Hwang_Hw5

Charles Hwang

11/26/2021

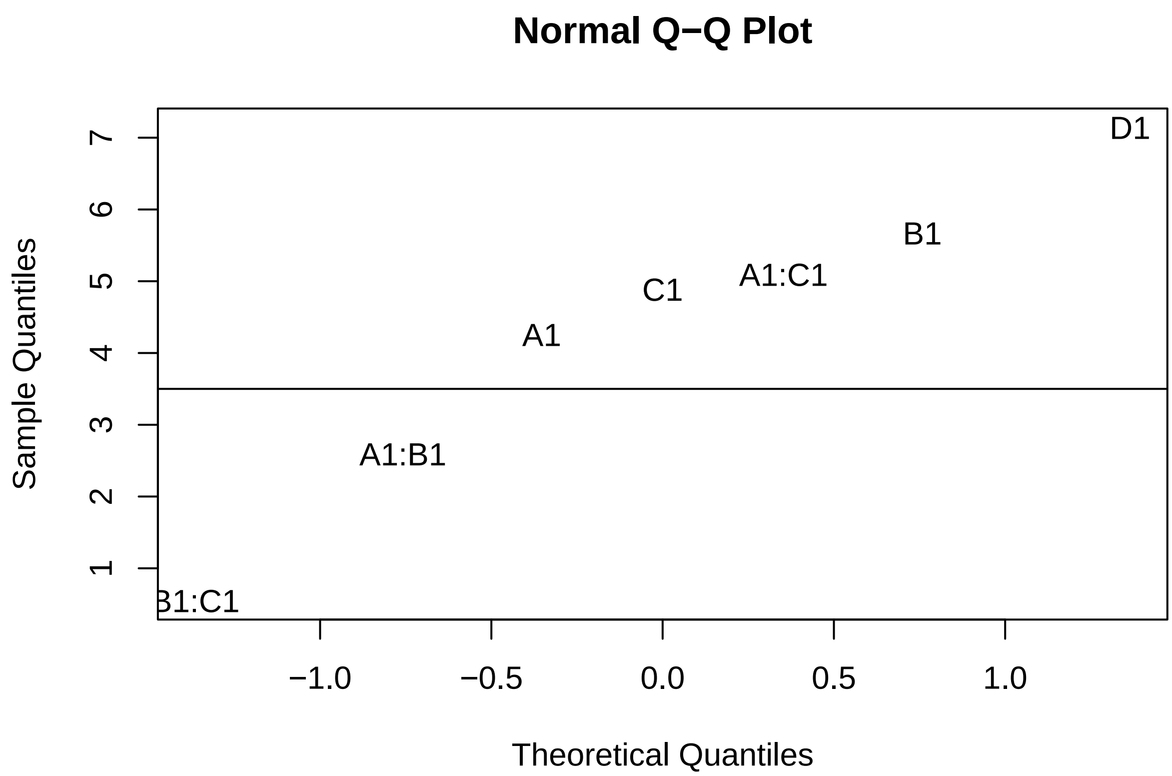
Problem 8.2

```r
rm(list=ls())
A<-as.factor(rep(c(-1,1),4))
B<-as.factor(rep(c(-1,1),2,each=2))
C<-as.factor(rep(c(-1,1),each=4))
D<-as.factor(as.numeric(as.character(A))*as.numeric(as.character(B))*as.numeric(as.character(C)))
length<-c(7.037,16.867,13.876,17.273,11.846,4.368,9.36,15.653) # (1), ad, bd, ab, cd, ac, bc, abcd
l<-lm(length~A*B*C*D)
anova(l)
```

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

```
## Analysis of Variance Table
##
## Response: length
##           Df Sum Sq Mean Sq F value Pr(>F)
## A          1 18.126  18.126     NaN    NaN
## B          1 32.176  32.176     NaN    NaN
## C          1 23.895  23.895     NaN    NaN
## D          1 51.025  51.025     NaN    NaN
## A:B        1  6.731   6.731     NaN    NaN
## A:C        1 25.963  25.963     NaN    NaN
## B:C        1  0.302   0.302     NaN    NaN
## Residuals  0  0.000     NaN
```

