Klausurensammlung Deskr. Statistik und Wahrscheinlichkeitsrechnung

apl. Prof. Dr. Martin Becker

26. März 2025

Inhaltsverzeichnis

1 Deskr. Statistik und Wahrscheinlichkeitsrechnung WS 2014/15

Aufgabe 1 (16 Punkte)

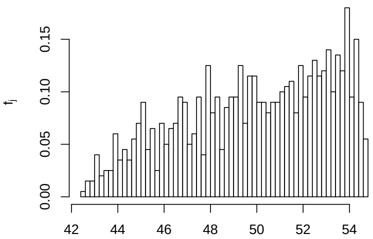
Markieren Sie jeweils mit einem Kreuz pro Aussage im betreffenden Kästchen, ob die unten stehenden Aussagen wahr oder falsch sind.
Richtige Antworten geben 2 Punkte, falsche Antworten und nicht bearbeitete Aussagen 0 Punkte (Aussagen mit zwei Kreuzen zählen als nicht bearbeitet!).

	wahr	falsch
1. Ist $\overline{x}$ der arithmetische Mittelwert des Merkmals X, dann sind stets genauso viele Urlisteneinträge x_i kleiner oder gleich $\overline{x}$ wie größer oder gleich $\overline{x}$.	☐	☐
2. Bei der Bildung der Klassen für die Klassierung eines Merkmals muss stets gewährleistet sein, dass jeder Urlisteneintrag in genau einer Klasse liegt.	☐	☐
3. Sind in einer Klausur mehr männliche als weibliche Prüfungsteilnehmer durchgefallen und haben mehr weibliche als männliche Prüflinge an der Klausur teilgenommen, so kann man daraus stets schließen, dass die Durchfallquote unter den männlichen Prüfungsteilnehmern größer ist als die unter den weiblichen Prüfungsteilnehmern.	☐	☐
4. Es seien $(\Omega, \mathcal{F}, P)$ ein Wahrscheinlichkeitsraum, $A, B, C \in \mathcal{F}$ drei Ereignisse mit $P(A) = P(B) = P(C) = \frac{1}{3}$. Dann gilt stets: $$P(A \cup B \cup C) = 1$$	☐	☐
5. Die Anzahl der Möglichkeiten, 3 aus 9 unterscheidbaren Objekten auszuwählen, entspricht der Anzahl der Möglichkeiten, 6 aus 9 unterscheidbaren Objekten auszuwählen.	☐	☐
6. Hat die Verteilungsfunktion F_X einer Zufallsvariablen X mindestens eine Sprungstelle, so kann X nicht stetig verteilt sein.	☐	☐
7. Die Summe von 4 stochastisch unabhängigen $N(10, 2^2)$-verteilten Zufallsvariablen ist $N(40, 8^2)$-verteilt.	☐	☐
8. Die Kovarianz $\mathrm{Cov}(X, Y)$ zweier Zufallsvariablen X und Y ist betragsmäßig stets größer als der Korrelationskoeffizient $\mathrm{Korr}(X, Y)$ der beiden Zufallsvariablen.	☐	☐

Aufgabe 2 (12 Punkte)

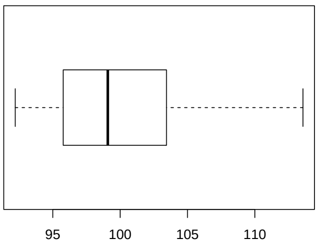
Markieren Sie jeweils die korrekte Antwort mit einem Kreuz im betreffenden Kästchen. Es ist jeweils genau ein Kreuz korrekt.
Richtige Antworten geben 3 Punkte, falsche Antworten und nicht bearbeitete Aufgabenteile 0 Punkte (Aufgabenteile mit mehr als einem Kreuz zählen als nicht bearbeitet!).

1. Kreuzen Sie an, auf welche Merkmalseigenschaften das folgende Histogramm *am ehesten* hindeutet:



(a) leptokurtisch und rechts**steil** ☐
(b) leptokurtisch und links**steil** ☐
(c) platykurtisch und rechts**steil** ☐
(d) platykurtisch und links**steil** ☐

2. Kreuzen Sie an, auf welche Merkmalseigenschaften der folgende Box-Plot *am ehesten* hindeutet:



(a) leptokurtisch und rechts**steil** ☐
(b) leptokurtisch und links**steil** ☐
(c) platykurtisch und rechts**steil** ☐
(d) platykurtisch und links**steil** ☐

3. Verteilungsfunktionen eindimensionaler Zufallsvariablen sind stets

(a) monoton wachsend und rechtsseitig stetig. ☐
(b) streng monoton wachsend und rechtsseitig stetig. ☐
(c) monoton wachsend und stetig. ☐
(d) streng monoton wachsend und stetig. ☐

4. Die Wahrscheinlichkeit, in dieser Klausuraufgabe (4 MC-Aufgabenteile mit jeweils genau einer korrekten Antwort aus 4 Antwortmöglichkeiten) durch *rein zufälliges* Ankreuzen jeweils einer Antwortmöglichkeit (jede Antwortmöglichkeit erhalte also eine Chance von 25%) keine einzige richtige Antwort zu markieren, beträgt (ggf. auf 2 Nachkommastellen gerundet):

(a) 25.00% ☐
(b) 31.64% ☐
(c) 68.36% ☐
(d) 75.00% ☐

Aufgabe 3 (1 + 4 + 5 + 3 = 13 Punkte)

Zu einem erhobenen Merkmal X sei die folgende empirische Verteilungsfunktion gegeben:

$$F(x) = \begin{cases} 0.000 & \text{für} \quad x < 1 \\ 0.100 & \text{für} \quad 1 \leq x < 3 \\ 0.350 & \text{für} \quad 3 \leq x < 5 \\ 0.575 & \text{für} \quad 5 \leq x < 7 \\ 0.825 & \text{für} \quad 7 \leq x < 9 \\ 0.950 & \text{für} \quad 9 \leq x < 11 \\ 1.000 & \text{für} \quad x \geq 11 \end{cases}$$

Außerdem sei die Länge der Urliste $n = 40$ bekannt.

(a) Geben Sie die Menge A der Merkmalsausprägungen an.

(b) Erstellen Sie eine Tabelle der absoluten und relativen Häufigkeiten.

(c) Berechnen Sie den arithmetischen Mittelwert und die empirische Varianz des Merkmals X.

(d) Bestimmen Sie ein unteres Quartil, ein oberes Quartil und den zugehörigen Interquartilsabstand des Merkmals X.

Aufgabe 4 (6 + 4 + 3 + 3 = 16 Punkte)

Zu einem kardinalskalierten Merkmal sei die folgende (zur einfacheren Bearbeitung der Aufgabe bereits sortierte) Urliste der Länge $n = 25$ gegeben:

13.76, 14.18, 14.48, 15.11, 16.27, 19.65, 20.99, 21.98, 25.33, 27.60, 29.75, 30.46, 30.77, 31.74, 33.27, 33.31, 35.26, 35.65, 36.36, 43.75, 45.55, 49.97, 54.53, 61.62, 64.66

(a) Führen Sie eine Klassierung der erhobenen Daten auf Grundlage der Klassen

$$K_1 = (10, 20],\ K_2 = (20, 30],\ K_3 = (30, 45],\ K_4 = (45, 70]$$

durch. Geben Sie insbesondere die jeweiligen Klassenbreiten, Klassenmitten, absoluten und relativen Klassenhäufigkeiten, Häufigkeitsdichten sowie die Werte der empirischen Verteilungsfunktion an den Klassengrenzen an.

(b) Stellen Sie die (approximative) empirische Verteilungsfunktion des Merkmals aus der Klassierung der Daten auf.

(c) Berechnen Sie aus den klassierten Daten den (approximativen) arithmetischen Mittelwert der Daten. Wie groß ist die relative Abweichung vom tatsächlichen (aus der Urliste bestimmten) Mittelwert von 32.24?

(d) Bestimmen Sie (unter Verwendung der bekannten Konvention zur eindeutigen Festlegung von Quantilen bei kardinalskalierten Merkmalen) *das* obere Quartil sowohl exakt aus der Urliste als auch approximativ mit Hilfe der Verteilungsfunktion für die klassierten Daten.

Aufgabe 5 (2 + 2 + 2 = 6 Punkte)

Bei der Herstellung von Terrassendielen tritt mit einer Wahrscheinlichkeit von 3% ein Fehler beim Zuschnitt der Dielen auf, mit einer Wahrscheinlichkeit von 2% ein Fehler beim Fräsen der Dielen und mit einer Wahrscheinlichkeit von 1.5% sowohl ein Fehler beim Zuschnitt der Dielen als auch ein Fehler beim Fräsen der Dielen. Berechnen Sie die Wahrscheinlichkeit dafür, dass

(a) höchstens einer der beiden Fehler,

(b) mindestens einer der beiden Fehler,

(c) ein Fehler beim Zuschnitt der Dielen, aber kein Fehler beim Fräsen der Dielen

auftritt.

Aufgabe 6 (6 + 2 + 2 = 10 Punkte)

Ein Hersteller von Tiefkühlfertiggerichten bezieht seine Frischfischlieferungen von den vier Großhändlern A, B, C und D. Dabei werden einzelne Lieferungen mit einer Wahrscheinlichkeit von 40% von Lieferant A, 35% von Lieferant B, 15% von Lieferant C und 10% von Lieferant D geliefert. Bei den anschließenden Qualitätskontrollen gibt es erfahrungsgemäß mit einer Wahrscheinlichkeit von 1% bei Lieferant A, 1.5% bei Lieferant B, 3% bei Lieferant C und 4% bei Lieferant D Anlass zu Beanstandungen.

(a) Wie groß ist die Wahrscheinlichkeit, dass eine zufällig ausgewählte Lieferung in der Qualitätskontrolle nicht beanstandet wird?

(b) Wie groß ist die Wahrscheinlichkeit, dass eine von der Qualitätskontrolle nicht beanstandete Lieferung von Großhändler B geliefert wurde?

(c) Sind die Ereignisse „Lieferung wird beanstandet“ und „Lieferung stammt von Großhändler B“ stochastisch unabhängig? Begründen Sie Ihre Antwort!

Aufgabe 7 (1 + 2 + 2 + 2 = 7 Punkte)

Die Anzahl der Unfälle pro Tag auf einem bestimmten Autobahnabschnitt lasse sich als eine Pois(0.25)-verteilte Zufallsvariable auffassen. Außerdem soll angenommen werden, dass die Anzahl der Unfälle pro Tag auf diesem Autobahnabschnitt für unterschiedliche Tage stochastisch unabhängig ist.

(a) Welchen Erwartungswert hat die Anzahl der Unfälle pro Tag auf diesem Autobahnabschnitt?

(b) Mit welcher Wahrscheinlichkeit ereignen sich an einem Tag auf diesem Autobahnabschnitt 0 Unfälle?

(c) Mit welcher Wahrscheinlichkeit ereignet sich an einem Tag auf diesem Autobahnabschnitt höchstens 1 Unfall?

(d) Welche Verteilung hat die Anzahl der Unfälle pro Woche auf diesem Autobahnabschnitt? Mit welcher Wahrscheinlichkeit ereignet sich in einer Woche auf diesem Autobahnabschnitt mindestens 1 Unfall?

Aufgabe 8 (5 + 2 + 6 + 4 = 17 Punkte)

Die Verteilung einer stetigen Zufallsvariablen X sei durch die folgende Dichtefunktion gegeben:

$$f_X(x) = \begin{cases} \frac{1}{4}x + \frac{1}{4} & \text{für } -1 \le x < 1 \\ -x + 2 & \text{für } 1 \le x \le 2 \\ 0 & \text{sonst} \end{cases}$$

(a) Berechnen Sie die Verteilungsfunktion F_X von X.

(b) Berechnen Sie $P(\{X < 0\})$ und $P\left(\left\{0 < X < \frac{3}{2}\right\}\right)$.

(c) Bestimmen Sie den Erwartungswert $\mathrm{E}(X)$.

(d) Bestimmen Sie das obere Quartil von X.

Aufgabe 9 (2 + 9 + 1 + 3 = 15 Punkte)

Gegeben sei die folgende Tabelle der gemeinsamen Wahrscheinlichkeitsverteilung zu einem zweidimensionalen diskreten Zufallsvektor (X, Y):

$X \setminus Y$	1	3	5	$p_{i\cdot}$
2	$\frac{1}{6}$	$\frac{1}{12}$	$\frac{1}{4}$	
4	$\frac{1}{3}$	$\frac{1}{6}$	0	
$p_{\cdot j}$				

(a) Ergänzen Sie die obige Tabelle (in den vorgesehenen Feldern) um ihre Randverteilungen.

(b) Berechnen Sie $\mathrm{E}(X)$, $\mathrm{E}(Y)$, $\mathrm{Var}(X)$, $\mathrm{Var}(Y)$, $\mathrm{Cov}(X,Y)$ sowie $\mathrm{Korr}(X,Y)$.

(c) Sind X und Y stochastisch unabhängig?

(d) Berechnen Sie $\mathrm{E}(2X - 6Y)$ sowie $\mathrm{Var}(2X - 6Y)$.

Aufgabe 10 (2 + 2 + 3 = 7 Punkte)

Ein freiberuflicher Netzwerktechniker benötigt für das Auflegen eines Netzwerkanschlusses im Mittel 16 Minuten bei einer Standardabweichung von 2 Minuten. Man kann davon ausgehen, dass die benötigten Zeitdauern für die einzelnen Anschlüsse nicht gegenseitig voneinander abhängen. In einem bestimmten Schaltschrank sind insgesamt 25 Anschlüsse aufzulegen.

(a) Geben Sie die den Erwartungswert sowie die Standardabweichung der gesamten Arbeitszeit (für alle 25 Netzwerkanschlüsse) an.

(b) Verwenden Sie den zentralen Grenzwertsatz, um näherungsweise zu bestimmen, mit welcher Wahrscheinlichkeit der Netzwerktechniker nicht länger als 7 Stunden bzw. 420 Minuten zum Auflegen aller Anschlüsse benötigt.

(c) Verwenden Sie den zentralen Grenzwertsatz, um näherungsweise ein 0.95-Quantil für die gesamte Arbeitszeit (für alle 25 Anschlüsse) zu bestimmen.

Hinweis: Verwenden Sie zur Bearbeitung von Aufgabenteil (b) und (c) die Tabelle zur Standardnormalverteilung auf Seite 120!

2 Deskr. Statistik und Wahrscheinlichkeitsrechnung SS 2015

Aufgabe 1 (16 Punkte)

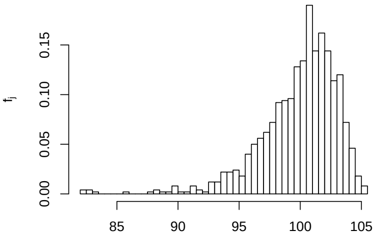
Markieren Sie jeweils mit einem Kreuz pro Aussage im betreffenden Kästchen, ob die unten stehenden Aussagen wahr oder falsch sind.
Richtige Antworten geben 2 Punkte, falsche Antworten und nicht bearbeitete Aussagen 0 Punkte (Aussagen mit zwei Kreuzen zählen als nicht bearbeitet!).

		wahr	falsch
1.	Ordinalskalierte Merkmale sind stets numerisch.	☐	☐
2.	Ist x_p ein p-Quantil eines kardinalskalierten Merkmals X (mit $0 < p < 1$), dann ist höchstens ein Anteil von $1 - p$ der Urlisteneinträge zum Merkmal X größer als x_p.	☐	☐
3.	Sind in einer Klausur mehr männliche als weibliche Prüfungsteilnehmer durchgefallen und haben mehr männliche als weibliche Prüflinge an der Klausur teilgenommen, so kann man daraus stets schließen, dass die Durchfallquote unter den männlichen Prüfungsteilnehmern größer ist als die unter den weiblichen Prüfungsteilnehmern.	☐	☐
4.	In einem Laplaceschen Wahrscheinlichkeitsraum kann jeder beliebigen Teilmenge der Ergebnismenge eine Wahrscheinlichkeit zugeordnet werden.	☐	☐
5.	Die Wahrscheinlichkeit, beim 5-maligen Würfeln mit einem (fairen) Würfel lauter unterschiedliche Punktzahlen zu erhalten, ist kleiner als 10%.	☐	☐
6.	Wenn Sie alle 8 Aufgabenteile dieser Aufgabe jeweils entweder mit *wahr* oder mit *falsch* beantworten, dann haben Sie zur Bearbeitung der Aufgabe insgesamt 8^2 Möglichkeiten.	☐	☐
7.	Ist (X, Y) zweidimensional normalverteilt, dann gilt stets: $$\lvert \mathrm{Korr}(X, Y)\rvert \leq 0.5$$	☐	☐
8.	Die Varianz der Summe von zwei Zufallsvariablen ist stets mindestens so groß wie die Summe der beiden einzelnen Varianzen.	☐	☐

Aufgabe 2 (12 Punkte)

Markieren Sie jeweils die korrekte Antwort mit einem Kreuz im betreffenden Kästchen. Es ist jeweils genau ein Kreuz korrekt.
Richtige Antworten geben 3 Punkte, falsche Antworten und nicht bearbeitete Aufgabenteile 0 Punkte (Aufgabenteile mit mehr als einem Kreuz zählen als nicht bearbeitet!).

1. Kreuzen Sie an, auf welche Merkmalseigenschaften das folgende Histogramm *am ehesten* hindeutet:



(a) leptokurtisch und rechts**steil** ☐
(b) leptokurtisch und links**steil** ☐
(c) platykurtisch und rechts**steil** ☐
(d) platykurtisch und links**steil** ☐

2. Die Ränge $\text{rg}(X)_1, \ldots, \text{rg}(X)_8$ zur (der Einfachheit halber sortierten) Urliste

$$3, \quad 5, \quad 7, \quad 7, \quad 9, \quad 9, \quad 9, \quad 12$$

des ordinalskalierten Merkmals X lauten:

(a) 3, 5, 7.5, 7.5, 9.5, 9.5, 9.5, 12 ☐
(b) 1, 2, 3, 3, 4, 4, 4, 5 ☐
(c) 1, 2, 3.5, 3.5, 6, 6, 6, 8 ☐
(d) 1, 2, 3.5, 3.5, 5.5, 5.5, 5.5, 7 ☐

3. Wenn Sie in dieser Klausuraufgabe (4 MC-Aufgabenteile mit jeweils genau einer korrekten Antwort aus 4 Antwortmöglichkeiten) alle Aufgabenteile durch *rein zufälliges* Ankreuzen jeweils einer Antwortmöglichkeit (jede Antwortmöglichkeit erhalte also jeweils eine Chance von 25%) bearbeiten, ist der Erwartungswert Ihrer Punktzahl beim aktuellen Bewertungsschema (jeweils 3 Punkte für die einzig richtige Antwort, 0 Punkte für eine falsche Antwort) gleich:

(a) 1 ☐
(b) 2 ☐
(c) 3 ☐
(d) 4 ☐

4. Sind X und Y zwei stochastisch unabhängige Poisson-verteilte Zufallsvariablen mit $E(X) = E(Y) = 8$, dann ist die Verteilung von $X + Y$

(a) im Allgemeinen keine Poisson-Verteilung. ☐
(b) eine Poisson-Verteilung mit Erwartungswert 4. ☐
(c) eine Poisson-Verteilung mit Erwartungswert 8. ☐
(d) eine Poisson-Verteilung mit Erwartungswert 16. ☐

Aufgabe 3 $(3 + 3 + 1 + 5 + 3 = 15$ Punkte)

Bei einer Umfrage wurden 40 Haushalte befragt, wie viele Tablets sie verwenden (Merkmal X). Das Ergebnis der Umfrage ist die folgende (bereits aufsteigend sortierte) Urliste zu X:

0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 3, 3, 4, 4, 4

(a) Erstellen Sie eine Tabelle mit den absoluten und relativen Häufigkeiten.

(b) Stellen Sie die zugehörige empirische Verteilungsfunktion auf.

(c) Wie groß ist der Anteil der Haushalte in der Umfrage, die höchstens 1 Tablet verwenden?

(d) Berechnen Sie den arithmetischen Mittelwert und die empirische Varianz des Merkmals X.

(e) Bestimmen Sie ein unteres Quartil, ein oberes Quartil und den zugehörigen Interquartilsabstand des Merkmals X.

Aufgabe 4 $(6 + 4 + 3 + 3 + 3 = 19$ Punkte)

Zu einem kardinalskalierten Merkmal sei die folgende (zur einfacheren Bearbeitung der Aufgabe bereits sortierte) Urliste der Länge $n = 40$ gegeben:

11.09, 12.57, 12.80, 13.90, 15.43, 17.14, 17.76, 18.68, 21.96, 22.19, 23.80, 25.42, 27.27, 28.30, 32.26, 32.83, 39.57, 39.70, 41.72, 42.72, 44.46, 46.47, 47.07, 47.08, 47.51, 49.24, 50.24, 50.69, 51.24, 53.44, 54.36, 54.86, 60.92, 62.38, 64.29, 64.75, 67.48, 69.50, 70.05, 77.20

(a) Führen Sie eine Klassierung der erhobenen Daten auf Grundlage der Klassen

$$K_1 = (10, 20],\ K_2 = (20, 30],\ K_3 = (30, 60],\ K_4 = (60, 90]$$

durch. Geben Sie insbesondere die jeweiligen Klassenbreiten, Klassenmitten, absoluten und relativen Klassenhäufigkeiten, Häufigkeitsdichten sowie die Werte der empirischen Verteilungsfunktion an den Klassengrenzen an.

(b) Stellen Sie die (approximative) empirische Verteilungsfunktion des Merkmals aus der Klassierung der Daten auf.

(c) Berechnen Sie aus den klassierten Daten den (approximativen) arithmetischen Mittelwert der Daten. Wie groß ist die relative Abweichung vom tatsächlichen (aus der Urliste bestimmten) Mittelwert von 40.758?

(d) Bestimmen Sie (aus der Urliste) die Anzahl von Merkmalswerten zwischen 25 und 50. Welche Näherung für diese Anzahl können Sie aus der in Teil (b) aufgestellten empirischen Verteilungsfunktion berechnen?

(e) Bestimmen Sie (unter Verwendung der bekannten Konvention zur eindeutigen Festlegung von Quantilen bei kardinalskalierten Merkmalen) *den* Median sowohl exakt aus der Urliste als auch approximativ mit Hilfe der Verteilungsfunktion für die klassierten Daten.

Aufgabe 5 (5 + 2 = 7 Punkte)

Bei einem Chip-Hersteller verteilt sich die Produktion eines bestimmten Mikrocontrollers auf insgesamt drei verschiedene Produktionslinien A, B und C. Dabei werden im Mittel 20% der Chips auf Linie A, 30% der Chips auf Linie B und 50% der Chips auf Linie C hergestellt. Aus den Ergebnissen der Qualitätssicherung ist bekannt, dass 3% der Chips aus Linie A, 2% der Chips aus Linie B und 1% der Chips aus Linie C fehlerhaft sind.

(a) Wie groß ist die Wahrscheinlichkeit, dass ein zufällig ausgewählter Chip nicht fehlerhaft ist?

(b) Wie groß ist die Wahrscheinlichkeit, dass ein fehlerhafter Chip auf der Linie C produziert wurde?

Aufgabe 6 (5 + 2 + 6 + 1 + 4 = 18 Punkte)

Die Verteilung einer stetigen Zufallsvariablen X sei durch die folgende Dichtefunktion gegeben:

$$f_X(x) = \begin{cases} \frac{1}{4}x + \frac{3}{4} & \text{für } -3 \leq x < -1 \\ \frac{1}{4}x + \frac{1}{4} & \text{für } -1 \leq x \leq 1 \\ 0 & \text{sonst} \end{cases}$$

(a) Berechnen Sie die Verteilungsfunktion F_X von X.

(b) Berechnen Sie $P\left(\left\{X < -\frac{3}{2}\right\}\right)$ und $P\left(\left\{X > \frac{1}{2}\right\}\right)$.

(c) Bestimmen Sie den Erwartungswert $\mathrm{E}(X)$.

(d) Ist X symmetrisch um ihren Erwartungswert verteilt (ohne Begründung)?

(e) Bestimmen Sie das obere Quartil von X.

Aufgabe 7 (2 + 1 + 2 = 5 Punkte)

Die Wartezeit zwischen zwei Übertragungsfehlern in einem Datennetzwerk lasse sich als eine exponentialverteilte Zufallsvariable auffassen. Im Mittel vergehen zwischen zwei Übertragungsfehlern 20 Sekunden.

(a) Welche Standardabweichung hat die Wartezeit zwischen zwei Übertragungsfehlern?

(b) Mit welcher Wahrscheinlichkeit beträgt die Wartezeit zwischen zwei Übertragungsfehlern mehr als 15 und weniger als 30 Sekunden?

(c) Berechnen Sie das 0.95-Quantil der Wartezeit zwischen zwei Übertragungsfehlern.

Aufgabe 8 (2 + 3 + 9 + 1 + 3 = 18 Punkte)

Gegeben sei die folgende Tabelle der gemeinsamen Wahrscheinlichkeitsverteilung zu einem zweidimensionalen diskreten Zufallsvektor (X, Y):

$X \setminus Y$	2	3	4	$p_{i\cdot}$
0	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{1}{8}$	
1	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{16}$	
2	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{16}$	
$p_{\cdot j}$				

(a) Ergänzen Sie die obige Tabelle (in den vorgesehenen Feldern) um ihre Randverteilungen.

(b) Geben Sie die bedingten Wahrscheinlichkeitsverteilungen von Y unter der Bedingung $X = x_i$ für alle $x_i \in T(X)$ über die zugehörigen (bedingten) Wahrscheinlichkeitsfunktionen an.

(c) Berechnen Sie $\mathrm{E}(X)$, $\mathrm{E}(Y)$, $\mathrm{Var}(X)$, $\mathrm{Var}(Y)$, $\mathrm{Cov}(X, Y)$ sowie $\mathrm{Korr}(X, Y)$.

(d) Sind X und Y stochastisch unabhängig?

(e) Berechnen Sie $\mathrm{E}\left(-3X + 4Y\right)$ sowie $\mathrm{Var}\left(-3X + 4Y\right)$.

Aufgabe 9 (2 + 4 + 4 = 10 Punkte)

Die Zufallsvariablen $X_1, \ldots, X_{432}$ seien unabhängig identisch $B(1, 0.25)$-verteilt. Die Summe der Zufallsvariablen X_i sei mit

$$Y := \sum_{i=1}^{432} X_i = X_1 + \ldots + X_{432}$$

bezeichnet.

(a) Geben Sie die (exakte) Verteilung von Y sowie deren Erwartungswert $\mathrm{E}(Y)$ und Varianz $\mathrm{Var}(Y)$ an.

(b) Verwenden Sie den zentralen Grenzwertsatz, um näherungsweise zu bestimmen, mit welcher Wahrscheinlichkeit Y Werte zwischen 100 und 120 annimmt.

(c) Verwenden Sie den zentralen Grenzwertsatz, um näherungsweise ein 0.90-Quantil von Y zu bestimmen.

Hinweis: Verwenden Sie zur Bearbeitung von Aufgabenteil (b) und (c) die Tabelle zur Standardnormalverteilung auf Seite 120!

3 Deskr. Statistik und Wahrscheinlichkeitsrechnung WS 2015/16

Aufgabe 1 (16 Punkte)

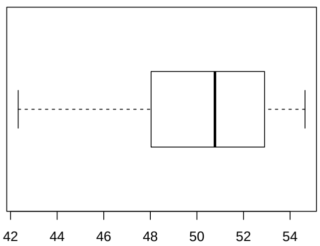
Markieren Sie jeweils mit einem Kreuz pro Aussage im betreffenden Kästchen, ob die unten stehenden Aussagen wahr oder falsch sind.
Richtige Antworten geben 2 Punkte, falsche Antworten und nicht bearbeitete Aussagen 0 Punkte (Aussagen mit zwei Kreuzen zählen als nicht bearbeitet!).

	wahr	falsch
1. Sind alle n Einträge einer Urliste der Länge n zu einem mindestens ordinalskalierten Merkmal X verschieden, so besteht auch die Urliste zum zugehörigen Rangmerkmal $\text{rg}(X)$ aus n unterschiedlichen Werten.	☐	☐
2. Ist a_i eine Ausprägung des Merkmals X und b_j eine Ausprägung des Merkmals Y auf derselben Menge von Merkmalsträgern, dann ist (a_i, b_j) stets eine Ausprägung des zugehörigen zweidimensionalen Merkmals (X, Y).	☐	☐
3. Für den korrigierten Pearsonschen Kontingenzkoeffizienten $C_{X,Y}^{\text{korr}}$ zweier Merkmale X und Y gilt stets $0 \leq C_{X,Y}^{\text{korr}} \leq 1$.	☐	☐
4. Beim gleichzeitigen Würfeln mit zwei (fairen) Würfeln ist es wahrscheinlicher, für die Summe der Punktzahlen den Wert 11 als den Wert 2 zu erhalten.	☐	☐
5. Wenn Sie alle 8 Aufgabenteile dieser Aufgabe rein zufällig mit *wahr* oder *falsch* beantworten, dann beträgt der Erwartungswert Ihrer in dieser Aufgabe erreichten Punktzahl 4.	☐	☐
6. Es seien $(\Omega, \mathcal{F}, P)$ ein Wahrscheinlichkeitsraum, $A, B, C \in \mathcal{F}$ drei Ereignisse mit $P(C) > 0$. Dann gilt: $P(A\|C) < P(B\|C) \quad \Rightarrow \quad P(A) < P(B)$	☐	☐
7. Für eine Zufallsvariable X gelte $P(X \in \{0, 8, 15\}) = 1$. Damit ist X eine diskrete Zufallsvariable.	☐	☐
8. Liegen alle Trägerpunkte eines zweidimensionalen Zufallsvektors (X, Y) auf einer Geraden mit der Steigung -0.8, so gilt $\text{Korr}(X, Y) = -0.8$.	☐	☐

Aufgabe 2 (12 Punkte)

Markieren Sie jeweils die korrekte Antwort mit einem Kreuz im betreffenden Kästchen. Es ist jeweils genau ein Kreuz korrekt.
Richtige Antworten geben 3 Punkte, falsche Antworten und nicht bearbeitete Aufgabenteile 0 Punkte (Aufgabenteile mit mehr als einem Kreuz zählen als nicht bearbeitet!).

1. Kreuzen Sie an, auf welche Merkmalseigenschaften der folgende Box-Plot *am ehesten* hindeutet:



(a) leptokurtisch und rechts**steil** ☐
(b) leptokurtisch und links**steil** ☐
(c) platykurtisch und rechts**steil** ☐
(d) platykurtisch und links**steil** ☐

2. Die 4-stellige Gewinnzahl einer Lotterie wird durch Ziehen **ohne Zurücklegen unter Berücksichtigung der Reihenfolge** aus einer Urne mit den Ziffern $\{1, 2, \ldots, 9\}$ gebildet. Dann beträgt die Anzahl der möglichen Gewinnzahlen insgesamt:

(a) $(9)_4 = \frac{9!}{5!}$ ☐
(b) $\binom{9}{4} = \frac{9!}{4! \cdot 5!}$ ☐
(c) 9^4 ☐
(d) 4^9 ☐

3. Die Anzahl der verschiedenen 4-stelligen Zahlen, die aus den Ziffern 3, 7, 9 und 9 gebildet werden können (eine der möglichen Zahlen ist also 9973), beträgt:

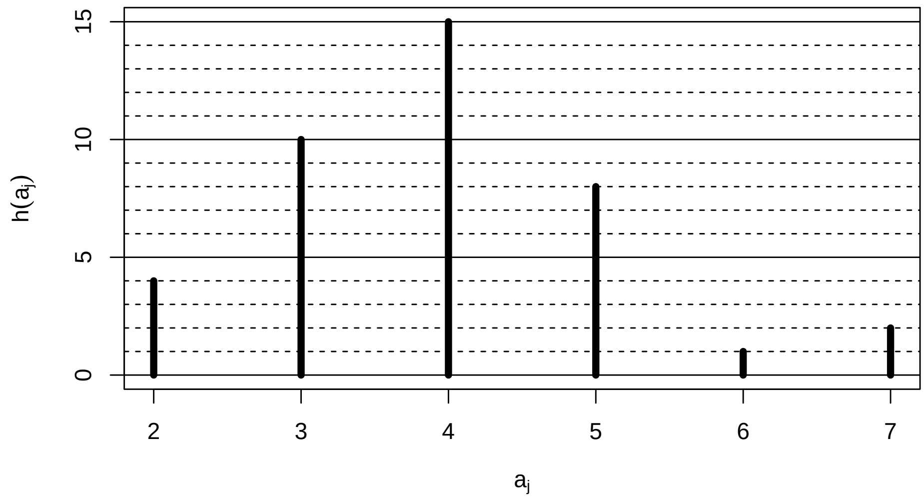
(a) $\frac{3799!}{3! \cdot 7! \cdot 9! \cdot 9!}$ ☐
(b) $\frac{4!}{1! \cdot 1! \cdot 2!}$ ☐
(c) $\frac{28!}{3! \cdot 7! \cdot 9! \cdot 9!}$ ☐
(d) $\frac{28!}{1! \cdot 1! \cdot 2!}$ ☐

4. Für $n \in \mathbb{N}$ seien die Zufallsvariablen $X_1, \ldots, X_n$ unabhängig identisch verteilt mit $\mathrm{E}(X_i) = \mu$ und $\mathrm{Var}(X_i) = \sigma^2$ für alle $i \in \{1, \ldots, n\}$. Dann gilt für $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$

(a) $\mathrm{E}(\overline{X}) = \frac{\mu}{n}$ und $\mathrm{Var}(\overline{X}) = \frac{\sigma^2}{n}$. ☐
(b) $\mathrm{E}(\overline{X}) = \mu$ und $\mathrm{Var}(\overline{X}) = \frac{\sigma^2}{n}$. ☐
(c) $\mathrm{E}(\overline{X}) = \frac{\mu}{n}$ und $\mathrm{Var}(\overline{X}) = \sigma^2$. ☐
(d) $\mathrm{E}(\overline{X}) = \mu$ und $\mathrm{Var}(\overline{X}) = \sigma^2$. ☐

Aufgabe 3 (4 + 5 + 3 + 3 = 15 Punkte)

Zu einem erhobenen Merkmal X sei das folgende Stabdiagramm gegeben:



(a) Erstellen Sie eine Tabelle der absoluten und relativen Häufigkeiten.

(b) Berechnen Sie den arithmetischen Mittelwert und die empirische Varianz des Merkmals X.

(c) Geben Sie die empirische Verteilungsfunktion des Merkmals X an.

(d) Berechnen Sie ein unteres Quartil, ein oberes Quartil und den zugehörigen Interquartilsabstand des Merkmals X.

Aufgabe 4 (6 + 4 + 3 + 3 = 16 Punkte)

Zu einem kardinalskalierten Merkmal sei die folgende (zur einfacheren Bearbeitung der Aufgabe bereits sortierte) Urliste der Länge $n = 30$ gegeben:

22.73, 23.57, 26.90, 29.07, 30.76, 30.77, 36.97, 37.46, 37.90, 38.26, 39.24, 39.52, 39.78, 39.81, 40.14, 41.10, 49.05, 52.59, 57.52, 58.16, 58.90, 59.32, 59.57, 60.60, 60.60, 61.24, 62.18, 65.46, 72.75, 73.65

(a) Führen Sie eine Klassierung der erhobenen Daten auf Grundlage der Klassen

$$K_1 = (5, 35],\ K_2 = (35, 55],\ K_3 = (55, 65],\ K_4 = (65, 75]$$

durch. Geben Sie insbesondere die jeweiligen Klassenbreiten, Klassenmitten, absoluten und relativen Klassenhäufigkeiten, Häufigkeitsdichten sowie die Werte der empirischen Verteilungsfunktion an den Klassengrenzen an.

(b) Stellen Sie die (approximative) empirische Verteilungsfunktion des Merkmals aus der Klassierung der Daten auf.

(c) Berechnen Sie aus den klassierten Daten den (approximativen) arithmetischen Mittelwert der Daten. Wie groß ist die relative Abweichung vom tatsächlichen (aus der Urliste bestimmten) Mittelwert von 46.852?

(d) Bestimmen Sie (unter Verwendung der bekannten Konvention zur eindeutigen Festlegung von Quantilen bei kardinalskalierten Merkmalen) *den* Median sowohl exakt aus der Urliste als auch approximativ mit Hilfe der Verteilungsfunktion für die klassierten Daten.

Aufgabe 5 $(1 + 2 + 2 = 5$ Punkte)

Eine Urne enthält 200 Kugeln, von denen 10 blau und gestreift, 90 rosa und gestreift, 30 blau und kariert sowie 70 rosa und kariert sind. Wie groß ist die Wahrscheinlichkeit dafür, dass

(a) eine rein zufällig aus der Urne entnommene Kugel blau und gestreift ist?

(b) eine rein zufällig aus der Urne entnommene Kugel rosa ist?

(c) eine rein zufällig aus der Urne entnommene Kugel kariert ist, wenn man weiß, dass sie rosa ist?

