Klausur
Grundlagen der Statistik
(SS 2002)

Bemerkungen:

- Die Klausur besteht aus **4 Aufgaben**, die alle zu lösen sind.
- Lösen Sie die Aufgaben **nur** auf den ausgeteilten Lösungsbögen.
- Vermerken Sie auf **jedem Blatt** Ihren Namen und Ihre Matrikelnummer.
- Schreiben Sie bitte **lesbar**.
- Achten Sie darauf, dass Ihr **Lösungsweg nachvollziehbar** ist.

Aufgabe 1 *(10 Punkte)*

Bewerten Sie jede dieser fünf Aussagen mit RICHTIG oder FALSCH und **begründen** Sie Ihre Entscheidung. Für (richtige) Bewertungen ohne Begründung gibt es keine Punkte.

1. Falls für die Zufallsvariablen X und Y gilt: $Var(X) = Var(Y) = 1$, so folgt daraus, dass ihre Korrelation gleich ihrer Kovarianz ist.
2. Likelihoodfunktionen geben die Wahrscheinlichkeit dafür an, dass ein (unbekannter) Parameter einen bestimmten Wert annimmt. Dabei ist sie maximal für den wahren Wert des Parameters.
3. Wenn für $A, B, C \subset \Omega$ mit $A \cap B \cap C \neq \emptyset$ gilt $\mathrm{P}(A) > \mathrm{P}(B) > \mathrm{P}(C) > 0$, dann gilt auch $\mathrm{P}(B|A) > \mathrm{P}(C|A)$.
4. Falls X eine binomialverteilte Zufallsvariable mit Parameter n und π ist, dann ist $\widehat{\pi} = \frac{X}{n}$ eine erwartungstreue Schätzung für π.
5. Die Wahrscheinlichkeit eine Realisation einer Zufallsvariablen zu beobachten, die mehr als drei Standardabweichungen vom Erwartungswert abweicht, ist gering. Dabei kann die Zufallsvariablen eine beliebige Verteilung besitzen.

Aufgabe 2 *(15 Punkte)*

Die Aufgabe besteht aus vier Teilaufgaben.

Der Student X. sammelt Kassenbelege. Alle vier bis sechs Wochen setzt er sich hin und wertet die Belege statistisch aus.

Heute ist es mal wieder so weit. In den letzten Wochen hat X. Kassenbelege mit folgende Beträgen (in EURO) gesammelt.

Beleg i	1	2	3	4	5	6	7	8	9	10	11	12
Betrag von Beleg i	103	6.5	5.5	41	2	5	11	13	7	4	10	32

Zum ersten Mal schreibt X. die Werte in EURO auf, da sein Supermarkt neuerdings die Preis nicht mehr in DM auszeichnet.

1. Berechnen Sie das arithmetische Mittel, die Standardabweichung sowie Median und Interquartilsabstand für die 12 Belege, die X. heute auswertet.

   Max 3 Punkte, je 0.5 Punkte für Mittelwert und Stadw., je 1 Punkte für IQR und Median
2. Fertigen Sie für die 12 vorliegenden Beträge einen Boxplot an. Charakterisieren Sie die Verteilung der Beträge.

   Max 3 Punkte, 2 Punkte für Plot und 1 Punkt für Interpretation

Da X. die Belege nicht aufhebt, sondern nach der Auswertung sorgsam dem Verwertungskreislauf zuführt, kennt er aus der DM–Zeit nur noch das arithmetische Mittel (24.87) und die Standardabweichung (25.05) in DM. Allerdings hat X. sich auch folgenden Tabelle aufgeschrieben

Betrag (in DM) von ... bis unter ...	Anzahl Belege
0 - 5	156
5 - 15	384
15 - 30	257
30 - 80	151
80 - 200	70

3. Berechnen Sie das arithmetischen Mittel über alle Kassenbelege (DM– und EURO–Zeit) in EURO sowie die entsprechende Standardabweichung.