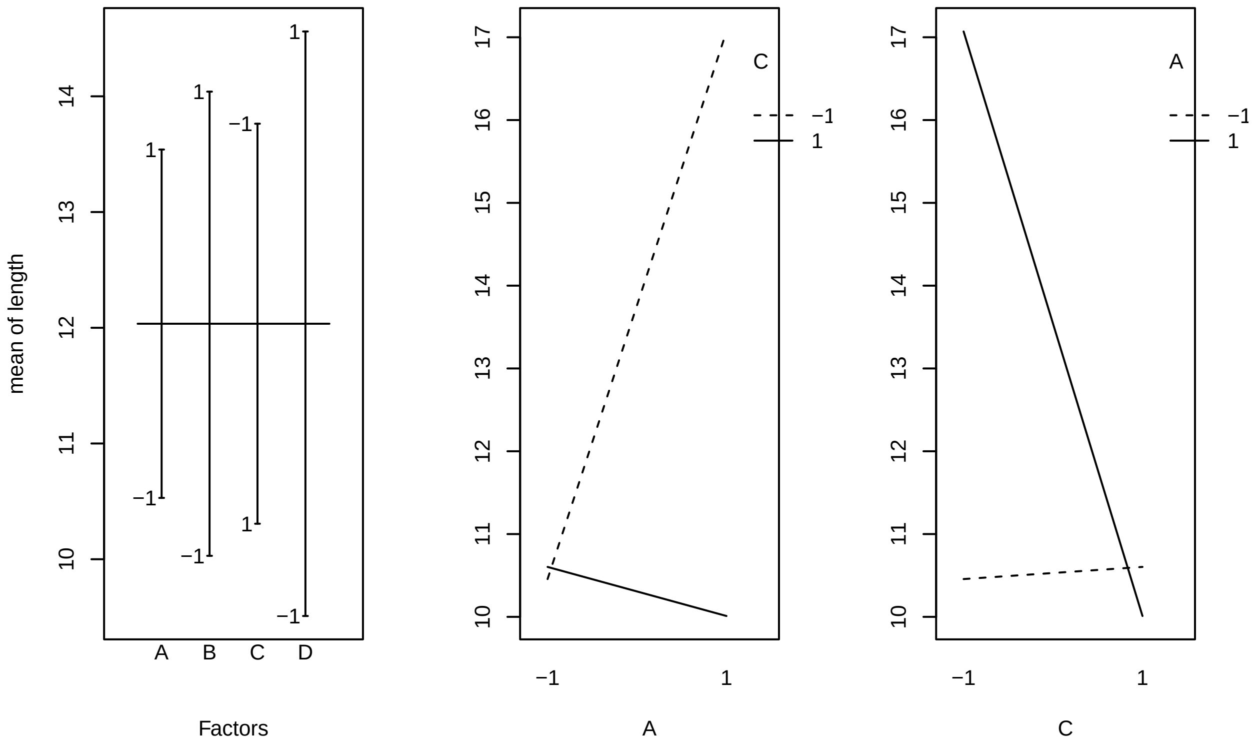
```r
qq<-qqnorm(abs(l$effects[-1]),type="n") # Remove variables
text(qq$x,qq$y,labels=names(abs(l$effects[-1])))
abline(h=3.5) # Arbitrary cutoff
```




```
m<-lm(length~A+B+C+D+A*C) # New model
anova(m)
```

```
## Analysis of Variance Table
##
## Response: length
##           Df Sum Sq Mean Sq F value  Pr(>F)
## A          1 18.126  18.126  5.1549 0.15119
## B          1 32.176  32.176  9.1505 0.09411 .
## C          1 23.895  23.895  6.7954 0.12102
## D          1 51.025  51.025 14.5110 0.06252 .
## A:C        1 25.963  25.963  7.3836 0.11295
## Residuals  2  7.033   3.516
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
par(mfrow=c(1,3))
plot.design(data.frame(A,B,C,D,length))
interaction.plot(A,C,length,ylab="")
interaction.plot(C,A,length,ylab="")
```




```
# The interaction plots indicate there is interaction between pouring temperature (A)
# and heat treatment method (C). However, the ANOVA output shows the interaction is
# not significant at the alpha = 0.05 level (p = 0.11295). Because of this discrepancy,
# we should still proceed in creating and running a reduced model without the interaction
n<-lm(length~A+B+C+D)      # term to see the differences in ANOVA output and results.
anova(n)
```

```
## Analysis of Variance Table
##
## Response: length
##           Df Sum Sq Mean Sq F value Pr(>F)
## A          1 18.126  18.126  1.6480 0.2894
## B          1 32.176  32.176  2.9255 0.1857
## C          1 23.895  23.895  2.1725 0.2369
## D          1 51.025  51.025  4.6392 0.1203
## Residuals  3 32.996  10.999
```

