

Wann zeigt auch der letzte Würfel eine Sechs?

NORBERT HENZE, KARLSRUHE, UND REIMUND VEHLING, HANNOVER

Zusammenfassung: *Drei echte, nicht unterscheidbare Würfel werden gleichzeitig geworfen. Diejenigen (eventuell vorhandenen) Würfel, die eine Sechs zeigen, werden beiseite gelegt und die unter Umständen noch verbleibenden Würfel erneut geworfen. Wiederum legt man etwaige Würfel, die eine Sechs zeigen, beiseite und wirft die übrigen Würfel erneut. Dieser stochastische Vorgang endet, wenn auch der letzte Würfel eine Sechs gezeigt hat. Die Zufallsvariable X beschreibe die Anzahl der dazu nötigen Würfe. Welche Verteilung besitzt X ? Wir lösen dieses Problem auf verschiedene Weisen. Als schlagkräftiges Hilfsmittel erweist sich ein Unabhängigkeitsargument. Mit diesem kann man auch die allgemeine Situation, dass n gleichartige, unter Umständen gefälschte Würfel gegeben sind, auf elegante Weise behandeln.*

1 Einleitung

Stochastik in der Schule lebt von einprägsamen Problemen aus der Erfahrungswelt von Schülerinnen und Schülern. Für das in der Zusammenfassung beschriebene Experiment sind verschiedene Fragestellungen denkbar, wie z. B.

- kann man getrost darauf wetten, dass 10 Würfe ausreichen, also $X \leq 10$ gilt?
- wie wahrscheinlich ist es, mehr als 15 Mal werfen zu müssen, also das Ereignis $\{X > 15\}$ eintritt?
- anhand des Experimentes soll ein Glücksspiel geplant und am Tag der offenen Tür durchgeführt werden, das am Ende dieses Tages einen Gewinn für die Klassenkasse erwarten lässt. Wie könnte ein solches Glücksspiel aussehen?

Offenbar benötigt man für jede dieser Fragestellungen die Wahrscheinlichkeiten $\mathbb{P}(X = k)$ zumindest für gewisse Werte von k . Auch der Erwartungswert von X könnte interessant sein.

So wie das Experiment in der Zusammenfassung beschrieben ist, würde es auch von Schülerinnen und Schülern durchgeführt werden. Im Hinblick auf eine elegante Modellierung, die unmittelbar auch zu einer Lösung eines allgemeineren Problems führt, wird jedoch eine falsche Fährte gelegt. Diese Fährte sugge-

riert, ausgetretene Pfade zu benutzen und die in der Schule fast ausschließlich zur Modellierung stochastischer Probleme eingesetzte Pfadregel zu verwenden.

In diesem Aufsatz zeigen wir, dass eine Tiefbohrung bei diesem recht überschaubaren Problem ganz unterschiedliche Lösungswege mit vielfältigen Vernetzungen zu anderen Gebieten mit sich bringt. Eine Verallgemeinerung führt schließlich zu Fragestellungen, die jenseits der üblichen Schulmathematik angesiedelt sind, aber für eine Lehrkraft von Interesse sein sollten. Das damit einhergehende Hintergrundwissen ist ausdrücklich erwünscht!

2 Experiment und Simulation

Da die Problemstellung ein Ausprobieren mit echten Würfeln geradezu herausfordert, sollten Schülerinnen und Schüler das *Warten auf die letzte Sechs* wiederholt spielen, um ein „stochastisches Gefühl“ für die Anzahl der dazu nötigen Würfe zu erhalten. Die Herausbildung eines solchen Gefühls kann nicht oft genug im Zentrum eines Stochastikunterrichts stehen!

Die Durchführung des Würfelspiels kann in kleinen Gruppen erfolgen. Die Zusammenfassung aller Ergebnisse zu einem gemeinsamen Kursexperiment ergibt dann eine erste Schätzung für die gesuchten Wahrscheinlichkeiten $\mathbb{P}(X = k)$. Schülerinnen und Schüler erhalten durch dieses Vorgehen auch die Möglichkeit, sich intensiver mit der Problemstellung auseinanderzusetzen, und erste Ideen für eine rechnerische Lösung entstehen unter Umständen sogar ganz beiläufig.

Statt realer Würfel kann für eine Simulation auch der in jedem Rechner vorhandene Zufallsgenerator benutzt werden. Zudem ist der Einsatz von GeoGebra möglich. Ein solches Vorgehen führt zu einer sinnvollen Verwendung von Tabellenkalkulation. Abbildung 1 zeigt das Ergebnis einer auf 500 Wiederholungen gründenden Simulation mit GeoGebra. Die zugehörige Datei ist vom zweiten Autor erhältlich.

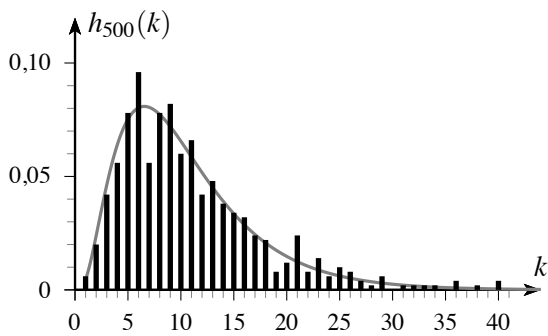


Abb. 1: Relative Häufigkeiten für das Eintreten des Ereignisses $\{X = k\}$ aus 500 Wiederholungen

Die dargestellten relativen Häufigkeiten können als Schätzungen der auf der eingezeichneten Kurve liegenden und in Abschnitt 5 ermittelten theoretischen Wahrscheinlichkeiten $\mathbb{P}(X = k)$ interpretiert werden. Natürlich kann eine solche Simulation auch *nach* etwaigen Berechnungen von Wahrscheinlichkeiten eingesetzt werden. Es ist schön zu sehen, wie sich mit zunehmender Anzahl an Wiederholungen das Bild der theoretischen Verteilung abzeichnet. Durch die unvermeidliche Fluktuation relativer Häufigkeiten verbleibt aber stets eine Diskrepanz zwischen relativer Häufigkeit und Wahrscheinlichkeit. Eine Konvergenz im Sinne der Analysis findet ja hier nicht statt!

