Notebook 3 Extra One-Way ANOVA Example (Coagulation)

```
# Extra example for one-way ANOVA - as referred to in Bronwyn Loong's slides

# Might need to reset the library location to access the faraway library
# .libPaths("H:/R")
library(faraway)
data(package="faraway")

coagulation
```

```
##    coag diet
## 1    62    A
## 2    60    A
## 3    63    A
## 4    59    A
## 5    63    B
## 6    67    B
## 7    71    B
## 8    64    B
## 9    65    B
## 10   66    B
## 11   68    C
## 12   66    C
## 13   71    C
## 14   67    C
## 15   68    C
## 16   68    C
## 17   56    D
## 18   62    D
## 19   60    D
## 20   61    D
## 21   63    D
## 22   64    D
## 23   63    D
## 24   59    D
```
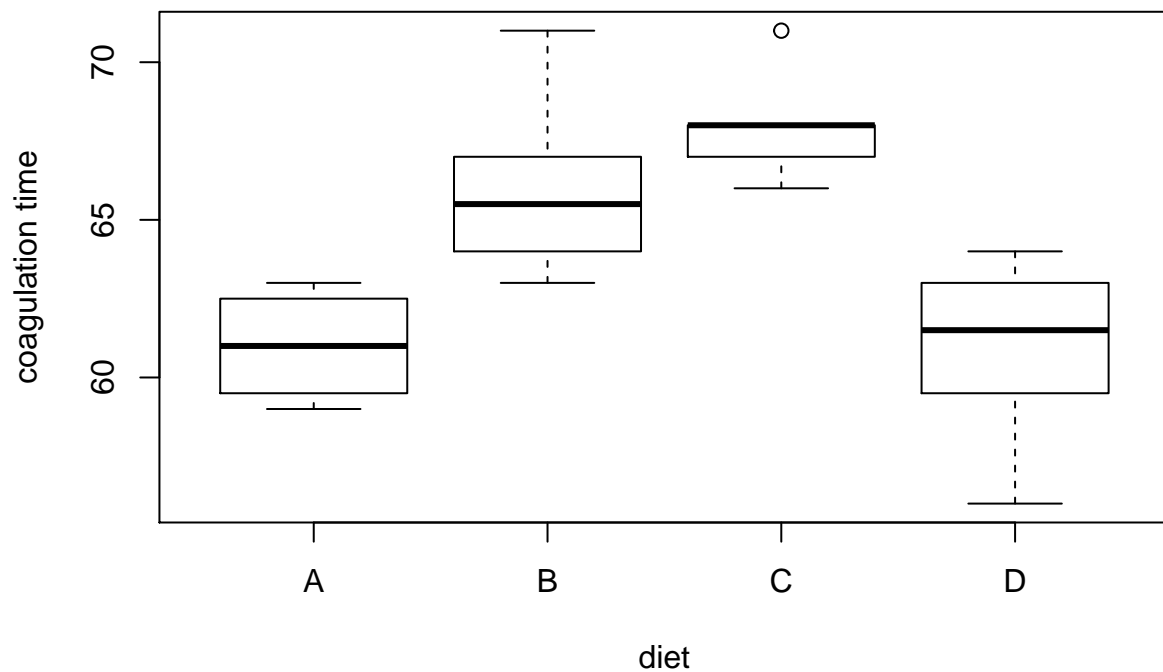
```
help(coagulation)
attach(coagulation)

plot(coag ~ diet, ylab="coagulation time")
```
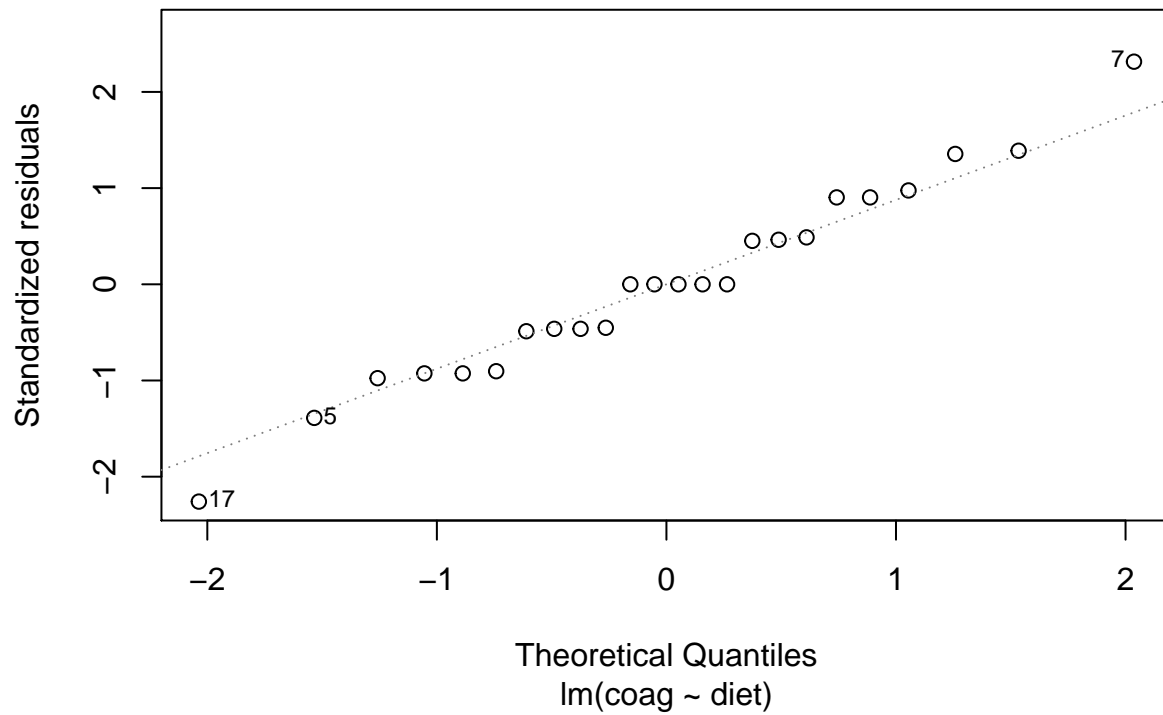
```
lvl.mns <- tapply(coag, diet, mean)
lvl.mns
```

```
##  A  B  C  D
## 61 66 68 61
```

```
# Fit default model - baseline (treatment) constraints (variety of Model II in slides):

g <- lm(coag ~ diet, coagulation)
plot(g)
```

