Klausur 2016, Fragen und Antworten

Diskrete Wahrscheinlichkeitstheorie (IN0018) (Technische Universität München)

Technische Universität München
Fakultät für Informatik
Lehrstuhl für Theoretische Informatik
Prof. Dr. Susanne Albers
M. Gottschau, D. Kraft, S. Schraink, R. Stotz

Sommersemester 2016
Lösungen der Klausur
3. August 2016

Diskrete Wahrscheinlichkeitstheorie

Name	Vorname	Matrikelnummer	
....	....	....	
Hörsaal	**Reihe**	**Sitzplatz**	**Unterschrift**
....	....	....	....

Allgemeine Hinweise zur Klausur

- Die Angabe der Klausur umfasst sechs Aufgaben und eine Tabelle der Standardnormalverteilung.
- Bitte schreiben Sie weder mit Bleistift, noch mit roter oder grüner Farbe.
- Außer Ihren Schreibutensilien und einem handgeschriebenen DIN-A4-Blatt sind keine Hilfsmittel erlaubt.
- Die Bearbeitungszeit beträgt 120 Minuten.

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

	A1	A2	A3	A4	A5	A6	Σ	Korrektor
Gesamt	6	7	5	7	8	7	40	
Erstkorrektur								
Zweitkorrektur								

Aufgabe 1 (1+1+2+2 Punkte)

Zur Vorbereitung auf die DWT-Klausur teilen sich die Studentinnen Anna, Bea und Carola zusammen mit ihren Kommilitonen David, Emil und Felix zufällig in 3 disjunkte Lerngruppen aus jeweils 2 Personen auf. Hierbei sind alle Aufteilungen gleich wahrscheinlich.

1. Zeigen Sie, dass es 15 Möglichkeiten gibt, die 6 Personen in Zweiergruppen zu partitionieren.
2. Bestimmen Sie die Wahrscheinlichkeit, mit der alle Lerngruppen gemischt sind, d.h. jede Lerngruppe aus einer Studentin und einem Studenten besteht.
3. Mit welcher Wahrscheinlichkeit sind alle Lerngruppen gemischt, falls Anna und David eine gemeinsame Lerngruppe bilden.
4. Zeigen oder widerlegen Sie die folgende Aussage: Die Vereinigung der beiden Ereignisse „Anna und David bilden eine Lerngruppe“ bzw. „Bea und Carola bilden eine Lerngruppe“ ist unabhängig vom Ereignis, dass alle Lerngruppen gemischt sind.

Lösungsvorschlag

1. Als Ergebnismenge Ω wählen wir alle Partitionen der 6 Studenten in Mengen der Kardinalität 2. Dafür nehmen wir vorerst an, dass die Lerngruppen unterscheidbar sind. Folglich gibt es $\binom{6}{2}$ Kombinationen aus 2 Studenten für die erste Lerngruppe, $\binom{4}{2}$ für die zweite Lerngruppe und $\binom{2}{2}$ für die dritte Lerngruppe. Um die Reihenfolge unter den Lerngruppen aufzuheben, teilen wir durch 3! und erhalten

$$|\Omega| = \frac{\binom{6}{2} \cdot \binom{4}{2} \cdot \binom{2}{2}}{3!} = \frac{90}{6} = 15.$$

2. Sei E das Ereignis, dass alle Lerngruppen gemischt sind. Gesucht ist die Wahrscheinlichkeit $\Pr[E]$. Da wir einen Laplace-Wahrscheinlichkeitsraum betrachten gilt $\Pr[E] = |E|/|\Omega|$. Man beachte, dass jedes Element in E einer bijektiven Abbildung von Jungen auf Mädchen entspricht. Somit gilt $|E| = 3!$ und es folgt, dass

$$\Pr[E] = \frac{3!}{15} = \frac{2}{5} = 0{,}4.$$

3. Bezeichne F das Ereignis, dass Anna und David in einer Lerngruppe sind. Gesucht ist die bedingte Wahrscheinlichkeit $\Pr[E \mid F]$. Sollten Anna und David bereits in einer Lerngruppe sein, so gibt es noch

