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

Name
Vorname
Matr.-Nr.:

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

09. März 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 (6)	3 (7)	4 (7)	5 (7)	6 (7)	$\sum$ (40)
Punkte							
Korrektor							

Gesamtpunktzahl	Note

Aufgabe 1 (6 Punkte)
Die Keimfähigkeit einer Sorte einer gartenbaulichen Kultur sei 80%, es gelte also P(„Samen keimt“) = 0.8. Die Keimfähigkeit eines Samens ist unabhängig von der Keimung jedes anderen Samens. Es werden 4 (durchnummerierte) Samen untersucht. Berechnen Sie die Wahrscheinlichkeiten der folgenden Ereignisse:

a) Alle 4 Samen keimen.

b) Der 1. und der 2. Samen keimen, der 3. und der 4. Samen keimen nicht.

c) Zwei Samen keimen und zwei keimen nicht.

d) Ein Samen keimt und drei keimen nicht.

e) Alle 4 Samen keimen nicht.

f) Mindestens ein Samen keimt.

g) Der 1. Samen keimt unter der Bedingung, dass mindestens ein Samen keimt.

Lösung zu Aufgabe 1
Wir definieren die Ereignisse A_i =„der i. Samen keimt“, $i = 1, \ldots, 4$. A_1, A_2, A_3, A_4 sind stochastisch unabhängig, und es gilt:

$$P(A_i) = 0.8, \quad P(A_i^c) = 1 - 0.8 = 0.2 \quad (i = 1, \ldots, 4).$$

Bezeichnet X die Anzahl der keimenden Samen unter den 4, so gilt $X \sim Bin(4, 0.8)$.

a) $P(A_1 \cap A_2 \cap A_3 \cap A_4) = P(A_1) \cdot P(A_2) \cdot P(A_3) \cdot P(A_4) = (0.8)^4 = 0.4096$ $\boxed{\text{1 P}}$

b) $P(A_1 \cap A_2 \cap A_3^c \cap A_4^c) = P(A_1) \cdot P(A_2) \cdot (1 - P(A_3)) \cdot (1 - P(A_4)) = 0.0256$ $\boxed{\text{1 P}}$

c) $P(X = 2) = \binom{4}{2} 0.8^2 \cdot 0.2^2 = 0.1536$ $\boxed{\text{1 P}}$

d) $P(X = 1) = \binom{4}{1} 0.8^1 \cdot 0.2^3 = 0.0256$ $\boxed{\text{1 P}}$

e) $P(X = 0) = 0.2^4 = 0.0016$ $\boxed{\text{0.5 P}}$

f) $P(X \geq 1) = 1 - P(X = 0) = 0.9984$ $\boxed{\text{0.5 P}}$

g) $P(A_1 | \{X \geq 1\}) = \dfrac{P(A_1 \cap \{X \geq 1\})}{P(X \geq 1)} = \dfrac{P(A_1)}{P(X \geq 1)} = \dfrac{0.8}{0.9984} = 0.8012$ $\boxed{\text{1 P}}$

Aufgabe 2 (6 Punkte)

Tom will im Karlsruher Zoo zur Eisbärenanlage. Er kann hierfür nach rechts (richtiger Weg) oder nach links (falscher Weg) gehen. Fragt er einen Besucher des Zoos mit Tageskarte nach dem Weg dorthin, so erhält er mit Wahrscheinlichkeit 2/3 die richtige Antwort und mit Wahrscheinlichkeit 1/3 eine falsche Antwort. Fragt er einen Dauerkartenbesitzer nach dem Weg dorthin, so erhält er stets die richtige Antwort. Antworten und Eintrittskarten von verschiedenen Personen sind unabhängig. Die Wahrscheinlichkeit, dass eine zufällig angesprochene Person eine Dauerkarte besitzt, sei 1/10.

a) Zeichnen Sie das zu dem 2-stufigen Experiment gehörende Baumdiagramm, und tragen Sie die Start- und Übergangswahrscheinlichkeiten ein.

b) Tom fragt einen Besucher $B1$ nach dem Weg zur Eisbärenanlage. Mit welcher Wahrscheinlichkeit erhält er eine richtige Antwort?

