Statistik für Wirtschaftswissenschaftler

Klausur am 22.01.2010, 14.00–17.00.

Bitte unbedingt beachten:

a) Gewertet werden alle 15 gestellten Aufgaben.

b) Lösungswege sind anzugeben. Die Angabe des Endergebnisses allein gilt nicht als Lösung. Da *keine* Taschenrechner zugelassen sind, brauchen Zahlenrechnungen, für die man normalerweise einen Taschenrechner benutzen würde, nicht durchgeführt zu werden.
Ausnahme: Zwischenergebnis, für das der Zahlenwert für die weitere Behandlung der Aufgabe unbedingt nötig ist. Dieser Zahlenwert kann aber dann durch Kopfrechnung ermittelt werden.
Ein Endergebnis ist vollständig, wenn zur Ermittlung des Zahlenwertes höchstens die Ausführung der elementaren Rechenoperationen (Addition, Subtraktion, Multiplikation, Division) und die Anwendung elementarer Funktionen ($\exp x (\equiv \mathrm{e}^{\mathrm{x}})$, $\ln x$, $\log x$, $\sin x$, $\cos x$, $\tan x$, $\arcsin x$, $\arccos x$, $\arctan x$, x^y, $\sqrt{x}, \sqrt[y]{x})$ nötig wäre. Z.B. wären $400 \cdot (1.004^{30} - 4)$ oder $\arctan(3.0/\sqrt{13.4})$ gültige Endergebnisse.
Die Bildung von $m!$ und des Binomialkoeffizienten z.B. gehören *nicht* zu den elementaren Rechenoperationen.

c) Zugelassene Hilfsmittel: 40 Seiten DIN A4 mit Sätzen, Definitionen und Formeln (einschließlich begleitender Text dazu), **aber ohne Aufgaben, ohne Lösungsvorschläge von Aufgaben und auch ohne Beispiele,**
Fremdsprachenwörterbücher ohne zusätzliche Einträge,
Tabelle der Standardnormalverteilung ohne zusätzliche Einträge ,
Tabelle der Quantile der χ^2–Verteilung ohne zusätzliche Einträge ,
Tabelle der Quantile der t–Verteilung ohne zusätzliche Einträge.

Weitere Hinweise:

a) Wer mindestens 45 Punkte erreicht hat, hat bestanden.

b) Weitere Infos finden Sie im Internet in dem File “allginfo.pdf” im Verzeichnis “`http://www.mathematik.uni-stuttgart.de/studium/infomat/WiSBa_Kolbe_SS09/`”.

Aufgabe 1 **9 Punkte**

Bei den Kleinbetrieben einer Stadt wurden die Meldung der Umsätze vom 02.01.10 in der folgenden Häufigkeitstabelle ausgewertet:

Klasse	Tagesumsatz von ... bis unter ... (Euro)	prozentuale Häufigkeit
1	0 – 400	20
2	400 – 500	30
3	500 – 600	20
4	600 – 900	30

a) Zeichnen Sie das Histogramm.

b) Bestimmen Sie die näherungsweise die Grenze y, für die die Tagesumsätze von 70% der Betriebe $\geq y$ sind?

c) Wieviel % der Betriebe haben näherungsweise einen Tagesumsatz von 450 Euro oder mehr?

Aufgabe 2 **5 Punkte**

In einem Betrieb wurde 2008 Artikel B durch Artikel B' ersetzt:

	2007		2008		2009	
Artikel	Stückpreis in Euro	Stückzahl	Stückpreis in Euro	Stückzahl	Stückpreis in Euro	Stückzahl
A	40.-	4000	50.–	4500	45.–	4800
B	30.–	1000	40.–	500	×	—
B'	×	—	80.–	600	82.–	1200

Beschreiben Sie die Preisentwicklung (nicht bei den einzelnen Artikeln, sondern bei dem Gesamtbetrieb) von 2007 nach 2008 und von 2007 nach 2009 durch die Bestimmung je eines geeigneten Indexes.

Aufgabe 3 **7 Punkte**

Ein Interviewer wählt aus den 500 Besuchern einer Veranstaltung nacheinander 6 zufällig für ein Interview aus, wobei er sich nicht merkt, wen er schon ausgewählt hat. Wie groß ist die Wahrscheinlichkeit, dass er mindestens einmal jemanden auswählt, den er schon befragt hat?

