# Assignment 16

Sachin Shubham

5/2/2021

Task 1

Two each of 40 samples of canned meat were stored at 2, 4.5, 21, and 38C for periods of 1,2,4,8, and 13 months, respectively. The two samples from each factor combination were randomly given to two taste panels who rated the samples on a continuous scale from 1 (excellent) to 8 (unacceptable). The data are given in Table 1; the two numbers in each combination are the ratings of panels 1 and 2, respectively. Analyze the data to ascertain the relationship of the quality of meat and temperature. Note that both factors have numeric levels.

Solution:

```
library(multcompView)

## Warning: package 'multcompView' was built under R version 4.0.4

library(lsmeans)

## Warning: package 'lsmeans' was built under R version 4.0.5

## Loading required package: emmeans

## Warning: package 'emmeans' was built under R version 4.0.5

## The 'lsmeans' package is now basically a front end for 'emmeans'.
## Users are encouraged to switch the rest of the way.
## See help('transition') for more information, including how to
## convert old 'lsmeans' objects and scripts to work with 'emmeans'.

library(ggplot2)

sample<-c(2.38,2.19,2.74,2.50,2.75,2.74,3.28,2.83,3.81,3.05,
          2.67,2.39,2.81,2.64,3.00,2.79,3.58,3.23,3.67,3.61,
          2.93,2.72,2.97,2.88,3.05,3.21,3.68,3.25,4.04,4.23,
          3.81,3.07,4.14,3.14,4.78,3.45,5.78,5.28,6.05,7.14)
time<-c(1,1,2,2,4,4,8,8,13,13,
        1,1,2,2,4,4,8,8,13,13,
        1,1,2,2,4,4,8,8,13,13,
        1,1,2,2,4,4,8,8,13,13)
temprature<-c(2,2,2,2,2,2,2,2,2,2,
              4.5,4.5,4.5,4.5,4.5,4.5,4.5,4.5,4.5,4.5,
              21,21,21,21,21,21,21,21,21,21,
              38,38,38,38,38,38,38,38,38,38)
```
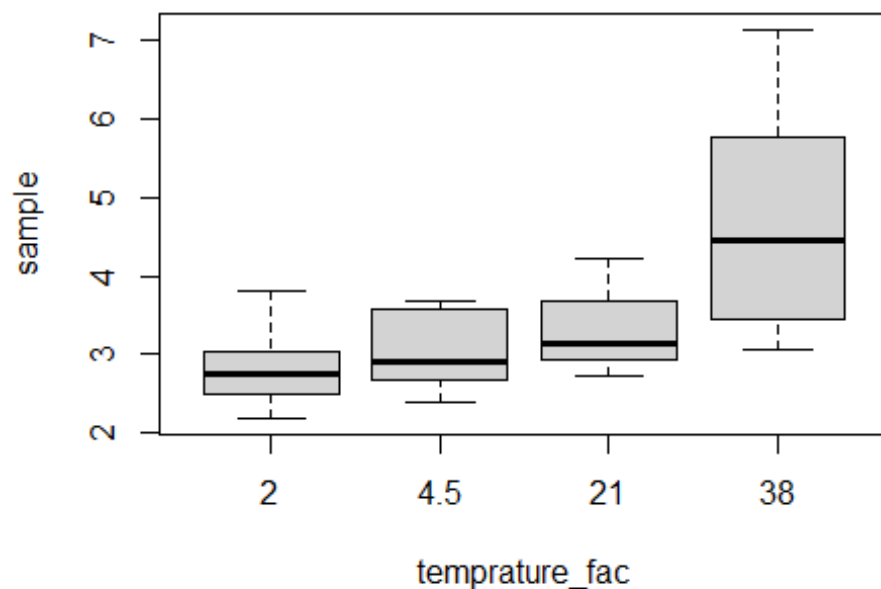
```r
time_fac<-as.factor(time)
temprature_fac<-as.factor(temprature)
df_task1<-data.frame(time_fac,temprature_fac,sample)

head(df_task1,5)

##   time_fac temprature_fac sample
## 1        1              2   2.38
## 2        1              2   2.19
## 3        2              2   2.74
## 4        2              2   2.50
## 5        4              2   2.75

boxplot(sample ~ temprature_fac)
```
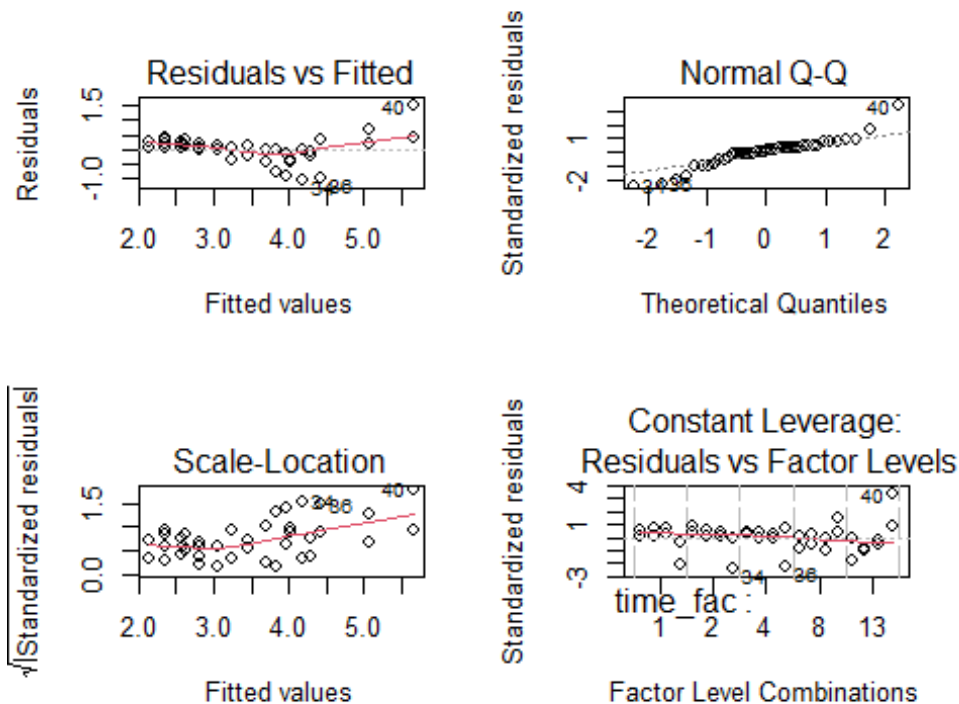


```r
model <- aov(sample ~ time_fac + temprature_fac,data = df_task1)
summary(model)

##                Df Sum Sq Mean Sq F value   Pr(>F)
## time_fac        4 15.272   3.818   15.29 4.31e-07 ***
## temprature_fac  3 20.544   6.848   27.43 5.65e-09 ***
## Residuals      32  7.988   0.250
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

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



