ST 411/511 Lab 7
Multiple Comparisons

Objectives for this Lab

- Write Tukey-Kramer simultaneous confidence intervals for all pairwise comparisons.
- Use Dunnett's procedure to compare all groups to a control group.
- Write a Scheffé confidence interval for a data-suggested comparison.
- Use a Bonferroni correction to write pre-planned simultaneous confidence intervals.

For reference, here is the form of all the confidence intervals:

$$\text{pt est} \pm \text{multiplier} \cdot \text{SE(pt est)} \tag{1}$$

and here are the formulas for the point estimate of a linear combination of population means and for the standard error of the point estimate:

$$g = C_1\overline{Y}_1 + C_2\overline{Y}_2 + \ldots + C_I\overline{Y}_I \qquad \text{SE}(g) = s_p\sqrt{\frac{C_1^2}{n_1} + \frac{C_2^2}{n_2} + \ldots + \frac{C_I^2}{n_I}}. \tag{2}$$

1. As usual, start up RStudio and open Lab7.R. Load the Sleuth3 and ggplot2 R packages.

```
> library(Sleuth3)
> library(ggplot2)
```

2. Perform an analysis of variance on the disability discrimination data. Save the aov object and get the ANOVA table, as in items 2(d) and 2(f) of Lab 5. Get the means and sample sizes as in item 4 of Lab 5.

```
> case0601_aov<-aov(Score~Handicap, data=case0601)
> anova(case0601_aov)
> with(case0601, unlist(lapply(split(Score,Handicap), mean)))
> with(case0601, unlist(lapply(split(Score,Handicap), length)))
```

3. Use Tukey-Kramer to write confidence intervals for all pairwise differences between means.

   For a pairwise comparison, the formula in (1) looks like

$$\overline{Y}_i - \overline{Y}_j \pm \text{multiplier} \cdot \text{SE}(\overline{Y}_i - \overline{Y}_j).$$

   (a) Calculate the standard error of $\overline{Y}_i - \overline{Y}_j$, using the formula in (2). Remember to get s_p from the ANOVA table, because we want to use *all* the data to estimate σ.

```
> (SE <- sqrt(2.6665) * sqrt(1/14 + 1/14))
```

   You should get about 0.617. Note that this is the same for all i and j in the disability discrimination study because the sample sizes are all $n_i = n_j = 14$. If the sample sizes were different, we would have to calculate the SE separately for each pair of means.

(b) The Tukey-Kramer multiplier requires a quantile from a studentized range distribution. The textbook's notation for this quantile is $q_{I,\text{df}}(1-\alpha)$ R's `qtukey()` function calculates this quantity. The function requires three arguments. The first is $1-\alpha = 0.95$ for a 95% confidence interval (*not* 0.975). The second argument is the number of groups, and the third argument is the residual degrees of freedom from the ANOVA table.

```
> qtukey(0.95, 5, 65)
```

You should get 3.968034. (Page 162 of the textbook gives 3.975 from Table A.5, which was omitted from the 3rd edition. Just ignore this. We'll use R to calculate $q_{I,\text{df}}(1-\alpha)$.)
The multiplier is $q_{I,\text{df}}(1-\alpha)$ divided by $\sqrt{2}$.

```
> (M <- qtukey(0.95, 5, 65) / sqrt(2))
```

Multiplier M should be approximately 2.806.

(c) Calculate Tukey-Kramer confidence interval for the difference between the mean scores for Hearing and Amputee populations.

```
> 4.05 - 4.428571 - M*SE
> 4.05 - 4.428571 + M*SE
```

You should get about $(-2.11, 1.35)$.

(d) There are ten pairwise differences among five groups. Tukey-Kramer controls the familywise confidence level for all of them, so typically, you would present all ten confidence intervals. You can calculate them one-by-one as in item 3(c) or use `TukeyHSD()`.

```
> TukeyHSD(case0601_aov)
```

(e) (Optional) If you're using R Markdown, here's a way to use the `xtable` R package to get a nicely formatted table of the Tukey confidence intervals. You will probably have to install the package. See item 4(a) below for the procedure.)
Put the following in a code chunk with these options:
`results='asis', echo=FALSE, warning=FALSE`.

```
> library(xtable)
> print(xtable(TukeyHSD(case0601_aov)$Handicap,
+              caption="95\\% Tukey Confidence Intervals"),
+       comment=FALSE, caption.placement="top")
```

The first argument to `xtable()` is the part of the `TukeyHSD()` output that contains the table.
The argument `caption="95\\% Tukey Confidence Intervals"` sets the table caption. You need two backslashes in front of the %, or you'll get funny output.
The entire `xtable()` command is the first argument to `print()`. The `print()` function gives you more control over formatting.

4. Use the R package `multcomp` to use Dunnett's procedure to compare all groups to a control. Dunnett's procedure, like Tukey-Kramer, is for pairwise differences, but unlike Tukey-Kramer, it works for comparing each group to a control group. Tukey-Kramer is for comparing *all* pairs of means.

