

Wie in der Einleitung erwähnt, nehmen wir an, dass der Leser mit der elementaren Wahrscheinlichkeitstheorie (s. z. B. das Lehrbuch von U. Krengel [10]) vertraut ist, auch wenn wir ihre Kenntnisse ab dem nächsten Kapitel nicht notwendig voraussetzen. Konkret bedeutet das additive Wahrscheinlichkeiten (s. Satz 1.2), speziell diskrete Modelle.

Wir wollen die Notwendigkeit, Wahrscheinlichkeitstheorie auf der Grundlage der Maßtheorie zu entwickeln, um gewisse Probleme behandeln zu können, am Beispiel des unendlichen Münzwurfs motivieren. Dies ist in gewissem Sinne das einfachste derartige Beispiel und von grundlegender Bedeutung. An ihm zeigen wir, wie weit man in der Wahrscheinlichkeitstheorie mit elementaren Methoden kommt, und überlegen uns, was für weitergehende Aussagen, in diesem Fall über das Konvergenzverhalten, benötigt wird. Schließlich stellen wir eine Beziehung her zu der gleichmäßigen Verteilung auf $[0,1]$, ein ebenfalls grundlegendes Beispiel.

Zur Konstruktion dieses Modells gehen wir von den bekannten diskreten Modellen für jeweils eine feste Anzahl von unabhängigen Würfeln einer idealen Münze aus.

Für $m \geq 1$ sei

$$\Omega_m = \{0,1\}^m \quad \text{und} \quad P_m(A_m) = \frac{|A_m|}{2^m} \quad \text{für } A_m \subset \Omega_m \quad (1.1)$$

mit „1“ stellvertretend für „Zahl“ und „0“ für „Wappen“.

Bei manchen Fragestellungen wie zum Beispiel der Verteilung der Anzahl der Würfe bis zum ersten Wurf mit Ausgang „Zahl“ oder dem Ruinproblem (s. Anfang von Kap. 10), ist die Anzahl der Würfe zwar endlich, hängt aber selbst vom Zufall ab. Für die Behandlung solcher Probleme stellen wir ein elementares Modell auf, das wir später erweitern werden.

Als Grundraum nehmen wir die Menge aller unendlichen 0-1-Folgen, also

$$\Omega = \{\omega = (\omega_n)_{n \geq 1} : \omega_n \in \{0,1\} \quad \text{für } n \geq 1\} = \{0,1\}^{\mathbb{N}}.$$

Bekanntlich ist Ω überabzählbar. Um Teilmengen von Ω eine Wahrscheinlichkeit zuzuordnen, die unabhängigen Würfeln einer idealen Münze entspricht, gehen wir von solchen

Teilmengen aus, die nur von endlich vielen Würfeln abhängen, wie zum Beispiel „unter den ersten 5 Würfeln tritt 1 höchstens 3-mal auf oder „der 3. und 6. Wurf hat das gleiche Ergebnis“. Solche Mengen nennt man Zylindermengen. Exakt lassen sie sich folgendermaßen definieren.

► **Definition 1.1** Eine Teilmenge $A \subset \{0,1\}^N$ heißt eine *Zylindermenge*, wenn ein $m \geq 1$ und eine Menge $A_m \subset \{0,1\}^m$ existiert, so dass $\{A = (\omega_n)_{n \geq 1} \in \{0,1\}^N : (\omega_1, \dots, \omega_m) \in A_m\}$ ist. Das System aller Zylindermengen werde mit $\mathcal{Z}$ bezeichnet.

Mit Hilfe der Projektionen auf die jeweils ersten m Komponenten

$$\pi_m: \{0,1\}^N \rightarrow \{0,1\}^m \quad \text{für } m \geq 1, \text{ definiert durch } \pi_m((\omega_n)_{n \geq 1}) = (\omega_1, \dots, \omega_m)$$

lassen sich Zylindermengen geeignet darstellen in der Form

$$\{(\omega_n)_{n \geq 1} \in \Omega : (\omega_1, \dots, \omega_m) \in A_m\} = \pi_m^{-1}(A_m).$$

Die Projektionen π_m ($m \geq 1$) sind surjektiv, wie man leicht durch eine beliebige Ergänzung eines gegebenen Elements aus $\{0,1\}^m$ zu einer unendlichen Folge sieht. Die Surjektivität wird sich im Folgenden als wichtig erweisen.