Residuals vs Fitted

lm(coag ~ diet)

Normal Q–Q

lm(coag ~ diet)

Scale–Location

lm(coag ~ diet)

Residuals vs Leverage
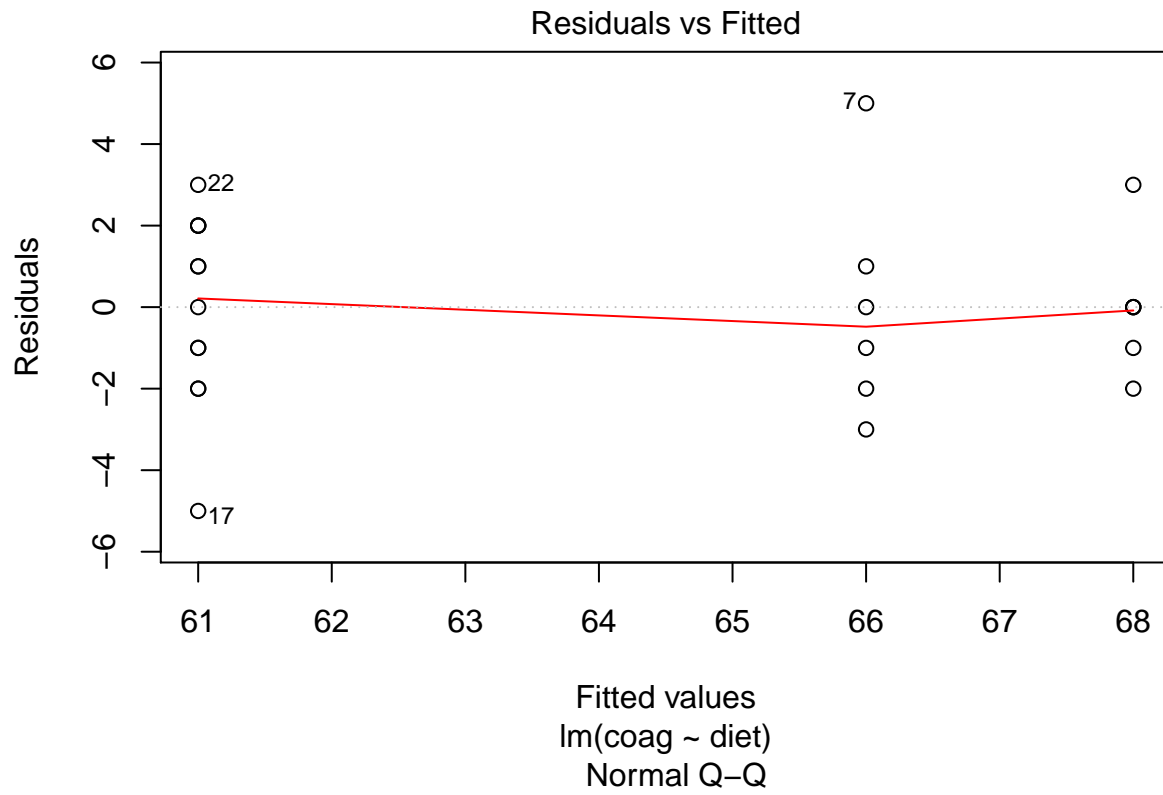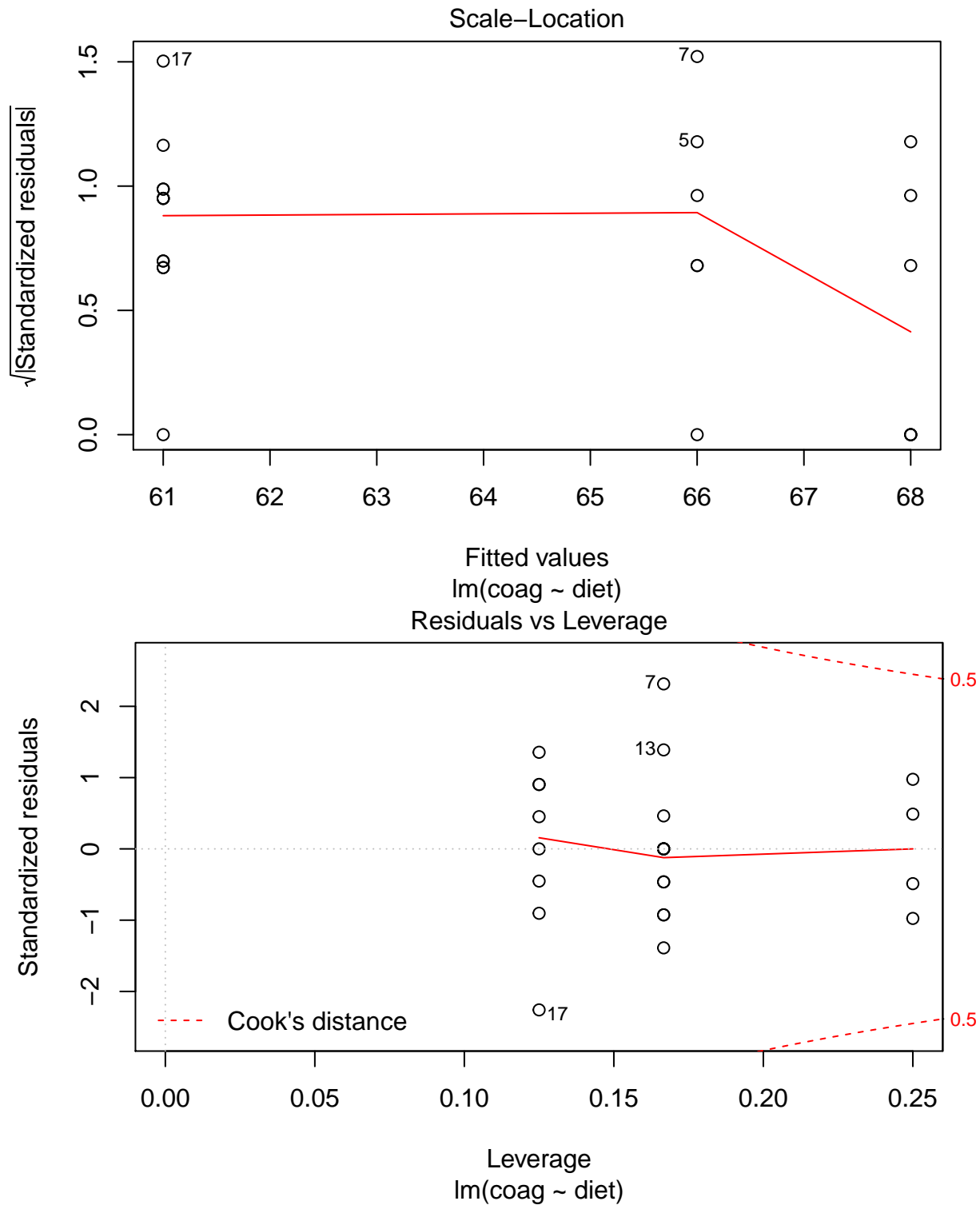
lm(coag ~ diet)

```
summary(g)
```

```
##
## Call:
## lm(formula = coag ~ diet, data = coagulation)
##
```

```
## Residuals:
##     Min     1Q Median     3Q    Max
##   -5.00  -1.25   0.00   1.25   5.00
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) 6.100e+01  1.183e+00  51.554  < 2e-16 ***
## dietB       5.000e+00  1.528e+00   3.273 0.003803 **
## dietC       7.000e+00  1.528e+00   4.583 0.000181 ***
## dietD       2.991e-15  1.449e+00   0.000 1.000000
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.366 on 20 degrees of freedom
## Multiple R-squared:  0.6706, Adjusted R-squared:  0.6212
## F-statistic: 13.57 on 3 and 20 DF,  p-value: 4.658e-05
```

```r
anova(g)
```

```
