

# Stat 207 HW4

Cheng Luo 912466499  
Fan Wu 912538518

February 2, 2015

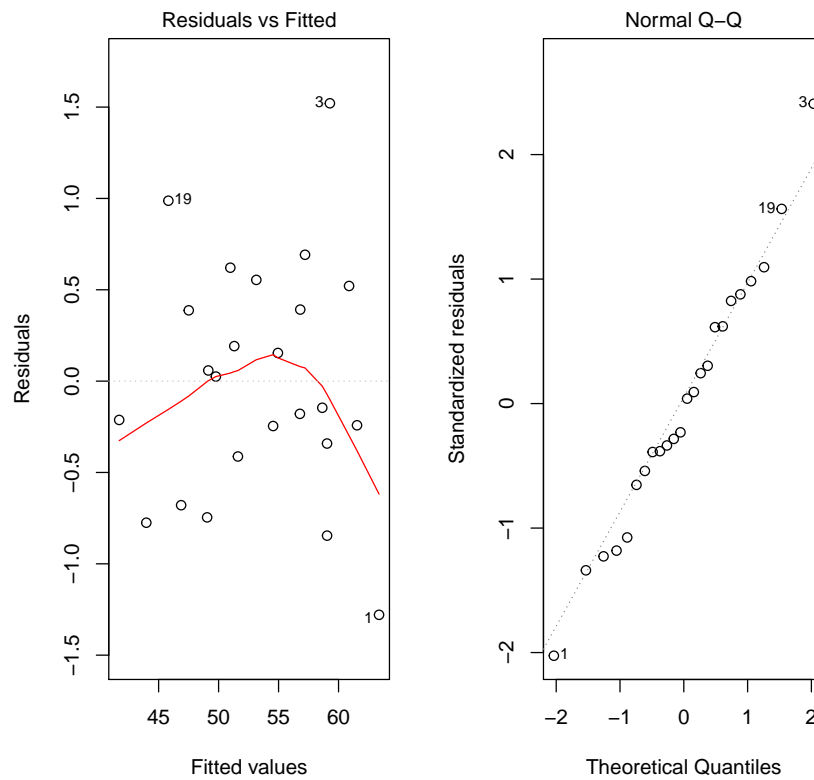


## 1 27.6

```
(a) dat = read.table("CH27PR06.txt")
names(dat) = c("Y", "S", "A")
dat1 = dat
dat$A = factor(dat$A)
dat$S = factor(dat$S)
r = length(unique(dat$S))
a = length(unique(dat$A))
model = aov(Y~ A + S, data = dat)
res = resid(model)
res
```

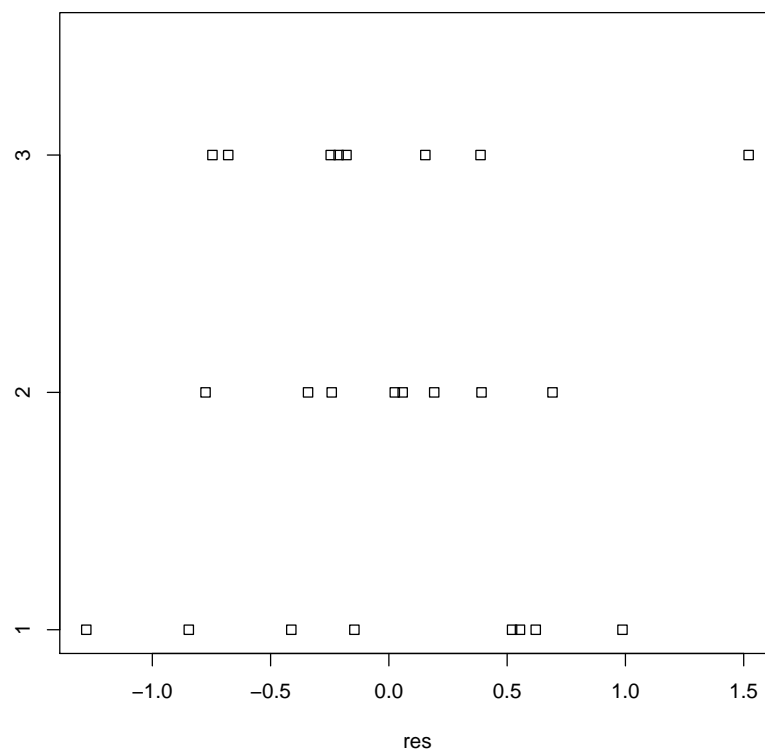
##	1	2	3	4	5	6
##	-1.27916667	-0.24166667	1.52083333	-0.84583333	0.69166667	0.15416667
##	7	8	9	10	11	12
##	0.62083333	0.05833333	-0.67916667	0.55416667	0.19166667	-0.74583333
##	13	14	15	16	17	18
##	0.52083333	-0.34166667	-0.17916667	-0.14583333	0.39166667	-0.24583333
##	19	20	21	22	23	24
##	0.98750000	-0.77500000	-0.21250000	-0.41250000	0.02500000	0.38750000

```
par(mfrow = c(1,2))
plot(model, which = 1)
plot(model, which = 2)
```



