

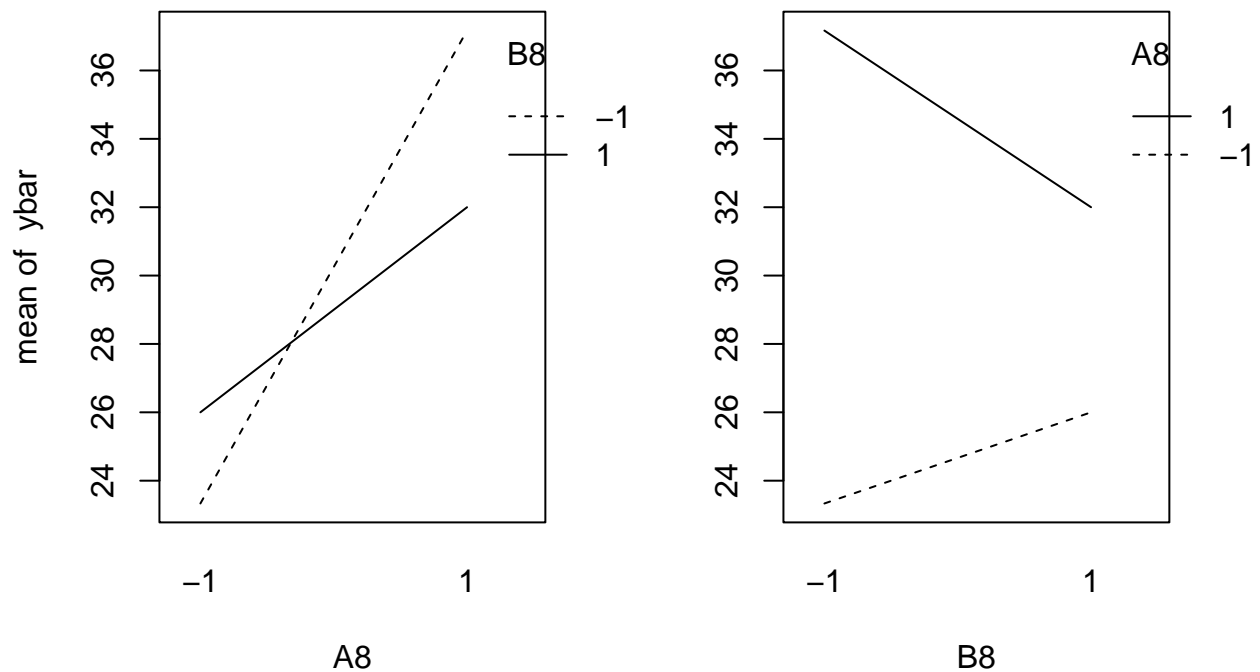
Hwang_Hw4

Charles Hwang

11/12/2021

Problem 6.8

```
rm(list=ls())
A8<-c(-1,-1,1,1)
B8<-c(-1,1,-1,1)
I<-c(21,25,37,31)
II<-c(22,26,39,34)
III<-c(23,24,38,29)
IV<-c(28,25,38,33)
V<-c(20,29,35,30)
VI<-c(26,27,36,35)
ybar<-apply(data.frame(I,II,III,IV,V,VI),1,sum)/6
par(mfrow=c(1,2))
interaction.plot(A8,B8,ybar)
interaction.plot(B8,A8,ybar,ylab="")
```



```
# There appears to be interaction between the two factors.
growth<-c(I,II,III,IV,V,VI)
factorA8<-as.factor(rep(A8,6))
factorB8<-as.factor(rep(B8,6))
g<-lm(growth~factorA8*factorB8)
```

```

anova(g)

## Analysis of Variance Table
##
## Response: growth
##           Df Sum Sq Mean Sq  F value    Pr(>F)
## factorA8      1 590.04   590.04 115.5057 9.291e-10 ***
## factorB8      1   9.38    9.38   1.8352 0.1906172
## factorA8:factorB8 1  92.04   92.04  18.0179 0.0003969 ***
## Residuals    20 102.17    5.11
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

# We reject the null hypotheses at the alpha = 0.05 level for factor A and the
# interaction. Even though we fail to reject the null hypothesis at the
# alpha = 0.05 level for factor B, it is significant because we rejected the null
# hypothesis for the interaction. There is sufficient evidence (p < 0.0001, p = 0.0004)
# that both factors and the interaction between them are significant.
shapiro.test(g$residuals)

##
##  Shapiro-Wilk normality test
##
## data:  g$residuals
## W = 0.96616, p-value = 0.5737
bartlett.test(g$residuals~factorA8)

##
##  Bartlett test of homogeneity of variances
##
## data:  g$residuals by factorA8
## Bartlett's K-squared = 0.62316, df = 1, p-value = 0.4299
bartlett.test(g$residuals~factorB8)

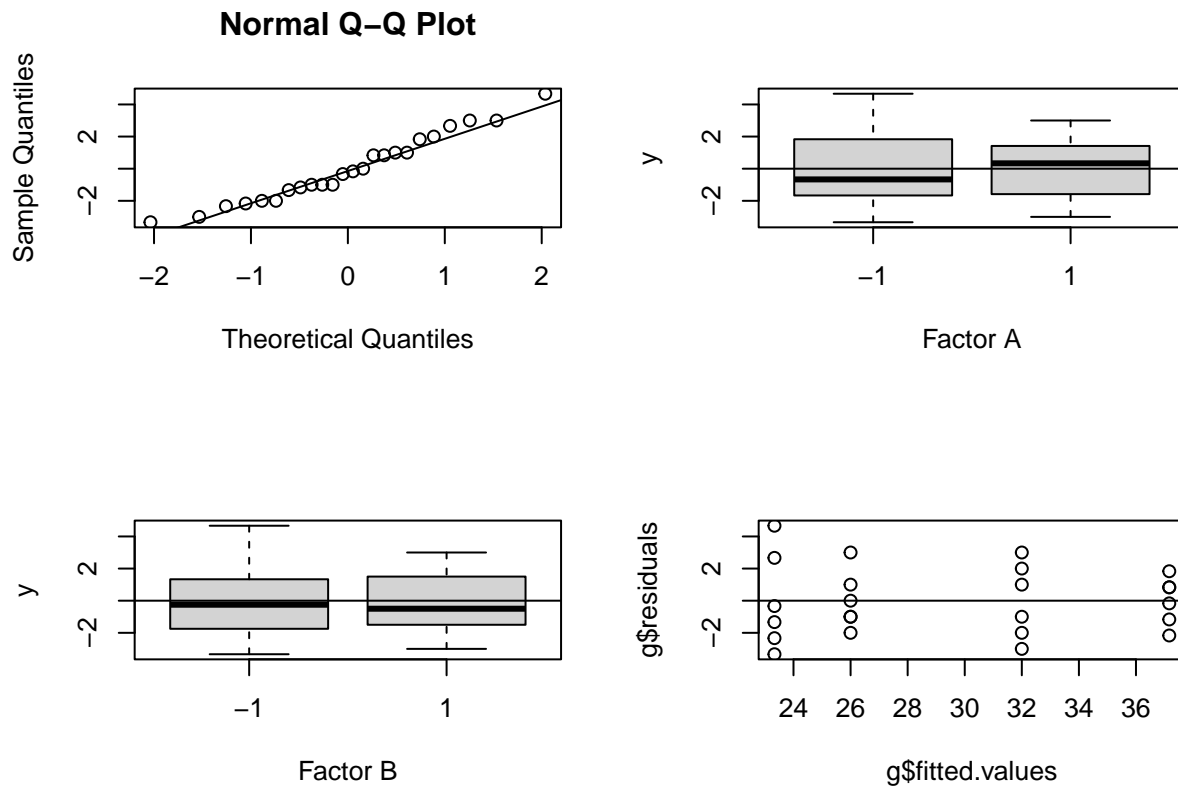
##
##  Bartlett test of homogeneity of variances
##
## data:  g$residuals by factorB8
## Bartlett's K-squared = 0.20427, df = 1, p-value = 0.6513
library(car)
leveneTest(g)

## Levene's Test for Homogeneity of Variance (center = median)
##           Df F value Pr(>F)
## group    3  1.1027 0.3712
##          20

# None of the null hypotheses were rejected, so the normality and
par(mfrow=c(2,2)) # equal variance assumptions appear to be met.
qqnorm(g$residuals)
qqline(g$residuals)
plot(factorA8,g$residuals,xlab="Factor A")
abline(h=0)
plot(factorB8,g$residuals,xlab="Factor B")

```

```
abline(h=0)
plot(g$fitted.values,g$residuals)
abline(h=0)
```



There is a slight variation in the Q-Q plot.

Problem 6.15

(a)

```
I<-c(7.037,14.707,11.635,17.273,10.403,4.368,9.360,13.440,8.561,16.867,13.876,19.824,11.846,6.125,11.19
II<-c(6.376,15.219,12.089,17.815,10.151,4.098,9.253,12.923,8.951,17.052,13.658,19.639,12.337,5.904,10.9
s<-apply(data.frame(I,II),1,sum)
names(s)<-c("(1)","a","b","ab","c","ac","bc","abc","d","ad","bd","abd","cd","acd","bcd","abcd")
eA<-(s["a"]+s["ab"]+s["ac"]+s["abc"]+s["ad"]+s["abd"]+s["acd"]+s["abcd"]-s["(1)"]-s["b"]-s["c"]-s["bc"]
eB<-(s["b"]+s["ab"]+s["bc"]+s["abc"]+s["bd"]+s["abd"]+s["bcd"]+s["abcd"]-s["(1)"]-s["a"]-s["c"]-s["ac"]
eC<-(s["c"]+s["ac"]+s["bc"]+s["abc"]+s["cd"]+s["acd"]+s["bcd"]+s["abcd"]-s["(1)"]-s["a"]-s["b"]-s["ab"]
eD<-(s["d"]+s["ad"]+s["bd"]+s["abd"]+s["cd"]+s["acd"]+s["bcd"]+s["abcd"]-s["(1)"]-s["a"]-s["b"]-s["ab"]
c(eA,eB,eC,eD)
```

```
##          a          b          c          d
##  3.018875  3.975875 -3.596250  1.957750
```

Factors A, B, and C appear to be relatively large.

(b)

```
A15<-as.factor(rep(c(-1,1),8))
B15<-as.factor(rep(c(-1,1),4,each=2))
C15<-as.factor(rep(c(-1,1),2,each=4))
D15<-as.factor(rep(c(-1,1),each=8))
```

```
length<-c(I,II)
factorA<-as.factor(rep(A15,2))
factorB<-as.factor(rep(B15,2))
factorC<-as.factor(rep(C15,2))
factorD<-as.factor(rep(D15,2))
l<-lm(length~factorA*factorB*factorC*factorD)
anova(l)

## Analysis of Variance Table
##
## Response: length
##
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
## factorA	1	72.909	72.909	898.3389	1.740e-15	***
## factorB	1	126.461	126.461	1558.1720	< 2.2e-16	***
## factorC	1	103.464	103.464	1274.8225	< 2.2e-16	***
## factorD	1	30.662	30.662	377.8021	1.485e-12	***
## factorA:factorB	1	29.927	29.927	368.7390	1.790e-12	***
## factorA:factorC	1	128.496	128.496	1583.2562	< 2.2e-16	***
## factorB:factorC	1	0.074	0.074	0.9084	0.3547	
## factorA:factorD	1	0.047	0.047	0.5769	0.4586	
## factorB:factorD	1	0.018	0.018	0.2201	0.6453	
## factorC:factorD	1	0.047	0.047	0.5825	0.4564	
## factorA:factorB:factorC	1	78.751	78.751	970.3255	9.485e-16	***
## factorA:factorB:factorD	1	0.077	0.077	0.9467	0.3450	
## factorA:factorC:factorD	1	0.003	0.003	0.0361	0.8518	
## factorB:factorC:factorD	1	0.010	0.010	0.1251	0.7282	
## factorA:factorB:factorC:factorD	1	0.002	0.002	0.0197	0.8902	
## Residuals	16	1.299	0.081			
## ---						
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1						

We reject the null hypotheses at the $\alpha = 0.05$ level for all four factors and the interactions between factors A and B; factors A and C; and factors A, B, and C. There is sufficient evidence ($p < 0.0001$) that all factors and all combinations of interactions between factors A, B, and C (except the interaction between B and C) are significant.

(c)

Model: $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1x_1 + \hat{\beta}_2x_2 + \hat{\beta}_3x_3 + \hat{\beta}_4x_4 + \hat{\beta}_{12}x_1x_2 + \hat{\beta}_{13}x_1x_3 + \hat{\beta}_{123}x_1x_2x_3$ (Page 247)

\hat{y} : response variable

$\hat{\beta}_0$: intercept term

$\hat{\beta}_1$: intercept term for x_1

x_1 : coded variable for pouring temperature (A)

$\hat{\beta}_2$: intercept term for x_2

x_2 : coded variable for titanium (B)

$\hat{\beta}_3$: intercept term for x_3

x_3 : coded variable for heat treatment method (C)

$\hat{\beta}_4$: intercept term for x_4

x_4 : coded variable for amount of grain refiner used (D)

$\hat{\beta}_{12}$: intercept term for interaction between x_1 and x_2

$\hat{\beta}_{13}$: intercept term for interaction between x_1 and x_3

$\hat{\beta}_{123}$: intercept term for interaction between x_1 , x_2 , and x_3

```
x1<-as.numeric(as.character(factorA))
x2<-as.numeric(as.character(factorB))
x3<-as.numeric(as.character(factorC))
x4<-as.numeric(as.character(factorD))
summary(lm(length~x1+x2+x3+x4+x1*x2*x3))

##
## Call:
## lm(formula = length ~ x1 + x2 + x3 + x4 + x1 * x2 * x3)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.3859 -0.1945  0.0245  0.1798  0.4069
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  11.98806    0.04518  265.369  <2e-16 ***
## x1           1.50944    0.04518   33.413  <2e-16 ***
## x2           1.98794    0.04518   44.005  <2e-16 ***
## x3          -1.79812    0.04518  -39.803  <2e-16 ***
## x4           0.97888    0.04518   21.668  <2e-16 ***
## x1:x2         0.96706    0.04518   21.407  <2e-16 ***
## x1:x3        -2.00388    0.04518  -44.358  <2e-16 ***
## x2:x3         0.04800    0.04518    1.063    0.299
## x1:x2:x3      1.56875    0.04518   34.726  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2555 on 23 degrees of freedom
## Multiple R-squared:  0.9974, Adjusted R-squared:  0.9965
## F-statistic: 1092 on 8 and 23 DF,  p-value: < 2.2e-16
```

(d)

```
shapiro.test(l$residuals)

##
##  Shapiro-Wilk normality test
##
## data:  l$residuals
## W = 0.93025, p-value = 0.0398

bartlett.test(l$residuals~factorA)

##
##  Bartlett test of homogeneity of variances
##
## data:  l$residuals by factorA
## Bartlett's K-squared = 0.048289, df = 1, p-value = 0.8261

bartlett.test(l$residuals~factorB)

