Calutta University
Semester 4

CC-9: STATISTICAL METHODS FOR PSYCHOLOGICAL RESEARCH-II

Unit 4: (12 hours): Hypothesis Testing for Categorical Variables and Inference about Frequencies

The Chi-Square as a Measure of Discrepancy between Expected and Observed Frequencies; Logic of the Chi-Square Test; Assumptions of Chi-Square; Interpretation of the Outcome of a Chi-Square Test

Testing Experimental Hypotheses:

The hypothesis proposed in a psychological experiment may be taking the form of general theory or specific inquiry. A specific hypothesis is ordinarily to be preferred to a general proposal, as the more definite and exact the query the greater the likelihood of a conclusive answer. Here, the significance of an obtained difference was tested against a null hypothesis.

The Hypothesis of "Chance":

Nature of the Null Hypothesis:

The null hypothesis is not confined to zero differences or to the differences between statistics. Other forms of this hypothesis assert the results found in an experiment do not differ significantly from results to be expected on a probability basis or stipulated in terms of some theory. It is ordinarily more useful than other hypotheses because it is exact. Hypotheses other than null can, to be sure, be stated exactly: for example, assert that a group which has received a special training will be 5 points on the average ahead of an untrained (control) group. But it is difficult to set up such precise expectations in many experiments. For this reason, it is usually better to test a null hypothesis, rather than some other.

Basic Concepts Concerning Testing of Hypotheses:

(a) *Null hypothesis and alternative hypothesis:* In the context of statistical analysis, the concepts of null hypothesis and alternative hypothesis are very common. If one wants to compare method *A* with method *B* about its superiority and if one proceeds on the assumption that both methods are equally good, then this assumption is termed as the null hypothesis. As against this, one may think that the method *A* is superior or the method *B* is inferior, one can then stating what is termed as alternative hypothesis. The null hypothesis is generally symbolized as H_0 and the alternative hypothesis as H_a. Suppose someone wants to test the hypothesis that the population mean (μ) is equal to the hypothesized mean (μ_{H0}) = 100.Then he or she would say that the null hypothesis is that the population mean is equal to the hypothesized mean 100 and symbolically it can be expressed as:

$$H_0: \mu = \mu_{H0} = 100$$

If the results of the sample do not support this null hypothesis, one should conclude that something else is true. Whatever he or she concludes rejecting the null hypothesis is known as alternative hypothesis. In other words, the set of alternatives to the null hypothesis is referred to as the alternative hypothesis. If one can accept H_0, then he or she can reject H_a and if one reject H_0, then he or she accept H_a. For H_0: $\mu = \mu_{H0} = 100$, one can consider three possible alternative hypotheses as follows:

Alternative hypothesis	To be read as follows
$H_a: \mu \neq \mu_{H0}$	(The alternative hypothesis is that the population mean is not equal to 100 i.e., it may be more or less than 100)
$H_a: \mu > \mu_{H0}$	(The alternative hypothesis is that the population mean is greater than 100)
$H_a: \mu < \mu_{H0}$	(The alternative hypothesis is that the population mean is less than 100)

Kothari, C.R. (1990). Research Methodology Methods and Techniques. Wishwa Prakashan. Calcutta, Second edition

The null hypothesis and the alternative hypothesis are chosen before the sample is drawn (the researcher must avoid the error of deriving hypotheses from the data that he collects and then testing the hypotheses from the same data). In the choice of null hypothesis, the following considerations are usually kept in view:

(a) Alternative hypothesis is usually the one which one wishes to prove and the null hypothesis is the one which one wishes to disprove. Thus, a null hypothesis represents the hypothesis somebody is trying to reject, and alternative hypothesis represents all other possibilities.

(b) If the rejection of a certain hypothesis when it is actually true involves great risk, it is taken as null hypothesis because then the probability of rejecting it when it is true is α (the level of significance) which is chosen very small.

(c) Null hypothesis should always be specific hypothesis i.e.; it should not state about or approximately a certain value.

Generally, in hypothesis testing it is better to proceed on the basis of null hypothesis, keeping the alternative hypothesis in view. Why so? The answer is that on the assumption that null hypothesis is true, one can assign the probabilities to different possible sample results, but this cannot be done if one proceeds with the alternative hypothesis. Hence the use of null hypothesis (at times also known as statistical hypothesis) is quite frequent.

(b) *The level of significance:* This is a very important concept in the context of hypothesis testing. It is always some percentage (usually 5%) which should be chosen with great care, thought and reason. In case someone considers the significance level at 5 per cent, then this implies that H_0 will be rejected when the sampling result (i.e., observed evidence) has a less than 0.05 probability of occurring if $H_0$is true. In other words, the 5 per cent level of significance means that researcher is willing to take as much as a 5 per cent risk of rejecting the null hypothesis when it (H_0) happens to be true. Thus, the significance level is the maximum value of the probability of rejecting H_0 when it is true and is usually determined in advance before testing the hypothesis.

(c) *Decision rule or test of hypothesis:* Given a hypothesis H_0 and an alternative hypothesis H_a, one can develop a rule which is known as decision rule according to which H_0 can be accepted (i.e., reject H_a) or rejected H_0 (i.e., accept H_a). For instance, if H_0 is that a certain lot is good (there are very few defective items in it) against H_a that the lot is not good (there are too many defective items in it), then one must decide the number of items to be tested and the criterion for accepting or rejecting the hypothesis. One might test 10 items in the lot and plan his or her decision saying that if there are none or only 1 defective item among the 10, he or she will accept H_0 otherwise will reject H_0 (or accept H_a). This sort of basis is known as decision rule.

(d) *Type I and Type II errors:* In the context of testing of hypotheses, there are basically two types of errors one can make. One may reject H_0 when H_0 is true and may accept H_0 when in fact H_0 is not true. The former is known as Type I error and the latter as Type II error. In other words, Type I error means rejection of hypothesis which should have been accepted and Type II error means accepting the hypothesis which should have been rejected. Type I error is denoted by α (alpha) known as α error, also called the level of significance of test; and Type II error is denoted by β (beta) known as β error. In a tabular form the said two errors can be presented as follows:

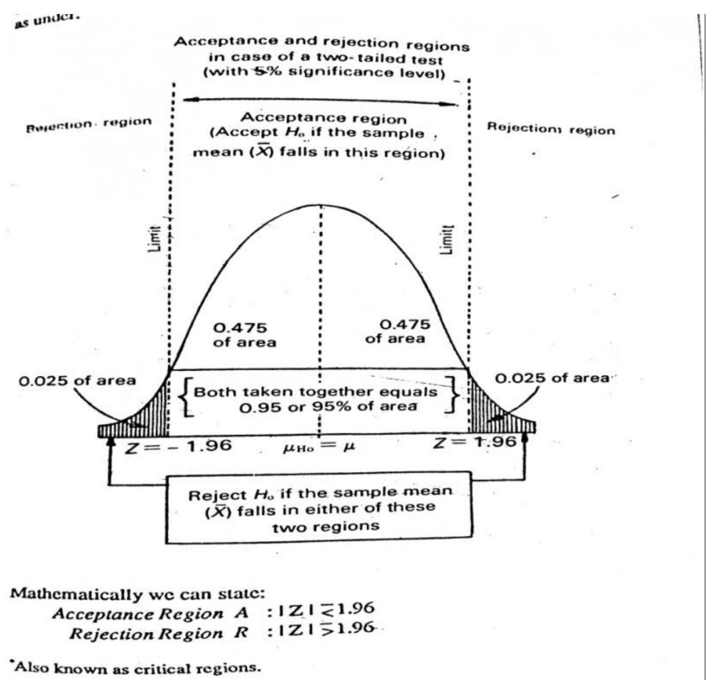
	Decision	
	Accept H_o	Reject H_o
H_o (true)	Correct Decision	Type I error (αerror)
H_o (false)	Type II error (β error)	Correct Decision

Kothari, C.R. (1990). Research Methodology Methods and Techniques. Wishwa Prakashan. Calcutta, Second edition

The probability of Type I error is usually determined in advance and is understood as the level of significance of testing the hypothesis. If type I error is fixed at 5 per cent, it means that there are about 5 chances in 100 that someone will reject H_0 when H_0 is true. One can control Type I error just by fixing it at a lower level. For instance, if one fixes it at 1 percent, he or she will say that the maximum probability of committing Type I error would only be 0.01.

