

Klausur
Wahrscheinlichkeitstheorie und Statistik
für Studierende der Informatik
vom 08.02.2010

Musterlösungen

A

Aufgabe A1

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

j	1	2	3	4	5	6	7	8
x_j	0.8	1.8	3	3.9	5	5.8	7	8
y_j	3.6	4.6	5	5.3	6.1	6.1	6.6	6.2

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

Lösung: Direkt aus den Daten ergibt sich gemäß Definition 1.8 und Paragraph 1.5 unter Ausnützung der Beziehung

$$\sum_{j=1}^n (x_j - \bar{x}) \cdot (y_j - \bar{y}) = \sum_{j=1}^n x_j \cdot y_j - n \cdot \bar{x} \cdot \bar{y}$$

$$\bar{x} = 4.41$$

$$s_x = 2.506$$

$$\bar{y} = 5.44$$

$$s_y = 1.007$$

$$r_{xy} = 0.9377$$

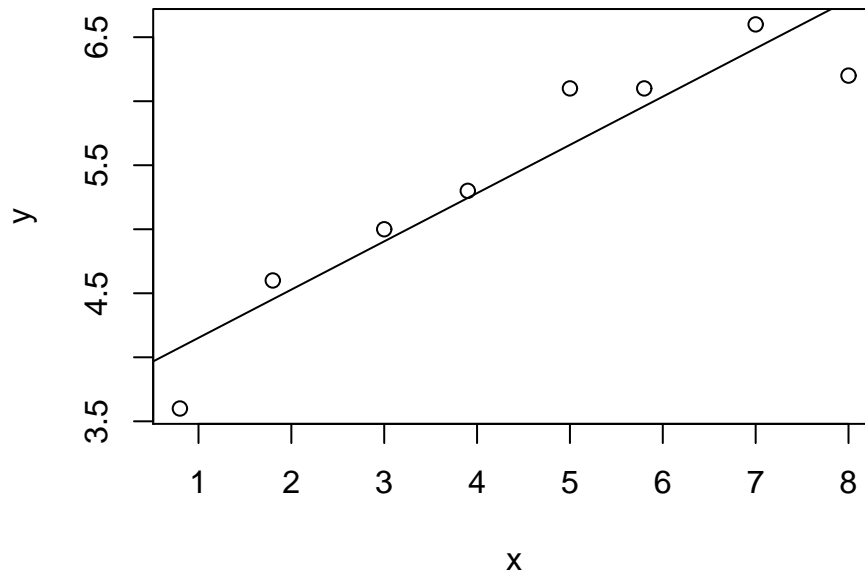
- b) Bestimmen Sie die zugehörige Regressionsgerade $y = a^* + b^* \cdot x$ von y auf x .

Lösung: Nach Paragraph 1.5 ist $b^* = r_{xy} \cdot \frac{s_y}{s_x}$ und $a^* = \bar{y} - b^* \cdot \bar{x}$, also

$$b^* = 0.377$$

$$a^* = 3.77$$

und die Regressionsgerade $y = 3.77 + 0.377 \cdot x$.



Punkte und Regressionsgerade $y = a^* + b^* \cdot x$

Für die Lösung der nächsten beiden Aufgabenteile benötigen wir die aufsteigend sortierten y -Werte. Es ist

$$y_{()} = (3.6, 4.6, 5, 5.3, 6.1, 6.1, 6.2, 6.6)$$

- c) Berechnen Sie das 0.15-getrimmte Stichprobenmittel $\bar{y}_{0.15}$ von (y_1, \dots, y_8) .
Lösung: Mit $k = [8 \cdot 0.15] = 1$ ergibt sich

$$\bar{y}_{0.15} = \frac{1}{8 - 2 \cdot 1} \cdot (y_{(2)} + \dots + y_{(7)}) = 5.55$$

- d) Bestimmen Sie das Stichproben-0.3-Quantil $\tilde{y}_{0.3}$ von (y_1, \dots, y_8) .
Lösung: Da $8 \cdot 0.3 = 2.4$ nicht ganzzahlig ist, ist mit $k = [2.4] = 2$

$$\tilde{y}_{0.3} = y_{(k+1)} = y_{(3)} = 5$$

Aufgabe A2

In einer Urne befinden sich 3 rote und 2 schwarze Kugeln. Es werden zunächst 2 Kugeln rein zufällig und ohne Zurücklegen gezogen.

