Prof. Dr. Jan Beyersmann
Matthias Weber

SS 2021
Übung: 28. Mai 2021, Abgabe: 27. Mai

Angewandte Stochastik 1 – Übungsblatt 3

Dieses Übungsblatt wird in der Übung am 28. Mai um 10:15 Uhr diskutiert. Bitte gebt eure Lösungen bis zum Donnerstag, 27. Mai um 12:00 Uhr über Moodle ab. Gebt eure Lösungen als exakt eine PDF-Datei ab, andere Dateiformate oder mehrere Dateien sind nicht erlaubt. Gebt die Lösungen nach Möglichkeit bitte in Gruppen von zwei oder drei Personen ab. Bitte beachtet auch die Hinweise zur Übung in Moodle.

Musterlösungen

Aufgabe 1 (2 + 2 = 4 Punkte) 2
Aufgabe 2 (1 + 2 + 2 = 5 Punkte) 3
Aufgabe 3 (2 + 2 = 4 Punkte) 5
Aufgabe 4 (2 Punkte) 6
Aufgabe 5 (2 Punkte) 7
Aufgabe 6 (3 Punkte) 8

Multiple Choice Aufgaben **9**

Aufgabe 7 (3 Bonuspunkte) 10
Aufgabe 8 (4 Bonuspunkte) 11

Aufgabe 1 (2 + 2 = 4 Punkte)

Eine Fußballmannschaft hat eine Siegchance von 75% je Spiel, falls ihr Kapitän in guter Form ist. Falls ihr Kapitän nicht gut in Form ist, dann betrage ihre Siegchance nur 40%. Bei 70% aller Spiele seiner Mannschaft sei der Kapitän in guter Form. Berechne die Wahrscheinlichkeit, dass

(a) die Mannschaft ein Spiel gewinnt,

Lösung:
Definiere die Ereignisse S = "Die Mannschaft gewinnt", G = "Der Kapitän ist gut in Form".

$$P(S) = P(S|G)P(G) + P(S|\bar{G})P(\bar{G}) = 0.75 \cdot 0.7 + 0.4 \cdot 0.3 = 0.645$$

(b) der Kapitän bei einem Spiel gut in Form ist, wenn die Mannschaft nicht gewinnt.

Lösung:

$$P(G|\bar{S}) = \frac{P(G \cap \bar{S})}{P(\bar{S})} = \frac{P(G \cap \bar{S})}{P(G)} \cdot \frac{P(G)}{P(\bar{S})} = P(\bar{S}|G) \cdot \frac{P(G)}{P(\bar{S})} = 0.25 \cdot \frac{0.7}{0.355} = 0.493$$

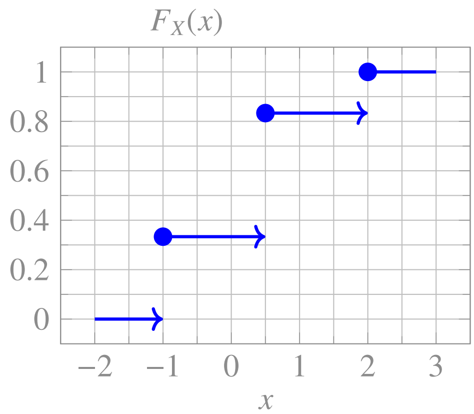
Aufgabe 2 (1 + 2 + 2 = 5 Punkte)

Gib die Verteilungsfunktionen folgender Zufallsvariablen an und skizziere sie:

(a) X, wobei $P(X = -1) = \frac{1}{3}$, $P(X = \frac{1}{2}) = \frac{1}{2}$ und $P(X = 2) = \frac{1}{6}$

Lösung:

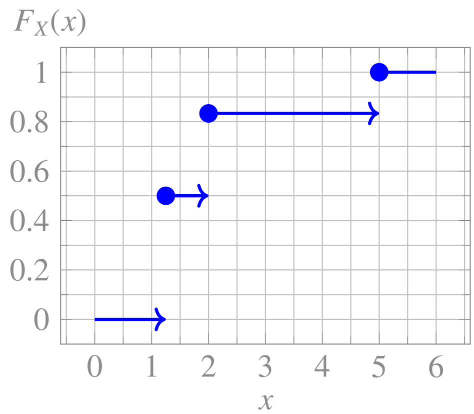
$$F_X(x) = \begin{cases} 0, & x < -1 \\ 1/3, & -1 \leq x < 1/2 \\ 5/6, & 1/2 \leq x < 2 \\ 1, & 2 \leq x \end{cases}$$



(b) $X^2 + 1$

Lösung:

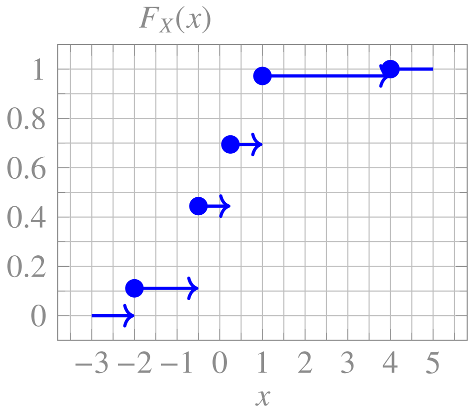
$$F_{X^2+1}(x) = \begin{cases} 0, & x < 1.25 \\ 1/2, & 1.25 \leq x < 2 \\ 5/6, & 2 \leq x < 5 \\ 1, & 5 \leq x \end{cases}$$



(c) $X \cdot X'$, wobei X und X' unabhängige Zufallsvariablen mit der Verteilungsfunktion aus Teil (a) sind.

Lösung:

$$F_{X\ cdotX'}(x) = \begin{cases} 0, & x < -2 \\ 4/36, & -2 \le x < -1/2 \\ 16/36, & -1/2 \le x < 1/4 \\ 25/36, & 1/4 \le x < 1 \\ 35/36, & 1 \le x < 4 \\ 1, & 4 \le x \end{cases}$$



Aufgabe 3 (2 + 2 = 4 Punkte)

Gegeben sein eine Zufallsvariable $X \sim Geo(p)$ für einen Parameter $p \in (0, 1)$ die der geometrischen Verteilung folgt.

(a) Berechne die Verteilungsfunktion der Zufallsvariable X.

Lösung:
Wir verwenden die Verteilung $P(X = k) = (1 - p)p^{k-1}$. Dann ist für $x \geq 1$

$$F_X(x) = F_X(\lfloor x \rfloor) = P(X \leq x) = \sum_{k=1}^{\lfloor x \rfloor} P(X = k) = (1-p) \sum_{k=0}^{\lfloor x \rfloor - 1} p^k = (1-p)\frac{1 - p^{\lfloor x \rfloor}}{1 - p} = 1-p^{\lfloor x \rfloor}$$

(b) Zeige, dass die geometrische Verteilung gedächtnislos ist, d.h., es gilt für $n, k = 1, 2, \ldots$:

$$P(X = n + k | X > k) = P(X = n)\,.$$

Lösung:

$$P(X = n + k | X > k) = \frac{P(X = n + k)}{1 - P(X \leq k)} = \frac{(1 - p)p^{n+k-1}}{p^k} = (1 - p)p^{n-1} = P(X = n)$$

Aufgabe 4 (2 Punkte)

Betrachte Aufgabe 4 von Übungsblatt 2. Berechne für die dort gegebenen Wahrscheinlichkeiten den so genannten *negativen Vorhersagewert*

$$P(\text{“eine Person ist nicht erkrankt”} \mid \text{“die Person wird negativ getestet”})\,.$$

Lösung:

Wir verwenden die Elementarereignisse $\omega_1 = (gesund, negativ)$, $\omega_2 = (gesund, positiv)$, $\omega_3 = (krank, negativ)$, $\omega_4 = (krank, positiv)$. *Dann suchen wir*

$$P(\{\omega_1, \omega_2\}|\{\omega_1, \omega_3\}) = \frac{P(\{\omega_1\})}{P(\{\omega_1, \omega_3\})} = 0.969/(0.969 + 0.001) \approx 0.9990$$

Aufgabe 5 (2 Punkte)

Gegeben sei eine Zufallsvariable X mit Verteilungsfunktion F_X und zwei Werte $a, b \in \mathbb{R}, a < b$. Drücke $P(a < X < b)$ und $P(a <= X < b)$ mit Hilfe von F_X aus.

