Koordinator
Konstantin Rusch

Wahrscheinlichkeit & Statistik

Lösungen 7

1. Seien X und Y zwei diskrete Zufallsvariablen mit folgender gemeinsamer Gewichtsfunktion:

$$p(j,k) = P[X=j, Y=k] = \begin{cases} C\left(\frac{1}{2}\right)^k & \text{für } k = 2,3,\dots \text{ und } j = 1,2,\dots,k-1 \\ 0 & \text{sonst.} \end{cases}$$

a) Bestimmen Sie die Konstante C.

b) Berechnen Sie die Gewichtsfunktionen p_X und p_Y der Randverteilungen von X und Y.

c) Berechnen Sie die bedingte Gewichtsfunktion $p_{X|Y}(j \mid k) = P[X=j \mid Y=k]$ von X, gegeben dass $Y=k$, sowie die bedingte Gewichtsfunktion $p_{Y|X}(k \mid j) = P[Y=k \mid X=j]$ von Y, gegeben dass $X=j$.

Lösung:

a) Da p eine Gewichtsfunktion sein soll, muss gelten

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

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 $p = \frac{1}{2}$.

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

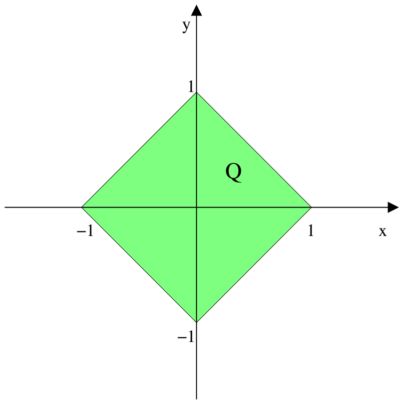
c)

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

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

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

2. Die gemeinsame Dichte $f(x, y)$ zweier Zufallsvariablen X, Y sei im Quadrat Q (vgl. Skizze) konstant und verschwinde ausserhalb von Q.



a) Bestimmen Sie die gemeinsame Dichte von (X, Y).

b) Bestimmen Sie die Randdichten f_X und f_Y der Zufallsvariablen X und Y.

c) Sind die Zufallsvariablen X und Y unabhängig? Begründen Sie ihre Antwort mit einem mathematischen Argument!

d) Was ist die Antwort in **c)**, wenn das Quadrat Q um 45 Grad gedreht wird?

Lösung:

a) Die Fläche von Q ist $4 \cdot \frac{1 \cdot 1}{2} = 2$. Daher ist die gemeinsame Dichte f gegeben durch

$$f(x, y) = \begin{cases} \frac{1}{2} & \text{falls } (x, y) \in Q, \\ 0 & \text{sonst.} \end{cases}$$

b) Für die Randdichte f_X sind 2 Fälle zu unterscheiden. Für $-1 \leq x \leq 0$ gilt

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y)dy = \int_{-1-x}^{1+x} \frac{1}{2}dy = \frac{1}{2}y\Big|_{-1-x}^{1+x} = \frac{1}{2}(1 + x + 1 + x) = 1 + x.$$

Siehe nächste Seite!

Für $0 \leq x \leq 1$ gilt

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)dy = \int_{-1+x}^{1-x} \frac{1}{2}dy = \frac{1}{2}y\Big|_{-1+x}^{1-x} = \frac{1}{2}(1-x+1-x) = 1-x.$$

Also ist

$$f_X(x) = \begin{cases} 1+x & \text{falls } -1 \leq x \leq 0, \\ 1-x & \text{falls } 0 \leq x \leq 1, \\ 0 & \text{sonst.} \end{cases}$$

Aufgrund der Symmetrie ist $f_Y = f_X$, d.h.

$$f_Y(y) = \begin{cases} 1+y & \text{falls } -1 \leq y \leq 0, \\ 1-y & \text{falls } 0 \leq y \leq 1, \\ 0 & \text{sonst.} \end{cases}$$

c) Wegen $f_X(x)f_Y(y) \not\equiv \frac{1}{2} = f(x,y)$ sind die Zufallsvariablen X und Y abhängig.

d) Dank der Symmetrie kann man immer noch schliessen, dass in diesem Fall die Randdichten gleich sind; zudem sind sie aber auch konstant auf dem Intervall $[-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}]$, d.h.

$$f_X(x) = f_Y(x) = \begin{cases} \frac{1}{\sqrt{2}} & \text{falls } -\frac{1}{\sqrt{2}} \leq x \leq \frac{1}{\sqrt{2}}, \\ 0 & \text{sonst.} \end{cases}$$

Wegen $f_X(x)f_Y(y) = \frac{1}{2} = f(x,y)$ sind die Zufallsvariablen X und Y dann unabhängig.

