Aufgabe 1

Ein Streamingdienst entwirft einen neuen Algorithmus zur Serienempfehlung. Für eine zufällig ausgewählte Nutzerin bzw. einen zufällig ausgewählten Nutzer gilt: Mit einer Wahrscheinlichkeit von 0,3 klickt die Person nach dem Einloggen mindestens eine empfohlene Serie an. Der Streamingdienst betrachtet 30 zufällig ausgewählte Nutzerinnen und Nutzer. Es wird angenommen, dass das Klickverhalten einer Person keinen Einfluss auf das Klickverhalten anderer Personen hat. Seien

X_i : „Die i-te Person klickt mindestens eine empfohlene Serie an" und

Y : „Anzahl der Personen, die mindestens eine empfohlene Serie anklicken".

a) Wie ist die Zufallsvariable X_i verteilt (Verteilungstyp und -parameter)?

$$X_i := \begin{cases} 1, & \text{falls die } i\text{-te Person mindestens eine empfohlene Serie anklickt,} \\ 0, & \text{sonst.} \end{cases}$$

mit $\mathbb{P}(X_i = 1) = 0{,}3$ also $X_i \sim \text{Bernoulli}(0{,}3)$.

b) Was bedeutet es, wenn man davon spricht, dass die Zufallsvariablen $X_1, X_2, ..., X_{30}$ „iid" sind? Ist diese Annahme laut Aufgabenstellung erfüllt?

Die Bezeichnung *iid* steht für *independent and identically distributed*. Das bedeutet, dass die Zufallsvariablen $X_1, \ldots, X_{30}$ stochastisch unabhängig sind und alle dieselbe Wahrscheinlichkeitsverteilung besitzen. Im Kontext der Aufgabe ist dies erfüllt, denn jede Person klickt *unabhängig* von den anderen *mit derselben Wahrscheinlichkeit* $\pi = 0{,}3$ mindestens eine empfohlene Serie an.

c) Definieren Sie die Zufallsvariable Y als Funktion der Zufallsvariablen $X_1,...,X_{30}$. Wie ist die Zufallsvariable Y verteilt (Verteilungstyp und -parameter)?

$Y := \sum_{i=1}^{30} X_i$, mit $X_i \overset{iid}{\sim} \text{Bernoulli}(0{,}3)$ für alle $i = 1, ..., 30$.
$\Rightarrow Y \sim \text{B}(30; 0{,}3)$.

d) Berechnen Sie die Wahrscheinlichkeit, dass

i) genau 5 Personen mindestens eine empfohlene Serie anklicken,

ii) höchstens 18 Personen mindestens eine empfohlene Serie anklicken,

iii) mindestens 9 Personnen mindestens eine empfohlene Serie anklicken,

vi) mehr als 10, aber weniger als 25 Personen mindestens eine empfohlene Serie anklicken.

i) $\mathbb{P}(Y = 5) = \binom{30}{5} \cdot 0{,}3^5 \cdot (1 - 0{,}3)^{30-5} = 0{,}0464$

ii) $\mathbb{P}(Y \leq 18) = \mathrm{F}(18) = 0{,}9998$

iii) $$\begin{aligned}\mathbb{P}(Y \geq 9) &= 1 - \mathbb{P}(Y < 9) = 1 - \mathbb{P}(Y \leq 8) \\ &= 1 - F(8) = 1 - 0{,}4315 = 0{,}5685\end{aligned}$$

iv) $$\begin{aligned}\mathbb{P}(10 < Y < 25) &= \mathbb{P}(Y < 25) - \mathbb{P}(Y \leq 10) = \mathbb{P}(Y \leq 24) - \mathbb{P}(Y \leq 10) \\ &= F(24) - F(10) = 1 - 0{,}7304 = 0{,}2696\end{aligned}$$

Aufgabe 2

Klickt ein Kunde in der App eines Lieferservices auf „Bestellung abschicken“, schlägt manchmal die Zahlung fehl, sodass der Zahlungsvorgang wiederholt werden muss. Die Wahrscheinlichkeit, dass ein einzelner Zahlungsversuch erfolgreich ist, beträgt 0,8. Die Ausgänge der Zahlungsversuche sind unabhängig. Sei

X : „Anzahl der Zahlungsversuche bis zur ersten erfolgreichen Zahlung“

a) Wie ist die Zufallsvariable X verteilt (Verteilungstyp und -parameter)?

