CHI-SQUARE (χ^2) TEST — DETAILED NOTES

1. What is Chi-Square (χ^2) Test?

Chi-Square test is a non■parametric statistical test used to examine whether there is a significant difference between observed frequencies and expected frequencies in categorical data.

It does not compare means. Instead, it works on frequencies (counts) and checks whether the observed data fits a theoretical or expected distribution.

2. When to Use Chi-Square Test

• When data is categorical (qualitative).

• When observations are given as frequencies or counts.

• When a theoretical or expected distribution is known.

• When sample size is sufficiently large.

3. Types of Chi-Square Test

(a) Goodness of Fit Test

Used to test whether observed frequencies follow a specified theoretical distribution.

(b) Test of Independence

Used to test whether two categorical variables are independent of each other.

4. Hypotheses

Null Hypothesis (H■): There is no significant difference between observed and expected frequencies.

Alternative Hypothesis (H■): There is a significant difference between observed and expected frequencies.

5. Formula

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

Where:
O = Observed frequency
E = Expected frequency

6. Conditions / Assumptions

• Observations must be independent.

• Expected frequency in each cell should be at least 5.

• Data should be in the form of counts, not percentages.

7. Degrees of Freedom

For Goodness of Fit test:

DOF = k – 1

Where k = number of categories.

8. Solved Example (Goodness of Fit)

A population is distributed according to the following weight categories:

<50 kg : 20% , 50–75 kg : 30% , >75 kg : 50%

A random sample of 500 people gives the following observed frequencies:

<50 kg: 140 , 50–75 kg: 160 , >75 kg: 200

Test whether the population distribution has changed at 5% level of significance.

Solution

Step 1: Hypotheses

H■: Observed frequencies follow the given population distribution.

H■: Observed frequencies do not follow the given distribution.

Step 2: Expected Frequencies

Expected frequency = (Percentage × Total sample size)

<50 kg: 0.20 × 500 = 100

50–75 kg: 0.30 × 500 = 150

>75 kg: 0.50 × 500 = 250

Step 3: Chi-Square Table Calculation

<50 kg: $(140–100)^2 / 100 = 16.00$
50–75 kg: $(160–150)^2 / 150 = 0.67$
>75 kg: $(200–250)^2 / 250 = 10.00$

$\chi^2 = 16.00 + 0.67 + 10.00 = 26.67$

Step 4: Degrees of Freedom

DOF = k – 1 = 3 – 1 = 2

Step 5: Decision

Critical χ^2 value at 5% significance and 2 DOF = 5.991

Since χ^2 calculated (26.67) > χ^2 critical (5.991), reject H■.

Conclusion

There is sufficient evidence at 5% level of significance to conclude that the population distribution has changed.

Exam Conclusion Line

Since χ^2 calculated is greater than χ^2 tabulated value, the null hypothesis is rejected.