

**WIRTSCHAFTSWISSENSCHAFTLICHES
PRÜFUNGSSEKRETARIAT**

FAKULTÄT FÜR EMPIRISCHE HUMANWISSENSCHAFTEN UND WIRTSCHAFTSWISSENSCHAFT
DER UNIVERSITÄT DES SAARLANDES

Von der/dem Studierenden auszufüllen (Bitte leserlich und in Blockschrift):

Deskriptive Statistik und Wahrscheinlichkeitsrechnung

Name der Prüfung:

Semester, dem die

Prüfung zugeordnet ist:

WS 2018/19

(z. B. WS 2015/2016, SS 2016)

(Prüfungen im Februar/April = WS; Prüfungen im August/Okttober = SS)

**Nachname, Vorname
der/des Studierenden:**

Matrikelnummer der/des

Studierenden:

Erklärung

Hiermit erkläre ich, dass mir die für diese Prüfung relevanten Zulassungsvoraussetzungen aus der für mich geltenden Prüfungsordnung bekannt sind.

Mir ist damit bewusst, dass diese Prüfungsleistung als nicht abgelegt zählt, wenn die hierfür vorgesehenen Zulassungsvoraussetzungen nicht erfüllt sind.

Mir ist bekannt, dass die Teilnahme an der Prüfung zudem die ordnungsgemäße Anmeldung zur Prüfung voraussetzt. Die Teilnahme bei versäumter Anmeldung hat die Ungültigkeit der Prüfung zur Folge.

Zudem ist mir bekannt, dass eine nicht bestandene Prüfung zweimal wiederholt werden kann. Die Wiederholung einer bestandenen Prüfung ist nicht zulässig.

Datum: _____

Unterschrift der/des Studierenden: _____

Von der Prüferin/Vom Prüfer auszufüllen:

Aufgabe	Punkte	Max. Punkte	Bemerkungen
1+2		28	
3		11	
4		19	
5		6	
6		8	
7		13	
8		7	
9		18	
10		10	
Summe		120	

bestanden

Note: _____

nicht bestanden

Unterschrift der Prüferin/des Prüfers: _____

KLAUSURHEFT ZUR
 BACHELOR-PRÜFUNG
 DESKRIPTIVE STATISTIK UND WAHRSCHEINLICHKEITSRECHNUNG
 WINTERSEMESTER 2018/19

PD Dr. Martin Becker

Hinweise für die Klausurteilnehmer

- Die Klausur besteht aus insgesamt 10 Aufgaben. Prüfen Sie die Vollständigkeit Ihres Exemplares nach; spätere Reklamationen können nicht berücksichtigt werden.
- Es sind insgesamt 120 Punkte (= 16 + 12 + 11 + 19 + 6 + 8 + 13 + 7 + 18 + 10) erreichbar.
- Als Hilfsmittel sind zugelassen: Taschenrechner (auch mit Grafikfähigkeit), 2 selbstgestellte DIN-A4 Blätter bzw. 4 selbstgestellte (einseitige) DIN-A4 Seiten. Benötigte Tabellen finden Sie am Ende dieses Klausurheftes.
- Bei der Korrektur werden **nur** die Lösungen auf diesen Blättern berücksichtigt.
- Bei mehreren Lösungsvorschlägen muss die gültige Lösung eindeutig gekennzeichnet sein.
- Mit Ausnahme der Multiple-Choice-Aufgaben muss der Lösungsweg klar ersichtlich sein.
- Alle Punkte innerhalb von Zahlen sind stets Dezimalpunkte (und keine Tausenderpunkte).

Bewertungsteil — Bitte nicht beschreiben						
Aufgabe	(a)	(b)	(c)	(d)	(e)	Σ
1		■	■	■	■	
2		■	■	■	■	
3					■	
4						
5				■	■	
6				■	■	
7				■	■	
8				■	■	
9						
10				■	■	
Σ						

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!).

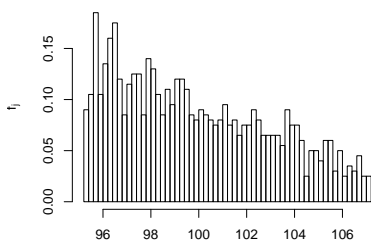
- | | wahr | falsch |
|---|--------------------------|--------------------------|
| 1. Alle Lagemaße, die für ordinalskalierte Merkmale berechnet werden können, können auch für nominalskalierte Merkmale berechnet werden. | <input type="checkbox"/> | <input type="checkbox"/> |
| 2. Betrachtet man die 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 $a = \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$. | <input type="checkbox"/> | <input type="checkbox"/> |
| 3. Es gibt Laplacesche Wahrscheinlichkeitsräume mit Ergebnismenge $\Omega = \mathbb{N}$. | <input type="checkbox"/> | <input type="checkbox"/> |
| 4. Die Wahrscheinlichkeit, mit drei fairen (sechseckigen) Würfeln bei gleichzeitigem Würfeln jeweils eine 2, eine 4 sowie eine 6 zu würfeln, beträgt (gerundet) 2.778%. | <input type="checkbox"/> | <input type="checkbox"/> |
| 5. Wenn Sie alle 8 Aufgabenteile dieser Aufgabe entweder mit <i>wahr</i> oder mit <i>falsch</i> beantworten, dann haben Sie insgesamt $\binom{8}{2}$ Möglichkeiten zur Bearbeitung der Aufgabe. | <input type="checkbox"/> | <input type="checkbox"/> |
| 6. Es seien (Ω, \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:
$P(A) \neq P(B) \implies P(A B) = P(B A) = 0$ | <input type="checkbox"/> | <input type="checkbox"/> |
| 7. Ist F_X die Verteilungsfunktion einer stetigen Zufallsvariablen X , so gilt für $a, b \in \mathbb{R}$ mit $a < b$ stets
$P(\{a < X < b\}) = \int_a^b F_X(x) dx .$ | <input type="checkbox"/> | <input type="checkbox"/> |
| 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). | <input type="checkbox"/> | <input type="checkbox"/> |

