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

Sommersemester 2010
Lösungsblatt 11
15. Juli 2010

Diskrete Wahrscheinlichkeitstheorie

Hausaufgabe 1 (5 Punkte)

Beweisen oder widerlegen Sie die folgenden Behauptungen.

1. Jede erwartungstreue Schätzvariable für einen Parameter ϑ schätzt den Erwartungswert von ϑ.
2. Bei echten Alternativtests ist die Summe der Wahrscheinlichkeit eines Fehlers 1.Art und der Wahrscheinlichkeit eines Fehlers 2. Art stets gleich 1.
3. Die Anwendung des Maximum-Likelihood-Prinzips setzt die Verfügbarkeit einer Stichprobe voraus.
4. Zu jeder stetigen Zufallsvariablen X mit Verteilung F_X gibt es ein 1-Quantil aus $\mathbb{R}$ von F_X.
5. Liegt beim Verfahren zum Test von Hypothesen H_0 der Wert der Testvariablen im Ablehnungsbereich, dann gibt die Fehlerwahrscheinlichkeit erster Art die Wahrscheinlichkeit an, mit der die Hypothese H_0 zutrifft.

Lösungsvorschlag

1. Falsch. Begründung: Parameter sind i.A. selbst keine Zufallsvariablen.
2. Falsch. Begründung: Echte Alternativtest werden mit dem Ziel konstruiert, beiden Fehlerarten kleine Wahrscheinlichkeit zu geben.
3. Wahr. Begründung: Es wird die Wahrscheinlichkeit $L(\vec{x}; 0)$ maximiert mit einer Stichprobe $\vec{x}$.
4. Falsch. Begründung: Das Quantil könnte auch ∞ sein. ∞ ist aber nicht aus $\mathbb{R}$.
5. Falsch. Begründung: Die Fehlerwahrscheinlichkeit bezeichnet die <u>maximale</u> Wahrscheinlichkeit eines Fehlers erster Art und ist damit eine Abschätzung von Wahrscheinlichkeiten, und zwar für den gesamten Parameterbereich der Nullhypothese.

Hausaufgabe 2 (5 Punkte)

1. Wir betrachten eine erwartungstreue, diskrete Schätzvariable X für einen Parameter α. Für die Verteilungsdichte f_X gelte $f_X(i) = \frac{e^{-3}(3^i)}{i!}$ für $i \in \mathbb{N}_0$.

   Mit welchem Wert wird α durch X geschätzt?

2. Für die Dichte f_X einer diskreten Schätzvariablen X für einen Parameter α gelte $f_X(i) = \binom{10}{i}(\frac{2}{9})^i(\frac{7}{9})^{10-i}$ und $W_X = \{0, 1, \ldots, 10\}$. Mit welchem Wert wird α durch X geschätzt, wenn der Bias $-\frac{2}{9}$ beträgt?

Lösungsvorschlag

1. X ist Poisson-verteilt mit $\lambda = 3$. Damit gilt $\mathbb{E}[X] = 3$, mithin

$$\alpha = \mathbb{E}[X] = 3\,.$$

2. Es gilt $\mathbb{E}[X - \alpha] = -\frac{2}{9}$. Da $X \sim \mathrm{Bin}(10, \frac{2}{9})$, folgen $\mathbb{E}[X] = \frac{20}{9}$ und

$$\alpha = -\mathbb{E}[X - \alpha] + \mathbb{E}[X] = \frac{2}{9} + \frac{20}{9} = \frac{22}{9}\,.$$

Hausaufgabe 3 (5 Punkte)

Sei $T = X_1 + X_2 + X_3$ eine Zufallsvariable, wobei X_1, X_2 und X_3 unabhängige Bernoulli-verteilte Zufallsvariablen mit gleicher Erfolgswahrscheinlichkeit p seien. Wir betrachten T als Stichprobenvariable zum Test der Hypothese $H_0 : p \geq \frac{1}{3}$ mit Ablehnungsbereich $\tilde{K} = \{0, 1, 2\}$.

1. Berechnen Sie die Verteilungsfunktion F_T von T!

2. Berechnen Sie die Fehlerwahrscheinlichkeit 1. Art!

3. Berechnen Sie die Fehlerwahrscheinlichkeit 2. Art unter der Annahme, dass $H_1 : p \leq \frac{1}{4}$ eine echte Alternative zu H_0 sei!

