Prof. Dr. Klimovsky Sommersemester 2014

Stochastik

Probeklausur

Aufgabe 1. Ein neuartiges Pflanzenschutzmittel schützt 80% der Laubbäume und 60% der Nadelbäume vor Schadorganismen. In einem Wald stehen 60% Laub- und 40% Nadelbäume, die alle mit dem Pflanzenschutzmittel gespritzt wurden.

(a) Wie groß ist die Wahrscheinlichkeit, dass ein zufällig ausgewählter Baum aus dem Wald Schadorganismen aufweist?

(b) Bestimmen Sie die Wahrscheinlichkeit, dass ein Baum mit Schadorganismen ein Laubbaum ist.

Lösung. Benutze die Bayes-Formell. Sei S das Ereignis, dass ein Baum Schadorganismen aufweist und L das Ereignis, dass ein Baum ein Laubbaum ist.

(a) $\mathbb{P}(S) = 0.2 \cdot 0.6 + 0.4 \cdot 0.4 = 0.28$.

(b) $\mathbb{P}(L|S) = \frac{0.2 \cdot 0.6}{0.28} = \frac{0.12}{0.28} = \frac{3}{7}$.

Aufgabe 2. Sei X eine ZV mit $\mathrm{Var}[X] = 0$. Geben Sie die Verteilungsfunktion von X an. *Hinweis:* Benutzen Sie die Tschebyscheff-Ungleichung.

Lösung. Tschebyscheff-Ungleichung $\Rightarrow \mathbb{P}\{|X - \mathbb{E}[X]| \geq \epsilon\} = 0 \ \forall \epsilon > 0 \Rightarrow X \overset{\text{f.s.}}{\equiv} \mathbb{E}[X] =: c$.

$$\Rightarrow F_X(x) = \begin{cases} 0, & x < c \\ 1, & x \geq c. \end{cases}$$

Aufgabe 3. Sei $S_n = X_1 + X_2 + \dots X_n$ die Position des Zufallswanderers nach n Schritten, wobei die Schritte $\{X_i\}_{i=1}^n$ u.i.v. ZVen sind. Sei $\mathbb{P}\{X_1 = -1\} = p$ und $\mathbb{P}\{X_1 = 1\} = 1 - p$ für $p \in (0, 1)$.

(a) Sei $p = 1/2$. Zeigen Sie, dass die Position des Zufallswanderers $S_n \underset{n\to\infty}{\overset{\mathbb{P}}{=}} \bar{o}(n^{\frac{1}{2}+\epsilon})$ für alle $\epsilon > 0$.

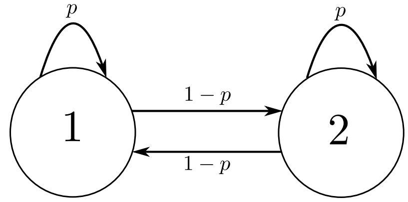
(b) Sei $p \neq 1/2$. Zeigen Sie, dass $S_n \underset{n\to\infty}{\overset{\text{f.s.}}{\sim}} (1 - 2p)n$.

Lösung.

(a) Es gilt $S_n/n^{\frac{1}{2}+\epsilon} = n^{-\epsilon} \cdot (S_n - \underbrace{\mathbb{E}[S_n]}_{=0})/\sqrt{n}$. Aus ZGS $\Rightarrow S_n/\sqrt{n}$ konvergiert in Verteilung gegen $Y \sim \mathcal{N}(0, 1/4)$. Es gilt $n^{-\epsilon} \underset{n\to\infty}{\longrightarrow} 0$. Aus den Rechenregeln für die Konvergenzarten von den ZVen folgt, dass $n^{-\epsilon} \cdot S_n/\sqrt{n}$ in Wahrscheinlichkeit gegen 0 als $n \to \infty$ konvergiert.

(b) Aus dem starken GGZ folgt, dass S_n/n gegen $\mathbb{E}[X_1] = 1 - 2p$ fast sicher als $n \to \infty$ konvergiert.

Aufgabe 4. Eine Markov-Kette ist gegeben durch den Graphen auf dem Zustandsraum $S = \{1, 2\}$ und den Übergangswahrscheinlichkeiten:



wobei $p \in (0,1)$.

(a) Geben Sie die Übergangsmatrix Q für diese Markov-Kette an.

(b) Geben Sie die stationäre Verteilung für diese Markov-Kette an.

(c) Rechnen Sie $\lim_{n\to\infty} Q^n$ aus.

Lösung.

