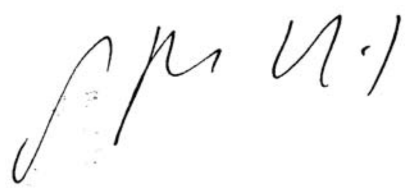
Vorwort des Herausgebers

MANFRED BOROVCNIK, KLAGENFURT

Lieber Leser, liebe Leserin!

Wir hätten Ihnen gerne das erste Heft dieses Jahrgangs mit den Ostereiern versteckt. So aber können Sie sich hoffentlich vor einem launischen Aprilwetter schützen und in den Angeboten schmökern. Der Reigen der Artikel ist weit gesteckt, diesmal sind nur deutsche Originalarbeiten enthalten. Alle Aufsätze nehmen Bezug auf den Unterricht: Einige thematisieren durchaus wichtige mathematische Beziehungen. Daten explorieren kommt genauso vor wie auch der Querbezug zu kausalem Denken und wissenschaftlicher Argumentation. Die Bibliographie mit der Besprechung von Neuerscheinungen rundet wie immer das Heft ab.
Geniessen Sie die Ideen, die Ihnen die Autoren anbieten, probieren Sie das eine oder andere im Unterricht selbst aus und geben Sie uns gelegentlich auch Rückmeldungen. Was immer Sie thematisieren, es kann unsere Zeitschrift lebendiger machen.

Es grüßt Sie Ihr



Mit RUNS den Zufall besser verstehen

PETER EICHELSBACHER, BOCHUM

Zusammenfassung: *Wir diskutieren ein interessantes Lehrexperiment im Rahmen des Modells des n-fachen unabhängigen Münzwurfs. Die Anzahl von Erfolgssträhnen ist dabei das entscheidende Maß für den Zufall. Das Experiment geht zurück auf Rényi. Es ermöglicht eine Vertiefung der Intuition zum Münzwurfexperiment und kann als eine sinnvolle Ergänzung zum schwachen Gesetz der großen Zahlen gesehen werden.*

1 Einleitung

Das klassische Experiment im Stochastikunterricht ist der n-fache Münzwurf. Werden die beiden möglichen Ausgänge eines einzelnen Wurfs mit 1 für Kopf und 0 für Zahl bezeichnet, so erzeugen wir eine Liste von n Zahlen $(w_1, \ldots, w_n) \in \{0,1\}^n$. Man möchte bekanntlich Gesetzmäßigkeiten dieser Listen studieren. So bestimmt man zum Beispiel die relative Anzahl der 1'sen in Bezug auf die Gesamtzahl der Münzwürfe. Bei einer fairen Münze erwartet man einen Wert, der nicht weit von $1/2$ entfernt ist, und man erwartet diesen Wert nahe $1/2$ umso eher, je größer n ist. Unsere Erwartung entspricht dem 1789 von Jakob Bernoulli formulierten und bewiesenen *schwachen Gesetz der großen Zahlen*. Wir rufen es weiter unten in formaler Sprache in Erinnerung.

Inwieweit gibt uns das schwache Gesetz Auskunft über den Verlauf des n-fachen Münzwurfs? Nun, zumindest erhalten wir - bei einer fair angenommenen Münze - Auskunft über die zu erwartende Anzahl von 0'en und 1'sen. Eine genauere Analyse des schwachen Gesetzes ermöglicht sehr präzise Aussagen über die Wahrscheinlichkeit sehr untypischer Listen aus $\{0,1\}^n$, siehe z. B. [1]. Doch wird die folgende Frage nicht erfaßt: Ist für $n = 12$

010011011010 oder 000011111001

typischer? Wie sieht eine durch einen Münzwurf erzeugte Liste aus? Wie sieht eine *zufällige* 0-1-Liste der Länge n des Münzwurfes aus? Um die Frage zu präzisieren, konfrontiere man eine Schulklasse mit dem folgenden Experiment:

2 Das Lehrexperiment

Man teile die Klasse in zwei gleichgroße Gruppen. In der einen Gruppe wird jeder gebeten, eine Münze 200 mal zu werfen und die 0-1-Sequenz zu notieren. In der zweiten Gruppe wird jeder gebeten, eine zufällige 0-1-Folge der Länge 200 zu notieren, als komme sie von einem Münzwurf. Die Münze soll aber nicht geworfen werden. Die Zettel werden eingesammelt und die Frage ist nun, ob man die Ergebnisse so auswerten kann, daß man beinahe jeden Zettel der richtigen Gruppe zuordnen kann. Wichtig ist, dieses Ziel erst dann zu verraten, wenn die Schüler die Aufgabe bereits ausgeführt haben!