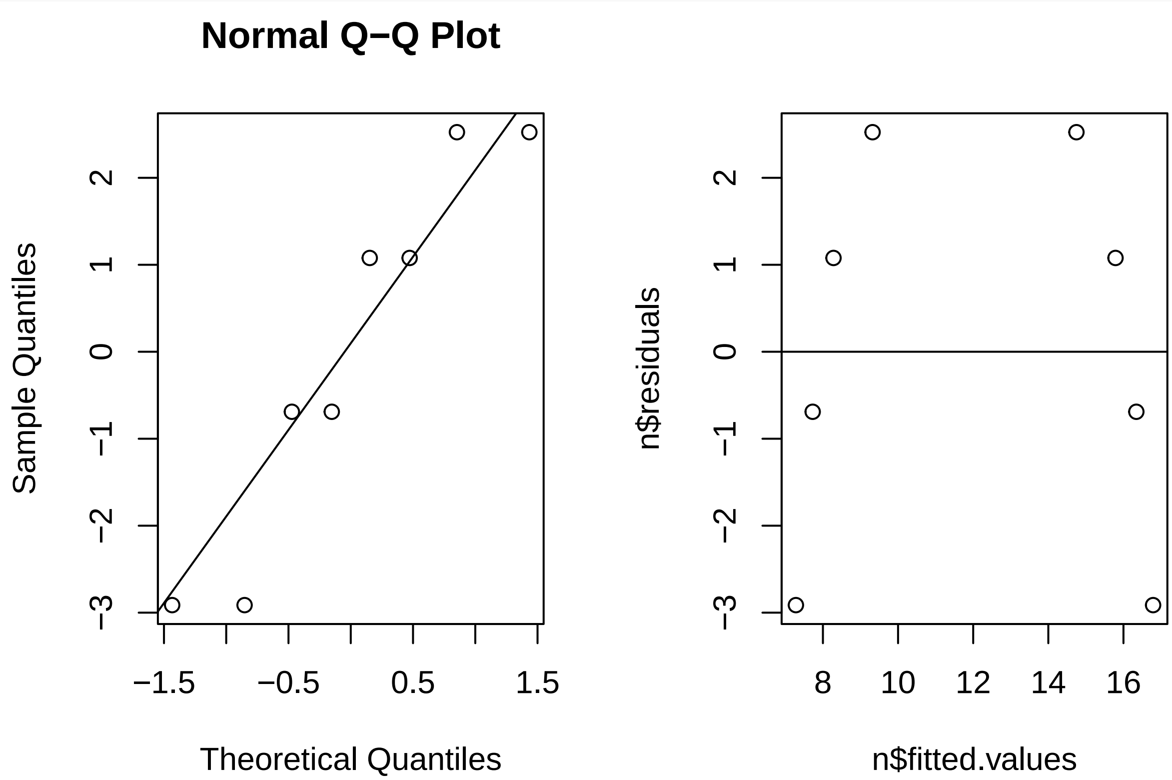
```
shapiro.test(n$residuals)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  n$residuals
## W = 0.89014, p-value = 0.2347
```

```
bartlett.test(n$residuals~A)
```

```
##
##  Bartlett test of homogeneity of variances
##
## data:  n$residuals by A
## Bartlett's K-squared = 0, df = 1, p-value = 1
```

```
bartlett.test(n$residuals~B)

##
##  Bartlett test of homogeneity of variances
##
## data:  n$residuals by B
## Bartlett's K-squared = 1.5226e-15, df = 1, p-value = 1

bartlett.test(n$residuals~C)

##
##  Bartlett test of homogeneity of variances
##
## data:  n$residuals by C
## Bartlett's K-squared = 1.5226e-15, df = 1, p-value = 1

bartlett.test(n$residuals~D)

##
##  Bartlett test of homogeneity of variances
##
## data:  n$residuals by D
## Bartlett's K-squared = 0, df = 1, p-value = 1

# None of the null hypotheses were rejected, so the normality and
par(mfrow=c(1,2))  # equal variance assumptions appear to be met.
qqnorm(n$residuals)
qqline(n$residuals)
plot(n$fitted.values,n$residuals)
abline(h=0)
```



```
# There appears to be some variation in the Q-Q plot. However, it is hard to accurately
# analyze and interpret the plots due to the small sample size.
```

