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

Name

Vorname

Matr.-Nr.:

Scheinklausur (Nachklausur) Stochastik 1 für Studierende des Lehramts und der Diplom-Pädagogik

04. Juni 2007

Diese Klausur hat bestanden, wer mindestens 14 Punkte erreicht.

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

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

Gesamtpunktzahl	Note

Aufgabe 1 (6 Punkte)
Louis Pasteur impfte im Jahr 1880 sechs Hühner mit einem Serum gegen Cholera. Als er diese und sechs ungeimpfte Hühner mit Cholera-Erregern infizierte, waren am nächsten Tag sechs Hühner tot.

a) Wäre die Impfung völlig wirkungslos gewesen, so wären die sechs toten Hühner eine rein zufällige Auswahl aus allen 12 infizierten Hühnern gewesen. Wie groß ist unter dieser Annahme die Wahrscheinlichkeit, daß die sechs toten Hühner gerade die sechs ungeimpften sind? (Dieser Fall trat tatsächlich ein.)

b) Sei X die Anzahl der geimpften Hühner unter den 6 toten Hühnern. Geben Sie unter der Annahme, daß die Impfung ohne jede Wirkung war, die Verteilung von X an.

c) Berechnen Sie unter der Annahme, daß die Impfung ohne jede Wirkung war, die Wahrscheinlichkeit dafür, daß sich unter den 6 toten Hühnern mindestens 2 geimpfte befinden.

Lösung zu Aufgabe 1

a) Es gibt $\boxed{2 \text{ P}}$

$$\binom{12}{6} = \frac{12!}{6!6!} = 924$$

Möglichkeiten, 6 aus 12 Hühnern auszuwählen. Bei einer rein zufälligen Auswahl von 6 aus 12 Hühnern ist die Wahrscheinlichkeit, daß die sechs toten Hühner gerade die sechs ungeimpften sind, folglich

$$\frac{1}{\binom{12}{6}} = \frac{1}{924} = 0.00108.$$

b) Unter der Annahme, dass die Impfung wirkungslos ist, sind die 6 toten Hühner eine zufällige Auswahl **ohne** Zurücklegen aus 12 Hühnern. Die Anzahl X der geimpften Hühner unter den 6 toten Hühnern ist dann hypergeometrisch verteilt ($X \sim Hyp(6,6,6)$) mit $\boxed{2 \text{ P}}$

$$P(X = k) = \frac{\binom{6}{k}\binom{6}{6-k}}{\binom{12}{6}}, \qquad k = 0, \ldots, 6.$$

c) Nach b) ist $\boxed{2 \text{ P}}$

$$P(X \geq 2) = 1 - P(X = 0) - P(X = 1) = 1 - \frac{1}{924} - \frac{6 \cdot 6}{924} = \frac{887}{924} = 0.960.$$

Aufgabe 2 (7 Punkte)

Ein zweistufiges Experiment werde wie folgt durchgeführt. Zunächst wird ein fairer Würfel geworfen und dabei die Augenzahl $i \in \{1, \ldots, 6\}$ erzielt. Nun werden in eine Urne i rote und $6 - i$ schwarze Kugeln gelegt. Abschließend werden fünf Kugeln aus dieser Urne ohne Zurücklegen gezogen.

a) Es sei A_i das Ereignis, dass die Augenzahl „i" gewürfelt wird und B das Ereignis, dass alle fünf gezogenen Kugeln rot sind. Bestimmen Sie die bedingten Wahrscheinlichkeiten $P(B|A_i)$ für $i = 1, \ldots, 6$.

b) Mit welcher Wahrscheinlichkeit sind alle fünf gezogenen Kugeln rot?

c) Wie groß ist die bedingte Wahrscheinlichkeit, dass eine „6" gewürfelt wurde, falls alle fünf gezogenen Kugeln rot sind?

d) Bei einem Spiel erhält man 5 Euro Gewinn (Auszahlung +5), wenn alle fünf gezogenen Kugeln rot sind, ansonsten muss man einen Euro zahlen (Auszahlung -1). Sei X die (zufällige) Auszahlung. Bestimmen Sie die Wahrscheinlichkeitsfunktion (Zähldichte) und den Erwartungswert von X.

