Chapter 11

Analysis of Variance

Comparing the golf clubs

- We would like to test if three different golf clubs yield different distances.
- We can randomly select five yield measurements from trials on an automated driving machine for each club.
- At the 0.05 significance level, is there a difference in mean distance?

Comparing the golf clubs

- Can we make use of any of those Hypothesis Testing tools that we have in our tool box?

Comparing the golf clubs

- Can we make use of any of those Hypothesis Testing tools that we have in our tool box?
- No we cannot! Why? Because those test for at most two population means… not three!

Comparing the golf clubs

- What should we do, then?
- We can make use of the, so called, Analysis of Variance to test for multiple population means.

The General ANOVA Setting

- You may control one or more **factors of interest**
- The *factor* is the basis of ANOVA analysis
  - Example: How much of a *factor* is the golf club in determining the distance?

The General ANOVA Setting

- You may control one or more **factors of interest**
  - Each **factor** contains two or more **levels** (e.g. categorical), which produce different **groups**
    - Example: Three different brand of golf club.

The General ANOVA Setting

- You may control one or more **factors of interest**
  - Each **factor** contains two or more **levels** (e.g. categorical), which produce different **groups**

- Confused?!? Well… think of each group as a sample from a different population

The General ANOVA Setting

- You may control one or more **factors of interest**
  - Each **factor** contains two or more **levels** (e.g. categorical), which produce different **groups**

  - Confused?!? Well… think of each group as a sample from a different population
  - Then, you may test the effects on the dependent variable by **comparing the groups**

Completely Randomized Design

- Completely randomized design is **an experiment with only one factor** (when we only change the golf club, and keep everything else the same)
- We use one-factor analysis of variance (ANOVA)

Completely Randomized Design

- Completely randomized design is **an experiment with only one factor** (when we only change the golf club, and keep everything else the same)
- We use one-factor analysis of variance (ANOVA)

Though ANOVA literally analyze **VARIATIONS**, the purpose of ANOVA is to reach conclusions about possible differences among the **MEANS** of each group.
We will make use of an F test to say something about population means!

One-way ANOVA

One-Way ANOVA

- Goal: Evaluate the difference among the means of three or more groups
- Assumptions
  - Populations are normally distributed
  - Populations have equal variances
  - Samples are randomly and independently drawn

Hypotheses of One-Way ANOVA

$$H_0 : \mu_1 = \mu_2 = \mu_3 = \cdots = \mu_c$$

- All population means are equal
- i.e., no factor effects (no variation in means among groups)

Hypotheses of One-Way ANOVA

$$H_0 : \mu_1 = \mu_2 = \mu_3 = \cdots = \mu_c$$

- All population means are equal
- i.e., no factor effects (no variation in means among groups)

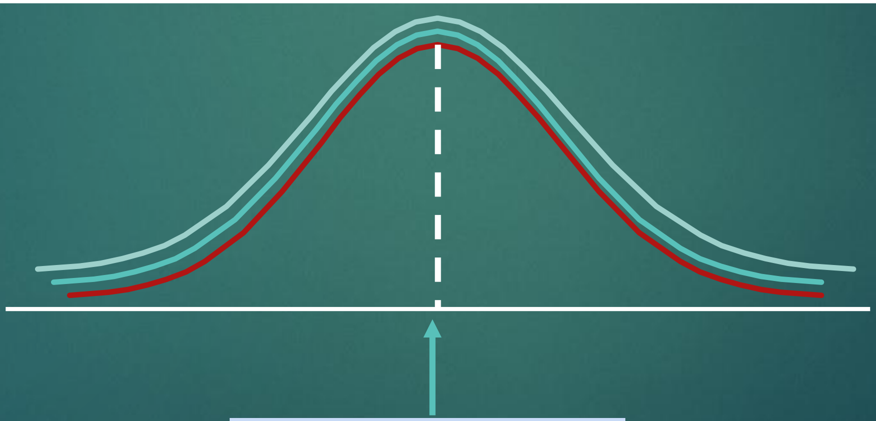
$$H_1 : \text{Not all of the population means are equal}$$

- At least one population mean is different
- i.e., there is a factor effect
- Does not mean that all population means are different (some pairs may be the same)

One-Way ANOVA

$$H_0 : \mu_1 = \mu_2 = \mu_3 = \cdots = \mu_c$$

When The Null Hypothesis is True
All Means are the same:
(No Factor Effect)



$$\mu_1 = \mu_2 = \mu_3$$

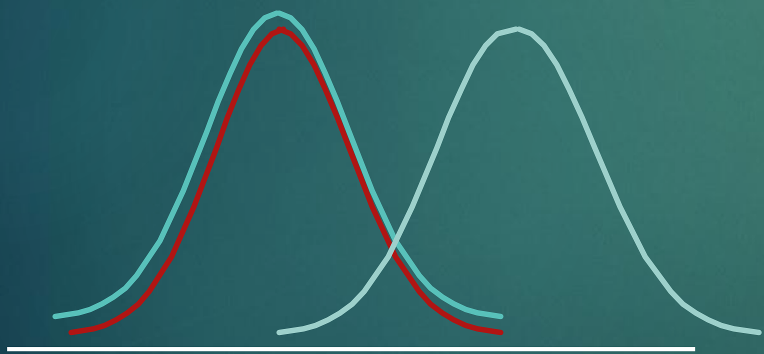
One-Way ANOVA

(continued)

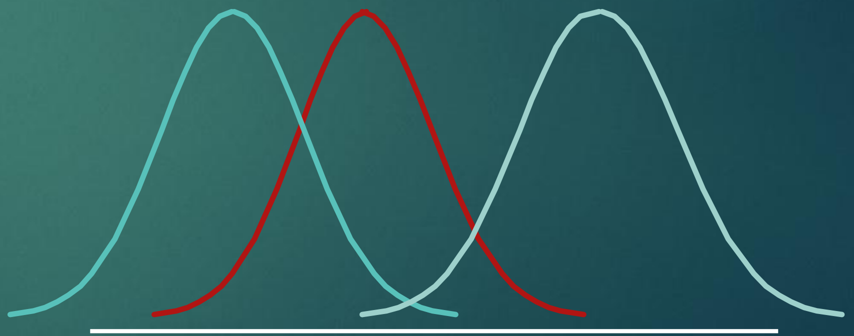
DCOVA

H_1 : Not all μ_j are equal

When The Null Hypothesis is rejected
At least one of the means is different
(Factor Effects are evident)



or



$\mu_1 = \mu_2 \neq \mu_3$

$\mu_1 \neq \mu_2 \neq \mu_3$

Partitioning the Variation

- Total variation can be split into two parts:

$$\mathbf{SST} = \mathrm{SSA} + \mathrm{SSW}$$

SST: Total Sum of Squares
(Total variation)
SSA: Sum of Squares Among Groups
(Among-group variation)
SSW = Sum of Squares Within Groups
(Within-group variation)

Partitioning the Variation

- Total variation can be split into two parts:

SST = SSA + SSW

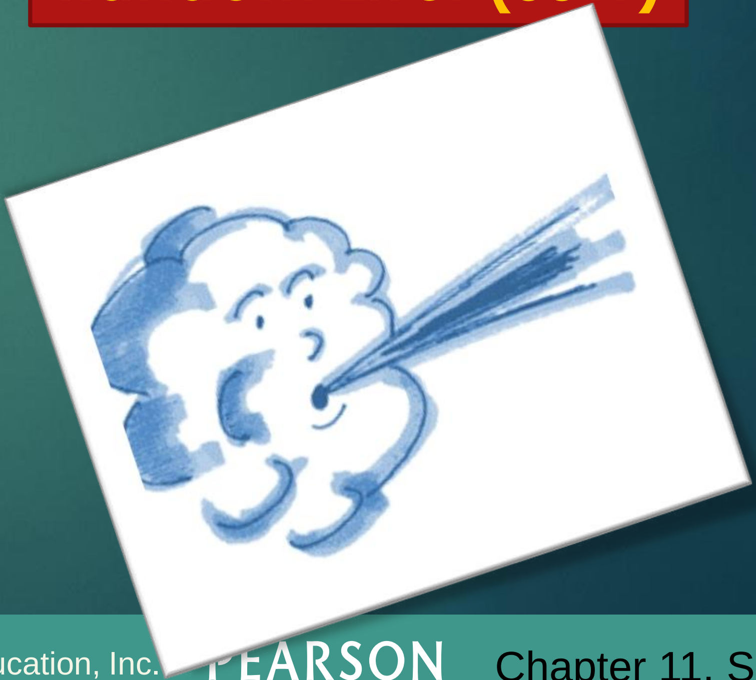
The aggregate variation (SST) can be partitioned into:

- The variations that are due to **differences among groups** (SSA: Differences in the distance caused by the choice of a particular golf club)
- The variations that are due to **random differences within a group** (SSW: Differences in the distance caused by random winds against the golf balls, given the choice of the golf club)

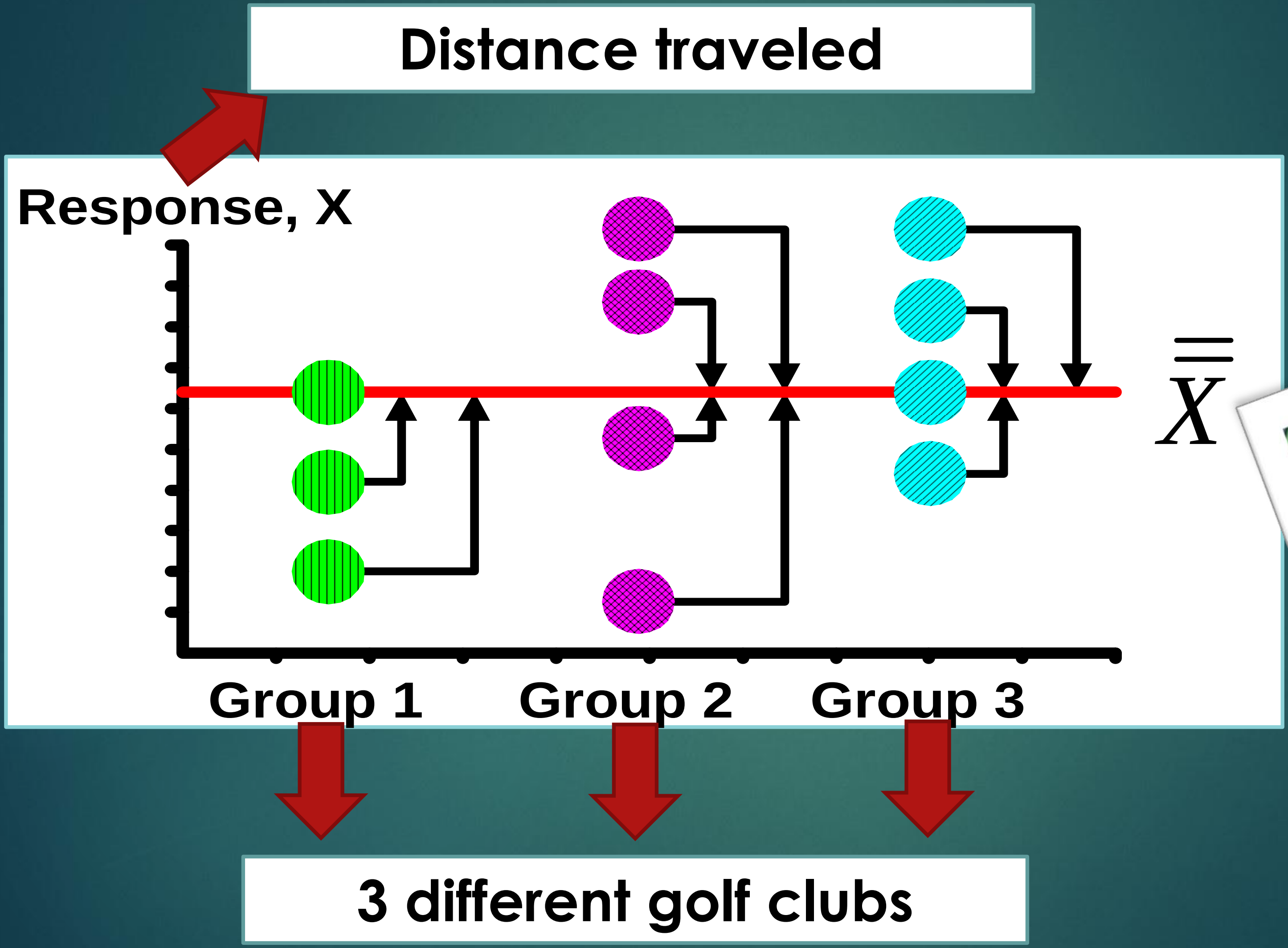
Partition of Total Variation

Total Variation (SST)

