UNABHÄNGIGE EREIGNISSE IN DISKRETEN WAHRSCHEINLICHKEITSMODELLEN

von **Dietmar Pfeifer**, Oldenburg

Zusammenfassung

Anhand konkreter Aufgaben aus verschiedenen Schulbüchern zur Stochastik wird die Problematik der Existenz unabhängiger Ereignisfolgen in diskreten Wahrscheinlichkeitsmodellen diskutiert. Insbesondere wird eine in diskreten Modellen gültige, einfache Abschätzung für die maximale Anzahl unabhängiger, *gleichwahrscheinlicher* Ereignisse angegeben. Hieraus folgt, daß viele der in der Schule im Zusammenhang mit Wartezeitproblemen gestellten Aufgaben im dort vorgegebenen Rahmen mathematisch nicht korrekt behandelbar sind. Zu einigen Aufgaben werden deshalb leichte Modifikationen vorgestellt, die diese Schwierigkeiten umgehen.

1. Einleitung

Der Begriff der stochastischen Unabhängigkeit von Ereignissen oder Zufallsgrößen spielt eine zentrale Rolle sowohl in der Wahrscheinlichkeitsrechnung — etwa im Zusammenhang mit dem Gesetz der großen Zahlen — als auch in der Statistik, z.B. bei der Behandlung von Tests (unabhängige Stichproben). Dementsprechend wird diese Thematik in allen Grund– und Leistungskursen zur Stochastik in praktisch allen Schulformen ausführlich behandelt. Der wohl größere Teil der hier verwendeten Lehrbücher beschränkt sich dabei auf eine endlich–diskrete Sichtweise, was zur Vermittlung grundlegender Kenntnisse sicher ausreicht (vgl. z.B. ALTHOFF [1]) und den Einsatz lediglich relativ elementarer Mittel erfordert. Methodische Schwierigkeiten ergeben sich allerdings dann, wenn darüber hinausgehend auch unendlich–diskrete Wahrscheinlichkeitsmodelle behandelt werden (vgl. z.B. BOSCH / WOLFF [4], S. 38ff) oder konkrete Aufgabenstellungen in den elementareren Unterrichtswerken faktisch nicht–endliche Betrachtungen erfordern.

Wir wollen zunächst kurz in Erinnerung rufen, was man unter der stochastischen Unabhängigkeit von Ereignissen versteht. Von den verschiedenen Zugangsmöglichkeiten wählen wir hier diejenige in BARTH / HALLER ([2], Definition 156.1), da sie für die nachfolgenden Argumentationen am besten geeignet ist.

Definition: Ereignisse $A_1, \ldots, A_n$ in einem diskreten Wahrscheinlichkeitsraum (Ω, P) heißen stochastisch unabhängig, wenn die folgenden 2^n Gleichungen gelten:

$$
\begin{aligned}
P(A_1 \cap A_2 \cap \ldots \cap A_n) &= P(A_1)P(A_2) \cdot \ldots \cdot P(A_n) \\
P(\overline{A_1} \cap A_2 \cap \ldots \cap A_n) &= P(\overline{A_1})P(A_2) \cdot \ldots \cdot P(A_n) \\
P(A_1 \cap \overline{A_2} \cap \ldots \cap A_n) &= P(A_1)P(\overline{A_2}) \cdot \ldots \cdot P(A_n) \\
&\vdots \\
P(A_1 \cap A_2 \cap \ldots \cap \overline{A_n}) &= P(A_1)P(A_2) \cdot \ldots \cdot P(\overline{A_n}) \\
P(\overline{A_1} \cap \overline{A_2} \cap \ldots \cap A_n) &= P(\overline{A_1})P(\overline{A_2}) \cdot \ldots \cdot P(A_n) \\
&\vdots \\
P(\overline{A_1} \cap \overline{A_2} \cap \ldots \cap \overline{A_n}) &= P(\overline{A_1})P(\overline{A_2}) \cdot \ldots \cdot P(\overline{A_n})
\end{aligned}
$$

oder kürzer:

$$
P\left(\bigcap_{k=1}^{n} B_k\right) = \prod_{k=1}^{n} P(B_k), \tag{1}
$$

wobei für die Mengen B_i wahlweise die Ereignisse A_i selbst oder deren Komplemente $\overline{A_i}$, $1 \leq i \leq n$, zu wählen sind.

Eine durch längere Rechnung nachzuprüfende Konsequenz hieraus ist u.a.:
Mit n Ereignissen $A_1, \ldots, A_n$ sind auch alle Auswahlen $A_{i_1}, A_{i_2}, \ldots, A_{i_k}$ von $k \leq n$ Ereignissen stochastisch unabhängig. (Diese hier aus der Definition abgeleitete Eigenschaft wird gelegentlich auch direkt zur Definition der stochastischen Unabhängigkeit verwendet, z.B. in LAUTER / RÜDIGER [7], S. 141, Def. 10.4, oder BOSCH / WOLF [4], S. 71, Def. 7.)

Anhand der beiden folgenden Beispiele wollen wir nun die eingangs angesprochene Problematik einmal exemplarisch verdeutlichen:

Drei Spieler A,B,C werfen abwechselnd in dieser Reihenfolge eine Münze. Gewonnen hat, wer zuerst Wappen wirft. Mit welcher Wahrscheinlichkeit gewinnt A (B;C)?

