed rank sum test to find the probability of observcing a value of T or more extreme. Most tables give both one-sided and two-sided p-values. If not, double the one-sided p-value to obtain the two-sided value.

It can be shown that the distribution of T is approximately normal (when n>20) with mean $\mu_T = \frac{n(n+1)}{4}$ and standard error $\delta_T = \sqrt{\frac{n(n+1)(2n+1}{24}}$

Thus, the test statistic under the Ho would be,

$$Z = \left| \frac{T - \mu_T}{\delta_T} \right| \sim Z(0,1)$$

We follow the same decision rule as before.

Dealing with ties :

These are two types of tied observations that may arise during the test procedure :

- Observations in the sample may be exactly equal to M (i.e. O in the case of paired differences). Ignore such observations and adjust n accordingly.
- Two or more observations/differences may be equal. If so, average the ranks across the tied observations and reduce the variance by $\frac{t^3 - t}{48}$ for each group of t tied ranks.

Example :

The following table shows the hours of relief provided by two analgesic drugs in 12 patients suffering from arthritis. Is there any evidence that the drug B provides longer relief than the drug A?

Patient:	1	2	3	4	5	6	7	8	9	10	11	12
Drug A :	2.0	3.6	2.6	2.6	7.3	3.4	14.9	6.6	2.3	2.0	6.8	8.5
Drug B :	3.5	5.7	2.9	2.4	9.9	3.3	16.7	6.0	3.8	4.0	9.1	20.9

Solution :

Here the null hypothesis to be tested is

Ho : There is no significant difference between the two drugs with respect to relief hour against the alternative.

H_1: Drug B Provides longer relief than the drug A.

Now, the differences between the two drugs with respect to relief hour are:

di ; (Drug B–Drug A) :

+1.5, +2.1, +0.3, –0.2, +2.6, –0.1, +1.8, –0.6, +1.5, +2.0, +2.3, +12.4

The ascending order of these differences (ignoring the sign) is given as,

0.1, 0.2, 0.3, 0.6, 1.5, 1.5, 1.8, 2.0, 2.1, 2.3, 2.6, 12.4

Now, let us rank these differences and label each rank with its sign according to the sign of di.

|di|	0.1	0.2	0.3	0.6	1.5	1.5
Rank	1	2	3	4	5.5	5.5
Sign	–	–	+	–	+	+

|di|	1.8	2.0	2.1	2.3	2.6	12.4
Rank	7	8	9	10	11	12
Sign	+	+	+	+	+	+

$\therefore \quad T_+$ = sum of the ranks of the positive dis

$= 3+5.5+5.5+7+8+9+10+11+12$

$= 71$

and T_- = sum of the ranks of the negative dis

$= 1+2+4$

$= 7$

$\therefore T_+ + T_- = 71+7=78$

and $\frac{n(n+1)}{2} = \frac{12 \times 13}{2} = 78$

i.e. $T_+ + T_- = \frac{n(n+1)}{2}$

Now, $T=\min(T_-, T_+) = 7$

We can use a normal approximation in this case. Here we have 2 tied ranks, so we must reduce the variance by

$$\frac{t^3 - t}{48} = \frac{2^3 - 2}{48} = 0.125$$

$\therefore$ We get,

$$Z = \left| \frac{T - \mu_T}{\delta_T} \right|$$

$$= \left| \frac{T - \frac{n(n+1)}{4}}{\sqrt{\frac{n(n+1)(2n+1)}{24} - \frac{t^3 - t}{48}}} \right|$$

$$= \left| \frac{7 - \frac{12 \times 13}{4}}{\sqrt{\frac{12 \times 13 \times 25}{24} - 0.125}} \right|$$

$$= \left| \frac{7 - 39}{\sqrt{162.5 - 0.125}} \right| = 2.511$$

The table value of Z at 5% level of significance for one-tailed test is 1.64

Conclusion : Since the calculated value of Z>1.64, the table value of Z, we may reject the null hypothesis at 5% level of significance and conclude that probably the drug B provides longer relief than the drug A.

Check Your Progress

1. What is ordered sample?
2. Non-Parametric test are called distribution free tests. Why?
3. Point out the uses of the Paired-Sample sign test and Wilcoxon signed Rank-Sum test.

Mann–Whitney U-test

Mann-Whitney U-test is the non-parametric test that helps us to determine whether two random samples have come from identical population or not. Usually this test is used when the data are ordinal and when the assumptions of the t-test are not met. The basic assumption of the test is that the distribution of the two populations are continuous with equal standard deviation.

