

**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:

SS 2019

(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		14	
4		16	
5		6	
6		9	
7		5	
8		18	
9		16	
10		8	
Summe		120	

bestanden

Note: _____

nicht bestanden

Unterschrift der Prüferin/des Prüfers: _____

KLAUSURHEFT ZUR
 BACHELOR-PRÜFUNG
 DESKRIPTIVE STATISTIK UND WAHRSCHEINLICHKEITSRECHNUNG
 SOMMERSEMESTER 2019

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 + 14 + 16 + 6 + 9 + 5 + 18 + 16 + 8) 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!).

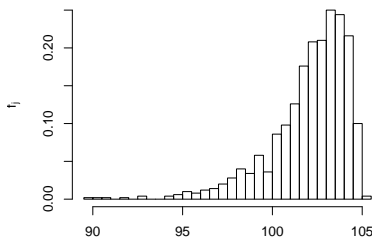
- | | wahr | falsch |
|---|--------------------------|--------------------------|
| 1. Numerische Merkmale sind stets kardinalskaliert. | <input type="checkbox"/> | <input type="checkbox"/> |
| 2. Bei empirischen Verteilungsfunktionen zu nicht klassierten Merkmalen X gibt es zu jeder Sprungstelle x (mindestens) einen passenden Eintrag $x_i = x$ in der zugehörigen Urliste x_1, \dots, x_n des Merkmals X . | <input type="checkbox"/> | <input type="checkbox"/> |
| 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]. | <input type="checkbox"/> | <input type="checkbox"/> |
| 4. Es seien (Ω, \mathcal{F}, P) ein Wahrscheinlichkeitsraum, $A, B, C \in \mathcal{F}$ drei Ereignisse mit $P(C) > 0$. Dann gilt stets:
$P(A C) + P(B C) \leq 2 \cdot P(A \cap B C)$ | <input type="checkbox"/> | <input type="checkbox"/> |
| 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 Untersucher bei positiver Diagnose dieses Tests mit einer Wahrscheinlichkeit von 97% tatsächlich erkrankt ist. | <input type="checkbox"/> | <input type="checkbox"/> |
| 6. Exponentialverteilte Zufallsvariablen sind stets leptokurtisch verteilt. | <input type="checkbox"/> | <input type="checkbox"/> |
| 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%. | <input type="checkbox"/> | <input type="checkbox"/> |
| 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 $\text{Var}(X_1 + X_2 - X_3) = 40$. | <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 Ränge $\text{rg}(X)_1, \dots, \text{rg}(X)_8$ zur (der Einfachheit halber sortierten) Urliste

2, 3, 5, 5, 9, 9, 11, 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?

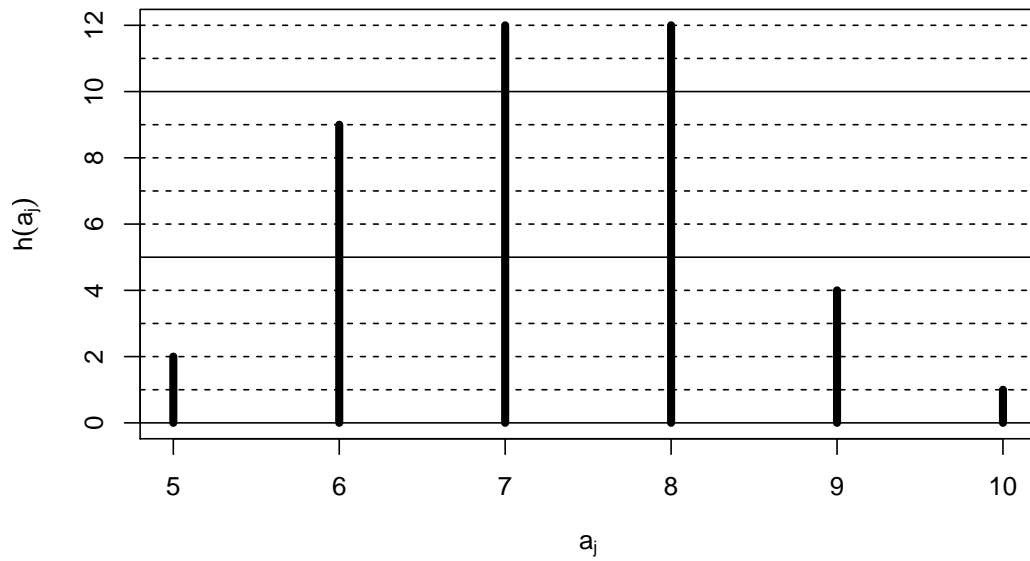
- (a) Zunahme um 10%
- (b) Zunahme um 5%
- (c) Zunahme um 0.5%
- (d) Keine Veränderung

