KAPITEL 10

Gesetz der großen Zahlen

10.1. Zwei Beispiele

Beispiel 10.1.1. Wir betrachten ein Bernoulli-Experiment, das unendlich oft wiederholt wird. Die Wahrscheinlichkeit für einen Erfolg sei p. Die Zufallsvariable, die den Ausgang des i-ten Experiments beschreibt, ist:

$$X_i = \begin{cases} 1, & \text{falls Experiment } i \text{ Erfolg,} \\ 0, & \text{sonst.} \end{cases}$$

Die Anzahl der Erfolge in den ersten n Experimenten ist $S_n = X_1 + \ldots + X_n$. Dann kann man erwarten, dass

$$\frac{S_n}{n} \underset{n\to\infty}{\longrightarrow} p.$$

Dies ist ein Spezialfall des sogenannten Gesetzes der großen Zahlen.

Beispiel 10.1.2. Ein fairer Würfel wird unendlich oft geworfen. Die Zufallsvariable X_i sei die Augenzahl bei Wurf i. Mit $S_n = X_1 + \ldots + X_n$ bezeichnen wir die Summe der Augenzahlen der ersten n Würfe. Der Erwartungswert von X_i ist

$$\mathbb{E}[X_i] = \frac{1}{6} \cdot (1 + \ldots + 6) = 3.5.$$

Man kann dann erwarten, dass

$$\frac{S_n}{n} \underset{n\to\infty}{\longrightarrow} 3.5.$$

Dies ist ein anderer Spezialfall des Gesetzes der großen Zahlen.

In beiden Aussagen geht es um Konvergenz von Folgen von Zufallsvariablen. Damit wir das Gesetz der großen Zahlen formulieren können, müssen wir definieren, was wir unter einer solchen Konvergenz verstehen. Es stellt sich heraus, dass es viele verschiedene Konvergenzbegriffe für Folgen von Zufallsvariablen gibt. Vier dieser Begriffe (*Konvergenz in Wahrscheinlichkeit*, *Konvergenz in L^p*, *fast sichere Konvergenz*, *Konvergenz in Verteilung*) werden im Folgenden behandelt.

10.2. Konvergenz in Wahrscheinlichkeit und L^2-Konvergenz

Definition 10.2.1. Eine Folge von Zufallsvariablen $Z_1, Z_2, \ldots : \Omega \to \mathbb{R}$ konvergiert *in Wahrscheinlichkeit* (oder *stochastisch*) gegen eine Zufallsvariable $Z : \Omega \to \mathbb{R}$, wenn für alle $\varepsilon > 0$ gilt:

$$\lim_{n\to\infty} \mathbb{P}[\{\omega \in \Omega : |Z_n(\omega) - Z(\omega)| > \varepsilon\}] = 0.$$

In einer vereinfachten Schreibweise:

$$\lim_{n\to\infty} \mathbb{P}[|Z_n - Z| > \varepsilon] = 0.$$

Bezeichnung. $Z_n \xrightarrow[N\to\infty]{P} Z$.

Definition 10.2.2. Eine Folge von Zufallsvariablen $Z_1, Z_2, \ldots : \Omega \to \mathbb{R}$ konvergiert in L^2 (oder *im quadratischen Mittel*) gegen eine Zufallsvariable $Z : \Omega \to \mathbb{R}$, wenn $Z_n \in L^2$, $Z \in L^2$ und

$$\lim_{n\to\infty} \mathbb{E}[(Z_n - Z)^2] = 0.$$

Bezeichnung. $Z_n \xrightarrow[n\to\infty]{L^2} Z$.

Im Folgenden werden wir zeigen, dass aus der L^2-Konvergenz die stochastische Konvergenz folgt, jedoch nicht umgekehrt.

10.3. Ungleichungen von Markow und Tschebyschew

Lemma 10.3.1. *Sei $Z \geq 0$ eine Zufallsvariable. Dann gilt für alle $a > 0$:*

$$\mathbb{P}[Z \geq a] \leq \frac{\mathbb{E}Z}{a}.$$

Beweis. Für die Zufallsvariable Z gilt:

$$Z(\omega) \geq \begin{cases} a, & \text{falls } Z(\omega) \geq a, \\ 0, & \text{falls } Z(\omega) < a. \end{cases}$$

Mit anderen Worten, $Z \geq a \cdot \mathbb{1}_{\{Z \geq a\}}$. Für den Erwartungswert von Z gilt somit:

$$\mathbb{E}Z \geq \mathbb{E}[a \cdot \mathbb{1}_{\{Z\geq a\}}] = a \cdot \mathbb{E}[\mathbb{1}_{\{Z\geq a\}}] = a \cdot \mathbb{P}[Z \geq a].$$

Durch Umformen erhält man die Ungleichung im Lemma. □

Satz 10.3.2 (Tschebyschew-Ungleichung). *Sei $X \in L^2$ eine Zufallsvariable. Dann gilt für alle $a > 0$:*

$$\mathbb{P}[|X - \mathbb{E}X| \geq a] \leq \frac{\operatorname{Var} X}{a^2}.$$

Beweis. Da $Z := (X - \mathbb{E}X)^2 \geq 0$, folgt mit Lemma 10.3.1:

$$\mathbb{P}[(X - \mathbb{E}X)^2 \geq a^2] \leq \frac{\mathbb{E}[(X - \mathbb{E}X)^2]}{a^2} = \frac{\operatorname{Var} X}{a^2}.$$

Die Ungleichung $(X - \mathbb{E}X)^2 \geq a^2$ gilt genau dann, wenn $|X - \mathbb{E}X| \geq a$. Damit erhält man die Tschebyschew-Ungleichung. □

Satz 10.3.3 (Markow-Ungleichung). *Sei $Z \geq 0$ eine Zufallsvariable und $f : [0, \infty) \to [0, \infty)$ eine monoton steigende Funktion. Dann gilt für alle $a > 0$:*

$$\mathbb{P}[Z \geq a] \leq \frac{\mathbb{E}[f(Z)]}{f(a)}.$$

Beweis. Da f monoton steigend ist, gilt für die Wahrscheinlichkeit:

$$\mathbb{P}[Z \geq a] = \mathbb{P}[f(Z) \geq f(a)] \leq \frac{\mathbb{E}[f(Z)]}{f(a)}.$$

Die letzte Ungleichung erhält man aus Lemma 10.3.1, indem man dort $f(Z)$ anstelle von Z einsetzt. □

Beispiel 10.3.4. Mit $f(a) = a^p$, wobei $p > 0$, erhält man die Ungleichung

$$\mathbb{P}[Z \geq a] \leq \frac{\mathbb{E}[Z^p]}{a^p}.$$

Satz 10.3.5. *Seien $Z_1, Z_2, \ldots$ und Z Zufallsvariablen mit $Z_n \xrightarrow{L^2} Z$. Dann gilt: $Z_n \xrightarrow{P} Z$.*

Beweis. Aus $Z_n \xrightarrow{L^2} Z$ folgt mit der Definition der L^2-Konvergenz, dass

$$\lim_{n\to\infty} \mathbb{E}[(Z_n - Z)^2] \underset{n\to\infty}{\longrightarrow} 0.$$

Sei $\varepsilon > 0$. Mit dem obigen Beispiel (mit $p = 2$) erhalten wir

$$\mathbb{P}[|Z_n - Z| \geq \varepsilon] \leq \frac{\mathbb{E}[(Z_n - Z)^2]}{\varepsilon^2} \underset{n\to\infty}{\longrightarrow} 0.$$

Wir haben gezeigt, dass

$$\lim_{n\to\infty} \mathbb{P}[|Z_n - Z| \geq \varepsilon] = 0$$

und somit $Z_n \xrightarrow{P} Z$. □

10.4. Schwaches Gesetz der großen Zahlen

Satz 10.4.1 (Schwaches Gesetz der großen Zahlen). *Seien $X_1, X_2, \ldots : \Omega \to \mathbb{R}$ unabhängige Zufallsvariablen mit $\mathbb{E}X_i = \mu$ und $\operatorname{Var} X_i = \sigma^2$ für alle $i \in \mathbb{N}$. Dann gilt:*

$$\frac{X_1 + \ldots + X_n}{n} \underset{n\to\infty}{\xrightarrow{L^2, P}} \mu$$

Das heißt, das arithmetische Mittel der Zufallsvariablen $X_1, \ldots, X_n$ konvergiert in L^2 bzw. in Wahrscheinlichkeit gegen den Erwartungswert.

Beweis. Sei $S_n = X_1 + \ldots + X_n$ die Summe der ersten n Zufallsvariablen. Dann gilt für den Erwartungswert und die Varianz der Summe

$$\mathbb{E}S_n = n\mu \text{ und } \operatorname{Var} S_n = n\sigma^2,$$

wobei im Fall der Varianz benutzt wird, dass die Zufallsvariablen unabhängig sind. Wir zeigen zuerst die L^2-Konvergenz:

$$\mathbb{E}\left[\left(\frac{S_n}{n} - \mu\right)^2\right] = \mathbb{E}\left[\left(\frac{S_n - \mathbb{E}S_n}{n}\right)^2\right] = \frac{1}{n^2} \cdot \operatorname{Var} S_n = \frac{1}{n^2} \cdot n\sigma^2 = \frac{\sigma^2}{n} \underset{n\to\infty}{\longrightarrow} 0.$$

Daraus fogt $\frac{S_n}{n} \xrightarrow{L^2} \mu$. Und mit Satz 10.3.5 folgt nun auch die stochastische Konvergenz: $\frac{S_n}{n} \xrightarrow[N\to\infty]{P} \mu$. □

Beispiel 10.4.2. Ein fairer Würfel wird unendlich oft geworfen. Es sei S_n die Augensumme nach n Würfen, also $S_n = X_1 + \ldots + X_n$, wobei die Zufallsvariable X_i die Augenzahl im i-ten Wurf beschreibt. Der Erwartungswert eines Wurfes ist $\mu = 3.5$. Mit Satz 10.4.1 folgt nun

$$\frac{S_n}{n} \xrightarrow[N\to\infty]{P} 3.5.$$

Man kann entsprechend der Definition der stochastischen Konvergenz z.B. mit $\varepsilon = 0.01$ das folgende Ergebnis erhalten:

$$\lim_{n\to\infty} \mathbb{P}\left[3,49 < \frac{S_n}{n} < 3,51\right] = 1.$$

Die Wahrscheinlichkeit, dass die mittlere Augenzahl außerhalb des Intervalls $(3,49, 3,51)$ liegt, geht gegen 0, wenn die Anzahl der Würfe gegen ∞ geht.

