One-Way Analysis of Variance: ANOVA

Dr. J. Kyle Roberts

Southern Methodist University
Simmons School of Education and Human Development
Department of Teaching and Learning
Background to ANOVA

- Recall from the Independent Samples $t$ test that we are testing to see if two means drawn from independent samples are statistically significantly different. With such, we are testing:

$$H_0 : \overline{X}_1 = \overline{X}_2$$

- where $\overline{X}_1$ and $\overline{X}_2$ are two sample means drawn from independent populations.

- While this is helpful for when $k = 2$, we must use alternative techniques when $k > 2$.

- In this case, we must use an $F$ test instead of the previously used $t$ test since we now have two sources of $df$s.
Hypothesis Test in One-Way ANOVA

- In the ANOVA, we refer to the number of independent variables as either “ways” or “factors.”
- The number of divisions of each “way” is referred to as “levels.”
- Therefore, an analysis in which we are testing for the mean difference of 3 recognized ethnicities on a single dependent variable would be referred to as a one-way ANOVA with 3 levels.
- The null hypothesis for this test could then be written as:
  \[ H_0 : \bar{Y}_1 = \bar{Y}_2 = \bar{Y}_3 \]

assuming \( \bar{Y}_1, \bar{Y}_2 \) and \( \bar{Y}_3 \) are random samples drawn from \( Y_{i1} \sim \mathcal{N}(\mu_1, \sigma^2) \), \( Y_{i2} \sim \mathcal{N}(\mu_2, \sigma^2) \) and \( Y_{i3} \sim \mathcal{N}(\mu_3, \sigma^2) \) distributions, respectively. Or alternatively as \( Y_{ij} \sim \mathcal{N}(\mu_j, \sigma^2) \).
The One-Way ANOVA Summary Table

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>p</th>
<th>( \eta^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between</td>
<td>( SS_t - SS_w )</td>
<td>( k - 1 )</td>
<td>( \frac{SS_b}{df_b} )</td>
<td>( \frac{MS_b}{MS_w} )</td>
<td>( \frac{SS_b}{SS_t} )</td>
<td></td>
</tr>
<tr>
<td>Within</td>
<td>( SS_t - SS_b )</td>
<td>( df_t - df_b )</td>
<td>( \frac{SS_w}{df_w} )</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>( SS_b + SS_w )</td>
<td>( n - 1 )</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- The statistical significance of \( F \) can be obtained by computing the \( F \)-critical value. Determining statistical significance follows the same pattern for the \( t \) test only we have two sources of \( df \): between and within.

- For a one-way ANOVA with 5 levels and 50 people, the critical value of \( F \) at \( \alpha = 0.05 \) would be:

```r
> qf(0.95, 4, 45)
[1] 2.578739
```
**One-Way ANOVA Practice**

- Fill in the Missing Values Below

\[ F_{-\text{crit}} = 2.690 \]

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>p</th>
<th>( \eta^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between</td>
<td>50</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Within</td>
<td></td>
<td>30</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Fill in the Missing Values Below

\[ F_{-\text{crit}} = 2.922 \]

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>F</th>
<th>( \eta^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between</td>
<td>10</td>
<td></td>
<td></td>
<td>0.15</td>
<td></td>
</tr>
<tr>
<td>Within</td>
<td></td>
<td>30</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>200</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Computing the Probability of $F$

- Based on our $F$ from the ANOVA Summary Table previously, we can compute the probability of matching or exceeding the probability that there are no differences between the groups/levels.
- We compute this in R by:

```r
> pf(7.5, 4, 30, lower = FALSE)
[1] 0.0002593994
> pf(1.765, 3, 30, lower = FALSE)
[1] 0.1750938
```
The “effects” form of the model

• An alternative representation of the model for testing mean differences among the way/factor is:

\[ Y_{ij} = \mu + \alpha_j + \epsilon_{ij}, \quad \epsilon_{ij} \sim \mathcal{N}(0, \sigma^2), \quad j = 1, \ldots, k, \quad i = 1, \ldots, n_j \]

This is a “signal + noise” form like the simple linear model.

• \( \alpha_j \) is called the effect of level \( j \) of the factor.