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

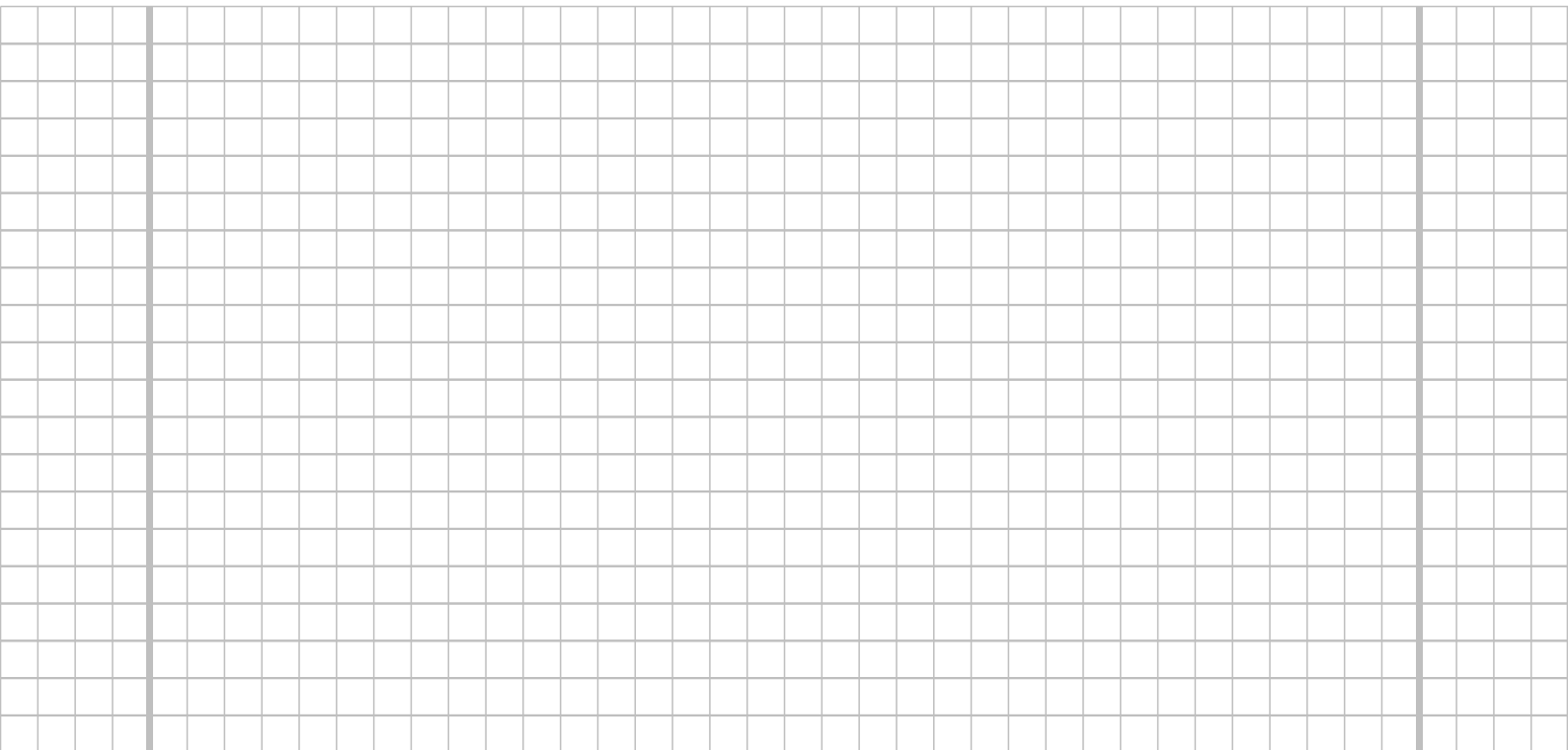
- (a) 24
- (b) 12.
- (c) 10.
- (d) 8.

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

Zu einem erhobenen Merkmal X sei das folgende Stabdiagramm gegeben:



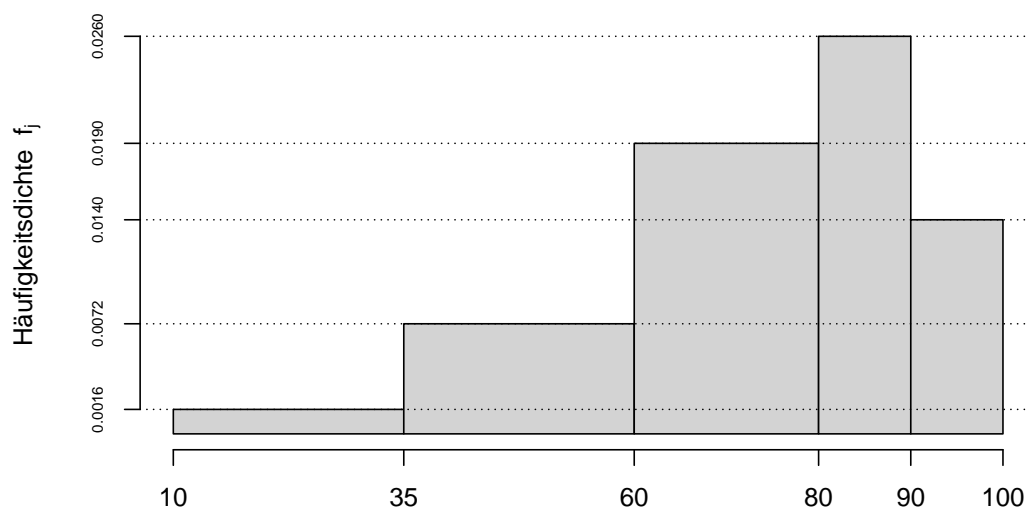
- (a) Erstellen Sie eine Tabelle der absoluten und relativen Häufigkeiten.
- (b) Berechnen Sie den arithmetischen Mittelwert und die empirische Standardabweichung des Merkmals X .
- (c) Geben Sie die empirische Verteilungsfunktion des Merkmals X an.
- (d) Wie groß ist der Anteil der Urlisteneinträge, die Werte von weniger als 9 annehmen?
- (e) Berechnen Sie einen Median des Merkmals X .