(a) Die Übergangsmatrix ist

$$Q := \begin{pmatrix} p & 1-p \\ 1-p & p \end{pmatrix}.$$

(b) Da Q symmetrisch ist, ist die Gleichverteilung $\pi = (\frac{1}{2}, \frac{1}{2})$ stationär für Q, d.h. $\pi = \pi Q$.

(c) $\lim_{n\to\infty} Q^n$ hat die stationäre Verteilung π in jeder Zeile stehen:

$$\lim_{n\to\infty} Q^n = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}. \tag{1}$$

Aufgabe 5. Sei $\alpha > 0$ ein gegebener Parameter. Die Dichtefunktion einer ZV X ist gegeben durch $f_X(x) := cx^{-\alpha-1}1_{[1,\infty)}$, wobei $c > 0$.

(a) Geben Sie die Konstante c an.

(b) Sei $n > 0$. Geben Sie die Menge aller Parametern α an, für die $\mathbb{E}[X^n] < \infty$ gilt.

Lösung.

(a) $\int_1^{+\infty} x^{-\alpha-1} \mathrm{d}x = -\frac{x^{-\alpha}}{\alpha}\big|_{x=1}^{x=+\infty} = \frac{1}{\alpha} \Rightarrow c = \alpha$.

(b) Es gilt

$$\mathbb{E}[X^n] = \int_1^{\infty} x^n \frac{\alpha}{x^{\alpha+1}} \mathrm{d}x = \int_1^{\infty} \frac{\alpha}{x^{\alpha-n+1}} \mathrm{d}x = \begin{cases} < \infty, & n < \alpha \\ = \infty, & \text{sonst.} \end{cases} \tag{2}$$

Somit ist die Antwort $\{\alpha \in \mathbb{R} : \alpha > n\}$.

Aufgabe 6. Sei $S_n = \sum_{i=1}^n X_i^2$, wobei $\{X_i\}_{i=1}^n$ eine Folge von u.i. $\mathcal{N}(0,1)$-verteilten ZVen ist. Geben Sie die Zahlenfolgen $\{a_n\}_{n=1}^{\infty}$ and $\{b_n\}_{n=1}^{\infty}$, so dass die Folge von ZV $\{a_n(S_n - b_n)\}_{n=1}^{\infty}$ als $n \to \infty$ in Verteilung gegen $\mathcal{N}(0,1)$ konvergiert.

Lösung. Benutze den *ZGS* für $Y_i = X_i^2$. Es gilt $\mu := \mathbb{E}[Y_1]$ und $\sigma^2 := \mathrm{Var}[Y_1] = \mathbb{E}[X_1^4] - \mathbb{E}[X_1^2]^2$. Aus der Formell für die *momenterzeugende Funktion* einer Gauß'schen ZV folgt $\mathbb{E}[X_1^4] = M_{X_1}^{(4)}(0) = 3$, $\mathbb{E}[X_1^2] = 1$. Somit $\sigma^2 = 2$. Schließlich folgt aus dem ZGS, dass $(Y_1 + Y_2 + \ldots Y_n - n)/(\sqrt{2n})$ in Verteilung gegen $\mathcal{N}(0,1)$ konvergiert. Demzufolge ist $a_n = 1/(\sqrt{2n})$ und $b_n = n$.

Aufgabe 7. Sei X_1 und X_2 u.i. $\mathcal{N}(0,1)$-verteilt. Definiere $\bar{X} := (X_1 + X_2)/2$. Zeigen Sie, dass die ZVen $\bar{X}$ und $X_1 - \bar{X}$ unkorreliert sind.

Lösung. Man Zeigt "zu Fuß" (mittels der Rechenregeln), dass $\mathrm{cov}[\bar{X}, X_1 - \bar{X}] = \mathbb{E}[\bar{X}(X_1 - \bar{X})] = 0$. Alternativ kann man so verfahren: $\mathbb{E}[\bar{X}(X_1 - \bar{X})] = \mathbb{E}[\mathbb{E}[\bar{X}(X_1 - \bar{X}) \mid \bar{X}]]$. Es gilt $\mathbb{E}[\bar{X}(X_1 - \bar{X}) \mid \bar{X}] = \bar{X}(\mathbb{E}[X_1 \mid \bar{X}] - \bar{X})$. Aus der Symmetrie und Unabhängigkeit folgt $\mathbb{E}[X_1 \mid \bar{X}] = \mathbb{E}[X_2 \mid \bar{X}]$. Demzufolge $\bar{X} = \mathbb{E}[\bar{X} \mid \bar{X}] = \frac{1}{2}(\mathbb{E}[X_1 \mid \bar{X}] + \mathbb{E}[X_2 \mid \bar{X}]) = \mathbb{E}[X_1 \mid \bar{X}]$. Somit ist $\mathbb{E}[X_1 \mid \bar{X}] - \bar{X} = \bar{X} - \bar{X} = 0$.

