# Homework 7 Ishita Dutta

Ishita Dutta

11/17/2021

1a) Age is a good blocking factor here because people of different age groups have different reactions to various fat diets and various amounts of fat absorption compared to each other.

1b)

```
Treatment_means=c(0.73,0.67,0.15, 0.86, 0.75, 0.21, 0.94, 0.81, 0.26, 1.40, 1.32, 0.75, 1.62, 1.41, 0.78
Y_fit=matrix(Treatment_means,nrow=5,byrow=T)
library(reshape2); library(ggplot2)
ggfit = melt(Y_fit, varnames = c("Ages", "Content"), value.name =  "tmean")
ggplot(data = ggfit, aes(x = Content, y = tmean, group = Ages, col = Ages)) +
  geom_point(size = 4) +
  geom_line(size = 2) +
  theme_bw() +
  ylab("Treatment Means")
```
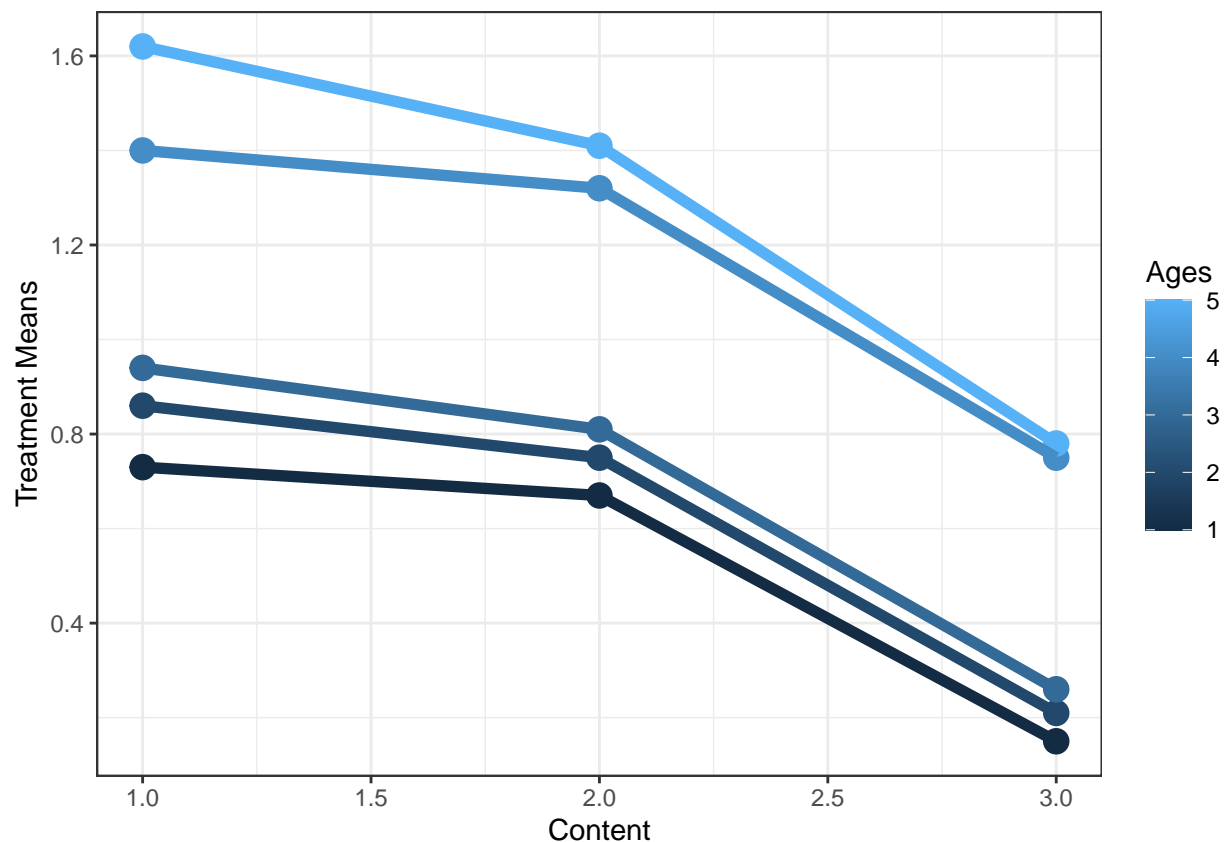

This plot points out the no interaction assumption of the RCBD model, as all of the lines are relatively

parallel and close to each other, especially within the lower age groups. In addition, we can also point out that there might be some changes between the Fairly Low diet and the Moderately low diet as there seems to be a dip in all 5 of the interaction plots when we look between these two treatments.

