Prof. Dr. Volker Schmidt
Björn Kriesche, M.Sc.

SS 2014
06. 05. 2014

Stochastik I - Übungsblatt 2

Abgabe am 13. 05. **vor** Beginn der Übung

Aufgabe 1 (2+3 Punkte)
Berechne das Stichprobenmittel und die Stichprobenvarianz der Stichprobe $x = (x_1, \ldots, x_n)$, falls

(a) $x = (7.22, 1.99, 13.02, 8.17, 10.47, 17.46, 11.78, 10.76, 12.07, 11.03)$,

(b) $x = (1, 2, \ldots, n)$.

Aufgabe 2 (4+2+3 Punkte)
Es sei $s_n^2 = s_n^2(x_1, \ldots, x_n)$ die in der Vorlesung eingeführte Stichprobenvarianz der Stichprobe $(x_1, \ldots, x_n)$. Außerdem seien $a > 0$ und $b \in \mathbb{R}$ beliebig.

(a) Zeige, dass $s_n^2(x_1 + b, \ldots, x_n + b) = s_n^2(x_1, \ldots, x_n)$ und $s_n^2(ax_1, \ldots, ax_n) = a^2 s_n^2(x_1, \ldots, x_n)$.

(b) Zeige, dass $s_n^2 = \dfrac{1}{n-1}\left(\sum_{i=1}^{n} x_i^2 - n\overline{x}_n^2\right)$.

(c) Betrachte die Funktion $f : \mathbb{R} \to [0, \infty)$ mit

$$f(y) = \sum_{i=1}^{n} (x_i - y)^2.$$

Für welchen Wert $y \in \mathbb{R}$ wird $f(y)$ minimal?

Aufgabe 3 (2+3+4 Punkte)
Sei $U_r : \Omega \to \mathbb{R}$ eine χ_r^2-verteilte Zufallsvariable und $\Phi(\cdot)$ die Verteilungsfunktion der Standardnormalverteilung.

(a) Zeige, dass $\lim_{r\to\infty} P\left(\dfrac{U_r - r}{\sqrt{2r}} \leq z\right) = \Phi(z)$.

(b) Verwende Teilaufgabe (a), um zu zeigen, dass $\lim_{r\to\infty} P\left(\sqrt{2U_r} - \sqrt{2r-1} \leq z\right) = \Phi(z)$.

(c) Sei nun $r = 25$. Berechne $P\left(U_{25} \leq 34.382\right)$

   (i) exakt (z.B. in **R** mit pchisq),

   (ii) approximativ laut der Formel in Aufgabe (a),

   (iii) approximativ laut der Formel in Aufgabe (b),

(iv) approximativ durch Simulation: Erzeuge mit **R** 1000 Realisierungen von U_{25} und schätze die gesuchte Wahrscheinlickeit als relative Häufigkeit des Ereignisses $\{U_{25} \leq 34.382\}$ (Hinweis: rchisq).

Vergleiche die exakte Wahrscheinlichkeit mit den approximativen Näherungswerten.

Aufgabe 4 (2+3 Punkte)

Bei einer konkreten Stichprobe $(x_1, \ldots, x_n)$, die eine Realisierung einer Stichprobe $(X_1, \ldots, X_n)$ von unabhängigen normalverteilten Zufallsvariablen ist, besagt die sogenannte k-Sigma-Regel, dass im Intervall $I_1 = [\overline{x}_n - s_n, \overline{x}_n + s_n]$ ca. 68%, im Intervall $I_2 = [\overline{x}_n - 2s_n, \overline{x}_n + 2s_n]$ ca. 95% und im Intervall $I_3 = [\overline{x}_n - 3s_n, \overline{x}_n + 3s_n]$ ca. 99% der Beobachtungen liegen.

(a) Schreibe in **R** zwei eigene Funktionen *mymean* und *myvar*, die jeweils das Stichprobenmittel bzw. die Stichprobenvarianz einer übergebenen Stichprobe berechnen.

(b) Simuliere 1000 Realisierungen einer $N(0, 1)$-verteilten Zufallsvariablen in **R** und schätze den Anteil der Beobachtungen in I_1, I_2 und I_3. Verwende zur Berechnung der Intervallgrenzen die in (a) definierten Funktionen. Vergleiche deine Ergebnisse mit der k-Sigma-Regel.