

# Palindrome - eine Forschungsreise mit offenem Ausgang

NORBERT HENZE, KARLSRUHE, UND REIMUND VEHLING, HANNOVER

**Zusammenfassung:** In dieser Arbeit geht es um Palindrome, also Wörter, die vorwärts und rückwärts gelesen das Gleiche ergeben. Was passiert, wenn Meister Zufall die Buchstaben erzeugt? Selbst im einfachsten „Bernoulli-Fall“, dass der Buchstabenvorrat nur aus den Zeichen 1 und 0 besteht, die unabhängig voneinander mit den Wahrscheinlichkeiten  $p$  bzw.  $1 - p$  gewählt werden, führen Fragen zu Palindromen zu einer aufregenden und völlig rezeptfreien Forschungsreise.

## 1 Einleitung

Wir betrachten unabhängige Bernoulli-Versuche mit gleicher Trefferwahrscheinlichkeit  $p$ , setzen  $q := 1 - p$  und schreiben 1 bzw. 0 für das Auftreten eines Treffers bzw. einer Niete. Die Ergebnisse von  $n$  hintereinander ausgeführten Bernoulli-Versuchen sind  $n$ -Tupel der Gestalt  $(a_1, a_2, \dots, a_n)$ , wobei  $a_j$  für das Ergebnis des  $j$ -ten Versuchs steht. Lässt man die Klammern und Kommata weg, so entsteht ein Wort  $a_1 a_2 \dots a_n$ . Im Fall  $n \geq 2$  heißt ein solches  $n$ -Tupel bzw. Wort *Palindrom*, falls anschaulich gesprochen das Gleiche herauskommt, wenn man es vorwärts oder rückwärts liest, falls also für jedes  $j \in \{1, \dots, n\}$  die Gleichung  $a_j = a_{n+1-j}$  erfüllt ist. Palindrome wie etwa 0110 oder 10101 sind somit durch eine Symmetrieeigenschaft gekennzeichnet. In diesem Aufsatz geht es um folgende spannende Fragen, die auch ganz ohne das Eintippen von Rechnerbefehlen zu einer aufregenden Forschungsreise führen können:

- Mit welcher Wahrscheinlichkeit entsteht nach  $n$  Bernoulli-Versuchen ein Palindrom?
- Wie wahrscheinlich ist es, dass man  $n$  Bernoulli-Versuche durchführen muss, damit *erstmal*s ein Palindrom entsteht?
- Erhält man mit Wahrscheinlichkeit eins irgendwann ein Palindrom, wenn man gedanklich beliebig viele Bernoulli-Versuche durchführen kann?

Natürlich ist das Auftreten von Palindromen nicht auf Bernoulli-Versuche beschränkt. Wirft man etwa wiederholt einen Würfel, so entsteht nach zwei Würfeln ein Palindrom, wenn wie 22 oder 55 ein Pasch geworfen wird, und 353 ist ein Palindrom, das sich nach drei Würfeln ergibt. Wir werden in Abschnitt 5 hierauf eingehen.

## 2 Einstiege mit Simulationen

Mit welcher Wahrscheinlichkeit entsteht nach  $2; 3; 4; 5; \dots; n$  Bernoulli-Versuchen ein Palindrom? Wird diese Frage als Einstiegsimpuls gewählt, sind vielfältige (natürliche) Differenzierungen möglich. So können Schülerinnen und Schüler zunächst spezielle Werte für  $n$  und  $p$  betrachten oder gleich versuchen, eine Antwort für allgemeines  $n$  und  $p$  zu finden. Es ist aber auch möglich, sich nach ersten Schätzungen einer Lösung durch Simulation zu nähern. Palindrome können nämlich erstaunlich einfach mithilfe digitaler Werkzeuge untersucht werden, wenn Listen- oder Vektoroperationen verfügbar sind. So bewirkt etwa der Befehl `rand(4)` beim grafikfähigen Taschenrechner TI-84+, dass vier (Pseudo)-Zufallszahlen zwischen 0 und 1 erzeugt und in einer Liste ausgegeben werden. Um eine Bernoulli-Folge der Länge  $n = 4$  mit Trefferwahrscheinlichkeit  $p = 0,8$  zu erhalten, genügt dann die Eingabe des Befehls `rand(4) <= 0.8`. Dadurch wird jedes Element  $\ell$  der Liste mit dem Wert 0,8 verglichen und eine 1 bzw. eine 0 ausgegeben, falls die Aussage  $\ell \leq 0,8$  wahr bzw. falsch ist. So bedeutet z. B. die Ausgabe 1110, dass in den ersten drei Versuchen jeweils ein Treffer und im vierten eine Niete erzielt wurde. Auch mithilfe von GeoGebra-CAS können Listen erzeugt und mit einem Befehl auf Gleichheit untersucht werden, siehe Abb. 1.

