

Lösungen zu Tutoriumsblatt 10

Aufgabe T1 (Minimum und Maximum von geometrisch verteilten Zufallsvariablen)

Es seien X_1, \dots, X_n unabhängige, identisch verteilte Zufallsvariablen mit $X_1 \sim G(p)$ für ein $p \in (0, 1)$. Wir definieren $Y := \min\{X_j \mid j \in [n]\}$ und $Z := \max\{X_j \mid j \in [n]\}$. Zeigen Sie:

(a) $Y \sim G(1 - (1 - p)^n)$

(b) $\mathbb{P}(Z = k) = (1 - (1 - p)^{k+1})^n - (1 - (1 - p)^k)^n$, $k \in \mathbb{N}$

Berechnen Sie außerdem den Erwartungswert $\mathbb{E}[Z]$.

Lösungsvorschlag:

(a) Zunächst ist $X_1 \sim G(p)$ äquivalent zu

$$\mathbb{P}(X_1 > k) = (1 - p)^{k+1}, \quad k \geq 0,$$

denn es gilt

$$\mathbb{P}(X_1 > k) = 1 - \mathbb{P}(X_1 \leq k) = 1 - \sum_{j=0}^k (1 - p)^j p = 1 - p \frac{1 - (1 - p)^{k+1}}{1 - (1 - p)} = (1 - p)^{k+1}.$$

Wegen der Unabhängigkeit der X_i gilt für $k \geq 0$

$$\mathbb{P}(Y > k) = \mathbb{P}(X_1 > k, \dots, X_n > k) = \mathbb{P}(X_1 > k) \cdot \dots \cdot \mathbb{P}(X_n > k)$$

und daher wegen der Verteilungsgleichheit

$$\mathbb{P}(Y > k) = \mathbb{P}(X_1 > k)^n = \left((1 - p)^{k+1}\right)^n = ((1 - p)^n)^{k+1} = (1 - (1 - (1 - p)^n))^{k+1}.$$

Das zeigt die Behauptung.

(b) Zunächst schreiben wir $\mathbb{P}(Z = k) = \mathbb{P}(Z \leq k) - \mathbb{P}(Z \leq k - 1)$ (*) für $k \in \mathbb{N}$ und berechnen die entsprechenden Wahrscheinlichkeiten separat. Aufgrund der Unabhängigkeit der X_i , deren identischer Verteilung und der Anfangsüberlegung aus Teil (a) erhalten wir

$$\mathbb{P}(Z \leq k) = \mathbb{P}(X_1 \leq k, \dots, X_n \leq k) = \mathbb{P}(X_1 \leq k) \cdot \dots \cdot \mathbb{P}(X_n \leq k) = \left(1 - (1 - p)^{k+1}\right)^n.$$

Analog erhalten wir $\mathbb{P}(Z \leq k - 1) = (1 - (1 - p)^k)^n$. Daraus folgt die Behauptung durch Einsetzen in (*).

Für die Berechnung von $\mathbb{E}[Z]$ schreiben wir $q := 1 - p$ und erhalten

$$\begin{aligned} \mathbb{E}[Z] &= \sum_{k=1}^{\infty} \mathbb{P}(Z \geq k) = \sum_{k=1}^{\infty} (1 - \mathbb{P}(Z \leq k-1)) = \sum_{k=0}^{\infty} (1 - \mathbb{P}(Z \leq k)) \\ &= \sum_{k=0}^{\infty} \left(1 - (1 - q^{k+1})^n\right) = \sum_{k=0}^{\infty} \left(1 - \sum_{j=0}^n \binom{n}{j} (-1)^j q^{j(k+1)}\right) \\ &= \sum_{k=0}^{\infty} \sum_{j=1}^n \binom{n}{j} (-1)^{j+1} q^{j(k+1)} \\ &= \sum_{j=1}^n \binom{n}{j} (-1)^{j+1} q^j \sum_{k=0}^{\infty} (q^j)^k \\ &= \sum_{j=1}^n \binom{n}{j} (-1)^{j+1} \frac{q^j}{1 - q^j}. \end{aligned}$$

Hier wurde im ersten Schritt Aufgabe T3 von Tutoriumsblatt 4 angewendet.