Sicher ist es nun interessant, die Schüler zu fragen, was sie erwarten. Sehen die Münzwurf-Listen der ersten Gruppe denen der zufällig gedachten Listen ähnlich? Kann man durch Diskussion ein Kriterium erarbeiten, nach dem man die Listen untersucht? Ich stelle die Behauptung auf, dass das schwache Gesetz der großen Zahlen hier nicht helfen wird.

Eine nicht nachweisbare These ist: die Intuition von zufälliger Liste aus $\{0,1\}^n$ läßt die Mitglieder der zweiten Gruppe in der Regel eine Liste notieren, die beinahe 50% 1'sen und somit beinahe 50% 0'en enthält. Diese Intuition ist vermutlich auch dann stark ausgeprägt, wenn man das schwache Gesetz der großen Zahlen noch nicht kennt. Ich gehe noch einen Schritt weiter: die genannte Intuition wird vermutlich bei beinahe allen Teilnehmern der zweiten Gruppe dazu führen, dass die 0'en und 1'sen in recht geordneter Abwechslung auftreten. Vermutlich wird man selten viele 0'en oder 1'sen aufeinander folgen lassen. Also eher wird man

100100110111001101

notieren anstatt

000001111110001111 .

Da der Münzwurf das Zufallsexperiment ist, führt die gleiche Intuition dazu, auch hier eher eine Liste der 1. Form als Ergebnis zu erwarten. Tatsächlich ist aber die Länge von Glückssträhnen (mehrere 1'sen hintereinander) bzw. von Pechsträhnen (mehrere 0'en hintereinander) ein Aspekt, dessen Analyse zu einer Überraschung führt. Bevor man die Überraschung aufdeckt, sollte man die Schüler ihre 0-1-Liste untersuchen lassen. Sehr genau, aber auch aufwendig, wäre die Auszählung von Glückssträhnen und Pechsträhnen (wir haben ja Listen der Länge 200 erstellen lassen). Im obigem Beispiel haben wir in Bezug auf die 1'en in der 1. Liste eine Sequenz 111, zwei Sequenzen 11; in der 2. Liste eine Sequenz 111111 und eine 1111. Als dann sollte man an der Tafel eine Liste erstellen, die für beide Gruppen getrennt Auskunft über die Häufigkeit von Glückssträhnen der Länge $k \in \mathbb{N}$ gibt. Nun wird der Gruppe ein Unterschied auffallen! Alternativ verrät man zunächst die Übersicht über die Glückssträhnenlängen und ihre Häufigkeiten noch nicht, sondern nimmt die Zettel, mischt sie, sortiert sie nach einem geheimen Kriterium und ordnet die Zettel damit den Gruppen zu. Die Überraschung wird perfekt sein, denn der Leiter wird in "fast allen" Fällen die Zettel der jeweilig richtigen Gruppe zuordnen können! Ist dies geschehen, sollte ein Blick auf die Übersicht an der Tafel folgen. Die Diskussion kann beginnen, was für ein Entscheidungskriterium der Leiter verwendet haben könnte. Und nun kann man die Katze aus dem Sack lassen:

Der Lehrer weiß, und dies wollen wir in diesem Aufsatz darstellen, dass in einer Münzwurfserie der Länge 200 Glückssträhnen der Länge 7 vorkommen (nicht sicher, aber eher häufig). Ein Schüler, der von Hand eine zufällige Serie notiert, scheut es - in der Regel - , mehr als 4 mal hintereinander das gleiche Symbol zu notieren. Der Lehrer wählt das folgende Kriterium:

Sind Glückssträhnen der Länge größer 5 vorhanden, ordnet er den Zettel der Gruppe zu, die die Münze 200 mal geworfen hat. Er liegt fast immer richtig. Die beschriebene Häufigkeitsübersicht an der Tafel wird dies vermutlich als ein "sinnvolles" Kriterium erscheinen lassen. Die Diskussion mit der Klasse sollte auch ergeben, dass die im folgenden zu begründende Zahl 7 natürlich mit der Anzahl der Experimente, es war 200, zusammenhängt.