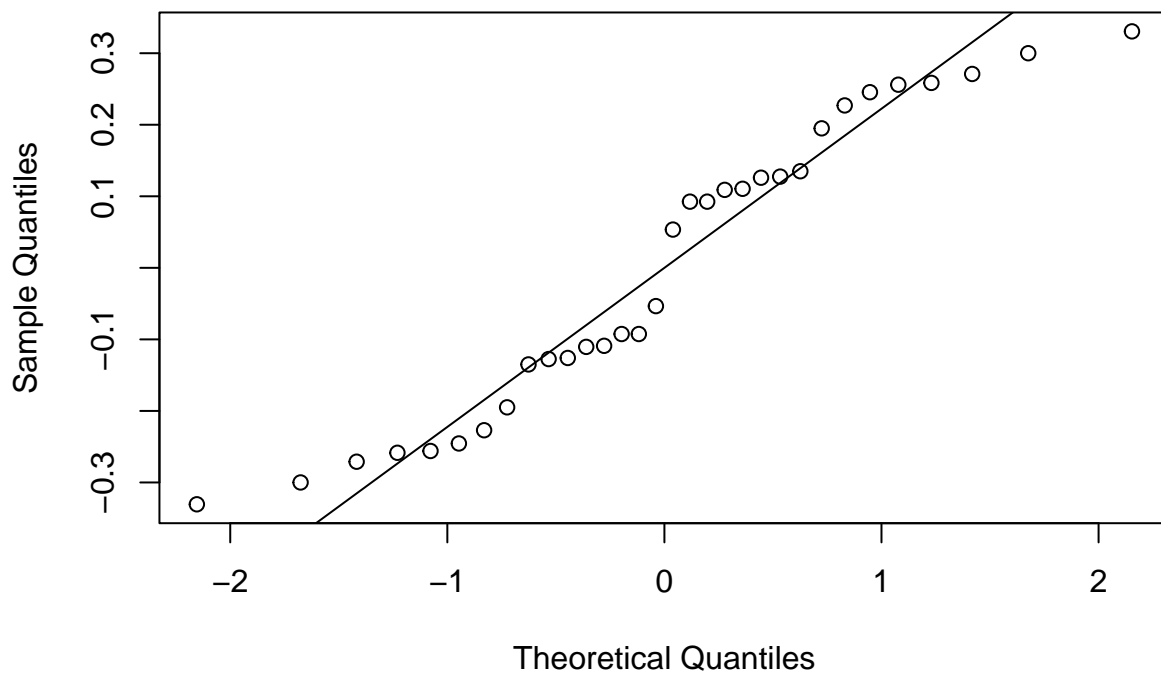
##
```

```
## Bartlett test of homogeneity of variances
##
## data: l$residuals by factorB
## Bartlett's K-squared = 0.0014914, df = 1, p-value = 0.9692
bartlett.test(l$residuals~factorC)

##
## Bartlett test of homogeneity of variances
##
## data: l$residuals by factorC
## Bartlett's K-squared = 0.24639, df = 1, p-value = 0.6196
bartlett.test(l$residuals~factorD)

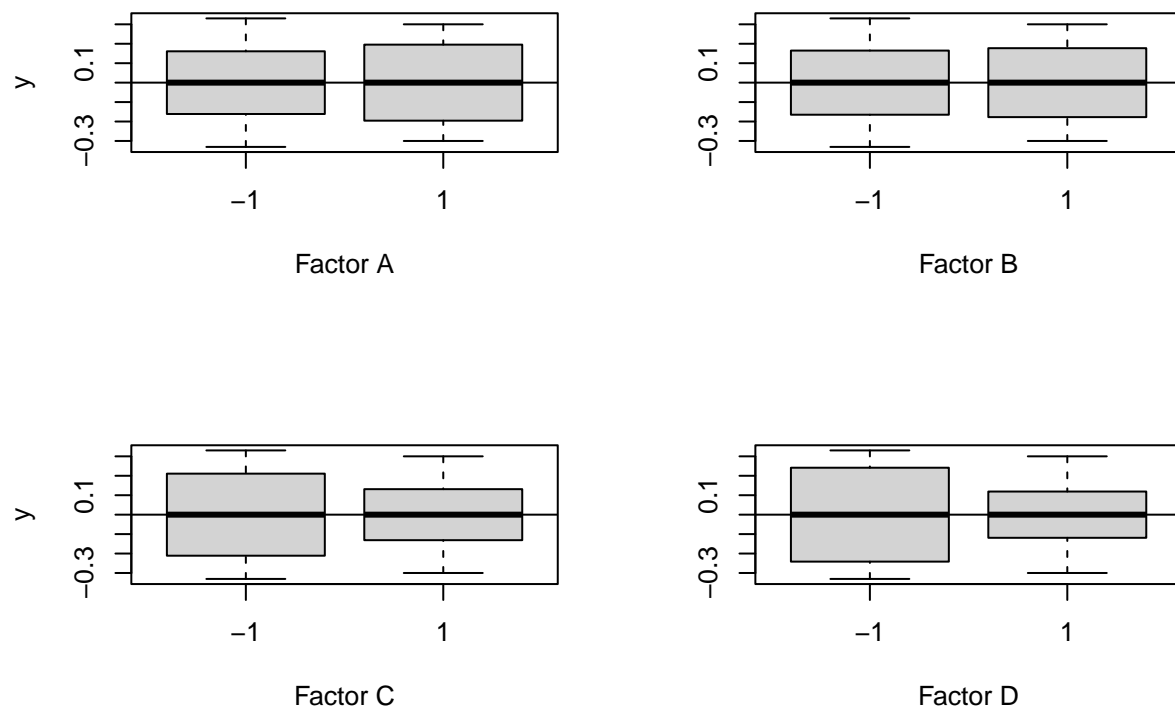
##
## Bartlett test of homogeneity of variances
##
## data: l$residuals by factorD
## Bartlett's K-squared = 0.88342, df = 1, p-value = 0.3473
# We should exercise some caution here as the null hypothesis for the Shapiro test has been
qqnorm(l$residuals) # rejected (p = 0.0398) and the assumption of normality may be violated.
qqline(l$residuals)
```

Normal Q–Q Plot



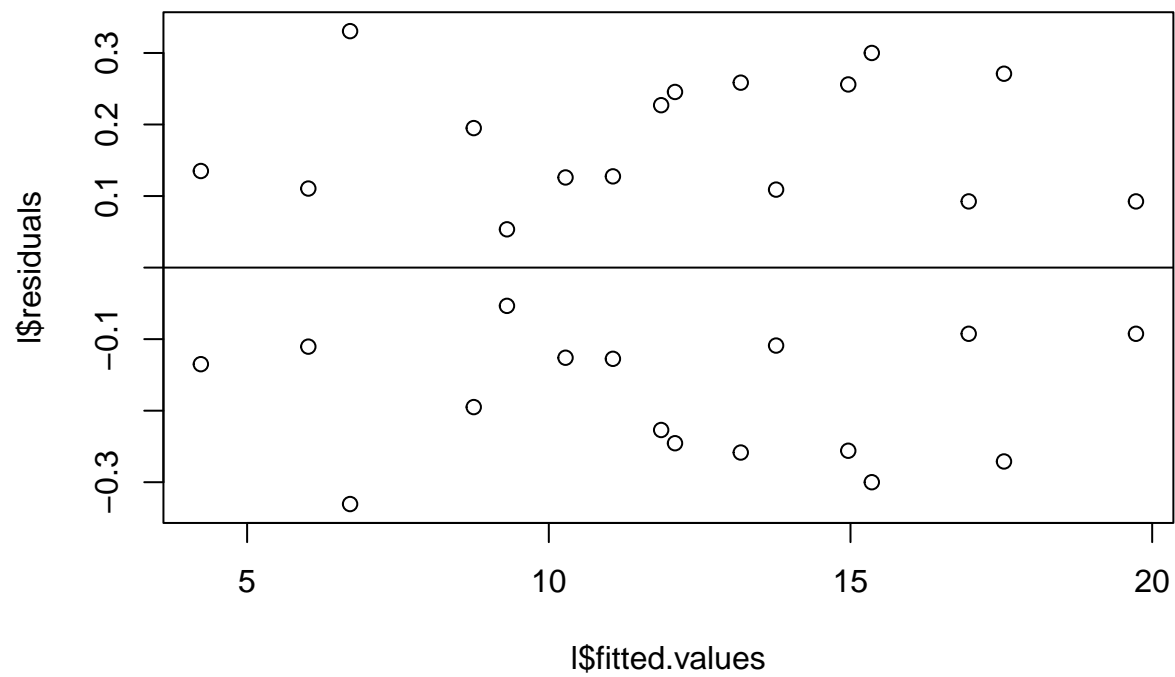
```
# There is a slight variation in the Q-Q plot.
par(mfrow=c(2,2))
plot(factorA,l$residuals,xlab="Factor A")
abline(h=0)
plot(factorB,l$residuals,xlab="Factor B",ylab="")
abline(h=0)
plot(factorC,l$residuals,xlab="Factor C")
```

```
abline(h=0)
plot(factorD,l$residuals,xlab="Factor D",ylab="")
abline(h=0)
```



(e)

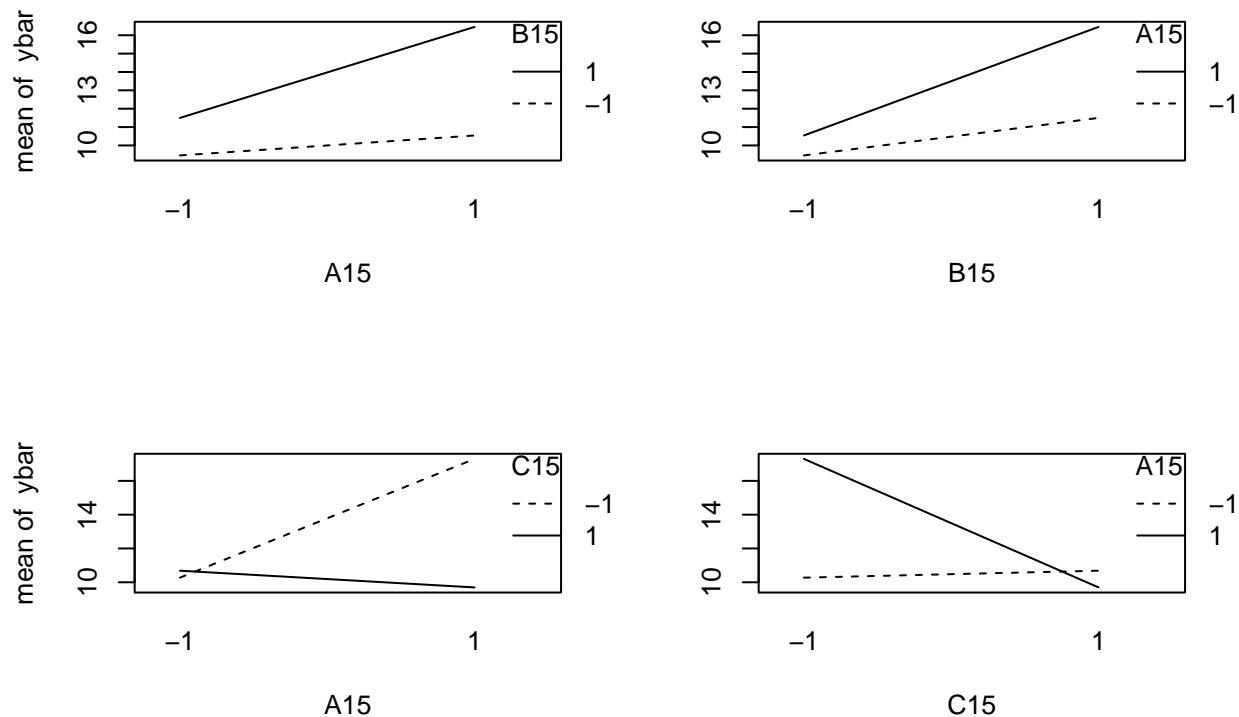
```
plot(l$fitted.values,l$residuals)
abline(h=0)
```



It does not appear that any of the factors affect the variability in cracking.

(f)

```
ybar<-apply(data.frame(I,II),1,sum)/2
par(mfrow=c(2,2))
interaction.plot(A15,B15,ybar)
interaction.plot(B15,A15,ybar,ylab="")
interaction.plot(A15,C15,ybar)
interaction.plot(C15,A15,ybar,ylab="")
```



There appears to be interaction between pouring temperature (A) and titanium content (B) as well as pouring temperature (A) and heat treatment method (C). This could mean these pairs of factors are related or associated with each other.

```
lm(length~x1+x2+x3+x4+x1*x2*x3)$coefficients
```

```
## (Intercept)      x1      x2      x3      x4      x1:x2
## 11.9880625  1.5094375  1.9879375 -1.7981250  0.9788750  0.9670625
##      x1:x3      x2:x3      x1:x2:x3
## -2.0038750  0.0480000  1.5687500
```

I would recommend to use the higher levels for temperature (A) and heat treatment method (C) and the lower methods for titanium content (B) and amount of grain refiner used (D) to decrease crack length. The coefficients calculated from the linear model show the interaction between factors A and C is inversely proportional to the response variable and both factors B and D are directly proportional to the response variable. We can see from Table P6.2 on page 294 that the replicates corresponding to treatment combination “ac” have the lowest values.

Problem 6.22

(a)


```

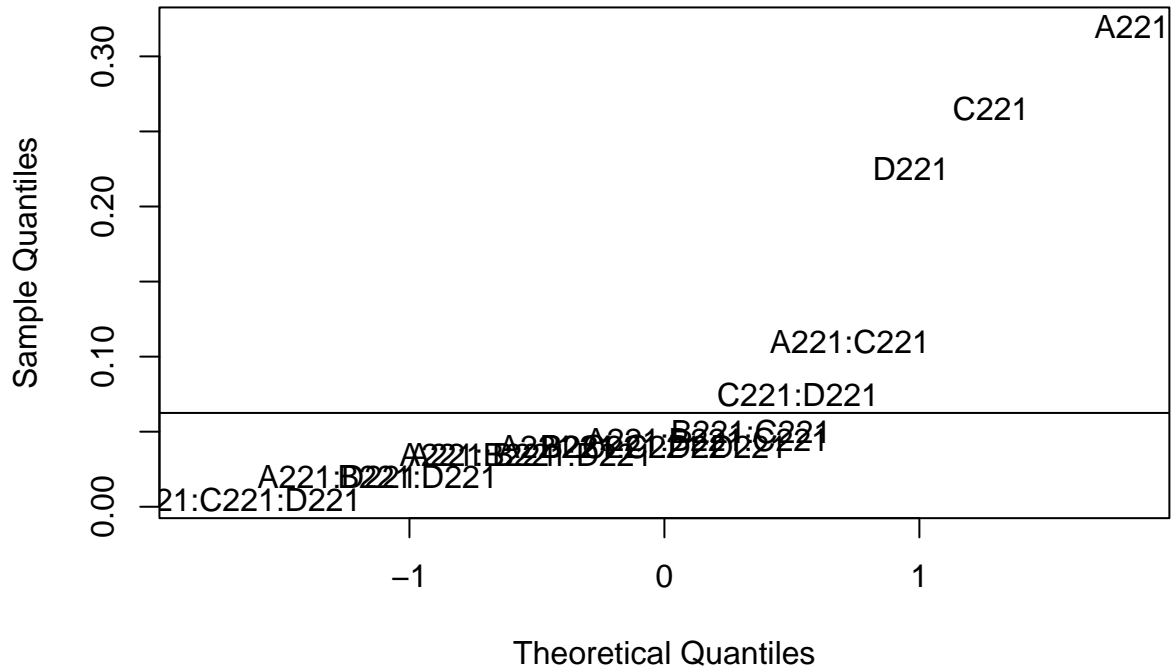
A22<-as.factor(rep(c(-1,1),8))
B22<-as.factor(rep(c(-1,1),4,each=2))
C22<-as.factor(rep(c(-1,1),2,each=4))
D22<-as.factor(rep(c(-1,1),each=8))
U<-c(0.75,0.98,0.72,0.98,0.63,0.67,0.65,0.8,0.6,0.81,0.63,0.79,0.56,0.65,0.55,0.69) # Rearranging data
c<-lm(U~A22*B22*C22*D22)
anova(c)