Aufgabe 6 $(6 + 2 + 2 = 10$ Punkte)

In einer Abteilung einer Finanzbehörde werden eingehende Einkommenssteuererklärungen zufällig auf die Mitarbeiter A, B, C und D aufgeteilt. Aufgrund unterschiedlicher Ausführungsgeschwindigkeiten werden 30% der Erklärungen von Mitarbeiter A, 30% der Erklärungen von Mitarbeiter B, 15% der Erklärungen von Mitarbeiter C und 25% der Erklärungen von Mitarbeiter D bearbeitet. Gegen die ausgestellten Steuerbescheide werden mit einer Wahrscheinlichkeit von 3% bei Mitarbeiter A, 4% bei Mitarbeiter B, 6% bei Mitarbeiter C und 5% bei Mitarbeiter D (erfolgreich) Einsprüche eingelegt.

(a) Wie groß ist die Wahrscheinlichkeit, dass gegen einen zufällig ausgewählten Einkommenssteuerbescheid ein (erfolgreicher) Einspruch eingelegt wird?

(b) Wie groß ist die Wahrscheinlichkeit, dass ein nicht (erfolgreich) per Einspruch beanstandeter Einkommenssteuerbescheid von Mitarbeiter B ausgestellt wurde?

(c) Sind die Ereignisse „Bescheid wird (erfolgreich) beanstandet“ und „Mitarbeiter B hat den Bescheid erstellt“ stochastisch unabhängig? Begründen Sie Ihre Antwort!

Aufgabe 7 $(3 + 2 + 12 + 1 + 4 = 22$ Punkte)

Die Verteilung einer stetigen Zufallsvariablen X sei durch die folgende Verteilungsfunktion gegeben:

$$F_X(x) = \begin{cases} 0 & \text{für } x \leq -2 \\ -\frac{1}{4}x^2 - \frac{1}{2}x & \text{für } -2 < x \leq -1 \\ \frac{3}{16}x^2 + \frac{3}{8}x + \frac{7}{16} & \text{für } -1 < x \leq 1 \\ 1 & \text{für } x > 1 \end{cases}$$

(a) Berechnen Sie eine Dichtefunktion f_X von X.

(b) Berechnen Sie $P(\{X < 0\})$ und $P(\{0 \leq X \leq 1\})$.

(c) Bestimmen Sie den Erwartungswert $\mathrm{E}(X)$ und die Varianz $\mathrm{Var}(X)$.

(d) Ist X symmetrisch um ihren Erwartungswert verteilt (ohne Begründung)?

(e) Bestimmen Sie den Median von X.

Aufgabe 8 $(2 + 9 + 1 + 3 = 15$ Punkte)

Gegeben sei die folgende Tabelle der gemeinsamen Wahrscheinlichkeitsverteilung zu einem zweidimensionalen diskreten Zufallsvektor (X, Y):

$X \setminus Y$	0	1	2	$p_{i\cdot}$
1	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{16}$	
2	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{4}$	
4	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{16}$	
$p_{\cdot j}$				

(a) Ergänzen Sie die obige Tabelle (in den vorgesehenen Feldern) um ihre Randverteilungen.

(b) Berechnen Sie $\mathrm{E}(X)$, $\mathrm{E}(Y)$, $\mathrm{Var}(X)$, $\mathrm{Var}(Y)$, $\mathrm{Cov}(X, Y)$ sowie $\mathrm{Korr}(X, Y)$.

(c) Sind X und Y stochastisch unabhängig?

(d) Berechnen Sie $\mathrm{E}(3X - 4Y)$ sowie $\mathrm{Var}(3X - 4Y)$.

Aufgabe 9 $(1 + 4 + 4 = 9$ Punkte)

In einem Hotel mit 300 Zimmern können Zimmerreservierungen bis zum Anreisetag kostenlos storniert werden. Man weiß aus Erfahrung, dass im Mittel 13% der reservierten Zimmer tatsächlich kurzfristig storniert werden. Um die Zahl der freien Zimmer möglichst gering zu halten, nimmt das Hotel daher mehr Zimmerreservierungen an als Zimmer im Hotel vorhanden sind.

(a) Wie ist die Anzahl Y der tatsächlich wegen Reservierungen benötigten (also nicht stornierten) Zimmer verteilt, wenn insgesamt 330 Zimmerreservierungen angenommen wurden und davon ausgegangen werden kann, dass das Stornierungsverhalten der Hotelgäste voneinander unabhängig ist?

(b) Berechnen Sie unter Verwendung des zentralen Grenzwertsatzes die Wahrscheinlichkeit dafür, dass bei 330 angenommenenen Reservierungen genügend Zimmer zur Verfügung stehen, um alle Hotelgäste, die reserviert und nicht storniert haben, auch im Hotel unterzubringen.

(c) Verwenden Sie den zentralen Grenzwertsatz, um näherungsweise ein 0.9-Quantil der Anzahl in Anspruch genommener Zimmerreservierungen Y zu bestimmen.

Hinweis: Verwenden Sie zur Bearbeitung von Aufgabenteil (b) und (c) die Tabelle zur Standardnormalverteilung auf Seite 120!

4 Deskr. Statistik und Wahrscheinlichkeitsrechnung SS 2016

Aufgabe 1 (16 Punkte)

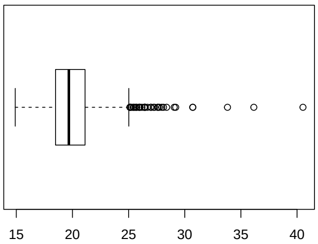
Markieren Sie jeweils mit einem Kreuz pro Aussage im betreffenden Kästchen, ob die unten stehenden Aussagen wahr oder falsch sind.
Richtige Antworten geben 2 Punkte, falsche Antworten und nicht bearbeitete Aussagen 0 Punkte (Aussagen mit zwei Kreuzen zählen als nicht bearbeitet!).

	wahr	falsch
1. Die Flächeninhalte der Rechtecke eines Histogramms addieren sich immer zu 1.	☐	☐
2. Ist (a_i, b_j) eine Ausprägung des zweidimensionalen Merkmals (X, Y), dann sind sowohl die relative Randhäufigkeit $r(a_i)$ des Merkmals X für a_i als auch die relative Randhäufigkeit $r(b_j)$ des Merkmals Y für b_j stets positiv.	☐	☐
3. Das Porto eines Standardbriefs hat sich in Deutschland Anfang letzten Jahres um 3.33% sowie Anfang diesen Jahres um 12.90% erhöht. Die durchschnittliche Preissteigerung der vergangenen beiden Jahre beträgt also (gerundet) 8.01%.	☐	☐
4. Es seien A, B und C drei Ereignisse in einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ mit $P(C) = 0.5$, $P(B\|C) = 0.6$ und $P(A\|B \cap C) = 0.4$. Damit gilt $P(A \cap B \cap C) = 0.12$.	☐	☐
5. Die Wahrscheinlichkeit, beim 5-maligen Würfeln mit einem (fairen) Würfel lauter unterschiedliche Punktzahlen zu erhalten, ist genauso groß wie beim 6-maligen Würfeln.	☐	☐
6. Für eine Zufallsvariable X gelte $P(X \in \{0, 8, 15\}) = 0$. Damit ist X eine stetige Zufallsvariable.	☐	☐
7. Die beiden Zufallsvariablen X und Y seien stochastisch unabhängig. Dann gilt (falls alle beteiligten Momente existieren) stets $\mathrm{E}(X \cdot Y) = \mathrm{E}(X) \cdot \mathrm{E}(Y)$.	☐	☐
8. Sind X und Y zwei Zufallsvariablen mit positiven Varianzen, so stimmt der Korrelationskoeffizient von X und Y mit der Kovarianz der standardisierten Zufallsvariablen überein, es gilt also: $\mathrm{Korr}(X, Y) = \mathrm{Cov}\left(\frac{X - \mathrm{E}(X)}{\sqrt{\mathrm{Var}(X)}}, \frac{Y - \mathrm{E}(Y)}{\sqrt{\mathrm{Var}(Y)}}\right)$	☐	☐

Aufgabe 2 (12 Punkte)

Markieren Sie jeweils die korrekte Antwort mit einem Kreuz im betreffenden Kästchen. Es ist jeweils genau ein Kreuz korrekt.
Richtige Antworten geben 3 Punkte, falsche Antworten und nicht bearbeitete Aufgabenteile 0 Punkte (Aufgabenteile mit mehr als einem Kreuz zählen als nicht bearbeitet!).

1. Kreuzen Sie an, auf welche Merkmalseigenschaften der folgende Box-Plot *am ehesten* hindeutet:



(a) leptokurtisch und rechts**steil** ☐
(b) leptokurtisch und links**steil** ☐
(c) platykurtisch und rechts**steil** ☐
(d) platykurtisch und links**steil** ☐

2. Für die ersten fünf Elfmeter eines Elfmeterschießens werden aus einer Mannschaft 5 von 11 Spielern als Schützen ausgewählt sowie die Reihenfolge festgelegt, in der diese 5 Schützen antreten. Die Anzahl der hierfür möglichen Konstellationen beträgt:

(a) 11^5 ☐
(b) 5^{11} ☐
(c) $\binom{11}{5} = \frac{11!}{5! \cdot 6!}$ ☐
(d) $(11)_5 = \frac{11!}{6!}$ ☐

3. Beim Zufallsexperiment des einmaligen Würfelns mit einem gewöhnlichen 6-seitigen Würfel

(a) sind $\{2\}$ und $\{5, 6\}$ jeweils Ergebnisse. ☐
(b) sind $\{2\}$ und $\{5, 6\}$ jeweils Ereignisse. ☐
(c) ist $\{2\}$ ein Ergebnis und $\{5, 6\}$ ein Ereignis. ☐
(d) ist $\{2\}$ ein Ereignis und $\{5, 6\}$ ein Ergebnis. ☐

4. Sind X und Y zwei stochastisch unabhängige Zufallsvariablen mit $X \sim B(20, 0.25)$ und $Y \sim B(30, 0.25)$, dann ist die Verteilung von $X + Y$ eine

(a) $B(50, 0.25)$-Verteilung. ☐
(b) $B(50, 0.5)$-Verteilung. ☐
(c) $B(25, 0.25)$-Verteilung. ☐
(d) $B(25, 0.5)$-Verteilung. ☐

Aufgabe 3 (1 + 4 + 1 + 5 + 2 = 13 Punkte)

Zu einem erhobenen Merkmal X sei die folgende empirische Verteilungsfunktion gegeben:

$$F(x) = \begin{cases} 0.00 & \text{für} \quad x < 2 \\ 0.08 & \text{für} \quad 2 \le x < 3 \\ 0.22 & \text{für} \quad 3 \le x < 4 \\ 0.50 & \text{für} \quad 4 \le x < 5 \\ 0.72 & \text{für} \quad 5 \le x < 6 \\ 0.96 & \text{für} \quad 6 \le x < 7 \\ 1.00 & \text{für} \quad x \ge 7 \end{cases}$$

Außerdem sei die Länge der Urliste $n = 50$ bekannt.

(a) Geben Sie die Menge A der Merkmalsausprägungen an.

(b) Erstellen Sie eine Tabelle der absoluten und relativen Häufigkeiten.

(c) Wie groß ist der Anteil der Urlisteneinträge, die Werte von mehr als 5 annehmen?

(d) Berechnen Sie den arithmetischen Mittelwert und die empirische Varianz des Merkmals X.

(e) Bestimmen Sie ein unteres Quartil und einen Median des Merkmals X.

Aufgabe 4 (6 + 4 + 3 + 3 + 3 = 19 Punkte)

Zu einem kardinalskalierten Merkmal sei die folgende (zur einfacheren Bearbeitung der Aufgabe bereits sortierte) Urliste der Länge $n = 25$ gegeben:

20.28, 28.78, 29.43, 35.63, 37.90, 39.63, 48.54, 51.15, 51.81, 53.52, 56.11, 59.21, 61.27, 62.47, 64.88, 65.02, 73.04, 73.72, 78.42, 79.65, 89.86, 91.70, 93.73, 97.13, 98.89

(a) Führen Sie eine Klassierung der erhobenen Daten auf Grundlage der Klassen

$$K_1 = (10, 40],\ K_2 = (40, 60],\ K_3 = (60, 80],\ K_4 = (80, 100]$$

durch. Geben Sie insbesondere die jeweiligen Klassenbreiten, Klassenmitten, absoluten und relativen Klassenhäufigkeiten, Häufigkeitsdichten sowie die Werte der empirischen Verteilungsfunktion an den Klassengrenzen an.

(b) Stellen Sie die (approximative) empirische Verteilungsfunktion des Merkmals aus der Klassierung der Daten auf.

(c) Berechnen Sie aus den klassierten Daten den (approximativen) arithmetischen Mittelwert der Daten. Wie groß ist die relative Abweichung vom tatsächlichen (aus der Urliste bestimmten) Mittelwert von 61.671?

(d) Bestimmen Sie (aus der Urliste) die Anzahl von Merkmalswerten zwischen 40 und 90. Welche Näherung für diese Anzahl können Sie aus der in Teil (b) aufgestellten empirischen Verteilungsfunktion berechnen?

(e) Bestimmen Sie (unter Verwendung der bekannten Konvention zur eindeutigen Festlegung von Quantilen bei kardinalskalierten Merkmalen) *den* Median sowohl exakt aus der Urliste als auch approximativ mit Hilfe der Verteilungsfunktion für die klassierten Daten.

Aufgabe 5 (6 + 2 + 1 = 9 Punkte)

In einer Fußballmannschaft sind die vier Mitspieler Andreas, Bastian, Christian und Daniel für die Ausführung von Eckstößen zuständig. Dabei werden im Mittel 20% der Eckstöße von Andreas, 30% der Eckstöße von Bastian, 30% der Eckstöße von Christian und 20% der Eckstöße von Daniel ausgeführt. Aus einer ausführlichen statistischen Auswertung ist bekannt, dass 7% der Eckstöße von Andreas, 12% der Eckstöße von Bastian, 10% der Eckstöße von Christian und 10% der Eckstöße von Daniel zu einem Tor führen.

(a) Wie groß ist die Wahrscheinlichkeit, dass ein zufällig ausgewählter Eckstoß zu einem Tor führt?

(b) Wie groß ist die Wahrscheinlichkeit, dass ein Eckstoß, der nicht zu einem Tor geführt hat, von Christian ausgeführt wurde?

(c) Sind die Ereignisse „Eckstoß führt zu einem Tor" und „Eckstoß wird von Christian ausgeführt" stochastisch unabhängig?

Aufgabe 6 (5 + 2 + 6 + 1 + 4 = 18 Punkte)

Die Verteilung einer stetigen Zufallsvariablen X sei durch die folgende Dichtefunktion gegeben:

$$f_X(x) = \begin{cases} \frac{1}{2}x + 1 & \text{für } -2 \leq x < -1 \\ -\frac{1}{6}x + \frac{1}{3} & \text{für } -1 \leq x \leq 2 \\ 0 & \text{sonst} \end{cases}$$

(a) Berechnen Sie die Verteilungsfunktion F_X von X.

(b) Berechnen Sie $P(\{X < 0\})$ und $P(\{-1 < X \leq 1\})$.

(c) Bestimmen Sie den Erwartungswert $\mathrm{E}(X)$.

(d) Ist X symmetrisch um ihren Erwartungswert verteilt (ohne Begründung)?

(e) Bestimmen Sie das obere Quartil von X.

Aufgabe 7 (2 + 3 = 5 Punkte)

Als Hausaufgabe im Fach Geschichte waren die Geburtsdaten von 25 bekannten Entdeckern auswendig zu lernen. Die Schülerin Ella Emsig hat 20 dieser Geburtsdaten auswendig gelernt (die Chance, bei den anderen 5 Geburtsdaten durch Raten eine richtige Antwort zu geben, sei gleich Null). Der Lehrer überprüft, ob Ella die Hausaufgabe ordentlich erledigt hat, indem er 4 Mal rein zufällig und unabhängig voneinander einen der Entdecker auswählt und die zugehörigen Geburtsdaten abfragt. Kann Ella mindestens zu 3 dieser 4 Entdecker die Geburtsdaten korrekt angeben, so ist die Überprüfung bestanden.

(a) Welche Verteilung besitzt die Anzahl der von Ella abgegebenen richtigen Antworten?

(b) Mit welcher Wahrscheinlichkeit besteht Ella die Überprüfung?

Aufgabe 8 $(2 + 3 + 9 + 1 + 3 = 18$ Punkte)

Gegeben sei die folgende Tabelle der gemeinsamen Wahrscheinlichkeitsverteilung zu einem zweidimensionalen diskreten Zufallsvektor (X, Y):

$X \setminus Y$	3	4	6	$p_{i\cdot}$
1	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{6}$	
2	$\frac{1}{6}$	$\frac{1}{6}$	0	
3	$\frac{1}{12}$	$\frac{1}{4}$	0	
$p_{\cdot j}$				

(a) Ergänzen Sie die obige Tabelle (in den vorgesehenen Feldern) um ihre Randverteilungen.

(b) Geben Sie die bedingten Wahrscheinlichkeitsverteilungen von Y unter der Bedingung $X = x_i$ für alle $x_i \in T(X)$ über die zugehörigen (bedingten) Wahrscheinlichkeitsfunktionen an.

(c) Berechnen Sie $\mathrm{E}(X)$, $\mathrm{E}(Y)$, $\mathrm{Var}(X)$, $\mathrm{Var}(Y)$, $\mathrm{Cov}(X, Y)$ sowie $\mathrm{Korr}(X, Y)$.

(d) Sind X und Y stochastisch unabhängig?

(e) Berechnen Sie $\mathrm{E}(-3X + 6Y)$ sowie $\mathrm{Var}(-3X + 6Y)$.

Aufgabe 9 $(2 + 4 + 4 = 10$ Punkte)

Die Zufallsvariablen $X_1, \ldots, X_{32}$ seien unabhängig identisch Pois(8)-verteilt. Die Summe der Zufallsvariablen X_i sei mit

$$Y := \sum_{i=1}^{32} X_i = X_1 + \ldots + X_{32}$$

bezeichnet.

(a) Geben Sie die (exakte) Verteilung von Y sowie deren Erwartungswert $\mathrm{E}(Y)$ und Varianz $\mathrm{Var}(Y)$ an.

(b) Verwenden Sie den zentralen Grenzwertsatz, um näherungsweise zu bestimmen, mit welcher Wahrscheinlichkeit Y Werte zwischen 240 und 260 annimmt.

(c) Verwenden Sie den zentralen Grenzwertsatz, um näherungsweise einen um den Erwartungswert von Y symmetrischen Bereich zu bestimmen, in dem sich Y mit einer Wahrscheinlichkeit von 0.95 realisiert.

Hinweis: Verwenden Sie zur Bearbeitung von Aufgabenteil (b) und (c) die Tabelle zur Standardnormalverteilung auf Seite 120!

5 Deskr. Statistik und Wahrscheinlichkeitsrechnung WS 2016/17

Aufgabe 1 (16 Punkte)

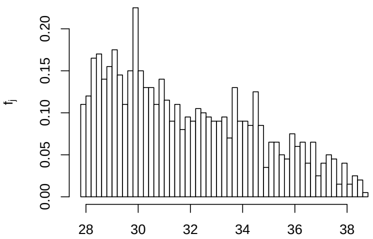
Markieren Sie jeweils mit einem Kreuz pro Aussage im betreffenden Kästchen, ob die unten stehenden Aussagen wahr oder falsch sind.
Richtige Antworten geben 2 Punkte, falsche Antworten und nicht bearbeitete Aussagen 0 Punkte (Aussagen mit zwei Kreuzen zählen als nicht bearbeitet!).

	wahr	falsch
1. Die Summe der Differenzen aller Urlisteneinträge x_i eines kardinalskalierten Merkmals von deren arithmetischem Mittel $\overline{x}$ ist stets 0, es gilt also (falls n die Länge der Urliste bezeichnet) stets $\sum_{i=1}^{n}(x_i - \overline{x}) = 0$.	☐	☐
2. In einem Histogramm entspricht die Höhe eines Rechtecks stets der relativen Häufigkeit der zugehörigen Klasse.	☐	☐
3. Bei linkssteilen (kardinalskalierten) Merkmalen ist der Median tendenziell größer als das arithmetische Mittel.	☐	☐
4. Die Wahrscheinlichkeit, beim 4-maligen Würfeln mit einem (fairen) Würfel lauter unterschiedliche Punktzahlen zu erhalten, ist größer als 25%.	☐	☐
5. Es seien $(\Omega, \mathcal{F}, P)$ ein Wahrscheinlichkeitsraum, $A, B \in \mathcal{F}$ zwei Ereignisse mit $P(A) = 0.5$ und $P(A \cup B) = 1$. Dann gilt stets $P(B) \geq 0.5$.	☐	☐
6. Sind X und Y Zufallsvariablen mit $\mathrm{E}(X) = 7$ und $\mathrm{E}(Y) = 5$, dann gilt $\mathrm{E}(X + Y) = 12$, auch wenn X und Y stochastisch abhängig sind.	☐	☐
7. Die Summe von 4 stochastisch unabhängigen $N(25, 3^2)$-verteilten Zufallsvariablen ist $N(100, 6^2)$-verteilt.	☐	☐
8. Für die Zufallsvariablen X und Y gelte $\mathrm{Cov}(X, Y) = 0$. Dann sind X und Y stets stochastisch unabhängig.	☐	☐

Aufgabe 2 (12 Punkte)

Markieren Sie jeweils die korrekte Antwort mit einem Kreuz im betreffenden Kästchen. Es ist jeweils genau ein Kreuz korrekt.
Richtige Antworten geben 3 Punkte, falsche Antworten und nicht bearbeitete Aufgabenteile 0 Punkte (Aufgabenteile mit mehr als einem Kreuz zählen als nicht bearbeitet!).

1. Kreuzen Sie an, auf welche Merkmalseigenschaften das folgende Histogramm *am ehesten* hindeutet:



(a) leptokurtisch und rechts**steil** ☐
(b) leptokurtisch und links**steil** ☐
(c) platykurtisch und rechts**steil** ☐
(d) platykurtisch und links**steil** ☐

2. Auf der Bank der Ersatzspieler einer Fußballmannschaft sitzen 7 Spieler. Wenn während des Fußballspiels 3 Ersatzspieler eingewechselt werden und die Reihenfolge der Einwechslungen keine Rolle spielen soll, so beträgt die Anzahl der verschiedenen Einwechslungsmöglichkeiten (für diese Mannschaft) insgesamt:

(a) $(7)_3 = \frac{7!}{4!}$ ☐
(b) $\binom{7}{3} = \frac{7!}{3! \cdot 4!}$ ☐
(c) 7^3 ☐
(d) 3^7 ☐

3. Die Ränge $\text{rg}(X)_1, \ldots, \text{rg}(X)_8$ zur (der Einfachheit halber sortierten) Urliste

$$3, \quad 4, \quad 4, \quad 6, \quad 6, \quad 6, \quad 7, \quad 10$$

des ordinalskalierten Merkmals X lauten:

(a) 1, 2, 2, 3, 3, 3, 4, 5 ☐
(b) 1, 2.5, 2.5, 5, 5, 5, 7, 8 ☐
(c) 1, 2.5, 2.5, 3.5, 3.5, 3.5, 6, 7 ☐
(d) 3, 4.5, 4.5, 6, 6, 6, 7, 10 ☐

4. Die Wahrscheinlichkeit, in dieser Klausuraufgabe (4 MC-Aufgabenteile mit jeweils genau einer korrekten Antwort aus 4 Antwortmöglichkeiten) durch *rein zufälliges* Ankreuzen jeweils einer Antwortmöglichkeit (jede Antwortmöglichkeit erhalte also eine Chance von 25%) genau zwei richtige Antworten zu markieren, beträgt (ggf. auf 2 Nachkommastellen gerundet):

   (a) 3.52% ☐
   (b) 12.50% ☐
   (c) 21.09% ☐
   (d) 25.00% ☐

Aufgabe 3 (4 + 1 + 5 + 2 = 12 Punkte)

Zu einem erhobenen Merkmal X sei die folgende empirische Verteilungsfunktion gegeben:

$$F(x) = \begin{cases} 0.00 & \text{für} \quad x < -1 \\ 0.08 & \text{für} \quad -1 \le x < 0 \\ 0.40 & \text{für} \quad 0 \le x < 1 \\ 0.66 & \text{für} \quad 1 \le x < 2 \\ 0.90 & \text{für} \quad 2 \le x < 3 \\ 1.00 & \text{für} \quad x \ge 3 \end{cases}$$

Außerdem sei die Länge der Urliste $n = 50$ bekannt.

(a) Erstellen Sie eine Tabelle der absoluten und relativen Häufigkeiten.

(b) Wie groß ist der Anteil der Urlisteneinträge, die Werte von mindestens 1 annehmen?

(c) Berechnen Sie den arithmetischen Mittelwert und die empirische Varianz des Merkmals X.

(d) Bestimmen Sie ein unteres Quartil und ein oberes Quartil des Merkmals X.

Aufgabe 4 (6 + 4 + 3 + 3 + 3 = 19 Punkte)

Zu einem kardinalskalierten Merkmal sei die folgende (zur einfacheren Bearbeitung der Aufgabe bereits sortierte) Urliste der Länge $n = 40$ gegeben:

5.50, 5.91, 8.23, 10.27, 10.52, 10.76, 15.96, 16.39, 16.98, 17.85, 18.83, 19.50, 19.78, 20.57, 23.84, 24.13, 24.39, 24.45, 25.29, 25.89, 27.03, 28.08, 30.87, 31.70, 31.90, 32.67, 38.35, 39.94, 42.31, 44.66, 45.47, 47.02, 47.52, 48.01, 49.10, 50.77, 51.20, 51.33, 52.05, 52.09

(a) Führen Sie eine Klassierung der erhobenen Daten auf Grundlage der Klassen

$$K_1 = (5, 15],\ K_2 = (15, 25],\ K_3 = (25, 40],\ K_4 = (40, 60]$$

durch. Geben Sie insbesondere die jeweiligen Klassenbreiten, Klassenmitten, absoluten und relativen Klassenhäufigkeiten, Häufigkeitsdichten sowie die Werte der empirischen Verteilungsfunktion an den Klassengrenzen an.

(b) Stellen Sie die (approximative) empirische Verteilungsfunktion des Merkmals aus der Klassierung der Daten auf.

(c) Berechnen Sie aus den klassierten Daten den (approximativen) arithmetischen Mittelwert der Daten. Wie groß ist die relative Abweichung vom tatsächlichen (aus der Urliste bestimmten) Mittelwert von 29.678?

(d) Bestimmen Sie (aus der Urliste) die Anzahl von Merkmalswerten zwischen 10 und 50. Welche Näherung für diese Anzahl können Sie aus der in Teil (b) aufgestellten empirischen Verteilungsfunktion berechnen?

(e) Bestimmen Sie (unter Verwendung der bekannten Konvention zur eindeutigen Festlegung von Quantilen bei kardinalskalierten Merkmalen) *den* Median sowohl exakt aus der Urliste als auch approximativ mit Hilfe der Verteilungsfunktion für die klassierten Daten.

Aufgabe 5 (2 + 2 + 2 = 6 Punkte)

Bei der Herstellung von Deckenpaneelen tritt mit einer Wahrscheinlichkeit von 4% ein Fehler beim Zuschnitt der Paneele auf, mit einer Wahrscheinlichkeit von 3.5% ein Fehler beim Laminieren der Paneele und mit einer Wahrscheinlichkeit von 2.5% sowohl ein Fehler beim Zuschnitt der Paneele als auch ein Fehler beim Laminieren der Paneele. Berechnen Sie die Wahrscheinlichkeit dafür, dass

(a) höchstens einer der beiden Fehler,

(b) mindestens einer der beiden Fehler,

(c) ein Fehler beim Zuschnitt der Paneele, aber kein Fehler beim Laminieren der Paneele

auftritt.

Aufgabe 6 (6 + 2 = 8 Punkte)

Ein Versandhaus beauftragt für den Versand seiner Sendungen einen von insgesamt vier verschiedenen Versand-Dienstleistern A, B, C und D. Dabei werden durchschnittlich 20% der Sendungen an Dienstleister A, 25% der Sendungen an Dienstleister B, 25% der Sendungen an Dienstleister C und 30% der Sendungen an Dienstleister D übergeben. Die umfangreiche Auswertung der vorhandenen Kunden-Feedbacks zu Qualität und Geschwindigkeit der Lieferung ergab, dass 97% der Lieferungen mit Dienstleister A, 96% der Lieferungen mit Dienstleister B, 97% der Lieferungen mit Dienstleister C und 98% der Lieferungen mit Dienstleister D nicht beanstandet werden.

(a) Wie groß ist die Wahrscheinlichkeit, dass eine zufällig ausgewählte Sendung einen Anlass zur Beanstandung durch den Kunden bietet?

(b) Wie groß ist die Wahrscheinlichkeit, dass eine nicht beanstandete Lieferung mit Dienstleister B versendet wurde?

Aufgabe 7 (5 + 2 + 6 + 1 + 4 = 18 Punkte)

Die Verteilung einer stetigen Zufallsvariablen X sei durch die folgende Dichtefunktion gegeben:

$$f_X(x) = \begin{cases} -\frac{1}{4}x + \frac{1}{4} & \text{für } -1 \leq x < 1 \\ x - 1 & \text{für } 1 \leq x \leq 2 \\ 0 & \text{sonst} \end{cases}$$

(a) Berechnen Sie die Verteilungsfunktion F_X von X.

(b) Berechnen Sie $P\left(\left\{X < -\frac{1}{2}\right\}\right)$ und $P\left(\left\{-\frac{1}{2} \leq X \leq \frac{1}{2}\right\}\right)$.

(c) Bestimmen Sie den Erwartungswert $\mathrm{E}(X)$.

(d) Ist X symmetrisch um ihren Erwartungswert verteilt (ohne Begründung)?

(e) Bestimmen Sie das obere Quartil von X.

Aufgabe 8 (2 + 1 + 2 = 5 Punkte)

Die Wartezeit zwischen zwei Angriffen auf einen Internetrouter lasse sich als eine exponentialverteilte Zufallsvariable auffassen. Im Mittel vergehen zwischen zwei Angriffen 5 Minuten.

(a) Welche Standardabweichung hat die Wartezeit zwischen zwei Angriffen?

(b) Mit welcher Wahrscheinlichkeit beträgt die Wartezeit zwischen zwei Angriffen weniger als 10 Minuten?

(c) Berechnen Sie das 0.90-Quantil der Wartezeit zwischen zwei Angriffen.

Aufgabe 9 (2 + 9 + 1 + 3 = 15 Punkte)

Gegeben sei die folgende Tabelle der gemeinsamen Wahrscheinlichkeitsverteilung zu einem zweidimensionalen diskreten Zufallsvektor (X, Y):

$X \setminus Y$	-1	0	1	$p_{i\cdot}$
2	$\frac{1}{6}$	$\frac{1}{8}$	$\frac{1}{24}$	
3	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{6}$	
4	0	$\frac{1}{24}$	$\frac{7}{24}$	
$p_{\cdot j}$				

(a) Ergänzen Sie die obige Tabelle (in den vorgesehenen Feldern) um ihre Randverteilungen.

(b) Berechnen Sie $\mathrm{E}(X)$, $\mathrm{E}(Y)$, $\mathrm{Var}(X)$, $\mathrm{Var}(Y)$, $\mathrm{Cov}(X, Y)$ sowie $\mathrm{Korr}(X, Y)$.

(c) Sind X und Y stochastisch unabhängig?

(d) Berechnen Sie $\mathrm{E}\,(2X - 4Y)$ sowie $\mathrm{Var}\,(2X - 4Y)$.

Aufgabe 10 (2 + 3 + 4 = 9 Punkte)

Ein Online-Händler bietet für 400 der an einem Tag eingehenden Bestellungen einen Express-Lieferservice an, der eine Abfertigung der Bestellung am nächsten Arbeitstag garantiert. Es ist davon auszugehen, dass die Zeitdauern zur Abfertigung einzelner Express-Bestellungen (in Stunden) unabhängig identisch verteilt sind mit einer mittleren Abfertigungsdauer von 0.2 Stunden bei einer Standardabweichung von 0.06 Stunden.

(a) Welchen Erwartungswert und welche Standardabweichung hat die Summe der Abfertigungsdauern von 400 Express-Bestellungen?

(b) Verwenden Sie den zentralen Grenzwertsatz, um die Wahrscheinlichkeit, dass 400 Express-Bestellungen in höchstens 82 (Mitarbeiter-)Stunden abgefertigt werden können, (näherungsweise) zu berechnen.

(c) Verwenden Sie den zentralen Grenzwertsatz, um näherungsweise einen um den zugehörigen Erwartungswert symmetrischen Bereich zu bestimmen, in dem sich die Gesamtabfertigungsdauer von 400 Express-Bestellungen mit einer Wahrscheinlichkeit von 0.95 realisiert.

Hinweis: Verwenden Sie zur Bearbeitung der Aufgabenteile (b) und (c) die Tabelle zur Standardnormalverteilung auf Seite 120!

6 Deskr. Statistik und Wahrscheinlichkeitsrechnung SS 2017

Aufgabe 1 (16 Punkte)

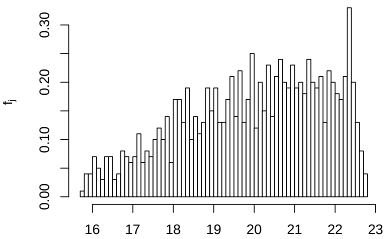
Markieren Sie jeweils mit einem Kreuz pro Aussage im betreffenden Kästchen, ob die unten stehenden Aussagen wahr oder falsch sind.
Richtige Antworten geben 2 Punkte, falsche Antworten und nicht bearbeitete Aussagen 0 Punkte (Aussagen mit zwei Kreuzen zählen als nicht bearbeitet!).

	wahr	falsch
1. In einem Histogramm sind die Höhen der Rechtecke stets proportional zu den relativen Häufigkeiten der zugehörigen Klassen.	☐	☐
2. Ist x_p ein p-Quantil eines kardinalskalierten Merkmals X (mit $0 < p < 1$), dann ist mindestens ein Anteil von $1 - p$ der Urlisteneinträge zum Merkmal X größer als oder gleich x_p.	☐	☐
3. In einer bestimmten Prüfung waren 65% aller durchgefallenen Prüflinge männlich, 35% weiblich. Daraus kann man schließen, dass der Anteil der durchgefallenen unter den männlichen Prüflingen größer ist als unter den weiblichen Prüflingen.	☐	☐
4. Es seien A, B und C drei Ereignisse in einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ mit $P(A) = 0.3$, $P(B) = 0.2$ und $P(C) = 0.4$. Damit gilt $P(A \cup B \cup C) \geq 0.9$.	☐	☐
5. Würfelt man fünf Mal mit einem fairen Würfel, so ist die Anzahl der Einsen geometrisch verteilt mit Parameter $p = 5 \cdot \frac{1}{6} = \frac{5}{6}$.	☐	☐
6. Für die stetige Zufallsvariable X mit Verteilungsfunktion F_X gelte $F_X(x) = 1 - F_X(-x)$ für alle $x \in \mathbb{R}$. Damit ist X symmetrisch um Null verteilt.	☐	☐
7. Die beiden Zufallsvariablen X und Y seien unabhängig und jeweils Poisson-verteilt mit Parameter $\lambda = 10$. Dann ist auch $X+Y$ Poisson-verteilt mit Parameter 10.	☐	☐
8. Sind X und Y zwei Zufallsvariablen mit positiven Varianzen, so ist der Korrelationskoeffizient von $2 \cdot X$ und $2 \cdot Y$ doppelt so groß wie der Korrelationskoeffizient von X und Y.	☐	☐

Aufgabe 2 (12 Punkte)

Markieren Sie jeweils die korrekte Antwort mit einem Kreuz im betreffenden Kästchen. Es ist jeweils genau ein Kreuz korrekt.
Richtige Antworten geben 3 Punkte, falsche Antworten und nicht bearbeitete Aufgabenteile 0 Punkte (Aufgabenteile mit mehr als einem Kreuz zählen als nicht bearbeitet!).

1. Kreuzen Sie an, auf welche Merkmalseigenschaften das folgende Histogramm *am ehesten* hindeutet:



(a) leptokurtisch und rechts**steil**

(b) leptokurtisch und links**steil**

(c) platykurtisch und rechts**steil**

(d) platykurtisch und links**steil**

2. Die Ränge $\mathrm{rg}(X)_1, \ldots, \mathrm{rg}(X)_8$ zur (der Einfachheit halber sortierten) Urliste

$$2, \quad 2, \quad 5, \quad 7, \quad 8, \quad 8, \quad 8, \quad 11$$

des ordinalskalierten Merkmals X lauten:

(a) 1.5, 1.5, 5, 7, 8.5, 8.5, 8.5, 11

(b) 1.5, 1.5, 2, 3, 4.5, 4.5, 4.5, 6

(c) 1.5, 1.5, 3, 4, 6, 6, 6, 8

(d) 1.5, 1.5, 3, 4, 6.5, 6.5, 6.5, 8

3. Es seien X und Y zwei Zufallsvariablen über dem Wahrscheinlichkeitraum $(\Omega, \mathcal{F}, P)$ mit existierenden positiven Varianzen $\mathrm{Var}(X)$ und $\mathrm{Var}(Y)$ sowie existierender Kovarianz $\mathrm{Cov}(X, Y)$. Dann gilt für die Unkorreliertheit von X und Y sowie die stochastische Unabhängigkeit von X und Y:

(a) Es gibt keinen Zusammenhang zwischen Unkorreliertheit und Unabhängigkeit.

(b) Aus der stochastischen Unabhängigkeit folgt die Unkorreliertheit.

