Initiation à la statistique avec R, code et compléments chapitre 10

Frédéric Bertrand et Myriam Maumy-Bertrand

11 décembre 2018

#Chapitre 10  
require(BioStatR)

## Loading required package: BioStatR

#page 402  
foret<-rep(1:3,c(10,10,10))  
hauteur<-c(23.4,24.4,24.6,24.9,25,26.2,26.1,24.8,25.5,25.8,18.9,21.1,21.1,  
 22.1,22.5,23.5,22.7,21.3,22.2,21.7,22.5,22.9,23.7,24,24,24.5,24.3,24.2,  
 23.4,23.9)  
foret<-factor(foret)  
arbre<-data.frame(foret,hauteur)  
rm(foret)  
rm(hauteur)  
arbre

## foret hauteur  
## 1 1 23.4  
## 2 1 24.4  
## 3 1 24.6  
## 4 1 24.9  
## 5 1 25.0  
## 6 1 26.2  
## 7 1 26.1  
## 8 1 24.8  
## 9 1 25.5  
## 10 1 25.8  
## 11 2 18.9  
## 12 2 21.1  
## 13 2 21.1  
## 14 2 22.1  
## 15 2 22.5  
## 16 2 23.5  
## 17 2 22.7  
## 18 2 21.3  
## 19 2 22.2  
## 20 2 21.7  
## 21 3 22.5  
## 22 3 22.9  
## 23 3 23.7  
## 24 3 24.0  
## 25 3 24.0  
## 26 3 24.5  
## 27 3 24.3  
## 28 3 24.2  
## 29 3 23.4  
## 30 3 23.9

moyennes<-tapply(arbre$hauteur,arbre$foret,mean)  
moyennes

## 1 2 3   
## 25.07 21.71 23.74

#page 403  
variances<-tapply(arbre$hauteur,arbre$foret,var)  
variances

## 1 2 3   
## 0.7356667 1.5565556 0.4026667

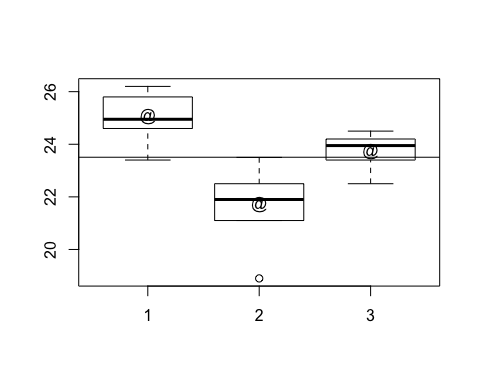
#page 404  
moy.g<-mean(arbre$hauteur)  
moy.g

## [1] 23.50667

mean(moyennes)

## [1] 23.50667

#page 405  
plot(arbre$foret,arbre$hauteur)  
points(1:3,moyennes,pch="@")  
abline(h=moy.g)



pdf("ch11fig101.pdf")  
plot(arbre$foret,arbre$hauteur)  
points(1:3,moyennes,pch="@")  
abline(h=moy.g)  
dev.off()

## quartz\_off\_screen   
## 2

#page 408  
options(contrasts=c("contr.sum","contr.poly"))  
modele1<-lm(hauteur~foret,data=arbre)  
anova(modele1)

## Analysis of Variance Table  
##   
## Response: hauteur  
## Df Sum Sq Mean Sq F value Pr(>F)   
## foret 2 57.265 28.6323 31.874 7.809e-08 \*\*\*  
## Residuals 27 24.254 0.8983   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

modele1\_aov<-aov(hauteur~foret,data=arbre)  
summary(modele1\_aov)

## Df Sum Sq Mean Sq F value Pr(>F)   
## foret 2 57.26 28.632 31.87 7.81e-08 \*\*\*  
## Residuals 27 24.25 0.898   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

#page 409  
options(contrasts=c("contr.sum","contr.poly"))  
  
#page 410  
residus<-residuals(modele1)  
shapiro.test(residus)

##   
## Shapiro-Wilk normality test  
##   
## data: residus  
## W = 0.95202, p-value = 0.1914

length(residus)

## [1] 30

#En plus : les résidus des deux modèles sont égaux  
all(residuals(modele1)==residuals(modele1\_aov))

## [1] TRUE

#page 412  
bartlett.test(residus~foret,data=arbre)

##   
## Bartlett test of homogeneity of variances  
##   
## data: residus by foret  
## Bartlett's K-squared = 3.8798, df = 2, p-value = 0.1437

coef(modele1)

## (Intercept) foret1 foret2   
## 23.506667 1.563333 -1.796667

#En plus : les coefficients des deux modèles sont égaux  
all(coef(modele1)==coef(modele1\_aov))

## [1] TRUE

#page 413  
-sum(coef(modele1)[2:3])

## [1] 0.2333333

dummy.coef(modele1)

## Full coefficients are   
##   
## (Intercept): 23.50667   
## foret: 1 2 3  
## 1.5633333 -1.7966667 0.2333333

#En plus : fonctionne aussi avec le modèle aov et introduction de la   
#fonction model.tables  
dummy.coef(modele1\_aov)