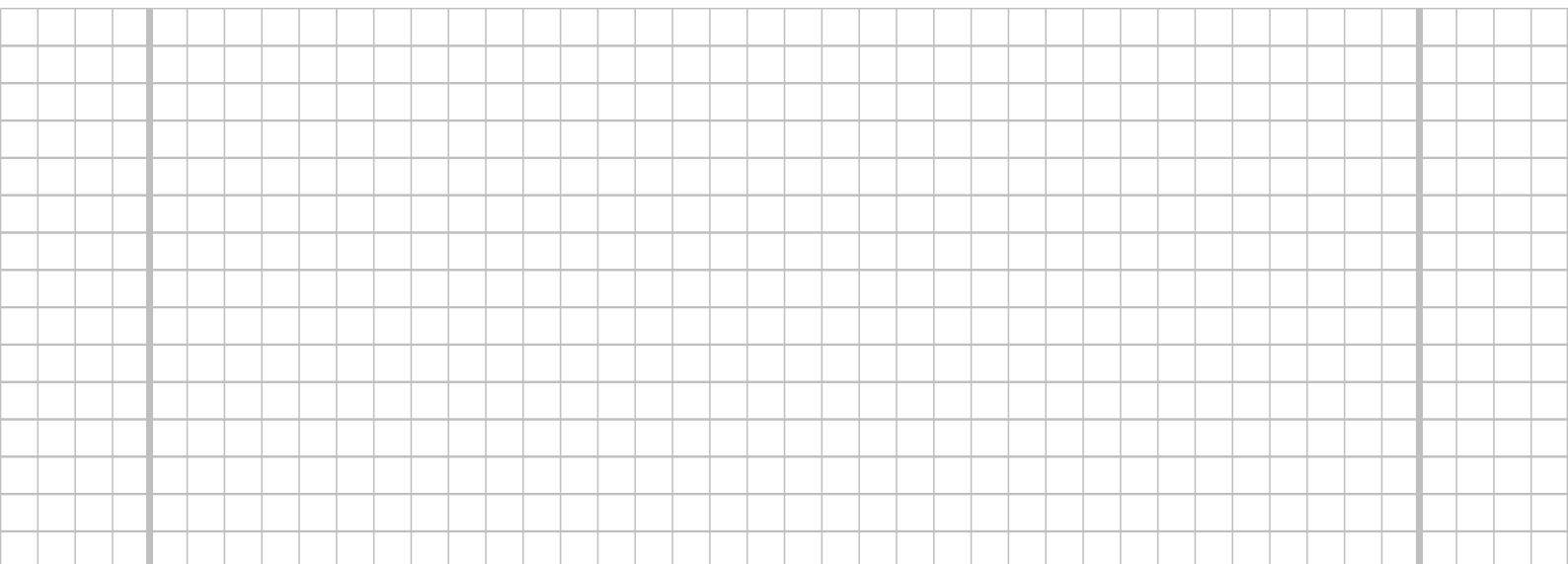


Aufgabe 4 (7 + 4 + 3 + 2 = 16 Punkte)

