BLOCK 3
NON-PARAMETRIC STATISTICS

ignou
THE PEOPLE'S
UNIVERSITY

UNIT 5 CHI-SQUARE*

Structure

5.0 OBJECTIVES

After reading this unit, you will be able to,

- explain what is chi-square;
- describe for the application of χ^2 test;
- explain chi-square distribution;
- discuss the properties of chi-square distribution;
- explain the application of chi-square; and
- discuss the precautions about using the chi-square test.

5.1 INTRODUCTION

In psychology some times the researcher is interested in such questions as to whether one type of soft drink preferred by the consumer is in any way influenced by the age of the consumer, and whether? Similarly among the four colours, red, green, indigo, yellow, is there any significant difference in the participants in regard to the preference of colours, etc.

* Dr. Vivek Belhekar, Faculty, Department of Psychology, Bombay University, Mumbai

To approach questions like these, we find the number of persons who preferred particular colour. Here we have the data in the form of frequencies, the number of cases falling into each of the four categories of colours. Here the researcher compares the observed frequencies characterising the choices of the 4 categories by a large number of individuals. Some prefer red, some white, some green and some yellow etc. Hypothetically speaking one may say that all colours are equally preferred and are equal choice for the individuals. But in reality there may be more number of persons choosing red or blue than yellow and similarly quite a few may choose white etc. Thus what is expected as choice which is equal for all colours may be seen differently when actual choice is made by the persons. Now we note the number of persons who had chosen red, blue, yellow and white respectively and note it in the table. Hypothetically we said that there will be the same number of persons choosing each of the 4 colours. This also we note in the table. Now we compare and see for each colour what is the difference between the actual number of choices made as compared to the hypothetical equal number. The differences thereof are noted. To find out if these differences are really of statistical significance amongst the persons in regard to choice of colours, we use a statistical technique called Chi-square. In the present unit, we will explain what is chi-square and also describe the assumptions for its application.

5.2 WHAT IS CHI- SQUARE?

This is one of the most important non-parametric statistics. As we see further that chi-square is used for several purposes.

The term ‘chi’ (is the Greek letter, it is pronounced ki). The chi-square test was originally developed by Karl Pearson in 1900 and is sometimes called the Pearson Chi-square. According to Garrett (1981), the difference between observed and expected frequencies are squared and divided by the expected number in each case and the sum of these quotients is chi-square‖. According to Guilford(1973) “By definition a $\chi 2$ is the sum of ratio (any number can be summed), each ratio is that between a squared discrepancy of difference and an expected frequency”.

On the basis of the above definitions it can be said that the discrepancy between observed and expected frequencies is expressed in term of a statistics named chi-square ($\chi 2$).

5.2.1 Assumptions for the Application of $\chi 2$Test

Before using chi-square as a test statistic to test a hypothesis, the following assumptions are necessary:

- Data should be in the form of frequencies. Although the chi-square test is conducted in terms of frequencies or data that can be readily transformed into frequency, it is best viewed conceptually as a test about proportions.

- The chi-square test is applied only to discrete data. However, any continuous data can be reduced to the categories in such a way that they can be treated as discrete data and then the application of chi-square is justified.
- The $\chi 2$ is applied when there is quantitative data.
- The observation should be independent. For example let us say that we wanted to find out the colour preference of 100 females. We select 100 subjects randomly, the preference of one individual is not predictable from that of any other. If there is a relationship between two variables or the participants are matched, then chi-square cannot be used to test. Correlated data like matched pair are not subjected to chi-square treatment.
- The sample drawn from the population about which inference is to be made should be random.
- If there are only two cells, the expected frequencies in each cell should be 5 or more. Because for observation less than 5, the value of $\chi 2$ shall be over estimated, resulting in the rejection of the null hypothesis.
- For more than two cells, no more than 20% of the expected values may be lesser than 5 and no expected value may be lesser than 1. If more than 20 percent of the cells have expectedfrequencies less than 5 then $\chi 2$ should not be applied.
- A sample with a sufficiently large size is assumed. If sample size is small then $\chi 2$ will yield an inaccurate inference. In general, larger the sample size, the less affected by chance is the observed distribution, and thus the more reliable the test of the hypothesis.
- The data should be expressed in original units, rather than in percentage or ratio form. Such precaution helps in comparison of attributes of interest.

5.2.2 $\chi 2$ Distribution

The term non-parametric does not mean that the population distribution under study has no parameters. All populations have certain parameters which define their distribution. The sampling distribution of $\chi 2$ is called $\chi 2$ distribution. Other hypothesis testing procedures, the calculated value of $\chi 2$ test statistic is compared with its critical (or table) value to know whether the null hypothesis is true. The decision of accepting a null hypothesis is based on how 'close' the sample results are to the expected results. If the null hypothesis is true, the observed frequencies will vary from their corresponding expected frequencies according to the influence of random sampling variation. The calculated value of $\chi 2$ will be smaller when agreement between observed frequencies and expected frequencies are smaller. The shape of a particular chi-square distribution depends on the number of degrees of freedom. Let us now see what is degrees of freedom? Let us say there are 5 scores, as for example, 25,35,45,55,65

The mean here is 45. Let us say that from this mean of 45 we try to see the difference in each of the other scores. While all the other 4 scores are taken for the difference, this 45 is not available for other calculations, as from this Mena=45, all other calculations are made. From this mean we note the differences in the remaining scores, and thus amongst the 5 scores one score is not movable or in other words the group of 5 subjects has lost one degree of freedom. Suppose we go for higher level calculations, then accordingly there will be more number of items that cannot be moved and thus we may have greater degrees of freedom. In a 3 × 3 table, we will be using one cell from the row and one cell from the column for the purpose of chi-square calculations. Thus the degrees of freedom will be (3 – 1) (3 – 1) = 4. That is (r- 1) (k-1) (r = row and k = column).

It may also be noted that chi-square assumes positive values only. As such a chi-square curve starting from the original (zero point) lies entirely to the right of the vertical axis.

If we draw a curve with degrees of freedom on the x axis and chi-square values on the y axis, the he shape of a chi-square distribution curve will be skewed for very small degree of freedom. As the degrees of freedom increase, the shape also changes. Eventually, for larger degrees of freedom, the curve looks similar to the curve of a normal distribution. A point worth nothing is that the total area under a chi-square distribution curve is 1.0 as is the case in all other continuous distribution curves.

Check Your Progress I

1) State any one assumption for the application of $\chi 2$ Test.

……

……

……

……

……

5.3 PROPERTIES OF CHI-SQUARE DISTRIBUTION

The following properties of chi-square test statistic are to be kept in mind when we analyse its sampling distribution.

- Chi-square is non-negative in value; it positively valued, because all discrepancies are squared, both positive and negative discrepancies make a positive contribution to the value of $\chi 2$.
- Chi-square will be zero only in the unusual event that each observed frequency exactly equals the corresponding expected frequency.

- It is not symmetrical; it is skewed to the right as mentioned earlier. The reason for this is because all the values are positive and hence so instead of having a symmetrical distribution, we have a distribution all on the positive side.
- Since density function of $\chi 2$ does not contain any parameter of population, Chi-square test statistic is referred to as a non-parametric test. Thus chi- square distribution does not depend upon the form of the parent population.
- There are many chi-square distributions. As with the t-distribution, there is a different chi-square distribution for the value of each degree of freedom. If we know the degrees of freedom and the area in the right tail of a chi-square distribution, we can find the value of chi-square from the table of chi-square. We give below two examples to show how the value of chi-square can be obtained from this table.

Example 1: Find the value of chi-square for 10 degree of freedom and an area of 0.05 in the right tail of the chi-square distribution curve.

Solution: In order to find the required value of chi-square, we first locate 10 in the column for degree of freedom (df) and 0.05 in the top row in table of chi-square. The required chi-square value is given by the entry at the intersection of the row for 10 and the column for 0.05. This value is 18.307.

Example 2: Find the value of chi-square for 20 degree of freedom and an area of 0.10 in the left tail of the chi-square distribution curve.

Solution: This example is different from the above one. Here, the area is in the left tail of the chi-square distribution curve is given. In such a case, we have to first find the area in the right tail. This is obtained as follows:

Area in the right tail= 1 - area in the left tail

Area in the right tail = 1 - .10 = .90

Now we can find, in the same manner as used earlier, the value of$\chi 2$. We locate 20 in the column for df and .90 on top row in the table of chi-square. The required value of $\chi 2$ is 12.443.

Check Your Progress II

1) State any one property of chi-square test statistic are to be kept in mind when we analyse its sampling distribution.

……

……

……

……

……

5. 4 APPLICATIONS OF CHI-SQUARE TEST

Chi-square is used for categorical data that is for data comprising unordered quantitative categories such as colours, political affiliation etc. The important applications of $\chi 2$test are as follows:

- Test of goodness of fit
- Test of independence
- Test for homogeneity

5.4.1 Test of Goodness of Fit

On several occasion a decision - maker needs to understand whether an actual sample distribution matches with a known theoretical distribution such as binomial, poisson, normal and so on. Similarly some time researcher compares the observed (sample) frequencies characterising the several categories of the distribution with those frequencies expected according to his or her hypothesis.

The $\chi 2$ test for goodness of fit is a statistical test of how well given data support an assumption about the distribution of a population. That is, whether it supports a random variable of interest. In other words, the test determines how well an assumed distribution fits the given data.

To apply this test, a particular theoretical distribution is first hypothesised for given population and then the test is carried out to determine whether or not the sample data could have come from the population of interest with the hypothesised theoretical distribution. The observed frequencies come from the observation of sample and the expected frequencies come from the hypothesised theoretical distribution. The goodness of fit test focusses on the difference between the observed frequencies and expected frequencies.

This can be explained with the help of following examples.

In a particular study, a researcher wants to study that among four brands of glucose biscuits, is there a difference in the proportion of consumers who prefer the taste of each. We formulate the null hypothesis that there is no differential preference among consumers for the four brand of glucose biscuits. Here we test the hypothesis that there are equal probability of individuals preferring each brand i.e. 1/4 = .25. Or if we take 100 persons to test this hypothesis, 25 each will prefer theoretically the 4 brands of biscuits.

The alternative hypothesis is that a preference exists for the first brand of the biscuits and the reminder are equally less preferred, or that the first two are preferred more than the second two and so on.

The use of chi-square tells us whether the relative frequencies observed in the several categories of our sample frequency distribution are in accordance with the set of frequencies hypothesised.

To test the null hypothesis in the above study, we might allow subjects to taste each of the four brand of glucose biscuits and then find their preference. We control possible extraneous influence, such as knowledge of brands name and order of presentation. Suppose we had as mentioned earlier randomly selected 100 subjects and that our observed frequencies of preference are as given in the table below:

	Brand A	**Brand B**	**Brand C**	**Brand D**
Observed Frequencies	20	18	30	32

We calculate the expected frequency for each category by multiplying the proportion hypothesised to characterise that category in the population by the sample size. According to null hypothesis, the expected proportionate preference for each biscuit is 1/4 and the frequencies of preference for each is therefore 25 (1/4) (100/4) = 25.

In any experiment, we anticipate that the observed frequency of choice will vary from the expected frequencies. We do not use these differences directly. One reason is that some differences are positive and some are negative. Thus they would cancel each other out. To get around this, we square each difference. But how much variation to expect? Some measures of discrepancy is required to test the null hypothesis, this can be tested by chi-square. We calculate the chi-square (the method of computation of chi-square will be explained in the next unit). If the obtained chi-square value is more than the critical values at 0.05 or 0.01 level than we reject the null hypothesis and if obtained value of chi-square is less than the given values then we retain the null hypothesis.

The obtained value of chi-square depends on the size of the discrepancy between observed frequency and expected frequencies. Larger the differences between observed frequencies and expected frequencies the larger will be the chi-square.

The size of the discrepancy relative to the magnitude of the expected frequency contribute in the determination of the value of chi-square. For example, if we toss a number of coins and find out how many times theheads‘ ortails ‘have appeared. To take a numerical example, when we toss coins in 12 times and 1000 times. Then according to null hypothesis our expected frequency will be 6, and 500 respectively. If we obtain 11 heads in 12 tosses the discrepancy is 5. However, if we obtain 505 heads in 1000 tosses, the discrepancy is also 5.