## Analysis of Variance Table
##
## Response: coag
##           Df Sum Sq Mean Sq F value     Pr(>F)
## diet       3    228    76.0  13.571 4.658e-05 ***
## Residuals 20    112     5.6
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
model.matrix(g)
```

```
##    (Intercept) dietB dietC dietD
## 1            1     0     0     0
## 2            1     0     0     0
## 3            1     0     0     0
## 4            1     0     0     0
## 5            1     1     0     0
## 6            1     1     0     0
## 7            1     1     0     0
## 8            1     1     0     0
## 9            1     1     0     0
## 10           1     1     0     0
## 11           1     0     1     0
## 12           1     0     1     0
## 13           1     0     1     0
## 14           1     0     1     0
## 15           1     0     1     0
## 16           1     0     1     0
## 17           1     0     0     1
## 18           1     0     0     1
## 19           1     0     0     1
## 20           1     0     0     1
## 21           1     0     0     1
## 22           1     0     0     1
## 23           1     0     0     1
## 24           1     0     0     1
```

```
## attr(,"assign")
## [1] 0 1 1 1
## attr(,"contrasts")
## attr(,"contrasts")$diet
## [1] "contr.treatment"
```

```r
coag.aov <- aov(coag~diet)
summary(coag.aov)
```

```
##             Df Sum Sq Mean Sq F value   Pr(>F)
## diet         3    228    76.0   13.57 4.66e-05 ***
## Residuals   20    112     5.6
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
# Using custom made dummy-variables:

diet1 <- with(coagulation,ifelse(diet=="B",1,0))
diet2 <- with(coagulation,ifelse(diet=="C",1,0))
diet3 <- with(coagulation,ifelse(diet=="D",1,0))
diets <- cbind(diet1,diet2,diet3)
diets
```

```
##       diet1 diet2 diet3
##  [1,]     0     0     0
##  [2,]     0     0     0
##  [3,]     0     0     0
##  [4,]     0     0     0
##  [5,]     1     0     0
##  [6,]     1     0     0
##  [7,]     1     0     0
##  [8,]     1     0     0
##  [9,]     1     0     0
## [10,]     1     0     0
## [11,]     0     1     0
## [12,]     0     1     0
## [13,]     0     1     0
## [14,]     0     1     0
## [15,]     0     1     0
## [16,]     0     1     0
## [17,]     0     0     1
## [18,]     0     0     1
## [19,]     0     0     1
## [20,]     0     0     1
## [21,]     0     0     1
## [22,]     0     0     1
## [23,]     0     0     1
## [24,]     0     0     1
```

```r
coag.lm <- lm(coag~diets,coagulation)
anova(coag.lm)
```

```
## Analysis of Variance Table
##
## Response: coag
##             Df Sum Sq Mean Sq F value   Pr(>F)
```
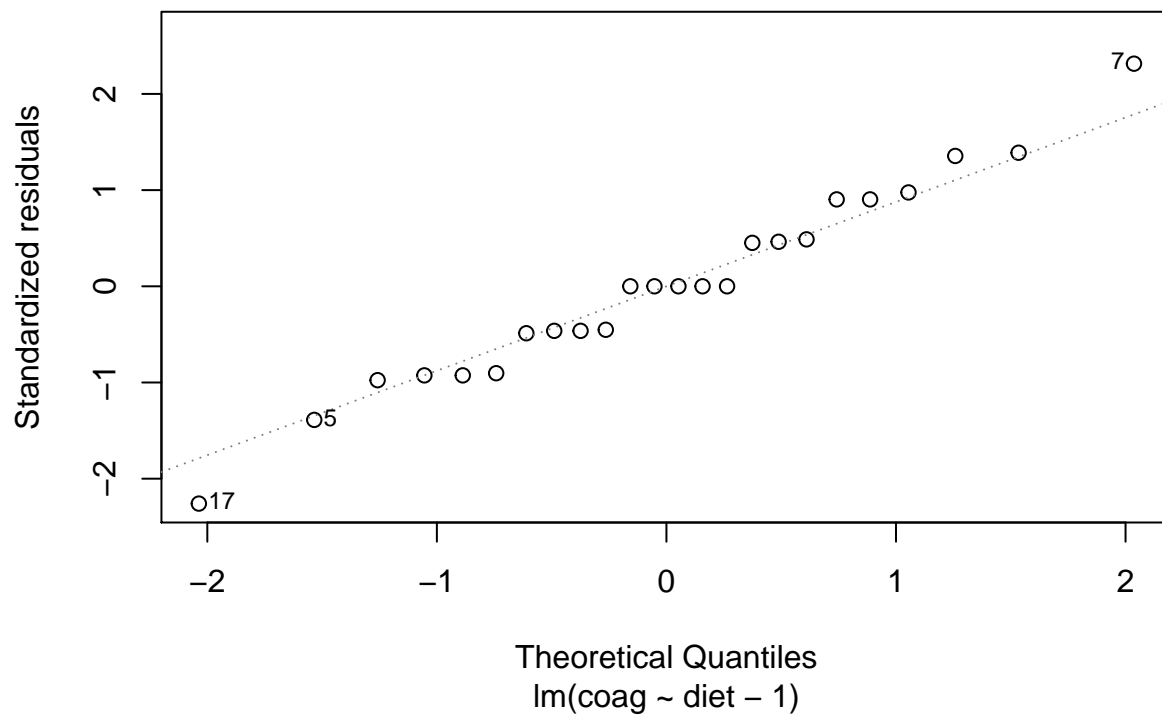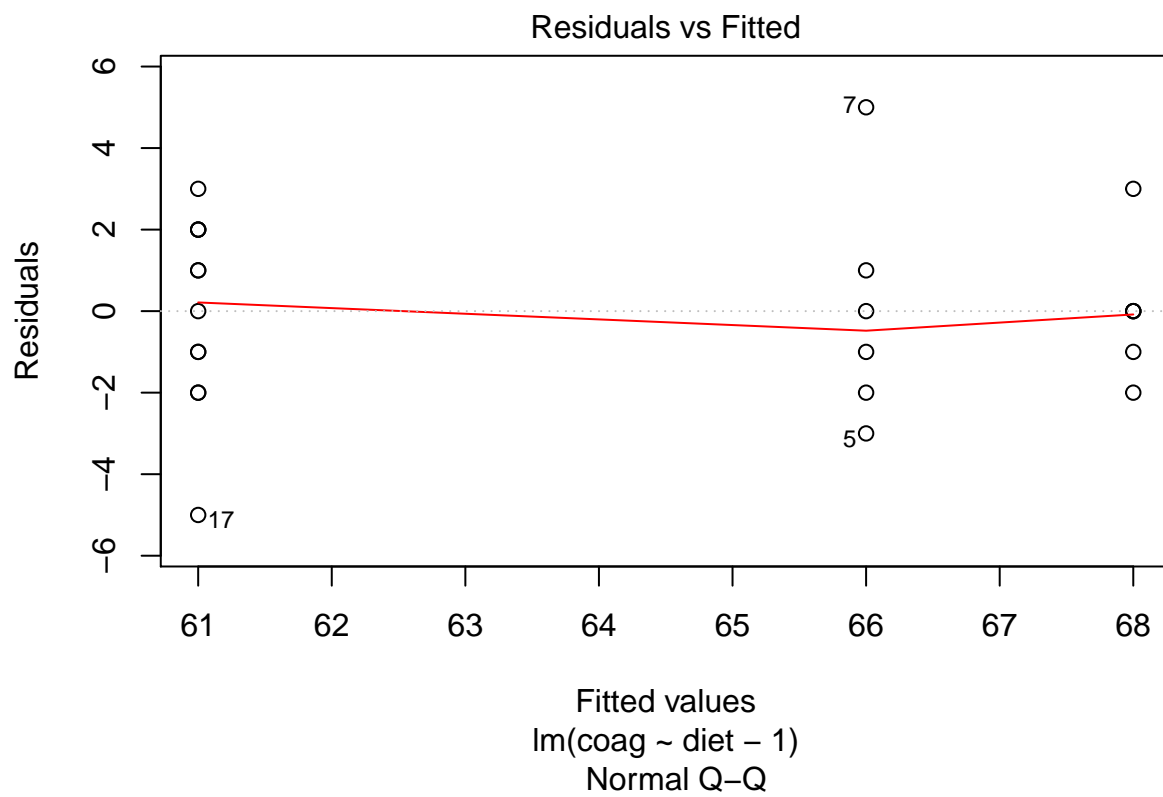
```
## diets       3     228     76.0   13.571 4.658e-05 ***
## Residuals 20     112      5.6
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```
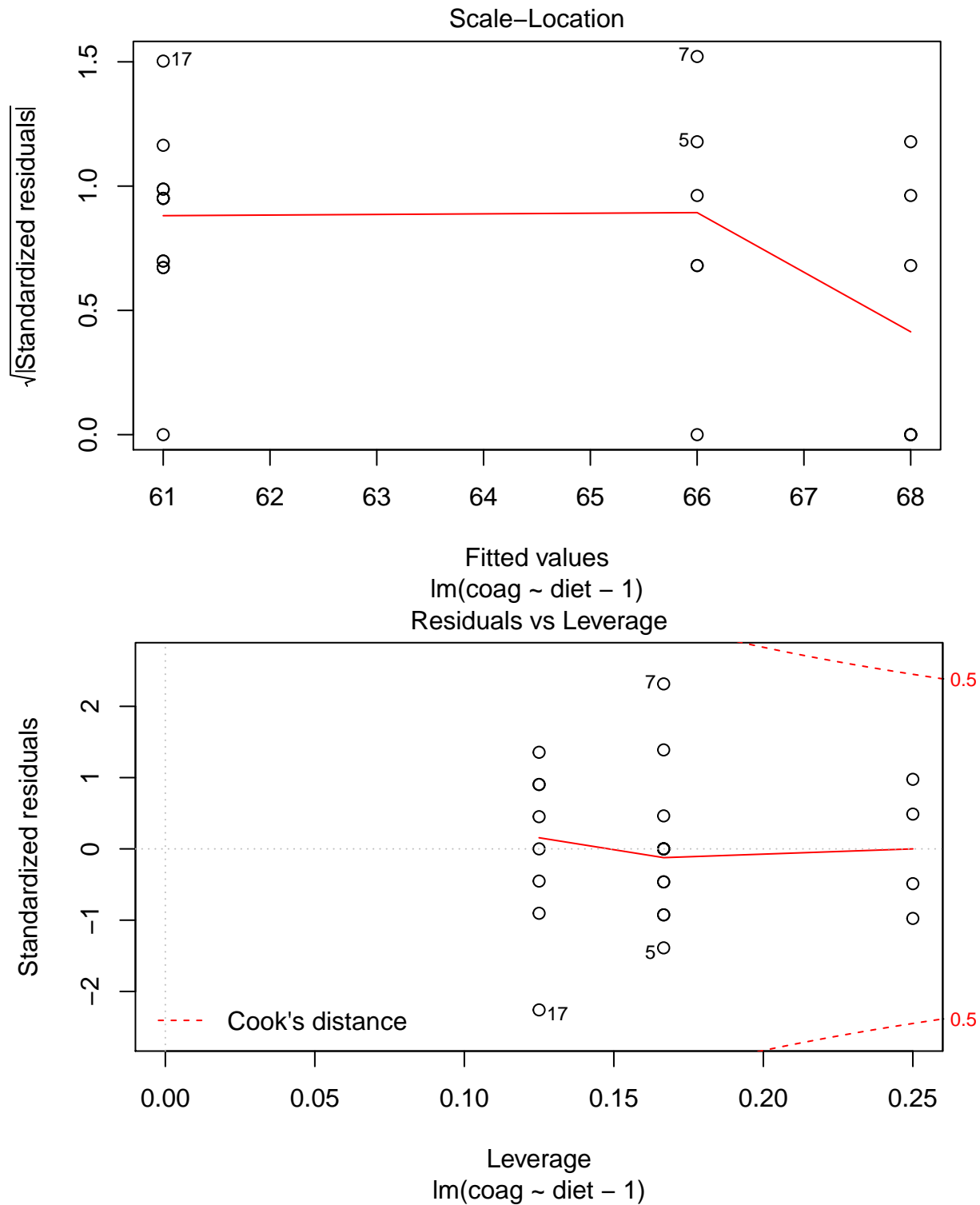
```r
# Fit Model I (also known as zero-intercept constraint)