## Full coefficients are   
##   
## (Intercept): 23.50667   
## foret: 1 2 3  
## 1.5633333 -1.7966667 0.2333333

model.tables(modele1\_aov)

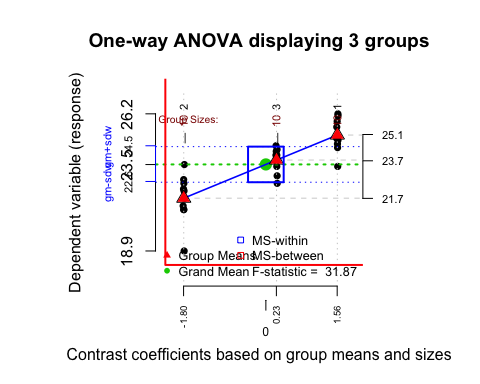
## Tables of effects  
##   
## foret   
## foret  
## 1 2 3   
## 1.5633 -1.7967 0.2333

if(!("granova" %in% rownames(installed.packages()))){  
 install.packages("granova")}  
library(granova)

## Loading required package: car

## Loading required package: carData

granova.1w(arbre$hauteur,arbre$foret)



## $grandsum  
## Grandmean df.bet df.with MS.bet MS.with   
## 23.51 2.00 27.00 28.63 0.90   
## F.stat F.prob SS.bet/SS.tot   
## 31.87 0.00 0.70   
##   
## $stats  
## Size Contrast Coef Wt'd Mean Mean Trim'd Mean Var. St. Dev.  
## 2 10 -1.80 21.71 21.71 21.82 1.56 1.25  
## 3 10 0.23 23.74 23.74 23.87 0.40 0.63  
## 1 10 1.56 25.07 25.07 25.10 0.74 0.86

pdf("chap10fig102.pdf")  
print(granova.1w(arbre$hauteur,arbre$foret))

## $grandsum  
## Grandmean df.bet df.with MS.bet MS.with   
## 23.51 2.00 27.00 28.63 0.90   
## F.stat F.prob SS.bet/SS.tot   
## 31.87 0.00 0.70   
##   
## $stats  
## Size Contrast Coef Wt'd Mean Mean Trim'd Mean Var. St. Dev.  
## 2 10 -1.80 21.71 21.71 21.82 1.56 1.25  
## 3 10 0.23 23.74 23.74 23.87 0.40 0.63  
## 1 10 1.56 25.07 25.07 25.10 0.74 0.86

dev.off()

## quartz\_off\_screen   
## 2

#page 415  
if(!("granovaGG" %in% rownames(installed.packages()))){  
 install.packages("granovaGG")}  
library(granovaGG)

## Loading required package: ggplot2

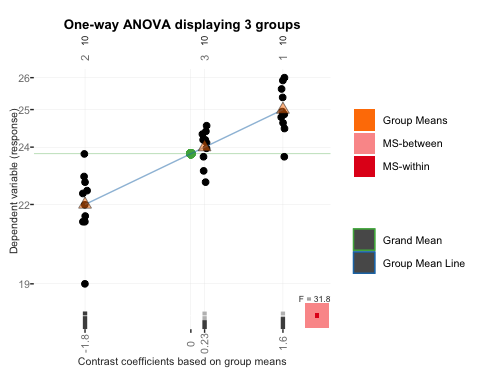
granovagg.1w(arbre$hauteur,arbre$foret)

##   
## By-group summary statistics for your input data (ordered by group means)

## group group.mean trimmed.mean contrast variance standard.deviation  
## 2 2 21.71 21.82 -1.80 1.56 1.25  
## 3 3 23.74 23.87 0.23 0.40 0.63  
## 1 1 25.07 25.10 1.56 0.74 0.86  
## group.size  
## 2 10  
## 3 10  
## 1 10

##   
## Below is a linear model summary of your input data

##   
## Call:  
## lm(formula = score ~ group, data = owp$data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -2.8100 -0.4550 0.0750 0.5425 1.7900   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 23.5067 0.1730 135.844 < 2e-16 \*\*\*  
## group1 1.5633 0.2447 6.388 7.66e-07 \*\*\*  
## group2 -1.7967 0.2447 -7.342 6.75e-08 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.9478 on 27 degrees of freedom  
## Multiple R-squared: 0.7025, Adjusted R-squared: 0.6804   
## F-statistic: 31.87 on 2 and 27 DF, p-value: 7.809e-08



pdf("chap10fig103.pdf")  
print(granovagg.1w(arbre$hauteur,arbre$foret))

##   
## By-group summary statistics for your input data (ordered by group means)

## group group.mean trimmed.mean contrast variance standard.deviation  
## 2 2 21.71 21.82 -1.80 1.56 1.25  
## 3 3 23.74 23.87 0.23 0.40 0.63  
## 1 1 25.07 25.10 1.56 0.74 0.86  
## group.size  
## 2 10  
## 3 10  
## 1 10

