Lösungen zu Tutoriumsblatt 8

Aufgabe T1 (Ungleichungen für Varianz und Standardabweichung)
Seien X und Y Zufallsvariablen auf einem diskreten Wahrscheinlichkeitsraum mit $\mathbb{E}[X^2] < \infty$ und $\mathbb{E}[Y^2] < \infty$.

(a) Gilt für die Varianz die Ungleichung

$$\mathbb{V}(X+Y) \leqslant \mathbb{V}(X) + \mathbb{V}(Y)?$$

Geben Sie einen Beweis oder ein explizites Gegenbeispiel an.

(b) Zeigen Sie:

$$\sigma(X+Y) \leqslant \sigma(X) + \sigma(Y).$$

Lösungsvorschlag:

(a) Wir geben ein Gegenbeispiel an. Sei $X \sim \mathrm{Bin}(1, \frac{1}{2})$ und $Y = X$. Dann gilt

$$1 = 4\mathbb{V}(X) = \mathbb{V}(2X) = \mathbb{V}(X+Y) > \mathbb{V}(X) + \mathbb{V}(Y) = 2\mathbb{V}(X) = \frac{1}{2}.$$

(b) Die Monotonie der Wurzelfunktion und die Cauchy-Schwarz-Ungleichung implizieren

$$\begin{aligned}
\sigma(X+Y) &= \sqrt{\mathbb{V}(X+Y)} \\
&= \sqrt{\mathbb{V}(X) + 2C(X,Y) + \mathbb{V}(Y)} \\
&\leqslant \sqrt{\mathbb{V}(X) + 2|C(X,Y)| + \mathbb{V}(Y)} \\
&\leqslant \sqrt{\mathbb{V}(X) + 2\sqrt{\mathbb{V}(X)\mathbb{V}(Y)} + \mathbb{V}(Y)} \\
&= \sqrt{\left(\sqrt{\mathbb{V}(X)} + \sqrt{\mathbb{V}(Y)}\right)^2} \\
&= \sqrt{\mathbb{V}(X)} + \sqrt{\mathbb{V}(Y)} = \sigma(X) + \sigma(Y).
\end{aligned}$$

Aufgabe T2 (Anwendung der Tschebyscheff-Ungleichung)
Sei X_n binomialverteilt mit Parametern $n \in \mathbb{N}$ und $p = \frac{1}{6}$.

(a) Berechnen Sie mithilfe der Tschebyscheff-Ungleichung Schranken für die folgenden Werte:

(i) $\mathbb{P}(18 \leqslant X_{120} \leqslant 22)$

(ii) $\mathbb{P}(180 \leqslant X_{1200} \leqslant 220)$

(b) Geben Sie für $i \in \{1,2\}$ jeweils, ebenfalls mithilfe der Tschebyscheff-Ungleichung, ein möglichst kleines $n_i \in \mathbb{N}$ an mit

$$\mathbb{P}\left(\left|\frac{1}{n_i}X_{n_i} - p\right| < 0.02\right) \geqslant 1 - \alpha_i, \qquad i \in \{1,2\},$$

wobei $\alpha_1 = 0.05$ und $\alpha_2 = 0.01$ ist.

Lösungsvorschlag:

(a) (i) Es gilt $X_{120} \sim \text{Bin}(120, \frac{1}{6})$ und somit $\mathbb{E}[X_{120}] = 120 \cdot \frac{1}{6} = 20$ und $\mathbb{V}(X_{120}) = 120 \cdot \frac{1}{6} \cdot \frac{5}{6} = \frac{50}{3}$. Mit der Tschebyscheff-Ungleichung erhalten wir

$$\mathbb{P}(18 \leqslant X_{120} \leqslant 22) = 1 - \mathbb{P}(|X_{120} - \mathbb{E}[X_{120}]| \geqslant 3) \geqslant 1 - \frac{\mathbb{V}(X_{120})}{3^2} = 1 - \frac{50}{27} = -\frac{23}{27}.$$

Da die untere Schranke hier negativ ist, liefert uns die Tschebyscheff-Ungleichung in diesem Fall keine neue Erkenntnis.

