Wann gleichen sich Treffer und Nieten erstmals aus?

NORBERT HENZE, KARLSRUHE

Zusammenfassung: *Eine faire Münze wird solange geworfen, bis erstmalig jede der beiden Seiten gleich oft oben lag. Die Anzahl der dazu nötigen Würfe sei mit W bezeichnet. In diesem Aufsatz wird die Verteilung von W hergeleitet. Überraschenderweise gilt $\mathbb{E}(W) = \infty$. Im Mittel wartet man also unendlich lange, bis sich ein Unentschieden zwischen Treffern und Nieten einstellt.*

1 Einleitung

Es ist nicht zu hoch gegriffen, unabhängige Bernoulli-Versuche mit zwei gleich wahrscheinlichen Ausgängen als die „Mutter allen Zufalls“ zu bezeichnen. Eine Einkleidung hierfür ist das wiederholte Werfen einer fairen Münze. Lassen Sie doch einmal Ihre Schülerinnen und Schüler schätzen, wie oft man eine solche Münze „im Mittel“ werfen muss, bis erstmals jede der beiden Seiten *Zahl* und *Wappen* gleich oft oben lag. Wissen Sie es? Vermutlich nicht, denn diese ganz natürliche Frage ist üblicherweise nicht Stoff einer einführenden Vorlesung in die Stochastik.

Bezeichnen wir *Zahl* als *Treffer* und *Wappen* als *Niete*, und codieren wir diese Fälle mit 1 bzw. 0, so stellt sich ein erstmaliger Ausgleich zwischen Treffern und Nieten genau dann nach nur zwei Würfen der Münze ein, wenn die beiden ersten Würfe von links nach rechts gelesen 10 oder 01 ergeben, und die Wahrscheinlichkeit hierfür ist $\frac{1}{2}$. Es gilt also

$$\mathbb{P}(W = 2) = \frac{1}{2}.$$

Hierbei gibt die Zufallsgröße W die Anzahl der Bernoulli-Versuche (Münzwürfe) an, bis erstmals gleich viele Treffer wie Nieten aufgetreten sind. In 50% aller Fälle gibt es also schon nach nur zwei Würfen einen solchen Gleichstand, und mit dieser Erkenntnis drängt sich die Vermutung geradezu auf, dass der Erwartungswert von W irgendwo im einstelligen Bereich liegt.

Hier kann es nicht schaden, Erfahrungen mit Meister Zufall zu sammeln. Diesbezüglich zeigen die Daten von Abb. 1 zeilenweise zu lesende Ergebnisse von Würfen einer handelsüblichen Münze, die ich selbst erhalten habe.

```
1 1 0 1 1 1 0 0 0 0 1 1 0 1 1 1 0 1 0 1 0 1 1 0 0
0 1 1 1 1 0 1 0 0 0 0 1 1 0 0 1 1 0 1 1 1 1 0 0 1
1 0 0 0 1 0 1 1 1 0 1 1 1 1 1 0 1 1 0 1 1 0 1 1 1
1 0 0 0 1 1 0 1 0 0 1 0 0 0 1 1 0 0 0 0 0 1 1 0 1
0 1 0 0 0 1 0 1 0 1 1 1 1 1 1 1 0 0 1 1 0 0 1 0 1
1 0 1 1 0 1 1 0 1 1 0 1 0 1 0 0 0 1 1 1 0 1 0 1 0
1 0 0 1 0 1 1 1 1 0 0 0 1 0 0 1 1 0 0 0 0 0 0 1 1
1 1 0 0 1 0 1 0 1 0 1 0 0 1 1 0 1 0 1 0 0 0 1 1 0
```

Abb. 1: Ergebnisse von 200 Münzwürfen (zeilenweise gelesen)

In diesem Fall tritt der erste Gleichstand zwischen Treffern und Nieten nach dem zehnten Wurf auf. Zählt man danach wieder neu, so stellt sich in 190 Würfen kein weiterer Gleichstand zwischen Treffern und Nieten ein! Das 100-malige Warten auf den ersten Gleichstand mithilfe von Computersimulationen ergab 49 mal den Wert 2 und 12 mal den Wert 4, aber unter anderem auch je einmal die Werte 50, 74, 116, 150, 154, 240, 1714 und 15910. Diese Beobachtung lässt aufhorchen und macht neugierig. Ist der Erwartungswert von W vielleicht deutlich größer als vermutet? Da ein solcher Gleichstand nur nach einer geraden Anzahl von Würfen auftreten kann, nimmt W die möglichen Werte $2, 4, 6, \ldots$ an, und der Erwartungswert berechnet sich über die Formel

$$\mathbb{E}(W) = \sum_{n=1}^{\infty} 2n \cdot \mathbb{P}(W = 2n). \tag{1}$$

Wir sind somit gefordert, einen Ausdruck für $\mathbb{P}(W = 2n)$ für jedes $n \geq 1$ herzuleiten, also die Verteilung von W zu bestimmen. Diesem Ziel dient unter anderem dieser Aufsatz.