Let n_1 and n_2 be the sizes of the samples taken from population I and population II respectively. The steps to be carry out for the test are as follows :

1. State the null hypothesis– in this case, Ho : the population means are equal,

   i.e. $\mu_1 = \mu_2$

against the alternative

   i.e. $H_1 : \mu_1 \neq \mu_2$

2. Rank all the n_1+n_2 observations, and arrange in ascending order.
3. Find R_1 and R_2, where R_i denotes the sum of ranks of the i^{th} sample (i=1, 2)

Now, for the sample sizes n_1 and n_2, the sum of R_1 and R_2 is simply the sum of first n_1+n_2 positive integers, which is given as

$$\frac{(n_1 + n_2)(n_1 + n_2 - 1)}{2}$$

This formula enables us to find R_2 if we know R_1 and vice versa. We obtain the following two statistics to make decisions :

$$U_1 = n_1 n_2 + \frac{n_1(n_1 + 1)}{2} - R_1$$

$$U_2 = n_1 n_2 + \frac{n_2(n_2 + 1)}{2} - R_2$$

For small samples, if both n_1 and n_2 are less than 10 special tables (statistical tables for Mann–Whitney U test) must be used. If U is smaller than the critical value, Ho can be related to the standard normal curve by the statistic

$$Z = \frac{U - \frac{n_1 n_2}{2}}{\sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}} \sim N(0,1)$$

Where, $\frac{n_1 n_2}{2} = \mu_U$

$$\sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}} = \delta_U$$

We will follow same decision rule as before.

Example :

Twenty-four applicants for a position are interviewed by three administrators and rated on a scale of 5 as to suitability for the position. Each applicant is given a 'suitability' score which is the sum of the three numbers. Although college education is not a requirement for the position, a personnel director felt that it might have some bearing on suitability for the position. Raters made their ratings on the basis of individual interviews and were not told the educational background of the applicants. Twelve of the applicants had completed at least two years of college. Use the Mann Whitney U-test to determine whether there was a difference in the scores of the two groups. Use a 0.05 level of significance, Group A had an educational background of less than two years of college, while group B had completed at least two years of college.

"Suitability" Scores

Group A	Group B
7	8
11	9
9	13
4	14
8	11
6	10
12	12

11	14
9	13
10	9
11	10
11	8

Solution : Here, the null hypothesis to be tested is

Ho : There is no significant difference in the scores of the two groups against the alternative.

H_1: There is a significant difference in the scores of the two groups.

Let us first prepare the following table

	Common Ranks	Observations
	1	4
	2	6
	3	7
	5	8
$\frac{4+5+6}{3}=5$	5	8
	5	8
	8.5	9
$\frac{7+8+9+10}{4}=8.5$	8.5	9
	8.5	9
	8.5	9
	12	10
$\frac{11+12+13}{3}=12$	12	10
	12	10
	16	11
	16	11
$\frac{14+15+16+17+18}{5}=16$	16	11
	16	11

16	11
$\frac{19+20}{2}=19.5$ 19.5	12
19.5	12
$\frac{21+22}{2}=21.5$ 21.5	13
21.5	13
23.5	14
23.5	14

Now, we allocate the ranks to the corresponding observations of the two groups as follows :

Group A	Corresponding Ranks	Group B	Corresponding Ranks
7	3	8	5
11	16	9	8.5
9	8.5	13	21.5
4	1	14	23.5
8	5	11	16
6	2	10	12
12	19.5	12	19.5
11	16	14	23.5
9	8.5	13	21.5
10	12	9	8.5
11	16	10	12
11	16	8	5
	R_1=123.5		R_2=176.5

$\therefore \quad R_1+R_2=123.5+176.5=300$

Again, $\frac{(n_1+n_2)(n_1+n_2+1)}{2}=\frac{(12+12)(12+12+1)}{2}=300=R_1+R_2$

This serves as a check for internal consistency.

Now, for n_1=12, n_2=12, R_1=123.5 & R_2=176.5

We have,

$$U_1 = n_1 n_2 + \frac{n_1(n_1+1)}{2} - R_1$$

$$= 12\times 12 + \frac{12(12+1)}{2} - 123.5$$

$$= 144+78-123.5$$

$$= 98.5$$

and

$$U_2 = n_1 n_2 + \frac{n_2(n_2+1)}{2} - R_2$$

$$= 12\times 12 + \frac{12(12+1)}{2} - 176.5$$

$$= 144+78-176.5$$

$$= 45.5$$

If the table value of U is available for n_1=12 and n_2=12, we may use it for decision making (Table : Critical Values of the Mann–Whitney U),

Otherwise, we may use the test statistic

$$Z = \left| \frac{U - \frac{n_1 n_2}{2}}{\sqrt{\frac{n_1.n_2(n_1+n_2+1)}{12}}} \right| \sim N(0,1)$$

Where U = min (U_1, U_2)

$$= \left| \frac{45.5 - \frac{12\times 12}{2}}{\sqrt{\frac{12\times 12(12+12+1)}{12}}} \right|$$

$$= \left| \frac{45.5-72}{17.3205} \right|$$

$$= 1.53$$

The tabulated value of Z at 5% level of significance is 1.96.