Im Stochastikunterricht ist es nicht schwierig, die Modellbildung und den Modellierungskreislauf ins Zentrum zu stellen. Hierzu gehört ganz wesentlich auch das Stellen von Fragen. In der vorliegenden Situation könnten derartige Fragen sein: Sind die einzelnen Würfel unabhängig voneinander? Gibt es Unterschiede, wenn die Würfel einzeln oder zusammen geworfen werden? Was ändert sich, wenn mehr als drei (oder nur zwei) Würfel zur Verfügung stehen?

3 Erste Wahrscheinlichkeiten

Da das Ereignis $\{X = 1\}$ genau dann eintritt, wenn der erste Wurf drei Sechsen ergibt, gilt

$$\mathbb{P}(X = 1) = \left(\frac{1}{6}\right)^3 = 0,0046296\dots$$

Dieses Resultat werden Schülerinnen und Schüler sicherlich schnell erhalten. Vielleicht ergeben sich an dieser Stelle aber auch schon Diskussionen, denn die drei Würfel waren ja als *nicht unterscheidbar* angenommen worden. Eine solche Diskussion könnte aber spätestens beim Problem, $\mathbb{P}(X = 2)$ zu bestimmen, auftreten. Da (erst) nach dem zweiten Wurf des

Bechers jeder der Würfel eine Sechs gezeigt haben soll, gibt es zur Beschreibung der dafür günstigen Abläufe die Fälle

$$(0, 3), (1, 2), (2, 1).$$

Dabei bedeutet allgemein ein Paar (i, j) , dass der erste Wurf i und der zweite j Sechsen ergeben hat. Soweit, so gut! Aber welche Wahrscheinlichkeiten besitzen diese drei Möglichkeiten (die man sich natürlich auch als Pfade in einem Baumdiagramm veranschaulichen kann)? Das Paar $(0, 3)$ besagt, dass der erste Wurf keine und der zweite drei Sechsen produziert. Nach der ersten Pfadregel ist die Wahrscheinlichkeit hierfür gleich

$$\left(\frac{5}{6}\right)^3 \cdot \left(\frac{1}{6}\right)^3 = \frac{125}{6^6}. \quad (1)$$

Damit sich der Ablauf $(1, 2)$ ergibt, muss beim ersten Wurf genau einer der drei Würfel eine Sechs zeigen, und beim zweiten, nur noch mit zwei Würfeln erfolgreichen Wurf müssen zwei Sechsen oben liegen. Mit Hilfe der Binomialverteilung und der ersten Pfadregel ergibt sich somit für den Ablauf $(1, 2)$ die Wahrscheinlichkeit

$$\binom{3}{1} \cdot \frac{1}{6} \cdot \left(\frac{5}{6}\right)^2 \cdot \left(\frac{1}{6}\right)^2 = \frac{75}{6^5}. \quad (2)$$

Da Schülerinnen und Schüler die Binomialverteilung als Verteilung der Trefferanzahl aus unabhängigen Versuchen kennengelernt haben, müssen sie bei obiger Herleitung über die Einsicht verfügen, dass sich der Zufall nicht beeinflussen lässt, wenn man die Würfel einzeln nacheinander wirft, was dann drei unabhängigen Versuchen entspricht. Mit dieser Vorstellung ist die Wahrscheinlichkeit für den Ablauf $(2, 1)$ durch

$$\binom{3}{2} \cdot \left(\frac{1}{6}\right)^2 \cdot \frac{5}{6} \cdot \frac{1}{6} = \frac{15}{6^4} \quad (3)$$

gegeben. Addition der in (1), (2) und (3) stehenden Werte liefert jetzt das Resultat

$$\mathbb{P}(X = 2) = \frac{1115}{6^6} = 0,023898\dots$$

Interessanter wird es, wenn auch noch die Wahrscheinlichkeit $\mathbb{P}(X = 3)$ bestimmt werden soll. Hierfür gibt es die sechs Abläufe

$$(0, 0, 3), (0, 1, 2), (0, 2, 1), (1, 0, 2), (1, 1, 1), (2, 0, 1).$$

Dabei bedeutet allgemein das Tripel (i, j, k) , dass der erste Wurf des Knobelbechers i , der zweite j und der dritte k Sechsen ergeben hat. Man beachte aber, dass etwa beim Ablauf $(1, 1, 1)$ der erste Wurf des Bechers mit drei und der zweite mit zwei Würfeln erfolgt, und

dass beim letzten Wurf nur noch ein Würfel vorhanden ist. Die Wahrscheinlichkeiten für diese Abläufe sind (in der aufgeführten Reihenfolge)

$$\begin{aligned} & \left(\frac{5}{6}\right)^3 \cdot \left(\frac{5}{6}\right)^3 \cdot \left(\frac{1}{6}\right)^3, \\ & \left(\frac{5}{6}\right)^3 \cdot \binom{3}{1} \cdot \frac{1}{6} \cdot \left(\frac{5}{6}\right)^2 \cdot \left(\frac{1}{6}\right)^2, \\ & \left(\frac{5}{6}\right)^3 \cdot \binom{3}{2} \cdot \left(\frac{1}{6}\right)^2 \cdot \frac{5}{6} \cdot \frac{1}{6}, \\ & \binom{3}{1} \cdot \frac{1}{6} \cdot \left(\frac{5}{6}\right)^2 \cdot \left(\frac{5}{6}\right)^2 \cdot \left(\frac{1}{6}\right)^2, \\ & \binom{3}{1} \cdot \frac{1}{6} \cdot \left(\frac{5}{6}\right)^2 \cdot \binom{2}{1} \cdot \frac{1}{6} \cdot \frac{5}{6} \cdot \frac{1}{6}, \\ & \binom{3}{2} \cdot \left(\frac{1}{6}\right)^2 \cdot \frac{5}{6} \cdot \frac{5}{6} \cdot \frac{1}{6}. \end{aligned}$$