(c) Aus der Unkorreliertheit folgt die stochastische Unabhängigkeit.

(d) Unkorreliertheit und stochastische Unabhängigkeit sind äquivalent.

4. Sind X_1, X_2 und X_3 drei stochastisch unabhängige Zufallsvariablen mit $X_1 \sim N(30, 5^2)$, $X_2 \sim N(50, 10^2)$ und $X_3 \sim N(70, 10^2)$, dann ist die Verteilung von $X_1 + X_2 + X_3$ eine

(a) $N(50, 25^2)$-Verteilung.

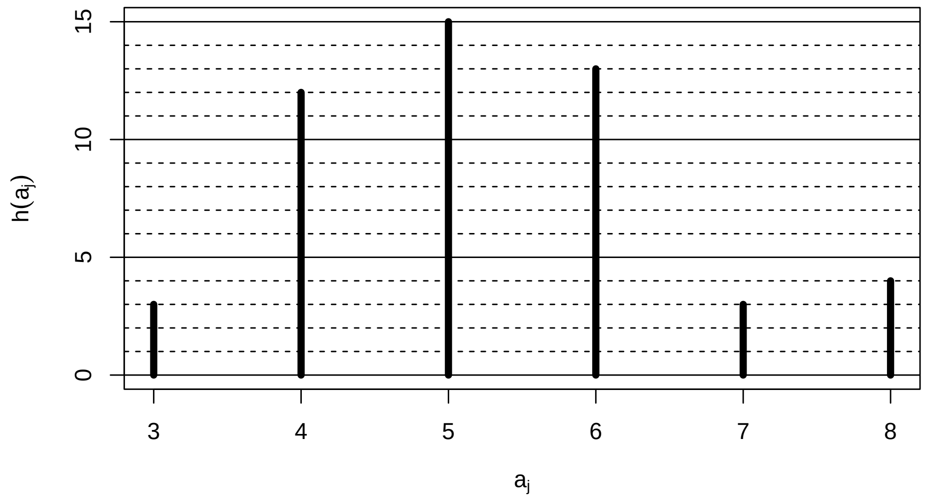
(b) $N(150, 25^2)$-Verteilung.

(c) $N(50, 15^2)$-Verteilung.

(d) $N(150, 15^2)$-Verteilung.

Aufgabe 3 (4 + 5 + 3 + 1 + 2 = 15 Punkte)

Zu einem erhobenen Merkmal X sei das folgende Stabdiagramm gegeben:



(a) Erstellen Sie eine Tabelle der absoluten und relativen Häufigkeiten.

(b) Berechnen Sie den arithmetischen Mittelwert und die empirische Varianz des Merkmals X.

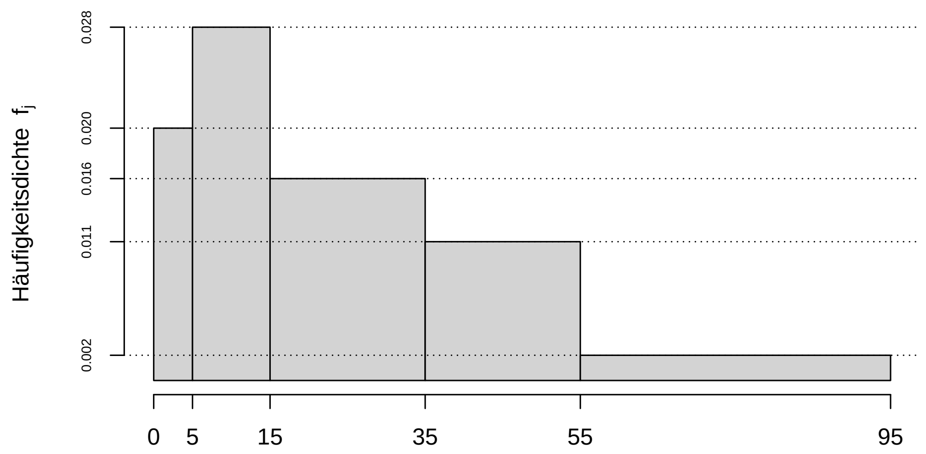
(c) Geben Sie die empirische Verteilungsfunktion des Merkmals X an.

(d) Wie groß ist der Anteil der Urlisteneinträge, die Werte von mehr als 5 annehmen?

(e) Berechnen Sie ein unteres Quartil und ein oberes Quartil des Merkmals X.

Aufgabe 4 (7 + 4 + 3 + 2 + 2 = 18 Punkte)

Gegeben sei das folgende Histogramm zur Klassierung einer Urliste der Länge $n = 50$:



(a) Rekonstruieren Sie die Klassierung der Daten aus dem Histogramm. Geben Sie insbesondere die jeweiligen Klassenbreiten, Klassenmitten, absoluten und relativen Klassenhäufigkeiten, Häufigkeitsdichten sowie die Werte der empirischen Verteilungsfunktion an den Klassengrenzen an.

(b) Stellen Sie die (approximative) empirische Verteilungsfunktion des Merkmals aus der Klassierung der Daten auf.

(c) Berechnen Sie aus den klassierten Daten den (approximativen) arithmetischen Mittelwert der Daten. Wie groß ist die relative Abweichung vom tatsächlichen (aus der Urliste bestimmten) Mittelwert von 27.638?

(d) Welche Näherung für die Anzahl von Merkmalswerten zwischen 20 und 70 können Sie unter Verwendung der approximativen Verteilungsfunktion für die klassierten Daten berechnen?

(e) Bestimmen Sie näherungsweise unter Verwendung der approximativen Verteilungsfunktion für die klassierten Daten den Median.

Aufgabe 5 (4 + 2 = 6 Punkte)

Der 23-köpfige Kader einer Fußballmannschaft bestehe aus drei Torwarten, acht Abwehrspielern, sieben Mittelfeldspielern und fünf Stürmern.

(a) Wie viele Möglichkeiten hat der Trainer, die Startelf – bestehend aus einem Torwart, vier Abwehrspielern, vier Mittelfeldspielern und zwei Stürmern – zusammenzustellen?

(b) Nehmen Sie nun an, dass der Trainer die Zusammenstellung gemäß Teil (a) rein zufällig bestimmt, und dass zwei der fünf Stürmer bereits mit einer gelben Karte vorbelastet sind. Wie groß ist die Wahrscheinlichkeit, dass genau diese beiden Stürmer bei der zufälligen Zusammenstellung der Startelf ausgewählt werden?

Hinweis: Sie können diesen Aufgabenteil auch ohne die Bearbeitung von Aufgabenteil (a) lösen!

Aufgabe 6 (6 + 2 + 2 = 10 Punkte)

Ein Hersteller von Tiefkühlfertiggerichten bezieht seine Frischfischlieferungen von den vier Großhändlern A, B, C und D. Dabei werden einzelne Lieferungen mit einer Wahrscheinlichkeit von 20% von Lieferant A, 50% von Lieferant B, 15% von Lieferant C und 15% von Lieferant D geliefert. Bei den anschließenden Qualitätskontrollen gibt es erfahrungsgemäß mit einer Wahrscheinlichkeit von 2.5% bei Lieferant A, 2% bei Lieferant B, 7% bei Lieferant C und 3% bei Lieferant D Anlass zu Beanstandungen.

(a) Wie groß ist die Wahrscheinlichkeit, dass eine zufällig ausgewählte Lieferung in der Qualitätskontrolle beanstandet wird?

(b) Wie groß ist die Wahrscheinlichkeit, dass eine von der Qualitätskontrolle nicht beanstandete Lieferung von Großhändler D geliefert wurde?

(c) Sind die Ereignisse „Lieferung wird beanstandet“ und „Lieferung stammt von Großhändler D“ stochastisch unabhängig? Begründen Sie Ihre Antwort!

Aufgabe 7 (5 + 2 + 6 + 4 = 17 Punkte)

Die Verteilung einer stetigen Zufallsvariablen X sei durch die folgende Dichtefunktion gegeben:

$$f_X(x) = \begin{cases} \frac{1}{8}x + \frac{1}{8} & \text{für } -1 \leq x < 1 \\ -\frac{1}{8}x + \frac{5}{8} & \text{für } 1 \leq x \leq 3 \\ 0 & \text{sonst} \end{cases}$$

(a) Berechnen Sie die Verteilungsfunktion F_X von X.

(b) Berechnen Sie $P(\{X > 0\})$ und $P(\{1 \leq X \leq 2\})$.

(c) Bestimmen Sie den Erwartungswert $\mathrm{E}(X)$.

(d) Bestimmen Sie den Median von X.

Aufgabe 8 (2 + 3 + 9 + 1 + 3 = 18 Punkte)

Gegeben sei die folgende Tabelle der gemeinsamen Wahrscheinlichkeitsverteilung zu einem zweidimensionalen diskreten Zufallsvektor (X, Y):

$X \setminus Y$	1	3	5	$p_{i\cdot}$
-1	0.05	0.08	0.12	
1	0.1	0.22	0.13	
3	0.2	0.05	0.05	
$p_{\cdot j}$				

(a) Ergänzen Sie die obige Tabelle (in den vorgesehenen Feldern) um ihre Randverteilungen.

(b) Geben Sie die bedingten Wahrscheinlichkeitsverteilungen von X unter der Bedingung $Y = y_j$ für alle $y_j \in T(Y)$ über die zugehörigen (bedingten) Wahrscheinlichkeitsfunktionen an.

(c) Berechnen Sie $\mathrm{E}(X)$, $\mathrm{E}(Y)$, $\mathrm{Var}(X)$, $\mathrm{Var}(Y)$, $\mathrm{Cov}(X, Y)$ sowie $\mathrm{Korr}(X, Y)$.

(d) Sind X und Y stochastisch unabhängig?

(e) Berechnen Sie $\mathrm{E}(3X - 2Y)$ sowie $\mathrm{Var}(3X - 2Y)$.

Aufgabe 9 (2 + 2 + 4 = 8 Punkte)

Ein freiberuflicher Netzwerktechniker benötigt für das Auflegen eines Netzwerkanschlusses im Mittel 12 Minuten bei einer Standardabweichung von 3 Minuten. Man kann davon ausgehen, dass die benötigten Zeitdauern für die einzelnen Anschlüsse nicht gegenseitig voneinander abhängen. In einem bestimmten Schaltschrank sind insgesamt 36 Anschlüsse aufzulegen.

(a) Geben Sie die den Erwartungswert sowie die Standardabweichung der gesamten Arbeitszeit (für alle 36 Netzwerkanschlüsse) an.

(b) Verwenden Sie den zentralen Grenzwertsatz, um näherungsweise zu bestimmen, mit welcher Wahrscheinlichkeit der Netzwerktechniker nicht länger als 7.5 Stunden bzw. 450 Minuten zum Auflegen aller Anschlüsse benötigt.

(c) Verwenden Sie den zentralen Grenzwertsatz, um näherungsweise einen um den zugehörigen Erwartungswert symmetrischen Bereich zu bestimmen, in dem sich die gesamte Arbeitszeit für 36 Netzwerkanschlüsse mit einer Wahrscheinlichkeit von 0.95 realisiert.

Hinweis: Verwenden Sie zur Bearbeitung von Aufgabenteil (b) und (c) die Tabelle zur Standardnormalverteilung auf Seite 120!

7 Deskr. Statistik und Wahrscheinlichkeitsrechnung WS 2017/18

Aufgabe 1 (16 Punkte)

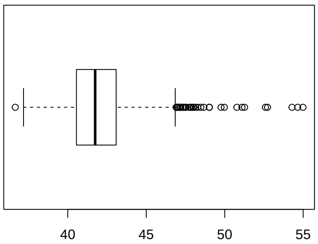
Markieren Sie jeweils mit einem Kreuz pro Aussage im betreffenden Kästchen, ob die unten stehenden Aussagen wahr oder falsch sind.
Richtige Antworten geben 2 Punkte, falsche Antworten und nicht bearbeitete Aussagen 0 Punkte (Aussagen mit zwei Kreuzen zählen als nicht bearbeitet!).

		wahr	falsch
1.	Stetige Merkmale sind stets kardinalskaliert.	☐	☐
2.	Empirische Verteilungsfunktionen F sind stets streng monoton wachsende Funktionen.	☐	☐
3.	Die jährlichen Inflationsraten in Deutschland betrugen in den Jahren 2014–2017 im Einzelnen 0.9%, 0.3%, 0.5% und 1.8%. Damit beträgt die durchschnittliche jährliche Inflationsrate in diesem Zeitraum (gerundet) 0.875%.	☐	☐
4.	Die Wahrscheinlichkeit, mit einem fairen (6-seitigen) Würfel dreimal nacheinander **keine** 6 zu würfeln, beträgt weniger als 60%.	☐	☐
5.	Wenn Sie bei der Bearbeitung von Aufgabe 2 dieser Klausur (Vier MC-Teilaufgaben mit jeweils vier Antwortmöglichkeiten) jeweils genau eine der vier möglichen Antworten ankreuzen, haben Sie insgesamt 256 Bearbeitungsmöglichkeiten.	☐	☐
6.	Es seien $(\Omega, \mathcal{F}, P)$ ein Wahrscheinlichkeitsraum, $A, B \in \mathcal{F}$ zwei Ereignisse mit $P(A) > 0$, $P(B) > 0$ sowie $P(A) = P(A\|B)$. Dann gilt auch $P(B) = P(B\|A)$.	☐	☐
7.	Es sei f_X eine Dichtefunktion zu einer stetigen Zufallsvariablen X. Dann gilt stets: $\int_{-\infty}^{\infty} f_X(x)dx = 1$	☐	☐
8.	Die Zufallsvariablen X, Y und Z seien stochastisch unabhängig, ferner gelte $X \sim B(10, 0.2)$, $Y \sim B(15, 0.2)$ und $Z \sim B(25, 0.2)$. Dann ist die Summe $S := X + Y + Z$ ebenfalls binomialverteilt, es gilt genauer $S \sim B(50, 0.2)$.	☐	☐

Aufgabe 2 (12 Punkte)

Markieren Sie jeweils die korrekte Antwort mit einem Kreuz im betreffenden Kästchen. Es ist jeweils genau ein Kreuz korrekt.
Richtige Antworten geben 3 Punkte, falsche Antworten und nicht bearbeitete Aufgabenteile 0 Punkte (Aufgabenteile mit mehr als einem Kreuz zählen als nicht bearbeitet!).

1. Kreuzen Sie an, auf welche Merkmalseigenschaften der folgende Box-Plot *am ehesten* hindeutet:



(a) leptokurtisch und rechts**steil** ☐
(b) leptokurtisch und links**steil** ☐
(c) platykurtisch und rechts**steil** ☐
(d) platykurtisch und links**steil** ☐

2. Die Wahrscheinlichkeit, durch eine rein zufällige Aneinanderreihung der Buchstaben E,K,K,O,R,R,R,T und U das Wort „KORREKTUR" zu erhalten, beträgt:

(a) $\frac{1!+2!+1!+3!+1!+1!}{9!}$ ☐

(b) $\frac{1}{1\cdot 2\cdot 1\cdot 3\cdot 1\cdot 1}$ ☐

(c) $\frac{1!\cdot 2!\cdot 1!\cdot 3!\cdot 1!\cdot 1!}{9!}$ ☐

(d) $\frac{1}{9!}$ ☐

3. Das Merkmal X des zweidimensionalen Merkmals (X,Y) sei ordinalskaliert, das Merkmal Y kardinalskaliert. Damit ist die Berechnung folgender Abhängigkeitsmaße zwischen X und Y immer möglich:

(a) Pearsonscher Korrelationskoeffizient und korrigierter Pearsonscher Kontingenzkoeffizient ☐
(b) Spearmanscher Rangkorrelationskoeffizient und korrigierter Pearsonscher Kontingenzkoeffizient ☐
(c) Nur Spearmanscher Rangkorrelationskoeffizient ☐
(d) Nur korrigierter Pearsonscher Kontingenzkoeffizient ☐

4. Der zweidimensionale diskrete Zufallsvektor (X,Y) besitze 8 Trägerpunkte, die alle auf einer Geraden mit Steigung 0.1 liegen. Dann gilt:

(a) $\mathrm{Korr}(X,Y) = -0.1$ ☐
(b) $\mathrm{Korr}(X,Y) = 0.1$ ☐
(c) $\mathrm{Korr}(X,Y) = 0.8$ ☐
(d) $\mathrm{Korr}(X,Y) = 1$ ☐

Aufgabe 3 (3 + 3 + 1 + 5 + 3 = 15 Punkte)

Bei einer Umfrage wurden 25 Personen befragt, wie viele Smartphones sie in den vergangenen fünf Jahren gekauft haben (Merkmal X). Das Ergebnis der Umfrage ist die folgende (bereits aufsteigend sortierte) Urliste zu X:

0, 0, 0, 1, 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 2, 2, 3, 3, 3, 3, 3, 3, 4, 4

(a) Erstellen Sie eine Tabelle mit den absoluten und relativen Häufigkeiten.

(b) Stellen Sie die zugehörige empirische Verteilungsfunktion auf.

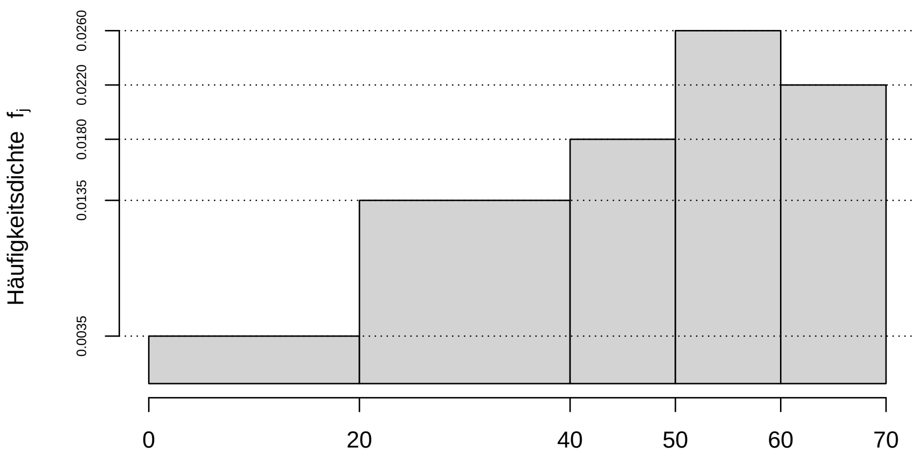
(c) Wie groß ist der Anteil der Personen in der Umfrage, die mindestens 3 Smartphones in den vergangenen fünf Jahren gekauft haben?

(d) Berechnen Sie den arithmetischen Mittelwert und die empirische Varianz des Merkmals X.

(e) Bestimmen Sie ein unteres Quartil, ein oberes Quartil und den zugehörigen Interquartilsabstand des Merkmals X.

Aufgabe 4 (7 + 4 + 3 + 2 + 2 = 18 Punkte)

Gegeben sei das folgende Histogramm zur Klassierung einer Urliste der Länge $n = 100$:



(a) Rekonstruieren Sie die Klassierung der Daten aus dem Histogramm. Geben Sie insbesondere die jeweiligen Klassenbreiten, Klassenmitten, absoluten und relativen Klassenhäufigkeiten, Häufigkeitsdichten sowie die Werte der empirischen Verteilungsfunktion an den Klassengrenzen an.

(b) Stellen Sie die (approximative) empirische Verteilungsfunktion des Merkmals aus der Klassierung der Daten auf.

(c) Berechnen Sie aus den klassierten Daten den (approximativen) arithmetischen Mittelwert der Daten. Wie groß ist die relative Abweichung vom tatsächlichen (aus der Urliste bestimmten) Mittelwert von 45.687?

(d) Welche Näherung für die Anzahl von Merkmalswerten zwischen 30 und 50 können Sie unter Verwendung der approximativen Verteilungsfunktion für die klassierten Daten berechnen?

(e) Bestimmen Sie näherungsweise unter Verwendung der approximativen Verteilungsfunktion für die klassierten Daten das untere Quartil.

Aufgabe 5 (1 + 2 + 2 = 5 Punkte)

Eine Urne enthält 30 gleichartige Kugeln, von denen 3 lila und gepunktet, 6 orange und gepunktet, 12 lila und ungemustert sowie 9 orange und ungemustert sind. Wie groß ist die Wahrscheinlichkeit dafür, dass

(a) eine rein zufällig aus der Urne entnommene Kugel orange und gepunktet ist?

(b) eine rein zufällig aus der Urne entnommene Kugel lila ist?

(c) eine rein zufällig aus der Urne entnommene Kugel ungemustert ist, wenn man weiß, dass sie lila ist?

Aufgabe 6 (5 + 2 + 1 = 8 Punkte)

Bei einem Chip-Hersteller verteilt sich die Produktion eines bestimmten Mikrocontrollers auf insgesamt drei verschiedene Produktionslinien A, B und C. Dabei werden im Mittel 15% der Chips auf Linie A, 40% der Chips auf Linie B und 45% der Chips auf Linie C hergestellt. Aus den Ergebnissen der Qualitätssicherung ist bekannt, dass 98% der Chips aus Linie A, 98.5% der Chips aus Linie B und 99% der Chips aus Linie C nicht fehlerhaft sind.

(a) Wie groß ist die Wahrscheinlichkeit, dass ein zufällig ausgewählter Chip nicht fehlerhaft ist?

(b) Wie groß ist die Wahrscheinlichkeit, dass ein fehlerhafter Chip auf der Linie C produziert wurde?

(c) Sind die Ereignisse „Chip ist fehlerhaft“ und „Chip wurde auf Linie C produziert“ stochastisch unabhängig?

Aufgabe 7 (5 + 2 + 6 + 4 = 17 Punkte)

Die Verteilung einer stetigen Zufallsvariablen X sei durch die folgende Dichtefunktion gegeben:

$$f_X(x) = \begin{cases} x + \frac{1}{2} & \text{für } -\frac{1}{2} \leq x < \frac{1}{2} \\ -4x + 4 & \text{für } \frac{1}{2} \leq x \leq 1 \\ 0 & \text{sonst} \end{cases}$$

(a) Berechnen Sie die Verteilungsfunktion F_X von X.

(b) Berechnen Sie $P(\{X < 0\})$ und $P\left(\left\{0 \leq X \leq \frac{3}{4}\right\}\right)$.

(c) Bestimmen Sie den Erwartungswert $\mathrm{E}(X)$.

(d) Bestimmen Sie das obere Quartil von X.

Aufgabe 8 (2 + 3 = 5 Punkte)

Als Hausaufgabe im Fach Chemie waren die Atommassen von 12 Elementen des Periodensystems auswendig zu lernen. Der Schüler Carl Clever hat 9 dieser Atommassen auswendig gelernt (die Chance, bei den anderen 3 Atommassen durch Raten eine richtige Antwort zu geben, sei gleich Null). Der Lehrer überprüft, ob Carl die Hausaufgabe ordentlich erledigt hat, indem er 3 Mal rein zufällig und unabhängig voneinander eines der Elemente auswählt und die zugehörigen Atommassen abfragt. Kann Carl mindestens zu 2 dieser 3 Elemente die Atommassen korrekt angeben, so ist die Überprüfung bestanden.

(a) Welche Verteilung besitzt die Anzahl der von Carl abgegebenen richtigen Antworten?

(b) Mit welcher Wahrscheinlichkeit besteht Carl die Überprüfung?

Aufgabe 9 (2 + 9 + 1 + 3 = 15 Punkte)

Gegeben sei die folgende Tabelle der gemeinsamen Wahrscheinlichkeitsverteilung zu einem zweidimensionalen diskreten Zufallsvektor (X, Y):

$X \setminus Y$	-1	0	2	$p_{i\cdot}$
2	0.07	0.17	0.06	
4	0.13	0.28	0.09	
6	0.1	0.05	0.05	
$p_{\cdot j}$				

(a) Ergänzen Sie die obige Tabelle (in den vorgesehenen Feldern) um ihre Randverteilungen.

(b) Berechnen Sie $\mathrm{E}(X)$, $\mathrm{E}(Y)$, $\mathrm{Var}(X)$, $\mathrm{Var}(Y)$, $\mathrm{Cov}(X, Y)$ sowie $\mathrm{Korr}(X, Y)$.

(c) Sind X und Y stochastisch unabhängig?

(d) Berechnen Sie $\mathrm{E}(4X - 5Y)$ sowie $\mathrm{Var}(4X - 5Y)$.

Aufgabe 10 (2 + 4 + 3 = 9 Punkte)

Die Zufallsvariablen $X_1, \ldots, X_{336}$ seien unabhängig identisch $B(1, 0.7)$-verteilt. Die Summe der Zufallsvariablen X_i sei mit

$$Y := \sum_{i=1}^{336} X_i = X_1 + \ldots + X_{336}$$

bezeichnet.

(a) Geben Sie die (exakte) Verteilung von Y sowie deren Erwartungswert $\mathrm{E}(Y)$ und Varianz $\mathrm{Var}(Y)$ an.

(b) Verwenden Sie den zentralen Grenzwertsatz, um näherungsweise zu bestimmen, mit welcher Wahrscheinlichkeit Y Werte zwischen 220 und 240 annimmt.

(c) Verwenden Sie den zentralen Grenzwertsatz, um näherungsweise ein 0.90-Quantil von Y zu bestimmen.

Hinweis: Verwenden Sie zur Bearbeitung von Aufgabenteil (b) und (c) die Tabelle zur Standardnormalverteilung auf Seite 120!

8 Deskr. Statistik und Wahrscheinlichkeitsrechnung SS 2018

Aufgabe 1 (16 Punkte)

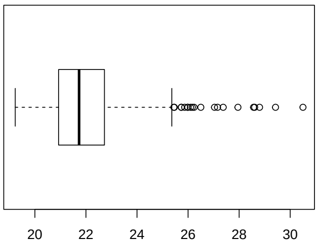
Markieren Sie jeweils mit einem Kreuz pro Aussage im betreffenden Kästchen, ob die unten stehenden Aussagen wahr oder falsch sind.
Richtige Antworten geben 2 Punkte, falsche Antworten und nicht bearbeitete Aussagen 0 Punkte (Aussagen mit zwei Kreuzen zählen als nicht bearbeitet!).

	wahr	falsch
1. Das arithmetische Mittel $\overline{\mathrm{rg}(X)}$ eines Rangmerkmals $\mathrm{rg}(X)$ hängt nur von der Länge der Urliste des Merkmals und nicht von den konkreten Merkmalsausprägungen ab.	☐	☐
2. Ist der korrigierte Pearsonsche Kontingenzkoeffizient von zwei Merkmalen X und Y positiv, dann gehen große Werte von X tendenziell auch mit großen Werten von Y einher.	☐	☐
3. Es seien A und B zwei Ereignisse in einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ mit $P(A) = 0.3$, $P(B) = 0.3$ und $P(A \cup B) = 0.4$. Damit gilt $P(A \cap B) = 0.1$.	☐	☐
4. Ist $(\Omega, \mathcal{F}, P)$ ein beliebiger Wahrscheinlichkeitsraum, so sind zwei Ereignisse $A, B \in \mathcal{F}$ genau dann stochastisch unabhängig, wenn $P(A \cap B) = P(A) + P(B)$ gilt.	☐	☐
5. In einem Laplaceschen Wahrscheinlichkeitsraum stimmen die Wahrscheinlichkeiten aller Elementarereignisse überein.	☐	☐
6. Für eine Zufallsvariable X über einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ gelte $P(X \in \{0, 8, 15\}) = 1$. Damit ist X eine diskrete Zufallsvariable.	☐	☐
7. Die Differenz einer $N(100, 6^2)$-verteilten Zufallsvariablen und einer hiervon stochastisch unabhängigen $N(50, 3^2)$-verteilten Zufallsvariablen ist $N(50, 5^2)$-verteilt.	☐	☐
8. Die Kovarianz zweier Zufallsvariablen ist betragsmäßig niemals größer als das Produkt ihrer beiden Standardabweichungen.	☐	☐

<u>**Aufgabe 2**</u> (12 Punkte)

Markieren Sie jeweils die korrekte Antwort mit einem Kreuz im betreffenden Kästchen. Es ist jeweils genau ein Kreuz korrekt.
Richtige Antworten geben 3 Punkte, falsche Antworten und nicht bearbeitete Aufgabenteile 0 Punkte (Aufgabenteile mit mehr als einem Kreuz zählen als nicht bearbeitet!).

1. Kreuzen Sie an, auf welche Merkmalseigenschaften der folgende Box-Plot *am ehesten* hindeutet:



(a) leptokurtisch und rechts**steil** ☐
(b) leptokurtisch und links**steil** ☐
(c) platykurtisch und rechts**steil** ☐
(d) platykurtisch und links**steil** ☐

2. Die Selbstkosten eines Produkts setzen sich (aktuell) aus 75% Herstellkosten, 15% Vertriebskosten sowie 10% Verwaltungskosten zusammen. Welche prozentuale Reduzierung der Selbstkosten ergibt sich (gerundet), wenn 5% der Herstellkosten, 20% der Vertriebskosten sowie 3% der Verwaltungskosten eingespart werden?

(a) 9.33% ☐
(b) 7.05% ☐
(c) 9.66% ☐
(d) 7.22% ☐

3. Es sei x_p ein p-Quantil des Merkmals X. Dann ist
(a) höchstens ein Anteil von p der Merkmalswerte von X höchstens so groß wie x_p und höchstens ein Anteil von $1-p$ mindestens so groß wie x_p. ☐
(b) mindestens ein Anteil von p der Merkmalswerte von X mindestens so groß wie x_p und mindestens ein Anteil von $1-p$ höchstens so groß wie x_p. ☐
(c) mindestens ein Anteil von p der Merkmalswerte von X höchstens so groß wie x_p und mindestens ein Anteil von $1-p$ mindestens so groß wie x_p. ☐
(d) höchstens ein Anteil von p der Merkmalswerte von X mindestens so groß wie x_p und höchstens ein Anteil von $1-p$ höchstens so groß wie x_p. ☐

4. Verteilungsfunktionen eindimensionaler Zufallsvariablen sind stets

(a) streng monoton wachsend und rechtsseitig stetig. ☐
(b) streng monoton wachsend und stetig. ☐
(c) monoton wachsend und rechtsseitig stetig. ☐
(d) monoton wachsend und stetig. ☐

Aufgabe 3 (4 + 1 + 5 + 1 = 11 Punkte)

Zu einem erhobenen Merkmal X sei die folgende empirische Verteilungsfunktion gegeben:

$$F(x) = \begin{cases} 0.00 & \text{für} \quad x < -3 \\ 0.06 & \text{für} \quad -3 \leq x < 0 \\ 0.28 & \text{für} \quad 0 \leq x < 3 \\ 0.50 & \text{für} \quad 3 \leq x < 6 \\ 0.86 & \text{für} \quad 6 \leq x < 9 \\ 0.96 & \text{für} \quad 9 \leq x < 12 \\ 1.00 & \text{für} \quad x \geq 12 \end{cases}$$

Außerdem sei die Länge der Urliste $n = 50$ bekannt.

(a) Erstellen Sie eine Tabelle der absoluten und relativen Häufigkeiten.

(b) Wie groß ist der Anteil der Urlisteneinträge, die Werte von weniger als 5 annehmen?

(c) Berechnen Sie den arithmetischen Mittelwert und die empirische Varianz des Merkmals X.

(d) Bestimmen Sie ein oberes Quartil des Merkmals X.

Aufgabe 4 (6 + 4 + 3 + 3 + 3 = 19 Punkte)

Zu einem kardinalskalierten Merkmal sei die folgende (zur einfacheren Bearbeitung der Aufgabe bereits sortierte) Urliste der Länge $n = 30$ gegeben:

13.11, 13.35, 14.90, 16.42, 17.25, 18.84, 20.74, 21.04, 21.86, 22.55, 22.76, 24.18, 25.68, 25.90, 27.16, 27.37, 27.49, 27.79, 28.35, 28.51, 29.01, 29.56, 29.80, 30.83, 31.23, 31.95, 32.30, 33.79, 34.30, 34.64

(a) Führen Sie eine Klassierung der erhobenen Daten auf Grundlage der Klassen

$$K_1 = (5, 15],\ K_2 = (15, 25],\ K_3 = (25, 31],\ K_4 = (31, 35]$$

durch. Geben Sie insbesondere die jeweiligen Klassenbreiten, Klassenmitten, absoluten und relativen Klassenhäufigkeiten, Häufigkeitsdichten sowie die Werte der empirischen Verteilungsfunktion an den Klassengrenzen an.

(b) Stellen Sie die (approximative) empirische Verteilungsfunktion des Merkmals aus der Klassierung der Daten auf.

(c) Berechnen Sie aus den klassierten Daten den (approximativen) arithmetischen Mittelwert der Daten. Wie groß ist die relative Abweichung vom tatsächlichen (aus der Urliste bestimmten) Mittelwert von 25.422?

(d) Bestimmen Sie (aus der Urliste) die Anzahl von Merkmalswerten zwischen 20 und 33. Welche Näherung für diese Anzahl können Sie aus der in Teil (b) aufgestellten empirischen Verteilungsfunktion berechnen?

(e) Bestimmen Sie (unter Verwendung der bekannten Konvention zur eindeutigen Festlegung von Quantilen bei kardinalskalierten Merkmalen) *den* Median sowohl exakt aus der Urliste als auch approximativ mit Hilfe der Verteilungsfunktion für die klassierten Daten.

Aufgabe 5 (1.5 + 1.5 + 1.5 + 1.5 = 6 Punkte)

Beim Einlaufen einer 11-köpfigen Fußballmannschaft auf den Platz laufen nacheinander in einer Reihe zuerst der Kapitän, dann der Torwart (falls dieser nicht auch der Kapitän ist) und anschließend die restlichen 9 (bzw. 10) Spieler ein.

(a) Wie viele verschiedene Reihenfolgen der Spieler gibt es beim Einlaufen, wenn der Torwart nicht gleichzeitig Kapitän der Mannschaft und die Reihenfolge der Spieler nach dem Torwart beliebig ist?

(b) Wie viele verschiedene Reihenfolgen gibt es, wenn der Torwart gleichzeitig Kapitän der Mannschaft und die Reihenfolge der anderen Spieler beliebig ist?

(c) Wie viele verschiedene Reihenfolgen gibt es, wenn der Torwart gleichzeitig Kapitän der Mannschaft ist und anschließend zunächst die 4 Abwehrspieler, dann die 4 Mittelfeldspieler und zuletzt die beiden Stürmer einlaufen sollen?

(d) Wie viele verschiedene Reihenfolgen gibt es, wenn der Torwart gleichzeitig Kapitän der Mannschaft ist und anschließend die Gruppen der 4 Abwehrspieler, der 4 Mittelfeldspieler und der beiden Stürmer zwar zusammenhängend einlaufen sollen, die Reihenfolge dieser Gruppen (Abwehr, Mittelfeld, Sturm) jedoch beliebig ist?

Aufgabe 6 (4 + 2 + 2 = 8 Punkte)

An einer seltenen Krankheit seien 2% der Bevölkerung einer bestimmten Altersgruppe erkrankt. Zum Einsatz in flächendeckenden Früherkennungsuntersuchungen existiere ein medizinisches Diagnoseverfahren, welches erkrankte Personen mit einer Wahrscheinlichkeit von 99% (korrekterweise) auch als krank einstuft, bei gesunden (bzw. nicht an dieser Krankheit erkrankten) Personen allerdings mit einer Wahrscheinlichkeit von 3% (fälschlicherweise) ebenfalls eine entsprechende Erkrankung diagnostiziert.

(a) Mit welcher Wahrscheinlichkeit wird ein Patient in der betrachteten Altersgruppe im Rahmen einer Früherkennungsuntersuchung als krank eingestuft?

(b) Mit welcher Wahrscheinlichkeit wird sich eine positive Diagnose bei einer Früherkennungsuntersuchung in der betrachteten Altersgruppe als falsch herausstellen?

(c) Wie groß ist die Wahrscheinlichkeit für eine richtige Diagnose bei einer Früherkennungsuntersuchung in der betrachteten Altersgruppe?

Aufgabe 7 (1 + 2 + 2 + 3 = 8 Punkte)

Die Anzahl der Flugzeugunfälle mit Todesfolge pro Jahr auf von U. S. Airlines betriebenen Flügen lasse sich als eine Pois(0.4)-verteilte Zufallsvariable auffassen. Außerdem soll angenommen werden, dass die Anzahl der Flugzeugunfälle mit Todesfolge pro Jahr auf von U. S. Airlines betriebenen Flügen für unterschiedliche Jahre stochastisch unabhängig ist.

(a) Welchen Erwartungswert hat die Anzahl der Flugzeugunfälle mit Todesfolge pro Jahr auf von U. S. Airlines betriebenen Flügen?

(b) Mit welcher Wahrscheinlichkeit ereignen sich in einem Jahr auf von U. S. Airlines betriebenen Flügen 0 Unfälle mit Todesfolge?

(c) Mit welcher Wahrscheinlichkeit ereignet sich in einem Jahr auf von U. S. Airlines betriebenen Flügen mehr als 1 Unfall mit Todesfolge?

(d) Welche Verteilung hat die Anzahl der Flugzeugunfälle mit Todesfolge pro Jahrzehnt auf von U. S. Airlines betriebenen Flügen? Mit welcher Wahrscheinlichkeit ereignet sich in einem Zehnjahreszeitraum auf von U. S. Airlines betriebenen Flügen mindestens 1 Unfall mit Todesfolge?

