Universität Salzburg		Florian Graf
	Machine Learning	
Übungsblatt **2**		25 Punkte

Aufgabe 1. *Grundlegende Definitionen* 3 P.

(a) Erklären Sie (informell) die Begriffe Zufallsvariable, Ereignis und Ereignisraum.

(b) Nennen Sie die Definition einer diskreten und stetigen Zufallsvariablen.

(c) Nennen Sie jeweils ein Beispiel einer diskreten und einer stetigen Zufallsvariable, sowie der zugehörigen Ereignisräumen.

Lösung 1.

(a) Eine Zufallsvariable modelliert eine zufallsbehaftete (Mess-)Größe. Die Werte, die die Zufallsvariable annehmen kann werden Ereignisse genannt und jedem Ereignis ist eine Wahrscheinlichkeit zugeordnet. Die Menge aller möglichen Ereignisse bildet den Ereignisraum.

(b) Eine diskrete Zufallsvariable ist eine Zufallsvariable mit diskretem Ereignisraum. Gebräuchlich ist auch die Definition als Zufallsvariable mit abzählbarem Ereignisraum. Eine stetige Zufallsvariable hat einen kontinuierlichen (=überabzählbaren) Ereignisraum.

(c) Diskrete Zufallsvariable; z.B. Münzwurf mit Ereignisraum {Kopf, Zahl}. Stetige Zufallsvariable; z.B. Zeit zwischen zwei Zerfallsereignissen einer radioaktiven Substanz mit Ereignisraum $\mathbb{R}^+$.

Aufgabe 2. *Diskrete Zufallsvariablen – Definitionen* 3 P.

Es sei X eine diskrete Zufallsvariable mit Ereignisraum $\mathcal{X}$.

(a) Es sei p die Wahrscheinlichkeitsfunktion von X. Nennen Sie die Definitionsmenge und die Zielmenge von p. Wie ist $p(x)$ definiert?

(b) Es sei A eine Teilmenge des Ereignisraums $\mathcal{X}$. Wie lässt sich $\Pr(X \in A)$, also die Wahrscheinlichkeit, dass X einen Wert in A annimmt, berechnen?

(c) Nennen Sie die Formel für den Erwartungswert von X und erklären Sie ihre Bedeutung.

(d) Nennen Sie die Formel für die Varianz von X und erklären Sie ihre Bedeutung.

(e) Es seien $a, b \in \mathbb{R}$ und $Y := aX + b$. Zeigen Sie, dass $\mathbb{E}[Y] = a\mathbb{E}[X] + b$. Berechnen Sie außerdem die Varianz $\mathbb{V}[Y]$ von Y.

Lösung 2.

(a)

$$\begin{aligned} p : \mathcal{X} &\to [0, 1] \\ x &\mapsto \Pr(X = x) \end{aligned}$$

(b) $\Pr(X \in A) = \sum_{x\in A} \Pr(X = x) = \sum_{x\in A} p(x)$. (Die Summe ist wohldefiniert, da $A \subset \mathcal{X}$ abzählbar ist.)

(c) $\mathbb{E}[X] = \sum_{x\in A} x \Pr(X = x) = \sum_{x\in A} xp(x)$. Der Erwartungswert ist ein gewichteter Mittelwert der Ereignisse von X, wobei jedes Ereignis gemäß seiner Wahrscheinlichkeit gewichtet wird.

(d) $\mathbb{V}[X] = \mathbb{E}[(X - \mathbb{E}[X])^2]$. Die Varianz quantifiziert die (quadratische) Streuung einer Zufallsvariable X um ihren Mittelwert.