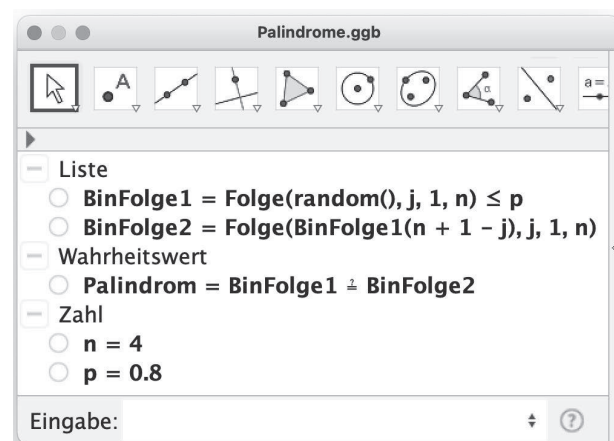


Abb. 1: Simulation mit GeoGebra-CAS

In einem Kurs lassen sich durch Gruppenexperimente für verschiedene Werte von  $n$  und  $p$  Schätzungen für die gesuchten Wahrscheinlichkeiten erhalten. Zudem hinterlässt die Erfahrung, dass man in bestimmten Fällen (z. B.  $0,2 \leq p \leq 0,8$  und  $n \geq 20$ ) sehr vie-

le Bernoulli-Versuche bis zum erstmaligen Auftreten eines Palindroms durchführen muss, einen bleibenden Eindruck.

Auch mithilfe der Statistik-Software R (siehe Core Team (2015)) lassen sich Schätzwerte für die gesuchten Wahrscheinlichkeiten ermitteln, wobei die Indikatorfunktion eine überzeugende Anwendung findet (siehe Abb. 2 für eine mögliche Umsetzung).

```
13 m<-5; n<-100000; p<-0.8; wert<-1:n
14 for (j in 1:n)
15 {
16   a<-sample(c(0,1),m,TRUE,c(1-p,p))
17   wert[j]<-sum(a[1:m]==a[m:1])==m
18 }
19 sum(wert)/n
20 print((2*p^2-2*p+1)^(round(m/2,0)))
```

Abb. 2: Simulation mit R

Mit dem *sample*-Befehl wird 100 000-mal eine (0,1)-Bernoulli-Folge der Länge  $m = 5$  mit  $p = 0,8$  und  $q = 0,2$  erzeugt und im Vektor  $a$  gespeichert. Die Eingabe  $a[m : 1]$  bewirkt eine Vertauschung der Reihenfolge der Elemente. Der Vergleich  $a[1 : m] == a[m : 1]$  liefert einen booleschen Vektor der Länge  $m$  mit den Einträgen TRUE bzw. FALSE. Liegt ein Palindrom vor, so besteht dieser Vektor nur aus Elementen, die den Wert TRUE haben. Da letzterer als 1 interpretiert wird, kann mithilfe des Summenbefehls *sum* und eines Vergleichs ein Palindrom identifiziert werden. Das Programm berechnet zusätzlich die (theoretische) Wahrscheinlichkeit. Unterschiede zur relativen Häufigkeit sind im Fall  $n = 100\,000$  nur in der dritten Nachkommastelle feststellbar.

Selbstverständlich muss die Lehrkraft entscheiden, ob der Einsatz von Simulationen aus didaktischer Sicht zielführend ist. Hierzu sollte sie mehrere digitale Werkzeuge und deren Vor- und Nachteile kennen, um eine fundierte Auswahlmöglichkeit zu haben. Natürlich können auch Realexperimente mit Münzen ( $p = 0,5$ ) oder Würfeln ( $p = \frac{1}{6}$ ) eingesetzt werden. Im nächsten Abschnitt wird die erste der in der Einleitung aufgeführten Fragen mithilfe theoretischer Betrachtungen beantwortet.

### 3 Palindrome nach $n$ Versuchen

Wie wahrscheinlich ist es, dass zwei unabhängige Bernoulli-Versuche das gleiche Ergebnis liefern, also ein Palindrom der Länge zwei entsteht? Da dafür entweder zwei Treffer oder zwei Nieten auftreten

müssen, ist diese Wahrscheinlichkeit gleich

$$\sigma := p^2 + q^2. \quad (1)$$

Damit sich ein Palindrom der Länge drei ergibt, ist der zweite Bernoulli-Versuch belanglos; es müssen nur der erste und der dritte Versuch den gleichen Ausgang haben. Da die Wahrscheinlichkeit dafür ebenfalls gleich  $\sigma$  ist, sind die Wahrscheinlichkeiten für das Auftreten eines Palindroms nach zwei bzw. nach drei Versuchen beide gleich  $\sigma$ .