$X \sim \mathrm{G}(\pi)$ mit $\pi = 0{,}8$

b) Berechnen Sie die Wahrscheinlichkeit, dass

i) die erste Zahlung bereits beim ersten Versuch erfolgreich ist,

ii) mindestens zwei Fehlversuche auftreten, bevor die Zahlung erfolgreich ist.

i) $\mathbb{P}(X = 1) = (1 - 0{,}8)^{1-1} \cdot 0{,}8^1 = 0{,}8.$

ii)
$$\begin{aligned}\mathbb{P}(X \geq 3) &= 1 - \mathbb{P}(X \leq 2)\\ &= 1 - (\mathbb{P}(X = 1) + \mathbb{P}(X = 2))\\ &= 1 - (0,8 + (1 - 0{,}8) \cdot 0{,}8) = 0{,}04\end{aligned}$$

c) Bestimmen Sie den Erwartungswert von X und interpretieren Sie diesen im Kontext der Aufgabe.

$\mathbb{E}(X) = \frac{1}{0{,}8} = 1{,}25$

Wird der Zahlungsprozess unendlich oft wiederholt, so sind im Mittel mit 1,25 fehlgeschlagenen Zahlungsversuchen pro Bestellung zu rechnen, bevor die Zahlung erfolgreich abgeschlossen wird.

d) Zeigen Sie durch einen formalen Rechenschritt, dass $\lim_{x\to\infty} F(x) = 1$, gilt, wobei F die Verteilungsfunktion von X beschreibt.

$$\begin{aligned}\lim_{x\to\infty} F(x) &= \lim_{x\to\infty} \mathbb{P}(X \leq x) = \lim_{x\to\infty} \sum_{k=1}^{x} \mathbb{P}(X = k) = \sum_{k=1}^{\infty} \mathbb{P}(X = k)\\ &= \sum_{k=1}^{\infty} (1 - 0{,}8)^{k-1} \cdot 0{,}8 = 0{,}8 \cdot \sum_{k=1}^{\infty} 0{,}2^{k-1} = 0{,}8 \cdot \sum_{k=0}^{\infty} 0{,}2^{k} = 0{,}8 \cdot \frac{1}{1 - 0{,}2} = 1\end{aligned}$$

(vgl. Screencast 27)

Aufgabe 3

Eine Person erhält einen Katalog mit 10 Prüfungsfragen, von denen 6 so schwer sind, dass sie von niemandem gelöst werden können. Die Person darf nun zufällig 3 Fragen aus dem Katalog für die Prüfung auswählen. Sei

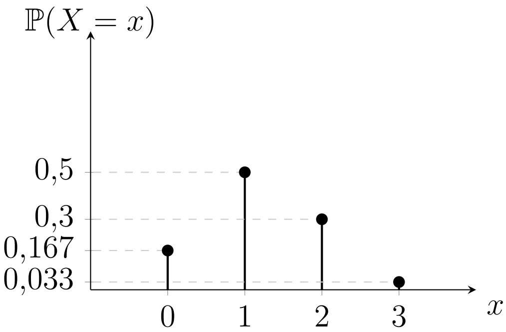
$$X : \text{„Anzahl der ausgewählten lösbaren Fragen"}.$$

a) Wie ist die Zufallsvariable X verteilt (Verteilungstyp und -parameter)?