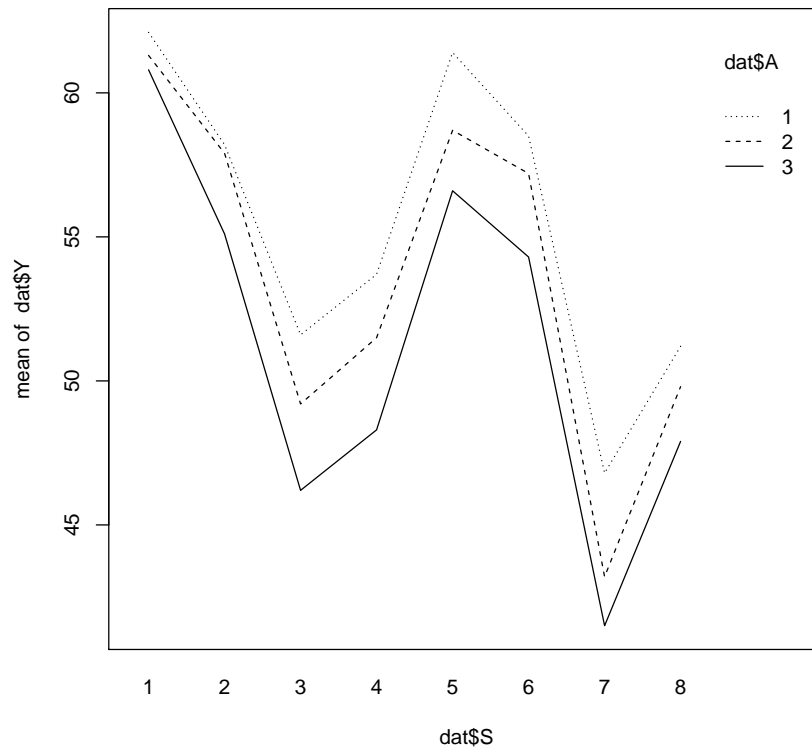
The residuals versus fitted values plots shows no sign for unequal variance. And the QQ-plot indicates approximately normal distribution, so that normality assumption seems to be reasonable, we can use repeated measures model here.

(b) `stripchart(res ~ dat$A, method = 'stack')`



These plots do not indicate any correlations of the error terms within a price level, and thus suggest that no interference effects are present.

(c) `interaction.plot(dat$S, dat$A, dat$Y)`



We see that the rating curves for the stores do not appear to exhibit substantial departures from being parallel, hence, the assumption of no interactions appear to be reasonable.

```
(d)  ab = dat1$S*dat1$A
      model1 = aov(Y~factor(A)+factor(S)+ab, data = dat1)
      anova(model1)

## Analysis of Variance Table
##
## Response: Y
##          Df Sum Sq Mean Sq  F value    Pr(>F)
## factor(A)  2   67.48   33.740   52.9429 5.653e-07 ***
## factor(S)  7  745.19  106.455  167.0410 1.056e-11 ***
## ab         1    1.29    1.288    2.0204   0.1787
## Residuals 13    8.28    0.637
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```

qf(1-0.01, 1, 13)

## [1] 9.073806

pf(2.0204, 1, 13, lower.tail = FALSE)

## [1] 0.1787417

```

$H_0: D = 0$   
 VS.  $H_1: D \neq 0$   
 $F^* = \frac{SSAB^*/1}{SSrem^*/(13)} = 1.288/0.637 = 2.0204$   
 we can reject  $H_0$  if  $F^* > F(1 - 0.01; 1, 13) = 9.073806$ , otherwise reject  $H_1$   
 so that reject  $H_1$  because  $F^* < 9.073806$ ,  
 therefore, our conclusion implies that D equals zero, so there's no  
 interactions, and P-value is 0.1787417.

## 2 27.7

(a)

```

model = aov(Y ~ A + Error(S/A), data = dat)
summary(model)

##
## Error: S
##           Df Sum Sq Mean Sq F value Pr(>F)
## Residuals  7  745.2   106.5
##
## Error: S:A
##           Df Sum Sq Mean Sq F value    Pr(>F)
## A           2   67.48   33.74   49.35 4.57e-07 ***
## Residuals 14    9.57    0.68
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

(b)

```

qf(1-0.05, 2, 14)

## [1] 3.738892

pf(49.35, 2, 14, lower.tail = FALSE)

