

**WIRTSCHAFTSWISSENSCHAFTLICHES
PRÜFUNGSSEKRETARIAT**

FAKULTÄT FÜR EMPIRISCHE HUMANWISSENSCHAFTEN UND WIRTSCHAFTSWISSENSCHAFT
DER UNIVERSITÄT DES SAARLANDES

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

Name der Prüfung: Deskriptive Statistik und Wahrscheinlichkeitsrechnung
Semester, dem die Prüfung zugeordnet ist: WS 2022/23 (z. B. WS 2019/2020, SS 2020)
(Prüfungen im Februar/April = WS; Prüfungen im August/Oktobre = 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		18	
2		15	
3		11	
4		19	
5		9	
6		17	
7		6	
8		17	
9		8	
<i>Summe</i>		120	

bestanden Note: _____

nicht bestanden Unterschrift der Prüferin/des Prüfers: _____

KLAUSURHEFT ZUR
 BACHELOR-PRÜFUNG
 DESKRIPTIVE STATISTIK UND WAHRSCHEINLICHKEITSRECHNUNG
 WINTERSEMESTER 2022/23

apl. Prof. Dr. Martin Becker

Hinweise für die Klausurteilnehmer

- Die Klausur besteht aus insgesamt 9 Aufgaben. Prüfen Sie die Vollständigkeit Ihres Exemplares nach; spätere Reklamationen können nicht berücksichtigt werden.
- Es sind insgesamt 120 Punkte (= 18 + 15 + 11 + 19 + 9 + 17 + 6 + 17 + 8) erreichbar.
- Als Hilfsmittel sind zugelassen: Taschenrechner (auch mit Grafikfähigkeit), 2 selbstgestaltete DIN-A4 Blätter bzw. 4 selbstgestaltete (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				■	■	
Σ						

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!).

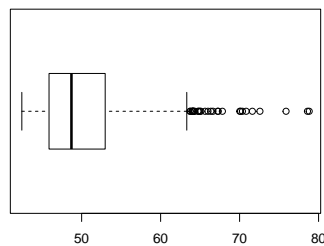
- | | wahr | falsch |
|---|--------------------------|--------------------------|
| 1. Empirische Verteilungsfunktionen F können streng monoton wachsend auf \mathbb{R} sein. | <input type="checkbox"/> | <input type="checkbox"/> |
| 2. Die jährlichen Inflationsraten in Deutschland betragen in den Jahren 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%. | <input type="checkbox"/> | <input type="checkbox"/> |
| 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. | <input type="checkbox"/> | <input type="checkbox"/> |
| 4. Sind (Ω, \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. | <input type="checkbox"/> | <input type="checkbox"/> |
| 5. Die Wahrscheinlichkeit, durch rein zufällige Anordnung der Buchstaben E, E, K, L, S, S das Wort KESSEL zu erhalten, beträgt (gerundet) 0.556%. | <input type="checkbox"/> | <input type="checkbox"/> |
| 6. Exponentialverteilte Zufallsvariablen sind nie symmetrisch. | <input type="checkbox"/> | <input type="checkbox"/> |
| 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. | <input type="checkbox"/> | <input type="checkbox"/> |
| 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). | <input type="checkbox"/> | <input type="checkbox"/> |
| 9. Sind X_1, \dots, X_n stochastisch unabhängige Zufallsvariablen mit $X_i \sim B(n_i, p_i)$, $i \in \{1, \dots, n\}$, dann genügt $Y = \sum_{i=1}^n X_i$ einer $B(n, p)$ -Verteilung mit $n = \sum_{i=1}^n n_i$ und $p = \sum_{i=1}^n p_i$. | <input type="checkbox"/> | <input type="checkbox"/> |

Aufgabe 2 (15 Punkte)

Markieren Sie jeweils die korrekte Antwort mit einem Kreuz im betreffenden Kästchen. Es ist jeweils genau ein Kreuz korrekt.

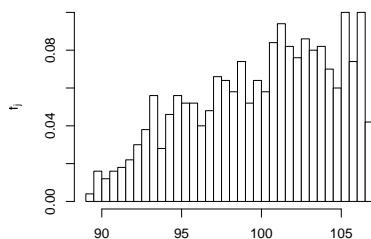
Richtige Antworten geben 3 Punkte, falsche Antworten und nicht bearbeitete Aufgabenteile 0 Punkte (Aufgabenteile mit mehr als einem Kreuz zählen als nicht bearbeitet!).

