Klausur Statistik II im Sommersemester 2004
Zweittermin
Prof. Dr. Joachim Grammig

- **Beginnen Sie jede Aufgabe auf einer neuen Seite.**
- **Schreiben Sie zu jeder Lösung den Verweis in eckigen Klammern [X1], so daß die Zuordnung möglich ist. Fehlt die Zuordnung, können keine Punkte vergeben werden.**
- **Machen Sie zu jeder Lösung den Lösungsweg deutlich, ansonsten können keine Punkte vergeben werden.**
- **Beachten Sie die Lösungshinweise.**

Aufgabe A:

Für die Kapazitätsplanung einen Call-Centers benötigen Sie Modell für die Wahrscheinlichkeitsverteilung der Dauer der Anrufe der Kunden (in Sekunden). Wir bezeichnen diese Zufallsvariable mit T. Wir nehmen an, daß T exponentialverteilt ist mit Parameter $\lambda = 0,005$.

[**A1**] Approximieren Sie unter Verwendung der Dichtefunktion der Exponentialverteilung (siehe Lösungshinweis) die Wahrscheinlichkeit für das Ereignis, daß ein Anruf zwischen 60 und 120 Sekunden dauert **(4 P)** und berechnen Sie den exakten Wert dieser Wahrscheinlichkeit unter Verwendung der Verteilungsfunktion (siehe Lösungshinweis) **(4 P)**

Lösungshinweise:

$$\begin{aligned} \text{Dichtefunktion der Exponentialverteilung: } f_T(t) &= \lambda e^{-\lambda t} \\ \text{Verteilungsfunktion der Exponentialverteilung: } F_T(t) &= 1 - e^{-\lambda t} \end{aligned}$$

[**A2**] Berechnen Sie unter Verwendung der momentenerzeugenden Funktion (siehe Lösungshinweis) den Erwartungswert und die Varianz der Zufallsvariable T. **(8 P)**

Lösungshinweis:

$$\text{MEF der Exponentialverteilung: } MEF(t) = \big(1 - \frac{t}{\lambda}\big)^{-1}$$

[**A3**] Angenommen, die Kosten pro Anruf hängen ab von der Dauer des Anrufs, und zwar in der Form $K(T) = 0,01T + 0,1$. Wie hoch sind die erwarteten Kosten pro Anruf $E(K(T))$? **(5 P)**
[**A4**] Angenommen, die Kosten pro Anruf sind gegeben mit $K(T) = 0,01T^2$. Wie sind die erwarteten Kosten $E(K(T))$ jetzt zu berechnen? (Ansatz genügt, nicht ausrechen!). **(5 P)**
[**A5**] Nehmen Sie an, daß der Parameter λ nicht bekannt ist. Beschreiben Sie ein Vorgehen, wie man diesen Parameter schätzen kann (Verfahren zur Datenerhebung, verwendetes Schätzverfahren) **(8 P)**.
[**A6**] In der Grundgesamtheit deutscher Studenten sei der Präferenzwert für ein Produkt (Zufallsvariable X) und Intelligenzquotient (Zufallsvariable Y) eines potentiellen Kunden gemeinsam normalverteilt, wobei $E(X) = 65$, $Var(X) = 225$, $E(Y) = 100$ und $Var(Y) = 900$. Die beiden Zufallsvariablen (Präferenz und IQ) werden als unabhängig angenommen. Berechnen Sie die gemeinsame Wahrscheinlichkeit, dass der Präferenzwert größer ist als $x = 50$ und der Intelligenzquotient größer gleich $y = 115$. **(5 P)**

Lösungshinweise:

Ausgewählte Werte aus der Verteilungsfunktion der Standardnormalverteilung

z	0,5	0,75	1	1,25	1,5
$F_Z(z)$	0,69	0,77	0,84	0,89	0,93

Aufgabe B:

Zufallsvariable X bezeichnet die Auszahlung einer Finanzanlage A in fünf Zuständen, die mit einer bestimmten Wahrscheinlichkeit eintreten können.