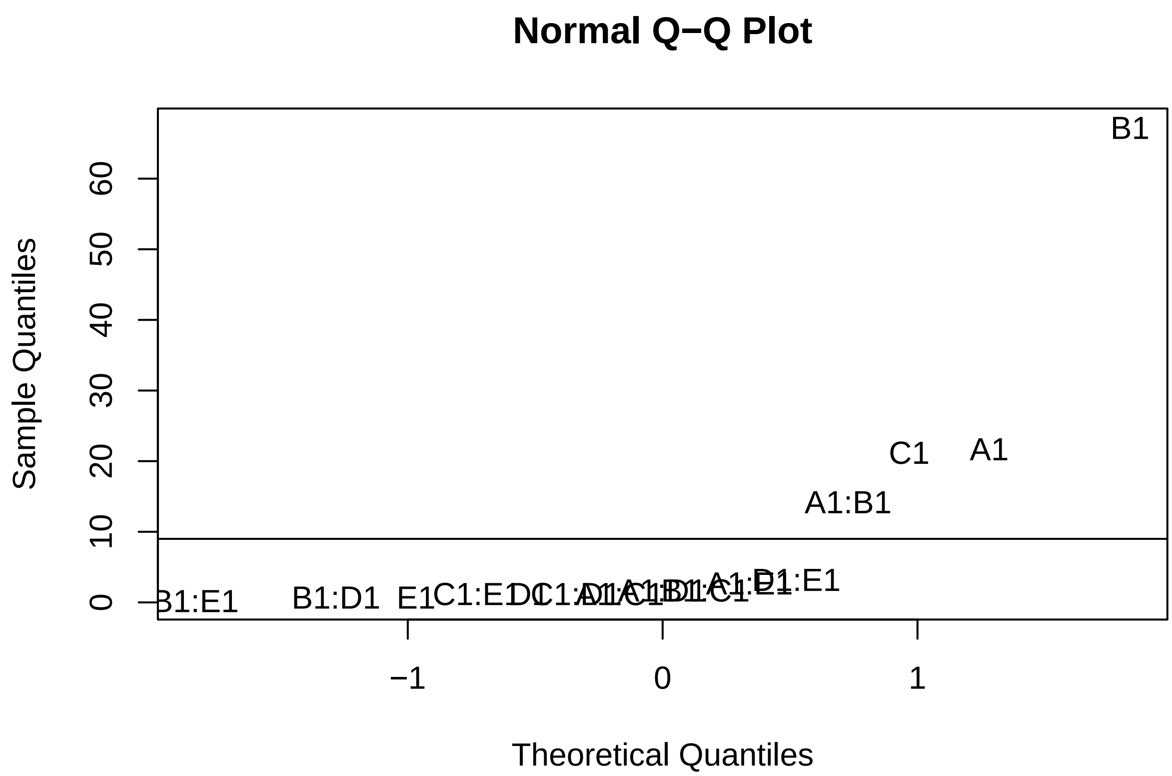
Problem 8.15

```
rm(list=ls()) # See Example 8.2 on pages 327-330
A<-as.factor(rep(c(-1,1),8))
B<-as.factor(rep(c(-1,1),4,each=2))
C<-as.factor(rep(c(-1,1),2,each=4))
D<-as.factor(rep(c(-1,1),each=8))
E<-as.factor(as.numeric(as.character(A))*as.numeric(as.character(B))*as.numeric(as.character(C))*as.nume
y<-c(8,9,34,52,16,22,45,60,8,10,30,50,15,21,44,63) # e, a, b, abe, c, ace, bce, abc, d,
s<-lm(y~A*B*C*D*E)                                 # ade, bde, abd, cde, acd, bcd, abcde
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)
## A          1  473.1   473.1     NaN    NaN
## B          1 4522.6  4522.6     NaN    NaN
## C          1  451.6   451.6     NaN    NaN
## D          1    1.6     1.6     NaN    NaN
## E          1    0.6     0.6     NaN    NaN
## A:B        1  203.1   203.1     NaN    NaN
## A:C        1    1.6     1.6     NaN    NaN
## B:C        1    3.1     3.1     NaN    NaN
## A:D        1    3.1     3.1     NaN    NaN
## B:D        1    0.6     0.6     NaN    NaN
## C:D        1    1.6     1.6     NaN    NaN
## A:E        1    7.6     7.6     NaN    NaN
## B:E        1    0.1     0.1     NaN    NaN
## C:E        1    1.6     1.6     NaN    NaN
## D:E        1   10.6    10.6     NaN    NaN
## Residuals  0    0.0     NaN
```

