LEARNING OBJECTIVES

- Understand the necessity of Analysis of Variance (ANOVA)
- Understand the assumptions of ANOVA
- Understand one way ANOVA and F–statistic

AMITY UNIVERSITY ONLINE
CAREERS OF TOMORROW

ANOVA

- ANOVA provides a method to compare the population means of more than two population/groups simultaneously.

- Analysing the variance between the means of each group and also variance within the group

- Hypothesis statement:
  - *H_0 = no differences between groups*
  - *H_1 = differences between groups*

(Note: We usually refer to the sub-populations as "groups" when doing ANOVA.)

AMITY UNIVERSITY ONLINE
CAREERS OF TOMORROW

GROUPS AND FACTOR LEVELS

- **Groups** are a part of the population also referred as sub population.
- **Factor:** These are quantities that affect the outcome, similar to independent variables. Factors are also referred to as "Treatment" due to origin of ANOVA in testing of fertilizer treatment on the yield.
- Different values of the factor constitute the levels.
- For example,

  Assuming we are studying the effect of discount rate on sales, discount rate is factor and the levels could be the various discount rates, namely 5%, 10%, 15%, and so on.

AMITY UNIVERSITY ONLINE
CAREERS OF TOMORROW

ASSUMPTIONS OF ANOVA

- Each group is approximately normal
- Check this by looking at histograms and/or normal quantile plots, or use assumptions.
- Can handle some non-normality, but not severe outliers.
- Standard deviations of each group are approximately equal
- Rule of thumb:

  Ratio of largest to smallest sample standard deviation must be less than 2:1.

AMITY UNIVERSITY
AMITY UNIVERSITY ONLINE
CAREERS OF TOMORROW

ANOVA TYPES

- One-way ANOVA
  - One factor with more than 2 levels
- Factorial ANOVAs
  - More than 1 factor
- Mixed design ANOVAs
  - Some factors independent, others related

AMITY UNIVERSITY ONLINE
CAREERS OF TOMORROW

ONE WAY ANALYSIS OF VARIANCE (ANOVA)

- Compares two types of variation to test equality of means
- Ratio of variances is comparison basis
- If treatment variation is significantly greater than random variation, then means are not equal
- Variation measures are obtained by 'partitioning' total variation

AMITY UNIVERSITY ONLINE
CAREERS OF TOMORROW

ONE WAY ANALYSIS OF VARIANCE (ANOVA) : Procedure

H0 = μ1 = μ2 = μ3 ………… = μk

HA = not all μ values are equal

The following measures are required to be computed, namely variation within the groups and between groups:

Step 1:

Sum of Squares of Total Variation (SST)

$$SST = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (Y_{ij} - \mu^2)$$

K: Number of groups

ni: Number of observations in a group I

Yij: Observation of j in group I

μi: Mean of group I

μ: Overall (Total) Mean

Mean Square Total (MST) Variation=SST/n-1, where n-1 is the degrees of freedom (df).

AMITY UNIVERSITY ONLINE
CAREERS OF TOMORROW

ONE WAY ANALYSIS OF VARIANCE (ANOVA) : Procedure

Step 1:

Sum of squares between groups (SSB):

$$SSB = \sum_{i=1}^{k} n_i * (\mu_i - \mu)^2$$

Mean Square between (MSB) Variation

$$MSB = \frac{SSB}{k-1}$$

Where k-1 is the degrees of freedom (df)

Step 3:

Sum of squares within groups (SSW):

$$SSW = \sum_{i=1}^{k} \sum_{j=1}^{n} (Y_{ij} - \mu^2)$$

Mean Square within (MSW) Variation:

$$MSW = \frac{SSW}{n-k}$$

SST = SSW + MSW

AMITY UNIVERSITY ONLINE
CAREERS OF TOMORROW

F-TEST

Statistic for testing null hypothesis is given by the F–distribution:

$$F = \frac{\frac{SSB}{k-1}}{\frac{SSW}{n-k}} = \frac{MSB}{MSW}$$

F-Test Result	Decision
Test Statistic (F) greater than Critical Value Fcr	Reject null hypothesis
Test Statistic (F) less than Critical Value Fcr	Failure to Reject null Hypothesis
p-value less than significance factor(α)	Reject null hypothesis
p-value greater than significance factor(α)	Failure to Reject null hypothesis

Test Statistic is a one tailed test (right tailed) as we need to find out whether the variation between groups is greater than variation within the group.

AMITY UNIVERSITY ONLINE
CAREERS OF TOMORROW

SUMMARY

- Number of factors decide the type of ANOVA test to be adopted for the experiment
- ANOVA will only tell whether there is a significant difference and gives no information on which mean(s) are different.
- To gain further knowledge of difference in means pairwise comparisons of the means has to be done
- Pairwise testing are likely to increase the Type I errors

AMITY UNIVERSITY ONLINE
CAREERS OF TOMORROW