Technische Universität München
Theoretische Informatik
Prof. J. Esparza / S. Kiefer / M. Luttenberger

Sommersemester 2009
Übungsblatt 6

LÖSUNG

Übungen zu Diskrete Wahrscheinlichkeitstheorie

Dieses Blatt ist am 15. Juni bis 10:15 abzugeben und wird am 15./16. Juni besprochen.

Beachten Sie: Alle Ergebnisse sind zu begründen!

Aufgabe 6.1 — 3P

Sei $X_1, X_2, \ldots$ eine Folge von unabhängigen (diskreten) Zufallsvariablen. Wir definieren für jedes $n \geq 1$ die Zufallsvariable $S_n = X_1 + \cdots + X_n$. In der Vorlesung wurde für den Fall, dass die X_i identisch verteilt sind, das sog. *Schwache Gesetz der Großen Zahlen* bewiesen, d.h., für alle $\delta > 0$ gilt, dass

$$\Pr\left[\left|\frac{S_n}{n} - \mathbb{E}X_1\right| \leq \delta\right] \text{ gegen 1 konvergiert für } n \to \infty.$$

Wenn die X_i verschieden verteilt sind, erfüllen die X_i das Schwache Gesetz der Großen Zahlen nicht unbedingt.

Zeigen Sie: Wenn $\Pr\left[X_i = 2^i - 1\right] = \Pr\left[X_i = -(2^i - 1)\right] = \frac{1}{2}$, dann erfüllen die X_i das Schwache Gesetz nicht und es gilt:

$$\Pr\left[\left|\frac{S_n}{n}\right| \geq 1\right] = 1 \qquad \text{für alle } n \geq 1.$$

Lösungsvorschlag: Es gilt

$$\begin{aligned}
& X_n = (X_1 + \cdots + X_n) - X_1 - X_2 - \cdots - X_{n-1} \\
\Longrightarrow\quad & |X_n| \leq |X_1 + \cdots + X_n| + |X_1| + \cdots + |X_{n-1}| \\
\Longrightarrow\quad & |X_1 + \cdots + X_n| \geq \underbrace{|X_n|}_{2^n-1} - (|X_1| + \cdots + |X_{n-1}|)
\end{aligned}$$

Mit

$$|X_1| + \cdots + |X_{n-1}| = \sum_{i=1}^{n-1} |X_i| = \sum_{i=1}^{n-1} (2^i - 1) = \underbrace{\sum_{i=1}^{n-1} 2^i}_{2^n-2} - \underbrace{\sum_{i=1}^{n-1} 1}_{n-1} = 2^n - n - 1$$

folgt

$$|X_1 + \cdots + X_n| \geq 2^n - 1 - (2^n - n - 1) = n$$

und daher

$$\left|\frac{S_n}{n}\right| \geq 1\,.$$

Aufgabe 6.2 — 1P+1P+1P

Sei X eine stetige Zufallsvariable mit Wertebereich $[-1, 1]$ und einer Dichte f_X der Form

$$f_X(x) = \begin{cases} ax + b & \text{wenn } |x| < 1 \\ 0 & \text{sonst} \end{cases}$$

(a) Zeigen Sie: $b = 1/2$.

(b) Zeigen Sie: $-1/2 \leq a \leq 1/2$.

(c) Zeigen Sie: $\mathbb{E}X = \frac{2}{3}a$ und damit auch $-1/3 \leq \mathbb{E}X \leq 1/3$.

Lösungsvorschlag:

(a)

$$1 = \int_{-\infty}^{\infty} f_X(x)\ dx = \int_{-1}^{1} (ax+b)\ dx = \left[\frac{1}{2}ax^2 + bx\right]_{-1}^{1} = \frac{1}{2}a - \frac{1}{2}a + b + b = 2b$$

Es folgt $b = \frac{1}{2}$.

(b) Aus $f_X(1) \geq 0$ folgt $a + \frac{1}{2} \geq 0$ und daher $a \geq -\frac{1}{2}$.
Aus $f_X(-1) \geq 0$ folgt $-a + \frac{1}{2} \geq 0$ und daher $a \leq \frac{1}{2}$.

