# Midterm Project

[Midterm Project: Stat 445/545: Analysis of Variance], Stat 445/545, Spring 2023, Behzad FallahiFard, UNM

AUTHOR
Behzad FallahiFard

PUBLISHED
March 26, 2023

```
potato<-read.table("potato.txt",header=TRUE)
n<-nrow(potato)
n
```

```
[1] 75
```

```
potato$regime<-factor(potato$regime,label=c("R","C"))
potato$variety<-factor(potato$variety,label=c("s1","s2"))
potato$temp<-factor(potato$temp,label=c("-4","-8"))
attach(potato)
potato[1:10,]
```

```
   variety regime temp   photo  leak
1       s1      R   -4  4.6386  2.25
2       s1      R   -4  5.7914  4.34
3       s1      R   -4 29.3515  4.25
4       s1      R   -4 18.4173  6.14
5       s1      R   -4  2.2556  2.38
6       s1      R   -8  8.2358 16.30
7       s1      R   -8  8.1972  5.24
8       s1      R   -8 -2.4482  3.25
9       s1      R   -8  5.2820  0.88
10      s1      R   -8 -5.1283  3.98
```

```
tapply(leak,regime,mean)
```

```
        R         C
21.114444  6.575128
```

```
tapply(leak,variety,mean)
```

```
       s1        s2
 6.281429 19.917500
```

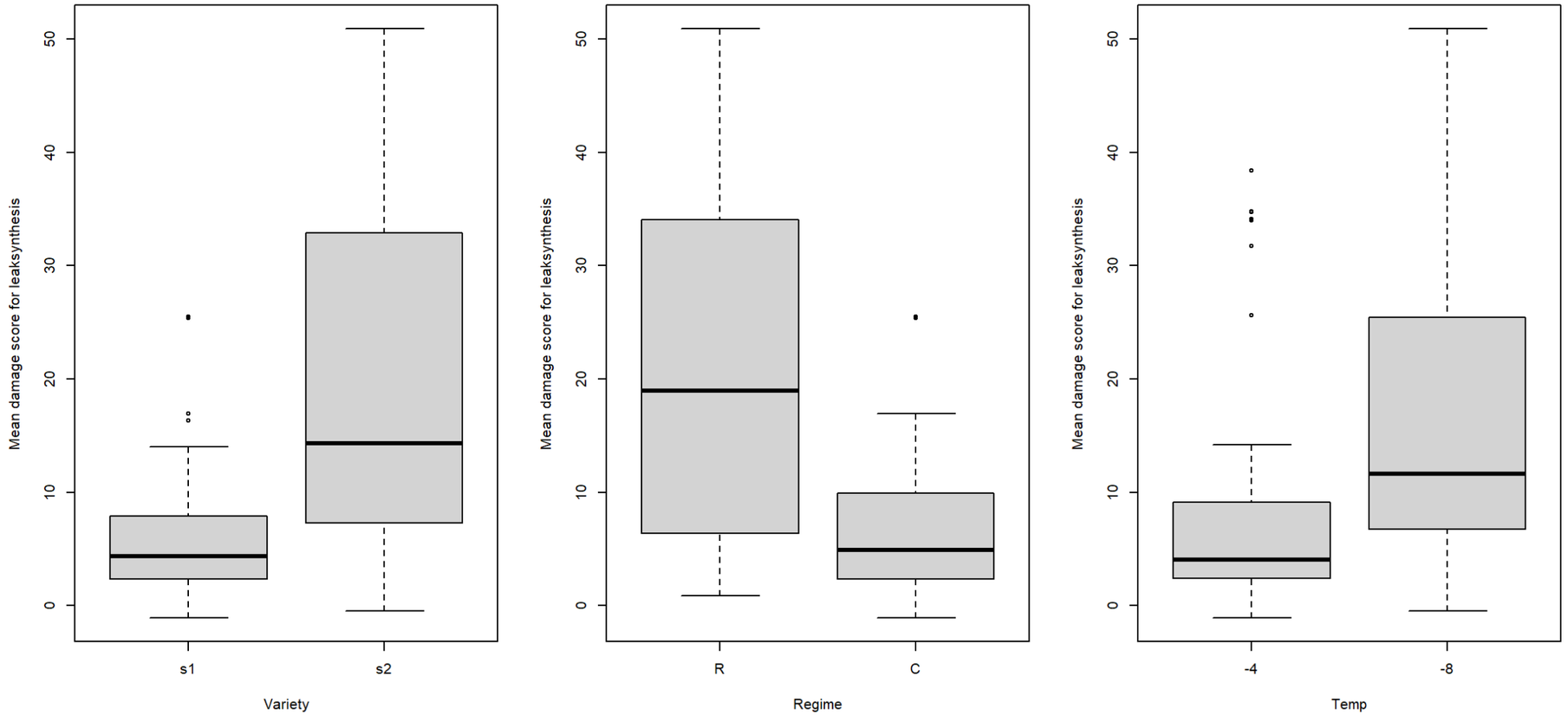
```
tapply(leak,temp,mean)
```

```
       -4        -8
 9.602703 17.401316
```

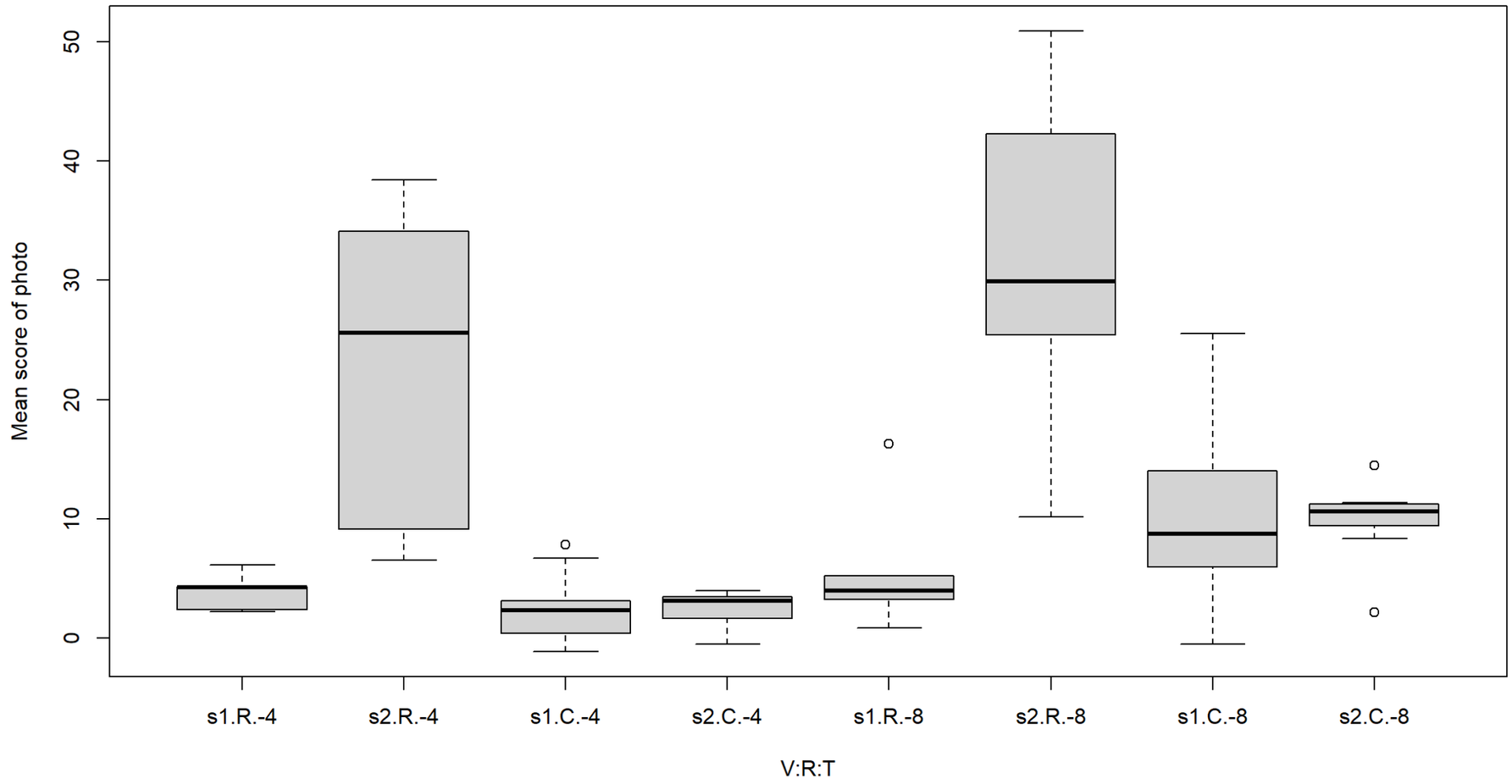
```
aggregate(leak~variety+regime+temp, data=potato, mean)
```

```
  variety regime temp      leak
1      s1      R   -4  3.872000
2      s2      R   -4 22.380000
3      s1      C   -4  2.339167
4      s2      C   -4  2.418571
5      s1      R   -8  5.930000
6      s2      R   -8 32.320769
7      s1      C   -8 10.982308
8      s2      C   -8  9.808571
```

