Klausur zur Einführung in die Wahrscheinlichkeitstheorie und Statistik

Prof. Dr. C. Löh/M. Blank 27. Juli 2012

Name: Vorname:

Matrikelnummer: ☐☐☐☐☐☐☐ Übungsleiter:

- Diese Klausur besteht aus 8 Seiten. Bitte überprüfen Sie, ob Sie alle Seiten erhalten haben.
- Bitte versehen Sie *alle* Seiten mit Ihrem Namen und Ihrer Matrikelnummer.
- Bitte schreiben Sie nicht Lösungen zu verschiedenen Aufgaben auf dasselbe Blatt.
- Sie haben zwei Stunden (= 120 Minuten) Zeit, um die Klausur zu bearbeiten; bitte legen Sie Ihren Studentenausweis und einen Lichtbildausweis zu Beginn der Klausur vor sich auf den Tisch und halten Sie die Ausweise bei der Abgabe bereit. Um Unruhe in den letzten Minuten zu vermeiden, geben Sie bitte entweder um 11:00 Uhr oder vor 10:40 Uhr ab.
- Die Klausur besteht aus 7 Aufgaben. Es können im Total 72 Punkte erreicht werden. Zum Bestehen genügen voraussichtlich 50% der Punkte.
- Es sind keinerlei Hilfsmittel wie Taschenrechner, Computer, Bücher, Vorlesungsmitschriften, Mobiltelephone etc. gestattet; Papier wird zur Verfügung gestellt. *Alle* Täuschungsversuche führen zum Ausschluss von der Klausur; die Klausur wird dann als nicht bestanden gewertet!

Viel Erfolg!

Aufgabe	1	2	3	4	5	6	7	Summe
Punkte maximal	12	12	12	10	10	10	6	72
erreichte Punkte								

Note: Unterschrift:

Aufgabe 1 (3+3+3+3 = 12 Punkte)**.** Beantworten Sie die folgenden Fragen; begründen Sie jeweils kurz Ihre Antwort (ca. ein bis drei Sätze).

1. Seien S_1 und S_2 zwei σ-Algebren auf einer Menge Ω. Ist dann auch $S_1 \cap S_2$ eine σ-Algebra auf Ω?
2. Sei (Ω, S, P) ein Wahrscheinlichkeitsraum und seien $A, B \in S$. Gilt dann bereits $P(A \cup B) \geq P(A) + P(B)$?
3. Sei $\lambda \in \mathbb{R}_{>0}$ und sei X eine $\operatorname{Exp}(\lambda)$-verteilte reellwertige Zufallsvariable auf einem Wahrscheinlichkeitsraum (Ω, S, P). Was ist $P(X \leq 1)$?
4. Ist
$$\begin{aligned} \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\longmapsto x^2 \cdot \chi_{[0,\sqrt[3]{3}]}(x) \end{aligned}$$
eine λ^1-Wahrscheinlichkeitsdichte auf $(\mathbb{R}, B(\mathbb{R}))$?

Lösung:

1. Ja, denn:
   - Die σ-Algebren S_1 und S_2 enthalten nach Definition Ω und $\emptyset$; also gilt $\Omega, \emptyset \in S_1 \cap S_2$.
   - Sei $A \in S_1 \cap S_2$. Da S_1 eine σ-Algebra ist, gilt $\Omega \setminus A \in S_1$; analog ist $\Omega \setminus A \in S_2$. Also gilt $\Omega \setminus A \in S_1 \cap S_2$.
   - Sei $(A_n)_{n\in\mathbb{N}}$ eine Folge in $S_1 \cap S_2$. Dann ist insbesondere $(A_n)_{n\in\mathbb{N}}$ eine Folge in der σ-Algebra S_1, und damit $\bigcup_{n\in\mathbb{N}} A_n \in S_1$. Analog folgt $\bigcup_{n\in\mathbb{N}} A_n \in S_2$. Also gilt $\bigcup_{n\in\mathbb{N}} A_n \in S_1 \cap S_2$.

   Daher ist $S_1 \cap S_2$ eine σ-Algebra.