= **Variation Due to Factor (SSA)** + **Variation Due to Random Error (SSW)**

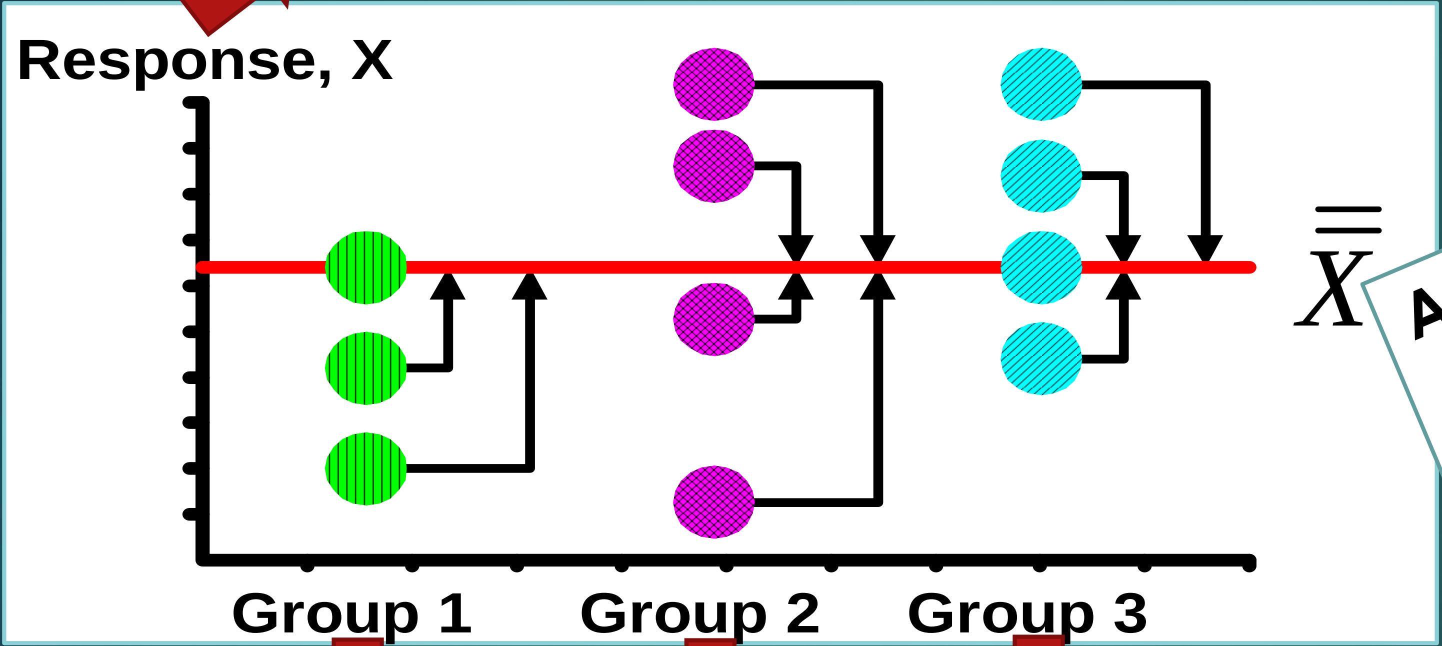


Total Variation



Total Variation

Distance traveled

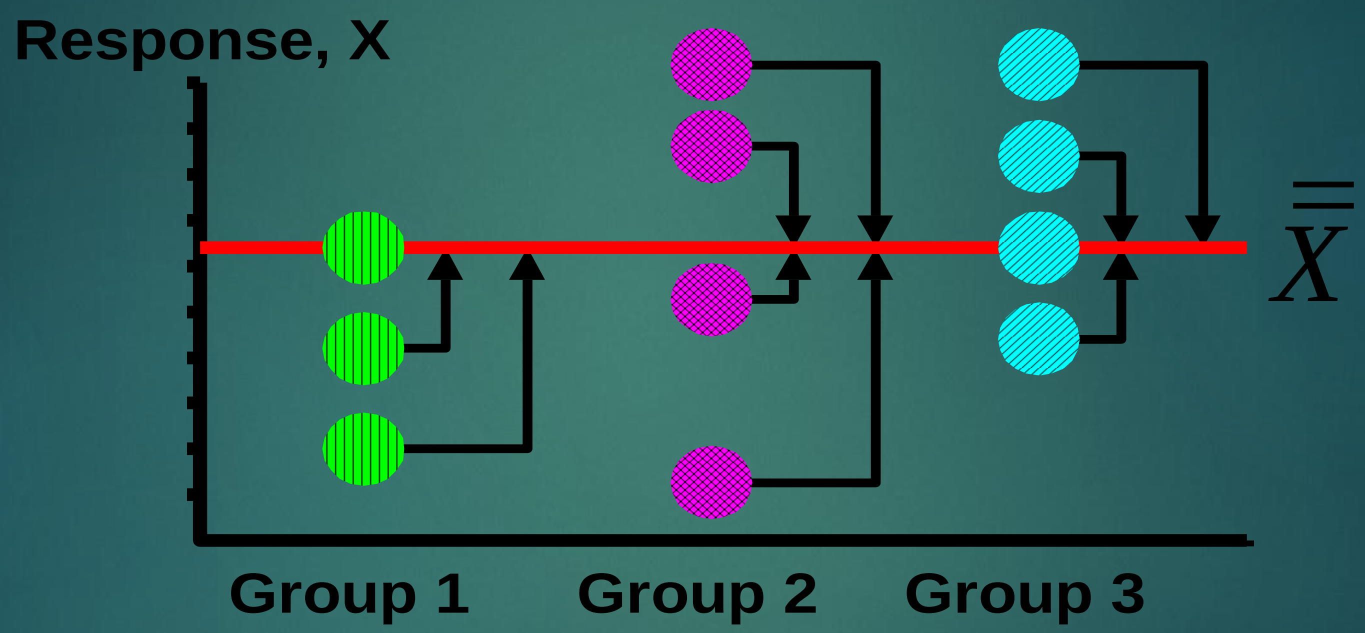


Average of distances traveled

3 different golf clubs

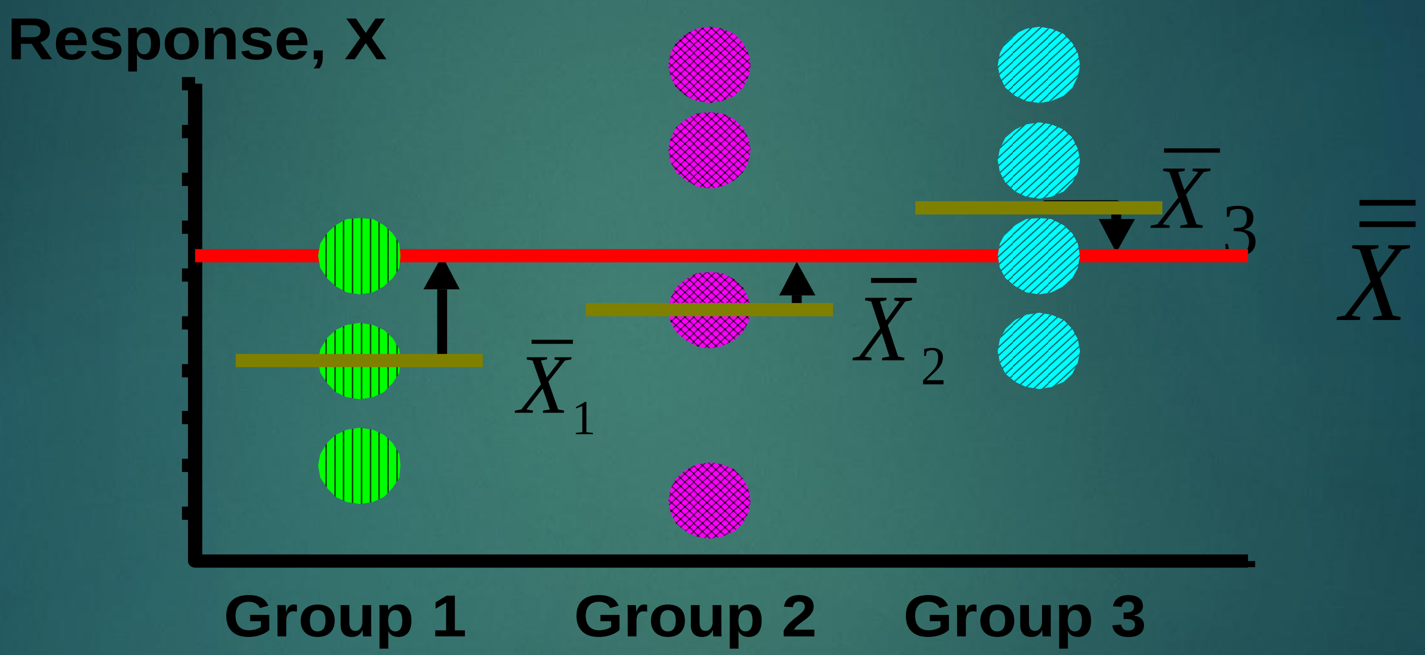
Total Variation,

Caused by **differences among groups** AND by **random errors**



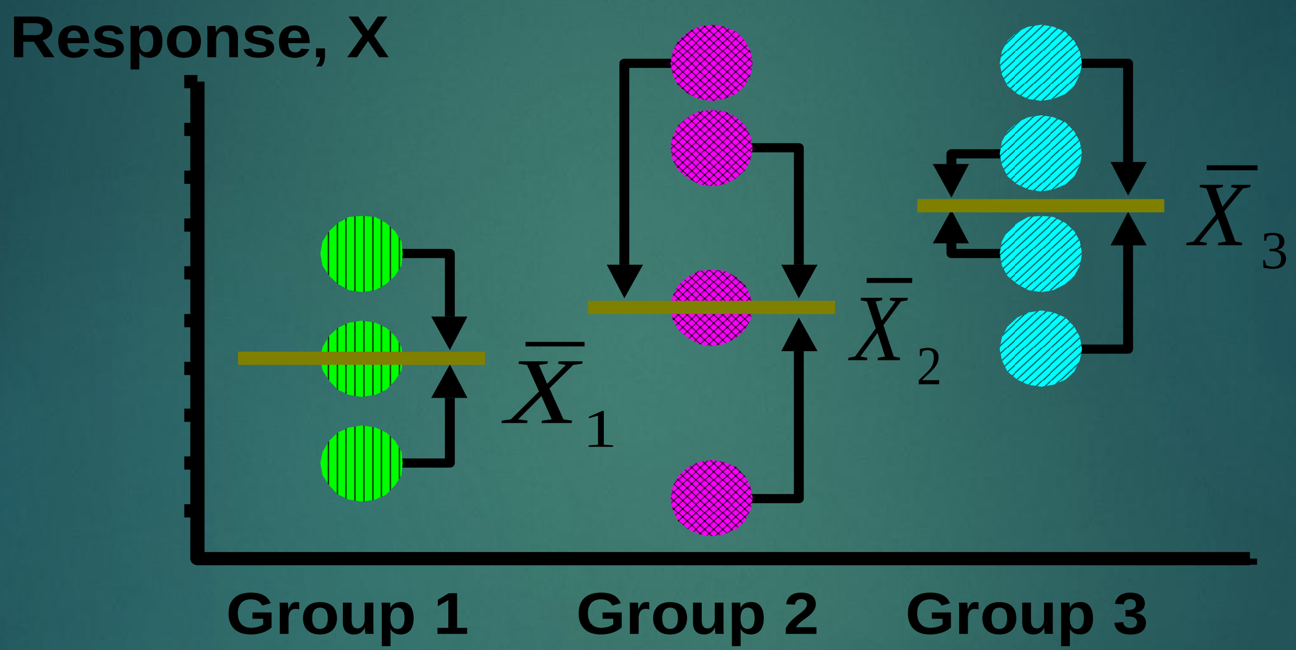
Among-Group Variation

Caused by **differences among groups**



Within-Group Variation

Caused by **random errors**



Total Sum of Squares

$$\boxed{SST} = SSA + SSW$$

$$SST = \sum_{j=1}^{c} \sum_{i=1}^{n_j} (X_{ij} - \bar{\bar{X}})^2$$

Where:

SST = Total sum of squares

c = number of groups (or levels)

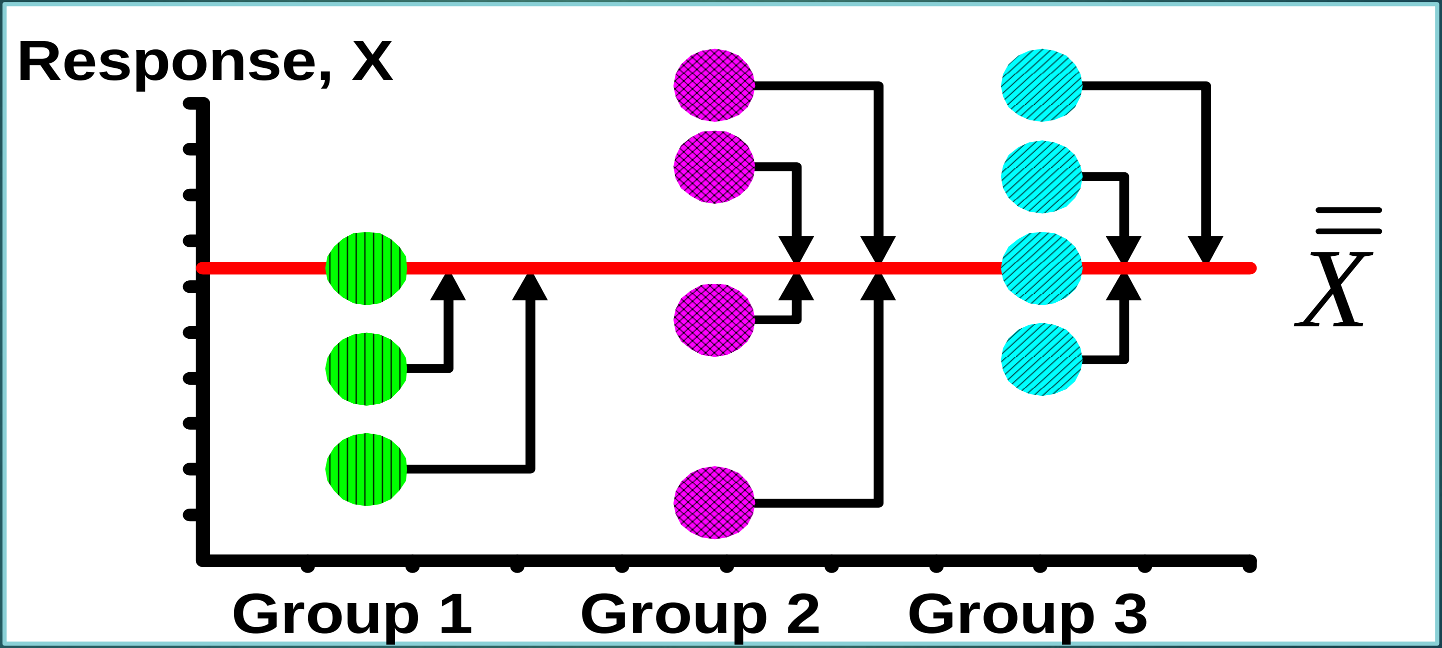
n_j = number of values in group j

X_{ij} = i^{th} observation from group j

$\bar{\bar{X}}$ = grand mean (mean of all data values)

Total Variation: SST

$$SST = (X_{11} - \bar{\bar{X}})^2 + (X_{12} - \bar{\bar{X}})^2 + \cdots + (X_{cn_c} - \bar{\bar{X}})^2$$



Among-Group Variation

$$SST = \boxed{SSA} + SSW$$

$$SSA = \sum_{j=1}^{c} n_j(\bar{X}_j - \bar{\bar{X}})^2$$

Where:

SSA = Sum of squares among groups

c = number of groups

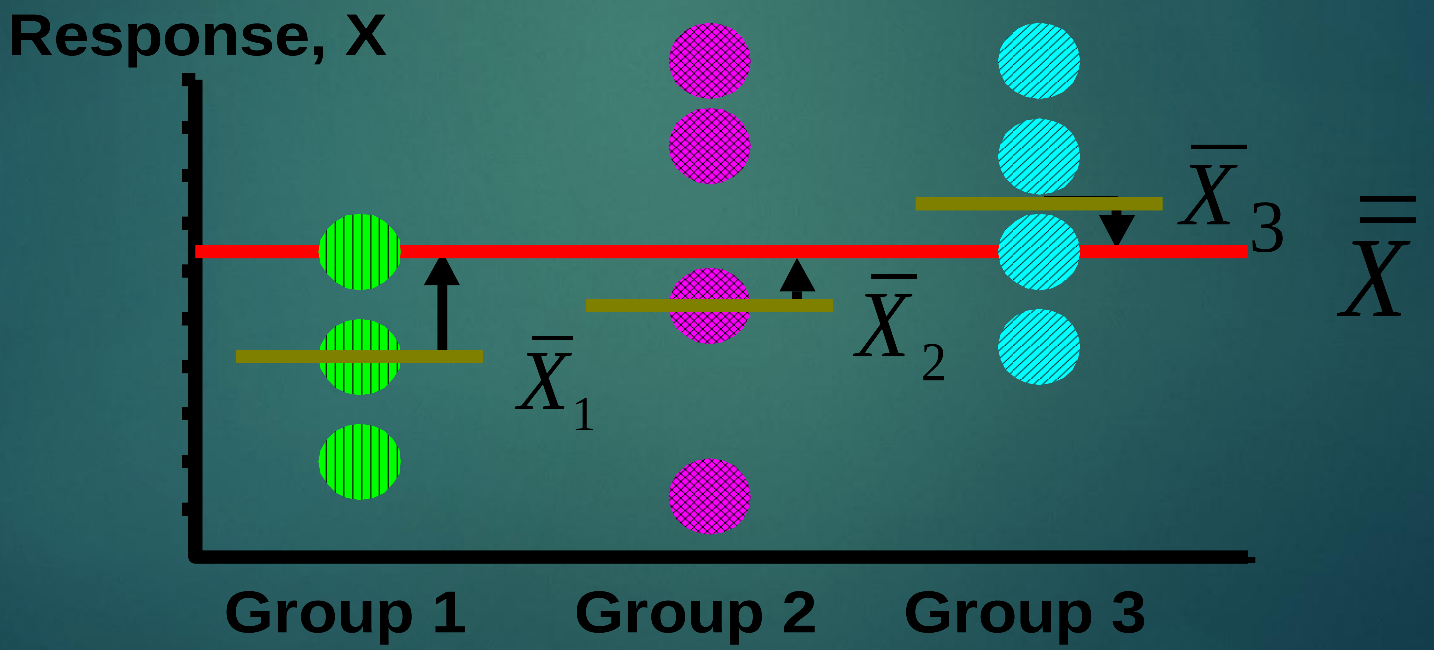
n_j = sample size from group j

$\bar{X}_j$ = sample mean from group j

$\bar{\bar{X}}$ = grand mean (mean of all data values)

