Karlsruher Institut für Technologie

Institut für Stochastik

Prof. Dr. Norbert Henze · Dipl.-Math. oec. Silke Weber

Name:

Vorname:

Matrikelnummer:

Lösungsvorschlag zur Klausur der Vorlesung

Grundlagen der Wahrscheinlichkeitstheorie und Statistik für Studierende der Informatik

Datum: 06. Februar 2016

Dauer: 90 Minuten

ACHTUNG:
Bei dieser Klausur werden **nur** diejenigen Ergebnisse gewertet, die in die vorgesehenen Kästchen eingetragen sind! Eine Begründung bzw. Herleitung der Ergebnisse wird nicht verlangt, es sei denn, dies wird explizit gefordert!
Vereinfachen Sie die Ergebnisse soweit wie möglich. Geben Sie die Ergebnisse so exakt wie möglich an (z.B. als Bruch), runden Sie ggf. auf 3 Nachkommastellen.

Als **Hilfsmittel** sind zugelassen:
Skriptum zur Vorlesung, Vorlesungsmitschrift, Taschenrechner (nicht vernetzbar), Wörterbuch

Hinreichend zum Bestehen der Klausur sind 20 Punkte.

VIEL ERFOLG!

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

p-Quantile $\chi^2_{k;p}$ der χ^2 – Verteilung mit k Freiheitsgraden

	p					
k	0.900	0.950	0.975	0.990	0.995	0.999
1	2.71	3.84	5.02	6.63	7.88	10.83
2	4.61	5.99	7.38	9.21	10.60	13.82
3	6.25	7.81	9.35	11.34	12.84	16.27
4	7.78	9.49	11.14	13.28	14.86	18.47
5	9.24	11.07	12.83	15.09	16.75	20.51
6	10.64	12.59	14.45	16.81	18.55	22.46
7	12.02	14.07	16.01	18.48	20.28	24.32
8	13.36	15.51	17.53	20.09	21.95	26.12
9	14.68	16.92	19.02	21.67	23.59	27.88
10	15.99	18.31	20.48	23.21	25.19	29.59
11	17.28	19.68	21.92	24.73	26.76	31.26
12	18.55	21.03	23.34	26.22	28.30	32.91
13	19.81	22.36	24.74	27.69	29.82	34.53
14	21.06	23.68	26.12	29.14	31.32	36.12
15	22.31	25.00	27.49	30.58	32.80	37.70
16	23.54	26.30	28.85	32.00	34.27	39.25
17	24.77	27.59	30.19	33.41	35.72	40.79
18	25.99	28.87	31.53	34.81	37.16	42.31
19	27.20	30.14	32.85	36.19	38.58	43.82
20	28.41	31.41	34.17	37.57	40.00	45.31
21	29.62	32.67	35.48	38.93	41.40	46.80
22	30.81	33.92	36.78	40.29	42.80	48.27
23	32.01	35.17	38.08	41.64	44.18	49.73
24	33.20	36.41	39.36	42.98	45.56	51.18
25	34.38	37.65	40.65	44.31	46.93	52.62
26	35.56	38.89	41.92	45.64	48.29	54.05
27	36.74	40.11	43.19	46.96	49.64	55.48
28	37.92	41.34	44.46	48.28	50.99	56.89
29	39.09	42.56	45.72	49.59	52.34	58.30
30	40.26	43.77	46.98	50.89	53.67	59.70
40	51.81	55.76	59.34	63.69	66.77	73.40
50	63.17	67.50	71.42	76.15	79.49	86.66
60	74.40	79.08	83.30	88.38	91.95	99.61
80	96.58	101.88	106.63	112.33	116.32	124.84
100	118.50	124.34	129.56	135.81	140.17	149.45

Aufgabe 1 (11.5 Punkte)

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

j	1	2	3	4	5	6	7	8	9	10
x_j	5	3.8	0.3	8.6	7.2	0.5	6	6.2	7.4	2
y_j	0.8	2	3.3	3.5	1.7	8.3	9.9	10.5	8	5.9