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



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

2. 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

3. Von den beiden Ereignissen A und B eines Wahrscheinlichkeitsraums (Ω, \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 Wahrscheinlichkeitsraum (Ω, \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:

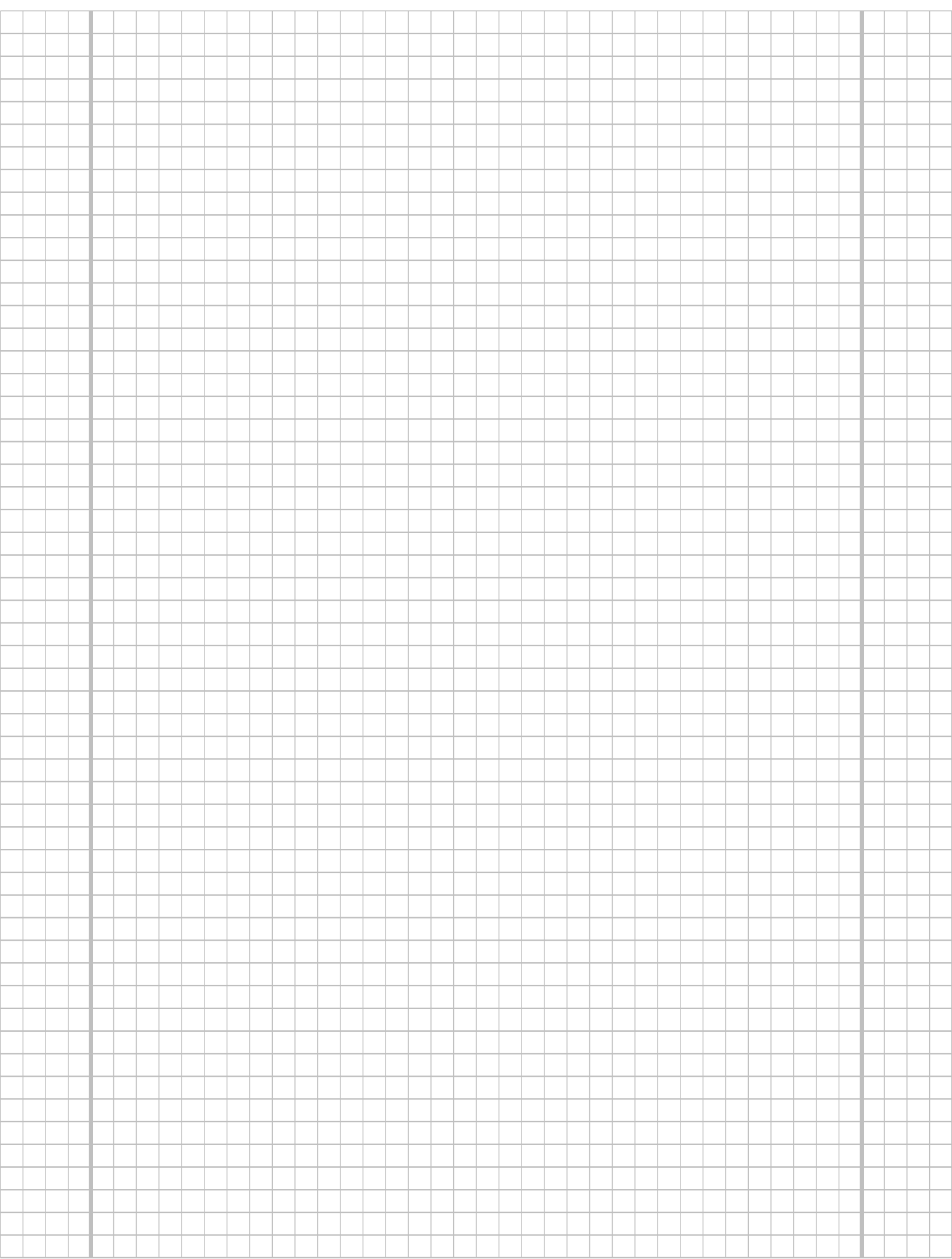
$$F(x) = \begin{cases} 0.000 & \text{für } x < 3 \\ 0.225 & \text{für } 3 \leq x < 4 \\ 0.575 & \text{für } 4 \leq x < 5 \\ 0.800 & \text{für } 5 \leq x < 6 \\ 0.950 & \text{für } 6 \leq x < 7 \\ 1.000 & \text{für } x \geq 7 \end{cases}$$

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:

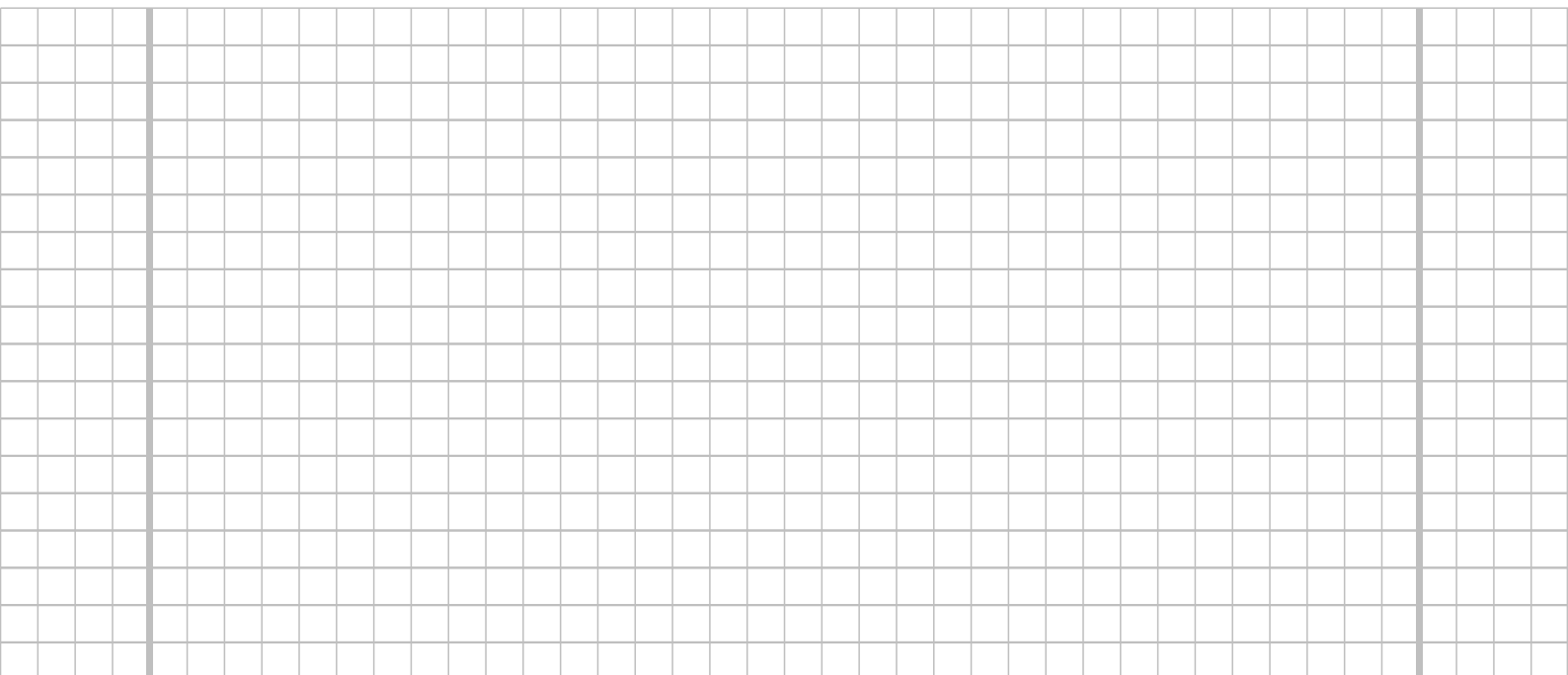
10.62, 11.15, 11.26, 11.65, 11.73, 14.40, 14.40, 15.21, 17.37, 18.89, 19.50, 19.63, 19.77, 20.34, 21.32, 22.21, 22.43, 23.42, 23.61, 24.10, 25.84, 26.27, 26.44, 26.53, 26.67, 27.12, 27.41, 28.04, 28.18, 28.26, 30.54, 30.77, 31.11, 31.26, 32.01, 32.49, 35.01, 35.24, 35.98, 36.05, 36.15, 36.47, 36.93, 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

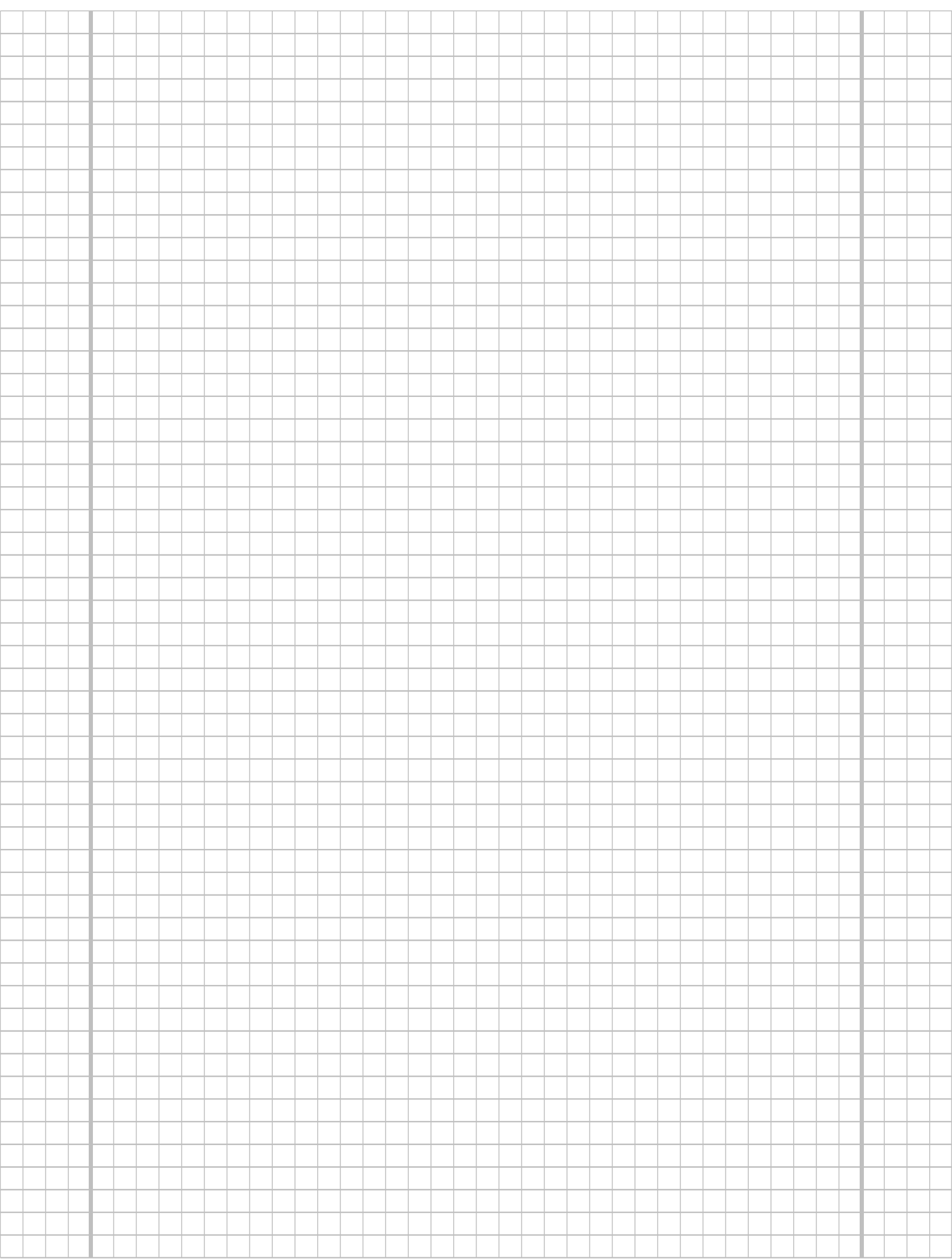
$$K_1 = (0, 15], K_2 = (15, 25], K_3 = (25, 35], K_4 = (35, 40]$$

