One-Way Between-Subjects Analysis of Variance (ANOVA)

- t-tests: 2 sample means and determined the probability that they were drawn from the same population.

- Analysis of Variance (ANOVA): Groups > 2
  - Like t-tests: ANOVA deals with differences between sample means.
  - Unlike t-tests it imposes no restriction on the number of means.

Why not just use t-tests?

- Imagine you have 3 groups: Just use a series of independent groups t-test to test whether there are differences?
  - 2 potential problems with this approach.
    1. Too many comparisons!
      - There would be 3 different t tests:

```markdown
Group 1 v. Group 2
Group 1 v. Group 3
Group 2 v. Group 3
```
• Number of t tests increases as the number of groups increases.
• Groups = 3, then we have 3 t tests
• 6 groups we would have to perform 15 different t tests.

Formula:

\[ C = \frac{(k)(k-1)}{2} \]

- 4 groups \[ C = \frac{4(4-1)}{2} = 6 \]

2. Inflated probability of making a Type I error.
• Probability of at least one Type I error increases as the number of t tests increases.
• alpha at .05 indicates we are willing to risk being wrong 5% of the time when we reject the Ho.
• With many t-tests we are no longer just sampling 1 t-value from the t distribution.
• Result: alpha is no longer equal to .05. Probability of making a Type I error has gone. Experimentwise error

With 3 groups the experiment-wise alpha = .14
Formula: $1 - (1-\alpha)^C$  
1 - (1 - .05)^3 = .14

4 groups: C = 6  
1 - (1 - .05)^6 = .26

**So, what’s the Advantage of an F-test?**

- **F-test**: one overall comparison
- Avoids increased Type I error probability.

**Definition of ANOVA**

- Analysis of variance (ANOVA) is a statistical method for comparing 2 or more groups/set of observations.
- Same circumstances as the independent groups t-test except more levels of the IV.

- Relationship between t and F.
  - $t^2 = F$

  - Example: $t = 2.00$
  - F value? $F = 4.00$
In our example what are the IV and DV?

- IV is the type of antismoking campaign (has 3 levels)
- the DV is the # of cigarettes smoked per week.

**Null and Alternative?**

- Ho: \( \mu_1 = \mu_2 = \mu_3 \)

- H1: \( \mu_1 \neq \mu_2 \neq \mu_3 \)

- What are we testing?
  - At least two of the means are different.
    
    (a) all of the means can be different from one another;
    
    (a) Alternatively, maybe social learning condition is different from antismoking pamphlets, but anti-smoking pamphlets and no treatment are not different.

ANOVA only tells us that there is a difference—Omnibus test
Assumptions of the F-Test

1. **Homogeneity of Variance:** Each of our populations has the same variance. In other words:

   \[ \sigma_1^2 = \sigma_2^2 = \sigma_3^2 \ldots \]

2. **Normality:** The scores for each condition are normally distributed around the population mean.

3. **Independence of observations:** Observations within an experimental treatment are independent of one another.

**Violate Assumptions?**

- ANOVA is a robust test
- Minimally affected by violations of population normality and homogeneity of variance so long as the samples are of equal size.
- **Big Problem:** violate homogeneity of variance and n-sizes are not equal
**Logic of ANOVA-Hypothesis testing**

A quick review:

1. Assume that Ho is true.
2. State an expected result based on this assumption
3. Compute a sample statistic (t-test; F-test etc…).
4. Treat the statistic as a score in some known sampling distribution.
5. If the statistic falls within the rejection region, reject the null hypothesis.
6. Otherwise, fail to reject the null hypothesis

What’s new with ANOVA?
- Basic Logic does not change!
- Difference: use different statistical test.
  - F-test (test of the ratio between two variances)

**Distinguishing Two types of Variability**

- Between-group variability and within group variability.

**Within group versus Between group.**

- Within group variability estimate: Individual differences, errors in measurement etc..
  - unexplained variability (sometimes called error variance).
• Between group variability: Same sources as within, but also includes any variability due to group differences.
  • experimental manipulation

• Different ways to estimate variance basis of the F-test.

• Rather than using means to evaluate the Ho, the F-test is based on the ratio of the variances.

• ANOVA: calculate $F$

• $F$ is the ratio of 2 independent $\sigma^2$ estimates.

