Pstat122_HW2

Celeste Herrera

4/22/2020

3.14 A pharmaceutical maunfacturer wants to investigate the bioactivity of a new drug. A completely randomized single-factor experiment was conducted with three dosage levels, and the following results were obtained.

$H_0 : \mu_{20} = \mu_{30} = \mu_{40}$

H_1 : At least one μ_i is different

(a) Is there evidence to indicate that the dosage level affects the bioactivity? $\alpha = .05$

```
experiment <-data.frame(observations=c(24,28,37,30,
                        37,44,31,35,
                        42,47,52,38),dosage= factor(rep(seq(20,40,by=10),each=4)))
experiment
```

```
##    observations dosage
## 1            24     20
## 2            28     20
## 3            37     20
## 4            30     20
## 5            37     30
## 6            44     30
## 7            31     30
## 8            35     30
## 9            42     40
## 10           47     40
## 11           52     40
## 12           38     40
```

```
dosage= factor(rep(seq(20,40,by=10),each=4))  # needed to be redefined for knitting purposes
n = 4  # The number values of observations
a = 3  # number of levels of dosage
N = n*a  # total number of observations
level = aov(observations~dosage, data= experiment)
summary(level)
```

```
##             Df Sum Sq Mean Sq F value Pr(>F)
## dosage       2  450.7  225.33   7.036 0.0145 *
## Residuals    9  288.2   32.03
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Since the p-value is $0.0145 < \alpha = .05$, reject the null hypothesis. There is sufficient evidence to suggest that there is a difference in the bioactivity at the $\alpha = 0.05$ level.

(b) If it is appropiate to do so, make comparisons between the parts of means. What conclusions can you draw?

$$H_1 : \mu_{20} = \mu_{30}$$

```
library(contrast)
fit1 <-lm(observations~dosage, data= experiment)
contrast(fit1, list(dosage = as.factor(20)),list(dosage = as.factor(30)))
```

```
## lm model parameter contrast
##
##   Contrast     S.E.     Lower    Upper     t df Pr(>|t|)
## 1       -7 4.001736 -16.05256 2.052555 -1.75  9   0.1142
```

From this contrast, we fail to reject the null hypothesis since the p-value is bigger than our $\alpha = .05$. Since we failed to reject, we can conclude the dosages of 20 and 30 do not have affect on the bioactivity.

$$H_2 : \mu_{20} = \mu_{40}$$

```
contrast(fit1, list(dosage = as.factor(20)),list(dosage = as.factor(40)))
```

