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

Sommersemester 2011
Lösungsblatt 8
8. Juli 2011

Diskrete Wahrscheinlichkeitstheorie

Hausaufgabe 1 (5 Punkte)

Sie werfen eine kreisförmige Münze mit Radius 0.5 auf ein Quadrat mit Seitenlänge 4, wobei der Mittelpunkt der Münze stets innerhalb des Quadrats zu liegen kommt.

Wie hoch ist die Wahrscheinlichkeit, dass die Münze über den Rand des Quadrats hinausragt, wenn wir eine Gleichverteilung des Mittelpunkts der Münze auf der Quadratfläche annehmen?

Lösung

Die Münze ragt genau dann nicht über den Rand des Quadrats Q hinaus, wenn der Mittelpunkt der Münze in einem abgeschlossenen Quadrat B mit Seitenlänge 3 liegt, das zentriert innerhalb von Q liegt.

Damit ist klar, dass die gesuchte Wahrscheinlichkeit Pr durch das Verhältnis der Flächen von $Q \setminus B$ zu Q gegeben ist.

Ergebnis: $\Pr = \frac{7}{16}$.

Bemerkung: Die Aufgabe kann auch durch Rechnung gelöst werden, wie folgt:
Wir bezeichnen $[a, b]_{\mathbb{R}} := \{x \in \mathbb{R}\,;\, a \leq x \leq b\}$.
Seien $Q = \{(x, y)\,;\, x, y \in [0, 4]_{\mathbb{R}}\}$ und $Q = \{(x, y)\,;\, x, y \in [0.5, 3.5]_{\mathbb{R}}\}$.
Man definiert zwei kontinuierliche Zufallsvariablen X und Y über einem Kolmogorov'schen Wahrscheinlichkeitsraum $\langle \Omega, \mathcal{A}, \Pr\rangle$ durch Angabe der 2-dimensionalen Dichte

$$f_{X,Y}(x, y) = \begin{cases} \frac{1}{16} & : (x, y) \in Q, \\ 0 & : \text{sonst.} \end{cases}$$

Dann gilt

$$\Pr[(x, y) \notin B] \;=\; 1 - \int_B f_{X,Y}(x, y)\, \mathrm{d}x\, \mathrm{d}y\,.$$

Hausaufgabe 2 (5 Punkte)

Seien X und Y kontinuierliche Zufallsvariablen mit gemeinsamer Dichtefunktion

$$f_{X,Y}(x, y) = \begin{cases} 6xy^2 & : \quad 0 \leq x \leq 1,\, 0 \leq y \leq 1 \\ 0 & : \quad \text{sonst} \end{cases}$$

1. Berechnen Sie die Randdichte $f_X(x)$.
2. Bestimmen Sie den Wert der Verteilungsfunktion $F_{X,Y}(\frac{1}{2}, \frac{1}{2})$.
3. Zeigen Sie die Unabhängigkeit der Variablen X und Y.

Lösung

1. $f_X(x) = 2x$. Berechnung für $0 \leq x \leq 1$:

$$\begin{aligned} f_X(x) &= \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,\mathrm{d}y \\ &= \int_0^1 6xy^2\,\mathrm{d}y \\ &= 2x \cdot \left[y^3\right]_0^1 \;=\; 2x\,. \end{aligned}$$

2. $F_{X,Y}(\frac{1}{2},\frac{1}{2}) = \frac{1}{32}$. Berechnung:

$$\begin{aligned} F_{X,Y}(\tfrac{1}{2},\tfrac{1}{2}) &= \int_0^{\frac{1}{2}}\left(\int_0^{\frac{1}{2}} 6xy^2\,\mathrm{d}y\right)\mathrm{d}x \\ &= \int_0^{\frac{1}{2}} 2x\cdot\left[y^3\right]_0^{\frac{1}{2}}\,\mathrm{d}x \\ &= \int_0^{\frac{1}{2}} 2x\cdot\frac{1}{8}\,\mathrm{d}x \\ &= \frac{1}{8}\cdot\left[x^2\right]_0^{\frac{1}{2}} \;=\; \frac{1}{32}\,. \end{aligned}$$