The value of chi-square depends on the number of discrepancies involved in its calculation. For example if we use three brands of biscuits not the four, there would be less discrepancy to contribute to the total of chi-square. This will influence the degrees of freedom, that is, when the number of

discrepancies are 4 the degrees of freedom will be 3 and when number of discrepancies are 3 the degree of freedom will be 2. When degrees of freedom is 4 and obtained chi-square is more than 9.48 then we reject the null hypothesis at 0.05 level. On the other hand when the degrees of freedom is 4 and obtained chi-square is 7.81 then we retain the null hypothesis at 0.05 level.

5.4.2 Test of Independence

In the above discussion so far, we have considered the application of chi-square only to 'one' variable case. We have considered testing whether the categories in an observed frequency distribution differ significantly from one another. The chi-square test has a much broader use in social science research; to test whether one observed frequency distribution significantly differ from another observed frequency. In other words, it can be said that chi-square is used for the analysis of bivariate frequency.

For example, social researcher is interested in surveying the attitudes of high school students concerning the importance of getting a college degree. She questions a sample of 60 senior high school students about whether they believe that college education is becoming more important, less important or staying the same and whether male students respond differently from female students ?

As we know that nominal and ordinal variables are generally presented in the form of a cross-tabulation. Specifically cross tabulation are used to compare the distribution of one variable often called dependent variable, across categories of some other variable the independent variable. In a cross tabulation the focus is on the difference between group — such as between males and females — in terms of the dependent variable — for example, opinion about the changing value of a college education. In the above example we want to study whether there exists gender differences in belief regarding the importance of a college degree - are statistically significant.

To study this question, we classify the data (observed frequencies) in a bivariate distribution. For each variable, the categories are mutually exclusive. The data is classified in the following table:

	More Important	**Less Important**	**About the Same**
Male	25	6	8
Female	10	4	7

Bivariate frequency distribution of two types are known as contingency tables. From such a table we may inquire what cell frequencies would be expected if the two are independent of each other in the population. The chi-square test may be used to compare the observed cell frequencies with those

expected under the null hypothesis of independence. If the (fo-fe) discrepancy are small, chi-square will be small suggesting that the two variable could be not different. Conversely, a large chi-square points toward a contingency relationship.

As in the case of the t ratio there is a sampling distribution for chi-square that can be used to estimate the probability of obtaining a significant chi-square value by chance alone rather than by actual population difference. The test is used to study the significance of difference between mean. On the other hand, chi-square is used to make comparison between frequencies rather than mean scores. As a result the null hypothesis for the chi-square test states that the populations do not differ with respect to frequencies of occurrence of a given characteristic. In general, the null hypothesis of independence for a contingency table is equivalent to hypothesising that in the population the relative frequencies for any row (across the categories of the column variable) are the same for all rows, or that in the population the relative frequencies for any column (across the categories of the two variables) are the same for all columns.

If the null hypothesis of independence is true at the population level, we should expect that random sampling will produce obtained value of chi-square that are in accord with the tabulated distribution of that statistic. If the hypothesis is false in any way, the obtained value of chi-square will tend to be larger than when the alternate hypothesis, stating that there will be a significant difference between the variables, is true.

In the case of a 2 × 2 contingency table, we can test for a difference between two frequencies or proportions from independent samples (just as with the 1 × 2 table, where we can test a hypothesis about a single proportion.

5.4.3 Test of Homogeneity

The test of homogeneity is useful in a case when we are interested in verifying, whether several populations are homogeneous with respect to some characteristic of interest. For example a movie producer is bringing out a new movie. In order to develop an advertising strategy he/she wants to determine whether the movie will appeal most to a particular age group or whether it will appeal equally to all age groups. We formulate the null hypothesis that the opinion of all age groups is same about the new movie. Hence, the test of homogeneity is useful in testing a null hypothesis that several populations are homogeneous with respect to a character. This test is different from the test of independence on account of the following reasons:

1) Instead of knowing whether two attributes are independent or not, we may like to know whether different samples come from the same population.

2) Instead of taking only one sample for this test two or more independent samples are drawn from each population.

3) To apply this test first a random sample is drawn from each population, and then in each sample the proportion falling in each category is determined. The sample data so obtained is arranged in contingency table. The procedure for testing of hypothesis is same as for test of independence.

Check Your Progress III

1) List the the important application of $\chi 2$ test.

……

……

……

……

……

5.5 PRECAUTIONS ABOUT USING THE CHI-SQUARE TEST

In order to use a chi-square test properly, one has to be extremely careful and keep in mind certain precautions:

1) A sample size should be large enough. If the expected frequencies are too small, the value of chi-square gets over estimated. To overcome this problem we must ensure that the observed frequency in any cell of the contingency table should not be less than 5.

2) When the calculated value of chi-square turns out to be more than the critical or theoretical value at a predetermined level of significance, we reject the null hypothesis. In contrast, when the chi-square value is less than the critical theoretical value, the null hypothesis is not rejected. However, when the chi-square value turns out to be zero, we have to be extremely careful to confirm that there is no difference, between the observed and expected frequencies. Such a situation sometimes arises on account of faulty method used in the collection of data.

Check Your Progress IV

1) Fill in the blanks

a) When the expected frequencies are too small, the value of chi-square gets _____.

b) When the chi-square value is less than the critical theoretical value, the null hypothesis is _____.

5.6 LET US SUM UP

To summarise, the chi-square test is the most commonly used non-parametric test. Non parametric techniques are used when a researcher works with ordinal or nominal data or with a small number of cases representing a highly asymmetrical underlying distribution. The chi-square test is used as a test of significance when we have data that are given in frequencies or categories. The chi-square test is used for two broad purposes. Firstly it is used as a test of goodness of fit and secondly as a test of independence. As a test of goodness of fit chi-square tells us whether the frequencies observed among the categories of a variable are tested to determine whether they differ from a set of expected hypothetical frequencies. As a test of independence chi-square is usually applied for testing the relationship between two variables in two way. First by testing the null hypothesis of independence and second by computing the value of contingency coefficient, a measurement of relationship existing between the two variable. When we have less than 5 observed frequencies in any cell then we use the Yate's correction.

5.7 REFERENCES

Kerlinger, Fred N. (1963). Foundation of Behavioural Research, (2nd Indian reprint) Surjeet Publication, New Delhi.

Garrett, H.E. (1971), Statistics in Psychology & Education, Bombay, Vakils, Seffer & Simoss Ltd.

Guilford, J.P. (1973), Fundamental Statistics in Psychology & Education, Newyork, McGraw Hill.

5.8 KEY WORDS

Chi-square distribution: A distribution with degree of freedom as the only parameter. It is skewed to the right for small degree of freedom.

Chi-square test: A statistical techniques used to test significance in the analysis of frequency distribution.

Contingency table: A table having rows and column wherein each row corresponds to a level of one variable and each column to a level of another variable.

Expected frequencies: The frequencies for different categories which are expected to occur on the assumption that the given hypothesis is true.

Observed frequencies: The frequencies actually obtained from the performance of an experiment.

5. 9 ANSWERS TO CHECK YOUR PROGRESS

Check Your Progress I

1) State any one assumption for the application of $\chi 2$ test.

   In $\chi 2$ test, data should be in the form of frequencies. Although the chi-square test is conducted in terms of frequencies or data that can be readily transformed into frequency, it is best viewed conceptually as a test about proportions.

Check Your Progress II

1) State any one property of chi-square test statistic are to be kept in mind when we analyse its sampling distribution.

   Chi-square is non-negative in value; it positively valued, because all discrepancies are squared, both positive and negative discrepancies make a positive contribution to the value of $\chi 2$.

Check Your Progress III

1) List the the important application of $\chi 2$ test.

   The important application of $\chi 2$ test are test of goodness of fit, test of independence and test for homogeneity.

Check Your Progress IV

1) Fill in the blanks

a) When the expected frequencies are too small, the value of chi-square gets over estimated.

b) When the chi-square value is less than the critical theoretical value, the null hypothesis is not rejected.

5.10 UNIT END QUESTIONS

1) Describe the properties of chi-square test.

2) What do you understand by the goodness of fit test ? Describe with examples.

3) What is the chi-square test for independence? Explain with examples.

4) What precautions would you keep in mind while using chi-square.

UNIT 6 COMPUTATION OF CHI-SQUARE*

Structure

6.0 OBJECTIVES

After reading this unit, you will be able to,

- identify chi-square as a test of goodness of fit;
- apply testing of equal probability hypothesis;
- use testing of normal probability hypothesis;
- calculate chi-square as a test of independence in contingency tables; and
- apply chi-square test when table entries are small.

6.1 INTRODUCTION

As we have discussed in last unit that chi-square is a commonly used non-parametric statistical test. This is used when we have data in the form of frequencies, ordinal or nominal levels of measurement. Chi-square test is used for different purposes. In the last unit you were given the conceptual knowledge of chi-square test and in this unit we will discuss the method of computation of chi-square.

6.2 COMPUTATION OF CHI- SQAURE

The chi-square test is used for comparing experimentally obtained results with those to beexpected. The computations of chi-square are discussed as follows:

* Dr. Vivek Belhekar, Faculty, Department of Psychology, Bombay University, Mumbai

6.2.1 Chi-square as a Test of Goodness of Fit

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. In Chi-Square goodness of fit test, the term goodness of fit is used in order to compare the observed sample distribution with the expected probability distribution. Chi-Square goodness of fit test determines how well theoretical distribution (such as normal, binomial, or Poisson) fits the empirical distribution. In Chi-Square goodness of fit test, sample data is divided into intervals. Then the numbers of points that fall into the interval are compared, with the expected numbers of points in each interval.

In regard to the procedure, set up the hypothesis for chi-square goodness of fit test, that is set up both null and alternative hypothesis.

a) **Null hypothesis:** In chi-square, goodness of fit test, the null hypothesis assumes that there is no significant difference between the observed and the expected value.

b) **Alternative hypothesis:** In chi-square, goodness of fittest, the alternative hypothesis assumes that there is a significant difference between the observed and the expected value.

   Calculate chi-square using the formula and find out for the given degrees of freedom if chi-square value is significant at 0.05 or 0.01 levels. If so reject the null hypothesis. If not significant accept the null hypothesis.

6.2.2 Testing Hypothesis of Equal Probability

The Chi-square test is a useful method of comparing experimentally obtained results with those to be expected theoretically on some hypothesis. The formula for calculating $\chi 2$ is

$$\chi 2 = \sum[(fo - fe)^2]/fe$$

Where;

*f*o= observed frequency of a phenomenon oreven which the experimenter is studying.

*f*e = expected frequency of the same phenomenon based on “no difference” or “null” hypotheses.

The use of the above formula can be illustrated by the following example.

Example: An attitude scale designed to measure attitude towards co-education was administered on 240 students. The data is given in table 6.1. The observed data is (*f*o) given in the first row of table 6.1. In the second row is the distribution of answer to be expected on the basis of null hypothesis (*f*e), if each answer is selected equally.

Table 6.1: Responses from participants regarding Attitude

Calculations	Favourable	Neutral	Unfavourable	Total
fo	70	50	120	240
fe	80	80	80	240
fo- fe	-10	-30	40	
$(fo- fe)^2$	100	900	1600	
$(fo- fe)^2/fe$	100/80	900/80	1600/80	
$\chi 2$	1.25	11.25	20	$\sum(fo- fe)^2/fe$ = 32.50

The formula of $\chi 2$ is

$\chi 2 = \sum[(fo- fe)^2]/fe$

$= 1.25 + 11.25 + 20$

$= 32.5$

df=(r-1)(c-1)
=(3-1)(2-1)

df = 2

Critical value of $\chi 2$ for df= 2 at 0.05 level = 5.59 (refer to statistical table given at the appendix of Statistics book). Critical value of $\chi 2$ at 0.01 level = 9.19 (refer to the table for chi-square given in the appendix of textbook on statistics) .

The computed value of $\chi 2$, i.e. 32.5 is above the critical values at 0.05 and 0.01significance levels, we conclude that $\chi 2$ is significant and we retain the null hypothesis.

