Homework 05 Answer Sheet

Psych 10C

Due: Sunday, October 23rd (by 11:59pm PT)

Zhenze Zhang

Submission Details

- Download *HW05AnswerSheet.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 *HW05AnswerSheet.Rmd* document.
- Once you have completed the assignment, click the *Knit* button to turn your completed answer document into a pdf file.
- Submit your HW05AnswerSheet.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 (5 points)

Perform a one-factor analysis of variance on the following data. There are 3 conditions, each containing 4 scores. Perform your analysis step-by-step using RStudio, but without using RStudio's *aov()* function.

```
## Condition 1 Data:  5 3 2 2
## Condition 2 Data:  5 0 1 2
## Condition 3 Data:  2 1 0 1
```

(a) (1 point) Calculate the between-conditions, within-conditions, and total degrees of freedom.

ANSWER

```
n1<- length(Cond1)
n2<- length(Cond2)
n3<- length(Cond3)
N<- n1+n2+n3
dfTotal<- N-1
```

```r
dfBetween<- 3-1
dfWithin<- dfTotal-dfBetween
cat("dfTotal:", dfTotal, "\ndfBetween:", dfBetween, "\ndfWithin:", dfWithin)
```

```
## dfTotal: 11
## dfBetween: 2
## dfWithin: 9
```

(b) (2 points) Calculate the between-conditions, within-conditions, and total sums of squares.

ANSWER:

```r
SumX<- sum(Cond1)+sum(Cond2)+sum(Cond3)
SumXsq<- sum(Cond1^2)+sum(Cond2^2)+sum(Cond3^2)
SumTsq<- sum(Cond1)^2+sum(Cond2)^2+sum(Cond3)^2
```

```r
SSTotal<- SumXsq-SumX^2/N
SSBetween<- SumTsq/n1-SumX^2/N
SSWithin<- SSTotal-SSBetween
cat("Total sums of squares:", SSTotal, "\nBetween-conditions:", SSBetween, "\nWithin-conditions:", SSWi
```

```
## Total sums of squares: 30
## Between-conditions: 8
## Within-conditions: 22
```

(c) (1 point) Calculate the between-conditions and within-conditions mean squares, the F-statistic, and the critical value of F. Use significance level $\alpha = .05$.

ANSWER:

```r
MSBetween<- SSBetween/dfBetween
MSWithin<- SSWithin/dfWithin
F<- MSBetween/MSWithin
CriticalF<- qf(0.05, dfBetween, dfWithin, lower.tail = FALSE)
cat("Between-conditions mean squares:", MSBetween,
    "\nWithin-conditions mean squares:", MSWithin, "\nF-statistic:", F,
    "\nCritical value of F:", CriticalF)
```

```
## Between-conditions mean squares: 4
## Within-conditions mean squares: 2.444444
## F-statistic: 1.636364
## Critical value of F: 4.256495
```

(d) (1 point) Based on your calculated F and critical F value, should we reject the null hypothesis?

ANSWER:
Since the F-value of 1.636364 is smaller than Critical value of F (4.256495), we fail to reject the null hypothesis.

Problem #2 (4 points)

Perform a one-factor analysis of variance on the following data. There are 4 conditions, each containing between 2 and 4 scores. Perform your analysis step-by-step using RStudio, but without using RStudio's *aov()* function.

```
## Condition 1 Data: 2 1 3
## Condition 2 Data: 1 5
## Condition 3 Data: 9 5 6 4
## Condition 4 Data: 3 4 5
```

(a) (1 point) Calculate the between-conditions, within-conditions, and total degrees of freedom.

ANSWER

```r
n1<- length(Cond1)
n2<- length(Cond2)
n3<- length(Cond3)
n4<- length(Cond4)
N2<- n1+n2+n3+n4
dfTot2<- N2-1
dfBet2<- 4-1
dfWith2<- dfTot2-dfBet2
cat("dfTotal:", dfTot2, "\ndfBetween:", dfBet2, "\ndfWithin:", dfWith2)
```

```
## dfTotal: 11
## dfBetween: 3
## dfWithin: 8
```

(b) (1 point) Calculate the between-conditions, within-conditions, and total sums of squares.