Fragen Sie ihre Schülerinnen und Schüler, mit welchen Wahrscheinlichkeiten die Zufallsgröße W die Werte 4 bzw. 6 annimmt, so werden Sie vermutlich schnell folgende Antworten erhalten: Das Ereignis $\{W = 4\}$ tritt genau dann ein, wenn die ersten vier Bernoulli-Versuche 1100 oder 0011 ergeben. Da jede dieser Vierersequenzen die Wahrscheinlichkeit $\frac{1}{16}$ besitzt, gilt

$$\mathbb{P}(W = 4) = \frac{2}{16} = \frac{1}{8}. \tag{2}$$

Darüber hinaus nimmt W genau dann den Wert 6 an, wenn für die Ausgänge der ersten sechs Münzwürfe einer der vier Fälle 111000, 110100, 000111 oder 001011 zutrifft. Es folgt also

$$\mathbb{P}(W = 6) = 4 \cdot \frac{1}{64} = \frac{1}{16}. \tag{3}$$

Stochastik in der Schule 42 (2022) 1, S. 14–20

Natürlich kann man auch noch $\mathbb{P}(W = 8)$ und $\mathbb{P}(W = 10)$ durch Abzählen günstiger Anfangsverläufe der Bernoulli-Versuche bestimmen, aber dieses Vorgehen liefert keinerlei Anhaltspunkt, wie ein Ausdruck für $\mathbb{P}(W = 2n)$ für allgemeines n aussehen könnte. Eine kleine strukturelle Einsicht, die man jedoch vielleicht schon aufgrund der Fälle $W = 2$, $W = 4$ und $W = 6$ gewonnen hat, ist die, dass $\mathbb{P}(W = 2n)$ von der Gestalt

$$\mathbb{P}(W = 2n) = 2 \cdot \frac{u_{2n}}{2^{2n}} \tag{4}$$

ist. Dabei bezeichnet u_{2n} die Anzahl der Sequenzen aus Einsen und Nullen der Länge $2n$, die *mit einer Eins beginnen* (Symmetrieargument!) und gleich viele Einsen wie Nullen aufweisen. Dabei darf aber im Fall $n \geq 2$ keine kürzere Anfangssequenz gerader Länge gleich viele Einsen wie Nullen enthalten.

Fragen Sie Ihre Schülerinnen und Schüler doch einmal nach dem Wert von u_8! Diese Aufgabe erfordert Fallunterscheidungen, und man muss stets daran denken, dass bei allen günstigen Sequenzen der Länge 8 von links nach rechts gelesen der Saldo zugunsten der Einsen immer positiv ausfallen muss und erst ganz am Ende ein Ausgleich zwischen Einsen und Nullen stattfindet. Hiermit schälen sich als einzige günstige Achtersequenzen

11110000, 11101000, 11100100,
11011000, 11010100

heraus. Also gilt $u_8 = 5$ und damit wegen (4)

$$\mathbb{P}(W = 8) = 2 \cdot \frac{5}{256} = \frac{5}{128}. \tag{5}$$

Im nächsten Abschnitt leiten wir eine geschlossene Formel für $\mathbb{P}(W = 2n)$ und damit für u_{2n} her. Dabei gehen wir einen lehrreichen kleinen Umweg.

2 Die Verteilung von W

Stellen Sie sich vor, eine Ihrer Schülerinnen fragt Sie: „Kann es nicht sein, dass sich nie ein Gleichstand zwischen Treffern und Nieten ergibt?“ Hier sollten Sie auf festem wahrscheinlichkeitstheoretischem Boden stehen und über das folgende Hintergrundwissen verfügen.

Ein stochastisches Modell für gedanklich beliebig oft unter gleichen, unabhängigen Bedingungen durchgeführte Münzwürfe sind stochastisch unabhängige Zufallsgrößen $X_1, X_2, \ldots$ mit der Eigenschaft

$$\mathbb{P}(X_j = 1) = \mathbb{P}(X_j = 0) = \frac{1}{2}, \quad j \geq 1. \tag{6}$$

Dabei modelliert X_j das Ergebnis des j-ten Münzwurfs, $j \geq 1$.

Ein gemeinsamer Grundraum, auf dem die X_j als Abbildungen definiert werden können, ist die überabzählbare Menge $\Omega = \{0,1\}^{\mathbb{N}}$ aller unendlichen Folgen $\omega = (\omega_1, \omega_2, \ldots)$ aus Nullen und Einsen. In dieser Folge steht das j-te Folgenglied für das Resultat des j-ten Münzwurfs, und es ist dann $X_j(\omega) := \omega_j$, $j \geq 1$, gesetzt. Die Wahrscheinlichkeit aller Folgen aus Ω, deren erste n Folgenglieder festgelegt sind, ist gleich $(1/2)^n$. Hieraus folgt, dass jede einzelne Folge die Wahrscheinlichkeit null besitzt; es gilt also $\mathbb{P}(\{\omega\}) = 0$ für jedes $\omega \in \Omega$.