(ii) Es gilt $X_{1200} \sim \text{Bin}(1200, \frac{1}{6})$ und somit $\mathbb{E}[X_{1200}] = 1200 \cdot \frac{1}{6} = 200$ und $\mathbb{V}(X_{1200}) = 1200 \cdot \frac{1}{6} \cdot \frac{5}{6} = \frac{500}{3}$. Mit der Tschebyscheff-Ungleichung erhalten wir

$$\mathbb{P}(180 \leqslant X_{1200} \leqslant 220) = 1 - \mathbb{P}(|X_{1200} - \mathbb{E}[X_{1200}]| \geqslant 21) \geqslant 1 - \frac{\mathbb{V}(X_{1200})}{21^2} = 1 - \frac{500}{3 \cdot 21^2} \approx 0.622.$$

Bemerkung: Man kann die obigen Wahrscheinlichkeiten natürlich explizit berechnen. Dadurch ergeben sich die Werte $\mathbb{P}(18 \leqslant X_{120} \leqslant 22) \approx 0.459$ und $\mathbb{P}(180 \leqslant X_{1200} \leqslant 220) \approx 0.8878$.

(b) Mit der Tschebyscheff-Ungleichung gilt

$$\mathbb{P}\left(\left|\frac{1}{n_i}X_{n_i} - p\right| < 0.02\right) = 1 - \mathbb{P}\left(\left|X_{n_i} - n_i p\right| \geqslant 0.02 n_i\right) \geqslant 1 - \frac{p(1-p)}{0.02^2 n_i} \overset{!}{\geqslant} 1 - \alpha_i.$$

Weiterhin gilt

$$1 - \frac{p(1-p)}{0.02^2 n_i} \geqslant 1 - \alpha_i \quad \Longleftrightarrow \quad n_i \geqslant \frac{p(1-p)}{0.02^2 \alpha_i}.$$

Die Forderung ist somit für $n_1 \geqslant 6945$ und $n_2 \geqslant 34723$ erfüllt.

Aufgabe T3 (Arithmetisches Mittel, Maximum und Minimum von unabhängigen Zufallsvariablen)

Es seien $X_1, \ldots, X_n$ unabhängige Zufallsvariablen mit gleicher Verteilung und $\mathbb{E}[X_1^2] < \infty$. Ferner seien $\mu := \mathbb{E}[X_1]$, $\sigma^2 := \mathbb{V}(X_1)$ und $\overline{X}_n := \frac{1}{n}\sum_{j=1}^n X_j$. Zeigen Sie:

(a) $\mathbb{E}\left[\overline{X}_n\right] = \mu$.

(b) $\mathbb{V}\left(\overline{X}_n\right) = \frac{\sigma^2}{n}$.

(c) $\mathbb{C}\left(X_j, \overline{X}_n\right) = \frac{\sigma^2}{n}$ (für jedes $j \in [n]$).

(d) $\rho\left(X_1 - 2X_2, \overline{X}_n\right) = -\frac{1}{\sqrt{5n}}$ (Pearson-Korrelationskoeffizient).

(e) $\mathbb{P}(\max(X_1, \ldots, X_n) \leqslant t) = (\mathbb{P}(X_1 \leqslant t))^n, t \in \mathbb{R}$.

(f) $\mathbb{P}(\min(X_1, \ldots, X_n) \leqslant t) = 1 - (1 - \mathbb{P}(X_1 \leqslant t))^n, t \in \mathbb{R}$.

Überlegen Sie sich, ob für die Aussagen in (a)-(f) jeweils alle der obigen Voraussetzungen notwendig sind. Geben Sie ggf. schwächere Voraussetzungen an.

Lösungsvorschlag:

(a) Der Erwartungswert ist linear, also gilt $\mathbb{E}\left[\overline{X}_n\right] = \frac{1}{n}\sum_{j=1}^n \mathbb{E}[X_j] = \frac{1}{n} \cdot n \cdot \mu = \mu$. Hier wird lediglich benötigt, dass die Zufallsvariablen $X_1, ..., X_n$ alle den selben Erwartungswert besitzen.

(b) Da die X_i alle unabhängig sind, gilt

$$\mathbb{V}\left(\overline{X}_n\right) = \mathbb{V}\left(\frac{1}{n}\sum_{j=1}^n X_j\right) = \frac{1}{n^2}\sum_{j=1}^n \mathbb{V}(X_j) = \frac{1}{n^2} \cdot n \cdot \sigma^2 = \frac{\sigma^2}{n}.$$

Es wird verwendet, dass die Zufallsvariablen paarweise unkorreliert sind und die Tatsache, dass $X_1, ..., X_n$ die selbe Varianz besitzen.

