Universität Karlsruhe (TH)
Institut für Stochastik
Dr. Bernhard Klar
Dipl.-Math. oec. Volker Baumstark

Name

Vorname

Matr.-Nr.:

Diplomvorprüfung Stochastik

21. September 2007

Diese Klausur hat bestanden, wer mindestens 16 Punkte erreicht.

Als Hilfsmittel sind nur zwei selbst erstellte DIN A4 Seiten sowie ein nicht programmierbarer Taschenrechner zugelassen!

Aufgabe	1 (7)	2 (6)	3 (7)	4 (6)	5 (7)	6 (7)	$\sum$ (40)
Punkte							
Korrektor							

Gesamtpunktzahl	Note

Aufgabe 1 (7 Punkte)

A, B, C seien Ereignisse mit $0 < P(A), P(B), P(C) < 1$. Man sagt, B zieht A an, falls $P(A|B) > P(A)$, und B stößt A ab, falls $P(A|B) < P(A)$. Zeigen Sie:

a) A zieht A an.

b) Gilt $A \cap B = \emptyset$, so zieht B das Ereignis A nicht an.

c) Zieht B das Ereignis A an, so zieht auch A das Ereignis B an.

d) Zieht B das Ereignis A an, so stößt B^c das Ereignis A ab.

e) Zieht B das Ereignis A an, und zieht C das Ereignis B an, so folgt i.Allg. nicht, dass C das Ereignis A anzieht.
Hinweis: Verwenden Sie Teil b).

Lösung zu Aufgabe 1

a) $P(A|A) = P(A)/P(A) = 1 > P(A)$. $\boxed{\text{1 P}}$

b) $P(A|B) = P(A \cap B)/P(B) = 0 < P(A)$. $\boxed{\text{0.5 P}}$

c) $P(B|A) = \frac{P(B\cap A)}{P(A)} = \frac{P(A\cap B)}{P(B)} \cdot \frac{P(B)}{P(A)} = P(A|B) \cdot \frac{P(B)}{P(A)} > P(A) \cdot \frac{P(B)}{P(A)} = P(B)$. $\boxed{\text{2 P}}$

d) Es gilt $\boxed{\text{2 P}}$

$$\begin{aligned} P(A|B^c) &= \frac{P(A \cap B^c)}{P(B^c)} = \frac{P(A) - P(A \cap B)}{P(B^c)} = \frac{P(A)}{P(B^c)} - \frac{P(A \cap B)}{P(B)} \cdot \frac{P(B)}{P(B^c)} \\ &< \frac{P(A)}{P(B^c)} - P(A) \cdot \frac{P(B)}{P(B^c)} = \frac{P(A)(1 - P(B))}{P(B^c)} = P(A). \end{aligned}$$

e) Gilt $A \cap C = \emptyset$, so zieht C das Ereignis A nicht an. Ein Beispiel ist $A = \{1, 2\}, B = \{2, 3\}, C = \{3, 4\}$ mit $P(\{i\}) = 1/10,\ i = 1, \dots, 4$. Dann gilt: $\boxed{\text{1.5 P}}$

$$\begin{aligned} P(A|B) = 1/2 &> P(A) = 1/5, \\ P(B|C) = 1/2 &> P(B) = 1/5, \\ P(A|C) = 0 &< P(C) = 1/5. \end{aligned}$$

Aufgabe 2 (6 Punkte)
In einem Team von 10 Fahrradfahrern sind 3 Fahrer gedopt. Rein zufällig wird vor einem Rennen eine Gruppe von 4 Testanden aus den 10 Fahrradfahrern ausgewählt und einer Dopingkontrolle unterzogen. Die Dopingkontrolle besteht aus einem Dopingtest, welcher bei einem nicht gedopten Fahrer stets das Ergebnis „nicht gedopt" liefert; bei einem gedopten Fahrer liefert der Test (unabhängig von den anderen Testergebnissen) mit Wahrscheinlichkeit $p = 0.9$ das Ergebnis „gedopt" und mit Wahrscheinlichkeit 0.1 das Ergebnis „nicht gedopt".

a) Es sei X die zufällige Anzahl gedopter Fahrer innerhalb der 4 Testanden. Geben Sie die Verteilung von X an.

b) Mit welcher Wahrscheinlichkeit wird bei allen drei gedopten Fahrern ein Test durchgeführt?

c) Mit welcher Wahrscheinlichkeit werden durch dieses Verfahren (Auswahl von 4 Testanden und anschließender Test) alle drei gedopten Fahrer des Dopings überführt?

d) Wird ein gedopter Fahrer durch die Dopingkontrolle des Dopings überführt, d.h. wird er als einer der vier Testanden ausgewählt und erhält er das Ergebnis „gedopt", so muss er 10000 Euro Strafe zahlen, d.h. er erhält eine Auszahlung von $Z = -10000$. Wird er nicht des Dopings überführt, so erhält er eine (durchschnittliche) Prämie von 5000 Euro, d.h. er erhält eine Auszahlung von $Z = 5000$. Berechnen Sie den Erwartungswert von Z.

