STAT5303HW2

Mingming Xu

9/25/2020

E4.4

This experiment has six treatments. There are 3 treatments about low moisture:particle board at low moisture, plywood at low moisture and solid fir at low moisture.There are 2 treatments about medium moisture: particle board at medium moisture and plywood at medium moisture.And one particle board at high moisture.

Because we are only interested in comparing low moisture to medium moisture,I will set up a linear contrast low moisture vs medium moisture with coefficients of these six treatments (1/3,1/3,1/3,-1/2,-1/2,0) with respsct to particle board at low moisture, plywood at low moisture,solid fir at low moisture,particle board at medium moisture,plywood at medium moisture and particle board at high moisture.

P4.2

```r
library(cfcdae)
```

```
## Warning: package 'cfcdae' was built under R version 4.0.2
```

```
## Registered S3 method overwritten by 'DoE.base':
##   method            from
##   factorize.factor conf.design
```

```r
data("FruitFlyLifespan")
attach(FruitFlyLifespan)
FruitFlyLifespan$companions<-as.factor(FruitFlyLifespan$companions)
fly.seperate<-lm(longevity~companions,data = FruitFlyLifespan)
##We check the model assumptions.
par(mfrow = c(1, 2))
plot(fly.seperate, which = c(1, 2)) ##assumptions are satisfied
```
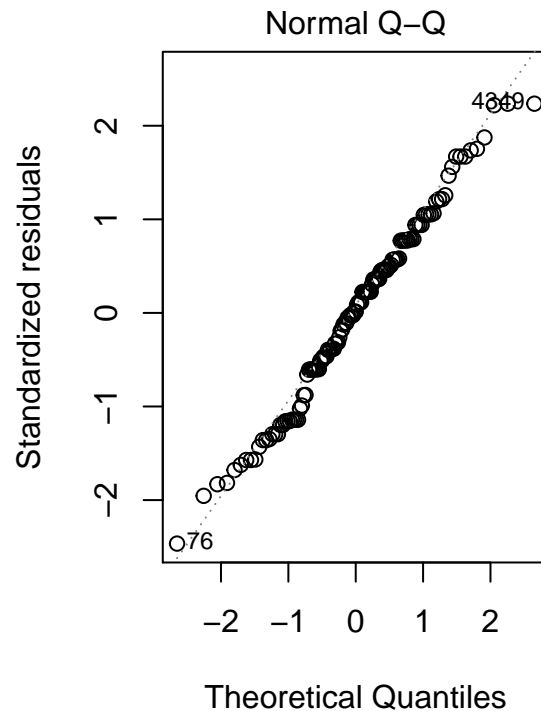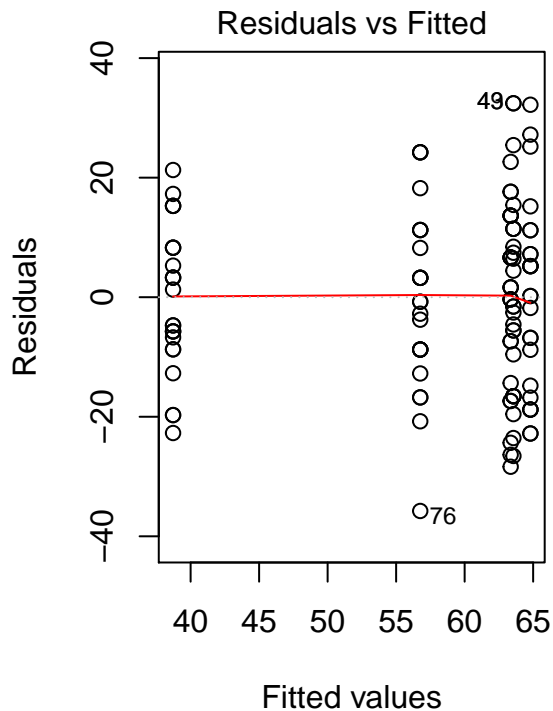
```r
summary(fly.seperate);model.effects(fly.seperate,"companions")
```

```
## 
## Call:
## lm(formula = longevity ~ companions, data = FruitFlyLifespan)
## 
## Residuals:
##    Min     1Q Median     3Q    Max
## -35.76  -8.76   0.20  11.20  32.44
## 
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   57.440      1.325  43.368  < 2e-16 ***
## companions1    6.120      2.649   2.310  0.02258 *
## companions2    7.360      2.649   2.778  0.00634 **
## companions3   -0.680      2.649  -0.257  0.79785
## companions4  -18.720      2.649  -7.067 1.13e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## Residual standard error: 14.81 on 120 degrees of freedom
## Multiple R-squared:  0.3121, Adjusted R-squared:  0.2892
## F-statistic: 13.61 on 4 and 120 DF,  p-value: 3.516e-09

## 1pregnant  1virgin 8pregnant  8virgin     none
##      6.12     7.36     -0.68   -18.72     5.92
```

```r
anova(fly.seperate)##Since the p-value is so small, there is sinificant diffrence between these treatmen
```