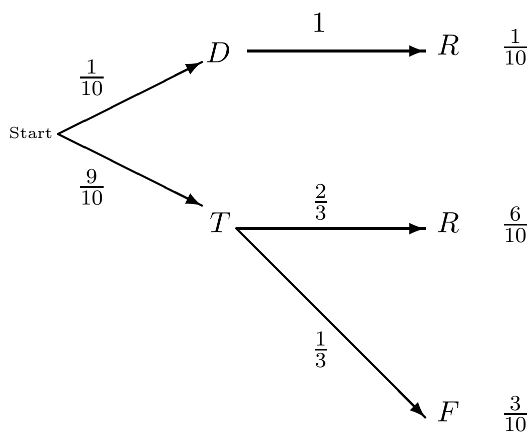
c) Wie groß ist die bedingte Wahrscheinlichkeit, dass der Besucher $B1$ eine Dauerkarte besitzt, wenn er die richtige Antwort gegeben hat?

d) Tom fragt einen weiteren Besucher $B2$ nach dem Weg zur Eisbärenanlage. Mit welcher Wahrscheinlichkeit geben $B1$ und $B2$ dieselbe Antwort?

Lösung zu Aufgabe 2

a) $D :=$ Dauerkarte, $T :=$ Tageskarte, $R :=$ Antwort richtig, $F :=$ Antwort falsch,

1.5 P



b) Der Index 1 beziehe sich auf $B1$. Aus obigem Diagramm lesen wir ab

$$P(R_1) = P(D_1 \cap R_1) + P(T_1 \cap R_1) = \frac{1}{10} + \frac{6}{10} = \frac{7}{10}.$$

1.5 P

c) Unser Diagramm und Teil a) liefern

$$P(D_1|R_1) = \frac{P(D_1 \cap R_1)}{P(R_1)} = \frac{1/10}{7/10} = \frac{1}{7}.$$

1 P

d) Der Index 2 beziehe sich auf $B2$. Es sei C das Ereignis, dass $B1$ und $B2$ dieselbe Antwort geben. Analog zu a) gilt

$$P(R_2) = \frac{7}{10}.$$

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

Die gesuchte Wahrscheinlichkeit berechnet sich wegen der Unabhängigkeit von R_1 und R_2 bzw. F_1 und F_2 zu

$$\begin{aligned} P(C) &= P(R_1 \cap R_2) + P(F_1 \cap F_2) \\ &= P(R_1)P(R_2) + P(F_1)P(F_2) = \left(\frac{7}{10}\right)^2 + \left(\frac{3}{10}\right)^2 = \frac{49+9}{100} = \frac{58}{100}. \end{aligned}$$

Aufgabe 3 (7 Punkte)
Sei X exponentialverteilt mit Parameter $\lambda > 0$, d.h. X besitze die Dichte

$$f(x) = \lambda \mathrm{e}^{-\lambda x}, \qquad x > 0.$$

a) Bestimmen Sie den Median von X.

b) Berechnen Sie die Verteilungsfunktion und Dichte der Zufallsvariablen $Y := e^{\alpha X}$, $\alpha > 0$.

c) Berechnen Sie den Erwartungswert von Y im Fall $0 < \alpha < \lambda$.

Lösung zu Aufgabe 3

a) Die Verteilungsfunktion F von X ist gegeben durch

$$F(x) = 1 - \mathrm{e}^{-\lambda x}, \qquad x > 0, \qquad \boxed{2\text{ P}}$$

und $F(x) = 0$ für $x \leq 0$. Wegen

$$F(x) = \frac{1}{2} \quad \Leftrightarrow \quad \mathrm{e}^{-\lambda x} = \frac{1}{2} \quad \Leftrightarrow \quad \lambda x = -\log\frac{1}{2} \quad \Leftrightarrow \quad x = \frac{1}{\lambda} \cdot \log 2$$

berechnet sich der Median von X zu $\dfrac{\log 2}{\lambda}$.