```
marginal = lsmeans(model,
                    ~ temprature_fac)
pairs(marginal,
       adjust="tukey")

##  contrast estimate    SE df t.ratio p.value
##  2 - 4.5     -0.212 0.223 32 -0.949  0.7789
##  2 - 21      -0.469 0.223 32 -2.099  0.1752
##  2 - 38      -1.837 0.223 32 -8.221  <.0001
##  4.5 - 21    -0.257 0.223 32 -1.150  0.6618
##  4.5 - 38    -1.625 0.223 32 -7.272  <.0001
##  21 - 38     -1.368 0.223 32 -6.122  <.0001
##
## Results are averaged over the levels of: time_fac
## P value adjustment: tukey method for comparing a family of 4 estimates

CLD = multcomp::cld(marginal,
                    alpha   = 0.05,
                    Letters = letters,  # Use lower-case letters for .group
                    adjust  = "sidak")  # Sidak-adjusted p-values
CLD

##  temprature_fac lsmean    SE df lower.CL upper.CL .group
##  2                2.83 0.158 32    2.41     3.24  a
```

```
##  4.5                3.04 0.158 32    2.62    3.46  a
##  21                 3.30 0.158 32    2.88    3.71  a
##  38                 4.66 0.158 32    4.25    5.08   b
##
## Results are averaged over the levels of: time_fac
## Confidence level used: 0.95
## Conf-level adjustment: sidak method for 4 estimates
## P value adjustment: sidak method for 6 tests
## significance level used: alpha = 0.05
```