## Warning in anova.lm(c): ANOVA F-tests on an essentially perfect fit are
## unreliable

## Analysis of Variance Table
##
## Response: U
##
##           Df    Sum Sq  Mean Sq F value Pr(>F)
## A22         1  0.102400  0.102400    NaN    NaN
## B22         1  0.001600  0.001600    NaN    NaN
## C22         1  0.070225  0.070225    NaN    NaN
## D22         1  0.050625  0.050625    NaN    NaN
## A22:B22      1  0.001225  0.001225    NaN    NaN
## A22:C22      1  0.012100  0.012100    NaN    NaN
## B22:C22      1  0.002500  0.002500    NaN    NaN
## A22:D22      1  0.000400  0.000400    NaN    NaN
## B22:D22      1  0.000400  0.000400    NaN    NaN
## C22:D22      1  0.005625  0.005625    NaN    NaN
## A22:B22:C22  1  0.002025  0.002025    NaN    NaN
## A22:B22:D22  1  0.001225  0.001225    NaN    NaN
## A22:C22:D22  1  0.001600  0.001600    NaN    NaN
## B22:C22:D22  1  0.001600  0.001600    NaN    NaN
## A22:B22:C22:D22 1  0.000025  0.000025    NaN    NaN
## Residuals    0  0.000000      NaN
qq<-qqnorm(abs(c$effects[-1]),type="n") # Remove variables
text(qq$x,qq$y,labels=names(abs(c$effects[-1])))
abline(h=0.0625) # Arbitrary cutoff

```

Normal Q-Q Plot



```
names(U)<-c("(1)", "a", "b", "ab", "c", "ac", "bc", "abc", "d", "ad", "bd", "abd", "cd", "acd", "bcd", "abcd")
eA<-(U["a"]+U["ab"]+U["ac"]+U["abc"]+U["ad"]+U["abd"]+U["acd"]+U["abcd"]-U["(1)"]-U["b"]-U["c"]-U["bc"])
eB<-(U["b"]+U["ab"]+U["bc"]+U["abc"]+U["bd"]+U["abd"]+U["bcd"]+U["abcd"]-U["(1)"]-U["a"]-U["c"]-U["ac"])
eC<-(U["c"]+U["ac"]+U["bc"]+U["abc"]+U["cd"]+U["acd"]+U["bcd"]+U["abcd"]-U["(1)"]-U["a"]-U["b"]-U["ab"])
eD<-(U["d"]+U["ad"]+U["bd"]+U["abd"]+U["cd"]+U["acd"]+U["bcd"]+U["abcd"]-U["(1)"]-U["a"]-U["b"]-U["ab"])
c(eA,eB,eC,eD)
```

```
##      a      b      c      d
## 0.1600 0.0200 -0.1325 -0.1125
```

Factors A, C, and D appear to significantly affect unused error correction (UEC).

(b)

```
d<-lm(U~A22*C22*D22) # New model
anova(d)
```

```
## Analysis of Variance Table
```

```
##
```

```
## Response: U
```

```
##      Df    Sum Sq Mean Sq F value    Pr(>F)
## A22      1 0.102400  0.102400  77.2830 2.202e-05 ***
## C22      1 0.070225  0.070225  53.0000 8.552e-05 ***
## D22      1 0.050625  0.050625  38.2075 0.0002648 ***
## A22:C22  1 0.012100  0.012100   9.1321 0.0165108 *
## A22:D22  1 0.000400  0.000400   0.3019 0.5977008
## C22:D22  1 0.005625  0.005625   4.2453 0.0733135 .
## A22:C22:D22 1 0.001600  0.001600   1.2075 0.3037858
## Residuals 8 0.010600  0.001325
```

```
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```

shapiro.test(d$residuals)

##
##  Shapiro-Wilk normality test
##
## data:  d$residuals
## W = 0.90176, p-value = 0.08575
bartlett.test(d$residuals~A22)

##
##  Bartlett test of homogeneity of variances
##
## data:  d$residuals by A22
## Bartlett's K-squared = 6.2043, df = 1, p-value = 0.01274
bartlett.test(d$residuals~C22)

##
##  Bartlett test of homogeneity of variances
##
## data:  d$residuals by C22
## Bartlett's K-squared = 6.4602, df = 1, p-value = 0.01103
bartlett.test(d$residuals~D22)

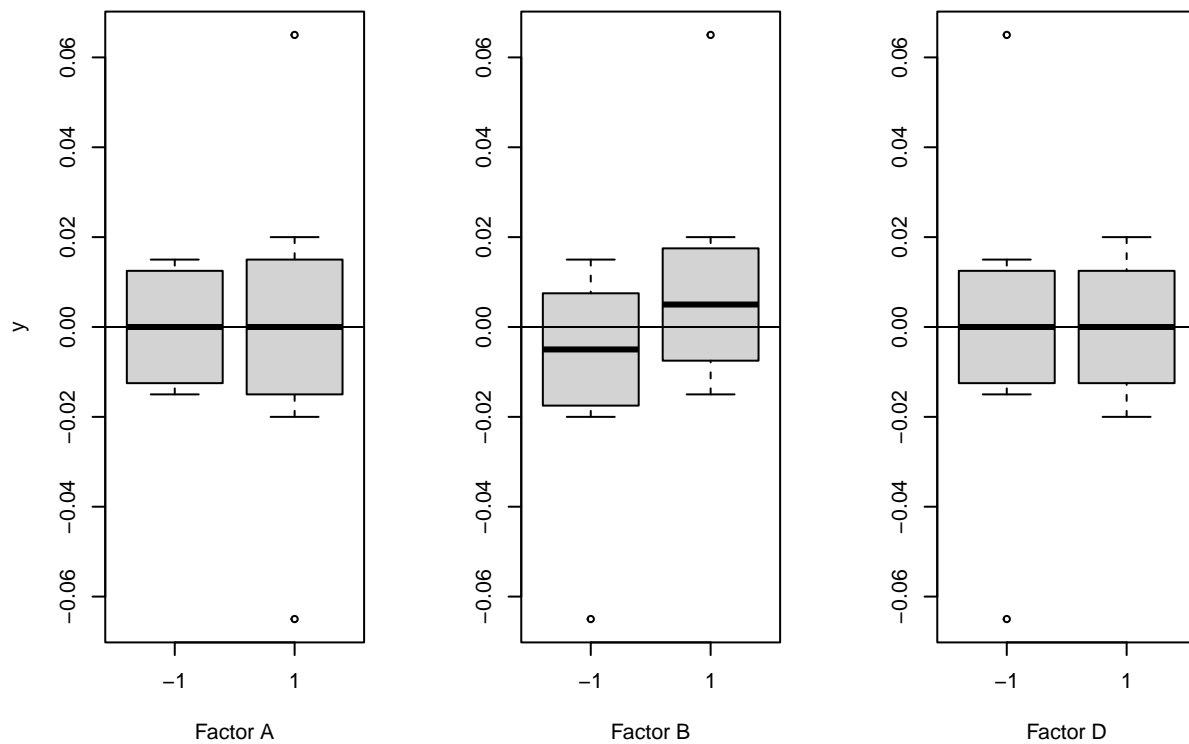
##
##  Bartlett test of homogeneity of variances
##
## data:  d$residuals by D22
## Bartlett's K-squared = 4.7149, df = 1, p-value = 0.0299
leveneTest(d)

## Warning in anova.lm(lm(resp ~ group)): ANOVA F-tests on an essentially perfect
## fit are unreliable

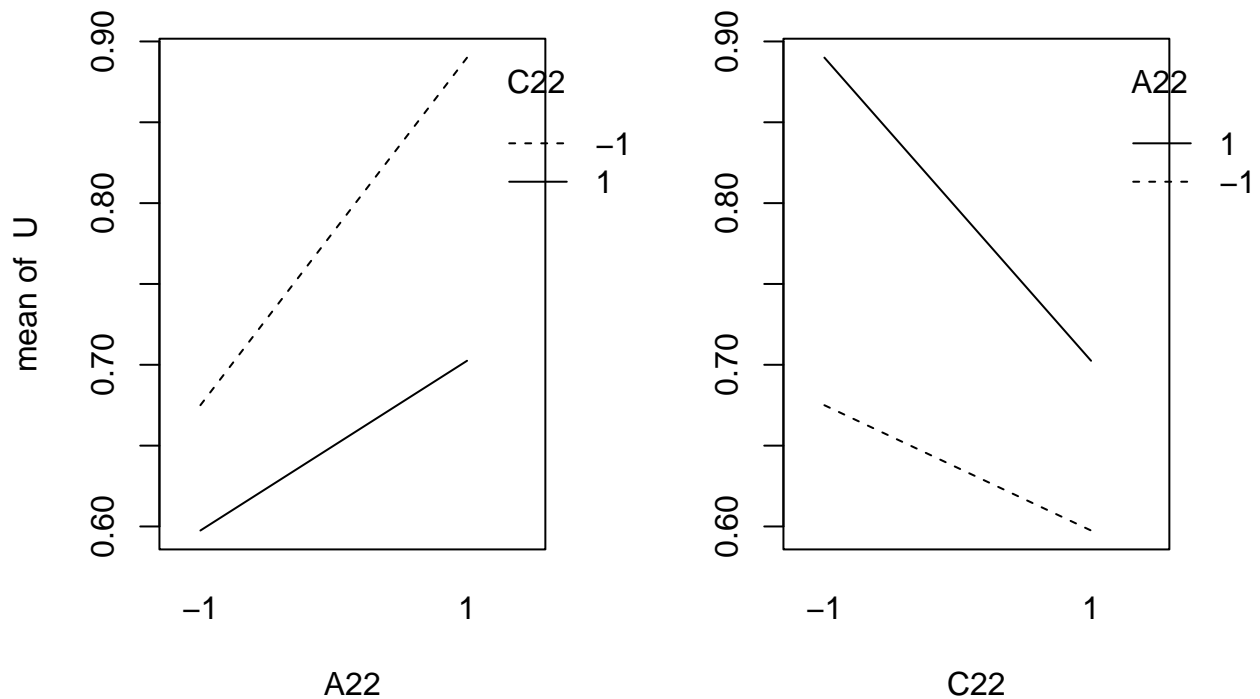
## Levene's Test for Homogeneity of Variance (center = median)
##      Df    F value    Pr(>F)
## group  7 4.8877e+29 < 2.2e-16 ***
##      8
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

# All of the null hypotheses for the Bartlett's tests and Levene's test were rejected.
# There are clear violations of the equal variance assumptions (p = 0.01274,
par(mfrow=c(1,3))      # p = 0.01103, p = 0.0299, p < 0.00001) in this model.
plot(A22,d$residuals,xlab="Factor A")
abline(h=0)
plot(B22,d$residuals,xlab="Factor B",ylab="")
abline(h=0)
plot(D22,d$residuals,xlab="Factor D",ylab="")
abline(h=0)

```



```
par(mfrow=c(1,2))
interaction.plot(A22,C22,U)
interaction.plot(C22,A22,U,ylab="")
```

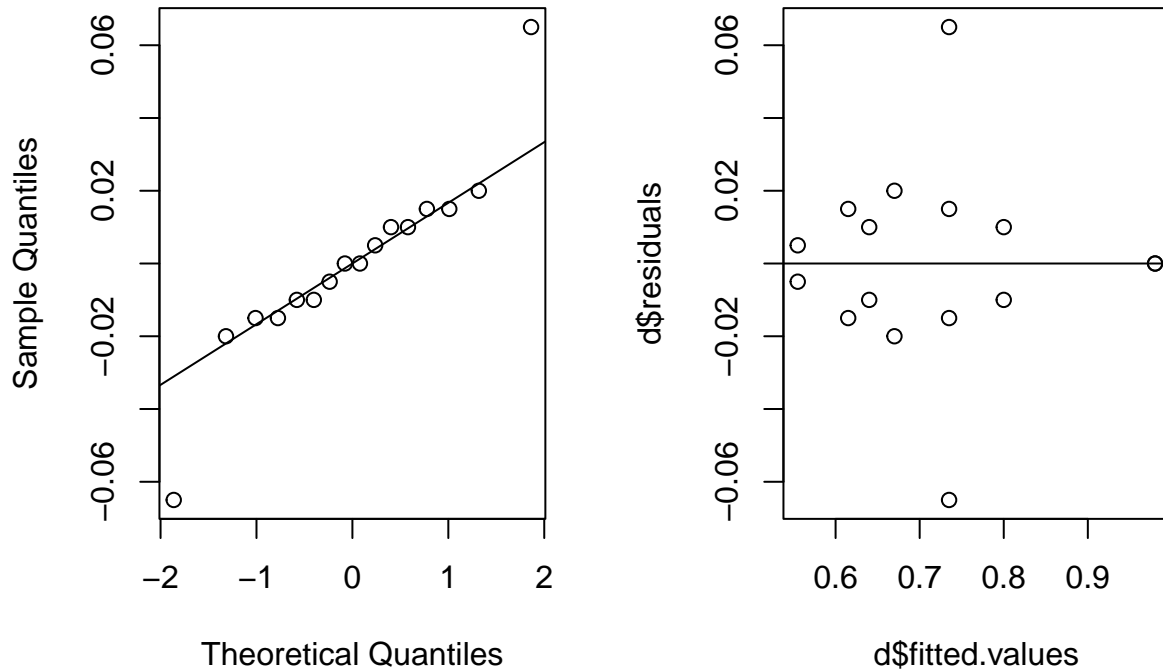


*# There appears to be interaction between laser power (A) and matrix cell size (C). This
could mean the two factors are related or associated with each other.*

```
qqnorm(d$residuals)
qqline(d$residuals)
```

```
plot(d$fitted.values,d$residuals)
abline(h=0)
```

Normal Q-Q Plot



There is a slight variation in the Q-Q plot and there appears to be a slight football effect in the residuals vs. fitted values plot. The range of the response variable is $[0,1]$ and the normality and equal variance assumptions were violated, so this is likely not the best model to use for this data.

Problem 6.26

(a)

```
A26<-as.factor(rep(c(-1,1),16))
B26<-as.factor(rep(c(-1,1),8,each=2))
C26<-as.factor(rep(c(-1,1),4,each=4))
D26<-as.factor(rep(c(-1,1),2,each=8))
E26<-as.factor(rep(c(-1,1),each=16))
y<-c(7,9,34,55,16,20,40,60,8,10,32,50,18,21,44,61,8,12,35,52,15,22,45,65,6,10,30,53,15,20,41,63)
s<-lm(y~A26*B26*C26*D26*E26)
names(y)<-c("(1)", "a", "b", "ab", "c", "ac", "bc", "abc", "d", "ad", "bd", "abd", "cd", "acd", "bcd", "abcd", "e", "ae",
eA<-(y["a"]+y["ab"]+y["ac"]+y["abc"]+y["ad"]+y["abd"]+y["acd"]+y["abcd"]+y["ae"]+y["abe"]+y["ace"]+y["a
eB<-(y["b"]+y["ab"]+y["bc"]+y["abc"]+y["bd"]+y["abd"]+y["bcd"]+y["abcd"]+y["be"]+y["abe"]+y["bce"]+y["a
eC<-(y["c"]+y["ac"]+y["bc"]+y["abc"]+y["cd"]+y["acd"]+y["bcd"]+y["abcd"]+y["ce"]+y["ace"]+y["bce"]+y["a
eD<-(y["d"]+y["ad"]+y["bd"]+y["abd"]+y["cd"]+y["acd"]+y["bcd"]+y["abcd"]+y["de"]+y["ade"]+y["bde"]+y["a
eE<-(y["e"]+y["ae"]+y["be"]+y["abe"]+y["ce"]+y["ace"]+y["bce"]+y["abce"]+y["de"]+y["ade"]+y["bde"]+y["a
c(eA,eB,eC,eD,eE)
```