Lösung zu Aufgabe 2

a) Es handelt sich um viermaliges Ziehen aus einer Urne mit 3 roten (gedopt) und 7 schwarzen (nicht gedopt) Kugeln ohne Zurücklegen. Es ist somit $\boxed{1\text{ P}}$

$$X \sim Hyp(4, 3, 7).$$

b) Es ist $\boxed{1.5\text{ P}}$

$$P(X = 3) = \frac{\binom{3}{3}\binom{7}{1}}{\binom{10}{4}} = \frac{1 \cdot 7 \cdot 4!}{10 \cdot 9 \cdot 8 \cdot 7} = \frac{1}{30}.$$

c) Es sei Y die Anzahl der Testergebnisse „gedopt". Die gesuchte W'keit berechnet sich zu $\boxed{1.5\text{ P}}$

$$\begin{aligned} P(\{Y = 3\} \cap \{X = 3\}) &= P(Y = 3 \,|\, X = 3) \cdot P(X = 3) \\ &= p^3 \cdot \frac{1}{30} = \left(\frac{9}{10}\right)^3 \cdot \frac{1}{30} = \frac{243}{10^4} = 0.0243. \end{aligned}$$

d) Die Wahrscheinlichkeit für einen gedopten Fahrer, des Dopings überführt zu werden, berechnet sich zu $\boxed{2\text{ P}}$

$$q := \frac{4}{10} \cdot \frac{9}{10} = 0.36.$$

Folglich ist

$$EZ = -q \cdot 10000 + (1 - q) \cdot 5000 = -3600 + 3200 = -400.$$

Aufgabe 3 (7 Punkte)
Es sei X $N(0,\sigma^2)$-verteilt mit $\sigma^2 > 0$, d.h. X besitzt die Dichte $f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{x^2}{2\sigma^2}\right)$ für $x \in \mathbb{R}$. Weiter sei $Y := e^X$.

a) Stellen Sie die Verteilungsfunktion von Y mit Hilfe der Verteilungsfunktion der Standardnormalverteilung dar und zeigen Sie, dass die Dichte von Y durch

$$g(y) \quad = \quad \frac{1}{y\sigma\sqrt{2\pi}} \exp\left(-\frac{(\log y)^2}{2\sigma^2}\right)$$

gegeben ist.

b) Bestimmen Sie den Erwartungswert von Y.
Hinweis: Transformationsformel für Erwartungswerte.

c) Bestimmen Sie den Median von Y.

d) Bestimmen Sie den Modalwert von Y, d.h. die Stelle, an der die Dichte von Y maximal wird.

Lösung zu Aufgabe 3

a) Sei $F(x) = \int_{-\infty}^{x} f(u)du = \Phi(x/\sigma)$ $(x \in \mathbb{R})$ die Verteilungsfunktion von X. Für $y \leq 0$ ist $G(y) = P(Y \leq y) = 0$ und somit auch $g(y) = 0$. Für $y > 0$ ist $\boxed{2.5 \text{ P}}$

$$\begin{aligned} G(y) &= P(Y \leq y) = P(e^X \leq y) = P(X \leq \log(y)) = F(\log(y)) = \Phi(\log(y)/\sigma), \\ g(y) &= G'(y) = \frac{1}{y} \cdot f(\log y) = \frac{1}{y\sigma\sqrt{2\pi}} \exp\left(-\frac{(\log y)^2}{2\sigma^2}\right). \end{aligned}$$

b) Mit der Transformationsformel für Erwartungswerte erhält man $\boxed{2 \text{ P}}$

$$\begin{aligned} EY = E\left(\mathrm{e}^X\right) &= \int_{-\infty}^{\infty} \mathrm{e}^x \frac{1}{\sigma\sqrt{2\pi}} \cdot \exp\left(-\frac{x^2}{2\sigma^2}\right) dx \\ &= \mathrm{e}^{\sigma^2/2} \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} \cdot \exp\left(-\frac{1}{2\sigma^2}(x^2 - 2\sigma^2 x + \sigma^4)\right) dx \\ &= \mathrm{e}^{\sigma^2/2}, \end{aligned}$$

da der Term unter dem Integral gerade die Dichte einer $N(\sigma^2, \sigma^2)$-Verteilung ist.