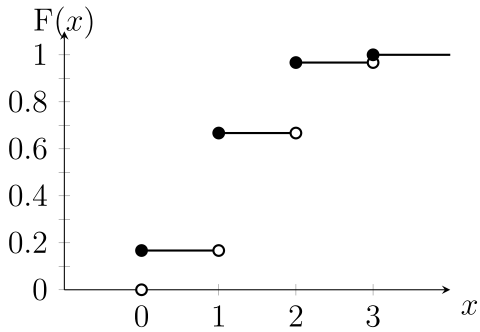
$X \sim \mathrm{H}(n; N; M)$ mit $n = 3, N = 10, M = 4$.

b) Skizzieren Sie die Wahrscheinlichkeitsfunktion und die Verteilungsfunktion von X.

Wahrscheinlichkeitsfunktion von X



Verteilungsfunktion von X



Nebenrechnungen:

$$\mathbb{P}(X = 0) = \frac{\binom{4}{0} \cdot \binom{10-4}{3-0}}{\binom{10}{3}} = \frac{1}{6} \approx 0{,}167 \quad \mathbb{P}(X = 1) = \frac{\binom{4}{1} \cdot \binom{10-4}{3-1}}{\binom{10}{3}} = \frac{1}{2} = 0{,}5$$

$$\mathbb{P}(X = 2) = \frac{\binom{4}{2} \cdot \binom{10-4}{3-2}}{\binom{10}{3}} = \frac{3}{10} = 0{,}3 \quad \mathbb{P}(X = 3) = 1 - \left(\frac{1}{6} + \frac{1}{2} + \frac{3}{10}\right) = \frac{1}{30} \approx 0{,}033$$

c) Berechnen Sie die Wahrscheinlichkeit, dass

   i) die Person alle 3 Fragen richtig lösen kann,

   ii) die Person genau 3 lösbare Fragen auswählt,

   iii) die Person mindestens eine lösbare Frage auswählt,

   iv) die zweite von der Person ausgewählte Frage lösbar ist.

i) Es ist keine Aussage möglich.

ii) $\mathbb{P}(X = 3) = \dfrac{1}{30}$

iii) $\mathbb{P}(X \geq 1) = 1 - \mathbb{P}(X < 1) = 1 - \mathbb{P}(X = 0) = 1 - \dfrac{1}{6} = \dfrac{5}{6}$

iv) Wir definieren die Ereignisse

$$\begin{aligned} B &:= \text{„Die zweite Frage ist lösbar."} \\ A &:= \text{„Die erste Frage ist lösbar"} \\ \bar{A} &:= \text{„Die erste Frage ist nicht lösbar"} \end{aligned}$$

Mit dem Satz der totalen Wahrscheinlichkeit folgt

$$\mathbb{P}(B) = \mathbb{P}(A) \cdot \mathbb{P}(B|A) + \mathbb{P}(\bar{A}) \cdot \mathbb{P}(B|\bar{A}) = \frac{4}{10} \cdot \frac{3}{9} + \frac{6}{10} \cdot \frac{4}{9} = 0{,}4$$

Aufgabe 4

Teil A

Erfahrungsgemäß erscheinen 10% der Fluggäste mit gültigem Ticket nicht zum Abflug, wobei das (Nicht-)Erscheinen einer Personen keinen Einfluss auf das (Nicht-)Erscheinen anderer Personen hat. Auf Basis dieser Information verkauft die Fluggesellschaft 29 Flugtickets für 26 verfügbare Sitzplätze eines Flugs von Mannheim nach Berlin. Sei

X : „Anzahl der erscheinenden Passagiere für Flug von Mannheim nach Berlin".

a) Wie ist die Zufallsvariable X verteilt (Verteilungstyp und -parameter)? Wieso ist die Annahme „*Das (Nicht-)Erscheinen einer Personen hat keinen Einfluss auf das (Nicht-)Erscheinen anderer Personen.*" hier wichtig?