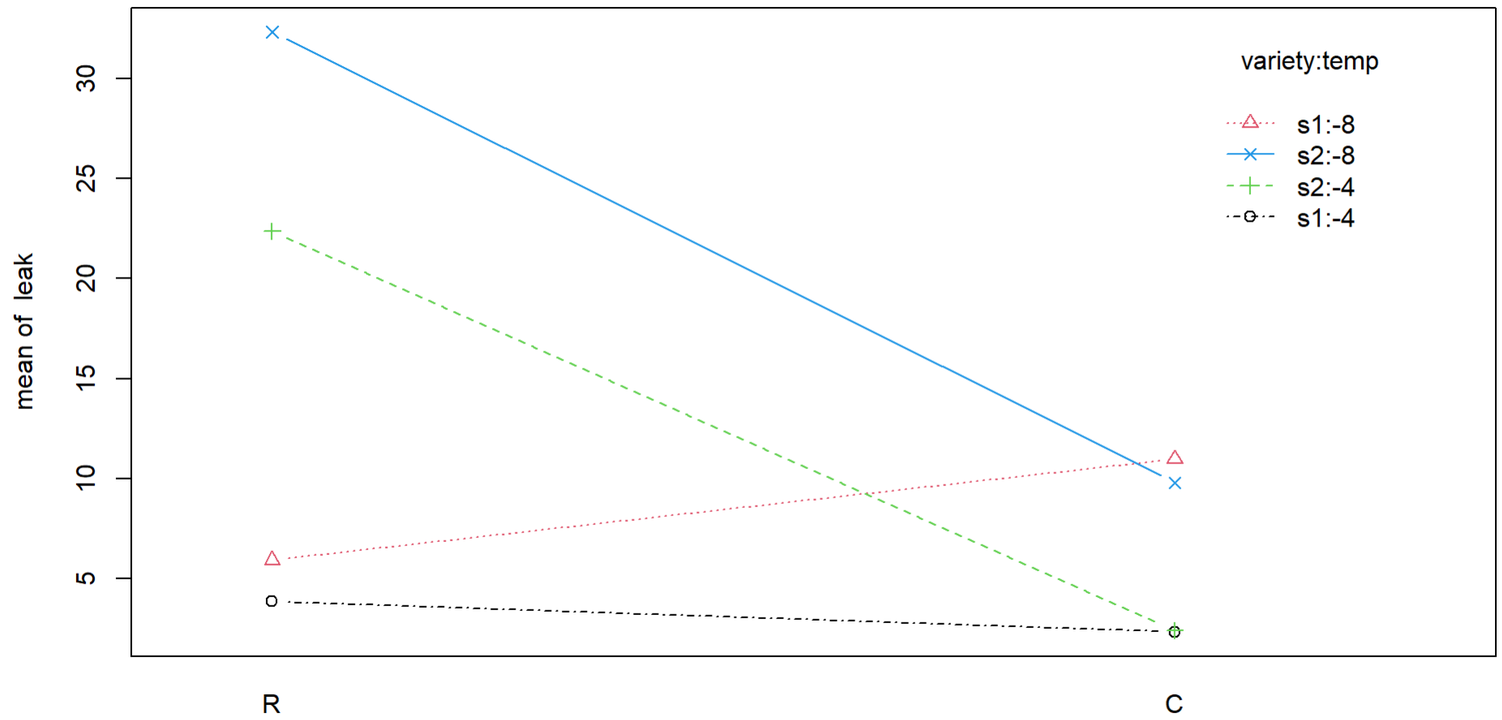
```
#boxplot
par(mfrow=c(1,3))
boxplot(leak ~ variety,
        data = potato,
        xlab = "Variety",
        ylab = "Mean damage score for leaksynthesis")
boxplot(leak ~ regime,
        data = potato,
        xlab = "Regime",
        ylab = "Mean damage score for leaksynthesis")
boxplot(leak ~ temp,
        data = potato,
        xlab = "Temp",
        ylab = "Mean damage score for leaksynthesis")
```



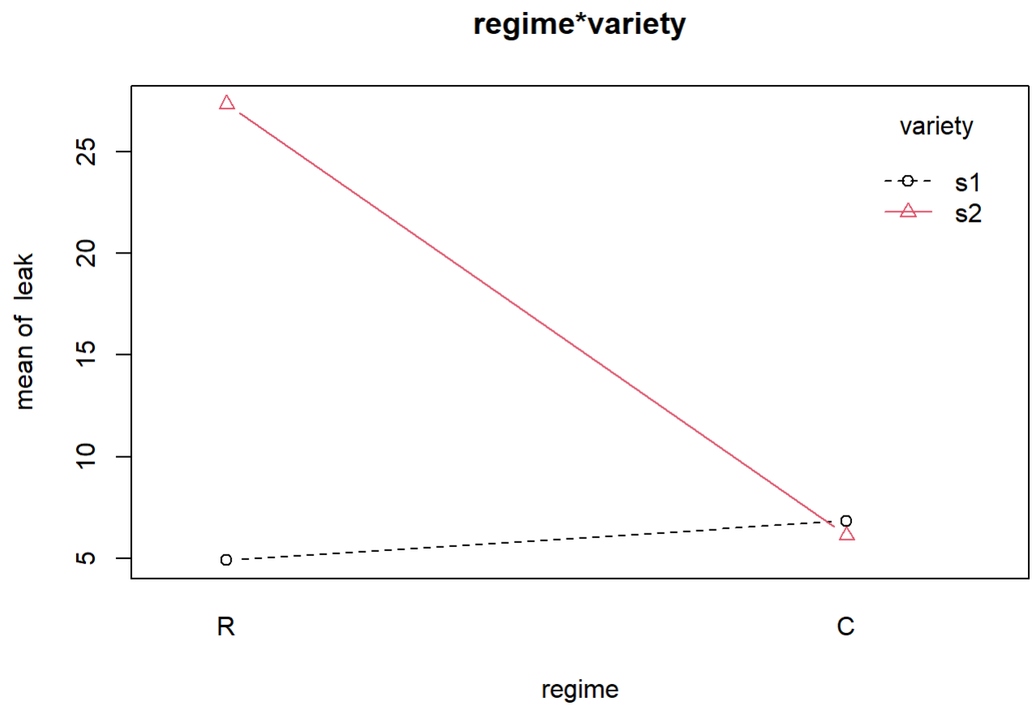

```
par(mfrow=c(1,1))
boxplot(leak ~ variety:regime:temp,
data = potato,
xlab = "V:R:T",
ylab = "Mean score of photo")
```




```
interaction.plot(regime,variety:temp,leak,type='b',
col=1:4, pch=1:4)
```



