

# Chi- square test

KUMARI SONA RANI  
ASSISTANT PROFESSOR  
DEPARTMENT OF BOTANY  
S.SINHA COLLEGE AURANGABAD

FOR M.Sc 2<sup>ND</sup> SEMESTER STUDENTS

A **chi-squared test** is basically a data analysis on the basis of observations of a random set of variables. Usually, it's a comparison of two statistical data sets. It is also represented as  $\chi^2$  test. This test was introduced by Karl Pearson in 1900 for categorical data analysis and distribution. So it was mentioned as Pearson's chi-squared test. With the assumption of the **null hypothesis** as true, this test is used to estimate how likely the observations that are made would be.

A hypothesis is a consideration, that a given condition or statement might be true, which we can test afterwards. Chi-squared tests are usually created from a sum of squared falsities or errors else via the sample variance

## Chi-Square Distribution

When we consider, the null speculation as true, the sampling distribution of the test statistic is called as **chi-squared distribution**. The chi-squared test helps to determine whether there is a notable difference between the normal frequencies and the observed frequencies in one or more classes or categories. It gives the probability of independent variables.

**Note:** Chi-squared test is applicable only for **categorical data**, such as men and women falling under the categories of Gender, Age, Height, etc.

Probability is all about chance or risk or uncertainty. It is the possibility of the outcome of the sample or the occurrence of an event. But when we talk about statistics, it is more about how we handle various data using different techniques. It helps to represent complicated data or bulk data in a very easy and understandable way. It describes the collection, analysis, interpretation, presentation, and organization of data. The concept of both **probability and statistics** is related to the chi-squared test.

## Chi-Square Test Formula

The chi-squared test is done to check if there is any difference between the observed value and expected value. The formula for chi-square can be written as;

$$\chi^2 = \sum \frac{(\text{Observed Value} - \text{Expected Value})^2}{\text{Expected Value}}$$

## Chi-Square Test of Independence

The chi-square test of independence also known as the chi-square test of association which is used to determine the association between the categorical variables. It is considered as a non-parametric test. It is mostly used to test statistical independence. The chi-square test of independence is not

appropriate when the categorical variables represent the pre-test and post-test observations. For this test, the data must meet the following requirements:

- Two categorical variables
- Relatively large sample size
- Categories of variables (two or more)
- Independence of observations