PSTAT 122 Final Exam

Justin Lee

6/11/2020

Problem 1

Hypothesis Test

$H_o : \mu_A = \mu_B = \mu_C = \mu_D = \mu_E = \mu_F$
vs.
H_A : not H_o

Setting up data and ANOVA

```
obs <- c(73, 75, 69, 68, 69, 80,
         70, 82, 67, 77, 63, 87,
         71, 82, 76, 76, 65, 86,
         65, 80, 74, 73, 65, 84,
         76, 86, 74, 71, 58, 78)
fert <- factor(rep(c('A', 'B', 'C', 'D', 'E', 'F'), times = 5) )


fert_a <- c(73, 70, 71, 65, 76)
fert_b <- c(75,82,82,80,86)
fert_c <- c(69,67,76,74,74)
fert_d <- c(68,77,76,73,71)
fert_e <- c(69,63,65,65,58)
fert_f <- c(80,87,86,84,78)

mean(fert_a)
```

```
## [1] 71
```

```
mean(fert_b)
```

```
## [1] 81
```

```
mean(fert_c)
```

```
## [1] 72
```

```
mean(fert_d)
```

```
## [1] 73
```

```r
mean(fert_e)
```

```
## [1] 64
```

```r
mean(fert_f)
```

```
## [1] 83
```

```r
ybar <- mean(obs)


n <- length(obs)


a <- aov(obs~fert)
summary(a)
```

```
##             Df Sum Sq Mean Sq F value   Pr(>F)
## fert         5   1220  244.00      16 5.78e-07 ***
## Residuals   24    366   15.25
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
alpha <- 0.05
qf(1-alpha , df1 = 5, df2 = 24)
```

```
## [1] 2.620654
```

As seen from the summary of the ANOVA tables above, we can see that the p value is extremely small.

Since P-value is 5.78e-07 and is less than 0.05. We can conclude that there is evidence to indicate that types of fertilizer effect yield at $\alpha = 0.05$.

$\hat{\mu} = 74$
$y_{ij} = \mu + \tau_i + \epsilon_{ij}$
$i = 1, 2, 3, 4, 5, 6$
$j = 1, 2, 3, 4, 5$

$$\sum_{i=1}^{6} \tau_i = 0$$

$\epsilon_{11}, \epsilon_{12}, ..., \epsilon_{65} \overset{iid}{\sim} N(0, \sigma^2)$