3. Mit

$$\begin{aligned} f_Y(y) &= \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,\mathrm{d}y \\ &= \int_0^1 6xy^2\,\mathrm{d}x \\ &= 3y^2\cdot\left[x^2\right]_0^1 \;=\; 3y^2\,. \end{aligned}$$

für alle $0 \leq y \leq 1$ folgt $f_{X,Y}(x,y) = f_X(x)\cdot f_Y(y)$ für alle $0 \leq x,y \leq 1$.
Ansonsten gilt $f_{X,Y}(x,y) = 0 = f_X(x)\cdot f_Y(y)$.

Hausaufgabe 3 (5 Punkte)

Die Lebensdauer T eines Rechners habe die folgende mit $\lambda > 0$ parametrisierte Dichte:

$$f_T(t) = \begin{cases} \lambda^2 t e^{-\lambda t} & : t \geq 0 \\ 0 & : t < 0 \end{cases}$$

1. Berechnen Sie $\Pr[T \leq \frac{2}{\lambda}]$.

2. Berechnen Sie $\mathbb{E}[T]$ in Abhängigkeit von λ.

Lösung

1. Wir wenden partielle Integration an wie folgt.

$$\begin{aligned}
\Pr\left[T \le \frac{2}{\lambda}\right] &= \int_0^{2/\lambda} f_T(t)\,\mathrm{d}t \\
&= \lambda^2 \int_0^{2/\lambda} te^{-\lambda t}\,\mathrm{d}t \\
&= \lambda^2 \cdot \left(\left[-\frac{1}{\lambda}te^{-\lambda t}\right]_0^{2/\lambda} + \left[-\frac{1}{\lambda}e^{-\lambda t}\right]_0^{2/\lambda}\right) \\
&= \lambda^2 \cdot \left(-\frac{1}{\lambda}\cdot\frac{2}{\lambda}e^{-2} - \frac{1}{\lambda^2}e^{-2} + \frac{1}{\lambda^2}\right) \\
&\approx 0{,}59399415\,.
\end{aligned}$$

2. Teilweise wie vorausgehend rechnen wir

$$\begin{aligned}
\mathbb{E}[T] &= \int_0^{\infty} t \cdot f_T(t)\,\mathrm{d}t \\
&= \int_0^{\infty} \lambda^2 t^2 e^{-\lambda t}\,\mathrm{d}t \\
&= \left[-\lambda t^2 e^{-\lambda t}\right]_0^{\infty} - \int_0^{\infty} -2\lambda te^{-\lambda t}\,\mathrm{d}t \\
&= 2\lambda \int_0^{\infty} te^{-\lambda t}\,\mathrm{d}t \\
&= 2\lambda \cdot \frac{1}{\lambda^2} \quad \text{(siehe oben)} \\
&= \frac{2}{\lambda}\,.
\end{aligned}$$

Hausaufgabe 4 (5 Punkte)

Wahr oder falsch? Begründen Sie Ihre Antwort!

1. Falls $X \sim \mathcal{N}(0,1)$ und $Y \sim \mathcal{N}(0,2)$ normalverteilt sind, dann folgt $\mathrm{Var}[X+Y] = 3$.
2. Seien X und Y standardnormalverteilt. Dann gilt $\Pr[X \le 0] = \Pr[Y \ge 0]$.
3. Bei kontinuierlichen Zufallsvariablen existiert stets der Erwartungswert.
4. Seien $X \sim \mathrm{Po}(1)$ und $Y \sim \mathrm{Bin}(n, \frac{1}{n})$. Dann gilt $|\Pr[X = 2n] - \Pr[Y = 2n]| < 2^{-n}$.
5. Sei X exponentialverteilt. Dann gilt $\Pr[X > 2 \mid X > 1] + \Pr[X \le 1] = 1$.

