11.3: Hypotheses in ANOVA

So far we have seen what ANOVA is used for, why we use it, and how we use it. Now we can turn to the formal hypotheses we will be testing. As with before, we have a null and a research hypothesis to lay out.

Research Hypotheses

Our research hypothesis for ANOVA is more complex with more than two groups. Let’s take a look at it and then dive deeper into what it means.

What the ANOVA tests is whether there is a difference between any one set of means, but usually we still have expected directions of what means we think will be bigger than what other means. Let's work out an example. Let's say that my IV is mindset, and the three groups (levels) are:

- Growth Mindset
- Mixed Mindset (some Growth ideas and some Fixed ideas)
- Fixed Mindset

If we are measuring passing rates in math, we could write this all out in one sentence and one line of symbols:

- Research Hypothesis: Students with Growth Mindset with have higher average passing rates in math than students with either a mixed mindset or Fixed Mindset, but Fixed Mindset will have similar average passing rates to students with mixed mindset.
- Symbols: $\overline{X}_G > \overline{X}_M = \overline{X}_F$

But it ends up being easier to write out each pair of means:

- Research Hypothesis: Students with Growth Mindset with have higher average passing rates in math than students with a mixed mindset. Students with Growth Mindset with have higher average passing rates in math than students with a Fixed Mindset. Students with a Fixed Mindset will have similar average passing rates to students with mixed mindset.
- Symbols:
  - $\overline{X}_G > \overline{X}_M$
  - $\overline{X}_G > \overline{X}_F$
  - $\overline{X}_M = \overline{X}_F$

What you might notice is that one of these looks like a null hypothesis (no difference between the means)! And that is okay, as long as the research hypothesis predicts that at least one mean will differ from at least one other mean. It doesn't matter what order you list these means in; it helps to match the research hypothesis, but it's really to help you conceptualize the relationships that you are predicting so put it in the order that makes the most sense to you!

Why is it better to list out each pair of means? Well, look at this research hypothesis:

- Research Hypothesis: Students with Growth Mindset with have a similar average passing rate in math as students with a mixed mindset. Students with Growth Mindset with have higher average passing rates in math than students with a Fixed Mindset. Students with a Fixed Mindset will have similar average passing rates to students with mixed mindset.
- Symbols:
  - $\overline{X}_G = \overline{X}_M$
  - $\overline{X}_G > \overline{X}_F$
  - $\overline{X}_M = \overline{X}_F$

If you try to write that out in one line of symbols, it'll get confusing because you won't be able to easily show all three predictions. And if you have more than three groups, many research hypotheses won't be able to be represented in one line.

Another reason that this makes more sense is that each mean will be statistically compared with each other mean if the ANOVA results end up rejecting the null hypothesis. If you set up your research hypotheses this way in the first place (in pairs of means), then these pairwise comparisons make more sense later.

Null Hypotheses

Our null hypothesis is still the idea of “no difference” in our data. Because we have multiple group means, we simply list them out as equal to each other:

- Null Hypothesis: Students with Growth Mindset, mixed mindset, and Fixed Mindset will have similar average passing rates in math .
- Symbols: $\overline{X}_G = \overline{X}_M = \overline{X}_F$

You can list them all out, as well, but it's less necessary with a null hypothesis:

- Research Hypothesis: Students with Growth Mindset with have a similar average passing rate in math as students with a mixed mindset. Students with Growth Mindset with have a similar average passing rates in math than students with a Fixed Mindset. Students with a Fixed Mindset will have similar average passing rates to students with mixed mindset.
- Symbols:
  - $\overline{X}_G = \overline{X}_M$
  - $\overline{X}_G = \overline{X}_F$
  - $\overline{X}_M = \overline{X}_F$

Null Hypothesis Significance Testing

In our studies so far, when we've calculated an inferential test statistics, like a t-score, what do we do next? Compare it to a critical value in a table! And that's the same thing that we do with our calculated F-value. We compare our calculated value to our critical value to determine if we retain or reject the null hypothesis that all of the means are similar.

$$\textbf{(Critical} < \textbf{Calculated)} = \textbf{Reject null} = \textbf{At least one mean is different from at least one other mean.} = \boldsymbol{p < .05}$$

$$\textbf{(Critical} > \textbf{Calculated)} = \textbf{Retain null} = \textbf{All of the means are similar.} = \boldsymbol{p > .05}$$

What does Rejecting the Null Hypothesis Mean for a Research Hypothesis with Three or More Groups?

