UNIT 16 ANALYSIS OF QUANTITATIVE DATA: INFERENTIAL STATISTICS BASED ON NON-PARAMETRIC TESTS

Structure

16.1 INTRODUCTION

In the previous unit you learnt about the use of parametric tests in making inferences about the means computed from large and small samples. The use of Z and T tests were also explained to you for testing the significance of the difference between the means of two large and small samples. The application and use of analysis of variance and co-variance for testing the differnce between the means of three or more samples were also discussed with the help of examples. The signifance of the Pearson's co-efficient of correlation using Fisher's Z conversion were also explained alongwith the the use of Z test for testing the signifance of the difference between Pearson's coefficients of correlation computed from two samples.

In the use of parametric tests for making statistical inferences, we need to take into account certain assumptions about the nature of the population distribution, and also the type of the measurement scale used ot quantify the data. In this unit you will learn about another category of tests which do not make stringent assumptions about the nature of the population distribution. This category of test is called distribution free or non-parametric tests. The use and application of various of non-parametric tests involving unrelated and related samples will be explained in this unit. These would include chi-square test, median test, Man-Whitney U test, sign test and Wilcoxon-matched pairs signed-ranks test.

16.2 OBJECTIVES

After studying this unit, you will be able to:

- explain the nature of non-parametric tests;
- state the use of non-parametric tests;
- draw statistical inference pertaining to unrelated samples using the: (i) chi-square test: (ii) median test; and (iii) man-whitney U test;
- make statistical inferences pertaining to related samples using the: (i) sign test and (ii) wilcoxon matched-pairs signed-ranks test; and
- test statistical signifance of spearman's correlation co-efficient, phi-co-efficient and contingency co-efficient.

16.3 NON-PARAMETRIC TESTS

In the last unit you learnt that parametric tests are generally quite robust and are useful even when some of their mathematical assumptions are violated. However, these tests are used only with the data based upon ratio or interval measurements. In case of counted or ranked data, we make use of non-parametric tests. It is argued that non-parametric tests have greater merit because their validity is not based upon assumptions about the nature of the population distribution, assumptions that are so frequently ignored or violated by researchers using parametric tests. It may be noted that non-parametric tests are less precise and have less power than the parametric tests.

The use of non-parametric tests do not make numerous or stringent assumption about the nature of the population distribution and hence they are called distribution free tests.

Non-parametric are used when:

1. The nature of the population from which samples are drawn is not known to be normal.
2. The variables are expressed in nominal form.
3. The data are measures which are ranked or expressed in numerical scores which have the strength of ranks.

16.4 STATISTICAL INFERENCE BASED ON NON-PARAMETRIC TESTS : UNRELATED SAMPLES

The most frequently non-parametric tests which are used in drawing statistical inferences in case of unrelated or independent samples are: (1) chi square test; (ii) median test; and (iii) man-whitney test. The use and application of these tests are discussed below:

16.4.1 The Chi Square ($\aleph^2$) Test

The chi square test is applied only to discrete data. The data that are counted rather than measured. It is a test of independence and is used to estimate the likelihood that some factor other chance accounts for the observed relationship. The chi square ($\aleph^2$) is not a measure of the degree of relationship between the variables under study, the chi square test merely evaluates the probability that the observed relationship results from chance. The basic assumption, as in case of other statistical significance, is that the sample observations have been randomly selected.

The formula for chi-square ($\aleph^2$) is:

$$\aleph^2 = \sum\left[\frac{(fo - fe)^2}{fe}\right]$$

In which

fo = frequency of occurrence of observed or experimentally determined facts.

fe = expected frequency of occurrence.

To evaluate the significance of chi square, we use the Table E of chi square presented in the Appendix with computed value of chi square and the appropriate number of degrees of freedom (df). The number of df = (r–1) (c–1) r is number of rows and c is the number of columns, in which the data are tabulated.