```
##      a      b      c      d      e
## 11.8125 35.0000  9.6875 -0.8125  0.4375
```

Factors A, B, and C appear to be relatively large.

(b)

```
anova(s)

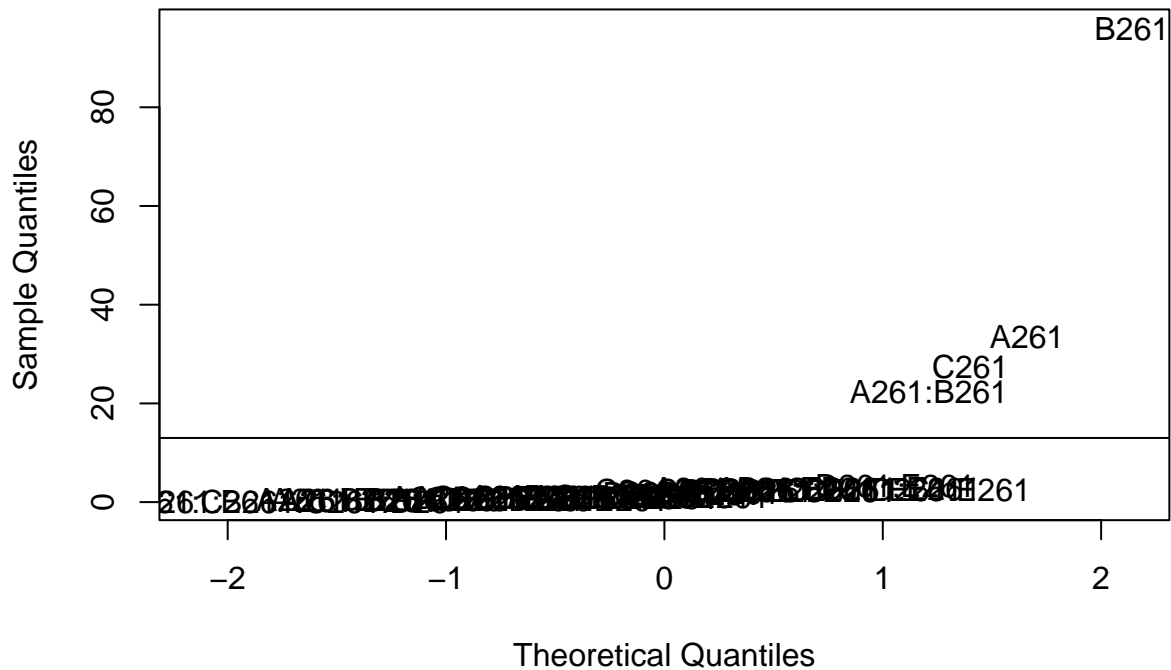
## Warning in anova.lm(s): ANOVA F-tests on an essentially perfect fit are
## unreliable

## Analysis of Variance Table
##
## Response: y
##
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
## A26	1	1116.3	1116.3	NaN	NaN
## B26	1	9214.0	9214.0	NaN	NaN
## C26	1	750.8	750.8	NaN	NaN
## D26	1	5.3	5.3	NaN	NaN
## E26	1	1.5	1.5	NaN	NaN
## A26:B26	1	504.0	504.0	NaN	NaN
## A26:C26	1	1.5	1.5	NaN	NaN
## B26:C26	1	0.0	0.0	NaN	NaN
## A26:D26	1	0.0	0.0	NaN	NaN
## B26:D26	1	3.8	3.8	NaN	NaN
## C26:D26	1	5.3	5.3	NaN	NaN
## A26:E26	1	7.0	7.0	NaN	NaN
## B26:E26	1	2.5	2.5	NaN	NaN
## C26:E26	1	0.8	0.8	NaN	NaN
## D26:E26	1	11.3	11.3	NaN	NaN
## A26:B26:C26	1	1.5	1.5	NaN	NaN
## A26:B26:D26	1	0.8	0.8	NaN	NaN
## A26:C26:D26	1	1.5	1.5	NaN	NaN
## B26:C26:D26	1	1.5	1.5	NaN	NaN
## A26:B26:E26	1	0.3	0.3	NaN	NaN
## A26:C26:E26	1	0.8	0.8	NaN	NaN
## B26:C26:E26	1	7.0	7.0	NaN	NaN
## A26:D26:E26	1	5.3	5.3	NaN	NaN
## B26:D26:E26	1	0.3	0.3	NaN	NaN
## C26:D26:E26	1	5.3	5.3	NaN	NaN
## A26:B26:C26:D26	1	0.0	0.0	NaN	NaN
## A26:B26:C26:E26	1	0.3	0.3	NaN	NaN
## A26:B26:D26:E26	1	7.0	7.0	NaN	NaN
## A26:C26:D26:E26	1	0.8	0.8	NaN	NaN
## B26:C26:D26:E26	1	7.0	7.0	NaN	NaN
## A26:B26:C26:D26:E26	1	0.3	0.3	NaN	NaN
## Residuals	0	0.0	NaN		

```
qq<-qqnorm(abs(s$effects[-1]),type="n") # Remove variables
text(qq$x,qq$y,labels=names(abs(s$effects[-1])))
abline(h=13) # Arbitrary cutoff
```

Normal Q-Q Plot



```
t<-lm(y~A26+B26+C26+A26*B26) # New model
anova(t)
```

```
## Analysis of Variance Table
```

```
##
```

```
## Response: y
```

```
##          Df Sum Sq Mean Sq F value    Pr(>F)
## A26         1 1116.3   1116.3   382.27 < 2.2e-16 ***
## B26         1 9214.0   9214.0  3155.34 < 2.2e-16 ***
## C26         1  750.8    750.8   257.10 2.534e-15 ***
## A26:B26     1  504.0    504.0   172.61 3.038e-13 ***
## Residuals 27    78.8      2.9
```

```
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
shapiro.test(t$residuals)
```

```
##
```

```
## Shapiro-Wilk normality test
```

```
##
```

```
## data:  t$residuals
```

```
## W = 0.96132, p-value = 0.2983
```

```
bartlett.test(t$residuals~A26)
```

```
##
```

```
## Bartlett test of homogeneity of variances
```

```
##
```

```
## data:  t$residuals by A26
```

```
## Bartlett's K-squared = 0.075734, df = 1, p-value = 0.7832
```

```
bartlett.test(t$residuals~B26)
```

```
##
## Bartlett test of homogeneity of variances
##
## data:  t$residuals by B26
## Bartlett's K-squared = 4.3671, df = 1, p-value = 0.03664
```

```
bartlett.test(t$residuals~C26)
```

```
##
## Bartlett test of homogeneity of variances
##
## data:  t$residuals by C26
## Bartlett's K-squared = 0.064506, df = 1, p-value = 0.7995
```

```
# We should exercise some caution here as the null hypothesis of the Bartlett's test for
# factor B has been rejected (p = 0.03664) and the assumption of equal variances for
# factor B may be violated.
```

(c)

Model: $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1x_1 + \hat{\beta}_2x_2 + \hat{\beta}_3x_3 + \hat{\beta}_{12}x_1x_2$ (Page 247)

\hat{y} : response variable

$\hat{\beta}_0$: intercept term

$\hat{\beta}_1$: intercept term for x_1

x_1 : coded variable for aperture setting (A)

$\hat{\beta}_2$: intercept term for x_2

x_2 : coded variable for exposure time (B)

$\hat{\beta}_3$: intercept term for x_3

x_3 : coded variable for development time (C)

$\hat{\beta}_{12}$: intercept term for interaction between x_1 and x_2

```
x1<-as.numeric(as.character(A26))
x2<-as.numeric(as.character(B26))
x3<-as.numeric(as.character(C26))
summary(lm(y~x1+x2+x3+x1*x2))
```

```
##
## Call:
## lm(formula = y ~ x1 + x2 + x3 + x1 * x2)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.7812 -1.2812 -0.3438  1.2500  2.7812
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   30.5312     0.3021  101.07 < 2e-16 ***
## x1             5.9063     0.3021   19.55 < 2e-16 ***
## x2            16.9688     0.3021   56.17 < 2e-16 ***
## x3             4.8438     0.3021   16.03 2.53e-15 ***
```

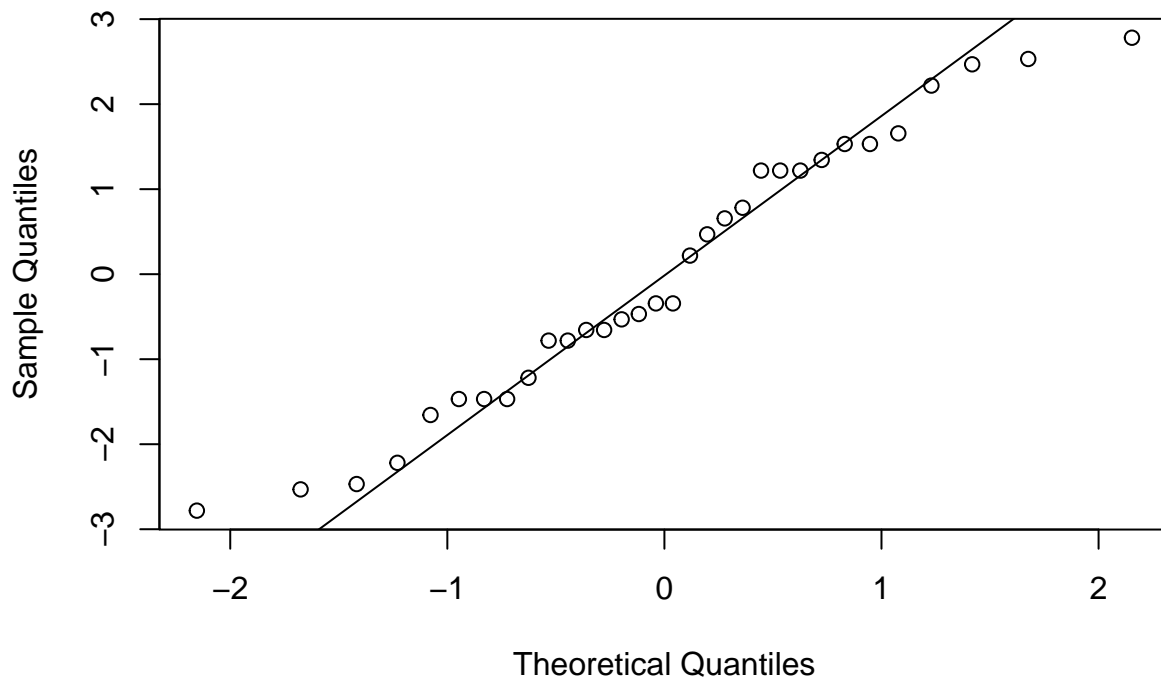


```
## x1:x2          3.9687      0.3021   13.14 3.04e-13 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.709 on 27 degrees of freedom
## Multiple R-squared:  0.9932, Adjusted R-squared:  0.9922
## F-statistic: 991.8 on 4 and 27 DF,  p-value: < 2.2e-16
```

(d)

```
qqnorm(t$residuals)
qqline(t$residuals)
```

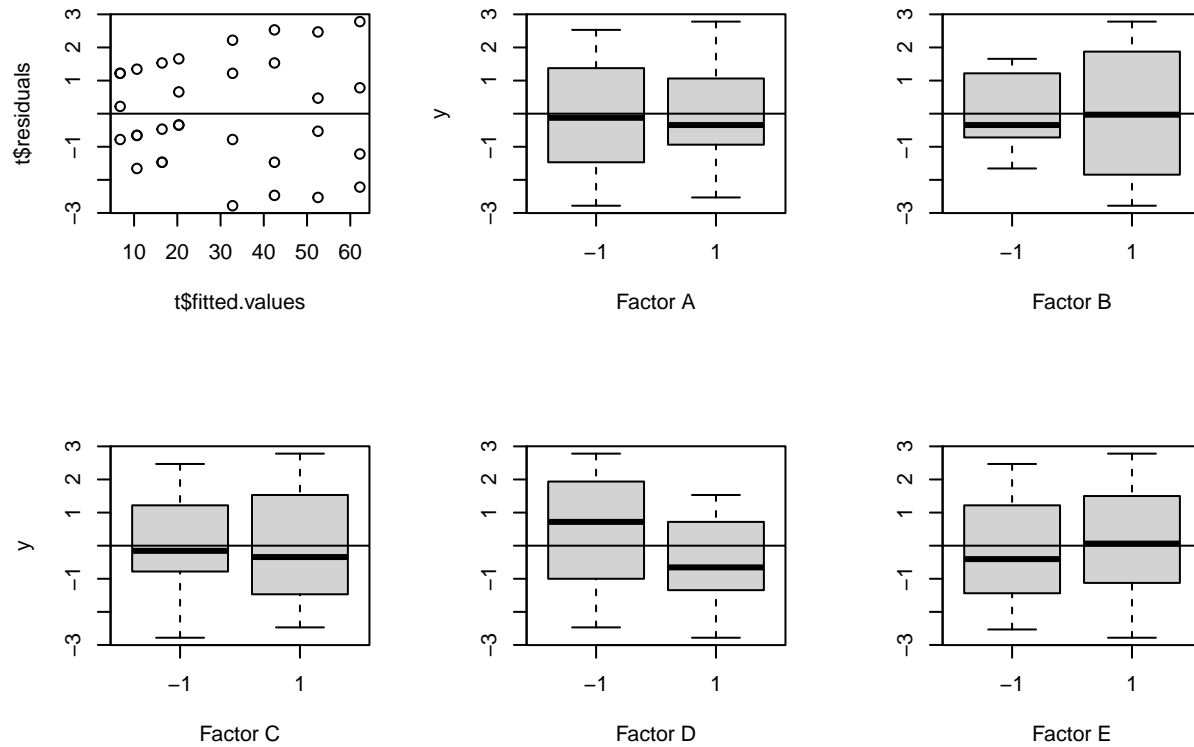
Normal Q-Q Plot



There is some variation in the Q-Q plot.

(e)

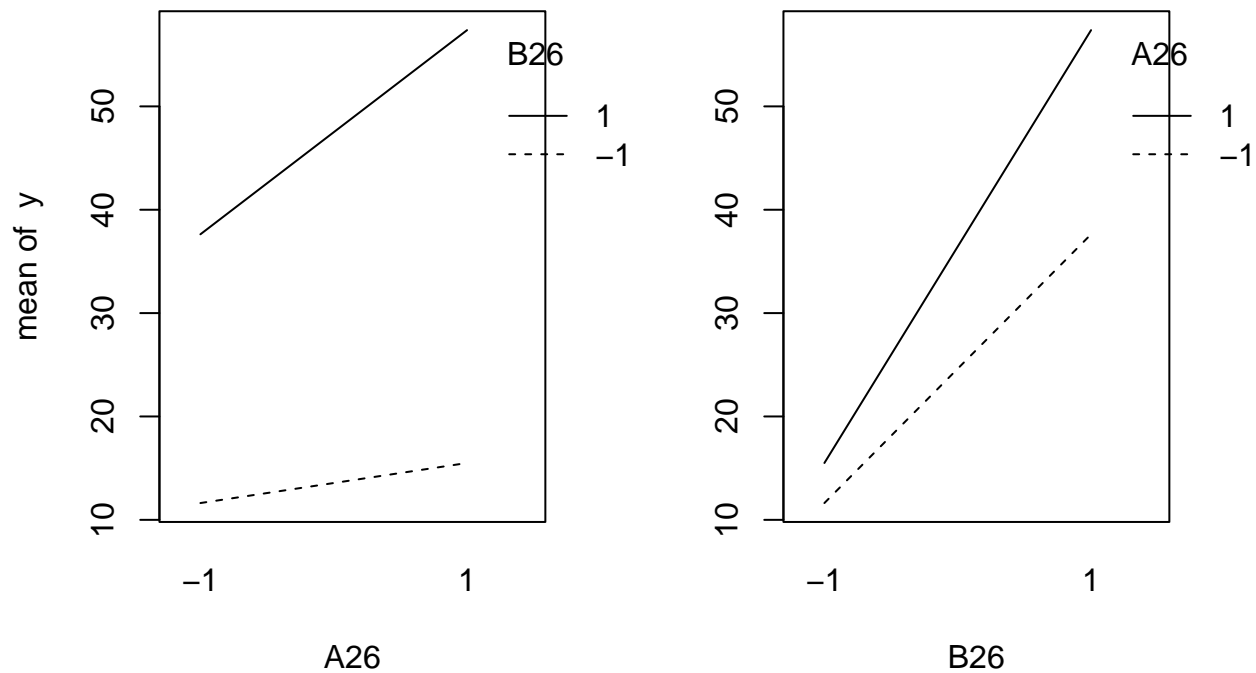
```
par(mfrow=c(2,3))
plot(t$fitted.values,t$residuals)
abline(h=0)
plot(A26,t$residuals,xlab="Factor A")
abline(h=0)
plot(B26,t$residuals,xlab="Factor B",ylab="")
abline(h=0)
plot(C26,t$residuals,xlab="Factor C")
abline(h=0)
plot(D26,t$residuals,xlab="Factor D",ylab="")
abline(h=0)
plot(E26,t$residuals,xlab="Factor E",ylab="")
abline(h=0)
```



There appears to be a megaphone effect in the residuals vs. fitted values plot.