Aufgabe T2 (Wartezeiten bei Würfelwürfen)

Ein Würfel wird solange geworfen, bis die erste „6“ auftritt. Bestimmen Sie die Wahrscheinlichkeit, dass davor genau ℓ mal die „3“ auftritt, $\ell \in \mathbb{N}_0$.

Hinweis: Die (allgemeinen) Binomialkoeffizienten $\binom{\alpha}{k} := \frac{\alpha(\alpha-1)\dots(\alpha-k+1)}{k!}$, $\alpha \in \mathbb{R}$, $k \in \mathbb{N}$, sind gegeben durch die Koeffizienten der Taylorreihe von $(1+x)^\alpha$, d.h.

$$(1+x)^\alpha = \sum_{k=0}^{\infty} \binom{\alpha}{k} x^k, \quad |x| < 1.$$

Die Konvergenz dieser Reihe (sog. *Binomialreihe*) für $|x| < 1$ folgt aus dem Quotientenkriterium. All dies dürfen Sie ohne Nachweis verwenden.

Lösungsvorschlag:

Es bezeichne X die zufällige Anzahl an „3“-en vor der ersten „6“. Falls im k -ten Wurf (mit $k > \ell$) die „6“ zum ersten mal geworfen wird, so müssen, damit das Ereignis $\{X = \ell\}$ eintritt, genau ℓ der $k-1$ vorherigen gewürfelten Zahlen eine „3“ sein. Die Wahrscheinlichkeit für das gesuchte Ereignis ist somit

$$\begin{aligned} \mathbb{P}(X = \ell) &= \sum_{k=\ell+1}^{\infty} \binom{k-1}{\ell} \left(\frac{1}{6}\right)^\ell \left(\frac{4}{6}\right)^{k-1-\ell} \cdot \frac{1}{6} \\ &= \left(\frac{1}{6}\right)^{\ell+1} \sum_{k=0}^{\infty} \binom{\ell+k}{\ell} \left(\frac{2}{3}\right)^k \\ &= \left(\frac{1}{6}\right)^{\ell+1} \sum_{k=0}^{\infty} \frac{(\ell+k)(\ell+k-1) \cdot \dots \cdot (\ell+1)}{k!} \left(\frac{2}{3}\right)^k \\ &= \left(\frac{1}{6}\right)^{\ell+1} \sum_{k=0}^{\infty} (-1)^k \frac{(-\ell-1)(-\ell-2) \cdot \dots \cdot (-\ell-k)}{k!} \left(\frac{2}{3}\right)^k \\ &\stackrel{(1)}{=} \left(\frac{1}{6}\right)^{\ell+1} \sum_{k=0}^{\infty} \binom{-\ell-1}{k} \left(-\frac{2}{3}\right)^k \\ &\stackrel{(2)}{=} \left(\frac{1}{6}\right)^{\ell+1} \left(1 - \frac{2}{3}\right)^{-\ell-1} \\ &= \left(\frac{1}{2}\right)^{\ell+1}. \end{aligned}$$

In (1) wurde die Definition des (allgemeinen) Binomialkoeffizienten verwendet. Im Anschluss daran wurde in (2) die Binomialreihe verwendet (siehe Hinweis).

Aufgabe T3 (Poissonverteilung zur Modellierung seltener Ereignisse)

Man beobachtet pro Jahr im Mittel ein Erdbeben der Mindeststärke 8 auf der Richter-Skala. Wir nehmen an, dass die Anzahl solcher Erdbeben pro Jahr eine Poisson-verteilte Zufallsvariable ist und dass die entsprechenden Anzahlen von Erdbeben in unterschiedlichen Jahren stochastisch unabhängig sind.

- Mit welcher Wahrscheinlichkeit gibt es im nächsten Jahr mehr als ein Erdbeben der Mindeststärke 8?
- Welche Verteilung besitzt die Anzahl X derjenigen Jahre (unter den nächsten 100 Jahren), in denen mehr als zwei solcher Erdbeben stattfinden?
- Wie viele Jahre mit mehr als zwei solcher Erdbeben kann man in den nächsten 100 Jahren erwarten?

