8. Übungsblatt - Lösungsskizzen

Aufgabe 29 (Satz von der konvergenten Reihe, 4 = 1.5 + 1 + 1.5 Punkte).

Beweisen Sie den **Satz von der konvergenten Reihe**: Sei $(\Omega, \mathcal{A}, \mathbb{P})$ ein Wahrscheinlichkeitsraum. Sei $(X_n)_{n\in\mathbb{N}}$ eine Folge von Zufallsvariablen aus $\mathscr{L}_1(\Omega, \mathcal{A}, \mathbb{P})$ mit $\sum_{n\in\mathbb{N}} \mathbb{E}(|X_n|) < \infty$. Dann gilt:

(i) $\mathbb{P}\left(\sum_{n\in\mathbb{N}} |X_n| < \infty\right) = 1$

(ii) $\sum_{n\in\mathbb{N}} X_n \in \mathscr{L}_1$

(iii) $\mathbb{E}\left(\sum_{n\in\mathbb{N}} X_n\right) = \sum_{n\in\mathbb{N}} \mathbb{E}(X_n)$

Lösung 29.

(i) Wir betrachten die Folge von Zufallsvariablen $Y_n := |X_n|$ aus $\overline{\mathcal{A}}^+$. Für diese Folge gilt nach Lemma 20.19 (bzw. Übungsaufgabe 27(a)) $\sum_{n\in\mathbb{N}} Y_n \in \overline{\mathcal{A}}^+$ und

$$\mathbb{E}\left(\sum_{n\in\mathbb{N}} Y_n\right) = \sum_{n\in\mathbb{N}} \mathbb{E}(Y_n).$$

Es gilt also $\mathbb{E}\left(\sum_{n\in\mathbb{N}} |X_n|\right) = \mathbb{E}\left(\sum_{n\in\mathbb{N}} Y_n\right) = \sum_{n\in\mathbb{N}} \mathbb{E}(Y_n) = \sum_{n\in\mathbb{N}} \mathbb{E}(|X_n|) < \infty$. Daraus folgt mit Lemma 20.13. (ii) bereits $\mathbb{P}(\sum_{n\in\mathbb{N}} |X_n| = \infty) = 0$ bzw. $\mathbb{P}(\sum_{n\in\mathbb{N}} |X_n| < \infty) = 1$.

(ii) Wir erinnern uns an die Definition $\mathscr{L}_1(\Omega, \mathcal{A}, \mathbb{P}) := \left\{Z \in \overline{\mathcal{A}} \mid \mathbb{E}(|Z|) < \infty\right\}$. Es gilt

$$\left|\sum_{n\in\mathbb{N}} X_n\right| \leq \sum_{n\in\mathbb{N}} |X_n| \quad \overset{\mathbb{E}\text{ monoton}}{\Longrightarrow} \quad \mathbb{E}\left(\left|\sum_{n\in\mathbb{N}} X_n\right|\right) \leq \mathbb{E}\left(\sum_{n\in\mathbb{N}} |X_n|\right) \overset{(i)}{<} \infty,$$

also $\sum_{n\in\mathbb{N}} X_n \in \mathscr{L}_1(\Omega, \mathcal{A}, \mathbb{P})$.

(iii) Wir betrachten die Partialsummen $S_k := \sum_{n=1}^k X_n$.
Es gilt dann offensichtlich $\lim_{k\to\infty} S_k = \lim_{k\to\infty} \sum_{n=1}^k X_n = \sum_{n\in\mathbb{N}} X_n =: S$ punktweise (D1) Außerdem existiert $Y = \sum_{n\in\mathbb{N}} |X_n| \in \mathscr{L}_1$ mit $\sup_{k\in\mathbb{N}} |S_k| \leq \sup_{k\in\mathbb{N}} \sum_{n=1}^k |X_n| = \sum_{n=1}^\infty |X_n| = Y$ (D2). Dann folgt aus dem Satz von der dominierten Konvergenz (Satz 21.17):

$$\mathbb{E}\left(\sum_{n\in\mathbb{N}} X_n\right) = \mathbb{E}(S) \overset{\text{dom. Konv.}}{=} \lim_{k\to\infty} \mathbb{E}(S_k) = \lim_{k\to\infty} \mathbb{E}\left(\sum_{n=1}^k X_n\right) \overset{\mathbb{E}\text{ lin.}}{=} \lim_{k\to\infty} \sum_{n=1}^k \mathbb{E}(X_n) = \sum_{n\in\mathbb{N}} \mathbb{E}(X_n)$$

Aufgabe 30 (Alternative Formeln für den Erwartungswert, 4 = 2 + 1 + 1 Punkte).

