Technische Universität München
Fakultät für Informatik
Lehrstuhl für Theoretische Informatik
Prof. Dr. Susanne Albers
Dennis Kraft, Richard Stotz

Sommersemester 2015
Lösungsblatt 12
20. Juli 2015

Diskrete Wahrscheinlichkeitstheorie

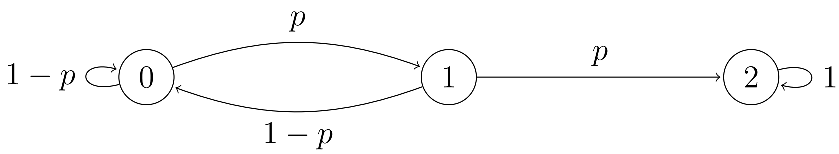
Tutoraufgabe 1

Nachdem Sie auf Übungsblatt 10 zu viele Experimente mit Wolpertingern durchgeführt haben, fühlen Sie sich von diesen possierlichen Tierchen verfolgt. Sie sind davon überzeugt, dass eine Katastrophe passieren wird, falls Sie an zwei aufeinanderfolgenden Tagen einen Wolpertinger sehen. Angenommen Ihnen erscheint jeden Tag unabhängig und mit gleicher Warhscheinlichkeit $p > 0$ ein Wolpertinger.

1. Modellieren Sie das Szenario durch eine Markov-Kette.
2. Wie viele Tage erwarten Sie bis zur Katastrophe wenn Sie am Vortag keinen Wolperdinger gesehen haben?
3. Wenn Sie einen Wolpertinger sehen, aber noch kein katastrophales Ereignis eingetreten ist, haben Sie einen gestressten Tag. Wie hoch ist an einem solchen Tag die Wahrscheinlichkeit, einen weiteren gestressten Tag zu erleben?

Lösungsvorschlag

Sei Zustand 0 ein Tag, an dem Ihnen keinen Wolpertinger begegnet, 1 ein Tag, an dem Sie zwar einen Wolpertinger sehen, allerdings nicht am Vortag, und 2 der Tag der Katastrophe. Das folgenden Übergangsdiagramm modelliert unser Szenario.



Wir suchen nun die erwartet Anzahl an Schritten um von Zustand 0 in Zustand 2 zu kommen und bezeichnen diese mit $h_{0,2}$. Außerdem sei die erwartet Anzahl an Schritten um von Zustand 1 Zustand 2 zu erreichen gegeben durch $h_{1,2}$. Nach der Formel aus der Vorlesung lassen sich $h_{0,2}$ und $h_{1,2}$ beschreiben durch

$$h_{0,2} = 1 + (1-p) \cdot h_{0,2} + p \cdot h_{1,2} \quad \text{und} \quad h_{1,2} = 1 + (1-p) \cdot h_{0,2} + 0 \cdot h_{1,2}$$

Setzen wir die zweite Formel in die Erste ein so erhalten wir

$$\begin{aligned} & h_{0,2} = 1 + (1-p) \cdot h_{0,2} + p \cdot (1 + (1-p) \cdot h_{0,2}) \\ \iff \quad & (1 - 1 + p - p + p^2) \cdot h_{0,2} = 1 + p \\ \iff \quad & h_{0,2} = \frac{1+p}{p^2}. \end{aligned}$$

Sie haben also erwartungsgemäß noch $\frac{1+p}{p^2}$ Tage bis zur Katastrophe. Kommen wir nun zum dritten Teil der Aufgabe. Die Wahrscheinlichkeit, nach einem gestressten Tag wieder einen solchen zu erleben, entspricht der Wahrscheinlichkeit, vom Zustand 1 zurück in 1 zu kommen. Sei $f_{i,j}$ die Wahrscheinlichkeit von Zustand i zu Zustand j zu gelangen. Für den Fall, dass i gleich j ist kürzen wir ab mit f_i. Wir sind also interessiert an f_1 und laut Vorlesung gilt

$$f_1 = 0 + p \cdot f_{2,1} + (1-p) \cdot f_{0,1}.$$

Da Zustand 2 absorbierend ist, ist die Wahrscheinlichkeit von 2 zurück nach 1 zu kommen gleich 0. Andererseits gilt

$$f_{0,1} = p + (1-p) \cdot f_{0,1} + 0 \cdot f_{2,1} \iff f_{0,1} = 1.$$

Wir setzen dieses Ergebnis in obige Gleichung ein und erhalten eine Wahrscheinlichkeit von $1-p$ für einen weiteren gestressten Tag.