Bemerkung 10.4.3. Das Wort "schwach" in der Bezeichnung "schwaches Gesetz der großen Zahlen" bezieht sich auf die Art der Konvergenz (in Wahrscheinlichkeit oder L^2). Im folgenden werden wir stärkere Versionen des Gesetzes der großen Zahlen beweisen.

10.5. Fast sichere Konvergenz

Definition 10.5.1. Sei $(\Omega, \mathcal{F}, \mathbb{P})$ ein Wahrscheinlichkeitsraum. Ein Ereignis $A \in \mathcal{F}$ heißt ein *fast sicheres* Ereignis, wenn $\mathbb{P}[A] = 1$.

Definition 10.5.2. Ein Ereignis $B \in \mathcal{F}$ heißt ein *fast unmögliches* Ereignis (oder ein *Nullereignis*), wenn $\mathbb{P}[B] = 0$.

Beispiel 10.5.3. Das Ereignis Ω ist fast sicher (und in Wirklichkeit, sogar sicher). Das Ereignis $\emptyset$ ist fast unmöglich (und in Wirklichkeit, sogar unmöglich). Das Komplement einer fast sicheren Ereignisses ist fast unmöglich und umgekehrt, das Komplement eines fast unmöglichen Ereignisses ist fast sicher.

Bemerkung 10.5.4. Man geht davon aus, dass ein Nullereignis in der Praxis nicht beobachtet werden kann. Es kann zwar sein, dass es im Wahrscheinlichkeitsraum gewisse Ausgänge gibt, die zum Eintreten dieses Ereignisses führen, diese bilden allerdings eine so "kleine" Menge, dass man diese Ausgänge nicht beobachten kann, egal, wie oft man das Experiment wiederholt. Analog geht man davon aus, dass ein fast sicheres Ereignis bei jeder Ausführung des Experiments beobachtet wird.

Bemerkung 10.5.5. Eine Vereinigung von abzählbar vielen Nullereignissen ist ein Nullereignis (auch dann, wenn die Ereignisse nicht disjunkt sind). Ein Schnitt von höchstens abzählbar vielen fast sicheren Ereignissen ist wieder ein fast sicheres Ereignis. Diese Aussagen gelten allerdings nicht für überabzählbare Schnitte bzw. Vereinigungen.

Beispiel 10.5.6. Man betrachte einen idealen Zufallsgenerator, der eine im Intervall $[0, 1]$ gleichverteilte Zahl erzeugt. Die Grundmenge dieses Experiments ist $\Omega = [0, 1]$, $\mathcal{F} = \mathcal{B}_{[0,1]}$ ist die Borel-σ-Algebra auf $[0, 1]$ und $\mathbb{P} = \lambda$ ist das Lebesgue-Maß.

Es sei $B = \{q_1, q_2, \ldots\} \subset [0,1]$ ein Ereignis, das aus abzählbar vielen Ausgängen besteht. Zum Beispiel, kann man das Ereignis

$$B = \mathbb{Q} \cap [0,1] = \text{“es wurde eine rationale Zahl erzeugt”}$$

betrachten. Für jedes q_i ist $\mathbb{P}[q_i] = 0$ und somit, mit der σ-Additivität,

$$\mathbb{P}[B] = \sum_{i=1}^{\infty} \mathbb{P}[q_i] = 0.$$

Daraus folgt, dass B fast unmöglich ist. Ein idealer Zufallsgenerator erzeugt also nie eine rationale Zahl. Natürlich erfüllt kein realer Zufallsgenerator (z.B. in einem Rechner) diese Anforderung. Im Gegenteil: Ein realer Zufallsgenerator erzeugt nur Zahlen mit einer endlichen Dezimaldarstellung, also nur rationale Zahlen.
Es sei nun $A = [0,1] \backslash B$ ein Ereignis, das ein Komplement einer abzählbaren Menge ist. Zum Beispiel kann man das Ereignis

$$A = [0,1] \backslash \mathbb{Q} = \text{“es wurde eine irrationale Zahl erzeugt”}$$

betrachten. Dann gilt für die Wahrscheinlichkeit von A: $\mathbb{P}[A] = 1 - \mathbb{P}[B] = 1 - 0 = 1$. Daraus folgt, dass A fast sicher ist. Ein idealer Zufallsgenerator erzeugt also nur irrationale Zahlen.

Definition 10.5.7. Eine Folge von Zufallsvariablen $Z_1, Z_2, \ldots : \Omega \to \mathbb{R}$ konvergiert *fast sicher* gegen eine Zufallsvariable $Z : \Omega \to \mathbb{R}$, wenn:

$$\mathbb{P}[\{\omega \in \Omega : \lim_{n\to\infty} Z_n(\omega) = Z(\omega)\}] = 1.$$

In einer vereinfachten Schreibweise:

$$\mathbb{P}\left[\lim_{n\to\infty} Z_n = Z\right] = 1.$$

Bezeichnung. $Z_n \xrightarrow[n\to\infty]{f.s.} Z$ oder $\lim_{n\to\infty} Z_n = Z$ f.s.

Bemerkung 10.5.8. Damit diese Definition Sinn macht, muss man zeigen, dass das Ereignis

$$A := \{\omega \in \Omega : \lim_{n\to\infty} Z_n(\omega) = Z(\omega)\}$$

messbar ist. Wir zeigen also, dass $A \in \mathcal{F}$. Die Gleichheit $\lim_{n\to\infty} Z_n(\omega) = Z(\omega)$ gilt genau dann, wenn für alle $\varepsilon > 0$ ein $k \in \mathbb{N}$ existiert, so dass für alle $n \geq k$:

$$|Z_n(\omega) - Z(\omega)| \leq \varepsilon.$$

Man kann sich auch nur auf Werte $\varepsilon = \frac{1}{m}$ mit $m \in \mathbb{N}$ einschränken. Dann kann man A folgendermaßen beschreiben:

$$\begin{aligned} A &= \left\{\omega \in \Omega : \forall m \in \mathbb{N}\ \exists k \in \mathbb{N}\ \forall n \geq k : |Z_n(\omega) - Z(\omega)| \leq \frac{1}{m}\right\} \\ &= \bigcap_{m\in\mathbb{N}} \bigcup_{k\in\mathbb{N}} \bigcap_{n\geq k} \left\{\omega \in \Omega : |Z_n(\omega) - Z(\omega)| \leq \frac{1}{m}\right\}. \end{aligned}$$

Dabei haben wir $\forall$ durch $\bigcap$ und $\exists$ durch $\bigcup$ ersetzt. Da $|Z_n - Z|$ eine messbare Funktion ist, sind alle Mengen der Form

$$\left\{\omega \in \Omega : |Z_n(\omega) - Z(\omega)| \leq \frac{1}{m}\right\}$$

messbare Ereignisse. Da die messbaren Ereignisse eine σ-Algebra bilden, sind auch beliebige abzählbare Schnitte und Vereinigungen dieser Mengen messbar, also auch A. Es folgt, dass $A \in \mathcal{F}$.

Satz 10.5.9. *Seien $Z_1, Z_2, \dots$ und Z Zufallsvariablen mit $Z_n \xrightarrow[n\to\infty]{f.s.} Z$. Dann gilt: $Z_n \xrightarrow[N\to\infty]{P} Z$.*

Beweis. Es gelte, dass $Z_n \xrightarrow[n\to\infty]{f.s.} Z$. Dann hat das Ereignisses

$$A = \{\omega \in \Omega : \lim_{n\to\infty} Z_n(\omega) = Z(\omega)\}$$

Wahrscheinlichkeit $\mathbb{P}[A] = 1$. Sei $\varepsilon > 0$. Wir definieren Ereignisse

$$A_k = \{\omega \in \Omega : \forall n \geq k : |Z_n(\omega) - Z(\omega)| \leq \varepsilon\}, \qquad k \in \mathbb{N}.$$

Es gilt $A \subset \cup_{k\in\mathbb{N}} A_k$. Aus $1 = \mathbb{P}[A] \leq \mathbb{P}[\cup_{k\in\mathbb{N}} A_k]$ folgt, dass

$$\mathbb{P}[\cup_{k\in\mathbb{N}} A_k] = 1.$$

Außerdem gilt $A_1 \subset A_2 \subset A_3 \subset \dots$. Mit der Stetigkeit des Wahrscheinlichkeitsmaßes folgt

$$\lim_{k\to\infty} \mathbb{P}[A_k] = \mathbb{P}[\cup_{k\in\mathbb{N}} A_k] = 1.$$

Es ist nun

$$\lim_{k\to\infty} \mathbb{P}[|Z_k - Z| \leq \varepsilon] = \lim_{k\to\infty} \mathbb{P}[\{\omega \in \Omega : |Z_k(\omega) - Z(\omega)| \leq \varepsilon\}] \geq \lim_{k\to\infty} \mathbb{P}[A_k] = 1.$$

Also ist $\lim_{k\to\infty} \mathbb{P}[|Z_k - Z| > \varepsilon] = 0$. Daraus folgt $Z_n \xrightarrow[N\to\infty]{P} Z$. □

Beispiel 10.5.10. Aus $Z_n \xrightarrow[N\to\infty]{P} Z$ folgt nicht $Z_n \xrightarrow[n\to\infty]{f.s.} Z$. Das zeigen wir an einem Beispiel. Die Zufallsvariablen $Z_1, Z_2, \dots$ seien unabhängig mit

$$\mathbb{P}[Z_n = 1] = \frac{1}{n^\alpha}\,, \quad \mathbb{P}[Z_n = 0] = 1 - \frac{1}{n^\alpha}.$$

Dabei sei $\alpha \in (0, 1]$ ein Parameter.