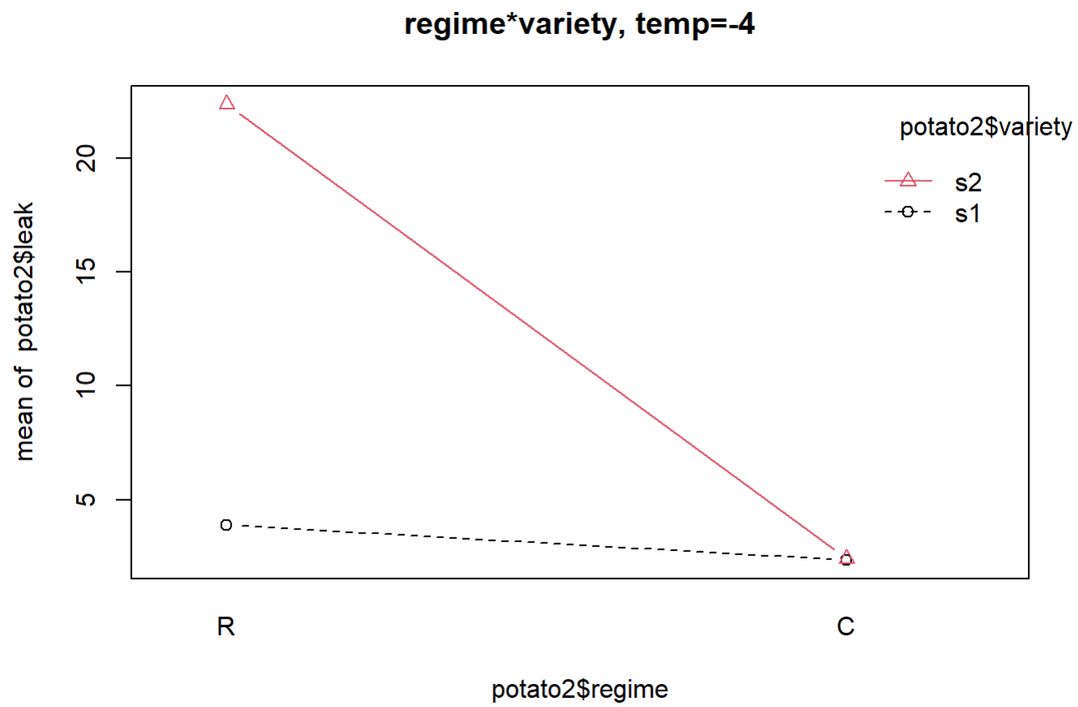

```
interaction.plot(regime,variety,leak,type='b',
col=1:2, pch=1:2,main="regime*variety")
```




```
potato2<-potato[temp==-4,]
nrow(potato2)
```

```
[1] 37
```

```
interaction.plot(potato2$regime,potato2$variety,potato2$leak,type='b',
col=1:2, pch=1:2,main="regime*variety, temp=-4")
```




```
##fit full model
myfit<-lm(leak~variety*regime*temp,
contrasts = c(variety=contr.sum, regime=contr.sum,temp=contr.sum))
library(car)
```

```
Loading required package: carData
```

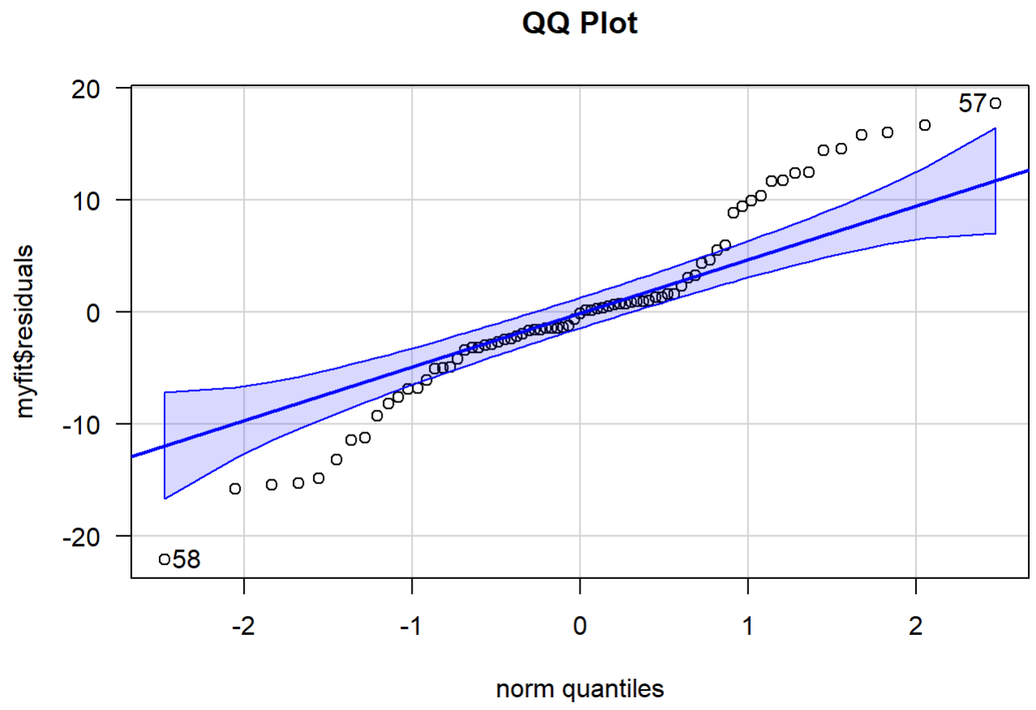
```
Anova(myfit,type=3)
```

```
Anova Table (Type III tests)

Response: leak
                  Sum Sq Df  F value    Pr(>F)
(Intercept)       8110.7  1 107.5408 1.441e-15 ***
variety           1919.2  1  25.4465 3.701e-06 ***
regime            1517.7  1  20.1233 2.921e-05 ***
temp               785.9  1  10.4207   0.00193 **
variety:regime    2115.8  1  28.0529 1.413e-06 ***
variety:temp        44.0  1   0.5829   0.44787
```

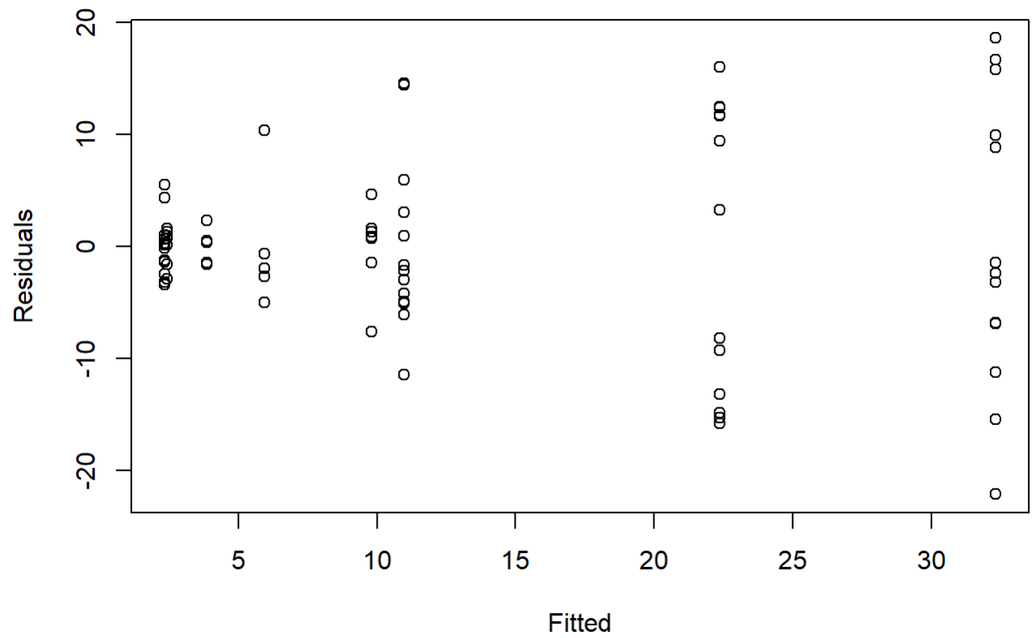
```
regime:temp            16.3  1   0.2158   0.64373
variety:regime:temp    83.5  1   1.1069   0.29654
Residuals            5053.1 67
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
qqPlot(myfit$residuals, las = 1, main="QQ Plot")
```




```
[1] 58 57
```

```
plot(myfit$fitted,myfit$res,xlab="Fitted",ylab="Residuals")
```



