Wahrscheinlichkeit und Statistik

Musterlösung Serie 4

1. Seien X_1, X_2 diskrete unabhängige Zufallsvariablen mit Werten $\mathbb{N}_0$.

   a) Zeigen Sie die Faltungsformel

$$\forall\, k \in \mathbb{N}_0: \qquad P[X_1 + X_2 = k] = \sum_{j=0}^{k} P[X_1 = j]P[X_2 = k - j]. \tag{1}$$

   b) Die Momenterzeugende Funktion einer diskreten Zufallsvariable X mit Werten in $\mathbb{N}_0$ ist definiert als

$$\forall\, s \in \mathbb{R}: \qquad M_X(s) = \sum_{k=0}^{\infty} e^{sk} P[X = k]. \tag{2}$$

   Zeigen sie

$$\forall\, s \in \mathbb{R}: \qquad M_{X_1+X_2}(s) = M_{X_1}(s) M_{X_2}(s). \tag{3}$$

Lösung:

a) Der Satz der totalen Wahrscheinlichkeit impliziert, dass für alle $k \in \mathbb{N}_0$ gilt

$$\begin{aligned}
P[X_1 + X_2 = k] &= \sum_{j=0}^{k} P[X_1 + X_2 = k | X_1 = j] P[X_1 = j] \\
&= \sum_{j=0}^{k} P[X_2 = k - j | X_1 = j] P[X_1 = j] \\
&= \sum_{j=0}^{k} P[X_1 = j] P[X_2 = k - j].
\end{aligned}$$

b) Mit Teilaufgabe a) kriegen wir, dass für alle $s \in \mathbb{R}$ gilt

$$\begin{aligned}
M_{X_1+X_2}(s) &= \sum_{k=0}^{\infty} e^{sk} P[X_1 + X_2 = k] \\
&= \sum_{k=0}^{\infty} e^{sk} \sum_{j=0}^{k} P[X_1 = j] P[X_2 = k - j] \\
&= \sum_{k=0}^{\infty} \sum_{j=0}^{k} e^{sj} P[X_1 = j] e^{s(k-j)} P[X_2 = k - j] \\
&= \left(\sum_{j=0}^{\infty} e^{sj} P[X_1 = j] \right) \left(\sum_{l=0}^{\infty} e^{sl} P[X_2 = l] \right) \\
&= M_{X_1}(s) M_{X_2}(s).
\end{aligned}$$

2. Wir betrachten eine mehrdimensionale symmetrische Irrfahrt. Seien $d, N \in \mathbb{N}^+$ und

$$\Omega = \big(\underbrace{\{-1,+1\} \times \ldots \times \{-1,+1\}}_{d-\text{mal}}\big)^N = \{(\omega_{i,j})_{1\le i\le d, 1\le j\le N} \,:\, \omega_{ij} = \pm 1\},$$
$$\mathcal{A} = \{A \,:\, A \subseteq \Omega\},$$
$$\mathbb{P} = \text{die Gleichverteilung auf } \Omega,$$
$$S_n = (S_n^{(1)}, \ldots, S_n^{(d)}),\ S_0 = (0,\ldots,0),$$

wobei

$$S_n^{(i)} = \sum_{j=1}^n \omega_{i,j} \qquad (1 \le i \le d).$$

Ferner sei R_n die Anzahl der Besuche in $(0,\ldots,0)$ bis zur Zeit n.

a) Interpretieren Sie $(S_0,\ldots,S_N)$ als Pfad einer Irrfahrt auf $\mathbb{Z}^d$. Zeichnen Sie insbesondere den Pfad, der zu

$$d = 2, N = 5 \quad \text{und}$$
$$\omega = ((1,1),\ (1,-1),\ (-1,-1),\ (1,1),\ (-1,1))$$

gehört.

b) Bestimmen Sie $\lim_{n\to\infty} \mathbb{P}(S_{2n} = (0,\ldots,0))$.
Hinweis: Benutzen Sie die Formel von Stirling.

(a) Benutzen Sie b) um zu entscheiden, ob

$$\lim_{n\to\infty} \mathbb{E}[R_n]$$

endlich oder unendlich ist.
Hinweis: Schreiben sie R_n als eine Summe von Indikatorfunktionen.

