Learning Objectives

By the end of lecture, students should be able to:

1. Apply ANOVA for comparison of means in more than 2 groups
2. Compute one way and two way ANOVA for a given data set
3. Interpret the results of ANOVA

ANOVA (Analysis of Variance)

- Is a statistical technique specially designed to compare more than 2 means
- Types
  - One way ANOVA
  - Two way ANOVA

Difference b/w one way & two way

- **One way ANOVA :**

  In one way ANOVA, we take into account only one variable, say the effect of different types of fertilizers on yield.

- **Two way ANOVA:** you could use a **two-way ANOVA** to understand whether there is an interaction between gender and educational level on test anxiety amongst university students, where gender (males/females) and education level (undergraduate/postgraduate) are your independent variables, and test anxiety is your dependent variable.

Data required

One way ANOVA or single factor ANOVA:

- Determines means of

≥ 3 independent groups

significantly different from one another.

- Only 1 independent variable (factor/grouping variable) with ≥3 levels
- Grouping variable- nominal
- Outcome variable- interval or ratio

Post hoc tests help determine where difference exist

Steps ANOVA

1.Define null & alternative hypotheses

2. State Alpha
3. Calculate degrees of Freedom
4. State decision rule

5. Calculate test statistic
- Calculate variance between samples
- Calculate variance within the samples
- Calculate ratio F
- If F is significant, perform *post hoc* test

6. State Results & conclusion

Calculate test statistic

- calculate variance between samples
- calculate variance within samples

Calculating variance between samples

1. Calculate the **mean** of each sample.
2. Calculate the **Grand average**
3. Take the difference between means of various samples & grand average.
4. Square these deviations & obtain total which will give sum of squares between samples **(SSC)**
5. Divide the total obtained in step 4 by the degrees of freedom to calculate the mean sum of square between samples **(MSC)**.

Calculating Variance within Samples

1. Calculate **mean** value of each sample
2. Take the deviations of the various items in a sample from the mean values of the respective samples.
3. Square these deviations & obtain total which gives the sum of square within the samples **(SSE)**
4. Divide the total obtained in 3rd step by the degrees of freedom to calculate the mean sum of squares within samples **(MSE).**

The mean sum of squares

Calculation of **MSC**- **M**ean sum of **S**quares between samples

$$MSC = \frac{SSC}{k-1}$$

Calculation of **MSE** Mean Sum Of Squares within samples

$$MSE = \frac{SSE}{n-k}$$

k= No of Samples, n= Total No of observations

Calculation of F statistic

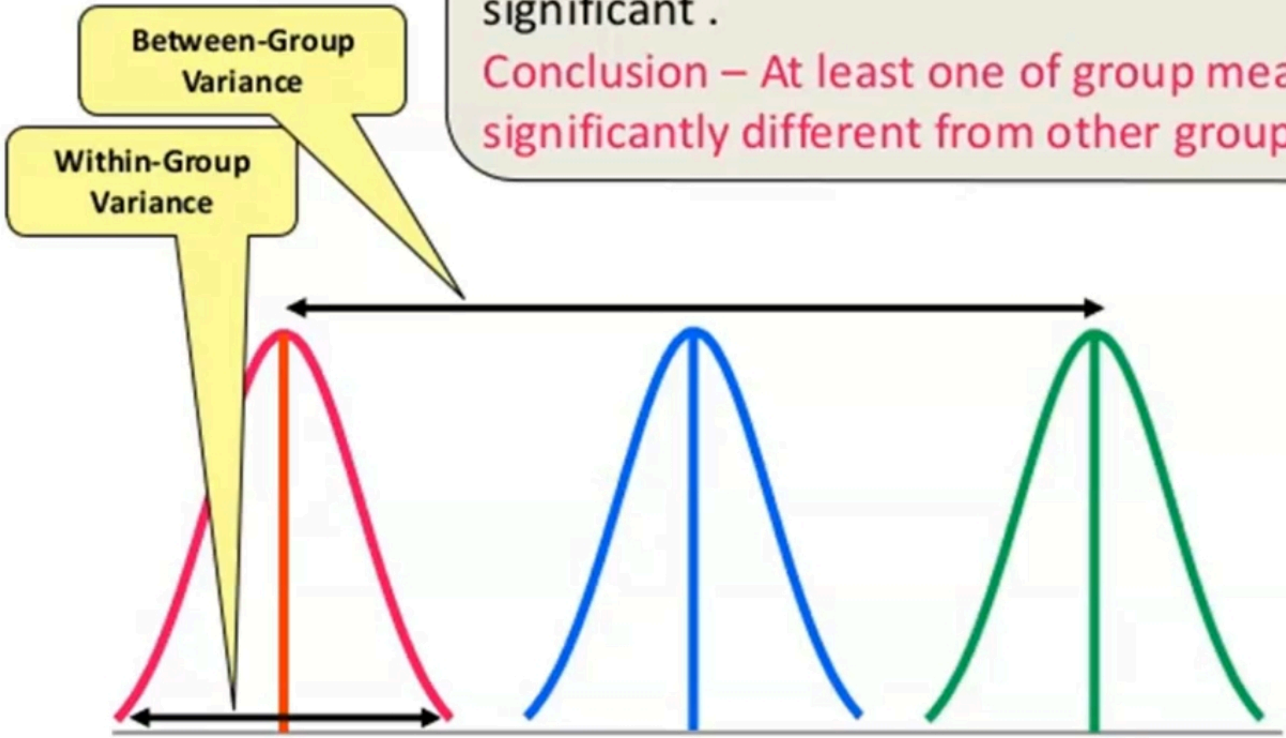
$$F = \frac{Variability\ between\ groups}{Variability\ within\ groups}$$