Dieses Modell für (gedanklich) unendlich viele Bernoulli-Versuche ist so reichhaltig, dass es ganze Bücher füllt (wobei meist allgemeiner $\mathbb{P}(X_j = 1) = p = 1 - \mathbb{P}(X_j = 0)$ für ein p mit $0 < p < 1$ angenommen wird), siehe z.B. Lesigne (2005) oder Henze (2018), Kapitel 2 und Abschnitte 4.1-4.4. Betrachtet man hier nur die ersten n Bernoulli-Versuche, so ist die als Anzahl der erzielten Treffer interpretierbare Zufallsgröße $X_1 + \ldots + X_n$ binomialverteilt mit Parametern n und p.

Mit dem obigen theoretischen Hintergrund können wir die *Wartezeit* W (im Sinne der Anzahl der nötigen Versuche) auf den ersten Gleichstand zwischen Treffern und Nieten formal als

$$W(\omega) := \inf\{2k : k \geq 1 \text{ und } X_1(\omega) + \ldots + X_{2k}(\omega) = k\},$$

$\omega \in \Omega$, definieren. Da die $\mathrm{Bin}(2k, 1/2)$-verteilte Zufallsgröße $X_1 + \ldots + X_{2k}$ die Anzahl der Treffer unter den ersten $2k$ Bernoulli-Versuchen angibt, bedeutet das Ereignis $\{X_1 + \ldots + X_{2k} = k\}$ ja gerade, dass man in $2k$ Versuchen gleich viele Treffer wie Nieten beobachtet hat.

Für diejenigen Folgen $\omega = (\omega_1, \omega_2, \ldots)$ aus Ω wie z.B. den konstanten Folgen $\omega = (1,1,1,\ldots)$ und $\omega = (0,0,0,\ldots)$, für die es kein k mit $X_1(\omega) + \ldots + X_{2k}(\omega) = k$ gibt, nimmt die Zufallsgröße W als Infimum über die leere Menge definitionsgemäß den Wert unendlich an. Wir werden aber sehen, dass $\mathbb{P}(W < \infty) = 1$ gilt, sodass das Ereignis $\{W = \infty\}$ nur mit der Wahrscheinlichkeit null auftritt. Der fragenden Schülerin kann man also entgegnen: „Ja, das kann rein theoretisch passieren, aber nur mit der Wahrscheinlichkeit null.“ Natürlich hätte man diese Antwort auch ohne jegliche Stochastikkenntnisse „aus dem Bauch heraus“ gegeben.

Um die Wahrscheinlichkeit $\mathbb{P}(W = 2n)$ zu erhalten, gehen wir einen kleinen, lehrreichen und recht oft auftretenden Umweg. Wir bestimmen zunächst die Wahrscheinlichkeit des Ereignisses, dass wir *länger als $2n$ Versuche* auf den ersten Gleichstand zwischen Treffern und Nieten warten müssen, also $\mathbb{P}(W > 2n)$. Da die Zufallsgröße W nur geradzahlige Werte annehmen kann, geht es im Folgenden also um die Wahrscheinlichkeit des Ereignisses $\{W \geq 2n+2\}$. Da das Ereignis $\{W \geq 2n\}$ die Vereinigung der unvereinbaren Ereignisse $\{W = 2n\}$ und $\{W \geq 2n+2\}$ ist, gilt

$$\mathbb{P}(W \geq 2n) = \mathbb{P}(W = 2n) + \mathbb{P}(W \geq 2n+2)$$

und folglich

$$\mathbb{P}(W = 2n) = \mathbb{P}(W \geq 2n) - \mathbb{P}(W \geq 2n+2). \quad (7)$$

Wir werden sehen, dass

$$\mathbb{P}(W \geq 2n+2) = \frac{\binom{2n}{n}}{2^{2n}}, \quad n \geq 1, \quad (8)$$

gilt, wobei diese Gleichung wegen $\binom{0}{0} = 1$, $2^0 = 1$ und $\mathbb{P}(W \geq 2) = 1$ auch für $n = 0$ erfüllt ist. Mithilfe von (7) erhält man dann die Verteilung von W durch elementares Rechnen mit Fakultäten (Übungsaufgabe für Ihre Schülerinnen und Schüler!) zu

$$\mathbb{P}(W = 2n) = \frac{\binom{2(n-1)}{n-1}}{2^{2(n-1)}} \cdot \frac{1}{2n}, \quad n \geq 1. \quad (9)$$

Als weitere Übungsaufgabe könnte man nachprüfen, ob diese allgemeine Formel die schon erhaltenen Spezialfälle (2), (3) und (5) bestätigt.

Entscheidend ist also nur die Herleitung von (8). Diese Gleichung ist insofern bemerkenswert, als der rechts stehende Quotient

$$\binom{2n}{n} \left(\frac{1}{2}\right)^n \left(1 - \frac{1}{2}\right)^{2n-n}$$

nach der Formel von Bernoulli gleich der Wahrscheinlichkeit ist, in $2n$ unabhängigen Bernoulli-Versuchen mit gleicher Trefferwahrscheinlichkeit $\frac{1}{2}$ genau n Treffer zu erzielen; es gilt also

$$\mathbb{P}(W \geq 2n+2) = \mathbb{P}(X_1 + \ldots + X_{2n} = n).$$

Um (8) herzuleiten, beschreiben wir die Ergebnisse der ersten $2n$ Bernoulli-Versuche als Wege in einem Koordinatensytem. Dabei wird auf der horizontalen Achse die Anzahl der Versuche aufgetragen, und es wird jeder Treffer als Aufwärtsschritt und jede Niete als Abwärtsschritt dargestellt. Abb. 1 zeigt einen solchen Weg in Form eines Polygonzuges.

