TU Berlin - Institut für Mathematik
Sommersemester 2024
Dozent: Dr. Nikolas Tapia
Assistentin: M.Sc. Claudia Drygala

Stochastik für Informatik(er) – Übung 8

Abgabe bis Freitag, den 21.06.2024 um 23:59

Hinweise zur Bearbeitung des Übungsblattes:

- Das Übungsblatt enthält Haus- und Tutoriumsaufgaben.
- Die Tutoriumsaufgaben werden in den Tutorien der KW 24 besprochen (10.06. - 16.06.).
- Die Abgabe der Hausaufgaben erfolgt über ISIS in festen Gruppen von 2-3 Personen. Die Gruppen bilden sich aus Studierenden, die das gleiche Tutorium besuchen bzw. mindestens dieselbe/denselben Tutor*in haben. Laden Sie Ihre handschriftlichen Lösungen (z.B. Scan Ihrer Lösungen oder Erstellung Ihrer Lösungen über Tablet) als eine PDF-Datei bei dem enstprechenden Übungsblatt hoch. LaTeX-Abgaben sind auch willkommen (in diesem Fall die kompilierte PDF)! Achten Sie darauf, dass die Abgaben gut lesbar und verständlich verfasst sind. Bitte Namen, Matrikelnummer und Übungsgruppe auf der Abgabe mit angeben!

Tutoriumsaufgaben

Tutoriumsaufgabe 8.1

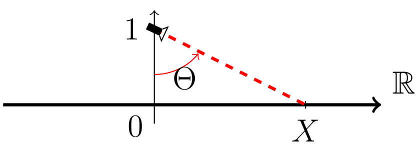
Bereits im 19. Jahrhundert wurde festgestellt, dass in Datensätzen Zahlen, die mit 1 beginnen, häufiger vorkommen als mit anderen Ziffern beginnende Zahlen.

(i) Sei Z gleichverteilt auf $[0,1]$. Berechnen Sie die kumulative Verteilungsfunktion und die Wahrscheinlichkeitsdichte von $Y := 10^Z$.

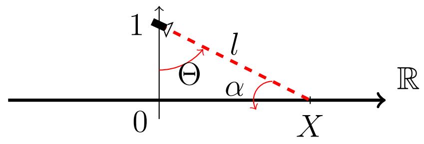
(ii) Mit welcher Wahrscheinlichkeit ist die führende Ziffer einer zufälligen Zahl Y eine Eins bzw. Neun unter der Annahme, dass die Nachkommastellen von $\log_{10} Y$ gleichverteilt auf $[0,1]$ sind?

Tutoriumsaufgabe 8.2

Ein Laser wird in einem Meter Höhe über dem Boden aufgehängt. Der Winkel, den er mit der Senkrechten bildet, ist zufällig, wird mit Θ bezeichnet und folgt der Gleichverteilung auf $[-\frac{\pi}{2}, \frac{\pi}{2}]$. Wir notieren mit X den Punkt, den der Laser auf dem Boden markiert hat (siehe die Abbildung unten).



Der geometrische Zusammenhang zwischen X und Θ ist nun wie folgt zu verstehen: Der Punkt X ist gegeben durch $X = \tan(\Theta)$. Seien α und l gegeben wie in der nachfolgenden Abbildung. So haben wir $1 = l\sin(\alpha)$ und $X = l\cos(\alpha)$.



Da das Dreieck rechtwinklig ist, haben wir $\alpha = \pi/2 - \Theta$ und

$$X = \frac{\cos(\alpha)}{\sin(\alpha)} = \frac{\cos(\pi/2 - \Theta)}{\sin(\pi/2 - \Theta)} = \frac{\sin(\Theta)}{\cos(\Theta)} = \tan(\Theta)$$

(i) Bestimmen Sie die Verteilungsfunktion von X und seine Dichte.

(ii) Prüfen Sie, ob $\mathbb{E}[X]$ existiert, und wenn ja, berechnen Sie diesen.