## [1] 4.564874e-07

```

$H_0$ : all  $\tau_j$  equal zero ( $j=1,2,3$ )  
 VS.  $H_1$ : not all  $\tau_j$  equal zero  
 $F^* = \frac{MSA}{MSE} = 33.74/0.68 = 49.35$   
 we can reject  $H_0$  if  $F^* > F(1 - 0.05; 2, 14) = 3.738892$ , otherwise reject  $H_1$   
 so that reject  $H_0$  because  $F^* > 3.738892$ ,  
 therefore, the mean sales of grapefruits differ for three price levels, and the  
 P-value is 4.57e-07

```

(c) means = with(dat, by(Y, A, mean))
      D1 = means[1] - means[2]
      D2 = means[1] - means[3]
      D3 = means[2] - means[3]
      tukey = 1/sqrt(2)*qtukey(0.95, a, (a-1)*(r-1))
      tukey

## [1] 2.61728

      mstr.s = 0.68
      s = sqrt(2*mstr.s/(r))
      s

## [1] 0.4123106

      c(D1-s*tukey, D1+s*tukey)

##           1           1
## 0.7583676 2.9166324

      c(D2-s*tukey, D2+s*tukey)

##           1           1
## 3.020868 5.179132

      c(D3-s*tukey, D3+s*tukey)

##           2           2
## 1.183368 3.341632

```



$$\bar{Y}_{1..} = 55.4375, \bar{Y}_{2..} = 53.6, \bar{Y}_{3..} = 51.3375$$

$$\hat{D}_1 = \bar{Y}_{1..} - \bar{Y}_{2..} = 1.8375, \hat{D}_2 = \bar{Y}_{1..} - \bar{Y}_{3..} = 4.1, \hat{D}_3 = \bar{Y}_{2..} - \bar{Y}_{3..} = 2.2625$$

Because we estimate all pairwise comparisons, so we use tukey procedure

$$S = \sqrt{\frac{MSTR.S}{r}} * 2 = 0.4123106, Tukey = \frac{1}{\sqrt{2}} qt_{\alpha/2}(1 - \alpha, a, (a - 1) * (r - 1)) = 2.61728$$

base on  $\hat{D}_i \pm S * Tukey$

$$0.7583676 \leq D_1 \leq 2.9166324$$

$$3.020868 \leq D_2 \leq 5.179132$$

$$1.183368 \leq D_3 \leq 3.341632$$

(d)

$$\begin{aligned} \hat{E} &= \frac{S_r^2}{MSTR.S} \\ &= \frac{(s-1)MSS + s(r-1) * MSTR.S}{(sr-1) * MSTR.S} \\ &= 48.36189 \end{aligned}$$

### 3 27.18

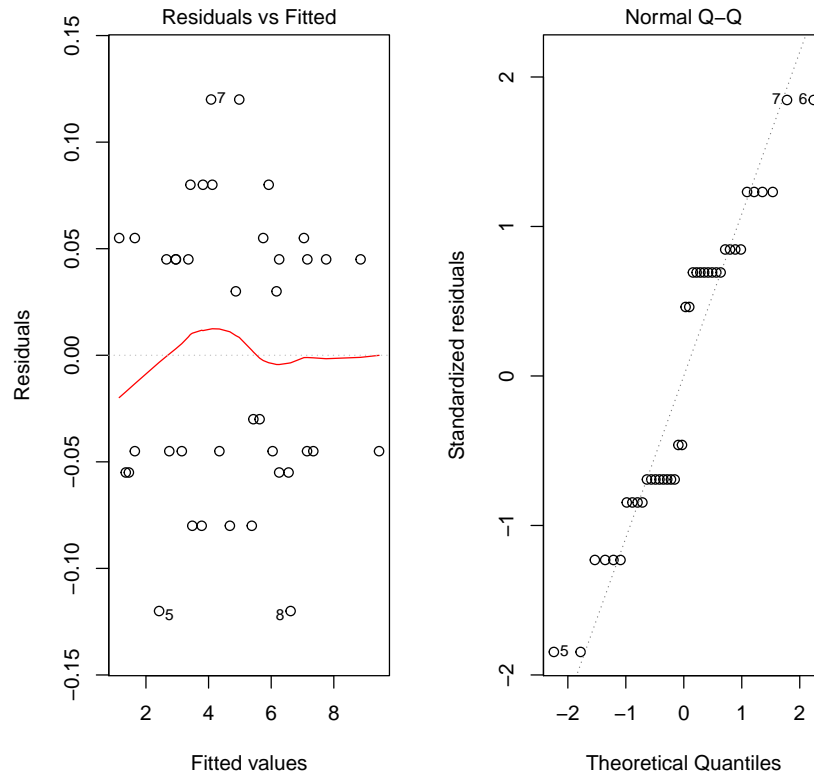
```
(a) dat = read.table("CH27PR18.txt")
names(dat) = c("Y", "S", "A", "B")
dat$S = factor(dat$S)
dat$A = factor(dat$A)
dat$B = factor(dat$B)
s = length(unique(dat$S))
a = length(unique(dat$A))
b = length(unique(dat$B))
model = aov(Y ~ S+A*B+A*S+B*S, data = dat)
resid(model)
```

##	1	2	3	4	5	6	7	8	9	10
##	-0.045	0.045	0.045	-0.045	-0.120	0.120	0.120	-0.120	0.080	-0.080
##	11	12	13	14	15	16	17	18	19	20
##	-0.080	0.080	-0.045	0.045	0.045	-0.045	0.080	-0.080	-0.080	0.080
##	21	22	23	24	25	26	27	28	29	30
##	0.055	-0.055	-0.055	0.055	0.030	-0.030	-0.030	0.030	-0.045	0.045
##	31	32	33	34	35	36	37	38	39	40
##	0.045	-0.045	0.055	-0.055	-0.055	0.055	-0.045	0.045	0.045	-0.045