But with a fixed sample size, *n*, when someone is trying to reduce Type I error, the probability of committing Type II error increases. Both types of errors cannot be reduced simultaneously. There is a trade-off between two types of errors which means that the probability of making one type of error can only be reduced if one is willing to increase the probability of making the other type of error. To deal with this trade-off in business situations, decision-makers decide the appropriate level of Type I error by examining the costs or penalties attached to both types of errors. If Type I error involves the time and trouble of reworking a batch of chemicals that should have been accepted, whereas Type II error means taking a chance that an entire group of users of this chemical compound will be poisoned, then in such a situation one should prefer a Type I error to a Type II error. As a result, one must set very high level for Type I error in one's testing technique of a given hypothesis. Hence, in the testing of hypothesis, one must make all possible effort to strike an adequate balance between Type I and Type II errors.

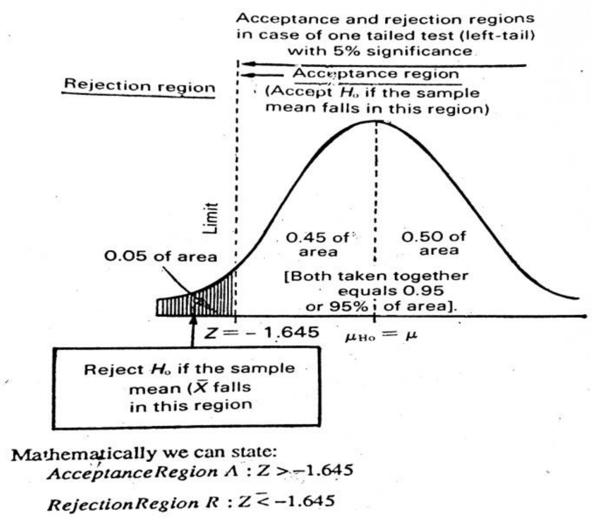
(e) *Two-tailed and One-tailed tests:* In the context of hypothesis testing, these two terms are quite important and must be clearly understood. A two-tailed test rejects the null hypothesis if, say, the sample mean is significantly higher or lower than the hypothesized value of the mean of the population. Such a test is appropriate when the null hypothesis is some specified value and the alternative hypothesis is a value not equal to the specified value of the null hypothesis. Symbolically, the two tailed test is appropriate when someone is having $H_0: \mu = \mu_{H0}$ $Ha: \mu \neq \mu_{H0}$ which may mean $\mu > \mu_{H0}$ or $\mu < \mu_{H0}$. Thus, in a two-tailed test, there are two rejection regions, one on each tail of the curve which can be illustrated as under:



Richard I. Levin, *Statistics for Management,* p. 247–248

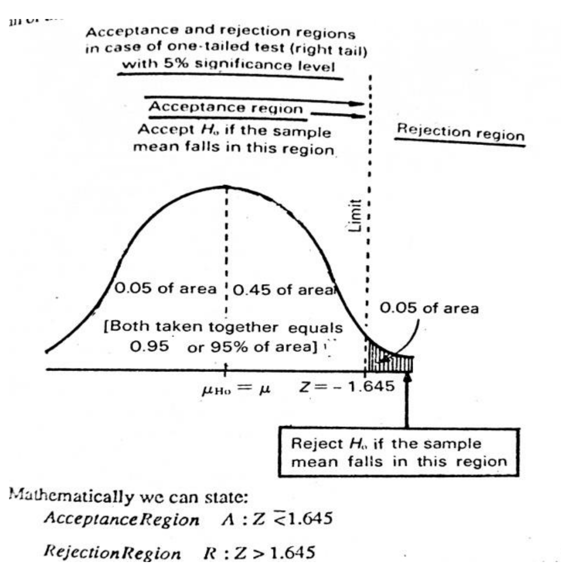
Kothari, C.R. (1990). Research Methodology Methods and Techniques. Wishwa Prakashan. Calcutta, Second edition

If the significance level is 5 per cent and the two-tailed test is to be applied, the probability of the rejection area will be 0.05 (equally splitted on both tails of the curve as 0.025) and that of the acceptance region will be 0.95 as shown in the above curve. If one considers$\mu = 100$and if the sample mean deviates significantly from 100 in either direction, then one should reject the null hypothesis; but if the sample mean does not deviate significantly from μ , in that case one should accept the null hypothesis.
But there are situations when only one-tailed test is considered appropriate. A *one-tailed test* would be used when it is required to test, say, whether the population mean is either lower than or higher than some hypothesized value. For instance, if H_0: $\mu = \mu_{H0}$ and Ha: $\mu < \mu_{H0}$, then one is interested in what is known as left-tailed test (wherein there is one rejection region only on the left tail) which can be illustrated as below:



Kothari, C.R. (1990). Research Methodology Methods and Techniques. Wishwa Prakashan. Calcutta, Second edition

If our $\mu = 100$ and if the sample mean deviates significantly from 100 in the lower direction, one should reject H_0, otherwise should accept H_0 at a certain level of significance. If the significance level in the given case is kept at 5%, then the rejection region will be equal to 0.05 of area in the left tail as has been shown in the above curve. In case H_0: $\mu = \mu_{H0}$ and H_a: $\mu > \mu_{H0}$, the main interest is in what is known as one-tailed test (right tail) and the rejection region will be on the right tail of the curve as shown below:



Kothari, C.R. (1990). Research Methodology Methods and Techniques. Wishwa Prakashan. Calcutta, Second edition

If our $\mu = 100$ and if the sample mean deviates significantly from 100 in the upward direction, one should reject H_0, otherwise one should accept the same. If in the given case the significance level is kept at 5%, then the rejection region will be equal to 0.05 of area in the right-tail as has been shown in the above curve.
It should always be remembered that accepting H_0on the basis of sample information does not constitute the proof that H_0 is true. It is only to say that there is no statistical evidence to reject it, but we are certainly not saying that H_0 is true (although we behave as if H_0 is true).

Hypothesis Testing for Comparing a Variance to some Hypothesized Population Variance:

The test is used for comparing a sample variance to some theoretical or hypothesized variance of population is different than z-test or the t-test. The test one can use for this purpose is known as chi-square test and the test statistic symbolized as χ^2, known as the chi-square value, is worked out. The chi-square value to test the null hypothesis viz., $H_0: \sigma 2 = \sigma 2$ worked out as under:

$$\chi^2 = \frac{\sigma_s^2}{\sigma_p^2}(n-1)$$

where σ_s^2 = variance of the sample

σ_p^2 = variance of the population

$(n-1)$ = degree of freedom, n being the number of items in the sample.

Then by comparing the calculated value of $\chi 2$ with its table value for $(n - 1)$ degrees of freedom at a given level of significance, one can either accept H_o or reject it. If the calculated value of $\chi 2$ is equal to or less than the table value, the null hypothesis is accepted; otherwise the null hypothesis is rejected. This test is based on chi-square distribution which is not symmetrical and all the values happen to be positive; one must simply know the degrees of freedom for using such a distribution.