gi <- lm(coag ~ diet-1)
summary(gi)
```

```
##
## Call:
## lm(formula = coag ~ diet - 1)
##
## Residuals:
##    Min    1Q Median     3Q    Max
##  -5.00  -1.25   0.00   1.25   5.00
##
## Coefficients:
##        Estimate Std. Error t value Pr(>|t|)
## dietA  61.0000     1.1832   51.55   <2e-16 ***
## dietB  66.0000     0.9661   68.32   <2e-16 ***
## dietC  68.0000     0.9661   70.39   <2e-16 ***
## dietD  61.0000     0.8367   72.91   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.366 on 20 degrees of freedom
## Multiple R-squared:  0.9989, Adjusted R-squared:  0.9986
## F-statistic:  4399 on 4 and 20 DF,  p-value: < 2.2e-16
```

```r
plot(gi)
```

Residuals vs Fitted

7○

○

○

○

○

○

○

○

○

○

○

○

○

5○

○17

Residuals

61 62 63 64 65 66 67 68

Fitted values
lm(coag ~ diet − 1)

Normal Q–Q

7○

○

○

○

○

○○○○

○○○

○○

○

○

○5

○17

Standardized residuals

−2 −1 0 1 2

Theoretical Quantiles
lm(coag ~ diet − 1)

8

Scale–Location

lm(coag ~ diet − 1)

Residuals vs Leverage

lm(coag ~ diet − 1)

```r
# This model gives same plots as before, but not the same ANOVA table and not the same
# overall F statistic:

anova(gi)

## Analysis of Variance Table
```

```
## 
## Response: coag
##           Df Sum Sq Mean Sq F value    Pr(>F)
## diet       4  98532 24633.0  4398.8 < 2.2e-16 ***
## Residuals 20    112     5.6
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
model.matrix(gi)
```

```
##    dietA dietB dietC dietD
## 1      1     0     0     0
## 2      1     0     0     0
## 3      1     0     0     0
## 4      1     0     0     0
## 5      0     1     0     0
## 6      0     1     0     0
## 7      0     1     0     0
## 8      0     1     0     0
## 9      0     1     0     0
## 10     0     1     0     0
## 11     0     0     1     0
## 12     0     0     1     0
## 13     0     0     1     0
## 14     0     0     1     0
## 15     0     0     1     0
## 16     0     0     1     0
## 17     0     0     0     1
## 18     0     0     0     1
## 19     0     0     0     1
## 20     0     0     0     1
## 21     0     0     0     1
## 22     0     0     0     1
## 23     0     0     0     1
## 24     0     0     0     1
## attr(,"assign")
## [1] 1 1 1 1
## attr(,"contrasts")
## attr(,"contrasts")$diet
## [1] "contr.treatment"
```