On the other hand if the computed value of $\chi 2$ is less than 5.59 or 9.19, then the null hypothesis is accepted and it may be concluded that there is no significant difference in the attitude of the participants.

However in our example we have found the chi-square value being greater than what is given in the table and so we reject the null hypothesis.

Steps for chi-square testing:

1) First set a null of hypothesis.

2) Collect the data and findout observed frequency.

3) Find out the expected frequencies by adding all the observed frequencies divided by number of categories (In Ist example 240/3=80, in II example 100/5=20).

4) Find out the difference between observed frequencies and expected frequencies (*fo- fe*).

5) Find out the square of the difference between observed and expected frequency $(fo- fe)^2$.

6) Divide it by expected frequency, $(fo- fe)^2/fe$ You will get aquotient.

7) Find out the sum of these quotients.

8) Determine the degree of freedom and find out the critical value of $\chi 2$ from table.

9) Compare the calculated and table value of $\chi 2$ and use the following decision rule.

Accept null hypothesis if $\chi 2$ is less than critical value given in table.

Reject the null hypothesis if calculated value of $\chi 2$ is more than what is given in the table under 0.05 or 0.01 significance levels.

Testing the Divergence of Observed Results from those Expected on the Hypothesis of a Normal Distribution

Some times the expected frequencies are determined on the basis of the normal distribution of observed frequencies in the entire population. The method as to how this is done is explained in the following examples.

Example 1: 180 school teachers were classified into six categories on the basis of their performance. Do these teachers differ significantly in their performance as compared to the one expected if performance is supposed to be distributed normally in our population of school teachers.

Categories 1	2	3	4	5	6	Total
No. of school teachers 15	25	45	50	35	10	180

Here, expected frequencies are to be determined on the basis of the normal distribution hypothesis. For this purpose we have to divide the base line of the curve into six equal segments of 1σ each. To mention here, the baseline in any normal distribution ranges from from -3σ to +3σ. The table below presents how many cases fall between the different areas.

Table 6.2 : Frequencies on the basis of Normal Distribution

Sr. No.	Category	% of cases falling between the area	No. of cases out of 180
1	+ 3.00 σ to 2.00 σ	2.28	3.6
2	+ 2.00 σ to +1.00 σ	13.59	25.2
3	+ 1.00 σ to M	34.13	61.2
4	M to - 1.00 σ	34.13	61.2
5	-1.00 σ to - 2.00 σ	13.59	25.2
6	- 2.00 σ to -1.00 σ	2.58	3.6

We will now test the null hypothesis by computing the values of $\chi 2$ givenin table.

Table 6.3 : Calculation of Chi-square step by step

	1	**2**	**3**	**4**	**5**	**6**	**Total**
fo	15	25	45	50	35	10	180
fe	3.6	25.2	61.2	61.2	25.2	3.6	180
fo- fe	11.4	-0.2	-16.2	-11.2	9.8	6.4	
$(fo- fe)^2$	129.96	0.04	262.44	125.44	96.04	40.96	
$(fo- fe)^2/fe$	36.1	0.001	4.29	2.05	3.81	11.38	

The formula of $\chi 2$ is

$$\chi 2=\sum[(fo- fe)^2]/fe$$

$$= 36.1+.001+4.29+2.05+3.81+11.38$$

$$=57.63$$

df =(r-1)(c-1)
=(6-1)(2-1)

df =5

The critical value of $\chi 2$ for df= 5 is 11.0700 at 0.05 level and 15.086 at 0.01 level (refer to the table for chi-square given in the appendix of book on statistics). Obtained values of $\chi 2$ is much greater than the critical values. Hence it can be said that $\chi 2$ is significant and we reject the null hypothesis.

Example 2: 200 salesmen were classified into five categories on the basis of their performance ranging from excellent performance to very poor performance. Does this evaluation differ significantly from the one expected from the normal distribution?

The observed frequencies are given in the following table.

Very good	Good	Average	Poor	Very poor
30	45	65	40	20

We have to find out the expected frequencies with the help of normal probability Curve. For this we have to divide the base line of the curve into five equal segments by 6/5=1.2

Table 6.4: Computation of Expected Frequencies on the basis of Normal Distribution

Sr. No.	Category	% of cases falling between the area	No. of cases out of 180
1	+ 3.00 σ to 1.8 σ	3.45	6.90 or 7
2	+ 1.8 σ to +.60 σ	23.84	47.68 or 48
3	+ .60 σ to -.60 σ	45.14	90.28 or 90
4	-.60 σ to - 1.8 σ	23.84	47.68 or 48
5	-1.8 σ to - 3.00 σ	3.45	6.90 or 7

We will now test the null hypothesis by computing the values of $\chi 2$ given in table.

Table 6.5: Calculation of Chi-square step by step

	Very good	Good	Average	Poor	Very poor	Total
fo	30	45	65	40	20	200
fe	7	48	90	48	7	
fo- fe	23	-3	-25	-8	13	
$(fo- fe)^2$	529	9	625	64	169	
$(fo- fe)^2/fe$	75.57	0.19	6.94	1.33	24.14	

$$\chi 2=\sum[(fo- fe)^2]/fe$$

$$\chi 2 = 75.57+.19+6.94+1.33+24.14$$

=108.17

df =(r-1)(c-1)

= (2-1) (5-1)
df =4

The critical value of $\chi 2$ at 4 df is 9.49 at 0.05 level and 13.29 at 0.01 level (refer to the table for chi-square given in the appendix of textbook on statistics) therefore obtained values of $\chi 2$ is much greater than the critical values. Hence it can be said that $\chi 2$ is significant and we reject the null hypothesis.

Steps for chi-square testing:

1) Formulate the null hypothesis.
2) Find out the observed frequencies.
3) To find out the expected frequencies first divide the base line of the normal curve (which is divided into 6 equal σ units) equal to the number of given categories of observed frequencies. For example in example 1 there are six categories therefore we have to divided the base line of the curve in 6 units. In example 2 there are 5 categories and we have to divide the base line of the curve in 5 units. Thus in the first case these can be obtained by 6/6= 1σ unit in example. In example 2 there are five categories and 5/6= 1.2 σ unit distance.
4) Now we have to decide the areas of normal curve as shown in table.
5) To decide the number of cases or percentage of cases, refer table of normal probability. These tables are given in the appendix of books on Statistics. In example 2, in the first category the area of normal curve is + 3 σ to + 1.8 σ. In table of normal distribution 0.4986 i.e. 49.86% cases lies between means and 3 σ and 0.4641, that is. 46.41% cases lies in between mean and 1.80 σ. Therefore 49.86-46.41=3.45% cases are there in the first category. In the second category the area of normal curve is + 1.80 σ to + 0.6 σ. There are .4641 i.e. 46.41% cases lies between mean and 1.8 σ and .2257 i.e. 22.57% cases lies in between mean and .6 σ . Therefore percentage of cases between + 1.80 σ to0.6 σ will be (46.41-22.57) 23.84.

   In the third category the areas of normal curve is + 0.6 σ to - 0.6 σ here .2257 means 22.57% cases lie on +0.6 σ and 22.57 lies on 0.6σ therefore (22.57+22.57) 45.14% cases lies in the third category. Because this is based on normal distributions which is symmetrically divided in two halves, therefore number of cases in fourth and fifth categories will be the same as in the second and first category.
6) No. of cases out of total number of cases can be found by dividing the obtained areas of normal curve by 100 and multiply by total number of cases. For example, in example 2, 3.45% of the cases falling in area I of

first category is 3.45. Therefore the number of cases out of 200 will be (3.45 × 200)/100=6.90. In this way find out fe. 7) After calculating the fe place the values in the table and find out the $\chi 2$with the help of given formula.

6.2.3 Chi-square as a Test of Independence

In addition to testing the agreement between observed frequencies and those expected from some hypothesis, that is, equal probability and normal probability, chi-square may also be applied to test the relationship between variables. In this we test whether two variables are dependent or independent to each other. The following table is known as contingency table. Contingency table is a double entry or two way table in which the possession by a group of varying degree of the characteristics is represented. The same person for example may be classified as on the basis of sex that is boys and girls or can be categorised in terms of the colour of their eyes, that is, brown or black etc. These attributes are noted in terms of the gender factor.

To cite another example, father and sons may be classified with respect to temperament or achievement and the relationship of the attributes in the groups may be studied.

In the following table there is a double entry or it is called a two way table in which 100 boys and 100 girls who express their view on the statement "There should be same syllabus in all the universities of the country," in terms of agree, neutral, disagree. The results are shown in the following table.

Table 6.6: Frequencies of respondents in terms of views expressed

Gender	Agree	Neutral	Disagree	Total
Boys	20	30	50	100
Girls	50	20	30	100
Total	70	50	80	200

Do you think that the opinion is influenced by the gender of the students.?

For this we present the data in contingency table form as above. The expected frequencies after being computed are written within brackets besides the respective observed frequencies. The contingency table and the process of computation of expected frequencies is given in the table 6.7.

Table 6.7: Contingency table with fo and fe for the data given in table 6.6

Gender	Agree	Neutral	Disagree	Total
Boys	20 (35)	30 (25)	50 (40)	100
Girls	50 (35)	20 (25)	30 (40)	100
Total	70	50	80	200

The method to find out the expected frequency for each cell is given below:

70 × 100/200 = 35 50 × 100/200 = 25 80 × 10/200=40
70 × 100/200 = 35 50 × 100/200 = 25 80 × 100/200=40
The value of $\chi 2$may be computed with the help of usual formula,

$$\chi 2=\sum[(fo-fe)^2]/fe$$

Table 6.8: Computation of $\chi 2$ from contingency table 6.7

fo	*fe*	*fo-fe*	$(fo-fe)^2$	$(fo-fe)^2/fe$
20	35	-15	225	6.43
30	25	5	25	1
50	40	10	100	2.5
20	35	-15	225	6.43
30	40	-5	25	1
20	40	-10	100	2.5
				$\chi 2=19.86$

The value of $\chi 2$ may be computed with the help of normal formula.

For df = 2, the critical value at 0.05 level is 5.99 and at 0.01 level 9.21. Our obtained value of $\chi 2$ is 19.86 . It is far higher than the table value. Therefore we reject the null hypothesis and conclude that gender influences the opinion.

Steps for chi-square testing:

1) Formulate the null hypothesis
2) Find out the expected values by the method shown in table.
3) Find out the difference between observed and expected values for each cell.
4) Square each difference and divide this in each cell by the expected frequency.
5) Add up these and the sum of these values gives $\chi 2$.

6.2.4 2× 2 Fold Contingency Table

When the contingency table is 2 x 2 fold $\chi 2$may be calculated without first computing the four expected frequencies. Example below illustrates the method.

Example: The mothers of 180 adolescents (some of them were graduates and others non- graduates) were asked whether they agree or disagree on a certain aspect of adolescent behaviour. Is the attitude of their mother related to their educational qualification? Data are arranged in a four fold table below:

	Agree	**Disagree**	
Graduate mother	30 (A)	50 (B)	80 (A+B)
Non graduate mother	70 (C)	30 (D)	100 (C+ D)
Total	100 (A +C)	80 (B +D)	180 (A+ B+ C+ D)

In four fold table Chi-square is calculated by the following formula:

$$\chi 2= N(AD-BC)^2/(A+B)(C+D)(A+C)(B+D)$$

Substitute for A,B,C,D, in the formula , we have

$$= 180\ (30\times30)-(50\times70)^2/80\times100\times100x80$$

$$=180(2600)^2/64000000$$

$$=19.01$$

$$\chi 2=19.01$$

df = (r-1)(c-1)

= (2-1)(2-1)

df = 1

On df = 1, the critical value at 0.05 level is 3.84 and on 0.01 is 6.63 (refer to the chi-square table given in the appendix of books on Statistics).

The obtained values is much higher therefore $\chi 2$is significant at .01 level, we reject the null hypothesis and conclude that attitude is related with education.

6.3 CHI-SQUARE TEST WHEN TABLE ENTRIES ARE SMALL (YATES CORRECTION)

When $\chi 2$ is computed from a table in which any experimental frequency is less than 5 then $\chi 2$ is not stable. Moreover in 2 × 2 contingency table (df=1) there are more chances of error. When any observed frequency is less than 5.