```
## lm model parameter contrast
##
##   Contrast     S.E.     Lower     Upper     t df Pr(>|t|)
## 1      -15 4.001736 -24.05256 -5.947445 -3.75  9   0.0046
```

From the contrast, we reject the null hypothesis because the p-value is less than $\alpha = .05$. So, we can conclude that the dosages of 20 and 40 do affect the bioactivity.

$$H_3 : \mu_{30} = \mu_{40}$$

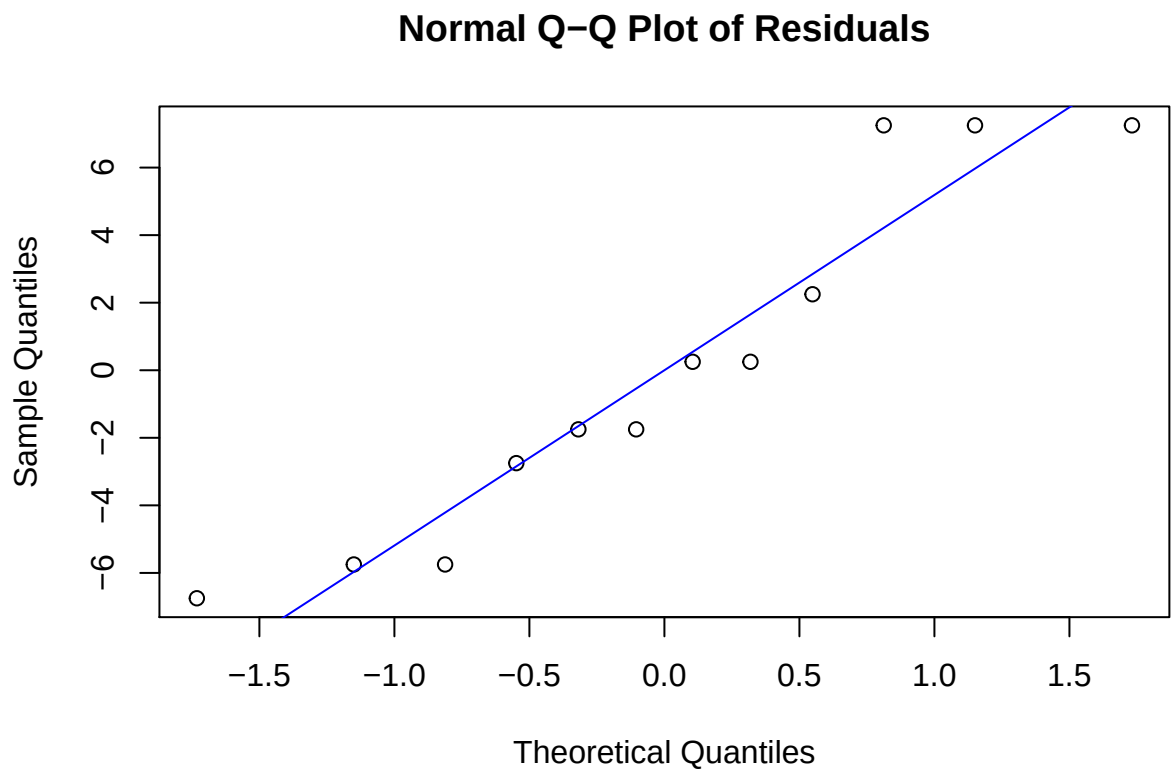
```
contrast(fit1, list(dosage = as.factor(30)),list(dosage = as.factor(40)))
```