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

```
##             Df Sum Sq Mean Sq F value   Pr(>F)
## companions   4  11939 2984.82  13.612 3.516e-09 ***
## Residuals  120  26314  219.28
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##contrast 1:
## I want to test if the males were supplied with receptive virgin female fruit flies per day have a sho
linear.contrast(fly.seperate,companions,c(0,0.5,0,0.5,-1))
```

```
##  estimates       se   t-value    p-value lower-ci  upper-ci
## 1    -11.6 3.627225 -3.198037 0.001769626 -18.78165 -4.418347
```

```
##Since p-value is 0.001769626, it shows a statistically sinificant difference so I can reject the null
##contrast 2:
## I want to test if the males were supplied with 8 receptive virgin female fruit flies per day have a s
linear.contrast(fly.seperate,companions,c(0,1,0,-1,0))
```

```
##  estimates       se  t-value    p-value lower-ci upper-ci
## 1     26.08 4.188358 6.226783 7.268917e-09 17.78734 34.37266
```

```
##Since p-value is 7.268917e-09 which is so small, it shows a statistically sinificant difference betwee
##contrast 3:
## I want to test if the males were supplied with receptive virgin female fruit flies per day have a sho
linear.contrast(fly.seperate,companions,c(0.5,-0.5,0.5,-0.5,0))
```

```
##  estimates       se  t-value    p-value lower-ci upper-ci
## 1       8.4 2.961617 2.836289 0.005358879 2.536205 14.26379
```

```
##Since p-value is 0.005358879,it cannt explian there is a statistically sinificant difference between
```

