INSTITUT FÜR MATHEMATISCHE STOCHASTIK SS 2006
UNIVERSITÄT KARLSRUHE
Priv.-Doz. Dr. D. Kadelka

Klausur

Wahrscheinlichkeitstheorie und Statistik
vom 18.9.2006

Musterlösungen

Aufgabe A1:

Gegeben sei eine Urliste mit den Paaren $(x_1, y_1), \ldots, (x_{10}, y_{10})$

j	1	2	3	4	5	6	7	8	9	10
x_j	0.1	0.8	4.3	4.7	4.7	5.3	6.2	8.1	9.8	14
y_j	5.1	7.7	5.7	2.9	3.4	−1.5	−1.7	−4.2	−2.1	−8.8

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

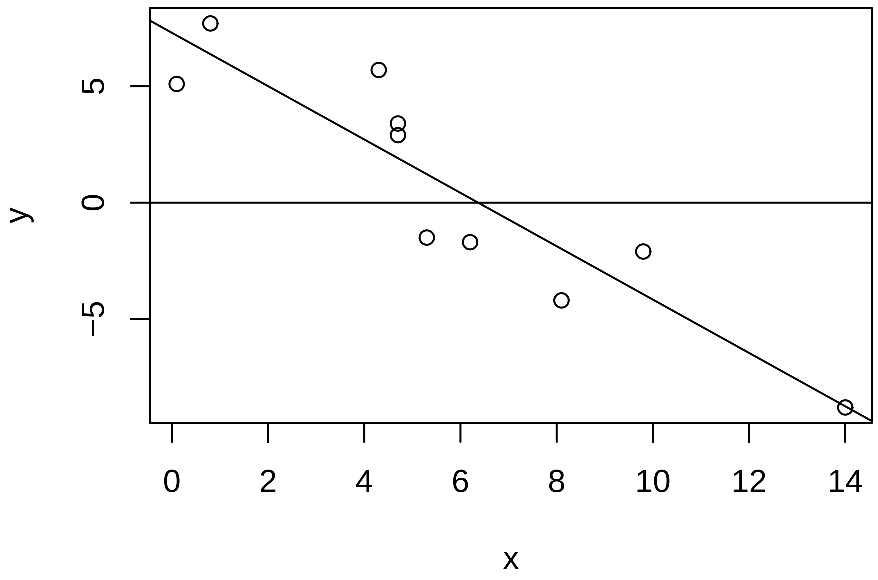
$$\sum_{j=1}^{n}(x_j - \bar{x}) \cdot (y_j - \bar{y}) = \sum_{j=1}^{n} x_j \cdot y_j - n \cdot \bar{x} \cdot \bar{y}$$

$$\begin{aligned} \bar{x} &= 5.8 & s_x &= 4.097 \\ \bar{y} &= 0.65 & s_y &= 5.144 \\ r_{xy} &= -0.913 \end{aligned}$$

b) Bestimmen Sie die zugehörige Regressionsgerade $y = a^* + b^* \cdot x$ von y auf x.
Lösung: Nach Paragraph 1.5 ist $b^* = r_{xy} \cdot \dfrac{s_y}{s_x}$ und $a^* = \bar{y} - b^* \cdot \bar{x}$, also

$$\begin{aligned} b^* &= -1.146 \\ a^* &= 7.3 \end{aligned}$$

und die Regressionsgerade $y = 7.3 - 1.146 \cdot x$.



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

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

$$y_{()} = (-8.8,\ -4.2,\ -2.1,\ -1.7,\ -1.5,\ 2.9,\ 3.4,\ 5.1,\ 5.7,\ 7.7)$$

c) Berechnen Sie das 0.1-getrimmte Stichprobenmittel $\bar{y}_{0.1}$ von $(y_1, \ldots, y_{10})$.
Lösung: Mit $k = [10 \cdot 0.1] = 1$ ergibt sich

$$\bar{y}_{0.1} = \frac{1}{10 - 2 \cdot 1} \cdot (y_{(2)} + \ldots + y_{(9)}) = 0.95$$

d) Berechnen Sie den Quartilsabstand von $(y_1, \ldots, y_{10})$.
Lösung: Da $0.25 \cdot 10 = 2.5$ und $0.75 \cdot 10 = 7.5$ beide nicht ganzzahlig sind, ergibt sich mit $k_1 = [2.5] = 2$ und $k_2 = [7.5] = 7$

$$\begin{aligned} \tilde{y}_{0.25} &= y_{(k_1+1)} = y_{(3)} = -2.1 \\ \tilde{y}_{0.75} &= y_{(k_2+1)} = y_{(8)} = 5.1 \end{aligned}$$

und damit der Quartilsabstand zu $\tilde{y}_{0.75} - \tilde{y}_{0.25} = 7.2$.