Tutoraufgabe 2

Sei X_t mit $t \geq 0$ eine endliche Markov-Kette über der Zustandsmenge S mit einer stationären Verteilung π. Beweisen oder widerlegen Sie die Behauptung $\lim_{t\to\infty} p_{ij}^{(t)} = \pi_j$ mit $i, j \in S$ für ergodische Markov-Ketten, irreduzible Markov-Ketten und aperiodische Markov-Ketten.

Lösungsvorschlag

Betrachten wir die Aussage $\lim_{t\to\infty} p_{i,j}^{(t)} = \pi_j$ zunächst für ergodische Markov-Ketten. Nach dem Fundamentalsatz für ergodische Markov-Ketten aus der Vorlesung gilt, dass die Verteilung q_t unabhängig von der Startverteilung q_0 gegen eine eindeutige stationäre Verteilung π konvergiert. Formaler ausgedrückt bedeutet das

$$\lim_{t\to\infty} q_t = \pi.$$

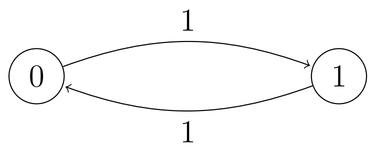
Wählen wir nun als Startzustand deterministisch i, also $(q_0)_i = 1$ und $(q_0)_j = 0$ für alle $j \neq i$, dann folgt aus der Definition von q_t, dass

$$(q_t)_j = \sum_{i\in S} (q_0)_i \cdot p_{i,j}^{(t)} = p_{i,j}^{(t)},$$

was unseren Beweis vervollständigt, da

$$\lim_{t\to\infty} p_{i,j}^{(t)} = \lim_{t\to\infty} (q_t)_j = \pi_j.$$

Sollte unsere Markov-Kette hingegen nicht ergodisch sein, so können wir den Fundamentalsatz nicht anwenden und es stellt sich heraus, dass weder Irreduzibilität noch Aperiodizität alleine stark genug sind, so dass unsere Behauptung hält. Für irreduzible Markov-Ketten betrachten wir das folgende Gegenbeispiel.



Zunächst überzeugen wir uns davon, dass es sich tatsächlich um eine irreduzible Markov-Kette handelt. Dies sieht man bspw. daran, dass das Übergangsdiagramm stark zusammenhängend ist. Betrachten wir nun die Übergangsmatrix P, so gilt für gerade t mit $t = 2k$, dass

$$P^t = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}^{2k} = \left(\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}^2\right)^k = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}^k = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

und für ungerade t mit $t = 2k+1$

$$P^t = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}^{2k+1} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}^{2k} \cdot \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Folglich kann $p_{i,j}^{(t)}$ nicht konvergieren was unserer Behauptung widerspricht. Betrachten wir nun ein Gegenbeispiel für eine aperiodische Markov-Ketten.

Nachdem jeder Zustand eine Schleife besitzt, handelt es sich offensichtlich um eine aperiodische Markov-Kette. Außerdem ist die Übergangsmatrix P die Einheitsmatrix und somit gilt

$$P^t = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}^t = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

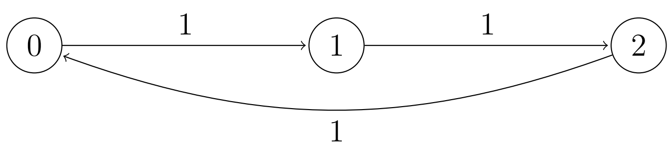
Der Grenzwert von $p_{i,j}^{(t)}$ ist also gleich 1 ist für $i = j$ und 0 für $i \neq j$. Dies ist allerdings ein Widerspruch zu der Annahme, dass $\lim_{n\to\infty} p_{i,j}^{(t)}$ unabhängig von i einen Wert π_j annimmt.

Tutoraufgabe 3

Beweisen oder widerlegen Sie die folgende Aussage. Handelt es sich bei X_t mit $t \geq 0$ um eine Markov-Kette über den Zuständen $S = \{0, \ldots, n\}$, so ist auch $X_t \bmod k$ eine Markov-Kette mit $R = \{0, \ldots, k-1\}$ Zuständen.

Lösungsvorschlag

Die Behauptung ist im Allgemeinen falsch, wie wir am folgenden Beispiel leicht sehen können. Sei X_t eine Markov-Kette mit den Zuständen $S = \{0, 1, 2\}$ und dem folgenden Übergangsdiagramm.