##   
## Below is a linear model summary of your input data

##   
## Call:  
## lm(formula = score ~ group, data = owp$data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -2.8100 -0.4550 0.0750 0.5425 1.7900   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 23.5067 0.1730 135.844 < 2e-16 \*\*\*  
## group1 1.5633 0.2447 6.388 7.66e-07 \*\*\*  
## group2 -1.7967 0.2447 -7.342 6.75e-08 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.9478 on 27 degrees of freedom  
## Multiple R-squared: 0.7025, Adjusted R-squared: 0.6804   
## F-statistic: 31.87 on 2 and 27 DF, p-value: 7.809e-08

dev.off()

## quartz\_off\_screen   
## 2

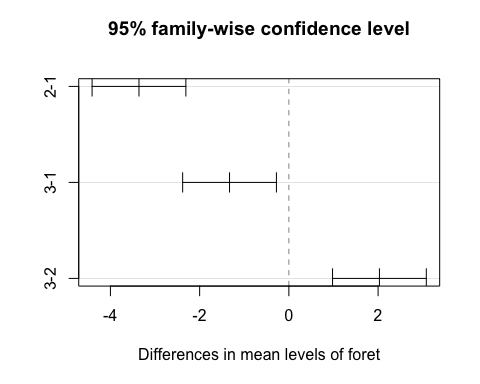
#page 418  
modele2<-aov(hauteur~foret,data=arbre)  
model.tables(modele2)

## Tables of effects  
##   
## foret   
## foret  
## 1 2 3   
## 1.5633 -1.7967 0.2333

TukeyHSD(modele2)

## Tukey multiple comparisons of means  
## 95% family-wise confidence level  
##   
## Fit: aov(formula = hauteur ~ foret, data = arbre)  
##   
## $foret  
## diff lwr upr p adj  
## 2-1 -3.36 -4.4109317 -2.3090683 0.0000000  
## 3-1 -1.33 -2.3809317 -0.2790683 0.0110330  
## 3-2 2.03 0.9790683 3.0809317 0.0001544

plot(TukeyHSD(modele2))



pdf("chap10fig104.pdf")  
plot(TukeyHSD(modele2))  
dev.off()

## quartz\_off\_screen   
## 2

#En plus : export des graphiques en niveaux de gris et aux formats .png ou .ps  
png("chap10fig102.png")  
granova.1w(arbre$hauteur,arbre$foret)

## $grandsum  
## Grandmean df.bet df.with MS.bet MS.with   
## 23.51 2.00 27.00 28.63 0.90   
## F.stat F.prob SS.bet/SS.tot   
## 31.87 0.00 0.70   
##   
## $stats  
## Size Contrast Coef Wt'd Mean Mean Trim'd Mean Var. St. Dev.  
## 2 10 -1.80 21.71 21.71 21.82 1.56 1.25  
## 3 10 0.23 23.74 23.74 23.87 0.40 0.63  
## 1 10 1.56 25.07 25.07 25.10 0.74 0.86

dev.off()

## quartz\_off\_screen   
## 2

postscript("chap10fig102.ps")  
granova.1w(arbre$hauteur,arbre$foret)

## $grandsum  
## Grandmean df.bet df.with MS.bet MS.with   
## 23.51 2.00 27.00 28.63 0.90   
## F.stat F.prob SS.bet/SS.tot   
## 31.87 0.00 0.70   
##   
## $stats  
## Size Contrast Coef Wt'd Mean Mean Trim'd Mean Var. St. Dev.  
## 2 10 -1.80 21.71 21.71 21.82 1.56 1.25  
## 3 10 0.23 23.74 23.74 23.87 0.40 0.63  
## 1 10 1.56 25.07 25.07 25.10 0.74 0.86

dev.off()

## quartz\_off\_screen   
## 2

pdf("chap10fig102bw.pdf",colormodel="gray")  
granova.1w(arbre$hauteur,arbre$foret)

## $grandsum  
## Grandmean df.bet df.with MS.bet MS.with   
## 23.51 2.00 27.00 28.63 0.90   
## F.stat F.prob SS.bet/SS.tot   
## 31.87 0.00 0.70   
##   
## $stats  
## Size Contrast Coef Wt'd Mean Mean Trim'd Mean Var. St. Dev.  
## 2 10 -1.80 21.71 21.71 21.82 1.56 1.25  
## 3 10 0.23 23.74 23.74 23.87 0.40 0.63  
## 1 10 1.56 25.07 25.07 25.10 0.74 0.86

dev.off()

## quartz\_off\_screen   
## 2

postscript("chap10fig102bw.ps",colormodel="gray")  
granova.1w(arbre$hauteur,arbre$foret)

## $grandsum  
## Grandmean df.bet df.with MS.bet MS.with   
## 23.51 2.00 27.00 28.63 0.90   
## F.stat F.prob SS.bet/SS.tot   
## 31.87 0.00 0.70   
##   
## $stats  
## Size Contrast Coef Wt'd Mean Mean Trim'd Mean Var. St. Dev.  
## 2 10 -1.80 21.71 21.71 21.82 1.56 1.25  
## 3 10 0.23 23.74 23.74 23.87 0.40 0.63  
## 1 10 1.56 25.07 25.07 25.10 0.74 0.86

dev.off()

