Technische Universität München
Fakultät für Informatik
Lehrstuhl für Theoretische Informatik
Prof. Dr. Susanne Albers
Dennis Kraft, Richard Stotz

Sommersemester 2015
Lösungsblatt 2
4. Mai 2015

Diskrete Wahrscheinlichkeitstheorie

Tutoraufgabe 1

Sie sind im Besitz zweier gezinkter Münzen a und b. Obwohl beide Münzen äußerlich identisch sind, zeigt a mit Wahrscheinlichkeit $\frac{2}{3}$ Kopf wohingegen b mit Wahrscheinlichkeit $\frac{1}{3}$ Kopf zeigt. Leider haben Sie vergessen, welche der Münzen welche ist. Um dies herauszufinden wählen Sie zufällig eine der beiden Münzen, wobei wir davon ausgehen, dass a und b gleich wahrscheinlich sind, und werfen diese n-mal.

1. Zeigen Sie, dass die Wahrscheinlichkeit für Kopf im i-ten Wurf $\frac{1}{2}$ ist.
2. Wie hoch ist die Wahrscheinlichkeit, dass Sie im n-ten Wurf Kopf werfen, wenn alle vorherigen Würfe bereits Kopf gezeigt haben? Wie verhält sich diese Wahrscheinlichkeit für n gegen unendlich?
3. Mit welcher Wahrscheinlichkeit haben Sie Münze a gewählt, wenn m Ihrer n Würfe Kopf gezeigt haben?

Lösungsvorschlag

Sei E das Ereignis, dass Sie Münze a gewählt haben. Die Wahrscheinlichkeit hierfür beträgt genau $\frac{1}{2}$. Des Weiteren bezeichnen wir das Ereignis, dass im i-ten Wurf Kopf gefallen ist, mit F_i. Nehmen wir nun an, dass Sie mit Münze a werfen. Dann ist die Wahrscheinlichkeit im i-ten Wurf Kopf zu erhalten $\frac{2}{3}$. Sollten Sie hingegen b gewählt haben, so ist diese Wahrscheinlichkeit lediglich $\frac{1}{3}$. Kombiniert man die beiden Ereignisse mit Hilfe des Satzes der totalen Wahrscheinlichkeit, so bestätigt sich der Wert von $\frac{1}{2}$ für $\Pr[F_i]$

$$\Pr[F_i] = \Pr[F_i \mid E] \cdot \Pr[E] + \Pr[F_i \mid \Omega \setminus E] \cdot \Pr[\Omega \setminus E] = \frac{1}{2} \cdot \left(\frac{2}{3} + \frac{1}{3}\right) = \frac{1}{2}.$$

In der zweiten Teilaufgabe ist nach der $\Pr[F_n \mid \bigcap_{i=1}^{n-1} F_i]$ gefragt. Gemäß der Definition bedingter Wahrscheinlichkeiten lässt sich dieser Ausdruck umformen zu

$$\Pr\left[F_n \;\middle|\; \bigcap_{i=1}^{n-1} F_i\right] = \frac{\Pr[\bigcap_{i=1}^{n} F_i]}{\Pr[\bigcap_{i=1}^{n-1} F_i]}.$$

Ähnlich wie vorher können wir nun die Wahrscheinlichkeit n-mal hintereinander Kopf zu werfen mit dem Satz der totalen Wahrscheinlichkeit aufspalten zu

$$\Pr\left[\bigcap_{i=1}^{n} F_i\right] = \Pr\left[\bigcap_{i=1}^{n} F_i \;\middle|\; E\right] \Pr[E] + \Pr\left[\bigcap_{i=1}^{n} F_i \;\middle|\; \Omega \setminus E\right] \cdot \Pr[\Omega \setminus E].$$

Wird mit a geworfen, so ist die gesuchte Wahrscheinlichkeit gleich $(\frac{2}{3})^n$. Andernfalls beträgt sie $(\frac{1}{3})^n$. Dasselbe Prinzip, lässt sich natürlich auch auf $n-1$ Würfe anwenden. Insgesamt ergibt sich daraus

$$\Pr\left[F_n \;\middle|\; \bigcap_{i=1}^{n-1} F_i\right] = \frac{(\frac{2}{3})^n \cdot \frac{1}{2} + (\frac{1}{3})^n \cdot \frac{1}{2}}{(\frac{2}{3})^{n-1} \cdot \frac{1}{2} + (\frac{1}{3})^{n-1} \cdot \frac{1}{2}} = \frac{(\frac{2}{3})^n + (\frac{1}{3})^n}{(\frac{2}{3})^{n-1} + (\frac{1}{3})^{n-1}}.$$