(LAUTER / RÜDIGER [7], S. 142, Aufgabe 12)

Zwei Laplace–Münzen werden beliebig oft geworfen, bis beide gleichzeitig Adler zeigen. Die Zufallsgröße X sei die Anzahl der dazu nötigen Würfe.

a) Gib einen gröberen Ergebnisraum an. Bestimme die Wahrscheinlichkeitsverteilung von X.

b) Berechne $E(X)$ und $\mathrm{Var}\,(X)$.

(BARTH / HALLER [2], S. 226, Aufgabe 24)

Als mögliche "Lösung" etwa der ersten Aufgabe würde die folgende Argumentation sicher von jeder Lehrkraft akzeptiert werden:

"Bezeichnet man mit A_n das Ereignis, daß beim n–ten Wurf Wappen erscheint und mit $\overline{A_n}$ das komplementäre Ereignis, daß im n–ten Wurf Zahl erscheint, sowie mit B_n das Ereignis, daß im n–ten Wurf *erstmalig* Wappen erscheint, also

$$B_n = \begin{cases} A_1, & n = 1 \\ \overline{A_1} \cap \ldots \cap \overline{A_{n-1}} \cap A_n, & n \geq 2, \end{cases}$$

so lassen sich die Ereignisse G_A, G_B, G_C, daß Spieler A,B,C gewinnt, mengentheoretisch ausdrücken durch

$$G_A = \bigcup_{n=0}^{\infty} B_{3n+1}, \quad G_B = \bigcup_{n=0}^{\infty} B_{3n+2}, \quad G_C = \bigcup_{n=0}^{\infty} B_{3n+3},$$

da jeder der Spieler nur jedes dritte Mal an die Reihe kommt und nur dann gewinnt, wenn vor ihm kein anderer Spieler bereits Wappen geworfen hat (d.h. z.B. Spieler A kann nur in den Runden 1, 4,7,11,... gewinnen). Wegen der Unabhängigkeit der einzelnen Würfe erhält man für $n \geq 2$

$$P(B_n) = P(\overline{A_1}) \cdot \ldots \cdot P(\overline{A_{n-1}}) \cdot P(A_n) = 2^{-n},$$

was wegen $B_1 = A_1$ auch noch für $n = 1$ richtig bleibt. Die Additivitätseigenschaft von P ergibt damit insgesamt unter Ausnutzung der Formel für die geometrische Reihe

$$P(G_A) = \sum_{n=0}^{\infty} P(B_{3n+1}) = \sum_{n=0}^{\infty} 2^{-(3n+1)} = 4/7,$$

$$P(G_B) = \sum_{n=0}^{\infty} P(B_{3n+2}) = \sum_{n=0}^{\infty} 2^{-(3n+2)} = 2/7,$$

$$P(G_C) = \sum_{n=0}^{\infty} P(B_{3n+3}) = \sum_{n=0}^{\infty} 2^{-(3n+3)} = 1/7."$$

Diese Aufgabe (wie entsprechend auch die andere) erfordert aber für eine mathematisch korrekte Lösung die Betrachtung einer ganzen *Folge*, also unendlich vieler stochastisch unabhängiger Ereignisse $A_1, A_2, A_3, \ldots$ (was bedeutet, daß je endlich viele von ihnen stochastisch unabhängig sind im Sinne der obigen Definition). Eine solche Folge existiert aber i.a. weder in einem endlich– noch unendlich–diskreten Rahmen, sondern erfordert die Betrachtung allgemeiner, überabzählbarer Meßräume im Sinne der mathematischen Maßtheorie. Die Angabe eines "gröberen Ergebnisraumes" im zweiten Beispiel ist also in korrekter Weise mit Schulmitteln gar nicht leistbar — allenfalls läßt sich hier heuristisch argumentieren wie z.B. in LAKOMA [6].

In Anbetracht dieser Tatsache erscheint es mir fraglich, ob Wartezeitaufgaben der obigen Art ohne entsprechende Erläuterungen überhaupt im Schulunterricht behandelt werden sollten, zumal es ohne weiteres möglich ist, durch einfache Modifikationen im endlich–diskreten Rahmen zu bleiben, ohne an Substanz wesentlich zu verlieren. Ein schönes Beispiel hierfür gibt ALTHOFF [1], S. 55, Aufgabe 31:

Max (Trefferwahrscheinlichkeit 0,75) und Moritz (Trefferwahrscheinlichkeit 0,6) führen einen Schießwettbewerb mit einem Fußball auf ein Loch in einer Torwand nach folgenden Regeln durch:

1. *Moritz und Max schießen abwechselnd, Moritz beginnt.*
2. *Der Wettbewerb ist beendet, wenn einer trifft oder jeder dreimal geschossen hat.*

a) *Mit welcher Wahrscheinlichkeit gewinnt Max?*

b) *Wie groß ist die Wahrscheinlichkeit, daß das Wettschießen unentschieden endet?*

c) *Geben Sie eine faire Wette an, die Max und Moritz vor dem Schießen abschließen können.*

Die Beschränkung auf drei Versuche ist dabei willkürlich und kann genau so gut durch eine beliebige (sogar unbestimmte) Anzahl (z.B. n) ersetzt werden; zur korrekten mathematischen Beschreibung werden dann nur noch endlich viele (nämlich $2n$) stochastisch unabhängige Ereignisse benötigt, deren Existenz in geeigneten endlich–diskreten Wahrscheinlichkeitsmodellen natürlich gesichert ist.

