Karlsruher Institut für Technologie

Institut für Stochastik

PD Dr. Steffen Winter

Celeste Mayer, M.Sc.

Lösungsvorschlag Klausur zur Vorlesung Wahrscheinlichkeitstheorie und Statistik für die Fachrichtungen Maschinenbau/ Informatik

Verteilungsfunktion $\Phi(x)$ der Standard – Normalverteilung $\mathcal{N}(0,1)$

x	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7703	0.7734	0.7764	0.7793	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.7	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.8	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Aufgabe 1 (11 Punkte)

In einem Labor wurden an 10 Proben ein Merkmal X gemessen und in einer Tabelle notiert. Im Anschluss wurde der Mittelwert berechnet $\overline{x} = 2.5$. Durch einen technischen Fehler gingen **nach** der Berechnung des Mittelwerts die zwei Messwerte x_9 und x_{10} verloren, sodass Ihnen ausschließlich der Mittelwert und die folgende unvollständige Tabelle bleiben:

j	1	2	3	4	5	6	7	8	9	10
x_j	2.7	0.3	2.1	4.9	1	1.7	4.5	2.8	?	?

a) Nehmen Sie an, dass x_9 und x_{10} unbekannt sind. Wie müssten die Werte gewählt werden, um die kleinstmögliche Stichprobenvarianz zu erhalten? Was ist in diesem Fall die Stichprobenvarianz? (3P)

$x_9 =$ 2.5 $\quad x_{10} =$ 2.5

$s_x^2 =$ 1.976

b) Aus sicherer Quelle erfahren Sie nun, dass $x_9 = 2.3$ gilt. Welchen Wert hat dann x_{10}? (2P)

$x_{10} =$ 2.7

Für die ersten 8 der oben erwähnten Proben wurde noch ein weiteres Merkmal Y ermittelt und es ergaben sich die folgenden Messwerte:

j	1	2	3	4	5	6	7	8
y_j	2.7	2	-0.5	-2	2.2	1.1	3.5	2.8

c) Geben Sie den Median, das obere Quartil und das untere Quartil von $y = (y_1, \dots, y_8)$ an. (3P)

Median = 2.1

oberes Quartil = 2.75 $\quad$ unteres Quartil = 0.3

d) Berechnen Sie den Mittelwert $\overline{y}$ und die Standardabweichung s_y der Stichprobe y sowie den empirischen Korrelationskoeffizienten r_{xy} von $x = (x_1, \dots, x_8)$ und y. Sie können voraussetzen, dass $s_x^2 = 2.54$ gilt. (3P)

$\overline{y} =$ 1.475 $\quad s_y =$ 1.862

$r_{xy} =$ -0.231

Lösungsvorschlag:

a) Es gilt

$$s_x^2 = \frac{1}{9}\sum_{j=1}^{10}(x_j - 2.5)^2.$$

Bemerke, dass $\frac{1}{8}\sum_{j=1}^{8} x_j = 2.5$ gilt. Damit muss $\frac{x_9+x_{10}}{2} = 2.5$ gelten, um zu garantieren, dass der Mittelwert $\overline{x} = 2.5$ ist. Weiter haben x_9 und x_{10} bei festem Mittelwert nur auf die Summanden $(x_9 - 2.5)^2$ bzw. $(x_{10} - 2.5)^2$ Einfluss. Obige Bedingung ist für $x_9 = x_{10} = 2.5$ erfüllt, was auch die Werte sind, welche die Summanden minimieren. Setzt man $x_1, \ldots, 2.5, 2.5$ in die Formel für die Stichprobenvarianz ein, so erhält man $s_x^2 = 1.976$.

b) Es gilt $2.5 = \overline{x} = \left(\frac{1}{10}\sum_{j=1}^{9} x_j + x_{10}\right)$. Umstellen nach x_{10} ergibt $x_{10} = 2.7$.