Addition dieser Werte liefert

$$\mathbb{P}(X = 3) = \frac{466075}{6^9} = 0,046248\dots$$

Man kann jetzt versuchen, induktiv vorzugehen, um $\mathbb{P}(X = 4)$ zu erhalten. Die möglichen Abläufe sind jetzt Quadrupel (4-Tupel) (i, j, k, ℓ) mit der analogen Bedeutung wie oben. Zerlegt man diese Abläufe nach den Werten der ersten Komponente i , also der Anzahl der Sechsen beim ersten Wurf des Bechers, so gibt es insgesamt 10 solcher Abläufe, die sich wie folgt aufteilen: Zunächst erhält man 6 Quadrupel mit jeweils $i = 0$, gefolgt von einem der obigen sechs Tripel. Des Weiteren gibt es die zu $i = 1$ gehörenden drei Quadrupel

$$(1, 0, 0, 2), (1, 0, 1, 1), (1, 1, 0, 1). \quad (4)$$

Zu guter Letzt darf man das Quadrupel $(2, 0, 0, 1)$ nicht vergessen, dessen Wahrscheinlichkeit nach der ersten Pfadregel

$$\binom{3}{2} \cdot \left(\frac{1}{6}\right)^2 \cdot \frac{5}{6} \cdot \left(\frac{5}{6}\right)^2 \cdot \frac{1}{6} \quad (5)$$

beträgt. Diese Regel liefert auch, dass die sechs Quadrupel mit $i = 0$ zusammen die Wahrscheinlichkeit

$$\left(\frac{5}{6}\right)^3 \cdot \mathbb{P}(X = 3) \quad (6)$$

besitzen, denn wegen $i = 0$ stellt sich ja die Situation nach dem ersten Wurf wie zu Beginn dar, nur mit dem Unterschied, dass bereits einmal geworfen wurde. Schreiben wir kurz

$$p_{i,\ell} = \binom{i}{\ell} \cdot \left(\frac{1}{6}\right)^\ell \cdot \left(\frac{5}{6}\right)^{i-\ell} \quad (7)$$

für $i \geq 1$ und $\ell = 0, 1, \dots, i$, so haben die in (4) stehenden Quadrupel (in der angegebenen Reihenfolge) die Wahrscheinlichkeiten

$$\begin{aligned} & p_{3,1} \cdot \left(\frac{25}{36}\right)^2 \cdot \left(\frac{1}{6}\right)^2, \\ & p_{3,1} \cdot \frac{25}{36} \cdot p_{2,1} \cdot \frac{1}{6}, \\ & p_{3,1} \cdot p_{2,1} \cdot \frac{5}{6} \cdot \frac{1}{6}. \end{aligned}$$

Addition dieser sowie der in (5) und (6) aufgeführten Werte liefert das Resultat

$$\begin{aligned} \mathbb{P}(X = 4) &= \left(\frac{5}{6}\right)^3 \cdot \mathbb{P}(X = 3) + \frac{125125}{3359232} \\ &= 0,064012\dots \end{aligned}$$

Im nächsten Abschnitt leiten wir eine allgemeine Rekursionsformel für $\mathbb{P}(X = k)$ her.

4 Eine Rekursionsformel für $\mathbb{P}(X = k)$

Nehmen wir an, wir würden $\mathbb{P}(X = k - 1)$ kennen. Können wir hieraus relativ schnell $\mathbb{P}(X = k)$ herleiten? Die folgenden Überlegungen zeigen, dass das möglich ist. Die das Ereignis $\{X = k\}$ bildenden Abläufe sind k -Tupel (a_1, \dots, a_k) , wobei a_j die Anzahl der Sechsen beim j -ten Wurf des Knobelbechers angibt. Natürlich sind hier nur gewisse k -Tupel möglich, denn es gilt zunächst $a_1 + \dots + a_k = 3$, denn insgesamt werden ja drei Sechsen geworfen. Des Weiteren gilt $a_k > 0$, denn im k -ten Wurf tritt auf jeden Fall die dritte Sechse auf. Wir unterscheiden im Folgenden alle möglichen Abläufe nach dem Wert der ersten Komponente a_1 des k -Tupels. Im Fall $a_1 = 0$ hat der erste Wurf des Bechers keine Sechsen ergeben. Danach stellt sich aber die Situation so dar, dass nach genau $k - 1$ Würfeln der letzte der drei Würfel eine Sechse zeigen soll. Nach der ersten Pfadregel besitzen also alle möglichen k -Tupel (a_1, \dots, a_k) mit $a_1 = 0$ zusammen die Wahrscheinlichkeit

$$\left(\frac{5}{6}\right)^3 \cdot \mathbb{P}(X = k - 1). \quad (8)$$

Als nächstes betrachten wir den Fall $a_1 = 2$, für den es nur das k -Tupel $(2, 0, \dots, 0, 1)$ mit $k - 2$ Nullen gibt. Mit der in (7) eingeführten Schreibweise ist die Wahrscheinlichkeit hierfür gleich

$$p_{3,2} \cdot \left(\frac{5}{6}\right)^{k-2} \cdot \frac{1}{6} = \frac{1}{60} \cdot \left(\frac{5}{6}\right)^k. \quad (9)$$

Von den verbleibenden möglichen k -Tupeln mit $a_1 = 1$ besitzt eines die Gestalt $(1, 0, \dots, 0, 2)$ mit $k - 2$ Nullen; seine Wahrscheinlichkeit ist durch

$$p_{3,1} \cdot \left(\frac{25}{36}\right)^{k-2} \cdot \left(\frac{1}{6}\right)^2 = \frac{1}{50} \cdot \left(\frac{25}{36}\right)^k \quad (10)$$

gegeben. Die anderen Tupel mit $a_1 = 1$ sind von der Gestalt

$$(1, \underbrace{0, \dots, 0}_{i-2}, 1, \underbrace{0, \dots, 0}_{k-1-i}, 1)$$

mit $i \in \{2, 3, \dots, k-1\}$. Der erste Wurf ergibt also eine Sechs, und bei den nächsten $i-2$ Würfeln des mit zwei Würfeln gefüllten Knobelbechers fällt jeweils keine Sechs. Danach erscheint genau eine Sechs, der sich $k-1-i$ Würfel anschließen, bei denen sich jeweils keine Sechs einstellt. Im letzten Wurf des Bechers tritt dann eine Sechs auf. Nach der ersten Pfadregel ist die Wahrscheinlichkeit des obigen Tupels durch

$$p_{3,1} \cdot \left(\frac{25}{36}\right)^{i-2} \cdot 2 \cdot \frac{1}{6} \cdot \frac{5}{6} \cdot \left(\frac{5}{6}\right)^{k-1-i} \cdot \frac{1}{6}$$

und somit durch

$$\frac{1}{36} \cdot \left(\frac{5}{6}\right)^{k+i-2} \quad (11)$$

gegeben. Summiert man diese Werte über i von 2 bis $k-1$, so entsteht der Ausdruck

$$\frac{1}{36} \cdot \left(\frac{5}{6}\right)^k \left\{ 1 + \frac{5}{6} + \left(\frac{5}{6}\right)^2 + \dots + \left(\frac{5}{6}\right)^{k-3} \right\}.$$