Will man trotzdem an einer Konzeption allgemeiner, unendlich–diskreter Wahrscheinlichkeitsmodelle im Unterricht festhalten — etwa im Rahmen von Leistungskursen oder speziellen Arbeitsgemeinschaften —, so kann man hier natürlich auch ohne Verwendung von Wartezeitmodellen auskommen; man denke etwa an die Approximation der Binomialverteilung (= endlich–diskret) durch die Poisson–Verteilung (= unendlich–diskret). Entsprechend könnte man die im Zusammenhang mit den obigen Wartezeitaufgaben auftretende *geometrische Verteilung* als approximative Verteilung einführen, indem man wie in der dritten Aufgabe den Versuchsumfang zunächst beschränkt und dann gegen ∞ gehen läßt — eine Vorgehensweise, die ähnlich auch in BARTH / HALLER [2], S. 225f, Aufgaben 21/22 zu finden ist. Allerdings ist es sehr wohl auch möglich, die Problematik der Nicht–Existenz gewisser unabhängiger Ereignisfolgen in diskreten Wahrscheinlichkeitsmodellen mit schulnahen Mitteln im Unterricht zu diskutieren — eine geeignete Vorgehensweise soll deshalb in dem nachfolgenden Abschnitt 2 genauer herausgearbeitet werden, wobei auch kurz auf einige weiterführende maßtheoretische Grundlagenfragen eingegangen wird. Abschnitt 3 ist der Untersuchung notwendiger und hinreichender Bedingungen für die Existenz *beliebiger* unabhängiger Ereignisfolgen in diskreten Wahrscheinlichkeitsmodellen gewidmet; Abschnitt 4 soll schließlich anhand einfacher Beispiele z.B. des Ziehens von Kugeln aus einer Urne die vorher diskutierten Kriterien anschaulich verdeutlichen sowie auf einige scheinbare Paradoxa in diesem Zusammenhang aufmerksam machen.

2. Wieviele unabhängige gleichwahrscheinliche Ereignisse gibt es in diskreten Wahrscheinlichkeitsmodellen?

Es ist unmittelbar einsichtig, daß in einem *endlich–diskreten* Wahrscheinlichkeitsmodell höchstens endlich viele stochastisch unabhängige Ereignisse existieren können: die Potenzmenge einer endlichen Menge — und damit sogar die Menge *aller* Ereignisse — ist ja selbst stets endlich (vgl. hierzu auch den ausführlichen Artikel über die Anzahl unabhängiger Ereignisse von EISENBERG / GHOSH [5]). Weniger einsichtig ist die Tatsache, daß diese Aussage für *gleichwahrscheinliche* unabhängige Ereignisse mit einer Eintrittswahrscheinlichkeit $p \in (0,1)$ sogar in *unendlich–diskreten* Wahrscheinlichkeitsmodellen gilt. (Der Fall $P(A) = p = 0$ bzw. $P(A) = p = 1$ läßt offensichtlich die triviale Ausnahmefolge $A, A, A, \ldots$ zu.) Genauer gilt hier:

Feststellung 1: In einem diskreten Wahrscheinlichkeitsmodell (Ω, P) gibt es stets nur *endlich viele* stochastisch unabhängige Ereignisse mit derselben Eintrittswahrscheinlichkeit p $(0 < p < 1)$.

Ihre Anzahl ist beschränkt durch die endliche Größe

$$N := \frac{\log\big(\max\{P(\{\omega\}) \mid \omega \in \Omega\}\big)}{\log\big(\max\{p, 1-p\}\big)}. \tag{2}$$

(Dabei bezeiche log den Logarithmus zu einer beliebigen Basis größer 1, also z.B. den Zehnerlogarithmus oder den natürlichen Logarithmus zur Basis e.)

Zum Nachweis der Richtigkeit dieser Aussage verwenden wir eine Modifikation des Beweises einer analogen Aussage in PLACHKY / BARINGHAUS / SCHMITZ [8], S. 174f.