   [*Häufige Fehler*: Manchmal wurde fälschlicherweise $S_1 \cap S_2$ als $\{A_1 \cap A_2 \mid A_1 \in S_1, A_2 \in S_2\}$ verstanden.]
2. Nein, denn: Für alle Wahrscheinlichkeitsräume (Ω, S, P) gilt zum Beispiel
$$P(\Omega) + P(\Omega) = 2 > 1 = P(\Omega) = P(\Omega \cup \Omega).$$

3. Wegen $P_X = \mathrm{Exp}(\lambda)$ gilt (nach dem Transformationssatz und der Definition von $\mathrm{Exp}(\lambda)$)

$$\begin{aligned}
P(\{X \le 1\}) &= \int \chi_{\{X\le 1\}}\, dP = \int \chi_{[0,1]} \circ X\, dP \\
&= \int \chi_{[0,1]}\, dP_X = \int \chi_{[0,1]}\, d(\mathrm{Exp}(\lambda)) \\
&= \int_{[0,1]} \lambda \cdot e^{-\lambda\cdot x}\, d\lambda^1(x) = \int_0^1 \lambda \cdot e^{-\lambda \cdot x}\, dx \\
&= \Big[-e^{-\lambda\cdot x}\Big]_{x=0}^{x=1} \\
&= 1 - e^{-\lambda}.
\end{aligned}$$

[Alternativ: Die ersten Gleichungen lassen sich zu $P(\{X \le 1\}) = \mathrm{Exp}((-\infty, 1])$ zusammenfassen. Außerdem kann statt dem Intervall $[0, 1]$ bzw. $(0, 1]$ auch das Intervall $(-\infty, 1]$ verwendet werden.]

[*Häufige Fehler*: Es wurden häufig elementare Rechenfehler gemacht oder die einzelnen Rechenschritte nicht gut begründet.]

4. Ja, denn: Als stückweise stetige Funktion ist $f\colon x \longmapsto x^2 \cdot \chi_{[0,\sqrt[3]{3}]}(x)$ messbar bzgl. $B(\mathbb{R})$. Da f auf $[0, \sqrt[3]{3}]$ stetig ist und außerhalb dieses Intervalls gleich 0 ist, ist f bzgl. λ^1 integrierbar. Dabei gilt

$$\begin{aligned}
\int x^2 \cdot \chi_{[0,\sqrt[3]{3}]}(x)\, d\lambda^1(x) &= \int_{[0,\sqrt[3]{3}]} x^2\, d\lambda^1(x) = \int_0^{\sqrt[3]{3}} x^2\, dx \\
&= \left[\frac{1}{3}\cdot x^3\right]_{x=0}^{x=\sqrt[3]{3}} \\
&= 1.
\end{aligned}$$

Da f keine negativen Werte annimmt, ist f also eine λ^1-Wahrscheinlichkeitsdichte.

[*Häufige Fehler*: In vielen Fällen wurde nur die Rechnung angegeben, aber es wurden die anderen nötigen Eigenschaften für Wahrscheinlichkeitsdichten nicht überprüft.]

Aufgabe 2 (3+3+3+3 = 12 Punkte)**.** Beantworten Sie die folgenden Fragen; begründen Sie jeweils kurz Ihre Antwort (ca. ein bis drei Sätze).

1. Sei (Ω, S, P) ein Wahrscheinlichkeitsraum und sei $(A_n)_{n\in\mathbb{N}}$ eine bezüglich P stochastisch unabhängige Folge in S. Ist dann auch $(A_{2\cdot n})_{n\in\mathbb{N}}$ bezüglich P stochastisch unabhängig?
2. Seien X und Y reellwertige stochastisch unabhängige Zufallsvariablen, die auf demselben Wahrscheinlichkeitsraum (Ω, S, P) definiert sind. Gilt dann bereits
$$P(X^2 + Y^2 = 0) = P(X = 0) \cdot P(Y = 0)\,?$$
3. Sei (Ω, S, P) ein Wahrscheinlichkeitsraum und sei $A \in S$. Gilt dann bereits $P(A \mid \Omega) = 1\,$?
4. Seien X und Y quadratintegrierbare reellwertige Zufallsvariablen, die auf demselben Wahrscheinlichkeitsraum definiert sind. Gilt dann bereits, dass
$$\operatorname{Cov}(X, Y) = \operatorname{Var}(X + Y)\,?$$

Lösung:

1. Ja, denn: Aus der Unabhängigkeit von $(A_n)_{n\in\mathbb{N}}$ bezüglich P folgt insbesondere für jede endliche Teilmenge $I \subset \{2 \cdot n \mid n \in \mathbb{N}\} \subset \mathbb{N}$, dass
$$P\Big(\bigcap_{i\in I} A_i\Big) = \prod_{i\in I} P(A_i).$$
Also ist $(A_{2\cdot n})_{n\in\mathbb{N}}$ eine bezüglich P stochastisch unabhängige Familie.
[*Häufige Fehler*: Oft wurde unsauber formuliert, welche endlichen Indexmengen betrachtet werden müssen.]
2. Ja, denn: Es gilt
$$P(X^2 + Y^2 = 0) = P(\{X = 0\} \cap \{Y = 0\}) = P(\{X = 0\}) \cdot P(\{Y = 0\}).$$
Dabei haben wir in der zweiten Gleichung die Unabhängigkeit von X und Y benutzt.

3. Nein, denn: Man betrachte etwa den Laplaceraum $(\{0,1\}, \mathrm{Pot}(\{0,1\}), U^d_{\{0,1\}})$ und das Ereignis $\{1\}$. Es gilt $P(\Omega) = 1 > 0$ und

$$P(A \mid \Omega) = \frac{P(A \cap \Omega)}{P(\Omega)} = \frac{P(A)}{1} = P(A) = \frac{1}{2} \neq 1.$$

[Analog für jedes andere Beispiel mit $P(A) < 1$, z.B. $A = \emptyset$.]

[*Häufige Fehler*: Manchmal wurde kein konkretes Gegenbeispiel gegeben.]

4. Nein, denn: Für $X = Y$ gilt etwa $\mathrm{Cov}(X,X) = \mathrm{Var}(X)$ und $\mathrm{Var}(X+X) = 4 \cdot \mathrm{Var}(X)$. Ist daher X beispielsweise $N(0,1)$-verteilt, so gilt

$$\mathrm{Cov}(X,X) = \mathrm{Var}(X) = 1 \neq 4 = \mathrm{Var}(X+X).$$

[Analog für jedes andere Beispiel mit $X = Y$ und $\mathrm{Var}(X) > 0$ und viele weitere Beispiele.]

[*Häufige Fehler*: Manchmal wurde kein konkretes Gegenbeispiel gegeben.]

Aufgabe 3 (3+3+3+3 = 12 Punkte)**.** Beantworten Sie die folgenden Fragen; begründen Sie jeweils kurz Ihre Antwort (ca. ein bis drei Sätze).

1. Sei (Ω, S, P) ein Wahrscheinlichkeitsraum, sei $(X_n)_{n\in\mathbb{N}}$ eine Folge reellwertiger Zufallsvariablen auf (Ω, S, P) und sei X eine reellwertige Zufallsvariable auf (Ω, S, P) mit $X_n \xrightarrow[n\to\infty]{\text{stoch}} X$. Konvergiert dann auch die Folge $((X_n - X)^2)_{n\in\mathbb{N}}$ stochastisch gegen 0?

2. Sei (Ω, S, P) ein Wahrscheinlichkeitsraum und sei $(X_n)_{n\in\mathbb{N}_{>0}}$ eine stochastisch unabhängige Folge identisch verteilter reellwertiger quadratintegrierbarer Zufallsvariablen auf (Ω, S, P) mit $\mathrm{Var}(X_1) > 0$. Gilt dann bereits
$$\frac{1}{\sqrt{n}} \cdot \sum_{k=1}^{n} \bigl(X_k - E(X_k)\bigr) \xrightarrow[n\to\infty]{\mathrm{d}} N\bigl(0, \mathrm{Var}(X_1)\bigr)\,?$$

3. Sei $(\Omega, S, (P_\vartheta)_{\vartheta\in\Theta})$ ein statistisches Modell und seien T und $\widetilde{T}$ erwartungstreue Schätzer für eine Abbildung $\tau\colon \Theta \longrightarrow \mathbb{R}$. Ist dann auch $1/3 \cdot T + 2/3 \cdot \widetilde{T}$ ein erwartungstreuer Schätzer für τ?