Wegen

$$(1 + x + x^2 + \dots + x^\ell)(1 - x) = 1 - x^{\ell+1} \quad (12)$$

nimmt die innerhalb der geschweiften Klammern stehende Summe den Wert

$$\frac{1 - (5/6)^{k-2}}{1 - 5/6} = 6 \cdot \left(1 - \left(\frac{5}{6}\right)^{k-2}\right)$$

an. Folglich ist die Summe der in (11) stehenden Werte über $i = 2, 3, \dots, k-1$ gleich

$$\frac{1}{6} \cdot \left(\frac{5}{6}\right)^k \left(1 - \left(\frac{5}{6}\right)^{k-2}\right).$$

Addiert man hierzu die in (8), (9) und (10) stehenden Ausdrücke, so ergibt sich nach direkter Rechnung die für jedes $k \geq 2$ geltende Rekursionsformel

$$\begin{aligned} \mathbb{P}(X = k) &= \left(\frac{5}{6}\right)^3 \cdot \mathbb{P}(X = k-1) \quad (13) \\ &+ \frac{11}{60} \left(\frac{5}{6}\right)^k - \frac{11}{50} \left(\frac{25}{36}\right)^k. \end{aligned}$$

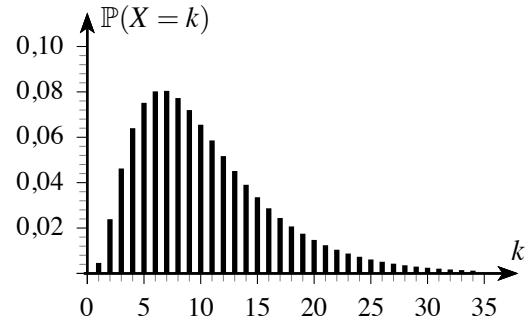


Abb. 2: Stabdiagramm der Verteilung von X

Abbildung 2 zeigt ein mithilfe obiger Rekursionsformel erstelltes Stabdiagramm der Verteilung von X . Es fällt auf, dass diese Verteilung „rechtsschief“ ist, d. h. die Wahrscheinlichkeiten werden zunächst schnell größer und fallen danach langsamer wieder ab. So wird die maximale Wahrscheinlichkeit $\mathbb{P}(X = k)$ für $k = 7$ angenommen, wohingegen wegen $\mathbb{P}(X \leq 8) = 0,4519\dots$ und $\mathbb{P}(X \leq 9) = 0,5239\dots$ erst für $k = 9$ mindestens die Hälfte der gesamten Wahrscheinlichkeitsmasse erreicht ist. Im nächsten Abschnitt werden wir sehen, dass der als physikalischer Schwerpunkt interpretierbare Erwartungswert von X durch $\mathbb{E}(X) = 10,555\dots$ gegeben ist, obwohl wegen $\mathbb{P}(X \leq 10) = 0,5895\dots$ unterhalb des Erwartungswertes schon fast 60% der gesamten Wahrscheinlichkeitsmasse aufgelaufen sind. Insbesondere kann man also in der Tat getrost auf das Eintreten des Ereignisses $\{X \leq 10\}$ wetten.

5 Eine explizite Darstellung für $\mathbb{P}(X = k)$

In diesem Abschnitt zeigen wir, wie man aus der Rekursionsformel (13) eine explizite Darstellung für $\mathbb{P}(X = k)$ erhalten kann. Setzen wir

$$q = \frac{5}{6}, \quad b = \frac{11}{60}, \quad c = \frac{11}{50}, \quad m_k = \mathbb{P}(X = k),$$

so nimmt (13) die Gestalt

$$m_k = q^3 m_{k-1} + bq^k - cq^{2k}, \quad k \geq 2, \quad (14)$$

mit $m_1 = 1/216$ an. Durch wiederholtes Einsetzen in (14) folgt

$$\begin{aligned} m_2 &= q^3 m_1 + bq^2 - cq^4, \\ m_3 &= q^3 (q^3 m_1 + bq^2 - cq^4) + bq^3 - cq^6 \\ &= q^6 m_1 + bq^3 (1 + q^2) - cq^6 (1 + q), \\ m_4 &= q^3 m_3 + bq^4 - cq^8 = \dots \\ &= q^9 m_1 + bq^4 (1 + q + q^2) - cq^8 (1 + q + q^2), \\ m_5 &= q^3 m_4 + bq^5 - cq^{10} = \dots \\ &= q^{12} m_1 + bq^5 (1 + q^2 + q^4 + q^6) - cq^{10} (1 + q + q^2 + q^3). \end{aligned}$$

Schreiben wir kurz

$$\begin{aligned} s_k &= 1 + q^2 + q^4 + \dots + q^{2(k-2)}, \\ t_k &= 1 + q + q^2 + \dots + q^{k-2}, \end{aligned}$$

so ergibt sich allgemein

$$m_k = q^{3(k-1)} m_1 + b q^k s_k - c q^{2k} t_k. \quad (15)$$

Mit (12) erhält man

$$s_k = \frac{1 - q^{2(k-1)}}{1 - q^2}, \quad t_k = \frac{1 - q^{k-1}}{1 - q}.$$

Setzt man die gewonnenen Ausdrücke in (15) ein, so liefert eine direkte Rechnung das gewünschte Resultat

$$\mathbb{P}(X = k) = \frac{3}{5} \cdot \left(\frac{5}{6}\right)^k - \frac{33}{25} \cdot \left(\frac{5}{6}\right)^{2k} + \frac{91}{125} \cdot \left(\frac{5}{6}\right)^{3k}.$$

Dieses kann auch einfacher mithilfe eines im Folgenden angebrachten Unabhängigkeitsargumentes erhalten werden.