$$|F| = \frac{\binom{4}{2} \cdot \binom{2}{2}}{2!} = \frac{6}{2} = 3$$

mögliche Aufteilungen für die verbliebenen Studenten. Außerdem sind für 2 der 3 Elemente aus F alle Lerngruppen gemischt, nämlich für $\{\{A, D\}, \{B, E\}, \{C, F\}\}$ und $\{\{A, D\}, \{B, F\}, \{C, E\}\}$. Daher gilt $|E \cap F| = 2$ und wir erhalten

$$\Pr[E \mid F] = \frac{\Pr[E \cap F]}{\Pr[F]} = \frac{|E \cap F|}{|F|} = \frac{2}{3} \approx 0{,}66667.$$

4. Sei G das Ereignis, dass Bea und Carola in einer Lerngruppe sind. Wir möchten wissen, ob die Ereignisse E und $F \cup G$ unabhängig sind. Aus Gründen der Symmetrie gilt $\Pr[F] = \Pr[G] = 3/15$. Außerdem ist die dritte Lerngruppe eindeutig bestimmt, falls sich sowohl Anna und David als auch Bea und Carola zu einer Lerngruppe zusammengefunden haben. Daher gilt $\Pr[F \cap G] = 1/15$. Gemäß der Siebformel ist die Wahrscheinlichkeit von $F \cup G$ gegeben durch

$$\Pr[F \cup G] = \Pr[F] + \Pr[G] - \Pr[F \cap G] = \frac{3}{15} + \frac{3}{15} - \frac{1}{15} = \frac{1}{3} \approx 0{,}33333.$$

Da die Ereignisse E und G nicht gleichzeitig eintreten können, gilt außerdem

$$\Pr[E \cap (F \cup G)] = \Pr[(E \cap F) \cup (E \cap G)] = \Pr[(E \cap F) \cup \emptyset] = \Pr[E \cap F] = \frac{2}{15} \approx 0{,}13333.$$

Insgesamt folgt, dass

$$\Pr[E \cap (F \cup G)] = \frac{2}{15} = \frac{2}{5} \cdot \frac{1}{3} = \Pr[E] \cdot \Pr[F \cup G],$$

womit wir die Unabhängigkeit von E und $F \cup G$ gezeigt haben.

Aufgabe 2 (1+3+3 Punkte)

Bei einem Marathon dürfen Teams bestehend aus 2 oder 4 Läufern antreten. Allerdings erreicht jeder Läufer unabhängig von den anderen Teilnehmern mit Wahrscheinlichkeit $0 \leq p \leq 1$ nicht das Ziel. Sollten echt weniger als die Hälfte der Läufer eines Teams im Ziel ankommen, so ist das Team disqualifiziert.

1. Bestimmen Sie in Abhängigkeit von p, mit welcher Wahrscheinlichkeit ein Zweierteam bzw. Viererteam disqualifiziert wird.

2. Für welche Werte von p ist die Wahrscheinlichkeit einer Disqualifikation für ein Zweierteam echt kleiner als für ein Viererteam.

3. Nach dem Lauf vermutet die Sportreporterin Anja, dass $p \leq 2/100$ gilt. Um diese Aussage zu überprüfen, ermittelt sie, wie viele der insgesamt 10000 Läufer das Ziel erreicht haben. Führen Sie einen geeigneten Hypothesentest durch und bestimmen Sie den größtmöglichen Ablehnungsbereich, mit dem Anjas Behauptung bei einem Signifikanzniveau von 0,05 abgelehnt werden kann. Nutzen Sie hierfür auch die Tabelle der Standardnormalverteilung auf dem letzten Blatt der Angabe.