Among-Group Variation: SSA

$$SSA = n_1(\bar{X}_1 - \bar{\bar{X}})^2 + n_2(\bar{X}_2 - \bar{\bar{X}})^2 + \cdots + n_c(\bar{X}_c - \bar{\bar{X}})^2$$



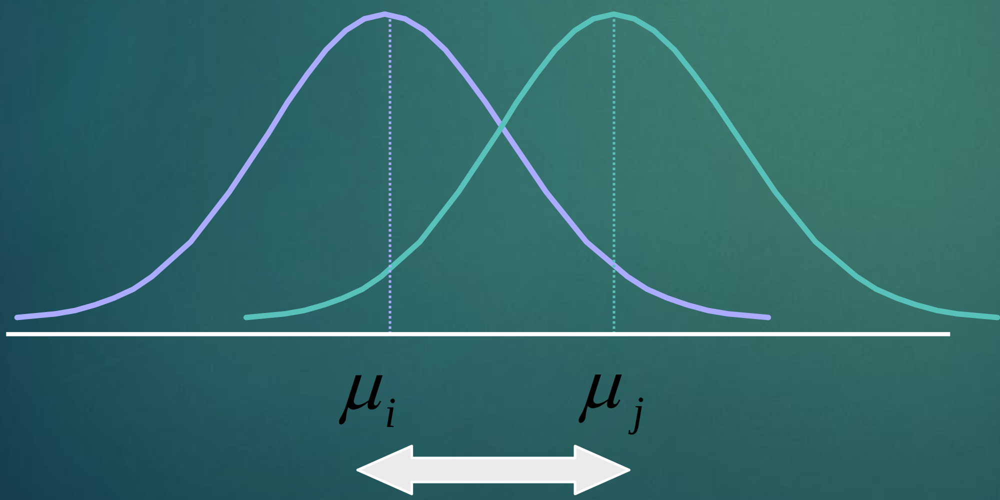
Among-Group Variation

$$SSA = \sum_{j=1}^{c} n_j (\bar{X}_j - \bar{\bar{X}})^2$$

$$MSA = \frac{SSA}{c-1}$$

Mean Square Among = SSA/degrees of freedom

Variation Due to
Differences Among Groups



Within-Group Variation

$$SST = SSA + \boxed{SSW}$$

$$SSW = \sum_{j=1}^{c} \sum_{i=1}^{n_j} (X_{ij} - \bar{X}_j)^2$$

Where:

SSW = Sum of squares within groups

c = number of groups

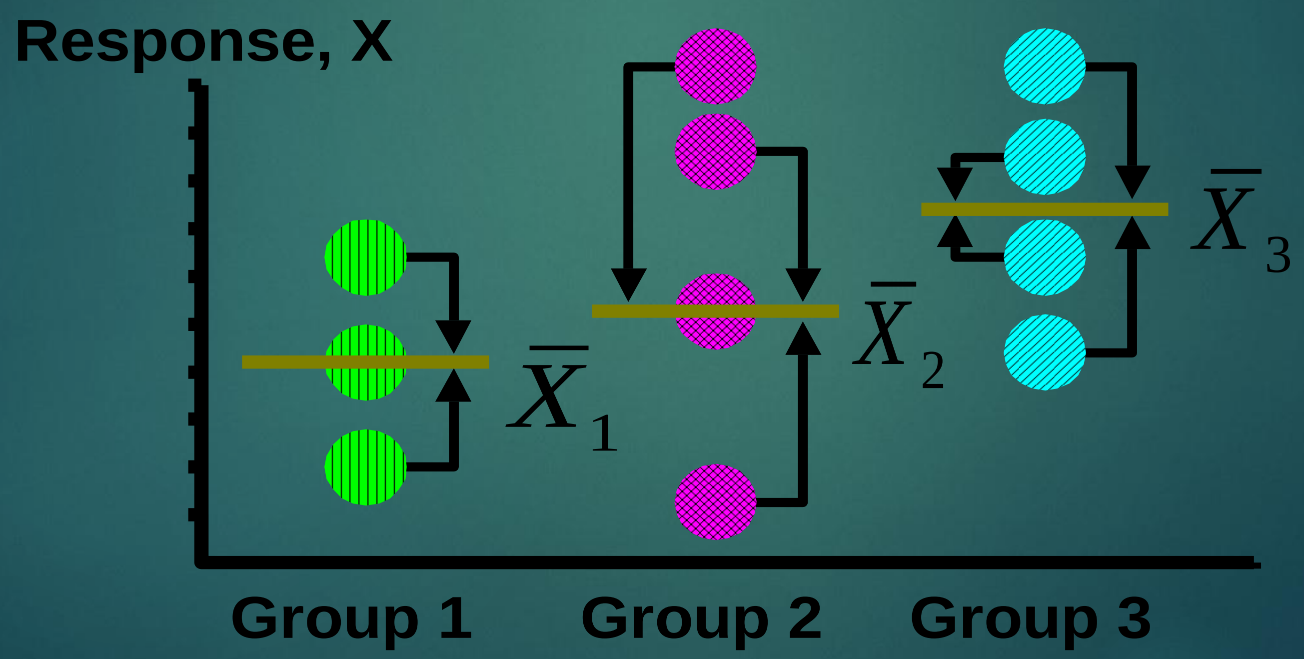
n_j = sample size from group j

$\bar{X}_j$ = sample mean from group j

X_{ij} = i^{th} observation in group j

Within-Group Variation

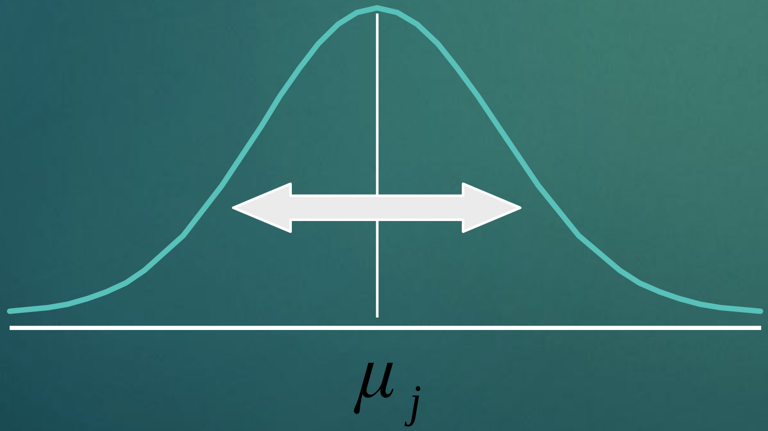
$$SSW = (X_{11} - \overline{X}_1)^2 + (X_{21} - \overline{X}_1)^2 + \cdots + (X_{cn_c} - \overline{X}_c)^2$$



Within-Group Variation

$$SSW = \sum_{j=1}^{c} \sum_{i=1}^{n_j} (X_{ij} - \overline{X}_j)^2$$

Summing the variation within each group and then adding over all groups



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

Mean Square Within = SSW/degrees of freedom

Obtaining the Mean Squares

The Mean Squares are obtained by dividing the various sum of squares by their associated degrees of freedom

$$MSA = \frac{SSA}{c-1}$$

Mean Square, Among (d.f. = c-1)

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

Mean Square, Within (d.f. = n-c)

$$MST = \frac{SST}{n-1}$$

Mean Square, Total (d.f. = n-1)

On the degrees of freedom

Mean Square, Among (MSA): **d.f. = c-1**

We are comparing the difference AMONG **c** different groups. Thus, the degree of freedom is **c-1**.