To illustrate the use of formula, let us consider the following data based on the judgements of 390 judges.The judgements have been classified into five categories taken to represent a continuum of opinion:

Categories

	I	II	III	IV	V	**Total**
Judgements	55	63	82	93	57	350

The hypothesis to be tested is 'equal probability hypothesis 'i.e. whether the judgments expressed in five categories differ significantly or not. For this we have to compute the distribution of answers to be expected on the equality or null hypothesis. Since the total judgements are 350 and the number of categories is 5, the expected judgements in each category would be 350/5 = 70. The data in respect of observed (fo) and expected frequencies (fe) alongwith the values (fo–fe), (fo–fe) 2 etc. can be arranged as under

	Categories					**Total**
	I	II	III	IV	V	
Observed Judgements (fo)	55	63	82	93	57	350
Expected Judgements (fe)	70	70	70	70	70	350
(fo–fe)	15	7	12	23	13	
$(fo–fe)^2$	225	49	144	529	169	
$\frac{(fo - fe)^2}{fe}$	3.21	0.70	2.06	7.56	2.41	

$$\aleph^2 = \sum\left(\frac{(fo - fe)^2}{fe}\right)$$

$$= 3.21 + 0.70 + 2.06 + 7.56 + 2.41$$

$$= 15.94$$

The degrees of freedom in the table may be calculated from the formula df = (r–1) (c–1) to be (5–1) (2–1) or 4.

Using chi square Table E in the Appendix, we find in row df = 4, $\aleph^2$ of 9.488 in the column headed .05. Since the obtained value of $\aleph^2$ = 15.94 is greater than the table value of 9.488, we reject the equal judgement hypothesis and conclude that judgements in terms of various categories differ significantly.

Suppose instead of the 'hypothesis of equality', we may wish to test the data expressed in various judgement categories against the hypothesis of a normal distribution. In that case our hypothesis may assert that the judgement frequencies which we have observed really follow the normal distribution instead of being equally probably. Using the data of the above example, we have to find out how many of the 350 (total of the categories of judgements) may be expected to fall in each categories on the hypothesis of a normal distribution. These are found by first dividing the base line of a normal curve (taken to extend over 6 σ) into 5 equal segments each of 1.20 σ each. From the normal table (Table A) of the Appendix, the proportion of the normal distribution to be found in each of these segments would be as follows:

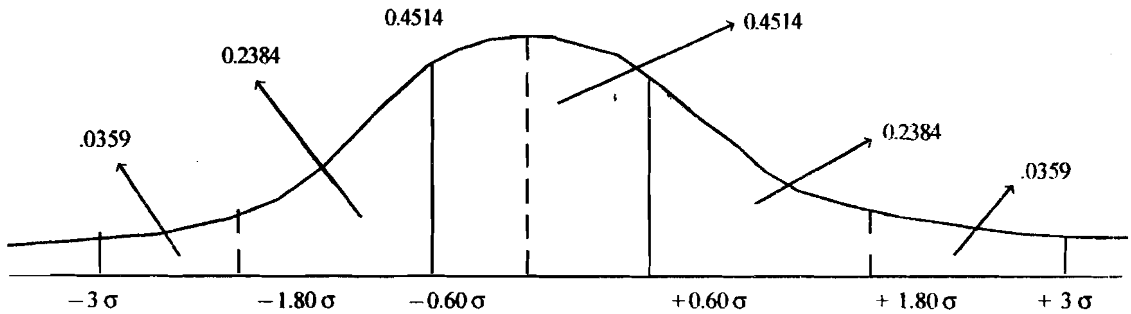

Fig.16.1: Normal Distribution of judgment frequencies.

Between + 3.00 σ and + 1.80 σ = .0359

+ 1.80 σ and + 0.60 σ = .2384

– 0.60 σ and – 0.60 σ = .4514

– 0.60 σ and – 1.80 σ = .2384

– 1.80 σ and – 3.00 σ = .0359

These proportions of 350 have been calculated as 12.56, 83, 44, 157.99, 83.44 and 12.56 and are entered in the row fe in following table:

	Categories					Total
	I	II	III	IV	V	
(fo)	55	63	82	93	57	350
(fe)	12.56	83.44	158.00	83.44	12.56	350
(fo–fe)	42.44	20.44	76	9.56	44.44	
$(fo-fe)^2$	1801.15	417.79	5776	91.39	1974.91	
$\frac{(fo-fe)^2}{fe}$	143.40	5.01	36.56	1.10	157.24	

$$\aleph^2 = \sum \left[\frac{(fo - fe)^2}{fe} \right]$$

$$= 143.40 + 5.01 + 36.56 + 1.10 + 157.24$$

$$= 343.31$$

The value of $\aleph^2$ in the Table E in the Appendix is 9.488 for df = 4 in the column headed by .05, which is less than the computed $\aleph^2$ value of 343.31. The difference between observed and expected values is so great that the hypothesis of normal distribution of judgement categories must be rejected.

