CHI-SQUARE TEST

(Dr. N. N. Mahto, Visthapit Mahavidyalaya, Balidih)

AMONGST the several tests used in statistics for judging the significance of the sample data, Chi Square test, developed by Karl Pearson, in 1900 is considered an important test . Chi Square symbolically written as x2(pronounced as Ki –square) is a statistical measure with the help of which it is possible to assess the significance of the different between the observed frequency and expected frequencies obtained from some hypothetical universe. Chi-square tests enable us to test whether more than two population proportions can be considered equal. In order that chi-square test may be applicable, both the frequencies must be grouped in the same way and the theoretical distribution must be adjusted to give the same total frequency which is equal to that of observed frequencies. Chi square is calculated with help of the following formula:

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

Where, fo = means of the observed frequency

Fe = means of the expected frequency

If two distributions (observed and theoretical) are exactly alike, as $\chi^2 = 0$ but generally due to sampling error Z is not equal to zero and we must know the sampling distribution of x2 so that we may be able to find the probability of an observed x2 being given by the a random sample from the hypothetical universe. Instead of working out of the probabilities we can use the ready tables which give the probabilities for given values of x2. Whether or not a calculated value of x2 is significant, can be ascertained by looking at the tabulated values x2(given) for given degree of freedom at the certain level of confidence (generally 5% level is taken) if the some time the x2 formula is stated under

$$\chi^2 = \sum(O- E)/E$$

where o= observed frequency

E= expected frequency.

Calculated value of x2 exceed the table value, the difference between the observed and expected frequencies is taken as significant but if the table value is more that the calculated value of x2, then the difference between the observed

Degrees of Freedom

Definition: The Degrees of Freedom refers to the number of values involved in the calculations that have the freedom to vary. In other words, the degrees of freedom, in general, can be defined as the total number of observations minus the number of independent constraints imposed on the observations.

commonly abbreviated as, df. The statistical formula to compute the value of degrees of **df= n-1** freedom is quite simple and is equal to the number of values in the data set minus one. Symbolically: degrees of freedom is usually denoted by a greek symbol v (mu) and is

df= n-1

Where n is the number of values in the data set or the sample size. The concept of df can be further understood through an illustration given below:

Suppose there is a data set X that includes the values: 10,20,30,40. First of all, we will calculate the mean of these values, which is equal to:

(10+20+30+40) /4 = 25.

Once the mean is calculated, apply the formula of degrees of freedom. As the number of values in the data set or sample size is 4, so,

df = 4-1=3.

Thus, this shows that there are three values in the data set that have the freedom to vary as long as the mean is 25.

Levels of Significance of Chi-Square Test:

The calculated values of χ^2 (Chi-square) are compared with the table values, to conclude whether the difference between expected and observed frequencies is due to the sampling fluctuations and as such significant or whether the difference is due to some other reason and as such significant. The divergence of theory and fact is always tested in terms of certain probabilities.

The probabilities indicate the extent of reliance that we can place on the conclusion drawn. The table values of χ^2 are available at various probability levels. These levels are called levels of significance. Usually the value of χ^2 at .05 and .01 level of significance for the given degrees of freedom is seen from the tables.

If the calculated value of χ^2 is greater than the tabulated value, it is said to be significant. In other words, the discrepancy between the observed and expected frequencies cannot be attributed to chance and we reject the null hypothesis.

Thus we conclude that the experiment does not support the theory. On the other hand if calculated value of χ^2 is less than the corresponding tabulated value then it is said to be non-significant at the required level of significance.

This implies that the discrepancy between observed values (experiment) and the expected values (theory) may be attributed to chance, i.e., fluctuations of sampling.

Chi-Square Test under Null Hypothesis:

Suppose we are given a set of observed frequencies obtained under some experiment and we want to test if the experimental results support a particular hypothesis or theory. Karl Pearson in 1990, developed a test for testing the significance of the discrepancy between experimental values and the theoretical values obtained under some theory or hypothesis.

This test is known as χ^2-test and is used to test if the deviation between observation (experiment) and theory may be attributed to chance (fluctuations of sampling) or if it is really due to the inadequacy of the theory to fit the observed data.

