INSTITUT FÜR MATHEMATISCHE STOCHASTIK SS 2005
UNIVERSITÄT KARLSRUHE
Dr. Martin Folkers
Priv.-Doz. Dr. D. Kadelka

Klausur

Wahrscheinlichkeitstheorie und Statistik
vom 26.7.2005

Musterlösungen

Aufgabe 1

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	-2.2	-0.9	0	0.8	2.1	3.5	4.3	5	5.7	7
y_j	8.9	8.2	4.9	3.3	2	0.7	1	-2.6	-1.5	-1.8

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

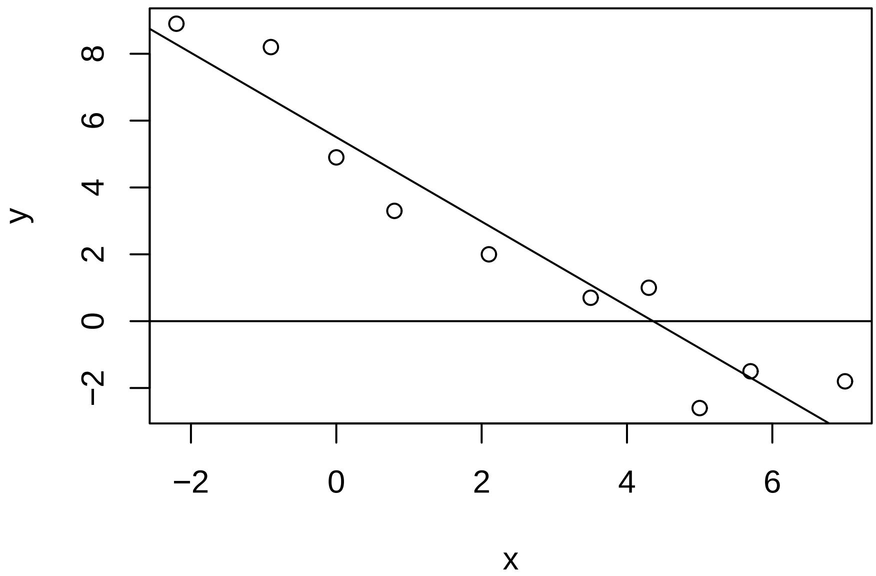
$$\sum_{j=1}^{n}(x_j - \bar{x}) \cdot (y_j - \bar{y}) = \sum_{j=1}^{n} x_j \cdot y_j - n \cdot \bar{x} \cdot \bar{y}$$

$$\begin{aligned} \bar{x} &= 2.53 & s_x &= 3.054 \\ \bar{y} &= 2.31 & s_y &= 4.024 \\ r_{xy} &= -0.9582 \end{aligned}$$

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

$$\begin{aligned} b^* &= -1.263 \\ a^* &= 5.5 \end{aligned}$$

und die Regressionsgerade $y = 5.5 - 1.263 \cdot x$.



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

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

$$y_{()} = (-2.6,\ -1.8,\ -1.5,\ 0.7,\ 1,\ 2,\ 3.3,\ 4.9,\ 8.2,\ 8.9)$$

c) Berechnen Sie das 0.2-getrimmte Stichprobenmittel $\bar{y}_{0.2}$ von $(y_1, \ldots, y_{10})$.
Lösung: Mit $k = [10 \cdot 0.2] = 2$ ergibt sich

$$\bar{y}_{0.2} = \frac{1}{10 - 2 \cdot 2} \cdot (y_{(3)} + \ldots + y_{(8)}) = 1.733$$

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 $k_2 = [7.5] = 7$

$$\begin{aligned} \tilde{y}_{0.25} &= y_{(k_1+1)} = y_{(3)} = -1.5 \\ \tilde{y}_{0.75} &= y_{(k_2+1)} = y_{(8)} = 4.9 \end{aligned}$$

und damit der Quartilsabstand zu $\tilde{y}_{0.75} - \tilde{y}_{0.25} = 6.4$.

Aufgabe 2

Die gemeinsame Zähldichte $f_{X,Y}(i,j)$ zweier Zufallszahlen X und Y ist in der folgenden Tabelle gegeben. So erhält man z.B. aus der Tabelle $\mathbb{P}(X=0, Y=1) = f_{X,Y}(0,1) = 0$, wenn man dort $i=0$ und $j=1$ setzt.

j \ i	-2	-1	0	1
0	0.05	0.05	0.05	0.05
1	0.05	0.10	0	0.05
2	0.05	0.10	0.05	0.05
3	0.15	0.15	0.05	0

a) Bestimmen Sie die Zähldichte f_Y von Y.

Lösung: Gemäß (6.5) erhält man die Zähldichte f_Y von Y als Zeilensumme der obigen Matrix. Es ist also

j	0	1	2	3
$f_Y(j)$	0.20	0.20	0.25	0.35

b) Berechnen Sie $\mathbb{E}Y$ und $\mathbb{E}Y^2$.