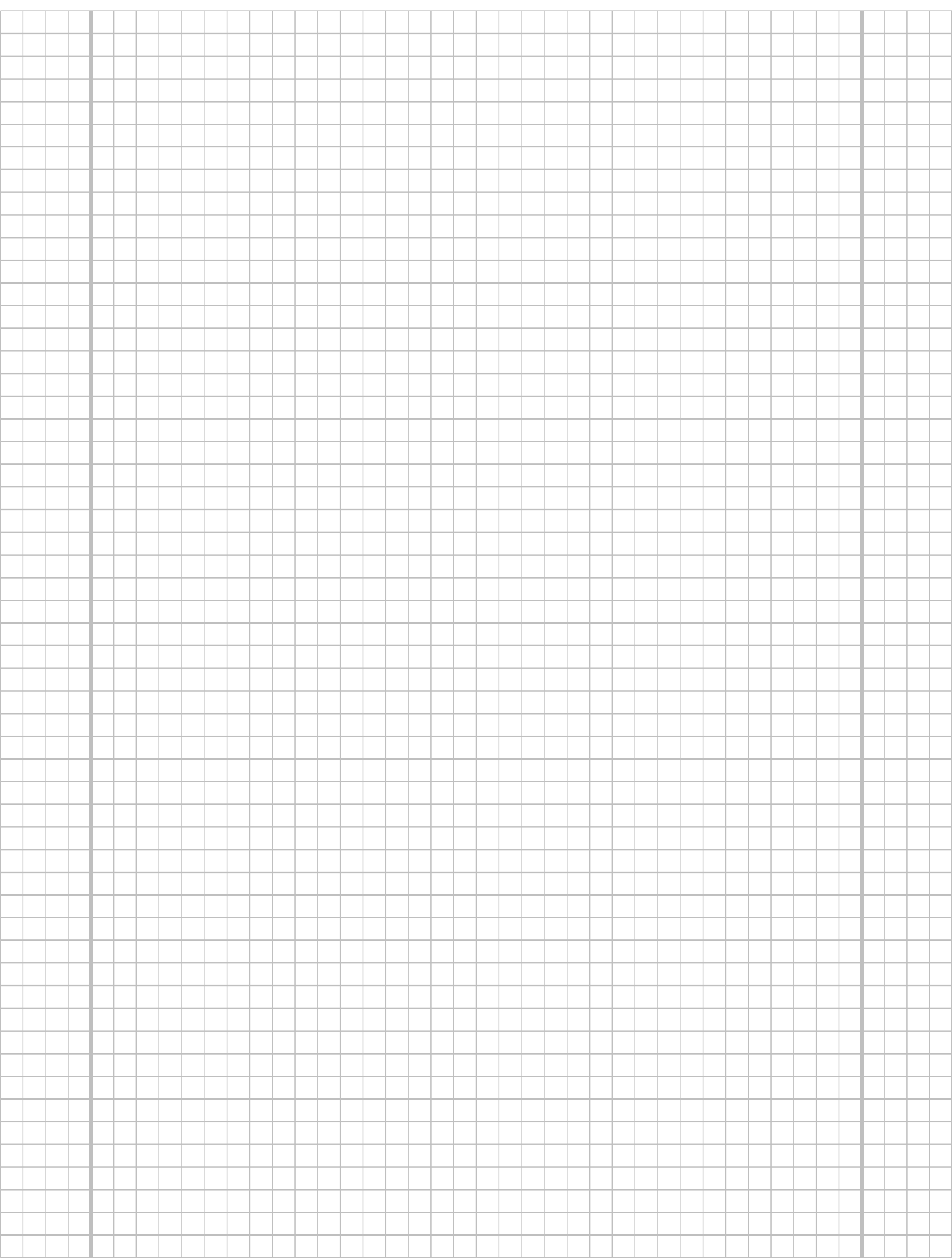
Gegeben sei das folgende Histogramm zur Klassierung einer Urliste der Länge $n = 50$:



- (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?

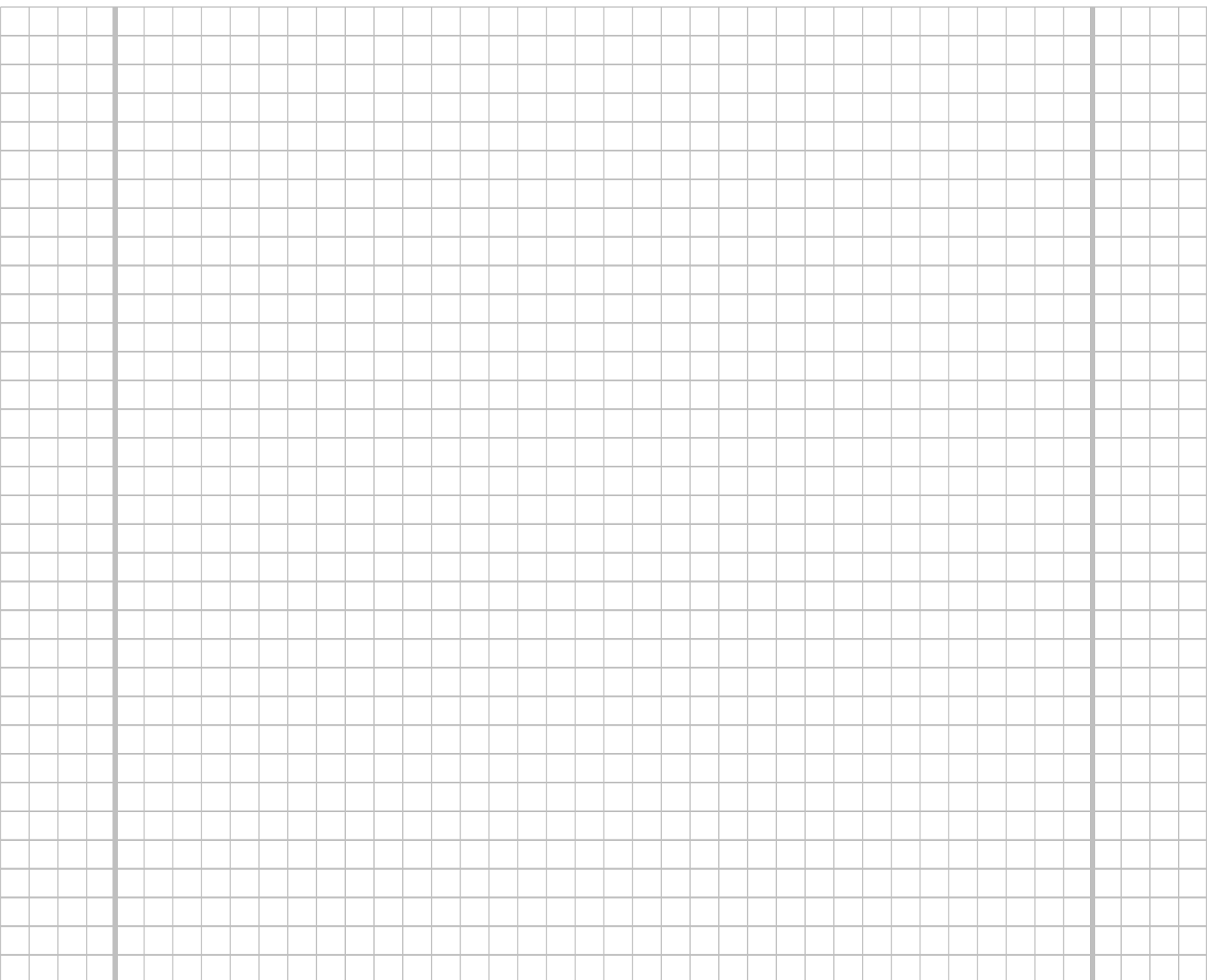




Aufgabe 6 (5 + 2 + 2 = 9 Punkte)

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!