c) Da $8 * 0.25 \in \mathbb{N}, 8 * 0.5 \in \mathbb{N}, 0.75 * 8 \in \mathbb{N}$ erhalten wir

$$\begin{aligned}
\widetilde{y}_{0.5} &= \frac{1}{2}(y_{(4)} + y_{(5)}) = \frac{1}{2}(2 + 2.2) = 2.1, \\
\widetilde{y}_{0.75} &= \frac{1}{2}(y_{(6)} + y_{(7)}) = \frac{1}{2}(2.7 + 2.8) = 2.75, \\
\widetilde{y}_{0.5} &= \frac{1}{2}(y_{(2)} + y_{(3)}) = \frac{1}{2}(-0.5 + 1.1) = 0.3.
\end{aligned}$$

d) Die Werte ergeben sich aus $\overline{y} = \frac{1}{8}\sum_{j=1}^{8} y_j$, $s_y = \sqrt{\frac{1}{7}\sum_{j=1}^{8}(y_j - \overline{y})^2}$ und $r_{xy} = \frac{\frac{1}{7}\sum_{j=1}^{8}(x_j-\overline{x})(y_j-\overline{y})}{s_x s_y}$.

Aufgabe 2 (11 Punkte)
In einem Online-Shop entdecken Sie ein Angebot: Sie können zu einem Vorzugspreis ein Überraschungspaket (Paket "Schokolade") mit 6 Tafeln Schokolade Ihres Lieblingsherstellers erwerben. (Bei einem Überraschungspaket werden Produkte rein zufällig zusammengestellt. Es ist insbesondere möglich, ein Produkt mehrmals zu erhalten.) Sie wissen, dass insgesamt 7 verschiedene Schokoladensorten im Handel sind, von denen Sie jedoch nur 4 Sorten mögen.

a) Wie groß ist die Wahrscheinlichkeit, dass Sie bei Bestellung eines Pakets "Schokolade" ausschließlich Sorten erhalten, die Sie mögen? (1P)

$$\text{Wahrscheinlichkeit} = \boxed{\left(\tfrac{4}{7}\right)^6}$$

b) Nehmen Sie an, dass Schokoladentafeln, die Sie nicht mögen, für Sie keinen Wert haben, und dass eine Tafel für gewöhnlich 1.50 Euro kostet. Wie viele Tafeln Schokolade sind im Mittel nach Ihrem Geschmack? Wieviel darf ein Paket maximal kosten, so dass es sich im Mittel für Sie lohnt, ein Paket zu bestellen? (3P)

$$\text{Anzahl Tafeln} = \boxed{\tfrac{24}{7}}$$

$$\text{Maximalpreis (in Euro)} = \boxed{\tfrac{36}{7}}$$

Sie entdecken nun ein weiteres Angebot (Überraschungspaket "Pralinen"), in dem eine sehr große, zufällige Anzahl Pralinen enthalten ist. Sie gehen davon aus, dass Sie im Mittel 3 der darin enthaltenen Pralinen nicht mögen. Nehmen Sie an, dass die zufällige Anzahl Y an Pralinen (in einem Überraschungspaket "Pralinen"), die Sie nicht mögen, Poisson-verteilt ist mit dem Parameter λ.

c) Welchen Wert wählen Sie dabei für den Parameter λ der Verteilung von Y? (1P)

$$\lambda = \boxed{3}$$

Sie bestellen nun je ein Überraschungspaket "Schokolade" und "Praline".

d) Wie groß ist die Wahrscheinlichkeit, dass Sie alle in Ihrem Einkauf enthaltenen Produkte mögen? (1P)

$$\text{Wahrscheinlichkeit} = \boxed{\left(\tfrac{4}{7}\right)^6 e^{-3}}$$

e) Wie groß ist die Wahrscheinlichkeit, dass in Ihrem Einkauf genau ein Produkt enthalten ist, das Sie nicht mögen? (1P)

$$\text{Wahrscheinlichkeit} = \boxed{\tfrac{15}{2}\left(\tfrac{4}{7}\right)^6 e^{-3}}$$

f) Wie groß ist die Wahrscheinlichkeit, dass in Ihrem Einkauf mindestens zwei Produkte enthalten sind, die Sie nicht mögen? (2P)