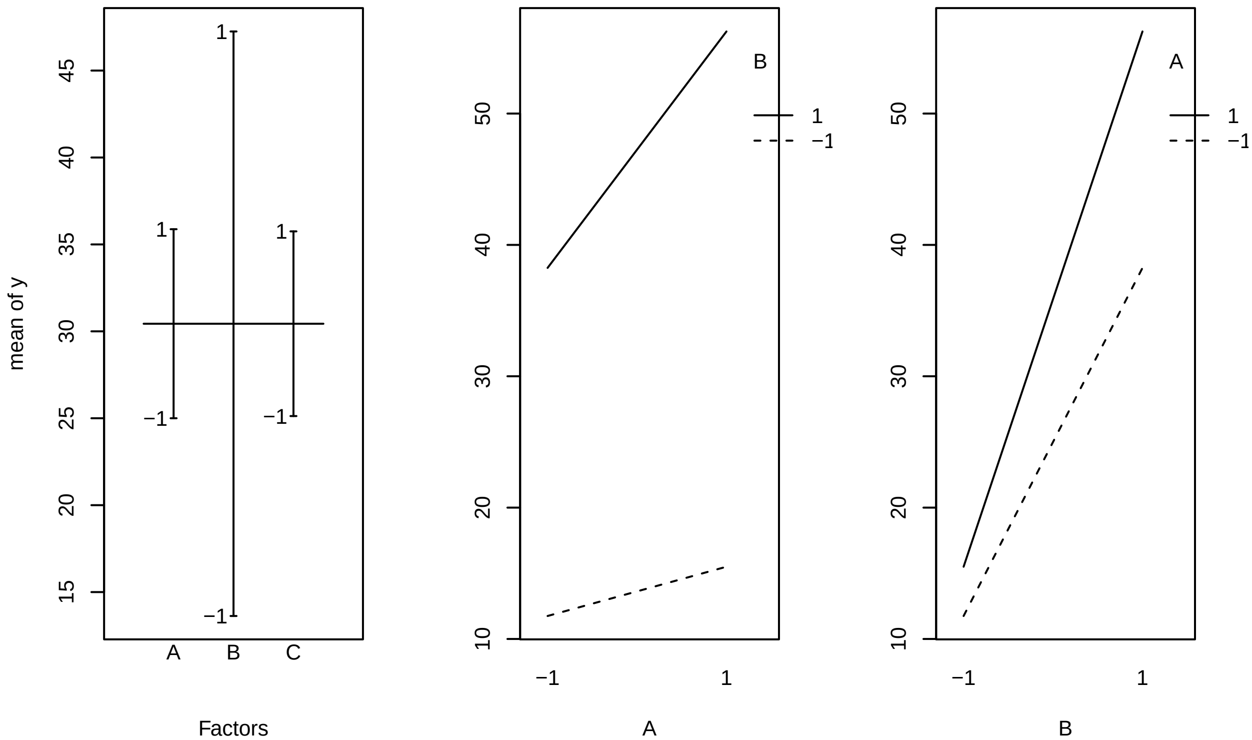
```
qq<-qqnorm(abs(s$effects[-1]),type="n") # Remove variables
text(qq$x,qq$y,labels=names(abs(s$effects[-1])))
abline(h=9) # Arbitrary cutoff
```




```
t<-lm(y~A*B+C) # New model
anova(t)
```

```
## Analysis of Variance Table
##
## Response: y
##           Df Sum Sq Mean Sq  F value    Pr(>F)
## A          1  473.1   473.1  164.219 5.905e-08 ***
## B          1 4522.6  4522.6 1569.963 3.208e-13 ***
## C          1  451.6   451.6  156.755 7.512e-08 ***
## A:B        1  203.1   203.1   70.491 4.115e-06 ***
## Residuals 11   31.7     2.9
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
par(mfrow=c(1,3))
plot.design(data.frame(A,B,C,y))
interaction.plot(A,B,y,ylab="")
interaction.plot(B,A,y,ylab="")
```




```
# There appears to be some interaction between aperture setting (A) and exposure time (B).
# This could mean the two factors are related or associated with each other.
shapiro.test(t$residuals)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  t$residuals
## W = 0.8996, p-value = 0.07919
```

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

```
##
##  Bartlett test of homogeneity of variances
##
## data:  t$residuals by A
## Bartlett's K-squared = 1.1788, df = 1, p-value = 0.2776
```

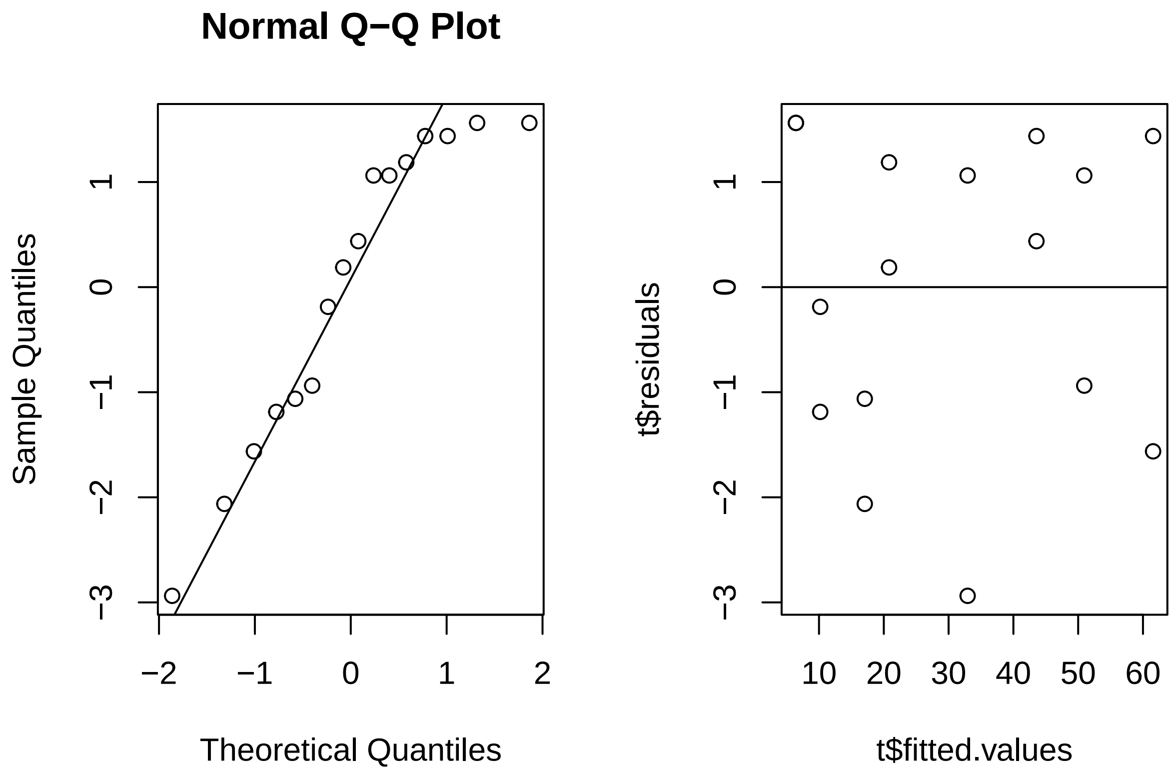
```
bartlett.test(t$residuals~B)
```

```
##
##  Bartlett test of homogeneity of variances
##
## data:  t$residuals by B
## Bartlett's K-squared = 0.19074, df = 1, p-value = 0.6623
```