(c)

$$\mathbb{E}X = \int_{-\infty}^{\infty} x f_X(x)\ dx = \int_{-1}^{1} (ax^2 + \frac{1}{2}x)\ dx = \left[\frac{1}{3}ax^3 + \frac{1}{4}x^2\right]_{1}^{1} = \frac{1}{3}a + \frac{1}{3}a + \frac{1}{4} - \frac{1}{4} = \frac{2}{3}a$$

Mit (b) folgt $\mathbb{E}X \in \left[-\frac{1}{3}, \frac{1}{3}\right]$.

Aufgabe 6.3 **3P**

Sei X eine beliebige stetige Zufallsvariable. Zeigen Sie, dass die Funktion

$$\phi(a) = \mathbb{E}[(X-a)^2]$$

ihr Minimum bei $a = \mathbb{E}X$ annimmt.
Hinweis: Leiten Sie z.B. die Funktion nach a ab.

Lösungsvorschlag: Es gilt

$$\phi(a) = \int_{\mathbb{R}} f_X(x)(x-a)^2\ dx$$

und daher

$$\begin{aligned}\phi'(a) &= -2\int_{\mathbb{R}} f_X(x)\cdot(x-a)\ dx\\ &= 2\left(\int_{\mathbb{R}} a f_X(x)\ dx - \int_{\mathbb{R}} f_X(x)x\ dx\right)\\ &= 2(a - \mathbb{E}X)\end{aligned}$$

Folglich ist $\phi(a)$ für $a \leq \mathbb{E}X$ monoton fallend und für $a \geq \mathbb{E}X$ monoton steigend.

Aufgabe 6.4 **2P+2P+2P**

Die Lebensdauer T einer Energiesparlampe hat die folgende (mit $\lambda > 0$ parametrisierte) Dichte:

$$f_T(t) = \begin{cases} \lambda^2 t e^{-\lambda t} & t \geq 0 \\ 0 & t < 0 \end{cases}$$

(a) Zeigen Sie, dass für alle $\lambda > 0$ die Funktion f_T tatsächlich eine Dichte ist, d.h., dass $\displaystyle\int_{-\infty}^{\infty} f_T(t)\ dt = 1$ gilt.

(b) Berechnen Sie $\Pr\left[T \leq \frac{4}{\lambda}\right]$.

(c) Berechnen Sie $\mathbb{E}T$ in Abhängigkeit von λ.

Lösungsvorschlag:

(a)

$$\begin{aligned}
\int_{-\infty}^{\infty} f_T(t)\ dt &= \lambda^2 \cdot \int_0^{\infty} te^{-\lambda t}\ dt \\
&= \lambda^2 \cdot \left(\left[-\frac{1}{\lambda} te^{-\lambda t} \right]_0^{\infty} - \int_0^{\infty} -\frac{1}{\lambda} e^{-\lambda t}\ dt \right) \\
&= \lambda^2 \cdot \left(0 + \left[-\frac{1}{\lambda^2} e^{-\lambda t} \right]_0^{\infty} \right) \\
&= 1
\end{aligned}$$

(b) Mit einer Rechnung ähnlich wie in (a) gilt:

$$\begin{aligned}
\Pr\left[T \leq \frac{4}{\lambda} \right] &= \int_0^{4/\lambda} f_T(t)\ dt \\
&= \lambda^2 \cdot \left(\left[-\frac{1}{\lambda} te^{-\lambda t} \right]_0^{4/\lambda} + \left[-\frac{1}{\lambda^2} e^{-\lambda t} \right]_0^{4/\lambda} \right) \\
&= \lambda^2 \cdot \left(-\frac{1}{\lambda} \cdot \frac{4}{\lambda} e^{-4} - \frac{1}{\lambda^2} e^{-4} + \frac{1}{\lambda^2} \right) \\
&= 1 - 5e^{-4} \approx 0.91
\end{aligned}$$

