# Übungen zu Diskrete Wahrscheinlichkeitstheorie

Dieses Blatt wird am besprochen.

## Aufgabe 2.1

In der Sendung „Wer wird Millionär" werden bekanntlich dem Kandidaten vier mögliche Antworten auf eine Frage gegeben, aus welchen er die korrekte Antwort bestimmen muss. Einmalig kann der Kandidat auch im Spielverlauf den sogenannten "50-50-Joker" einsetzen, um zwei falsche Antworten zu identifizieren und aus dem Spiel zu nehmen.

Wir gehen von einem Kandidaten aus, der mit Wahrscheinlichkeit $p$ die korrekte Antwort errät. Wenn er somit nicht den 50-50-Joker benutzt, sondern seinem Wissen vertraut, so gewinnt er mit Wahrscheinlichkeit $p$.

Nehmen Sie an, dass der Kandidat sich für eine Antwort entschieden hat, nun jedoch den "50-50-Joker" verwendet, um sich abzusichern, und dabei folgende Strategie verfolgt:

i) Wird durch den 50-50-Joker die vom Kandidaten gewählte Antwort nicht entfernt, so bleibt er bei dieser;

ii) ansonsten wählt er zufällig aus den beiden verbleibenden Antworten.

Ihre Aufgaben sind nun:

a) Analysieren Sie, mit welcher Wahrscheinlichkeit der Kandidat unter Verwendung dieser Strategie gewinnt.

b) Bestimmen Sie auch seine Gewinnwahrscheinlichkeit, wenn er sich stattdessen – ähnlich der optimalen Strategie im „Ziegenproblem" – im Fall i) für die andere verbleibende Antwort entscheidet.

c) Vergleichen Sie beide Strategien, welche ist die bessere?

## Lösungsvorschlag

a) Mit W'keit $p$ entscheidet sich der Kandidat für die korrekte Antwort. Damit gewinnt er in diesem Fall mit der ersten Strategie mit (bed.) W'keit $p$.

Entscheidet er sich jedoch für die falsche Antwort (mit W'keit $1-p$), so entfernt der Joker mit (bed.) W'keit 2/3 seine gewählte Anwort, womit er sich dann mit (bed.) W'keit 1/2 zufällig für die richtige unter den verbleibenden zwei entscheidet.

Insgesamt ergibt sich somit die W'keit $p + (1-p)\frac{2}{3}\frac{1}{2} = \frac{1}{3} + \frac{2}{3}p$. Der Joker sichert somit tatsächlich den Spieler ab – unter Verwendung dieser Strategie, da $p < \frac{1}{3} + \frac{2}{3}p$ für $p < 1$.

b) Wegen dem Umentscheiden des Kandidatens kann dieser nur gewinnen, falls er zunächst die falsche Antwort wählt (W'keit $1-p$). Entfernt der Joker nun die gewählte Antwort (mit W'keit 2/3), so gewinnt er mit (bed.) W'keit 1/2 (insgesamt $(1-p)\frac{1}{3}$); bleibt seine gewählte (falsche) Antwort jedoch stehen (mit W'keit 1/3), so gewinnt er mit bed. W'keit 1, da er sich dann für die korrekte Antwort umentscheidet.

Insgesamt ergibt sich also die Gewinnw'keit $\frac{2}{3}(1-p)$.

c) Hat der Kandidat eine Trefferchance von $p \leq 1/4$ (er überschätzt sein Wissen, versteht die Frage falsch, o.ä.), so ist die zweite Strategie für ihn die bessere; ansonsten fährt er für $p > 1/4$ mit der ersten Strategie stets besser.

## Aufgabe 2.2

Alice möchte Bob eine Nachricht schicken. Die Nachricht besteht aus $n$ zufälligen $a$s und $b$s, d.h. sie ist ein Wort $w \in \{a,b\}^n$. Um die Nachricht zu übermitteln, hat sie einen Kanal zur Verfügung, der nur Nullen und Einsen überträgt. Außerdem ist der Kanal rauschbehaftet, d.h. es geht kein Bit verloren, aber mit einer gewissen Wahrscheinlichkeit $f \in (0, \frac{1}{2})$ überträgt der Kanal den falschen Wert.