Beweis: $A_1, \ldots, A_n$ seien stochastisch unabhängige Ereignisse in dem diskreten Wahrscheinlichkeitsmodell (Ω, P) mit $0 < P(A_i) = p < 1,\ i = 1, 2, \ldots, n$. Offensichtlich ist jedes der 2^n Produkte auf der rechten Seite von Gleichung (1) durch die Zahl r^n mit $r := \max\{p, 1-p\}$ beschränkt, da stets entweder $P(B_i) = p \leq r$ oder $P(B_i) = 1-p \leq r$ für $i = 1, 2, \ldots, n$ gilt. Sei nun $\omega \in \Omega$ beliebig gewählt. Für jedes $i = 1, 2, \ldots, n$ gilt dann entweder $\omega \in A_i$ oder $\omega \in \overline{A_i}$, d.h. ω liegt in genau einer der 2^n Schnittmengen, die in der linken Seite von Formel (1) auftreten. Es folgt also $P(\{\omega\}) \leq r^n$ für alle $\omega \in \Omega$ aufgrund der Monotonie–Eigenschaft von Wahrscheinlichkeiten (vgl. BOSCH / WOLF [3], S. 35, Satz 3; LAUTER / RÜDIGER [8], S 80, Satz 5.6) und damit auch $s := \sup\{P(\{\omega\}) \mid \omega \in \Omega\} = \max\{P(\{\omega\}) \mid \omega \in \Omega\} \leq r^n$. (Wenn Ω eine endliche Menge ist, ist dies trivial; ist Ω abzählbar–unendlich, so folgt die Gleichheit von Supremum und Maximum aus der Tatsache, daß wegen der Konvergenz der Reihe $1 = \sum_{\omega \in \Omega} P(\{\omega\})$ die Folge $\{P(\{\omega\})\}_{\omega \in \Omega}$ eine Nullfolge ist.) Ferner muß $0 < s < 1$ gelten: die erste Ungleichung wegen $\sum_{\omega \in \Omega} P(\{\omega\}) = 1$, die zweite wegen $r < 1$ und damit auch $r^n < 1$. Durch Logarithmieren gelangt man also zu der Ungleichung $\log s \leq n \cdot \log r$ und somit — wegen $\log r < 0$ — zu $n \leq \dfrac{\log s}{\log r} = N$, wie behauptet. ■

Die beiden eingangs vorgestellten Aufgaben lassen sich also konsequenterweise tatsächlich nicht korrekt in dem diskreten Rahmen, wie er in der Schule üblich (und auch nur möglich) ist, behandeln.

Wir wollen an dieser Stelle kurz auf einige maßtheoretische Grundlagenfragen zu sprechen kommen, auf die man unweigerlich stößt, wenn man das Problem in einem allgemeineren Rahmen doch noch zu lösen versucht. Zunächst einmal benötigt man offensichtlich eine *überabzählbare* Grundmenge Ω zur Konstruktion gewisser unabhängiger Ereignisfolgen. Die Ausführungen zum Abschnitt *Zufallszahlen* in BARTH / HALLER [2], S. 51 zeigen dabei im Prinzip, wie z.B. die Menge $\Omega = [0,1]$ in natürlicher Weise mit dem Problem der Existenz solcher Folgen unabhängiger Ereignisse zusammenhängt: definiert man nämlich für eine Folge $\{A_n\}_{n\in\mathbf{N}}$ stochastisch unabhängiger Ereignisse mit Eintrittswahrscheinlichkeit $P(A_n) = p$ die Zufallsgrößen

$$X_n(\omega) = \begin{cases} 1, & \text{falls } \omega \in A_n \\ 0, & \text{falls } \omega \notin A_n, \end{cases} \tag{3}$$

so erhält man eine *Bernoulli–Kette* $\{X_n\}_{n\in\mathbf{N}}$ mit Parameter p; denkt man sich deren Realisationen als den Nachkomma–Anteil der Dualzahldarstellung einer reellen Zahl $Z(\omega)$, also formal

$$Z(\omega) = \sum_{n=1}^{\infty} \frac{X_n(\omega)}{2^n},$$

so ist offensichtlich, daß jede so gewonnene Zahl $Z(\omega)$ im Intervall [0,1] liegt, wobei der rechte Endpunkt 1 erreicht wird, falls *alle* Ereignisse eintreten, und der linke Endpunkt 0, falls *keines* der Ereignisse eintritt. Da umgekehrt jede reelle Zahl $Z(\omega)$ im Intervall [0,1] eine entsprechende Dualzahldarstellung besitzt (eventuell mit unendlich vielen Nullen), gehört zu jeder solchen Zahl $Z(\omega)$ auch eine mögliche Realisation der Bernoulli–Kette. Für den Fall $p = 1/2$, den man durch wiederholtes Werfen einer Laplace–Münze realisieren kann (vgl. hierzu auch LAKOMA [6], S. 16), kann man auf diese Weise über die Werte der Zufallsgröße Z eine *Gleichverteilung* über dem Intervall [0,1] simulieren (vgl. BOSCH / WOLF [4], S. 56ff, Abschnitt 3.4, dort "geometrische Wahrscheinlichkeit" genannt; siehe auch BILLINGSLEY [3], S. 3ff). Umgekehrt erzeugt das "Ziehen" einer reellen Zahl Z aus [0,1] nach einer Gleichverteilung durch ihren Nachkomma–Anteil eine unabhängige Folge von Einsen und Nullen (Bernoulli–Kette mit Parameter $p = 1/2$).