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

$$\sum_{j=1}^{10} x_j = 47, \quad \sum_{j=1}^{10} y_j = 53.9, \quad \sum_{j=1}^{10} x_j^2 = 298.78, \quad \sum_{j=1}^{10} y_j^2 = 406.63, \quad \sum_{j=1}^{10} x_j y_j = 254.58.$$

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

c) Bestimmen Sie das 0.33-getrimmte Stichprobenmittel von $(x_1, \ldots, x_{10})$.

d) Bestimmen Sie den Median, das untere und obere Quartil sowie den Quartilsabstand von $(x_1, \ldots, x_{10})$.

e) Welche Kenngrößen ändern sich nicht, wenn x_9 durch 74 ersetzt wird?

Lösungsvorschlag:

a) Es ist (mit den Hinweisen und $n = 10$)

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{10}\sum_{i=1}^{10} x_i = \frac{1}{10} \cdot 47 = 4.7, \qquad 0.5P$$

$$\overline{y} = \frac{1}{n}\sum_{i=1}^{n} y_i = \frac{1}{10}\sum_{i=1}^{10} y_i = \frac{1}{10} \cdot 53.9 = 5.39, \qquad 0.5P$$

$$s_x = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2} = \sqrt{\frac{1}{n-1}\left(\sum_{i=1}^{n} x_i^2 - n\overline{x}^2\right)} = \sqrt{\frac{1}{9}(298.78 - 10 \cdot 4.7^2)} \approx 2.942, \qquad 1P$$

$$s_y = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(y_i - \overline{y})^2} = \sqrt{\frac{1}{n-1}\left(\sum_{i=1}^{n} y_i^2 - n\overline{y}^2\right)} = \sqrt{\frac{1}{9}(406.63 - 10 \cdot 5.39^2)} \approx 3.592, \qquad 1P$$

$$r_{xy} = \frac{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})(y - \overline{y}))}{s_x s_y} = \frac{\frac{1}{n-1}\left(\sum_{i=1}^{n} x_i y_i - n\overline{x}\overline{y}\right)}{s_x s_y} = \frac{\frac{1}{9}(254.58 - 10 \cdot 4.7 \cdot 5.39)}{2.942 \cdot 3.592} \approx 0.013. \qquad 1P$$

b) Es gilt

$$b^* = r_{xy}\frac{s_y}{s_x} = 0.013 \cdot \frac{3.592}{2.942} \approx 0.016, \qquad 1P$$
$$a^* = \overline{y} - b^*\overline{x} = 5.39 - 0.016 \cdot 4.7 \approx 5.315. \qquad 1P$$

Dann erhält man

$$y = 5.315 + 0.016x.$$

c) Die geordnete Stichprobe ist $x_{()} = (0.3, 0.5, 2, 3.8, 5, 6, 6.2, 7.2, 7.4, 8.6)$ und $k = \lfloor n\alpha \rfloor = \lfloor 10 \cdot 0.33 \rfloor = \lfloor 3.3 \rfloor = 3$. Damit ist das 0.33-getrimmte Mittel

$$\overline{x}_{0.33} = \frac{1}{4}(3.8 + 5 + 6 + 6.2) = 5.25 \qquad \color{red}{1P}.$$

d) Der Median ist das 0.5-Quantil. Da $k = \lfloor n\alpha \rfloor = \lfloor 5 \rfloor = 5$ und $n\alpha \in \mathbb{N}$ gilt

$$\tilde{x}_{0.5} = \frac{1}{2}(x_{(5)} + x_{(6)}) = 5.5. \qquad \color{red}{1P}$$