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

```
##
##  Bartlett test of homogeneity of variances
##
## data:  t$residuals by C
## Bartlett's K-squared = 0.13311, df = 1, p-value = 0.7152
```

```
# We should exercise some caution here as the null hypothesis of the Shapiro-Wilk test is
# close to being rejected at the alpha = 0.05 level (p = 0.07919). The assumption of
par(mfrow=c(1,2))                                    # normality may be slightly violated.
qqnorm(t$residuals)
qqline(t$residuals)
plot(t$fitted.values,t$residuals)
abline(h=0)
```




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

Problem 13.2

(a)

```
rm(list=ls())
part<-as.factor(rep(1:10,each=9))
ins<-as.factor(rep(1:3,10,each=3))
ti<-c(37,38,37,41,41,40,41,42,41,42,41,43,42,42,42,43,42,43,30,31,31,31,31,31,29,30,28,42,43,42,43,43,4
c<-lm(ti~part*ins)
anova(c)
```

```
## Analysis of Variance Table
##
## Response: ti
##            Df Sum Sq Mean Sq  F value    Pr(>F)
## part        9 3936.0  437.33 855.6425 < 2.2e-16 ***
## ins         2   39.3   19.63  38.4130 1.817e-11 ***
## part:ins   18   48.5    2.70   5.2729 5.060e-07 ***
## Residuals 60   30.7    0.51
## ---
```

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

We can see the p-value for the interaction term is significant at the $\alpha = 0.05$ level (p < 0.000001).

(b)

```
MSA<-anova(c)[1,3]
MSB<-anova(c)[2,3]
MSAB<-anova(c)[3,3]
MSE<-anova(c)[4,3]
a<-nlevels(part)
b<-nlevels(ins)
n<-length(ti)/a/b
fp<-data.frame(c(MSA/MSAB,1-pf(MSA/MSAB,a-1,(a-1)*(b-1))),c(MSB/MSAB,1-pf(MSB/MSAB,b-1,(a-1)*(b-1))),c(M
names(fp)<-c("A","B","AB")
row.names(fp)<-(c("F","p"))
fp["F",]<-format(fp["F",],scientific=FALSE)
fp
```

```
##                      A                   B                    AB
## F             162.2703            7.284929              5.272947
## p 2.22044604925031e-15 0.00480960887996595 5.06009005918351e-07
```

```
est<-data.frame(c(MSE),c((MSA-MSE)/(b*n)),c((MSB-MSE)/(a*n)),c((MSAB-MSE)/n),c(MSE+(MSB-MSE)/(a*n)))
names(est)<-c("sigma^2","sigma^2_tau","sigma^2_beta","sigma^2_taubeta","sigma^2_gauge")
row.names(est)<-("Estimate")
est
```

```
##            sigma^2 sigma^2_tau sigma^2_beta sigma^2_taubeta sigma^2_gauge
## Estimate 0.5111111    48.53525    0.6374074       0.7279835      1.148519
```

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

```
## 
##  Shapiro-Wilk normality test
## 
## data:  c$residuals
## W = 0.9551, p-value = 0.0035
```

```
bartlett.test(c$residuals~part)
```

```
## 
##  Bartlett test of homogeneity of variances
## 
## data:  c$residuals by part
## Bartlett's K-squared = 13.15, df = 9, p-value = 0.1559
```

```
bartlett.test(c$residuals~ins)
```

