

Elementare Zugänge zum Problem der vollständigen Serie

HEINZ HAAKE, MINDEN

Zusammenfassung: Das Problem der durchschnittlichen Wartezeit auf eine vollständige Serie von Ereignissen wird in der Schulmathematik allenfalls für den Fall aufgegriffen, dass die einzelnen Ereignisse alle gleich wahrscheinlich sind. Durch Betrachtung exemplarischer Häufigkeiten als heuristisches Hilfsmittel werden durchschnittliche Wartezeiten für das Vorliegen einer vollständigen Serie auch bei Ereignissen mit unterschiedlichen Wahrscheinlichkeiten bestimmt.

1 Einleitung

Bei zahlreichen Zufallsphänomenen des alltäglichen Lebens fragt der unbefangene Beobachter weniger nach Wahrscheinlichkeiten, sondern eher nach Erwartungswerten: „Wie oft muss ich im Durchschnitt würfeln, bis zum ersten Mal eine „6“ auftaucht? Wie oft muss ich im Durchschnitt Lotto spielen, bis ich zum ersten Mal sechs Richtige habe? Wie viele Cornflakes-Packungen muss ich im Durchschnitt kaufen, bis ich einen vollständigen Satz aller darin enthaltenen Sammelfiguren besitze?“ In der Literatur werden diese Fragestellungen häufig als Wartezeitprobleme bezeichnet, obwohl oftmals nicht die Zeitspanne, sondern die durchschnittliche Anzahl der Versuche bis zum Eintreten eines oder mehrerer Ereignisse gesucht ist.

Im Unterricht der Sekundarstufe I werden solche Wartezeitprobleme kaum behandelt¹, und allenfalls in der Sekundarstufe II aufgegriffen. Folgt man dem üblichen formalen Aufbau, so scheint in der Tat die Behandlung solcher Fragestellungen frühestens in der Sekundarstufe II möglich. Betrachten wir dazu den Zufallsversuch „Warten auf den ersten Erfolg“ (Standardbeispiel: Wie viele Versuche sind durchschnittlich bis zum erstmaligen Würfeln einer „6“ erforderlich?): Ein Ereignis E trete mit Wahrscheinlichkeit p ein. Die Zufallsgröße X beschreibe die Anzahl der Versuche bis zum erstmaligen Auftreten des Ereignisses E . Bekanntlich ist X dann geometrisch verteilt mit

$$P(X = k) = (1 - p)^{k-1} \cdot p,$$

¹Ein stark den Modellierungsaspekt betonender Unterrichtsvorschlag wurde im Rahmen des Sinus-Transferprojekts entwickelt: <http://db.learnline.de/angebote/materialdatenbank/nutzersicht/materialeintrag.jsp?matId=282>.

²Bei den folgenden Überlegungen werden die mittleren Wartezeiten ohne Rückgriff auf die zugehörige Wahrscheinlichkeitsverteilung bestimmt, vgl. auch Büchler und Henn (2005), S. 261.

und für die durchschnittliche Anzahl der Versuche bis zum Eintreten des Erfolgs gilt

$$E(X) = \sum_{k=1}^{\infty} k \cdot P(X = k) = \frac{1}{p}.$$

Zur Lösung des Problems werden demnach die Begriffe Zufallsgröße und Erwartungswert benötigt. Zusätzlich bereitet die Summation der geometrischen Reihe nicht unerhebliche Schwierigkeiten, es sei denn, wir überlassen diese Rechnung einem Computer-Algebra-System.

Beim Übergang vom „Warten auf den ersten Erfolg“ zum „Warten auf eine vollständige Serie“ (Standardbeispiel: Wie oft muss durchschnittlich gewürfelt werden, bis jede Augenzahl mindestens einmal aufgetreten ist?), ist die Bestimmung der zugehörigen Wahrscheinlichkeitsverteilung allenfalls in einem Leistungskurs zu bewältigen². Ein gangbarer Weg wird in Althoff (2000) beschrieben. Völlig außerhalb der schulischen Möglichkeiten scheint das Problem der vollständigen Serie bei Ereignissen mit unterschiedlichen Wahrscheinlichkeiten zu liegen (Standardbeispiel: Wie oft muss durchschnittlich ein gezinkter Würfel geworfen werden, bis jede Augenzahl mindestens einmal aufgetreten ist?) Ein Zugang zu diesem Problem findet sich bei Henze (1998).