– bitte wenden –

Aufgabe 4 **6 Punkte**

Die Umsatzdaten eines Unternehmens seien in Quartalswerten y_i angegeben, und zwar ab dem 1.Quartal 2007. Die ersten 5 Werte der Zeitreihe (beginnend also mit dem 1.Quartal 2007) sind $y_1 = 155$, $y_2 = 34$, $y_3 = 66$, $y_4 = 57$, $y_5 = 11$. Zur Vorbereitung der Schätzung der Saisonnormale wurden von den Werten y_i – soweit möglich – die zugehörigen Werte des gleitenden Durchschnitts abgezogen, was dann folgende Differenzen d_i (in Mio.Euro) ergab:

	Quartal			
Jahr	I	II	III	IV
2007	×	×	6	3
2008	-7	9	2	1
2009	-3	5	×	×

Bestimmen Sie dazu die ersten 5 Schätzwerte (beginnend also mit dem 1.Quartal 2007) der saisonbereinigte Zeitreihe.

Aufgabe 5 **8 Punkte**

Zu drei Merkmalen liegen Daten aus 4 Beobachtungen vor.

i	1	2	3	4
x_i	3	10	0	3
y_i	2	1	-4	1
z_i	4	5	-1	2

Geben Sie für die Regressionsebene $z = a_1 + b_1 x + c_1 y$ ein lineares Gleichungssystem für a_1, b_1 und c_1 an. Es sind also die Elemente der Koeffizientenmatrix und die Koordinaten des Störvektors zu bestimmen. Eine Lösung des lineares Gleichungssystems ist *nicht* verlangt.

Aufgabe 6 **5 Punkte**

$$f(x) := \begin{cases} \frac{28}{9} \cdot x^{1/3} \cdot (1-x) & \text{für } 0 \le x \le 1, \\ 0 & \text{sonst,} \end{cases}$$

sei die Wahrscheinlichkeitsdichte einer stetigen Zufallsvariablen X. Bestimmen Sie den Erwartungswert und die Varianz von X.

Hinweis: Hilfsformel zur Bestimmung der Integrale: $\frac{d}{dx} x^\alpha = \alpha x^{\alpha-1}$.

Aufgabe 7 **6 Punkte**

Eine automatische Feuermeldeanlage gebe mit der Wahrscheinlichkeit 0.980 richtigen und mit der Wahrscheinlichkeit 0.220 falschen Alarm. Die Wahrscheinlichkeit für einen Brand sei 0.003. Wie groß ist bei Feueralarm die Wahrscheinlichkeit, dass es tatsächlich brennt?

Geben Sie dabei die Ereignisse an, deren Wahrscheinlichkeiten bzw. bedingte Wahrscheinlichkeiten in die Berechnungsformel für die gesuchte Wahrscheinlichkeit eingehen.

Aufgabe 8 **9 Punkte**

Die Reaktionszeit eines Autofahrers auf visuelle Reize ist normalverteilt mit einem Erwartungswert von 0.4 Sekunden und einer Standardabweichung von 0.05 Sekunden.

(i) Bestimmen Sie die Wahrscheinlichkeit, dass er für eine Reaktion mehr als 0.5 Sekunden braucht.

(ii) Mit welcher Wahrscheinlichkeit benötigt er für seine Reaktion zwischen 0.4 und 0.5 Sekunden?

(iii) Welche Reaktionszeit wird mit einer 90%- Wahrscheinlichkeit überschritten?

Aufgabe 9 **8 Punkte**

Sei X die Zahl der fehlerhaften Bits, welche in eine Telekommunikationsleitung ankommen. Angenommen, X ist binomialverteilt mit $p = 0.001$ und 1000 Bits werden übertragen. Bestimmen Sie:

(i) $P(X = 1)$;

(ii) $P(X \geq 1)$;

(iii) $P(X \leq 2)$;

(iv) Erwartungswert und Varianz von X.

Aufgabe 10 **5 Punkte**

Aus früheren Erfahrungen ist es bekannt, dass die Produktivität eines chemischen Prozesses $N(\mu, 3)$-verteilt ist. In den letzten 5 Tagen hat die entsprechende Chemieanlage einen Produktivitätsdurchschnittswert von
$(x_1 + x_2 + x_3 + x_4 + x_5)/5 = 90.48$

geliefert. Bestimmen Sie ein 95%-Konfidenzintervall für μ.

Aufgabe 11 **6 Punkte**

Ein Biologe untersucht eine Spezies von Käfern, welche normalerweise in Tälern leben. Hoch in den Bergen entdeckt er eine Käferpopulation aus dieser Spezies und untersucht $n = 25$ Käfer um festzustellen, ob sich die Bergkäfer von den Talkäfern unterscheiden. Eines der relevanten Merkmale ist dabei die Länge der schwarzen Flecken auf deren Flügeln. Bekanntlich ist letztere für Talkäfer normalverteilt, mit einem Erwartungswert von $\mu = 3.14$. Aus den Messungen findet der Biologe für seine Bergkäfer eine Fleckenlänge von durchschnittlich $\bar{x} = 3.23$ mit einer geschätzten Standardabweichung von

$$\sqrt{\frac{1}{24}\sum_{i=1}^{25}(x_i - \overline{x})^2} = 0.214\,.$$

Er nimmt an, dass die Fleckenlänge bei den Bergkäfern ebenfalls normalverteilt ist.