Aufgabe 8 (3 + 2 + 6 + 1 + 4 = 16 Punkte)

Die Verteilung einer stetigen Zufallsvariablen X sei durch die folgende Verteilungsfunktion gegeben:

$$F_X(x) = \begin{cases} 0 & \text{für } x \leq -2 \\ \frac{1}{12}x^2 + \frac{1}{3}x + \frac{1}{3} & \text{für } -2 < x \leq 0 \\ -\frac{1}{24}x^2 + \frac{1}{3}x + \frac{1}{3} & \text{für } 0 < x \leq 4 \\ 1 & \text{für } x > 4 \end{cases}$$

(a) Berechnen Sie eine Dichtefunktion f_X von X.

(b) Berechnen Sie $P(\{X > 0\})$ und $P(\{1 \leq X \leq 3\})$.

(c) Bestimmen Sie den Erwartungswert $\mathrm{E}(X)$.

(d) Ist X symmetrisch um ihren Erwartungswert verteilt (ohne Begründung)?

(e) Bestimmen Sie das obere Quartil von X.

Aufgabe 9 (2 + 8 + 1 + 2 + 3 = 16 Punkte)

Ein fairer Würfel wird zweimal geworfen. Es seien X die Anzahl der Würfe mit einer Augenzahl ≤ 4 sowie Y die Anzahl der Würfe mit einer Augenzahl ≥ 4.

(a) Welcher Verteilung genügen X und Y (jeweils)?

(b) Die gemeinsame Verteilung von (X, Y) ist gegeben durch:

$X \setminus Y$	0	1	2
0	0	0	$\frac{1}{9}$
1	0	$\frac{1}{3}$	$\frac{1}{9}$
2	$\frac{1}{4}$	$\frac{1}{6}$	$\frac{1}{36}$

Bestimmen Sie $\mathrm{E}(X)$, $\mathrm{E}(Y)$, $\mathrm{Var}(X)$, $\mathrm{Var}(Y)$, $\mathrm{Cov}(X, Y)$ sowie $\mathrm{Korr}(X, Y)$.

Hinweis: Beachten Sie, dass Sie die diesen Aufgabenteil unter Verwendung der Ergebnisse aus Teil (a) zum Teil recht schnell und insbesondere vollständig ohne die Bestimmung der Randwahrscheinlichkeiten von X und Y lösen können!

(c) Mit welcher Wahrscheinlichkeit nehmen sowohl X als auch Y Werte von mindestens 1 an?

(d) Sind X und Y stochastisch unabhängig? Begründen Sie Ihre Antwort.

(e) Berechnen Sie $\mathrm{E}(6X - 4Y)$ sowie $\mathrm{Var}(6X - 4Y)$.

Aufgabe 10 (3 + 2 + 3 = 8 Punkte)

Aufgrund langjähriger Aufzeichnungen über entsprechende Wahlbeteiligungen gehe man davon aus, dass sich die 14800 wahlberechtigten Studierenden bei einer anstehenden Senatswahl unabhängig voneinander jeweils mit einer Wahrscheinlichkeit von 7.5% dazu entschließen, ihr Wahlrecht auch auszuüben und einen Stimmzettel auszufüllen.

(a) Wie ist die Anzahl der ausgefüllten Stimmzettel Y exakt verteilt? Geben Sie auch den Erwartungswert $\mathrm{E}(Y)$ sowie die Varianz $\mathrm{Var}(Y)$ der Anzahl der ausgefüllten Stimmzettel an.

(b) Die Wahlleitung entschließt sich aus ökologischen und ökonomischen Gründen, zunächst nur 1175 Stimmzettel für die Wahl auszudrucken und weitere Stimmzettel erst bei Bedarf nachzudrucken, falls mehr als die zunächst gedruckten Stimmzettel benötigt werden sollten. Mit welcher (mit dem zentralen Grenzwertsatz näherungsweise zu bestimmenden) Wahrscheinlichkeit müssen keine Stimmzettel nachgedruckt werden?

(c) Verwenden Sie den zentralen Grenzwertsatz, um näherungsweise die (kleinste) Anzahl von Stimmzetteln zu bestimmen, die mit einer Wahrscheinlichkeit von (mindestens) 99.5% ausreichend ist.

Hinweis: Verwenden Sie zur Bearbeitung von Aufgabenteil (b) und (c) die Tabelle zur Standardnormalverteilung auf Seite 120!

9 Deskr. Statistik und Wahrscheinlichkeitsrechnung WS 2018/19

Aufgabe 1 (16 Punkte)

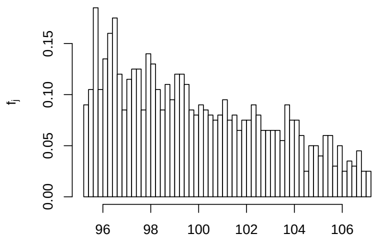
Markieren Sie jeweils mit einem Kreuz pro Aussage im betreffenden Kästchen, ob die unten stehenden Aussagen wahr oder falsch sind.
Richtige Antworten geben 2 Punkte, falsche Antworten und nicht bearbeitete Aussagen 0 Punkte (Aussagen mit zwei Kreuzen zählen als nicht bearbeitet!).

	wahr	falsch
1. Alle Lagemaße, die für ordinalskalierte Merkmale berechnet werden können, können auch für nominalskalierte Merkmale berechnet werden.	☐	☐
2. Betrachtet man die Summe der quadrierten Abweichungen aller n Urlisteneinträge x_i eines kardinalskalierten Merkmals X von einer beliebigen reellen Zahl a, so wird diese Summe am kleinsten für $a = \overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$.	☐	☐
3. Es gibt Laplacesche Wahrscheinlichkeitsräume mit Ergebnismenge $\Omega = \mathbb{N}$.	☐	☐
4. Die Wahrscheinlichkeit, mit drei fairen (sechsseitigen) Würfeln bei gleichzeitigem Würfeln jeweils eine 2, eine 4 sowie eine 6 zu würfeln, beträgt (gerundet) 2.778%.	☐	☐
5. Wenn Sie alle 8 Aufgabenteile dieser Aufgabe entweder mit *wahr* oder mit *falsch* beantworten, dann haben Sie insgesamt $\binom{8}{2}$ Möglichkeiten zur Bearbeitung der Aufgabe.	☐	☐
6. Es seien $(\Omega, \mathcal{F}, P)$ ein Wahrscheinlichkeitsraum, $A, B \in \mathcal{F}$ zwei Ereignisse mit $P(A) > 0$, $P(B) > 0$ sowie $P(A\|B) = P(B\|A)$. Dann gilt: $$P(A) \neq P(B) \quad \Longrightarrow \quad P(A\|B) = P(B\|A) = 0$$	☐	☐
7. Ist F_X die Verteilungsfunktion einer stetigen Zufallsvariablen X, so gilt für $a, b \in \mathbb{R}$ mit $a < b$ stets $$P(\{a < X < b\}) = \int_a^b F_X(x)dx\ .$$	☐	☐
8. Ist die Kovarianz zweier Zufallsvariablen X und Y positiv, so ist die Varianz der Differenz von X und Y kleiner als die Summe der einzelnen Varianzen (von X und Y).	☐	☐

Aufgabe 2 (12 Punkte)

Markieren Sie jeweils die korrekte Antwort mit einem Kreuz im betreffenden Kästchen. Es ist jeweils genau ein Kreuz korrekt.
Richtige Antworten geben 3 Punkte, falsche Antworten und nicht bearbeitete Aufgabenteile 0 Punkte (Aufgabenteile mit mehr als einem Kreuz zählen als nicht bearbeitet!).

1. Kreuzen Sie an, auf welche Merkmalseigenschaften das folgende Histogramm *am ehesten* hindeutet:



(a) leptokurtisch und rechts**steil** ☐

(b) leptokurtisch und links**steil** ☐

(c) platykurtisch und rechts**steil** ☐

(d) platykurtisch und links**steil** ☐

2. Die Wahrscheinlichkeit, durch eine rein zufällige Anordnung der Ziffern 3,3,3,6,6,6 die Zahl 636363 zu erhalten, beträgt:

(a) $\frac{3! \cdot 3!}{6!}$ ☐

(b) $\frac{1}{3! \cdot 3!}$ ☐

(c) $\frac{3!}{6! \cdot 6!}$ ☐

(d) $\frac{3!}{3^6}$ ☐

3. Die Ränge $\text{rg}(X)_1, \dots, \text{rg}(X)_8$ zur (der Einfachheit halber sortierten) Urliste

5, 7, 10, 10, 10, 12, 12, 15

des ordinalskalierten Merkmals X lauten:

(a) 1.5, 1.5, 3.5, 3.5, 3.5, 6.5, 6.5, 8 ☐

(b) 1, 2, 3.5, 3.5, 3.5, 6.5, 6.5, 8 ☐

(c) 1, 2, 3, 3, 3, 4, 4, 5 ☐

(d) 1, 2, 4, 4, 4, 6.5, 6.5, 8 ☐

4. Sind X und Y zwei stochastisch unabhängige Zufallsvariablen mit $X \sim B(100, 0.2)$ und $Y \sim B(100, 0.2)$, dann ist die Verteilung von $X + Y$ eine

(a) $B(100, 0.2)$-Verteilung. ☐

(b) $B(100, 0.4)$-Verteilung. ☐

(c) $B(200, 0.2)$-Verteilung. ☐

(d) $B(200, 0.4)$-Verteilung. ☐

Aufgabe 3 (4 + 1 + 5 + 1 = 11 Punkte)

Zu einem erhobenen Merkmal X sei die folgende empirische Verteilungsfunktion gegeben:

$$F(x) = \begin{cases} 0.00 & \text{für} \quad x < 2 \\ 0.12 & \text{für} \quad 2 \leq x < 3 \\ 0.38 & \text{für} \quad 3 \leq x < 4 \\ 0.68 & \text{für} \quad 4 \leq x < 5 \\ 0.88 & \text{für} \quad 5 \leq x < 6 \\ 0.97 & \text{für} \quad 6 \leq x < 7 \\ 1.00 & \text{für} \quad x \geq 7 \end{cases}$$

Außerdem sei die Länge der Urliste $n = 100$ bekannt.

(a) Erstellen Sie eine Tabelle der absoluten und relativen Häufigkeiten.

(b) Wie groß ist der Anteil der Urlisteneinträge, die Werte von mehr als 3 annehmen?

(c) Berechnen Sie den arithmetischen Mittelwert und die empirische Varianz des Merkmals X.

(d) Bestimmen Sie einen Median des Merkmals X.

Aufgabe 4 (6 + 4 + 3 + 3 + 3 = 19 Punkte)

Zu einem kardinalskalierten Merkmal sei die folgende (zur einfacheren Bearbeitung der Aufgabe bereits sortierte) Urliste der Länge $n = 40$ gegeben:

19.47, 19.63, 20.24, 20.56, 28.27, 29.29, 30.77, 33.97, 36.80, 39.17, 42.96, 47.32, 47.39, 49.86, 51.99, 59.47, 64.66, 64.87, 67.07, 69.79, 70.40, 71.27, 78.11, 78.74, 79.65, 80.19, 80.52, 81.15, 81.65, 82.52, 84.35, 84.64, 85.30, 88.21, 88.21, 88.85, 90.20, 90.70, 96.65, 96.70

(a) Führen Sie eine Klassierung der erhobenen Daten auf Grundlage der Klassen

$$K_1 = (10, 40],\ K_2 = (40, 60],\ K_3 = (60, 80],\ K_4 = (80, 100]$$

durch. Geben Sie insbesondere die jeweiligen Klassenbreiten, Klassenmitten, absoluten und relativen Klassenhäufigkeiten, Häufigkeitsdichten sowie die Werte der empirischen Verteilungsfunktion an den Klassengrenzen an.

(b) Stellen Sie die (approximative) empirische Verteilungsfunktion des Merkmals aus der Klassierung der Daten auf.

(c) Berechnen Sie aus den klassierten Daten den (approximativen) arithmetischen Mittelwert der Daten. Wie groß ist die relative Abweichung vom tatsächlichen (aus der Urliste bestimmten) Mittelwert von 63.039?

(d) Bestimmen Sie (aus der Urliste) die Anzahl von Merkmalswerten zwischen 25 und 60. Welche Näherung für diese Anzahl können Sie aus der in Teil (b) aufgestellten empirischen Verteilungsfunktion berechnen?

(e) Bestimmen Sie (unter Verwendung der bekannten Konvention zur eindeutigen Festlegung von Quantilen bei kardinalskalierten Merkmalen) *den* Median sowohl exakt aus der Urliste als auch approximativ mit Hilfe der Verteilungsfunktion für die klassierten Daten.

Aufgabe 5 $(2 + 2 + 2 = 6$ Punkte)

Eine Lostrommel enthält 20 (gleichartige) Kugeln, die von 1 bis 20 durchnummeriert sind. Es wird einmalig rein zufällig eine der 20 Kugeln gezogen.

(a) Geben Sie einen geeigneten Wahrscheinlichkeitsraum zur Beschreibung dieses Zufallsexperiments an.

(b) Wie groß ist die Wahrscheinlichkeit, eine Kugel mit einer geraden Zahl zu ziehen? Wie groß ist die Wahrscheinlichkeit, eine Kugel mit einer Zahl kleiner oder gleich 10 zu ziehen?

(c) Sind die Ereignisse „Kugel mit einer geraden Zahl“ und „Kugel mit einer Zahl kleiner oder gleich 10“ aus Teil (b) stochastisch unabhängig? Begründen Sie Ihre Antwort!

Aufgabe 6 $(4 + 2 + 2 = 8$ Punkte)

An einer seltenen Krankheit seien 1% der Bevölkerung einer bestimmten Altersgruppe erkrankt. Zum Einsatz in flächendeckenden Früherkennungsuntersuchungen existiere ein medizinisches Diagnoseverfahren, welches erkrankte Personen mit einer Wahrscheinlichkeit von 98% (korrekterweise) auch als krank einstuft, bei gesunden (bzw. nicht an dieser Krankheit erkrankten) Personen allerdings mit einer Wahrscheinlichkeit von 4% (fälschlicherweise) ebenfalls eine entsprechende Erkrankung diagnostiziert.

(a) Mit welcher Wahrscheinlichkeit wird ein Patient in der betrachteten Altersgruppe im Rahmen einer Früherkennungsuntersuchung als krank eingestuft?

(b) Mit welcher Wahrscheinlichkeit wird sich eine positive Diagnose bei einer Früherkennungsuntersuchung in der betrachteten Altersgruppe als falsch herausstellen?

(c) Wie groß ist die Wahrscheinlichkeit für eine richtige Diagnose bei einer Früherkennungsuntersuchung in der betrachteten Altersgruppe?

Aufgabe 7 $(3 + 6 + 4 = 13$ Punkte)

Die Verteilung einer stetigen Zufallsvariablen X sei durch die folgende Verteilungsfunktion gegeben:

$$F_X(x) = \begin{cases} 0 & \text{für } x \leq 0 \\ -\frac{1}{2}x^2 + x & \text{für } 0 < x \leq 1 \\ \frac{1}{8}x^2 - \frac{1}{4}x + \frac{5}{8} & \text{für } 1 < x \leq 3 \\ 1 & \text{für } x > 3 \end{cases}$$

(a) Berechnen Sie eine Dichtefunktion f_X von X.

(b) Bestimmen Sie den Erwartungswert $\mathrm{E}(X)$.

(c) Bestimmen Sie das obere Quartil von X.

Aufgabe 8 $(3 + 2 + 2 = 7$ Punkte)

Bei einem bekannten Brettspiel muss (mit einem fairen sechsseitigen Würfel) zunächst eine Sechs gewürfelt werden, um eine Spielfigur auf das Startfeld des (eigentlichen) Spielfeldes zu stellen. Dazu hat man zu Beginn des Spiels (bevor man die erste Sechs gewürfelt hat) jeweils bis zu 3 Versuche pro Runde.

(a) Wie oft muss man im Mittel würfeln (einschließlich des erfolgreichen Wurfs!), bis man zum ersten Mal eine Sechs gewürfelt hat?

(b) Wie groß ist die Wahrscheinlichkeit, dass man innerhalb der ersten beiden Runden (also nach spätestens 6 Würfen) die erste Sechs gewürfelt hat?

(c) Wie groß ist die Wahrscheinlichkeit, dass man auch nach 3 Runden (also insgesamt 9 Würfen) noch keine Sechs gewürfelt hat?

Aufgabe 9 $(2 + 3 + 9 + 1 + 3 = 18$ Punkte)

Gegeben sei die folgende Tabelle der gemeinsamen Wahrscheinlichkeitsverteilung zu einem zweidimensionalen diskreten Zufallsvektor (X, Y):

$X \setminus Y$	-2	0	2	$p_{i\cdot}$
1	$\frac{2}{25}$	$\frac{4}{25}$	$\frac{4}{25}$	
3	$\frac{2}{25}$	$\frac{1}{25}$	$\frac{2}{25}$	
5	$\frac{6}{25}$	0	$\frac{4}{25}$	
$p_{\cdot j}$				

(a) Ergänzen Sie die obige Tabelle (in den vorgesehenen Feldern) um ihre Randverteilungen.

(b) Geben Sie die bedingten Wahrscheinlichkeitsverteilungen von Y unter der Bedingung $X = x_i$ für alle $x_i \in T(X)$ über die zugehörigen (bedingten) Wahrscheinlichkeitsfunktionen an.

(c) Berechnen Sie $\mathrm{E}(X)$, $\mathrm{E}(Y)$, $\mathrm{Var}(X)$, $\mathrm{Var}(Y)$, $\mathrm{Cov}(X, Y)$ sowie $\mathrm{Korr}(X, Y)$.

(d) Sind X und Y stochastisch unabhängig?

(e) Berechnen Sie $\mathrm{E}(5X - 3Y)$ sowie $\mathrm{Var}(5X - 3Y)$.

Aufgabe 10 (2 + 4 + 4 = 10 Punkte)

Die Zufallsvariablen $X_1, \ldots, X_{49}$ seien unabhängig identisch Pois(4)-verteilt. Die Summe der Zufallsvariablen X_i sei mit

$$Y := \sum_{i=1}^{49} X_i = X_1 + \ldots + X_{49}$$

bezeichnet.

(a) Geben Sie die (exakte) Verteilung von Y sowie deren Erwartungswert $\mathrm{E}(Y)$ und Varianz $\mathrm{Var}(Y)$ an.

(b) Verwenden Sie den zentralen Grenzwertsatz, um näherungsweise zu bestimmen, mit welcher Wahrscheinlichkeit Y Werte zwischen 175 und 210 annimmt.

(c) Verwenden Sie den zentralen Grenzwertsatz, um näherungsweise einen um den Erwartungswert von Y symmetrischen Bereich zu bestimmen, in dem sich Y mit einer Wahrscheinlichkeit von 0.90 realisiert.

Hinweis: Verwenden Sie zur Bearbeitung von Aufgabenteil (b) und (c) die Tabelle zur Standardnormalverteilung auf Seite 120!

10 Deskr. Statistik und Wahrscheinlichkeitsrechnung SS 2019

Aufgabe 1 (16 Punkte)

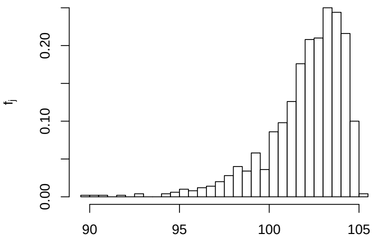
Markieren Sie jeweils mit einem Kreuz pro Aussage im betreffenden Kästchen, ob die unten stehenden Aussagen wahr oder falsch sind.
Richtige Antworten geben 2 Punkte, falsche Antworten und nicht bearbeitete Aussagen 0 Punkte (Aussagen mit zwei Kreuzen zählen als nicht bearbeitet!).

	wahr	falsch
1. Numerische Merkmale sind stets kardinalskaliert.	☐	☐
2. Bei empirischen Verteilungsfunktionen zu nicht klassierten Merkmalen X gibt es zu jeder Sprungstelle x (mindestens) einen passenden Eintrag $x_i = x$ in der zugehörigen Urliste $x_1, \ldots, x_n$ des Merkmals X.	☐	☐
3. Wenn Sie eine Teilstrecke von 400 [km] mit einem Durchschnittsverbrauch von 7 [l/100 km] zurücklegen und eine weitere Teilstrecke von 600 [km] mit einem Durchschnittsverbrauch von 6 [l/100 km], dann haben Sie auf der Gesamtstrecke von 1000 [km] einen (Gesamt-)Durchschnittsverbrauch von 6.4 [l/100 km].	☐	☐
4. Es seien $(\Omega, \mathcal{F}, P)$ ein Wahrscheinlichkeitsraum, $A, B, C \in \mathcal{F}$ drei Ereignisse mit $P(C) > 0$. Dann gilt stets: $$P(A\|C) + P(B\|C) \leq 2 \cdot P(A \cap B\|C)$$	☐	☐
5. Ein medizinisches Diagnoseverfahren zur Früherkennung einer seltenen Erkrankung führt mit einer Wahrscheinlichkeit von 3% (fälschlicherweise) zu einer positiven Diagnose, obwohl die Erkrankung tatsächlich nicht vorliegt. Das bedeutet, dass man als Untersuchter bei positiver Diagnose dieses Tests mit einer Wahrscheinlichkeit von 97% tatsächlich erkrankt ist.	☐	☐
6. Exponentialverteilte Zufallsvariablen sind stets leptokurtisch verteilt.	☐	☐
7. Die Wahrscheinlichkeit, bei einem zweimaligen Würfelwurf mit einem fairen (6-seitigen) Würfel im zweiten Wurf eine höhere Punktzahl zu würfeln als im ersten Wurf, beträgt (gerundet) 33.3%.	☐	☐
8. Seien X_1, X_2 und X_3 stochastisch unabhängige Zufallsvariablen mit $\text{Var}(X_1) = 30$, $\text{Var}(X_2) = 20$, sowie $\text{Var}(X_3) = 10$. Dann gilt $\text{Var}(X_1 + X_2 - X_3) = 40$.	☐	☐

Aufgabe 2 (12 Punkte)

Markieren Sie jeweils die korrekte Antwort mit einem Kreuz im betreffenden Kästchen. Es ist jeweils genau ein Kreuz korrekt.
Richtige Antworten geben 3 Punkte, falsche Antworten und nicht bearbeitete Aufgabenteile 0 Punkte (Aufgabenteile mit mehr als einem Kreuz zählen als nicht bearbeitet!).

1. Kreuzen Sie an, auf welche Merkmalseigenschaften das folgende Histogramm *am ehesten* hindeutet:



(a) leptokurtisch und rechts**steil** ☐
(b) leptokurtisch und links**steil** ☐
(c) platykurtisch und rechts**steil** ☐
(d) platykurtisch und links**steil** ☐

2. Die Ränge $\mathrm{rg}(X)_1, \ldots, \mathrm{rg}(X)_8$ zur (der Einfachheit halber sortierten) Urliste

$$2, \quad 3, \quad 5, \quad 5, \quad 9, \quad 9, \quad 11, \quad 14$$

des ordinalskalierten Merkmals X lauten:

(a) 1, 2, 3, 3, 4, 4, 5, 6 ☐
(b) 1, 2, 3.5, 3.5, 4.5, 4.5, 7, 8 ☐
(c) 1, 2, 3, 3, 5, 5, 7, 8 ☐
(d) 1, 2, 3.5, 3.5, 5.5, 5.5, 7, 8 ☐

3. In Deutschland gilt man als armutsgefährdet, wenn das eigene Nettoäquivalenzeinkommen geringer ist als 60% des Medians aller Nettoäquivalenzeinkommen. Wie verändert sich der Anteil der Armutsgefährdeten in Deutschland, wenn sich das Nettoäquivalenzeinkommen der einkommensstärksten 10% der Bevölkerung um 5% erhöht?

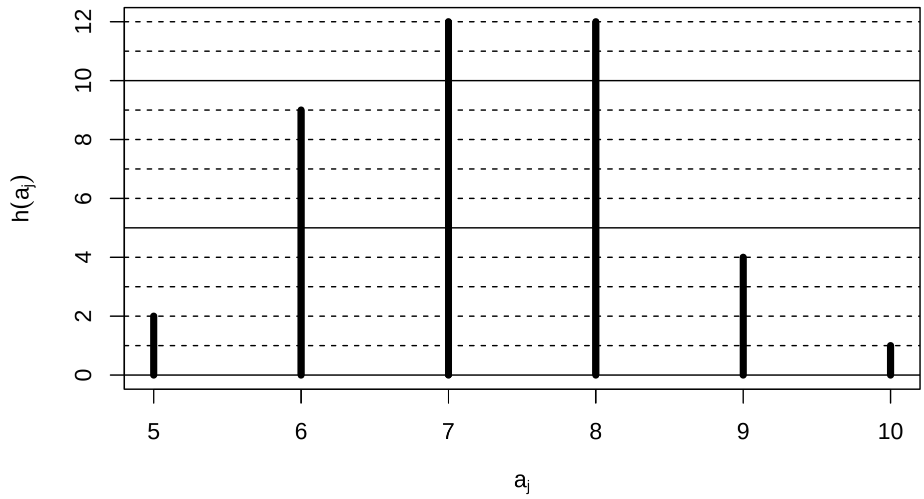
(a) Zunahme um 10% ☐
(b) Zunahme um 5% ☐
(c) Zunahme um 0.5% ☐
(d) Keine Veränderung ☐

4. Bei der Aufstellung eines Listenvorschlags zu einer Kommunalwahl sollen die 3 weiblichen und 2 männlichen Kanditat(inn)en so auf die 5 Listenplätze verteilt werden, dass sich die Geschlechter jeweils abwechseln. Damit beträgt die Anzahl der möglichen Listenanordnungen

(a) 24 ☐
(b) 12. ☐
(c) 10. ☐
(d) 8. ☐

Aufgabe 3 (4 + 5 + 3 + 1 + 1 = 14 Punkte)

Zu einem erhobenen Merkmal X sei das folgende Stabdiagramm gegeben:



(a) Erstellen Sie eine Tabelle der absoluten und relativen Häufigkeiten.

(b) Berechnen Sie den arithmetischen Mittelwert und die empirische Standardabweichung des Merkmals X.

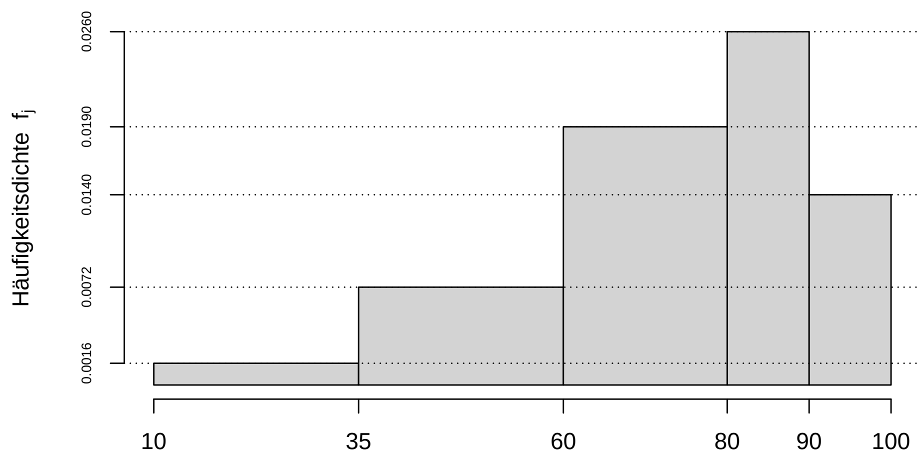
(c) Geben Sie die empirische Verteilungsfunktion des Merkmals X an.

(d) Wie groß ist der Anteil der Urlisteneinträge, die Werte von weniger als 9 annehmen?

(e) Berechnen Sie einen Median des Merkmals X.

Aufgabe 4 (7 + 4 + 3 + 2 = 16 Punkte)

Gegeben sei das folgende Histogramm zur Klassierung einer Urliste der Länge $n = 50$:



(a) Rekonstruieren Sie die Klassierung der Daten aus dem Histogramm. Geben Sie insbesondere die jeweiligen Klassenbreiten, Klassenmitten, absoluten und relativen Klassenhäufigkeiten, Häufigkeitsdichten sowie die Werte der empirischen Verteilungsfunktion an den Klassengrenzen an.

(b) Stellen Sie die (approximative) empirische Verteilungsfunktion des Merkmals aus der Klassierung der Daten auf.

(c) Berechnen Sie aus den klassierten Daten den (approximativen) arithmetischen Mittelwert der Daten. Wie groß ist die relative Abweichung vom tatsächlichen (aus der Urliste bestimmten) Mittelwert von 71.584?

(d) Welche Näherung für die Anzahl von Merkmalswerten zwischen 50 und 80 können Sie unter Verwendung der approximativen Verteilungsfunktion für die klassierten Daten berechnen?

Aufgabe 5 (4 + 2 = 6 Punkte)

Von den Besuchern eines 2-tägigen Musikfestivals sind 30% nur am 1. Tag sowie 20% nur am 2. Tag auf dem Festivalgelände anwesend (die restlichen 50% sind an beiden Tagen auf dem Festivalgelände anwesend).

(a) Wie wahrscheinlich ist es, dass ein (zufällig ausgewählter) Besucher, der am 1. Tag auf dem Festivalgelände anwesend war, auch am 2. Tag auf dem Festivalgelände anwesend ist?

(b) Wie wahrscheinlich ist es, dass ein (zufällig ausgewählter) Besucher, der am 2. Tag auf dem Festivalgelände anwesend ist, nicht schon am 1. Tag auf dem Festivalgelände anwesend war?

Aufgabe 6 (5 + 2 + 2 = 9 Punkte)

Ein Sanitärinstallationsbetrieb verwendet für die Ausführung von Warmwasser-Installationen drei unterschiedlichen Systeme A, B und C von verschiedenen Herstellern. Dabei werden Warmwasser-Installationen mit einer Wahrscheinlichkeit von 10% mit System A, 20% mit System B und 70% mit System C ausgeführt. Bei den anschließenden Druckprüfungen gibt es erfahrungsgemäß mit einer Wahrscheinlichkeit von 95% bei System A, 96% bei System B und 98% bei System C keine Undichtigkeiten.

(a) Wie groß ist die Wahrscheinlichkeit, dass eine zufällig ausgewählte Warmwasser-Installation bei der Druckprüfung undicht ist?

(b) Wie groß ist die Wahrscheinlichkeit, dass eine bei der Druckprüfung nicht undichte Installation mit System B ausgeführt wurde?

(c) Sind die Ereignisse „Installation ist undicht“ und „System B wurde verwendet“ stochastisch unabhängig? Begründen Sie Ihre Antwort!

Aufgabe 7 (2 + 1 + 2 = 5 Punkte)

Die Wartezeit zwischen zwei Unfällen auf einem stark frequentierten Autobahnabschnitt lasse sich als eine exponentialverteilte Zufallsvariable auffassen. Im Mittel vergehen zwischen zwei Unfällen 4 Tage.

(a) Welche Varianz hat die Wartezeit zwischen zwei Unfällen?

(b) Mit welcher Wahrscheinlichkeit beträgt die Wartezeit zwischen zwei Unfällen mehr als 2 und weniger als 5 Tage?

(c) Berechnen Sie das 0.95-Quantil der Wartezeit zwischen zwei Unfällen.

Aufgabe 8 (5 + 2 + 6 + 1 + 4 = 18 Punkte)

Die Verteilung einer stetigen Zufallsvariablen X sei durch die folgende Dichtefunktion gegeben:

$$f_X(x) = \begin{cases} \frac{1}{3}x + \frac{1}{3} & \text{für } -1 \leq x < 1 \\ -\frac{2}{3}x + \frac{4}{3} & \text{für } 1 \leq x \leq 2 \\ 0 & \text{sonst} \end{cases}$$

(a) Berechnen Sie die Verteilungsfunktion F_X von X.

(b) Berechnen Sie $P\left(\left\{X > \frac{1}{2}\right\}\right)$ und $P\left(\left\{\frac{1}{2} \leq X \leq \frac{3}{2}\right\}\right)$.

(c) Bestimmen Sie den Erwartungswert $\mathrm{E}(X)$.

(d) Ist X symmetrisch um ihren Erwartungswert verteilt (ohne Begründung)?

(e) Bestimmen Sie den Median von X.

Aufgabe 9 (2 + 8 + 1 + 2 + 3 = 16 Punkte)

Ein fairer Würfel wird zweimal geworfen. Es seien X die Anzahl der Würfe mit einer Augenzahl ≤ 3 sowie Y die Anzahl der Würfe mit einer geraden Augenzahl.

(a) Welcher Verteilung genügen X und Y (jeweils)?

(b) Die gemeinsame Verteilung von (X, Y) ist gegeben durch:

$X \setminus Y$	0	1	2
0	$\frac{1}{36}$	$\frac{1}{9}$	$\frac{1}{9}$
1	$\frac{1}{9}$	$\frac{5}{18}$	$\frac{1}{9}$
2	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{36}$

Bestimmen Sie $\mathrm{E}(X)$, $\mathrm{E}(Y)$, $\mathrm{Var}(X)$, $\mathrm{Var}(Y)$, $\mathrm{Cov}(X, Y)$ sowie $\mathrm{Korr}(X, Y)$.

Hinweis: Beachten Sie, dass Sie die diesen Aufgabenteil unter Verwendung der Ergebnisse aus Teil (a) zum Teil recht schnell und insbesondere vollständig ohne die Bestimmung der Randwahrscheinlichkeiten von X und Y lösen können!

(c) Mit welcher Wahrscheinlichkeit nehmen sowohl X als auch Y Werte von mindestens 1 an?

(d) Sind X und Y stochastisch unabhängig? Begründen Sie Ihre Antwort.

(e) Berechnen Sie $\mathrm{E}(3X - 2Y)$ sowie $\mathrm{Var}(3X - 2Y)$.

Aufgabe 10 (2 + 2 + 4 = 8 Punkte)

Ein Online-Händler bietet für 625 der an einem Tag eingehenden Bestellungen einen Express-Lieferservice an, der eine Abfertigung der Bestellung am nächsten Arbeitstag garantiert. Es ist davon auszugehen, dass die Zeitdauern zur Abfertigung einzelner Express-Bestellungen (in Stunden) unabhängig identisch verteilt sind mit einer mittleren Abfertigungsdauer von 0.2 Stunden bei einer Standardabweichung von 0.08 Stunden.

(a) Welchen Erwartungswert und welche Standardabweichung hat die Summe der Abfertigungsdauern von 625 Express-Bestellungen?

(b) Verwenden Sie den zentralen Grenzwertsatz, um die Wahrscheinlichkeit, dass 625 Express-Bestellungen in höchstens 128 (Mitarbeiter-)Stunden abgefertigt werden können, (näherungsweise) zu berechnen.

(c) Verwenden Sie den zentralen Grenzwertsatz, um näherungsweise einen um den zugehörigen Erwartungswert symmetrischen Bereich zu bestimmen, in dem sich die Gesamtabfertigungsdauer von 625 Express-Bestellungen mit einer Wahrscheinlichkeit von 0.95 realisiert.

Hinweis: Verwenden Sie zur Bearbeitung der Aufgabenteile (b) und (c) die Tabelle zur Standardnormalverteilung auf Seite 120!

11 Deskr. Statistik und Wahrscheinlichkeitsrechnung WS 2019/20

Aufgabe 1 (16 Punkte)

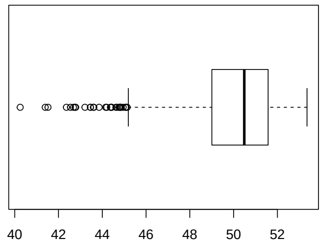
Markieren Sie jeweils mit einem Kreuz pro Aussage im betreffenden Kästchen, ob die unten stehenden Aussagen wahr oder falsch sind.
Richtige Antworten geben 2 Punkte, falsche Antworten und nicht bearbeitete Aussagen 0 Punkte (Aussagen mit zwei Kreuzen zählen als nicht bearbeitet!).

	wahr	falsch
1. Die Höhen der Rechtecke eines Histogramms haben stets die Summe 1.	☐	☐
2. Verdoppelt man alle Urlisteneinträge einer Urliste der Länge $n = 99$ eines kardinalskalierten Merkmals, so verdoppelt sich stets auch der Median des Merkmals.	☐	☐
3. Hat sich der Preis eines bestimmten Produkts in drei aufeinanderfolgenden Jahren jeweils um 2.00%, 3.00% und 7.00% erhöht, so beträgt die durchschnittliche jährliche Preissteigerung dieses Produkts (ggf. gerundet) 4.00%.	☐	☐
4. In einem Laplaceschen Wahrscheinlichkeitsraum mit einer 77-elementigen Ergebnismenge Ω gilt für beliebige Ereignisse $A \subseteq \Omega$ stets $P(A) \neq P(\overline{A})$.	☐	☐
5. Es seien $(\Omega, \mathcal{F}, P)$ ein Wahrscheinlichkeitsraum, $A, B, C \in \mathcal{F}$ drei Ereignisse mit $P(C) > 0$. Dann gilt stets: $$P(A\vert C) + P(B\vert C) \leq 2 \cdot P(A \cup B\vert C)$$	☐	☐
6. Wenn Sie alle 8 Aufgabenteile dieser Aufgabe rein zufällig entweder mit *wahr* oder mit *falsch* beantworten, dann ist die Wahrscheinlichkeit, den Wahrheitsgehalt aller Aussagen korrekt zu bewerten, größer als 1%.	☐	☐
7. Exponentialverteilte Zufallsvariablen sind stets rechtssteil verteilt.	☐	☐
8. Die Zufallsvariablen X und Y seien unkorreliert. Dann sind X und Y auch stochastisch unabhängig.	☐	☐

Aufgabe 2 (12 Punkte)

Markieren Sie jeweils die korrekte Antwort mit einem Kreuz im betreffenden Kästchen. Es ist jeweils genau ein Kreuz korrekt.
Richtige Antworten geben 3 Punkte, falsche Antworten und nicht bearbeitete Aufgabenteile 0 Punkte (Aufgabenteile mit mehr als einem Kreuz zählen als nicht bearbeitet!).