Wenn zur Übermittlung der Nachricht $m$ Bits übertragen werden, bezeichnen wir die von Alice gesendeten Bits mit $x_1 \cdots x_m$ und die von Bob empfangenen Bits mit $y_1 \cdots y_m$. Der Ergebnismenge $\Omega$ bei diesem Experiment ist die Menge der möglichen Folgen $x_1 \cdots x_m y_1 \cdots y_m$ ausgesandter und empfangener Bits. Die bedingten Wahrscheinlichkeiten für jedes Bit $i = 1, \ldots, m$ sind unabhängig voneinander:

$$\begin{array}{rclrcl} \Pr[y_i = 0 \mid x_i = 0] & = & 1-f & \Pr[y_i = 0 \mid x_i = 1] & = & f \\ \Pr[y_i = 1 \mid x_i = 0] & = & f & \Pr[y_i = 1 \mid x_i = 1] & = & 1-f \end{array}$$

a) Nehmen Sie an, dass Alice für jedes $a$ eine 0 schickt, für jedes $b$ eine 1, d.h. es gilt $m = n$. Bob setzt die erhaltene Nachricht dementsprechend wieder zusammen. Das Ereignis $E$ enthalte alle Elementarereignisse, bei denen Bob die Nachricht korrekt erhält. Berechnen Sie $\Pr[E]$.

b) Um sich besser gegen Übermitlungsfehler abzusichern, ändert Alice die Übertragungsmethode: Für jedes $a$, das sie senden will, sendet sie nun 000 über den Kanal, entsprechend sendet sie 111 für ein $b$ (d.h. es gilt $m = 3n$). Bob interpretiert eine Folge von drei Bits als $a$, wenn sie mehr Nullen als Einsen enthält, sonst als $b$. Sei $E$ analog zu a) definiert. Berechnen Sie erneut $\Pr[E]$.

c) Zeigen Sie, dass die Dekodierung, die Bob in b) verwendet, in folgendem Sinne optimal ist: Er interpretiert einen beliebigen Drei-Bit-Block $z_1z_2z_3$ genau dann als ein $a$, wenn $\Pr[x_1x_2x_3 = 000 \mid y_1y_2y_3 = z_1z_2z_3] > \Pr[x_1x_2x_3 = 111 \mid y_1y_2y_3 = z_1z_2z_3]$ gilt.

## Lösungsvorschlag

a) Für eine beliebige Nachricht $b_1 \ldots b_n \in \{0,1\}^n$ gilt

$$\begin{array}{rl} & \Pr[x_1 \ldots x_n = b_1 \ldots b_n \wedge y_1 \ldots y_n = b_1 \ldots b_n] \\ = & \Pr[x_1 \ldots x_n = b_1 \ldots b_n]\Pr[y_1 \ldots y_n = b_1 \ldots b_n \mid x_1 \ldots x_n = b_1 \ldots b_n] \\ = & \frac{1}{2^n} \prod_{i=1}^n \Pr[y_i = b_i \mid x_i = b_i] \\ = & \frac{1}{2^n}(1-f)^n. \end{array}$$

bzw.

$$\Pr[x_1 \ldots x_n = b_1 \ldots b_n \wedge y_1 \ldots y_n = b_1 \ldots b_n \mid x_1 \ldots x_n = b_1 \ldots b_n] = (1-f)^n.$$

Damit ist die W'keit, dass eine Nachricht der Länge $n$ fehlerfrei übertragen wird

$$\begin{array}{rl} & \Pr[\{b_1 \ldots b_n b_1 \ldots b_n \mid b_i \in \{0,1\}\}] \\ = & \frac{1}{2^n}(1-f)^n \left|\{b_1 \ldots b_n b_1 \ldots b_n \mid b_i \in \{0,1\}\}\right| \\ = & (1-f)^n \end{array}$$

b) Das Ereignis, dass Bob ein $a$ empfängt, sei mit $B = a$ abgekürzt, entsprechend bezeichne $A = b$ das Ereignis, dass Alice ein $a$ sendet, d.h. 000 über den Kanal verschickt.