Lösung:

$$\begin{aligned}
\mathbb{E}Y &= \sum_{j=0}^{3} j \cdot f_Y(j) = 0 \cdot 0.20 + 1 \cdot 0.20 + 2 \cdot 0.25 + 3 \cdot 0.35 = 1.75 \\
\mathbb{E}Y^2 &= \sum_{j=0}^{3} j^2 \cdot f_Y(j) = 0 \cdot 0.20 + 1^2 \cdot 0.20 + 2^2 \cdot 0.25 + 3^2 \cdot 0.35 = 4.35
\end{aligned}$$

c) Bestimmen Sie die bedingten Wahrscheinlichkeiten $\mathbb{P}(X=i \mid Y=1)$ von X unter der Bedingung $Y=1$ für $i=-2,\ldots,1$.

Lösung: Nach Definition 10.5 und wegen b) gilt für $i=-2,\ldots,1$

$$\mathbb{P}(X=i \mid Y=1) = \frac{\mathbb{P}(X=i, Y=-1)}{\mathbb{P}(Y=1)} = \frac{f_{X,Y}(i,1)}{f_Y(1)} = 5 \cdot f_{X,Y}(i,1)$$

und damit

i	-2	-1	0	1
$\mathbb{P}(X=i \mid Y=1)$	0.25	0.50	0.00	0.25

d) Berechnen Sie $\mathbb{P}(X \leq 0 \mid Y=1)$.

Lösung: Mit c) erhält man

$$\begin{aligned}
\mathbb{P}(X \leq 0 \mid Y=1) &= \mathbb{P}(X=-2 \mid Y=1) + \mathbb{P}(X=-1 \mid Y=1) + \mathbb{P}(X=0 \mid Y=1) \\
&= 0.25 + 0.50 + 0.00 = 0.75.
\end{aligned}$$

e) Bestimmen Sie $\mathbb{P}(X+Y=2)$.

Lösung: Es ist $\{X+Y=2\} = \{X=-1, Y=3\} + \{X=0, Y=2\} + \{X=1, Y=1\}$ und damit

$$\begin{aligned} \mathbb{P}(X+Y=2) &= \mathbb{P}(X=-1, Y=3) + \mathbb{P}(X=0, Y=2) + \mathbb{P}(X=1, Y=1) \\ &= 0.15 + 0.05 + 0.05 \;=\; 0.25 \end{aligned}$$

Aufgabe 3

X sei eine normalverteilte Zufallsvariable mit dem Erwartungswert 3 und der Varianz 4 und Y eine normalverteilte Zufallsvariable mit dem Erwartungswert 6 und der Varianz 9. Die Zufallsvariablen X und Y seien stochastisch unabhängig.

a) Bestimmen Sie $\mathbb{P}(Y \leq 12)$.

Lösung: Wegen $Y \sim \mathcal{N}(\mu, \sigma^2)$ mit $\mu := \mathbb{E}Y = 6$ und $\sigma^2 := V(Y) = 9$ und (9.6) gilt

$$\mathbb{P}(Y \leq 12) = \Phi_{\mu,\sigma^2}(12) = \Phi\left(\frac{12-\mu}{\sigma}\right) = \Phi\left(\frac{12-6}{3}\right) = \Phi(2).$$

Wegen $\Phi(2) = 0.9772$ (aus Tabelle A.1) folgt

$$\mathbb{P}(Y \leq 12) = 0.9772.$$

b) Bestimmen Sie das 0.95-Quantil $q_{0.95}$ der Zufallsvariablen Y.

Lösung: $q_{0.95}$ ist die Lösung q der Gleichung

$$F_Y(q) = \Phi_{6,9}(q) = \Phi\left(\frac{q-6}{3}\right) = 0.95.$$

Wegen $\Phi(1.6449) = 0.95$ (aus Bemerkung 12.20) gilt also

$$\frac{q-6}{3} = 1.6449$$

und damit

$$q_{0.95} = q = 1.6449 \cdot 3 + 6 = 10.9347.$$

c) Es seien die Zufallsvariablen $W := aX + Y$ und $Z := X - aY$ für ein $a \in \mathbb{R}$ gegeben. Berechnen Sie die Varianzen $V(W)$ und $V(Z)$ in Abhängigkeit von a .

Hinweis: Nach Aufgabenstellung ist V(X)=4 und V(Y)=9.