2 Ein heuristischer Zugang mittels absoluter Häufigkeiten

In Wassner, Krauss und Martignon (2002) zeigen die Autoren, dass Aussagen über bedingte Wahrscheinlichkeiten dann am besten verstanden werden, wenn die zugehörigen Daten in Form von absoluten Häufigkeiten vorliegen oder in solche umgewandelt werden (vgl. auch Henze, 2004, S. 107 f.). In der folgenden Aufgabensequenz wird zunächst bei fast allen Lösungen auf der Basis absoluter Häufigkeiten argumentiert und das Ergebnis als relative Häufigkeit notiert. In der anschließenden Formalisierung werden die relativen Häufigkeiten in Wahrscheinlichkeiten übersetzt. Grundlage aller Argumentationen ist weniger der stochastische Formalismus, sondern

vielmehr das empirische Gesetz über die Stabilisierung relativer Häufigkeiten.

Bei der Formalisierung sollen folgende Schreibweisen verwendet werden: Es seien $E_i, i = 1, \dots, n$, paarweise disjunkte Ereignisse eines Ereignisraums Ω mit $P(E_i) = p_i$. Es bezeichne $t(E_i)$ die durchschnittliche Wartezeit (i.e. die durchschnittliche Anzahl der Versuche) bis zum erstmaligen Eintreten des Ereignisses E_i . Des Weiteren bezeichne z. B. $t(E_1, E_2, E_3)$ die Wartezeit, bis alle drei Ereignisse E_1, E_2 und E_3 eingetreten sind.

2.1 Wartezeit auf ein Ereignis

Aufgabe 1

Die Firma Allogs legt fünf Sechstel der von ihr produzierten Cornflakes-Packungen eine Sammelfigur bei. Wie viele Packungen muss man im Durchschnitt kaufen, um eine Sammelfigur zu erhalten?

Lösung 1a)

Bei 600 Käufen werden sich im Durchschnitt 500 Erfolge einstellen. Also sind pro Erfolg durchschnittlich $600/500 = 1,2$ Käufe erforderlich.

Lösung 1b)

Wenn 600 Käufer 600 Packungen kaufen, dann werden nur durchschnittlich 500 von ihnen eine Sammelfigur erwerben. Beim Kauf von 720 Packungen bekommt im Durchschnitt jeder eine Figur. Also muss jeder der 600 Käufer durchschnittlich 1,2 Packungen kaufen.

Lösung 1c)

Es sei t die durchschnittliche Anzahl der zu kaufenden Packungen. Dann benötigen 600 Käufer im Durchschnitt $600 \cdot t$ Käufe. Bei den ersten 600 Käufen werden im Durchschnitt 500 Käufer Erfolg haben, die übrigen 100 müssen ihr Glück erneut versuchen und benötigen dafür $100 \cdot t$ Versuche. Insgesamt ergibt sich daraus die Gleichung

$$600 \cdot t = 600 + 100 \cdot t,$$

aus welcher sich wieder $t = 1,2$ ergibt.

Formalisierung 1

³Diese Situation lässt sich sehr einfach mit Hilfe eines Laplace-Würfels simulieren.

Wenn das Ereignis E mit Wahrscheinlichkeit $P(E)$ eintritt, dann gilt für die durchschnittliche Wartezeit

$$t(E) = \frac{1}{P(E)}.$$

2.2 Wartezeit auf eine vollständige Serie bei gleichwahrscheinlichen Ereignissen

Aufgabe 2

Jede Cornflakes-Packung der Firma Bellogs enthält genau eine von sechs Sammelfiguren, die alle mit der gleichen Wahrscheinlichkeit vorkommen. Wie viele Packungen muss ein Käufer im Durchschnitt erwerben, bis er einen vollständigen Sammelfigurensatz besitzt? ³

Lösung 2

Der erste Packungskauf führt mit Sicherheit zu einer neuen Figur. Mit der Wahrscheinlichkeit $\frac{5}{6}$ führt der nächste Kauf zu einer weiteren Figur, die mit der ersten Figur nicht übereinstimmt. Nach Formalisierung 1 beträgt die durchschnittliche Wartezeit auf die weitere Figur $\frac{6}{5}$. Die weiteren Wartezeiten betragen dann sukzessive $\frac{6}{4}, \frac{6}{3}, \frac{6}{2}$ und $\frac{6}{1}$.