3. Ein häufig benutztes Modell in der Rückversicherung zur Abdeckung grosser Schäden ist ein sogenannter "Excess-of-Loss" Vertrag. Gegen Bezahlung einer Prämie verpflichtet sich dabei die Rückversicherungsgesellschaft, allfällige Schäden, welche ein bestimmtes Level von x_0 CHF übersteigen, zu übernehmen.
Um die Höhe der Prämie zu bestimmen, untersucht die Gesellschaft die Grossschäden (Schäden $> x_0$) des letzten Jahres. In guter Näherung können solche Grossschäden durch eine *Pareto-verteilte* Zufallsvariable X modelliert werden, d.h. für einen Parameter $\alpha > 0$ ist die Verteilungsfunktion von X gegeben durch

$$F_X(x; x_0, \alpha) = \begin{cases} 1 - \left(\frac{x}{x_0}\right)^{-\alpha}, & x \geq x_0, \\ 0, & \text{sonst.} \end{cases}$$

a) Wie teuer ist ein einzelner Grossschaden im Mittel?

b) Die Netto-Jahresprämie berechnet sich durch $P_{net} = E[X]E[N]$. Dabei werde durch $N \sim \text{Poisson}(\lambda)$ die (zufällige) Anzahl der Grossschäden pro Jahr modelliert. Wie gross ist die Prämie P_{net} für $\alpha = 2,\ x_0 = 2 \cdot 10^6$ CHF und $\lambda = 3$?

c) Warum wird in obigem Modell in der Regel $\alpha > 1$ vorausgesetzt?

Lösung: Sei X Pareto-verteilt mit den Parametern x_0 und $\alpha > 0$.

a) Nach Definition des Erwartungswertes für stetige Verteilungen gilt $E[X] = \int_{-\infty}^{\infty} x f_X(x) dx$. Die Dichte f_X erhält man durch Ableiten der Verteilungsfunktion von X:

$$f_X(x) = F'_X(x) = \begin{cases} \alpha x_0^\alpha x^{-\alpha-1} & \text{für } x \geq x_0, \alpha > 0, \\ 0 & \text{sonst.} \end{cases}$$

(Bemerkung: $\alpha > 0$ muss gelten, weil sonst $f_X(x)$ keine Dichte ist.) Um $E[X]$ zu berechnen unterscheiden wir zwei Fälle. Für $\alpha \neq 1$ gilt

$$\begin{aligned} E[X] &= \int_{-\infty}^{\infty} x f_X(x) dx = \int_{x_0}^{\infty} x \alpha x_0^\alpha x^{-\alpha-1} dx \\ &= \alpha x_0^\alpha \int_{x_0}^{\infty} x^{-\alpha} dx = \alpha x_0^\alpha \left[\frac{1}{-\alpha+1} x^{-\alpha+1} \right]_{x_0}^{\infty} \\ &= \begin{cases} x_0 \frac{\alpha}{\alpha-1} < \infty & \text{falls } \alpha > 1, \\ \infty & \text{falls } \alpha < 1. \end{cases} \end{aligned}$$

Im Grenzfall $\alpha = 1$ gilt $E[X] = x_0 \int_{x_0}^{\infty} \frac{1}{x} dx = \infty$.

b) Da $N \sim \text{Poisson}(\lambda)$ gilt $E[N] = \lambda$. Für die Nettoprämie P_{net} gilt dann $P_{net} = E[X]E[N] = x_0 \frac{\alpha}{\alpha-1} \lambda = 2 \cdot 10^6 \cdot 2 \cdot 3 = 12 \cdot 10^6$, also 12 Millionen CHF.

c) Falls $\alpha \leq 1$, so ist $E[X] = +\infty$ (siehe **a)**). Eine Rückversicherung würde deshalb für keine noch so grosse Prämie Grossschäden versichern, die mit $\alpha \leq 1$ modelliert werden (müssen).

4. Bei den folgenden 6 Fragen ist jeweils genau eine Antwort richtig.

a) Sie erhalten eine grosse Lieferung von 10 verschiedenen Materialien. Aus Erfahrung wissen wir, dass im Schnitt 5% der Materialien mangelhaft sind. Nehmen Sie an, dass die Materialien unabhängig voneinander sind. Dann gilt:

1. Die Wahrscheinlichkeit, dass genau zwei Materialien mangelhaft sind, ist
$$\binom{10}{2} 0.05^8 (1 - 0.05)^2.$$
2. Die Wahrscheinlichkeit, dass genau zwei Materialien mangelhaft sind, ist
$$\binom{10}{2} 0.05^2 (1 - 0.05)^8.$$
3. Die Wahrscheinlichkeit, dass höchstens zwei Materialien mangelhaft sind, ist
$$\binom{10}{2} 0.05^2 (1 - 0.05)^8.$$
4. Wenn wir 20 Materialien anschauen, gibt es darunter sicher eine mangelhafte.

b) Sei X die Anzahl der Einsen, die Sie in 10 unabhängigen Würfen mit einem fairen Würfel werfen. Welche Verteilung kommt für X in Frage?

