ETH Zürich HS 2015 D-INFK
Prof. Dr. P. Embrechts

Wahrscheinlichkeit und Statistik

Lösungen Serie 5

Lösung 5-1. (a) Wir berechnen $P[X = 100]$ und $P[X = 250]$:

$$\begin{aligned} P[X = 100] &= P[X = 100, Y = 0] + P[X = 100, Y = 100] + P[X = 100, Y = 200] \\ &= 0.20 + 0.10 + 0.40 = 0.7 \\ P[X = 250] &= P[X = 250, Y = 0] + P[X = 250, Y = 100] + P[X = 250, Y = 200] \\ &= 0.05 + 0.15 + 0.10 = 0.3 \end{aligned}$$

und es folgt

$$p_X(x_i) = \begin{cases} 0.7 & \text{falls } x_i = 100, \\ \\ 0.3 & \text{falls } x_i = 250. \end{cases}$$

(b) Es gilt $p(x_i, y_j) \neq p_X(x_i)p_Y(y_j)$ für $x_i \in \{100, 250\},\ y_j \in \{0, 100, 200\}$. Um dies einzusehen, berechnen wir zuerst die Gewichtsfunktion $p_Y(y_j)$. Die Zufallsvariable Y nimmt die Werte 0, 100 und 200 an. Es gilt

$$P[Y = 0] = P[X = 100, Y = 0] + P[X = 250, Y = 0] = 0.20 + 0.05 = 0.25$$

$$P[Y = 100] = P[X = 100, Y = 100] + P[X = 250, Y = 100] = 0.10 + 0.15 = 0.25$$

$$P[Y = 200] = P[X = 100, Y = 200] + P[X = 250, Y = 200] = 0.40 + 0.10 = 0.50$$

und somit

$$p_Y(y_j) = \begin{cases} 0.25 & \text{falls } y_j = 0, \\ 0.25 & \text{falls } y_j = 100, \\ 0.50 & \text{falls } y_j = 200. \end{cases}$$

Beachte: $\sum_{x_i \in \mathcal{W}(X)} p_X(x_i) = 1$, dasselbe gilt für $\sum_{y_j \in \mathcal{W}(Y)} p_Y(y_j) = 1$. Nun gilt aber (zum Beispiel):

$$p(100, 0) = 0.20 \neq 0.175 = 0.7 \cdot 0.25 = p_X(100)p_Y(0).$$

(c) Wir wissen, dass der Erwartungswert $E[Z_1] = -19$ ist. Somit folgt mit $Z_1 = X_1 - 3X_2 + X_4$:

$$E[Z_1] = E[X_1] - 3E[X_2] + E[X_4] = -19.$$

Da $E[X_1]$, $E[X_2]$ bekannt sind, können wir die obige Gleichung nach $E[X_4]$ auflösen. Daraus folgt $E[X_4] = 3$.

(d) Wir können $\mathrm{Var}[X_2]$ nicht berechnen, da wir die Kovarianz $\mathrm{Cov}(X_i, X_j)$, $i, j = 1, 2, 3$ nicht kennen.

Bemerkung: Falls die Zufallsvariablen $X_1, \ldots, X_3$ stochastisch unabhängig wären, dann wäre es möglich, $\mathrm{Var}[X_2]$ zu berechnen. In diesem Fall gilt $\mathrm{Var}[Z_2] = \mathrm{Var}[X_1] + \mathrm{Var}[2X_2] + \mathrm{Var}[-2X_3] = \mathrm{Var}[X_1] + 4\mathrm{Var}[X_2] + 4\mathrm{Var}[X_3] = 29$. Da $\mathrm{Var}[X_1]$ und $\mathrm{Var}[X_3]$ bekannt sind, erhalten wir $\mathrm{Var}[X_2] = 2$.

(e) Die Kovarianz stochastisch unabhängiger Zufallsvariablen gleich 0 ist.

Lösung 5-2. (a) Wenn das k-te Segment defekt ist und $k \in \{1, \ldots, 19\}$, dann brauchen wir genau k Inspektionen. Ist das Leck im letzte Segment, so wissen wir das auch nach 19 Inspektionen. Also gilt $P[X^1 = k] = \frac{1}{20}$ für $k \in \{1, \ldots, 18\}$ und $P[X^1 = 19] = \frac{2}{20}$.

$$E[X] = \sum_{k=1}^{19} k \cdot P[X = k] = \frac{1}{20} \sum_{k=1}^{18} k + \frac{19}{10} = \frac{18 \cdot 19}{20 \cdot 2} + \frac{19}{10} = 10.45.$$

$$\begin{aligned} \mathrm{Var}[X] &= E[(X^1)^2] - (E[X^1])^2 = \frac{1}{20} \sum_{k=1}^{18} k^2 + \frac{19^2}{10} - (E[X])^2 \\ &= \frac{18 \cdot 19 \cdot 37}{20 \cdot 6} + \frac{361}{10} - (E[X])^2 = 32.348\ldots \end{aligned}$$

(b) *Idee:* Die Inspektion wird jeweils in der “Mitte” des Gebiets durchgeführt, in dem sich das Leck noch befinden kann. Damit wird das Gebiet in jedem Schritt“halbiert”.