(c) Die Kovarianz ist bilinear und X_i ist unabhängig von X_j für $i \neq j$. Also:

$$\begin{aligned}\mathbb{C}\left(X_j, \overline{X}_n\right) = \mathbb{C}\left(X_j, \frac{1}{n}\sum_{i=1}^{n} X_i\right) &= \frac{1}{n}\sum_{j=1}^{n} \underbrace{\mathbb{C}(X_j, X_i)}_{=0 \text{ für } i \neq j} \\ &= \frac{1}{n}\mathbb{C}(X_j, X_j) \\ &= \frac{1}{n}\mathbb{V}(X_j) = \frac{\sigma^2}{n}.\end{aligned}$$

Hier wurde lediglich die paarweise Unkorreliertheit der Zufallsvariablen verwendet, die aus der paarweisen Unabhängigkeit folgt. Zudem wird benötigt, dass alle Zufallsvariablen die selbe Varianz besitzen. Es wir jedoch nicht die Unabhängigkeit aller Zufallsvariablen benötigt.

(d) Nach Definition ist $\rho\left(X_1 - 2X_2, \overline{X}_n\right) = \frac{\mathbb{C}(X_1 - 2X_2, \overline{X}_n)}{\sqrt{\mathbb{V}(X_1 - 2X_2)\mathbb{V}(\overline{X}_n)}}$. Wir berechnen die noch fehlenden Größen:

$$\begin{aligned}\mathbb{C}(X_1 - 2X_2, \overline{X}_n) &= \mathbb{C}(X_1, \overline{X}_n) - 2\mathbb{C}(X_2, \overline{X}_n) \overset{(c)}{=} -\frac{\sigma^2}{n} \\ \mathbb{V}(X_1 - 2X_2) &= \mathbb{V}(X_1) + \mathbb{V}(-2X_2) = \mathbb{V}(X_1) + 4\mathbb{V}(X_2) = 5\sigma^2\end{aligned}$$

Beachte dabei, dass auch X_1 und $-2X_2$ unabhängig sind (Blockungslemma!). Somit:

$$\rho\left(X_1 - 2X_2, \overline{X}_n\right) = \frac{-\frac{\sigma^2}{n}}{\sqrt{5\sigma^2 \cdot \frac{\sigma^2}{n}}} = -\frac{1}{\sqrt{5n}}.$$

Diese Aussage benötig die selben Voraussetzungen wie Teil (c).

(e) Es gilt

$$\begin{aligned}\mathbb{P}(\max(X_1, \ldots, X_n) \leqslant t) &= \mathbb{P}(X_1 \leqslant t, \ldots, X_n \leqslant t) \\ &\overset{\text{Unabh.}}{=} \mathbb{P}(X_1 \leqslant t) \cdot \ldots \cdot \mathbb{P}(X_n \leqslant t) \\ &\overset{\text{identisch vert.}}{=} (\mathbb{P}(X_1 \leqslant t))^n.\end{aligned}$$

Wie in den Umformungen zu erkennen ist, wird die Unabhängigkeit aller Zufallsvariablen und deren identische Verteilung benötigt.

(f) Hier gilt

$$\begin{aligned}\mathbb{P}(\min(X_1, \ldots, X_n) \leqslant t) &= 1 - \mathbb{P}(\min(X_1, \ldots, X_n) > t) \\ &= 1 - \mathbb{P}(X_1 > t, \ldots, X_n > t) \\ &\overset{\text{Unabh.}}{=} 1 - \mathbb{P}(X_1 > t) \cdots \mathbb{P}(X_n > t) \\ &\overset{\text{identisch vert.}}{=} 1 - \mathbb{P}(X_1 > t)^n \\ &= 1 - (1 - \mathbb{P}(X_1 \leqslant t))^n.\end{aligned}$$

Teil (e) benötigt die selben Voraussetzungen wie Teil (f).

Aufgabe T4 (Doppelter Würfelwurf und Maximum)

Wir betrachten den doppelten Würfelwurf (mit fairen Würfeln). Es bezeichne X_1 das Ergebnis des ersten Würfels, X_2 das Ergebnis des zweiten Würfels und $Z = \max\{X_1, X_2\}$.

(a) Berechnen Sie den Erwartungswert von Z sowie die Varianz von X_1.