Dann gilt

$$\begin{array}{rl} & \Pr[A = a, B = a] \\ = & \Pr[x_1x_2x_3 = 000, y_1y_2y_3 \in \{000, 001, 010, 100\}] \\ = & \Pr[x_1x_2x_3 = 000, y_1y_2y_3 = 000] \\ + & \Pr[x_1x_2x_3 = 000, y_1y_2y_3 = 001] + \Pr[x_1x_2x_3 = 000, y_1y_2y_3 = 010] + \Pr[x_1x_2x_3 = 000, y_1y_2y_3 = 100] \\ = & \frac{(1-f)^3 + 3(1-f)^2 f}{8}. \end{array}$$

bzw.

$$\Pr[B = a \mid A = a] = (1-f)^3 + 3(1-f)^2 f.$$

Entsprechend folgt auch $\Pr[B = b \mid A = b] = (1-f)^3 + 3(1-f)^2 f$.

Es folgt

$$\begin{aligned}\Pr[E] &= \Pr[\{A = w, B = w \mid w \in \{a,b\}^n\}] \\ &= \textstyle\sum_{w\in\{a,b\}^n} \Pr[B = w \mid A = w]\Pr[A = w] \\ &= \textstyle\sum_{w\in\{a,b\}^n} \frac{1}{2^n} \prod_{i=1}^n \Pr[B = w_i \mid A = w_i] \\ &= \left((1-f)^3 + 3(1-f)^2 f\right)^n .\end{aligned}$$

c) Man muss einfach zeigen, dass Bob $b_1b_2b_3$ genau dann als ein $a$ interpretiert, wenn

$$\Pr[x_1x_2x_3 = 000|y_1y_2y_3 = b_1b_2b_3] > \Pr[x_1x_2x_3 = 111|y_1y_2y_3 = b_1b_2b_3]$$

gilt.

Letzteres schreibt man am besten als (mit der Konvention $\frac{1}{0} > 1$)

$$\begin{aligned} & \tfrac{\Pr[x_1x_2x_3=000|y_1y_2y_3=b_1b_2b_3]}{\Pr[x_1x_2x_3=111|y_1y_2y_3=b_1b_2b_3]} > 1 \\ \Leftrightarrow\ & \frac{\Pr[y_1y_2y_3=b_1b_2b_3|x_1x_2x_3=000]\frac{\Pr[x_1x_2x_3=000]}{\Pr[y_1y_2y_3=b_1b_2b_3]}}{\Pr[y_1y_2y_3=b_1b_2b_3|x_1x_2x_3=111]\frac{\Pr[x_1x_2x_3=111]}{\Pr[y_1y_2y_3=b_1b_2b_3]}} > 1 \\ \Leftrightarrow\ & \tfrac{\Pr[y_1y_2y_3=b_1b_2b_3|x_1x_2x_3=000]}{\Pr[y_1y_2y_3=b_1b_2b_3|x_1x_2x_3=111]} \tfrac{\Pr[A=a]}{\Pr[A=b]} > 1 \\ \Leftrightarrow\ & \tfrac{\Pr[y_1y_2y_3=b_1b_2b_3|x_1x_2x_3=000]}{\Pr[y_1y_2y_3=b_1b_2b_3|x_1x_2x_3=111]} > 1 \end{aligned}$$

Nun kann man die acht Fälle nacheinander überprüfen, z.b.:

$b_1b_2b_3 = 101$ Bob interpretiert dies als $b$.

Andererseits gilt auch

$$\frac{\Pr[y_1y_2y_3 = 101 \mid x_1x_2x_3 = 000]}{\Pr[y_1y_2y_3 = 101 \mid x_1x_2x_3 = 111]} = \frac{(1-f)f^2}{(1-f)^2 f} = \frac{f}{1-f}.$$

Wegen $0 < f < 0.5$ ist letzteres $< 1$. Also wird 101 auch bzgl. des zweiten Kriteriums als $b$ interpretiert.

