Chapter IV

Limit Theorems

Given: a sequence of random variables X_n, $n \in \mathbb{N}$, on a probability space $(\Omega, \mathfrak{A}, P)$. Put

$$S_n = \sum_{i=1}^{n} X_i, \qquad n \in \mathbb{N}.$$

For instance, S_n might be the cumulative gain after n trials or (one of the coordinates of) the position of a particle after n collisions.

Question: Convergence of S_n/a_n for suitable weights $0 < a_n \uparrow \infty$ in a suitable sense? Particular case: $a_n = n$.

1 Zero-One Laws

Definition 1. For σ-algebras $\mathfrak{A}_n \subset \mathfrak{A}$, $n \in \mathbb{N}$, the corresponding *tail σ-algebra* is

$$\mathfrak{A}_\infty = \bigcap_{n \in \mathbb{N}} \sigma\Big(\bigcup_{m \geq n} \mathfrak{A}_m \Big),$$

and $A \in \mathfrak{A}_\infty$ is called a *tail (terminal) event*.

Example 1. Let $\mathfrak{A}_n = \sigma(X_n)$. Put $\mathfrak{C} = \bigotimes_{i=1}^{\infty} \mathfrak{B}$. Then

$$\mathfrak{A}_\infty = \bigcap_{n \in \mathbb{N}} \sigma(\{X_m : m \geq n\})$$

and

$$A \in \mathfrak{A}_\infty \quad \Leftrightarrow \quad \forall\, n \in \mathbb{N}\ \exists\, C \in \mathfrak{C} : \ A = \{(X_n, X_{n+1}, \dots) \in C\}.$$

For instance,

$$\{(S_n)_{n \in \mathbb{N}} \text{ converges}\}, \ \{(S_n/a_n)_{n \in \mathbb{N}} \text{ converges}\} \in \mathfrak{A}_\infty,$$

and the function $\liminf_{n\to\infty} S_n/a_n$ is $\mathfrak{A}_\infty$-$\overline{\mathfrak{B}}$-measurable. However, S_n as well as $\liminf_{n\to\infty} S_n$ are not $\mathfrak{A}_\infty$-$\overline{\mathfrak{B}}$-measurable, in general. Analogously for the lim sup's.

Theorem 1 (Kolmogorov's Zero-One Law)**.** Let $(\mathfrak{A}_n)_{n\in\mathbb{N}}$ be an independent sequence of σ-algebras $\mathfrak{A}_n \subset \mathfrak{A}$. Then

$$\forall\, A \in \mathfrak{A}_\infty : \; P(A) \in \{0,1\}\,.$$

Proof. We show that $\mathfrak{A}_\infty$ and $\mathfrak{A}_\infty$ are independent (*terminology*), which implies $P(A) = P(A) \cdot P(A)$ for every $A \in \mathfrak{A}_\infty$. Put

$$\overline{\mathfrak{A}}_n = \sigma(\mathfrak{A}_1 \cup \cdots \cup \mathfrak{A}_n).$$

Note that $\mathfrak{A}_\infty \subset \sigma(\mathfrak{A}_{n+1} \cup \dots)$. By Corollary III.5.1 and Remark III.5.1.(i)

$$\overline{\mathfrak{A}}_n, \mathfrak{A}_\infty \text{ independent},$$

and therefore $\bigcup_{n\in\mathbb{N}} \overline{\mathfrak{A}}_n$ and $\mathfrak{A}_\infty$ are independent, too. Thus, by Theorem III.5.1,

$$\sigma\Big(\bigcup_{n\in\mathbb{N}} \overline{\mathfrak{A}}_n\Big), \mathfrak{A}_\infty \text{ independent.}$$

Finally,

$$\mathfrak{A}_\infty \subset \sigma\Big(\bigcup_{n\in\mathbb{N}} \mathfrak{A}_n\Big) = \sigma\Big(\bigcup_{n\in\mathbb{N}} \overline{\mathfrak{A}}_n\Big).$$

□

Corollary 1. Let $X \in \overline{\mathfrak{Z}}(\Omega, \mathfrak{A}_\infty)$. Under the assumptions of Theorem 1, X is constant P-a.s.

Remark 1. Assume that $(X_n)_{n\in\mathbb{N}}$ is independent. Then

$$P(\{(S_n)_{n\in\mathbb{N}} \text{ converges}\}),\; P(\{(S_n/a_n)_{n\in\mathbb{N}} \text{ converges}\}) \in \{0,1\}.$$

In case of convergence P-a.s., $\lim_{n\to\infty} S_n/a_n$ is constant P-a.s.