Lösungsvorschlag

1. T ist binomialverteilt mit

$$f_T(i) = \binom{3}{i}p^i(1-p)^{3-i} \text{ und } W_T = \{0, 1, 2, 3\}.$$

Allgemein gilt

$$\begin{aligned}
F_T(i) &= \Pr[T \leq i].\\
F_T(0) &= (1-p)^3,\\
F_T(1) &= (1-p)^3 + 3p(1-p)^2\\
&= 1 - 3p^2 + 2p^3,\\
F_T(2) &= 1 - p^3,\\
F_T(3) &= 1.
\end{aligned}$$

2. $$\begin{aligned}\alpha_1 &= \max_{p\in H_0} \Pr[T \in \tilde{K}] \\ &= \max_{p\geq \frac{1}{3}} F_T(2) \\ &= \max_{p\geq\frac{1}{3}} (1-p^3) \\ &= 1 - \left(\tfrac{1}{3}\right)^3 = \tfrac{26}{27}.\end{aligned}$$

3. $$\begin{aligned}\alpha_2 &= \max_{p\in H_1} \Pr[T \notin \tilde{K}] \\ &= \max_{p\leq \frac{1}{4}} f_T(3) \\ &= \max_{p\leq\frac{1}{4}} p^3 \\ &= \left(\tfrac{1}{4}\right)^3 = \tfrac{1}{64}.\end{aligned}$$

Hausaufgabe 4 (5 Punkte)

Sei X eine Zufallsvariable mit unbekanntem Erwartungswert μ und einer Standardabweichung $\sigma_X = 2$. Wir verwenden $\bar{X} = \frac{1}{n}\sum_{i=1}^n X_i$ als Schätzvariable für μ, wobei die X_i unabhängige Wiederholungen von X seien. Wir setzen voraus, dass $\bar{X}$ für $n \geq 1500$ normalverteilt ist.

1. Berechnen Sie die Varianz von $\bar{X}$!

2. Leiten Sie eine möglichst kleine untere Schranke n_0 für n her, so dass für alle $n \geq n_0$ gilt
$$\Pr[|\bar{X} - \mu| < 0{,}1] \geq 0{,}9\,.$$
Für das 0,95-Quantil der Standardnormalverteilung ist dabei $z_{0{,}95} \approx 1{,}65$ zu verwenden.

3. Nun sei $n = 2500$. Geben Sie ein Konfidenzintervall für μ zum Konfidenzniveau 0,9 an!

Lösungsvorschlag

1. Es gilt
$$\begin{aligned}\mathrm{Var}[\bar{X}] &= \mathrm{Var}[\tfrac{1}{n}\textstyle\sum_{i=1}^n X_i] \\ &= \tfrac{1}{n^2}\textstyle\sum_{i=1}^n \mathrm{Var}[X_i] \\ &= \tfrac{1}{n^2}\cdot n\cdot \sigma_X^2 = \tfrac{4}{n}.\end{aligned}$$

2. Wir setzen $\bar{\sigma}^2 = \mathrm{Var}[\bar{X}]$. Dann gilt $\bar{\sigma} = \frac{2}{\sqrt{n}}$.
$$\begin{aligned}\Pr[|\bar{X}-\mu| < 0{,}1] &= \Pr[-0{,}1 < \bar{X} - \mu < 0{,}1] \\ &\approx \Phi\left(\frac{0{,}1}{\bar{\sigma}}\right) - \Phi\left(-\frac{0{,}1}{\bar{\sigma}}\right) \\ &= 2\cdot\Phi\left(\frac{0{,}1}{\bar{\sigma}}\right) - 1\,.\end{aligned}$$
Es gilt $2\cdot\Phi\left(\frac{0{,}1}{\bar{\sigma}}\right) - 1 \geq 0,9$ genau dann, wenn $\Phi\left(\frac{0{,}1}{\bar{\sigma}}\right) \geq 0,95\,.$

Mit dem Quantil $z_{0{,}95} \approx 1,65$ erhalten wir für n die Ungleichung

$$\frac{0{,}1}{\bar{\sigma}} = \frac{\sqrt{n}}{20} \geq 1,65\,,$$

d. h.
$$n \geq 20^2 \cdot 1,65^2 = 1089.$$

Da die Normalverteilung $n \geq 1500$ voraussetzt, folgt

$$n \geq \max\{1500, 1089\} = 1500.$$

3. Für $n = 2500$ und einen zu bestimmenden Parameter c betrachten wir $\Pr[|\bar{X} - \mu| < c]$

   und erhalten die Gleichung

   $$\frac{c}{\bar{\sigma}} = \frac{c\sqrt{2500}}{2} = 1,65$$

   mit der Lösung $c = \frac{3{,}3}{50} = 0,066$.

