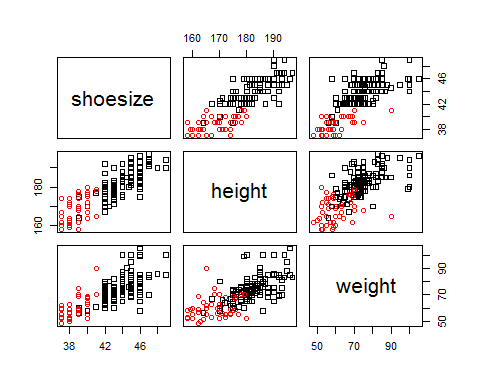
Exercise1.R

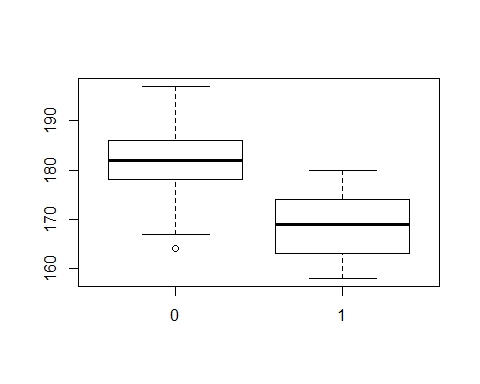
B-C-Herbert

2019-11-20

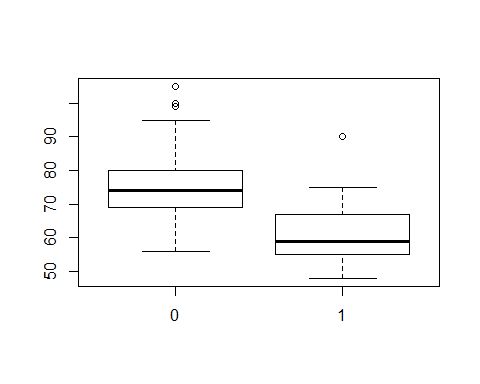
### Exercise 1  
  
remove(list = ls())  
  
students2008 = read.table("students2008.txt", header = TRUE, dec = ",")  
students2008 = students2008[c(1,4,35,36)]  
students2008 = na.omit(students2008)  
attach(students2008)  
  
# a)  
  
pairs(students2008[,2:4], col = sex+1, pch = sex)



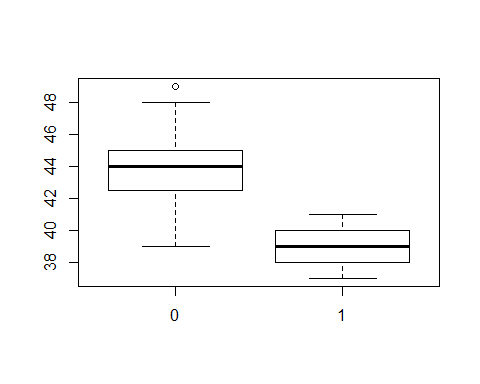
# b)   
  
boxplot(height ~ sex)



boxplot(weight ~ sex)



boxplot(shoesize ~ sex)



female = subset(students2008, sex == 1)  
male = subset(students2008, sex == 0)  
  
# c)  
  
colMeans(male)

## sex shoesize height weight   
## 0.00000 43.97000 182.57273 74.98182

colMeans(female)

## sex shoesize height weight   
## 1.00000 38.76087 168.47826 60.41304

cov(male[,2:4])

## shoesize height weight  
## shoesize 3.521936 8.614587 10.40954  
## height 8.614587 44.485488 37.01968  
## weight 10.409541 37.019683 94.93545

cov(female[,2:4])

## shoesize height weight  
## shoesize 1.474879 4.383575 5.63430  
## height 4.383575 40.655072 12.57585  
## weight 5.634300 12.575845 65.00338

# d)  
  
percMale = (dim(male)[1])/(dim(students2008)[1])  
percMale

## [1] 0.7051282

percFem = (dim(female)[1])/(dim(students2008)[1])  
percFem

## [1] 0.2948718

# e)  
  
plot(height, weight, col = sex+1, pch = sex, xlim = c(150,200), ylim = c(40,110) )  
c2 = qchisq(0.95,2)  
meanmalehei= colMeans(male)[3]  
meanmalewei = colMeans(male)[4]  
meanfemalehei = colMeans(female)[3]  
meanfemalewei = colMeans(female)[4]  
  