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.





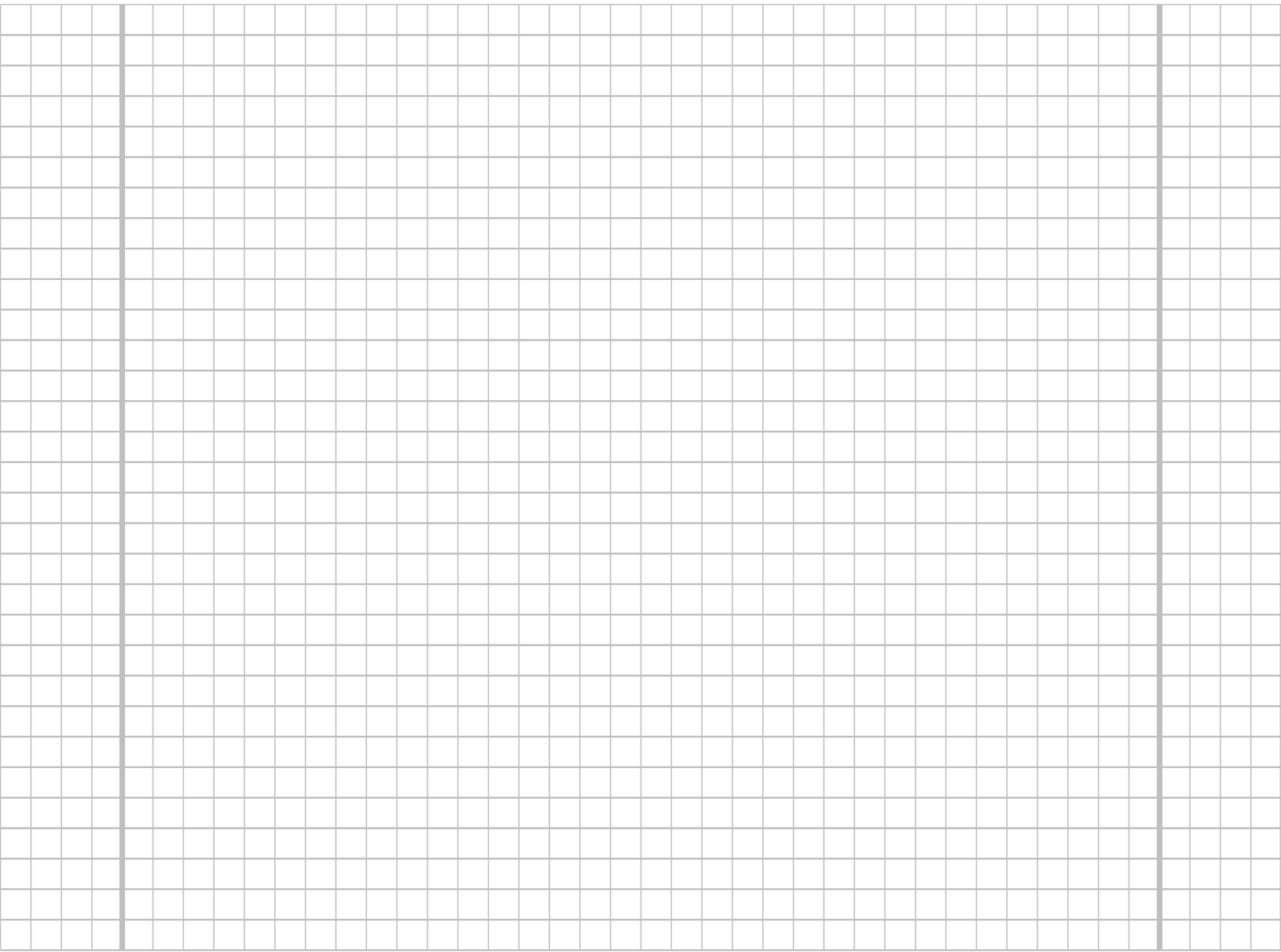




Aufgabe 5 (6 + 2 + 1 = 9 Punkte)

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?