Lösung

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

1. Falsch! Begründung: Das ist allgemein nur richtig, wenn X und Y unabhängig sind.

2. Wahr! Es gilt $\Pr[X \leq 0] = \frac{1}{2} = \Pr[Y \leq 0] = 1 - \Pr[Y \leq 0] = \Pr[Y \geq 0]$.

3. Falsch! Begründung: Durch Vertauschung von Abschnitten der Dichtefunktion können den kleinen Dichtewerten beliebig hohe Werte der Variablen zugeordnet werden.

4. Wahr! Begründung: $\Pr[Y = 2n] = 0$ und $\Pr[X = 2n] = \frac{e^{-1}}{(2n)!} < 2^{-n}$.

5. Wahr! Begründung: Da die Exponentialverteilung gedächtnislos ist, gilt $\Pr[X > 2 \mid X > 1] = \Pr[X > 1] = 1 - \Pr[X \leq 1]$.

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

In welcher Weise hängt der „Random Walk“ mit Tutoraufgabe 1 zusammen?

Lösung

Elementar.

Vorbereitung 2

In einem Unfallkrankenhaus treffen im Schnitt alle 20 Minuten Patienten zur Behandlung ein. Die Zeit zwischen zwei Behandlungsfällen sei exponentialverteilt mit Parameter $\frac{1}{20}$. Wenn 1 Stunde lang kein Patient eingetroffen ist, macht das Personal Ruhepause. Wir wollen wissen, welcher Zeitabstand zwischen zwei Ruhepausen zu erwarten ist.
Seien $T_1, T_2, \ldots$ die Zeitspannen zwischen dem Eintreffen zweier Behandlungsfälle und W die Wartezeit bis zur nächsten Ruhepause.

1. Geben Sie $\mathbb{E}[T_1 \mid T_1 \geq 60]$ an.

2. Geben Sie $\mathbb{E}[W \mid T_1 \geq 60]$ an.

3. Berechnen Sie $\mathbb{E}[W]$.

Lösung

Vorüberlegungen:

a) Es wird offenbar eine Folge von Zeitpunkten z_i betrachtet,
zu denen Patienten im Krankenhaus eintreffen.

b) Die Zeitdifferenzen $T_i = z_i - z_{i-1}$ werden als exponentialverteilt angenommen. Dies bedeutet, dass diese T_i nicht davon abhängen, wie lange noch kein Patient eingetroffen ist (Gedächtnislosigkeit).

Der Parameter λ bedeutet „Anzahl der Patienten pro Zeiteinheit" im Durchschnitt, hier also $\frac{1}{20}$ Patient pro Zeiteinheit als Erwartungswert.

Alle T_i sind unabhängig, d. h. die Menge der T_i ist unabhängig.

c) Falls $T_i > 60$, dann gibt es eine Ruhepause.

d) Die Frage ist, wie lange man warten muss, bis erstmalig $T_N > 60$ festgestellt wird?

Die Wartezeit ist dann

$$W = 60 + \sum_{j=1}^{N-1} T_j$$

oder

$$W = 60 + W'$$

mit

$$W' = \sum_{j=1}^{N-1} T_j\,.$$

1. Sei $\lambda = \frac{1}{20}$. Dann gilt $\mathbb{E}[T_1] = \frac{1}{\lambda} = 20$.

   Da die Exponentialverteilung gedächtnislos ist, gilt

   $$\mathbb{E}[T_1 \mid T_1 \geq 60] = 60 + \mathbb{E}[T_1] = 80\,.$$

   Wir zeigen die Gleichung direkt durch Berechnung wie folgt:

   T_1 ist exponentialverteilt:

   $$F_{T_1}(t) = \begin{cases} 1 - e^{-\lambda t} & : t \geq 0 \\ 0 & : \text{sonst}\,. \end{cases}$$

