zu Aufgabe 1:

a) Für den Träger der diskreten Zufallsgröße X gilt gemäß (§ 0.20)

$$\mathcal{S}_X = \{x \in \mathcal{X} \mid p_X(x) > 0\} = \{1, 2, 3\}.$$

b) Für den Erwartungswert der diskreten Zufallsgröße X gilt gemäß (§ 0.26)

$$\mathbb{E}\,(X) = \sum_{x \in \mathcal{X}} x\, p_X(x) = 1 \cdot \tfrac{1}{3} + 2 \cdot \tfrac{1}{2} + 3 \cdot \tfrac{1}{6} + 4 \cdot 0 = \tfrac{11}{6}.$$

Bemerkung: Da für alle $x \in \mathcal{X} \setminus \mathcal{S}_X$ der Zusammenhang $p_X(x) = 0$ gilt, können wir den Erwartungswert der diskreten Zufallsgröße X sowohl mit Hilfe des Trägers $\mathcal{S}_X$ als auch mit Hilfe des Alphabets $\mathcal{X}$ berechnen.

---

c) *Vorbemerkung:* Durch Transformation der diskreten Zufallsgröße X mit der Funktion g (die auf dem kompletten Alphabet $\mathcal{X}$ definiert ist und damit natürlich auch überall auf dem Träger $\mathcal{S}_X$) entsteht eine neue diskrete Zufallsgröße: die Zufallsgröße $Y = g(X)$. Bei der Transformation einer diskreten Zufallsgröße ändern sich i. A. der Träger / das Alphabet und auch die Wahrscheinlichkeitsfunktion der Zufallsgröße. Für den Träger $\mathcal{S}_Y$ der Zufallsgröße $Y = g(X)$ gilt

$$\mathcal{S}_Y = \{g(x) \in \mathbb{R} \mid x \in \mathcal{S}_X\}.$$

Für das Alphabet $\mathcal{Y}$ der Zufallsgröße Y vereinbaren wir[1]

$$\mathcal{Y} = \{g(x) \in \mathbb{R} \mid x \in \mathcal{X}\}.$$

Werden durch die Transformation mit der Funktion g verschiedene Elemente des Trägers $\mathcal{S}_X$ / des Alphabets $\mathcal{X}$ der Zufallsgröße X auf den gleichen Wert abgebildet, addieren sich die zugehörigen Wahrscheinlichkeiten, d. h. für die Wahrscheinlichkeitsfunktion p_Y der Zufallsgröße Y gilt

$$\forall\, y \in \mathcal{Y}\colon\; p_Y(y) = \sum_{x \in \mathcal{X}\colon\, g(x) = y} p_X(x).$$

Für den Erwartungswert der diskreten Zufallsgröße Y gilt gemäß (§ 0.26)

$$\mathbb{E}\,(Y) = \sum_{y \in \mathcal{Y}} y\, p_Y(y), \qquad \text{(Variante 1)}$$

aber auch

$$\mathbb{E}\,(Y) = \mathbb{E}\,(g(X)) = \sum_{x \in \mathcal{X}} g(x)\, p_X(x). \qquad \text{(Variante 2)}$$

Für Variante 1 müssen wir das Alphabet $\mathcal{Y}$ und die Wahrscheinlichkeitsfunktion p_Y der transformierten Zufallsgröße Y explizit berechnen. Wir erhalten[2]

$$\mathcal{Y} = \{1^2, 2^2, 3^2, 4^2\} = \{1, 4, 9, 16\}$$

und

y	1	4	9	16
$p_Y(y)$	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{6}$	0

[1] *Hinweis:* Genau genommen ist nur der Träger $\mathcal{S}_Y$ der Zufallsgröße Y eindeutig bestimmt, da sich dieser aus den transformierten Werten des Trägers $\mathcal{S}_X$ der Zufallsgröße X ergibt. Als Alphabet könnten wir eine beliebige Menge $\mathcal{Y}$ bezeichnen, die den Träger $\mathcal{S}_Y$ vollständig enthält.