ANSWER:

```r
SumX2<- sum(Cond1)+sum(Cond2)+sum(Cond3)+sum(Cond4)
SumXsq2<- sum(Cond1^2)+sum(Cond2^2)+sum(Cond3^2)+sum(Cond4^2)
SumTsq2<- sum(Cond1)^2+sum(Cond2)^2+sum(Cond3)^2+sum(Cond4)^2
```

```r
SSTot2<- SumXsq2 - SumX2^2/ N2
SSBet2<- (sum(Cond1)^2/n1 + sum(Cond2)^2/n2 + sum(Cond3)^2/n3 + sum(Cond4)^2/n4) - SumX2^2/ N2
SSWith2<- SSTot2-SSBet2
cat("Total sums of squares:", SSTot2, "\nBetween-conditions:", SSBet2, "\nWithin-conditions:", SSWith2)
```

```
## Total sums of squares: 56
## Between-conditions: 30
## Within-conditions: 26
```

(c) (1 point) Calculate the between-conditions and within-conditions mean squares, the F-statistic, and the critical value of F. Use significance level $\alpha = .05$.

ANSWER:

```
MSBet2<- SSBet2/dfBet2
MSWith2<- SSWith2/dfWith2
F2<- MSBet2/MSWith2
CriticalF2<- qf(0.05, dfBet2, dfWith2, lower.tail = FALSE)
cat("Between-conditions mean squares:", MSBet2,
    "\nWithin-conditions mean squares:", MSWith2, "\nF-statistic:", F2,
    "\nCritical value of F:", CriticalF2)
```

```
## Between-conditions mean squares: 10
## Within-conditions mean squares: 3.25
## F-statistic: 3.076923
## Critical value of F: 4.066181
```

(d) (1 point) Based on your calculated F and critical F value, should we reject the null hypothesis?

ANSWER:
Since the F value is smaller than the criticla F value, we fail to reject the null hypothesis.

Problem #3 (5 points)

It might be predicted that consumer buying behavior would vary with the location at which the product is displayed in the store. We therefore look at the purchases of well-known brands of candy bars when they are displayed on a shelf behind the counter, on the counter next to the cash register, on a special stand next to the counter, or in the "snacks" aisle of the store. The dependent variable is the number of candy bars sold per day. The data is shown below. Perform your analysis step-by-step using RStudio, but without using RStudio's *aov()* function.

```
## Behind Counter: 15 23 18 16 25 29 17
## On Counter: 24 14 15 19 30 26 18
## On Stand: 10 5 8 13 6 10 12
## Snack Aisle: 15 13 10 17 18 11 15
```

(a) (1 point) Calculate the between-conditions, within-conditions, and total degrees of freedom.

ANSWER

```r
n1<- length(BehindCounter)
n2<- length(OnCounter)
n3<- length(OnStand)
n4<- length(SnackAisle)
N<- n1+n2+n3+n4
```

```r
dfBet3<- 4-1
dfTot3<- N-1
dfWith3<- dfTot3-dfBet3
cat("dfTotal:", dfTot3, "\ndfBetween:", dfBet3, "\ndfWithin:", dfWith3)
```

```
## dfTotal: 27
## dfBetween: 3
## dfWithin: 24
```

(b) (2 points) Calculate the between-conditions, within-conditions, and total sums of squares.

ANSWER:

```r
SumX3<- sum(BehindCounter)+sum(OnCounter)+sum(OnStand)+sum(SnackAisle)
SumXsq3<- sum(BehindCounter^2)+sum(OnCounter^2)+sum(OnStand^2)+sum(SnackAisle^2)
SumTsq3<- sum(BehindCounter)^2+sum(OnCounter)^2+sum(OnStand)^2+sum(SnackAisle)^2
```

```r
SSTot3<- SumXsq3-SumX3^2/N
SSBet3<- SumTsq3/n1 - SumX3^2/N
SSWith3<- SSTot3-SSBet3
cat("Total sums of squares:", SSTot3, "\nBetween-conditions:", SSBet3, "\nWithin-conditions:", SSWith3)
```

```
## Total sums of squares: 1141.429
## Between-conditions: 655.1429
## Within-conditions: 486.2857
```

(c) (1 point) Calculate the between-conditions and within-conditions mean squares, the F-statistic, and the critical value of F. Use significance level $\alpha = .05$.

