TU Berlin, Institut für Mathematik Prof. Dr. Noemi Kurt
SoSe 19 András Tóbiás

Stochastik für Informatik, (9LP)
Klausur

30. September 2019

Musterlösung für die Einsicht

Aufgabe 1.1 *10 Punkte*

Ein Signal wird über einen von drei Kanälen A, B, C gesendet. Kanal A wird mit einer Wahrscheinlichkeit von 40% ausgewählt, B und C jeweils mit 30%. Bei Kanal A kommt es mit einer Wahrscheinlichkeit von 10 % zu einem Übertragungsfehler, bei B mit 20% und bei C mit 15%. Für mehrere Signale sind sowohl die Auswahl des Kanals als auch das Auftreten von Übertragungsfehlern unabhängig voneinander.

(a) Stellen Sie die Situation als Baum dar. Schreiben Sie die im Aufgabentext gegebenen Wahrscheinlichkeiten an die richtigen Stellen im Baum.

(b) Mit welcher Wahrscheinlichkeit kommt ein gesendetes Signal korrekt an?

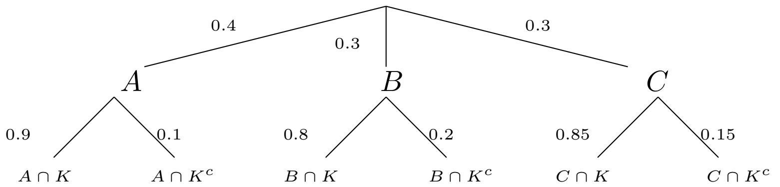
(c) Wenn ein Signal korrekt ankommt, mit welcher Wahrscheinlichkeit wurde es über Kanal B geschickt?

Es werden nun zwei verschiedene Signale gesendet. Berechnen Sie die Wahrscheinlichkeiten für die folgenden Ereignisse:

(d) Mindestens ein Signal wird fehlerhaft übertragen.

(e) Beide Signale werden über denselben Kanal übertragen.

Lösungsskizze zu Aufgabe 1.1. a) Z.B. Notation: A, B, C: der jeweilige Kanal wird verwendet, K: Signal kommt korrekt an.



b) Formel von der Gesamtwahrscheinlichkeit liefert (Gegenwahrscheinlichkeit verwenden)

$$\mathbb{P}(K) = \mathbb{P}(K|A)\mathbb{P}(A) + \mathbb{P}(K|B)\mathbb{P}(B) + \mathbb{P}(K|C)\mathbb{P}(C) = 0.9 \cdot 0.4 + 0.8 \cdot 0.3 + 0.85 \cdot 0.3 = 0.855.$$

c) Bayes-Formel liefert

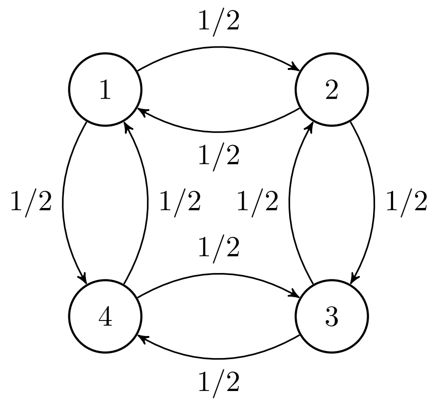
$$\mathbb{P}(B|K) = \frac{\mathbb{P}(K|B)\mathbb{P}(B)}{\mathbb{P}(K)} = \frac{0.8 \cdot 0.3}{0.855} = 0.281.$$

d)Unabhängigkeit: $\mathbb{P}(\text{beide korrekt}) = \mathbb{P}(K)^2 = 0.731$, $\mathbb{P}(\text{mind. einer fehlerhaft}) = 1 - \mathbb{P}(K)^2 = 0.269$. Alternativ: $\mathbb{P}(\text{mind. einer fehlerhaft}) = 2\mathbb{P}(K^c) + \mathbb{P}(K^c)^2 = 1 - 0.731 = 0.269$.
Lösung "$\mathbb{P}(\text{mind. ein fehlerhaft}) = \mathbb{P}(K^c) + \mathbb{P}(K^c)^2$" gibt einen Punkt.

d)Unabhängigkeit: $\mathbb{P}(A)^2 + \mathbb{P}(B)^2 + \mathbb{P}(C)^2 = 0.34$.