Under the Null Hypothesis we state that there is no significant difference between the observed (experimental) and the theoretical or hypothetical values, i.e., there is a good compatibility between theory and experiment.

The equation for chi-square (χ^2) is stated as follows:

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

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

f_e = expected frequency of occurrence on some hypothesis.

Thus chi-square is the sum of the values obtained by dividing the square of the difference between observed and expected frequencies by the expected frequencies in each case. In other words 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.

Several illustrations of the chi-square test will clarify the discussion given above.

The differences of f_o and f_e are written always + ve.

1. Testing the divergence of observed results from those expected on the hypothesis of equal probability (null hypothesis):

Example 1:

Ninety-six subjects are asked to express their attitude towards the proposition "Should AIDS education be integrated in the curriculum of Higher secondary stage" by marking F (favourable), I (indifferent) or U (unfavourable).

It was observed that 48 marked 'F', 24 'I' and 24 'U':

(i) Test whether the observed results diverge significantly from the results to be expected if there are no preferences in the group.

(ii) Test the hypothesis that "there is no difference between preferences in the group".

(iii) Interpret the findings.

Solution:

Following steps may be followed for the computation of x^2 and drawing the conclusions:

Step 1:

Compute the expected frequencies (f_e) corresponding to the observed frequencies in each case under some theory or hypothesis.

In our example the theory is of equal probability (null hypothesis). In the second row the distribution of answers to be expected on the null hypothesis is selected equally.

	Favourable	Indifferent	Unfavourable	
Observed (f_o)	48	24	24	96
Expected (f_e)	32	32	32	96
($f_o - f_e$)	16	8	8	
$(f_o - f_e)^2$	256	64	64	
$\frac{(f_o - f_e)^2}{f_e}$	8	2	2	

$$\chi^2 = \Sigma\left[\frac{(f_o - f_e)^2}{f_e}\right] = 12 \qquad df = 2 \qquad \text{P is less than .01}$$

Step 2:

Compute the deviations ($f_o - f_e$) for each frequency. Each of these differences is squared and divided by its f_e (256/32, 64/32 and 64/32).

Step 3:

Add these values to compute:

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

From Table E
Tabulated χ^2 with 2 df at .01 level = 9.21

$\left(\frac{256}{32} + \frac{64}{32} + \frac{64}{32}\right)$ to give $\chi^2 = 8 + 2 + 2 = 12$

Step 4:

The degrees of freedom in the table is calculated from the formula df = (r – 1) (c – 1) to be (3 – 1) (2 – 1) or 2.

Step 5:

Look up the calculated (critical) values of χ^2 for 2 df at certain level of significance, usually 5% or 1%.

With df = 2, the χ^2 value to be significant at .01 level is 9.21 (Table E). The obtained χ^2 value of 12 > 9.21.

i. Hence the marked divergence is significant.

ii. The null hypothesis is rejected.

iii. We conclude that our group really favours the proposition.

We reject the “equal answer” hypothesis and conclude that our group favours the proposition.

Example 2:

The number of automobile accidents per week in a certain community were as follows:

12, 8, 20, 2, 14, 10, 15, 6, 9, 4

Are these frequencies in agreement with the belief that accident conditions were the same during this 10-week period?

Solution:

Null Hypothesis—Set up the null hypothesis that the given frequencies (of number of accidents per week in a certain community) are consistent with the belief that the accident conditions were same during the 10-week period.

Since the total number of accidents over the 10 weeks are:

$$12 + 8 + 20 + 2 + 14 + 10 + 15 + 6 + 9 + 4 = 100.$$

Under the null hypothesis, these accidents should be uniformly distributed over the 10-week period and hence the expected number of accidents for each of the 10 weeks are 100/10 = 10.

Computation of χ^2

Week	Observed No. of accidents (f_o)	Expected No. of accidents (f_e)	($f_o - f_e$)	$(f_o - f_e)^2$	$\frac{(f_o - f_e)^2}{f_e}$
1.	12	10	2	4	0.4
2.	8	10	2	4	0.4
3.	20	10	10	100	10.0
4.	2	10	8	64	6.4
5.	14	10	4	16	1.6
6.	10	10	0	0	0.0
7.	15	10	5	25	2.5
8.	6	10	4	16	1.6
9.	9	10	1	1	0.1
10.	4	10	6	36	3.6
	100	100			26.6

$$\therefore \quad \chi^2 = \Sigma\left[\frac{(f_o - f_e)^2}{f_e}\right] = 26.6$$

$df = 9$.P is less than .01

From Table E
Tabulated χ^2 with 9 df
at .05 level = 16.916
at .01 level = 21.666

Since calculated value of $\chi^2 = 26.6$ is greater than the tabulated value, 21.666. It is significant and the null hypothesis rejected at .01 level of significance. Hence we conclude that the accident conditions are certainly not uniform (same) over the 10-week period.