## quartz\_off\_screen   
## 2

png("chap10fig103.png")  
granovagg.1w(arbre$hauteur,arbre$foret)

##   
## By-group summary statistics for your input data (ordered by group means)

## group group.mean trimmed.mean contrast variance standard.deviation  
## 2 2 21.71 21.82 -1.80 1.56 1.25  
## 3 3 23.74 23.87 0.23 0.40 0.63  
## 1 1 25.07 25.10 1.56 0.74 0.86  
## group.size  
## 2 10  
## 3 10  
## 1 10

##   
## Below is a linear model summary of your input data

##   
## Call:  
## lm(formula = score ~ group, data = owp$data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -2.8100 -0.4550 0.0750 0.5425 1.7900   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 23.5067 0.1730 135.844 < 2e-16 \*\*\*  
## group1 1.5633 0.2447 6.388 7.66e-07 \*\*\*  
## group2 -1.7967 0.2447 -7.342 6.75e-08 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.9478 on 27 degrees of freedom  
## Multiple R-squared: 0.7025, Adjusted R-squared: 0.6804   
## F-statistic: 31.87 on 2 and 27 DF, p-value: 7.809e-08

dev.off()

## quartz\_off\_screen   
## 2

postscript("chap10fig103.ps")  
granovagg.1w(arbre$hauteur,arbre$foret)

##   
## By-group summary statistics for your input data (ordered by group means)

## group group.mean trimmed.mean contrast variance standard.deviation  
## 2 2 21.71 21.82 -1.80 1.56 1.25  
## 3 3 23.74 23.87 0.23 0.40 0.63  
## 1 1 25.07 25.10 1.56 0.74 0.86  
## group.size  
## 2 10  
## 3 10  
## 1 10

##   
## Below is a linear model summary of your input data

##   
## Call:  
## lm(formula = score ~ group, data = owp$data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -2.8100 -0.4550 0.0750 0.5425 1.7900   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 23.5067 0.1730 135.844 < 2e-16 \*\*\*  
## group1 1.5633 0.2447 6.388 7.66e-07 \*\*\*  
## group2 -1.7967 0.2447 -7.342 6.75e-08 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.9478 on 27 degrees of freedom  
## Multiple R-squared: 0.7025, Adjusted R-squared: 0.6804   
## F-statistic: 31.87 on 2 and 27 DF, p-value: 7.809e-08

## Warning in grid.Call.graphics(C\_segments, x$x0, x$y0, x$x1, x$y1, x$arrow):  
## semi-transparency is not supported on this device: reported only once per  
## page

dev.off()

## quartz\_off\_screen   
## 2

pdf("chap10fig103bw.pdf",colormodel="gray")  
granovagg.1w(arbre$hauteur,arbre$foret)

##   
## By-group summary statistics for your input data (ordered by group means)

## group group.mean trimmed.mean contrast variance standard.deviation  
## 2 2 21.71 21.82 -1.80 1.56 1.25  
## 3 3 23.74 23.87 0.23 0.40 0.63  
## 1 1 25.07 25.10 1.56 0.74 0.86  
## group.size  
## 2 10  
## 3 10  
## 1 10

##   
## Below is a linear model summary of your input data

##   
## Call:  
## lm(formula = score ~ group, data = owp$data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -2.8100 -0.4550 0.0750 0.5425 1.7900   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 23.5067 0.1730 135.844 < 2e-16 \*\*\*  
## group1 1.5633 0.2447 6.388 7.66e-07 \*\*\*  
## group2 -1.7967 0.2447 -7.342 6.75e-08 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.9478 on 27 degrees of freedom  
## Multiple R-squared: 0.7025, Adjusted R-squared: 0.6804   
## F-statistic: 31.87 on 2 and 27 DF, p-value: 7.809e-08

dev.off()

## quartz\_off\_screen   
## 2

postscript("chap10fig103bw.ps",colormodel="gray")  
granovagg.1w(arbre$hauteur,arbre$foret)

##   
## By-group summary statistics for your input data (ordered by group means)

## group group.mean trimmed.mean contrast variance standard.deviation  
## 2 2 21.71 21.82 -1.80 1.56 1.25  
## 3 3 23.74 23.87 0.23 0.40 0.63  
## 1 1 25.07 25.10 1.56 0.74 0.86  
## group.size  
## 2 10  
## 3 10  
## 1 10

##   
## Below is a linear model summary of your input data

##   
## Call:  
## lm(formula = score ~ group, data = owp$data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -2.8100 -0.4550 0.0750 0.5425 1.7900   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 23.5067 0.1730 135.844 < 2e-16 \*\*\*  
## group1 1.5633 0.2447 6.388 7.66e-07 \*\*\*  
## group2 -1.7967 0.2447 -7.342 6.75e-08 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.9478 on 27 degrees of freedom  
## Multiple R-squared: 0.7025, Adjusted R-squared: 0.6804   
## F-statistic: 31.87 on 2 and 27 DF, p-value: 7.809e-08