In these situations we make a correction for continuity known as Yates" correction.

To apply the Yates correction we subtract 0.05 from the absolute value of the difference between observed and expected frequency.

What is the reason for applying Yates correction and what is its effect upon $\chi 2$can be illustrated from the following example.

Example: A student is asked to respond on 10 objective type questions requiring response in 'yes' or 'no'. 7 times the student says 'yes' and 3 times the student says 'No'. Is this result different from what would have been obtained by merely guessing?

Table 6.9: Computation of $\chi 2$ with Yates Correction-1

	Yes	**No**	**Total**
Fo	7	3	10
Fe	5	5	10
fo- fe	2	-2	
Correction (*fo- fe*- 0.5)	1.5	1.5	
$(fo- fe)^2$	2.25	2.25	
$(fo- fe)^2/fe$	0.45	0.45	

If we do not apply the Yates correction then the value of $\chi 2$ will be different. This can be illustrated in the following table.

Table 6.10: Computation of $\chi 2$ with Yates Correction-1-1

	Yes	**No**	**Total**
fo	7	3	10
fe	5	5	10
fo- fe	2	-2	
$(fo- fe)^2$	4	4	
$(fo- fe)^2/fe$	4/5 = 0.20	4/5 = 0.20	

$$\chi 2=\sum[(fo- fe)^2/fe]$$

$$= .2+.2 =$$

$$0.4$$

df =(3-1)(2-1)

df=2

For df= 2, the critical values (refer to the table for chi-square given in the appendix of book on statistics) at 0.05 is 5.991 and on 0.01 is 9.210. In these examples although both the $\chi 2$are not significant but if we do not apply the correction then the probability of significance of $\chi 2$increased. We can conclude that failure to use the correction causes the probability of its being called significant considerably increased.

Yates Correction in 2×2 contingency

Example: 60 Students 50 from urban areas and 10 from rural area were administered attitude scale to measure the attitude towards dating. Some of them expressed positive attitudes and some expressed negative attitudes. Whether the attitudes of the students will be effected by their belonging to rural or urban areas. The data are given in the following table.

Table 6.11: Computation of $\chi 2$ with Yates Correction-1-1-1

	Positive	**Negative**	**Total**
Urban	20 (A)	30 (B)	50
Rural	30 (C)	7 (D)	10
Total	23 (A +C)	37 (B+D)	N= 60 (A+ B+ C+ D)

In this type of 2×2 table you can calculate chi-square in a different method which is given below:

$$\chi 2= N(|AD-BC|- N/2)^2/(A+B)(C+D)(A+C)(B+D)$$

$$= 60(|20\times7-30\times3|) -60/2]^2/ (50)(10)(23)(37)$$

$$= 60(|140-90|-30]^2/425500$$

$$=60\times400/425500$$

$$=24000/425500$$

$$= .056$$

Entering the table for chi-square we find that for df= 1 the value of Chi-square at 0.05 level is be 3.841. The obtained value is less than the table value and hence we retain the null hypothesis and conclude that the attitude

towards dating is not influenced by the person belonging to rural or urban areas.

Check Your Progress I

1) Complete the sentences:

a) The chi-square test is a useful method ……

……

……

b) Contingency table is a double entry or two way table in which ….

……

……

c) When the contingency table is 2 x 2 fold $\chi 2$may be calculated …..

……

……

d) To apply the Yates correction …….

……

……

6.4 LET US SUM UP

The chi-square test is used for two broad purposes. It is used as a test of goodness of fit and second as a test of independence. As a test of goodness of fit $\chi 2$ tries to determine how the observed frequencies are different from the frequency expected theoretically by hypothesis of equal probability and hypothesis of normal probability. The formula of $\chi 2$ is $\chi 2=\sum[(fo- fe)^2/fe]$. As a test of independence $\chi 2$ is applied for testing the relationship between two variables in two ways. In the 2 × 2 contingency table there are four cells A,B,C,D, and $\chi 2$ is calculated with the help of formula, $\chi 2= N(AD-BC)^2/(CA+B)(C+D)(A+C)(B+D)$. When in any cells entries are less than 5, Yate's correction is applied and the formula used to calculate is $\chi 2= N(|AD-BC|)-N/2/(A+B(C+D)(A+C)(B+D)$. The vertical lines |AD-BC| mean that the difference is to be taken as positive.

5.5 REFERENCES

Kerlinger, Fred N. (1963). Foundation of Behavioural Research, (2nd Indian reprint) Surjeet Publication, New Delhi.

Garrett, H.E. (1971), Statistics in Psychology & Education, Bombay, Vakils, Seffer & Simoss Ltd.

Guilford, J.P. (1973), Fundamental Statistics in Psychology & Education, Newyork, McGraw Hill.

6.6 KEY WORDS

Critical value: Table value according to the given df value. 32 Expected Frequency : The expected scores in a given class.

Expected frequencies: The frequencies for different categories which are expected to occur on the assumption that the given hypothesis is true.

Observed frequencies: The frequencies actually obtained from the performance of an experiment.

Yate's correction: The correction of -.5 in expected frequency.

Contingency table: A table having rows and columns where in each row corresponds to a level of one variable and each column to a level of another variable.

6.7 ANSWERS TO CHECK YOUR PROGRESS

Check Your Progress I

1) Complete the sentences:

a) The chi-square test is a useful method of comparing experimentally obtained results with those to be expected theoretically on some hypothesis.

b) Contingency table is a double entry or two way table in which the possession by a group of varying degree of the characteristics is represented.

c) When the contingency table is 2 x 2 fold $\chi 2$ may be calculated without first computing the four expected frequencies.

d) To apply the Yates correction we subtract 0.05 from the absolute value of thedifference between observed and expected frequency.

6.8 UNIT END QUESTIONS

1) Describe the properties of chi-square test.

2) What do you understand by the goodness of fit test ? Describe with examples.

3) What is the chi-square test for independence ? Explain with examples.

4) What precautions would you keep in mind while using chi-square.

5) Write short notes on: Yates correction and expected frequencies in contingency table.

UNIT 7 MANN WHITNEY U TEST*

Structure

7.0 OBJECTIVES

After reading this unit, you will be able to

- discuss Mann Whitney U test; and
- compute Mann Whitney U test for small sample, moderately large sample and large sample.

7.1 INTRODUCTION

A researcher wanted to study if significant difference exists between junior and senior managers with regard to their work motivation. The researcher collected data using standardised psychological test for work motivation and then had to carry out statistical analysis of the data. Initially, the researcher felt that independent t test cannot be used as the assumptions of parametric tests were not fulfilled. Therefore, the researcher decided to use Mann Whitney U test.

In unit 3 of this course, we discussed about t test. t test is parametric statistical technique that us used in order to measure the difference between two sample means. When the assumptions of parametric tests that we discussed in unit 2, are not met, it is not possible for us to use t test. In such a case the non-parametric counterpart of independent t test that is Mann Whitney U test can be effectively used in order to measure the difference between two sample means.

* Prof. Suhas Shetgovekar, Faculty, Discipline of Psychology, SOSS, IGNOU, Delhi

Thus, in the present unit, we will try to understand what is Mann Whitney U test and also learn how to compute it.

7.2 MANN WHITNEY U TEST

Mann Whitney U test can be explained as a powerful non-parametric test and an effective alternative to independent t test. It can be used adequately with variables that are continuous and discrete and can be effectively used with small sample. Though in terms of power, it is less powerful compared to independent t test, but its power is higher than other Non-parametric tests like median test (that will be discussed in the next unit) and Wilcoxon rank sum test.

Mann Whitney U test can be applied to find out whether two sample subgroups have been taken from a same population. It also helps in determining the distribution of data with regard to its median (Veeraraghavan and Shetgovekar, 2016).

The assumptions of non-parametric tests are applicable to Mann Whitney U test. Let us now discuss some of the assumptions of Mann Whitney U test (Mohanty and Misra, 2016):

1) The observations need to be independent.
2) The dependent variable that is taken in the research needs to display continuity.
3) The test is used when the two sample subgroups are independent and the data is ordinal. In case if the data is in interval or ratio form then the same is converted in to rank order.
4) The sample sub groups need to be independent and not correlated.

Check Your Progress I

1) State any one assumption of Mann Whitney U Test.

……

……

……

……

……

7.3 COMPUTATION OF MANN WHITNEY U TEST FOR SMALL SAMPLE

A school counsellor wanted to study whether gender difference exists in students with regard to mathematical ability. The same size was quiet sample as each of the sample subgroup had less than 8 participants. In such a

situation Mann Whitney U test for small sample can be computed. The data is given below, let us learn how Mann Whitney U can be computed in this context with the help of steps:

Table 7.1: Mathematical Ability Scores of Male and Female Students

Males (A)	Females (B)
6	10
13	16
7	12
11	14
5	17
N = 5	N = 5

Step 1: Null hypothesis is to be formulated.

The null hypothesis is formulated as follows:

H_o = There is no significant difference between males and females with regard to mathematical ability.

Step 2: Arranging the scores obtained by males and females in an ascending order, that is, from lowest to highest.

The codes for their groups, A for males and B for females also need to mentioned.

Table 7.1: Scores arranged in Ascending Order

5	6	7	10	11	12	13	14	16	17
A	**A**	**A**	**B**	**A**	**B**	**A**	**B**	**B**	**B**

Step 3: U is computed.

This is done by determining the number of scores for A (males) that are falling below the scores of B (females). This is done as follows:

For the score in group B of 10, the number of scores from group A that fall below 10 = 3

For the score in group B of 12, the number of scores from group A that fall below 12= 4

For the score in group B of 14, the number of scores from group A that fall below 14 = 5

For the score in group B of 16, the number of scores from group A that fall below 16 = 5

For the score in group B of 17, the number of scores from group A that fall below 17= 5

Thus, $U = 3 + 4 + 5 + 5 + 5 = 22$

Let us also compute U'

For the score in group A of 5, the number of scores from group B that fall below 5 = 0

For the score in group A of 6, the number of scores from group B that fall below 6 = 0

For the score in group A of 7 , the number of scores from group B that fall below 7 = 0

For the score in group A of 11, the number of scores from group B that fall below 11 = 1

For the score in group A of 13, the number of scores from group B that fall below 13 = 2

$$U' = 0 + 0 + 0 + 1 + 2 = 3$$

Thus, U = 22 and U' = 3

Further, the total of U and U' needs to be equal to $N_1 N_2$.

Thus,

$$U + U' = N_1 N_2$$

Thus,

$$U' = N_1 \times N_2 - U$$

$$= 5 \times 5 - 22$$

$$= 25 - 22$$

$$= 3$$

The value obtained is 3 and it is same as the value obtained earlier using the formula.

Step 4: Interpretation of the U value

To interpret the obtained U, the table for U is to be referred to, that can be found in appendix of books on Statistics. In our example $N_1 = 5$ and $N_2 = 5$. And the U = 3. The table value in this regard is .579. The U obtained by us is more than the table value and thus null hypothesis can be accepted and it can be said that the samples have been drawn from the same population.

Check Your Progress II

1) List the steps in computation of Mann Whitney U test for small sample.

……

……

……

7.4 COMPUTATION OF MANN WHITNEY U TEST FOR MODERATELY LARGE SAMPLES

In the previous section, we discussed about how Mann Whitney U test can be computed for small sample. In this section, we will discuss how it is computed for a moderately large sample, that is a sample size that falls between 9 and 20.

Let us now understand how to compute Mann Whitney U test for moderately large sample with the help of an example.

A researcher was carrying out a study on how two different teaching methods, lecture and group discussion have an effect on academic performance of the students. The first group (Group A) consisted of 11 students and were taught using lecture method and the other group (Group B) consisted of 9 students and were taught using group discussion method. The date obtained is given in table 7.3.

Table 7.3: Academic Performance of the Students

Group A	Group B
49	51
39	48
65	36
62	37
34	51
45	56
46	59
65	78
65	83
66	
85	
N= 11	N= 9

Step 1: Formulate the Null hypothesis.

H_o = There will be no significant difference between the two groups with regard to their academic performance.

Step 2:Ranks are to be given to the combined scores from both groups.

