Chi-Square Test for Categorical Data

Learning Objectives

- - Understand categorical data
- - Explain Chi-square test
- - Perform Chi-square test
- - Interpret results

What is Categorical Data?

- - Data divided into categories
- - Non-numerical values
- - Examples: gender, blood group, disease status

Statistical Tests for Categorical Data

- - Chi-square test
- - Fisher's exact test
- - Used for frequency data

What is the Chi-Square Test?

- - Non-parametric test
- - Compares observed vs expected frequencies
- - Tests association between variables

Types of Chi-Square Tests

- - Chi-square test of independence
- - Chi-square goodness-of-fit test

Chi-Square Test of Independence

- - Determines relationship between two categorical variables
- - Uses contingency tables

Observed vs Expected Frequencies

- Observed (O): actual data
- Expected (E): theoretical values
- E = (Row total × Column total) / Grand total

Chi-Square Formula

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

- Calculated for each cell

Degrees of Freedom

- df = (rows − 1) × (columns − 1)
- Used to find critical value

Assumptions of Chi-Square Test

- - Random sample
- - Independent observations
- - Expected frequency ≥ 5

Example (Medical Research)

- Disease status vs Gender
- Is there an association?

Decision Rule

- - Compare χ^2 calculated with χ^2 critical
- - Or use p-value
- $p < 0.05 \rightarrow$ significant

Applications in Biology & Medicine

- - Epidemiology
- - Clinical studies
- - Genetics

Summary

- - Used for categorical data
- - Tests association
- - Simple and widely used