Lösung zu Aufgabe 2

a) Das Ereignis B tritt nur ein, falls die gewürfelte Augenzahl 5 oder 6 beträgt. Es gelten also $P(B|A_i) = 0$ für $i = 1, \ldots, 4$ und $\boxed{2\text{ P}}$

$$P(B|A_6) = 1, \qquad P(B|A_5) = \frac{1}{\binom{6}{5}} = \frac{1}{6}.$$

b) Die Formel von der totalen Wahrscheinlichkeit liefert $\boxed{1.5\text{ P}}$

$$P(B) = \frac{1}{6} \cdot 1 + \frac{1}{6} \cdot \frac{1}{6} = \frac{7}{36}.$$

c) Es gelten $\boxed{1.5\text{ P}}$

$$P(B|A_6) = 1, \qquad P(A_6) = \frac{1}{6}.$$

Die gesuchte bedingte Wahrscheinlichkeit berechnet sich zu

$$P(A_6|B) = \frac{P(B|A_6) \cdot P(A_6)}{P(B)} = \frac{1}{6} \cdot \frac{36}{7} = \frac{6}{7}.$$

d) Nach b) ist die Wahrscheinlichkeitsfunktion von X gegeben durch $\boxed{2\text{ P}}$

$$P(X = 5) = \frac{7}{36}, \qquad P(X = -1) = \frac{29}{36}.$$

Damit ist

$$EX = 5 \cdot P(X = 5) + (-1) \cdot P(X = -1) = \frac{1}{6}.$$

Aufgabe 3 (7 Punkte)
Die Zufallsvariable X sei standardnormalverteilt, die Wahrscheinlichkeitsdichte von X ist also gegeben durch $f(x) = \frac{1}{\sqrt{2\pi}} \exp(-x^2/2),\ x \in \mathbb{R}$.

a) Zeigen Sie, dass $|X|$ die Dichte

$$f(x) := \sqrt{\frac{2}{\pi}} \cdot \exp\left(-\frac{x^2}{2}\right), \qquad x > 0,$$

und $f(x) = 0$ für $x \leq 0$ besitzt.

b) Sind X und $|X|$ unabhängig?

c) Berechnen Sie den Erwartungswert und die Varianz von $|X|$.
Hinweis: Verwenden Sie für die Berechnung des Erwartungswertes die Substitution $y := x^2/2$.

d) Berechnen Sie die Kovarianz von X und $|X|$.

Lösung zu Aufgabe 3

a) Für $t \geq 0$ gilt $\boxed{2\text{ P}}$

$$\begin{aligned} P(|X| \leq t) = P(-t \leq X \leq t) &= \frac{1}{\sqrt{2\pi}} \int_{-t}^{t} \exp\left(-\frac{x^2}{2}\right) dx \\ &= 2 \cdot \frac{1}{\sqrt{2\pi}} \int_0^t \exp\left(-\frac{x^2}{2}\right) dx = \sqrt{\frac{2}{\pi}} \int_0^t \exp\left(-\frac{x^2}{2}\right) dx. \end{aligned}$$

Hieraus folgt die Behauptung.

b) Die Zufallsvariablen X und $|X|$ sind nicht unabhängig. Es gelten beispielsweise $\boxed{1\text{ P}}$

$$P(-1 < X < 1,\, |X| > 1) = 0 \qquad \text{und} \qquad P(-1 < X < 1) \cdot P(|X| > 1) > 0.$$

c) Der Erwartungswert von $|X|$ berechnet sich mittels Substitution $y := x^2/2$ zu $\boxed{2.5\text{ P}}$

$$\begin{aligned} E|X| &= \int_0^\infty x \cdot \sqrt{\frac{2}{\pi}} \cdot \exp\left(-\frac{x^2}{2}\right) dx \\ &= \sqrt{\frac{2}{\pi}} \int_0^\infty \exp(-y)\, dy = \sqrt{\frac{2}{\pi}}. \end{aligned}$$

