Multiple Comparisons

October 18, 2019

After the ANOVA

For an ANOVA,

$$\begin{aligned} H_0 &: \quad \mu_1 = \mu_2 = \cdots = \mu_k \\ H_A &: \quad \mu_i \neq \mu_j \quad \text{for at least one pair } (i, j) \end{aligned}$$

If we reject H_0, we know that at least one mean differs... but we don't know where those differences lie.

Multiple Comparisons

Consider an ANOVA with three groups. If we reject H_0, there are three comparisons to make:

- group 1 and group 2
- group 1 and group 3
- group 2 and group 3

Multiple Comparisons

Most of this builds on techniques you already know!

- We compare groups using a two-sample t-test.
- But we need to modify the significance level.
- We also use a pooled standard deviation estimate.

Example

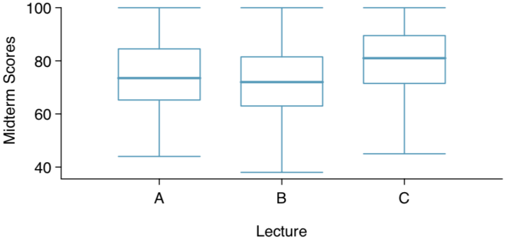
- A university offers 3 lectures for an introductory psychology course.
- A single professor offers 8am, 10am, and 3pm lectures.
- We want to know if the average midterm scores differ between these lectures.

We already wrote down hypotheses for this ANOVA.

Example

Are the ANOVA conditions satisfied?

Class i	A	B	C
n_i	58	55	51
$\bar{x}_i$	75.1	72.0	78.9
s_i	13.8	13.9	13.1



Example

Here is part of the ANOVA for this data (`R` output):

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
lecture			645.06		0.0330
Residuals			185.16		

Let's fill in the rest. What can we conclude?

Example

So at least one pair of means differ... but which?

- We need to correct for Type I error before running our t-tests.

Pooled Standard Error

Pooled standard error may be calculated as follows:

$$s_{pooled} = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2 + \cdots + (n_k - 1)s_k^2}{n - k}}$$

where $n = n_1 + n_2 + \cdots + n_k$ is the total number of observations.

Pooled Standard Error

If each group's sample size is equal,

$$s_{pooled} = \sqrt{\frac{s_1^2 + s_2^2 + \cdots + s_k^2}{k}}$$

The degrees of freedom for these t-tests will be $n - k$.

Example

Class i	A	B	C
n_i	58	55	51
$\bar{x}_i$	75.1	72.0	78.9
s_i	13.8	13.9	13.1

Let's calculate the pooled standard error for our exams.

Example

R also provides the pooled standard deviation estimate with the ANOVA output.

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
lecture	2	1290.11	645.06	3.48	0.0330
Residuals	161	29810.12	185.16		

$s_{pooled} = 13.61$ on $df = 161$

Significance Level Adjustments

- The final adjustment is to modify the significance level.
- When we do many pairwise comparisons, we increase our chances of Type I error.
- This correction will adjust our probability of Type I error.
- Adjusted significance levels will help ensure that the Type I error is no greater than α.

The Bonferroni Correction

Testing many pairs of groups is called **multiple comparisons**. The **Bonferroni correction** sets a new significance level α^*:

$$\alpha^* = \alpha/K$$

where K is the number of comparisons.

Example

Complete the pairwise comparisons for the three lectures.

Post ANOVA: Other Situations

If we fail to reject H_0 in an ANOVA

- No pairwise comparisons are necessary.
- (None will be significant.)

If we reject H_0 in an ANOVA

- Sometimes our pairwise comparisons won't show any significance.
- This does not invalidate the ANOVA results!

Multiple Comparisons

The Bonferroni correction is one method of many! Others include

- Tukey's Honest Significant Difference
- Scheffe's Method
- and others.

We will not learn these in detail.