Aufgabe 1.2 *10 Punkte*

Sei $(X_n)_{n\in\mathbb{N}_0}$ eine Markov-Kette auf $S = \{1, 2, 3, 4\}$ mit Übergangsgraph



(a) Ist die Markov-Kette irreduzibel? Ist sie aperiodisch? (Begründen Sie).

(b) Bestimmen Sie alle möglicherweise vorhandenen invarianten Verteilungen.

Sei nun die Startverteilung $\nu = (1, 0, 0, 0)$, d.h. $X_0 = 1$.

(c) Geben Sie die Verteilung von X_1 und X_2 an.

(d) Geben Sie $\mathbb{P}(X_n = 1)$ und $\mathbb{P}(X_n = 4)$ jeweils allgemein für gerade und ungerade $n \in \mathbb{N}$ an. Was können Sie über $\lim_{n\to\infty} \mathbb{P}(X_n = 1)$ aussagen?

Lösungsskizze zu Aufgabe 1.2. a) Irreduzibel ja denn jeder Zustand kann von jedem aus erreicht werden, aperiodisch nein, denn man kann nur in einer geraden Anzahl Schritten zurückkommen (Periode 2).

b) Entweder: Da die Matrix symmetrisch ist, müssen alle π_i gleich sein. Weil die Summe $= 1$ sein muss folgt $\pi = (1/4, 1/4, 1/4, 1/4)$. Alternativ durch Lösen des entsprechenden Gleichungssystems aus $(P - I)^T \pi = 0$, mit $(P - I)^T = \begin{pmatrix} -1 & 1/2 & 0 & 1/2 \\ 1/2 & -1 & 1/2 & 0 \\ 0 & 1/2 & -1 & 1/2 \\ 1/2 & 0 & 1/2 & -1 \end{pmatrix}$ und $\pi_1 + \pi_2 + \pi_3 + \pi_4 = 1$ ergibt die Lösung.

c) Da man in 1 startet, ist $\mathbb{P}(X_1 = i) = 1/2$ für $i \in \{2, 4\}$ und 0 sonst (es gibt verschiedene Möglichkeiten, das zu notieren), $\mathbb{P}(X_1 = i) = 1/2$ für $i \in \{1, 3\}$ und 0 sonst.

d) Wegen der Periodizität gelten $\mathbb{P}(X_n = 1) = 1/2$ bei geradem n und 0 sonst, und $\mathbb{P}(X_n = 4) = 1/2$ bei ungeradem n und 0 sonst. Somit kann $\lim_{n\to\infty} \mathbb{P}(X_n = i)$ nicht existieren.

Aufgabe 1.3 — *10 Punkte*

Es sei $f : \mathbb{R} \to \mathbb{R}$ gegeben durch

$$f(x) = \begin{cases} a\mathrm{e}^{-x}, & x \in [0, 2] \\ 0 & \text{sonst.} \end{cases}$$

(a) Bestimmen Sie $a \in \mathbb{R}$ so, dass es sich bei f um eine Dichte einer Zufallsvariablen X handelt.

(b) Berechnen Sie die Verteilungsfunktion von X.

(c) Berechnen Sie $\mathbb{P}(1 \leq X^2 \leq 4)$.

(d) Berechnen Sie $\mathbb{E}[\mathrm{e}^X]$.

Hinweis: Falls Sie Aufgabe (a) nicht lösen konnten, können Sie in die weiteren Aufgabenteile in Abhängigkeit von a lösen.