Ferner gilt

$$E(|X|^2) = EX^2 = 1,$$

folglich

$$V(|X|) = E(|X|^2) - E(|X|)^2 = 1 - \frac{2}{\pi} = \frac{\pi - 2}{\pi}.$$

d) Zunächst gilt $\boxed{1.5\text{ P}}$

$$E\left(X \cdot |X|\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x \cdot |x| \cdot \exp\left(-\frac{x^2}{2}\right) dx = 0,$$

da der Integrand punktsymmetrisch ist. Wegen $EX = 0$ erhalten wir

$$C(X, |X|) = E(X \cdot |X|) - EX \cdot E|X| = 0.$$

Aufgabe 4 (7 Punkte)
X und Y seien die zufälligen Wartezeiten von zwei Kunden A und B, die an unterschiedlichen Kassen stehen. Wir nehmen an, dass X und Y stochastisch unabhängig und jeweils exponentialverteilt mit Parameter 1 sind; die Wahrscheinlichkeitsdichte von X bzw. Y ist also $f(x) = \exp(-x)$ für $x > 0$, und $f(x) = 0$ für $x \leq 0$.

a) Mit welcher Wahrscheinlichkeit ist die Wartezeit von Kunde A größer als 2?

b) Geben Sie die gemeinsame Dichte von X und Y, also die Wahrscheinlichkeitsdichte des Vektors (X, Y), an.

c) Wie groß ist die Wahrscheinlichkeit, dass die Wartezeit sowohl von Kunde A wie auch von Kunde B größer als 2 ist?

d) Wie groß ist die Wahrscheinlichkeit, dass die längere der beiden Wartezeiten größer als 2 ist?

e) Wie groß ist die Wahrscheinlichkeit, dass Kunde B mindestens doppelt so lange wie Kunde A wartet?

Lösung zu Aufgabe 4

a) $P(X > 2) = \int_2^\infty f(x)dx \; = \; |-e^{-x}|_2^\infty \; = \; e^{-2}$ $(= 0.135)$.

b) Wegen der Unabhängigkeit von X und Y ist 1 P

$$h(x,y) \quad = \quad f(x) \cdot f(y) \quad = \begin{cases} e^{-x-y}, & x, y > 0 \\ 0, & \text{sonst} \end{cases}$$ 1 P

(eine) Dichte von (X, Y).

c) Unter Verwendung von b) folgt

$$P(X > 2, Y > 2) = \int_2^\infty \int_2^\infty h(x,y)d(x,y) \quad = \quad \left(\int_2^\infty f(x)dx \right)^2 \quad = \quad e^{-4} \; (= 0.018).$$ 1 P

Alternativ: Die Unabhängigkeit von X, Y liefert

$$P(X > 2, Y > 2) = P(X > 2) \cdot P(Y > 2) \; = \; e^{-4}.$$

d) Es ist

$$\begin{aligned} P(\max(X,Y) > 2) \quad &= \quad 1 - P(X \leq 2, Y \leq 2) \; = \; 1 - P(X \leq 2) \cdot P(Y \leq 2) \\ &= \quad 1 - \left(1 - e^{-2}\right)^2 \; = \; e^{-2}\left(2 - e^{-2}\right) \; (= 0.252). \end{aligned}$$ 1.5 P