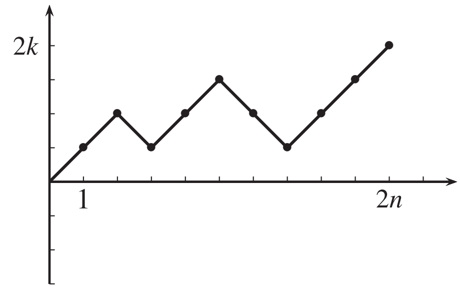

Abb. 1: Bernoulli-Sequenz der Länge $2n$ als Weg

Obwohl dieser Weg der Länge $2n$ den konkreten Fall $2n = 10$ und die Sequenz 1101100111 repräsentiert, beschreibe er allgemein den Fall, dass es im Laufe von $2n$ Bernoulli-Versuchen nie zu einem Ausgleich von Treffern und Nieten kommt (der Weg trifft nicht die horizontale Achse), dass der erste Versuch einen Treffer ergibt (der Weg beginnt mir einem Aufwärtsschritt), und dass am Ende die Differenz zwischen der Anzahl der Nieten und der Treffer gleich $2k$ ist (der Weg endet im Punkt $(2n, 2k)$).

Codiert man eine Niete nicht mit 0, sondern mit -1, setzt man also $Y_j := 2X_j - 1$, $j \geq 1$, so gilt

$$\sum_{j=1}^{2n} Y_j = 0 \iff \sum_{j=1}^{2n} X_j = n.$$

Diese Codierung wird häufig verwendet, denn wenn man die Punkte $(j, Y_1 + \ldots + Y_j)$, $j \geq 1$, in einem Koordinatensystem aufträgt und zu einem Polygonzug verbindet, so liefert jeder Gleichstand zwischen Treffern und Nieten eine Nullstelle, siehe z.B. Henze (2018), Kapitel 2.

Das Ereignis $\{W \geq 2n+2\}$ tritt genau dann ein, wenn in den ersten $2n$ Bernoulli-Versuchen kein Ausgleich zwischen Treffern und Nieten stattfindet, also der zugehörige Weg nie die horizontale Achse trifft. Insgesamt gibt es 2^{2n} gleich wahrscheinliche Wege der Länge $2n$, und aus Symmetriegründen (Spiegelung an der horizontalen Achse!) ist die Anzahl der für das Ereignis $\{W \geq 2n+2\}$ günstigen Wege gleich dem *Doppelten der Anzahl der positiven Wege*, die wie der in Abb. 1 dargestellte Weg mit einem Aufwärtsschritt beginnen und danach nie die horizontale Achse treffen. Alle positiven Wege enden aber in genau einem der Punkte $(2n, 2k)$ mit $k \in \{1, 2, \ldots, n\}$.

Schreiben wir kurz $P_{n,k}$ für die Anzahl der positiven Wege der Länge $2n$, die im Punkt $(2n, 2k)$ enden, so gilt also

$$\mathbb{P}(W \geq 2n+2) = \frac{2}{2^{2n}} \cdot \sum_{k=1}^{n} P_{n,k}. \quad (10)$$

Wie viele positive Wege von $(0,0)$ nach $(2n,2k)$ gibt es? Um diese Frage zu beantworten, bestimmen wir zunächst die mit $A_{n,k}$ bezeichnete Anzahl *aller* von $(0,0)$ nach $(2n,2k)$ verlaufenden Wege und ziehen davon die mit $B_{n,k}$ bezeichnete Anzahl aller Wege von $(0,0)$ nach $(2n,2k)$ ab, die nach dem Start *mindesten einmal die horizontale Achse treffen*. Es gilt also

$$P_{n,k} = A_{n,k} - B_{n,k}. \tag{11}$$

Da jeder Weg von $(0,0)$ nach $(2n,2k)$ genau $n+k$ Aufwärts- und $n-k$ Abwärtsschritte durchläuft und wir von den insgesamt $2n$ Zeitpunkten $0,1,2,\dots,2n-1$ genau $n+k$ auswählen müssen, zu denen die Aufwärtsschritte stattfinden (für den Weg in Abb. 1 sind das die Zeitpunkte 0, 1, 3, 4, 7, 8 und 9), gilt

$$A_{n,k} = \binom{2n}{n+k}. \tag{12}$$

Jeder Weg von $(0,0)$ nach $(2n,2k)$, der nach seinem Start mindestens einmal die horizontale Achse trifft, weist entweder zu Beginn einen Abwärtsschritt auf (Fall 1), oder er beginnt mit einem Aufwärtsschritt und trifft dann in seinem Verlauf mindestens einmal die horizontale Achse (Fall 2). Die Wege zu Fall 1 lassen sich unmittelbar abzählen, denn nach dem ersten Abwärtsschritt durchläuft der Weg $2n-1$ Schritte, von denen $n+k$ Aufwärtsschritte und $n-k-1$ Abwärtsschritte sind. Die Anzahl der Wege zu Fall 1 ist also gleich dem Binomialkoeffizienten $\binom{2n-1}{n+k}$.