   Verteilung der bedingten Variablen $T_1 \mid T_1 \geq 0$:

   $$\begin{aligned} F_{(T_1|T_1 \geq 60)}(t) &= \Pr[(T_1 \mid T_1 \geq 0) \leq t] \\ &= \frac{\Pr[T_1 \leq t, T_1 \geq 60]}{\Pr[T_1 \geq 60]} \end{aligned}$$

   Für $t \geq 60$ folgt

   $$\begin{aligned} F_{(T_1|T_1 \geq 60)}(t) &= \frac{(1 - e^{-\lambda \cdot t}) - (1 - e^{-\lambda \cdot 60})}{e^{-\lambda \cdot 60}} \\ &= 1 - e^{-\lambda \cdot (t-60)}\,. \end{aligned}$$

Damit ist die Variable $T' = (T_1 \mid T_1 \geq 60) - 60$ gleichverteilt wie T_1, d. h. entsprechend exponentialverteilt.

Es folgt

$$\begin{aligned} \mathbb{E}[(T_1 \mid T_1 \geq 60)] &= \mathbb{E}[T' + 60] \\ &= \mathbb{E}[T_1] + 60 \\ &= 80 \,. \end{aligned}$$

2. Offenbar gilt

$$\mathbb{E}[W \mid T_1 \geq 60] = 60 \,.$$

3. Es gilt $W = 60 + \sum_{j=1}^{N-1} T_j$ oder $W = 60 + W'$ mit $W' = \sum_{j=1}^{N-1} T_j$.

N ist eine geometrisch verteilte Zufallsvariable mit Erfolgswahrscheinlichkeit $p = \Pr[T \geq 60]$. Es gilt

$$p = \Pr[T \geq 60] = e^{-3} \,.$$

Berechnung von $\mathbb{E}[W']$:

Wir setzen $T = T_1$.

$$\begin{aligned} \mathbb{E}[W'] &= \sum_{n=1}^{\infty} \mathbb{E}[W' | N = n] \cdot p(1-p)^{n-1} \\ &= \sum_{n=1}^{\infty} \mathbb{E}[T | T \leq 60] \cdot (n-1) \cdot p(1-p)^{n-1} \\ &= \mathbb{E}[T | T \leq 60] \cdot \mathbb{E}[N-1]. \end{aligned}$$

Es gilt

$$\mathbb{E}[N-1] = e^3 - 1 \,.$$

$\mathbb{E}[T|T \leq 60]$ erhalten wir aus den Gleichungen

$$\begin{aligned} \mathbb{E}[T] &= \mathbb{E}[T|T \leq 60] \cdot \Pr[T \leq 60] \; + \; \mathbb{E}[T|T \geq 60] \cdot \Pr[T \geq 60] \\ &= \mathbb{E}[T|T \leq 60] \cdot (1 - e^{-3}) + 80 \cdot e^{-3} \\ &= 20 \,. \end{aligned}$$

Es folgt

$$\mathbb{E}[T|T \leq 60] = \frac{20 - 80 \cdot e^{-3}}{1 - e^{-3}} = \frac{20 \cdot e^3 - 80}{e^3 - 1} \,.$$

Ergebnis:

$$\begin{aligned} \mathbb{E}[W'] &= 20 \cdot e^3 - 80 \,, \\ \mathbb{E}[W] &= 20 \cdot e^3 - 80 + 60 \\ &\approx 382 \text{ (Minuten)} \,. \end{aligned}$$

Vorbereitung 3

Bei einem Einwahlserver für $n = 10^3$ Teilnehmer nehmen wir an, dass zu einem festen Zeitpunkt jeder Teilnehmer mit Wahrscheinlichkeit $p = 0,05$ Zugriff auf den Server wünscht.

Berechnen Sie eine Näherung der Wahrscheinlichkeit, mit der gleichzeitig mehr als 55 Verbindungswünsche auftreten? Approximieren Sie dabei die Binomialverteilung durch die entsprechende Normalverteilung und benutzen Sie ggf. geeignete Tabellen für die Werte der Standardnormalverteilung.