Den Zylindermengen ordnen wir die Wahrscheinlichkeiten zu, die endlich vielen Würfeln entspricht. Wir definieren daher auf $\mathcal{Z}$ die Wahrscheinlichkeit P durch

$$P(\pi_m^{-1}(A_m)) = P_m(A_m) \tag{1.2}$$

mit den durch (1.1) definierten Wahrscheinlichkeiten P_m ($m > 1$).

Da sich Zylindermengen auf verschiedene Weisen in der Form $\pi_m^{-1}(A_m)$ darstellen lassen, müssen wir zeigen, dass diese Definition eindeutig ist. Ferner beweisen wir, dass $\mathcal{Z}$ eine Algebra und P eine additive Wahrscheinlichkeit auf $\mathcal{Z}$ ist.

► **Satz 1.2**

1. $\mathcal{Z}$ ist eine Algebra, d. h. es gilt
 - i) $\Omega \in \mathcal{Z}$
 - ii) $A \in \mathcal{Z} \Rightarrow A^c \in \mathcal{Z}$
 - iii) $A, B \in \mathcal{Z} \Rightarrow A \cup B \in \mathcal{Z}$
2. P ist durch (1.2) auf $\mathcal{Z}$ eindeutig definiert und eine additive Wahrscheinlichkeit, d. h. es gilt
 - i) $0 \leq P(A) \leq 1$ für $A \in \mathcal{Z}$, $P(\Omega) = 1$
 - ii) $P(A \cup B) = P(A) + P(B)$ für disjunkte $A, B \in \mathcal{Z}$.

Zum Beweis der Eindeutigkeit überlegen wir uns, inwieweit die Darstellung von Zylindermengen auf verschiedene Weisen möglich ist. Das folgende Ergebnis werden wir auch für andere Zwecke benötigen.

► **Lemma 1.3** Für $A_m \subset \{0, 1\}^m$ und $B_k \subset \{0, 1\}^k$ mit $m \geq k$ gilt:

$$\pi_m^{-1}(A_m) = \pi_k^{-1}(B_k) \Leftrightarrow A_m = B_k \times \{0, 1\}^{m-k}.$$

Beweis: Wir behandeln zunächst den Spezialfall $m = k$, d. h. wir zeigen:

$$\pi_m^{-1}(A_m) = \pi_m^{-1}(B_m) \Leftrightarrow A_m = B_m.$$

Die Richtung $\Leftarrow$ ist in diesem Fall trivial. Zum Beweis von $\Rightarrow$ zeigen wir allgemein:

► **Lemma 1.4** Für beliebige Mengen M, N und Abbildungen $f: M \rightarrow N$ gilt

1. $f(f^{-1}(A)) \subset A$ für alle Mengen $A \subset N$
2. Für surjektive Abbildungen f ist $f(f^{-1}(A)) = A$ für alle Mengen $A \subset N$.

Beweis

1. $x \in f^{-1}(A)$ bedeutet $f(x) \in A$. Daraus folgt $f(f^{-1}(A)) \subset A$.
2. Sei $y \in A$. Wegen der Surjektivität von f existiert ein $x \in M$ mit $f(x) = y$. Da $y \in A$ ist, ist $x \in f^{-1}(A)$ und damit $y = f(x) \in f(f^{-1}(A))$. Es folgt $A \subset f(f^{-1}(A))$ und mit 1. die Gleichheit.

Durch Anwendung von Lemma 1.4.2 auf die surjektiven Abbildungen π_m folgt $A_m = \pi_m(\pi_m^{-1}(A_m))$ und $B_m = \pi_m(\pi_m^{-1}(B_m))$ und damit Lemma 1.3 für den Fall $m = k$.

Für beliebiges $m \geq k$ und $B_k \subset \{0, 1\}^k$ zeigen wir

$$\pi_k^{-1}(B_k) = \pi_m^{-1}(B_k \times \{0, 1\}^{m-k}). \quad (1.3)$$

Mit dem soeben bewiesenen Spezialfall folgt daraus Lemma 1.3.