2. Testing the divergence of observed results from those expected on the hypothesis of a normal distribution:

The hypothesis, instead of being equally probable, may follow the normal distribution. An example illustrates how this hypothesis may be tested by chi-square.

Example 3:

Two hundred salesmen have been classified into three groups very good, satisfactory, and poor—by consensus of sales managers.

Does this distribution of rating differ significantly from that to be expected if selling ability is normally distributed in our population of salesmen?

	Good	Satisfactory	Poor	Total
Observed (f_o)	76	96	28	200
Expected (f_e)	32	136	32	200
$(f_o - f_e)$	44	40	4	
$(f_o - f_e)^2$	1936	1600	16	
$\frac{(f_o - f_e)^2}{f_e}$	60.50	11.76	0.50	

$$\chi^2 = \Sigma\left[\frac{(f_o - f_e)^2}{f_e}\right] = 60.50 + 11.76 + .50 = 72.76$$

We set up the hypothesis that selling ability is normally distributed. The normal curve extends from – 3σ to + 3σ. If the selling ability is normally distributed the base line can be divided into three equal segments, i.e.

Rating	σ range between	% in Table A	% of 200 or (f_e)
Good	+ 3.00 σ and + 1.00 σ	16%	32
Satisfactory	+ 1.00 σ and – 1.00 σ	68%	136
Poor	– 1.00 σ and – 3.00 σ	16%	32
		100%	200

(+ 1σ to + 3σ), (- 1σ to + 1σ) and (- 3σ to – 1σ) representing good, satisfactory and poor salesmen respectively. By referring Table A we find that 16% of cases lie between + 1σ and +3σ, 68% in between – 1σ and + 1σ and 16% in between – 3σ and – 1σ. In case of our problem 16% of 200 = 32 and 68% of 200 = 136.

df= 2. P is less than .01

The calculated χ^2 = 72.76

The calculated χ^2 of 72.76 > 9.21. Hence P is less than .01.

From Table E
Tabulated χ^2 for $2df$ at .01 level = 9.21

∴ The discrepancy between observed frequencies and expected frequencies is quite significant. On this ground the hypothesis of a normal distribution of selling ability in this group must be rejected. Hence we conclude that the distribution of ratings differ from that to be expected.

3. Chi-square test when our expectations are based on predetermined results:

Example 4:

In an experiment on breeding of peas a researcher obtained the following data:

The theory predicts the proportion of beans, in four groups A, B, C and D should be 9: 3: 3: 1. In an experiment among 1,600 beans, the numbers in four groups were 882, 313, 287 and 118. Does the experiment results support the genetic theory? (Test at .05 level).

Solution:

We set up the null hypothesis that there is no significant difference between the experimental values and the theory. In other words there is good correspondence between theory and experiment, i.e., the theory supports the experiment.

Category	Expected frequency (f_e)
A	$\frac{9}{16} \times 1600 = 900$
B	$\frac{3}{16} \times 1600 = 300$
C	$\frac{3}{16} \times 1600 = 300$
D	$\frac{1}{16} \times 1600 = 100$

$9 + 3 + 3 + 1 = 16$

Computation of $\chi 2$

	A	B	C	D
Observed frequency f_o	882	313	287	118
Expected frequency f_e	900	300	300	100
$(f_o - f_e)$	18	13	13	18
$(f_o - f_e)^2$	324	169	169	324
$\frac{(f_o - f_e)^2}{f_e}$	.360	.563	.563	3.240

$$\Sigma \left[\frac{(f_o - f_e)^2}{f_e}\right] = .360 + .563 + .563 + 3.240$$

$$= 4.726$$

$df = 3$ P is near about .20

The calculated $\chi^2 = 4.726$

From Table E
Tabulated χ^2 for $3df$
at .05 level = 7.81

Since the calculated χ^2 value of 4.726 < 7.81, it is not significant. Hence null hypothesis may be accepted at .05 level of significance and we may conclude that the experimental results support the genetic theory.