1c) H0 -> D = 0, Ha -> D != 0 At significance level 0.01, reject H0 if and only if F* > F(1 - alpha, 1, ab - a - b)

```
Y_fit2<-Y_fit
a=5
b=3
n=1
Y_bar_A = apply(Y_fit2, 1, mean)
Y_bar_B = apply(Y_fit2, 2, mean)
Y_bar = mean(Y_fit2)
SSA = n*b*sum((Y_bar_A-Y_bar)^2)
SSB = n*a*sum((Y_bar_B-Y_bar)^2)
SSTO= n*sum((Y_fit2-Y_bar)^2)
SSAB = SSTO-SSA-SSB
SSAB_star<-(t(Y_bar_A-Y_bar)%*%Y_fit2%*%(Y_bar_B-Y_bar))^2/(sum((Y_bar_A-Y_bar)^2)*sum((Y_bar_B-Y_bar)^2
SSRem<-SSTO-SSA-SSB-SSAB_star
df_rem<-a*b-a-b
F_tukey<-(SSAB_star/1)/(SSRem/df_rem)
F_tukey
```

```
##           [,1]
## [1,] 6.445275
```

```
qf(0.99,1,df_rem)
```

```
## [1] 12.24638
```

```
Y_fit3 = read.table("diet.txt")
```

As F* < F(0.99, 1, ab - a - b), we fail to reject H0 and keep the status quo that there is no significant interaction effect, which is what we have visualized in the graph.

1d)

```
data <- read.table("diet.txt")
```

```
table(data[,1:3])
```

```
## , , V3 = 1
##
##       V2
## V1     1 2 3 4 5
##    0.15 0 0 0 0 0
##    0.21 0 0 0 0 0
##    0.26 0 0 0 0 0
##    0.67 0 0 0 0 0
##    0.73 1 0 0 0 0
##    0.75 0 0 0 0 0
```

2

```
##    0.78 0 0 0 0 0
##    0.81 0 0 0 0 0
##    0.86 0 1 0 0 0
##    0.94 0 0 1 0 0
##    1.32 0 0 0 0 0
##    1.4  0 0 0 1 0
##    1.41 0 0 0 0 0
##    1.62 0 0 0 0 1
##
## , , V3 = 2
##
##         V2
## V1     1 2 3 4 5
##    0.15 0 0 0 0 0
##    0.21 0 0 0 0 0
##    0.26 0 0 0 0 0
##    0.67 1 0 0 0 0
##    0.73 0 0 0 0 0
##    0.75 0 1 0 0 0
##    0.78 0 0 0 0 0
##    0.81 0 0 1 0 0
##    0.86 0 0 0 0 0
##    0.94 0 0 0 0 0
##    1.32 0 0 0 1 0
##    1.4  0 0 0 0 0
##    1.41 0 0 0 0 1
##    1.62 0 0 0 0 0
##
## , , V3 = 3
##
##         V2
## V1     1 2 3 4 5
##    0.15 1 0 0 0 0
##    0.21 0 1 0 0 0
##    0.26 0 0 1 0 0
##    0.67 0 0 0 0 0
##    0.73 0 0 0 0 0
##    0.75 0 0 0 1 0
##    0.78 0 0 0 0 1
##    0.81 0 0 0 0 0
##    0.86 0 0 0 0 0
##    0.94 0 0 0 0 0
##    1.32 0 0 0 0 0
##    1.4  0 0 0 0 0
##    1.41 0 0 0 0 0
##    1.62 0 0 0 0 0
```

```r
a = 5; b = 3; n = 1
Y = array(0, c(a, b, n))
for(i in 1:nrow(data)){
  Y[data[i,2], data[i,3], data[i,4]] = data[i, 1]
}
Y
```

```
## , , 1
##
##      [,1] [,2] [,3]
## [1,]    0    0    0
## [2,]    0    0    0
## [3,]    0    0    0
## [4,]    0    0    0
## [5,]    0    0    0
```

```
#Fitted  values
Y_fit = apply(Y, c(1,2), mean)
rownames(Y_fit) = paste0("A(i=", 1:5, ")")
colnames(Y_fit) = paste0("B(i=", 1:3, ")")

library(knitr)
knitr::kable(Y_fit)
```

