

Klausur (Maschineningenieure)
Wahrscheinlichkeitstheorie und Statistik
vom 3.3.2009
Musterlösungen

Aufgabe 1:

Gegeben sei eine Urliste mit den Paaren $(x_1, y_1), \dots, (x_{11}, y_{11})$

| j | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|-------|-----|------|------|-----|-----|-----|------|-----|------|------|------|
| x_j | -3 | -1.7 | -1.2 | 0.1 | 0.9 | 2.1 | 3.1 | 3.9 | 5.5 | 5.9 | 7.2 |
| y_j | 5.2 | 16.6 | 5.9 | 6.5 | 6.9 | 2.1 | -3.3 | 0.1 | -8.4 | -2.6 | -7.5 |

- 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

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

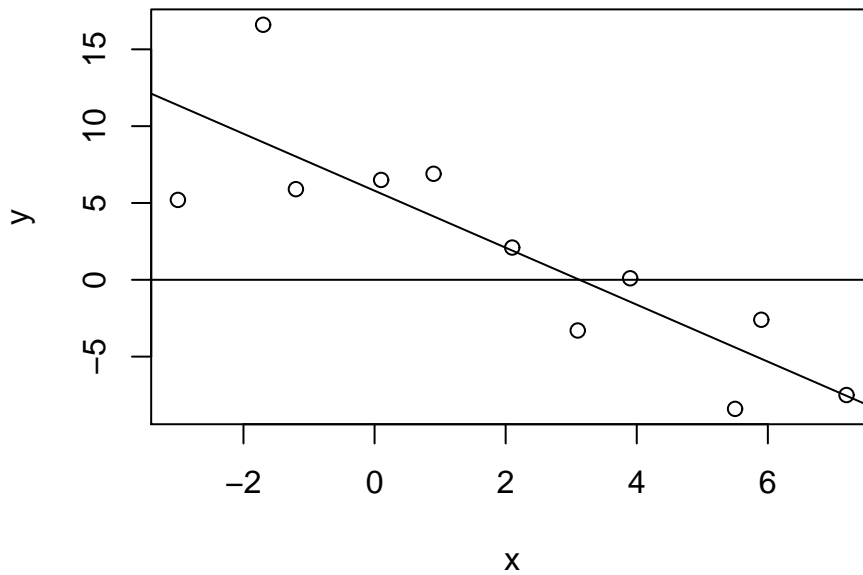
$$\begin{aligned}\bar{x} &= 2.07 & s_x &= 3.362 \\ \bar{y} &= 1.95 & s_y &= 7.308 \\ r_{xy} &= -0.853\end{aligned}$$

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

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

$$\begin{aligned}b^* &= -1.854 \\ a^* &= 5.8\end{aligned}$$

und die Regressionsgerade $y = 5.8 - 1.854 \cdot x$.



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

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

$$y_{()} = (-8.4, -7.5, -3.3, -2.6, 0.1, 2.1, 5.2, 5.9, 6.5, 6.9, 16.6)$$

- c) Berechnen Sie das 0.2-getrimmte Stichprobenmittel $\bar{y}_{0.2}$ von (y_1, \dots, y_{11}) .

Lösung: Mit $k = \lceil 11 \cdot 0.2 \rceil = 2$ ergibt sich

$$\bar{y}_{0.2} = \frac{1}{11 - 2 \cdot 2} \cdot (y_{(3)} + \dots + y_{(9)}) = 1.986$$

- d) Bestimmen Sie das Stichproben-0.2-Quantil $\tilde{y}_{0.2}$ von (y_1, \dots, y_{11}) .

Lösung: Da $11 \cdot 0.2 = 2.2$ nicht ganzzahlig ist, ist mit $k = \lceil 2.2 \rceil = 2$

$$\tilde{y}_{0.2} = y_{(k+1)} = y_{(3)} = -3.3$$

- e) Berechnen Sie den Quartilsabstand von (y_1, \dots, y_{11}) .