(Reference: Kothari, C.R. (1990). Research Methodology Methods and Techniques. Wishwa Prakashan. Calcutta, Second edition)

Difference between Categorical, Ordinal and Numerical Variables:

A categorical variable (sometimes called a nominal variable) is one that has two or more categories, but there is no intrinsic ordering to the categories. For example, gender is a categorical variable having two categories (male and female) and there is no intrinsic ordering to the categories. Hair color is also a categorical variable having a number of categories (blonde, brown, brunette, red, etc.) and again, there is no agreed way to order these from highest to lowest. A purely categorical variable is one that simply allows anyone to assign categories but one cannot clearly order the categories.

An ordinal variable is similar to a categorical variable. The difference between the two is that there is a clear ordering of the categories. Now consider a variable like educational experience (with values such as elementary school graduate, high school graduate, some college and college graduate). These also can be ordered as elementary school, high school, some college, and college graduate. Even though one can order these from lowest to highest, the spacing between the values may not be the same across the levels of the variables. Say one can assign scores 1, 2, 3 and 4 to these four levels of educational experience and then compare the difference in education between categories one and two with the difference in educational experience between categories two and three, or the difference between categories three and four. The difference between categories one and two (elementary and high school) is probably much bigger than the difference between categories two and three (high school and some college).

Numerical variable is similar to an ordinal variable, except that the intervals between the values of the numerical variable are equally spaced. For example, suppose you have a variable such as annual income that is measured in dollars, and we have three people who make $10,000, $15,000 and $20,000. The second person makes $5,000 more than the first person and $5,000 less than the third person, and the size of these intervals is the same.

The chi-square (χ^2), test is used to determine whether an association (or relationship) between two categorical variables in a sample is likely to reflect a real association between these two variables in the population. The sample data is used to calculate a single number (or test statistic), the size of which reflects the probability (p-value) that the observed association between the two variables has occurred by chance, i.e. due to sampling error.

Concept of Nonparametric Statistics:

In statistical test two kinds of assertions are involved, viz., an assertion directly related to the purpose of investigation and other assertion to make a probability statement. The former is an assertion to be tested and is technically called a hypothesis, whereas other assertion is called a model, it is known as distribution-free or non-parametric test. Under nonparametric or distribution free test, it is difficult to assume that a particular distribution is applicable or that a certain value is attached to a parameter of the population. Nonparametric statistics require few assumptions, no estimate of parameter in their computation, and no normal distribution of the variables in the population.

Non-parametric (distribution-free) statistics (i) require very few assumptions; (ii) do not require normal distributions of the variables in the population. (iii) do not use any precomputed statistic as an estimate of parameter in the computation, (iv) can be used for very small samples and (v) are computed by much simpler methods; but (vi) their powers are lower than their parametric counterparts so long as the assumptions for the latter are fulfilled.

Differences between Parametric and Nonparametric Statistics:

Parametric Statistics	**Nonparametric Statistics**
1. It serves as estimates of corresponding parameters. Their computations require the use of precomputed statistics as estimates of parameter	1. It requires very few assumptions. Do not use any precomputed statistics as an estimate of parameter in the computation
2. They are interpreted with reference to specific population distributions of the variables such as the normal and t distributions	2. Do not require normal distributions of the variables in the population
3. It cannot be applied for too small samples	3. It can be used for very small samples
4. They cannot be used for discrete, nominal or ordinal variables, and non-normal distribution but can be applied when the variables are in interval and ratio scales	4. It can be used for discrete, nominal or ordinal variables

Das, D. & Das, A. (Latest edition). Statistics in Biology and Psychology. Academic Publishers, Calcutta

Chi Square (χ^2) Test:

The Chi Square test is an important test amongst the several tests of significance developed by statisticians. Chi Square symbolically written as (χ^2), is a statistical measure used in the context of sampling analysis for comparing a variance to a theoretical variance. As a non-parametric test, it "can be used to determine if categorical data shows dependency or two classifications are independent. It can also be used to make comparisons between theoretical populations and actual data when categories are used". Thus, the chi-square test is applicable in large number of problems. The test is, in fact, a technique through the use of which it is possible for all researchers to (i) test the goodness of fit; (ii) test the significance of association between two attributes, and (iii) test the homogeneity or the significance of population variance.

Nonparametric chi square χ^2 test explores the significance of deviation of an experimentally observed frequency distribution from a proposed frequency distribution and, therefore, constitutes an analysis of frequencies.
The chi-square test represents a useful method of comparing experimentally obtained results with those to be expected theoretically on some hypothesis. The equation for chi-square (χ^2) is stated as follows:

$$\chi^2 = \Sigma \left[\frac{(f_o - f_e)^2}{f_e} \right] \qquad (69)$$

(chi-square formula for testing agreement between observed and expected results)

in which f_o = frequency of occurrence of observed or experimentally determined facts;
f_e = expected frequency of occurrence on some hypothesis.

df=(r-1) (c-1) that is (row-1) (column-1)

The differences between observed and expected frequencies are squared and divided by the expected number in each case, and the sum of these quotients is χ^2. The more closely the observed results approximate to the expected, the smaller the chi-square and the closer agreement between observed data and hypothesis being tested. Contrariwise, the larger the chi-square the greater the probability of a real divergence of experimentally observed from expected results. The number of *df* = (r-1) (c-1) in which r is the number of rows and c is the number of columns in which the data are tabulated.

Logic of Chi Square (χ^2) Test:

Chi-square test is applied to test the goodness of fit to verify the distribution of observed data with assumed theoretical distribution. It is a measure to study the divergence of actual and expected frequencies. It has great use in statistics, especially in sampling studies, where one can expect a doubted coincidence between actual and expected frequencies, and the extent to which the difference can be ignored, because of fluctuations in sampling. If there is no difference between the actual and expected frequencies, χ^2 is zero. Thus, it is used to describe the discrepancy between theory and observation.

Characteristics of Chi Square (χ^2) Test:

- Test is based on events or frequencies and not on the parameters like mean and standard deviation
- To draw inferences, this test is applied, specially testing the hypothesis but not useful to estimation
- It can be used between entire set of observed and expected frequencies
- For every increase in the number of degrees of freedom, a new χ^2 is formed

- The test can also be applied to a complex contingency table with several classes and highly useful in research
- It is a non-parametric test as no rigid assumptions are necessary in regard to the type of population, no need of parameter values and relatively less mathematical details are involved

Assumptions of Chi Square (χ^2) Test:

- It requires no assumption for normality of the population distribution of variable(s)
- Uses no precomputed statistic as an estimate of parameter in its computation
- It is applicable to very small sample
- Can be used also for discrete, nominal or ordinal variables. The chi-square (χ^2), test is used to determine whether an association (or relationship) between two categorical variables in a sample is likely to reflect a real association between these two variables in the population, due to this reason chi-square is used with data in the form of frequencies, or data that can be readily transformed into frequencies. This includes proportions or probabilities.

Calculation of the Chi-square (χ^2) goodness-of-fit-test:

As a test of goodness-of-fit, χ^2 test enables us to see how well does the assumed theoretical distribution fit to the observed data? When some theoretical distribution is fitted to the given data, it is our general interest to know as how well this distribution fits with the observed data. The chi-square test can give answer to this. If calculated value of χ^2 is less than the table value at a certain level of significance, the fit is considered to be a good one which means that the divergence between the observed and expected frequencies is attributed to fluctuations of sampling. But if the calculated value of χ^2 is greater than its table value, the fit is not considered to be a good one.

Chi-square test for Goodness of fit:

This test is used to explore how far a distribution of observed frequencies (f_o) fits with theoretical distribution such as the normal distribution, a binomial distribution, a Mendelian phenotype distribution, and a distribution proposed by hypothesis of equal probability. The f_e values are computed here on the basis of the proposed theoretical distribution. Theχ^2computed from the $(f_o$-$f_e)$ values proves significant if it equals or exceeds the critical χ^2 for the chosen level of significance; in such a case, the f_o distribution differs significantly from the proposed distribution. A computed χ^2 indicates that the f_o distribution fits with proposed distribution and does not differ significantly from the latter.