	B(i=1)	B(i=2)	B(i=3)
A(i=1)	0	0	0
A(i=2)	0	0	0
A(i=3)	0	0	0
A(i=4)	0	0	0
A(i=5)	0	0	0

```
#Residuals
e = array(0, c(a,b,n))
for(i in 1:n) e[,,i] = Y[,,i] - Y_fit

e
```
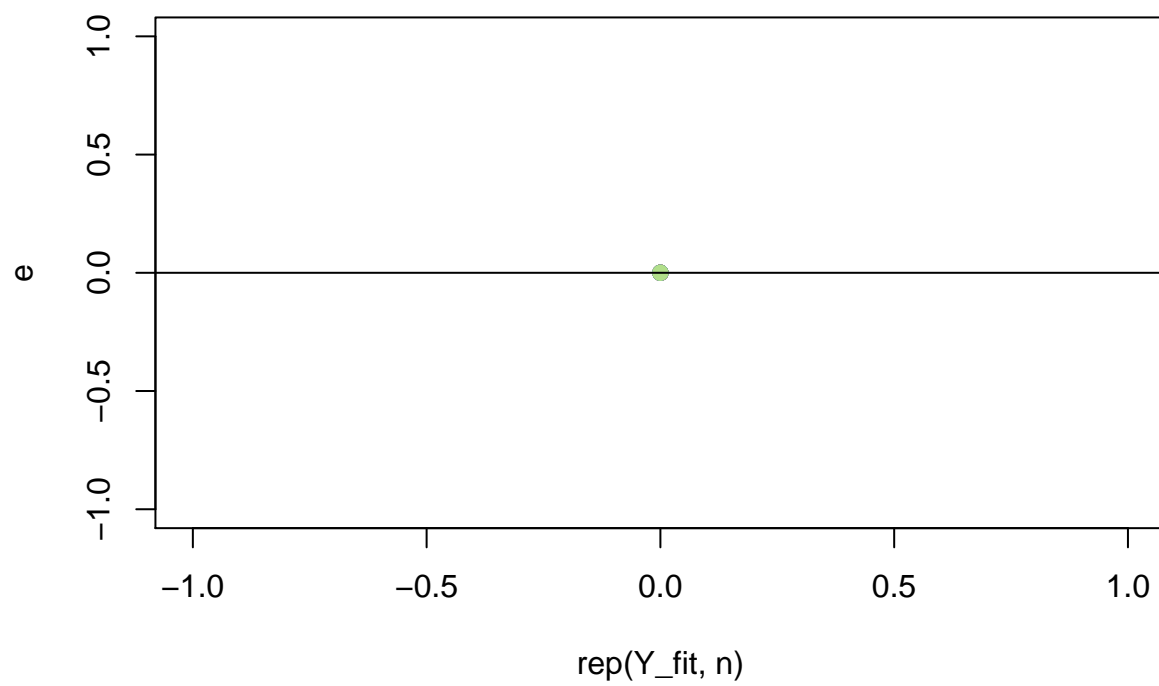
```
## , , 1
##
##      [,1] [,2] [,3]
## [1,]    0    0    0
## [2,]    0    0    0
## [3,]    0    0    0
## [4,]    0    0    0
## [5,]    0    0    0
```

```
#Residuals v.s. fitted
library(RColorBrewer)
number_of_treatment_means = a*b
colors <- rep(brewer.pal(n = number_of_treatment_means, name = "Paired"), n)
```

```
## Warning in brewer.pal(n = number_of_treatment_means, name = "Paired"): n too large, allowed maximum
## Returning the palette you asked for with that many colors
```

```
plot(rep(Y_fit, n), e, col = colors, pch = 19)
abline(h = 0)
```

```
#QQ plot
qqnorm(e, col = colors, pch = 19)
qqline(e)
```

