2. Martingales

Consider a game of flipping a coin infinitely and independently. Suppose that the head appears with probability p, while the tail comes up with probability $1-p$. Each time a bet, say d dollars, is set before the coin is flipped. As a result, the gambler wins the bet on heads but loses it otherwise. For the gambler, it is crucial to know the probability of ruin when the gambler starts with s dollars?

Let $Z_1, Z_2, ...$ be a sequence of i.i.d. random variables with $\mathbb{P}(Z_1 = 1) = p$ and $\mathbb{P}(Z_1 = -1) = 1-p$ and b_n be the bet set for the nth flip. Assume that b_n is a function of $Z_1, ..., Z_{n-1}$. Let S_n be the total asset after the nth flip. Clearly, one has

$$S_0 = d, \quad S_n = S_{n-1} + Z_n b_n(Z_1, ..., Z_{n-1}), \quad \forall n \geq 1, \tag{2.1}$$

where $0 < b_{n+1} \leq S_n$. By setting $\tau := \inf\{n \geq 1 : S_n = 0\}$ with $\inf \emptyset := \infty$, the ruin probability is given by $\mathbb{P}(\tau < \infty)$. In some computations, we obtain

$$\mathbb{E}(S_{n+1}|S_1, ..., S_n) \begin{cases} = S_n \text{ a.s.} & \text{for } p = 1/2 \\ < S_n \text{ a.s.} & \text{for } p < 1/2 \\ > S_n \text{ a.s.} & \text{for } p > 1/2 \end{cases}, \quad \text{when } S_n > 0.$$

The gambler may conclude from the above observation that the sequence of games $Z_1, Z_2, ...$ is *fair* if $p = 1/2$, *unfavorable* if $p < 1/2$ and *favorable* if $p > 1/2$.

2.1. **Definitions and properties.**

Definition 2.1. A sequence of σ-fields, $(\mathcal{F}_n)_{n=0}^{\infty}$, is a *filtration* if

$$\mathcal{F}_0 \subset \mathcal{F}_1 \subset \mathcal{F}_2 \subset \cdots.$$

A process $(X_n)_{n=0}^{\infty}$ is *adapted* to a filtration $(\mathcal{F}_n)_{n=0}^{\infty}$ if X_n is $\mathcal{F}_n$-measurable for all $n \geq 0$. We briefly say that X_n is adapted to $\mathcal{F}_n$.

Definition 2.2. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $(\mathcal{F}_n)_{n=0}^{\infty}$ be a filtration with $\mathcal{F}_n \subset \mathcal{F}$. A stochastic process, $(X_n)_{n=0}^{\infty}$, defined on $(\Omega, \mathcal{F})$ is a *martingale* (resp. *submartingale* and *supermartingale*) w.r.t. $(\mathcal{F}_n)_{n=0}^{\infty}$, or briefly X_n is a martingale (resp. submartingale and supermartingale) w.r.t. $\mathcal{F}_n$ if

(1) $(X_n)_{n=0}^{\infty}$ is adapted to $(\mathcal{F}_n)_{n=0}^{\infty}$;
(2) $\mathbb{E}|X_n| < \infty$ for all $n \geq 0$;
(3) $\mathbb{E}(X_{n+1}|\mathcal{F}_n) = X_n$ a.s. (resp. $\geq X_n$ a.s. and $\leq X_n$ a.s.) for all $n \geq 0$.

In the case that $\mathcal{F}_n = \mathcal{F}(X_0, ..., X_n)$ for $n \geq 0$, we briefly call $(X_n)_{n=0}^{\infty}$ or X_n a martingale (resp. submartingale and supermartingale) when (2) and (3) hold.

Remark 2.1. If $(X_n)_{n=0}^{\infty}$ is a submartingale w.r.t. $(\mathcal{F}_n)_{n=0}^{\infty}$, then $(-X_n)_{n=0}^{\infty}$ is a supermartingale w.r.t. $(\mathcal{F}_n)_{n=0}^{\infty}$. Besides, $(X_n)_{n=0}^{\infty}$ is a martingale w.r.t. $(\mathcal{F}_n)_{n=0}^{\infty}$ if and only if $(X_n)_{n=0}^{\infty}$ is both a submartingale and a supermartingale w.r.t. $(\mathcal{F}_n)_{n=0}^{\infty}$

Example 2.1. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and let X_n be a sequence of i.i.d. random variables on Ω satisfying $\mathbb{P}(X_1 = 1) = p$ and $\mathbb{P}(X_1 = -1) = 1 - p$ with $p \in (0, 1)$. Fix $d \in \mathbb{R}$ and set

$$S_0 = d, \quad S_n = d + X_1 + X_2 + \cdots + X_n, \quad \forall n \geq 1,$$

and

$$\mathcal{F}_0 = \{\emptyset, \Omega\}, \quad \mathcal{F}_n = \mathcal{F}(X_1, ..., X_n), \quad \forall n \geq 1.$$

Then, S_n is a martingale (resp. submartingale and supermartingale) w.r.t. $\mathcal{F}_n$ if $p = 1/2$ (resp. $p > 1/2$ and $p < 1/2$).

Example 2.2. Let X be a random variable with $\mathbb{E}|X| < \infty$ and $(\mathcal{F}_n)_{n=0}^{\infty}$ be a filtration. Then, $Y_n = \mathbb{E}(X|\mathcal{F}_n)$ is a martingale w.r.t. $\mathcal{F}_n$.

Proposition 2.1. *Suppose X_n is a martingale (resp. submartingale and supermartingale) with respect to $\mathcal{F}_n$. Then,*

1. $\mathcal{F}(X_0, X_1, ..., X_n) \subset \mathcal{F}_n$ *for* $n \geq 0$.
2. *For any filtration $\mathcal{G}_n$ satisfying*
$$\mathcal{F}(X_0, ..., X_n) \subset \mathcal{G}_n \subset \mathcal{F}_n,$$
X_n is a martingale (resp. submartingale and supermartingale) with respect to $\mathcal{G}_n$. In particular, X_n is a martingale (resp. submartingale and supermartingale).

Proof. Immediately from Proposition 1.4(4). □

Remark 2.2. Why we call them supermartingales? Such a denomination can be related to the superharmonic function f on $\mathbb{R}^d$, which is defined by

$$f(x) \geq \frac{1}{|B(0,r)|} \int_{B(x,r)} f(y) dy \quad \forall x \in \mathbb{R},\ r > 0,$$

where $B(x,r)$ is the open ball in $\mathbb{R}^d$ with radius r and center x, and $|B(0,r)|$ is the Lebesgue measure of $B(0,r)$.

Exercise 2.1. Let $X_1, X_2, ...$ be i.i.d. random elements with values on $\mathbb{R}^d$ and uniform distributed on $B(0,1)$. Set $X_0 = x$ and $S_n = X_0 + \cdots + X_n$ for $n \geq 0$. Prove that $f(S_n)$ is a supermartingale w.r.t. $\mathcal{F}(X_0, ..., X_n)$ for any superharmonic function f on $\mathbb{R}^d$.

Exercise 2.2. Let X_n be a submartingale w.r.t. $\mathcal{F}_n$. Prove that

1. $\mathbb{E}X_n \leq \mathbb{E}X_{n+1}$ for all $n \geq 0$.
2. $\sup_n \mathbb{E}|X_n| < \infty$ if and only if $\sup_n \mathbb{E}X_n^+ < \infty$, where $t^+ = t \vee 0$.

Derive similar statements for supermatringales?

Theorem 2.2. *Let X_n be a random variable and $\mathcal{F}_n$ be a filtration. Then, X_n is a martingale (resp. submartingale and supermartingale) w.r.t. $\mathcal{F}_n$ if and only if*

$$\mathbb{E}(X_n|\mathcal{F}_m) \overset{\text{a.s.}}{=} X_m \quad \forall n > m. \quad (\text{resp.} \geq X_m \text{ a.s. and} \leq X_m \text{ a.s.})$$

Proof. We prove the case for martingales, while the other cases can be shown in a similar way. The sufficiency for martingales is clear by choosing $m = n-1$. For the necessity, assume that X_n is a martingale w.r.t. $\mathcal{F}_n$ and let $n = m + k$. Clearly, the theorem holds for $k = 1$. Inductively, if $\mathbb{E}(X_{m+k}|\mathcal{F}_m) \overset{\text{a.s.}}{=} X_m$, then

$$\mathbb{E}(X_{m+k+1}|\mathcal{F}_m) \overset{\text{a.s.}}{=} \mathbb{E}(\mathbb{E}(X_{m+k+1}|\mathcal{F}_{m+k})|\mathcal{F}_m) \overset{\text{a.s.}}{=} X_m.$$

□

Theorem 2.3. *Assume that X_n is a martingale w.r.t. $\mathcal{F}_n$ and φ is a convex function on $\mathbb{R}$ satisfying $\mathbb{E}|\varphi(X_n)| < \infty$ for all $n \geq 0$. Then, $\varphi(X_n)$ is a submartingale w.r.t. $\mathcal{F}_n$.*

Proof. Since φ is convex, one has

$$\mathbb{E}(\varphi(X_n)|\mathcal{C}) \overset{\text{a.s.}}{\geq} \varphi(\mathbb{E}(X_n|\mathcal{C}))$$

for any σ-field $\mathcal{C} \subset \mathcal{F}$. This leads to

$$\mathbb{E}(\varphi(X_{n+1})|\mathcal{F}_n) \overset{\text{a.s.}}{\geq} \varphi(\mathbb{E}(X_{n+1}|\mathcal{F}_n)) \overset{\text{a.s.}}{=} \varphi(X_n).$$

□

Exercise 2.3. Let X_n be a submartingale w.r.t. $\mathcal{F}_n$. Assume that X_n takes values on an interval I for all n and φ is a non-decreasing convex function on I satisfying $\mathbb{E}|\varphi(X_n)| < \infty$. Show that $\varphi(X_n)$ is a submartingale w.r.t. $\mathcal{F}_n$. Give a similar statement for supermartingales?