Smale = cov(male[,3:4])  
Sfemale = cov(female[,3:4])  
  
SmaleInv = solve(Smale)  
aSM11 = SmaleInv[1,1]  
aSM22 = SmaleInv[2,2]  
aSM12 = SmaleInv[1,2]  
SfemaleInv = solve(Sfemale)  
aSF11 = SfemaleInv[1,1]  
aSF22 = SfemaleInv[2,2]  
aSF12 = SfemaleInv[1,2]  
  
x1 = seq(150,210, le = 1000)  
x2 = seq(40,110, le = 1000)  
f = function(x1, x2)  
{  
 aSM11\*(x1-meanmalehei)^2+aSM22\*(x2-meanmalewei)^2+2\*aSM12\*(x1-meanmalehei)\*(x2-meanmalewei)  
}  
z = outer(x1,x2,f)  
contour(x1, x2, z, asp = 1, levels = c2, add = TRUE, drawlabels = FALSE)  
  
ff = function(x1,x2)  
{   
 aSF11\*(x1-meanfemalehei)^2+aSF22\*(x2-meanfemalewei)^2+2\*aSF12\*(x1-meanfemalehei)\*(x2-meanfemalewei)  
}  
y = outer(x1,x2,ff)  
contour(x1, x2, y, asp = 1, levels = c2, add = TRUE, drawlabels = FALSE)  
  
m = c(meanmalehei,meanmalewei)  
n = length(male[,3])  
sum(mahalanobis(male[,3:4],m,Smale)<=c2)/n

## [1] 0.9181818

# ~91,82%  
  
mf = c(meanfemalehei,meanfemalewei)  
nf = length(female[,3])  
sum(mahalanobis(female[,3:4],mf,Sfemale)<=c2)/nf

## [1] 0.9782609

# ~97,83%  
  
# f)  
X = as.matrix(students2008[,3:4])  
Sp = ((109)/(154))\*Smale+((45)/(154))\*Sfemale   
x1bar = matrix(mf,2,1)  
x2bar = matrix(m,2,1)  
  
lhs = X%\*%solve(Sp)%\*%(x1bar-x2bar)-(0.5\*t(x1bar-x2bar)%\*%solve(Sp)%\*%(x1bar+x2bar))[1,1]  
rhs = log(percMale/percFem)  
pi1 = c(lhs>=rhs)  
  
countFemasMale = c(students2008[1] == 1 & pi1 == FALSE)  
countFemasFemale = c(students2008[1] == 1 & pi1 == TRUE)  
countMaleasFem = c(students2008[1] == 0 & pi1 == TRUE)  
countMaleasMale = c(students2008[1] == 0 & pi1 == FALSE)  
  
sum(countMaleasFem)

## [1] 6

sum(countMaleasMale)

## [1] 104

sum(countFemasFemale)

## [1] 36

sum(countFemasMale)

## [1] 10

APER = ((sum(countFemasMale)+sum(countMaleasFem))/dim(students2008)[1])  
APER

## [1] 0.1025641

# ~ 10.26%  
  
# g)  
X2 = as.matrix(students2008[,2:4])  
Smale2 = cov(male[,2:4])  
Sfemale2 = cov(female[,2:4])  
meanfemalesho = colMeans(female)[2]  
meanmalesho = colMeans(male)[2]  
m2 = c(meanmalesho,meanmalehei,meanmalewei)  
mf2 = c(meanfemalesho,meanfemalehei,meanfemalewei)  
x1bar2 = matrix(mf2,3,1)  
x2bar2 = matrix(m2,3,1)  
Sp2 = ((109)/(154))\*Smale2+((45)/(154))\*Sfemale2  
lhs2 = X2%\*%solve(Sp2)%\*%(x1bar2-x2bar2)-(0.5\*t(x1bar2-x2bar2)%\*%solve(Sp2)%\*%(x1bar2+x2bar2))[1,1]  
rhs = log(percMale/percFem)  
  
pi12 = c(lhs2<=rhs)  
  
countFemasMale2 = c(students2008[1] == 1 & pi12 == FALSE)  
countFemasFemale2 = c(students2008[1] == 1 & pi12 == TRUE)  
countMaleasFem2 = c(students2008[1] == 0 & pi12 == TRUE)  
countMaleasMale2 = c(students2008[1] == 0 & pi12 == FALSE)  
  
sum(countMaleasFem2)