Tutoriumsaufgabe 8.3

In einer Mensa ist der Reis an jedem Werktag unabhängig voneinander mit Wahrscheinlichkeit $2/5$ verkocht.

(i) Wir betrachten einen Zeitraum von 100 Werktagen. Sei X die Anzahl der Tage, an denen der Reis verkocht ist. Was ist die Verteilung von X? Berechenen $\mathbb{E}[X]$ und $\sigma(X)$.

(ii) Berechnen Sie die Wahrscheinlichkeit des folgenden Ereignisses mit einer geeigneten Approximation: der Reis in der Mensa ist an minimal 40 und maximal 45 in 100 Werktagen verkocht.

Hinweis: Wenden Sie die Standardisierung auf X an und lesen Sie die gesuchten Werte in der angehängten Tabelle ab.

Tutoriumsaufgabe 8.1

Bereits im 19. Jahrhundert wurde festgestellt, dass in Datensätzen Zahlen, die mit 1 beginnen, häufiger vorkommen als mit anderen Ziffern beginnende Zahlen.

(i) Sei Z gleichverteilt auf $[0,1]$. Berechnen Sie die kumulative Verteilungsfunktion und die Wahrscheinlichkeitsdichte von $Y := 10^Z$.

(ii) Mit welcher Wahrscheinlichkeit ist die führende Ziffer einer zufälligen Zahl Y eine Eins bzw. Neun unter der Annahme, dass die Nachkommastellen von $\log_{10} Y$ gleichverteilt auf $[0,1]$ sind?

(i) $Z \sim \text{Uniform}([0,1])$ $\quad Y = 10^Z$

$F_Y(x) = \mathbb{P}(Y \leq x) = \mathbb{P}(10^Z \leq x)$

für $x > 0$: $\mathbb{P}(Z \leq \log_{10} x) = \mathbb{P}(0 \leq Z \leq 1, Z \leq \log_{10} x)$

$$= \begin{cases} 0 & 0 < x \leq 1 \\ \frac{|[0, \log_{10} x]|}{|[0,1]|} & 1 \leq x < 10 \\ 1 & x \geq 10 \end{cases}$$

für $x \leq 0$, $F_Y(x) = 0$

$$F_Y(x) = \begin{cases} 0 & x \leq 1 \\ \log_{10} x & 1 < x \leq 10 \\ 1 & x \geq 10 \end{cases}$$

Dichte von Y:

$$f_Y(x) = F_Y'(x) = \begin{cases} \frac{1}{x \ln 10} & \text{für } x \in [1, 10] \\ 0 & \text{sonst} \end{cases}$$

(ii) $\mathbb{P}(1 \leq Y < 2) = \mathbb{P}(Y \leq 2) - \mathbb{P}(Y \leq 1)$

Hinweis: Sei eine ZV eine stetige Verteilung

$\mathbb{P}(X = x) = 0$ für alle $x \in \mathbb{R}$

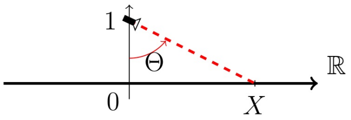
$$= \mathbb{P}(\log_{10} Y \leq \log_{10} 2) - \mathbb{P}(\log_{10} Y \leq \underbrace{\log_{10} 1}_{=0})$$

$$F_{\text{uniform}([0,1])}(\log_{10} 2) = \frac{\log_{10} 2 - 0}{1 - 0} = \log_{10} 2 \approx 0{,}3$$

$\mathbb{P}(9 \leq Y \leq 10) = \dots = 0.046$

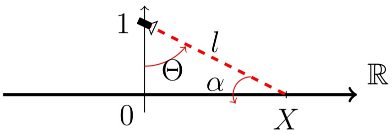
Tutoriumsaufgabe 8.2

Ein Laser wird in einem Meter Höhe über dem Boden aufgehängt. Der Winkel, den er mit der Senkrechten bildet, ist zufällig, wird mit Θ bezeichnet und folgt der Gleichverteilung auf $[-\frac{\pi}{2}, \frac{\pi}{2}]$. Wir notieren mit X den Punkt, den der Laser auf dem Boden markiert hat (siehe die Abbildung unten).