(1) Es gilt $Z_n \xrightarrow[N\to\infty]{P} 0$, denn für alle $\varepsilon > 0$ gilt

$$\mathbb{P}[|Z_n| > \varepsilon] \leq \mathbb{P}[Z_n = 1] = \frac{1}{n^\alpha} \xrightarrow[n\to\infty]{} 0.$$

(2) Allerdings gilt nicht $Z_n \xrightarrow[n\to\infty]{f.s.} 0$. Es ist $\sum_{n=1}^{\infty} \frac{1}{n^\alpha} = \infty$ (da $\alpha \leq 1$). Mit dem Lemma von Borel–Cantelli (Teil 2) folgt

$$\mathbb{P}[Z_n = 1 \text{ für unendlich viele } n] = 1.$$

Daraus folgt $\mathbb{P}[\lim_{n\to\infty} Z_n \neq 0] = 1$ und deshalb gilt nicht $Z_n \xrightarrow[n\to\infty]{f.s.} 0$.

10.6. Starkes Gesetz der großen Zahlen: Erste Version

Satz 10.6.1 (Starkes Gesetz der großen Zahlen: Erste Version). *Seien $X_1, X_2, \dots$ unabhängige Zufallsvariablen auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbb{P})$ mit $\mathbb{E}[X_n^2] < \infty$ für alle $n \in \mathbb{N}$. Außerdem sei*

$$\sum_{n=1}^{\infty} \frac{\operatorname{Var} X_n}{n^2} < \infty.$$

Für die Summen $S_n = X_1 + \ldots + X_n$ gilt dann

$$\frac{S_n - \mathbb{E} S_n}{n} \xrightarrow[n\to\infty]{f.s.} 0.$$

Bemerkung 10.6.2. Spezialfall: Seien $X_1, X_2, \dots$ unabhängige Zufallsvariablen mit $\mathbb{E} X_n = \mu$ und $\operatorname{Var} X_n = \sigma^2$. Dann gilt:

$$\sum_{n=1}^{\infty} \frac{\operatorname{Var} X_n}{n^2} = \sigma^2 \cdot \sum_{n=1}^{\infty} \frac{1}{n^2} < \infty.$$

Satz 10.6.1 besagt nun, dass $\frac{S_n - n\mu}{n} \xrightarrow[n\to\infty]{f.s.} 0$ und daraus folgt, dass

$$\frac{S_n}{n} \xrightarrow[n\to\infty]{f.s.} \mu.$$

Die fast sichere Konvergenz impliziert die stochastische Konvergenz (ist also stärker), daher die Bezeichung starkes Gesetz der großen Zahlen.

Im Folgenden werden wir Satz 10.6.1 beweisen. Dafür benötigen wir einige Hilfsmittel.

Satz 10.6.3 (Kolmogorow-Ungleichung). *Seien $X_1, X_2, \dots, X_n$ unabhängige Zufallsvariablen mit Erwartungswert $\mathbb{E} X_i = 0$ und $X_i \in L^2$ für $i = 1, \dots, n$. Sei $S_i = X_1 + \ldots + X_i$ die Summe der ersten i Zufallsvariablen. Dann gilt für alle $a > 0$ die Ungleichung*

$$\mathbb{P}[\max\{|S_1|, |S_2|, \dots, |S_n|\} \geq a] \leq \frac{\operatorname{Var} S_n}{a^2}.$$

Bemerkung 10.6.4. Die Tschebyschew-Ungleichung ist schwächer. Mit der kann man lediglich eine der Summen abschätzen, z.B.

$$\mathbb{P}[|S_n| \geq a] \leq \frac{\operatorname{Var} S_n}{a^2}.$$

Beweis von Satz 10.6.3. Schritt 1. Wir definieren das Ereignis:

$$A := \left\{ \max_{k=1,\dots,n} |S_k| \geq a \right\}$$

und, für $k = 1, \dots, n$, die Ereignisse

$$\begin{aligned} A_k &:= \{|S_1| < a, |S_2| < a, \dots, |S_{k-1}| < a, |S_k| \geq a\} \\ &= \text{“Erster Austritt aus dem Intervall } (-a, a) \text{ findet zum Zeitpunkt } k \text{ statt.”} \end{aligned}$$

Es gilt $A = \cup_{k=1}^{n} A_k$. (Ereignis A tritt genau dann ein, wenn es ein k gibt, sodass der erste Austritt aus dem Intervall $(-a, a)$ zum Zeitpunkt k stattfindet). Außerdem sind $A_1, \dots, A_n$

disjunkt. (Wenn der erste Austritt zum Zeitpunkt k stattfindet, kann er nicht zu einem anderen Zeitpunkt stattfinden). Also ist

$$\mathbb{P}[A] = \sum_{k=1}^{n} \mathbb{P}[A_k].$$

Schritt 2. Die Varianz von S_n ist: $\operatorname{Var} S_n = \mathbb{E}[S_n^2] - (\mathbb{E}S_n)^2 = \mathbb{E}[S_n^2]$, denn $\mathbb{E}S_n = \mathbb{E}[X_1 + \ldots + X_n] = \mathbb{E}X_1 + \ldots \mathbb{E}X_n = 0$. Wir betrachten also $\mathbb{E}[S_n^2]$:

$$\mathbb{E}[S_n^2] \geq \mathbb{E}[S_n^2 \cdot \mathbb{1}_A] = \mathbb{E}\left[\sum_{k=1}^{n} S_n^2 \cdot \mathbb{1}_{A_k}\right] = \sum_{k=1}^{n} \mathbb{E}[S_n^2 \cdot \mathbb{1}_{A_k}].$$

Für die Summanden auf der rechten Seite gilt

$$\begin{aligned}
\mathbb{E}[S_n^2 \cdot \mathbb{1}_{A_k}] &= \mathbb{E}[(S_k + S_n - S_k)^2 \cdot \mathbb{1}_{A_k}] \\
&= \mathbb{E}[S_k^2 \cdot \mathbb{1}_{A_k}] + 2 \cdot \mathbb{E}[S_k \cdot (S_n - S_k) \cdot \mathbb{1}_{A_k}] + \mathbb{E}[(S_n - S_k)^2 \cdot \mathbb{1}_{A_k}] \\
&= (1) + (2) + (3).
\end{aligned}$$

Schritt 3. Die drei Terme auf der rechten Seite kann man folgendermaßen abschätzen. Der erste Term:

$$(1) = \mathbb{E}[S_k^2 \cdot \mathbb{1}_{A_k}] \geq \mathbb{E}[a^2 \cdot \mathbb{1}_{A_k}] = a^2 \cdot \mathbb{P}[A_k],$$

denn wenn A_k eintritt, ist $\mathbb{1}_{A_k} = 1$ und $|S_k| \geq a$.
Der zweite Term:

$$(2) = 2 \cdot \mathbb{E}[S_k \cdot \mathbb{1}_{A_k} \cdot (S_n - S_k)] = 2 \cdot \mathbb{E}[S_k \cdot \mathbb{1}_{A_k}] \cdot \mathbb{E}[S_n - S_k] = 0.$$

Die Umformung gilt, da die Zufallsvariable $S_k \cdot \mathbb{1}_{A_k}$ nur von $X_1, \ldots, X_k$ und die Zufallsvariable $(S_n - S_k)$ nur von $X_{k+1}, \ldots, X_n$ abhängt, also sind diese zwei Zufallsvariablen unabhängig. Außerdem gilt $\mathbb{E}[S_n - S_k] = 0$, da $\mathbb{E}S_n = \mathbb{E}S_k = 0$.
Schließlich ist der dritte Term trivialerweise nichtnegativ:

$$(3) \geq 0.$$

Schritt 4. Somit erhalten wir

$$\mathbb{E}[S_n^2 \cdot \mathbb{1}_{A_k}] = (1) + (2) + (3) \geq a^2 \cdot \mathbb{P}[A_k].$$

Betrachte wir nun wieder die Varianz von S_n:

$$\operatorname{Var} S_n = \mathbb{E}[S_n^2] \geq \sum_{k=1}^{n} \mathbb{E}[S_n^2 \cdot \mathbb{1}_{A_k}] \geq \sum_{k=1}^{n} a^2 \cdot \mathbb{P}[A_k].$$

Damit folgt: $\mathbb{P}[A] \leq \frac{\operatorname{Var} S_n}{a^2}$. □

Beweis von Satz 10.6.1. Schritt 1. O.B.d.A. können wir annehmen, dass $\mathbb{E}X_n = 0$ für alle $n \in \mathbb{N}$. Sonst betrachte man die zentrierten Zufallsvariablen

$$\widetilde{X_n} = X_n - \mathbb{E}X_n, \quad \widetilde{S_n} = \widetilde{X_1} + \ldots + \widetilde{X_n}.$$

Es gilt dann $\mathbb{E}\widetilde{X_n} = 0$ und

$$\sum_{n=1}^{\infty} \frac{\operatorname{Var} \widetilde{X_n}}{n^2} = \sum_{n=1}^{\infty} \frac{\operatorname{Var} X_n}{n^2} < \infty,$$

denn $\operatorname{Var} \widetilde{X_n} = \operatorname{Var}[X_n - \mathbb{E}X_n] = \operatorname{Var} X_n$, da $\mathbb{E}X_n$ eine Konstante ist. Außerdem ist

$$\frac{\widetilde{S_n} - \mathbb{E}\widetilde{S_n}}{n} = \frac{\widetilde{S_n}}{n} = \frac{S_n - \mathbb{E}S_n}{n}.$$

Somit gilt Satz 10.6.1 für die Zufallsvariablen X_n genau dann, wenn er für die Zufallsvariablen $\widetilde{X}_n$ gilt.