Normal Q–Q Plot
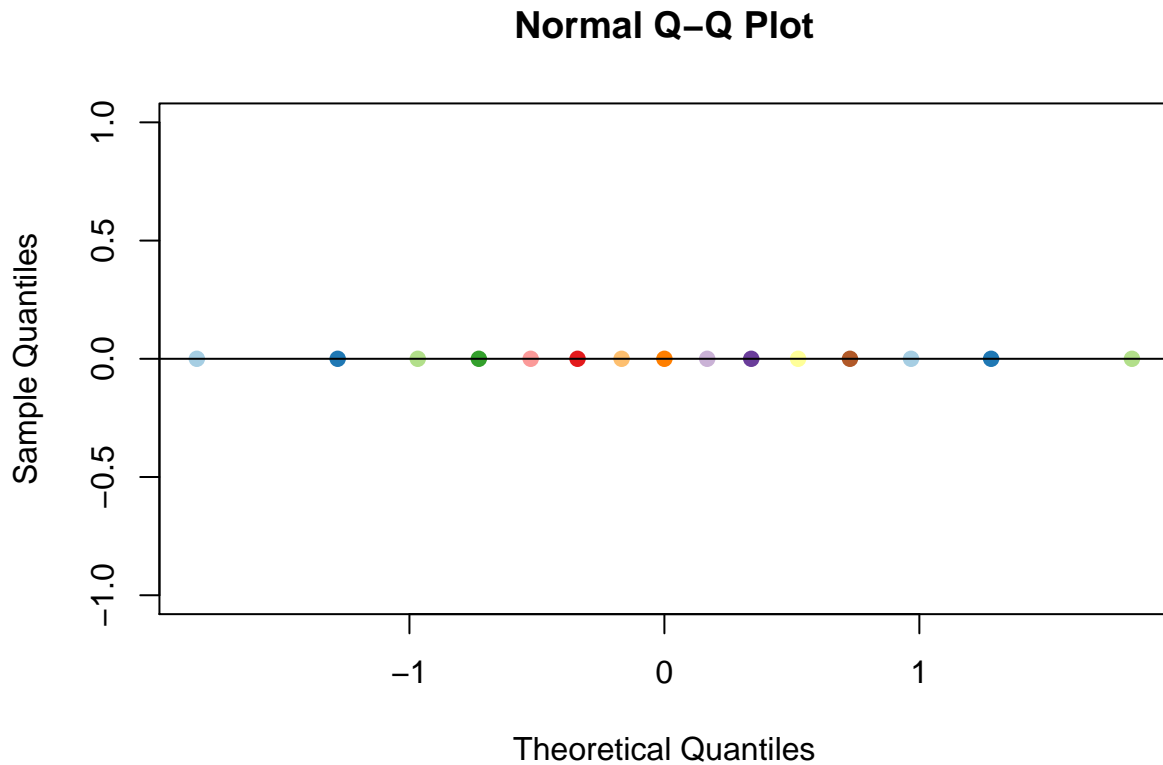
Residuals on qqline, showing a normal distribution among the values.

2a)

```r
block   <- factor(rep(1:5, times = 3))
variety <- rep(c("Extremely Low", "Fairly Low", "Moderately Low"), each = 5)
#yield  <- c(133.25, 113.25, 86.75, 108, 95.5, 90.25,
            #129.75, 121.25, 118.5, 95, 85.25, 109,
            #143, 87.25,  82.5, 91.5, 92, 89.5)
oat <- data.frame(block, variety, Treatment_means)
fit.oat <- aov(Treatment_means ~ block + variety)
summary(fit.oat)
```

```
##             Df Sum Sq Mean Sq F value Pr(>F)
## block        4 0.0881  0.0220   0.097  0.981
## variety      2 0.8478  0.4239   1.861  0.217
## Residuals    8 1.8226  0.2278
```

2b) H0 –> D = 0, Ha –> D != 0 At significance level 0.01, reject H0 if and only if $F^* > F(1 - alpha, 1, ab - a - b)$

```r
Fcrit = 0.4239 / 0.2278
Fcrit
```

```
## [1] 1.860843
```

6

```r
qf(0.95, 2, 8)
```

```
## [1] 4.45897
```

Fail to reject null hypothesis. There is no apparent interaction between the lipid levels over the three diets.

2c)

3a) We should use a single factor anova model for this data set as the independent variable is our chosen machine. The 10 coils chosen from that machine should be viewed as 10 repetitions of an experiment.

3b)

```r
coils = read.table("coil.txt")
Values = coils[,1]
Machine = coils[,2]
one.way <- aov(Values ~ Machine)
summary(one.way)
```

```
##             Df Sum Sq Mean Sq F value Pr(>F)
## Machine      1   58.3   58.32   2.765  0.105
## Residuals   38  801.6   21.09
```

3c) Ho -> u1 = u2 = u3 = u4 Ha -> u1 != u2 != u3 != u4 Reject if p-value < alpha = 0.1 p-val = 0.105 from the ANOVA table above. When testing at significance level 0.1, this means we fail to reject the Ho, meaning that the four machines produce similarly wound coils(ie. there is no major statistical difference in the coils as a result of the machines).