Lösungsvorschlag

1. Seien X und Y Zufallsvariablen, die zählen, wie viele Läufer in einem Zweierteam bzw. Viererteam das Ziel erreichen. Aufgrund der Unabhängigkeit zwischen den Läufern sind X und Y mit den Parametern p und 2 bzw. p und 4 binomialverteilt. Daher gilt
$$\Pr[X < 2/2] = \Pr[X = 0] = \binom{2}{0} \cdot p^2 \cdot (1-p)^0 = p^2$$
und
$$\Pr[Y < 4/2] = \Pr[Y \leq 1] = \binom{4}{0} \cdot p^4 \cdot (1-p)^0 + \binom{4}{1} \cdot p^3 \cdot (1-p) = -3p^4 + 4p^3.$$

2. Gesucht sind alle Werte für p, sodass $\Pr[X < 2/2] < \Pr[Y < 4/2]$. Durch Umstellen und Einsetzen erhalten wir
$$0 < \Pr[Y < 4/2] - \Pr[X < 2/2] = -3p^4 + 4p^3 - p^2.$$
Wir müssen also die Werte von p zwischen 0 und 1 bestimmen, für die das Polynom $-3p^4+4p^3-p^2$ positiv ist. Offensichtlich hat $-3p^4+4p^3-p^2$ eine doppelte Nullstelle bei $p = 0$. Faktorisieren wir entsprechend, so erhalten wir den Term $p^2 \cdot (3p^2-4p+1)$. Das Polynom $3p^2 - 4p + 1$ hat wiederum Nullstellen an den Punkten
$$p = \frac{4 + \sqrt{16 - 4 \cdot 3 \cdot 1}}{6} = 1$$
und
$$p = \frac{4 - \sqrt{16 - 4 \cdot 3 \cdot 1}}{6} = \frac{1}{3}.$$
Da $-3p^4+4p^3-p^2$ für $p \to \infty$ gegen $-\infty$ geht, kann $-3p^4+4p^3-p^2$ ausschließlich im Intervall $(1/3, 1)$ positiv sein. Somit ist die Wahrscheinlichkeit einer Disqualifikation für Zweierteams genau dann echt kleiner als für Viererteams, wenn $1/3 < p < 1$ gilt.

3. Um Anjas Nullhypothese zu testen, nämlich $H_0 : p_0 \leq 0{,}02$, führen wir einen approximativen Binomialtest durch. Dies ist möglich, da die Anzahl der Läufer, die nicht ins Ziel kommen, binomialverteilt ist und wir bei 10000 Läufern davon ausgehen können, dass die Anzahl der Stichproben hinreichend groß ist, um mit der Normalverteilung zu approximieren. Die Testgröße lautet

$$Z = \frac{h - 10000 \cdot 0{,}02}{\sqrt{10000 \cdot 0{,}02 \cdot (1 - 0{,}02)}} = \frac{h - 200}{14},$$

wobei h die Anzahl der Läufer angibt, die es nicht ins Ziel geschafft haben. Die Nullhypothese kann mit einem Signifikanzniveau von 0,05 abgelehnt werden, wenn $Z > z_{1-0{,}05}$ gilt. Wir müssen h also so wählen, dass

$$h > 200 + 14 z_{0{,}95} \approx 200 + 14 \cdot 1{,}65 = 223{,}1.$$

Das 0,95-Quantil der Standardnormalverteilung haben wir hierbei der Tabelle aus dem Anhang entnommen. Bei einem Signifikanzniveau von 0,05 erhalten wir demnach den größtmöglichen Ablehnungsbereich, wenn wir Anjas Hypothese für $h \geq 224$ ablehnen.

Aufgabe 3 (2+3 Punkte)

Der Waschbär Andreas ernährt sich am liebsten von Kirschen. Pro Tag erbeutet Andreas eine zufällige Anzahl X der köstlichen Früchte. Die wahrscheinlichkeitserzeugende Funktion von X ist definiert als:

$$G_X(s) = \exp\bigl(5 \cdot (s^2 - 1)\bigr).$$

1. Zeigen Sie, dass der Erwartungswert und die Varianz von X gegeben sind durch $\mathbb{E}[X] = 10$ und $\mathrm{Var}[X] = 20$.

2. Nutzen Sie die Chebyshev-Ungleichung um zu zeigen, dass die Wahrscheinlichkeit, mit der Andreas mindestens 20 Kirschen an einem Tag findet, höchstens 1/5 ist.

