11.6: Practice on Mindset Data

Okay, here's your time to practice the whole process of a Between Groups ANOVA, including each Sum of Squares formula and the pairwise comparisons!

Scenario

Dr. MO helped student researchers try to improve the mindset of their fellow community college students (If you need a refresher, check out the Growth Mindset section, or search online for "growth mindset"). Three different professors (all women) tried three different interventions (shown below) to see if they could improve mindset by the end of the semester. The three intervention levels were:

- No Intervention: This was the comparison control group. It was a behavioral statistics course (like you are taking right now!). The instructor had students complete the Mindset Quiz as part of class related to surveys and data collection, but the professor never explained what Growth Mindset was or how it helps students succeed. No class activities or assignments were related to Growth Mindset.
- Minimal Intervention: This was an Early Childhood Education class that explained what Growth Mindset was, and how it could help both the students themselves, and the children that they would be taking care of or teaching. The course had a few assignments and class activities related to Growth Mindset.
- Super Intervention: This was a History course. This professor explained Growth Mindset and how it could help them succeed, as well as weekly readings, discussions, and activities related to Growth Mindset.

Students who enrolled in these courses were participants in one of the three different interventions, and filled out the Mindset Quiz at the end of the semester. Mindset Quiz scores can range from 20 to 60, with higher scores showing that the student has more Growth Mindset (and less Fixed Mindset). At the end of the semester, 10 students in each course completed the Mindset Quiz and had attended most classes throughout the semester to experience any activities related to mindset (N = 30).

After reading the scenario, can you describe the sample, population, IV, and DV?

? Exercise 11.6.1

Answer the following questions to understand the variables and groups that we are working with.

1. Who is the sample?
2. Who do might be the population?
3. What is the IV (groups being compared)?
4. What is the DV (quantitative variable being measured)?

Answer

1. 30 community college students.
2. Community college students? Community college students with women professors?
3. Intervention levels: None, Minimal, Super
4. Mindset Quiz

Step 1: State the Hypotheses

Using the following the research question and the descriptive statistics in Table 11.6.1, what could be the research hypothesis?

> **Research Question**: The student researchers were interested to know if Growth Mindset could be taught, and how much time the instructor needed to devote to these intervention activities to improve growth mindset.

Table 11.6.1- Descriptive Statistics for each IV level and the Total

	N:	Mean:	SD:	Median:
No Intervention	10	42.70	8.18	43.00
Minimal Intervention	10	44.40	5.87	43.00
Super Intervention	10	51.10	4.93	50.50

	N:	Mean:	SD:	Median:
Total	30	46.07	7.25	48.00

Note

You are encourages to use the raw data found in Table 11.6.2 in Example 11.6.2 to practice calculating the means, standard deviations, and medians yourself!

What's your research hypothesis? Remember, the research hypothesis should predict how all levels of the IV related to all other levels of the IV.

Exercise 11.6.2

What is a research hypothesis in words and symbols?

Answer

- Research Hypothesis: Students in the No Intervention group average score will be lower on the Mindset Quiz than both students in the Minimal Intervention Group and the Super Intervention group. Similarly, students in the Minimal Intervention group average score will be lower than students in the Super Intervention group.
- Symbols:
  - $\overline{X}_N < \overline{X}_M$
  - $\overline{X}_N < \overline{X}_S$
  - $\overline{X}_M < \overline{X}_S$

  Your research hypothesis might be slightly different than this one. Just make sure that at least one mean is predicted to be different than one other mean.

What is the null hypothesis with this scenario?

Exercise 11.6.3

State the null hypothesis in words and symbols.

Answer

- Null Hypothesis: Students in the No Intervention group will have a similar average score on the Mindset Quiz as both students in the Minimal Intervention Group and the Super Intervention group. Similarly, students in the Minimal Intervention group will have a similar average score as students in the Super Intervention group.
- Symbols:
  - $\overline{X}_N = \overline{X}_M$
  - $\overline{X}_N = \overline{X}_S$
  - $\overline{X}_M = \overline{X}_S$

Step 2: Find the Critical Values

This step might be easier after you've completed the ANOVA Summary Table because you will have the Degrees of Freedom for both groups, but let's try it now.

Exercise 11.6.3

Using the Critical Values of F Table by going to the page or finding the link in the Common Critical Values page at the end of the textbook, what is the critical value at α = 0.05 with three groups and a total of 30 people?

Answer

Critical F(2,27)=3.35

The first Degree of Freedom (2), for the numerator (df_B) is found through: k–1, with k being the number of groups.