(c)

$$\begin{aligned}
\mathbb{E}T &= \int_0^{\infty} t f_T(t)\ dt \\
&= \int_0^{\infty} \lambda^2 t^2 e^{-\lambda t}\ dt \\
&= \underbrace{\left[-\lambda t^2 e^{-\lambda t} \right]_0^{\infty}}_{0} - \int_0^{\infty} -2\lambda te^{-\lambda t}\ dt \\
&= 2\lambda \int_0^{\infty} te^{-\lambda t}\ dt \\
&= 2\lambda \cdot \frac{1}{\lambda^2} \qquad \text{(wie in (a) gerechnet)} \\
&= \frac{2}{\lambda}
\end{aligned}$$

Aufgabe 6.5 **1P+2P+2P**

(a) Es sei Ω eine Ergebnismenge. Weiterhin sei eine Menge $\mathcal{E}$ von Ereignissen gegegeben (d.h., $\mathcal{E} \subseteq 2^{\Omega}$). Wir nehmen an, dass $\mathcal{E}$ gerade die Ereignisse enthält, an denen wir prinzipiell interessiert sind. Wir suchen daher eine σ-Algebra $\mathcal{A}$ über Ω mit $\mathcal{E} \subseteq \mathcal{A}$. Dies wirft die Frage auf, ob es eine kleinste σ-Algebra über Ω, die $\mathcal{E}$ enthält. Hierfür definiert man:

$$\sigma_\Omega(\mathcal{E}) := \bigcap \{\mathcal{A} \mid \mathcal{A} \text{ ist } \sigma\text{-Algebra über } \Omega \text{ mit } \mathcal{E} \subseteq \mathcal{A}\}.$$

Zeigen Sie, dass $\sigma_\Omega(\mathcal{E})$ eine σ-Algebra über Ω ist mit $\mathcal{E} \subseteq \sigma_\Omega(\mathcal{E})$, und für jede andere σ-Algebra über Ω mit $\mathcal{E} \subseteq \mathcal{A}$ gilt $\sigma_\Omega(\mathcal{E}) \subseteq \mathcal{A}$.

Bemerkungen:

- Man nennt $\sigma_\Omega(\mathcal{E})$ auch die kleinste von $\mathcal{E}$ erzeugte σ-Algebra über Ω.
- Wenn es auf $\sigma_\Omega(\mathcal{E})$ kein W'keitsmaß gibt, so kann es auch auf keiner anderen σ-Algebra über Ω, die $\mathcal{E}$ enthält, ein W'keitsmaß geben.
- Die Borelschen Mengen über $\mathbb{R}$ sind gerade die kleinste σ-Algebra, die von den geschlossenen Intervallen erzeugt wird.
- Die Borelschen Mengen $\mathcal{B}(\mathbb{R}^2)$ über $\mathbb{R}^2$ sind definiert durch

$$\mathcal{B}(\mathbb{R}^2) := \sigma_{\mathbb{R}^2}(\{[a,b] \times [c,d] \mid a,b,c,d \in \mathbb{R} \wedge a < b \wedge c < d\}).$$

(b) Zeigen Sie, dass die Menge $\{(x, y) \in \mathbb{R}^2 \mid x + y \leq 1\}$ in $\mathcal{B}(\mathbb{R}^2)$ enthalten ist.

(c) Zeigen Sie, dass die Menge $K_1 = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1\}$ in $\mathcal{B}(\mathbb{R}^2)$ enthalten ist.

Lösungsvorschlag:

(a) Man überprüft zunächst, dass die Axiome (E1) bis (E3) auf Folie 236 von $\sigma_\Omega(\mathcal{E})$ erfüllt werden:

(E1) gilt, da Ω in jeder der betrachteten σ-Algebren $\mathcal{A}$ liegt.

(E2) gilt, da jedes $A \in \sigma(\mathcal{E})$ in jeder der betrachteten σ-Algebren $\mathcal{A}$ liegt, womit ebenfalls $\bar{A}$ in jeder dieser σ-Algebren, und daher im Schnitt liegt.

(E3) folgt analog, da jede Menge A_i in jeder betrachteten σ-Algebra liegt, und somit auch deren Vereinigung.

Da auch $\mathcal{E}$ in jeder betrachteten σ-Algebra enthalten ist, liegt auch $\mathcal{E}$ in deren Schnitt, also $\mathcal{E} \subseteq \sigma_\Omega(\mathcal{E})$.