x	-150	-50	0	100	300
$P(X = x)$	0,2	0,1	0,4	0,2	0,1

[**B1**] Berechnen Sie $E(X)$ und $Var(X)$ und stellen Sie die Verteilungsfunktion $F_X(x)$ grafisch dar (Achsen beschriften!). Interpretieren Sie einen Wert der Verteilungsfunktion an einer beliebigen Stelle. **(10 P)**
[**B2**] Ein Finanzinstitut entwickelt nun eine (derivative) Finanzanlage B, welche die folgende Auszahlung Z liefert: Wenn die Auszahlung von Finanzanlage X weniger als 0 beträgt, wird ein Euro ausgezahlt, ansonsten ist die Auszahlung gleich Null. Stellen Sie die Wahrscheinlichkeitsfunktion und die Verteilungsfunktion der Zufallsvariable Z grafisch dar. **(5 P)** Berechnen Sie den Erwartungswert und die Varianz der Auszahlung Z. **(5 P)**
[**B3**] Berechnen Sie die Kovarianz der Zufallsvariablen X und Z, $Cov(X, Z)$ und den Korrelationskoeffizienten **(10 P)**.

Aufgabe C:

Die risikobehafteten Auszahlungen (in Euro) von drei Investitionsobjekten in drei Regionen seien bezeichnet mit X, Y, Z. Die Auszahlungen sind zum Investitionszeitpunkt noch unbekannt, d.h. es sind Zufallsvariablen. Ein Investor stellt sich nun aus diesen Investitionsobjekten ein Portfolio zusammen. Er beteiligt sich mit 50 % am ersten Investitionsobjekt (er erhält somit 50 % der Auszahlung X). Am zweiten Investitionsobjekt beteiligt er sich zu 100% (d.h. er erhält 100 % der Auszahlung Y). Von der Auszahlung Z erwirbt er 25 %. Die Erwartungswerte der drei Zufallsvariablen werden in folgendem Vektor gesammelt: Auszahlungsvektors beträgt:

$$E(\underline{X}) = \big(E(X), E(Y), E(Z)\big)' = \big(50, 20, 100\big)'.$$

Die Kovarianzen zwischen den Auszahlungen betragen:

$$Cov(\underline{X}) = \begin{pmatrix} Var(X) & Cov(X,Y) & Cov(X,Z) \\ Cov(X,Y) & Var(Y) & Cov(Y,Z) \\ Cov(X,Z) & Cov(Y,Z) & Var(Z) \end{pmatrix} = \begin{pmatrix} 1600 & 640 & 800 \\ 640 & 1600 & 2800 \\ 800 & 2800 & 10000 \end{pmatrix}$$

[**C1**] Berechnen Sie den Erwartungswert und die Standardabweichung der Auszahlung des Portfolios.
(15 P)
[**C2**] Außerdem unterstellen wir, daß X, Y, Z gemeinsam normalverteilt sind. Berechnen Sie das 5 % Quantil der Wahrscheinlichkeitsverteilung der Auszahlung des Investor-Portfolios. **(5 P)**

Lösungshinweise:

Für einen Vektor von reellen Zahlen $\underline{a}$ und einen Vektor von Zufallsvariablen $\underline{X}$ gilt: $E(\underline{a}'\underline{X}) = \underline{a}'E(\underline{X})$ und $Var(\underline{a}'\underline{X}) = \underline{a}'Cov(\underline{X})\underline{a}$

Verteilungsfunktion der Standardnormalverteilung

z	1,2816	1,6449	1,9600	2,3263
$F_Z(z)$	0,90	0,95	0,975	0,99

Aufgabe D:

Interpretieren Sie die folgenden Testergebnisse. Beachten Sie dabei die Angaben in den Lösungshinweisen.