The second Degree of Freedom (27), for the denominator (df_W) is found through $N-k(30-3=27)$. A common mistake is to use the Degrees of Freedom of the numerator instead of the number of groups..

Step 3: Compute the Test Statistic

If you will never have to calculate the Sums of Squares by hand, skip this part and just fill in the ANOVA Summary Table (Table 11.6.4) at the end of this section. If you are practicing the Sums of Squares, each Sum of Square will have it's own Example. Heads up, to do all of these can take about an hour!

Let's go!

Example 11.6.1

Calculate the Between Groups Sums of Squares.

Solution

$$SS_B = \sum_{EachGroup}\left[\left(\overline{X}_{group}-\overline{X}_T\right)^2 * (n_{group})\right]$$

The $\sum_{EachGroup}$ means that you do everything following that for each intervention level, then add them all together. Let's start with what's inside the brackets for the No Intervention group.

$\sum_{EachGroup}$

$$SS_B = \sum_{EachGroup}\left[\left(\overline{X}_{group}-\overline{X}_T\right)^2 * (n_{group})\right]$$

$$\text{No Intervention} = \left[\left(\overline{X}_{group}-\overline{X}_T\right)^2 * (n_{group})\right]$$

$\overline{X}_{group}$ is asking for the mean of the group that we're looking at, so the No Intervention group right now. $\overline{X}_T$ is asking for the total mean, the mean for all 30 scores. Both of these means were provided in Table 11.6.1. The number of scores in the group that we're looking at right now is what n_{group} is asking.

So let's plug those all in!

$$\text{No Intervention} = \left[(42.70-46.07)^2 * (10)\right]$$

$$\text{No Intervention} = \left[(-3.37)^2 * (10)\right]$$

The mean of the group minus the mean of the sample should be negative, but that negative sign goes away when we square it:

$$\text{No Intervention} = [(11.36) * (10)]$$

$$\text{No Intervention} = [113.60]$$

And let's do that process two more times, once for the Minimal Intervention Group and once for the Super Intervention group, and then we'll add those three numbers together to get the SS_B.

$$\text{Minimal Intervention} = \left[\left(\overline{X}_{group}-\overline{X}_T\right)^2 * (n_{group})\right]$$

$$\text{Minimal Intervention} = \left[(44.40-46.07)^2 * (10)\right]$$

$$\text{Minimal Intervention} = \left[(-1.67)^2 * (10)\right]$$

$$\text{Minimal Intervention} = [(2.79) * (10)]$$

$$\text{Minimal Intervention} = [27.90]$$

Why don't you try to do it on your own for the Super Intervention group?

$$\text{Super Intervention} = \left[\left(\overline{X}_{group} - \overline{X}_T \right)^2 * (n_{group}) \right]$$

$$\text{Super Intervention} = [253.10]$$

Next step, add them all together! It's easy to forget this step, but the Sum of Squares ends up to be one number, so when you get lost or forget the next step, look back at the full formula:

$$SS_B = \sum_{EachGroup} \left[\left(\overline{X}_{group} - \overline{X}_T \right)^2 * (n_{group}) \right]$$

$$SS_B = \sum_{EachGroup} = [113.60 + 27.90 + 253.10]$$

$$SS_B = 394.60$$

You did it! Only two more Sums of Squares to go!

Example 11.6.2

Calculate the Within (Error) Sum of Squares.

Solution

$$SS_W = \sum \left[\left(X - \overline{X}_{group} \right)^2 \right]$$

As before, complete the calculations in the brackets first for each group, then add them all together. To complete the calculations in the brackets, it's easiest to use a table with all of the values since you subtract the mean of the group that you're working with from each score. Do you remember doing this with standard deviations?

Table 11.6.2 shows the group mean subtracted from each score in the column to the right of the raw scores, then that is squared in the column to the next right. The squared values are then summed for each group.

Table 11.6.2- Group Mean Subtracted from Each Score, Squared, then Summed

No Intervention	**minus group mean**	**squared**	Minimal Intervention	**minus group mean**	**squared**	Super Intervention	**minus group mean**	**squared**
30	-12.70	161.29	37	-7.40	54.76	43	-8.10	65.61
35	-7.70	59.29	38	-6.40	40.96	45	-6.10	37.21
36	-6.70	44.89	39	-5.40	29.16	48	-3.10	9.61
36	-6.70	44.89	42	-2.40	5.76	50	-1.10	1.21
43	0.30	0.09	42	-2.40	5.76	50	-1.10	1.21
43	0.30	0.09	44	-0.40	0.16	51	-0.10	0.01
49	6.30	39.69	48	3.60	12.96	54	2.90	8.41
49	6.30	39.69	48	3.60	12.96	56	4.90	24.01
53	10.30	106.09	53	8.60	73.96	56	4.90	24.01
53	10.30	106.09	53	8.60	73.96	58	6.90	47.61