• This means that an individual’s score (\( Y_{ij} \)) can be thought of as the sum of the grand mean (\( \mu \)) plus that individual group’s deviation around the grand mean (\( \alpha_j \)) and their own deviation around their group mean (\( \epsilon_{ij} \)).
The Null Hypothesis for One-Way ANOVA

- In the effects form we write the null and alternative hypotheses as

\[ H_0 : \alpha_1 = \alpha_2 = \cdots = \alpha_k = 0 \]
\[ H_a : \text{at least one } \alpha_i \neq 0 \]

- Note that in order to reject \( H_0 \) we only need to have one mean different from the other means.

- This null may also be written as

\[ H_0 : \bar{Y}_1 = \bar{Y}_2 = \cdots = \bar{Y}_k \]
\[ H_a : \text{at least one } \bar{Y}_j \neq \text{ any other } \bar{Y}_j \]
Computation of Sums of Squares

- Computation for $SS_{total}$.

$$SS_t = \sum_{i=1}^{N} (Y_{ij} - \bar{Y})^2$$

- Computation for $SS_{between}$.

$$SS_b = \sum_{j=1}^{k} n_j (\bar{Y}_j - \bar{Y})^2$$

- Computation for $SS_{within}$.

$$SS_w = \sum_{i=1}^{N} (Y_{ij} - \bar{Y}_j)^2$$
Graphical Representation of Computation of $SS$
Understanding $\eta^2$ as a Measure of Effect Size

• Recall that for the $t$ test, we measure the size of the effect of the mean difference as a function of the standardized difference between the two means. This distance we called $d$ or $\Delta$.

• In the case of the ANOVA, we cannot compute a standardized difference of movement since we have no basis by which we would compute the difference in a case where $k > 2$.

• Instead we are going to compute the amount of variance in the dependent variable that is “explained” by the grouping variable.

• We compute this explained variance as $\eta^2 = SS_b/SS_t$, or as a ratio of the $SS_t$ that is in $SS_b$ (thus we can think of this as a percent since $SS_b$ will never exceed $SS_t$).
Understanding $\eta^2$ cont.

- Consider the following two cases where $k = 3$.
- Assessment 1

---

- Assessment 2

---
Running ANOVA in R

- Consider the following dataset:
- Ethnicity 1-8,7,6,7,9,11,13
- Ethnicity 2-11,13,14,18,17,14,12,15
- Ethnicity 3-14,13,15,15,20,21,22

```r
> ethdata <- data.frame(ethn = factor(rep(1:3, c(7, 8, 7))), score = c(8, 7, 6, 7, 9, 11, 13, 11, 13, 14, 18, 17, 14, 12, 15, 14, 13, 15, 15, 20, 21, 22))
> tapply(ethdata$score, ethdata$ethn, mean)
   1     2     3
8.714286 14.250000 17.142857
> tapply(ethdata$score, ethdata$ethn, sd)
   1     2     3
2.497618 2.375470 3.716117
```
Assumptions for ANOVA

• Relatively the same number of people in each level
• Normality in the population for each of the levels
• Homogeneity of variance

\[ H_0 : \sigma_1^2 = \sigma_2^2 = \cdots = \sigma_k^2 \]

> bartlett.test(score ~ ethn, ethdata)

Bartlett test of homogeneity of variances
data:  score by ethn
Bartlett’s K-squared = 1.5034, df = 2, p-value = 0.4716
Graphing Ethnicity Data

```r
> boxplot(score ~ ethn, ethdata, ylab = "Ethnicity",
+          xlab = "Score", horizontal = TRUE)
```

```r
> print(dotplot(ethn ~ score, ethdata, type = c("p",
+              "a")))
```

![Boxplot of Score by Ethnicity](image1)

![Dotplot of Ethnicity by Score](image2)
> m1 <- aov(score ~ ethn, ethdata)
> anova(m1)

Analysis of Variance Table
Response: score

<table>
<thead>
<tr>
<th>Df</th>
<th>Sum Sq</th>
<th>Mean Sq</th>
<th>F value</th>
<th>Pr(&gt;F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ethn</td>
<td>2</td>
<td>257.53</td>
<td>128.77</td>
<td>15.312</td>
</tr>
<tr>
<td>Residuals</td>
<td>19</td>
<td>159.79</td>
<td>8.41</td>
<td></td>
</tr>
</tbody>
</table>

• Note that R puts the ANOVA summary table in a slightly different format than we will report. There is no “Total” row and there is no $\eta^2$.