Lösung: Da $0.25 \cdot 11 = 2.75$ und $0.75 \cdot 11 = 8.25$ beide nicht ganzzahlig sind, ergibt sich mit $k_1 = \lceil 2.75 \rceil = 2$ und $k_2 = \lceil 8.25 \rceil = 8$

$$\tilde{y}_{0.25} = y_{(k_1+1)} = y_{(3)} = -3.3$$

$$\tilde{y}_{0.75} = y_{(k_2+1)} = y_{(9)} = 6.5$$

und damit der Quartilsabstand zu $\tilde{y}_{0.75} - \tilde{y}_{0.25} = 9.8$.

Aufgabe 2

Ein LKW wird insgesamt n -mal durch einen Kran mit Kies beladen. Bei jedem Beladungsvorgang i erfasst der Kran unabhängig von den anderen Beladungsvorgängen eine zufällige Masse X_i an Kies, die normalverteilt ist mit Mittelwert $\mu = 250$ (in kg) und Varianz $\sigma^2 = 2500$ (in kg^2), wobei $i = 1, \dots, n$.

- Welche Verteilung besitzt S_n , die zufällige Gesamtmasse an Kies bei n unabhängigen Beladungsvorgängen?
- Der LKW selbst hat eine Masse von 4000 kg. Welche Verteilung besitzt G_n , die Masse des LKW's inklusive der Beladung nach n unabhängigen Beladungsvorgängen?
- Wie groß ist die Wahrscheinlichkeit, dass die Gesamtmasse S_{24} mindestens 5700 kg beträgt, wenn der LKW mit $n = 24$ Kranfüllungen beladen wird.
- Wie groß muss n mindestens sein, damit $\mathbb{P}(S_n \geq 6000) \geq 0.95$ gilt?

Hinweis:

| | |
|-----|---|
| n | $5 \cdot \sqrt{n} - \frac{120}{\sqrt{n}}$ |
| 24 | 0.00 |
| 25 | 1.00 |
| 26 | 1.96 |
| 27 | 2.89 |

Lösung: $S_n = X_1 + \dots + X_n$ ist die zufällige Gesamtmasse.

- Nach Voraussetzung gilt $X_i \sim \mathcal{N}(250, 2500)$ für $i = 1, \dots, n$. Da X_1, \dots, X_n stochastisch unabhängig sind, gilt nach der Faltungsformel (Skriptum, 11.16) $S_n \sim \mathcal{N}(250 \cdot n, n \cdot 2500)$.
- Wegen $G_n = S_n + 4000$ und wegen a) gilt $G_n \sim \mathcal{N}(4000 + 250 \cdot n, n \cdot 2500)$.
- Für den Fall $n = 24$ gilt $S_n \sim \mathcal{N}(n \cdot 250, n \cdot 2500) = \mathcal{N}(6000, 60000)$. Daher ist die gesuchte Wahrscheinlichkeit

$$\begin{aligned}\mathbb{P}(S_{24} \geq 5700) &= 1 - \mathbb{P}(S_{24} < 5700) \stackrel{(*)}{=} 1 - \mathbb{P}(S_{24} \leq 5700) = 1 - \Phi\left(\frac{5700 - 6000}{\sqrt{60000}}\right) \\ &= 1 - \Phi(-1.2247) \approx 1 - \Phi(-1.22) = \Phi(1.22) = 0.8888\end{aligned}$$

An Stelle (*) wurde Satz 8.11 verwendet und ausgenützt, dass S_{24} eine stetige Zufallsvariable ist.

- Gesucht ist das kleinste n , so dass $\mathbb{P}(S_n \geq 6000) \geq 0.95$ gilt, d.h. wie in b), dass

$$\begin{aligned}\mathbb{P}(S_n \geq 6000) &= 1 - \mathbb{P}(S_n \leq 6000) = 1 - \Phi\left(\frac{6000 - 250n}{\sqrt{2500n}}\right) = 1 - \Phi\left(\frac{6000 - 250n}{50 \cdot \sqrt{n}}\right) \\ &= 1 - \Phi\left(\frac{120 - 5n}{\sqrt{n}}\right) = 1 - \Phi\left(\frac{120}{\sqrt{n}} - 5\sqrt{n}\right) \stackrel{(**)}{=} \Phi\left(5 \cdot \sqrt{n} - \frac{120}{\sqrt{n}}\right) \geq 0.95\end{aligned}$$

gilt, wobei an Stelle (**) Skriptum, (9.5) ausgenützt wurde. Wegen dem Hinweis und wegen $\Phi(0) = 0.5$, $\Phi(1) = 0.8413$, $\Phi(1.96) = 0.975 > 0.95$ gilt also $\mathbb{P}(S_n \geq 6000) \geq 0.95$ für $n = 26, 27, \dots$ n muss also mindestens 26 sein.