## Warning in grid.Call.graphics(C\_segments, x$x0, x$y0, x$x1, x$y1, x$arrow):  
## semi-transparency is not supported on this device: reported only once per  
## page

dev.off()

## quartz\_off\_screen   
## 2

#page 425  
#Exercice 10.1  
#1)  
options(contrasts=c(unordered="contr.sum", ordered="contr.poly"))  
  
#page 426  
#2)  
variete<-rep(1:6,c(5,5,5,5,5,5))  
vitamine<-c(93.6,95.3,96,93.7,96.2,95.3,96.9,95.8,97.3,97.7,94.5,97,97.8,97,  
98.3,98.8,98.2,97.8,97.2,97.9,94.6,97.8,98,95,98.9,93.2,94.4,93.8,95.6,94.8)  
variete<-factor(variete)  
exo1<-data.frame(variete,vitamine)  
modele1<-aov(vitamine~variete,data=exo1)  
residus1<-residuals(modele1)  
shapiro.test(residus1)

##   
## Shapiro-Wilk normality test  
##   
## data: residus1  
## W = 0.9563, p-value = 0.2485

length(residus1)

## [1] 30

bartlett.test(residus1~variete,data=exo1)

##   
## Bartlett test of homogeneity of variances  
##   
## data: residus1 by variete  
## Bartlett's K-squared = 5.6023, df = 5, p-value = 0.3469

#page 427  
#3)  
modele1

## Call:  
## aov(formula = vitamine ~ variete, data = exo1)  
##   
## Terms:  
## variete Residuals  
## Sum of Squares 45.836 38.512  
## Deg. of Freedom 5 24  
##   
## Residual standard error: 1.266754  
## Estimated effects may be unbalanced

summary(modele1)

## Df Sum Sq Mean Sq F value Pr(>F)   
## variete 5 45.84 9.167 5.713 0.00131 \*\*  
## Residuals 24 38.51 1.605   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

#page 428  
#4)  
granovagg.1w(vitamine,group=variete)

##   
## By-group summary statistics for your input data (ordered by group means)

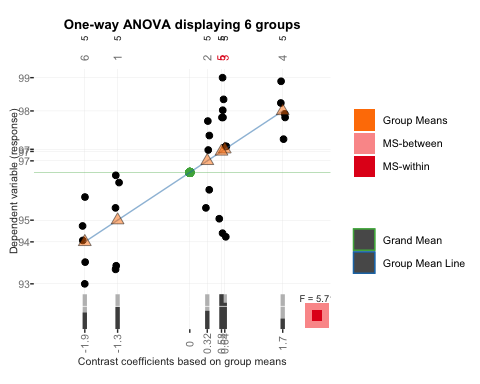
## group group.mean trimmed.mean contrast variance standard.deviation  
## 6 6 94.36 94.33 -1.92 0.85 0.92  
## 1 1 94.96 95.00 -1.32 1.54 1.24  
## 2 2 96.60 96.67 0.32 1.03 1.01  
## 5 5 96.86 96.93 0.58 3.73 1.93  
## 3 3 96.92 97.27 0.64 2.14 1.46  
## 4 4 97.98 97.97 1.70 0.34 0.58  
## group.size  
## 6 5  
## 1 5  
## 2 5  
## 5 5  
## 3 5  
## 4 5

##   
## The following groups are likely to be overplotted

## group group.mean contrast  
## 5 5 96.86 0.58  
## 3 3 96.92 0.64

##   
## Below is a linear model summary of your input data

##   
## Call:  
## lm(formula = score ~ group, data = owp$data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -2.420 -0.795 0.150 0.925 2.040   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 96.2800 0.2313 416.298 < 2e-16 \*\*\*  
## group1 -1.3200 0.5172 -2.552 0.01748 \*   
## group2 0.3200 0.5172 0.619 0.54189   
## group3 0.6400 0.5172 1.238 0.22785   
## group4 1.7000 0.5172 3.287 0.00311 \*\*   
## group5 0.5800 0.5172 1.122 0.27316   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 1.267 on 24 degrees of freedom  
## Multiple R-squared: 0.5434, Adjusted R-squared: 0.4483   
## F-statistic: 5.713 on 5 and 24 DF, p-value: 0.001311



pdf("chap10fig105.pdf")  
granovagg.1w(vitamine,group=variete)

##   
## By-group summary statistics for your input data (ordered by group means)

## group group.mean trimmed.mean contrast variance standard.deviation  
## 6 6 94.36 94.33 -1.92 0.85 0.92  
## 1 1 94.96 95.00 -1.32 1.54 1.24  
## 2 2 96.60 96.67 0.32 1.03 1.01  
## 5 5 96.86 96.93 0.58 3.73 1.93  
## 3 3 96.92 97.27 0.64 2.14 1.46  
## 4 4 97.98 97.97 1.70 0.34 0.58  
## group.size  
## 6 5  
## 1 5  
## 2 5  
## 5 5  
## 3 5  
## 4 5