Lösung

Sei X die Anzahl der Verbindungswünsche. X ist näherungsweise normalverteilt. Wir nehmen also an $X \sim \mathcal{N}(\mu, \sigma^2)$ mit $\mu = np = 50$ und $\sigma^2 = np(1-p) = 47,5$. Damit folgt

$$\begin{aligned}
\Pr[X > 55] &= 1 - \Pr[X \le 55] \\
&= 1 - \Pr\left[\frac{X-\mu}{\sigma} \le \frac{55-50}{\sqrt{47,5}}\right] \\
&= 1 - \Phi\left(\frac{55-50}{\sqrt{47,5}}\right) \\
&\approx 0,2343\,.
\end{aligned}$$

Vorbereitung 4

Sei X eine binomialverteilte Zufallsvariable mit $X \sim \mathrm{Bin}(2000, 0.05)$. Wir nehmen an, dass die Voraussetzungen sowohl für eine Approximation von entsprechenden Verteilungen mit Poisson-Verteilung als auch mit Normalverteilung vorliegen.
Berechnen Sie approximativ

1. $\Pr[X = 110]\,,$
2. $\Pr[X > 110]\,.$

Begründen Sie jeweils die Wahl einer der Approximationen.

Lösung

1. Für die Wahrscheinlichkeit des Auftretens eines bestimmten Wertes einer diskreten Zufallsvariablen, wie hier $X = 110$, ist es naheliegend die approximierende diskrete Verteilung zu verwenden. In diesem Fall wählen wir deshalb die Poisson-Verteilung.

   Es gilt zunächst $E[X] = 2000 \cdot 0.05 = 100$. Mit $\lambda = n \cdot p_n = 100$ gilt

   $$\lim_{n\to\infty} b(k; n, p_n) = \lim_{n\to\infty} \binom{n}{k} \cdot p_n^k \cdot (1-p_n)^{n-k} = e^{-\lambda} \cdot \frac{\lambda^k}{k!}$$

   und wir erhalten mit $k = 110$

   $$\Pr[X\!=\!110] \;=\; b(110; 2000, 0.05) \;\approx\; e^{-100} \cdot \frac{100^{110}}{110!}\;.$$

   Die Berechnung erfordert nun die Anwendung der Stirling-Formel auf die Fakultätsfunktion.

   $$k! \;\approx\; \sqrt{2\pi k} \cdot \left(\frac{k}{e}\right)^k\;.$$

Damit rechnen wir

$$\begin{aligned} e^{-100} \cdot \frac{100^{110}}{110!} &= e^{-100} \cdot \frac{100^{110}}{\sqrt{2\pi \cdot 110}} \cdot \left(\frac{e}{110}\right)^{110} \\ &= \frac{e^{10}}{\sqrt{2\pi \cdot 110}} \cdot \left(\frac{100}{110}\right)^{110} \\ &\approx 0.0234\ldots \end{aligned}$$

Ergebnis: $\Pr[X=110] \approx 0.0234$.

Bemerkung: In praktischen Anwendungen müss nun eine Abschätzung der Genauigkeit der Berechnungen folgen.

2. Es giltt $\Pr[X > 110] = 1 - \Pr[X \leq 110]$. D.h., wir müssen lediglich den Wert der Verteilungsfunktion der binomialen Dichtefunktion an der Stelle 110 berechnen.

   Allerdings haben wir die Verteilungsfunktion der Poisson-Dichte noch nicht zur Verfügung. Schon aus diesem Grund bietet sich die Normalverteilung für eine Approximation an. In diesem Fall steht uns die Verteilungsfunktion Φ für die Standardnormalverteilung zur Verfügung.

   Wir berechnen die Wahrscheinlichkeit $\Pr[X \leq 110]$ mit $\mathbb{E}[X] = 2000 \cdot 0.05 = 100$ und $\sigma = \sqrt{np(1-p)} = \sqrt{100 \cdot 0.95} = \sqrt{95}$ wie folgt.

$$\begin{aligned} \Pr[X \leq 110] &= F_X(110) \\ &\approx \Phi\left(\frac{k - np}{\sqrt{np(1-p)}}\right) \\ &= \Phi\left(\frac{110 - 100}{\sqrt{95}}\right) \\ &\approx \Phi(1.025978) \approx 0.847 \end{aligned}$$

   Ergebnis: $\Pr[X > 110] \approx 1 - 0.847 = 0.153$.