4. Sei $(\Omega, S, (P_\vartheta)_{\vartheta\in\Theta})$ ein statistisches Modell. Ist $\Omega \longrightarrow \mathbb{R}, \omega \longmapsto 0$ ein gleichmäßig bester erwartungstreuer Schätzer für $\Theta \longrightarrow \mathbb{R}, \vartheta \longmapsto 0$ auf diesem Modell?

Lösung:

1. Ja, denn: Sei $\varepsilon \in \mathbb{R}_{>0}$. Dann gilt wegen der Monotonie der Wurzelfunktion auf $\mathbb{R}_{>0}$ und wegen $X_n \xrightarrow[n\to\infty]{\text{stoch}} X$:
$$0 = \lim_{n\to\infty} P(|X_n - X| > \sqrt{\varepsilon}) = \lim_{n\to\infty} P((X_n - X)^2 > \varepsilon).$$
(Insbesondere existiert der letzte Grenzwert). Also konvergiert $((X_n - X)^2)_{n\in\mathbb{N}}$ stochastisch gegen 0.

   [*Häufige Fehler*: In einigen Fällen fehlte das Monotonieargument. Oft wurde aus stochastischer Konvergenz fälschlicherweise auf punktweise Konvergenz geschlossen.]

2. Ja, denn: Für alle $n \in \mathbb{N}_{>0}$ setzen wir
$$S_n := \frac{1}{\sqrt{n}} \cdot \sum_{k=1}^{n} \bigl(X_k - E(X_k)\bigr)$$

und

$$S_n^* := \frac{1}{\sqrt{\operatorname{Var}(X_1)}} \cdot S_n = \frac{1}{\sqrt{n}} \cdot \sum_{k=1}^{n} \frac{1}{\sqrt{\operatorname{Var} X_k}} \cdot \bigl(X_k - E(X_k)\bigr).$$

Die Folge $(X_n)_{n\in\mathbb{N}_{>0}}$ erfüllt die Voraussetzungen des zentralen Grenzwertsatzes; mit dem zentralen Grenzwertsatz folgt, dass

$$S_n^* \xrightarrow[n\to\infty]{\mathrm{d}} X,$$

wobei X eine $N(0,1)$-verteilte Zufallsvariable ist. Damit erhalten wir

$$S_n = \sqrt{\operatorname{Var} X_1} \cdot S_n^* \xrightarrow[n\to\infty]{\mathrm{d}} \sqrt{\operatorname{Var} X_1} \cdot X$$

[denn: ist Y eine reellwertige Zufallsvariable und ist $c \in \mathbb{R}_{>0}$, so gilt für alle $x \in \mathbb{R}$, dass $F_{c\cdot Y}(x) = F_Y(1/c \cdot x)$; alternativ kann man diese Vererbungseigenschaft von Verteilungskonvergenz auch mit Hilfe der Charakterisierung über schwache Konvergenz einsehen].

Aus dem Transformationsverhalten der Normalverteilungen folgt, dass die Zufallsvariable $\sqrt{\operatorname{Var} X_1} \cdot X$ die Verteilung $N(0, \sqrt{\operatorname{Var}(X_1)}^2 \cdot 1) = N(0, \operatorname{Var}(X_1))$ besitzt.

[*Häufige Fehler*: Viele haben das Wahrscheinlichkeitsmaß $N(0,1)$ und die zugehörige definierende λ^1-Wahrscheinlichkeitsdichte nicht sauber auseinandergehalten. Oft wurde fälschlicherweise behauptet, dass man daraus, dass der zentrale Grenzwertsatz nicht mit der betrachteten Aussage übereinstimmt, schließen kann, dass die betrachtete Aussage falsch sein muss.]

3. Ja, denn: Für alle $\vartheta \in \Theta$ gilt mit der Linearität des Erwartungswertes (insbesondere existiert der erste Erwartungswert)

$$E_{P_\vartheta}\Bigl(\frac{1}{3} \cdot T + \frac{2}{3} \cdot \widetilde{T}\Bigr) = \frac{1}{3} \cdot E_{P_\vartheta}(T) + \frac{2}{3} \cdot E_{P_\vartheta}(\widetilde{T}) = \frac{1}{3} \cdot \tau(\vartheta) + \frac{2}{3} \cdot \tau(\vartheta) = \tau(\vartheta).$$