Nach Definition der Projektionen gilt:

$$\begin{aligned} (\omega_n)_{n \geq 1} \in \pi_k^{-1}(B_k) &\Leftrightarrow (\omega_1, \dots, \omega_k) \in B_k \\ (\omega_n)_{n \geq 1} \in \pi_m^{-1}(B_k \times \{0, 1\}^{m-k}) &\Leftrightarrow (\omega_1, \dots, \omega_k, \omega_{k+1}, \dots, \omega_m) \in B_k \times \{0, 1\}^{m-k}. \end{aligned}$$

Da die rechten Seiten offensichtlich äquivalent sind, folgt (1.3).

Beweis von Satz 1.2: Der Beweis beruht im wesentlichen darauf, dass das Urbild mit allen Mengenoperationen vertauscht.

1. i) $\Omega \in \mathcal{Z}$, da $\Omega = \pi_m^{-1}(\Omega_m)$ für ein beliebiges $m \geq 1$ ist.
- ii) Aus $A = \pi_m^{-1}(A_m) \in \mathcal{Z}$ folgt $A^c = (\pi_m^{-1}(A_m))^c = \pi_m^{-1}(A_m^c) \in \mathcal{Z}$.
- iii) Seien $A = \pi_m^{-1}(A_m)$, $B = \pi_k^{-1}(B_k) \in \mathcal{Z}$. Aus Symmetriegründen können wir ohne Einschränkung $m \geq k$ annehmen. Nach Lemma 1.3 lässt sich B auch als $B = \pi_m^{-1}(B_m)$ mit $B_m = B_k \times \{0, 1\}^{m-k}$ darstellen. Es folgt $A \cup B = \pi_m^{-1}(A_m) \cup \pi_m^{-1}(B_m) = \pi_m^{-1}(A_m \cup B_m) \in \mathcal{Z}$.
2. Zum Beweis der Eindeutigkeit sei $\pi_m^{-1}(A_m) = \pi_k^{-1}(B_k)$, wobei wir wieder $m \geq k$ annehmen können. Nach Lemma 1.3 ist $A_m = B_k \times \{0, 1\}^{m-k}$ und daher $|A_m| = |B_k| \cdot 2^{m-k}$. Es folgt

$$P_m(A_m) = \frac{|A_m|}{2^m} = \frac{|B_k| \cdot 2^{m-k}}{2^m} = \frac{|B_k|}{2^k} = P_k(B_k)$$

und damit die Eindeutigkeit der Definition von P .

- i) Aus $A_m \subset \Omega_m$ ($m \geq 1$) folgt $|A_m| \leq |\Omega_m| = 2^m$ und damit $0 \leq P(A) = P_m(A_m) \leq 1$ für $A = \pi_m^{-1}(A_m) \in \mathcal{Z}$ sowie $P(\Omega) = 1$ für $A_m = \Omega_m$.
- ii) Wie im Beweis von 1 iii) seien $A = \pi_m^{-1}(A_m)$, $B = \pi_k^{-1}(B_k) = \pi_m^{-1}(B_m) \in \mathcal{Z}$ mit $m \geq k$. Sind A und B paarweise disjunkt, dann sind auch A_m und B_m paarweise disjunkt. Denn aus $\pi_m^{-1}(A_m \cap B_m) = A \cap B = \emptyset$ folgt mit Lemma 1.4, dass $A_m \cap B_m = \pi_m(\pi_m^{-1}(A_m \cap B_m)) = \emptyset$ ist.

Die Additivität von P folgt damit aus der Additivität von P_m :

$$P(A \cup B) = P_m(A_m \cup B_m) = P_m(A_m) + P_m(B_m) = P(A) + P(B).$$

Bei wiederholten unabhängigen Würfeln einer idealen Münze erwartet man, dass die relative Häufigkeit der Anzahl der Würfe mit Ausgang „Zahl“ gegen $\frac{1}{2}$ konvergiert. Man kann jedoch leicht Folgen von „Zahl“ und „Wappen“ angeben, für die das nicht gilt. Es zeigt sich aber, dass diese Folgen stochastisch eine Ausnahme bilden, wobei wir präzisieren müssen, in welchem Sinne das gilt.