Aufgabe A2

Die gemeinsame Zähldichte $f_{X,Y}(i,j)$ von zwei diskreten Zufallsvariablen X und Y ist in untenstehender Tabelle angegeben.

j \ i	-1	0	1
-2	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{16}$
-1	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$
1	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{8}$
2	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$

a) Berechnen Sie die Zähldichten f_X von X und f_Y von Y.

b) Berechnen Sie die bedingte Wahrscheinlichkeit $\mathbb{P}(Y > 0 \mid X = 1)$.

c) (Bei diesem Aufgabenteil wird bei einer falschen Antwort ein Punkt abgezogen!) Sind X und Y stochastisch unabhängig?

d) Berechnen Sie die Erwartungswerte $\mathbb{E}X$ und $\mathbb{E}Y$, die zweiten Momente $\mathbb{E}X^2$ und $\mathbb{E}Y^2$ und die Varianzen $V(X)$ und $V(Y)$.

Lösung:

a) Die Zähldichten f_X bzw. f_Y erhält man aus der Tabelle durch Bilden der Spaltensummen bzw. Zeilensummen:

j \ i	-1	0	1	$f_Y(j)$
-2	$2/16$	$2/16$	$1/16$	$5/16$
-1	$1/16$	$1/16$	$1/16$	$3/16$
1	$2/16$	$1/16$	$2/16$	$5/16$
2	$1/16$	$1/16$	$1/16$	$3/16$
$f_X(i)$	$6/16$	$5/16$	$5/16$	

b) Nach Definition gilt allgemein $\mathbb{P}(A \mid B) = \dfrac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$ für zwei Ereignisse A und B. Hier ist $A = \{Y > 0\} = \{Y = 1\} + \{Y = 2\}$, $B = \{X = 1\}$ und $A \cap B = \{X = 1, Y = 1\} + \{X = 1, Y = 2\}$. Damit gilt

- $\mathbb{P}(A \cap B) = \mathbb{P}(X = 1, Y = 1) + \mathbb{P}(X = 1, Y = 2) = f_{X,Y}(1,1) + f_{X,Y}(1,2) = \frac{3}{16}$,
- $\mathbb{P}(B) = \mathbb{P}(X = 1) = f_X(1) = \frac{5}{16}$, insgesamt

$$\mathbb{P}(Y > 0 \mid X = 1) = \mathbb{P}(A \mid B) = \frac{\frac{3}{16}}{\frac{5}{16}} = \frac{3}{5}.$$

c) Wären X und Y stochastisch unabhängig, so müsste

$$\mathbb{P}(Y > 0 \mid X > 0) = \frac{3}{5} = \mathbb{P}(Y > 0) = \frac{3+5}{16} = \frac{1}{2}$$

gelten. Da dies falsch ist, können X und Y nicht unabhängig sein.

d) Wegen b) gilt

$$\begin{aligned}
\mathbb{E}X &= \sum_{i=-1}^{1} i \cdot f_X(i) = -1 \cdot \frac{6}{16} + 0 \cdot \frac{5}{16} + 1 \cdot \frac{5}{16} = -\frac{1}{16} \\
\mathbb{E}X^2 &= \sum_{i=-1}^{1} i^2 \cdot f_X(i) = 1 \cdot \frac{6}{16} + 0 \cdot \frac{5}{16} + 1 \cdot \frac{5}{16} = \frac{11}{16} \\
V(X) &= \mathbb{E}X^2 - (\mathbb{E}X)^2 = \frac{11}{16} - \left(\frac{1}{16}\right)^2 = \frac{175}{256} = 0.6836 \\
\mathbb{E}Y &= \sum_{j=-2}^{2} j \cdot f_Y(j) = -2 \cdot \frac{5}{16} - 1 \cdot \frac{3}{16} + 1 \cdot \frac{5}{16} + 2 \cdot \frac{3}{16} = -\frac{1}{8} \\
\mathbb{E}Y^2 &= \sum_{j=-2}^{2} j^2 \cdot f_Y(j) = 4 \cdot \frac{5}{16} + 1 \cdot \frac{3}{16} + 1 \cdot \frac{5}{16} + 4 \cdot \frac{3}{16} = \frac{5}{2} \\
V(Y) &= \mathbb{E}Y^2 - (\mathbb{E}Y)^2 = \frac{159}{64} = 2.4844
\end{aligned}$$