Wahrscheinlichkeit $=$ $\boxed{1 - \frac{17}{2}\left(\frac{4}{7}\right)^6 e^{-3}}$

g) Wie groß ist die Wahrscheinlichkeit, dass, wenn Sie genau ein unbeliebtes Produkt im Einkauf haben, dieses eine Praline ist? (2P)

Wahrscheinlichkeit $=$ $\boxed{\frac{2}{5}}$

Lösungsvorschlag:

a) Es liegt ein Treffer-Niete-Modell vor, bei dem die Trefferwahrscheinlichkeit (= Wahrscheinlichkeit dafür, dass eine Tafel gemocht wird) $p = \frac{4}{7}$ ist. Da das Paket 6 zufällige (also voneinander unabhängige) Tafeln enthält, gilt für die Zufallsvariable X, welche die Anzahl guter Tafeln innerhalb eines Paketes angibt

$$X \sim \text{Bin}(6, \frac{4}{7}).$$

Gesucht ist $\mathbb{P}(X = 6) = \left(\frac{4}{7}\right)^6$.

b) Im Mittel erhält man $\mathbb{E}[X] = 6\frac{4}{7}$ gute Tafeln. Damit ist der Maximalpreis

$$\mathbb{E}[1.5X] = 1.5 \cdot \mathbb{E}[X] = \frac{3}{2} \cdot \frac{24}{7} = \frac{36}{7}.$$

c) Bei der Poisson-Verteilung entspricht der Parameter dem Erwartungswert. In diesem Szenario ist also $\lambda = 3$ die Anzahl der erwarteten Pralinen, die nicht gemocht werden.

d) Sei D das Ereignis, dass alle Tafeln Schokolade und alle Pralinen gemocht werden. Wegen der Unabhängigkeit der Pakete ergibt sich

$$\mathbb{P}(D) = \mathbb{P}(X = 6, Y = 0) = \mathbb{P}(X = 6)\mathbb{P}(Y = 0) = \left(\frac{4}{7}\right)^6 e^{-3} \approx 0.002.$$

e) Für die Wahrscheinlichkeit des beschriebenen Ereignisses E gilt

$$\begin{aligned}\mathbb{P}(E) =& \mathbb{P}(\{X = 5, Y = 0\} \cup \{X = 6, Y = 1\}) \\ =& \mathbb{P}(X = 5)\mathbb{P}(Y = 0) + \mathbb{P}(X = 6)\mathbb{P}(Y = 1) \\ =& \left(\frac{4}{7}\right)^6 e^{-3}3 + 6\left(\frac{4}{7}\right)^5 \frac{3}{7}e^{-3} \\ =& \frac{15}{2}\left(\frac{4}{7}\right)^6 e^{-3} \approx 0.0130.\end{aligned}$$

f) Für das beschriebene Ereignis F gilt $F = \Omega \setminus (D \cup E)$, wobei die Ereignisse D und E disjunkt sind. Demnach ist die gesuchte Wahrscheinlichkeit

$$\mathbb{P}(F) = 1 - \frac{15}{2}\left(\frac{4}{7}\right)^6 e^{-3} - \left(\frac{4}{7}\right)^6 e^{-3} = 1 - \frac{17}{2}\left(\frac{4}{7}\right)^6 e^{-3} \approx 0.985.$$

g) Gesucht ist die Wahrscheinlichkeit (vergleiche mit e))

$$\begin{aligned}
\mathbb{P}(Y=1|E) &= \mathbb{P}(Y=1|\{X=5,Y=0\}\cup\{X=6,Y=1\}) \\
&= \frac{\mathbb{P}(X=6,Y=1)}{\mathbb{P}(E)} \\
&= \frac{\mathbb{P}(X=6)\mathbb{P}(Y=1)}{\mathbb{P}(E)} \\
&= \frac{\left(\frac{4}{7}\right)^6 e^{-3}3}{\frac{15}{2}\left(\frac{4}{7}\right)^6 e^{-3}} = \frac{2}{5}.
\end{aligned}$$