   Das gesuchte Konfidenzintervall für μ ist demnach

   $$\bar{X} - 0,066 < \mu < \bar{X} + 0,066\,.$$

Hinweis: *Die als freiwillig gekennzeichneten Aufgaben werden nicht bewertet.*

Freiwillige Aufgabe 1

In einer großen Stadt gibt es N Taxis, die mit den Nummern $1, \ldots, N$ beschriftet sind. Wir stehen an der Straße und beobachten die Taxis, dabei notieren wir uns deren Nummer (Wiederholungen werden ignoriert). Sei $x_1, \ldots, x_n$ die aufsteigend sortierte Folge der beobachteten Nummern. Nun wollen wir die Anzahl N der Taxis schätzen.

1. Was ist der Maximum Likelihood Schätzer für N? Ist dieser erwartungstreu?

2. Geben Sie einen Schätzer für die Anzahl der Taxis an, der N dadurch abschätzt, dass er versucht, die Größe der (ev. nicht beobachteten) Lücke $x_n + 1, \ldots, N$ oberhalb von x_n abzuschätzen. Ist dieser Schätzer erwartungstreu?

Lösungsvorschlag

1. Es gilt offensichtlich $N \geq x_n$, und übrigens auch $x_n \geq n$. Die Wahrscheinlichkeit $\Pr_N[\{x_1, \ldots x_n\}]$ ("Wahrscheinlichkeit, $\{x_1, \ldots x_n\}$ zu beobachten, wenn es N Taxis gibt") ist $\binom{N}{n}^{-1}$, denn bei N Taxis ist jede Teilmenge der Mächtigkeit n gleichwahrscheinlich. Für einen Maximum-Likelihood Schätzer gilt es, $\binom{N}{n}^{-1}$ in Abhängigkeit von N zu maximieren, was genau dann der Fall ist, wenn N möglichst klein geschätzt wird, was für $\widehat{N}(\{x_1, \ldots x_n\}) = x_n$ der Fall ist. Man schätzt also die Anzahl N der Taxis durch die höchste beobachtete Nummer x_n ab. Dieser Schätzer ist offensichtlich nicht erwartungstreu, weil man oft zu niedrig, aber nie zu hoch schätzt.

   Wir wollen die Aussage zur Erwartungstreue nun präzisieren und definieren zu gegebenem Stichprobenumfang n und Stichprobenvariablen X_i die Schätzvariable $X = \max\{X_1, X_2, \ldots, X_n\}$ für den Parameter N. Man beachte, dass hier die Stichprobenvariablen einmal nicht voneinander unabhängig sind, weil die Notierung der

Taxinummern einer Auswahl ohne Zurücklegen entspricht, um die Wiederholung von Nummern zu vermeiden.

Jede Teilmenge $\{x_1, \ldots, x_n\} \subseteq [N]$ ist gleichwahrscheinlich mit $\Pr_N[\{x_1, \ldots x_n\}] = \binom{N}{n}^{-1}$. Zu gegebenem x_n mit $n \leq x_n \leq N$, gibt es $\binom{x_n - 1}{n-1}$ Teilmengen $\{x_1, x_2, \ldots, x_{n-1}\} \subseteq [N]$ mit $x_i < x_n$. Wir erhalten also

$$
\begin{aligned}
\mathbb{E}[X] &= \sum_{\{x_1,\ldots,x_n\}\subseteq[N]} x_n \cdot \Pr_N[\{x_1, \ldots, x_n\}] \\
&= \sum_{x=n}^{N} x \cdot \binom{N}{n}^{-1} \cdot \binom{x-1}{n-1} \\
&= \sum_{x=n}^{N} x \cdot \frac{n!}{N^{\underline{n}}} \cdot \frac{(i-1)^{\underline{x-1}}}{(n-1)!} \\
&= \frac{n}{N^{\underline{n}}} \sum_{x=n}^{N} x^{\underline{n}} \\
&= \frac{n}{N^{\underline{n}}} \cdot \left[\frac{x^{\underline{n+1}}}{n+1} \right]_n^{N+1} \\
&= \frac{n}{N^{\underline{n}}} \cdot \left[\frac{(N+1)^{\underline{n+1}}}{n+1} - \frac{n^{\underline{n+1}}}{n+1} \right] \\
&= \frac{nN+n}{n+1} = N - \frac{N-n}{n+1} .
\end{aligned}
$$