Wir kommen nun zur mathematischen Analyse dieses Lehrexperiments:

3 Eine Heuristik

Gegeben sei der Ereignisraum $\Omega = \{0,1\}^n$. Für ein $\omega \in \Omega$ bezeichne $X_i(\omega) = \omega_i$ den Ausgang des i'ten Teilexperiments. Setzen wir $p(\omega) = p^k(1-p)^{n-k}$ mit $0 < p < 1$, wenn k die Anzahl der 1'en in ω bezeichnet, so ist

$$P(X_i = 1) = \sum_{\omega : X_i(\omega)=1} p(\omega) = p$$

und $P(X_i = 0) = 1 - p$. Die Zufallsgrößen $X_1, \dots, X_n$ beschreiben also die n Ausgänge des Münzwurfexperiments. Wir untersuchen nun das Auftreten von Runs. Hierbei ist formal jede maximale Teilfolge von einander benachbarten gleichen Symbolen ein *Run*. So sind zum Beispiel in

1011110000001001

4 1-Runs, einer der Länge 4 und drei der Länge 1.

Eine Aneinanderreihung von m Einsen (an einer beliebigen Position startend) tritt mit Wahrscheinlichkeit p^m auf. Es können den m Einsen natürlich ohne Unterbrechung noch weitere Einsen folgen. Eine solche Aneinanderreihung ist also immer eine von mindestens m Einsen. Wir bezeichnen sie in der Folge als m-Folge. Wir betrachten ein heuristisches Argument: Ein Aneinanderreihung von m Eisen kann an Position 1, an Position $2, \dots,$ an Position $n-m+1$ beginnen. Es gibt also in $(\omega_1, \dots, \omega_n)$ $n-m+1$ mögliche Positionen, jede so positionierte m-Folge von Einsen tritt mit Wahrscheinlichkeit p^m auf. Es bezeichne Y_n die Anzahl der m-Folgen in $(\omega_1, \dots, \omega_n)$. Wir erwarten "im Mittel" $(n-m+1)\,p^m$ viele solcher m-Folgen. Wenn wir m fixieren und uns n groß denken, ist der Erwartungswert von Y_n approximativ np^m (wir berechen den Erwartungswert von Y_n weiter unten genau). Es bezeichne nun $R_n(\omega)$ die Länge des längsten 1-Runs für jedes $\omega \in \Omega$. R_n ist natürlich eine Zufallsgröße. R_n entspricht der längsten vorkommenden Aneinanderreihung von Einsen. Im Mittel gibt es somit $\approx np^{R_n}$ längste 1-Runs. Wenn wir nun - im Rahmen dieser Heuristik - annehmen, dass es im Mittel nur genau einen längsten 1-Run gibt, so erfüllt R_n unter dieser Annahme angenähert

$$1 \approx n\,p^{R_n}.$$

Denken wir uns das $\approx$ als Gleichung, so können wir nach R_n auflösen:

$$\log 1 \approx \log n + R_n \log p \qquad \text{oder} \qquad R_n \approx \frac{\log n}{\log 1/p},$$

wobei wir ohne Einschränkung den Logarithmus naturalis (Basis e) nehmen. Wenn wir in der obigen Heuristik keinen Fehler gemacht haben, so haben wir angenähert eine Auskunft über die Größe von R_n hergeleitet. Werfen wir nun wie im Lehrexperiment 200 mal die faire Münze $(n = 200\,, p = 1/2)$, so folgt

$$R_{200} \approx \frac{\log 200}{\log 2} \approx 7.64\,.$$

Der längste Run in einem 200-fachen Münzwurf ist also eine Zufallsgröße, die in einem zu klärenden Sinn typischerweise Werte bei 7 annimmt. Somit haben wir über einen einfachen Gedankenweg die eigentliche Idee für das Kriterium bereits dargestellt. Im Nachhinein verstehen wir nun auch, warum das Experiment 200 Münfwürfe vorsieht. Dazu ein paar Vergleichswerte für $\frac{\log n}{\log 1/p}$ für festes $p = \frac{1}{2}$:

$$n = 50 \qquad \log n/\log 2 \approx 5.64$$

$$n = 100 \qquad \log n/\log 2 \approx 6.64$$

$$n = 150 \qquad \log n/\log 2 \approx 7.23$$

$$n = 200 \qquad \log n/\log 2 \approx 7.64$$

$$n = 300 \qquad \log n/\log 2 \approx 8.23$$

Sicher ist das Nachvollziehen der Heuristik schon ein wesentlicher Schritt, um das Lehrexperiment zu verstehen. Doch sollte man nach der experimentellen und der heuristischen Phase nun dazu übergeghen, formal mit m-Folgen und mit 1-Runs zu rechnen.

4 Die Aussagen in der Stochastik-Sprache

Dazu berechnen wir den Erwartungswert von Y_n (die Anzahl der Positionen, an denen mindestens m Einsen in Folge folgen) genau und schließen zur Übung noch an, die Zufallsgröße Z_n der Anzahl aller Runs in n Würfen zu untersuchen. Dann wollen wir die Aussage

$$\frac{R_n}{\log n} \approx \frac{1}{\log 1/p}$$

mathematisch präzise fassen. Es wird sich ein *neues schwaches Gesetz der großen Zahlen* ergeben.