Die durchschnittliche Wartezeit t auf einen vollständigen Figurensatz beträgt demnach

$$t = \frac{6}{6} + \frac{6}{5} + \frac{6}{4} + \frac{6}{3} + \frac{6}{2} + \frac{6}{1} = 14,7.$$

Formalisierung 2

Sind E_1, \dots, E_n Ereignisse mit $P(E_i) = \frac{1}{n}$ für alle $i, i = 1, \dots, n$, dann gilt

$$t(E_1, \dots, E_n) = n \sum_{i=1}^n \frac{1}{i}.$$

2.3 Wartezeit auf eine vollständige Serie bei zwei Ereignissen mit unterschiedlichen Wahrscheinlichkeiten

Aufgabe 3

Von den Cornflakes-Packungen der Firma Cellogs enthalten zwei Drittel die Figur 1 und ein Drittel die Figur 2. Wie viele Packungen muss ein Sammler im Durchschnitt kaufen, bis er beide Figuren besitzt?

Lösung 3

Wieder gehen wir von 600 Käufern aus. Im Durchschnitt werden 400 Käufer die Figur 1 beim ersten

Kauf erhalten. Anschließend braucht von diesen 400 Käufern jeder durchschnittlich drei Versuche, um Figur 2 zu erwerben. Die 200 Käufer, die beim ersten Kauf die Figur 2 erhalten haben, benötigen in der Folge durchschnittlich je 1,5 Versuche, um Figur 1 zu sammeln. Insgesamt sind also durchschnittlich $600 + 400 \cdot 3 + 200 \cdot 1,5 = 2100$ Käufe erforderlich, was einen Durchschnittswert von 3,5 Käufen für jeden Sammler ergibt. Schreiben wir die obige Überlegung in der Form

$$\frac{600}{600} + \frac{400}{600} \cdot \frac{3}{1} + \frac{200}{600} \cdot \frac{3}{2} = 3,5,$$

so ergibt sich daraus

Formalisierung 3

$$\begin{aligned} t(E_1, E_2) &= 1 + p_1 \cdot \frac{1}{p_2} + p_2 \cdot \frac{1}{p_1} \\ &= 1 + p_1 \cdot t(E_2) + p_2 \cdot t(E_1) \end{aligned}$$

2.4 Wartezeit auf zwei Ereignisse mit unterschiedlichen Wahrscheinlichkeiten

Aufgabe 4

Von den Cornflakes-Packungen der Firma Dellogs enthält die Hälfte der Packungen Figur 1 und ein Drittel der Packungen Figur 2. Im verbleibenden Sechstel der Packungen befindet sich keine Figur.

Lösung 4

Ausgehend von 600 Käufern sind zunächst durchschnittlich 720 Versuche erforderlich, um eine der beiden Figuren zu erhalten. In 360 von diesen 720 Käufen wird Figur 1 in der Packung vorkommen. Für jeden dieser Fälle sind dann noch durchschnittlich 3 Käufe erforderlich, um die zweite Figur zu erhalten. In 240 Fällen wird zunächst Figur 2 erworben. Zum Sammeln von Figur 1 sind dann im Durchschnitt noch 2 Käufe notwendig. Insgesamt sind also $720 + 360 \cdot 3 + 240 \cdot 2 = 2280$, also 3,8 Käufe im Durchschnitt erforderlich. Die Darstellung

$$\frac{720}{600} + \frac{1}{2} \cdot \frac{720}{600} \cdot 3 + \frac{1}{3} \cdot \frac{720}{600} \cdot 2$$

führt zur

Formalisierung 4

$$\begin{aligned} t(E_1, E_2) &= \frac{1}{p_1 + p_2} + \frac{p_1}{p_1 + p_2} \cdot \frac{1}{p_2} + \frac{p_2}{p_1 + p_2} \cdot \frac{1}{p_1} \\ &= \frac{1}{p_1 + p_2} + \frac{p_1}{p_1 + p_2} t(E_2) + \frac{p_2}{p_1 + p_2} t(E_1) \\ &= \frac{1 + p_1 \cdot t(E_2) + p_2 \cdot t(E_1)}{p_1 + p_2} \end{aligned}$$

2.5 Wartezeit auf drei Ereignisse mit unterschiedlichen Wahrscheinlichkeiten

Aufgabe 5

Von den Cornflakes-Packungen der Firma Ellogs enthält ein Viertel der Packungen Figur 1, ein Sechstel der Packungen Figur 2 und ein Zwölftel der Packungen Figur 3. In der Hälfte der Packungen befindet sich keine Figur.