Corollary 2.4. *Assume that X_n is a martingale w.r.t. $\mathcal{F}_n$. For $p \in [1, \infty)$, if $\mathbb{E}|X_n|^p < \infty$, then $|X_n|^p$ is a submartingale.*

Corollary 2.5. *For $s, t \in \mathbb{R}$, let $s \vee t$ and $s \wedge t$ be the maximum and minimum of s, t. Fix $a \in \mathbb{R}$.*

(1) *If X_n is a submartingale w.r.t. $\mathcal{F}_n$, then $X_n \vee a$ is a submartingale w.r.t. $\mathcal{F}_n$.*
(2) *If X_n is a supermartingale w.r.t. $\mathcal{F}_n$, then $X_n \wedge a$ is a supermartingale w.r.t. $\mathcal{F}_n$.*

Definition 2.3. A sequence of random variables H_n is *predictable* w.r.t. a filtration $\mathcal{F}_n$ if H_n is $\mathcal{F}_{n-1}$-measurable for $n \geq 1$.

Theorem 2.6. *Let X_n be a martingale (resp. supermartingale and submartingale) w.r.t. $\mathcal{F}_n$, H_n be predictable w.r.t. $\mathcal{F}_n$ and set*

$$S_0 = 0, \quad S_n = \sum_{i=1}^{n} H_i(X_i - X_{i-1}), \quad \forall n \geq 1.$$

Assume that H_n is nonnegative and bounded with probability one. Then, S_n is a martingale (resp. supermartingale and submartingale) w.r.t. $\mathcal{F}_n$.

Proof. Suppose that X_n is a submartingale w.r.t. $\mathcal{F}_n$. By the triangle inequality, one has $\mathbb{E}|S_n| < \infty$. This implies

$$\mathbb{E}(S_{n+1}|\mathcal{F}_n) \overset{\text{a.s.}}{=} S_n + \mathbb{E}(H_{n+1}(X_{n+1} - X_n)|\mathcal{F}_n) \overset{\text{a.s.}}{=} S_n + H_{n+1}\mathbb{E}(X_{n+1} - X_n|\mathcal{F}_n) \overset{\text{a.s.}}{\geq} S_n.$$

Other cases can be proved in a similar way and omitted. □

Remark 2.3. In the above theorem, $H_n \geq 0$ is not required for the case of martingales.

Remark 2.4. Referring to (2.1), Theorem 2.6 says that if the game is "unfavorable", i.e. the gambling system is a supermartingale, then the asset function is always a supermartingale whatever predictable strategy is applied.

Definition 2.4. A stopping time τ for a filtration $\mathcal{F}_n$ is defined to be a random variable taking values on $\{0, 1, 2, ...\} \cup \{\infty\}$ satisfying $\{\tau = n\} \in \mathcal{F}_n$ for all n.

Remark 2.5. Equivalently, τ is a stopping time for $\mathcal{F}_n$ if $\{\tau \leq n\} \in \mathcal{F}_n$ for all $n \geq 0$.

Theorem 2.7. *If X_n is a martingale (resp. supermartingale and submartingale) w.r.t. $\mathcal{F}_n$ and τ is a stopping time for $\mathcal{F}_n$, then $X_{\tau \wedge n}$ is a martingale (resp. supermartingale and submartingale) w.r.t. $\mathcal{F}_n$.*

Proof. For $n \geq 1$, set $H_n = \mathbf{1}_{\{\tau \geq n\}}$ and $S_n = \sum_{i=1}^{n} H_i(X_i - X_{i-1})$. It is clear that H_n is predictable w.r.t. $\mathcal{F}_n$. By Theorem 2.6, if X_n is a martingale w.r.t. $\mathcal{F}_n$, then so is S_n. The desired property is then given by the identity $S_n = X_{\tau \wedge n} - X_0$. □

2.2. **Optional sampling theorem.** In this subsection, we introduce the preservation of martingales through optional samplings.

Definition 2.5. For any process $X_0, X_1, X_2, ...$, we call a sequence $\tau_0, \tau_1, \tau_2, ...$ *sampling variables* of $(X_n)_{n=0}^{\infty}$ if τ_n is a random variable taking values on non-negative integers and satisfying

(1) $0 \leq \tau_0 \leq \tau_1 \leq \tau_2 \leq \cdots$;
(2) $\{\tau_k = j\} \in \mathcal{F}(X_0, ..., X_j)$ for all $j, k < \infty$.

Theorem 2.8 (Optional sampling theorem)**.** *Let $(X_n)_{n=0}^\infty$ be a process, $(\tau_n)_{n=0}^\infty$ be sampling variables of $(X_n)_{n=0}^\infty$ and $Y_n = X_{\tau_n}$. Assume that X_n is a martingale (resp. submartingale and supermartingale), $\mathbb{E}|Y_n| < \infty$ and*

$$\liminf_{k\to\infty} \int_{\{\tau_n > k\}} |X_k| d\mathbb{P} = 0 \quad \forall n \geq 1. \tag{2.2}$$

Then, Y_n is a martingale (resp. submartingale and supermartingale).

Proof. We prove the case of martingale by showing

$$\int_A Y_{n+1} d\mathbb{P} = \int_A Y_n d\mathbb{P}, \quad \forall A \in \mathcal{F}(Y_0, ..., Y_n).$$

Let $A \in \mathcal{F}(Y_0, ..., Y_n)$ and set $D_j = A \cap \{\tau_n = j\}$. To finish the proof, it suffices to show

$$\int_{D_j} Y_{n+1} d\mathbb{P} = \int_{D_j} Y_n d\mathbb{P} \quad \forall j \geq 0.$$

By writing $A = \{(Y_0, ..., Y_n) \in B\}$ for some $B \in \mathcal{B}(\mathbb{R}^{n+1})$, one has

$$\begin{aligned} D_j &= \{(Y_0, ..., Y_n) \in B,\, \tau_n = j\} \\ &= \bigcup_{0 \leq j_0 \leq \cdots \leq j_n = j} \{(X_{j_0}, ..., X_{j_n}) \in B,\, \tau_0 = j_0, ..., \tau_{n-1} = j_{n-1}, \tau_n = j_n\} \end{aligned}$$

This implies $D_j \in \mathcal{F}(X_0, ..., X_j)$ for all $j \geq 0$. Set $\int_{D_j} Y_{n+1} d\mathbb{P} = I_1(k) + I_2(k)$ for $k \geq j$, where

$$I_1(k) = \sum_{i=j}^{k} \int_{D_j \cap \{\tau_{n+1} = i\}} X_i d\mathbb{P} + \int_{D_j \cap \{\tau_{n+1} > k\}} X_k d\mathbb{P}$$

and

$$I_2(k) = \int_{D_j \cap \{\tau_{n+1} > k\}} Y_{n+1} d\mathbb{P} - \int_{D_j \cap \{\tau_{n+1} > k\}} X_k d\mathbb{P}.$$

Note that $\{\tau_{n+1} > k-1\} = \{\tau_{n+1} \leq k-1\}^c \in \mathcal{F}(X_0, ..., X_{k-1})$. By Theorem 2.2, this implies

$$\int_{D_j \cap \{\tau_{n+1} = k\}} X_k d\mathbb{P} + \int_{D_j \cap \{\tau_{n+1} > k\}} X_k d\mathbb{P} = \int_{D_j \cap \{\tau_{n+1} > k-1\}} X_{k-1} d\mathbb{P}, \quad \forall k > j,$$

or, equivalently, $I_1(k) = I_1(k-1)$ for $k > j$. Since $D_j \subset \{\tau_n \geq j\}$, we have

$$I_1(k) = I_1(j) = \int_{D_j} X_j d\mathbb{P} = \int_{D_j} Y_n d\mathbb{P}.$$

Combining all above gives

$$\int_{D_j} Y_{n+1} d\mathbb{P} - \int_{D_j} Y_n d\mathbb{P} = I_2(k) \quad \forall k > j.$$

By (2.2), one may choose a subsequence $(m_l)_{l=1}^\infty$ such that $\mathbb{E}(|X_{m_l}| \mathbf{1}_{\{\tau_{n+1} > m_l\}}) \to 0$. In addition with $\mathbb{E}|Y_{n+1}| < \infty$, this implies $I_2(m_l) \to 0$. □

Corollary 2.9. *Let $(X_n)_{n=0}^\infty$ be a submartingale, $(\tau_n)_{n=0}^\infty$ be sampling variables of $(X_n)_{n=0}^\infty$ and $Y_n = X_{\tau_n}$. Assume that $\sup_n \mathbb{E}X_n^+ < \infty$ and*

$$\liminf_{k\to\infty} \int_{\{\tau_n > k\}} |X_k| d\mathbb{P} = 0 \quad \forall n \geq 0. \tag{2.3}$$

Then Y_n is a submartingale. Moreover, one has

$$\mathbb{E}X_0 \leq \mathbb{E}Y_n \leq \sup_{n\geq 0} \mathbb{E}X_n, \quad \mathbb{E}|Y_n| \leq 2\sup_{n\geq 0} \mathbb{E}X_n^+ - \mathbb{E}X_0. \tag{2.4}$$

Proof. By Theorem 2.8, if the third inequality in (2.4) holds, then Y_n is a submartingale. Note that, by Theorem 2.7, $(X_{\tau_n\wedge m})_{m=0}^\infty$ is a submartingale and, hence,

$$\mathbb{E}X^-_{\tau_n\wedge m} \le \mathbb{E}X^+_{\tau_n\wedge m} - \mathbb{E}X_0, \quad \forall n, m \ge 0.$$

Next, fix $n \ge 0$. By (2.3), we may select a subsequence $(m_k)_{k=1}^\infty$ of $\mathbb{N}$ with $m_0 = 0$ such that

$$\lim_{k\to\infty} \int_{\{\tau_n > m_k\}} |X_{m_k}| d\mathbb{P} = 0.$$

Since X_n^+ is a submartingale, one has

$$\mathbb{E}Y_n^+ = \sum_{j=0}^\infty \int_{\{\tau_n=j\}} X_j^+ d\mathbb{P} \le \limsup_{k\to\infty} \sum_{j=0}^{m_k} \int_{\{\tau_n=j\}} X_{m_k}^+ d\mathbb{P} = \limsup_{k\to\infty} \mathbb{E}X_{m_k}^+ = \sup_{n\ge1} \mathbb{E}X_n^+ < \infty.$$