Es sei $Y_i^{(m)} := \prod_{j=i}^{i+m-1} X_j = X_i X_{i+1} \cdots X_{i+m-1}$. Diese kompliziert anmutende Zufallsgröße nimmt nur zwei Werte an: den Wert 1 nur in dem Fall, in dem *alle* $X_i, X_{i+1}, \dots, X_{i+m-1}$ den Wert 1 annehmen, also den Wert 1, wenn eine m Folge in Position i startet, also ein 1-Run der Länge $\geq m$ - in Position i startend - vorliegt. Wir betrachten nun

$$Y_n := \sum_{i=1}^{n-m+1} Y_i^{(m)}.$$

Nun ist der Erwartungswert linear:

$$E(Y_n) = \sum_{i=1}^{n-m+1} E(Y_i^{(m)})$$

und nach Definition des Erwartungswerts ist

$$\begin{aligned} E(Y_i^{(m)}) &= 1 \cdot P(Y_i^{(m)} = 1) + 0 \cdot P(Y_i^{(m)} = 0) \\ &= P(Y_i^{(m)} = 1) \\ &= P(X_i = 1, X_{i+1} = 1, \ldots, X_{i+m-1} = 1) = p^m \end{aligned}$$

für alle $i = 1, \ldots, n-m+1$. Also ist $E(Y_n) = (n-m+1)p^m$, wie oben bereits behauptet. Für $m = 7$ ist im Fall $n = 200$ und $p = \frac{1}{2}$ ist $E(Y_{200}) \sim 200(\frac{1}{2})^7 > 1$.

Sei Z_n die Anzahl aller Runs in n Würfen. Wir wollen $E(Z_n)$ bestimmen. Eine sehr lehrreiche Darstellung von Z_n ist die folgende: in $\Omega = \{0,1\}^n$ betrachte $A_i = \{\omega = (\omega_1, \ldots, \omega_n) \in \Omega \colon \omega_i = 1\}$, $1 \leq i \leq n$. Dann ist $P(A_i) = p$ und $A_1, \ldots, A_n$ sind unabhängige Ereignisse. Sei nun

$$B_j = (A_{j-1} \cap A_j^c) \cup (A_{j-1}^c \cap A_j)$$

für $2 \leq j \leq n$. $A_{j-1} \cap A_j^c$ ist das Ereignis, dass an Position $j-1$ eine 1 und an Position j eine 0 realisiert wurde, $A_{j-1}^c \cap A_j$ das Ereignis, an Position $j-1$ eine 0 und an Position j eine 1 realisiert zu haben. B_j beschreibt also das Ereignis zwischen Position $j-1$ und j einen Wechsel des Wertes vorliegen zu haben. Nun ist die Anzahl der Wechsel zwischen den Werten 0 und 1 bei n Würfen gleich $Z_n - 1$, also gilt

$$Z_n = 1 + \sum_{j=2}^{n} 1_{B_j},$$

wobei

$$1_{B_j}(\omega) = \begin{cases} 1 & \omega \in B_j, \\ 0 & \text{sonst.} \end{cases}$$

Also ist $E(Z_n) = 1 + \sum_{j=2}^n E(1_{B_j})$. Nun ist $E(1_{B_j}) = 1 \cdot P(B_j) + 0 \cdot P(B_j^c)$, wobei $B_j^c = \Omega \backslash B_j$ bezeichne. Also ist $E(1_{B_j}) = P(B_j) = p(1-p) + (1-p)p = 2p(1-p)$. Wir erhalten somit

$$E(Z_n) = 1 + \sum_{j=2}^{n} 2p(1-p) = 1 + 2(n-1)p(1-p).$$

Für $p = \frac{1}{2}$ und $n = 200$ ist $E(Z_{200}) = 1 + 2 \cdot 199 \cdot \frac{1}{4} \approx 100$. Eine etwas kompliziertere Aufgabe ist die Bestimmung der Varianz von Z_n. Die Berechnung verlangt die Kenntnis von Kovarianzen und wird hier nicht vorgeführt. Es gilt

$$V(Z_n) = (4n-6)p(1-p) + 4(5-3n)p^2(1-p)^2,$$

wobei $V(Z_n)$ die Varianz bezeichne.