Lösung 5

Auch hier betrachten wir 600 Käufer. Da nur die Hälfte der Packungen eine Figur enthält, sind durchschnittlich 1200 Käufe notwendig, um eine der drei Figuren zu erhalten. Bei durchschnittlich 300 dieser 1200 Käufe erhalten die Käufer Figur 1. Anschließend sind noch durchschnittlich $300 \cdot t(E_2, E_3)$ Käufe zum Erwerb der Figuren 2 und 3 erforderlich. Nach der obigen Formalisierung 4 gilt $t(E_2, E_3) = 14$. In ca. 200 von den 1200 Käufen wird Figur 2 als erste Figur erworben. Dann sind noch $200 \cdot t(E_1, E_3) = 200 \cdot 13$ Käufe notwendig, um die Figuren 1 und 3 zu erhalten. In ca. 100 von den 1200 Käufen wird zunächst Figur 3 gesammelt. Anschließend sind noch $100 \cdot t(E_1, E_2) = 100 \cdot \frac{38}{5}$ Käufe zum Kauf der Figuren 1 und 2 erforderlich. Insgesamt ergibt sich die Summe $1200 + 300 \cdot 14 + 200 \cdot 13 + 100 \cdot \frac{38}{5} = 8760$, was durchschnittlich 14,6 Versuchen pro Käufer entspricht. Aus der Darstellung

$$\frac{1200}{600} + \frac{1}{4} \cdot \frac{1200}{600} \cdot 14 + \frac{1}{6} \cdot \frac{1200}{600} \cdot 13 + \frac{1}{12} \cdot \frac{1200}{600} \cdot \frac{38}{5}$$

folgt die

Formalisierung 5

$$\begin{aligned} t(E_1, E_2, E_3) &= \\ &= \frac{1 + p_1 \cdot t(E_2, E_3) + p_2 \cdot t(E_1, E_3) + p_3 \cdot t(E_1, E_2)}{p_1 + p_2 + p_3} \end{aligned}$$

2.6 Wartezeit auf das Eintreffen von n Ereignissen

Aus den obigen Formalisierungen 4 und 5 lässt sich unschwer folgende Rekursionsformel erraten:

Formalisierung 6

$$t(E_1, \dots, E_n) = \frac{1 + \sum_{i=1}^n p_i \cdot t(E_1, \dots, \cancel{E_i}, \dots, E_n)}{\sum_{i=1}^n p_i}$$

Diese Formel findet sich auch in Boneh und Hofri (1997), S.43 (7).

3 Ergänzungen und Vertiefungen

3.1 Eine vereinfachte Darstellung

Bei den obigen Überlegungen trat bei der Wartezeit auf zwei Ereignisse die Wartezeit $t(E_1 \cup E_2)$ als „Zwischenstation“ auf. Für zwei disjunkte Ereignisse E_1 und E_2 mit den Wahrscheinlichkeiten p_1 und p_2 gilt $t(E_1 \cup E_2) = \frac{1}{p_1 + p_2}$. Bilden wir die Differenz aus den Wartezeiten $t(E_1 \cup E_2)$ und $t(E_1)$, so erhalten wir:

$$t(E_1) - t(E_1 \cup E_2) = \frac{1}{p_1} - \frac{1}{p_1 + p_2} = \frac{p_2}{p_1 + p_2} \cdot \frac{1}{p_1}$$

Analog gilt

$$t(E_2) - t(E_1 \cup E_2) = \frac{1}{p_2} - \frac{1}{p_1 + p_2} = \frac{p_1}{p_1 + p_2} \cdot \frac{1}{p_2}$$

Diese Terme tauchten bereits bei der Formalisierung 4 auf. Sie lassen sich auch als mittlere Wartezeit auf das Ereignis E_1 bzw. E_2 interpretieren unter der Bedingung, dass das Ereignis $E_1 \cup E_2$ eingetreten ist. Deuten wir außerdem die Terme $\frac{p_1}{p_1 + p_2}$