4. Ja, denn: Als konstante Abbildung ist die Funktion $T\colon \Omega \longrightarrow \mathbb{R}$, $\omega \longmapsto 0$ messbar und bezüglich jedem Wahrscheinlichkeitsmaß quadratintegrierbar. Weiter gilt für alle $\vartheta \in \Theta$, dass

$$E_{P_\vartheta}(T) = 0;$$

d.h. der Schätzer T ist erwartungstreu bzgl. $\tau\colon \Omega \longrightarrow \mathbb{R}$, $\vartheta \longmapsto 0$. Außerdem gilt $\operatorname{Var}_{P_\vartheta}(T) = 0$ für alle $\vartheta \in \Theta$ und Varianzen sind immer nicht-negativ; also ist T ein gleichmäßig bester Schätzer für τ.

Aufgabe 4 (5 + 4 + 1 = 10 Punkte)**.**

1. Formulieren Sie das starke Gesetz der großen Zahlen!
2. Definieren Sie die im starken Gesetz der großen Zahlen auftretende Konvergenzart!
3. Auf welchem Satz beruht der Beweis (aus der Vorlesung) des starken Gesetzes der großen Zahlen?

Lösung:

1. Starkes Gesetz der großen Zahlen: Sei (Ω, S, P) ein Wahrscheinlichkeitsraum und sei $(X_n)_{n\in\mathbb{N}_{>0}}$ eine Folge quadratintegrierbarer paarweise unkorrelierter reellwertiger Zufallsvariablen auf (Ω, S, P). Die Folge $(\mathrm{Var}(X_n))_{n\in\mathbb{N}_{>0}}$ sei beschränkt. Dann gilt
$$\frac{1}{n}\cdot\sum_{k=1}^{n}(X_k - E(X_k)) \xrightarrow[n\to\infty]{P\text{-f.s.}} 0.$$
[*Häufige Fehler*: Manchmal wurde die Beschränktheitsbedingung an die Varianzen unsauber bzw. inkorrekt formuliert.]
2. Sei (Ω, S, P) ein Wahrscheinlichkeitsraum und sei $(X_n)_{n\in\mathbb{N}}$ eine Folge reellwertiger Zufallsvariablen auf (Ω, S, P). Die Folge $(X_n)_{n\in\mathbb{N}}$ *konvergiert P-fast sicher* gegen eine reellwertige Zufallsvariable X auf (Ω, S, P), wenn
$$P\big(\{\omega\in\Omega \mid \lim_{n\to\infty} X_n(\omega) = X(\omega)\}\big) = 1.$$
3. Der Beweis (aus der Vorlesung) beruht auf dem Lemma von Borel-Cantelli.

Aufgabe 5 (2 + 4 + 4 = 10 Punkte)**.** Wir betrachten eine unendliche Folge von Würfen derselben Münze; dabei sei angenommen, dass sich die Münzwürfe nicht gegenseitig beeinflussen und dass bei der Münze mit Wahrscheinlichkeit $p \in (0,1)$ „Kopf" fällt.

1. Modellieren Sie diese Situation geeignet. Erklären Sie Ihr Modell und Ihre Modellannahmen.
2. Bestimmen Sie (in Abhängigkeit von p) die Wahrscheinlichkeit dafür, dass in den ersten beiden Würfen genau einmal „Kopf" fällt.
3. Bestimmen Sie (in Abhängigkeit von p) die Wahrscheinlichkeit dafür, dass in allen Würfen „Kopf" fällt.

Lösung:

1. Wir betrachten das Produktmodell [alternativ kann man natürlich auch das Produkt über $\mathbb{N}$ betrachten]
$$(\{0,1\}, \mathrm{Pot}(\{0,1\}), B(1,p))^{\otimes \mathbb{N}_{>0}},$$
d.h. wir modellieren das Problem durch ein (unendliches) Produkt von Bernoulli-Räumen zum Paramter $p \in (0,1)$. Unsere Modellannahmen sind daher, dass wir jeden einzelnen Münzwurf durch einen Bernoulli-Raum modellieren, wobei wir 1 als „Kopf" interpretieren und annehmen, dass „Kopf" in jedem einzelnen Wurf mit (fester) Wahrscheinlichkeit p fällt.