Aufgabe 3 (9 Punkte)

Es sei X eine normalverteilte Zufallsvariable mit Erwartungswert 1 und Varianz 4. Weiter sei Y eine von X unabhängige Zufallsvariable mit $Y \sim \mathcal{N}(0,1)$.

a) Wie ist die Zufallsvariable $Z := X + 2Y$ verteilt? (2P)

$$X + 2Y \sim \boxed{\mathcal{N}(1,8)}$$

b) Berechnen Sie die Kovarianz von X und Z. (1P)

$$C(X,Z) = \boxed{4}$$

c) Berechnen Sie: (2P)

$$\mathbb{P}(X \geq 2) = \boxed{1 - \Phi(\tfrac{1}{2}) = 0.309} \qquad \mathbb{P}(X \leq 0 | Y \leq 2) = \boxed{0.309}$$

d) Bestimmen Sie ein $t \in \mathbb{R}$ so, dass $\mathbb{P}(|X - 1| \leq t) = 0.39$ gilt. Es genügt, wenn Sie das Ergebnis auf 2 Nachkommastellen genau angeben. (2P)

$$t = \boxed{1.02}$$

e) Geben die Wahrscheinlichkeit dafür an, dass das Maximum von X und Y kleiner als 2 ist. (2P)

$$\mathbb{P}(\max\{X,Y\} \leq 2) = \boxed{0.676}$$

Lösungsvorschlag:

a) Wegen der Unabhängigkeit von X und Y gilt mit den Rechenregeln für die Normalverteilung

$$X + 2Y \sim \mathcal{N}(1 + 0, 4 + 2^2 \cdot 1) = \mathcal{N}(1,8).$$

b) Mit der Unabhängigkeit von X und Y gilt mit den Rechenregeln für die Kovarianz

$$C(X,Z) = C(X, X+2Y) = C(X,X) + 2C(X,Y) = C(X,X) = V(X) = 4.$$

c) Es ist

$$\mathbb{P}(X \geq 2) = 1 - \mathbb{P}(X \leq 2) = 1 - \mathbb{P}\left(\frac{X - \mathbb{E}[X]}{\sqrt{V(X)}} \leq \frac{2-1}{\sqrt{4}}\right) = 1 - \Phi(\frac{1}{2}) = 0.309.$$

d) Es ist

$$\mathbb{P}(|X-1| \le t) = \mathbb{P}(-t-1 \le X \le t-1) = \mathbb{P}(-t/2 \le \frac{X - \mathbb{E}[X]}{\sqrt{V(X)}} \le t/2)$$
$$= \Phi(t/2) - \Phi(-t/2) = 2\Phi(t/2) - 1 \overset{!}{=} 0.39.$$

Wir erhalten $0.6950 = \frac{1.39}{2} = \Phi(t/2)$. Damit ist mit der Normalverteilungstabelle $t = 2\Phi^{-1}(0.6950) = 2 \cdot 0.501 = 1.02$.

e) Es ist wegen der Unabhängigkeit und c)

$$\mathbb{P}(\max\{X,Y\} \le 2) = \mathbb{P}(X \le 2, Y \le 2) = \mathbb{P}(X \le 2)\mathbb{P}(Y \le 2) = \Phi(\frac{1}{2})\Phi(2) = 0.676.$$

Aufgabe 4 (10 Punkte)
Es seien X und Y zwei stetige Zufallsvariablen, deren gemeinsame Dichte $f_{X,Y}$ gegeben ist durch

$$f_{X,Y}(x,y) = c \cdot \left(\frac{1}{x^2} + y^2\right) \mathbf{1}\{x \in [1,3], y \in [1,2]\}, \qquad (x,y) \in \mathbb{R}^2,$$

wobei $c \in \mathbb{R}$ eine geeignete Konstante ist.