Wir haben nun ein paar Erkenntnisse über Erwartungswerte für Runs gesammelt und fragen nun, inwieweit unsere Heuristik

$$R_n \approx \frac{\log n}{\log 1/p}$$

mathematisch präzisiert werden kann. Dazu erinnern wir zunächst an das gewöhnliche schwache Gesetz der großen Zahlen: Es sei $S_n := X_1 + \cdots + X_n$, so ist $E(S_n) = np$ und $E(\frac{S_n}{n}) = p$ und es gilt für jedes $\varepsilon > 0$:

$$\lim_{n \to \infty} P\left(\left| \frac{S_n}{n} - p \right| > \varepsilon \right) = 0. \tag{1}$$

Die mathematische Präzisierung unserer Überlegungen mündet nun ebenfalls in eine Formulierung wie unter (1). Tatsächlich gilt:

Theorem 2 (Erdös, Rényi; 1970) *Für jedes* $\varepsilon > 0$ *gilt beim n-fachen unabhängigen Münzwurf mit Erfolgswahrscheinlichkeit* $0 < p < 1$*:*

$$\lim_{n \to \infty} P\left(\left| \frac{R_n}{\log n} - \frac{1}{\log 1/p} \right| > \varepsilon \right) = 0.$$

Dieses Gesetz wird in der Originalarbeit [2] aus dem Jahre 1970 “a new law of large numbers”, ein neues Gesetz der großen Zahlen, genannt. $R_n/\log n$ ist mit großer Wahrscheinlichkeit in der Nähe des Wertes $(\log(1/p))^{-1}$. In der Arbeit von Erdös und Rényi wird indirekt ein starkes Gesetz der großen Zahlen hergeleitet. Genauer besagt dies, dass $R_n/\log n$ fast sicher gegen $(\log(1/p))^{-1}$ konvergiert. Wir führen dies hier nicht weiter aus, da der Begriff der fast sicheren Konvergenz in der Schule keine Rolle spielt. Weiter ist zu bemerken, dass in der Arbeit 1970 tatsächlich mehr gezeigt wurde. Dort betrachtet man die Summe von Bernoulli-verteilten Zufallsgrößen über einen Fensterbereich der Größenordnung $\log_2 n$, der richtig skaliert noch ein starkes Gesetz liefert. Auch diesen Teil größerer Abstraktion führen wir hier nicht weiter aus.

Für eine Unterrichtsreihe sollte man nun an den Beweis von (1) erinnern und daraufhinweisen, dass R_n, die Länge des längsten Runs, eine kompliziertere Darstellung hat als z. B. $S_n = X_1 + \cdots + X_n$.

Das ist quasi das Hauptproblem! Es gibt keine geschlossene Darstellung z. B. als Summe von einfacheren Zufallsgrößen und somit ist die Berechnung der Varianz von R_n, die nach der Tschebyschev-Ungleichung gebraucht wird, unklar. (Wir haben bisher noch nicht einmal $E(R_n)$ explizit bestimmt!) Wir geben im Anhang einen wesentlichen Beweisschritt an und darüberhinausgehende Literatur und empfehlen das Studium des Anhangs zum Beispiel auch dem Schüler oder der Schülerin, der/die sich im Rahmen einer Facharbeit mit dem Thema *Runs* beschäftigen könnte.

Zusammenfassend empfehle ich obiges Lehrexperiment zur Schärfung von Intuition und der Kenntnis von Gesetzmäßigkeiten beim n-fachen Münzwurf.

5 Kombinatorisches

In diesem Abschnitt betrachten wir die Frage, wie groß die Wahrscheinlichkeit ist, bei sagen wir 17 Einsen und 11 Nullen genau 7 1-Runs und 5 0-Runs vorzufinden. Dies ist eine durchaus knifflige kombinatorische Aufgabe. Wir folgen den Ausführungen in [3]. Zunächst benötigt man die folgende Hilfsaussage:

Die Zahl der Möglichkeiten, n ununterscheidbare Kugeln auf N Zellen so zu verteilen, dass keine Zelle ganz leer bleibt, ist $\binom{n-1}{N-1}$.

Bevor wir die Hilfsaussage beweisen, betrachten wir die Anwendung. Es liege eine Folge von d Einsen und h Nullen vor. Da die Einsen sowie die Nullen ununterscheidbar sind, gibt es $\binom{d+h}{d}$ Anordnungen für eine solche Folge. Wir wollen nun auszählen, wieviele davon sagen wir r 1-Runs und s 0-Runs haben. Es ist klar, dass nach Definition eines Runs als einer maximalen Teilfolge von einander benachbarten gleichen Symbolen sich die 1-Runs und 0-Runs abwechseln. Also ist s um 1 größer oder kleiner als r oder gleichgroß: $r-1 \leq s \leq r+1$. Man möchte nun die vorhandenen d Einsen auf r Runs verteilen, also in Übersetzung der obigen Hilfsaussage auf r Zellen, so dass keine Zelle ganz leer bleibt - also wirklich ein 1-Run (einer nicht bestimmten Länge) vorliegt. Es gibt also $\binom{d-1}{r-1}$ Möglichkeiten, die Längen der 1-Runs festzulegen. Analog gibt es $\binom{h-1}{s-1}$ Möglichkeiten die Längen der 0-Runs festzulegen. Also gibt es insgesamt