The scores of both the groups are combined and then arranged in an ascending order and ranks are assigned to them. Rank 1 is allotted to the lowest score and the next above it will be given a rank 2 and so on. In case of tied scores, the procedure similar to the one that we discussed in BPCC104 while computing Spearman's rho is used. Thus, the re are two 51 scores, the ranks allotted are (9+ 10)/ 2= 9.5 and the next score gets rank 11. Further, there are three 65 score. Thus, (14 + 15+ 16)/ 3= 15. Thus, the rank assigned to this score is 15 and the next score gets rank 17.

Table 7.4: Academic Performance of the Students

Group 1 (Lecture Method)	Rank (R1)	Group 2 (Audio-video method)	Rank (R2)
49	8	51	**9.5 (tied ranks)**
39	4	48	7
65	**15 (tied ranks)**	36	2
62	13	37	3
34	1	51	**9.5 (tied ranks)**
45	5	56	11
46	6	59	12
65	**15 (tied ranks)**	78	18
65	**15 (tied ranks)**	83	19
66	17		
85	20		
N= 11	**$\sum R_1 = 119$**	N= 9	**$\sum R_2 = 91$**

Step 3: Compute the summation of the ranks in each group

Thus R_1= 119and R_2= 91. (R1 + R2 = 210)

The same can be checked as using a formula as follows:

$$N (N +1)/ 2$$

where $N = N_1 + N_2$

$= 20 (20 +1)/2$

$= 20 \times 21/ 2$

$= 420/ 2$

$= 210$

The value obtained is 210 and is equal to the value we obtained after adding $R_1$and R_2.

Step 4: Compute U

U can be computed using the following formula

$$U = N_1 N_2 + [N_1 (N_1 + 1)]/ 2 - \sum R_1$$

The formula forU' is as follows:

$$U' = N_1 N_2 + [N_2 (N_2 + 1)/2] - \sum R_2$$

Let us now compute U using the values given in table 7.3

$$U = N_1 N_2 + N_1(N_1 + 1) / 2 - \sum R_1$$

$= 11 \times 9 + [11 (11 +1)] /2 - 119$

$= 99+ 11 (12) / 2- 119$

$= 99 + 132 /2 - 119$

$= 99 + 66 - 119$

$= 165- 119 = 46$

Step 5: Find out U' using the formula given in the box above.

$$U' = N_1 N_2 + [N_2 (N_2 + 1)/2] - \sum R_2$$

$= 11 \times 9+ [9(9+1) / 2] – 91$

$= 99 + (9 \times 10) / 2] – 91$

$= 99+ 90 / 2 – 91$

$= 99 + 45 – 91$

$= 144 – 91 = 53$

Thus, U = 46 and U' = 53

Further, the total of U and U' needs to be equal to $N_1 N_2$.

$U' = N_1 N_2 – U$

$=11 \times 9 - 46$

$= 99- 46$

$= 53$

Step 6: Interpretation of U

In order to interpret the obtained U, we need to refer to the table value that is given in the table for Mann Whitney U in appendix of any Statistics book. Referring to such a table, the table value with the large sample as 11 and small sample as 9 is 18 at 0.01 level of significance and 27 at 0.05 level of significance. The obtained U is 46 is higher than the table value. Thus, the null hypothesis is accepted and it can be said that there is no significant difference in the academic performance with regard to the two teaching methods.

Check Your Progress III

1) What is the formula to compute U.

……

……

……

……

……

……

……

……

……

7.5 COMPUTATION OF MANN WHITNEY U TEST FOR LARGE SAMPLES

A large sample is a sample that is more than 20. The distribution for such a sample could appear similar to normal distribution and thus, in this case the formula for z can be employed in order to find out if significant difference exists between the two groups.

Let us now try to understand the computation of Mann Whitney U test for large samples with the help of an example.

A researcher wanted to find if significant difference exists in attitude towards environment of individuals in rural and urban areas. The data collected is give below in table

Table 7.5: Scores on Attitude tiowards Environment along with their ranks .

Individuals from rural area	**Rank 1**	**Individuals from urban area**	**Rank 2**
42	9	85	28
45	10	109	37
46	11	87	29
95	34	88	30
23	1	90	31
94	33	91	32
33	6	26	2
35	7	100	35
32	5	102	36
40	8	27	3
30	4	110	38
83	27	76	26
72	24	73	25
69	23	47	13
47	13	50	16
59	22	57	21
47	13	50	16
51	18	56	20
54	19	50	16
120	39	121	40
N= 20	$\sum R_1$ =326	N= 20	$\sum R_2$ =494

Step 1:The data collected in tabulated as is give in table 7.5.

The data is to be organised in a table form with scores as obtained by participants in each group.

Step 2: Rank the scores

As was done in previous section. In similar manner, ranks are to be given to the scores. Rank 1 is given to the lowest score and the last score gets the highest rank. Tied ranks are also given as was discussed earlier.

Step 3: Ranks for each group are summated.

Thus $\sum R_1 = 326$ and $\sum R_2 = 494$

$R_1 + R_2 = 326 + 494 = 820$

This can be checked using the formula as follows:

$$N (N +1)/ 2$$

$$\text{where } N = N_1 + N_2$$

$$= 40(40 +1)/2$$

$$= 40 \times 41/ 2$$

$$= 1640/ 2$$

$$= 820$$

The value obtained is 210 and is equal to the value we obtained after adding R_1 and R_2.

Step 4: Compute U

U can be computed using the following formula

$$U = N_1 N_2 + [N_1 (N_1 + 1)]/ 2 - \sum R_1$$

The formula for U' is as follows:

$$U' = N_1 N_2 + [N_2 (N_2 + 1)/2] - \sum R_2$$

Let us now compute U using the values given in table 7.3

$$U = N_1 N_2 + N_1(N_1 + 1) / 2 - \sum R_1$$

$$= 20 \times 20 + [20(20 +1)] /2 - 326$$

$$= 400 + 20 (21) / 2 - 326$$

$$= 400 + 420 /2 - 326$$

$$= 400 + 210 - 326$$

$$= 610 - 326 = 284$$

Step 5: Find out U' using the formula given in the box above.

$$U' = \mathbf{N_1 N_2 + [N_2 (N_2 + 1)/2] - \sum R_2}$$

$$= 20 \times 20 + [20(20 +1)] /2 - 494$$

$$= 400+ 20 (21) / 2- 494$$

$$= 400 + 420 /2 - 494$$

$$= 400 + 210 - 494$$

$$= 610- 494= 116$$

Thus, U = 284 and U' = 116

Further, the total of U and U' needs to be equal to $N_1 N_2$.

$$U' = N_1 N_2 – U$$

$$=20 \times 20 - 284$$

$$= 400- 284$$

$$= 116$$

Step 5: Compute z

Now we need to compute z with the help of the following formula:

$$\mathbf{z = U – [(N_1 N_2) /2] / \sqrt{\{(N_1)(N_2)(N_1+N_2+1)/ 12\}}}$$

$$= 284 – [20x20) / 2] / \sqrt{[(20) (20)(20+20+1)/12]}$$

$$= 284 - [400 /2] / \sqrt{[400x 41/12]}$$

$$= 284 – 200 / \sqrt{1366.67}$$

$$= 84 / 36.97$$

$$= 2.27$$

In place of U, you could have taken U'. The z value obtained in either case wold remain same.

$$\mathbf{z = –U' [(N_1N_2) /2] / \sqrt{\{(N_1)(N_2)(N_1+N_2+1)/ 12\}}}$$

$$= 116– [20x20) / 2] / \sqrt{[(20) (20)(20+20+1)/12]}$$

$$= 116 - [400 /2] / \sqrt{[400x 41/12]}$$

$$= 116 – 200 / \sqrt{1366.67}$$

$$= 84 / 36.97$$

$$= 2.27$$

Thus, the z value is obtained as 2.27.

Step 6: Interpretation of z value

To interpret the z value we will be required to refer to the table value from the table for z value that is given in appendix of any Statistics book.

In this context, it can be said that if the z value is ≥1.96, U is interpreted as significant at 0.05 level of significance (for two tailed test) and thus null hypothesis is rejected. If the obtained z value is ≥2.58, U is interpreted as significant at 0.01 level of significance (for two tailed test) and thus null hypothesis is rejected. With regard to one tailed test the z value needs to be is ≥ 1.64 at 0.05 level of significance and ≥ 2.33 at 0.01 level of significance order to reject the null hypothesis.

The z value obtained by us is 2.27 which is less than ≥2.58 for 0.01 level of significance but of more than 1.96, thus it can be said that U is significant at 0.05 level and null hypothesis can be rejected.

This indicates that the null hypothesis is rejected and one may conclude that there is a significant difference in the attitude towards environment of individuals belonging to rural and urban area of residence.

Check Your Progress IV

1) What is the formula for z?

……

……

……

……

……

7.6 LET US SUM UP

In the present unit, we discussed about Mann Whitney U test. Mann Whitney U test can be explained as a powerful Non-parametric test and an effective alternative to independent t test. It can be used adequately with variables that are continuous and discrete and can be effectively used with small sample. Though in terms of power, it is less powerful compared to independent t test, but its power is higher than other non-parametric tests like median test (that will be discussed in the next unit) and Wilcoxon rank sum test. After understanding the test, we also discussed about its assumptions. Further, we learnt to compute Mann Whitney U test for small sample, moderately large sample and large sample with the help of steps and examples.

7.7 REFERENCES

King, Bruce. M; Minium, Edward.W. (2008).*Statistical Reasoning in the Behavioural Sciences*. Delhi: John Wiley and Sons, Ltd.

Mangal, S. K. (2002). *Statistics in Psychology and Education*.new Delhi: Phi Learning Private Limited.

Minium, E. W., King, B. M., & Bear, G. (2001).*Statistical Reasoning in Psychology and Education*. Singapore: John-Wiley.

Mohanty, B and Misra, S. (2016). *Statistics for Behavioural and Social Sciences*. Delhi: Sage.

Veeraraghavan, V and Shetgovekar, S. (2016). *Textbook of Parametric and Non-parametric Statistics*. Delhi: Sage.

7.8 KEY WORD

Mann Whitney U Test: Mann Whitney U test can be explained as a powerful Non-parametric test and an effective alternative to independent t test. It can be used adequately with variables that are continuous and discrete and can be effectively used with small sample.

7.9 ANSWERS TO CHECK YOUR PROGRESS

Check Your Progress I

1) State any one assumption of Mann Whitney U Test.

   The observations need to be independent.

Check Your Progress II

1) List the steps in computation of Mann Whitney U test for small sample.

   The steps in computation of Mann Whitney U test for small sample are:

   Step 1: Null hypothesis is to be formulated.

   Step 2: Arranging the scores obtained by males and females in an ascending order, that is from lowest to highest.

   Step 3: U is computed.

   Step 4: Interpretation of the U value

Check Your Progress III

1) What is the formula to compute U.

   $U = N_1 N_2 + [N_1 (N_1 + 1)]/ 2 - \sum R_1$

Check Your progress IV

1) What is the formula for z?

   $z = U - [(N1N2) /2] / \sqrt{\{(N1)(N2)(N1+N2+1)/ 12\}}$

7.10 UNIT END QUESTIONS

1) Explain Mann Whitney U test with a focus on its assumptions.

2) Compute Mann Whitney U test for the following data:

Data A	Data B
4	9
6	3
7	16
2	8

3) A researcher was carrying out a study on emotional intelligence of early and late adolescents. The date is given as follows: Compute Mann Whitney U test for the same.

Early Adolescents	Late Adolescents
23	56
23	65
45	43
43	45
37	65
43	56
67	
87	
65	
45	

UNIT 8 MEDIAN TEST *

Structure

8.0 OBJECTIVES

After reading this unit, you will be able to,

- discuss median test and its assumptions; and
- compute one sample median test, compute median test with two independent samples and median test for more than two independent samples.

8.1 INTRODUCTION

A researcher wanted to study the median academic scores obtained by students in 9th standard class of a school. Yet another researcher wanted to compare the median of scores obtained by individuals in urban and rural areas on attitude towards environment.