1. Kreuzen Sie an, auf welche Merkmalseigenschaften der folgende Box-Plot *am ehesten* hindeutet:



   (a) leptokurtisch und rechts**steil** ☐

   (b) leptokurtisch und links**steil** ☐

   (c) platykurtisch und rechts**steil** ☐

   (d) platykurtisch und links**steil** ☐

2. Aus der Menge von 9 Einzelbewerbern für den Vorsitz einer politischen Partei soll auf einem außerordentlichen Parteitag eine Doppelspitze gebildet werden, also eine Wahl von 2 (gleichberechtigten) Bewerbern als Vorsitzende erfolgen. Hierfür gibt es insgesamt

   (a) $(9)_2 = \frac{9!}{7!}$ Möglichkeiten. ☐

   (b) 9^2 Möglichkeiten. ☐

   (c) $\binom{9}{2} = \frac{9!}{2! \cdot 7!}$ Möglichkeiten. ☐

   (d) 2^9 Möglichkeiten. ☐

3. Sind X_1, X_2 und X_3 drei stochastisch unabhängige Zufallsvariablen mit $X_1 \sim N(20, 8^2)$, $X_2 \sim N(40, 8^2)$ und $X_3 \sim N(60, 14^2)$, dann ist die Verteilung von $X_1 + X_2 + X_3$ eine

   (a) $N(120, 30^2)$-Verteilung. ☐

   (b) $N(120, 18^2)$-Verteilung. ☐

   (c) $N(60, 30^2)$-Verteilung. ☐

   (d) $N(60, 18^2)$-Verteilung. ☐

4. Das Merkmal X des zweidimensionalen Merkmals (X, Y) sei ordinalskaliert, das Merkmal Y lediglich nominalskaliert. Damit ist die Berechnung folgender Abhängigkeitsmaße zwischen X und Y immer möglich:

   (a) Pearsonscher Korrelationskoeffizient und korrigierter Pearsonscher Kontingenzkoeffizient ☐

   (b) Spearmanscher Rangkorrelationskoeffizient und korrigierter Pearsonscher Kontingenzkoeffizient ☐

   (c) Nur Spearmanscher Rangkorrelationskoeffizient ☐

   (d) Nur korrigierter Pearsonscher Kontingenzkoeffizient ☐

Aufgabe 3 $(3 + 3 + 1 + 2 + 1 = 10$ Punkte)

Bei einer Umfrage wurden 50 Haushalte befragt, wie viele Smart Speaker sie in ihrem Haushalt installiert haben (Merkmal X). Das Ergebnis der Umfrage ist die folgende (bereits aufsteigend sortierte) Urliste zu X:

0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 2, 2, 2, 2, 3, 3, 3, 3, 4, 4, 4, 4, 4, 5, 5, 5

(a) Erstellen Sie eine Tabelle mit den absoluten und relativen Häufigkeiten.

(b) Stellen Sie die zugehörige empirische Verteilungsfunktion auf.

(c) Wie groß ist der Anteil der Haushalte in der Umfrage, die mehr als 1 Smart Speaker in ihrem Haushalt installiert haben?

(d) Berechnen Sie den arithmetischen Mittelwert des Merkmals X.

(e) Bestimmen Sie ein oberes Quartil des Merkmals X.

Aufgabe 4 $(6 + 4 + 3 + 3 + 3 = 19$ Punkte)

Zu einem kardinalskalierten Merkmal sei die folgende (zur einfacheren Bearbeitung der Aufgabe bereits sortierte) Urliste der Länge $n = 40$ gegeben:

22.64, 22.68, 23.82, 24.64, 25.35, 25.50, 26.00, 27.79, 27.80, 28.37, 30.06, 30.32, 30.76, 30.88, 30.95, 32.08, 33.22, 33.79, 34.89, 35.70, 35.73, 36.95, 39.25, 40.05, 40.19, 40.41, 41.00, 43.54, 44.03, 49.35, 50.70, 51.30, 52.41, 54.58, 55.49, 58.54, 59.86, 63.97, 65.20, 72.72

(a) Führen Sie eine Klassierung der erhobenen Daten auf Grundlage der Klassen

$$K_1 = (20, 30],\ K_2 = (30, 40],\ K_3 = (40, 60],\ K_4 = (60, 90]$$

durch. Geben Sie insbesondere die jeweiligen Klassenbreiten, Klassenmitten, absoluten und relativen Klassenhäufigkeiten, Häufigkeitsdichten sowie die Werte der empirischen Verteilungsfunktion an den Klassengrenzen an.

(b) Stellen Sie die (approximative) empirische Verteilungsfunktion des Merkmals aus der Klassierung der Daten auf.

(c) Berechnen Sie aus den klassierten Daten den (approximativen) arithmetischen Mittelwert der Daten. Wie groß ist die relative Abweichung vom tatsächlichen (aus der Urliste bestimmten) Mittelwert von 39.313?

(d) Bestimmen Sie (aus der Urliste) die Anzahl von Merkmalswerten zwischen 35 und 60. Welche Näherung für diese Anzahl können Sie aus der in Teil (b) aufgestellten empirischen Verteilungsfunktion berechnen?

(e) Bestimmen Sie (unter Verwendung der bekannten Konvention zur eindeutigen Festlegung von Quantilen bei kardinalskalierten Merkmalen) *das* obere Quartil sowohl exakt aus der Urliste als auch approximativ mit Hilfe der Verteilungsfunktion für die klassierten Daten.

Aufgabe 5 (2 + 2 + 2 = 6 Punkte)

Bei der Herstellung von Rauspundbrettern tritt mit einer Wahrscheinlichkeit von 2% ein Fehler beim Zuschnitt der Bretter auf, mit einer Wahrscheinlichkeit von 1.5% ein Fehler beim Fräsen der Bretter und mit einer Wahrscheinlichkeit von 1% sowohl ein Fehler beim Zuschnitt der Bretter als auch ein Fehler beim Fräsen der Bretter. Berechnen Sie die Wahrscheinlichkeit dafür, dass

(a) höchstens einer der beiden Fehler,

(b) mindestens einer der beiden Fehler,

(c) ein Fehler beim Zuschnitt der Bretter, aber kein Fehler beim Fräsen der Bretter

auftritt.

Aufgabe 6 (6 + 2 + 2 = 10 Punkte)

Ein Versandhaus beauftragt für den Versand seiner Sendungen einen von insgesamt vier verschiedenen Versand-Dienstleistern A, B, C und D. Dabei werden durchschnittlich 15% der Sendungen an Dienstleister A, 25% der Sendungen an Dienstleister B, 30% der Sendungen an Dienstleister C und 30% der Sendungen an Dienstleister D übergeben. Die umfangreiche Auswertung der vorhandenen Kunden-Feedbacks zu Qualität und Geschwindigkeit der Lieferung ergab, dass 5% der Lieferungen mit Dienstleister A, 5% der Lieferungen mit Dienstleister B, 2% der Lieferungen mit Dienstleister C und 2% der Lieferungen mit Dienstleister D beanstandet werden.

(a) Wie groß ist die Wahrscheinlichkeit, dass eine zufällig ausgewählte Sendung keinen Anlass zur Beanstandung durch den Kunden bietet?

(b) Wie groß ist die Wahrscheinlichkeit, dass eine beanstandete Lieferung mit Dienstleister B versendet wurde?

(c) Sind die Ereignisse „Lieferung wird beanstandet“ und „Dienstleister B war mit dem Versand beauftragt“ stochastisch unabhängig? Begründen Sie Ihre Antwort!

Aufgabe 7 (5 + 2 + 6 + 1 + 4 = 18 Punkte)

Die Verteilung einer stetigen Zufallsvariablen X sei durch die folgende Dichtefunktion gegeben:

$$f_X(x) = \begin{cases} \frac{1}{6}x + \frac{1}{2} & \text{für } -2 \leq x < 0 \\ -\frac{1}{6}x + \frac{1}{3} & \text{für } 0 \leq x \leq 2 \\ 0 & \text{sonst} \end{cases}$$

(a) Berechnen Sie die Verteilungsfunktion F_X von X.

(b) Berechnen Sie $P(\{X < -1\})$ und $P(\{-1 < X < 1\})$.

(c) Bestimmen Sie den Erwartungswert $\mathrm{E}(X)$.

(d) Ist X symmetrisch um ihren Erwartungswert verteilt (ohne Begründung)?

(e) Bestimmen Sie den Median von X.

Aufgabe 8 (3 + 3 = 6 Punkte)

Die Anzahl der Störfälle pro Jahr in einem bestimmten Atomkraftwerk (AKW) lasse sich als eine Pois(0.2)-verteilte Zufallsvariable auffassen. Außerdem soll angenommen werden, dass die Anzahl der Störfälle pro Jahr in dem betreffenden AKW für unterschiedliche Jahre stochastisch unabhängig ist.

(a) Mit welcher Wahrscheinlichkeit ereignet sich in einem Jahr in dem betreffenden AKW höchstens 1 Störfall?

(b) Welche Verteilung hat die Anzahl der Störfälle pro Jahrzehnt in dem betreffenden AKW? Mit welcher Wahrscheinlichkeit ereignet sich in einem Zehnjahreszeitraum in dem betreffenden AKW mindestens 1 Störfall?

Aufgabe 9 (2 + 9 + 1 + 3 = 15 Punkte)

Gegeben sei die folgende Tabelle der gemeinsamen Wahrscheinlichkeitsverteilung zu einem zweidimensionalen diskreten Zufallsvektor (X, Y):

$X \setminus Y$	0	1	2	$p_{i\cdot}$
0	0.2	0.1	0.2	
2	0.05	0.1	0.1	
4	0.1	0.1	0.05	
$p_{\cdot j}$				

(a) Ergänzen Sie die obige Tabelle (in den vorgesehenen Feldern) um ihre Randverteilungen.

(b) Berechnen Sie $\mathrm{E}(X)$, $\mathrm{E}(Y)$, $\mathrm{Var}(X)$, $\mathrm{Var}(Y)$, $\mathrm{Cov}(X, Y)$ sowie $\mathrm{Korr}(X, Y)$.

(c) Sind X und Y stochastisch unabhängig?

(d) Berechnen Sie $\mathrm{E}\left(2X + 4Y\right)$ sowie $\mathrm{Var}\left(2X + 4Y\right)$.

Aufgabe 10 (3 + 2 + 3 = 8 Punkte)

In einem Hotel mit 375 Zimmern können Zimmerreservierungen bis zum Anreisetag kostenlos storniert werden. Man weiß aus Erfahrung, dass im Mittel 10% der reservierten Zimmer tatsächlich kurzfristig storniert werden. Um die Zahl der freien Zimmer möglichst gering zu halten, nimmt das Hotel daher mehr Zimmerreservierungen an als Zimmer im Hotel vorhanden sind.

(a) Wie ist die Anzahl Y der tatsächlich wegen Reservierungen benötigten (also nicht stornierten) Zimmer verteilt, wenn insgesamt 400 Zimmerreservierungen angenommen wurden und davon ausgegangen werden kann, dass das Stornierungsverhalten der Hotelgäste voneinander unabhängig ist? Geben Sie auch den Erwartungswert und die Varianz von Y an.

(b) Berechnen Sie unter Verwendung des zentralen Grenzwertsatzes die Wahrscheinlichkeit dafür, dass bei 400 angenommenenen Reservierungen genügend Zimmer zur Verfügung stehen, um alle Hotelgäste, die reserviert und nicht storniert haben, auch im Hotel unterzubringen.

(c) Verwenden Sie den zentralen Grenzwertsatz, um näherungsweise ein 0.95-Quantil der Anzahl in Anspruch genommener Zimmerreservierungen Y zu bestimmen.

Hinweis: Verwenden Sie zur Bearbeitung von Aufgabenteil (b) und (c) die Tabelle zur Standardnormalverteilung auf Seite 120!

12 Deskr. Statistik und Wahrscheinlichkeitsrechnung SS 2020

Aufgabe 1 (16 Punkte)

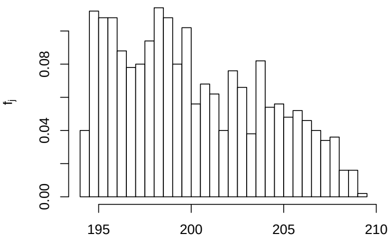
Markieren Sie jeweils mit einem Kreuz pro Aussage im betreffenden Kästchen, ob die unten stehenden Aussagen wahr oder falsch sind.
Richtige Antworten geben 2 Punkte, falsche Antworten und nicht bearbeitete Aussagen 0 Punkte (Aussagen mit zwei Kreuzen zählen als nicht bearbeitet!).

	wahr	falsch
1. Bei einer Urliste der Länge n zu einem ordinalskalierten Merkmal stimmt die Summe der Werte des zugehörigen Rangmerkmals – auch wenn Bindungen vorliegen – stets mit der Summe der ersten n natürlichen Zahlen überein.	☐	☐
2. Wenn es in drei aufeinanderfolgenden Jahren tarifliche Lohnerhöhungen von jeweils 2.5%, 3.2% beziehungsweise 5.1% gibt, dann beträgt die mittlere Lohnerhöhung pro Jahr in diesem Zeitraum (gegebenenfalls gerundet) 3.59%.	☐	☐
3. Für den korrigierten Pearsonschen Kontingenzkoeffizienten $C_{X,Y}^{\text{korr}}$ zweier Merkmale X und Y gelte $C_{X,Y}^{\text{korr}} = 0$. Dann sind X und Y stets unabhängig.	☐	☐
4. Es seien A und B zwei Ereignisse in einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ mit $P(A) = 0.3$, $P(B) = 0.5$ und $P(A \cup B) = 0.7$. Damit gilt $P(A \cap B) = 0.1$.	☐	☐
5. Es seien $(\Omega, \mathcal{F}, P)$ ein Wahrscheinlichkeitsraum, $A, B, C \in \mathcal{F}$ drei Ereignisse mit $P(C) > 0$. Dann gilt: $$P(A) = P(B) \qquad \Rightarrow \qquad P(A\|C) = P(B\|C)$$	☐	☐
6. Werden zwei faire (sechsseitige) Würfel gleichzeitig geworfen, so erhält man mit einer Wahrscheinlichkeit von $\frac{15}{21}$ zwei unterschiedliche Augenzahlen.	☐	☐
7. Die Summe von vier stochastisch unabhängigen $N(16, 2^2)$-verteilten Zufallsvariablen ist $N(64, 4^2)$-verteilt.	☐	☐
8. Ist die Kovarianz zweier Zufallsvariablen X und Y negativ, so ist die Varianz der Summe von X und Y kleiner als die Summe der einzelnen Varianzen (von X und Y).	☐	☐

Aufgabe 2 (12 Punkte)

Markieren Sie jeweils die korrekte Antwort mit einem Kreuz im betreffenden Kästchen. Es ist jeweils genau ein Kreuz korrekt.
Richtige Antworten geben 3 Punkte, falsche Antworten und nicht bearbeitete Aufgabenteile 0 Punkte (Aufgabenteile mit mehr als einem Kreuz zählen als nicht bearbeitet!).

1. Kreuzen Sie an, auf welche Merkmalseigenschaften das folgende Histogramm *am ehesten* hindeutet:



(a) leptokurtisch und rechts**steil** ☐
(b) leptokurtisch und links**steil** ☐
(c) platykurtisch und rechts**steil** ☐
(d) platykurtisch und links**steil** ☐

2. Die Anzahl der verschiedenen (sechsstelligen) Zahlen, die aus den Ziffern 2,2,2,5,5,7 gebildet werden können, beträgt:

(a) $\frac{3^6}{3 \cdot 2 \cdot 1}$ ☐
(b) $\frac{6!}{3! \cdot 2! \cdot 1!}$ ☐
(c) $2^3 \cdot 5^2 \cdot 7^1$ ☐
(d) $3! \cdot 2! \cdot 1!$ ☐

3. Wenn Sie an einem schönen Sommertag besonders viel Appetit auf Eis haben und daher nacheinander gleich drei verschiedene Eisspezialitäten aus der beachtenswerten Gefriertruhen-Auswahl von acht Sorten als Nachtisch verspeisen möchten, so haben Sie zur Wahl der Sorten und Reihenfolge beim Verzehr insgesamt

(a) 3^8 Möglichkeiten. ☐
(b) $(8)_3 = \frac{8!}{5!}$ Möglichkeiten. ☐
(c) $\binom{8}{3} = \frac{8!}{3! \cdot 5!}$ Möglichkeiten. ☐
(d) 8^3 Möglichkeiten. ☐

4. Beim Zufallsexperiment des einmaligen Würfelns mit einem gewöhnlichen sechsseitigen Würfel

(a) sind $\{1, 6\}$ und 4 jeweils Ergebnisse. ☐
(b) sind $\{1, 6\}$ und 4 jeweils Ereignisse. ☐
(c) ist $\{1, 6\}$ ein Ergebnis und 4 ein Ereignis. ☐
(d) ist $\{1, 6\}$ ein Ereignis und 4 ein Ergebnis. ☐

Aufgabe 3 (3 + 3 + 1 + 5 = 12 Punkte)

Bei einer Umfrage wurden 40 Personen befragt, wie viele Smartspeaker sie in den vergangenen drei Jahren gekauft haben (Merkmal X). Das Ergebnis der Umfrage ist die folgende (bereits aufsteigend sortierte) Urliste zu X:

0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 2, 2, 2, 3, 3, 3, 3, 3, 4, 4, 4, 4, 4, 4, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5, 5

(a) Erstellen Sie eine Tabelle mit den absoluten und relativen Häufigkeiten.

(b) Stellen Sie die zugehörige empirische Verteilungsfunktion auf.

(c) Wie groß ist der Anteil der Personen in der Umfrage, die mindestens 3 Smartspeaker in den vergangenen drei Jahren gekauft haben?

(d) Berechnen Sie den arithmetischen Mittelwert und die empirische Varianz des Merkmals X.

Aufgabe 4 (6 + 4 + 3 + 3 = 16 Punkte)

Zu einem kardinalskalierten Merkmal sei die folgende (zur einfacheren Bearbeitung der Aufgabe bereits sortierte) Urliste der Länge $n = 40$ gegeben:

6.12, 7.58, 7.76, 8.09, 9.18, 9.33, 9.94, 9.94, 10.57, 10.68, 11.00, 11.19, 11.28, 11.33, 11.48, 11.59, 11.60, 13.46, 13.72, 13.83, 14.15, 14.54, 14.96, 15.82, 16.11, 16.38, 16.95, 17.13, 17.67, 18.12, 18.50, 18.95, 19.08, 20.15, 22.61, 23.12, 23.49, 26.06, 27.12, 27.87

(a) Führen Sie eine Klassierung der erhobenen Daten auf Grundlage der Klassen

$$K_1 = (5, 10],\ K_2 = (10, 15],\ K_3 = (15, 25],\ K_4 = (25, 35]$$

durch. Geben Sie insbesondere die jeweiligen Klassenbreiten, Klassenmitten, absoluten und relativen Klassenhäufigkeiten, Häufigkeitsdichten sowie die Werte der empirischen Verteilungsfunktion an den Klassengrenzen an.

(b) Stellen Sie die (approximative) empirische Verteilungsfunktion des Merkmals aus der Klassierung der Daten auf.

(c) Bestimmen Sie (aus der Urliste) die Anzahl von Merkmalswerten zwischen 10 und 30. Welche Näherung für diese Anzahl können Sie aus der in Teil (b) aufgestellten empirischen Verteilungsfunktion berechnen?

(d) Bestimmen Sie (unter Verwendung der bekannten Konvention zur eindeutigen Festlegung von Quantilen bei kardinalskalierten Merkmalen) *das* obere Quartil sowohl exakt aus der Urliste als auch approximativ mit Hilfe der Verteilungsfunktion für die klassierten Daten.

Aufgabe 5 (4 + 2 + 2 = 8 Punkte)

Um eine überstandene Infektion an einer neuartigen Viruserkrankung zu identifizieren, wird ein Antikörpertest verwendet, der laut Herstellerangabe eine tatsächlich überstandene Infektion auch mit einer Wahrscheinlichkeit von 99% erkennt. Gleichzeitig zeigt der

Antikörpertest aber auch mit einer Wahrscheinlichkeit von 2% fälschlicherweise eine überstandene Infektion an, obwohl tatsächlich keine Infektion vorlag. Gehen Sie (außer von der Korrektheit der Herstellerangaben) nun (zunächst) davon aus, dass der Antikörpertest in einer Region flächendeckend eingesetzt werden soll und bereits 3% der Bevölkerung der betreffenden Region die Virusinfektion überstanden haben.

(a) Mit welcher Wahrscheinlichkeit wird der Antikörpertest bei einer (zufällig ausgewählten) untersuchten Person eine überstandene Infektion anzeigen?

(b) Mit welcher Wahrscheinlichkeit hat eine (mit dem Antikörpertest) positiv getestete Person tatsächlich noch gar keine Infektion mit dem neuartigen Virus überstanden?

(c) Wie ändern sich die Ergebnisse aus Teil (a) und Teil (b), wenn in der betreffenden Region bereits 10% der Bevölkerung die Virusinfektion überstanden haben?

Aufgabe 6 (3 + 3 + 3 + 3 = 12 Punkte)

An einer hochinfektiösen Erkrankung seien zu einem bestimmten Zeitpunkt 0.5% der Bevölkerung eines Landes erkrankt. Im Folgenden soll die Entwicklung des Infektionsgeschehens untersucht werden, wenn sich Personengruppen innerhalb geschlossener Räumlichkeiten treffen. Gehen Sie dazu (näherungsweise) davon aus, dass die Anzahl infizierter Personen in einer Gruppe von n Personen $B(n, 0.005)$-verteilt ist. Ferner soll davon ausgegangen werden, dass nach einem Treffen sämtliche Personen einer Gruppe infiziert sind, falls mindestens ein Teilnehmer des Treffens vorher bereits infiziert war (und andernfalls niemand in dieser Gruppe).

(a) Wie groß ist die Wahrscheinlichkeit (in Abhängigkeit von n), dass mindestens eine Person in einer Gruppe der Größe n (vor dem Treffen) infiziert ist?

(b) Wie groß ist bei einer Gruppe der Größe n der Erwartungswert (in Abhängigkeit von n) der Anzahl der nach dem Treffen (in einer geschlossenen Räumlichkeit) infizierten Personen?

(c) Vergleichen Sie die erwartete Gesamtanzahl der nach den Treffen infizierten Personen zwischen einem Treffen mit 100 Personen und 20 Treffen mit jeweils 5 Personen, indem Sie den Erwartungswert aus Teil (b) für $n = 100$ mit dem 20-fachen des Erwartungswerts für $n = 5$ vergleichen. Ist aus dieser Sicht eine Beschränkung der Personenzahl für Treffen innerhalb geschlossener Räumlichkeiten sinnvoll?

(d) Welchen Erwartungswert erhalten Sie für die Anzahl der nach dem Treffen infizierten Personen für $n = 100$, wenn Sie zur Näherung der Binomialverteilung in Teil (a) und (b) eine geeignete Poisson-Verteilung verwenden?

Aufgabe 7 (5 + 4 + 2 + 6 + 1 = 18 Punkte)

Die Verteilung einer stetigen Zufallsvariablen X sei durch die folgende Dichtefunktion gegeben:

$$f_X(x) = \begin{cases} \frac{1}{8}x & \text{für } 1 \leq x < 3 \\ -\frac{1}{8}x + \frac{3}{4} & \text{für } 3 \leq x \leq 5 \\ 0 & \text{sonst} \end{cases}$$

(a) Berechnen Sie die Verteilungsfunktion F_X von X.

(b) Bestimmen Sie das obere Quartil von X.

(c) Berechnen Sie $P\left(\left\{X > \frac{3}{2}\right\}\right)$ und $P\left(\left\{\frac{3}{2} \leq X \leq 4\right\}\right)$.

(d) Bestimmen Sie den Erwartungswert $\mathrm{E}(X)$.

(e) Ist X symmetrisch um ihren Erwartungswert verteilt (ohne Begründung)?

Aufgabe 8 (2 + 3 + 9 + 3 = 17 Punkte)

Gegeben sei die folgende Tabelle der gemeinsamen Wahrscheinlichkeitsverteilung zu einem zweidimensionalen diskreten Zufallsvektor (X, Y):

$X \setminus Y$	2	3	4	$p_{i\cdot}$
-2	$\frac{1}{25}$	$\frac{1}{15}$	$\frac{1}{12}$	
1	$\frac{4}{25}$	$\frac{1}{15}$	$\frac{11}{60}$	
4	$\frac{1}{5}$	$\frac{1}{15}$	$\frac{2}{15}$	
$p_{\cdot j}$				

(a) Ergänzen Sie die obige Tabelle (in den vorgesehenen Feldern) um ihre Randverteilungen.

(b) Geben Sie die bedingten Wahrscheinlichkeitsverteilungen von X unter der Bedingung $Y = y_j$ für alle $y_j \in T(Y)$ über die zugehörigen (bedingten) Wahrscheinlichkeitsfunktionen an.

(c) Berechnen Sie $\mathrm{E}(X)$, $\mathrm{E}(Y)$, $\mathrm{Var}(X)$, $\mathrm{Var}(Y)$, $\mathrm{Cov}(X, Y)$ sowie $\mathrm{Korr}(X, Y)$.

(d) Berechnen Sie $\mathrm{E}\,(5X - 3Y)$ sowie $\mathrm{Var}\,(5X - 3Y)$.

Aufgabe 9 (3 + 2 + 4 = 9 Punkte)

Aufgrund langjähriger Aufzeichnungen über entsprechende Wahlbeteiligungen gehe man davon aus, dass sich die 14500 wahlberechtigten Studierenden bei einer anstehenden Senatswahl unabhängig voneinander jeweils mit einer Wahrscheinlichkeit von 10% dazu entschließen, ihr Wahlrecht auch auszuüben und einen Stimmzettel auszufüllen.

(a) Wie ist die Anzahl der ausgefüllten Stimmzettel Y exakt verteilt? Geben Sie auch den Erwartungswert $\mathrm{E}(Y)$ sowie die Varianz $\mathrm{Var}(Y)$ der Anzahl der ausgefüllten Stimmzettel an.

(b) Die Wahlleitung entschließt sich aus ökologischen und ökonomischen Gründen, zunächst nur 1500 Stimmzettel für die Wahl auszudrucken und weitere Stimmzettel erst bei Bedarf nachzudrucken, falls mehr als die zunächst gedruckten Stimmzettel benötigt werden sollten. Mit welcher (mit dem zentralen Grenzwertsatz näherungsweise zu bestimmenden) Wahrscheinlichkeit müssen keine Stimmzettel nachgedruckt werden?

(c) Verwenden Sie den zentralen Grenzwertsatz, um näherungsweise einen um die erwartete Anzahl ausgefüllter Stimmzettel symmetrischen Bereich zu bestimmen, in dem sich die Anzahl ausgefüllter Stimmzettel mit einer Wahrscheinlichkeit von 0.99 realisiert.

Hinweis: Verwenden Sie zur Bearbeitung von Aufgabenteil (b) und (c) die Tabelle zur Standardnormalverteilung auf Seite 120!

13 Deskr. Statistik und Wahrscheinlichkeitsrechnung WS 2020/21

<u>**Aufgabe 1**</u> (16 Punkte)

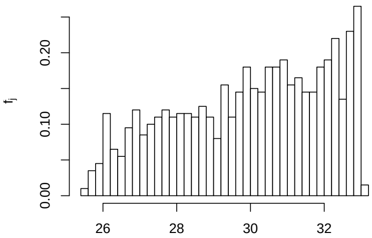
Markieren Sie jeweils mit einem Kreuz pro Aussage im betreffenden Kästchen, ob die unten stehenden Aussagen wahr oder falsch sind.
Richtige Antworten geben 2 Punkte, falsche Antworten und nicht bearbeitete Aussagen 0 Punkte (Aussagen mit zwei Kreuzen zählen als nicht bearbeitet!).

	wahr	falsch
1. Bei einer Urliste der Länge n zu einem ordinalskalierten Merkmal stimmt das arithmetische Mittel der Werte des zugehörigen Rangmerkmals – auch wenn Bindungen vorliegen – stets mit dem arithmetischen Mittel der ersten n natürlichen Zahlen überein.	☐	☐
2. In einem Laplaceschen Wahrscheinlichkeitsraum gilt für zwei Ereignisse A und B stets $P(A \cap B) = P(A) \cdot P(B)$.	☐	☐
3. Wenn ein Schnelltest auf eine infektiöse Krankheit bei Vorliegen einer akuten Infektion mit einer Wahrscheinlichkeit von 98% (korrekterweise) ein positives Resultat liefert, dann ist das gleichbedeutend damit, dass man bei Vorliegen eines negativen Testergebnisses mit einer Wahrscheinlichkeit von 2% dennoch akut infiziert ist.	☐	☐
4. Wenn Sie zweimal nacheinander mit einem fairen (sechsseitigen) Würfel würfeln, dann ist die zweite gewürfelte Zahl mit einer Wahrscheinlichkeit von $\frac{18}{36} = \frac{1}{2}$ mindestens so groß wie die erste gewürfelte Zahl.	☐	☐
5. Wenn Sie in Aufgabe 2 dieser Klausur in allen 4 Aufgabenteilen jeweils rein zufällig genau eine der 4 Antwortmöglichkeiten ankreuzen, dann erzielen Sie dort mit einer Wahrscheinlichkeit von $\frac{1}{16}$ die volle Punktzahl.	☐	☐
6. Hat die Verteilungsfunktion F_X einer Zufallsvariablen X mindestens eine Sprungstelle, so kann X nicht diskret verteilt sein.	☐	☐
7. Exponentialverteilte Zufallsvariablen sind stets linkssteil verteilt.	☐	☐
8. Die Summe vier unabhängiger $B(25, 0.05)$-verteilter Zufallsvariablen ist $B(100, 0.2)$-verteilt.	☐	☐

Aufgabe 2 (12 Punkte)

Markieren Sie jeweils die korrekte Antwort mit einem Kreuz im betreffenden Kästchen. Es ist jeweils genau ein Kreuz korrekt.
Richtige Antworten geben 3 Punkte, falsche Antworten und nicht bearbeitete Aufgabenteile 0 Punkte (Aufgabenteile mit mehr als einem Kreuz zählen als nicht bearbeitet!).

1. Kreuzen Sie an, auf welche Merkmalseigenschaften das folgende Histogramm *am ehesten* hindeutet:



(a) leptokurtisch und rechts**steil** ☐
(b) leptokurtisch und links**steil** ☐
(c) platykurtisch und rechts**steil** ☐
(d) platykurtisch und links**steil** ☐

2. Auf der Bank der Ersatzspieler einer Fußballmannschaft sitzen 9 Spieler. Wenn während des Fußballspiels 5 Ersatzspieler eingewechselt werden und die Reihenfolge der Einwechslungen keine Rolle spielen soll, so beträgt die Anzahl der verschiedenen Einwechslungsmöglichkeiten (für diese Mannschaft) insgesamt:

(a) 9^5 ☐
(b) 5^9 ☐
(c) $(9)_5 = \frac{9!}{4!}$ ☐
(d) $\binom{9}{5} = \frac{9!}{5! \cdot 4!}$ ☐

3. Verteilungsfunktionen eindimensionaler Zufallsvariablen sind stets

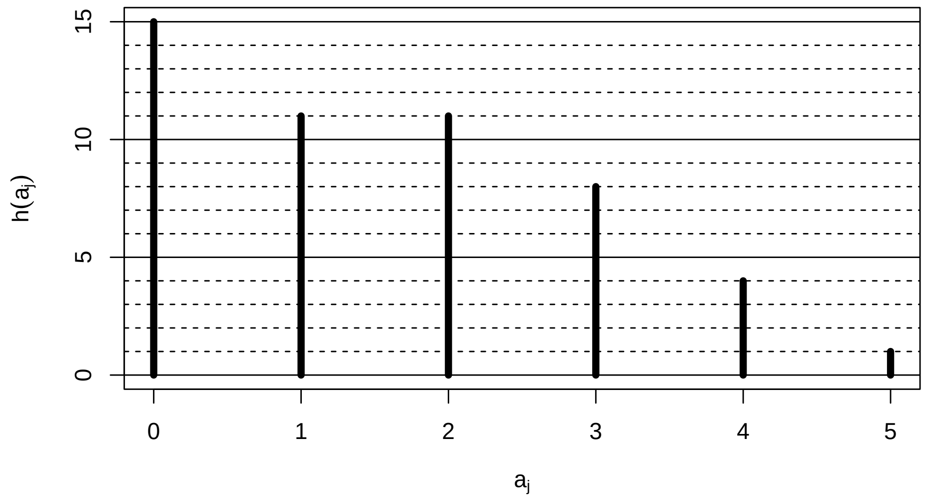
(a) monoton wachsend und rechtsseitig stetig. ☐
(b) monoton wachsend und stetig. ☐
(c) streng monoton wachsend und rechtsseitig stetig. ☐
(d) streng monoton wachsend und stetig. ☐

4. Der zweidimensionale diskrete Zufallsvektor (X, Y) besitze 7 Trägerpunkte, die alle auf einer Geraden mit Steigung -0.1 liegen. Dann gilt:

(a) $\text{Korr}(X, Y) = -0.1$ ☐
(b) $\text{Korr}(X, Y) = -0.7$ ☐
(c) $\text{Korr}(X, Y) = +0.7$ ☐
(d) $\text{Korr}(X, Y) = -1$ ☐

Aufgabe 3 (4 + 5 + 3 + 1 + 1 = 14 Punkte)

Zu einem erhobenen Merkmal X sei das folgende Stabdiagramm gegeben:



(a) Erstellen Sie eine Tabelle der absoluten und relativen Häufigkeiten.

(b) Berechnen Sie den arithmetischen Mittelwert und die empirische Standardabweichung des Merkmals X.

(c) Geben Sie die empirische Verteilungsfunktion des Merkmals X an.

(d) Wie groß ist der Anteil der Urlisteneinträge, die Werte von weniger als 2 annehmen?

(e) Bestimmen Sie ein oberes Quartil des Merkmals X.

Aufgabe 4 (6 + 4 + 3 + 3 = 16 Punkte)

Zu einem kardinalskalierten Merkmal sei die folgende (zur einfacheren Bearbeitung der Aufgabe bereits sortierte) Urliste der Länge $n = 40$ gegeben:

15.83, 18.98, 19.55, 20.41, 21.02, 21.38, 21.48, 23.25, 23.32, 23.68, 27.02, 27.19, 28.10, 28.78, 29.33, 29.92, 30.10, 31.35, 31.42, 32.32, 32.78, 33.69, 33.90, 34.10, 34.20, 34.49, 34.60, 35.42, 36.62, 36.96, 37.08, 37.09, 37.83, 38.12, 38.33, 38.68, 38.69, 38.84, 38.93, 39.29

(a) Führen Sie eine Klassierung der erhobenen Daten auf Grundlage der Klassen

$$K_1 = (10, 20],\ K_2 = (20, 30],\ K_3 = (30, 35],\ K_4 = (35, 40]$$

durch. Geben Sie insbesondere die jeweiligen Klassenbreiten, Klassenmitten, absoluten und relativen Klassenhäufigkeiten, Häufigkeitsdichten sowie die Werte der empirischen Verteilungsfunktion an den Klassengrenzen an.

(b) Stellen Sie die (approximative) empirische Verteilungsfunktion des Merkmals aus der Klassierung der Daten auf.

(c) Bestimmen Sie (aus der Urliste) die Anzahl von Merkmalswerten zwischen 25 und 35. Welche Näherung für diese Anzahl können Sie aus der in Teil (b) aufgestellten empirischen Verteilungsfunktion berechnen?

(d) Bestimmen Sie (unter Verwendung der bekannten Konvention zur eindeutigen Festlegung von Quantilen bei kardinalskalierten Merkmalen) *den* Median sowohl exakt aus der Urliste als auch approximativ mit Hilfe der Verteilungsfunktion für die klassierten Daten.

Aufgabe 5 (1 + 2 + 2 = 5 Punkte)

Eine Urne enthält 50 gleichartige Kugeln, von denen 8 schwarz und gepunktet, 12 weiß und gepunktet, 12 schwarz und kariert sowie 18 weiß und kariert sind. Wie groß ist die Wahrscheinlichkeit dafür, dass

(a) eine rein zufällig aus der Urne entnommene Kugel schwarz und kariert ist?

(b) eine rein zufällig aus der Urne entnommene Kugel weiß ist?

(c) eine rein zufällig aus der Urne entnommene Kugel gepunktet ist, wenn man weiß, dass sie schwarz ist?

