UNIT 13 ONE-SAMPLE TESTS

Structure

13.1 INTRODUCTION

In the previous block, we have described the procedure for testing various hypotheses involving population parameter(s) such as mean(s), proportion(s), variance(s), etc., together with a large number of examples. What we were doing there can be summarised in following two points:

- We were making an assumption regarding the form of the distribution function of the parent population from which the sample has been drawn is known to us, and
- We were testing a statistical hypothesis regarding population parameters under study such as mean(s) or variance(s) or proportion(s), etc.

Those types of tests where we deal above two points are known as **"parametric tests"**.

A parametric test is a test which is based on the fundamental assumption that the form of parent population is known and involved parameter(s).

A non-parametric test is a test that does not make any assumption regarding the form of the parent population from which the sample is drawn.

But in many real life problems particularly in social and behavioural sciences the requirements of parametric tests do not meet. Therefore, in such situations, the parametric tests are not applicable. Thus, statisticians discovered various tests and methods which are independent of the form of population distribution and also applicable when the observations are not measured in interval scale that is observations are measured in ordinal scale or nominal scale. These tests are known as **"Non-parametric tests"** or **"Distribution Free Tests"**. The name distribution free test is suggested, since the distribution of the test statistic is free (independent) from the form of population distribution. In addition to it the non-parametric test is suggested, since these tests are developed for such hypothesis testing which are not involving the population parameter(s).

These tests are based on the **"order statistics"** theory which has the property that "The distribution of the area under the density function between any two ordered observations is independent of the form of the density function". Therefore, in these tests we use median, range, quartile, etc., in our hypothesis.

Assumptions Associated with Non-parametric Tests:

Although the non-parametric tests are not based on the assumption that the form of parent population is known but these tests require some basic assumptions which are listed below:

(i) Sample observations are independent.

(ii) The variable under study is continuous.

(iii) Lower order moments exist.

Obviously, these assumptions are fewer and weaker than those associated with parametric tests.

This unit is divided into eight sections. Section 13.1 is described the need of non-parametric tests. The advantages and disadvantages of the non-parametric tests over parametric tests are described in Section 13.2. In Section 13.3 and 13.4, we discuss sign test and Wilcoxon signed-rank test for one-sample which are generally used when assumption(s) of t-test is (are) not fulfilled. To test the randomness of the data we use run test so in Section 13.5 we explore the run test. In Section 13.6, we describe the Kolmogorov-Smirnov test for goodness of fit. Unit ends by providing summary of what we have discussed in this unit in Section 13.7 and solution of exercises in Section 13.8.

Objectives

After completion of this unit, you should be able to:

- describe the need of non-parametric tests;
- describe the applications of non-parametric tests in various fields with their advantages and disadvantages over parametric tests;
- describe the one-sample non-parametric tests;
- describe an appropriate one-sample non-parametric test for different situations;
- perform the test which is used in place of t-test for mean when assumption(s) of the t-test is (are) not fulfilled, that is, sign test and Wilcoxon signed-rank test;
- perform the test for testing the randomness of the data, that is, run test; and
- perform the Kolmogorov-Smirnov test for goodness of fit.

13.2 ADVANTAGES, DISADVANTAGES AND TYPES OF NON-PARAMETRIC TESTS

In previous section, we have already said that non-parametric tests do the job for us when assumptions of parametric tests do not meet. With this advantage and some other advantages, non parametric test also have some disadvantages. Both are listed in the following table:

Advantages	Disadvantages
The non-parametric tests do not make any assumption regarding the form of the parent population from which the sample is drawn.	When the assumptions of parametric tests are fulfilled then parametric tests are more powerful than non-parametric tests. For example, Wilcoxon test has approximately 95% power compare to t-test. It means that if null hypothesis is wrong then Wilcoxon test requires 5% more observations in the sample compare to t-test to achieve same power. As an example, Wilcoxon test requires 20 observations compare to t-test which requires 19 observations to achieve same power.
Non-parametric tests have	They do not use all the information available in terms of

simple calculation compare to parametric tests because they generally use ranks rather than actual values of the sample observations.	sample observations because they generally use ranks rather than actual values of the sample observations.
Because of using ranks they are also applicable even if data is available in ordinal measurement scale.	In non-parametric tests the required sample size reduces if samples observations are tied as we will see during calculation.

Now, you can answer the following exercise.

E1) Write any three advantages and disadvantages of non-parametric tests compare to parametric tests?

Types of Non-parametric Tests

After discussing the advantages and disadvantages of the non-parametric tests, now, you are interesting to know type of non-parametric tests. Broadly, the non-parametric tests are classified into three categories as:

1. One-sample Tests
2. Two-sample Tests
3. k-sample Tests.

In this unit, we will discuss one-sample tests and other type of tests will be discussed in subsequent units of this block.

Types of One-sample Tests

Some of the commonly used one-sample tests are listed below:

1. Sign Test
2. Wilcoxon Signed-Rank Test
3. Run Test
4. Kolmogorov-Smirnov Goodness of Fit Test

Let us discuss these tests one by one in subsequent sections.

13.3 SIGN TEST

The sign test is the simplest of the non-parametric tests. It is called the sign test because this test is based on signs as plus and minus as you will see when we proceed through an example. The sign test is used as an alternative of the t-test for testing population median instead of population mean under the circumstances when the parent population is not normal. Further it also works if data is available in ordinal scale while for t-test we needed data at least in interval scale.

The hypothesis concerning the median cannot be considered as non-parametric testing problem in strict sense (median is itself a parameter) but it is taken as a valid non-parametric testing problem since the distribution of the test statistic does not involve the parent population

Assumptions

This test works under following assumptions:

(i) The sample is selected from the population with unknown median.

(ii) The variable under study is continuous.

(iii) The variable under study is measured on at least ordinal scale.

Let us discuss general procedure of this test:

Let $X_1, X_2, ..., X_{n'}$ be a set of n' random observations arranged in the order in which they occur taken from the parent population having unknown median $\tilde{\mu}$. Suppose, we wish to test the hypothesis about the specified value $\tilde{\mu}_0$ of population median $\tilde{\mu}$. So we can take the null and alternative hypotheses as

$$H_0 : \tilde{\mu} = \tilde{\mu}_0 \text{ and } H_1 : \tilde{\mu} \neq \tilde{\mu}_0 \qquad [\text{for two-tailed test}]$$

or

$$\left.\begin{array}{l} H_0 : \tilde{\mu} \leq \tilde{\mu}_0 \text{ and } H_1 : \tilde{\mu} > \tilde{\mu}_0 \\ H_0 : \tilde{\mu} \geq \tilde{\mu}_0 \text{ and } H_1 : \tilde{\mu} < \tilde{\mu}_0 \end{array}\right\} \qquad [\text{for one-tailed test}]$$

After setting null and alternative hypotheses, sign test involves following steps:

Step 1: First of all, we convert the given observations into a sequence of plus and minus signs. For this, we subtract the postulated value of median $(\tilde{\mu}_0)$ from each observation, that is, we obtain the differences $X_i - \tilde{\mu}_0$ for all observations and check their signs. Another way, we compare each observation X_i with $\tilde{\mu}_0$. If X_i is greater than $\tilde{\mu}_0$, that is, $X_i > \tilde{\mu}_0$, then we replace the observation X_i by a plus sign and if X_i is less than $\tilde{\mu}_0$, that is, $X_i < \tilde{\mu}_0$, then we replace X_i by a minus sign. But when the observation X_i which is equal to $\tilde{\mu}_0$ give no information in terms of plus or minus signs so we exclude all such observations from the analysis part. Due to such observations our sample size reduces and let reduced sample size be denoted by n.

Step 2: After that count the number of plus signs and number of minus signs. Suppose they are denoted by S^+ and S^- respectively.

Step 3: When null hypothesis is true and the population is dichotomised on the basis of postulated value of median $\tilde{\mu}_0$ then we expect that the number of plus signs (success) and number of minus signs (failure) approximately equal. And number of plus signs or number of minus signs follows binomial distribution (n, p = 0.5). For convenient consider the smaller number of plus or minus signs. If plus sign (S^+) is less than minus sign (S^-) then we will take plus sign (S^+) as success and minus sign (S^-) as failure. Similarly, if minus sign (S^-) is less than plus sign (S^+) then we will assume minus sign (S^-) as success and plus sign (S^+) as failure.

The critical values of this test are not generally in tabular form and slightly difficult to obtain whereas p-value is easy to obtained with the help of cumulative binomial table so we use concept of p-value for take the decision about the null hypothesis.

Step 4: To take the decision about the null hypothesis, we use concept of p-value. We have discussed p-value in Unit 9 of this course. For p-value we determine the probability that test statistic is less than or equal to the calculated value of test statistic i.e. actually observed plus or minus signs. Since distribution of number of plus or minus signs is binomial (n, p = 0.5) therefore, this probability can be obtained with the help of **Table I** given in Appendix at the end of this block which provide the cumulative binomial probability and compare this probability with the level of significance (α). Here, test statistic depends upon the alternative hypothesis so the following cases arise:

For one-tailed test:

Case I: When $H_0: \tilde{\mu} \le \tilde{\mu}_0$ and $H_1: \tilde{\mu} > \tilde{\mu}_0$ (right-tailed test)

In this case, we expect that number of minus signs (S^-) is smaller than number of plus signs (S^+) therefore, the test statistic(S) is the number of minus signs (S^-). The p-value is determined as

$$\text{p-value} = P\left[S \le S^-\right]$$

If p-value is less than or equal to α, that is, $P\left[S \le S^-\right] \le \alpha$ then we reject the null hypothesis at α level of significance and if the p-value is greater than α then we do not reject the null hypothesis.

Case II: When $H_0: \tilde{\mu} \ge \tilde{\mu}_0$ and $H_1: \tilde{\mu} < \tilde{\mu}_0$ (left-tailed test)

In this case, we expect that number of plus signs (S^+) is smaller than number of minus signs (S^-) therefore, the test statistic(S) is the number of plus signs (S^+). The p-value is determined as

$$\text{p-value} = P\left[S \le S^+\right]$$

If p-value is less than or equal to α, that is, $P\left[S \le S^+\right] \le \alpha$ then we reject the null hypothesis at α level of significance and if the p-value is greater than α then we do not reject the null hypothesis.