(a) Sei $(\Omega, \mathcal{A}, \mathbb{P})$ ein Wahrscheinlichkeitsraum und sei $X \in \overline{\mathcal{A}}^+$ eine Zufallsvariable. Zeigen Sie, dass dann gilt:

$$\mathbb{E}(X) = \int_0^\infty \mathbb{P}(X > y) \, \mathrm{d}y$$

Hinweis: *Schreiben Sie $\mathbb{P}(X > y)$ als Integral und verwenden Sie den Satz von Fubini.*

(b) Berechnen Sie mittels der Formel aus (a) den Erwartungswert einer exponentialverteilten Zufallsvariable $X \sim \mathrm{Exp}_\lambda$ mit Parameter $\lambda > 0$.

(c) Sei $X : \Omega \longrightarrow \mathbb{N} = \{1, 2, 3, ...\}$ eine diskrete Zufallsvariable. Die Formel für den Erwartungswert aus (a) wird dann zu

$$\mathbb{E}X = \sum_{n=1}^\infty \mathbb{P}(X \geq n).$$

Berechnen Sie unter Verwendung dieser Formel den Erwartungswert im Falle, dass $X \sim \mathrm{Geo}_p$ eine geometrisch verteilte Zufallsvariable mit Parameter $p \in (0, 1)$ ist.

Lösung 30.

(a) Wir machen eine Fallunterscheidung:

- Gilt $\mathbb{P}(X = +\infty) > 0$, so sind sowohl $\mathbb{E}(X)$ als auch $\int_0^\infty \mathbb{P}(X > y) \, \mathrm{d}y$ unendlich.
- Gilt $\mathbb{P}(X = +\infty) = 0$, also $\mathbb{P}(X < \infty) = 1$, so betrachten wir $Y := X\mathbb{1}_{\{X<\infty\}} \in \mathcal{A}^+$ und es gilt $\mathbb{E}(X) = \mathbb{E}(Y)$. Dann gilt (wir schreiben $x = \int_{[0,x]} 1 \, \mathrm{d}y = \int_{\mathbb{R}^+} \mathbb{1}_{\{y<x\}} \, \mathrm{d}y$):

$$\begin{aligned}\mathbb{E}(Y) &\overset{\text{Def}}{=} \int_{\mathbb{R}^+} x\mathbb{P}^Y(\mathrm{d}x) = \int_{\mathbb{R}^+} \int_{\mathbb{R}^+} \mathbb{1}_{\{y<x\}} \, \mathrm{d}y\mathbb{P}^Y(\mathrm{d}x) \\ &\overset{\text{Fubini}}{=} \int_{\mathbb{R}^+} \int_{\mathbb{R}^+} \mathbb{1}_{\{y<x\}}\mathbb{P}^Y(\mathrm{d}x) \, \mathrm{d}y \\ &= \int_0^\infty \mathbb{P}(Y > y) \, \mathrm{d}y = \int_0^\infty \mathbb{P}(X > y) \, \mathrm{d}y\end{aligned}$$

wobei wir den Satz von Fubini anwenden können, da alle Ausdrücke im Integral ≥ 0 sind.