LSD TEST

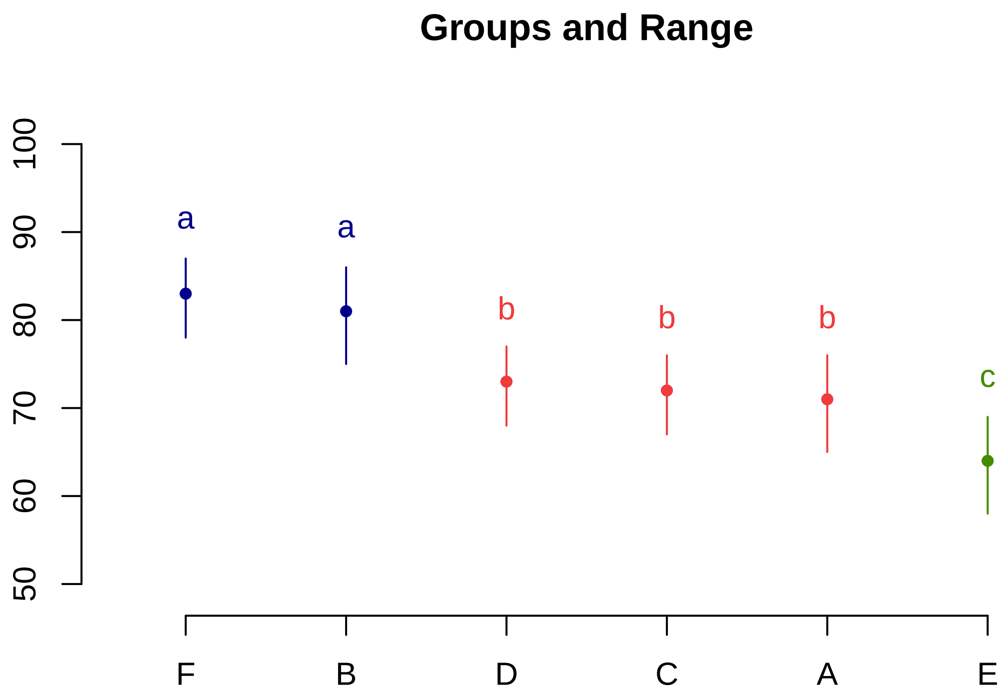
```r
library(agricolae)
```

```
## Warning: package 'agricolae' was built under R version 3.6.3
```

```r
LSD.test(a,"fert", console = T)
```

```
##
## Study: a ~ "fert"
##
## LSD t Test for obs
##
## Mean Square Error:  15.25
##
## fert,  means and individual ( 95 %) CI
##
##   obs      std r      LCL      UCL Min Max
## A  71 4.062019 5 67.39556 74.60444  65  76
## B  81 4.000000 5 77.39556 84.60444  75  86
## C  72 3.807887 5 68.39556 75.60444  67  76
## D  73 3.674235 5 69.39556 76.60444  68  77
## E  64 4.000000 5 60.39556 67.60444  58  69
## F  83 3.872983 5 79.39556 86.60444  78  87
##
## Alpha: 0.05 ; DF Error: 24
## Critical Value of t: 2.063899
##
## least Significant Difference: 5.097453
##
## Treatments with the same letter are not significantly different.
##
##   obs groups
## F  83      a
## B  81      a
## D  73      b
## C  72      b
## A  71      b
## E  64      c
```

```r
plot(LSD.test(a,"fert"))
```