Geeignete Grenzwertbetrachtungen zeigen, daß für eine solche "Gleichverteilung" über dem Intervall [0,1] — bezeichnen wir sie mit Q — jedenfalls die Eigenschaft

$$Q\big((a,b]\big) = b - a, \quad 0 \leq a < b \leq 1, \tag{4}$$

gelten muß (vgl. BOSCH / WOLF [4], S. 56, Definition 2). Es erhebt sich dann konsequenterweise die Frage, ob wie in den diskreten Wahrscheinlichkeitsmodellen auf der ganzen Potenzmenge $\mathcal{P}(\Omega)$ von $\Omega = [0,1]$ eine Wahrscheinlichkeitsverteilung existiert (im Sinne der Kolomogoroff'schen Axiomatik), die speziell die Eigenschaft (4) besitzt. Leider läßt sich diese Frage ohne weitere axiomatische Annahmen der (nicht–naiven) Mengenlehre nicht beantworten; unter Hinzuziehung der sogenannten Kontinuumshypothese *("Jede überabzählbare Teilmenge von* **R** *besitzt dieselbe Mächtigkeit wie* **R**") läßt sich jedoch beweisen, daß eine solche Wahrscheinlichkeitsverteilung *nicht* existieren kann (vgl. dazu PLACHKY [9], S. 5ff). Will man dennoch sinnvoll Wahrscheinlichkeitsrechnung in derartigen Situationen betreiben, muß man zwangsläufig den Definitionsbereich der Wahrscheinlichkeitsverteilung geeignet verkleinern — und damit allgemeiner σ–Algebren zur axiomatischen Modellbildung in Kauf nehmen.

Das anschaulich so einfache Beispiel des potentiell unbegrenzten Münzwurfs wie in den beiden ersten Aufgaben entpuppt sich also hier letztendlich als eine wesentliche Motivation zur konsequenten Einführung der Kolmogoroff'schen Axiomatik mit all ihren Folgeerscheinungen — und ist damit wesentlich überzeugender als die üblichen Begründungsversuche in Schulbüchern (z.B. ALTHOFF [1], S. 30; BARTH / HALLER [2], S. 80; BOSCH / WOLF [4], S. 34; LAUTER / RÜDIGER [7], S. 78), abgesehen davon, daß in diskreten Wahrscheinlichkeitsmodellen eine solch abstrakte Axiomatik gar nicht zwingend notwendig ist (man definiere etwa eine Wahrscheinlichkeitsverteilung vermöge einer Festlegung von Elementarwahrscheinlichkeiten $p_\omega \geq 0$ $(\omega \in \Omega)$ mit der Eigenschaft $\sum_{\omega\in\Omega} p_\omega = 1$ und dann allgemeiner $P(A) = \sum_{\omega\in A} p_\omega$ $(A \subseteq \Omega)$).

Ich sehe hier insgesamt einen didaktischen Widerspruch: auf der einen Seite werden wesentliche Aspekte der mathematischen Modellbildung aus den Lehrplänen ausgeklammert, obwohl man sie — wie die obigen Ausführungen zeigen — wenigstens zum Teil durchaus im schulischen Unterricht behandeln kann, auf der anderen Seite wird fast durchweg eine Axiomatik propagiert, die ihre Berechtigung eigentlich nur aus den Widersprüchlichkeiten gerade dieser Modellbildung bezieht ...

3. Zur Existenz beliebiger unabhängiger Ereignisse in diskreten Wahrscheinlichkeitsmodellen

Wie steht es nun überhaupt mit der Existenz von Folgen nicht–trivialer unabhängiger Ereignisse in diskreten Wahrscheinlichkeitsmodellen? Wenn es solche gibt, können sie — bis auf den Fall 0 oder 1 — nach der obigen Feststellung jedenfalls nicht dieselbe Eintrittswahrscheinlichkeit besitzen. Eine geringfügige Modifikation der obigen Argumentation liefert denn auch leicht eine *notwendige* Bedingung für die Existenz solcher Ereignisfolgen:

Feststellung 2: In einem diskreten Wahrscheinlichkeitsmodell gibt es höchstens dann eine Folge $\{A_n\}_{n\in\mathbf{N}}$ stochastisch unabhängiger Ereignisse, wenn für die zugehörigen Eintrittswahrscheinlichkeiten $p_n = P(A_n)$ gilt:

$$\sum_{n=1}^{\infty} \min\{p_n, 1-p_n\} < \infty. \tag{5}$$

Beweis: Zum Nachweis der Richtigkeit dieser Feststellung können wir uns auf den Fall $0 < p_n < 1$ beschränken. Bezeichnen wir jetzt $r_i = \max\{p_i, 1-p_i\}$, $i = 1, 2, \ldots, n$ und wieder $s = \max\{P(\{\omega\}) \mid \omega \in \Omega\}$, so zeigt dieselbe Rechnung wie bei Feststellung 1, daß diesmal $0 < s \leq \prod_{i=1}^{n} r_i$ für alle natürlichen Zahlen n gelten muß. Eine erneute Anwendung des (natürlichen) Logarithmus führt also zu der Ungleichung

$$-\log\left(\prod_{i=1}^{n} r_i\right) = -\sum_{i=1}^{n} \log r_i \leq -\log s < \infty. \tag{6}$$

Man hat jetzt nur noch zu beachten, daß für den natürlichen Logarithmus die Ungleichung $-\log r_i \geq 1 - r_i = \min\{1-p_i, p_i\}$ gilt (man betrachte etwa die Tangente im Punkt 1 an die Funktion $f(x) = -\log x$); mit (6) erhält man hieraus sofort auch (5). ∎