Note that

$$\mathbb{E}Y_n^+ = \lim_{k\to\infty} \left(\sum_{j=0}^{m_k} \int_{\{\tau_n=j\}} X_j^+ d\mathbb{P} + \int_{\{\tau_n>m_k\}} X_{m_k}^+ d\mathbb{P} \right) = \lim_{k\to\infty} \mathbb{E}X^+_{\tau_n\wedge m_k} \tag{2.5}$$

and, similarly,

$$\mathbb{E}Y_n^- = \lim_{k\to\infty} \mathbb{E}X^-_{\tau_n\wedge m_k} \le \limsup_{k\to\infty} \mathbb{E}X^+_{\tau_n\wedge m_k} - \mathbb{E}X_0 = \mathbb{E}Y_n^+ - \mathbb{E}X_0. \tag{2.6}$$

Putting all above together yields

$$\mathbb{E}|Y_n| \le 2\mathbb{E}Y_n^+ - \mathbb{E}X_0 \le 2\sup_{n\ge0} \mathbb{E}X_n^+ - \mathbb{E}X_0.$$

This proves the third inequality of (2.4).

By Theorem 2.8, $(Y_n)_{n=0}^\infty$ is a submartingale and, by (2.5) and (2.6), one has

$$\mathbb{E}Y_n = \lim_{k\to\infty} \mathbb{E}X_{\tau_n\wedge m_k} \ge \mathbb{E}X_{\tau_n\wedge 0} \ge \mathbb{E}X_0.$$

Since $(X_n)_{n=0}^\infty$ is a submartingale and τ_n is a stopping time for $(X_n)_{n=0}^\infty$, we have

$$\begin{aligned}
\mathbb{E}X_{\tau_n\wedge m_k} &= \sum_{j=0}^{m_k} \int_{\{\tau_n=j\}} X_j d\mathbb{P} + \int_{\{\tau_n>m_k\}} X_{m_k} d\mathbb{P} \\
&\le \sum_{j=0}^{m_k} \int_{\{\tau_n=j\}} X_{m_k} d\mathbb{P} + \int_{\{\tau_n>m_k\}} X_{m_k} d\mathbb{P} = \mathbb{E}X_{m_k} \le \sup_{n\ge0} \mathbb{E}X_n.
\end{aligned}$$

Passing k to the infinity gives the second inequality of (2.4). □

The following proposition should be considered as a generalization of Kolmogorov's inequality and Chebyshev's inequality.

Proposition 2.10 (Doob's inequality)**.** *Let X_n be a submartingale. For $\epsilon > 0$, one has*

$$\mathbb{P}\left(\max_{0\le j\le k} X_j > \epsilon \right) \le \frac{\mathbb{E}X_k^+}{\epsilon}, \quad \mathbb{P}\left(\min_{0\le j\le k} X_j < -\epsilon \right) \le \frac{\mathbb{E}X_k^+ - \mathbb{E}X_0}{\epsilon}.$$

Proof. For the first inequality, let $\tau_0, \tau_1, ...$ be sampling variables defined by

$$\tau_0 = \inf\{j \le k : X_j > \epsilon\}$$

and $\tau_0 = k$ if the infimum is taken over an empty set. For $n \ge 1$, set $\tau_n = k$. Then, $Y_n := X_{\tau_n}$ satisfies

$$\mathbb{E}|Y_n| \le \sum_{j=0}^k \mathbb{E}|X_j| < \infty \quad \forall n \ge 0, \qquad \int_{\{\tau_n > \ell\}} |X_\ell| d\mathbb{P} = 0 \quad \forall \ell > k.$$

By the optional sampling theorem, Y_n is a submartingale and

$$\mathbb{P}\left(\max_{0\le j\le k} X_j > \epsilon\right) = \mathbb{P}(Y_0 > \epsilon) \le \frac{\int_{\{Y_0>\epsilon\}} Y_0 d\mathbb{P}}{\epsilon} \le \frac{\int_{\{Y_0>\epsilon\}} Y_1 d\mathbb{P}}{\epsilon} \le \frac{\mathbb{E}X_k^+}{\epsilon}.$$

For the second inequality, we set $\tau'_0 = 0$ and $\tau'_n = k$ for $n \ge 2$. For $n = 1$, set

$$\tau'_1 = \inf\{j \le k : X_j < -\epsilon\}$$

and $\tau'_1 = k$ if the infimum is taken over an empty set. Let $Z_n = X_{\tau'_n}$. As before, we have

$$\mathbb{E}|Z_n| < \infty \quad \forall n \ge 1, \qquad \int_{\{\tau'_n > \ell\}} |X_\ell| d\mathbb{P} = 0 \quad \forall \ell > k.$$

By Theorem 2.8, Z_n is a submartingale and

$$\begin{aligned}\mathbb{E}X_0 \le \mathbb{E}Z_1 &= \int_{\{Z_1\ge-\epsilon\}} Z_1 d\mathbb{P} + \int_{\{Z_1<-\epsilon\}} Z_1 d\mathbb{P} \\ &\le \int_{\{Z_1\ge-\epsilon\}} X_k d\mathbb{P} - \epsilon\mathbb{P}\left(\min_{0\le j\le k} X_j < -\epsilon\right) \le \mathbb{E}X_k^+ - \epsilon\mathbb{P}\left(\min_{0\le j\le k} X_j < -\epsilon\right).\end{aligned}$$

□

Exercise 2.4. Prove that if $X_0, X_1, ...$ is a martingale satisfying $\mathbb{E}X_n^2 < \infty$, then

$$\mathbb{P}\left(\max_{0\le j\le k} |X_j| > \epsilon\right) \le \frac{\mathbb{E}X_k^2}{\epsilon^2} \quad \forall \epsilon > 0.$$

2.3. **Martingale convergence theorem.** We use Theorem 2.6 to derive the convergence theorem of martingales.

Lemma 2.11 (Upcrossing lemma)**.** *Let X_n be a submartingale and a, b be real numbers satisfying $a < b$. Set $N_0 = -1$ and, for $k \ge 1$, define*

$$N_{2k-1} = \inf\{j > N_{2k-2} : X_j \le a\}, \quad N_{2k} = \inf\{j > N_{2k-1} : X_j \ge b\},$$

where $\inf\emptyset := \infty$. For $n \ge 1$, set $U_n = \sup\{k \ge 0 : N_{2k} \le n\}$. Then, $(b-a)\mathbb{E}U_n \le \mathbb{E}(X_n - a)^+ - \mathbb{E}(X_0 - a)^+$.

Proof. First, it is easy to see that $(N_j)_{j=0}^\infty$ are stopping times for $\mathcal{F}(X_0, ..., X_n)$. Let $Y_n = X_n \vee a$. Since X_n is a submartingale, Y_n is a submartingale. For $m \ge 0$, set

$$H_m = \begin{cases} 1 & \text{if } N_{2k-1} < m \le N_{2k} \text{ for some } k \\ 0 & \text{o.w.} \end{cases}$$

Observe that

$$\{H_m = 1\} = \{N_{2k-1} < m \le N_{2k}\} = \{N_{2k-1} < m\} \cap \{N_{2k} < m\}^c \in \mathcal{F}(X_0, ..., X_{m-1}).$$

This implies that H_n is predictable w.r.t. $\mathcal{F}(X_0, ..., X_n)$. For $n \ge 0$, set $S_0 = S'_0 = 0$ and

$$S_n = \sum_{i=1}^n H_i(Y_i - Y_{i-1}), \quad S'_n = \sum_{i=1}^n (1 - H_i)(Y_i - Y_{i-1}).$$

By Theorem 2.6, S_n, S'_n are submartingales. Note that $S_n \ge (b-a)U_n$ and $Y_n - Y_0 = S_n + S'_n$. This implies $\mathbb{E}S'_n \ge \mathbb{E}S'_0 = 0$ and

$$(b-a)\mathbb{E}U_n \le \mathbb{E}S_n \le \mathbb{E}Y_n - \mathbb{E}Y_0 = \mathbb{E}(X_n - a)^+ - \mathbb{E}(X_0 - a)^+.$$

□

Theorem 2.12 (Martingale convergence theorem)**.** *Let X_n be a submartingale satisfying $\sup_n \mathbb{E}X_n^+ < \infty$. Then, X_n converges a.s. to some random variable X with $\mathbb{E}|X| < \infty$.*

Proof. Let U_n be the random variable defined in Lemma 2.11 and $U = \sup\{k \geq 0 : N_{2k} < \infty\}$. Note that U_n is non-decreasing, non-negative and converges to U. By the monotone convergence theorem and the upcrossing lemma, we have

$$\mathbb{E}U = \lim_{n\to\infty} \mathbb{E}U_n \leq \frac{1}{b-a}\left(|a| + \sup_{n\geq 0}\mathbb{E}X_n^+\right) < \infty, \quad \forall a < b.$$

This implies $U < \infty$ a.s. for any $a < b$. As a result, the following event

$$\bigcup_{a,b\in\mathbb{Q},\, a<b} \left\{\omega \in \Omega : \liminf_{n\to\infty} X_n(\omega) < a < b < \limsup_{n\to\infty} X_n(\omega)\right\}$$

has probability 0 and then $\liminf_n X_n = \limsup_n X_n$ a.s.. Set $X = \lim_n X_n$. By Fatou's lemma, we have

$$\mathbb{E}|X| = \mathbb{E}\liminf_{n\to\infty} |X_n| \leq \liminf_{n\to\infty} \mathbb{E}|X_n| \leq \sup_{n\geq 0}\mathbb{E}|X_n| < \infty.$$

□

Remark 2.6. For any process $(X_t)_{t\in[S,T]}$ with $S < \infty$ and $T \in (S,\infty]$, set

$$U_{a,b} = \sup\{k : N_{2k}(a,b) < \infty\},$$

where $N_0(a,b) = S$ and

$$N_{2k-1}(a,b) = \inf\{N_{2k-2}(a,b) < t < T : X_t \leq a\}$$

and

$$N_{2k}(a,b) = \inf\{N_{2k-1}(a,b) < t < T : X_t \geq b\}.$$

If $U_{a,b} < \infty$ a.s. for all rational numbers $a < b$, then $\limsup_{t\to T} X_t = \liminf_{t\to T} X_t$ a.s., but the converse is generally not true. However, in the case of $\{0, 1, 2...\}$, $\limsup_n X_n = \liminf_n X_n$ a.s. if and only if $U_{a,b} < \infty$ a.s. all rational numbers $a < b$.