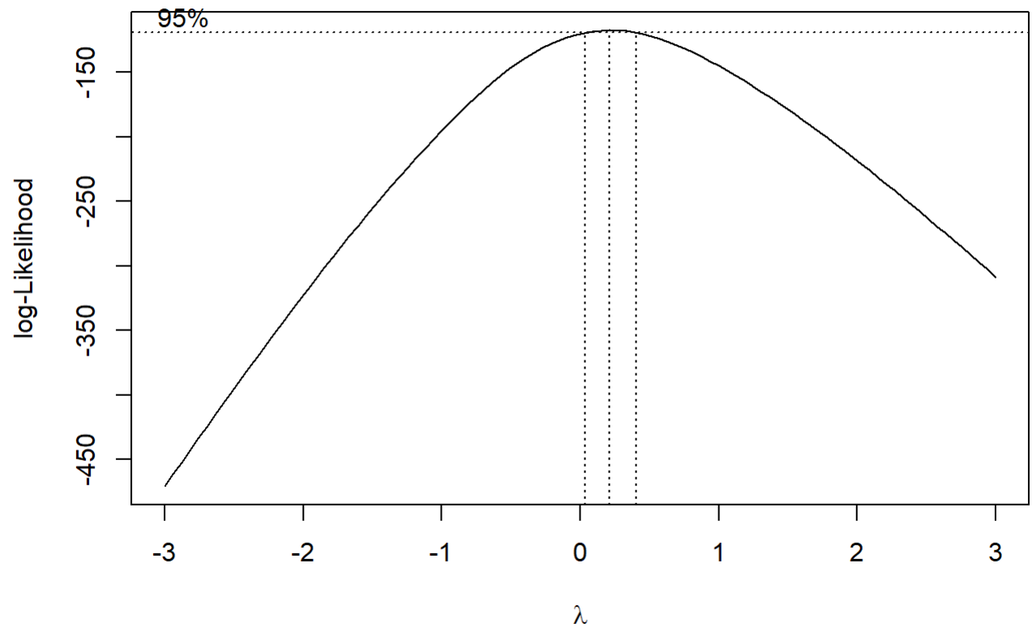

```
outlierTest(myfit)
```

```
No Studentized residuals with Bonferroni p < 0.05
Largest |rstudent|:
    rstudent unadjusted p-value Bonferroni p
58 -2.785452          0.0069699      0.52274
```

```
min(leak) #-8.5655
```

```
[1] -1.11
```

```
leak2<- leak+2 #make response positive
myfit2<-lm(leak2~variety*regime*temp,contrasts = c(variety=contr.sum,
regime=contr.sum,temp=contr.sum))
library(MASS)
par(mfrow=c(1,1))
BC<-boxcox(myfit2, lambda = seq(-3, 3, length = 10))
```




#————————————————————————————

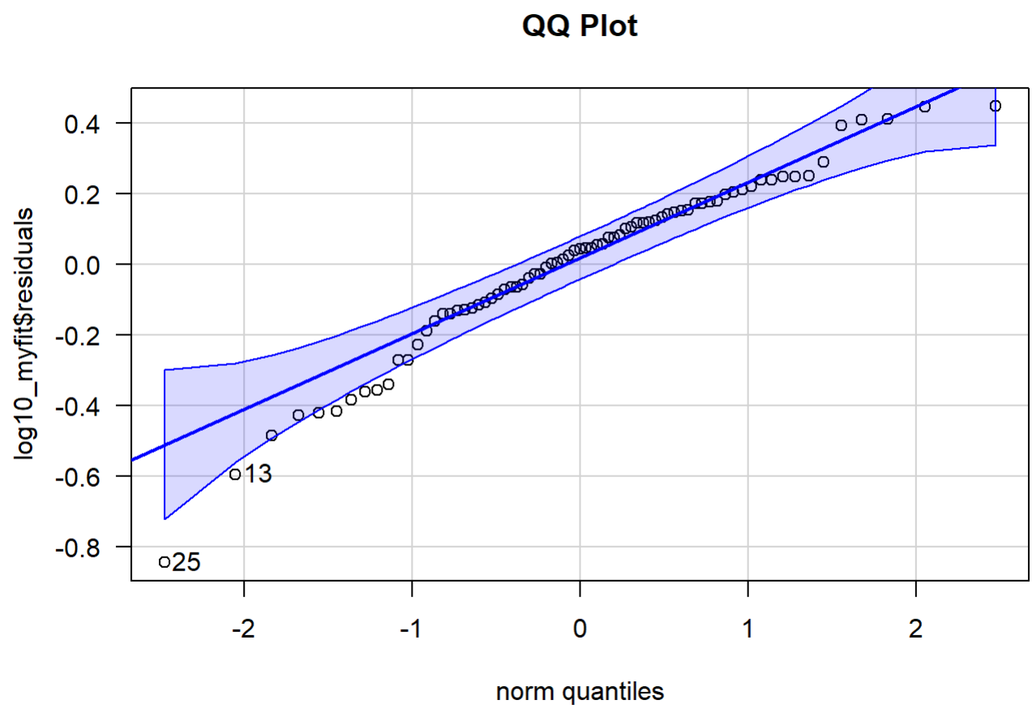
```
##fit full model
log10_myfit<-lm(log10(leak2)~variety*regime*temp,
contrasts = c(variety=contr.sum, regime=contr.sum,temp=contr.sum))
library(car)
Anova(log10_myfit,type=3)
```

```
Anova Table (Type III tests)

Response: log10(leak2)
                   Sum Sq Df  F value    Pr(>F)
(Intercept)        57.929  1 850.0605 < 2.2e-16 ***
variety             1.757  1  25.7774 3.270e-06 ***
regime              1.369  1  20.0921 2.958e-05 ***
temp                1.351  1  19.8309 3.286e-05 ***
variety:regime      1.368  1  20.0795 2.973e-05 ***
variety:temp        0.006  1   0.0915   0.76324
regime:temp         0.452  1   6.6396   0.01218 *
variety:regime:temp 0.029  1   0.4285   0.51498
```

```
Residuals             4.566 67
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
qqPlot(log10_myfit$residuals, las = 1, main="QQ Plot")
```