Let us use chi square test to the data, which represent the number of boys and the number of girls who chose each of the three possible answers to an item on a personality inventory, to test whether the item differentiates significantly between boys and girls.

	Yes	**No**	**Undecided**	**Total**
Boys	14	66	10	90
Girls	27	66	7	100
Total	41	132	17	190

For each of the observed frequency in the table, let us compute the expected frequency in the following way:

Row 1 (Boys): $\dfrac{41 \times 90}{190} = 19.42;\ \dfrac{132 \times 90}{190} = 62.53;\ \dfrac{17 \times 90}{190} = 8.05$

Row 2 (Girls): $\dfrac{41 \times 100}{190} = 21.58;\ \dfrac{132 \times 100}{190} = 62.47;\ \dfrac{17 \times 100}{190} = 8.95$

The data in respect of observed and expected frequencies are arranged in the following table. The values in parentheses within the different cells are expected frequencies.

Responses

	Yes	**No**	**Undecided**	**Total**
Boys	14 (19.42)	66 (62.53)	10 (8.05)	90
Girls	27 (21.58)	66 (69.47)	7 (8.95)	100
Total	41	132	17	190

The hypothesis to be tested is the null hypothesis namely, that the item does not differentiate between the groups of boys and girls.

Using the formula of $\aleph^2$

$$\aleph^2 = \sum \left[\frac{(fo - fe)^2}{fe} \right]$$

$$= \frac{(14-19.42)^2}{19.42} + \frac{(66-62.53)^2}{62.53} + \frac{(10-8.05)^2}{8.05}$$

$$= \frac{(27-21.58)^2}{21.58} + \frac{(66-69.47)^2}{69.47} + \frac{(17-8.95)^2}{8.95}$$

$$= 1.51 + 0.19 + 0.47 + 1.36 + 0.17 + 7.24$$

$$= 10.94$$

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

The $\aleph^2$ critical values for 2 df as given in the Table E are 5.991 and 9.210 respectively for .05 and .01 levels of signifance and the obtained value 10.94 of $\aleph^2$ is higher than these values. This indicates that the item of the personality inventory differentiate between boys and girls and the null hypothesis is rejected.

Check Your Progress

Notes : a) Space is given below for your answer.

b) Compare your answer with the one given at the end of this unit.

1. The following judgements were classified into six categories taken to represent a continuum of opinion:

	Categories						
	I	II	III	IV	V	VI	**Total**
Judgements	48	61	82	91	57	45	384

Test the given distribution versus normal distribution hypothesis.

....

....

2. The following table represents the number of boys and the number of girls who choose each of the possible answers to an item in an attitude scale.

	Strongly Approve	Approve	Indifferent	Disapprove	Strongly Disapprove	**Total**
Boys	25	30	10	25	10	100
Girls	10	15	5	15	15	60
Total	35	45	15	40	25	160

Do these date indicate a significant sex difference in attitude towards this question?

....

....

16.4.2 The Median Test

The median test is used for testing whether two independent samples differ in central tendencies. It gives information as to whether it is likely that two independent samples have been drawn from populations with the same median. It is particularly useful when even the measurements for the two samples are expressed in an ordinal scale.

In using the median test, we first calculate the combined median for all measures (scores) in both samples. Then both sets of scores at the combined median are dichotomized and the data are set in a 2×2 table presented below:

Table for Use of Median Test

	Group I	Group II	Total
No. of measures (scores) above combined Median	A	B	A + B
No. of measures (scores) below combined Median	C	D	C + D
Total	A + C	B + D	

Under the null hypothesis, we would expect about half of each group's (scores) to be above the combined median and about half to be below , that is, we would expect frequencies A and C to be about equal, and frequencies B and D to be about equal. In order to test this hypothesis, we calculate $\aleph^2$using the following formula:

$$\aleph^2 = \frac{N\left[\,|AD - BC| - \frac{N}{2}\right]^2}{(A+B)(C+D)(A+C)(B+D)}$$