(b) Zeigen Sie, dass $\mathbb{E}[X_1 \cdot Z] = \frac{616}{36}$ gilt und berechnen Sie die Kovarianz von X_1 und Z.

(c) Was ist die beste affine Vorhersagefunktion für Z durch X_1 (siehe Satz 12.10)?

Lösungsvorschlag:

Da wir einen fairen, doppelten Würfelwurf betrachten, wählen wir den Grundraum $\Omega = \{1, \ldots, 6\}^2$ und als Wahrscheinlichkeitsmaß die Gleichverteilung $\mathbb{P}$.

(a) Es gilt

$$\begin{aligned}
\{Z=1\} &= \{(1,1)\},\\
\{Z=2\} &= \{(1,2),(2,2),(2,1)\},\\
\{Z=3\} &= \{(1,3),(2,3),(3,3),(3,2),(3,1)\},\\
\{Z=4\} &= \{(1,4),(2,4),(3,4),(4,4),(4,3),(4,2),(4,1)\},\\
\{Z=5\} &= \{(1,5),\dots,(5,5),\dots,(5,1)\},\\
\{Z=6\} &= \{(1,6),\dots,(6,6),\dots,(6,1)\}
\end{aligned}$$

und deswegen

$$\begin{aligned}
\mathbb{E}[Z] &= 1\cdot\frac{1}{36}+2\cdot\frac{3}{36}+3\cdot\frac{5}{36}+4\cdot\frac{7}{36}+5\cdot\frac{9}{36}+6\cdot\frac{11}{36}\\
&= \frac{161}{36}.
\end{aligned}$$

Mit

$$\mathbb{E}[X_1^2]=\frac{1}{6}(1+4+9+16+25+36)=\frac{91}{6}$$

folgt

$$\mathbb{V}(X_1)=\frac{91}{6}-\frac{49}{4}=\frac{182-147}{12}=\frac{35}{12}.$$

(b) Es gilt

$$\begin{aligned}
\mathbb{E}[X_1\cdot Z] &= \sum_{i=1}^{6}\sum_{j=1}^{6} i\cdot\max\{i,j\}\cdot\mathbb{P}(X_1=i,X_2=j)\\
&= \frac{1}{36}(1\cdot1+1\cdot2+1\cdot3+1\cdot4+1\cdot5+1\cdot6\\
&\quad +2\cdot2+2\cdot2+2\cdot3+2\cdot4+2\cdot5+2\cdot6\\
&\quad +3\cdot3+3\cdot3+3\cdot3+3\cdot4+3\cdot5+3\cdot6\\
&\quad +4\cdot4+4\cdot4+4\cdot4+4\cdot4+4\cdot5+4\cdot6\\
&\quad +5\cdot5+5\cdot5+5\cdot5+5\cdot5+5\cdot5+5\cdot6\\
&\quad +6\cdot6\cdot6)\\
&= \frac{616}{36}.
\end{aligned}$$

Damit können wir leicht die Kovarianz berechnen:

$$\mathbb{C}(X_1,Z)=\mathbb{E}[X_1\cdot Z]-\mathbb{E}[X_1]\cdot\mathbb{E}[Z]=\frac{616}{36}-\frac{7}{2}\cdot\frac{161}{36}=\frac{1232-1127}{72}=\frac{105}{72}.$$

(c) Nach Satz 12.10 ist die beste affine Vorhersage von Z durch X_1 gegeben durch

$$a^*+b^*X_1$$

mit

$$b^*=\frac{\mathbb{C}(X_1,Z)}{\mathbb{V}(X_1)},\quad a^*=\mathbb{E}Z-b^*\mathbb{E}(X_1).$$

Hier ist

$$b^*=\frac{105\cdot12}{72\cdot35}=\frac{1}{2}$$

und

$$a^*=\frac{161}{36}-\frac{1}{2}\cdot\frac{7}{2}=\frac{49}{18}.$$

X_1	1	2	3	4	5	6
$a^*+b^*X_1$	3.22	3.72	4.22	4.72	5.22	5.72

Aufgabe T5 (Lineare Regression)

(a) Es seien X_1, X_2 unabhängige, binomialverteilte Zufallsvariablen mit den Parametern n_1 und p, beziehungsweise n_2 und p mit $n_1, n_2 \in \mathbb{N}$ und $p \in [0,1]$. Weiter sei $Y = X_1 + X_2$. Bestimmen Sie Parameter $a, b \in \mathbb{R}$, sodass $f(a,b) = \mathbb{E}[(Y - a - bX_1)^2]$ minimal wird. Geben Sie das Minimum M^* an.