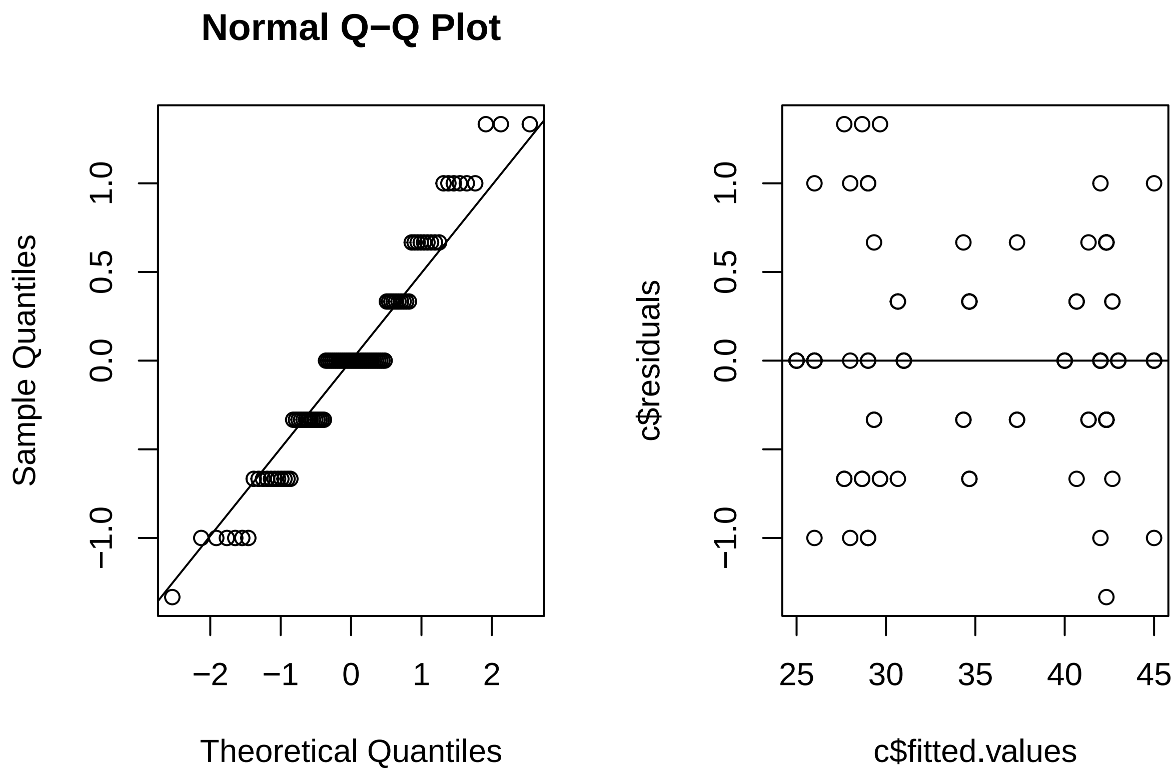
```
## 
##  Bartlett test of homogeneity of variances
## 
## data:  c$residuals by ins
## Bartlett's K-squared = 3.1022, df = 2, p-value = 0.212
```

```
library(car)
leveneTest(c)
```

```
## Levene's Test for Homogeneity of Variance (center = median)
##       Df F value Pr(>F)
```

```
## group 29  0.5436 0.9626
##         60
```

```
# We should exercise caution here as the null hypothesis for the Shapiro-Wilk test has
# been rejected (p = 0.0035) and the assumption of normality may be violated.
par(mfrow=c(1,2))
qqnorm(c$residuals)
qqline(c$residuals)
plot(c$fitted.values,c$residuals)
abline(h=0)
```




```
# There is some variation in the residuals vs. fitted values plot.
```

(c)

```
rm(list=ls()) # https://sakai.luc.edu/portal/directtool/3f2340da-9af7-4db6-8c07-3e89c943e075
```

Problem 13.5

```
temp<-as.factor(rep(c(800,825,850),times=6))
pos<-as.factor(rep(1:2,each=9))
density<-c(570,1063,565,565,1080,510,583,1043,590,528,988,526,547,1026,538,521,1004,532)
d<-lm(density~temp*pos)
anova(d)
```

```
## Analysis of Variance Table
##
## Response: density
##             Df Sum Sq Mean Sq  F value    Pr(>F)
## temp         2 945342  472671 1056.117 3.25e-14 ***
```

```
## pos         1   7160    7160   15.998 0.001762 **
## temp:pos    2    818     409    0.914 0.427110
## Residuals 12   5371     448
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
MSA<-anova(d)[1,3] # We can see the p-value for the interaction term is not
MSB<-anova(d)[2,3] # significant at the alpha = 0.05 level (p = 0.42711).
MSAB<-anova(d)[3,3]
MSE<-anova(d)[4,3]
a<-nlevels(temp)
n<-length(density)/a/nlevels(pos)
fpA<-data.frame(c(MSA/MSAB),c(1-pf(MSA/MSAB,anova(d)[1,1],anova(d)[3,1])))
names(fpA)<-c("F","p")
row.names(fpA)<-("A")
fpA
```

```
##          F            p
## A 1155.518 0.0008646645
```

```
est<-data.frame(c((MSB-MSE)/(a*n)),c((MSAB-MSE)/n))
names(est)<-c("sigma^2_beta","sigma^2_taubeta")
row.names(est)<-("Estimate")
est # Page 576 provides an example with a negative estimate for sigma^2_taubeta.
```

```
##          sigma^2_beta sigma^2_taubeta
## Estimate     745.8333       -12.83333
```

```
e<-lm(density~temp+pos) # Reduced model without interaction term
anova(e)
```

```
## Analysis of Variance Table
##
## Response: density
##           Df Sum Sq Mean Sq  F value    Pr(>F)
## temp       2 945342  472671 1069.257 4.924e-16 ***
## pos        1   7160    7160   16.197  0.001254 **
## Residuals 14   6189     442
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
cat("Reduced model estimate of sigma^2_beta:",(anova(e)[2,3]-anova(e)[3,3])/(a*n),"\n")
```

```
## Reduced model estimate of sigma^2_beta: 746.4444
```

```
# We can see this estimate for sigma^2_beta is very similar to the one in the original model.
shapiro.test(e$residuals)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  e$residuals
## W = 0.91252, p-value = 0.09529
```

```
bartlett.test(e$residuals~temp)
```