$\Theta \in [-\frac{\pi}{2}, \frac{\pi}{2}]$

Der geometrische Zusammenhang zwischen X und Θ ist nun wie folgt zu verstehen: Der Punkt X ist gegeben durch $X = \tan(\Theta)$. Seien α und l gegeben wie in der nachfolgenden Abbildung. So haben wir $1 = l\sin(\alpha)$ und $X = l\cos(\alpha)$.



Da das Dreieck rechtwinklig ist, haben wir $\alpha = \pi/2 - \Theta$ und

$$X = \frac{\cos(\alpha)}{\sin(\alpha)} = \frac{\cos(\pi/2 - \Theta)}{\sin(\pi/2 - \Theta)} = \frac{\sin(\Theta)}{\cos(\Theta)} = \tan(\Theta)$$

(i) Bestimmen Sie die Verteilungsfunktion von X und seine Dichte.

(ii) Prüfen Sie, ob $\mathbb{E}[X]$ existiert, und wenn ja, berechnen Sie diesen.

(i) Wdh. Gleichverteilung: $Z \sim \text{Uniform}[a,b]$ falls $\mathbb{P}(Z \in A) = \frac{|A|}{|[a,b]|}$

Verteilungsfunktion: $\mathbb{P}(Z \le x) = \mathbb{P}(a \le Z \le b, Z \le x)$

$$= \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & a \le x < b \\ 1 & x \ge b \end{cases}$$

$$f_Z(z) = \begin{cases} 0 & \text{sonst} \\ \frac{1}{b-a} & a \le x < b \end{cases}$$

$$F_X(x) = \mathbb{P}(X \le x) = \mathbb{P}(\tan(\Theta) \le x) = \mathbb{P}(\Theta \le \arctan x)$$

$$= \begin{cases} 0 & \arctan x < -\frac{\pi}{2} \Rightarrow \\ \frac{\arctan(x) - (-\frac{\pi}{2})}{\frac{\pi}{2} - (-\frac{\pi}{2})} = \frac{\arctan(x)}{\pi} + \frac{1}{2} & -\frac{\pi}{2} \le \arctan x < \frac{\pi}{2} \Rightarrow -\infty < x < \infty \\ 1 & \text{sonst} \end{cases}$$

$$f_X(x) = F_X'(x) = \frac{1}{\pi} \frac{1}{1+x^2}$$

(ii) Formel ist $E[X] = \int_{-\infty}^{\infty} x f_X(x)\,dx = \frac{1}{\pi} \int \frac{x}{1+x^2}\,dx = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{1}{2(1+x^2)}\,d(1+x^2) = \frac{1}{2\pi} \ln(1+x^2)$

↑ gleich

oder Alternativ: $E[X] = E[\tan\Theta] = \int_{-\infty}^{\infty} \tan x\, f_\Theta(x)\,dx = \frac{1}{\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \tan(x)\,dx$

$$\frac{1}{\pi} \int \frac{x}{1+x^2}\,dx = \frac{1}{\pi} \Big(\lim_{a \to -\infty} \int_a^0 \frac{x}{1+x^2}\,dx + \lim_{b \to \infty} \int_0^b \frac{x}{1+x^2}\,dx$$

E existiert, falls diese Grenzwerte existieren. es gilt: $\int_0^\infty \frac{x}{1+x^2}\,dx = \lim_{b\to\infty} \int_0^b \frac{x}{1+x^2}\,dx \ge \lim_{b\to\infty} \int_1^b \frac{x}{1+x^2}\,dx$

$$\ge \lim_{b\to\infty} \int_1^b \frac{x}{2x^2}\,dx$$

$$= \frac{1}{2} \lim_{b\to\infty} \int_1^b \frac{1}{x}\,dx$$

$$= \frac{1}{2} \lim_{b\to\infty} |\ln b| = \infty$$

direkt: $\int_0^\infty \frac{x}{1+x^2}\,dx = \int \frac{1}{2(1+x^2)}\,d(1+x^2) = \lim_{b\to\infty} \frac{1}{2}\ln(1+x^2)\Big|_0^b = \infty$

