Homework 06 Answer Sheet

Psych 10C
Due: Sunday, October 30th (by 11:59pm PT)

Zhenze Zhang

Submission Details

- Download *HW06AnswerSheet.Rmd* from the Canvas course space and open it RStudio.
- Enter your name in the *author* field at the top of the document.
- Complete the assignment by entering your answers in your *HW06AnswerSheet.Rmd* document.
- Once you have completed the assignment, click the *Knit* button to turn your completed answer document into a pdf file.
- Submit your HW06AnswerSheet.pdf file only (no other formats are acceptable) before the assignment's deadline.

Problems

For each problem, show/describe all of your work.

Problem #1 (8 points)

A research study was conducted to examine the clinical efficacy of a new antidepressant. Depressed patients were randomly assigned to one of three groups: a placebo group, a group that received a low dose of the drug, and a group that received a moderate dose of the drug. After four weeks of treatment, the patients completed the Beck Depression Inventory. The higher the score, the more depressed the patient.

Load the data from our course website (HW06_DrugData.csv) into a data frame named *DrugData*. The Beck Depression Inventory scores will be contained in the *Score* variable of the data frame (DrugData$Scores) and the conditions (i.e. dosage of drug) will be contained in the *Dose* variable (DrugData$Dose). The three "Dose" levels are "Placebo", "Low", and "Moderate".

(a) (2 points) Perform an ANOVA using RStudio's *aov()* function. Use an $\alpha = .05$. Use the *summary()* function to show the results of the ANOVA. Report whether or not the null hypothesis should be rejected.

ANSWER:

```
DrugData<- read.csv("/Users/zzze/Downloads/HW06_DrugData.csv")
my.anova1<- aov(Score~ Dose, DrugData)
summary(my.anova1)
```

```
##             Df Sum Sq Mean Sq F value  Pr(>F)
## Dose         2 1484.9   742.5   11.27 0.00176 **
## Residuals   12  790.8    65.9
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

- Since our p-value is less than 0.05, we can reject the null hypothesis.

(b) (2 points) If F is significant, use RStudio to perform a Tukey test of all pairs of means in the study. Compute your values step-by-step (i.e. do not use the *TukeyHSD()* function) with $\alpha = .05$. Report HSD, all pairwise mean differences, and interpret the results.

ANSWER:

```
qf(0.05, 2, 12, lower.tail = FALSE)
```

```
## [1] 3.885294
```

- The F value 11.27 is higher than critical F value 3.885294, therefore it is significant.

```
n1<- 5 #number of samples in each condition
MSWith1<- 65.9 #residuals of mean square
k1<- 3 #number of conditions
q1<- qtukey(0.95, 3, 12)
HSD1<- q1*sqrt(MSWith1/ n1)
HSD1
```

```
## [1] 13.69734
```

```
Drug1Mean<- mean(DrugData[DrugData$Dose == 'Placebo', 'Score'])
Drug2Mean<- mean(DrugData[DrugData$Dose == 'Low', 'Score'])
Drug3Mean<- mean(DrugData[DrugData $Dose == 'Moderate', 'Score'])

D1.D2<- Drug1Mean- Drug2Mean
D1.D3<- Drug1Mean- Drug3Mean
D2.D3<- Drug2Mean- Drug3Mean
cat("Placebo-Low:", D1.D2, "\nPlacebo-Moderate:", D1.D3,
    "\nLow-Moderate:", D2.D3)
```

```
## Placebo-Low: 17.6
## Placebo-Moderate: 23.4
## Low-Moderate: 5.8
```

- The HSD we got is 13. 69734. Comparisons among the means for each dose using Tukey's HSD (13.69734) showed that significantly different numbers of depressed were yielded by both Placebo and Low dose, and Placebo and Moderate dose.

(c) (2 points) Calculate confidence intervals for each of the comparisons.

ANSWER:

```
cat("Placebo-Low:[", D1.D2- HSD1, ",", D1.D2 + HSD1, "]")
```

```
## Placebo-Low:[ 3.902657 , 31.29734 ]
```

```
cat("\nPlacebo-Moderate:[", D1.D3- HSD1, ",", D1.D3 +HSD1, "]")
```

```
##
## Placebo-Moderate:[ 9.702657 , 37.09734 ]
```

```
cat("\nLow-Moderate:[", D2.D3- HSD1, ",", D2.D3+ HSD1, "]")
```

```
##
## Low-Moderate:[ -7.897343 , 19.49734 ]
```

(d) (2 points) Perform a Scheffe test, comparing the Placebo group to the Low and Moderate groups combined. Interpret the results.

ANSWER:

```
coeff1<- c(2, -1, -1)
TPlacebo<- sum(DrugData[DrugData $Dose == "Placebo", "Score"])
TLow <- sum(DrugData[DrugData $Dose == "Low", "Score"])
TModerate<- sum(DrugData[DrugData $Dose == "Moderate", "Score"])
```

```
SScomp1<- sum(coeff1 * c(TPlacebo, TLow, TModerate))^2/ (n1 * sum(coeff1 ^2))
SScomp1
```

```
## [1] 1400.833
```

```
MScomp1<- SScomp1
F1<- MScomp1/MSWith1
F1
```

```
## [1] 21.25695
```

```
dfBet1<- 2
dfWith1<- 12
FCritical1<- qf( 0.05, dfBet1, dfWith1, lower.tail = FALSE)
FCritical1
```

```
## [1] 3.885294
```

```
Fprime1<- (k1-1) * FCritical1
Fprime1
```

```
## [1] 7.770588
```

- Since the F statistic (21.25695) exceeds our F' value(7.770588), we reject the null hypothesis, finding a significant difference between Placebo group compared with the Low group and Moderate group combined.

Problem #2 (9 points)

Four fabrics are tested for flammability at the National Bureau of Standards. Burn times (in seconds) are recorded after a paper tab is ignited on the hem of a dress made of each fabric.

Load the data from our course website (HW06_FabricData.csv) into a data frame named *FabricData*. The burn times will be contained in the *BurnTime* variable of the data frame (FabricData$BurnTime) and the conditions (i.e. fabric number) will be contained in the *Fabric* variable (FabricData$Fabric).

```
FabricData<- read.csv("/Users/zzze/Downloads/HW06_FabricData.csv")
```

(a) (2 points) Perform an ANOVA using RStudio's *aov()* function. Use an $\alpha =$ .05. Use the *summary()* function to show the results of the ANOVA. Report whether or not the null hypothesis should be rejected.

ANSWER

```
my.anova2<- aov(BurnTime~ Fabric ,FabricData)
summary(my.anova2)
```

```
##             Df Sum Sq Mean Sq F value   Pr(>F)
## Fabric       3 120.50   40.17   13.89 0.000102 ***
## Residuals   16  46.26    2.89
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

- Since the p-value is less than 0.05, we can reject the null hypothesis.

(b) (2 points) If F is significant, use RStudio to perform a Tukey test of all pairs of means in the study. Compute your values step-by-step (i.e. do not use the *TukeyHSD()* function) with $\alpha = .05$. Report HSD, all pairwise mean differences, and interpret the results.