Da $0.25 \cdot 10 \notin \mathbb{N}$ und $0.75 \cdot 10 \notin \mathbb{N}$, berechnen sich das obere und untere Quartil zu

$$\tilde{x}_{0.25} = x_{(3)} = 2, \quad \tilde{x}_{0.75} = x_{(8)} = 7.2, \qquad \color{red}{1P}$$

und der Quartilsabstand ist

$$\tilde{x}_{0.75} - \tilde{x}_{0.25} = 7.2 - 2 = 5.2. \qquad \color{red}{0.5P}$$

e)
- 0.33-getrimmtes Stichprobenmittel
- Median
- unteres/oberes Quartil 2 P
- Quartilsabstand

Summe: 11.5 P

Aufgabe 2 (9 Punkte)

Es seien X eine Zufallsvariable mit Werten in $\{0,1\}$ und Y eine Zufallsvariable, die die Werte $-1,0$ und 1 annehmen kann. Die folgende Tabelle gibt die gemeinsame Verteilung $\mathbb{P}(X=i, Y=j)$ des Zufallsvektors (X,Y) für die Werte $i=0,1$ und $j=-1,0,1$ an.

	$j=-1$	$j=0$	$j=1$
$i=0$	$\frac{1}{8}$	$\frac{1}{4}$	0
$i=1$	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{1}{4}$

a) Bestimmen Sie die Marginalverteilung von Y (d.h. $\mathbb{P}(Y=j)$ für $j\in\{-1,0,1\}$) sowie den Erwartungswert $\mathbb{E}(Y)$.

b) Wie groß ist die bedingte Wahrscheinlichkeit $\mathbb{P}(Y=-1|X=1)$?

c) Berechnen Sie die Wahrscheinlichkeit $\mathbb{P}(X=1, Y>-1)$.

d) Sei $T := X\cdot Y$. Berechnen Sie $\mathbb{E}(T)$, $\mathbb{E}(T^2)$ sowie $\mathbb{V}(T)$.

e) Sind X und Y unabhängig? Begründen Sie Ihre Antwort.

Lösungsvorschlag:

a) Es ist $\mathbb{P}(Y=-1) = 1/4 = \mathbb{P}(Y=1)$ sowie $\mathbb{P}(Y=0)=1/2$. Hieraus folgt

$$\mathbb{E}(Y) \;=\; (-1)\cdot\frac{1}{4} + 0\cdot\frac{1}{2} + 1\cdot\frac{1}{4} \;=\; 0.$$

1.5+1 P = 2.5 P

b) Es ist

$$\mathbb{P}(Y=-1|X=1) = \frac{\mathbb{P}(X=1, Y=-1)}{\mathbb{P}(X=1)} = \frac{1/8}{5/8} = \frac{1}{5}.$$

1 P

c) Es gilt $\mathbb{P}(X=1, Y>-1) = \mathbb{P}(X=1, Y=0) + \mathbb{P}(X=1, Y=1) = 1/4 + 1/4 = 1/2$.
1 P

d) Das Produkt $T = XY$ nimmt die Werte $-1,0,1$ an, und zwar mit den Wahrscheinlichkeiten

$$\begin{aligned}\mathbb{P}(T=-1) &= \mathbb{P}(X=1, Y=-1) = \frac{1}{8},\\ \mathbb{P}(T=1) &= \mathbb{P}(X=1, Y=1) = \frac{1}{4}, \quad \mathbb{P}(T=0) \;=\; \frac{5}{8}.\end{aligned}$$

Es folgt

$$\begin{aligned}
\mathbb{E}(T) &= (-1)\cdot\frac{1}{8}+1\cdot\frac{1}{4}=\frac{1}{8},\\
\mathbb{E}(T^2) &= (-1)^2\cdot\frac{1}{8}+1^2\cdot\frac{1}{4}=\frac{3}{8},\\
\mathbb{V}(T) &= \mathbb{E}(T^2)-(\mathbb{E}(T))^2=\frac{3}{8}-\left(\frac{1}{8}\right)^2=\frac{23}{64}.
\end{aligned}$$

