Das Stimmzettelproblem

NORBERT HENZE, KARLSRUHE

Zusammenfassung: *Bei einer Klassensprecherwahl wurden 14 Stimmen für Anja und 11 für Bernd abgegeben, sodass Anja die Wahl gewonnen hat. Es stellte sich heraus, dass Anja während der gesamten Stimmauszählung führte. Wie wahrscheinlich ist das bei diesem Endergebnis? Die Antwort auf diese Frage, nämlich 0.12, führt uns auf eine spannende rezeptfreie Reise in das Reich der Binomialkoeffizienten, und ein wenig Geometrie ist auch dabei.*

1 Problem und mögliche Lösung

Das oben beschriebene klassische Problem (siehe z.B. André (1887)) ist in der englischsprachigen Literatur als *ballot problem* wohlbekannt (Feller (1968), S. 69; Renault (2007)). Abstrahieren wir von den in der Zusammenfassung genannten Personen und Anzahlen, so hat zwischen zwei Kandidaten A und B eine geheime Wahl stattgefunden, bei der A gewonnen hat, und zwar mit a Stimmen gegenüber b Stimmen von B. Wie groß ist die Wahrscheinlichkeit des mit C bezeichneten Ereignisses, dass Kandidat A während der gesamten Stimmauszählung führte?

Eine Lösung dieses Problems auf Kursstufenniveau könnte wie folgt aussehen: Wir beschreiben die Auszählungsverläufe als Wege in einem rechtwinkligen Koordinatensystem, indem wir jede Stimme für A bzw. für B als Aufwärts- bzw. Abwärtsschritt notieren. Jeder Auszählungsverlauf entspricht dann genau einem Weg, der vom Punkt $(0,0)$ zum Punkt $(a+b, a-b)$ verläuft, siehe Abb. 1.

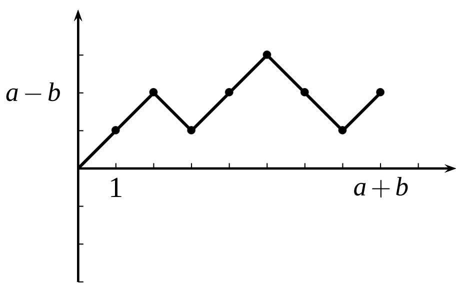

Abb. 1: Auszählungsverlauf als Weg (hier führt A immer)

Bei dem Weg in Abb. 1 führt Kandidat A während der gesamten Stimmauszählung, die wir auch in der Form (A,A,B,A,A,B,B,A) hätten notieren können. Bei dem in Abb. 2 dargestellten Auszählungsverlauf führt jedoch A *nicht* immer, da die erste ausgezählte Stimme an den Kandidaten B geht.

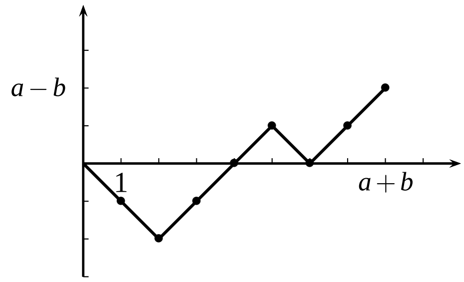

Abb. 2: Auszählungsverlauf, bei dem A nicht immer führt

Wir nehmen an, dass alle Auszählungsverläufe gleich wahrscheinlich sind, arbeiten also mit einem Laplace-Modell auf der Menge aller Auszählungsverläufe. Wie viele Auszählungsverläufe, also Wege von $(0,0)$ nach $(a+b, a-b)$, gibt es insgesamt? Setzen wir kurz $n := a+b$, so ist ein Weg eindeutig durch ein n-Tupel $(c_1, \ldots, c_n)$ beschrieben, wobei genau a der c_j's gleich A und b der C_j's gleich B sind. So ist dieses Tupel für den Abb. 2 gezeigten Weg gleich (B,B,A,A,A,B,A,A). Wir müssen also a der n Plätze im Tupel auswählen und mit A's besetzen. Hierfür gibt es $\binom{n}{a}$ Möglichkeiten; die anderen Plätze im Tupel sind dann mit den b B's besetzt.

Es bleibt also nur übrig, die Anzahl der für das Eintreten des Ereignisses C günstigen Wege zu bestimmen. Hier hilft es, das Gegenereignis $D := \overline{C}$, also „A führt nicht immer“, zu betrachten. Die entscheidenden Idee besteht darin, dieses Gegenereignis durch Fallunterscheidung in zwei sich ausschließende Ereignisse zu zerlegen. Der erste Fall ist der, dass der erste Stimmzettel auf B lautet (Ereignis D_1), und der zweite, dass der erste Stimmzettel auf A lautet und A *nicht* während der gesamten Stimmauszählung führt (Ereignis D_2). Es gilt dann

$$\mathbb{P}(D) = \mathbb{P}(D_1) + \mathbb{P}(D_2). \qquad (1)$$

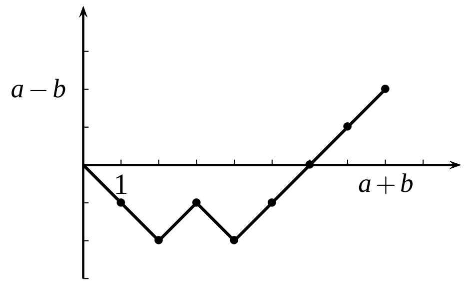

Abb. 3: Weg, für den das Ereignis D_1 eintritt

Abb. 3 zeigt einen Weg, für den das Ereignis D_1 eintritt. Wie viele für das Eintreten des Ereignisses D_1 günstige Wege gibt es? Da ein solcher Weg am Anfang einen Abwärtsschritt macht, ist der erste Platz im n-tupel aus a A's und b B's, welches den Weg festlegt, mit einem B belegt. Es gibt also für das Ereignis D_1 so viele günstige Wege, wie es Möglichkeiten gibt, von den insgesamt $n-1$ Plätzen Nr. 2 bis Nr. n im n-Tupel a Plätze auszuwählen und diese jeweils mit A zu besetzen (die verbleibenden Plätze erhalten natürlich jeweils ein B). Diese Anzahl ist der Binomialkoeffizient $\binom{n-1}{a}$. Es gilt also

$$\mathbb{P}(D_1) = \frac{\binom{n-1}{a}}{\binom{n}{a}} = \frac{n-a}{n}. \tag{2}$$

Wie sieht es mit dem Ereignis D_2 aus? Abb. 4 zeigt einen Weg, für den dieses Ereignis eintritt.

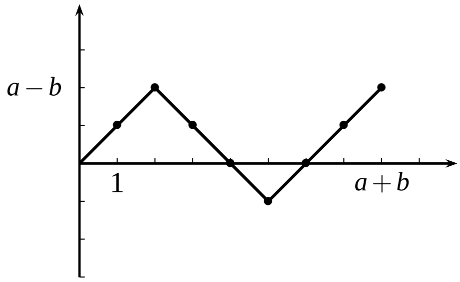

Abb. 4: Weg, für den das Ereignis D_2 eintritt

Jeder Weg, für den das Ereignis D_2 eintritt, macht zuerst einen Aufwärtsschritt, aber – wenn wir die horizontale Achse als Zeitachse deuten – irgendwann zeitlich danach trifft dieser Weg *mindestens einmal* diese Achse. Konsequenterweise gibt es dann im zeitlichen Verlauf einen *ersten Zeitpunkt*, zu dem die horizontale Achse getroffen wird. Eine geniale, als *Spiegelungsprinzip* bekannte und gemeinhin Désiré André (1840–1918) zugeschriebene Idee besteht darin, den Anfangsteil des Weges bis zum ersten Schnittpunkt mit der Zeitachse an dieser Achse zu spiegeln, wobei der sich daran anschließende Teilweg nicht verändert wird. Für den Weg in Abb. 4 entsteht auf diese Weise der in Abb. 5 gezeigte Weg.

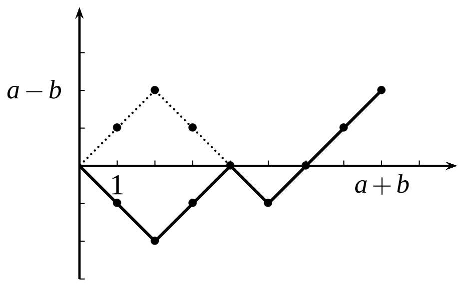

Abb. 5: Aus dem Weg aus Abb. 4 durch Spiegelung hervorgehender Weg

Das Schöne ist, dass durch diese Spiegelungsvorschrift ein Weg entstanden ist, für den das Ereignis D_1 eintritt, und wenn wir mit einem anderen Weg aus D_2 gestartet wären, hätte sich auch ein anderer Weg aus D_2 ergeben (die Spiegelungsvorschrift ist eine injektive Abbildung!). Diese Abbildung ist aber auch surjektiv und damit insgesamt bijektiv, also eineindeutig, denn wenn wir mit einem beliebigen Weg aus D_1 starten, geht ja der erste Schritt nach unten, und da der Weg am Ende in der positiven Höhe $a-b$ endet, muss er im zeitlichen Verlauf mindestens einmal (und damit auch ein erstes Mal!) die horizontale Achse treffen. Wir machen die oben eingeführte Spiegelungsabbildung jetzt einfach rückgängig, indem wir den von $(0,0)$ bis zum erstmaligen Treffen der Zeitachse verlaufenden Teilweg an dieser Achse spiegeln und den sich anschließenden Teilweg unverändert lassen. Auf diese Weise erhält man einen Weg aus D_2, und speziell vom Weg in Abb. 5 ausgehend denjenigen von Abb. 4.

Folglich sind die Anzahlen der Wege aus D_1 und D_2 gleich, woraus $\mathbb{P}(D_1) = \mathbb{P}(D_2)$ und damit wegen (1) und (2)

$$\mathbb{P}(D) = 2 \cdot \frac{n-a}{n}$$

folgt. Wegen $\mathbb{P}(C) = 1 - \mathbb{P}(D)$ und $n = a+b$ ergibt sich dann die Antwort auf die eingangs gestellte Frage zu

$$\mathbb{P}(C) = 1 - 2 \cdot \frac{b}{a+b} = \frac{a-b}{a+b}.$$

Die gesuchte Warscheinlichkeit ist also gleich der Steigung der durch die Punkte $(0,0)$ und $(a+b, a-b)$ verlaufenden Geraden. Für das in der Zusammenfassung auftretende Zahlenbeispiel erhält man den dort angegebenen Wert

$$\mathbb{P}(C) = \frac{14-11}{14+11} = \frac{3}{25} = \frac{12}{100} = 0.12.$$

2 Varianten der Fragestellung

Das Faszinierende an der Mathematik ist ja, dass es „hinterm Horizont immer weiter geht" und sich mit jedem gelösten Problem neue Fragen auftun. Wie groß ist denn die Wahrscheinlichkeit, dass während des gesamten Auszählungsprozesses Kandidat A immer *mindestens so viele Stimmen* wie Kandidat B erhalten hat? (Man könnte natürlich auch danach fragen, mit welcher Wahrscheinlichkeit A gegenüber B ab dem zweiten erfassten Stimmzettel immer mit mindestens zwei Stimmen Vorsprung führte.) Bezeichnen wir dieses Ereignis mit E, so empfiehlt sich aufgrund der bisherigen Überlegungen, auch hier zum Gegenereignis $F := \overline{E}$ überzugehen, das (nur) auf zwei verschiedene Weisen eintreten kann. Entweder der erste Stimmzettel lautet auf B (Ereignis F_1), oder der erste Stimmzettel lautet auf A, und A hat während des gesamten Auszählungsprozesses *nicht* immer mindestens so viele Stimmen wie B. Da die Ereignisse F_1 und D_1 identisch sind, gilt nach (2)

$$\mathbb{P}(F_1) = \frac{n-a}{n}.$$

Das Ereignis F_2 tritt genau dann ein, wenn der den Auszählungsprozess beschreibende Weg nach dem ersten Aufwärtsschritt in seinem weiteren Verlauf mindestens einmal die Höhe -1 trifft (Abb. 6).

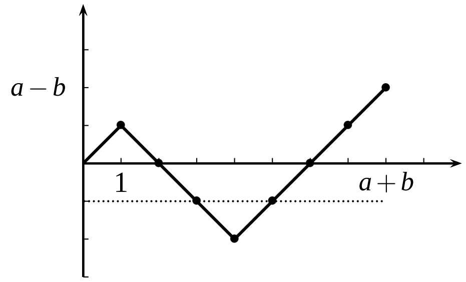

Abb. 6: Weg, für den das Ereignis F_2 eintritt

Spiegelt man das Anfangstück des Weges bis zum *erstmaligen* Erreichen der in Abb. 6 gestrichelt eingezeichneten Höhe -1 und lässt das sich anschließende Teilstück unverändert, so entsteht der in Abb. 7 dargestellte, vom Punkt $(0,-2)$ zum Punkt $(a+b, a-b)$ verlaufende *und einen ersten Abwärtsschritt beinhaltende* Weg, und diese Spiegelungsvorschrift ist offenbar eine injektive Abbildung.

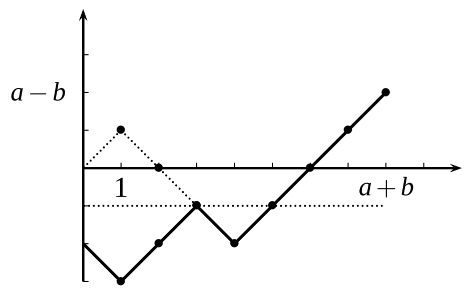

Abb. 7: Der Weg aus Abb. 6 nach Spiegelung

Da der von $(1,-3)$ nach $(a+b, a-b)$ verlaufende Weg $a+1$ Aufwärts- und $b-2$ Abwärtsschritte durchläuft, gibt es $\binom{n-1}{a+1}$ Wege von $(1,-3)$ nach $(a+b, a-b)$. Jeder solche Weg trifft in seinem zeitlichen Verlauf irgendwann *erstmalig* die Höhe -1, sodass Spiegelung des Teilweges vor diesem Treffpunkt (inklusive des ersten Abwärtsschrittes) unter Beibehaltung des sich anschließenden Weges die Bijektivität der Spiegelungsabbildung zeigt. Es folgt

$$\mathbb{P}(F_2) = \frac{\binom{n-1}{a+1}}{\binom{n}{a}} = \frac{b(b-1)}{(a+b)(a+1)}$$

und damit wegen $\mathbb{P}(E) = 1 - \mathbb{P}(F_1) - \mathbb{P}(F_2)$

$$\mathbb{P}(E) = \frac{a-b+1}{a+1}.$$

Wie man leicht nachrechnet, gilt (wie es sein muss!) $\mathbb{P}(E) > \mathbb{P}(C)$, und für das Zahlenbeispiel der Klassensprecherwahl ergibt sich speziell $\mathbb{P}(E) = 4/15 \approx 0.267$.

3 Schlussbemerkungen

Stochastik lebt von spannenden konkreten Fragen. Eine davon ist das Stimmzettelproblem. Für den Kursstufenunterricht ist dieses Problem insofern von Interesse, weil es sowohl einen Modellierungsaspekt als auch das weitgehend aus den Schulbüchern verschwundene Begründen (ehemals: Beweisen) beinhaltet. Besonders schön ist, dass durch das Spiegelungsprinzip die Eineindeutigkeit einer Abbildung eingesehen wird, die letztlich zu einer eleganten Lösung des Problems führt. Man sieht auch, wie wichtig es ist, den Binomialkoeffizienten $\binom{n}{k}$ als Anzahl der Möglichkeiten, aus n Dingen k auszuwählen, verinnerlicht zu haben. In dieser Hinsicht könnte Henze (2019) für interessierte Schülerinnen und Schüler eine „Offenbarung“ sein, weil u.a. der allgemeine binomische Lehrsatz nichts weiter als ein Abfallprodukt einer begrifflichen Definition des Binomialkoeffizienten ist.

Danksagung: Ich danke den beiden Gutachtern für wertvolle Hinweise.

Literatur

André, D. (1887). Solution directe du problème résolu par M. Bertrand. *Compt. Rend. Acad. Sci. paris* 105, 436–437.

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

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

Renault, M. (2007). Four proofs of the ballot theorem. *Mathem. Magazine* 80(5), 345–352.

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`