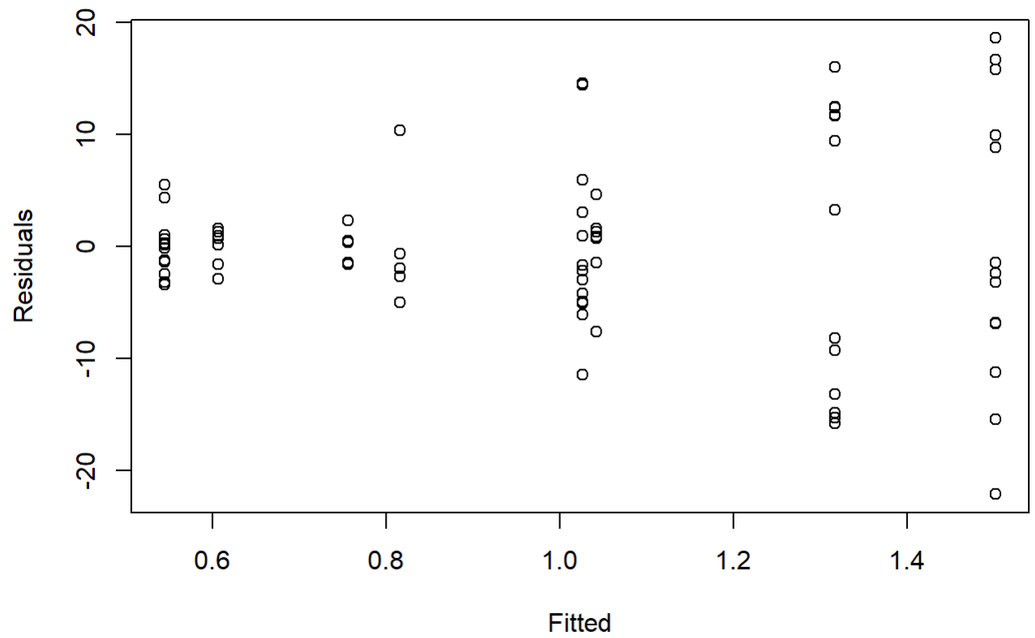

```
[1] 25 13
```

```
shapiro.test(log10_myfit$resid)
```

```
    Shapiro-Wilk normality test

data:  log10_myfit$resid
W = 0.95957, p-value = 0.01729
```

```
plot(log10_myfit$fitted,myfit$res,xlab="Fitted",ylab="Residuals")
```



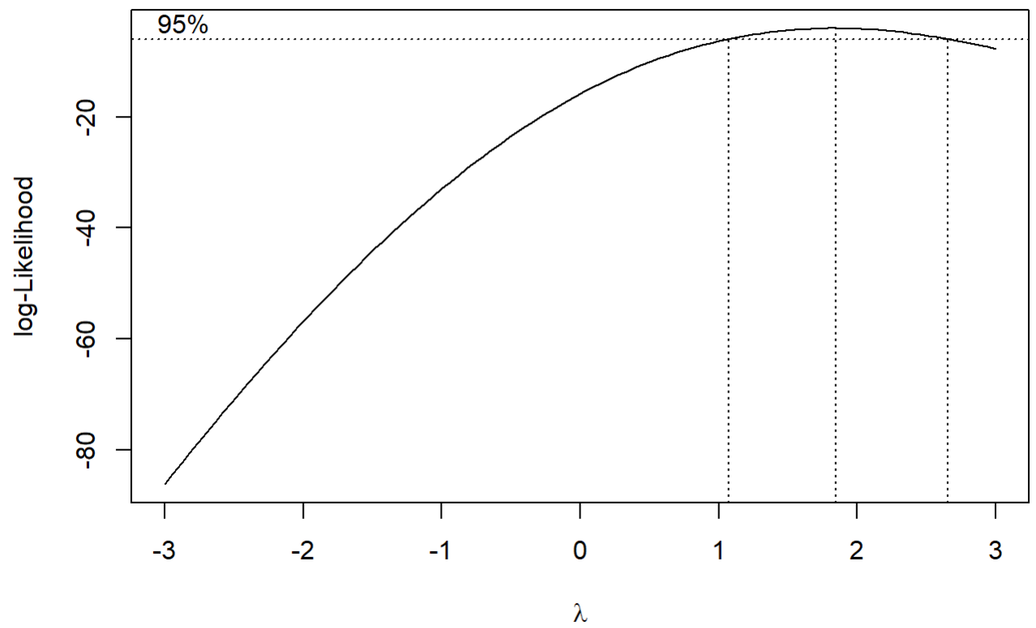

```
outlierTest(log10_myfit)
```

```
    rstudent unadjusted p-value Bonferroni p
25 -3.667084         0.00049135     0.036851
```

```
min(log10(leak2))
```

```
[1] -0.05060999
```

```
logleak21 <- log10(leak2)+1
log10p1_myfit<-lm(logleak21 ~ variety*regime*temp,
contrasts = c(variety=contr.sum, regime=contr.sum,temp=contr.sum))
par(mfrow=c(1,1))
BC2<-boxcox(log10p1_myfit, lambda = seq(-3, 3, length = 10))
```



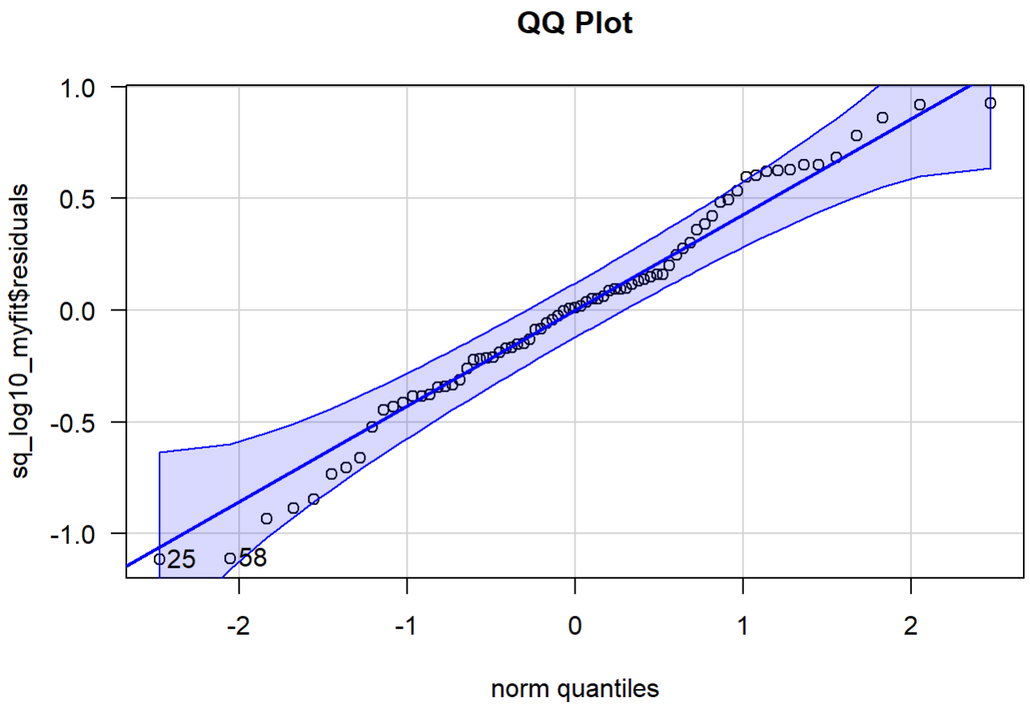

```
##fit full model
#sq_log10_leak <- (log10(leak2)+1)^0.5 + (log10(leak2)+2)^0.5
sq_log10_leak <- (log10(leak2))^2
sq_log10_myfit<-lm(sq_log10_leak~variety*regime*temp,
contrasts = c(variety=contr.sum, regime=contr.sum,temp=contr.sum))
library(car)
Anova(sq_log10_myfit,type=3)
```

```
Anova Table (Type III tests)

Response: sq_log10_leak
                  Sum Sq Df  F value    Pr(>F)
(Intercept)       71.795  1 293.9062 < 2.2e-16 ***
variety            7.642  1  31.2853 4.451e-07 ***
regime             5.473  1  22.4060 1.184e-05 ***
temp               4.436  1  18.1602 6.496e-05 ***
variety:regime     7.757  1  31.7553 3.775e-07 ***
variety:temp       0.082  1   0.3350   0.56467
regime:temp        0.694  1   2.8395   0.09663 .
```

```
variety:regime:temp  0.148  1   0.6051   0.43939
Residuals           16.367 67
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
qqPlot(sq_log10_myfit$residuals, las = 1, main="QQ Plot")
```