b) Wegen $X > 0$ ist $Y = e^{\alpha X} > 1$. Für $t > 1$ gilt

$$\begin{aligned} P(Y \leq t) &= P(\alpha X \leq \log t) = P\left(X \leq \frac{\log t}{\alpha}\right) \\ &= 1 - \exp\left(-\lambda \cdot \frac{\log t}{\alpha}\right) \\ &= 1 - t^{-\lambda/\alpha}. \end{aligned} \qquad \boxed{3\text{ P}}$$

Eine Dichte für Y ist folglich durch

$$f(t) := \frac{dP(Y \leq t)}{dt} \cdot \mathbf{1}\{t > 1\} = \frac{\lambda}{\alpha} \cdot t^{-\lambda/\alpha - 1} \cdot \mathbf{1}\{t > 1\}$$

gegeben.

c) Nach der Transformationsformel für Erwartungswerte gilt für $0 < \alpha < \lambda$

$$EY = \int_0^\infty e^{\alpha x} \lambda e^{-\lambda x} dx \;=\; \int_0^\infty \lambda e^{-(\lambda - \alpha)x} dx \;=\; \frac{\lambda}{\lambda - \alpha}. \qquad \boxed{2\text{ P}}$$

Aufgabe 4 (7 Punkte)
Es sei (U, W) ein zweidimensionaler Zufallsvektor mit Dichte

$$f(u,w) = \frac{u}{w^2} \mathrm{e}^{-u/w} \, \mathbf{1}_{\{u>0, 0<w<1\}}.$$

a) Zeigen Sie, dass $U \sim \mathrm{Exp}(1)$ und $W \sim \mathrm{U}(0,1)$ gilt.

b) Berechnen Sie die Kovarianz $\mathrm{C}(U,W)$ und den Korrelationskoeffizienten $\rho(U,W)$ zwischen U und W.

   Hinweis: Sie können ohne Beweis $E(UW) = 2/3$ und $V(U) = 1$ verwenden.

c) Sind U und W unabhängig?

Lösung zu Aufgabe 4

a) Die Dichte f_W von W ergibt sich durch

$$f_W(w) = \int_0^\infty \frac{u}{w^2} \mathrm{e}^{-\frac{u}{w}} \mathrm{d}u = \int_0^\infty x \mathrm{e}^{-x} \mathrm{d}x = 1, \qquad w \in (0,1),$$

2.5 P

und $f_W(w) = 0$ sonst. Somit ist W auf $(0,1)$ gleichverteilt.

Ferner gilt für $u > 0$

$$f_U(u) = \int_0^1 \frac{u}{w^2} \mathrm{e}^{-\frac{u}{w}} dw = \int_u^\infty \mathrm{e}^{-x} dx = \mathrm{e}^{-u}.$$

Somit ist U exponentialverteilt mit Parameter 1.

b) Für $U \sim \mathrm{Exp}(1)$ bzw. $W \sim \mathrm{U}(0,1)$ gilt

$$EU = 1, \qquad EW = \frac{1}{2}.$$

3.5 P

Ferner ist

$$E(W^2) = \int_0^1 w^2 dw = \frac{1}{3},$$

folglich

$$V(W) = E(W^2) - (EW)^2 = \frac{1}{3} - \frac{1}{4} = \frac{1}{12}.$$

Somit

$$\mathrm{C}(U,W) = E(UW) - EUEW = \frac{2}{3} - \frac{1}{2} = \frac{1}{6}$$

und

$$\rho(U,W) = \frac{\mathrm{C}(U,W)}{\sqrt{V(U)V(W)}} = \frac{\frac{1}{6}}{\frac{1}{2\sqrt{3}}} = \frac{1}{\sqrt{3}}.$$

c) Aus $\mathrm{C}(U,W) \neq 0$ folgt, dass U und W nicht unabhängig sind.