Aufgabe A3

Jeden Morgen treffen in einer Transportfirma die Aufträge für den Tag ein. Die Firma beobachtet, dass die zufällige Anzahl N der Aufträge für einen Tag eine Poisson-Verteilung $Po(1.9)$ besitzt.

a) Bestimmen Sie $\mathbb{E}N$.

b) Berechnen Sie $P(N \leq k)$ für $k = 0, 1, 2$.

c) Für jeden Auftrag braucht ein Fahrzeug genau einen Tag. Es werden nur soviel Anträge angenommen, wie pro Tag ausgeführt werden können. Wieviele Fahrzeuge b muss die Firma mindestens zur Verfügung haben, damit mit Wahrscheinlichkeit ≥ 0.6 kein Auftrag abgelehnt werden muss?

Hinweis: Verwenden Sie Teil b).

d) Seien $N_1, N_2, \ldots, N_5$ die zufälligen Anzahlen von Aufträgen an den 5 Arbeitstagen einer Woche. Sei G die zufällige Anzahl von Aufträgen, die innerhalb einer Woche an diesen 5 Arbeitstagen in der Firma eintreffen. Bestimmen Sie eine geeignete Formel für G unter Verwendung von $N_1, \ldots, N_5$.

e) Welche Verteilung besitzt G, wenn $N_1, \ldots, N_5$ stochastisch unabhängig sind? Bestimmen Sie den Erwartungswert und die Varianz von G.

Lösung:

a) Ist N eine Zufallsvariable mit Verteilung $Po(\alpha)$, so gilt $\mathbb{E}N = \alpha$. Nach Voraussetzung gilt daher $\mathbb{E}N = 1.9$.

b) Allgemein gilt $\mathbb{P}(N \leq k) = \mathbb{P}(N = 0)+\mathbb{P}(N = 1)+\ldots+\mathbb{P}(N = k)$. Wegen $N \sim Po(1.9)$ gilt

$$\mathbb{P}(N = k) = e^{-1.9} \cdot \frac{1.9^k}{k!}$$

und damit

k	$\mathbb{P}(N = k)$	$\mathbb{P}(N \leq k)$
0	0.1496	0.1496
1	0.2842	0.4338
2	0.2700	0.7038

c) Sei b die Anzahl der zur Verfügung stehenden Fahrzeuge. Gilt $N \leq b$, so stehen genügend Fahrzeuge zur Verfügung und kein Auftrag braucht abgelehnt zu werden. Ist dagegen $N > b$, so muss mindestens ein Auftrag abgelehnt werden. Daher ist b so zu bestimmen, dass

$$\mathbb{P}(N \leq b) \geq 0.6$$

erfüllt ist.

Aus Teil b) folgt, dass für $b = 1$ $\mathbb{P}(N \leq b) = 0.4338 < 0.6$ ist. 1 Fahrzeug reicht daher nicht aus. Für $b = 2$ gilt dagegen $\mathbb{P}(N \leq b) = 0.7038 > 0.6$. 2 Fahrzeuge, natürlich auch mehr, reichen daher aus.

Es müssen daher mindestens 2 Fahrzeuge vorhanden sein.

d) Eine geeignete Formel für G ist

$$G = N_1 + N_2 + N_3 + N_4 + N_5.$$

e) Da jede der stochastisch unabhängigen Zufallsvariablen $N_1, \ldots, N_5$ die Verteilung $Po(1.9)$ besitzt, gilt nach der Faltungsformel

$$G \sim Po(1.9 + 1.9 + 1.9 + 1.9 + 1.9) = Po(9.5).$$

Damit gilt dann auch

- $\mathbb{E}G = 9.5$ und
- $V(G) = 9.5$.