Corollary 2.13. *If X_n is a submartingale uniformly bounded from above, then X_n converges a.s. to some random variable X with $\mathbb{E}|X| < \infty$ and $\mathbb{E}X \geq \mathbb{E}X_0$.*

Proof. The almost sure convergence of X_n comes immediately from Theorem 2.12. To see the last inequality, let $M > 0$ be a constant such that $\sup_n X_n \leq M$ a.s.. Clearly, $X \leq M$ and, by Fatou's lemma, one has

$$0 \leq \mathbb{E}(M - X) \leq \liminf_{n\to\infty} \mathbb{E}(M - X_n) \leq M - \mathbb{E}X_0 < \infty.$$

This implies $\mathbb{E}X \geq \mathbb{E}X_0$. □

Example 2.3. (Martingales that converge a.s. but not in L^1) Let $X_1, X_2, ...$ be i.i.d. random variables with $\mathbb{P}(X_1 = 1) = \mathbb{P}(X_1 = -1) = 1/2$. Let $d \in \mathbb{Z}^+$ and set

$$S_n = d + \sum_{i=1}^{n} X_i, \quad \forall n \geq 1.$$

Clearly, S_n is a martingale. Consider the stopping time $N = \min\{n \geq 0 : S_n = 0\}$ and set $Y_n = S_{N\wedge n}$. By Theorem 2.7, Y_n is a martingale. Since Y_n is nonnegative, Corollary 2.13 implies that Y_n converges almost surely to some random variable Y satisfying $\mathbb{E}|Y| < \infty$. Note that $|S_{n+1} - S_n| = 1$ for all $n \geq 1$. This implies $\{Y_n \text{ converges}\} \subset \{N < \infty\}$ and, hence,

$$\mathbb{P}(N < \infty) = 1, \quad Y = \lim_{n\to\infty} S_{N\wedge n} = S_N = 0 \quad a.s..$$

But, $\mathbb{E}Y_n = EY_0 = d > 0$ and this means that Y_n does not converge in L^1.

Example 2.4 (Martingales that converge in probability but not a.s.). Let $\Omega = [0, 1)$, $\mathcal{F}$ be the Borel σ-field over Ω and $\mathbb{P}$ be the Lebesgue measure on $(\Omega, \mathcal{F})$. Let $X_0 \equiv 0$ and $X_1, X_2, X_3, ...$ be defined iteratively as follows. For $n \geq 0$, let $a_{n,1} = 0$, $a_{n,i_n} = 1$ and let $a_{n,2}, ..., a_{n,i_n-1}$ be discontinuous points of X_n. If $X_n(a_{n,j}) = 0$, set

$$X_{n+1}(t) = \mathbf{1}_{[a_{n,j}, a_{n,j}+\epsilon)}(t) - \mathbf{1}_{[a_{n,j}+\epsilon, a_{n,j}+2\epsilon)}(t), \quad \forall t \in [a_{n,j}, a_{n,j+1}),$$

where $\epsilon = \frac{a_{n,j+1} - a_{n,j}}{2(n+1)}$. If $X_n(a_{n,j}) \neq 0$, set

$$X_{n+1}(t) = (n+1) X_n(a_{n,j}) \mathbf{1}_{[a_{n,j}, a_{n,j}+\epsilon)}(t), \quad \forall t \in [a_{n,j}, a_{n,j+1}),$$

where $\epsilon = \frac{a_{n,j+1} - a_{n,j}}{n+1}$. This process has $X_1 = \mathbf{1}_{[0,1/2)} - \mathbf{1}_{[1/2,1)}$ and $X_2 = 2(\mathbf{1}_{[0,1/4)} - \mathbf{1}_{[1/2,3/4)})$. One can show by induction that X_n is an integer-valued martingale.

Let $A_n = \{a_{n,j} : \forall 1 \leq j \leq i_n, n \geq 1\}$ and $A = \bigcup_n A_n$. Clearly, $A_n \subset A_{n+1}$ and $|A_{n+1}| = 2|A_n| - 1$. This implies that A is a countable set. Observe that, for $t \in A^c$, if $X_m(x) = 0$, then $|X_n(t)| = 1$ for some $n > m$. If $X_m(t) \neq 0$, then $X_n(t) = 0$ for some $n > m$. This implies that X_n diverges on A^c. As $\mathbb{P}(X_n \neq 0) = 1/n$, X_n converges to 0 in probability.

Example 2.5 (Martingales that tend to infinity). Let $X_1, X_2, ...$ be independent random variables with

$$\mathbb{P}(X_n = n) = n^{-2}, \quad \mathbb{P}(X_n = -(n - n^{-1})^{-1}) = 1 - n^{-2}.$$

and set $S_n = X_1 + \cdots + X_n$. As $\mathbb{E}X_n = 0$, S_n forms a martingale. By the Borel-Cantelli lemma, $\mathbb{P}(X_n = n \text{ i.o.}) = 0$. This implies that, with probability 1,

$$X_n = -\frac{1}{n - 1/n} \leq -\frac{1}{n} \quad \text{for } n \text{ large enough.}$$

Hence, $S_n \to -\infty$ almost surely.

Exercise 2.5. Let $X_1, X_2, ...$ be independent random variable satisfying $\mathbb{E}X_n = 0$ and $\mathbb{E}X_n^2 < \infty$. For $n \geq 1$, set $S_n = \sum_{i=1}^n X_i$ and $s_n^2 = \sum_{i=1}^n \mathbb{E}X_i^2$. Show that

(1) $S_n^2 - s_n^2$ is a martingale;
(2) If s_n converges, then S_n converges almost surely.

Exercise 2.6. Let $X_1, X_2, ...$ be i.i.d. non-negative random variables satisfying $\mathbb{E}X_1 = 1$ and $\mathbb{P}(X_1 = 1) < 1$. For $n \geq 1$, set $Y_n = \prod_{i=1}^n X_i$. Show that

(1) Y_n is a martingale,
(2) Y_n converges 0 a.s.,
(3) $\frac{1}{n} \log Y_n$ converges a.s. to $c \in [-\infty, 0)$.

2.4. **Branching processes.** Let $\{\xi_i^n : n \geq 0, i \geq 0\}$ be a family of i.i.d. nonnegative integer valued random variables. Set $Z_0 = 1$ and define, for $n \geq 0$,

$$Z_{n+1} := \sum_{i=1}^{Z_n} \xi_i^{n+1} \quad \text{if } Z_n > 0, \quad Z_{n+1} := 0 \quad \text{if } Z_n = 0.$$

In the above setting, Z_n is called a *Galton-Watson* process. An idea behind this definition is that Z_n refers to the number of people at the nth generation and each member gives birth independently to an identically distributed number of children. After that, all individuals of the nth generation pass away and the number of new-born offsprings amounts to Z_{n+1}. To analyze this process, we set $p_k = \mathbb{P}(\xi_i^n = k)$ and call $(p_k)_{k=0}^\infty$ the *offspring distribution.*

Proposition 2.14. *Let $\mathcal{F}_n = \mathcal{F}(\xi_i^m, i \geq 0, 0 \leq m \leq n)$ and $\mu = \mathbb{E}\xi_i^n$. Assume that $\mu \in (0, \infty)$. Then Z_n/μ^n is a martingale w.r.t. $\mathcal{F}_n$.*

Proof. It is clear that Z_n is adapted to $\mathcal{F}_n$. To see $\mathbb{E}Z_n < \infty$, observe that $\mathbb{E}Z_0 = 1 < \infty$ and, for $n \geq 0$,

$$\mathbb{E}Z_{n+1} = \sum_{k=1}^{\infty} \mathbb{E}\left(\sum_{i=1}^{k} \xi_i^{n+1}\mathbf{1}_{\{Z_n=k\}}\right) = \sum_{i=1}^{\infty} k\mu\mathbb{P}(Z_n = k) = \mu\mathbb{E}Z_n < \infty,$$

where the first equality uses the monotone convergence theorem. By induction, we obtain $\mathbb{E}Z_n = \mu^n < \infty$, which implies

$$\mathbb{E}(Z_{n+1}|\mathcal{F}_n) = \sum_{k=1}^{\infty}\sum_{i=1}^{k} \mathbb{E}(\xi_i^{n+1}\mathbf{1}_{\{Z_n=k\}}|\mathcal{F}_n) = \sum_{k=1}^{\infty} k\mu\mathbf{1}_{\{Z_n=k\}} = \mu Z_n.$$

This proves that Z_n/μ^n is a martingale w.r.t. $\mathcal{F}_n$. □

Concerning the extinction of spices, we define the stopping time $T = \inf\{n \geq 0 : Z_n = 0\}$. Note that, when $0 < \mu < \infty$, Z_n/μ^n is nonnegative with mean 1. By the martingale convergence theorem, Z_n/μ^n converges almost surely to some integrable random variable.

