Name: James Tsai
Section: MSDS6371-401
Date: 10/11/15

1. **Problem 13, page 171. Handicap Study. Use the Bonferroni method to construct simultaneous confidence intervals for $\mu_2 - \mu_3$, $\mu_2 - \mu_5$, and $\mu_3 - \mu_5$ (to see whether there are differences in attitude toward the mobility type of handicaps).**

Summary Statistics

Pooled SD	$S_p = 1.6329$	65 d.f.
Amputee	$\bar{X}_2 = 4.4286$	
Crutches	$\bar{X}_3 = 5.9214$	
Wheelchair	$\bar{X}_5 = 5.3429$	

Assumption: We are using the pooled SD and df from the textbook. We are assuming 3 pairs of hypothesis tests, therefore setting K=3.

PART A) Amputee, Crutches

Estimate the linear combination

$g = \bar{X}_2 - \bar{X}_3 = 4.4286 - 5.9214 = -1.4928$

Find the standard error of the estimate (using pooled SD)

$SE(g) = 1.6329\sqrt{\frac{1}{14} + \frac{1}{14}} = 0.617$

Construct the 95% confidence interval

Df = 65, K = 3, $\alpha = .05$

Bonferroni adjustment for alpha = $\alpha/3 = 0.0167$

$t_{0.0167/2,65} = 2.457$

$E = (2.457)(0.617) = 1.516$

$-1.4928 \pm 1.516 \rightarrow (-3.0088, 0.0232)$

Conclusion

For the difference between the attitude toward amputees and those with crutches, we fail to reject H_0 since zero is contained in the confidence interval.

PART B) Amputee, Wheelchair

Estimate the linear combination

$g = \bar{X}_2 - \bar{X}_5 = 4.4286 - 5.3429 = -0.9143$

Find the standard error of the estimate (using pooled SD)

$SE(g) = 1.6329\sqrt{\frac{1}{14}+\frac{1}{14}} = 0.617$

Construct the 95% confidence interval

Df = 65, K = 3, $\alpha = .05$

Bonferroni adjustment for alpha = $\alpha/3 = 0.0167$

$t_{0.0167/2,65} = 2.457$

E = (2.457)(0.617) = 1.516

$-0.9143 \pm 1.516 \rightarrow (-2.4303,\ 0.6017)$

Conclusion

For the difference between the attitude toward amputees and those in wheelchairs, we fail to reject H_0 since zero is contained in the confidence interval.

PART C) Crutches, Wheelchair

Estimate the linear combination

$g = \bar{X}_3 - \bar{X}_5 = 5.9214 - 5.3429 = 0.5785$

Find the standard error of the estimate (using pooled SD)

$SE(g) = 1.6329\sqrt{\frac{1}{14}+\frac{1}{14}} = 0.617$

Construct the 95% confidence interval

Df = 65, K = 3, $\alpha = .05$

Bonferroni adjustment for alpha = $\alpha/3 = 0.0167$

$t_{0.0167/2,65} = 2.457$

E = (2.457)(0.617) = 1.516

$0.5785 \pm 1.516 \rightarrow (-0.9375,\ 2.0945)$

Conclusion

For the difference between the attitude toward those in crutches and those in wheelchairs, we fail to reject H_0 since zero is contained in the confidence interval.

2. **Problem 14, page 171. Handicap Study. Examine these data with your available statistical computer package. See what multiple comparison procedures are available within the one-way analysis of variance procedure. Verify the 95% confidence interval half-widths in Display 6.6.**

DISPLAY 6.6 Summary of 95% confidence interval procedures for differences between treatment means in the handicap study

		Difference with . . .			
Group	Average	Hearing	Amputee	Control	Wheelchair
Crutches	5.921	1.871	1.492	1.021	0.578
Wheelchair	5.343	1.293	0.914	0.443	
Control	4.900	0.850	0.471		
Amputee	4.429	0.379			
Hearing	4.050				

A confidence interval is centered at a difference with half-width given by one of the procedures.

Procedure	95% interval half-width
LSD	1.233
Dunnett	1.545 (for comparisons with control only)
Tukey–Kramer	1.735
Bonferroni	1.794
Scheffé	1.957

Using SAS, we ANOVA procedure “PROC GLM” with the parameters LSD, DUNNET, TUKEY, BON, and SCHEFFE to verify the “95% confidence interval half-width” in display 6.6. The value we can use from the SAS output to verify this is “Least Significance Difference”. Below is a table summarizing the findings and verifying the SAS code and output with the textbook.

	95% CI half-width (from Textbook)	**Least Significance Difference (from SAS)**
LSD	1.233	1.2326
DUNNET	1.545	1.5449
TUKEY	1.735	1.7317
BONFERRONI	1.794	1.7936
SCHEFFE	1.957	1.9568

SAS code and outputs (mean diffs and CI’s not captured)

```
DATA HANDICAP;
INFILE 'C:\Users\james\Documents\My SAS Files\9.4\Ch 6
Handicap Data.csv' DLM=',' FIRSTOBS=2;
INPUT SCORE HANDICAP $;
RUN;

PROC GLM DATA=HANDICAP;
CLASS HANDICAP;
MODEL SCORE=HANDICAP;
MEANS HANDICAP / HOVTEST = BF LSD CLDIFF;
RUN;
```

The SAS System

The GLM Procedure

t Tests (LSD) for SCORE

Note: This test controls the Type I comparisonwise error rate, not the experimentwise error rate.