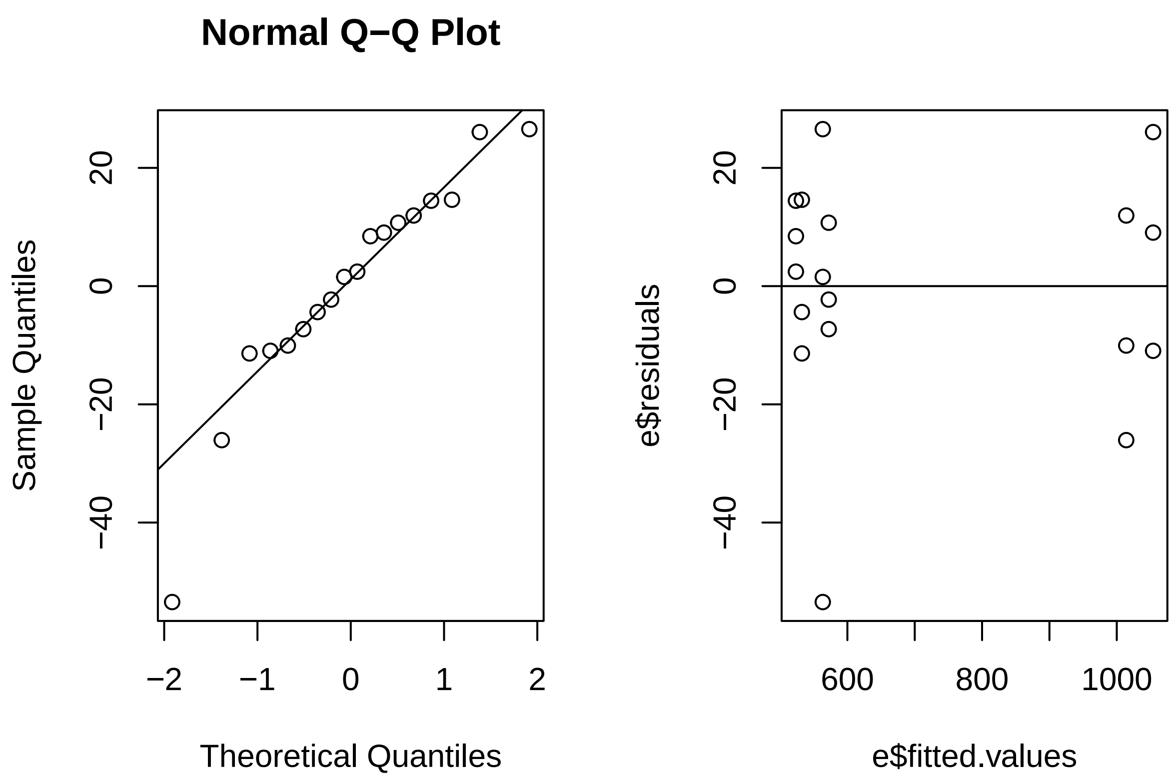
```
##
##  Bartlett test of homogeneity of variances
```

```
##
## data:  e$residuals by temp
## Bartlett's K-squared = 3.9449, df = 2, p-value = 0.1391
```

```r
bartlett.test(e$residuals~pos)
```

```
##
##  Bartlett test of homogeneity of variances
##
## data:  e$residuals by pos
## Bartlett's K-squared = 2.0896, df = 1, p-value = 0.1483
```

```r
# We should exercise some caution here as the null hypothesis of the Shapiro-Wilk test is
# close to being rejected at the alpha = 0.05 level (p = 0.09529). The assumption of
par(mfrow=c(1,2))                                # normality may be slightly violated.
qqnorm(e$residuals)
qqline(e$residuals)
plot(e$fitted.values,e$residuals)
abline(h=0)
```




```r
# There is a slight variation in the Q-Q plot.
```