[2] Für den Träger $\mathcal{S}_Y$ der Zufallsgröße Y würden wir $\mathcal{S}_Y = \{1^2, 2^2, 3^2\} = \{1, 4, 9\}$ erhalten.

und schließlich

$$\begin{aligned}\mathbb{E}(Y) &= \sum_{y\in\mathcal{Y}} y\, p_Y(y)\\ &= 1\cdot\tfrac{1}{3} + 4\cdot\tfrac{1}{2} + 9\cdot\tfrac{1}{6} + 16\cdot 0 = \tfrac{23}{6}.\end{aligned}$$

Für Variante 2 berechnen wir direkt

$$\begin{aligned}\mathbb{E}(Y) = \mathbb{E}(g(X)) &= \sum_{x\in\mathcal{X}} g(x)\, p_X(x)\\ = \mathbb{E}\left(X^2\right) \quad &= \sum_{x\in\mathcal{X}} x^2\, p_X(x)\\ &= 1^2\cdot\tfrac{1}{3} + 2^2\cdot\tfrac{1}{2} + 3^2\cdot\tfrac{1}{6} + 4^2\cdot 0 = \tfrac{23}{6}.\end{aligned}$$

Für die Varianz der diskreten Zufallsgröße X gilt gemäß (§ 0.29)

$$\operatorname{var}(X) = \mathbb{E}\left((X - \mathbb{E}(X))^2\right).$$

Alternativ können wir auch einfach den Verschiebungssatz (§ 0.31) 3) anwenden:

$$\begin{aligned}\operatorname{var}(X) &= \mathbb{E}\left(X^2\right) - (\mathbb{E}(X))^2\\ &= \tfrac{23}{6} - \left(\tfrac{11}{6}\right)^2 = \tfrac{23}{6} - \tfrac{121}{36} = \tfrac{17}{36}.\end{aligned}$$

d) Die Erläuterungen aus Teilaufgabe c) gelten analog auch für die transformierte Zufallsgröße Z, die durch Transformation mit der Funktion $\tilde{g}$ aus der diskreten Zufallsgröße X entsteht. Für den Erwartungswert der diskreten Zufallsgröße Z gilt gemäß (§ 0.26)

$$\begin{aligned}\mathbb{E}(Z) &= \mathbb{E}(\tilde{g}(X)) = \sum_{x\in\mathcal{X}} \tilde{g}(x)\, p_X(x)\\ &= \mathbb{E}\left(-\log_2\left(p_X(X)\right)\right) = -\sum_{x\in\mathcal{X}} \log_2\left(p_X(x)\right) p_X(x)\\ &= -\left(\tfrac{1}{3}\log_2\left(\tfrac{1}{3}\right) + \tfrac{1}{2}\log_2\left(\tfrac{1}{2}\right) + \tfrac{1}{6}\log_2\left(\tfrac{1}{6}\right) + \underbrace{0\,\log_2 0}_{=0 \quad \text{siehe (§ 1.2)}}\right)\\ &\qquad\qquad \text{(Stichwort: Regel von Bernoulli-L'Hôpital)}\\ &= \tfrac{1}{2}\log_2(3) + \tfrac{2}{3} = \tfrac{1}{6}\log_2(432) \approx 1.4591.\end{aligned}$$

(Darstellung des Ergebnisses ist abhängig von der Anwendung der Logarithmengesetze)

Hinweis: Auch eine Wahrscheinlichkeitsfunktion ist eine „normale“ Funktion und kann zur Transformation von Zufallsgrößen verwendet werden oder wie hier Bestandteil der Transformationsfunktion sein.