  $F = \frac{\text{variance b/n sample means}}{\text{variance expected by chance (error)}}$

  $F = \frac{\text{Differences among treatment means}}{\text{Differences among Ss treated alike}}$

If Ho is true:

  $F = \frac{\text{Sampling error + Measurement error}}{\text{Sampling error + measurement error}} = 1$
If Ho is false:

\[ F = \text{Sampling error} + \text{effect of IV} + \text{Measurement error} > 1 \]
\[ \text{Sampling error} + \text{Measurement error} \]

Variance estimates for ANOVA (see section 16.2 Howell)

- Analysis of Variance: Two independent measures of variability
- Population variability can be estimated in two ways.
  - First, for any given treatment, the variance of the scores that compose that treatment could be estimated and used as an estimate of the underlying population variance.

\[ s^2 \] is used to estimate \( \sigma^2 \)

Multiple groups from the same underlying population:
- Assume homogeneity of variance. Thus, multiple groups represent multiple estimates of the common population variance.
  - We can take the average:
    - \( \Sigma s^2/k \)
  - Result? MSerror (or the Mswithin). Denominator of the F-ratio.
• Does not depend on the truth of Ho.

Second estimate:
• assumes Ho is true.
• If Ho is true each group is an independent samples from the same population.

central limit theorem

• The variance of means drawn from the same population equals the variance of the population divided by the sample size.

\[ S_x^2 = \sigma^2/n \]

• We want to estimate the population variance.
• Result is the following formula:

\[ N * S_x^2 = \sigma^2 \]

• MS_{treatment} or MS_{between} variance component.

Now have two estimates of the population variance \( \sigma^2 \)

1. MS error-independent of the truth of Ho.
2. The second is only an estimate of \( \sigma^2 \) as long as the Null hypothesis is true..
Summarize the logic of ANOVA—
1. Calculate two estimates of the population variance

2. If the two estimates agree, we have no reason to reject Ho.
   • If they disagree: conclude that underlying treatment inflated our second estimate.

If Ho is true:
\[
F = \frac{\text{Sampling error} + \text{Measurement error}}{\text{Sampling error} + \text{measurement error}} = 1
\]

If Ho is false:
\[
F = \frac{\text{Sampling error} + \text{effect of IV} + \text{Measurement error}}{\text{Sampling error} + \text{Measurement error}} > 1
\]
Calculations
DATA SET (Handout)

An example of ANOVA Calculations

Class Example: Three junior high schools participated in an experiment on smoking prevention.

<table>
<thead>
<tr>
<th>School #1</th>
<th>School #2</th>
<th>School #3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Social Learning</td>
<td>Anti-Smoking Pamphlets</td>
<td>No Treatment (control)</td>
</tr>
<tr>
<td>X₁</td>
<td>X₁²</td>
<td>X₂</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>9</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>9</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>16</td>
<td>5</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>25</td>
<td>5</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>6</td>
</tr>
</tbody>
</table>

| 23 | 71 | 36 | 158 | 48 | 268 |

n₁ = 10  n₂ = 10  n₃ = 10

\[ \bar{X}_1 = 2.3 \]
\[ \bar{X}_2 = 3.6 \]
\[ \bar{X}_3 = 4.8 \]
To calculate F we will need:

- \( \text{MS}_{\text{error}} \) and \( \text{MS}_{\text{between}} \) To obtain these we need:

**Total Sums of Squares** (\( \text{SS}_{\text{total}} \))

**Error Sums of Squares** (\( \text{SS}_{\text{error}} \))

**Group (between) Sums of Squares** (\( \text{SS}_{\text{between}} \))

**Total degrees of freedom** (\( df_{\text{total}} \))

**Error degrees of freedom** (\( df_{\text{error}} \))

**Group degrees of freedom** (\( df_{\text{between}} \))

**Notation:**

Observation \( X_{ij} \) (observation \( i \) in condition \( j \))

Call the mean of any condition \( \bar{X}_j \)

\( N = \) Total number of subjects
\( n = \) number in a group

The overall grand mean of the data: \( \bar{X}_{..} \)

The sum of the x’s for a single group \( T_j \)
Our data:
\[ T_1 = 23; T_2 = 36; T_3 = 48 \]