Lösungsvorschlag:

Sei Y_i die Anzahl der auftretenden Erdbeben im i -ten Folgejahr. Dann gilt $Y_i \sim \text{Po}(\lambda)$ mit noch unbekanntem Parameter λ , und Y_1, Y_2, \dots sind unabhängig. Wir wissen, dass man im Mittel ein schweres Erdbeben beobachtet, d.h. $\mathbb{E}[Y_i] \stackrel{!}{=} 1$. Wegen $\mathbb{E}[Y_i] = \lambda$ (siehe Satz 16.3) folgt direkt, dass $\lambda = 1$ gilt.

- (a) Gesucht ist die Wahrscheinlichkeit, dass $Y_1 > 1$. Es gilt

$$\begin{aligned}\mathbb{P}(Y_1 > 1) &= 1 - \mathbb{P}(Y_1 \leq 1) = 1 - \mathbb{P}(Y_1 = 1) - \mathbb{P}(Y_1 = 0) \\ &= 1 - e^{-1} \frac{1}{1!} - e^{-1} \frac{1}{0!} \\ &= 1 - 2e^{-1} \approx 0.2642.\end{aligned}$$

- (b) Sei A_i das Ereignis, dass im i -ten Folgejahr mehr als zwei solcher Erdbeben stattfinden, $1 \leq i \leq 100$. Die Ereignisse A_1, \dots, A_{100} sind unabhängig und besitzen die gleiche Wahrscheinlichkeit

$$\mathbb{P}(A_1) = \mathbb{P}(Y_1 > 2) = 1 - \mathbb{P}(Y_1 \leq 2) = 1 - e^{-1} \frac{1}{2!} - e^{-1} \frac{1}{1!} - e^{-1} \frac{1}{0!} = 1 - \frac{5}{2e} \approx 0.08.$$

Wegen $X = \sum_{i=1}^{100} \mathbb{1}_{A_i}$ gilt $X \sim \text{Bin}(100, 1 - \frac{5}{2e})$.

- (c) Da $X \sim \text{Bin}(100, 1 - \frac{5}{2e})$ gilt $\mathbb{E}[X] = 100 \cdot (1 - \frac{5}{2e}) \approx 8$.

Aufgabe T4 (Eigenschaften der Poissonverteilung)

Es seien $\lambda, \mu > 0$ und $X \sim \text{Po}(\lambda)$ und $Y \sim \text{Po}(\mu)$ unabhängige Zufallsvariablen. Zeigen Sie, dass dann $X + Y \sim \text{Po}(\lambda + \mu)$ gilt.

Lösungsvorschlag:

Mit der Faltungsformel für unabhängige Zufallsvariablen (Satz 11.8) erhalten wir für $k \in \mathbb{N}_0$

$$\begin{aligned}\mathbb{P}(X + Y = k) &= \sum_{j=0}^k \mathbb{P}(X = j) \mathbb{P}(Y = k - j) = \sum_{j=0}^k e^{-\lambda} \frac{\lambda^j}{j!} e^{-\mu} \frac{\mu^{(k-j)}}{(k-j)!} \\ &= e^{-(\lambda+\mu)} \frac{(\lambda+\mu)^k}{k!} \sum_{j=0}^k \frac{k!}{j!(k-j)!} \frac{\lambda^j}{(\lambda+\mu)^j} \frac{\mu^{(k-j)}}{(\lambda+\mu)^{k-j}}.\end{aligned}$$

Mit $p := \frac{\lambda}{\lambda+\mu}$ erhalten wir also

$$\mathbb{P}(X + Y = k) = e^{-(\lambda+\mu)} \frac{(\lambda+\mu)^k}{k!} \sum_{j=0}^k \binom{k}{j} p^j (1-p)^{k-j} = e^{-(\lambda+\mu)} \frac{(\lambda+\mu)^k}{k!}.$$

Damit besitzt $X + Y$ eine Poissonverteilung mit Parameter $\lambda + \mu > 0$.

Aufgabe T5 (Erzeugende Funktionen I)

Sei $Y \sim \text{Nb}(r, p)$ und $Z \sim U_{[n]}$ für $r \geq 1, p \in (0, 1), n \in \mathbb{N}$.