Lösung:
Da X und Y stochastisch unabhängig sind, folgt mit den Rechenregeln für die Varianz (siehe Satz 12.11 c) und Satz 12.23 f))

$$\begin{aligned} V(W) &= a^2 \cdot V(X) + V(Y) = a^2 \cdot 4 + 9 = 9 + 4a^2 \\ V(Z) &= V(X) + (-a)^2 \cdot V(Y) = 4 + a^2 \cdot 9 = 4 + 9a^2 \end{aligned}$$

d) Berechnen Sie die Kovarianz $C(W, Z)$ und den Korrelationskoeffizienten $\rho(W, Z)$ der Zufallsvariablen W und Z in Abhängigkeit von a.

Lösung: Wegen Satz 12.23 und der Unabhängigkeit von X und Y, also $C(X, Y) = C(Y, X) = 0$, gilt

$$\begin{aligned} C(W, Z) &= C(aX + Y, X - aY) \\ &= C(aX, X) + C(aX, -aY) + C(Y, X) + C(Y, -aY) \\ &= aC(X, X) - aC(X, Y) + C(Y, X) - aC(Y, Y) \\ &= aV(X) - aV(Y) \ = \ -5a. \end{aligned}$$

Nach Definition 12.22 b) des Korrelationskoeffizienten gilt

$$\rho(W, Z) = \frac{C(W, Z)}{\sqrt{V(W)V(Z)}} = \frac{-5a}{\sqrt{(9 + 4a^2)(4 + 9a^2)}}$$

e) Für welche a sind W und Z unkorreliert?

Lösung: Mit d) folgt

$$W \text{ und } Z \text{ unkorreliert } \iff \rho(W, Z) = 0 \iff a = 0.$$

Aufgabe 4

Ein Gerät bestehe aus 5 gleichartigen Komponenten K_1, K_2, K_3, K_4, K_5. Die zufälligen Lebensdauern $X_1, \ldots, X_5$ der einzelnen Komponenten seien unabhängig und $\mathcal{U}(0, 5)$-verteilt, die Gleichverteilung auf $(0, 5)$. Sei $0 \leq t \leq 5$ fest.

a) Bestimmen Sie für $0 < t < 5$ die Wahrscheinlichkeit p_t, dass Komponente K_1 zum Zeitpunkt t bereits ausgefallen ist.

b) Es sei Y_t die zufällige Anzahl der Komponenten, die zum Zeitpunkt t bereits ausgefallen sind. Welche Verteilung besitzt Y_t? Geben Sie die Parameter dieser Verteilung an!

c) Das Gerät ist intakt, solange höchstens zwei Komponenten ausgefallen sind.

   c_1) Drücken Sie das Ereignis $A_t :=$ „Das Gerät ist zum Zeitpunkt t intakt“ mit Hilfe von Y_t aus.

   c_2) Wie groß ist die Wahrscheinlichkeit $\mathbb{P}(A_t)$, dass das Gerät zum Zeitpunkt t noch intakt ist?

d) Berechnen Sie die Wahrscheinlichkeit $\mathbb{P}(A_3)$ (auf 3 Nachkommastellen genau), dass das Gerät zum Zeitpunkt $t = 3$ intakt ist.

Lösung:

a) Komponente K_1 ist zum Zeitpunkt t genau dann schon ausgefallen, wenn $X_1 \leq t$ gilt. Daher

$$p_t = \mathbb{P}(X_1 \leq t) = F_{X_1}(t) = \frac{t}{5}.$$

b) Es liegt ein Laplace-Experiment vor mit Erfolgswahrscheinlichkeit p_t, das 5-mal durchgeführt wird. Daher gilt

$$Y_t \sim Bin(5, p_t) = Bin\left(5, \frac{t}{5}\right).$$

c) Das Gerät ist zum Zeitpunkt t intakt, wenn dann höchstens zwei Komponenten ausgefallen sind. Es gilt daher

c$_1$) $A_t = \{Y_t \leq 2\}$ und

c$_2$) $\mathbb{P}(A_t) = \mathbb{P}(Y_t \leq 2) = f_{Y_t}(0) + f_{Y_t}(1) + f_{Y_t}(2)$
$= \binom{5}{0} \cdot p_t^0 \cdot (1-p_t)^5 + \binom{5}{1} \cdot p_t^1 \cdot (1-p_t)^4 + \binom{5}{2} \cdot p_t^2 \cdot (1-p_t)^3$
$= (1-p_t)^5 + 5 \cdot p_t \cdot (1-p_t)^4 + 10 \cdot p_t^2 \cdot (1-p_t)^3$

d) Wegen c) ist die Wahrscheinlichkeit, dass das Gerät zum Zeitpunkt $t = 3$ intakt ist, mit $p_3 = 0.6$ gerade

$$\mathbb{P}(A_3) = 0.4^5 + 5 \cdot 0.4^4 \cdot 0.6 + 10 \cdot 0.4^3 \cdot 0.6^2 = 0.31744$$