ANSWER:

```r
MSBet3<- SSBet3/dfBet3
MSWith3<- SSWith3/dfWith3
F3<- MSBet3/MSWith3
CriticalF3<-qf(0.05, dfBet3, dfWith3, lower.tail = FALSE)
cat("Between-conditions mean squares:", MSBet3,
    "\nWithin-conditions mean squares:", MSWith3, "\nF-statistic:", F3,
    "\nCritical value of F:", CriticalF3)
```

```
## Between-conditions mean squares: 218.381
## Within-conditions mean squares: 20.2619
## F-statistic: 10.77791
## Critical value of F: 3.008787
```

(d) (1 point) Based on your calculated F and critical F value, should we reject the null hypothesis (that the location where candy bars are displayed in store makes no difference)?

ANSWER:
Since the F value is larger than our critical value, we can reject the null hypothesis.

Problem #4 (3 points)

Perform an analysis of variance on the data from Problem #3 using RStudio's *aov()* function. Load the data from our course website (HW05_CandyData.csv) into a data frame named *CandyData*. The number of candy bars sold will be contained in the *NumSold* variable of the data frame (CandyData$NumSold) and the conditions will be contained in the *Location* variable (CandyData$Location).

(a) (1 point) Run the *aov()* function to perform your analysis, saving the output to your own variable. Display the contents of your output variable.

ANSWER:

```
CandyData<- read.csv("/Users/zzze/Downloads/HW05_CandyData.csv")
my.anova<- aov(CandyData$NumSold ~ CandyData$Location)
my.anova
```

```
## Call:
##    aov(formula = CandyData$NumSold ~ CandyData$Location)
##
## Terms:
##                 CandyData$Location Residuals
## Sum of Squares            655.1429  486.2857
## Deg. of Freedom                  3        24
##
## Residual standard error: 4.501323
## Estimated effects may be unbalanced
```

(b) (1 point) Use the *summary()* function to display your saved output from the *aov()* function.

```
summary(my.anova)
```

```
##                    Df Sum Sq Mean Sq F value   Pr(>F)
## CandyData$Location  3  655.1  218.38   10.78 0.000112 ***
## Residuals          24  486.3   20.26
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

(c) (1 point) Verify that your conclusion is the same as in Problem #3.

F value is the same as we got from Q3. Since the p-value <.05, we can reject the null hypothesis.

Problem #5 (3 points)

An experiment is conducted to determine the soil moisture deficit resulting from varying amounts of residual timber left after cutting trees in a forest. The three conditions are:

"None": no timber left

"2000BF": 2000 board feet left ("board feet" is a measurement of timber volume)

"8000BF": 8000 board feet left

Perform an analysis of variance using RStudio's aov() function. Load the data from our course website (HW05_MoistureData.csv) into a data frame named *MoistureData*. The soil moisture deficit will be contained in the *Deficit* variable of the data frame (MoistureData$Deficit) and the conditions (i.e. level of remaining timber) will be contained in the *RemTimber* variable (MoistureData$RemTimber).

(a) (1 point) Run the *aov()* function to perform your analysis, saving the output to your own variable. Display the contents of your output variable.

ANSWER:

```
MoistureData<- read.csv("/Users/zzze/Downloads/HW05_MoistureData.csv")
my.anova2<- aov(Deficit ~ RemTimber, MoistureData)
my.anova2
```

```
## Call:
##    aov(formula = Deficit ~ RemTimber, data = MoistureData)
##
## Terms:
##                 RemTimber Residuals
## Sum of Squares   5.279128  1.772483
## Deg. of Freedom         2        14
##
## Residual standard error: 0.3558173
## Estimated effects may be unbalanced
```

(b) (1 point) Use the *summary()* function to display your saved output from the *aov()* function.

```
summary(my.anova2)
```

```
##             Df Sum Sq Mean Sq F value   Pr(>F)
## RemTimber    2  5.279  2.6396   20.85 6.34e-05 ***
## Residuals   14  1.772  0.1266
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

(c) (1 point) Based on the above analysis, should we reject the null hypothesis (that the amount of timber remaining has no effect on soil moisture deficit)?

Based on the fact aht p-value<0.05, we can reject the null hypothesis.