3 P

e) X und Y sind nicht stochastisch unabhängig, da

$$\mathbb{P}(X=0,Y=1)=0\neq\mathbb{P}(X=0)\cdot\mathbb{P}(Y=1)=\frac{3}{8}\cdot\frac{1}{4}.$$

Eine alternative Begründung verwendet, dass im Fall der Unabhängigkeit von X und Y die Gleichheit $\mathbb{E}(XY)=\mathbb{E}(X)\cdot\mathbb{E}(Y)$ gelten müsste, was wegen $\mathbb{E}(XY)=1/2$ und $\mathbb{E}(Y)=0$ nicht der Fall ist. 1.5 P

Summe: 9 P

Aufgabe 3 (8 Punkte)

Gegeben seien zwei Urnen U_1 und U_2. Urne U_1 enthalte drei rote und drei schwarze Kugeln, Urne U_2 zwei rote und vier schwarze Kugeln. Es wird eine der Urnen rein zufällig ausgewählt. Aus *dieser* Urne werden dann rein zufällig zwei Kugeln *mit einem Griff* gezogen. Die Zufallsvariable X bezeichne die Anzahl der gezogenen roten Kugeln. Sei A_j das Ereignis, dass aus Urne U_j gezogen wurde ($j = 1, 2$).

a) Welche (bekannte) Verteilung besitzt X unter der Bedingung A_1?

b) Bestimmen Sie $\mathbb{P}(X = k|A_j)$ für $k \in \{0, 1, 2\}$ und $j \in \{1, 2\}$.

c) Bestimmen Sie $\mathbb{P}(X = k)$ für $k = 0, 1, 2$.

d) Bestimmen Sie die bedingte Wahrscheinlichkeit $\mathbb{P}(A_1|X = 0)$.

Lösungsvorschlag:

a) Das Ziehen zweier Kugeln mit einem Griff ist äquivalent dazu, zwei Kugeln rein zufällig ohne Zurücklegen zu ziehen. Unter der Bedingung A_1 wird aus einer Urne mit drei roten und drei schwarzen Kugeln gezogen. Die bedingte Verteilung von X unter der Bedingung A_1 ist also die hypergeometrische Verteilung Hyp(2,3,3). 1 P

b) Es gelten

$$\mathbb{P}(X = k|A_1) = \frac{\binom{3}{k}\binom{3}{2-k}}{\binom{6}{2}}, \qquad \mathbb{P}(X = k|A_2) = \frac{\binom{2}{k}\binom{4}{2-k}}{\binom{6}{2}},$$

$k = 0, 1, 2$. Wegen $\binom{6}{2} = 15$ ergeben sich die Zahlenwerte

$$\begin{aligned}
\mathbb{P}(X = 0|A_1) &= \frac{1 \cdot 3}{15} = \frac{1}{5}, \ \mathbb{P}(X = 1|A_1) = \frac{3 \cdot 3}{15} = \frac{3}{5}, \ \mathbb{P}(X = 2|A_1) = \frac{3 \cdot 1}{15} = \frac{1}{5}, \\
\mathbb{P}(X = 0|A_2) &= \frac{1 \cdot 6}{15} = \frac{2}{5}, \ \mathbb{P}(X = 1|A_2) = \frac{2 \cdot 4}{15} = \frac{8}{15}, \ \mathbb{P}(X = 2|A_2) = \frac{1 \cdot 1}{15} = \frac{1}{15}.
\end{aligned}$$

3 P

c) Nach der Formel von der totalen Wahrscheinlichkeit gilt

$$\mathbb{P}(X = k) \;=\; \mathbb{P}(A_1) \cdot \mathbb{P}(X = k|A_1) + \mathbb{P}(A_2) \cdot \mathbb{P}(X = k|A_2).$$