- Berechnen Sie Erwartungswert und Varianz von Y mit Hilfe der erzeugenden Funktion.
- Berechnen Sie die erzeugende Funktion g_Z von Z .
- Es sei X eine \mathbb{N}_0 -wertige Zufallsvariable mit erzeugender Funktion

$$g_X(t) = \frac{1}{5}t(2 + 3t^2), \quad t \in [-1, 1].$$

Bestimmen Sie die Verteilung von X .

Lösungsvorschlag:

- Aus der Vorlesung wissen wir

$$g_Y(t) = \left(\frac{p}{1 - (1-p)t} \right)^r, \quad |t| < \frac{1}{1-p}.$$

Deshalb folgen

$$g'_Y(t) = r \cdot p^r (1 - t(1-p))^{-r-1} \cdot (1-p)$$

und

$$g''_Y(t) = r(r+1)p^r (1 - t(1-p))^{-r-2} (1-p)^2.$$

Mit Hilfe der Folgerung aus Satz 17.6 erhalten wir

$$\mathbb{E}[Y] = g'_Y(1) = r \cdot p^r \cdot p^{-r-1} \cdot (1-p) = r \frac{1-p}{p}.$$

Da außerdem

$$g''_Y(1) = r(r+1)p^{-2}(1-p)^2.$$

gilt, liefert die Folgerung aus Satz 17.6 für die Varianz von Y

$$\begin{aligned} \mathbb{V}(Y) &= g''_Y(1) + g'_Y(1) - (g'_Y(1))^2 \\ &= r(r+1)p^{-2}(1-p)^2 + rp^{-1}(1-p) - r^2p^{-2}(1-p)^2 \\ &= \frac{r(1-p)}{p} \cdot ((r+1)p^{-1}(1-p) + 1 - rp^{-1}(1-p)) \\ &= \frac{r(1-p)}{p^2}. \end{aligned}$$

- Für $t \in (-1, 1)$ gilt

$$g_Z(t) = \sum_{j=1}^n \mathbb{P}(Z=j)t^j = \sum_{j=1}^n n^{-1}t^j = \frac{t}{n} \sum_{j=1}^n t^{j-1} = \frac{t}{n} \sum_{j=0}^{n-1} t^j = \frac{t - t^{n+1}}{n(1-t)}.$$

- Wir verwenden die erzeugende Funktion und deren Ableitungen, jeweils ausgewertet an der Stelle $t = 0$, um die Verteilung von X zu bestimmen. Dazu verwenden wir Bemerkung (3) nach Definition 17.1, die besagt, dass für alle $n \in \mathbb{N}_0$

$$\mathbb{P}(X = n) = \frac{g_X^{(n)}(0)}{n!} \tag{1}$$

gilt. Also bestimmen wir die Ableitungen von g_X und erhalten

$$\begin{aligned} g'_X(t) &= \frac{2}{5} + \frac{9}{5}t^2, \\ g''_X(t) &= \frac{18}{5}t, \\ g_X^{(3)}(t) &= \frac{18}{5}, \\ g_X^{(k)}(t) &\equiv 0 \quad \text{für alle } k \geq 4 \text{ und } t \in [-1, 1]. \end{aligned}$$

Dann ist die Verteilung von X mit Hilfe von (1) gegeben durch

$$\begin{aligned}\mathbb{P}(X = 0) &= g_X(0) = 0, \\ \mathbb{P}(X = 1) &= g_X'(0) = \frac{2}{5}, \\ \mathbb{P}(X = 2) &= \frac{g_X''(0)}{2} = 0, \\ \mathbb{P}(X = 3) &= \frac{g_X^{(3)}(0)}{6} = \frac{18}{6 \cdot 5} = \frac{3}{5}, \\ \mathbb{P}(X = k) &= 0 \quad \text{für } k \geq 4.\end{aligned}$$

Alternativ können wir die Verteilung direkt anhand der Koeffizienten ablesen, denn nach Definition der erzeugenden Funktion ist

$$\sum_{k=0}^{\infty} \mathbb{P}(X = k)t^k = g_X(t) = \frac{2}{5}t + \frac{3}{5}t^3,$$

sodass wir $\mathbb{P}(X = 1) = \frac{2}{5}$ und $\mathbb{P}(X = 3) = \frac{3}{5}$ direkt ablesen können. Da keine weiteren Potenzen von t auftauchen, erkennen wir außerdem, dass $\mathbb{P}(X = k) = 0$ ist für alle $k \notin \{1, 3\}$.