P4.3

```
data("PoliticalAds")
attach(PoliticalAds)
condition<-as.factor(condition)
levels(condition)
```

```
## [1] "Neg ed."        "Pos ed."        "Trans & neg ed." "Trans & pos ed."
## [5] "Trans only"
```

```
pads<-lm(response~condition)
##We check the model assumptions.
par(mfrow = c(1, 2))
plot(pads, which = c(1, 2)) ##assumptions are satisfied
```
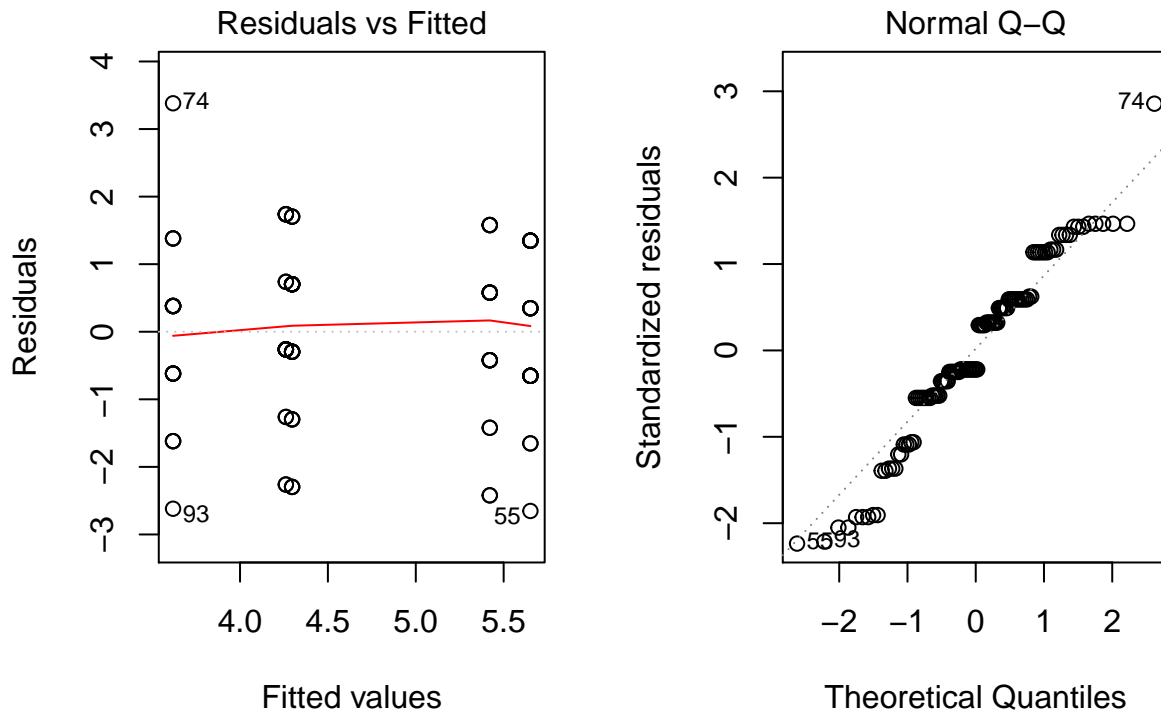
```r
anova(pads);model.effects(pads,"condition")
```

```
## Analysis of Variance Table
##
## Response: response
##             Df  Sum Sq Mean Sq F value    Pr(>F)
## condition    4  63.524  15.881  10.796 2.123e-07 ***
## Residuals  108 158.866   1.471
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

##          Neg ed.        Pos ed. Trans & neg ed. Trans & pos ed.     Trans only
##        0.7711646     -1.0308404       1.0022859      -0.3535917     -0.3890184
```

From the ANOVA table, the p-value 2.123e-07 is so small so it shows that there is a statistically sinificant diffence between these 5 treatments.So,it implies that the media covarage affect perception.

```r
##I will set up a linear contrast to compare Trans vs no Trans
linear.contrast(pads,condition,c(-0.5,-0.5,1/3,1/3,1/3))
```

```
##   estimates        se t-value  p-value   lower-ci  upper-ci
## 1 0.2163965 0.2390239 0.905334 0.3673037 -0.2573904 0.6901833
```

There is no enough evidence that shows the actual transcript changes the response to media reporting.

```r
##I will set up a linear contrast to compare  pos.ad vs neg.ad
linear.contrast(pads,condition,c(-0.5,0.5,-0.5,0.5,0))
```

```
##    estimates        se   t-value      p-value   lower-ci upper-ci
## 1 -1.578941 0.2578289 -6.123989 1.505443e-08 -2.090003 -1.06788
```

Because the p-value 1.505443e-08 is so small so it has enough evidence to sho that there is a statistically sinificant diffence between positive reporting and negative reporting in improving preception.

P4.6

```
2*0+-1*1+-1*-1
```

[1] 0

The result is 0 so these two contrast are orthgonal. These two contrasts are statistically independent. In this experiment,there are 3 group.Then by Pythagorean theorem, these two contrasts' sum squares partition the sum of squares for treatments with 2 df.So the sum of squares for the contrast with coefficients (0, 1, -1) is 100-80=20.