Here, we take p-value as the probability less than or equal to observed value of the test statistic in each case because in this test we consider the smaller sign so critical region lies in lower tail.

For two-tailed test: When $H_0: \tilde{\mu} = \tilde{\mu}_0$ and $H_1: \tilde{\mu} \ne \tilde{\mu}_0$

For two tailed test, the test statistic(S) is the smaller of number of plus signs (S^+) and minus signs (S^-), that is,

$$S = \min\left\{S^+, S^-\right\}$$

and approximate p-value is determined as

$$\text{p-value} = 2P\left[S \le S^+\right]; \quad \text{if } S^+ \text{ is small}$$

$$\text{p-value} = 2P\left[S \le S^-\right]; \quad \text{if } S^- \text{ is small}$$

If p-value is less than or equal to α, then we reject the null hypothesis at α level of significance and if the p-value is greater than α then we do not reject the null hypothesis.

For large sample $(n > 20)$:

For a large sample size n greater than 20, we use normal approximation to binomial distribution with mean

$$E(S) = np = n \times \frac{1}{2} = \frac{n}{2} \qquad \ldots (1)$$

and variance

$$Var(S) = npq = n \times \frac{1}{2} \times \frac{1}{2} = \frac{n}{4} \qquad \ldots (2)$$

Therefore in this case, we use normal test (i.e. Z-test which is described in Unit 10 of Block 3 of this course). The test statistic of Z-test is given by

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

$$= \frac{S - \frac{n}{2}}{\sqrt{\frac{n}{4}}} \sim N(0,1) \qquad \text{[Using equations (1) and (2)]} \qquad \ldots (3)$$

After that, we calculate the value of test statistic Z and compare it with the critical value given in **Table 10.1** of the Unit 10 of this course at prefixed level of significance α. Take the decision about the null hypothesis as described in the Section 10.2 of Unit 10 of this course.

Note 1: Here, you may shock on the point that we used n > 20 for large sample instead of n > 30, because binomial distribution with p = 0.5 approximated by normal distribution less than 30.

Let us do some examples to become more user friendly with this test.

Example 1: The breaking strength (in pounds) of a random sample of 10 ropes made by a manufacturer is given by

163 165 165 160 171 158 151 162 169 172

Use the sign test to test the manufacturer's claim that the average breaking strength of a rope is greater than 160 pounds at 5% level of significance.

Solution: Here, distribution of the population of the breaking strengths of the ropes is not given. So the assumption of normality for t-test is not fulfilled. Also sample size is small so we can not use Z-test. So we go for sign test.

Here, we want to test the manufacturer's claim that the average (median) breaking strength $(\tilde{\mu})$ of a rope is greater than 160 pounds. So the claim is $\tilde{\mu} > 160$ and its complement is $\tilde{\mu} \le 160$. Since complement contains the equality sign so we can take the complement as the null hypothesis and the claim as the alternative hypothesis. Thus,

$$H_0 : \tilde{\mu} \le \tilde{\mu}_0 = 160 \text{ and } H_1 : \tilde{\mu} > 160$$

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

For applying sign test, we compare each observation with $\tilde{\mu}_0 (= 160)$ and replacing each observation greater than 160 with a plus sign and each observation less than 160 with a minus sign and discarding the one observation which equals to 160, we get

+ + + + − − + + +

By counting, we have

S^+ = number of plus signs = 7

S^- = number of mimus signs = 2

n = total number of plus and minus signs = 9

Since alternative hypothesis is right-tailed so the test statistic (S) is the number of minus signs (S^-), that is,

S = number of minus signs (S^-) = 2

Here, n = 9(< 20) so it is a case of small sample. Thus, to take the decision about the null hypothesis, we determine p-value with the help of **Table I** given in Appendix at the end of this block. Here, n = 9, p = 0.5 and r = 2. Thus, we have

$$\text{p-value} = P\left[S \le 2\right]$$
$$= 0.0899$$

Since p-value = $0.0899 > 0.05\left(=\alpha\right)$. So we do not reject the null hypothesis and reject the alternative hypothesis i.e. we reject the manufacturer's claim at 5% level of significance.

Thus, we conclude that the sample provide us sufficient evidence against the claim so manufacturer's claim that the breaking strength of a rope is greater than 160 pounds is not true.

Example 2: An economist believes that the median starting salary for a computer programmer in a certain city is Rs 12,000. To verify this claim, 12 computer programmers with similar backgrounds who recently got the jobs were randomly selected. Their starting salaries (in Rs) were 18000, 15000, 12000, 10000, 13000, 12000, 10000, 16000, 11000, 9000, 10000 and 9000. At 1% level of significance give your conclusion where you reached after using sign test.

Solution: Here, distribution of the salaries of the computer programmers is not given. So the assumption of normality for t-test is not fulfilled. Also sample size is small so we can not use Z-test. So we go for sign test.

Here, we want to test that the median starting salary $(\tilde{\mu})$ for a computer programmer in a certain city is Rs 12000. So our claim is $\tilde{\mu} = 12000$ and it complement is $\tilde{\mu} \ne 12000.$ Since claim contains the equality sign so we can take the claim as the null hypothesis and the complement as the alternative hypothesis. Thus,

$$H_0 : \tilde{\mu} = \tilde{\mu}_0 = 12000 \text{ and } H_1 : \tilde{\mu} \ne 12000$$

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

For applying sign test, we replacing each value greater than $\tilde{\mu}_0 (= 12000)$ with a plus sign and each value less than 12000 with a minus sign and discarding the two values which are equal to 12000, we get

$$+ \quad + \quad - \quad + \quad - \quad + \quad - \quad - \quad - \quad -$$

By counting, we have

S^+ = number of plus signs = 4

S^- = number of minus signs = 6

n = total number of plus and minus signs = 10

Since alternative hypothesis is two-tailed so the test statistic(S) is the minimum of number of plus signs (S^+) and minus signs (S^-), that is,

$$S = \min\left\{S^+, S^-\right\} = \min\left\{4, 6\right\} = 4$$

Here, n =10 (< 20), so it is a case of small sample. Thus, to take the decision about the null hypothesis, we determine p-value with the help of **Table I** given in Appendix at the end of this block. Here, n = 10, p = 0.5 and r = 4. Thus, we have

$$\text{p-value} = P\left[S \le 4\right] = 0.3770$$

Since p-value = 0.37707 > 0.01(α) so we do not reject the null hypothesis i.e. we support the claim at 1% level of significance.

Thus, we conclude that the sample fails to provide us sufficient evidence against the claim so we may assume that median salary of the newly appointed computer programmers is Rs 12000.

Example 3: The following data give the milk production (in thousand kg) in full cream by 40 different dairies:

17	15	20	29	19	19	22	25	27	9
24	20	17	6	24	14	15	23	24	26
19	23	28	19	16	22	24	17	20	13
19	10	23	18	31	13	20	17	24	14

Use the sign test to test that median $(\tilde{\mu})$ production of milk in dairies is 21.5 thousand kg at 1% level of significance.

Solution: Here, distribution of the milk production of the different dairies is not given. So the assumption of normality for t-test is not satisfied. So we go for sign test.

Here, we want to test that median production $(\tilde{\mu})$ of milk in dairies is 21.5 thousand kg. So our claim is $\tilde{\mu} = 21.5$ and it complement is $\tilde{\mu} \neq 21.5.$ Since claim contains the equality sign so we can take the claim as the null hypothesis and the complement as the alternative hypothesis. Thus,

$$H_0 : \tilde{\mu} = \tilde{\mu}_0 = 21.5 \text{ and } H_1 : \tilde{\mu} \neq 21.5$$

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

For applying sign test, we replacing each value greater than 21.5 (median value) with a plus sign and each value less than 21.5 with a minus sign, we get

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

By counting, we have

$$S^+ = \text{number of plus signs} = 16$$

$$S^- = \text{number of minus signs} = 24$$

n = total number of plus and minus signs = 40

Since alternative hypothesis is two-tailed so the test statistic (S) is given by

$$S = \min\left\{S^+, S^-\right\} = \min\left\{16, 24\right\} = 16$$

Here, n = 40 (> 20), so it is a case of large sample. In this case, we use Z-test. The test statistic of Z-test is given by

$$Z = \frac{S - \frac{n}{2}}{\sqrt{\frac{n}{4}}} \sim N(0,1)$$

$$= \frac{16 - \frac{40}{2}}{\sqrt{\frac{40}{4}}}$$

$$= \frac{16 - 20}{\sqrt{10}} = -1.27$$

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

Since calculated value of Z (= −1.27) is greater than the critical value (= − 2.58) and less than the critical value (= 2.58), that means it lies in non-rejection region, so we do not reject the null hypothesis i.e. we support the claim at 1% level of significance.

Decision according to p-value:

Since test is two-tailed, therefore,

p-value = $2P\left[Z \geq |z|\right] = 2P\left[Z \geq 1.27\right]$

$$= 2\left[0.5 - P\left[0 \leq Z \leq 1.27\right]\right]$$

$$= 2\left(0.5 - 0.3980\right) = 0.204$$

Since p-value (= 0.204) is greater than $\alpha (= 0.01)$ so we do not reject the null hypothesis i.e. we support the claim at1% level of significance.

Thus, we conclude that the sample fails to provide us sufficient evidence against the claim so we may assume that the median production of milk in dairies is 21.5 thousand kg.

Now, you can try the following exercises.

E2) Give one difference between t-test and sign test for one-sample.

E3) A Metropolitan Area Road Transport Authority claims that the average (median) waiting time on a well travelled rout is 20 minutes. A random sample of 12 passengers showed weighting times of 22, 13, 17, 14, 25, 26, 19, 20, 22, 30, 10 and15 minutes. Test the Metropolitan Area Road Transport Authority (MARTA) claim's that average (median) waiting time on a well traveled rout is 20 minutes by sign test at 1% level of significance.

E4) The following data shows the weight (in kg) of a random sample of 30 cadets of a centre:

49	46	57	40	45	57	50	65	34	50
58	46	47	42	53	67	40	52	49	66
53	48	68	40	53	49	61	54	48	38

Use sign test to examine whether the median weight of all the cadets of the centre is 50 kg at 5% level of significance.