   Max 5 Punkte, 2 Punkte für arithemtische Mittel und 3 Punkte für die Standardabweichung
4. Aktualisiere n Sie die obige Tabelle mit den gruppierten Beobachtungen und fertigen Sie das zugehörige Histogramm (in DM) an. Charakterisieren Sie die Verteilung der Beträge.

HINWEIS: Rechnen Sie mit 1 EURO = 1.96 DM um.

Aufgabe 3 *(20 Punkte)*

Die Aufgabe besteht aus drei Teilaufgaben mit insgesamt zehn Unterpunkten.

Zwei Freunde treffen sich täglich zum Basketballspielen. Aus Erfahrung wissen die beiden, dass A. mit einer Wahrscheinlichkeit von 80 Prozent den Korb trifft und B. mit einer Wahrscheinlichkeit von 70 Prozent. Dabei treffen die beiden Freunde unabhängig voneinander und unabhängig von ihren eigenen Treffern.

1. Jeder Freund wirft einmal auf den Korb. Wie groß ist die Wahrscheinlichkeit, daß

   (a) A. trifft, aber B. nicht (E1)?

   (b) A. und B. treffen?

   (c) A nicht trifft und B. trifft (E2)?

   (d) beide den Korb nicht treffen?

   (e) beide gleich häufig treffen (E3)?

   Max 5 Punkte, je ein Punkt.

2. Häufig werfen die beiden dreimal hintereinander auf den Korb und derjenige, der häufiger trifft, hat gewonnen (Dreiwurfspiel). Sei nun X die Zufallsvariable, die zählt, wieviel Treffer A. mehr bei drei Würfen erzielt als B. (z.B. $X = 2$, falls A. zwei Treffer mehr als B. erzielt — d.h. (E1,E1,E3) oder (E1,E3,E1) usw). Negative Realisationen kennzeichnen Dreiwurfspiele, bei denen B. häufiger trifft.

   Die beiden haben schon für fast alle Werte von X die Wahrscheinlichkeiten berechnet

x	3	2	1	0	-1	-2	-3
$\mathrm{P}(X = x)$	0.013824	0.107136	0.300960		0.175560	0.036456	

   (a) Leiten Sie die Werte für $\mathrm{P}(X = 0)$ und $\mathrm{P}(X = -3)$ her.
   (Falls Sie die Werte nicht berechnen können, dann rechnen Sie in den folgenden Unterpunkten mit $\mathrm{P}(X = 0) = 0.363320$ und $\mathrm{P}(X = -3) = 0.002744$)

   Max 3 Punkte, zwei Punkte für X=0 und 1 Punkte für X=-3.

   (b) Berechnen Sie den Erwartungswert von X und geben Sie die Formel zur Berechnung der Varianz von X an.

   MAX 2 Punkte, ein Punkt E-wert, ein Punkt Varianzformel

   (c) Interpretieren Sie den Erwartungswert des Ausgangs des Dreiwurfspiels.

   MAX 1Punkte.

   (d) Jeder der beiden Freunde soll nun angeben, wieviel er von den nächsten zehn Dreiwurfspielen gewinnt. A. sagt, dass er drei bis fünf Spiele gewinnen wird. Welchen möglichst kurzen Bereich muß B. angeben, damit er eine ähnlich zuverlässige Prognose abgibt?

   MAX 6 Punkte, 1 Punkte Prognoseniveau gefragt, 1 Punkt Binomialverteilung mit Parameter, 1 Punkt PN gerechnet, drei Punkte für PI für B.

3. Eines Tages erscheint ein Unbekannter auf dem Basketballplatz. Die beiden Freunde beobachten, wie er einmal elf Treffer hintereinander erzielt bis er zum erstenmal nicht trifft, einmal sieben und einmal neun. Nach den dreißig Würfen verschwindet der Basketballer wieder.

   (a) Wie hoch sollten die beiden Freunde die Trefferwahrscheinlichkeit des Unbekannten schätzen? Welche Annahme ist dazu notwendig.