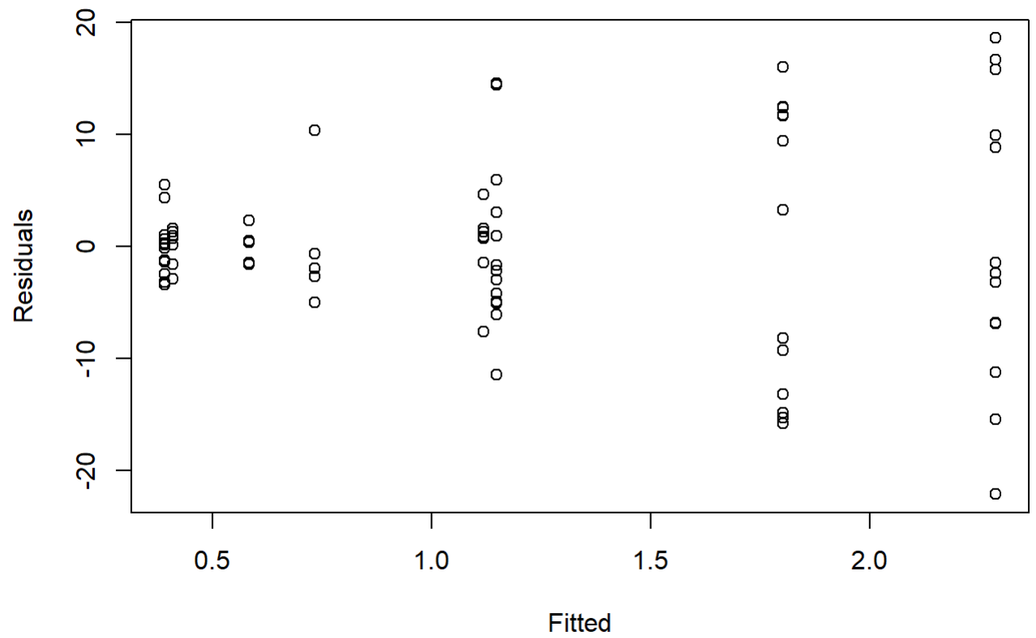

```
[1] 25 58
```

```
shapiro.test(sq_log10_myfit$resid)
```

```
    Shapiro-Wilk normality test

data:  sq_log10_myfit$resid
W = 0.98143, p-value = 0.3392
```

```
plot(sq_log10_myfit$fitted,myfit$res,xlab="Fitted",ylab="Residuals")
```




```
leveneTest(sq_log10_myfit)
```

```
Levene's Test for Homogeneity of Variance (center = median)
      Df F value Pr(>F)
group  7  2.7261  0.015 *
      67
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
outlierTest(sq_log10_myfit)
```

```
No Studentized residuals with Bonferroni p < 0.05
Largest |rstudent|:
    rstudent unadjusted p-value Bonferroni p
25 -2.434653           0.017618           NA
```

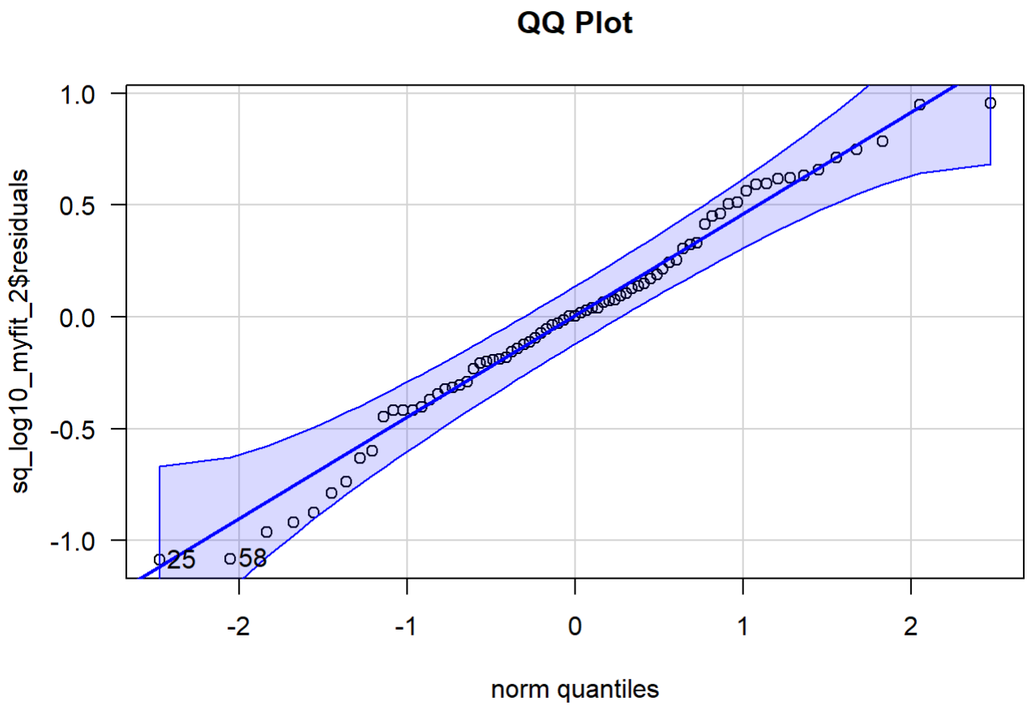
```
sq_log10_myfit_2<-lm(sq_log10_leak~variety+regime+temp+variety:regime+variety:temp +temp:regime,
contrasts = c(variety=contr.sum, regime=contr.sum,temp=contr.sum))
```

```
library(car)
Anova(sq_log10_myfit_2,type=3)
```

```
Anova Table (Type III tests)

Response: sq_log10_leak
               Sum Sq Df  F value    Pr(>F)
(Intercept)    71.840  1 295.8073 < 2.2e-16 ***
variety         7.629  1  31.4132 4.111e-07 ***
regime          5.462  1  22.4903 1.122e-05 ***
temp            5.869  1  24.1676 5.867e-06 ***
variety:regime  7.771  1  31.9987 3.347e-07 ***
variety:temp    0.061  1   0.2492    0.6192
regime:temp     0.628  1   2.5843    0.1126
Residuals      16.515 68
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
qqPlot(sq_log10_myfit_2$residuals, las = 1, main="QQ Plot")
```



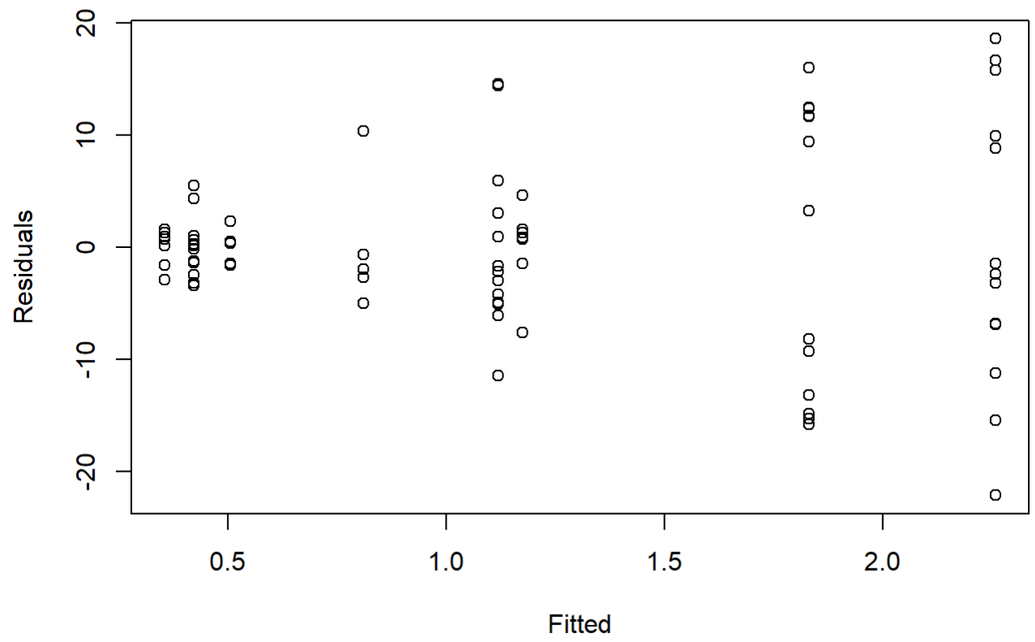
```
[1] 25 58
```

```
shapiro.test(sq_log10_myfit_2$resid)
```

```
    Shapiro-Wilk normality test

data:  sq_log10_myfit_2$resid
W = 0.98236, p-value = 0.3811
```

```
plot(sq_log10_myfit_2$fitted,myfit$res,xlab="Fitted",ylab="Residuals")
```




```
#leveneTest(sq_log10_myfit_2)
outlierTest(sq_log10_myfit_2)
```

```
No Studentized residuals with Bonferroni p < 0.05
Largest |rstudent|:
```