Lösungsskizze zu Aufgabe 1.3. a)

$$1 \overset{!}{=} \int_0^2 ae^{-x} dx = \left[-ae^{-x}\right]_0^2 = a(1 - e^{-2}),$$

somit folgt $a = \frac{1}{1-e^{-2}} \approx 1.1565$.

b)

$$F_X(t) = \begin{cases} 0 & t \leq 0 \\ \int_0^t ae^{-x}dx = [-ae^{-x}]_0^t = a(1-e^{-t})\Big(= \frac{1-e^{-t}}{1-e^{-2}}\Big) & 0 \leq t \leq 2 \\ 1 & t \geq 2. \end{cases}$$

c) $\mathbb{P}(1 \leq X^2 \leq 4) = \mathbb{P}(1 \leq X \leq 2) = 1 - a(1-e^{-1}) \approx 0.2690$

d)

$$\mathbb{E}[e^X] = \int_0^2 e^x ae^{-x}dx = a\int_0^2 1dx = [ax]_0^2 = 2a \approx 2.3130.$$

Aufgabe 1.4 *10 Punkte*

Die gemeinsame Verteilung von zwei Zufallsvariablen X, Y mit Werten in $\{1,2,3,4\}$ sowie die Randverteilung von Y sind in der folgenden Tabelle gegeben.

X \ Y	1	2	3	4
1	0.24	0.12	0.04	0
2	0.24	0.12	0	0.04
3	0.06	0.03	0.005	0.005
4	0.06	0.03	0.005	0.005
p_Y	0.6	0.3	0.05	0.05

(a) Berechnen Sie die Erwartungswerte von Y und Y^2.

(b) Sind X und Y unabhängig? Warum?

(c) Berechnen Sie die Kovarianz von Y und Y^2.

(d) Sei $g\colon \{1,2,3,4\} \to \{1,2,3\}$ eine Funktion definiert durch $g(1) = 1$, $g(2) = 2$, $g(3) = g(4) = 3$. Geben Sie die gemeinsame Verteilung der Zufallsvariablen $g(X)$ und $g(Y)$ sowie die Randverteilungen von $g(X)$ und $g(Y)$ an.

Lösungsskizze zu Aufgabe 1.4.

(a) $\mathbb{E}Y = 0.6 + 2 \cdot 0.3 + 3 \cdot 0.05 + 4 \cdot 0.05 = 1.55$.

$$\mathbb{E}[Y^2] = \sum_{i=1}^4 i^2\mathbb{P}(Y = i) = 0.6 + 4 \cdot 0.3 + 9 \cdot 0.05 + 16 \cdot 0.05 = 3.05$$

(b) Nein, z.B. $\mathbb{P}(X = 3, Y = 2) = 0 \neq \mathbb{P}(X = 3) \cdot \mathbb{P}(Y = 2)$. (Man kann entweder die Randverteilung von X berechnen oder sagen, dass $\mathbb{P}(X = 3, Y = 2)$ oder $\mathbb{P}(X = 4, Y = 1)$ gleich 0 ist und die entsprechende Randwahrscheinlichkeiten nicht null sind.)

(c) Es gilt

$$\mathrm{cov}(Y, Y^2) = \mathbb{E}[Y \cdot Y^2] - \mathbb{E}[Y]\mathbb{E}[Y^2] = \mathbb{E}[Y^3] - \mathbb{E}[Y]\mathbb{E}[Y^2].$$

$\mathbb{E}[Y]$ und $\mathbb{E}[Y^2]$ wurden schon berechnet. Weiterhin gilt

$$\mathbb{E}[Y^3] = \sum_{i=1}^4 i^3\mathbb{P}(Y = i) = 0.6 + 8 \cdot 0.3 + 27 \cdot 0.05 + 64 \cdot 0.05 = 7.55.$$

Dies liefert $\mathrm{cov}(Y, Y^2) = 7.55 - 4.7275 = 2.8225$.