Tutoraufgabe 1

Wir werfen eine unfaire Münze, die mit Wahrscheinlichkeit $\frac{1}{3}$ „Kopf" und mit Wahrscheinlichkeit $\frac{2}{3}$ „Zahl" zeigt. Wir betrachten für eine Folge von Würfen $W_1, W_2, \ldots$ das Ereignis $E(i,k)$, beim i-ten Wurf weniger als $(\frac{i}{2}+k)$-mal „Kopf" geworfen zu haben.
Widerlegen Sie:

$$\forall k > 0 : \lim_{n\to\infty} \Pr\left[\bigcap_{i=1}^{n} E(i,k)\right] = 0\,.$$

Lösung

Wir widerlegen die Behauptung und begründen zunächst den Zusammenhang der Aufgabe mit rekurrenten Ereignissen, insbesondere mit dem „Random Walk".

Wir betrachten Zufallsvariable K_i und Z_i für die Anzahl der Ergebnisse „Kopf" bzw. „Zahl", die bei i Würfen eingetreten sind. Beide Variablen sind als eine Summe von Indikatorvariablen binomialverteilt. Allerdings sind beide voneinander abhängig, denn es gilt $K_i + Z_i = i$.

Sei $D_i = K_i - Z_i$. Es gilt $D_i = K_i - (i - K_i) = 2K_i - i$ mit $W_{D_i} \subseteq [-i, i]$. Die Dichte von D_i ergibt sich aus der Dichte von K_i wie folgt.

$$f_{D_i}(j) = \begin{cases} f_{K_i}\left(\frac{j+i}{2}\right) & : -i \le j \le i \text{ und } j+i \text{ gerade} \\ 0 & : \text{ sonst} \end{cases}$$

Die Ereignisfolge $H_k = [D_k = 0]$, $k = 1, 2, \ldots$ ist rekurrent mit

$$\Pr[H_k] = \begin{cases} \binom{k}{k/2}\left(\frac{1}{3}\right)^{k/2}\left(\frac{2}{3}\right)^{k/2} & : \quad k \text{ gerade} \\ 0 & : \quad \text{sonst} \end{cases}$$

Wir betrachten die erzeugende Funktion $H(s) = \sum_{k=0}^{\infty} h_k s^k$ der Auftrittswahrscheinlichkeiten und zeigen, dass $H(1)$ konvergiert. Mit Korollar 83 der Vorlesung folgt dann, dass die Ereignisse H_k mit positiver Wahrscheinlichkeit nie eintreten. In der Sprache des Random Walk bedeutet dies, dass bei der Brown'schen Bewegung ein Teilchen nicht mit Wahrscheinlichkeit 1 zum Start zurückkehrt, falls die Wahrscheinlichkeiten der Links- bzw. Rechtsbewegung ungleich 1/2 sind.

Wir schätzen $H(1)$ wie folgt ab und benutzen dabei die Stirlingformel für die Näherung $\binom{2k}{k} \approx c_k \frac{2^{2k}}{\sqrt{k}}$ mit $c_k < c$ für ein existierendes positives c.

$$\begin{aligned}
H(1) &= \sum_{k=0}^{\infty} h_k \\
&= \sum_{k \in 2\mathbb{N}_0} \binom{k}{k/2} \left(\frac{1}{3}\right)^{k/2} \left(\frac{2}{3}\right)^{k/2} \\
&= \sum_{k=0}^{\infty} \binom{2k}{k} \left(\frac{1}{3}\right)^{k} \left(\frac{2}{3}\right)^{k} \\
&\approx \sum_{k=0}^{\infty} c_k \frac{2^{2k}}{\sqrt{k}} \left(\frac{2}{9}\right)^{k} = \sum_{k=0}^{\infty} c_k \frac{2^{2k}}{\sqrt{k}} \left(\frac{\sqrt{2}}{3}\right)^{2k} = \sum_{k=0}^{\infty} c_k \frac{1}{\sqrt{k}} \left(\frac{8}{9}\right)^{k} \\
&< c \cdot \sum_{k=0}^{\infty} \left(\frac{8}{9}\right)^{k} \\
&= c \cdot 9 < \infty .
\end{aligned}$$