Aufgabe 5

Die maximale Windgeschwindigkeit an einem bestimmten Ort und an einem bestimmten Tag sei (gemessen in geeigneten Einheiten) eine Zufallsvariable $X > 0$ mit der Dichte

$$f_\vartheta(t) \;=\; \begin{cases} 0 & , \quad t \leq 0 \\ \frac{1}{\vartheta \cdot t^2} \cdot e^{-1/(\vartheta \cdot t)} & , \quad t > 0 \end{cases}$$

für ein festes, unbekanntes $\vartheta > 0$.

a) Bestimmen Sie den Erwartungswert $\mathbb{E}_\vartheta(1/X)$ und die Varianz $V_\vartheta(1/X)$.
Hinweis: Die Zufallsvariable $1/X$ hat die Verteilung $Exp(1/\vartheta)$.

b) Bestimmen Sie die Loglikelihood-Funktion $M_x(\vartheta)$ zur Stichprobe $x = (x_1, \ldots, x_n)$, wobei $x_i > 0$ für $i = 1, \ldots, n$.

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

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

Lösung:

a) Setze $Y := 1/X$. Nach dem Hinweis gilt $Y \sim Exp(1/\vartheta)$ und damit wegen der Tabellen auf den Seiten 124 und 128

$$\mathbb{E}(1/X) = \mathbb{E}Y = \frac{1}{1/\vartheta} = \vartheta$$

und

$$V(1/X) = V(Y) = \frac{1}{(1/\vartheta)^2} = \vartheta^2.$$

b) Wegen

$$\ln(f_\vartheta(t)) = -\ln(\vartheta) - 2 \cdot \ln(t) - \frac{1}{\vartheta t}$$

bestimmt sich die Loglikelihood-Funktion zu

$$\begin{aligned} M_x(\vartheta) &= \sum_{j=1}^{n} \ln(f_\vartheta(x_j)) \;=\; \sum_{j=1}^{n} \left[-\ln(\vartheta) - 2 \cdot \ln(x_j) - \frac{1}{\vartheta x_j} \right] \\ &= -n \cdot \ln(\vartheta) - 2 \cdot \sum_{j=1}^{n} \ln(x_j) - \frac{1}{\vartheta} \cdot \sum_{j=1}^{n} \frac{1}{x_j}. \end{aligned}$$

c) Ihre Ableitung ist

$$M_x'(\vartheta) \;=\; -\frac{n}{\vartheta} + \frac{1}{\vartheta^2} \cdot \sum_{j=1}^{n} \frac{1}{x_j} \;=\; \frac{n}{\vartheta^2} \cdot \left(-\vartheta + \frac{1}{n} \cdot \sum_{j=1}^{n} \frac{1}{x_j} \right).$$

d) Wegen

$$M_x'(\vartheta) \left\{ \begin{matrix} > \\ = \\ < \end{matrix} \right\} 0 \qquad \Longleftrightarrow \qquad \vartheta \left\{ \begin{matrix} < \\ = \\ > \end{matrix} \right\} \frac{1}{n} \cdot \sum_{j=1}^{n} \frac{1}{x_j} =: \hat{\vartheta}(x)$$

ist M_x unterhalb von $\hat{\vartheta}(x)$ steigend, oberhalb von $\hat{\vartheta}(x)$ fallend, also

$$\hat{\vartheta}(x) = \frac{1}{n} \cdot \sum_{j=1}^{n} \frac{1}{x_j}$$

der gesuchte Maximum-Likelihood-Schätzer.

Aufgabe 6

Eine Firma besitzt 17 gleichartige Computer, welche unabhängig voneinander und unter gleichen Bedingungen 4 Jahre eingesetzt worden waren. Dabei waren 12 dieser Computer während der vier Jahre ohne Reparatur einsatzfähig, während die übrigen Computer mindestens einmal ausfielen.
Es sei p die Wahrscheinlichkeit, dass ein Computer dieser Bauart 4 Jahre ohne Reparatur einsatzfähig ist. Geben Sie ein (zweiseitiges) Konfidenzintervall zum Konfidenzniveau 0.9 für die Wahrscheinlichkeit p an.

Lösung: Nach Voraussetzung kann wie in Beispiel 18.5 ein ideales Zufallsexperiment mit den zwei möglichen Ergebnissen „Niemals repariert“ (Treffer) und „Mindestens einmal ausgefallen“ (Niete) und der Trefferwahrscheinlichkeit $\vartheta := p$ angesehen werden. Nach diesem Beispiel und wegen 18.6 ist das gesuchte Konfidenzintervall $[l(x), L(x)]$, wobei die Konfidenzgrenzen $l(x)$ und $L(x)$ für $x = 12$ und $n - x = 5$ und $1 - \alpha = 0.9$ aus Tabelle A.4 entnommen werden. Dies ergibt das Konfidenzintervall

$$\mathcal{C}_1(x) = [0.478\ ,\ 0.876]\ .$$