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 rechtssteil
- (b) leptokurtisch und linkssteil
- (c) platykurtisch und rechtssteil
- (d) platykurtisch und linkssteil

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.

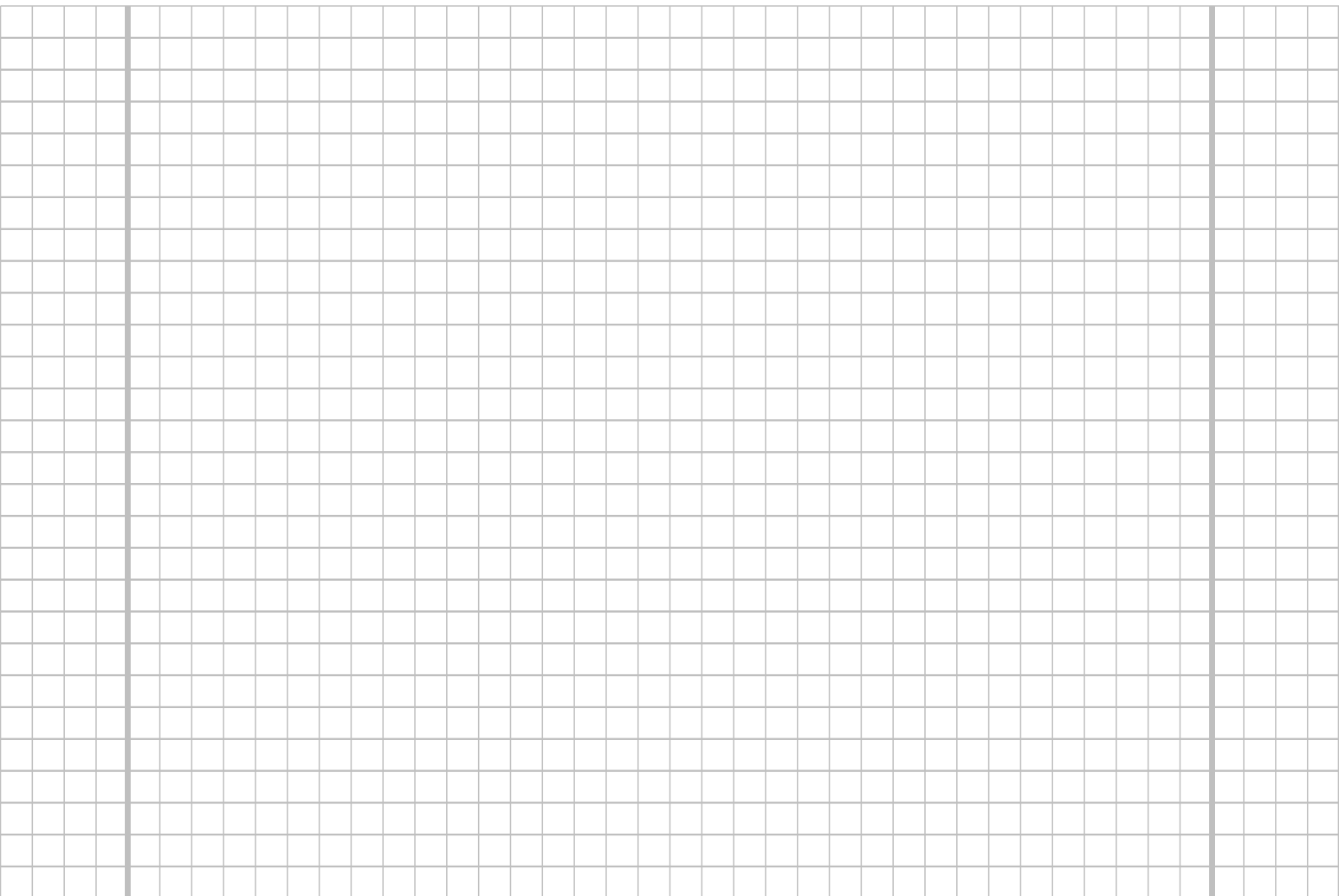
Aufgabe 3 (4 + 1 + 5 + 1 = 11 Punkte)

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

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

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:

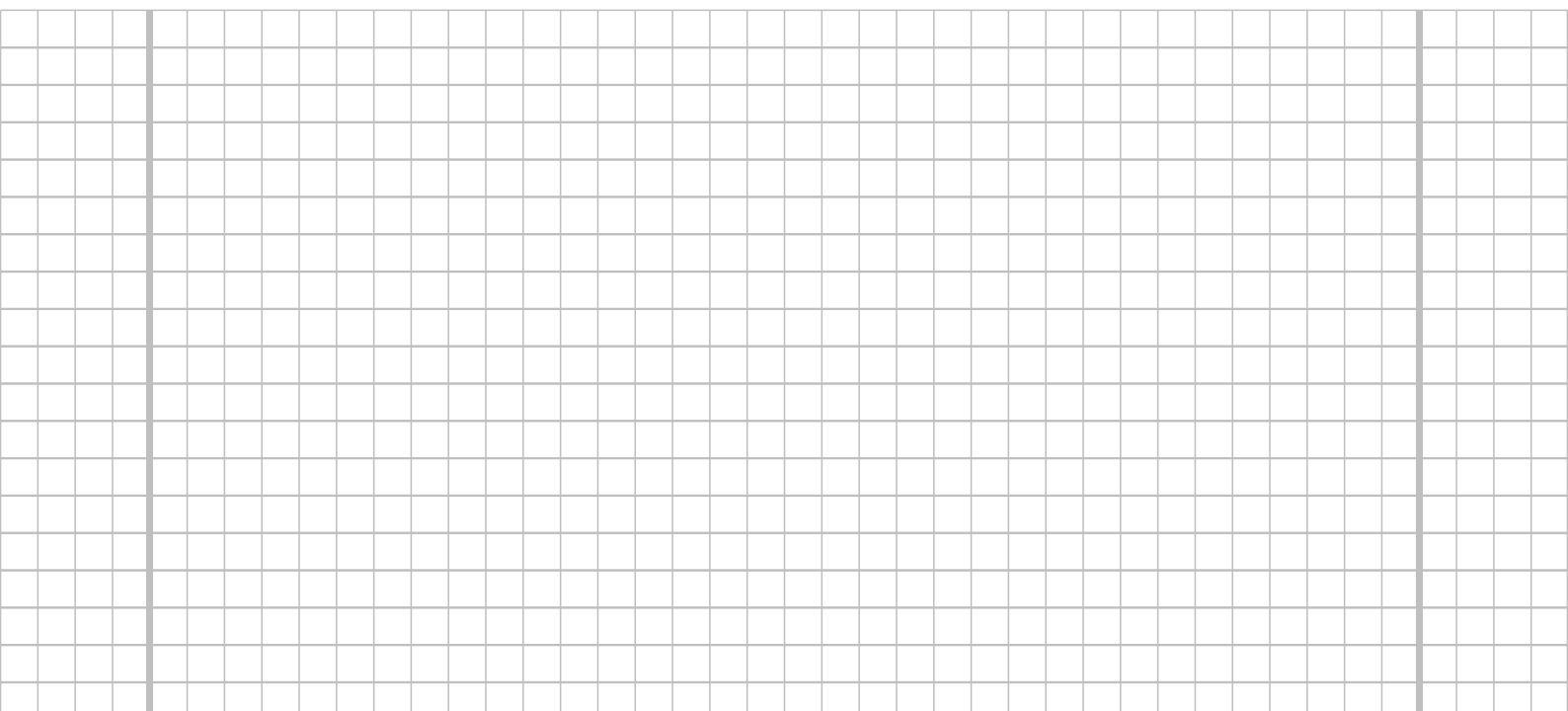
19.47, 19.63, 20.24, 20.56, 28.27, 29.29, 30.77, 33.97, 36.80, 39.17, 42.96, 47.32, 47.39, 49.86, 51.99, 59.47, 64.66, 64.87, 67.07, 69.79, 70.40, 71.27, 78.11, 78.74, 79.65, 80.19, 80.52, 81.15, 81.65, 82.52, 84.35, 84.64, 85.30, 88.21, 88.21, 88.85, 90.20, 90.70, 96.65, 96.70

