

**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:** SS 2020 (z. B. WS 2015/2016, SS 2016)  
(Prüfungen im Februar/April = WS; Prüfungen im August/Oktober = 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		16	
2		12	
3		12	
4		16	
5		8	
6		12	
7		18	
8		17	
9		9	
<i>Summe</i>		120	

bestanden

Note: \_\_\_\_\_

nicht bestanden

Unterschrift der Prüferin/des Prüfers: \_\_\_\_\_

KLAUSURHEFT ZUR  
BACHELOR-PRÜFUNG  
DESKRIPTIVE STATISTIK UND WAHRSCHEINLICHKEITSRECHNUNG  
SOMMERSEMESTER 2020

PD 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 ( $= 16 + 12 + 12 + 16 + 8 + 12 + 18 + 17 + 9$ ) erreichbar.
- Als Hilfsmittel sind zugelassen: Taschenrechner (auch mit Grafikfähigkeit), 2 selbsterstellte DIN-A4 Blätter bzw. 4 selbsterstellte (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)	$\Sigma$
1		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
2		<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
3					<input type="checkbox"/>	
4					<input type="checkbox"/>	
5				<input type="checkbox"/>	<input type="checkbox"/>	
6					<input type="checkbox"/>	
7						
8					<input type="checkbox"/>	
9				<input type="checkbox"/>	<input type="checkbox"/>	
$\Sigma$						

**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:

$$P(A) = P(B) \quad \Rightarrow \quad P(A|C) = P(B|C)$$

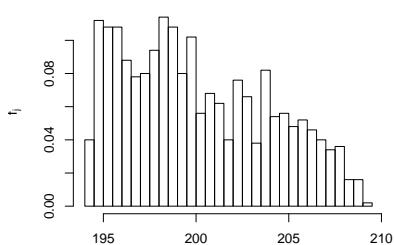
6. Werden zwei faire (sechsseitige) Würfel gleichzeitig geworfen, so erhält man mit einer Wahrscheinlichkeit von  $\frac{15}{21}$  zwei unterschiedliche Augenzahlen.
7. Die Summe von vier stochastisch unabhängigen  $N(16, 2^2)$ -verteilten Zufallsvariablen ist  $N(64, 4^2)$ -verteilt.
8. Ist die Kovarianz 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 rechtssteil
- (b) leptokurtisch und linkssteil
- (c) platykurtisch und rechtssteil
- (d) platykurtisch und linkssteil

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 Würfels 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, 0, 0, 0, 0, 0, 2, 2, 2, 3, 3, 3, 3, 4, 4, 4, 4, 4, 4, 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:

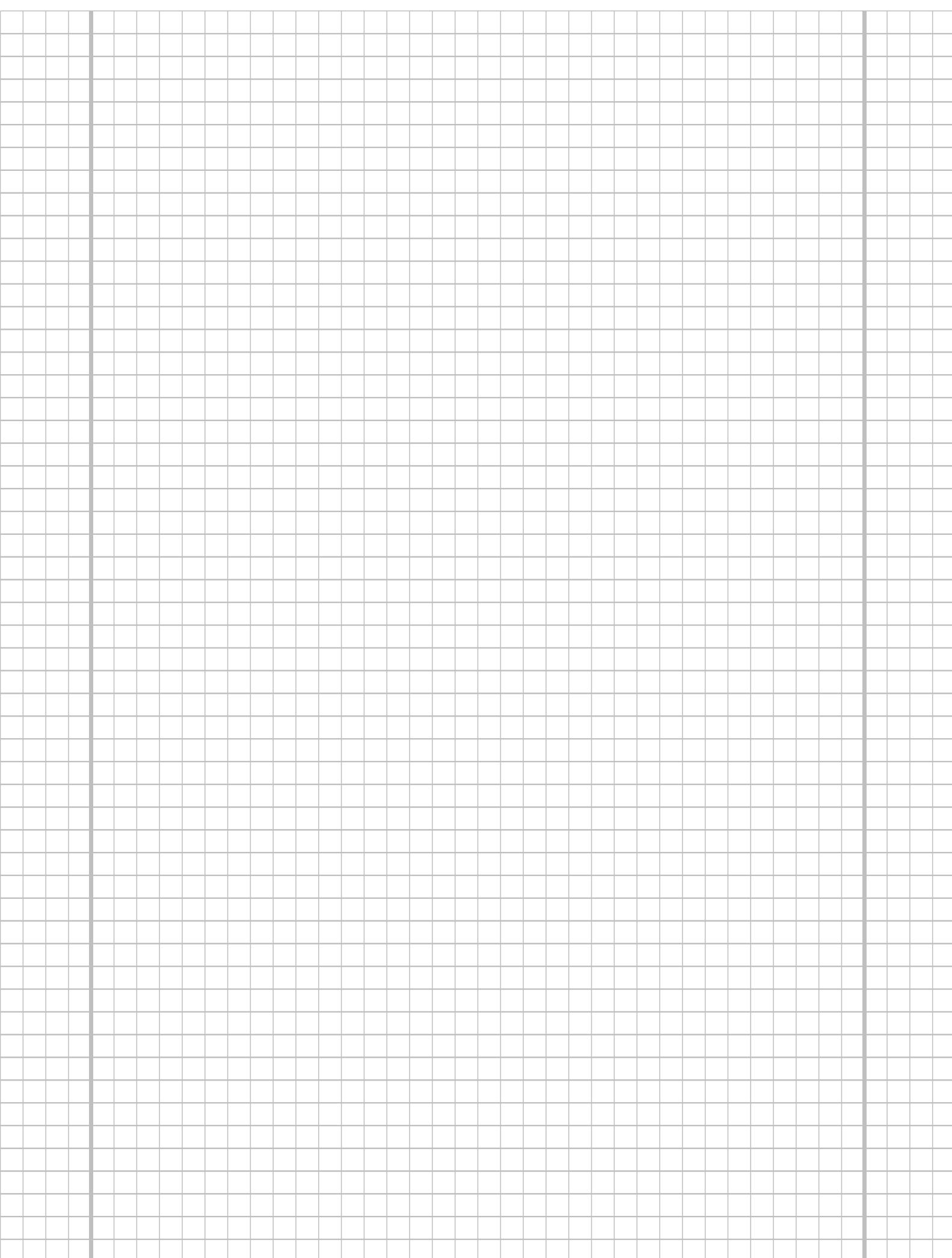
6.12, 7.58, 7.76, 8.09, 9.18, 9.33, 9.94, 9.94, 10.57, 10.68, 11.00, 11.19, 11.28, 11.33, 11.48, 11.59, 11.60, 13.46, 13.72, 13.83, 14.15, 14.54, 14.96, 15.82, 16.11, 16.38, 16.95, 17.13, 17.67, 18.12, 18.50, 18.95, 19.08, 20.15, 22.61, 23.12, 23.49, 26.06, 27.12, 27.87

- (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  $P(\{X > \frac{3}{2}\})$  und  $P(\{\frac{3}{2} \leq X \leq 4\})$ .
- (d) Bestimmen Sie den Erwartungswert  $E(X)$ .
- (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	$p_i$
-2	$\frac{1}{25}$	$\frac{1}{15}$	$\frac{1}{12}$	
1	$\frac{4}{25}$	$\frac{1}{15}$	$\frac{11}{60}$	
4	$\frac{1}{5}$	$\frac{1}{15}$	$\frac{2}{15}$	
$p_{\cdot j}$				

- Ergänzen Sie die obige Tabelle (in den vorgesehenen Feldern) um ihre Randverteilungen.
- 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.
- Berechnen Sie  $E(X)$ ,  $E(Y)$ ,  $\text{Var}(X)$ ,  $\text{Var}(Y)$ ,  $\text{Cov}(X, Y)$  sowie  $\text{Korr}(X, Y)$ .
- Berechnen Sie  $E(5X - 3Y)$  sowie  $\text{Var}(5X - 3Y)$ .







**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  $E(Y)$  sowie die Varianz  $\text{Var}(Y)$  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 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