Technische Universität München
Fakultät für Informatik
Lehrstuhl für Effiziente Algorithmen
Prof. Dr. Ernst W. Mayr
Dr. Werner Meixner

Sommersemester 2011
Lösungen der Klausur
6. August 2011

Diskrete Wahrscheinlichkeitstheorie

Name	Vorname	Studiengang	Matrikelnummer
....	....	☐ Diplom ☐ Inform. ☐ Bachelor ☐ BioInf. ☐ Lehramt ☐ WirtInf.	....

Hörsaal	Reihe	Sitzplatz	Unterschrift
....	....	....	....

Code:

Allgemeine Hinweise

- Bitte füllen Sie obige Felder in Druckbuchstaben aus und unterschreiben Sie!
- Bitte schreiben Sie nicht mit Bleistift oder in roter/grüner Farbe!
- Die Arbeitszeit beträgt 90 Minuten.
- Alle Antworten sind in die geheftete Angabe auf den jeweiligen Seiten (bzw. Rückseiten) der betreffenden Aufgaben einzutragen. Auf dem Schmierblattbogen können Sie Nebenrechnungen machen. Der Schmierblattbogen muss ebenfalls abgegeben werden, wird aber in der Regel nicht bewertet.

Hörsaal verlassen von ...... bis ...... / von ...... bis ......
Vorzeitig abgegeben um ......
Besondere Bemerkungen:

	A1	A2	A3	A4	A5	Σ	Korrektor
Erstkorrektur							
Zweitkorrektur							

Aufgabe 1 (6 Punkte)

Wahr oder falsch? Begründen Sie Ihre Antwort!

1. Die Menge $\mathcal{B}(\mathbb{R})$ der Borelschen Mengen über $\mathbb{R}$ ist abzählbar.
2. Jede aperiodische Übergangsmatrix ist irreduzibel.
3. Der Maximum-Likelihood-Schätzwert für den Parameter δ einer Verteilungsdichte $f(x;\delta)$ stellt den wahrscheinlichsten Wert für δ dar.
4. Die Summe $X_1 + X_2$ unabhängiger exponentialverteilter Zufallsvariablen X_1 und X_2 ist ebenfalls exponentialverteilt.
5. Für die gemeinsame Verteilung $F_{X,Y}$ zweier kontinuierlicher Zufallsvariablen X und Y gilt $F_{X,Y}(x,y) = \Pr[X \leq x, Y \leq y]$.
6. Es gibt unendliche Markov-Ketten mit diskreter Zeit, für die alle Zustände transient sind.

Lösung

Für die richtige Antwort und für die richtige Begründung gibt es jeweils einen $\frac{1}{2}$ Punkt.

1. Falsch! Begründung: $\mathcal{B}(\mathbb{R})$ enthält alle Intervalle $[a,b]$ mit $a,b \in \mathbb{R}$.
2. Falsch! Eine geeignete Diagonalmatrix ist aperiodisch und nicht irreduzibel.
3. Falsch! Begründung: δ ist keine Zufallsvariable.
4. Falsch! Verweis auf mehrere Übungsaufgaben möglich.
5. Wahr! Definition!
6. Wahr! Begründung: $p_{i(i+1)} = 1$.

Aufgabe 2 (8 Punkte)

Wir betrachten einen Zufallsprozess, bei dem gleichverteilte Punkte aus dem Bereich $B = \{(x,y) \in \mathbb{R}^2 \,;\, |x| \leq 1 - y^2\}$ ausgewählt werden. Die Koordinaten der Punkte sind also gegeben durch kontinuierliche Zufallsvariablen X, Y über $\mathbb{R}^2$.

1. Berechnen Sie die gemeinsame Dichtefunktion $f_{X,Y}$!
   Welchen Zahlenwert besitzt $f_{X,Y}(0,0)$?

2. Berechnen Sie für alle $y \leq 0$ die Randverteilung $F_Y(y)$!
   Welchen Zahlenwert besitzt $F_Y(-\frac{1}{2})$?

3. Sind X und Y unabhängig? Begründen Sie Ihre Antwort!

Lösung

1. $F_B = 4 \cdot \int_0^1 (1 - y^2)\,\mathrm{d}\,y = 4 \cdot \frac{2}{3}$.
   $f_{X,Y}(x,y) = \frac{3}{8}$ für alle $(x,y) \in B$,
   $f_{X,Y}(x,y) = 0$ für alle $(x,y) \notin B$ (3 P.)