4. The Chi-square test when table entries are small:

When table entries are small and when table is 2 x 2 fold, i.e., df = 1, χ^2 is subject to considerable error unless a correction for continuity (called Yates' Correction) is made.

Example 5:

Forty rats were offered opportunity to choose between two routes. It was found that 13 chose lighted routes (i.e., routes with more illumination) and 27 chose dark routes.

(i) Test the hypothesis that illumination makes no difference in the rats' preference for routes (Test at .05 level).

(ii) Test whether the rats have a preference towards dark routes.

Solution:

If illumination makes no difference in preference for routes i.e., if H_0 be true, the proportionate preference would be 1/2 for each route (i.e., 20).

In our example we are to subtract .5 from each ($f_o - f_e$) difference for the following reason:

In 2×2 fold tables, especially when entries are small, the χ^2 curve is not continuous. Hence, the deviation of 27 from 20 must be written as 6.5 (26.5 – 20) instead of 7(27 – 20), as 26.5 is the lower limit of 27 in a continuous series. In like manner the deviation of 13 from 20 must be taken from the upper limit of 13, namely, 13.5.

The data can be tabulated as follows:

	Dark routes	Lighted routes	Total
Observed frequencies f_o	27	13	40
Expected frequencies f_e	20	20	40
$(f_o - f_e)$	7	7	
$[(f_o - f_e) - .5]$	6.5	6.5	
$[(f_o - f_e) - .5]^2$	42.25	42.25	
$\frac{[(f_o - f_e) - .5]^2}{f_e}$	2.11	2.11	

$\therefore \chi^2 = 2.11 + 2.11 = 4.22.$

When the expected entries in 2 x 2 fold table are the same as in our problem the formula for chi-square may be written in a somewhat shorter form as follows:

$$\chi^2 = \frac{2[(f_o - f_e) - .5]^2}{f_e} \quad \text{...(56)}$$

$$= \frac{2(6.5)^2}{20} = \frac{2 \times 42.25}{20} = 4.22$$

From Table E

The tabulated value of χ^2 for 1 *df* at .05 level = 3.841.

df = 1 P is .043 (by interpolation)

Calculated $\chi^2 = 4.22$

the null hypothesis is rejected at .05 level. Apparently light or dark is a factor in the rats' choice for routes. (i) The critical value of χ^2 at .05 level is 3.841. The obtained χ^2 of 4.22 is more than 3.841. Hence

(ii) In our example we have to make a one-tailed test. Entering table E we find that χ^2 of 4.22 has a P = .043 (by interpolation).

$\therefore$ P/2 = .0215 or 2%. In other words there are 2 chances in 100 that such a divergence would occur.

Hence we mark the divergence to be significant at 02 level.

Therefore, we conclude that the rats have a preference for dark routes.

5. The Chi-square test of independence in contingency tables:

Sometimes we may encounter situations which require us to test whether there is any relationship (or association) between two variables or attributes. In other words χ^2 can be made when we wish to investigate the relationship between traits or attributes which can be classified into two or more categories.

For example, we may be required to test whether the eye-colour of father is associated with the eye-colour of sons, whether the socio-economic status of the family is associated with the preference of different brands of a commodity, whether the education of couple and family size are related, whether a particular vaccine has a controlling effect on a particular disease etc.

To make a test we prepare a contingency table end to calculate f_e (expected frequency) for each cell of the contingency table and then compute χ^2 by using formula:

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

Null hypothesis:

χ^2 is calculated with an assumption that the two attributes are independent of each other, i.e. there is no relationship between the two attributes.

The calculation of expected frequency of a cell is as follows:

$$f_e \text{ of a cell} = \frac{\text{Row Total} \times \text{Column Total}}{\text{Grand total}}$$

Example 6:

In a certain sample of 2,000 families 1,400 families are consumers of tea where 1236 are Hindu families and 164 are non-Hindu. And 600 families are not consumers of tea where 564 are Hindu families and 36 are non-Hindu. Use χ^2 – test and state whether there is any significant difference between consumption of tea among Hindu and non-Hindu families.