Bemerkung: Den Erwartungswert $\mathbb{E}\left(-\log_2\left(p_X(X)\right)\right)$ werden wir zukünftig als Entropie $H(X)$ der diskreten Zufallsgröße X bezeichnen. Bei Verwendung des Logarithmus zur Basis 2 ergibt sich die Pseudo-Einheit „bit“.

zu Aufgabe 2:

a) Im Fall (1) erhält man die Wahrscheinlichkeitsfunktionen p_X und p_Y der Zufallsgrößen X bzw. Y als Randwahrscheinlichkeitsfunktionen aus der gemeinsamen Wahrscheinlichkeitsfunktion $p_{X,Y}$ gemäß (§ 0.45), d. h. es gilt

$$\forall\, x \in \mathcal{X}\colon \quad p_X(x) = \sum_{y \in \mathcal{Y}} p_{X,Y}(x,y) \qquad \text{und} \qquad \forall\, y \in \mathcal{Y}\colon \quad p_Y(y) = \sum_{x \in \mathcal{X}} p_{X,Y}(x,y).$$

Im Fall (2) kann die Berechnung der Wahrscheinlichkeitsfunktionen $\tilde{p}_X$ und $\tilde{p}_Y$ analog zu Fall (1) erfolgen.

Fall (1):

$p_{X,Y}(x,y)$		$y=1$	$y=2$	$y=3$	$p_X(x)$
x	1	$\frac{1}{2}$	$\frac{1}{8}$	0	$\frac{5}{8}$
	2	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{4}$	$\frac{3}{8}$
$p_Y(y)$		$\frac{9}{16}$	$\frac{3}{16}$	$\frac{4}{16}$	

Fall (2):

$\tilde{p}_{X,Y}(x,y)$		$y=1$	$y=2$	$y=3$	$\tilde{p}_X(x)$
x	1	$\frac{1}{9}$	$\frac{1}{6}$	$\frac{1}{18}$	$\frac{1}{3}$
	2	$\frac{2}{9}$	$\frac{1}{3}$	$\frac{1}{9}$	$\frac{2}{3}$
$\tilde{p}_Y(y)$		$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{6}$	

b) Wir berechnen für alle Paare $(x,y) \in \mathcal{X} \times \mathcal{Y}$ das Produkt der Randwahrscheinlichkeiten.

Fall (1):

$p_X(x)\, p_Y(y)$		$y=1$	$y=2$	$y=3$	$p_X(x)$
x	1	$\frac{45}{128}$	$\frac{15}{128}$	$\frac{20}{128}$	$\frac{5}{8}$
	2	$\frac{27}{128}$	$\frac{9}{128}$	$\frac{12}{128}$	$\frac{3}{8}$
$p_Y(y)$		$\frac{9}{16}$	$\frac{3}{16}$	$\frac{4}{16}$	

Fall (2):

$\tilde{p}_X(x)\, \tilde{p}_Y(y)$		$y=1$	$y=2$	$y=3$	$\tilde{p}_X(x)$
x	1	$\frac{1}{9}$	$\frac{1}{6}$	$\frac{1}{18}$	$\frac{1}{3}$
	2	$\frac{2}{9}$	$\frac{1}{3}$	$\frac{1}{9}$	$\frac{2}{3}$
$\tilde{p}_Y(y)$		$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{6}$	

Im Fall (1) gilt $p_{X,Y} \neq p_X\, p_Y$, d. h. gemäß (§ 0.56) sind die Zufallsgrößen X und Y nicht (stochastisch) unabhängig.

Im Fall (2) gilt $\tilde{p}_{X,Y} = \tilde{p}_X\, \tilde{p}_Y$, d. h. gemäß (§ 0.56) sind die Zufallsgrößen X und Y (stochastisch) unabhängig.

Hinweis zur Schreibweise:
Bei den obigen Aussagen haben wir die Schreibweise mit Funktionen (p_X usw.) gewählt. Wir können alternativ auch eine Schreibweise mit Funktionswerten ($p_X(x)$ usw.) verwenden, müssen dann aber ergänzen, welche Funktionswerte (alle oder nur ausgewählte) wir meinen:

$$\begin{aligned} & p_{X,Y} \neq p_X\, p_Y \\ \Leftrightarrow\quad & \exists (x,y) \in \mathcal{X} \times \mathcal{Y}\colon \quad p_{X,Y}(x,y) \neq p_X(x)\, p_Y(y), \end{aligned}$$

$$\begin{aligned} & \tilde{p}_{X,Y} = \tilde{p}_X\, \tilde{p}_Y \\ \Leftrightarrow\quad & \forall (x,y) \in \mathcal{X} \times \mathcal{Y}\colon \quad \tilde{p}_{X,Y}(x,y) = \tilde{p}_X(x)\, \tilde{p}_Y(y). \end{aligned}$$

zu Aufgabe 3:

Für den Erwartungswert der Zufallsgröße $Z := a \sum\limits_{k=1}^{n} X_k$ mit $a \in \mathbb{R}$ gilt

$$
\begin{aligned}
\mathbb{E}\left(Z\right) &= \mathbb{E}\left(a \sum_{k=1}^{n} X_k\right) && \\
&= a\,\mathbb{E}\left(\sum_{k=1}^{n} X_k\right) && (\S\,0.31)\ 1)\ \text{Multiplikation mit Konstante} \\
&= a \sum_{k=1}^{n} \mathbb{E}\left(X_k\right) && (\S\,0.54)\ 1)\ \text{Linearität des Erwartungswerts} \\
&= a\,n\,\mathbb{E}\left(X_1\right) && \begin{array}{l}\text{identische Verteilung der Zufallsgrößen } X_1, X_2, \ldots, X_n \\ \Rightarrow \mathbb{E}\left(X_1\right) = \mathbb{E}\left(X_2\right) = \ldots = \mathbb{E}\left(X_n\right)\end{array} \\
&= a\,n\,\mu.
\end{aligned}
$$

Für die Varianz der Zufallsgröße $Z := a \sum\limits_{k=1}^{n} X_k$ mit $a \in \mathbb{R}$ gilt

$$
\begin{aligned}
\operatorname{var}\left(Z\right) &= \operatorname{var}\left(a \sum_{k=1}^{n} X_k\right) && \\
&= a^2 \operatorname{var}\left(\sum_{k=1}^{n} X_k\right) && (\S\,0.31)\ 4)\ \text{Multiplikation mit Konstante} \\
&= a^2 \sum_{k=1}^{n} \operatorname{var}\left(X_k\right) && \begin{array}{l}(\S\,0.51)\ 5)\ \text{ohne Kovarianzterme, denn es gilt:} \\ \text{Unabhängigkeit der Zufallsgrößen } X_1, X_2, \ldots, X_n \\ \Rightarrow \text{(paarweise) Unkorreliertheit der Zufallsgrößen} \\ \Rightarrow \text{Kovarianzterme sind alle Null}\end{array} \\
&= a^2\,n \operatorname{var}\left(X_1\right) && \begin{array}{l}\text{identische Verteilung der Zufallsgrößen } X_1, X_2, \ldots, X_n \\ \Rightarrow \operatorname{var}\left(X_1\right) = \operatorname{var}\left(X_2\right) = \ldots = \operatorname{var}\left(X_n\right)\end{array} \\
&= a^2\,n\,\sigma^2.
\end{aligned}
$$

Hinweis: Aus der Unabhängigkeit der Zufallsgrößen $X_1, X_2, \ldots, X_n$ folgt deren paarweise Unabhängigkeit und daraus folgt dann deren (paarweise) Unkorreliertheit, siehe (§ 0.57), aber die Umkehrung dieser Aussagen gilt i. A. nicht!

zu Aufgabe 4:

Vorbemerkung: Die Definition einer *konvexen Menge* und erläuternde Bemerkungen dazu finden Sie im Anhang des Skripts in (A.8) und (A.9).

---

a) (1) Die dargestellte Menge ist *konvex*, da für *jede beliebige* Auswahl von zwei Punkten aus dieser Menge gilt, dass auch die *komplette* Verbindungsstrecke zwischen diesen beiden Punkten in der Menge liegt.

(2) Die dargestellte Menge ist *nicht konvex*, da es Punktpaare in dieser Menge gibt, für die gilt, dass *nicht* die *komplette* Verbindungsstrecke zwischen diesen beiden Punkten in der Menge liegt.