Alpha	0.05
Error Degrees of Freedom	65
Error Mean Square	2.666484
Critical Value of t	1.99714
Least Significant Difference	1.2326

```
PROC GLM DATA=HANDICAP;
CLASS HANDICAP;
MODEL SCORE=HANDICAP;
MEANS HANDICAP / HOVTEST = BF DUNNETT CLDIFF;
RUN;
```

The SAS System

The GLM Procedure

Dunnett's t Tests for SCORE

Note: This test controls the Type I experimentwise error for comparisons of all treatments against a control.

Alpha	0.05
Error Degrees of Freedom	65
Error Mean Square	2.666484
Critical Value of Dunnett's t	2.50316
Minimum Significant Difference	1.5449

```
PROC GLM DATA=HANDICAP;
CLASS HANDICAP;
MODEL SCORE=HANDICAP;
MEANS HANDICAP / HOVTEST = BF TUKEY CLDIFF;
RUN;
```

The SAS System

The GLM Procedure

Tukey's Studentized Range (HSD) Test for SCORE

Note: This test controls the Type I experimentwise error rate.

Alpha	0.05
Error Degrees of Freedom	65
Error Mean Square	2.666484
Critical Value of Studentized Range	3.96804
Minimum Significant Difference	1.7317

```
PROC GLM DATA=HANDICAP;
CLASS HANDICAP;
MODEL SCORE=HANDICAP;
MEANS HANDICAP / HOVTEST = BF BON CLDIFF;
RUN;
```

The SAS System

The GLM Procedure

Bonferroni (Dunn) t Tests for SCORE

Note: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than Tukey's for all pairwise comparisons.

Alpha	0.05
Error Degrees of Freedom	65
Error Mean Square	2.666484
Critical Value of t	2.90602
Minimum Significant Difference	1.7936

```
PROC GLM DATA=HANDICAP;
CLASS HANDICAP;
MODEL SCORE=HANDICAP;
MEANS HANDICAP / HOVTEST = BF SCHEFFE CLDIFF;
RUN;
```

The SAS System

The GLM Procedure

Scheffe's Test for SCORE

Note: This test controls the Type I experimentwise error rate, but it generally has a higher Type II error rate than Tukey's for all pairwise comparisons.

Alpha	0.05
Error Degrees of Freedom	65
Error Mean Square	2.666484
Critical Value of F	2.51304
Minimum Significant Difference	1.9568

3. **Problem 21, page 172. Reconsider the data problem of Exercise 5.25 concerning the distributions of annual incomes in 2005 for Americans in each of the five education categories. (a) Use the Tukey-Kramer procedure to compare every group to every other group. Which pairs of means differ and by how many dollars (or by what percent)? (Use *p*-values and confidence intervals in your answer.) (b) User the Dunnett procedure to compare every other group to the group with 12 years of education. Which group means apparently differ from the mean for those 12 years of education and by how many dollars (or by what percent)? (Use *p*-values and confidence intervals in your answer.)**

PART A) Education/Income2005, Tukey-Kramer procedure

SAS code and outputs

```
DATA EDUCATION;
INFILE 'C:\Users\james\Documents\My SAS
Files\9.4\ex0525.csv' DLM=',' FIRSTOBS=2;
INPUT SUBJECT EDUC $ INCOME2005;
RUN;

PROC SORT DATA=EDUCATION OUT=EDUCATION_SORTED; BY EDUC;

PROC GLM DATA=EDUCATION_SORTED;
CLASS EDUC;
MODEL INCOME2005=EDUC;
LSMEANS EDUC / PDIFF CL ADJUST=TUKEY;
RUN;
```

The GLM Procedure
Least Squares Means
Adjustment for Multiple Comparisons: Tukey-Kramer

EDUC	INCOME2005 LSMEAN	LSMEAN Number
12	36864.8961	1
13-15	44875.9568	2
16	69996.9729	3
<12	28301.4485	4
>16	76855.4626	5

Least Squares Means for effect EDUC
Pr > |t| for H0: LSMean(i)=LSMean(j)
Dependent Variable: INCOME2005

i/j	1	2	3	4	5
1		0.0026	<.0001	0.2031	<.0001
2	0.0026		<.0001	0.0006	<.0001
3	<.0001	<.0001		<.0001	0.1861
4	0.2031	0.0006	<.0001		<.0001
5	<.0001	<.0001	0.1861	<.0001	

EDUC	INCOME2005 LSMEAN	95% Confidence Limits	
12	36865	34175	39555
13-15	44876	41501	48251
16	69997	65733	74261
<12	28301	20934	35669
>16	76855	72413	81298