Wegen $\mathbb{P}(A_1) = \mathbb{P}(A_2) = 1/2$ folgt mithilfe von b)

$$\begin{aligned}
\mathbb{P}(X = 0) &= \frac{1}{2} \cdot \left(\frac{3+6}{15}\right) = \frac{9}{30} = \frac{3}{10}, \ \mathbb{P}(X = 1) = \frac{1}{2} \cdot \left(\frac{9+8}{15}\right) = \frac{17}{30}, \\
\mathbb{P}(X = 2) &= \frac{1}{2} \cdot \left(\frac{3+1}{15}\right) = \frac{4}{30} = \frac{2}{15}.
\end{aligned}$$

3 P

d) Nach der Bayes-Formel gilt

$$\begin{aligned}\mathbb{P}(A_1|X=0) &= \frac{\mathbb{P}(A_1)\cdot\mathbb{P}(X=0|A_1)}{\mathbb{P}(A_1)\cdot\mathbb{P}(X=0|A_1)+\mathbb{P}(A_2)\cdot\mathbb{P}(X=0|A_2)}\\ &= \frac{\frac{1}{2}\cdot\frac{3}{15}}{\frac{1}{2}\cdot\frac{3}{15}+\frac{1}{2}\cdot\frac{6}{15}} = \frac{1}{3}.\end{aligned}$$

1 P

Summe: 8 P

Aufgabe 4 (9.5 Punkte)

Die Zufallsvariable X habe die Dichte

$$f_\vartheta(t) \;=\; \frac{1}{2\vartheta}\exp\left(-\frac{|t|}{\vartheta}\right), \qquad t \in \mathbb{R}.$$

Dabei ist $\vartheta \in \Theta := (0,\infty)$ ein unbekannter Parameter. Der Parameter ϑ soll aufgrund einer Stichprobe $x = (x_1,\ldots,x_n)$ geschätzt werden, wobei $x_1,\ldots,x_n$ Realisierungen von unabhängigen Zufallsvariablen $X_1,\ldots,X_n$ mit gleicher Verteilung wie X sind.

a) Wie lautet die Likelihood-Funktion für ϑ zur Stichprobe x?

b) Wie lautet die Loglikelihood-Funktion für ϑ zur Stichprobe x?

c) Geben Sie die Ableitung der Loglikelihood-Funktion für ϑ zur Stichprobe x an.

d) Geben Sie den Maximum-Likelihood-Schätzer für ϑ an.

e) Ist $\widehat{\vartheta}_n$ ein erwartungstreuer Schätzer für ϑ? Begründen Sie Ihre Antwort.
Hinweis auch für f): Verwenden Sie die Symmetrie der Dichte f_ϑ um 0.

f) Wie lautet die Varianz $\mathbb{V}_\vartheta(\widehat{\vartheta}_n)$ von $\widehat{\vartheta}_n$?

g) Ist die Schätzfolge $(\widehat{\vartheta}_n)_{n\geq 1}$ konsistent für ϑ? Begründen Sie Ihre Antwort.

Lösungsvorschlag:

a)

$$L_x(\vartheta) = \left(\frac{1}{2\vartheta}\right)^n \exp\left(-\frac{1}{\vartheta}\sum_{j=1}^n |x_j|\right). \qquad 2P$$

b)

$$M_x(\vartheta) = \log L_x(\vartheta) = -n\log 2 - n\log\vartheta - \frac{1}{\vartheta}\sum_{j=1}^n |x_j|. \qquad 0.5P$$

c)

$$M_x'(\vartheta) = -\frac{n}{\vartheta} + \left(\frac{1}{\vartheta}\right)^2 \sum_{j=1}^n |x_j| \qquad 0.5P$$

d) Die Gleichung

$$M_x'(\vartheta) = -\frac{n}{\vartheta} + \left(\frac{1}{\vartheta}\right)^2 \sum_{j=1}^n |x_j| \stackrel{!}{=} 0 \qquad (1)$$

ist gleichbedeutend mit $\vartheta = \frac{1}{n}\sum_{j=1}^n |x_j|$. Wegen

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

wechselt M_x' an der Stelle $\frac{1}{n}\sum_{j=1}^n |x_j|$ das Vorzeichen von Plus nach Minus, und so liegt dort in der Tat ein lokales Maximum vor, was wegen der Eindeutigkeit der Lösung von (1) auch ein globales Maximum ist.