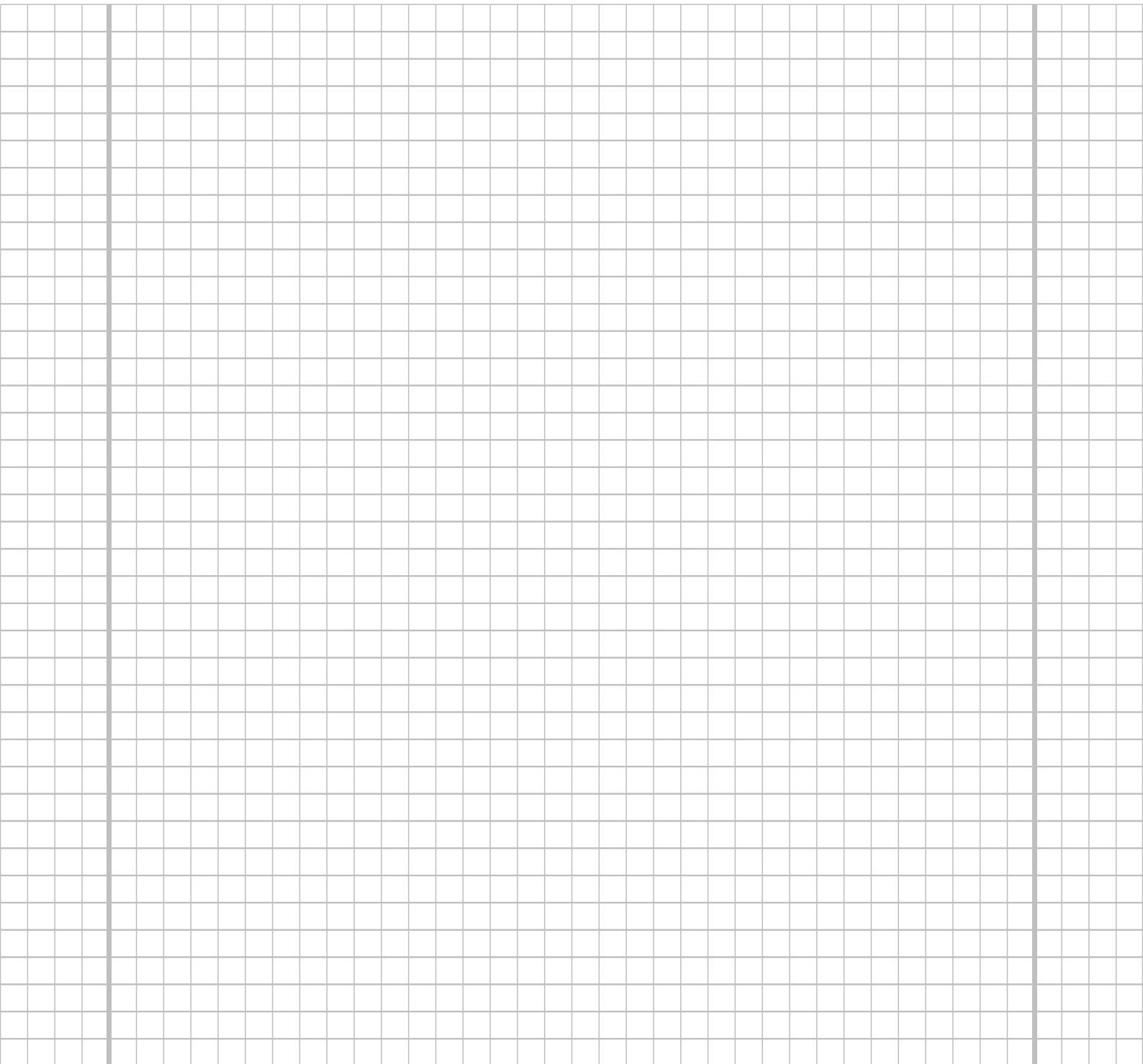




Aufgabe 7 (2 + 1 + 2 = 5 Punkte)

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.