Da das häufige Auftreten von Kopf ein Indiz dafür ist, dass Sie Münze a gewählt haben, sollte $\Pr[E_n \mid \bigcap_{i=1}^{n-1} E_i]$ für n gegen unendlich gleich $\frac{2}{3}$ sein. In der Tat lässt sich dies einfach nachrechnen

$$\lim_{n\to\infty} \frac{(\frac{2}{3})^n + (\frac{1}{3})^n}{(\frac{2}{3})^{n-1} + (\frac{1}{3})^{n-1}} = \lim_{n\to\infty} \frac{\frac{2}{3} + \frac{1}{3}2^{1-n}}{1 + 2^{1-n}} = \frac{2}{3}.$$

Kommen wir nun zum dritten und letzten Punkt. Gesucht ist die Wahrscheinlichkeit $\Pr[E \mid G_m]$, wobei G_m das Ereignis bezeichnet, dass exakt m der n Würfe Kopf zeigen. Mit dem Satz von Bayes können wir die Bedingung umkehren zu

$$\Pr[E \mid G_m] = \frac{\Pr[G_m \mid E] \cdot \Pr[E]}{\Pr[G_m \mid E] \cdot \Pr[E] + \Pr[F_m \mid \Omega \setminus E] \cdot \Pr[\Omega \setminus E]}.$$

Die Wahrscheinlichkeit von G_m unter der Bedingung, dass a bzw. b gewählt wurde ist gegeben durch

$$\binom{n}{m} \cdot \left(\frac{2}{3}\right)^m \cdot \left(\frac{1}{3}\right)^{n-m} \text{ bwz. } \binom{n}{m} \cdot \left(\frac{1}{3}\right)^m \cdot \left(\frac{2}{3}\right)^{n-m}.$$

Wenn wir diese Wahrscheinlichkeiten einsetzen erhalten wir schließlich

$$\begin{aligned}
\Pr[E \mid G_m] &= \frac{\binom{n}{m} \cdot (\frac{2}{3})^m \cdot (\frac{1}{3})^{n-m} \cdot \frac{1}{2}}{\binom{n}{m} \cdot (\frac{2}{3})^m \cdot (\frac{1}{3})^{n-m} \cdot \frac{1}{2} + \binom{n}{m} \cdot (\frac{1}{3})^m \cdot (\frac{2}{3})^{n-m} \cdot \frac{1}{2}} \\
&= \frac{(\frac{2}{3})^m \cdot (\frac{1}{3})^{n-m}}{(\frac{2}{3})^m \cdot (\frac{1}{3})^{n-m} + (\frac{1}{3})^m \cdot (\frac{2}{3})^{n-m}} \\
&= \frac{1}{1 + (\frac{1}{3})^m \cdot (\frac{2}{3})^{n-m} \cdot (\frac{2}{3})^{-m} \cdot (\frac{1}{3})^{-n+m}} \\
&= \frac{1}{1 + 2^{n-2m}}.
\end{aligned}$$

Tutoraufgabe 2

Eine Familie von Ereignissen E_i wird paarweise unabhängig genannt, wenn für alle Indizes $i \neq j$ die Wahrscheinlichkeit $\Pr[E_i \cap E_j]$ gleich dem Produkt $\Pr[E_i] \cdot \Pr[E_j]$ ist. Beweisen oder widerlegen Sie, dass eine Familie paarweise unabhängiger Ereignisse unabhängig ist.