```
## lm model parameter contrast
##
##   Contrast     S.E.     Lower    Upper  t df Pr(>|t|)
## 1       -8 4.001736 -17.05256 1.052555 -2  9   0.0767
```

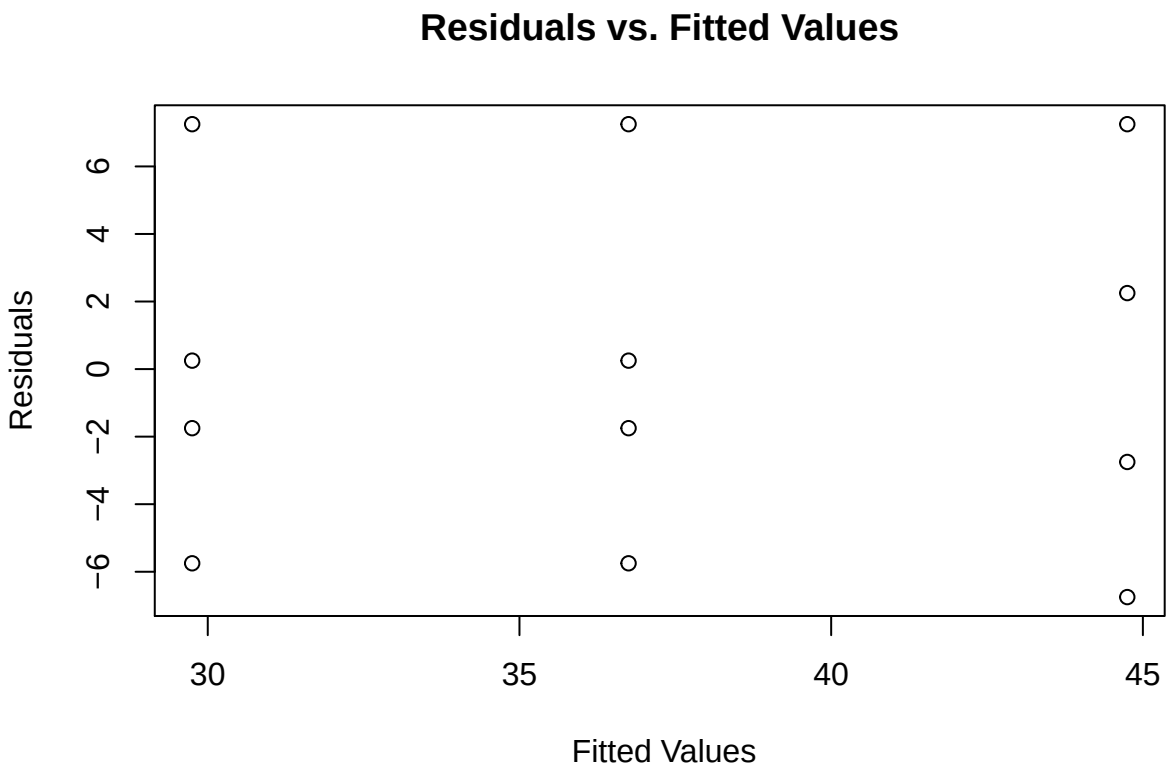
From the contrast, we fail to reject the null hypothesis since the p-value is bigger than our $\alpha = .05$. Since we failed to reject, we can conclude the dosages of 30 and 40 do not have affect on the bioactivity.

(c)Analyze the residuals from the experiment and comment on model adequacy

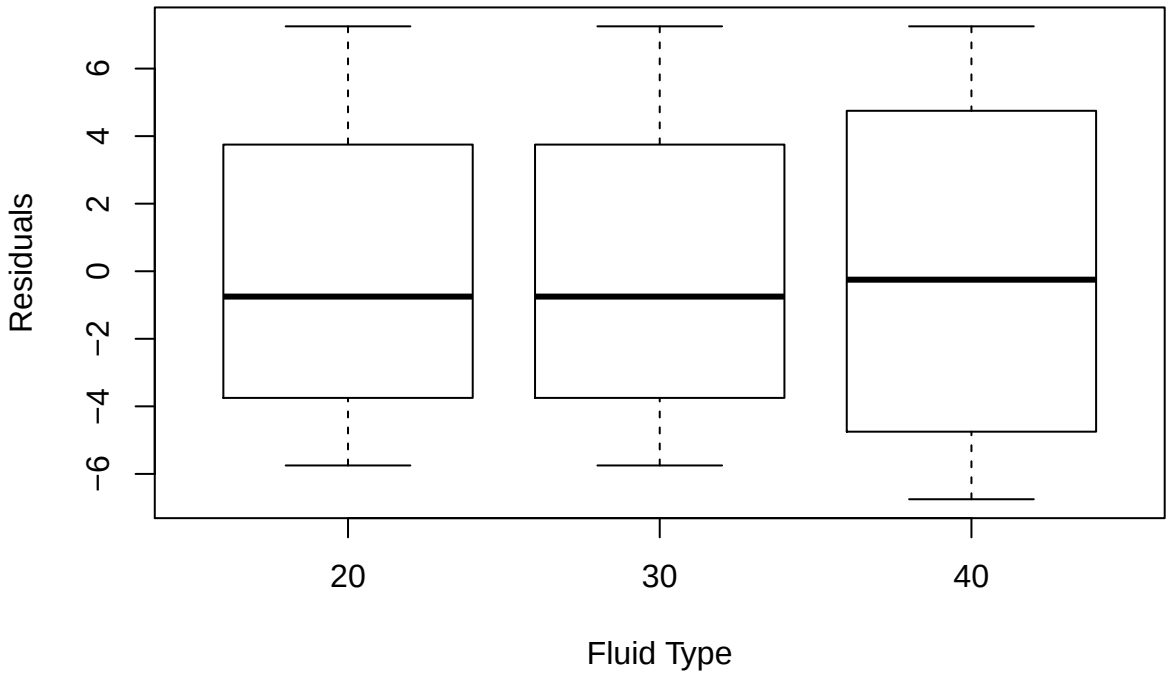
```
qqnorm(level$residuals, main = 'Normal Q-Q Plot of Residuals')
qqline(level$residuals, col= 'blue')
```




```
plot(level$residuals~level$fitted.values, main = "Residuals vs. Fitted Values", xlab = " Fitted Values"
```