6 Ein Unabhängigkeitsargument

Was ändert sich im Hinblick auf die Verteilung von X , wenn wir die drei Würfel unterschiedlich färben würden, um sie unterscheiden zu können? „Natürlich nichts“, würden Schülerinnen und Schüler hier sofort antworten. Der Zufall ließe sich davon nicht beeinflussen. Nun gut, dann sprechen wir im Folgenden von Würfel 1, Würfel 2 und Würfel 3. Warum sollte man die Würfel dann zusammen in einen Knobelbecher tun? Nehmen wir doch für jeden Würfel einen eigenen Becher und werfen jeweils so lange, bis eine Sechs erscheint. Man erkennt an dieser veränderten Versuchsanordnung auch, dass man die Becher gar nicht gleichzeitig werfen muss. Drei verschiedene Personen könnten in getrennten Räumen und auch zeitlich getrennt voneinander jeweils so lange werfen, bis sich die erste Sechs einstellt. Bezeichnet W_j die dafür nötige Anzahl der Würfe für Würfel j , so nimmt die interessierende Zufallsvariable X die Gestalt

$$X = \max(W_1, W_2, W_3)$$

an. Diese Darstellung ist der Schlüssel für weitere Überlegungen und höhere Einsichten. Nach der Pfadregel gilt

$$\mathbb{P}(W_j > k) = \left(\frac{5}{6}\right)^k, \quad k \geq 1,$$

denn das Ereignis $\{W_j > k\}$ tritt ja genau dann ein, wenn die ersten k Würfe mit Würfel j jeweils keine

Sechs ergeben. Geht man zum Gegenereignis über, so folgt

$$\mathbb{P}(W_j \leq k) = 1 - \left(\frac{5}{6}\right)^k, \quad k \geq 1,$$

und zwar für jedes $j = 1, 2, 3$. Nun ist das Maximum von drei Zahlen höchstens gleich k , wenn jede dieser Zahlen höchstens gleich k ist. Da die Würfel sich gegenseitig nicht beeinflussen (W_1, W_2 und W_3 sind stochastisch unabhängige Zufallsvariablen!) gilt

$$\begin{aligned} \mathbb{P}(X \leq k) &= \mathbb{P}(W_1 \leq k, W_2 \leq k, W_3 \leq k) \\ &= \mathbb{P}(W_1 \leq k) \cdot \mathbb{P}(W_2 \leq k) \cdot \mathbb{P}(W_3 \leq k) \\ &= \left(1 - \left(\frac{5}{6}\right)^k\right)^3. \end{aligned} \quad (16)$$

Natürlich würden die Schülerinnen und Schüler die Gültigkeit des zweiten Gleichheitszeichens mit der vertrauten ersten Pfadregel einsehen.

Wegen

$$\mathbb{P}(X > 15) = 1 - \left(1 - \left(\frac{5}{6}\right)^{15}\right)^3 = 0,1823\dots$$

ist jetzt auch die zweite in der Einleitung aufgeworfene Frage beantwortet.

Aus Darstellung (16) erhält man zudem unmittelbar die Verteilung von X , denn die Ereignisse $\{X \leq k-1\}$ und $\{X = k\}$ schließen sich aus und ergeben vereinigt das Ereignis $\{X \leq k\}$. Es gilt also

$$\mathbb{P}(X = k) = \mathbb{P}(X \leq k) - \mathbb{P}(X \leq k-1) \quad (17)$$

und folglich

$$\mathbb{P}(X = k) = \left(1 - \left(\frac{5}{6}\right)^k\right)^3 - \left(1 - \left(\frac{5}{6}\right)^{k-1}\right)^3$$

für jedes $k = 1, 2, \dots$. Durch Anwendung der binomischen Formel und Ausklammern ergibt sich nach kurzer Rechnung die bereits bekannte Darstellung

$$\mathbb{P}(X = k) = \frac{3}{5} \cdot \left(\frac{5}{6}\right)^k - \frac{33}{25} \cdot \left(\frac{5}{6}\right)^{2k} + \frac{91}{125} \cdot \left(\frac{5}{6}\right)^{3k}.$$

Um hieraus den Erwartungswert von X zu erhalten, müsste man gemäß der Formel

$$\begin{aligned} \mathbb{E}(X) &= 1 \cdot \mathbb{P}(X = 1) + 2 \cdot \mathbb{P}(X = 2) + \dots \\ &= \sum_{m=1}^{\infty} m \cdot \mathbb{P}(X = m) \end{aligned}$$

für allgemeines p mit $0 < p < 1$ wissen, was beim Aufsummieren von

$$1 + 2p + 3p^2 + 4p^3 + \dots + kp^{k-1}$$

passiert, wenn k über alle Grenzen wächst. Die Lehrkraft sollte hier natürlich die geometrische Reihe

$$\sum_{m=0}^{\infty} p^m = \frac{1}{1-p}, \quad |p| < 1,$$

und deren erste Ableitung

$$\sum_{m=0}^{\infty} mp^{m-1} = \frac{1}{(1-p)^2} \quad (18)$$

kennen. Hiermit folgt aus der obigen Darstellung von $\mathbb{P}(X = k)$ schnell das Ergebnis

$$\begin{aligned} \mathbb{E}(X) &= \frac{1}{2} \cdot \frac{1}{(1-\frac{5}{6})^2} - \frac{33}{36} \frac{1}{(1-\frac{25}{36})^2} \\ &\quad + \frac{91}{216} \cdot \frac{1}{(1-\frac{125}{216})^2} \end{aligned}$$

und somit

$$\mathbb{E}(X) = \frac{10566}{1001} = 10,555444. \quad (19)$$

Die zu diesem Resultat führende entscheidende Formel (18) kann aber wie folgt auch von Schülerinnen und Schülern eingesehen werden: Bildet man die Summe

$$1 + 2p + 3p^2 + 4p^3 + \dots + kp^{k-1} \quad (20)$$

und multipliziert diese mit $1 - p$, so entsteht der Ausdruck $1 + p + p^2 + \dots + p^{k-1} - kp^k$. Eine weitere Multiplikation mit $1 - p$ liefert das Resultat $1 - (k+1)p^k + kp^{k+1}$. Es gilt also

$$1 + 2p + 3p^2 + \dots + kp^{k-1} = \frac{1 - (k+1)p^k + kp^{k+1}}{(1-p)^2}.$$

Wenn man die Schülerinnen und Schüler jetzt noch davon überzeugen kann, dass die durch $b_k = (k+1)p^k$ definiert Folge (b_k) bei wachsendem k gegen Null strebt, wäre die Gleichung (18) unmittelbar einsichtig geworden, denn es gilt ja $kp^{k+1} = p^2 b_{k-1}$. Diese Überzeugung sollte aber aus der Darstellung

$$0 < b_{k+1} = \left(1 + \frac{1}{k+1}\right) p \cdot b_k$$

resultieren. Da $p < 1$ gilt, gibt es ein $\varepsilon > 0$ mit $p \leq 1 - \varepsilon$. Ist $k \geq 1/\varepsilon - 1$, so ist der Vorfaktor vor b_k in obiger Gleichung höchstens gleich $(1 + \varepsilon)(1 - \varepsilon) =$

$1 - \varepsilon^2$ und somit strikt kleiner als Eins. Hieraus folgt, dass (b_k) bei wachsendem k gegen Null strebt.