Lösungsvorschlag

1. Der Erwartungswert von X ist gegeben durch

   $$\mathbb{E}[X] = G'_X(1) = \exp\bigl(5 \cdot (1^2 - 1)\bigr) \cdot (10 \cdot 1) = 10.$$

   Außerdem berechnet sich die zweite Ableitung der Funktion G_X an der Stelle 1 zu

   $$G''_X(1) = \exp\bigl(5 \cdot (1^2 - 1)\bigr) \cdot (10 \cdot 1)^2 + \exp\bigl(5 \cdot (1^2 - 1)\bigr) \cdot 10 = 110.$$

   Die Varianz von X lautet somit

   $$\mathrm{Var}[X] = G''_X(1) + G'_X(1) - \bigl(G'_X(1)\bigr)^2 = 110 + 10 - 10^2 = 20.$$

2. Die Wahrscheinlichkeit, mit der Andreas mindestens 20 Kirschen an einem Tag findet, lässt sich abschätzen mit

   $$\Pr[X \geq 20] = \Pr[X - \mathbb{E}[X] \geq 10] \leq \Pr[|X - \mathbb{E}[X]| \geq 10].$$

   Die obere Ungleichung ist gültig, da $X - \mathbb{E}[X] \geq 10$ das Ereignis $|X - \mathbb{E}[X]| \geq 10$ impliziert. Gemäß der Chebyshev-Ungleichung können wir weiter abschätzen zu

   $$\Pr[|X - \mathbb{E}[X]| \geq 10] \leq \frac{\mathrm{Var}[X]}{10^2} = \frac{20}{100} = \frac{1}{5}.$$

   Insgesamt haben wir somit gezeigt, dass $\Pr[X \geq 20] \leq 1/5$ gilt.

Aufgabe 4 (3+3+1 Punkte)

Der unpünktliche Student Adam hat sich mit dem noch unpünktlicheren Studenten Ben an der Uni verabredet. Sei X eine Zufallsvariable, die Adams Verspätung in Stunden angibt. Ferner sei Y Bens Verspätung in Stunden. Die gemeinsame Dichte der beiden Zufallsvariablen ist beschrieben durch:

$$f_{X,Y}(x,y) = \begin{cases} (e^{-x+y})/(e-c) & \text{falls } x,y \in [0,1] \text{ und } x \leq y \\ 0 & \text{sonst} \end{cases}.$$

1. Finden Sie einen Wert für c, sodass $f_{X,Y}$ eine zulässige Dichtefunktion ist.
2. Stellen Sie die Randdichten f_X und f_Y auf.
3. Untersuchen Sie die Zufallsvariablen X und Y auf Unabhängigkeit.

Lösungsvorschlag

1. Damit es sich bei der gegebenen Funktion um eine zulässige Dichte handelt, muss $f_{X,Y}(x,y) \geq 0$ sowie $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(s,t)\,ds\,dt = 1$ gelten. Die erste Bedingung ist für $c < e$ trivialerweise erfüllt. Für die zweite Bedingung betrachten wir das Integral

   $$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(s,t)\,ds\,dt = \int_0^1\int_0^t \frac{e^{-s+t}}{e-c}\,ds\,dt = \frac{1}{e-c}\cdot\int_0^1\int_0^t e^{-s+t}\,ds\,dt.$$

   Die Stammfunktion von e^{-s+t} bezüglich s lautet $-e^{-s+t}$. Wir können somit das innere Integral auflösen zu

   $$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(s,t)\,ds\,dt = \frac{1}{e-c}\cdot\int_0^1 \left[-e^{-s+t}\right]_0^t dt = \frac{1}{e-c}\cdot\int_0^1 -1+e^t\,dt.$$

   Die Stammfunktion von $e^t - 1$ bezüglich t ist wiederum gegeben durch $e^t - t$. Letztendlich erhalten wir

   $$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(s,t)\,ds\,dt = \frac{1}{e-c}\cdot\left[e^t - t\right]_0^1 = \frac{(e-1)-(1-0)}{e-c} = \frac{e-2}{e-c}.$$

   Damit dieser Term 1 wird, wählen wir $c = 2$. Nachdem außerdem $2 < e$ gilt, ist $f_{X,Y}$ für $c = 2$ eine zulässige Dichte.