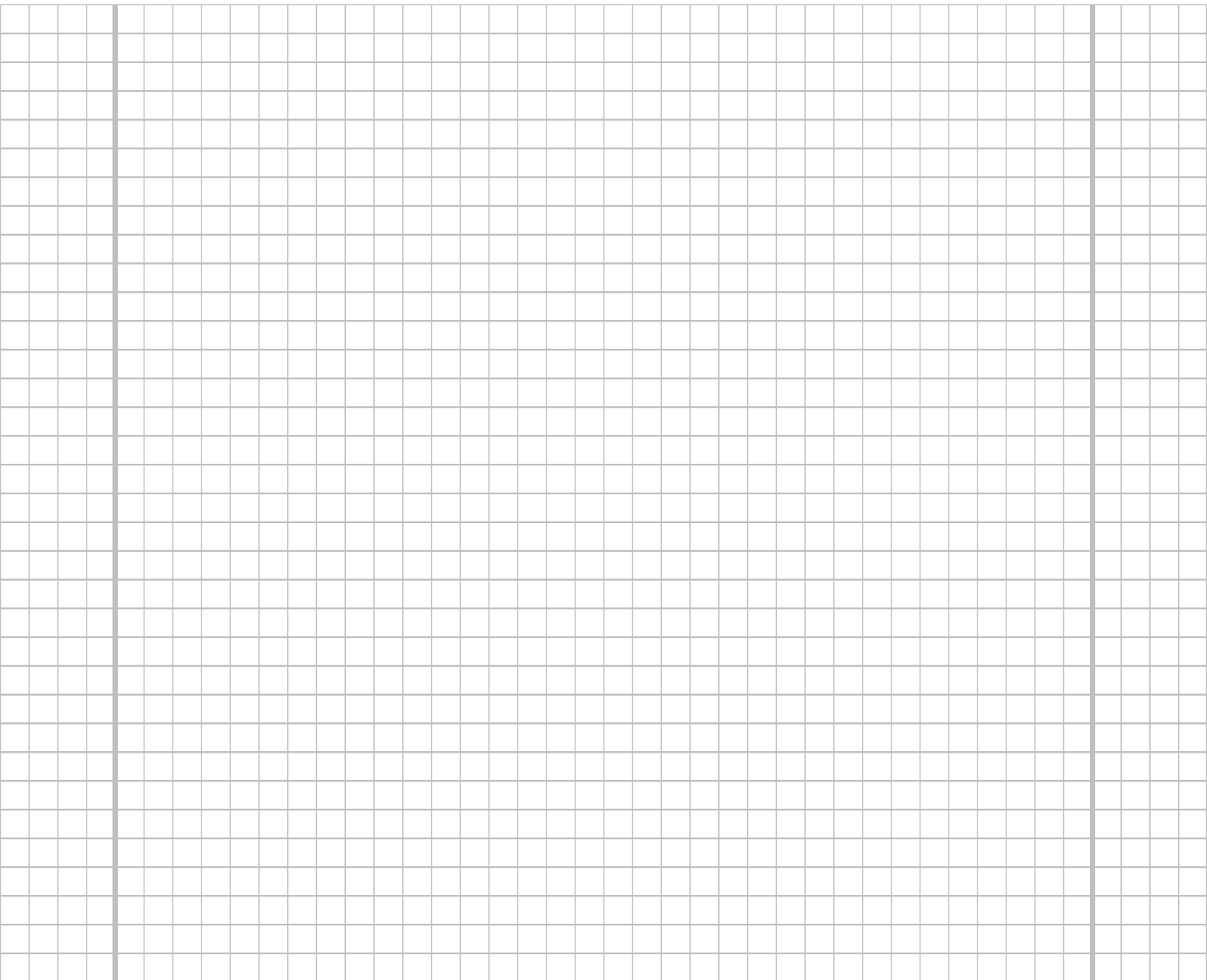


Aufgabe 8 (5 + 2 + 6 + 1 + 4 = 18 Punkte)