Nur im Falle $n = N$ gilt $\mathbb{E}[X] = N$, ansonsten gilt $\mathbb{E}[X] \neq N$.

2. Wir wählen nun den folgenden Ansatz: Zwischen den beobachteten x_i haben wir "Lücken". Wir versuchen nun, dadurch die Größe der Lücke von x_n zu N durch die mittlere Lückengröße in $\{x_1, \ldots x_n\}$ abzuschätzen. Beachten Sie, dass unmittelbar aufeinanderfolgende Nummern $x_{i+1} = x_i + 1$ eine Lücke mit Wert 0 bedeutet. Wenn wir $x_0 = 0$ definieren, ist die mittlere Lückengröße

$$\frac{1}{n}\left(((x_1 - x_0) - 1) + ((x_2 - x_1) - 1) + \cdots + ((x_n - x_{n-1}) - 1)\right) = \frac{x_n - n}{n}$$

und führt zum Schätzer

$$X_1(\{x_1, \ldots x_n\}) = x_n + \frac{x_n - n}{n}.$$

Dieser Schätzer ist in der Tat erwartungstreu: Sei $X := \max\{X_1, \ldots X_n\}$. Der Erwartungswert des Schätzers, wenn die Anzahl der Taxis gleich N ist, ist

$$\mathbb{E}[X_1] = \mathbb{E}\left[X + \frac{X-n}{n}\right] = \frac{n+1}{n} \cdot \mathbb{E}[X] - 1\,,$$

woraus unmittelbar

$$\mathbb{E}[X_1] = \frac{n+1}{n} \cdot \left(N + \frac{n-N}{1+n}\right) - 1 = N$$

folgt.

***Hinweis:** Auf den Übungsblättern in diesem Semester wird es grundsätzlich die drei Aufgabentypen Vorbereitungsaufgabe, Tutoraufgabe und Hausaufgabe geben. Die als Vorbereitung bezeichneten Aufgaben dienen der häuslichen Vorbereitung der Tutoraufgaben. Tutoraufgaben werden in den Übungsgruppen bearbeitet. Dabei wird die Lösung der Vorbereitungsaufgaben vorausgesetzt. Die Vorbereitungsaufgaben werden in der Zentralübung unterstützt.*

Vorbereitung 1

Wiederholung rekurrente Ereignisfolgen, Hausaufgabe 1, Blatt 9.

Tutoraufgabe 1

Auf zwei unabhängigen Servern stehe ein Web-Dienst zur Verfügung. Es soll festgestellt werden, welcher Server schnellere Antwortzeiten liefert.

Dazu werden $n = 1000$ Anfragen an die Server geschickt und es wird festgestellt, von welchem Server die Antwort zuerst eintrifft. Dabei gehen wir davon aus, dass Pakete nicht gleichzeitig empfangen werden können. In 540 Fällen antwortet Server A vor Server B.

Wir wählen als Nullhypothese H_0 die Aussage, dass Server B im Mittel schneller ist als Server A.

Kann man für einen entsprechenden statistischen Test auf einem Signifikanzniveau von $\alpha = 0,04$ die Nullhypothese annehmen? Formulieren Sie hierzu den Test und weisen Sie Ihre Behauptung nach.

Lösungsvorschlag

Sei X eine Bernoulli-verteilte Zufallsvariable mit Wert 1 genau dann, wenn A schneller antwortet als B. Wir schreiben $p = \Pr[X = 1]$. Seien X_i für $i \in \{1, 2, \ldots, n\}$ unabhängige Wiederholungen von X mit $n = 1000$. Dann ist die Testgröße $T = \sum_{i=1}^{n} X_i$ binomialverteilt. Wir formulieren einen approximativen Binomialtest.

Nullhypothese H_0: $p < \frac{1}{2}$, Alternative H_1: $p \geq \frac{1}{2}$.

