Revisiting Multiple Comparisons

October 23, 2019

Week 5

- Lab B (Wed/Thurs) is cancelled.
- Lab A will be dedicated to midterm review.
- There will be no lab due.
- (But we might have something small for you to turn in during your Lab A.)

Post Hoc Tests

Another term for multiple comparisons is **post hoc tests** (analyses done after an ANOVA).

- For a factorial experiment, we have three possible sets of comparisons.
  1. Means for factor A
  2. Means for factor B
  3. Means for the factor level combinations (relating to interaction)

Post Hoc Tests

- For the previous example, we have factor level combinations
  1. Supervisor 1 with Day Shift
  2. Supervisor 1 with Swing Shift
  3. Supervisor 1 with Night Shift
  4. Supervisor 2 with Day Shift
  5. Supervisor 2 with Swing Shift
  6. Supervisor 2 with Night Shift
- This results in $\frac{k(k-1)}{2} = \frac{6 \times 5}{2} = 15$ possible pairs.

Revisiting Multiple Comparisons

While the Bonferroni correction is effective and a standard approach in many fields, it represents a "worse case scenario" approach.

- This means it can sometimes be too aggressive.
- Naturally, this may not always be ideal.
- We want other options!

Tukey's Honest Significant Difference

For Tukey's method for paired comparisons

- The Type I error will be α.
- The ANOVA assumptions are necessary.
  - But if we do these tests post-ANOVA, these are already satisfied.
- In addition, we must have independent sample means and equal group sizes.

Tukey's Honest Significant Difference

We will compare a value ω to differences in population means.

- This represents the **honest significant difference**.
- If $|\bar{x}_i - \bar{x}_j| > \omega$, we conclude that μ_i is different from μ_j.

Tukey's Honest Significant Difference

$$\omega = q_\alpha(k, df)\left(\frac{s}{\sqrt{n_t}}\right)$$

where

- k = number of treatments (factor level combinations)
- s^2 = MSE, the estimate of the common variance σ^2
- df = degrees of freedom for s^2 =MSE
- n_t = the number of observations in each treatment
- $q_\alpha(k, df)$ comes from Tukey's table of critical values

Example

Suppose you want to make pairwise comparisons for an ANOVA

- $k = 5$ means
- $\alpha = 0.05$
- s^2 has 9 df

Tukey's Table of Critical Values

A Partial Reproduction of Table 11(a) in Appendix I; Upper 5% Points

df	2	3	4	5	6	7	8	9	10	11	12
1	17.97	26.98	32.82	37.08	40.41	43.12	45.40	47.36	49.07	50.59	51.96
2	6.08	8.33	9.80	10.88	11.74	12.44	13.03	13.54	13.99	14.39	14.75
3	4.50	5.91	6.82	7.50	8.04	8.48	8.85	9.18	9.46	9.72	9.95
4	3.93	5.04	5.76	6.29	6.71	7.05	7.35	7.60	7.83	8.03	8.21
5	3.64	4.60	5.22	5.67	6.03	6.33	6.58	6.80	6.99	7.17	7.32
6	3.46	4.34	4.90	5.30	5.63	5.90	6.12	6.32	6.49	6.65	6.79
7	3.34	4.16	4.68	5.06	5.36	5.61	5.82	6.00	6.16	6.30	6.43
8	3.26	4.04	4.53	4.89	5.17	5.40	5.60	5.77	5.92	6.05	6.18
9	3.20	3.95	4.41	4.76	5.02	5.24	5.43	5.59	5.74	5.87	5.98
10	3.15	3.88	4.33	4.65	4.91	5.12	5.30	5.46	5.60	5.72	5.83
11	3.11	3.82	4.26	4.57	4.82	5.03	5.20	5.35	5.49	5.61	5.71
12	3.08	3.77	4.20	4.51	4.75	4.95	5.12	5.27	5.39	5.51	5.61

Example

We have an experiment to determine the effect of nutrition on attention span of elementary school students.

- 15 students were randomly assigned to each of three meal plans:
    - no breakfast
    - light breakfast
    - full breakfast
- Attention spans were recorded during a morning reading.

Example

The ANOVA table for this experiment (from R) is:

```
> summary(aov(span~trt))
            Df Sum Sq Mean Sq F value Pr(>F)
trt          2  58.53  29.267   4.933 0.0273 *
Residuals   12  71.20   5.933
```

What can we conclude?

Example

Calculate Tukey's yardstick for this ANOVA.

Tukey's Table of Critical Values

A Partial Reproduction of Table 11(a) in Appendix I; Upper 5% Points

df	2	3	4	5	6	7	8	9	10	11	12
1	17.97	26.98	32.82	37.08	40.41	43.12	45.40	47.36	49.07	50.59	51.96
2	6.08	8.33	9.80	10.88	11.74	12.44	13.03	13.54	13.99	14.39	14.75
3	4.50	5.91	6.82	7.50	8.04	8.48	8.85	9.18	9.46	9.72	9.95
4	3.93	5.04	5.76	6.29	6.71	7.05	7.35	7.60	7.83	8.03	8.21
5	3.64	4.60	5.22	5.67	6.03	6.33	6.58	6.80	6.99	7.17	7.32
6	3.46	4.34	4.90	5.30	5.63	5.90	6.12	6.32	6.49	6.65	6.79
7	3.34	4.16	4.68	5.06	5.36	5.61	5.82	6.00	6.16	6.30	6.43
8	3.26	4.04	4.53	4.89	5.17	5.40	5.60	5.77	5.92	6.05	6.18
9	3.20	3.95	4.41	4.76	5.02	5.24	5.43	5.59	5.74	5.87	5.98
10	3.15	3.88	4.33	4.65	4.91	5.12	5.30	5.46	5.60	5.72	5.83
11	3.11	3.82	4.26	4.57	4.82	5.03	5.20	5.35	5.49	5.61	5.71
12	3.08	3.77	4.20	4.51	4.75	4.95	5.12	5.27	5.39	5.51	5.61

Example

Treatment	Mean	Standard Deviation
No Breakfast	9.4	2.30
Light Breakfast	14	2.55
Full Breakfast	13	2.50

Tukey's Method in R

```
> TukeyHSD(aov(span~trt))
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = span ~ trt)

$trt
            diff       lwr        upr     p adj
light-full  1.0 -3.110011  5.1100111 0.7963670
none-full  -3.6 -7.710011  0.5100111 0.0886624
none-light -4.6 -8.710011 -0.4899889 0.0284289
```

Example: Tukey's Method for More Complex ANOVAs

We will bring our example back to the supervisor and shift problem.

- We know there is a difference between the two supervisors.
- We will use Tukey's approach to compare each treatment (factor level combination).

Example

The ANOVA we found last class was

Source	*df*	SS	MS	F
Supervisor (A)	1	19208	19208	26.68
Shift (B)	2	247	123.5	0.17
Interaction (AB)	2	81127	40563.5	56.34
Error	12	8640	720	
Total	17	109222		

Example

Our treatment means looked like

	Shift		
Supervisor	Day	Swing	Night
1	602	498	450
2	487	602	657

There are $k = 6$ treatments.

Example

	s2night	s1day	s2swing	s1swing	s2day	s1night
s2night	-	55	55	159	170	207
s1day	-	-	0	104	115	152
s2swing	-	-	-	104	115	152
s1swing	-	-	-	-	11	48
s2day	-	-	-	-	-	37
s1night	-	-	-	-	-	-