Alternative Lösung:

$$\begin{aligned} P(\max(X,Y) > 2) \quad &= \quad P(X > 2 \text{ oder } Y > 2) \\ &= \quad P(X > 2) + P(Y > 2) - P(X > 2 \text{ und } Y > 2)) \; = \; e^{-2}\left(2 - e^{-2}\right). \end{aligned}$$

e) Es ist mit $B = \{(x,y) \in \mathcal{R}^2 : y \geq 2x\}$

$$\begin{aligned} P(Y \geq 2X) \quad &= \quad P((X,Y) \in B) \; = \; \int_B h(x,y) \; d(x,y) \\ &= \quad \int_0^\infty \left(\int_{2x}^\infty f(y) \; dy \right) f(x) \; dx = \int_0^\infty e^{-2x} e^{-x} dx = \left| -\frac{e^{-3x}}{3} \right|_0^\infty = \frac{1}{3}. \end{aligned}$$ 2.5 P

Aufgabe 5 (6 Punkte)
Die Zufallsvariablen $X_1, \dots, X_n$ seien unabhängig und jeweils $N(0, \vartheta)$-verteilt, dabei ist $\vartheta > 0$ unbekannt. Die Dichte von X_1 ist also gegeben durch

$$f(x; \vartheta) = \frac{1}{\sqrt{2\pi\vartheta}} \exp\left(-\frac{x^2}{2\vartheta}\right), \qquad x \in \mathbb{R}.$$

a) Bestimmen Sie den Maximum-Likelihood-Schätzer für den unbekannten Parameter ϑ.

b) Welcher der beiden Schätzer $S_n = \frac{1}{n}\sum_{i=1}^n X_i^2$ und $T_n = \frac{1}{n-1}\sum_{i=1}^n X_i^2$ ist erwartungstreu, welcher ist asymptotisch erwartungstreu?

Lösung zu Aufgabe 5

a) Für $x = (x_1, \dots, x_n)$ gilt:

$$L_x(\vartheta) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\vartheta}} \exp\left(-\frac{x_i^2}{2\vartheta}\right) = \left(\frac{1}{2\pi\vartheta}\right)^{n/2} \exp\left(-\frac{1}{2\vartheta}\sum_{i=1}^n x_i^2\right).$$

$\boxed{4 \text{ P}}$

Folglich ist

$$\log L_x(\vartheta) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log\vartheta - \frac{1}{2\vartheta}\sum_{i=1}^n x_i^2$$

und

$$\frac{d}{d\vartheta}\log L_x(\vartheta) = -\frac{n}{2\vartheta} + \frac{1}{2\vartheta^2}\sum_{i=1}^n x_i^2 = 0 \quad \Leftrightarrow \quad \vartheta = \frac{1}{n}\sum_{i=1}^n x_i^2.$$

Weiter gilt mit $\hat{\vartheta}_n := \frac{1}{n}\sum_{i=1}^n x_i^2$

$$\frac{d^2}{d\vartheta^2}\log L_x(\hat{\vartheta}_n) = \frac{n}{2\hat{\vartheta}_n^2} - \frac{n}{\hat{\vartheta}_n^2} < 0.$$

Somit liegt eine Maximalstelle vor, und

$$\hat{\vartheta}_n(X_1, \dots, X_n) = \frac{1}{n}\sum_{i=1}^n X_i^2$$

ist der gesuchte Maximum-Likelihood-Schätzer für ϑ.

b) Aus $E_\vartheta X_1 = 0$ folgt $E_\vartheta X_1^2 = V_\vartheta(X_1) = \vartheta$ und somit

$$E_\vartheta S_n = \frac{1}{n}\sum_{i=1}^n E_\vartheta\left(X_i^2\right) = \vartheta$$

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

für alle $\vartheta > 0$. S_n ist also erwartungstreu. Wegen

$$E_\vartheta T_n = \frac{n}{n-1}E_\vartheta S_n = \frac{n}{n-1}\vartheta$$

ist T_n asymptotisch erwartungstreu, aber nicht erwartungstreu.