Solution:

The above data can be arranged in the form of a 2 x 2 contingency table as given below:

	Hindu	Non-Hindu	Total
Families consuming tea	(I) 1236	(II) 164	1400
Families not consuming tea	(III) 564	(IV) 36	600
Grand Total	1800	200	2000

We set up the null hypothesis (H_0) that the two attributes viz., 'consumption of tea' and the 'community' are independent. In other words, there is no significant difference between the consumption of tea among Hindu and non-Hindu families.

Calculation of (f_e) :

	Hindu	Non-Hindu	Total
Families consuming tea	(I) $\frac{1800 \times 1400}{2000}$ $= 1260$	(II) $\frac{200 \times 1400}{2000}$ $= 140$	1400
Families not consuming tea	(III) $\frac{1800 \times 600}{2000}$ $= 540$	(IV) $\frac{200 \times 600}{2000}$ $= 60$	600
Total	1800	200	2000

f_e of each cell $= \frac{\text{Row Total} \times \text{Column Total}}{\text{Grand total}}$

Calculation of χ^2

Cells	f_o	f_e	$(f_o - f_e)$	$(f_o - f_e)^2$	$\frac{(f_o - f_e)^2}{f_e}$
I	1236	1260	24	576	0.4571
II	164	140	24	576	4.1143
III	564	540	24	576	1.0667
IV	36	60	24	576	9.6000

$(f_o - f_e)$ is written disregarding sign. $\chi^2 = 15.2381$

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

P is less than .01

Calculated $\chi^2 = 15.2381$

From Table E
Tabulated value of χ^2 for 1 *df* at
.05 level = 3.841
.01 level = 6.635

Since the calculated value of χ^2, viz., 15.24 is much greater than the tabulated value of χ^2 at .01 level of significance; the value of χ^2 is highly significant and null hypothesis is rejected.

Hence we conclude that the two communities (Hindu and Non-Hindus) differ significantly as regards the consumption of tea among them.

Example 7:

The table given below shows the data obtained during an epidemic of cholera.

	Attacked	Non Attacked	Total
Inoculated	31	469	500
Not Inoculated	185	1315	1500
Total	216	1784	2000

Test the effectiveness of inoculation in preventing the attack of cholera.

Solution:

We set up the null hypothesis (H_0) that the two attributes viz., inoculation and absence of attack from cholera are not associated. These two attributes in the given table are independent.

	Attacked	Non Attacked	Total
Inoculated	(I) 31	(II) 469	500
Not Inoculated	(III) 185	(IV)1315	1500
Total	216	1784	2000

Basing on our hypothesis we can calculate the expected frequencies as follows:

Calculation of (f_e):

	Attacked	Not Attacked	Total
Inoculated	(I) $\frac{500 \times 216}{2000} = 54$	(II) $\frac{500 \times 1784}{2000} = 446$	500
Not Inoculated	(III) $\frac{1500 \times 216}{2000} = 162$	(IV) $\frac{1500 \times 1784}{2000} = 1338$	1500
Total			2000

$$f_e \text{ of each cell} = \frac{\text{Row Total} \times \text{Column Total}}{\text{Grand Total}}$$

Calculation of χ^2

Cells	f_o	f_e	$(f_o - f_e)$	$(f_o - f_e)^2$	$\frac{(f_o - f_e)^2}{f_e}$
I	31	54	23	529	9.796
II	469	446	23	529	1.186
III	185	162	23	529	3.265
IV	1315	1338	23	529	0.395

$$\chi^2 = \Sigma\left[\frac{(f_o - f_e)^2}{f_e}\right] = 9.796 + 1.186 + 3.265 + 0.395 = 14.642$$

$df = (2 - 1)(2 - 1) = 1$. P is less than .01

Calculated $\chi^2 = 14.64$

From Table E

Tabulated value of χ for 1 df at .05 level = 3.841

The five percent value of χ^2 for 1 df is 3.841, which is much less than the calculated value of χ^2. So in the light of this, conclusion is evident that the hypothesis is incorrect and inoculation and absence of attack from cholera are associated.