2. Im Fall $x < 0$ bzw. $x > 1$ ist die Randdichte von X trivialerweise 0. Für alle anderen Werte von x, das heißt $0 \leq x \leq 1$ ist f_X gegeben durch

   $$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,t)\,dt = \int_x^1 \frac{e^{-x+t}}{e-2}\,dt = \frac{1}{e-2}\cdot\int_x^1 e^{-x+t}\,dt.$$

   Die Stammfunktion von e^{-x+t} bezüglich t lautet e^{-x+t} und wir können weiter auflösen zu

   $$f_X(x) = \frac{1}{e-2}\cdot\left[e^{-x+t}\right]_x^1 = \frac{e^{-x+1}-e^{-x+x}}{e-2} = \frac{e^{1-x}-1}{e-2}.$$

Ähnlich berechnet sich die Randdichte von Y im Intervall $0 \leq y \leq 1$ zu

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(s,y)\,ds = \int_0^y \frac{e^{-s+y}}{e-2}\,ds = \frac{1}{e-2} \cdot \int_0^y e^{-s+y}\,ds.$$

Die Stammfunktion von e^{-s+y} bezüglich s ist uns bereits aus der ersten Teilaufgabe bekannt. Sie lautet $-e^{-s+y}$ und es folgt

$$f_Y(y) = \frac{1}{e-2} \cdot \left[-e^{-s+y}\right]_0^y = \frac{-e^{-y+y} + e^{-0+y}}{e-2} = \frac{e^y - 1}{e-2}.$$

Insgesamt erhalten wir die Randdichten

$$f_X(x) = \begin{cases} (e^{1-x} - 1)/(e-2) & \text{falls } 0 \leq x \leq 1 \\ 0 & \text{sonst} \end{cases}$$

und

$$f_Y(y) = \begin{cases} (e^{y} - 1)/(e-2) & \text{falls } 0 \leq y \leq 1 \\ 0 & \text{sonst} \end{cases}.$$

3. Damit zwei Zufallsvariablen unabhängig sind, muss die gemeinsame Dichte das Produkt der Randdichten sein. Für X und Y ist dies offensichtlich nicht der Fall, da bspw. $f_X(3/4) > 0$ und $f_Y(1/4) > 0$ gilt obwohl $f_{X,Y}(3/4, 1/4) = 0$. Die gemeinsame Dichte ist in diesem Fall 0 weil $1/4 > 3/4$.

Aufgabe 5 (4+2+2 Punkte)

Ada bricht einen Schokoriegel der Länge ℓ an einer zufälligen Stelle X in zwei Teile, wobei X auf dem Intervall $[0, \ell]$ gleichverteilt ist. Das größere Stück behält Ada für sich. Das kleinere Stück, dessen Länge durch die Zufallsvariable Y beschrieben ist, gibt sie ihrer Schwester Bianca.

1. Zeigen Sie, dass die Verteilungsfunktion von Y gegeben ist durch:

$$F_Y(y) = \begin{cases} 0 & \text{falls } y < 0 \\ 2y/\ell & \text{falls } 0 \leq y \leq \ell/2 \\ 1 & \text{falls } \ell/2 < y \end{cases}.$$

2. Bestimmen Sie die Dichtefunktion und den Erwartungswert von Y.

3. Bianca möchte die ihr unbekannte Länge des Schokoriegels anhand der Länge ihres Stücks abschätzen. Konstruieren Sie mit Hilfe von Y einen erwartungstreuen Schätzer und einen Maximum-Likelihood-Schätzer für ℓ.

Lösungsvorschlag

1. Um die Verteilungsfunktion von Y zu bestimmen betrachten wir zunächst die Wahrscheinlichkeit $\Pr[Y \leq y]$. In den Fällen $y \leq 0$ bzw. $\ell/2 \leq y$ gilt trivialerweise $\Pr[Y \leq y] = 0$ bzw. $\Pr[Y \leq y] = 1$. Sei daher $0 \leq y \leq \ell/2$.