In both these case, the researchers can use median test in order to analyse the data collected by them. In the previous unit, we discussed about Mann-Whitney U test. In the present unit, we will discuss about median test that can be used in similar way as the Mann Whitney U test. We will discuss median test and its assumptions. And will further explain the computation of one sample median test. compute median test with two independent samples and median test for more than two independent samples.

* Prof. Suhas Shetgovekar, Faculty, Discipline of Psychology, SOSS, IGNOU, Delhi

8.2 MEDIAN TEST

It can be said that median test is a non-parametric alternative of t test. It can be employed to know if the median of a certain data is same as that of assumed median in which case one sample median test can be used. It can also be used to know whether two or more samples that are independent and equal or not equal with regard to their size, have different median or not. Thus, medians of two sample sub groups are compared and these samples can be equal or unequal in size.

Median test is a non-parametric statistic and thus it can be used with distribution that is normal or not normal. Further it can be used with variables having either continuous and discontinuous measurements. It can also be conveniently used with small sample. though when compared with Mann-Whitney U test, it is less powerful.

We need to remember here that as such the assumptions of Non-parametric test are applicable here as well. Let us now look at the assumptions of the median test (Mohanty and Misra, 2016) :

1) The scores in the samples (two or more) can be measured in terms of ordinal scale of measurement.
2) Every score in the sample occurs in a random manner.
3) Every score is independent of other scores.
4) The samples (two or more) need to be independent and not correlated to each other.

Check Your Progress I

1) When can the median test be used?

……

……

……

……

……

8.3 COMPUTATION OF ONE SAMPLE MEDIAN TEST

One sample median test is used in order to know if the median of certain sample is equal to the population median that is assumed (veeraraghavan and Shetgovekar, 2016).

Let us now try to understand the computation of one sample median test with the help of steps and example.

An industrial psychologist wanted to know the average years of experience of newly appointed junior managers in the organisation. The data collected is given as follows:

2	3	6	4	3	2	4	6	10	9	7	8	3	2

Step 1: Null hypothesis is to be formulated.

H_0 = The median years of experience of junior managers will be the same as the assumed median years of experience.

Step 2: Assuming median sample years of experience as 5

Step 3: Coding the scores

The scoreshaving lower value than the assumed median are to be given the code as '0' and the scores having value more than the assumed median

Those having value more than the assumed median will be given a code as '1'. The values that are the same as that of the assumed median will not be considered.

2	3	6	4	3	2	4	6	10	9	7	8	3	2
0	0	1	0	0	0	0	1	1	1	1	1	0	0

Step 4: Rearranging the data

The data is now rearranged in terms of frequency of scores above assumed median and below assumed median as can be seen below:

Above Assumed Median	Below Assumed Median
6	8

Step 5: Frequency expected (*fe*) is computed using the formula:

N= 14, that is the total number of participants.

$$fe = 1/2 \text{x } N$$

$$= 1/2 \text{ x } 14$$

$$= 7$$

Thus

Frequency observed (*fo*)	Frequency expected (*fe*)
6	7
8	7

Step 6: Chi Square is computed using the following formula:

$$\chi 2 = \Sigma\ [(fo - fe)^2]/fe$$

Frequency observed (*fo*)	Frequency expected (*fe*)	*fo-fe*	$(fo-fe)^2$	$(fo-fe)^2/fe$
6	7	-1	1	0.14
8	7	1	1	0.14
				X^2=0.28

The X^2 =0.28

Step 7: Interpretation

In order to interpret the obtained chi- square, we need to first know the degree of freedom that will be

df= (r-1) (k-1)

r means rows

k means columns.

Thus,

= (2-1) (2-1)

= 1

Thus, df= 1

Now we can refer to the chi- square table that is given in appendix of any Statistics book. The critical value of chi square for df= 1, is 6.635 at 0.01 level of significance and 3.841at 0.05 level of significance. The chi- square obtained by us is 0.28 which is lower than the table value at both the levels. Thus the null hypothesis is accepted and it can be said that there is no significsnt difference between the assumed median and the sample median years of experience of the newly appointed junior managers.

Check Your Progress II

1) List the steps in computation of one sample median test.

……

……

……

……

……

8.4 COMPUTATION OF MEDIAN TEST FOR TWO INDEPENDENT SAMPLES

As was mentioned earlier, median test can also be used in order to know whether two samples that are independent differ in terms of their median or not. In the present section, we will learn how to compute median test for two independent samples and we will do so with the help of steps and an example.

The scores obtained by employees of public and private banks on work motivation are given as follows

Employees from private banks	12	13	13	15	17	18	19	21	22	21	11	23
Employees from public banks	23	12	12	14	14	21	23	21	23	21	20	

Step1: Formulate null hypothesis.

The scores obtained by the employees from private and public sector banks are not different from the common median.

Step 2: Computation of common median

A common median is to be computed for both the groups.

$$n= n_1 + n_2$$

$$= 12 +11= 23$$

$$\text{Median} = n=1/ 2^{th} \text{ score}$$

$$= 23 + 1/2$$

$$= 12^{th} \text{ score}$$

Now we need to arrange the scores from both the groups in an ascending order and then identify the 12th score

11, 12, 12, 12, 13, 13, 14, 15, 17, 18, **19**, 20, 21, 21, 21,21, 21, 22, 23, 23 , 23, 23

The 12^{th} score is 19. Thus, the common median is obtained as 19.

Step 3: Assignment of signs to the scores.

The scores in each group are assigned positive (+) or negative (-) signs depending on whether they fall above or below the median. If above median, positive sign is given and if below or equal to the median, then negative sign is given.

Employees from private banks	-	-	-	-	-	-	-	+	+	+	-	+
Employees from public banks	+	-	-	-	-	+	+	+	+	+	+	

Step 4: 2 x 2 contingency table is created.

Such a table is created so that chi square can be computed.

Employees	**Positive signs**	**Negative signs**	**Total**
Private banks	4	8	12
Public banks	7	4	11
Total	11	12	23

Step 5: Compute chi- square

$$\Sigma[(f_0 - f_e)^2 / f_e]$$

For that we first need to compute the frequency expected (f_e) as follows:

f_e for cell 1: (12 × 11)/23 = 132/23 = 5.74

f_e for cell 2: (11 × 11)/23 = 121/23 = 5.74

f_e for cell 4: (12 × 12)/23 = 144/23 = 6.26

f_e for cell 5: (12 × 12)/23 = 132/23 = 5.74

Frequency observed (f_o)	**Frequency expected (f_e)**	$f_o - f_e$	$(f_o - f_e)^2$	$(f_o - f_e)^2/f_e$
4	5.74	-1.74	3.0276	0.53
7	5.26	1.74	3.0276	0.57
8	6.26	1.74	3.0276	0.48
4	5.74	-1.74	3.0276	0.53
				$X^2 = 2.11$

The $\chi 2$ obtained is 2.11

Calculate the degrees of freedom applying the formula (r-1)(k-1). The df for the data in the table is = (2-1)(2-1) = 1 × 1 = 1

Step 6: Interpretations

Now we can refer to the chi- square table that is given in appendix of any Statistics book. The critical value of chi square for df= 1, is 6.635 at 0.01 level of significance and 3.841at 0.05 level of significance. The chi- square obtained by us is 2.11 which is lower than the table value at both the levels. Thus the null hypothesis is accepted and its can be said that the scores obtained by the employees from private and public sector banks are are not different from the common median.

Check Your Progress III

1) State the formula for chi- square while computing median test for two independent samples.

8.5 COMPUTATION OF MEDIAN TEST FOR MORE THAN TWO INDEPENDENT SAMPLES

Median test can also be computed for more than two independent samples. The procedure for computation is similar to that of median test for two independent sample, though the formula used for computation of chi-square will differ. Let us try to understand the computation with the help of steps and example.

The scores obtained by three groups (Groups A, B and C) of students on achievement motivation are given as follows

Group A	4	27	6	4	26	8	9	2	7	3
Group B	3	6	9	12	16	20	21	26	5	29
Group C	4	8	10	14	15	18	23	12	15	17

Step1: Formulate null hypothesis.

The scores obtained by the three groups of students are not different from the common median.

Step 2: Computation of common median

A common median is to be computed for the three groups.

$$n= n_1 + n_2 + n3$$

$$= 10 +10 +10 = 30$$

$$\text{Median} = n=1/ 2^{th} \text{ score}$$

$$= 30 + 1/2$$

$$= 15.5^{th} \text{ score}$$

Now we need to arrange the scores from the three groups in an ascending order and then identify the 15.5th score

2, 3, 3, 7, 4, 4, 4, 5, 6, 6, 8, 8, 9, 9, **10, 12**, 12, 14, 15,15, 17, 16, 18, 20, 21, 23, 26 26, 27, 29

The median is the 15,5th score , that is an average of 15th and 16th score. In our example, the 15th score is 10 and t16h score is 12. The average of these scores would be (10 +12)/2+ 11. Thus, the common median is 11.

Step 3: Assignment of signs to the scores.

The scores in each group are assigned positive (+) or negative (-) signs depending on whether they fall above or below the median. If above median, positive sign is given and if below or equal to the median, then negative sign is given.

Group A	-	+	-	-	+	-	-	-	-	-
Group B	-	-	-	+	+	+	+	+	-	+
Group C	-	-	-	+	+	+	+	+	+	+

Step 4: 2 x 2 contingency table is created.

Such a table is created so that chi square can be computed.

	Positive sign	**Negative sign**	**Total**
Group A	2	8	10
Group B	6	4	10
Group C	7	3	10
Total	15	15	n= 30

Step 5: Compute chi- square

The formula used is as follows:

$$x^2 = \sum[(f_0 - f_e)^2 / f_e]$$

For that we first need to compute the frequency expected (f_e) as follows:

$f_e$for cell 1: (10 x 15)/ 30 = 150/ 30 = 5

f_e for cell 2: (10 x 15)/ 30= 150/ 30 = 5

$f_e$for cell 3: (10 x 15)/ 30= 150/ 30 = 5

f_e for cell 4: (10 x 15)/ 30= 150/ 30 = 5

f_e for cell 5: (10 x 15)/ 30= 150/ 30 = 5

$f_e$for cell 6: (10 x 15)/ 30 = 150/ 30 = 5

Frequency observed (*fo*)	Frequency expected (*fe*)	*fo-fe*	$(fo-fe)^2$	$(fo-fe)^2/fe$
2	5	-3	9	1.8
6	5	1	1	0.2
7	5	2	4	0.8
8	5	3	9	1.8
4	5	-1	1	0.2
3	5	-2	4	0.8
				$X^2 = 5.6$

The chi-square obtained is 5.6

Calculate the degrees of freedom applying the formula (r-1)(k-1). The df for the data in the table is = (3-1)(2-1) = 2 x 1= 2

Step 6 Interpretation

We need to refer to the table value for chi square that is given in any statistics book.

For df = 2 the table value is 5.99 for 0.05 level of significance and 13.82 at 0.01 level of significance. The chi square obtained by us is 5.6 which is lower than the table values at both the levels of significance. Thus, the null hypothesis that the scores obtained by the three groups of students are not different from the common median.is accepted.

Check Your Progress IV

1) State the formula for chi- square while computing median test for more than two independent samples.

8.6 LET US SUM UP

To summarise, median test is a non parametric alternative of t test. It can be employed to know if the median of a certain data is same as that of assumed median in which case one sample median test can be used. It can also be used

to know whether two or more samples that are independent and equal or not equal with regard to their size, have different median or not. Thus, medians of two sample sub groups are compared and these samples can be equal or unequal in size. The assumptions for median test were also discussed. The unit then focused on the computation one sample median test. compute median test with two independent samples and median test for more than two independent samples.

8.7 REFERENCES

King, Bruce. M; Minium, Edward.W. (2008).*Statistical Reasoning in the Behavioural Sciences*. Delhi: John Wiley and Sons, Ltd.

Mangal, S. K. (2002). *Statistics in Psychology and Education*.new Delhi: Phi Learning Private Limited.

Minium, E. W., King, B. M., & Bear, G. (2001).*Statistical Reasoning in Psychology and Education*. Singapore: John-Wiley.

Mohanty, B and Misra, S. (2016). *Statistics for Behavioural and Social Sciences*. Delhi: Sage.

