Untitled

Kah Meng Soh

2/6/2022

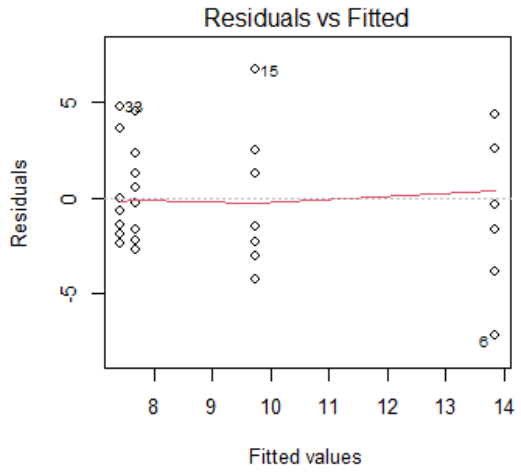
```
data=read.table("http://www.stat.umn.edu/~gary/book/fcdae.data/ex6.3",header=
TRUE)
attach(data)
drug=as.factor(drug)
m=lm(dose~drug)
anova(m)
```

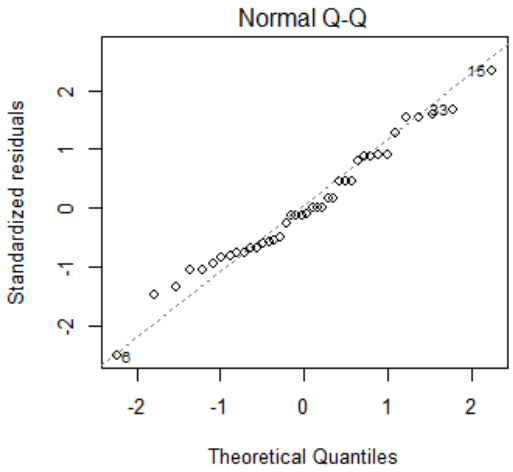
```
## Analysis of Variance Table
##
## Response: dose
##           Df Sum Sq Mean Sq F value    Pr(>F)
## drug       3 265.48  88.494  9.8651 6.906e-05 ***
## Residuals 36 322.93   8.970
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

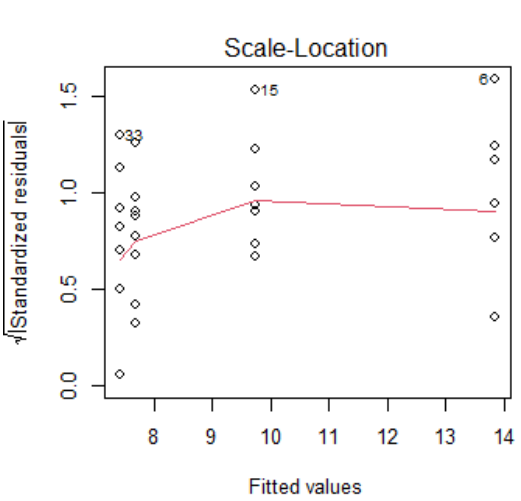
```
summary(m)
```

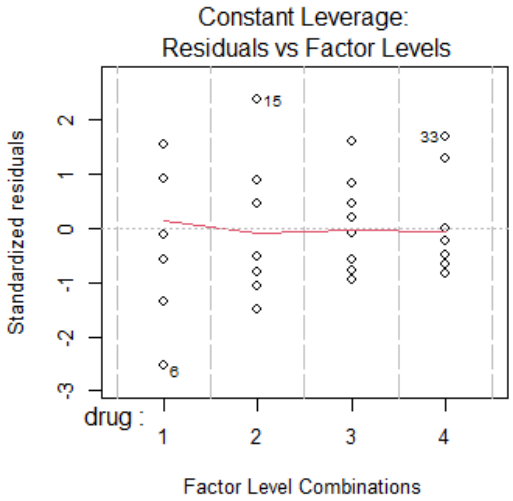
```
##
## Call:
## lm(formula = dose ~ drug)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -7.160 -1.982 -0.330  2.342  6.670
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  13.8600     0.9471  14.634  < 2e-16 ***
## drug2        -4.1300     1.3394  -3.083  0.00392 **
## drug3        -6.1600     1.3394  -4.599 5.08e-05 ***
## drug4        -6.4500     1.3394  -4.815 2.64e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.995 on 36 degrees of freedom
## Multiple R-squared:  0.4512, Adjusted R-squared:  0.4054
## F-statistic: 9.865 on 3 and 36 DF,  p-value: 6.906e-05
```

```
par(mfrow=c(2,2))
plot(m)
```










```
#From the residual vs fitted, the slope is horizontal with some part slanted,
therefore I believe the error variance to be equal, and from the Normal Q-Q
plot, many points are on the straight line, therefore I believe that error
follows normality assumption. Hence, we considered it valid to use ANOVA.
#Based on the treatment mean, the most effective drug is drug 4 and the least
effective drug is drug 1

m=aov(dose ~ drug)
cis=TukeyHSD(m, which="drug", ordered=T, conf.level=0.95)
cis

##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##     factor levels have been ordered
##
## Fit: aov(formula = dose ~ drug)
##
```

```
## $drug
##     diff        lwr       upr     p adj
## 3-4 0.29 -3.3173943  3.897394 0.9963493
## 2-4 2.32 -1.2873943  5.927394 0.3224748
## 1-4 6.45  2.8426057 10.057394 0.0001496
## 2-3 2.03 -1.5773943  5.637394 0.4390060
## 1-3 6.16  2.5526057  9.767394 0.0002860
## 1-2 4.13  0.5226057  7.737394 0.0195604
```

```
#Use MCB to find subset of treatments that are most similar to the most
effective drug with an error rate of 0.05.
m=lm(dose~drug)
anova(m)
```

```
## Analysis of Variance Table
##
## Response: dose
##           Df Sum Sq Mean Sq F value    Pr(>F)
## drug       3 265.48  88.494  9.8651 6.906e-05 ***
## Residuals 36 322.93   8.970
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
k=4-1
nu=40-4
smallest_trt_mu=7.41
dunnet_crit_value=2.13
se=sqrt(8.97/5)
best_cutoff=smallest_trt_mu+dunnet_crit_value*se
best_cutoff
```

```
## [1] 10.26293
```

```
#All treatments with group mean less than 10.26 will be in the 95%  best
subset grouping, so drug 2,3,4 are less than the cutoff value, so they are
the best subset which cannot be distinguished from the most effective drug.
```