Wirtschaftswissenschaftliches Prüfungsamt

Bachelor-Prüfung
Deskriptive Statistik und Wahrscheinlichkeitsrechnung
Wintersemester 2010/11

Aufgabenstellung und Ergebnisse

Dr. Martin Becker

Hinweise für die Klausurteilnehmer

- Kleben Sie bitte **sofort** Ihr Namensschild auf obige Markierung!
- 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 + 13 + 16 + 8 + 8 + 8 + 14 + 15 + 10) erreichbar.
- Als Hilfsmittel sind zugelassen: Taschenrechner (ohne 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.

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

Aufgabe 1 (16 Punkte)

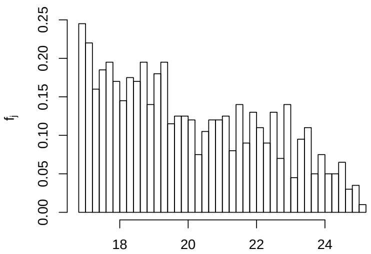
Markieren Sie jeweils mit einem Kreuz pro Aussage im betreffenden Kästchen, ob die unten stehenden Aussagen wahr oder falsch sind.
Richtige Antworten geben +2 Punkte, falsche Antworten −1 Punkt, nicht bearbeitete Aussagen 0 Punkte (Aussagen mit zwei Kreuzen zählen als nicht bearbeitet!).
Die Aufgabe wird insgesamt mit mindestens 0 Punkten bewertet!

Nr.	Aussage	wahr	falsch
1.	Quantitative Merkmale sind immer kardinalskaliert.	■	□
2.	Empirische Verteilungsfunktionen F sind stets stetige Funktionen.	□	■
3.	Jeder Median eines kardinalskalierten Merkmals kommt mindestens einmal in der Urliste vor.	□	■
4.	Sei (X, Y) ein zweidimensionales Merkmal, X und Y seien kardinalskaliert. Dann gilt für den Korrelationskoeffizienten $r_{X,Y}$ nach Bravais-Pearson stets $-1 \leq r_{X,Y} \leq 1$.	■	□
5.	Es seien $(\Omega, \mathcal{F}, P)$ ein Wahrscheinlichkeitsraum, $A, B, C \in \mathcal{F}$ drei Ereignisse mit $P(C) > 0$. Dann gilt: $P(A\|C) > P(B\|C) \quad \Rightarrow \quad P(A) > P(B)$	□	■
6.	Diskrete Zufallsvariablen haben immer endlich viele Trägerpunkte.	□	■
7.	Es sei f_X eine Dichtefunktion zu einer stetigen Zufallsvariablen X. Dann gilt stets: $\int_{-\infty}^{\infty} f_X(x)dx = 1$	■	□
8.	Sind X und Y unkorrelierte Zufallsvariablen, so gilt $\mathrm{E}(X+Y) = \mathrm{E}(X) + \mathrm{E}(Y) - \mathrm{E}(X) \cdot \mathrm{E}(Y)$.	□	■

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 −1 Punkt, nicht bearbeitete Aufgabenteile 0 Punkte (Aufgabenteile mit mehr als einem Kreuz zählen als nicht bearbeitet!).
Die Aufgabe wird insgesamt mit mindestens 0 Punkten bewertet!

1. Kreuzen Sie an, auf welche Merkmalseigenschaften das folgende Histogramm *am ehesten* hindeutet:



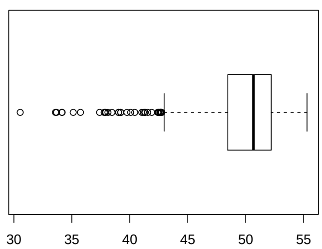
(a) leptokurtisch und linksschief ☐

(b) leptokurtisch und rechtsschief ☐

(c) platykurtisch und linksschief ☐

(d) platykurtisch und rechtsschief ■

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



(a) leptokurtisch und linksschief ■

(b) leptokurtisch und rechtsschief ☐

(c) platykurtisch und linksschief ☐

(d) platykurtisch und rechtsschief ☐

3. 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) Nur Spearmanscher Rangkorrelationskoeffizient ☐