$H(1)$ konvergiert also, d.h., $\Pr[Z = \infty] \neq 0$, wobei Z die Schritte bis zum erstmaligen Eintreten eines Ereignisses aus der Folge H_k zählt.

Wir widerlegen nun die betreffende Gleichung für $k = 1$.

Das Ereignis $E(i, k)$ ist gleichbedeutend mit $K_i < \frac{i}{2} + k$, d. h. $D_i = 2K_i - i < i + 2k - i = 2k$. In der Sprache der Brown'schen Bewegung für $d = 1$ bedeutet dies, dass das Teilchen nach dem i-ten Schritt auf der linken Seite von $2k$ liegt. Das Ereignis

$$E_n = \bigcap_{i=1}^{n} E(i, k)$$

heißt dann, dass sich das Teilchen vom Start bis zum n-ten Schritt stets auf der linken Seite von $2k$ bzw. 2 befindet.

Wir haben berechnet, dass das Teilchen mit positiver Wahrscheinlichkeit nicht mehr zum Nullpunkt zurückkehrt. Da die Wahrscheinlichkeit für die Linksbewegung größer ist als die Wahrscheinlichkeit einer Rechtsbewegung, gibt es auch eine positive Wahrscheinlichkeit, dass das Teilchen nach links geht und dann nie mehr zum Nullpunkt zurückkehrt. Dies heißt aber, dass das Teilchen mit positiver Wahrscheinlichkeit auf der linken Seite von 2 bleibt, im Widerspruch zu $\lim_{n\to\infty} \Pr[E_n] = 0$.

Tutoraufgabe 2

Wir benutzen die Funktion $h(t) = 0.027 + 0.0025 \cdot (t - 40)^2$ für $t \in \mathbb{R}$, um die „Sterblichkeitsrate" durch Lungenkrebs von Kettenraucherinnen abzuschätzen, die mindestens $t \geq 40$ Jahre alt sind. Ihre Lebensdauer sei X und es gelte

$$\Pr[X > t \mid X > 40] = \exp\left(-\int_{40}^{t} h(s)\mathrm{d}s\right) .$$

Wie groß ist die Wahrscheinlichkeit, dass eine 45-jährige Kettenraucherin mindestens 50 Jahre alt wird?

Lösung

Zu berechnen ist offenbar $\Pr[X \geq 50 \mid X \geq 45]$.
Bei stetigen Zufallsvariablen gilt $\Pr[X \geq 50 \mid X \geq 45] = \Pr[X > 50 \mid X > 45]$.
Für $t \geq s \geq 40$ berechnen wir allgemein die folgende bedingte Wahrscheinlichkeit.

$$\begin{aligned}
\Pr[X > t \mid X > s] &= \frac{\Pr[(X > t) \cap (X > s)]}{\Pr[X > s]} \\
&= \frac{\Pr[X > t]}{\Pr[X > s]} \\
&= \frac{\Pr[X > t \mid X > 40]}{\Pr[X > s \mid X > 40]} \\
&= \exp\left(- \int_{40}^{t} h(x)\ \mathrm{d}x + \int_{40}^{s} h(x)\ \mathrm{d}x \right) \\
&= \exp\left(- \int_{s}^{t} h(x)\ \mathrm{d}x \right) .
\end{aligned}$$