c) Gesucht ist die Stelle y^*, für die $G(y^*) = 1/2$ gilt. Nach a) gilt $G(1) = F(0) = \Phi(0) = 1/2$; somit ist $y^* = 1$ der Median von Y. $\boxed{1 \text{ P}}$

d) Mit g aus Teil a) ergibt sich $\boxed{1.5 \text{ P}}$

$$\begin{aligned} g'(y) &= \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(\log y)^2}{2\sigma^2}\right) \left(\frac{-1}{y^2} + \frac{1}{y} \cdot \frac{-\log y}{y\sigma^2}\right) \\ &= \frac{-1}{y^2\sigma^3\sqrt{2\pi}} \exp\left(-\frac{(\log y)^2}{2\sigma^2}\right) \left(\sigma^2 + \log y\right). \end{aligned}$$

Dies ist 0 für $\log y = -\sigma^2$, größer 0 für $\log y < -\sigma^2$, und kleiner 0 für $\log y > -\sigma^2$. Somit ist $y = \mathrm{e}^{-\sigma^2}$ eindeutige Maximalstelle von g, d.h. $\mathrm{e}^{-\sigma^2}$ ist der Modalwert von Y.

Aufgabe 4 (6 Punkte)
Die gemeinsame Dichte der Zufallsvariablen X und Y sei gegeben durch

$$h(x,y) = \begin{cases} 1+\alpha(1-2x)(1-2y), & 0 \le x,y \le 1 \\ 0, & \text{sonst} \end{cases},$$

wobei $-1 \le \alpha \le 1$ sei. Für $0 \le x, y \le 1$ gilt dann

$$P(X \le x, Y \le y) = xy\,(1+\alpha(1-x)(1-y)).$$

Hinweis: In den folgenden Aufgabenteilen muss kein Integral berechnet werden!

a) Zeigen Sie, dass X und Y jeweils gleichverteilt auf dem Intervall $[0,1]$ sind.

b) Berechnen Sie die Wahrscheinlichkeit $P(X > 1/2, Y > 1/2)$.

c) Bestimmen Sie die bedingte Wahrscheinlichkeit $P(X \le x | Y \le y) \quad (0 < x, y \le 1)$.

d) Für welche Werte von α sind X und Y stochastisch unabhängig?

e) Für welche Werte von α wird der Korrelationskoeffizient zwischen X und Y minimal bzw. maximal?
Hinweis: Verwenden Sie Teil a) und $\int_0^1 \int_0^1 xy\,h(x,y)\,dxdy = (1+\alpha/9)/4$.

Lösung zu Aufgabe 4

a) Wegen $P(X \le x) = P(X \le x, Y \le 1) = x$ für $x \in (0,1)$ ist X gleichverteilt auf $[0,1]$. Analog folgt dies für Y.

$\boxed{1\text{ P}}$

b) Es gilt (Skizze!)

$$\begin{aligned} P(X > 1/2, Y > 1/2) &= 1 - P(0 \le X \le 1, 0 \le Y \le 1/2) - P(0 \le X \le 1/2, 0 \le Y \le 1) \\ &\quad + P(0 \le X \le 1/2, 0 \le Y \le 1/2) \\ &= 1 - 1/2 - 1/2 + 1/4 \cdot (1+\alpha(1-1/2)(1-1/2)) \\ &= (1+\alpha/4)/4. \end{aligned}$$

$\boxed{2\text{ P}}$

c) Es gilt

$$P(X \le x | Y \le y) = \frac{P(X \le x, Y \le y)}{P(Y \le y)} = x\,(1+\alpha(1-x)(1-y)).$$

$\boxed{1\text{ P}}$

d) Genau für $\alpha = 0$ gilt

$$h(x,y) = 1 = f(x) \cdot g(y) \qquad (0 \le x, y \le 1),$$

$\boxed{1\text{ P}}$

wobei f mit $f(x) = 1$ $(0 \le x \le 1)$ bzw. g mit $g(y) = 1$ $(0 \le y \le 1)$ eine Dichte von X bzw. Y ist. Genau für $\alpha = 0$ sind X und Y daher stochastisch unabhängig.