(3) *konvex*, Begründung wie in (1)

(4) *nicht konvex*, Begründung wie in (2)

(5) *nicht konvex*, Begründung wie in (2)

(6) *konvex*, Begründung wie in (1)

(7) *nicht konvex*, Begründung wie in (2)

(8) *nicht konvex*, Begründung wie in (2)

---

b) *Vorbemerkung:* Gemäß (A.8) gilt: Die Menge $\mathcal{A}$ heißt konvex, falls für alle Vektoren $x, y \in \mathcal{A}$ und jedes $\lambda \in [0,1]$

$$\lambda \cdot x + (1-\lambda) \cdot y \in \mathcal{A}$$

gilt, d. h. auch jede mit Hilfe von λ gebildete Linearkombination aus x und y (= „Verbindungsstrecke zwischen x und y") liegt vollständig in der Menge $\mathcal{A}$.

Wir betrachten zwei beliebige Vektoren $x, y \in \mathcal{A}$. Das sind Vektoren der Länge n, die die Eigenschaften der Menge $\mathcal{A}$ erfüllen, d. h. es gilt

$$x = (x_1, x_2, \ldots, x_n) \text{ mit } x_k \in \mathbb{R} \text{ und } x_k \leq a_k \text{ für } k = 1, 2, \ldots n \quad \text{sowie}$$
$$y = (y_1, y_2, \ldots, y_n) \text{ mit } y_k \in \mathbb{R} \text{ und } y_k \leq a_k \text{ für } k = 1, 2, \ldots n.$$

Jeder Vektor z, der sich mit Hilfe von $\lambda \in [0,1]$ als Linearkombination von x und y bilden lässt, soll auch wieder der Menge $\mathcal{A}$ angehören und alle damit verbundenen Eigenschaften besitzen, d. h. es soll

$$\forall\, x, y \in \mathcal{A},\ \lambda \in [0,1]: \quad z := \lambda \cdot x + (1-\lambda) \cdot y \in \mathcal{A}$$

gelten. Zu zeigen ist also

$$z = (z_1, z_2, \ldots, z_n) \text{ mit } z_k \in \mathbb{R} \text{ und } z_k \leq a_k \text{ für } k = 1, 2, \ldots n.$$

- Offensichtlich ist auch z ein Vektor der Länge n mit reellwertigen Komponenten, da die einzelnen Vektorkomponenten für $k = 1, 2, \ldots n$ durch

$$z_k = \lambda\, x_k + (1-\lambda)\, y_k$$

gegeben sind, wobei $z_k \in \mathbb{R}$ direkt aus $x_k, y_k \in \mathbb{R}$ und $\lambda \in [0,1]$ folgt.

- Außerdem gilt

$$z_k = \lambda \underbrace{x_k}_{\leq a_k} + (1-\lambda) \underbrace{y_k}_{\leq a_k} \leq \lambda\, a_k + (1-\lambda)\, a_k = a_k.$$

Wir können also für alle $x, y \in \mathcal{A}$ und jedes $\lambda \in [0,1]$ schlussfolgern, dass $z = \lambda \cdot x + (1-\lambda) \cdot y$ *alle* Eigenschaften der Menge $\mathcal{A}$ erfüllt, d. h. es gilt $z \in \mathcal{A}$.

Damit ist die Menge $\mathcal{A}$ konvex.

zu Aufgabe 5:

Vorbemerkung: Die Definition einer *konvexen Funktion* und erläuternde Bemerkungen dazu finden Sie im Anhang des Skripts in (A.8) und (A.9).

Die Untersuchung der Konvexität / Konkavität der Funktionen g_1, g_2 und g_3 kann auf zwei verschiedene Arten erfolgen.

Variante 1 (Überprüfung gemäß Definition) – am Beispiel der Funktion g_1:

Gemäß (A.8) gilt: Die Funktion g_1 ist auf dem (noch nicht näher bestimmten) Intervall (a, b) konvex, falls für alle $x, y \in (a, b)$ und jedes $\lambda \in [0, 1]$

$$g_1(\lambda\, x + (1-\lambda)\, y) \leq \lambda\, g_1(x) + (1-\lambda)\, g_1(y)$$

gilt, d. h. für zwei beliebige Punkte auf dem Funktionsgraphen (im betrachteten Intervall) gilt, dass die Funktion zwischen diesen beiden Punkten nicht oberhalb der direkten „Verbindungsstrecke“ zwischen den beiden Punkten verläuft. Gilt das umgekehrte Relationszeichen, ist die Funktion g_1 auf dem Intervall (a, b) konkav.