```

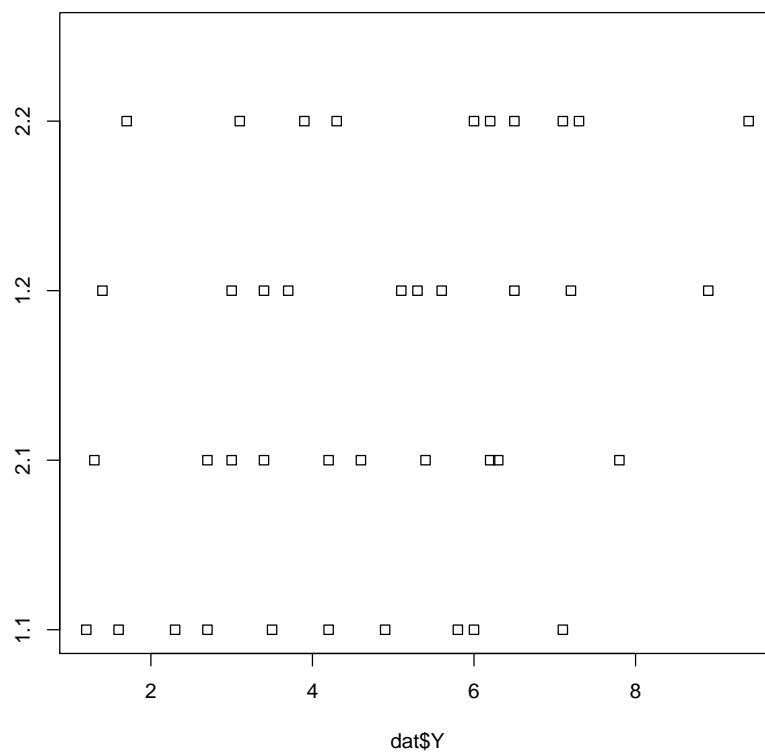
par(mfrow = c(1,2))
plot(model, which = 1)
plot(model, which = 2)

```



The residuals versus fitted values plots shows no sign for unequal variance. And the QQ-plot indicates approximately normal distribution with slightly light tail, so that normality assumption seems to be reasonable, we can use model here.

(b) `stripchart(dat$Y ~ dat$A*dat$B, method="stack")`



These plots do not indicate any correlations of the error terms within AB, and thus suggest that no interference effects are present.

## 4 27.19

```
(a) model = aov(Y~ A*B+A*S+B*S+ Error(S/(A*B)), data = dat)
      summary(model)

##
## Error: S
##   Df Sum Sq Mean Sq
## S   9  154.6    17.18
##
## Error: S:A
##   Df Sum Sq Mean Sq
## A   1   3.025    3.0250
```

```
## A:S  9  2.035  0.2261
##
## Error: S:B
##      Df Sum Sq Mean Sq
## B      1 11.449   11.449
## B:S  9  5.061    0.562
##
## Error: S:A:B
##      Df Sum Sq Mean Sq F value Pr(>F)
## A:B      1  0.001  0.00100   0.053  0.823
## Residuals  9  0.169  0.01878
```

- (b) From the plot, we know that treatment interaction effects are not present, but two treatment main effects are present.

(c) `qf(1-0.005, 1, 9)`

```
## [1] 13.61361

pf(0.053, 1, 9, lower.tail = FALSE)

## [1] 0.8230705
```

$H_0$ : all  $(\alpha\beta)_{jk}$  equal zero

VS.  $H_1$ : not all  $(\alpha\beta)_{jk}$  equal zero

$$F^* = \frac{MSAB}{MSABS} = 0.001/0.01878 = 0.053$$

we can reject  $H_0$  if  $F^* > F(1 - 0.005; 1, 9) = 13.61361$ , otherwise reject  $H_1$

so that reject  $H_1$  because  $F^* < 13.61361$ ,

therefore, there's no two treatment interaction effect, and the P-value is 0.823

(d) `qf(1-0.05, 1, 9)`