The classical formula, based on $(f_o - f_e)$ values, is used in computing the χ^2.

$$\chi^2 = \Sigma \frac{(f_o - f_e)^2}{f_e}.$$

The alternative formula, avoiding the use of f_e values, cannot be applied as no contingency table is involved. Yates' correction has to be applied if f_e of any class falls below 5 and the *df* amounts 1 only.

Degrees of freedom:

The *df* of the computed χ^2 depends in this test on the nature of the proposed theoretical distribution and is limited to the number of the classes in the distribution whose frequencies are free to vary. Distributions like Mendelian 9:3:3:1 phenotype distribution and equal probability hypothesis distribution do not involve any parameter; so, only one *df* is lost keeping the sample size (n) constant- for such distributions, *df*=k-1, where k is the number of classes in the distribution. For distributions involving estimates of one or more parameters, the *df* is further lowered by the number of parameters involved. For example, for a normal distribution, df amounts to k-3, because three degrees of freedom are lost in keeping n, μ and σ constant; for binomial and Poisson distributions, *df* amounts to k-2, because n and μ are to be kept constant.

(Reference: Das, D. & Das, A. (Latest edition). Statistics in Biology and Psychology. Academic Publishers, Calcutta)

Yates' Correction:

If any computed f_e is less than 5 and the χ^2 has the *df* of 1 only, Yates' correction has to be applied. Yates' correction brings each (f_o-f_e) closer to zero by 0.5-this means the subtraction of 0.5 from each positive (f_o-f_e) and the addition of 0.5 to each negative (f_o-f_e). The corrected (f_o-f_e) values are used for computing χ^2. In other words,

$$\chi^2 = \Sigma \frac{(|f_o - f_e| - 0.5)^2}{f_e}$$

where the bars on two sides of f_o-f_e indicate that all values of (f_o-f_e) are taken as positive, ignoring their algebraic signs.

Alternative Formula for Fourfold Contingency Table:

For 2×2-fold contingency tables, the following alternative method may be used avoiding the computation of $(f_o - f_e)$ values.

(a) For 2×2-fold contingency table, the classes of the two variables are arranged exactly like that for working out the phi-coefficient. Thus, the cell A at the top right corner and the cell D at the bottom left corner contain concordant cases belonging to either to the high value of classes of both variables (cell A) or to the low value classes of both (cell D); the cell B at the top left corner and the cell C at the bottom right corner contain discordant cases belonging to the high value class of one variable and low value of the class of the other, or vice-versa.
(b) The f_o values for different class combinations are then entered in the respective cells; marginal totals are also worked out for each row or column and entered as the f_r or f_c. Evidently, the f_r for the cells A and B amounts to the total of their cell frequencies, viz., A+B, and that for the cells C and D is given by the sum of their frequencies, viz., C+D; similarly, the f_c for the B and D amounts to B+D, while the f_c for cells A and C is the sum of their frequencies, viz., A+C.
(c) χ^2 is then computed using the products of pairs of cell frequencies (AD and BC) and the marginal totals.

$$\chi^2 = \frac{n(AD - BC)^2}{(A+B)(A+C)(B+D)(C+D)}.$$

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

If it seems that any f_e may be lower than 5, this is checked by computing the lowest f_e possible for any cell, using the smaller of the f_r values and the smaller of the f_c values.

$$\text{lowest } f_e = \frac{(\text{smaller } f_r) \times (\text{smaller } f_c)}{n}.$$

When f_e is found to be less than 5, Yates' correction is applied by changing the computational formula:

$$\chi^2 = \frac{n(|AD - BC| - n/2)^2}{(A+B)(A+C)(B+D)(C+D)}$$

where the bars on two sides of AD-BC indicate that (AD-BC) is taken as positive irrespective of its algebraic sign.

(Reference: Das, D. & Das, A. (Latest edition). Statistics in Biology and Psychology. Academic Publishers, Calcutta)

Another example of 2×2-fold contingency tables:

When the table is 2×2-fold χ^2 may be calculated without first computing the four expected frequencies-the four independence values. One example by explaining the method is given below:

All of the sixth-grade children in a public-school system a given a standard achievement test on arithmetic. A sample of 40 boys, drawn at random from the sixth-grade population, showed 23 at or above the national norm in the test and 17 below the national norm. A random sample of 50 sixth-grade girls showed 22 at or above national norm and 28 below. Are the boys really better than the girls in arithmetic? Data are arranged in a fourfold table as follows:

	below norm	at or above norm	
Boys	(A) 17	(B) 23	(A + B) 40
Girls	(C) 28	(D) 22	(C + D) 50
	(A + C) 45	(B + D) 45	N 90

In a fourfold table, chi square is given by the following formula.•

$$\chi^2 = \frac{N(AD - BC)^2}{(A+B)(C+D)(A+C)(B+D)} \quad (71)$$

(*chi square in a fourfold contingency table*)

Substituting for *A, B, C, D,* in the formula, we have

$$\chi^2 = \frac{90(374 - 644)^2}{40 \times 50 \times 45 \times 45} = 1.62$$

and for df=1, P is larger than .20. χ^2 is not significant and there is no evidence that the table entries really vary from expectation, i.e., that there is a true sex difference in arithmetic.
When entries in a fourfold table are quite small (for example, 5 or less) Yates' correction for continuity should be applied. The corrected formula is as follows:

$$\chi^2_c = \frac{N(|AD - BC| - N/2)^2}{(A+B)(C+D)(A+C)(B+D)} \quad (72)$$

(χ^2 *for 2 × 2 fold table, corrected for continuity*)

The vertical lines $|AD\text{-}BC|$ means that the difference is to be taken as positive. Here the above formula is applied to the data of the table. Substituting for N, A, B, C and D it is found that a value somewhat smaller than the χ^2 of 1.62 obtained without correction. Yates' correction will always reduce the size of χ^2. It should be used when entries are small; as it is here that its effect may be crucial. If the χ^2 is barely

$$\chi^2_c = \frac{90(|374 - 644| - 45)^2}{40 \times 50 \times 45 \times 45} = 1.12$$

significant, χ^2_c will fall may below the level set for significance. However, if χ^2 is not significant, χ^2_c will be even less so.

(Reference: Garrett, H.E. (1981). Statistics in Psychology and education. VakilsFetTer & Simons Ltd)

Another example of chi-square test when table entries small (less than 5):

When the entries in a table are fairly large, χ^2 gives an estimate of divergence from hypothesis which is close to that obtained by other measures of probability. But χ^2 is not stable when computed from a table in which any experimental frequency is **less than 5**. Moreover, when the table is 2×2 fold (when *df*=1), χ^2 is subject to considerable error unless a correction for continuity (called Yates' correction) is made. Reasons for making this correction and its effect upon χ^2 can best be seen by working through the examples following.

For example, an observer gave 7 correct judgments in ten trials. The probability of a right judgment was ½ in each instance, so that the expected number of correct judgments was 5. Test our subject's deviation from the null hypothesis by computing chi-square and compare the *P* with that found by direct calculation.

