

Wirtschaftswissenschaftliches Prüfungsamt
 BACHELOR-PRÜFUNG
 DESKRIPTIVE STATISTIK UND WAHRSCHEINLICHSRECHNUNG
 SOMMERSEMESTER 2012

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 + 12 + 16 + 7 + 6 + 5 + 23 + 13 + 10) 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.

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 -1 Punkt, nicht bearbeitete Aussagen 0 Punkte (Aussagen mit zwei Kreuzen zählen als nicht bearbeitet!).

Die Aufgabe wird insgesamt mit mindestens 0 Punkten bewertet!

- | | wahr | falsch |
|---|-------------------------------------|-------------------------------------|
| 1. Jedes nominalskalierte Merkmal besitzt auch die Eigenschaften ordinalskaliertes Merkmale. | <input type="checkbox"/> | <input checked="" type="checkbox"/> |
| 2. Die absoluten Klassenhäufigkeiten von klassierten Merkmalen erhält man stets als Produkt der jeweiligen relativen Klassenhäufigkeit und der zugehörigen Klassenbreite. | <input type="checkbox"/> | <input checked="" type="checkbox"/> |
| 3. Empirische Verteilungsfunktionen F sind stets monoton wachsende Funktionen. | <input checked="" type="checkbox"/> | <input type="checkbox"/> |
| 4. Ist (Ω, \mathcal{F}, P) ein beliebiger Wahrscheinlichkeitsraum und sind $A, B \in \mathcal{F}$ mit $P(B) > 0$, so gilt stets:
$P(A B) + P(\bar{A} B) = 1$ | <input checked="" type="checkbox"/> | <input type="checkbox"/> |
| 5. Wenn Sie genau 5 Aufgabenteile dieser Aufgabe bearbeiten, dann haben Sie $\binom{8}{5}$ Möglichkeiten für die Auswahl der Aufgabenteile. | <input checked="" type="checkbox"/> | <input type="checkbox"/> |
| 6. Existiert der Erwartungswert $E(X)$ einer Zufallsvariablen X und ist X symmetrisch um $a \in \mathbb{R}$, dann gilt stets $a = E(X)$. | <input checked="" type="checkbox"/> | <input type="checkbox"/> |
| 7. Sind X und Y Zufallsvariablen über einem Wahrscheinlichkeitsraum (Ω, \mathcal{F}, P) , so gilt stets
$E(X \cdot Y) = E(X) \cdot E(Y) .$ | <input type="checkbox"/> | <input checked="" type="checkbox"/> |
| 8. Summen unabhängig Poisson-verteilter Zufallsvariablen sind wieder Poisson-verteilt. | <input checked="" type="checkbox"/> | <input type="checkbox"/> |

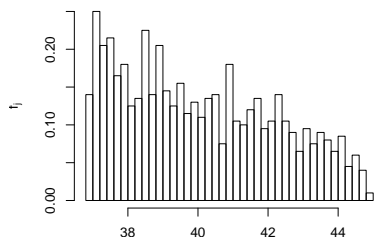
Aufgabe 2 (12 Punkte)

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 -1 Punkt, nicht bearbeitete Aufgabenteile 0 Punkte (Aufgabenteile mit mehr als einem Kreuz zählen als nicht bearbeitet!).

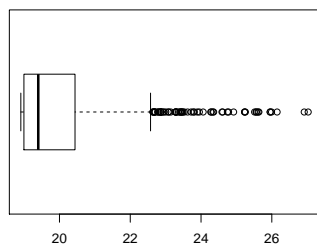
Die Aufgabe wird insgesamt mit mindestens 0 Punkten bewertet!

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



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

2. 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

3. Sind X und Y zwei stochastisch unabhängige Zufallsvariablen mit $X \sim N(50, 3^2)$ und $Y \sim N(30, 4^2)$, dann ist die Verteilung von $X + Y$ eine

- (a) $N(80, 5^2)$ -Verteilung.
- (b) $N(80, 7^2)$ -Verteilung.
- (c) $N(40, 5^2)$ -Verteilung.
- (d) $N(40, 7^2)$ -Verteilung.

4. 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 4 (6 + 3 + 4 + 3 = 16 Punkte)

Zu einem kardinalskalierten Merkmal sei die folgende (zur einfacheren Bearbeitung der Aufgabe bereits sortierte) Urliste der Länge $n = 30$ gegeben:

16.81, 20.76, 24.93, 30.59, 33.95, 34.11, 34.69, 35.81, 37.62, 39.78, 40.75, 43.47,
45.90, 47.25, 48.21, 51.01, 51.39, 52.29, 53.74, 54.85, 57.78, 58.44, 58.80, 59.71,
60.02, 64.97, 65.17, 65.24, 66.35, 67.96

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

$$K_1 = (10, 30], K_2 = (30, 50], K_3 = (50, 60], K_4 = (60, 70]$$

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 47.412?
- (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 25 und 50. Welche Näherung für diese Anzahl können Sie aus der in Teil (c) aufgestellten empirischen Verteilungsfunktion berechnen?

