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

$$\overline{Y}_i - \overline{Y}_j \pm \text{multiplier} \cdot s_p\sqrt{\frac{1}{n_i} + \frac{1}{n_j}}.$$

   (a) The multiplier for simultaneous 95% Tukey-Kramer confidence intervals is derived from the *studentized range distribution*, and can be calculated using the `qtukey()` function in R as

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

for the disability discrimination study. See page 10 of Outline 6 for details. The first argument to `qtukey()` is the desired familywise confidence level (95%), the second argument is the number of groups, and the third argument is the residual degrees of freedom from the ANOVA table in item 2. You should get 2.805824. The Tukey-Kramer intervals are all calculated as

$$\overline{Y}_i - \overline{Y}_j \pm 2.805824 \cdot s_p \sqrt{\frac{1}{n_i} + \frac{1}{n_j}}.$$

(b) Instead of calculating the pairwise confidence intervals by hand using the multiplier from 3(a), we'll use R's `tukeyHSD()` function which will calculate Tukey-Kramer confidence intervals for all pairwise comparisons.

```
> TukeyHSD(case0601_aov)
```

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.

(c) (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.

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

The ANOVA table doesn't depend on the ordering of the groups, so we don't need to recreate that by running the `anova()` function on the new `aov` object.

(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)
```

The first argument to `glht` is the `aov` object. The `linfct` ("linear function") argument specifies we want Dunnett's procedure multiple comparison procedure (`mcp`). The syntax within `mcp()` has the grouping variable on the left of the `=`.

At the end of the `confint()` output, you'll see a table that contains the point estimates of the comparisons to the control (`Estimate`) and the lower and upper Dunnett's confidence bounds (`lwr` and `upr`).

(e) (Optional) RMarkdown users can reuse the code from item 3(c) above to get a nicely-formated table of Dunnett's intervals. The table information is in the variable `confint(case0601_glht)$confint`, and we need to edit the table caption.

```
> print(xtable(confint(case0601_glht)$confint,
+              caption="95\\% Dunnett's Confidence Intervals"),
+        comment=FALSE, caption.placement="top")
```

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% Scheffé 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 (*not* 0.975).

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

You should get $M = 3.170514$.

(b) Calculate a 95% Scheffé confidence interval for the difference between the average of the Crutches and Wheelchair groups and the average of the Amputee and Hearing groups. See Display 6.4 on page 155 of the textbook or on page 6 of Outline 6 for an application of formulas 2 calculating the point estimate and standard error.

```
> 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 $(0.009, 2.777)$.

Display 6.4 on page 155 of the textbook or on page 6 of Outline 6 calculates a 95% t-confidence interval for the difference between the average of the Crutches and Wheelchair groups and the average of the Amputee and Hearing groups as $(0.521, 2.264)$, which is narrower than our Scheffé confidence interval. The Scheffé confidence interval must be wider because it accommodates estimation of all possible comparisons. That's a lot of confidence intervals, and each one has a chance to miss the truth.

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

Here, $k = 3$ because we want three Bonferroni confidence intervals.

(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) (Optional) RMarkdown users will need to construct the table to give to `xtable()`. One way to do it is to build a data frame containing the information.

```
> Bonf_table <- data.frame(
+   row.names=c("None - avg. of the others",
+               "Hearing - avg. of Amputee, Crutches, and Wheelchair",
+               "Crutches - avg. of Amputee, Hearing, and Wheelchair"),
+   lwr=c(-1.23, -2.42, 0.08),
+   upr=c(1.16, 0.06, 2.55))
> print(xtable(Bonf_table,
+              caption="95\\% Bonferroni Confidence Intervals"),
+              comment=FALSE, caption.placement="top")
```

(e) 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. The Tukey-Kramer multiplier from 3(a) is only 2.805824. 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.