Bereits in diesem elementaren Modell können wir das schwache Gesetz der großen Zahlen formulieren und beweisen. Dazu definieren wir auf Ω die Zufallsvariablen, d. h. Abbildungen X_k ($k \geq 1$) durch $X_k((\omega_n)_{n \geq 1}) = \omega_k$ und $S_n = X_1 + \dots + X_n$ ($n \geq 1$). S_n ist die Anzahl der Würfe mit Ausgang „Zahl“ unter den ersten n Würfeln. Das schwache Gesetz der großen Zahlen besagt, dass für jedes $\varepsilon > 0$

$$P\left(\left|\frac{S_n}{n} - \frac{1}{2}\right| \geq \varepsilon\right) \rightarrow 0 \quad \text{für } n \rightarrow \infty \quad (1.4)$$

konvergiert.

Offensichtlich ist $\left\{\left|\frac{S_n}{n} - \frac{1}{2}\right| \geq \varepsilon\right\} \in \mathcal{Z}$ und $P\left(\left|\frac{S_n}{n} - \frac{1}{2}\right| \geq \varepsilon\right)$ daher definiert.

Zum Beweis von (1.4) schätzen wir für $P\left(\left|\frac{S_n}{n} - \frac{1}{2}\right| \geq \varepsilon\right)$ für $\varepsilon > 0$ und $n \geq 1$ mit der Tschebychev'schen Ungleichung ab.

Für Zufallsvariable X mit endlicher Varianz $V(X)$ gilt für alle $\varepsilon > 0$:

$$P(|X - EX| \geq \varepsilon) \leq \frac{V(X)}{\varepsilon^2}. \quad (1.5)$$

Da es uns in diesem Kapitel nur darauf ankommt, zu zeigen, was man mit elementaren Methoden beweisen kann, verzichten wir an dieser Stelle auf den für Zufallsvariable mit diskreter Verteilung elementaren Beweis der Tschebychev'schen Ungleichung. In Kap. 4 werden wir sie in der verallgemeinerten Form 4.43 beweisen.

Es folgt für jedes $\varepsilon > 0$

$$P\left(\left|\frac{S_n}{n} - \frac{1}{2}\right| \geq \varepsilon\right) \leq \frac{V\left(\frac{S_n}{n}\right)}{\varepsilon^2} = \frac{V(X_1)}{n\varepsilon^2} = \frac{1}{4n\varepsilon^2} \rightarrow 0 \quad (1.6)$$

für $n \rightarrow \infty$ und damit das schwache Gesetz der großen Zahlen (1.4).

Es besagt, dass die Wahrscheinlichkeit von größeren Abweichungen der Differenz der relativen Häufigkeit $\frac{S_n}{n}$ von $\frac{1}{2}$ als eine beliebig kleine positive Schranke beliebig klein wird für hinreichend großes n . Es bedeutet jedoch nicht, dass die Wahrscheinlichkeit, dass $\frac{S_n}{n}$ nicht gegen $\frac{1}{2}$ konvergiert, beliebig klein und damit gleich 0 ist. Es könnte zu zufälligen Zeiten immer seltener, aber immer mal wieder Abweichungen geben (s. Übung 1.1 bzw. Gegenbeispiel 1 zu Satz 3.17). Dass das hier jedoch nicht der Fall ist, dass also die Wahrscheinlichkeit der Ausnahmemenge $N = \left\{\frac{S_n}{n} \not\rightarrow \frac{1}{2} \text{ für } n \rightarrow \infty\right\}$ gleich 0 ist, behauptet das starke Gesetz der großen Zahlen. Das starke und schwache Gesetz der großen Zahlen unterscheiden sich in der Konvergenzart. In Kap. 3 werden wir genauer auf den Unterschied zwischen diesen Konvergenzarten eingehen. An dieser Stelle begnügen wir uns mit dem Hinweis, dass die Menge N keine Zylindermenge ist, $P(N)$ also nicht definiert ist.