Wir betrachten die Differenz der beiden Seiten der Ungleichung und erhalten:

$$\begin{aligned}
& \overbrace{\lambda\, g_1(x) + (1-\lambda)\, g_1(y)}^{\text{rechte Seite}} \qquad \overbrace{- g_1(\lambda\, x + (1-\lambda)\, y)}^{\text{linke Seite}} \\
&= \lambda\,(x^2 - 5) + (1-\lambda)\,(y^2-5) \qquad - \big((\lambda\, x + (1-\lambda)\, y)^2 - 5\big) \\
&= \lambda\, x^2 - 5\,\lambda + (1-\lambda)\, y^2 - 5\,(1-\lambda) - \lambda^2\, x^2 - 2\,\lambda\,(1-\lambda)\, x\, y - (1-\lambda)^2\, y^2 + 5 \\
&= \lambda\,(1-\lambda)\, x^2 + \lambda\,(1-\lambda)\, y^2 - 2\,\lambda\,(1-\lambda)\, x\, y \\
&= \underbrace{\lambda}_{\geq 0}\ \underbrace{(1-\lambda)}_{\geq 0}\ \underbrace{(x-y)^2}_{\geq 0} \geq 0 \quad \text{für alle } x, y \in \mathbb{R} \text{ und jedes } \lambda \in [0,1].
\end{aligned}$$

Daraus können wir schlussfolgern, dass die Funktion g_1 konvex auf $\mathbb{R}$ ist. Das Intervall (a, b) ist in diesem Fall also $(-\infty, \infty)$.

Bemerkung: Die Funktion g_1 ist sogar streng konvex auf $\mathbb{R}$, da die obige Ungleichung für $x \neq y$ nur für $\lambda = 0$ oder $\lambda = 1$ mit Gleichheit erfüllt ist.

Variante 2 (Überprüfung mittels zweiter Ableitung):

Da die reellwertigen Funktionen g_1, g_2 und g_3 nur von einer Variablen abhängen und auf ihren Definitionsbereichen jeweils zweimal differenzierbar sind, kann man auch mit Hilfe des Vorzeichens der zweiten Ableitung entscheiden, in welchen Bereichen sie konvex oder konkav sind, siehe (A.9). Damit ergeben sich die folgenden Eigenschaften:

$$\begin{aligned}
g_1''(x) &= 2 && > 0 \quad \text{für } x \in \mathbb{R} && \Rightarrow \quad g_1 \text{ ist (streng) konvex auf } \mathbb{R}, \\
g_2''(x) &= -\frac{1}{x^2} && < 0 \quad \text{für } x \in \mathbb{R}_+^* && \Rightarrow \quad g_2 \text{ ist (streng) konkav auf } \mathbb{R}_+^*, \\
g_3''(x) &= 6x - 12 && \begin{cases} \geq 0 \quad \text{für } x \geq 2 & \Rightarrow \quad g_3 \text{ ist konvex auf } [2, \infty), \\ < 0 \quad \text{für } x < 2 & \Rightarrow \quad g_3 \text{ ist (streng) konkav auf } (-\infty, 2). \end{cases}
\end{aligned}$$

Bemerkungen:

- Auf $\mathbb{R}$ ist die Funktion g_3 weder konvex noch konkav. Nur auf den eingeschränkten Mengen $[2, \infty)$ und $(-\infty, 2)$ gilt jeweils eine der beiden Eigenschaften.

- Alternativ kann man für die Funktion g_3 auch die Intervalle $(2, \infty)$ und $(-\infty, 2]$ wählen und die Schlussfolgerungen bezüglich der strengen Konvexität / Konkavität entsprechend anpassen.

Streng genommen müsste man noch die Definitionsbereiche der Funktionen auf Konvexität überprüfen. Offensichtlich ist dies aber bei allen Funktionen erfüllt.