Universität Konstanz
Fachbereich Mathematik und Statistik
Prof. Dr. J. Beran
WS 2007/08
16. Oktober 2007

1. Computerübung zur Mathematischen Statistik

Aufgabe 1 (Vektoren, Statistikfunktionen)

(a) Erzeugen Sie einen Vektor a, der die Zahlen von 1 bis 5 enthält.

(b) Erzeugen Sie einen Vektor b, der die geraden Zahlen von 2 bis 10 enthält.

(c) Sei $c = [11, 2, 6, 3, 7]$. Berechnen Sie $a + (b * c)$ elementweise.

(d) Berchnen Sie das Skalarprodukt von a und b und das Produkt $a * b'$.

(e) Verbinden Sie a, b und c zu einem Vektor d. Wie viele Elemente hat dieser neue Vektor? Wie lauten die Elemente 7 bis 10?

(f) Bestimmen Sie arithmetisches Mittel, Maximum, Minimum und die Quantile von d. Mit welchem Befehl erhält man alles auf einmal?

Aufgabe 2 (Konstruktion von Matrizen) Benutzen Sie in dieser Aufgabe den Befehl *matrix*.

(a) Erzeugen Sie einen Vektor aus 20 Zahlen zwischen 0 und 5 und bilden Sie daraus eine Matrix $M1$ mit 4 Zeilen und 5 Spalten.

(b) Bilden Sie nun mit Zahlen der Art (3,6,9,...) eine Matrix $M2$, die dieselbe Größe wie oben hat. Ordnen sie nun aber die Zahlen zeilenweise an.

(c) Transponieren Sie nun $M1$. Führen Sie eine Matrizenmultiplikation mit $M2$ durch.

Aufgabe 3 (Verteilungen) Aus welchen Verteilungen könnten die folgenden simulierten Daten stammen. Der Befehl *hist* ist hier hilfreich. Berechnen Sie auch jeweils arithmetisches Mittel, Varianz und Schiefe.

(a) data1.txt

(b) data2.txt

(c) data3.txt

(d) data4.txt

Die Datensätze finden sie auf http://www.math.uni-konstanz.de/∼beran/MathstatistikWS07.html.

Aufgabe 4 (Gesetz der Großen Zahlen und ZGS) Sei $(X_k)_{k\in\mathbb{N}}$ eine Folge unabhängig identisch verteilter Zufallsvariablen und

$$\overline{X}_n = \frac{1}{n}\sum_{k=1}^{n} X_k \quad \text{für} \quad n \in \mathbb{N}.$$

Simulieren Sie die Zufallsvariable $\overline{X}_n$ für wachsendes n $(n = 10, 100, 1000, 10000)$, wobei die X_k

(a) auf $[0, 1]$ uniformverteilt sind,

(b) exponentialverteilt mit Parameter $\lambda = 1$ sind,

(c) Standard-Cauchy-verteilt sind.

Stellen Sie Ihre Ergebnisse grafisch dar und erläutern Sie diese. Benutzen Sie die Befehle *par(mfrow=c(4,3))* und *hist*. Veranschaulichen Sie weiterhin den Zentralen Grenzwertsatz, indem Sie in den Fällen (a) und (b) für alle n das Histogramm und die Q-Q-Plots (*qqnorm*) von

$$\sqrt{n}\frac{(\overline{X}_n - E(X_0))}{\sqrt{Var(X_0)}}$$

plotten.

Aufgabe 5 (Erwartungstreue)

(a) Simulieren Sie die Zufallsvariable $\overline{X}_n = \frac{1}{n}\sum_{k=1}^{n} X_k$ mit i.i.d. $X_k \sim \mathcal{N}(\mu, \sigma^2)$. Veranschaulichen Sie die Verteilung von $\overline{X}_n$ durch ein Histogramm.

(b) Nun sollen die simulierten Werte von $\overline{X}_n$ quadriert werden. Veranschaulichen Sie, dass $\overline{X}_n$ erwartungstreu für μ ist, $\overline{X}_n^2$ hingegen nicht für μ^2 , d.h. also $E(\overline{X}_n) = \mu$, $E(\overline{X}_n^2) \neq \mu^2$. Verwenden Sie verschiedene Werte für μ und σ^2.

(c) Veranschaulichen Sie weiterhin, dass der Schätzer

$$U_n = \frac{1}{\binom{n}{2}} \sum_{1 \leq k < l \leq n} X_k X_l$$

erwartungstreu für μ^2 ist.