	Right	Wrong	
Observed (f_o)	7	3	10
Expected (f_e)	5	5	10
$(f_o - f_e)$	2	2	
Correction (−.5)	1.5	1.5	
$(f_o - f_e)^2$	2.25	2.25	
$\frac{(f_o - f_e)^2}{f_e}$	.45	.45	

$\chi^2 = .90$
$df = 1$
$P = .356$ (by interpolation in Table E)
$1/2P = .178$

Calculations in the above table indicate a correction which consists in subtracting .5 from each $(f_o\text{-}f_e)$ difference. In applying the χ^2 test it is assumed that adjacent frequencies are connected by a continuous and smooth curve (like normal curve) and are not discrete numbers. But in 2×2-fold tables, especially when entries are small, the χ^2 is not continuous. Hence, the deviation of 7 from 5 must be written as 1.5

(6.5-5) instead of 2 (7-5), as 6.5 is the lower limit of 7 in a continuous series. In like manner the deviation of 3 from 5 must be taken from the upper limit of 3, namely, 3.5. Still another change in procedure must be made in order to have the probability obtained from χ^2 agree with the direct determination of probability. P in the χ^2 table gives the probability of 7 or more right answers and of 3 or fewer right answers, i.e., takes account of both ends of the probability curve. We must have to consider ½ of *P*, therefore, if we want the probability of 7 or more right answers. If we repeated the test, we should expect a score of 7 or better about 17 times in 100 trials. It is clear, therefore, that the obtained score is not significant and does not refute the null hypothesis.

(Reference: Garrett, H.E. (1981). Statistics in Psychology and education. VakilsFetTer& Simons Ltd)

Another example of Chi-square when frequencies are small (less than 10):

In case of application of chi-square to a problem with *df*=1 and when f_e frequency is **less than10,** one should apply a modification known as Yates' correction continuity. This correction consisting in reducing by .5 each obtained frequency that is greater than expected and increasing by the same amount each frequency that is less than expected. This has the effect of reducing the amount of each difference between obtained and expected frequency to the extent of .5. The result is a reduction in the size of chi-square.

The correction is needed because a computed chi-square, being based on frequencies (which are whole numbers), varies in discrete jumps, whereas the chi-square table, representing the distributions of chi-square, gives values from a continuous scale. When frequencies are large, this connection is relatively unimportant, but when they are small, a change of .5 is of some consequence. The correction is particularly important when chi-square turns out to be near a point of division critical regions.

Suppose, in a public-opinion poll conducted some years ago, attitude towards radio newscasts were sampled. Some 43 interviewees in one sample were asked the question, "Do you find it easier to listen to news than to read it?" The samples had been stratified into higher and lower socio-economic status, 19 being in the former and 24 in the latter. The numbers responding "Yes" to the question in the two groups were 10 and 20 respectively. The problem to be investigated was whether there was a real difference between the two groups between the opinions on the question.

In the following table it is seen that two of the expected frequencies are less than 10. The computation is done first without Yates' correction and then with Yates' correction to see what difference it will make in conclusion?

TABLE 11.4 Computation of chi square for responses of two socioeconomic groups to preference for radio news to reading a newspaper

Response	*Obtained frequencies*			*Expected frequencies*		
	Lower	*Higher*	*Both*	*Lower*	*Higher*	*Both*
Yes	**20**	**10**	**30**	**16.74**	**13.26**	**30**
No	**4**	**9**	**13**	**7.26**	**5.74**	**13**
Both	**24**	**19**	**43**	**24**	**19**	**43**

Without the correction, the cell deviations would all equal 3.26. This value squared is 10.63. After applying the formula and solving it, the chi-square value equals to 4.76, which is significant between the

0.05 and 0.01 levels. With correction, the cell deviation in all cells is 2.76, which squared is 6.72. Hence, chi-square becomes 3.43 and fails to reach the 0.05 level of significance. The correction will not always make a difference of this kind in the conclusion, but it should be used in a problem like this.

It should be noted that the correction of .5 is applied to all cells in the table even though only one and two frequencies are small. Here the expected frequencies must be low that determine whether the correction should be applied, not low observed frequencies. It is also applied only to instances of 1 *df*, including 2×2 and 1×2 tables. In larger tables the need for correction is not so great, and it would be complicated to apply. It is also possible to combine categories in such a way as to get rid of small expected frequencies.

(Reference: Guilford, J. P (1981): Fundamental in Statistics in Psychology and Education. McGraw-Hill. Sixth Edition)

Another example of the same is given below:

In the experimental group 4 scores are above and 10 below the common median instead of the 7 above and 7 below to be expected by chance. In the control group, 12 scores are above and 6 below the common median instead of the expected 9 in each category. The frequencies are entered in the following table and χ^2 by the formula of Yates' correction.

	Below Median	Above Median	Total
Experimental	10	4	14
Control	6	12	18
	16	16	32

The χ^2 is, when corrected,

$$\chi^2_c = \frac{32(|120 - 24| - 32/2)^2}{16 \times 16 \times 18 \times 14}$$
$$= 3.17$$

A χ^2 of 3.17 with1 degree of freedom yields a *P* which lies at .08, about midway between .05 and .10. Here the question is to find out whether the median of the experimental group was significantly lower than that of the control. For this hypothesis, a one-tailed test, *P*/2, is approximately .04 and χ^2_c is significant at 0.05 level. If the hypothesis been that the two groups differ without specifying the direction, then we would definitely have had a two-tailed test and χ^2 would have been marked not significant. Conclusion tentatively indicates that the drug produces reduction in tremor. But owing to the small samples and lack of a highly significant finding, the clinical psychologist would almost certainly repeat the experiment-perhaps several times.

χ^2 is generally applicable in the median test. However, when N_1 and N_2 are small (e.g., less than about 10) the χ^2 test is not accurate and the exact method of computing probability is used.

(Reference: Garrett, H.E. (1981). Statistics in Psychology and education. VakilsFetTer & Simons Ltd)

Chi Square for 'Two-Way' Classification Variables-Contingency Table Analysis:

Chi-Square Test of Independence:

An association may exist between two variables if the change, in value or quality, of one variable is accompanied by a similar or opposite change of the other in the same individual. Absence of association is called the independence of two variables. A χ^2 test of independence explores whether or not two variables are significantly independent of each other; in other words, it explores the existence of any significant association between the variables. But it can (i) neither measure the magnitude and the direction of the association, (ii) nor predict a cause-and-effect relationship between the variables.

The data consist of frequencies of paired observations of the two variables, often discrete or nominal, distributed in different combinations of their classes. χ^2 is computed in the following way:

(a) The frequencies of paired observations are first arranged in a contingency table showing the association between two variables in their combined distribution. A contingency table is a two-way classification table with the classes of one variable arranged in r number of rows and those of other in the c number of columns. Each cell of the table represents a specific combination of two classes, one of each variable. According to the number of rows and columns, i.e., the number of classes of the variables, the table is an r × c fold table. For example, it is a 3×2 -fold contingency table where one variable has three classes as against two classes of the other. The degrees of freedom (*df*) of the computed chi-square χ^2 are given by: *df*=(r-1) (c-1).

The observed frequencies (f_o) of cases, belonging to the specific combinations of classes of the variables, are entered in the corresponding cells of the table. The total frequency of a row or column, f_r or f_c respectively in the marginal column or row; these total frequencies are called respective marginal frequencies. Where n is the total frequency or sample size, $\sum f_r = \sum f_c = n$.

(b) The expected frequency (f_e) of each cell is computed on the basis of H_O and entered against the corresponding f_o – the H_O for a test of independence proposes that the variables are independent of each other, that there is no significant distributions of f_o and f_e scores, and that any (f_o-f_e) difference has resulted from mere chances of random sampling and would not have occurred if the entire population was studied instead of a random sample. However, to keep the sample size (n) constant, the f_e values are directly computed, according to the following formula in terms of the H_o, for only as many randomly chosen cells as the *df* of the χ^2.

$$f_e = \frac{f_r . f_c}{n},$$

Where f_r and f_c are marginal total frequencies of respectively the row and the column to which the given cell belongs. The f_e of each of the remaining cells is computed by subtracting the already computed f_e values from the marginal totals.

