BETWEEN-SUBJECTS, ONE-WAY ANOVA

ANALYSIS OF VARIANCE (ANOVA)

We have learned statistical tests for analyzing the effect of an EV on a DV by:

- Comparing a single sample to a population
  - Single Sample z-Test
  - Single Sample t-Test
- Comparing two samples to each other
  - Independent Samples t-Test
  - Paired Samples t-Test

- Next, we'll learn a statistical test for analyzing the effect of an EV on a DV by comparing **more than two samples**
  - Analysis of Variance, or ANOVA for short

TYPES OF ANOVA

- The type of ANOVA depends on the **number of explanatory variables** *and* whether the samples are **independent** or **related**

- **Between-Subjects, One-Way ANOVA**
  - The participants in each sample are *independent* of each other
  - There is one (1) explanatory variable
  - More than two samples are being compared

- **Repeated-Measures, One-Way ANOVA**
  - The participants in each sample are *related* to each other
  - There is one (1) explanatory variable
  - More than two samples are being compared

THE ISSUE WITH COMPARING 2+ SAMPLES

Example: Say a researcher is interested in testing whether there is a difference in the artistic abilities of people who are left-handed, right-handed, or ambidextrous.

- Why not just perform multiple t-Tests?
  - Each time you perform a statistical test using sample data, there is a chance of making a **Type I Error**

- Each time you perform a statistical test, there is a 5% chance of making a Type I error
  - Performing multiple tests increases the chances of making a Type I Error

ANOVA

- The Analysis of Variance (ANOVA) test allows you to compare **all** samples to each other using a *single* test

- The **Between-Subjects, One-Way ANOVA** does this by comparing the following measures of variability:
  - Between Group Variability
  - Within Group Variability

- Reminder: *Variability* refers to how much variation there is among a set of scores

BETWEEN-SUBJECTS, ONE-WAY ANOVA

- *Thought Question*: The explanatory variable is handedness, and the dependent variable is level of artistic ability (from 1 to 10)

Left-Handed	Right-Handed	Ambidextrous
7	3	4
5	6	4
6	2	5
8	4	7

$M_1 = 6.50$ $M_2 = 3.75$ $M_3 = 5.00$

- Why aren't all of the scores ***within*** a condition the same?
  - **Individual Differences** (all the things about individuals that make them different from one another)
- Why do the sample means ***between*** conditions differ?
  - **Individual Differences AND Treatment Effect**

BETWEEN-SUBJECTS, ONE-WAY ANOVA

- **Within-group variability:** the variability among the scores *within* each sample
  - This variability is due to **individual differences**, which are characteristics that vary between individuals

- **Between-group variability:** the variability among the scores *between* each sample
  - This variability is due to **individual differences PLUS treatment effect**
  - **Treatment effect** refers to the effect of the explanatory variable on the dependent variable

BETWEEN-SUBJECTS, ONE-WAY ANOVA

Between-Subjects One-WAY ANOVA Test Statistic:

$$F = \frac{\text{Between Group Variability}}{\text{Within Group Variability}}$$

- Conceptually, this ratio is measuring:

$$F = \frac{\text{Individual Differences} + \text{Treatment Effect}}{\text{Individual Differences}}$$

- *Question*: If there is **NO** effect of the EV on the DV (meaning no treatment effect), what will F be equal to?
  - When there is no treatment effect, F ≈ 1
  - When there is a treatment effect, F should be *greater than* 1

CONDUCTING A BETWEEN-SUBJECTS, ONE-WAY ANOVA

Example: A researcher is interested in how the type of feedback students receive affects their performance on a math test. The researcher randomly assigns 12 participants to receive either positive, negative, or neutral feedback about their math skills. After receiving the feedback, each participant takes a ten-item math test and the researcher measures the number of questions each participant gets correct.

DV: Math Performance

EV: Type of feedback

- Positive
- Negative
- Neutral

Positive	Negative	Neutral
7	3	5
7	4	6
8	3	5
6	2	4
$M_1 = 7.00$ $s = 0.82$	$M_2 = 3.00$ $s = 0.82$	$M_3 = 5.00$ $s = 0.82$

WHEN TO CONDUCT A BETWEEN-SUBJECTS, ONE-WAY ANOVA

Step 1: Choose the correct test

- A between-subjects, one-way ANOVA is used to compare
  - The effect of *one* explanatory variable on a dependent variable
  - When there are more than two samples
  - The participants in each sample are *independent* of one another

