COMPUTERMATHEMATIK
Laborübungen zur R-Programmierung

http://www.statistik.tuwien.ac.at/lv-guide/

BLATT I — SOMMERSEMESTER 2011

1) Man erzeuge einen normalverteilte Stichprobe vom Umfang 100, $X \sim N(0,1)$. Für eine passende Klasseneinteilung (ca. 8 Klassen) soll ein Histogramm und ein Kreisdiagramm erstellt werden.

2) Man erzeuge einen exponentialverteilte Stichprobe $X \sim Ex_\lambda$ vom Umfang 150. Danach vergleiche man graphisch (in einer Zeichnung) die empirische Verteilungsfunktion aus den erzeugten Daten mit der Verteilungsfunktion der Exponentialverteilung Ex_λ. (Parameterwerte $\lambda = 3$ bzw. $\lambda = 0.2$)

3) $p_i = P[X = i]$ bezeichne die Punktwahrscheinlichkeit einer Poissonverteilung $X \sim P_5$ Erstellen Sie eine 6×8 Matrix $A = (a_{i,j})$ mit

$$a_{i,j} = \begin{cases} p_i . p_j & \text{wenn } i > j \\ p_i + p_j & \text{sonst} \end{cases}$$

Für diese Matrix sollen folgende Operationen durchgeführt werden:
a) Berechnung der Vektoren von Zeilen- bzw. Spaltensummen
b) Sortieren der Matrix nach der zweiten Spalte (aufsteigend geordnet)
c) Berechnung von $A * A^\top$

4) Für die Verteilungsfunktion

$$F(x) = \begin{cases} 0 & \text{für } x < 1 \\ \frac{x^2}{9} & \text{für } 1 \leq x \leq 3 \\ 1 & \text{für } x > 3 \end{cases}$$

sollen Zufallszahlen erzeugt werden. Man erstelle ein Skript für die verallgemeinerte Inverse $F^{-1}(.)$. Mit stetig gleichverteilten Werten U_i , $U_i \sim U_{0,1}$ sind dann

$$X_i = F^{-1}(U_i)$$

nach der Verteilungsfunktion $F(.)$ verteilt.

5) Man analysiere folgenden *R*-Code:

```
function(){ x <- round(rexp(100,0), 2)
            y <- scan("DataBsp5")
  print(A <- matrix(c(var(x),cov(x,y),var(y),cor(x,y)),ncol=2))
  qv = c(0.1,0.3,0.5,0.8,0.9,1.1)
  xs <- (x-mean(x))/sd(x) ; ys <- (y-mean(y))/sd(y)
  xq <- quantile(xs,qv, type=2) , yq <- quantile(ys,qv, type=2)
  plot(xq,yq, xlim=c(-1.5,1.5), ylim=c(-1.5,1.5) )
  xr <- rank(x) ;  yr <- rank(y)
  xrs <- (xr -mean(xr))/sd(xr) ; yrs <- (yr -mean(yr))/sd(yr)
  xrq <- quantile(xrs,qv, type=2)
```

```
yrq <- quantile(yrs,qv, type=2)
print(Ar <- matrix(c(var(xr),cov(xr,yr),var(yr),cor(xr,yr)),ncol=2))
points(xrq,yrq, pch=20,col="red") }
```

a) Man korrigiere den Code und ersetze oder entferne ungeeignete Parameter.
b) Man erkläre die Berechnungen der Prozedur.

Ergänzungsaufgaben

E - 6) Erzeugen Sie eine Stichprobe vom Umfang 200 einer zweidimensionalen normalverteilten Stochastischen Größe (X, Y). Dabei soll $X \sim N(0, 2)$, $Y \sim N(3, 5)$ gelten und die Korrelation gleich $\rho_{X,Y} = -0.3$ sein.

Hinweis: Die Aufgabe kann auf verschiedene Arten gelöst werden.
a) Man gehe von unabhängigen normalverteilten Werten $X_0 \sim N(0, 1)$, $Y_0 \sim N(0, 1)$ aus und bestimme zuerst Parameter θ_i (ohne Programm) sodaß

$$X = \theta_1 + \theta_2 X_0$$

$$Y = \theta_3 X_0 + \theta_4 Y_0 + \theta_5$$

die geforderten Mittelwerte, Varianzen und Korrelation hat . (Eigentlich ist nur ein Parameter zu berechnen, etwa muß $\theta_1 = 0$, $\theta_2 = \sqrt{2}$, $\theta_5 = 3$ sein.)
oder
b) Man verwende direkt das R-Programm für eine multivariate Normalverteilung.

E - 7) Mit den Werten aus Beispiel E - 6 soll eine Regressionsgerade

$$y_i = \alpha + \beta \, x_i$$

berechnet werden. Man schätze die Parameter α, β, das Bestimmtheitsmaß und erstelle einen Plot der Punkte und Regressionsgerade.