Der zweite Fall ist in Abb. 2 veranschaulicht.

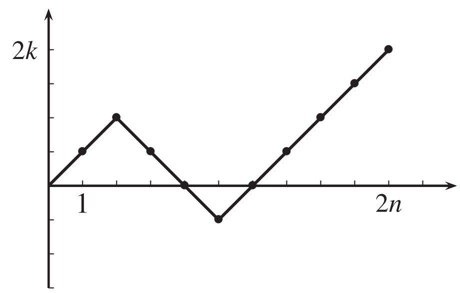

Abb. 2: Weg zu Fall 2

Um die zu Fall 2 gehörenden Wege abzuzählen, verwenden wir eine dem französischen Mathematiker Désiré André (1840–1918) zugeschriebene Methode, die als *Spiegelungsprinzip* bezeichnet wird (siehe z.B. Henze (2018), S. 9 ff.). Hierzu ordnen wir einem Weg wie in Abb. 2 einen neuen Weg zu, indem wir den Teilweg bis zum ersten Treffpunkt mit der horizontalen Achse an dieser Achse spiegeln und den sich anschließenden Teilweg unverändert lassen. Hierdurch ergibt sich der in Abb. 3 gezeigte Weg.

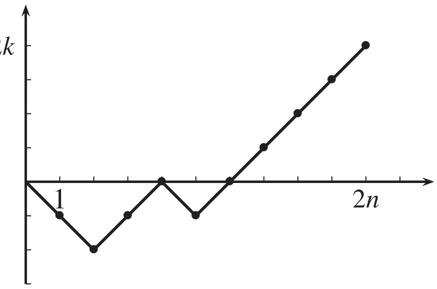

Abb. 3: Der Weg aus Abb. 2 nach (Teil-)Spiegelung

Der durch diese Vorschrift erhaltene Weg durchläuft zu Beginn einen Abwärtsschritt und endet im Punkt $(2n,2k)$; er ist also ein Weg, auf den Fall 1 zutrifft. Offenbar gehen bei dieser Spiegelungsvorschrift verschiedene Wege auf verschiedene Weg über (die Abbildung ist injektiv). Die Spiegelungsvorschrift ist aber auch eine surjektive Abbildung (und damit eine bijektive Abbildung zwischen den Wegen der Fälle 1 und 2), denn jeder Weg, für den Fall 1 zutrifft, kann durch Spiegelung seines Anfangteils bis zum erstmaligen Treffpunkt mit der horizontalen Achse (unter Beibehaltung des sich anschließenden Teilweges) in einen Weg überführt werden, für den Fall 2 zutrifft. Die beiden Fälle 1 und 2 beinhalten also gleich viele Wege, sodass

$$B_{n,k} = 2\binom{2n-1}{n+k}$$

folgt. Zusammen mit (11) und (12) ergibt sich also

$$P_{n,k} = \binom{2n}{n+k} - 2\binom{2n-1}{n+k}, \quad k = 1,\dots,n. \tag{13}$$

Dabei verschwindet der Subtrahend für $k=n$.

Im Hinblick auf (10) müssen wir diese Binomialkoeffizienten noch über k summieren. Wegen

$$\sum_{j=0}^{2n} \binom{2n}{j} = 2^{2n}$$

und der Symmetriebeziehung $\binom{2n}{j} = \binom{2n}{2n-j}$ folgt

$$2\sum_{k=1}^{n} \binom{2n}{n+k} + \binom{2n}{n} = 2^{2n}$$

und somit

$$\sum_{k=1}^{n} \binom{2n}{n+k} = \frac{1}{2}\cdot 2^{2n} - \frac{1}{2}\cdot\binom{2n}{n}. \tag{14}$$

Für den Subtrahenden in (13) machen wir uns zunutze, dass bei ungeradem ℓ jeder der Binomialkoeffizienten $\binom{\ell}{j}$, $j \in \{0,\dots,\ell\}$, genau zweimal in einer Zeile des Pascalschen Dreiecks. auftritt. Spalten wir die beiden größten (mittleren) Binomialkoeffizienten in dieser Zeile ab, so folgt für $\ell = 2n-1$

$$2\sum_{k=1}^{n}\binom{2n-1}{n+k} + 2\binom{2n-1}{n} = 2^{2n-1}$$

und somit

$$2\sum_{k=1}^{n}\binom{2n-1}{n+k} = 2^{2n-1} - 2\binom{2n-1}{n}. \qquad (15)$$

Mit (13), (14) und (15) ergibt sich also

$$\begin{aligned}\sum_{k=1}^{n} P_{n,k} &= 2\binom{2n-1}{n} - \frac{1}{2}\binom{2n}{n} \\ &= \frac{1}{2}\binom{2n}{n},\end{aligned}$$

wobei der Nachweis des letzten Gleichheitszeichens eine weitere kleine Übungsaufgabe für Ihre Schülerinnen und Schüler sein könnte. Angesichts von (10) haben wir also (8) und damit auch (9) erhalten. Aus (9) und (4) folgt als Abfallprodukt zudem die Darstellung

$$u_{2n} = \frac{1}{n}\binom{2(n-1)}{n-1}, \quad n \geq 1,$$

und damit speziell der schon durch direktes Abzählen gewonnene Wert $u_8 = 5$.