$\int_0^\infty \frac{x}{1+x^2}\,dx$ existiert nicht $\Rightarrow \mathbb{E}[X]$ existiert nicht

Bsp eines ZV mit nicht existierende E-Wert

$X \sim \text{Zipf}(\alpha)$ für $\alpha \in (1,2]$

$\mathbb{E}[X] = \frac{1}{Z(\alpha)} \sum_{k=1}^\infty k \cdot \frac{1}{k^\alpha} = \frac{1}{Z(\alpha)} \sum_{k=1}^\infty \frac{1}{k^{\alpha-1}} = \infty$

Normalverteilung $X \sim \mathcal{N}(\mu, \delta^2)$ für $\mu \in \mathbb{R}$, $\delta^2 > 0$ falls X die Dichte $f_X(x) = \dots$

Kenngrößen: $\mathbb{E}[X] = \mu$ $\mathbb{V}[X] = \delta^2$

Zentrale Grenzwertsatz Sei (X_n) eine Folge von i.i.d (unabhängig identisch verteilt) ZVen. Sei $\mathbb{E}[X_1] = \mu$, $\mathbb{V}[X_1] = \delta^2 > 0$

wir def. $S_n = \frac{\sum_{i=1}^n X_i - n\mu}{\sqrt{n}\,\delta}$

dann gilt: $\lim_{n\to\infty} F_{S_n}(x) = F_{\mathcal{N}(0,1)}(x)$ für $x \in \mathbb{R}$

Tutoriumsaufgabe 8.3

In einer Mensa ist der Reis an jedem Werktag unabhängig voneinander mit Wahrscheinlichkeit $2/5$ verkocht.

(i) Wir betrachten einen Zeitraum von 100 Werktagen. Sei X die Anzahl der Tage, an denen der Reis verkocht ist. Was ist die Verteilung von X? Berechenen $\mathbb{E}[X]$ und $\sigma(X)$.

(ii) Berechnen Sie die Wahrscheinlichkeit des folgenden Ereignisses mit einer geeigneten Approximation: der Reis in der Mensa ist an minimal 40 und maximal 45 in 100 Werktagen verkocht.

Hinweis: Wenden Sie die Standardisierung auf X an und lesen Sie die gesuchten Werte in der angehängten Tabelle ab.

(i) $X \sim \text{Bin}(n = 100, p = \frac{2}{5})$

$\mathbb{E}[X] = np = 40$ $\mathbb{V}[X] = np(1-p) = 40 \cdot \frac{3}{5} = 24$ $\rightarrow \delta = \sqrt{24}$

(ii) $S_{100} = \frac{X - 40}{\sqrt{24}}$ $\Rightarrow X = \sqrt{24}\, S_{100} + 40$

$\mathbb{P}(40 \le X \le 45) = \mathbb{P}(40 \le \sqrt{24}\, S_{100} + 40 \le 45)$

$= \mathbb{P}(0 \le S_{100} \le \frac{5}{\sqrt{24}})$

$= F_{S_{100}}(\frac{5}{\sqrt{24}}) - F_{S_{100}}(0)$

$= F_{\mathcal{N}(0,1)}(1{,}01) - \frac{1}{2}$

$= 0{,}846136 - \frac{1}{2} = 0{,}346136$

Hausaufgaben

Hausaufgabe 8.1 (5=2+2+1 Punkte)

Die Funktion $f : \mathbb{R} \to [0, \infty)$ sei gegeben durch

$$f(x) = \begin{cases} c_1, & -2 \leq x < 0; \\ c_2 x, & 0 \leq x \leq 1; \\ 0, & \text{sonst.} \end{cases}$$

(i) Bestimmen Sie die Konstanten c_1 und c_2, so dass f eine Wahrscheinlichkeitsdichte ist, und die zugehörige Zufallsvariable den Erwartungswert 0 hat.