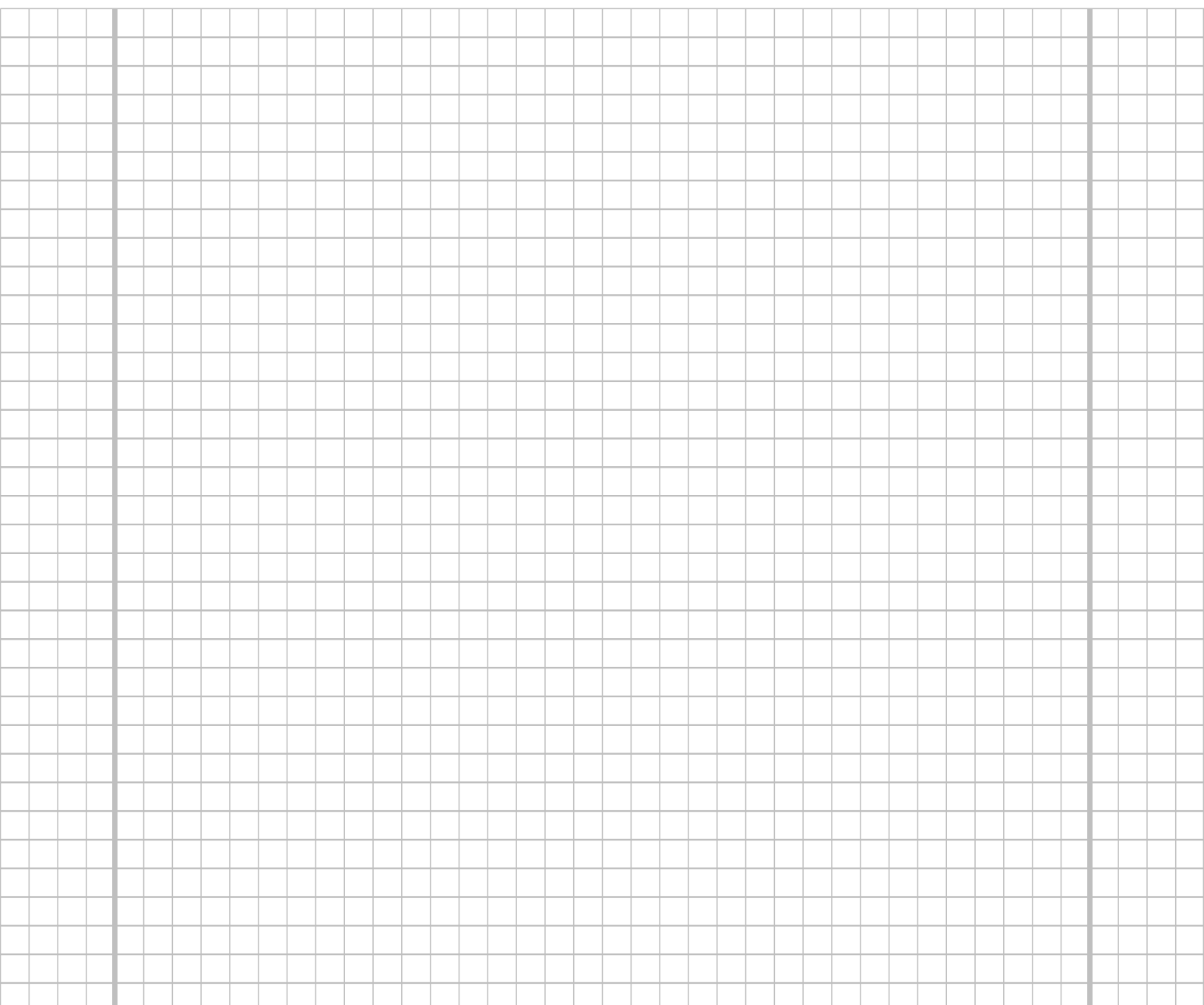


Aufgabe 6 (5 + 2 + 6 + 4 = 17 Punkte)