Betrachten wir nun die Zufallsvariablen $X_t \bmod 2$. Befinden wir uns zum Zeitpunkt $i-2$ in Zustand 2, so sind wir nach einem Schritt in Zustand 0 und im zweiten Schritt automatisch in Zustand 1. Es gilt also

$$\Pr[X_t \bmod 2 = 1 \mid X_{t-2} \bmod 2 = 0, X_{t-1} \bmod 2 = 0] = 1.$$

Starten wir hingegen in Zustand 1, so führt uns ein Schritt zu Zustand 2 und ein weiterer Schritt zu Zustand 0 und es folgt

$$\Pr[X_t \bmod 2 = 1 \mid X_{t-2} \bmod 2 = 1, X_{t-1} \bmod 2 = 0] = 0.$$

Obwohl in beiden Fällen $X_{t-1} \bmod 2$ den gleichen Wert hat, nämlich 0, erhalten wir unterschiedliche Wahrscheinlichkeiten für $X_t \bmod 2 = 1$. Folglich hängt der Wert der Zufallsvariablen $X_t \bmod 2$ nicht notwendigerweise von der unmittelbar vorherigen Zufallsvariable $X_{t-1} \bmod 2$ ab, sondern auch von $X_{t-2} \bmod 2$, was bedeutet, dass sich $X_t \bmod 2$ nicht als Markov-Kette darstellen lässt.

Hausaufgabe 1 (5 Punkte)

Sei X_i für $i \geq 0$ eine Familie von unabhängigen Zufallsvariablen, die jeweils mit Wahrscheinlichkeit $\frac{1}{2}$ die Werte 1 und -1 annehmen. Des Weiteren sei $Y_t = (\sum_{i=0}^{t} X_i) \bmod 4$ mit $t \geq 0$ eine Markov-Kette über den Zuständen $S = \{0, 1, 2, 3\}$. Zeichnen Sie ein Übergangsdiagramm für Y_t und geben Sie mit kurzer Begründung an, ob es sich um eine ergodische Markov-Kette handelt.

Lösungsvorschlag

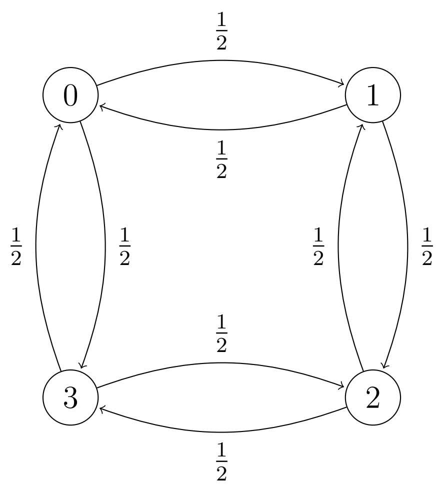
Angenommen wir befinden uns zum Zeitpunkt t in Zustand i. Die Wahrscheinlichkeit, dass der Folgezustand entweder $(i+1) \bmod 4$ oder $(i-1) \bmod 4$ ist, beträgt somit jeweils $\frac{1}{2}$. Um dies zu beweisen betrachten wir $p_{i,(i+1) \bmod 4}$ und formen um zu

$$\begin{aligned}
p_{i,(i+1) \bmod 4} &= \Pr[Y_{t+1} = (i+1) \bmod 4 \mid Y_t = i] \\
&= \Pr\left[\left(\sum_{i=0}^{t+1} X_i\right) \bmod 4 = (i+1) \bmod 4 \middle| Y_t = i\right] \\
&= \Pr\left[\left(\left(\sum_{i=0}^{t} X_i\right) \bmod 4 + X_{t+1}\right) \bmod 4 = (i+1) \bmod 4 \middle| Y_t = i\right] \\
&= \Pr\left[(Y_t + X_{t+1}) \bmod 4 = (i+1) \bmod 4 | Y_t = i\right] \\
&= \Pr[(s + X_{t+1}) \bmod 4 = (i+1) \bmod 4 \mid Y_t = i] \\
&= \Pr[X_{t+1} \bmod 4 = 1 \mid Y_t = i] \\
&= \Pr[X_{t+1} = 1 \mid Y_t = i],
\end{aligned}$$

wobei der letzte Umformungsschritt gültig ist, da der Wertebereich von X_{t+1} lediglich 1 und -1 umfasst. Nachdem X_{t+1} unabhängig von allen X_i mit $i \leq t$ ist, und somit insbesondere unabhängig von Y_t, gilt außerdem