Entsprechend folgt auch in den restlichen sieben Fällen, dass beide Kriterien dasselbe Ergebnis liefern.

**Aufgabe 2.3**

Professor Esparza hält eine Reihe von mündlichen Prüfungen ab. Wenn ein Student die Prüfung besteht, verbessert sich seine Laune, und der nächste Student besteht mit Wahrscheinlichkeit $p_1$, wobei $p_1 > \frac{1}{2}$ ist. Wenn ein Student die Prüfung nicht besteht, so verdüstert sich seine Miene, und der nächste Student besteht nur noch mit Wahrscheinlichkeit $p_2 < \frac{1}{2}$. Morgens ist Professor Esparza stets gut gelaunt, so dass der erste Student mit Sicherheit besteht.

Mit $A_n$ bezeichnen wir die Wahrscheinlichkeit, dass der $n$-te Student die Prüfung besteht; das Komplement $A_n^c$ bezeichnet dementsprechend den Fall, dass der $n$-te Student durchfällt. Als Abkürzung sei $q_n := \Pr[A_n]$ festgelegt.

(a) Stellen Sie die bedingten Wahrscheinlichkeiten $\Pr[A_{n+1} \mid A_n]$, $\Pr[A_{n+1}^c \mid A_n]$, $\Pr[A_{n+1} \mid A_n^c]$ und $\Pr[A_{n+1}^c \mid A_n^c]$ für $n \geq 1$ auf.

(b) Berechnen Sie mit Hilfe der bedingten Wahrscheinlichkeiten aus a) $q_{n+1}$ in Abhängigkeit von $q_n$.

(c) Lösen Sie die in b) erhaltene Rekursionsgleichung auf, d.h. ermitteln Sie eine geschlossene Form für $q_n$. Berechnen Sie außerdem $q := \lim_{n\to\infty} q_n$, d.h. den Wert, gegen den die Bestehenswahrscheinlichkeit konvergiert.

(d) Wie ändert sich $q$, falls der erste Student stets durchfällt?

## Lösungsvorschlag

Wir gehen davon aus, dass Prof. Esparza insgesamt $N > 1$ Prüfungen abhält. Als Ergebnismenge kann dann

$$\Omega = \{0,1\}^N$$

gewählt werden, wobei die $n$.te Komponente des Elementarereignisses $(s_1, \ldots, s_N) \in \Omega$ genau dann gleich 1 ist, falls der $i$.te Student die Prüfung bestanden hat

Die W'keit des Elementarereignisses $(s_1, s_2, \ldots, s_N)$ entsprechend der "Pfadregel" für mehrstufige W'keitsexperimente zu

$$\Pr[(s_1,\ldots,s_N)] = s_1 \cdot \prod_{n=2}^{N} \left(s_{n-1}s_n p_1 + s_{n-1}(1-s_n)(1-p_1) + (1-s_{n-1})s_n p_2 + (1-s_{n-1})(1-s_n)(1-p_2)\right).$$

a) Nach Aufgabenstellung gilt

$$q_1 = \Pr[A_1] = 1$$

und für $n \geq 1$

$$\Pr[A_{n+1} \mid A_n] = p_1, \quad \Pr[A_{n+1}^c \mid A_n] = 1-p_1, \quad \Pr[A_{n+1} \mid A_n^c] = p_2, \quad \Pr[A_{n+1}^c \mid A_n^c] = 1-p_2.$$

b) Für $n \geq 1$

$$\begin{aligned} q_{n+1} = \Pr[A_{n+1}] &= \Pr[A_{n+1} \mid A_n]\Pr[A_n] + \Pr[A_{n+1} \mid A_n^c]\Pr[A_n^c] \\ &= p_1 q_n + p_2(1-q_n) \\ &= (p_1 - p_2)q_n + p_2 \end{aligned}$$

c) Sei $\Delta p = p_1 - p_2$. Es gilt dann

$$\begin{aligned} \Pr[A_1] &= 1 \\ \Pr[A_2] &= \Delta p + p_2 = p_1 \\ \Pr[A_3] &= \Delta p\, p_1 + p_2 \\ \Pr[A_4] &= \Delta p^2 p_1 + p_2 \Delta p\, p_2 + p_2 \\ \Pr[A_5] &= \Delta p^3 p_1 + \Delta p^2 p_2 + \Delta p\, p_2 + p_2. \end{aligned}$$

– Für $p_1 = 1$ folgt sofort mittels Induktion, dass stets $\Pr[A_n] = 1$.