Aufgabe T6 (Erzeugende Funktionen II)

Sei X eine Zufallsvariable mit Werten in $\{0, 1, 2\}$ und $\mathbb{E}[X] = 1$ sowie $\mathbb{E}[X^2] = \frac{3}{2}$. Zeigen Sie:

(a) Die erzeugende Funktion g_X von X ist gegeben durch

$$g_X(t) = \frac{1}{4}(1+t)^2, \quad |t| \leq 1.$$

(b) Es gibt unabhängige Zufallsvariablen X_1, X_2 mit identischer Verteilung, sodass $X_1 + X_2$ die selbe Verteilung wie X besitzt. Wie heißt die Verteilung von X ?

(c) Die erzeugende Funktion g_{3X} von $3X$ ist gegeben durch

$$g_{3X}(t) = g_X(t^3), \quad |t| \leq 1.$$

(d) Berechnen Sie mithilfe der erzeugenden Funktion den Erwartungswert $\mathbb{E}[3X(3X - 1)]$.

Lösungsvorschlag:

(a) Wir definieren $p_i := \mathbb{P}(X = i)$ für $i \in \{0, 1, 2\}$. Nach Voraussetzung gilt

$$\begin{aligned}1 &= \mathbb{E}[X] = p_1 + 2p_2 \\ \frac{3}{2} &= \mathbb{E}[X^2] = p_1 + 4p_2.\end{aligned}$$

Daraus folgt

$$p_2 = \frac{1}{4}, \quad p_1 = \frac{1}{2}, \quad p_0 = \frac{1}{4}$$

Somit gilt für $|t| \leq 1$

$$g_X(t) = \sum_{k=0}^2 t^k p_k = p_0 + t p_1 + t^2 p_2 = \frac{1}{4} + \frac{1}{2}t + \frac{1}{4}t^2 = \frac{1}{4}(1 + 2t + t^2) = \frac{1}{4}(1+t)^2.$$

(b) Nach Aufgabenteil (a) wissen wir

$$g_X(t) = \frac{1}{4}(1+t)^2 = \frac{1}{2}(1+t) \frac{1}{2}(1+t), \quad |t| \leq 1,$$

wobei die Funktion $h(t) := \frac{1}{2}(1+t)$ die erzeugende Funktion einer $\text{Bin}(1, \frac{1}{2})$ -verteilten Zufallsvariable (bzw. $U(\{0, 1\})$ -verteilten Zufallsvariable) ist. Sind X_1, X_2 stochastisch unabhängig und je $\text{Bin}(1, \frac{1}{2})$ -verteilt, so gilt nach der Multiplikationsformel

$$g_{X_1+X_2}(t) = g_{X_1}(t) g_{X_2}(t) = \frac{1}{2}(1+t) \frac{1}{2}(1+t) = \frac{1}{4}(1+t)^2 = g_X(t), \quad |t| \leq 1.$$

Da die Verteilung einer \mathbb{N}_0 -wertigen Zufallsvariablen eindeutig durch ihre erzeugende Funktion festgelegt ist (Eindeutigkeitssatz), folgt die Verteilungsgleichheit von X und $X_1 + X_2$. Also gilt $X \sim \text{Bin}(2, \frac{1}{2})$.

(c) Es gilt

$$g_{3X}(t) = \sum_{k=0}^2 t^{3k} p_k = \sum_{k=0}^2 (t^3)^k p_k = g_X(t^3).$$

(d) Nach dem Satz 17.6 (angewandt auf die Zufallsvariable $3X$) gilt

$$\mathbb{E}[3X(3X-1)] = g_{3X}''(1-).$$

Wir benötigen also die zweite Ableitung von g_{3X} :

$$\begin{aligned} g_{3X}(t) &= \frac{1}{4}(1+t^3)^2, \\ g_{3X}'(t) &= \frac{3t^2}{2}(1+t^3) = \frac{3}{2}t^2 + \frac{3}{2}t^5, \\ g_{3X}''(t) &= 3t + \frac{15}{2}t^4. \end{aligned}$$

Somit ist

$$\mathbb{E}[3X(3X-1)] = g_{3X}''(1-) = g_{3X}''(1) = \frac{21}{2}.$$