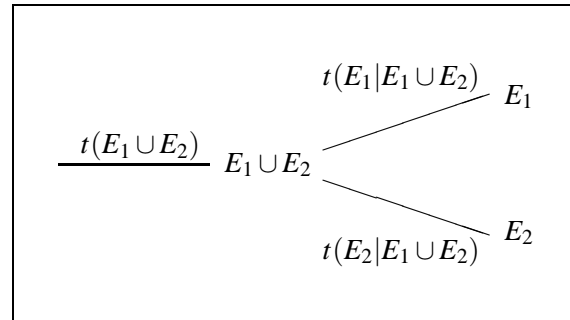
bzw. $\frac{p_2}{p_1 + p_2}$ als die bedingten Wahrscheinlichkeiten $P(E_1|E_1 \cup E_2)$ bzw. $P(E_2|E_1 \cup E_2)$, so ergibt sich folgendes Ergebnis: die „bedingte Wartezeit“ $t(E_i|E_1 \cup E_2)$ ist das Produkt aus der bedingten Wahrscheinlichkeit $P(E_i|E_1 \cup E_2)$ und der Wartezeit auf das Ereignis E_i , $i \in \{1, 2\}$. Die obigen bedingten Wartezeiten sind Beispiele für bedingte Erwartungswerte. Eine systematische Einführung in die Theorie der bedingten Erwartungswerte findet sich in den Arbeiten von Kilian (1987) und Humenberger (2000).

Ersetzen wir in Formalisierung 4 die bedingten Wartezeiten $\frac{p_2}{p_1 + p_2} \cdot \frac{1}{p_1}$ und $\frac{p_1}{p_1 + p_2} \cdot \frac{1}{p_2}$ durch die Ter-

me $\frac{1}{p_1} - \frac{1}{p_1 + p_2}$ und $\frac{1}{p_2} - \frac{1}{p_1 + p_2}$, so erhalten wir:

$$\begin{aligned} t(E_1, E_2) &= \frac{1}{p_1} + \frac{1}{p_2} - \frac{1}{p_1 + p_2} \\ &= t(E_1) + t(E_2) - t(E_1 \cup E_2) \end{aligned}$$

Dieses auf algebraischem Weg gewonnene Resultat lässt sich mit Hilfe des folgenden Diagramms stochastisch interpretieren:



Die Pfade zu den Ereignissen E_1 und E_2 besitzen einen gemeinsamen zum Ereignis $E_1 \cup E_2$ führenden Teilpfad. Die Wartezeit zu diesem Teilpfad wird bei der Addition der Wartezeiten $t(E_1)$ und $t(E_2)$ zweifach gezählt. Die Subtraktion der Wartezeit $t(E_1 \cup E_2)$ sorgt für die gleiche Gewichtung aller Teilpfade. Diese Situation lässt sich mit Hilfe des Ein- und Ausschlussprinzips verallgemeinern. Nach Boneh und Hofri (1997), S. 42,(1) gilt:

$$\begin{aligned} t(E_1, \dots, E_n) &= \sum_{1 \leq i \leq n} \frac{1}{p_i} - \sum_{1 \leq i < j \leq n} \frac{1}{p_i + p_j} \\ &+ \sum_{1 \leq i < j < k \leq n} \frac{1}{p_i + p_j + p_k} - \dots \end{aligned}$$

3.2 Verallgemeinerung auf nicht disjunkte Ereignisse

In diesem Abschnitt betrachten wir das Sammelbilderproblem für zwei nicht notwendig disjunkte Ereignisse.

Aufgabe 6

Die Firma Fellogs legt der Hälfte ihrer Packungen die Sammelfigur 1 bei, ein Drittel der Packungen enthält die Sammelfigur 2 und ein Sechstel sogar beide Sammelfiguren. Wie viele Packungen sind im Durchschnitt bis zum Erwerb beider Sammelfiguren erforderlich?