$$\binom{d-1}{r-1}\binom{h-1}{s-1}$$

Möglichkeiten, die Längen aller Runs festzulegen. Nun unterscheiden wir den Fall $r = s+1$, in dem der erste Run ein 1-Run sein muß, vom Fall $r = s-1$, in dem der erste Run ein 0-Run ist. In diesen Fällen liegt nach der Wahl der Längen aller 0-Runs und 1-Runs die ganze Folge der $d+h$ Werte eindeutig fest. Die gesuchte Wahrscheinlichkeit ist dann

$$\frac{\binom{d-1}{r-1}\binom{h-1}{s-1}}{\binom{d+h}{d}}.$$

Im Fall $r = s$ wird diese Wahrscheinlichkeit mit 2 multipliziert, da man in diesem Fall mit einer 1 oder einer 0 starten kann. Das obige Zahlenbeispiel liefert $\binom{16}{6}\binom{10}{4}/\binom{28}{17}$.

Nun zum Beweis der Hilfsaussage: Ein schöner Trick ist, die N Zellen zunächst mit je einer Kugel zu füllen. Damit ist jede Zelle gefüllt und es verbleibt die Aufgabe, die übrigen $n-N$ Kugeln, die ununterscheidbar sind, auf die Zellen bei möglicher Mehrfachbesetzung zu verteilen. Diese Situation entspricht dem Urnenproblem, aus einer Stichprobe vom Umfang N (Zellen) $n-N$ ohne Reihenfolge (ununterscheidbare Kugeln) und mit Zurücklegen (mit Mehrfachbesetzung der Zelle) zu ziehen. Dafür gibt es bekanntlich

$$\binom{(n-N)+N-1}{n-N} = \binom{n-1}{N-1}$$

Möglichkeiten, was zu zeigen war.

6 Anhang

1. Wir betrachten hier zwei sich leicht unterscheidende Beweise von Theorem 2. In beiden Beweisen betrachten wir das Ereignis

$$\frac{R_n}{\log n} - \frac{1}{\log 1/p} > \varepsilon.$$

Der Fall

$$\frac{R_n}{\log n} - \frac{1}{\log 1/p} < -\varepsilon$$

geht analog. Es gilt

$$\left\{\frac{R_n}{\log n} - \frac{1}{\log 1/p} > \varepsilon\right\} = \left\{R_n > \varepsilon \log n + \frac{\log n}{\log 1/p}\right\}.$$

Der längste 1-Run ist hier also größer oder gleich $\varepsilon \log n + \frac{\log n}{\log 1/p}$. Da die Länge eine natürliche Zahl ist, können wir auch zur unteren Gauß-Klammer [] übergehen. Nun teilen wir den Bereich der Länge n in disjunkte Blöcke der Länge

$$t = \left[\left(\varepsilon + (\log(1/p))^{-1}\right) \log n\right] + 1.$$

Dann impliziert das Ereignis, dass der längste 1-Run größer oder gleich $\varepsilon \log n + \frac{\log n}{\log 1/p}$ ist, dass in mindestens einem der Blöcke alle Positionen mit einer 1 belegt sind, was mit Wahrscheinlichkeit p^t geschieht. Somit ist

$$P\left\{ \frac{R_n}{\log n} - \frac{1}{\log 1/p} > \varepsilon \right\} \leq [n/t]\, p^t \leq n\, p^t.$$

Dafür müssen wir allerdings n hinreichend groß wählen, damit der "Restteil" der n Positionen bei der Blockzerlegung in Blöcke der Länge t vernachlässigbar klein wird. Wir werden weiter unten sehen, dass das Einsetzen der Definition von t liefert:

$$\lim_{n\to\infty} n\, p^t = 0.$$

Damit ist der erste Beweis vorgeführt. Dies ist der typische Beweis in der Literatur, der aber den Punkt des Restteils für den Schulbereich etwas im Dunkeln lässt, da hier ungewohnte Argumente auftauchen.