```
## [1] 5.117355

pf(13.37904, 1, 9, lower.tail = FALSE)

## [1] 0.005253961
```

$H_0$ : all  $\alpha_i$  equal zero( $i=1,2$ )

VS.  $H_1$ :not all  $\alpha_i$  equal zero

$$F^* = \frac{MSA}{MSAS} = 3.0250/0.2261 = 13.37904$$

we can reject  $H_0$  if  $F^* > F(1 - 0.05; 1, 9) = 5.117355$ , otherwise reject  $H_1$

so that reject  $H_0$  because  $F^* > 5.117355$ ,

therefore, factor A main effect is present, and the P-value is 0.005253961

```
qf(1-0.05, 1, 9)

## [1] 5.117355

pf(20.37189, 1, 9, lower.tail = FALSE)

## [1] 0.001460304
```

$H_0$ : all  $\beta_j$  equal zero( $j=1,2$ )

VS.  $H_1$ :not all  $\beta_j$  equal zero

$$F^* = \frac{MSB}{MSBS} = 11.449/0.562 = 20.37189$$

we can reject  $H_0$  if  $F^* > F(1 - 0.05; 1, 9) = 5.117355$ , otherwise reject  $H_1$

so that reject  $H_0$  because  $F^* > 5.117355$ ,

therefore, factor B main effect is present, and the P-value is 0.001460304

```
(e) means = with(dat, by(Y, list(A,B) , mean))
L1 = means[2,1] - means[1,1]
L2 = means[1,2] - means[1,1]
L3 = means[2,1] - means[1,2]
L4 = means[2,2] - means[1,1]
B = qt(1-0.05/(2*4), 9)
B

## [1] 3.110935

msabs = 0.01878
S = sqrt(2*msabs/(s))
S

## [1] 0.06128621

c(L1-S*B, L1+S*B)
```

```
## [1] 0.3693426 0.7506574

c(L2-S*B, L2+S*B)

## [1] 0.8893426 1.2706574

c(L3-S*B, L3+S*B)

## [1] -0.7106574 -0.3293426

c(L4-S*B, L4+S*B)

## [1] 1.429343 1.810657
```

$$\begin{aligned}
\bar{Y}_{.11} &= 3.936, \bar{Y}_{.12} = 5.01, \\
\bar{Y}_{.21} &= 4.49, \bar{Y}_{.22} = 5.55 \\
\hat{L}_1 &= \bar{Y}_{.21} - \bar{Y}_{.11} = .56, \hat{L}_2 = \bar{Y}_{.12} - \bar{Y}_{.21} = 1.08, \\
\hat{L}_3 &= \bar{Y}_{.21} - \bar{Y}_{.12} = -.52, \hat{L}_4 = \bar{Y}_{.22} - \bar{Y}_{.11} = 1.62 \\
S &= \sqrt{\frac{MSABS}{s}} * 2 = 0.06128621, B = t(1 - \alpha/(2 * 4), (a - 1)(b - 1)(s - 1)) = 3.110935 \\
&\text{base on } \hat{L}_i \pm S * B \\
0.3693426 &\leq L_1 \leq 0.7506574 \\
0.8893426 &\leq L_2 \leq 1.2706574 \\
-0.7106574 &\leq L_3 \leq -0.3293426 \\
1.429343 &\leq L_4 \leq 1.810657
\end{aligned}$$

The treatment of high dose of both drugs has the most significant reduction in pain intensity compared to the treatments of only one drug has high dose , and a significant difference exists in the mean effects of two drugs used alone.

## 5 27.20

```
(a) dat = read.table("CH27PR20.txt")
names(dat) = c("Y", "S", "A", "B")
dat$S = factor(dat$S)
dat$A = factor(dat$A)
dat$B = factor(dat$B)
```

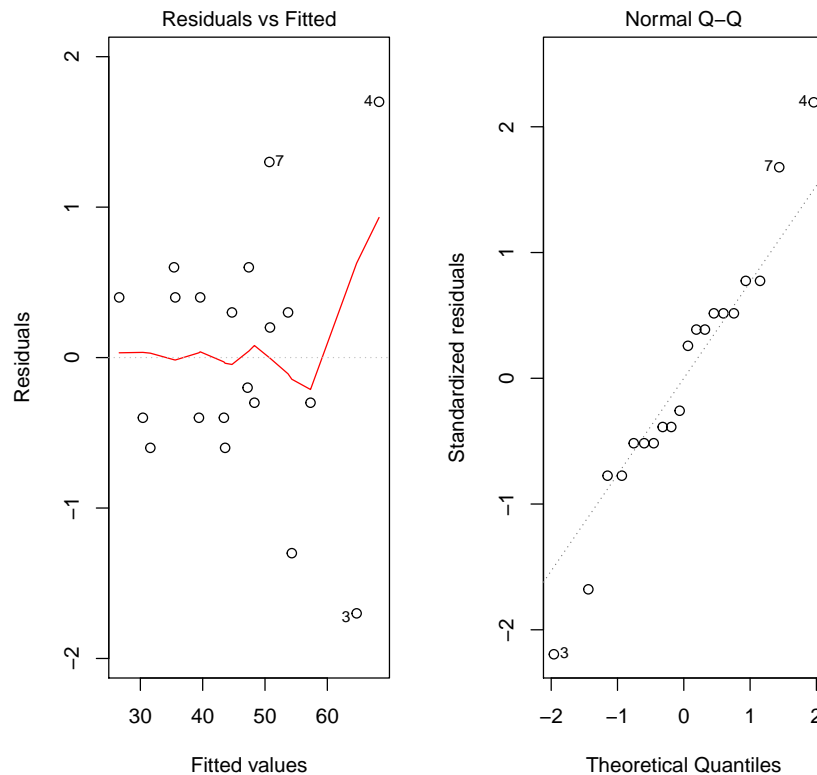
```

