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 2019/20 (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+2		28	
3		10	
4		19	
5		6	
6		10	
7		18	
8		6	
9		15	
10		8	
Summe		120	

◯ *bestanden* *Note:* _____

◯ *nicht bestanden* *Unterschrift der Prüferin/des Prüfers:* ________________________

Klausurheft zur

Bachelor-Prüfung
Deskriptive Statistik und Wahrscheinlichkeitsrechnung
Wintersemester 2019/20

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 + 10 + 19 + 6 + 10 + 18 + 6 + 15 + 8) 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 den Seiten 1–25 berücksichtigt. Die Zusammenfassung der Aufgabenstellungen am Ende des Klausurhefts darf abgetrennt und muss nicht abgegeben werden.
- 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)

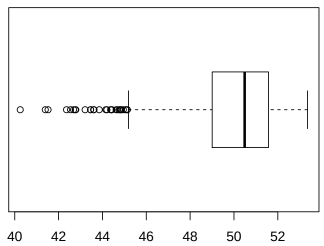
Markieren Sie jeweils mit einem Kreuz pro Aussage im betreffenden Kästchen, ob die unten stehenden Aussagen wahr oder falsch sind.
Richtige Antworten geben 2 Punkte, falsche Antworten und nicht bearbeitete Aussagen 0 Punkte (Aussagen mit zwei Kreuzen zählen als nicht bearbeitet!).

	wahr	falsch
1. 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: $$P(A\|C) + P(B\|C) \leq 2 \cdot P(A \cup B\|C)$$	☐	☐
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 $\mathrm{E}\,(2X + 4Y)$ sowie $\mathrm{Var}\,(2X + 4Y)$.

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 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