(b) Die Verteilungsfunktion von $X \sim \mathrm{Exp}_\lambda$ lautet

$$\mathbb{F}^X(x) = 1 - e^{-\lambda x}, \qquad x \geq 0.$$

Damit folgt

$$\mathbb{E}X = \int_0^\infty (1 - \mathbb{F}^X(y)) \, \mathrm{d}y = \int_0^\infty e^{-\lambda y} \, \mathrm{d}y = \Big[- \frac{1}{\lambda} \cdot e^{-\lambda y} \Big]_0^\infty = \frac{1}{\lambda}.$$

(c) Für eine geometrisch verteilte Zufallsvariable $X \sim \mathrm{Geo}_p$ gilt:

$$\mathbb{P}(X = k) = p \cdot (1 - p)^{k-1}, \qquad k \in \mathbb{N} = \{1, 2, 3, ...\}.$$

Es folgt

$$\mathbb{P}(X \leq n-1) = \sum_{k=1}^{n-1} \mathbb{P}(X = k) = p \cdot \sum_{k=0}^{n-2} (1-p)^k = p \cdot \frac{1-(1-p)^{n-1}}{1-(1-p)} = 1-(1-p)^{n-1},$$

d.h. $\mathbb{P}(X \geq n) = 1 - \mathbb{P}(X \leq n-1) = (1-p)^{n-1}$. Damit haben wir

$$\mathbb{E}X = \sum_{n=1}^{\infty} \mathbb{P}(X \geq n) = \sum_{n=1}^{\infty} (1-p)^{n-1} = \sum_{n=0}^{\infty} (1-p)^n = \frac{1}{1-(1-p)} = \frac{1}{p}.$$

Aufgabe 31 (Eigenschaften der (Ko-)Varianz, 4 = 1 + 2 + 1 Punkte).
Beweisen Sie Lemma 24.03 aus der Vorlesung:
Seien $X, Y, Z \in \mathscr{L}_2$ und $a, b \in \mathbb{R}$.

(a) Es gilt $\mathbb{V}\mathrm{ar}(X) = 0$ genau dann, wenn $\mathbb{P}(X = \mathbb{E}(X)) = 1$.

(b) $\mathbb{C}\mathrm{ov} : \mathscr{L}_2 \times \mathscr{L}_2 \longrightarrow \mathbb{R}^+$ ist eine positiv semi-definite, symmetrische Bilinearform, d.h.

- ▶ (**symmetrisch**) $\mathbb{C}\mathrm{ov}(X, Y) = \mathbb{C}\mathrm{ov}(Y, X)$
- ▶ (**linear**) $\mathbb{C}\mathrm{ov}(aX + bY, Z) = a\mathbb{C}\mathrm{ov}(X, Z) + b\mathbb{C}\mathrm{ov}(Y, Z)$
- ▶ (**positiv semi-definit**) $\mathbb{C}\mathrm{ov}(X, X) \geq 0$.

und es gilt $\mathbb{C}\mathrm{ov}(a, X) = 0$.

(c) $\mathbb{V}\mathrm{ar} : \mathscr{L}_2 \longrightarrow \mathbb{R}^+$ ist die von $\mathbb{C}\mathrm{ov}$ induzierte quadratische Form, sodass

- ▶ $\mathbb{V}\mathrm{ar}(aX + b) = a^2\mathbb{V}\mathrm{ar}(X)$
- ▶ $\mathbb{V}\mathrm{ar}(X + Y) = \mathbb{V}\mathrm{ar}(X) + \mathbb{V}\mathrm{ar}(Y) + 2\mathbb{C}\mathrm{ov}(X, Y)$

gelten.

Lösung 31.

(a) Wir erinnern uns an Lemma 23.05. (i) für $p = 1$: Für eine Zufallsvariable $U \in \overline{\mathcal{A}}$ gilt: $\mathbb{E}\,|U| = 0 \iff \mathbb{P}(U = 0) = 1$. Damit folgern wir:

$$\begin{aligned} & \mathbb{V}\mathrm{ar}(X) = 0 \\ \overset{\text{Def.}}{\iff} \; & \mathbb{E}\left((X - \mathbb{E}(X))^2\right) = 0 \\ \overset{23.03.(i)}{\iff} \; & \mathbb{P}\left((X - \mathbb{E}(X))^2 = 0\right) = 1 \\ \iff \; & \mathbb{P}\left(X = \mathbb{E}(X)\right) = 1. \end{aligned}$$

