TU Berlin, Institut für Mathematik Prof. Dr. Noemi Kurt
SoSe 18 Lukas Wessels

Stochastik für Informatik, (9LP)

Musterlösung

31. Juli 2018

Aufgabe 1 *10 Punkte*

Gegeben sei die Matrix

$$P = \begin{bmatrix} 0 & a & 0.6 \\ b & 0 & 0.7 \\ 0.2 & 0.8 & c \end{bmatrix}.$$

(a) Bestimmen Sie die Zahlen a, b, c so, dass die Matrix stochastisch ist. Zeichnen Sie den zugehörigen Übergangsgraphen, und geben Sie an, ob die zugehörige Markov-Kette irreduzibel ist.

(b) Bestimmen Sie alle gegebenenfalls vorhandenen invarianten Verteilungen der Markov-Kette zur stochastischen Matrix P.

(c) Gegeben seien zusätzlich Sprungraten $\lambda_1 = \lambda_2 = 2$ und $\lambda_3 = 1$. Geben Sie die Q-Matrix der zugehörigen Markov-Kette in stetiger Zeit an, und erklären Sie kurz die Bedeutung der Sprungraten für das Verhalten der Markovkette.

Lösungsskizze zu Aufgabe 1.

(a) Da die Zeilensummen jeweils 1 ergeben müssen, folgt $a = 0.4, b = 0.3, c = 0$. *Skizze.* Die Markov-Kette ist irreduzibel, da man von jedem Zustand aus jeden anderen Zustand erreichen kann (sieht man an der Skizze, bzw. direkt an den Übergangswahrscheinlichkeiten, denn $p_{i,j} \neq 0$ für jedes Paar $i \neq j$.)

(b) Zu lösen ist das Gleichungssystem $\pi^T P = \pi^T, \pi_1 + \pi_2 + \pi_3 = 1$, bzw. $(P - I)^T \pi = 0, \pi_1 + \pi_2 + \pi_3 = 1$. Dies führt auf

$$\begin{aligned} 0.3\pi_2 + 0.2\pi_3 &= \pi_1 \\ 0.4\pi_1 + 0.8\pi_3 &= \pi_2 \\ 0.6\pi_1 + 0.7\pi_2 &= \pi_3 \\ \pi_1 + \pi_2 + \pi_3 &= 1. \end{aligned}$$

Lösen mittels Gauss-Algorithmus oder mittels direkter Rechnungen führt auf die eindeutige Lösung $\pi = (1/5, 2/5, 2/5)^T = (0.2, 0.4, 0.4)^T$.

(c) Mit der Formel $q_{i,i} = -\lambda_i$ und $q_{i,j} = \lambda_i p_{i,j}, i \neq j$, erhält man

$$Q = \begin{bmatrix} -2 & 0.8 & 1.2 \\ 0.6 & -2 & 1.4 \\ 0.2 & 0.8 & -1 \end{bmatrix}.$$

Die Rate λ_i gibt an, dass die Kette im Zustand i eine exponentialverteilte Zeit mit Parameter λ_i wartet, bevor sie den nächsten Sprung ausführt.

Aufgabe 2 *10 Punkte*

Die gemeinsame Verteilung von X und Y sowie die Randverteilungen von X und Y sind in folgender unvollständiger Tabelle gegeben:

$X \backslash Y$	-1	0	1	$\mathbb{P}(X=k)$
-1		$3/20$	0	
0	$1/4$	$1/4$		
1	$1/20$			$1/5$
$\mathbb{P}(Y=k)$			$1/4$	

Die Zufallsvariable X habe Erwartungswert 0.

(a) Bestimmen Sie $\mathbb{P}(X=-1)$ und $\mathbb{P}(X=0)$.
Hinweis: Benutzen Sie die Information über $\mathbb{E}[X]$.

(b) Vervollständigen Sie die gegebene Tabelle.

(c) Bestimmen Sie die bedingte Verteilung von X gegeben $Y=1$.

(d) Bestimmen Sie die gemeinsame Verteilung von X^2 und Y^2.

(e) Sind X^2 und Y^2 unabhängig? Begründen Sie Ihre Antwort.

Lösungsskizze zu Aufgabe 2.

(a) Aus $\mathbb{E}[X]=0$ folgt $\mathbb{P}(X=-1)=1/5$ und $\mathbb{P}(X=0)=3/5$.

(b) Die Lösung ist in der Tabelle:

$X \backslash Y$	-1	0	1	p_X	$p_{X\mid Y=1}$
-1	$1/20$	$3/20$	0	$1/5$	0
0	$1/4$	$1/4$	$1/10$	$3/5$	$2/5$
1	$1/20$	0	$3/20$	$1/5$	$3/5$
p_Y	$7/20$	$2/5$	$1/4$		

(c) siehe Tabelle in (b).