Conditions for the Validity of Chi-Square Test:

The Chi-square test statistic can be used if the following conditions are satisfied:

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

2. The sample observations should be independent. This implies that no individual item should be included twice or more in the sample.

3. The constraints on the cell frequencies, if any, should be linear (i.e., they should not involve square and higher powers of the frequencies) such as $\sum f_o = \sum f_e = N$.
4. No theoretical frequency should be small. Small is a relative term. Preferably each theoretical frequency should be larger than 10 but in any case not less than 5.

If any theoretical frequency is less than 5 then we cannot apply χ^2 -test as such. In that case we use the technique of "pooling" which consists in adding the frequencies which are less than 5 with the preceding or succeeding frequency (frequencies) so that the resulting sum is greater than 5 and adjust for the degrees of freedom accordingly.
5. The given distribution should not be replaced by relative frequencies or proportions but the data should be given in original units.

6. Yates' correction should be applied in special circumstances when df = 1 (i.e. in 2 x 2 tables) and when the cell entries are small.

7. χ^2-test is mostly used as a non-directional test (i.e. we make a two-tailed test.). However, there may be cases when χ^2 tests can be employed in making a one-tailed test.
In one-tailed test we double the P-value. For example with df = 1, the critical value of χ^2 at 05 level is 2.706 (2.706 is the value written under. 10 level) and the critical value of; χ^2 at .01 level is 5.412 (the value is written under the .02 level).

The Additive Property of Chi-Square Test:

χ^2 has a very useful property of addition. If a number of sample studies have been conducted in the same field then the results can be pooled together for obtaining an accurate idea about the real position.

Suppose ten experiments have been conducted to test whether a particular vaccine is effective against a particular disease. Now here we shall have ten different values of χ^2 and ten different values of df.

We can add the ten χ^2 to obtain one value and similarly ten values of df can also be added together. Thus, we shall have one value of χ^2 and one value of degrees of freedom. Now we can test the results of all these ten experiments combined together and find out the value of P.

Suppose five independent experiments have been conducted in a particular field. Suppose in each case there was one df and following values of χ^2 were obtained.

Experiment Number	Value of χ^2	df
1.	4.3	1
2.	5.7	1
3.	2.1	1
4.	3.9	1
5.	8.3	1

Now at 5% level of significance (or for P – .05) the value χ^2 for one df is 3.841. From the calculated values of χ^2 given above we notice that in only one ease i.e., experiment No. 3 the observed value of χ^2 is less than the tabulated value of 3.841.

It means that so far as this experiment is concerned the difference is insignificant but in the remaining four cases the calculated value of χ^2 is more than 3.841 and as such at 5% level of significance the difference between the expected and the actual frequencies is significant.

If we add all the values of χ^2 we get (4.3 + 5.7 + 2.1 + 3.9 + 8.3) or 24.3. The total of the degrees of freedom is 5. It means that the calculated value of χ^2 for 5 df is 24.3.

If we look in the table of χ^2 we shall find that at 5% level of significance for 5 df the value of χ^2 is 11.070. The calculated value of χ^2 which is 24.3 is much higher than the tabulated value and as such we can conclude that the difference between observed and expected frequencies is significant one.

Even if we take 1% level of significance (or P = .01) the table value of χ^2 is only 15.086. Thus the probability of getting a value of χ^2 equal to or more than 24.3 as a result of sampling fluctuations is much less than even .01 or in other words the difference is significant.

Applications of Chi-Test:

The applications of χ^2-test statistic can be discussed as stated below:

1. Testing the divergence of observed results from expected results when our expectations are based on the hypothesis of equal probability.

2. Chi-square test when expectations are based on normal distribution.

3. Chi-square test when our expectations are based on predetermined results.

4. Correction for discontinuity or Yates' correction in calculating χ^2.

5. Chi-square test of independence in contingency tables.

Uses of Chi-Square Test:

1. Although test is conducted in terms of frequencies it can be best viewed conceptually as a test about proportions.

2. χ^2 test is used in testing hypothesis and is not useful for estimation.

3. Chi-square test can be applied to complex contingency table with several classes.

4. Chi-square test has a very useful property i.e., 'the additive property'. If a number of sample studies are conducted in the same field, the results can be pooled together. This means that χ^2-values can