1 P

Aufgabe 5 (7 Punkte)
Die Zufallsvariablen $X_1, \ldots, X_n$ seien unabhängig und identisch verteilt mit

$$P(X_1 = k) = \binom{k+r-1}{k} \left(\frac{r}{r+\mu}\right)^r \left(\frac{\mu}{r+\mu}\right)^k, \qquad k \in \mathbb{N}_0.$$

Dabei ist $r \in \mathbb{N}$ bekannt und $\mu > 0$ unbekannt.

a) Zeigen Sie, dass

$$\hat{\mu}_n = \frac{1}{n} \sum_{i=1}^n X_i$$

der Maximum-Likelihood-Schätzer für μ ist, falls $\sum_{i=1}^n X_i > 0$ ist.

b) Ist $\hat{\mu}_n$ erwartungstreu für μ?

c) Berechnen Sie die Varianz von $\hat{\mu}_n$ und die mittlere quadratische Abweichung

$$MQA_{\hat{\mu}_n}(\mu) = E_\mu(\hat{\mu}_n - \mu)^2.$$

d) Ist die Folge $(\hat{\mu}_n)_{n\in\mathbb{N}}$ konsistent für μ?

Hinweis: Sie können ohne Beweis $E_\mu X_1 = \mu$ und $V_\mu(X_1) = \dfrac{\mu(\mu+r)}{r}$ benutzen.

Lösung zu Aufgabe 5

a) Für $x = (x_1, \ldots, x_n)$, wobei $x_1, \ldots, x_n \in \mathbb{N}_0$, gilt:

$$\begin{aligned} L_x(\mu) &= \prod_{i=1}^n \binom{x_i+r-1}{x_i} \left(\frac{r}{r+\mu}\right)^r \left(\frac{\mu}{r+\mu}\right)^{x_i} \\ &= \left(\frac{r}{r+\mu}\right)^{nr} \left(\frac{\mu}{r+\mu}\right)^{\sum_{i=1}^n x_i} \underbrace{\prod_{i=1}^n \binom{x_i+r-1}{x_i}}_{=:c}. \end{aligned}$$

4 P

Folglich

$$\log L_x(\mu) = nr(\log r - \log(\mu+r)) + \sum_{i=1}^n x_i(\log\mu - \log(r+\mu)) + \log c$$

und

$$\begin{aligned} \frac{d}{d\mu} \log L_x(\mu) &= -\frac{nr}{\mu+r} + \sum_{i=1}^n \overbrace{\left(\frac{1}{\mu} - \frac{1}{\mu+r}\right)}^{=\frac{r}{\mu(\mu+r)}} \cdot x_i \\ &= \frac{nr}{\mu+r}\left(-1 + \bar{x}_n \cdot \frac{1}{\mu}\right) = 0 \quad \Leftrightarrow \quad \mu = \bar{x}_n, \end{aligned}$$

sowie

$$\frac{d^2}{d^2\mu} \log L_x(\bar{x}_n) = \frac{nr}{\bar{x}_n + r} \cdot \left(-\frac{1}{\bar{x}_n}\right) < 0,$$

falls $\bar{x}_n > 0$. Folglich ist

$$\hat{\mu}_n := \bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

der gesuchte Maximum-Likelihood-Schätzer, falls $\sum_{i=1}^{n} X_i > 0$ ist.

b) Wegen $E_\mu \bar{X}_n = \mu$ ist $\hat{\mu}_n$ erwartungstreu.

c) Die mittlere quadratische Abweichung stimmt wegen der Erwartungstreue von $\hat{\mu}_n$ mit $V_\mu(\hat{\mu}_n)$ übereinstimmt und es gilt [1 P]

$$\mathrm{MAQ}_{\hat{\mu}_n}(\mu) = V_\mu(\hat{\mu}_n) = \frac{V_\mu(X_1)}{n} = \frac{\mu(r+\mu)}{nr}.$$ [1 P]

d) Die Folge $(\hat{\mu}_n)_{n\in\mathbb{N}}$ ist konsistent: Für $\varepsilon > 0$ folgt aus der Tschebyscheff-Ungleichung

$$P_\mu(|\hat{\mu}_n - \mu| \geq \varepsilon) \leq \frac{V_\mu(\hat{\mu}_n)}{\varepsilon^2} = \frac{\mu(r+\mu)}{nr\varepsilon^2} \to 0$$ [1 P]

für $n \to \infty$.