Theorem 2.15. *If $\mu \in (0, 1)$, then $\mathbb{P}(T < \infty) = 1$ and $Z_n/\mu^n \to 0$ almost surely.*

Proof. By Proposition 2.14, $\mathbb{E}Z_n = \mu^n\mathbb{E}Z_0 = \mu^n$. This implies

$$\mathbb{P}(Z_n > 0) = \mathbb{P}(Z_n \geq 1) \leq \mathbb{E}Z_n = \mu^n \to 0 \quad \text{as } n \to \infty,$$

which means that Z_n converges to 0 in probability. Let k_n be a subsequence such that Z_{k_n} converges to 0 almost surely. As Z_n is integer-valued and $Z_n = 0$ implies $Z_{n+1} = 0$, we may conclude that $\mathbb{P}(Z_n = 0 \text{ for some } n \geq 0) = 1$ or $\mathbb{P}(Z_n/\mu^n = 0 \text{ for some } n \geq 0) = 1$. □

Theorem 2.16. *If $\mu = 1$ and $p_1 < 1$, then $\mathbb{P}(T < \infty) = 1$.*

Proof. By Proposition 2.14, Z_n is a martingale and, by the martingale convergence theorem, Z_n converges almost surely to some integrable random variable Z_∞. Since Z_n is integer valued, we have $\mathbb{P}(Z_n = Z_\infty \text{ for } n \text{ large enough}) = 1$. To finish the proof, it remains to show that $\mathbb{P}(Z_\infty = 0) = 1$. Note that, for $k > 0$,

$$\mathbb{P}(Z_{n+1} = k|Z_n = k, Z_{n-1} = a_{n-1}, ..., Z_1 = a_1) = \mathbb{P}(\xi_1^{n+1} + \cdots + \xi_k^{n+1} = k) =: c_k < 1,$$

where the last inequality uses the assumption of $\mu = 1$ and $p_1 < 1$. This implies

$$\mathbb{P}(Z_n = k, \forall n \geq N) = \lim_{\ell\to\infty} \mathbb{P}(Z_{N+\ell} = \cdots = Z_N = k) = \mathbb{P}(Z_N = k) \lim_{\ell\to\infty} c_k^\ell = 0, \quad \forall N > 0.$$

As a result, we obtain $\mathbb{P}(Z_\infty = k) = \mathbb{P}(Z_n = k \text{ for } n \text{ large enough}) = 0$ for $k > 0$. □

Remark 2.7. Note that, when $\mu = 1$, $p_1 < 1$ is equivalent to $p_0 > 0$.

Theorem 2.17. *If $\mu \in (1, \infty)$, then $\mathbb{P}(T = \infty) > 0$.*

Proof. Consider the following *generating function* of the offspring distribution.

$$\varphi(s) = p_0 + \sum_{k=1}^{\infty} p_k s^k, \quad \forall |s| \leq 1.$$

Clearly, φ is analytic on $(-1, 1)$ and

$$\varphi'(s) = p_1 + \sum_{k=2}^{\infty} kp_k s^{k-1}, \quad \varphi''(s) = 2p_2 + \sum_{k=3}^{\infty} k(k-1)p_k s^{k-2}, \quad \forall |s| < 1.$$

As $\mu > 1$, we must have $p_k > 0$ for some $k \geq 2$. This implies that φ is increasing and convex on $(0, 1)$ and

$$\lim_{s<1,s\to 1} \frac{\varphi(s) - \varphi(1)}{s - 1} = \lim_{s<1,s\to 1} \sum_{k=1}^{\infty} p_k(1 + s + \cdots + s^{k-1}) = \mu > 1.$$

Since $\varphi(1) = 1$ and $\varphi(0) = p_0 \in [0, 1)$, there is $\rho \in [0, 1)$ such that $\varphi(\rho) = \rho$. Moreover, one has $\varphi(x) > x$ for $x \in (0, \rho)$ and $\varphi(x) < x$ for $x \in (\rho, 1)$.

Next, we set $\theta_m = \mathbb{P}(Z_m = 0)$. Obviously, $\theta_0 = 0$. Note that each individual of the offspring generation can be regarded as an ancestor of a Galton-Watson process. This implies

$$\theta_m = p_0 + \sum_{k=1}^{\infty} p_k \theta_{m-1}^k = \varphi(\theta_{m-1}) \quad \forall m \geq 1.$$

If $\rho = 0$, then $\theta_m = 0$ for all m. If $\rho > 0$, then $0 \leq \theta_m < \theta_{m+1} < \rho$. Let θ_∞ be the limit of φ_m. By the continuity of φ, one has $\varphi(\theta_\infty) = \theta_\infty$ and this implies $\theta_\infty = \rho$. Consequently, we achieve

$$\mathbb{P}(T < \infty) = \lim_{m\to\infty} \mathbb{P}(Z_m = 0) = \rho < 1.$$

□

Corollary 2.18. *For $p_0 > 0$, the probability that the species is finally extinct equals the smallest fixed point of the generating function of the offspring distribution on $[0, 1]$.*

Proof. The proof for $\mu > 1$ has been given in the proof of Theorem 2.17. Let $\varphi(s) = p_0 + \sum_{k=1}^{\infty} p_k s^k$. We first assume $\mu \in [0, 1)$. In this case, it is clear that $\varphi'(1) < 1$ and this implies that φ has only one fixed point on $[0, 1]$, which is 1.

Next, we consider the case $\mu = 1$ and $p_1 < 1$. Clearly, $p_0 > 0$ and $p_k > 0$ for some $k \geq 2$. This implies that φ is strictly convex on $(0, 1)$ and, hence, 1 is the unique fixed point of φ. When $p_1 = 1$, all points in $[0, 1]$ are fixed points of φ and 0 is the smallest one, while $\mathbb{P}(T < \infty) = 0$. □

Remark 2.8. Galton and Watson who invented the process that bears their names were interested in the survival of family names. Suppose that each family has exactly m children but their sex is determined by a fair coin. In 1800s, only male children kept the family name. The male offsprings then lead to an offspring distribution

$$p_k = \binom{m}{k} 2^{-m} \quad \forall 0 \leq k \leq m.$$

By Theorems 2.15-2.17, a family name disappears when $m \leq 2$ and has a positive probability to survive forever when $m > 2$.

2.5. **Uniform integrability.** In this subsection, we use $\mathbb{E}(X; A)$ to denote $\int_A X d\mathbb{P}$. If $A = \{X \in B\}$ for some $B \in \mathcal{B}(\mathbb{R})$, we simple write $\mathbb{E}(X; X \in B)$ for $\mathbb{E}(X; \{X \in B\})$.

Definition 2.6. A family of random variables $\{X_i : i \in I\}$ is *uniformly integrable* if

$$\lim_{x\to\infty} \sup_{i\in I} \mathbb{E}(|X_i|; |X_i| > x) = 0.$$

Exercise 2.7. Show that $(X_n)_{n=0}^{\infty}$ is uniformly integrable if and only if $\mathbb{E}|X_n| < \infty$ for all n and $\limsup_n \mathbb{E}(|X_n|; |X_n| > x) \to 0$ as $x \to \infty$.

Exercise 2.8. Prove the following statements.

(1) If $(X_i)_{i\in I}$ is uniformly integrable, then $\sup_{i\in I} \mathbb{E}|X_i| < \infty$.
(2) If $|X_i| \leq X$ for all $i \in I$ and $\mathbb{E}X < \infty$, then $(X_i)_{i\in I}$ is uniformly integrable.

Proposition 2.19. *If X is a random variable on $(\Omega, \mathcal{F}, \mathbb{P})$ with $\mathbb{E}|X| < \infty$, then the family $\{\mathbb{E}(X|\mathcal{G}) : \mathcal{G}$ is a sub-σ-field of $\mathcal{F}\}$ is uniformly integrable.*

Proof. Let $\epsilon > 0$. Since X is integrable, one may select $\delta > 0$ such that $\mathbb{E}(|X|; A) < \epsilon$ for all $A \in \mathcal{F}$ satisfying $\mathbb{P}(A) < \delta$. Note that, for any $x > \mathbb{E}|X|/\delta$, one may use the Markov inequality and Jensen's inequality to get

$$\mathbb{P}(\mathbb{E}(|X||\mathcal{G}) > x) \le \frac{\mathbb{E}(\mathbb{E}(|X||\mathcal{G}))}{x} = \frac{\mathbb{E}|X|}{x} < \delta,$$

and

$$\mathbb{E}(|\mathbb{E}(X|\mathcal{G})|; |\mathbb{E}(X|\mathcal{G})| > x) \le \mathbb{E}(\mathbb{E}(|X||\mathcal{G}); \mathbb{E}(|X||\mathcal{G}) > x) = \mathbb{E}(|X|; \mathbb{E}(|X||\mathcal{G}) > x) < \epsilon,$$

for any sub-σ-field $\mathcal{G} \subset \mathcal{F}$. This implies $\sup_{\mathcal{G} \subset \mathcal{F}} \mathbb{E}(|\mathbb{E}(X|\mathcal{G})|; |\mathbb{E}(X|\mathcal{G})| > x) \le \epsilon$ for $x > \mathbb{E}|X|/\delta$. □

Exercise 2.9. Let $\varphi : [0, \infty) \to [0, \infty)$ be a function satisfying $\varphi(x)/x \to \infty$ as $x \to \infty$. (For instance, $\varphi(x) = x^p$ with $p \in (1, \infty)$ and $\varphi(x) = x \log^+ x := x(\log x)^+$.) Prove that if $\sup_i \mathbb{E}\varphi(|X_i|) < \infty$, then $(X_i)_{i \in I}$ is uniformly integrable.

Proposition 2.20. *Let X_n, X be real-valued random variables. Assume that X_n converges to X in probability and $\mathbb{E}|X_n| < \infty$. Then, the following are equivalent.*

(1) $(X_n)_{n=0}^\infty$ *is uniformly integrable.*
(2) X_n *converges to* X *in* L^1.
(3) $\mathbb{E}|X_n| \to \mathbb{E}|X|$ *with* $\mathbb{E}|X| < \infty$.

