FOR 215

Assignment 1 (Give in by Monday 19.9. 14:00)

Exercise 1

a) Form three separate vectors from scalars 4,5 and 33; -48, 0 and 45; 7, 3 and 1

b) Combine vectors into one 3x3 matrix

c) Name rows of the matrix as r1, r2 and r3. Name columns of the matrix as A, B and C

Exercise 2

Generate the following vectors with functions rep() and seq()

a) [0 0 0 0 0 1 1 1 1 1 2 2 2 2 2 3 3 3 3 3 4 4 4 4 4]

b) [1 2 3 4 5 1 2 3 4 5 1 2 3 4 5 1 2 3 4 5 1 2 3 4 5]

Exercise 3

Solar radiation was measured in a greenhouse at eight different times. Observations were 11.1, 10.6, 6.3,

8.8, 10.7, 11.2, 8.9 and 12.2 units.

a) Save measurements as variable solar.radiation (set value with a vector containing all

measurements)

b) Define mean, median and variance of variable solar.radiation (with pre-determined R-functions)

c) Add value 10 to every observation and name the new vector as sr10. Define mean, median and

variance of modified vector. Which of the determined values changed from the previous and how

much?

d) Multiply the original observations by -2 and name the resulting vector as srm2. What happened to

mean, median and variance?

e) Variance is generally determined with following formulas:

ଵ

௡

(ݔ − ௜ݔ) ∑

௡ ଶ

௜ୀଵ and ଵ

௡ିଵ

(ݔ − ௜ݔ) ∑

௡ ଶ

௜ୀଵ

Which of the previous two is used in R-function var()?

Exercise 4

Generate 25 random integers from interval [1,100] and place them in row-wise in 5x5 matrice.

a) Name the rows as r1-r5 and the columns as c1-c5

b) Determine the sum of numbers in the matrice

c) Substract the minimum of the matrice from every other figure in the Matrix

d) Print the figure that is in column 3 of row 4

Exercise 5

Search the internet for information on function getEventProb.

a) To which R-package does the function belong?

b) What does the function actually do?

c) With which parameters can you control the function?

Download the package to your computer and activate it from the R-library. Then, answer to following

questions with help of the tools found in the package.

d) What are the odds that a throw with two 6-sided dice results sum of 7?

e) What are the odds that a throw with two 10-sided dice results two tens? (the sum of dice is 20)?

#clear memory

rm(list = ls())

setwd("C:/Users/oyeda/Desktop/R\_COURSE/assignment4")

#load the data

data<-read.table("puudata.txt", header = T, sep = "\t")

#Exercise 1

#Form separate histograms for diameters (LPM) of spruces and pines

#(tree species = PUULAJI, pine = 1 and spruce = 2) in the first crown layer

#(canopy cover layer = LATVKERROS). Which of the two is visually closer

#to normal distribution?

par(mfcol=c(1,2))

pineLayer1<- subset(data, data$PUULAJI==1 & data$LATVKERROS==1)

pine.hist<- hist(pineLayer1$LPM, main = "Histogram of diameter of Pine",

col = "violet")

spruceLayer1<-subset(data, data$PUULAJI==2 & data$LATVKERROS==1)

spruce.hist<-hist(data$LPM, main = "Histogram of diameter of Spruce",

col="yellow")

#Answer: Pine apperas to be more normally distributed

#Exercise 2

#Plot tree height (PITUUS) as function of tree diameter (LPM)

#(height on the y-axis and dbh on the x-axis)

#a) for spruces in the first canopy cover layer

?plot

plot(spruceLayer1$LPM, spruceLayer1$PITUUS,

xlab = "DBH", ylab = "Height", main="Spruce in Crown Layer 2", col="red")

#plot(x=spruceLayer1$LPM, y=spruceLayer1$PITUUS) alternatively

#b) for pines in the second canopy cover layer

pineLayer2<- subset(data, data$PUULAJI==1 & data$LATVKERROS==2)

plot(pineLayer2$PITUUS, pineLayer2$LPM, main = "Pines in Crown Layer 2",

xlab = "DBH", ylab = "Height", col="brown")

#Exercise 3

#Form a box-and-whiskers plot for heights of conifers (species 1 and 2)

#in first and second crown layer.

?par

par(mfrow=c(2,2), bg="white")

boxplot(spruceLayer1$PITUUS, ylab="height", main="Spruce Crown Layer 1")

spruceLayer2<-subset(data, data$PUULAJI==2 & data$LATVKERROS==2)

boxplot(spruceLayer2$PITUUS, ylab="height", main="Spruce Crown Layer 2")

boxplot(pineLayer1$PITUUS, ylab="height", main="Pine Crown Layer 1")

boxplot(pineLayer2$PITUUS, ylab="height", main="Pine Crown Layer 2")

#What can you tell from the difference between height

#distributions of pines and spruces?

#The Spruce in the crown layer one appear to be more varying and

#less varying in the second crown layer.

#Pine, on the other hand, is less varying in the first crown layer but

#more varying in the second crown layer.

#-Exercise 4

#Plot the diameter, height, height of living crown (ELAVALARAJA), and

#width of the crown (LATVUSLEV) for silver birches (PUULAJI = 3)

#in the first canopy cover layer with pairs() function.

s.birchLayer1<- subset(data, data$PUULAJI==3 & data$LATVKERROS==1)

#simple pairing but I used a mmore detailed pairer after this

pairs(~LPM + PITUUS + ELAVALARAJA + LATVUSLEV, data = s.birchLayer1)

## put (absolute) correlations on the upper panels,

## with size proportional to the correlations.

panel.cor <- function(x, y, digits = 2, prefix = "", cex.cor, ...)

{

usr <- par("usr"); on.exit(par(usr))

par(usr = c(0, 1, 0, 1))

r <- abs(cor(x, y))

txt <- format(c(r, 0.123456789), digits = digits)[1]

txt <- paste0(prefix, txt)

if(missing(cex.cor)) cex.cor <- 0.8/strwidth(txt)

text(0.5, 0.5, txt, cex = cex.cor \* r)

}

pairs(~LPM+PITUUS+ELAVALARAJA+LATVUSLEV, data=s.birchLayer1,

lower.panel=panel.smooth, upper.panel=panel.cor,

pch=20, main="Birch Scatterplot Matrix")

?pairs

#Choose the two variables that seem to have the most linear relationship.

#Examine the distribution #of these two variables by drawing a QQ-plot.

#What can you tell from the plot?

#diameter(LPM) and height(PITUUS), have the most linear relationship

#with a correation of 0.79

# Plot #3: similar plot using ggplot2

?qqplot

par(mfrow=c(1,2))

qqnorm(s.birchLayer1$LPM, main = "DBH Q-Q Plot")

qqline(s.birchLayer1$LPM, col="red")

qqnorm(s.birchLayer1$PITUUS, main = "Height Q-Q Plot")

qqline(s.birchLayer1$PITUUS, col="red")

#diameter(LPM) is not normally distributed but skewed to the right

#height, appears to be normally distributed

#a<-qqplot(birchLayer1$LPM, birchLayer1$PITUUS, plot.it = TRUE,

# xlab = deparse(substitute(birchLayer1$LPM)),

# ylab = deparse(substitute(birchLayer1$PITUUS)))

?shapiro.test

shapiro.test(s.birchLayer1$LPM)

shapiro.test(s.birchLayer1$PITUUS)

?plot

Assignment 5