Hinweis: Evtl. ist die Einführung einer Zufallsvariablen X , die die zufällige Anzahl der gezogenen roten Kugeln angibt, hilfreich.

a) Mit welcher Wahrscheinlichkeit sind die beiden gezogenen Kugeln schwarz?

Lösung:

$$\mathbb{P}(\text{beide Kugeln schwarz}) = \frac{2}{5} \cdot \frac{1}{4} = \frac{1}{10}.$$

Alternativ: Die oben definierte Zufallsvariable X ist $Hyp(2, 3, 2)$ -verteilt. Gesucht ist dann

$$\mathbb{P}(X = 0) = \frac{\binom{3}{0} \cdot \binom{2}{2}}{\binom{5}{2}} = \frac{1}{10}.$$

b) Mit welcher Wahrscheinlichkeit haben die beiden gezogenen Kugeln unterschiedliche Farben?

Lösung:

$$\mathbb{P}(\text{Kugeln haben unterschiedliche Farben}) = \frac{3}{5} \cdot \frac{2}{4} + \frac{2}{5} \cdot \frac{3}{4} = \frac{3}{10} + \frac{3}{10} = \frac{6}{10}.$$

c) Sind die beiden gezogenen Kugeln gleichfarbig, werden zwei rote Kugeln in die Urne gelegt. Haben die beiden gezogenen Kugeln unterschiedliche Farben, werden zwei schwarze Kugeln in die Urne gelegt. Die Zufallsvariable Y bezeichne die Anzahl der schwarzen Kugeln, die danach in der Urne liegen. Bestimmen Sie die Wahrscheinlichkeiten $\mathbb{P}(Y = k)$ für $k = 0, 1, 2, 3$.

Lösung: Die Zufallsvariable Y nimmt die Werte 0, 2, 3 an: Es gilt $Y = 0$ bzw. $Y = 2$, falls die beiden gezogenen Kugeln schwarz bzw. rot sind, und $Y = 3$, falls die beiden gezogenen Kugeln unterschiedliche Farben haben. Somit gilt

$$\mathbb{P}(Y = 0) = \frac{1}{10}, \quad \mathbb{P}(Y = 1) = 0, \quad \mathbb{P}(Y = 3) = \frac{6}{10}, \quad \mathbb{P}(Y = 2) = 1 - \frac{1}{10} - \frac{6}{10} = \frac{3}{10}.$$

Alternativ: Es gilt

$$\begin{aligned} \mathbb{P}(Y = 0) &= \mathbb{P}(X = 0) = \frac{1}{10}, \\ \mathbb{P}(Y = 3) &= \mathbb{P}(X = 1) = \frac{6}{10}, \\ \mathbb{P}(Y = 2) &= \mathbb{P}(X = 2) = 1 - \frac{7}{10} = \frac{3}{10}. \end{aligned}$$

d) Anschließend wird nach gutem Mischen rein zufällig eine Kugel gezogen. Wie groß ist die Wahrscheinlichkeit, dass diese Kugel schwarz ist?

Lösung: Mit dem Satz über die totale Wahrscheinlichkeit erhält man

$$\begin{aligned} \mathbb{P}(\text{es wird eine schwarze Kugel gezogen}) &= \mathbb{P}(\text{Kugel schwarz} \mid Y = 0) \cdot \mathbb{P}(Y = 0) \\ &\quad + \mathbb{P}(\text{Kugel schwarz} \mid Y = 2) \cdot \mathbb{P}(Y = 2) \\ &\quad + \mathbb{P}(\text{Kugel schwarz} \mid Y = 3) \cdot \mathbb{P}(Y = 3) \\ &= 0 \cdot \frac{1}{10} + \frac{2}{5} \cdot \frac{3}{10} + \frac{3}{5} \cdot \frac{6}{10} \\ &= \frac{0 + 6 + 18}{50} = \frac{12}{25}. \end{aligned}$$

- e) Wie groß ist die bedingte Wahrscheinlichkeit p , dass die beiden zuerst gezogenen Kugeln rot sind, unter der Bedingung, dass die zuletzt gezogene Kugel schwarz ist?