Ist man nur am Ergebnis interessiert, kann auch ein CAS zum Einsatz kommen. So führen z. B. mit dem CAS in GeoGebra die drei folgenden Eingaben zum exakten Wert von $\mathbb{E}(X)$:

$$f(k) := (1 - (5/6)^k)^3 - (1 - (5/6)^{(k-1)})^3$$

$$E(n) := \text{Summe}[k * f(k), k, 1, n]$$

$$\text{Grenzwert}[E(n), \infty]$$

7 Der allgemeine Fall

Liegen anstelle von drei Würfeln n gleichartige nicht notwendig echte Würfel vor und bezeichnen wir die Wahrscheinlichkeit für das Auftreten einer Sechs mit p , so unterscheiden wir die Würfel wieder gedanklich und nennen sie „Würfel 1“, ..., „Würfel n “. Modelliert die Zufallsvariable W_j die Anzahl der Würfe, bis sich bei Würfel j die erste Sechs einstellt, so beschreibt

$$X_n = \max(W_1, \dots, W_n) \quad (21)$$

die Anzahl der Würfe, bis auch der letzte Würfel eine Sechs zeigt. Dabei haben wir die bislang als X bezeichnete Zufallsvariable mit dem Index n versehen, um die Anzahl der Würfel hervorzuheben. Wegen

$$\mathbb{P}(W_j > k) = (1-p)^k, \quad k \geq 0,$$

und der Tatsache, dass das Ereignis $\{X_n \leq k\}$ eintritt, wenn jedes der Ereignisse $\{W_1 \leq k\}, \dots, \{W_n \leq k\}$ eintritt, folgt

$$\mathbb{P}(X_n \leq k) = (1 - (1-p)^k)^n$$

und hieraus durch Differenzbildung wie oben

$$\begin{aligned} \mathbb{P}(X_n = k) &= \mathbb{P}(X_n \leq k) - \mathbb{P}(X_n \leq k-1) \\ &= (1 - (1-p)^k)^n - (1 - (1-p)^{k-1})^n, \end{aligned}$$

$k = 1, 2, \dots$ Mit der binomischen Formel

$$(x+y)^n = \sum_{j=0}^n \binom{n}{j} x^j y^{n-j}$$

ergibt sich hieraus die für die Berechnung des Erwartungswertes von X nützliche Darstellung

$$\mathbb{P}(X_n = k) = \sum_{i=1}^n \binom{n}{i} (-1)^i (1-p)^{(k-1)i} ((1-p)^i - 1).$$

Abbildung 3 zeigt ein Stabdiagramm der Verteilung von X_6 für $p = 1/6$. Verglichen mit dem in Abbildung 2 dargestellten Fall $n = 3$ dreier Würfel sind die größten Wahrscheinlichkeiten etwas kleiner

geworden, und das Stabdiagramm ist „etwas nach rechts gerückt“, was nicht weiter verwunderlich ist, denn man wartet ja jetzt darauf, dass auch der letzte von *sechs* Würfeln eine Sechs zeigt. An der schon im Zusammenhang mit Abbildung 2 beobachteten *Rechtsschiefe* hat sich jedoch nichts geändert.

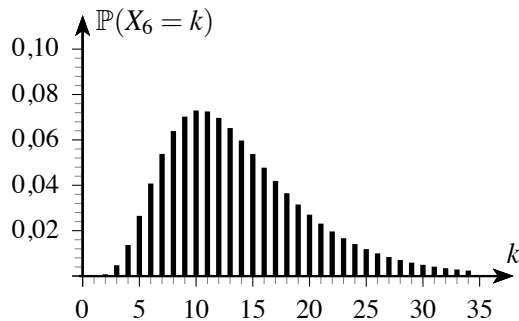


Abb. 3: Stabdiagramm der Verteilung von X_6

Unter Verwendung von (18) erhält man auch den Erwartungswert von X_n . Es gilt

$$\begin{aligned} \mathbb{E}(X_n) &= \sum_{k=1}^{\infty} k \cdot \mathbb{P}(X_n = k) \\ &= \sum_{i=1}^n \binom{n}{i} \frac{(-1)^{i-1}}{1 - (1-p)^i}. \end{aligned}$$

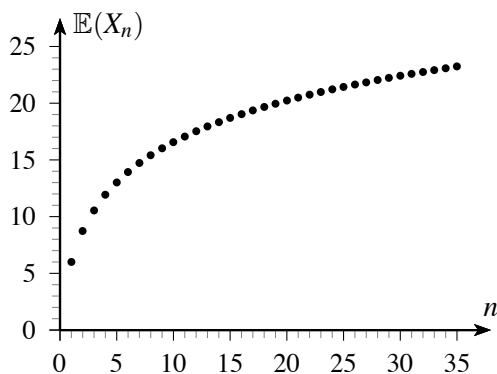


Abb. 4: $\mathbb{E}(X_n)$ in Abhängigkeit von n

Abbildung 4 zeigt den Erwartungswert von X_n als Funktion von n . Offenbar wächst $\mathbb{E}(X_n)$ wesentlich langsamer als n .

Die Darstellung (21) liefert auch unmittelbar die Verteilung von X_{n+1} aus der Verteilung von X_n , denn es gilt

$$\begin{aligned} X_{n+1} &= \max(W_1, \dots, W_n, W_{n+1}) \\ &= \max(X_n, W_{n+1}) \end{aligned}$$

und damit wegen der stochastischen Unabhängigkeit von X_n (als Funktion von W_1, \dots, W_n) und W_{n+1} das

Resultat

$$\begin{aligned} \mathbb{P}(X_{n+1} \leq k) &= \mathbb{P}(X_n \leq k, W_{n+1} \leq k) \\ &= \mathbb{P}(X_n \leq k) \cdot \mathbb{P}(W_{n+1} \leq k) \\ &= (1 - (1-p)^k) \cdot \mathbb{P}(X_n \leq k). \end{aligned}$$

Da X_n nach (21) ein Maximum aus unabhängigen und identisch verteilten Zufallsvariablen ist, könnte man (analog zu vergleichbaren Situationen, siehe z. B. Henze (2016a)) versuchen, geeignete reelle Folgen (u_n) und (v_n) zu finden, sodass die Zufallsvariable $u_n(X_n - v_n)$ beim Grenzübergang $n \rightarrow \infty$ eine nichtausgeartete Limesverteilung besitzt. Dass dieser Versuch (wegen der zugrunde liegenden geometrischen Verteilung) zum Scheitern verurteilt ist, zeigt Beispiel 3.1.5. in Embrechts, Klüppelberg und Mikosch (1997).