Abb. 4 zeigt ein Stabdiagramm der Verteilung von W. Mit eingezeichnet ist der relativ große Wert $\mathbb{P}(W \geq 22) \approx 0,176$.

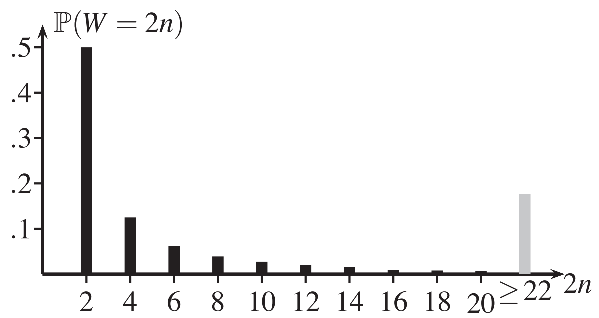

Abb. 4: Stabdiagramm der Verteilung von W

Die Wahrscheinlichkeiten $\mathbb{P}(W = 2n)$ genügen der Rekursionsformel

$$\mathbb{P}(W = 2(n+1)) = \frac{2n-1}{2n+2} \cdot \mathbb{P}(W = 2n), \quad n \geq 1,$$

mit der Anfangsbedingung $\mathbb{P}(W = 2) = \frac{1}{2}$. Auch der Nachweis dieser Beziehung könnte eine Übungsaufgabe für Ihre Schülerinnen und Schüler sein.

3 Zum Erwartungswert von W

Wir kommen jetzt zur eingangs gestellten Frage, wie lange es im Mittel dauert, bis sich Treffer und Nieten erstmals ausgleichen und damit zum Erwartungswert von W.

Nach (1) und (9) gilt

$$\mathbb{E}(W) = \sum_{n=1}^{\infty} \frac{\binom{2(n-1)}{n-1}}{2^{2(n-1)}} = 1 + \sum_{n=1}^{\infty} \frac{\binom{2n}{n}}{2^{2n}}.$$

Eine spannende Frage ist natürlich, ob die rechts stehende Reihe konvergiert oder divergiert. Hier kommt Analysis ins Spiel. Aufgrund der verblüffenden Limesbeziehung

$$\lim_{n\to\infty}\left\{\frac{\binom{2n}{n}}{2^{2n}}\sqrt{n}\right\} = \frac{1}{\sqrt{\pi}} \qquad (16)$$

(Begründung folgt!) erhalten wir mit (8) nicht nur, dass

$$\lim_{n\to\infty} \mathbb{P}(W \geq 2n+2) = 0$$

und somit $\mathbb{P}(W = \infty) = 0$ gilt, sondern auch die Beziehung $\mathbb{E}(W) = \infty$. Nach (16) gilt ja für genügend großes n die Ungleichung

$$\frac{\binom{2n}{n}}{2^{2n}} \geq \frac{1}{2\sqrt{n}\sqrt{\pi}},$$

und da die Reihe $\sum_{n=1}^{\infty} n^{-1/2}$ divergiert, divergiert auch die Reihe $\sum_{n=1}^{\infty}\binom{2n}{n}2^{-2n}$.

Eine Begründung von (16) in einem Leistungskurs könnte so aussehen (siehe hierzu auch das Erklärvideo Henze (2019b)):

Mit $\sin^n x := (\sin x)^n$, $\cos^n x := (\cos x)^n$, $n \geq 0$, sei

$$I_n := \int_0^{\pi/2} \sin^n x \,\mathrm{d}x, \quad n \geq 0.$$

Dann gelten $I_0 = \frac{\pi}{2}$ (nur hierdurch kommt π ins Spiel!) und $I_1 = 1$. Für $k \geq 2$ liefert partielle Integration

$$\begin{aligned} I_k &= \int_0^{\pi/2} \sin^{k-1} x \sin x \,\mathrm{d}x \\ &= \Big[-\sin^{k-1} x \cos x\Big]_0^{\pi/2} + (k-1)\int_0^{\pi/2} \sin^{k-2} x \cos^2 x \mathrm{d}x. \end{aligned}$$

Hier verschwindet der erste Summand, und mit $\cos^2 x = 1 - \sin^2 x$ folgt $I_k = (k-1)I_{k-2} - (k-1)I_k$ und damit

$$I_k = \frac{k-1}{k} \cdot I_{k-2}, \quad k \geq 2.$$

Wegen $I_0 = \frac{\pi}{2}$ und $I_1 = 1$ ergibt sich

$$\begin{aligned} I_{2n} &= \frac{2n-1}{2n} \cdot \frac{2n-3}{2n-2} \cdot \ldots \cdot \frac{3}{4} \cdot \frac{1}{2} \cdot \frac{\pi}{2}, \\ I_{2n+1} &= \frac{2n}{2n+1} \cdot \frac{2n-2}{2n-1} \cdot \ldots \cdot \frac{4}{5} \cdot \frac{2}{3} \cdot 1, \end{aligned}$$

und mit $I_{2n+1} \leq I_{2n} \leq I_{2n-1}$ erhalten wir dann

$$\begin{aligned} \frac{2 \cdot 4 \cdot \ldots \cdot (2n)}{3 \cdot 5 \cdot \ldots \cdot (2n+1)} &\leq \frac{1 \cdot 3 \cdot (2n-1)}{2 \cdot 4 \cdot \ldots \cdot (2n)} \cdot \frac{\pi}{2} \\ &\leq \frac{2 \cdot 4 \cdot \ldots \cdot (2n-2)}{3 \cdot 5 \cdot \ldots \cdot (2n-1)}. \end{aligned}$$