13.4 WILCOXON SIGNED-RANK TEST

In the previous section, we discussed sign test. We recall that the sign test utilizes only information whether the difference between each observation and the postulated value of the median is positive or negative, that is, it considers only sign of differences. This test is fine if the information about the observation sample is available in ordinal scale only. But if the measurement of the observation is available interval or ratio scales then choice of sign test is not recommended. Because this test do not take into account the information available in terms of the magnitude of the differences. To overcome this drawback of the sign test Wilcoxon signed-rank test do the job for us, which takes into account the information of signs as well as of magnitude of differences. Since Wilcoxon signed-rank test use more information than sign test so it is more powerful than the sign test. The Wilcoxon signed-rank test is also used as an alternative of the t-test.

Assumptions

Wilcoxon signed-rank test requires following assumptions to work:

(i) The sample is selected from the population with unknown median.

(ii) The sampled population is symmetric about its median.

(iii) The variable under study is continuous.

(iv) The variable under study is measured on at least interval scale.

Let us discuss general procedure of this test:

Let $X_1, X_2, ..., X_{n'}$ be a random sample of size n' from a continuous symmetric population. Let $\tilde{\mu}$ be median of the population. Here, we wish to test the hypothesis about the specified value $\tilde{\mu}_0$ of population median $\tilde{\mu}$. So we can take the null and alternative hypotheses as

$$H_0 : \tilde{\mu} = \tilde{\mu}_0 \text{ and } H_1 : \tilde{\mu} \neq \tilde{\mu}_0 \qquad [\text{for two-tailed test}]$$

or

$$\left.\begin{array}{l} H_0 : \tilde{\mu} \leq \tilde{\mu}_0 \text{ and } H_1 : \tilde{\mu} > \tilde{\mu}_0 \\ H_0 : \tilde{\mu} \geq \tilde{\mu}_0 \text{ and } H_1 : \tilde{\mu} < \tilde{\mu}_0 \end{array}\right\} \qquad [\text{for one-tailed test}]$$

After setting null and alternative hypotheses, Wilcoxon signed-rank test involves following steps:

Step 1: We subtract $\tilde{\mu}_0$ from each observation and obtain the difference d_i with their plus and minus sign as

$$d_i = X_i - \tilde{\mu}_0 \quad \text{for all observations}$$

But when the observation X_i equal to $\tilde{\mu}_0$ give no information in terms of signs as well as magnitude so we exclude all such observations from the analysis part. Due to such observations our sample size reduces and let reduced sample size be denoted by n.

Step 2: After that, we find the absolute value of these d_i's obtained in Step 1 as $|d_1|, |d_2|, ..., |d_n|$.

Step 3: In this step, we are ranked $|d_i|$'s (obtained in Step 2) with respect to their magnitudes from smallest to largest, that is, the rank 1 is given to the smallest of $|d_i|$'s, rank 2 is given to the second smallest and so on up to the largest $|d_i|$'s. If several values are same (tied), we assign each the average of ranks they would have received if there were no repetition.

For more detailed about repeated (tied) ranks please go through Section 7.4 of the Unit 2 of MST-002.

Step 4: Now assign the signs to the ranks which the original differences have.

Step 5: Finally, we calculate the sum of the positive ranks (T^+) and sum of negative ranks (T^-) separately.

Under H_0, we expect approximately equal number of positive and negative ranks. And under the assumption that population under study is symmetric about its median we expect that sum of the positive ranks (T^+) and sum of negative ranks (T^-) are equal.

Step 6: Decision Rule:

To take the decision about the null hypothesis, the test statistic is the smaller of T^+ and T^-. And the test statistic is compared with the critical (tabulated) value for a given level of significance (α) under the condition that the null hypothesis is true. **Table II** given in Appendix at the end of this block provides the critical values of test statistic at α level of significance for both one-tailed and two-tailed tests. Here, test statistic depends upon the alternative hypothesis so the following cases arise:

For one-tailed test:

Case I: When $H_0: \tilde{\mu} \le \tilde{\mu}_0$ and $H_1: \tilde{\mu} > \tilde{\mu}_0$ (right-tailed test)

In this case, we expect that sum of negative ranks (T^-) is smaller than sum of positive ranks (T^+) therefore, the test statistic (T) is the sum of negative ranks (T^-).

If computed value of test statistic (T) is less than or equal to the critical value T_α at α level of significance, that is, $T \le T_\alpha$ then we reject the null hypothesis at α level of significance, otherwise we do not reject the null hypothesis.

Here, we consider that if test statistic less than or equal to critical value of the test statistic then we reject null hypothesis in each case because in this test we consider the smaller sum so critical region lies in lower tail.

Case II: When $H_0: \tilde{\mu} \ge \tilde{\mu}_0$ and $H_1: \tilde{\mu} < \tilde{\mu}_0$ (left-tailed test)

In this case, we expect that sum of positive ranks (T^+) is smaller than sum of negative ranks (T^-) therefore, the test statistic (T) is the sum of positive ranks (T^+).

If computed value of test statistic (T) is less than or equal to the critical value T_α at α level of significance, that is, $T \le T_\alpha$ then we reject the null hypothesis at α level of significance, otherwise we do not reject the null hypothesis.

For two-tailed test:

When $H_0: \tilde{\mu} = \tilde{\mu}_0$ and $H_1: \tilde{\mu} \ne \tilde{\mu}_0$

In this case, the test statistic (T) is the smaller of sum of positive ranks (T^+) and sum of negative ranks (T^-), that is,

$$T = \min\left\{T^+, T^-\right\}$$

If computed value of test statistic (T) is less than or equal to the critical value $T_{\alpha/2}$ at α level of significance , that is, $T \le T_{\alpha/2}$ then we reject the null hypothesis at α level of significance, otherwise we do not reject the null hypothesis.

For large sample (n > 25):

For a large sample size n greater than 25, the distribution of test statistic (T) approximated by a normal distribution with mean

$$E(T) = \frac{n(n+1)}{4} \qquad \ldots (4)$$

and variance

$$Var(T) = \frac{n(n+1)(2n+1)}{24} \qquad \ldots (5)$$

The proof of mean and variance of T is beyond the scope of this course.

Therefore in this case, we use normal test (Z-test). The test statistic of Z-test is given by

$$Z = \frac{T - E(T)}{SE(T)} = \frac{T - E(T)}{\sqrt{Var(T)}}$$

$$= \frac{T - \frac{n(n+1)}{4}}{\sqrt{\frac{n(n+1)(2n+1)}{24}}} \sim N(0,1) \left[\begin{array}{c}\text{Using equations} \\ \text{(4) and (5)}\end{array}\right] \qquad \ldots (6)$$

After that, we calculate the value of test statistic Z and compare it with the critical value given in **Table 10.1** at prefixed level of significance α. Take the decision about the null hypothesis as described in the Section 10.2 of Unit 10 of this course.

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

Example 4: A random sample of 15 children of one month or older shows the following pulse rates (beats per minute):

119, 120, 125, 122, 118, 117, 126, 114, 115, 126, 121, 120, 124, 127, 126

Assuming that the distribution of pulse rate is symmetric about its median and continuous, is there evidence to suggest that the median pulse rate of one month or older children is 120 beats per minute at 5% level of significance?

Solution: Here, distribution of pulse rate is not given. So the assumption of normality for t-test is not fulfilled although all the assumptions of Wilcoxon signed-rank test hold. So we go for this test.

Here, we want to test that median pulse rate $(\tilde{\mu})$ of children of one month or older is 120 beats per minute. So our claim is $\tilde{\mu} = 120$ and it complement is $\tilde{\mu} \neq 120.$ Since claim contains the equality sign so we can take the claim as the null hypothesis and the complement as the alternative hypothesis. Thus,

$$H_0 : \tilde{\mu} = \tilde{\mu}_0 = 120 \quad [\text{median pluse rate is equal to 120}]$$

$$H_1 : \tilde{\mu} \neq 120 \quad [\text{median pluse rate is not equal to 120}]$$

Since the alternative hypothesis is two-tailed so the test is two-tailed and the test statistic is the smaller of sum of positive ranks (T^+) and sum of negative ranks (T^-), that is,

$$T = \min\{T^+, T^-\}$$

Calculation for T:

S. No.	Beats per Minute (X)	Difference $d = X - \tilde{\mu}$ $= X - 120$	Absolute Value of Difference $\lvert d \rvert$	Rank of $\lvert d \rvert$	Signed Rank
1	119	−1	1	1.5	−1.5
2	120	Tie	---	---	---
3	125	5	5	7.5	7.5
4	122	2	2	3.5	3.5
5	118	−2	2	3.5	−3.5
6	117	−3	3	5	−5
7	126	6	6	10.5	10.5
8	114	−6	6	10.5	−10.5
9	115	−5	5	7.5	−7.5
10	126	6	6	10.5	10.5
11	121	1	1	1.5	1.5
12	120	Tie	---	---	---
13	124	4	4	6	6
14	127	7	7	13	13
15	126	6	6	10.5	10.5

From the above calculations, we have

$T^+ = 7.5 + 3.5 + 10.5 + 10.5 + 1.5 + 6 + 13 + 10.5 = 63.0$

$T^- = 1.5 + 3.5 + 5 + 10.5 + 7.5 = 28.0$

n = number of non-zero d_i's =13

Putting the values in test statistic, we have

$$T = \min\left\{T^+, T^-\right\} = \min\left\{63.0, 28.0\right\} = 28.0$$

The critical (tabulated) value for two-tailed test corresponding n = 13 at 5% level of significance is 18.

Since calculated value of test statistic T (= 28.0) is greater than the critical value (= 18) so we do not reject the null hypothesis i.e. we support the claim at 5% level of significance.

Thus, we conclude that the sample fails to provide us sufficient evidence against the claim so we may assume that median pulse rate of children of one month or older is 120 beats per minute.

Example 5: The following data show the weight (in kg) of a random sample of 30 cadets of a college:

49	46	57	37	45	57	50	65	34	50	58
46	47	42	53	67	40	52	49	66	53	48
68	40	53	49	61	54	48	38			

Assume that distribution of the weight of the cadets is symmetric about its median, test to examine whether the median weight of all the cadets of the college is 50 kg by useing Wilcoxon signed-rank at 5% level of significance.

Solution: Here, distribution of weight of the cadets is not given. So the assumption of normality for t-test is not fulfilled. We are given to us that distribution of the weight of the cadets is symmetric about its median and assumption of continuity holds because the characteristic weight is continuous in nature. So we go for Wilcoxon signed-rank test.