(f)

```
par(mfrow=c(1,2))
interaction.plot(A26,B26,y)
interaction.plot(B26,A26,y,ylab="")
```



There appears to be interaction between aperture setting (A) and exposure time (B). This
could mean the two factors are related or associated with each other.

(g)

I would recommend to use the higher levels for aperture setting (A), exposure time (B), and development time (C) to increase yield. The coefficients calculated from the linear model show these three factors and the interaction between factors A and B are directly proportional to the response variable. We can see from the bottom row of the data on page 297 that the replicates corresponding to treatment combinations with the higher levels (abc, abcd, abce, abcde) have the greatest values.

(h)

```
sk<-data.frame(A26[1:2^3],B26[1:2^3],as.factor(as.numeric(as.character(A26))[1:2^3]*as.numeric(as.character(B26))[1:2^3]))
names(sk)<-c("A","B","AB","C","y") # 2^k design, k = 3
sk
```

```
##      A  B AB  C  y
## (1) -1 -1  1 -1  7
## a    1 -1 -1 -1  9
## b   -1  1 -1 -1 34
## ab    1  1  1 -1 55
## c   -1 -1  1  1 16
## ac    1 -1 -1  1 20
## bc   -1  1 -1  1 40
## abc    1  1  1  1 60

aA<-mean(c(y["a"],y["ab"],y["ac"],y["abc"]))
rA<-max(c(y["a"],y["ab"],y["ac"],y["abc"]))-min(c(y["a"],y["ab"],y["ac"],y["abc"]))
aB<-mean(c(y["b"],y["ab"],y["bc"],y["abc"]))
rB<-max(c(y["b"],y["ab"],y["bc"],y["abc"]))-min(c(y["b"],y["ab"],y["bc"],y["abc"]))
aAB<-mean(c(y["ab"],y["abc"]))
rAB<-max(c(y["ab"],y["abc"]))-min(c(y["ab"],y["abc"]))
aC<-mean(c(y["c"],y["ac"],y["bc"],y["abc"]))
rC<-max(c(y["c"],y["ac"],y["bc"],y["abc"]))-min(c(y["c"],y["ac"],y["bc"],y["abc"]))
ar<-data.frame(c(aA,rA),c(aB,rB),c(aAB,rAB),c(aC,rC))
names(ar)<-c("A","B","AB","C")
row.names(ar)<-c("Average","Range")
ar
```

```
##      A      B      AB  C
## Average 36 47.25 57.5 34
## Range   51 26.00  5.0 44
```

No, this “sketch” does not aid in interpreting the results of this experiment. I have no idea how this is relevant to the original data.

Problem 7.3

```
Block<-as.factor(rep(c("I","II"),each=16)) # Treating replicates as blocks
m<-lm(length~Block+factorA*factorB*factorC*factorD) # See Problem 6.15(b) for data
anova(m)
```

```
## Analysis of Variance Table
```

```
##
```

```
## Response: length
```

```
##
```

```
## Block
```

```
## factorA
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Block	1	0.016	0.016	0.1853	0.6730
factorA	1	72.909	72.909	852.5942	1.240e-14 ***

```
## factorB          1 126.461 126.461 1478.8275 < 2.2e-16 ***
## factorC          1 103.464 103.464 1209.9066 9.310e-16 ***
## factorD          1  30.662  30.662  358.5639 6.957e-12 ***
## factorA:factorB  1  29.927  29.927  349.9622 8.289e-12 ***
## factorA:factorC  1 128.496 128.496 1502.6344 < 2.2e-16 ***
## factorB:factorC  1   0.074   0.074   0.8622   0.3678
## factorA:factorD  1   0.047   0.047   0.5475   0.4708
## factorB:factorD  1   0.018   0.018   0.2089   0.6542
## factorC:factorD  1   0.047   0.047   0.5529   0.4686
## factorA:factorB:factorC  1  78.751  78.751  920.9150 7.017e-15 ***
## factorA:factorB:factorD  1   0.077   0.077   0.8985   0.3582
## factorA:factorC:factorD  1   0.003   0.003   0.0342   0.8557
## factorB:factorC:factorD  1   0.010   0.010   0.1187   0.7352
## factorA:factorB:factorC:factorD  1   0.002   0.002   0.0187   0.8931
## Residuals        15   1.283   0.086
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# We reject the null hypotheses at the alpha = 0.05 level for all three factors and the
# interactions between factors A and B; factors A and C; and factors A, B, and C. There is
# sufficient evidence (p < 0.0001) that all factors and all combinations of interactions
# between factors A, B, and C (except the interaction between B and C) are significant.
shapiro.test(m$residuals)
```

```
##
## Shapiro-Wilk normality test
##
## data:  m$residuals
## W = 0.92271, p-value = 0.02465
```

```
bartlett.test(m$residuals~factorA)
```

```
##
## Bartlett test of homogeneity of variances
##
## data:  m$residuals by factorA
## Bartlett's K-squared = 0.028907, df = 1, p-value = 0.865
```

```
bartlett.test(m$residuals~factorB)
```

```
##
## Bartlett test of homogeneity of variances
##
## data:  m$residuals by factorB
## Bartlett's K-squared = 0.032149, df = 1, p-value = 0.8577
```

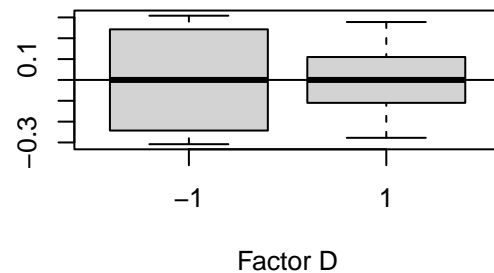
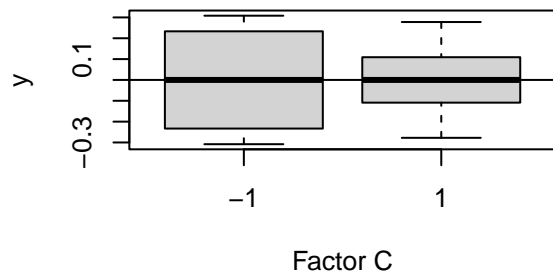
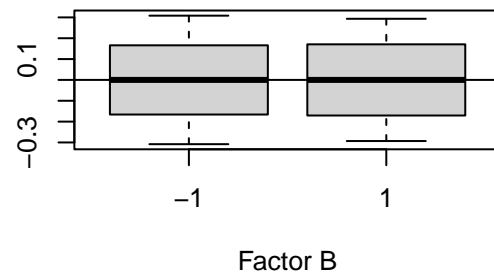
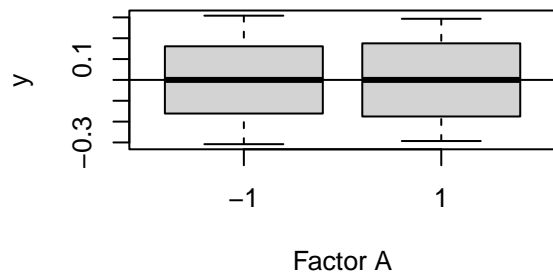
```
bartlett.test(m$residuals~factorC)
```

```
##
## Bartlett test of homogeneity of variances
##
## data:  m$residuals by factorC
## Bartlett's K-squared = 0.76615, df = 1, p-value = 0.3814
```

```
bartlett.test(m$residuals~factorD)
```

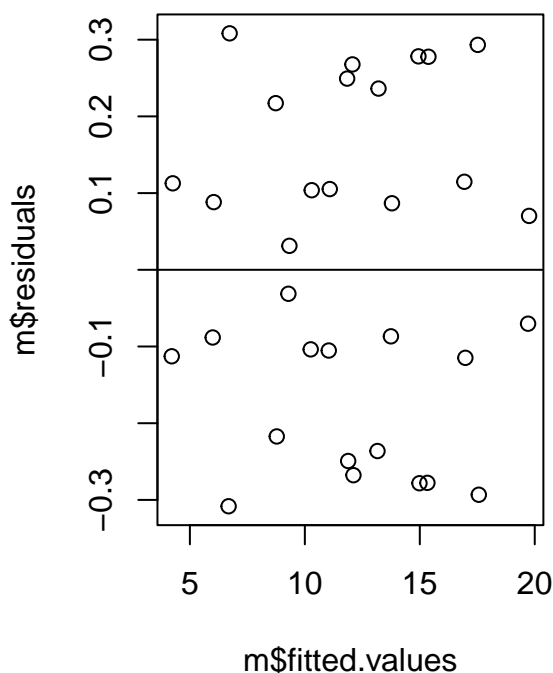
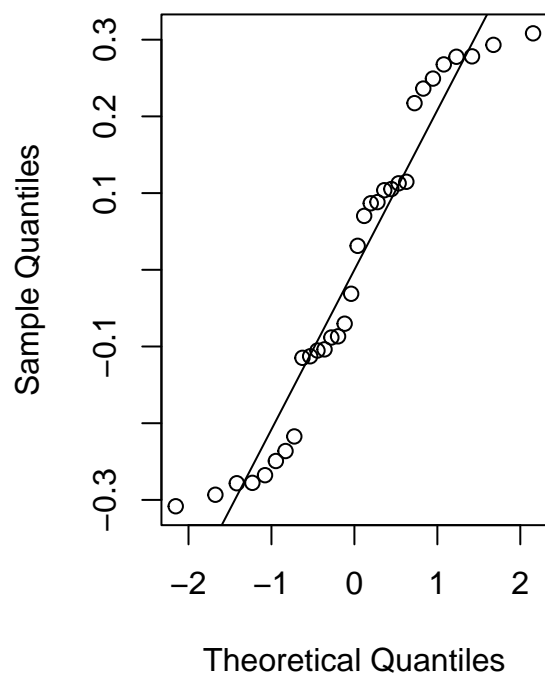
```
##
## Bartlett test of homogeneity of variances
```

```
##
## data: m$residuals by factorD
## Bartlett's K-squared = 0.93641, df = 1, p-value = 0.3332
# We should exercise some caution here as the null hypothesis for the Shapiro test has been
par(mfrow=c(2,2)) # rejected (p = 0.02465) and the assumption of normality may be violated.
plot(factorA,m$residuals,xlab="Factor A") # See Problem 6.15(f) for interaction plots
abline(h=0)
plot(factorB,m$residuals,xlab="Factor B",ylab="")
abline(h=0)
plot(factorC,m$residuals,xlab="Factor C")
abline(h=0)
plot(factorD,m$residuals,xlab="Factor D",ylab="")
abline(h=0)
```



```
par(mfrow=c(1,2))
qqnorm(m$residuals)
qqline(m$residuals)
plot(m$fitted.values,m$residuals)
abline(h=0)
```

Normal Q-Q Plot



There is variation in the Q-Q plot.

Problem 7.7

```
ABCDE<-as.numeric(as.character(A26))*as.numeric(as.character(B26))*as.numeric(as.character(C26))*as.numeric(as.character(D26))
u<-lm(y~A26*B26*C26*D26+A26*B26*C26*E26+A26*B26*D26*E26+A26*C26*D26*E26+B26*C26*D26*E26+ABCDE)
anova(u)
```

```
## Warning in anova.lm(u): ANOVA F-tests on an essentially perfect fit are
## unreliable
```

```
## Analysis of Variance Table
```

```
##
```

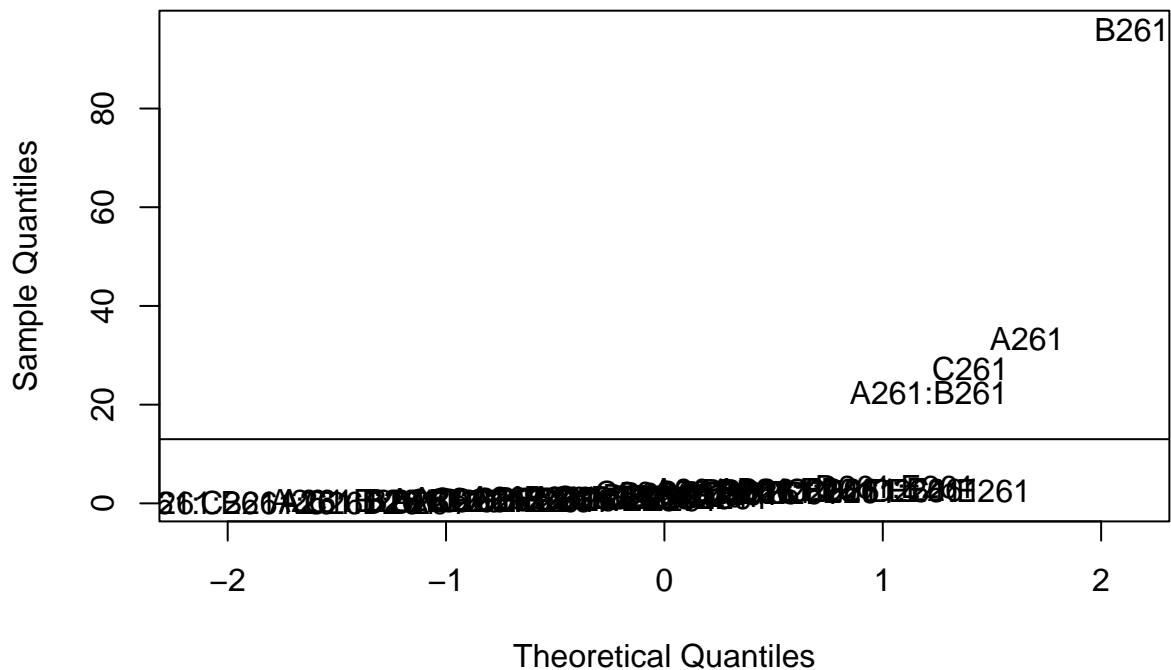
```
## Response: y
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
A26	1	1116.3	1116.3	NaN	NaN
B26	1	9214.0	9214.0	NaN	NaN
C26	1	750.8	750.8	NaN	NaN
D26	1	5.3	5.3	NaN	NaN
E26	1	1.5	1.5	NaN	NaN
ABCDE	1	0.3	0.3	NaN	NaN
A26:B26	1	504.0	504.0	NaN	NaN
A26:C26	1	1.5	1.5	NaN	NaN
B26:C26	1	0.0	0.0	NaN	NaN
A26:D26	1	0.0	0.0	NaN	NaN
B26:D26	1	3.8	3.8	NaN	NaN
C26:D26	1	5.3	5.3	NaN	NaN
A26:E26	1	7.0	7.0	NaN	NaN
B26:E26	1	2.5	2.5	NaN	NaN
C26:E26	1	0.8	0.8	NaN	NaN

```
## D26:E26      1  11.3  11.3   NaN   NaN
## A26:B26:C26  1   1.5   1.5   NaN   NaN
## A26:B26:D26  1   0.8   0.8   NaN   NaN
## A26:C26:D26  1   1.5   1.5   NaN   NaN
## B26:C26:D26  1   1.5   1.5   NaN   NaN
## A26:B26:E26  1   0.3   0.3   NaN   NaN
## A26:C26:E26  1   0.8   0.8   NaN   NaN
## B26:C26:E26  1   7.0   7.0   NaN   NaN
## A26:D26:E26  1   5.3   5.3   NaN   NaN
## B26:D26:E26  1   0.3   0.3   NaN   NaN
## C26:D26:E26  1   5.3   5.3   NaN   NaN
## A26:B26:C26:D26 1   0.0   0.0   NaN   NaN
## A26:B26:C26:E26 1   0.3   0.3   NaN   NaN
## A26:B26:D26:E26 1   7.0   7.0   NaN   NaN
## A26:C26:D26:E26 1   0.8   0.8   NaN   NaN
## B26:C26:D26:E26 1   7.0   7.0   NaN   NaN
## Residuals    0   0.0   NaN
```

```
qq<-qqnorm(abs(u$effects[-1]),type="n") # Remove variables
text(qq$x,qq$y,labels=names(abs(u$effects[-1])))
abline(h=13) # Arbitrary cutoff
```