Lösungsvorschlag

Obwohl die Definition von Unabhängigkeit trivialerweise auch paarweise Unabhängigkeit impliziert, gilt die Gegenrichtung im Allgemeinen nicht. Als Gegenbeispiel betrachten wir einen diskreten Wahrscheinlichkeitsraum $(\Omega, \Pr)$ mit der Ergebnismenge $\{\omega_1, \omega_2, \omega_3, \omega_4\}$, wobei jedes Elementarereignis gleich wahrscheinlich ist. Zusätzlich definieren wir die Ereignisse E_1, E_2 und E_3 als

$$E_1 = \{\omega_1, \omega_4\},\ E_2 = \{\omega_2, \omega_4\} \text{ und } E_3 = \{\omega_3, \omega_4\}.$$

Man beachte, dass die Wahrscheinlichkeit der einzelnen Ereignisse jeweils $\frac{1}{2}$ ist. Außerdem haben alle Paare von Ereignissen genau ein Elementarereignis gemeinsam, nämlich ω_4. Die Wahrscheinlichkeit eines Schnitts zweier Ereignisse ist somit stets $\frac{1}{4}$. Folglich sind die Ereignisse paarweise unabhängig. Andererseits gilt

$$\Pr[E_1 \cap E_2 \cap E_3] = \Pr[\omega_4] = \frac{1}{4} \neq \frac{1}{8} = \Pr[E_1] \cdot \Pr[E_2] \cdot \Pr[E_3],$$

weshalb die Ereignisse insgesamt nicht unabhängig sind.

Tutoraufgabe 3

Da die Übungen in Diskreter Wahrscheinlichkeitstheorie aufeinander aufbauen, hängt der Lernerfolg eines Studenten unter anderem von seinen bisherigen Leistungen ab. Sollte der Student in der letzten Woche seine Hausaufgaben erfolgreich gelöst haben, so gehen wir davon aus, dass er dies auch in der aktuellen Woche mit Wahrscheinlichkeit $\frac{3}{4}$ schafft. Ist er in der vorherigen Woche jedoch nicht erfolgreich, so sinken seine Chancen auf $\frac{1}{4}$. Beweisen Sie, dass ein Student, der in der ersten Woche mit Wahrscheinlichkeit $\frac{3}{4}$ erfolgreich ist, seine Hausaufgaben auch in Woche n mit Wahrscheinlichkeit $(\frac{1}{2})^{n+1} + \frac{1}{2}$ erfolgreich bearbeitet.

Lösungsvorschlag

Sei E_i das Ereignis, dass der Student in Woche i erfolgreich seine Hausaufgaben bearbeitet hat. Gemäß dem dem Satz der totalen Wahrscheinlichkeit, lässt sich $\Pr[E_i]$ rekursiv darstellen als

$$\begin{aligned}
\Pr[E_i] &= \Pr[E_i \mid E_{i-1}] \cdot \Pr[E_{i-1}] + \Pr[E_i \mid \Omega \setminus E_{i-1}] \cdot \Pr[\Omega \setminus E_{i-1}] \\
&= \frac{3}{4} \cdot \Pr[E_{i-1}] + \frac{1}{4} \cdot (1 - \Pr[E_{i-1}]) \\
&= \frac{1}{2} \cdot \Pr[E_{i-1}] + \frac{1}{4}
\end{aligned}$$

für alle $i > 1$. Durch Induktion können wir nun zeigen, dass die Wahrscheinlichkeit von $\Pr[n]$ tatsächlich $(\frac{1}{2})^{n+1} + \frac{1}{2}$ entspricht. Da die Erfolgschancen des Studenten in der ersten Woche $\frac{3}{4}$ betragen, ist der Basisfall für $n = 1$ trivialerweise erfüllt

$$\Pr[E_1] = \frac{3}{4} = \left(\frac{1}{2}\right)^{1+1} + \frac{1}{2}.$$

Nehmen wir nun an, dass wir Gleichheit für n bereits bewiesen haben und schließen daraus auf Gleichheit für den Fall $n + 1$

$$\Pr[E_{n+1}] = \frac{1}{2} \cdot \Pr[E_n] + \frac{1}{4} \stackrel{\text{I.H.}}{=} \frac{1}{2} \cdot \left(\left(\frac{1}{2}\right)^{n+1} + \frac{1}{2}\right) + \frac{1}{4} = \left(\frac{1}{2}\right)^{n+2} + \frac{1}{2}.$$

Folglich ist die Erfolgswahrscheinlichkeit des Studenten in Woche n wie vermutet $(\frac{1}{2})^{n+1} + \frac{1}{2}$.