(ii) Berechnen Sie die Varianz von der zugehörigen Zufallsvariable mit der Dichte f.

(iii) Skizzieren Sie die Dichte f und die zu f gehörige kumulative Verteilungsfunktion F.

Hausaufgabe 8.2 (5=1+3+1 Punkte)

Seien X, Y zwei unabhängige standardnormalverteilte Zufallsvariabeln.

(i) Bestimmen Sie die gemeinsame Dichte $f_{(X,Y)}$.

(ii) Bestimmen Sie die Verteilungsfunktion von $R = X^2 + Y^2$ und seine Dichte.

Hinweis: Das zweidimensionale Integral einer Funktion $g(x^2 + y^2)$ über die Kreisfläche $B(0, \sqrt{r}) := \{(x, y) \in \mathbb{R}^2 | x^2 + y^2 \leq r\}$ kann wie folgt durch ein eindimensionales Integral ausgedrückt werden:

$$\int_{B(0,\sqrt{r})} g(x^2 + y^2)\, dxdy = 2\pi \int_0^{\sqrt{r}} g(\rho^2)\rho\, d\rho$$

(iii) Geben Sie $\mathbb{E}[R]$ und $\mathrm{Var}[R]$ an.

Hausaufgabe 8.3 (4=1+1+2 Punkte)

Die Körpergröße der Mitglieder eines Eishockey-Teams der Männer ist normalverteilt mit Erwartungswert 180 cm und Standardabweichung 6 cm.

(i) Sei X die Zufallsvariable, welche die Körpergröße eines Team-Mitglieds darstellt. Mit welcher Wahrscheinlichkeit ist ein Team-Mitglied bei diesem Kurs größer als 192 cm?

(ii) *Prüfung der Modellierung:* Ist die Wahrscheinlichkeit, dass ein Team-Mitglied eine negative Körpergröße hat ($X < 0$), kleiner als 0.0003?

(iii) Angenommen, die Körpergrößen von zwei unterschiedlichen Team-Mitgliedern sind unabhängig. Wie groß ist die Wahrscheinlichkeit, dass die Summe der beiden Körpergrößen mindestens 372 cm ist?

Hinweis: Wenden Sie die Standardisierung auf X an und lesen Sie die gesuchten Werte in der angehängten Tabelle ab.

Hausaufgabe 8.4 (6=3+3 Punkte)

Seien $Y_1, Y_2, \ldots$ unabhängig und exponential-verteilt mit Parameter $\lambda = 1$. Sei $u > 0$ sowie

$$X_k := \begin{cases} 1 & \text{falls } Y_k > u \\ 0 & \text{sonst.} \end{cases}$$

(i) Bestimmen Sie (mit Begründung)

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} X_k \ ,$$

, wobei der Grenzwert bezüglich stochastischer Konvergenz zu verstehen ist.

Hinweis: Stochastische Konvergenz, wie in der Vorlesung beim Gesetz der großen Zahlen definiert worden ist.

(ii) Sei nun $u = \ln(2)$. Zeigen Sie mit Hilfe des zentralen Grenzwertsatzes, dass

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{2}{\sqrt{n}} \sum_{k=1}^{n} X_k \leq \sqrt{n} \right) = \frac{1}{2}.$$

z_α	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.00	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.10	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.20	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.30	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.40	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.50	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.60	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.70	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.80	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.90	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.00	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.10	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.20	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.30	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.40	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.50	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.60	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.70	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.80	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.90	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.00	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.10	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.20	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.30	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.40	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.50	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.60	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.70	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.80	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.90	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.00	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.10	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.20	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.30	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.40	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998