(c)The difference (f_o-f_e) is worked out for each cell and entered in the table.

(d)Each such difference is squared to give the corresponding $(f_o-f_e)^2$.

(e)Each of the latter is divided by the corresponding f_e to get the ratio $(f_o-f_e)^2/f_e$. The sums of all these ratios are used in working out of χ^2.

$$\chi^2 = \Sigma \frac{(f_o - f_e)^2}{f_e}.$$

(Reference: Das, D. & Das, A. (Latest edition). Statistics in Biology and Psychology. Academic Publishers, Calcutta)

An example of Chi-square test of Independence in Contingency Table:

It is seen that χ^2may be employed to test the agreement between observed results and those expected on some hypothesis. A further useful application of χ^2 can be made when it is required to investigate the relationship between traits or attributes which can be classified into two or more categories. In the following two-way table the possession of a group by varying degrees of two characteristics is represented. In the following table 413 persons have been classified as to "eyedness" and "handedness". Eyedness, or eye dominance is described as left-eyed, ambiocular or right-eyed; handedness as left-handed, ambidextrous, or right-handed.

	Left-eyed	Ambiocular	Right-eyed	Totals
Left-handed	(35.4) 34	(58.5) 62	(30.0) 28	124
Ambidextrous	(21.4) 27	(35.4) 28	(18.2) 20	75
Right-handed	(61.1) 57	(101.0) 105	(51.8) 52	214
Totals	118	195	100	413

I. Calculation of independence values (f_e):

$$\frac{118 \times 124}{413} = 35.4 \qquad \frac{195 \times 124}{413} = 58.5 \qquad \frac{100 \times 124}{413} = 30.0$$

$$\frac{118 \times 75}{413} = 21.4 \qquad \frac{195 \times 75}{413} = 35.4 \qquad \frac{100 \times 75}{413} = 18.2$$

$$\frac{118 \times 214}{413} = 61.1 \qquad \frac{195 \times 214}{413} = 101.0 \qquad \frac{100 \times 214}{413} = 51.8$$

II. Calculation of χ^2:

$(-1.4)^2 \div 35.4 = .055$ $\quad (3.5)^2 \div 58.5 = .209$ $\quad (-2.0)^2 \div 30 = .133$

$(5.6)^2 \div 21.4 = 1.465$ $\quad (-7.4)^2 \div 35.4 = 1.547$ $\quad (1.8)^2 \div 18.2 = .178$

$(-4.1)^2 \div 61.1 = .275$ $\quad (4.0)^2 \div 101.0 = .158$ $\quad (.20)^2 \div 51.8 = .001$

$\chi^2 = 4.02$ $\quad df = 4$ $\quad$ *P* lies between .30 and .50

The hypothesis to be tested is the null hypothesis, namely, that handedness and eyedness are essentially unrelated or independent. In order to compute χ^2 it is necessary to calculate an "independence value" for each cell in the contingency table. Independence values are represented by the figures in parentheses within the different cells; they give the number of people whom one should expect to find possessing the designated eyedness and handedness combinations in the absence of any real association. In the table, there are 118 left-eyed and 124 left-handed persons. If there were no association between left-eyedness and left-handedness, one should expect to find, by chance, 118×124/413 or 35.4 individuals in the group who are left-eyed and left-handed. It is already revealed that 118/413 of the entire group are left-eyed. This proportion left-eyed individual should hold for any sub-group, if there is no dependence of handedness or eyedness. Hence, 118/413 or 28.6% of the 124 left-handed individuals, i.e., 35.4, should also be left-eyed.

When the expected or independence values have been computed, one finds the difference between observed and expected values for each cell square each difference and divide in each instance by independence value. The sum of these quotients gives χ^2. In the present problem χ^2=4.02 and *df*= (3-1) (3-1) or 4. Here *P* lies between .30 and .50 and hence χ^2 is not significant. The observed results are close to those to be expected on the hypothesis of the independence and there is no evidence of any real association between eyedness and handedness within the group.

(Reference: Garrett, H.E. (1981). Statistics in Psychology and education. VakilsFetTer & Simons Ltd)

The Additive Property of χ^2:

When several χ^2s' have been computed from independent experiments, these may be summed to give a new chi-square with *df*= the sum of separate *df*'s. The fact that chi-square may be added to provide an overall test of a hypothesis is important in many experimental studies. Suppose, in an arithmetic achievement test, the boys did slightly better than girls, but the chi-square of 1.62 is not large enough to indicate a superiority of boys over girls. The calculation is given below:

	below norm	at or above norm	
Boys	(A) 17	(B) 23	(A + B) 40
Girls	(C) 28	(D) 22	(C + D) 50
	(A + C) 45	(B + D) 45	N 90

In a fourfold table, chi square is given by the following formula.•

$$\chi^2 = \frac{N(AD - BC)^2}{(A+B)(C+D)(A+C)(B+D)} \quad (71)$$

(chi square in a fourfold contingency table)

Substituting for *A, B, C, D*, in the formula, we have

$$\chi^2 = \frac{90(374 - 644)^2}{40 \times 50 \times 45 \times 45} = 1.62$$

and for *df*=1, P is larger than .20. χ^2 is not significant and there is no evidence that the table entries really vary from expectation, i.e., that there is a true sex difference in arithmetic.
Suppose that three repetitions of this experiment are carried out, in each instance groups of boys and girls being drawn independently at random from the sixth grade and listed as scoring "at or above" or "below" the national norm. Further that the three chi-squares from these tables are 2.71, 5.39, and .15, in each case boys being somewhat better than the girls. No one can combine these four results to get an overall test of the significance of this sex difference in arithmetic. Adding these three χ^2's to the 1.62, one can get a total χ^2 of 9.87 with 4 *df*'s. Now the χ^2 is significant at 0.05 levels, and one can be reasonably sure that sixth-grade boys are, on the average, better than sixth -grade girls in arithmetic. It will be noted that four experiments taken in an aggregate yield a significant result, although one of the χ^2's (5.39) is itself significant. Combining the data from several experiments will often yield a conclusive result, when separate experiments, taken alone, provide only indications.

(Reference: Garrett, H.E. (1981). Statistics in Psychology and education. VakilsFetTer & Simons Ltd)

Interpretation of the Outcome of a Chi-Square Test:

Significance of the Computed χ^2:

The H_o contends that the computed χ^2 is not significant and has resulted from mere chances of random sampling. To test the H_o by a two-tail test, the computed χ^2 is compared to the critical χ^2 with the computed *df* and for the chosen level of significance (α). If the computed χ^2 exceeds or equals the $\chi^2_{\alpha(df),}$ the probability P of obtaining the computed χ^2 by chance is respectively lower than or equal to α and may be considered too low $(P \leq \alpha)$. The H_o is then rejected and the computed χ^2 is considered significant-the variables are then considered as to have a significant association with each other. But if the computed χ^2 is lower than the critical $\chi^2_{\alpha(df)}$, *P* exceeds α and is too high to justify the rejection of $H_o(P>\alpha)$; the computed χ^2 is not then not significant- the variables have no significant association with each other.

For one-tail χ^2 test, α for a given critical χ^2 is half that of an identical critical χ^2 for two-tail tests. In other words, the one-tail critical χ^2 for a given significance level is identical with two-tail critical χ^2 for double that significance level. Thus, the one-tail critical χ^2 (*df*=4) for 0.05 level amounts 7.78 which is identical with two-tail critical χ^2 (*df*=4) for 0.10 level.