##   
## The following groups are likely to be overplotted

## group group.mean contrast  
## 5 5 96.86 0.58  
## 3 3 96.92 0.64

##   
## Below is a linear model summary of your input data

##   
## Call:  
## lm(formula = score ~ group, data = owp$data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -2.420 -0.795 0.150 0.925 2.040   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 96.2800 0.2313 416.298 < 2e-16 \*\*\*  
## group1 -1.3200 0.5172 -2.552 0.01748 \*   
## group2 0.3200 0.5172 0.619 0.54189   
## group3 0.6400 0.5172 1.238 0.22785   
## group4 1.7000 0.5172 3.287 0.00311 \*\*   
## group5 0.5800 0.5172 1.122 0.27316   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 1.267 on 24 degrees of freedom  
## Multiple R-squared: 0.5434, Adjusted R-squared: 0.4483   
## F-statistic: 5.713 on 5 and 24 DF, p-value: 0.001311

dev.off()

## quartz\_off\_screen   
## 2

#page 430  
#6)  
Tukey1 <- TukeyHSD(modele1, conf.level = 0.95)  
Tukey1

## Tukey multiple comparisons of means  
## 95% family-wise confidence level  
##   
## Fit: aov(formula = vitamine ~ variete, data = exo1)  
##   
## $variete  
## diff lwr upr p adj  
## 2-1 1.64 -0.8371495 4.11714947 0.3468489  
## 3-1 1.96 -0.5171495 4.43714947 0.1803668  
## 4-1 3.02 0.5428505 5.49714947 0.0107386  
## 5-1 1.90 -0.5771495 4.37714947 0.2058535  
## 6-1 -0.60 -3.0771495 1.87714947 0.9733815  
## 3-2 0.32 -2.1571495 2.79714947 0.9985151  
## 4-2 1.38 -1.0971495 3.85714947 0.5310572  
## 5-2 0.26 -2.2171495 2.73714947 0.9994551  
## 6-2 -2.24 -4.7171495 0.23714947 0.0926651  
## 4-3 1.06 -1.4171495 3.53714947 0.7697394  
## 5-3 -0.06 -2.5371495 2.41714947 0.9999996  
## 6-3 -2.56 -5.0371495 -0.08285053 0.0399329  
## 5-4 -1.12 -3.5971495 1.35714947 0.7278111  
## 6-4 -3.62 -6.0971495 -1.14285053 0.0017510  
## 6-5 -2.50 -4.9771495 -0.02285053 0.0470143

#page 431  
#4)  
if(!("multcomp" %in% rownames(installed.packages()))){  
 install.packages("multcomp")}  
library(multcomp)

## Loading required package: mvtnorm

## Loading required package: survival

## Loading required package: TH.data

## Loading required package: MASS

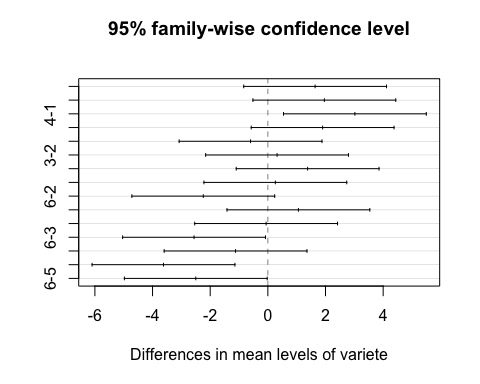
##   
## Attaching package: 'TH.data'

## The following object is masked from 'package:MASS':  
##   
## geyser

wht = glht(modele1, linfct = mcp(variete = "Tukey"))  
cld(wht)

## 1 2 3 4 5 6   
## "ab" "ac" "bc" "c" "bc" "a"

plot(Tukey1)



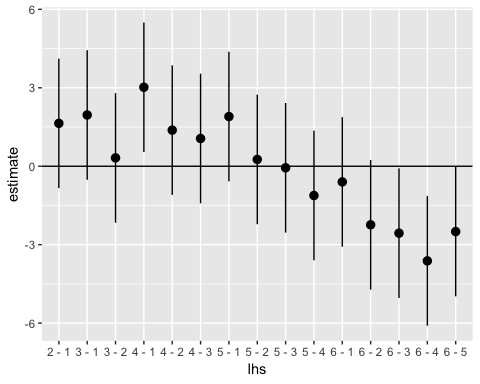
pdf("chap10fig106.pdf")  
plot(Tukey1)  
dev.off()

## quartz\_off\_screen   
## 2

#page 432  
CI <- confint(wht)  
fortify(CI)