Lösung 6

Es gilt $P(E_1) = \frac{2}{3}$, $P(E_2) = \frac{1}{2}$ und $P(E_1 \cap E_2) = \frac{1}{6}$. Auch diese Aufgabe soll zunächst durch die exemplarische Betrachtung von 600 Käufern gelöst werden: Die ersten 600 Käufe führen mit Sicherheit zu mindestens einer Sammelfigur. Durchschnittlich 100 Käufer haben sogar mit einem Kauf beide Figuren erworben. Die 300 Käufer, die beim ersten Kauf lediglich Figur 1 erworben haben, benötigen im Durchschnitt zwei Käufe zum Erwerb von Figur 2. Die 200 Käufer, die beim ersten Kauf Figur 2 erhalten haben, brauchen anschließend noch durchschnittlich 1,5 Versuche, um Figur 1 zu erwerben. Insgesamt sind im Durchschnitt $600 + 300 \cdot 2 + 200 \cdot 1,5 = 1500$ Käufe notwendig, was im Mittel 2,5 Versuchen pro Käufer entspricht.

Zur Formalisierung dieses Problems greifen wir auf Formalisierung 4 aus Abschnitt 2.4 zurück. Wegen der Nicht-Disjunktheit der Ereignisse sind der Nenner $p_1 + p_2$ durch den Ausdruck $P(E_1) + P(E_2) - P(E_1 \cap E_2)$ und die Zählerfaktoren p_1 und p_2 durch jeweils $P(E_1) - P(E_1 \cap E_2)$ und $P(E_2) - P(E_1 \cap E_2)$ zu ersetzen, denn bei Eintreten des Ereignisses $E_1 \cap E_2$ sind keine weiteren Versuche erforderlich. Damit erhalten wir insgesamt

$$t(E_1, E_2) = \frac{1}{P(E_1) + P(E_2) - P(E_1 \cap E_2)} + \frac{(P(E_1) - P(E_1 \cap E_2)) \cdot t(E_2)}{P(E_1) + P(E_2) - P(E_1 \cap E_2)} + \frac{(P(E_2) - P(E_1 \cap E_2)) \cdot t(E_1)}{P(E_1) + P(E_2) - P(E_1 \cap E_2)}$$

Diese recht unübersichtliche Formel lässt sich vereinfachen, nachdem sich durch „Addition der Null“ Kürzungsmöglichkeiten ergeben:

$$\begin{aligned} & \frac{1}{P(E_1) + P(E_2) - P(E_1 \cap E_2)} \\ & + \frac{(P(E_1) - P(E_1 \cap E_2) + (P(E_2) - P(E_1 \cap E_2))) \cdot \frac{1}{P(E_2)} - \frac{P(E_2)}{P(E_2)}}{P(E_1) + P(E_2) - P(E_1 \cap E_2)} \\ & + \frac{(P(E_2) - P(E_1 \cap E_2) + P(E_1)) \cdot \frac{1}{P(E_1)} - \frac{P(E_1)}{P(E_1)}}{P(E_1) + P(E_2) - P(E_1 \cap E_2)} \\ & = t(E_1) + t(E_2) - t(E_1 \cup E_2) \end{aligned}$$

Die in Abschnitt 3.1 entwickelte Beziehung

⁴Intuitiv ist dieses Resultat auch aus der Betrachtung des obigen Diagramms zu erwarten.

$$t(E_1, E_2) = t(E_1) + t(E_2) - t(E_1 \cup E_2)$$

gilt demnach auch für nicht disjunkte Ereignisse⁴.

4 Anmerkungen zur Berechnung der Wartezeiten

4.1 Ereignisse mit gleichen Wahrscheinlichkeiten

Falls alle Ereignisse gleich wahrscheinlich sind, so lässt sich die Wartezeit gemäß Formalisierung 2 berechnen. Steht dabei zur Bestimmung der Summe kein Rechner zur Verfügung, dann lässt sich die harmonische Reihe $\sum_{i=1}^n \frac{1}{i}$ für große n durch $\log(n) + \gamma$ approximieren, mit der Eulerschen Konstante $\gamma = 0,577\dots$. Für große n gilt demnach

$$t(E_1, \dots, E_n) \approx n \cdot (\log(n) + \gamma).$$

4.2 Ereignisse mit unterschiedlichen Wahrscheinlichkeiten

Weisen die Ereignisse E_1, \dots, E_n unterschiedliche Wahrscheinlichkeiten auf, so lassen sich für eine kleine Anzahl n von Ereignissen, etwa $n \leq 3$, die Wartezeiten auf eine vollständige Serie recht einfach ohne Hilfe eines Rechners gemäß Formalisierung 5 bestimmen.