Normal Q-Q Plot



```
v<-lm(y~A26+B26+C26+A26*B26+ABCDE) # New model
anova(v)
```

```
## Analysis of Variance Table
```

```
##
```

```
## Response: y
```

```
##          Df Sum Sq Mean Sq  F value    Pr(>F)
## A26        1 1116.3  1116.3   369.4296 < 2.2e-16 ***
## B26        1  9214.0  9214.0  3049.3532 < 2.2e-16 ***
```

```
## C26      1  750.8   750.8  248.4686 8.027e-15 ***
## ABCDE    1    0.3     0.3    0.0931   0.7627
## A26:B26   1  504.0   504.0  166.8075 8.080e-13 ***
## Residuals 26   78.6     3.0
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

# We reject the null hypotheses at the alpha = 0.05 level for all three factors and
# the interaction between factors A and B. There is sufficient evidence (p < 0.0001)
# that all factors and the interaction between factors A and B are significant.
shapiro.test(v$residuals)

##
##  Shapiro-Wilk normality test
##
## data:  v$residuals
## W = 0.95496, p-value = 0.199
bartlett.test(v$residuals~A26)

##
##  Bartlett test of homogeneity of variances
##
## data:  v$residuals by A26
## Bartlett's K-squared = 0.17093, df = 1, p-value = 0.6793
bartlett.test(v$residuals~B26)

##
##  Bartlett test of homogeneity of variances
##
## data:  v$residuals by B26
## Bartlett's K-squared = 4.1683, df = 1, p-value = 0.04119
bartlett.test(v$residuals~C26)

##
##  Bartlett test of homogeneity of variances
##
## data:  v$residuals by C26
## Bartlett's K-squared = 0.15369, df = 1, p-value = 0.695
bartlett.test(v$residuals~D26)

##
##  Bartlett test of homogeneity of variances
##
## data:  v$residuals by D26
## Bartlett's K-squared = 1.024, df = 1, p-value = 0.3116
bartlett.test(v$residuals~E26)

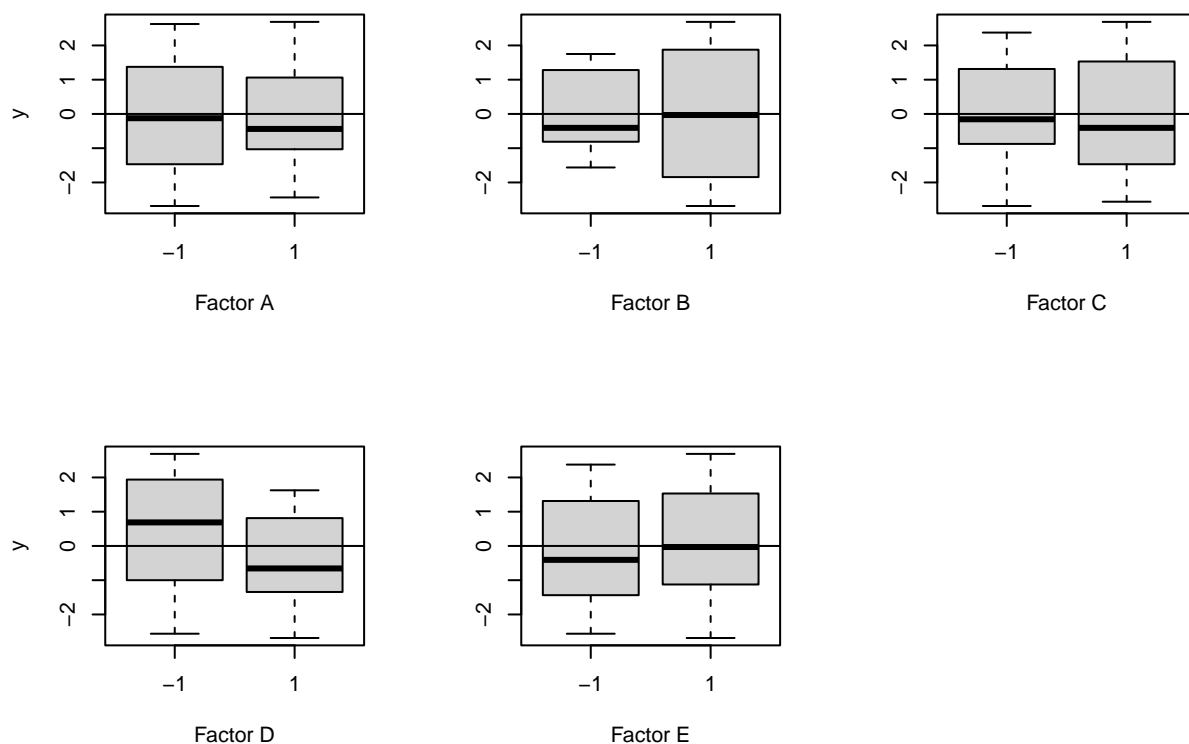
##
##  Bartlett test of homogeneity of variances
##
## data:  v$residuals by E26
## Bartlett's K-squared = 0.041686, df = 1, p-value = 0.8382
```



```

# We should exercise some caution here as the null hypothesis of the Bartlett's
# test for factor B has been rejected ( $p = 0.04119$ ) and the assumption of equal
par(mfrow=c(2,3)) # variances for factor B may be violated.
plot(A26,v$residuals,xlab="Factor A") # See Problem 6.26(f) for interaction plots
abline(h=0)
plot(B26,v$residuals,xlab="Factor B",ylab="")
abline(h=0)
plot(C26,v$residuals,xlab="Factor C",ylab="")
abline(h=0)
plot(D26,v$residuals,xlab="Factor D")
abline(h=0)
plot(E26,v$residuals,xlab="Factor E",ylab="")
abline(h=0)
par(mfrow=c(1,2))

```

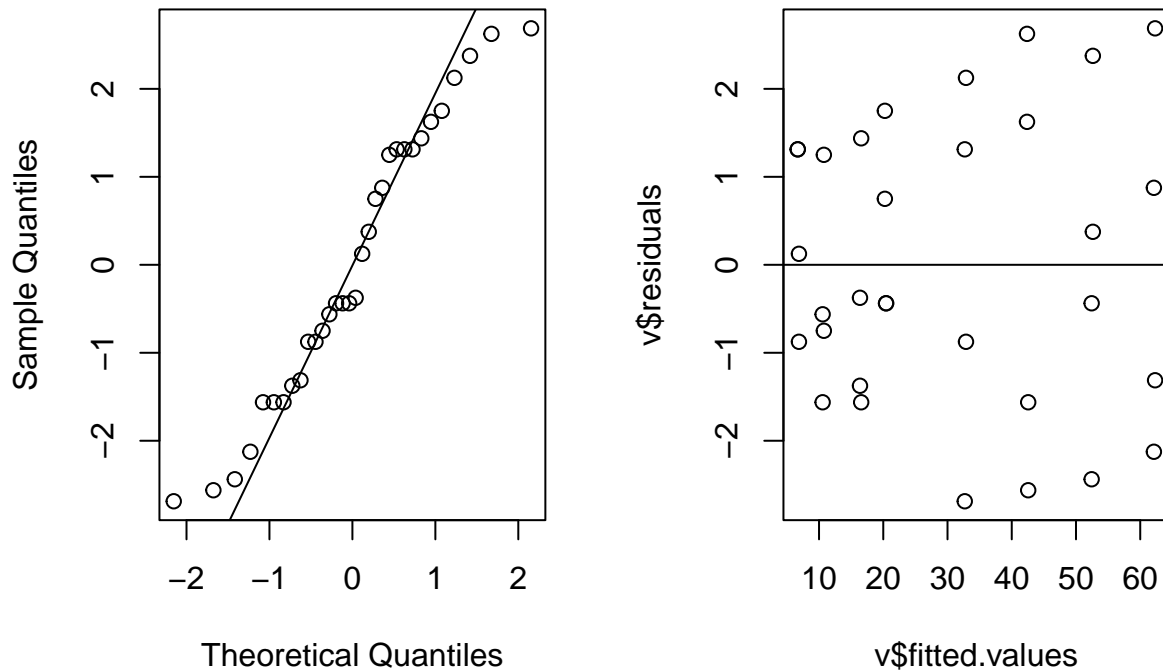


```

qqnorm(v$residuals)
qqline(v$residuals)
plot(v$fitted.values,v$residuals)
abline(h=0)

```

Normal Q-Q Plot



*# There is some variation in the Q-Q plot and a megaphone effect in the
residuals vs. fitted values plot.*

Problem 7.9

```
ACDE<-as.numeric(as.character(A26))*as.numeric(as.character(C26))*as.numeric(as.character(D26))*as.nu
BCD<-as.numeric(as.character(B26))*as.numeric(as.character(C26))*as.numeric(as.character(D26))
ABE<-as.numeric(as.character(A26))*as.numeric(as.character(B26))*as.numeric(as.character(E26))
blocks<-factor(levels=c(1:4))
for(i in 1:32) {
  if(ACDE[i]==-1 & BCD[i]==-1) blocks[i]=1
  if(ACDE[i]==1 & BCD[i]==-1) blocks[i]=2
  if(ACDE[i]==-1 & BCD[i]==1) blocks[i]=3
  if(ACDE[i]==1 & BCD[i]==1) blocks[i]=4
}
w<-lm(y~blocks+A26*B26*C26*D26*E26)
anova(w)
```

```
## Warning in anova.lm(w): ANOVA F-tests on an essentially perfect fit are
## unreliable
```

```
## Analysis of Variance Table
```

```
##
```

```
## Response: y
```

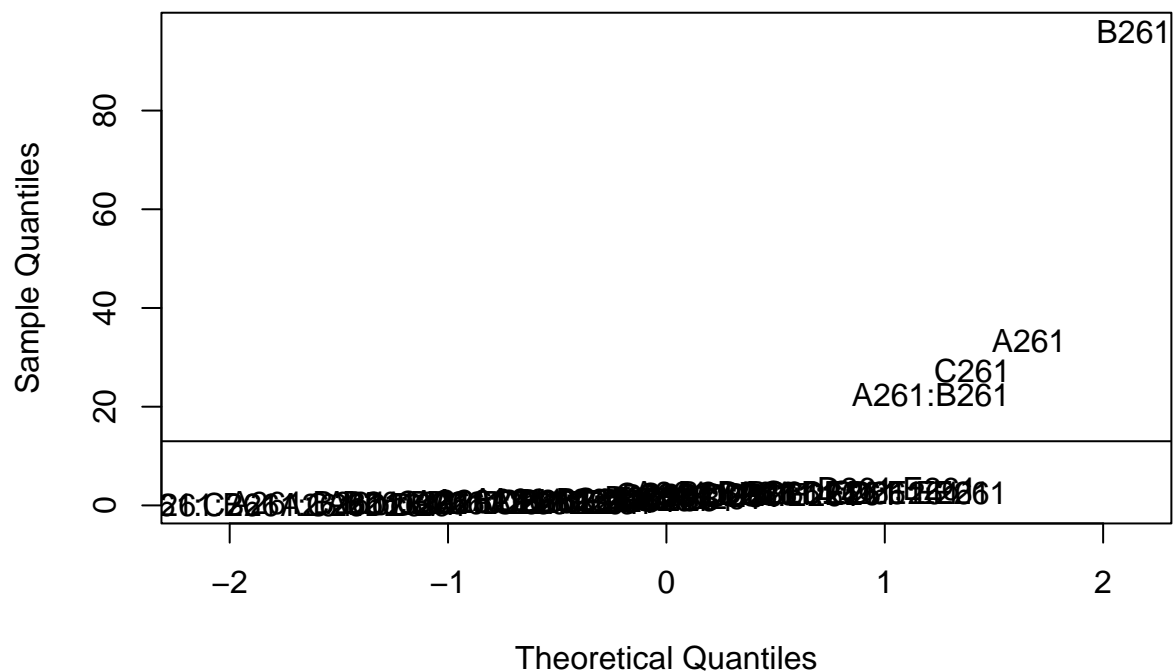
```
##
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
## blocks	3	2.6	0.9	NaN	NaN
## A26	1	1116.3	1116.3	NaN	NaN
## B26	1	9214.0	9214.0	NaN	NaN
## C26	1	750.8	750.8	NaN	NaN
## D26	1	5.3	5.3	NaN	NaN

## E26	1	1.5	1.5	NaN	NaN
## A26:B26	1	504.0	504.0	NaN	NaN
## A26:C26	1	1.5	1.5	NaN	NaN
## B26:C26	1	0.0	0.0	NaN	NaN
## A26:D26	1	0.0	0.0	NaN	NaN
## B26:D26	1	3.8	3.8	NaN	NaN
## C26:D26	1	5.3	5.3	NaN	NaN
## A26:E26	1	7.0	7.0	NaN	NaN
## B26:E26	1	2.5	2.5	NaN	NaN
## C26:E26	1	0.8	0.8	NaN	NaN
## D26:E26	1	11.3	11.3	NaN	NaN
## A26:B26:C26	1	1.5	1.5	NaN	NaN
## A26:B26:D26	1	0.8	0.8	NaN	NaN
## A26:C26:D26	1	1.5	1.5	NaN	NaN
## A26:C26:E26	1	0.8	0.8	NaN	NaN
## B26:C26:E26	1	7.0	7.0	NaN	NaN
## A26:D26:E26	1	5.3	5.3	NaN	NaN
## B26:D26:E26	1	0.3	0.3	NaN	NaN
## C26:D26:E26	1	5.3	5.3	NaN	NaN
## A26:B26:C26:D26	1	0.0	0.0	NaN	NaN
## A26:B26:C26:E26	1	0.3	0.3	NaN	NaN
## A26:B26:D26:E26	1	7.0	7.0	NaN	NaN
## B26:C26:D26:E26	1	7.0	7.0	NaN	NaN
## A26:B26:C26:D26:E26	1	0.3	0.3	NaN	NaN
## Residuals	0	0.0	NaN		