Das starke Gesetz der großen Zahlen rechtfertigt die Interpretation von $\frac{1}{2}$ als Wahrscheinlichkeit und damit den Modellansatz (s. auch das allgemeine Modell von Bernoulli-Experimenten mit beliebiger Wahrscheinlichkeit (Übung 1.3)). Wir werden auf seine diesbezügliche Bedeutung in Kap. 6, in dem wir das starke Gesetz der großen Zahlen allgemein behandeln, genauer eingehen und in diesem Zusammenhang auch die Beziehung zwischen Phänomenen der Realität und entsprechenden mathematischen Modellen diskutieren.

Jetzt wollen wir versuchen, das starke Gesetz der großen Zahlen so weit wie möglich mit elementaren Methoden zu begründen und uns überlegen, was uns dann zu einem exakten Beweis noch fehlt.

Dazu überdecken wir die Ausnahmemenge N mit Mengen möglichst kleiner Wahrscheinlichkeit. Die folgende Konstruktion ist speziell auf dieses Beispiel zugeschnitten.

Im 1. Schritt zeigen wir, dass es genügt, die Teilfolge $(S_{m^2})_{m \geq 1}$ zu betrachten, indem wir für $\omega \in \Omega$ beweisen:

$$\frac{S_n(\omega)}{n} \rightarrow \frac{1}{2} \quad \text{für } n \rightarrow \infty \Leftrightarrow \frac{S_{m^2}(\omega)}{m^2} \rightarrow \frac{1}{2} \quad \text{für } m \rightarrow \infty. \quad (1.7)$$

Die Richtung $\Rightarrow$ ist als die Konvergenz einer Teilfolge trivial.

Beweis von $\Leftarrow$: Zu $n \geq 1$ existiert ein $m \geq 1$ mit $m^2 \leq n < (m+1)^2$. Mit $0 \leq n - m^2 < 2m + 1$ bzw. $0 \leq n - m^2 \leq 2m$ folgt die Richtung $\Leftarrow$ aus den folgenden Abschätzungen für $\omega = (\omega_n)_{n \geq 1} \in \Omega$:

$$\begin{aligned} \left| \frac{S_n(\omega)}{n} - \frac{S_{m^2}(\omega)}{m^2} \right| &= \left| \left(\frac{1}{n} - \frac{1}{m^2} \right) \sum_{k=1}^{m^2} \omega_k + \frac{1}{n} \sum_{k=m^2+1}^n \omega_k \right| \leq \left(\frac{1}{m^2} - \frac{1}{n} \right) m^2 + \frac{n - m^2}{n} \\ &= 2 \frac{n - m^2}{n} \leq \frac{4m}{m^2} = \frac{4}{m} \rightarrow 0 \quad \text{für } m \rightarrow \infty. \end{aligned}$$

Mit (1.7) lässt sich die Ausnahmemenge N auch als $N = \left\{ \frac{S_{m^2}}{m^2} \not\rightarrow \frac{1}{2} \text{ für } m \rightarrow \infty \right\}$ darstellen.

Im 2. Schritt konstruieren wir eine Überdeckung von N mit Mengen möglichst kleiner Wahrscheinlichkeit.

Sei dazu $(\varepsilon_m)_{m \geq 1}$ eine zunächst beliebige Folge mit $\varepsilon_m \downarrow 0$ für $m \rightarrow \infty$ die wir später geeignet wählen werden. Für $m \geq 1$ sei $N_m = \left\{ \left| \frac{S_{m^2}}{m^2} - \frac{1}{2} \right| \geq \varepsilon_m \right\}$. Es ist $N_m \in \mathcal{Z}$.

Wir zeigen:

1. Für alle $k \geq 1$ ist $N \subset \cup_{m=k}^{\infty} N_m$
2. Die Folge $(\varepsilon_m)_{m \geq 1}$ mit $\varepsilon_m \downarrow 0$ kann so gewählt werden, dass $\sum_{m=1}^{\infty} P(N_m) < \infty$ ist.

Beweis 1: Wir beweisen die Behauptung in der äquivalenten Form $N^c \supset \cap_{m=k}^{\infty} (N_m)^c$ für alle $k \geq 1$. Sei daher $k \geq 1$ und $\omega \in \cap_{m=k}^{\infty} (N_m)^c$. Nach Definition von N_m ist $\left| \frac{S_{m^2}(\omega)}{m^2} - \frac{1}{2} \right| < \varepsilon_m$ für alle $m \geq k$. Aus $\varepsilon_m \downarrow 0$ folgt $\frac{S_{m^2}(\omega)}{m^2} \rightarrow \frac{1}{2}$ und damit $\omega \in N^c$.