Definition 2. Let $A_n \in \mathfrak{A}$ for $n \in \mathbb{N}$. Then

$$\liminf_{n\to\infty} A_n = \bigcup_{n\in\mathbb{N}} \bigcap_{m\geq n} A_m, \qquad \limsup_{n\to\infty} A_n = \bigcap_{n\in\mathbb{N}} \bigcup_{m\geq n} A_m.$$

Remark 2.

(i) $\Big(\liminf_{n\to\infty} A_n\Big)^c = \limsup_{n\to\infty} A_n^c$.

(ii) $P\Big(\liminf_{n\to\infty} A_n\Big) \leq \liminf_{n\to\infty} P(A_n) \leq \limsup_{n\to\infty} P(A_n) \leq P\Big(\limsup_{n\to\infty} A_n\Big)$.

(iii) If $(A_n)_{n\in\mathbb{N}}$ is independent, then $P\Big(\limsup_{n\to\infty} A_n\Big) \in \{0,1\}$ (*Borel's Zero-One Law*).

Proof: Übung 10.1

Theorem 2 (Borel-Cantelli Lemma)**.** Let $A = \limsup_{n\to\infty} A_n$ with $A_n \in \mathfrak{A}$.

(i) If $\sum_{n=1}^{\infty} P(A_n) < \infty$ then $P(A) = 0$.

(ii) If $\sum_{n=1}^{\infty} P(A_n) = \infty$ and $(A_n)_{n\in\mathbb{N}}$ is independent, then $P(A) = 1$.

Proof. Ad (i):

$$P(A) \le P\Big(\bigcup_{m\ge n} A_m\Big) \le \sum_{m=n}^{\infty} P(A_m).$$

By assumption, the right-hand side tends to zero as n tends to ∞.
Ad (ii): We have

$$P(A^c) = P(\liminf_{n\to\infty} A_n^c) \le \sum_{n=1}^{\infty} P\Big(\bigcap_{m\ge n} A_m^c\Big).$$

Use $1 - x \le \exp(-x)$ for $x \ge 0$ to obtain

$$P\Big(\bigcap_{m=n}^{\ell} A_m^c\Big) = \prod_{m=n}^{\ell}(1 - P(A_m)) \le \prod_{m=n}^{\ell} \exp(-P(A_m)) = \exp(-\sum_{m=n}^{\ell} P(A_m)).$$

By assumption, the right-hand side tends to zero as ℓ tends to ∞. Thus $P(A^c) = 0$. □

Example 2. A fair coin is tossed an infinite number of times. Determine the probability that 0 occurs twice in a row infinitely often. Model: $(X_n)_{n\in\mathbb{N}}$ is independent and

$$P(\{X_n = 0\}) = P(\{X_n = 1\}) = 1/2, \qquad n \in \mathbb{N}.$$

Put

$$A_n = \{X_n = X_{n+1} = 0\}.$$

Then $(A_{2n})_{n\in\mathbb{N}}$ is independent and $P(A_{2n}) = 1/4$. Thus $P(\limsup_{n\to\infty} A_n) = 1$.

Remark 3. A stronger version of Theorem 2.(ii) requires only pairwise independence, see Bauer (1996, p. 70).

Example 3. Let $(X_n)_{n\in\mathbb{N}}$ be independent with

$$P(\{X_n = 1\}) = p = 1 - P(\{X_n = -1\}), \qquad n \in \mathbb{N},$$

with some constant $p \in [0, 1]$. Put

$$A = \limsup_{n\to\infty} \{S_n = 0\},$$

and note that

$$A \notin \mathfrak{A}_\infty = \bigcap_{n\in\mathbb{N}} \sigma(\{X_m : m \ge n\}).$$

Clearly

$$1/2 \cdot (S_n + n) \sim B(n, p).$$

Use Stirling's Formula

$$n! \approx \left(\frac{n}{e}\right)^n \sqrt{2\pi n}$$

to obtain

$$P(\{S_{2n} = 0\}) = \binom{2n}{n} \cdot p^n \cdot (1-p)^n \approx \frac{r^n}{\sqrt{\pi n}},$$

where $r = 4p \cdot (1-p) \in [0,1]$.
Suppose that

$$p \neq 1/2.$$

Then $r < 1$, and therefore

$$\sum_{n=0}^{\infty} P(\{S_n = 0\}) = \sum_{n=0}^{\infty} P(\{S_{2n} = 0\}) < \infty.$$