$$\text{F- statistic} = \frac{MSC}{MSE}$$

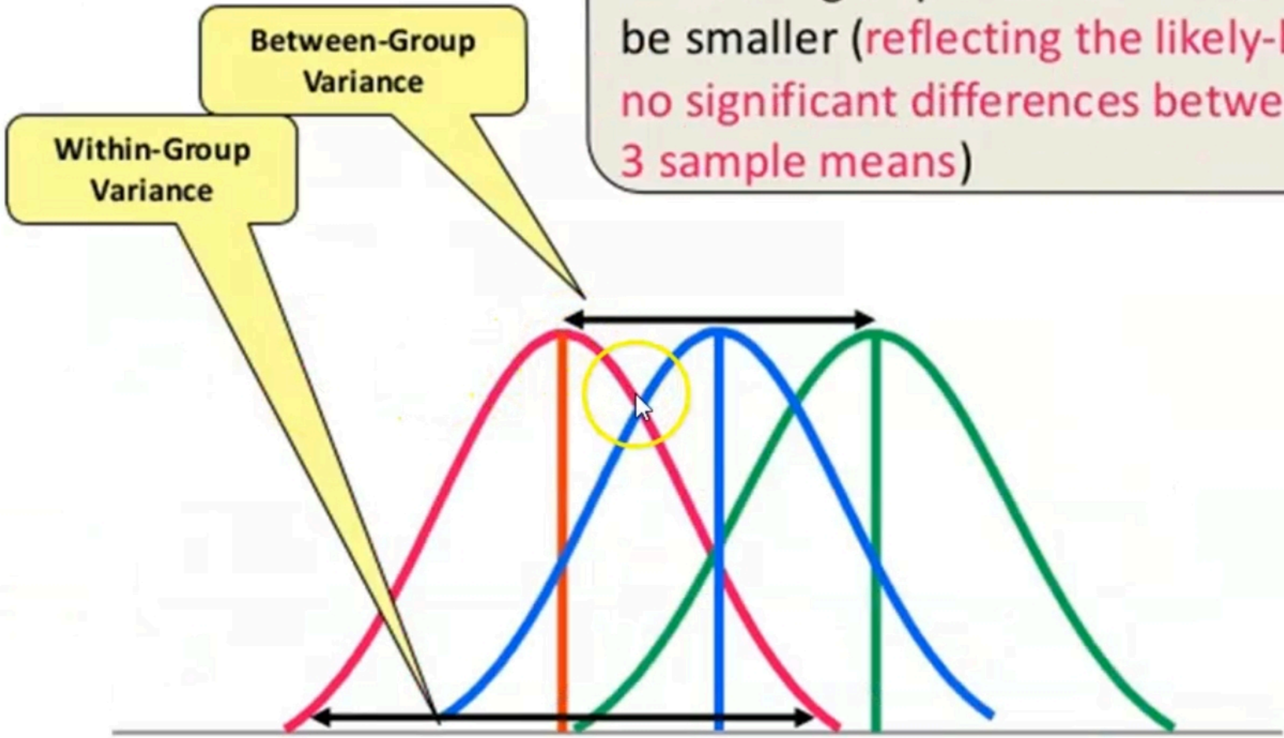
Compare the F-statistic value with F(critical) value which is obtained by looking for it in F distribution tables against degrees of freedom. The calculated value of F > table value H0 is rejected

Between-group variance is large relative to the within-group variance, so F statistic will be larger & > critical value, therefore statistically significant .

