Technische Universität München
Institut für Informatik
Prof. Dr. Angelika Steger
Martin Marciniszyn
Alexander Offtermatt-Souza
Jan Remy

Wintersemester 2003/2004
Nachholklausur

Diskrete Strukturen II

Aufgabe 1. Was trifft zu? (1 + 1 + 1 + 1 + 2 = 6 Punkte)

Entscheiden Sie, welche der folgenden Aussagen zutreffen. Begründen Sie Ihre Antwort, indem Sie gegebenenfalls rechnen, die Vorlesung bzw. das Buch zitieren oder Gegenbeispiele angeben.

(a) Für zwei Zufallsvariablen X und Y gilt: $\mathrm{Var}[X+Y] = \mathrm{Var}[X] + \mathrm{Var}[Y]$

(b) Für jede normalverteilte Zufallsvariable X und alle $x, y \geq 0$ gilt:
$\Pr[X > x + y \mid X > y] = \Pr[X > x]$.

(c) Um zu beweisen, dass eine endliche Markov-Kette eine stationäre Verteilung hat, genügt es zuzeigen, dass sie irreduzibel und aperiodisch ist.

(d) Die Matrix

$$\begin{pmatrix} 0 & \nu_{12} & 0 \\ \nu_{21} & 0 & \nu_{23} \\ \nu_{31} & 0 & 0 \end{pmatrix} \qquad \nu_{12}, \nu_{21}, \nu_{23}, \nu_{31} > 0,$$

stellt die Übergangsraten einer Markovkette mit kontinuierlicher Zeit dar. Diese Kette hat eine nicht triviale stationäre Verteilung, die unabhängig vom Startzustand ist.

(e) Sind zwei Ereignisse A, B disjunkt mit $\Pr[A], \Pr[B] > 0$, dann sind A und B abhängig.

Aufgabe 2. Bedingte Wahrscheinlichkeit (5 Punkte)

Zu untersuchen ist ein medizinischer Test, der entweder ein positives Resultat T oder ein negatives $\overline{T}$ liefert. Er soll Aufschluss über das Vorliegen einer Erkrankung K geben. Es ist bekannt, dass ein falsch positives Ergebnis mit der Wahrscheinlichkeit $\Pr\left[T \mid \overline{K}\right]$ und ein falsch negatives mit $\Pr\left[\overline{T} \mid K\right]$ auftritt. Man kennt außerdem die so genannte Prävalenz der Erkrankung $\Pr[K]$, d.h. die Wahrscheinlichkeit, mit der eine zufällig aus der Bevölkerung ausgewählte Person die Erkrankung hat. Leiten Sie die Formel her, mit der sich aus den gegebenen Größen die Wahrscheinlichkeit für das tatsächliche Vorliegen der Erkrankung bei einem Patienten berechnen lässt, bei dem der Test positiv gewesen ist.

Aufgabe 3. Teilbarkeit (2 + 3 = 5 Punkte)

Aus der Menge $\{1, \ldots, 144\}$ wird eine Zahl k zufällig, gleichverteilt gezogen. Bestimmen Sie die Wahrscheinlichkeit, dass

(a) k *nicht* durch 3 teilbar ist.

(b) k weder durch 3, durch 4 noch durch 6 teilbar ist.

Aufgabe 4. Zufallsgraph (2 + 4 + 5 = 11 Punkte)

Sei G ein ungerichteter einfacher Graph auf der Knotenmenge $V = \{1, 2, \ldots, n\}$, wobei für jede der $\binom{n}{2}$ möglichen Kanten unabhängig mit Wahrscheinlichkeit p entschieden wird, ob sie in G existiert oder nicht.

(a) Bestimmen Sie die erwartete Anzahl von Dreiecken X in G in Abhängigkeit von der Kantenwahrscheinlichkeit p. (Tipp: Verwenden Sie Indikatorvariablen.)

(b) Sei $p = \frac{c}{n}$ für eine Konstante $c > 0$. Zeigen Sie $\Pr[X = 0] \geq 1 - \frac{c^3}{6}$.

(c) Sei $p = \frac{c}{n}$ für eine Konstante $c > 1$. Zeigen Sie $\Pr[X = 0] \leq (1 + \frac{9}{2} \cdot \frac{c^3}{n})/\mathbb{E}[X]$.
(Tipp: Verwenden Sie $\Pr[X = 0] \leq \Pr[|X - \mathbb{E}[X]| \geq \mathbb{E}[X]]$ und die Ungleichung von Chebyshev.)

Aufgabe 5. Markovketten (1 + 3 + 4 = 8 Punkte)

Die Matrix

$$\begin{pmatrix} a & 1-a & 0 \\ b & 0 & 1-b \\ 0 & 1 & 0 \end{pmatrix} \qquad\qquad 0 < a, b < 1,$$

beschreibe die Übergangswahrscheinlichkeiten einer diskreten Markovkette mit drei Zuständen.

(a) Zeichnen Sie das Zustandsübergangsdiagramm der Markovkette.

(b) Begründen Sie, warum die Markovkette eine eindeutige stationäre Verteilung besitzt, und bestimmen Sie diese.

(c) Bestimmen Sie die erwarteten Hitting Times h_{13} vom Zustand 1 nach 3 und h_{23} vom Zustand 2 nach 3.

Aufgabe 6. Datenkompression (3 + 2 = 5 Punkte)

Ein neues Kompressionsverfahren A_{neu} soll experimentell mit einem Standardverfahren A_{std} verglichen werden. Dazu werden 10 000 Worte gleicher Länge, die zufällig gleichverteilt und unabhängig erzeugt wurden, jeweils mit beiden Verfahren komprimiert, und die resultierenden Längen miteinander verglichen.

(a) Wählen Sie einen Test, mit dem sich auf einem vorgegebenen Signifikanzniveau α statistisch belegen lässt, dass A_{neu} bezüglich der Wortlänge besser ist als A_{std}. Definieren Sie die Zufallsvariablen und stellen Sie die Testgröße auf.

(b) A_{neu} schlägt A_{std} in 5 124 Fällen. Lässt sich daraus bei einem Niveau $\alpha = 5\%$ auf eine signifikante Überlegenheit von A_{neu} schließen? (Quantile der Standardnormalverteilung: $z_{0,95} = 1,6449$, $z_{0,975} = 1,9600$)

Viel Erfolg!