Musterlösung

Wahrscheinlichkeitstheorie und Statistik (BSc D-ITET)

1. Die Aufgabe handelt von dreimaligem Ziehen aus der Menge $\{1,2,3,4,5,6\}$ ohne Zurücklegen. Wir wählen als Grundraum

$$\Omega \;=\; \{\{x_1,x_2,x_3\} \mid x_1,x_2,x_3 \in \{1,2,3,4,5,6\},\ x_i \neq x_j \text{ für } i \neq j\},$$

mit $|\Omega| = \binom{6}{3} = 20$.

a) Die günstigen Fälle werden beschrieben durch $A = \{\{6,x_1,x_2\} \mid x_1,x_2 \in \{1,2,3,4,5\},\ x_1 \neq x_2\}$. Es gilt $|A| = \binom{5}{2} = 10$. Es folgt

$$P[6 \text{ wurde gewählt}] \;=\; \frac{|A|}{|\Omega|} \;=\; \frac{1}{2}.$$

Sei G_i das Ereignis, genau i gerade Zahlen zu ziehen, $i = 1,...,3$. Es gilt

$$\begin{aligned}
G_1 &= \{\{x_1,x_2,x_3\} \mid x_1 \in \{2,4,6\},\ x_2,x_3 \in \{1,3,5\},\ x_2 \neq x_3\};\\
G_2 &= \{\{x_1,x_2,x_3\} \mid x_1 \in \{1,3,5\},\ x_2,x_3 \in \{2,4,6\},\ x_2 \neq x_3\};\\
G_3 &= \{\{2,4,6\}\}.
\end{aligned}$$

Für die Kardinalitäten gilt $|G_1| = |G_2| = 3 \times \binom{3}{2} = 9$ sowie $|G_3| = 1$. Für den Erwartungswert von X folgt somit

$$\begin{aligned}
E[X] &= \sum_{i=1}^{3} iP[X=i] \;=\; P[G_1] + 2P[G_2] + 3P[G_3]\\
&= \frac{9}{20} + 2\frac{9}{20} + 3\frac{1}{20} \;=\; \frac{30}{20} \;=\; \frac{3}{2}.
\end{aligned}$$

b) Nach Definition der bedingten Wahrscheinlichkeit gilt

$$P[X=2 \mid 6 \text{ wurde gewählt}] \;=\; \frac{P[\{6 \text{ wurde gewählt}\} \cap \{X=2\}]}{P[6 \text{ wurde gewählt}]} \;=\; \frac{|A \cap G_2|}{|A|}.$$

Wegen $A \cap G_2 = \{\{6,x_1,x_2\} \mid x_1 \in \{2,4\},\ x_2 \in \{1,3,5\}\}$ mit $|A \cap G_2| = 2 \times 3 = 6$ folgt

$$P[X=2 \mid 6 \text{ wurde gewählt}] \;=\; \frac{6}{10} \;=\; \frac{3}{5}.$$

c) Sei Y_1 die erste gezogene Zahl und Y_2 die zweite gezogene Zahl. Nach dem Satz der totalen Wahrscheinlichkeit gilt

$$\begin{aligned}
P[Y_1 + Y_2 \text{ gerade}] &= P[Y_2 \text{ ungerade} \mid Y_1 \text{ ungerade}]\frac{2}{3}\\
&\quad + P[Y_2 \text{ gerade} \mid Y_1 \text{ gerade}]\frac{1}{3}\\
&= \frac{2}{3}\frac{2}{3} + \frac{1}{2}\frac{1}{3} \;=\; \frac{4}{9} + \frac{1}{6} \;=\; \frac{16+6}{36} \;=\; \frac{11}{18}.
\end{aligned}$$

2. **a)** Wegen der Unabhängigkeit von H und R gilt

$$E[V] \;=\; E[\pi H R^2] \;=\; \pi E[HR^2] \;=\; \pi E[H]E[R^2] \;=\; \frac{\pi}{2}E[R^2] \;=\; \pi,$$

wobei wir verwendet haben, dass $E[H] = \int_0^1 x dx = 1/2$ gilt. Analog gilt für die den Erwartungswert von S,

$$\begin{aligned} E[S] &= E[2\pi(HR + R^2)] \;=\; 2\pi E[HR] + 2\pi E[R^2] \;=\; 2\pi E[H]E[R] + 2\pi E[R^2] \\ &= \pi E[R] + 4\pi \;=\; 5\pi. \end{aligned}$$

Es gilt daher auch

$$\begin{aligned} \operatorname{cov}(S,V) &= E[SV] - E[S]E[V] \;=\; 2\pi^2 E\left[H^2R^3 + HR^4\right] - 5\pi^2 \\ &= 2\pi^2 E\left[H^2\right] E\left[R^3\right] + 2\pi^2 E\left[H\right] E\left[R^4\right] - 5\pi^2 \\ &= 12\pi^2 E\left[H^2\right] + 24\pi^2 - 5\pi^2 \\ &= 4\pi^2 + 19\pi^2 \;=\; 23\pi^2. \end{aligned}$$

b) Die beste lineare Prognose von V durch R ist gegeben durch

$$Z \;=\; \frac{\operatorname{cov}(V,R)}{\operatorname{Var}(R)}(R - E[R]) + E[V] \;=\; \operatorname{cov}(V,R)(R-1) + \pi,$$

da $E[R] = 1$, $\operatorname{Var}(R) = E[R^2] - 1 = 1$ und $E[V] = \pi$. Es gilt

$$\begin{aligned} \operatorname{cov}(V,R) &= E[VR] - E[V]E[R] \;=\; \pi E\left[HR^3\right] - \pi \\ &= \pi E\left[H\right] E\left[R^3\right] - \pi \\ &= 3\pi - \pi \;=\; 2\pi. \end{aligned}$$

Somit folgt

$$Z \;=\; \operatorname{cov}(V,R)(R-1) + \pi \;=\; 2\pi(R-1) + \pi \;=\; 2\pi R - \pi.$$

c) Wir bemerken zuerst, dass die gemeinsame Dichte von H und R wegen der Unabhängigkeit gegeben ist durch

$$f_{H,R}(h,r) = f_H(h) f_R(r),$$

für $h, r \in \mathbb{R}$, wobei f_H und f_R die Dichten von H bzw. R beschreiben.

