ANOVA (Analysis of Variance)

THEORY

ANOVA is a statistical technique used to test whether the means of three or more populations are equal.

It compares variation between groups with variation within groups.

If the variation between groups is significantly larger than the variation within groups, we conclude that at least one group mean is different.

Variables in ANOVA

Factor (Independent Variable): A categorical variable that forms different groups.

Levels: Different categories of the factor.

Dependent Variable: A numerical variable whose mean is compared across groups.

Assumptions of ANOVA

1. Sampling distribution of the mean is normal.

2. Absence of extreme outliers.

3. Homogeneity of variance (equal population variances).

4. Observations are independent and randomly selected.

Hypotheses

Null Hypothesis (H■): μ■ = μ■ = μ■ = ... = μ■

Alternative Hypothesis (H■): At least one population mean is different.

FORMULAS WITH MEANING

Group Mean (X■■)

X■■ = (Sum of observations in group i) / (Number of observations in group i)

Meaning: Average value of each group.

Grand Mean (X■)

X■ = ΣX / N

Meaning: Overall average of all observations combined.

Between-Group Sum of Squares (SSB)

$SSB = \Sigma\ n_{\blacksquare}\ (X_{\blacksquare\blacksquare} - X_{\blacksquare})^2$

Meaning: Measures how far group means are from the grand mean (variation between groups).

Within-Group Sum of Squares (SSW)

$SSW = \Sigma\Sigma\ (X_{\blacksquare\blacksquare} - X_{\blacksquare\blacksquare})^2$

Meaning: Measures natural variation within each group.

Total Sum of Squares (SST)

$SST = \Sigma\ (X_{\blacksquare\blacksquare} - X_{\blacksquare})^2$

Meaning: Total variation in the data.

Relationship: SST = SSB + SSW

Degrees of Freedom

Between groups: $dfB = k - 1$

Within groups: $dfW = N - k$

Mean Squares

$MSB = SSB / dfB \rightarrow$ Between-group variance

$MSW = SSW / dfW \rightarrow$ Within-group variance

F-Statistic

$F = MSB / MSW$

Meaning: Ratio of signal (between groups) to noise (within groups).

PROBLEM (ONE-WAY ANOVA)

Marks obtained by students using three teaching methods are given below:

Method A: 50, 55, 60

Method B: 60, 65, 70

Method C: 70, 75, 80

Test at 5% level of significance whether the mean marks of the three teaching methods are equal.

SOLUTION

Step 1: H■: μA = μB = μC

H■: At least one mean is different.

Step 2: Group Means

X■A = 55, X■B = 65, X■C = 75

Step 3: Grand Mean

X■ = 65

Step 4: Between-Group Sum of Squares

SSB = 3[(55–65)² + (65–65)² + (75–65)²] = 600

Step 5: Within-Group Sum of Squares

Group A: (50–55)² + (55–55)² + (60–55)² = 50

Group B: (60–65)² + (65–65)² + (70–65)² = 50

Group C: (70–75)² + (75–75)² + (80–75)² = 50

SSW = 150

Step 6: Degrees of Freedom

dfB = 2, dfW = 6

Step 7: Mean Squares

MSB = 300, MSW = 25

Step 8: F-Statistic

F = 300 / 25 = 12

Step 9: Conclusion

Since calculated F is greater than critical F, reject H■.

Conclusion: At least one teaching method has a different mean score.