hw\_2.R

Nadim

Wed Jan 23 22:40:28 2019

dat<- read.table("https://www.stat.ncsu.edu/people/maity/courses/st537-S2019/data/T4-3.DAT", header=F)  
colnames(dat) <- c("x1", "x2", "x3", "x4", "d2")  
n <- nrow(dat)  
p <- ncol(dat)  
head(dat)

## x1 x2 x3 x4 d2  
## 1 1889 1651 1561 1778 0.60  
## 2 2403 2048 2087 2197 5.48  
## 3 2119 1700 1815 2222 7.62  
## 4 1645 1627 1110 1533 5.21  
## 5 1976 1916 1614 1883 1.40  
## 6 1712 1712 1439 1546 2.22

######Question 1#######  
i<-which(dat$d2==max(dat$d2))  
dat <- dat[-c(i),]  
j<-which(dat$d2==max(dat$d2))  
dat<- dat[-c(j),]  
dat$d2

## [1] 0.60 5.48 7.62 5.21 1.40 2.22 4.99 1.49 0.77 1.93 0.46 2.70 0.13 1.08  
## [15] 3.50 3.99 1.36 1.46 9.90 5.06 0.80 2.54 4.58 3.40 2.38 3.00 6.28 2.58

max(dat$d2)

## [1] 9.9

apply(dat[,1:4], 2, shapiro.test)

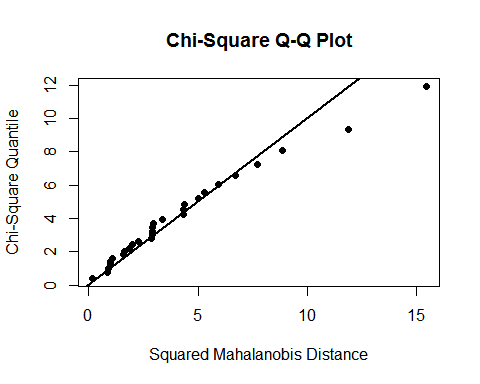
## $x1  
##   
## Shapiro-Wilk normality test  
##   
## data: newX[, i]  
## W = 0.98469, p-value = 0.9439  
##   
##   
## $x2  
##   
## Shapiro-Wilk normality test  
##   
## data: newX[, i]  
## W = 0.96636, p-value = 0.4871  
##   
##   
## $x3  
##   
## Shapiro-Wilk normality test  
##   
## data: newX[, i]  
## W = 0.96717, p-value = 0.507  
##   
##   
## $x4  
##   
## Shapiro-Wilk normality test  
##   
## data: newX[, i]  
## W = 0.96598, p-value = 0.4779

library(MVN)

## Warning: package 'MVN' was built under R version 3.5.2

## sROC 0.1-2 loaded

mvn(dat[, 1:4], mvnTest = "royston", multivariatePlot = "qq")



## $multivariateNormality  
## Test H p value MVN  
## 1 Royston 1.098338 0.6271166 YES  
##   
## $univariateNormality  
## Test Variable Statistic p value Normality  
## 1 Shapiro-Wilk x1 0.9847 0.9439 YES   
## 2 Shapiro-Wilk x2 0.9664 0.4871 YES   
## 3 Shapiro-Wilk x3 0.9672 0.5070 YES   
## 4 Shapiro-Wilk x4 0.9660 0.4779 YES   
##   
## $Descriptives  
## n Mean Std.Dev Median Min Max 25th 75th Skew  
## x1 28 1865.929 262.1619 1857.5 1325 2403 1711.50 2049.75 0.08994538  
## x2 28 1697.964 244.8618 1663.0 1170 2301 1593.25 1847.50 0.39091767  
## x3 28 1488.643 253.1536 1466.0 1002 2087 1307.25 1617.25 0.49661284  
## x4 28 1710.250 277.9986 1674.5 1176 2234 1528.75 1876.25 0.25921958  
## Kurtosis  
## x1 -0.5084972  
## x2 0.1961808  
## x3 0.0516768  
## x4 -0.6484407

chisquare.plot <- function(x,mark){  
 p <- ncol(x)  
 n<-nrow(x)  
 xbar<- colMeans(x)  
 s <- cov(x)  
   
 ###Mahalonobis Dist  
   
 x.cent <- scale(x,center =T, scale = F)  
 d2<- diag(x.cent%\*%solve(s)%\*%t(x.cent))  
   
 qchi <- qchisq((1:n-0.5)/n, df =p)  
 sortd <- sort(d2)  
   
 plot(qchi,sortd,pch=19,xlab="Chi-squared quantiles",ylab="Mahalonobis squared distance",main="chi-square Q-Q plot")  
   
 points(qchi[(n-mark+1):n],sortd[(n-mark+1):n],cex=3,col="#990000")  
   
 return((sortd[(n-mark+1):n]))  
 }  
  
 a<-chisquare.plot(x=dat[,1:4],mark = 2)  
 a[1]

## 21   
## 11.45119

a[2]

## 3   
## 14.89312

# plot(dat$x1,dat$x2 ,  
 # xlim = c(1000, 3500), ylim = c(800, 3000),   
 # pch=19, col = c("steelblue"))  
 abline(0,1,col="blue")