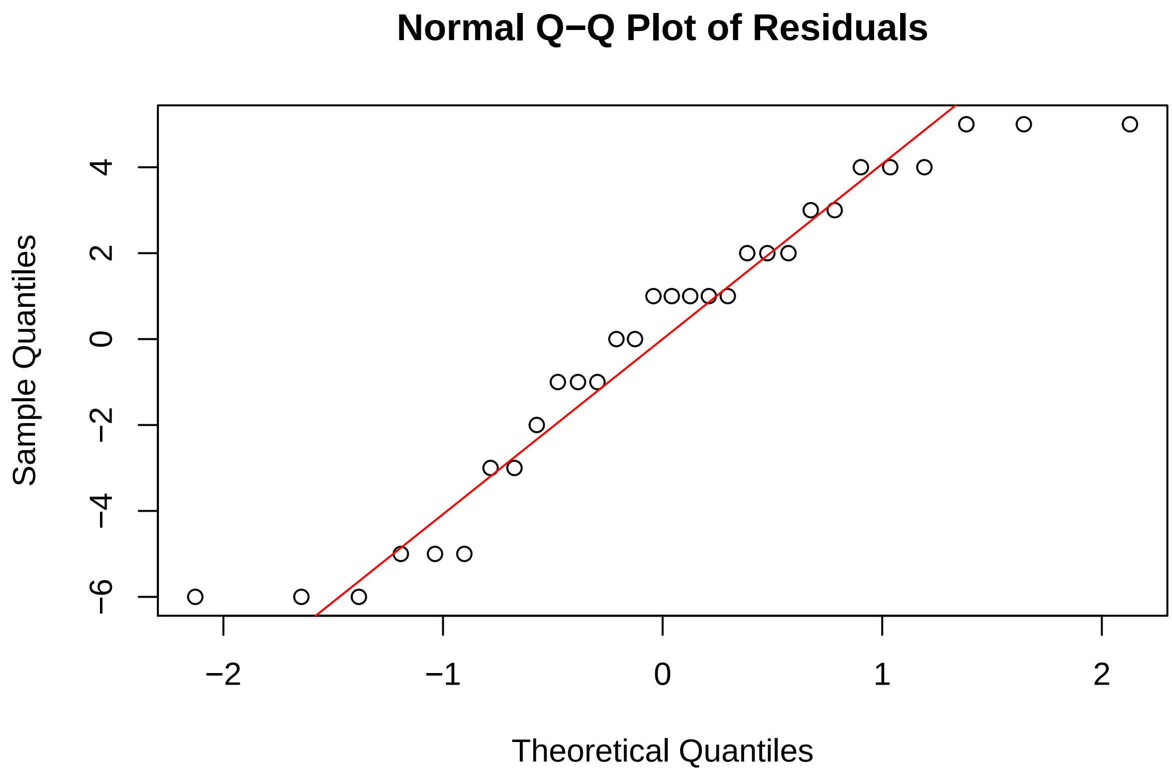


From the LSD test we can first see that fertilizer F and B are not significantly different. Also fertilizer D,C, and A are not significantly different.

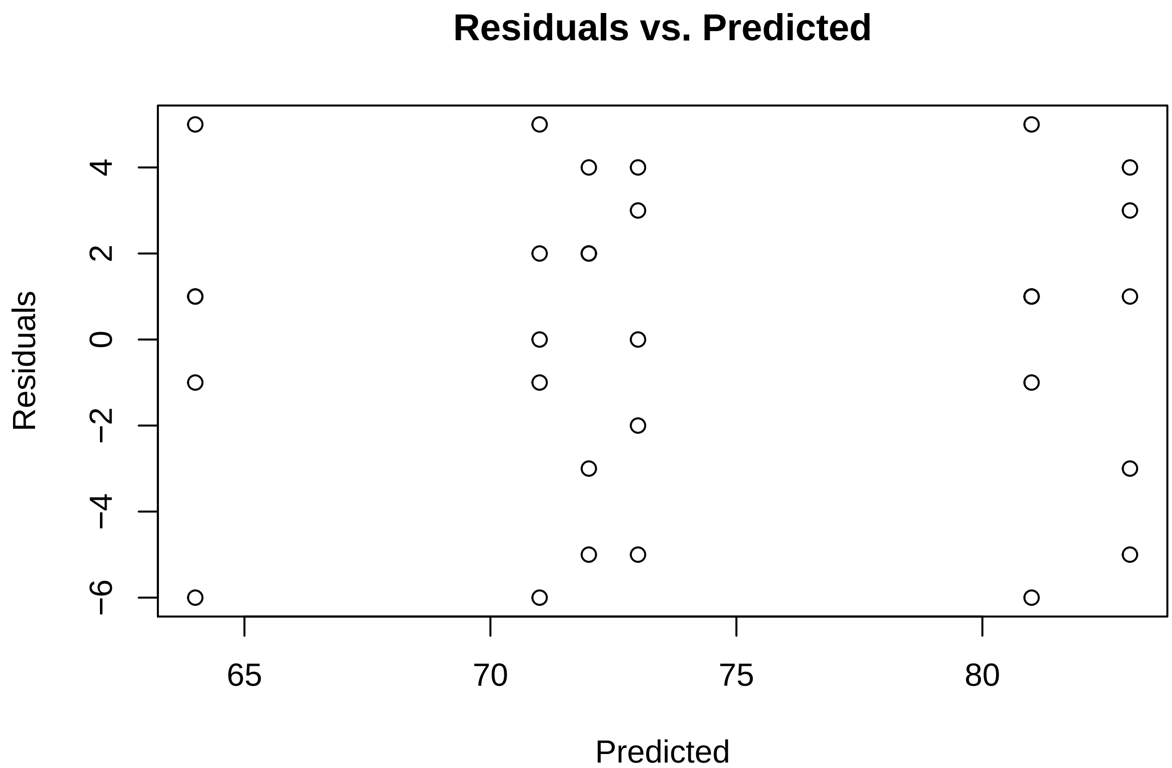
If we're going for the highest production of wheat, fertilizer F seems like the way to go with an average of 83 bushels of wheat produced with a standard deviation of 3.872983.