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

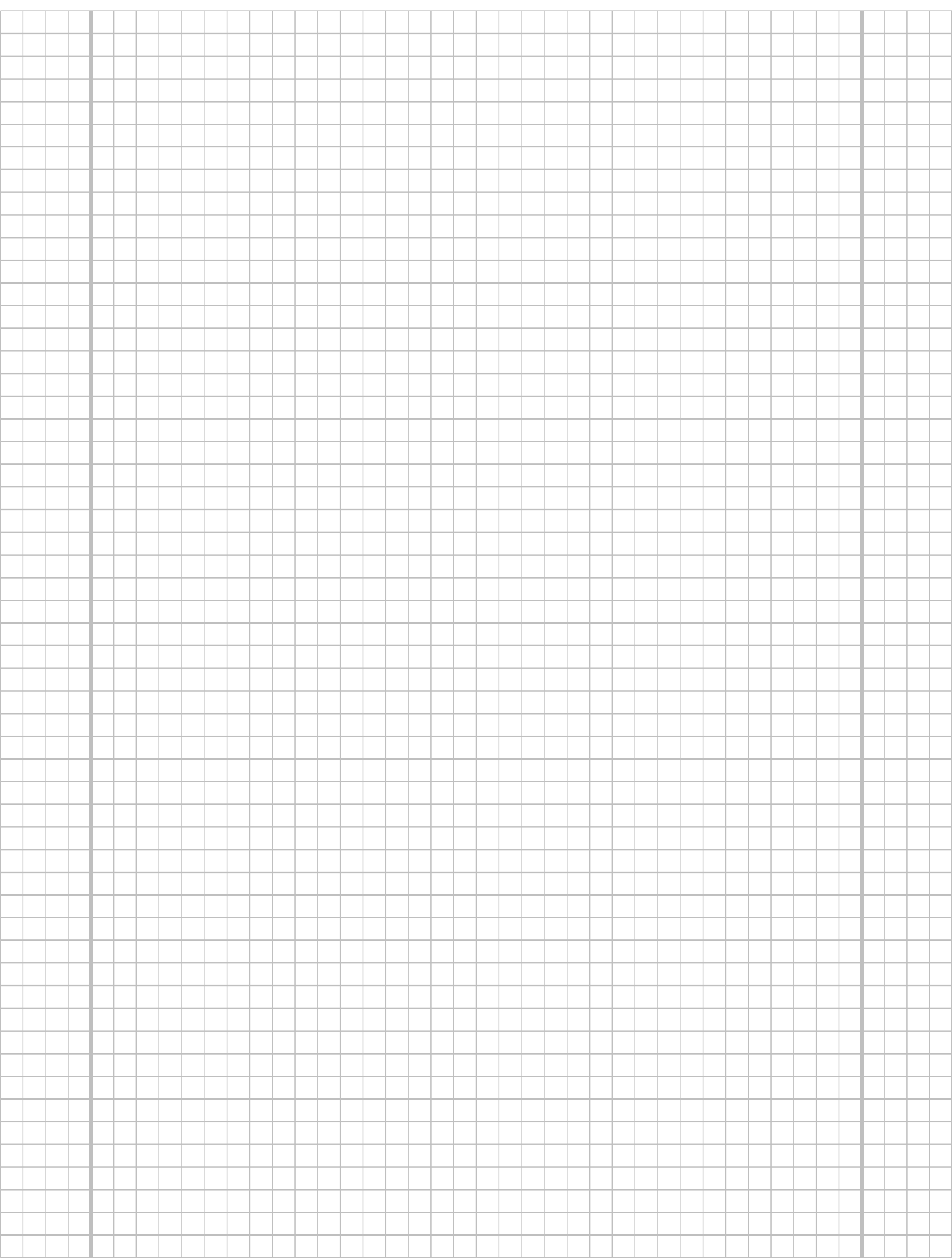
$$K_1 = (10, 40], K_2 = (40, 60], K_3 = (60, 80], K_4 = (80, 100]$$

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!

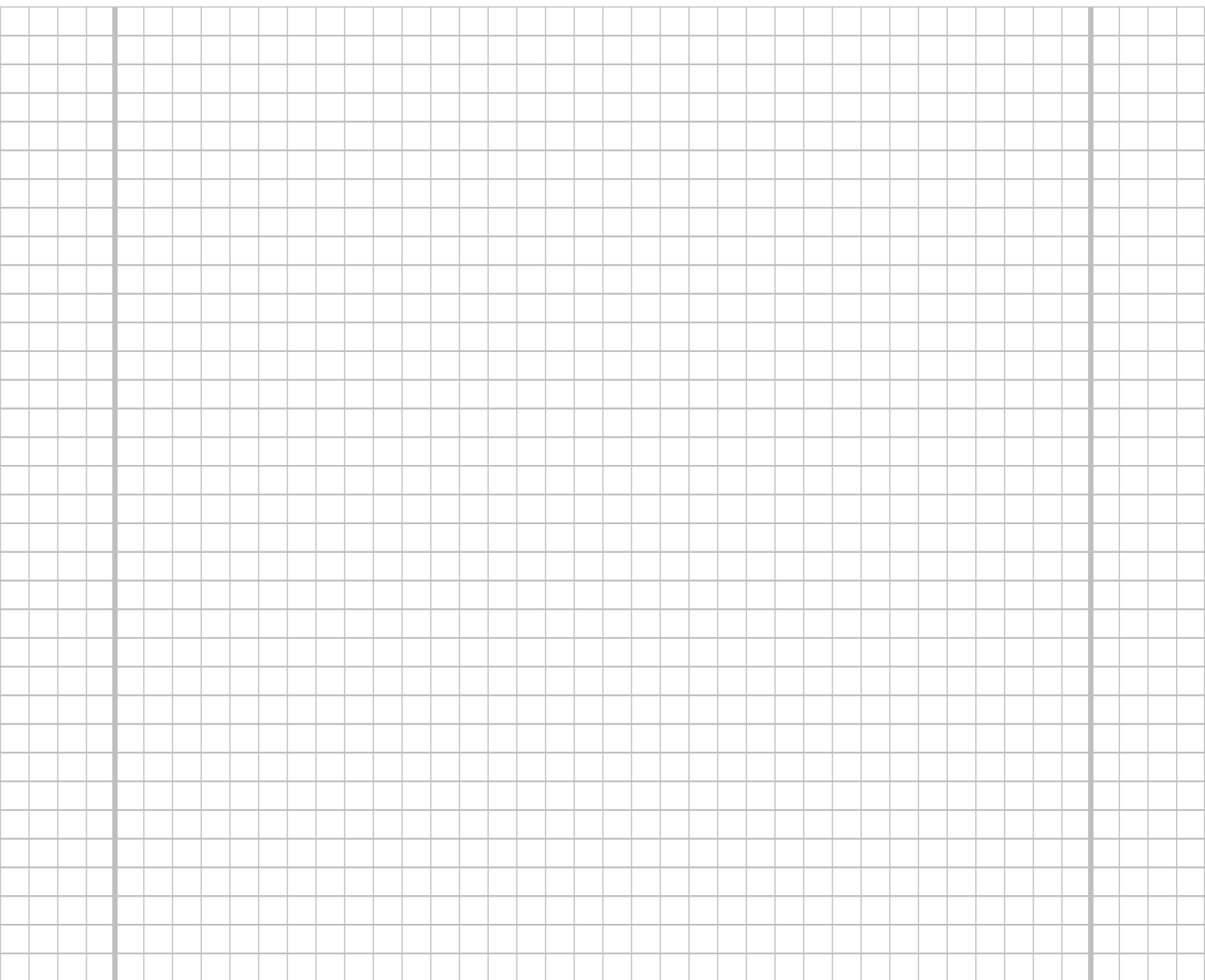




Aufgabe 6 (4 + 2 + 2 = 8 Punkte)

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?