Erstaunlicherweise ist die Bedingung (5) auch *hinreichend* für die Modellierbarkeit stochastisch unabhängiger Ereignisfolgen in diskreten Wahrscheinlichkeitsmodellen. Allerdings ist die zugehörige Konstruktion nicht ganz so einfach wie die obige Argumentation. Sie beruht darauf, daß mit Wahrscheinlichkeit 1 unter (5) in der Folge der Ereignisse $\{A_n\}_{n\in\mathbf{N}}$ entweder die Ereignisse selbst oder deren Komplemente nur *endlich oft* eintreten (der interessierte Leser sei auf die Ausführungen in BILLINGSLEY [3], S. 581, Nr. 4.19, verwiesen). Die letztgenannte Tatsache läßt sich sogar anschaulich noch relativ einfach wie folgt verstehen: Betrachtet man lediglich Ereignisse mit Eintrittswahrscheinlichkeiten $p_n \leq 1/2$ — dies ist keine wesentliche Einschränkung, da man in unabhängigen Ereignisfolgen A_n auch durch $\overline{A_n}$ ersetzen darf, ohne die Unabhängigkeit zu verletzen — so ist der Wert der Reihe in (5) gerade gleich dem Erwartungswert der (zufälligen) Anzahl S der innerhalb der Folge eingetretenen Ereignisse. Mit der im vorigen Abschnitt in (3) gewählten Konstruktion von Bernoulli–Ketten ergibt sich nämlich die Darstellung $S = \sum_{n=1}^{\infty} X_n$, wobei wieder $X_n = 1$ oder 0 gewählt wird je nachdem, ob das Ereignis A_n eintritt oder nicht. Es folgt demnach

$$E(S) = \sum_{n=1}^{\infty} E(X_n) = \sum_{n=1}^{\infty} P(A_n) = \sum_{n=1}^{\infty} p_n < \infty,$$

wie behauptet. Da im Mittel also nur endlich viele der Ereignisse eintreten, ist entsprechend auch die Anzahl S der eintretenden Ereignisse selbst — jedenfalls mit Wahrscheinlichkeit 1 — endlich.

4. Zusammenfassung und Ausblick

In diesem Abschnitt wollen wir uns abschließend mit einigen einfachen Beispielen beschäftigen, die die eher theoretischen Ausführungen der beiden vorangegangenen Abschnitte veranschaulichen sollen.

Betrachten wir dazu einmal das folgende einfache Experiment: In einer Urne befindet sich zu Beginn je eine weiße und eine schwarze Kugel. Nach der n–ten Ziehung einer Kugel aus der Urne wird diese mit noch weiteren $c_n \geq 0$ weißen Kugeln in die Urne zurückgelegt. Gefragt wird nach der Nummer derjenigen Ziehung, bei der erstmalig die schwarze (alternativ: eine weiße) Kugel erscheint.

Kann dieses Experiment im Rahmen diskreter Wahrscheinlichkeitsmodelle behandelt werden? Die vorigen Ausführungen zeigen, daß die Beantwortung dieser Frage mit dem Wachstumsverhalten der neu hinzukommenden Kugeln in der Urne zusammenhängt — und zwar ausschließlich mit diesem. Bezeichnet man mit A_n das Ereignis, daß im n–ten Zug die schwarze Kugel gezogen wird, so erhält man leicht

$$P(A_n) = p_n = \frac{1}{\sum_{i=0}^{n-1} c_i}, \tag{7}$$

wobei abkürzend $c_0 = 2$ gewählt wurde. Wir wollen hier exemplarisch einmal drei Fälle betrachten:

1. $c_n = 0$: Dies ist die in Abschnitt 3 betrachtete Situation mit $p_n = p = 1/2$. Eine Modellierung im diskreten Rahmen ist nicht möglich.

2. $c_n = 1$: Es ist $P(A_n) = p_n = \frac{1}{n+1}$. Die Reihe $\sum_{n=1}^{\infty} p_n$ ist *divergent*; eine Modellierung im diskreten Rahmen bleibt unmöglich.

3. $c_n = 2n + 1$: Hier ergibt sich $P(A_n) = p_n = \frac{1}{n^2+1}$. Die Reihe $\sum_{n=1}^{\infty} p_n$ ist *konvergent*; eine Modellierung im diskreten Rahmen ist also möglich.

Trotz der vordergründigen Gleichartigkeit der Experimente (Ziehen von Kugeln aus einer Urne) unterscheiden sich die zugehörigen stochastischen Modelle also sehr wesentlich voneinander: bei zu "langsamer" Vermehrung der Kugeln ist eine diskrete Modellierung nicht möglich, bei genügend "schneller" Vermehrung dagegen sehr wohl! Bezogen auf die Aufgabenbeispiele in Abschnitt 1 bedeutet dies, daß z.B. die folgenden Modifikationen durchaus in einem diskreten Wahrscheinlichkeitsmodell (und damit mit schulischen Mitteln korrekt) behandelbar sind:

Drei Spieler A,B,C werfen abwechselnd in dieser Reihenfolge Münzen, und zwar beim ersten Wurf zwei, dann drei gleichzeitig, dann vier gleichzeitig usw. Gewonnen hat derjenige, in dessen Wurf zuerst mindestens zweimal Wappen erscheint. Mit welcher Wahrscheinlichkeit gewinnt A (B;C)?