Aufgabe 6 (7 Punkte)

Das durchschnittliche Geburtsgewicht von Kindern von Nichtraucherinnen beträgt 3500g. In einer Studie soll gezeigt werden, dass das Geburtsgewicht von Kindern, deren Mütter starke Raucherinnen sind, im Durchschnitt niedriger ist als 3500g. Dazu wird das Geburtsgewicht von n Kindern gemessen, deren Mütter während der Schwangerschaft täglich mehr als 20 Zigaretten rauchten. Wir nehmen dabei an, dass die Geburtsgewichte als Realisierungen unabhängiger Zufallsvariablen mit gleicher Normalverteilung $N(\mu, 350^2)$ modelliert werden können.

a) Formulieren Sie eine geeignete Hypothese und Alternative und geben Sie ein Testverfahren basierend auf der Testgröße $T_n = \sqrt{n}(\bar{X}_n - 3500)/350$ an, das geeignet ist, die obige Behauptung „statistisch nachzuweisen".
(Hierbei ist $\bar{X}_n = 1/n \sum_{i=1}^{n} X_i$ der arithmetische Mittelwert der Geburtsgewichte X_i).

   Wie muss dabei der kritische Wert gewählt werden, wenn der Stichprobenumfang $n = 20$ und das Signifikanzniveau $\alpha = 5\%$ betragen soll?

b) Zu welchem Ergebnis kommt der Test, wenn bei den 20 Kindern ein durchschnittliches Geburtsgewicht von $\bar{x}_{20}$ =3200g gemessen wurde?

c) Mit welcher Wahrscheinlichkeit wird mit der in a) ermittelten Entscheidungsregel eine Fehlentscheidung getroffen, wenn im Fall $n = 20$ das durchschnittliches Geburtsgewicht in Wirklichkeit 3400g beträgt?

Einige Werte der Verteilungsfunktion Φ der Normalverteilung $N(0, 1)$:

t	0	0.37	1.28	1.65	1.96	2.33	2.58
$\Phi(t)$	0.5	0.64	0.9	0.95	0.975	0.99	0.995

Lösung zu Aufgabe 6

a) Man muss als Hypothese $H_0 : \mu \geq 3500$ und als Alternative $H_1 : \mu < 3500$ wählen. Nur dann ist bei Ablehnung von H_0 die Aussage "bei Kindern von starken Raucherinnen ist das Geburtsgewicht im Mittel niedriger als 3500g" statistisch gesichert. 2.5 P

   Da kleine Werte von T_n gegen H_0 sprechen, und da T_n unter H_0 standardnormalverteilt ist, muss H_0 zum Niveau $\alpha = 0.05$ verworfen werden, wenn $T_n \leq z_{0.05} = \Phi^{-1}(0.05) = -1.65$ ist. Andernfalls besteht kein Einwand gegen H_0.

b) Wegen $T_{20} = \sqrt{20}(\bar{X}_{20} - 3500)/350 = -3.83 < z_{0.05}$ wird H_0 wird auf dem 5%–Niveau abgelehnt. 1.5 P

c) Ist in Wirklichkeit $\mu = 3400$, so liegt eine Fehlentscheidung vor, wenn H_0 nicht verworfen wird. Da jetzt die Größe $\frac{\sqrt{n}(\bar{X}_n - 3400)}{350}$ standardnormalverteilt ist, gilt 3 P

$$
\begin{aligned}
P_\mu(\text{Test verwirft } H_0 \text{ nicht}) &= P_\mu\left(\frac{\sqrt{n}(\bar{X}_n - 3500)}{350} > z_{0.05}\right) \\
&= 1 - P_\mu\left(\frac{\sqrt{n}(\bar{X}_n - 3500)}{350} \leq z_{0.05}\right)
\end{aligned}
$$

$$\begin{aligned} &= 1 - P_\mu\left(\frac{\sqrt{20}(\bar{X}_{20} - 3400)}{350} \le z_{0.05} + \frac{\sqrt{20}(3500 - 3400)}{350}\right) \\ &= 1 - \Phi(-1.65 + 1.28) \;=\; 1 - \Phi(-0.37) \;=\; \Phi(0.37) \;=\; 0.64. \end{aligned}$$