Lösung:

a) Die gesuchte Irrfahrt ist gegeben durch

$$S_0 = (0,0) \qquad S_1 = (1,1) \qquad S_2 = (2,0) \qquad S_3 = (1,-1) \qquad S_4 = (2,0) \qquad S_5 = (1,1).$$

b) Die Anzahl $|\{S_{2n} = (0,\ldots,0)\}|$ der Pfade der d-dimensionalen Irrfahrt welche nach $2n$ Schritten in $(0,\ldots,0)$ enden, ist gleich a^d, wobei a die Anzahl der Pfade der eindimensionalen Irrfahrt welche nach $2n$ Schritten in 0 enden, darstellt[1]. Aus Proposition 3.2 im Skript ist $a = \binom{2n}{n}$ und somit

$$\begin{aligned}\mathbb{P}\left(S_{2n} = (0,\ldots,0)\right) &= \frac{|\{S_{2n} = (0,\ldots,0)\}|}{|\Omega|} \\ &= \frac{\binom{2n}{n}^d}{2^{2nd}} = \left(\binom{2n}{n} 2^{-2n}\right)^d \\ &\sim \underbrace{(\pi n)^{\frac{-d}{2}}}_{\to 0\ (n\to\infty)},\end{aligned}$$

[1] Das wird klarer wenn man die mehrdimensionale Irrfahrt als d-faches kartesisches Produkt der eindimensionalen Irrfahrt schreibt, d.h. Ω in der Form $\underbrace{\{-1,+1\}^N \times \cdots \times \{-1,+1\}^N}_{d-\text{mal}}$ darstellt.

wobei die letzte Approximation von der Stirling Formel stammt, siehe Skript Formel über (3.6).

c) **Asymptotisches Verhalten von $\mathbb{E}[R_n]$:**
Es gilt

$$\mathbb{E}[R_n] = \mathbb{E}\left[\sum_{i=1}^{\lfloor \frac{n}{2} \rfloor} 1_{\{S_{2i}=(0,\ldots,0)\}}\right] = \sum_{i=1}^{\lfloor \frac{n}{2} \rfloor} \mathbb{E}\left[1_{\{S_{2i}=(0,\ldots,0)\}}\right]$$
$$= \sum_{i=1}^{\lfloor \frac{n}{2} \rfloor} \mathbb{P}(S_{2i} = (0,\ldots,0)) = \sum_{i=1}^{\lfloor \frac{n}{2} \rfloor} \left(\binom{2i}{i} 2^{-2i}\right)^d.$$

Betrachte nun $d = 2$, $d = 1$. Aus der Stirling Formel folgt, dass es zu jedem $\epsilon > 0$ ein $n \in \mathbb{N}$ gibt mit $\frac{\left(\binom{2i}{i}2^{-2i}\right)^d}{(\pi i)^{\frac{-d}{2}}} \geq 1 - \epsilon$ für alle $i \geq n$. Wir wählen also ein $\epsilon \in (0, 1)$ mit entsprechenden n_ϵ, sodass $\left(\binom{2i}{i}2^{-2i}\right)^d \geq (1 - \epsilon)(\pi i)^{\frac{-d}{2}}$ für alle $i \geq n_\epsilon$. Nun gilt

$$\lim_{n\to\infty} \mathbb{E}[R_n] = \sum_{i=1}^{\infty} \left(\binom{2i}{i} 2^{-2i}\right)^d \geq \sum_{i=n_\epsilon}^{\infty} \left(\binom{2i}{i} 2^{-2i}\right)^d$$
$$\geq (1-\epsilon)\frac{1}{\pi^{\frac{d}{2}}} \sum_{i=n_\epsilon}^{\infty} \frac{1}{i^{\frac{d}{2}}} \overset{d=2}{=} (1-\epsilon)\frac{1}{\pi} \sum_{i=n_\epsilon}^{\infty} \frac{1}{i} = \infty.$$

Der letzte Schritt funktioniert analog für $d = 1$.
Sei nun $d \geq 3$. Aus der Stirling Formel folgt, dass es zu jedem $\epsilon > 0$ ein $n \in \mathbb{N}$ gibt mit $\frac{\left(\binom{2i}{i}2^{-2i}\right)^d}{(\pi i)^{\frac{-d}{2}}} \leq 1 + \epsilon$ für alle $i \geq n$. Wir wählen also ein $\epsilon > 0$ mit entsprechenden n_ϵ, sodass $\left(\binom{2i}{i}2^{-2i}\right)^d \leq (1 + \epsilon)(\pi i)^{\frac{-d}{2}}$ für alle $i \geq n_\epsilon$. Nun gilt