Schritt 2. Sei also $\mathbb{E}S_n = 0$ für alle $n \in \mathbb{N}$. Wir zeigen, dass

$$\frac{|S_k|}{k} \xrightarrow[k\to\infty]{f.s.} 0.$$

Dazu definieren wir die Zufallsvariablen

$$U_n := \max_{k=1,\ldots,2^n} |S_k|, \qquad n \in \mathbb{N}.$$

Nun gibt es für alle $k \in \mathbb{N}$ genau ein $n \in \mathbb{N}$ mit $2^{n-1} \le k < 2^n$. Daraus folgt, dass

$$\frac{|S_k|}{k} \le \frac{U_n}{k} \le \frac{U_n}{2^{n-1}} = 2 \cdot \frac{U_n}{2^n},$$

also reicht es zu zeigen, dass

$$\frac{U_n}{2^n} \xrightarrow[n\to\infty]{f.s.} 0.$$

Schritt 3. Sei $\varepsilon > 0$ fest. Betrachte das Ereignis $A_n(\varepsilon) = \left\{\frac{U_n}{2^n} > \varepsilon\right\}$. Wir zeigen nun, dass dieses Ereignis mit Wahrscheinlichkeit 1 nur endlich oft eintritt, also

$$\mathbb{P}[\text{“}A_n(\varepsilon) \text{ tritt für unendlich viele } n \in \mathbb{N} \text{ ein”}] = \mathbb{P}[\limsup_{n\to\infty} A_n(\varepsilon)] = 0.$$

Dafür benutzen wir den ersten Teil des Lemmas von Borel–Cantelli. Mit der Kolmogorov-Ungleichung erhalten wir

$$\sum_{n=1}^{\infty} \mathbb{P}[A_n(\varepsilon)] = \sum_{n=1}^{\infty} \mathbb{P}[U_n > 2^n\varepsilon] = \sum_{n=1}^{\infty} \mathbb{P}\left[\max_{k=1,\ldots,2^n} |S_k| > 2^n\varepsilon\right] \le \sum_{n=1}^{\infty} \frac{\operatorname{Var} S_{2^n}}{2^{2n} \cdot \varepsilon^2}.$$

Wir definieren noch $\sigma_k^2 = \operatorname{Var} X_k$. Somit gilt

$$\begin{aligned}
\sum_{n=1}^{\infty} \mathbb{P}[A_n(\varepsilon)] &\leq \frac{1}{\varepsilon^2} \cdot \sum_{n=1}^{\infty} \frac{\sigma_1^2 + \sigma_2^2 + \ldots + \sigma_{2^n}^2}{2^{2n}} \quad \text{(gilt, da } X_1, X_2, \ldots \text{ unabhängig)} \\
&= \frac{1}{\varepsilon^2} \cdot \sum_{k=1}^{\infty} \sigma_k^2 \cdot \sum_{n: 2^n \geq k} \frac{1}{2^{2n}} \\
&\leq \frac{1}{\varepsilon^2} \cdot \sum_{k=1}^{\infty} \sigma_k^2 \cdot \frac{1}{k^2} \cdot \sum_{m=1}^{\infty} \frac{1}{4^m} \\
&= \frac{4}{3} \cdot \frac{1}{\varepsilon^2} \sum_{k=1}^{\infty} \frac{\sigma_k^2}{k^2} \\
&< \infty. \quad \text{(gilt nach Voraussetzung)}
\end{aligned}$$

Mit dem Lemma von Borel–Cantelli folgt, dass $\mathbb{P}[\limsup_{n\to\infty} A_n(\varepsilon)] = 0$.

Schritt 4. Setzt man nun $\varepsilon = \frac{1}{m}$ mit $m \in \mathbb{N}$, so gilt für alle $m \in \mathbb{N}$:

$$P\left[\limsup_{n\to\infty} \frac{U_n}{2^n} > \frac{1}{m}\right] = 0.$$

Da eine Vereinigung von abzählbar vielen Nullereignisses wieder ein Nullereignis ist, gilt somit auch

$$\mathbb{P}\left[\limsup_{n\to\infty} \frac{U_n}{2^n} > 0\right] = 0.$$

Daraus folgt:

$$\mathbb{P}\left[\lim_{n\to\infty} \frac{U_n}{2^n} = 0\right] = 1$$

und somit $\frac{U_n}{2^n} \xrightarrow[n\to\infty]{f.s.} 0$. □

10.7. Starkes Gesetz der Großen Zahlen: Zweite Version

In der ersten Version des starken Gesetzes der großen Zahlen haben wir vorausgesetzt, dass die Varianzen der Zufallsvariablen endlich sein sollen. In der Behauptung des Satzes, nämlich

$$\frac{S_n - \mathbb{E}S_n}{n} \xrightarrow[n\to\infty]{f.s.} 0$$

taucht die Varianz allerdings nicht auf. Man kann sich deshalb fragen, ob die Voraussetzung der Endlichkeit der Varianz nicht überflüssig ist. Das ist in der Tat der Fall, wie der nächste Satz zeigt.

Satz 10.7.1 (Starkes Gesetz der großen Zahlen: Zweite Version). *Seien $X_1, X_2, \ldots$ unabhängige und identisch verteilte Zufallsvariablen auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbb{P})$. Die Zufallsvariablen seien außerdem integrierbar, d.h. $\mathbb{E}|X_1| < \infty$. Dann gilt:*

$$\frac{X_1 + \ldots + X_n}{n} \xrightarrow[n\to\infty]{f.s.} \mathbb{E}[X_1]$$

Das heißt, das arithmetische Mittel der Zufallsvariablen konvergiert fast sicher gegen den Erwartungswert.

Bemerkung 10.7.2. Die Zufallsvariablen $X_1, X_2, \dots$ heißen identisch verteilt, wenn ihre Verteilungsfunktionen gleich sind, d.h.

$$F_{X_1}(t) = F_{X_2}(t) = \dots \quad \text{für alle } t \in \mathbb{R}.$$

Identisch verteilte Zufallsvariablen haben den gleichen Erwartungswert: $\mathbb{E}[X_1] = \mathbb{E}[X_2] = \dots$.

Für den Beweis von Satz 10.7.1 benötigen wir mehrere Lemmata.

Lemma 10.7.3. *Seien $X_n, Y_n, X, Y : \Omega \to \mathbb{R}$ Zufallsvariablen mit $X_n \xrightarrow[n\to\infty]{f.s.} X$ und $Y_n \xrightarrow[n\to\infty]{f.s.} Y$. Dann gilt:*

$$\begin{aligned} X_n + Y_n &\xrightarrow[n\to\infty]{f.s.} X + Y, \\ X_n \cdot Y_n &\xrightarrow[n\to\infty]{f.s.} X \cdot Y, \\ a \cdot X_n &\xrightarrow[n\to\infty]{f.s.} a \cdot X, \quad a \in \mathbb{R}. \end{aligned}$$

Beweis. Wir beweisen nur die erste Aussage. Dazu definieren wir zwei Ereignisse:

$$\begin{aligned} A &:= \{\omega \in \Omega : \lim_{n\to\infty} X_n(\omega) = X(\omega)\}, \\ B &:= \{\omega \in \Omega : \lim_{n\to\infty} Y_n(\omega) = Y(\omega)\}. \end{aligned}$$

Es gilt mit der Definition der fast sicheren Konvergenz:

$$X_n \xrightarrow[n\to\infty]{f.s.} X \Rightarrow \mathbb{P}[A] = 1 \text{ und } Y_n \xrightarrow[n\to\infty]{f.s.} Y \Rightarrow \mathbb{P}[B] = 1.$$

Daraus folgt:

$$\mathbb{P}[A^c] = 0 \text{ und } \mathbb{P}[B^c] = 0.$$

Für Ausgänge aus dem Schnitt von A und B, also $\omega \in A \cap B$, gilt:

$$\lim_{n\to\infty} (X_n(\omega) + Y_n(\omega)) = X(\omega) + Y(\omega).$$

Für die Wahrscheinlichkeit des Ereignisses $A \cap B$ gilt:

$$\mathbb{P}[A \cap B] = 1 - \mathbb{P}[(A \cap B)^c] = 1 - \mathbb{P}[A^c \cup B^c] \geq 1 - \mathbb{P}[A^c] - \mathbb{P}[B^c] = 1,$$

wobei wir im zweiten Schritt die Regeln von de Morgan und im dritten Schritt die Subadditivität benutzt haben. Das Ereignis $A \cap B$ tritt also fast sicher ein. Damit gilt:

$$X_n + Y_n \xrightarrow[n\to\infty]{f.s.} X + Y.$$

Die anderen Aussagen werden analog bewiesen. □

Lemma 10.7.4. *Sei $Z \geq 0$ eine Zufallsvariable. Dann gilt:*

$$\sum_{n=1}^{\infty} \mathbb{P}[Z \geq n] \leq \mathbb{E}Z \leq \sum_{n=0}^{\infty} \mathbb{P}[Z \geq n].$$

Beweis. Wir beweisen nur die Ungleichung $\sum_{n=1}^{\infty} \mathbb{P}[Z \geq n] \leq \mathbb{E}Z$. Die andere Ungleichung kann man analog beweisen. Sei $p_n = \mathbb{P}[n \leq Z < n+1]$. Es gilt

$$\mathbb{E}Z = \mathbb{E}\left[\sum_{n=0}^{\infty} Z \cdot \mathbb{1}_{n \leq Z < n+1}\right] \geq \mathbb{E}\left[\sum_{n=0}^{\infty} n \cdot \mathbb{1}_{n \leq Z < n+1}\right] = \sum_{n=0}^{\infty} n p_n = \sum_{n=1}^{\infty} n p_n.$$

Es gilt außerdem

$$\begin{array}{lcccccc} \mathbb{P}[Z \geq 1] & = & p_1+ & p_2+ & p_3+ & p_4+ & \ldots, \\ \mathbb{P}[Z \geq 2] & = & & p_2+ & p_3+ & p_4+ & \ldots, \\ \mathbb{P}[Z \geq 3] & = & & & p_3+ & p_4+ & \ldots, \\ \mathbb{P}[Z \geq 4] & = & & & & p_4+ & \ldots, \end{array}$$

Addiert man diese Ungleichungen, so erhält man $\sum_{n=1}^{\infty} \mathbb{P}[Z \geq n] = \sum_{n=1}^{\infty} n p_n$. Kombiniert man beide Resultate, so erhält man die gewünschte Ungleichung. □

Lemma 10.7.5. *Für alle $k \geq 2$ gilt:* $\sum_{n=k}^{\infty} \frac{1}{n^2} \leq \frac{2}{k}$.