ANSWER:

```r
qf(0.05, 3, 16, lower.tail = FALSE)
```

```
## [1] 3.238872
```

- Since the F value 13.89 is larger than critical F 3.238872, it is significant.

```r
n2<- 5 #number of samples in each condition
MSWith2<- 2.89 #residual
q2<- qtukey(.95, 4, 16)
HSD2<- q2* sqrt(MSWith2/n2)
HSD2
```

```
## [1] 3.076095
```

```r
Fab1Mean<- mean(FabricData[FabricData$Fabric == "Fab1", "BurnTime"])
Fab2Mean<- mean(FabricData[FabricData$Fabric == "Fab2", "BurnTime"])
Fab3Mean<- mean(FabricData[FabricData$Fabric == "Fab3", "BurnTime"])
Fab4Mean<- mean(FabricData[FabricData$Fabric == "Fab4", "BurnTime"])

Fab1.Fab2<- Fab1Mean - Fab2Mean
Fab1.Fab3<- Fab1Mean - Fab3Mean
Fab1.Fab4<- Fab1Mean - Fab4Mean
Fab2.Fab3<- Fab2Mean - Fab3Mean
Fab2.Fab4<- Fab2Mean - Fab4Mean
Fab3.Fab4<- Fab3Mean - Fab4Mean
cat("fab1-fab2:",Fab1.Fab2,"\nfab1-fab3:",Fab1.Fab3,"\nfab1-fab4:",Fab1.Fab4,
    "\nfab2-fab3:",Fab2.Fab3,"\nfab2-fab4:",Fab2.Fab4,"\nfab3-fab4:", Fab3.Fab4)
```

```
## fab1-fab2: 5.02
## fab1-fab3: 6.54
## fab1-fab4: 4.8
## fab2-fab3: 1.52
## fab2-fab4: -0.22
## fab3-fab4: -1.74
```

- The HSD we got is 3.076095. Comparisons among the means for each dose using Tukey's HSD (3.076095) showed that significantly different flammability were yield by fab1 and fab2, fab1 and fab3, fab1 and fab4.

(c) (2 points) Calculate confidence intervals for each of the comparisons.

ANSWER:

```r
cat("fab1-fab2:[", Fab1.Fab2- HSD2, ",", Fab1.Fab2 + HSD2, "]")
```

```
## fab1-fab2:[ 1.943905 , 8.096095 ]
```

```r
cat("\nfab1-fab3:[", Fab1.Fab3- HSD2, ",", Fab1.Fab3 +HSD2, "]")
```

```
##
## fab1-fab3:[ 3.463905 , 9.616095 ]
```

```r
cat("\nfab1-fab4:[", Fab1.Fab4- HSD2, ",", Fab1.Fab4+ HSD2, "]")
```

```
##
## fab1-fab4:[ 1.723905 , 7.876095 ]
```

```r
cat("\nfab2-fab3:[", Fab2.Fab3- HSD2, ",", Fab2.Fab3+ HSD2, "]")
```

```
##
## fab2-fab3:[ -1.556095 , 4.596095 ]
```

```r
cat("\nfab2-fab4:[", Fab2.Fab4- HSD2, ",", Fab2.Fab4+ HSD2, "]")
```

```
##
## fab2-fab4:[ -3.296095 , 2.856095 ]
```

```r
cat("\nfab3-fab4:[", Fab3.Fab4- HSD2, ",", Fab3.Fab4+ HSD2, "]")
```

```
##
## fab3-fab4:[ -4.816095 , 1.336095 ]
```

(d) (1 point) Perform a Tukey test using the *TukeyHSD()* function with $\alpha = .05$. Confirm that the results are the same as you found with your Tukey Test in parts (b-c).

ANSWER:

```r
TukeyHSD(my.anova2, conf.level = 0.95)
```

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = BurnTime ~ Fabric, data = FabricData)
##
## $Fabric
##            diff      lwr      upr     p adj
## Fab2-Fab1 -5.02 -8.09676 -1.94324 0.0013227
## Fab3-Fab1 -6.54 -9.61676 -3.46324 0.0000851
## Fab4-Fab1 -4.80 -7.87676 -1.72324 0.0019981
## Fab3-Fab2 -1.52 -4.59676  1.55676 0.5094118
## Fab4-Fab2  0.22 -2.85676  3.29676 0.9968426
## Fab4-Fab3  1.74 -1.33676  4.81676 0.3968476
```

- The results are the same as I got with Tukey test in part b and c.

(e) (2 points) Perform a Scheffe test, comparing Fabric 4 to Fabrics 2 and 3 combined.

ANSWER:

```
coeff2<- c(0, -1, -1, 2)
TFab1<- sum(FabricData[FabricData$Fabric == "Fab1", "BurnTime"])
TFab2<- sum(FabricData[FabricData$Fabric == "Fab2", "BurnTime"])
TFab3<- sum(FabricData[FabricData$Fabric == "Fab3", "BurnTime"])
TFab4<- sum(FabricData[FabricData$Fabric == "Fab4", "BurnTime"])

SScomp2<- sum(coeff2* c(TFab1, TFab2, TFab3, TFab4))^2/ (n2* sum(coeff2^2))
MSComp2<- SScomp2

F2<- MSComp2/MSWith2
F2
```

```
## [1] 1.107728
```

```
k2<- 4
dfBet2<- 3
dfWith2<- 16
FCritical2<- qf(0.05, dfBet2, dfWith2, lower.tail = FALSE)
FCritical2
```

```
## [1] 3.238872
```

```
Fprime2<- (k2-1)* FCritical2
Fprime2
```

```
## [1] 9.716615
```

- Since our F value (1.107728) does not exceed our F' value (9.716615), we cannot reject the null hypothesis, showing there is no significant difference between the Fabric 4 and Fabric 2 & 3.

(f) (2 points) Perform a Scheffe test, comparing Fabric 1 to Fabrics 2, 3, and 4 combined.

ANSWER:

```
coeff3<- c(3, -1, -1, -1)
SScomp3<- sum(coeff3* c(TFab1, TFab2, TFab3, TFab4))^2/ (n2* sum(coeff3^2))
MSComp3<- SScomp3

F3<- MSComp3/MSWith2
F3
```

```
## [1] 38.58847
```

- Since our F value (38.5887) exceeds our F' value (9.716615), we reject the null hypothesis, showing there is a significant difference between the Fabric 1 and Fabric 2 & 3 & 4 combined.

Problem #3 (3 points)

In an effort to improve the quality of recording tapes, the effects of four kinds of coatings A, B, C, D on the reproducing quality of sound are compared. Suppose that the measurements of sound distortion given in the table below are from tapes treated with the four coatings.

Load the data from our course website (HW06_TapeData.csv) into a data frame named *TapeData*. The sound distortion scores will be contained in the *Score* variable of the data frame (TapeData$Scores) and the conditions (i.e. type of coating) will be contained in the *Coating* variable (TapeData$Coating).

```
TapeData<- read.csv("/Users/zzze/Downloads/HW06_TapeData.csv")
```

(a) (2 points) Perform an ANOVA using RStudio's *aov()* function. Use an $\alpha = .05$. Use the *summary()* function to show the results of the ANOVA. Report whether or not the null hypothesis should be rejected.

ANSWER:

```
my.anova3<- aov(Score~Coating, TapeData)
summary(my.anova3)
```

```
##             Df Sum Sq Mean Sq F value Pr(>F)
## Coating      3     68  22.667    4.34 0.0181 *
## Residuals   18     94   5.222
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

- Since our p-value is less than 0.05, we should reject our null hypothesis.

(b) (1 point) Perform a Tukey test using the *TukeyHSD()* function with $\alpha = .05$. Report all significantly different tape coatings.

ANSWER:

```
TukeyHSD(my.anova3, conf.level = 0.95)
```

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = Score ~ Coating, data = TapeData)
##
## $Coating
##     diff        lwr      upr     p adj
## B-A    5  0.6673794 9.332621 0.0205494
## C-A    4  0.2181787 7.781821 0.0360310
```

```
## D-A     3 -0.9109306 6.910931 0.1700724
## C-B    -1 -5.0481978 3.048198 0.8964449
## D-B    -2 -6.1690661 2.169066 0.5413330
## D-C    -1 -4.5932831 2.593283 0.8595773
```

- From the table above, we can see that the any p-value that is less than 0.05 is significant. Therefore, significantly different reproducing quality of sounds were yield by Group B & A and Group C & A