Aufgabe 3

Eine Urne enthalte 10 gleichgroße Kugeln in den Farben rot und blau, die 20 g oder 50 g wiegen, entsprechend den Anzahlen:

| | rot | blau |
|------|-----|------|
| 20 g | 4 | 3 |
| 50 g | 2 | 1 |

Man bestimme die Wahrscheinlichkeit dafür, dass eine „rein zufällig“ herausgezogene Kugel

- rot ist,
- blau ist und 50 g wiegt,
- 50 g wiegt.
- Berechnen Sie die bedingte Wahrscheinlichkeit, dass eine „rein zufällig“ herausgezogene Kugel rot ist, wenn bekannt ist, dass sie 50 g wiegt.

Wie groß ist die Wahrscheinlichkeit dafür, dass bei „rein zufälliger“ aufeinanderfolgender Entnahme von Kugeln zuerst eine blaue zu 50 g, dann eine beliebig schwere rote, dann eine beliebig schwere blaue und zuletzt eine blaue Kugel zu 50 g gezogen wird, wenn die entnommenen Kugeln

- jeweils vor der nächsten Ziehung wieder zurückgelegt werden,
- nicht wieder zurückgelegt werden?

Hinweis: Die auftretenden Brüche brauchen nicht ausgerechnet zu werden.

Lösung: Sei A das Ereignis, dass eine rote Kugel gezogen wird und B das Ereignis, dass die gezogene Kugel 50 g wiegt.

- Da in der Urne insgesamt 10 Kugeln sind und davon 6 rot sind, gilt $\mathbb{P}(A) = \frac{6}{10} = 0.6$.
- Da in der Urne insgesamt 10 Kugeln sind und davon 1 blau ist und 50 g wiegt, ist die gesuchte Wahrscheinlichkeit $\mathbb{P}(A^c \cap B) = \frac{1}{10} = 0.1$.
- Da in der Urne insgesamt 10 Kugeln sind und davon 3 20 g wiegen, ist die gesuchte Wahrscheinlichkeit $\mathbb{P}(B) = \frac{3}{10} = 0.3$.
- Gesucht ist die bedingte Wahrscheinlichkeit $\mathbb{P}(A | B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$. $A \cap B$ ist das Ereignis, dass die gezogene Kugel rot ist und 50 g wiegt. Da in der Urne 2 derartige Kugeln sind, gilt $\mathbb{P}(A \cap B) = \frac{2}{10} = 0.2$ und damit

$$\mathbb{P}(A | B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{2/10}{3/10} = \frac{2}{3}$$

- Da die entnommenen Kugeln jedesmal zurückgelegt werden, liegt ein ideales Zufallsexperiment vor, das viermal durchgeführt wird und die gesuchte Wahrscheinlichkeit ist wegen der Produktregel

$$\mathbb{P}(A^c \cap B) \cdot \mathbb{P}(A) \cdot \mathbb{P}(A^c) \cdot \mathbb{P}(A^c \cap B) = \frac{1}{10} \cdot \frac{6}{10} \cdot \frac{4}{10} \cdot \frac{1}{10} = \frac{24}{10000}$$

- f) Die gesuchte Wahrscheinlichkeit ist 0, da es unmöglich ist, zweimal eine blaue Kugel zu 50 g zu ziehen, wenn die gezogene Kugel nicht zurückgelegt wird und insgesamt nur eine derartige Kugel vorhanden ist.