Analyzing Residuals and model adequacy

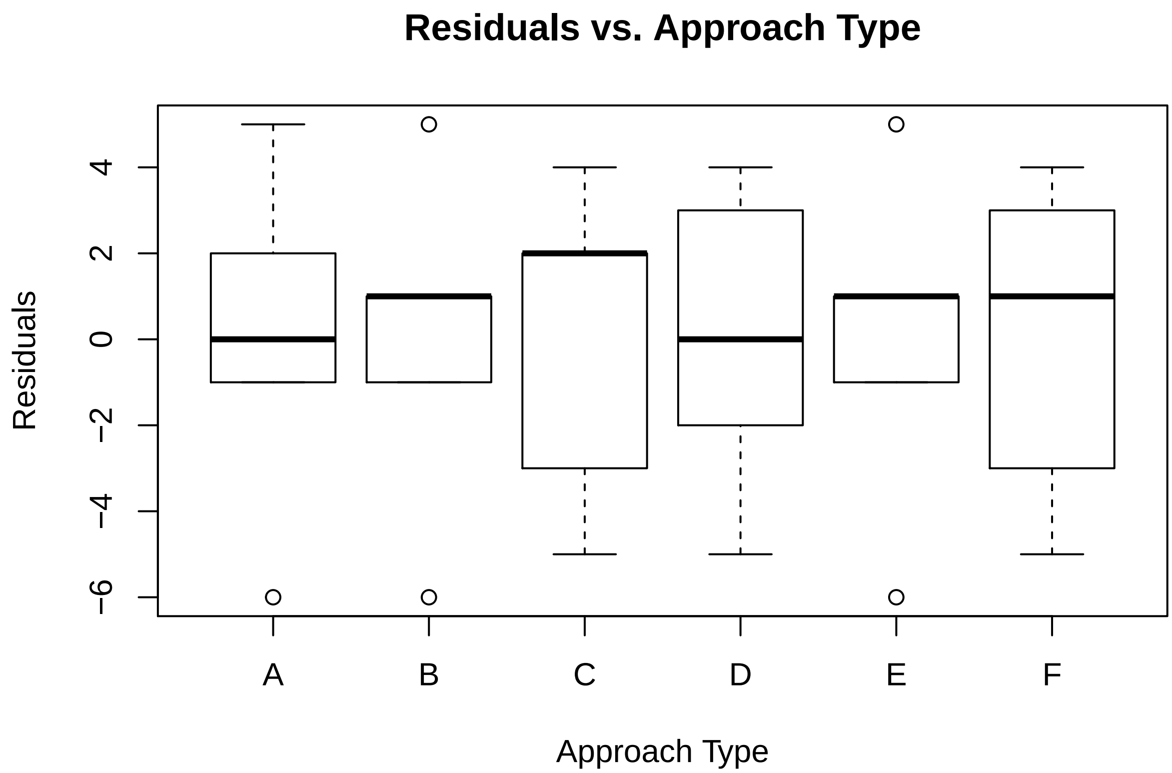
```
qqnorm(a$residuals, main = "Normal Q-Q Plot of Residuals")
qqline(a$residuals, col="red")
```




```r
plot(a$residuals~a$fitted.values,main="Residuals vs. Predicted",xlab="Predicted",ylab="Residuals")
```




```r
plot(a$residuals~fert,main="Residuals vs. Approach Type",xlab="Approach Type",ylab="Residuals")
```



Looking at the graphs and analyzing the residuals from this experiment, we can see that there is nothing unusual in the residual plots. The normal Q-Q plots of residuals show that the points follow the line.