   Wir machen nunmehr eine Fallunterscheidung bezüglich X. Für $X \leq \ell/2$ gilt $Y = X$. Anderenfalls, für $X > \ell/2$, gilt $Y = \ell - X$. Da beide Fälle disjunkt sind können wir die Wahrscheinlichkeit $\Pr[Y \leq y]$ umformen zu

$$\begin{aligned} \Pr[Y \leq y] &= \Pr[Y \leq y, X \leq \ell/2] + \Pr[Y \leq y, X > \ell/2] \\ &= \Pr[X \leq y, X \leq \ell/2] + \Pr[n - X \leq y, X > \ell/2]. \end{aligned}$$

   Aus $0 \leq y \leq \ell/2$ schließen wir einerseits, dass

$$\Pr[X \leq y, X \leq \ell/2] = \Pr[X \leq y] = \frac{y}{\ell},$$

   und andererseits, dass

$$\Pr[\ell - X \leq y, X > \ell/2] = \Pr[X > \ell - y, X > \ell/2] = \Pr[X > \ell - y] = 1 - \frac{\ell - y}{\ell}.$$

   Insgesamt erhalten wir demnach eine Wahrscheinlichkeit von

$$\Pr[Y \leq y] = \frac{y}{\ell} + 1 - \frac{\ell - y}{\ell} = \frac{2y}{\ell}$$

   und die Verteilungsfunktion von Y lautet

$$F_Y(y) = \begin{cases} 0 & \text{falls } y < 0 \\ 2y/\ell & \text{falls } 0 \leq y \leq \ell/2 \\ 1 & \text{falls } \ell/2 < y \end{cases}.$$

2. Um die Dichtefunktion von Y zu bestimmen, leiten wir F_Y nach y ab und setzen die Funktion an den nicht differenzierbaren Stellen $y = 0$ und $y = \ell/2$ links bzw. rechtsseitig stetig fort

$$f_Y(y) = F_Y'(y) = \begin{cases} 2/\ell & \text{falls } 0 \leq y \leq \ell/2 \\ 0 & \text{sonst} \end{cases}.$$

An der Dichtefunktion erkennen wir, dass Y auf dem Intervall $[0, \ell/2]$ gleichverteilt ist. Der Erwartungswert von Y ist demnach

$$\mathbb{E}[Y] = \frac{0 + \ell/2}{2} = \frac{\ell}{4}.$$

3. Aus dem Erwartungswert von Y können wir unmittelbar einen erwartungstreuen Schätzer für ℓ bestimmen. Als Schätzer wählen wir $4Y$,

$$\mathbb{E}[4Y] = 4 \cdot \mathbb{E}[Y] = 4 \cdot \frac{\ell}{4} = \ell.$$

Für den Maximum-Likelihood-Schätzer gehen wir davon aus, dass $y > 0$ die konkrete Länge von Biancas Stück des Schokoriegels ist. Gemäß der Dichtefunktion aus der ersten Teilaufgabe ist die Likelihood-Funktion von ℓ gegeben durch

$$L(y; n) = \begin{cases} 2/\ell & \text{falls } \ell \geq 2y \\ 0 & \text{sonst} \end{cases}.$$

Da $2/\ell$ auf dem Intervall $(0, \infty)$ streng monoton fällt, ist die Likelihood-Funktion für $\ell = 2y$ maximal. Somit ist $2Y$ ein Maximum-Likelihood-Schätzer für ℓ.

Aufgabe 6 (2+3+2 Punkte)

Der Wetterfrosch Alice hat festgestellt, dass auf einen Regentag mit Wahrscheinlichkeit 1/2 ein Sonnentag und sonst ein Regentag folgt. Auf einen Sonnentag folgt hingegen mit Wahrscheinlichkeit 3/4 ein Sonnentag und sonst ein Regentag. Sei X_t eine Zufallsvariable, sodass $X_t = 1$ gilt, falls der t-te Tag sonnig ist. Ist der t-te Tag hingegen regnerisch, so gilt $X_t = 0$.