Step 2: Assumptions

TABLE 10.4 Assumptions for One-Way ANOVA

	Assumption	Robustness
1. Random samples	Each sample is a random sample from its population.	Robust
2. Independence of cases	Each case is not influenced by other cases in the sample.	Not robust
3. Normality	The dependent variable is normally distributed in each population.	Robust, especially if sample size (N) is large and the n's are about equal.
4. Homogeneity of variance	The degree of variability in the populations is equivalent.	Robust, especially if sample size (N) is large and the n's are about equal

THE NULL AND ALTERNATIVE HYPOTHESES

Step 3: Hypotheses

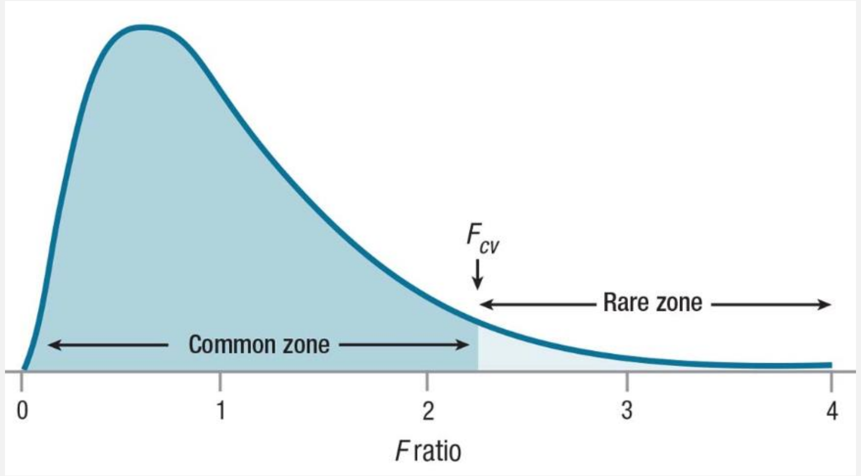
- The null hypothesis states the expected relationship between population means if there is ***no*** effect of the EV on the DV
  - H_0: $\mu_1 = \mu_2 = \mu_3$

- The alternative hypothesis states the expected relationship between population means if there is an effect of the EV on the DV
  - H_1: At least one population mean is different from the others.

THE DECISION RULE

Step 4: Decision Rule

a) The sampling distribution for an ANOVA is an F-distribution made up of the F-ratios that would occur when the null is true.



$$F = \frac{\text{Between Group Variability}}{\text{Within Group Variability}}$$

b) Specify alpha (typically .05 or .01)

THE DECISION RULE

c) Find the F-critical value on an F-critical value table using:

- Alpha and degrees of freedom
  - Numerator degrees of freedom
  - Denominator degrees of freedom

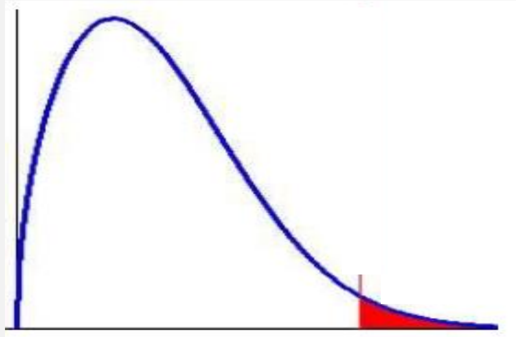
$$F = \frac{\text{Between Group Variability}}{\text{Within Group Variability}}$$

- $df_{\text{Between}} = (k - 1)$ ← Numerator df
- $df_{\text{Within}} = (N - k)$ ← Denominator df

k = the number of groups (aka, samples)
N = the total number of scores

For our example:

- $df_{\text{Btw}} = 3 - 1 = 2$
- $df_{\text{Wth}} = 12 - 3 = 9$
- $Fcv = 4.256$



$F_{\text{CV}} = 4.256$

CALCULATING THE TEST STATISTIC

Step 5: Calculate the test statistic

- The formula for calculating the test-statistic for a Between-Subjects, One-Way ANOVA is:

$$F = \frac{MS_{\text{Between}}}{MS_{\text{Within}}} \Leftarrow \frac{\text{Between groups variance}}{\text{Within groups variance}}$$

- **Note: *MS* stands for "Mean Square" which is just another term for "Variance"**

Formula for MS_{Between}

$$MS_{\text{Between}} = \frac{SS_{\text{Between}}}{df_{\text{Between}}}$$

Formula for MS_{Within}

$$MS_{\text{Within}} = \frac{SS_{\text{Within}}}{df_{\text{Within}}}$$

SOLVING FOR THE NUMERATOR OF THE F-RATIO: MS-BETWEEN

Formula for MS_{Between}

$$MS_{\text{Between}} = \frac{SS_{\text{Between}}}{df_{\text{Between}}}$$

Formula for SS_{Between}

$$SS_{\text{Between}} = \Sigma N_{\text{Group}}(M_{\text{Group}} - M_{\text{grand}})^2$$