a) Bestimmen Sie die Konstante c in der obigen Dichte $f_{X,Y}$. (2P)

$$c = \boxed{\tfrac{3}{16}}$$

b) Geben Sie die Wahrscheinlichkeit an, dass $X \geq 1$ gilt. (1P)

$$\mathbb{P}(X \geq 1) = \boxed{1}$$

c) Geben Sie die Dichten f_X und f_Y der Zufallsvariablen X bzw. Y an. (2P)

$$f_X(x) = \boxed{\left(\tfrac{3}{16x^2} + \tfrac{7}{16}\right) \mathbf{1} x \in [1,3]}$$

$$f_Y(y) = \boxed{\left(\tfrac{1}{8} + \tfrac{3}{8}y^2\right) \mathbf{1} y \in [1,2]}$$

d) Berechnen Sie $\mathbb{E}[X]$, $\mathbb{E}[Y]$ und $\mathbb{E}[X+Y]$: (3P)

$$\mathbb{E}[X] = \boxed{\tfrac{3}{16}\log(3) + \tfrac{7}{4}} \qquad \mathbb{E}[Y] = \boxed{\tfrac{51}{32}}$$

$$\mathbb{E}[X+Y] = \boxed{\tfrac{3}{16}\log(3) + \tfrac{107}{32}}$$

e) Woran erkennt man, dass X und Y stochastisch abhängig sind? (2P)

In einer Umgebung von $(x,y) = (2, 1.5)$ ist $f_X(x) f_Y(y) \neq f_{x,y}(x,y)$.

Lösungsvorschlag:

a) Da $f_{X,Y}$ für alle $c \geq 0$ nicht negativ ist, muss ein positives c gefunden werden, sodass

$$\int_1^2 \int_1^3 f_{X,Y}(x,y)dxdy = 1$$

gilt. Es ist

$$\begin{aligned}
\int_1^2 \int_1^3 f_{X,Y}(x,y)dxdy &= c\int_1^2 \int_1^3 \frac{1}{x^2} + y^2 dxdy \\
&= c\int_1^2 \left[-\frac{1}{x} + xy^2\right]_1^3 dy \\
&= c\int_1^2 (-\frac{1}{3} + 3y^2 + 1 - y^2)dy \\
&= c\left[\frac{2}{3}y + \frac{2}{3}y^3\right]_1^2 \\
&= c\left(\frac{4}{3} + \frac{16}{3} - \frac{2}{3} - \frac{2}{3}\right) = \frac{16}{3}c.
\end{aligned}$$

Damit ist $c = \frac{3}{16}$.

b) Da X nur Werte annimmt, die mindestens 1 sind, ist die gesuchte Wahrscheinlichkeit 1.

c) Die Marginaldichten erhält man durch Integration bzgl. einer Komponente. Es ist für $x \in [1,3]$

$$\begin{aligned}
f_X(x) &= \int_1^2 \frac{3}{16}\left(\frac{1}{x^2} + y^2\right) dy = \frac{3}{16}\left[\frac{y}{x^2} + \frac{1}{3}y^3\right]_{y=1}^2 \\
&= \frac{3}{16}\left(\frac{2}{x^2} + \frac{8}{3} - \frac{1}{x^2} - \frac{1}{3}\right) \\
&= \frac{3}{16}\frac{1}{x^2} + \frac{7}{16}.
\end{aligned}$$