Sei nun $\mathcal{A}$ eine beliebige σ-Algebra über Ω mit $\mathcal{E} \subseteq \mathcal{A}$. Dann gilt nach Definition sofort $\sigma_\Omega(\mathcal{E}) \subseteq \mathcal{A}$.

(b) todo

(c) Sei $K_r = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq r\}$ und $L_r = \mathbb{R}^2 \setminus K_r = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 > r$. Wir zeigen, dass $L_1 \in \mathcal{B}(\mathbb{R}^2)$, da dann sofort $K_1 \in \mathcal{B}(\mathbb{R}^2)$ folgt.

Sei $\epsilon > 0$. Dann gilt für jeden Punkt $(x, y) \in L_{1+\epsilon}$, dass der Quader $[x \pm \epsilon/2] \times [y \pm \epsilon/2]$ noch in L_1 enthalten ist, da jeder dieser Quader einen Außenradius von $\sqrt{2}/2\epsilon (\leq \epsilon)$ hat, d.h., jeder Punkt in $[x \pm \epsilon/2] \times [y \pm \epsilon/2]$ hat noch mindestens Distanz $1 + \epsilon - \sqrt{2}/2\epsilon > 1$ vom Ursprung.

Es gilt also

$$L_1 \supseteq A_\epsilon := \bigcup_{(x,y) \in L_{1+\epsilon} \cap \mathbb{Q}^2} [x \pm \epsilon/2] \times [y \pm \epsilon/2].$$

Da $\mathbb{Q}^2 \cap L_{1+\epsilon}$ abzählbar ist, gilt $A_\epsilon \subseteq \mathcal{B}(R^2)$ und somit auch $A := \bigcup_{n \in \mathbb{N}} A_{1/n} \subseteq \mathcal{B}(\mathbb{R}^2)$.

Da $\mathbb{Q}^2$ dicht in $\mathbb{R}^2$ liegt, existiert zu jedem $(x, y) \in L_{1+1/n}$ ein $(p, q) \in \mathbb{Q}^2$ im Abstand $\leq 1/(2n)$. Insbesondere gilt dann also $(x, y) \in [p \pm 1/(2n)] \times [q \pm 1/(2n)]$, d.h., $L_{1+1/n} \subseteq A_{1/n}$ und daher auch

$$\bigcup_{n \in \mathbb{N}} L_{1+1/n} \subseteq \bigcup_{n \in \mathbb{N}_0} A_{1/n} = A.$$

Sei nun $(x, y) \in L_1$. Da L_1 offen ist, existiert ein $\epsilon > 0$, so dass $x^2 + y^2 = 1 + \epsilon$. Wähle $n \in \mathbb{N}$ so groß, dass $1/n < \epsilon$. Dann gilt $x^2 + y^2 > 1 + 1/n$, d.h., $(x, y) \in L_{1+1/n}$ und damit auch $(x, y) \in A$. Es folgt $L_1 = A$.

Aufgabe 6.6 Unkorrigierte Zusatzaufgabe 0P

Sei $X_1, X_2, \ldots$ eine Folge von unabhängigen (diskreten) Zufallsvariablen mit $\mathbb{E}X_i = 0$ für alle $i \in \{1, 2, \ldots\}$. Wir definieren für jedes $n \geq 1$ die Zufallsvariable $S_n = X_1 + \cdots + X_n$. In der Vorlesung wurde für den Fall, dass die X_i identisch verteilt sind, das sog. *Schwache Gesetz der Großen Zahlen* bewiesen, d.h., für alle $\delta > 0$ gilt, dass

$$\Pr\left[\left|\frac{S_n}{n}\right| \leq \delta\right] \text{ gegen 1 konvergiert für } n \to \infty.$$

Tatsächlich gilt für identisch verteilte X_i aber sogar das (schwerer zu beweisende) *Starke Gesetz der Großen Zahlen*, d.h.

$$\Pr\left[\frac{S_n}{n} \text{ konvergiert gegen 0 für } n \to \infty\right] = 1$$

Ziel der Aufgabe ist zu sehen, dass das nicht dasselbe ist.
Hinweis: Alle Teilaufgaben können unabhängig voneinander gelöst werden.

