Karlsruher Institut für Technologie

Institut für Stochastik
Prof. Dr. Daniel Hug

Name:

Vorname:

Matr.-Nr.:

Klausur zum Fach

GRUNDLAGEN DER WAHRSCHEINLICHKEITSTHEORIE UND STATISTIK

für Studierende der INFORMATIK

Datum: **08. Februar 2012**

Dauer: **90 Minuten**

Achtung:
Bei dieser Klausur werden **nur** diejenigen Ergebnisse gewertet, die in die vorgesehenen Kästchen auf dem extra ausgegebenen

Lösungsblatt

eingetragen sind! Die Herleitung wird **nicht** bewertet, es sei denn, dies wird auf dem Lösungsblatt explizit gefordert!

Die Aufgabenblätter werden nicht abgegeben und korrigiert!
Diese Klausur hat bestanden, wer mindestens 18 Punkte von 50 möglichen Punkten erreicht.

Aufgabe 1 (10 Punkte)

Gegeben sei eine Urliste mit den Paaren $(x_1, y_1), \ldots, (x_8, y_8)$

j	1	2	3	4	5	6	7	8
x_j	1.5	1.7	2.7	2.7	3.1	3.2	3.2	3.4
y_j	2.6	2.9	3.4	3.7	3.9	4.2	4.5	4.6

a) Berechnen Sie die Stichprobenmittel $\bar{x}$, $\bar{y}$, die Stichproben-Standardabweichungen s_x, s_y und den empirischen Korrelationskoeffizienten r_{xy}.

Hinweis:

$$\sum_{j=1}^{8} x_j = 21.5,\ \sum_{j=1}^{8} x_j^2 = 61.37,\ \sum_{j=1}^{8} y_j = 29.8,\ \sum_{j=1}^{8} y_j^2 = 114.68,\ \sum_{j=1}^{8} x_j \cdot y_j = 83.57.$$

b) Bestimmen Sie die zugehörige Regressionsgerade $y = a^* + b^* \cdot x$ von y auf x.

c) Berechnen Sie das 0.2-getrimmte Stichprobenmittel $\bar{y}_{0.2}$ von $(y_1, \ldots, y_8)$.

d) Bestimmen Sie das Stichproben-0.3-Quantil $\tilde{y}_{0.3}$ von $(y_1, \ldots, y_8)$.

Aufgabe 2 (10 Punkte)
Es seien X eine Zufallsvariable mit Werten in $\{-1, 1, 2\}$ und

$$\mathbb{P}(X = 1) = \frac{2}{3}, \qquad \mathbb{P}(X = -1) = \mathbb{P}(X = 2) = \frac{1}{6}$$

sowie Y eine Zufallsvariable mit Werten in $\{0, 1\}$ und den bedingten Wahrscheinlichkeiten

$$\mathbb{P}(Y = 0 \mid X = -1) = \frac{1}{2}, \qquad \mathbb{P}(Y = 0 \mid X = 1) = \frac{1}{4}, \qquad \mathbb{P}(Y = 1 \mid X = 2) = \frac{3}{4}.$$

a) **Hinweis:** beachten Sie, dass $\mathbb{P}(Y = 0 \mid X = -1) + \mathbb{P}(Y = 1 \mid X = -1) = 1$ gilt.
Berechnen Sie $\mathbb{P}(Y = 1 \mid X = -1)$, $\mathbb{P}(Y = 1 \mid X = 1)$ und $\mathbb{P}(Y = 0 \mid X = 2)$.

b) Berechnen Sie die gemeinsame Verteilung von X und Y sowie die Verteilung von Y, d.h. $\mathbb{P}(X = x, Y = y)$ und $\mathbb{P}(Y = y)$ für alle $x \in \{-1, 1, 2\}$ und $y \in \{0, 1\}$.

c) Bestimmen Sie die Erwartungswerte $\mathbb{E}(X)$, $\mathbb{E}(Y)$, $\mathbb{E}(X + Y)$.

d) Berechnen Sie $\mathbb{E}(X \cdot Y)$ und die Varianz $V(X)$ von X.

e) Sind X und Y stochastisch unabhängig? Begründen Sie Ihre Antwort.