Lösung:

$$\begin{aligned} p &= \mathbb{P}(\text{die beiden zuerst gezogenen Kugeln rot} \mid \text{die zuletzt gezogene Kugel ist schwarz}) \\ &= \mathbb{P}(\text{die ersten zwei Kugeln rot, die dritte schwarz}) / \mathbb{P}(\text{die zuletzt gezogene Kugel ist schwarz}) \\ &= \frac{\frac{3}{5} \cdot \frac{2}{4} \cdot \frac{2}{5}}{\frac{12}{25}} = \frac{3}{12} = \frac{1}{4}. \end{aligned}$$

Aufgabe A3

Auf einer Metallhobelmaschine werden ringförmige Platten mit zufälligem Außendurchmesser X und Innendurchmesser Y hergestellt. X mit Normalverteilung $\mathcal{N}(80, 20)$ und Y mit Normalverteilung $\mathcal{N}(40, 16)$ seien stochastisch unabhängig. Falls die Ringbreite $U = \frac{X-Y}{2}$ in das Intervall $[14, 26]$ fällt, so wird die Platte hergestellt, ansonsten wird der Vorgang abgebrochen.

- a) Wie groß ist die Wahrscheinlichkeit, dass X zwischen 78 und 83 liegt?

Lösung:

$$\begin{aligned}\mathbb{P}(78 \leq X \leq 83) &= \Phi\left(\frac{83-80}{\sqrt{20}}\right) - \Phi\left(\frac{78-80}{\sqrt{20}}\right) \\ &= \Phi(0.67) - \Phi(-0.45) = \Phi(0.67) + \Phi(0.45) - 1 \\ &= 0.7486 + 0.6736 - 1 = 0.4222.\end{aligned}$$

- b) Geben Sie den Erwartungswert und die Varianz von U an. Welche Verteilung hat U ? Wie groß ist die Wahrscheinlichkeit, dass eine Platte hergestellt wird?

Lösung: Es gilt $\mathbb{E}(U) = (80 - 40)/2 = 20$; wegen der Unabhängigkeit von X und Y gilt weiter $V(U) = (20 + 16)/4 = 9$. Nach der Faltungsformel ist $U \mathcal{N}(20, 9)$ -verteilt, es ist also $\sigma := \sqrt{V(U)} = \sqrt{9} = 3$. Somit folgt

$$\begin{aligned}\mathbb{P}(14 \leq U \leq 26) &= \mathbb{P}(|U - 20| \leq 6) \\ &= \mathbb{P}(|U - 20| \leq 2\sigma) \\ &= 2 \cdot \Phi(2) - 1 = 2 \cdot 0.9772 - 1 = 0.9544.\end{aligned}$$

- c) Sei $W = \frac{X+Y}{2}$. Welche Verteilung besitzt W ?

Lösung: Es gilt $W \sim \mathcal{N}(60, 9)$.

- d) Berechnen Sie Kovarianz und Korrelationskoeffizient von U und W .

Lösung:

$$\begin{aligned}C(U, W) &= C\left(\frac{X-Y}{2}, \frac{X+Y}{2}\right) \\ &= \frac{1}{4} \cdot [C(X, X) + C(X, Y) - C(Y, X) - C(Y, Y)] \\ &= \frac{1}{4} \cdot (V(X) - V(Y)) = \frac{1}{4} \cdot (20 - 16) = 1.\end{aligned}$$

Weiter gilt

$$\rho(U, W) = \frac{C(U, W)}{\sqrt{V(U) \cdot V(W)}} = \frac{1}{9}.$$

- e) Die zufällige Fläche der ringförmigen Platte ist $F = \pi \cdot U \cdot W = \frac{\pi}{4}(X^2 - Y^2)$. Bestimmen Sie den Erwartungswert von F .

Lösung:

$$\begin{aligned}\mathbb{E}F &= \mathbb{E}(\pi \cdot U \cdot W) \\ &= \pi \cdot [C(U, W) + \mathbb{E}U \cdot \mathbb{E}W] \\ &= \pi \cdot (1 + 20 \cdot 60) = 1201\pi.\end{aligned}$$

Alternativ:

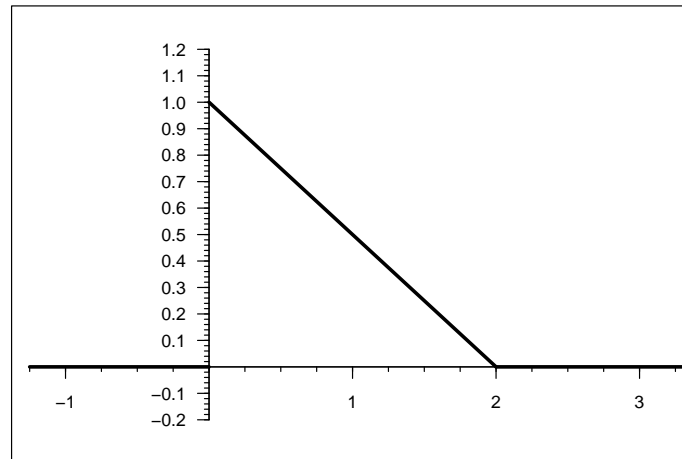
$$\begin{aligned}\mathbb{E}F &= \frac{\pi}{4} \cdot \mathbb{E}(X^2 - Y^2) \\ &= \frac{\pi}{4} \cdot \mathbb{E}(V(X) + \mathbb{E}(X^2) - V(Y) - \mathbb{E}(Y^2)) \\ &= \frac{\pi}{4} \cdot (20 + 6400 - 16 - 1600) = 1201\pi.\end{aligned}$$

Aufgabe A4

Die Dichte f einer stetigen Zufallsvariablen X ist gegeben durch

$$f(x) = \begin{cases} (2-x)/2, & \text{falls } 0 \leq x \leq 2, \\ 0, & \text{sonst.} \end{cases}$$

Das nachstehende Schaubild zeigt f :



- a) Bestimmen Sie den Erwartungswert von X .

Lösung: Es gilt

$$\mathbb{E}X = \int_{-\infty}^{\infty} x \cdot f(x) dx = \int_0^2 (x - x^2/2) dx = \left[\frac{x^2}{2} - \frac{x^3}{6} \right]_0^2 = \frac{2}{3}.$$

- b) Bestimmen Sie das zweite Moment $\mathbb{E}(X^2)$ und die Varianz $V(X)$ von X .

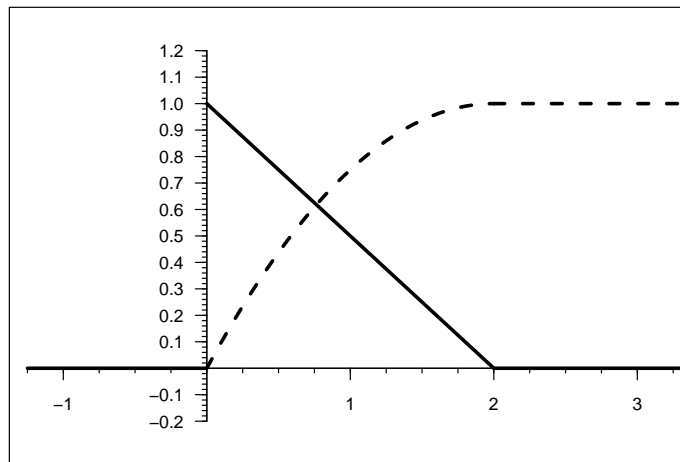
Lösung: Es ist

$$\mathbb{E}(X^2) = \int_{-\infty}^{\infty} x^2 \cdot f(x) dx = \int_0^2 (x^2 - x^3/2) dx = \left[\frac{x^3}{3} - \frac{x^4}{8} \right]_0^2 = \frac{2}{3},$$
$$V(X) = \mathbb{E}(X^2) - (\mathbb{E}X)^2 = \frac{2}{9}.$$

- c) Bestimmen Sie die Verteilungsfunktion F zur Dichte f , und zeichnen Sie den Graphen von F in das Schaubild ein.

Lösung: Es gilt

$$F(x) = \int_{-\infty}^x f(t) dt = \begin{cases} 0, & \text{falls } x < 0 \\ x - \frac{x^2}{4}, & \text{falls } 0 \leq x \leq 2, \\ 1, & \text{falls } x > 2. \end{cases}$$



d) Bestimmen Sie den Median von F .

Lösung: Es gilt

$$x - \frac{x^2}{4} \stackrel{!}{=} \frac{1}{2} \Leftrightarrow x^2 - 4 \cdot x + 2 = 0 \Leftrightarrow x_{1/2} = 2 \pm \sqrt{4 - 2} = 2 \pm \sqrt{2}.$$

Folglich muss das 0.5-Quantil von F gleich $2 - \sqrt{2} = 0.5858$ sein.