## lhs rhs estimate lwr upr  
## 1 2 - 1 0 1.64 -0.8382462 4.11824617  
## 2 3 - 1 0 1.96 -0.5182462 4.43824617  
## 3 4 - 1 0 3.02 0.5417538 5.49824617  
## 4 5 - 1 0 1.90 -0.5782462 4.37824617  
## 5 6 - 1 0 -0.60 -3.0782462 1.87824617  
## 6 3 - 2 0 0.32 -2.1582462 2.79824617  
## 7 4 - 2 0 1.38 -1.0982462 3.85824617  
## 8 5 - 2 0 0.26 -2.2182462 2.73824617  
## 9 6 - 2 0 -2.24 -4.7182462 0.23824617  
## 10 4 - 3 0 1.06 -1.4182462 3.53824617  
## 11 5 - 3 0 -0.06 -2.5382462 2.41824617  
## 12 6 - 3 0 -2.56 -5.0382462 -0.08175383  
## 13 5 - 4 0 -1.12 -3.5982462 1.35824617  
## 14 6 - 4 0 -3.62 -6.0982462 -1.14175383  
## 15 6 - 5 0 -2.50 -4.9782462 -0.02175383

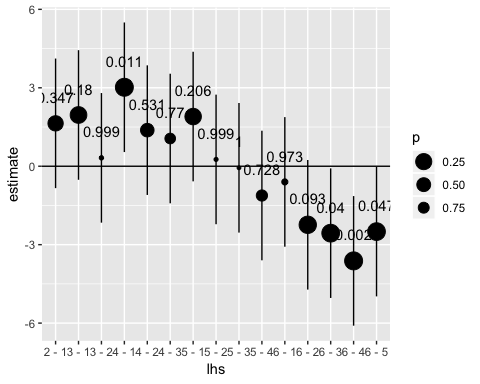
ggplot(CI,aes(lhs,estimate,ymin=lwr,ymax=upr))+geom\_pointrange()+  
 geom\_hline(yintercept = 0)



pdf("chap10fig107.pdf")  
print(ggplot(CI,aes(lhs,estimate,ymin=lwr,ymax=upr))+geom\_pointrange()+  
 geom\_hline(yintercept = 0))  
dev.off()

## quartz\_off\_screen   
## 2

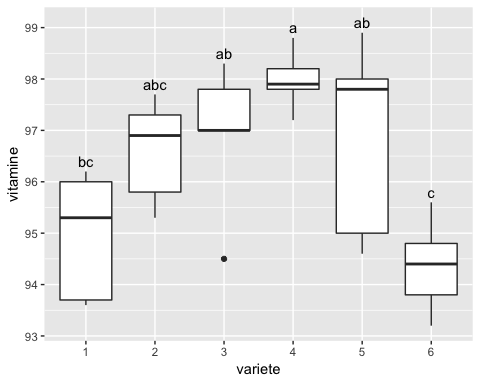
ggplot(aes(lhs,estimate),data=fortify(summary(wht))) +   
 geom\_linerange(aes(ymin=lwr,ymax=upr),data=CI) +   
 geom\_text(aes(y=estimate+1,label=round(p,3)))+geom\_hline(yintercept = 0) +  
 geom\_point(aes(size=p),data=summary(wht)) +scale\_size(trans="reverse")



pdf("chap10fig108.pdf")  
ggplot(aes(lhs,estimate),data=fortify(summary(wht))) +   
 geom\_linerange(aes(ymin=lwr,ymax=upr),data=CI) +   
 geom\_text(aes(y=estimate+1,label=round(p,3)))+geom\_hline(yintercept = 0) +  
 geom\_point(aes(size=p),data=summary(wht)) +scale\_size(trans="reverse")  
dev.off()

## quartz\_off\_screen   
## 2

#page 433  
if(!("multcompView" %in% rownames(installed.packages()))){  
 install.packages("multcompView")}  
library(multcompView)  
if(!("plyr" %in% rownames(installed.packages()))){install.packages("plyr")}  
library(plyr)  
generate\_label\_df <- function(HSD,flev){  
Tukey.levels <- HSD[[flev]][,4]  
Tukey.labels <- multcompLetters(Tukey.levels)['Letters']  
plot.labels <- names(Tukey.labels[['Letters']])  
boxplot.df <- ddply(exo1, flev, function (x) max(fivenum(x$vitamine)) + 0.2)  
plot.levels <- data.frame(plot.labels, labels = Tukey.labels[['Letters']],  
 stringsAsFactors = FALSE)  
labels.df <- merge(plot.levels, boxplot.df, by.x = 'plot.labels', by.y = flev,  
 sort = FALSE)  
return(labels.df)  
}  
  
#page 434  
p\_base <- ggplot(exo1,aes(x=variete,y=vitamine)) + geom\_boxplot() +   
 geom\_text(data = generate\_label\_df(Tukey1, 'variete'), aes(x = plot.labels,   
 y = V1, label = labels))  
p\_base



pdf("chap10fig109.pdf")  
print(p\_base)  
dev.off()

## quartz\_off\_screen   
## 2

#page 435  
#Exercice 10.1  
#2)  
traitement<-rep(1:5,c(7,7,7,7,7))  
taux<-c(4.5,2.5,6,4.5,3,5.5,3.5,7.5,3,2.5,4,2,4,5.5,8,6.5,6,3.5,5,  
 7,5,2,7.5,4,2.5,5,3.5,6.5,6.5,5.5,6,4.5,4,7,5.5)  
traitement<-factor(traitement)  
exo2<-data.frame(traitement,taux)  
modele2<-aov(taux~traitement,data=exo2)  
residus2<-residuals(modele2)  
shapiro.test(residus2)