Let us illustrate the use of this formula with the help of the following example:

Twenty male and fifteen female teacher educators of a teacher training institute were asked to express their attitude towards teacher education programmes offered through distance mode at the B. Ed. Level. Both the groups were administrated an attitude scale and common median attitude score was computed. The number of cases from both groups falling above and below the median score is shown in the following table:

Distribution of Male and Female Teacher Educators Below and Above the Common Median Attitude Score

	Below Median	Above Median	Total
Female Teachers Educators	9	6	15
Male Teachers Educators	6	14	20
Total	15	20	35

Using the formula for $\aleph^2$

$$\aleph^2 = \frac{35\left[\,|(9)(14) - (6)(6)| - \frac{35}{2}\right]^2}{(15)(20)(15)(20)}$$

$$= \frac{35(|126 - 36| - 17.5)^2}{(300)(300)}$$

$$= \frac{35(|126 - 36| - 17.5)^2}{90000}$$

$$= \frac{35(90 - 17.5)^2}{90000}$$

$$= \frac{35(72.5)^2}{90000}$$

$$= 2.044$$

The critical χ^2 value for (2–1) (2–1) df or 1 df is 3.84 for. .05 level. Since the obtained χ^2 value of 2.044 is less than 3.84, the null hypothesis is retained and we may conclude that there is no difference in the attitude of male and female teacher educators towards teacher education programmes at the B.Ed. Level.

16.4.3 The Mann-Whitney U Test

The Mann-Whitney U test is more useful than the Median test. It is a most useful alternative to the parametric t test when the parametric assumptions cannot be met and when the measurements are expressed in ordinal scale values.

Suppose N_1 is the number of individuals in one of the two independent groups and N_2 the number of the individuals in the other. In using Mann-Whitney U test, we first combine the measures or scores from both groups, and rank these in order of increasing size. In this ranking, we have to consider the algebraic sign, that is, the lowest ranks are assigned to the largest negative numbers, if any. The ranks of each sample group are then summed individually and represented as ΣR_1 and ΣR_2.

There are two Us: U1 and U2 which are calculated using the following formula:

$$U_1 = N_1 N_2 + \frac{N_1 (N_1 + 1)}{2} - \sum R_1$$

$$U_2 = N_1 N_2 + \frac{N_2 (N_1 + 1)}{2} - \sum R_2$$

in which:

N_1 = number in one group

N_2 = number in second group

$\sum R_1$ = sum of ranks in one group

$\sum R_2$ = sum of ranks in second group

The two U's are related by the equation:

$$U_1 = N_1 N_2 - U_2$$

Thus only one U needs to be calculated, for the other can be easily determined by this education.

The Z value of U can be computed by the formula:

$$Z = \frac{U - \frac{N_1 N_2}{2}}{\sqrt{\frac{(N_1)(N_2)(N_1 + N_2 + 1)}{12}}}$$

It does not matter which U (the larger or smaller) is used in the computation of Z. The sign of Z will depend on which U is used, but the numerical value will be identical.

The following example used by Koul (1997) illustrated the application of Mann-Whitney U test in which a researcher wished to evaluate the effectiveness of micro-teaching and simulation in developing certain teachings skills among student-teachers of a college into two groups A and B by randomly assigning 20 to each of the groups. Group A was trained in various skills of teaching through micro- teaching and the group B was trained through simulation technique. After a period of two months training, the student-teachers were rated in the teaching skills by supervisors. The rating scores of the student teachers are given in table 16.1:

Table 16.1: Scores of Student Teachers

Group	Rank	Group B	Rank
90	39	46	4
78	29	42	1
75	25.5	65	15
72	22	61	12
75	25.5	64	14
83	33.5	82	32
73	23	69	19
80	31	66	16
74	24	56	9
67	17	48	6
63	13	68	18
45	3	44	2
55	8	85	36
84	35	83	33.5
89	38	71	21
77	28	87	37
70	20	76	27
58	10	50	7
47	5	59	11
92	40	79	30
N_1=20	$\sum R_1$=469.50	N_2=20	$\sum R_2$=350.50

All rating measures are ranked from lowest to highest and the Mann-Whitney U test is used to test the null hypothesis at the .05 of significance using the formula of U_1 and U_2.