Aufgabe 4

Ein Autohändler verkauft je Tag X Autos des Typs **1** und Y Autos des Typs **2**. Die Tabelle zeigt die gemeinsame Zähldichte $f_{X,Y}(i, j) = \mathbb{P}(X = i, Y = j)$ von X und Y für $i, j = 0, 1, 2$.

| | | | | |
|-----|---|-----|-----|-----|
| | | j | | |
| | | 0 | 1 | 2 |
| i | 0 | 0.1 | 0.1 | 0.0 |
| | 1 | 0.1 | 0.3 | 0.1 |
| | 2 | 0.0 | 0.2 | 0.1 |

Es ist also z.B. $\mathbb{P}(X = 2, Y = 1) = f_{X,Y}(2, 1) = 0.2$.

- Bestimmen Sie die Zähldichten f_X von X und f_Y von Y .
- Berechnen Sie die bedingte Wahrscheinlichkeit $\mathbb{P}(X \geq 1 \mid Y = 1)$.
- Sind X und Y stochastisch unabhängig?
- Berechnen Sie die Erwartungswerte $\mathbb{E}X$, $\mathbb{E}Y$ und die Varianzen $V(X)$ und $V(Y)$.
- Berechnen Sie den Korrelationskoeffizienten $\rho(X, Y)$ von X und Y . Sind X und Y positiv, negativ oder unkorreliert?

Lösung:

- Die Bestimmung der Zähldichte von X bzw. Y geschieht durch Summation der Wahrscheinlichkeiten der gemeinsamen Verteilung über die möglichen Werte von Y bzw. X (vergl. Skriptum Beispiel 6.7 und Tabelle 6.1)

| | | | | | |
|-----|----------|----------|-----|-----|----------|
| | | j | | | |
| | | 0 | 1 | 2 | Σ |
| i | 0 | 0.1 | 0.1 | 0.0 | 0.2 |
| | 1 | 0.1 | 0.3 | 0.1 | 0.5 |
| | 2 | 0.0 | 0.2 | 0.1 | 0.3 |
| | Σ | 0.2 | 0.6 | 0.2 | 1.0 |
| | | $f_Y(j)$ | | | $f_X(i)$ |

- Nach Definition gilt $\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$, falls $\mathbb{P}(B) > 0$. Da $A \rightarrow \mathbb{P}(A|B)$ ein Wahrscheinlichkeitsmaß ist (Satz 10.11), gilt

$$\begin{aligned} \mathbb{P}(X \geq 1 \mid Y = 1) &= \mathbb{P}(X = 1 \mid Y = 1) + \mathbb{P}(X = 2 \mid Y = 1) \\ &= \frac{\mathbb{P}(X = 1, Y = 1)}{\mathbb{P}(Y = 1)} + \frac{\mathbb{P}(X = 2, Y = 1)}{\mathbb{P}(Y = 1)} = \frac{0.3}{0.6} + \frac{0.2}{0.6} = \frac{5}{6} = 0.8333 \end{aligned}$$

- X und Y sind nicht stochastisch unabhängig, denn $\mathbb{P}(X = 0, Y = 2) = 0.0 \neq 0.2 \cdot 0.2 = \mathbb{P}(X = 0) \cdot \mathbb{P}(Y = 2)$.

d) Es gilt

$$\begin{aligned}\mathbb{E}X &= 0 \cdot f_X(0) + 1 \cdot f_X(1) + 2 \cdot f_X(2) = 0 \cdot 0.2 + 1 \cdot 0.5 + 2 \cdot 0.3 = 1.1 \\ \mathbb{E}X^2 &= 0^2 \cdot f_X(0) + 1^2 \cdot f_X(1) + 2^2 \cdot f_X(2) = 0 \cdot 0.2 + 1 \cdot 0.5 + 4 \cdot 0.3 = 1.7\end{aligned}$$

und damit dann $V(X) = \mathbb{E}X^2 - (\mathbb{E}X)^2 = 1.7 - 1.1^2 = 0.49$.