Wir betrachten nun (nicht identisch verteilte) X_i mit der Dichte

$$\Pr[X_i = i] = \Pr[X_i = -i] = \frac{1}{2i \log_2 i}, \quad \Pr[X_i = 0] = 1 - \frac{1}{i \log_2 i} \quad \text{für } i \geq 2$$

und $\Pr[X_1 = 0] = 1$. In den folgenden drei Teilaufgaben zeigen wir, dass diese X_i das Schwache Gesetz erfüllen.

(a) Zeigen Sie:

$$\mathrm{Var}\left[\frac{S_n}{n}\right] = \frac{1}{n^2} \cdot \sum_{i=2}^{n} \frac{i}{\log i}$$

(b) Verwenden Sie (a), um zu zeigen, dass $\mathrm{Var}\left[\frac{S_n}{n}\right]$ gegen 0 konvergiert für $n \to \infty$.

Hinweis: Betrachten Sie nur Quadratzahlen n und zerlegen Sie die Summe in $\sum_{i=2}^{\sqrt{n}-1}$ und $\sum_{i=\sqrt{n}}^{n}$.

(c) Verwenden Sie (b) und die Chebyshev-Ungleichung, um zu zeigen, dass die X_i das Schwache Gesetz erfüllen.

Im Rest der Aufgabe zeigen wir, dass diese X_i das Starke Gesetz **nicht** erfüllen.

(d) Betrachten Sie eine beliebige Folge $A_1, A_2, \ldots$ unabhängiger Ereignisse mit $\Pr[A_i] = a_i$. Geben Sie einen Ausdruck für $\Pr[$"Kein A_i mit $i \geq r$ geschieht"$]$ an und benutzen Sie anschließend die Ungleichung $1 - x \leq e^{-x}$, um zu zeigen:

$$\Pr[\text{“Kein } A_i \text{ mit } i \geq r \text{ geschieht”}] \leq e^{-\sum_{i=r}^{\infty} a_i}$$

(e) Zeigen Sie mit (d) folgende Version des Borel-Cantelli-Lemmas:
Wenn $\sum_{i=1}^{\infty} a_i$ divergiert, dann ist $\Pr[$"Nur endlich viele A_i geschehen"$] = 0$.

(f) Sei nun A_i das Ereignis "$|X_i| \geq i$". Geben Sie $\Pr[A_i] = a_i$ an und zeigen Sie mit dem Integralkriterium (siehe Wikipedia), dass $\sum_{i=1}^{\infty} a_i$ divergiert. Mit (e) folgt dann also $\Pr[$"Nur endlich viele A_i geschehen"$] = 0$.

(g) Zeigen Sie mit (f), dass die X_i das Starke Gesetz nicht erfüllen, indem Sie zeigen:

$$\Pr\left[\frac{S_n}{n} \text{ konvergiert gegen 0 für } n \to \infty\right] = 0$$

Hinweis: Angenommen, $\frac{S_n}{n} \to 0$. Dann folgt $\frac{X_n}{n} \to 0$ wegen $\frac{X_n}{n} = \frac{S_n}{n} - \frac{n-1}{n} \cdot \frac{S_{n-1}}{n-1}$. Setzen Sie von hier fort.

Lösungsvorschlag:

(a) Da die X_i unabhängig sind, gilt

$$\mathrm{Var}\left[\frac{S_n}{n}\right] = \frac{1}{n^2}\mathrm{Var}[S_n] = \frac{1}{n^2}\mathrm{Var}[X_1 + \cdots X_n] \overset{X_i \text{ unabh.}}{=} \frac{1}{n^2} \cdot \sum_{i=1}^{n} \mathrm{Var}[X_i]$$

Weiterhin gilt $\mathrm{Var}[X_1] = 0$ und für $i \geq 2$:

$$\mathrm{Var}[X_i] = \mathbb{E}[(X_i)^2] - \underbrace{(\mathbb{E}[X_i])^2}_{0} = \mathbb{E}[(X_i)^2] = i^2 \cdot \Pr[X_i = i] + (-i)^2 \cdot \Pr[X_i = -i] + 0^2 \cdot \Pr[X_i = 0] = i^2 \cdot \frac{1}{i \log i} = \frac{i}{\log i}$$