Weiter ist für $y \in [1,2]$ (vergleiche mit a)):

$$f_Y(y) = \int_1^3 \frac{3}{16}\left(\frac{1}{x^2} + y^2\right) dx = \frac{3}{16}\left(\frac{2}{3} + 2y^2\right) = \frac{1}{8} + \frac{3}{8}y^2.$$

d) Die Erwartungswerte ergeben sich durch

$$\begin{aligned}
\mathbb{E}[X] &= \int_1^3 f_X(x)xdx = \int_1^3 \frac{3}{16}\frac{1}{x} + \frac{7}{16}xdx \\
&= \left[\frac{3}{16}\log(x) + \frac{7}{32}x^2\right]_1^3 = \frac{3}{16}\log(3) + \frac{63}{32} - \frac{7}{32} \\
&= \frac{3}{16}\log(3) + \frac{7}{4}, \\
\mathbb{E}[Y] &= \int_1^2 \frac{1}{8}y + \frac{3}{8}y^3 dy = \left[\frac{1}{16}y^2 + \frac{3}{32}y^4\right]_1^2 \\
&= \frac{1}{4} + \frac{3}{2} - \frac{1}{16} - \frac{3}{32} = \frac{51}{32}, \\
\mathbb{E}[X+Y] &= \mathbb{E}[X] + \mathbb{E}[Y] = \frac{3}{16}\log(3) + \frac{7}{4} + \frac{51}{32} \\
&= \frac{3}{16}\log(3) + \frac{107}{32}.
\end{aligned}$$

e) Man muss zeigen, dass $f_X(x)f_Y(y)$ keine Dichte von (X,Y) ist. Eine Möglichkeit dafür ist es zu zeigen, dass es eine Menge M mit $\mathbb{P}((X,Y)\in M)\neq 0$ gibt für die

$$f_{X,Y}(x,y)\neq f_X(x)f_Y(y) \quad \text{für alle } (x,y)\in M$$

gilt. Wegen der Stetigkeit von $f_{X,Y}$, f_X und f_Y genügt es, die Ungleichheit in einem Punkt zu zeigen, an dem $f_{X,Y}$ strikt positiv ist. Betrachtet man $(x,y)=(2,1.5)$ so gilt

$$f_{X,Y}(2,1.5)=\frac{3}{16}\left(\frac{1}{4}+\frac{9}{4}\right)=\frac{15}{32}=\frac{960}{2048}.$$

Weiter ist

$$f_X(2)f_Y(1.5)=\left(\frac{3}{16\cdot 4}+\frac{7}{16}\right)\left(\frac{1}{8}+\frac{3}{8}\frac{9}{4}\right)=\frac{31^2}{64\cdot 32}=\frac{961}{2048}.$$

Alternativ kann man zeigen, dass $C(X,Y)\neq 0$ gilt.

Aufgabe 5 (9 Punkte)
Es soll der unbekannte Parameter $\vartheta > 0$ einer stetigen Verteilung mit der Dichte

$$f_\vartheta(t) = \begin{cases} \frac{2}{\vartheta} \cdot \exp\left(-\frac{2t}{\vartheta}\right), & t > 0, \\ 0, & t \leq 0 \end{cases}$$

bestimmt werden.

a) Geben Sie die zur Stichprobe $x = (x_1, \dots, x_n)$ gehörende Likelihood-Funktion $L_x(\vartheta)$ an und berechnen Sie die Loglikelihood-Funktion $M_x(\vartheta)$. Dabei dürfen Sie voraussetzen, dass $x_i > 0$ für $i = 1, \dots, n$ gilt. (2P)

$$L_x(\vartheta) = \boxed{\left(\tfrac{2}{\vartheta}\right)^n \exp\left(-\tfrac{2}{\vartheta}\textstyle\sum_{j=1}^n x_j\right)}$$

$$M_x(\vartheta) = \boxed{-\tfrac{2}{\vartheta}\left(\textstyle\sum_{j=1}^n x_j\right) + n\log\left(\tfrac{2}{\vartheta}\right)}$$

b) Bestimmen Sie den Maximum-Likelihood-Schätzer $\hat{\vartheta}(x)$ für ϑ. Gehen Sie folgendermaßen vor:

b1) Geben Sie zunächst die Ableitung von $M_x(\vartheta)$ an. (1P)

$$M_x'(\vartheta) = \boxed{\tfrac{2}{\vartheta^2}\left(\textstyle\sum_{j=1}^n x_j\right) - \tfrac{n}{\vartheta}}$$

b2) Geben Sie nun den Maximum-Likelihood Schätzer für ϑ an. (1P)

$$\hat{\vartheta}(x) = \boxed{\tfrac{2}{n}\textstyle\sum_{j=1}^n x_j}$$

b3) Begründen Sie, dass der von Ihnen unter b2) angegebene Schätzer tatsächlich der ML-Schätzer für ϑ ist. (3P)