1. Bernoulli verteilt mit Parameter $p = 1/6$.
2. Geometrisch verteilt mit Parameter $p = 1/6$.
3. Binomialverteilt mit Parameter $n = 10$ und $p = 1/6$.
4. Poisson verteilt mit Parameter $\lambda = 5/3$.

c) Sei $X \sim \text{Poisson}(\lambda)$ mit $\lambda > 0$. Dann gilt:

1. $P(X > 5) = 1 - P(X < 5)$.
2. $P(X \geq 1 | X \leq 1) = \lambda/(\lambda + 1)$.
3. $2X \sim \text{Poisson}(2\lambda)$.

d) Betrachte eine normalverteilte Zufallsvariable $X \sim N(\mu, \sigma^2)$ mit $\mu = 1$ und $\sigma^2 = 3$. Dann gilt:

1. $P(X \leq 0) < P(X \geq 3)$.
2. Die Fläche unter der Dichte im Intervall $[1, 1 + \sqrt{3}]$ ist etwa 66%.
3. Die Fläche unter der Dichte im Intervall $[1 - 3, 1 + 3]$ ist etwa 66%.
4. Die Fläche unter der Dichte im Intervall $[1 - 2\sqrt{3}, 1 + 2\sqrt{3}]$ ist etwa 95%.

e) Die stetige Zufallsvariable X hat die Verteilungsfunktion

$$F_X(x) = \frac{1}{1 + \exp(-x)}, \quad x \in \mathbb{R}.$$

Dann gilt:

1. Es ist $P(X = 0) = 0.5$.
2. Für $x \geq 0$ gilt $P(X > x) = \frac{\exp(-x)}{1+\exp(-x)}$.
3. Die Dichte von X ist $\frac{-\exp(-x)}{(1+\exp(-x))^2}$, $x \in \mathbb{R}$.

Lösung:

a) 2. Die zweite Binomialformel ist richtig. Die letzte Antwort ist falsch:

$$P(\text{keine Materialie ist mangelhaft}) = \binom{20}{0} 0.05^0 0.95^{20} = 0.95^{20} \approx 0.35.$$

b) 3.

c) 2. Die erste Antwort ist falsch: $P(X > 5) = 1 - P(X \leq 5)$. Weil X diskret ist, ist $1 - P(X \leq 5) = 1 - P(X < 5) - P(X = 5) \neq 1 - P(X < 5)$. Die zweite Antwort ist richtig:

$$\begin{aligned} P(X \geq 1 | X \leq 1) &= P(X \geq 1 \text{ und } X \leq 1)/P(X \leq 1) = P(X = 1)/P(X \leq 1) \\ &= \frac{\lambda \exp(-\lambda)}{\exp(-\lambda) + \lambda \exp(-\lambda)} = \frac{\lambda}{\lambda + 1}. \end{aligned}$$

Die letzte Antwort ist falsch: Der Wertebereich von $2X$ ist $\{0, 2, 4, 6, \ldots\}$. Das zeigt, dass $2X$ nicht Poisson verteilt ist.

d) 4. ist korrekt. Da für $X \sim \mathcal{N}(1, 3)$ gilt, dass $(X - 1)/\sqrt{3} \sim \mathcal{N}(0, 1)$, bekommen wir

$$\begin{aligned} P[1 - 2\sqrt{3} \leq X \leq 1 + 2\sqrt{3}] &= P[-2 < (X - 1)/\sqrt{3} < 2] \\ &= \Phi(2) - \Phi(-2) = 2\Phi(2) - 1 \approx 0.9544. \end{aligned}$$

e) 2. Die erste Antwort ist falsch: $F_X(0) = 0.5 = P(X \leq 0)$ und $P(X = 0) = 0$, weil X stetig ist. Die zweite Antwort ist richtig: Für $x \geq 0$ gilt $P(X > x) = 1 - P(X \leq x) = 1 - \frac{1}{1+\exp(-x)} = \frac{\exp(-x)}{1+\exp(-x)}$. Die dritte Antwort ist falsch: Die Dichte von X ist nicht negativ.