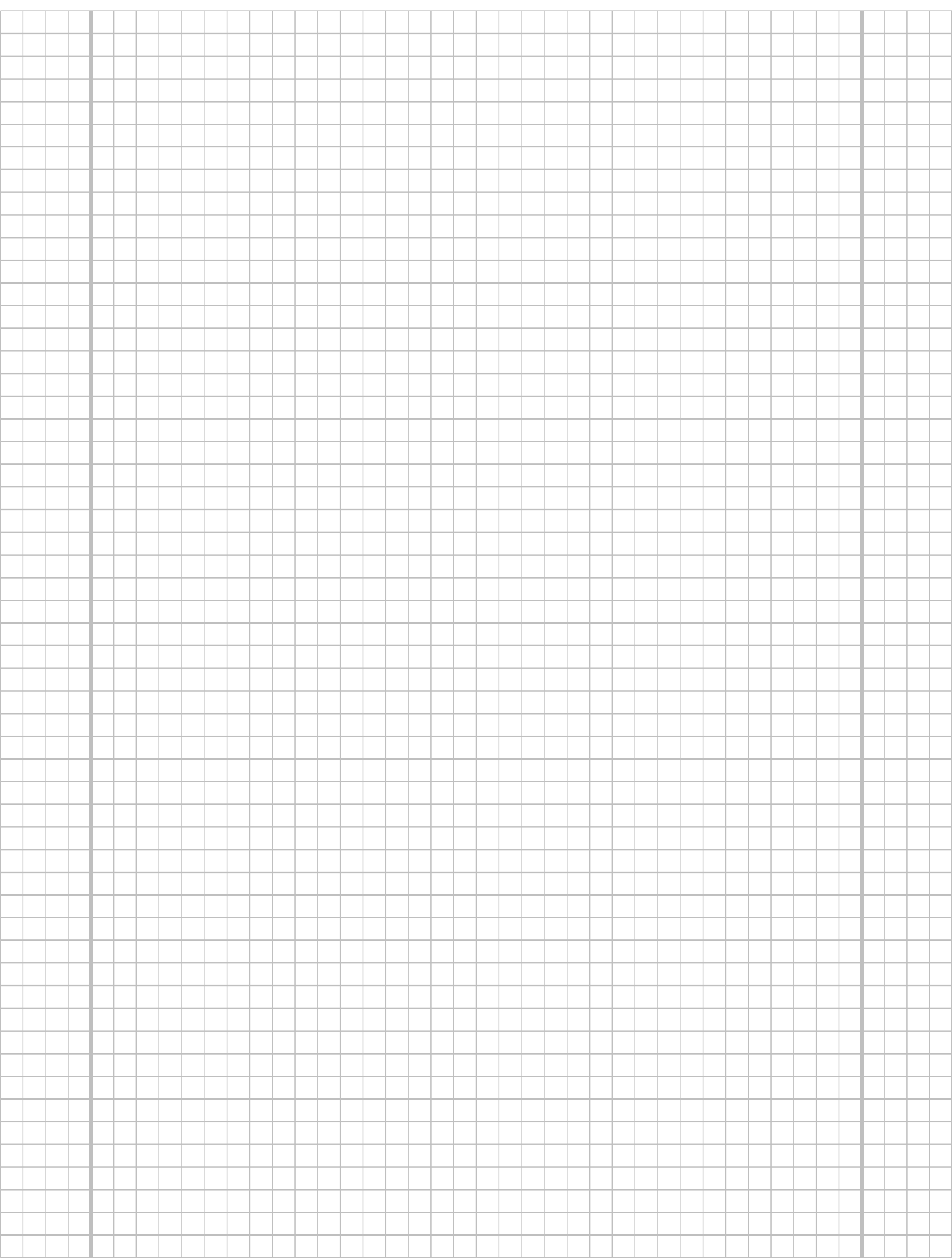
Die Verteilung einer stetigen Zufallsvariablen X sei durch die folgende Dichtefunktion gegeben:

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

- (a) Berechnen Sie die Verteilungsfunktion F_X von X .
- (b) Berechnen Sie $P(\{X < -\frac{1}{2}\})$ und $P(\{-\frac{1}{2} \leq X \leq \frac{1}{2}\})$.
- (c) Bestimmen Sie den Erwartungswert $E(X)$.
- (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 $\text{Pois}(0.1)$ -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)