Here, we want to test that median weight $(\tilde{\mu})$ of all the cadets of a college is 50 kg. So our claim is $\tilde{\mu} = 50$ and it complement is $\tilde{\mu} \neq 50$. Since claim contains the equality sign so we can take the claim as the null hypothesis and the complement as the alternative hypothesis. Thus,

$$H_0 : \tilde{\mu} = \tilde{\mu}_0 = 50 \quad [\text{median weight is equal to } 50\,\text{kg}]$$

$$H_1 : \tilde{\mu} \neq 50 \quad [\text{median weight is not equal to } 50\,\text{kg}]$$

Since the alternative hypothesis is two-tailed so the test is two-tailed test and the test statistic is the smaller of sum of positive ranks (T^+) and sum of negative ranks (T^-), that is,

$$T = \min\{T^+, T^-\}$$

Calculation for T:

S. No.	Weight (in kg) (X)	Difference $d = X - \tilde{\mu}$ $= X - 50$	Absolute Value of Difference $\lvert d \rvert$	Rank of $\lvert d \rvert$	Signed Rank
1	49	−1	1	2	−2
2	46	−4	4	12	−12
3	57	7	7	15.5	15.5
4	37	−13	13	23	−23
5	45	−5	5	14	−14
6	57	7	7	15.5	15.5
7	50	Tie	---	---	---
8	65	15	15	24	24
9	34	−16	16	25.5	−25.5
10	50	Tie	---	---	---
11	58	8	8	17.5	17.5
12	46	−4	4	12	−12
13	47	−3	3	8.5	−8.5
14	42	−8	8	17.5	−17.5
15	53	3	3	8.5	8.5
16	67	17	17	27	27
17	40	−10	10	19.5	−19.5
18	52	2	2	5	5
19	49	−1	1	2	−2
20	66	16	16	25.5	25.5
21	53	3	3	8.5	8.5
22	48	−2	2	5	−5
23	68	18	18	28	28
24	40	−10	10	19.5	−19.5
25	53	3	3	8.5	8.5
26	49	−1	1	2	−2
27	61	11	11	21	21
28	54	4	4	12	12
29	48	−2	2	5	−5
30	38	−12	12	22	−22

From the above calculations, we have

$T^+ = 15.5 + 15.5 + 24 + 17.5 + 8.5 + 27 + 5 + 25.5 + 8.5 + 28 + 8.5 + 21 + 12$

$= 216.5$

T^{-} = 2 + 12 + 23 + 14 + 25.5 + 12 + 8.5 + 17.5 + 19.5 + 2 + 5 + 19.5 + 2 + 5 + 22

= 189.5

n = number of non-zero d_i's = 28

Putting the values in test statistic, we have

$$T = \min\left\{T^{+}, T^{-}\right\} = \min\left\{216.5, 189.5\right\} = 189.5$$

Also n = 28 (> 25) therefore, it is the case of large sample. So in this case, we use Z-test. The test statistic of Z-test is given by

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

where, $E(T) = \frac{n(n+1)}{4} = \frac{28(28+1)}{4} = 203$ and

$$SE(T) = \sqrt{\frac{n(n+1)(2n+1)}{24}} = \sqrt{\frac{28(28+1)(2\times 28+1)}{24}} = 43.91$$

Putting the values in test statistic Z, we have

$$Z = \frac{189.5 - 203}{43.91} = -0.31$$

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

Since calculated value of Z (= −0.31) is greater than the critical value (= − 1.96) and less than the critical value (= 1.96), that means it lies in non-rejection region, so we do not reject the null hypothesis i.e. we support the claim.

Thus, we conclude that the sample fails to provide us sufficient evidence against the claim so we may assume that the median weight of all the cadets of the college is 50 kg.

Now, you can try the following exercises in the same manner.

E5) Give one main difference between sign test and Wilcoxon signed-rank test.

E6) The observations of a random sample of size 10 from a distribution which is continuous and symmetric about its median are given below:

20.2, 24.1, 21.3, 17.2, 19.8, 16.5, 21.8, 18.7, 17.1, 19.9

Use Wilcoxon test to test the hypothesis that the sample is taken from a population having median greater than 18 at 5% level of significance.

E7) The breaking strength (in pounds) of a random sample of 29 rods made by a manufacturer is given as follows:

19	30	28	35	23	37	36	32	40	24
31	33	21	36	30	32	26	40	38	28
30	41	22	43	25	28	24	37	17	

On the basis of this sample test the manufacturer's claim that the average (median) breaking strength of the rope is 30 pounds by Wilcoxon singed-rank test at 1% level of significance. Assume that breaking strength of the rods is symmetric about its median.

A run is a succession of identical letters or symbols that is followed and preceded by a different letter or by no letter at all.

13.5 RUN TEST

One of the fundamental assumptions of the parametric test is that the observed data are random and test statistic and the subsequent analysis are based on this assumption. It is always better to check whether this assumption is true or not. A very simple tool for checking this assumption is run test. This section is devoted to throw light on the run test. Before discussing the run test first we have to explain what we mean by a "run".

A run in observations, is defined as a sequence of letters or symbols of one kind, immediately preceded and succeeded by letters of other kind or no letters. For example a sequence of two letters H and T as given below:

HHTHTTTHTHHHTTT

In this sequence, we start with first letter H and go up to other kind of letter, that is, T. In this way, we get first run of two H's. Then we start with this T and go up to other kind of letter, that is, H. Then we get a run of one T and so on and finally a run of three T's. In all, we see that there are eight runs as shown below:

$$\underbrace{HH}_{1}\underbrace{T}_{2}\underbrace{H}_{3}\underbrace{TTT}_{4}\underbrace{H}_{5}\underbrace{T}_{6}\underbrace{HHH}_{7}\underbrace{TTT}_{8}$$

Under run test, randomness of observations is judged by number of runs in the observed sequence. Too few runs indicate that there is some clustering or trend and too large runs indicate that there is some kind of repeated or cycles according to some pattern.

For example, the following sequence of H's and T's is obtained when tossing a coin 20 times:

$$\underbrace{HHHHHHHHHH}_{1}\underbrace{TTTTTTTTTT}_{2}$$

In this sequence, we have 2 runs-a run of 10 heads followed by a run of 10 tails. Since similar items tend to cluster together, therefore, such a sequence could not be considered random even though as would theoretically be expected 10 heads and 10 tails out of 20 tosses. In another example, suppose the following sequence has been obtained:

$$\underbrace{H}_{1}\underbrace{T}_{2}\underbrace{H}_{3}\underbrace{T}_{4}\underbrace{H}_{5}\underbrace{T}_{6}\underbrace{H}_{7}\underbrace{T}_{8}\underbrace{H}_{9}\underbrace{T}_{10}\underbrace{H}_{11}\underbrace{T}_{12}\underbrace{H}_{13}\underbrace{T}_{14}\underbrace{H}_{15}\underbrace{T}_{16}\underbrace{H}_{17}\underbrace{T}_{18}\underbrace{H}_{19}\underbrace{T}_{20}$$

In this sequence there are 20 runs-10 runs of one head each and 10 runs of one tail each. This sequence also could not be considered random because too many runs indicate a pattern.

On the other hand, if the sequence of H's and T's is thoroughly mixed, such as,

$$\underbrace{H}_{1}\underbrace{TTT}_{2}\underbrace{HH}_{3}\underbrace{TTT}_{4}\underbrace{H}_{5}\underbrace{T}_{6}\underbrace{H}_{7}\underbrace{TT}_{8}\underbrace{HHH}_{9}\underbrace{T}_{10}\underbrace{HH}_{11}$$

where neither the number of runs too small nor too large, this type of sequence may be considered random. Therefore, in testing for randomness, **ordering** or **positioning** of the items in the sequence is essential, **not** the frequency of the items of each type.

Assumptions

Run test make the following assumptions:

(i) Observed data should be such that we can categorise the observations into two mutually exclusive types.

(ii) The variable under study is continuous.

Let $X_1, X_2, ..., X_{n'}$ be a set of n' observations arranged in the order in which they occur. Generally, we are interested to test whether a population or a sample or a sequence is random or not. So here we consider only two-tailed case. Thus, we can take the null and alternative hypotheses as

H_0 : The observations are random

H_1 : The observations are not random [two-tailed test]

Let us discuss the general procedure of this test:

Step 1: First of all, we check the form of the given dada that the given data are in symbolical form such as sequence of H and T, A and B, etc. or in the numeric form. If the data in symbolical form then it is ok, but if data in numeric form then first we convert numeric data in symbolical form. For this, we calculate median of the given observations by using either of the following formula (discussed in Unit 1 of the course MST-002.) given below:

$$\text{Median} = \left(\frac{n'+1}{2}\right)^{th} \text{observation}; \qquad \text{if n' is odd}$$

$$\text{Median} = \frac{1}{2}\left[\left(\frac{n'}{2}\right)^{th} \text{observation} + \left(\frac{n'}{2}+1\right)^{th} \text{observation}\right]; \quad \text{if n' is even}$$

provided observations should be either in ascending or descending order of magnitude.

After that, we replace the observations which are above the median by a symbol 'A' (say) and the observation which are below the median by a symbol 'B' (say) without altering the observed order. The observations which are equal to median are discarded form the analysis and let reduced size of the sample denoted by n.

Step 2: Counts number of times first symbol (A) occurs and denote it by n_1.

Step 3: Counts number of times second symbol (B) occurs and denote it by n_2 where, $n_1 + n_2 = n$.

Step 4: For testing the null hypothesis, the test statistic is the total number of runs so in this step we count total number of runs in the sequence of symbols and denote it by R.

Step 5: Obtain critical values of test statistic corresponding n_1, n_2 at α % level of significance under the condition that null hypothesis is true. Part I and Part II of **Table III** in the Appendix respectively provide lower (R_L) and upper (R_U) critical values of the number of runs for a given combination of n_1 and n_2 at 5% level of significance.

Note1: Generally, critical values for run test are available at 5% level of significance so we test our hypotheses for 5% level of significance.

Step 6: **Decision Rule:**

To take the decision about null hypothesis, the test statistic is compared with the critical (tabulated) values.