Analog

$$\begin{aligned}\mathbb{E}Y &= 0 \cdot f_Y(0) + 1 \cdot f_Y(1) + 2 \cdot f_Y(2) = 0 \cdot 0.2 + 1 \cdot 0.6 + 2 \cdot 0.2 = 1.0 \\ \mathbb{E}Y^2 &= 0^2 \cdot f_Y(0) + 1^2 \cdot f_Y(1) + 2^2 \cdot f_Y(2) = 0 \cdot 0.2 + 1 \cdot 0.6 + 4 \cdot 0.2 = 1.4\end{aligned}$$

und damit dann $V(Y) = \mathbb{E}Y^2 - (\mathbb{E}Y)^2 = 1.4 - 1.0^2 = 0.4$.

e) Es gilt $C(X, Y) = \mathbb{E}(X \cdot Y) - \mathbb{E}X \cdot \mathbb{E}Y$. Aus der Tabelle ergibt sich

$$\begin{aligned}\mathbb{E}(X \cdot Y) &= 1 \cdot 1 \cdot f_{X,Y}(1, 1) + 1 \cdot 2 \cdot f_{X,Y}(1, 2) + 2 \cdot 1 \cdot f_{X,Y}(2, 1) + 2 \cdot 2 \cdot f_{X,Y}(2, 2) \\ &= 1 \cdot 0.3 + 2 \cdot 0.1 + 2 \cdot 0.2 + 4 \cdot 0.1 = 1.3\end{aligned}$$

und damit $C(X, Y) = 1.3 - 1.1 \cdot 1.0 = 0.2$, also

$$\rho(X, Y) = \frac{C(X, Y)}{\sqrt{V(X) \cdot V(Y)}} = \frac{0.2}{\sqrt{0.49 \cdot 0.4}} = 0.4518$$

Wegen $\rho(X, Y) > 0$ sind X und Y positiv korreliert.

Aufgabe 5

Eine Zufallsvariable X habe die Verteilung $Exp(\lambda)$ mit dem Parameter $\lambda > 0$ und es sei $Y = \frac{X}{1+X}$ mit festem Parameter $\lambda > 0$.

a) Bestimmen Sie den Median von X für den Spezialfall $\lambda = \ln(2)$.

b) Skizzieren Sie die Funktion $x \rightarrow \frac{x}{1+x}$ für $0 \leq x \leq 10$. Welche Werte nimmt also die Zufallsvariable Y an?

c) Berechnen Sie den Erwartungswert $\mathbb{E} \frac{Y^2}{(1-Y)^2}$.

Hinweis: $\frac{Y^2}{(1-Y)^2}$ ist eine einfache Funktion der Zufallsvariablen X .

d) Berechnen Sie $\mathbb{P}(Y \leq t)$ für $0 < t < 1$ und begründen Sie, dass Y die Verteilungsfunktion

$$t \rightarrow F_Y(t) := \begin{cases} 0 & , t \leq 0 \\ 1 - e^{-\frac{\lambda t}{1-t}} & , 0 < t < 1 \\ 1 & , t \geq 1 \end{cases}$$

besitzt.

e) Bestimmen Sie die Dichte f_Y von Y .

Lösung:

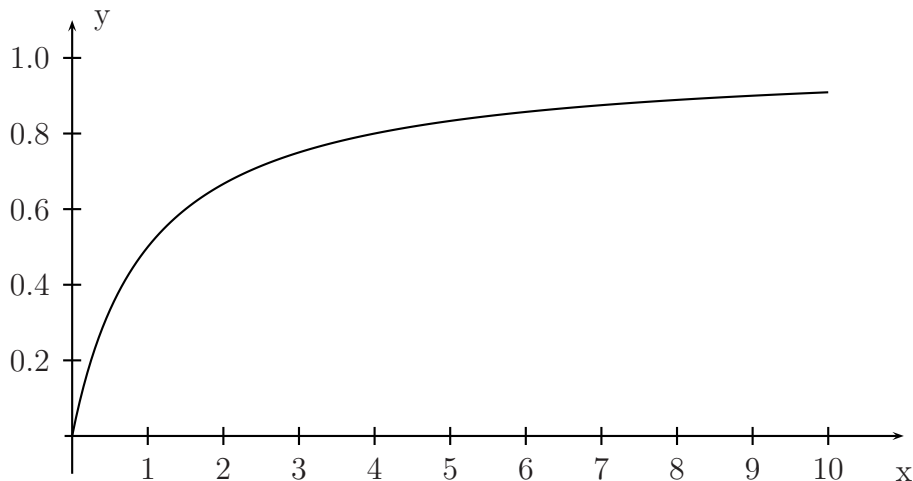
- a) Nach Beispiel 12.21 besitzt die Zufallsvariable X allgemein das p -Quantil

$$t_p = -\frac{1}{\lambda} \cdot \ln(1 - p).$$

Für den Median ist $p = 1/2$. Speziell für $\lambda = \ln(2)$ ergibt sich

$$t_{1/2} = -\frac{1}{\ln(2)} \cdot \ln(1 - 1/2) = -\frac{1}{\ln(2)} \cdot \ln(1/2) = 1.$$