(Reference: Das, D. & Das, A. (Latest edition). Statistics in Biology and Psychology. Academic Publishers, Calcutta)

Brief outline of the points that needs to be mentioned in Interpretation:

- **Significance Level:** The results of the Chi-square χ^2 test needs to be interpreted on the basis of levels of significance (0.05/0.01 levels of significance), that is, whether the obtained value is significant or not significant at 0.05 or 0.01 levels.
- **Hypothesis testing:** Testing of null hypothesis is considered as another important aspect of the Chi-square χ^2 test. Here, the null hypothesis is accepted or rejected should be mentioned.
- **Probability:** Here the findings needs to be related with probability (whether the association between the variables comes due to chance or not)
- **Variables mentioned:** Results must be interpreted on the basis of the variables mentioned in the problem

References:

Das, D. & Das, A. (Latest edition). Statistics in Biology and Psychology. Academic Publishers, Calcutta

Garrett, H.E. (1981). Statistics in Psychology and education. VakilsFetTer & Simons Ltd.

Guilford, J. P: Fundamental in Statistics in Psychology and Education

Kothari, C.R. (1990). Research Methodology Methods and Techniques. Wishwa Prakashan Calcutta, Second edition.

CC-9: STATISTICAL METHODS FOR PSYCHOLOGICAL RESEARCH-II

Practicum: On Computation of Chi Square (χ^2) test

(Two problems to be written in the practical files: One with Yates Correction and another without Yates Correction)

Introduction to Nonparametric Statistics:

In statistical test two kinds of assertions are involved, viz., an assertion directly related to the purpose of investigation and other assertion to make a probability statement. The former is an assertion to be tested and is technically called a hypothesis, whereas other assertion is called a model, it is known as distribution-free or non-parametric test. Under nonparametric or distribution free test, it is difficult to assume that a particular distribution is applicable or that a certain value is attached to a parameter of the population. Nonparametric statistics require few assumptions, no estimate of parameter in their computation, and no normal distribution of the variables in the population.

Parametric statistics serve as an estimation of the corresponding parameters. Their computations require the use of precomputed statistics as estimates of parameters. Moreover, they are interpreted with reference to specific population distributions of the variables such as normal and *t* distributions. Thus, they cannot be used for too small samples, nominal, ordinal and discrete variables, and non-normal distributions.

According to Bradley the characteristics of nonparametric statistics are as follows:

(a) Speed of Application: When sample size is small and moderate, distribution free methods are generally faster than parametric techniques.

(b) Scope of Application: Since nonparametric tests are based on fewer and less elaborate assumptions than parametric tests, nonparametric techniques can be correctly applied to a much larger class of population.

(c)Type of Measurement of require: Distribution-free statistical test usually require nominal or ordinal data.

(d) Influence on Sample Size: When the sample size is less or equal to 10, distribution free statistical tests are easier and quicker, though less efficient. At such sample sizes the parametric assumptions may not be satisfied for this, nonparametric tests are most appropriate. As the sample size increases the nonparametric tests become more laborious and time consuming and frequently become a much less efficient statistical test.

(e) Susceptibility to violations of assumptions: Since the assumptions are fewer and less elaborate with nonparametric statistical test, they are less susceptible to violation.

(f) In terms of mathematical criterion of statistical efficiency, distribution free tests are often superior or equal to their parametric counterparts particularly when, the assumptions of nonparametric tests are fulfilled but the assumptions of parametric tests are not. If both tests are applied when all assumptions of parametric tests are made, distribution free statistics are only slightly less efficient, when sample size is small.

According to Moses, nonparametric statistics are:

(a) easier to apply; (b) applicable to ranked data; (c) usable to two sets of observation coming from different population; (d) the only alternative when sample size is small, (e) useful at a specified significant level, (f) less efficient statistics, (g) not applicable to normal distribution in the variables of the population, (h) not included any precomputed statistic as an estimate of parameter in the computation.

Differences between Parametric and Nonparametric Statistics:

Parametric Statistics	**Nonparametric Statistics**
1. It serves as estimates of corresponding parameters. Their computations require the use of precomputed statistics as estimates of parameter	1. It requires very few assumptions. Do not use any precomputed statistics as an estimate of parameter in the computation
2. They are interpreted with reference to specific population distributions of the variables such as the normal and t distributions	2. Do not require normal distributions of the variables in the population
3. It cannot be applied for too small samples	3. It can be used for very small samples
4. They cannot be used for discrete, nominal or ordinal variables, and non-normal distribution but can be applied when the variables are in interval and ratio scales	4. It can be used for discrete, nominal or ordinal variables

The Chi-square (χ^2) is an important test among the several tests of significance developed by statistician. As a nonparametric test it is used to determine if categorical data shows dependence or two classifications are independent. It can also be used to make comparison between theoretical population and actual data when categories are used. Thus, the chi-square test is applicable to a large number of problems. The test is, in fact the technique through the use of which it is possible for all researchers who:

(a) test the goodness of fit; (b) test the significance of association between two attributes, (c) test the homogeneity or significance of population variance.

Chi-square (χ^2) is an important nonparametric test and as such no rigid assumptions are necessary in respect of the type of population. As a nonparametric test, Chi-square (χ^2) can be used as a goodness of fit and as a test of independence.

Nonparametric Chi-square (χ^2) test explores the significance of deviation of an experimentally observed frequency distribution from a proposed frequency distribution and, therefore, constitutes an analysis of frequencies.

The chi-square test represents a useful method of comparing experimentally obtained results with those to be expected theoretically on some hypothesis. The equation for chi-square (χ^2) is stated as follows:

$$\chi^2 = \Sigma \left[\frac{(f_o - f_e)^2}{f_e} \right] \qquad (69)$$

(chi-square formula for testing agreement between observed and expected results)

in which f_o = frequency of occurrence of observed or experimentally determined facts;

f_e = expected frequency of occurrence on some hypothesis.

df=(r-1) (c-1) that is (row-1) (column-1)

The differences between observed and expected frequencies are squared and divided by the expected number in each case, and the sum of these quotients is χ^2. The more closely the observed results approximate to the expected, the smaller the chi-square and the closer agreement between observed data and hypothesis being tested. Contrariwise, the larger the chi-square the greater the probability of a real divergence of experimentally observed from expected results.

Assumptions of Chi-square Test:

- It requires no assumption for normality of the population distribution of variable(s)
- Uses no precomputed statistic as an estimate of parameter in its computation
- It is applicable to very small sample
- Can be used also for discrete, nominal or ordinal variables. The chi-square (χ^2), test is used to determine whether an association (or relationship) between two categorical variables in a sample is likely to reflect a real association between these two variables in the population, due to this reason chi-square is used with data in the form of frequencies, or data that can be readily transformed into frequencies. This includes proportions or probabilities.

Chi-square test for Goodness of fit:

This test is used to explore how far a distribution of observed frequencies (f_o) fits with theoretical distribution such as the normal distribution, a binomial distribution, a Mendelian phenotype distribution, and a distribution proposed by hypothesis of equal probability. The f_e values are computed here on the basis of the proposed theoretical distribution. The χ^2 computed from the (f_o-f_e) values proves significant if it equals or exceeds the critical χ^2 for the chosen level of significance; in such a case, the f_o distribution differs significantly from the proposed distribution. A computed χ^2 indicates that the f_o distribution fits with proposed distribution and does not differ significantly from the latter.

The classical formula, based on (f_o-f_e) values, is used in computing the χ^2.

$$\chi^2 = \Sigma \frac{(f_o - f_e)^2}{f_e}.$$

The alternative formula, avoiding the use of f_e values, cannot be applied as no contingency table is involved. Yates' correction has to be applied if f_e of any class falls below 5 and the *df* amounts 1 only.

Chi-Square Test of Independence:

An association may exist between two variables if the change, in value or quality, of one variable is accompanied by a similar or opposite change of the other in the same individual. Absence of association is called the independence of two variables. A χ^2 test of independence explores whether or not two variables are significantly independent of each other; in other words, it explores the existence of any significant association between the variables. But it can (i) neither measure the magnitude and the direction of the association, (ii) nor predict a cause-and-effect relationship between the variables.