   Weiter entspricht die Verwendung des Produktes der Annahme, dass die einzelnen Münzwürfe unabhängig sind. [Und die Indexmenge unseres Produktes entspricht der Aufzählung der einzelnen Würfe.]

   [Alternativ kann man diese Situation zum Beispiel auch durch eine stochastisch unabhängige Folge von $B(1,p)$-verteilten Zufallsvariablen modellieren.]

   [*Häufige Fehler*: Oft wurden nur endliche Produkte statt Produkte über $\mathbb{N}_{>0}$ oder $\mathbb{N}$ betrachtet. Manchmal wurde dann außerdem die Anzahl der Würfe, in denen „Kopf" fällt, modelliert, statt die Folge der Wurfergebnisse.]

2. Sei $p \in (0,1)$ und $P := B(1,p)^{\otimes \mathbb{N}_{>0}}$. Die Wahrscheinlichkeit dafür, dass in den ersten beiden Würfen genau einmal „Kopf" fällt lässt sich in unserem Modell

berechnen als (wobei π_1 bzw. π_2 die Projektionen auf die erste bzw. zweite Komponente bezeichnen)

$$\begin{aligned}
&P\big((\{\pi_1 = 1\} \cap \{\pi_2 = 0\}) \cup (\{\pi_1 = 0\} \cap \{\pi_2 = 1\})\big) \\
&= P\big(\{\pi_1 = 1\} \cap \{\pi_2 = 0\}\big) + P\big(\{\pi_1 = 0\} \cap \{\pi_2 = 1\}\big) \\
&\hspace{16em}\text{(die Ereignisse sind disjunkt)} \\
&= P(\{\pi_1 = 1\}) \cdot P(\{\pi_2 = 0\}) + P(\{\pi_1 = 0\}) \cdot P(\{\pi_2 = 1\}) \\
&\hspace{16em}(\pi_1 \text{ und } \pi_2 \text{ unabhängig}) \\
&= p \cdot (1-p) + (1-p) \cdot p \\
&= 2 \cdot p \cdot (1-p).
\end{aligned}$$

[Alternativ: Da π_1 und π_2 unabhängig und $B(1,p)$-verteilt sind, ist $\pi_1 + \pi_2$ eine $B(2,p)$ verteilte Zufallsvariable und es gilt

$$\begin{aligned}
P\big((\{\pi_1 = 1\} \cap \{\pi_2 = 0\}) \cup (\{\pi_1 = 0\} \cap \{\pi_2 = 1\})\big) &= P\big(\{\pi_1 + \pi_2 = 1\}\big) \\
&= 2 \cdot p \cdot (p-1). \text{]}
\end{aligned}$$

[*Häufige Fehler*: Einige haben für diesen Teil ein anderes Modell gewählt als im ersten Teil. Dabei wurden oft auch (ohne Begründung) nur noch die ersten beiden Würfe betrachtet statt der gesamten Folge.]

3. Sei $p \in (0,1)$ und $P := B(1,p)^{\otimes \mathbb{N}_{>0}}$. Dann lässt sich die gesuchte Wahrscheinlichkeit berechnen durch

$$\begin{aligned}
P\Big(\bigcap_{i \in \mathbb{N}} \{\pi_i = 1\}\Big) &= \lim_{n \to \infty} P\Big(\bigcap_{i=1}^{n} \{\pi_i = 1\}\Big) && (\sigma\text{-Stetigkeit von } P) \\
&= \lim_{n \to \infty} p^n && (\text{Unabhängigkeit der Folge } (\pi_i)_{i \in \mathbb{N}_{>0}}) \\
&= 0.
\end{aligned}$$

[Es gibt viele alternative Möglichkeiten dies zu zeigen.]

Aufgabe 6 ($5+2+3=10$ Punkte)**.** Sei $n \in \mathbb{N}_{>1}$. In einer Kiste befinden sich n Toaster, von denen eine gewisse Anzahl $d \in \{0,\dots,n\}$ defekt sei. Aus dieser Kiste wird zufällig (gleichverteilt) eine Teilmenge von genau zwei Toastern ausgewählt.

1. Bestimmen Sie (in Abhängigkeit von d) die Wahrscheinlichkeitsverteilung P_d der Anzahl der defekten Toaster unter den beiden zufällig ausgewählten Toastern. Erklären Sie dabei Ihr Modell und Ihre Modellannahmen.