Daher gilt

$$\Pr[X > 50 \mid X > 45] \ = \ \exp\left(- \int_{45}^{50} 0.027 + 0.0025(t-40)^2\ \mathrm{d}t \right) ,$$

woraus wegen

$$\int_{45}^{50} 0.027 + 0.0025(t-40)^2\ \mathrm{d}t \ = \ 0.135 + 2.1875/3 \ = \ 0.86417\bar{6}$$

folgt, dass gilt

$$\Pr[X \geq 50 \mid X \geq 45] \ = \ \exp(-0.86417\bar{6}) \ \approx \ 42.1402575\ldots\% .$$

Tutoraufgabe 3

Sei $X = \sum_{i=1}^{2000} X_i$ die Summe der Augenzahlen, wenn man 2000-mal mit einem idealen Würfel würfelt.

1. Berechnen Sie näherungsweise $\Pr[7000 \leq X \leq 7100]$!
2. Wie groß muss man Δ wählen, damit $\Pr[7000 - \Delta \leq X \leq 7000 + \Delta] \approx \frac{1}{2}$ gilt?

Approximieren Sie die entsprechenden Wahrscheinlichkeiten mit Hilfe des Zentralen Grenzwertsatzes.

Lösung

1. Zunächst gilt für jedes i

   $$\mu = \mathbb{E}[X] = 2000 \cdot \mathbb{E}[X_i] = 2000 \cdot \frac{7}{2} = 7000$$

   und

   $$\mathrm{Var}[X] = 2000 \cdot \mathrm{Var}[X_i] = 2000 \cdot \frac{35}{12} = \frac{17500}{3} .$$

Mit

$$\sigma = \sqrt{\frac{17500}{3}} = 50 \cdot \sqrt{\frac{7}{3}}\,.$$

folgt

$$\begin{aligned}
\Pr[7000 \le X \le 7100] &= \Pr\left[\frac{7000-\mu}{\sigma} \le \frac{X-\mu}{\sigma} \le \frac{7100-\mu}{\sigma}\right] \\
&= \Pr\left[0 \le \frac{X-\mu}{\sigma} \le \sqrt{\frac{12}{7}}\right] \\
&\approx \Phi(1,309307341...) - \Phi(0) \\
&\approx 0,90478... - 0,5 \\
&= 0,40478... \quad .
\end{aligned}$$

Man beachte, dass $\frac{X-\mu}{\sigma}$ als standardnormalverteilt betrachtet und der Funktionswert $\Phi(1,309307341...)$ durch Interpolation aus den Tabellenwerten approximiert wurde. Außerdem wurde $0 < \frac{X-\mu}{\sigma}$ durch $0 \le \frac{X-\mu}{\sigma}$ ersetzt, weil Φ stetig ist.

2. Man beachte im Folgenden u. a. die Verwendung der Gleichung $\Phi(-z) = 1 - \Phi(z)$. Außerdem benützen wir wieder die Stetigkeit von Φ.

$$\begin{aligned}
\Pr[7000 - \Delta \le X \le 7000 + \Delta] &= \Pr\left[\frac{-\Delta}{\sigma} \le \frac{X-\mu}{\sigma} \le \frac{\Delta}{\sigma}\right] \\
&\approx \Phi\left(\frac{\Delta}{\sigma}\right) - \Phi\left(-\frac{\Delta}{\sigma}\right) \\
&= 2 \cdot \Phi\left(\frac{\Delta}{\sigma}\right) - 1\,.
\end{aligned}$$

Wir setzen nun $2 \cdot \Phi\left(\frac{\Delta}{\sigma}\right) - 1 = \frac{1}{2}$. Das bedeutet

$$\Phi\left(\frac{\Delta}{\sigma}\right) = 0,75\,.$$

Mit dem Quantil

$$z_{0,75} \approx 0,6745$$

erhalten wir

$$\Delta = z_{0,75} \cdot \sigma \approx 0,6745 \cdot 50 \cdot \sqrt{\frac{7}{3}} = 51,51788... \quad .$$

Bei der Bestimmung des Quantils aus der Tabelle von Φ haben wir eine Interpolation angewendet.