s = length(unique(dat$S))
a = length(unique(dat$A))
b = length(unique(dat$B))
model = aov(Y~ A*B+(A/S), data = dat)
resid(model)

##      1      2      3      4      5      6      7      8      9     10     11     12     13     14     15
## -0.6   0.6 -1.7   1.7   0.4 -0.4   1.3 -1.3 -0.6   0.6   0.3 -0.3   0.4 -0.4 -0.2
##      16     17     18     19     20
##   0.2   0.4 -0.4   0.3 -0.3

par(mfrow = c(1,2))
plot(model, which = 1)
plot(model, which = 2)

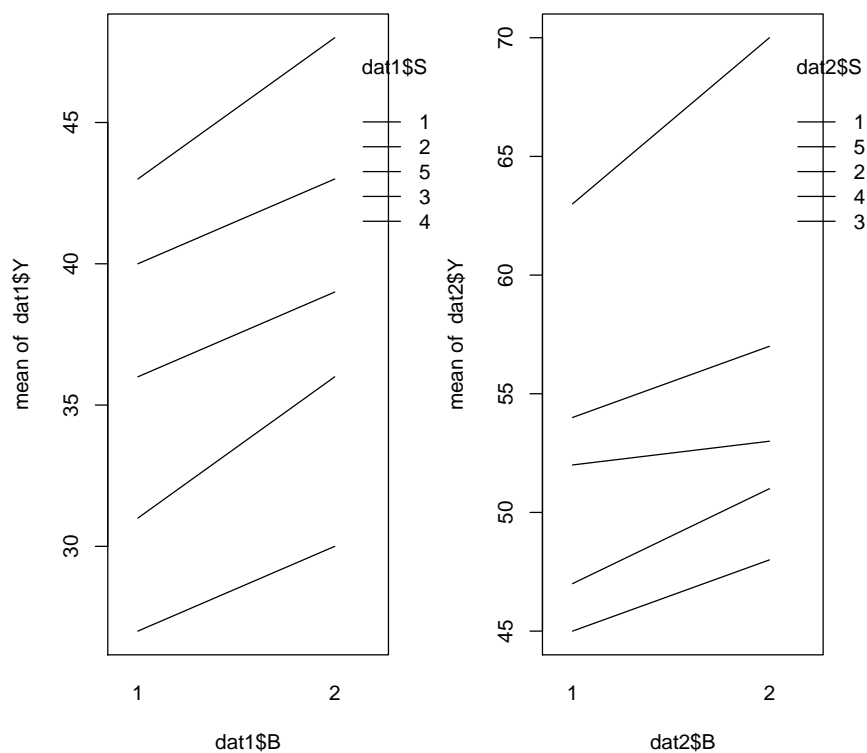
```



The residuals versus fitted values plots shows no sign for unequal variance. And the QQ-plot indicates approximately normal distribution with

slightly heavy tail, so that normality assumption seems to be reasonable, we can use model here.

```
(b) dat1 = dat[ which(dat$A == 1), ]
      dat2 = dat[ which(dat$A == 2), ]
      par( mfrow = c(1, 2))
      interaction.plot(dat1$B, dat1$S, dat1$Y, lty = 1)
      interaction.plot(dat2$B, dat2$S, dat2$Y, lty = 1)
```



There's no evidence any interactions between field and treatments, so split-plot design can be used here.

## 6 27.21

```
(a) model = aov(Y~ A*B+Error(A/S), data = dat)
      summary(model)
```



```
##
## Error: A
##   Df Sum Sq Mean Sq
## A  1   1394    1394
##
## Error: A:S
##           Df Sum Sq Mean Sq F value Pr(>F)
## Residuals  8  837.6    104.7
##
## Error: Within
##           Df Sum Sq Mean Sq F value    Pr(>F)
## B           1  68.45    68.45  45.633 0.000144 ***
## A:B          1   0.05     0.05   0.033 0.859674
## Residuals  8  12.00     1.50
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

- (b) From the plot, we know that treatment interaction effects are not present, but two treatment main effects are present.