Formula for df_{Between}

$$df_{\text{Between}} = (k - 1)$$

From our Example:

M_{Group}	M_{Group} - M_{Grand}	$(M_{\text{Group}} - M_{\text{Grand}})^2$	$N_{\text{Group}}(M_{\text{Group}} - M_{\text{grand}})^2$
7	2	4	16
3	-2	4	16
5	0	0	0

$M = 5$

$$SS_{\text{Between}} = \Sigma N_{\text{Group}}(M_{\text{Group}} - M_{\text{grand}})^2 = 32$$

SOLVING FOR THE NUMERATOR OF THE F-RATIO: MS-BETWEEN

Example cont.:

M_{Group}	M_{Group} - M_{Grand}	$(M_{\text{Group}} - M_{\text{Grand}})^2$	$N_{\text{Group}}(M_{\text{Group}} - M_{\text{grand}})^2$
7	2	4	16
3	-2	4	16
5	0	0	0

$M = 5$

$SS_{\text{Between}} = \Sigma N_{\text{Group}}(M_{\text{Group}} - M_{\text{grand}})^2 = 32$

$df_{\text{Between}} = k - 1 = 3 - 1 = 2$

- Solve for MS_{Between}
  - $MS_{\text{Between}} = \dfrac{SS_{\text{Between}}}{df_{\text{Between}}} = \dfrac{32}{2} = 16$

SOLVING FOR THE DENOMINATOR OF THE F-RATIO: MS-WITHIN

Formula for MS_{Within}

$$MS_{\text{Within}} = \frac{SS_{\text{Within}}}{df_{\text{Within}}}$$

Formula for SS_{Within}

$$SS_{\text{Within}} = \Sigma(X_{\text{Group}} - M_{\text{Group}})^2$$

Formula for df_{Within}

$$df_{\text{Within}} = (N - k)$$

From our Example:

Positive	Negative	Neutral
7	3	5
7	4	6
8	3	5
6	2	4

$M_1 = 7.00$ $M_2 = 3.00$ $M_3 = 5.00$

SOLVING FOR THE DENOMINATOR OF THE F-RATIO: MS-WITHIN

Positive	**Group Means**	**$(X_{Group} - M_{Group})$**	**$(X_{Group} - M_{Group})^2$**
7		**0**	**0**
7		0	0
8	$M_1 = 7.00$	1	1
6		-1	1
Negative			
3		0	0
4	$M_2 = 3.00$	1	1
3		0	0
2		-1	1
Neutral			
5		0	0
6	$M_3 = 5.00$	1	1
5		0	0
4		-1	1

$$SS_{Within} = \Sigma(X_{Group} - M_{group})^2 = 6$$

SOLVING FOR THE DENOMINATOR OF THE F-RATIO: MS-WITHIN

$SS_{\text{Within}} = \Sigma(\text{X}_{\text{Group}} - M_{\text{group}})^2 = 6$

$df_{\text{Within}} = N - k \;\; = 12 - 3 = 9$

- Solve for MS_{Within}
  - $MS_{\text{Within}} = \dfrac{SS_{\text{Within}}}{df_{Within}} = \dfrac{6}{9} = 0.67$

CALCULATING THE TEST STATISTIC

- The formula for calculating the test-statistic for a Between-Subjects, One-Way ANOVA is:

$$F = \frac{MS_{\text{Between}}}{MS_{\text{Within}}}$$

$$MS_{\text{Between}} = \frac{SS_{\text{Between}}}{df_{\text{Between}}} = 16$$

$$MS_{\text{Within}} = \frac{SS_{\text{Within}}}{df_{Within}} = 0.67$$

- Solve for our obtained F-ratio:

- $F = \frac{MS_{\text{Between}}}{MS_{\text{Within}}} = \frac{16}{0.67} = 23.88$

SUMMARY SOURCE TABLE OF CALCULATIONS

Source	SS	DF	MS	F-obtained	F-Critical	Prob of H_o
Between	$\Sigma N_{Group}(M_{Group} - M_{grand})^2$	$k - 1$	$\frac{SS_{Between}}{df_{Between}}$	$\frac{MS_{Between}}{MS_{Within}}$	$df_{Between} = (k - 1)$ $df_{Within} = (N - k)$	$p < .05$ *or* $p > .05$
Within	$\Sigma(X_{Group} - M_{Group})^2$	$N - k$	$\frac{SS_{Within}}{df_{Within}}$			
Total	$\Sigma(X - M_{Grand})^2$	$N - 1$				