Problem 2

Hypothesis Test

$H_0 : (\tau_\beta)_{ij} = 0$ vs. $H_A : (\tau_\beta)_{ij} \neq 0$

Setting up data and ANOVA

```
obs <- c(147, 153, 182, 178, 128, 124,
         157, 161, 144, 150, 121, 119,
         132, 126, 156, 162, 109, 107)

temp <- factor(rep(c(15, 70, 125),each = 2, times = 3))

type <- factor(rep(c('I','II', 'III'),each = 6))


twoway.df <- data.frame(obs = obs,  type = type, temp = temp)
xtabs(~ type + temp, data= twoway.df)
```

```
##       temp
## type  15 70 125
```

```
##   I    2  2   2
##   II   2  2   2
##   III  2  2   2
```

```r
twoway.df
```

```
##    obs type temp
## 1  147    I   15
## 2  153    I   15
## 3  182    I   70
## 4  178    I   70
## 5  128    I  125
## 6  124    I  125
## 7  157   II   15
## 8  161   II   15
## 9  144   II   70
## 10 150   II   70
## 11 121   II  125
## 12 119   II  125
## 13 132  III   15
## 14 126  III   15
## 15 156  III   70
## 16 162  III   70
## 17 109  III  125
## 18 107  III  125
```

```r
mean(obs)
```

```
## [1] 142
```

```r
meas.aov <- aov(obs~temp*type)
summary(meas.aov)
```

```
##             Df Sum Sq Mean Sq F value   Pr(>F)
## temp         2   5952  2976.0   267.8 9.58e-09 ***
## type         2   1200   600.0    54.0 9.71e-06 ***
## temp:type    4   1200   300.0    27.0 5.00e-05 ***
## Residuals    9    100    11.1
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
qf(1-alpha,4,9)
```

```
## [1] 3.633089
```

asdfasdfasdfasdfasdfasdfasdfasdfasdfasdfasdfaaaaaaaaaaaaaaaaaaaaaaa

$F_{4,9,0.05} = 3.633089$

Model

$Y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \epsilon_{ijk}$

$\mu = 142$
$\sum_{i=1}^{a} \tau_i = 0$
$\sum_{j=1}^{b} \beta_j = 0$
$\sum_{j=1}^{b} (\tau\beta)_{ij} = 0$
$\epsilon ijk \overset{iid}{\sim} N(0, \sigma^2)$

Interaction

$H_0 : (\tau\beta)_{ij} = 0 \ for \ all \ i, j \ vs. H_1 : at \ least \ one \ (\tau\beta)_{ij} \neq 0$

$Under \ H_0 : \frac{MS_{AB}}{MS_E} \sim F_{(a-1)(b-1), ab(n-1)}$

$\frac{300}{11.1} = 27.0 > F_{4,9,0.05} = 3.633089$

Reject null hypothesis because there are significant interaction.

Main Effect of Temperature (Factor A)

$H_0 : \tau_i = 0$ for all i vs.$H_1 : not \ H_0$

$Under \ H_0 : \frac{MS_A}{MS_E} \sim F_{a-1, ab(n-1)}$

```r
qf(0.95,2,4)
```

```
## [1] 6.944272
```

$\frac{2976}{11.1} = 267.8 > F_{2,4,0.05} = 6.944272$

Reject null hypothesis and conclude that temperature is a significant factor that affects battery life.

Main Effect of material type (Factor B)

$H_0 : \beta_j = 0$ fot all j vs.$H_1 : not \ H_0$

$Under \ H_0 : \frac{MS_B}{MS_E} \sim F_{b-1, ab(n-1)}$

$\frac{600}{11.1} = 54.05405 > F_{2,4,0.05} = 6.944272$

Reject null hypothesis and conclude that material type does affect battery life.