Laplace–Münzen werden nach dem folgenden Schema geworfen: im ersten Wurf zwei, dann drei gleichzeitig, dann vier gleichzeitig usw., bis wenigstens zwei von *ihnen gleichzeitig Adler zeigen. Die Zufallsgröße X sei die Anzahl der dazu nötigen Würfe.*

a) *Gib einen gröberen Ergebnisraum an. Bestimme die Wahrscheinlichkeitsverteilung von X.*

b) *Berechne $E(X)$ und* $\operatorname{Var}(X)$.

Eine korrekte Lösung etwa der ersten dieser beiden Aufgaben gestaltet sich dann z.B. so:

Bezeichnet man mit A_n das Ereignis, daß beim n–ten Wurf mindestens zweimal Wappen erscheint und mit $\overline{A_n}$ das dazu komplementäre Ereignis, sowie mit B_n das Ereignis, daß im n–ten Wurf *erstmalig* Wappen erscheint, also

$$B_n = \begin{cases} A_1, & n = 1 \\ \overline{A_1} \cap \ldots \cap \overline{A_{n-1}} \cap A_n, & n \geq 2, \end{cases}$$

so lassen sich die Ereignisse G_A, G_B, G_C, daß Spieler A,B,C gewinnt, mengentheoretisch wieder ausdrücken durch

$$G_A = \bigcup_{n=0}^{\infty} B_{3n+1}, \quad G_B = \bigcup_{n=0}^{\infty} B_{3n+2}, \quad G_C = \bigcup_{n=0}^{\infty} B_{3n+3},$$

da jeder der Spieler nur jedes dritte Mal an die Reihe kommt und nur dann gewinnt, wenn vor ihm kein anderer Spieler bereits mindestens zweimal Wappen in einem Wurf erzielt hat (d.h. z.B. Spieler A kann wieder nur in den Runden 1, 4,7,11,... gewinnen). Das Ereignis $\overline{A_n}$ tritt dabei genau dann ein, wenn keine oder genau eine der $n+1$ Münzen Wappen zeigt; die Wahrscheinlichkeit hierfür beträgt also $1 - p_n = P\big(\overline{A_n}\big) = (n+2)2^{-(n+1)}$ mit $\sum_{n=1}^{\infty}(1-p_n) = \sum_{n=1}^{\infty}(n+2)2^{-(n+1)} = 2 < \infty$. (Im Mittel tritt im Verlauf des unbegrenzten Münzwurfs dieser Fall also genau zwei Mal auf.) Wegen der Unabhängigkeit der einzelnen Würfe erhält man somit

$$P(B_n) = P(\overline{A_1}) \cdot \ldots \cdot P(\overline{A_{n-1}}) \cdot P(A_n) = \frac{(n+1)!}{2^{n(n+1)/2}} - \frac{(n+2)!}{2^{(n+1)(n+2)/2}} \quad (n \geq 2);$$
$$P(B_1) = 1/4.$$

Die Additivitätseigenschaft von P ergibt damit insgesamt (numerische Näherungswerte)

$$P(G_A) = \sum_{n=0}^{\infty} P(B_{3n+1}) = 0.34536,$$
$$P(G_B) = \sum_{n=0}^{\infty} P(B_{3n+2}) = 0.39457,$$
$$P(G_C) = \sum_{n=0}^{\infty} P(B_{3n+3}) = 0.26007.$$

Man beachte, daß bei dieser Modifikation Spieler B die größeren Gewinnchancen hat! Für diejenigen, die diese Ergebnisse praktisch überprüfen möchten, empfehle ich einmal das folgende BASIC–Programm auszuprobieren (1=Wappen, 0=Zahl):

```
10 CLS
20 DIM X(16)
30 DIM G(3)
40 RANDOMIZE TIMER
50 M=1
60 I=0
70 FOR N=2 TO 16
80 S=0
90 PRINT "Wurf Nr.";N-1;":  ";
100 FOR K=1 TO N
110 X(K)=INT(RND+.5)
120 S=S+X(K)
130 PRINT X(K);" ";
140 NEXT K
150 IF S>=2 THEN GOTO 180
160 PRINT
170 NEXT N
180 Z=N MOD 3
190 ON Z GOTO 270,210
200 GOTO 240
210 A$="A"
220 G(1)=G(1)+1
230 GOTO 290
240 A$="B"
250 G(2)=G(2)+1
260 GOTO 290
270 A$="C"
280 G(3)=G(3)+1
290 PRINT
300 PRINT "Spieler ";A$;" gewinnt!"
310 PRINT
320 PRINT
330 PRINT TAB(1);"A: ";G(1)/M; TAB(20);"B: ";G(2)/M;
335 PRINT TAB(40);"C: ";G(3)/M;TAB(60);M;" Spiele"
                                        (relative Gewinnhäufigkeiten)
340 PRINT
350 M=M+1
360 B$=INKEY$
370 IF B$="q" THEN SYSTEM                            (Abbruch)
380 IF I=1 THEN GOTO 410
390 IF B$="r" THEN I=1                       (Non–Stop–Betrieb)
```

```
400 IF B$="" THEN GOTO 360
410 IF B$="s" THEN I=0                    (Einzelschritt–Betrieb)
420 GOTO 70
```

(Gute Simulationsergebnisse werden bereits bei Spielanzahlen von `M` ab etwa 1000 erzielt.)