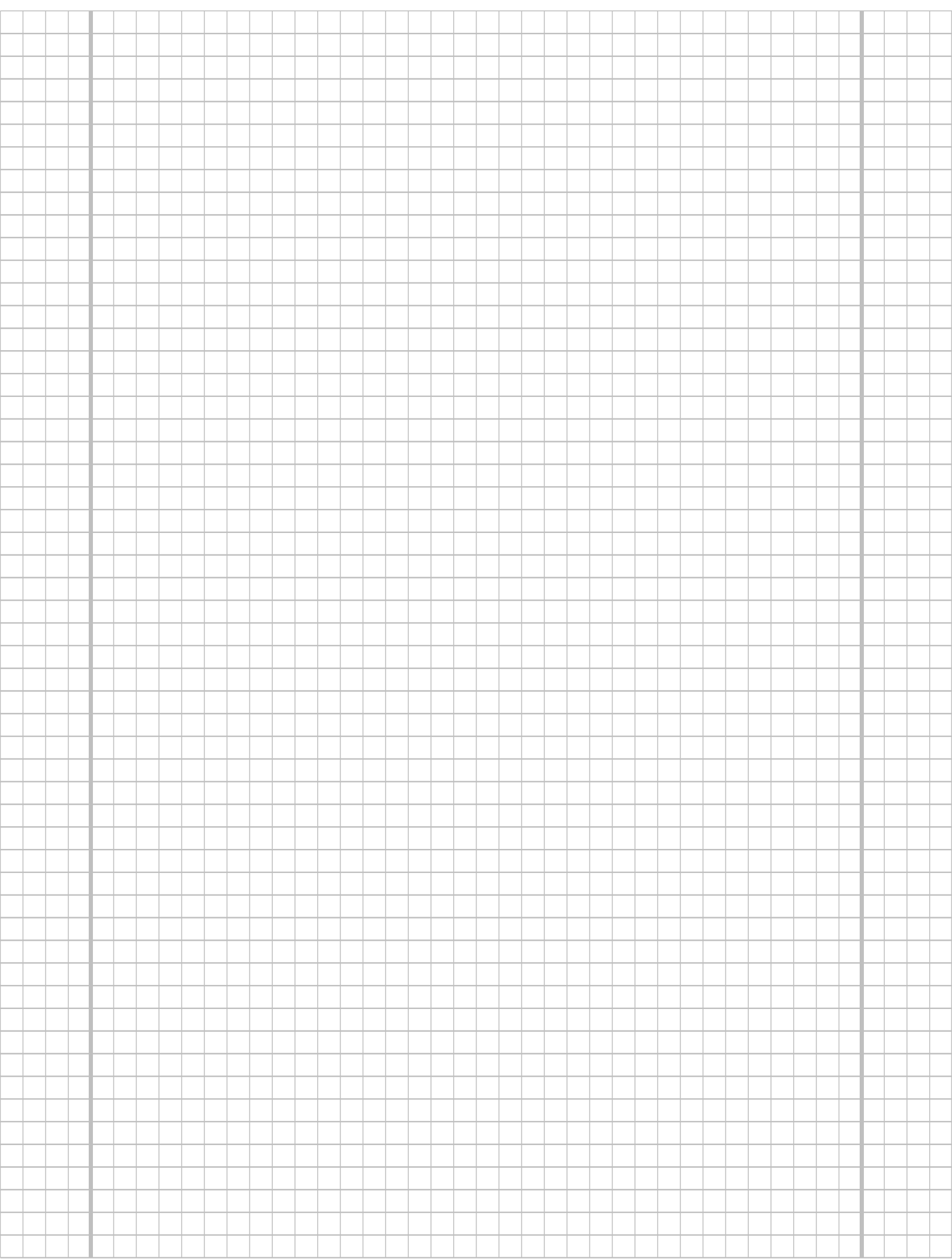
Die Verteilung einer stetigen Zufallsvariablen X sei durch die folgende Dichtefunktion gegeben:

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

- (a) Berechnen Sie die Verteilungsfunktion F_X von X .
- (b) Berechnen Sie $P(\{X > \frac{1}{2}\})$ und $P(\{\frac{1}{2} \leq X \leq \frac{3}{2}\})$.
- (c) Bestimmen Sie den Erwartungswert $E(X)$.
- (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	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{36}$

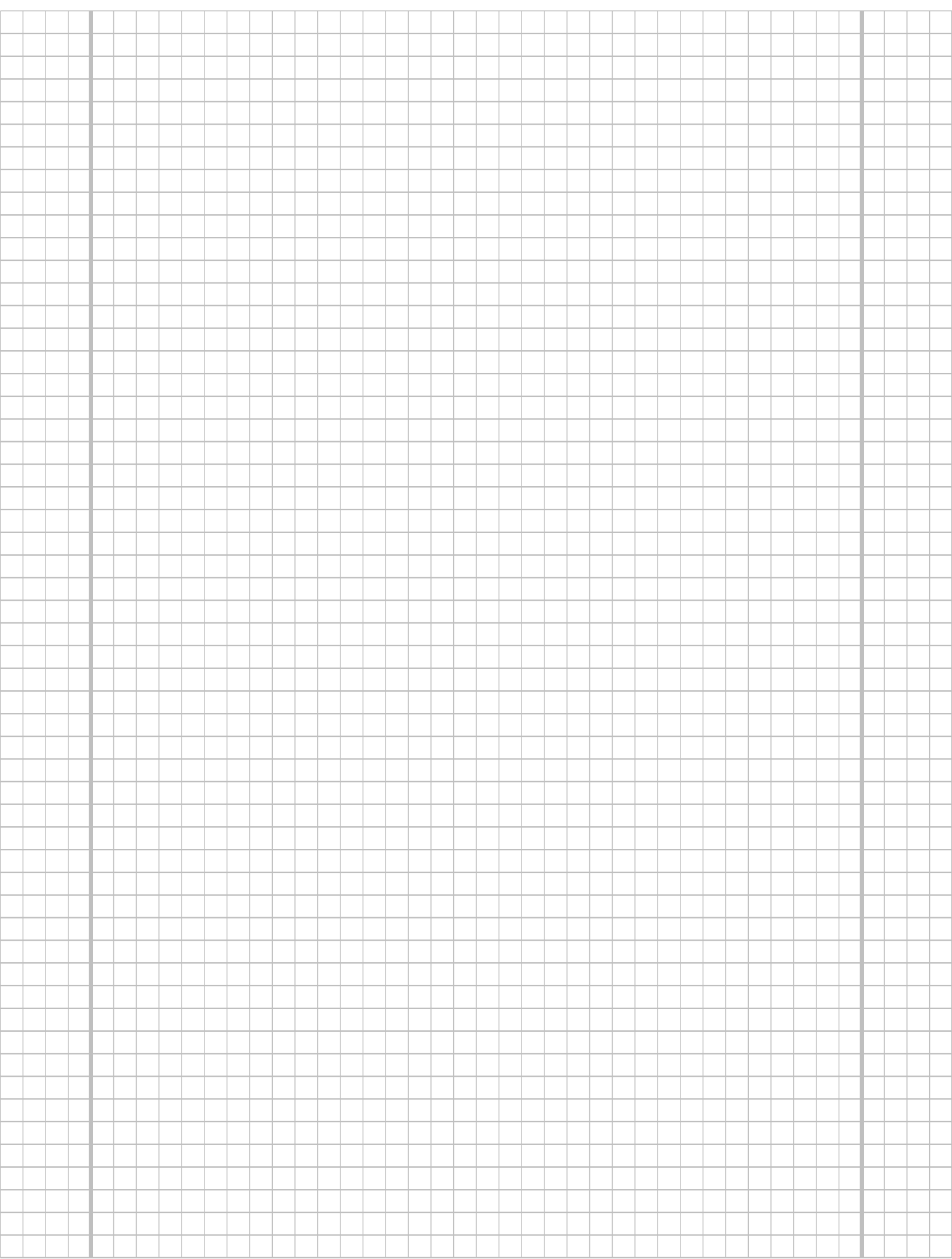
Bestimmen Sie $E(X)$, $E(Y)$, $\text{Var}(X)$, $\text{Var}(Y)$, $\text{Cov}(X, Y)$ sowie $\text{Korr}(X, Y)$.