N_{Group} = the number of scores within a group
M_{Group} = the sample mean for a group
M_{Grand} = the mean of *all* the scores

X_{Group} = the scores within a group
M_{Group} = the sample mean for a group

X = an individual score
M_{Grand} = the mean of *all* the scores

$SS_{Total} = SS_{Between} + SS_{Within}$

$df_{Total} = df_{Between} + df_{Within}$

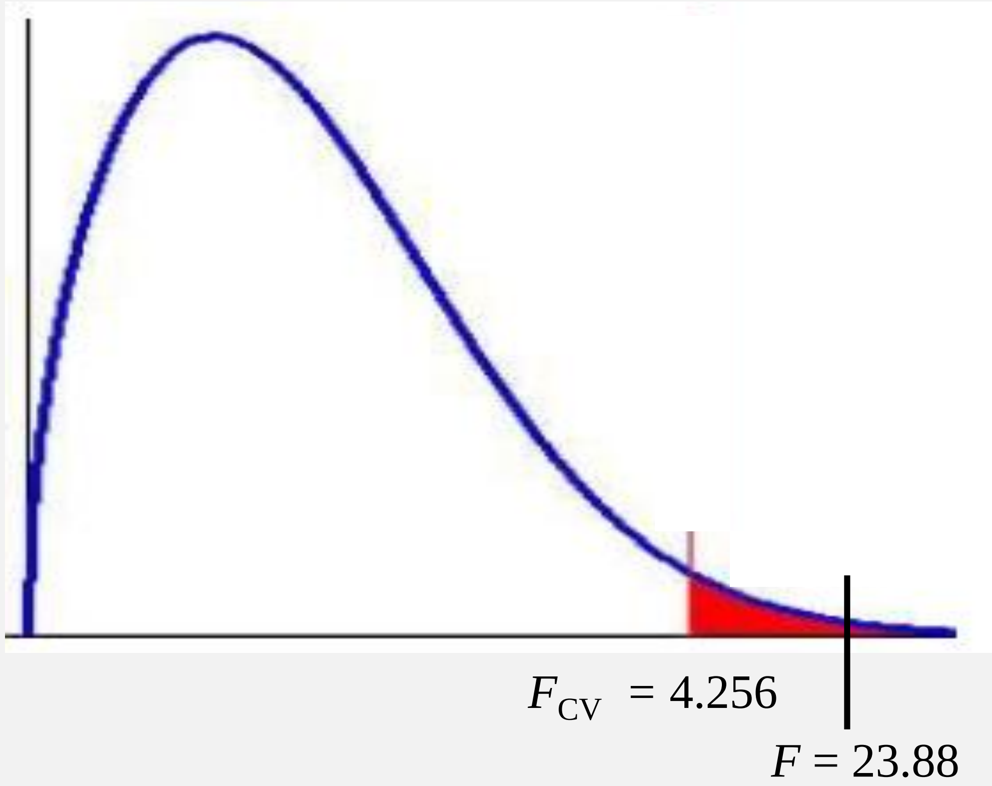
INTERPRETATION & EFFECT SIZE

Step 6: Interpretation

a) Decide whether to reject or fail to reject the null hypothesis

F-critical value:
- $F_{\text{CV}} = 4.256$

Obtained F-statistic:
- $F = 23.88$



INTERPRETATION & EFFECT SIZE

b) If the omnibus test (the overall ANOVA) is significant, that is an indication that…

- **H1: At least one population mean is different from the others.**

- In order to know which *particular* means are significantly different from each other, we need to perform a **post-hoc test**!

Positive	Negative	Neutral
7	3	5
7	4	6
8	3	5
6	2	4

$M_1 = 7.00$ $M_2 = 3.00$ $M_3 = 5.00$

INTERPRETATION & EFFECT SIZE

c) Effect Size

- The formula for r^2, the percentage of variability in the dependent variable accounted for by the explanatory variable, is:

$$r^2 = \frac{SS_{Between}}{SS_{Total}} \text{ x } 100$$

Note: Another term for r^2 is η^2 (eta-squared)

For our example:

r^2 = (32 / (32 + 6)) x 100 = (32 / 38) x 100 = 84.21%

STEP 7: POST-HOC TESTS

- If the omnibus test (the overall ANOVA) is significant, that tells us that *at least one mean is different from another mean*.
  - How do we tell *which* means are different from each other?

- When the omnibus ANOVA is significant, perform a post-hoc test
  - A post-hoc test can be used to make multiple comparisons (i.e., M_1 vs M_2, M_2 vs M_3, etc.) <u>without</u> inflating Type I Error by making an adjustment to α

- There are a variety of types of post-hoc tests
  - Scheffé, Bonferroni-Dunn, Tukey HSD… and more!