8 Ein Spiel zum Tag der offenen Tür

Ein in der Einleitung angesprochenes, auf der Anzahl X der Würfe bis zum Auftreten der letzten Sechs gründendes Spiel für den Tag der offenen Tür könnte wie folgt aussehen: Man zahlt einen Euro Einsatz und gewinnt im Fall $X = 1$ stolze 25 Euro. Im Fall $X = 2$ erhält man 10 Euro, und im Fall $X = 3$ werden noch 5 Euro als Gewinn ausbezahlt. Hat man Pech und muss mehr als 15 mal werfen, bis die Sechs erscheint, gibt es den Einsatz zurück. Bezeichnet die Zufallsvariable G den Gewinn bei diesem Spiel (von dem noch der Einsatz abzuziehen ist), so nimmt G die Werte 25, 10, 5 und 1 an. Nach früheren Überlegungen sind die Wahrscheinlichkeiten für diese Ausgänge durch

$$\begin{aligned} \mathbb{P}(G = 25) &= \mathbb{P}(X = 1) = \frac{1}{216}, \\ \mathbb{P}(G = 10) &= \mathbb{P}(X = 2) = \frac{1115}{6^6}, \\ \mathbb{P}(G = 5) &= \mathbb{P}(X = 3) = \frac{466075}{6^9}, \\ \mathbb{P}(G = 1) &= \mathbb{P}(X > 15) = 1 - \left(1 - \left(\frac{5}{6}\right)^{15}\right)^3 \end{aligned}$$

gegeben. Eine direkte Rechnung ergibt

$$\begin{aligned} \mathbb{E}(G) &= 25 \cdot \mathbb{P}(G = 25) + 10 \cdot \mathbb{P}(G = 10) \\ &\quad + 5 \cdot \mathbb{P}(G = 5) + 1 \cdot \mathbb{P}(G = 1) \\ &= 0,7683\dots \end{aligned}$$

Pro Spiel werden also auf die Dauer im Mittel knapp 77 Cent ausbezahlt und folglich ca. 23 Cent für die Klassenkasse einbehalten.

9 Weitere Wartezeitprobleme

Die behandelte Fragestellung lässt sich unter dem Oberbegriff *Wartezeitprobleme* einordnen. Bei einem stochastischen Wartezeitproblem wird zu jedem der Zeitpunkte $1, 2, \dots$ ein Zufallsexperiment durchgeführt, wobei man auf das *erstmalige Eintreten* eines vorher definierten Ereignisses wartet. Die interessierende Zufallsvariable ist dann die Anzahl der nötigen Experimente, bis dieses Ereignis eintritt.

Einfache Wartezeitprobleme sind das Warten auf den ersten oder allgemeiner den r -ten Treffer in einer Bernoulli-Kette (siehe z.B. Henze (2016b), Kapitel 23). Schwierigere Fragen ergeben sich beim Warten auf ein bestimmtes Muster wie z.B. 101 in einer Bernoulli-Kette (siehe etwa Hombas (1997), Humenberger (2000a), Humenberger (2000b) oder Henze (2001)). Das paradoxe Phänomen eines unendlichen Erwartungswertes beim Warten auf erstmals gleich viele Treffer wie Nieten in einer Bernoulli-Kette mit Trefferwahrscheinlichkeit $1/2$ wird ausführlich in Henze (2013) behandelt. Colwell und Gillett (1984) untersuchten den Erwartungswert der Anzahl der Versuche in einer Bernoulli-Kette, bis sich erstmals m direkt aufeinander folgende Treffer oder n direkt aufeinander folgende Nieten einstellen; für den Spezialfall $m = n = 2$ siehe auch Dorner (2017).

Beliebte Wartezeitprobleme im Fächer-Modell sind das Warten auf eine vollständige Serie (siehe etwa Jäger und Schupp (1987), Treiber (1988), Boneh und Hofri (1997), Althoff (2000), Haake (2006) oder Henze (2016a)) oder das Warten auf die erste Kollision, die gern als Doppelgeburtstag interpretiert wird, siehe etwa Schrage (1990), Barth und Haller (2012), Arnold und Glaß (2013), Riehl (2014) oder Henze (2015).

Ein interessantes Wartezeitproblem entsteht auch beim rein zufälligen Ziehen ohne Zurücklegen aus einer Urne mit r roten und s schwarzen Kugeln. Die interessierende Zufallsvariable X ist hier die Anzahl der nötigen Züge, bis *jede* rote Kugel gezogen wurde. Es gilt $\mathbb{E}(X) = r(s+1)/(r+1)$, was für Schülerinnen und Schüler auf den ersten Blick verwunderlich ist. So ergibt sich im Spezialfall $r = 15, s = 75$ der vermeintlich hohe Wert $\mathbb{E}(X) = 85,3125$, obwohl doch nur 15 der insgesamt 90 Kugeln gezogen werden müssen, siehe Henze und Humenberger (2011).

10 Fazit

Wer hätte gedacht, dass ein einfaches Würfelspiel vielfältige Lösungswege mit Querverbindungen zur Algebra und zur Analysis mit sich bringt? Wie Abbildung 5 zeigt, sind aber auch Vernetzungen mit der Linearen Algebra möglich.

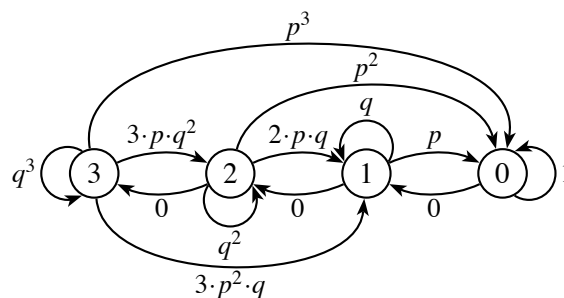


Abb. 5: Prozessdiagramm; $p = \frac{1}{6}, q = \frac{5}{6}$

Hier symbolisieren die vier Zustände 3, 2, 1 und 0 die vier möglichen Anzahlen an Würfeln, die nach der Regel benutzt werden. Zu Beginn wird dreimal gewürfelt. Dann können Durchführungen mit zwei und einem Würfel auftreten. Das Spiel ist zu Ende, wenn kein Würfel mehr vorhanden ist, also Zustand 0 erreicht wird. In dieser Sichtweise liegt eine Markov-Kette vor, sodass die Verteilung von X auch mithilfe stochastischer Matrizen erhalten werden kann.