Aufgabe A4

Es sei X eine normalverteilte Zufallsvariable mit Erwartungswert 0 und der Varianz 16, die Zufallsvariable Y sei normalverteilt mit Erwartungswert 4 und der Varianz 1. X und Y seien stochastisch unabhängig.

a) Bestimmen Sie die Verteilung der Zufallsvariablen $Z := \dfrac{X}{2} + 1$ und die Verteilung der Zufallsvariablen $W := 5 - Y$.

b) Sind die Zufallsvariablen $U := X^3 \cdot \sin(X)$ und $V := 2\pi \cdot \frac{Y}{e^{Y^2}}$ stochastisch unabhängig?

c) Welche Verteilung besitzt die Zufallsvariable $\dfrac{X}{2} + 1 - Y$?

d) Bestimmen Sie $\mathbb{P}(Y \leq 5)$.

e) Bestimmen Sie das 0.975-Quantil t der Zufallsvariablen Z.

Lösung:

a) Nach Satz 9.7 besitzt $Z = \frac{1}{2} \cdot X + 1$ die Verteilung $\mathcal{N}(1, \frac{1}{2^2} \cdot 16) = \mathcal{N}(1, 4)$ und $W = 5 - Y = (-Y) + 5$ die Verteilung $\mathcal{N}(5 - 4, (-1)^2 \cdot 1) = \mathcal{N}(1, 1)$.

b) Als Funktionen der unabhängigen Zufallsvariablen X und Y sind nach dem Blockungslemma 11.9 auch U und V stochastisch unabhängig.

c) Da auch die Zufallsvariablen $\frac{X}{2} + 1$ und $-Y$ stochastisch unabhängig sind und da $\frac{X}{2} + 1$ die Verteilung $\mathcal{N}(1, 4)$ und $-Y$ die Verteilung $\mathcal{N}(-4, (-1)^2 \cdot 1) = \mathcal{N}(-4, 1)$ besitzt, gilt nach 11.16

$$\frac{X}{2} + 1 - Y \sim \mathcal{N}(1 - 4, 4 + 1) = \mathcal{N}(-3, 5).$$

d) Wegen Satz 9.6 gilt $\mathbb{P}(Y \leq 5) = \Phi_{4,1}(5) = \Phi\left(\dfrac{5-4}{1}\right) = \Phi(1) = 0.8413.$

e) Gesucht ist die Lösung $t := t_{0.975}(Z)$ von $\mathbb{P}(Z \leq t) = 0.975$, also

$$0.975 = \Phi_{1,4}(t) = \Phi\left(\frac{t-1}{2}\right).$$

Nach 12.20 d) ist $\Phi(1.9600) = 0.975$, also $\frac{t-1}{2} = 1.96$ und damit

$$t = t_{0.975}(Z) = 2 \cdot 1.96 + 1 = 4.92.$$

Aufgabe A5

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

$$f_\vartheta(x) = \begin{cases} 5\vartheta^5 \cdot x^4 \cdot e^{-\vartheta^5 x^5} & , x > 0 \\ 0 & , x \leq 0 \end{cases}$$

bestimmt werden.

a) Berechnen Sie zur Stichprobe $x = (x_1, \ldots, x_n)$ mit $x_1 > 0, \ldots, x_n > 0$ die Loglikelihood-Funktion $M_x(\vartheta)$.

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

c) Bestimmen Sie den Momentenschätzer $\hat{\vartheta}_1(x)$ zur Stichprobe x.
Hinweis: Sie dürfen ohne Nachweis verwenden, dass $\mathbb{E}X = 0.9182/\vartheta$ für jede Zufallsvariable X mit Dichte f_ϑ gilt.

d) Bestimmen Sie den Maximum-Likelihood-Schätzwert $\hat{\vartheta}(x)$ und den Momenten-Schätzwert $\hat{\vartheta}_1(x)$ für die Stichprobe

$$x = (0.9\ , 1.2\ , 0.5\ , 0.9\ , 0.9\ , 0.8\ , 1.4\ , 1.4\ , 0.8\ , 0.8).$$

Hinweis: $\sum_{i=1}^{10} x_i = 9.6$, $\sum_{i=1}^{10} x_i^2 = 9.96$ und $\sum_{i=1}^{10} x_i^5 = 16.031$.