Wir wollen abschließend noch einmal kurz auf den Fall 2 im Beispiel mit dem Ziehen der Kugeln aus der Urne zurückkommen. Es ergibt sich nämlich hier die scheinbar paradoxe Situation, daß zwar die schwarze Kugel wegen der Divergenz der Reihe in (4) im Mittel unendlich oft gezogen wird, die Wahrscheinlichkeit für eine Ziehung aber gegen Null strebt! Für die Verteilung des Zeitpunkts Y des ersten Ziehens der schwarzen Kugel erhält man hier übrigens:

$$P(Y=n) = p_n \prod_{i=1}^{n-1}(1-p_i) = \frac{1}{n+1}\prod_{i=1}^{n-1}\frac{i}{i+1} = \frac{1}{n(n+1)}, \quad n \in \mathbf{N},$$

mit dem Erwartungswert

$$E(Y) = \sum_{n=1}^{\infty} n \cdot \frac{1}{n(n+1)} = \sum_{n=1}^{\infty}\frac{1}{n+1} = \infty;$$

d.h. man muß im Mittel unendlich lange warten, bis die schwarze Kugel überhaupt gezogen wird, obwohl sie doch nach der vorherigen Aussage im Verlauf des Experiments sogar unendlich häufig zu erwarten ist!

Welche Konsequenzen kann und soll man aus all diesen seltsamen Phänomenen für den Unterricht ziehen? Maßtheoretische Argumentationen wird man dort wohl kaum ausführlich einbeziehen können, deshalb bieten sich m.E. hier grundsätzlich nur zwei verschiedene, didaktisch sinnvolle und mathematisch vertretbare Alternativen an:

a) Verzicht auf die Behandlung unabhängiger Experimente mit potentiell unbegrenztem Versuchsumfang; Betrachtung unendlich–diskreter Verteilungen nur als praktisch brauchbare Approximationen endlich–diskreter Verteilungen.

b) Diskussion unabhängiger Experimente mit potentiell unbegrenztem Versuchsumfang nur bei gleichzeitigem Hinweis auf deren Nicht–Modellierbarkeit im diskreten Rahmen; Ableitung formaler "Lösungen" auf der Basis der Kolmogoroff'schen Axiomatik im Bewußtsein der Tatsache, daß durch Einbettung in einen geeigneten Meßraum (=Grundmenge versehen mit einer σ–Algebra) diese auch tatsächlich exakt sind.

Die letztere Vorgehensweise scheint mir allerdings geeigneter, etwas mehr das kritische Bewußtsein dafür zu schulen, daß nicht alle anschaulich einfachen (nicht nur Zufalls–)Vorgänge auch entsprechend einfach mathematisch zu beschreiben bzw.

zu modellieren sind. Gerade in der heutigen Zeit, wo angesichts fortschreitender Verbreitung von Computern immer häufiger die Bedeutung der Mathematik im allgemeinen öffentlich diskutiert wird, sollte man m.E. stärker auch in der Schule mathematische Grundlagenfragen behandeln, sofern dies inhaltlich möglich ist. Eine Überbetonung formalistischer Aspekte von Mathematik halte ich demgegenüber für wenig hilfreich, solange nicht in ausreichendem Maße zugleich auch inhaltliche Begründungen für deren Notwendigkeit gegeben werden.

5. Literatur

[1] ALTHOFF, H. (1985): *Wahrscheinlichkeitsrechnung und Statistik.* J.B. Metzler, Stuttgart.

[2] BARTH, F. und HALLER, R. (1985): *Stochastik Leistungskurs.* 3. Aufl., Ehrenwirth, München.

[3] BILLINGSLEY, P. (1986): *Probability and Measure.* 2. Aufl., Wiley, N.Y.

[4] BOSCH, K. und WOLFF, H. (1980): *Leistungskurs Wahrscheinlichkeitsrechnung und Statistik.* Westermann, Braunschweig.

[5] EISENBERG, B. und GHOSH, B.K. (1987): *Independent events in a discrete uniform probability space.* The American Statistician, Vol. 41, No. 1.

[6] LAKOMA, E. (1990): *Lokale Modelle im Stochastik–Unterricht.* Stochastik in der Schule, Heft 3, 8 – 21.

[7] LAUTER, J. und RÜDIGER, K. (1988): *Mathematik Sekundarstufe II. Wahrscheinlichkeitsrechnung und Statistik.* Cornelsen – Schwann, Düsseldorf.

[8] PLACHKY, D., BARINGHAUS, L. und SCHMITZ, N. (1983): *Stochastik I.* Eine elementare Einführung in Grundbegriffe der Wahrscheinlichkeitsrechnung und Statistik. 2. Aufl., Akademische Verlagsgesellschaft, Wiesbaden.

[9] PLACHKY, D. (1981): *Stochastik II.* Eine maßtheoretische Einführung in Grundbegriffe der Wahrscheinlichkeitstheorie und mathematischen Statistik. Akademische Verlagsgesellschaft, Wiesbaden.