$$\Pr[X_{t+1} = 1 \mid Y_t = i] = \Pr[X_{t+1} = 1] = \frac{1}{2}.$$

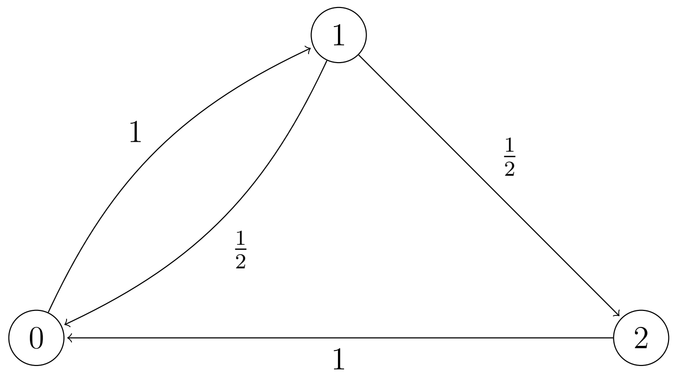
Der Beweis für $p_{i,(i-1) \bmod 4}$ verläuft analog. Das Übergangsdiagramm unserer Markov-Kette ist als ein Kreis, wobei man von jedem Zustand i die Nachbarzustände $(i+1) \bmod 4$ und $(i-1) \bmod 4$ jeweils mit Wahrscheinlichkeit $\frac{1}{2}$ erreicht.



Am Übergangsdiagramm sehen wir auch, dass es sich nicht um eine ergodische Markov-Kette handeln kann. Zwar ist das Übergangsdiagramm stark zusammenhängend, was bedeutet, dass es sich um eine irreduzible Markov-Kette handelt. Andererseits gibt es lediglich gerade Kreise im Übergangsdiagramm, was wiederum bedeutet, dass jeder Zustand eine Periode von mindestens 2 hat. Folglich ist die Markov-Kette nicht aperiodisch und somit auch nicht ergodisch.

Hausaufgabe 2 (5 Punkte)

Berechnen Sie alle stationären Verteilungen π einer Markov-Kette mit dem folgendem Übergangsdiagramm.



Lösungsvorschlag

Eine Verteilung π ist stationär wenn $\pi \cdot P = \pi$ gilt, wobei P die Übergangsmatrix der entsprechenden Markov-Kette ist. In unserer Aufgabe ist P also

$$P = \begin{pmatrix} 0 & 1 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 1 & 0 & 0 \end{pmatrix},$$

wodurch wir das folgende Gleichungssystem für π erhalten

$$\begin{aligned} \frac{1}{2} \cdot \pi_2 + \pi_3 &= \pi_1 \\ \pi_1 &= \pi_2 \\ \frac{1}{2} \cdot \pi_2 &= \pi_3. \end{aligned}$$

Aus der zweiten und dritten Gleichung können wir herauslesen, dass $\pi_1 = \pi_2 = 2\pi_3$ gelten muss. Damit es sich bei π um einen Verteilungsvektor handelt, muss außerdem die Summe $\pi_1 + \pi_2 + \pi_3$ gleich 1 sein. Die einzige Verteilung, die alle diese Bedingungen erfüllt lautet demnach $\pi = (\frac{2}{5}, \frac{2}{5}, \frac{1}{5})$.

Hausaufgabe 3 (5 Punkte)

Wahr oder Falsch? Begründen Sie Ihre Antworten.

1. Seien X und Y unabhängige standardnormalverteilte Zufallsvariablen, dann ist der Durchschnitt $\frac{X+Y}{2}$ ebenfalls standardnormalverteilt.

2. Die Zufallsvariablen X_1 bis X_n seien unabhängig und identisch verteilt mit Standardabweichung σ und Erwartungswert μ. Des Weiteren sei $Y_n = \sum_{i=1}^{n} X_i$ für alle $n \in \mathbb{N}$. Die Zufallsvariable $Z_n = \frac{Y_n - n\mu}{\sigma\sqrt{n}}$ ist asymptotisch standardnormalverteilt.

3. Es existieren zwei kontinuierliche Dichtefunktionen $f(x)$ und $g(x)$, die keiner Gleichverteilung entsprechen, so dass $f(x) \cdot g(x)$ ebenfalls eine Dichtefunktion ist.