Lösung:

$$P(a < X < b) = \lim_{x \to b} F_X(x) - F_X(a)$$

und

$$P(a \leq X < b) = \lim_{x \to b-} F_X(x) - \lim_{x \to a-} F_X(x)$$

Aufgabe 6 (3 Punkte)

Welche der Plots in Abbildung 1 zeigen Verteilungsfunktionen? Begründe deine Antwort, falls ein Plot keine Verteilungsfunktion zeigt. Gib für die skizzierten diskreten Verteilungen die Zähldichten $\{p_k\}$ an.

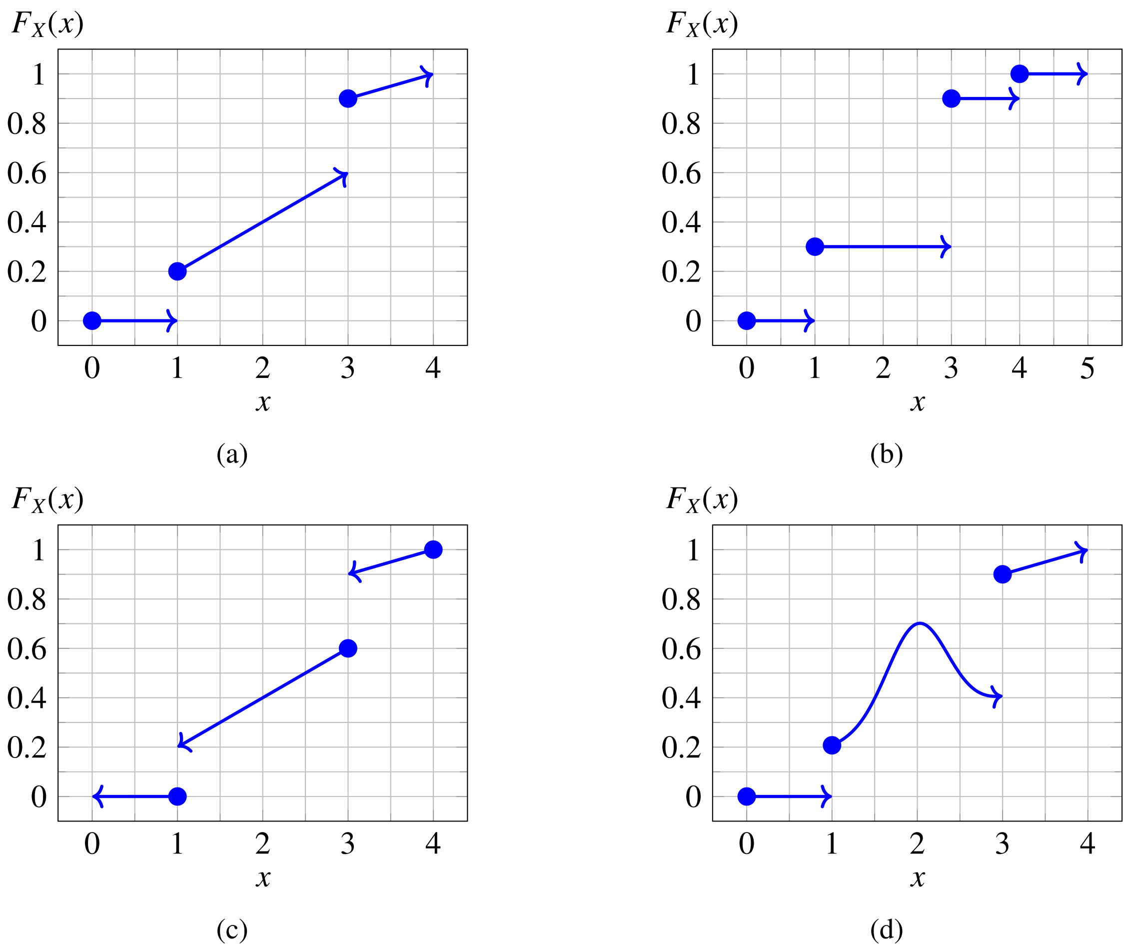


Abbildung 1: Welche Plots zeigen Verteilungsfunktionen?