The Borel-Cantelli Lemma implies

$$P(A) = 0.$$

Suppose that

$$p = 1/2.$$

Then

$$\sum_{n=0}^{\infty} P(\{S_n = 0\}) = \sum_{n=0}^{\infty} P(\{S_{2n} = 0\}) = \infty,$$

but $(\{S_n = 0\})_{n\in\mathbb{N}}$ is not independent. Using the Central Limit Theorem (De Moivre-Laplace), one can show that $P(A) = 1$, see Übung 10.2.

2 Strong Law of Large Numbers

Definition 1. $(X_n)_{n\in\mathbb{N}}$ *independent and identically distributed (i.i.d.)* if $(X_n)_{n\in\mathbb{N}}$ is independent and

$$\forall\, n \in \mathbb{N}: \ P_{X_n} = P_{X_1}.$$

Throughout this section: $(X_n)_{n\in\mathbb{N}}$ independent.
Put

$$C = \big\{(S_n)_{n\in\mathbb{N}} \text{ converges in } \mathbb{R}\big\}.$$

By Remark 1, $P(C) \in \{0,1\}$.
First we provide sufficient conditions for $P(C) = 1$ to hold.

Theorem 1 (Hajek-Rényi inequality)**.** If

$$b_1 \geq b_2 \geq \cdots \geq b_n > 0$$

and

$$\forall\, i \in \{1,\dots,n\}: \ X_i \in \mathfrak{L}^2 \wedge \mathrm{E}(X_i) = 0,$$

then

$$P\Big(\Big\{\sup_{1\le k\le n} b_k \cdot |S_k| \ge 1\Big\}\Big) \le \sum_{i=1}^n b_i^2 \cdot \mathrm{Var}(X_i).$$

In particular, for $b_1 = \dots = b_n = 1/\varepsilon > 0$ *(Kolmogorov's inequality)*

$$P\Big(\Big\{\sup_{1\le k\le n} |S_k| \ge \varepsilon\Big\}\Big) \le \frac{1}{\varepsilon^2} \cdot \mathrm{Var}(S_n).$$

Proof. See Gänssler, Stute (1977, p. 98) for the Hajek-Rényi inequality. Here: the Kolmogorov inequality. Let $1 \le k \le n$. We show that

$$\forall\, B \in \sigma(\{X_1,\dots,X_k\}) : \int_B S_k^2 \, dP \le \int_B S_n^2 \, dP. \tag{1}$$

Note that

$$S_n^2 = (S_n - S_k)^2 + 2S_nS_k - S_k^2 = (S_n - S_k)^2 + 2S_k(S_n - S_k) + S_k^2.$$

Moreover, for $B \in \sigma(\{X_1,\dots,X_k\})$,

$$1_B \cdot S_k \text{ is } \sigma(\{X_1,\dots,X_k\})\text{-}\mathfrak{B}\text{-measurable,}$$
$$S_n - S_k \text{ is } \sigma(\{X_{k+1},\dots,X_n\})\text{-}\mathfrak{B}\text{-measurable,}$$

see Theorem II.2.8. Use Theorem III.5.4 to obtain

$$1_B \cdot S_k,\ S_n - S_k \text{ independent.}$$

Hence Theorem III.5.6 yields

$$\mathrm{E}(1_B \cdot S_k \cdot (S_n - S_k)) = \mathrm{E}(1_B \cdot S_k) \cdot \mathrm{E}(S_n - S_k) = 0,$$

and thereby

$$\mathrm{E}(1_B \cdot S_n^2) \ge 2 \cdot \mathrm{E}(1_B \cdot S_k \cdot (S_n - S_k)) + \mathrm{E}(1_B \cdot S_k^2) = \mathrm{E}(1_B \cdot S_k^2).$$

This completes the proof of (1).