E5.5

```
library(cfcdae)
data("OrangePulpSilage")
attach(OrangePulpSilage)
treatment<-as.factor(treatment)
orange.seperate<-lm(moisture~treatment)
##We check the model assumptions.
par(mfrow = c(1, 2))
plot(orange.seperate, which = c(1, 2)) ##assumptions are satisfied
```
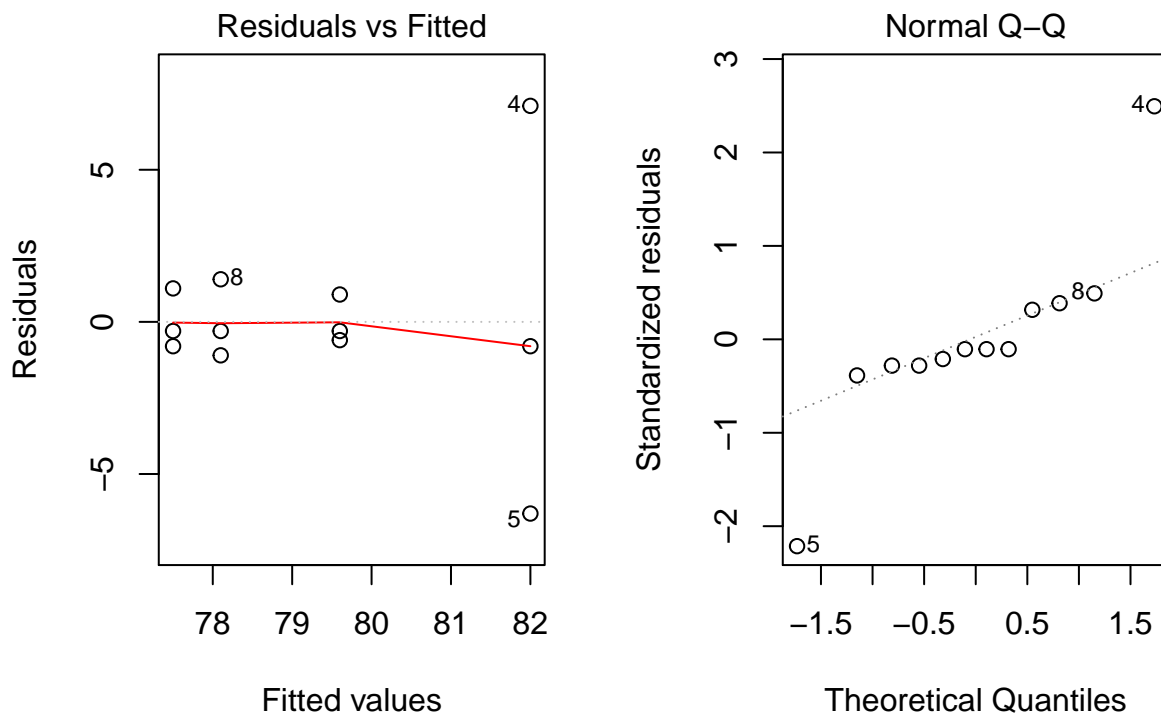


```
levels(treatment)
```

[1] "BeetPulp"   "Control"   "FormicAcid" "NaCl"

```
compare.to.control(orange.seperate,treatment,control=2,conf=0.95)
```

```
##                      difference     lower     upper
##   BeetPulp - Control        0.6 -7.582884  8.782884
##   NaCl - Control            2.1 -6.082884 10.282884
##   FormicAcid - Control      4.5 -3.682884 12.682884
```

P5.1

```
library(cfcdae)
data("FruitFlyLifespan")
```

```
attach(FruitFlyLifespan)

## The following objects are masked from FruitFlyLifespan (pos = 5):
##
##     companions, longevity

FruitFlyLifespan$companions<-as.factor(FruitFlyLifespan$companions)
fly.seperate<-lm(longevity~companions,data = FruitFlyLifespan)
fly.tukey<-pairwise(fly.seperate,companions)
fly.tukey

##
## Pairwise comparisons ( hsd ) of companions
##                         estimate signif diff     lower      upper
##    1pregnant - 1virgin     -1.24   11.60047 -12.840468  10.360468
##    1pregnant - 8pregnant    6.80   11.60047  -4.800468  18.400468
## * 1pregnant - 8virgin      24.84   11.60047  13.239532  36.440468
##    1pregnant - none         0.20   11.60047 -11.400468  11.800468
##    1virgin - 8pregnant      8.04   11.60047  -3.560468  19.640468
## * 1virgin - 8virgin        26.08   11.60047  14.479532  37.680468
##    1virgin - none           1.44   11.60047 -10.160468  13.040468
## * 8pregnant - 8virgin      18.04   11.60047   6.439532  29.640468
##    8pregnant - none        -6.60   11.60047 -18.200468   5.000468
## * 8virgin - none          -24.64   11.60047 -36.240468 -13.039532

sidelines(fly.tukey)

##
## 8virgin   -18.72
## 8pregnant  -0.68 |
## none        5.92 |
## 1pregnant   6.12 |
## 1virgin     7.36 |
```