(b)

$$\begin{aligned}
\sum_{i=2}^{n} \frac{i}{\log i} &= \sum_{i=2}^{\sqrt{n}-1} \frac{i}{\log i} + \sum_{\sqrt{n}}^{n} \frac{i}{\log i} \\
&= \underbrace{\frac{1}{\log 2}}_{1} \sum_{i=2}^{\sqrt{n}-1} i \quad + \underbrace{\frac{1}{\log \sqrt{n}}}_{\frac{2}{\log n}} \cdot \sum_{i=\sqrt{n}}^{n} i \\
&\leq (\sqrt{n})^2 + \frac{2}{\log n} \cdot n^2
\end{aligned}$$

Folglich gilt

$$\frac{1}{n^2} \cdot \sum_{i=2}^{n} \frac{i}{\log i} \to 0 \quad \text{für } n \to \infty.$$

(c) Es ist zu zeigen, dass für alle $\delta > 0$

$$\Pr\left[\left|\frac{S_n}{n}\right| > \delta\right] \text{ gegen 0 konvergiert für } n \to \infty.$$

Mit der Chebyshev-Ungleichung haben wir

$$\Pr\left[\left|\frac{S_n}{n}\right| > \delta\right] \leq \Pr\left[\left|\frac{S_n}{n}\right| \geq \delta\right] \leq \frac{\mathrm{Var}\left[\frac{S_n}{n}\right]}{\delta^2},$$

was mit (b) gegen 0 geht.

(d) Da die A_i unabhängig sind, gilt

$$\Pr[\text{“Kein } A_i \text{ mit } i \geq r \text{ geschieht”}] = \prod_{i=r}^{\infty}(1 - a_i) \leq \prod_{i=r}^{\infty} e^{-a_i} = e^{-\sum_{i=r}^{\infty} a_i}$$

(e) Die Reihe $\sum_{i=1}^{\infty} a_i$ divergiere.
Dann divergiert auch für alle r die Reihe $\sum_{i=r}^{\infty} a_i$.
Mit (d) folgt für alle r:

$$\Pr[\text{“Kein } A_i \text{ mit } i \geq r \text{ geschieht”}] \leq e^{-\infty} = 0\,.$$

Mit dem

$$\begin{aligned}
\Pr[\text{“Endlich viele } A_i \text{ geschehen”}] &= \Pr[\text{“Es gibt ein } r\text{, sodass kein } A_i \text{ mit } i \geq r \text{ geschieht”}] \\
&= \Pr\left[\bigcup_{r=1}^{\infty} \text{“Kein } A_i \text{ mit } i \geq r \text{ geschieht”}\right] \\
&\leq \sum_{r=1}^{\infty} \underbrace{\Pr[\text{“Kein } A_i \text{ mit } i \geq r \text{ geschieht”}]}_{0} \\
&= 0
\end{aligned}$$

(f) Es gilt $\Pr[A_i] = \Pr[|X_i| \geq i] = \frac{1}{i \log i}$ für $i \geq 2$ und $\Pr[A_1] = 0$. Folglich ist zu zeigen, dass

$$\sum_{i=2}^{\infty} \frac{1}{i \log i}$$

divergiert. Mit dem Integralkriterium genügt es zu zeigen, dass das entsprechende Integral divergiert:

$$\begin{aligned}
\int_2^{\infty} \frac{1}{i \cdot \log i}\, di &= \ln 2 \cdot \int_2^{\infty} \frac{1}{i \cdot \ln i}\, di \\
&= \ln 2 \cdot [\ln(\ln i)]_2^{\infty} \\
&= \infty
\end{aligned}$$

(g) Aus $\frac{X_n}{n} \to 0$ folgt wegen der Definition von Konvergenz, dass $A_i =$ “$|X_i| \geq i$” nur für endlich viele i geschieht. Dieses Ereignis hat nach (f) die W'keit 0. Folglich hat auch “$\frac{S_n}{n} \to 0$” die W'keit 0.