- b) Direkt aus der nachfolgenden Figur folgt, dass Y alle Werte zwischen 0 und 1 annimmt.



- c) Es gilt $\frac{Y^2}{(1-Y)^2} = \frac{\left(\frac{X}{1+X}\right)^2}{\left(1 - \frac{X}{1+X}\right)^2} = \frac{\frac{X^2}{(1+X)^2}}{\frac{1}{(1+X)^2}} = X^2$. Da X die Verteilung $Exp(\lambda)$ besitzt, folgt aus Beispiel 12.16 a) (mit $\alpha = 1$ und $\beta = \lambda$)

$$\mathbb{E} \frac{Y^2}{(1-Y)^2} = \mathbb{E} X^2 = \frac{2}{\lambda^2}.$$

Alternativ erhalte man das Resultat auch mit $\mathbb{E} X^2 = V(X) + (\mathbb{E} X)^2 = \frac{1}{\lambda^2} + \left(\frac{1}{\lambda}\right)^2$.

- d) Da Y Werte zwischen 0 und 1 annimmt, gilt $F_Y(t) = \mathbb{P}(Y \leq t) = 0$ für $t \leq 0$ und $F_Y(t) = \mathbb{P}(Y \leq t) = \mathbb{P}(Y \leq 1) = 1$ für $t \geq 1$. Für $0 < t < 1$ gilt schließlich

$$\begin{aligned} F_Y(t) &= \mathbb{P}(Y \leq t) = \mathbb{P}\left(\frac{X}{1+X} \leq t\right) = \mathbb{P}(X \leq t \cdot (1+X)) = \mathbb{P}(X - t \cdot X \leq t) \\ &= \mathbb{P}((1-t) \cdot X \leq t) = \mathbb{P}\left(X \leq \frac{t}{1-t}\right) = 1 - e^{-\lambda \cdot \frac{t}{1-t}}, \end{aligned}$$

da X die Verteilungsfunktion $x \rightarrow 1 - e^{-\lambda \cdot x}$, $x > 0$ besitzt.

- e) Nach dem Hinweis ist die Verteilungsfunktion F_Y stetig. Aus d) folgt, dass F_Y bis auf eventuell die Stellen 0 und 1 auch stetig differenzierbar ist, so dass Satz 8.12 anwendbar ist. Da $F_Y(t)$ für $t \leq 0$ und auch für $t \geq 1$ konstant ist, folgt $f_Y(y) = 0$ für $y \leq 0$ und für $y \geq 1$. Für $0 < y < 1$ folgt schließlich mit der Kettenregel

$$f_Y(y) = F'_Y(y) = \lambda \cdot \left(\frac{y}{1-y}\right)' \cdot e^{-\lambda \cdot \frac{y}{1-y}} = \frac{\lambda}{(1-y)^2} \cdot e^{-\lambda \cdot \frac{y}{1-y}}.$$

Aufgabe 6

Ein Merkmal habe die Dichte der Erlang-Verteilung $E(3, 1/\vartheta)$

$$t \rightarrow f_\vartheta(t) := \frac{t^2}{2\vartheta^3} \cdot e^{-t/\vartheta}, \quad t > 0,$$

wobei $\vartheta \in \Theta := (0, \infty)$ ein unbekannter Parameter ist. ϑ soll aufgrund einer unabhängigen Stichprobe $x = (x_1, \dots, x_n)$ geschätzt werden, wobei $x_1 > 0, \dots, x_n > 0$ vorausgesetzt wird.