1. Modellieren Sie die Markov-Kette $\{X_t\}_{t\geq 0}$. Skizzieren Sie insbesondere den Übergangsgraphen und geben Sie die Übergangsmatrix an.
2. Angenommen am 0-ten Tag ist es sonnig. Beweisen Sie durch Induktion über t, dass

$$q_t = \left(\frac{1-4^{-t}}{3}, \frac{2+4^{-t}}{3}\right).$$

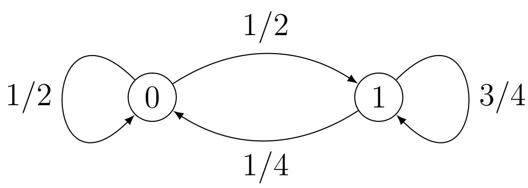
3. Bestimmen Sie die stationären Verteilungen der Markov-Kette $\{X_t\}_{t\geq 0}$.

Lösungsvorschlag

1. Da die Wahrscheinlichkeit $\Pr[X_t = 0]$ bzw. $\Pr[X_t = 1]$ nur vom vorherigen Tag abhängen, können wir das Wetter am t-ten Tag durch eine Markov-Kette $\{X_t\}_{t\geq 0}$ über der Zustandsmenge $S = \{0, 1\}$ modellieren. Mit den Übergangswahrscheinlichkeiten aus der Angabe erhalten wir als Übergangsmatrix

$$P = \begin{pmatrix} 1/2 & 1/2 \\ 1/4 & 3/4 \end{pmatrix}.$$

Graphisch lässt sich die Markov-Kette durch den folgenden Übergangsgraphen darstellen:



2. Wie gefordert beweisen wir die Identität aus der Angabe mittels einer Induktion über t. Wir beginnen mit der Induktionsbasis $t = 0$. Laut Angabe ist der heutige Tag sonnig. Die Startverteilung q_0 ist daher gegeben durch

$$q_0 = (0, 1) = \left(\frac{1-4^{-0}}{3}, \frac{2+4^{-0}}{3}\right).$$

Für den Induktionsschritt sei $t > 0$ und wir nehmen an, dass

$$q_{t-1} = \left(\frac{1-4^{-(t-1)}}{3}, \frac{2+4^{-(t-1)}}{3}\right)$$

bereits bewiesen wurde. Wir möchten nun zeigen, dass diese Identität auch für t gilt. Aus der Definition von q_t folgt

$$q_t = q_{t-1} \cdot P = q_{t-1} \cdot \begin{pmatrix} 1/2 & 1/2 \\ 1/4 & 3/4 \end{pmatrix}.$$

Setzen wir nun die Induktionshypothese ein, so erhalten wir

$$q_t = \Big(\frac{1 - 4^{-(t-1)}}{3}, \frac{2 + 4^{-(t-1)}}{3}\Big) \cdot \begin{pmatrix} 1/2 & 1/2 \\ 1/4 & 3/4 \end{pmatrix}.$$

Für $(q_t)_0$ ergibt sich hieraus ein Wert von

$$(q_t)_0 = \frac{1 - 4^{-(t-1)}}{3} \cdot \frac{1}{2} + \frac{2 + 4^{-(t-1)}}{3} \cdot \frac{1}{4} = \frac{4 - 4^{-(t-1)}}{12} = \frac{1 - 4^{-t}}{3}.$$

Ähnlich berechnet sich $(q_t)_1$ zu

$$(q_t)_1 = \frac{1 - 4^{-(t-1)}}{3} \cdot \frac{1}{2} + \frac{2 + 4^{-(t-1)}}{3} \cdot \frac{3}{4} = \frac{8 + 4^{-(t-1)}}{12} = \frac{2 + 4^{-t}}{3}.$$

Insgesamt haben wir somit gezeigt, dass

$$q_t = \Big(\frac{1 - 4^{-t}}{3}, \frac{2 + 4^{-t}}{3}\Big).$$

3. Da das Übergangsdiagramm der Markov-Kette stark zusammenhängend ist, handelt es sich um eine irreduzible Markov-Kette. Sie ist zudem aperiodisch, da jeder Zustand eine Schleife auf sich selbst hat. Insgesamt ist die Markov-Kette ergodisch und der Fundamental Satz für ergodische Markov-Ketten kann angewendet werden. Dieser besagt, dass unabhängig der Startverteilung $\lim_{t\to\infty} q_t = \pi$ gilt, wobei π die eindeutige stationäre Verteilung der Markov-Kette ist. Aus der zweiten Teilaufgabe schlussfolgern wir, dass $\{X_t\}_{t\geq 0}$ genau eine stationäre Verteilung besitzt, nämlich

$$\lim_{t\to\infty} \Big(\frac{1 - 4^{-t}}{3}, \frac{2 + 4^{-t}}{3}\Big) = \Big(\frac{1}{3}, \frac{2}{3}\Big).$$