Da wir $T = 540$ als Testwert bereits kennen, könnten wir sofort mit dem Standard-Binomialtest aus der Vorlesung prüfen, ob H_0 mit Signifikanzniveau, d. h., maximaler Irrtumswahrscheinlichkeit $\alpha = 0.04$ ablehnen können. Und dies ist in der Tat so. Aber das war nicht die Frage. Gefragt war, ob wir H_0 mit Signifikanzniveau $\alpha = 0.04$ annehmen können.

Um diese Frage korrekt mit den Ablehnungskriterien des Standard-Binomialtest aus der Vorlesung beantworten zu können, müssen wir H_0 und H_1 quasi vertauschen. Wir setzen $H_0' = H_1$ und $H_1' = H_0$, d. h.

Nullhypothese H_0': $p \geq \frac{1}{2}$, Alternative H_1': $p < \frac{1}{2}$,

und wenden den approximativen Binomialtest wie folgt an. Sei mit $p_0 = \frac{1}{2}$ und $T = h$

$$Z = \frac{h - np_0}{\sqrt{np_0(1-p_0)}} = \frac{540 - 500}{\sqrt{250}} \approx 2.5298\ldots\ .$$

Das Ablehnungskriterium für H_0' mit $\alpha = 0.04$ ist $Z < z_{0.04}$. Dieses Kriterium ist aber wegen $Z = \frac{540-500}{\sqrt{250}} > z_{0.04}$ nicht erfüllt. Auf diesem Weg kann also nicht hergeleitet werden, dass H_0 angenommen werden kann mit Signifikanzniveau $\alpha = 0.04$. Es gilt ganz im Gegenteil, dass H_0 mit Signifikanzniveau $\alpha = 0.04$ abgelehnt werden kann.

Tutoraufgabe 2

Seien $(X_t)_{t\in\mathbb{N}_0}$ die Zufallsvariablen einer zeithomogenen Markov-Kette über den Zuständen $Q = \{0,1,2\}$ mit Übergangsmatrix

$$P = (p_{i,j}) = \begin{pmatrix} 0,25 & 0,25 & 0,5 \\ 0,25 & 0,25 & 0,5 \\ 0,25 & 0,25 & 0,5 \end{pmatrix}.$$

Die Dichtefunktion von X_0, d. h., die Startverteilung der Markov-Kette sei $q_0 = (s_0, s_1, s_2)$.

1. Berechnen Sie die Dichtefunktion q_1 von X_1.
2. Bestimmen Sie die Menge aller stationären Startverteilungen.
3. Beweisen Sie die Unabhängigkeit der beiden Variablen X_0 und X_1.

   Dabei sind X_0 und X_1 als Zufallsvariable über dem zugeordneten Wahrscheinlichkeitsraum $\langle\Omega, \Pr\rangle$ zu betrachten mit

$$\Omega = \{(x_0,x_1) : x_0, x_1 \in Q\}, \quad \Pr[(x_0,x_1)] = (q_0)_{x_0} \cdot \Pr[X_1 = x_1 | X_0 = x_0],$$
$$X_0((x_0,x_1)) = x_0 \qquad \text{und} \qquad X_1((x_0,x_1)) = x_1.$$

Lösungsvorschlag

1.
$$\begin{aligned} q_1 &= q_0 \cdot P \\ &= (\sum_{i=0}^{2} s_i \cdot 0{,}25,\ \sum_{i=0}^{2} s_i \cdot 0{,}25,\ \sum_{i=0}^{2} s_i \cdot 0{,}5) \\ &= (0{,}25\,,\ 0{,}25\,,\ 0{,}5) \end{aligned}$$

2. Für stationäre Lösungen (s_0, s_1, s_2) muss gelten
$$(s_0, s_1, s_2) = (s_0, s_1, s_2) \cdot P$$
   mit Nebenbedingung
$$\textstyle\sum_{i=0}^{2} s_i = 1.$$
   Wegen
$$(s_0, s_1, s_2) \cdot P = (0{,}25\,,\ 0{,}25\,,\ 0{,}5)$$
   folgt
$$(s_0, s_1, s_2) = (0{,}25\,,\ 0{,}25\,,\ 0{,}5).$$

3. Es seien q_0 und q_1 die diskreten Verteilungen von X_0 und X_1. Für die Unabhängigkeit von X_0 und X_1 genügt der Nachweis der Gleichung
$$\Pr[(x_0,x_1)] = (q_0)_{x_0} \cdot (q_1)_{x_1}$$
   für alle $x_0, x_1 \in \{0,1,2\}$.

