Fachbereich Mathematik
Prof. Dr. K. Ritter
B. Debrabant
T. Wagner
Dr. M. Doering

TECHNISCHE UNIVERSITÄT DARMSTADT

SS 2007 20./21.06.2007

Einführung in die Statistik für WInf, LaB, CE, Inf BSc etc.

5. Übung, Lösungsvorschlag

Gruppenübungen

Aufgabe G13

a) Wir erläutern Beispiel VI.31. Zunächst ist $(Z_i)_{i\in\mathbb{N}}$ mit $Z_i := (Y_i + 1)/2$ eine unabhängige, identisch verteilte Folge von $\mathbf{B}(1,p)$-verteilten Zufallsvariablen. Damit gilt für $n \in \mathbb{N}$

$$\sum_{i=1}^{n} Z_i \sim \mathbf{B}(n,p),$$

also $\sum_{i=1}^{n}(Y_i + 1)/2 = (S_n + n)/2 \sim \mathbf{B}(n,p)$. Daraus folgt für die Verteilung von S_n

$$\begin{aligned} P(\{S_n = x\}) &= P\left(\left\{\frac{S_n + n}{2} = \frac{x+n}{2}\right\}\right) \\ &= \binom{n}{(x+n)/2} p^{(x+n)/2}(1-p)^{(n-x)/2}, \end{aligned}$$

falls $x \in \{-n, -n+2, \dots, n-2, n\}$ und $P(\{S_n = x\}) = 0$ sonst (vgl. auch Satz IV.36). Ferner ergibt sich

$$\begin{aligned} \mathrm{E}(S_n) &= \mathrm{E}\left(2\cdot\frac{S_n+n}{2} - n\right) = 2np - n = n(2p-1), \\ \mathrm{Var}(S_n) &= \mathrm{Var}\left(2\cdot\frac{S_n+n}{2} - n\right) = 4\,\mathrm{Var}((S_n+n)/2) = 4np(1-p). \end{aligned}$$

b) Zunächst stellen wir fest, dass $E(S_n/n) = E(Y_1) = 2p - 1$. Gesucht ist also

$$P(\{|S_n/n - (2p-1)| < 0.1\}).$$

Beispiel VI.30 zeigt, dass

$$P(\{|S_n/n - (2p-1)| < 0.1\}) = \sum_{k\in K}\binom{n}{k}p^k(1-p)^{n-k}$$

für $K = \{k \in \{0,1,\dots,n\} : |k - np| < n/2 \cdot 0.1\}$.

p	n	K	$P(\{\lvert S_n/n-(2p-1)\rvert<0.1\})$
1/2	20	$\{10\}$	0.1762
1/2	50	$\{23,\ldots,27\}$	0.5201
1/2	100	$\{46,\ldots,54\}$	0.6318
1/4	20	$\{5\}$	0.2023
1/4	50	$\{11,\ldots,14\}$	0.4859
1/4	100	$\{21,\ldots,29\}$	0.7016

c) Schwaches Gesetz der großen Zahlen (vgl. Satz VI.32): Für jedes $\epsilon > 0$ gilt

$$\lim_{n\to\infty} P(\{|S_n/n-(2p-1)|\geq\epsilon\})=0.$$

Starkes Gesetz der großen Zahlen: Für fast alle $\omega\in\Omega$ gilt

$$\lim_{n\to\infty} S_n(\omega)/n = 2p-1.$$

d) Wir gehen analog zu Beispiel VI.50 vor. Gesucht ist eine Folge $(c_n)_{n\in\mathbb{N}}$ in $]0,\infty[$ so, dass

$$\lim_{n\to\infty} P(\{|S_n-n(2p-1)|\leq c_n\})=\alpha.$$

Für $p=3/4$ ergibt sich $\mathrm{E}(S_n)=n/2$ sowie $\mathrm{Var}(S_n)=3n/4$ (s. Teil a)). Damit erhält man die standardisierte Summenvariable

$$S_n^*=\frac{S_n-n/2}{\sqrt{3n/4}}.$$

Der Zentrale Grenzwertsatz (Satz VI.46) liefert nun:

$$\begin{aligned}\lim_{n\to\infty} P(\{|S_n-n/2|\leq c_n\}) &= \lim_{n\to\infty} P\left(\left\{|S_n^*|\leq\frac{c_n}{\sqrt{3n}/2}\right\}\right)\\ &= \Phi\left(\frac{2c_n}{\sqrt{3n}}\right)-\Phi\left(\frac{-2c_n}{\sqrt{3n}}\right)\\ &= 2\Phi\left(\frac{2c_n}{\sqrt{3n}}\right)-1.\end{aligned}$$