$$\lim_{n\to\infty} \mathbb{E}[R_n] = \sum_{i=1}^{\infty} \left(\binom{2i}{i} 2^{-2i}\right)^d = \underbrace{\sum_{i=1}^{n_\epsilon - 1} \left(\binom{2i}{i} 2^{-2i}\right)^d}_{=:c<\infty} + \sum_{i=n_\epsilon}^{\infty} \left(\binom{2i}{i} 2^{-2i}\right)^d$$
$$\leq c + (1+\epsilon)\frac{1}{\pi^{\frac{d}{2}}} \sum_{i=n_\epsilon}^{\infty} \frac{1}{i^{\frac{d}{2}}} \overset{d\geq 3}{=}: C < \infty,$$

denn wie wir wissen ist $\sum_{i=1}^{\infty} \frac{1}{i^\alpha} < \infty$ für alle $\alpha > 1$.
Es ist plausibel (und kann auch bewiesen werden), dass $\mathbb{E}[R_n]$ genau dann gegen unendlich konvergiert, wenn die Wahrscheinlichkeit, nach $(0,\ldots,0)$ in maximal n Schritten zurückzukehren, gegen 1 konvergiert. Die Irrfahrt ist also rekurrent für $d = 1, 2$ und transient für $d > 2$.

3. **Das Ruinproblem.** Arno und Benno spielen folgendes Spiel: In jeder Runde wird eine faire Münze geworfen. Erscheint Kopf, so zahlt Benno einen Franken an Arno, und umgekehrt bei Zahl. Arnos bzw. Bennos Vermögen vor der ersten Runde beläuft sich auf a bzw. b Franken (a, $b \in \mathbb{N}$). Das Spiel geht zu Ende, wenn einer der beiden kein Geld mehr hat, spätestens aber nach n Runden. Es bezeichne p_n bzw. q_n die Wahrscheinlichkeit, dass Arno bzw. Benno nach diesem Spiel ruiniert ist.

a) Zeigen Sie: $\lim_{n\to\infty}(p_n + q_n) = 1$.

b) Berechnen Sie den Grenzwert $p = \lim_{n\to\infty} p_n$.
Hinweis: Betrachten Sie die Stoppzeit

$$T_n := \min\Big\{n, \min\big\{k > 0 | S_k \in \{-a, b\}\big\}\Big\}$$

und wenden Sie den Stoppsatz (Satz 3.12) an.

Lösung:

Betrachte die Irrfahrt $S_n = \sum_{k=1}^n X_k$, wobei

$$X_k := \begin{cases} +1 & \text{Arno gewinnt die } k\text{-te Runde} \\ -1 & \text{Benno gewinnt die } k\text{-te Runde} \end{cases}.$$

a) Es gilt:

$$\mathbb{P}\left(S_k \in (-a,b)\ \forall k \le n\right) \le \mathbb{P}\left(S_k \in (-\infty,b)\ \forall k \le n\right) \xrightarrow[n\to\infty]{} 0,$$

wobei die letzte Konvergenzaussage folgt aus Korollar 3.5 im Skript mit $T_b = \inf\{n > 0 \,|\, S_n = b\}$. Also ist

$$\begin{aligned} p_n + q_n &= \mathbb{P}\left(\text{Arno am Ende ruiniert}\right) + \mathbb{P}\left(\text{Benno am Ende ruiniert}\right) \\ &= \mathbb{P}\left(\text{Arno oder Benno am Ende ruiniert}\right) \\ &= 1 - \mathbb{P}\left(\text{weder Arno noch Benno am Ende ruiniert}\right) \\ &= 1 - \underbrace{\mathbb{P}\left(S_k \in (-a,b)\ \forall k \le n\right)}_{\xrightarrow[n\to\infty]{} 0} \xrightarrow[n\to\infty]{} 1. \end{aligned}$$

b) Für die Stoppzeit $T_n := n \wedge \inf\{k \ge 1 \,|\, S_k = -a \text{ oder } S_k = b\}$ beschreibt die Zufallsvariable S_{T_n} definiert durch

$$S_{T_n} : \Omega \to \mathbb{Z} \qquad\qquad \omega \mapsto S_{T_n(\omega)}(\omega)$$

Arnos Gewinn am Spielschluss. Nach Satz 3.12 im Skript gilt ja $\mathbb{E}\left[S_{T_n}\right] = 0$. Die Folge p_n konvergiert, da sie beschränkt ist und monoton wächst. Letzteres gilt, da die Anzahl der Pfade die im Ruin von Arno enden pro Schritt mindestens verdoppelt werden, während sich die Anzahl aller Pfade insgesamt genau verdoppelt. Wir setzen $p := \lim_{n\to\infty} p_n$ sowie $q := \lim_{n\to\infty} q_n = \lim_{n\to\infty}(q_n + p_n) - p_n \overset{a)}{=} 1 - p$. Wenn A bzw. B die Ereignisse bezeichnen, dass Arno bzw. Benno am Spielschluss ruiniert ist, so folgt

$$0 = \mathbb{E}(S_{T_n}) = \sum_{k=-a}^{b} k\mathbb{P}(S_{T_n} = k) = -ap_n + bq_n + \sum_{k=-a+1}^{b-1} k\mathbb{P}(S_{T_n} = k).$$