Hausaufgabe 1 (5 Punkte)

Nehmen wir an, Sie züchten eine Bakterienkolonie in einer Petrischale. Von einer Generation auf die nächste kann sich jedes Bakterium entweder in zwei neue Bakterien teilen oder es stirbt. Beide Ereignisse treten mit der gleichen Wahrscheinlichkeit und unabhängig vom Verhalten der anderen Bakterien ein. Wie groß ist die Wahrscheinlichkeit, dass Ihre Kolonie in der vierten Generation nicht ausgestorben ist, wenn die nullte Generation aus einer Bakterie besteht? Protokollieren Sie Ihre Rechnung.

Lösungsvorschlag

Sei $E_{n,m}$ das Ereignis, dass die Kolonie in der n-ten Generation aus m Bakterien beseht. Für $n \geq 1$ lässt sich die Wahrscheinlichkeit $\Pr[E_{n,m}]$ rekursive aus dem Satz der totalen Wahrscheinlichkeit bestimmen

$$\Pr[E_{n,m}] = \sum_{i=0}^{\infty} \Pr[E_{n,m}|E_{n-1,i}] \cdot \Pr[E_{n-1,i}].$$

Die Wahrscheinlichkeit, dass eine Kolonie in Generation n aus m Bakterien besteht wenn die Elterngeneration aus i Bakterien besteht ist wiederum gegeben durch

$$\Pr[E_{n,m}|E_{n-1,i}] = \binom{i}{\frac{m}{2}} \cdot \left(\frac{1}{2}\right)^i$$

für $i \geq \frac{m}{2}$ und gerade m. Andernfalls ist die Wahrscheinlichkeit 0. Da unsere Kolonie in der nullten Generation aus einer einzigen Bakterie besteht und sich die Anzahl der Bakterien mit jedem Generationswechsel höchstens verdoppelt, ist die maximale Größe einer Population in der n-ten Generation gleich 2^n. Zusätzliche lässt sich beobachten, dass jede Generation abgesehen von der nullten aus einer geraden Anzahl von Bakterien bestehen muss. Für die erste bis dritte Generationen berechnen wir somit die folgenden Wahrscheinlichkeitswerte.

$\Pr[E_{n,m}]$	1	2	3
0	$\frac{1}{2}$	$\frac{5}{8}$	$\frac{89}{128}$
2	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{5}{32}$
4	0	$\frac{1}{8}$	$\frac{7}{64}$
6	0	0	$\frac{1}{32}$
8	0	0	$\frac{1}{128}$

Aus dieser Tabelle lässt sich nun unmittelbar die Wahrscheinlichkeit mit der die Bakterienkolonie in der vierten Generation noch nicht ausgestorben ist herleiten

$$\Pr[\Omega \setminus E_{4,0}] = 1 - \left(\frac{89}{1 \cdot 128} + \frac{5}{4 \cdot 32} + \frac{7}{16 \cdot 64} + \frac{1}{64 \cdot 32} + \frac{1}{256 \cdot 128}\right) = \frac{8463}{32768} \approx 0{,}25827.$$

Hausaufgabe 2 (5 Punkte)

Wir betrachten einen Hashfunktion, die n Schlüssel auf m Hashwerte abbildet. Vereinfachend gehen wir davon aus, dass der Hashwert eines Schlüssel zufällig gewählt wird, wobei jeder Hashwert gleich wahrscheinlich ist. Sollte der selbe Hashwert mehreren Schlüsseln zugeordnet werden, sprechen wir von einer Kollision. Für effizientes Hashing ist es wichtig, Kollisionen zwischen Schlüsseln zu vermeiden. Da dies für $n > m$ nicht möglich ist, betrachten wir im Folgenden stets den Fall $n \leq m$.

1. Wir nummerieren die Schlüssel von 1 bis n und bezeichnen E_i als das Ereignis, dass der i-te Schlüssel mit keinem Schlüssel eines kleineren Index kollidiert. Bestimmen Sie $\Pr[E_1]$ und $\Pr[E_i \mid \bigcap_{j=1}^{i-1} E_j]$ für $i > 1$.

2. Zeigen Sie, dass die Wahrscheinlichkeit mit der es zu keiner Kollisionen kommt mindestens $(1 - \frac{n-1}{m})^{n-1}$ ist.