Lösung:
a) und b) zeigen Verteilungsfunktionen. c) ist nicht rechtsseitig stetig, also keine Verteilungsfunktion d) ist nicht monoton wachsend, also ebenfalls keine Verteilungsfunktion

b) ist die Verteilungsfunktion einer diskreten Verteilung, da hier eine Stufenfunktion zu sehen ist. Die Verteilung ist $\{p_1, p_3, p_4\} = \{0.3, 0.6, 0.1\}$ *(und* $p_k = 0$ *für alle anderen k)*

Multiple Choice Aufgaben

Bei den folgenden Multiple Choice Aufgaben sind die richtigen Aussagen anzugeben. Dabei kann genau eine Aussage zutreffen, mehrere Aussagen können zutreffen oder aber keine der Aussagen trifft zu. Die Aufgabe ist richtig gelöst, wenn alle zutreffenden Aussagen angegeben wurden, jedoch keine nicht zutreffende Aussage angegeben wurde.

Für die Bewertung des Übungsblatts gibt jede richtige Antwort 1 Punkt, jede falsche Antwort gibt 1 Punkt Abzug. Je Aufgabe können minimal 0 Punkte erreicht werden.

Aufgabe 7 (3 Bonuspunkte)

In einer Studie von Vanderpump et al. (Clinical Endocrinology 1995, 43, 55–69) wurden Todesereignisse innerhalb von 10 Jahren bei RaucherInnen (139 von 582) und NichtraucherInnen (230 von 732) erhoben. Berechne die bedingten Wahrscheinlichkeiten für Tod nach 10 Jahren im Laplaceschen Wahrscheinlichkeitsraum, bedingt auf RaucherInnenstatus, und gib an, welche Aussagen zutreffen.

(a) Die bedingten Wahrscheinlichkeit für Tod nach 10 Jahren ist höher für RaucherInnen als für NichtraucherInnen. **Lsg:** falsch

(b) Deine Analyse zeigt, dass das Rauchverhalten einen kausalen Einfluss auf die Sterbewahrscheinlichkeit hat. **Lsg:** falsch

(c) Die Ereignisse *Tod nach 10 Jahren* und *RaucherIn* sind statistisch unabhängig. **Lsg:** falsch

---

Lösung:

$$P(Tod|Rauchen) = 139/582 = 0.239$$

und

$$P(Tod|nicht\ Rauchen) = 230/732 = 0.314$$

also sind a) und c) falsch.
b) ist ebenfalls falsch, da diese Analyse keine Kausalität zeigen kann.

Aufgabe 8 (4 Bonuspunkte)

Drei SpielerInnen X, Y und Z spielen Skat und haben jeweils 10 zufällige Karten aus einem gemeinsamen, 32 Karten umfassenden Skatdeck (jeweils 7, 8, 9, 10, Bube, Dame, König, Ass in den 4 Farben Karo, Herz, Pik und Kreuz) erhalten. Die verbleibenden beiden Karten bilden den sogenannten *Skat*. Betrachte die folgenden Ereignisse:

- A = "Im Skat liegt exakt ein Bube"
- B = "X erhält mehr Buben als Y"
- C = "X erhält 4 Buben"
- D = "X erhält 3 Buben"
- E = "Im Skat liegt mindestens ein Bube"

Gib an, welche Aussagen korrekt sind:

(a) $P(C) < P(D|A)$ **Lsg:** wahr

(b) $P(C) > P(D|E)$ **Lsg:** falsch

(c) A und B sind unabhängig. **Lsg:** falsch

(d) D und B sind unabhängig. **Lsg:** falsch

Lösung:

$$P(C) = \frac{\binom{4}{4}\binom{28}{6}}{\binom{32}{10}} = 0.005839822$$

$$P(D|A) = \frac{\binom{3}{3}\binom{27}{7}}{\binom{30}{10}} = 0.02955665$$

Definiere A' = "Im Skat liegen exakt zwei Buben". Dann sind A und A' disjunkt.

$$P(A) = \frac{\binom{4}{1}\binom{28}{1}}{\binom{32}{2}} = 0.2258065$$

$$P(A') = \frac{\binom{4}{2}\binom{28}{0}}{\binom{32}{2}} = 0.01209677$$

$$P(D|A') = 0$$

$$
\begin{aligned}
P(D|E) &= \frac{P(D \cap (A \cup A'))}{P(A \cup A')} \\
&= \frac{P((D \cap A) \cup (D \cap A')))}{P(A \cup A')} \\
&= \frac{P(D|A)P(A) + P(X|A')P(A')}{P(A \cup A')} \\
&= \frac{0.02955665 * 0.2258065}{0.2258065 + 0.01209677} \\
&= 0.02805377
\end{aligned}
$$

Also ist a) richtig und b) falsch.

d) ist offensichtlich falsch, da $D \subset B$ (und $P(B) \neq 1$).

Definiere B' = "Y erhält mehr Buben als X". Aus Symmetrie folgt $P(B) = P(B')$ und außerdem gilt $B'' := \Omega \setminus (B \cap B')$ = "X und Y erhalten gleich viele Buben". Also ist $P(B) = \frac{1-P(B'')}{2}$ und die Aussage c) reduziert sich auf "A und B'' sind unabhängig". Bezeichne $\beta(X)$ die Anzahl der Buben die Spieler X erhält. Es ist

$$
\begin{aligned}
P(B'') &= P(\beta(X) = 0 \wedge \beta(Y) = 0) + P(\beta(X) = 1 \wedge \beta(Y) = 1) + P(\beta(X) = 2 \wedge \beta(Y) = 2) \\
&= \frac{\binom{4}{4}\binom{28}{8}}{\binom{32}{12}} + P(\beta(Y) = 1|\beta(X) = 1)P(\beta(X) = 1) + P(\beta(Y) = 2|\beta(X) = 2)P(\beta(X) = 2) \\
&= 0.01376529 + \frac{\binom{3}{1}\binom{19}{9}}{\binom{22}{10}} \cdot \frac{\binom{4}{1}\binom{28}{9}}{\binom{32}{10}} + \frac{\binom{2}{2}\binom{20}{8}}{\binom{22}{10}} \cdot \frac{\binom{4}{2}\binom{32}{8}}{\binom{32}{10}} \\
&= 0.01376529 + 0.1835373 + 0.05631257 = 0.2536152
\end{aligned}
$$

und

$$
\begin{aligned}
P(B''|A) &= P(\beta(X) = 0 \wedge \beta(Y) = 0|A) + P(\beta(X) = 1 \wedge \beta(Y) = 1|A) + P(\beta(X) = 2 \wedge \beta(Y) = 2|A) \\
&= P(\beta(Y) = 0|A \wedge \beta(X) = 0)P(\beta(X) = 0|A) \\
&+ P(\beta(Y) = 1|A \wedge \beta(X) = 1)P(\beta(X) = 1|A) \\
&+ 0 \\
&= \frac{\binom{3}{0}\binom{17}{10}}{\binom{20}{10}} \cdot \frac{\binom{3}{0}\binom{27}{10}}{\binom{30}{10}} + \frac{\binom{2}{1}\binom{18}{9}}{\binom{20}{10}} \cdot \frac{\binom{3}{1}\binom{27}{9}}{\binom{30}{10}} \\
&= 0.02955665 + 0.2463054 = 0.2758621
\end{aligned}
$$

Also ist Aussage c) falsch.