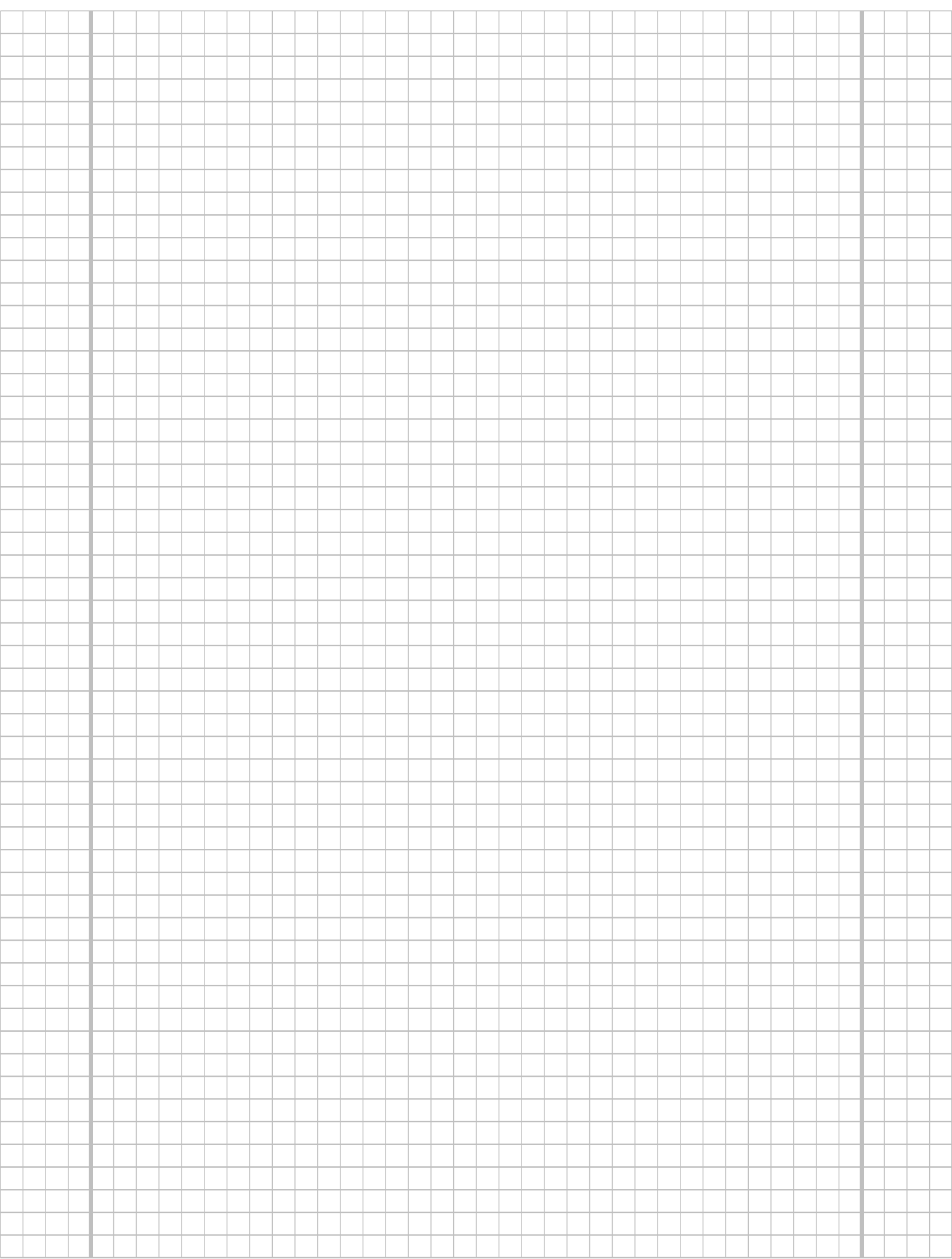
Gegeben sei die folgende Tabelle der gemeinsamen Wahrscheinlichkeitsverteilung zu einem zweidimensionalen diskreten Zufallsvektor (X, Y) :

$X \setminus Y$	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 $E(X)$, $E(Y)$, $\text{Var}(X)$, $\text{Var}(Y)$, $\text{Cov}(X, Y)$ sowie $\text{Korr}(X, Y)$.
- (d) Berechnen Sie $E(-2X + 4Y)$ sowie $\text{Var}(-2X + 4Y)$.