Proof. For (1)⇒(2), let $\varphi_M(x) = (x \wedge M) \vee (-M)$ for $M > 0$. Clearly, $|\varphi_M(x) - x| = (|x| - M)^+ \le |x|\mathbf{1}_{[M,\infty)}(|x|)$. By the triangle inequality, one has

$$|X_n - X| \le |X_n - \varphi_M(X_n)| + |\varphi_M(X_n) - \varphi_M(X)| + |\varphi_M(X) - X|.$$

To prove $\mathbb{E}|X_n - X| \to 0$, it is equivalent to show that, for any subsequence k_n, there exists a further subsequence k'_n such that $\mathbb{E}|X_{k'_n} - X| \to 0$. Let k_n be a subsequence of $\{1, 2, ...\}$. Since $X_n \to X$ in probability, we may choose a subsequence k'_n of k_n such that $X_{k'_n}$ converges to X a.s.. By the Lebesgue dominated convergence theorem and Fatou's lemma, this implies

$$\lim_{n\to\infty} \mathbb{E}|\varphi_M(X_{k'_n}) - \varphi_M(X)| = 0, \quad \forall M > 0, \quad \mathbb{E}|X| \le \liminf_{n\to\infty} \mathbb{E}|X_{k'_n}| < \infty,$$

where the second inequality uses the uniform integrability of $(X_n)_{n=1}^\infty$. Putting all above together yields

$$\limsup_{n\to\infty} \mathbb{E}|X_{k'_n} - X| \le \sup_{n\ge1} \mathbb{E}(|X_{k'_n}|; |X_{k'_n}| > M) + \mathbb{E}(|X|; |X| > M).$$

Passing $M \to \infty$ gives $\mathbb{E}|X_{k'_n} - X| \to 0$.

(2)⇒(3) is obvious. For (3)⇒(1), let ψ_M be a piecewise linear function on $[0, \infty)$ defined by

$$\psi_M(x) = \begin{cases} x & \text{for } x \in [0, M-1] \\ (M-1)(M-x) & \text{for } x \in [M-1, M] \\ 0 & \text{for } x \ge M \end{cases}$$

Obviously, $x\mathbf{1}_{(M,\infty)}(x) \le x - \psi_M(x) \le x\mathbf{1}_{(M-1,\infty)}(x)$ for $x \ge 0$. A similar argument as before implies $\mathbb{E}\psi_M(|X_n|) \to \mathbb{E}\psi_M(|X|)$. This leads to

$$\begin{aligned} \limsup_{n\to\infty} \mathbb{E}(|X_n|; |X_n| > M) &\le \limsup_{n\to\infty} [\mathbb{E}|X_n| - \mathbb{E}\psi_M(|X_n|)] \\ &= \mathbb{E}|X| - \mathbb{E}\psi_M(|X|) \le \mathbb{E}(|X|; |X| > M-1). \end{aligned}$$

Since $\mathbb{E}|X| < \infty$, letting $M \to \infty$ gives (1). □

Theorem 2.21. *Let X_n be a submartingale. Then, the following are equivalent.*

(1) $(X_n)_{n=0}^{\infty}$ *is uniformly integrable;*
(2) X_n *converges a.s. and in* L^1*;*
(3) X_n *converges in* L^1.

Proof. First, suppose (1) holds. By Exercise 2.8 and Theorem 2.12, X_n converges a.s. and, by Proposition 2.20, X_n converges in L^1. This proves (2). (2)⇒(3) is obvious and (3)⇒(1) follows immediately from Proposition 2.20. □

Proposition 2.22. *If X_n is a martingale w.r.t. $\mathcal{F}_n$ and X_n converges to X_∞ in L^1, then $X_n = \mathbb{E}(X_\infty|\mathcal{F}_n)$ for $n \geq 0$.*

Proof. By Theorem 2.2, $\mathbb{E}(X_n|\mathcal{F}_m) = X_m$ for all $n > m$. Since X_n converges to X_∞ in L^1, $\mathbb{E}(X_n|\mathcal{F}_m) \to \mathbb{E}(X|\mathcal{F}_m)$ in L^1 as $n \to \infty$. This implies $X_m = \mathbb{E}(X_\infty|\mathcal{F}_m)$ almost surely. □

Corollary 2.23. *For any martingale X_n w.r.t. $\mathcal{F}_n$, the following are equivalent.*

(1) $(X_n)_{n=0}^{\infty}$ *is uniformly integrable;*
(2) X_n *converges a.s. and in* L^1*;*
(3) X_n *converges in* L^1*;*
(4) *There exists X with $\mathbb{E}|X| < \infty$ such that $X_n = \mathbb{E}(X|\mathcal{F}_n)$ for all n.*

Proof. The equivalence of (1), (2) and (3) is given by Theorem 2.21. (3)⇒(4) is given by Proposition 2.22 and (4)⇒(1) is implied by Proposition 2.19. □

Theorem 2.24. *Let $(\Omega, \mathcal{F}, P)$ be a probability space, X be a random variable on Ω and $\mathcal{F}_n$ be a filtration in $\mathcal{F}$. Set $\mathcal{F}_\infty = \sigma(\bigcup_n \mathcal{F}_n)$ and assume $\mathbb{E}|X| < \infty$. Then, $\mathbb{E}(X|\mathcal{F}_n)$ converges to $\mathbb{E}(X|\mathcal{F}_\infty)$ almost surely and in L^1.*

Proof. Set $Y_n = \mathbb{E}(X|\mathcal{F}_n)$. It is clear that, for $n > m$,

$$\mathbb{E}(Y_n|\mathcal{F}_m) = \mathbb{E}(\mathbb{E}(X|\mathcal{F}_n)|\mathcal{F}_m) = \mathbb{E}(X|\mathcal{F}_m) = Y_m.$$

This implies that Y_n is a martingale w.r.t. $\mathcal{F}_n$. By Proposition 2.19, $(Y_n)_{n=0}^{\infty}$ is uniformly integrable and, by Theorem 2.21, Y_n converges almost surely and in L^1 to some random variable Y_∞. By Proposition 2.22, $\mathbb{E}(X|\mathcal{F}_n) = \mathbb{E}(Y_\infty|\mathcal{F}_n)$ a.s. for all n. Consider the following class.

$$\mathcal{D} = \{A \in \mathcal{F}_\infty : \mathbb{E}(X; A) = \mathbb{E}(Y_\infty; A)\}.$$

Note that $\mathcal{D}$ is a λ-system, $\bigcup_n \mathcal{F}_n$ is a π-system and $\bigcup_n \mathcal{F}_n \subset \mathcal{D}$. By the $\pi - \lambda$ lemma, $\mathcal{D} = \mathcal{F}_\infty$. As Y_∞ is $\mathcal{F}_\infty$-measurable, we obtain $Y_\infty = \mathbb{E}(X|\mathcal{F}_\infty)$. □

Corollary 2.25 (Lévy's zero-one law)**.** *Let $\mathcal{F}_n$ be a filtration in $\mathcal{F}$ and $A \in \mathcal{F}_\infty = \sigma(\bigcup_n \mathcal{F}_n)$. Then, $\mathbb{E}(\mathbf{1}_A|\mathcal{F}_n)$ converges almost surely to $\mathbf{1}_A$.*

Remark 2.9. Let $X_1, X_2, ...$ be independent random variables and $\mathcal{F}_n = \mathcal{F}(X_1, ..., X_n)$. Let A be a tail event, that is, $A \in \cap_n \mathcal{F}(X_n, X_{n+1}, ...)$. Then, $\mathbb{E}(\mathbf{1}_A|\mathcal{F}_n) = \mathbb{P}(A)$ almost surely. By Lévy's zero-one law, we have $\mathbb{P}(A) = \mathbf{1}_A$ almost surely. This implies $\mathbb{P}(A) \in \{0, 1\}$, which is exactly Kolmogorov's zero-one law.

Exercise 2.10. Let X be a random variable on the probability space $([0,1), \mathcal{B}([0,1)), \mathbb{P})$, where $\mathbb{P}$ is the Lebesgue measure on $[0, 1)$. Assume that $\mathbb{E}|X| < \infty$ and set, for $n \geq 1$,

$$X_n(x) = 2^n \mathbb{E}(X; [k2^{-n}, (k+1)2^{-n})), \quad \forall x \in [k2^{-n}, (k+1)2^{-n}),\ 0 \leq k < 2^n.$$

Prove that X_n converges to X almost surely and in L^1. Using a variant of this exercise, one may conclude that, for any integrable function f on $\mathbb{R}$ and $\epsilon > 0$, there is a step function g such that $\int_{\mathbb{R}} |f - g| dx < \epsilon$. *Hint:* Apply Theorem 2.24 on the following filtrations

$$\mathcal{F}_n = \mathcal{F}\left(\left\{[k2^{-n}, (k+1)2^{-n}) : 0 \leq k < 2^n\right\}\right), \quad \mathcal{F}_\infty = \mathcal{B}([0,1)).$$

Exercise 2.11. Let $([0,1),\mathcal{B}([0,1)),\mathbb{P})$ be as in Exercise 2.10. Let X be a Lipschitz continuous function on $[0,1)$ and set, for $n \geq 1$,

$$X_n = \sum_{k=0}^{2^n-1} \frac{X((k+1)2^{-n}) - X(k2^{-n})}{2^{-n}} \mathbf{1}_{\{[k2^{-n},(k+1)2^{-n})\}}.$$

Prove that X_n is a martingale, X_n converges to X_∞ a.s. and in L^1 and

$$X(b) - X(a) = \mathbb{E}(X_\infty; [a,b)) \quad \forall 0 \leq a < b < 1.$$

This exercise says that Lipschitz continuous functions are absolutely continuous.

Exercise 2.12. Let $X_1, X_2, ...$ be a process taking values on $[0,\infty)$ and having 0 as an absorbing states, that is, $X_n = 0$ implies $X_m = 0$ for all $m > n$. Let $D = \{X_n = 0 \text{ for some } n > 0\}$ and assume that, for any $x > 0$, there is $\delta(x) > 0$ such that

$$\mathbb{P}(D|X_1,...,X_n) \overset{a.s.}{\geq} \delta(x) \quad \text{on } \{X_n \leq x\}, \quad \forall n \geq 1. \tag{2.7}$$

Prove that $\mathbb{P}(D \cup \{\lim_n X_n = \infty\}) = 1$. *Hint:* Use Lévy's zero-one law or Exercise 2.13.