2. Die Anzahl der defekten Toaster in dieser Kiste sei unbekannt und soll mit Hilfe einer zufällig (gleichverteilt) ausgewählten Teilmenge von genau zwei Toastern geschätzt werden. Modellieren Sie diese Situation durch ein geeignetes statistisches Modell und erklären Sie Ihr Modell.

3. Geben Sie für dieses statistische Modell einen erwartungstreuen Schätzer für die Anzahl der defekten Toaster an (und begründen Sie Ihre Antwort).

   Hinweis. Betrachten Sie ein geeignetes skalares Vielfaches der Anzahl der defekten Toaster unter den zwei ausgewählten Toastern …

Lösung:

1. Sei $\Omega := \{A \subset \{1,\dots,n\} \mid |A| = 2\}$. Wir betrachten den Laplaceraum $(\Omega, \operatorname{Pot}(\Omega), P)$ mit $P := U^d_\Omega$, sowie die Zufallsvariable

$$\begin{aligned} T_d\colon (\Omega, \operatorname{Pot}(\Omega), P) &\longrightarrow (\{0,1,2\}, \operatorname{Pot}(\{0,1,2\})) \\ \omega &\longmapsto |\omega \cap \{1,\dots,d\}|. \end{aligned}$$

   Dann induziert T_d die gesuchte Verteilung auf $(\{0,1,2\}, \operatorname{Pot}(\{0,1,2\}))$. Wir erhalten

$$P_{T_d}(\{0\}) = \frac{\left|\binom{\{d+1,\dots,n\}}{2}\right|}{\left|\binom{\{1,\dots,n\}}{2}\right|} = \frac{\binom{n-d}{2}}{\binom{n}{2}} = \frac{(n-d)\cdot(n-d-1)}{n\cdot(n-1)}$$

$$P_{T_d}(\{2\}) = \frac{\left|\binom{\{1,\dots,d\}}{2}\right|}{\left|\binom{\{1,\dots,n\}}{2}\right|} = \frac{\binom{d}{2}}{\binom{n}{2}} = \frac{d\cdot(d-1)}{n\cdot(n-1)}$$

$$\begin{aligned} P_{T_d}(\{1\}) &= 1 - P_{T_d}(\{0\}) - P_{T_d}(\{2\}) \\ &= \frac{n\cdot(n-1) - (n-d)\cdot(n-d-1) - d\cdot(d-1)}{n\cdot(n-1)} \\ &= 2\cdot\frac{n\cdot d - d^2}{n\cdot(n-1)} \\ \Big[&= \frac{\binom{d}{1}\cdot\binom{n-d}{1}}{\binom{n}{2}}. \Big] \end{aligned}$$

In unserem Modell entspricht jede Zahl in $\{1,\ldots,n\}$ einem der n Toaster, wobei die Elemente der Teilmenge $\{1,\ldots,d\}$ den defekten Toastern entsprechen. Der Modellraum beschreibt das Auswählen zweier Toaster über zweielementige Teilmengen von $\{1,\ldots,n\}$. Die Annahme, dass jedes der Paare mit gleicher Wahrscheinlichkeit gezogen werden soll, entspricht im Modell der Verwendung der Gleichverteilung. Die Zufallsvariable T_d gibt schließlich die Zahl der defekten Toaster in der Stichprobe an.

[Alternativ kann man zum Beispiel als Modell direkt einen geeigneten Laplaceraum verwenden und die Modellierung ohne Zufallsvariable formulieren.]

[*Häufige Fehler*: Oft wurde fälschlicherweise angenommen, dass man die beiden Toaster ohne weiteres durch unabhängiges einzelnes Ziehen von Toastern modellieren kann; man beachte jedoch, dass der erste und der zweite Toaster nicht übereinstimmen dürfen und daher eine Unabhängigkeitsannahme ohne weiteres nicht gerechtfertigt ist. Außerdem wurde in einigen Fällen implizit angenommen, dass es einen „ersten“ und einen “zweiten“ Toaster gibt (auch das passt nicht ohne weiteres zur beschriebenen Situation). Viele haben versucht, „elementar“ – ohne exaktes Modell – zu argumentieren.]

