Homework7

Jyoti Chaudhary

April 8, 2018

Problem 1

This is the case of two way Anova without interaction. Here Factor A = portfolio and Factor B = Advisor. k = 3, b = 4.

```
data2 <- read.csv(paste(getwd(),"/XR12095.csv",sep = ""),header=T)


library(reshape2)
data3 <- melt(data2)
```

```
## Using Advisor as id variables
```

```
data3
```

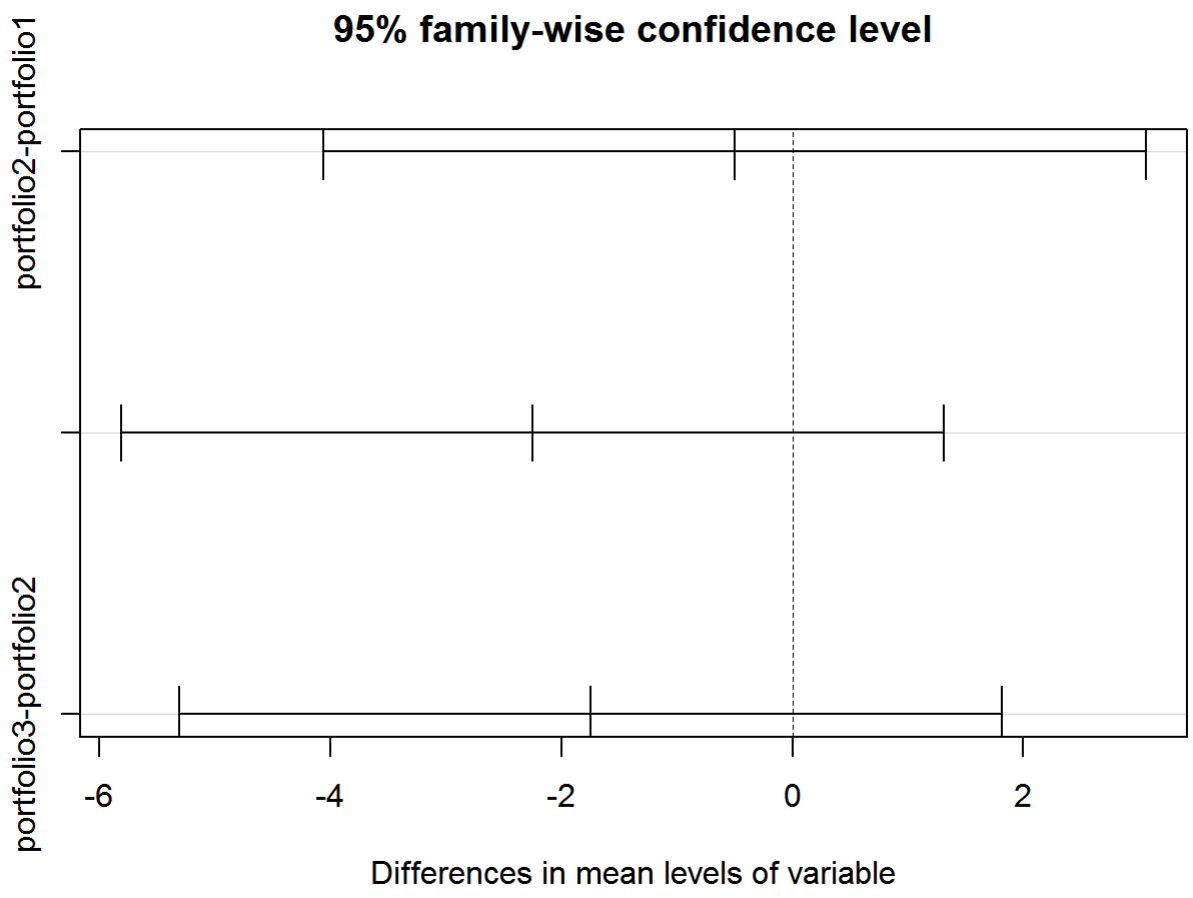
```
##     Advisor   variable value
## 1  advisorA portfolio1     8
## 2  advisorB portfolio1    12
## 3  advisorC portfolio1     8
## 4  advisorD portfolio1    15
## 5  advisorA portfolio2     8
## 6  advisorB portfolio2    10
## 7  advisorC portfolio2    11
## 8  advisorD portfolio2    12
## 9  advisorA portfolio3     5
## 10 advisorB portfolio3     8
## 11 advisorC portfolio3    10
## 12 advisorD portfolio3    11
```

```
mdl1 <- aov(value ~ Advisor + variable, data = data3)

# Now lets compare the portfolios
TukeyHSD(mdl1, conf.level = 0.95, "variable")
```