Theorem 2.26. *Let $\mathcal{F}_n$ be a filtration and $\mathcal{F}_\infty = \sigma(\bigcup \mathcal{F}_n)$. Assume that X_n converges a.s. to X_∞ and $|X_n| \leq Y$ with $\mathbb{E}Y < \infty$. Then, $\mathbb{E}(X_n|\mathcal{F}_n)$ converges a.s. to $\mathbb{E}(X_\infty|\mathcal{F}_\infty)$.*

Proof. For $N \geq 1$, set $Z_N = \sup\{|X_n - X_m| : n, m \geq N\}$. Then, $Z_N \leq 2Y$. As $|X_n - X_\infty| \leq Z_N$ almost surely for $n \geq N$, one may use Theorem 2.24 to derive

$$\limsup_{n\to\infty} \mathbb{E}(|X_n - X_\infty||\mathcal{F}_n) \leq \limsup_{n\to\infty} \mathbb{E}(Z_N|\mathcal{F}_n) = \mathbb{E}(Z_N|\mathcal{F}_\infty) \quad \forall N > 0,$$

By the dominated convergence theorem, $\mathbb{E}(Z_N|\mathcal{F}_\infty) \to 0$ a.s. and this implies

$$|\mathbb{E}(X_n|\mathcal{F}_n) - \mathbb{E}(X_\infty|\mathcal{F}_n)| \leq \mathbb{E}(|X_n - X_\infty||\mathcal{F}_n) \overset{a.s.}{\to} 0 \quad \text{as } n \to \infty.$$

The desired limit is then given by the fact of $\mathbb{E}(X_\infty|\mathcal{F}_n) \to \mathbb{E}(X_\infty|\mathcal{F}_\infty)$ a.s.. □

Exercise 2.13. Let $X_1, X_2, ...$ be a process and $A, B \in \mathcal{B}(\mathbb{R})$. Suppose

$$\mathbb{P}(X_m \in B \text{ for some } m > n|X_1,...,X_n) \geq \delta > 0 \quad \text{a.s. on } \{X_n \in A\}.$$

Prove that $\{X_n \in A \text{ i.o.}\} \subset \{X_n \in B \text{ i.o.}\}$ a.s..

Exercise 2.14. Let $\mathcal{F}_n$ be a filtration and $\mathcal{F}_\infty = \sigma(\bigcup_n \mathcal{F}_n)$. Show that

1. If $\mathbb{E}|X| < \infty$ and $X_n \to X$ in L^1, then $\mathbb{E}(X_n|\mathcal{F}_n) \to \mathbb{E}(X|\mathcal{F}_\infty)$ in L^1;
2. If $(X_n)_{n=1}^\infty$ is uniformly integrable and converges to X almost surely, then $\mathbb{E}(X_n|\mathcal{F}) \to \mathbb{E}(X|\mathcal{F})$ in L^1 for any σ-field $\mathcal{F}$.

The following example shows that Theorem 2.26 can fail when $|X_n| \leq Y$ with $\mathbb{E}Y < \infty$ is replaced by the uniform integrability of $(X_n)_{n=0}^\infty$.

Example 2.6. Consider independent random variables $X_1, X_2, ..., Y_1, Y_2, ...$ defined by

$$\mathbb{P}(X_n = 1) = 1/n, \ \mathbb{P}(X_n = 0) = 1 - 1/n,$$

and

$$\mathbb{P}(Y_n = n) = 1/n, \ \mathbb{P}(Y_n = 0) = 1 - 1/n.$$

Set $Z_n = X_nY_n$, $\mathcal{F}_n = \mathcal{F}(X_1,...,X_n)$ and $\mathcal{F}_\infty = \mathcal{F}(X_1,X_2,...)$. Then Z_n is uniformly integrable since

$$\mathbb{E}(|Z_n|; |Z_n| > x) \leq \mathbb{E}(X_nY_n; X_n = 1, Y_n = n) = 1/n \quad \forall n, x > 0.$$

Note that $\mathbb{P}(Z_n > 0) = 1/n^2$. By the Borel-Cantelli lemma, we have $\mathbb{P}(Z_n > 0 \text{ i.o.}) = 0$ and this implies $Z_n \to 0$ almost surely. Note that $\mathbb{E}(Z_n|\mathcal{F}_n) = X_n\mathbb{E}Y_n = X_n$. As $\mathbb{P}(X_n = 1 \text{ i.o.}) = \mathbb{P}(X_n = 0 \text{ i.o.}) = 1$, $\mathbb{E}(Z_n|\mathcal{F}_n)$ diverges almost surely. In fact, X_n and, thus $\mathbb{E}(Z_n|\mathcal{F}_n)$, converges to 0 in L^1.

2.6. **Backward martingales.**

Definition 2.7. Let $(\mathcal{F}_n)_{n\le 0}$ be a non-decreasing sequence of σ-field, that is, $\mathcal{F}_n \subset \mathcal{F}_{n+1}$ for $n < 0$. A process indexed by non-positive integers $(X_n)_{n\le 0}$ is a *backward martingale* (*submartingale*) w.r.t. $(\mathcal{F}_n)_{n\le 0}$, if X_n is $\mathcal{F}_n$-measurable, $\mathbb{E}|X_n| < \infty$ and $\mathbb{E}(X_{n+1}|\mathcal{F}_n) = X_n$ $(\ge X_n)$ for $n < 0$.

Remark 2.10. Clearly, if X_n is a backward martingale w.r.t. $\mathcal{F}_n$, then $X_n = \mathbb{E}(X_{n+1}|\mathcal{F}_n) = \mathbb{E}(X_0|\mathcal{F}_n)$. By Proposition 2.19, $(X_n)_{n\le 0}$ is uniformly integrable.

Theorem 2.27. *Let X_n be a backward submartingale w.r.t. $\mathcal{F}_n$. Then, $\liminf_n X_n = \limsup_n X_n < \infty$ a.s. and, by setting $X_{-\infty} := \lim_n X_n$ and $\mathcal{F}_{-\infty} = \bigcap_{n\le 0} \mathcal{F}_n$, the following are equivalent.*

1. *X_n is uniformly integrable.*
2. *$X_n \to X_{-\infty}$ a.s. and in L^1.*
3. *$X_n \to X_{-\infty}$ in L^1.*
4. *$\mathbb{E}|X_{-\infty}| < \infty$ and $\mathbb{E}(X_n|\mathcal{F}_{-\infty}) \ge X_{-\infty}$ for all $n \le 0$.*
5. *$\lim_n \mathbb{E}X_n > -\infty$.*

In particular, if X_n is a backward martingale w.r.t. $\mathcal{F}_n$, then $X_{-\infty} = \mathbb{E}(X_n|\mathcal{F}_{-\infty})$ for all n.

Proof. For $n \ge 0$, let $U_n(a, b)$ be the upcrossing number of (a, b) by $X_{-n}, ..., X_0$. Clearly, $U_n(a, b)$ is also the downcrossing number of (a, b) by $X_0, X_{-1}, ..., X_{-n}$ and thus non-decreasing. Set $U = \lim_n U_n$. By Lemma 2.11, $(b - a)\mathbb{E}U_n \le \mathbb{E}(X_0 - a)^+$. By the monotone convergence theorem, passing n to ∞ leads to $\mathbb{P}(U(a, b) < \infty) = 1$ for any $a < b$ and, hence, $\liminf_n X_n = \limsup_n X_n$ almost surely. Set $X_{-\infty} = \lim_n X_n$. To show $X_{-\infty} < \infty$ a.s., it suffices to proved that $X^+_{-\infty} < \infty$ a.s. or further $\mathbb{E}X^+_{-\infty} < \infty$. As $X^+_{-\infty} = \lim_n X^+_n$ a.s., one may apply Fatou's lemma to conclude

$$\mathbb{E}X^+_{-\infty} \le \liminf_{n\to-\infty} \mathbb{E}X^+_n \le \mathbb{E}X^+_0 < \infty,$$

where the second inequality uses the fact that X^+_n is a backward submartingale.

For (1)⇒(2), assume that X_n is uniformly integrable. By Exercise 2.8, $\sup_n \mathbb{E}|X_n| < \infty$. Again, by Fatou's lemma, one has

$$\mathbb{E}|X_{-\infty}| \le \liminf_{n\to-\infty} \mathbb{E}|X_n| < \sup_{n\le 0} \mathbb{E}|X_n| < \infty.$$

This implies $X_n \to X_{-\infty}$ a.s. and then in probability. By Proposition 2.20, $X_n \to X_{-\infty}$ in L^1. (2)⇒(3) is clear. For (3)⇒(4), note that

$$\mathbb{E}(X_{n+1}|\mathcal{F}_{-\infty}) = \mathbb{E}(\mathbb{E}(X_{n+1}|\mathcal{F}_n)|\mathcal{F}_{-\infty}) \ge \mathbb{E}(X_n|\mathcal{F}_{-\infty}) \tag{2.8}$$

and

$$\mathbb{E}|\mathbb{E}(X_n|\mathcal{F}_{-\infty}) - X_{-\infty}| \le \mathbb{E}|X_n - X_{-\infty}| \to 0, \quad \text{as } n \to -\infty.$$

This implies $\mathbb{E}(X_n|\mathcal{F}_{-\infty}) \to X_{-\infty}$ in L^1 and then in probability. By selecting a subsequence k_n such that $\mathbb{E}(X_{k_n}|\mathcal{F}_{-\infty}) \to X_{-\infty}$ a.s, one may use (2.8) to conclude $\mathbb{E}(X_n|\mathcal{F}_{-\infty}) \ge X_{-\infty}$ a.s..