Beweis 2: Für $n = m^2$ liefert die Tschebychev'sche Ungleichung (1.6)

$$P(N_m) = P\left(\left|\frac{S_{m^2}}{m^2} - \frac{1}{2}\right| \geq \varepsilon_m\right) \leq \frac{1}{4m^2\varepsilon_m^2}.$$

Wählen wir zum Beispiel $\varepsilon_m = m^{-\alpha}$ mit $0 < \alpha < \frac{1}{2}$, dann konvergiert $\varepsilon_m \downarrow 0$ und es ist

$$\sum_{m=1}^{\infty} P(N_m) \leq \sum_{m=1}^{\infty} \frac{1}{4m^2\varepsilon_m^2} = \sum_{m=1}^{\infty} \frac{1}{4m^{2-2\alpha}} < \infty.$$

Mit einer solchen Folge $(\varepsilon_m)_{m \geq 1}$ und den dazu gehörenden Zylindermengen $(N_m)_{m \geq 1}$ erhalten wir zu jedem $\varepsilon > 0$ ein $k \geq 1$ mit $N \subset \cup_{m=k}^{\infty} N_m$, so dass $\sum_{m=k}^{\infty} P(N_m) \leq \varepsilon$ ist.

Wir können damit die Ausnahmemenge N für jedes $\varepsilon > 0$ mit abzählbar vielen Zylindermengen überdecken, deren Wahrscheinlichkeiten eine Gesamtsumme $\leq \varepsilon$ hat. Hat damit N Wahrscheinlichkeit 0 und haben wir damit das starke Gesetz der großen Zahlen bewiesen? Natürlich nicht! Denn zunächst einmal ist $P(N)$ nach wie vor nicht definiert. Wir müssen P also als erstes auf eine größere Klasse, die N enthält, fortsetzen. N lässt sich leicht

aus Zylindermengen mit Hilfe abzählbarer Vereinigungen und Durchschnitten darstellen (Übung 1.2; s. auch Kap. 3). Wir müssen P daher auf ein Mengensystem fortsetzen, das die Zylindermengen enthält und außer den Eigenschaften einer Algebra auch abzählbare Vereinigungen und Durchschnitte enthält, also eine σ -Algebra ist.

Wir können dann zeigen, dass $P(N) = 0$ ist, wenn wir aus $\sum_{m=k}^{\infty} P(N_m) \leq \varepsilon$ schließen können, dass $P(\cup_{m=k}^{\infty} N_m) \leq \varepsilon$ ist. Wir wollen uns überlegen, welche Eigenschaft der Fortsetzung von P diesen Schritt erlaubt.

Für endliche Vereinigungen folgt aus der Additivität von P auf $\mathcal{Z}$ für $j \geq k \geq 1$

$$P\left(\bigcup_{m=k}^j N_m\right) \leq \sum_{m=k}^j P(N_m) \leq \sum_{m=k}^{\infty} P(N_m).$$

Dabei haben wir die Subadditivität von P benutzt, die leicht elementar zu beweisen ist (s. Satz 2.14). Ist k genügend groß mit $\sum_{m=k}^{\infty} P(N_m) \leq \varepsilon$, so ist $P(\cup_{m=k}^j N_m) \leq \varepsilon$ für alle $j \geq k$.

Hieraus folgt dann $P(\cup_{m=k}^{\infty} N_m) \leq \varepsilon$, wenn für aufsteigende Folgen $(A_n)_{n \geq 1}$ mit Vereinigung A gilt, dass $P(A_n) \rightarrow P(A)$ für $n \rightarrow \infty$ konvergiert. Wir werden in Satz 2.19 zeigen, dass für eine additive Wahrscheinlichkeit diese Stetigkeitseigenschaft äquivalent zur σ -Additivität ist. Das bedeutet, dass Eigenschaft 2 ii) von Satz 1.2 entsprechend auch für abzählbare disjunkte Vereinigungen gilt.