$\hat{\vartheta}$ ist eine Nullstelle von M_x' und es gilt $M_x''(\hat{\vartheta}) < 0$.

c) Geben Sie den Momentenschätzer für ϑ an. (2P)

Momentenschätzer : $\boxed{\tfrac{2}{n}\textstyle\sum_{j=1}^n x_j}$

Lösungsvorschlag:

a) Die Likelihoodfunktion ist

$$L_x(\vartheta) = \prod_{j=1}^{n} f_\vartheta(x_j) = \prod_{j=1}^{n} \frac{2}{\vartheta} \cdot \exp\left(-\frac{2x_j}{\vartheta}\right) = \left(\frac{2}{\vartheta}\right)^n \exp\left(-\frac{2}{\vartheta}\sum_{j=1}^{n} x_j\right),$$

weshalb die Loglikelihood-Funktion die Darstellung

$$M_x(\vartheta) = \log\left(\left(\frac{2}{\vartheta}\right)^n \exp\left(-\frac{2}{\vartheta}\sum_{j=1}^{n} x_j\right)\right) = -\frac{2}{\vartheta}\left(\sum_{j=1}^{n} x_j\right) + n\log(\frac{2}{\vartheta})$$

hat.

b) b1) Die Ableitung der in a) berechneten Funktion ist

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

b2)+b3) Um den Maximum-Likelihood-Schätzer zu bestimmen, müssen die Nullstellen der Funktion in b1) ermittelt werden. Es ist

$$\begin{aligned} M_x'(\vartheta) = 0 &\Longleftrightarrow \frac{2}{\vartheta}\left(\sum_{j=1}^{n} x_j\right) - n = 0 \\ &\Longleftrightarrow \vartheta = \frac{2}{n}\sum_{j=1}^{n} x_j. \end{aligned}$$

Um zu garantieren, dass der kritische Punkt eine Maximalstelle von M_x ist, bestimmt man die zweite Ableitung und verifiziert, dass diese im kritischen Punkt strikt negativ ist:

$$M_x''(\vartheta) = -\frac{4}{\vartheta^3}\left(\sum_{j=1}^{n} x_j\right) + \frac{n}{\vartheta^2}.$$

Einsetzen von $\hat{\vartheta} = \frac{2}{n}\sum_{j=1}^{n} x_j$ ergibt

$$\begin{aligned} M_x''(\hat{\vartheta}) &= -\frac{4}{\hat{\vartheta}^3}\left(\sum_{j=1}^{n} x_j\right) + \frac{n}{\hat{\vartheta}^2} \\ &= -\frac{4}{\left(\frac{2}{n}\sum_{j=1}^{n} x_j\right)^3}\left(\sum_{j=1}^{n} x_j\right) + \frac{n}{\left(\frac{2}{n}\sum_{j=1}^{n} x_j\right)^2} \\ &= \frac{-2n^4+1}{4n\left(\sum_{j=1}^{n} x_j\right)^2}, \end{aligned}$$

was wegen $n \geq 1$ strikt kleiner als 0 ist. Damit ist $\hat{\vartheta}$ der Maximum-Likelihood-Schätzer.

c) Da $\mathbb{E}[X] = \vartheta/2$, entspricht der MLE dem Momentenschätzer. Die Eigenschaft $\mathbb{E}[X] = \vartheta/2$ erhält man da f_ϑ die Dichte einer Exponentialverteilung mit Parameter $\frac{\vartheta}{2}$ ist. Erkennt man dies nicht, so kann man alternativ

$$\int_0^\infty x f_\vartheta(x)dx = \vartheta/2$$

berechnen.