Beweis. $\sum_{n=k}^{\infty} \frac{1}{n^2} \leq \int\limits_{k-1}^{\infty} \frac{1}{x^2}\, \mathrm{d}x = \frac{1}{k-1} \leq \frac{2}{k}$. □

Lemma 10.7.6 (Cesaro). *Sei $x_1, x_2, \ldots$ eine Folge in reeller Zahlen mit einem endlichen Grenzwert $\lim_{n\to\infty} x_n = x$. Dann gilt:*

$$\lim_{n\to\infty} \frac{x_1 + \ldots + x_n}{n} = x.$$

Beweis. Schritt 0. Wir können o.B.d.A. annehmen, dass $x = 0$. Sonst kann man $\widetilde{x}_n := x_n - x$ betrachten.

Schritt 1. Sei $\varepsilon > 0$. Aus $\lim_{n\to\infty} x_n = 0$ folgt, dass es ein $K \in \mathbb{N}$ gibt, so dass für alle $k \geq K$ gilt:

$$|x_k| < \frac{\varepsilon}{2}.$$

Nachdem K gewählt wurde, können wir ein $N \in \mathbb{N}$ finden, so dass für alle $n \geq N$ gilt:

$$\frac{|x_1 + \ldots + x_K|}{n} < \frac{\varepsilon}{2}.$$

Das liegt daran, dass der Zähler konstant ist, während der Nenner beliebig groß gewählt werden kann.

Schritt 2. Sei nun $n \geq \max\{K, N\}$. Dann gilt:

$$\begin{aligned} \left|\frac{x_1 + \ldots + x_n}{n}\right| &\leq \left|\frac{x_1 + \ldots + x_K}{n}\right| + \frac{1}{n} \sum_{k=K+1}^{n} |x_k| \\ &\leq \frac{\varepsilon}{2} + \frac{1}{n} \sum_{k=K+1}^{n} \frac{\varepsilon}{2} \\ &\leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \\ &= \varepsilon. \end{aligned}$$

Da $\varepsilon > 0$ beliebig ist, folgt daraus, dass $\lim_{n\to\infty} \frac{x_1+\ldots+x_n}{n} = 0$. $\square$

Schließlich können wir das starke Gesetz der großen Zahlen (zweite Version) beweisen.
Beweis von Satz 10.7.1. Seien $X_1, X_2, \ldots : \Omega \to \mathbb{R}$ unabhängige, identisch verteilte Zufallsvariablen mit $\mathbb{E}|X_1| < \infty$. Wir zeigen, dass dann gilt:

$$\frac{S_n}{n} := \frac{X_1 + \ldots + X_n}{n} \xrightarrow[n\to\infty]{f.s.} \mathbb{E}[X_1].$$

Schritt 1. Definiere

$$X_n' := X_n \cdot \mathbb{1}_{|X_n|<n} \text{ und } X_n'' := X_n \cdot \mathbb{1}_{|X_n|\geq n},$$

dann ist $X_n' + X_n'' = X_n$. Außerdem sei

$$S_n' := X_1' + \ldots + X_n' \text{ und } S_n'' := X_1'' + \ldots + X_n''.$$

Es gilt:

(1) $X_1', X_2', \ldots$ sind unabhängig,
(2) $X_1'', X_2'', \ldots$ sind unabhängig.

Schritt 2. Unser Ziel ist es, zu zeigen, dass, dass

$$\frac{S_n}{n} - \mathbb{E}X_1 \xrightarrow[n\to\infty]{f.s.} 0.$$

Es gilt:

$$\frac{S_n}{n} - \mathbb{E}X_1 = \frac{S_n - \mathbb{E}S_n}{n} = \frac{S_n' - \mathbb{E}S_n'}{n} + \frac{S_n''}{n} - \frac{\mathbb{E}S_n''}{n} =: a_n + b_n - c_n$$

wobei a_n, b_n, c_n Zufallsvariablen sind. Nach Lemma 10.7.3 reicht es zu zeigen, dass

$$a_n, b_n, c_n \xrightarrow[n\to\infty]{f.s.} 0.$$

Schritt 3. Wir zeigen, dass $a_n = \frac{S_n' - \mathbb{E}S_n'}{n} \xrightarrow[n\to\infty]{f.s.} 0$, indem wir Satz 10.6.1 benutzen.

$$\begin{aligned}
\sum_{n\geq 1} \frac{\operatorname{Var} X_n'}{n^2} &\leq \sum_{n\geq 1} \frac{\mathbb{E}[X_n'^2]}{n^2} \\
&= \sum_{n\geq 1} \frac{1}{n^2} \cdot \mathbb{E}[X_n^2 \cdot \mathbb{1}_{|X_n|<n}] \\
&= \sum_{n\geq 1} \frac{1}{n^2} \cdot \mathbb{E}[X_1^2 \cdot \mathbb{1}_{|X_1|<n}] \quad \text{(gilt, da } X_n \text{ identisch verteilt)} \\
&= \sum_{n\geq 1} \frac{1}{n^2} \cdot \mathbb{E}\left[X_1^2 \cdot \sum_{k=1}^{n} \mathbb{1}_{A_k}\right],
\end{aligned}$$

wobei $A_k = \{k-1 \le X_1 < k\}$. Nun vertauschen wir die beiden Summen:

$$\begin{aligned}\sum_{n\ge1}\frac{\operatorname{Var} X_n'}{n^2} &\le \sum_{n\ge1}\frac{1}{n^2}\cdot\sum_{k=1}^{n}\mathbb{E}[X_1^2\cdot \mathbb{1}_{A_k}]\\ &= \sum_{k\ge1}\mathbb{E}[X_1^2\cdot \mathbb{1}_{A_k}]\cdot\sum_{n\ge k}\frac{1}{n^2}\\ &\le 2\cdot\sum_{k\ge1}\frac{1}{k}\cdot\mathbb{E}[X_1^2\cdot \mathbb{1}_{A_k}],\end{aligned}$$

wobei wir Lemma 10.7.5 benutzt haben: $\sum_{n\ge k}\frac{1}{n^2}\le\frac{2}{k}$. Es folgt, dass

$$\begin{aligned}\sum_{n\ge1}\frac{\operatorname{Var} X_n'}{n^2} &\le 2\cdot\sum_{k\ge1}\mathbb{E}\left[\frac{|X_1|}{k}\cdot|X_1|\cdot\mathbb{1}_{A_k}\right] \quad \left(\text{wobei } \frac{|X_1|}{k}\le 1, \text{ falls } A_k \text{ eintritt}\right)\\ &\le 2\cdot\sum_{k\ge1}\mathbb{E}[|X_1|\cdot\mathbb{1}_{A_k}]\\ &= 2\cdot\mathbb{E}\left[\sum_{k\ge1}|X_1|\cdot\mathbb{1}_{A_k}\right] \quad (\text{gilt mit monotoner Konvergenz})\\ &= 2\cdot\mathbb{E}|X_1|\\ &< \infty \quad (\text{gilt nach Voraussetzung})\end{aligned}$$

Mit Satz 10.6.1 folgt nun $a_n \xrightarrow[n\to\infty]{f.s.} 0$.

Schritt 4. Wir zeigen, dass $b_n = \frac{S_n''}{n} \xrightarrow[n\to\infty]{f.s.} 0$, indem wir den ersten Teil des Lemmas von Borel–Cantelli benutzen. Es gilt:

$$\begin{aligned}\sum_{n\ge1}\mathbb{P}[X_n''\ne0] &= \sum_{n\ge1}\mathbb{P}[|X_n|\ge n]\\ &= \sum_{n\ge1}\mathbb{P}[|X_1|\ge n] \quad (\text{gilt, da } X_n \text{ identisch verteilt})\\ &\le \mathbb{E}|X_1| \quad (\text{gilt mit Lemma 10.7.4})\\ &< \infty.\end{aligned}$$

Mit dem Lemma von Borel–Cantelli folgt nun:

$$\mathbb{P}\left[\limsup_{n\to\infty}\{X_n''\ne0\}\right] = \mathbb{P}[X_n''\ne 0 \text{ für unendlich viele } n] = 0.$$

Daraus folgt:

$$\mathbb{P}[X_n''\ne 0 \text{ für endlich viele } n] = \mathbb{P}[A] = 1.$$

Für $\omega\in A$ gilt dann, dass $X_n''(\omega)\ne0$ für endlich viele Werte von n. Also ist $S_n''(\omega)$ konstant ab irgendeinem n und damit gilt:

$$\lim_{n\to\infty}\frac{S_n''(\omega)}{n} = 0.$$

Da $\mathbb{P}[A] = 1$, folgt $b_n \xrightarrow[n\to\infty]{f.s.} 0$.