(d) Die Lösung ist in der Tabelle:

$X^2 \backslash Y^2$	0	1	p_{X^2}
0	$1/4$	$7/20$	$12/20$
1	$3/20$	$1/4$	$8/20$
p_{Y^2}	$8/20$	$12/20$	

(e) X^2 und Y^2 sind nicht unabhängig, da $p_{X^2}(0)p_{Y^2}(0) \neq \mathbb{P}(X=0, Y=0)$.

Aufgabe 3 *10 Punkte*

Eine Zeichenfolge wird über einen stark gestörten Kanal übermittelt. Jedes Zeichen kommt unabhängig von den anderen Zeichen mit Wahrscheinlichkeit $p=0.85$ korrekt an, andernfalls kommt es falsch an.

(a) Es werden insgesamt n Zeichen übermittelt. Sei X die Anzahl korrekt übermittelter Zeichen. Welche Verteilung hat X? Geben Sie außerdem $\mathbb{E}[X]$ und $\mathbb{V}(X)$ in Abhängigkeit von n an.

(b) Für $n=10$ bestimme man die Wahrscheinlichkeit, dass höchstens ein Zeichen falsch ankommt.

(c) Für $n=100$ bestimme man mittels Normalapproximation die Wahrscheinlichkeit, dass höchstens 10% der Zeichen falsch ankommt. Warum darf diese Approximation vewendet werden?

Lösungsskizze zu Aufgabe 3.

(a) $X \sim \mathrm{Bin}(n, 0.85)$,
$\mathbb{E}[X]=0.85n$,
$\mathrm{var}(X)=n\cdot 0.85\cdot 0.15=n\cdot 0.1275$.

(b) Mit X aus (a) folgt

$$\begin{aligned}\mathbb{P}(\text{höchstens 1 Zeichen falsch}) &= \mathbb{P}(X \geq 9) = \mathbb{P}(X = 9) + \mathbb{P}(X = 10)\\ &= \binom{10}{9} 0.85^9 \cdot 0.15 + \binom{10}{10} 0.85^{10} \approx 0.3474 + 0.1969 = 0.5443.\end{aligned}$$

(c) 10% von $n = 100$ sind ≤ 10 falsche Zeichen, also $X > 100 - 10 = 90$ (für X definiert wie in (a)). Es gilt $\mathbb{E}[X] = 85$ und $\sigma = \sqrt{12.75} \approx 3.571$. Somit folgt, (nach dem Zentralen Grenzwertsatz bzw. wie in Beispiel 7.6) für $Y \sim \mathcal{N}(0,1)$

$$\begin{aligned}\mathbb{P}(X \geq 89) =& \mathbb{P}\left(\frac{X-\mu}{\sigma} \geq \frac{89-\mu}{\sigma}\right) \approx \mathbb{P}\left(Y \geq \frac{89-85}{3.571}\right)\\ =& 1 - \mathbb{P}(Y \leq \frac{4}{3.571}) = 1 - \Phi_{0,1}(1.120) = 1 - 0.8686 = 0.1314.\end{aligned}$$

Die Approximation darf wegen des zentralen Grenzwertsatzes verwendet werden, und weil die Faustregel $np \geq 5$ und $n(1-p) \geq 5$ erfüllt ist.

Alternative Rechnung unter der Verwendung der 1/2-Korrektur (Satz 7.7 im Skript): $X \geq 89.5$, also

$$\begin{aligned}\mathbb{P}(X \geq 89.5) =& \mathbb{P}\left(\frac{X-\mu}{\sigma} \geq \frac{89.5-\mu}{\sigma}\right) \approx \mathbb{P}\left(Y \geq \frac{89.5-85}{3.571}\right)\\ =& 1 - \mathbb{P}(Y \leq \frac{4.5}{3.571}) = 1 - \Phi_{0,1}(1.260) = 1 - 0.8962 = 0.1038.\end{aligned}$$

Aufgabe 4 *10 Punkte*

Lena hat eine Trickspielmünze gefunden und vermutet, dass sie nicht fair ist. Sie möchte einen χ^2-Test auf Gleichverteilung durchführen, um sich von ihrer Vermutung zu überzeugen.

(a) Stellen Sie eine geeignete Nullhypothese H_0 und Alternativhypothese H_A auf.

(b) Bestimmen Sie den Freiheitsgrad und das für den Test relevante Quantil zum Fehlerniveau $\alpha = 0.05$.

(c) Sei $k \in \{0, 1, \ldots, 10\}$ die Anzahl der Köpfe bei 10 Münzwürfen. Berechnen Sie die Teststatistik (=Testwert) in Abhängigkeit von k.

(d) Für welche Werte von k lässt sich Lenas Vermutung zum Fehlerniveau $\alpha = 0.05$ stützen?