$$U_1 = (20)(20) + \frac{(20)(21)}{2} - 469.50$$

$$= 140.50$$

$$U_2 = (20)(20) + \frac{(20)(21)}{2} - 350.50$$

$$= 259.50$$

Using the equation $U_1 = N_1N_2 - U_2$, we check:

$$140.50 = 400 - 259.50$$

$$140.50 = 140.50$$

$$Z = \frac{140.50 - \frac{400}{2}}{\sqrt{\frac{(20)(20)(41)}{12}}}$$

$$= -1.61$$

The obtained Z value of −1.61 does not exceed the Z critical value of 1.96 at .05 level, the null hypothesis is accepted. It may be concluded that micro teaching approach and simulation technique are equally effective in developing certain teaching skills among student teachers.

Check Your Progress

Notes : a) Space is given below for your answer.

b) Compare your answer with the one given at the end of this unit.

3. In answering a questionnaire the following scores were achieved by 10 men and 20 women:

   Men: 22, 31, 38, 47, 48, 48, 49, 50, 52, 61

   Women: 22, 23, 25, 25, 13, 33, 34, 35, 37, 40, 41, 42, 43, 44, 44, 46, 48, 53, 54

   Do men and women differ significantly in their answers to this questionnaire? Apply Median test by taking the Median = 41.5.

   ....

   ....

   ....

   ....

4. The performance scores of the students taught by method A and method B are given below:

Method A	Method B
50	49
60	90
89	88
94	76
82	92
75	81
63	55
52	64
97	84
95	51
83	47
80	70
77	66
80	69
88	87
78	74
85	71
79	61
72	55
68	73

Apply Mann-Whitney U test and test the significance between the performance of the students taught by method A and method B.

....

....

....

16.5 STATISTICAL INFERENCE BASED ON NON-PARAMETRIC TEST: RELATED SAMPLES

Various tests are used in drawing statistical inferences in case of related samples. In this section we shall confine our discussion to the use of Sign Test and Wilcoxon Matched-Paris Signed-Ranks Test Only.

16.5.1 The Sign Test

The sign test is the simplest test of significance in the category of non-parametric tests. It makes use of plus and minus signs rather than quantitative measures as its data. It is particularly useful in situations in which quantitative measurement is impossible or inconvenient, but on the basis of superior or inferior performance it is possible to rank with respect to each other, the two members of each pair.

The sign test is used either in the case of single sample from which observations are obtained under two experimental conditions are obtained under two experimental conditions and the researcher wants to establish that two conditions are different or to the case of two equivalent samples in which the subjects are matched with respect to the relevant extraneous variables. The use of this test does not make any assumption about the form of the distribution of differences. The only assumption underlying this test is that the variable under investigation has a continuous distribution.

If the number of the individuals in the single sample or in each of the equivalent or related samples is less or equal to 25, we make use of the 'Table of Probabilities Associated with Values as Small as Observed Values of x (number of fewer signs) in the Binomial Test' (Siegal, 1956).

When the number of individuals in the group is larger than 25, the normal approximation to the binomial distribution is used. The significance of difference is tested by calculating the value of Z given by the formula:

$$Z = \frac{(X \pm .5) - \frac{1}{2}N}{\frac{1}{2}\sqrt{N}}$$

Where $(x + .5)$ is used when $x < 1/2$ N and $(x–5)$ is used when $x > 1/2$ N. The significance of the obtained value of Z is determined by reference to the normal Table A.

16.5.2 The Wilcoxon Matched – Pairs Signed – Ranks Test

The Wilcoxon matched-pairs signed – ranks test is more powerful than the sign test because it tests not only direction but also magnitude of differences within pairs of matched groups. This test, like the sign test, deals with dependent groups made up of matched pairs of individuals and is not applicable to independent groups. The null hypothesis would assume that the direction and magnitude of pair difference would be about the same.

The application of wilcoxon matched-pairs signed-ranks test involves the following steps:

1. Let d_i be the difference scores for any matched pair, representing the difference between a pair's scores under two treatments A and B. There would be one d_i for each pair of scores.