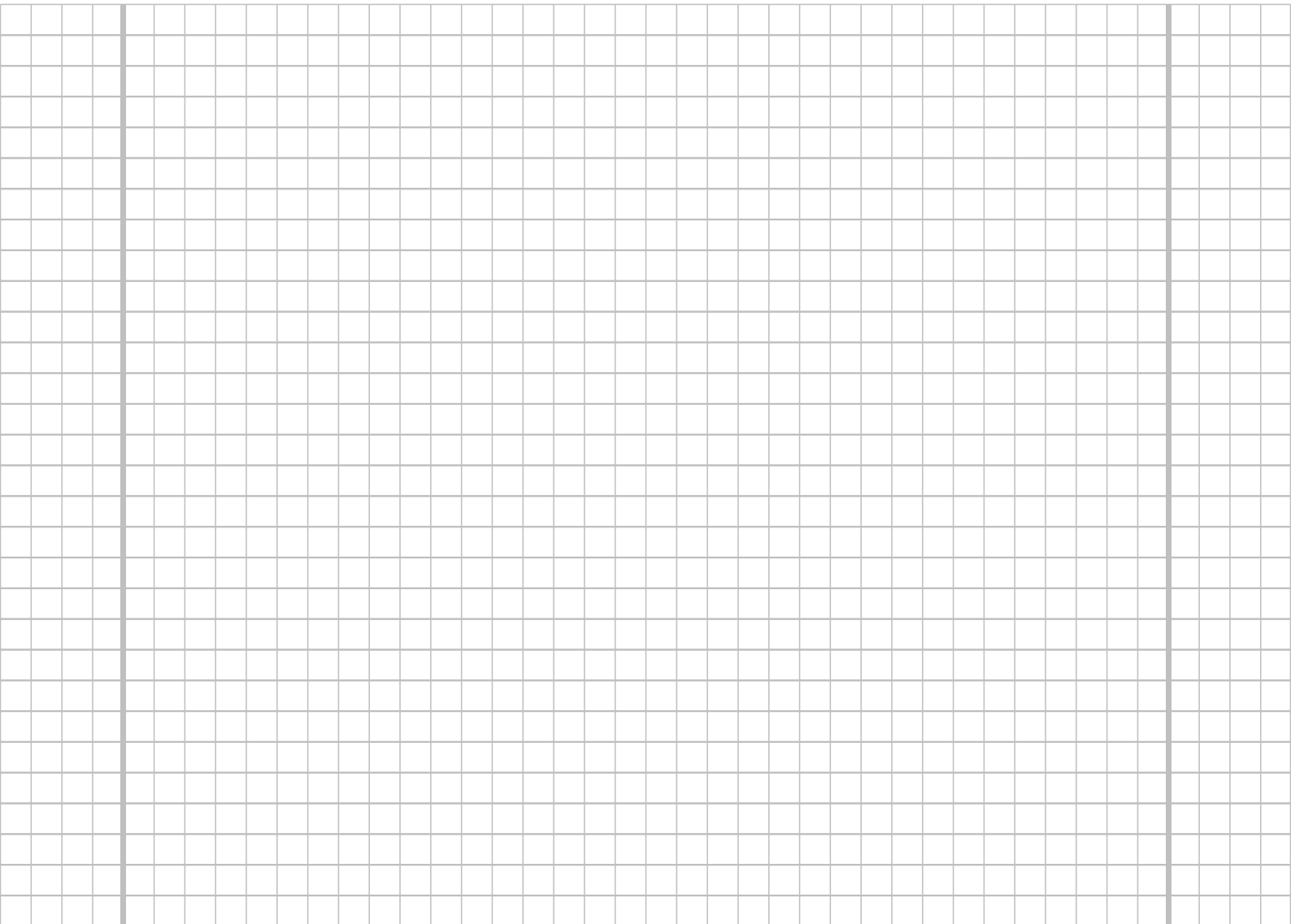


Aufgabe 9 (2 + 2 + 4 = 8 Punkte)

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 28!





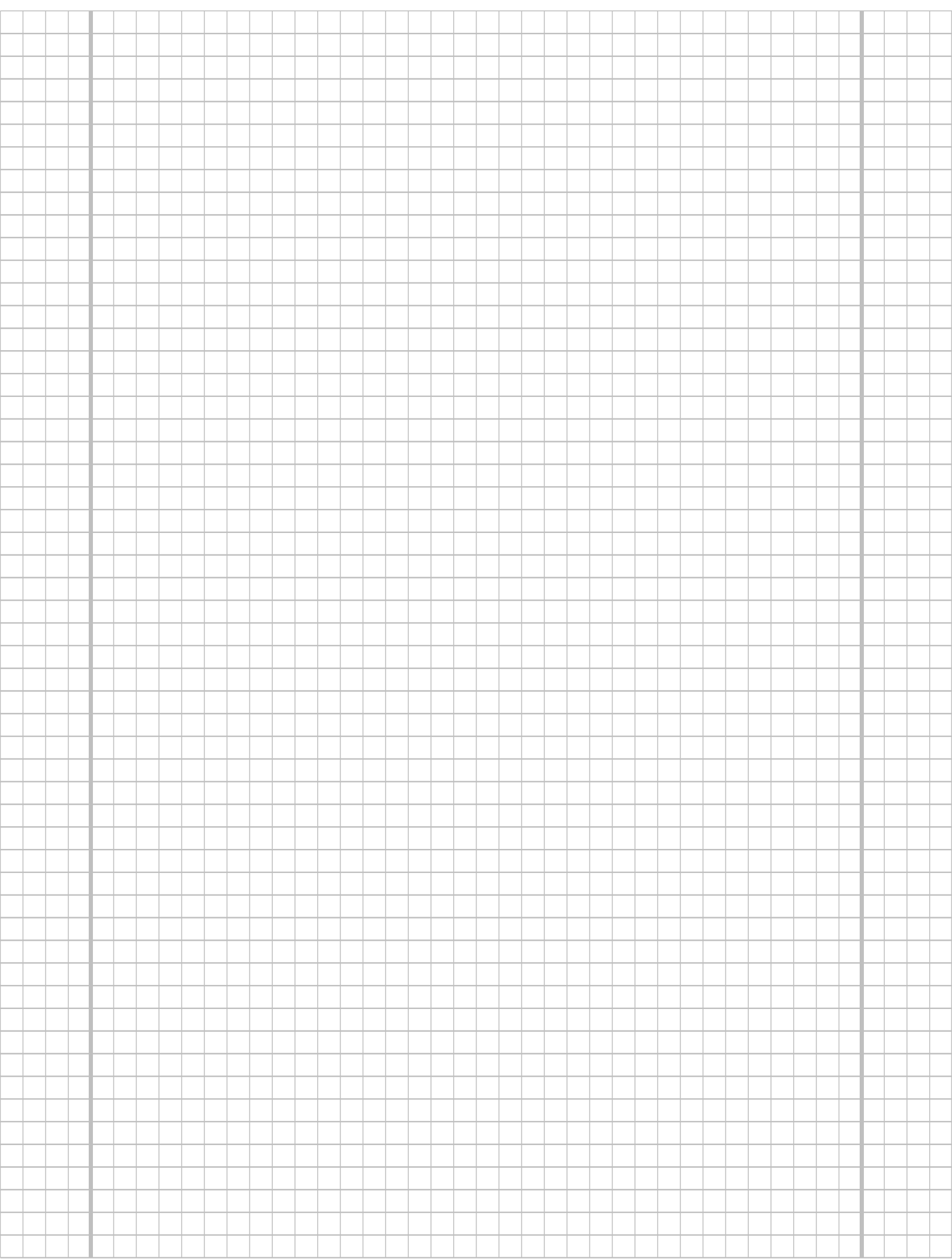




Tabelle zur Standardnormalverteilung

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

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

	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