```
plot(level$residuals~dosage, main = "Residuals vs. Dosage", xlab= "Fluid Type", ylab= "Residuals")
```



Considering that the Normal Q-Q plot gave a good approximation along the diagonal, we can assume that the normality assumption is satisfied. Although it was safe to note that there is no type of given pattern for the residuals vs. fit plot it is safe to assume that the equal variance assumption is satisied. And lastly there is a consideration of the data being collected randomly during a randomized experiment, it is safe to asssume that the independence assumption is completely satisfied.

3.17 A regional opera company has tried three approaches to solicit donations from 24 potentials sponsors. The 24 potential sponsors were randomly divided into three groups of eight, and one approach wass used for each group. The dollar amounts of the resulting contributions are shown in the following table.

(a) Do the data indicate that there is a difference in results obtained from the three different apporaches? Use $\alpha = .05$

$H_0 : \mu_{approach1} = \mu_{approach2} = \mu_{approach3}$

H_1 : At least one μ_i is different

```
Contributions<-c(1000,1500,1200,1800,1600,1100,1000,1250,
   1500,1800,2000,1200,2000,1700,1800,1900,
   900,1000,1200,1500,1200,1550,1000,1100)
Approach <- factor(rep(1:3,each=8))
cbind(Approach,Contributions)
```

```
##       Approach Contributions
##  [1,]        1          1000
##  [2,]        1          1500
##  [3,]        1          1200
##  [4,]        1          1800
##  [5,]        1          1600
##  [6,]        1          1100
```

```
##  [7,]        1          1000
##  [8,]        1          1250
##  [9,]        2          1500
## [10,]        2          1800
## [11,]        2          2000
## [12,]        2          1200
## [13,]        2          2000
## [14,]        2          1700
## [15,]        2          1800
## [16,]        2          1900
## [17,]        3           900
## [18,]        3          1000
## [19,]        3          1200
## [20,]        3          1500
## [21,]        3          1200
## [22,]        3          1550
## [23,]        3          1000
## [24,]        3          1100
```

```
a<- aov(Contributions~Approach)
a
```

```
## Call:
##    aov(formula = Contributions ~ Approach)
##
## Terms:
##                 Approach Residuals
## Sum of Squares   1362708   1520625
## Deg. of Freedom        2        21
##
## Residual standard error: 269.0924
## Estimated effects may be unbalanced
```

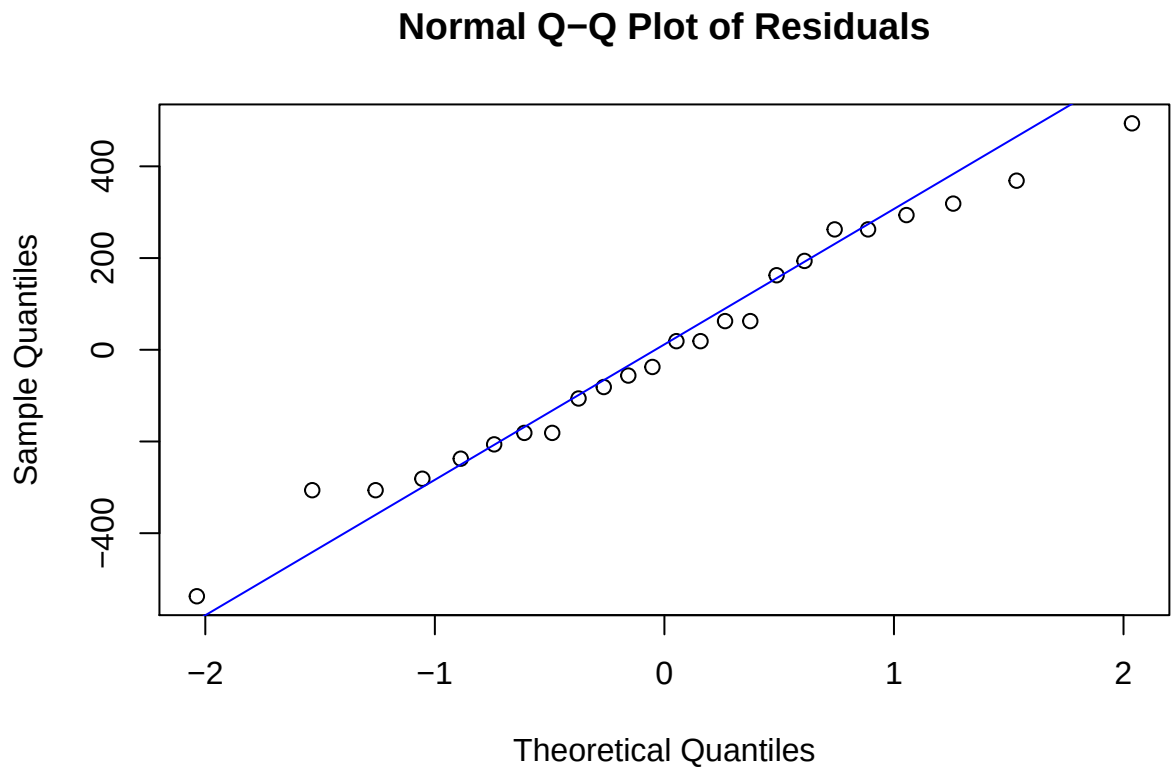
```
summary(a)
```