3. Wie groß sollte m mindestens gewählt werden, damit beim Einfügen von 1000 Schlüsseln die Wahrscheinlichkeit einer Kollision höchstens $\frac{1}{10}$ ist?

Lösungsvorschlag

Da der Hashwert des ersten Schlüssels noch nicht vergeben sein kann gilt $\Pr[E_1] = 1$. Das Ereignis $\bigcap_{j=1}^{i-1} E_j$ entspricht der Situation, dass die ersten $i-1$ Schlüssel allesamt unterschiedliche Hashwerte haben. Damit sind für den i-ten Schlüssel noch $m-(i-1)$ Hashwert frei was bedeutet, dass

$$\Pr\left[E_i \,\middle|\, \bigcap_{j=1}^{i-1} E_j\right] = \frac{m-(i-1)}{m} = 1 - \frac{i-1}{m},$$

für $i > 1$. Kommen wir nun zum zweiten Punkt. Die Wahrscheinlichkeit einer kollisionsfreien Zuordnung lässt sich darstellen als $\Pr[\bigcap_{i=1}^{n} E_i]$. Gemäß des Multiplikationssatzes formen wir diese Wahrscheinlichkeit um zu

$$\Pr\left[\bigcap_{i=1}^{n} E_i\right] = \Pr[E_1] \cdot \prod_{i=2}^{n} \Pr\left[E_i \,\middle|\, \bigcap_{j=1}^{i-1} E_j\right] = \prod_{i=2}^{n} \Pr\left[E_i \,\middle|\, \bigcap_{j=1}^{i-1} E_j\right].$$

Mit den Ergebnissen der ersten Teilaufgabe und der Beobachtung, dass der Term $1 - \frac{i-1}{m}$ für $i = n$ minimal ist, können wir nun die geforderte Abschätzung treffen

$$\Pr\left[\bigcap_{i=1}^{n} E_i\right] = \prod_{i=2}^{n} \left(1 - \frac{i-1}{m}\right) \geq \left(1 - \frac{n-1}{m}\right)^{n-1}.$$

Für die letzte Teilaufgabe setzen wir $n = 1000$ und suchen ein m, dass die folgenden Ungleichungen erfüllt

$$\Pr\left[\bigcap_{i=1}^{1000} E_i\right] \geq \left(1 - \frac{999}{m}\right)^{999} \geq \frac{9}{10}.$$

Durch Umformen der zweiten Ungleichung ergibt sich

$$m \geq 999 \cdot \left(1 - \left(\frac{9}{10}\right)^{\frac{1}{999}}\right)^{-1} \approx 9472748{,}13787.$$

Um sicherzustellen, dass eine Kollision höchstens mit Wahrscheinlichkeit $\frac{1}{10}$ auftritt, können wir für m also den Wert 9472749 wählen.

Hausaufgabe 3 (5 Punkte)

Ein Tutor für Diskrete Wahrscheinlichkeitstheorie betreut zwei Übungsgruppen, eine Vormittags und eine Nachmittags. In der Vormittagsgruppe studieren 12 der Teilnehmer Informatik und 3 Games Engineering. Die Nachmittagsgruppe besteht hingegen aus 11 Informatikstudenten, 4 Games Engineering Studenten und 2 Bioinformatikstudenten. Angenommen wir wählen zufällig und mit gleicher Wahrscheinlichkeit jeweils einen Teilnehmer aus beiden Gruppen und erhalten dabei sowohl einen Informatiker als auch einen Games Engineerer. Mit welcher Wahrscheinlichkeit ist der Informatikstudent aus der Vormittagsgruppe?

Lösungsvorschlag

Um unsere Notation zu vereinfachen bezeichnen wir zufällig eine der Übungsgruppen mit a und die andere mit b. Des Weiteren definieren wir E als das Ereignis, dass a die Vormittagsgruppe ist, und F als das Ereignis, dass ein Informatikstudent aus a und ein Games Engineering Student aus b gewählt wird. Gesucht ist also die Wahrscheinlichkeit $\Pr[E \mid F]$. Mit dem Satz von Bayes erhalten wir

$$Pr[E \mid F] = \frac{Pr[F \mid E] \cdot \Pr[E]}{Pr[F \mid E] \cdot \Pr[E] + Pr[F \mid \Omega \setminus E] \cdot \Pr[\Omega \setminus E]}.$$