Veeraraghavan, V and Shetgovekar, S. (2016). *Textbook of Parametric and Non-parametric Statistics*. Delhi: Sage.

8.8 KEY WORD

Median test: Median test is a non parametric alternative of t test. It can be employed to know if the median of a certain data is same as that of assumed median in which case one sample median test can be used. It can also be used to know whether two or more samples that are independent and equal or not equal with regard to their size, have different median or not.

8.9 ANSWERS TO CHECK YOUR PROGRESS

Check Your Progress I

1) When can the median test be used?

   Median can be used to know if the median of a certain data is same as that of assumed median in which case one sample median test can be used. It can also be used to know whether two or more samples that are independent and equal or not equal with regard to their size, have different median or not.

Check Your Progress II

1) List the steps in computation of one sample median test.

   The steps in computation of one sample median test are:

Step 1: Null hypothesis is to be formulated.

Step 2: Assuming median

Step 3: Coding the scores

Step 4: Rearranging the data

Step 5:Frequency expected (*fe*) is computed

Step 6: Chi Square is computed

Step 7: Interpretation

Check Your Progress III

1) State the formula for chi- square while computing median test for two independent samples.

$$\sum[(f_0 - f_e)^2 / f_e]$$

Check Your Progress IV

1) State the formula for chi- square while computing median test for more than two independent samples.

$$\sum[(f_0 - f_e)^2 / f_e]$$

8.10 UNIT END QUESTIONS

1) Explain median test with a focus on its assumptions.
2) Compute one sample median test for the following data:

34	32	22	34	43	45	56	54	56	43	22	36	43	33

3) Explain the computation of median test for two independent samples with the help of steps and examples.
4) Compute median test for more than two independent samples for the following dats:

Sample A	4	5	4	6	8	22	14	45	32	2
Sample B	6	7	45	5	6	32	16	6	4	7
Sample C	4	8	3	6	44	34	32	4	10	12
Sample D	3	3	4	6	5	12	12	8	9	13

UNIT 9 KRUSKAL- WALLIS ANALYSIS OF VARIANCE*

9.0 OBJECTIVES

After reading this unit, you will be able to:

- explain Kruskal- Wallis ANOVA;
- enumerate the conditions when this test can be applied;
- compute Kruskal-Wallis ANOVA test; and
- Compare Kruskal- Wallis ANOVA with one way ANOVA of parametric test.

9.1 INTRODUCTION

So far in previous unit, we studied about statistical tests when we wish to compare two groups (t test if data is from a normal population, Mann-Whitney U test if there are no assumptions about the distribution of the data), but what if there are more than two groups that require comparison? One may think that we may apply the same tests in that condition too. Like for example, if there are three groups say A, B, and C, one may see the difference between A&B, B&C and A&C. This may not look so cumbersome. Now, think if we need to compare 5 groups, A, B, C, D, E, the number for comparison tests we need to do would be 10 (A&B, A&C, A&D, A&E, B&C, B&D, B&E, C&D, C&E, D&E). And what if we need to compare 6 groups? Number of two sample test in these cases become too cumbersome and may not be feasible at all. This may further lead to

* Ms. Sonia Puar, Faculty, Amity University, Noida

unnecessary calculations and also give rise to type I error. The answer in these cases when we have more than two groups (>2 groups) to be compared is to conduct Analysis of Variance .

9.2 KRUSKAL-WALLIS ANOVA TEST

Let us revise ANOVA before we go on to discuss about Kruskal- Wallis ANOVA test.

The term Analysis of Variance (for which the acronym ANOVA is often employed) describes a group of inferential statistical procedures developed by the British statistician Sir Ronald Fisher. Analysis of Variance is all about examining the amount of variability in a y (response) variable and trying to understand where that variability is coming from. One way that you can use ANOVA is to compare several populations regarding some quantitative variable, y. The populations you want to compare constitute different groups (denoted by an x variable), such as political affiliations, age groups, or different brands of a product. ANOVA is also particularly suitable for situations involving an experiment where you apply certain treatments (x) to subjects, and you measure a response (y).

Null hypothesis (H0): Population means are equal. There will be no difference in the population means.

Alternative hypothesis (H1): Population means are not equal. There will be difference in the means of the different populations.

The logic used in ANOVA to compare means of multiple groups is similar to that used with the t-test to compare means of two independent groups. When one way ANOVA is applied to the special case of two groups, this one way ANOVA gives identical results as the t-test.

Not surprisingly, the assumptions needed for the t-test are also needed for ANOVA. We need to assume:

1) random, independent sampling from the k populations;
2) normal population distributions;
3) equal variances within the k populations.

Assumption 1 is crucial for any inferential statistic. As with the t-test, Assumptions 2 and 3 can be relaxed when large samples are used, and Assumption 3 can be relaxed when the sample sizes are roughly the same for each group even for small samples. (If there are extreme outliers or errors in the data, we need to deal with them first.)

9.2.1 Kruskal- Wallis ANOVA Test

When there are more than two groups or k number of groups to be compared, ANOVA is utilised but, again since ANOVA is a parametric statistics and

requires assumption of normality as a key assumption, we need to also be aware of its non-parametric counterpart. The Kruskal- Wallis test compares the medians of several (more than two) populations to see whether they are all the same or not. The Kruskal- Wallis test is a non-parametric analogue to ANOVA. It can be viewed as ANOVA based on rank transformed data.

That is, the initial data are transformed to their associated ranks before being subjected to ANOVA. In other words, it's like ANOVA, except that it is computed with medians and not means. It can also be viewed as a test of medians.

The null and alternative hypotheses may be stated as:

H_0: the population medians are equal

H_1: the population medians differ

The Kruskal- Wallis one way Analysis of Variance by ranks (Kruskal, 1952) and (Kruskal and Wallis, 1952) is employed with ordinal (rank order) data in a hypothesis testing situation involving a design with two or more independent samples.The test is an extension of the Mann-Whitney U test (Test 12) to a design involving more than two independent samples and, when k = 2, the Kruskal- Wallis one way Analysis of Variance by ranks will yield a result that is equivalent to that obtained with the Mann-Whitney U test.

If the result of the Kruskal- Wallis one way Analysis of Variance by ranks is significant, it indicates there is a significant difference between at least two of the sample medians in the set of k medians. As a result of the latter, the researcher can conclude there is a high likelihood that at least two of the samples represent populations with different median values.

In employing the Kruskal- Wallis one way Analysis of Variance by ranks one of the following is true with regard to the rank order data that are evaluated:

a) The data are in a rank-order format, since it is the only format in which scores are available; or

b) The data have been transformed into a rank-order format from an interval/ratio format, since the researcher has reason to believe that one or more of the assumptions of the single-factor between-subjects Analysis of Variance (which is the parametric analog of the Kruskal-Wallis test) are saliently violated.

It should be noted that when a researcher decides to transform a set of interval/ ratio data into ranks, information is sacrificed. This latter fact accounts for why there is reluctance among some researchers to employ non-parametric tests such as the Kruskal- Wallis oneway Analysis of Variance by ranks, even if there is reason to believe that one or more of the assumptions of the single factor between subjects Analysis of Variance have beenviolated.

Various sources [Conover (1980, 1999), Daniel (1990), and Marascuilo and McSweeney (1977)] note that the Kruskal-Wallis one way Analysis of Variance by ranks is based on the following assumptions:

a) Each sample has been randomly selected from the population it represents;

b) The k samples are independent of one another;

c) The dependent variable (which is subsequently ranked) is a continuous random variable. In truth, this assumption, which is common to many non-parametric tests, is often not adhered to, in that such tests are often employed with a dependent variable which represents a discrete random variable; and

d) The underlying distributions from which the samples are derived are identical in shape.

The shapes of the underlying population distributions, however, do not have to be normal.

Maxwell and Delaney (1990) point out that the assumption of identically shaped distributions implies equal dispersion of data within each distribution. Because of this, they note that, like the single factor between subjects Analysis of Variance , the Kruskal- Wallis one way Analysis of Variance by ranks assumes homogeneity of variance with respect to the underlying population distribution. Because the latter assumption is not generally acknowledged for the Kruskal- Wallis one way Analysis of Variance by ranks, it is not uncommon for sources to state that violation of the homogeneity of variance assumption justifies use of the Kruskal- Wallis one way Analysis of Variance by ranks in lieu of the single factor between subjects Analysis of Variance.

It should be pointed out, however, that there is some empirical research which suggests that the sampling distribution for the Kruskal- Wallis test statistic is not as affected by violation of the homogeneity of variance assumption as is the F distribution (which is the sampling distribution for the single-factor between- subjects Analysis of Variance).

One reason cited by various sources for employing the Kruskal- Wallis one way Analysis of Variance by ranks is that by virtue of ranking interval/ratio data a researcher can reduce or eliminate the impact of outliers. In t test for two independent samples, since outliers can dramatically influence variability, they can be responsible for heterogeneity of variance between two or more samples. In addition, outliers can have a dramatic impact on the value of a sample mean.

Check Your Progress I

1) List any one assumption of Kruskal- Wallis one way ANOVA.

9.3 COMPUTATION OF KRUSKAL-WALLISANOVA

Let us us now understand the computation of Kruskal-Wallis one way ANOVA with the help of steps and example.

Step 1: Rank all the numbers in the entire data set from smallest to largest (using all samples combined); in the case of ties, use the average of the ranks that the values would have normally been given.

Step 2: Total the ranks for each of the samples; call those totals T1, T2, . . ., Tk, where k is the number of groups or populations.

Step 3: Calculate the Kruskal-Wallis test statistic with the help of following formula:

$$H=[12/N(N+1)][\Sigma((\Sigma R)2\ /n)]–3(N+1)$$

N = the total number of cases

n = the number of cases in a given group

$(\Sigma R)^2$ = the sum of the ranks squared for a given group of subjects

Step 4: Find the *p*-value.

Step 5: Make your conclusion about whether you can reject Ho by examining the p-value.

Example of a Small Sample: In a Study, 12 participants were divided into three groups of 4 each, they were subjected to three different conditions, A (Low Noise), B (Average Noise), and C (Loud Noise). They were given a test and the errors committed by them on the test were noted and are given in the table below.

Participant No.	Condition A (Low Noise)	Participant No.	Condition B (Average Noise)	Participant No.	Condition C (Loud Noise)
1	3	5	2	9	10
2	5	6	7	10	8
3	6	7	9	11	7
4	3	8	8	12	11

The researcher wishes to know whether these three conditions differ amongst themselves. and there are no assumptions of the probability. To apply Kruskal- Wallis test, following steps would be taken:

Step 1: Rank all the numbers in the entire data set from smallest to largest (using all samples combined); in the case of ties, use the average of the ranks that the values would have normally been given.

Condition A	Ranks T 1	Condition B	Ranks T2	Condition C	Ranks T3
3	2.5	2	1	10	11
5	4	7	6.5	8	8.5
6	5	9	10	7	6.5
3	2.5	8	8.5	11	12
	ΣT1 = 14		ΣT2 = 26		ΣT3 = 38

Step 2: Total the ranks for each of the samples; call those totals $T1, T2, \ldots, Tk$, where k is the number of populations. T1 =14, T2=26, T3=38

Step3: Calculate H

$$H = [\ 12 / N\ (N+1)\ ]\ [\ \Sigma((\Sigma R)^2 / n)\ ] - 3(N + 1)$$

N = 12

n = 4

$(\Sigma R)^2 = (14+ 26+ 38)^2 = 6084$

$$H= [12/\ 12\ (12 + 1)\ ]\ [\ (14^2/4) + (26^2/4) + (38^2/4)] - 3\ (12+ 1)$$

$$H= [12/156]\ [49 + 169 + 361] - 39$$

$$H= (0.076 \times 579) - 39$$

$$H= 44.525 - 39$$

$$H= 5.537$$

Step 4: Find the p-value.

Since the groups are three and number of items in each group are 4, therefore looking in table for Kruskal- Wallis given in appendix of any Statistics book, (k=3, sample size of 4,4,4) it can be seen that the critical value is 5.692 (ab = 0.05) level of significance.

Step 5: Make your conclusion about whether you can reject Ho by examining the p-value.