```r
# Fit Model II - sum contrasts constraint - by resetting the default constraints:

options(contrasts=c("contr.sum","contr.poly"))

gs <- lm(coag ~ diet)
summary(gs)
```

```
## 
## Call:
## lm(formula = coag ~ diet)
## 
## Residuals:
##    Min     1Q Median     3Q    Max
##  -5.00  -1.25   0.00   1.25   5.00
```

```
## 
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  64.0000     0.4979 128.537  < 2e-16 ***
## diet1        -3.0000     0.9736  -3.081 0.005889 **
## diet2         2.0000     0.8453   2.366 0.028195 *
## diet3         4.0000     0.8453   4.732 0.000128 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## Residual standard error: 2.366 on 20 degrees of freedom
## Multiple R-squared:  0.6706, Adjusted R-squared:  0.6212
## F-statistic: 13.57 on 3 and 20 DF,  p-value: 4.658e-05
```

```r
#diagnostic checks
qqnorm(residuals(g))
```

Normal Q–Q Plot
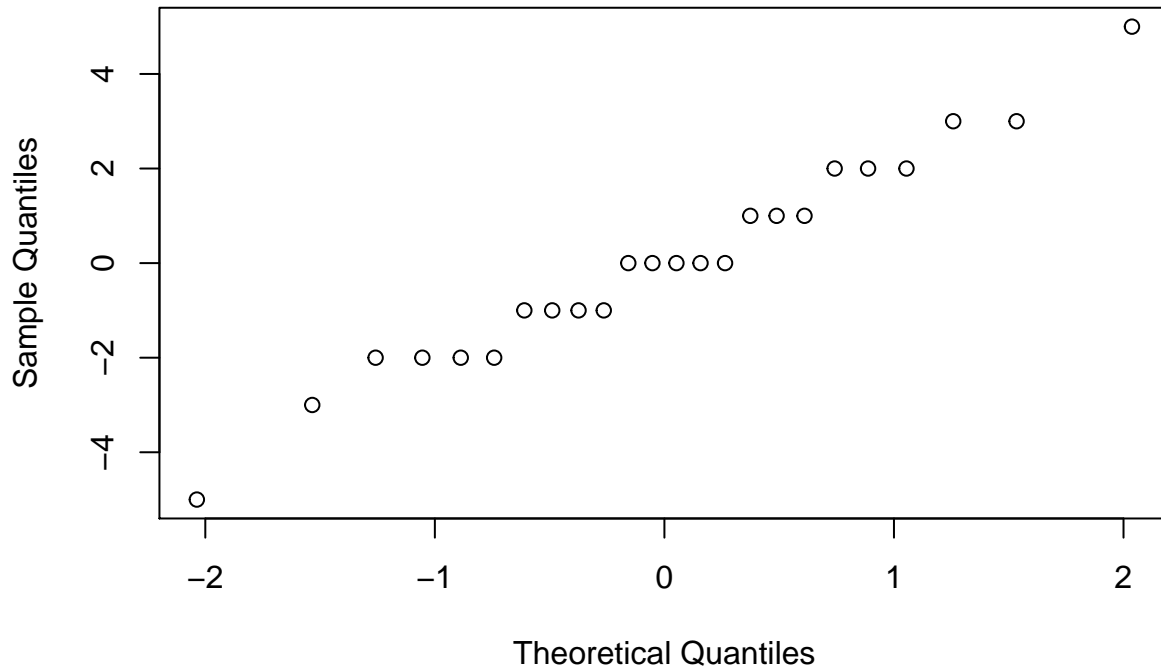

```r
# This command used to store the following picture to a file for later use in the slides:
# pdf("Pic3.pdf")