Nur unter dieser Annahme können wir die Bernoulli-Zufallsvariablen, die das Erscheinen jeder einzelnen Person mit Erfolgswahrscheinlichkeit 0,9 modellieren, als *unabhängig* betrachten. Daraus folgt dann

$$X \sim \mathrm{B}(29; 0{,}9).$$

b) Bestimmen Sie die Wahrscheinlichkeit, dass alle erscheinenden Fluggäste einen Sitzplatz erhalten.

Sei $H := 29 - X$ (H : „Anzahl der nicht erscheinenden Passagiere für Flug von Mannheim nach Berlin"). Mit der *Symmetrieeigenschaft* der Binomialverteilung gilt nun $H \sim \mathrm{B}(29; 0{,}1)$. Also gilt

$$\begin{aligned}
\mathbb{P}(X \leq 26) &= \mathbb{P}(29 - H \leq 26), \quad \text{weil } H := 29 - X \Leftrightarrow X = 29 - H \\
&= \mathbb{P}(29 - H - 29 \leq 26 - 29) \\
&= \mathbb{P}(-H \leq -3) = \mathbb{P}(H \geq 3) \\
&= 1 - \mathbb{P}(H < 3) = 1 - \mathbb{P}(H \leq 2) = 1 - F_H(2) = 1 - 0{,}435 = 0{,}565,
\end{aligned}$$

wobei F_H die Verteilungsfunktion von H ist.

c) Die Fluggesellschaft bietet zusätzlich einen Flug von Mannheim nach Sylt an, bei dem 30 Tickets für 26 Plätze verkauft wurden. Ob die Passagiere für den Flug von Mannheim nach Berlin erscheinen oder nicht, hat keinen Einfluss auf das (Nicht-) Erscheinen der Passagiere des Flugs von Mannheim nach Sylt. Bestimmen Sie die Verteilung (Verteilungstyp und -parameter) der Zufallsvariable

G : „Anzahl erscheinender Passagiere beider Flüge".

Sei $\tilde{X}$: „Anzahl der erscheinenden Passagiere für Flug von Mannheim nach Sylt" mit $\tilde{X} \sim \mathrm{B}(30; 0{,}9)$. Per Annahme gilt, dass $\tilde{X}$ und X stochastisch unabhängig sind. Mit der *Additionseigenschaft (Reproduktionseigenschaft) der Binomialverteilung* folgt

$$\tilde{X} + X = G \sim \mathrm{B}(59; 0{,}9).$$

Teil B

Für die Fluggesellschaft ist nun ausschließlich relevant, ob es zu einer Überbuchung auf dem Flug nach Berlin kommt oder nicht.

d) Definieren Sie auf Basis der Zufallsvariable X eine geeignete Zufallsvariable D, die das Auftreten einer Überbuchung modelliert. Wie ist die Zufallsvariable D verteilt (Verteilungstyp und -parameter)?

$$D := \begin{cases} 1, & \text{falls } X > 26 \text{ (Überbuchung)}, \\ 0, & \text{falls } X \leq 26 \text{ (keine Überbuchung)}. \end{cases}$$

$$\mathbb{P}(D = 1) = \mathbb{P}(X > 26) = 1 - \mathbb{P}(X \leq 26) \overset{b)}{=} 0,435 \Rightarrow D \sim \mathrm{Bernoulli}(0{,}435).$$

e) Berechnen Sie den Erwartungswert von D und interpretieren Sie diesen im Kontext der Aufgabe.

$\mathbb{E}(D) = 0 \cdot \mathbb{P}(D = 0) + 1 \cdot \mathbb{P}(D = 1) = 0,435$

Wendet die Fluggesellschaft diese Buchungspolitik (für 26 Sitzplätze 29 verkaufte Tickets) auf eine sehr große Anzahl von Flügen von Mannheim nach Berlin an, so ist im Mittel bei etwa 43,5% der Flüge eine Überbuchung zu erwarten.