```
qq<-qqnorm(abs(w$effects[-1]),type="n") # Remove variables
text(qq$x,qq$y,labels=names(abs(w$effects[-1])))
abline(h=13) # Arbitrary cutoff
```

Normal Q-Q Plot



```

x<-lm(y~A26+B26+C26+A26*B26+blocks) # New model
anova(x)

## Analysis of Variance Table
##
## Response: y
##           Df Sum Sq Mean Sq  F value    Pr(>F)
## A26         1 1116.3   1116.3   351.3541 7.758e-16 ***
## B26         1 9214.0   9214.0 2900.1541 < 2.2e-16 ***
## C26         1  750.8    750.8   236.3115 6.357e-14 ***
## blocks      3    2.6     0.9    0.2721   0.8449
## A26:B26     1  504.0    504.0   158.6459 4.556e-12 ***
## Residuals  24   76.2     3.2
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

# We reject the null hypotheses at the alpha = 0.05 level for all three factors and
# the interaction between factors A and B. There is sufficient evidence (p < 0.0001)
# that all factors and the interaction between factors A and B are significant.
shapiro.test(x$residuals)

##
## Shapiro-Wilk normality test
##
## data:  x$residuals
## W = 0.95601, p-value = 0.2129
bartlett.test(x$residuals~A26)

##
## Bartlett test of homogeneity of variances
##
## data:  x$residuals by A26
## Bartlett's K-squared = 0.14589, df = 1, p-value = 0.7025
bartlett.test(x$residuals~B26)

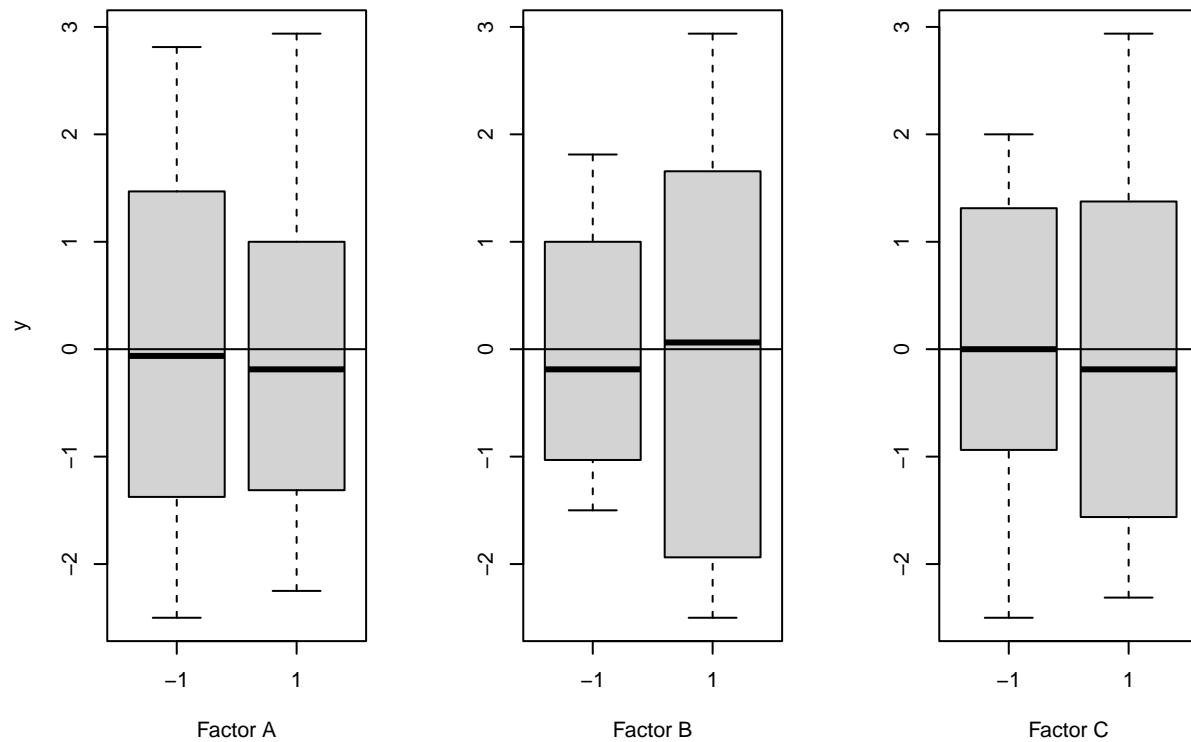
##
## Bartlett test of homogeneity of variances
##
## data:  x$residuals by B26
## Bartlett's K-squared = 3.7484, df = 1, p-value = 0.05286
bartlett.test(x$residuals~C26)

##
## Bartlett test of homogeneity of variances
##
## data:  x$residuals by C26
## Bartlett's K-squared = 0.54885, df = 1, p-value = 0.4588

# We should exercise some caution here as the null hypothesis of the Bartlett's
# test for factor B is close to being rejected (p = 0.05286). The assumption of
par(mfrow=c(1,3)) # equal variances for factor B may be slightly violated.
plot(A26,x$residuals,xlab="Factor A") # See Problem 6.26(f) for interaction plots
abline(h=0)
plot(B26,x$residuals,xlab="Factor B",ylab="")
abline(h=0)

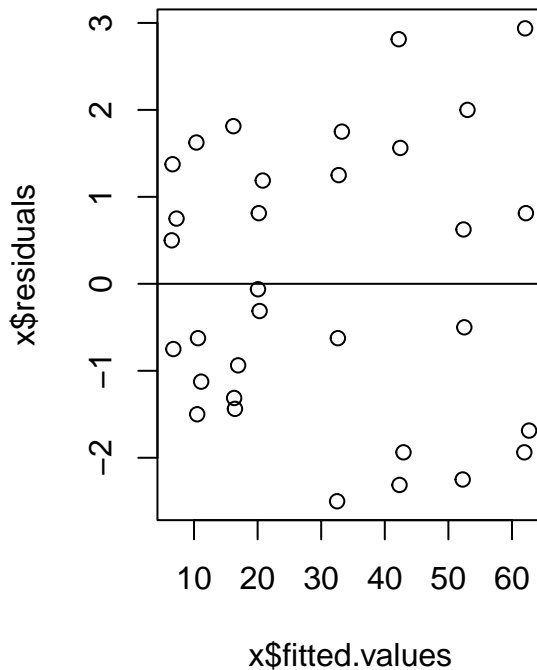
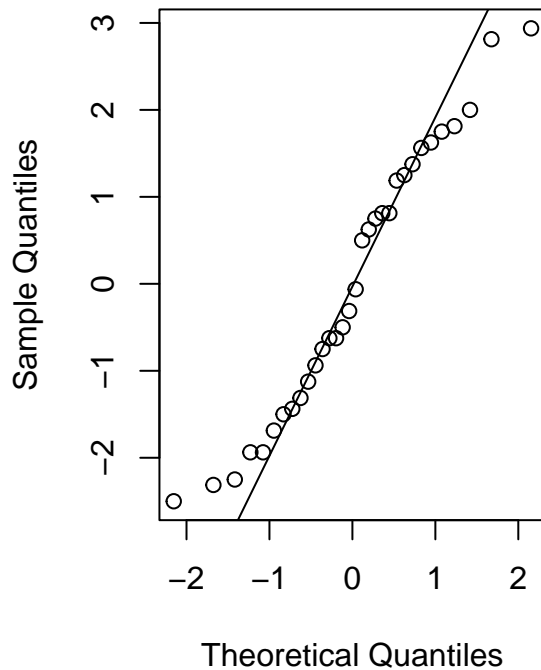
```

```
plot(C26,x$residuals,xlab="Factor C",ylab="")
abline(h=0)
```



```
par(mfrow=c(1,2))
qqnorm(x$residuals)
qqline(x$residuals)
plot(x$fitted.values,x$residuals)
abline(h=0)
```

Normal Q-Q Plot



*# There is some variation in the Q-Q plot and a slight megaphone effect in the
residuals vs. fitted values plot.*

Problem 7.27

```
A27<-as.factor(rep(c(-1,1),4))
B27<-as.factor(rep(c(-1,1),2,each=2))
C27<-as.factor(rep(c(-1,1),each=4))
ABC<-as.numeric(as.character(A27))*as.numeric(as.character(B27))*as.numeric(as.character(C27))
BC<-as.numeric(as.character(B27))*as.numeric(as.character(C27))
Rep<-as.factor(rep(c("I","II","III"),each=8))
Block<-as.factor(c(rep(ABC/2+3/2,2),BC/2+7/2))
a<-1:(2^3*3) # Random data
factorA27<-as.factor(rep(A27,3))
factorB27<-as.factor(rep(B27,3))
factorC27<-as.factor(rep(C27,3))
b<-lm(a~Rep+Block%in%Rep+factorA27*factorB27*factorC27)
anova(b)
```

```
## Warning in anova.lm(b): ANOVA F-tests on an essentially perfect fit are
## unreliable
```

```
## Analysis of Variance Table
```

```
##
```

```
## Response: a
```

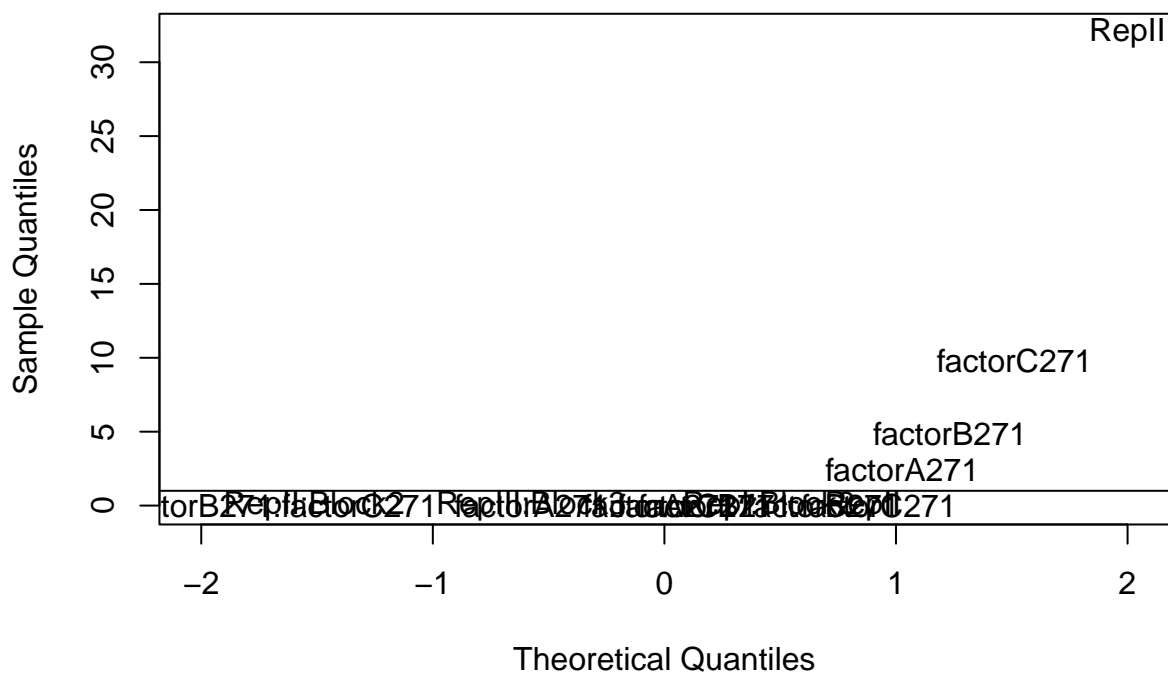
```
##
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
## Rep	2	1024	512	8.5240e+32	<2e-16 ***
## factorA27	1	6	6	9.9890e+30	<2e-16 ***
## factorB27	1	24	24	3.9956e+31	<2e-16 ***
## factorC27	1	96	96	1.5982e+32	<2e-16 ***

```
## Rep:Block          3      0      0 6.6800e-01 0.5891
## factorA27:factorB27 1      0      0 1.5188e+00 0.2435
## factorA27:factorC27 1      0      0 8.9260e-01 0.3651
## factorB27:factorC27 1      0      0 1.8940e+00 0.1961
## factorA27:factorB27:factorC27 1      0      0 1.0300e-02 0.9211
## Residuals         11      0      0
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

qq<-qqnorm(abs(b$effects[-1]),type="n") # Remove variables
text(qq$x,qq$y,labels=names(abs(b$effects[-1])))
abline(h=1) # Arbitrary cutoff
```

Normal Q-Q Plot



```
c<-lm(a~Rep+Block%in%Rep+factorA27+factorB27+factorC27) # New model
anova(c)

## Warning in anova.lm(c): ANOVA F-tests on an essentially perfect fit are
## unreliable

## Analysis of Variance Table
##
## Response: a
##          Df Sum Sq Mean Sq    F value Pr(>F)
## Rep       2   1024      512 8.3483e+32 <2e-16 ***
## factorA27 1     6         6 9.7832e+30 <2e-16 ***
## factorB27 1    24        24 3.9133e+31 <2e-16 ***
## factorC27 1    96        96 1.5653e+32 <2e-16 ***
## Rep:Block 3     0         0 6.5420e-01 0.5927
## Residuals 15     0         0
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```

# We reject the null hypotheses at the alpha = 0.05 level for all three factors. There
# is sufficient evidence (p < 0.0001) that all factors are significant.
shapiro.test(c$residuals)

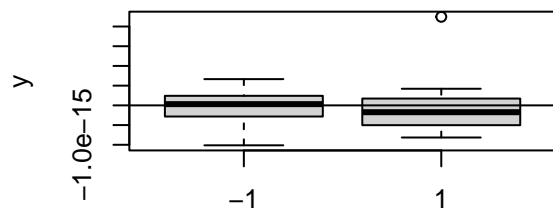
##
##  Shapiro-Wilk normality test
##
## data:  c$residuals
## W = 0.85505, p-value = 0.002716
bartlett.test(c$residuals~factorA27)

##
##  Bartlett test of homogeneity of variances
##
## data:  c$residuals by factorA27
## Bartlett's K-squared = 3.0011, df = 1, p-value = 0.08321
bartlett.test(c$residuals~factorB27)

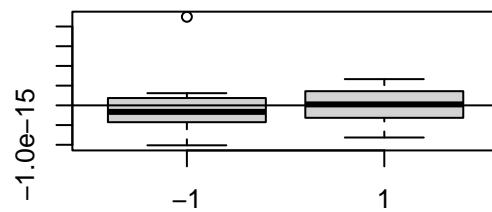
##
##  Bartlett test of homogeneity of variances
##
## data:  c$residuals by factorB27
## Bartlett's K-squared = 3.3318, df = 1, p-value = 0.06795
bartlett.test(c$residuals~factorC27)

##
##  Bartlett test of homogeneity of variances
##
## data:  c$residuals by factorC27
## Bartlett's K-squared = 1.022, df = 1, p-value = 0.312
# We should exercise some caution here as the null hypothesis for the Shapiro-Wilk test
# has been rejected (p = 0.002716) and the assumption of normality may be violated. The
# null hypotheses of the Bartlett's tests for factors A and B are also close to being
# rejected (p = 0.08321, p = 0.06795) and the assumption of equal variances for
par(mfrow=c(2,2)) # factors A and B may be slightly violated.
plot(factorA27,c$residuals,xlab="Factor A") # No interaction terms
abline(h=0)
plot(factorB27,c$residuals,xlab="Factor B",ylab="")
abline(h=0)
plot(factorC27,c$residuals,xlab="Factor C")
abline(h=0)
plot(Block,c$residuals,xlab="Block",ylab="")
abline(h=0)

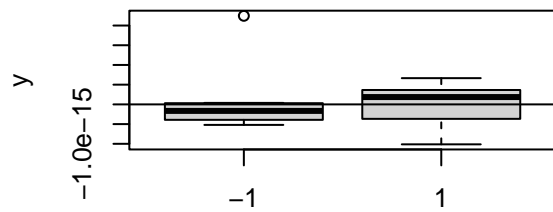
```