Lösung:

a) Wegen Definition 17.6 gilt

$$\begin{aligned} L_x(\vartheta) &= f_\vartheta(x_1) \cdot \ldots \cdot f_\vartheta(x_n) \\ &= 5\vartheta^5 \cdot x_1^4 \cdot e^{-\vartheta^5 x_1^5} \cdot \ldots \cdot 5\vartheta^5 \cdot x_n^4 \cdot e^{-\vartheta^5 x_n^5} = 5^n \cdot (\vartheta^5)^n \cdot \prod_{i=1}^n x_i^4 \cdot e^{-\vartheta^5 \cdot \sum_{i=1}^n x_i^5} \\ &= 5^n \cdot \prod_{i=1}^n x_i^4 \cdot \vartheta^{5n} \cdot e^{-\vartheta^5 \cdot \sum_{i=1}^n x_i^5}. \end{aligned}$$

Nach Logarithmieren ergibt sich

$$M_x(\vartheta) = \ln\left(L_x(\vartheta)\right) = n \cdot \ln(5) + \sum_{i=1}^n 4\ln(x_i) + 5n \cdot \ln(\vartheta) - \vartheta^5 \cdot \sum_{i=1}^n x_i^5$$

Alternativ kann man hier auch $M_x(\vartheta) = \sum_{i=1}^n \ln(f_\vartheta(x_i))$ ausnützen. Mit $\ln(f_\vartheta(x) = \ln(5) + 5\ln(\vartheta) + 4\ln(x) - \vartheta^5 x^5$ ergibt sich für $M_x(\vartheta)$ der gleiche Wert.

b) Eine (globale) Maximumstelle von M_x erhalten wir durch Differenzieren.

$$M_x'(\vartheta) = \frac{5n}{\vartheta} - 5 \cdot \vartheta^4 \cdot \sum_{i=1}^n x_i^5 = 5n \cdot \vartheta^4 \cdot \left(\frac{1}{\vartheta^5} - \frac{1}{n} \sum_{i=1}^n x_i^5 \right)$$

Es ist $M_x'(\vartheta) \left\{ \begin{matrix} > \\ = \\ < \end{matrix} \right\} 0$ genau dann, wenn $\frac{1}{\vartheta^5} \left\{ \begin{matrix} > \\ = \\ < \end{matrix} \right\} \frac{1}{n} \sum_{i=1}^n x_i^5$. Dies gilt genau dann, wenn $\vartheta \left\{ \begin{matrix} < \\ = \\ > \end{matrix} \right\} \left(\frac{n}{\sum_{i=1}^n x_i^5} \right)^{1/5} =: \hat{\vartheta}(x)$. Dies bedeutet, dass die Funktion $\vartheta \to M_x(\vartheta)$ unterhalb von $\hat{\vartheta}(x)$ zuerst steigt und ab $\hat{\vartheta}(x)$ fällt. An der Stelle $\hat{\vartheta}(x)$ liegt also eine (globale) Maximumstelle von M_x und damit ist $\hat{\vartheta}(x)$ der gesuchte Maximum-Likelihood-Schätzer.

c) Den Momentenschätzer $\hat{\vartheta}_1(x)$ erhalten wir durch Gleichsetzen von $\bar{x}$ und $\mathbb{E}X$. Aus der Gleichung

$$0.9182/\hat{\vartheta}_1(x) = \mathbb{E}X = \bar{x}$$

ergibt sich

$$\hat{\vartheta}_1(x) = \frac{0.9182}{\bar{x}}.$$

d) Es ist $\bar{x} = 0.96$, $\sum_{i=1}^{10} x_i^5 = 16.031$ und damit

- $\hat{\vartheta}(x) = \left(\frac{10}{16.031}\right)^{1/5} = 0.9099$ und
- $\hat{\vartheta}_1(x) = \frac{0.9182}{\bar{x}} = \frac{0.9182}{0.96} = 0.9565.$

Bem.: Die Werte entstammen einer Stichprobe mit $\vartheta = 0.82$.