Genau genommen kann das Gebiet natürlich nur dann halbiert werden, wenn es $2n$ Segmente enthält. Verbleiben noch $2n + 1$ Segmente, so soll die Inspektion zwischen dem n-ten und dem $(n + 1)$-ten dieser Segmente durchgeführt werden. Wenn das Leck sich in Segment k befindet, braucht man mit dieser Strategie gerade mal 4 Messungen für

$$k \in \{1, 2, 3, 6, 7, 8, 11, 12, 13, 16, 17, 18\}$$

und 5 Messungen für

$$k \in \{4, 5, 9, 10, 14, 15, 19, 20\}.$$

Also gilt $P[Y = 4] = \frac{3}{5}$, $P[Y = 5] = \frac{2}{5}$ und

$$E[Y] = 4\frac{3}{5} + 5\frac{2}{5} = 4.4.$$

$$\mathrm{Var}[X] = E[Y^2] - (E[Y])^2 = 16\frac{3}{5} + 25\frac{2}{5} - 4.4^2 = 0.24.$$

Lösung 5-3. (a) Da p eine Gewichtsfunktion sein soll, muss gelten

$$\begin{aligned} 1 &\overset{!}{=} \sum_{k=2}^{\infty} \sum_{j=1}^{k-1} P[X = j, Y = k] = C \sum_{k=2}^{\infty} \sum_{j=1}^{k-1} \left(\frac{1}{2}\right)^k = C \sum_{j=1}^{\infty} \sum_{k=j+1}^{\infty} \left(\frac{1}{2}\right)^k \\ &= C \sum_{j=1}^{\infty} \left(\frac{1}{2}\right)^{j+1} \sum_{k=0}^{\infty} \left(\frac{1}{2}\right)^k = C \sum_{j=1}^{\infty} \frac{\left(\frac{1}{2}\right)^{j+1}}{1 - \frac{1}{2}} = C \sum_{j=1}^{\infty} \left(\frac{1}{2}\right)^j = C \frac{\frac{1}{2}}{1 - \frac{1}{2}} = C. \end{aligned}$$

Also ist $C = 1$.

(b)

$$p_X(j) = \sum_{k=j+1}^{\infty} p(j, k) = \sum_{k=j+1}^{\infty} \left(\frac{1}{2}\right)^k = \left(\frac{1}{2}\right)^{j+1} \sum_{k=0}^{\infty} \left(\frac{1}{2}\right)^k = \left(\frac{1}{2}\right)^j \quad \text{für } j \geq 1;$$

also ist X geometrisch verteilt mit Parameter $p = \frac{1}{2}$.

$$p_Y(k) = \sum_{j=1}^{k-1} p(j, k) = (k - 1) \left(\frac{1}{2}\right)^k \quad \text{für } k \geq 2.$$

Siehe nächstes Blatt!

(c)

$$p_{X|Y}(j\,|\,k) = \frac{p(j,k)}{p_Y(k)} = \frac{\left(\frac{1}{2}\right)^k}{(k-1)\left(\frac{1}{2}\right)^k} = \frac{1}{k-1} \text{ für } k \geq 2,\ j = 1, \dots, k-1,$$

d.h. gegeben $Y = k$ ist X gleichverteilt auf $\{1, \dots, k-1\}$.

$$p_{Y|X}(k\,|\,j) = \frac{p(j,k)}{p_X(j)} = \frac{\left(\frac{1}{2}\right)^k}{\left(\frac{1}{2}\right)^j} = \left(\frac{1}{2}\right)^{k-j}, \text{ für } j \geq 1,\ k > j.$$

Lösung Challenge Serie 5. Wir bezeichnen mit X_n die (zufällige) Anzahl Passagiere die gestört werden bei n Passagieren. Ferner sei

$$A_n := \{\text{der Sitzplatz des } n\text{-ten Passagiers ist besetzt}\}.$$

Bezeiche $E_n := E[X_n]$ die erwartete Anzahl Passagiere die gestört wurden. Dann gilt die Gleichung

$$E[X_n | A_n^c] = 1 + E_{n-1}.$$

Das Argument geht wie folgt: Gegeben, dass der n-te Passagier den Sitzplatz wechselt (also $+1$), steht nun der $(n-1)$-te Passagier genau vor dem gleichen Problem wie der n-te Passagier vorher und damit noch $+E_{n-1}$ erwartete Sitzwechsel.

Desweiteren gilt auch $E[X_n \mathbb{1}_{A_n}] = 0$, da der n-te Passagier keinen weiteren Passagieren stört, wenn sein Platz schon frei ist. Insgesamt bekommen wir

$$E_n = E[X_n \mathbb{1}_{A_n}] + E[X_n \mathbb{1}_{A_n^c}] = E[X_n | A_n^c] P[A_n^c] = (1 + E_{n-1}) P[A_n^c].$$

Klar ist, dass $P[A_n^c] = \frac{(n-1)!}{n!} = \frac{n-1}{n}$ und die obige rekursive Gleichung ergibt $E_n = \frac{n-1}{2}$. Also in unserem Fall $E_{100} = 99/2$.