Ergebnisse:

- (a) Klassierung:

Nr.	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	(10, 30]	20	20	3	0.1	0.0050	0.1
2	(30, 50]	20	40	12	0.4	0.0200	0.5
3	(50, 60]	10	55	9	0.3	0.0300	0.8
4	(60, 70]	10	65	6	0.2	0.0200	1.0

- (b) Mittelwert (näherungsweise): 47.5
relative Abweichung vom exakten Wert: 0.001856 bzw. 0.1856%
- (c) (Approximative) Verteilungsfunktion:

$$F(x) = \begin{cases} 0 & \text{für } x \leq 10 \\ 0.005 \cdot (x - 10) & \text{für } 10 < x \leq 30 \\ 0.1 + 0.02 \cdot (x - 30) & \text{für } 30 < x \leq 50 \\ 0.5 + 0.03 \cdot (x - 50) & \text{für } 50 < x \leq 60 \\ 0.8 + 0.02 \cdot (x - 60) & \text{für } 60 < x \leq 70 \\ 1 & \text{für } x > 70 \end{cases}$$

- (d) Anzahl (aus Urliste): 12
Mit emp. Verteilungsfunktion genäherte Anzahl: 12.75

Aufgabe 5 (5 + 2 = 7 Punkte)

Ein Reifenhersteller verteilt seine Reifenproduktion auf drei Werke (A,B,C), wobei 50% aller Reifen in Werk A, 15% aller Reifen in Werk B sowie 35% aller Reifen in Werk C produziert werden. Aufgrund langjähriger Erfahrungen aus der Qualitätskontrolle ist bekannt, dass es bei 2% aller in Werk A, bei 5% aller in Werk B sowie bei 3% aller in Werk C produzierten Reifen zu einem Fabrikationsfehler kommt.

- (a) Wie groß ist die Wahrscheinlichkeit, dass bei einem zufällig ausgewählten Reifen ein Fabrikationsfehler auftritt?
- (b) Ein bestimmter Reifen weist einen Fabrikationsfehler auf. Wie groß ist die Wahrscheinlichkeit, dass dieser Reifen in Werk C produziert wurde?

Ergebnisse:

- (a) 0.028
- (b) 0.375

Aufgabe 6 (6 Punkte)

Bei einem Würfelspiel werden 5 (faire) Würfel gleichzeitig geworfen. Wie groß ist die Wahrscheinlichkeit, dabei eine „große Straße“, also die Punktzahlkombinationen 1,2,3,4,5 oder 2,3,4,5,6 (nicht unbedingt in dieser „Reihenfolge“), zu würfeln?

Hinweis: Achten Sie darauf, für die Modellierung des Zufallsexperiments einen Laplace-schen Wahrscheinlichkeitsraum zu verwenden.

Ergebnisse:

Gesuchte Wahrscheinlichkeit: 3.086%

Aufgabe 7 (2 + 3 = 5 Punkte)

Als Hausaufgabe im Fach Musik waren die Geburtsdaten 20 berühmter Komponisten auswendig zu lernen. Der Schüler Fritz Fleißig hat 16 dieser Daten auswendig gelernt (die Chance, bei den anderen 4 Geburtstagen durch Raten eine richtige Antwort zu geben, sei gleich Null). Der Lehrer überprüft, ob Fritz die Hausaufgabe ordentlich erledigt hat, indem er 5 Mal rein zufällig und unabhängig voneinander jeweils einen Komponisten auswählt und das zugehörige Geburtsdatum abfragt. Kann Fritz mindestens zu 3 dieser 5 Komponisten das Geburtsdatum korrekt angeben, so hat er die Überprüfung bestanden.

- (a) Welche Verteilung besitzt die Anzahl der von Fritz abgegebenen richtigen Antworten?
- (b) Mit welcher Wahrscheinlichkeit besteht er die Überprüfung (formelmäßige Angabe mit der Verteilungsfunktion der Verteilung aus Teil (a) genügt)?