Aufgabe 3 (10 Punkte)

An einer E-Mail-Adresse kommen erwünschte E-Mails und Spam-Mails an. Jede ankommende E-Mail sei unabhängig von den vorangehenden mit der Wahrscheinlichkeit $p \in (0,1)$ **erwünscht**. Sei S_n die zufällige Anzahl von **nicht erwünschten** Spam-Mails bei n eingegangenen E-Mails.

a) Welche Verteilung hat die Zufallsvariable S_n?

Berechnen Sie $\mathbb{P}(S_n \geq 2)$ für $p = 0.7$ und $n = 10$.

b) Berechnen Sie den Erwartungswert $\mathbb{E}(S_n)$ und die Varianz $V(S_n)$ für $p = 0.7$ und $n = 10$.

c) Schätzen Sie für $p = 0.7 = 1 - q$, $n = 500$ und $\varepsilon = 0.05$ die Wahrscheinlichkeit

$$\mathbb{P}\left(\left|\frac{1}{n} \cdot S_n - q\right| \geq \varepsilon\right)$$

nach oben mit Hilfe der Tschebyschev-Ungleichung ab.

d) Sei nun $p \in (0,1)$ wieder beliebig. Ferner bezeichne T_k die zufällige Anzahl an Spam-Mails bis zur k-ten erwünschten E-Mail, $k = 1, 2$.

Welche Verteilungen haben die Zufallsvariablen T_1 beziehungsweise T_2?

Geben Sie $\mathbb{E}(T_1)$ und $\mathbb{E}(T_2)$ an.

e) Berechnen Sie die Wahrscheinlichkeit $\mathbb{P}(T_1 \geq 2)$ für allgemeines p sowie für $p = 0.7$.

Aufgabe 4 (10 Punkte)

Die Zufallsvariable X besitze die Verteilung $\mathcal{N}(2,1)$. Weiter sei $Y := 1 - 2X$.

a) Berechnen Sie den Erwartungswert $\mathbb{E}(Y)$ und die Varianz $V(Y)$.

b) Welche Verteilung hat Y?

c) Drücken Sie die Wahrscheinlichkeit $\mathbb{P}(-1 \leq Y \leq 3)$ mit Hilfe der Verteilungsfunktion Φ der Standardnormalverteilung $\mathcal{N}(0,1)$ aus.

d) Bestimmen Sie die Kovarianz $C(X,Y)$ und den Korrelationskoeffizienten $\rho(X,Y)$.

e) Sei U eine auf $[0,1]$ gleichverteilte Zufallsvariable, das heißt $U \sim \mathcal{U}(0,1)$. Welche Verteilung hat die Zufallsvariable $2 \cdot \Phi^{-1}(U) + 1$? Begründen Sie Ihre Antwort.

Aufgabe 5 (10 Punkte)

Es soll der unbekannte Parameter $\vartheta > 0$ für die Verteilung mit der Dichte

$$f_\vartheta(t) = \begin{cases} \dfrac{4t^3}{\vartheta} \cdot \exp\left(-\dfrac{1}{\vartheta} \cdot t^4\right), & t > 0, \\ 0, & \text{sonst} \end{cases}$$

bestimmt werden.

a) Geben Sie die zur Stichprobe $x = (x_1, \dots, x_n)$ gehörende Likelihood-Funktion $L_x(\vartheta)$ und die Loglikelihood-Funktion $M_x(\vartheta)$ an.

b) Berechnen Sie die Ableitung $M_x'(\vartheta)$.

c) Bestimmen Sie einen Maximum-Likelihood-Schätzer $\hat{\vartheta}(x)$ für ϑ zur Stichprobe x.

d) Bestimmen Sie zunächst die Verteilungsfunktion von X_1, wenn X_1 die Dichte f_ϑ hat.
Zeigen Sie dann, dass X_1^4 einer Exp$(\frac{1}{\vartheta})$-Verteilung folgt.

e) Ist $\hat{\vartheta}$ ein erwartungstreuer Schätzer für ϑ?