2. Für $-1 \leq y \leq 0$ gilt
   $F_Y(y) = \frac{3}{8} \cdot 2 \cdot \int_{-1}^{y} (1 - y^2)\,\mathrm{d}\,y = \frac{3}{4}y - \frac{1}{4}y^3 + \frac{1}{2}$.
   Für $y < -1$ gilt $F_Y(y) = 0$.
   $F_Y(-\frac{1}{2}) = \frac{5}{32}$. (3 P.)

3. Nein! Begründung: $F_{X,Y}(-\frac{3}{4}, -\frac{1}{2}) = 0 \neq F_X(-\frac{3}{4}) \cdot F_Y(-\frac{1}{2})$.
   (2 P.)

Aufgabe 3 (8 Punkte)

Wir betrachten einen Zähler Z, der bei Start von 0 beginnend in Schritten von 1 beliebig hoch zählen kann. Die für einen beliebigen Zählschritt i benötigte Zeit $0 < T_i \in \mathbb{R}$ (gemessen in vorgegebenen Zeiteinheiten) sei jeweils unabhängig exponentialverteilt mit gleichem Parameter $\lambda = \frac{1}{2}$.

1. Wie groß ist, vom Start aus gerechnet, die erwartete Zeit, bis der Zähler auf 3 hochzählt?

2. Berechnen Sie die Wahrscheinlichkeit p, dass der Zähler zum Hochzählen auf den Wert 2 nicht länger als 10 Zeiteinheiten benötigt.

   Geben Sie dabei den Wert von p durch einen arithmetischen Ausdruck an.

3. Sei $Z(t)$ diejenige Zufallsvariable, die für den Zeitpunkt t den Zählerstand angibt.

   Bestimmen Sie für $t = 10$ die Dichtefunktion $f_{Z(t)}$ und geben Sie für $\Pr[Z(10) = 4]$ einen arithmetischen Ausdruck an.

Hinweis: In arithmetischen Ausdrücken dürfen bekannte Funktionen der Kombinatorik und die Exponentialfunktion unausgewertet verwendet werden.

Lösung

1. $3 \cdot \mathbb{E}[T_1] = 3 \cdot \frac{1}{\lambda} = 6\,.$

   (2 P.)

2. (siehe Übungsblatt 10)

   $p = F_{T_1+T_2}(10) = 1 - e^{-\lambda t} - \lambda t e^{-\lambda t} = 1 - e^{-5} - 5e^{-5} = 1 - 6e^{-5}\,.$

   (3 P.)

3. $Z(t)$ ist Poisson-verteilt mit Parameter $\lambda t = 5$.

   $\Pr[Z(10) = 4] = \frac{(\lambda t)^4}{4!} e^{-\lambda t} = \frac{(5)^4}{4!} e^{-5}\,.$

   (3 P.)

Aufgabe 4 (8 Punkte)

Sei $X_1, X_2, \ldots, X_i, \ldots$ eine Folge von unabhängigen Bernoulli-verteilten Zufallsvariablen mit gleicher Erfolgswahrscheinlichkeit p, wobei $0 < p < 1$ gelte.

1. Bestimmen Sie für alle $n \in \mathbb{N}$ reelle Zahlen a_n, b_n, so dass für $Z_n = a_n \cdot (\sum_{i=1}^{n} X_i) + b_n$ gilt: $\mathbb{E}[Z_n] = 0$ und $\mathrm{Var}[Z_n] = 1\,.$

2. Welchen Zahlenwert besitzt $\lim_{n\to+\infty} \Pr[Z_n \leq 0]$? Begründen Sie Ihre Antwort!

3. Sei $n = 1000$. Wir nehmen an, dass n genügend groß ist, um einen approximativen Binomialtest für p ausreichend genau durchführen zu können. Dabei sei die Hypothese $H_0 : p \leq \frac{1}{100}$ mit trivialer Alternative $H_1 : p > \frac{1}{100}$ gegeben.

   Der Test liefere $\sum_{i=1}^{n} X_i = 25$.

   Kann die Hypothese H_0 mit Signifikanz $\alpha = 0{,}001$ abgelehnt werden, wenn wir annehmen, dass für das Quantil $z_{1-\alpha}$ der Wert $z_{1-\alpha} = 3{,}1$ gilt? Begründung!

Lösung

1. Gleichungen:

   $\mathbb{E}[Z_n] = a_n \cdot n \cdot p + b_n = 0 \quad$ und $\quad \mathrm{Var}[Z_n] = a_n^2 \cdot n \cdot p(1-p) = 1\,.$

   Daraus: $\quad a_n = \frac{1}{\sqrt{np(1-p)}} \quad$ und $\quad b_n = -\sqrt{\frac{np}{1-p}}\,.$

   (3 P.)