(c)

```
qf(1-0.05, 1, 8)

## [1] 5.317655

pf(0.033, 1, 8, lower.tail = FALSE)

## [1] 0.8603687
```

$H_0$ : all  $(\alpha\beta)_{jk}$  equal zero

VS.  $H_1$ : not all  $(\alpha\beta)_{jk}$  equal zero

$$F^* = \frac{MS_{AB}}{MS_{B.W(A)}} = 0.05/1.5 = 0.033$$

we can reject  $H_0$  if  $F^* > F(1 - 0.05; 1, 8) = 5.317655$ , otherwise reject  $H_1$

so that reject  $H_1$  because  $F^* < 5.317655$ ,

therefore, there's no two treatment interaction effect, and the P-value is 0.8603687

(d)

```
qf(1-0.05, 1, 8)

## [1] 5.317655

pf(13.31423, 1, 8, lower.tail = FALSE)

## [1] 0.006505011
```

$H_0$ : all  $\alpha_i$  equal zero( $i=1,2$ )

VS.  $H_1$ :not all  $\alpha_i$  equal zero

$$F^* = \frac{MSA}{MSW(A)} = 1394/104.7 = 13.31423$$

we can reject  $H_0$  if  $F^* > F(1 - 0.05; 1, 8) = 5.317655$ , otherwise reject  $H_1$

so that reject  $H_0$  because  $F^* > 5.317655$ ,

therefore, factor A effect is present, and the P-value is 0.00650501

```
qf(1-0.05, 1, 8)

## [1] 5.317655

pf(45.63333, 1, 8, lower.tail = FALSE)

## [1] 0.0001442759
```

$H_0$ : all  $\beta_j$  equal zero( $j=1,2$ )

VS.  $H_1$ :not all  $\beta_j$  equal zero

$$F^* = \frac{MSB}{MSB.W(A)} = 68.45/1.5 = 45.63333$$

we can reject  $H_0$  if  $F^* > F(1 - 0.05; 1, 8) = 5.317655$ , otherwise reject  $H_1$

so that reject  $H_0$  because  $F^* > 5.317655$ ,

therefore, factor B main effect is present, and the P-value is 0.0001442759

```
(e) means1 = with(dat, by(Y, A , mean))
      means2 = with(dat, by(Y, B , mean))
      L1 = means1[1] - means1[2]
      L2 = means2[1] - means2[2]
      B = qt(1-0.1/(2*2), 8)
      B

## [1] 2.306004

      msb.wa = 1.5
      mswa = 104.7
      S1 = sqrt(2*mswa/(b*s))
      S2 = sqrt(2*msb.wa/(a*s))
      S1

## [1] 4.576024
```

```

S2

## [1] 0.5477226

c(L1-S1*B, L1+S1*B)

##          1          1
## -27.252331  -6.147669

c(L2-S2*B, L2+S2*B)

##          1          1
## -4.96305  -2.43695

```

$$\begin{aligned}
\bar{Y}_{.1} &= 37.3, \bar{Y}_{.2} = 54, \\
\bar{Y}_{.1} &= 43.8, \bar{Y}_{.2} = 47.5 \\
\hat{L}_1 &= \bar{Y}_{.1} - \bar{Y}_{.2} = -16.7, \hat{L}_2 = \bar{Y}_{..1} - \bar{Y}_{..2} = -3.7 \\
B &= t(1 - \alpha/(2 * 2), a(b - 1)(s - 1)) = 2.306004 \\
S1 &= \sqrt{\frac{MSW(A)}{bs}} * 2 = 4.576024, S2 = \sqrt{\frac{MSB.W(A)}{as}} * 2 = 0.5477226 \\
&\text{base on } \hat{L}_i \pm S * B \\
-27.252331 &\leq L_1 \leq -6.147669 \\
-4.96305 &\leq L_2 \leq -2.43695
\end{aligned}$$

Irrigation method 2 is significantly better than Irrigation method 1, and fertilizer 2 is significantly better than fertilizer 1.