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 $E(3X - 2Y)$ sowie $\text{Var}(3X - 2Y)$.









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 26!

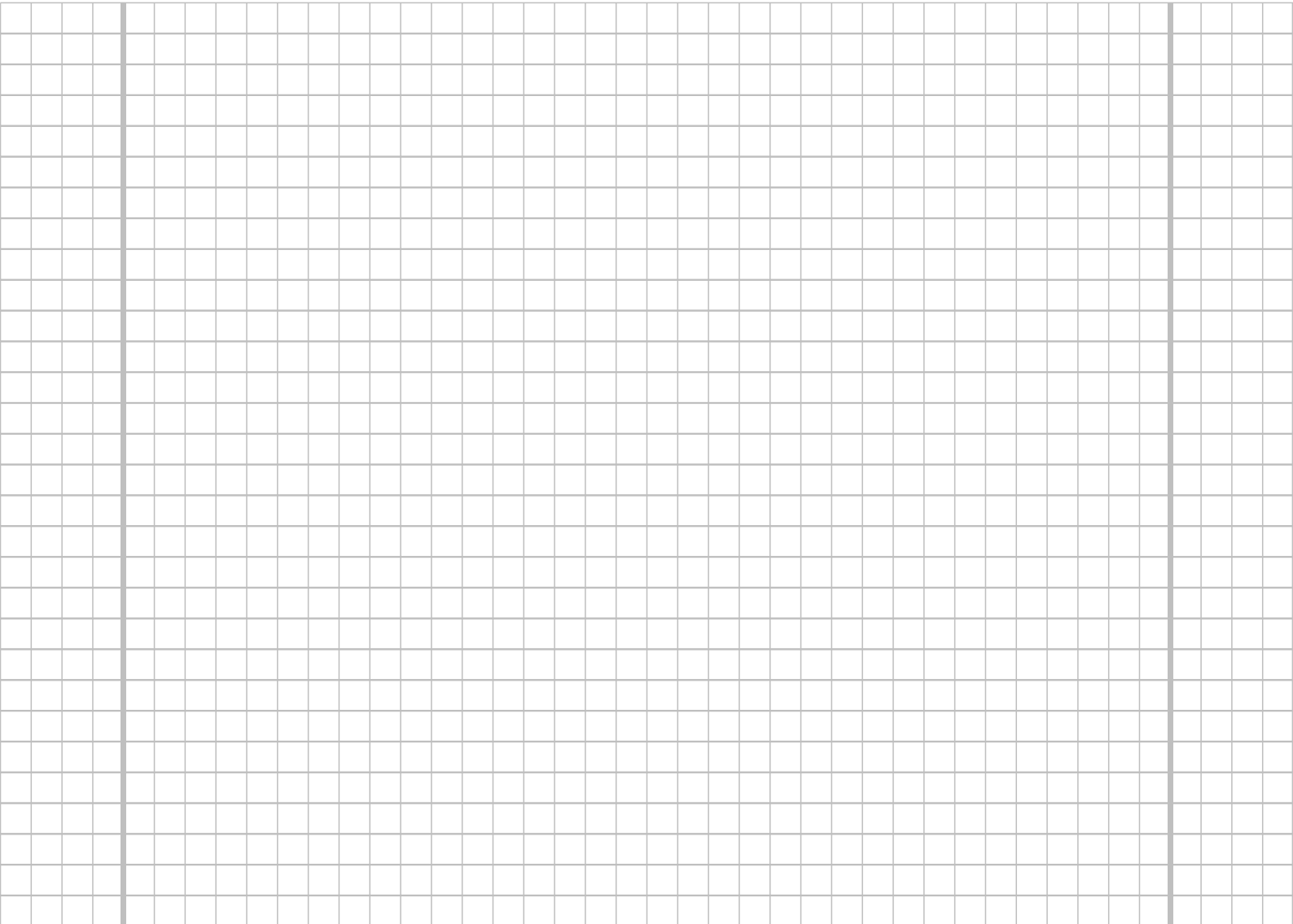




Tabelle zur Standardnormalverteilung

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

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

	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