Der zweite Weg ist etwas umständlicher, formal aber klarer. Dazu formalisieren wir (zum ersten Mal) die Zufallsgröße

$$R_n = \max\{l-k\colon\ 0 \leq k < l \leq n\colon\ S_l - S_k = l-k\}\,.$$

Hierbei setzt man $S_0 = 0$. Bis zum "Zeitpunkt" n tritt also ein 1-Run der Länge r oder länger ein. Wir betrachten nun den ersten Zeitpunkt des Eintritts eines 1- Runs der Länge r oder länger. Formalisiert ist dies

$$T_r = \inf\{l\colon\ S_l - S_k = l-k \quad \text{für ein} \quad 0 \leq k \leq l-r\}\,.$$

Tritt das Ereignis im Fall $k = l-r$ auf, so ist T_r der kleinste Wert l, für den $S_l - S_{l-r} = r$ gilt. Denken wir an die Definition von S_n, so folgt $S_l - S_{l-r} = X_{l-r+1} + \cdots + X_l$ den Wert r hat, was aber nur geht, wenn alle X_i hier den Wert 1 annehmen. Ist $k < l-r$, liegt ein 1-Run der Länge $> r$ vor. Somit ist $\{R_n \geq r\} = \{T_r \leq n\}$. Nun ist für $m \in \mathbb{N}$

$$\{T_r \leq m\} \subset \bigcup_{k=0}^{m-r} \bigcup_{l=k+r}^{m} \{S_l - S_k = l-k\}$$

und dies ist enthalten in

$$\bigcup_{k=0}^{m-1} \bigcup_{l=k+r}^{m} \{S_l - S_k = l-k\}.$$

Weiter ist

$$P\left(\frac{S_l - S_k}{l-k} = 1\right) = P\left(\frac{S_{l-k}}{l-k} = 1\right) = p^{l-k},$$

und somit

$$P(T_r \leq m) \leq m \sum_{n=r}^{m} p^n = m p^r \sum_{n=0}^{m-r} p^n\,.$$

Hier ist $\sum_{n=0}^{m-r} p^n$ endlich (geometrische Partialsumme).

Der wirklich aufwendige Teil des Beweises ist die Umschreibung unseres Ereignisses mit Hilfe der Zufallsgrößen T_r und der Abschätzung mit Hilfe der Ereignisse $\{S_l - S_k = l-k\}$, deren Wahrscheinlichkeit wir einfach berechnen können. Tatsächlich folgt nun

$$P\left(R_n \geq \varepsilon \log n + \frac{\log n}{\log 1/p}\right) = P\Big(T_r \leq n\Big) \leq n p^r c$$

mit $r = \varepsilon \log n + \frac{\log n}{\log 1/p}$ und c ist eine Konstante.

Nun ist

$$\lim_{n\to\infty} n p^{\varepsilon \log n} p^{\log n (\log 1/p)^{-1}} = \lim_{n\to\infty} p^{\varepsilon \log n} = 0,$$

denn es ist $0 < p < 1$, $\varepsilon > 0$ und

$$\frac{\log n}{\log 1/p} = \log_{1/p} n \quad \text{sowie} \quad p^{\log_{1/p} n} = \frac{1}{n}\,.$$

Die letzten beiden Rechenzeilen liefern für den ersten Beweisweg die Rechung nach.

2. Abschliessend sei eine Erweiterung zu dem 1970 von Erdös und Rényi erzielten Resultat erwähnt. Es gibt einen interessanten Zusammenhang zu einer verbesserten Tschebychev-Ungleichung, so wie sie zum Beispiel in [Eich'99] dargestellt wurde. Dort findet man für $\alpha > 0$

$$P\left(\frac{S_n}{n} - p \geq \alpha\right) \leq \exp(-nH(\alpha + p|p)) \qquad (3)$$

mit

$$H(x|p) := x \log \frac{x}{p} + (1-x) \log \frac{1-x}{1-p}$$

für $x \in (0,1)$. Die stetige Fortsetzung in $x = 1$ liefert $H(1|p) = \log 1/p$ (hierbei ist $0 \log 0 = 0$). Betrachtet man nun zu einem Intervall $J = [a,b]$ mit $0 \leq a \leq b \leq 1$

$$R_n^J = \max\left\{l-k\colon\ \frac{S_l - S_k}{l-k} \in J\right\},$$

so gilt

$$\lim_{n\to\infty} P\left(\left|\frac{R_n^J}{\log n} - \frac{1}{H_J}\right| > \varepsilon\right) = 0$$

mit $H_J := \inf_{x \in J} H(x|p)$. Für $J = \{1\}$ ist dies die Situation bei den 1-Runs. Für den allgemeineren Fall nutzt man im Beweis (3). In der Fachliteratur wird ein Gesetz wie für R_n oder R_n^J *Erdös-Rényi-Gesetz* genannt. Sie bekommen zunehmend Bedeutung in Anwendungen wie zum Beispiel der mathematischen Biologie [4].