So, the results shows that there is no significant difference between the treatments 8pregnant,1pregnant,1virgin and none. However,it also shows that 8virgin has significant difference with other treatment groups.

P5.2

```
library(cfcdae)
data("Autoclaving")
attach(Autoclaving)

## The following object is masked _by_ .GlobalEnv:
##
##     treatment

## The following object is masked from OrangePulpSilage:
##
##     treatment

auto.seperate<-lm(modulus~treatment,data = Autoclaving)
##We check the model assumptions.
par(mfrow = c(1, 2))
plot(auto.seperate, which = c(1, 2)) ##assumptions are satisfied
```
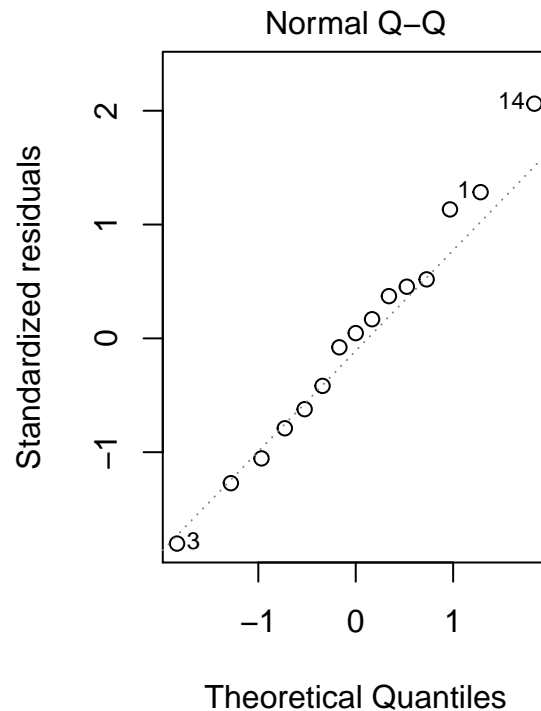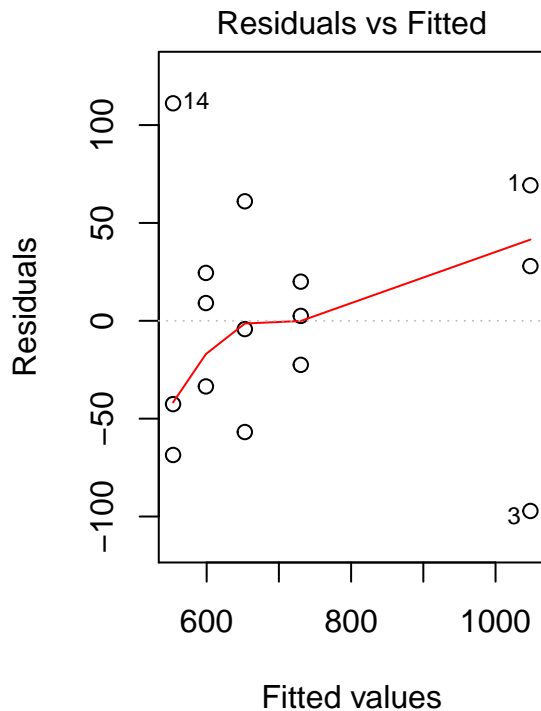
Residuals vs Fitted