(b) Nur korrigierter Pearsonscher Kontingenzkoeffizient ■

(c) Pearsonscher Korrelationskoeffizient und korrigierter Pearsonscher Kontingenzkoeffizient ☐

(d) Spearmanscher Rangkorrelationskoeffizient und korrigierter Pearsonscher Kontingenzkoeffizient ☐

4. Die Anzahl (verschiedener) r-elementiger Teilmengen, die man aus einer n-elementigen Menge bilden kann, ist (für $0 \leq r \leq n$) gegeben durch

(a) n^r ☐

(b) r^n ☐

(c) $\dfrac{n!}{r!}$ ☐

(d) $\dbinom{n}{r}$ ■

Aufgabe 3 (2 + 4 + 4 + 3 = 13 Punkte)

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

$$F(x) = \begin{cases} 0.000 & \text{für} \quad x < 2 \\ 0.025 & \text{für} \quad 2 \leq x < 4 \\ 0.175 & \text{für} \quad 4 \leq x < 6 \\ 0.375 & \text{für} \quad 6 \leq x < 8 \\ 0.725 & \text{für} \quad 8 \leq x < 10 \\ 0.950 & \text{für} \quad 10 \leq x < 12 \\ 1.000 & \text{für} \quad x \geq 12 \end{cases}$$

Außerdem sei die Länge der Urliste $n = 40$ bekannt.

(a) Geben Sie die Menge A der Merkmalsausprägungen an.

(b) Erstellen Sie eine Tabelle der relativen und absoluten Häufigkeiten.

(c) Berechnen Sie den arithmetischen Mittelwert und die empirische Varianz des Merkmals X.

(d) Berechnen Sie das untere und obere Quartil des Merkmals X sowie den zugehörigen Interquartilsabstand.

Ergebnisse:

(a) $A = \{2, 4, 6, 8, 10, 12\}$

(b) Häufigkeitstabelle (mit relativen und absoluten Häufigkeiten):

a_j	2	4	6	8	10	12	Σ
$r(a_j)$	0.025	0.150	0.200	0.350	0.225	0.050	1.000
$h(a_j)$	1	6	8	14	9	2	40

(c) $\overline{x} = 7.5$, $s^2 = 5.55$.

(d) $x_{0.25} = 6$, $x_{0.75} = 10$, IQA: 4

Aufgabe 4 $(6 + 3 + 4 + 3 = 16$ Punkte)

Zu einem kardinalskalierten Merkmal sei die folgende (zur einfacheren Bearbeitung der Aufgabe bereits sortierte) Urliste der Länge $n = 25$ gegeben:

3.47, 5.75, 7.00, 10.01, 11.89, 12.76, 15.64, 16.07, 17.74, 17.76, 19.56, 19.72, 23.31, 23.45, 23.81, 25.19, 25.71, 27.94, 30.44, 31.00, 32.02, 33.73, 34.28, 34.36, 36.89

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

$$K_1 = (0, 10],\ K_2 = (10, 20],\ K_3 = (20, 30],\ K_4 = (30, 40]$$

durch. Geben Sie insbesondere die jeweiligen Klassenbreiten, Klassenmitten, absoluten und relativen Klassenhäufigkeiten, Häufigkeitsdichten sowie die Werte der empirischen Verteilungsfunktion an den Klassengrenzen an.

(b) 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 21.58?

(c) Stellen Sie die (approximative) empirische Verteilungsfunktion des Merkmals aus der Klassierung der Daten auf.

(d) Bestimmen Sie (aus der Urliste) die Anzahl von Merkmalswerten zwischen 15 und 30. Welche Näherung für diese Anzahl können Sie aus der in Teil (c) aufgestellten empirischen Verteilungsfunktion berechnen?

