Chapter 11
Martingale Convergence Theorems and Their Applications

We became familiar with martingales $X = (X_n)_{n \in \mathbb{N}_0}$ as fair games and found that under certain transformations (optional stopping, discrete stochastic integral) martingales turn into martingales. In this chapter, we will see that under weak conditions (non-negativity or uniform integrability) martingales converge almost surely. Furthermore, the martingale structure implies L^p-convergence under assumptions that are (formally) weaker than those of Chapter 7. The basic ideas of this chapter are Doob's inequality (Theorem 11.2) and the upcrossing inequality (Lemma 11.3).

11.1 Doob's Inequality

With Kolmogorov's inequality (Theorem 5.28), we became acquainted with an inequality that bounds the probability of large values of the maximum of a square integrable process with independent centered increments. Here we want to improve this inequality in two directions. On the one hand, we replace the independent increments by the assumption that the process of partial sums is a martingale. On the other hand, we can manage with less than second moments; alternatively, we can get better bounds if we have higher moments.

Let $I \subset \mathbb{N}_0$ and let $X = (X_n)_{n \in I}$ be a stochastic process. For $n \in \mathbb{N}$, we denote

$$X_n^* = \sup\{X_k : k \le n\} \quad \text{and} \quad |X|_n^* = \sup\big\{|X_k| : k \le n\big\}.$$

Lemma 11.1 *If X is a submartingale, then, for all $\lambda > 0$,*

$$\lambda \mathbf{P}\big[X_n^* \ge \lambda\big] \le \mathbf{E}[X_n \mathbb{1}_{\{X_n^* \ge \lambda\}}] \le \mathbf{E}\big[|X_n| \mathbb{1}_{\{X_n^* \ge \lambda\}}\big].$$

Proof The second inequality is trivial. For the first one, let

$$\tau := \inf\{k \in I : X_k \ge \lambda\} \wedge n.$$

A. Klenke, *Probability Theory*, Universitext, DOI 10.1007/978-1-4471-5361-0_11,

By Theorem 10.11 (optional sampling theorem),

$$\begin{aligned}\mathbf{E}[X_n] \ge \mathbf{E}[X_\tau] &= \mathbf{E}[X_\tau \mathbb{1}_{\{X_n^* \ge \lambda\}}] + \mathbf{E}[X_\tau \mathbb{1}_{\{X_n^* < \lambda\}}] \\ &\ge \lambda \mathbf{P}\big[X_n^* \ge \lambda\big] + \mathbf{E}[X_n \mathbb{1}_{\{X_n^* < \lambda\}}].\end{aligned}$$

(Note that $\tau = n$ if $X_n^* < \lambda$.) Now subtract $\mathbf{E}[X_n \mathbb{1}_{\{X_n^* < \lambda\}}]$. □

Theorem 11.2 (Doob's L^p-inequality) *Let X be a martingale or a positive submartingale.*

(i) *For any $p \ge 1$ and $\lambda > 0$,*

$$\lambda^p \mathbf{P}\big[|X|_n^* \ge \lambda\big] \le \mathbf{E}\big[|X_n|^p\big].$$

(ii) *For any $p > 1$,*

$$\mathbf{E}\big[|X_n|^p\big] \le \mathbf{E}\big[\big(|X|_n^*\big)^p\big] \le \left(\frac{p}{p-1}\right)^p \mathbf{E}\big[|X_n|^p\big].$$

Proof We follow the proof in [144].

(i) By Theorem 9.35, $(|X_n|^p)_{n\in I}$ is a submartingale, and the claim follows by Lemma 11.1.

(ii) The first inequality is trivial. For the second inequality, we may assume that $\mathbf{E}[|X_n|^p] < \infty$. Note that, by Lemma 11.1,

$$\lambda \mathbf{P}\big[|X|_n^* \ge \lambda\big] \le \mathbf{E}\big[|X_n| \mathbb{1}_{\{|X|_n^* \ge \lambda\}}\big].$$

Hence, for any $K > 0$,

$$\begin{aligned}\mathbf{E}\big[\big(|X|_n^* \wedge K\big)^p\big] &= \mathbf{E}\left[\int_0^{|X|_n^* \wedge K} p\lambda^{p-1}\, d\lambda\right] \\ &= \mathbf{E}\left[\int_0^K p\lambda^{p-1} \mathbb{1}_{\{|X|_n^* \ge \lambda\}}\, d\lambda\right] \\ &= \int_0^K p\lambda^{p-1} \mathbf{P}\big[|X|_n^* \ge \lambda\big]\, d\lambda \\ &\le \int_0^K p\lambda^{p-2} \mathbf{E}\big[|X_n| \mathbb{1}_{\{|X|_n^* \ge \lambda\}}\big]\, d\lambda \\ &= p\mathbf{E}\left[|X_n| \int_0^{|X|_n^* \wedge K} \lambda^{p-2}\, d\lambda\right] \\ &= \frac{p}{p-1} \mathbf{E}\big[|X_n| \cdot \big(|X|_n^* \wedge K\big)^{p-1}\big].\end{aligned}$$

Hölder's inequality then yields

$$\mathbf{E}\big[\big(|X|_n^* \wedge K\big)^p\big] \le \frac{p}{p-1}\mathbf{E}\big[\big(|X|_n^* \wedge K\big)^p\big]^{(p-1)/p} \cdot \mathbf{E}\big[|X_n|^p\big]^{1/p}.$$

We raise both sides to the pth power and divide by $\mathbf{E}[(|X|_n^* \wedge K)^p]^{p-1}$ (here we need the truncation at K to make sure we divide by a finite number) to obtain

$$\mathbf{E}\big[\big(|X|_n^* \wedge K\big)^p\big] \le \left(\frac{p}{p-1}\right)^p \mathbf{E}\big[|X_n|^p\big].$$

Finally, let $K \to \infty$. □

Exercise 11.1.1 Let $(X_n)_{n\in\mathbb{N}_0}$ be a submartingale or a supermartingale. Use Theorem 11.2 and Doob's decomposition to show that, for all $n \in \mathbb{N}$ and $\lambda > 0$,

$$\lambda \mathbf{P}\big[|X|_n^* \ge \lambda\big] \le 12\mathbf{E}\big[|X_0|\big] + 9\mathbf{E}\big[|X_n|\big].$$

11.2 Martingale Convergence Theorems

In this section, we present the usual martingale convergence theorems and give a few small examples. We start with the core of the martingale convergence theorems, the so-called upcrossing inequality.

Let $\mathbb{F} = (\mathcal{F}_n)_{n\in\mathbb{N}_0}$ be a filtration and $\mathcal{F}_\infty = \sigma\big(\bigcup_{n\in\mathbb{N}_0} \mathcal{F}_n\big)$. Let $(X_n)_{n\in\mathbb{N}_0}$ be real-valued and adapted to $\mathbb{F}$. Let $a, b \in \mathbb{R}$ with $a < b$. If we think of X as a stock price, it would be a sensible trading strategy to buy the stock when its price has fallen below a and to sell it when it exceeds b at least if we knew for sure that the price would always rise above the level b again. Each time the price makes such an *upcrossing* from a to b, we make a profit of at least $b - a$. If we get a bound on the maximal profit we can make, dividing it by $b - a$ gives a bound on the maximal number of such upcrossings. If this number is finite for all $a < b$, then the price has to converge as $n \to \infty$.

Let us get into the details. Define stopping times $\sigma_0 \equiv 0$ and

$$\tau_k := \inf\{n \ge \sigma_{k-1} : X_n \le a\} \quad \text{for } k \in \mathbb{N},$$

$$\sigma_k := \inf\{n \ge \tau_k : X_n \ge b\} \quad \text{for } k \in \mathbb{N}.$$

Note that $\tau_k = \infty$ if $\sigma_{k-1} = \infty$, and $\sigma_k = \infty$ if $\tau_k = \infty$. We say that X has its kth *upcrossing* over $[a, b]$ between τ_k and σ_k if $\sigma_k < \infty$. For $n \in \mathbb{N}$, define

$$U_n^{a,b} := \sup\{k \in \mathbb{N}_0 : \sigma_k \le n\}$$

as the number of upcrossings over $[a, b]$ until time n.

Lemma 11.3 (Upcrossing inequality) *Let $(X_n)_{n\in\mathbb{N}_0}$ be a submartingale. Then*

$$\mathbf{E}\big[U_n^{a,b}\big] \le \frac{\mathbf{E}[(X_n-a)^+] - \mathbf{E}[(X_0-a)^+]}{b-a}.$$

Proof Recall the discrete stochastic integral $H \cdot X$ from Definition 9.37. Formally, the intimated trading strategy H is described for $m \in \mathbb{N}_0$ by

$$H_m := \begin{cases} 1, & \text{if } m \in \{\tau_k+1, \ldots, \sigma_k\} \text{ for some } k \in \mathbb{N}, \\ 0, & \text{else.} \end{cases}$$

H is nonnegative and predictable since, for all $m \in \mathbb{N}$,

$$\{H_m = 1\} = \bigcup_{k=1}^{\infty} \big(\{\tau_k \le m-1\} \cap \{\sigma_k > m-1\}\big),$$

and each of the events is in $\mathcal{F}_{m-1}$. Define $Y = \max(X, a)$. If $k \in \mathbb{N}$ and $\sigma_k < \infty$, then clearly $Y_{\sigma_i} - Y_{\tau_i} = Y_{\sigma_i} - a \ge b-a$ for all $i \le k$; hence

$$(H \cdot Y)_{\sigma_k} = \sum_{i=1}^{k} \sum_{j=\tau_i+1}^{\sigma_i} (Y_j - Y_{j-1}) = \sum_{i=1}^{k} (Y_{\sigma_i} - Y_{\tau_i}) \ge k(b-a).$$

For $j \in \{\sigma_k, \ldots, \tau_{k+1}\}$, we have $(H \cdot Y)_j = (H \cdot Y)_{\sigma_k}$. On the other hand, for $j \in \{\tau_k+1, \ldots, \sigma_k\}$, we have $(H \cdot Y)_j \ge (H \cdot Y)_{\tau_k} = (H \cdot Y)_{\sigma_{k-1}}$. Hence $(H \cdot Y)_n \ge (b-a)U_n^{a,b}$ for all $n \in \mathbb{N}$.

By Corollary 9.34, Y is a submartingale, and (by Theorem 9.39) so are $H \cdot Y$ and $(1-H) \cdot Y$. Now $Y_n - Y_0 = (1 \cdot Y)_n = (H \cdot Y)_n + ((1-H) \cdot Y)_n$; hence

$$\mathbf{E}[Y_n - Y_0] \ge \mathbf{E}\big[(H \cdot Y)_n\big] \ge (b-a)\mathbf{E}\big[U_n^{a,b}\big]. \qquad \square$$

Theorem 11.4 (Martingale convergence theorem) *Let $(X_n)_{n\in\mathbb{N}_0}$ be a submartingale with $\sup\{\mathbf{E}[X_n^+] : n \ge 0\} < \infty$. Then there exists an $\mathcal{F}_\infty$-measurable random variable X_∞ with $\mathbf{E}[|X_\infty|] < \infty$ and $X_n \xrightarrow{n\to\infty} X_\infty$ almost surely.*

Proof Let $a < b$. Since $\mathbf{E}[(X_n-a)^+] \le |a| + \mathbf{E}[X_n^+]$, by Lemma 11.3,

$$\mathbf{E}\big[U_n^{a,b}\big] \le \frac{|a| + \mathbf{E}[X_n^+]}{b-a}.$$

Manifestly, the monotone limit $U^{a,b} := \lim_{n\to\infty} U_n^{a,b}$ exists. By assumption, we have $\mathbf{E}[U^{a,b}] = \lim_{n\to\infty} \mathbf{E}[U_n^{a,b}] < \infty$. In particular, $\mathbf{P}[U^{a,b} < \infty] = 1$. Define the $\mathcal{F}_\infty$-measurable events

$$C^{a,b} = \Big\{\liminf_{n\to\infty} X_n < a\Big\} \cap \Big\{\limsup_{n\to\infty} X_n > b\Big\} \subset \big\{U^{a,b} = \infty\big\}$$

and

$$C = \bigcup_{\substack{a,b \in \mathbb{Q} \\ a<b}} C^{a,b}.$$

Then $\mathbf{P}[C^{a,b}] = 0$ and thus also $\mathbf{P}[C] = 0$. However, by construction, $(X_n)_{n \in \mathbb{N}}$ is convergent on C^c. Hence there exists the almost sure limit $X_\infty = \lim_{n\to\infty} X_n$. Each X_n is $\mathcal{F}_\infty$-measurable; hence X_∞ also is $\mathcal{F}_\infty$-measurable.

By Fatou's lemma,

$$\mathbf{E}\big[X_\infty^+\big] \le \sup\big\{\mathbf{E}\big[X_n^+\big] : n \ge 0\big\} < \infty.$$

On the other hand (since X is a submartingale), again by Fatou's lemma,

$$\begin{aligned}\mathbf{E}\big[X_\infty^-\big] &\le \liminf_{n\to\infty} \mathbf{E}\big[X_n^-\big] = \liminf_{n\to\infty}\big(\mathbf{E}\big[X_n^+\big] - \mathbf{E}[X_n]\big) \\ &\le \sup\big\{\mathbf{E}\big[X_n^+\big] : n \in \mathbb{N}_0\big\} - \mathbf{E}[X_0] < \infty.\end{aligned}$$

□

Corollary 11.5 *If X is a nonnegative supermartingale, then there is an $\mathcal{F}_\infty$-measurable random variable $X_\infty \ge 0$ with $\mathbf{E}[X_\infty] \le \mathbf{E}[X_0]$ and $X_n \overset{n\to\infty}{\longrightarrow} X_\infty$ a.s.*

Proof The preceding theorem with $(-X)$ establishes X_∞ as the almost sure limit. Fatou's lemma yields

$$\mathbf{E}[X_\infty] \le \liminf_{n\to\infty} \mathbf{E}[X_n] \le \mathbf{E}[X_0].$$

□

Example 11.6 Let S_n be the account balance in the Petersburg game after the nth round (see Example 9.40). Then S is a martingale and $S_n \le 1$ almost surely for any n. Hence the assumptions of Theorem 11.4 are fulfilled and $(S_n)_{n \in \mathbb{N}_0}$ converges to a finite random variable almost surely for $n \to \infty$. Since the account changes as long as stakes are put up (that is, as long as $S_n < 1$), we get $\lim_{n\to\infty} S_n = 1$ almost surely.

Since $\mathbf{E}[S_n] = 0$ for all $n \in \mathbb{N}_0$, this convergence cannot hold in L^1. This observation tallies with the fact that S is not uniformly integrable. ◊

For uniformly integrable martingales, a stronger convergence theorem holds.

Theorem 11.7 (Convergence theorem for uniformly integrable martingales) *Let $(X_n)_{n \in \mathbb{N}_0}$ be a uniformly integrable $\mathbb{F}$- (sub-, super-) martingale. Then there exists an $\mathcal{F}_\infty$-measurable integrable random variable X_∞ with $X_n \overset{n\to\infty}{\longrightarrow} X_\infty$ a.s. and in L^1. Furthermore:*

- $X_n = \mathbf{E}[X_\infty \mid \mathcal{F}_n]$ *for all $n \in \mathbb{N}$ if X is a martingale.*
- $X_n \le \mathbf{E}[X_\infty \mid \mathcal{F}_n]$ *for all $n \in \mathbb{N}$ if X is a submartingale.*
- $X_n \ge \mathbf{E}[X_\infty \mid \mathcal{F}_n]$ *for all $n \in \mathbb{N}$ if X is a supermartingale.*

Remark 11.8 The statement of Theorem 11.7 can be reformulated as: The process $(X_n)_{n\in\mathbb{N}_0\cup\{\infty\}}$ is a (sub-, super-) martingale with respect to $(\mathcal{F}_n)_{n\in\mathbb{N}_0\cup\{\infty\}}$. ◊

Proof We give the proof for the case where X is a submartingale. Uniform integrability implies $\sup\{\mathbf{E}[X_n^+] : n \geq 0\} < \infty$. By Theorem 11.4, the almost sure limit X_∞ exists. Hence $\mathbf{E}[|X_n - X_\infty|] \overset{n\to\infty}{\longrightarrow} 0$ by Theorem 6.25. By Corollary 8.21, the L^1-convergence of (X_n) implies the L^1-convergence of the conditional expectations: $\mathbf{E}[|\mathbf{E}[X_n \mid \mathcal{F}_m] - \mathbf{E}[X_\infty \mid \mathcal{F}_m]|] \overset{n\to\infty}{\longrightarrow} 0$. Thus, by the triangle inequality,

$$\begin{aligned}&\left|\mathbf{E}\left[\left(\mathbf{E}[X_\infty \mid \mathcal{F}_m] - X_m\right)^-\right] - \mathbf{E}\left[\left(\mathbf{E}[X_n \mid \mathcal{F}_m] - X_m\right)^-\right]\right| \\ &\quad\leq \mathbf{E}\left[\left|\mathbf{E}[X_\infty \mid \mathcal{F}_m] - \mathbf{E}[X_n \mid \mathcal{F}_m]\right|\right] \overset{n\to\infty}{\longrightarrow} 0.\end{aligned}$$

As X is a submartingale, we have $(\mathbf{E}[X_n \mid \mathcal{F}_m] - X_m)^- = 0$ for $n \geq m$. Therefore, $\mathbf{E}[(\mathbf{E}[X_\infty \mid \mathcal{F}_m] - X_m)^-] = 0$ and thus $\mathbf{E}[X_\infty \mid \mathcal{F}_m] - X_m \geq 0$ almost surely. □

Corollary 11.9 *Let $X \geq 0$ be a martingale and let $X_\infty = \lim_{n\to\infty} X_n$. Then $\mathbf{E}[X_\infty] = \mathbf{E}[X_0]$ if and only if X is uniformly integrable.*

Proof This is a direct consequence of Theorem 6.25. □

Let $p \in [1, \infty)$. A real-valued stochastic process $(X_i)_{i\in I}$ is called L^p-bounded if $\sup_{i\in I} \mathbf{E}[|X_i|^p] < \infty$ (Definition 6.20). In general, for $(|X_i|^p)_{i\in I}$ to be uniformly integrable it is not enough that $(X_i)_{i\in I}$ be L^p-bounded. However, if X is a martingale and if $p > 1$, then Doob's inequality implies that the statements are equivalent. In particular, in this case, almost sure convergence implies convergence in L^p.

Theorem 11.10 (L^p-convergence theorem for martingales) *Let $p > 1$ and let $(X_n)_{n\in\mathbb{N}_0}$ be an L^p-bounded martingale. Then there exists an $\mathcal{F}_\infty$-measurable random variable X_∞ with $\mathbf{E}[|X_\infty|^p] < \infty$ and $X_n \overset{n\to\infty}{\longrightarrow} X_\infty$ almost surely and in L^p. In particular, $(|X_n|^p)_{n\in\mathbb{N}_0}$ is uniformly integrable.*

Proof By Corollary 6.21, X is uniformly integrable. Hence the almost sure limit X_∞ exists. By Doob's inequality (Theorem 11.2), for all $n \in \mathbb{N}$,

$$\mathbf{E}\left[\sup\left\{|X_k|^p : k \leq n\right\}\right] \leq \left(\frac{p}{p-1}\right)^p \mathbf{E}\left[|X_n|^p\right].$$

Therefore,

$$\mathbf{E}\left[\sup\left\{|X_k|^p : k \in \mathbb{N}_0\right\}\right] \leq \left(\frac{p}{p-1}\right)^p \sup\left\{\mathbf{E}\left[|X_n|^p\right] : n \in \mathbb{N}_0\right\} < \infty.$$

Hence, in particular, $(|X_n|^p)_{n\in\mathbb{N}_0}$ is uniformly integrable.

Since $|X_n - X_\infty|^p \le 2^p \sup\{|X_n|^p : n \in \mathbb{N}_0\}$, dominated convergence yields

$$\mathbf{E}\big[|X_\infty|^p\big] < \infty \quad \text{and} \quad \mathbf{E}\big[|X_n - X_\infty|^p\big] \overset{n\to\infty}{\longrightarrow} 0. \qquad \square$$

For the case of square integrable martingales, there is a convenient criterion for L^2-boundedness that we record as a corollary (see Definition 10.3).

Corollary 11.11 *Let X be a square integrable martingale with square variation process $\langle X\rangle$. Then the following four statements are equivalent*:

(i) $\sup_{n\in\mathbb{N}} \mathbf{E}[X_n^2] < \infty$.
(ii) $\lim_{n\to\infty} \mathbf{E}[\langle X\rangle_n] < \infty$.
(iii) *X converges in L^2.*
(iv) *X converges almost surely and in L^2.*

Proof "(i) $\iff$ (ii)" Since $\mathbf{Var}[X_n - X_0] = \mathbf{E}[\langle X\rangle_n]$ (see Theorem 10.4), X is bounded in L^2 if and only if (ii) holds.

"(iv) $\implies$ (iii) $\implies$ (i)" This is trivial.

"(i) $\implies$ (iv)" This is the statement of Theorem 11.10. $\square$

Remark 11.12 In general, the statement of Theorem 11.10 fails for $p = 1$. See Exercise 11.2.1. $\Diamond$

Lemma 11.13 *Let X be a square integrable martingale with square variation process $\langle X\rangle$, and let τ be a stopping time. Then the stopped process X^τ has square variation process $\langle X^\tau\rangle = \langle X\rangle^\tau := (\langle X\rangle_{\tau\wedge n})_{n\in\mathbb{N}_0}$.*

Proof This is left as an exercise. $\square$

If in Corollary 11.11 we do not assume that the *expectations* of the square variation are bounded but only that the square variation is *almost surely* bounded, then we still get that X converges almost surely (albeit not in L^2).

Theorem 11.14 *If X is a square integrable martingale with $\sup_{n\in\mathbb{N}}\langle X\rangle_n < \infty$ almost surely, then X converges almost surely.*

Proof Without loss of generality, we can assume that $X_0 = 0$, otherwise consider the martingale $(X_n - X_0)_{n\in\mathbb{N}_0}$, which has the same square variation process. For $K > 0$, let

$$\tau_K := \inf\big\{n \in \mathbb{N} : \langle X\rangle_{n+1} \ge K\big\}.$$

This is a stopping time since $\langle X\rangle$ is predictable. Evidently, $\sup_{n\in\mathbb{N}}\langle X\rangle_{\tau_K\wedge n} \le K$ almost surely. By Corollary 11.11, the stopped process X^{τ_K} converges almost surely (and in L^2) to a random variable that we denote by $X_\infty^{\tau_K}$. By assumption, $\mathbf{P}[\tau_K = \infty] \to 1$ for $K \to \infty$; hence X converges almost surely. $\square$

Example 11.15 Let X be a symmetric simple random walk on $\mathbb{Z}$. That is, $X_n = \sum_{k=1}^n R_k$, where $R_1, R_2, \ldots$ are i.i.d. and $\sim\mathrm{Rad}_{1/2}$:

$$\mathbf{P}[R_1 = 1] = \mathbf{P}[R_1 = -1] = \frac{1}{2}.$$

Then X is a martingale; however, $\limsup_{n\to\infty} X_n = \infty$ and $\liminf_{n\to\infty} X_n = -\infty$. Therefore, X does not even converge improperly. By the martingale convergence theorem, this is consonant with the fact that X is not uniformly integrable. ◊

Example 11.16 (Voter model, due to [28, 75]) Consider a simple model that describes the behavior of opportunistic voters who are capable of only one out of two opinions, say 0 and 1. Let $\Lambda \subset \mathbb{Z}^d$ be a set that we interpret as the sites at each of which there is one voter. For simplicity, assume that $\Lambda = \{0, \ldots, L-1\}^d$ for some $L \in \mathbb{N}$. Let $x(i) \in \{0, 1\}$ be the opinion of the voter at site $i \in \Lambda$ and denote by $x \in \{0, 1\}^\Lambda$ a generic state of the whole population. We now assume that the individual opinions may change at discrete time steps. At any time $n \in \mathbb{N}_0$, one site I_n out of Λ is chosen at random and the individual at that site reconsiders his or her opinion. To this end, the voter chooses a neighbor $I_n + N_n \in \Lambda$ (with *periodic boundary conditions*; that is, with addition modulo L in each coordinate) at random and adopts his or her opinion. We thus get a random sequence $(X_n)_{n\in\mathbb{N}_0}$ of states in $\{0, 1\}^\Lambda$ that represents the random evolution of the opinions of the whole colony. See Fig. 11.1 for a computer simulation of the voter model.

For a formal description of this model, let $(I_n)_{n\in\mathbb{N}}$ and $(N_n)_{n\in\mathbb{N}}$ be independent random variables. For any $n \in \mathbb{N}$, I_n is uniformly distributed on Λ and N_n is uniformly distributed on the set $\mathcal{N} := \{i \in \mathbb{Z}^d : \|i\|_2 = 1\}$ of the $2d$ nearest neighbors of the origin. Furthermore, $x = X_0 \in \{0, 1\}^\Lambda$ is the initial state. The states at later times are defined inductively by

$$X_n(i) = \begin{cases} X_{n-1}(i), & \text{if } I_n \neq i, \\ X_{n-1}(I_n + N_n), & \text{if } I_n = i. \end{cases}$$

Of course, the behavior over small periods of time is determined by the perils of randomness. However, in the long run, we might see certain patterns. To be more specific, the question is: In the long run, will there be a consensus of all individuals or will competing opinions persist?

Let $M_n := \sum_{i\in\Lambda} X_n(i)$ be the total number of individuals of opinion 1 at time n. Let $\mathbb{F}$ be the filtration $\mathbb{F} = (\mathcal{F}_n)_{n\in\mathbb{N}_0}$, where $\mathcal{F}_n = \sigma(I_k, N_k : k \le n)$ for all $n \in \mathbb{N}_0$. Then M is adapted to $\mathbb{F}$ and

$$\begin{aligned}\mathbf{E}[M_n \mid \mathcal{F}_{n-1}] &= M_{n-1} - \mathbf{E}\big[X_{n-1}(I_n) \mid \mathcal{F}_{n-1}\big] + \mathbf{E}\big[X_{n-1}(I_n + N_n) \mid \mathcal{F}_{n-1}\big] \\ &= M_{n-1} - \sum_{i\in\Lambda} \mathbf{P}[I_n = i] X_{n-1}(i) + \sum_{i\in\Lambda} \mathbf{P}[I_n + N_n = i] X_{n-1}(i) \\ &= M_{n-1}\end{aligned}$$

since $\mathbf{P}[I_n = i] = \mathbf{P}[I_n + N_n = i] = L^{-d}$ for all $i \in \Lambda$. Hence M is a bounded $\mathbb{F}$-martingale and thus converges almost surely and in L^1 to a random variable M_∞.

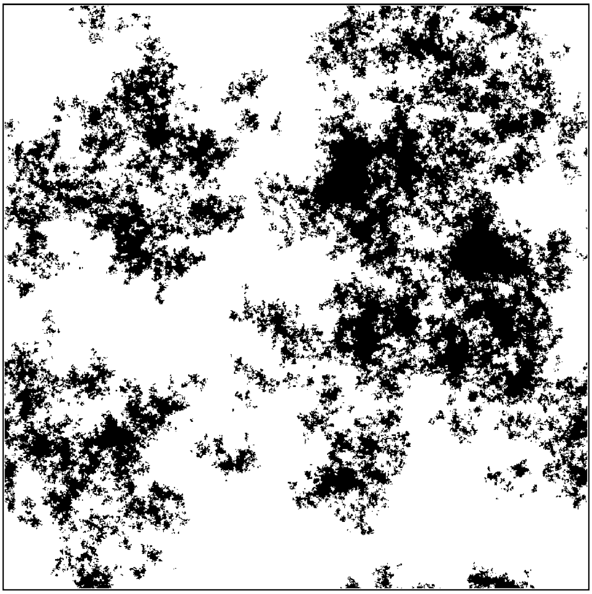

Fig. 11.1 Snapshot of a voter model on an 800×800 torus. The *black dots* are the Ones

Since M takes only integer values, there is a (random) n_0 such that $M_n = M_{n_0}$ for all $n \geq n_0$. However, then also $X_n = X_{n_0}$ for all $n \geq n_0$. Manifestly, no state x with $x \not\equiv 0$ and $x \not\equiv 1$ is stable. In fact, if x is not constant and if $i, j \in \Lambda$ are neighbors with $x(i) \neq x(j)$, then

$$\mathbf{P}[X_n \neq X_{n-1} \mid X_{n-1} = x] \geq \mathbf{P}[I_{n-1} = i, N_{n-1} = j - i] = L^{-d}(2d)^{-1}.$$

This implies $M_\infty \in \{0, L^d\}$. Now $\mathbf{E}[M_\infty] = M_0$; hence we have

$$\mathbf{P}\big[M_\infty = L^d\big] = \frac{M_0}{L^d} \quad \text{and} \quad \mathbf{P}[M_\infty = 0] = 1 - \frac{M_0}{L^d}.$$

Thus, eventually there will be a consensus of all individuals, and the probability that the surviving opinion is $e \in \{0, 1\}$ is the initial frequency of opinion e.

We could argue more formally to show that only the constant states are stable: Let $\langle M \rangle$ be the square variation process of M. Then

$$\begin{aligned}\langle M \rangle_n &= \sum_{k=1}^{n} \mathbf{E}[\mathbb{1}_{\{M_k \neq M_{k-1}\}} \mid \mathcal{F}_{k-1}] \\ &= \sum_{k=1}^{n} \mathbf{P}\big[X_{k-1}(I_k) \neq X_{k-1}(I_k + N_k) \,\big|\, \mathcal{F}_{k-1}\big].\end{aligned}$$

Hence

$$\begin{aligned} L^{2d} &\geq \mathbf{Var}[M_n] = \mathbf{E}\big[\langle M\rangle_n\big] \\ &= \sum_{k=1}^{n} \mathbf{P}\big[X_{k-1}(I_k) \neq X_{k-1}(I_k+N_k)\big] \\ &\geq (2d)^{-1}L^{-d}\sum_{k=1}^{n}\mathbf{P}\big[M_{k-1}\notin\{0,L^d\}\big]. \end{aligned}$$

Therefore, $\sum_{k=1}^{\infty}\mathbf{P}[M_{k-1}\notin\{0,L^d\}] \leq 2dL^{3d} < \infty$, and so, by the Borel–Cantelli lemma, $M_\infty \in \{0, L^d\}$. ◊

Example 11.17 (Radon–Nikodym theorem) With the aid of the martingale convergence theorem, we give an alternative proof of the Radon–Nikodym theorem (Corollary 7.34).

Let $(\Omega, \mathcal{F}, \mathbf{P})$ be a probability space and let Q be another probability measure on $(\Omega, \mathcal{A})$. We assume that $\mathcal{F}$ is countably generated; that is, there exist countably many sets $A_1, A_2, \ldots \in \mathcal{F}$ such that $\mathcal{F} = \sigma(\{A_1, A_2, \ldots\})$. For example, this is the case if $\mathcal{F}$ is the Borel σ-algebra on a Polish space. For the case $\Omega = \mathbb{R}^d$, one could take the open balls with rational radii, centered at points with rational coordinates (compare Remark 1.24).

We construct a filtration $\mathbb{F} = (\mathcal{F}_n)_{n\in\mathbb{N}}$ by letting $\mathcal{F}_n := \sigma(\{A_1, \ldots, A_n\})$. Evidently, $\#\mathcal{F}_n < \infty$ for all $n \in \mathbb{N}$. More precisely, there exists a unique finite subset $Z_n \subset \mathcal{F}_n \setminus \{\emptyset\}$ such that $B = \biguplus_{\substack{C\in Z_n \\ C\subset B}} C$ for any $B \in \mathcal{F}_n$. Z_n decomposes $\mathcal{F}_n$ into its "atoms". Finally, define a stochastic process $(X_n)_{n\in\mathbb{N}}$ by

$$X_n := \sum_{C\in Z_n:\mathbf{P}[C]>0} \frac{Q(C)}{\mathbf{P}[C]}\mathbb{1}_C.$$

Clearly, X is adapted to $\mathbb{F}$. Let $B \in \mathcal{F}_n$ and $m \geq n$. For any $C \in Z_m$, either $C \cap B = \emptyset$ or $C \subset B$. Hence

$$\mathbf{E}[X_m \mathbb{1}_B] = \sum_{C\in Z_m:\mathbf{P}[C]>0}\frac{Q(C)}{\mathbf{P}[C]}\mathbf{P}[C\cap B] = \sum_{C\in Z_m:C\subset B} Q(C) = Q(B). \tag{11.1}$$

In particular, X is an $\mathbb{F}$-martingale.

Now assume that Q is absolutely continuous with respect to P. By Example 7.39, this implies that X is uniformly integrable. By the martingale convergence theorem, X converges $\mathbf{P}$-almost surely and in $L^1(\mathbf{P})$ to a random variable X_∞. By (11.1), we have $\mathbf{E}[X_\infty \mathbb{1}_B] = Q(B)$ for all $B \in \bigcup_{n\in\mathbb{N}} \mathcal{F}_n$ and thus also for all $B \in \mathcal{F}$. Therefore, X_∞ is the Radon–Nikodym density of Q with respect to $\mathbf{P}$.

Note that for this proof we did not presume the existence of conditional expectations (rather we constructed them explicitly for finite σ-algebras); that is, we did not resort to the Radon–Nikodym theorem in a hidden way.

It could be objected that this argument works only for probability measures. However, this flaw can easily be remedied. Let μ and ν be arbitrary (but nonzero) σ-finite measures. Then there exist measurable functions $g, h : \Omega \to (0, \infty)$ with $\int g\, d\mu = 1$ and $\int h\, d\nu = 1$. Define $\mathbf{P} = g\mu$ and $Q = h\nu$. Clearly, $Q \ll \mathbf{P}$ if $\nu \ll \mu$. In this case, $\frac{g}{h} X_\infty$ is a version of the Radon–Nikodym derivative $\frac{d\nu}{d\mu}$.

The restriction that $\mathcal{F}$ is countably generated can also be dropped. Using the approximation theorems for measures, it can be shown that there is always a countably generated σ-algebra $\mathcal{G} \subset \mathcal{F}$ such that for any $A \in \mathcal{F}$, there is a $B \in \mathcal{G}$ with $\mathbf{P}[A \triangle B] = 0$. This can be employed to prove the general case. We do not give the details but refer to [169, Chapter 14.13]. $\Diamond$

Exercise 11.2.1 For $p = 1$, the statement of Theorem 11.10 may fail. Give an example of a nonnegative martingale X with $\mathbf{E}[X_n] = 1$ for all $n \in \mathbb{N}$ but such that $X_n \overset{n\to\infty}{\longrightarrow} 0$ almost surely.

Exercise 11.2.2 Let $X_1, X_2, \ldots$ be independent, square integrable random variables with $\sum_{n=1}^{\infty} \frac{1}{n^2} \mathbf{Var}[X_n] < \infty$. Use the martingale convergence theorem to show the strong law of large numbers for $(X_n)_{n\in\mathbb{N}}$.

Exercise 11.2.3 Give an example of a square integrable martingale that converges almost surely but not in L^2.

Exercise 11.2.4 Show that in Theorem 11.14 the converse implication may fail. That is, there exists a square integrable martingale X that converges almost surely but without $\lim_{n\to\infty} \langle X \rangle_n < \infty$ almost surely.

Exercise 11.2.5 Show the following converse of Theorem 11.14. Let $L > 0$ and let $(X_n)_{n\in\mathbb{N}}$ be a martingale with the property

$$|X_{n+1} - X_n| \le L \quad \text{a.s.} \tag{11.2}$$

Define the events

$$\begin{aligned}
C &:= \left\{ (X_n)_{n\in\mathbb{N}} \text{ converges as } n \to \infty \right\}, \\
A^+ &:= \left\{ \limsup_{n\to\infty} X_n < \infty \right\}, \\
A^- &:= \left\{ \liminf_{n\to\infty} X_n > -\infty \right\}, \\
F &:= \left\{ \sup_{n\in\mathbb{N}} \langle X \rangle_n < \infty \right\}.
\end{aligned}$$

Show that

$$C = A^+ = A^- = F \pmod{\mathbf{P}}.$$

Here equality of events (mod **P**) means that the events differ at most by a **P**-null set (see Definition 1.68(iii)).

Hint: Use the stopping times $\sigma_K = \inf\{n \in \mathbb{N} : |X_n| \geq K\}$; $\sigma_K^{\pm} = \inf\{n \in \mathbb{N} : \pm X_n \geq K\}$ and τ_K as in the proof of Theorem 11.14.

Exercise 11.2.6 Let the notation be as in Exercise 11.2.5. However, instead of (11.2) we make the weaker assumption

$$\mathbf{E}\Big[\sup_{n\in\mathbb{N}} |X_{n+1} - X_n|\Big] < \infty. \tag{11.3}$$

Show that

$$C = A^+ = A^- \pmod{\mathbf{P}}.$$

Hint: Use suitable stopping times ϱ_K and apply the martingale convergence theorem (Theorem 11.4) to the stopped process X^{ϱ_K}.

Exercise 11.2.7 (Conditional Borel–Cantelli lemma) Let $(\mathcal{F}_n)_{n\in\mathbb{N}_0}$ be a filtration and let $(A_n)_{n\in\mathbb{N}}$ be events with $A_n \in \mathcal{F}_n$ for all $n \in \mathbb{N}$. Define $A_\infty = \{\sum_{n=1}^{\infty} \mathbf{P}[A_n \mid \mathcal{F}_{n-1}] = \infty\}$ and $A^* = \limsup_{n\to\infty} A_n$. Show the conditional Borel–Cantelli lemma: $\mathbf{P}[A_\infty \triangle A^*] = 0$.

Hint: Apply Exercise 11.2.5 to $X_n = \sum_{n=1}^{\infty} (\mathbb{1}_{A_n} - \mathbf{P}[A_n \mid \mathcal{F}_{n-1}])$.

Exercise 11.2.8 Let $p \in [0, 1]$ and let $X = (X_n)_{n\in\mathbb{N}_0}$ be a stochastic process with values in $[0, 1]$. Assume that for all $n \in \mathbb{N}_0$, given $X_0, \ldots, X_n$, we have

$$X_{n+1} = \begin{cases} 1 - p + pX_n & \text{with probability } X_n, \\ pX_n & \text{with probability } 1 - X_n. \end{cases}$$

Show that X is a martingale that converges almost surely. Compute the distribution of the almost sure limit $\lim_{n\to\infty} X_n$.

Exercise 11.2.9 Let $f \in \mathcal{L}^1(\lambda)$, where λ is the restriction of the Lebesgue measure to $[0, 1]$. Let $I_{n,k} = [k2^{-n}, (k + 1)2^{-n})$ for $n \in \mathbb{N}$ and $k = 0, \ldots, 2^n - 1$. Define $f_n : [0, 1] \to \mathbb{R}$ by

$$f_n(x) = 2^n \int_{I_{k,n}} f \, d\lambda, \quad \text{if } k \text{ is chosen such that } x \in I_{k,n}.$$

Show that $f_n(x) \overset{n\to\infty}{\longrightarrow} f(x)$ for λ-almost all $x \in [0, 1]$.

Exercise 11.2.10 Assume that $\mathbb{F} = (\mathcal{F}_n)_{n\in\mathbb{N}}$ is a filtration on the probability space $(\Omega, \mathcal{A}, \mathbf{P})$. Let $\mathcal{F}_\infty := \sigma(\mathcal{F}_n : n \in \mathbb{N})$, and let $\mathcal{M}$ be the vector space of uniformly integrable $\mathbb{F}$-martingales. Show that the map $\Phi : \mathcal{L}^1(\mathcal{F}_\infty) \to \mathcal{M}$, $X_\infty \mapsto (\mathbf{E}[X_\infty \mid \mathcal{F}_n])_{n\in\mathbb{N}}$ is an isomorphism of vector spaces.

11.3 Example: Branching Process

Let $p = (p_k)_{k\in\mathbb{N}_0}$ be a probability vector on $\mathbb{N}_0$ and let $(Z_n)_{n\in\mathbb{N}_0}$ be the Galton–Watson process with one ancestor and offspring distribution p (see Definition 3.9). For convenience, we recall the construction of Z. Let $(X_{n,i})_{n\in\mathbb{N}_0, i\in\mathbb{N}}$ be i.i.d. random variables with $\mathbf{P}[X_{1,1} = k] = p_k$ for $k \in \mathbb{N}_0$. Let $Z_0 = 1$ and inductively define

$$Z_{n+1} = \sum_{i=1}^{Z_n} X_{n,i} \quad \text{for } n \in \mathbb{N}_0.$$

We interpret Z_n as the size of a population at time n and $X_{n,i}$ as the number of offspring of the ith individual of the nth generation.

Let $m := \mathbf{E}[X_{1,1}] < \infty$ be the expected number of offspring of an individual and let $\sigma^2 := \mathbf{Var}[X_{1,1}] \in (0, \infty)$ be its variance. Let $\mathcal{F}_n := \sigma(X_{k,i} : k < n, i \in \mathbb{N})$. Then Z is adapted to $\mathbb{F}$. Define $W_n = m^{-n} Z_n$.

Lemma 11.18 *W is a martingale. In particular,* $\mathbf{E}[Z_n] = m^n$ *for all* $n \in \mathbb{N}$.

Proof We compute the conditional expectation for $n \in \mathbb{N}_0$:

$$\begin{aligned}
\mathbf{E}[W_{n+1} \mid \mathcal{F}_n] &= m^{-(n+1)}\mathbf{E}[Z_{n+1} \mid \mathcal{F}_n] \\
&= m^{-(n+1)}\mathbf{E}\left[\sum_{i=1}^{Z_n} X_{n,i} \middle| \mathcal{F}_n\right] \\
&= m^{-(n+1)} \sum_{k=1}^{\infty} \mathbf{E}[\mathbb{1}_{\{Z_n=k\}} k \cdot X_{n,i} \mid \mathcal{F}_n] \\
&= m^{-n} \sum_{k=1}^{\infty} \mathbf{E}[k \cdot \mathbb{1}_{\{Z_n=k\}} \mid \mathcal{F}_n] \\
&= m^{-n} Z_n = W_n.
\end{aligned}$$

□

Theorem 11.19 *Let* $\mathbf{Var}[X_{1,1}] \in (0, \infty)$. *The a.s. limit* $W_\infty = \lim_{n\to\infty} W_n$ *exists and*

$$m > 1 \quad \Longleftrightarrow \quad \mathbf{E}[W_\infty] = 1 \quad \Longleftrightarrow \quad \mathbf{E}[W_\infty] > 0.$$

Proof W_∞ exists since $W \geq 0$ is a martingale. If $m \leq 1$, then $(Z_n)_{n\in\mathbb{N}}$ converges a.s. to some random variable Z_∞. Note that Z_∞ is the only choice since $\sigma^2 > 0$.

Now let $m > 1$. Since $\mathbf{E}[Z_{n-1}] = m^{n-1}$ (Lemma 11.18), by the Blackwell–Girshick formula (Theorem 5.10),

$$\begin{aligned}
\mathbf{Var}[W_n] &= m^{-2n}\big(\sigma^2\mathbf{E}[Z_{n-1}] + m^2\mathbf{Var}[Z_{n-1}]\big) \\
&= \sigma^2 m^{-(n+1)} + \mathbf{Var}[W_{n-1}].
\end{aligned}$$

Inductively, we get $\mathbf{Var}[W_n] = \sigma^2 \sum_{k=2}^{n+1} m^{-k} \leq \frac{\sigma^2 m}{m-1} < \infty$. Hence W is bounded in L^2 and Theorem 11.10 yields $W_n \to W_\infty$ in L^2 and thus in L^1. In particular, $\mathbf{E}[W_\infty] = \mathbf{E}[W_0] = 1$. □

The proof of Theorem 11.19 was simple due to the assumption of finite variance of the offspring distribution. However, there is a much stronger statement that here we can only quote (see [96], and see [110] for a modern proof).

Theorem 11.20 (Kesten-Stigum (1966)) *Let $m > 1$. Then*

$$\mathbf{E}[W_\infty] = 1 \quad \iff \quad \mathbf{E}[W_\infty] > 0 \quad \iff \quad \mathbf{E}\big[X_{1,1} \log(X_{1,1})^+\big] < \infty.$$