3 golf clubs → df=2

On the degrees of freedom

Mean Square, Among (MSA): **d.f. = c-1**

We are comparing the difference AMONG **c** different groups. Thus, the degree of freedom is **c-1**.

3 golf clubs → df=2

Mean Square, Within (MSW): **d.f. = n-c**

Each group generate **n_j-1** degree of freedom and there are **c** different groups. Thus, the degree of freedom is equal to:

df_MSW= $n_1-1+ n_2-1+ n_3-1+ n_4-1+\ldots+ n_j-1$

$n_1+ n_2+ n_3+ n_4+\ldots+ n_j=n$

There are c number of 1s too.

Thus, **df_MSW=n-c**

3 golf clubs, 5 mechanical shots each → df=15-3=12

On the degrees of freedom

Mean Square, Among (MSA): **d.f. = c-1**

We are comparing the difference AMONG **c** different groups. Thus, the degree of freedom is **c-1**.

3 golf clubs → df=2

Mean Square, Within (MSW): **d.f. = n-c**

Each group generate **n_j-1** degree of freedom and there are **c** different groups. Thus, the degree of freedom is equal to:

df_MSW= $n_1-1+n_2-1+n_3-1+n_4-1+\ldots+n_j-1$

$n_1+n_2+n_3+n_4+\ldots+n_j=n$

There are c number of 1s too.

Thus, **df_MSW=n-c**

3 golf clubs, 5 mechanical shots each → df=15-3=12

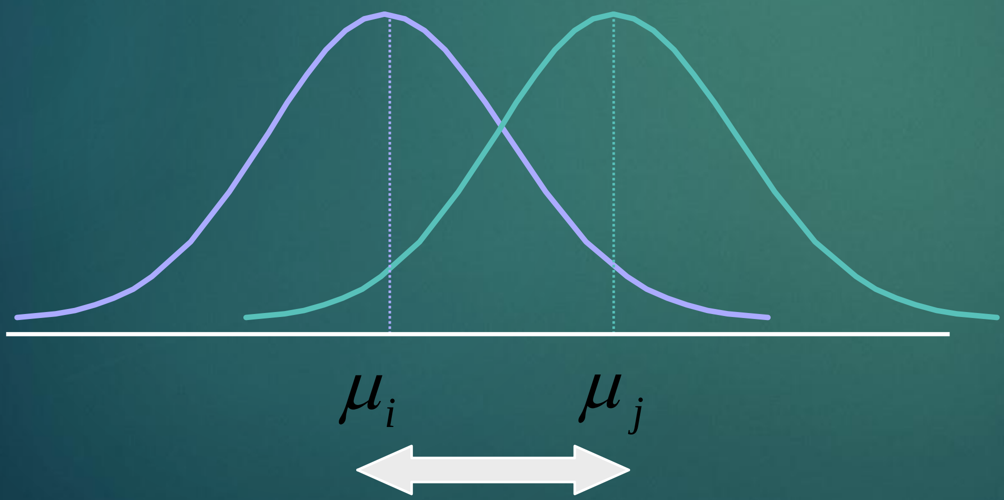
Mean Square, Total (MST): **d.f. = n-1**

Given the grand mean, the degree of freedom is **n-1**

Among-Group Variation

$$SSA = \sum_{j=1}^{c} n_j(\bar{X}_j - \bar{\bar{X}})^2$$

Variation Due to
Differences Among Groups



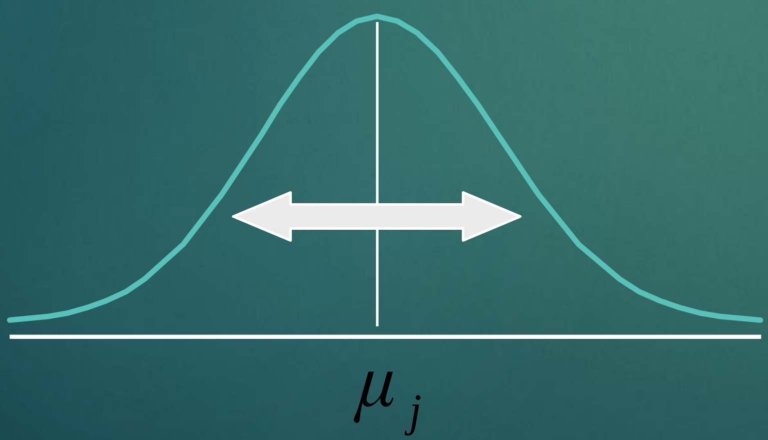
$$MSA = \frac{SSA}{c-1}$$

Mean Square Among = SSA/degrees of freedom

Within-Group Variation

$$SSW = \sum_{j=1}^{c} \sum_{i=1}^{n_j} (X_{ij} - \bar{X}_j)^2$$

Summing the variation within each group and then adding over all groups



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

Mean Square Within = SSW/degrees of freedom

One-Way ANOVA Table

Source of Variation	Degrees of Freedom	Sum Of Squares	Mean Square (Variance)	F
Among Groups	c - 1	SSA	$MSA = \frac{SSA}{c - 1}$	$F_{STAT} = \frac{MSA}{MSW}$
Within Groups	n - c	SSW	$MSW = \frac{SSW}{n - c}$	
Total	n – 1	SST		

c = number of groups
n = sum of the sample sizes from all groups
df = degrees of freedom

One-Way ANOVA F Test Statistic

H_0: $\mu_1 = \mu_2 = \cdots = \mu_c$

H_1: At least two population means are different

- Test statistic

$$F_{STAT} = \frac{MSA}{MSW}$$

MSA is mean squares among groups

MSW is mean squares within groups

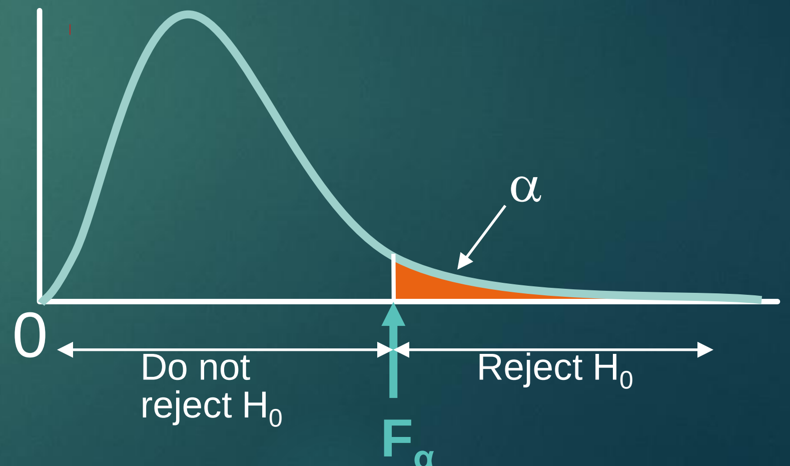
- Degrees of freedom
  - $df_1 = c - 1$ (c = number of groups)
  - $df_2 = n - c$ (n = sum of sample sizes from all populations)

Interpreting One-Way ANOVA F Statistic

- The F statistic is the ratio of the among estimate of variance and the within estimate of variance
  - The ratio must always be positive
  - $df_1 = c - 1$ will typically be small
  - $df_2 = n - c$ will typically be large

Decision Rule:

Reject H_0 if $F_{STAT} > F_\alpha$, otherwise do not reject H_0



One-Way ANOVA F Test Example

We wanted to test if three different golf clubs yield different distances.