Aufgabe 6 (6 + 2 + 2 = 10 Punkte)

Ein Hersteller von Tiefkühlfertiggerichten bezieht seine Frischfischlieferungen von den vier Großhändlern A, B, C und D. Dabei werden einzelne Lieferungen mit einer Wahrscheinlichkeit von 15% von Lieferant A, 15% von Lieferant B, 50% von Lieferant C und 20% von Lieferant D geliefert. Bei den anschließenden Qualitätskontrollen gibt es erfahrungsgemäß mit einer Wahrscheinlichkeit von 97% bei Lieferant A, 96% bei Lieferant B, 98% bei Lieferant C und 97% bei Lieferant D keinen Anlass zu Beanstandungen.

(a) Wie groß ist die Wahrscheinlichkeit, dass eine zufällig ausgewählte Lieferung in der Qualitätskontrolle nicht beanstandet wird?

(b) Wie groß ist die Wahrscheinlichkeit, dass eine von der Qualitätskontrolle beanstandete Lieferung von Großhändler C geliefert wurde?

(c) Sind die Ereignisse „Lieferung wird beanstandet“ und „Lieferung stammt von Großhändler C“ stochastisch unabhängig? Begründen Sie Ihre Antwort!

Aufgabe 7 (3 + 2 + 6 + 1 + 4 = 16 Punkte)

Die Verteilung einer stetigen Zufallsvariablen X sei durch die folgende Verteilungsfunktion gegeben:

$$F_X(x) = \begin{cases} 0 & \text{für } x \leq -1 \\ \frac{1}{3}x^2 + \frac{2}{3}x + \frac{1}{3} & \text{für } -1 < x \leq 0 \\ -\frac{1}{6}x^2 + \frac{2}{3}x + \frac{1}{3} & \text{für } 0 < x \leq 2 \\ 1 & \text{für } x > 2 \end{cases}$$

(a) Berechnen Sie eine Dichtefunktion f_X von X.

(b) Berechnen Sie $P(\{X > 1\})$ und $P(\{0 < X \leq 1\})$.

(c) Bestimmen Sie den Erwartungswert $\mathrm{E}(X)$.

(d) Ist X symmetrisch um ihren Erwartungswert verteilt (ohne Begründung)?

(e) Bestimmen Sie den Median von X.

Aufgabe 8 (2 + 3 = 5 Punkte)

Als Hausaufgabe im Fach Bildende Kunst waren die Geburtsorte von 20 Bildhauern auswendig zu lernen. Die Schülerin Elsa Emsig hat 17 dieser Geburtsorte auswendig gelernt (die Chance, bei den anderen 3 Geburtsorten durch Raten eine richtige Antwort zu geben, sei gleich Null). Der Lehrer überprüft, ob Elsa die Hausaufgabe ordentlich erledigt hat, indem er 4 Mal rein zufällig und unabhängig voneinander einen der Bildhauer auswählt und die zugehörigen Geburtsorte abfragt. Kann Elsa mindestens zu 3 dieser 4 Bildhauer die Geburtsorte korrekt angeben, so ist die Überprüfung bestanden.

(a) Welche Verteilung besitzt die Anzahl der von Elsa abgegebenen richtigen Antworten?

(b) Mit welcher Wahrscheinlichkeit besteht Elsa die Überprüfung?

Aufgabe 9 (2 + 3 + 9 + 1 + 3 = 18 Punkte)

Gegeben sei die folgende Tabelle der gemeinsamen Wahrscheinlichkeitsverteilung zu einem zweidimensionalen diskreten Zufallsvektor (X, Y):

$X \setminus Y$	-1	0	2	$p_{i\cdot}$
1	0.04	0.16	0.2	
3	0.08	0.16	0.16	
5	0.08	0.08	0.04	
$p_{\cdot j}$				

(a) Ergänzen Sie die obige Tabelle (in den vorgesehenen Feldern) um ihre Randverteilungen.

(b) Geben Sie die bedingten Wahrscheinlichkeitsverteilungen von X unter der Bedingung $Y = y_j$ für alle $y_j \in T(Y)$ über die zugehörigen (bedingten) Wahrscheinlichkeitsfunktionen an.

(c) Berechnen Sie $\mathrm{E}(X)$, $\mathrm{E}(Y)$, $\mathrm{Var}(X)$, $\mathrm{Var}(Y)$, $\mathrm{Cov}(X, Y)$ sowie $\mathrm{Korr}(X, Y)$.

(d) Sind X und Y stochastisch unabhängig?

(e) Berechnen Sie $\mathrm{E}(4X - 3Y)$ sowie $\mathrm{Var}(4X - 3Y)$.

Aufgabe 10 (2 + 2 + 4 = 8 Punkte)

Ein freiberuflicher Netzwerktechniker benötigt für das Auflegen eines Netzwerkanschlusses im Mittel 10 Minuten bei einer Standardabweichung von 2 Minuten. Man kann davon ausgehen, dass die benötigten Zeitdauern für die einzelnen Anschlüsse nicht gegenseitig voneinander abhängen. In einem bestimmten Schaltschrank sind insgesamt 36 Anschlüsse aufzulegen.

(a) Geben Sie die den Erwartungswert sowie die Standardabweichung der gesamten Arbeitszeit (für alle 36 Netzwerkanschlüsse) an.

(b) Verwenden Sie den zentralen Grenzwertsatz, um näherungsweise zu bestimmen, mit welcher Wahrscheinlichkeit der Netzwerktechniker nicht länger als 6.25 Stunden bzw. 375 Minuten zum Auflegen aller Anschlüsse benötigt.

(c) Verwenden Sie den zentralen Grenzwertsatz, um näherungsweise einen um den zugehörigen Erwartungswert symmetrischen Bereich zu bestimmen, in dem sich die gesamte Arbeitszeit für 36 Netzwerkanschlüsse mit einer Wahrscheinlichkeit von 0.95 realisiert.

Hinweis: Verwenden Sie zur Bearbeitung von Aufgabenteil (b) und (c) die Tabelle zur Standardnormalverteilung auf Seite 120!

14 Deskr. Statistik und Wahrscheinlichkeitsrechnung SS 2021

Aufgabe 1 (16 Punkte)

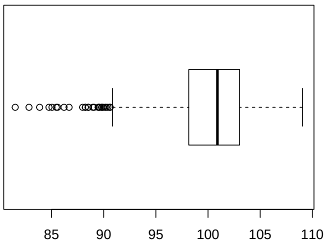
Markieren Sie jeweils mit einem Kreuz pro Aussage im betreffenden Kästchen, ob die unten stehenden Aussagen wahr oder falsch sind.
Richtige Antworten geben 2 Punkte, falsche Antworten und nicht bearbeitete Aussagen 0 Punkte (Aussagen mit zwei Kreuzen zählen als nicht bearbeitet!).

	wahr	falsch
1. Jeder Median eines Merkmals X kommt mindestens einmal in der Urliste zu X vor.	☐	☐
2. Ist (a_i, b_j) eine Ausprägung des zweidimensionalen Merkmals (X, Y), dann ist die gemeinsame relative Häufigkeit $r(a_i, b_j)$ dieser Ausprägung niemals größer als die zugehörige relative Randhäufigkeit $r(a_i)$ der Ausprägung a_i des Merkmals X.	☐	☐
3. Es seien A und B zwei Ereignisse in einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$. Dann gilt stets $P(A) \geq P(A \cap B)$.	☐	☐
4. Ist $(\Omega, \mathcal{F}, P)$ ein beliebiger Wahrscheinlichkeitsraum und sind $A, B \in \mathcal{F}$ mit $P(A) > 0$ und $P(B) > 0$, so gilt stets: $$P(A\|B) + P(B\|A) = 1$$	☐	☐
5. Die Anzahl der Fehlversuche, bevor man zum ersten Mal mit einem 6-seitigen fairen Würfel eine 5 oder 6 würfelt, ist Geom $\left(\frac{1}{3}\right)$-verteilt.	☐	☐
6. Die Summe von drei unabhängigen Pois(4)-verteilten Zufallsvariablen ist Pois(12)-verteilt.	☐	☐
7. Der Korrelationskoeffizient $\text{Korr}(X, Y)$ zweier Zufallsvariablen X und Y ist betragsmäßig stets höchstens so groß wie die zugehörige Kovarianz $\text{Cov}(X, Y)$ der beiden Zufallsvariablen.	☐	☐
8. Die Zufallsvariablen X_1 und X_2 seien unkorreliert und für die Standardabweichungen gelte $\text{Sd}(X_1) = 3$ sowie $\text{Sd}(X_2) = 4$. Dann gilt für die Standardabweichung der Summe $X_1 + X_2$ $\text{Sd}(X_1 + X_2) = 7$.	☐	☐

Aufgabe 2 (12 Punkte)

Markieren Sie jeweils die korrekte Antwort mit einem Kreuz im betreffenden Kästchen. Es ist jeweils genau ein Kreuz korrekt.
Richtige Antworten geben 3 Punkte, falsche Antworten und nicht bearbeitete Aufgabenteile 0 Punkte (Aufgabenteile mit mehr als einem Kreuz zählen als nicht bearbeitet!).

1. Kreuzen Sie an, auf welche Merkmalseigenschaften der folgende Box-Plot *am ehesten* hindeutet:



(a) leptokurtisch und rechts**steil** ☐
(b) leptokurtisch und links**steil** ☐
(c) platykurtisch und rechts**steil** ☐
(d) platykurtisch und links**steil** ☐

2. Die Ränge $\text{rg}(X)_1, \ldots, \text{rg}(X)_8$ zur (der Einfachheit halber sortierten) Urliste

$$4, \quad 4, \quad 4, \quad 8, \quad 9, \quad 9, \quad 10, \quad 12$$

des ordinalskalierten Merkmals X lauten:

(a) 2, 2, 2, 4, 5.5, 5.5, 7, 8 ☐
(b) 1, 1, 1, 2, 3, 3, 4, 5 ☐
(c) 1, 2, 3, 4, 5, 6, 7, 8 ☐
(d) 1.5, 1.5, 1.5, 4, 5.5, 5.5, 7, 8 ☐

3. Wenn Sie 3-mal mit Zurücklegen aus einer Urne mit 6 unterscheidbaren Kugeln ziehen, dann beträgt die Wahrscheinlichkeit dafür, dass alle gezogenen Kugeln unterschiedlich sind:

(a) $\frac{3!}{6^3}$ ☐
(b) $\frac{(6)_3}{6^3}$ ☐
(c) $\frac{6!}{3^6}$ ☐
(d) $\frac{3!}{3^6}$ ☐

4. Von den 200 Studierenden, die in einem bestimmten Semester sowohl an der Mathematik- als auch an der Statistikklausur teilgenommen haben, haben 160 Studierende die Statistiklausur bestanden, 140 Studierende haben die Mathematikklausur bestanden und 130 Studierende haben sowohl die Statistik- als auch die Mathematikklausur bestanden. Mit welcher Wahrscheinlichkeit hat ein(e) zufällig ausgewählte(r) Studierende(r) wenigstens eine der beiden Klausuren bestanden?

   (a) 75% ☐

   (b) 80% ☐

   (c) 85% ☐

   (d) 90% ☐

Aufgabe 3 (4 + 1 + 5 + 1 = 11 Punkte)

Zu einem erhobenen Merkmal X sei die folgende empirische Verteilungsfunktion gegeben:

$$F(x) = \begin{cases} 0.00 & \text{für} \quad x < 2 \\ 0.02 & \text{für} \quad 2 \leq x < 4 \\ 0.36 & \text{für} \quad 4 \leq x < 6 \\ 0.64 & \text{für} \quad 6 \leq x < 8 \\ 0.84 & \text{für} \quad 8 \leq x < 10 \\ 0.94 & \text{für} \quad 10 \leq x < 12 \\ 1.00 & \text{für} \quad x \geq 12 \end{cases}$$

Außerdem sei die Länge der Urliste $n = 50$ bekannt.

(a) Erstellen Sie eine Tabelle der absoluten und relativen Häufigkeiten.

(b) Wie groß ist der Anteil der Urlisteneinträge, die Werte von weniger als 5 annehmen?

(c) Berechnen Sie den arithmetischen Mittelwert und die empirische Varianz des Merkmals X.

(d) Bestimmen Sie einen Median des Merkmals X.

Aufgabe 4 (6 + 4 + 3 + 3 + 3 = 19 Punkte)

Zu einem kardinalskalierten Merkmal sei die folgende (zur einfacheren Bearbeitung der Aufgabe bereits sortierte) Urliste der Länge $n = 30$ gegeben:

11.70, 12.50, 13.00, 14.93, 16.42, 17.03, 17.36, 17.68, 17.73, 18.62, 18.75, 19.06, 19.15, 19.21, 19.24, 19.37, 19.71, 19.91, 20.30, 20.49, 20.66, 20.81, 20.84, 20.94, 21.27, 21.40, 21.52, 21.61, 21.81, 21.99

(a) Führen Sie eine Klassierung der erhobenen Daten auf Grundlage der Klassen

$$K_1 = (10, 14],\ K_2 = (14, 18],\ K_3 = (18, 20],\ K_4 = (20, 22]$$

durch. Geben Sie insbesondere die jeweiligen Klassenbreiten, Klassenmitten, absoluten und relativen Klassenhäufigkeiten, Häufigkeitsdichten sowie die Werte der empirischen Verteilungsfunktion an den Klassengrenzen an.

(b) Stellen Sie die (approximative) empirische Verteilungsfunktion des Merkmals aus der Klassierung der Daten auf.

(c) Berechnen Sie aus den klassierten Daten den (approximativen) arithmetischen Mittelwert der Daten. Wie groß ist die relative Abweichung vom tatsächlichen (aus der Urliste bestimmten) Mittelwert von 18.834?

(d) Bestimmen Sie (aus der Urliste) die Anzahl von Merkmalswerten zwischen 12 und 18. Welche Näherung für diese Anzahl können Sie aus der in Teil (b) aufgestellten empirischen Verteilungsfunktion berechnen?

(e) Bestimmen Sie (unter Verwendung der bekannten Konvention zur eindeutigen Festlegung von Quantilen bei kardinalskalierten Merkmalen) *das* obere Quartil sowohl exakt aus der Urliste als auch approximativ mit Hilfe der Verteilungsfunktion für die klassierten Daten.

Aufgabe 5 (4 + 2 + 2 = 8 Punkte)

Zur Unterstützung der Bekämpfung einer Pandemie werden auch Schnelltests eingesetzt, die im Vergleich zu aufwändigeren Testverfahren eine geringere Zuverlässigkeit aufweisen. Gehen Sie im Folgenden davon aus, dass diese Tests eine tatsächlich vorhandene Infektion (bei asymptomatischen Personen) nur mit einer Wahrscheinlichkeit von 80% erkennen, während bei nicht infizierten (asymptomatischen) Personen fälschlicherweise mit einer Wahrscheinlichkeit von 0.15% ein positives Testergebnis entsteht. Gehen Sie außerdem davon aus, dass in einer zum betreffenden Zeitpunkt vergleichsweise günstigen Infektionslage die Wahrscheinlichkeit des tatsächlichen Vorliegens einer Infektion (bei asymptomatischen Personen) nur 0.05% beträgt.

(a) Mit welcher Wahrscheinlichkeit wird ein Schnelltest bei einer (zufällig ausgewählten und asymptomatischen) untersuchten Person ein positives Ergebnis liefern?

(b) Mit welcher Wahrscheinlichkeit ist eine (mit einem Schnelltest) positiv getestete (asymptomatische) Person tatsächlich gar nicht infiziert?

(c) Vor dem Hintergrund des Ergebnisses aus Teil (b) wird der Nutzen der Schnelltests bei der aktuellen Infektionslage in Frage gestellt, indem pauschalisiert ein „hoher Anteil falscher Testergebnisse" beklagt wird. Entkräftigen Sie diese pauschalisierte Aussage durch die Berechnung der Wahrscheinlichkeit dafür, dass eine (asymptomatische) Person in der beschriebenen Infektionslage ein *korrektes* Testergebnis erhält.

Aufgabe 6 (2 + 1 + 2 = 5 Punkte)

Die Lebensdauer bis zum Ausfall einer Server-Festplatte lasse sich als eine exponentialverteilte Zufallsvariable auffassen. Im Mittel vergehen bis zum Ausfall 8 Jahre.

(a) Welche Standardabweichung hat die Lebensdauer bis zum Ausfall?

(b) Mit welcher Wahrscheinlichkeit beträgt die Lebensdauer bis zum Ausfall mehr als 10 Jahre?

(c) Berechnen Sie das 0.80-Quantil der Lebensdauer bis zum Ausfall.

Aufgabe 7 (5 + 2 + 5 + 6 + 3 + 2 = 23 Punkte)

Die Verteilung einer stetigen Zufallsvariablen X sei durch die folgende Dichtefunktion gegeben:

$$f_X(x) = \begin{cases} \frac{1}{6}x + \frac{1}{3} & \text{für } -2 \leq x < 1 \\ -\frac{1}{2}x + 1 & \text{für } 1 \leq x \leq 2 \\ 0 & \text{sonst} \end{cases}$$

(a) Berechnen Sie die Verteilungsfunktion F_X von X.

(b) Berechnen Sie $P\left(\left\{X > \frac{3}{2}\right\}\right)$ und $P\left(\left\{-1 \leq X \leq \frac{3}{2}\right\}\right)$.

(c) Bestimmen Sie den Median und das obere Quartil von X.

(d) Bestimmen Sie den Erwartungswert $\mathrm{E}(X)$.

(e) Bestimmen Sie eine Dichtefunktion von $Y := 2 \cdot X + 4$.

(f) Bestimmen Sie Erwartungswert und Varianz von $Y := 2 \cdot X + 4$.
Hinweis: $\mathrm{Var}(X) = 0.7\overline{2}$

Aufgabe 8 (2 + 3 + 9 + 3 = 17 Punkte)

Gegeben sei die folgende Tabelle der gemeinsamen Wahrscheinlichkeitsverteilung zu einem zweidimensionalen diskreten Zufallsvektor (X, Y):

$X \setminus Y$	-2	0	4	$p_{i\cdot}$
1	0.06	0.06	0.18	
2	0.14	0.28	0.08	
4	0.1	0.06	0.04	
$p_{\cdot j}$				

(a) Ergänzen Sie die obige Tabelle (in den vorgesehenen Feldern) um ihre Randverteilungen.

(b) Geben Sie die bedingten Wahrscheinlichkeitsverteilungen von Y unter der Bedingung $X = x_i$ für alle $x_i \in T(X)$ über die zugehörigen (bedingten) Wahrscheinlichkeitsfunktionen an.

(c) Berechnen Sie $\mathrm{E}(X)$, $\mathrm{E}(Y)$, $\mathrm{Var}(X)$, $\mathrm{Var}(Y)$, $\mathrm{Cov}(X, Y)$ sowie $\mathrm{Korr}(X, Y)$.

(d) Berechnen Sie $\mathrm{E}(-3X + 2Y)$ sowie $\mathrm{Var}(-3X + 2Y)$.

Aufgabe 9 (3 + 2 + 4 = 9 Punkte)

In einem (fiktiven) Impfzentrum stehen täglich 360 Impfdosen zur Verfügung. Auf Basis der Erfahrungen aus der jüngeren Vergangenheit soll davon ausgegangen werden, dass vergebene Impftermine mit einer Wahrscheinlichkeit von 10% (ohne rechtzeitige Absage) nicht in Anspruch genommen werden. Für die individuelle Inanspruchnahme der Impftermine soll dabei Unabhängigkeit unterstellt werden. Um die Verschwendung ungenutzter Impfdosen zu reduzieren, plant das Impfzentrum, täglich zusätzliche Feierabendtermine zu vergeben, zu denen eine Impfung in Aussicht gestellt, aber nicht garantiert wird.

(a) Wie ist die Anzahl der (täglich) nach Ablauf der regulären Impftermine *übrig gebliebenen* Impfdosen Y exakt verteilt? Geben Sie auch den Erwartungswert $\mathrm{E}(Y)$ sowie die Varianz $\mathrm{Var}(Y)$ der Anzahl der übrig gebliebenen Impfdosen an.

(b) Das Impfzentrum entschließt sich dazu, täglich 30 zusätzliche Feierabendtermine zu vergeben. Mit welcher (näherungsweise mit einem geeigneten Grenzwertsatz zu bestimmenden) Wahrscheinlichkeit reichen die übrig gebliebenen Impfdosen aus, um alle 30 Impfwilligen der Feierabendtermine auch tatsächlich impfen zu können?

(c) Verwenden Sie einen geeigneten Grenzwertsatz, um näherungsweise einen um den zugehörigen Erwartungswert symmetrischen Bereich zu bestimmen, in dem sich die Anzahl der (täglich) übrig gebliebenen Impfdosen mit einer Wahrscheinlichkeit von 0.95 realisiert. *(Die Grenzen des Bereichs müssen nicht gerundet werden!)*

Hinweis: Verwenden Sie zur Bearbeitung der Aufgabe die Tabelle zur Standardnormalverteilung auf Seite 120!

15 Deskr. Statistik und Wahrscheinlichkeitsrechnung WS 2021/22

Aufgabe 1 (16 Punkte)

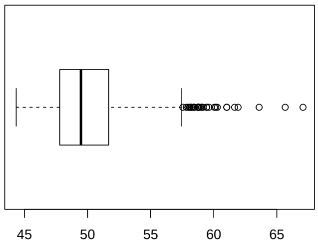
Markieren Sie jeweils mit einem Kreuz pro Aussage im betreffenden Kästchen, ob die unten stehenden Aussagen wahr oder falsch sind.
Richtige Antworten geben 2 Punkte, falsche Antworten und nicht bearbeitete Aussagen 0 Punkte (Aussagen mit zwei Kreuzen zählen als nicht bearbeitet!).

	wahr	falsch
1. Kommt keine andere Ausprägung eines Merkmals X häufiger in der Urliste vor als die Ausprägung a_j, so gehört a_j stets zu den Modalwerten des Merkmals X.	☐	☐
2. Ist (a_1, b_1) eine Ausprägung des zweidimensionalen Merkmals (X, Y) und gilt für die relativen Randhäufigkeiten $r(a_1)$ und $r(b_1)$ sowie die gemeinsame relative Häufigkeit $r(a_1, b_1)$ der Zusammenhang $r(a_1, b_1) = r(a_1) \cdot r(b_1)$, so sind X und Y stets unabhängig.	☐	☐
3. Sei $(\Omega, \mathcal{F}, P)$ ein beliebiger Wahrscheinlichkeitsraum. Dann gilt für zwei Ereignisse $A, B \in \mathcal{F}$ stets $P(A \cup B) \leq P(A) + P(B)$.	☐	☐
4. Es seien $(\Omega, \mathcal{F}, P)$ ein Wahrscheinlichkeitsraum und $A, B, C \in \mathcal{F}$ drei Ereignisse mit positiven Eintrittswahrscheinlichkeiten. Gilt außerdem $P(A) > P(B)$, so gilt auch stets $P(A\|C) > P(B\|C)$.	☐	☐
5. Ist X eine Zufallsvariable mit $\mathrm{Var}(X) = 4$, so gilt für die Zufallsvariable $Y := 2X + 8$ stets $\mathrm{Var}(Y) = 16$.	☐	☐
6. Stetig gleichverteilte Zufallsvariablen sind stets platykurtisch verteilt.	☐	☐
7. Stimmen die Varianzen der beiden Zufallsvariablen X und Y überein, so gilt auch stets $\mathrm{Korr}(X, Y) = \mathrm{Cov}(X, Y)$.	☐	☐
8. Für $i \in \{1, \ldots, 4\}$ seien X_i normalverteilte Zufallsvariablen mit $\mathrm{E}(X_i) = 5$ sowie $\mathrm{Sd}(X_i) = 2$, zudem gelte $\mathrm{Cov}(X_i, X_j) = 0$ für alle $i, j \in \{1, \ldots, 4\}$. Für die Summe $Y := X_1 + X_2 + X_3 + X_4$ gilt dann stets $Y \sim N(20, 4^2)$.	☐	☐

Aufgabe 2 (12 Punkte)

Markieren Sie jeweils die korrekte Antwort mit einem Kreuz im betreffenden Kästchen. Es ist jeweils genau ein Kreuz korrekt.
Richtige Antworten geben 3 Punkte, falsche Antworten und nicht bearbeitete Aufgabenteile 0 Punkte (Aufgabenteile mit mehr als einem Kreuz zählen als nicht bearbeitet!).

1. Kreuzen Sie an, auf welche Merkmalseigenschaften der folgende Box-Plot *am ehesten* hindeutet:



(a) leptokurtisch und rechts**steil** ☐
(b) leptokurtisch und links**steil** ☐
(c) platykurtisch und rechts**steil** ☐
(d) platykurtisch und links**steil** ☐

2. Die Wahrscheinlichkeit, in dieser Klausuraufgabe (4 MC-Aufgabenteile mit jeweils genau einer korrekten Antwort aus 4 Antwortmöglichkeiten) durch *rein zufälliges* Ankreuzen jeweils einer Antwortmöglichkeit (jede Antwortmöglichkeit erhalte also eine Chance von 25%) genau in einem Aufgabenteil die richtige Antwort zu markieren, beträgt (ggf. auf 2 Nachkommastellen gerundet):

(a) 25.00% ☐
(b) 37.25% ☐
(c) 42.19% ☐
(d) 50.00% ☐

3. Bei der Aufstellung eines Listenvorschlags zu einer Gremienwahl sollen die 4 weiblichen und 3 männlichen Kanditat(inn)en so auf die 7 Listenplätze verteilt werden, dass sich die Geschlechter jeweils abwechseln. Damit beträgt die Anzahl der möglichen Listenanordnungen

(a) 12 ☐
(b) 35. ☐
(c) 72. ☐
(d) 144. ☐

4. Für $n \in \mathbb{N}$ seien die Zufallsvariablen $X_1, \ldots, X_n$ unabhängig identisch verteilt mit $\mathrm{E}(X_i) = \mu$ und $\mathrm{Var}(X_i) = \sigma^2$ für alle $i \in \{1, \ldots, n\}$. Dann gilt für $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$

(a) $\mathrm{E}(\overline{X}) = \frac{\mu}{n}$ und $\mathrm{Var}(\overline{X}) = \frac{\sigma^2}{n}$. ☐
(b) $\mathrm{E}(\overline{X}) = \mu$ und $\mathrm{Var}(\overline{X}) = \frac{\sigma^2}{n}$. ☐
(c) $\mathrm{E}(\overline{X}) = \frac{\mu}{n}$ und $\mathrm{Var}(\overline{X}) = \sigma^2$. ☐
(d) $\mathrm{E}(\overline{X}) = \mu$ und $\mathrm{Var}(\overline{X}) = \sigma^2$. ☐

Aufgabe 3 (3 + 3 + 1 + 5 + 1 = 13 Punkte)

Bei einer Umfrage wurden 50 Personen befragt, wie viele Streaming-Dienste sie abonniert haben (Merkmal X). Das Ergebnis der Umfrage ist die folgende (bereits aufsteigend sortierte) Urliste zu X:

0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 3, 3, 3, 4, 4, 4, 4

(a) Erstellen Sie eine Tabelle mit den absoluten und relativen Häufigkeiten.

(b) Stellen Sie die zugehörige empirische Verteilungsfunktion auf.

(c) Wie groß ist der Anteil der Personen in der Umfrage, die mindestens 2 Streaming-Dienste abonniert haben?

(d) Berechnen Sie den arithmetischen Mittelwert und die empirische Varianz des Merkmals X.

(e) Bestimmen Sie ein oberes Quartil des Merkmals X.

Aufgabe 4 (6 + 4 + 3 + 3 + 3 = 19 Punkte)

Zu einem kardinalskalierten Merkmal sei die folgende (zur einfacheren Bearbeitung der Aufgabe bereits sortierte) Urliste der Länge $n = 40$ gegeben:

6.15, 7.03, 7.90, 8.05, 8.10, 9.37, 9.58, 9.94, 10.45, 11.33, 11.75, 13.96, 14.57, 15.02, 15.26, 16.45, 16.45, 16.46, 17.04, 19.62, 21.55, 21.84, 22.45, 23.00, 24.18, 27.70, 31.02, 34.87, 36.34, 36.43, 37.34, 37.36, 37.87, 39.58, 41.31, 45.79, 46.95, 48.50, 53.22, 56.93

(a) Führen Sie eine Klassierung der erhobenen Daten auf Grundlage der Klassen

$$K_1 = (5, 15],\ K_2 = (15, 25],\ K_3 = (25, 45],\ K_4 = (45, 65]$$

durch. Geben Sie insbesondere die jeweiligen Klassenbreiten, Klassenmitten, absoluten und relativen Klassenhäufigkeiten, Häufigkeitsdichten sowie die Werte der empirischen Verteilungsfunktion an den Klassengrenzen an.

(b) Stellen Sie die (approximative) empirische Verteilungsfunktion des Merkmals aus der Klassierung der Daten auf.

(c) Berechnen Sie aus den klassierten Daten den (approximativen) arithmetischen Mittelwert der Daten. Wie groß ist die relative Abweichung vom tatsächlichen (aus der Urliste bestimmten) Mittelwert von 24.218?

(d) Bestimmen Sie (aus der Urliste) die Anzahl von Merkmalswerten zwischen 25 und 40. Welche Näherung für diese Anzahl können Sie aus der in Teil (b) aufgestellten empirischen Verteilungsfunktion berechnen?

(e) Bestimmen Sie (unter Verwendung der bekannten Konvention zur eindeutigen Festlegung von Quantilen bei kardinalskalierten Merkmalen) *das* obere Quartil sowohl exakt aus der Urliste als auch approximativ mit Hilfe der Verteilungsfunktion für die klassierten Daten.

Aufgabe 5 (4 + 2 = 6 Punkte)

Von den Unternehmen, die sich auf einer 2-tägigen Messe präsentieren, sind 10% nur am 1. Tag sowie 15% nur am 2. Tag vertreten (die restlichen 75% sind an beiden Tagen vertreten).

(a) Wie wahrscheinlich ist es, dass ein (zufällig ausgewähltes) Unternehmen, das am 1. Tag vertreten war, auch am 2. Tag vertreten ist?

(b) Wie wahrscheinlich ist es, dass ein (zufällig ausgewähltes) Unternehmen, das am 2. Tag vertreten ist, nicht schon am 1. Tag vertreten war?

Aufgabe 6 (5 + 2 = 7 Punkte)

Bei einem Chip-Hersteller verteilt sich die Produktion eines bestimmten Mikrocontrollers auf insgesamt drei verschiedene Produktionslinien A, B und C. Dabei werden im Mittel 25% der Chips auf Linie A, 60% der Chips auf Linie B und 15% der Chips auf Linie C hergestellt. Aus den Ergebnissen der Qualitätssicherung ist bekannt, dass 98% der Chips aus Linie A, 99% der Chips aus Linie B und 97% der Chips aus Linie C nicht fehlerhaft sind.

(a) Wie groß ist die Wahrscheinlichkeit, dass ein zufällig ausgewählter Chip fehlerhaft ist?

(b) Wie groß ist die Wahrscheinlichkeit, dass ein fehlerhafter Chip auf der Linie A produziert wurde?

Aufgabe 7 (3 + 2 + 6 + 1 + 5 + 3 = 20 Punkte)

Die Verteilung einer stetigen Zufallsvariablen X sei durch die folgende Verteilungsfunktion gegeben:

$$F_X(x) = \begin{cases} 0 & \text{für } x \leq 0 \\ \frac{1}{8}x^2 + \frac{1}{8}x & \text{für } 0 < x \leq 1 \\ -\frac{1}{4}x^2 + \frac{3}{2}x - 1 & \text{für } 1 < x \leq 2 \\ 1 & \text{für } x > 2 \end{cases}$$

(a) Berechnen Sie eine Dichtefunktion f_X von X.

(b) Berechnen Sie $P\left(\left\{X > \frac{1}{2}\right\}\right)$ und $P\left(\left\{\frac{1}{2} < X \leq \frac{3}{2}\right\}\right)$.

(c) Bestimmen Sie den Erwartungswert $\mathrm{E}(X)$.

(d) Ist X symmetrisch um ihren Erwartungswert verteilt (ohne Begründung)?

(e) Bestimmen Sie das untere Quartil und den Median von X.

(f) Bestimmen Sie die Verteilungsfunktion von $Y := 2 \cdot X - 3$.

Aufgabe 8 (2 + 3 + 9 + 1 + 3 = 18 Punkte)

Gegeben sei die folgende Tabelle der gemeinsamen Wahrscheinlichkeitsverteilung zu einem zweidimensionalen diskreten Zufallsvektor (X, Y):

$X \setminus Y$	1	2	3	$p_{i\cdot}$
-4	0.1	0.14	0.06	
0	0.08	0.08	0.04	
2	0.12	0.18	0.2	
$p_{\cdot j}$				

(a) Ergänzen Sie die obige Tabelle (in den vorgesehenen Feldern) um ihre Randverteilungen.

(b) Geben Sie die bedingten Wahrscheinlichkeitsverteilungen von Y unter der Bedingung $X = x_i$ für alle $x_i \in T(X)$ über die zugehörigen (bedingten) Wahrscheinlichkeitsfunktionen an.

(c) Berechnen Sie $\mathrm{E}(X)$, $\mathrm{E}(Y)$, $\mathrm{Var}(X)$, $\mathrm{Var}(Y)$, $\mathrm{Cov}(X, Y)$ sowie $\mathrm{Korr}(X, Y)$.

(d) Sind X und Y stochastisch unabhängig?

(e) Berechnen Sie $\mathrm{E}(5X - 2Y)$ sowie $\mathrm{Var}(5X - 2Y)$.

Aufgabe 9 (2 + 4 + 3 = 9 Punkte)

Ein Holzschnitzer stellt Weihnachtsfiguren als Unikate in Handarbeit her. Aus langjähriger Erfahrung weiß er, dass er pro Figur im Mittel 2 Stunden bei einer Standardabweichung von 0.3 Stunden benötigt. Man kann davon ausgehen, dass die einzelnen Herstellungszeiten nicht gegenseitig voneinander abhängen. In einem bestimmten Jahr nimmt der Schnitzer vor Weihnachten Bestellungen über insgesamt 64 Figuren an.

(a) Geben Sie die den Erwartungswert sowie die Standardabweichung der gesamten Herstellungszeit (für alle 64 Figuren) an.

(b) Verwenden Sie den zentralen Grenzwertsatz, um näherungsweise zu bestimmen, mit welcher Wahrscheinlichkeit der Schnitzer insgesamt zwischen 124 und 134 Stunden zur Herstellung aller Figuren benötigt.

(c) Verwenden Sie den zentralen Grenzwertsatz, um näherungsweise ein 0.95-Quantil für die gesamte Herstellungszeit (für alle 64 Figuren) zu bestimmen.

Hinweis: Verwenden Sie zur Bearbeitung von Aufgabenteil (b) und (c) die Tabelle zur Standardnormalverteilung auf Seite 120!

16 Deskr. Statistik und Wahrscheinlichkeitsrechnung SS 2022

Aufgabe 1 (18 Punkte)

Markieren Sie jeweils mit einem Kreuz pro Aussage im betreffenden Kästchen, ob die unten stehenden Aussagen wahr oder falsch sind.
Richtige Antworten geben 2 Punkte, falsche Antworten und nicht bearbeitete Aussagen 0 Punkte (Aussagen mit zwei Kreuzen zählen als nicht bearbeitet!).

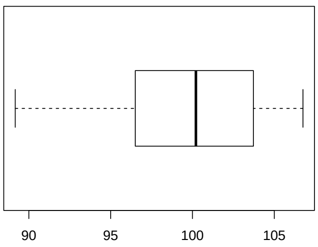
	wahr	falsch
1. Ein Merkmal mit der Urliste $$1, 2, 3, 3, 4, 5, 5, 6, 7$$ hat einen einzigen bzw. eindeutig bestimmten Modalwert.	☐	☐
2. Ist (a_1, b_1) eine Ausprägung des zweidimensionalen Merkmals (X, Y) und gilt für die relativen Randhäufigkeiten $r(a_1)$ und $r(b_1)$ sowie die gemeinsame relative Häufigkeit $r(a_1, b_1)$ der Zusammenhang $r(a_1, b_1) \neq r(a_1) \cdot r(b_1)$, so sind X und Y stets abhängig.	☐	☐
3. Sei $(\Omega, \mathcal{F}, P)$ ein beliebiger Wahrscheinlichkeitsraum. Dann gilt für zwei Ereignisse $A, B \in \mathcal{F}$ stets $2 \cdot P(A \cap B) \leq P(A) + P(B)$.	☐	☐
4. Ist $(\Omega, \mathcal{F}, P)$ ein beliebiger Wahrscheinlichkeitsraum und sind $A, B \in \mathcal{F}$ Ereignisse mit $P(A) > 0$ und $P(B) > 0$, so gilt stets: $$P(A\|B) \cdot P(B) = P(B\|A) \cdot P(A)$$	☐	☐
5. In einer bestimmten Kalenderwoche waren 72% aller hospitalisierten Covid-Patienten vollständig geimpft. Hieraus kann man schließen, dass in dieser Kalenderwoche der Anteil der hospitalisierten Covid-Patienten unter den vollständig geimpften Personen größer war als unter den nicht vollständig geimpften.	☐	☐
6. Die stetige Zufallsvariable X mit der Dichtefunktion $$f_X(x) = \begin{cases} \lvert x \rvert & \text{für } -1 \leq x \leq 1 \\ 0 & \text{sonst} \end{cases}$$ ist symmetrisch um ihren Erwartungswert 0 verteilt.	☐	☐
7. Ist X eine Zufallsvariable mit $\text{Var}(X) = 3$, so gilt für die Zufallsvariable $Y := -2X + 12$ stets $\text{Var}(Y) = 6$.	☐	☐
8. Für zwei Zufallsvariablen X und Y über einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ mit $\text{Var}(X) = 2$ und $\text{Var}(Y) = 8$ gilt stets $\lvert \text{Cov}(X, Y) \rvert \leq 4$.	☐	☐
9. Die Zufallsvariablen X und Y seien stochastisch unabhängig, ferner gelte $X \sim N(10, 2^2)$, $Y \sim N(20, 3^2)$. Dann ist die Summe $S := X + Y$ normalverteilt mit $\text{E}(S) = 30$ und $\text{Var}(S) = 5^2$.	☐	☐

<u>**Aufgabe 2**</u> (15 Punkte)

Markieren Sie jeweils die korrekte Antwort mit einem Kreuz im betreffenden Kästchen. Es ist jeweils genau ein Kreuz korrekt.