Since the critical value is more than the obtained value we accept the null hypothesis that all the three conditions A (Low Noise), B (Average Noise), and C (Loud Noise), do not differ from each other, therefore, in the said experiment there was no differences in the groups performance based on the noise level.

Check Your Progress II

1) What is the formula for Kruskal-Wallis test statistic?

9.4 CONSIDERATIONS FOR LARGE SAMPLE

When the number of sample increases, the table H is unable to give us with the critical values, like for example it gives critical values up to 8 samples when k=3, 4 when k=4, and 3 samples when k=5, therefore as the sample increases table H is not of use for the critical value. In such a case we resort to chi square table for getting our information on the critical value taking degrees of freedom (k – 1).

Exact tables of the Kruskal-Wallis distribution: Although an exact probability value can be computed for obtaining a configuration of ranks which is equivalent to or more extreme than the configuration observed in the data evaluated with the Kruskal-Wallis oneway Analysis of Variance by ranks, the chi-square distribution is generally employed to estimate the latter probability. As the values of k and N increase, the chi-square distribution provides a more accurate estimate of the exact Kruskal-Wallis distribution. Although most sources employ the chi-square approximation regardless of the values of k and N, some sources recommend that exact tables be employed under certain conditions. Beyer (1968), Daniel, and Siegel and Castellan (1988) provide exact Kruskal-Wallis probabilities for whenever k= 3 and the number of participants in any of the samples is five or less. Use of the chi-square distribution for small sample sizes will generally result in a slight decrease in the Power of a test (i.e., there is a higher likelihood of retaining a false null hypothesis). Thus, for small sample sizes the tabled critical chi- square value should, in actuality, are a little lower than the value of Table H.

Worked Example for a large sample: A state court administrator asked the 24 court coordinators in the state's three largest counties to rate their relative need for training in case flow management on a Likert scale (1 to 7).

1 = no training need

7 = critical training need

Training Need of Court Coordinators

District A	District B	District C
3	7	4
1	6	2
3	5	5
1	7	1
5	3	6
4	1	7
4	6	
2	4	
	4	
	5	

Step 1: Rank order the total groups' Likert scores from lowest to highest.

If tied scores are encountered, sum the tied positions and divide by the number of tied scores. Assign this rank to each of the tied scores.

Scores & Ranks Across the Three Counties

Ratings	Ranks	Ratings	Ranks
1	2.5	4	12
1	2.5	4	12
1	2.5	5	16.5
1	2.5	5	16.5
2	5.5	5	16.5

2	5.5	5	16.5
3	8	6	20
3	8	6	20
3	8	6	20
4	12	7	23
4	12	7	23
4	12	7	23

Calculating the ranks of tied scores

Example: Three court administrators rated their need for training as a 3. These threescores occupy the rank positions 7, 8, & 9. (7 + 8 + 9) / 3 = 8

Step 2: Sum the ranks for each group and square the sums

District A		**District B**		**District C**	
Rating	Rank	Rating	Rank	Rating	Rank
3	8	7	23	4	12
1	2.5	6	20	2	5.5
3	8	5	16.5	5	16.5
1	2.5	7	23	1	2.5
5	16.5	3	8	6	20
4	12	1	2.5	7	23
4	12	6	20		
2	5.5	4	12		
		4	12		
		5	16.5		
Σ R	67.0		153.5		79.5
$(\Sigma R)^2$	4489		23562.25		6320.25

Step3: CalculateH

$$H = [\ 12 / N\ (N+1)\ ]\ [\ \Sigma((\Sigma R)^2 / n)\ ] - 3(N + 1)$$

$$H = [\ 12 / 24\ (24+1)\ ][4489 / 8 + 23562.25 / 10 + 6320.25 / 6] - 3\ (24 + 1)$$

$$H = (0.02)\,(3970.725) - (75)$$

$$H = 4.42$$

$$df = (k - 1) = (3 - 1) = 2$$

Step 4: Interpretation

The critical chi-square table value of H for $\alpha = 0.05$, and df = 2, is 5.991

Since $4.42 < 5.991$, the null hypothesis is accepted. There is *no difference* in the training needs of the court coordinators in the three counties.

Check Your Progress III

1) Fill in the blanks

a) As the values of k and N increase, the …… provides a more accurate estimate of the exact Kruskal-Wallis distribution.

b) Use of the chi-square distribution for …… will generally result in a slight decrease in the power of the test

9.5 COMPARISON OF A NOVA AND KRUSKAL-WALLIS ANOVA TEST

The Kruskal- Wall is (KW) ANOVA is the non-parametric equivalent of a one-way ANOVA. As it does not assume normality, the KW ANOVA tests the null hypothesis of no difference between three or more group medians, against the alternative hypothesis that a significant difference exists between the medians. The KW ANOVA is basically an extension of the Wilcoxon-Mann-Whitney (WMW) 2 sample test, and so has the same assumptions: 1) the groups have the same spreads; and 2) the data distributions have the same shape.

ANOVA compares means of different population to indicate the similarity between the populations KW ANOVAcompares medians of these populations. ANOVA compares the data itself, KWANOVA, converts data into ranks and then does its computation, in this respect Kruskal- Wallis ANOVA is knows as Kruskal- Wallis Analysis of Variance by Ranks.

Let’s look at the Example to see how their calculations differ:

Three groups 1, 2, and 3, performed a task, we want to see whether they differ or not.

Group 1	Group 2	Group 3		Group1 (Ranks)	Group2 (Ranks)	Group3 (Ranks)
3	6	9		1	3.5	10
5	7	10		2	5.5	13
6	8	11		3.5	7.5	15.5
7	9	12		5.5	10	17
8	10	15		7.5	13	18
9	10			10	13	
	11				15.5	
38	61	57	**Total**	29.5	68	73.5

ANOVA

$n_1=6$ $n_2=7$ $n_3=5$

$\Sigma X_1= 38$ $\Sigma X_2=61$ $\Sigma X_3 =57$

$\Sigma(X_1)^2=264$ $\Sigma (X_2)^2=551$ $\Sigma (X_3)^2 =671$

$SS_{total} = (264 + 551 + 671) – [(38 + 61+ 57)^2 / 18] =134$

$SS_{Between\ Group} = (382/6) + (612/7) + (572/5) – [(38 + 61+ 57)2 / 18]$

$SS_{Within\ Group} = [264 - (382/6)] + [551- (612/7)] + [671- (572/5)] = 63.962$

Source of Variation	**SS**	**df**	**MS**	**F ratio**	**F critical Value**	**Test Decision**
Between Groups	70.038	2	35.019	8.213	3.68	Reject H_0
Within Groups	63.962	15	4.264			
Total	134.000	17				

Kruskal-Wallis H test:

$$H = [12 / 18(18+1)] [(29.52/6) + (682/7) + (73.52/5)] – [3 (18+1)]$$

$$H= 66.177 – 57 = 9.177$$

Chi Square for Degrees of freedom 2 (3 – 1) is 5.99, Therefore reject the Ho

In both the case, ANOVA or Kruskal-Wallis ANOVA, we will reject the Null Hypothesis,and state that the three groups differ.

F ratio as a function of H: The Fisher's F or F ratio or ANOVA one way variance is equivalent to H test or Kruskal-Wallis test or Kruskal-Wallis ANOVA, or ANOVA by rank order. This can also be seen in book by Iman and Conover (1981).

Where the rank transform statistics states:

$$F= [\{(k-1)/(N-k)\} \{((N-1)/H)-1\}]-1$$

If We see from the above mentioned example F was 8.213 and H was 9.177

$$F= [\{(3-1)/(18-3)\} \{((18-1)/9.177)-1\}]-1 \quad F= [(2/15) \{(17/9.177)-1\}]-1$$

$$F= [0.133 \times (1.852-1)] -1 = 0.1214-1$$

$$F= 8.231$$

Check Your Progress IV

1) Fill in the blanks:

a) The Kruskal- Wallis (KW)ANOVAis the non-parametric equivalent of a ……...

b) The …….. is basically an extension of the Wilcoxon-Mann-Whitney (WMW) 2 sample test, and so has the same assumptions: 1) the groups have the same spreads; and 2) the data distributions have the same shape.

c) ANOVA compares …….. to indicate the similarity between the populations KWANOVA compares medians of these populations.

9.6 LET US SUM UP

Kruskal-Wallis Analysis of Variance (KW ANOVA) is a non-parametric analogue of ANOVA one way variance for independent sample. Kruskal-Wallis ANOVA is used when there are more than 2 groups ($k > 2$). The assumption of KW ANOVA are that: a) Each sample has been randomly selected from the population it represents; b) The k samples are independent of one another; c) The dependent variable (which is subsequently ranked) is a continuous random variable. In truth, this assumption, which is common to many non-parametric tests, is often not adhered to, in that such tests are often employed with a dependent variable which represents a discrete random variable; and d) the underlying distributions from which the samples are derived are identical in shape.

The ANOVA or F test and KW ANOVA Or H test are equivalent to each other and can be appropriately used depending upon the type of population in question.

9.7 REFERENCES

Daniel, W. W. (1990) Applied Non-parametric Statistics, 2d ed. Boston: PWS- Kent.

Iman, R. L., and W. J. Conover (1981), Rank transformations as a bridge between parametric and non-parametric statistics, The American Statistician, 35, 124–129.

Johnson, Morrell, and Schick (1992), Two-Sample Non-parametric Estimation and Confidence Intervals Under Truncation, Biometrics, 48, 1043-1056.

Leach, C. (1979). Introduction to statistics: A non-parametric approach for the social sciences. Chichester: John Wiley & Sons

Lehman, E. L. (1975). Non-parametric statistical methods based on ranks. San Francisco: Holden-Day.

Siegel S. and Castellan N.J. (1988) Non-parametric Statistics for the Behavioral Sciences (2nd edition). New York: McGraw Hill.

Wampold BE & Drew CJ. (1990) Theory and application of statistics. New York: McGraw-Hill.

9.8 KEY WORD

Kruskal-Wallis One way Analysis of Variance (KW ANOVA): Kruskal-Wallis One way Analysis of Variance (KW ANOVA) is a Non-parametric Analogue of ANOVA one way variance for independent sample. Kruskal-Wallis ANOVA is used when there are more than 2 groups ($k > 2$).

9.9 ANSWERS TO CHECK YOUR PROGRESS

Check Your Progress I

1) List any one assumption of Kruskal-Wallis One Way ANOVA.

   One of the assumptions of Kruskal-Wallis One Way ANOVA is:

   Each sample has been randomly selected from the population it represents;

Check Your Progress II

1) What is the formula for Kruskal-Wallis test statistic?

$$H=[12/N(N+1)][\Sigma((\Sigma R)2\ /n)]–3(N+1)$$

Check Your Progress III

1) Fill in the blanks:

a) As the values of k and N increase, the chi-square distribution provides a

more accurate estimate of the exact Kruskal-Wallis distribution.

b) Use of the chi-square distribution for small sample sizes will generally result in a slight decrease in the power of the test

Check Your Progress IV

1) Fill in the blanks:

b) The Kruskal-Wallis (KW)ANOVAis the non-parametric equivalent of a one-way ANOVA.

c) The KWANOVA is basically an extension of the Wilcoxon-Mann-Whitney (WMW) 2 sample test, and so has the same assumptions: 1) the groups have the same spreads; and 2) the data distributions have the same shape.

d) ANOVA compares means of different populations to indicate the similarity between the populations KWANOVA compares medians of these populations.

9.10 UNIT END QUESTIONS

1) Explain Kruskal–Wallis ANOVA test.

2) A firm wishes to compare four programs for training workers to perform a certain manual task. Twenty new employees are assigned to the training programs, with 5 in each program. At the end of the training period, a test is conducted to see how quickly trainees can perform the task. The number of times the task is performed per minute is recorded for each trainee, with the following results:

Observation	Program 1	Program 2	Program 3	Program 4
1	9	10	12	9
2	12	6	14	8
3	14	9	11	11
4	11	9	13	7
5	13	10	11	8

Calculate H, and report your results appropriately.

3) Compare between ANOVA and Kruskal-Wall is ANOVA Test

ignou
THE PEOPLE'S
UNIVERSITY