MAX 3 Punkte, ein Punkte binomial, ein Punkt Schätzer, ein Punkt Annahme.

Aufgabe 4 (20 Punkte)

Diese Aufgabe besteht aus fünf Teilaufgaben mit insgesamt fünf Unterpunkten.

In einem Schmelzofen sollen Gold und Kupfer getrennt werden. Dazu muss der Ofen auf jeden Fall eine Temperatur von weniger als 1083 °C haben, da dies der Schmelzpunkt von Kupfer ist. Der Schmelzpunkt von Gold liegt bei 1064 °C.

Um die Temperatur im Schmelzofen zu bestimmen, wird ein Messsonde benutzt. Der Messwerte X ist normalverteilt mit Erwartungswert μ und Varianz $\sigma^2 = 25$. Dabei ist μ die tatsächliche Temperatur im Schmelzofen.

Der Schmelzofen ist betriebsbereit, wenn die Temperatur unter den Schmelzpunkt des Kupfers liegt. Die Entscheidung, ob der Ofen betriebsbreit ist, wird mit Hilfe der Messsonde bestimmt. Dabei wird so vorgegangen, dass der Ofen als betriebsbereit **erklärt** wird, wenn die Messsonde einen Messwert zwischen 1064 und 1070 °C anzeigt.

1. Wie groß ist die Wahrscheinlichkeit, dass bei diesem Vorgehen beide Metalle geschmolzen werden, d.h. die Temperatur mindestens 1083 °C beträgt?

   MAX 3 Punkte. Ein Punkt für die "maximale" Wahrscheinlichkeit, ein Punkt für die Standardisierung, ein Punkt für die weitere Berechnung.

2. Wie groß ist die Wahrscheinlichkeit, dass die Temperatur im Ofen bei diesem Vorgehen den Schmelzpunkt des Goldes nicht überschreitet.

   MAX 3 Punkte: Ein Punkte für die "maximale" Wahrscheinlichkeit, ein PUnkt für die Standardisierung, ein Punkt für die Weitere Berechnung.

3. Der Schmelzofenbetreiber erkennt, dass die Entscheidung über die Betriebsbereitschaft des Ofens Ähnlichkeiten mit Testentscheidung hat. Er betont, dass das Schmelzen des Kupfers zu verhindern ist. Als kritische Region sieht er den Bereich von 1064 und 1070 °C an.

   (a) Formulieren Sie die Hypothese des Testproblems Betriebsbereitschaft.
   MAX 1 Punkt: saubere Formulierung

   (b) Begründen Sie die Wahl der Hypothesen.
   MAX 2 Punkte: ein Punkt klare Argumentation, ein Punkt richtige Argumentation

   (c) Geben Sie die Prüfgröße und deren Verteilung unter der Nullhypothese an.
   MAX 2 Punkte, 1 Punkt Prüfgröße, 1 Punkt Verteilung

   (d) Berechnen Sie das Testniveau. Benutzen Sie dabei die vom Schmelzofenbesitzer angegebene kritische Region.
   MAX 2 Punkte, ein Punkt für die maximale Fehler erster Art, ein Punkt für die Verbindung zu 1.

   (e) Welchen Wert nimmt die Gütefunktion am Schmelzpunkt des Goldes an?
   MAX 2 Punkte, ein Punkt für die Gütefunktion, ein Punkt für die Verbindung zu 2.

4. Ist es möglich eine Wahrscheinlichkeit dafür anzugeben, dass die Temperatur im Hochofen zwischen 1064 und 1083 °C liegt?

   MAX 2 Punkte

5. Geben Sie einen Bereich an, in dem die Schmelzofentemperatur mit hoher Sicherheit liegt, wenn die Messsonde 1068 °C anzeigt.

   MAX 3 PUnkte, 2 Punkte fürs Konfidenzintervall, 1 Punkte für ein nicht beleidigendes $\alpha \in (0.8, 0.99)$.