Ergebnisse:

(a) Klassierung:

Nr. j	Klasse $K_j = (k_{j-1}, k_j]$	Klassen-breite b_j	Klassen-mitte m_j	absolute Häufigkeit h_j	relative Häufigkeit $r_j = \frac{h_j}{n}$	Häufigkeits-dichte $f_j = \frac{r_j}{b_j}$	Verteilungs-funktion $F(k_j)$
1	$(0, 10]$	10	5	3	$\frac{3}{25}$	0.0120	0.12
2	$(10, 20]$	10	15	9	$\frac{9}{25}$	0.0360	0.48
3	$(20, 30]$	10	25	6	$\frac{6}{25}$	0.0240	0.72
4	$(30, 40]$	10	35	7	$\frac{7}{25}$	0.0280	1.00

(b) Mittelwert (näherungsweise): 21.8
relative Abweichung vom exakten Wert: 0.0102 bzw. 1.02%

(c) (Approximative) Verteilungsfunktion:

$$F(x) = \begin{cases} 0 & \text{für } x < 0 \\ 0.012 \cdot x & \text{für } 0 \le x < 10 \\ 0.12 + 0.036 \cdot (x - 10) & \text{für } 10 \le x < 20 \\ 0.48 + 0.024 \cdot (x - 20) & \text{für } 20 \le x < 30 \\ 0.72 + 0.028 \cdot (x - 30) & \text{für } 30 \le x < 40 \\ 1 & \text{für } x \ge 40 \end{cases}$$

(d) Anzahl (aus Urliste): 12
Mit emp. Verteilungsfunktion genäherte Anzahl: 10.5

Aufgabe 5 (2 + 2 + 2 + 2 = 8 Punkte)

Eine Urne enthält 80 Kugeln, von denen 10 blau und kariert, 5 grün und kariert, 40 blau und gestreift sowie 25 grün und gestreift sind.
Wie groß ist die Wahrscheinlichkeit dafür, dass eine zufällig aus der Urne entnommene Kugel

(a) grün und gestreift ist?

(b) blau ist?

(c) kariert ist, wenn man weiß, dass sie blau ist?

(d) Wie groß ist die Wahrscheinlichkeit, bei dreimaligem zufälligen Ziehen *mit Zurücklegen* zunächst eine blaue, dann eine grüne und zuletzt eine karierte Kugel zu ziehen?

Ergebnisse:

(a) $\frac{5}{16}$

(b) $\frac{5}{8}$

(c) $\frac{1}{5}$

(d) $\frac{45}{1024}$

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

In einem Autozulieferbetrieb werden 40% der Dieseleinspritzpumpen während der Frühschicht, 40% der Dieseleinspritzpumpen während der Spätschicht und die restlichen 20% der Dieseleinspritzpumpen während der Nachtschicht produziert. Es ist bekannt, dass der Anteil defekter Dieseleinspritzpumpen für die Frühschicht 2%, für die Spätschicht 3% und für die Nachtschicht 8% beträgt.

(a) Wie groß ist die Wahrscheinlichkeit, dass eine in diesem Autozulieferbetrieb produzierte Dieseleinspritzpumpe defekt ist?

(b) Wie groß ist die Wahrscheinlichkeit, dass eine Dieseleinspritzpumpe, von der bekannt ist, dass sie defekt ist, während der Frühschicht produziert wurde?

(c) Sind die Ereignisse „Dieseleinspritzpumpe wurde in der Frühschicht produziert“ bzw. „Dieseleinspritzpumpe ist defekt“ stochastisch unabhängig? Begründen Sie Ihre Antwort!

Ergebnisse:

(a) 0.036

(b) $0.\overline{2}$

(c) Nein.

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

In einer Multiple-Choice-Aufgabe mit 4 Aufgabenteilen mit jeweils 4 Antwortmöglichkeiten ist jeweils genau eine eine Antwortmöglichkeit richtig. Für korrekte Antworten werden +3 Punkte, für falsche Antworten −1 Punkt vergeben. Nehmen Sie an, dass ein unvorbereiteter Prüfling bei jedem der 4 Aufgabenteile rein zufällig eine der 4 Antwortmöglichkeiten ankreuzt.

(a) Sei X die Zufallsvariable, die die **Anzahl der korrekten Antworten** des Prüflings beschreibt. Wie ist X verteilt? Wie groß ist der Erwartungswert von X?

(b) Die Anzahl der falschen Antworten ist offensichtlich gleich $4 - X$. Bestimmen Sie die bei der MC-Aufgabe **erreichte Punktzahl** Y in Abhängigkeit von X. Wie groß ist der Erwartungswert von Y?

(c) Geben Sie die Wahrscheinlichkeit, eine negative **Punktzahl** zu erhalten, an.

Ergebnisse:

(a) $X \sim B(4, 0.25)$, $\mathrm{E}(X) = 1$.