Schritt 5. Es bleibt noch zu zeigen, dass $c_n = \frac{\mathbb{E}S''_n}{n} \xrightarrow[n\to\infty]{f.s.} 0$. Es gilt:

$$\mathbb{E}X''_n = \mathbb{E}[X_n \cdot \mathbb{1}_{|X_n|\geq n}] = \mathbb{E}[X_1 \cdot \mathbb{1}_{|X_1|\geq n}] \quad \text{(da } X_i \text{ identisch verteilt)}.$$

Nun gilt für alle $\omega \in \Omega$:

$$|X_1(\omega)| \cdot \mathbb{1}_{|X_1(\omega)|\geq n} \xrightarrow[n\to\infty]{\text{monoton}} 0.$$

Mit dem Satz von der monotonen Konvergenz erhält man:

$$\mathbb{E}[|X_1| \cdot \mathbb{1}_{|X_1|\geq n}] \xrightarrow[n\to\infty]{} 0.$$

Es folgt:

$$|\mathbb{E}X''_n| = |\mathbb{E}[X_1 \cdot \mathbb{1}_{|X_1|\geq n}]| \leq \mathbb{E}[|X_1| \cdot \mathbb{1}_{|X_1|\geq n}] \xrightarrow[n\to\infty]{} 0 \quad \text{(da } |\mathbb{E}Z| \leq \mathbb{E}|Z|\text{)}.$$

Also gilt:

$$\mathbb{E}X''_n \xrightarrow[n\to\infty]{} 0.$$

Mit Lemma 10.7.6 folgt dann:

$$c_n = \frac{\mathbb{E}X''_1 + \ldots + \mathbb{E}X''_n}{n} \xrightarrow[n\to\infty]{f.s.} 0 \quad \text{(und sogar sicher)}.$$

Schritt 6. Fügt man die Ergebnisse für a_n, b_n, c_n zusammen, so erhält man:

$$\frac{S_n}{n} - \mathbb{E}X_1 = a_n + b_n + c_n \xrightarrow[n\to\infty]{f.s.} 0.$$

Das ist die gewünschte Konvergenz. □

10.8. Der Fall eines unendlichen Erwartungswerts

In allen Versionen des Gesetzes der großen Zahlen haben wir vorausgesetzt, dass der Erwartungswert existiert. Der nächste Satz zeigt, was passiert, wenn der Erwartungswert nicht existiert. In diesem Fall gilt das Gesetz der großen Zahlen nämlich nicht.

Satz 10.8.1. *Seien $X_1, X_2, \ldots$ unabhängige, identisch verteilte Zufallsvariablen mit $\mathbb{E}|X_1| = +\infty$ auf einem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbb{P})$. Dann gilt:*

$$\mathbb{P}\left[\lim_{n\to\infty} \frac{X_1 + \ldots + X_n}{n} \textit{ exitiert und ist endlich}\right] = 0.$$

Beweis. Schritt 1. Wir definieren das Ereignis

$$A := \left\{\omega \in \Omega : L(\omega) := \lim_{n\to\infty} \frac{S_n(\omega)}{n} \text{ existiert und ist endlich}\right\},$$

wobei wieder $S_n = X_1 + \ldots + X_n$. Dieses Ereignis ist messbar (Übung). Für $\omega \in A$ gilt nun:

$$\frac{X_n(\omega)}{n} = \frac{S_n(\omega)}{n} - \frac{S_{n-1}(\omega)}{n} = \frac{S_n(\omega)}{n} - \frac{S_{n-1}(\omega)}{n-1} \cdot \frac{n-1}{n} \xrightarrow[n\to\infty]{} 0,$$

denn
$$\frac{S_n(\omega)}{n} \underset{n\to\infty}{\longrightarrow} L(\omega), \quad \frac{S_{n-1}(\omega)}{n-1} \underset{n\to\infty}{\longrightarrow} L(\omega), \quad \frac{n-1}{n} \underset{n\to\infty}{\longrightarrow} 1.$$
Nun betrachten wir das Ereignis
$$B := \left\{\omega \in \Omega : \lim_{n\to\infty} \frac{X_n(\omega)}{n} = 0\right\}.$$
Wir haben gezeigt: falls $\omega \in A$, dann ist auch $\omega \in B$. Daraus folgt $A \subset B$.

Schritt 2. Als nächstes definieren wir eine Folge von Ereignissen:
$$C_n := \{\omega \in \Omega : |X_n| \geq n\}.$$
Daraus erhalten wir:
$$C = (\limsup_{n\to\infty} C_n)^c = \{\omega \in \Omega : |X_n(\omega)| \geq n \text{ für endlich viele } n\}.$$
Falls $\omega \in B$, dann ist auch $\omega \in C$. Daraus folgt $A \subset B \subset C$.

Schritt 3. Wir wollen zeigen, dass $\mathbb{P}[A] = 0$. Dafür reicht es nun zu zeigen, dass $\mathbb{P}[C] = 0$ bzw. $\mathbb{P}[\limsup C_n] = 1$. Wir benutzen dazu den zweiten Teil des Lemmas von Borel–Cantelli:
$$\begin{aligned}
\sum_{n\geq 1} \mathbb{P}[C_n] &= \sum_{n\geq 1} \mathbb{P}[|X_n| \geq n] \\
&= \sum_{n\geq 1} \mathbb{P}[|X_1| \geq n] \quad \text{(gilt, da } X_i \text{ identisch verteilt)} \\
&\geq \mathbb{E}|X_1| - 1 \quad \text{(gilt mit Lemma 10.7.4, da } |X_1| \geq 0) \\
&= \infty.
\end{aligned}$$
Da die Ereignisse C_n unabhängig sind, folgt mit dem zweiten Teil des Lemma von Borel–Cantelli, dass
$$\mathbb{P}[\limsup_{n\to\infty} C_n] = 1.$$
Daraus folgt dann $\mathbb{P}[C] = 0$. Da $A \subset C$, folgt auch $\mathbb{P}[A] = 0$. □

Beispiel 10.8.2. Seien $X_1, X_2, \ldots$ unabhängige, Cauchy-verteilte Zufallsvariablen mit Dichte $f_{X_k}(t) = \frac{1}{\pi} \cdot \frac{1}{1+t^2}$, $t \in \mathbb{R}$. Dann gilt für den Erwartungswert: $\mathbb{E}|X_k| = \infty$. Mit Satz 10.8.1 folgt nun, dass
$$\mathbb{P}\left[\lim_{n\to\infty} \frac{S_n}{n} \text{ existiert und ist endlich}\right] = 0.$$
Das starke Gesetz der großen Zahlen gilt also für Cauchy-verteilte Zufallsvariablen nicht.

10.9. Anwendungen des Gesetzes der großen Zahlen

Die nächsten 3 Beispiele sind Spezialfälle der sogenannten Monte-Carlo-Methode.

Berechnung von π. Man kann die Zahl π mit stochastischen Mitteln berechnen. Diese Methode ist jedoch sehr zeitaufwendig bzw. ungenau. Man erzeuge n Punkte bzw. Zufallsvektoren $Z_1, Z_2, \ldots, Z_n$, die unabhängig und gleichverteilt auf dem Quadrat $[-1, 1]^2$ sind.

Nun betrachtet man alle Punkte, die im Einheitskreis liegen, deren Bertag also kleiner 1 ist. Die Anzahl dieser Punkte ist

$$S_n := \sum_{k=1}^{n} \mathbb{1}_{|Z_k|\le 1}.$$

Für den Erwartungswert der Indikatorfunktion gilt

$$\mathbb{E}\mathbb{1}_{|Z_1|\le 1} = \mathbb{P}[|Z_1| \le 1] = \frac{\lambda(\text{Kreis})}{\lambda(\text{Quadrat})} = \frac{\pi}{4}.$$

Also dürfen wir das starke Gesetz der großen Zahlen benutzen:

$$\frac{1}{n}\sum_{k=1}^{n} \mathbb{1}_{|Z_k|\le 1} \xrightarrow[n\to\infty]{f.s.} \frac{\pi}{4}.$$

Um die Zahl π zu berechnen, zählen wir also die Punkte, die im Kreis liegen und teilen diese Anzahl durch die Anzahl aller Punkte. Bei sehr großem n sollte das Verhältnis die Zahl $\pi/4$ approximieren. Den Fehler dieser Approximation bestimmen wir später mit dem zentralen Grenzwertsatz.

Monte-Carlo-Integration. Man betrachte eine integrierbare Funktion $\varphi : [0,1]^d \to \mathbb{R}$. Das Integral der Funktion bezeichnen wir mit

$$I := \int_{[0,1]^d} \varphi(x)\mathrm{d}x = \int_0^1 \ldots \int_0^1 \varphi(x_1,\ldots,x_d)\mathrm{d}x_1\ldots\mathrm{d}x_d.$$

Für kleines d, etwa $d=1$ oder 2, kann man zur Berechnung des Integrals z.B. die folgende Formel benutzen:

$$I \approx \frac{1}{n}\cdot\sum_{k_1=0}^{n-1}\ldots\sum_{k_d=0}^{n-1}\varphi\left(\frac{k_1}{n},\ldots,\frac{k_d}{n}\right) \text{ für großes } n.$$

Wenn allerdings die Dimension d groß ist, so besteht die Summe auf der rechten Seite aus n^d Summanden, was sehr viel sein kann. Man denke z.B. an den Fall $d = 100$, wo selbst für $n = 2$ die Anzahl der Summanden $2^{100} > 10^{30}$ ist. In diesem Fall wird die oben beschriebene Methode ineffizient und man kann stattdessen die sogenannte Monte–Carlo Methode benutzen.