Setzen wir kurz

$$\begin{aligned} a_n &:= \frac{2^2 \cdot 4^2 \cdot \ldots \cdot (2n)^2}{1^2 \cdot 3^2 \cdot \ldots \cdot (2n-1)^2} \cdot \frac{1}{2n+1}, \\ b_n &:= \frac{2^2 \cdot 4^2 \cdot \ldots \cdot (2n)^2}{1^2 \cdot 3^2 \cdot \ldots \cdot (2n-1)^2} \cdot \frac{1}{2n} \end{aligned}$$

sowie $c_n := a_n(2n+1)$, so folgt

$$a_n \leq \frac{\pi}{2} \leq b_n$$

sowie

$$\begin{aligned} 0 &\leq b_n - \frac{\pi}{2} \leq b_n - a_n \\ &= c_n \left(\frac{1}{2n} - \frac{1}{2n+1} \right) = \frac{c_n}{2n(2n+1)} \\ &= \frac{a_n}{2n}. \end{aligned}$$

Wegen $\lim_{n\to\infty}(b_n - a_n) = 0$ und $a_n \leq \pi/2$ folgt $\lim_{n\to\infty} b_n = \pi/2$, was nach Multiplikation mit 2 die berühmte Produktdarstellung

$$\pi = \lim_{n\to\infty} \frac{2^2 \cdot 4^2 \cdot \ldots \cdot (2n)^2}{1^2 \cdot 3^2 \cdot \ldots \cdot (2n-1)^2} \cdot \frac{1}{n}$$

der Kreiszahl π von John Wallis (1616–1703) liefert.

Zieht man hier die Wurzel und geht zum Reziproken über, so ergibt sich

$$\frac{1}{\sqrt{\pi}} = \lim_{n\to\infty} \frac{1 \cdot 3 \cdot \ldots \cdot (2n-1)}{2 \cdot 4 \cdot \ldots \cdot (2n)} \cdot \sqrt{n}.$$

Wenn man jetzt noch den Bruch mit $2^n n!$ erweitert, so folgt (16).

4 Schlussbemerkungen

Stochastik lebt von spannenden konkreten Problemen, die zudem noch mit Überraschungen aufwarten. Hierzu passen die in diesem Aufsatz aufgeworfenen Fragen im Zusammenhang mit Bernoulli-Versuchen. Ein unendlicher Erwartungswert für die Anzahl der Bernoulli-Versuche bis zum ersten Gleichstand zwischen Treffern und Nieten ist völlig kontraintuitiv, zumal die Anzahl der Versuche bis zum erstmaligen Erreichen eines vorgegebenen noch so langen Musters aus Einsen und Nullen einen endlichen Erwartungswert besitzt, siehe z.B. Henze (2020b).

Eine andere nahe liegende Frage ist, wann bei unabhängigen Bernoulli-Versuchen mit gleicher Trefferwahrscheinlichkeit $1/2$ erstmals mehr Treffer als Nieten aufgetreten sind. Bezeichnen wir die dazu nötige Anzahl von Versuchen mit V, so nimmt V für die Daten aus Abb. 1 den Wert 1 an, da der erste Versuch einen Treffer ergeben hat. Schülerinnen und Schüler können auch hier schnell erkennen, dass

$$\mathbb{P}(V=1) = \frac{1}{2},\ \mathbb{P}(V=3) = \frac{1}{8},\ \mathbb{P}(V=5) = \frac{2}{32}$$

gelten, denn für den Fall $V = 3$ gibt es nur die Sequenz 011, und das Ereignis $\{V = 5\}$ tritt für die beiden Sequenzen 01011 und 00111 aus Treffern und Nieten ein.

Nachdem wir $\mathbb{E}(W) = \infty$ gezeigt haben, können wir jetzt ohne jegliche Rechnung das frappierende Resultat einsehen, dass auch V einen unendlichen Erwartungswert besitzt. Warum? Zunächst ist aus Symmetriegründen klar, dass man gedanklich Treffer und Nieten vertauschen kann und dass somit V die gleiche Verteilung wie die Anzahl der Versuche besitzt, bis erstmalig mehr Nieten als Treffer aufgetreten sind.