(a) Install and load the `multcomp` package. This is the installation procedure described in item 6(a) of Lab 1: From the Packages pane, click the Install button to open the Install Packages dialog box. Then type "multcomp" in the Packages line, and click "Install" at the bottom of the dialog box. Then load the package into R's library.

```
> library(multcomp)
```

(b) Dunnett's procedure requires you to select the control group. For the disability discrimination study, this is the `None` group. For the function we will use in `multcomp`, the control group needs to be the first group. Use `relevel()` to tell R to put `None` first.

```
> summary(case0601$Handicap) # Check the original ordering.
> case0601$Handicap <- relevel(case0601$Handicap, "None") # Put None first.
> summary(case0601$Handicap) # Check to make sure of the order.
```

(c) Since we reordered `Handicap`, we need to recreate the `aov` object. The ANOVA table doesn't depend on the ordering of the groups, so we don't need to recreate that.

```
> case0601_aov <- aov(Score~Handicap, data=case0601)
```

(d) Now use the `glht()` ("general linear hypothesis test") and `confint()` functions from the `multcomp` package to get the four Dunnett's confidence intervals.

```
> case0601_glht<- glht(case0601_aov, linfct=mcp(Handicap="Dunnett"))
> confint(case0601_glht)
```

At the end of the output, you'll see a table that contains the point estimate ("Estimate") and the lower and upper Dunnett's confidence bounds ("lwr" and "upr").

5. Write a Scheffé confidence interval for a data-suggested comparison. Scheffé is the ***only*** multiple comparison procedure that allows this because the Scheffé multiplier is appropriate for all possible linear contrasts. A *linear contrast* is a linear combination $\gamma = C_1\mu_1 + C_2\mu_2 + \ldots + C_I\mu_I$ where the C_i's sum to 0. All our comparisons have been contrasts.

(a) Scheffé's procedure is based on an F distribution. We can get the F quantiles from R's `qf()` function. This function takes three arguments. The first is $1 - \alpha = 0.95$ for a 95% confidence interval. The second and third are the numerator and denominator degrees of freedom from the ANOVA table. The numerator degrees of freedom are the extra or "model" degrees of freedom. The denominator degrees of freedom are the residual degrees of freedom.

```
> qf(0.95, 4, 65)
```

You should get 2.51304. In the notation of the *Sleuth*, this is $F_{4,65}(1-0.05)$. The multiplier for Scheffé's procedure is $\sqrt{(I-1)F_{(I-1),\mathrm{df}}(1-\alpha)}$, where I is the number of groups, df is the residual (denominator) degrees of freedom from the ANOVA table, and $1 - \alpha = 0.95$ for a 95% confidence interval (again, *not* 0.975).

```
> (M<-sqrt(4 * qf(0.95, 4, 65)))
```

You should get $M = 3.170514$.

(b) Calculate a Scheffé confidence interval for the difference between the average of the Crutches and Wheelchair groups and the average of the Amputee and Hearing groups. Note that we can't use SE from item 3(a), since this isn't a pairwise difference. Refer to the formula for SE(g) on page 154 of the text, and see Display 6.4 (but note that the interval in Display 6.4 is not a Scheffé confidence interval).

```
> SE <- sqrt(2.6665) * sqrt((0.5)^2/14 + (0.5)^2/14 + (0.5)^2/14 + (0.5)^2/14)
> (5.921429+5.342857 )/2 - (4.428571+4.05)/2 - M*SE
> (5.921429+5.342857 )/2 - (4.428571+4.05)/2 + M*SE
```

You should get about $(0.009, 2.777)$.

6. The Bonferroni correction is a very general multiple comparison procedure and can be used for any combination of several comparisons or tests. You need to know in advance that you will be making k comparisons. Suppose in the design phase of the handicap discrimination study, the researchers had decided to make the following three comparisons:

   - None vs. the average of the others
   - Hearing vs. the average of Amputee, Crutches, and Wheelchair
   - Crutches vs. the average of Amputee, Hearing, and Wheelchair

   (a) Calculate the Bonferroni-corrected multiplier for three simultaneous 95% confidence intervals from `qt()` :

   ```
   > alpha <- 0.05/3 # Set Bonferroni alpha to nominal alpha divided by k.
   > (M <- qt(1-alpha/2, 65))
   ```

   The multiplier should be 2.457515.

   (b) Calculate the first interval.

   ```
   > pt_est <- 4.9 - (4.428571+5.921429+4.05+5.342857)/4
   > SE <- sqrt(2.6665)*sqrt(1/14 + 4*(0.25)^2/14)
   > pt_est - M*SE
   > pt_est + M*SE
   ```

   Your interval should be approximately $(-1.23, 1.16)$.

   (c) Mimic the code in item 6(b) to calculate the other two intervals. You'll need to recalculate the standard errors. The intervals should be approximately $(-2.42, 0.06)$ and $(0.08, 2.55)$.

   (d) Because Bonferroni is so general, you could use it to make all ten pairwise comparisons. The Bonferroni-adjusted 95% multiplier would be the usual t-quantile but with $\alpha = 0.05$ divided by $k = 10$:

   ```
   > qt(1-(0.05/10)/2,65)
   ```

   You should get about 2.906. Compare this to the Tukey-Kramer multiplier in item 3(b) to see that Tukey-Kramer is preferable because it yields shorter intervals. This makes sense because Bonferonni is a general procedure whereas Tukey-Kramer is tailored for comparing all pairs of means.