(b)
- ▶ $\mathbb{C}\mathrm{ov}(X, Y) \overset{\text{Def.}}{=} \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) = \mathbb{E}(YX) - \mathbb{E}(Y)\mathbb{E}(X) \overset{\text{Def.}}{=} \mathbb{C}\mathrm{ov}(Y, X)$
- ▶ $\mathbb{C}\mathrm{ov}(aX+bY, Z) \overset{\text{Def.}}{=} \mathbb{E}\left((aX + bY)Z\right) - \mathbb{E}\left((aX + bY)\right)\mathbb{E}(Z) \overset{\mathbb{E}\text{ lin.}}{=} a\mathbb{E}(XZ) + b\mathbb{E}(YZ) - a\mathbb{E}(X)\mathbb{E}(Z) - b\mathbb{E}(Y)\mathbb{E}(Z) = a\left(\mathbb{E}(XZ) - \mathbb{E}(X)\mathbb{E}(Z)\right) + b\left(\mathbb{E}(YZ) - \mathbb{E}(Y)\mathbb{E}(Z)\right) = a\mathbb{C}\mathrm{ov}(X, Z) + b\mathbb{C}\mathrm{ov}(Y, Z)$
- ▶ Es gilt $\mathbb{C}\mathrm{ov}(X, X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$. Die Funktion $\phi(x) := x^2$ ist konvex, d.h. mit der **Ungleichung von Jensen** (21.12) folgt: $\phi(\mathbb{E}(X)) = (\mathbb{E}(X))^2 \leq \mathbb{E}(X^2) = \mathbb{E}(\phi(X))$. Dann gilt auch $\mathbb{C}\mathrm{ov}(X, X) \geq 0$.

- $\mathbb{C}\text{ov}(a, X) = \mathbb{E}(aX) - \mathbb{E}(a)\mathbb{E}(X) = a\mathbb{E}(X) - a\mathbb{E}(X) = 0.$

(c)
- $\mathbb{V}\text{ar}(aX + b) \overset{\text{Def}}{=} \mathbb{E}\left((aX + b - \mathbb{E}(aX + b))^2\right) \overset{\mathbb{E}\text{ lin}}{=} \mathbb{E}\left((a(X - \mathbb{E}(X)) + b - b)^2\right) \overset{\mathbb{E}\text{ lin.}}{=} a^2\mathbb{E}\left((X - \mathbb{E}(X))^2\right) = a^2\mathbb{V}\text{ar}(X)$
- $\mathbb{V}\text{ar}(X + Y) \overset{\text{Def}}{=} \mathbb{E}\left((X + Y)^2\right) - \left(\mathbb{E}(X + Y)\right)^2 \overset{\text{ausmult.}}{=} \mathbb{E}(X^2) + 2\mathbb{E}(XY) + \mathbb{E}(Y^2) - \left(\mathbb{E}(X)\right)^2 - 2\mathbb{E}(X)\mathbb{E}(Y) - \left(\mathbb{E}Y\right)^2 = \mathbb{V}\text{ar}(X) + \mathbb{V}\text{ar}(Y) + 2\left(\mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)\right) = \mathbb{V}\text{ar}(X) + \mathbb{V}\text{ar}(Y) + 2\mathbb{C}\text{ov}(X, Y)$

Aufgabe 32 (Enten, Jäger*innen und die Tschebycheff Ungl., 4 = 1.5 + 1.5 + 1 Punkte).

Seien $m, n \in \mathbb{N}$. Eine Gruppe von n (perfekten) Jäger*innen schießt auf m Enten, wobei sich jede Jäger*in ihr Opfer zufällig und unabhängig von den anderen Jäger*innen auswählt. Insbesondere kann eine Ente also von mehreren Jäger*innen ausgewählt werden. Sei X die Anzahl der bei diesem Massaker überlebenden Enten.