Nach vier Versuchen entsteht ein Palindrom genau dann, wenn der erste und der vierte Versuch gleich ausgehen (die Wahrscheinlichkeit hierfür ist gleich  $\sigma$ ), und wenn der zweite und der dritte Versuch das gleiche Ergebnis zeigen. Da die Wahrscheinlichkeit dafür ebenfalls gleich  $\sigma$  ist, ergibt sich wegen der stochastischen Unabhängigkeit von Ereignissen, die sich auf verschiedene Bernoulli-Versuche beziehen, dass die Wahrscheinlichkeit für ein Palindrom nach 4 Versuchen gleich  $\sigma^2$  ist. Der gleiche Wert ergibt sich für die Wahrscheinlichkeit eines Palindroms nach 5 Versuchen, da dafür die dritte „Spiegelstelle“ irrelevant ist.

Mit diesen Einsichten ist die Wahrscheinlichkeit für ein Palindrom nach  $2k$  oder  $2k + 1$  Bernoulli-Versuchen identisch, nämlich gleich  $\sigma^k$ , denn es müssen für jedes  $j \in \{1, \dots, k\}$  der  $j$ -te und der  $j$ -letzte Versuch den gleichen Ausgang haben. Bezeichnet  $A_n$  das Ereignis, dass nach  $n$  Versuchen ein Palindrom entsteht, so können wir mithilfe der durch  $\lfloor x \rfloor := \max\{k \in \mathbb{Z} : k \leq x\}$  für jede reelle Zahl  $x$  definierten *Abrundungsfunktion* oder *unteren Gauß-Klammer* die Fälle, dass  $n$  gerade oder ungerade ist, in der Formel

$$\mathbb{P}(A_n) = (p^2 + q^2)^{\lfloor n/2 \rfloor}, \quad n \geq 2, \quad (2)$$

zusammenfassen.

Im Unterricht wird dieser elegante Lösungsweg sicherlich am Ende des Lösungsprozesses stehen. Zuerst wird man sich Schritt für Schritt an (2) herantasten, also für einige Werte von  $n$  die zum Ereignis  $A_n$  gehörende Ergebnismenge  $A_n$  sowie die Wahrscheinlichkeit  $\mathbb{P}(A_n)$  ermitteln, z. B.  $A_4 = \{1111, 0110, 1001, 0000\}$  und  $\mathbb{P}(A_4) = p^4 + 2p^2q^2 + q^4 = (p^2 + q^2)^2$ . Danach ergibt sich ganz natürlich die Frage, ob es noch einen anderen Lösungsweg gibt, der vielleicht auch das WARUM klärt. Solche (seltenen) Anlässe, in denen das Argumentieren im Vordergrund steht, sollten immer wieder aufgegriffen werden.

Abb. 3 zeigt die Wahrscheinlichkeiten  $\mathbb{P}(A_n)$  als Funktion von  $p$  für  $n \leq 7$ . Da  $\mathbb{P}(A_n) = \sigma^{\lfloor n/2 \rfloor}$  eine Potenz von  $\sigma$  und  $\sigma$  nach (1) symmetrisch in  $p$  und  $q$  ist, sind die drei Graphen in Abb. 3 achsensymmetrisch zur Geraden mit der Gleichung  $p = 0,5$ . Zudem wird  $\mathbb{P}(A_n)$  für  $p = \frac{1}{2}$  minimal, denn es gilt wegen  $q = 1 - p$

$$\sigma = 2 \left( p - \frac{1}{2} \right)^2 + \frac{1}{2},$$

und somit wird auch jede Potenz von  $\sigma$  für  $p = \frac{1}{2}$  minimal.