- Bestimmen Sie den Maximum-Likelihood-Schätzer $\hat{\vartheta}(x_1, \dots, x_n)$ für ϑ .
- Berechnen Sie den Erwartungswert von $\hat{\vartheta}(X) = \hat{\vartheta}(X_1, \dots, X_n)$ bei Vorliegen von ϑ für den Fall, dass die unabhängigen Stichprobenvariablen X_1, \dots, X_n die Erlang-Verteilung $E(3, 1/\vartheta)$ besitzen.

Hinweis: Bestimmen Sie zuerst $\mathbb{E}_\vartheta(X_j)$ für $j = 1, \dots, n$.

- Welche der drei folgenden Eigenschaften sind für den Maximum-Likelihood-Schätzer $\hat{\vartheta}(x_1, \dots, x_n)$ erfüllt?
 - $\hat{\vartheta}(x_1, \dots, x_n)$ ist erwartungstreu für ϑ .
 - Die Schätzfolge $(\hat{\vartheta}(x_1, \dots, x_n))_{n \in \mathbb{N}}$ ist asymptotisch erwartungstreu für ϑ .
 - Die Schätzfolge $(\hat{\vartheta}(x_1, \dots, x_n))_{n \in \mathbb{N}}$ ist konsistent für ϑ .Begründen Sie Ihre Antwort! Bei Aufgabe c₃) dürfen Sie vom Erfülltsein geeigneter Regularitätsvoraussetzungen ausgehen.

Lösung:

- Es ist

$$\log(f_\vartheta(t)) = 2 \log(t) - \log(2) - 3 \log(\vartheta) - \frac{t}{\vartheta}, \quad t > 0,$$

also

$$\begin{aligned} M_x(\vartheta) &= \sum_{j=1}^n \log(f_\vartheta(x_j)) = \sum_{j=1}^n \left(2 \log(x_j) - \log(2) - 3 \log(\vartheta) - \frac{x_j}{\vartheta} \right) \\ &= 2 \sum_{j=1}^n \log(x_j) - n \log(2) - 3n \log(\vartheta) - \frac{1}{\vartheta} \sum_{j=1}^n x_j. \end{aligned}$$

Es gilt

$$M'_x(\vartheta) = -\frac{3n}{\vartheta} + \frac{1}{\vartheta^2} \sum_{j=1}^n x_j = -\frac{3n}{\vartheta^2} \left(\vartheta - \frac{1}{3n} \sum_{j=1}^n x_j \right) \begin{matrix} > \\ = 0 \\ < \end{matrix}$$

genau dann, wenn (beachte $-\frac{3n}{\vartheta^2} < 0$)

$$\vartheta \begin{matrix} < \\ = \\ > \end{matrix} \frac{1}{3n} \sum_{j=1}^n x_j =: \hat{\vartheta}(x).$$

Daher ist $\hat{\vartheta}(x)$ die einzige Maximumstelle von $\vartheta \rightarrow M_x(\vartheta)$, insbesondere ist

$$(x_1, \dots, x_n) \rightarrow \hat{\vartheta}(x) = \frac{1}{3n} \sum_{j=1}^n x_j$$

der gesuchte Maximum-Likelihood-Schätzer für ϑ .

- b) Da die Stichprobenvariablen X_1, \dots, X_n die Erlang-Verteilung $\Gamma(3, 1/\vartheta)$ (eine spezielle Gamma-Verteilung) besitzen, gilt nach der Tabelle vor Satz 12.8 $\mathbb{E}_\vartheta X_j = 3\vartheta$ und damit

$$\mathbb{E}_\vartheta \hat{\vartheta}(X) = \mathbb{E}_\vartheta \left(\frac{1}{3n} \sum_{j=1}^n X_j \right) = \frac{1}{3n} \sum_{j=1}^n \mathbb{E}_\vartheta X_j = \frac{1}{3n} \cdot n \cdot 3\vartheta = \vartheta.$$