## [1] 107

sum(countMaleasMale2)

## [1] 3

sum(countFemasFemale2)

## [1] 2

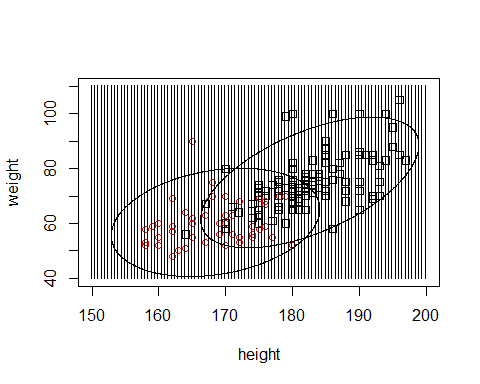
sum(countFemasMale2)

## [1] 44

APER2 = (2+10)/dim(students2008)[1]  
APER2

## [1] 0.07692308

# ~ 7,69%  
  
### Exercise 2  
  
require(utils)  
points(expand.grid(height = seq(150,200, 0.5), weigth = seq(40, 110, 0.5)), cex = 0.1)



### Exercise 3  
  
# f)  
  
students2009 = read.table("students2009.txt", header = TRUE, dec = ",")  
students2009 = students2009[c(1,4,35,36)]  
students2009 = na.omit(students2009)  
X3 = as.matrix(students2009[,3:4])  
X3 = matrix(as.numeric(X3), 130,2)  
  
lhs3 = X3%\*%solve(Sp)%\*%(x1bar-x2bar)-(0.5\*t(x1bar-x2bar)%\*%solve(Sp)%\*%(x1bar+x2bar))[1,1]  
rhs = log(percMale/percFem)  
pi3 = c(lhs3>=rhs)  
  
countFemasMale3 = c(students2009[1] == 1 & pi3 == FALSE)  
countFemasFemale3 = c(students2009[1] == 1 & pi3 == TRUE)  
countMaleasFem3 = c(students2009[1] == 0 & pi3 == TRUE)  
countMaleasMale3 = c(students2009[1] == 0 & pi3 == FALSE)  
  
sum(countMaleasFem3)

## [1] 7

# Anzahl der Männer die Frauen zugeordnet wurden: 7  
sum(countMaleasMale3)

## [1] 93

# Anzahl der Männer die Männern zugeordnet wurden: 93  
sum(countFemasFemale3)

## [1] 20

# Anzahl der Frauen die Frauen zugeordnet wurden: 20  
sum(countFemasMale3)

## [1] 10

# Anzahl der Frauen die Männern zugeordnet wurden: 10  
APER3 = ((sum(countFemasMale3)+sum(countMaleasFem3))/dim(students2009)[1])  
APER3

## [1] 0.1307692

# g)   
  
X4 = as.matrix(students2009[,2:4])  
X4 = matrix(as.numeric(X3), 130,3)

## Warning in matrix(as.numeric(X3), 130, 3): Datenlänge [260] ist kein Teiler  
## oder Vielfaches der Anzahl der Spalten [3]

lhs4 = X4%\*%solve(Sp2)%\*%(x1bar2-x2bar2)-(0.5\*t(x1bar2-x2bar2)%\*%solve(Sp2)%\*%(x1bar2+x2bar2))[1,1]  
pi4 = c(lhs4>=rhs)  
  
countFemasMale4 = c(students2009[1] == 1 & pi4 == FALSE)  
countFemasFemale4 = c(students2009[1] == 1 & pi4 == TRUE)  
countMaleasFem4 = c(students2009[1] == 0 & pi4 == TRUE)  
countMaleasMale4 = c(students2009[1] == 0 & pi4 == FALSE)  
  
sum(countMaleasFem4)

## [1] 0

sum(countMaleasMale4)

## [1] 100

sum(countFemasFemale4)

## [1] 0

sum(countFemasMale4)

## [1] 30

APER4 = ((sum(countFemasMale4)+sum(countMaleasFem4))/dim(students2009)[1])  
APER4

## [1] 0.2307692