(b) $Y = 4X - 4$, $\mathrm{E}(Y) = 0$.

(c) $P\{Y < 0\} = 0.3164062$.

Aufgabe 8 (4 + 3 + 6 + 1 = 14 Punkte)

Die Verteilung der stetigen Zufallsvariablen X sei gegeben durch die Dichtefunktion

$$f_X : \mathbb{R} \to \mathbb{R}; f_X(x) = \begin{cases} \frac{1}{12}(4 + 2x - x^2) & \text{für } 0 \leq x \leq 3 \\ 0 & \text{sonst} \end{cases} .$$

(a) Berechnen Sie die Verteilungsfunktion F_X von X.

(b) Berechnen Sie $P\{X \leq 1\}$, $P\{X \geq 2\}$ sowie $P\{1 \leq X \leq 3\}$.

(c) Berechnen Sie $\mathrm{E}(X)$ und $\mathrm{Var}(X)$.

(d) Ist X symmetrisch um ihren Erwartungswert verteilt?

Ergebnisse:

(a) $F_X(x) = \begin{cases} 0 & \text{für } x < 0 \\ \frac{x}{3} + \frac{x^2}{12} - \frac{x^3}{36} & \text{für } 0 \leq x \leq 3 \\ 1 & \text{für } x > 3 \end{cases}$

(b) Man erhält:

- $P\{X \leq 1\} = 0.3\overline{8}$
- $P\{X \geq 2\} = 0.\overline{2}$
- $P\{1 \leq X \leq 3\} = 0.6\overline{1}$

(c) $\mathrm{E}(X) = 1.3125$, $\mathrm{Var}(X) = 0.60234$.

(d) Nein.

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	$p_{i\cdot}$
2	0.1	0.2	0.1	
4	0.1	0.25	0.25	
$p_{\cdot j}$				

(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 $\mathrm{E}(3X + 2Y)$ sowie $\mathrm{Var}(3X + 2Y)$.

Ergebnisse:

(a) Ergänzte Tabelle:

$X \setminus Y$	0	1	2	$p_{i\cdot}$
2	0.1	0.2	0.1	0.4
4	0.1	0.25	0.25	0.6
$p_{\cdot j}$	0.2	0.45	0.35	1

(b) Es gilt:

- $\mathrm{E}(X) = 3.2$
- $\mathrm{E}(Y) = 1.15$
- $\mathrm{Var}(X) = 0.96$
- $\mathrm{Var}(Y) = 0.5275$
- $\mathrm{Cov}(X, Y) = 0.12$
- $\mathrm{Korr}(X, Y) = 0.16863$

(c) $\mathrm{Korr}(X, Y) \neq 0 \quad \Rightarrow \quad X, Y$ nicht stochastisch unabhängig.

(d) $\mathrm{E}(3 \cdot X + 2 \cdot Y) = 11.9$

$\mathrm{Var}(3 \cdot X + 2 \cdot Y) = 12.19$

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

Die Zufallsvariablen $X_1, \ldots, X_{144}$ seien unabhängig identisch $B(1, 0.5)$-verteilt. Die Summe der Zufallsvariablen X_i sei mit

$$Z := \sum_{i=1}^{144} X_i = X_1 + \ldots + X_{144}$$

bezeichnet.

(a) Geben Sie die (exakte) Verteilung von Z sowie deren Erwartungswert $\mathrm{E}(Z)$ und Varianz $\mathrm{Var}(Z)$ an.

(b) Verwenden Sie den zentralen Grenzwertsatz, um näherungsweise zu bestimmen, mit welcher Wahrscheinlichkeit Z Werte zwischen 66 und 84 annimmt.

(c) Verwenden Sie den zentralen Grenzwertsatz, um näherungsweise ein 0.75-Quantil von Z zu bestimmen.

Hinweis: Verwenden Sie zur Bearbeitung von Aufgabenteil (b) und (c) die Tabelle zur Standardnormalverteilung auf Seite 11!

Ergebnisse:

(a) $Z \sim \mathbf{B}(144, 0.5)$, $\mathrm{E}(Z) = 72$, $\mathrm{Var}(Z) = 36$.

(b) $P\{66 \leq Z \leq 84\} = 0.8185$

(c) $z_{0.75} = 76.02$

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