Fassen wir zusammen. Zum Beweis des starken Gesetzes der großen Zahlen müssen wir P auf eine σ -Algebra, die $\mathcal{Z}$ enthält, zu einer σ -additiven Mengenfunktion fortsetzen.

Was wir schon jetzt exakt bewiesen haben, können wir mit der folgenden Definition als Satz formulieren.

► **Definition 1.5** Eine Menge $N \subset \Omega$ heißt *P -vernachlässigbar*, wenn es zu jedem $\varepsilon > 0$ Mengen $N_m \in \mathcal{Z}$ ($m \geq 1$) gibt, so dass $N \subset \cup_{m=1}^{\infty} N_m$ und $\sum_{m=1}^{\infty} P(N_m) \leq \varepsilon$ ist.

In unserem Beispiel erhält man zu gegebenem $\varepsilon > 0$ diese Mengen N_m ($m \geq 1$) aus den ursprünglichen Mengen N_m ($m \geq 1$) des Beweises durch entsprechendes Weglassen von endlich vielen N_m .

► **Satz 1.6** Die Menge $\{\frac{S_n}{n} \neq \frac{1}{2} \text{ für } n \rightarrow \infty\}$ ist P -vernachlässigbar.

Die Definition von P -vernachlässigbar lässt sich unmittelbar auf beliebige additive Wahrscheinlichkeiten übertragen (s. z. B. den folgenden Satz 1.7).

Zum Schluss dieses Kapitels wollen wir mit Hilfe der Dualbruchentwicklung eine Beziehung zu der gleichmäßigen Verteilung auf $[0,1)$ herstellen.

Jede Zahl $x \in [0,1)$ lässt sich bekanntlich in einen Dualbruch $x = \sum_{n=1}^{\infty} \frac{\omega_n}{2^n}$ mit $\omega_n \in \{0,1\}$ für $n \geq 1$ entwickeln. Diese Entwicklung lässt sich durch $(\omega_n)_{n \geq 1} \in \Omega$ darstellen. Sie ist eindeutig, wenn man ausschließt, dass nur endlich viele $\omega_n = 0$ sind. Auf diese Weise erhält man für $n \geq 1$ Abbildungen $d_n: [0,1) \rightarrow \{0,1\}$, die durch $d_n(x) = \omega_n$ definiert sind,

und die wir als Folge zu der Abbildung $D: [0, 1) \rightarrow \Omega$ mit $D(x) = (d_n(x))_{n \geq 1}$ zusammensetzen.

Teilmengen von $[0, 1)$, auf denen einige der „Ziffern“ d_n konstant sind, sind zum Beispiel

$$\begin{aligned}\{x : d_1(x) = 0\} &= \left[0, \frac{1}{2}\right) \\ \{x : d_1(x) = 1\} &= \left[\frac{1}{2}, 1\right) \\ \{x : d_1(x) = d_2(x) = 0\} &= \left[0, \frac{1}{4}\right)\end{aligned}$$

und allgemein für $\omega_1, \dots, \omega_m \in \{0, 1\}$ ($m \geq 1$)

$$\{x : d_1(x) = \omega_1, \dots, d_m(x) = \omega_m\} = \left[\sum_{k=1}^m \frac{\omega_k}{2^k}, \sum_{k=1}^m \frac{\omega_k}{2^k} + \frac{1}{2^m} \right). \quad (1.8)$$

Sei Q die gleichmäßige Verteilung auf $[0, 1)$, die von $Q([a, b]) = b - a$ für $0 \leq a \leq b \leq 1$ ausgeht. Man kann Q leicht additiv auf endliche disjunkte Vereinigungen solcher Intervalle, die eine Algebra bilden, fortsetzen (s. Beispiel 2 von Abschn. 2.2). Q ist ein Modell für die zufällige Wahl einer reellen Zahl aus $[0, 1)$ mit stationärer Verteilung, z. B. auf einem Glücksrad mit auf 1 normiertem Gesamtwinkel.