– Wir nehmen daher $p_1 < 1$ im Weiteren an. Wegen $0 \leq p_2 < \frac{1}{2} < p_1 < 1$ folgt dann $0 \leq \Delta p < 1$. Mittels Induktion folgt dann für $4 \leq n \leq N$

$$\Pr[A_n] = \Delta p^{n-2} p_1 + p_2 \sum_{l=0}^{n-3} \Delta p^l.$$

Mit $\sum_{l=0}^{m} g^i = \frac{1-g^{m+1}}{1-g}$ (für $|g| < 1$) lässt sich das also als

$$\begin{aligned} \Pr[A_n] &= \Delta p^{n-2} p_1 + \tfrac{1-\Delta p^{n-2}}{1-\Delta p} p_2 \\ &= \Delta p^{n-2}\left(p_1 - \tfrac{p_2}{1-\Delta p}\right) + \tfrac{p_2}{1-\Delta p} \\ &= \Delta p^{n-2} \tfrac{p_1 - \Delta p\, p_1 - p_2}{1-\Delta p} + \tfrac{p_2}{1-\Delta p} \\ &= \Delta p^{n-1} \tfrac{1-p_1}{1-\Delta p} + \tfrac{p_2}{1-\Delta p} \end{aligned}$$

schreiben. Man überprüft leicht, dass diese Formel sogar für $n = 1, 2$ gilt.

Wegen $\Delta p < 1$ gilt dann $\Delta p^n \xrightarrow{n\to\infty} 0$, also

$$q_n = \Pr[A_n] \xrightarrow{n\to\infty} \frac{p_2}{1-\Delta p}.$$

d) Sei wieder $\Delta p = p_1 - p_2$. Es gilt dann

$$\begin{aligned} \Pr[A_1] &= 0 \\ \Pr[A_2] &= p_2 \\ \Pr[A_3] &= \Delta p\, p_2 + p_2 \\ \Pr[A_4] &= \Delta p^2 p_2 + p_2 \Delta p\, p_2 + p_2 \\ \Pr[A_5] &= \Delta p^3 p_2 + \Delta p^2 p_2 + \Delta p\, p_2 + p_2. \end{aligned}$$

Mittels Induktion folgt (mit $0 \leq p_2 < \frac{1}{2}$ )

$$\Pr[A_n] = p_2 \frac{1-\Delta p^{n-1}}{1-\Delta p}$$

für $n \geq 2$.

Es folgt wieder $\Pr[A_n] \xrightarrow{n\to\infty} \frac{p_2}{1-\Delta p}$.

Anmerkung: Man kann das System auch als Markov-Kette mit den Zuständen "Bestanden" und "Durchgefallen" auffassen. Dann ist $p_1$ die W'keit von "Bestanden" nach "Bestanden" zu wechseln, $1 - p_1$ die Transitionsw'keit "Bestanden" nach "Durchgefallen", usw. .

Dass der erste Student durchfällt/besteht, legt einzig den Startzustand der Markov-Kette fest. Da die Kette jedoch irreduzibel (= stark zusammenhängend) und aperiodisch (hier: jeder Zustand besitzt einen Kreis der Länge 1) ist, ist die stationäre Verteilung (= Grenzw'keit, dass man sich in einem bestimmten Zustand befindet) eindeutig bestimmt - und $\lim_{n\to\infty} \Pr[A_n]$ ist gerade diese Grenzw'keit für den Zustand "Bestanden"; daher spielt die Startverteilung keine Rolle.