(a) Berechnen Sie Erwartungswert von X.
Hinweis: *Nummerieren Sie die Enten von 1 bis m und definieren Sie das Ereignis $A_i :=$ "Die i-te Ente überlebt". Drücken Sie X durch die Zufallsvariablen $X_i := \mathbb{1}_{A_i}$ aus. Nutzen Sie dann die Linearität des Erwartungswerts und ermitteln Sie $\mathbb{E}X_i$.*

(b) Berechnen Sie die Varianz von X.
Hinweis: *Benutzen Sie die Formel $\mathbb{V}ar(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$. Nun müssen Sie sich Gedanken über den Erwartungswert $\mathbb{E}[X_iX_j]$ machen.*

(b) Sei nun $m = n = 50$. Nutzen Sie die **Tschebyscheff-Ungleichung** (24.08) und die Ergebnisse aus (a), um ein Intervall $[m_1, m_2]$, $m_1, m_2 \in \mathbb{N}$, anzugeben, in dem mit einer Wahrscheinlichkeit von 90% die Anzahl der überlebenden Enten liegt.
Anmerkung: *Wir suchen also ein 90%-Konfidenzintervall.*

Lösung 32. (a) Wir nutzen die Notationen aus dem Hinweis. Dann gilt

$$X = \sum_{i=1}^{m} X_i,$$

und die erwartete Anzahl an überlebenden Enten ist:

$$\mathbb{E}(X) = \mathbb{E}\left(\sum_{i=1}^{m} X_i\right) = \sum_{i=1}^{m} \mathbb{E}(X_i) = \sum_{i=1}^{m} \mathbb{P}(A_i).$$

Um den Erwartungswert zu berechnen, müssen wir uns also Gedanken über die Wahrscheinlichkeit von A_i machen. Es ist

$$\begin{aligned}
\mathbb{P}(A_i) &= \mathbb{P}(i\text{-te Ente überlebt}) \\
&= \mathbb{P}(\text{Kein Jäger schießt auf Ente } i) \\
&= \mathbb{P}(\text{Jäger 1 schießt nicht auf Ente } i, \text{ ..., Jäger } n \text{ schießt nicht auf Ente } i) \\
&\overset{(*)}{=} \mathbb{P}(\text{Jäger 1 schießt nicht auf Ente } i) \cdot ... \cdot \mathbb{P}(\text{Jäger } n \text{ schießt nicht auf Ente } i)
\end{aligned}$$

(Es gilt $(*)$, weil die Jäger unabhängig auf die Enten schießen)
Die Jäger wählen jede Ente mit gleicher Wahrscheinlichkeit als Ziel aus. Daher ist

$$\mathbb{P}(\text{Jäger } j \text{ schießt nicht auf Ente } i) = \frac{m-1}{m}.$$

Damit erhalten wir:

$$\mathbb{P}(A_i) = \left(\frac{m-1}{m}\right)^n,$$

und folglich für die erwartete Anzahl an überlebenden Enten:

$$\underline{\mathbb{E}(X) = \sum_{i=1}^m \mathbb{E}(X_i) = \sum_{i=1}^m \left(\frac{m-1}{m}\right)^n = m \cdot \left(\frac{m-1}{m}\right)^n.}$$

(b) Für die Varianz nutzen wir die Formel $\mathbb{V}\mathrm{ar}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$. Da $\mathbb{E}X$ bereits bekannt ist, ist nur noch $\mathbb{E}[X^2]$ zu berechnen. Es gilt wegen der Linearität des Erwartungswerts:

$$\mathbb{E}(X^2) = \mathbb{E}\left(\sum_{i,j=1}^m X_i X_j\right) = \sum_{i=1}^m \mathbb{E}(X_i^2) + \sum_{i,j=1, i\neq j}^m \mathbb{E}(X_i X_j).$$