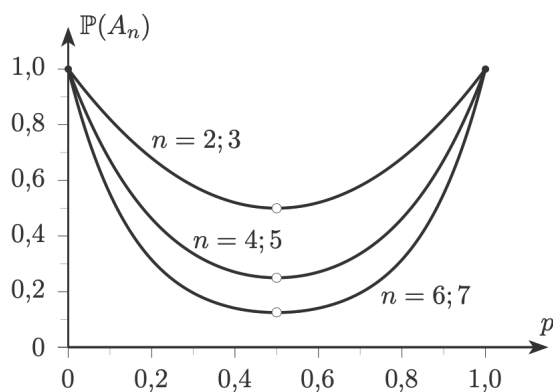


Abb. 3:  $\mathbb{P}(A_n)$  als Funktion von  $p$  für  $n \leq 7$

Die Achsensymmetrie der obigen Graphen kann rein begrifflich eingesehen werden, indem man von jedem Palindrom  $(a_1, \dots, a_n)$  der Länge  $n$  zu dessen „geflippter“ Version  $(1 - a_1, \dots, 1 - a_n)$  übergeht, bei der Einsen durch Nullen ersetzt werden und umgekehrt. Hat das Ausgangspalindrom  $\ell$  Einsen, so ist seine Wahrscheinlichkeit gleich  $p^\ell q^{n-\ell}$ , und die des geflippten Palindroms ist gleich  $q^\ell p^{n-\ell}$ . Die Summe beider Wahrscheinlichkeiten ändert sich nicht, wenn  $p$  und  $q$  vertauscht werden. Zusätzlich zur Minimum- und Symmetrieeigenschaft von  $\mathbb{P}(A_n)$  sollte auch klar sein, dass die Werte  $p = 1$  und  $p = 0$  zu  $\mathbb{P}(A_n) = 1$  führen.

#### 4 Warten auf das erste Palindrom

Im Folgenden geht es um die mit  $W$  bezeichnete zufällige Anzahl der Bernoulli-Versuche, bis zum ersten Mal ein Palindrom auftritt. Hier gibt es eine faustdicke Überraschung. Abb. 4 zeigt das Ergebnis einer auf 10000 Wiederholungen fußenden Simulation für den Erwartungswert von  $W$  in Abhängigkeit von  $p$ . Dabei wurden die Werte  $p = 0$ ;  $p = 0,025$ ;  $p = 0,05$ ; ...;  $p = 0,975$  und  $p = 1$  gewählt.

Zeigen Sie doch einfach einmal Ihren Schülerinnen und Schülern diese erstaunliche Grafik. Sie werden daraufhin bestimmt vermuten, dass der Erwartungswert von  $W$  unabhängig von  $p$  konstant ist und viel-

leicht immer den Wert 3 besitzt, aber kurioserweise scheint für die Werte  $p = 0$  und  $p = 1$  der Wert 2 herauszukommen.

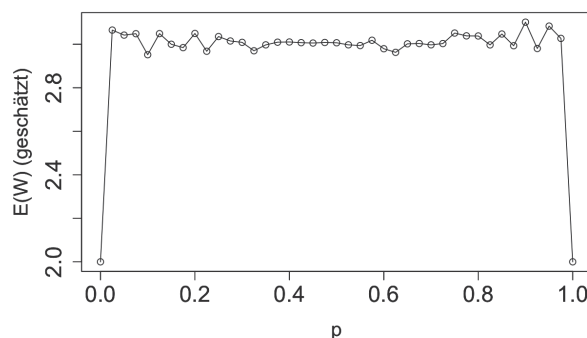


Abb. 4: Simulationen für  $\mathbb{E}(W)$ ; 10000 Wiederholungen

Da sich selbst für die sehr nahe bei eins liegenden Erfolgswahrscheinlichkeit  $p = 0,9999$  (für die die Anzahl der Wiederholungen auf  $10^6$  erhöht wurde) der auf drei Nachkommastellen gerundete Schätzwert 2,92 für  $\mathbb{E}(W)$  ergab, liegt vielleicht ein hochinteressantes „Unstetigkeitsphänomen“ für einen Erwartungswert vor. Dieses Phänomen „schreit“ geradezu danach, theoretisch beleuchtet zu werden, und das soll jetzt geschehen.

Gilt  $p = 1$  oder  $p = 0$ , so ergibt sich  $\mathbb{P}(W = 2) = 1$ , denn die Bernoulli-Folge beginnt mit Wahrscheinlichkeit eins mit 11 bzw. mit 00. Insbesondere folgt  $\mathbb{E}(W) = 2$ . In allen anderen Fällen, insbesondere also auch für  $p = 0,9999$ , gilt jedoch  $\mathbb{E}(W) = 3$ , und zwar völlig unabhängig von  $p$ . Warum?

Um diesen frappierenden Sachverhalt einzusehen, muss man nur die beiden Fälle unterscheiden, dass der erste Bernoulli-Versuch einen Treffer bzw. eine Niete ergibt. Ist der erste Versuch ein Treffer, so tritt erstmalig ein Palindrom auf, wenn sich der zweite Treffer einstellt, und beginnen die Bernoulli-Versuche mit einer Niete, so wartet man für das Entstehen eines Palindroms auf die zweite Niete. Da der Erwartungswert der Anzahl der Bernoulli-Versuche bis zum ersten Treffer bzw. bis zur ersten Niete (einschließlich des Versuchs, der den Treffer bzw. die Niete ergibt), gleich  $\frac{1}{p}$  bzw. gleich  $\frac{1}{q}$  ist, und da man den ersten Versuch mitzählen muss, gilt die schon in Engel (1976), S. 26, aufgeführte Gleichung

$$\mathbb{E}(W) = p \left( 1 + \frac{1}{p} \right) + q \left( 1 + \frac{1}{q} \right) = 3. \quad (3)$$

Man kann diese Gleichung auch mithilfe der über die Formel von der totalen Wahrscheinlichkeit (zweiten

Pfadregel) erhältlichen und durch

$$\begin{aligned}\mathbb{P}(W = k) &= p\mathbb{P}(W = k|1) + q\mathbb{P}(W = k|0) \\ &= p^2 q^{k-2} + q^2 p^{k-2}, \quad k \geq 2,\end{aligned}$$

gegebenen Verteilung von  $W$  herleiten, denn es gilt

$$\mathbb{E}(W) = \sum_{k=2}^{\infty} k\mathbb{P}(W = k). \quad (4)$$

Diese Darstellung von  $\mathbb{E}(W)$  führt zu den beiden unendlichen Reihen

$$\begin{aligned}2p^2 + 3p^2 q + 4p^2 q^2 \cdots &= p^2 \sum_{k=0}^{\infty} (k+2)q^k, \\ 2q^2 + 3q^2 p + 4q^2 p^2 \cdots &= q^2 \sum_{k=0}^{\infty} (k+2)p^k\end{aligned}$$

und damit auf die geometrische Reihe und deren Ableitung. Leider gehört die geometrische Reihe in der Schule momentan nicht zum Pflichtkanon. An dieser Stelle muss dann ein CAS-Programm benutzt werden, was unseres Erachtens legitim ist. Schaffen es die Schülerinnen und Schüler bis zur Darstellung in (4), haben sie schon viel Mathematik betrieben. Die Ermittlung eines Terms für die Summe kann dann auch ein CAS-Programm übernehmen. Abb. 5 zeigt in Zeile 2 die Berechnung von  $\mathbb{E}(W)$  mithilfe von GeoGebra-CAS. Hierbei wurde (nur) 1:1 die Definition des Erwartungswertes mit dem ermittelten Term für  $\mathbb{P}(W = k)$  in die Zeile 2 eingegeben. Sogar die Berechnung der zugehörigen Varianz  $\mathbb{V}(W)$  mit

$$\mathbb{V}(W) = \sum_{k=2}^{\infty} (k - \mathbb{E}(W))^2 \cdot \mathbb{P}(W = k) \quad (5)$$

ist auf Knopfdruck möglich.

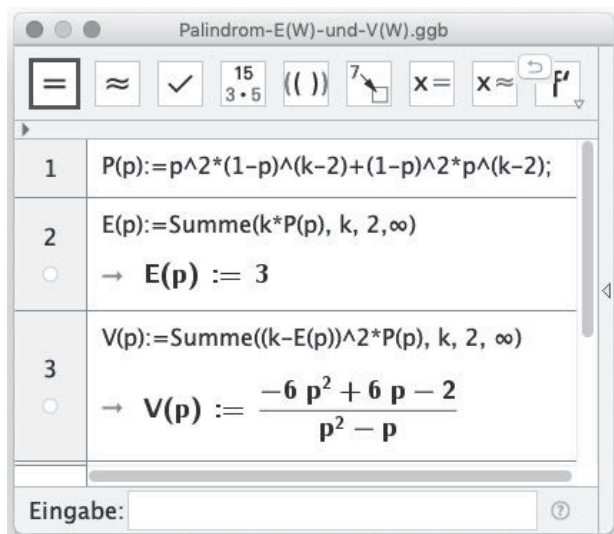


Abb. 5:  $\mathbb{E}(W)$  und  $\mathbb{V}(W)$ : Berechnung mit GeoGebra-CAS

Die Umsetzung von (5) steht in Zeile 3 von Abb. 5. Da sofort der zugehörige Graph gezeichnet werden kann, ergibt sich als Vermutung  $\mathbb{V}(W) \geq 2$ . Diese Vermutung kann auf Schulniveau auch begründet werden. Mit  $q := 1 - p$  ergibt sich

$$\mathbb{V}(W) = \frac{2}{pq} - 6. \quad (6)$$

Wegen  $pq \leq 0,25$  gilt also  $\mathbb{V}(W) \geq 2$ , und das Minimum wird für  $p = 0,5$  angenommen. Dass die Varianz von  $W$  bei Annäherung von  $p$  an einen der Werte 0 oder 1 über alle Grenzen wächst, wird in Abb. 4 in entsprechenden stärkeren Schwankungen der Schätzwerte für  $\mathbb{E}(W)$  sichtbar.

Die Ermittlung der Varianz (s. (6)) ohne CAS-Einsatz ist auf Schulniveau sehr anspruchsvoll und höchstens etwas für besonders interessierte Schülerinnen und Schüler - die es aber immer noch gibt!

$$\begin{aligned}\mathbb{V}(W) &= \sum_{k=2}^{\infty} (k - \mathbb{E}(W))^2 \cdot \mathbb{P}(W = k) \\ &= \sum_{k=2}^{\infty} k^2 \cdot \mathbb{P}(W = k) - 2(\mathbb{E}(W))^2 + (\mathbb{E}(W))^2 \\ &= \sum_{k=2}^{\infty} k^2 \cdot \mathbb{P}(W = k) - 9 \\ &= \frac{3p^2 - 3p - 2}{p^2 - p} - 9 \\ &= \frac{2}{pq} - 6.\end{aligned}$$

Die vorletzte Zeile ergibt sich entweder mit einem CAS-Programm oder mithilfe der Formel

$$\sum_{j=2}^{\infty} j(j-1)x^{j-2} = \frac{2}{(1-x)^3}, \quad |x| < 1,$$

für die zweite Ableitung der geometrischen Reihe und ein wenig Termumformung.

Abschließend sei gesagt, dass die Zufallsvariable  $W$  eng mit der in Ebner und Henze (2016) untersuchten und mit  $Y$  bezeichneten Länge des zweiten Runs in einer Bernoulli-Folge verknüpft ist. Unter der Bedingung  $W \geq 3$ , dass die beiden ersten Bernoulli-Versuche verschiedene Ausgänge haben, gilt  $W = Y + 2$ .

## 5 Palindrome bei Versuchen mit mehr als zwei Ausgängen

Palindrome sind bei Wörtern mit Buchstaben aus jedem Alphabet möglich. So ist etwa das Wort OTTO ein Palindrom der Länge vier mit Buchstaben

aus dem deutschen Alphabet. Anstelle von Buchstaben wählen wir die Zahlen  $1, 2, \dots, s$ , wobei  $s \geq 2$  gelte. Damit ist der bislang betrachtete Fall  $s = 2$  eingeschlossen (wobei wir jetzt nur 1 anstelle 0 und 2 anstelle von 1 setzen). Damit der Zufall ins Spiel kommt, diene ein Versuch mit  $s$  möglichen Ausgängen, die mit 1 bis  $s$  durchnummeriert seien. Dabei möge der Ausgang  $j$  die Wahrscheinlichkeit  $p_j$  besitzen ( $j = 1, \dots, s$ ,  $p_1 + \dots + p_s = 1$ ). Ein Beispiel hierfür ist der Würfelwurf, für den  $s = 6$  gilt und der Ausgang  $j$  für die Augenzahl  $j$  steht. Bei einem fairen Würfel gilt zudem  $p_1 = \dots = p_6 = \frac{1}{6}$ .

Wird der Versuch  $n$ -mal in unabhängiger Folge durchgeführt, so sind die in zeitlicher Reihenfolge notierten Ergebnisse  $n$ -Tupel  $(a_1, \dots, a_n)$ . Dabei steht  $a_j$  wie früher für das Ergebnis des  $j$ -ten Versuchs. Ein *Palindrom* ist wie bisher für  $n \geq 2$  durch die Bedingung  $a_j = a_{n+1-j}$  für jedes  $j \in \{1, \dots, n\}$  charakterisiert, und wir schreiben unverändert  $\mathbb{P}(A_n)$  für die Wahrscheinlichkeit, dass nach  $n$  Versuchen ein Palindrom entsteht.

Diese Wahrscheinlichkeit fällt fast wie Schuppen von den Augen, wenn man sich klarmacht, mit welcher Wahrscheinlichkeit irgendwelche zwei (unabhängig voneinander durchgeführte) Versuche den gleichen Ausgang haben. Da sich zweimal unabhängig voneinander eines der Ergebnisse  $1, 2, \dots, s$  einstellen muss, ist diese Wahrscheinlichkeit gleich

$$w := p_1^2 + \dots + p_s^2. \quad (7)$$

Ist  $n = 2k$  mit  $k \in \mathbb{N}$  eine gerade Zahl, so entsteht ein Palindrom genau dann, wenn  $k$ -mal unabhängig voneinander zwei Versuchsergebnisse gleich sind, nämlich die des  $j$ -ten und des  $(2k+1-j)$ -ten Versuchs für  $j = 1, \dots, k$ . Es folgt somit  $\mathbb{P}(A_{2k}) = w^k$ , und da mit den gleichen Überlegungen wie früher im Fall  $n = 2k+1$  mit  $k \in \mathbb{N}$  der  $(k+1)$ -te Versuch für die Entstehung eines Palindroms keine Rolle spielt, gilt  $\mathbb{P}(A_{2k+1}) = w^k$ . Mit der Abrundungsfunktion  $\lfloor x \rfloor$  können beide Fälle in der Form

$$\mathbb{P}(A_n) = (p_1^2 + \dots + p_s^2)^{\lfloor n/2 \rfloor}, \quad n \geq 2,$$

zusammengefasst werden.

Wir wollen uns noch überlegen, dass diese Wahrscheinlichkeit für den Gleichverteilungsfall  $p_1 = \dots = p_s = \frac{1}{s}$  minimal wird. Diese Einsicht lässt sich auf Schulniveau gewinnen, indem die Änderungen vom Gleichverteilungsfall betrachtet werden. Mit  $p_i = \frac{1}{s} + \varepsilon_i$ ,  $i \in \{1, \dots, s\}$ , und  $\varepsilon_1 + \dots + \varepsilon_s = 0$

folgt:

$$\begin{aligned} p_1^2 + \dots + p_s^2 &= \left(\frac{1}{s} + \varepsilon_1\right)^2 + \dots + \left(\frac{1}{s} + \varepsilon_s\right)^2 \\ &= \frac{1}{s} + 2 \cdot \frac{1}{s} \cdot \sum_{i=1}^s \varepsilon_i + \sum_{i=1}^s \varepsilon_i^2 \\ &= \frac{1}{s} + \sum_{i=1}^s \varepsilon_i^2 \\ &\geq \frac{1}{s}. \end{aligned}$$

Diese untere Schranke wird (nur) im Gleichverteilungsfall angenommen, und somit wird auch  $\mathbb{P}(A_n)$  als Potenz von  $w (= p_1^2 + \dots + p_s^2)$  minimal.

## 6 Ein offenes Problem

Als höherer Gesichtspunkt für Lehrkräfte sei noch ein offenes mathematisches Problem vorgestellt. Dieses entsteht, wenn man nach der *Verteilung* der Anzahl  $X$  der Palindrome in einer gedanklich unendlich langen Folge unabhängiger Bernoulli-Versuche mit gleicher Trefferwahrscheinlichkeit  $p$ ,  $0 < p < 1$ , fragt. Dabei wird ein Palindrom stets vom ersten Bernoulli-Versuch an betrachtet, d.h., es handelt sich um die Zufallsgröße

$$X = \mathbf{1}\{A_2\} + \mathbf{1}\{A_3\} + \dots$$

Hier gibt allgemein die Indikatorfunktion  $\mathbf{1}\{A\}$  eines Ereignisses  $A$  an, ob  $A$  eintritt ( $\mathbf{1}\{A\} = 1$ ) oder nicht ( $\mathbf{1}\{A\} = 0$ ). Abb. 6 zeigt Schätzwerte für  $\mathbb{P}(X = k)$  im Fall  $p = 0,8$  aufgrund einer Simulation.

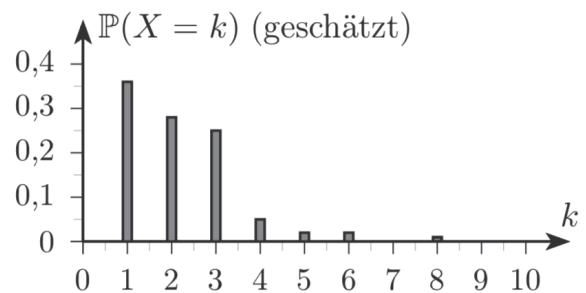


Abb. 6: Schätzwerte für  $\mathbb{P}(X = k)$  für  $p = 0,8$  aufgrund von 400 Wiederholungen aus einer Bernoulli-Folge der Länge 10000

Anscheinend gilt  $\mathbb{P}(X = 0) = 0$ , was sich auch schnell einsehen lässt. Es gibt nämlich nur zwei unendlich lange Bernoulli-Folgen ohne Palindrom, und zwar diejenige, die mit einer Eins beginnt, auf die unendlich viele Nullen folgen, und diejenige, die mit einer Null startet und dann mit unendlich vielen Einsen weitergeht. Jede dieser beiden Folgen hat die Wahrscheinlichkeit null, denn das Ereignis, dass auf eine Eins  $n$  Nullen folgen, hat die Wahrscheinlichkeit

$pq^n$ . Flippt man Einsen und Nullen, so hat das betreffende Ereignis die Wahrscheinlichkeit  $qp^n$ , und beide Wahrscheinlichkeiten konvergieren für  $n \rightarrow \infty$  gegen null.

Obwohl die Verteilung von  $X$  unbekannt ist, ist der Erwartungswert von  $X$  unmittelbar erhältlich, denn es gilt mit (2) und  $\mathbb{E}(\mathbf{1}\{A\}) = \mathbb{P}(A)$

$$\begin{aligned}\mathbb{E}(X) &= \sum_{n=2}^{\infty} \mathbb{P}(A_n) = \sum_{n=2}^{\infty} (p^2 + q^2)^{\lfloor n/2 \rfloor} \\ &= 2 \cdot \sum_{k=1}^{\infty} (p^2 + q^2)^k \\ &= 2 \cdot \left( \frac{1}{1 - (p^2 + q^2)} - 1 \right) \\ &= \frac{p}{q} + \frac{q}{p}.\end{aligned}$$

Dabei gilt das erste Gleichheitszeichen aufgrund des sog. Satzes von der monotonen Konvergenz (siehe z. B. Henze (2019), S. 335).

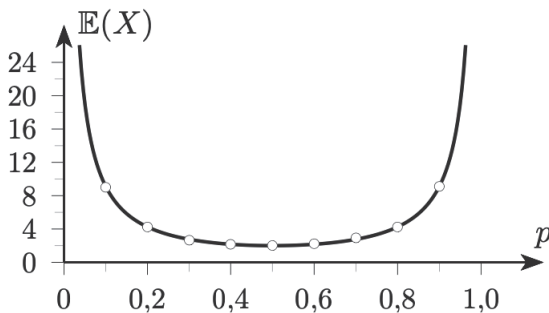


Abb. 7: Graph von  $\mathbb{E}(X)$  und Schätzwerte

Abb. 7 zeigt den Graphen des Erwartungswertes von  $X$  als Funktion von  $p$  sowie Simulationsergebnisse für  $0,1 \leq p \leq 0,9$  in 0,1-Schritten, dargestellt als Punkte. Hierfür wurden Bernoulli-Folgen der Länge  $n = 10000$  mit jeweils 500 Wiederholungen für die Schätzwerte betrachtet. Dass  $\mathbb{E}(X)$  für  $p = 0,5$  minimal wird und den kleinsten Wert 2 liefert, ist eine schöne Anwendung im Umfeld von Optimierungsaufgaben. Eine einfache Begründung besteht darin,  $x := \frac{p}{q}$  zu setzen und die durch  $f(x) = x + \frac{1}{x}$  für  $0 < x < 1$  definierte Funktion  $f$  zu betrachten. Es ergibt sich die Extremstelle  $x = 1$  und somit  $p = q$ . Mit der zusätzlichen Bedingung  $p + q = 1$  folgt  $p = q = 0,5$  und mit  $f(1) = 2$  das Minimum 2. Alternativ kann man auch damit argumentieren, dass das arithmetische Mittel mindestens gleich dem geometrischen Mittel ist. Damit gilt

$$\frac{1}{2} \left( x + \frac{1}{x} \right) \geq \sqrt{x \cdot \frac{1}{x}} = 1.$$

Dadurch, dass der Erwartungswert von  $X$  endlich ist, folgt, dass  $\mathbb{P}(X < \infty) = 1$  gilt. Man erhält also

mit Wahrscheinlichkeit eins irgendwann ein Palindrom, wenn man beliebig viele Bernoulli-Versuche durchführen kann. Somit erfährt die in der Einleitung gestellte dritte Frage eine positive Antwort.

## 7 Fazit

Diese Arbeit ist ein weiteres Plädoyer dafür, bei Bernoulli-Versuchen nicht nur die Anzahl der erzielten Treffer zu betrachten. So gewinnt man bei Fragen zu Palindromen in einem fruchtbaren Wechselspiel zwischen Theorie und Simulation vielfältige Einsichten über den Zufall. Wir würden uns freuen, wenn diese Arbeit zu eigenen Überlegungen führt, die noch nicht in dem Artikel beleuchtet sind. Lassen Sie doch auch einmal ihre Schülerinnen und Schüler weitere Problemstellungen finden. Der zweite Autor hat damit immer gute Erfahrungen sammeln können. Die Motivation ist dadurch streng monoton gewachsen, Problemstellungen zu bearbeiten.

**Danksagung:** Wir danken Joachim Engel und den beiden Gutachtern bzw. Gutachterinnen für wertvolle Hinweise.

## Literatur

- Core Team, R. (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>.
- Ebner, B., und Henze, N. (2016). Runs in Bernoulli-Ketten. *Stochastik in der Schule* 36 (3), S. 21–27.
- Engel, A. (1976). Wahrscheinlichkeitsrechnung und Statistik. Band 2. Ernst Klett Verlag, Stuttgart.
- Henze, N. (2019). Stochastik: Eine Einführung mit Grundzügen der Maßtheorie. Springer Spektrum, Heidelberg.

Anschriften der Verfasser:

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](mailto:Henze@kit.edu)

Reimund Vehling  
Studienseminar für das Lehramt an Gymnasien  
Hannover I  
[vehling@icloud.com](mailto:vehling@icloud.com)