If the observed number of runs(R) is either less than or equal to the lower critical value (R_L) or greater than or equal to the upper critical value (R_U), that is, if $R \le R_L$ or $R \ge R_U$ then we reject the null hypothesis at 5% level of significance.

If R lies between R_L and R_U, that is, $R_L < R < R_U$, then we do not reject the null hypothesis at 5% level of significance.

For large sample (n > 40):

For a sample size n greater than 40 or either n_1 or n_2 exceeds 20, the statistic R is approximately normally distributed with mean

$$E(R) = \frac{2n_1n_2}{n} + 1 \qquad \text{... (7)}$$

and variance

$$Var(R) = \frac{2n_1n_1(2n_1n_2 - n)}{n^2(n-1)} \qquad \text{... (8)}$$

The proof of mean and variance of R is beyond the scope of this course.

Therefore in this case, we use normal test (Z-test). The test statistic of Z-test is given by

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

$$Z = \frac{R - \frac{2n_1n_2}{n} - 1}{\sqrt{\frac{2n_1n_2(2n_1n_2 - n)}{n^2(n-1)}}} \quad \text{[Using equations (7) and (8)]} \qquad \text{... (9)}$$

After that, we calculate the value of test statistic Z and compare it with the critical value given in **Table 10.1** at prefixed level of significance α. Take the decision about the null hypothesis as described in the Section 10.2 of Unit 10 of this course.

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

Example 6: A sequence of Heads (H) and Tails (T) in tossing of a coin 16 times is given below:

HTTTHTHTHHTHTTHH

(i) Count the number of runs.

(ii) Test whether the Heads and Tails occur in random order.

Solution: Here, we are given a sequence of Heads of Tails as

HTTTHTHTHHTHTTHH

(i) For counting the number of runs, we start with first letter H and go up to other kind of letter, that is, T. In this way, we get first run of one H. Then we start with this T and go up to other kind of letter, that is, H then we get a run of three T's and so on in the last there is a run of two H's. In this way, the total number of runs is 11 as shown below:

$$\underbrace{H}_{1}\underbrace{TTT}_{2}\underbrace{H}_{3}\underbrace{T}_{4}\underbrace{H}_{5}\underbrace{T}_{6}\underbrace{HH}_{7}\underbrace{T}_{8}\underbrace{H}_{9}\underbrace{TT}_{10}\underbrace{HH}_{11}$$

(ii) Here, we want to test the randomness of the sequence so we go for run test.

We want to test that the given sequence of H and T is in random order. So our claim is "the given sequence of H and T is in random order" and its complement is "the given sequence of H and T is not in random order". So we can take claim as the null hypothesis and complement as the alternative hypothesis. Thus,

H_0 : Given sequence of Heads (H) and Tails (T) is in random order

H_1 : Given sequence of Heads (H) and Tails (T) is not in random order

Here by counting, we have

n_1 = number of times H occurs = 8

n_2 = number of times T occurs = 8

R = total number of runs in given sequence = 11

For run test, the test statistic is the total number of runs in the given sequence.

Since $n_1 = 8$ (< 20) and $n_2 = 8$ (< 20), so it is a small sample case.

The lower and upper critical values of runs corresponding to $n_1(=8)$ and $n_2(=8)$ at 5% level of significance are $R_L = 4$ and $R_U = 14$.

Since observed number of runs (R = 11) lies between critical values of runs, $R_L = 4$ and $R_U = 14$, so we do not reject the null hypothesis i.e. we support the claim at 5% level of significance.

Thus, we conclude that the given sequence of H and T is in random order.

Example 7: A machine produced a number of defective items per day. To check the randomness of the defective items, a sample of defective items produced per day over a period of 20 days is taken and found following results:

13	17	14	20	18	17	16	14	19	21
18	20	17	13	14	19	20	15	19	17

Test at 5% level of significance that the machine produced defective items in random order.

Solution: Here, we want to test the randomness of the defective items produced per day so we go for run test.

We want to test that machine produced defective items in random order. So our claim is "the machine produced defective items are in random order" and its complement is "the machine produced defective items are not in random order". So we can take claim as the null hypothesis and complement as the alternative hypothesis. Thus,

H_0 : The machine produced defective items are in random order

H_1 : The machine produced defective items are not in random order

For applying run test, we need given observations in a sequence of two symbols. In order to convert theses observations into a sequence of two symbols first of all, we have to calculate median of the given observations. For obtaining median, we have to arrange the given observations in ascending or descending order of magnitude. So arrange the observations in ascending order of magnitude, we have

13 13 14 14 14 15 16 17 17 17 17 18 18 19 19 19 20 20 20 21

Number of observations $(n') = 20$ (even) therefore the median can be obtained as

$$\text{Median} = \frac{1}{2}\left[\left(\frac{n'}{2}\right)^{th} \text{observation} + \left(\frac{n'}{2}+1\right)^{th} \text{observation}\right]$$

$$= \frac{1}{2}\left[\left(\frac{20}{2}\right)^{th} \text{observation} + \left(\frac{20}{2}+1\right)^{th} \text{observation}\right]$$

$$= \frac{1}{2}\left[(10)^{th} \text{observation} + (11)^{th} \text{observation}\right]$$

$$= \frac{1}{2}(17+17) = 17$$

Now, we replace the given observations by the symbol 'A' if it is above the median value (= 17) and by symbol 'B' if it is below the median value and discard the observation equal to median value. In this way, the sequence obtained by discarding 4 observations which are equal to median value 17 is given below:

BBAABBAAAABBAABA

For applying run test, we count the number of runs. We see that there is a first run of two B's then there is a run of two A's and so on in the last there is a run of single A as shown below:

$$\underbrace{BB}_{1}\underbrace{AA}_{2}\underbrace{BB}_{3}\underbrace{AAAA}_{4}\underbrace{BB}_{5}\underbrace{AA}_{6}\underbrace{B}_{7}\underbrace{A}_{8}$$

By counting, we have

n_1 = number of symbol A occurs = 9

n_2 = number of symbol B occurs = 7

R = total number of runs in above sequence = 8

For run test, the test statistic is the total number of runs in the given sequence.

Since $n_1 = 9(< 20)$ and $n_2 = 7$ (< 20) so it is a small sample case.

The lower and upper critical values of runs corresponding to $n_1 = 9$ and $n_2 = 7$ and at 5% level of significance are $R_L = 4$ and $R_U = 14$.

Since observed number of runs (R = 9) lies between critical values of runs, $R_L = 4$, and $R_U = 14$ so we do not reject the null hypothesis i.e. we support the claim at 5% level of significance.

Thus, we conclude that the sample fails to provide us sufficient evidence against the claim so we may assume that the machine produced defective items are in random order.

Example 8: A sample of 36 tools produced by a machine shows the following sequence of good (G) and bad (B) tools:

GGGBGGGGGBBBGGGGGGBBGGGGGGGGGBBGGGGG

Test the machine produced good (G) and bad (B) tools in random order at 5% level of significance.

Solution: Here, we want to test the randomness of the sequence so we go for run test.

We want to test that machine produced good (G) and bad (B) tools in random order. So our claim is "the machine produced good and bad tools are in random order" and its complement is "the machine produced good and bad tools are not in random order". So we can take claim as the null hypothesis and complement as the alternative hypothesis. Thus,

H_0 : The machine produced good and bad tools are in random order

H_1 : The machine produced good and bad tools are not in random order

For applying run test, we count the number of runs.

We see that there is a first run of three G's then there is a run of single B and so on in the last there is a run of five G's as shown below:

$$\underbrace{GGG}_{1}\underbrace{B}_{2}\underbrace{GGGGG}_{3}\underbrace{BBB}_{4}\underbrace{GGGGGG}_{5}\underbrace{BB}_{6}\underbrace{GGGGGGGGG}_{7}\underbrace{BB}_{8}\underbrace{GGGGG}_{9}$$

By counting, we have

n_1 = number of symbol A occurs = 28

n_2 = number of symbol B occurs = 8

R = total number of runs in above sequence = 9

Since n_1 = 28 (>20) therefore, it is the case of large sample so we use Z-test. The test statistic of Z-test is given by

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

where, $E(R) = \frac{2n_1n_2}{n} + 1 = \frac{2 \times 28 \times 8}{36} + 1 = 13.44$ and

$$SE(R) = \sqrt{\frac{2n_1n_2(2n_1n_2 - n)}{n^2(n-1)}} = \sqrt{\frac{2 \times 28 \times 8(2 \times 28 \times 8 - 36)}{36^2(36-1)}}$$

$$= \sqrt{4.07} = 2.02$$

Putting the values in test statistic Z, we have

$$Z = \frac{9 - 13.44}{2.02} = -2.20$$

The critical (tabulated) values for two-tailed test at 5% level of significance are $\pm z_{\alpha/2} = \pm z_{0.005} = \pm 1.96$.

Since calculated value of Z (= −2.20) is less than critical values (= ± 1.96), that means it lies in rejection region so we reject the null hypothesis i.e. we reject the claim at 5% level of significance.

Thus, we conclude that the sample provides us sufficient evidence against the claim so the machine produced good and bad tools are not in random order.

Now, you can try the following exercises.

E8) The following observations were taken from a table of random numbers from a distribution F(x) with median 0:

0.464	0.137	2.455	– 0.323	– 0.068
0.907	–0.513	– 0.525	0.595	0.886
– 0.482	1.678	– 0.057	–1.229	– 0.486
–1.787	–0.261	1.237	1.046	0.962

Test whether the data are random at 5% level of significance?

E9) The cars that enter in a parking are classified either Foreign-made (F) or Indian-made (I). To check that the car entre in the parking are in random order, a sample of first 40 cars that entered in the parking is taken and found the following sequence:

IIFFFIFIFFFFIIFIIIFFIFFIFFFFIIFFFIFFIIII

Test that the cars enter in the parking are in random order at 1% level of significance.

13.6 KOLMOGOROV-SMIRNOV GOODNESS OF FIT TEST

This test has its name on the names of its discovers A. N. Kolmogorov and N.V. Smirnov. It is a simple non-parametric test for testing whether data follow a specified or assumed distribution or sample has come from a specified or assumed distribution or there is a significant difference between an observed distribution and a theoretical distribution. Therefore, it is a measure of goodness of fit to a theoretical distribution. We also use chi-square test for goodness of fit (will be described in Unit 16 of this block). The main difference between chi-square test and Kolmogorov-Smirnov (K-S) test is that the chi-square test is designed for categorical data whereas K-S test is designed for the continuous data.