Nach (1.8) ist $Q(x : d_1(x) = \omega_1, \dots, d_m(x) = \omega_m) = \frac{1}{2^m}$ und mit der Additivität folgt

$$Q(x : (d_1(x), \dots, d_m(x)) \in A_m) = \frac{|A_m|}{2^m} \quad \text{für } A_m \subset \Omega_m.$$

Unter Q hat D die Verteilung P . Denn aus $\{x : (d_1(x), \dots, d_m(x)) \in A_m\} = \{x : \pi_m(D(x)) \in A_m\} = \{x : D(x) \in \pi_m^{-1}(A_m)\} = D^{-1}(\pi_m^{-1}(A_m))$ folgt nach (1.2):

$$Q(\{x : D(x) \in A\}) = P(A) \quad \text{für } A \in \mathcal{Z}.$$

Damit lassen sich aus Aussagen über P entsprechende über Q ableiten. Was Satz 1.6 betrifft, liefert er, da

$$D^{-1}(N) = \left\{ x : \frac{1}{n} \sum_{k=1}^n \frac{d_k(x)}{2^k} \not\rightarrow \frac{1}{2} \quad \text{für } n \rightarrow \infty \right\} \quad \text{für } N = \left\{ \frac{S_n}{n} \not\rightarrow \frac{1}{2} \quad \text{für } n \rightarrow \infty \right\}$$

ist:

► **Satz 1.7 (Borel)** Die Menge aller $x \in [0, 1)$ mit $\frac{1}{n} \sum_{k=1}^n \frac{d_k(x)}{2^k} \not\rightarrow \frac{1}{2}$ für $n \rightarrow \infty$ ist Q -vernachlässigbar.

Satz 1.7 stellt ein Beispiel einer nicht-trivialen Fortsetzung von Q dar.

1.1 Übungen

- 1.1* Man gebe eine Folge von Zufallsvariablen $(X_n)_{n \geq 1}$ an mit:
1. für jedes $\varepsilon > 0$ konvergiert $P(|X_n| \geq \varepsilon) \rightarrow 0$ für $n \rightarrow \infty$
 2. $P(X_n \rightarrow 0 \text{ für } n \rightarrow \infty) = 0$
- 1.2 Im Modell des unendlichen Münzwurfs stelle man $\left\{ \frac{S_n}{n} \not\rightarrow \frac{1}{2} \text{ für } n \rightarrow \infty \right\}$ mit Hilfe abzählbarer Vereinigungen und Durchschnitte aus Zylindermengen dar.
- 1.3 Ein Bernoulli-Experiment ist eine Verallgemeinerung des idealen Münzwurfs auf ein Zufallsexperiment mit zwei möglichen Ausgängen mit i. A. verschiedenen Wahrscheinlichkeiten. Ein Modell ist gegeben durch $\Omega_1 = \{0, 1\}^1$ mit Wahrscheinlichkeit p für den Ausgang 1 für ein p mit $0 \leq p \leq 1$ und $q = 1 - p$ für den Ausgang 0. Man konstruiere ein Modell für unendlich viele unabhängige Wiederholungen eines Bernoulli-Experiments.
- 1.4 Man beweise das starke Gesetz der großen Zahlen in der Form von Satz 1.6 für unendlich viele unabhängige Wiederholungen eines Bernoulli-Experiments.
- 1.5 Sei $b \geq 2$ eine natürliche Zahl. Jede Zahl $x \in [0, 1)$ lässt sich als Bruch zur Basis b darstellen:

$$x = \sum_{n=1}^{\infty} \frac{d_n(x)}{b^n} \quad (d \in \{0, 1, \dots, b-1\}).$$

Die Darstellung ist eindeutig, wenn man ausschließt, dass die Folge $(d_n(x))_{n \geq 1}$ schließlich nur aus $(b-1)$ besteht.

Man beweise für jedes $d \in \{0, 1, \dots, b-1\}$ ein starkes Gesetz der großen Zahlen für die relative Häufigkeit des Auftretens der Ziffer d .

- 1.6 a) Man zeige: Sind die Mengen $A_n \subset \Omega$ ($n \geq 1$) P -vernachlässigbar, dann ist auch $\bigcup_{n=1}^{\infty} A_n$ P -vernachlässigbar.
- b) Was folgt daraus für das Beispiel von Übung 1.5 für alle b und d ?

Wahrscheinlichkeitstheorie und Stochastische Prozesse

Mürmann, M.

2014, XII, 428 S., Softcover

ISBN: 978-3-642-38159-1