Bei einem Signifikanzniveau von $\alpha = 0.05$, testen Sie die Hypothese, dass sich (bzgl. Fleckenlänge auf Flügeln) die Bergkäfer nicht von ihren Kollegen, die im Tal leben, unterscheiden.

Hinweis. $0.428 \cdot 2.064 = 0.8834$.

Aufgabe 12 **7 Punkte**

Im Rahmen einer großangelegten Studie über Frauen und deren Familienleben interessiert u.a. das Alter von Frauen bei der Geburt des ersten Kindes. Es wird vermutet, dass das Durchschnittsalter Erstgebärender bei über 25 Jahren liegt.

Zur Überprüfung dieser Hypothese werden 64 Mütter zufällig ausgewählt und nach ihrem Alter bei der Geburt ihres ersten Kindes befragt. Es ergab sich ein Durchschnittsalter von $\bar{x} = 27$.

Überprüfen Sie die Hypothese

$$H_0: \qquad \mu \leq 25$$

gegen die Alternative

$$H_1: \qquad \mu > 25.$$

Benutzen Sie dabei für H_0 eine Irrtumswahrscheinlichkeit von $\alpha = 0.01$ und für H_1 eine Irrtumwahrscheinlichkeit von $\beta = 0.05$. Gehen Sie davon aus, dass das Alter Erstgebährender normalverteilt ist, mit einer Standardabweichnung von 4 Jahren.

Aufgabe 13 **6 Punkte**

Bestimmen Sie die in der folgenden Tabelle noch fehlenden Wahrscheinlichkeiten der gemeinsamen Verteilung und der Randverteilungen, wenn die beiden ZV X und Y *unabhängig* sind:

$\downarrow X \mid Y \rightarrow$	-1	0	3	
0	0.15	$*$	$*$	$*$
1	$*$	$*$	$*$	0.40
3	$*$	0.03	$*$	$*$
	$*$	0.10	$*$	

Aufgabe 14 **8 Punkte**

Eine Lieferung von 300 000 Stück soll höchstens 5% defekte Stücke enthalten. Für die Wahrscheinlichkeit p für die Ziehung eines defekten Stückes soll also die Hypothese $H_0 : p \leq 0.05$ (etwa) gegen die Hypothese $H_1 : p \geq 0.051$ gestestet werden, und zwar mit 0.10 für die Wahrscheinlichkeit für einen Fehler 1. und 2. Art.

a) Es werden 60 zufällige Ziehungen m.Z. durchgeführt. Zu welchem Testergebnis kommen Sie, wenn zweimal ein defektes Stück gezogen wird? *Benötigte Rechenergebnisse:* $0.10/(0.949^{60}) = 2.312\,,\quad 60 \cdot 51/949 = 3.224$ und $30 \cdot 59 \cdot (51^2)/(949^2) = 5.112$

b) Über welchem Wert u muss bei einer evtl. zweiten Stichprobe der Stichprobenumfang liegen, damit die Hypothese H_0 überhaupt mit ausreichender Sicherheit abgelehnt werden kann?

Aufgabe 15 **7 Punkte**

Die drei Schichten A,B und C, in die die Grundgesamtheit der 10^4 Kommunen eines Staates aufgeteilt wird, haben die Anteile von 20%, 30% bzw. 50%. Das Durchschnittssteueraufkommen μ aller Kommunen soll geschätzt werden. Dazu wird eine Stichprobe vom Umfang 400 gezogen, und zwar auf zweierlei Art. Als Einheit war bei allen vorgegebenen Größen Millionen Euro gewählt worden.

a) In jeder Schicht wird eine (streng zufällige) Teilstichprobe von dem Umfang gezogen, der zu dem Anteil der Schicht an der Grundgesamtheit propotional ist. Dabei erhält man als Durchschnittswert der Steueraufkommen aller Kommunen der Teilstichprobe aus Schicht A 100, aus Schicht B 60 und aus Schicht C 20. Bestimmen Sie daraus einen Schätzwert für μ. Bestimmen Sie außerdem die Umfänge der Teilstichproben.

b) Eine neue Stichprobe vom Umfang 400 soll nach dem Prinzip der optimalen Stichprobe gezogen werden, wobei die modifizierten Standardabweichungen in den einzelnen Schichten bekannt sind: $\tilde{\sigma}_1 = 45$, $\tilde{\sigma}_2 = 20$ bzw. $\tilde{\sigma}_3 = 2$ für Schicht A, B bzw. C. Bestimmen Sie die Umfänge der Teilstichproben aus den einzelnen Schichten!