TP4

Guillaume, Julio, Pierre-Yves

15 mai 2019

##Exercice 1 :

1) (a)

tab<-read.table("hotdogs.txt",header=TRUE)  
head(tab)

## Type Calories Sodium  
## 1 1 186 495  
## 2 1 181 477  
## 3 1 176 425  
## 4 1 149 322  
## 5 1 184 482  
## 6 1 190 587

str(tab)

## 'data.frame': 55 obs. of 3 variables:  
## $ Type : int 1 1 1 1 1 1 1 1 1 1 ...  
## $ Calories: int 186 181 176 149 184 190 158 139 175 148 ...  
## $ Sodium : int 495 477 425 322 482 587 370 322 479 375 ...

1. nbre d’observation par type

table(tab$Type)

##   
## 1 2 3 4   
## 20 17 17 1

tab=tab[-which(tab$Type==4),]  
table(tab$Type)

##   
## 1 2 3   
## 20 17 17

On enlève l’échantillon unique dont le type est 4 (soja).

tab$Type=as.factor(tab$Type)

2] (a)

summary(tab$Calories)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 86.0 132.0 145.0 145.4 172.8 195.0

summary(tab$Sodium)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 144.0 362.5 405.0 424.8 503.5 645.0

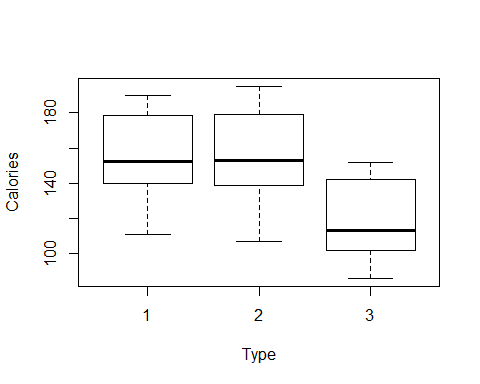
by(tab$Calories,tab$Type,function(x) summary(x))

## tab$Type: 1  
## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 111.0 140.5 152.5 156.8 177.2 190.0   
## --------------------------------------------------------   
## tab$Type: 2  
## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 107.0 139.0 153.0 158.7 179.0 195.0   
## --------------------------------------------------------   
## tab$Type: 3  
## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 86.0 102.0 113.0 118.8 142.0 152.0

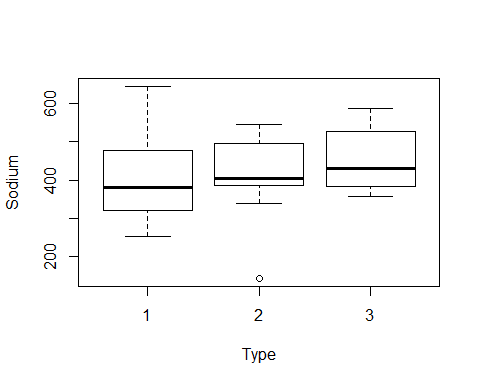
by(tab$Sodium,tab$Type,function(x) summary(x))

## tab$Type: 1  
## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 253.0 321.2 380.5 401.1 477.5 645.0   
## --------------------------------------------------------   
## tab$Type: 2  
## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 144.0 386.0 405.0 418.5 496.0 545.0   
## --------------------------------------------------------   
## tab$Type: 3  
## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 357 383 430 459 528 588

plot(Calories~Type,data=tab)



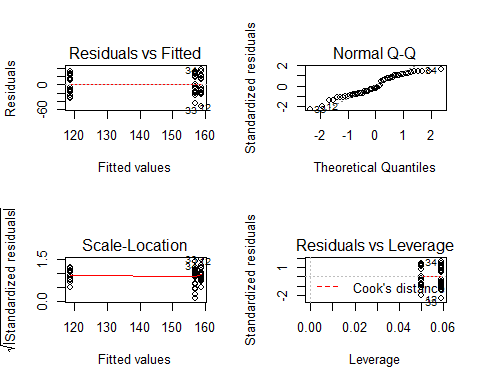
plot(Sodium~Type,data=tab)



1. Le type 3 contient moins de calories mais plus de sodium. Les types 1 et 2 sont similaire au niveaux des calories mais au niveau du sodium le type 2 est plus sal?.

3] (a) modèle 2 des effets et avec (effet du niveau du facteur) et terme d’erreur (indépendant et de loi normale centrée)

mod1=lm(Calories~Type,data=tab)  
par(mfrow=c(2,2))  
plot(mod1)

  
Les hypothèses sont vérifiées (résidus suivant une loi normale). (c)

anova(mod1)

## Analysis of Variance Table  
##   
## Response: Calories  
## Df Sum Sq Mean Sq F value Pr(>F)   
## Type 2 17692 8846.1 16.074 3.862e-06 \*\*\*  
## Residuals 51 28067 550.3   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Test si les (] égaux et nuls) égaux entre eux ou un diff?rent pour au moins un couple

Sum Sq Mean Sq Fvalue

Type

Résidus

summary(mod1)

##   
## Call:  
## lm(formula = Calories ~ Type, data = tab)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -51.706 -18.492 -5.278 22.500 36.294   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 156.850 5.246 29.901 < 2e-16 \*\*\*  
## Type2 1.856 7.739 0.240 0.811   
## Type3 -38.085 7.739 -4.921 9.39e-06 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 23.46 on 51 degrees of freedom  
## Multiple R-squared: 0.3866, Adjusted R-squared: 0.3626   
## F-statistic: 16.07 on 2 and 51 DF, p-value: 3.862e-06

Estimate Std. Error t value Pr(>|t|)

(Intercept) Type 2 Type 3

Contrainte =0

Type1 :

Type2 :

R² =0,38

library(lsmeans)

## Loading required package: emmeans

## 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'.

lsmeans(mod1,pairwise~Type,data=tab,adjust="bonferroni")

## $lsmeans  
## Type lsmean SE df lower.CL upper.CL  
## 1 157 5.25 51 146 167  
## 2 159 5.69 51 147 170  
## 3 119 5.69 51 107 130  
##   
## Confidence level used: 0.95   
##   
## $contrasts  
## contrast estimate SE df t.ratio p.value  
## 1 - 2 -1.86 7.74 51 -0.240 1.0000   
## 1 - 3 38.09 7.74 51 4.921 <.0001   
## 2 - 3 39.94 8.05 51 4.964 <.0001   
##   
## P value adjustment: bonferroni method for 3 tests

lsmean : nombre moyen de catégorie pour chaque type

contrasts

contrast estimate SE df z.ratio p.value (proba critique avec =5%)

1-2

1-3

2-3