Least Squares Means for Effect EDUC

i	j	Difference Between Means	Simultaneous 95% Confidence Limits for LSMean(i)-LSMean(j)	
1	2	-8011.060712	-14020	-2002.370298
1	3	-33132	-40151	-26113
1	4	8563.447549	-2355.449497	19482
1	5	-39991	-47221	-32760
2	3	-25121	-32692	-17550
2	4	16575	5292.914210	27856
2	5	-31980	-39747	-24212
3	4	41696	29845	53546
3	5	-6858.489660	-15431	1714.215306
4	5	-48554	-60531	-36577

For those with greater than 16 years of education, there is a difference of means of $31,980 as compared to those with only 13 to 15 years of education. With a *p*-value of less than 0.0001, we have a 95% confidence interval at α = .05, that the group with greater than 16 years of education have average earnings between $24,212 and $39,747 greater than those with only 13 to 15 years of education.

For those with greater than 16 years of education, there is a difference of means of $39,991 as compared to those with only 12 years of education. With a *p*-value of less than 0.0001, we have a 95% confidence interval at α = .05, that the group with greater than 16 years of education have average earnings between $32,760 and $47,221 greater than those with only 12 years of education.

For those with greater than 16 years of education, there is a difference of means of $48,554 as compared to those with less than 12 years of education. With a *p*-value of less than 0.0001, we have a 95% confidence interval at α = .05, that the group with greater than 16 years of education have average earnings between $36,577 and $60,531 greater than those with less than 12 years of education.

For those with 16 years of education, there is a difference of means of $25,121 as compared to those with 13 to 15 years of education. With a *p*-value of less than 0.0001, we have a 95% confidence interval at α = .05, that the group with greater than 16 years of education have average earnings between $17,550 and $32,692 greater than those with 13 to 15 years of education.

For those with 16 years of education, there is a difference of means of $33,132 as compared to those with 12 years of education. With a *p*-value of less than 0.0001, we have a 95% confidence interval at α = .05, that the group with greater than 16 years of education have average earnings between $26,113 and $40,151 greater than those with 12 years of education.

For those with 16 years of education, there is a difference of means of $41,696 as compared to those with less than 12 years of education. With a *p*-value of less than 0.0001, we have a 95% confidence interval at α = .05, that the group with greater than 16 years of education have average earnings between $29845 and $53546 greater than those with less than 12 years of education.

For those with 13 to 15 years of education, there is a difference of means of $8,011 as compared to those with 12 years of education. With a *p*-value of 0.0026, we have a 95% confidence interval at α = .05, that the group with 13 to 15 years of education have average earnings between $2,002 and $14,020 greater than those with 12 years of education.

For those with 13 to 15 years of education, there is a difference of means of $16,575 as compared to those with less than12 years of education. With a *p*-value of 0.0006, we have a 95% confidence interval at α = .05, that the group

with 13 to 15 years of education have average earnings between $5293 and $27,856 greater than those with less than 12 years of education.

PART B) Education/Income2005, Dunnett procedure

```
PROC GLM DATA=EDUCATION_SORTED;
CLASS EDUC;
MODEL INCOME2005=EDUC;
LSMEANS EDUC / PDIFF CL ADJUST=DUNNETT;
RUN;
```

The SAS System

The GLM Procedure
Least Squares Means
Adjustment for Multiple Comparisons: Dunnett

EDUC	INCOME2005 LSMEAN	H0:LSMean=Control Pr > |t|
12	36864.8961	
13-15	44875.9568	0.0011
16	69996.9729	<.0001
<12	28301.4485	0.1180
>16	76855.4626	<.0001

EDUC	INCOME2005 LSMEAN	95% Confidence Limits	
12	36865	34175	39555
13-15	44876	41501	48251
16	69997	65733	74261
<12	28301	20934	35669
>16	76855	72413	81298

Least Squares Means for Effect EDUC

i	j	Difference Between Means	Simultaneous 95% Confidence Limits for LSMean(i)-LSMean(j)	
2	1	8011.060712	2550.534144	13472
3	1	33132	26754	39511
4	1	-8563.447549	-18486	1359.334839
5	1	39991	33420	46561

The three groups means that apparently differ from the mean for those with 12 years of education are those with greater than 16 years of education, those with 16 years of education, and those with 13 to 15 years of education. The *p*-values corresponding to these comparisons with the control group (12 years of education) are <.0001, <.0001, and 0.0011, respectively.

For those with greater than 16 years of education, we have a 95% confidence interval at α = .05, that their average earnings are between $33,420 and $46,561 greater than those with only 12 years of education.

For those with 16 years of education, we have a 95% confidence interval at α = .05, that their average earnings are between $26,754 and $39,511 greater than those with only 12 years of education.

For those with 13 to 15 years of education, we have a 95% confidence interval at α = .05, that their average earnings are between $2,551 and $13,472 greater than those with only 12 years of education.