Normal Q-Q

```
anova(auto.seperate);model.effects(auto.seperate,treatment)
```

```
## Analysis of Variance Table
##
## Response: modulus
##            Df Sum Sq Mean Sq F value    Pr(>F)
## treatment   4 464143  116036   26.63 2.594e-05 ***
## Residuals  10  43574    4357
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

##      121_10     121_20     135_10     135_20        None
##    13.54000  -63.92667 -117.82667 -163.29333   331.50667
```

```
##By the anova table, the small p-value explains that there is significant difference between these trea
boxplot(modulus~treatment,data = Autoclaving)
##I will choose the per comparison error rate.
pairwise(auto.seperate,treatment,confidence = 0.99,type="lsd")
```

```
##
## Pairwise comparisons ( lsd ) of treatment
##                    estimate signif diff       lower       upper
##   121_10 - 121_20   77.46667    170.8151  -93.348421   248.2818
##   121_10 - 135_10  131.36667    170.8151  -39.448421   302.1818
## * 121_10 - 135_20  176.83333    170.8151    6.018245   347.6484
## * 121_10 - None    -317.96667   170.8151 -488.781755  -147.1516
##   121_20 - 135_10   53.90000    170.8151 -116.915088   224.7151
##   121_20 - 135_20   99.36667    170.8151  -71.448421   270.1818
## * 121_20 - None    -395.43333   170.8151 -566.248421  -224.6182
##   135_10 - 135_20   45.46667    170.8151 -125.348421   216.2818
## * 135_10 - None    -449.33333   170.8151 -620.148421  -278.5182
## * 135_20 - None    -494.80000   170.8151 -665.615088  -323.9849
```

modulus

121_10   135_10   None

treatment
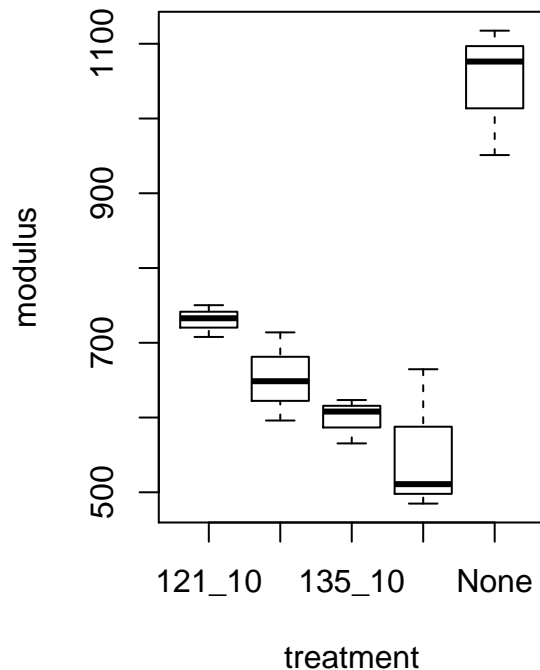
P5.6

pros:

per comparison error rate:For these 10,000 locations,this error rate just does individual test for each null hypothesis ignoring the multiple testing. Each individual comparison is controlled at ϵ.

false discovery rate: Since the units 10,000 is a little large,it may cause that we might have many incorrect rejections.With this error rate, we allow more incorrect rejections of null hypopthesis as the number of true rejections increases, but the ratio is limited. And the value of ϵcan be flexibly selected according to needs.

strong familywise error rate:It controls that probability of making any false rejections is at ϵ or under it,regardless of the number of correct rejections made.

cons:

per comparison error rate: It makes no correction for multiple comparisons, therefore there may be many Type I errors.

false discovery rate: It may increase more Type I errors.

strong familywise error rate: The value of ϵ is traditionally more fixed.

So in this experiment ,at each location, we test the null hypothesis hat the expected gamma ray emission rate is equal to the known background rate and clean up the locations that are significantly diffrent. So,I would like to use per comparison error rate because it controls each individual test under the error rate.

P5.9

In this experiment, one student has a p-value less than 5% for telepathy, and one student has a p-value less than 5% for precognition.If we choose 0.05 as the significance level,in this case,p-value $< 5\%$.Then, we conclude that the null hypothesis that any single person could exhibit any kind of ESP is false and should be rejected.