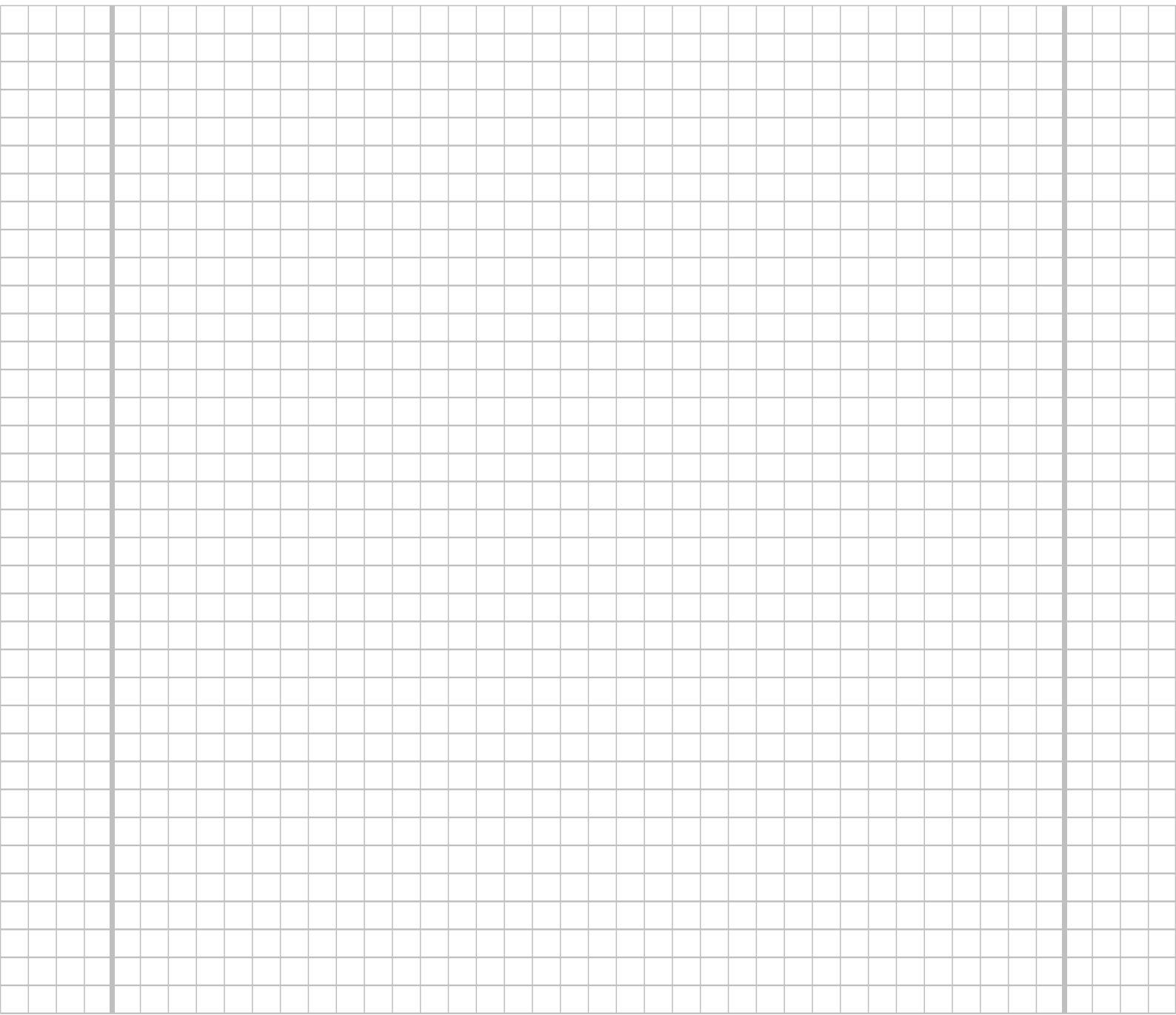


Aufgabe 7 (3 + 6 + 4 = 13 Punkte)