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = value ~ Advisor + variable, data = data3)
##
## $variable
##                          diff       lwr      upr     p adj
## portfolio2-portfolio1 -0.50 -4.061343 3.061343 0.9043374
## portfolio3-portfolio1 -2.25 -5.811343 1.311343 0.2084893
## portfolio3-portfolio2 -1.75 -5.311343 1.811343 0.3522676
```

```
plot(TukeyHSD(mdl1,"variable"))
```



From the output p-values and the plot, we can determine that there is no signficant difference between portfolios at 5% level.

```
with(data3, pairwise.t.test(x=value, g=variable, p.adjust="none")) ## FISHER'S LSD
```

```
##
##  Pairwise comparisons using t tests with pooled SD
##
## data:  value and variable
##
##            portfolio1 portfolio2
## portfolio2 0.80       -
## portfolio3 0.27       0.38
##
## P value adjustment method: none
```

None of the portfolio comparisons are statistically significant. Hence we conclude that there is no significant difference between the expected annual rate of return between portfolios at 5% level.

Problem 2:

This is one-way anova test.

```
data1 <- read.csv(paste(getwd(),"/XR12096.csv",sep = ""),header=T)

Stacked_data1 <- stack(data1)

mdl2 <- aov(values ~ ind, data = Stacked_data1)
```

Lets compare the brands using 0.025 level.

```
TukeyHSD(mdl2, conf.level = 0.975, "ind")
```

```
##   Tukey multiple comparisons of means
##     97.5% family-wise confidence level
##
## Fit: aov(formula = values ~ ind, data = Stacked_data1)
##
## $ind
##                     diff        lwr       upr     p adj
## brandB-brandA  4.583333 -1.5769755 10.743642 0.1249145
## brandC-brandA  1.666667 -4.4936422  7.826976 0.8315865
## brandD-brandA  5.483333 -0.6769755 11.643642 0.0513716
## brandC-brandB -2.916667 -9.0769755  3.243642 0.4663531
## brandD-brandB  0.900000 -5.2603088  7.060309 0.9674269
## brandD-brandC  3.816667 -2.3436422  9.976976 0.2440724
```

Looking at the p-value, we can further determine that at 0.025 level, there is no statistically significant difference between the breaking strength of different brand types.

```
with(Stacked_data1, pairwise.t.test(x=values, g=ind, p.adjust="none")) ## FISHER'S LSD
```

```
##
##  Pairwise comparisons using t tests with pooled SD
##
## data:  values and ind
##
##        brandA brandB brandC
## brandB 0.030  -      -
## brandC 0.407  0.154  -
## brandD 0.011  0.652  0.067
##
## P value adjustment method: none
```

Except for brandD-brandA, none of the other portfolio comparisons are statistically significant. For brandD-brandA, there is statistically significant difference between the breaking strength. For all other brand comparisons, we conclude that there is no significant difference between the breaking strength at 2.5% level.

Problem 3:

```
data4 <- read.csv(paste(getwd(),"/XR12099.csv",sep = ""),header=T)

data5 <- melt(data4)
```

```
## Using Style as id variables
```

```
mdl3 <- aov(value ~ Style * variable, data = data5)

summary(mdl3)
```

```
##                Df Sum Sq Mean Sq F value Pr(>F)
## Style           1   12.0   12.00   2.250 0.1843
## variable        2   64.5   32.25   6.047 0.0365 *
## Style:variable  2    9.5    4.75   0.891 0.4585
## Residuals       6   32.0    5.33
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
TukeyHSD(mdl3,"Style")
```

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = value ~ Style * variable, data = data5)
##
## $Style
##                 diff       lwr      upr     p adj
## TypSty2-TypSty1   -2 -5.262549 1.262549 0.1842807
```