Let's randomly select five measurements from trials on an automated driving machine for each club. At the 0.05 significance level, is there a difference in mean distance?

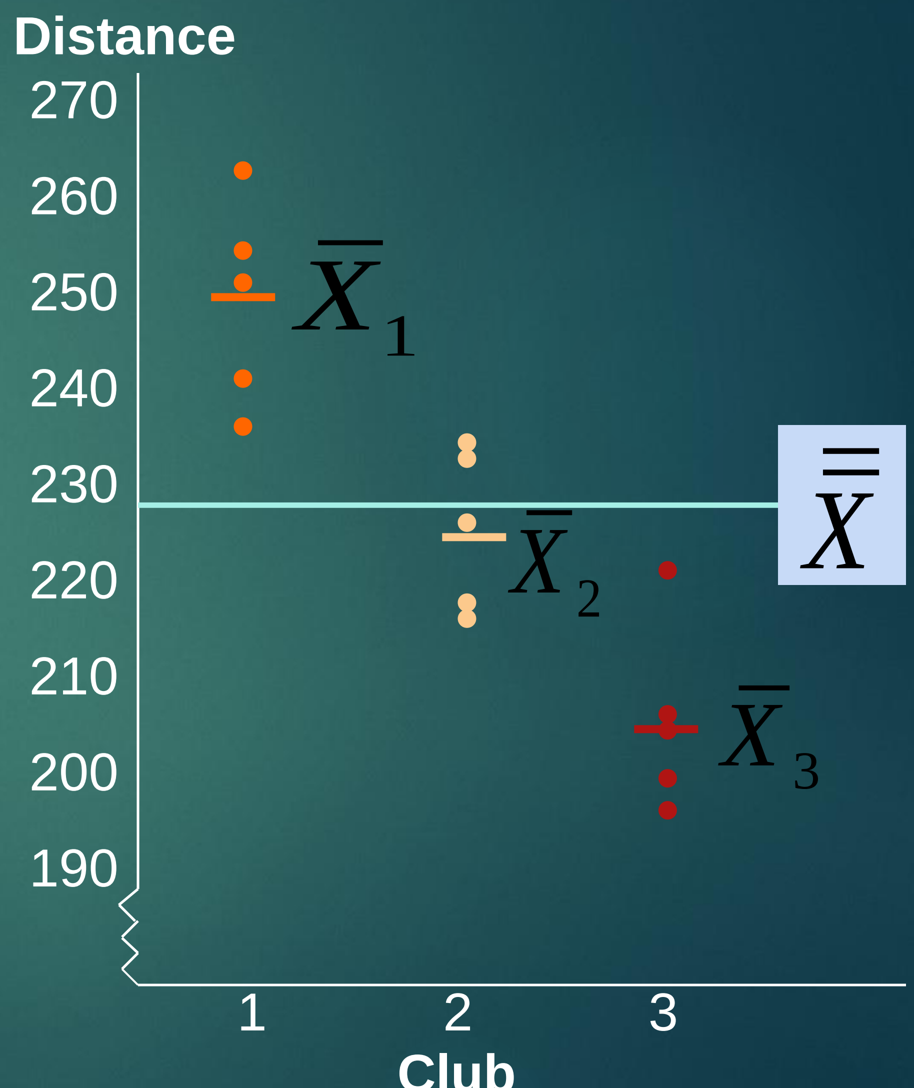
Club 1	Club 2	Club 3
254	234	200
263	218	222
241	235	197
237	227	206
251	216	204

One-Way ANOVA F Test
Scatterplot

Club 1	Club 2	Club 3
254	234	200
263	218	222
241	235	197
237	227	206
251	216	204

$\bar{x}_1 = 249.2$	$\bar{x}_2 = 226.0$	$\bar{x}_3 = 205.8$

$\bar{\bar{x}} = 227.0$



One-Way ANOVA F Test
F Stat

Club 1	Club 2	Club 3
254	234	200
263	218	222
241	235	197
237	227	206
251	216	204

$\bar{X}_1 = 249.2$ $n_1 = 5$

$\bar{X}_2 = 226.0$ $n_2 = 5$

$\bar{X}_3 = 205.8$ $n_3 = 5$

$\bar{\bar{X}} = 227.0$ $n = 15$

$c = 3$

One-Way ANOVA F Test
F Stat

Club 1	Club 2	Club 3
254	234	200
263	218	222
241	235	197
237	227	206
251	216	204

$\bar{X}_1 = 249.2$ $\quad n_1 = 5$

$\bar{X}_2 = 226.0$ $\quad n_2 = 5$

$\bar{X}_3 = 205.8$ $\quad n_3 = 5$

$\bar{\bar{X}} = 227.0$ $\quad n = 15$

$c = 3$

$$SSA = 5\,(249.2 - 227)^2 + 5\,(226 - 227)^2 + 5\,(205.8 - 227)^2 = 4716.4$$

$$SSW = (254 - 249.2)^2 + (263 - 249.2)^2 + \cdots + (204 - 205.8)^2 = 1119.6$$

One-Way ANOVA F Test
F Stat

Club 1	Club 2	Club 3
254	234	200
263	218	222
241	235	197
237	227	206
251	216	204

$\bar{X}_1 = 249.2$ $n_1 = 5$

$\bar{X}_2 = 226.0$ $n_2 = 5$

$\bar{X}_3 = 205.8$ $n_3 = 5$

$\bar{\bar{X}} = 227.0$ $n = 15$

$c = 3$

SSA = 5 (249.2 – 227)^2 + 5 (226 – 227)^2 + 5 (205.8 – 227)^2 = 4716.4

SSW = (254 – 249.2)^2 + (263 – 249.2)^2 +⋯+ (204 – 205.8)^2 = 1119.6

MSA = 4716.4 / (3-1) = 2358.2

MSW = 1119.6 / (15-3) = 93.3

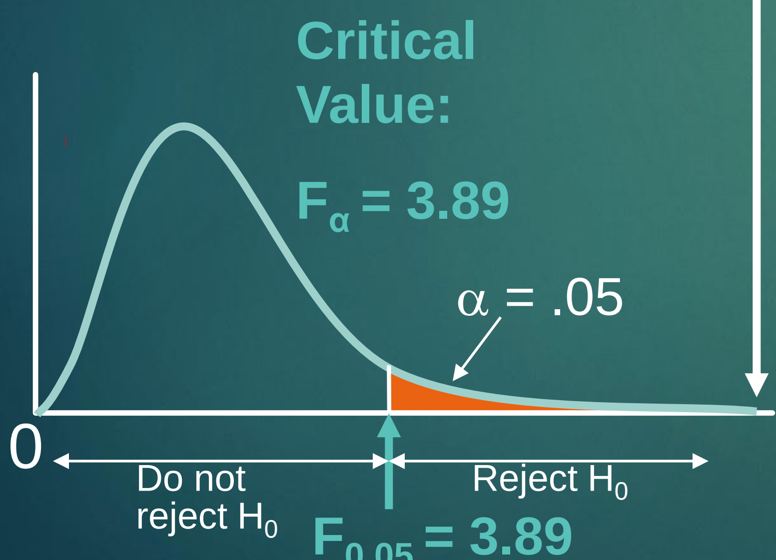
$$F_{STAT} = \frac{2358.2}{93.3} = 25.275$$

One-Way ANOVA F Test
Comparison

H_0: $\mu_1 = \mu_2 = \mu_3$
H_1: μ_j not all equal

$\alpha = 0.05$

$df_1 = 2$ $\quad df_2 = 12$



Test Statistic:

$$F_{STAT} = \frac{MSA}{MSW} = \frac{2358.2}{93.3} = 25.275$$

Decision:

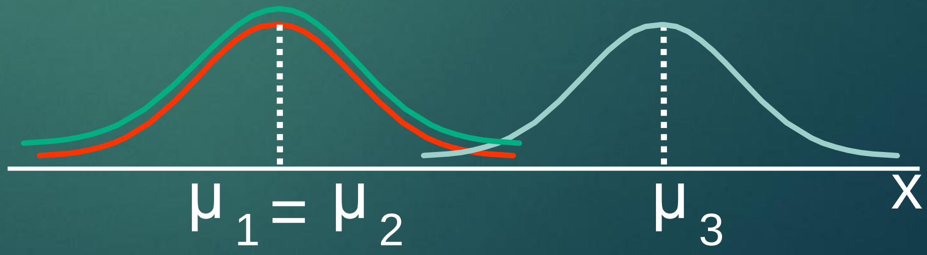
Reject H_0 at $\alpha = 0.05$

Conclusion:

There is evidence that at least one μ_j differs from the rest

The Tukey-Kramer Procedure

- Tells which population means are significantly different
  - e.g.: $\mu_1 = \mu_2 \neq \mu_3$
  - Done after rejection of equal means in ANOVA
- Allows paired comparisons
  - Compare absolute mean differences with critical range

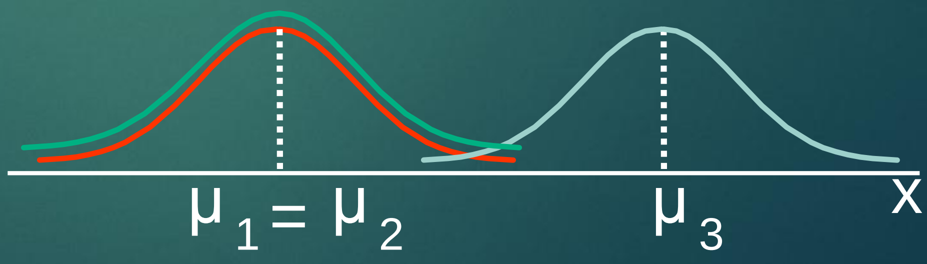


The Tukey-Kramer Procedure

- What do we need for this test?
  - We need to compute the **absolute value** of the **paired differences in sample means**

i.e.: $|\bar{X}_1 - \bar{X}_2|$, $|\bar{X}_1 - \bar{X}_3|$, and $|\bar{X}_2 - \bar{X}_3|$

  - We need to compare the magnitude of each difference to a **Critical Range**

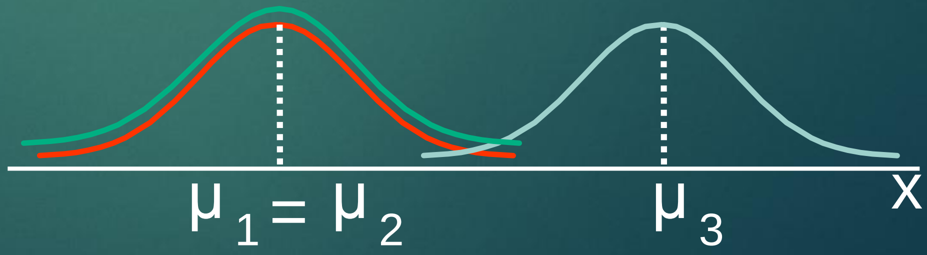


The Tukey-Kramer Procedure

- What do we need for this test?
  - We need to compute the **absolute value** of the **paired differences in sample means**

i.e.: $|\overline{X}_1 - \overline{X}_2|$, $|\overline{X}_1 - \overline{X}_3|$, and $|\overline{X}_2 - \overline{X}_3|$

  - We need to compare the magnitude of each difference to a **Critical Range**
  - **Decision**: If the absolute value of the paired difference is **greater** than the critical range, there exists enough **evidence of difference in the means of the pair** (e.g., μ_1 and μ_3 below). Otherwise, we cannot reject the null hypothesis that the population means in the pair are the



Tukey-Kramer Critical Range

$$\text{Critical Range} = Q_{\alpha}\sqrt{\frac{\text{MSW}}{2}\left(\frac{1}{n_j}+\frac{1}{n_{j'}}\right)}$$

where,

Q_a = Upper Tail Critical Value from **Studentized Range Distribution** (beyond the scope of this course) with c and n - c degrees of freedom (see appendix E.7 table)

MSW = Mean Square Within

n_j and **$n_{j'}$** = Sample sizes from groups j and j'

The Tukey-Kramer Procedure: Example

Club 1	Club 2	Club 3
254	234	200
263	218	222
241	235	197
237	227	206
251	216	204

1. Compute absolute mean differences:

$$|\bar{x}_1 - \bar{x}_2| = |249.2 - 226.0| = 23.2$$

$$|\bar{x}_1 - \bar{x}_3| = |249.2 - 205.8| = 43.4$$

$$|\bar{x}_2 - \bar{x}_3| = |226.0 - 205.8| = 20.2$$

The Tukey-Kramer Procedure: Example

Club 1	**Club 2**	**Club 3**
254	234	200
263	218	222
241	235	197
237	227	206
251	216	204

1. Compute absolute mean differences:

$$\left|\bar{x}_1 - \bar{x}_2\right| = \left|249.2 - 226.0\right| = 23.2$$

$$\left|\bar{x}_1 - \bar{x}_3\right| = \left|249.2 - 205.8\right| = 43.4$$

$$\left|\bar{x}_2 - \bar{x}_3\right| = \left|226.0 - 205.8\right| = 20.2$$

2. Find the Q_α value (from the table in appendix E.7 with $c = 3$ and $(n - c) = (15 - 3) = 12$ degrees of freedom):

$$Q_\alpha = 3.77$$

The Tukey-Kramer Procedure: Example

3. Compute Critical Range:

$$\text{Critical Range} = Q_{\alpha}\sqrt{\frac{\text{MSW}}{2}\left(\frac{1}{n_j} + \frac{1}{n_{j'}}\right)} = 3.77\sqrt{\frac{93.3}{2}\left(\frac{1}{5} + \frac{1}{5}\right)} = 16.285$$

The Tukey-Kramer Procedure: Example

3. Compute Critical Range:

$$\text{Critical Range} = Q_{\alpha}\sqrt{\frac{MSW}{2}\left(\frac{1}{n_j}+\frac{1}{n_{j'}}\right)} = 3.77\sqrt{\frac{93.3}{2}\left(\frac{1}{5}+\frac{1}{5}\right)} = 16.285$$

4. Compare:

$$\left|\bar{X}_1 - \bar{X}_2\right| = 23.2$$

$$\left|\bar{X}_1 - \bar{X}_3\right| = 43.4$$

$$\left|\bar{X}_2 - \bar{X}_3\right| = 20.2$$

The Tukey-Kramer Procedure: Example

$$\text{Critical Range} = Q_\alpha \sqrt{\frac{\text{MSW}}{2}\left(\frac{1}{n_j} + \frac{1}{n_{j'}}\right)} = 3.77\sqrt{\frac{93.3}{2}\left(\frac{1}{5} + \frac{1}{5}\right)} = 16.285$$

All of the absolute mean differences are **greater** than the critical range. Therefore **there is a significant difference between each pair of means** at 5% level of significance.

$$\left|\bar{x}_1 - \bar{x}_2\right| = 23.2$$

$$\left|\bar{x}_1 - \bar{x}_3\right| = 43.4$$

$$\left|\bar{x}_2 - \bar{x}_3\right| = 20.2$$

FOR MORE: READ FROM P. 394 TO P. 404.

NOTE: THESE MATERIALS ARE SUBJECT TO THE FINAL EXAM