with(coagulation,plot(jitter(fitted(g)),residuals(g),xlab="Fitted",ylab="Residuals",pch=as.character(di
abline(h=0)
```
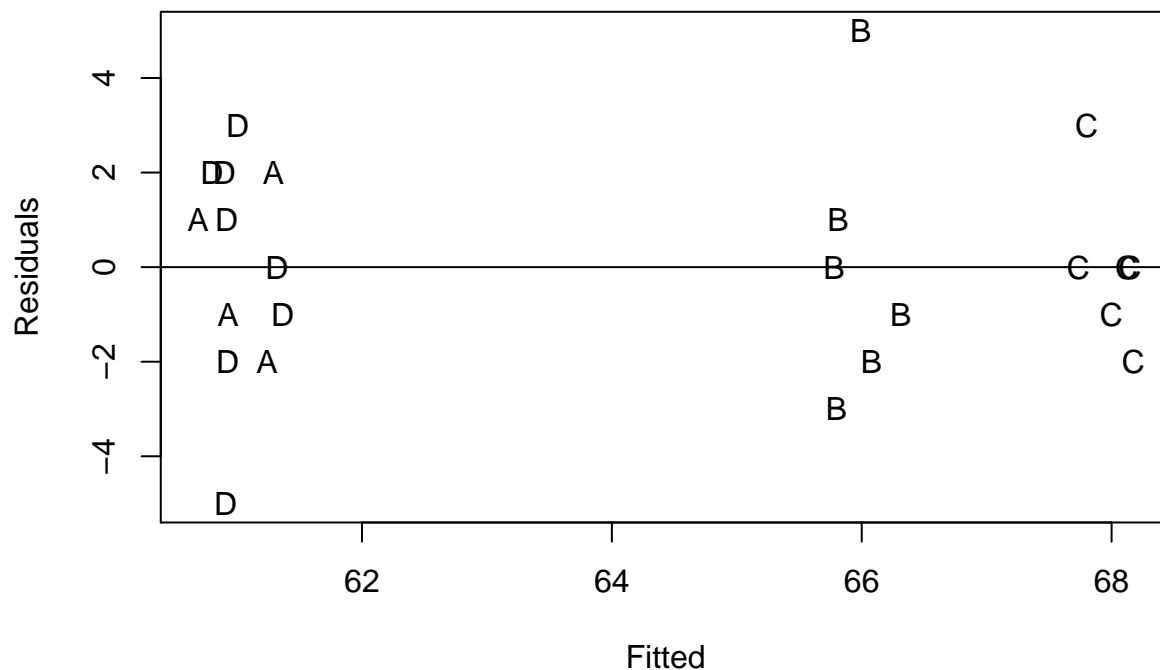
```
# Note if you use pdf("Pic3.pdf"), plot does not appear in R, but is stored in the pdf file
# instead - you then need to switch the device off to finishing writing to the external file:
# dev.off()

with(coagulation, var(coag[diet=="A"]))
```

```
## [1] 3.333333
```

```
with(coagulation, var(coag[diet=="B"]))
```

```
## [1] 8
```

```
with(coagulation, var(coag[diet=="C"]))
```

```
## [1] 2.8
```

```
with(coagulation, var(coag[diet=="D"]))
```

```
## [1] 6.857143
```

```
# Pairwise comparisons:

h1 <- c(-1,1,0,0)
h2 <- c(-1,0,1,0)
h3 <- c(-1,0,0,1)
h4 <- c(0,1,-1,0)
h5 <- c(0,1,0,-1)
h6 <- c(0,0,-1,1)
MSE<-summary(g)$sigma^2
ni<-with(coagulation,tapply(coag,diet,length))
sd1 <- sqrt(MSE)*sqrt(sum((h1^2)/ni))
sd2 <- sqrt(MSE)*sqrt(sum((h2^2)/ni))
sd3 <- sqrt(MSE)*sqrt(sum((h3^2)/ni))
sd4 <- sqrt(MSE)*sqrt(sum((h4^2)/ni))
sd5 <- sqrt(MSE)*sqrt(sum((h5^2)/ni))
```

```r
sd6 <- sqrt(MSE)*sqrt(sum((h6^2)/ni))
t1hat <- t(h1)%*%lvl.mns
t2hat <- t(h2)%*%lvl.mns
t3hat <- t(h3)%*%lvl.mns
t4hat <- t(h4)%*%lvl.mns
t5hat <- t(h5)%*%lvl.mns
t6hat <- t(h6)%*%lvl.mns
ests <- c(t1hat,t2hat,t3hat,t4hat,t5hat,t6hat)
sds <- c(sd1,sd2,sd3,sd4,sd6,sd6)
uppers <- ests+qt(1-0.05/12,coag.aov$df.residual)*sds
lowers <- ests-qt(1-0.05/12,coag.aov$df.residual)*sds
ints <- cbind(uppers,ests,lowers)
ints
```

```
##           uppers ests       lowers
## [1,]   9.471248    5    0.5287517
## [2,]  11.471248    7    2.5287517
## [3,]   4.241799    0   -4.2417986
## [4,]   1.999206   -2   -5.9992061
## [5,]   8.740915    5    1.2590853
## [6,]  -3.259085   -7  -10.7409147
```

```r
# Special protein test:

h <- c(0.5,-0.5,-0.5,0.5)
sd <- sqrt(MSE)*sqrt(sum((h^2)/ni))
est <- t(h)%*%lvl.mns
upper<-est+qt(0.975,coag.aov$df.residual)*sd
lower<-est-qt(0.975,coag.aov$df.residual)*sd
c(lower,est,upper)
```

```
## [1] -8.077254 -6.000000 -3.922746
```