```
##             Df  Sum Sq Mean Sq F value  Pr(>F)
## Approach     2 1362708  681354    9.41 0.00121 **
## Residuals   21 1520625   72411
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

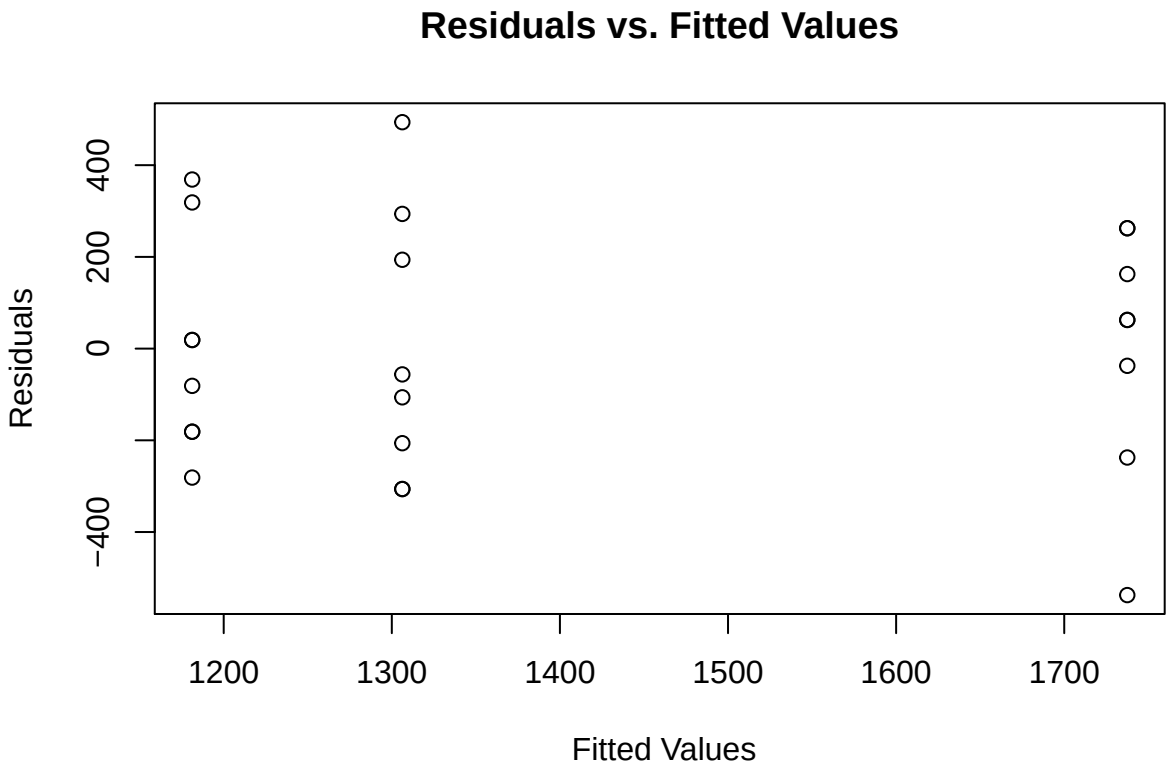
Since the p-value is .00121 $< \alpha = .05$, reject the null hypothes. There is sufficent evidence to suggest there is a difference in results obtained from the three different apporaches.