```r
#plot for LS Mean
ggplot(CLD,aes(x= temprature_fac ,y= lsmean,
               label = .group))+
  geom_point(shape  = 15,
             size   = 4) +

  geom_errorbar(aes(ymin  =  lower.CL,
                    ymax  =  upper.CL),
                width =  0.2,
                size  =  0.7) +

  theme_bw() +
  theme(axis.title   = element_text(face = "bold"),
        axis.text    = element_text(face = "bold"),
        plot.caption = element_text(hjust = 0)) +
  ylab("Mean Quality of the meat")+
  xlab("Temperature")
```
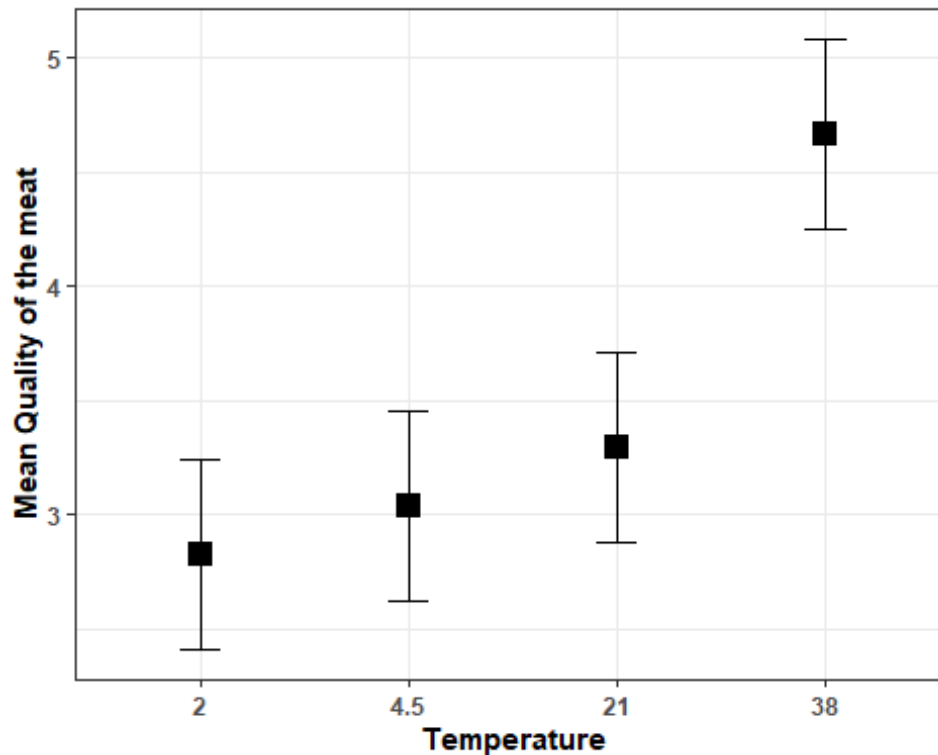
Conclusion:

Our study finds that, the quality of the meat is significantly related to the temperature the canned meat was stored as p-value << 0.05 and F-value >> 1.

Task 2

An agriculture experiment was conducted to compare four varieties of sweet potatoes. The experiment was conducted in a completely randomized design with varieties as the treatment. The response variable was yield in tons per acre. The data are given in Table 2. Test for a difference in distributions of yields using the Kruskal-Wallis test. (Use $\alpha = 0.01$.)

Solution

```
Values<-
c(8.3,9.4,9.1,9.1,9.0,8.9,8.9,9.1,9.0,8.1,8.2,8.8,8.4,8.3,10.1,10.0,9.6,9.3,9
.8,9.5,9.4,7.8,8.2,8.1,7.9,7.7,8.0,8.1)
Variety<-
c("Variety_A","Variety_A","Variety_A","Variety_A","Variety_A","Variety_A","Va
riety_A","Variety_B","Variety_B","Variety_B","Variety_B","Variety_B","Variety
_B","Variety_B","Variety_C","Variety_C","Variety_C","Variety_C","Variety_C","
Variety_C","Variety_C","Variety_D","Variety_D","Variety_D","Variety_D","Varie
ty_D","Variety_D","Variety_D")

df_task2<-data.frame(Values,Variety)

head(df_task2,5)
```

```
##   Values   Variety
## 1    8.3 Variety_A
## 2    9.4 Variety_A
## 3    9.1 Variety_A
## 4    9.1 Variety_A
## 5    9.0 Variety_A

kruskal.test(Values ~ Variety, data = df_task2)

##
##  Kruskal-Wallis rank sum test
##
## data:  Values by Variety
## Kruskal-Wallis chi-squared = 22.683, df = 3, p-value = 4.701e-05
```

Conclusion:

Our study finds that the Kruskal-Wallis p-value is << 0.01 so the difference in distributions is considered statistically significant.Therefore we reject the null hypothesis and accept the alternative hypothesis that the the distributions have no difference.