   Wir rechnen
$$\begin{aligned} (q_0)_{x_0} \cdot (q_1)_{x_1} &= (q_0)_{x_0} \cdot (q_0 P)_{x_1} \\ &= (q_0)_{x_0} \cdot \left(\sum_{i=0}^{2} (q_0)_i\, p_{i,x_1} \right) \\ &= (q_0)_{x_0} \cdot (p_{x_0,x_1}) \\ &= \Pr[(x_0,x_1)]. \end{aligned}$$

Tutoraufgabe 3

1. Wir betrachten Markov-Ketten M mit 6 Zuständen. Wie viele transiente Zustände kann M höchstens besitzen? Begründung!

2. Wie viele stationäre Verteilungen besitzt die Markovkette mit Übergangsmatrix

$$P = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix} ?$$

   Begründung!

3. Gegeben sei eine Markov-Kette M mit Zustandsmenge $\mathcal{S} = \{0, 1, 2, \ldots\}$, $p_{n,(n+1)} = 2/3$ und $p_{n,0} = 1/3$ für alle $n \in \mathcal{S}$. Wie groß ist die Wahrscheinlichkeit, sich nach langer Zeit im Zustand i zu befinden?

Lösungsvorschlag

1. Hätte M 6 transiente Zustände, dann würden alle 6 Zustände mit positiver Wahrscheinlichkeit verlassen werden können bei $n > 6$ Übergängen. Dies ist nicht möglich, weil sich mindestens 1 Zustand nach 6 Übergängen wiederholen muss.

   Andererseits gibt es ein Beispiel einer Markovkette, mit einem einzigen absorbierenden Zustand, der von allen anderen 5 Zuständen aus erreichbar ist.

   Im Ergebnis ist die maximale Zahl transienter Zustände gleich 5.

2. Eine stationäre Verteilung $\pi = (c_1, c_2)$ erfüllt zum einen die Gleichung

$$c_1 + c_2 = 1$$

   und ist auch ein Linkseigenvektor von P zum Eigenwert 1, erfüllt also auch die Gleichung

$$\pi = \pi \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix} .$$

   Die Lösung ergibt sich eindeutig mit

$$c_1 = \frac{2}{5} \quad \text{und} \quad c_2 = \frac{3}{5} .$$

   Man kann auch ohne Rechnung sehen, dass es höchstens eine einzige Lösung geben kann. Gäbe es noch eine zweite stationäre Lösung, dann müssten alle Verteilungen stationär sein, denn der Raum der Eigenvektoren zum Eigenwert 1 wäre dann 2-dimensional. Offensichtlich aber ist z. B. $q = (1, 0)$ nicht stationär.

3. Sei $M = (X_t)_{t\in\mathbb{N}_0}$ und $(q_t)_{t\in\mathbb{N}_0}$ die Folge der entsprechenden Verteilungen der X_t. Für eine beliebige Anfangsverteilung q_0 gilt bereits $q_t(0) = \Pr[X_t = 0] = \frac{1}{3}$ für alle $t > 0$. Dies zeigt die folgende Rechnung.

$$\begin{aligned} q_t(0) &= \sum_{n=0}^{\infty} q_{t-1}(n) \cdot p_{n,0} \\ &= \underbrace{(q_{t-1}(0) + q_{t-1}(1) + q_{t-1}(2) + \dots)}_{=1} \cdot \frac{1}{3} \\ &= \frac{1}{3}\,. \end{aligned}$$

Die Wahrscheinlichkeit $q_t(n)$, dass X_t im Zustand n ist, ist für alle $n \geq 1$ gegeben durch

$$q_t(n) = q_{t-1}(n-1) \cdot p_{n-1,n} = q_{t-1}(n-1) \cdot \frac{2}{3}\,.$$

Wir erhalten

$$(\forall t \geq n) \ \left[p_t(n) = \frac{1}{3}\left(\frac{2}{3}\right)^n \right]\,.$$

Wenn man also den Zeitpunkt t genügend groß wählt, dann wird der Zustand n mit der berechneten Wahrscheinlichkeit angenommen.