Aufgabe 8. Ein Meinungsforschungsinstitut möchte eine Umfrage machen um den Prozentsatz aller Wähler einer bestimmten Partei mit der Genauigkeit von $\pm 2\%$ abzuschätzen. Die Irrtumswahrscheinlichkeit soll dabei maximal $\alpha = 0.05$ sein. Benutzen Sie das Binomialmodell und die Normalapproximation um die Mindestanzahl der befragten Personen festzulegen.

Die Quantile der Normalverteilung kann man der folgenden Tabelle entnehmen

p	0	0.5	0.9	0.95	0.975	0.99	0.995	1
$\Phi^{-1}(p)$	$-\infty$	0	1.282	1.645	1.960	2.326	2.576	$+\infty$

wobei $\Phi(x)$ die Verteilungsfunktion von $\mathcal{N}(0,1)$ ist.

Lösung. Sei $\{X_i\}_{i=1}^{\infty}$ u.i. Bernoulli-verteilte ZVen mit dem unbekannten Parameter p. Somit ist $\mathbb{E}[X_1] = p$. In Formeln muss das n so gewählt werden, dass

$$\mathbb{P}\{|\bar{X}_n - p| \geq 0.02\} \leq 0.05.$$

Sei $\epsilon := 0.02$. Dann soll

$$\mathbb{P}\{|\bar{X}_n - p| \geq \epsilon\} \leq \alpha.$$

gelten.

Um das n zu finden verfahren wir genau so, wie in der Vorlesungsbeispiel über eine soziologische Umfrage. Sei $Y_i := \frac{1}{n}(X_i - p)$. Sommit ist $\mathbb{E}[Y_i] = 0$. Ferner gilt

$$\begin{aligned}
\alpha &= \mathbb{P}\{|\bar{X}_n - p| \geq \epsilon\} = \mathbb{P}\{|\sum_{i=1}^{n} Y_i| \geq \epsilon\} \\
&= \mathbb{P}\{\sum_{i=1}^{n} Y_i \geq \epsilon\} + \mathbb{P}\{\sum_{i=1}^{n} Y_i \leq -\epsilon\} \\
&\approx [\text{Symmetrie der Normalverteilung}] \approx 2\mathbb{P}\{\sum_{i=1}^{n} Y_i \geq \epsilon\} \\
&= 2(1 - \mathbb{P}\{\sum_{i=1}^{n} Y_i \leq \epsilon\}).
\end{aligned} \tag{3}$$

Somit soll für das n

$$\mathbb{P}\left\{\sum_{i=1}^{n} Y_i \leq \epsilon\right\} \geq 1 - \alpha/2 = 0.975.$$

sein. Für die Normalapproximation müssen wir den Erwartungswert und die Varianz ausrechnen. Es gilt $\mathbb{E}[Y_1] = 0$, $\mathrm{Var}[Y_1] = p(1-p)/n^2$. Äquivalent soll

$$\mathbb{P}\left\{\frac{1}{\sqrt{\mathrm{Var}[Y_1] \cdot n}} \sum_{i=1}^{n} Y_i \leq \frac{\epsilon}{\sqrt{\mathrm{Var}[Y_1] \cdot n}}\right\} \geq 0.975$$

sein. Nach Normalapproximation/ZGS ist die ZV $\frac{1}{\sqrt{\mathrm{Var}[Y_1] \cdot n}} \sum_{i=1}^{n} Y_i$ approximativ $\mathcal{N}(0,1)$-verteilt. Somit ist das relevante Quantil aus der Tabelle $q := \Phi^{-1}(0.975) \approx 1.96$. Demzufolge muss das n so gewählt werden, dass

$$
\begin{aligned}
&\frac{\epsilon}{\sqrt{\mathrm{Var}[Y_1] \cdot n}} \geq 1.96 \\
\Leftrightarrow\ &\frac{\epsilon\sqrt{n}}{\sqrt{p(1-p)}} \geq 1.96 \\
\Leftrightarrow\ &n \geq \frac{(1.96)^2(1-p)p}{\epsilon^2} \\
\Leftarrow\ &\left[\max_{p\in[0,1]}(1-p)p = \frac{1}{4}\right] \\
\Leftarrow\ &n \geq \frac{(1.96)^2}{4\epsilon^2} = \frac{(1.96)^2}{4(0.02)^2} = 2401.
\end{aligned}
$$