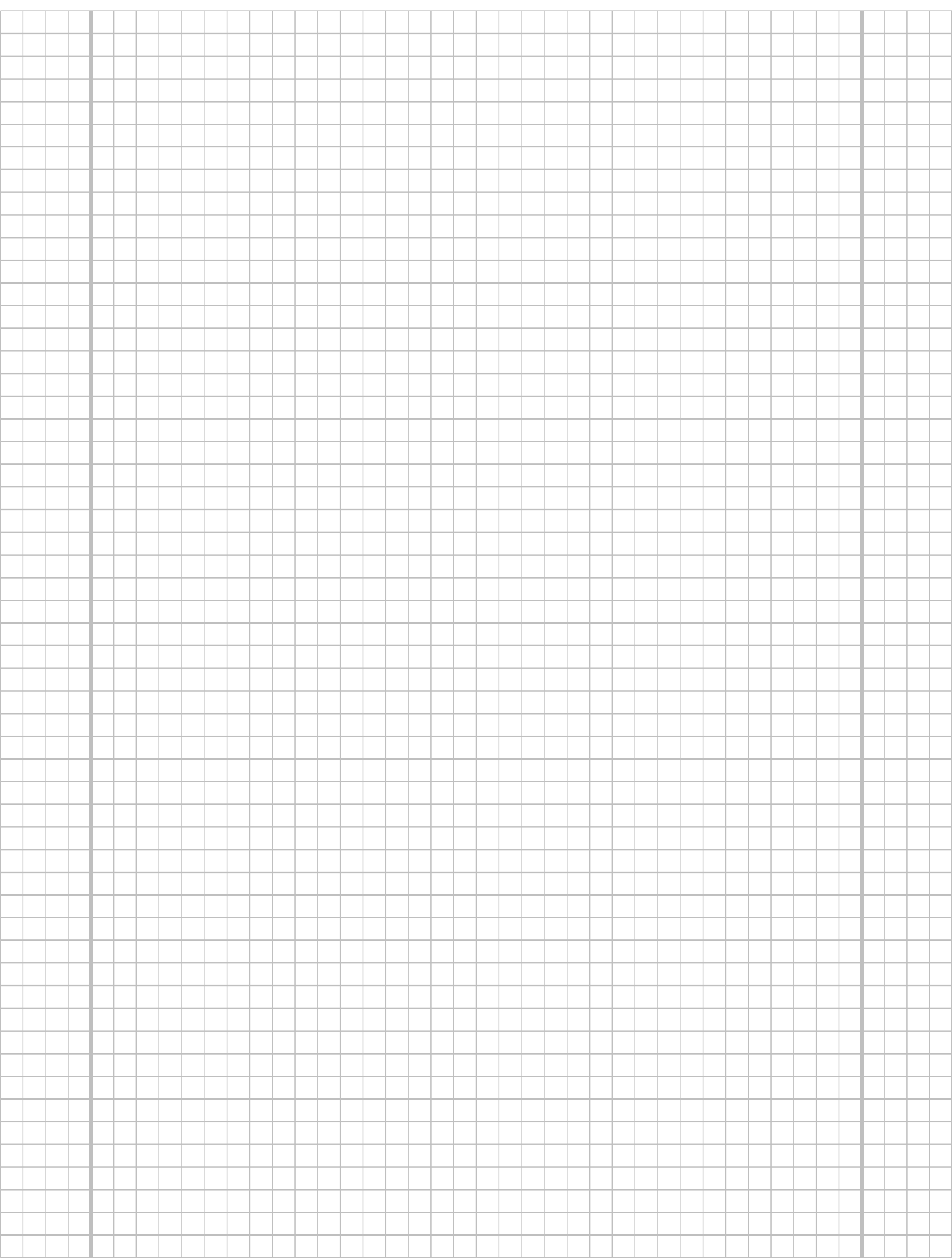
Die Verteilung einer stetigen Zufallsvariablen X sei durch die folgende Verteilungsfunktion gegeben:

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

- (a) Berechnen Sie eine Dichtefunktion f_X von X .
- (b) Bestimmen Sie den Erwartungswert $E(X)$.
- (c) Bestimmen Sie das obere Quartil von X .





