M.A. Economics

Semester – II Paper CC-09 (Statistical Methods)

Module V: **Chi-Square (χ^2) Test**

Dr. Sweta Sharan, Assistant Professor, Department of Economics, Magadh Mahila College, Patna University, Email:swetasharan87@gmail.com,Ph:9308489824

Introduction

The chi-square testis one of the most widely used non-parametric test in statistical works. It is denoted by Greek letter chi (χ^2). The test was first used by Karl Pearson in the year 1900. The quantity (χ^2) describes the magnitude of the discrepancy between theory and observation. The value of χ^2 is positive and its upper limit is infinity. Since χ^2 is a non-parametric test it involves no assumption about the original form of distribution from which the observation comes.

It is defined as: $\chi^2 = \sum \frac{(O-E)^2}{E}$

Where O refers to observed frequencies and E refers to the expected frequencies.

Expected frequency (E) is calculated using the formula

$$E = \frac{RT \times CT}{N}$$

Where, RT = row total for the row containing the cell
CT = column total for the column containing the cell
N = total number of observations

Degrees of Freedom

Degrees of freedom refer to number of classes to which the values can be assigned arbitrarily or at will without violating the restrictions or limitations placed. Symbolically it can be denoted as follows:

ϑ (nu) = n – k

where k refers to the number of independent constraints

Example: If we have to choose five numbers whose total is 100, we are free to choose any four numbers only, the fifth number is fixed by virtue of the total being 100 as the fifth number must be equal to 100 minus the total of four numbers selected. if the four numbers selected are 20,35,15 and10, then the fifth number must be 100 – (20+35+15+10) which is 20.

Properties of χ^2 test

- The sum of observed and expected frequency is always zero
- The χ^2 depends only on the set of observed and expected frequencies and not on degrees of freedom
- It is a limiting approximation of multinomial distribution
- χ^2 distribution can be applied to discrete random variables also in addition to continuous variable

The Chi-Square distribution

For large sample sizes, the sampling probability distribution of χ^2can be closely approximated by a continuous curve known as the Chi-square distribution. The probability function of χ^2 is given by

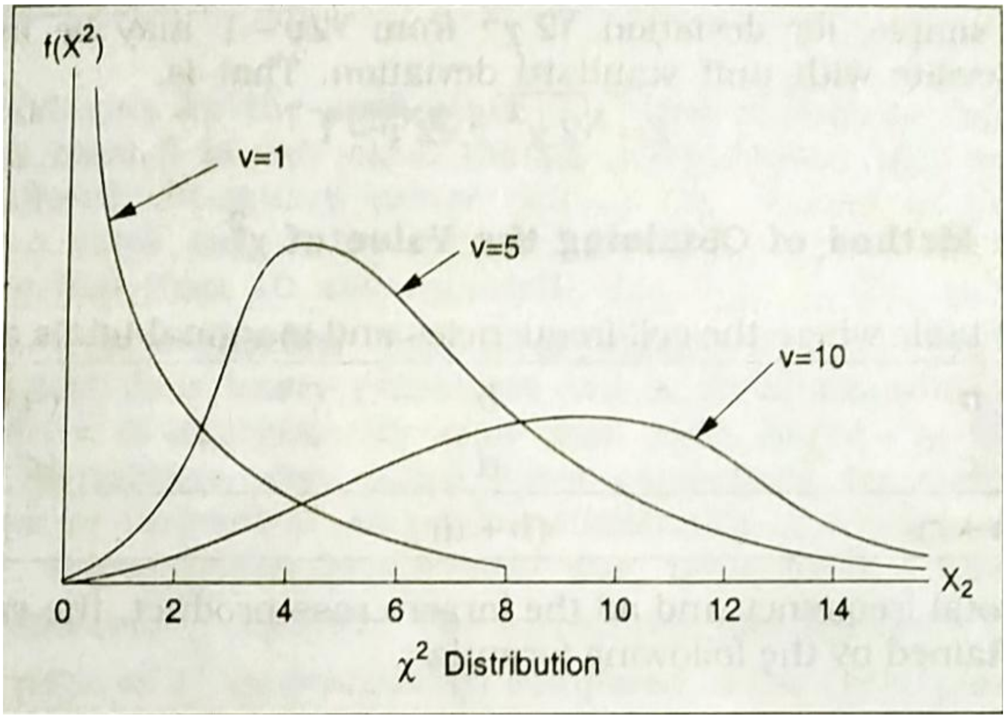
$$\mathbf{f(\chi^2) = C\ (\chi^2)^{(\vartheta/2 - 1)}\ x\ e^{-\chi 2/2}}$$

where e = 2.71828

ϑ = number of degrees of freedom

C = a constant depending only on ϑ

The χ^2 distribution has only one parameter i.e. degrees of freedom. For very small degrees of freedom, the chi-square is severely skewed to the right. As the number of degrees of freedom increases, the curve rapidly becomes more symmetrical. For larger values of ϑ the chi-square distribution is closely approximated by normal curve. Total area under the curve in each chi-square distribution is equal 1.0.



Uses of Chi-Square Test

i) χ^2 test as a test of independence: With the help of χ^2 test it can be found out whether two or more attributes are associated or not.

ii) χ^2 test as a test of goodness of fit: χ^2 test is very popularly known as test of goodness of fit because it enables in ascertaining how appropriately the theoretical distributions like Binomial, Poisson, Normal etc fits the empirical distributions.

iii) χ^2 test as a test of homogeneity: This test is an extension of test of independence. This test helps in determining whether two or more independent random samples are drawn from same population or from different population.

An example:

1. In an anti-malarial campaign in certain area, quinine was administered to 812 persons out of a total population of 3248. The number of fever cases is shown below:

Treatment	Fever	No Fever	Total
Quinine	20	792	812
No quinine	220	2216	2436

Total	240	3008	3248

Discuss the usefulness of quinine in checking malaria.

Solution: Let us take the hypothesis that quinine is not effective in checking malaria.

Applying χ^2 test:

Expectation (E_{11}) column wise (first column) first element of the above table = $\frac{240 \times 812}{3248}$ = 60

Expectation (E_{21}) column wise (first column) second element of the above table = $\frac{240 \times 2436}{3248}$ = 180

Expectation (E_{31}) for the column wise (first column) third element of the above table = $\frac{240 \times 3248}{3248}$ = 240

Similarly expectations will be calculated for second column

The bale of expected frequency shall be

60	752	812
180	2256	2436
240	3008	3248

Using the formula for chi-square test

O	E	$(0 – E)^2$	$(O – E)^2/E$
20	60	1600	26.667
220	182	1600	8.889
792	752	1600	2.128
2216	2256	1600	0.709
			38.393

$$\chi^2 = \sum \frac{(O-E)^2}{E}$$

and degrees of freedom = (r – 1)(c – 1) = (2-1)(2-1) = 1

Table value of χ^2 for degrees of freedom 1 at 5% level of significance is 3.84. Since the calculated value is greater than table value so the hypothesis is rejected. Hence we conclude that quinine is effective in checking malaria.