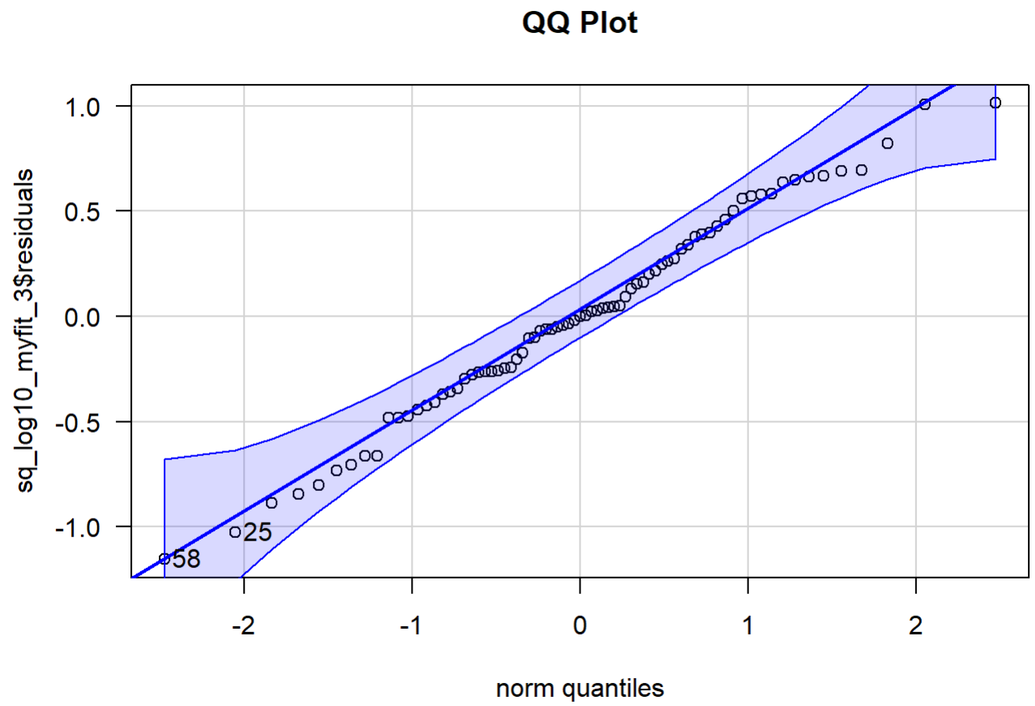
```
    rstudent unadjusted p-value Bonferroni p
25 -2.362637           0.021057           NA
```

```
sq_log10_myfit_3<-lm(sq_log10_leak~variety+regime+temp+variety:regime,
contrasts = c(variety=contr.sum, regime=contr.sum,temp=contr.sum))
library(car)
Anova(sq_log10_myfit_3,type=3)
```

```
Anova Table (Type III tests)

Response: sq_log10_leak
               Sum Sq Df F value    Pr(>F)
(Intercept)    71.933  1 293.704 < 2.2e-16 ***
variety         7.603  1  31.042 4.388e-07 ***
regime          5.440  1  22.210 1.205e-05 ***
temp            6.177  1  25.219 3.750e-06 ***
variety:regime  7.800  1  31.847 3.297e-07 ***
Residuals      17.144 70
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
qqPlot(sq_log10_myfit_3$residuals, las = 1, main="QQ Plot")
```



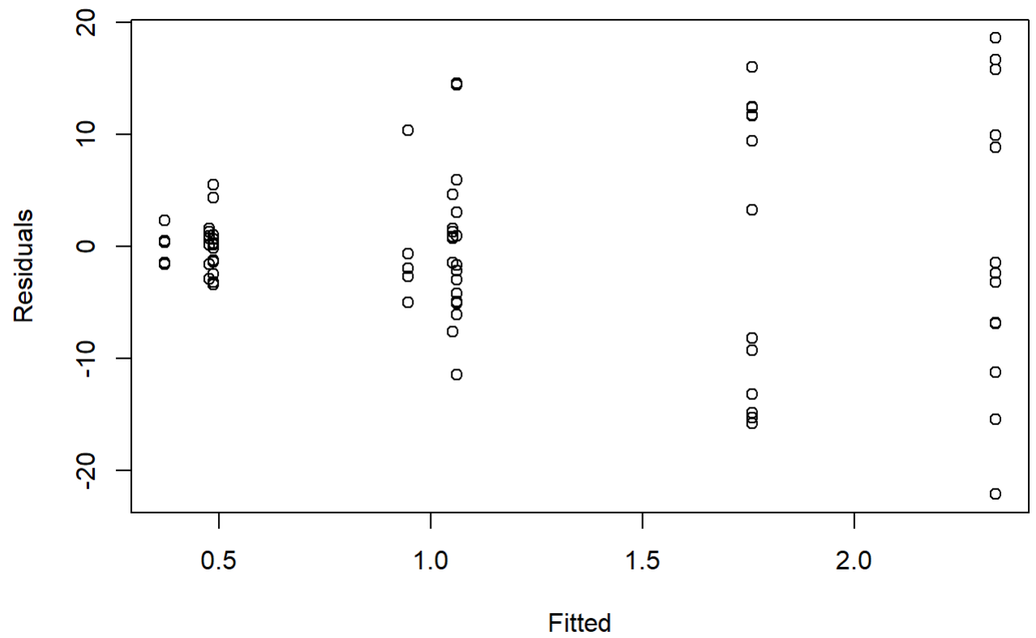

```
[1] 58 25
```

```
shapiro.test(sq_log10_myfit_3$resid)
```

```
    Shapiro-Wilk normality test

data:  sq_log10_myfit_3$resid
W = 0.98941, p-value = 0.7918
```

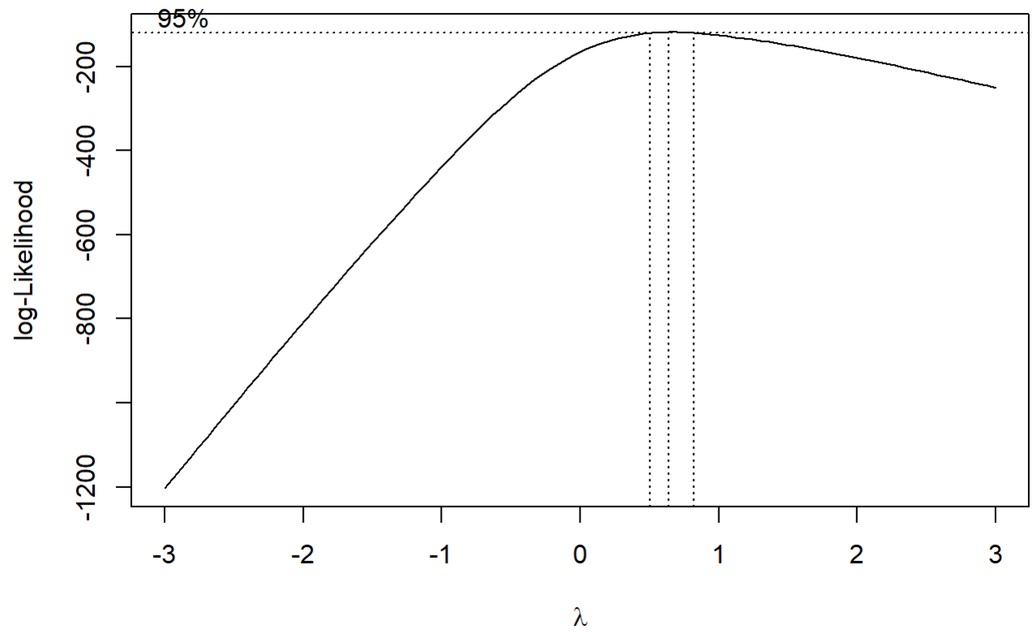
```
plot(sq_log10_myfit_3$fitted,myfit$res,xlab="Fitted",ylab="Residuals")
```




```
#leveneTest(sq_log10_myfit_2)
outlierTest(sq_log10_myfit_3)
```

```
No Studentized residuals with Bonferroni p < 0.05
Largest |rstudent|:
    rstudent unadjusted p-value Bonferroni p
58 -2.482032           0.015502           NA
```

```
BC3<-boxcox(sq_log10_myfit_3, lambda = seq(-3, 3, length = 10))
```