oder:

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

$$M_x''\left(\frac{1}{n}\sum_{j=1}^{n}|x_j|\right) = \frac{1}{\vartheta^2}(1-2n) < 0, \text{ da } n \geq 1.$$

Somit ist

$$\widehat{\vartheta}_n = \frac{1}{n}\sum_{j=1}^{n}|X_j|$$

der Maximum-Likelihood-Schätzer für ϑ. 2 P

e) Wegen der Linearität der Erwartungswertbildung und der identischen Verteilung von $X, X_1, \ldots, X_n$ gilt

$$\mathbb{E}_\vartheta(\widehat{\vartheta}_n) = \frac{1}{n}\sum_{j=1}^{n}\mathbb{E}(|X_j|) = \mathbb{E}_\vartheta(|X|).$$

Weiter ist wegen der Symmetrie der Dichte f_ϑ um 0 und $|t| = t$ für $t \geq 0$ sowie der Substitution $u = t/\vartheta$

$$\mathbb{E}_\vartheta(|X|) = \int_{-\infty}^{\infty}|t|f_\vartheta(t)\,\mathrm{d}t = \frac{1}{\vartheta}\int_0^\infty t\mathrm{e}^{-t/\vartheta}\,\mathrm{d}t = \vartheta\int_0^\infty u\mathrm{e}^{-u}\,\mathrm{d}u = \vartheta.$$

Folglich ist der Schätzer $\widehat{\vartheta}_n$ erwartungstreu für ϑ. 1.5 P

f) Wegen $\mathbb{V}(aY) = a^2\mathbb{V}(Y)$, der Additivität der Varianz bei unabhängigen Zufallsvariablen und der identischen Verteilung von $X, X_1, \ldots, X_n$ gilt zunächst

$$\mathbb{V}_\vartheta(\widehat{\vartheta}_n) = \frac{1}{n^2}\sum_{j=1}^{n}\mathbb{V}_\vartheta(|X_j|) = \frac{\mathbb{V}_\vartheta(|X|)}{n}.$$

Weiter ist

$$\mathbb{V}_\vartheta(|X|) = \mathbb{E}_\vartheta(|X|^2) - (\mathbb{E}_\vartheta(|X|))^2 = \mathbb{E}_\vartheta(X^2) - \vartheta^2.$$

Mit der Symmetrie der Dichte f_ϑ um 0 und der Substitution $u = t/\vartheta$ folgt

$$\mathbb{E}_\vartheta(X^2) = \frac{1}{\vartheta}\int_0^\infty t^2\mathrm{e}^{-t/\vartheta}\,\mathrm{d}t = \vartheta^2\int_0^\infty u^2\mathrm{e}^{-u}\,\mathrm{d}u = 2\vartheta^2$$

und damit $\mathbb{V}_\vartheta(\widehat{\vartheta}_n) = \vartheta^2/n$. 1.5 P

f) Nach der Tschebyschew-Ungleichung gilt für jedes $\varepsilon > 0$ und jedes ϑ

$$\mathbb{P}_\vartheta\left(|\widehat{\vartheta}_n - \vartheta| \geq \varepsilon\right) \leq \frac{\mathbb{V}_\vartheta(\widehat{\vartheta}_n)}{\varepsilon^2} = \frac{\vartheta^2}{n\varepsilon^2} \to 0 \text{ für } n \to \infty.$$

Also ist die Schätzfolge $(\widehat{\vartheta}_n)_{n\geq 1}$ konsistent für ϑ.