Alternativ folgt die Konsistenz direkt aus dem Schwachen Gesetz großer Zahlen.

Aufgabe 6 (7 Punkte)

In einer Studie soll gezeigt werden, dass bei trainierten Personen der mittlere systolische Blutdruck niedriger als der Normalwert von 80mmHG ist. Dazu wird bei n Sportlern der Blutdruck gemessen, wobei wir annehmen, dass die Blutdruckwerte als Realisierungen unabhängiger Zufallsvariablen mit gleicher Normalverteilung $N(\mu, 16)$ modelliert werden können.

a) Formulieren Sie eine geeignete Hypothese und Alternative und geben Sie ein Testverfahren basierend auf der Testgröße $T_n = \sqrt{n}(\bar{X}_n - 80)/4$ an, das geeignet ist, die obige Behauptung „statistisch nachzuweisen“.

   Wie muss dabei der kritische Wert gewählt werden, wenn der Stichprobenumfang $n = 30$ und das Signifikanzniveau $\alpha = 1\%$ betragen soll?

b) Was ist das Testergebnis, wenn bei den 30 Sportlern $\bar{x}_{30}$ =78mmHG gemessen wurde.

c) Mit welcher Wahrscheinlichkeit wird mit der in a) ermittelten Entscheidungsregel eine Fehlentscheidung getroffen, wenn im Fall $n = 30$ der mittlere systolische Blutdruck in Wirklichkeit 78mmHG beträgt?

d) Was könnte der Leiter der Studie tun, um die Fehlerwahrscheinlichkeit in c) zu verringern?

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

t	0	0.41	1.28	1.65	1.96	2.33	2.58
$\Phi(t)$	0.5	0.66	0.9	0.95	0.975	0.99	0.995

Lösung zu Aufgabe 6

a) Man muss als Hypothese $H_0 : \vartheta \geq 80$ und als Alternative $H_1 : \vartheta < 80$ wählen. Nur dann ist bei Ablehnung von H_0 die Aussage “bei trainierten Personen ist der mittlere systolische Blutdruck niedriger als 80mmHG“ 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 $\alpha = 0.01$ verworfen werden, wenn $T_n \leq z_{0.01} = \Phi^{-1}(0.01) = -2.33$ ist. Andernfalls besteht kein Einwand gegen H_0.

b) Wegen $T_{30} = \sqrt{30}(\bar{X}_{30} - 80)/4 = -2.74 < z_{0.01}$ wird H_0 wird auf dem 1%–Niveau abgelehnt.

c) Ist in Wirklichkeit $\mu = 78$, so liegt eine Fehlentscheidung vor, wenn H_0 nicht verworfen wird. 1 P

$$\begin{aligned}
P_\mu(\text{Test verwirft} H_0 \text{ nicht}) &= P_\mu\left(\frac{\sqrt{n}(\bar{X}_n - 80)}{4} > z_{0.01}\right) \\
&= 1 - P_\mu\left(\frac{\sqrt{n}(\bar{X}_n - 80)}{4} \leq z_{0.01}\right) \\
&= 1 - P_\mu\left(\frac{\sqrt{30}(\bar{X}_{30} - 78)}{4} \leq z_{0.01} + \frac{\sqrt{30}(80 - 78)}{4}\right) \\
&= 1 - \Phi(-2.33 + 2.74) \;=\; 1 - \Phi(0.41) \;=\; 0.34.
\end{aligned}$$

2.5 P

d) Der Studienleiter könnte den Stichprobenumfang erhöhen, denn dann nimmt die Güte des Tests zu, es verringert sich also die Wahrscheinlichkeit, dass bei Vorliegen der Alternative die Hypothese H_0 fälschlicherweise nicht abgelehnt wird. $\boxed{1\ \text{P}}$

Eine weitere Möglichkeit wäre, den Fehler 1. Art anzuheben (etwa auf $\alpha = 5\%$).