Put

$$A_k = \bigcap_{\ell=1}^{k-1} \{|S_\ell| < \varepsilon\} \cap \{|S_k| \ge \varepsilon\}.$$

Then $A_k \in \sigma(\{X_1,\dots,X_k\})$, and by (1)

$$\begin{aligned}\varepsilon^2 \cdot P\Big(\Big\{\sup_{1\le k\le n} |S_k| \ge \varepsilon\Big\}\Big) &= \varepsilon^2 \cdot \sum_{k=1}^n P(A_k) \le \sum_{k=1}^n \int_{A_k} S_k^2 \, dP \\ &\le \sum_{k=1}^n \int_{A_k} S_n^2 \, dP \le \int_\Omega S_n^2 \, dP = \mathrm{Var}(S_n).\end{aligned}$$

□

Theorem 2. If

$$\forall\, n \in \mathbb{N}: \; X_n \in \mathfrak{L}^2 \wedge \mathrm{E}(X_n) = 0$$

and

$$\sum_{i=1}^{\infty} \mathrm{Var}(X_i) < \infty,$$

then

$$P(C) = 1.$$

Proof. Clearly

$$\omega \in C \quad \Leftrightarrow \quad \forall\, \varepsilon > 0 \; \exists\, n \in \mathbb{N} \; \forall\, k \in \mathbb{N} \; |S_{n+k}(\omega) - S_n(\omega)| < \varepsilon.$$

Put

$$M = \inf_{n \in \mathbb{N}} \sup_{k \in \mathbb{N}} |S_{n+k} - S_n|.$$

Then

$$C = \{M = 0\}.$$

Let $\varepsilon > 0$. For every $n \in \mathbb{N}$

$$\{M > \varepsilon\} \subset \Big\{ \sup_{k \in \mathbb{N}} |S_{n+k} - S_n| > \varepsilon \Big\},$$

and

$$\Big\{ \sup_{1 \le k \le r} |S_{n+k} - S_n| > \varepsilon \Big\} \uparrow \Big\{ \sup_{k \in \mathbb{N}} |S_{n+k} - S_n| > \varepsilon \Big\}$$

as r tends to ∞. Hence

$$P(\{M > \varepsilon\}) \le \lim_{r \to \infty} P\Big(\Big\{ \sup_{1 \le k \le r} |S_{n+k} - S_n| > \varepsilon \Big\} \Big),$$

and Kolmogorov's inequality yields

$$P\Big(\Big\{ \sup_{1 \le k \le r} |S_{n+k} - S_n| > \varepsilon \Big\} \Big) \le \frac{1}{\varepsilon^2} \cdot \sum_{i=n+1}^{n+r} \mathrm{Var}(X_i) \le \frac{1}{\varepsilon^2} \cdot \sum_{i=n+1}^{\infty} \mathrm{Var}(X_i).$$

Thus $P(\{M > \varepsilon\}) = 0$ for every $\varepsilon > 0$, which implies $P(\{M > 0\}) = 0$. □

Example 1. Let $(Y_n)_{n \in \mathbb{N}}$ be i.i.d. with $P_{Y_1} = 1/2 \cdot (\varepsilon_1 + \varepsilon_{-1})$. Then $\mathrm{E}(Y_n) = 0$ and $\mathrm{Var}(Y_n) = 1$, so that $\sum_{i=1}^{\infty} Y_i \cdot \frac{1}{i}$ converges P-a.s.

In the sequel: $0 < a_n \uparrow \infty$.
We now study convergence almost surely of $(S_n/a_n)_{n \in \mathbb{N}}$.

Lemma 1 (Kronecker's Lemma)**.** For every sequence $(x_n)_{n \in \mathbb{N}}$ in $\mathbb{R}$

$$\sum_{i=1}^{\infty} \frac{x_i}{a_i} \text{ converges} \quad \Rightarrow \quad \lim_{n \to \infty} \frac{1}{a_n} \cdot \sum_{i=1}^{n} x_i = 0.$$

Proof. Put $c = \sum_{i=1}^{\infty} x_i/a_i$ and $c_n = \sum_{i=1}^{n} x_i/a_i$. It is straightforward to verify that

$$\frac{1}{a_n} \cdot \sum_{i=1}^{n} x_i = c_n - \frac{1}{a_n} \cdot \sum_{i=2}^{n} (a_i - a_{i-1}) \cdot c_{i-1}.$$

Moreover, since $a_{i-1} \leq a_i$ and $\lim_{i\to\infty} a_i = \infty$,

$$c = \lim_{n\to\infty} \frac{1}{a_n} \cdot \sum_{i=2}^{n} (a_i - a_{i-1}) \cdot c_{i-1}.$$

□

Theorem 3 (Strong Law of Large Numbers, $\mathfrak{L}^2$ Case)**.** If

$$\forall\, n \in \mathbb{N}: \; X_n \in \mathfrak{L}^2 \quad \wedge \quad \sum_{i=1}^{\infty} \frac{1}{a_i^2} \cdot \mathrm{Var}(X_i) < \infty \tag{2}$$

then

$$\frac{1}{a_n} \cdot \sum_{i=1}^{n} (X_i - \mathrm{E}(X_i)) \xrightarrow{P\text{-a.s.}} 0.$$

Proof. Put $Y_n = 1/a_n \cdot (X_n - \mathrm{E}(X_n))$. Then $\mathrm{E}(Y_n) = 0$ and $(Y_n)_{n\in\mathbb{N}}$ is independent. Moreover,

$$\sum_{i=1}^{\infty} \mathrm{Var}(Y_i) = \sum_{i=1}^{\infty} \frac{1}{a_i^2} \cdot \mathrm{Var}(X_i) < \infty.$$

Thus $\sum_{i=1}^{\infty} Y_i$ converges P-a.s. due to Theorem 2. Apply Lemma 1. □

Remark 1. In particular, if $(X_n)_{n\in\mathbb{N}}$ is i.i.d. and $X_1 \in \mathfrak{L}^2$, then Theorem 3 with $a_n = n$ implies

$$\frac{1}{n} \cdot \sum_{i=1}^{n} X_i \xrightarrow{P\text{-a.s.}} \mathrm{E}(X_1),$$

see Einführung in die Stochastik. In fact, this conclusion already holds if $X_1 \in \mathfrak{L}^1$, see Theorem 4 below.

Remark 2. Assume

$$\sup_{n\in\mathbb{N}} \mathrm{Var}(X_n) < \infty.$$

Then another possible choice of a_n in Theorem 3 is

$$a_n = \sqrt{n} \cdot (\log n)^{1/2+\varepsilon}$$

for any $\varepsilon > 0$, and we have

$$\lim_{n\to\infty} \frac{S_n - \mathrm{E}(S_n)}{a_n} = 0 \;\; P\text{-a.s.}$$

Note that $\lim_{n\to\infty} a_n/n = 0$. Precise description of the fluctuation of $S_n(\omega)$ for P-a.e. $\omega \in \Omega$: law of the iterated logarithm, see Section 6. See also `Übung` 10.2.

Lemma 2. Let $U_i,\ V_i,\ W \in \mathfrak{Z}(\Omega, \mathfrak{A})$ such that

$$\sum_{i=1}^{\infty} P(\{U_i \neq V_i\}) < \infty.$$

Then

$$\frac{1}{n} \cdot \sum_{i=1}^{n} U_i \xrightarrow{P\text{-a.s.}} W \quad \Leftrightarrow \quad \frac{1}{n} \cdot \sum_{i=1}^{n} V_i \xrightarrow{P\text{-a.s.}} W.$$

Proof. The Borel-Cantelli Lemma implies $P(\limsup_{i\to\infty}\{U_i \neq V_i\}) = 0$. □

Lemma 3. For $X \in \mathfrak{Z}_+(\Omega, \mathfrak{A})$

$$\mathrm{E}(X) \le \sum_{k=0}^{\infty} P(\{X > k\}) \le \mathrm{E}(X) + 1.$$

(Cf. Corollary II.8.2.)

Proof. We have

$$\mathrm{E}(X) = \sum_{k=1}^{\infty} \int_{\{k-1<X\le k\}} X\, dP,$$

and therefore

$$\mathrm{E}(X) \le \sum_{k=1}^{\infty} k \cdot P(\{k-1 < X \le k\}) = \sum_{k=0}^{\infty} P(\{X > k\})$$

as well as

$$\mathrm{E}(X) \ge \sum_{k=1}^{\infty} (k-1) \cdot P(\{k-1 < X \le k\}) \ge \sum_{k=0}^{\infty} P(\{X > k\}) - 1.$$

□

Theorem 4 (Strong Law of Large Numbers, i.i.d. Case)**.** Let $(X_n)_{n\in\mathbb{N}}$ be i.i.d. Then

$$\exists\, Z \in \mathfrak{Z}(\Omega, \mathfrak{A}) : \ \frac{1}{n} \cdot S_n \xrightarrow{P\text{-a.s.}} Z \quad \Leftrightarrow \quad X_1 \in \mathfrak{L}^1,$$

in which case $Z = \mathrm{E}(X_1)$ P-a.s.

Proof. '⇒': Clearly

$$P(\{|X_1| > n\}) = P(A_n)$$

where

$$A_n = \{|X_n| > n\}.$$

Note that

$$\frac{1}{n} \cdot X_n = \frac{1}{n} \cdot S_n - \frac{n-1}{n} \cdot \frac{1}{n-1} \cdot S_{n-1} \xrightarrow{P\text{-a.s.}} 0.$$

Hence

$$P(\limsup_{n\to\infty} A_n) = 0.$$

Since $(A_n)_{n\in\mathbb{N}}$ is independent, the Borel-Cantelli Lemma implies

$$\sum_{n=1}^{\infty} P(A_n) < \infty.$$

Use Lemma 3 to obtain $\mathrm{E}(|X_1|) < \infty$.
'$\Leftarrow$': Consider the truncated random variables

$$Y_n = \begin{cases} X_n & \text{if } |X_n| < n \\ 0 & \text{otherwise.} \end{cases}$$

We have

$$\sum_{i=1}^{\infty} \frac{1}{i^2} \cdot \mathrm{Var}(Y_i) < \infty. \tag{3}$$

Proof: Observe that

$$\begin{aligned} \mathrm{Var}(Y_i) \le \mathrm{E}(Y_i^2) &= \sum_{k=1}^{i} \mathrm{E}(Y_i^2 \cdot 1_{[k-1,k[} \circ |Y_i|) \\ &= \sum_{k=1}^{i} \mathrm{E}(X_i^2 \cdot 1_{[k-1,k[} \circ |X_i|) \le \sum_{k=1}^{i} k^2 \cdot P(\{k-1 \le |X_1| < k\}). \end{aligned}$$

Thus

$$\begin{aligned} \sum_{i=1}^{\infty} \frac{1}{i^2} \cdot \mathrm{Var}(Y_i) &\le \sum_{k=1}^{\infty} k^2 \cdot P(\{k-1 \le |X_1| < k\}) \cdot \sum_{i=k}^{\infty} \frac{1}{i^2} \\ &\le 2 \cdot \sum_{k=1}^{\infty} k \cdot P(\{k-1 \le |X_1| < k\}) \le 2 \cdot (\mathrm{E}(|X_1|) + 1) < \infty, \end{aligned}$$

cf. the proof of Lemma 3.
Moreover,

$$\sum_{i=1}^{\infty} P(\{X_i \neq Y_i\}) < \infty, \tag{4}$$

since, by Lemma 3,

$$\sum_{i=1}^{\infty} P(\{X_i \neq Y_i\}) = \sum_{i=1}^{\infty} P(\{|X_i| \ge i\}) \le \sum_{i=0}^{\infty} P(\{|X_1| > i\}) \le \mathrm{E}(|X_1|) + 1 < \infty.$$

Furthermore,

$$\lim_{n\to\infty} \mathrm{E}(Y_n) = \mathrm{E}(X_1), \tag{5}$$

according to the dominated convergence theorem.
We obtain

$$\frac{1}{n} \cdot \sum_{i=1}^{n} (Y_i - \mathrm{E}(Y_i)) \xrightarrow{P\text{-a.s.}} 0$$

from Theorem 3 and (3). Due to (5)

$$\lim_{n\to\infty} \frac{1}{n} \cdot \sum_{i=1}^{n} (\mathrm{E}(Y_i) - \mathrm{E}(X_i)) = 0.$$

Thus

$$\frac{1}{n} \cdot \sum_{i=1}^{n} (Y_i - \mathrm{E}(X_i)) \overset{P\text{-a.s.}}{\longrightarrow} 0.$$

Finally, by Lemma 2 and (4)

$$\frac{1}{n} \cdot \sum_{i=1}^{n} (X_i - \mathrm{E}(X_i)) \overset{P\text{-a.s.}}{\longrightarrow} 0.$$

□

Theorem 5. Let $(X_n)_{n\in\mathbb{N}}$ be i.i.d.

(i) If $\mathrm{E}(X_1^-) < \infty \wedge \mathrm{E}(X_1^+) = \infty$ then

$$\frac{1}{n} \cdot S_n \overset{P\text{-a.s.}}{\longrightarrow} \infty.$$

(ii) If $\mathrm{E}(|X_1|) = \infty$ then

$$\limsup_{n\to\infty} \left| \frac{1}{n} \cdot S_n \right| = \infty \; P\text{-a.s.}$$

Proof. (i) follows from Theorem 4, and (ii) is an application of the Borel-Cantelli Lemma, see Gänssler, Stute (1977, p. 131). □

Remark 3. We already have $S_n/n \overset{P\text{-a.s.}}{\longrightarrow} \mathrm{E}(X_1)$ if the random variables X_n are identically distributed, P-integrable, and pairwise independent. See Bauer (1996, §12).

Remark 4. The basic idea of *Monte-Carlo algorithms*: to compute a quantity $a \in \mathbb{R}$

(i) find a probability measure μ on $(\mathbb{R}, \mathfrak{B})$ such that $\int_{\mathbb{R}} x\, \mu(dx) = a$,

(ii) take an i.i.d. sequence $(X_n)_{n\in\mathbb{N}}$ with $P_{X_1} = \mu$ and approximate a by $1/n \cdot S_n(\omega)$.

Clearly S_n/n is an *unbiased estimator* for a, i.e.,

$$E\Big(\frac{1}{n} \cdot S_n\Big) = a.$$

Due to the Strong Law of Large Numbers S_n/n converges almost surely to a. If $X_1 \in \mathfrak{L}^2$, then

$$\mathrm{E}\Big(\frac{1}{n} \cdot S_n - a\Big)^2 = \mathrm{Var}\Big(\frac{1}{n} \cdot S_n - a\Big) = \mathrm{Var}\Big(\frac{1}{n} \cdot \sum_{i=1}^{n} (X_i - a)\Big) = \frac{1}{n} \cdot \mathrm{Var}(X_1),$$

i.e., the variance of X_1 is the key quantity for the error of the Monte Carlo algorithm in the mean square sense. Moreover,

$$\frac{1}{n-1} \cdot \sum_{i=1}^{n} (X_i - S_n/n)^2 \xrightarrow{P\text{-a.s.}} \mathrm{Var}(X_1)$$

provides a simple estimator for this variance, see Einführung in die Stochastik.

Applications: see, e.g., Übung 10.3 and 10.4.

Remark 5. Let $(X_n)_{n\in\mathbb{N}}$ be i.i.d. with $\mu = P_{X_1}$ and corresponding distribution function $F = F_{X_1}$. Suppose that μ is unknown, but observations $X_1(\omega), \dots, X_n(\omega)$ are available for 'estimation of μ'.

Fix $C \in \mathfrak{B}$. Due to Theorem 4

$$\frac{1}{n} \cdot \sum_{i=1}^{n} 1_C \circ X_i \xrightarrow{P\text{-a.s.}} \mu(C).$$

The particular case $C =]-\infty, x]$ leads to the definitions

$$F_n(x,\omega) = \frac{1}{n} \cdot |\{i \in \{1,\dots,n\} : X_i(\omega) \le x\}|, \qquad x \in \mathbb{R},$$

and

$$\mu_n(\cdot,\omega) = \frac{1}{n} \cdot \sum_{i=1}^{n} \varepsilon_{X_i(\omega)}$$

of the *empirical distribution function* $F_n(\cdot,\omega)$ and the *empirical distribution* $\mu_n(\cdot,\omega)$, resp. We obtain

$$\forall x \in \mathbb{R}\ \exists A \in \mathfrak{A} :\ P(A) = 1 \wedge \Big(\forall \omega \in A :\ \lim_{n\to\infty} F_n(x,\omega) = F(x)\Big).$$

Therefore

$$\exists A \in \mathfrak{A} :\ P(A) = 1 \wedge \Big(\forall q \in \mathbb{Q}\ \forall \omega \in A :\ \lim_{n\to\infty} F_n(q,\omega) = F(q)\Big),$$

which implies

$$\exists A \in \mathfrak{A} :\ P(A) = 1 \wedge \Big(\forall \omega \in A :\ \mu_n(\cdot,\omega) \xrightarrow{\mathrm{w}} \mu\Big),$$

see Helly's Theorem (ii), p. 61, and Theorem III.3.2.

A refined analysis yields the *Glivenko-Cantelli Theorem*

$$\exists A \in \mathfrak{A} :\ P(A) = 1 \wedge \Big(\forall \omega \in A :\ \lim_{n\to\infty} \sup_{x\in\mathbb{R}} |F_n(x,\omega) - F(x)| = 0\Big),$$

see Einführung in die Stochastik.