POST-HOC TESTS: TUKEY HSD

To perform the Tukey HSD post-hoc test:

- Calculate HSD. Find q using the Tukey HSD Post-Hoc Test Table

$$HSD = q\sqrt{\frac{MS_{\text{Within}}}{n}}$$

TABLE 10.12 Part of Appendix Table 5, Table of Values of q for Use in the Tukey *HSD* Post-Hoc Test

	k		
df	**2**	**3**	**4**
2	6.09	8.33	9.80
3	4.50	5.91	6.83
4	3.93	5.04	5.76
5	3.64	4.60	5.22
6	3.46	4.34	4.90
7	3.34	4.17	4.68
8	3.26	4.04	4.53
9	3.20	3.95	4.42
10	3.15	3.88	4.33

A q value is needed to calculate a Tukey *HSD* post-hoc test. The q value for a between-subjects, one-way ANOVA is found at the intersection of the column for k, the number of groups in the ANOVA, and the row where $df = df_{\text{Within}}$.

- q = the q-value from the Tukey post-hoc test table
- $MS_{\text{Within}} = MS_{\text{Within}}$ from the omnibus ANOVA test
- n = sample size of the smallest group

POST-HOC TESTS: TUKEY HSD

Tukey HSD post-hoc test cont.:

- The value of *HSD* is the amount by which two means must differ in order to be significantly different from one another
- Example:
  - $HSD = 3.95\sqrt{\frac{0.67}{4}} = 1.62$
  - $M_1 - M_2 = 7 - 3 = 4$
    - Significantly different
  - $M_1 - M_3 = 7 - 5 = 2$
    - Significantly different
  - $M_2 - M_3 = 5 - 3 = -2$
    - Significantly different

Positive	**Negative**	**Neutral**
7	3	5
7	4	6
8	3	5
6	2	4

$M_1 = 7.00$ $M_2 = 3.00$ $M_3 = 5.00$

APA STYLE SUMMARY

Using a between-subjects, one-way ANOVA, we found that there was a significant difference in math performance among people who were given positive (M = 7.00, s = 0.82), negative (M = 3.00, s = .82) and neutral (M = 5.00, s = 0.82) feedback, $F(2, 9)$ = 23.88, p < .001, r^2 = 84.21%. Specifically, using a Tukey HSD post-hoc test, we found that people given positive feedback had significantly higher math scores compared to people given negative feedback, p < .001, and compared to people given neutral feedback, p = .018. People given negative feedback also had significantly lower math scores compared to people given neutral feedback, p = .018.

PRACTICE PROBLEMS

A consumer researcher gave consumers a sample shampoo. After using the shampoo, each consumer used an interval-level scale from 0 to 100 to rate their satisfaction with it (higher scores meant greater satisfaction). Consumers were randomly assigned to three groups: (1) receive a store brand shampoo in a bottle labeled store brand, (2) receive a premium brand shampoo in a bottle labeled premium brand, or (3) receive store brand shampoo in a bottle labeled premium brand. Here are the participants’ final satisfaction scores:

Store Brand, Store Bottle	Premium Brand, Premium Bottle	Store Brand, Premium Bottle
70	85	85
65	90	80
65	95	90
60	90	85
M = 65.00 s = 4.08	M = 90.00 s = 4.08	M = 85.00 s = 4.08

Using the six steps of hypothesis testing and the appropriate test, see what the researcher should conclude about the effect of the shampoo condition on satisfaction ratings.

PRACTICE PROBLEMS

Question: Which of the following means are significantly different from each other?

a) Store Brand/Store Bottle and Premium Brand/Premium Bottle
b) Store Brand/Store Bottle and Store Brand/Premium Bottle
c) Premium Brand/Premium Bottle and Store Brand/Premium Bottle
d) Both a and b
e) Both a and c
f) Both b and c

PRACTICE PROBLEMS

Question: Which of the following means are significantly different from each other?

a) Store Brand/Store Bottle and Premium Brand/Premium Bottle
b) Store Brand/Store Bottle and Store Brand/Premium Bottle
c) Premium Brand/Premium Bottle and Store Brand/Premium Bottle
d) Both a and b
e) Both a and c
f) Both b and c

WEEK 5 REMINDERS

- Homework 4
  - Due Monday, 7/26, 11:59pm

- Jamovi HW5
  - Due Thursday, 7/29, 11:59pm

- Midterm 2
  - This Wednesday, July 28th
  - Available from 8:00am to 11:59pm
  - Time limit: 90 minutes