## 7 27.22

$$\begin{aligned}
SSTO &= \sum_i \sum_j (Y_{ij} - \bar{Y}_{..})^2 \\
&= \sum_i \sum_j (Y_{ij} - \bar{Y}_{i.} + \bar{Y}_{i.} - \bar{Y}_{..})^2 \\
&= \sum_i \sum_j (Y_{ij} - \bar{Y}_{i.})^2 + \sum_i \sum_j (\bar{Y}_{i.} - \bar{Y}_{..})^2 + 2 \sum_i \sum_j (Y_{ij} - \bar{Y}_{i.})(\bar{Y}_{i.} - \bar{Y}_{..}) \\
&= \sum_i \sum_j (Y_{ij} - \bar{Y}_{i.})^2 + \sum_i \sum_j (\bar{Y}_{i.} - \bar{Y}_{..})^2 + 2 \sum_i [(\bar{Y}_{i.} - \bar{Y}_{..}) * \sum_j (Y_{ij} - \bar{Y}_{i.})] \\
&\quad (\text{since } \sum_j (Y_{ij} - \bar{Y}_{i.}) = 0) \\
&= \sum_i \sum_j (Y_{ij} - \bar{Y}_{i.})^2 + r \sum_i (\bar{Y}_{i.} - \bar{Y}_{..})^2 \\
&= SSS + SSW
\end{aligned}$$

## 8 Extra Problem

Base on:

$$\begin{aligned}
Y_{ij} &= \mu_{..} + \rho_i + \tau_j + \epsilon_{ij} \\
\bar{Y}_{i.} &= \mu_{..} + \rho_i + \bar{\epsilon}_{i.} \\
\bar{Y}_{.j} &= \mu_{..} + \bar{\rho}_{.} + \tau_j + \bar{\epsilon}_{.j} \\
\bar{Y}_{..} &= \mu_{..} + \bar{\rho}_{.} + \bar{\epsilon}_{..}
\end{aligned}$$

So that:

$$\begin{aligned}
E(SSS) &= r \sum_i (E(\bar{Y}_{i.} - \bar{Y}_{..})^2) \\
&= r \sum_i (Var(\bar{Y}_{i.} - \bar{Y}_{..}) + [E(\bar{Y}_{i.} - \bar{Y}_{..})]^2) \\
&\quad (\text{since } E(\bar{Y}_{i.} - \bar{Y}_{..}) = 0) \\
&= r \sum_i (Var(\bar{Y}_{i.}) + Var(\bar{Y}_{..}) - 2Cov(\bar{Y}_{i.}, \bar{Y}_{..})) \\
&= r \sum_i ((\sigma_\rho^2 + \frac{\sigma^2}{r}) + (\frac{\sigma_\rho^2}{s} + \frac{\sigma^2}{rs}) - 2 * (Cov(\rho_i, \bar{\rho}_{..}) + Cov(\bar{\epsilon}_{i.}, \bar{\rho}_{..}) + Cov(\rho_i, \bar{\epsilon}_{..}) + Cov(\bar{\epsilon}_{i.}, \bar{\epsilon}_{..}))) \\
&= rs * ((\sigma_\rho^2 + \frac{\sigma^2}{r}) + (\frac{\sigma_\rho^2}{s} + \frac{\sigma^2}{rs}) - 2 * (\frac{\sigma_\rho^2}{s} + 0 + 0 + \frac{\sigma^2}{rs})) \\
&= (s-1)\sigma^2 + r(s-1)\sigma_\rho^2 \\
E(MSS) &= \frac{E(SSS)}{s-1} \\
&= \sigma^2 + r\sigma_\rho^2
\end{aligned}$$

Likewise:

$$\begin{aligned}
E(SSTR) &= s \sum_j (E(\bar{Y}_{.j} - \bar{Y}_{..})^2) \\
&= s \sum_j (Var(\bar{Y}_{.j} - \bar{Y}_{..}) + [E(\bar{Y}_{.j} - \bar{Y}_{..})]^2) \\
&= s \sum_j (Var(\bar{Y}_{.j}) + Var(\bar{Y}_{..}) - 2Cov(\bar{Y}_{.j}, \bar{Y}_{..}) + \tau_j) \\
&= s \sum_j ((\frac{\sigma_\rho^2}{s} + \frac{\sigma^2}{s}) + (\frac{\sigma_\rho^2}{s} + \frac{\sigma^2}{rs}) - 2 * (\frac{\sigma_\rho^2}{s} + 0 + 0 + \frac{\sigma^2}{rs}) + \tau_j) \\
&= (r-1)\sigma^2 + s \sum_j \tau_j^2 \\
E(MSTR) &= \frac{E(SSTR)}{r-1} \\
&= \sigma^2 + s \frac{\sum_j \tau_j^2}{r-1}
\end{aligned}$$

$$\begin{aligned}
E(SSTR.S) &= \sum_i \sum_j (E(Y_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y}_{..})^2) \\
&= \sum_i \sum_j (Var(Y_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y}_{..}) + [E(Y_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y}_{..})]^2) \\
&= \sum_i \sum_j (Var(Y_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y}_{..})) \\
&= (r-1)(s-1)\sigma^2 \\
E(MSTR.S) &= \frac{E(SSTR.S)}{(r-1)(s-1)} \\
&= \sigma^2
\end{aligned}$$