4. Die Exponentialverteilung ist die einzige kontinuierliche Verteilung mit Wertebereich in den positiven reellen Zahlen, die gedächtnislos ist.

5. Sind X, Y zwei unabhängige und identisch verteilte Zufallsvariablen, dann gilt $\mathbb{E}\left[\frac{X}{Y}\right]$ ist gleich $\frac{\mathbb{E}[X]}{\mathbb{E}[Y]}$.

Lösungsvorschlag

1. **Falsch:** Nach Vorlesung gilt $\frac{X+Y}{2} \sim N(0, \frac{1}{2})$.

2. **Falsch:** Sei X_i diskret mit $\Pr[X_1 = 0] = 1$. Dann ist Z_n wegen $\sigma = 0$ nicht definiert.

3. **Wahr:** Sei beispielsweise $f(x) = 2x$ auf dem Intervall $[0, 1]$ und 0 überall sonst. Sei außerdem $g(x) = \frac{3x}{2}$ auf dem Intervall $[0, \frac{2}{\sqrt{3}}]$ und 0 überall sonst. Dann ist

$$f(x) \cdot g(x) = \begin{cases} 3x^2 & \text{falls } 0 \leq x \leq 1 \\ 0 & \text{sonst} \end{cases}.$$

   Da diese Funktion positiv ist und ihr Integral über die reellen Zahlen 1 ergibt, ist $f(x) \cdot g(x)$ eine Dichtefunktion.

4. **Wahr:** Siehe Vorlesung Satz 98.

5. **Falsch:** Seien X und Y unabhängige und identisch verteilte Zufallsvariablen, die jeweils mit Wahrscheinilhckeit $\frac{1}{2}$ die Werte 1 und 2 annehmen. Da X und Y den selben Erwartungswert haben ist $\frac{\mathbb{E}[X]}{\mathbb{E}[Y]}$ gleich 1. Andererseits gelten folgende Wahrscheinlichkeiten für $\frac{X}{Y}$.

$$\begin{aligned} \Pr\left[\frac{X}{Y} = \frac{1}{2}\right] &= \Pr[X = 1, Y = 2] = \frac{1}{4}, \\ \Pr\left[\frac{X}{Y} = 1\right] &= \Pr[X = 1, Y = 1] + \Pr[X = 2, Y = 2] = \frac{1}{2}, \\ \Pr\left[\frac{X}{Y} = 2\right] &= \Pr[X = 2, Y = 1] = \frac{1}{4}. \end{aligned}$$

   Damit ist $\mathbb{E}[\frac{X}{Y}]$ gleich $\frac{9}{8}$ und nicht wie gefordert 1.

Hausaufgabe 4 (5 Punkte)

Wahr oder Falsch? Begründen Sie Ihre Antworten.

1. Sei X eine Zufallsvariable mit Varianz σ^2 und sei $\bar{X}$ das Stichprobenmittel über n unabhängigen Stichprobenvariablen X_i, dann ist $\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$ ein erwartungstreuer Schätzer für σ^2.
2. Das 0,5-Quantil der Standardnormalverteilung ist 0.
3. Für jede erwartungstreue Schätzvariable X ist der Bias gleich dem Erwartungswert von X.
4. Bei echten Alternativtests ist die Summe der Wahrscheinlichkeit eines Fehlers erster Art und der Wahrscheinlichkeit eines Fehlers zweiter Art stets 1.
5. Falls ein Maximum-Likelihood-Schätzwert existiert, so ist er eindeutig.

Lösungsvorschlag

1. **Falsch:** Nach Vorlesung ist $\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$ nicht erwartungstreu.
2. **Wahr:** Folgt aus der Symmetrie der Standardnormalverteilung.
3. **Falsch:** Die Aussage gilt nur, falls der Erwartungswert 0 ist.
4. **Falsch:** Siehe Blatt 11, Tutoraufgabe 1.
5. **Falsch:** Sei $f(x)$ eine diskrete Dichtefunktion mit Parameter $\theta \in \mathbb{R}$ so dass

$$f(x) = \begin{cases} \frac{1}{2} & \text{falls } x = \theta + 1 \\ \frac{1}{2} & \text{falls } x = \theta - 1 \\ 0 & \text{sonst} \end{cases}.$$

   Für eine einzelne Stichprobe x gilt nun, dass sowohl $\theta = x - 1$ also auch $\theta = x + 1$ die Likelihood-Funktion maximieren, was der Eindeutigkeit widerspricht.