Remember when we rejected the null hypothesis when comparing two means with a t-test that we didn't have to do any additional comparisons; rejecting the null hypothesis with a t-test tells us that the two means are statistically significantly different, which means that the bigger mean was statistically significantly bigger. All we had to do was make sure that the means were in the direction that the research hypothesis predicted.

Unfortunately, with three or more group means, we do have to do additional statistical comparisons to test which means are statistically significantly different from which other means. The ANOVA only tells us that at least one mean is different from one other mean. So, rejecting the null hypothesis doesn't really tell us whether our research hypothesis is (fully) supported, partially supported, or not supported. When the null hypothesis is rejected, we will know that a difference exists somewhere, but we will not know where that difference is. Is Growth Mindset different from mixed mindset and Fixed Mindset, but mixed and Fixed are the same? Is Growth Mindset different from both mixed and Fixed Mindset? Are all three of them different from each other? And even if the means are different, are they different in the hypothesized direction? Does Growth Mindset always have a higher mean? We will come back to this issue later and see how to find out specific differences. For now, just remember that an ANOVA tests for any difference in group means, and it does not matter where that difference occurs. We must follow-up with any significant ANOVA to see which means are different from each other, and whether those mean differences (fully) support, partially support, or do not support the research hypothesis.

Table 11.3.1- Interpreting Null Hypotheses

REJECT THE NULL HYPOTHESIS	RETAIN THE NULL HYPOTHESIS
Small p-values **(p<.05)**	**Large p-values** **(p>.05)**
A small p-value means a small probability that ***all*** of the means are similar. **Suggesting that *at least one of the means* is different from at least one other mean…**	A large p-value means a large probability that all of the means are similar.

REJECT THE NULL HYPOTHESIS	RETAIN THE NULL HYPOTHESIS
We conclude that: • **At least one mean is different from one other mean.** • At least one group is *not* from the same population as the other groups.	We conclude that • The means for all of the groups are similar. • All of the groups *are* from the same population.
The calculated F is further from zero (more extreme) than the **critical F**. In other words, the calculated F is *bigger* than the **critical F**. (Draw the standard normal curve and mark the calculated F and the **critical F** to help visualize this.)	The calculated F is closer to zero (less extreme) than the **critical F**. In other words, the calculated F is *smaller* than the **critical F.** (Draw the standard normal curve and mark the calculated F and the **critical F** to help visualize this.)
Reject the null hypothesis (which says that all of the means are similar).	Retain (or fail to reject) the null hypothesis (which says that the all of the means are similar).

Support the Research Hypothesis? MAYBE.
Look at the actual means:
• *Support the Research Hypothesis if the means are in the directions that were hypothesized.*
  ◦ *The mean of the group that you said would be bigger, really is bigger;*
  ◦ *The mean of the group that you said would be smaller really is smaller;*
  ◦ *The means of the groups that you said would be similar are actually similar.*
• *Partial support of the Research Hypothesis if some of the means are in the directions that were hypothesized, but some aren't.*
• *Do not support the Research Hypothesis if none of the means are in the direction that were hypothesized.*
	Do not support the Research Hypothesis (because all of the means are similar).

Statistical sentence: F(df) = F-calc, p<.05 (fill in the df and the calculated F)	Statistical sentence: F(df) = F-calc, p>.05 (fill in the df and the calculated F)