Um die Dichte von T zu bestimmen, bestimmen wir zuerst die Verteilungsfunktion $F_T(x)$ von T. Es ist $F(x) = 0$ für $x \leq 0$. Für $x > 0$ gilt

$$\begin{aligned} F_T(x) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} 1_{\{h \leq xr\}} f_{H,R}(h,r) dh dr \\ &= \int_{-\infty}^{\infty} f_H(h) \left(\int_0^{\infty} 1_{\{h/x \leq r\}} e^{-r} dr\right) dh \\ &= \int_0^1 \left(-e^{-r}\big|_{h/x}^{\infty}\right) dh \\ &= \int_0^1 e^{-h/x} dh \;=\; -xe^{-h/x}\Big|_0^1 \;=\; x\left(1 - e^{-1/x}\right). \end{aligned}$$

Für die Dichte von T erhalten wir schliesslich $f_T(x) = 0$ für $x \leq 0$ und

$$f_T(x) \;=\; \frac{d}{dx}F_T(x) \;=\; \left(1 - e^{-1/x}\right) + x\left(-e^{-1/x}/x^2\right) \;=\; 1 - e^{-1/x} - e^{-1/x}/x,$$

für $x > 0$.

3. **a)** Es muss gelten

$$1 \;=\; \int f_\vartheta(x)dx \;=\; c(\vartheta)\int_0^\infty \frac{1}{(1+x)^{\vartheta+1}}dx \;=\; \left.\frac{c(\vartheta)}{-\vartheta(1+x)^\vartheta}\right|_0^\infty \;=\; \frac{c(\vartheta)}{\vartheta}.$$

Somit gilt $c(\vartheta)=\vartheta$. Für den Erwartungswert von X gilt

$$\begin{aligned}
E[X] &= \int_0^\infty x\frac{\vartheta}{(1+x)^{\vartheta+1}}dx \\
&= \int_0^\infty (1+x)\frac{\vartheta}{(1+x)^{\vartheta+1}}dx - \int_0^\infty \frac{\vartheta}{(1+x)^{\vartheta+1}}dx \\
&= \int_0^\infty \frac{\vartheta}{(1+x)^{\vartheta}}dx - 1 \\
&= \left.\frac{\vartheta}{-(\vartheta-1)(1+x)^{\vartheta-1}}\right|_0^\infty - 1 \\
&= \frac{\vartheta}{\vartheta-1} - 1 \;=\; \frac{1}{\vartheta-1}.
\end{aligned}$$

b) Es gilt $\log(1+X)>0$, also $F(y)=0$ für $y\le 0$, wobei F die Verteilungsfunktion von $\log(1+X)$ ist. Für $y>0$ gilt

$$\begin{aligned}
F(y) &= P[\log(1+X)\le y] \;=\; P\left[X\le e^y-1\right] \;=\; \int_0^{e^y-1}\frac{\vartheta}{(1+x)^{\vartheta+1}}dx \\
&= \left.\frac{1}{-(1+x)^\vartheta}\right|_0^{e^y-1} \\
&= 1-\frac{1}{e^{\vartheta y}}.
\end{aligned}$$

$\log(1+X)$ hat also eine Exponentialverteilung mit Parameter ϑ.

c) Die Likelihood-Funktion ist für $x_1,\dots,x_n>0$,

$$L(\vartheta;x_1,\dots,x_n) \;=\; \prod_{i=1}^n f_\vartheta(x_i) \;=\; \vartheta^n\prod_{i=1}^n\frac{1}{(1+x_i)^{\vartheta+1}}.$$

Die log-Likelihood-Funktion ist also

$$\ell(\vartheta;x_1,\dots,x_n) \;=\; n\log(\vartheta)-(\vartheta+1)\sum_{i=1}^n\log(1+x_i).$$

Wir leiten diese Funktion nach ϑ ab und setzten den Ausdruck gleich 0.

$$0 \;=\; \frac{n}{\vartheta}-\sum_{i=1}^n\log(1+x_i).$$

Der Maximum-Likelihood Schätzer ist also gegeben durch

$$\hat\vartheta \;=\; \frac{1}{\frac{1}{n}\sum_{i=1}^n\log(1+X_i)},$$

wobei $X_1,\dots,X_n$ unabhängige Zufallsvariablen sind, welche je die gleiche Verteilung wie X haben.

Es gilt wegen der Unabhängigkeit, dass

$$\begin{aligned}
\mathrm{Var}(1/\hat\vartheta) &= \mathrm{Var}\left(\frac{1}{n}\sum_{i=1}^n\log(1+X_i)\right) \;=\; \frac{1}{n^2}\sum_{i=1}^n\mathrm{Var}\left(\log(1+X_i)\right) \\
&= \frac{1}{n^2}\sum_{i=1}^n\frac{1}{\vartheta^2} \;=\; \frac{1}{n\vartheta^2},
\end{aligned}$$

wobei wir verwendet haben, dass $\log(1+X)$ eine Exponentialverteilung mit Parameter ϑ hat.