2. Wir modellieren das Problem (unter Verwendung der Notation aus Aufgabenteil 1) durch das statistische Modell

$$(\{0,1,2\}, \mathrm{Pot}(\{0,1,2\}), (P_{T_d})_{d\in\{0,\ldots,n\}}).$$

Als Werte für die Anzahl d der defekten Toaster werden nun alle Elemente von $\{0,\ldots,n\}$ betrachtet, die Modellierung der Wahrscheinlichkeitsräume für ein bestimmtes d erfolgt wie in Aufgabenteil 1.

[*Häufige Fehler*: Viele haben für diesen Teil ein Modell gewählt, das nichts mit dem gewählten Modell aus dem ersten Teil zu tun hat(!).]

3. Wir definieren einen Schätzer durch die messbare (!) Abbildung

$$T\colon (\{0,1,2\}, \mathrm{Pot}(\{0,1,2\})) \longrightarrow (\mathbb{R}, B(\mathbb{R}))$$
$$\omega \longmapsto \omega \cdot \frac{n}{2}.$$

Für jedes $d \in \{0,\dots,n\}$ ist T bezüglich P_{T_d} integrierbar (es ist $\{0,1,2\}$ eine endliche Menge) und es gilt

$$E_{P_{T_d}}(T) = \frac{n}{2} \cdot \left(2 \cdot \frac{d\cdot(d-1)}{n\cdot(n-1)} + 1 \cdot \frac{2\cdot d\cdot(n-d)}{n\cdot(n-1)}\right) = d.$$

Also ist T ein erwartungstreuer Schätzer für die Anzahl d der defekten Toaster.

[*Häufige Fehler*: Auch hier wurde oft ein weiteres Modell eingeführt – statt das aus dem ersten/zweiten Teil weiterzuverwenden. Viele der betrachteten Funktionen waren nicht wohldefiniert (da in der Definition der unbekannte Parameter verwendet wurde).]

Aufgabe 7 (6 Punkte)**.** Sei (Ω, S, P) ein Wahrscheinlichkeitsraum und sei $X\colon \Omega \longrightarrow \mathbb{N}$ eine $(\mathbb{N}, \mathrm{Pot}(\mathbb{N}))$-wertige Zufallsvariable auf (Ω, S, P), für die $P(X = n) > 0$ für alle $n \in \mathbb{N}$ gilt. Außerdem gelte für alle $n, k \in \mathbb{N}$, dass

$$P(X \geq k + n \mid X \geq n) = P(X \geq k).$$

Zeigen Sie, dass P_X eine geometrische Verteilung ist.

Lösung: Sei $q := P(X \geq 1)$. Nach Voraussetzung ist $q = P(X \geq 1) \geq P(X = 1) > 0$ und $q = 1 - P(X = 0) < 1$, und es gilt für jedes $n \in \mathbb{N}$, dass

$$q = P(X \geq n + 1 \mid X \geq n).$$

Damit folgt (da $P(X \geq n) > 0$)

$$\begin{aligned} P(X \geq n+1) &= \frac{P(X \geq n+1)}{P(X \geq n)} \cdot P(X \geq n) \\ &= P(X \geq n+1 \mid X \geq n) \cdot P(X \geq n) \\ &= q \cdot P(X \geq n) \end{aligned}$$

für alle $n \in \mathbb{N}$ und induktiv erhalten wir

$$P(X \geq n) = q^n$$

für alle $n \in \mathbb{N}$. Also gilt für alle $n \in \mathbb{N}$, dass

$$P(X = n) = P(X \geq n) - P(X \geq n+1) = q^n - q^{n+1} = q^n \cdot (1 - q).$$

Somit ist X geometrisch verteilt zum Paramter $1-q$ [d.h. zum Parameter $P(X = 0)$].

[*Häufige Fehler*: Die Exponentialverteilung wird durch Gedächtnislosigkeit auf $(\mathbb{R}, B(\mathbb{R}))$ charakterisiert, nicht auf $(\mathbb{N}, \mathrm{Pot}(\mathbb{N}))$(!). In einigen Fällen wurde die zu zeigende Implikation mit ihrer Umkehrung verwechselt. Fast alle, die im Prinzip einen korrekten Beweis gegeben haben, haben nicht argumentiert, warum der angegebene Parameter tatsächlich in $(0, 1)$ liegt.]