Lösungsskizze zu Aufgabe 4.

(a)

$$\begin{aligned}&H_0: \text{ Münze ist fair (folgt einer Gleichverteilung)},\\ &H_A: \text{ Münze ist nicht fair (folgt keiner Gleichverteilung)}.\end{aligned}$$

(b) Der Freiheitsgrad ist $f = 1$. Das relevante Quantil ist $\chi^2_{1-\alpha,f} = \chi^2_{0.95,1} = 3,84$.

(c) Die Teststatistik ergibt sich als

$$\chi^2 = \sum_{i=1}^{2} \frac{(N_i - F_i)^2}{F_i} = \frac{(k-5)^2}{5} + \frac{(10-k-5)^2}{5} = \frac{2}{5}(k-5)^2.$$

(d) Die Nullhypothese wird verworfen, falls

$$\begin{aligned}&\chi^2 > \chi^2_{1-\alpha,f}\\ \iff& \frac{2}{5}(k-5)^2 > 3.84\\ \iff& |k-5| > \sqrt{\frac{5}{2} * 3.84} \approx 3,098\\ \iff& k \in \{0, 1, 9, 10\}.\end{aligned}$$

Aufgabe 5 *10 Punkte*

Gegeben sei die Funktion $f_\theta : \mathbb{R} \to \mathbb{R}$ mit den Parametern $a, b \in \mathbb{R}$ und

$$f_\theta(x) = \begin{cases} a(b+\theta x^3) & \text{, falls } -1 \leq x \leq 1 \\ 0 & \text{, sonst} \end{cases},$$

wobei $\theta \in [-1,1]$ ein unbekannter Parameter sei.

(a) Bestimmen Sie alle Paare $(a,b) \in \mathbb{R}^2$, sodass es sich bei f_θ für alle $\theta \in [-1,1]$ um eine Dichte handelt.

Sei von nun an $a = \frac{1}{2}$ und $b = 1$ in der Definition von f_θ. Seien $X_1, X_2, \ldots, X_n$, $n \in \mathbb{N}$, unabhängig und identisch verteilte Zufallsvariablen mit Dichte f_θ.

(b) Berechnen Sie den Erwartungswert der Zufallsvariablen X_1.

(c) Zeigen Sie, dass der Schätzer $\theta_n(X_1, \ldots, X_n) := \frac{1}{n}\sum_{k=1}^n X_k$ für den unbekannten Parameter θ, nicht erwartungstreu ist.

(d) Geben Sie einen erwartungstreuen Schätzer $\tilde{\theta}(X_1, \ldots, X_n)$ für den unbekannten Parameter θ an.

Lösungsskizze zu Aufgabe 5.

(a) Es muss für alle $\theta \in [-1,1]$ gelten

$$\int_{-\infty}^{\infty} f_\theta(x)\mathrm{d}x = 1,$$

also

$$\int_{-1}^{1} a(b+\theta x^3)\mathrm{d}x = 2ab = 1.$$

Daraus folgt schon $a, b \neq 0$ und $a = 1/(2b)$. Weiterhin muss gelten $f_\theta(x) \geq 0$ für alle $x \in \mathbb{R}$, $\theta \in [-1,1]$, also

$$a(b+\theta x^3) \geq 0 \quad \forall x, \theta \in [-1,1].$$

1. Fall: $a > 0$:

$$\begin{aligned} & b+\theta x^3 \geq 0 \quad \forall x, \theta \in [-1,1] \\ \iff & b \geq -\theta x^3 \quad \forall x, \theta \in [-1,1] \\ \iff & b \geq 1. \end{aligned}$$

2. Fall: $a < 0$:

$$\begin{aligned} & b+\theta x^3 \leq 0 \quad \forall x, \theta \in [-1,1] \\ \iff & b \leq -\theta x^3 \quad \forall x, \theta \in [-1,1] \\ \iff & b \leq -1. \end{aligned}$$

Also ist die Lösungsmenge

$$\left\{(a,b) \in \mathbb{R}^2 | b \geq 1, a = 1/(2b)\right\} \cup \left\{(a,b) \in \mathbb{R}^2 | b \leq -1, a = 1/(2b)\right\}.$$

(b)

$$\mathbb{E}[X_1] = \int_{-\infty}^{\infty} x f_\theta(x)\mathrm{d}x = \frac{1}{2}\int_{-1}^{1} x + \theta x^4 \mathrm{d}x = \frac{\theta}{5}.$$

(c)

$$\mathbb{E}[\theta_n(X_1, \ldots, X_n)] = \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^n X_i\right] = \frac{1}{n}\sum_{i=1}^n \mathbb{E}[X_i] = \frac{\theta}{5}.$$

(d) $\tilde{\theta}(X_1, \ldots, X_n) := 5\theta_n(X_1, \ldots, X_n)$.