2. Delete all such pairs for which $d_i = 0$

3. Rank all the d_i's without regard to sign, giving rank 1 to the smallest difference d_i rank 2 the next smallest, etc. If two or more d_i's are of the same size, assign the same rank to such tied cases. The rank assigned would be average of the ranks which would have been assigned if the d_i's had differed slightly

from each other. For example, if three pairs yield d_i's of –1, –1 and +1, then each pair would be assigned the rank of 2, for $\frac{1+2+3}{2} = 2$, and next d_i on order would be assigned the rank of 4 because ranks 1, 2 and 3 have already been exhausted.

4. Indicate which ranks arose from negative d_i's and which ranks arose from positive d_i's by affixing to each rank the sign of difference.

5. Sum the ranks for the positive differences and sum the ranks for the negative differences. Under the null hypothesis we would expect the two sums to be equal. In other words, if the sum of the positive ranks equals the sum of the negative ranks, we would conclude that the treatments A and B are not different. But if the sum of the positive ranks is very much different from the sum of the negative ranks, we would infer that the treatment A differs from treatment B and thus we would reject the null hypothesis.

Let us illustrate the application of Wilcoxon test with the help of the following example used by Koul (1997). Suppose a group of 26 delinquent children were initially rated for their social adjustment by psychiatrist and sent to a juvenile jail. After a year they were rated again by a psychiatrist for social adjustment and then initial and final adjustment rating scores were compared. The rating data are presented in the following Table 16.2:

Table 16.2: Rating Scores of Delinquent Children

Post-Rating Score	Pre-Rating Score	d	Rank of d	Rank with less Frequent Sign
50	47	3	7.5	
54	52	2	5.5	
62	63	–1	–2.5	2.5
39	31	8	17.5	
56	52	4	10	
51	42	9	19	
58	51	7	16	
60	49	11	23	
48	42	6	14	
46	42	4	10	
42	40	2	5.5	
53	56	–3	–7.5	7.5
40	36	4	10	
51	50	1	2.5	
56	42	19	26	
60	48	12	24	
57	47	10	21	
43	48	–5	–12	12
45	37	8	17.5	
52	39	13	25	
61	60	1	2.5	
62	52	10	21	
48	42	6	14	
39	45	–6	–14	14
41	42	–1	–25	25
39	49	–10	–21	21

The null hypothesis that there was no difference in initial and final adjustment rating scores of the group was tested at .05 level of significance using the following formula:

$$Z = \frac{T - \frac{N(N+1)}{4}}{\sqrt{\frac{N(N+1)(2N+1)}{24}}}$$

in which

N = number of pairs ranked

T = sum of ranks of the smaller of the like-signed ranks

In the example, T, the smaller of sums of the like-signed ranks = 2.5 + 7.5 + 12.0 + 14.0 + 2.5 + 21 = 59.5.

$$Z = \frac{59.5 - \frac{(26)(26+1)}{4}}{\sqrt{\frac{(26)(26+1)(52+1)}{24}}}$$

$$= 2.95$$

Since the obtained Z value of 2.95 exceeds Z critical value of 1.96 at .05 level, the null hypothesis is rejected and we may conclude that the environment in juvenile jail has considerable improved the social adjustment of delinquent children.

Check Your Progress

Notes : a) Space is given below for your answer.

b) Compare your answer with the one given at the end of this unit.

5. List the uses of : (i) sign test and (ii) Wilcoxon matched-pairs signed-ranked test.

....

....

....

....

16.6 STATISTICAL INFERENCE REGARDING CORRELATION USING NON-PARAMETRIC DATA

The correlations which are computed from the measurements based on nominal (enumerative) and ordinal (ranking) data give rise to spearman's rho (ρ), phi (ϕ) contingence (C) coefficients. In this section, we will discuss the procedure for testing the statistical significance of these coefficients.

16.6.1 Significance of Spearman's Rho (ρ) Correlation Coefficient

There is no generally accepted formula for estimating the standard error of ρ which we need for testing its significance and determine its confidence limits. However, we can test the null hypothesis that the two variables under study are not associated in the population and that the observed value of rho (ρ) differs zero only by chance, in two ways.

1. When the size of sample N is from 4 to 30, the interpretation is best made by the aid of Table L given in the Appendix, in which are given ρ coefficients significant at .05 and .01 levels of confidence. This is a one tailed table, that is, the stated probabilities apply when the observed value of ρ is in the predicted direction, either positive or negative. For a one tailed test, if an observed value of ρ equals or exceeds the value of ρ shown in the Table L, the observed value is significant at the level indicated.