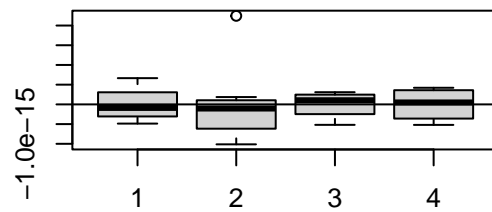
Factor A



Factor B



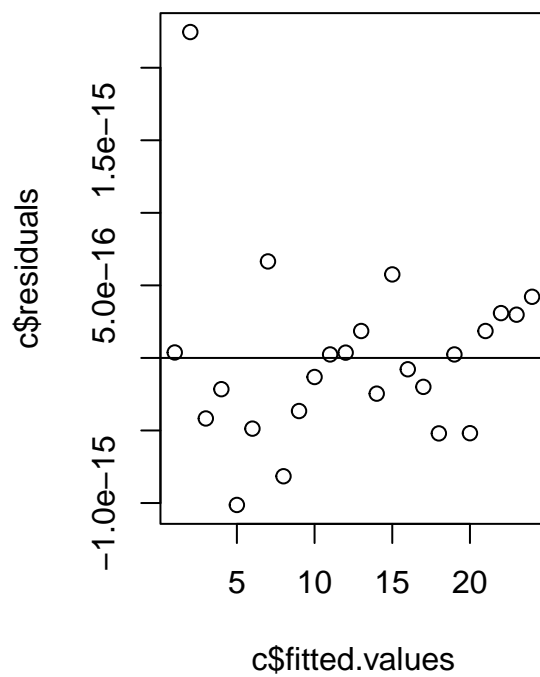
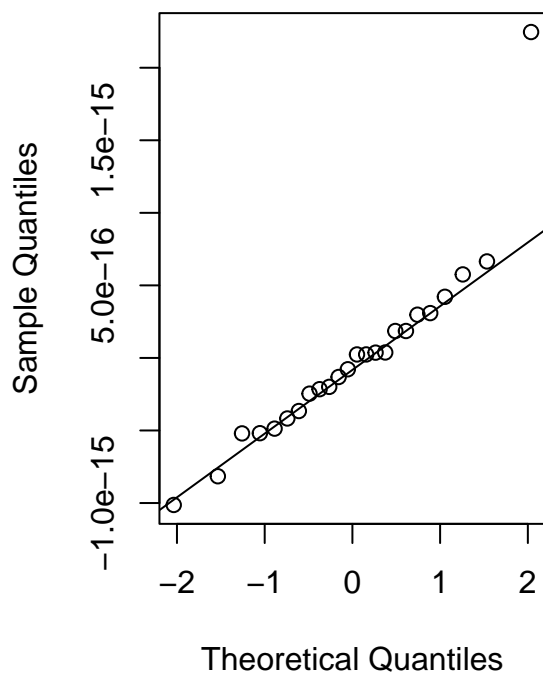
Factor C



Block

```
par(mfrow=c(1,2))
qqnorm(c$residuals)
qqline(c$residuals)
plot(c$fitted.values,c$residuals)
abline(h=0)
```

Normal Q-Q Plot



```
# There does not appear to be any distinct pattern in either the Q-Q plot or
# residuals vs. fitted values plot.
```

Problem 7.20 (Extra Credit)

```
A20<-as.factor(rep(c(-1,1),32)) # 2^6 factorial
B20<-as.factor(rep(c(-1,1),16,each=2))
C20<-as.factor(rep(c(-1,1),8,each=4))
D20<-as.factor(rep(c(-1,1),4,each=8))
E20<-as.factor(rep(c(-1,1),2,each=16))
F20<-as.factor(rep(c(-1,1),each=32)) # Choosing arbitrary blocks
ABC<-as.numeric(as.character(A20))*as.numeric(as.character(B20))*as.numeric(as.character(C20))
DEF<-as.numeric(as.character(D20))*as.numeric(as.character(E20))*as.numeric(as.character(F20))
blocks<-factor(levels=c(1:4))
for(i in 1:64) {
  if(ABC[i]==-1 & DEF[i]==-1) blocks[i]=1
  if(ABC[i]==1 & DEF[i]==-1) blocks[i]=2
  if(ABC[i]==-1 & DEF[i]==1) blocks[i]=3
  if(ABC[i]==1 & DEF[i]==1) blocks[i]=4
}
aa<-1:2^6 # Random data
bb<-lm(aa~blocks+A20*B20*C20*D20*E20*F20)
anova(bb)
```

```
## Warning in anova.lm(bb): ANOVA F-tests on an essentially perfect fit are
## unreliable
```

```
## Analysis of Variance Table
```

```
##
```

```
## Response: aa
```

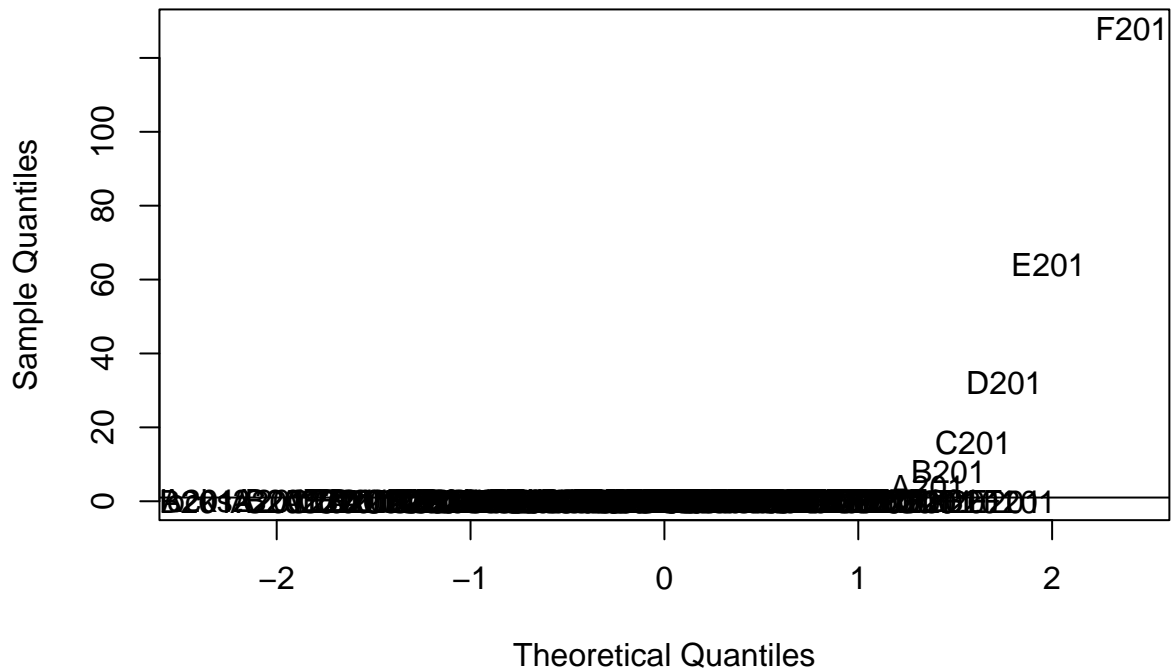
```
##
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
## blocks	3	0	0	NaN	NaN
## A20	1	16	16	NaN	NaN
## B20	1	64	64	NaN	NaN
## C20	1	256	256	NaN	NaN
## D20	1	1024	1024	NaN	NaN
## E20	1	4096	4096	NaN	NaN
## F20	1	16384	16384	NaN	NaN
## A20:B20	1	0	0	NaN	NaN
## A20:C20	1	0	0	NaN	NaN
## B20:C20	1	0	0	NaN	NaN
## A20:D20	1	0	0	NaN	NaN
## B20:D20	1	0	0	NaN	NaN
## C20:D20	1	0	0	NaN	NaN
## A20:E20	1	0	0	NaN	NaN
## B20:E20	1	0	0	NaN	NaN
## C20:E20	1	0	0	NaN	NaN
## D20:E20	1	0	0	NaN	NaN
## A20:F20	1	0	0	NaN	NaN
## B20:F20	1	0	0	NaN	NaN
## C20:F20	1	0	0	NaN	NaN
## D20:F20	1	0	0	NaN	NaN
## E20:F20	1	0	0	NaN	NaN
## A20:B20:D20	1	0	0	NaN	NaN

## A20:C20:D20	1	0	0	NaN	NaN
## B20:C20:D20	1	0	0	NaN	NaN
## A20:B20:E20	1	0	0	NaN	NaN
## A20:C20:E20	1	0	0	NaN	NaN
## B20:C20:E20	1	0	0	NaN	NaN
## A20:D20:E20	1	0	0	NaN	NaN
## B20:D20:E20	1	0	0	NaN	NaN
## C20:D20:E20	1	0	0	NaN	NaN
## A20:B20:F20	1	0	0	NaN	NaN
## A20:C20:F20	1	0	0	NaN	NaN
## B20:C20:F20	1	0	0	NaN	NaN
## A20:D20:F20	1	0	0	NaN	NaN
## B20:D20:F20	1	0	0	NaN	NaN
## C20:D20:F20	1	0	0	NaN	NaN
## A20:E20:F20	1	0	0	NaN	NaN
## B20:E20:F20	1	0	0	NaN	NaN
## C20:E20:F20	1	0	0	NaN	NaN
## A20:B20:C20:D20	1	0	0	NaN	NaN
## A20:B20:C20:E20	1	0	0	NaN	NaN
## A20:B20:D20:E20	1	0	0	NaN	NaN
## A20:C20:D20:E20	1	0	0	NaN	NaN
## B20:C20:D20:E20	1	0	0	NaN	NaN
## A20:B20:C20:F20	1	0	0	NaN	NaN
## A20:B20:D20:F20	1	0	0	NaN	NaN
## A20:C20:D20:F20	1	0	0	NaN	NaN
## B20:C20:D20:F20	1	0	0	NaN	NaN
## A20:B20:E20:F20	1	0	0	NaN	NaN
## A20:C20:E20:F20	1	0	0	NaN	NaN
## B20:C20:E20:F20	1	0	0	NaN	NaN
## A20:D20:E20:F20	1	0	0	NaN	NaN
## B20:D20:E20:F20	1	0	0	NaN	NaN
## C20:D20:E20:F20	1	0	0	NaN	NaN
## A20:B20:C20:D20:E20	1	0	0	NaN	NaN
## A20:B20:C20:D20:F20	1	0	0	NaN	NaN
## A20:B20:C20:E20:F20	1	0	0	NaN	NaN
## A20:B20:D20:E20:F20	1	0	0	NaN	NaN
## A20:C20:D20:E20:F20	1	0	0	NaN	NaN
## B20:C20:D20:E20:F20	1	0	0	NaN	NaN
## Residuals	0	0	NaN		

```
qq<-qqnorm(abs(bb$effects[-1]),type="n") # Remove variables
text(qq$x,qq$y,labels=names(abs(bb$effects[-1])))
abline(h=1) # Arbitrary cutoff
```

Normal Q-Q Plot



```
cc<-lm(aa~blocks+A20+B20+C20+D20+E20+F20) # New model
anova(cc)
```

```
## Warning in anova.lm(cc): ANOVA F-tests on an essentially perfect fit are
## unreliable
```

```
## Analysis of Variance Table
```

```
##
```

```
## Response: aa
```

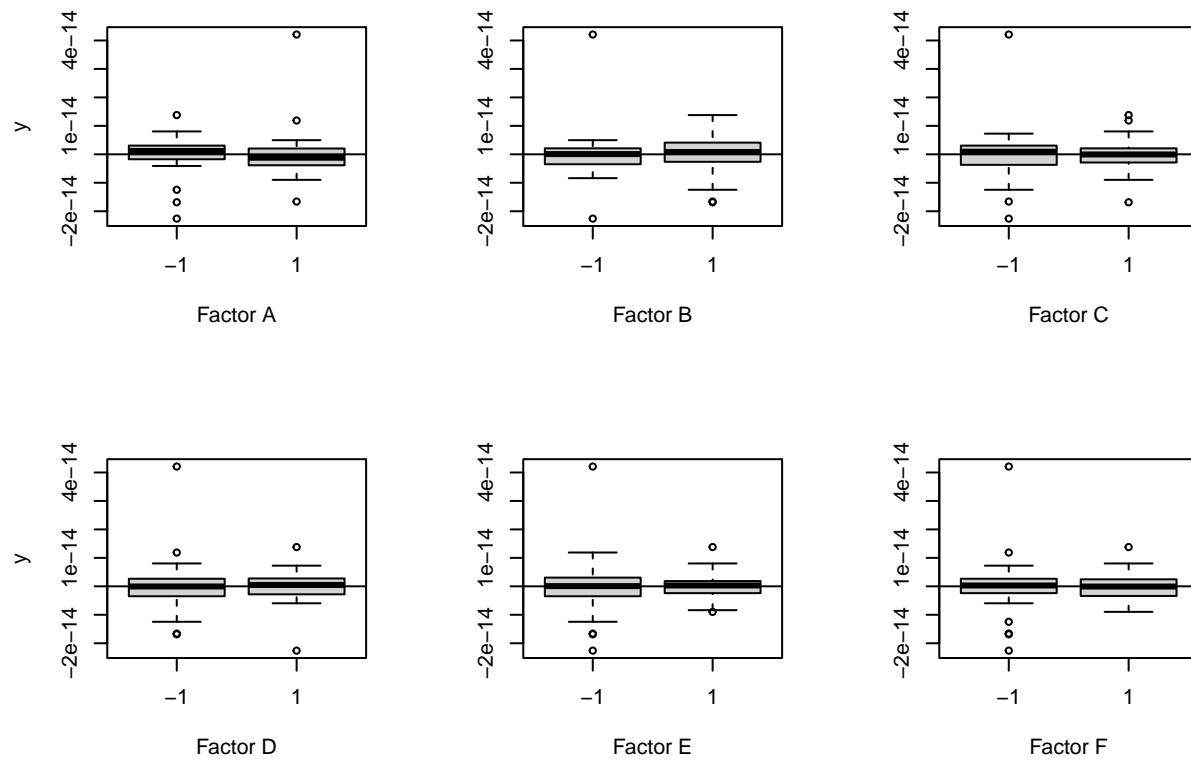
	Df	Sum Sq	Mean Sq	F value	Pr(>F)
blocks	3	0	0	3.7983e+00	0.01523 *
A20	1	16	16	2.1242e+29	< 2e-16 ***
B20	1	64	64	8.4969e+29	< 2e-16 ***
C20	1	256	256	3.3988e+30	< 2e-16 ***
D20	1	1024	1024	1.3595e+31	< 2e-16 ***
E20	1	4096	4096	5.4380e+31	< 2e-16 ***
F20	1	16384	16384	2.1752e+32	< 2e-16 ***
Residuals	54	0	0		

```
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
par(mfrow=c(2,3))
plot(A20,cc$residuals,xlab="Factor A") # No interaction terms
abline(h=0)
plot(B20,cc$residuals,xlab="Factor B",ylab="")
abline(h=0)
plot(C20,cc$residuals,xlab="Factor C",ylab="")
abline(h=0)
plot(D20,cc$residuals,xlab="Factor D")
abline(h=0)
plot(E20,cc$residuals,xlab="Factor E",ylab="")
```

```
abline(h=0)
plot(F20, cc$residuals, xlab="Factor F", ylab="")
abline(h=0)
```



```
par(mfrow=c(1,2))
qqnorm(cc$residuals)
qqline(cc$residuals)
plot(cc$fitted.values, cc$residuals)
abline(h=0)
```

Normal Q-Q Plot