Gleichsetzen mit α führt zu

$$\begin{aligned} &\Phi\left(\frac{2c_n}{\sqrt{3n}}\right)=(1+\alpha)/2\\ \Leftrightarrow\quad & c_n=u_{(1+\alpha)/2}\cdot\frac{\sqrt{3n}}{2}.\end{aligned}$$

Aufgabe G14

a) Zunächst stellen wir fest, dass $\mathrm{E}^\vartheta(X_1)=\vartheta$ gilt (vgl. Beispiel VI.6). Damit können wir Satz VII.7 anwenden oder die Schätzfunktion zu Fuß auf Erwartungstreue überprüfen (s. Beweis desselben Satzes). Dabei verwenden wir die Linearität des Erwartungswertes (vgl. Bemerkung VII.1).

$$\begin{aligned}\mathrm{E}^\vartheta(g_n(X)) &= \mathrm{E}^\vartheta\left(\frac{1}{n}(X_1+\ldots+X_n)\right)=\frac{1}{n}\mathrm{E}^\vartheta(X_1+\ldots+X_n)\\ &= \frac{1}{n}(\mathrm{E}^\vartheta(X_1)+\ldots+\mathrm{E}^\vartheta(X_n))=\frac{1}{n}\Big(n\cdot\mathrm{E}^\vartheta(X_1)\Big)\\ &= \mathrm{E}^\vartheta(X_1)=\vartheta.\end{aligned}$$

b) Für die Varianz erhalten wir wegen $\mathrm{Var}^\vartheta(X_1) = \frac{(2\vartheta-0)^2}{12} = \frac{\vartheta^2}{3}$ analog

$$\begin{aligned}
\mathrm{Var}^\vartheta(g_n(X)) &= \mathrm{Var}^\vartheta\left(\frac{1}{n}(X_1 + \ldots + X_n)\right) = \left(\frac{1}{n}\right)^2 \mathrm{Var}^\vartheta(X_1 + \ldots + X_n) \\
&= \left(\frac{1}{n}\right)^2 \Big(\mathrm{Var}^\vartheta(X_1) + \ldots + \mathrm{Var}^\vartheta(X_n)\Big) = \left(\frac{1}{n}\right)^2 \cdot n\,\mathrm{Var}^\vartheta(X_1) \\
&= \frac{\mathrm{Var}^\vartheta(X_1)}{n} = \frac{\vartheta^2}{3n},
\end{aligned}$$

wobei sowohl Satz VI.22 (ii) als auch der Satz von Bienaymé benutzt wird. Aus Bemerkung VII.15 erhalten wir für den Quadratmittel-Fehler

$$R^\vartheta(g_n;\gamma) = \mathrm{Var}^\vartheta(g_n(X)) = \frac{\vartheta^2}{3n}.$$

c) Die Folge ist stark konsistent. Begründung: Satz VII.7 bzw. das starke Gesetz der großen Zahlen.

d) Wegen $\mathrm{Var}^\vartheta(X_1) = \frac{\vartheta^2}{3} = \gamma(\vartheta)$ liefert Satz VII.9 die Schätzfunktion, die das Gewünschte leistet, nämlich:

$$v_n(x) = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x}_n)^2.$$

Aufgabe G15

Die Zufallsvariable X_i beschreibe die Bedienungsdauer der i-ten Person ($i = 1, \ldots, 40$). Dann kann $(X_i)_{i\in\mathbb{N}}$ als unabhängige, identisch verteilte Folge angenommen werden. Aus $X_1 \sim \mathbf{Exp}(\lambda)$ und $\mathrm{E}(X_1) = 5/6$ folgt $\lambda = 6/5$ (s. Beispiel VI.6). Die Summe $S := X_1 + \ldots + X_{40}$ ist nach dem Zentralen Grenzwertsatz näherungsweise normalverteilt, und zwar mit den Parametern (verwende Beispiel VI.6)

$$\begin{aligned}
\mu &= \mathrm{E}(S) = 40\,\mathrm{E}(X_1) = 33\,\frac{1}{3}, \\
\sigma^2 &= \mathrm{Var}(S) = 40\,\mathrm{Var}(X_1) = 27\,\frac{7}{9}.
\end{aligned}$$

Damit folgt

$$\begin{aligned}
P(\{S \le 35\}) &\approx P\left(\left\{S^* \le \frac{35 - 33\,\frac{1}{3}}{\sqrt{27\,\frac{7}{9}}}\right\}\right) \\
&\approx \Phi(0.32) = 0.6255.
\end{aligned}$$