Beachte, dass $\mathbb{E}(X_i^2) = \mathbb{E}(\mathbb{1}_{A_i}^2) = \mathbb{E}(\mathbb{1}_{A_i}) = \mathbb{P}(A_i) = \left(\frac{m-1}{m}\right)^n$ (s.o.).
Für $\mathbb{E}(X_i X_j) = \mathbb{P}(A_i \cap A_j)$ berechnen wir wie oben:

$$\begin{aligned}
& \mathbb{P}(A_i \cap A_j) \\
= \; & \mathbb{P}(i\text{-te Ente und } j\text{-te Ente überlebt}) \\
= \; & \mathbb{P}(\text{Kein Jäger schießt auf Enten } i, j) \\
= \; & \mathbb{P}(\text{Jäger 1 schießt nicht auf Enten } i, j, \text{ ..., Jäger } n \text{ schießt nicht auf Enten } i, j) \\
= \; & \mathbb{P}(\text{Jäger 1 schießt nicht auf Enten } i, j) \cdot \ldots \cdot \mathbb{P}(\text{Jäger } n \text{ schießt nicht auf Enten } i, j) \\
= \; & \left(\frac{m-2}{m}\right)^n.
\end{aligned}$$

Insgesamt erhalten wir:

$$\mathbb{E}(X^2) = m \cdot \left(\frac{m-1}{m}\right)^n + m \cdot (m-1) \cdot \left(\frac{m-2}{m}\right)^n,$$

und damit

$$\mathbb{V}\mathrm{ar}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = m \cdot \left(\frac{m-1}{m}\right)^n + m \cdot (m-1) \cdot \left(\frac{m-2}{m}\right)^n - \left(m \cdot \left(\frac{m-1}{m}\right)^n\right)^2.$$

(c) Die Tschebyscheff-Ungleichung lautet: Für beliebiges $\varepsilon > 0$ gilt:

$$\mathbb{P}(|X - \mathbb{E}(X)| \geq \varepsilon) \leq \frac{\mathbb{V}\mathrm{ar}(X)}{\varepsilon^2}.$$

Durch das Bilden des Gegenereignisses erhalten wir:

$$\mathbb{P}\Big(X \in (\mathbb{E}(X) - \varepsilon, \mathbb{E}(X) + \varepsilon)\Big) = \mathbb{P}(|X - \mathbb{E}(X)| < \varepsilon) \geq 1 - \frac{\mathbb{V}\mathrm{ar}(X)}{\varepsilon^2}.$$

Diese Ungleichung ist wie folgt zu interpretieren: Mit einer Wahrscheinlichkeit von mindestens $1 - \frac{\mathbb{V}\mathrm{ar}(X)}{\varepsilon^2}$ liegt X, die Anzahl der überlebenden Enten, im Intervall $(\mathbb{E}(X) - \varepsilon, \mathbb{E}(X) + \varepsilon)$. Wir wollen erreichen, dass

$$1 - \frac{\mathbb{V}\mathrm{ar}(X)}{\varepsilon^2} \geq 0.9 \quad \Leftrightarrow \quad 0.1 \geq \frac{\mathbb{V}\mathrm{ar}(X)}{\varepsilon^2} \quad \Leftrightarrow \quad \varepsilon \geq \sqrt{10\mathbb{V}\mathrm{ar}(X)}. \tag{1}$$

Für $n = m = 50$ gilt

$$\mathbb{E}(X) \approx 18.21, \qquad \mathbb{V}\text{ar}(X) \approx 4.88$$

Damit müssen wir gemäß (1) also $\varepsilon \geq 6.99$ wählen. Wir erhalten das Intervall

$$(\mathbb{E}(X) - \varepsilon, \mathbb{E}(X) + \varepsilon) = (11.22, 25.20).$$

Entsprechend besitzt das Intervall $[11, 26]$ die gewünschten Eigenschaften aus der Aufgabenstellung. Mit einer Wahrscheinlichkeit von mindestens 90% werden also zwischen 11 und 26 Enten überleben.
Anmerkung: *Tatsächlich ist hier sogar das Intervall* $[12, 25]$ *ausreichend, weil* $X \in (11.22, 25.20)$ *im Falle von* X *diskret bedeutet, dass* $X \in \{12, ..., 25\}$. *Würde man aber* X *mit einer stetigen Verteilung modellieren, müsste man tatsächlich* $[11, 26]$ *wählen.*

Abgabe:
In Zweiergruppen, bis spätestens Montag, den **25. Januar 2021, 09:00 Uhr**.

Homepage der Vorlesung:
`https://sip.math.uni-heidelberg.de/vl/ews-ws20/`