

Einführung in die angewandte Stochastik

1. Übung

Aufgabe 1

Die von einer Bäckerei angebotenen Brotsorten unterscheiden sich u.a. in den Ausprägungen folgender Merkmale:

- Preis pro Stück, Backtemperatur, Zuckerzusatz (ja/nein), Haltbarkeit, Produktionszahl pro Tag, Name, interne Produktnummer.

Geben Sie für jedes Merkmal mögliche Ausprägungen an und ordnen Sie es jeweils den folgenden Merkmalstypen zu:

- nominal, ordinal, diskret, stetig, dichotom, verhältnisskaliert, intervallskaliert.

Aufgabe 2

Bei einer Umfrage in einem Verein werden 50 Mitglieder zu ihrem Familienstand befragt. Die Antwortmöglichkeiten sind *ledig(l)*, *verheiratet(vh)*, *geschieden(g)* sowie *verwitwet(vw)*. Es werden folgende Antworten gegeben:

vh, vw, l, vw, g, vh, vh, l, g, vw, vw, l, vh, vw, l, vw, l, vh, vh, vh,
vh, l, g, vw, vh, vw, vh, vw, vh, vh, l, g, g, vh, l, g, vw, vh, vw, vh,
vw, vh, vw, g, vh, g, vh, vw, vh, vh.

- Bestimmen Sie die absoluten und relativen Häufigkeiten der einzelnen Ausprägungen des Merkmals Familienstand.
- Stellen Sie die absoluten Häufigkeiten in einem Stabdiagramm dar.
- Stellen Sie die relativen Häufigkeiten in einem Kreisdiagramm dar.

Aufgabe 3

Mit dem Ziel, Information über den Zeitaufwand im Bachelorstudium zu erhalten, werden 200 Studierende eines Studiengangs nach der Wochenstundenzahl der von ihnen im letzten Semester regelmäßig besuchten Veranstaltungen befragt. Es ergibt sich folgendes Bild:

Semesterwochenstunden	0-5	6-12	13-16	17-20	21-25	26-32
absolute Häufigkeit	9	16	28	72	47	28

- Stellen Sie die zugehörige absoluten Häufigkeiten sinnvoll grafisch dar. Beachten Sie dabei die Wahl der Klassengrenzen!
- Die selben Daten wurden in einer anderen Studie verwertet, in der die genaue Stundenzahl keine besondere Rolle spielte und daher wie folgt zusammengefasst wurde:

Semesterwochenstunden	genaue Anzahl
wenig	0-12
durchschnittlich	13-20
viel	21-32

Zeichnen Sie zu diesen Daten jeweils ein Säulen- und Kreisdiagramm.

Einführung in die angewandte Stochastik

Lösung zur 1. Übung

Lösung Aufgabe 1

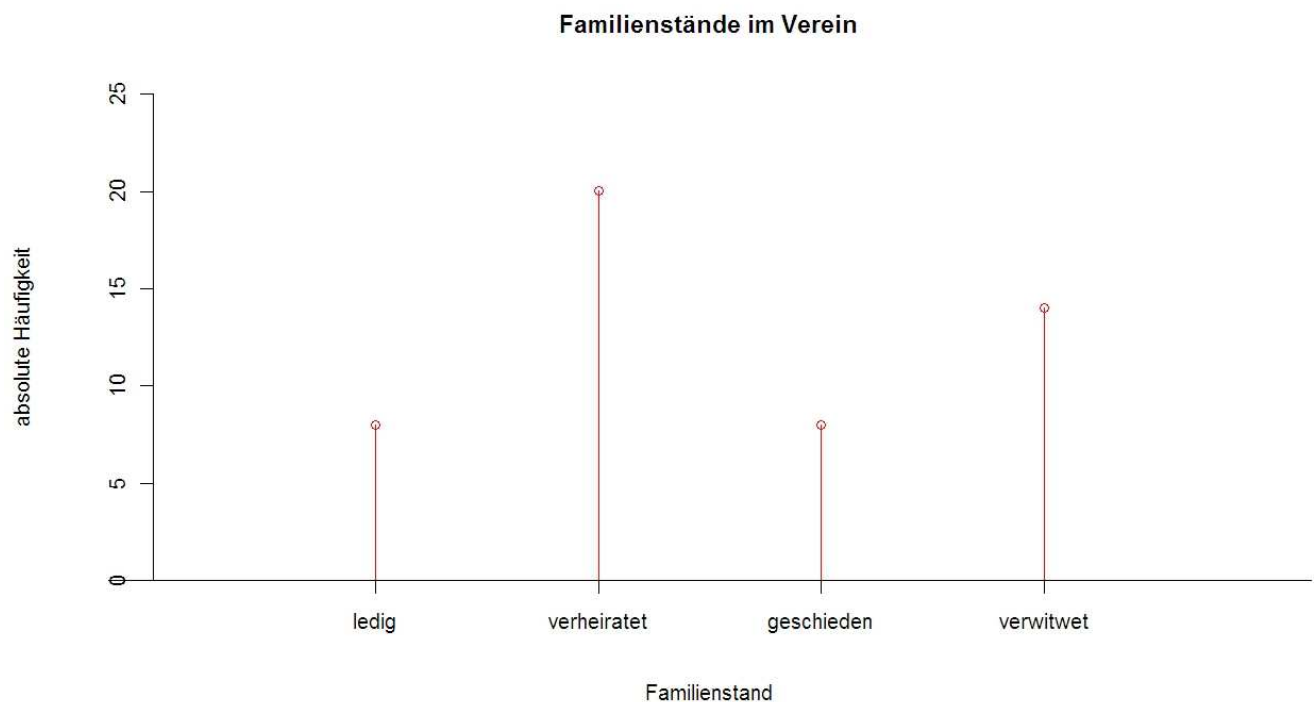
Merkmal	mögliche Ausprägung	Merkmalstyp
Preis pro Stück	2 Euro, 400 Cent etc.	diskret, verhältnisskaliert
Backtemperatur	250 Grad Celsius, 120 Grad Fahrenheit etc.	stetig, intervallskaliert
Zuckerzusatz	ja, nein	nominal, dichotom
Haltbarkeit	2 Tage, 3 Monate etc.	diskret, verhältnisskaliert
Produktionszahl pro Tag	100 Stück, 30 Stück etc.	diskret, verhältnisskaliert
Name	Vollkornbrot, Sonnenblumenkernbrot etc.	nominal
interne Produktnummer	126, 56 etc.	nominal

Lösung Aufgabe 2

(a)

Ausprägung	absolute Häufigkeit	relative Häufigkeit
ledig (l)	8	0.16
verheiratet (vh)	20	0.40
geschieden (g)	8	0.16
verwitwet (vw)	14	0.28

(b)

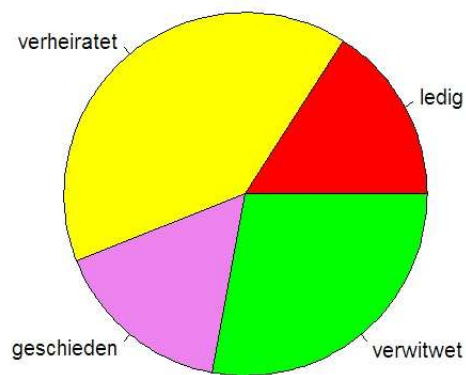


(c)

Benötigte Werte für die Erstellung des Kreisdiagramms:

Ausprägung	absolute Häufigkeit	relative Häufigkeit	Winkel im Kreisdiagramm
ledig	8	0.16	57,6°
verheiratet	20	0.40	144°
geschieden	8	0.16	57,6°
verwitwet	14	0.28	100,8°

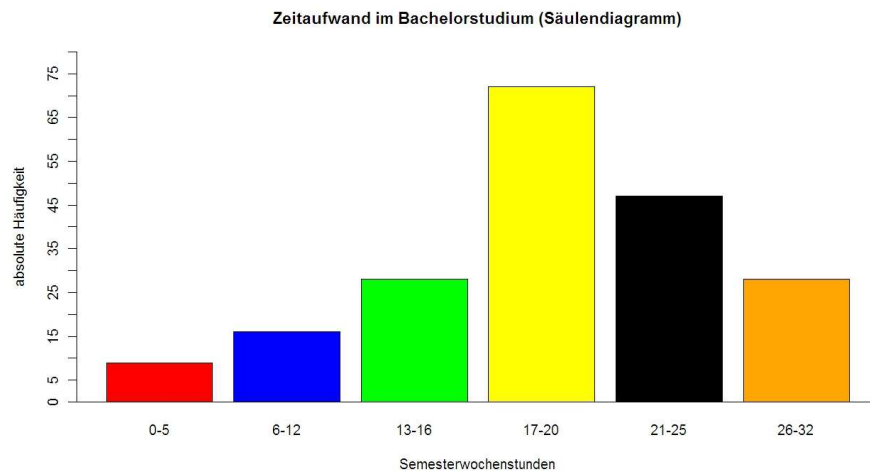
Familienstände im Verein



Lösung Aufgabe 3

(a)

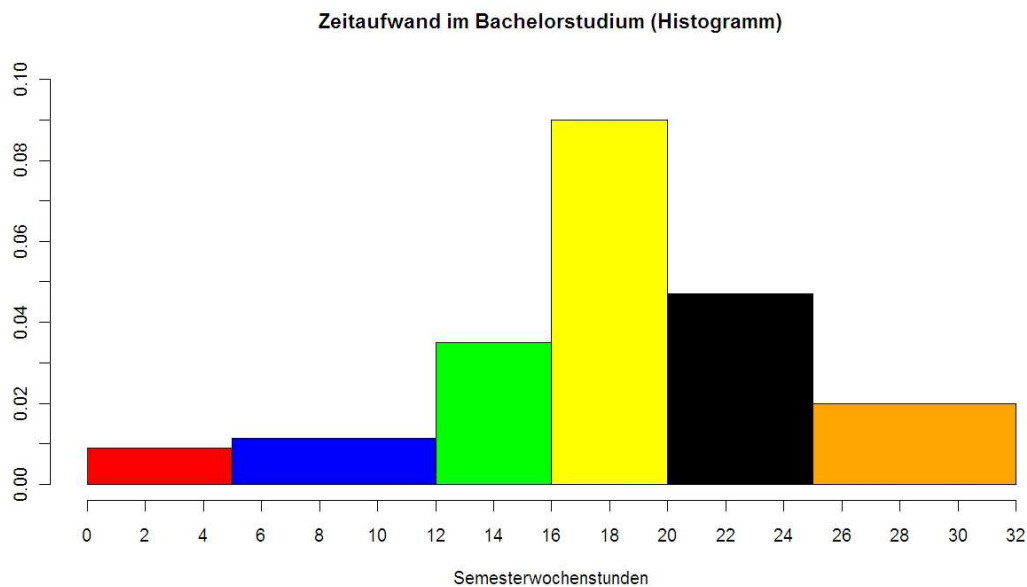
Erste Möglichkeit: Säulendiagramm



Zweite Möglichkeit: Histogramm

Die dafür benötigten Größen lauten:

Klasse	Klassenbreite	absolute Häufigkeit	relative Häufigkeit	Höhe
[0,5]	5	9	0.045	0.009
(5,12]	7	16	0.08	0.0114
(12,16]	4	28	0.14	0.035
(16,20]	4	72	0.36	0.09
(20,25]	5	47	0.235	0.047
(25,32]	7	28	0.14	0.02

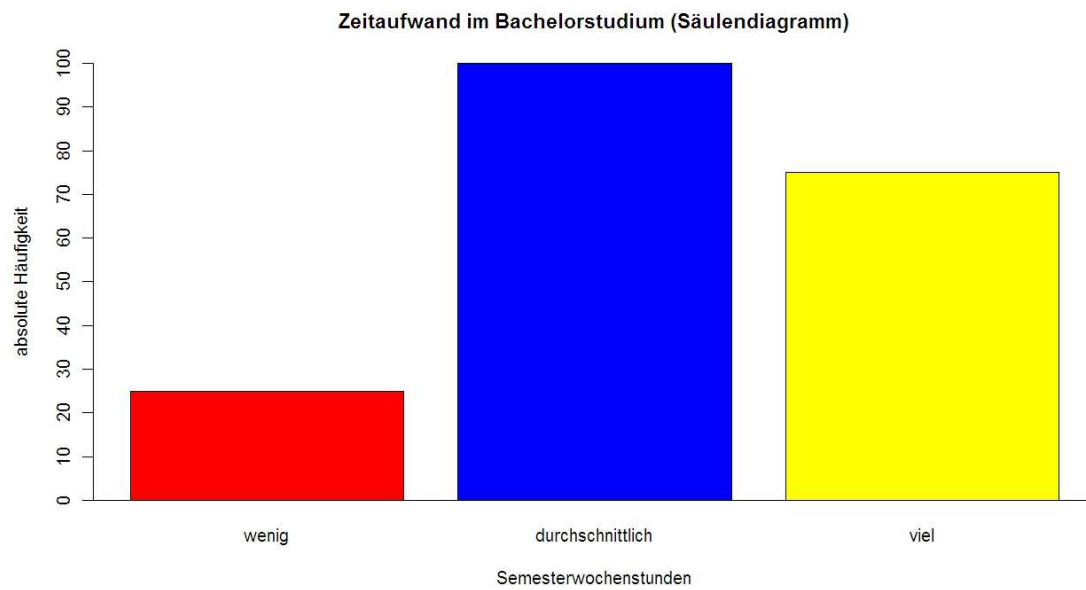


(b)

Benötigte Werte für die Erstellung der Diagramme:

Ausprägung	absolute Häufigkeit	relative Häufigkeit	Winkel im Kreisdiagramm
wenig	25	0.125	45°
durchschnittlich	100	0.5	180°
viel	75	0.375	135°

Säulendiagramm



Kreisdiagramm

Zeitaufwand im Bachelorstudium (Kreisdiagramm)



Einführung in die angewandte Stochastik

2. Übung

Aufgabe 4

Bei einem Quiz haben 15 Leute 8 Fragen beantwortet. Für die Personen ergaben sich jeweils folgende Anzahlen richtiger Antworten:

$$x_1 = 3, x_2 = 5, x_3 = 5, x_4 = 7, x_5 = 4, x_6 = 8, x_7 = 2, x_8 = 4, \\ x_9 = 4, x_{10} = 6, x_{11} = 2, x_{12} = 7, x_{13} = 6, x_{14} = 4, x_{15} = 6.$$

- (i) Ordnen Sie die Daten, und berechnen Sie die Ränge.
- (ii) Bestimmen Sie die empirische Verteilungsfunktion F_{15} zu dem vorliegenden Datensatz und zeichnen Sie ihren Graphen.
- (iii) Bestimmen Sie den Modus. Kann man diesen direkt aus dem Graphen der empirischen Verteilungsfunktion ablesen?
- (iv) Bestimmen Sie den Median, und vergleichen Sie ihn mit dem Modus.
- (v) Berechnen Sie das untere und obere Quartil sowie das 2-te Dezantil.

Aufgabe 5

Sei $x_1, \dots, x_n$, $n \in \mathbb{N}$, ein metrischer Datensatz. Zeigen Sie, dass der Median $\tilde{x}_{0.5}$ die folgende Abbildung minimiert:

$$g : \mathbb{R} \longrightarrow [0, \infty), \quad g(t) = \sum_{i=1}^n |t - x_i|.$$

Hinweis: Unterscheiden Sie die Fälle n ungerade, n gerade.

Aufgabe 6

Seien $x_1, \dots, x_n \in \mathbb{R}$ ein Datensatz und $a, b \in \mathbb{R}$.

- (i) Zeigen Sie, dass für den linear transformierten Datensatz $y_1, \dots, y_n$ mit $y_i = a + bx_i$, $i = 1, \dots, n$, gilt:

$$s_y = |b|s_x, \quad d_y = |b|d_x.$$

Zeigen Sie weiter, dass unter den Annahme $x_1, \dots, x_n > 0$ und $a = 0, b > 0$ für die Variationskoeffizienten gilt: $V_y = V_x$.

Hinweis: Verwenden Sie ohne Beweis, dass $\tilde{y}_{0.5} = a + b\tilde{x}_{0.5}$ gilt.

- (ii) Zeigen Sie:

$$s_x^2 = \overline{x^2} - \bar{x}^2, \quad \text{wobei } \bar{x}^2 = \frac{1}{n} \sum_{i=1}^n x_i^2.$$

Aufgabe 7

Bei einem Test, der mit den 32 SchülerInnen einer Klasse durchgeführt wurde, ergaben sich folgende Messwerte für die zum Lösen einer Rechenaufgabe benötigte Zeit (in Sekunden):

$$\begin{aligned}x_1 &= 7.2, x_2 = 5.6, x_3 = 2.9, x_4 = 4.1, x_5 = 3.9, x_6 = 9.1, x_7 = 5.2, x_8 = 5.2, \\x_9 &= 2.8, x_{10} = 5.1, x_{11} = 6.7, x_{12} = 8.1, x_{13} = 4.1, x_{14} = 4.8, x_{15} = 8.9, x_{16} = 3.0, \\x_{17} &= 7.7, x_{18} = 6.7, x_{19} = 8.0, x_{20} = 4.6, x_{21} = 4.8, x_{22} = 5.0, x_{23} = 7.0, x_{24} = 5.3, \\x_{25} &= 6.6, x_{26} = 8.8, x_{27} = 5.9, x_{28} = 6.8, x_{29} = 8.9, x_{30} = 6.1, x_{31} = 5.4, x_{32} = 4.7.\end{aligned}$$

- (i) Berechnen Sie für diesen Datensatz Modus, Quartilsmitte, 5%-Quantil, 90%-Quantil, arithmetisches Mittel, Spannweite, Quartilsabstand, Quartilskoeffizient, mittlere absolute Abweichung, Varianz, Standardabweichung und Variationskoeffizient.
- (ii) Erstellen Sie mit den Ergebnissen aus (i) einen Box-Plot für den Datensatz.
- (iii) Erstellen Sie ein Histogramm zur Kategorisierung $K_1^* = [1.5, 2.5]$, $K_2^* = (2.5, 3.5]$, $K_3^* = (3.5, 4.5]$, $K_4^* = (4.5, 5.5]$, $K_5^* = (5.5, 6.5]$, $K_6^* = (6.5, 7]$, $K_7^* = (7, 7.5]$, $K_8^* = (7.5, 8.5]$, $K_9^* = (8.5, 9.5]$.

Einführung in die angewandte Stochastik

Lösung zur 2. Übung

Lösung Aufgabe 4

(a) Rangwertreihe und Ränge:

i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
x_i	3	5	5	7	4	8	2	4	4	6	2	7	6	4	6
$x_{(i)}$	2	2	3	4	4	4	4	5	5	6	6	6	7	7	8
$R(x_i)$	3	8.5	8.5	13.5	5.5	15	1.5	5.5	5.5	11	1.5	13.5	11	5.5	11

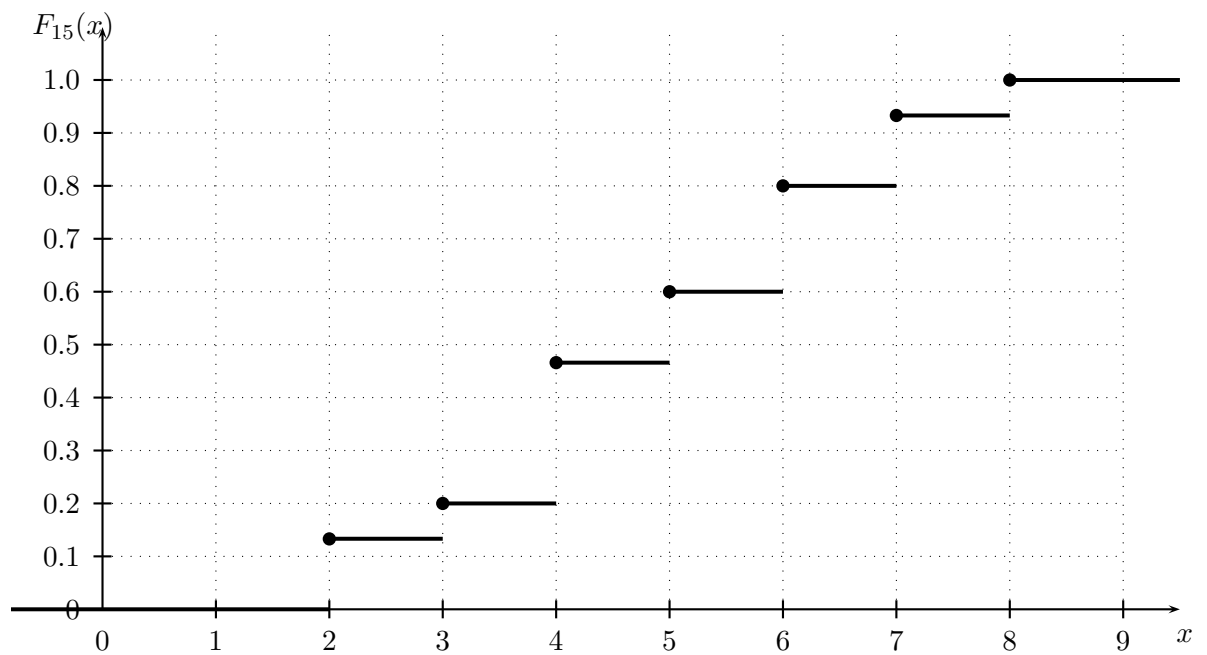
(b) Merkmalsausprägungen: $u_i = i + 1, i = 1, \dots, 7$.

Häufigkeiten:

i	1	2	3	4	5	6	7
u_i	2	3	4	5	6	7	8
n_i	2	1	4	2	3	2	1
f_i	$\frac{2}{15}$	$\frac{1}{15}$	$\frac{4}{15}$	$\frac{2}{15}$	$\frac{3}{15}$	$\frac{2}{15}$	$\frac{1}{15}$
$f_1 + \dots + f_i$	$\frac{2}{15}$	$\frac{3}{15}$	$\frac{7}{15}$	$\frac{9}{15}$	$\frac{12}{15}$	$\frac{14}{15}$	$\frac{15}{15}$

Empirische Verteilungsfunktion F_{15} :

$$F_{15}(x) = \begin{cases} 0, & x < 2, \\ \frac{2}{15}, & 2 \leq x < 3, \\ 0.2, & 3 \leq x < 4, \\ \frac{7}{15}, & 4 \leq x < 5, \\ 0.6, & 5 \leq x < 6, \\ 0.8, & 6 \leq x < 7, \\ \frac{14}{15}, & 7 \leq x < 8, \\ 1, & 8 \leq x. \end{cases}$$



Lösung Aufgabe 5

Seien $x_1, \dots, x_n \in \mathbb{R}$ ein Datensatz und $a, b \in \mathbb{R}$.

z.z.: $\tilde{x}_{0,5}$ minimiert die Abbildung $g: \mathbb{R} \rightarrow [0, \infty)$ definiert durch

$$g(t) = \sum_{i=1}^n |t - x_i|.$$

Beweis:

1. Fall: n ungerade, d.h. $n = 2k - 1$ für ein $k \in \mathbb{N}$.

Nach Vorlesung gilt $\tilde{x}_{0,5} = x_{(\frac{n+1}{2})} = x_{(k)}$. Weiter gilt für alle $t \in \mathbb{R}$:

$$\begin{aligned} g(\tilde{x}_{0,5}) &= \sum_{i=1}^n |x_{(k)} - x_i| = \sum_{i=1}^n |x_{(k)} - x_{(i)}| \\ &= \sum_{i=1}^{k-1} (x_{(k)} - x_{(i)}) + \sum_{i=k+1}^n (x_{(i)} - x_{(k)}) \\ &= \sum_{i=1}^{k-1} (t - x_{(i)}) + \sum_{i=k+1}^n (x_{(i)} - t) \\ &\leq \sum_{i=1}^n |t - x_{(i)}| + \sum_{i=k+1}^n |x_{(i)} - t| \\ &\stackrel{(*)}{\leq} \sum_{i=1}^n |t - x_{(i)}| = \sum_{i=1}^n |t - x_i| = g(t) \end{aligned}$$

Somit ist $\tilde{x}_{0,5}$ globale Minimalstelle von g .

Wegen (*) gilt, dass $g(\tilde{x}_{0,5}) < g(t) \quad \forall t \in \mathbb{R}, \quad t \neq \tilde{x}_{0,5}$. $\tilde{x}_{0,5}$ ist folglich die eindeutig bestimmte globale Minimalstelle von g .

2. Fall: n gerade, d.h. $n = 2k$ für ein $k \in \mathbb{N}$.

Nach Vorlesung gilt: $\tilde{x}_{0,5} = \frac{1}{2}(x_{(k)} + x_{(k+1)})$. Für $t \in \mathbb{R}$ gilt:

$$\begin{aligned} g(\tilde{x}_{0,5}) &= \sum_{i=1}^n |\tilde{x}_{0,5} - x_i| \stackrel{\text{s.o.}}{=} \sum_{i=1}^k (\tilde{x}_{0,5} - x_{(i)}) + \sum_{i=k+1}^n (x_{(i)} - \tilde{x}_{0,5}) \\ &= \sum_{i=1}^k (t - x_{(i)}) + \sum_{i=k+1}^n (x_{(i)} - t) \\ &\leq \sum_{i=1}^n |t - x_i| = g(t). \end{aligned}$$

Daraus folgt, dass $\tilde{x}_{0,5}$ eine globale Minimalstelle von g ist.

In diesem Fall ist jedoch jedes $x \in [x_{(k)}, x_{(k+1)}]$ eine globale Minimalstelle von g (konkrete Darstellung von $\tilde{x}_{0,5}$ geht in obige Rechnung nicht ein!).

Lösung Aufgabe 6

Seien $x_1, \dots, x_n \in \mathbb{R}$ ein Datensatz und $a, b \in \mathbb{R}$.

(i) **z.z.:** Für $y_1, \dots, y_n$ mit $y_i = a + bx_i$, $i = 1, \dots, n$, gilt:

$$s_y = |b|s_x \quad \text{und} \quad d_y = |b|d_x.$$

Zunächst gilt $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i = \frac{1}{n} \sum_{i=1}^n (a + bx_i) = a + b\bar{x}$.

$$\begin{aligned} \implies s_y^2 &= \frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2 = \frac{1}{n} \sum_{i=1}^n (a + bx_i - (a + b\bar{x}))^2 \\ &= \frac{1}{n} \sum_{i=1}^n b^2 (x_i - \bar{x})^2 = b^2 s_x^2 \quad \implies \quad s_y = |b|s_x \\ d_y &= \frac{1}{n} \sum_{i=1}^n |y_i - \tilde{y}_{0,5}| \stackrel{\text{Hinweis}}{=} \frac{1}{n} \sum_{i=1}^n |a + bx_i - (a + b\tilde{x}_{0,5})| \\ &= \frac{1}{n} \sum_{i=1}^n |b| |x_i - \tilde{x}_{0,5}| = |b|d_x \end{aligned}$$

z.z.: Für $x_1, \dots, x_n > 0$, $a = 0$ und $b > 0$ gilt: $V_y = V_x$.

Wegen $x_i > 0, b > 0$ gilt $y_i = bx_i > 0 \quad \forall i = 1, \dots, n$.

Somit sind die Variationskoeffizienten für $x_1, \dots, x_n$ bzw. $y_1, \dots, y_n$ definiert und es gilt:

$$V_y = \frac{s_y}{\bar{y}} \stackrel{\text{s.o.}}{=} \frac{|b|s_x}{b\bar{x}} \stackrel{b>0}{=} \frac{bs_x}{b\bar{x}} = \frac{s_x}{\bar{x}} = V_x.$$

□

(ii) **z.z.:** $s_x^2 = \overline{x^2} - \bar{x}^2$, wobei $\overline{x^2} = \frac{1}{n} \sum_{i=1}^n x_i^2$.

Beweis:

$$\begin{aligned} s_x^2 &\stackrel{\text{Def.}}{=} \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 = \frac{1}{n} \sum_{i=1}^n (x_i^2 - 2x_i\bar{x} + \bar{x}^2) \\ &= \frac{1}{n} \sum_{i=1}^n x_i^2 - 2\bar{x} \frac{1}{n} \sum_{i=1}^n x_i + \bar{x}^2 \\ &= \overline{x^2} - 2\bar{x}^2 + \bar{x}^2 = \overline{x^2} - \bar{x}^2. \end{aligned}$$

□

Lösung Aufgabe 7

- (i) Um die angegebenen Kenngrößen berechnen zu können, bestimmen wir zunächst die Rangwertreihe des Datensatzes:

i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
x_i	7.2	5.6	2.9	4.1	3.9	9.1	5.2	5.2	2.8	5.1	6.7	8.1	4.1	4.8	8.9	3.0
$x_{(i)}$	2.8	2.9	3.0	3.9	4.1	4.1	4.6	4.7	4.8	4.8	5.0	5.1	5.2	5.2	5.3	5.4

i	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
x_i	7.7	6.7	8.0	4.6	4.8	5.0	7.0	5.3	6.6	8.8	5.9	6.8	8.9	6.1	5.4	4.7
$x_{(i)}$	5.6	5.9	6.1	6.6	6.7	6.7	6.8	7.0	7.2	7.7	8.0	8.1	8.8	8.9	8.9	9.1

Damit ergibt sich:

Modus

$$x_{mod} \in \{4.1, 4.8, 5.2, 6.7, 8.9\} \longrightarrow \text{Modus hier nicht eindeutig bestimmt.}$$

Quartilsmitte

Zunächst Bestimmung des oberen bzw. unteren Quartils:

- $n \cdot 0.25 = 8 \in \mathbb{N} \implies \tilde{x}_{0.25} = \frac{1}{2}(x_{(8)} + x_{(9)}) = 4.75$
- $n \cdot 0.75 = 24 \in \mathbb{N} \implies \tilde{x}_{0.75} = \frac{1}{2}(x_{(24)} + x_{(25)}) = 7.1$

$$\implies Q_{\text{mitte}} = \frac{1}{2}(\tilde{x}_{0.25} + \tilde{x}_{0.75}) = \frac{1}{2} \cdot 11.85 = 5.925$$

5%-Quantil

$$n \cdot 0.05 = 1.6 \notin \mathbb{N} \implies \tilde{x}_{0.05} = x_{(2)} = 2.9$$

90%-Quantil

$$n \cdot 0.9 = 28.8 \notin \mathbb{N} \implies \tilde{x}_{0.9} = x_{(29)} = 8.8$$

arithmetisches Mittel

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i = \frac{1}{32} \sum_{i=1}^{32} x_i = 5.90625$$

Spannweite

$$R = x_{(32)} - x_{(1)} = 9.1 - 2.8 = 6.3$$

Quartilsabstand & -koeffizient

$$Q = \tilde{x}_{0.75} - \tilde{x}_{0.25} \stackrel{s.o.}{=} 7.1 - 4.75 = 2.35$$

$$Q_{\text{koeff}} = \frac{Q}{Q_{\text{mitte}}} \approx 0.397$$

mittlere absolute Abweichung

$$\tilde{x}_{0.5} = \frac{1}{2}(x_{(16)} + x_{(17)}) = 5.5 \longrightarrow d_x = \frac{1}{32} \sum_{i=1}^{32} |x_i - 5.5| = 1.475$$

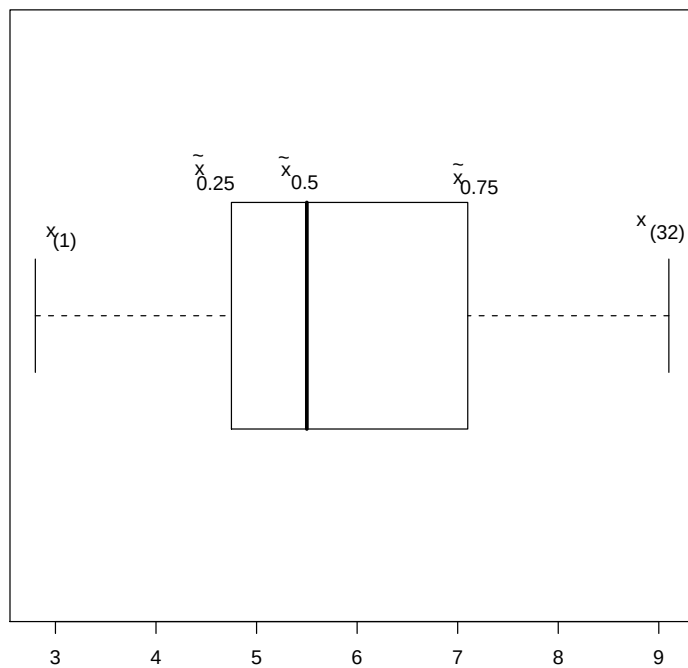
Varianz & Standardabweichung

$$s_x^2 = \frac{1}{32} \sum_{i=1}^{32} (x_i - \bar{x})^2 = \overline{x^2} - \bar{x}^2 \approx 3.151 \rightarrow s_x \approx 1.775$$

Variationskoeffizient

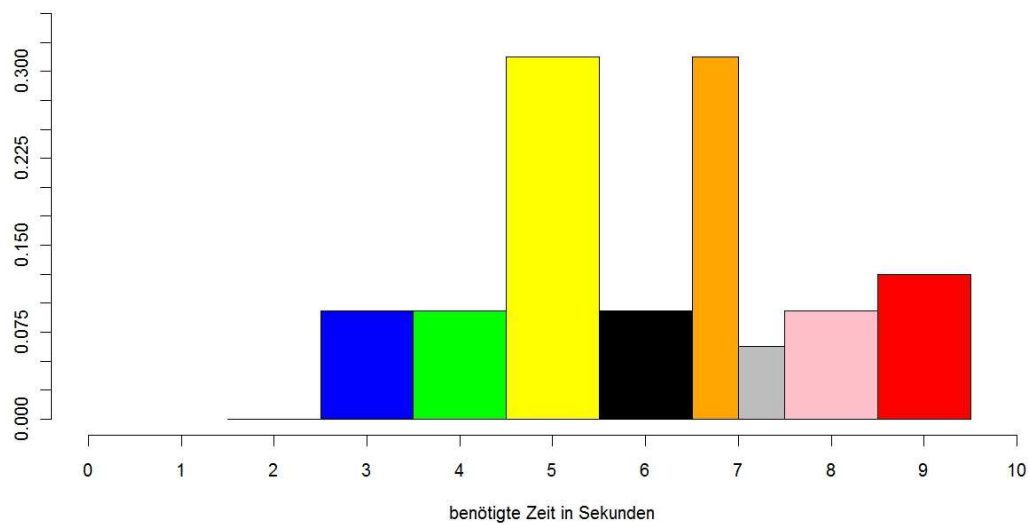
$$V = \frac{s_x}{\bar{x}} \stackrel{s.o.}{\approx} 0.301$$

(ii) Boxplot zum gegebenen Datensatz:



(iii) Histogramm:

Benötigte Zeit zum Lösen einer Rechenaufgabe (Histogramm)



Einführung in die angewandte Stochastik

3. Übung

Aufgabe 8

- (a) Zeigen Sie: Für Zahlen $a_1, \dots, a_n \in \mathbb{R}$ und $b_1, \dots, b_n \in \mathbb{R}$ gilt die Ungleichung

$$\left(\sum_{i=1}^n a_i b_i \right)^2 \leq \sum_{i=1}^n a_i^2 \cdot \sum_{i=1}^n b_i^2.$$

In dieser Ungleichung liegt Gleichheit genau dann vor, wenn ein $c \in \mathbb{R}$ existiert, so dass entweder

$$a_i = c \cdot b_i \text{ für alle } i \in \{1, \dots, n\} \quad \text{oder} \quad b_i = c \cdot a_i \text{ für alle } i \in \{1, \dots, n\}.$$

- (b) Zeigen Sie: Sind $(x_1, y_1), \dots, (x_n, y_n)$ ein gepaarter Datensatz zum bivariaten Merkmal (X, Y) und $s_x > 0$ bzw. $s_y > 0$ die Standardabweichungen von $x_1, \dots, x_n$ bzw. $y_1, \dots, y_n$, so gilt für den Bravais-Pearson-Korrelationskoeffizient r_{xy} :

- (i) Aus $r_{xy} = -1$ folgt, es gibt eine negative Zahl $b < 0$ und eine reelle Zahl $a \in \mathbb{R}$ mit

$$y_i = a + bx_i, \quad i = 1, \dots, n.$$

- (ii) Aus $r_{xy} = 1$ folgt, es gibt eine positive Zahl $b > 0$ und eine reelle Zahl $a \in \mathbb{R}$ mit

$$y_i = a + bx_i, \quad i = 1, \dots, n.$$

Aufgabe 9

Für eine Produktionsanlage soll ein lineares Kostenmodell aufgrund der Daten aus vergangenen Perioden angepasst werden. Diese Daten bestehen aus produzierten Mengen x_i (in 1000 Stück) und entstandenen Kosten y_i (in 1000 Euro), $i = 1, \dots, 6$, in sechs Perioden:

Menge x_i	7	9	6	11	8	7
Kosten y_i	16.5	22.0	15.0	24.5	19.0	17.0

Schätzen Sie die Regressionsgerade $f(x) = a + bx$, und berechnen Sie das Bestimmtheitsmaß.

Einführung in die angewandte Stochastik

Lösung zur 3. Übung

Lösung Aufgabe 8

(a) Seien

$$A := \sqrt{\sum_{i=1}^n a_i^2} \quad \text{und} \quad B := \sqrt{\sum_{i=1}^n b_i^2}.$$

O.E. gelte $A, B > 0$. Setze für $i \in \{1, 2, \dots, n\}$

$$\alpha_i := \frac{a_i}{A} \quad \text{und} \quad \beta_i := \frac{b_i}{B}.$$

Dann gilt wegen $\sum_{i=1}^n \alpha_i^2 = \sum_{i=1}^n \beta_i^2 = 1$:

$$0 \leq \sum_{i=1}^n (\alpha_i \pm \beta_i)^2 = \sum_{i=1}^n \alpha_i^2 + \sum_{i=1}^n \beta_i^2 \pm 2 \sum_{i=1}^n \alpha_i \beta_i = 2 \left(1 \pm \sum_{i=1}^n \alpha_i \beta_i \right).$$

Damit haben wir

$$-1 \leq \sum_{i=1}^n \alpha_i \beta_i \leq 1$$

mit

$$-1 = \sum_{i=1}^n \alpha_i \beta_i \iff \sum_{i=1}^n (\alpha_i + \beta_i)^2 = 0 \iff \alpha_i = -\beta_i \quad \forall i \in \{1, 2, \dots, n\}$$

und

$$1 = \sum_{i=1}^n \alpha_i \beta_i \iff \sum_{i=1}^n (\alpha_i - \beta_i)^2 = 0 \iff \alpha_i = \beta_i \quad \forall i \in \{1, 2, \dots, n\}.$$

D.h.:

$$\begin{aligned} -1 \leq \sum_{i=1}^n \alpha_i \beta_i \leq 1 &\iff 0 \leq \left(\sum_{i=1}^n \alpha_i \beta_i \right)^2 \leq 1 \\ &\iff 0 \leq \left(\sum_{i=1}^n a_i b_i \right)^2 \leq A^2 B^2 \\ &\iff 0 \leq \left(\sum_{i=1}^n a_i b_i \right)^2 \leq \sum_{i=1}^n a_i^2 \cdot \sum_{i=1}^n b_i^2 \end{aligned}$$

mit

$$\left(\sum_{i=1}^n a_i b_i \right)^2 = \sum_{i=1}^n a_i^2 \cdot \sum_{i=1}^n b_i^2 \iff \left(\sum_{i=1}^n \alpha_i \beta_i \right)^2 = 1,$$

also

$$\alpha_i = \beta_i \quad \forall i \in \{1, 2, \dots, n\} \quad \text{oder} \quad \alpha_i = -\beta_i \quad \forall i \in \{1, 2, \dots, n\},$$

was gleichbedeutend ist mit

$$a_i = \frac{A}{B} \cdot b_i \quad \forall i \in \{1, 2, \dots, n\} \quad \text{oder} \quad a_i = -\frac{A}{B} \cdot b_i \quad \forall i \in \{1, 2, \dots, n\}.$$

(b) (i) Nach Voraussetzung gilt $r_{xy} = -1$, d.h.:

$$\begin{aligned} -1 &= \frac{s_{xy}}{s_x s_y} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \\ \implies \left(\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) \right)^2 &= \sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2. \end{aligned}$$

Aus Teil (a) folgt, es existiert ein $c \in \mathbb{R}$ mit

$$y_i - \bar{y} = c(x_i - \bar{x}) \quad \forall i \in \{1, 2, \dots, n\}.$$

Setze also $b = c$ und $a = \bar{y} - c\bar{x}$, dann gilt

$$y_i = a + bx_i, \quad i = 1, \dots, n.$$

Wäre hierbei $b \geq 0$, so erhielten wir die offenbar falsche Ungleichung

$$-1 = \frac{s_{xy}}{s_x s_y} = \frac{bs_{xx}}{s_x s_y} \geq 0,$$

so dass alles gezeigt ist.

(ii) Analog zu (i).

Lösung Aufgabe 9

Die geschätzte Regressionsgerade ist gegeben durch

$$\hat{f}(x) = \hat{a} + \hat{b}x,$$

wobei sich $\hat{a}$ und $\hat{b}$ wie folgt berechnen lassen:

$$\hat{b} = \frac{s_{xy}}{s_x^2} \quad \text{und} \quad \hat{a} = \bar{y} - \hat{b} \cdot \bar{x}.$$

Die zur Berechnung benötigten Werte sind gegeben durch:

Beobachtung	x_i	y_i	x_i^2	y_i^2	$x_i y_i$
1	7	16.5	49	272,25	115.5
2	9	22	81	484	198
3	6	15	36	225	90
4	11	24.5	121	600.25	269.5
5	8	19	64	361	152
6	7	17	49	289	119
Mittelwert	8	19	$\frac{200}{3}$	$\frac{4463}{12}$	$\frac{472}{3}$

Damit ist $\hat{b}$ gegeben durch

$$\hat{b} = \frac{s_{xy}}{s_x^2} = \frac{\overline{xy} - \bar{x} \cdot \bar{y}}{\overline{x^2} - (\bar{x})^2} = \frac{\frac{472}{3} - 8 \cdot 19}{\frac{200}{3} - 8^2} = \frac{\frac{16}{3}}{\frac{8}{3}} = 2$$

und $\hat{a}$ durch

$$\hat{a} = \bar{y} - \hat{b} \cdot \bar{x} = 19 - 2 \cdot 8 = 3.$$

Die geschätzte Regressionsgerade lautet also

$$\hat{f}(x) = 3 + 2x,$$

und das Bestimmtheitsmaß berechnet sich zu

$$B_{xy} = r_{xy}^2 = \left(\frac{s_{xy}}{s_x s_y} \right)^2 = \frac{(\overline{xy} - \bar{x} \cdot \bar{y})^2}{\left(\overline{x^2} - (\bar{x})^2 \right) \left(\overline{y^2} - (\bar{y})^2 \right)} = \frac{\left(\frac{16}{3} \right)^2}{\frac{8}{3} \cdot \left(\frac{4463}{12} - 19^2 \right)} = \frac{128}{131} \approx 0.9770.$$

Einführung in die angewandte Stochastik

4. Übung

Aufgabe 10

Bei einer medizinischen Untersuchung wird der gesundheitliche Zustand von Rauchern und Nichtrauchern verglichen. Dabei werden die an der Untersuchung teilnehmenden Personen in Männer und Frauen sowie in Raucher und Nichtraucher eingeteilt. Die Menge der Männer wird mit A bezeichnet, die Menge der Raucher mit B .

Welche Personen werden durch die folgenden Mengen beschrieben?

- (i) B^c
- (ii) $A \cap B$
- (iii) $B \setminus A$
- (iv) $(A \cup B)^c$
- (v) $(A \cup B^c) \cap (A^c \cup B)$

Aufgabe 11

In einem Wahrscheinlichkeitsraum $(\Omega, \mathfrak{A}, P)$ seien von den Ereignissen A , B und C die folgenden Wahrscheinlichkeiten bekannt:

- (i) $P(B) = \frac{7}{20}$,
- (ii) $P(C^c) = \frac{7}{10}$,
- (iii) $P(A) = \frac{3}{10}$,
- (iv) $P(A^c \cap C) = \frac{1}{4}$,
- (v) $P(A \cap B) = \frac{1}{10}$,
- (vi) $P(A \cap B \cap C) = \frac{1}{20}$,
- (vii) $P((A \cup B) \cap C) = \frac{3}{20}$.

Bestimmen Sie die Wahrscheinlichkeiten der nachfolgenden Ereignisse:

- (a) $A \cup B$,
- (b) $A^c \cup C$,
- (c) $A \cap C$,
- (d) $A \cap B^c \cap C$,
- (e) $B \cap C$,
- (f) $A \cup B \cup C$.

Aufgabe 12

Es werden gleichzeitig ein roter und ein schwarzer Würfel geworfen (beides unverfälschte, jeweils sechsseitige Würfel mit den Augenzahlen $1, \dots, 6$).

- (a) Geben Sie zu diesem Zufallsexperiment einen geeigneten Grundraum Ω an, und beschreiben Sie die folgenden Ereignisse als Teilmengen von Ω :
 - (i) *Die Augensumme beträgt 8,*
 - (ii) *Die Augensumme ist kleiner als 5,*
 - (iii) *Beide Augenzahlen betragen jeweils höchstens 3,*
 - (iv) *Die Augenzahl des roten Würfels ist geringer als die des schwarzen Würfels.*
- (b) Berechnen Sie für die Ereignisse aus (a) die zugehörigen Wahrscheinlichkeiten.

Aufgabe 13

Eine faire Münze (d.h. die Auftretenswahrscheinlichkeiten von Kopf und Zahl sind identisch) werde sieben Mal hintereinander geworfen.

- (a) Geben Sie einen geeigneten Wahrscheinlichkeitsraum für dieses Experiment an.
- (b) Wie viele mögliche Versuchsausgänge gibt es?
- (c) Beschreiben Sie die nachstehenden Ereignisse als Teilmengen von Ω und berechnen Sie ihre Wahrscheinlichkeiten:
 - (i) *Genau sechs der sieben Münzwürfe ergeben Zahl.*
 - (ii) *Mindestens sechs der sieben Münzwürfe ergeben Zahl.*
 - (iii) *In zwei aufeinander folgenden Würfen fällt Kopf oder in zwei aufeinander folgenden Würfen fällt Zahl.*

Aufgabe 14

Eine Gruppe von k Personen sei hinsichtlich ihrer Geburtstage in einem Jahr (kein Schaltjahr, d.h. 365 Tage) zufällig zusammengesetzt. Bestimmen Sie die Wahrscheinlichkeit dafür, dass mindestens zwei der k Personen am selben Tag des Jahres Geburtstag haben.

Hinweis: Betrachten Sie das Komplementärereignis.

Einführung in die angewandte Stochastik

Lösung zur 4. Übung

Lösung Aufgabe 10

Welche Personen werden durch die folgenden Mengen beschrieben?

- (a) $B^c \hat{=} \text{Nichtraucher}$
- (b) $A \cap B \hat{=} \text{männliche Raucher}$
- (c) $B \setminus A = B \cap A^c \hat{=} \text{weibliche Raucher}$
- (d) $(A \cup B)^c = A^c \cap B^c \hat{=} \text{weibliche Nichtraucher}$
- (e) Es gilt

$$\begin{aligned}(A \cup B^c) \cap (A^c \cup B) &= (A \cap (A^c \cup B)) \cup (B^c \cap (A^c \cup B)) \\ &= \underbrace{(A \cap A^c)}_{=\emptyset} \cup (A \cap B) \cup (B^c \cap A^c) \cup \underbrace{(B^c \cap B)}_{=\emptyset} \\ &= (A \cap B) \cup (A^c \cap B^c).\end{aligned}$$

Es handelt sich also um die Menge aller männlichen Raucher und weiblichen Nichtraucher.

Lösung Aufgabe 11

(a) $P(A \cup B) = P(A) + P(B) - P(A \cap B) \stackrel{(i),(iii),(v)}{=} \frac{3}{10} + \frac{7}{20} - \frac{1}{10} = \frac{11}{20},$

(b)

$$\begin{aligned}P(A^c \cup C) &= P(A^c) + P(C) - P(A^c \cap C) = 1 - P(A) + 1 - P(C^c) - P(A^c \cap C) \\ &\stackrel{(ii),(iii),(iv)}{=} 1 - \frac{3}{10} + 1 - \frac{7}{10} - \frac{1}{4} = \frac{3}{4},\end{aligned}$$

(c) $P(A \cap C) = P(C) - P(A^c \cap C) = 1 - P(C^c) - P(A^c \cap C) \stackrel{(ii),(iv)}{=} 1 - \frac{7}{10} - \frac{1}{4} = \frac{1}{20},$

(d) $P(A \cap B^c \cap C) = P(A \cap C) - P(A \cap B \cap C) \stackrel{c),(vi)}{=} \frac{1}{20} - \frac{1}{20} = 0,$

(e)

$$\begin{aligned}P((A \cup B) \cap C) &= P((A \cap C) \cup (B \cap C)) \\ &= P(A \cap C) + P(B \cap C) - P(A \cap C \cap B \cap C), \\ \implies P(B \cap C) &= P((A \cup B) \cap C) - P(A \cap C) + P(A \cap B \cap C) \\ &\stackrel{(vii),(c),(vi)}{=} \frac{3}{20} - \frac{1}{20} + \frac{1}{20} = \frac{3}{20},\end{aligned}$$

(f)

$$P(A \cup B \cup C) = P((A \cup B) \cup C) \stackrel{(a),(ii),(vii)}{=} \frac{P(A \cup B) + P(C) - P((A \cup B) \cap C)}{\frac{11}{20} + 1 - \frac{7}{10} - \frac{3}{20}} = \frac{7}{10}.$$

Lösung Aufgabe 12

(a) Grundraum

$$\Omega = \{(\omega_1, \omega_2) \mid \omega_i \in \{1, \dots, 6\}, i \in \{1, 2\}\}$$

mit der Interpretation ($1 \leq i, j \leq 6$):

$\omega_1 = i$ bedeutet *"Roter Würfel zeigt die Zahl i"*,

$\omega_2 = j$ bedeutet *"Schwarzer Würfel zeigt die Zahl j"*.

Für die Anzahl der Elemente in Ω gilt $|\Omega| = 36$.

Menge A für *"Die Augensumme beträgt 8"*

$$\begin{aligned} A &= \{(\omega_1, \omega_2) \in \Omega \mid \omega_1 + \omega_2 = 8\} \\ &= \{(2, 6), (3, 5), (4, 4), (5, 3), (6, 2)\} \\ |A| &= 5 \end{aligned}$$

Menge B für *"Die Augensumme ist kleiner als 5"*

$$\begin{aligned} B &= \{(\omega_1, \omega_2) \in \Omega \mid \omega_1 + \omega_2 < 5\} \\ &= \{(1, 1), (1, 2), (1, 3), (2, 1), (2, 2), (3, 1)\} \\ |B| &= 6 \end{aligned}$$

Menge C für *"Beide Augenzahlen betragen jeweils höchstens 3"*

$$\begin{aligned} C &= \{(\omega_1, \omega_2) \in \Omega \mid \omega_1 \leq 3, \omega_2 \leq 3\} \\ &= \{(1, 1), (1, 2), (1, 3), (2, 1), (2, 2), (2, 3), (3, 1), (3, 2), (3, 3)\} \\ |C| &= 9 \end{aligned}$$

Menge D für *"Die Augenzahl des roten Würfels ist geringer als die des schwarzen Würfels"*

$$\begin{aligned} D &= \{(\omega_1, \omega_2) \in \Omega \mid \omega_1 < \omega_2\} \\ &= \{(1, 2), (1, 3), (1, 4), (1, 5), (1, 6), (2, 3), (2, 4), (2, 5), (2, 6), (3, 4), (3, 5), (3, 6), \\ &\quad (4, 5), (4, 6), (5, 6)\} \\ |D| &= 15 \end{aligned}$$

(b) Beschreibung mit Laplace-Raum (unverfälschte Würfel!), d.h.

$$P(\{\omega\}) = \frac{1}{|\Omega|} = \frac{1}{36} \quad \text{für alle } \omega \in \Omega.$$

Daher ergibt sich

$$\begin{aligned} P(A) &= \frac{|A|}{|\Omega|} = \frac{5}{36}, & P(B) &= \frac{|B|}{|\Omega|} = \frac{6}{36} = \frac{1}{6}, \\ P(C) &= \frac{|C|}{|\Omega|} = \frac{9}{36} = \frac{1}{4}, & P(D) &= \frac{|D|}{|\Omega|} = \frac{15}{36} = \frac{5}{12}. \end{aligned}$$

Lösung Aufgabe 13

(a) Grundraum

$$\Omega = \{(\omega_1, \dots, \omega_7) \mid \omega_i \in \{0, 1\}, i \in \{1, \dots, 7\}\}$$

mit der Interpretation ($1 \leq i \leq 7$):

$\omega_i = 0$ bedeutet "In Wurf i erscheint Kopf",

$\omega_i = 1$ bedeutet "In Wurf i erscheint Zahl".

Damit ist durch

$$P(A) = \frac{|A|}{|\Omega|}, \quad A \subseteq \Omega$$

ein Laplace-Raum auf Ω gegeben.

(b) Das Experiment kann als 7-maliges Ziehen aus einer Urne mit 2 unterscheidbaren Kugeln modelliert werden, wobei die Reihenfolge, in der die Kugeln gezogen werden, beachtet wird und gezogene Kugeln wieder in die Urne zurückgelegt werden. Gemäß Vorlesung gibt es

$$|\Omega| = 2^7 = 128$$

mögliche Versuchsausgänge.

(c) (i) Ereignis A für "Genau sechs der sieben Münzwürfe ergeben Zahl":

$$\begin{aligned} A &= \{(\omega_1, \dots, \omega_7) \in \Omega \mid \exists! i \in \{1, \dots, 7\} : \omega_i = 0, \omega_j = 1 \forall j \in \{1, \dots, 7\} \setminus \{i\}\} \\ &= \{(0, 1, \dots, 1), (1, 0, 1, \dots, 1), \dots, (1, \dots, 1, 0)\}. \end{aligned}$$

Somit gilt:

$$P(A) = \frac{|A|}{|\Omega|} = \frac{7}{128}.$$

(ii) Ereignis B für "Mindestens sechs der sieben Münzwürfe ergeben Zahl":

$$B = A \cup \{(1, \dots, 1)\} = \{(0, 1, \dots, 1), (1, 0, 1, \dots, 1), \dots, (1, \dots, 1, 0), (1, \dots, 1, 1)\}.$$

Somit gilt:

$$P(B) = \frac{|B|}{|\Omega|} = \frac{8}{128} = \frac{1}{16}.$$

- (iii) Ereignis C für *"In zwei aufeinander folgenden Würfeln fällt Kopf oder in zwei aufeinander folgenden Würfeln fällt Zahl"*:

$$C = \{(\omega_1, \dots, \omega_7) \in \Omega \mid \exists i \in \{1, \dots, 6\} : \omega_i = \omega_{i+1} = 0\} \\ \cup \{(\omega_1, \dots, \omega_7) \in \Omega \mid \exists j \in \{1, \dots, 6\} : \omega_j = \omega_{j+1} = 1\}.$$

C besitzt das komplementäre Ereignis C^C *"Keine zwei aufeinander folgende Würfe zeigen dasselbe Symbol"* und es gilt:

$$C^C = \{(1, 0, 1, 0, 1, 0, 1), (0, 1, 0, 1, 0, 1, 0)\}.$$

Somit gilt:

$$P(C) = 1 - P(C^C) = 1 - \frac{|C^C|}{|\Omega|} = 1 - \frac{2}{128} = \frac{63}{64}.$$

Lösung Aufgabe 14

Betrachte Murmelmodell: Verteile k unterscheidbare Murmeln (Personen) auf 365 Zellen (Tage des Jahres), wobei Mehrfachbelegungen zugelassen sind

Beschreibung mittels des Grundraums

$$\Omega = \{(\omega_1, \dots, \omega_k) \mid \omega_i \in \{1, \dots, 365\}, 1 \leq i \leq k\}$$

mit der Interpretation ($1 \leq i \leq k, 1 \leq l \leq 365$):

$\omega_i = l$ bedeutet *"Die Person i hat am l -ten Tag des Jahres Geburtstag"*

Dann wird durch

$$P(A) = \frac{|A|}{|\Omega|}, \quad A \subset \Omega,$$

ein Laplace-Raum auf Ω definiert (vgl. B 1.11). Außerdem gilt $|\Omega| = 365^k$.

Gesucht ist die Wahrscheinlichkeit für das Ereignis *"Mindestens 2 der Personen haben am selben Tag des Jahres Geburtstag"*. Dieses Ereignis wird durch die Menge

$$E = \{(\omega_1, \dots, \omega_k) \in \Omega \mid \text{es gibt } i, j \in \{1, \dots, k\}, i \neq j, \text{ mit } \omega_i = \omega_j\}$$

beschrieben. Daher lässt sich das Komplementäreignis *"Alle Personen haben an verschiedenen Tagen des Jahres Geburtstag"* durch das Komplement

$$E^c = \{(\omega_1, \dots, \omega_k) \in \Omega \mid \text{für alle } i, j \in \{1, \dots, k\}, i \neq j, \text{ gilt } \omega_i \neq \omega_j\}$$

beschreiben. Aus B 1.13 folgt, dass $|E^c| = \frac{365!}{(365-k)!}$ gilt. Wegen $|E| + |E^c| = |\Omega|$ ergibt sich damit für die gesuchte Wahrscheinlichkeit

$$P(E) = \frac{|E|}{|\Omega|} = \frac{|\Omega| - |E^c|}{|\Omega|} = 1 - \frac{|E^c|}{|\Omega|} = 1 - \frac{365!}{365^k(365-k)!}.$$

Einführung in die angewandte Stochastik

5. Übung

Aufgabe 15

- (a) Zeigen Sie, dass folgende Mengensysteme σ -Algebren über $\Omega \neq \emptyset$ sind:
- (i) $\mathfrak{A} = \{\emptyset, \Omega\}$,
 - (ii) $\mathfrak{A} = \{\emptyset, A, A^c, \Omega\}$, $\emptyset \neq A \subset \Omega$, $A \neq \Omega$,
 - (iii) $\mathfrak{A} = \mathfrak{Pot}(\Omega)$,
 - (iv) $\mathfrak{A} = \{A \subset \Omega \mid A \text{ höchstens abzählbar oder } A^c \text{ höchstens abzählbar}\}$.
- (b) Weisen Sie nach, dass die nachstehenden Funktionen Zähl- bzw. Riemannsdichten auf dem jeweils angegebenen Träger darstellen:
- (i) $p_k = p(1-p)^k$, $k \in \mathbb{N}_0$, für ein $p \in (0, 1)$ (Geometrische Verteilung $geo(p)$),
 - (ii) $p_k = \frac{\lambda^k}{k!} e^{-\lambda}$, $k \in \mathbb{N}_0$, für ein $\lambda > 0$ (Poisson-Verteilung $po(\lambda)$),
 - (iii) $f(x) = \lambda e^{-\lambda x}$, $x \geq 0$, für ein $\lambda > 0$ (Exponentialverteilung $Exp(\lambda)$),
 - (iv) $f(x) = \frac{\alpha}{x^{\alpha+1}}$, $x \geq 1$, für ein $\alpha > 0$ (Pareto-Verteilung $Par(\alpha)$).

Aufgabe 16

Eine Softwarefirma beschäftigt drei Programmierer P_1, P_2 und P_3 . Von P_1 wurden 230, von P_2 690 und von P_3 460 Programmierungen im vergangenen Jahr vorgenommen. Hierbei haben bei

- | | |
|---------|--|
| P_1 : | 12% aller Programme mindestens zwei Programmierfehler,
40% aller Programme genau einen Programmierfehler, |
| P_2 : | 15% aller Programme genau einen Programmierfehler,
70% aller Programme keinen Programmierfehler, |
| P_3 : | 75% aller Programme keinen Programmierfehler,
10% aller Programme mindestens zwei Programmierfehler. |

Die Softwarefirma wählt aus allen geschriebenen Programmen zufällig eines aus.

- (a) Mit welcher Wahrscheinlichkeit hat das ausgewählte Programm keine Programmierfehler?
- (b) Die Softwarefirma stellt fest, dass das Programm genau einen Programmierfehler aufweist. Mit welcher Wahrscheinlichkeit stammt es vom Programmierer P_2 ?

Aufgabe 17

Ein fairer Würfel werde zweimal geworfen.

- (a) Zeigen Sie, dass für alle $k \in \{1, \dots, 6\}$ gilt: Die Ereignisse

„1. Wurf zeigt Ziffer k “ und
„Die Summe der Ziffern beider Würfe ist durch drei teilbar“

sind stochastisch unabhängig.

- (b) Gibt es ein $k \in \{1, \dots, 6\}$, so dass die Ereignisse

„1. Wurf zeigt Ziffer k “ und
„Das Produkt der Ziffern beider Würfe ist durch drei teilbar“

stochastisch unabhängig sind?

Aufgabe 18

Ein Software-Unternehmen stellt im Internet kostenlos ein Statistik-Lernprogramm zum Download zur Verfügung. Es gibt zwei verschiedene Möglichkeiten das Programm zu installieren:

- (1) Das Programm wird in 2 Datenpakete zerlegt und übermittelt. Nach dem Download wird das Programm installiert, falls keines der beiden Datenpakete fehlerhaft ist.
- (2) Das Programm wird in 4 Datenpakete zerlegt und übermittelt. Nach dem Download wird das Programm installiert, falls höchstens eines der Datenpakete fehlerhaft ist.

Nehmen Sie an, dass unabhängig von der Größe eines Datenpaketes die Fehlerrate bzw. die Fehlerwahrscheinlichkeit bei seiner Übermittlung $1-p$ für ein $p \in (0, 1)$ betrage und dass die Übermittlung der einzelnen Datenpakete unabhängig voneinander erfolge. Geben Sie ein geeignetes stochastisches Modell für diese Situation an und berechnen Sie in Abhängigkeit von der Intaktwahrscheinlichkeit p , welche der beiden Varianten zuverlässiger zu einer Installation der Software führt.

Einführung in die angewandte Stochastik

6. Übung

Aufgabe 19

- (a) Auf einen Server werden täglich mehrere Hackerangriffe verübt. Die Anzahl der an einem Tag registrierten Angriffe kann durch eine Poisson-verteilte Zufallsvariable mit Parameter $\lambda = 6$ beschrieben werden.

Berechnen Sie die Wahrscheinlichkeit dafür, dass an einem Tag

- (i) genau fünf Angriffe,
- (ii) mindestens vier Angriffe

verzeichnet werden.

- (b) Die Lebensdauer (in Stunden) von bestimmten Bauteilen werde durch eine mit Parameter $\lambda > 0$ exponentialverteilte Zufallsvariable X beschrieben.

- (i) Berechnen Sie für $\lambda = \frac{1}{800}$ die Wahrscheinlichkeit dafür, dass die Lebensdauer eines derartigen Bauteils

- (I) höchstens 300 Stunden,
 - (II) mehr als 120 Stunden,
 - (III) zwischen 240 und 360 Stunden
- beträgt.

- (ii) Für welchen Wert des Parameters λ ergibt sich eine Lebensdauerverteilung, bei der mit Wahrscheinlichkeit 0,99 die Lebensdauer eines derartigen Bauteils mindestens 100 Stunden beträgt?

Aufgabe 20

Die Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ sei gegeben durch

$$f(x) = \begin{cases} \frac{c}{\sqrt{x+2}} & , \quad 2 \leq x \leq 7 , \\ 0 & , \quad \text{sonst} \end{cases}$$

mit einer Konstanten $c \in \mathbb{R}$.

- (i) Bestimmen Sie c so, dass f Dichtefunktion einer stetigen Zufallsvariablen X ist.
- (ii) Bestimmen Sie die Verteilungsfunktion F^X der Zufallsvariablen X .

Aufgabe 21

Seien X und Y diskrete Zufallsvariablen auf einem Wahrscheinlichkeitsraum $(\Omega, \mathfrak{A}, P)$ mit $P(X = k) = P(Y = k + 1) = 1/4$ für $k \in \{0, 1, 2, 3\}$.

- (a) Bestimmen Sie Verteilung von $X + Y$ unter der Annahme, dass X und Y stochastisch unabhängig sind.
- (b) Wie sieht die Verteilung von $X + Y$ aus, falls X und Y *nicht* stochastisch unabhängig sind, sondern die gemeinsame Verteilung

$$P(X = k, Y = k + 1) = 1/4, \quad k \in \{0, 1, 2, 3\},$$

besitzen?

Aufgabe 22

Seien X und Y zwei diskrete Zufallsvariablen, wobei X die Werte $-1, 0$ und 1 und Y die Werte $1, 2$ und 3 annimmt. Die zugehörigen Wahrscheinlichkeiten $p_{ij} = P(X = i, Y = j)$ für $i \in \{-1, 0, 1\}$, $j \in \{1, 2, 3\}$, sind in der folgenden Tabelle angegeben:

$p_{ij} = P(X = i, Y = j)$		j		
		1	2	3
i	-1	1/20	1/5	0
	0	1/5	1/5	1/10
	1	1/10	1/10	1/20

- (a) Bestimmen Sie die Zähldichten der beiden Randverteilungen, und entscheiden Sie, ob die Zufallsvariablen X und Y stochastisch unabhängig sind.
- (b) Bestimmen Sie die bedingte (Zähl-) Dichte von Y bei gegebenem $X = 1$.
- (c) Berechnen Sie die bedingte Wahrscheinlichkeit $P(X = 0 | Y \geq 2)$.
- (d) Bestimmen Sie die Wahrscheinlichkeitsverteilung der Zufallsvariablen $Z = X + Y$.

Einführung in die angewandte Stochastik

Lösung zur 6. Übung

Lösung Aufgabe 19

- (a) Die Anzahl der pro Tag registrierten Hackerangriffe wird beschrieben durch eine Poisson-verteilte ZV X mit Parameter $\lambda = 6$ ($X \sim po(6)$). Nach B 2.5 ist die zugehörige Zähldichte definiert durch

$$p_k = P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda} = \frac{6^k}{k!} e^{-6}, \quad k \in \mathbb{N}_0.$$

Folglich gilt:

- (i) Die Wahrscheinlichkeit für genau fünf registrierte Angriffe ist gegeben durch

$$P(X = 5) = p_5 = \frac{6^5}{5!} e^{-6} \approx 0,161.$$

- (ii) Die Wahrscheinlichkeit für mindestens vier registrierte Angriffe ist gleich

$$\begin{aligned} P(X \geq 4) &= 1 - P(X < 4) = 1 - P(X \leq 3) \\ &= 1 - e^{-6} \sum_{k=0}^3 \frac{6^k}{k!} = 1 - e^{-6}(1 + 6 + 18 + 36) \approx 0,849. \end{aligned}$$

- (b) Sei $X \sim Exp(\lambda)$ mit $\lambda > 0$. Dann ist nach B 3.7 die zugehörige Verteilungsfunktion gegeben durch:

$$F^X(x) = \begin{cases} 1 - e^{-\lambda x}, & x > 0, \\ 0, & x \leq 0. \end{cases}$$

- (i) Sei $\lambda = \frac{1}{800}$. Dann gilt:

- (I) $P(X \leq 300) = F^X(300) = 1 - e^{-3/8} \approx 0,313$
(II) $P(X > 120) = 1 - P(X \leq 120) = 1 - F^X(120) = e^{-3/20} \approx 0,861$
(III) Da die Verteilungsfunktion von X stetig ist (vgl. C 2.4), folgt

$$\begin{aligned} P(240 < X < 360) &= P(240 \leq X \leq 360) = P(240 < X \leq 360) \\ &= F^X(360) - F^X(240) = 1 - e^{-9/20} - (1 - e^{-3/10}) \approx 0,103. \end{aligned}$$

- (ii) Gesucht wird $\lambda > 0$ mit $P(X \geq 100) = 0,99$. Die Lösung ergibt sich folgendermaßen:

$$\begin{aligned} P(X \geq 100) &= P(X > 100) = 1 - P(X \leq 100) \\ &= 1 - F^X(100) = e^{-100\lambda} \stackrel{!}{=} 0,99 \\ \iff -100\lambda &= \ln(0,99) \\ \iff \lambda &= -\frac{\ln(0,99)}{100} \approx 0,0001005 = 1,005 \cdot 10^{-4}. \end{aligned}$$

Lösung Aufgabe 20

Die gegebene Funktion $f: \mathbb{R} \rightarrow \mathbb{R}$ ist Dichtefunktion einer stetigen Zufallsvariable X , falls (vgl. B 3.4):

(a) $f(x) \geq 0$ für alle $x \in \mathbb{R}$,

(b) $\int_{-\infty}^{\infty} f(x)dx = 1$.

(i) Aus (a) folgt $c \geq 0$. Aus (b) ergibt sich

$$1 = \int_{-\infty}^{\infty} f(x)dx = \int_2^7 \frac{c}{\sqrt{x+2}} dx = 2c \sqrt{(x+2)} \Big|_2^7 = 2c,$$

d.h. für $c = \frac{1}{2}$ ist f eine Dichtefunktion.

(ii) Die Verteilungsfunktion $F^X : \mathbb{R} \rightarrow [0, 1]$ von X ist gegeben durch (vgl. C 2.2(ii)):

$$F^X(t) = \int_{-\infty}^t f(x)dx, \quad t \in \mathbb{R}.$$

Mit $c = \frac{1}{2}$ folgt

- für $t < 2$:

$$F^X(t) = \int_{-\infty}^t f(x)dx = \int_{-\infty}^t 0 dx = 0.$$

- für $2 \leq t < 7$:

$$F^X(t) = \int_2^t \frac{1}{2\sqrt{x+2}} dx = \sqrt{(x+2)} \Big|_2^t = \sqrt{t+2} - 2.$$

- für $t \geq 7$:

$$F^X(t) = \int_2^7 \frac{1}{2\sqrt{x+2}} dx = \sqrt{9} - 2 = 1.$$

Zusammengefasst ergibt sich also für die Verteilungsfunktion:

$$F^X(t) = \begin{cases} 0, & t < 2, \\ \sqrt{t+2} - 2, & 2 \leq t < 7, \\ 1, & t \geq 7. \end{cases}$$

Lösung Aufgabe 21

Es gilt allgemein für diskrete Zufallsvariablen X und Y mit Werten in $\mathbb{Z}$ (vgl. Beweis zu Satz C 1.13):

$$P(X + Y = k) = \sum_{j \in \mathbb{Z}} P(X = j, Y = k - j), \quad k \in \mathbb{Z}.$$

Seien X und Y Zufallsvariablen mit $P(X = k) = P(Y = k + 1) = 1/4$ für $k \in \{0, 1, 2, 3\}$.

(a) Aufgrund der Unabhängigkeit gilt in diesem Fall (vgl. Satz C 1.13)

$$P(X + Y = k) = \sum_{j \in \mathbb{Z}} P(X = j)P(Y = k - j), \quad k \in \mathbb{Z}.$$

Es folgt (beachte die Träger der Verteilungen von X und Y):

$$P(X + Y = 1) = P(X = 0)P(Y = 1) = \frac{1}{16}$$

$$P(X + Y = 2) = P(X = 1)P(Y = 1) + P(X = 0)P(Y = 2) = \frac{1}{8}$$

$$P(X + Y = 3) = P(X = 0)P(Y = 3) + P(X = 1)P(Y = 2) + P(X = 2)P(Y = 1) = \frac{3}{16}$$

$$P(X + Y = 4) = P(X = 0)P(Y = 4) + P(X = 1)P(Y = 3) \\ + P(X = 2)P(Y = 2) + P(X = 3)P(Y = 1) = \frac{1}{4}$$

$$P(X + Y = 5) = P(X = 1)P(Y = 4) + P(X = 2)P(Y = 3) + P(X = 3)P(Y = 2) = \frac{3}{16}$$

$$P(X + Y = 6) = P(X = 2)P(Y = 4) + P(X = 3)P(Y = 3) = \frac{1}{8}$$

$$P(X + Y = 7) = P(X = 3)P(Y = 4) = \frac{1}{16}$$

Aufgrund der Träger von X und Y nimmt $X + Y$ nur Werte in der Menge $\{1, \dots, 7\}$ mit positiver Wahrscheinlichkeit an.

(b) Gilt für die gemeinsame Verteilung

$$P(X = k, Y = k + 1) = 1/4, \quad k \in \{0, 1, 2, 3\},$$

so folgt

$$P(X + Y = 1) = P(X = 0, Y = 1) = \frac{1}{4},$$

$$P(X + Y = 3) = P(X = 1, Y = 2) = \frac{1}{4},$$

$$P(X + Y = 5) = P(X = 2, Y = 3) = \frac{1}{4},$$

$$P(X + Y = 7) = P(X = 3, Y = 4) = \frac{1}{4}.$$

Die Summe $X + Y$ nimmt in diesem Fall nur Werte in der Menge $\{1, 3, 5, 7\}$ mit positiver Wahrscheinlichkeit an.

Lösung Aufgabe 22

(a) Es gilt

$$P(X = i) = \sum_{j=1}^3 P(X = i, Y = j) = p_{i1} + p_{i2} + p_{i3}, \quad i \in \{-1, 0, 1\},$$

$$P(Y = j) = \sum_{i=-1}^1 P(X = i, Y = j) = p_{-1j} + p_{0j} + p_{1j}, \quad j \in \{1, 2, 3\}.$$

Hiermit ergibt sich

$$\begin{aligned}P(X = -1) &= \frac{1}{20} + \frac{1}{5} + 0 = \frac{1}{4}, \\P(X = 0) &= \frac{1}{5} + \frac{1}{5} + \frac{1}{10} = \frac{1}{2}, \\P(X = 1) &= \frac{1}{10} + \frac{1}{10} + \frac{1}{20} = \frac{1}{4}, \\P(Y = 1) &= \frac{1}{20} + \frac{1}{5} + \frac{1}{10} = \frac{7}{20}, \\P(Y = 2) &= \frac{1}{5} + \frac{1}{5} + \frac{1}{10} = \frac{1}{2}, \\P(Y = 3) &= 0 + \frac{1}{10} + \frac{1}{20} = \frac{3}{20}.\end{aligned}$$

Die Zufallsvariablen X und Y sind nicht stochastisch unabhängig, da beispielsweise

$$P(X = -1, Y = 3) = 0 \neq \frac{1}{4} \cdot \frac{3}{20} = P(X = -1)P(Y = 3).$$

(b) Wegen

$$P(Y = j|X = 1) = \frac{P(X = 1, Y = j)}{P(X = 1)}, \quad j \in \{1, 2, 3\},$$

folgt aus der Aufgabenstellung und (a):

$$\begin{aligned}P(Y = 1|X = 1) &= \frac{\frac{1}{10}}{\frac{1}{4}} = \frac{2}{5}, \\P(Y = 2|X = 1) &= \frac{\frac{1}{10}}{\frac{1}{4}} = \frac{2}{5}, \\P(Y = 3|X = 1) &= \frac{\frac{1}{20}}{\frac{1}{4}} = \frac{1}{5}.\end{aligned}$$

(c) Es gilt

$$\begin{aligned}P(X = 0|Y \geq 2) &= \frac{P(X = 0, Y \geq 2)}{P(Y \geq 2)} = \frac{P(X = 0, Y = 2) + P(X = 0, Y = 3)}{P(Y = 2) + P(Y = 3)} \\&= \frac{\frac{1}{5} + \frac{1}{10}}{\frac{1}{2} + \frac{3}{20}} = \frac{6}{13}.\end{aligned}$$

(d) Die Verteilung der Zufallsvariablen $Z = X + Y$ hat den Träger $\{0, 1, 2, 3, 4\}$. Es gilt:

$$\begin{aligned}P(Z = 0) &= P(X = -1, Y = 1) = p_{-11} = \frac{1}{20}, \\P(Z = 1) &= P(X = -1, Y = 2) + P(X = 0, Y = 1) = p_{-12} + p_{01} = \frac{2}{5}, \\P(Z = 2) &= P(X = -1, Y = 3) + P(X = 0, Y = 2) + P(X = 1, Y = 1) \\&= p_{-13} + p_{02} + p_{11} = \frac{3}{10}, \\P(Z = 3) &= P(X = 0, Y = 3) + P(X = 1, Y = 2) = p_{03} + p_{12} = \frac{1}{5}, \\P(Z = 4) &= P(X = 1, Y = 3) = p_{13} = \frac{1}{20}.\end{aligned}$$

Einführung in die angewandte Stochastik

7. Übung

Aufgabe 23

Seien X und Y Zufallsvariablen auf einem Wahrscheinlichkeitsraum $(\Omega, \mathfrak{A}, P)$ mit $X \sim \Gamma(\alpha, \beta)$, $\alpha, \beta > 0$ und $Y \sim R(0, 1)$.

- (a) Zeigen Sie, dass für $a > 0$ gilt: $aX \sim \Gamma(\alpha/a, \beta)$.
- (b) Bestimmen Sie die Verteilung der Zufallsvariablen $Z = -\ln(1 - Y)$.

Aufgabe 24

Gegeben sei die Situation aus Aufgabe 22.

- (a) Berechnen Sie EX und EY .
- (b) Berechnen Sie $Var X$ und $Var Y$.
- (c) Berechnen Sie $Kov(X, Y)$ und $Korr(X, Y)$.

Aufgabe 25

Die Wahrscheinlichkeitsverteilung eines zweidimensionalen diskreten Zufallsvektors (X, Y) sei durch die folgende Tabelle der Wahrscheinlichkeiten

$$P(X = i, Y = j), \quad 1 \leq i \leq 2, \quad 1 \leq j \leq 5,$$

gegeben.

$X = i \backslash Y = j$	1	2	3	4	5
1	0	0,1	?	0,1	0,2
2	0,1	?	0,1	0,2	0,1
$P(Y = j)$	0,1	0,2	0,1	?	0,3

- (a) Vervollständigen Sie die Wahrscheinlichkeitstabelle.
- (b) Bestimmen Sie die Randverteilung von X . Sind X und Y stochastisch unabhängig?
- (c) Berechnen Sie EX , $Var X$ und $E\left(\frac{1}{X}\right)$.

Einführung in die angewandte Stochastik

Lösung zur 7. Übung

Lösung Aufgabe 23

- (i) Die Behauptung lässt sich analog zu Beispiel C 4.1(i) zeigen. Hier wird eine Herleitung mittels Satz C 4.2 gewählt.

Seien $a > 0$ und $X \sim \Gamma(\alpha, \beta)$, $\alpha, \beta > 0$. Betrachte $g : (0, \infty) \rightarrow (0, \infty)$ mit $g(x) = ax$ für $x > 0$. Die bijektive Funktion g erfüllt die Voraussetzungen von Satz C 4.2. Insbesondere gilt $g^{-1}(y) = \frac{1}{a}y$ und $(g^{-1})'(y) = \frac{1}{a}$ für $y > 0$. Es folgt für die Dichte der Zufallsvariable $Y = g(X) = aX$:

$$\begin{aligned} f^{aX}(y) &= f^Y(y) \\ &\stackrel{\text{C 4.2}}{=} |(g^{-1})'(y)| f^X(g^{-1}(y)) \\ &= \frac{1}{a} f^X\left(\frac{y}{a}\right) \\ &\stackrel{X \sim \Gamma(\alpha, \beta)}{=} \frac{1}{a} \frac{\alpha^\beta}{\Gamma(\beta)} \left(\frac{y}{a}\right)^{\beta-1} \exp\left(-\alpha \frac{y}{a}\right) \\ &= \frac{\left(\frac{\alpha}{a}\right)^\beta}{\Gamma(\beta)} y^{\beta-1} \exp\left(-\frac{\alpha}{a} y\right), \quad y > 0. \end{aligned}$$

Hierbei handelt es sich um die Dichte einer Gamma-Verteilung mit den Parametern α/a und β (vgl. B 3.9). Also gilt $aX \sim \Gamma(\alpha/a, \beta)$.

- (ii) Die Aussage lässt sich ebenfalls mit Satz C 4.2 zeigen. Hier wird ähnlich wie in Beispiel C 4.1(i) (bzw. mit Lemma C 2.7(iv)) argumentiert.

Betrachte $h : (0, 1) \rightarrow (0, \infty)$ mit $h(y) = -\ln(1-y)$ für $y \in (0, 1)$. Die Funktion h ist streng monoton wachsend und stetig. Die Umkehrfunktion von h ist $h^{-1}(z) = 1 - e^{-z}$, $z > 0$. Also gilt für die Verteilung der Zufallsvariable $Z = h(Y)$, falls $Y \sim R(0, 1)$ (vgl. B 3.6):

$$\begin{aligned} P(Z \leq z) &= P(h(Y) \leq z) = P(\ln(1-Y) \geq -z) \\ &= P(1-Y \geq e^{-z}) = P(Y \leq 1 - e^{-z}) \\ &= P(Y \leq h^{-1}(z)) \stackrel{Y \sim R(0,1)}{=} h^{-1}(z) \\ &= 1 - e^{-z}, \quad z > 0. \end{aligned}$$

Also folgt Z einer Exponentialverteilung $Exp(1)$. (Da h mit der Quantilfunktion der Verteilungsfunktion der Exponentialverteilung übereinstimmt, kann die Aussage auch direkt aus Lemma C 2.7(iv) gefolgert werden.)

Lösung Aufgabe 24

(a) Mit Hilfe der Resultate aus Aufgabe 22(a) ergibt sich:

$$E(X) = \sum_{i=-1}^1 i P(X = i) = -P(X = -1) + P(X = 1) = -\frac{1}{4} + \frac{1}{4} = 0,$$

$$E(Y) = \sum_{j=1}^3 j P(Y = j) = P(Y = 1) + 2P(Y = 2) + 3P(Y = 3) = \frac{7}{20} + 2 \cdot \frac{1}{2} + 3 \cdot \frac{3}{20} = \frac{9}{5}.$$

(b) Es gilt

$$E(X^2) = \sum_{i=-1}^1 i^2 P(X = i) = P(X = -1) + P(X = 1) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2},$$

$$\begin{aligned} E(Y^2) &= \sum_{j=1}^3 j^2 P(Y = j) = P(Y = 1) + 4P(Y = 2) + 9P(Y = 3) \\ &= \frac{7}{20} + 4 \cdot \frac{1}{2} + 9 \cdot \frac{3}{20} = \frac{37}{10}. \end{aligned}$$

Es folgt

$$\text{Var}(X) = E(X^2) - (E(X))^2 = \frac{1}{2} - 0^2 = \frac{1}{2},$$

$$\text{Var}(Y) = E(Y^2) - (E(Y))^2 = \frac{37}{10} - \left(\frac{9}{5}\right)^2 = \frac{23}{50}.$$

(c) Es gilt

$$\begin{aligned} E(X \cdot Y) &= \sum_{i=-1}^1 \sum_{j=1}^3 i j P(X = i, Y = j) \\ &= -P(X = -1, Y = 1) - 2P(X = -1, Y = 2) + P(X = 1, Y = 1) \\ &\quad + 2P(X = 1, Y = 2) + 3P(X = 1, Y = 3) \\ &= -\frac{1}{20} - 2 \cdot \frac{1}{5} + \frac{1}{10} + 2 \cdot \frac{1}{10} + 3 \cdot \frac{1}{20} = 0. \end{aligned}$$

Also folgt

$$\text{Kov}(X, Y) = E(X \cdot Y) - E(X)E(Y) = 0 - 0 \cdot \frac{9}{5} = 0$$

und damit auch $\text{Korr}(X, Y) = 0$ (vgl. C 5.17(ii)), d.h. X und Y sind unkorreliert.

Beachte: Laut Aufgabe 22(a) sind X und Y *nicht* stochastisch unabhängig. Dies ist ein weiteres Beispiel dafür, dass aus Unkorreliertheit nicht die stochastische Unabhängigkeit folgt (vgl. auch C 5.20).

Lösung Aufgabe 25

(a) Die vollständig ausgefüllte Tabelle sieht folgendermaßen aus:

$X = i \backslash Y = j$	1	2	3	4	5
1	0	0,1	0	0,1	0,2
2	0,1	0,1	0,1	0,2	0,1
$P(Y = j)$	0,1	0,2	0,1	0,3	0,3

(b) Es gilt

$$P(X = i) = \sum_{j=1}^5 P(X = i, Y = j), \quad 1 \leq i \leq 2.$$

Also ergibt sich mit den Angaben der Tabelle aus (a):

$$P(X = 1) = 0 + 0,1 + 0 + 0,1 + 0,2 = 0,4,$$

$$P(X = 2) = 0,1 + 0,1 + 0,1 + 0,2 + 0,1 = 0,6.$$

Die Zufallsvariablen X und Y sind nicht stochastisch unabhängig, denn

$$P(X = 1, Y = 1) = 0 \neq 0,4 \cdot 0,1 = P(X = 1)P(Y = 1).$$

(c) Es gilt

$$E(X) = P(X = 1) + 2P(X = 2) = 0,4 + 2 \cdot 0,6 = 1,6,$$

$$E(X^2) = P(X = 1) + 2^2P(X = 2) = 0,4 + 4 \cdot 0,6 = 2,8,$$

$$Var(X) = E(X^2) - (E(X))^2 = 2,8 - (1,6)^2 = 0,24.$$

Außerdem gilt

$$E\left(\frac{1}{X}\right) = \sum_{i=1}^2 \frac{1}{i} P(X = i) = P(X = 1) + \frac{1}{2} \cdot P(X = 2) = 0,4 + \frac{1}{2} \cdot 0,6 = 0,7.$$

Einführung in die angewandte Stochastik

8. Übung

Aufgabe 26

Die Riemann-Dichte $f^{X,Y} : \mathbb{R}^2 \rightarrow \mathbb{R}$ der zweidimensionalen stetigen Zufallsvariablen (X, Y) sei gegeben durch:

$$f^{X,Y}(x, y) = \begin{cases} 12(1-x)xy & , \text{ falls } x, y \in [0, 1] , \\ 0 & , \text{ sonst .} \end{cases}$$

- (a) Bestimmen Sie die zugehörigen Randdichten f^X und f^Y .
- (b) Berechnen Sie EX , EY , $Var X$ und $Var Y$.
- (c) Berechnen Sie $Kov(X, Y)$. Sind die Zufallsvariablen X und Y stochastisch unabhängig?

Aufgabe 27

Seien X und Y unkorrelierte Zufallsvariablen mit $X \sim R(-2, 4)$ und $Y \sim \text{Exp}(\frac{1}{2})$. Weiter seien $U = X - 2Y$ und $V = 3X + Y - 5$.

- (a) Berechnen Sie EU und EV .
- (b) Berechnen Sie $Var U$ und $Var V$.
- (c) Berechnen Sie $Kov(U, X)$.

Hinweis: Sie können die Resultate aus den Beispielen C 5.2 und C 5.13 verwenden.

Aufgabe 28

Sei X eine Zufallsvariable auf dem Wahrscheinlichkeitsraum $(\Omega, \mathfrak{A}, P)$ mit

$$EX = 1 \quad \text{und} \quad Var(X) = 2.$$

Zeigen Sie: $P(X \in (-1, 3)) \geq \frac{1}{2}$.

Aufgabe 29

Sei $X \sim \text{po}(\lambda)$ mit $\lambda > 0$.

- (a) Berechnen Sie die zugehörige wahrscheinlichkeitserzeugende Funktion g_X .
- (b) Berechnen Sie mit Hilfe der in (a) bestimmten erzeugenden Funktion g_X den Erwartungswert EX und die Varianz $Var(X)$ der Zufallsvariablen X .
- (c) Sei $Y \sim \text{po}(\mu)$ mit $\mu > 0$ stochastisch unabhängig von X . Bestimmen Sie die Verteilung von $X + Y$ mit Hilfe von Aufgabenteil (a).

Einführung in die angewandte Stochastik

Lösung zur 8. Übung

Lösung Aufgabe 26

(a) Es gilt

$$\begin{aligned}f^X(x) &= \int_{-\infty}^{\infty} f^{X,Y}(x,y) dy = \int_0^1 12(1-x)xy dy \\&= 12(1-x)x \left(\frac{1}{2}y^2 \right) \Big|_0^1 = 6(1-x)x \quad \text{für } x \in [0,1], \\f^X(x) &= 0 \quad \text{für } x \notin [0,1], \\f^Y(y) &= \int_{-\infty}^{\infty} f^{X,Y}(x,y) dx = \int_0^1 12(1-x)xy dx \\&= 12y \left(\frac{1}{2}x^2 - \frac{1}{3}x^3 \right) \Big|_0^1 = 12y \left(\frac{1}{2} - \frac{1}{3} \right) = 2y \quad \text{für } y \in [0,1], \\f^Y(y) &= 0 \quad \text{für } y \notin [0,1].\end{aligned}$$

(b) Es gilt

$$\begin{aligned}E(X) &= \int_{-\infty}^{\infty} xf^X(x) dx = \int_0^1 6(1-x)x^2 dx = 6 \left(\frac{1}{3}x^3 - \frac{1}{4}x^4 \right) \Big|_0^1 = \frac{1}{2}, \\E(Y) &= \int_{-\infty}^{\infty} yf^Y(y) dy = \int_0^1 2y^2 dy = 2 \left(\frac{1}{3}y^3 \right) \Big|_0^1 = \frac{2}{3}.\end{aligned}$$

Analog ergibt sich

$$E(X^2) = \int_{-\infty}^{\infty} x^2 f^X(x) dx = \frac{3}{10}, \quad E(Y^2) = \int_{-\infty}^{\infty} y^2 f^Y(y) dy = \frac{1}{2},$$

und daraus folgt

$$\begin{aligned}\text{Var}(X) &= E(X^2) - (E(X))^2 = \frac{3}{10} - \frac{1}{4} = \frac{1}{20}, \\ \text{Var}(Y) &= E(Y^2) - (E(Y))^2 = \frac{1}{2} - \frac{4}{9} = \frac{1}{18}.\end{aligned}$$

(c) Es gilt

$$\begin{aligned}E(XY) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f^{X,Y}(x,y) dx dy \\&= \int_0^1 \int_0^1 12(1-x)x^2 y^2 dx dy = \int_0^1 12y^2 \left(\frac{1}{3}x^3 - \frac{1}{4}x^4 \right) \Big|_0^1 dy \\&= \int_0^1 y^2 dy = \frac{1}{3}.\end{aligned}$$

Daher folgt

$$\text{Kov}(X,Y) = E(XY) - E(X)E(Y) = \frac{1}{3} - \frac{1}{2} \cdot \frac{2}{3} = 0.$$

Die Zufallsvariablen X und Y sind stochastisch unabhängig, denn es gilt (siehe (a))

$$f^{X,Y}(x, y) = 6(1-x)x \cdot 2y = f^X(x)f^Y(y) \quad \text{für } x, y \in [0, 1],$$

und

$$f^{X,Y}(x, y) = 0 = f^X(x)f^Y(y), \quad \text{falls } x \notin [0, 1] \text{ oder } y \notin [0, 1].$$

Also ist die Bedingung aus Satz C 3.14 erfüllt.

Beachte: Aus der Tatsache $Kov(X, Y) = 0$ kann *nicht* auf die stochastische Unabhängigkeit von X und Y geschlossen werden (vgl. C 5.20 und Aufgabe 24(c))!

Lösung Aufgabe 27

Laut C 5.2 und C 5.13 gilt für $X \sim R(-2, 4)$ und $Y \sim \text{Exp}(\frac{1}{2})$:

$$\begin{aligned} EX &= \frac{-2+4}{2} = 1, & Var X &= \frac{(4-(-2))^2}{12} = \frac{36}{12} = 3, \\ EY &= \frac{1}{\frac{1}{2}} = 2, & Var Y &= \frac{1}{(\frac{1}{2})^2} = 4. \end{aligned}$$

Laut Aufgabenstellung sind X und Y unkorreliert. Unter Anwendung von Eigenschaften aus C 5.7, C 5.12 und C 5.18 folgt:

(a)

$$\begin{aligned} EU &= E(X - 2Y) = EX - 2EY = 1 - 4 = -3, \\ EV &= E(3X + Y - 5) = 3EX + EY - 5 = 3 + 2 - 5 = 0. \end{aligned}$$

(b)

$$\begin{aligned} Var U &= Var(X - 2Y) \stackrel{X,Y \text{ unkor.}}{=} 1^2 \cdot Var X + (-2)^2 \cdot Var Y \\ &= Var X + 4 \cdot Var Y = 3 + 16 = 19, \\ Var V &= Var(3X + Y - 5) = Var(3X + Y) \\ &\stackrel{X,Y \text{ unkor.}}{=} 3^2 \cdot Var X + 1^2 \cdot Var Y = 9 Var X + Var Y \\ &= 27 + 4 = 31. \end{aligned}$$

(c) Wegen $Kov(X, Y) = E(XY) - E(X)E(Y) = 0$ gilt $E(XY) = E(X)E(Y)$. Damit folgt

$$\begin{aligned} Kov(U, X) &= Kov(X - 2Y, X) \\ &= E((X - 2Y)X) - E(X - 2Y)E(X) \\ &= E(X^2) - 2E(XY) - (E(X) - 2E(Y))E(X) \\ &= E(X^2) - 2E(X)E(Y) - (E(X) - 2E(Y))E(X) \\ &= E(X^2) - (E(X))^2 = Var(X) = 3. \end{aligned}$$

Bemerkung:

Allgemein kann für Zufallsvariablen $X_1, \dots, X_n$, $Y_1, \dots, Y_m$ und $a_1, \dots, a_n, b_1, \dots, b_m \in \mathbb{R}$ gezeigt werden:

$$Kov\left(\sum_{i=1}^n a_i X_i, \sum_{j=1}^m b_j Y_j\right) = \sum_{i=1}^n \sum_{j=1}^m a_i b_j Kov(X_i, Y_j),$$

falls die einzelnen Kovarianzen alle (als reelle Zahlen) existieren.

Hiermit lässt sich Teil (c) auch folgendermaßen lösen:

$$Kov(U, X) = Kov(X - 2Y, X) \stackrel{\text{Bem.}}{=} \underbrace{Kov(X, X)}_{=Var X} - 2 \cdot \underbrace{Kov(Y, X)}_{=0, \text{ da } X, Y \text{ unkor.}} = Var X = 3.$$

Lösung Aufgabe 28

Mit Hilfe der Tschebyscheff-Ungleichung gilt (vgl. Satz C 5.22.(iii)):

$$\begin{aligned} P(X \in (-1, 3)) &= P(-1 < X < 3) \\ &= P(-2 < X - EX < 2) \\ &= P(|X - EX| < 2) \\ &= 1 - P(|X - EX| \geq 2) \\ &\geq 1 - \frac{Var(X)}{2^2} \\ &= 1 - \frac{2}{4} \\ &= \frac{1}{2}. \end{aligned}$$

Lösung Aufgabe 29

Sei $X \sim \text{po}(\lambda)$ mit $\lambda > 0$, d.h. $P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}$, $k \in \mathbb{N}_0$.

(a) Für die erzeugenden Funktion erhält man:

$$\begin{aligned} g_X(t) &= \sum_{k=0}^{\infty} P(X = k) \cdot t^k = \sum_{k=0}^{\infty} e^{-\lambda} \frac{\lambda^k}{k!} \cdot t^k = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda t)^k}{k!} \\ &= e^{-\lambda} \cdot e^{\lambda t} = e^{\lambda(t-1)}, \quad t \in \mathbb{R}. \end{aligned}$$

(b) Für $k \in \mathbb{N}$ gilt:

$$g_X^{(k)}(1) = E(X(X-1) \cdot \dots \cdot (X-k+1)).$$

Mit (a) folgern wir:

$$\begin{aligned} g_X^{(1)}(t) &= \lambda e^{\lambda(t-1)} & \text{und damit} & \quad EX = g_X^{(1)}(1) = \lambda, \\ g_X^{(2)}(t) &= \lambda^2 e^{\lambda(t-1)} & \text{und damit} & \quad E(X(X-1)) = \lambda^2. \end{aligned}$$

Weiter folgt für die Varianz von X :

$$\begin{aligned} Var(X) &= E(X^2) - (EX)^2 \\ &= E(X(X-1)) + E(X) - (EX)^2 \\ &= g_X^{(2)}(1) + g_X^{(1)}(1) - \left(g_X^{(1)}(1)\right)^2 \\ &= \lambda^2 + \lambda - \lambda^2 = \lambda. \end{aligned}$$

(c) Für stochastisch unabhängige Zufallsvariablen X und Y gilt:

$$g_{X+Y}(t) = g_X(t) \cdot g_Y(t) = e^{\lambda(t-1)} \cdot e^{\mu(t-1)} = e^{(\lambda+\mu)(t-1)}.$$

Damit folgt mit Bemerkung C 6.4.(i):

$$X + Y \sim \text{po}(\lambda + \mu).$$

Einführung in die angewandte Stochastik

9. Übung

Aufgabe 30

Seien X und Y zwei Zufallsvariablen auf einem Wahrscheinlichkeitsraum $(\Omega, \mathfrak{A}, P)$ mit gemeinsamer Riemannndichtefunktion

$$f^{X,Y}(x, y) = \begin{cases} \frac{y^s e^{-(x+1)y}}{(s-1)!} & , \text{ falls } x, y > 0, \\ 0 & , \text{ sonst,} \end{cases}$$

wobei $s \in \mathbb{N}$ gelte.

- (a) Bestimmen Sie die Randdichte f^Y von Y und geben Sie die Verteilung von Y an.
- (b) Sei $y > 0$ gegeben. Bestimmen Sie die bedingte Dichte $f^{X|Y=y}$ von X unter (der Hypothese) $Y = y$. Um die Dichte welcher Verteilung handelt es sich?
- (c) Berechnen Sie für $y > 0$ die bedingte Erwartung $E(X|Y = y)$ von X unter (der Hypothese) $Y = y$.

Aufgabe 31

Eine Bank betreibt in einer Region insgesamt 200 Geldautomaten, von denen jeder (unabhängig von den übrigen Automaten) mit Wahrscheinlichkeit 0,05 aufgrund einer Störung innerhalb einer Woche mindestens einmal ausfällt.

Für die Einrichtung eines ständigen Wartungsdienstes sei die Wahrscheinlichkeit von Interesse, dass die Anzahl der Geldautomaten, die in einer Woche mindestens eine derartige Störung aufweisen, mindestens 5 und höchstens 15 beträgt.

- (a) Schätzen Sie die gesuchte Wahrscheinlichkeit mittels der Tschebyscheff-Ungleichung nach unten ab.
- (b) Für eine Folge $(X_n)_{n \in \mathbb{N}}$ von iid Zufallsvariablen mit $EX_1 = \mu$ und $Var X_1 = \sigma^2 < \infty$ gilt:

$$\lim_{n \rightarrow \infty} P\left(\sqrt{n} \frac{\bar{X}_n - \mu}{\sigma} \leq t\right) = \Phi(t)$$

wobei $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$ und Φ die Verteilungsfunktion der Standardnormalverteilung. Diese Aussage ist bekannt als Zentraler Grenzwertsatz. Man sagt $\sqrt{n} \frac{\bar{X}_n - \mu}{\sigma}$ hat asymptotisch eine $N(0, 1)$ -Verteilung. Berechnen Sie nun die gesuchte Wahrscheinlichkeit approximativ mit Hilfe des Zentralen Grenzwertsatzes.

Hinweis:

Sie können Funktionswerte und Quantile der Standardnormalverteilung aus Tabellen in der Literatur entnehmen (siehe z.B. Formelsammlung *Statistik griffbereit*).

Aufgabe 32

Die Zufallsvariablen $X_1, \dots, X_n$ seien stochastisch unabhängig und jeweils auf dem Intervall $[0, b]$ Rechteck-verteilt mit unbekannter oberer Intervallgrenze $b > 0$. Um b aus zugehörigen Beobachtungen $x_1, \dots, x_n$ zu schätzen, betrachtet man folgende Schätzfunktion:

$$\hat{b}_n = 2 \bar{X}_n = \frac{2}{n} \sum_{i=1}^n X_i .$$

- (a) Ist $\hat{b}_n$ ein erwartungstreuer Schätzer für den Parameter b ?
- (b) Berechnen Sie die Varianz von $\hat{b}_n$.

Einführung in die angewandte Stochastik

Lösung zur 9. Übung

Lösung Aufgabe 30

(a) Für $y > 0$ gilt:

$$\begin{aligned} f^Y(y) &= \int_{-\infty}^{\infty} f^{X,Y}(x,y) dx = \int_0^{\infty} \frac{y^s e^{-(x+1)y}}{(s-1)!} dx = \frac{y^s e^{-y}}{(s-1)!} \lim_{a \rightarrow \infty} \int_0^a e^{-xy} dx \\ &= \frac{y^s e^{-y}}{(s-1)!} \lim_{a \rightarrow \infty} \left[-\frac{e^{-xy}}{y} \right]_{x=0}^{x=a} = \frac{y^s e^{-y}}{(s-1)!} \left(-\lim_{a \rightarrow \infty} \frac{e^{-ay}}{y} + \frac{1}{y} \right) = \frac{y^{s-1} e^{-y}}{(s-1)!}. \end{aligned}$$

Für $y \leq 0$ erhalten wir

$$f^Y(y) = \int_{-\infty}^{\infty} \underbrace{f^{X,Y}(x,y)}_{=0 \text{ nach Vor.}} dx = \int_{-\infty}^{\infty} 0 dx = 0.$$

Y besitzt somit eine Gammaverteilung $\Gamma(1, s)$ mit Parametern 1 und s (beachte: $\Gamma(s) = (s-1)!$, $s \in \mathbb{N}$).

(b) Für gegebenes $y > 0$ gilt nach (a) $f^Y(y) > 0$ und damit (vgl. Def. C 7.1) für $x > 0$

$$f^{X|Y=y}(x) = \frac{f^{X,Y}(x,y)}{f^Y(y)} \stackrel{(a), Vor.}{=} \frac{\frac{y^s e^{-(x+1)y}}{(s-1)!}}{\frac{y^{s-1} e^{-y}}{(s-1)!}} = y e^{-yx}.$$

Für $x \leq 0$ erhalten wir

$$f^{X|Y=y}(x) = \frac{\overbrace{f^{X,Y}(x,y)}^{=0 \text{ nach Vor.}}}{f^Y(y)} = 0.$$

Folglich handelt es sich bei $f^{X|Y=y}$ um die Dichte der Exponentialverteilung $Exp(y)$ mit Parameter y .

(c) Mit C 7.4 und der Regel der partiellen Integration erhalten wir für gegebenes $y > 0$:

$$\begin{aligned} E[X|Y=y] &= \int x f^{X|Y=y}(x) dx \stackrel{(b)}{=} \lim_{a \rightarrow \infty} \int_0^a x y e^{-xy} dx \\ &= \lim_{a \rightarrow \infty} [x \cdot (-e^{-xy})]_{x=0}^{x=a} + \lim_{a \rightarrow \infty} \int_0^a e^{-xy} dx \\ &= \left(\lim_{a \rightarrow \infty} (-a e^{-ya}) + 0 \right) + \lim_{a \rightarrow \infty} \int_0^a e^{-xy} dx \\ &= \lim_{a \rightarrow \infty} \left[-\frac{e^{-xy}}{y} \right]_{x=0}^{x=a} = \lim_{a \rightarrow \infty} \frac{1}{y} (1 - e^{-ay}) = \frac{1}{y}. \end{aligned}$$

Lösung Aufgabe 31

Zum Modell: Die Störungen werden beschrieben durch Zufallsvariablen $X_1, \dots, X_{200}$ mit der Interpretation

$$X_i = \begin{cases} 1, & i\text{-ter Automat fällt mindestens einmal aus,} \\ 0, & \text{sonst,} \end{cases} \quad i = 1, \dots, 200.$$

Gemäß Aufgabenstellung sind $X_1, \dots, X_{200}$ stochastisch unabhängig mit $X_i \sim \text{bin}(1, 0,05), i = 1, \dots, 200$. Also gilt gemäß C 1.14(i):

$$S_{200} = \sum_{i=1}^{200} X_i \sim \text{bin}(200, 0,05).$$

Mit $n = 200$, $p = 0,05 = 1/20$ folgt gemäß C 5.2 und C 5.13:

$$E(S_n) = np = 200 \cdot \frac{1}{20} = 10, \quad \text{Var}(S_n) = np(1-p) = 200 \cdot \frac{1}{20} \cdot \frac{19}{20} = 9,5.$$

Damit lassen sich nun beide Aufgabenteile lösen:

(a) Zunächst gilt

$$\begin{aligned} P(5 \leq S_n \leq 15) &= P(-5 \leq S_n - 10 \leq 5) \\ &= P(|S_n - E(S_n)| \leq 5) \\ &\geq P(|S_n - E(S_n)| < 5) \\ &= 1 - P(|S_n - E(S_n)| \geq 5). \end{aligned}$$

Aus der Tschebyscheff-Ungleichung (vgl. C 5.22(iii)) folgt dann

$$P(|S_n - E(S_n)| \geq 5) \leq \frac{\text{Var}(S_n)}{5^2}$$

und daher

$$\begin{aligned} P(5 \leq S_n \leq 15) &\geq 1 - P(|S_n - E(S_n)| \geq 5) \\ &\geq 1 - \frac{\text{Var}(S_n)}{5^2} \\ &= 1 - \frac{9,5}{25} = 0,62. \end{aligned}$$

(b) Beachte, dass mit den obigen Bezeichnungen

$$\sqrt{n} \frac{\bar{X}_n - \mu}{\sigma} = \frac{S_n - E(S_n)}{\sqrt{\text{Var}(S_n)}}.$$

Daher gilt

$$\begin{aligned} P(5 \leq S_n \leq 15) &\stackrel{S_n \in \mathbb{N}_0}{=} P(S_n \leq 15) - P(S_n \leq 4) \\ &\stackrel{\text{Standardisierung}}{=} P\left(\frac{S_n - 10}{\sqrt{9,5}} \leq \frac{15 - 10}{\sqrt{9,5}}\right) - P\left(\frac{S_n - 10}{\sqrt{9,5}} \leq \frac{4 - 10}{\sqrt{9,5}}\right) \end{aligned}$$

Der Zentrale Grenzwertsatz (siehe C 8.6) liefert nun

$$P\left(\frac{S_n - 10}{\sqrt{9,5}} \leq t\right) \approx \Phi(t), \quad t \in \mathbb{R}.$$

Damit folgt

$$\begin{aligned}P(5 \leq S_n \leq 15) &\approx \Phi\left(\frac{15-10}{\sqrt{9,5}}\right) - \Phi\left(\frac{4-10}{\sqrt{9,5}}\right) \\&\approx \Phi(1,62) - \Phi(-1,95) \\&= \Phi(1,62) - (1 - \Phi(1,95)).\end{aligned}$$

Eine Tabelle der Funktionswerte der Standardnormalverteilung (siehe z.B. Formelsammlung *Statistik griffbereit*, 4. Auflage, S. 33) liefert $\Phi(1,62) \approx 0,947$ und $\Phi(1,95) \approx 0,974$. Damit folgt schließlich

$$P(5 \leq S_n \leq 15) \approx \Phi(1,62) - (1 - \Phi(1,95)) \approx 0,947 - 1 + 0,974 = 0,921.$$

Lösung Aufgabe 33

Seien $X_1, \dots, X_n$ unabhängig und identisch verteilt mit $X_1 \sim R(0, b)$ mit $b > 0$.

(a) Laut C 5.2(iii) gilt $EX_i = \frac{b}{2}$, $i = 1, \dots, n$. Hiermit folgt für $b > 0$ und $n \in \mathbb{N}$:

$$\begin{aligned}E_b \hat{b}_n &= E_b \left(\frac{2}{n} \sum_{i=1}^n X_i \right) \\&= \frac{2}{n} \sum_{i=1}^n \underbrace{EX_i}_{=b/2} \\&= \frac{2}{n} n \frac{b}{2} = b.\end{aligned}$$

Somit ist $\hat{b}_n$ erwartungstreu für b .

(b) Laut C 5.13(iii) gilt $Var(X_i) = \frac{b^2}{12}$, $i = 1, \dots, n$. Hiermit folgt für $b > 0$ und $n \in \mathbb{N}$:

$$\begin{aligned}Var_b \hat{b}_n &= Var_b \left(\frac{2}{n} \sum_{i=1}^n X_i \right) = \frac{4}{n^2} \sum_{i=1}^n \underbrace{Var X_i}_{b^2/12} \\&= \frac{4}{n^2} n \frac{b^2}{12} = \frac{b^2}{3n}.\end{aligned}$$

Einführung in die angewandte Stochastik

10. Übung

Aufgabe 33

Die Zufallsvariablen $X_1, \dots, X_n$ seien stochastisch unabhängig und jeweils Pareto-verteilt mit (unbekanntem) Parameter $\alpha > 0$. Gemäß Vorlesung sind dann die zugehörige Dichtefunktion f_α und die zugehörige Verteilungsfunktion F_α der Zufallsvariablen X_i für $i \in \{1, \dots, n\}$ in Abhängigkeit vom Parameter α gegeben durch

$$f_\alpha(x) = \begin{cases} \frac{\alpha}{x^{\alpha+1}} & , \quad x \geq 1, \\ 0 & , \quad x < 1 \end{cases} \quad \text{bzw.} \quad F_\alpha(x) = \begin{cases} 1 - \frac{1}{x^\alpha} & , \quad x \geq 1, \\ 0 & , \quad x < 1. \end{cases}$$

- (a) Bestimmen Sie eine Maximum-Likelihood-Schätzung $\hat{\alpha}$ für den Parameter α .
(b) Berechnen Sie den aus (a) resultierenden Schätzwert zu folgenden Daten:

1.65 , 2.97 , 3.52 , 3.07 , 2.06 , 2.41 , 3.84 , 5.48 , 1.86 , 4.90 , 7.36 , 1.58 .

- (c) Berechnen Sie die Wahrscheinlichkeit dafür, dass X_1 den Wert $x = 2$ übersteigt, indem Sie den unbekannten Parameter α durch den in (b) berechneten Schätzwert ersetzen.

Aufgabe 34

Eine Fluggesellschaft möchte wissen, wie hoch der Anteil p der Passagiere ist, die ihren Flug nicht antreten. Hierzu soll ein Konfidenzintervall für p bestimmt werden.

Die Überprüfung von 1000 zufällig ausgewählten Passagieren ergibt, dass 74 von ihnen den Flug nicht angetreten haben. Bestimmen Sie anhand dieses Ergebnisses ein approximatives zweiseitiges Konfidenzintervall für p zum Konfidenzniveau 90%.

Aufgabe 35

Auf 12 Versuchsflächen wurde eine neue Weizensorte angebaut. Die einzelnen Flächen erbrachten die folgenden Hektarerträge (in t):

3.56 , 3.37 , 3.78 , 3.12 , 3.72 , 3.41 , 3.56 , 3.66 , 3.71 , 3.49 , 3.56 , 3.40 .

Aus Erfahrung ist bekannt, dass diese Hektarerträge als Realisationen stochastisch unabhängiger, jeweils $N(\mu, \sigma^2)$ -verteilter Zufallsvariablen mit $\mu \in \mathbb{R}$ und $\sigma > 0$ angesehen werden können.

- (a) Ermitteln Sie ein zweiseitiges Konfidenzintervall für den Erwartungswert μ zum Konfidenzniveau 0.9 bei
- (i) bekannter Varianz $\sigma^2 = 0.0324$,
 - (ii) unbekannter Varianz σ^2 .

- (b) Ermitteln Sie ein einseitiges unteres 99%–Konfidenzintervall für
- die Varianz σ^2 ,
 - die Standardabweichung σ .
- (c) Wie groß müsste die Anzahl der Versuchsflächen mindestens sein, um bei bekannter Varianz $\sigma^2 = 0.0324$ ein zweiseitiges 90%–Konfidenzintervall für den (unbekannten) Erwartungswert μ angeben zu können, dessen Länge höchstens 0.1 (t) beträgt?

Aufgabe 36

Für Geschwindigkeitsmessungen im Straßenverkehr stehen ein mobiles Radarsystem S_1 und eine festinstallierte (geeichte) Messanlage S_2 zur Verfügung. Durch Vergleich der Messergebnisse soll die Tauglichkeit der mobilen Anlage S_1 für den alltäglichen Einsatz überprüft werden.

Hierzu wurde mit beiden Systemen jeweils mehrfach die Geschwindigkeit eines Fahrzeugs gemessen, dass mit konstanter Geschwindigkeit von 50 km/h den Messpunkt passierte. Es wurden 12 Messungen mit dem mobilen System S_1 und 8 Messungen mit dem System S_2 durchgeführt. Die zugehörigen gemessenen Geschwindigkeiten (in km/h) sind in der folgenden Tabelle angegeben:

Nr. der Messung	1	2	3	4	5	6	7	8	9	10	11	12
Messanlage S_1	52.3	53.9	51.0	49.3	49.9	47.6	49.2	50.9	48.1	50.3	53.9	48.4
Messanlage S_2	48.1	50.8	53.3	52.0	49.9	48.4	52.1	50.2				

Nehmen Sie an, dass die mit System S_1 bzw. S_2 gemessenen Geschwindigkeiten $x_1, \dots, x_{12}$ und $y_1, \dots, y_8$ als Realisationen stochastisch unabhängiger Zufallsvariablen $X_1, \dots, X_{12}, Y_1, \dots, Y_8$ mit $X_i \sim N(\mu_1, \sigma^2)$ und $Y_j \sim N(\mu_2, \sigma^2)$ für $i \in \{1, \dots, 12\}$, $j \in \{1, \dots, 8\}$ aufgefasst werden können, wobei $\mu_1, \mu_2 \in \mathbb{R}$ und $\sigma^2 > 0$ jeweils unbekannt seien.

Geben Sie ein zweiseitiges Konfidenzintervall für die Differenz $d = \mu_1 - \mu_2$ der erwarteten Geschwindigkeitsmessungen μ_1 und μ_2 zum Konfidenzniveau 0.95 an.

Aufgabe 37

In einer Supermarktfiliale werden innerhalb eines festgelegten Zeitraumes jeweils die Angebotspreise $x_1, \dots, x_7$ und die zugehörigen Absatzmengen $y_1, \dots, y_7$ eines bestimmten Artikels notiert:

i	1	2	3	4	5	6	7
Preis x_i (in €/kg)	2.49	2.68	2.62	2.51	2.84	2.65	2.76
Absatzmenge y_i (in 100 kg)	6.7	6.2	5.8	6.2	5.4	6.5	5.9

Es sei das übliche Modell der linearen Regression zugrunde gelegt, d.h.

$$Y_i = a + bx_i + \varepsilon_i, \quad i = 1, \dots, 7.$$

Hierbei seien die Fehler $\varepsilon_1, \dots, \varepsilon_7$ stochastisch unabhängig, jeweils $N(0, \sigma^2)$ -verteilt mit unbekannter Varianz σ^2 .

- Schätzen Sie die Regressionsgerade $f(x) = a + bx$. Geben Sie insbesondere eine Schätzung an für die Absatzmenge bei einem Angebotspreis von 2,70 €/kg.
- Bestimmen Sie zum Konfidenzniveau $1 - \alpha = 0.9$ ein zweiseitiges Konfidenzintervall für die Varianz σ^2 .

Einführung in die angewandte Stochastik

Lösung zur 10. Übung

Lösung Aufgabe 33

(a) Seien $x_1, \dots, x_n$ Realisationen von $X_1, \dots, X_n$ mit $x_i \geq 1$ für alle $i \in \{1, \dots, n\}$ (sonst gilt $f_\alpha(x_i) = 0$). Dann ist die zugehörige Likelihood-Funktion aufgrund der stochastischen Unabhängigkeit der Zufallsvariablen $X_1, \dots, X_n$ gegeben durch

$$L(\alpha|x_1, \dots, x_n) = \prod_{i=1}^n f_\alpha(x_i) = \prod_{i=1}^n \frac{\alpha}{x_i^{\alpha+1}} = \alpha^n \prod_{i=1}^n x_i^{-(\alpha+1)}, \quad \alpha > 0.$$

Betrachte Maximierung der Log-Likelihood-Funktion:

$$\begin{aligned} l(\alpha) &= \ln(L(\alpha|x_1, \dots, x_n)) \\ &= \ln(\alpha^n) + \ln\left(\prod_{i=1}^n x_i^{-(\alpha+1)}\right) \\ &= n \ln(\alpha) + \sum_{i=1}^n (-(\alpha+1)) \ln(x_i) = n \ln(\alpha) - (\alpha+1) \sum_{i=1}^n \ln(x_i), \quad \alpha > 0. \end{aligned}$$

Die Funktion l ist differenzierbar auf $(0, \infty)$ mit

$$l'(\alpha) = \frac{n}{\alpha} - \sum_{i=1}^n \ln(x_i), \quad \alpha > 0.$$

Also gilt

$$l'(\alpha) \begin{cases} > \\ = \\ < \end{cases} 0 \iff \alpha \begin{cases} < \\ = \\ > \end{cases} \frac{n}{\sum_{i=1}^n \ln(x_i)}$$

für $\alpha > 0$. Für

$$\hat{\alpha} = \frac{n}{\sum_{i=1}^n \ln(x_i)}$$

gilt daher: l ist streng monoton wachsend auf $(0, \hat{\alpha}]$ und streng monoton fallend auf $[\hat{\alpha}, \infty)$. Damit ist $\hat{\alpha}$ die (eindeutige) ML-Schätzung für α (bei geg. $x_1, \dots, x_n$).

(b) Einsetzen der $n = 12$ gegebenen Datenwerte liefert den folgenden Schätzwert für α :

$$\hat{\alpha} = \frac{12}{\sum_{i=1}^{12} \ln(x_i)} \approx 0,903.$$

(c) Ersetzen von α durch den Schätzwert $\hat{\alpha} \approx 0.903$ liefert:

$$P_{\hat{\alpha}}(X_1 > 2) = 1 - F_{\hat{\alpha}}(2) \approx 1 - \left(1 - \frac{1}{2^{0,903}}\right) = \frac{1}{2^{0,903}} \approx 0,535.$$

Lösung Aufgabe 34

Zur Beschreibung der Situation betrachte stochastisch unabhängige Zufallsvariablen $X_1, \dots, X_n$, $n = 1000$, mit der Interpretation

$$X_i = \begin{cases} 1, & i\text{-ter Passagier tritt Flug nicht an,} \\ 0, & i\text{-ter Passagier tritt Flug an.} \end{cases}$$

Dann gilt $X_i \sim \text{bin}(1, p)$, $i = 1, \dots, n$ mit $p \in (0, 1)$ unbekannt.

Gemäß Vorlesung ist in dieser Situation ein approximatives zweiseitiges $(1 - \alpha)$ -Konfidenzintervall für p gegeben durch

$$K = \left[\hat{p} - \frac{u_{1-\alpha/2}}{\sqrt{n}} \sqrt{\hat{p}(1-\hat{p})}, \hat{p} + \frac{u_{1-\alpha/2}}{\sqrt{n}} \sqrt{\hat{p}(1-\hat{p})} \right].$$

Hier bezeichnet $u_{1-\alpha/2}$ das $(1 - \alpha/2)$ -Quantil der Standardnormalverteilung $N(0, 1)$ und $\hat{p} = \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ die übliche erwartungstreue Punktschätzung für p .

Einsetzen der vorgegebenen Zahlenwerte führt zum gewünschten Konfidenzintervall: Es gilt

$$n = 1000, \quad \hat{p} = \bar{x} = \frac{74}{1000} = 0.074, \quad 1 - \alpha = 0.9.$$

Insbesondere folgt $\alpha = 0.1$ und damit $1 - \alpha/2 = 0.95$. Einer Tabelle mit Quantilen der Standardnormalverteilung lässt sich entnehmen:

$$u_{1-\alpha/2} = u_{0.95} \approx 1.64.$$

Damit ergibt sich für das gesuchte Konfidenzintervall:

$$\begin{aligned} K &\approx \left[0.074 - \frac{1.64}{\sqrt{1000}} \sqrt{0.074 \cdot 0.926}, 0.074 + \frac{1.64}{\sqrt{1000}} \sqrt{0.074 \cdot 0.926} \right] \\ &\approx [0.060, 0.088]. \end{aligned}$$

Lösung Aufgabe 35

- (a) (i) Gemäß D 5.6 ist ein zweiseitiges Konfidenzintervall für μ zum Konfidenz-Niveau $1 - \alpha$ bei bekannter Varianz σ^2 gegeben durch:

$$K_1 = \left[\bar{x} - u_{1-\alpha/2} \frac{\sigma}{\sqrt{n}}, \bar{x} + u_{1-\alpha/2} \frac{\sigma}{\sqrt{n}} \right],$$

wobei $u_{1-\alpha/2}$ das $(1 - \alpha/2)$ -Quantil der Standardnormalverteilung $N(0, 1)$ bezeichnet.

Einsetzen der vorgegebenen Zahlenwerte liefert:

$$n = 12,$$

$$1 - \alpha = 0.9 \Leftrightarrow \alpha = 0.1 \Leftrightarrow 1 - \alpha/2 = 0.95$$

$$\Rightarrow u_{1-\alpha/2} = u_{0.95} \approx 1.64,$$

$$\sigma^2 = 0.0324$$

$$\Rightarrow \sigma = \sqrt{\sigma^2} = 0.18,$$

$$\bar{x} = \frac{1}{12} \sum_{i=1}^{12} x_i \approx 3.528.$$

Daher ist das gesuchte Konfidenzintervall gegeben durch

$$K_1 \approx [3.528 - 1.64 \frac{0.18}{\sqrt{12}}, 3.528 + 1.64 \frac{0.18}{\sqrt{12}}] \approx [3.443, 3.613].$$

- (ii) Ein zweiseitiges Konfidenzintervall für μ zum Konfidenz-Niveau $1 - \alpha$ bei unbekannter Varianz σ^2 ist gemäß D 5.8 gegeben durch

$$K_2 = \left[\bar{x} - t_{1-\alpha/2}(n-1) \frac{\hat{\sigma}}{\sqrt{n}}, \bar{x} + t_{1-\alpha/2}(n-1) \frac{\hat{\sigma}}{\sqrt{n}} \right]$$

Hierbei bezeichne $t_{1-\alpha/2}(n-1)$ das $(1 - \alpha/2)$ -Quantil zu $t(n-1)$ (t -Verteilung mit $n-1$ Freiheitsgraden) und $\hat{\sigma}^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 = \frac{1}{n-1} (\sum_{i=1}^n x_i^2 - n\bar{x}^2)$ die erwartungstreue Punktschätzung für σ^2 .

Einsetzen der vorgegebenen Zahlenwerte liefert:

$$\begin{aligned} n &= 12, \\ 1 - \alpha/2 &= 0.95 \\ \Rightarrow t_{1-\alpha/2}(n-1) &= t_{0.95}(11) \stackrel{Tab.}{\approx} 1.796, \\ \bar{x} &\approx 3.528, \\ \hat{\sigma} &= \sqrt{\frac{1}{11} \left(\sum_{i=1}^{12} (x_i - \bar{x})^2 \right)} \approx 0.185. \end{aligned}$$

Daher ist das gesuchte Konfidenzintervall gegeben durch

$$K_2 \approx [3.528 - 1.796 \frac{0.185}{\sqrt{12}}, 3.528 + 1.796 \frac{0.185}{\sqrt{12}}] \approx [3.432, 3.623].$$

- (b) (i) Gemäß D 5.9 ist ein einseitiges unteres Konfidenzintervall für σ^2 bei unbekanntem μ zum Konfidenz-Niveau $1 - \alpha$ gegeben durch

$$K_3 = \left[0, \frac{n-1}{\chi_\alpha^2(n-1)} \hat{\sigma}^2 \right]$$

wobei $\hat{\sigma}^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 = \frac{1}{n-1} (\sum_{i=1}^n x_i^2 - n\bar{x}^2)$ die erwartungstreue Punktschätzung für σ^2 bezeichne.

Einsetzen der vorgegebenen Zahlenwerte liefert:

$$\begin{aligned} n &= 12, \\ \alpha &= 0.01 \Rightarrow \chi_\alpha^2(n-1) = \chi_{0.01}^2(11) \stackrel{Tab.}{\approx} 3.05, \\ \hat{\sigma}^2 &\stackrel{vgl. (a)(ii)}{\approx} 0.0342. \end{aligned}$$

Daher ist das gesuchte Konfidenzintervall gegeben durch

$$K_3 = [0, \frac{n-1}{\chi_\alpha^2(n-1)} \hat{\sigma}^2] \approx [0, 0.123].$$

- (ii) Gemäß D 5.11 erhält man ein einseitiges unteres Konfidenzintervall für σ zum Konfidenz-Niveau $1 - \alpha = 99\%$ durch Ziehen der Quadratwurzel aus den Intervallgrenzen des Konfidenzintervalls aus (a):

$$K_4 = \left[\sqrt{0}, \sqrt{\frac{n-1}{\chi_\alpha^2(n-1)} \hat{\sigma}^2} \right] \stackrel{(i)}{\approx} [0, \sqrt{0.123}] \approx [0, 0.351].$$

- (c) Versuchsplanung: Laut D 5.7 hat das zweiseitiges Konfidenzintervall für μ zum Konfidenzniveau $1 - \alpha$ bei bekannter Varianz σ^2 höchstens die Länge $\mathfrak{L}_0$, falls

$$n \geq \frac{4\sigma^2 u_{1-\alpha/2}^2}{\mathfrak{L}_0^2}.$$

Einsetzen der vorgegebenen Zahlenwerte $\mathfrak{L}_0 = 0.1, \sigma^2 = 0.0324, u_{1-\alpha/2} = u_{0.95} \approx 1.64$ liefert:

$$n \geq \frac{4 \cdot 0.0324 \cdot 1.64^2}{0.01} \approx 34.86.$$

Also muss $n \geq 35$ gelten, damit die Länge des Konfidenzintervalls höchstens 0.1 beträgt.

Lösung Aufgabe 36

Gemäß Modell D 5.12 und Verfahren D 5.13 der Vorlesung ist ein zweiseitiges $(1-\alpha)$ -Konfidenzintervall für die Differenz der Erwartungswerte $\Delta = \mu_1 - \mu_2$ bei unbekannter (gleicher) Varianz σ^2 gegeben durch:

$$K = \left[\hat{\Delta} - t_{1-\alpha/2}(n_1 + n_2 - 2) \hat{\sigma}_{pool} \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}, \hat{\Delta} + t_{1-\alpha/2}(n_1 + n_2 - 2) \hat{\sigma}_{pool} \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \right],$$

wobei $\hat{\Delta} = \bar{x} - \bar{y}$ eine Punktschätzung für die Differenz der Erwartungswerte Δ ,

$$\hat{\sigma}_{pool}^2 = \frac{1}{n_1 + n_2 - 2} \left(\sum_{i=1}^{n_1} (x_i - \bar{x})^2 + \sum_{j=1}^{n_2} (y_j - \bar{y})^2 \right) = \frac{n_1 - 1}{n_1 + n_2 - 2} \hat{\sigma}_1^2 + \frac{n_2 - 1}{n_1 + n_2 - 2} \hat{\sigma}_2^2$$

eine kombinierte Varianzschätzung mit den Stichprobenvarianzen $\hat{\sigma}_1^2$ und $\hat{\sigma}_2^2$ der Stichproben $x_1, \dots, x_{n_1}$ und $y_1, \dots, y_{n_2}$ sowie $t_{1-\alpha/2}(n_1 + n_2 - 2)$ das $(1 - \alpha/2)$ -Quantil der $t(n_1 + n_2 - 2)$ -Verteilung sind.

In der hier vorliegenden Situation gilt:

$$n_1 = 12, \quad n_2 = 8,$$

$$1 - \alpha = 0.95 \quad \Leftrightarrow \quad 1 - \frac{\alpha}{2} = 0.975 \quad \Rightarrow \quad t_{1-\alpha/2}(n_1 + n_2 - 2) = t_{0.975}(18) \approx 2.101,$$

$$\bar{x} = 50.4, \quad \bar{y} = 50.6 \quad \Rightarrow \quad \hat{\Delta} = -0.2,$$

$$\hat{\sigma}_1^2 \approx 4.433, \quad \hat{\sigma}_2^2 \approx 3.326 \quad \Rightarrow \quad \hat{\sigma}_{pool} \approx \sqrt{\frac{11 \cdot 4.433 + 7 \cdot 3.326}{18}} \approx 2.0006.$$

Nach Einsetzen dieser Werte erhalten wir als zweiseitiges Konfidenzintervall für $\Delta = \mu_1 - \mu_2$ zum Konfidenzniveau 0.95:

$$K = \left[-0.2 - 2.101 \cdot 2.0006 \cdot \sqrt{\frac{1}{12} + \frac{1}{8}}, -0.2 + 2.101 \cdot 2.0006 \cdot \sqrt{\frac{1}{12} + \frac{1}{8}} \right] \approx [-2.027, 1.627].$$

Lösung Aufgabe 37

- (a) Gemäß D 7.4 der Vorlesung sind die mittels der Kleinsten Quadrate-Methode hergeleiteten Schätzfunktionen für die Steigung b und den Achsenabschnitt a gegeben durch

$$\hat{b} = \frac{s_{xY}}{s_x^2} \approx \frac{-0.0361}{0.0136} \approx -2.654.$$

und

$$\hat{a} = \bar{Y} - \hat{b}\bar{x} \approx 6.1 + 2.654 \cdot 2.65 \approx 13.133.$$

Hieraus ergibt sich die geschätzte Regressionsgerade

$$\widehat{f(x)} = \hat{a} + \hat{b}x \approx 13.133 - 2.654x, \quad x \in \mathbb{R}.$$

Speziell für $x = 2.7$ (Eur/Kg) erhält man die geschätzte Absatzmenge

$$\widehat{f(2.7)} = 13.133 - 2.654 \cdot 2.7 \approx 5.967.$$

- (b) Gemäß D 7.13 (iii) der Vorlesung ist ein zweiseitiges $(1-\alpha)$ -Konfidenz-Intervall für σ^2 gegeben durch

$$K = \left[\frac{(n-2)\hat{\sigma}^2}{\chi_{1-\alpha/2}^2(n-2)}, \frac{(n-2)\hat{\sigma}^2}{\chi_{\alpha/2}^2(n-2)} \right],$$

wobei $\chi_{\beta}^2(n-2)$ als β -Quantil zu $\chi^2(n-2)$ und (vgl. D 7.7 der Vorlesung)

$$\hat{\sigma}^2 = \frac{1}{n-2} \sum_{i=1}^n (Y_i - \hat{a} - \hat{b}x_i)^2 = \frac{n}{n-2} s_Y^2 (1 - r_{xY}^2), \quad r_{xY}^2 = \frac{s_{xY}^2}{s_x^2 s_Y^2}.$$

Zahlenwerte:

$$n = 7, \quad \hat{\sigma}^2 = \frac{7}{5} s_Y^2 \left(1 - \frac{s_{xY}^2}{s_x^2 s_Y^2}\right) \approx \frac{7}{5} 0.1657 \left(1 - \frac{(-0.0361)^2}{0.0136 \cdot 0.1657}\right) \approx 0.098,$$

$$1 - \alpha = 0.9 \quad \Rightarrow \quad \chi_{1-\alpha/2}^2(n-2) = \chi_{0.95}^2(5) \approx 11.07, \quad \chi_{\alpha/2}^2(n-2) = \chi_{0.05}^2(5) \approx 1.15.$$

Damit:

$$K \approx \left[\frac{5 \cdot 0.098}{11.07}, \frac{5 \cdot 0.098}{1.15} \right] \approx [0.044, 0.426].$$

Einführung in die angewandte Stochastik

1. Test

Vor- und Nachname: _____ Matr.-Nr.: _____

Bitte kreuzen Sie, in welcher Fragestunde Sie Ihren Test zurückerhalten möchten:

A Mo 8.15-9.45 h - V

B Mo 10.00-11.30 h - AHI

C Di 13.15-14.45 h - AHII

D Mi 13.15-14.45 h - AHI

E Do 11.45-13.15 h - V

F Fr 13.30-15.00 h - H 201

- (a) 20 Kunden eines Fahrzeugherstellers wurden danach gefragt, nach wie vielen gefahrenen Kilometern sie zum ersten Mal eine Panne am Fahrzeug hatten. Dabei wurden folgende Angaben (in tausend Kilometern) gemacht:

Kunde i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
x_i	10	55	100	15	55	80	150	80	100	55	70	80	10	130	70	55	80	40	10	100
$x_{(i)}$	10	10	10	15	40	55	55	55	55	70	70	80	80	80	80	100	100	100	130	150
$R(x_i)$																				

- (i) Berechnen Sie die Ränge. Nutzen Sie dazu den Platz in obiger Tabelle.

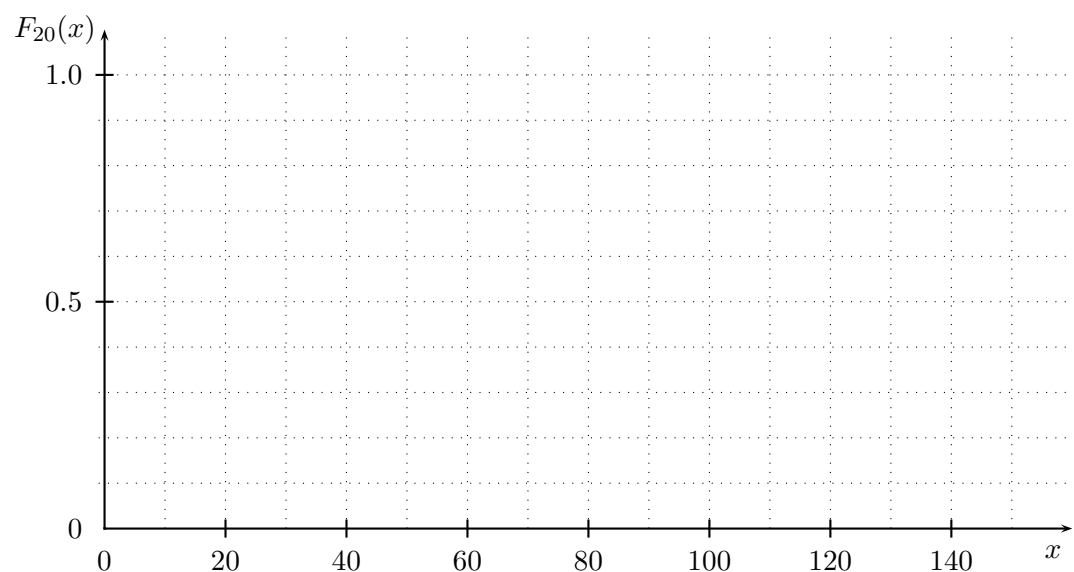
6 Punkte

- (ii) Geben Sie den Modus, den empirischen Median sowie das untere und obere Quartil an.

4 Punkte

- (iii) Bestimmen Sie die empirische Verteilungsfunktion F_{20} der Daten und zeichnen Sie diese in das vorgegebene Koordinatensystem ein.

12 Punkte



- (iv) Wie groß ist die relative Häufigkeit dafür, mit einem Produkt dieses Herstellers die erste Panne
- * in den ersten 20 000 Kilometern,
 - * nach den ersten 70 000 Kilometern,
 - * zwischen 41 000 und 90 000 Kilometern
- zu erleiden?

4 Punkte

- (v) Kennzeichnen Sie in der Zeichnung aus (iii), wie die Anteile aus (iv) dort abgelesen werden können.

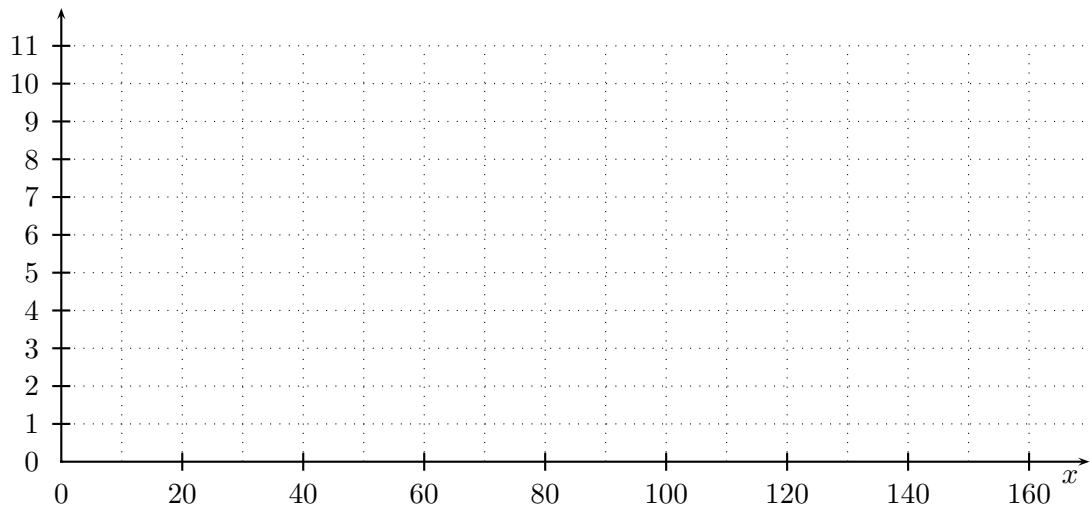
3 Punkte

- (b) Berechnen Sie zu der Klassierung

$$K_1 = [0, 40] , \quad K_2 = (40, 70] , \quad K_3 = (70, 100] , \quad K_4 = (100, 160]$$

die relativen Klassenhäufigkeiten $f(K_i)$, die Klassenbreiten b_i und die Höhen der Rechtecke im zugehörigen Histogramm h_i , $i = 1, 2, 3$. Zeichnen Sie anschließend das zugehörige Histogramm in das unten stehende Koordinatensystem ein.

6 Punkte



Einführung in die angewandte Stochastik

2. Test

Vor- und Nachname: _____ Matr.-Nr.: _____

Bitte kreuzen Sie an, in welcher Fragestunde Sie Ihren Test zurückerhalten möchten:

A Mo 8.15-9.45 h - V

B Mo 10.00-11.30 h - AHI

C Di 13.15-14.45 h - AHII

D Mi 13.15-14.45 h - AHI

E Do 11.45-13.15 h - V

F Fr 13.30-15.00 h - H 201

Aufgabe 1

Ein Statistiker hat bei einer Untersuchung folgende Paare (x_i, y_i) von Messwerten erhalten:

i	1	2	3	4	5
x_i	2	3	4	5	6
y_i	7	4	3	4	7

- (a) Berechnen Sie die Regressionsgerade von y bzgl. x .

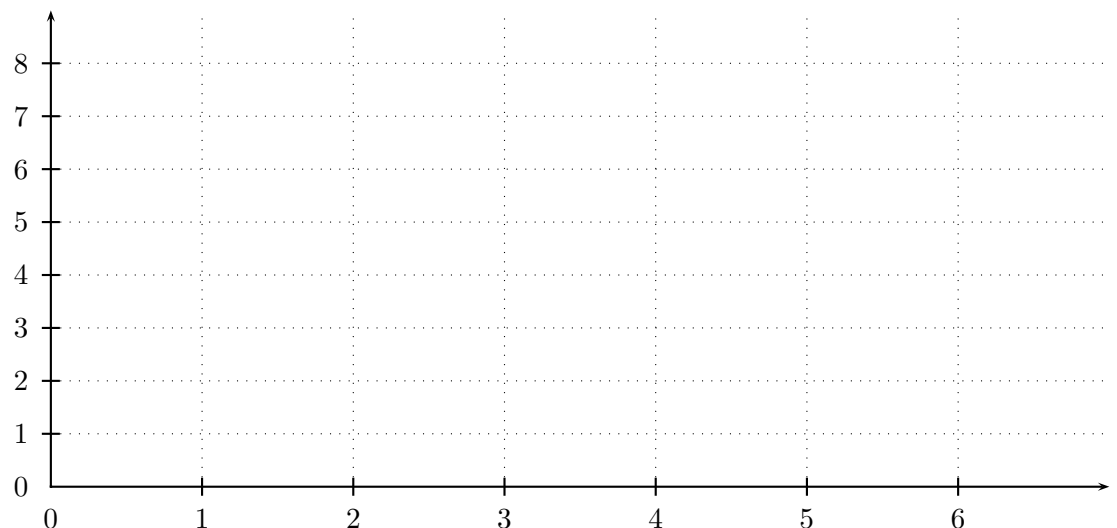
5 Punkte

- (b) Sind aufgrund der Beobachtungswerte die zugrundeliegenden Merkmale X und Y - gemessen am Bravais-Pearson-Korrelationskoeffizienten - positiv, negativ oder unkorreliert?

2 Punkte

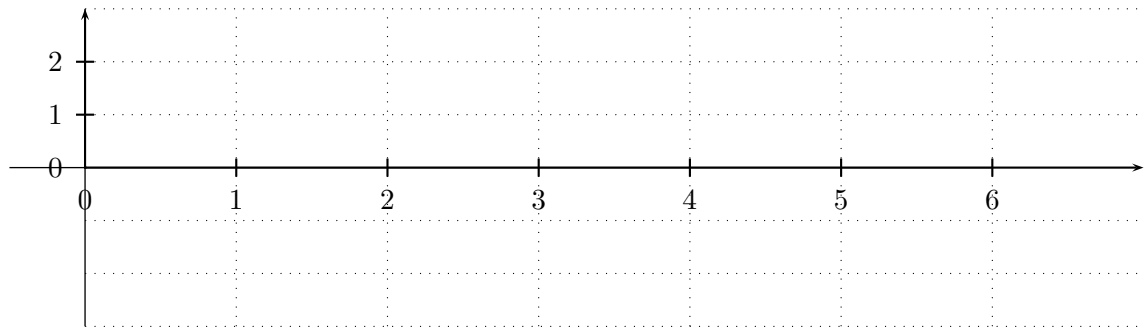
- (c) Zeichnen Sie die Messpunkte (x_i, y_i) und die Regressionsgerade in das folgende Diagramm ein.

2 Punkte



- (d) Berechnen Sie die Residuen und zeichnen Sie den Residualplot.

4 Punkte



- (e) Besteht ein starker, ein schwacher oder kein erkennbarer formelmässiger Zusammenhang zwischen den x -Werten und y -Werten? (Betrachten Sie (c).) Erklären Sie den scheinbaren Widerspruch zu (b).

2 Punkte

Aufgabe 2

- (a) Ein symmetrischer (sechseckiger) Würfel wird zweimal hintereinander geworfen.
 (i) Geben Sie für dieses Zufallsexperiment einen geeigneten Wahrscheinlichkeitsraum $(\Omega, \mathfrak{A}, P)$ an.

2 Punkte

- (ii) Beschreiben Sie die folgenden Ereignisse als Teilmengen von Ω .

A: Es wird zweimal die gleiche Augenzahl geworfen.

B: Die Differenz 'Höhere Augenzahl minus niedrigere Augenzahl' beträgt 2.

2 Punkte

- (iii) Berechnen Sie die Wahrscheinlichkeiten für das Eintreten der Ereignisse aus (ii).

2 Punkte

- (b) Beim Würfeln eines unsymmetrischen (sechseckigen) Würfels sind zu den Ereignissen

A: Es wird eine gerade Augenzahl geworfen,

B: Es wird eine Augenzahl kleiner oder gleich 3 geworfen,

C: Es wird eine 1 oder 4 geworfen,

die folgenden Wahrscheinlichkeiten bekannt:

$$P(A) = \frac{5}{8}, \quad P(C) = \frac{5}{12}, \quad P(A \cup B) = \frac{23}{24}, \quad P(A \cap B) = \frac{1}{6}, \quad P(A \cap C) = \frac{1}{3}.$$

Berechnen Sie die Elementarwahrscheinlichkeiten $P(\{\omega\})$ für $\omega \in \Omega = \{1, \dots, 6\}$.

7 Punkte

Einführung in die angewandte Stochastik

3. Test

Vor- und Nachname: _____ Matr.-Nr.: _____

Ihre erarbeitete Lösung der Aufgaben können Sie am 24. Juni in der Globalübung abgeben. Dort werden eine Woche später (am 1. Juli) dann auch die korrigierten Lösungen zurückgegeben.

Die Bearbeitung der Aufgaben sollte in etwa 45 Minuten in Anspruch nehmen.

Aufgabe 1

Ein Chiphersteller hat drei Werke A , B und C , die jeweils 35%, 15% und 50% der Gesamtproduktion herstellen. Die Wahrscheinlichkeit, dass ein produzierter Chip nicht defekt ist, sei 0,75 bei Fabrik A , 0,95 bei Fabrik B und 0,85 bei Fabrik C .

- Mit welcher Wahrscheinlichkeit ist ein beliebiger Chip defekt?
- Ein Kunde erhält einen defekten Chip. Mit welcher Wahrscheinlichkeit kam dieser aus Fabrik C ?

Aufgabe 2

Gegeben sei die folgende Wahrscheinlichkeitstabelle eines diskreten Zufallsvektors (X, Y) mit der Wahrscheinlichkeitsverteilung $p_{ij} = P(X = i, Y = j)$ für $i = 0, 1$ und $j = -1, 0, 1$.

$X = i \backslash Y = j$	-1	0	1	$P(X = i)$
0				$\frac{1}{3}$
1				
$P(Y = j)$	$\frac{1}{4}$	$\frac{1}{2}$		

- Berechnen Sie die fehlenden Werte dieser Tabelle so, dass die Zufallsvariablen X und Y stochastisch unabhängig sind.
- Bestimmen Sie in der Situation aus (i) die bedingte Zähdichte von X bei gegebenem $Y = 0$.
- Bestimmen Sie in der Situation aus (i) die Verteilung der Zufallsvariablen $Z = X + Y$.
- Bestimmen Sie die bedingte Wahrscheinlichkeit $P(Z \leq 1 | X = 1)$.
- Berechnen Sie $E(Z)$ und $Var(Z)$.

Einführung in die angewandte Stochastik

4. Test

Vor- und Nachname: _____ Matr.-Nr.: _____

Ihre erarbeitete Lösung der Aufgaben können Sie am 08. Juli in der Globalübung abgeben. Dort werden eine Woche später (am 15. Juli) dann auch die korrigierten Lösungen zurückgegeben.

Die Bearbeitung der Aufgaben sollte in etwa 45 Minuten in Anspruch nehmen.

Aufgabe 1

Die Studenten A und B verabreden sich zwischen 12 und 13 Uhr in der Mensa. Sie erscheinen unabhängig voneinander, wobei die Zeitpunkte ihres Eintreffens sich durch unabhängige, in $[12, 13]$ rechteckverteilte Zufallsgrößen X und Y beschreiben lassen. Wie groß ist die Wahrscheinlichkeit dafür,

- (a) dass beide vor 12 Uhr 30 eintreffen?
- (b) dass A vor 12 Uhr 15 und B nach 12 Uhr 45 eintrifft?
- (c) dass A vor B eintrifft?

Aufgabe 2

Eine faire Münze (Kopf/Zahl) werde 100 mal unabhängig geworfen. Die Anzahl der geworfenen Köpfe sei mit X bezeichnet.

- (a) Geben Sie ein geeignetes Modell für die gegebene Situation an und bestimmen Sie den Erwartungswert und die Varianz von X .
- (b) Bestimmen Sie mit Hilfe der Tschebyscheff-Ungleichung eine Schranke für die Wahrscheinlichkeit, dass die Anzahl der geworfenen Köpfe (echt) zwischen 40 und 60 liegt.

Aufgabe 3

Die Zufallsvariablen $X_1, \dots, X_n$ seien stochastisch unabhängig und identisch verteilt mit der Dichte

$$f_\lambda(x) = \begin{cases} 2\lambda x \cdot e^{-\lambda x^2}, & x > 0, \\ 0, & \text{sonst,} \end{cases}$$

mit Parameter $\lambda > 0$.

Bestimmen Sie einen Maximum-Likelihood-Schätzer $\hat{\lambda}(X_1, \dots, X_n)$ für λ .