e) Nach Teil a) hängen die Momente von X bzw. Y nicht von α ab. Nach dem Hinweis gilt $E(XY) = (1+\alpha/9)/4$; dieser Wert und somit die Kovarianz $C(X,Y) = E(XY) - EXEY$ bzw. die Korrelation $\rho(X,Y) = C(X,Y)/\sqrt{V(X)V(Y)}$ wird minimal für $\alpha = -1$ und maximal für $\alpha = +1$. $\boxed{1 \text{ P}}$

Bemerkung: Wegen $EX = EY = 1/2$ ist die minimale bzw. maximale Kovarianz gegeben durch -1/36 bzw. +1/36. Aus $V(X) = V(Y) = 1/12$ folgt weiter, dass die minimale bzw. maximale Korrelation -1/3 bzw. +1/3 ist.

Aufgabe 5 (7 Punkte)
Es seien $X_1, \ldots, X_n, \ldots$ uanbhängige Zufallsvariablen, wobei X_i $N(\mu, \sigma_i^2)$-verteilt ist $(i = 1, 2, \ldots)$. Dabei ist $\mu \in \mathbb{R}$ unbekannt, die $\sigma_i^2 > 0$ sind bekannt.

a) Zeigen Sie, dass $T_n = \frac{1}{n}\sum_{i=1}^n X_i$ ein erwartungstreuer Schätzer für μ ist.

b) Berechnen Sie die mittlere quadratische Abweichung $\mathrm{MQA}_{T_n}(\mu) = E_\mu(T_n - \mu)^2$.

c) Geben Sie eine hinreichende Bedingung dafür an, dass $(T_n)_{n\geq 1}$ eine konsistente Schätzfolge für μ ist.

d) Geben Sie die gemeinsame Dichte von $X_1, \ldots, X_n$ an, und bestimmen Sie den Maximum-Likelihood-Schätzer $\hat{\mu}_n$ für μ.

Lösung zu Aufgabe 5

a) Es gilt $\boxed{1\ \text{P}}$

$$E_\mu T_n = \frac{1}{n}\sum_{i=1}^n E_\mu(X_i) = \frac{1}{n}\sum_{i=1}^n \mu = \mu.$$

b) Unter Verwendung von a) rechnen wir $\boxed{1.5\ \text{P}}$

$$MQA_{T_n}(\mu) = E_\mu(T_n - E_\mu T_n)^2 = V_\mu(T_n) = \frac{1}{n^2}\sum_{i=1}^n V_\mu(X_i) = \frac{1}{n^2}\sum_{i=1}^n \sigma_i^2.$$

c) Unter Verwendung von a) und b) gilt nach der Tschebyscheff-Ungleichung für $\varepsilon > 0$ $\boxed{1.5\ \text{P}}$

$$P_\mu(|T_n - \mu| \geq \varepsilon) = P_\mu(|T_n - E_\mu T_n| \geq \varepsilon) \leq \frac{V_\mu(T_n)}{\varepsilon^2} = \frac{1}{\varepsilon^2 n^2}\sum_{i=1}^n \sigma_i^2.$$

Gilt nun z.B. $\frac{1}{n}\sum_{i=1}^n \sigma_i^2 \leq K$, so folgt

$$P_\mu(|T_n - \mu| \geq \varepsilon) \leq \frac{K}{n\varepsilon^2} \to 0$$

für $n \to \infty$, d.h. (T_n) ist konsistent für μ.

d) Zunächst gilt $\boxed{3\ \text{P}}$

$$L_x(\mu) := f_\mu(x_1, \ldots, x_n) = \prod_{i=1}^n \left\{\frac{1}{\sqrt{2\pi}\sigma_i}\right\} \exp\left(-\frac{1}{2}\sum_{i=1}^n \left(\frac{x_i - \mu}{\sigma_i}\right)^2\right).$$

Folglich gilt

$$\frac{d}{d\mu}\log L_x(\mu) = \frac{d}{d\mu}\left(-\frac{1}{2}\sum_{i=1}^n \left(\frac{x_i - \mu}{\sigma_i}\right)^2\right) = \sum_{i=1}^n \frac{x_i - \mu}{\sigma_i^2} = \sum_{i=1}^n \frac{x_i}{\sigma_i^2} - \mu\sum_{i=1}^n \frac{1}{\sigma_i^2} = 0$$

genau dann, wenn

$$\mu = \frac{\sum_{i=1}^n x_i/\sigma_i^2}{\sum_{i=1}^n 1/\sigma_i^2}.$$

Dies ist wegen

$$\frac{d^2}{d\mu^2} \log L_x(\mu) = -\sum_{i=1}^{n} \frac{1}{\sigma_i^2} < 0$$

die Maximalstelle von $\log L_x$. Somit ist $\hat{\mu}_n = \frac{\sum_{i=1}^n X_i/\sigma_i^2}{\sum_{i=1}^n 1/\sigma_i^2}$ der gesuchte ML-Schätzer für μ.