oder: Aus

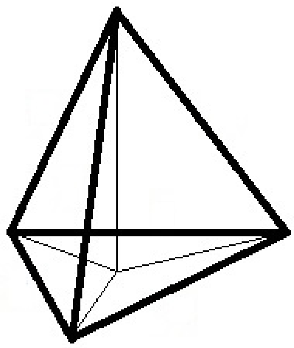
$$\mathbb{V}_\vartheta(\widehat{\vartheta}_n) = \frac{\vartheta^2}{n} \xrightarrow{n\to\infty} 0$$

und Erwartungstreue von $\widehat{\vartheta}_n$ folgt Konsistenz von $\widehat{\vartheta}_n$ für ϑ. 1.5 P

Summe: 9.5 P

Aufgabe 5 (12 Punkte)

Eine Gruppe von Studenten verabredet sich, um ein Gesellschaftsspiel zu spielen, bei dem ein Tetraeder („Vier-Flächner“) verwendet wird (siehe Abbildung).



a) Fabian, einer der Studenten, glaubt nicht an die Echtheit des Tetraeders (d.h. jede Seite liegt mit gleicher Wahrscheinlichkeit unten) und führt deshalb folgendes Experiment durch. Er nummeriert die Seiten des Tetraeders von 1 bis 4 und wirft es 200 mal in unabhängiger Folge. Dann notiert er sich die Häufigkeit, mit der die Zahlen 1 bis 4 jeweils unten liegen. Es ergeben sich die Häufigkeiten 46, 65, 54, 35 für die Zahlen $1, \dots, 4$. Nun möchte Fabian einen Test durchführen, der seinen Zweifel an der Echtheit des Tetraeders statistisch nachweist.

   i) Wie wählt Fabian die Hypothese bzw. Alternative? Welchen aus der Vorlesung bekannten Test muss Fabian anwenden?

   ii) Geben Sie den kritischen Wert an, wenn Fabian einen Test zum Niveau 5% durchführen möchte?

   iii) Wie lautet das Testverfahren?

   iv) Berechnen Sie die Teststatistik aus den Daten.
   Kann Fabian die Hypothese verwerfen, wenn er eine Wahrscheinlichkeit für den Fehler 1. Art von 5% toleriert?

   v) Ändert sich Fabians Entscheidung, wenn er einen Test zum niedrigeren Niveau 1% durchführt?

b) Gehen Sie im Folgenden von der Echtheit des Tetraeders aus. Dieses wird n mal in unabhängiger Folge geworfen.

   Sei X_i die Zahl, die nach dem i-ten Wurf unten liegt und

$$Y_n := \sum_{i=1}^{n} X_i \tag{2}$$

   die Zahlen-Summe nach n Würfen.

   i) Berechnen Sie den Erwartungswert und die Varianz von Y_n.

   ii) Das Tetraeder wird nun 1000 mal geworfen. Approximieren Sie mit Hilfe des Zentralen Grenzwertsatzes die Wahrscheinlichkeit

$$\mathbb{P}(2400 \leq Y_{1000} \leq 2600). \tag{3}$$

Lösungsvorschlag:

a) i) Wir sind hier in der Situation eines Chi-Quadrat-Anpassungstests, der aus der Vorlesung bekannt ist. Da Fabian die Behauptung $\mathbb{P}(X_i = k) \neq \frac{1}{4}$ für mindestens ein $k \in \{1, \ldots, 4\}$ statistisch nachweisen möchte, ist die Hypothese bzw. Alternative

$$H_0 : p_j = \frac{1}{4}, \quad \forall j = 1, \ldots, 4 \qquad\qquad H_1 : p_j \neq \frac{1}{4} \quad \text{für mind. ein } j \in \{1, \ldots, 4\}$$

2 P

ii) Der kritische Wert ist $\chi^2_{s-1;1-\alpha}$, das $(1-\alpha)-$Quantil der χ^2-Verteilung mit $s-1$ Freiheitsgraden.