Genauso wichtig wie Querverbindungen innerhalb der Mathematik sind aber auch die Vernetzungen zwischen Schul- und Hochschulmathematik. Angehende Lehrerinnen und Lehrer sollten unbedingt anhand geeigneter Beispiele das an einer Universität Gelernte mit der Schulmathematik verknüpfen. Eine solche Verknüpfung könnte mit dazu beitragen, die Kluft zwischen gymnasialer und universitärer Mathematik zu verringern, weil die Einsicht wächst, dass „höhere Mathematik“ einen konkreten Nutzen für den schulischen Alltag stiften kann.

Wir hoffen, dass die Leserinnen und Leser angeregt werden, bei Problemstellungen selbst einmal tiefer zu bohren, Fragen zu stellen und verschiedene Lösungswege zu suchen. Man muss sich bisweilen nicht nur auf das eingeführte Schulbuch verlassen!

Im Unterricht kann der hier dargelegte Weg natürlich nur unter gewissen Reduktionen beschränkt werden, wobei der Phantasie (fast) keine Grenzen gesetzt sind. So ist es z. B. möglich, die Fragestellung auch zur Binnendifferenzierung einzusetzen: Wieso sollten nicht in einem Kurs verschiedene Gruppen die drei Fälle $k = 1, 2, 3$ getrennt betrachten? Übrigens eignen sich Problemstellungen dieser Art auch sehr gut für Seminararbeiten.

Leider gibt es schon genügend „BLABLA-Aufgaben“ zur Binomialverteilung, bei denen der (oft mit einer an den Haaren herbeigezogenen vermeintlichen Einkleidung versehene) Aufgabentext keinerlei Rolle spielt und nur nach n , p und k Ausschau gehalten werden muss, um die Werte dann in einen Rechner einzugeben. Die in Grundkursen fast ausschließlich behandelte Binomialverteilung kommt hier zwar auch zum Einsatz, erfordert aber deutlich mehr Nachdenken. Da ist es gut, mit Schülerinnen und Schülern auch einmal eine für stochastische Extremwertprobleme typische rechtsschiefe Verteilung zu behandeln, die zudem einen Erwartungswert besitzt, der nicht einfach mit $n \cdot p$ zu berechnen ist.

An einigen Stellen wurde (hoffentlich) deutlich, dass ein Rechnereinsatz sinnvoll ist – sowohl bei einer Simulation (Tabellenkalkulation) als auch bei bestimmten Berechnungen (CAS-Einsatz).

Danksagung: Die Autoren danken den beiden Gutachtern für diverse Verbesserungsvorschläge.

Literatur

- Althoff, H. (2000): Zur Berechnung der Wahrscheinlichkeit für das Vorliegen einer vollständigen Serie. *Stochastik in der Schule* **20** (1), S. 18–20.
- Arnold, M. und Glaß, W. (2013): Simple Approximation Formulas of the Birthday Problem. *Amer. Mathem. Monthly* **120**, 645–648.
- Barth, F. und Haller, R. (2012): Besetzungen und Geburtstage. *Stochastik in der Schule* **32** (3), 20–27.
- Boneh, A. und Hofri, M. (1997): The Coupon Collector Problem revisited – A Survey of engineering Problems and computational Methods. *Stoch. Models* **13**, 39–66.
- Colwell, D.J. und Gillett, J.R. (1984): Runs in Bernoulli trials. *Math. Gaz.* **68**, 196–199.
- Dorner, Ch. (2017): Einfache Wartezeitprobleme. *Stochastik in der Schule* **37** (3), 8–15.
- Embrechts, P., Klüppelberg, C., und Mikosch, Th. (1997): Modeling extremal Events. 1. Auflage. Berlin u.a.: Springer.
- Haake, H. (2006): Elementare Zugänge zum Problem der vollständigen Serie. *Stochastik in der Schule* **26** (3), S. 28–33.
- Henze, N. (2001): Muster in Bernoulli-Ketten. *Stochastik in der Schule* **21** (2), S. 2–10.
- Henze, N., und Humenberger, H. (2011): Stochastische Überraschungen beim Spiel Bingo. *Stochastik in der Schule* **31** (3), S. 2–11.
- Henze, N. (2013): Irrfahrten und andere Zufälle. Heidelberg: Springer Spektrum.
- Henze, N. (2015): Stochastische Extremwertprobleme im Fächer-Modell I: Minima von Wartezeiten und Kollisionsprobleme. *Stochastik in der Schule* **35** (3), S. 24–30.
- Henze, N. (2016a): Stochastische Extremwertprobleme im Fächer-Modell II: Maxima von Wartezeiten und Sammelbilderprobleme. *Stochastik in der Schule* **36** (1), S. 2–9.
- Henze, N. (2016b): Stochastik für Einsteiger. 11. Auflage. Heidelberg: Springer Spektrum.
- Hombas, V.C. (1997): Waiting time and expected waiting time – paradoxical situations. *Amer. Statist.* **51**, 130–133.
- Humenberger, H. (2000a): Überraschendes bei Münzwurfserien. *Stochastik in der Schule* **20** (1), 4–17.
- Humenberger, H. (2000b): Kopf-Adler-Muster in Münzwurfserien, unendliche Reihen und Fibonacci-Zahlen. *Stochastik in der Schule* **20** (3), 15–22.
- Jäger, J. und Schupp, H. (1987): Wann sind alle Kästchen besetzt? Oder: Das Problem der vollständigen Serie am Galton-Brett. *Did. der Mathem.* **15**, 37–48.
- Riehl, G. (2014): Alte Geburtstagsprobleme – neu gelöst. *Mathem. Semesterber.* **61**, S. 215–232.
- Schrage, G. (1990): Ein Geburtstagsproblem. *Mathem. Semesterber.* **61**, 251–257.
- Treiber, D. (1988): Zur Wartezeit auf eine vollständige Serie. *Didakt. Mathem.* **16**, 235–237.

Anschriften der Verfasser:

Prof. Dr. Norbert Henze
 Institut für Stochastik
 Karlsruher Institut für Technologie (KIT)
 Englerstr. 2
 76131 Karlsruhe
 Henze@kit.edu

Reimund Vehling
 Studienseminar für das Lehramt an Gymnasien
 Hannover I
 vehling@icloud.com