Assumptions

This test works under the following assumptions:

(i) The sample is randomly selected from some unknown distribution.

(ii) The observations are independent.

(iii) The variable under study is continuous.

(iv) The variable under study is measured on at least ordinal scale.

Let $X_1, X_2,..., X_n$ be a random sample from a population with unknown continuous distribution function F(x). Generally, we are interested to test whether data follow a specified distribution $F_0(x)$ or a sample has come from a specified or assumed distribution or not. So here we consider only two-tailed case. Thus, we can take the null and alternative hypotheses as

H_0: Data follow a specified distribution

H_1: Data do not follow a specified distribution

In symbolical form

$$H_0 : F(x) = F_0(x) \quad \text{for all values of x}$$

$$H_1 : F(x) \neq F_0(x) \quad \text{for at least one value of x} \quad [\text{two-tailed test}]$$

We summarise the Kolmogorov-Smirnov goodness of fit test in the following steps:

Step 1: K-S goodness of fit test is based on the comparison of empirical (observed or sample) cumulative distribution function with the theoretical (specified or population) cumulative distribution function. Therefore first of all, we compute the empirical cumulative distribution function, say, S(x) from the sample data which is defined as the proportion of sample observations less than or equal to some value x, that is,

S(x) = proportion of sample observations less than or equal to some value x

$$S(x) = \frac{\text{The number of sample observation less than or equal to x}}{\text{Total number of observations}}$$

Step 2: In this Step, we find the theoretical cumulative distribution function $F_0(x)$ for all possible values of x on the basis of specified distribution.

Step 3: After finding the empirical and theoretical cumulative distribution functions for all possible values of x, we find the deviation between the empirical and theoretical cumulative distribution functions for all x. That is,

$$S(x) - F_0(x) \quad \text{for all x}$$

Step 4: Since S(x) is the statistical image of the population distribution function F(x) so if null hypothesis is true then the difference between S(x) and F(x) should be small for all values of x. Thus, in all deviations obtained in Step 3, we find the point(s) at which the two functions show the maximum deviation.

The test statistic is denoted by D_n and it is the greatest vertical deviation between S(x) and $F_0(x)$, that is,

$$D_n = \sup_x \left| S(x) - F_0(x) \right|$$

which is read as "D_n equals the supreme over all x, of the absolute value of the difference $S(x) - F_0(x)$"

Step 5: Obtain critical value of test statistic at α % level of significance under the condition that null hypothesis is true. **Table IV** in Appendix provides the critical values at α level of significance.

Step 6: Decision Rule:

To take the decision about the null hypothesis, the test statistic (calculated in Step 4) is compared with the critical (tabulated) value (obtained in Step 5) for a given level of significance (α).

If computed value of D_n is greater than or equal to critical value $D_{n,\alpha}$ at α level of significance, that is, if $D_n \geq D_{n,\alpha}$ then we reject the null hypothesis at α level of significance, otherwise we do not reject H_0.

For large sample (n > 40):

For a sample size n greater than 40, the critical value of test statistic at given α level of significance, is approximated by the formula given in last row of **Table IV** in the Appendix. As the critical value for two-tailed test at 5% level of significance is calculated by the formula given by

$$D_{n,0.05} = \frac{1.36}{\sqrt{n}}$$

where, n is the sample size.

Let us do some examples to become more user friendly with the calculations explained above:

Example 9: A coin is tossed 400 times in sets of four. The frequencies of getting 0 to 4 tails are:

Number of Tails	0	1	2	3	4	Total
Frequency	23	95	164	92	26	400

Examine the fairness of coin using K-S test at 5% level of significance.

Solution: Here, we want to test that the coin is fair. We know that if a coin is fair then number of tails follows the binomial distribution with parameter n and p = 1/2. So we can test that the number of tails follows the binomial distribution. If $F_0(x)$ denotes the cumulative distribution function of binomial distribution with parameter n = 4 and p = 1/2 then our claim is $F(x) = F_0(x)$ and its complement is $F(x) \neq F_0(x)$. Since the claim contains the equality sign so we can take claim as the null hypothesis and complement as the alternative hypothesis. Thus

$$H_0 : F(x) = F_0(x) \text{ for all values of x}$$

$$H_1 : F(x) \neq F_0(x) \text{ for at least one value of x}$$

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

$$D_n = \sup_x \left| S(x) - F_0(x) \right|$$

where, S(x) is the empirical distribution function based on the observed sample.

For obtaining the empirical cumulative distribution function S(x), first we obtain cumulative frequencies and then by the definition of empirical distribution function we obtain S(x) as

$$S(x) = \frac{\text{The number of sample observations} \leq x}{\text{Total number of observations}}$$

Therefore, at x = 0, 1, 2, 3, 4, we have

$$S(0) = \frac{23}{400} = 0.0575,\ S(1) = \frac{118}{400} = 0.2950,\ S(2) = \frac{282}{400} = 0.705,$$

$$S(3) = \frac{374}{400} = 0.935,\quad S(4) = \frac{400}{400} = 1$$

The theoretical distribution can be obtained by the definition of cumulative distribution function as

$$F_0(x) = P[X \le x]$$

In this case, our theoretical distribution is binomial distribution with parameter n = 4 and p = 1/2. Therefore,

i.e. $$F_0(x) = P[X \le x] = {}^4C_x \left(\frac{1}{2}\right)^x \left(\frac{1}{2}\right)^{4-x}$$

$$= {}^4C_x \left(\frac{1}{2}\right)^4, x = 0,1,2,3,4$$

With the help of **Table I** given in Appendix at the end of this block we can find the values of $F_0(x)$ at different values of x.

Thus, at x = 0, we have

$$F_0(0) = P[X \le 0] = 0.0625$$

Similarly,

$$\text{At } x = 1, F_0(1) = P[X \le 1] = 0.3125$$

$$\text{At } x = 2, F_0(2) = P[X \le 2] = 0.6875$$

$$\text{At } x = 3, F_0(3) = P[X \le 3] = 0.9375$$

$$\text{At } x = 4, F_0(4) = P[X \le 4] = 1$$

Calculation for $|S(x) - F_0(x)|$:

Number of Tails (x)	**Observed Frequency**	**Cumulative Observed Frequency**	**Observed Distribution Function S(x) = (3)/n**	**Theoretical Distribution Function**	**Absolute Difference $\|S(x) - F_0(x)\|$**
(1)	(2)	(3)	(4)	(5)	(6)
0	23	23	0.0575	0.0625	0.0050
1	95	118	0.2950	0.3125	0.0175
2	164	282	0.7050	0.6875	0.0175
3	92	374	0.9350	0.9375	0.0025
4	26	400	1	1	0

From the above calculation, we have

$$D_n = \sup_{0 \le x \le 4} |S(x) - F_0(x)| = 0.0175$$

Here, n = 400 > 40, therefore, critical value of test statistic can be obtained by the formula given by

$$D_{n,0.05} = \frac{1.36}{\sqrt{n}}$$

$$= \frac{1.36}{\sqrt{400}} = 0.068$$

Since calculated value of test statistic D_n (= 0.0175) is less than critical value $D_{n,0.05}$ (= 0.068) so we do not reject the null hypothesis i.e. we support the claim at 5% level of significance.

Thus, we conclude that the sample fails to provide us sufficient evidence against the claim so we may assume that number of tails follows the binomial distribution or the coin is fair or unbiased.

The Kolmogorov-Smirnov test for goodness of fit is applied only when all the parameters of the fitting distribution are known. If all parameter (s) of the fitted distribution is (are) not known then we use chi-square goodness of fit

Example 10: The following data were obtained from a table of random numbers of normal distribution with mean 3 and standard deviation 1:

2.1, 1.9, 3.2, 2.8, 1.0, 5.1, 4.2, 3.6, 3.9, 2.7

Test by K-S test that the data have come form the same normal distribution at 5% level of significance.

Solution: Here, we want to test that the data have come from normal distribution with mean 3 and variance 1. If $F_0(x)$ is the cumulative distribution function of N(3,1) then our claim is $F(x) = F_0(x)$ and its complement is $F(x) \neq F_0(x)$. Since the claim contains the equality sign so we can take claim as the null hypothesis and complement as the alternative hypothesis. Thus

$$H_0 : F(x) = F_0(x) \text{ for all values of x}$$

$$H_1 : F(x) \neq F_0(x) \text{ for at least one value of x}$$

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

$$D_n = \sup_x |S(x) - F_0(x)|$$

where, S(x) is the empirical distribution function based on the observed sample.

For obtaining the empirical cumulative distribution function S(x), first we obtain cumulative frequencies and then by the definition of empirical distribution function we obtain S(x) as

$$S(x) = \frac{\text{The number of sample observations} \leq x}{\text{Total number of observations}}$$

Here, theoretical cumulative distribution function $F_0(x)$ can be obtained with the help of the method as described in Unit 4 of MST-003. Therefore, by the definition of distribution function, we have

$$F_0(x) = P[X \leq x] = P\left[\frac{X-\mu}{\sigma} \leq \frac{x-\mu}{\sigma}\right] = P[Z \leq z]$$

where, $P[Z \leq z]$ can be to obtained with he help of **Table I** given in Appendix at the end of the Block 1 of this course.