Die Anzahl W der Versuche bis zum ersten Gleichstand zwischen Treffern und Nieten lässt sich aber additiv zerlegen in 1 (das ist der zu zählende erste Versuch, der uns einen Treffer bzw. eine Niete liefert) und V, denn danach warten wir entweder darauf, dass erstmals mehr Nieten als Treffer auftreten oder umgekehrt, und die dafür nötige Anzahl der Versuche hat – ganz egal, was der erste Versuch ergeben hat – die gleiche Verteilung wie V. Es gilt also die durch eine Tilde beschriebene Verteilungsgleichheit

$$W \sim 1 + V. \tag{17}$$

Hiermit ist zum einen klar, dass $\mathbb{E}(V) = \infty$ gelten muss, denn andernfalls wäre ja aufgrund der Additivität der Erwartungswertbildung auch $\mathbb{E}(W) < \infty$. Zum anderen erhält man auch die Verteilung von V, denn mit (9) und (17) folgt für jedes $n \geq 1$

$$\begin{aligned} \mathbb{P}(V = 2n-1) &= \mathbb{P}(W = 2n) \\ &= \frac{\binom{2(n-1)}{n-1}}{2^{2(n-1)}} \cdot \frac{1}{2n}. \end{aligned}$$

Setzt man hier $n = 1$, $n = 2$ und $n = 3$, so ergeben sich mit $0! = 1$ (wie es sein muss) die oben stehenden Werte für $\mathbb{P}(V = 1), \mathbb{P}(V = 3)$ und $\mathbb{P}(V = 5)$. Die Fragen nach den Verteilungen von W und V hängen unmittelbar mit dem sog. Hauptlemma für symmetrische Irrfahrten zusammen, siehe z.B. Feller (1968), S. 76, Henze (2018), S. 13, oder das Erklärvideo Henze (2020a).

Was die Behandlung der in diesem Aufsatz gestellten Fragen in einem Leistungskurs betrifft, ist die Herleitung der Verteilung von W insofern elementar, als nur Binomialkoeffizienten und das Spiegelungsprinzip eine Rolle spielen. Es reicht aber nicht zu wissen, wie der Binomialkoeffizient $\binom{n}{k}$ mithilfe eines Taschenrechners erhalten werden kann, sondern man muss seine begriffliche Bedeutung als Anzahl der Möglichkeiten, aus n Objekten k auszuwählen, verinnerlicht haben, siehe hierzu z.B. Henze (2019a).

Natürlich ist der Nachweis der Unendlichkeit des Erwartungswertes von W mithilfe der Limesbeziehung (16) an der oberen Grenze dessen, was man mit interessierten Schülerinnen und Schülern in einem Leistungskurs machen kann, aber es muss auch über den Tellerrand der im Zusammenhang mit Bernoulli-Versuchen fast ausschließlich auftretenden Binomialverteilung „Futter für das Fordern“ geben!

Obwohl die eingangs gestellte Frage im Zusammenhang mit Bernoulli-Versuchen fast auf der Hand liegt, dürften die in diesem Aufsatz vorgestellten Ergebnisse auch für viele Lehrkräfte neu sein, aber Hintergrundwissen ist ausdrücklich erwünscht! Wer jetzt neugierig geworden ist und mehr zu Überraschungen bei Bernoulli-Versuchen erfahren möchte, wird z.B. bei Feller (1968), Chapter 3, oder Henze (2018), Kapitel 2, fündig.

Abschließend sei betont, dass das zur Herleitung der Verteilung von W verwendete Spiegelungsprinzip ein schlagkräftiges Hilfsmittel auch für andere stochastische Fragestellungen mit direktem Schulbezug ist (siehe z.B. Henze (2021)).

Danksagung: Ich danke beiden Gutachter(inne)n sowie Philipp Ullmann für wertvolle Hinweise.

Literatur

Feller, W. (1968). An Introduction to Probability Theory and Its Applications (1968). Band 1, 3. Auflage. J. Wiley, New York.

Henze, N. (2018). Irrfahrten – Faszination der random walks. 2. Auflage. Springer Spektrum, Wiesbaden.

Henze, N. (2019a). Binomialkoeffizienten und Pascalsches Dreieck. Erklärvideo. DIVA, KIT. DOI:10.5445/DIVA/2019-978.

Henze, N. (2019b). Die Wallis-Produktdarstellung für die Kreiszahl π. Erklärvideo. DIVA, KIT. DOI:10.5445/DIVA/2019-971.

Henze, N. (2020a). Irrfahrten auf den ganzen Zahlen: Hauptlemma. Erklärvideo. DIVA, KIT. DOI:10.5445/IR/1000118604.

Henze, N. (2020b). Muster in Bernoulli-Versuchen: Erwartungswerte I. Erklärvideo. DIVA, KIT. DOI:10.5445/IR/1000122609.

Henze, N. (2021). Das Stimmzettelproblem. *Stochastik in der Schule* 41(2), S.26–28.

Lesigne, E. (2005). Heads or Tails: An Introduction to Limit Theorems in Probability. Student Mathematical Library Vo. 28. American Mathematical Society. Providence, Rhode Island.

Anschrift des Verfassers:

Prof. i.R. Dr. Norbert Henze
KIT Distinguished Senior Fellow
Institut für Stochastik
Karlsruher Institut für Technologie (KIT)
Englerstr. 2, 76131 Karlsruhe
Henze@kit.edu