Ergebnisse:

- (a) Verteilung: $B(5, 0.8)$
- (b) Gesuchte Wahrscheinlichkeit: $1 - F_{B(5,0.8)}(3 - 0) = 1 - F_{B(5,0.8)}(2) (= 94.208\%)$

Aufgabe 8 (3 + 2 + 12 + 4 + 2 = 23 Punkte)

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

$$F_X(x) = \begin{cases} 0 & \text{für } x \leq -2 \\ \frac{1}{6}x^2 + \frac{2}{3}x + \frac{2}{3} & \text{für } -2 < x \leq 0 \\ -\frac{1}{3}x^2 + \frac{2}{3}x + \frac{2}{3} & \text{für } 0 < x \leq 1 \\ 1 & \text{für } x > 1 \end{cases}$$

- (a) Berechnen Sie eine Dichtefunktion f_X von X .
- (b) Berechnen Sie $P(\{X < -1\})$ sowie $P(\{X > \frac{1}{2}\})$.
- (c) Bestimmen Sie den Erwartungswert $E(X)$ und die Varianz $\text{Var}(X)$.
- (d) Bestimmen Sie ein unteres Quartil von X .
- (e) Bestimmen Sie Erwartungswert und Varianz von $Y := -3 \cdot X + 2$.

Ergebnisse (ohne Begründung/Rechenweg):

- (a) Dichtefunktion von X :

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

- (b) $P(\{X < -1\}) = \frac{1}{6}$, $P(\{X > \frac{1}{2}\}) = \frac{1}{12}$
- (c) $E(X) = -\frac{1}{3}$, $\text{Var}(X) = \frac{7}{18}$
- (d) Unteres Quartil: $x_{0.25} = -0.7753$
- (e) $E(Y) = 3$, $\text{Var}(Y) = \frac{7}{2}$

Aufgabe 9 (4 + 2 + 4 + 3 = 13 Punkte)

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

$X \setminus Y$	1	3	5
2	0.05	0	0.15
3	0.1	0.1	0.1
5	0.2	0.1	0.2

- (a) 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.
- (b) Sind X und Y stochastisch unabhängig? Begründen Sie Ihre Antwort!
- (c) Berechnen Sie unter Verwendung der Ergebnisse

$$E(X) = 3.8, \quad E(Y) = 3.2, \quad E(X^2) = 16, \quad E(Y^2) = 13.4, \quad E(X \cdot Y) = 11.8$$

die Varianzen von X und Y sowie die Kovarianz und den Korrelationskoeffizienten zwischen X und Y .

- (d) Berechnen Sie $E(3X + 4Y)$ sowie $\text{Var}(3X + 4Y)$.

Ergebnisse:

- (a) Tabelle der bedingten Wahrscheinlichkeitsfunktionswerte von $X|Y = y_j, j \in \{1, 2, 3\}$:

x_i	2	3	5
$p_{X Y=1}(x_i)$	$\frac{1}{7}$	$\frac{2}{7}$	$\frac{4}{7}$
$p_{X Y=3}(x_i)$	0	$\frac{1}{2}$	$\frac{1}{2}$
$p_{X Y=5}(x_i)$	$\frac{1}{3}$	$\frac{2}{9}$	$\frac{4}{9}$

- (b) Nein, X und Y sind stochastisch abhängig, da die bedingten Verteilungen aus Teil (a) nicht alle miteinander übereinstimmen.
- (c) $\text{Var}(X) = 1.56, \text{Var}(Y) = 3.16, \text{Cov}(X, Y) = -\frac{9}{25}, \text{Korr}(X, Y) = -0.1621$
- (d) $E(3 \cdot X + 4 \cdot Y) = 24.2, \text{Var}(3 \cdot X + 4 \cdot Y) = 55.96$

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

Ein Online-Händler bietet für 100 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) stochastisch unabhängig stetig gleichverteilt auf dem Intervall $[0.05, 0.35]$ (also $\text{Unif}(0.05, 0.35)$ -verteilt) sind.

- (a) Welchen Erwartungswert und welche Standardabweichung hat die Summe der Abfertigungsdauern von 100 Express-Bestellungen?
- (b) Verwenden Sie den zentralen Grenzwertsatz, um die Wahrscheinlichkeit, dass 100 Express-Bestellungen in höchstens 21 (Mitarbeiter-)Stunden abgefertigt werden können, (näherungsweise) zu berechnen.
- (c) Verwenden Sie den zentralen Grenzwertsatz, um das 97.5%-Quantil der Gesamtabfertigungsdauer von 100 Express-Bestellungen (näherungsweise) zu bestimmen.

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

Ergebnisse:

Es bezeichne Z die Gesamtabfertigungsdauer von 100 Express-Bestellungen.

- (a) $E(Z) = 20$, $\sqrt{\text{Var}(Z)} = 0.866$.
- (b) $P\{Z \leq 21\} = 0.8749$
- (c) $z_{0.975} = 21.6974$

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