(e) Der Ereignisraum von $Y = aX + b$ ist $\mathcal{Y} = \{ax + b | \, x \in \mathcal{X}\}$. Außerdem gilt für $x \in \mathcal{X}$, dass $\Pr(Y = ax + b) = \Pr(aX + b = ax + b) = \Pr(X = x)$, was $p_Y(ax + b) = p_X(x)$ impliziert. Somit ist

$$\mathbb{E}[Y] = \sum_{y \in Y} y p_Y(y) = \sum_{x \in \mathcal{X}} (ax + b) p_Y(ax + b) = \sum_{x \in \mathcal{X}} (ax + b) p_X(x) = a \underbrace{\sum_{x \in \mathcal{X}} x p_X(x)}_{=\mathbb{E}[X]} + b \underbrace{\sum_{x \in \mathcal{X}} p_X(x)}_{=\Pr(X \in \mathcal{X}) = 1} = a\mathbb{E}[X] + b \ .$$

Bezüglich der Varianz folgt sofort, dass

$$\mathbb{V}[Y] = \mathbb{E}[(aX + b - (a\mathbb{E}[X] + b))^2] = \mathbb{E}[a^2(X - \mathbb{E}[X])^2] = a^2\mathbb{E}[(X - \mathbb{E}[X])^2] = a^2\mathbb{V}[X] \ .$$

Aufgabe 3. *Diskrete Zufallsvariablen – Beispiele* 6 P.

(a) Es sei $X \sim \mathrm{Ber}(\theta)$, d.h., die Zufallsvariable X folgt einer Bernoulliverteilung mit Parameter $\theta \in (0, 1)$.

(i) Nennen Sie den Ereignisraum von X und die Wahrscheinlichkeitsfunktion von X.

(ii) Berechnen Sie den Erwartungswert und die Varianz von X.

(b) Es sei $X \sim \mathrm{Cat}(\boldsymbol{\theta})$, d.h., die Zufallsvariable X folgt einer Kategorischen Verteilung mit Parameter $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_C)$, wobei $\theta_i \in (0, 1)$ für alle $i = 1, \ldots, C$ und $\sum_{c=1}^{C} = 1$.

(i) Nennen Sie den Ereignisraum von X und die Wahrscheinlichkeitsfunktion von X in der 'one-hot' Schreibweise.

(ii) Berechnen Sie den Erwartungswert und die Varianz von X.

(c) Es sei $X \sim \mathrm{Bin}(N, \theta)$, d.h., die Zufallsvariable X folgt einer Binomialverteilung mit Parametern $N \in \mathbb{N}$ und $\theta \in (0, 1)$.

(i) Nennen Sie den Ereignisraum von X und die Wahrscheinlichkeitsfunktion von X.

(ii) Berechnen Sie den Erwartungswert und die Varianz von X.

Lösung 3.

(a) Es sei $X \sim \mathrm{Ber}(\theta)$, dann ist

$$\begin{aligned} p : \{0, 1\} &\longrightarrow [0, 1], \\ x &\longmapsto \left.\begin{cases} \theta & \text{falls } x = 1, \\ 1 - \theta & \text{falls } x = 0 \end{cases}\right\} = \theta^x \theta^{1-x} \end{aligned}$$

Somit ist

$$\mathbb{E}[X] = \sum_{x=0}^{1} x p(x) = 0 \cdot (1 - \theta) + 1 \cdot \theta = \theta$$

und

$$\mathbb{V}[X] = \sum_{x=0}^{1} (x - \mathbb{E}[X])^2 p(x) = (-\theta)^2 \cdot (1 - \theta) + (1 - \theta)^2 \cdot \theta = \theta(1 - \theta)(\theta + 1 - \theta) = \theta(1 - \theta) \ .$$

(b) Es sei $X \sim \mathrm{Cat}(\boldsymbol{\theta})$, dann ist

$$\begin{aligned} p : \{\mathbf{e}_1, \ldots, \mathbf{e}_C\} &\longrightarrow [0, 1], \\ x &\longmapsto \left.\begin{cases} \theta_1 & \text{falls } x = \mathbf{e}_1, \\ \ldots & \\ \theta_C & \text{falls } x = \mathbf{e}_C \end{cases}\right\} = \prod_{i=1}^{C} \theta_i^{1\{x = \mathbf{e}_i\}} \end{aligned}$$

wobei $\mathbf{e}_i$ der i-te standard Baisvektor von $\mathbb{R}^C$ ist.

Folglich gilt

$$\mathbb{E}[X] = \sum_{i=1}^{C} \mathbf{e}_i p(\mathbf{e}_i) = \sum_{i=1}^{C} \mathbf{e}_i \theta_i = \begin{pmatrix} \theta_1 \\ \vdots \\ \theta_C \end{pmatrix} = \boldsymbol{\theta}$$

und

$$\mathbb{V}[X] = \sum_{i=1}^{C} (\mathbf{e}_i - \mathbb{E}[X])^2 p(\mathbf{e}_i) = \sum_{i=1}^{C} (\mathbf{e}_i - \boldsymbol{\theta})^2 \cdot \theta_i \ ,$$

wobei das Quadrat Komponentenweise zu verstehen ist. Somit hat $\mathbb{V}[X]$ Komponenten

$$\begin{aligned}
\mathbb{V}[X]_k &= \sum_{i=1}^{C} (\delta_{ik} - \theta_k)^2 \cdot \theta_i \\
&= (1-\theta_k)^2 \theta_k + \theta_k^2 \sum_{i \neq k} \theta_i = (1-\theta_k)^2 \theta_k + \theta_k^2 (1-\theta_k) \\
&= (1-\theta_k)\theta_k (1 - \theta_k + \theta_k) = (1-\theta_k)\theta_k
\end{aligned}$$

und es gilt $\mathbb{V}[X] = \boldsymbol{\theta}(1-\boldsymbol{\theta})$, mit komponentenweisem Produkt. Alternativ ist auch eine Herleitung mithilfe des Verschiebungssatzes $\mathbb{V}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$ und $\mathbb{E}[X^2] = \mathbb{E}[X] = \boldsymbol{\theta}$ möglich.

(c) Es sei $X \sim \mathrm{Bin}(N, \theta)$, dann ist

$$\begin{aligned}
p : \{0, \dots, N\} &\to [0,1], \\
n &\mapsto \binom{N}{n} \theta^n (1-\theta)^{N-n} \ .
\end{aligned}$$

Zu berechnen ist

$$\mathbb{E}[X] = \sum_{n=0}^{N} n \binom{N}{n} \theta^n (1-\theta)^{N-n} = \sum_{n=1}^{N} n \binom{N}{n} \theta^n (1-\theta)^{N-n} \ .$$

Es gilt

$$n\binom{N}{n} = \frac{nN!}{n!(N-n)!} = \frac{N!}{(n-1)!(N-n)!} = \frac{N\,(N-1)!}{(n-1)!(N-n)!} = N\binom{N-1}{n-1} \ ,$$

woraus folgt dass

$$\begin{aligned}
\mathbb{E}[X] &= \sum_{n=1}^{N} n \binom{N}{n} \theta^n (1-\theta)^{N-n} \\
&= \sum_{n=1}^{N} N \binom{N-1}{n-1} \theta^n (1-\theta)^{N-n} \\
&= \sum_{n=0}^{N-1} N \binom{N-1}{n} \theta^{n+1} (1-\theta)^{N-1-n} \\
&= N\theta \underbrace{\sum_{n=0}^{N-1} \binom{N-1}{n} \theta^n (1-\theta)^{N-1-n}}_{=(\theta+(1-\theta))^{N-1}=1} \\
&= N\theta \ .
\end{aligned}$$

Ähnlich ist

$$n(n-1)\binom{N}{n} = N(N-1)\binom{N-2}{n-2}$$

und daher

$$\begin{aligned}
\mathbb{E}[X^2] - \mathbb{E}[X] &= \mathbb{E}[X(X-1)] \\
&= \sum_{n=2}^{N} n(n-1)\binom{N}{n}\theta^n(1-\theta)^{N-n} \\
&= \sum_{n=2}^{N} N(N-1)\binom{N-2}{n-2}\theta^n(1-\theta)^{N-n} \\
&= \sum_{n=2}^{N} N(N-1)\binom{N-2}{n-2}\theta^n(1-\theta)^{N-n} \\
&= \sum_{n=0}^{N-2} N(N-1)\binom{N-2}{n}\theta^{n+2}(1-\theta)^{N-2-n} \\
&= N(N-1)\theta^2 \quad .
\end{aligned}$$

Daraus folgt

$$\begin{aligned}
\mathbb{V}[X] &= \mathbb{E}[X^2] - \mathbb{E}[X]^2 = (\mathbb{E}[X^2] - \mathbb{E}[X]) - (\mathbb{E}[X]^2 - \mathbb{E}[X]) \\
&= N(N-1)\theta^2 - N^2\theta^2 + N\theta = -N\theta^2 + N\theta = N\theta(1-\theta) \quad .
\end{aligned}$$

Aufgabe 4. *Stetige Zufallsvariablen – Definitionen* 3 P.

Es sei X eine stetige Zufallsvariable mit Ereignisraum $\mathcal{X}$.

(a) Es sei P die kumulative Wahrscheinlichkeitsfunktion von X. Nennen Sie die Definitionsmenge und die Zielmenge von P. Wie ist $P(x)$ definiert?

(b) Wie ist die Wahrscheinlichkeitsdichte von X definiert?

(c) Es sei A eine Teilmenge des Ereignisraums $\mathcal{X}$. Wie lässt sich $\Pr(X \in A)$, also die Wahrscheinlichkeit, dass X einen Wert in A annimmt, berechnen?

(d) Wie lautet die Formel für den Erwartungswert von X.

(e) Wie lautet die Formel für die Varianz von X.

(f) Es seien $a, b \in \mathbb{R}$ und $Y := aX + b$. Zeigen Sie, dass $\mathbb{E}[Y] = a\mathbb{E}[X] + b$.

Lösung 4.

Es sei X eine stetige Zufallsvariable mit Ereignisraum $\mathcal{X}$.

(a)

$$\begin{aligned}
P : \mathbb{R} &\longrightarrow [0, 1] \\
x &\longmapsto \Pr(X \leq x)
\end{aligned}$$

(b) Die Wahrscheinlichkeitsdichte p von X ist definiert als

$$\begin{aligned}
p : \mathcal{X} &\rightarrow \mathbb{R}^+ \\
x &\mapsto \frac{d}{dx}P(x) = \frac{d}{dx}\Pr(X \leq x)
\end{aligned}$$

(c) Die Wahrscheinlichkeit $\Pr(X \in A)$ lässt sich berechnen durch $\Pr(X \in A) = \int_A p(x)dx = \int_A \frac{d}{dx}\Pr(X \leq x)dx$.

(d) Die Formel für den Erwartungswert von X lautet $\mathbb{E}[X] = \int_{\mathcal{X}} xp(x)$.

(e) Die Formel für die Varianz von X ist $\mathbb{V}[X] = \mathbb{E}[(X - \mathbb{E}[X])^2] = \int_{\mathcal{X}}(x - \mathbb{E}[X])^2 p(x)dx$.

(f) Nach Definition ist $\mathbb{E}[Y] = \int_{\mathcal{Y}} y p_Y(y)$. Um das Integral zu berechnen, müssen der Ereignisraum $\mathcal{Y}$ und die Dichte p_Y von $Y = aX + b$ in Abhängigkeit von $\mathcal{X}$ und X dargestellt werden.

Die Zufallsvariable $Y := aX + b$ hat den Ereignisraum $\mathcal{Y} = \{ax + b : x \in \mathcal{X}\}$. Außerdem ist $P_Y(y) = \Pr(Y \leq y) = \Pr(aX + b \leq y)$, sodass für alle $x \in \mathcal{X}$ die Gleichung $P_Y(ax + b) = P_X(x)$ gilt. Folglich ist

$$p_X(x) = \frac{d}{dx} P_X(x) = \frac{d}{dx} P_Y(ax + b) = \frac{dP_Y(ax + b)}{d(ax + b)} \frac{d(ax + b)}{dx} = p_Y(ax + b)a$$

und daher

$$\mathbb{E}[Y] = \int_{\mathcal{Y}} y p_Y(y) dy \overset{y=ax+b}{=} \int_{\mathcal{X}} (ax + b) \underbrace{p_Y(ax + b)a}_{=p_X(x)} dx = \int_{\mathcal{X}} (ax + b) p_X(x) dx$$

$$= a \underbrace{\int_{\mathcal{X}} x p_X(x) dx}_{=\mathbb{E}[X]} + b \underbrace{\int_{\mathcal{X}} x p_X(x) dx}_{=\Pr(X \in \mathcal{X})=1} = a\mathbb{E}[X] + b \ .$$

Aufgabe 5. *Stetige Zufallsvariablen – Beispiele* 6 P.

(a) Es sei $X \sim \mathcal{N}(\mu, \sigma^2)$, d.h., die Zufallsvariable X folgt einer (univariaten) Normalverteilung mit Parametern $\mu \in \mathbb{R}$ und $\sigma^2 > 0$.

 (i) Nennen Sie den Ereignisraum von X und die Wahrscheinlichkeitsdichte von X.

 (ii) Berechnen Sie den Erwartungswert und die Varianz von X.

(b) Es sei $X \sim \mathcal{U}(a, b)$, d.h., die Zufallsvariable X folgt einer stetigen Gleichverteilung auf dem Interval $[a, b] \subset \mathbb{R}$.

 (i) Nennen Sie den Ereignisraum von X und die Wahrscheinlichkeitsdichte von X.

 (ii) Berechnen Sie den Erwartungswert und die Varianz von X.

(c) Es sei $X \sim \text{Expon}(\lambda)$, d.h., die Zufallsvariable X folgt einer Exponentialverteilung mit Parameter $\lambda > 0$.

 (i) Nennen Sie den Ereignisraum von X und die Wahrscheinlichkeitsdichte von X.

 (ii) Berechnen Sie den Erwartungswert und die Varianz von X.

Lösung 5.

(a) Es sei $X \sim \mathcal{N}(\mu, \sigma^2)$, dann ist

$$\begin{aligned} p : \mathbb{R} &\to \mathbb{R}^+ \\ x &\mapsto \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \ . \end{aligned}$$

Also ist

$$\begin{aligned} \mathbb{E}[X] &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{\mathbb{R}} x e^{-\frac{(x-\mu)^2}{2\sigma^2}} \overset{y=x-\mu}{=} \frac{1}{\sqrt{2\pi\sigma^2}} \int_{\mathbb{R}} (y + x) e^{-\frac{y^2}{2\sigma^2}} dy \\ &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{\mathbb{R}} y e^{-\frac{y^2}{2\sigma^2}} + \frac{\mu}{\sqrt{2\pi\sigma^2}} \int_{\mathbb{R}} e^{-\frac{y^2}{2\sigma^2}} \ . \end{aligned}$$

Da die Funktion $f : y \mapsto y e^{-\frac{y^2}{2\sigma^2}}$ schiefsymmetrisch ist (es gilt $f(-y) = -f(y)$) verschwindet das erste Integral. Somit ist

$$\mathbb{E}[X] = \frac{\mu}{\sqrt{2\pi\sigma^2}} \int_{\mathbb{R}} e^{-\frac{y^2}{2\sigma^2}} = \mu \Pr(X \in \mathbb{R}) = \mu \ .$$

Für die Varianz gilt

$$\mathbb{V}[X] = \mathbb{E}[(X - \mu)^2] = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{\mathbb{R}} (x - \mu)^2 e^{-\frac{(x-\mu)^2}{2\sigma^2}} \overset{y=x-\mu}{=} \frac{1}{\sqrt{2\pi\sigma^2}} \int_{\mathbb{R}} y^2 e^{-\frac{y^2}{2\sigma^2}} \ .$$

Es ist

$$y^2 e^{-\frac{y^2}{2\sigma^2}} = y\frac{d}{dy}\left(-\sigma^2 e^{-\frac{y^2}{2\sigma^2}}\right) = \frac{d}{dy}\left(-y\sigma^2 e^{-\frac{y^2}{2\sigma^2}}\right) + \left(\frac{d}{dy}y\right)\sigma^2 e^{-\frac{y^2}{2\sigma^2}} \quad ,$$

sodass (mit partieller Integration) folgt dass

$$\mathbb{V}[X] = \frac{1}{\sqrt{2\pi\sigma^2}}\int_{\mathbb{R}} y^2 e^{-\frac{y^2}{2\sigma^2}} = \frac{1}{\sqrt{2\pi\sigma^2}}\underbrace{\left[-y\sigma^2 e^{-\frac{y^2}{2\sigma^2}}\right]_{-\infty}^{\infty}}_{=0} + \sigma^2 \underbrace{\frac{1}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^{\infty} e^{-\frac{y^2}{2\sigma^2}} dy}_{=\Pr(X-\mu\in\mathbb{R})=1} = \sigma^2 \quad .$$

(b) Es sei $X \sim \mathcal{U}(a,b)$, dann ist

$$\begin{aligned} p : [a,b] &\to \mathbb{R}^+ \\ x &\mapsto \frac{1}{b-a} \quad . \end{aligned}$$

Somit ist

$$\mathbb{E}[X] = \int_a^b \frac{x}{b-a}dx = \frac{1}{2(b-a)}(b^2-a^2) = \frac{b+a}{2}$$

und

$$\mathbb{V}[X] = \frac{1}{b-a}\int_a^b \left(x - \frac{a+b}{2}\right)^2 dx \overset{y=x-(a+b)/2}{=} \frac{1}{b-a}\int_{-\frac{b-a}{2}}^{\frac{b-a}{2}} y^2 dy = \frac{2}{3(b-a)}\left(\frac{b-a}{2}\right)^3 = \frac{1}{12}(b-a)^2 \quad .$$

(c) Es sei $X \sim \text{Expon}(\lambda)$, dann ist

$$\begin{aligned} p : \mathbb{R}^+ &\to \mathbb{R}^+ \\ x &\mapsto \lambda e^{-\lambda x} \quad . \end{aligned}$$

Somit ist

$$E[X] = \int_0^\infty \lambda x e^{-\lambda x} dx \overset{y=\lambda x}{=} \int_0^\infty y e^{-y}\frac{dy}{\lambda} \quad .$$

Nun ist $ye^{-y} = -y\frac{d}{dy}e^{-y}$, sodass (mit partieller Integration) folgt dass

$$\begin{aligned} E[X] &= -\frac{1}{\lambda}\int_0^\infty y\left(\frac{d}{dy}e^{-y}\right)dy = -\frac{1}{\lambda}\int_0^\infty \frac{d}{dy}\left(ye^{-y}\right)dy + \frac{1}{\lambda}\int_0^\infty \left(\frac{d}{dy}y\right)e^{-y}dy \\ &= -\frac{1}{\lambda}\left[ye^{-y}\right]_0^\infty + \frac{1}{\lambda}\int_0^\infty e^{-y}dy = \frac{1}{\lambda} \quad . \end{aligned}$$

Ähnlich ist

$$\begin{aligned} E[X^2] &\overset{y=\lambda x}{=} \frac{1}{\lambda^2}\int_0^\infty y^2 e^{-y}dy \\ &\overset{\text{part. int}}{=} -\frac{1}{\lambda^2}\left[y^2 e^{-y}\right]_0^\infty + \frac{1}{\lambda^2}\int_0^\infty 2ye^{-y}dy \\ &= \frac{2}{\lambda}\mathbb{E}[X] = \frac{2}{\lambda^2} \quad , \end{aligned}$$

sodass

$$\mathbb{V}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2} \quad .$$

Aufgabe 6. *Bedingte Verteilungen – Definitionen* 4 P.

Es seien X und Y zwei diskrete Zufallsvariablen mit Ereignisräumen $\mathcal{X}$ und $\mathcal{Y}$.

(a) Es sei $p_{X,Y}$ die gemeinsame Wahrscheinlichkeitsfunktion von X und Y. Nennen Sie die Definitionsmenge und die Zielmenge von $p_{X,Y}$. Wie ist $p_{X,Y}$ definiert?

(b) Wie ist die Randverteilung von X definiert?

(c) Wie ist die bedingte Verteilung von X gegeben $Y = y$ definiert?

(d) Wie ist die Unabhängigkeit zweier Zufallsvariablen X und Y definiert.

(e) Nennen Sie ein Beispiel zweier Zufallsvariablen die der gleichen Verteilung folgen und unabhängig sind. Nennen Sie auch ein Beispiel zweier Zufallsvariablen die der gleichen Verteilung folgen und nicht unabhängig sind.

(f) Wie ist die Kovarianz $\mathrm{Cov}(X, Y)$ zweier Zufallsvariablen X und Y definiert?

(g) Nun seien X und Y unabhängig. Berechnen Sie $E[XY]$ und $\mathrm{Cov}[X, Y]$.

Lösung 6.

(a) Es ist

$$\begin{aligned} p_{X,Y} : \mathcal{X} \times \mathcal{Y} &\to [0, 1] \\ (x, y) &\mapsto \Pr(X = x \text{ und } Y = y) \end{aligned}$$

(b) Die Randverteilung P_X von X ist definiert als $p_X(x) = \sum_{y \in \mathcal{Y}} p_{X,Y}(x, y) = \Pr(X = x)$

(c) Die bedingte Verteilung $p_{X|Y}$ von X gegeben $Y = y$ ist definiert als $p_{X|Y}(x|y) = \frac{p_{X,Y}(x,y)}{p_Y(y)} = \frac{\Pr(X=x \text{ und } Y=y)}{\Pr(Y=y)}$.

(d) Zwei diskrete Zufallsvariablen X und Y sind (paarweise) unabhängig, falls für alle $(x, y) \in \mathcal{X} \times \mathcal{Y}$ gilt dass $p_{X,Y}(x, y) = p_X(x)p_Y(y)$.

(e) **Gleichverteilt und unabhängig.** Werfen zweier Würfel; X und Y sind die Augenzahl von jeweils einem Würfel.

Gleichverteilt aber abhängig. Werfen eines Würfels; X ist die Augenzahl, $Y = 7 - X$.

(f) Die Kovarianz $\mathrm{Cov}(X, Y)$ zweier Zufallsvariablen X und Y ist definiert als $\mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])]$.

(g) Allgemein ist

$$\mathbb{E}[XY] = \sum_{(x,y) \in \mathcal{X} \times \mathcal{Y}} xy\, p_{X,Y}(x, y) = \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} xy\, p_{X,Y}(x, y) \;.$$

Aus der Definition der Unabhängigkeit folgt dann

$$\mathbb{E}[XY] = \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} xy\, p_{X,Y}(x, y) = \sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} xy\, p_X(x) p_Y(y) = \left(\sum_{x \in \mathcal{X}} x\, p_X(x) \right) \left(\sum_{y \in \mathcal{Y}} y\, p_Y(y) \right) = \mathbb{E}[X]\mathbb{E}[Y] \;.$$

Folglich ist

$$\mathrm{Cov}[X, Y] = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])] = \mathbb{E}[XY - X\mathbb{E}[Y] - Y\mathbb{E}[X] + \mathbb{E}[X]\mathbb{E}[Y]] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = 0 \;.$$