- c) Direkt aus b) und der Definition der Erwartungstreue in Definition 17.11 folgt, dass $\hat{\vartheta}(x_1, \dots, x_n)$ erwartungstreu für ϑ ist. Damit gilt c_1). Wegen

$$\lim_{n \rightarrow \infty} \mathbb{E}_\vartheta(\hat{\vartheta}(X_1, \dots, X_n)) = \lim_{n \rightarrow \infty} \vartheta = \vartheta$$

ist die Schätzfolge $(\hat{\vartheta}(x_1, \dots, x_n))_{n \in \mathbb{N}}$ damit auch asymptotisch erwartungstreu für ϑ und es gilt c_2). Um zu begründen, dass die Schätzfolge $(\hat{\vartheta}(x_1, \dots, x_n))_{n \in \mathbb{N}}$ auch konsistent für ϑ ist, kann man darauf verweisen, dass Maximum-Likelihood-Schätzer unter schwachen Regularitätsvoraussetzungen (die hier erfüllt sind und was wir nicht zeigen wollen) konsistent sind. Eine präzisere Begründung (für die Interessierten) erhält man durch Verweis auf Satz 17.19a):

Der Schätzer \bar{x} ist konsistent für $\gamma(\vartheta) := \mathbb{E}_\vartheta(X_1) = 3\vartheta$ und daher gilt

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbb{P}_\vartheta(|\hat{\vartheta}(X_1, \dots, X_n) - \vartheta| \geq \epsilon) &= \lim_{n \rightarrow \infty} \mathbb{P}_\vartheta(|3\hat{\vartheta}(X_1, \dots, X_n) - 3\vartheta| \geq 3\epsilon) \\ &= \lim_{n \rightarrow \infty} \mathbb{P}_\vartheta(|\bar{X} - \gamma(\vartheta)| \geq 3\epsilon) = 0 \end{aligned}$$

für beliebiges $\epsilon > 0$.

Aufgabe 7

Eine Firma erhält eine Lieferung mit $N = 500$ Bauteilen. Zur Überprüfung der Qualität werden 50 der Bauteile überprüft. Dabei stellt sich heraus, dass 9 der überprüften Bauteile unbrauchbar sind.

- a) Bestimmen Sie ein zweiseitiges Konfidenzintervall zum Konfidenzniveau $1 - \alpha = 0.95$ für den Anteil unbrauchbarer Bauteile an der Lieferung.
- b) Welches zweiseitige Konfidenzintervall für den Anteil unbrauchbarer Bauteile ergäbe sich zum Konfidenzniveau $1 - \alpha = 0.95$, wenn Sie von einer „praktisch“ unendlich großen (d.h. $N = \infty$) Lieferung ausgehen?

Lösung:

- a) Hier ist Beispiel 18.12 (Anteils-Schätzung bei endlicher Population) aus dem Skriptum anwendbar. In der Lieferung sind $N = 500$ Bauteile, von denen r (r unbekannt) unbrauchbar, die übrigen brauchbar sind. Es ist dann $T_n = 9/50 = 0.18$ der relative Anteil der unbrauchbaren Bauteile in der zufälligen Stichprobe vom Umfang $n = 50$. Mit $\vartheta := r/N$ (= unbekannter Anteil der unbrauchbaren Bauteile), $h := u_{1-\alpha/2} = 1.9600$ und dem Endlichkeitskorrektur-Faktor $\gamma := 1 - \frac{n-1}{N-1} = 0.9018$ bilden dann

$$\frac{T_n + \frac{h^2}{2n}\gamma \pm \frac{h}{\sqrt{n}} \cdot \sqrt{T_n(1-T_n)\gamma + \frac{h^2}{4n}\gamma^2}}{1 + \frac{h^2}{n}\gamma}$$

den oberen bzw. unteren Endpunkt eines approximativen $(1 - \alpha)$ -Konfidenzintervalles für $\vartheta (= \frac{r}{N})$.

Setzt man die Werte ein, so ergibt sich für den unbekanntem Anteil r/N das Konfidenzintervall

$$\mathcal{C}(x) = [0.1008, 0.3007] .$$

- b) Für den Fall $N = \infty$ ist der Endlichkeitskorrektur-Faktor $\gamma = 1$. Einsetzen in die obige Formel ergibt dann für den unbekanntem Anteil unbrauchbarer Teile das Konfidenzintervall

$$\mathcal{C}(x) = [0.0977, 0.3080] .$$