Aufgabe 6 (7 Punkte)

Es seien $X_1, \dots, X_8$ unabhängig und identisch verteilt mit $P(X_1 = 1) = \vartheta$, $P(X_1 = 0) = 1 - \vartheta$. Für den unbekannten Parameter ϑ gelte $\vartheta \in [1/2, 1) =: \Theta$. Zum Testen der Nullhypothese $H_0 : \vartheta = 1/2$ gegen die einseitige Alternative $H_1 : \vartheta > 1/2$ betrachte man die folgenden Tests: Test A verwirft H_0 bei 8 Treffern, d.h. wenn $X_i = 1$ für $i = 1, \dots, 8$ gilt, ansonsten wird H_0 nicht verworfen. Test B verwirft H_0 bei mindestens 7 Treffern, ansonsten wird H_0 nicht verworfen.

a) Sind Test A bzw. B Tests zum Niveau $\alpha = 0.05$?

b) Geben Sie die Gütefunktionen der beiden Tests an.

c) Wie groß ist der Fehler 2. Art bei den beiden Tests, falls in Wirklichkeit $\vartheta = 0.75$ gilt?

d) $X_1, \dots, X_8$ seien wie oben verteilt, es sei aber $\Theta = (0, 1)$. Geben Sie einen sinnvollen Test zum Niveau 0.05 für $H_0 : \vartheta = 1/2$ gegen die zweiseitige Alternative $H_2 : \vartheta \neq 1/2$ an.

Lösung zu Aufgabe 6

a) Unter H_0 gilt $X_1 + \ldots + X_8 \sim Bin(8, 1/2)$. Somit beträgt der Fehler 1. Art von Test A bzw. B $\boxed{2.5 \text{ P}}$

$$\begin{aligned} P_{1/2}(\text{Test A verwirft } H_0) &= P_{1/2}\left(\sum_i X_i = 8\right) = \frac{1}{2^8} = 0.0039, \\ P_{1/2}(\text{Test B verwirft } H_0) &= P_{1/2}\left(\sum_i X_i \geq 7\right) = 8\frac{1}{2^8} + \frac{1}{2^8} = 0.035. \end{aligned}$$

Beide Tests sind also Tests zum Niveau 0.05.

b) Wegen $X_1 + \ldots + X_8 \sim Bin(8, \vartheta)$ ist die Gütefunktion von Test A bzw. B gegeben durch $\boxed{1.5 \text{ P}}$

$$\begin{aligned} G_A(\vartheta) &= P_\vartheta(\text{Test A verwirft } H_0) = P_\vartheta\left(\sum_i X_i = 8\right) = \vartheta^8, \\ G_B(\vartheta) &= P_\vartheta(\text{Test B verwirft } H_0) = P_\vartheta\left(\sum_i X_i \geq 7\right) = 8\vartheta^7(1-\vartheta) + \vartheta^8 = \vartheta^7(8 - 7\vartheta). \end{aligned}$$

c) Gilt $\theta = 0.75$, so ist der Fehler 2. Art bei Test A bzw. B gegeben durch $\boxed{1.5 \text{ P}}$

$$\begin{aligned} P_{3/4}(\text{Test A verwirft } H_0 \text{ nicht}) &= 1 - G_A(3/4) = 1 - (3/4)^8 = 0.90, \\ P_{3/4}(\text{Test B verwirft } H_0 \text{ nicht}) &= 1 - G_B(3/4) = 1 - (3/4)^7(8 - 21/4) = 0.63. \end{aligned}$$

d) Ein sinnvoller Test für $H_0 : \vartheta = 1/2$ gegen $H_2 : \vartheta \neq 1/2$ verwirft H_0, falls $\sum_i X_i = 0$ oder $\sum_i X_i = 8$. Nach a) beträgt der Fehler 1. Art für diesen Test $2 \cdot 0.0039 = 0.0078$, es ist also ein Test zum Niveau 0.05. $\boxed{1.5 \text{ P}}$
Ein Test, der H_0 verwirft, falls $\sum_i X_i \leq 1$ oder $\sum_i X_i \geq 7$, ist dagegen kein Test zum Niveau 0.05.