2. When N is 10 or large, the significance of an obtained ρ under null hypothesis may tested by the formula:

$$t = \rho \sqrt{\frac{N-2}{1-\rho^2}}$$

The interpretation of the obtained value of t is made with the use of Table C of the Appendix, using (N–2) degrees of freedom (df).

16.6.2 Significance of Phi (ϕ) Correlation Coefficient

The significance of ϕ coefficient is determined using the relationship of ϕ to $\aleph^2$ as:

$$\aleph^2 = N\phi^2$$

This formula helps us to test an obtained ϕ against the null hypothesis. First we convert ϕ to an equivalent $\aleph^2$ and then test the significance of $\aleph^2$ by referring to the chi-square Table E. If $\aleph^2$ comes out to be significant for a particular level of confidence, the corresponding value of s is also significant.

16.6.3 Significance of Contingency Coefficient (C)

The significane of contingency coefficient C is also determined through the relationship which it bears to $\aleph^2$ using the formula:

$$C = \sqrt{\frac{\aleph^2}{N+\aleph^2}}$$

This formula helps us to test the significance of the obtained value of C coefficient against the null hypothesis by first converting C to $\aleph^2$. The interpretation of an obtained chi-square is made with the use of chi-square Table E. If chi-square is significant at a particular level of confidence, C is also significant.

Check Your Progress

Notes : a) Space is given below for your answer.

b) Compare your answer with the one given at the end of this unit.

6. Test the significance of rho (ρ) 0.76 for N = 2

....

....

7. The coefficient of contingency between father's eye colour and son's eye colour computed on the basis of 4×4 contingency table came out to be 0.46. Test its significance at .05 level.

....

....

16.7 LET US SUM UP

1. The use of non-parametric tests do not make stringent assumptions about the nature of the population distribution. Non parametric tests are distribution free tests.
2. Non-parametric tests are used when: (i) the nature of the population, from which samples are drawn, is not known to be normal; (ii) the variables are expressed in nominal scale of measurement; and (iii) the data are measures which are ranked or expressed in numerical scores which have the strength of ranks.
3. Chi-square test, median and Mann-Whitney U test are most frequently non-parametric tests of significance which we use in case of unrelated or independent samples.
4. In case of related or dependent samples, we make use of sign test and Wilcoxon-matched-pairs signed-ranks test.
5. The procedure for testing the significance of rho (ρ), phi (ϕ) and contingency (C) coefficients of correlations were also discussed.

16.8 UNIT-END ACTIVITIES

1. Discuss the uses of non-parametric tests.
2. Describe the uses of chi square test, median test and Man-Whitney U test.
3. Illustrate the use of Wilcoxon matched-pairs signed-ranks test with the help of an example.

16.9 SUGGESTED READINGS

Garrett, H.E. (1962): *Statistics in Psychology and Education*. Bombay: Allied Pacific Pvt. Ltd.

Guilford, J.P. (1965): *Fundamental Statistics in Psychology and Education*. New York: McGraw Hill Book Company.

Koul, Lokesh (1997): *Methodology of Educational Research*. New Delhi; Vikas Publishing House Pvt. Ltd. (Third Revised Edition).

Siegal, S. (1956): *Non-Parametric Statistics for Behavioural Sciences*. Tokyo: McGraw Hill Hoga Kusna Ltd.

16.10 ANSWERS TO CHECK YOUR PROGRESS

1. χ^2 = 346, the deviation from the normal distribution is significant.

2. χ^2 = 7.03, no signifance sex difference in attitude towards the question.

3. No, $\chi^2 = 1.35$

4. No, $Z = -1.61$

5. i) Sign test is particularly useful in the situations is which quantitative measurements is impossible or impracticable, on the basis of superior or inferior performance. It is applicable either to the case of single sample from which observations are obtained under two experimental conditions and one wished to establish that two conditions are different or to the case of the equivalent samples in which the subjects are matched with respect to the relevant extraneous variables.

   ii) Wilcoxon test is more powerful than the sign test because it tests not only direction but also magnitude of differences within pairs of matched groups.

6. Significant at .01 level.

7. Not significant at .05 level.