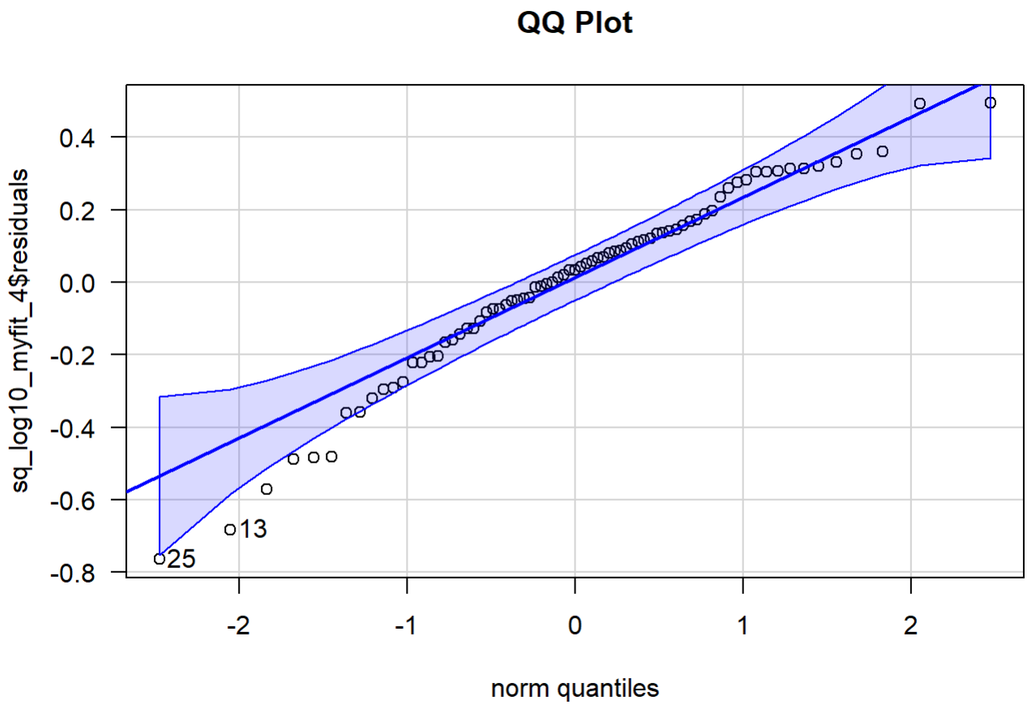



```
sq_log10_myfit_4<-lm(log10(leak2)~variety+regime+temp+variety:regime,
contrasts = c(variety=contr.sum, regime=contr.sum,temp=contr.sum))
library(car)
Anova(sq_log10_myfit_4,type=3)
```

```
Anova Table (Type III tests)

Response: log10(leak2)
               Sum Sq Df F value    Pr(>F)
(Intercept)    58.041  1 802.480 < 2.2e-16 ***
variety         1.739  1  24.046 5.880e-06 ***
regime          1.354  1  18.718 4.938e-05 ***
temp            1.847  1  25.535 3.326e-06 ***
variety:regime  1.384  1  19.140 4.149e-05 ***
Residuals       5.063 70
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
qqPlot(sq_log10_myfit_4$residuals, las = 1, main="QQ Plot")
```



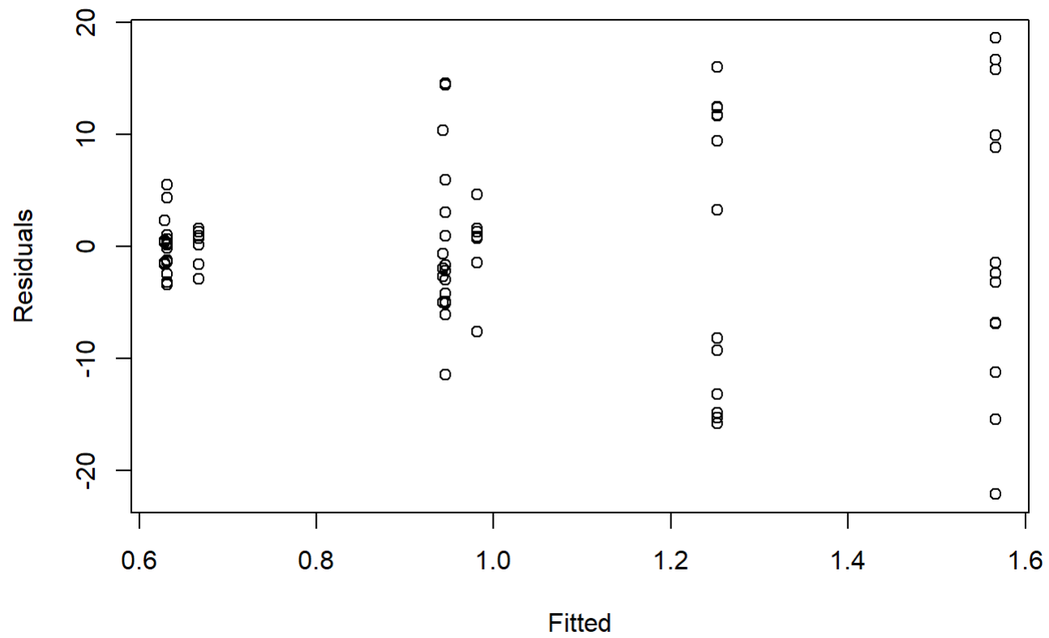

```
[1] 25 13
```

```
shapiro.test(sq_log10_myfit_4$resid)
```

```
    Shapiro-Wilk normality test

data:  sq_log10_myfit_4$resid
W = 0.96617, p-value = 0.04235
```

```
plot(sq_log10_myfit_4$fitted,myfit$res,xlab="Fitted",ylab="Residuals")
```




```
#leveneTest(sq_log10_myfit_2)
outlierTest(sq_log10_myfit_4)
```

```
No Studentized residuals with Bonferroni p < 0.05
Largest |rstudent|:
   rstudent unadjusted p-value Bonferroni p
25 -3.09169          0.0028706      0.21529
```

```
#BC3<-boxcox(sq_log10_myfit_4, lambda = seq(-3, 3, length = 10))
```

```
library(lmtest)
```

```
Loading required package: zoo

Warning: package 'zoo' was built under R version 4.2.2


Attaching package: 'zoo'
```

```
The following objects are masked from 'package:base':

    as.Date, as.Date.numeric
```

```
#perform Breusch-Pagan test
bptest(sq_log10_myfit_4)
```

```
    studentized Breusch-Pagan test

data:  sq_log10_myfit_4
BP = 4.1786, df = 4, p-value = 0.3824
```