Using 0.05 level of significance, we can conclude that Type style 1 and type style 2 has no significant difference in the effect on publication readability.

```
TukeyHSD(mdl3,"variable")
```

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = value ~ Style * variable, data = data5)
##
## $variable
##                   diff        lwr        upr     p adj
## TypMed-TypLight -5.25 -10.260471 -0.2395288 0.0417878
## TypDrk-TypLight -4.50  -9.510471  0.5104712 0.0738189
## TypDrk-TypMed    0.75  -4.260471  5.7604712 0.8922271
```

Using 0.05 level of significance, we can conclude that Only medium and light darkness levels have statistically significant difference in the effect on publication readability.

```
TukeyHSD(mdl3,"Style:variable")
```

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = value ~ Style * variable, data = data5)
##
## $`Style:variable`
##                                     diff        lwr       upr     p adj
## TypSty2:TypLight-TypSty1:TypLight -0.5  -9.691062  8.691062 0.9998933
## TypSty1:TypMed-TypSty1:TypLight   -5.0 -14.191062  4.191062 0.3670304
## TypSty2:TypMed-TypSty1:TypLight   -6.0 -15.191062  3.191062 0.2297666
## TypSty1:TypDrk-TypSty1:TypLight   -2.5 -11.691062  6.691062 0.8724125
## TypSty2:TypDrk-TypSty1:TypLight   -7.0 -16.191062  2.191062 0.1415192
## TypSty1:TypMed-TypSty2:TypLight   -4.5 -13.691062  4.691062 0.4569213
## TypSty2:TypMed-TypSty2:TypLight   -5.5 -14.691062  3.691062 0.2914508
## TypSty1:TypDrk-TypSty2:TypLight   -2.0 -11.191062  7.191062 0.9420843
## TypSty2:TypDrk-TypSty2:TypLight   -6.5 -15.691062  2.691062 0.1804384
## TypSty2:TypMed-TypSty1:TypMed     -1.0 -10.191062  8.191062 0.9970263
## TypSty1:TypDrk-TypSty1:TypMed      2.5  -6.691062 11.691062 0.8724125
## TypSty2:TypDrk-TypSty1:TypMed     -2.0 -11.191062  7.191062 0.9420843
## TypSty1:TypDrk-TypSty2:TypMed      3.5  -5.691062 12.691062 0.6693236
## TypSty2:TypDrk-TypSty2:TypMed     -1.0 -10.191062  8.191062 0.9970263
## TypSty2:TypDrk-TypSty1:TypDrk     -4.5 -13.691062  4.691062 0.4569213
```

Using 0.05 level of significance, we can conclude that the interaction effect of style type and darkness level do not have statistically significant difference in the effect on publication readability.

Problem 4

```
data6 <- read.csv(paste(getwd(),"/LtnSqr.csv",sep = ""),header=T)

mdl4 <- aov(Yield ~ Peanuts + Column + Row, data = data6)


TukeyHSD(mdl4,"Peanuts")
```

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = Yield ~ Peanuts + Column + Row, data = data6)
##
## $Peanuts
##       diff        lwr      upr     p adj
## B-A  4.075 -0.8189415 8.968942 0.0990478
## C-A  0.825 -4.0689415 5.718942 0.9334796
## D-A  0.275 -4.6189415 5.168942 0.9971031
## C-B -3.250 -8.1439415 1.643942 0.2001869
## D-B -3.800 -8.6939415 1.093942 0.1252130
## D-C -0.550 -5.4439415 4.343942 0.9782233
```

Using 0.05 level of significance, we can conclude that the comparison in difference of yield of various peanuts varities is not statistically significant.