(b) Seien die Zufallsvariablen X, Y auf dem gemeinsamen Wahrscheinlichkeitsraum $(\Omega, \mathbb{P})$ gegeben, wobei deren gemeinsame Verteilung durch die unten angegebene Tabelle bestimmt ist. Bestimmen Sie $c \in [0, \frac{1}{2}]$, sodass

$$M_c^* := \min_{a,b\in\mathbb{R}} \mathbb{E}[(Y - a - bX)^2]$$

maximal wird. Welche Werte ergeben sich dabei für die optimalen Parameter a_c^*, b_c^* und M_c^*? Von welcher Art ist hierbei die Korrelation zwischen X und Y?

j \ k	1	2	$\mathbb{P}(Y = j)$
1	c	$\frac{1}{2} - c$	$\frac{1}{2}$
2	$\frac{1}{2} - c$	c	$\frac{1}{2}$
$\mathbb{P}(X = k)$	$\frac{1}{2}$	$\frac{1}{2}$	1

Lösungsvorschlag:

(a) Hierzu verwenden wir Satz 12.10 der Vorlesung. Dieser besagt, dass die optimalen Parameter gegeben sind durch

$$b^* = \frac{\mathbb{C}(Y, X_1)}{\mathbb{V}(X_1)}, \quad a^* = \mathbb{E}[Y] - b^*\mathbb{E}[X_1].$$

Die Erwartungswerte und Varianzen sind bekannt:

$$\mathbb{E}[Y] = (n_1 + n_2)p, \quad , \mathbb{V}(Y) = (n_1 + n_2)p(1-p), \quad \mathbb{E}[X_1] = n_1 p, \quad \mathbb{V}(X_1) = n_1 p(1-p).$$

Für die Kovarianz ergibt sich

$$\mathbb{C}(Y, X_1) = \mathbb{C}(X_1 + X_2, X_1) = \mathbb{C}(X_1, X_1) + \mathbb{C}(X_2, X_1) = \mathbb{V}(X_1) = n_1 p(1-p).$$

Dies ergibt für die Wahl der Parameter

$$b^* = 1, \quad a^* = pn_1 + pn_2 - pn_1 = pn_2 (= \mathbb{E}[X_2]).$$

Ebenfalls mit 12.10 erhält man den Minimalwert

$$M^\star = \mathbb{V}(Y) - \frac{\mathbb{C}(Y, X_1)^2}{\mathbb{V}(X_1)} = (n_1 + n_2)p(1-p) - \frac{(n_1 p(1-p))^2}{n_1 p(1-p)} = n_2 p(1-p).$$

Dies entspricht gerade der Varianz von X_2.

(b) Zunächst berechnen wir die Erwartungswerte und (Ko-)Varianzen, die wir zur Bestimmung der optimalen Parameter a_c^* und b_c^* mithilfe von Satz 12.10 benötigen.

$$\mathbb{E}[X] = \frac{1}{2} \cdot 1 + \frac{1}{2} \cdot 2 = \frac{3}{2} = \mathbb{E}[Y],$$

$$\mathbb{V}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \frac{1}{2} \cdot 1 + \frac{1}{2} \cdot 2^2 - \left(\frac{3}{2}\right)^2 = \frac{1}{4} = \mathbb{V}(Y),$$

$$\mathbb{C}(X, Y) = \mathbb{E}[X \cdot Y] - \mathbb{E}[X]\mathbb{E}[Y] = c \cdot 1 + 2 \cdot \left(\frac{1}{2} - c\right) \cdot 2 + c \cdot 4 - \frac{3}{2} \cdot \frac{3}{2} = c - \frac{1}{4}.$$

Mit Hilfe von Satz 12.10 erhalten wir dann

$$b_c^* = \frac{\mathbb{C}(Y,X)}{\mathbb{V}(X)} = \frac{c-\frac{1}{4}}{\frac{1}{4}} = 4c-1,$$

$$a_c^* = \mathbb{E}(Y) - b^\star \mathbb{E}(X) = \frac{3}{2} - \frac{3}{2} \cdot b^* = 3-6c.$$

Ebenso ergibt sich

$$M_c^* = \mathbb{V}(Y) - \frac{\mathbb{C}(Y,X)^2}{\mathbb{V}(X)} = \frac{1}{4} - \frac{(c-\frac{1}{4})^2}{\frac{1}{4}} = 2c-4c^2.$$

Dieser Ausdruck wird maximal für $c = \frac{1}{4}$, sodass $b^* = 0$, $a^* = \frac{3}{2}$, $M^* = \frac{1}{4}$ gilt. Zudem sind X und Y für $c = \frac{1}{4}$ unkorreliert.