Diese Tabelle der Standardnormalverteilung enthält die Werte von $\Phi(x)$ für $0 \leq x \leq 2{,}99$. Beispielsweise gilt $\Phi(1{,}55) \approx 0{,}939$.

$\Phi(x)$	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,500	0,504	0,508	0,512	0,516	0,520	0,524	0,528	0,532	0,536
0,1	0,540	0,544	0,548	0,552	0,556	0,560	0,564	0,567	0,571	0,575
0,2	0,579	0,583	0,587	0,591	0,595	0,599	0,603	0,606	0,610	0,614
0,3	0,618	0,622	0,626	0,629	0,633	0,637	0,641	0,644	0,648	0,652
0,4	0,655	0,659	0,663	0,666	0,670	0,674	0,677	0,681	0,684	0,688
0,5	0,691	0,695	0,698	0,702	0,705	0,709	0,712	0,716	0,719	0,722
0,6	0,726	0,729	0,732	0,736	0,739	0,742	0,745	0,749	0,752	0,755
0,7	0,758	0,761	0,764	0,767	0,770	0,773	0,776	0,779	0,782	0,785
0,8	0,788	0,791	0,794	0,797	0,800	0,802	0,805	0,808	0,811	0,813
0,9	0,816	0,819	0,821	0,824	0,826	0,829	0,831	0,834	0,836	0,839
1,0	0,841	0,844	0,846	0,848	0,851	0,853	0,855	0,858	0,860	0,862
1,1	0,864	0,867	0,869	0,871	0,873	0,875	0,877	0,879	0,881	0,883
1,2	0,885	0,887	0,889	0,891	0,893	0,894	0,896	0,898	0,900	0,901
1,3	0,903	0,905	0,907	0,908	0,910	0,911	0,913	0,915	0,916	0,918
1,4	0,919	0,921	0,922	0,924	0,925	0,926	0,928	0,929	0,931	0,932
1,5	0,933	0,934	0,936	0,937	0,938	0,939	0,941	0,942	0,943	0,944
1,6	0,945	0,946	0,947	0,948	0,949	0,951	0,952	0,953	0,954	0,954
1,7	0,955	0,956	0,957	0,958	0,959	0,960	0,961	0,962	0,962	0,963
1,8	0,964	0,965	0,966	0,966	0,967	0,968	0,969	0,969	0,970	0,971
1,9	0,971	0,972	0,973	0,973	0,974	0,974	0,975	0,976	0,976	0,977
2,0	0,977	0,978	0,978	0,979	0,979	0,980	0,980	0,981	0,981	0,982
2,1	0,982	0,983	0,983	0,983	0,984	0,984	0,985	0,985	0,985	0,986
2,2	0,986	0,986	0,987	0,987	0,987	0,988	0,988	0,988	0,989	0,989
2,3	0,989	0,990	0,990	0,990	0,990	0,991	0,991	0,991	0,991	0,992
2,4	0,992	0,992	0,992	0,992	0,993	0,993	0,993	0,993	0,993	0,994
2,5	0,994	0,994	0,994	0,994	0,994	0,995	0,995	0,995	0,995	0,995
2,6	0,995	0,995	0,996	0,996	0,996	0,996	0,996	0,996	0,996	0,996
2,7	0,997	0,997	0,997	0,997	0,997	0,997	0,997	0,997	0,997	0,997
2,8	0,997	0,998	0,998	0,998	0,998	0,998	0,998	0,998	0,998	0,998
2,9	0,998	0,998	0,998	0,998	0,998	0,998	0,998	0,999	0,999	0,999