(b) Analyze the residuals from this experiment and comment on model adequacy.

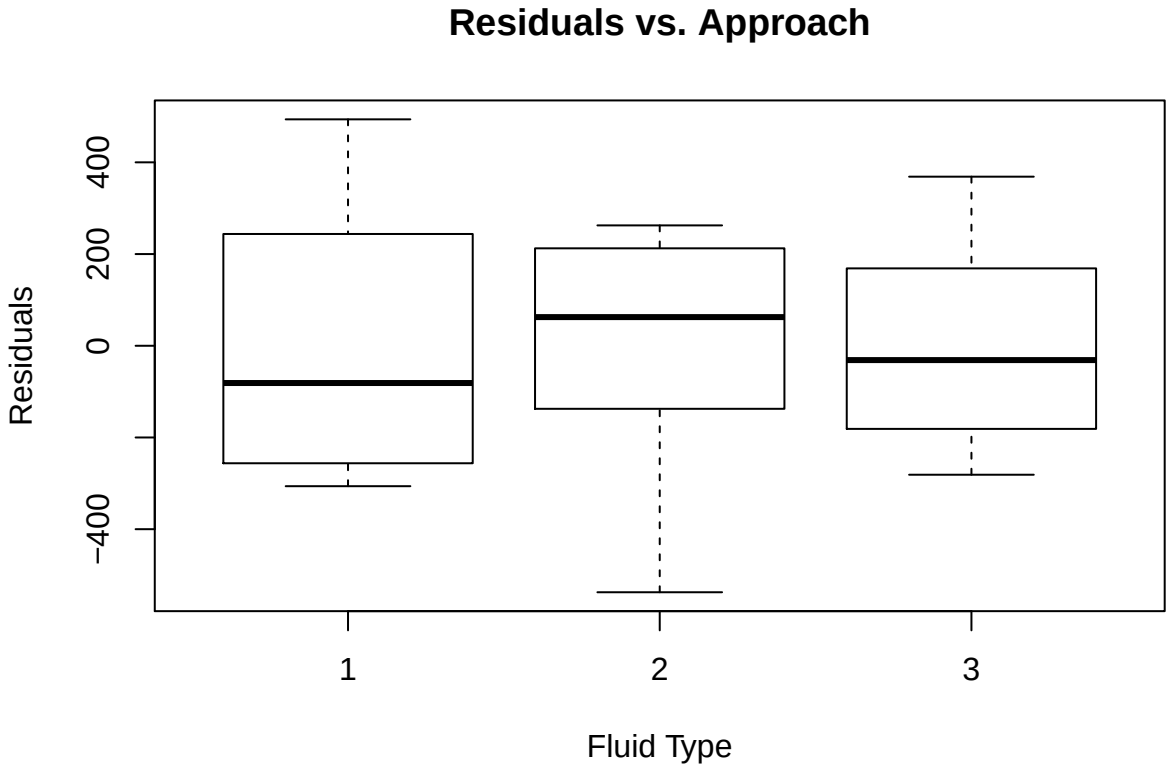
```
qqnorm(a$residuals, main = "Normal Q-Q Plot of Residuals")
qqline(a$residuals, col= "blue")
```




```
plot(a$residuals~a$fitted.values, main = "Residuals vs. Fitted Values", xlab = " Fitted Values", ylab =
```




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



Considering that the Normal Q-Q plot gave a good approximation along the diagonal, we can assume that the normality assumption is satisfied. Although it was safe to note that there is no type of given pattern for the residuals vs. fit plot it is safe to assume that the equal variance assumption is satisied. And lastly there is a consideration of the data being collected randomly during a randomized experiment, it is safe to asssume that the independence assumption is completely satisfied.