• For this case $\eta^2$ is computed as $\frac{257.53}{257.53 + 159.79} = 0.617$. 
Testing Pairwise Comparisons

- If we reject $H_0$ for the F test, we conclude that there are significant differences between the groups. Usually we follow up and determine which groups are significantly different.
- If we test all possible pairs of groups for significant differences we will perform $\binom{k}{2} = k(k - 1)/2$ separate tests. We say we are doing *multiple comparisons*.
- Using $t$ tests without any adjustment for the multiple comparisons inflates the probability of declaring a significant difference when there isn’t one.
- There are several techniques for adjusting the tests. One of the most common is Tukey’s “honest significant difference” (function *TukeyHSD*), which is based on the Studentized range.
Tukey’s HSD for ANOVA

- The Tukey’s HSD provides a correction factor to the pairwise comparisons such that the p-value is slightly inflated.
- These adjustments are based on the number of comparisons.

```r
> TukeyHSD(m1)

Tukey multiple comparisons of means
95% family-wise confidence level
Fit: aov(formula = score ~ ethn, data = ethdata)

$ethn

               diff      lwr      upr   p adj
2-1  5.535714  1.722823  9.348606 0.004245
3-1  8.428571  4.490634 12.366509 0.000086
3-2  2.892857 -0.920034  6.705749 0.158296
```
Plotting Tukey HSD

> plot(TukeyHSD(m1))

95% family–wise confidence level

Differences in mean levels of ethn
Proving the Importance of the Way

• Consider the following dataset in which there are two separate ways $g_1$ and $g_2$ for the $dv$.

```r
newtrial <- data.frame(dv = c(1:9, 11), g1 = rep(1:2, + 5), g2 = rep(1:2, each = 5))
```

• Run two separate ANOVAS for both $g_1$ and $g_2$ on $dv$. Why is it that the first ANOVA was not statistically significant and the second one was when we used the *same* dependent variable?
In-Class Practice Example with 5 Levels

- Create the following dataset in R.
- Test all assumptions and run all appropriate post-hoc tests.

<table>
<thead>
<tr>
<th>Program 1</th>
<th>Program 2</th>
<th>Program 3</th>
<th>Program 4</th>
<th>Program 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>32</td>
<td>31</td>
<td>43</td>
<td>44</td>
</tr>
<tr>
<td>33</td>
<td>35</td>
<td>34</td>
<td>47</td>
<td>50</td>
</tr>
<tr>
<td>31</td>
<td>28</td>
<td>33</td>
<td>53</td>
<td>50</td>
</tr>
<tr>
<td>25</td>
<td>29</td>
<td>32</td>
<td>54</td>
<td>49</td>
</tr>
<tr>
<td>26</td>
<td>19</td>
<td>29</td>
<td>52</td>
<td>47</td>
</tr>
<tr>
<td>29</td>
<td>20</td>
<td>30</td>
<td>55</td>
<td>49</td>
</tr>
<tr>
<td>29</td>
<td>20</td>
<td>31</td>
<td>45</td>
<td>49</td>
</tr>
<tr>
<td>31</td>
<td></td>
<td>31</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Out of Class Homework Assignment

• Create heuristic data for a one-way ANOVA with 4 levels. You must have at least 10 people in each level.

1. Create the first ANOVA such that results are statistically significant and in which the post-hoc tests reveal:

\[ \bar{Y}_1 \neq \bar{Y}_2, \bar{Y}_1 \neq \bar{Y}_3, \bar{Y}_1 \neq \bar{Y}_4 \]

but

\[ \bar{Y}_2 = \bar{Y}_3 = \bar{Y}_4 \]

2. Create the second ANOVA such that results are statistically significant and in which the post-hoc tests reveal:

\[ \bar{Y}_1 = \bar{Y}_2 \text{ and } \bar{Y}_3 = \bar{Y}_4 \]

but

\[ \bar{Y}_1 \text{ and } \bar{Y}_2 \neq \bar{Y}_3 \text{ and } \bar{Y}_4 \]