No Intervention	minus group mean	squared	Minimal Intervention	minus group mean	squared	Super Intervention	minus group mean	squared
	Sum:	**602.10**		**Sum:**	**310.40**		**Sum:**	**218.90**
$\overline{X}_N$ = **42.7**	N/A	N/A	$\overline{X}_M$ = **44.4**	N/A	N/A	$\overline{X}_S$ = **51.1**	N/A	N/A

Now that we have the sum of each squared score of the subtraction, we can finish the formula:

$$SS_W = \sum_{EachGroup} \left[\sum \left(\left(X - \overline{X}_{group} \right)^2 \right) \right]$$

$$SS_W = \sum_{EachGroup} = [602.10 + 310.40 + 218.90]$$

$$SS_W = 1131.40$$

And, on to the final Sum of Squares!

Example 11.6.3

Calculate the Total Sum of Squares.

Solution

$$SS_T = \sum \left[\left(X - \overline{X}_T \right)^2 \right]$$

This formula is also saying to subtract a mean from each score, but this time we should be subtracting the Total mean ($\overline{X}_T$ = 46.07, found in Table 11.6.1). This is again easiest to compute in a table.

Table 11.6.3 shows the Total mean subtracted from each score in the column to the right of the raw scores, then that is squared in the column to the next right. The squared values are then summed for *all* of the scores.

Table 11.6.3- Total Mean Subtracted from Each Score, Squared, then Summed

IV Levels	Mindset Quiz Scores	minus Total mean	squared
No Intervention	30	-16.07	258.24
No Intervention	35	-11.07	122.54
No Intervention	36	-10.07	101.40
No Intervention	36	-10.07	101.40
No Intervention	43	-3.07	9.42
No Intervention	43	-3.07	9.42
No Intervention	49	2.93	8.58
No Intervention	49	2.93	8.58
No Intervention	53	6.93	48.02
No Intervention	53	6.93	48.02
Minimal Intervention	37	-9.07	82.26
Minimal Intervention	38	-8.07	65.12
Minimal Intervention	39	-7.07	49.98
Minimal Intervention	42	-4.07	16.56

IV Levels	Mindset Quiz Scores	minus Total mean	squared
Minimal Intervention	42	-4.07	16.56
Minimal Intervention	44	-2.07	4.28
Minimal Intervention	48	1.93	3.72
Minimal Intervention	48	1.93	3.72
Minimal Intervention	53	6.93	48.02
Minimal Intervention	53	6.93	48.02
Super Intervention	43	-3.07	9.42
Super Intervention	45	-1.07	1.14
Super Intervention	48	1.93	3.72
Super Intervention	50	3.93	15.44
Super Intervention	50	3.93	15.44
Super Intervention	51	4.93	24.30
Super Intervention	54	7.93	62.88
Super Intervention	56	9.93	98.60
Super Intervention	56	9.93	98.60
Super Intervention	58	11.93	142.32
σ	N/A	$\sigma = 1525.87$	

$$SS_T = \sum \left[\left(X - \overline{X}_T \right)^2 \right] = 1525.87$$

Yay! You finished the Sum of Squares! But before we move on, let's do the computation check to see if we did the Sums of Squares correctly: $SS_T = SS_B + SS_W$

Let's check:

$$SS_T = SS_B + SS_W = 394.40 + 1131.40 = 1526.00$$

This is close to our SS_T of 1525.87, so Dr. MO checked to make sure all of the calculations were done correctly with both a calculator and with Excel; it appears that this slight discrepancy is due to rounding differences.

But we're not done yet! Now, plug those three Sums of Squares into the ANOVA Summary Table (Table 11.6.4) so that you can fill out the rest of the table to calculate the final ANOVA F-value.

Table 11.6.4- ANOVA Summary Table with SS only

Source	SS	df	MS	F
Between	394.60			
Within	1131.40			
Total	1525.87			

Example 11.6.4

Fill out the rest of the ANOVA Summary Table.

Solution

Table 11.6.5- ANOVA Summary Table with Formulas

Source	SS	df	MS	F
Between	394.60	$k-1=3–1=2$	$\frac{SS_B}{df_B}=\frac{394.60}{2}=197.30$	$\frac{MS_B}{MS_W}=\frac{197.30}{41.90}=4.71$
Within	1131.40	$N-k=30–3=27$	$\frac{SS_W}{df_W}=\frac{1131.40}{27}=41.90$	leave blank
Total	1525.87	$N-1=30–1=29$	leave blank	leave blank

If you do a computation check (which you should) of the degrees of freedom, you find that:

$$df_T = df_B + df_W = 2 + 27 = 29$$

Since we did that part correctly, we can move on to the next step of the process.