Aufgabe A5

Die zufällige Lebensdauer X eines Produktes besitze die Dichte

$$f_{\vartheta}(t) = \begin{cases} 2 \cdot \vartheta \cdot t \cdot \exp(-\vartheta \cdot t^2), & t > 0, \\ 0, & \text{sonst,} \end{cases}$$

und die Verteilungsfunktion

$$F_{\vartheta}(t) = \begin{cases} 1 - \exp(-\vartheta \cdot t^2), & t > 0, \\ 0, & \text{sonst,} \end{cases}$$

mit einem unbekanntem Parameter $\vartheta > 0$.

Hinweis: Es gilt $\mathbb{E}X = \frac{1}{2}\sqrt{\frac{\pi}{\vartheta}}$ und $V(X) = \frac{4-\pi}{4\vartheta}$.

- a) Bestimmen Sie die Likelihood-Funktion $L_x(\vartheta)$ und die Log-Likelihood-Funktion $M_x(\vartheta)$ zur Stichprobe $x = (x_1, \dots, x_n)$, wobei $x_i > 0$ für $i = 1, \dots, n$ gelte.

Lösung: Die Likelihood-Funktion ist

$$L_x(\vartheta) = \prod_{i=1}^n f_{\vartheta}(x_i) = (2\vartheta)^n \exp\left(-\vartheta \sum_{i=1}^n x_i^2\right) \cdot \prod_{i=1}^n x_i.$$

Weiter gilt

$$\log(f_{\vartheta}(t)) = \log \vartheta + \log(2t) - \vartheta t^2, \quad t > 0$$

somit ergibt sich die Loglikelihood-Funktion zu

$$\begin{aligned} M_x(\vartheta) &= \sum_{i=1}^n \log(f_{\vartheta}(x_i)) = \sum_{i=1}^n (\log \vartheta + \log(2x_i) - \vartheta x_i^2) \\ &= n \log \vartheta + \sum_{i=1}^n \log(2x_i) - \vartheta \sum_{i=1}^n x_i^2. \end{aligned}$$

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

Lösung: Die Ableitung der Loglikelihood-Funktion nach ϑ ist

$$M'_x(\vartheta) = \frac{n}{\vartheta} - \sum_{i=1}^n x_i^2.$$

Damit gilt $M'_x(\vartheta) = 0$ für $\vartheta = n / \sum_{i=1}^n x_i^2$. Da für die zweite Ableitung

$$M''_x(\vartheta) = -\frac{n}{\vartheta^2} < 0$$

gilt, erhält man den Maximum-Likelihood-Schätzer

$$\hat{\vartheta}(x) = \frac{n}{\sum_{i=1}^n x_i^2}.$$

- c) Bestimmen Sie den Momentenschätzer $\tilde{\vartheta}(x)$ für ϑ zur Stichprobe $x = (x_1, \dots, x_n)$.
Lösung: Aus der Gleichung

$$\frac{1}{2} \sqrt{\frac{\pi}{\vartheta}} = \mathbb{E}X = \bar{x}$$

ergibt sich

$$\tilde{\vartheta}(x) = \frac{\pi}{4\bar{x}^2}.$$

- d) X_1, \dots, X_n seien unabhängige Zufallsvariablen mit Dichte f_ϑ . Bestimmen Sie Konstanten a und k so, dass $\frac{1}{n} \sum_{i=1}^n X_i^2$ ein erwartungstreuer Schätzer für $\gamma(\vartheta) = \frac{a}{\vartheta^k}$ ist.

Lösung:

$$\begin{aligned} \mathbb{E} \left(\frac{1}{n} \sum_{i=1}^n X_i^2 \right) &= \frac{1}{n} \sum_{i=1}^n \mathbb{E}X_i^2 = \mathbb{E}X_1^2 \\ &= V(X_1) + (\mathbb{E}X_1)^2 = \frac{4 - \pi}{4\vartheta} + \frac{\pi}{4\vartheta} = \frac{1}{\vartheta}. \end{aligned}$$

Der Schätzer ist also erwartungstreu für $\gamma(\vartheta) = 1/\vartheta$, es muss also $a = 1$ und $k = 1$ gewählt werden.

- e) Bestimmen Sie die Verteilung von $X_1^2/2$.

Lösung: Für $x > 0$ ist

$$\mathbb{P}(X_1^2/2 \leq x) = \mathbb{P}(X_1 \leq \sqrt{2x}) = 1 - \exp(-\vartheta \cdot 2x),$$

d.h. $X_1^2/2$ ist exponentialverteilt mit Parameter 2ϑ .