Man erzeuge n zufällige Punkte bzw. Zufallsvektoren $Z_1, Z_2, \ldots, Z_n$, die unabhängig und gleichverteilt auf dem "Quader" $[0,1]^d$ sind. Da die Funktion φ integrierbar ist, gilt

$$\mathbb{E}[\varphi(Z_1)] = \int_{\mathbb{R}^d}\varphi(x)\cdot f_{Z_1}(x)\mathrm{d}x = \int_{[0,1]^d}\varphi(x)\mathrm{d}x = I.$$

Also können wir das starke Gesetz der großen Zahlen benutzen:

$$\frac{1}{n}\sum_{k=1}^{n}\varphi(Z_k) \xrightarrow[n\to\infty]{f.s.} I.$$

Man kann also das Integral I durch das arithmetische Mittel $\frac{1}{n}\sum_{k=1}^{n}\varphi(Z_k)$ approximieren. Später werden wir mithilfe des zentralen Grenzwertsatzes den Fehler dieser Approximation bestimmen.

Empirische Definition der Wahrscheinlichkeit. Bei einem Experiment mit Grundmenge Ω möchte man die Wahrscheinlichkeit eines Ereignisses $A \subset \Omega$ bestimmen. Dazu wiederholt man das Experiment n-mal und erhält die Ausgänge $\omega_1, \ldots, \omega_n \in \Omega$. Die Anzahl der Augänge, die in A liegen, sei

$$N_n(A) = \sum_{k=1}^{n} \mathbb{1}_{\omega_k \in A}.$$

Wir können auf die Indikatorfunktionen $\mathbb{1}_{\omega_k \in A}$ das starke Gesetz der großen Zahlen anwenden:

$$\frac{N_n(A)}{n} = \frac{1}{n}\sum_{k=1}^{n} \mathbb{1}_{\omega_k \in A} \xrightarrow[n\to\infty]{f.s.} \mathbb{E}\mathbb{1}_{\omega_1 \in A} = \mathbb{P}[A].$$

In Worten: Die relative Häufigkeit eines Ereignisses konvergiert fast sicher gegen seine Wahrscheinlichkeit, wenn die Anzahl der Experimente gegen ∞ geht.

Beispiel 10.9.1 (Befragung). Wir betrachten eine Population von N Personen, die eine Frage mit "ja" oder "nein" beantworten. Die Zahl N sei bekannt. Es sei N_0 (bzw. N_1) die Anzahl der Personen, die die Frage mit "nein" (bzw. mit "ja") beantworten. Dabei sind N_0 und N_1 unbekannt. Es gilt $N_0 + N_1 = N$.
Wir kann man N_0 und N_1 bestimmen, ohne die ganze Population zu befragen? Es werden zufällig n Personen aus der Population mit Zurücklegen ausgewählt und befragt, wobei n groß aber viel kleiner als N ist. Es sei A_k das Ereignis "die k-te Person sagt ja". Nach dem starken Gesetz der großen Zahlen gilt

$$\frac{1}{n} \cdot \sum_{k=1}^{n} \mathbb{1}_{A_k} \xrightarrow[n\to\infty]{f.s.} \mathbb{E}\mathbb{1}_{A_1} = \mathbb{P}[A_1] = \frac{N_1}{N}.$$

Damit kann N_1 geschätzt werden.

Approximationssatz von Weierstraß. Als nächstes werden wir einen Satz aus der Analysis mit Mitteln aus der Wahrscheinlichkeitstheorie beweisen.

Satz 10.9.2 (Weierstraß). *Sei $f : [0,1] \to \mathbb{R}$ eine stetige Funktion. Dann kann man zu jedem gegebenen $\varepsilon > 0$ ein Polynom P mit*

$$\max_{x\in[0,1]} |f(x) - P(x)| \le \varepsilon$$

konstruieren. Das heißt, jede stetige Funktion auf einem Intervall kann beliebig genau durch Polynome approximiert werden.

Beweis. Schritt 1. Seien $X_1, X_2, \ldots$ unabhängige Zufallsvariablen mit $X_i \sim \operatorname{Bern}(p)$, d.h.

$$\mathbb{P}[X_i = 1] = p, \quad \mathbb{P}[X_i = 0] = 1 - p, \quad p \in [0,1].$$

Dann ist $S_n := X_1 + \ldots + X_n \sim \operatorname{Bin}(n,p)$, also $\mathbb{P}[S_n = k] = \binom{n}{k} p^k (1-p)^{n-k}$. Insbesondere gilt für den Erwartungswert und die Varianz

$$\mathbb{E}S_n = n\mathbb{E}X_1 = np, \quad \operatorname{Var} S_n = np(1-p).$$

Wir definieren eine Funktionenfolge:

$$f_n(p) := \mathbb{E}\left[f\left(\frac{S_n}{n}\right)\right] = \sum_{k=0}^{n} f\left(\frac{k}{n}\right)\mathbb{P}[S_n = k] = \sum_{k=0}^{n} f\left(\frac{k}{n}\right)\binom{n}{k}p^k(1-p)^{n-k}.$$

Somit ist jedes f_n ein Polynom vom Grad n in p. Es wird *Bernstein-Polynom* genannt.

SCHRITT 2. Die Idee des Beweises ist nun: für $n \to \infty$ besagt das Gesetz der großen Zahlen, dass $\frac{S_n}{n} \approx p$ und damit auch $f\left(\frac{S_n}{n}\right) \approx f(p)$. Wendet man den Erwartungswert auf die beiden Funktionswerte an, so erhält man $f_n(p) \approx f(p)$. Somit approximiert das Polynom f_n die Funktion f, jedenfalls wenn n groß ist. Nun präzisieren wir diese Idee.

SCHRITT 3. Da f stetig ist, folgt:

(1) f ist beschränkt, d.h. es existiert $M > 0$ mit $|f(x)| < M$ für alle $x \in [0,1]$.
(2) f ist sogar gleichmäßig stetig, d.h. für jedes $\varepsilon > 0$ existiert ein $\delta > 0$ so dass für alle $x, y \in [0,1]$ mit $|x-y| \leq \delta$ gilt, dass $|f(x) - f(y)| < \varepsilon$.

SCHRITT 4. Sei $\varepsilon > 0$ fest vorgegeben. Dazu konstruieren wir ein $\delta > 0$ wie in Schritt 3. Es gilt

$$|f_n(p) - f(p)| = \left|\mathbb{E}\left[f\left(\frac{S_n}{n}\right) - f(p)\right]\right| \leq \mathbb{E}\left|f\left(\frac{S_n}{n}\right) - f(p)\right| \leq \mathbb{E}\left[\varepsilon + 2M \cdot \mathbb{1}_{\left|\frac{S_n}{n} - p\right| > \delta}\right],$$

wobei wir die Fallunterscheidung $|\frac{S_n}{n} - p| \leq \delta$ bzw. $|\frac{S_n}{n} - p| > \delta$ gemacht haben. Im ersten Fall haben wir die gleichmäßige Stetigkeit von f benutzt, im zweiten Fall die Beschränktheit von f. Es gilt also

$$|f_n(p) - f(p)| \leq \varepsilon + 2M \cdot \mathbb{P}\left[\left|\frac{S_n}{n} - p\right| > \delta\right] \leq \varepsilon + 2M \cdot \frac{\operatorname{Var}\frac{S_n}{n}}{\delta^2},$$

wobei wir die Tschebyschew-Ungleichung benutzt haben. Mit der Formel $\operatorname{Var} S_n = np(1-p)$ erhalten wir nun für hinreichend großes n

$$|f_n(p) - f(p)| \leq \varepsilon + 2M \cdot \frac{np(1-p)}{n^2\delta^2} \leq \varepsilon + 2M \cdot \frac{1}{n\delta^2} \leq 2\varepsilon.$$

Dabei folgt die letzte Ungleichung aus der Tatsache, dass $2M \cdot \frac{1}{n\delta^2}$ für $n \to \infty$ gegen 0 konvergiert, also $< \varepsilon$ für hinreichend großes n ist. Wir haben ein approximierendes Polynom mit einem Approximationsfehler von 2ε konstruiert. Da jedoch ε beliebig war, ist der Satz damit bewiesen. □

Normale Zahlen. Sei $\omega \in (0,1)$ eine reelle Zahl mit der Dezimaldarstellung

$$\omega = 0.x_1x_2x_3\ldots.$$

Dabei ist $x_k \in \{0, \ldots, 9\}$ die k-te Ziffer von ω.

DEFINITION 10.9.3. Eine Zahl $\omega \in (0,1)$ heißt *normal*, wenn für jedes $l \in \mathbb{N}$ und für jede Ziffernfolge $(a_1, \ldots, a_l) \in \{0, \ldots, 9\}^l$ der Länge l gilt:

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n} \mathbb{1}_{X_{i+1}=a_1,\ \ldots, X_{i+l}=a_l} = 10^{-l}.$$

In Worten: Jede Ziffernfolge der Länge l kommt in der Dezimaldarstellung von ω mit Häufigkeit 10^{-l} vor, und das gilt für jedes l.

Definition 10.9.4. Eine Zahl $\omega \in (0,1)$ heißt *einfach normal*, wenn die obige Bedingung für $l = 1$ erfüllt ist, d.h., wenn in der Dezimaldarstellung von ω jede Ziffer mit einer Häufigkeit von $1/10$ vorkommt.

Beispiel 10.9.5. Die Zahl $\frac{1}{3} = 0{,}3333\ldots$ ist nicht normal. Allgemein sind rationale Zahlen nicht normal, da sie eine periodische Dezimalbruchentwicklung besitzen.