Step 4: Make the Decision

Based on the completed ANOVA Summary Table (Table 11.6.5), our calculated F-score is 4.71. If you remember all the way back to Step 2, we found our critical F-score to be 3.35.

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

(Critical $>$ Calculated) = Retain null = All of the means are similar. = $p > .05$

Since our critical value is smaller than our calculated value, we reject the null hypothesis that all of the means are similar, which means that at least one mean is different from at least one other mean. That's not specific enough to evaluate whether our research hypothesis is correct or not, so on to pairwise comparisons!

Pairwise Comparisons

Let's start with finding the differences between each pair of means.

Example 11.6.5

What are the mean differences for each pair of means?

Solution

$$\overline{X}_N - \overline{X}_M = 42.70–44.40 = -1.7$$

$$\overline{X}_N - \overline{X}_S = 42.7–51.10 = -8.4$$

$$\overline{X}_M - \overline{X}_S = 44.40–51.10 = -6.7$$

Great! Now, let's compute Tukey's HSD to see if any of these mean differences are big enough to say that they are statistically significantly different.

Example 11.6.6

Using the q-value for our degrees of freedom from the Alpha = 0.05 table from Real-Statistics.com's q table, compute Tukey's HSD. If you will never have to calculate a post-hoc analysis by hand, skip this part and use the calculated Tukey's HSD to make your decisions and write-up the conclusion.

Solution

From the critical q table, we find that for our study with 3 groups and the degrees of freedom for the denominator (df_W = 27), the correct q-score to plug into the formula is 3.506. The other information is in the completed ANOVA Summary Table (Table 11.6.5).

$$HSD = q * \sqrt{\frac{MSw}{n_{group}}}$$

$$HSD = 3.506 * \sqrt{\frac{41.90}{10}}$$

$$HSD = 3.506 * \sqrt{4.19}$$

$$HSD = 3.506 * 2.05$$

$$HSD = 7.18$$

Now what?

Well, you have a critical value (Tukey's HSD) and some mean differences, let's compare them!

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

(Critical $>$ Calculated) = Retain null = All of the means are similar. = $p > .05$

Example 11.6.7

Using the Tukey's HSD of 7.18, which means are statistically significantly different from each other?

Solution

Our pairwise comparison critical value is 7.18. Of the absolute value of the mean differences, that means that the Super Intervention was statistically significantly higher than the No Intervention group's average score on the Mindset Quiz, but none of the other pairs of means were significantly different.

$$\overline{X}_N - \overline{X}_M = 42.70 - 44.40 = -1.7$$

$$\overline{X}_N - \overline{X}_S = 42.7 - 51.10 = -8.4$$

$$\overline{X}_M - \overline{X}_S = 44.40 - 51.10 = -6.7$$

How does this relate to the research hypothesis?

- $\overline{X}_N < \overline{X}_M$: This was NOT supported. Although the mean of the Minimal Intervention group was larger than the group with No Intervention, it was not significantly larger (because the critical value of Tukey's HSD of 7.18 was bigger than the absolute value of the mean difference between these two groups of -1.7)
- $\overline{X}_N < \overline{X}_S$: This was supported.
- $\overline{X}_M < \overline{X}_S$: This was NOT supported. The means were not different enough to reject the null hypothesis that the means are from different populations, so we conclude that the Minimal Intervention group had similar Mindset Quiz scores as the Super Intervention group.

This means that our research hypothesis was *partially supported*.

Write-Up

Okay, here's the big finish!

Exercise 11.6.4

Write a conclusion to describe the results of the analysis. Don't forget to include the four components necessary in any report of results.

Answer

The researchers hypothesized that students in the No Intervention group will score lower on average on the Mindset Quiz than both students in the Minimal Intervention Group and the Super Intervention group. Similarly, students in the Minimal Intervention group will score lower on average than students in the Super Intervention group.

This research hypothesis was partially supported ($F(2,27) = 4.71$, $p < 0.05$). The Super Intervention group's average Mindset Quiz score ($M = 51.10$) was higher than the No Intervention group ($M = 42.70$), but the the Minimal Intervention group's ($M = 44.40$) average Mindset Quiz scores were not statistically significantly different than either the Super Intervention group or the No Intervention group.

Did Dr. MO's write-up include all of the components?

You did it! Take a break and reward yourself!

Next up, we'll look at what to do if you don't think that your distribution is normally distributed when you have three or more groups…
