Post hoc analysis for ANOVA

Rick Balkin, Ph.D., LPC
Department of Counseling
Texas A&M University-Commerce
Rick_balkin@tamu-commerce.edu
What is a post hoc analysis?

*Post hoc analyses* are the statistical tests conducted to indicate exactly where statistically significant differences exist. They are only conducted when ANOVA results indicated statistical significance.
What is a post hoc analysis?

- When statistically significant differences are found in an ANOVA, what is really known is that statistically significant differences exist among the groups. There may be significant differences between two groups only, or significant differences may exist between more of the groups.
What is a post hoc analysis?

- In the example in your notepack, the results of $F$-test are indicative of statistically significant results among the four groups. Exactly what groups are different or where the statistically significant differences lies are not explained in ANOVA results. Further analysis is necessary.
What is a post hoc analysis?

- As a reminder, post hoc analyses are only conducted when ANOVA results are statistically significant.
- In a post hoc analysis, each group mean is compared to all other group means.
Methods of post hoc analysis

- There are many different ways to conduct post hoc analyses, but the most common and widely accepted method is the Tukey method of multiple comparisons (TMC).
- The TMC tests all possible pairwise comparisons.
- Compared to other post hoc methods, the TMC is strongest in controlling for type I error.

Balkin, R. S. (2008).
TMC

As mentioned before, the TMC tests all possible pairwise comparisons:

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

where \( C \) is the number of pairwise comparisons, and \( j \) is the number of groups. Thus, the number of pairwise comparisons for the ANOVA example is

\[ C = \frac{4(4 - 1)}{2} = 6 \]
So, six pairwise comparisons are possible among the four group means:

1. group 1 to group 2
2. group 1 to group 3
3. group 1 to group 4
4. group 2 to group 3
5. group 2 to group 4
6. group 3 to group 4
In order to determine statistically significant differences are evident in each pairwise comparison, the absolute value of the mean differences is divided by the standard error term:

\[ q = \frac{|\bar{X}_j - \bar{X}_K|}{\sqrt{\frac{MS_w}{2} \left( \frac{1}{n_j} + \frac{1}{n_K} \right)}} \]

where \( \bar{X}_j \) and \( \bar{X}_K \) are the group means being compared, \( MS_w \) is the mean square within, and \( n_j \) and \( n_k \) are the sample sizes in each comparison group.
When conducting a TMC by hand calculations, it is best to start with the two groups that have the largest mean differences.

When a group comparison is not statistically significant, there is no need to test further pairwise comparisons with smaller mean differences.

You can learn more about the hand computation of the TMC on p. 14 of the notepack.
### Multiple Comparisons

**Dependent Variable: Score**

**Tukey HSD**

<table>
<thead>
<tr>
<th>(I) Group</th>
<th>(J) Group</th>
<th>Mean Difference (I - J)</th>
<th>Std. Error</th>
<th>Sig.</th>
<th>95% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>-3.00</td>
<td>1.153</td>
<td>.082</td>
<td>-6.30 - .30</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>-1.00</td>
<td>1.153</td>
<td>.822</td>
<td>-4.30 - 2.30</td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>3.60(*)</td>
<td>1.153</td>
<td>.030</td>
<td>.30 - 6.90</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>3.00</td>
<td>1.153</td>
<td>.082</td>
<td>-.30 - 6.30</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>2.00</td>
<td>1.153</td>
<td>.339</td>
<td>-1.30 - 5.30</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>6.60(*)</td>
<td>1.153</td>
<td>.000</td>
<td>3.30 - 9.90</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>1.00</td>
<td>1.153</td>
<td>.822</td>
<td>-2.30 - 4.30</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>-2.00</td>
<td>1.153</td>
<td>.339</td>
<td>-5.30 - 1.30</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>4.60(*)</td>
<td>1.153</td>
<td>.005</td>
<td>1.30 - 7.90</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>-3.60(*)</td>
<td>1.153</td>
<td>.030</td>
<td>-6.90 - -.30</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>-6.60(*)</td>
<td>1.153</td>
<td>.000</td>
<td>-9.90 - -3.30</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>-4.60(*)</td>
<td>1.153</td>
<td>.005</td>
<td>-7.90 - 1.30</td>
</tr>
</tbody>
</table>

Based on observed means.

* The mean difference is significant at the .05 level.
A priori analyses

- A priori analyses refer to the researcher’s decision to evaluate only specific groups for comparison BEFORE the data is even collected.

- Reasons for this might include:
  - The desire to compare to only one group
  - The desire to combine groups and compare to other groups.
A priori analysis

- For example, let’s say I have 3 treatment groups and 1 control group, and I only want to compare the treatments to the control group. So, my group comparisons are as follows:
  - tx 1: control
  - tx 2: control
  - tx 3: control

- Remember that with fewer tests I increase statistical power, that is, I increase the likelihood of finding statistical significance and decrease the likelihood of making a type I error.
A priori analysis

- When I am only comparing one group to another, it is called a simple contrast comparison.
- When I combine groups, it is known as a complex contrast comparison.
- For example, I combine the means of all three treatment groups and compare to the control group.
A priori analysis

- The most powerful type of a priori analysis is called a planned orthogonal comparison.
- When contrast comparisons are conducted and
  - no overlapping variance is identified in the model
  - and the number of comparison is J-1 contrasts
  then we have an orthogonal design.
- While this method increases power, conditions requiring contrasts to be planned and orthogonal restricts its utility.
A priori analysis

- The interested reader can learn more about this type of design from the textbook and pages 15-17 in the notepack.
- These analyses are uncommon in the social sciences, and there are usually less complex ways to compare groups, such as a $t$-test.