Grand mean: \[ \frac{\Sigma X_{tot}}{N} = \frac{107}{30} = 3.57 \]

### Partitioning the variance

<table>
<thead>
<tr>
<th>Computational Formula</th>
<th>Definitional Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>( S_{total} = )</td>
<td>( \frac{\left( \Sigma X \right)^2}{\Sigma X^2 - \frac{N}{n}} )</td>
</tr>
</tbody>
</table>

- **\( S_{total} \)**: total variation in all of the scores
- **Compare to variance formula. It’s exactly the same as top component (SS).**

<table>
<thead>
<tr>
<th>Computational Formula</th>
<th>Definitional Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>( S_{between} = )</td>
<td>( \sum \frac{T_i^2}{n} - \frac{(\Sigma X)^2}{N} )</td>
</tr>
</tbody>
</table>

- **Note**: Terms between, treatment, and group used interchangeably.
- **Definitional formula you can see that you are assessing the variability of the group means around the grand mean for all of the groups.**
### Computational Formula vs. Definitional Formula

<table>
<thead>
<tr>
<th>Computational Formula</th>
<th>Definitional Formual</th>
</tr>
</thead>
<tbody>
<tr>
<td>$SS_{\text{within}} =$</td>
<td>$\sum \frac{T_i^2}{n} \quad \Sigma X^2 \quad \Sigma(X_{ij} - \bar{X}_j)^2$</td>
</tr>
</tbody>
</table>

Note: You will often see this term referred as within or error.
- $SS_{\text{within}}$ calculates the degree to which individual scores vary from their group means.

### Back to our data:

<table>
<thead>
<tr>
<th>Computational Formula</th>
<th>Calculations</th>
</tr>
</thead>
<tbody>
<tr>
<td>$SS_{\text{total}} =$</td>
<td>$\frac{(\Sigma X)^2}{\Sigma X^2 - \frac{N}{n}} \quad \Sigma 71 + 158 + 268 - \frac{(\Sigma 23 + 36 + 48)^2}{30}$</td>
</tr>
<tr>
<td></td>
<td>$497 - \frac{11449}{30} \quad = 115.37$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Computational Formula</th>
<th>Calculation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$SS_{\text{between}} =$</td>
<td>$\frac{\sum T_i^2}{n} \quad \frac{(\Sigma X)^2}{N} \quad \frac{\Sigma 529 + 1296 + 2304}{10} - \frac{(107)}{30}$</td>
</tr>
<tr>
<td></td>
<td>$412.9 - 381.63 \quad = 31.27$</td>
</tr>
</tbody>
</table>
Alternatively, use the definitional formula:

\[ n \Sigma (X_j - \bar{X}_..)^2 = 10 \Sigma (2.3 - 3.56)^2 + (3.6 - 3.56)^2 + (4.8 - 3.56)^2 \]

\[ = 10 \Sigma 1.5876 + .0016 + 1.5376 = 31.27 \]

- Finally, calculate \( SS_{\text{within}} = \)

<table>
<thead>
<tr>
<th>Computational Formula</th>
<th>Calculation</th>
</tr>
</thead>
<tbody>
<tr>
<td>( SS_{\text{within}} = ) ( \Sigma X^2 - \Sigma \left( \frac{Tj^2}{n} \right) )</td>
<td>( \Sigma 71 + 158 + 268 - \Sigma \left( \frac{(23)^2 + (36)^2 + (48)^2}{10} \right) )</td>
</tr>
<tr>
<td></td>
<td>( 497 - (52.9 + 129.6 + 230.4) )</td>
</tr>
<tr>
<td></td>
<td>( = 84.1 )</td>
</tr>
</tbody>
</table>

- Alternatively, given that \( SS_{\text{total}} = SS_{\text{between}} + SS_{\text{within}} \) we could calculate \( SS_{\text{within}} \) as follows:

\[ 115.37 = 31.27 + SS_{\text{within}} \]

Therefore: \( SS_{\text{within}} = 115.37 - 31.27 = 84.1 \)
Four important points:

1. If you know any two you can calculate the third:

\[ SS_{\text{total}} = SS_{\text{between}} + SS_{\text{within}} \]

2. Check your calculations:

\[ SS_{\text{total}} = SS_{\text{between}} + SS_{\text{within}} \]

- The elements on the right side equal the left side \((SS_{\text{total}})\)

3. SS values can not be negative.

4. \(SS_{\text{between}}\) and \(SS_{\text{within}}\) will each be less than the \(SS_{\text{total}}\)

- \(SS_{\text{within}}\) generally should not be greater than \(SS_{\text{between}}\).