Für größere Anzahlen n lassen sich die Wartezeiten mit Hilfe eines Computers unter Verwendung der Rekursionsformel aus Formalisierung 6 berechnen. Das folgende Programm zeigt eine Implementation unter Maple:

```
t:= proc(p::list)
if nops(p)=1 then 1/p[1]
else
(1+add(p[i]*t(subsop(i=NULL,p)),i=1..nops(p)))
/add(p[i],i=1..nops(p))
fi
end;
```

Die Prozedur `t` hat als Parameter die Liste `p` der Wahrscheinlichkeiten der Ereignisse E_1 bis E_n . Zum Beispiel wird Aufgabe 5 wie folgt gelöst:

```
> t([1/4,1/6,1/12]);
```

$$\frac{73}{5}$$

Diese Programme haben nur für $n \leq 9$ kurze Laufzeiten, da durch die Rekursion ein $n!$ -facher Prozeduraufruf erfolgt.

Falls nur eines der n Ereignisse von der Gleichverteilung abweicht, lässt sich die Zahl der rekursi-

ven Prozeduraufrufe auf $n - 1$ verringern: Ohne Beschränkung der Allgemeinheit gelte $P(E_i) = p_i =: p$ für $i = 1, \dots, n - 1$ und $P(E_n) = q$. Dann lässt sich die Rekursionsformel aus Formalisierung 6 wegen

$$\sum_{i=1}^{n-1} p_i \cdot t(E_1, \dots, E_i, \dots, E_n) = (n-1) \cdot p \cdot t(E_2, \dots, E_n)$$

und

$$t(E_1, \dots, E_{n-1}) = \frac{1}{p} \cdot \sum_{i=1}^{n-1} \frac{1}{i}$$

in folgender Form schreiben:

$$t(E_1, \dots, E_n) =$$

$$\frac{1 + (n-1) \cdot p \cdot t(E_2, \dots, E_n) + \frac{q}{p} \cdot \sum_{i=1}^{n-1} \frac{1}{i}}{(n-1) \cdot p + q}$$

Es folgt die Umsetzung für Maple:

```
t:= proc(n::integer,p,q)
if n=2 then (1+p/q+q/p)/(p+q)
else (1+(n-1)*p*t(n-1,p,q)
+q*sum(1/i,i=1..n-1)/p)/((n-1)*p+q)
fi
end;
```

Beispiel: Für $n = 300$, $q = \frac{1}{600}$ und $p = \frac{1 - \frac{1}{600}}{299}$ ergibt sich als mittlere Wartezeit:

```
> evalf(t(300,599/600/299,1/600));
```

1911.512497

Zum Vergleich: Bei Gleichverteilung ($p = q = \frac{1}{300}$) ergibt sich

```
> evalf(t(300,1/300,1/300));
```

1884.799164

Literatur

- Althoff, H. (2000): Zur Berechnung der Wahrscheinlichkeiten für das Vorliegen einer vollständigen Serie (Sammelbilderproblem) In: Stochastik in der Schule 20, 1 S. 18-20
- Boneh, A. und Hofri, M. (1997): The Coupon-Collector Problem Revisited – A Survey of Engineering Problems and Computational Methods In: Communications in Statistics – Stoch. Models 13, S. 39-66.
- Büchter, A. und Henn H.-W. (2005): Elementare Stochastik, Springer-Verlag
- Henze, N. (1998): Wartezeitprobleme, unveröffentlichtes Manuskript zu einer Lehrerfortbildung an der Universität Bielefeld am 18.2.1998
- Henze, N. (2004): Stochastik für Einsteiger, 5. Auflage, Vieweg Verlag
- Humenberger, H. (2000): Bedingte Erwartungswerte In: Stochastik in der Schule 20, 1 S. 20-27
- Kilian, H. (1987): Bedingte Erwartungswerte im Stochastikunterricht In: Stochastik in der Schule 7, 3 S. 24-45
- Wassner, C., Krauss, S. und Martignon, L. (2002): Muss der Satz von Bayes schwer verständlich sein? In: Praxis der Mathematik Heft 1/44 S. 12-16

Danksagung: Ich danke dem mir unbekanntem Gutachter für seine wertvollen Hinweise.

Anschrift des Verfassers

Heinz Haake

Dützer Dorfstr. 22

32429 Minden

Heinz.Haake@t-online.de