Literatur

[1] Eichelsbacher, P., (1999): Wie gut ist die Tschebychev-Ungleichung? Mathematik in der Schule.

[2] Erdös, P. and Rényi, A. (1970): On a new law of large numbers, Journ. Analyse Mathématique, 23, 103-111.

[3] Krengel, U. (1991): Einführung in die Wahrscheinlichkeitstheorie und Statistik, Vieweg Studiem, Aufbaukurs Mathematik, Braunschweig, 3. erweiterte Auflage.

[4] Waterman, M. (1995): Introduction to computational biology, Capman & Hill.

Anschrift des Verfassers
Peter Eichelsbacher
Fakultät für Mathematik
Ruhr-Universität Bochum
NA 3 / 68
44780 Bochum
`peter.eichelsbacher@ruhr-uni-bochum.de`

Eine Formel von de Moivre

PETER EICHELSBACHER, BOCHUM

Zusammenfassung: *Wir diskutieren eine in der Literatur fast in Vergessenheit geratene Formel von de Moivre zur Berechnung des absoluten Fehlers zwischen der relativen Häufigkeit und dem Erwartungswert in einer Bernoulli–Kette. Dabei geben wir eine historische Übersicht über die Entwicklung dieser Formel Bezug nehmend auf einen Artikel von Persi Diaconis und Sandy Zabell in Statistical Science.*

1 Einleitung

Das sicher am häufigsten im Stochastikunterricht behandelte Zufallsexperiment ist das des n-maligen unabhängigen Münzwurfs. Gegeben sei $\Omega = \{E, M\}$, wobei E für Erfolg und M für Mißerfolg steht. Die Wahrscheinlichkeit für Erfolg sei $0 < p < 1$ (der Stochastiker interessiert sich nicht für den Fall $p = 0$ oder $p = 1$). Dann setze $\Omega_n = \{E, M\}^n$, die Menge der E-M-Folgen der Länge n. Die Wahrscheinlichkeit für jedes $\omega = (\omega_1, \ldots, \omega_n) \in \Omega_n$ sei gegeben durch $p(\omega) = p^k (1-p)^{n-k}$, wobei k die Anzahl der E's in der Folge $\omega_1, \ldots, \omega_n$ bezeichnet. X sei die Anzahl der Erfolge in der Bernoulli-Kette der Länge n zum Parameter p. Darunter versteht man nichts anderes als den gerade beschriebenen Wahrscheinlichkeitsraum (Ω_n, p). Setze $X_i = 1$, falls $\omega_i = 1$ und $X_i = 0$ sonst, $1 \le i \le n$, so folgt $X = X_1 + X_2 + \cdots + X_n$. Es gilt

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k} =: b(k; n, p)$$

für $k \in \{0, \ldots, n\}$. Die Zufallsgröße mit obiger Verteilung heißt binomialverteilt mit Parametern n und p. Es gilt $\mathbb{E}(X) = np$ und $V(X) = np(1-p)$, wobei $\mathbb{E}(X)$ den Erwartungswert von X bezeichnet und $V(X)$ die Varianz von X. Eine gute Darstellung findet man zum Beispiel im Buch von Barth und Haller [1]. Mit diesem Modell der Bernoulli-Kette arbeitet man auch in der *Statistik*, wo es darum geht, aus bekanntem n und einer Beobachtung von sagen wir k Erfolgen Rückschlüsse auf den unbekannten Parameter p zu ziehen. Nun kann man Schülern schnell klar machen, dass die sogenannte *naive Schätzung* k/n (in der Sprache der Zufallsgrößen X/n) ein vernünftiger Schätzer für p ist. Es gilt

$$\mathbb{E}(X/n) = np/n = p.$$

Im Mittel kommt also bei dieser Schätzung der wahre Parameter raus (man nennt das in der Fachsprache *Erwartungstreue* des Schätzers X/n). Die Erwartungstreue ist quasi ein erstes Indiz für die Tauglichkeit der Wahl X/n. Weiter kann man plausibel machen, dass der Fehler dieser Schätzung in der Form $X/n - p$ dargestellt werden sollte. Da aber $X/n - p$ eine Zufallsgröße ist, kann man den mittleren Fehler bestimmen, nun gilt aber bekanntlich $\mathbb{E}(X/n - p) = \mathbb{E}(X/n) - p = p - p = 0$. Das einfache Subtrahieren liefert also immer einen "Nullfehler". Somit ist $X/n - p$ kein geieigneter Kandidat für eine Maßzahl