The Additive Property of χ^2:

When several χ^2s' have been computed from independent experiments, these may be summed to give a new chi-square with *df*= the sum of separate *df*'s. The fact that chi-square may be added to provide an overall test of a hypothesis is important in many experimental studies. Combining the data from several experiments will often yield a conclusive result, when separate experiments, taken alone, provide only indications.

For example

Problem No: 1

General Problem: On Nonparametric Statistics

Specific Problem: The following table represents possessions of two characteristics of varying degrees by a group. In total 413 persons have been classified as to "eyedness" and "handedness". Do these data handedness and eyednesses are essentially independent? Test the hypothesis at 0.05 level.

	Left-eyed	Ambiocular	Right-eyed	Totals
Left-handed	34	62	28	124
Ambidextrous	27	28	20	75
Right-handed	57	105	52	214
Totals	118	195	100	413

Concept: Nonparametric chi-square (χ^2) test explores the significance of deviation of an experimentally observed frequency distribution from a proposed frequency distribution and, therefore constitutes an analysis of frequencies. When the data consists of frequencies in discrete categories, the chi-square test may be used to determine the significance of association between two independent groups. The hypothesis being tested is that the two groups differ or associated with respect to some characteristics and therefore, with respect to the relative frequency with which group members fall into several categories. To test the hypothesis, one must count the number of cases from each group which fall in various categories and compare the proportion of cases from the other group. The focus of the test is whether the differences in proportion exceed those expected due to the chance or the difference or association is genuine.

Assumptions: (a) It requires no assumption for normality of the population distribution of variable(s) (b) Uses no precomputed statistic as an estimate of parameter in its computation (c) It is applicable to very small sample (d) Can be used also for discrete, nominal or ordinal variables. The chi-square (χ^2), test is used to determine whether an association (or relationship) between two categorical variables in a sample is likely to reflect a real association between these two variables in the population, due to this reason chi-square is used with data in the form of frequencies, or data that can be readily transformed into frequencies. This includes proportions or probabilities.

Null Hypothesis: There is no significant association between the two characteristics namely handedness and eyedness. Any such association is due to chance alone.

Statistic used for Testing of Hypothesis: A two-tail Chi-square (χ^2) test will be used here for testing the hypothesis. The relevant formula is as follows:

$$\chi^2 = \Sigma \left[\frac{(f_o - f_e)^2}{f_e} \right] \quad (69)$$

(chi-square formula for testing agreement between observed and expected results)

in which f_o = frequency of occurrence of observed or experimentally determined facts;

f_e = expected frequency of occurrence on some hypothesis.

df=(r-1) (c-1) that is (row-1) (column-1)

Statistical Treatment:

Calculation needs to be done

Interpretation and Conclusion: The obtained chi-square (χ^2) value has been found as -----. The critical value of chi-square with *df*=------------ at 0.05 level is----------. As the obtained chi-square value is lesser/higher than the critical value, therefore it can be said that the obtained chi-square value is not significant/significant at 0.05 level. So, the probability of obtaining such association by chance is greater/lesser than 0.05. It may also be said that the null hypothesis is rejected/accepted and the alternative hypothesis is rejected/accepted. In other words, there is no significant association/significant association between handedness and eyedness. It means that the two characteristics eyedness and handedness are independent/not independent of each other.

For example

Problem No: 2

General Problem: On Nonparametric Statistics

Specific Problem: Out of 15 diabetic subjects, 8 were found to be suffering from hypercholesterolemia while the rest had normal serum cholesterol. Out of 10 nondiabetics, 2 had high serum cholesterol while the rest had normal serum cholesterol level. Is there any significant association between hypercholesterolemia and diabetes?

Concept: Nonparametric chi-square (χ^2) test explores the significance of deviation of an experimentally observed frequency distribution from a proposed frequency distribution and, therefore constitutes an analysis of frequencies. When the data consists of frequencies in discrete categories, the chi-square test may be used to determine the significance of association between two independent groups. The hypothesis is being tested is that the two groups differ or associated with respect to some characteristics and therefore, with respect to the relative frequency with which group members fall into several categories. To test the hypothesis, one must count the number of cases from each group which fall in various categories and compare the proportion of cases from the other group. The focus of the test is whether the differences in proportion exceed those expected due to the chance or the difference or association is genuine.

Assumptions:(a)It requires no assumption for normality of the population distribution of variable(s) (b) Uses no precomputed statistic as an estimate of parameter in its computation (c) It is applicable to very small sample (d) Can be used also for discrete, nominal or ordinal variables. The chi-square (χ^2), test is used to determine whether an association (or relationship) between two categorical variables in a sample is likely to reflect a real association between these two variables in the population, due to this reason chi-square is used with data in the form of frequencies, or data that can be readily transformed into frequencies. This includes proportions or probabilities.

Null Hypothesis: There is no significant association between hypercholesterolemia and diabetes. Any such association is due to chance alone.

Statistic used for Testing of Hypothesis: A two-tail Chi-square (χ^2) test will be used here for testing the hypothesis. The relevant formula is as follows:

$$\chi^2 = \Sigma \left[\frac{(f_o - f_e)^2}{f_e} \right] \qquad (69)$$

(*chi-square formula for testing agreement between observed and expected results*)

in which f_o = frequency of occurrence of observed or experimentally determined facts;

f_e = expected frequency of occurrence on some hypothesis.

df=(r-1) (c-1) that is (row-1) (column-1)

In this context, however it may be noted that, if any computed f_e is less than 5 and the χ^2 has the *df* of 1 only, Yates' correction has to be applied. Yates' correction brings each (f_o- f_e) closer to zero by 0.5-this means the subtraction of 0.5 from each positive (f_o- f_e) and the addition of 0.5 to each negative (f_o- f_e). The corrected (f_o- f_e) values are used for computing χ^2. In other words,

$$\chi^2 = \Sigma \frac{(|f_o - f_e| - 0.5)^2}{f_e}$$

Where the bars on the two sides of f_o- f_e indicate that all values of (f_o- f_e) are taken as positive, ignoring their algebraic sigs.

Alternative formula:

When any f_e is less than 5, Yates' correction is applied by changing the computational formula:

$$\chi^2 = \frac{n(|AD - BC| - n/2)^2}{(A + B)(A + C)(B + D)(C + D)}$$

Where the bars on the two sides of AD-BC indicate that (AD-BC) is take n as positive irrespective of its algebraic sign.

The correction is needed because a computed chi-square based on frequencies (which are whole numbers), varies in discrete jumps, whereas the chi-square table, representing the distribution of chi-square gives values from a continuous scale. When frequencies are large, this correction is relatively unimportant, but when they are small, a chance of 0.5 is some consequence. The correction is particularly important when chi-square turn out to be near a point of division between critical regions.

Statistical Treatment:

	Non-Hypercholesterolemic	Hypercholesterolemic	Total
Diabetic	7	8	15
Nondiabetic	8	2	10
Total	15	10	25(n)

Calculation needs to done

Interpretation and Conclusion: The obtained chi-square (χ^2) value has been found as ------. The critical value of chi-square with *df*=1 at 0.05 level is----------. As the obtained chi-square value is lesser/higher than the critical value, therefore it can be said that the obtained chi-square value is not significant/significant at 0.05 level. So, the probability of obtaining such association by chance is greater/lesser than 0.05. It may also be said that the null hypothesis is rejected/accepted and the alternative hypothesis is rejected/accepted. In other words, there is no significant association/significant association between hypercholesterolemia and diabetes.

* The statistics mentioned here are just considered as an example to show the format of writing, but for practical file any statistical calculation (One with Yates' Correction and another without Yates' Correction) can be done according to the decision taken into the workshop.