Conclusion – At least one of group means is significantly different from other group means



Within-group variance is larger, and the between-group variance smaller, so F will be smaller (reflecting the likely-hood of no significant differences between these 3 sample means)



Post-hoc Tests

- Used to determine which mean or group of means is/are significantly different from the others (**significant *F***)
- Depending upon research design & research question:

✓ **Bonferroni** (more powerful)

Only some pairs of sample means are to be tested

Desired alpha level is divided by no. of comparisons

✓ **Tukey's HSD Procedure**

when all pairs of sample means are to be tested

✓ **Scheffe's Procedure** (when sample sizes are unequal)

Example- one way ANOVA

Example: 3 samples obtained from normal populations with equal variances. Test the hypothesis that sample means are equal

8	7	12
10	5	9
7	10	13
14	9	12
11	9	14

1.Null hypothesis –
No significant difference in the means of 3 samples

2. State Alpha i.e 0.05

3. Calculate degrees of Freedom
k-1 & n-k = 2 & 12

4. State decision rule
Table value of F at 5% level of significance for d.f 2 & 12 is 3.88
The calculated value of F > 3.88 ,H0 will be rejected

5. Calculate test statistic

Variance BETWEEN samples (M1=10, M2=8,M3=12)

Sum of squares between samples (SSC) =

n1 (**M1** – Grand avg)^2 + n2 (**M2**– Grand avg)^2 + n3(**M3**– Grand avg)^2

$$5\,(10 - 10)^2 + 5\,(8 - 10)^2 + 5\,(12 - 10)^2 = 40$$

Calculation of **M**ean sum of **S**quares between samples (**MSC**)

$$MSC = \frac{SSC}{k-1} = \frac{40}{2} = 20$$

k= No of Samples, n= Total No of observations

Variance WITH IN samples ($M1=10, M2=8, M3=12$)

X1	$(X1 - M1)^2$	X2	$(X2 - M2)^2$	X3	$(X3 - M3)^2$
8	4	7	1	12	0
10	0	5	9	9	9
7	9	10	4	13	1
14	16	9	1	12	0
11	1	9	1	14	4
	30		16		14

Sum of squares within samples (SSE) = 30 + 16 +14 = 60

Calculation of Mean Sum Of Squares within samples **(MSE)**

$$MSE = \frac{SSE}{n-k} = \frac{60}{12} = 5$$

Calculation of ratio F

$$F = \frac{\text{Variability between groups}}{\text{Variability within groups}}$$

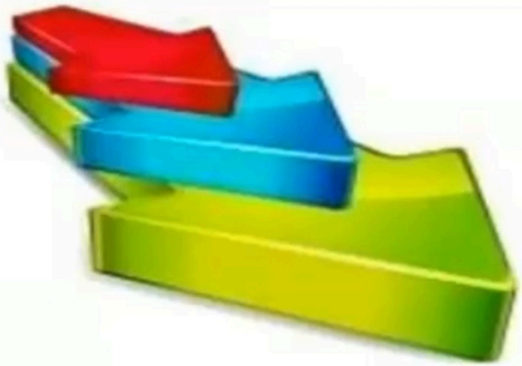
$$\text{F- statistic} = \frac{MSC}{MSE} = 20/5 = 4$$

The Table value of F at 5% level of significance for d.f 2 & 12 is 3.88
The calculated value of F > table value
H0 is rejected. Hence there is significant difference in sample means

Two Way ANOVA

Data required

- When 2 independent variables (Nominal/categorical) have an effect on one dependent variable (ordinal or ratio measurement scale)



- Compares relative influences on Dependent Variable
- Examine interactions between independent variables
- Just as we had Sums of Squares and Mean Squares in **One-way ANOVA**, we have the same in **Two-way ANOVA.**

Two way ANOVA

Include tests of three null hypotheses:

1) Means of observations grouped by one factor are same;

2) Means of observations grouped by the other factor are the same; and

3) There is no interaction between the two factors. The interaction test tells whether the effects of one factor depend on the other factor

Example-

we have test score of boys & girls in age group of 10 yr,11yr & 12 yr. If we want to study the effect of gender & age on score.

Two independent factors- Gender, Age

Dependent factor - Test score

Calculate Degrees of Freedom for

- D.f between samples = K-1
- D.f within samples = n- k
- D.f subjects=r -1
- D.f error= d.f within- d.f subjects
- D.f total = n-1

State decision rule

If calculated value of F >table value of F, reject Ho

Calculate test statistic (f= MS bw/ MS error)

	SS	DF	MS	F
Between				
Within -subjects - error				
Total				

State Results & conclusion