Thus, when x = 1.0, $z = \frac{x-\mu}{\sigma} = \frac{1.0-3}{1} = -2.0$ and

$$F_0(1.0) = P[Z \leq z] = P[Z \leq -2.0] = 0.5 - P[-2.0 < Z < 0]$$

$$= 0.5 - P[0 < Z < -2.0]$$

$$= 0.5 - 0.4772 = 0.0228$$

Similarly, when x = 3.2, $z = \frac{3.2-3}{1} = 0.2$ and

$$F_0(3.2) = P[Z \le z] = P[Z \le 0.2]$$
$$= 0.5 + P[0 < Z \le 0.2]$$
$$= 0.5 + 0.0793 = 0.5793 \text{ and so on.}$$

Calculation for $|S(x) - F_0(x)|$:

X	Observed Frequency	Cumulative Observed Frequency	Observed Distribution Function S(x)	$Z = \frac{X-\mu}{\sigma} = \frac{X-3}{1}$	Theoretical Distribution Function $F_0(x)$	Absolute Difference $\lvert S(x)-F_0(x)\rvert$
1.0	1	1	1/10 = 0.1	–2.0	0.0228	0.0772
1.9	1	2	2/10 = 0.2	–1.1	0.1562	0.0438
2.1	1	3	3/10 = 0.3	–0.9	0.1841	0.1159
2.7	1	4	4/10 = 0.4	–0.3	0.3821	0.0179
2.8	1	5	5/10 = 0.5	–0.2	0.4207	0.0793
3.2	1	6	6/10 = 0.6	0.2	0.5793	0.0207
3.6	1	7	7/10 = 0.7	0.6	0.7257	0.0257
3.9	1	8	8/10 = 0.8	0.9	0.8159	0.0159
4.2	1	9	9/10 = 0.9	1.2	0.8849	0.0151
5.1	1	10	10/10 = 1.0	2.1	0.9821	0.0179

From the above calculations, we have

$$D_n = \sup_x |S(x) - F_0(x)| = 0.1159$$

The critical value of test statistic at 5% level of significance is 0.409.

Since calculated value of test statistic D_n (= 0.1159) is less than critical value (= 0.409) so we do not reject the null hypothesis i.e. we support the claim at 5% level of significance.

Thus, we conclude that the sample fails to provide us sufficient evidence against the claim so we may assume that the data have come from the normal distribution with mean 3 and standard deviation 1.

Now, you can try the following exercises.

E10) Write the main difference between chi-square test and Kolmogorov-Smirnov (K-S) test for goodness of fit.

E11) The following data were obtained from a table of random numbers of normal distribution with mean 5.6 and standard deviation 1.2:

6.0	6.6	5.9	5.0	6.2	4.8
5.5	6.3	5.9	5.8	6.6	6.2

Test the hypothesis by K-S test that the data have come from the same normal distribution at 5% level of significance.

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

13.7 SUMMARY

In this unit, we have discussed following points:

1. Need of non-parametric tests.
2. Applications of non-parametric tests in various fields.
3. Advantages and disadvantages of non-parametric tests over parametric tests.
4. The one-sample non-parametric tests.
5. The suitable non-parametric test is appropriate for a particular situation.
6. Test the hypothesis when we cannot make any assumption about the distribution from which we are sampling.
7. The tests which are used in place of t-test for mean when assumption(s) of the t-test is (are) not fulfilled, that is, sign test and Wilcoxon signed-rank test.
8. The test for randomness of data, that is, run test.
9. The Kolmogorov-Smirnov test for goodness of fit.

13.8 SOLUTIONS /ANSWERS

E1) Refer Section 13.2.

E2) The main difference between t-test and sign test is that the t-test based on the assumption that the parent population is normal where as sign test is not based on such assumption.

E3) Here, distribution of the waiting time on a well travelled rout is not given. So the assumption of normality for t-test is not fulfilled. Also sample size is small so we can not use Z-test. So we go for sign test.

Here, we wish to test the claim of Metropolitan Area Road Transport Authority that the average (median) waiting time $(\tilde{\mu})$ on a well travelled rout is 20 minutes. So our claim is $\tilde{\mu} = 20$ and it complement is $\tilde{\mu} \neq 20$. Since claim contains the equality sign so we can take the claim as the null hypothesis and the complement as the alternative hypothesis. Thus,

$$H_0 : \tilde{\mu} = \tilde{\mu}_0 = 20 \text{ and } H_1 : \tilde{\mu} \neq 20$$

For applying sign test, we replacing each observation greater than 20 (median value) with a plus sign and each observation less than 20 with a minus sign and discarding the one observation which equal to 20, we get

$$+ \quad - \quad - \quad - \quad + \quad + \quad - \quad + \quad + \quad - \quad -$$

By counting, we have

S^+ = number of plus signs = 5

S^- = number of mines signs = 6

n = total number of plus and minus signs = 11

Since alternative hypothesis is two-tailed so the test statistic (S) is given by

$$S = \min\{S^{+}, S^{-}\}$$

$$= \min\{5, 6\} = 5$$

Here, n = 11(< 20) so it is a case of small sample. Thus, to take the decision about the null hypothesis, we determine p-value with the help of **Table I** given in Appendix at the end of this block. Here, n = 11, p = 0.5 and r = 5. Thus, we have

$$\text{p-value } = P\left[S \le 5\right] = 0.50$$

Since p-value = 0.50 > 0.01 (α). So we do not reject the null hypothesis i.e. we support the claim at 1% level of significance.

Thus, we conclude that the sample fails to provide us sufficient evidence against the claim so we may assume that the claim of Metropolitan Area Road Transport Authority that the average waiting time is 20 minutes is true.

E4) Here, distribution of the weight of the cadets is not given. So the assumption of normality for t-test is not fulfilled. So we go for sign test.

Here, we want to test that median weight $(\tilde{\mu})$ of all the cadets of a college is 50 kg. So our claim is $\tilde{\mu} = 50$ and it complement is $\tilde{\mu} \ne 50$. Since claim contains the equality sign so we can take the claim as the null hypothesis and the complement as the alternative hypothesis. Thus,

$$H_0 : \tilde{\mu} = \tilde{\mu}_0 = 50 \quad \left[\text{median weight is equal to 50 kg}\right]$$

$$H_1 : \tilde{\mu} \ne 50 \quad \left[\text{median weight is not equal to 50 kg}\right]$$

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

For applying sign test, we replacing each observation greater than 50 (median value) with a plus sign and each observation less than 50 with a minus sign and discarding the two observations which are equal to 50, we get

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

By counting, we have

S^{+} = number of plus signs = 13

S^{-} = number of minus signs = 15

n = total number of plus and minus signs = 28

Since alternative hypothesis is two-tailed so the test statistic(S) is the minimum of number of plus signs (S^{+}) and minus signs (S^{-}), that is,

$$S = \min\{S^{+}, S^{-}\} = \min\{13, 15\} = 13$$

Here, n = 28 (> 20), so it is a case of large sample. In this case. we the use Z-test. The test-statistic of Z-test is given by

$$Z = \frac{S - \frac{n}{2}}{\sqrt{\frac{n}{4}}} = \frac{13 - \frac{28}{2}}{\sqrt{\frac{28}{4}}}$$

$$= \frac{13 - 14}{\sqrt{7}} = 0.38$$

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

Since calculated value of test statistic Z (= 0.38) is greater than the critical value (= − 1.96) and less than the critical value (= 1.96), that means it lies in non-rejection region, so we do not reject the null hypothesis i.e. we support the claim at 5% level of significance.

Thus, we conclude that the samples do not provide us sufficient evidence against the claim so we may assume that the median weight of all the cadets of the centre is 50 kg.

E5) The main difference between sign test and Wilcoxon signed-rank test is that the sign test utilizes only information whether the difference between observation and the postulated value of the median is positive or negative, that is, it considers only sign of differences but ignores the magnitude of these differences whereas the Wilcoxon signed-rank test uses both the signs as well as the magnitude of the differences.

E6) We are given that characteristic under study is continuous and population is symmetric about its median and other assumptions of Wilcoxon signed-rank test hold. So we can apply this test.

Here, we want to test that sample is taken from the population having median $(\tilde{\mu})$ greater than 18. So our claim is $\tilde{\mu} > 18$ and it complement is $\tilde{\mu} \le 18.$ Since complement contains the equality sign so we can take the complement as the null hypothesis and the claim as the alternative hypothesis. Thus,

$$H_0 : \tilde{\mu} \le \tilde{\mu}_0 = 18 \text{ and } H_1 : \tilde{\mu} > 18$$

Since the alternative hypothesis is right-tailed so the test is right-tailed test and the test statistic (T) is the sum negative ranks (T^-).

Calculation for T:

S. No.	Sample (X)	Difference $d = X - \tilde{\mu} = X - 18$	Absolute Value of Difference $\|d\|$	Rank of $\|d\|$	Signed Rank
1	20.2	2.2	2.2	7	7
2	24.1	6.1	6.1	10	10
3	21.3	3.3	3.3	8	8
4	17.2	−0.8	0.8	2	−2
5	19.8	1.8	1.8	5	5
6	16.5	−1.5	1.5	4	−4
7	21.8	3.8	3.8	9	9
8	18.7	0.7	0.7	1	1
9	17.1	−0.9	0.9	3	−3
10	19.9	1.9	1.9	6	6

Therefore, we have

$T^+ = 7 + 10 + 8 + 5 + 9 + 1 + 6 = 46$

$T^- = 2 + 4 + 3 = 9$

n = number of non-zero d_i's =10

So the value of test statistic T = 9

The critical value of test statistic T for right-tailed test corresponding n = 10 at 5% level of significance is 11.

Since calculated value of test statistic T (= 9) is less than the critical value (= 11) so we reject the null hypothesis and support the alternative hypothesis i.e. we support the claim at 5% level of significance.

Thus, we conclude that the sample fails to provide us sufficient evidence against the claim so we may assume that the sample is taken from the population with median greater than 18.

E7) Here all the assumptions of Wilcoxon signed-rank test hold. So we go for this test.