Wir beachten dass $\{S_{T_n} = k\} = \{S_n = k, T_n = n\}$ für $k \in \{-a+1, \ldots, b-1\}$, so $\mathbb{P}(S_{T_n} = k) = \mathbb{P}(S_n = k, T_n = n) \le 1 - (p_n + q_n)$ für $k \in \{-a+1, \ldots, b-1\}$. Deshalb,

$$\Big| \sum_{k=-a+1}^{b-1} k\mathbb{P}(S_{T_n} = k) \Big| \le (1 - (p_n + q_n)) \sum_{k=-a+1}^{b-1} | k |$$

Die Summe auf der rechten Seite begrenzt ist durch $f(a,b)$, und $p_n+q_n \to 1$ aus Aufgabenteil a). Deshalb $\sum_{k=-a+1}^{b-1} k\mathbb{P}(S_{T_n} = k) \to 0$ as $n \to \infty$, sodann

$$0 = \lim_{n\to\infty} -ap_n + bq_n + \sum_{k=-a+1}^{b-1} k\mathbb{P}(S_{T_n} = k) = -ap + b(1-p),$$

deshalb

$$p = \frac{b}{a+b}.$$

4. Betrachte das folgende Spielsystem V : "Start mit dem Einsatz 1 Franken; sukzessives Verdoppeln des Einsatzes bis zur ersten 1, dann Spielabbruch; jedoch spätestens Abbruch nach n Spielen."

   a) Berechne die Verteilung des Gewinns $(V \cdot S)_n$ bei diesem Spielsystem: Wie gross ist die Wahrscheinlichkeit, einen Verlust zu erleiden?

   b) Berechne den Erwartungswert und die Varianz des Gewinns.

Lösung:

a) Betrachte die Stoppzeit $T := \min\{1 \le k \le n \mid X_k = +1\}$ (beachte, dass $\min\emptyset = \infty$) und das Spielsystem

$$V_k := 2^{k-1}1_{\{T\ge k\}}.$$

Es gilt

$$\begin{aligned}
(V\cdot S)_n &= \sum_{k=1}^{n} V_k X_k = \sum_{k=1}^{n} 2^{k-1}1_{\{T\ge k\}}X_k \\
&= \left(\sum_{k=1}^{T} 2^{k-1}X_k\right)1_{\{T<\infty\}} - \left(\sum_{k=1}^{n} 2^{k-1}\right)1_{\{T=\infty\}} \\
&= \left(\sum_{k=1}^{T-1} 2^{k-1}(-1) + 2^{T-1}\right)1_{\{T<\infty\}} + (1-2^n)1_{\{T=\infty\}} \\
&= (1-2^{T-1}+2^{T-1})1_{\{T<\infty\}} + (1-2^n)1_{\{T=\infty\}} \\
&= 1_{\{T<\infty\}} + (1-2^n)1_{\{T=\infty\}}.
\end{aligned}$$

d.h. $(V\cdot S)_n$ nimmt nur die Werte 1 und $1-2^n$ an. Insbesondere gilt

$$\mathbb{P}[(V\cdot S)_n = 1-2^n] = \mathbb{P}[T=\infty] = \mathbb{P}[\bigcap_{i=1}^{n}\{X_i=-1\}] = 2^{-n},$$

$$\mathbb{P}[(V\cdot S)_n = 1] = \mathbb{P}[T<\infty] = 1-\mathbb{P}[T=\infty] = 1-2^{-n}.$$

Also ist die Wahrscheinlichkeit eines Verlustes gleich 2^{-n}.

b) Der erwartete Gewinn ist gleich

$$\mathbb{E}[(V \cdot S)_n] = (1 - 2^n)2^{-n} + 1(1 - 2^{-n}) = 0.$$

Aufgrund von Aufgabe a) hat man

$$\begin{aligned}
\mathrm{Var}[(V \cdot S)_n] &= \mathbb{E}[(V \cdot S)_n^2] - \mathbb{E}[(V \cdot S)_n]^2 = \mathbb{E}[(V \cdot S)_n^2] \\
&= (1 - 2^n)^2 2^{-n} + 1^2(1 - 2^{-n}) \\
&= 2^{-n} - 2 + 2^n + 1 - 2^{-n} = 2^n - 1.
\end{aligned}$$