Beispiel 10.9.6. Unbewiesene Vermutung: Die Zahlen

$$\begin{aligned}\pi &= 3.14159265358979323846264338327950288419716939937 51\ldots,\\ e &= 2.71828182845904523536028747135266249775724709370 00\ldots,\\ \log 2 &= 0.69314718055994530941723212145817656807550013436 02\ldots\end{aligned}$$

sind normal.

Beispiel 10.9.7. Unbewiesene Vermutung: Jede algebraische Zahl ist normal, sofern sie nicht rational ist. Zum Beispiel ist die Zahl

$$\sqrt{2} = 1.41421356237309504880168872420969807856967187537 69\ldots$$

normal.

Beispiel 10.9.8. Die Champernowne-Zahl entsteht durch die Aneinanderreihung aller natürlichen Zahlen in ihrer Dezimaldarstellung:

$$0.123\ldots 91011\ldots 192021\ldots 99100101\ldots$$

Man kann zeigen, dass diese Zahl normal ist.

Satz 10.9.9 (Borel, 1909). *Sei $A \subset [0,1]$ die Menge der normalen Zahlen. Dann ist das Lebesgue-Maß von A gleich 1. D.h., fast jede Zahl ist normal.*

Beweis. Schritt 1. Wir betrachten den Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, \mathbb{P})$, wobei $\Omega = (0,1)$, $\mathcal{F}$ die σ-Algebra der Borel-Teilmengen von $(0,1)$ und $\mathbb{P}$ das Lebesgue-Maß ist. Für alle $k \in \mathbb{N}$ sei

$$X_k(\omega) = \omega_k$$

die k-te Ziffer von $\omega \in (0,1)$. Dann ist X_k eine Zufallsvariable auf Ω.
Wir zeigen nun, dass die Zufallsvariablen $X_1, X_2, \ldots$ unabhängig und gleichverteilt auf $\{0, \ldots, 9\}$ sind. Betrachte für ein beliebiges $m \in \mathbb{N}$ und eine Ziffernfolge $(b_1, \ldots, b_m) \in \{0, \ldots, 9\}^m$ die Menge

$$\{\omega \in (0,1) : X_1(\omega) = b_1, \ \ldots, X_m(\omega) = b_m\}.$$

Diese Menge ist ein Intervall mit Endpunkten

$$0.b_1b_2\ldots b_m000\ldots \ \text{ und } 0.b_1b_2\ldots b_m999\ldots.$$

Die Länge des Intervalls ist somit 10^{-m}. Es gilt somit

$$\mathbb{P}[X_1 = b_1, \ldots, X_m = b_m] = 10^{-m}.$$

Sei nun $i \in \{1, \ldots, m\}$ fest und betrachte die Menge

$$\{\omega \in (0,1) : X_i(\omega) = b_i\}.$$

Das ist die Menge aller ω, die in der Dezimaldarstellung an der i-ten Stelle die Ziffer b_i haben. Für die ersten $i-1$ Stellen gibt es 10^{i-1} Möglichkeiten, die Ziffern zu wählen. Diese Menge ist somit eine Vereinigung von 10^{i-1} Intervallen der Länge 10^{-i}. Somit ergibt sich für die Wahrscheinlichkeit dieser Menge:

$$\mathbb{P}[X_i = b_i] = 10^{i-1} \cdot 10^{-i} = \frac{1}{10}.$$

Daraus folgt, dass die Zufallsvariable X_i gleichverteilt auf $\{0, \dots, 9\}$ ist. Außerdem gilt für alle $m \in \mathbb{N}$ und alle $(b_1, \dots, b_m) \in \{0, \dots, 9\}^m$

$$\mathbb{P}[X_1 = b_1, \dots, X_m = b_m] = 10^{-m} = \mathbb{P}[X_1 = b_1] \cdot \ldots \cdot \mathbb{P}[X_m = b_m].$$

Daraus folgt, dass die Zufallsvariablen $X_1, X_2, \dots$ unabhängig sind.

Schritt 2. Die Idee ist nun, das starke Gesetz der großen Zahlen folgendermaßen zu benutzen. Es gilt z.B. für jede Ziffer $a \in \{0, 1, \dots, 9\}$:

$$\frac{1}{n} \sum_{k=1}^{n} \mathbb{1}_{X_k = a} \xrightarrow[n\to\infty]{f.s.} \mathbb{E}[\mathbb{1}_{X_1 = a}] = \mathbb{P}[X_1 = a] = \frac{1}{10}.$$

Das zeigt schon mal, dass fast jede Zahl einfach normal ist. Wir werden nun diese Argumentation auf Ziffernfolgen beliebiger Länge l ausweiten.

Schritt 3. Wir betrachten eine Ziffernfolge $(a_1, \dots, a_l) \in \{0, \dots, 9\}^l$ der Länge l. Wir definieren die Zufallsvariablen

$$Z_i^{(0)} := \mathbb{1}_{X_{i\cdot l+1} = a_1,\ \dots,\ X_{i \cdot l + l} = a_l}, \quad i = 0, 1, 2, \dots,$$

und, allgemeiner,

$$Z_i^{(j)} := \mathbb{1}_{X_{i\cdot l+j+1} = a_1,\ \dots,\ X_{i \cdot l + j + l} = a_l}, \quad i = 0, 1, 2, \dots, \quad j = 0, \dots, l-1.$$

Diese Zufallsvariablen sind folgendermaßen zu verstehen: Man unterteile alle Ziffern einer Zahl ω in "Pakete" der Länge l, wobei das erste "Paket" an der Stelle $j+1$ startet, und betrachte das i-te "Paket". Entspricht dieses der Ziffernfolge $(a_1, \dots, a_l)$, so ist $Z_i^{(j)} = 1$. Sonst ist $Z_i^{(j)} = 0$.

Für alle $j \in \{0, \dots, l-1\}$ sind die Zufallsvariablen $Z_0^{(j)}, Z_1^{(j)}, Z_2^{(j)}, \dots$ unabhängig und identisch verteilt. Mit dem starken Gesetz der großen Zahlen erhalten wir folgendes: Es gibt eine Menge $B(l, j, a_1, \dots, a_l) \subset (0, 1)$ mit Lebesgue-Maß 0, so dass für alle $\omega \notin B(l, j, a_1, \dots, a_l)$ gilt:

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n} Z_i^{(j)}(\omega) = \mathbb{E} Z_0^{(j)} = 10^{-l}.$$

Die Häufigkeit der Ziffernfolge $(a_1, \dots, a_l)$ ist also 10^{-l}, allerdings werden dabei nur "Pakete" berücksichtigt, die an den Stellen $j+1, j+1+l, j+1+2l, \dots$ anfangen. Da dies aber für jedes $j \in \{0, \dots, l-1\}$ gilt, erhalten wir, dass für alle $\omega \notin \bigcup_{j=0}^{l-1} B(l, j, a_1, \dots, a_l)$

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n} \mathbb{1}_{X_{k+1} = a_1,\ \dots,\ X_{k+l} = a_l}(\omega) = 10^{-l}.$$

Nun definieren wir die "Ausnahmemenge"

$$B := \bigcup_{l\in\mathbb{N}} \bigcup_{j=0}^{l-1} \bigcup_{a_1,\ldots,a_l\in\{0,\ldots,9\}} B(l,j,a_1,\ldots,a_l) \subset (0,1).$$

Jede Zahl ω, die nicht in B liegt, ist normal. Da B eine abzählbare Vereinigung von Nullmengen ist, ist das Lebesgue-Maß von B gleich 0. Das Komplement von B hat also Lebesgue-Maß 1, und alle Zahlen darin sind normal. Somit ist das Lebesgue-Maß der Menge der normalen Zahlen gleich 1. □

Erneuerungssatz. Man betrachte ein Gerät (etwa eine Glühbirne), dessen Lebensdauer eine Zufallsvariable $X_1 > 0$ ist. Sobald das Gerät kaputt geht, wird es durch ein neues Gerät ersetzt, dessen Lebensdauer eine Zufallsvariable $X_2 > 0$ ist. Sobald das zweite Gerät kaputt geht, wird es durch ein drittes Gerät ersetzt, usw. Wir nehmen an, dass $X_1, X_2, \ldots$ (die Lebensdauern der Geräte) unabhängige Zufallsvariablen sind, da ein Gerät nichts von der Lebensdauer eines anderen wissen kann. Außerdem nehmen wir an, dass $X_1, X_2, \ldots$ identisch verteilt sind, etwa weil die Geräte vom gleichen Hersteller sind bzw. die gleiche Qualität haben.
Seien also $X_1, X_2, \ldots : \Omega \to \mathbb{R}$ unabhängige und identisch verteilte Zufallsvariablen mit positiven Werten, d.h. $\mathbb{P}[X_k > 0] = 1$. Die Zufallsvariable $S_n = X_1 + \ldots + X_n$ heißt die n-te *Erneuerungszeit*. Zum Zeitpunkt S_n geht das n-te Gerät kaputt und wird durch das $n+1$-te Gerät ersetzt. Es gilt $0 < S_1 < S_2 < \ldots$. Mit

$$N(T) = \sum_{n=1}^{\infty} \mathbb{1}_{S_n < T}, \qquad T > 0,$$

bezeichnen wir die Anzahl der Erneuerungen im Zeitintervall $(0, T)$.

Satz 10.9.10 (Erneuerungssatz). *Es gilt:*

$$\frac{N(T)}{T} \xrightarrow[T\to\infty]{f.s.} \frac{1}{\mathbb{E}X_1}.$$

Wir werden den Satz nicht beweisen. Die Aussage ist allerdings sehr intuitiv: die Lebensdauer eines Geräts ist im Durchschnitt gleich $\mathbb{E}X_1$, also verbraucht man im Zeitintervall $(0, T)$ ungefähr $T/\mathbb{E}X_1$ Geräte, jedenfalls dann, wenn T groß ist. Es gilt also die Approximation

$$\frac{N(T)}{T} \approx \frac{1}{\mathbb{E}X_1}.$$