2. $\lim_{n\to+\infty} \Pr[Z_n \leq 0] = \frac{1}{2}\,.$

   Begründung: die Verteilung von Z_n konvergiert gegen die Standardnormalverteilung Φ mit $\Phi(0) = \frac{1}{2}\,.$

   (2 P.)

3. Antwort: Ja.

   Ablehnungskriterium: $Z_n > z_{1-\alpha} = 3{,}1$.

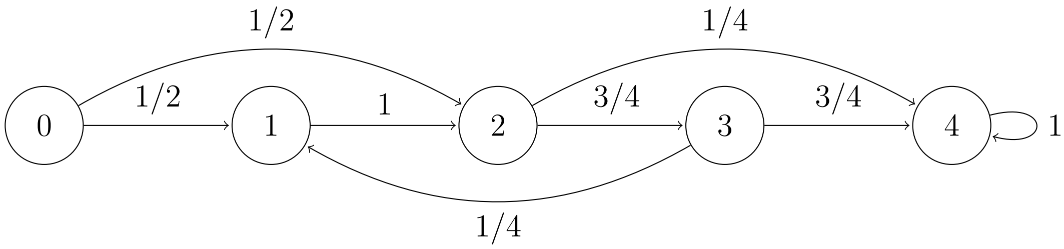
   Mit $p = \frac{1}{100}$ und $\sum_{i=1}^{n} X_i = 25$ gilt

$$\begin{aligned} Z_n &= a_n \cdot \left(\sum_{i=1}^{n} X_i\right) + b_n \\ &= \frac{25 - 1000 \cdot \frac{1}{100}}{\sqrt{1000 \cdot \frac{1}{100} \cdot \frac{99}{100}}} \\ &= 5 \cdot \sqrt{\frac{10}{11}}\,. \end{aligned}$$

   Es gilt $5 \cdot \sqrt{\frac{10}{11}} > 3{,}1\,.$ (3 P.)

Aufgabe 5 (10 Punkte)

Sei $(X_t)_{t\in\mathbb{N}_0}$ eine endliche (zeit)homogene Markov-Kette mit diskreter Zeit über der Zustandsmenge $S = \{0,1,2,3,4\}$. Die positiven Übergangswahrscheinlichkeiten seien durch das folgende Übergangsdiagramm gegeben:



1. Geben Sie die Menge der transienten Zustände der Markov-Kette an.

2. Sei T_{01} die Übergangszeit vom Zustand 0 in den Zustand 1.
   Bestimmen Sie $\Pr[T_{01} = n]$ für alle $n \in \mathbb{N}$!

3. Berechnen Sie die Ankunftswahrscheinlichkeit f_{01}!

4. Berechnen Sie die erwartete Übergangszeit h_{14}!

Lösung

1. $\{0,1,2,3\}$. (1 P.)

2. $\Pr[T_{01} = 1] = \frac{1}{2}$, $\Pr[T_{01} = 3] = \frac{1}{2} \cdot \frac{3}{4} \cdot \frac{1}{4} = \frac{3}{32}$.
   Für alle übrigen n gilt $\Pr[T_{01} = n] = 0$. (3 P.)

3. Aus der Dichte von T_{01} oder mit Hilfe eines Gleichungssystems:

$$\begin{aligned} f_{01} &= p_{01} + p_{00}f_{01} + p_{02}f_{21} + p_{03}f_{31} + p_{04}f_{41} \\ &= \tfrac{1}{2} + \tfrac{1}{2}f_{21}\,, \\ f_{21} &= \tfrac{3}{4}f_{31}\,, \\ f_{31} &= \tfrac{1}{4}\,. \end{aligned}$$

   Es folgt $f_{21} = \frac{3}{16}$, $f_{01} = \frac{19}{32}$. (3 P.)

4. Mit Hilfe eines Gleichungssystem:

$$\begin{aligned} h_{14} &= 1 + p_{10}h_{04} + p_{11}h_{14} + p_{12}h_{24} + p_{13}h_{34} \\ &= 1 + h_{24}\,, \\ h_{24} &= 1 + \tfrac{3}{4}h_{34}\,, \\ h_{34} &= 1 + \tfrac{1}{4}h_{14}\,. \end{aligned}$$

   Es folgt $h_{34} = \frac{24}{13}$, $h_{24} = \frac{31}{13}$, $h_{14} = \frac{44}{13}$.

   (3 P.)