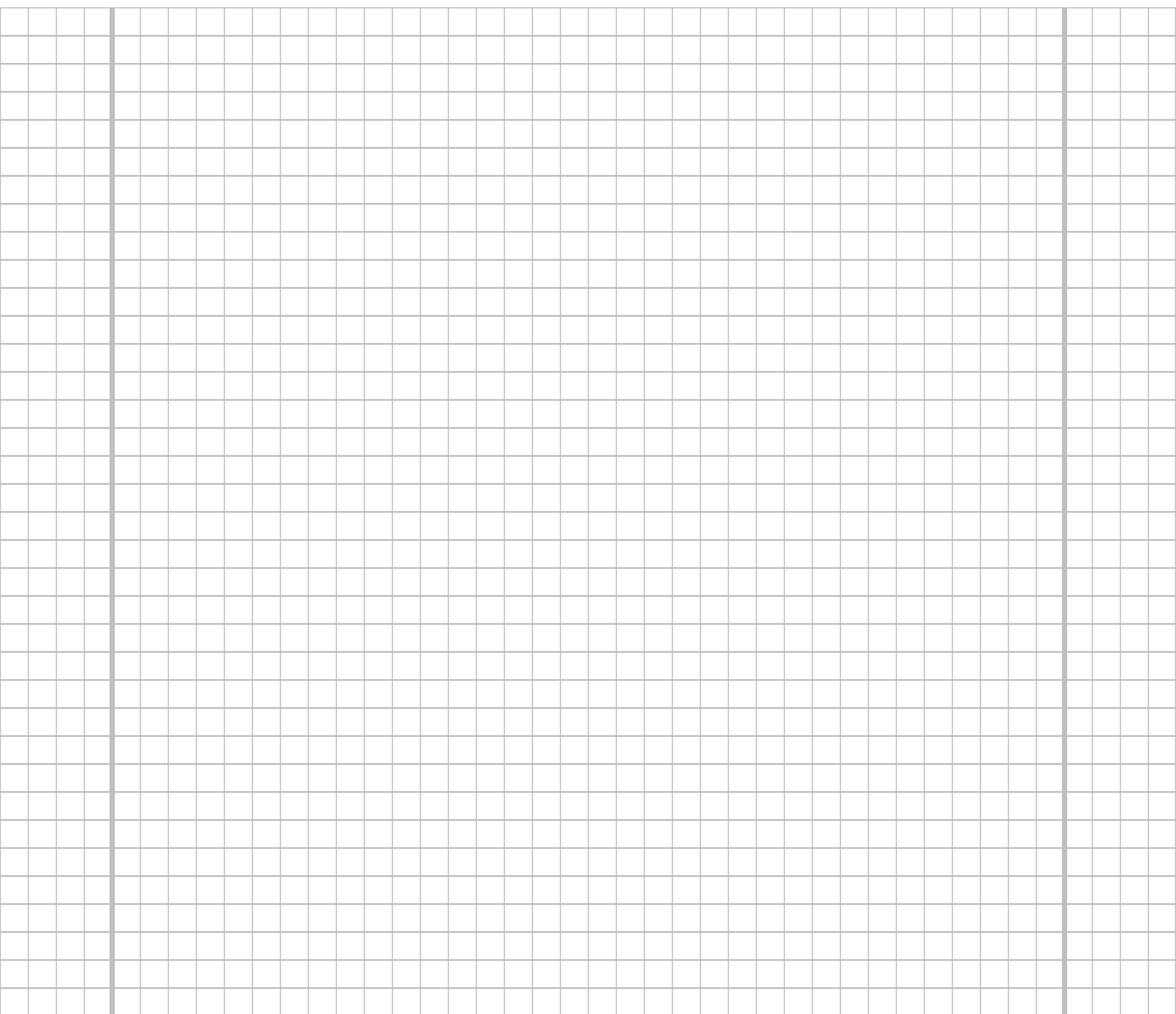




Aufgabe 8 (3 + 2 + 2 = 7 Punkte)

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





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

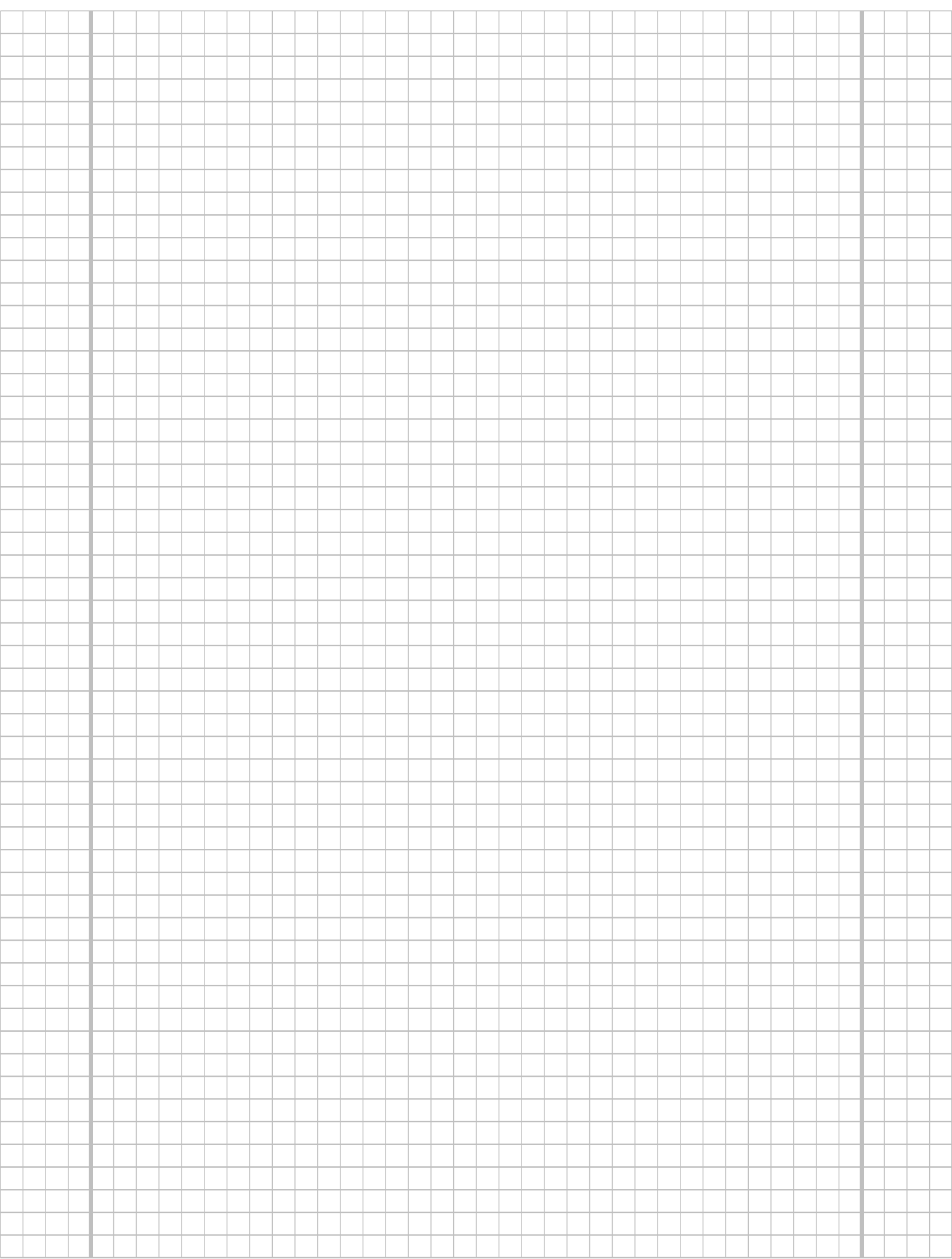
Gegeben sei die folgende Tabelle der gemeinsamen Wahrscheinlichkeitsverteilung zu einem zweidimensionalen diskreten Zufallsvektor (X, Y) :