If Ho is true:
\[ F = \frac{\text{Sampling error + Measurement error}}{\text{Sampling error + measurement error}} = 1 \]

If Ho is false:
\[ F = \frac{\text{Sampling error + effect of IV + Measurement error}}{\text{Sampling error + Measurement error}} > 1 \]
**MS**

- Calculate the Mean Squares.
- Divide SS by degrees of freedom.

**Reminder:**
Variance is sum of squares divided by the degrees of freedom:

\[ S^2 = \frac{SS}{N - 1} \]

**Computational formula for Variance:**

\[ S^2 = \frac{(\sum X)^2}{\sum X^2 - \frac{N}{N-1}} \]  
\[ \text{denominator} \]

Top of this formula no different than the formula we developed for \( SS_{\text{total}} \) earlier:

\[ SS_{\text{total}} = \frac{(\sum X)^2}{\sum X^2 - \frac{N}{N}} \]

**Degrees of Freedom.**
- Total degrees of freedom are \( N-1 \)—\( N \) is the total number of observations.
**Degrees of freedom between**
- $k-1$, where $k$ is the number of treatments.
- The total variability between our groups is based on $k$ scores, therefore we have $k-1$.
- $\text{MS}_{\text{between}}$:
  \[
  \text{MS}_{\text{between}} = \frac{\text{SS}_{\text{between}}}{k-1} = \frac{31.27}{3-1} = 15.633
  \]

**Degrees of freedom within**
- Calculated in one of two ways.
  - Simply subtract: Total df - Between df = Within df).
    - e.g., $29 - 2 = 27$
  - Alternatively: $k(n-1)$.
    - e.g., $3(10-1) = 27$

Calculate $\text{MS}_{\text{within}}$:
\[
\text{MS}_{\text{within}} = \frac{\text{SS}_{\text{within}}}{k(n-1)}
\]

With our data:
\[
\text{MS}_{\text{within}} = \frac{84.100}{27} = 3.115
\]
Calculating F

\[ \frac{\text{MS}_{\text{between}}}{\text{MS}_{\text{within}}} = F = \frac{15.633}{3.115} = 5.019 \]

Is F significant?

- F- table (Tables E3 and E4).
  - alpha and df between and df within.

  - E.g., F (2, 27) degrees of freedom. Critical value = 3.34.

- F exceeds the critical value reject the null hypothesis. If it does not “fail to reject the null hypothesis”.

The Summary Table

- Always construct an ANOVA source table or a Summary Table:

  - Summarizes the series of calculations: Sources of variation, degrees of freedom, mean squares, and finally the F-statistic.

Sources of variation:
• The first column: sum of squares.

• Total sum of squares and two portions: Between and within

**Degrees of Freedom**
• degrees of freedom column: allocation of the total number of degrees of freedom between sources.

**Mean Squares**
• simply derived by dividing SS by respective degrees of freedom.

\[ F \]
• divide the MS\(_{\text{between}}\)/MS\(_{\text{error}}\)

For our data:

<table>
<thead>
<tr>
<th>SOURCE</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between</td>
<td>31.3</td>
<td>2</td>
<td>15.65</td>
<td>5.03</td>
</tr>
<tr>
<td>Within</td>
<td>84.1</td>
<td>27</td>
<td>3.11</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>115.4</td>
<td>29</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
What does a significant F tell us?

- Tests the null hypothesis against all possible alternatives.

E.g.,
Ho: \( \mu_1 = \mu_2 = \mu_3 \)

Large number of alternative hypotheses. For instance:

1. \( \mu_1 \neq \mu_2 \neq \mu_3 \)
2. \( \mu_1 = \mu_2 \neq \mu_3 \)
3. \( \mu_1 \neq \mu_2 = \mu_3 \)

- Conclude that there is some difference, it’s just not clear where that difference lies.

- How can we tell where the difference lies?

  - Follow it up with more statistical tests (post hocs and contrasts).