**Chi-Square Test: Meaning, Applications and Uses | Statistics**

**Meaning of Chi-Square Test:**

The Chi-square (χ2) test represents a useful method of comparing experimentally obtained results with those to be expected theoretically on some hypothesis.

Thus, Chi-square is a measure of actual divergence of the observed and expected frequencies. It is obvious that the importance of such a measure would be very great in sampling studies where we have invariably to study the divergence between theory and fact.

Chi-square as we have seen is a measure of divergence between the expected and observed frequencies and as such if there is no difference between expected and observed frequencies the value of Chi-square is 0.

If there is a difference between the observed and the expected frequencies, then the value of Chi-square would be more than 0. That is, the larger the Chi-square the greater the probability of a real divergence of experimentally observed from expected results.

If the calculated value of chi-square is very small as compared to its table value it indicates that the divergence between actual and expected frequencies is very little and consequently the fit is good. If, on the other hand, the calculated value of chi-square is very big as compared to its table value it indicates that the divergence between expected and observed frequencies is very great and consequently the fit is poor.

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

Thus in 2 x 2 table degrees of freedom are (2 – 1) (2 – 1) or 1. Similarly in 3 x 3 table, degrees of freedom are (3 – 1) (3 – 1) or 4 and in 3 x 4 table the degrees of freedom are (3 – 1) (4 – 1) or 6.

**Levels of Significance of Chi-Square Test:**

The calculated values of χ2 (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 χ2 are available at various probability levels. These levels are called levels of significance. Usually, the value of χ2 at .05 and .01 level of significance for the given degrees of freedom is seen from the tables.

If the calculated value of χ2 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 χ2 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 χ2-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 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 (χ2) is stated as follows:**



in which fo = frequency of occurrence of observed or experimentally determined facts fe = 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 χ2.

Several illustrations of the chi-square test will clarify the discussion given above. The differences of fo and fe 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).

1. **It was observed that 48 marked ‘F’, 24 ‘I’ and 24 ‘U’:**
2. (i) Test whether the observed results diverge significantly from the results to be expected if there are no preferences in the group.
3. (ii) Test the hypothesis that “there is no difference between preferences in the group”.
4. (iii) Interpret the findings.

**Solution:**

1. **Following steps may be followed for the computation of x2 and drawing the conclusions:**
2. **Step 1:**

Compute the expected frequencies (fe) 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.

Table

Description automatically generated

**Step 2:**

Compute the deviations (fo – fe) for each frequency. Each of these differences is squared and divided by its fe (256/32, 64/32 and 64/32).

**Add these values to compute:**

Diagram

Description automatically generated with low confidence

**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 χ2 for 2 df at certain level of significance, usually 5% or 1%.

With df = 2, the χ2 value to be significant at .01 level is 9.21 (Table E). The obtained χ2 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 favors the proposition.

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