Here, we want to test the manufacturer's claim that the average (median) breaking strength $(\tilde{\mu})$ of the rods is 30 pounds. So our claim is $\tilde{\mu} = 30$ and it complement is $\tilde{\mu} \neq 30.$ So we can take claim as the null hypothesis and complement as the alternative hypothesis. Thus,

$$H_0 : \tilde{\mu} = \tilde{\mu}_0 = 30 \quad [\text{median breaking strength is } 30]$$

$$H_1 : \tilde{\mu} \neq 30 \quad [\text{median breaking strength is not } 30]$$

Since the alternative hypothesis is two-tailed so the test is two-tailed test and the test statistic is the smaller of sum of positive ranks (T^+) and sum of negative ranks (T^-), that is,

$$T = \min\{T^+, T^-\}$$

Calculation for T:

S. No.	Breaking Strength (X)	Difference $d = X - \tilde{\mu}$ $= X - 30$	Absolute Value of Difference $\lvert d \rvert$	Rank of $\lvert d \rvert$	Signed Rank
1	19	−11	11	23.5	−23.5
2	30	Tie	---	---	---
3	28	−2	2	4	−4
4	35	5	5	9.5	9.5
5	23	−7	7	16	−16
6	37	7	7	16	16
7	36	6	6	12.5	12.5
8	32	2	2	4	4
9	40	10	10	21.5	21.5
10	24	−6	6	11.5	−11.5
11	31	1	1	1	1
12	33	3	3	7	7

13	21	−9	9	20	−20
14	36	6	6	12.5	12.5
15	30	Tie	---	---	---
16	32	2	2	4	4
17	26	−4	4	8	−8
18	40	10	10	21.5	21.5
19	38	8	8	18.5	18.5
20	28	−2	2	4	4
21	30	Tie	---	---	---
22	41	11	11	23.5	23.5
23	22	−8	8	18.5	−18.5
24	43	13	13	25.5	25.5
25	25	−5	5	9.5	−9.5
26	28	−2	2	4	−4
27	24	−6	6	125	−12.5
28	37	7	7	16	16
29	17	−13	13	25.5	−25.5

From the above calculation, we have

$T^+ = 9.5 + 16 + 12.5 + 4 + 21.5 + 1 + 7 + 12.5 + 4 + 21.5 + 18.5 + 23.5 + 25.5 + 16 = 193.0$

$T^- = 23.5 + 4 + 16 + 12.5 + 20 + 8 + 4 + 18.5 + 9.5 + 4 + 12.5 + 25.5$

$= 158.0$

n = number of non-zero d_i's = 29

Putting the values in test statistic, we have

$$T = \min\{T^+, T^-\} = \min\{193.0, 158.0\} = 158.0$$

Also n = 29 (> 20), therefore, it is the case of large sample. So in this case, we use Z-test. The test statistic of Z-test is given by

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

where, $E(T) = \frac{n(n+1)}{4} = \frac{29(29+1)}{4} = 217.5$ and

$$SE(T) = \sqrt{\frac{n(n+1)(2n+1)}{24}}$$

$$= \sqrt{\frac{29(29+1)(2\times 29+1)}{24}} = 46.25$$

Putting the values in test statistic Z, we have

$$Z = \frac{158.0 - 217.5}{46.25} = -1.29$$

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

Since calculated value of test statistic Z (= −1.29) is greater than the critical value (= − 2.58) and less than the critical value (= 2.58), that means it lies in non-rejection region so we do not reject the null hypothesis i.e. we support the claim at 1% level of significance.

Thus, we conclude that the sample fails to provide us sufficient evidence against the claim so we may assume that the manufacturer's claim that the average (median) breaking strength of the rods is 30 pounds is true.

E8) Here, we want to test the randomness of the given data so we go for run test.

We want to test that given data are in random order. So our claim is "the data are in random order" and its complement is "the data are not in random order". So we can take claim as the null hypothesis and complement as the alternative hypothesis. Thus,

H_0 : The data are in random order

H_1 : The data are not in random order

For applying run test, we need given observations in a sequence of two symbols. In order to convert given observations into a sequence of two symbols, first of all we replace the given observations by the symbol A if it is above the median value (= 0) and by symbol B if it is below the median value and discard the observation equal to median value. In this way, we get a sequence of symbols A's and B's which is given below:

$$\underbrace{AAA}_{1}\underbrace{BB}_{2}\underbrace{A}_{3}\underbrace{BB}_{4}\underbrace{AA}_{5}\underbrace{B}_{6}\underbrace{A}_{7}\underbrace{BBBBB}_{8}\underbrace{AAA}_{9}$$

Here, we see that there is a first run of three A's then there is a run of two B's and so on. In the last there is a run of three A's.

Here by counting, we have

n_1 = number of symbol A occurs = 10

n_2 = number of symbol B occurs = 10

R = total number of runs in above sequence = 9

For run test the test statistic is the total number of runs in the given sequence.

Since $n_1 = 10(< 20)$ and $n_2 = 10(< 20)$, so it is a small sample case.

The lower and upper critical values of runs corresponding to n_1 and n_2 at 5% level of significance are $R_L = 6, \quad R_U = 16$.

Since observed number of runs (R = 9) lies between critical values of runs, $R_L = 6$ and $R_U = 16$ so we do not reject the null hypothesis i.e. we support the claim at 5% level of significance.

Thus, we conclude that the given data are in random order.

E9) Here, we want to test the randomness of the cars entered in the parking so we go for run test.

Also we wish to test that the cars entered in the parking are in random order. So our claim is "the cars that enter in the parking are in random order" and its complement is "the cars that enter in the parking are not in random order”. So we can take claim as the null hypothesis and complement as the alternative hypothesis. Thus,

H_0: The cars that enter in the parking are in random order

H_1: The cars that enter in the parking are not in random order

For applying run test, we count the number of runs in the given sequence.

$$\underbrace{II}_{1}\underbrace{FFF}_{2}\underbrace{I}_{3}\underbrace{F}_{4}\underbrace{I}_{5}\underbrace{FFFF}_{6}\underbrace{II}_{7}\underbrace{F}_{8}\underbrace{III}_{9}\underbrace{FF}_{10}\underbrace{I}_{11}\underbrace{FF}_{12}\underbrace{I}_{13}\underbrace{FFFF}_{14}\underbrace{II}_{15}\underbrace{FFF}_{16}\underbrace{I}_{17}\underbrace{FF}_{18}\underbrace{IIII}_{19}$$

We see that there is a first run of two I’s then there is a run of three F’s and so on in the last there is a run of four I’s. By counting, we have

n_1 = number of symbol I occurs = 18

n_2 = number of symbol F occurs = 22

R = total number of runs in above sequence = 19

Since $n_2 = 22(>20)$ therefore, we use Z-test. The test statistic of Z-test is given by

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

where $E(R) = \frac{2n_1n_2}{n} + 1$

$$= \frac{2 \times 18 \times 22}{40} + 1 = 20.8 \text{ and}$$

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

$$= \sqrt{\frac{2 \times 18 \times 22(2 \times 18 \times 22 - 40)}{40^2(40-1)}} = \sqrt{9.54} = 3.08$$

Putting the values in test statistic Z, we have

$$Z = \frac{19 - 20.8}{3.08} = -0.58$$

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

Since calculated value of Z (= − 0.58) is greater than critical value (= − 2.58) and less than critical value (= 2.58), that means it lies in non-rejection region, so we do not reject the null hypothesis i.e. we support the claim at 1% level of significance.

Thus, we conclude that the sample fails to provide us sufficient evidence against the claim so we may assume that the cars enter in the parking lot are in random order.

E10) The main difference between chi-square test and Kolmogorov-Smirnov (K-S) test is that the chi-square test is designed for categorical data whereas K-S test is designed for the continuous data.

E11) Here, we want to test that the data have come from normal distribution with mean 5.6 and standard deviation 1.2. If $F_0(x)$ is the cumulative distribution function of this normal distribution then our claim is $F(x) = F_0(x)$ and its complement is $F(x) \neq F_0(x)$. Since the claim contains the equality sign so we can take claim as the null hypothesis and complement as the alternative hypothesis. Thus

$$H_0 : F(x) = F_0(x) \text{ for all values of x}$$

$$H_1 : F(x) \neq F_0(x) \text{ for at least one value of x}$$

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

$$D_n = \sup_x \left| S(x) - F_0(x) \right|$$

where, S(x) is the empirical cumulative distribution function based on the observed sample.

For obtaining the empirical cumulative distribution function S(x), first we obtain cumulative frequencies and then by the definition of empirical cumulative distribution function we obtain S(x) as

$$S(x) = \frac{\text{The number of sample observations} \leq x}{\text{Total number of observations}}$$

Here, theoretical cumulative distribution function $F_0(x)$ can be obtained with the help of the method as described in Unit 4 of Block 4 of the course MST-003. Therefore, by the definition of distribution function, we have

$$F_0(x) = P[X \leq x] = P\left[\frac{X-\mu}{\sigma} \leq \frac{x-\mu}{\sigma}\right] = P[Z \leq z]$$

where, $P[Z \leq z]$ can be to obtained with he help of **Table I** given in Appendix at the end of the Block 1 of this course.

Thus, when x = 4.8, $z = \frac{x-\mu}{\sigma} = \frac{4.8-5.6}{1.2} = -0.67$ and

$$F_0(4.8) = P[Z \leq z] = P[Z \leq -0.67]$$

$$= 0.5 - P[0 < Z < 0.67]$$

$$= 0.5 - 0.2486$$

$$= 0.2514$$

Similarly, when x = 5.8, $z = \frac{5.8-5.6}{1.2} = 0.17$ and

$$F_0(5.8) = P[Z \leq z] = P[Z \leq 0.17]$$

$$= 0.5 + P[0 < Z \leq 0.17]$$

$$= 0.5 + 0.0675 = 0.5675 \text{ and so on.}$$

Calculation for $|S(x) - F_0(x)|$:

X	Observed Frequency	Cumulative Observed Frequency	Observed Distribution Function S(x)	$Z = \frac{X - \mu}{\sigma} = \frac{X - 5.6}{1.2}$	Theoretical Distribution Function $F_0(x)$	Absolute Difference $\lvert S(x) - F_0(x)\rvert$
4.8	1	1	1/12 = 0.0833	−0.67	0.2514	0.1681
5.0	1	2	2/12 = 0.1667	−0.50	0.3085	0.1418
5.5	1	3	3/12 = 0.2500	−0.08	0.4681	0.2181
5.8	1	4	4/12 = 0.3333	0.17	0.5675	0.2342
5.9	2	6	6/12 = 0.5000	0.25	0.5987	0.0987
6.0	1	7	7/12 = 0.5833	0.33	0.6293	0.0460
6.2	2	9	9/12 = 0.7500	0.50	0.6915	0.0585
6.3	1	10	10/12 = 0.8333	0.58	0.7190	0.1143
6.6	2	12	12/12 = 1.0000	0.83	0.7967	0.2033

From the above calculations, we have

$$D_n = \sup_x |S(x) - F_0(x)| = 0.2342$$

The critical value of test statistic D_n corresponding n = 12 at 5% level of significance is 0.375.

Since calculated value of test statistic D_n (= 0.2342) is less than critical value (= 0.375) so we do not reject the null hypothesis i.e. we support the claim at 5% level of significance.

Thus, we conclude that the sample fails to provide us sufficient evidence against the claim so we may assume that data have come from the normal distribution with mean 5.6 and standard deviation 1.2.