Richtige Antworten geben 3 Punkte, falsche Antworten und nicht bearbeitete Aufgabenteile 0 Punkte (Aufgabenteile mit mehr als einem Kreuz zählen als nicht bearbeitet!).

1. Kreuzen Sie an, auf welche Merkmalseigenschaften der folgende Box-Plot *am ehesten* hindeutet:



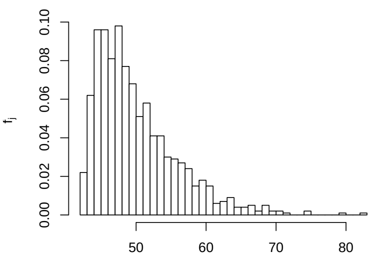
(a) leptokurtisch und rechts**steil** ☐

(b) leptokurtisch und links**steil** ☐

(c) platykurtisch und rechts**steil** ☐

(d) platykurtisch und links**steil** ☐

2. Kreuzen Sie an, auf welche Merkmalseigenschaften das folgende Histogramm *am ehesten* hindeutet:



(a) leptokurtisch und rechts**steil** ☐

(b) leptokurtisch und links**steil** ☐

(c) platykurtisch und rechts**steil** ☐

(d) platykurtisch und links**steil** ☐

3. Die Ränge $\text{rg}(X)_1, \ldots, \text{rg}(X)_8$ zur (der Einfachheit halber sortierten) Urliste

sehr gut, sehr gut, gut, gut, gut, befriedigend, ausreichend, ausreichend

des ordinalskalierten Merkmals X lauten:

(a) 1, 1, 3, 3, 3, 5, 7, 8 ☐

(b) 1, 2, 3, 4, 5, 6, 7, 8 ☐

(c) 1.5, 1.5, 3.5, 3.5, 3.5, 6.5, 7.5, 7.5 ☐

(d) 1.5, 1.5, 4, 4, 4, 6, 7.5, 7.5 ☐

4. Eine Big Band hat insgesamt 25 Musiktitel in ihrem Repertoire. In einem Konzert sollen 15 dieser 25 Titel zu Gehör gebracht werden, wobei kein Titel mehrfach gespielt werden soll. Wie viele Möglichkeiten hat der Bandleader zur Zusammenstellung des Konzertprogramms, wenn auch die Reihenfolge der Titel von Bedeutung ist?

   (a) 25^{15} Möglichkeiten ☐
   (b) 15^{25} Möglichkeiten ☐
   (c) $(25)_{15} = \frac{25!}{10!}$ Möglichkeiten ☐
   (d) $\binom{25}{15} = \frac{25!}{15! \cdot 10!}$ Möglichkeiten ☐

5. Die Wahrscheinlichkeit, in dieser Klausuraufgabe (5 MC-Aufgabenteile mit jeweils genau einer korrekten Antwort aus 4 Antwortmöglichkeiten) durch *rein zufälliges* Ankreuzen jeweils einer Antwortmöglichkeit (jede Antwortmöglichkeit erhalte also eine Chance von 25%) keine einzige richtige Antwort zu markieren, beträgt (ggf. auf 2 Nachkommastellen gerundet):

   (a) 20.00% ☐
   (b) 23.73% ☐
   (c) 40.96% ☐
   (d) 59.04% ☐

Aufgabe 3 (3 + 3 + 1 + 5 + 1 = 13 Punkte)

Bei einer Umfrage wurden 40 Haushalte befragt, wie viele Fernseher sie in den vergangenen 10 Jahren angeschafft haben (Merkmal X). Das Ergebnis der Umfrage ist die folgende (bereits aufsteigend sortierte) Urliste zu X:

0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 2, 2,
2, 2, 2, 2, 2, 2, 3, 3, 4

(a) Erstellen Sie eine Tabelle mit den absoluten und relativen Häufigkeiten.

(b) Stellen Sie die zugehörige empirische Verteilungsfunktion auf.

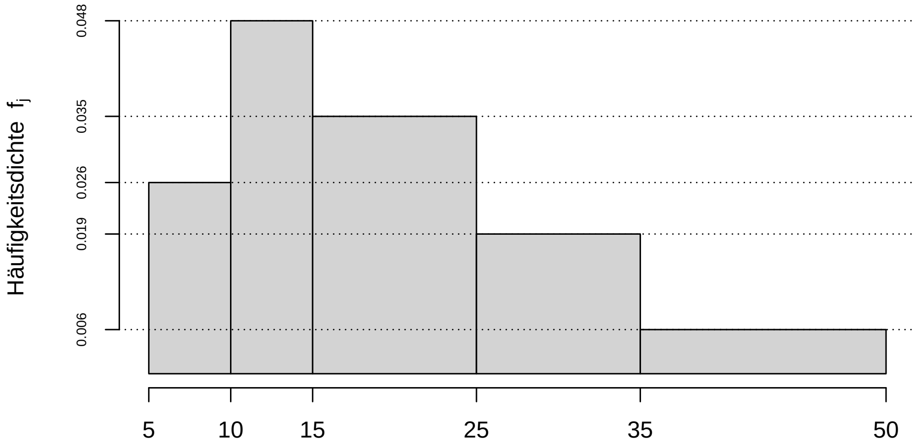
(c) Wie groß ist der Anteil der Haushalte in der Umfrage, die höchstens 2 Fernseher in den vergangenen 10 Jahren angeschafft haben?

(d) Berechnen Sie den arithmetischen Mittelwert und die empirische Varianz des Merkmals X.

(e) Bestimmen Sie einen Median des Merkmals X.

Aufgabe 4 (7 + 4 + 3 + 2 + 2 = 18 Punkte)

Gegeben sei das folgende Histogramm zur Klassierung einer Urliste der Länge $n = 100$:



(a) Rekonstruieren Sie die Klassierung der Daten aus dem Histogramm. Geben Sie insbesondere die jeweiligen Klassenbreiten, Klassenmitten, absoluten und relativen Klassenhäufigkeiten, Häufigkeitsdichten sowie die Werte der empirischen Verteilungsfunktion an den Klassengrenzen an.

(b) Stellen Sie die (approximative) empirische Verteilungsfunktion des Merkmals aus der Klassierung der Daten auf.

(c) Berechnen Sie aus den klassierten Daten den (approximativen) arithmetischen Mittelwert der Daten. Wie groß ist die relative Abweichung vom tatsächlichen (aus der Urliste bestimmten) Mittelwert von 20.092?

(d) Welche Näherung für die Anzahl von Merkmalswerten zwischen 10 und 40 können Sie unter Verwendung der approximativen Verteilungsfunktion für die klassierten Daten berechnen?

(e) Bestimmen Sie näherungsweise unter Verwendung der approximativen Verteilungsfunktion für die klassierten Daten das untere Quartil.

Aufgabe 5 (2 + 2 + 2 = 6 Punkte)

Eine Lostrommel enthält 10 (gleichartige) Kugeln, die von 1 bis 10 durchnummeriert sind. Es wird einmalig rein zufällig eine der 10 Kugeln gezogen.

(a) Geben Sie einen geeigneten Wahrscheinlichkeitsraum zur Beschreibung dieses Zufallsexperiments an.

(b) Wie groß ist die Wahrscheinlichkeit, eine Kugel mit einer geraden Zahl zu ziehen? Wie groß ist die Wahrscheinlichkeit, eine Kugel mit einer Zahl kleiner oder gleich 5 zu ziehen?

(c) Sind die Ereignisse „Kugel mit einer geraden Zahl" und „Kugel mit einer Zahl kleiner oder gleich 5" aus Teil (b) stochastisch unabhängig? Begründen Sie Ihre Antwort!

Aufgabe 6 (5 + 2 = 7 Punkte)

Ein Sanitärinstallationsbetrieb verwendet für die Ausführung von Warmwasser-Installationen drei unterschiedlichen Systeme A, B und C von verschiedenen Herstellern. Dabei werden Warmwasser-Installationen mit einer Wahrscheinlichkeit von 40% mit System A, 35% mit System B und 25% mit System C ausgeführt. Bei den anschließenden Druckprüfungen gibt es erfahrungsgemäß mit einer Wahrscheinlichkeit von 2% bei System A, 2.5% bei System B und 3% bei System C Undichtigkeiten.

(a) Wie groß ist die Wahrscheinlichkeit, dass eine zufällig ausgewählte Warmwasser-Installation bei der Druckprüfung nicht undicht ist?

(b) Wie groß ist die Wahrscheinlichkeit, dass eine bei der Druckprüfung nicht undichte Installation mit System C ausgeführt wurde?

Aufgabe 7 (3 + 2 + 5 + 6 + 2 = 18 Punkte)

Die Verteilung einer stetigen Zufallsvariablen X sei durch die folgende Verteilungsfunktion gegeben:

$$F_X(x) = \begin{cases} 0 & \text{für } x \leq -2 \\ -\frac{1}{12}x^2 + \frac{1}{6}x + \frac{2}{3} & \text{für } -2 < x \leq 1 \\ \frac{1}{4}x^2 - \frac{1}{2}x + 1 & \text{für } 1 < x \leq 2 \\ 1 & \text{für } x > 2 \end{cases}$$

(a) Berechnen Sie eine Dichtefunktion f_X von X.

(b) Berechnen Sie $P(\{-1 \leq X \leq 0\})$ und $P(\{0 < X < 1\})$.

(c) Bestimmen Sie den Median und das obere Quartil von X.

(d) Bestimmen Sie den Erwartungswert $\mathrm{E}(X)$.

(e) Bestimmen Sie Erwartungswert und Varianz von $Y := -2 \cdot X + 3$.
Hinweis: $\mathrm{Var}(X) = 1.7\overline{2}$

Aufgabe 8 (2 + 8 + 1 + 2 + 3 = 16 Punkte)

Ein fairer (sechsseitiger) Würfel wird zweimal geworfen. Es seien X die Anzahl der Würfe mit einer Augenzahl ≤ 3 sowie Y die Anzahl der Würfe mit einer ungeraden Augenzahl.

(a) Welcher Verteilung genügen X und Y (jeweils)?

(b) Die gemeinsame Verteilung von (X, Y) ist gegeben durch:

$X \setminus Y$	0	1	2
0	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{36}$
1	$\frac{1}{9}$	$\frac{5}{18}$	$\frac{1}{9}$
2	$\frac{1}{36}$	$\frac{1}{9}$	$\frac{1}{9}$

Bestimmen Sie $\mathrm{E}(X)$, $\mathrm{E}(Y)$, $\mathrm{Var}(X)$, $\mathrm{Var}(Y)$, $\mathrm{Cov}(X,Y)$ sowie $\mathrm{Korr}(X,Y)$.

Hinweis: Beachten Sie, dass Sie die diesen Aufgabenteil unter Verwendung der Ergebnisse aus Teil (a) zum Teil recht schnell und insbesondere vollständig ohne die Bestimmung der Randwahrscheinlichkeiten von X und Y lösen können!

(c) Mit welcher Wahrscheinlichkeit nehmen sowohl X als auch Y Werte von mindestens 1 an?

(d) Sind X und Y stochastisch unabhängig? Begründen Sie Ihre Antwort.

(e) Berechnen Sie $\mathrm{E}\,(2X - 4Y)$ sowie $\mathrm{Var}\,(2X - 4Y)$.

Aufgabe 9 (2 + 4 + 3 = 9 Punkte)

Die Zufallsvariablen $X_1, \ldots, X_{48}$ seien unabhängig identisch Pois(3)-verteilt. Die Summe der Zufallsvariablen X_i sei mit

$$Y := \sum_{i=1}^{48} X_i = X_1 + \ldots + X_{48}$$

bezeichnet.

(a) Geben Sie die (exakte) Verteilung von Y sowie deren Erwartungswert $\mathrm{E}(Y)$ und Varianz $\mathrm{Var}(Y)$ an.

(b) Verwenden Sie den zentralen Grenzwertsatz, um näherungsweise zu bestimmen, mit welcher Wahrscheinlichkeit Y Werte zwischen 140 und 150 annimmt.

(c) Verwenden Sie den zentralen Grenzwertsatz, um näherungsweise ein 0.975-Quantil von Y zu bestimmen.

Hinweis: Verwenden Sie zur Bearbeitung von Aufgabenteil (b) und (c) die Tabelle zur Standardnormalverteilung auf Seite 120!

17 Deskr. Statistik und Wahrscheinlichkeitsrechnung WS 2022/23

Aufgabe 1 (18 Punkte)

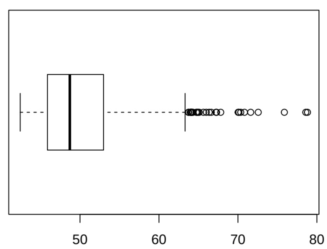
Markieren Sie jeweils mit einem Kreuz pro Aussage im betreffenden Kästchen, ob die unten stehenden Aussagen wahr oder falsch sind.
Richtige Antworten geben 2 Punkte, falsche Antworten und nicht bearbeitete Aussagen 0 Punkte (Aussagen mit zwei Kreuzen zählen als nicht bearbeitet!).

		wahr	falsch
1.	Empirische Verteilungsfunktionen F können streng monoton wachsend auf $\mathbb{R}$ sein.	☐	☐
2.	Die jährlichen Inflationsraten in Deutschland betrugen in den Jahren 2019–2022 im Einzelnen 1.4%, 0.5%, 3.1% und 6.9%. Damit beträgt die durchschnittliche jährliche Inflationsrate in diesem Zeitraum (gerundet) 2.975%.	☐	☐
3.	Sind in einer Klausur mehr weibliche als männliche Prüfungsteilnehmende durchgefallen und haben mehr männliche als weibliche Prüflinge an der Klausur teilgenommen, so kann man daraus stets schließen, dass die Durchfallquote unter den weiblichen Prüflingen größer ist als die unter den männlichen.	☐	☐
4.	Sind $(\Omega, \mathcal{F}, P)$ ein Wahrscheinlichkeitsraum und $A, B \in \mathcal{F}$ zwei Ereignisse mit $P(A \cap B) = P(A) + P(B)$, dann muss stets sowohl $P(A) = 0$ als auch $P(B) = 0$ gelten.	☐	☐
5.	Die Wahrscheinlichkeit, durch rein zufällige Anordnung der Buchstaben E, E, K, L, S, S das Wort KESSEL zu erhalten, beträgt (gerundet) 0.556%.	☐	☐
6.	Exponentialverteilte Zufallsvariablen sind nie symmetrisch.	☐	☐
7.	Wird bei allen 9 Aufgabenteilen dieser Aufgabe jeweils rein zufällig (entweder) “wahr” oder “falsch” angekreuzt, so beträgt der Erwartungswert für die erreichte Punktzahl 9.	☐	☐
8.	Ist die Kovarianz zweier Zufallsvariablen X und Y (mit existierenden Varianzen) positiv, so ist die Varianz der Differenz von X und Y kleiner als die Summe der einzelnen Varianzen (von X und Y).	☐	☐
9.	Sind $X_1, \ldots, X_n$ stochastisch unabhängige Zufallsvariablen mit $X_i \sim B(n_i, p_i)$, $i \in \{1, \ldots, n\}$, dann genügt $Y = \sum_{i=1}^n X_i$ einer $B(n, p)$-Verteilung mit $n = \sum_{i=1}^n n_i$ und $p = \sum_{i=1}^n p_i$.	☐	☐

Aufgabe 2 (15 Punkte)

Markieren Sie jeweils die korrekte Antwort mit einem Kreuz im betreffenden Kästchen. Es ist jeweils genau ein Kreuz korrekt.
Richtige Antworten geben 3 Punkte, falsche Antworten und nicht bearbeitete Aufgabenteile 0 Punkte (Aufgabenteile mit mehr als einem Kreuz zählen als nicht bearbeitet!).

1. Kreuzen Sie an, auf welche Merkmalseigenschaften der folgende Box-Plot *am ehesten* hindeutet:



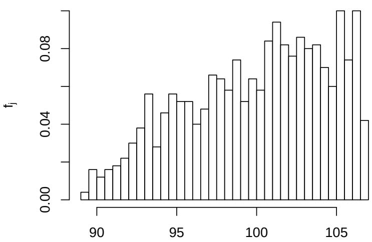
(a) leptokurtisch und rechts**steil** ☐

(b) leptokurtisch und links**steil** ☐

(c) platykurtisch und rechts**steil** ☐

(d) platykurtisch und links**steil** ☐

2. Kreuzen Sie an, auf welche Merkmalseigenschaften das folgende Histogramm *am ehesten* hindeutet:



(a) leptokurtisch und rechts**steil** ☐

(b) leptokurtisch und links**steil** ☐

(c) platykurtisch und rechts**steil** ☐

(d) platykurtisch und links**steil** ☐

3. Von den beiden Ereignissen A und B eines Wahrscheinlichkeitsraums $(\Omega, \mathcal{F}, P)$ sei bekannt, dass

- mit einer Wahrscheinlichkeit von $\frac{1}{3}$ Ereignis A, aber nicht Ereignis B eintritt,
- mit einer Wahrscheinlichkeit von $\frac{1}{4}$ Ereignis B, aber nicht Ereignis A eintritt,
- (mindestens) eines der beiden Ereignisse stets eintritt.

Damit

(a) tritt das Ereignis B mit der Wahrscheinlichkeit $\frac{1}{3}$ ein. ☐

(b) tritt das Ereignis B mit der Wahrscheinlichkeit $\frac{2}{3}$ ein. ☐

(c) tritt das Ereignis B mit der Wahrscheinlichkeit $\frac{3}{4}$ ein. ☐

(d) kann die Wahrscheinlichkeit von B nicht bestimmt werden. ☐

4. Während einer Pressekonferenz sollen die 8 Vorstandsmitglieder (4 Frauen und 4 Männer) eines Unternehmens nebeneinander an einem langen Tisch (mit 8 Stühlen) sitzen. Wie viele Möglichkeiten zur Anordnung der 8 Vorstandsmitglieder gibt es hierbei, wenn sich die Geschlechter jeweils abwechseln sollen?

   (a) 1152 Möglichkeiten ☐
   (b) 576 Möglichkeiten ☐
   (c) 48 Möglichkeiten ☐
   (d) 24 Möglichkeiten ☐

5. Es seien X und Y zwei Zufallsvariablen über dem Wahrscheinlichkeitraum $(\Omega, \mathcal{F}, P)$ mit existierenden positiven Varianzen $\text{Var}(X)$ und $\text{Var}(Y)$ sowie existierender Kovarianz $\text{Cov}(X, Y)$. Dann gilt für die Unkorreliertheit von X und Y sowie die stochastische Unabhängigkeit von X und Y:

   (a) Es gibt keinen Zusammenhang zwischen Unkorreliertheit und Unabhängigkeit. ☐
   (b) Aus der Unkorreliertheit folgt die stochastische Unabhängigkeit. ☐
   (c) Aus der stochastischen Unabhängigkeit folgt die Unkorreliertheit. ☐
   (d) Unkorreliertheit und stochastische Unabhängigkeit sind äquivalent. ☐

Aufgabe 3 (4 + 1 + 5 + 1 = 11 Punkte)

Zu einem erhobenen Merkmal X sei die folgende empirische Verteilungsfunktion gegeben:

$$F(x) = \begin{cases} 0.000 & \text{für} \quad x < 3 \\ 0.225 & \text{für} \quad 3 \leq x < 4 \\ 0.575 & \text{für} \quad 4 \leq x < 5 \\ 0.800 & \text{für} \quad 5 \leq x < 6 \\ 0.950 & \text{für} \quad 6 \leq x < 7 \\ 1.000 & \text{für} \quad x \geq 7 \end{cases}$$

Außerdem sei die Länge der Urliste $n = 40$ bekannt.

(a) Erstellen Sie eine Tabelle der absoluten und relativen Häufigkeiten.

(b) Wie groß ist der Anteil der Urlisteneinträge, die Werte von mehr als 4 annehmen?

(c) Berechnen Sie den arithmetischen Mittelwert und die empirische Varianz des Merkmals X.

(d) Bestimmen Sie ein unteres Quartil des Merkmals X.

Aufgabe 4 (6 + 4 + 3 + 3 + 3 = 19 Punkte)

Zu einem kardinalskalierten Merkmal sei die folgende (zur einfacheren Bearbeitung der Aufgabe bereits sortierte) Urliste der Länge $n = 50$ gegeben:

10.62, 11.15, 11.26, 11.65, 11.73, 14.40, 14.40, 15.21, 17.37, 18.89, 19.50, 19.63, 19.77, 20.34, 21.32, 22.21, 22.43, 23.42, 23.61, 24.10, 25.84, 26.27, 26.44, 26.53, 26.67, 27.12, 27.41, 28.04, 28.18, 28.26, 30.54, 30.77, 31.11, 31.26, 32.01, 32.49, 35.01, 35.24, 35.98, 36.05, 36.15, 36.47, 36.93, 37.06, 37.32, 37.46, 38.49, 39.06, 39.70, 39.87

(a) Führen Sie eine Klassierung der erhobenen Daten auf Grundlage der Klassen

$$K_1 = (0, 15],\ K_2 = (15, 25],\ K_3 = (25, 35],\ K_4 = (35, 40]$$

durch. Geben Sie insbesondere die jeweiligen Klassenbreiten, Klassenmitten, absoluten und relativen Klassenhäufigkeiten, Häufigkeitsdichten sowie die Werte der empirischen Verteilungsfunktion an den Klassengrenzen an.

(b) Stellen Sie die (approximative) empirische Verteilungsfunktion des Merkmals aus der Klassierung der Daten auf.

(c) Berechnen Sie aus den klassierten Daten den (approximativen) arithmetischen Mittelwert der Daten. Wie groß ist die relative Abweichung vom tatsächlichen (aus der Urliste bestimmten) Mittelwert von 26.655?

(d) Bestimmen Sie (aus der Urliste) die Anzahl von Merkmalswerten zwischen 25 und 30. Welche Näherung für diese Anzahl können Sie aus der in Teil (b) aufgestellten empirischen Verteilungsfunktion berechnen?

(e) Bestimmen Sie (unter Verwendung der bekannten Konvention zur eindeutigen Festlegung von Quantilen bei kardinalskalierten Merkmalen) *den* Median sowohl exakt aus der Urliste als auch approximativ mit Hilfe der Verteilungsfunktion für die klassierten Daten.

Aufgabe 5 (6 + 2 + 1 = 9 Punkte)

In einer Fußballmannschaft sind die vier Mitspielerinnen Andrea, Beatrice, Christina und Dana für die Ausführung von Eckstößen zuständig. Dabei werden Eckstöße mit einer Wahrscheinlichkeit von 40% von Andrea, mit einer Wahrscheinlichkeit von 30% von Beatrice, mit einer Wahrscheinlichkeit von 20% von Christina und mit einer Wahrscheinlichkeit von 10% von Dana ausgeführt. Aus einer ausführlichen statistischen Auswertung ist bekannt, dass Eckstöße von Andrea mit einer Wahrscheinlichkeit von 13%, Eckstöße von Beatrice mit einer Wahrscheinlichkeit von 10%, Eckstöße von Christina mit einer Wahrscheinlichkeit von 7% und Eckstöße von Dana mit einer Wahrscheinlichkeit von 4% zu einem Tor führen.

(a) Wie groß ist die Wahrscheinlichkeit, dass ein zufällig ausgewählter Eckstoß zu einem Tor führt?

(b) Wie groß ist die Wahrscheinlichkeit, dass ein Eckstoß, der nicht zu einem Tor geführt hat, von Beatrice ausgeführt wurde?

(c) Sind die Ereignisse „Eckstoß führt zu einem Tor“ und „Eckstoß wird von Beatrice ausgeführt“ stochastisch unabhängig?

Aufgabe 6 (5 + 2 + 6 + 4 = 17 Punkte)

Die Verteilung einer stetigen Zufallsvariablen X sei durch die folgende Dichtefunktion gegeben:

$$f_X(x) = \begin{cases} \frac{1}{4}x + \frac{1}{4} & \text{für } -1 \leq x < 1 \\ -x + 2 & \text{für } 1 \leq x \leq 2 \\ 0 & \text{sonst} \end{cases}$$

(a) Berechnen Sie die Verteilungsfunktion F_X von X.

(b) Berechnen Sie $P\left(\left\{X < -\frac{1}{2}\right\}\right)$ und $P\left(\left\{-\frac{1}{2} \leq X \leq \frac{1}{2}\right\}\right)$.

(c) Bestimmen Sie den Erwartungswert $\mathrm{E}(X)$.

(d) Bestimmen Sie das obere Quartil von X.

Aufgabe 7 (3 + 3 = 6 Punkte)

Die Anzahl der Netzwerkausfälle pro Monat in einem bestimmten Rechenzentrum lasse sich als eine Pois(0.1)-verteilte Zufallsvariable auffassen. Außerdem soll angenommen werden, dass die Anzahl der Netzwerkausfälle pro Monat in dem betreffenden Rechenzentrum für unterschiedliche Monate stochastisch unabhängig ist.

(a) Mit welcher Wahrscheinlichkeit ereignet sich in einem Monat in dem betreffenden Rechenzentrum höchstens 1 Netzwerkausfall?

(b) Welche Verteilung hat die Anzahl der Netzwerkausfälle pro Jahr in dem betreffenden Rechenzentrum? Mit welcher Wahrscheinlichkeit ereignet sich in einem Jahreszeitraum in dem betreffenden Rechenzentrum mindestens 1 Netzwerkausfall?

Aufgabe 8 (2 + 3 + 9 + 3 = 17 Punkte)

Gegeben sei die folgende Tabelle der gemeinsamen Wahrscheinlichkeitsverteilung zu einem zweidimensionalen diskreten Zufallsvektor (X, Y):

$X \setminus Y$	2	3	4	$p_{i\cdot}$
-1	$\frac{3}{40}$	$\frac{1}{8}$	$\frac{1}{20}$	
1	$\frac{3}{20}$	$\frac{1}{5}$	$\frac{1}{40}$	
2	$\frac{1}{40}$	$\frac{3}{10}$	$\frac{1}{20}$	
$p_{\cdot j}$				

(a) Ergänzen Sie die obige Tabelle (in den vorgesehenen Feldern) um ihre Randverteilungen.

(b) Geben Sie die bedingten Wahrscheinlichkeitsverteilungen von X unter der Bedingung $Y = y_j$ für alle $y_j \in T(Y)$ über die zugehörigen (bedingten) Wahrscheinlichkeitsfunktionen an.

(c) Berechnen Sie $\mathrm{E}(X)$, $\mathrm{E}(Y)$, $\mathrm{Var}(X)$, $\mathrm{Var}(Y)$, $\mathrm{Cov}(X, Y)$ sowie $\mathrm{Korr}(X, Y)$.

(d) Berechnen Sie $\mathrm{E}\left(-2X + 4Y\right)$ sowie $\mathrm{Var}\left(-2X + 4Y\right)$.

Aufgabe 9 (2 + 2 + 4 = 8 Punkte)

Ein Online-Händler bietet für 225 der an einem Tag eingehenden Bestellungen einen Express-Lieferservice an, der eine Abfertigung der Bestellung am nächsten Arbeitstag garantiert. Es ist davon auszugehen, dass die Zeitdauern zur Abfertigung einzelner Express-Bestellungen (in Stunden) unabhängig identisch verteilt sind mit einer mittleren Abfertigungsdauer von 0.16 Stunden bei einer Standardabweichung von 0.05 Stunden.

(a) Welchen Erwartungswert und welche Standardabweichung hat die Summe der Abfertigungsdauern von 225 Express-Bestellungen?

(b) Verwenden Sie den zentralen Grenzwertsatz, um die Wahrscheinlichkeit, dass 225 Express-Bestellungen in höchstens 37.5 (Mitarbeiter-)Stunden abgefertigt werden können, (näherungsweise) zu berechnen.

(c) Verwenden Sie den zentralen Grenzwertsatz, um näherungsweise einen um den zugehörigen Erwartungswert symmetrischen Bereich zu bestimmen, in dem sich die Gesamtabfertigungsdauer von 225 Express-Bestellungen mit einer Wahrscheinlichkeit von 0.95 realisiert.

Hinweis: Verwenden Sie zur Bearbeitung der Aufgabenteile (b) und (c) die Tabelle zur Standardnormalverteilung auf Seite 120!

18 Deskr. Statistik und Wahrscheinlichkeitsrechnung SS 2023

Aufgabe 1 (18 Punkte)

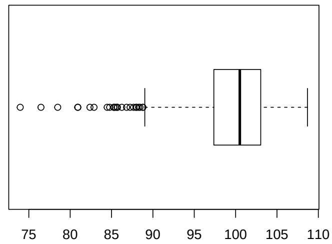
Markieren Sie jeweils mit einem Kreuz pro Aussage im betreffenden Kästchen, ob die unten stehenden Aussagen wahr oder falsch sind.
Richtige Antworten geben 2 Punkte, falsche Antworten und nicht bearbeitete Aussagen 0 Punkte (Aussagen mit zwei Kreuzen zählen als nicht bearbeitet!).

	wahr	falsch
1. Das stetige Merkmal mit der (sortierten) Urliste $$1.2, 2.4, 3.6, 3.9, 4.3, 5.1, 5.5, 6.8$$ hat $x_{0.5} = 3.9$ als einzigen Median.	☐	☐
2. Mit dem (korrigierten) Pearsonschen Kontingenzkoeffizienten kann die Richtung der Abhängigkeit zwischen zwei Merkmalen gemessen werden.	☐	☐
3. Es seien A und B zwei Ereignisse in einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ mit $P(A) = 0.3$, $P(B) = 0.2$ und $P(A \cup B) = 0.4$. Damit gilt $P(A \cap B) = 0.1$.	☐	☐
4. Die Wahrscheinlichkeit, beim (gleichzeitigen) Werfen von fünf fairen (sechsseitigen) Würfeln ausschließlich gerade Punktzahlen zu erhalten, beträgt 6.25%.	☐	☐
5. Wenn man bei der Vorbereitung auf eine Prüfung zwei der möglichen 13 Themengebiete ausspart, so hat man für die Auswahl der ausgesparten Themengebiete 78 Möglichkeiten.	☐	☐
6. Die Anzahl der Trägerpunkte diskreter Zufallsvariablen ist stets endlich.	☐	☐
7. Geometrisch verteilte Zufallsvariablen können symmetrisch (um ihren Erwartungswert) sein.	☐	☐
8. Für zwei stochastisch unabhängige Zufallsvariablen X und Y gilt (bei Existenz der entsprechenden Momente) stets $\mathrm{E}(X \cdot Y) = \mathrm{E}(X) \cdot \mathrm{E}(Y)$.	☐	☐
9. Seien X_1, X_2 und X_3 stochastisch unabhängige Zufallsvariablen mit $\mathrm{Var}(X_1) = 40$, $\mathrm{Var}(X_2) = 30$, sowie $\mathrm{Var}(X_3) = 20$. Dann gilt $\mathrm{Var}(X_1 + X_2 - X_3) = 50$.	☐	☐

Aufgabe 2 (15 Punkte)

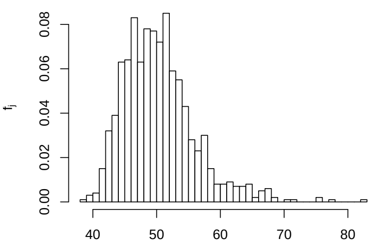
Markieren Sie jeweils die korrekte Antwort mit einem Kreuz im betreffenden Kästchen. Es ist jeweils genau ein Kreuz korrekt.
Richtige Antworten geben 3 Punkte, falsche Antworten und nicht bearbeitete Aufgabenteile 0 Punkte (Aufgabenteile mit mehr als einem Kreuz zählen als nicht bearbeitet!).

1. Kreuzen Sie an, auf welche Merkmalseigenschaften der folgende Box-Plot *am ehesten* hindeutet:



(a) leptokurtisch und rechts**steil** ☐
(b) leptokurtisch und links**steil** ☐
(c) platykurtisch und rechts**steil** ☐
(d) platykurtisch und links**steil** ☐

2. Kreuzen Sie an, auf welche Merkmalseigenschaften das folgende Histogramm *am ehesten* hindeutet:



(a) leptokurtisch und rechts**steil** ☐
(b) leptokurtisch und links**steil** ☐
(c) platykurtisch und rechts**steil** ☐
(d) platykurtisch und links**steil** ☐

3. Von den Personen, die an einer 2-tägigen Veranstaltung teilgenommen haben, waren 80% am ersten und 60% am zweiten Veranstaltungstag anwesend. Wie groß ist die (gegebenenfalls gerundete) Wahrscheinlichkeit, dass eine zufällig ausgewählte am zweiten Tag anwesende Person auch schon am ersten Tag anwesend war?

(a) 40% ☐
(b) 60% ☐
(c) 66.67% ☐
(d) 100% ☐

4. Vor einem Spiel einer Fußballweltmeisterschaft laufen gemeinsam mit den 11 Spielern einer Fußballmannschaft auch 11 sogenannte Einlaufkinder auf das Spielfeld, so dass beim Abspielen der Nationalhymne vor jedem der 11 Spieler jeweils ein Einlaufkind steht. Damit gibt es für die Zuordnung, welches der 11 Kinder vor welchem der 11 Nationalspieler dieser Mannschaft steht, insgesamt

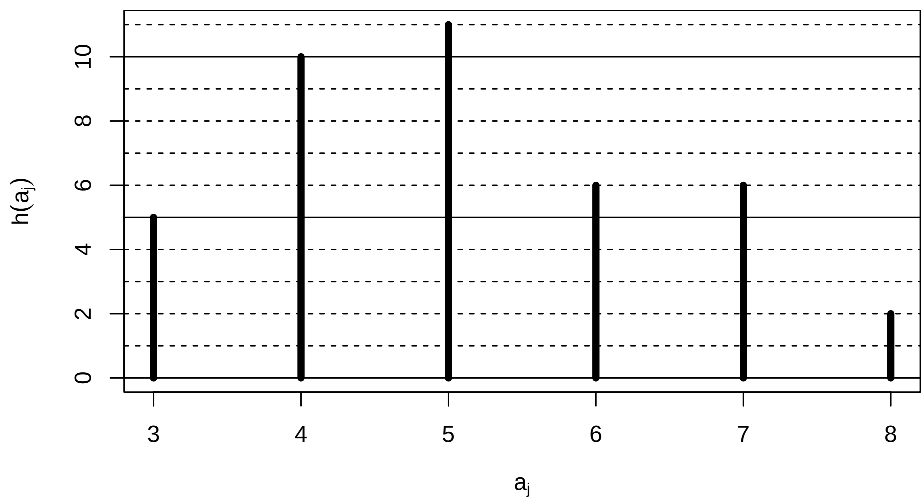
   (a) $22!$ Möglichkeiten. ☐
   (b) $11!$ Möglichkeiten. ☐
   (c) $(22)_{11}$ Möglichkeiten. ☐
   (d) $\binom{22}{11}$ Möglichkeiten. ☐

5. Sind X und Y zwei stochastisch unabhängige Zufallsvariablen mit $X \sim B(50, 0.1)$ und $Y \sim B(50, 0.1)$, dann ist die Verteilung von $X + Y$ eine

   (a) $B(50, 0.2)$-Verteilung. ☐
   (b) $B(50, 0.1)$-Verteilung. ☐
   (c) $B(100, 0.2)$-Verteilung. ☐
   (d) $B(100, 0.1)$-Verteilung. ☐

Aufgabe 3 (2 + 4 + 5 + 3 + 1 + 1 = 16 Punkte)

Zu einem erhobenen Merkmal X sei das folgende Stabdiagramm gegeben:



(a) Geben Sie die Menge A der Merkmalsausprägungen und die Länge n der Urliste an.

(b) Erstellen Sie eine Tabelle der absoluten und relativen Häufigkeiten.

(c) Berechnen Sie den arithmetischen Mittelwert und die empirische Standardabweichung des Merkmals X.

(d) Geben Sie die empirische Verteilungsfunktion des Merkmals X an.

(e) Wie groß ist der Anteil der Urlisteneinträge, die Werte von höchstens 5 annehmen?

(f) Berechnen Sie ein oberes Quartil des Merkmals X.