Conclusion : Since the calculated value of $Z < 1.96$, the tabulated value of Z, we may accept the null hypothesis at 5% level of significance and conclude that probably there is no significant difference in the scores of the two groups.

Runs Test for Randomness

Run test of randomness is a non-parametric test that is used to know the randomness in data. It is alternative test to test auto-correlation in the data. Run test of randomness is basically based on the run. A run is defined as a sequence of identical symbols which are preceded and followed by different or no symbols at all. Run test of randomness assumes that the mean and variance are constant and the probability is independent.

Test Procedure

The first step in the runss test is to count the number of runs in the data sequence. For example, suppose that a sequence of two symbols, A and B, occured as follows :

ABAABABBBAAABBA

The number of runs in the above sequence are 9.

It should be noted that when there are n observation, where each is denoted by either symbol A or by B, the possible number of runs would lie between and including 2 to n.

Now to test randomness of a sample, we set the null hypothesis as :

Ho : the sequence was produced in a random manner, i.e. the sample is a random.

against the alternative

H_1 : the sample is not random.

As per the null hypothesis, if we get a significantly large or small number of runs, Ho is rejected.

Now, to construct the test statistic, let us assume that n_1 be the number of symbol of one type and n_2 be the number of symbols of other type, so that $n=n_1+n_2$ is the total number of observations in the sequence.

Again, let R be the number of runs in the sequence.

Using theory of algebra, it can be shown that R is a random variable having mean $\mu_R = \frac{2n_1n_2}{n} + 1$ and standard error

$$\delta_R = \sqrt{\frac{2n_1n_2(2n_1n_2 - n}{n^2(n-1)}}$$

For large sample runs test (when $n_1 > 10$ and $n_2 > 10$), the distribution of R can be approminated by a normal distribution.

In this case, the test statistic would be

$$Z = \left| \frac{R - \mu_R}{\delta_R} \right| \sim N(0,1)$$

We reject Ho, if calculated Z>tabulated Z at α % of level of significance. Otherwise we accept Ho.

For a small sample runs test, there are tables to determine critical values that depend on values of n_1 and n_2.

Example :

The weights (gms) of 31 apples have been collected from a consignment and are as follows :

106, 107, 76, 82, 106, 107, 115, 93, 187, 95, 123, 125, 111, 92, 86, 70, 127, 68, 130, 129, 139, 119, 115, 128, 100, 186, 84, 99, 113, 204, 111

Test the hypothesis that the sample is random.

Solution : Let us denote the increase in the successive observation by a plus (+) sign and the decrease of successive observation by a minus (–) sing. From the given data, we can have the following sequence of plus (+) and minus (–) signs.

+ – + + + + – + – + + – – – – + – + – +

– – + – + – + + + –

Let, n_1 = the number of plus sign

n_2 = the num ber of minus sign

R = the number runs

Here, the null hypothesis to be tested is

Ho : the sample is random

against the alternative

H_1 : the sample is not random

From, the sequence of plus and minus sign, we have

$n_1=16, \quad n_2=14$

and R=20

$\therefore \quad n = n_1+n_2=16+14=30$

$$\text{Thus, } \mu_R = \frac{2n_1n_2}{n}+1$$

$$= \frac{2\times16\times14}{n}+1$$

$$= \frac{2\times16\times14}{n}+1 = 15.93$$

$$\text{and} \qquad \delta_R = \sqrt{\frac{2\times16\times14(2\times16\times14-30)}{900\times29}} = 2.68$$

$\therefore$ the test statistic is given as

$$Z = \left|\frac{R-\mu_R}{\delta_R}\right| \sim N(0,1)$$

$$= \left|\frac{20-15.93}{2.68}\right| = 1.52$$

The table value of Z at 5% level of significance is 1.96.

Conclusion : Since the calculated value of Z<1.96, the table value of Z, we may accept the null hypothesis at 5% level of significance and conclude that probably the sample is random.

5.4 Median Test for Randomness

Median test for randomness is a kind of non-parametric test. It is a test based upon the number of runs above and below the median of the sample. If the sample is random, the successive observations of the sample are expected to be above or below median and consequently the number of runs (R) would be large.