##   
## Shapiro-Wilk normality test  
##   
## data: residus2  
## W = 0.97436, p-value = 0.5734

length(residus2)

## [1] 35

bartlett.test(residus2~traitement,data=exo2)

##   
## Bartlett test of homogeneity of variances  
##   
## data: residus2 by traitement  
## Bartlett's K-squared = 3.1361, df = 4, p-value = 0.5353

#page 436  
#3)  
modele1<-lm(taux~traitement,data=exo2)  
anova(modele1)

## Analysis of Variance Table  
##   
## Response: taux  
## Df Sum Sq Mean Sq F value Pr(>F)  
## traitement 4 19.043 4.7607 1.8687 0.1419  
## Residuals 30 76.429 2.5476

#4)  
power.anova.test(5,7,19.043,76.42857)

##   
## Balanced one-way analysis of variance power calculation   
##   
## groups = 5  
## n = 7  
## between.var = 19.043  
## within.var = 76.42857  
## sig.level = 0.05  
## power = 0.4684833  
##   
## NOTE: n is number in each group

#page 437  
power.anova.test(groups=5,between.var=19.043,within.var=76.42857,power=.80)

##   
## Balanced one-way analysis of variance power calculation   
##   
## groups = 5  
## n = 12.96035  
## between.var = 19.043  
## within.var = 76.42857  
## sig.level = 0.05  
## power = 0.8  
##   
## NOTE: n is number in each group

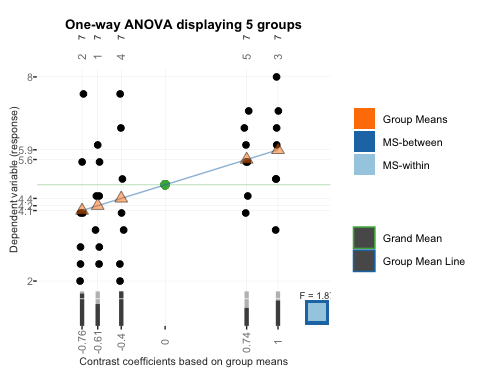
granovagg.1w(taux,group=traitement)

##   
## By-group summary statistics for your input data (ordered by group means)

## group group.mean trimmed.mean contrast variance standard.deviation  
## 2 2 4.07 3.8 -0.76 3.62 1.90  
## 1 1 4.21 4.2 -0.61 1.65 1.29  
## 4 4 4.43 4.3 -0.40 4.12 2.03  
## 5 5 5.57 5.6 0.74 1.12 1.06  
## 3 3 5.86 5.9 1.03 2.23 1.49  
## group.size  
## 2 7  
## 1 7  
## 4 7  
## 5 7  
## 3 7

##   
## Below is a linear model summary of your input data

##   
## Call:  
## lm(formula = score ~ group, data = owp$data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -2.4286 -1.0714 -0.0714 1.0357 3.4286   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 4.8286 0.2698 17.897 <2e-16 \*\*\*  
## group1 -0.6143 0.5396 -1.138 0.2639   
## group2 -0.7571 0.5396 -1.403 0.1708   
## group3 1.0286 0.5396 1.906 0.0662 .   
## group4 -0.4000 0.5396 -0.741 0.4643   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 1.596 on 30 degrees of freedom  
## Multiple R-squared: 0.1995, Adjusted R-squared: 0.09272   
## F-statistic: 1.869 on 4 and 30 DF, p-value: 0.1419



pdf("chap10fig1010.pdf",colormodel="gray")  
granovagg.1w(taux,group=traitement)

##   
## By-group summary statistics for your input data (ordered by group means)

## group group.mean trimmed.mean contrast variance standard.deviation  
## 2 2 4.07 3.8 -0.76 3.62 1.90  
## 1 1 4.21 4.2 -0.61 1.65 1.29  
## 4 4 4.43 4.3 -0.40 4.12 2.03  
## 5 5 5.57 5.6 0.74 1.12 1.06  
## 3 3 5.86 5.9 1.03 2.23 1.49  
## group.size  
## 2 7  
## 1 7  
## 4 7  
## 5 7  
## 3 7

##   
## Below is a linear model summary of your input data

##   
## Call:  
## lm(formula = score ~ group, data = owp$data)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -2.4286 -1.0714 -0.0714 1.0357 3.4286   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 4.8286 0.2698 17.897 <2e-16 \*\*\*  
## group1 -0.6143 0.5396 -1.138 0.2639   
## group2 -0.7571 0.5396 -1.403 0.1708   
## group3 1.0286 0.5396 1.906 0.0662 .   
## group4 -0.4000 0.5396 -0.741 0.4643   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 1.596 on 30 degrees of freedom  
## Multiple R-squared: 0.1995, Adjusted R-squared: 0.09272   
## F-statistic: 1.869 on 4 and 30 DF, p-value: 0.1419

dev.off()

## quartz\_off\_screen   
## 2