(d)

$g(X)$ \ $g(Y)$	1	2	3	$p_{g(X)}$
1	0.24	0.12	0.04	0.4
2	0.24	0.12	0.04	0.4
3	0.12	0.06	0.02	0.2
$p_{g(Y)}$	0.6	0.3	0.1	

Aufgabe 1.5 *10 Punkte*

Sei X eine Zufallsvariable mit Dichte

$$f_\theta(x) = \theta^2 x\,\mathrm{e}^{-\theta x}, \qquad x > 0,$$

wobei $\theta > 0$ ein unbekannter Parameter ist.

(a) Ermitteln Sie einen Schätzer für den Parameter θ mittels der Maximum-Likelihood-Methode und geben Sie die Schätzung an.

(b) Sei U gleichverteilt auf $[0, 1]$. Geben Sie eine Funktion $g\colon [0, 1] \to [1, 4]$ an, sodass $g(U)$ auf $[1, 4]$ gleichverteilt ist.

(c) Sei $\theta = 1$ bei obiger Dichte f_θ. Geben Sie den Monte-Carlo-Schätzer für die Wahrscheinlichkeit $\mathbb{P}(1 \le X \le 4)$ mithilfe von 100 unabhängigen und auf $[1, 4]$ gleichverteilten Zufallsvariablen $U_1, \ldots, U_{100}$ an.

Lösungsskizze zu Aufgabe 1.5.

(a) Die Likelihood-Funktion ist

$$L(x_1, \ldots, x_n; \theta) = \theta^{2n} \prod_{i=1}^{n} (x_i\,\mathrm{e}^{-\theta x_i}) \Big(= \theta^{2n} \Big(\prod_{i=1}^{n} x_i\Big)\,\mathrm{e}^{-\theta \sum_{i=1}^{n} x_i} \Big).$$

Die log-Likelihood-Funktion ist

$$l(x_1, \ldots, x_n; \theta) = 2n \log \theta + \sum_{i=1}^{n} \ln x_i - \theta \sum_{i=1}^{n} x_i.$$

Ihre Ableitung

$$\frac{\mathrm{d}}{\mathrm{d}\theta} l(x_1, \ldots, x_n; \theta) = \frac{2n}{\theta} - \sum_{i=1}^{n} x_i$$

ist genau dann null, wenn

$$\theta = \frac{2n}{\sum_{i=1}^{n} x_i}.$$

Das ist eine Maximumstelle, denn

$$\frac{\mathrm{d}^2}{\mathrm{d}\theta^2} l(x_1, \ldots, x_n; \theta) = \frac{-2n}{\theta^2}$$

ist negativ für jedes $\theta > 0$ (genug zu sagen: „überall negativ" oder „negativ für $\theta = \frac{2n}{\sum_{i=1}^{n} x_i}$". Alternativ kann mann checken, dass das Vorzeichen der Ableitung aus positiv zu negativ wächselt), also ist der Schätzer

$$\widetilde{\theta} = \frac{2n}{\sum_{i=1}^{n} X_i} \Big(= \frac{2n}{\bar{\mu}} \Big).$$

(b) $3U_1 + 1, \ldots, 3U_n + 1$ sind u.i.v. gleichverteilt auf $[1, 4]$, also ist $g\colon x \mapsto 3x + 1$ eine solche Funktion.

(c) Sei $\theta = 1$. Zu approximieren ist das Integral

$$\mathbb{P}(1 \le X \le 4) = \int_1^4 f_1(x)\mathrm{d}x = \int_1^4 x\,\mathrm{e}^{-x}\,\mathrm{d}x.$$

Damit ist der Monte-Carlo-Schätzer wie folgt gegeben

$$\frac{b-a}{n} \sum_{i=1}^{n} f_1(U_i) = \frac{3}{100} \sum_{i=1}^{100} U_i \exp(-U_i).$$