Mit $s = 4$ und $\alpha = 0.05$ ergibt sich als kritischer Wert $\chi^2_{3;0.95} = 7.81$. 1 P

iii) Der Test lautet

$$\begin{aligned} H_0 \text{ wird verworfen, } & \text{ falls } \chi^2_{200}(k_1, \ldots, k_4) \geq \chi^2_{3;0.95} = 7.81 \\ \text{kein Widerspruch zu } H_0, & \text{ falls } \chi^2_{200}(k_1, \ldots, k_4) < \chi^2_{3;0.95} = 7.81. \end{aligned}$$

1 P

iv) Die Testgröße ist

$$\chi^2_n(k_1, \ldots, k_s) = \sum_{i=1}^{s} \frac{(k_i - n \cdot \pi_i)^2}{n \cdot \pi_i}.$$

Mit $n = 200$, $s = 4$, $(k_1, \ldots, k_4) = (46, 65, 54, 35)$ und $\pi_i = \frac{1}{4}$ für $1 \leq i \leq 4$ ergibt sich

$$\chi^2_{200}(46, 65, 54, 35) = \sum_{i=1}^{4} \frac{(k_i - 200 \cdot \frac{1}{4})^2}{200 \cdot \frac{1}{4}} = 9.64.$$

Nun gilt

$$\chi^2_{3;0.95} = 7.81 < 9.64 = \chi^2_{200}(46, 65, 54, 35).$$

Damit wird H_0 abgelehnt. 2 P

v) Es gilt nun

$$\chi^2_{3;0.99} = 11.34 > 9.64 = \chi^2_{200}(46, 65, 54, 35).$$

Also wird H_0 nicht abgelehnt. 2 P

b) Da wir nun annehmen, dass das Tetraeder regelmäßig ist, gilt

$$\mathbb{P}(X_i = k) = \frac{1}{4}, \quad \text{für } k = 1, \ldots, 4 \text{ und } i \in \mathbb{N}.$$

i) Es ist

$$\begin{aligned} \mathbb{E}[Y_n] = \mathbb{E}\left[\sum_{i=1}^{n} X_i\right] &= \sum_{i=1}^{n} \mathbb{E}[X_i] = n \cdot \mathbb{E}[X_1] \\ &= n \cdot \frac{1}{4} \cdot (1 + 2 + 3 + 4) = \frac{5}{2} n \qquad 1P \end{aligned}$$

und da die X_i unabhängig sind

$$
\begin{aligned}
\mathbb{V}(Y_n) = \mathbb{V}\left(\sum_{i=1}^{n} X_i\right) &= \sum_{i=1}^{n} \mathbb{V}[X_i] = n \cdot \mathbb{V}(X_1) \\
&= n \cdot (\mathbb{E}(X_1^2) - \mathbb{E}[X_1]^2) = n \cdot \left(\frac{1}{4}(1+4+9+16) - \left(\frac{5}{2}\right)^2\right) \\
&= n\left(\frac{30}{4} - \frac{25}{4}\right) = \frac{5}{4}n. \qquad \color{red}{1P}
\end{aligned}
$$

ii) Es gilt
$\mathbb{E}(Y_{1000}) = 2500$ und $\mathbb{V}(Y_{1000}) = 1250$. Also

$$
\begin{aligned}
\mathbb{P}(2400 \leq Y_{1000} \leq 2600) &= \mathbb{P}\left(\frac{2400-2500}{\sqrt{1250}} \leq \frac{Y_{1000}-2500}{\sqrt{1250}} \leq \frac{2600-2500}{\sqrt{1250}}\right) \\
&\approx \Phi\left(2.83\right) - \Phi\left(-2.83\right) \\
&= \Phi(2.83) - (1 - \Phi(2.83)) = 2\Phi(2.83) - 1 \\
&= 2 \cdot 0.9977 - 1 = 0.9954 \qquad \color{red}{2P}
\end{aligned}
$$

Summe: 12

Punkte insgesamt:
Informatiker: 50 P