Aufgabe 4 (6 + 4 + 3 + 3 + 3 = 19 Punkte)

Zu einem kardinalskalierten Merkmal sei die folgende (zur einfacheren Bearbeitung der Aufgabe bereits sortierte) Urliste der Länge $n = 40$ gegeben:

10.30, 14.45, 15.05, 15.46, 16.36, 16.60, 17.12, 17.76, 20.39, 20.68, 21.95, 22.35, 22.47, 22.60, 23.86, 24.55, 24.83, 25.29, 25.46, 25.67, 28.52, 28.72, 28.77, 29.38, 29.57, 29.87, 30.04, 30.19, 30.26, 30.32, 30.46, 30.85, 31.04, 32.00, 32.44, 32.77, 33.21, 33.22, 33.89, 34.89

(a) Führen Sie eine Klassierung der erhobenen Daten auf Grundlage der Klassen

$$K_1 = (5, 15],\ K_2 = (15, 25],\ K_3 = (25, 30],\ K_4 = (30, 35]$$

durch. Geben Sie insbesondere die jeweiligen Klassenbreiten, Klassenmitten, absoluten und relativen Klassenhäufigkeiten, Häufigkeitsdichten sowie die Werte der empirischen Verteilungsfunktion an den Klassengrenzen an.

(b) Stellen Sie die (approximative) empirische Verteilungsfunktion des Merkmals aus der Klassierung der Daten auf.

(c) Berechnen Sie aus den klassierten Daten den (approximativen) arithmetischen Mittelwert der Daten. Wie groß ist die relative Abweichung vom tatsächlichen (aus der Urliste bestimmten) Mittelwert von 25.59?

(d) Bestimmen Sie (aus der Urliste) die Anzahl von Merkmalswerten zwischen 10 und 25. Welche Näherung für diese Anzahl können Sie aus der in Teil (b) aufgestellten empirischen Verteilungsfunktion berechnen?

(e) Bestimmen Sie (unter Verwendung der bekannten Konvention zur eindeutigen Festlegung von Quantilen bei kardinalskalierten Merkmalen) *den* Median sowohl exakt aus der Urliste als auch approximativ mit Hilfe der Verteilungsfunktion für die klassierten Daten.

Aufgabe 5 (5 + 2 + 2 = 9 Punkte)

Ein Versandhaus beauftragt für den Versand seiner Sendungen einen von insgesamt drei verschiedenen Versand-Dienstleistern A, B und C. Dabei werden durchschnittlich 45% der Sendungen an Dienstleister A, 25% der Sendungen an Dienstleister B und 30% der Sendungen an Dienstleister C übergeben. Die umfangreiche Auswertung der vorhandenen Kunden-Feedbacks zu Qualität und Geschwindigkeit der Lieferung ergab, dass 4% der Lieferungen mit Dienstleister A, 5% der Lieferungen mit Dienstleister B und 3% der Lieferungen mit Dienstleister C beanstandet werden.

(a) Wie groß ist die Wahrscheinlichkeit, dass eine zufällig ausgewählte Sendung keinen Anlass zur Beanstandung durch den Kunden bietet?

(b) Wie groß ist die Wahrscheinlichkeit, dass eine beanstandete Lieferung mit Dienstleister C versendet wurde?

(c) Sind die Ereignisse „Lieferung wird beanstandet“ und „Dienstleister C war mit dem Versand beauftragt“ stochastisch unabhängig? Begründen Sie Ihre Antwort!

Aufgabe 6 (5 + 2 + 6 + 4 = 17 Punkte)

Die Verteilung einer stetigen Zufallsvariablen X sei durch die folgende Dichtefunktion gegeben:

$$f_X(x) = \begin{cases} \frac{1}{20}x + \frac{1}{4} & \text{für } -2 \leq x < 0 \\ -\frac{2}{15}x + \frac{2}{5} & \text{für } 0 \leq x \leq 3 \\ 0 & \text{sonst} \end{cases}$$

(a) Berechnen Sie die Verteilungsfunktion F_X von X.

(b) Berechnen Sie $P(\{-1 \leq X \leq 0\})$ und $P(\{0 < X < 1\})$.

(c) Bestimmen Sie den Erwartungswert $\mathrm{E}(X)$.

(d) Bestimmen Sie den Median von X.

Aufgabe 7 (2 + 3 = 5 Punkte)

Als Hausaufgabe im Fach Politik waren die Geburtsdaten von 30 bekannten Politikern auswendig zu lernen. Der Schüler Samuel Schlauberger hat 27 dieser Geburtsdaten auswendig gelernt (die Chance, bei den anderen 3 Geburtsdaten durch Raten eine richtige Antwort zu geben, sei gleich Null). Der Lehrer überprüft, ob Samuel die Hausaufgabe ordentlich erledigt hat, indem er 5 Mal rein zufällig und unabhängig voneinander einen der Politiker auswählt und die zugehörigen Geburtsdaten abfragt. Kann Samuel mindestens zu 4 dieser 5 Politiker die Geburtsdaten korrekt angeben, so ist die Überprüfung bestanden.

(a) Welche Verteilung besitzt die Anzahl der von Samuel abgegebenen richtigen Antworten?

(b) Mit welcher Wahrscheinlichkeit besteht Samuel die Überprüfung?

Aufgabe 8 (2 + 3 + 5 + 3 = 13 Punkte)

Gegeben sei die folgende Tabelle der gemeinsamen Wahrscheinlichkeitsverteilung zu einem zweidimensionalen diskreten Zufallsvektor (X, Y):

$X \setminus Y$	-1	0	1	$p_{i\cdot}$
1	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{16}$	
2	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{4}$	
4	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{16}$	
$p_{\cdot j}$				

(a) Ergänzen Sie die obige Tabelle (in den vorgesehenen Feldern) um ihre Randverteilungen.

(b) Geben Sie die bedingten Wahrscheinlichkeitsverteilungen von X unter der Bedingung $Y = y_j$ für alle $y_j \in T(Y)$ über die zugehörigen (bedingten) Wahrscheinlichkeitsfunktionen an.

(c) Berechnen Sie mit Hilfe der Zwischenergebnisse

$$\mathrm{E}(X) = \frac{9}{4}, \mathrm{E}(Y) = \frac{1}{8}, \mathrm{E}(X^2) = \frac{25}{4} \text{ und } \mathrm{E}(Y^2) = \frac{5}{8}$$

die Größen $\mathrm{Var}(X)$, $\mathrm{Var}(Y)$, $\mathrm{Cov}(X, Y)$ sowie $\mathrm{Korr}(X, Y)$.

(d) Berechnen Sie $\mathrm{E}\left(2X - 4Y\right)$ sowie $\mathrm{Var}\left(2X - 4Y\right)$.

Aufgabe 9 $\quad$ ($2 + 2 + 4 = 8$ Punkte)

Ein freiberuflicher Netzwerktechniker benötigt für das Auflegen eines Netzwerkanschlusses im Mittel 11 Minuten bei einer Standardabweichung von 3 Minuten. Man kann davon ausgehen, dass die benötigten Zeitdauern für die einzelnen Anschlüsse nicht gegenseitig voneinander abhängen. In einem bestimmten Schaltschrank sind insgesamt 64 Anschlüsse aufzulegen.

(a) Geben Sie die den Erwartungswert sowie die Standardabweichung der gesamten Arbeitszeit (für alle 64 Netzwerkanschlüsse) an.

(b) Verwenden Sie den zentralen Grenzwertsatz, um näherungsweise zu bestimmen, mit welcher Wahrscheinlichkeit der Netzwerktechniker nicht länger als 12 Stunden bzw. 720 Minuten zum Auflegen aller Anschlüsse benötigt.

(c) Verwenden Sie den zentralen Grenzwertsatz, um näherungsweise einen um den zugehörigen Erwartungswert symmetrischen Bereich zu bestimmen, in dem sich die gesamte Arbeitszeit für 64 Netzwerkanschlüsse mit einer Wahrscheinlichkeit von 0.95 realisiert.

Hinweis: Verwenden Sie zur Bearbeitung von Aufgabenteil (b) und (c) die Tabelle zur Standardnormalverteilung auf Seite 120!

19 Deskr. Statistik und Wahrscheinlichkeitsrechnung WS 2023/24

Aufgabe 1 (18 Punkte)

Markieren Sie jeweils mit einem Kreuz pro Aussage im betreffenden Kästchen, ob die unten stehenden Aussagen wahr oder falsch sind.
Richtige Antworten geben 2 Punkte, falsche Antworten und nicht bearbeitete Aussagen 0 Punkte (Aussagen mit zwei Kreuzen zählen als nicht bearbeitet!).

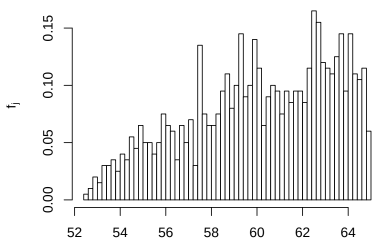
	wahr	falsch
1. Beim Modus (oder Modalwert) eines kardinalskalierten Merkmals X handelt es sich um die größte angenommene Merkmalsausprägung von X.	☐	☐
2. Am Vorzeichen des Bravais-Pearsonschen Korrelationskoeffizienten zweier Merkmale X und Y kann man erkennen, ob X auf Y wirkt oder umgekehrt (Y auf X).	☐	☐
3. Es seien A und B zwei Ereignisse in einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ mit $P(B \backslash A) = 0.3$ und $P(A \backslash B) = 0.4$. Dann gilt stets $P(A \cup B) \geq 0.7$.	☐	☐
4. Die Wahrscheinlichkeit, dass sich bei rein zufälliger Anordnung der Ziffern 1, 1, 1, 3, 3 und 7 die Zahl 317131 ergibt, beträgt (ggf. gerundet) 1.667%.	☐	☐
5. Es seien A, B und C drei Ereignisse in einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$ mit $P(C) > 0$. Gilt sowohl $P(A\|C) < P(A)$ als auch $P(B\|C) < P(B)$, dann gilt stets auch $P(A \cap B\|C) < P(A \cap B)$.	☐	☐
6. Verteilungsfunktionen stetiger Zufallsvariablen sind stets streng monoton wachsend auf $\mathbb{R}$.	☐	☐
7. Binomialverteilte Zufallsvariablen sind nie symmetrisch (um ihren Erwartungswert) verteilt.	☐	☐
8. Die Funktion $p_X : \mathbb{R} \to \mathbb{R}$ mit $p_X(x) = \begin{cases} \frac{1}{10} & \text{für } x \in \{11, 12, \ldots, 20\} \\ 0 & \text{sonst} \end{cases}$ erfüllt alle Voraussetzungen für Wahrscheinlichkeitsfunktionen diskreter Zufallsvariablen.	☐	☐
9. Existieren alle beteiligten Momente, so gilt: Ist die Kovarianz zweier Zufallsvariablen X und Y negativ, dann ist die Varianz der Differenz von X und Y größer als die Summe der einzelnen Varianzen (von X und Y).	☐	☐

Aufgabe 2 (15 Punkte)

Markieren Sie jeweils die korrekte Antwort mit einem Kreuz im betreffenden Kästchen. Es ist jeweils genau ein Kreuz korrekt.

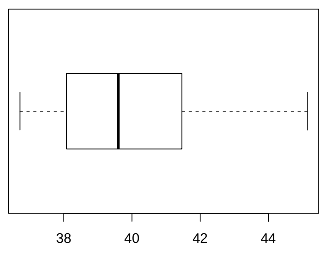
Richtige Antworten geben 3 Punkte, falsche Antworten und nicht bearbeitete Aufgabenteile 0 Punkte (Aufgabenteile mit mehr als einem Kreuz zählen als nicht bearbeitet!).

1. Kreuzen Sie an, auf welche Merkmalseigenschaften das folgende Histogramm *am ehesten* hindeutet:



(a) leptokurtisch und rechts**steil** ☐
(b) leptokurtisch und links**steil** ☐
(c) platykurtisch und rechts**steil** ☐
(d) platykurtisch und links**steil** ☐

2. Kreuzen Sie an, auf welche Merkmalseigenschaften der folgende Box-Plot *am ehesten* hindeutet:



(a) leptokurtisch und rechts**steil** ☐
(b) leptokurtisch und links**steil** ☐
(c) platykurtisch und rechts**steil** ☐
(d) platykurtisch und links**steil** ☐

3. Die Ränge $\mathrm{rg}(X)_1, \ldots, \mathrm{rg}(X)_8$ zur (der Einfachheit halber sortierten) Urliste

sehr gut, sehr gut, sehr gut, gut, gut, befriedigend, befriedigend, ausreichend

des ordinalskalierten Merkmals X lauten:

(a) 1, 1, 1, 4, 4, 6, 6, 8 ☐
(b) 1, 1, 1, 2, 2, 4, 3, 4 ☐
(c) 2, 2, 2, 4.5, 4.5, 6.5, 6.5, 8 ☐
(d) 1.5, 1.5, 1.5, 4.5, 4.5, 6.5, 6.5, 8 ☐

4. In einer Eisdiele werden insgesamt 13 verschiedene Eissorten als Kugeln für Eiswaffeln angeboten. Wie viele Möglichkeiten haben Sie zur Zusammenstellung einer Eiswaffel mit 3 Kugeln, wenn Sie nicht mehrere Kugeln derselben Sorte in der Waffel haben möchten und die Reihenfolge, in der die Kugeln in der Waffel angeordnet sind, egal ist?

   (a) 3^{13} Möglichkeiten. ☐

   (b) 13^3 Möglichkeiten. ☐

   (c) $(13)_3 = \frac{13!}{10!}$ Möglichkeiten. ☐

   (d) $\binom{13}{3} = \frac{13!}{3! \cdot 10!}$ Möglichkeiten. ☐

5. Sind X_1, X_2 und X_3 drei stochastisch unabhängige Zufallsvariablen mit $X_1 \sim N(30, 6^2)$, $X_2 \sim N(70, 12^2)$ und $X_3 \sim N(50, 4^2)$, dann ist die Verteilung von $X_1 + X_2 + X_3$ eine

   (a) $N(50, 22^2)$-Verteilung. ☐

   (b) $N(50, 14^2)$-Verteilung. ☐

   (c) $N(150, 22^2)$-Verteilung. ☐

   (d) $N(150, 14^2)$-Verteilung. ☐

Aufgabe 3 (3 + 3 + 1 + 5 + 1 = 13 Punkte)

Bei einer Umfrage wurden 50 Personen befragt, wie viele Tagesgeldkonten sie in den vergangenen 12 Monaten neu eröffnet haben (Merkmal X). Das Ergebnis der Umfrage ist die folgende (bereits aufsteigend sortierte) Urliste zu X:

0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,
1, 1, 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 3

(a) Erstellen Sie eine Tabelle mit den absoluten und relativen Häufigkeiten.

(b) Stellen Sie die zugehörige empirische Verteilungsfunktion auf.

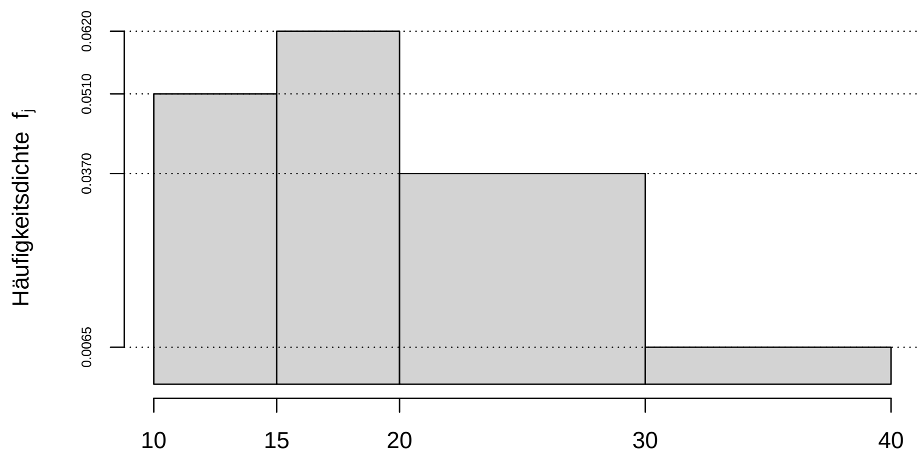
(c) Wie groß ist der Anteil der Personen in der Umfrage, die mindestens 1 Tagesgeldkonto in den vergangenen 12 Monaten neu eröffnet haben?

(d) Berechnen Sie den arithmetischen Mittelwert und die empirische Varianz des Merkmals X.

(e) Bestimmen Sie ein oberes Quartil des Merkmals X.

Aufgabe 4 (6 + 4 + 3 + 2 + 2 = 17 Punkte)

Gegeben sei das folgende Histogramm zur Klassierung einer Urliste der Länge $n = 200$:



(a) Rekonstruieren Sie die Klassierung der Daten aus dem Histogramm. Geben Sie insbesondere die jeweiligen Klassenbreiten, Klassenmitten, absoluten und relativen Klassenhäufigkeiten, Häufigkeitsdichten sowie die Werte der empirischen Verteilungsfunktion an den Klassengrenzen an.

(b) Stellen Sie die (approximative) empirische Verteilungsfunktion des Merkmals aus der Klassierung der Daten auf.

(c) Berechnen Sie aus den klassierten Daten den (approximativen) arithmetischen Mittelwert der Daten. Wie groß ist die relative Abweichung vom tatsächlichen (aus der Urliste bestimmten) Mittelwert von 19.895?

(d) Welche Näherung für die Anzahl von Merkmalswerten zwischen 20 und 35 können Sie unter Verwendung der approximativen Verteilungsfunktion für die klassierten Daten berechnen?

(e) Bestimmen Sie näherungsweise unter Verwendung der approximativen Verteilungsfunktion für die klassierten Daten das obere Quartil.

Aufgabe 5 (1 + 2 + 2 + 2 = 7 Punkte)

Eine Urne enthält 128 gleichartige Kugeln, von denen 8 blau und gestreift, 32 rot und gestreift, 64 blau und gepunktet sowie 24 rot und gepunktet sind. Wie groß ist die Wahrscheinlichkeit dafür, dass

(a) eine rein zufällig aus der Urne entnommene Kugel rot und gestreift ist?

(b) eine rein zufällig aus der Urne entnommene Kugel blau ist?

(c) eine rein zufällig aus der Urne entnommene Kugel gepunktet ist, wenn man weiß, dass sie rot ist?

(d) bei dreimaligem rein zufälligen Ziehen *mit Zurücklegen* die erste Kugel blau, die zweite Kugel rot und die letzte Kugel gestreift ist?

Aufgabe 6 $(4 + 2 + 2 = 8$ Punkte)

An einer seltenen Krankheit seien 1.5% der Bevölkerung einer bestimmten Altersgruppe erkrankt. Zum Einsatz in flächendeckenden Früherkennungsuntersuchungen existiere ein medizinisches Diagnoseverfahren, welches erkrankte Personen mit einer Wahrscheinlichkeit von 96% (korrekterweise) auch als krank einstuft, bei gesunden (bzw. nicht an dieser Krankheit erkrankten) Personen allerdings mit einer Wahrscheinlichkeit von 3% (fälschlicherweise) ebenfalls eine entsprechende Erkrankung diagnostiziert.

(a) Mit welcher Wahrscheinlichkeit wird ein Patient in der betrachteten Altersgruppe im Rahmen einer Früherkennungsuntersuchung als krank eingestuft?

(b) Mit welcher Wahrscheinlichkeit wird sich eine positive Diagnose bei einer Früherkennungsuntersuchung in der betrachteten Altersgruppe als falsch herausstellen?

(c) Wie groß ist die Wahrscheinlichkeit für eine richtige Diagnose bei einer Früherkennungsuntersuchung in der betrachteten Altersgruppe?

Aufgabe 7 $(5 + 1 + 6 + 4 = 16$ Punkte)

Die Verteilung einer stetigen Zufallsvariablen X sei durch die folgende Dichtefunktion gegeben:

$$f_X(x) = \begin{cases} -\frac{1}{4}x + \frac{1}{2} & \text{für } 0 \leq x < 2 \\ \frac{1}{6}x - \frac{1}{4} & \text{für } 2 \leq x \leq 4 \\ 0 & \text{sonst} \end{cases}$$

(a) Berechnen Sie die Verteilungsfunktion F_X von X.

(b) Berechnen Sie $P(\{X < 1\})$.

(c) Bestimmen Sie den Erwartungswert $\mathrm{E}(X)$.

(d) Bestimmen Sie das untere Quartil von X.

Aufgabe 8 $(2 + 3 + 9 + 3 = 17$ Punkte)

Gegeben sei die folgende Tabelle der gemeinsamen Wahrscheinlichkeitsverteilung zu einem zweidimensionalen diskreten Zufallsvektor (X, Y):

$X \setminus Y$	-2	0	2	$p_{i\cdot}$
1	$\frac{1}{40}$	$\frac{3}{20}$	$\frac{3}{40}$	
2	$\frac{3}{40}$	$\frac{1}{5}$	$\frac{1}{10}$	
4	$\frac{1}{10}$	$\frac{1}{5}$	$\frac{3}{40}$	
$p_{\cdot j}$				

(a) Ergänzen Sie die obige Tabelle (in den vorgesehenen Feldern) um ihre Randverteilungen.

(b) Geben Sie die bedingten Wahrscheinlichkeitsverteilungen von Y unter der Bedingung $X = x_i$ für alle $x_i \in T(X)$ über die zugehörigen (bedingten) Wahrscheinlichkeitsfunktionen an.

(c) Berechnen Sie $\mathrm{E}(X)$, $\mathrm{E}(Y)$, $\mathrm{Var}(X)$, $\mathrm{Var}(Y)$, $\mathrm{Cov}(X,Y)$ sowie $\mathrm{Korr}(X,Y)$.

(d) Berechnen Sie $\mathrm{E}\,(4X - 2Y)$ sowie $\mathrm{Var}\,(4X - 2Y)$.

Aufgabe 9 (2 + 4 + 3 = 9 Punkte)

Die Zufallsvariablen $X_1, \ldots, X_{625}$ seien unabhängig identisch $B(1, 0.8)$-verteilt. Die Summe der Zufallsvariablen X_i sei mit

$$Y := \sum_{i=1}^{625} X_i = X_1 + \ldots + X_{625}$$

bezeichnet.

(a) Geben Sie die (exakte) Verteilung von Y sowie deren Erwartungswert $\mathrm{E}(Y)$ und Varianz $\mathrm{Var}(Y)$ an.

(b) Verwenden Sie den zentralen Grenzwertsatz, um näherungsweise zu bestimmen, mit welcher Wahrscheinlichkeit Y Werte zwischen 485 und 508 annimmt.

(c) Verwenden Sie den zentralen Grenzwertsatz, um näherungsweise ein 0.90-Quantil von Y zu bestimmen.

Hinweis: Verwenden Sie zur Bearbeitung von Aufgabenteil (b) und (c) die Tabelle zur Standardnormalverteilung auf Seite 120!

20 Deskr. Statistik und Wahrscheinlichkeitsrechnung SS 2024

Aufgabe 1 (18 Punkte)

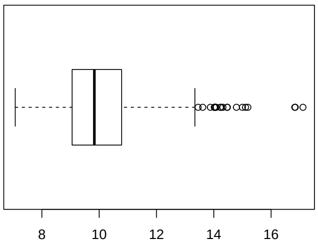
Markieren Sie jeweils mit einem Kreuz pro Aussage im betreffenden Kästchen, ob die unten stehenden Aussagen wahr oder falsch sind.
Richtige Antworten geben 2 Punkte, falsche Antworten und nicht bearbeitete Aussagen 0 Punkte (Aussagen mit zwei Kreuzen zählen als nicht bearbeitet!).

	wahr	falsch
1. Die Summe der Flächeninhalte aller Rechtecke eines Histogramms beträgt stets 1.	☐	☐
2. Wenn Sie mit einem Elektrofahrzeug 300 km bei einem durchschnittlichen Verbrauch von 20 kWh pro 100 km und anschließend 200 km bei einem durchschnittlichen Verbrauch von 15 kWh pro 100 km zurücklegen, so beträgt der Gesamtdurchschnitt des Verbrauchs 18 kWh pro 100 km.	☐	☐
3. Modalwerte eines Merkmals X können nur dann bestimmt werden, wenn das Merkmal X ordinal- oder kardinalskaliert ist.	☐	☐
4. Wenn die Aufstellung einer 4×400-m-Staffel aus der Auswahl von 4 aus 6 Kadersportlerinnen sowie der Festlegung der Reihenfolge für die ausgewählten 4 Sportlerinnen besteht, so gibt es für die Aufstellung insgesamt 360 Möglichkeiten.	☐	☐
5. In einem Laplaceschen Wahrscheinlichkeitsraum ist die Anzahl der Elemente in Ω stets endlich.	☐	☐
6. Seien $(\Omega, \mathcal{F}, P)$ ein beliebiger Wahrscheinlichkeitsraum und $A, B \in \mathcal{F}$ mit $0 < P(A) < 1$. Dann gilt stets: $$P(B\|A) + P(B\|\overline{A}) = 1$$	☐	☐
7. Binomialverteilte Zufallsvariablen sind stets symmetrisch (um ihren Erwartungswert) verteilt.	☐	☐
8. Gilt $\mathrm{Cov}(X, Y) = 0$ für zwei Zufallsvariablen X und Y, so gilt stets auch $\mathrm{E}(X \cdot Y) = \mathrm{E}(X) \cdot \mathrm{E}(Y)$.	☐	☐
9. Liegen alle Trägerpunkte eines zweidimensionalen diskreten Zufallsvektors (X, Y) auf einer Geraden mit der Steigung 0.3, so gilt $\mathrm{Korr}(X, Y) = 0.3$.	☐	☐

Aufgabe 2 (15 Punkte)

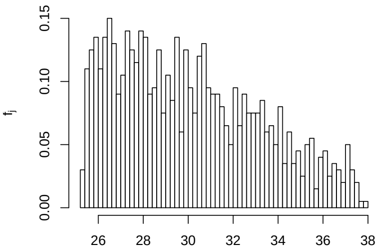
Markieren Sie jeweils die korrekte Antwort mit einem Kreuz im betreffenden Kästchen. Es ist jeweils genau ein Kreuz korrekt.
Richtige Antworten geben 3 Punkte, falsche Antworten und nicht bearbeitete Aufgabenteile 0 Punkte (Aufgabenteile mit mehr als einem Kreuz zählen als nicht bearbeitet!).

1. Kreuzen Sie an, auf welche Merkmalseigenschaften der folgende Box-Plot *am ehesten* hindeutet:



(a) leptokurtisch und rechts**steil** ☐
(b) leptokurtisch und links**steil** ☐
(c) platykurtisch und rechts**steil** ☐
(d) platykurtisch und links**steil** ☐

2. Kreuzen Sie an, auf welche Merkmalseigenschaften das folgende Histogramm *am ehesten* hindeutet:



(a) leptokurtisch und rechts**steil** ☐
(b) leptokurtisch und links**steil** ☐
(c) platykurtisch und rechts**steil** ☐
(d) platykurtisch und links**steil** ☐

3. Die Ränge $\text{rg}(X)_1, \ldots, \text{rg}(X)_8$ zur (der Einfachheit halber sortierten) Urliste

sehr zufrieden, sehr zufrieden, zufrieden, zufrieden,
weniger zufrieden, unzufrieden, unzufrieden, unzufrieden

des ordinalskalierten Merkmals X lauten:

(a) 1, 1, 3, 3, 5, 6, 6, 6 ☐
(b) 1, 1, 2, 2, 3, 4, 4, 4 ☐
(c) 1.5, 1.5, 3.5, 3.5, 4, 6.5, 6.5, 6.5 ☐
(d) 1.5, 1.5, 3.5, 3.5, 5, 7, 7, 7 ☐

4. Die Anzahl der verschiedenen (sechsstelligen) Zahlen, die aus den Ziffern 3, 3, 5, 5, 8, 9 gebildet werden können, beträgt:

(a)	180	☐
(b)	256	☐
(c)	720	☐
(d)	16200	☐

5. Sind X_1, X_2 und X_3 drei stochastisch unabhängige Zufallsvariablen mit $X_1 \sim N(40, 6^2)$, $X_2 \sim N(40, 9^2)$ und $X_3 \sim N(40, 2^2)$, dann ist die Verteilung von $X_1 + X_2 + X_3$ eine

(a)	$N(40, 11^2)$-Verteilung.	☐
(b)	$N(40, 17^2)$-Verteilung.	☐
(c)	$N(120, 11^2)$-Verteilung.	☐
(d)	$N(120, 17^2)$-Verteilung.	☐

Aufgabe 3 (4 + 1 + 5 + 1 = 11 Punkte)

Zu einem erhobenen Merkmal X sei die folgende empirische Verteilungsfunktion gegeben:

$$F(x) = \begin{cases} 0.00 & \text{für} \quad x < 1 \\ 0.14 & \text{für} \quad 1 \leq x < 3 \\ 0.56 & \text{für} \quad 3 \leq x < 5 \\ 0.80 & \text{für} \quad 5 \leq x < 7 \\ 0.92 & \text{für} \quad 7 \leq x < 9 \\ 0.98 & \text{für} \quad 9 \leq x < 11 \\ 1.00 & \text{für} \quad x \geq 11 \end{cases}$$

Außerdem sei die Länge der Urliste $n = 50$ bekannt.

(a) Erstellen Sie eine Tabelle der absoluten und relativen Häufigkeiten.

(b) Wie groß ist der Anteil der Urlisteneinträge, die Werte von weniger als 7 annehmen?

(c) Berechnen Sie den arithmetischen Mittelwert und die empirische Varianz des Merkmals X.

(d) Bestimmen Sie einen Median des Merkmals X.

Aufgabe 4 (6 + 4 + 3 + 3 + 3 = 19 Punkte)

Zu einem kardinalskalierten Merkmal sei die folgende (zur einfacheren Bearbeitung der Aufgabe bereits sortierte) Urliste der Länge $n = 40$ gegeben:

11.70, 13.29, 15.84, 18.11, 18.50, 18.97, 22.60, 25.97, 28.00, 29.55, 30.85, 31.83, 32.40, 34.77, 36.13, 39.02, 43.25, 45.80, 46.16, 46.42, 47.66, 48.76, 50.40, 51.25, 55.03, 58.24, 58.41, 61.45, 61.68, 64.21, 64.68, 65.15, 65.19, 70.37, 72.18, 76.02, 77.70, 78.69, 83.57, 86.68

(a) Führen Sie eine Klassierung der erhobenen Daten auf Grundlage der Klassen

$$K_1 = (10, 20],\ K_2 = (20, 40],\ K_3 = (40, 60],\ K_4 = (60, 100]$$

durch. Geben Sie insbesondere die jeweiligen Klassenbreiten, Klassenmitten, absoluten und relativen Klassenhäufigkeiten, Häufigkeitsdichten sowie die Werte der empirischen Verteilungsfunktion an den Klassengrenzen an.

(b) Stellen Sie die (approximative) empirische Verteilungsfunktion des Merkmals aus der Klassierung der Daten auf.

(c) Berechnen Sie aus den klassierten Daten den (approximativen) arithmetischen Mittelwert der Daten. Wie groß ist die relative Abweichung vom tatsächlichen (aus der Urliste bestimmten) Mittelwert von 47.162?

(d) Bestimmen Sie (aus der Urliste) die Anzahl von Merkmalswerten zwischen 20 und 50. Welche Näherung für diese Anzahl können Sie aus der in Teil (b) aufgestellten empirischen Verteilungsfunktion berechnen?

(e) Bestimmen Sie (unter Verwendung der bekannten Konvention zur eindeutigen Festlegung von Quantilen bei kardinalskalierten Merkmalen) *das* untere Quartil sowohl exakt aus der Urliste als auch approximativ mit Hilfe der Verteilungsfunktion für die klassierten Daten.

Aufgabe 5 (4 + 2 = 6 Punkte)

Von den Unternehmen, die sich auf einer 2-tägigen Messe präsentieren, sind 5% nur am 1. Tag sowie 10% nur am 2. Tag vertreten (die restlichen 85% sind an beiden Tagen vertreten).

(a) Wie wahrscheinlich ist es, dass ein (zufällig ausgewähltes) Unternehmen, das am 1. Tag vertreten war, auch am 2. Tag vertreten ist?

(b) Wie wahrscheinlich ist es, dass ein (zufällig ausgewähltes) Unternehmen, das am 2. Tag vertreten ist, nicht schon am 1. Tag vertreten war?

Aufgabe 6 (6 + 2 + 1 = 9 Punkte)

In einer Fußballmannschaft sind die vier Mitspielerinnen Anna, Beate, Christine und Diana für die Ausführung von Eckstößen zuständig. Dabei werden Eckstöße mit einer Wahrscheinlichkeit von 35% von Anna, mit einer Wahrscheinlichkeit von 35% von Beate, mit einer Wahrscheinlichkeit von 20% von Christine und mit einer Wahrscheinlichkeit von 10% von Diana ausgeführt. Aus einer ausführlichen statistischen Auswertung ist bekannt, dass Eckstöße von Anna mit einer Wahrscheinlichkeit von 18%, Eckstöße von Beate mit einer Wahrscheinlichkeit von 15%, Eckstöße von Christine mit einer Wahrscheinlichkeit von 12% und Eckstöße von Diana mit einer Wahrscheinlichkeit von 9% zu einem Tor führen.

(a) Wie groß ist die Wahrscheinlichkeit, dass ein zufällig ausgewählter Eckstoß nicht zu einem Tor führt?

(b) Wie groß ist die Wahrscheinlichkeit, dass ein Eckstoß, der zu einem Tor geführt hat, von Christine ausgeführt wurde?

(c) Sind die Ereignisse „Eckstoß führt zu einem Tor“ und „Eckstoß wird von Christine ausgeführt“ stochastisch unabhängig?

Aufgabe 7 (3 + 2 + 6 + 1 + 4 = 16 Punkte)

Die Verteilung einer stetigen Zufallsvariablen X sei durch die folgende Verteilungsfunktion gegeben:

$$F_X(x) = \begin{cases} 0 & \text{für } x \leq -4 \\ \frac{1}{32}x^2 + \frac{1}{4}x + \frac{1}{2} & \text{für } -4 < x \leq 0 \\ -\frac{1}{16}x^2 + \frac{3}{8}x + \frac{1}{2} & \text{für } 0 < x \leq 2 \\ 1 & \text{für } x > 2 \end{cases}$$

(a) Berechnen Sie eine Dichtefunktion f_X von X.

(b) Berechnen Sie $P(\{X > -2\})$ und $P(\{-2 < X \leq 1\})$.

(c) Bestimmen Sie den Erwartungswert $\mathrm{E}(X)$.

(d) Ist X symmetrisch um ihren Erwartungswert verteilt (ohne Begründung)?

(e) Bestimmen Sie das obere Quartil von X.

Aufgabe 8 (2 + 3 + 9 + 1 + 3 = 18 Punkte)

Gegeben sei die folgende Tabelle der gemeinsamen Wahrscheinlichkeitsverteilung zu einem zweidimensionalen diskreten Zufallsvektor (X, Y):

$X \setminus Y$	-2	0	2	$p_{i\cdot}$
2	0.16	0.04	0.05	
4	0.40	0.15	0.05	
6	0.04	0.01	0.10	
$p_{\cdot j}$				

(a) Ergänzen Sie die obige Tabelle (in den vorgesehenen Feldern) um ihre Randverteilungen.

(b) Geben Sie die bedingten Wahrscheinlichkeitsverteilungen von Y unter der Bedingung $X = x_i$ für alle $x_i \in T(X)$ über die zugehörigen (bedingten) Wahrscheinlichkeitsfunktionen an.

(c) Berechnen Sie $\mathrm{E}(X)$, $\mathrm{E}(Y)$, $\mathrm{Var}(X)$, $\mathrm{Var}(Y)$, $\mathrm{Cov}(X, Y)$ sowie $\mathrm{Korr}(X, Y)$.

(d) Sind X und Y stochastisch unabhängig?

(e) Berechnen Sie $\mathrm{E}(3X-2Y)$ sowie $\mathrm{Var}(3X-2Y)$.

Aufgabe 9 (3 + 2 + 3 = 8 Punkte)

Aufgrund langjähriger Aufzeichnungen über entsprechende Wahlbeteiligungen gehe man davon aus, dass sich die 16600 wahlberechtigten Studierenden bei einer anstehenden Senatswahl unabhängig voneinander jeweils mit einer Wahrscheinlichkeit von 7.5% dazu entschließen, ihr Wahlrecht auch auszuüben und einen Stimmzettel auszufüllen.

(a) Wie ist die Anzahl der ausgefüllten Stimmzettel Y exakt verteilt? Geben Sie auch den Erwartungswert $\mathrm{E}(Y)$ sowie die Varianz $\mathrm{Var}(Y)$ der Anzahl der ausgefüllten Stimmzettel an.

(b) Die Wahlleitung entschließt sich aus ökologischen und ökonomischen Gründen, zunächst nur 1300 Stimmzettel für die Wahl auszudrucken und weitere Stimmzettel erst bei Bedarf nachzudrucken, falls mehr als die zunächst gedruckten Stimmzettel benötigt werden sollten. Mit welcher (mit dem zentralen Grenzwertsatz näherungsweise zu bestimmenden) Wahrscheinlichkeit müssen keine Stimmzettel nachgedruckt werden?

(c) Verwenden Sie den zentralen Grenzwertsatz, um näherungsweise die (kleinste) Anzahl von Stimmzetteln zu bestimmen, die mit einer Wahrscheinlichkeit von (mindestens) 99% ausreichend ist.

Hinweis: Verwenden Sie zur Bearbeitung von Aufgabenteil (b) und (c) die Tabelle zur Standardnormalverteilung auf Seite 120!

Tabelle zur Standardnormalverteilung

$$F_{N(0,1)}(x) = \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}t^2} dt$$

$$\Phi(x) = 1 - \Phi(-x)$$

	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998