$X \setminus Y$	-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 $E(X)$, $E(Y)$, $\text{Var}(X)$, $\text{Var}(Y)$, $\text{Cov}(X, Y)$ sowie $\text{Korr}(X, Y)$.
- (d) Sind X und Y stochastisch unabhängig?
- (e) Berechnen Sie $E(5X - 3Y)$ sowie $\text{Var}(5X - 3Y)$.









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

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

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

bezeichnet.

- (a) Geben Sie die (exakte) Verteilung von Y sowie deren Erwartungswert $E(Y)$ und Varianz $\text{Var}(Y)$ 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 26!

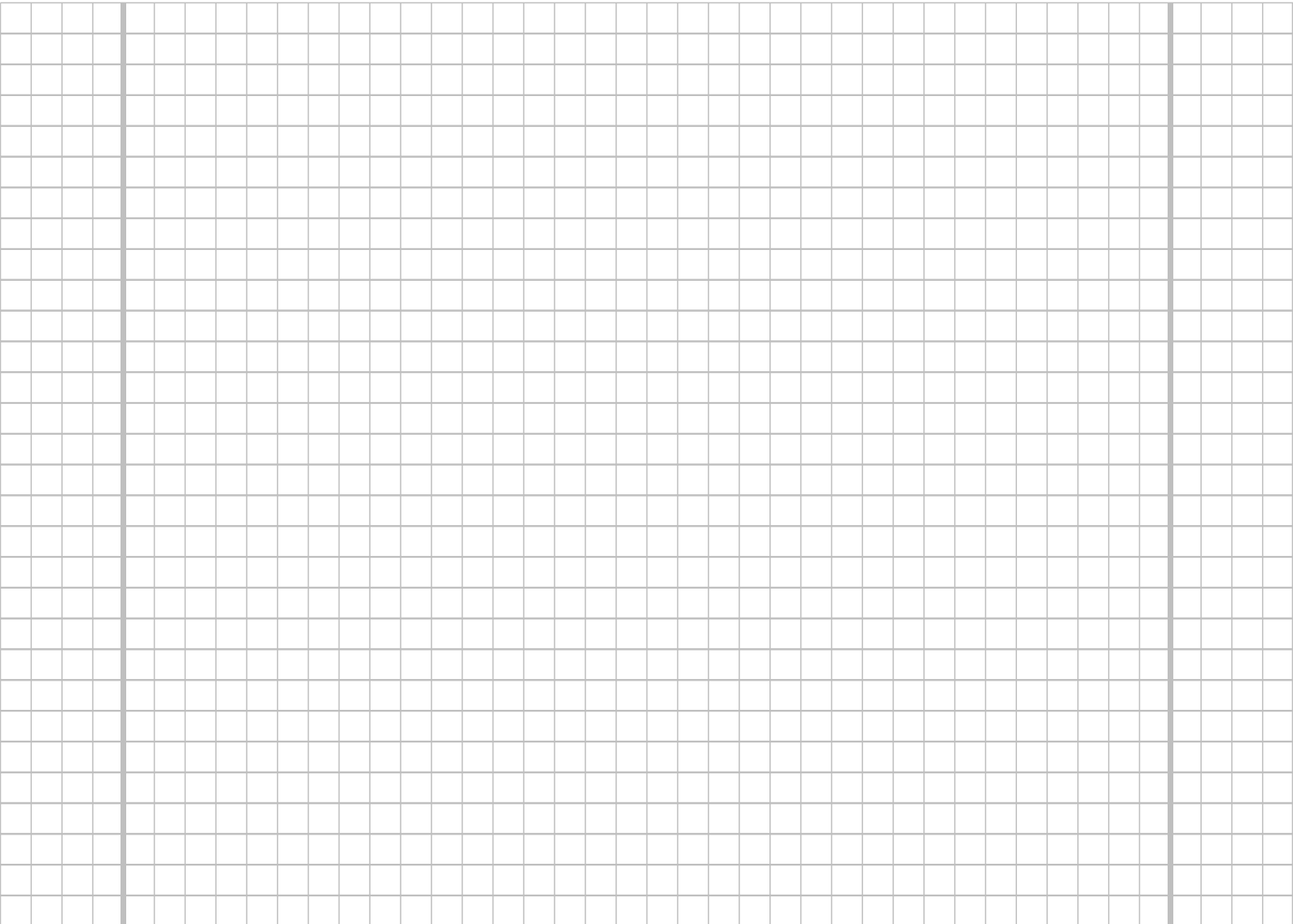




Tabelle zur Standardnormalverteilung

$$F_{N(0,1)}(x) = \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{1}{2}t^2} dt$$

$$\Phi(x) = 1 - \Phi(-x)$$

	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