Aufgabe T6 (Einstieg in die deskriptive Statistik)

a) Gegeben sei eine reellwertige Stichprobe $x = (x_1, \ldots, x_n)$ vom Umfang $n \in \mathbb{N}$. Zeigen Sie die folgende Beziehung für die Stichprobenvarianz s_x^2:

$$s_x^2 = \frac{1}{n-1}\left(\sum_{j=1}^n x_j^2 \;-\; n(\overline{x})^2\right).$$

b) Nun betrachten wir konkreter die folgende Stichprobe $x = (x_1, \ldots, x_{20})$ mit

$$\begin{aligned} x = (&19.23, 20.72, 20.14, 19.93, 19.25, 19.22, 19.13, 20.06, 19.96, 19.58,\\ &19.76, 20.80, 20.08, 19.95, 20.07, 20.07, 19.20, 19.24, 19.29, 20.58). \end{aligned}$$

(i) Bestimmen Sie das arithmetische Mittel $\bar{x}$, die Stichprobenvarianz s_x^2 und den Stichproben-Variationskoeffizienten v_x.

(ii) Zeichnen Sie ein Histogramm der Stichprobe x mit 5 Klassen, die äquidistant über das Intervall $(19.0, 21.0]$ verteilt sind.

Lösungsvorschlag:

(a) Nach Definition gilt $s_x^2 = \frac{1}{n-1}\sum_{j=1}^n (x_j - \overline{x})^2$. Wir müssen also nur zeigen, dass

$$\sum_{j=1}^n (x_j - \overline{x})^2 = \sum_{j=1}^n x_j^2 \;-\; n(\overline{x})^2$$

gilt:

$$\begin{aligned} \sum_{j=1}^n (x_j - \overline{x})^2 &= \sum_{j=1}^n \left(x_j^2 - 2x_j\overline{x} + (\overline{x})^2\right)\\ &= \sum_{j=1}^n x_j^2 \;+\; \sum_{j=1}^n (-2)\overline{x}x_j \;+\; \sum_{j=1}^n (\overline{x})^2\\ &= \sum_{j=1}^n x_j^2 \;-\; 2\overline{x} \cdot \sum_{j=1}^n x_j \;+\; n \cdot (\overline{x})^2\\ &= \sum_{j=1}^n x_j^2 \;-\; 2\overline{x} \cdot n \cdot \overline{x} \;+\; n \cdot (\overline{x})^2 = \sum_{j=1}^n x_j^2 \;-\; n \cdot (\overline{x})^2. \end{aligned}$$

(b) (i) Es gelten

$$\bar{x} = \frac{1}{20}\sum_{j=1}^{20} x_j = 19.813,$$

$$s_x^2 = \frac{1}{19}\left(\sum_{j=1}^{20} x_j^2 - 20 \cdot 19.813^2\right) = 0.2801,$$

$$s_x = \sqrt{s_x^2} = 0.5292,$$

$$v_x = \frac{s_x}{\bar{x}} = 0.0267.$$

(ii) Als Klasseneinteilung wählen wir im Bereich 19.0 bis 21.0 fünf äquidistante Klassen der Form $(19.0, 19.4]$, $(19.4, 19.8]$, $(19.8, 20.2]$, $(20.2, 20.6]$, $(20.6, 21.0]$ mit Intervalllänge 0.4. Die relativen Häufigkeiten innerhalb dieser Klassen sind $h_1 = \frac{7}{20}$, $h_2 = \frac{1}{10}$, $h_3 = \frac{2}{5}$, $h_4 = \frac{1}{20}$, $h_5 = \frac{1}{10}$. Daher ergeben sich die Balkenhöhen im Histogramm durch $d_k = h_k/0.4$, also $d_1 = \frac{7}{8}$, $d_2 = \frac{1}{4}$, $d_3 = 1$, $d_4 = \frac{1}{8}$, $d_5 = \frac{1}{4}$. Das Histogramm hat damit die folgende Form: