Runs in Bernoulli-Ketten

BRUNO EBNER UND NORBERT HENZE, KARLSRUHE

Zusammenfassung: *Es seien X und Y die Längen des ersten bzw. zweiten Runs in einer Bernoulli-Kette mit Trefferwahrscheinlichkeit p. Dabei ist ein Run als Sequenz maximaler Länge gleicher aufeinander folgender Symbole (1 oder 0) definiert. Wir geben die Verteilungen von X und Y sowie die gemeinsame Verteilung von X und Y an und berichten von typischen Schwierigkeiten, die Lehramtsstudierende insbesondere bei der Herleitung der Verteilung von Y haben. Interessanterweise gilt $\mathbb{E}(Y)=2$, unabhängig von der Trefferwahrscheinlichkeit p. Die gemeinsame Verteilung der Längen des ersten, zweiten, dritten, … Runs besitzt eine Symmetrieeigenschaft, die zu weiteren Einsichten führt. So ist etwa die Länge jedes Runs mit ungerader (bzw. gerader) Nummer verteilt wie X (bzw. Y).*

1 Einleitung

Bernoulli-Ketten bilden einen Kern des gymnasialen Stochastikunterrichts. Sie modellieren die Situation unabhängiger, gleichartiger Versuche mit den beiden möglichen Ausgängen Treffer und Niete. In einer Bernoulli-Kette mit Trefferwahrscheinlichkeit p, $0 < p < 1$, bezeichnen X und Y die Längen des ersten bzw. zweiten Runs. Dabei ist ein Run als Sequenz maximaler Länge gleicher aufeinander folgender Symbole 1 (für Treffer) und 0 (für Niete) definiert. Beginnt die Bernoulli-Kette etwa mit 001110…, so nehmen X und Y die Werte 2 bzw. 3 an. In Feller (1970), S. 238, findet sich die Übungsaufgabe, die Verteilungen von X und Y sowie die zugehörigen Erwartungswerte und Varianzen und auch die gemeinsame Verteilung von X und Y zu bestimmen.

In einer Übungsaufgabe für Studierende des Lehramts Mathematik an Gymnasien im Rahmen einer Vorlesung *Einführung in die Stochastik* sollte (mit der Abkürzung $q = 1-p$) Folgendes gezeigt werden:

a) $\mathbb{P}(X=k) \;=\; p^k q + q^k p, \quad k \geq 1,$

b) $\mathbb{E}(X) \;=\; \dfrac{1}{pq} - 2,$

c) $\mathbb{P}(Y=k) \;=\; p^2 q^{k-1} + q^2 p^{k-1},\ k \geq 1,$

d) $\mathbb{E}(Y) \;=\; 2,$

e) $\mathbb{P}(X=k, Y=\ell) \;=\; p^{k+1} q^\ell + q^{k+1} p^\ell,\ k,\ell \geq 1,$

f) X, Y stochastisch unabhängig $\Longleftrightarrow p = \dfrac{1}{2}.$

Wir werden die Lösungen zu allen Teilaufgaben vorstellen und höhere Gesichtspunkte ansprechen.

2 Die Verteilung der Länge des ersten Runs

Aufgabenteil a) bereitete den Studierenden keinerlei Schwierigkeiten. Das Ereignis $\{X = k\}$ tritt genau dann ein, wenn die Bernoulli-Kette entweder mit k Treffern startet und dann eine Niete folgt oder mit k Nieten beginnt, die durch einen Treffer abgelöst werden. Wegen der stochastische Unabhängigkeit der Ergebnisse der einzelnen Versuche besitzen diese Möglichkeiten die Wahrscheinlichkeiten $p^k q$ bzw. $q^k p$, woraus unmittelbar a) folgt. Abbildung 1 zeigt ein Stabdiagramm der Verteilung von X im Fall $p = 1/4$.

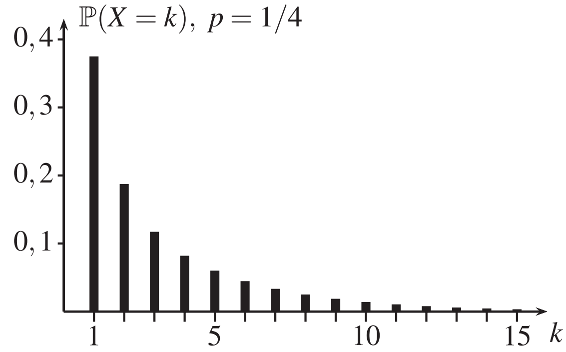

Abb. 1: Stabdiagramm der Verteilung von X, $p = 1/4$

Intuitiv ist zu erwarten, dass im Vergleich zum Fall $p = 1/2$ der erste Run „im Mittel länger wird", wenn man p vergrößert oder verkleinert. Diese Vermutung wird durch das in Abb. 2 gezeigte Stabdiagramm der Verteilung von X im Fall $p = 1/6$ erhärtet.

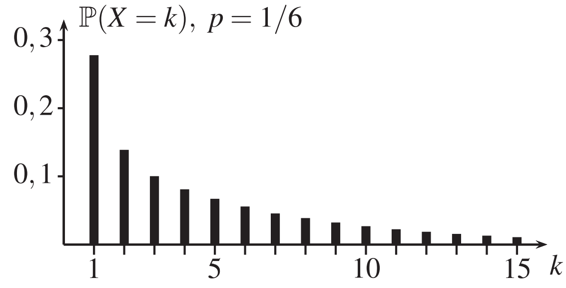

Abb. 2: Stabdiagramm der Verteilung von X, $p = 1/6$

Stochastik in der Schule 36 (2016) 3, S. 20–26

Es ist deutlich zu erkennen, dass der als physikalischer Schwerpunkt der durch das Stabdiagramm gegebenen Masseverteilung interpretierbare Erwartungswert von X im Fall $p = 1/6$ im Vergleich zu $p = 1/4$ größer geworden ist.

Diesen Erwartungswert von X kann man z.B. mithilfe der geometrischen Reihe

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

und deren Ableitung

$$\sum_{j=1}^{\infty} j x^{j-1} = \frac{\mathrm{d}}{\mathrm{d}x}\left(\sum_{j=0}^{\infty} x^j\right) = \frac{1}{(1-x)^2}, \quad |x| < 1,$$

erhalten: Es gilt

$$\begin{aligned}
\mathbb{E}(X) &= \sum_{k=1}^{\infty} k\,\mathbb{P}(X=k) \\
&= pq\sum_{k=1}^{\infty} k\,p^{k-1} + pq\sum_{k=1}^{\infty} k\,q^{k-1} \\
&= \frac{pq}{p^2} + \frac{pq}{q^2} = \frac{q}{p} + \frac{p}{q} \\
&= \frac{1}{pq} - 2.
\end{aligned}$$

Wegen $pq \leq 1/2$ nimmt dieser Erwartungswert in der Tat sein Minimum ($= 2$) für $p = 1/2$ an, was die zuvor geäußerte Vermutung bestätigt.

Obiges Ergebnis lässt sich auch auf anderem Wege erhalten, wenn man weiß, dass der Erwartungswert der Anzahl aller Versuche bis zum ersten Treffer (einschließlich des Versuchs, der diesen Treffer ergibt) in einer Bernoulli-Kette mit Trefferwahrscheinlichkeit τ gleich $1/\tau$ ist. Führt man nämlich die Ereignisse A_1 bzw. A_0 ein, dass der erste Versuch einen Treffer bzw. eine Niete ergibt, so ist der bedingte Erwartungswert von X unter der Bedingung A_1 gleich $1/q$, denn man wartet nach dem ersten Treffer auf die erste Niete. In gleicher Weise ist der bedingte Erwartungswert von X unter der Bedingung A_0 gleich $1/p$, und die Formel vom totalen Erwartungswert (s. z.B. Henze (2013), S. 207 oder Humenberger (2000)) liefert

$$\begin{aligned}
\mathbb{E}(X) &= \mathbb{E}[X|A_1]\cdot p + \mathbb{E}[X|A_0]\cdot q \\
&= \frac{p}{q} + \frac{q}{p} = \frac{1}{pq} - 2.
\end{aligned}$$

Mit der Fallunterscheidung danach, ob die Bernoulli-Kette mit einer 1 oder einer 0 beginnt, ergibt sich sogar, dass für jedes k die Wahrscheinlichkeit $\mathbb{P}(X \geq k)$, dass der erste Run die Mindestlänge k besitzt, für $p = 1/2$ ein Minimum annimmt (man spricht dann davon, dass die Verteilung von X im Fall $p = 1/2$ *stochastisch minimal wird*). Zum Beweis beachte man, dass nach der Formel von der totalen Wahrscheinlichkeit

$$\mathbb{P}(X \geq k) = p\,\mathbb{P}(X \geq k|A_1) + q\,\mathbb{P}(X \geq k|A_0)$$

gilt. Wegen $\mathbb{P}(X \geq k|A_1) = p^{k-1}$ und $\mathbb{P}(X \geq k|A_0) = q^{k-1}$ (die Bedingungen besagen, dass bereits ein Treffer bzw. eine Niete aufgetreten ist!) folgt dann

$$\mathbb{P}(X \geq k) = p^k + q^k, \quad k \geq 1.$$

Kürzen wir die rechte Seite mit $f_k(p)$ ab, so gelten offenbar $f_1(p) = 1$, $0 < p < 1$, sowie

$$f_2(p) = p^2 + q^2 = 1 - 2pq \geq \frac{1}{2}$$

mit Gleichheit (nur) für $p = 1/2$. Im Fall $k \geq 3$ bilden wir die Ableitung

$$f_k'(p) = kp^{k-1} - k(1-p)^{k-1}$$

und erhalten $f_k'(p) = 0 \Longleftrightarrow p = 1/2$. Da $f_k(p)$ für $p \to 0$ und $p \to 1$ gegen 1 strebt und die Ableitung $f_k'(p)$ beim Übergang von $p < 1/2$ zu $p > 1/2$ das Vorzeichen von Minus nach Plus wechselt, besitzt die Funktion f_k für $k \geq 2$ auf dem Intervall $(0,1)$ in der Tat ein eindeutiges Minimum an der Stelle $p = 1/2$ (die Positivität der zweiten Ableitung $f_k''(p) = k(k-1)(p^{k-2} + q^{k-2})$ auf $(0,1)$ zeigt, dass f_k dort sogar strikt konvex ist).

An dieser Stelle sei auch die Varianz von X angegeben. Unter Verwendung der zweiten Ableitung

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

einer geometrischen Reihe erhält man zunächst

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

und damit wegen

$$y := \frac{p}{q} + \frac{q}{p} = \frac{(p+q)^2 - 2pq}{pq} = \frac{1}{pq} - 2$$

$$\begin{aligned}
\mathbb{V}(X) &= \mathbb{E}X(X-1) + \mathbb{E}(X) - (\mathbb{E}X)^2 \\
&= 2\left(\frac{p}{q}\right)^2 + 2\left(\frac{q}{p}\right)^2 + y - y^2 \\
&= 2y^2 - 4 + y - y^2 \\
&= \left(\frac{1}{pq}\right)^2 - \frac{3}{pq} - 2.
\end{aligned}$$

Wie der Erwartungswert hängt also auch die Varianz von X nur vom Produkt pq ab, und wie $\mathbb{E}(X)$ wird auch $\mathbb{V}(X)$ im Fall $p = 1/2$ minimal. Letztere Behauptung ergibt sich durch eine einfache Diskussion der Funktion $g(x) = 1/x^2 - 3/x - 2$ auf dem Intervall $(0, 1/4]$ (man beachte die Ungleichung $x := pq \leq 1/4$): Es gilt $g(x) \to \infty$ für $x \to 0$ sowie $g(1/4) = 2$. Ferner gilt

$$g'(x) = -\frac{2}{x^3} + \frac{3}{x^2} = \frac{1}{x^2}\left(3 - \frac{2}{x}\right)$$

und damit $g'(x) < 0$ auf dem Intervall $(0, 1/4]$. Dies zeigt, dass g an der Stelle $x = 1/4$ – was $p = 1/2$ entspricht – ein Minimum besitzt.

Die Fallunterscheidung danach, ob der erste Versuch ein Treffer oder eine Niete ist, mündet in einer eleganten Darstellung für X, wenn man die Indikatorvariable $U := \mathbf{1}\{A_1\}$ einführt, dass die Bernoulli-Kette mit 1 beginnt. Es gilt dann $\mathbb{P}(U = 1) = p = 1 - \mathbb{P}(U = 0)$. Beschreiben die Zufallsvariablen N und T die Anzahl der Versuche bis zum Auftreten der ersten Niete bzw. des ersten Treffers in einer Bernoulli-Kette mit Trefferwahrscheinlichkeit p, so besitzt die Länge X des ersten Runs die Darstellung

$$X \;=\; U \cdot N + (1-U) \cdot T, \tag{2}$$

wobei U, N und T stochastisch unabhängig sind.

Darstellung (2) besagt, dass die Verteilung von X eine *Mischung* der Verteilungen von N und T mit den *Mischungswahrscheinlichkeiten* p bzw. q ist. Sie liefert zugleich eine bequeme Möglichkeit, den Erwartungswert von X zu berechnen: Wegen der Unabhängigkeit von U, N und T gilt

$$\begin{aligned} \mathbb{E}(X) &= \mathbb{E}U \cdot \mathbb{E}N + \mathbb{E}(1-U) \cdot \mathbb{E}T \\ &= p \cdot \frac{1}{q} + q \cdot \frac{1}{p}, \end{aligned}$$

was schon auf anderem Wege gezeigt wurde.

Da U nur die Werte 0 und 1 annimmt, gelten $U^2 = U$, $(1-U)^2 = 1-U$ und $U(1-U) = 0$. Somit folgt aus (2)

$$X^2 \;=\; U \cdot N^2 + (1-U) \cdot T^2,$$

was wegen $\mathbb{V}(X) = \mathbb{E}(X^2) - (\mathbb{E}X)^2$ eine weitere Möglichkeit zur Berechnung der Varianz von X eröffnet.

3 Die Verteilung der Länge des zweiten Runs

Teilaufgabe c), also die Herleitung der Gleichung

$$\mathbb{P}(Y = k) = p^2(1-p)^{k-1} + (1-p)^2 p^{k-1},\; k \geq 1,$$

bereitete den Lehramtsstudierenden vergleichsweise große Schwierigkeiten. Im Speziellen ging es um das begriffliche Verständnis der Exponenten $k-1$ (die Studierenden hatten hier den Exponenten k erwartet). Was bei der Bestimmung der Verteilung der Länge Y des zweiten Runs nicht berücksichtigt wurde, ist die Tatsache, dass für den mit Wahrscheinlichkeit p auftretenden Fall, dass der erste Versuch einen Treffer ergibt, der zweite Run *automatisch mit einer Niete beginnt* und dann genau $k-1$ Nieten und danach ein Treffer folgen müssen, damit sich das Ereignis $\{Y = k\}$ einstellt. Für den komplementären, mit Wahrscheinlichkeit q eintretenden Fall, dass die Bernoulli-Kette mit einer Niete startet, beginnt der zweite Run automatisch mit einem Treffer, und es müssen danach genau $k-1$ Treffer und dann eine Niete folgen, damit Y den Wert k annimmt.

Die falsche Überlegung der Studierenden war hier eher: Die Wahrscheinlichkeit setzt sich zusammen aus p für den *letzten Treffer im ersten Run*, dann $(1-p)^k$ für die k Nieten des zweiten Runs und wieder ein p, um den zweiten Run abzuschließen. Die Abbildungen 3 und 4 zeigen Stabdiagramme der Verteilung von Y für die Fälle $p = 1/4$ bzw. $p = 1/6$.

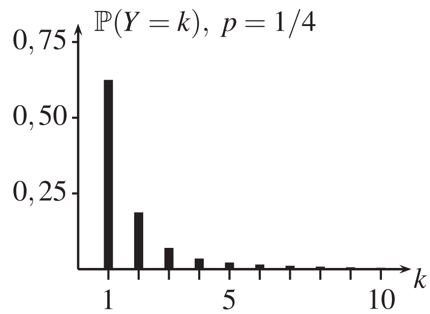

Abb. 3: Stabdiagramm der Verteilung von Y, $p = 1/4$

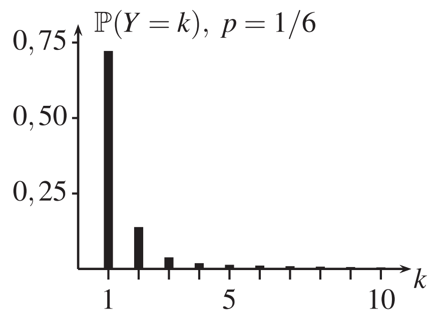

Abb. 4: Stabdiagramm der Verteilung von Y, $p = 1/6$

Vergleicht man Abb. 3 mit Abb. 1 sowie Abb. 4 mit Abb. 2 (man beachte den veränderten Maßstab!), so springt zunächst ins Auge, dass die Verteilung von Y viel stärker bei kleineren Werten konzentriert ist als die Verteilung der Länge X des ersten Runs. Insbesondere ist die Wahrscheinlichkeit $\mathbb{P}(Y=1)$, dass der zweite Run die kleinstmögliche Länge 1 besitzt, deutlich größer als $\mathbb{P}(X=1)$. In der Tat gilt nach a) und c)

$$\begin{aligned}\mathbb{P}(X=1) &= 2pq \le \frac{1}{2},\\ \mathbb{P}(Y=1) &= p^2+q^2 = 1-2pq \ge \frac{1}{2},\end{aligned}$$

wobei die jeweiligen Schranken nur für $p=1/2$ angenommen werden.

Eine interessante Beobachtung ergibt sich bei der Betrachtung der Wahrscheinlichkeit

$$\mathbb{P}(Y\ge k) = pq^{k-1}+qp^{k-1},$$

dass der zweite Run die *Mindestlänge k besitzt*. Diese Wahrscheinlichkeit ergibt sich am einfachsten aus der Überlegung, dass für den mit Wahrscheinlichkeit p bzw. q eintretenden Fall, dass die Bernoulli-Kette mit einer 1 bzw. 0 beginnt, der zweite Run automatisch mit einer 0 bzw. 1 anfängt und dann zum Erreichen einer Mindestlänge k noch $k-1$ Nullen bzw. $k-1$ Einsen folgen müssen. Alternativ kann man unter Verwendung der geometrischen Reihe die Wahrscheinlichkeiten $\mathbb{P}(Y=j)$ über $j\ge k$ summieren. Speziell gelten also

$$\begin{aligned}\mathbb{P}(Y\ge 2) &= 2pq,\\ \mathbb{P}(Y\ge 3) &= pq^2+qp^2 = pq,\\ \mathbb{P}(Y\ge 4) &= pq^3+qp^3 = pq(1-2pq).\end{aligned}$$

In jedem dieser Fälle wird $\mathbb{P}(Y\ge k)$ im Fall $p=1/2$ maximal, und die Wahrscheinlichkeit, einen zweiten Run der Mindestlänge 2, 3 oder 4 zu beobachten, wächst im Bereich $0<p\le 1/2$ streng monoton. Betrachtet man jedoch die Wahrscheinlichkeit

$$\begin{aligned}\mathbb{P}(Y\ge 5) &= pq^4+qp^4 = pq(q^3+p^3)\\ &= pq\left((p+q)^3-3pq^2-3q^2p\right)\\ &= pq(1-3pq),\end{aligned}$$

so hängt – weil die Funktion $h(x)=x(1-3x)$ im Bereich $0<x\le 1/4$ für $x=1/6$ ihr (striktes) Maximum annimmt – die Wahrscheinlichkeit, dass sich ein zweiter Run der Mindestlänge 5 einstellt, *nicht monoton von p ab*. Löst man die quadratische Gleichung $p(1-p)=1/6$ nach p auf, so ergibt sich p zu

$$p = \frac{1}{2}\pm\frac{1}{2\sqrt{3}}.$$

Die Wahrscheinlichkeit $\mathbb{P}(Y\ge 5)$ wird also für die Werte $p=0.211324\cdots$ und $p=0.788675\cdots$ maximal. Abb. 5 zeigt den Graphen der Wahrscheinlichkeit $\mathbb{P}(Y\ge k)$ als Funktion von p für $k=4$ (gepunktet), $k=5$ (durchgezogen) und $k=6$ (gestrichelt). Dabei wurde die Abhängigkeit von p durch Indizierung mit p hervorgehoben.

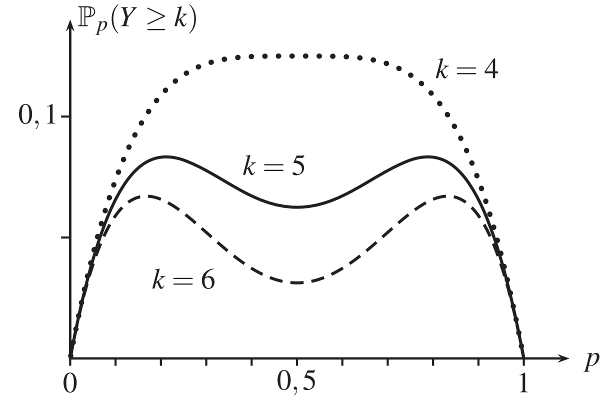

Abb. 5: $\mathbb{P}_p(Y\ge k)$ als Funktion von p

Man sieht, dass für $k=4$ (gerade noch) ein Maximalwert für $p=1/2$ vorliegt. Für $k=5$ und $k=6$ nimmt die Wahrscheinlichkeit $\mathbb{P}_p(Y\ge k)$ jedoch für $p=1/2$ ein lokales *Minimum* an. Wer selbst ein wenig rechnen möchte, wird feststellen, dass die Wahrscheinlichkeit

$$\mathbb{P}_p(Y\ge 6) = pq-4(pq)^2+2(pq)^3$$

an den Stellen $p=0,5\pm 0,332234\cdots$ jeweils ein lokales Maximum annimmt. Diese ergeben sich, indem man zunächst die im Intervall $(0,1/4]$ liegende Nullstelle der Funktion $f(x)=x-4x^2+2x^3$ zu

$$x_0 = \frac{4-\sqrt{10}}{6} = 0,1396203\cdots$$

bestimmt und dann die quadratische Gleichung $p(1-p)=x_0$ nach p auflöst, was zu den Lösungen

$$p = \frac{1}{2}\pm\sqrt{\frac{1}{4}-x_0}$$

führt.

Der Erwartungswert von Y (Teilaufgabe d)) ergibt sich aus der Verteilung zu

$$\begin{aligned}\mathbb{E}(Y) &= \sum_{k=1}^{\infty} k\,\mathbb{P}(Y=k)\\ &= p^2\sum_{k=1}^{\infty} k\,q^{k-1}+q^2\sum_{k=1}^{\infty} k\,p^{k-1}\\ &= p^2\cdot\frac{1}{p^2}+q^2\cdot\frac{1}{q^2} = 2.\end{aligned}$$

Dieses Ergebnis ist auf den ersten Blick überraschend, denn warum sollte der zweite Run – völlig unabhängig von der Erfolgswahrscheinlichkeit p – im Mittel immer die Länge 2 besitzen? Die folgende, mithilfe der Formel vom totalen Erwartungswert erfolgende Herleitung des Erwartungswertes von Y macht das Resultat zumindest ein wenig plausibel: Ist der erste Versuch ein Treffer, so beginnt der zweite Run automatisch mit einer Niete. Danach wartet man (als wenn die Bernoulli-Kette „bei Null starten würde“) auf den ersten Treffer. Zählt man diesen mit, so ist die Verteilung der Länge des zweiten Runs gleich der Anzahl der Nieten vor dem ersten Treffer (einschließlich des Versuchs, der den Treffer ergibt), und der Erwartungswert der Anzahl der Versuche hierfür ist gleich $1/p$. Es gilt also $\mathbb{E}[Y|A_1] = 1/p$. In gleicher Weise gilt $\mathbb{E}[Y|A_0] = 1/q$, und die Formel vom totalen Erwartungswert ergibt

$$\mathbb{E}(Y) = \mathbb{E}[Y|A_1] \cdot p + \mathbb{E}[Y|A_0] \cdot q = 2.$$

Anhand dieser Darstellung erkennt man etwa, dass für ein sehr kleines p das Ereignis A_1 zwar sehr selten (nämlich mit der Wahrscheinlichkeit p) eintritt. Wenn es aber eintritt, besteht der zweite Run aus den mit der großen Wahrscheinlichkeit q auftretenden Nieten, und der bedingte Erwartungswert $\mathbb{E}[Y|A_1]$ ist gleich $1/p$, also sehr groß. Ebenso argumentiert man in Bezug auf das Ereignis A_0.

1	26	3	1	25	1	7	1
6	1	12	1	7	1	9	1
3	1	4	1	3	1	16	1
4	1	2	1	9	1	3	1
21	2	7	1	1	1	11	1
4	1	12	1	9	1	5	2
1	1	5	1	3	1	1	1
2	1	2	1	9	1	15	1
2	1	7	1	6	1	3	1
6	1	2	1	10	1	2	7
2	1	1	1	3	1	17	1
14	1	29	1	28	1	10	1
1	2	46	1	32	1	6	2
8	1	34	1	27	1	1	27
1	3	27	1	1	1	4	1
23	2	8	1	10	1	17	1
2	1	3	2	24	2	3	1
1	1	70	1	2	1	13	1
1	1	5	1	1	9	1	1
2	1	4	1	3	1	30	1

Tab. 1: 80 simulierte Längen des ersten Runs (links) und des zweiten Runs (rechts), $p = 0,1$.

Tabelle 1 zeigt dieses Phänomen anhand einer Simulation der Längen von erstem und zweiten Run für den Fall $p = 0,1$, wobei der Anfang einer Bernoulli-Kette 80 mal beobachtet wurde. Völlig im Einklang mit obiger Betrachtung fällt auf, dass der zweite Run selten länger als Eins ist, aber in den wenigen Fällen, in denen die Bernoulli-Kette mit dem wenig wahrscheinlichen Treffer beginnt, auch extrem lang sein kann. Dieser Umstand trifft schon auf die erste Realisierung zu, bei der auf einen kürzestmöglichen ersten Run ein zweiter Run der Länge 26 folgt.

Die Varianz von Y ergibt sich mithilfe von

$$\begin{aligned}
\mathbb{E}Y(Y-1) &= \sum_{k=2}^{\infty} k(k-1)\mathbb{P}(Y=k) \\
&= p^2 q \sum_{k=2}^{\infty} k(k-1) q^{k-2} \\
&\qquad + q^2 p \sum_{k=2}^{\infty} k(k-1) p^{k-2} \\
&= p^2 q \cdot \frac{2}{p^3} + q^2 p \cdot \frac{2}{q^3} \\
&= 2\left(\frac{q}{p} + \frac{p}{q}\right)
\end{aligned}$$

zu

$$\mathbb{V}(Y) = 2\left(\frac{1}{pq} - 3\right).$$

Insbesondere nimmt ebenso wie die Varianz der Länge des ersten Runs auch die Varianz von Y im Fall $p = 1/2$ den Minimalwert 2 an. In gewisser Weise ist somit der gleichwahrscheinliche Fall „am wenigsten spannend“.

Die obigen Überlegungen zeigen auch, dass analog zu (2) die Zufallsvariable Y in der Form

$$Y \;=\; U \cdot T + (1-U) \cdot N \tag{3}$$

dargestellt werden kann. Dabei sind U, T und N stochastisch unabhängig und besitzen die nach (2) gegebenen Bedeutungen. Der entscheidende Unterschied zu (2) besteht darin, dass die Faktoren von U und $1-U$ vertauscht sind! Analog zu oben liefert (3) unmittelbar den Erwartungswert von Y, denn es gilt

$$\mathbb{E}(Y) = \mathbb{E}U \cdot \mathbb{E}T + \mathbb{E}(1-U) \cdot \mathbb{E}N = \frac{p}{p} + \frac{q}{q} = 2.$$

4 Gemeinsame Verteilung von X und Y

Die Gleichung

$$\mathbb{P}(X = k, Y = \ell) = p^{k+1}(1-p)^{\ell} + (1-p)^{k+1} p^{\ell}$$

für die gemeinsame Verteilung von X und Y folgt ganz analog zu früher angestellten Überlegungen, wenn man nach den beiden Fällen unterscheidet, dass der erste Versuch einen Treffer oder eine Niete ergibt. Unter der (mit Wahrscheinlichkeit p eintretenden) Bedingung, dass die Bernoulli-Kette mit einem Treffer beginnt, müssen $k-1$ Treffer und danach ℓ Nieten und dann wieder ein Treffer folgen, damit das Ereignis $\{X=k, Y=\ell\}$ eintritt. Startet die Bernoulli-Kette mit einer Niete (was mit Wahrscheinlichkeit q passiert), so müssen $k-1$ Nieten und ℓ Treffer sowie danach eine Niete folgen, damit sich das Ereignis $\{X=k, Y=\ell\}$ einstellt. Nach der Formel von der totalen Wahrscheinlichkeit erhält man

$$\begin{aligned}\mathbb{P}(X=k, Y=\ell) &= \mathbb{P}(X=k, Y=\ell | A_1) \cdot p \\ &\quad + \mathbb{P}(X=k, Y=\ell | A_0) \cdot q \\ &= (p^{k-1} q^{\ell} p) \cdot p + (q^{k-1} p^{\ell} q) \cdot q\end{aligned}$$

und damit die Behauptung.

Aus der gemeinsamen Verteilung ergeben sich wegen

$$\begin{aligned}\mathbb{P}(X=k) &= \sum_{\ell=1}^{\infty} \mathbb{P}(X=k, Y=\ell), \\ \mathbb{P}(Y=\ell) &= \sum_{k=1}^{\infty} \mathbb{P}(X=k, Y=\ell)\end{aligned}$$

nicht nur die bereits hergeleiteten Verteilungen von X und Y, sondern auch die Kovarianz und der Korrelationskoeffizient zwischen X und Y. Intutiv ist zu erwarten, dass X und Y im Fall $p \neq 1/2$ negativ korreliert sind, denn ein langer erster Run (der auf $p > 1/2$ oder $p < 1/2$ hindeutet) sollte mit einem eher kurzen zweiten Run einhergehen und umgekehrt. Aus der gemeinsamen Verteilung erhält man mithilfe von (1) nach direkter Rechnung

$$\begin{aligned}\mathbb{E}(XY) &= \sum_{k=1}^{\infty} \sum_{\ell=1}^{\infty} k\ell \mathbb{P}(X=k, Y=\ell) \\ &= \frac{1}{q} + \frac{1}{p}\end{aligned}$$

und damit die Kovarianz zwischen X und Y zu

$$\begin{aligned}\mathrm{C}(X,Y) &= \mathbb{E}(XY) - \mathbb{E}X \cdot \mathbb{E}Y \\ &= \frac{1}{p} + \frac{1}{q} - \left(\frac{1}{pq} - 2\right) \cdot 2 \\ &= 4 - \frac{1}{pq}.\end{aligned}$$

Wegen $pq \leq 1/4$ folgt in der Tat $\mathrm{C}(X,Y) \leq 0$, wobei das Gleichheitszeichen nur im Fall $p = 1/2$ eintritt. Hieraus ergibt sich sofort, dass die Unabhängigkeit von X und Y die Gleichheit $p = 1/2$ nach sich zieht, denn unabhängige Zufallsvariablen sind notwendigerweise unkorreliert. Umgekehrt impliziert $p = 1/2$ die Unabhängigkeit von X und Y, denn dann nimmt die gemeinsame Verteilung die spezielle Gestalt

$$\mathbb{P}(X=k, Y=\ell) = \left(\frac{1}{2}\right)^{k+\ell}, \qquad k, \ell \geq 1,$$

an, und es gilt $\mathbb{P}(X=k) = (1/2)^k$, $\mathbb{P}(Y=\ell) = (1/2)^{\ell}$. Damit ist auch Teilaufgabe f) gelöst.

5 Höhere Gesichtspunkte

Bezeichnet allgemein R_j die Länge des j-ten Runs in einer Bernoulli-Kette mit Trefferwahrscheinlichkeit p, $0 < p < 1$, so ergibt sich für festes $m \geq 1$ die gemeinsame Verteilung von $R_1, \ldots, R_{2m}$ zu

$$\begin{aligned}&\mathbb{P}(R_1 = k_1, R_2 = \ell_1, \ldots, R_{2m-1} = k_m, R_{2m} = \ell_m) \\ &= p^{k+1} q^{\ell} + q^{k+1} p^{\ell}.\end{aligned}$$

Dabei sind $k_1, \ell_1, \ldots, k_m, \ell_m$ natürliche Zahlen, und es wurde kurz

$$k = \sum_{j=1}^{m} k_j, \ \ell = \sum_{j=1}^{m} \ell_j$$

gesetzt. Dieses Resultat erschließt sich, wenn man die beiden Fälle, dass die Bernoulli-Kette mit einem Treffer oder einer Niete startet, unterscheidet. Im ersten Fall folgen auf k_1 Treffer ℓ_1 Nieten, dann wieder k_2 Treffer, dann ℓ_2 Nieten usw., und die letzte Sequenz von ℓ_m Nieten wird mit einem Treffer abgeschlossen. Die Wahrscheinlichkeit hierfür ist

$$p^{k_1} q^{\ell_1} p^{k_2} q^{\ell_2} \cdot \ldots \cdot p^{k_m} q^{\ell_m} \cdot p \ = \ p^{k+1} q^{\ell}.$$

In gleicher Weise liefert der zweite Fall die Wahrscheinlichkeit $q^{k+1} p^{\ell}$.

Durch Summation ergeben sich aus der gemeinsamen Verteilung Marginalverteilungen für Komponenten von $(R_1, \ldots, R_{2m})$. So folgt etwa im Fall $m = 2$ aus

$$\begin{aligned}&\mathbb{P}(R_1 = k_1, R_2 = \ell_1, R_3 = k_2, R_4 = \ell_2) \\ &= p^{k_1+k_2+1} q^{\ell_1+\ell_2} + q^{k_1+k_2+1} p^{\ell_1+\ell_2}\end{aligned}$$

durch Summation über ℓ_1 und ℓ_2 die gemeinsame Verteilung der Längen des ersten und dritten Runs zu

$$\mathbb{P}(R_1 = k_1, R_3 = k_2) = q^2 p^{k_1+k_2-1} + p^2 q^{k_1+k_2-1},$$

$k_1, k_2 \geq 1$. Hieraus liest man nicht nur ab, dass R_1 und R_3 die gleiche Verteilung besitzen, sondern man kann auch direkt

$$\mathbb{E}(R_1 R_3) = \sum_{k_1, k_2 = 1}^{\infty} k_1 k_2 \mathbb{P}(R_1 = k_1, R_3 = k_2) = \frac{p}{q^2} + \frac{q}{p^2}$$

und damit

$$C(R_1,R_3) = \frac{p}{q} + \frac{q}{p} - 2 = \frac{1}{pq} - 4 = -C(R_1,R_2) \ (!)$$

erhalten. Plausiblerweise sind also die Längen des ersten und dritten Runs im Fall $p \neq 1/2$ positiv korreliert.

Aus der Gestalt der gemeinsamen Verteilung von $(R_1,\ldots,R_{2m})$ ergibt sich unmittelbar, dass diese Verteilung symmetrisch in $R_1,R_3,\ldots,R_{2m-1}$ und auch symmetrisch in $R_2,\ldots,R_{2m}$ ist. Insbesondere sind $R_3,R_5,\ldots,R_{2m-1}$ jeweils verteilt wie R_1, und $R_4,R_6,\ldots,R_{2m}$ sind jeweils verteilt wie R_2. Zudem hängen die gemeinsamen Verteilungen von R_1 und R_{2j} bzw. von R_1 und R_{2j-1} für jedes $j \geq 1$ nicht von j ab. Insbesondere gelten $C(R_1,R_2) = C(R_1,R_{2j})$, $j \geq 1$, und $C(R_1,R_3) = C(R_1,R_{2j+1})$, $j \geq 1$.

All diese Resultate sich begrifflich unmittelbar einsichtig, wenn man eine Fallunterscheidung nach dem Beginn der Bernoulli-Kette vornimmt. Mit Wahrscheinlichkeit p besteht der erste Run aus Einsen, und nachdem ein Wechsel zum zweiten, aus Nullen bestehenden Run erfolgt ist, beginnt der dritte Run mit einer Eins, und danach verhält sich die Bernoulli-Kette stochastisch so, als wäre sie zeitlich nach dem allerersten Treffer „neu geboren“. Gleiches gilt für die Situation nach dem Beginn des fünften, siebten, ... Runs. Für den Fall, dass der erste Versuch eine Niete ergibt, argumentiert man genauso, nur dass jetzt die Runs mit den ungeraden Nummern mit einer 0 beginnen.

6 Abschließende Bemerkungen

Es gibt eine umfangreiche Literatur über Runs in Bernoulli-Ketten. Hierbei interessiert meist die Verteilung der Länge des längstens Runs (aus Treffern oder Nieten), siehe z.B. Eichelsbacher (2002), Henze (1998), Schilling (1990), Schilling (2012) und die dort angegebene Literatur. Runs sind spezielle Muster aus Nullen und Einsen. Für stochastische Probleme im Zusammenhang mit Mustern in Bernoulli-Ketten siehe z.B. Henze (2001) oder Humenberger (2000).

Danksagung: Die Autoren danken den Gutachtern für diverse Verbesserungsvorschläge.

Literatur

Eichelsbacher, P. (2002): Mit RUNS den Zufall besser verstehen. In: *Stochastik in der Schule* 22, S. 2–8.

Feller, W. (1970): An Introduction to Probability Theory and Its Applications. Vol.1, 3. Auflage, Wiley, New York.

Henze, N. (1998): Stochastische Extremwertprobleme oder wie banal ist die Sensation? In: *Mitt. Math. Ges. Hamb.* 17, S. 51–74.

Henze, N. (2001): Muster in Bernoulli-Ketten. In: *Stochastik in der Schule* 21, S.2 – 10.

Henze, N. (2013): Stochastik für Einsteiger. 10. Auflage: Verlag Springer Spektrum. Heidelberg.

Humenberger, H. (2000): Überraschendes bei Münzwurfserien. In: *Stochastik in der Schule* 20, S. 4–17.

Schilling, M.F. (1990): The longest run of heads. In: *The College Mathematics Journal* 21, S. 196–207.

Schilling, M.F. (2012): The surprising predictability of Long Runs. In: *Mathematics Magazine* 85, S. 141–149.

Anschriften der Verfasser:

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