[**D1**] Mit Hilfe des sogenannten Jarque-Bera Tests kann man die Nullhypothese testen, daß eine Zufallsvariable normalverteilt ist. Die Alternativhypothese lautet, daß die Zufallsvariable nicht normalverteilt ist. Zur Konstruktion der Teststatistik wird eine Zufallsstichprobe aus der Grundgesamtheit benötigt. Unter der Nullhypothese ist die Jarque-Bera Statistik χ^2 verteilt mit 2 Freiheitsgraden. Der Test ist einseitig, d.h. man verwirft die Nullhypothese für große Werte der Teststatistik. Die Berechnung der Jarque-Bera Teststatistik auf Basis einer Zufallstichprobe von logarithmierten individuellen Einkommensdaten liefert einen Wert von 10,9. Interpretieren Sie dieses Testergebnis. **(6 P)**

[D2] Die Jarque-Bera Teststatistik wurde außerdem auf Basis einer weiteren Zufallstichprobe berechnet, in welcher der IQ des Befragten erhoben wurde. Für diese Daten ergibt sich ein $p-$Wert (empirisches Signifikanzniveau) der Jarque-Bera Statistik von 0,12. Interpretieren Sie dieses Testergebnis. **(6 P)**

[D3] Der Kruskal-Wallis-Test wird dazu verwendet, die Nullhypothese zu testen, daß zwei oder mehr Zufallstichproben aus Verteilungen gezogen wurden, in der die interessierende Zufallsvariable den gleichen Erwartungswert besitzt. So könnte man z.B. testen, ob die erwartete Präferenz für ein Produkt (z.B. Weissbier) in 2 Verkaufsregionen (z.B. Schleswig-Holstein und Bayern) identisch ist. Zur Konstruktion benötigt man Zufallsstichproben aus den beiden Verkaufsregionen. Unter der Nullhypothese folgt die Teststatistik einer Verteilung, die in vielen Lehrbüchern tabelliert ist. Der Test ist als einseitiger Test formuliert, d.h. wir verwerfen für große Werte der Teststatistik. In einer Marketingstudie wurden in Zufallstichproben Daten zu den Präferenzwerten für das Produkt in den o.g. Regionen erhoben und die Kruskal-Wallis-Teststatistik berechnet. Es ergab sich ein $p-$Wert (empirisches Signifikanzniveau) von 0,001. Interpretieren Sie dieses Testergebnis. **(6 P)**

[D4] Lo und MacKinlay haben einen Test entwickelt, mit dem die Null-Hypothese getestet werden kann, daß Aktienrenditen mit Hilfe der Historie der Renditen nicht prognostizierbar sind. Die Alternativhypothese lautet, daß Prognostizierbarkeit vorliegt. Unter der Nullhypothese ist die Lo/MacKinlay Teststatistik standardnormalverteilt. Zur Konstruktion der Teststatistik benötigt man eine Zeitreihe von Aktienrenditen. Der Test ist zweiseitig. Unter Verwendung von Kurszeitreihen der Börse New York wurde für eine Zeitreihe von Aktienrenditen ein Wert der Lo/MacKinlay Teststatistik von -1,s56 ermittelt. Interpretieren Sie dieses Testergebnis. **(6 P)**

[D5] Bei statistischen Hypothesentests spricht man von einer Wahrscheinlichkeit für einen Fehler 1. Art bzw. der Wahrscheinlichkeit für einen Fehler 2. Art und der Macht eines Tests. Was versteht man unter diesen Begriffen? Erläutern Sie den unvermeidlichen Trade-Off zwischen den beiden Fehlerwahrscheinlichkeiten. **(7 P)**

Lösungshinweise:

Verteilungsfunktion der χ^2-Verteilung mit 2 Freiheitsgraden:

$F(x)$	0,500	0,90	0,950	0,975	0,990	0,995
X	1,39	4,61	5,99	7,38	9,21	10,6

Verteilungsfunktion der Standardnormalverteilung

z	1,2816	1,6449	1,9600	2,3263
$F_Z(z)$	0,90	0,95	0,975	0,99