Da unsere Benennung der Gruppen willkürlich erfolgt, ist die Wahrscheinlichkeit $\Pr[E]$ gleich $\frac{1}{2}$. Außerdem ist die Wahrscheinlichkeit mit der ein Informatiker aus a und ein Games Engineerer aus b gezogen wird unter der Bedingung, dass es sich bei a um die Vormittagsgruppe bzw. Nachmittagsgruppe handelt, gegeben durch

$$\Pr[F \mid E] = \frac{12}{12+3} \cdot \frac{4}{11+4+2} = \frac{16}{85} \text{ bzw. } \Pr[F \mid \Omega \setminus E] = \frac{3}{12+3} \cdot \frac{11}{11+4+2} = \frac{11}{85}.$$

Wenn wir diese Werte nun in unsere Formel einsetzen erhalten wir

$$\Pr[E \mid F] = \frac{\frac{16}{85} \cdot \frac{1}{2}}{\frac{16}{85} \cdot \frac{1}{2} + \frac{11}{85} \cdot (1 - \frac{1}{2})} = \frac{16}{27} \approx 0{,}59259.$$

Hausaufgabe 4 (5 Punkte)

Sie werfen eine Münze, die Kopf und Zahl jeweils mit der gleichen Wahrscheinlichkeit zeigt, n mal nacheinander, wobei $n \geq 2$ ist. Sei E das Ereignis, dass sowohl Kopf als auch Zahl geworfen wird und F das Ereignis, dass höchstens einmal Kopf geworfen wird. Zeigen Sie, dass die Ereignisse E und F ausschließlich für $n = 3$ unabhängig sind.

Hinweis: Die Ungleichung $2n + 2 < 2^n$ für $n > 3$ darf ohne Beweis verwendet werden.

Lösungsvorschlag

Die Ereignisse E und F sind nach Definition unabhängig, wenn die folgende Bedingung erfüllt ist:

$$\Pr[E \cap F] = \Pr[E] \cdot \Pr[F].$$

In unserem Fall besteht der Schnitt von E und F aus allen Folgen von mindestens 2 Münzwürfen, die exakt einmal Kopf zeigen. Da es dafür insgesamt n Möglichkeiten gibt, tritt $E \cap F$ mit Wahrscheinlichkeit $n(\frac{1}{2})^n$ ein. Das Ereignis F enthält zusätzlich noch den Fall, dass ausschließlich Zahl geworfen wird, woraus sich eine Wahrscheinlichkeit von

$(n+1)(\frac{1}{2})^n$ ergibt. Schließlich müssen wir noch die Wahrscheinlichkeit von E bestimmen. Dies gelingt am einfachsten über das komplementäre Ereignis, also jenen Wurfsequenzen, die entweder nur aus Kopf oder nur aus Zahl bestehen. Die Wahrscheinlichkeit von E ist demnach $1 - 2(\frac{1}{2})^n$. Wenn wir diese Werte nun in unsere Gleichung einsetzen erhalten wir durch Umformung

$$
\begin{aligned}
& n\left(\frac{1}{2}\right)^n = \left(1 - 2\left(\frac{1}{2}\right)^n\right) \cdot (n+1)\left(\frac{1}{2}\right)^n \\
\Longleftrightarrow \quad & n = \left(1 - 2\left(\frac{1}{2}\right)^n\right) \cdot (n+1) \\
\Longleftrightarrow \quad & n = n + 1 - (2n+2)\left(\frac{1}{2}\right)^n \\
\Longleftrightarrow \quad & 1 = (2n+2) \cdot \left(\frac{1}{2}\right)^n \\
\Longleftrightarrow \quad & 2n + 2 = 2^n.
\end{aligned}
$$

Für $n = 3$ lässt sich leicht nachrechnen, dass die Gleichung erfüllt ist. Für $n = 2$ ist die Gleichung dagegen nicht erfüllt. Da die Exponentialfunktion 2^n für alle $n > 3$ echt größer als $2n + 2$ ist, können auch keine weiteren n die Gleichung erfüllen. Somit folgt, dass E und F genau dann unabhängig sind wenn $n = 3$ ist.