The test of hypothesis in this case would be a one tailed test. The null hypothesis to be tested is,

Ho : the sample is random i.e. number of runs $R \geq \mu_R$

against the alternative

H_1 : the sample is not random i.e. number of runs $R < \mu_R$

It can be shown that R follows normal distribution with mean $\mu_R = \frac{n+2}{2}$

and standard error $\delta_R = \sqrt{\frac{n(n-2)}{4(n-1)}}$, where n denotes the number of observations (excluding the observations that are equal to median value) in the sequences.

The test statistic is given as,

$$Z = \left| \frac{R - \mu_R}{\delta_R} \right| \sim N(0,1)$$

We follow the same decision rule as before,

Example :

The weights (gms) of 31 apples have been collected from a consignment and are as follows :

106, 107, 76, 82, 106, 107, 115, 93, 187, 95, 123,

125, 111, 92, 86, 70, 127, 68, 130, 129, 139, 119,

115, 128, 100, 186, 84, 99, 113, 204, 111

Test the hypothesis that the sample is random.

Solution : Let us arrange the observations in ascending order as follows :

68, 70, 76, 82, 84, 86, 92, 93, 95, 99, 100, 106, 106,

107, 107, 111, 111, 113, 115, 115, 119, 123, 125, 127,

128, 129, 130, 139, 186, 187, 204

$\therefore$ the median of the sequence is 111.

Let, L denotes that an observation is lower than median and H denotes that an observation is higher than it.

Then, the given sequence may be written in L and H in the following way :

L L L L L L H L H L H H L L L H L H H H H H H L H L L H H (ignoring the observations that are equal to median)

Thus, we have, R = the number of runs = 14

n = number of observations in the sequence (ignoring the observation that are equal to median)

= 29

Here, the null hypothesis to be tested is,

Ho : the sample is random i.e. $R \geq \mu_R$ against the alternative $H_1 : R < \mu_R$

Now, $\delta_R = \frac{n+2}{2} = \frac{29+2}{2} = 15.5$

and $\delta_R = \sqrt{\frac{n(n-2)}{4(n-1)}} = \sqrt{\frac{29 \times 27}{4 \times 28}} = 2.64$

$\therefore$ the test statistics is given as,

$$Z = \left| \frac{R - \mu_R}{\delta_R} \right| \sim N(0,1)$$

$$= \left| \frac{14 - 15.5}{2.64} \right|$$

$$= 0.568$$

The table value of Z at 5% level of significance for one tailed test is 1.645.

Conclusion : Since calculated Z<1.645, the tabulated value of Z, we may accept the null hypothesis at 5% level of significance and conclude that probably the sample is random.

Check Your Progress

1. What is Ordered sample?
2. Non Paramatic Tests are called distribution free tests. Why?
3. Point out the uses of the paired sample sign test and wilcoxon signed rank sum test.

5.5 Summing Up

Non Paramatric tests are methods of statistical analysis thatb do not require the nature of the population. Moreover, no assumptions are made about the parameters and that is why, such tests are also termed as 'distribution-free tests'.Thee are several advantages and disadvantages of using such tests. The most popular non parametric test are namely ,the Paired-sample sign test, Wilcoxon Signed Rank-Sum Test, Mann Whitney U-Test, Runs Test for Randomness and Median test for randomness.

The Paired Sample Sign Test is used to test for consistent differences between pairs of observations. Again the Wilcoxon Signed rank sum test is applied to test whether the median of a distribution is equal to a specific value or not. On the other hand, Mann-Whitney U-Test helps us to determine whether two random samples have the same population or not .Another two tests namely- Runs Test for randomness and Median Test for randomness are used to know the randomness in data.

5.6 Model Questions

1. What are non parametric tests? Explain Briefly
2. Point out the advantages and disadvantages of non parametric tests.
3. Define the paired sample sign test
4. Write the steps to carry out the Wilcoxon signed Rank Sum Test
5. What is Mann Whitney U -Test. How the test is carried out?
6. Discuss the tests used for randomness in data.

5.7 References and Suggested Readings

1. Hogg, Tanis,Rao; Probability and statistical inference; Pearson
2. Bhuyan K.C; Probability Distribution Theory and Statistical Inference; New Cemtral Book Agency(P) Ltd.
3. Elhance D.N, Elhance Veena, Agarwar B.M,;Fundamentals of Statistics; Kitab Mahal
4. Gupta S.C.,Kapoor V.K.; Fundamentals of Mathematical Statistics ; Sultan Chand & Sons
5. Bhardwaj R.S.; Business Statistics ; Excel Books
6. Choudhury L, Sarma R, Deka M, Gogoi S.J.; An Introduction to Statistics; L. Choudhury.

******