(4)⇒(5) is clear. For (5)⇒(1), set $L = -\inf_n \mathbb{E}X_n$. Note that, for $x > 0$,

$$\mathbb{P}(|X_n| > x) \le x^{-1}\mathbb{E}|X_n| = x^{-1}(2\mathbb{E}X^+_n - \mathbb{E}X_n) \le x^{-1}(2\mathbb{E}X^+_0 + L), \quad \forall n \le 0.$$

Let $\epsilon > 0$. As $\mathbb{E}X_n \to -L$, we may select $N < 0$ such that $\mathbb{E}X_N - \mathbb{E}X_n < \epsilon/2$ for $n \le N$. This implies that, for $n \le N$,

$$\begin{aligned}\mathbb{E}(|X_n|; X_n < -x) &= -\mathbb{E}(X_n; X_n < -x) = -\mathbb{E}X_n + \mathbb{E}(X_n; X_n \ge -x)\\ &\le -\mathbb{E}X_n + \mathbb{E}(X_N; X_n \ge -x) = \mathbb{E}X_N - \mathbb{E}X_n - \mathbb{E}(X_N; X_n < -x)\\ &\le \epsilon/2 + \mathbb{E}(|X_N|; |X_n| > x).\end{aligned}$$

Also, one has

$$\mathbb{E}(|X_n|; X_n > x) = \mathbb{E}(X_n^+; X_n > x) \le \mathbb{E}(X_0^+; X_n > x) \le \mathbb{E}(X_0^+; |X_n| > x).$$

As $\sup_n \mathbb{P}(|X_n| > x) \to 0$, we may choose x such that

$$\sup_n \mathbb{E}(|X_N|; |X_n| > x) < \epsilon/4, \quad \sup_n \mathbb{E}(X_0^+; |X_n| > x) < \epsilon/4.$$

Consequently, the above discussion leads to $\sup_{n\le N} \mathbb{E}(|X_n|; |X_n| > x) < \epsilon$, as desired.

When X_n is a backward martingale, Remark 2.10 implies that X_n is uniformly integrable. By the L^1-convergence, we have

$$\mathbb{E}(X_{-\infty}; A) = \lim_{n\to-\infty} \mathbb{E}(X_n; A) = \mathbb{E}(X_0; A), \quad \forall A \in \mathcal{F}_{-\infty}.$$

As $X_{-\infty}$ is $\mathcal{F}_{-\infty}$-measurable, this implies $X_{-\infty} = \mathbb{E}(X_0|\mathcal{F}_{-\infty})$. □

Exercise 2.15. Let $(\mathcal{F}_n)_{n\le 0}$ be a non-decreasing sequence of σ-fields and $\mathcal{F}_{-\infty} = \bigcap_{n\le 0} \mathcal{F}_n$. Prove that if X is a random variable satisfying $\mathbb{E}|X| < \infty$, then $\mathbb{E}(X|\mathcal{F}_n)$ converges a.s. and in L^1 to $\mathbb{E}(X|\mathcal{F}_{-\infty})$ as $n \to -\infty$.

Exercise 2.16. Let $\mathcal{F}_n \subset \mathcal{F}_{n+1}$ for $n < 0$ and set $\mathcal{F}_{-\infty} = \bigcap_{n\le 0} \mathcal{F}_n$. Suppose that X_n converges a.s to X and $|X_n| \le Y$ and $\mathbb{E}Y < \infty$. Prove that $\mathbb{E}(X_n|\mathcal{F}_n)$ converges a.s. to $\mathbb{E}(X|\mathcal{F}_{-\infty})$ as $n \to -\infty$.

Let $(S, \mathcal{G}, \mu)$ be a probability space and set

$$\Omega = S \times S \times \cdots, \quad \mathcal{F} = \mathcal{G} \otimes \mathcal{G} \otimes \cdots, \quad \mathbb{P} = \mu \times \mu \times \cdots. \tag{2.9}$$

For any permutation of $\mathbb{N}$ and $\omega \in \Omega$, define $\pi(\omega) := (\omega_{\pi(n)})_{n=1}^{\infty}$. An event $A \in \mathcal{F}$ is called permutable if $\pi(A) := \{\pi(\omega) : \omega \in A\} = A$ for any finite permutation π. A sub-σ-field $\mathcal{E}$ of $\mathcal{F}$ is called exchangeable if $\mathcal{E}$ is generated by permutable events. In fact, all events in $\mathcal{E}$ are permutable. By the $\pi - \lambda$ lemma, it is easy to check $\mathbb{P}(\pi(A)) = \mathbb{P}(A)$ for $A \in \mathcal{F}$ and π is any finite permutation.

Theorem 2.28 (The Hewitt-Savage zero-one law)**.** *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be the probability space defined in* (2.9) *and $\mathcal{E}$ is an exchangeable σ-field. Then, $\mathbb{P}(A) \in \{0, 1\}$ for all $A \in \mathcal{E}$.*

Proof. Let $A \in \mathcal{E}$ and set $\mathcal{G}_{n+1} = \mathcal{G}_n \times \mathcal{G}$. As $\bigcup_{n=1}^{\infty} \{B \times \Omega : B \in \mathcal{G}_n\}$ is a field generating $\mathcal{F}$, one may select $A_n = B_n \times \Omega$ with $B_n \in \mathcal{G}_n$ such that $\mathbb{P}(A_n \Delta A) \to 0$. This implies $\mathbb{P}(A_n) \to \mathbb{P}(A)$. Let π be the following permutation

$$(1, n + 1)(2, n + 2) \cdots (n, 2n)$$

and let $A_n' = \pi(A) = S^n \times B_n \times S^{\infty}$. Clearly, one has $\mathbb{P}(A_n \Delta A) = \mathbb{P}(\pi(A_n \Delta A)) = \mathbb{P}(A_n' \Delta A)$. Note that $B \setminus C \subset (B \setminus D) \cup (D \setminus C)$ for any set D, which implies $B\Delta C \subset (B\Delta D) \cup (D\Delta C)$. Immediately, this yields

$$\mathbb{P}(A_n \Delta A_n') \le \mathbb{P}(A_n \Delta A) + \mathbb{P}(A_n' \Delta A) = 2\mathbb{P}(A_n \Delta A) \to 0, \quad \text{as } n \to \infty.$$

Thus, $|\mathbb{P}(A_n) - \mathbb{P}(A_n \cap A_n')| \le \mathbb{P}(A_n \Delta A_n') \to 0$ and then $\mathbb{P}(A_n \cap A_n') \to \mathbb{P}(A)$. Since A_n and A_n' are independent, we have $\mathbb{P}(A_n \cap A_n') = \mathbb{P}(A_n)\mathbb{P}(A_n') \to \mathbb{P}(A)^2$. Consequently, $\mathbb{P}(A) \in \{0, 1\}$. □

Example 2.7 (The law of large numbers). Let $X_1, X_2, ...$ be i.i.d. random variables with $E|X_1| < \infty$. Set $S_n = X_1 + X_2 + \cdots + X_n$, $Y_{-n} = S_n/n$ and

$$\mathcal{F}_{-n} = \mathcal{F}(S_n, S_{n+1}, ...) = \mathcal{F}(S_n, X_{n+1}, X_{n+1}, ...).$$

Note that

$$\mathbb{E}(X_{n+1}|\mathcal{F}_{-n-1}) = \frac{S_{n+1}}{n+1} = Y_{-n-1} \quad \text{a.s..}$$

This implies

$$\mathbb{E}(Y_{-n}|\mathcal{F}_{-n-1}) = \frac{\mathbb{E}(S_{n+1}|\mathcal{F}_{-n-1}) - \mathbb{E}(X_{n+1}|\mathcal{F}_{-n-1})}{n} \overset{a.s.}{=} \frac{(n+1)Y_{-n-1} - Y_{-n-1}}{n} = Y_{-n-1},$$

and, hence, $(Y_n)_{n\le 0}$ is a backward martingale. By Theorem 2.27, S_n/n converges a.s. to $\mathbb{E}(X_1|\mathcal{F}_{-\infty})$. The strong law of large numbers is then given by the Hewitt-Savage 0-1 law.

2.7. **Doob's inequality and L^p-convergence.** In this subsection, we consider the L^p-convergence of submartingales for $p \in (1,\infty)$. Recall that, in the proof of Doob's inequality, we have in fact derived the following fact. For any nonnegative submartingale $X_0, ..., X_n$, one has

$$\mathbb{P}(\bar{X}_n > \epsilon) \le \epsilon^{-1}\mathbb{E}(X_n; \bar{X}_n > \epsilon) \le \epsilon^{-1}\mathbb{E}X_n, \quad \forall \epsilon > 0, \tag{2.10}$$

where $\bar{X}_n = \max\{X_0, ..., X_n\}$.

Example 2.8. Let $\xi_1, \xi_2, ...$ be independent random variables satisfying $\mathbb{E}\xi_n = 0$ and $\sigma_n^2 = \mathbb{E}\xi_n^2 < \infty$. Set $S_n = \xi_1 + \cdots + \xi_n$ and $X_n = S_n^2$. Then, X_n is a submartingale. Applying Doob's inequality with $\epsilon = x^2$, we obtain Kolmogorov's inequality, that is,

$$\mathbb{P}\left(\max_{1\le m\le n} |S_m| > x\right) \le \frac{\mathrm{Var}(S_n)}{x^2}.$$

Exercise 2.17. Let $\xi_1, \xi_2, ...$ be independent and satisfy $|\xi_n| \le K$ for all n. Set $S_n = \xi_1 + \cdots + \xi_n$. Prove that if $\mathbb{E}\xi_n = 0$, then

$$\mathbb{P}\left(\max_{1\le m\le n} |S_m| \le x\right) \le \frac{(x+K)^2}{\mathrm{Var}(S_n)}$$

Hint: Use the fact that $S_n^2 - s_n^2$ is a martingale, where $s_n^2 = \sigma_1^2 + \cdots + \sigma_n^2$ and $\sigma_n^2 = \mathbb{E}\xi_n^2$.

Exercise 2.18. Let X_n be a martingale satisfying $X_0 \equiv 0$ and $\mathbb{E}X_n^2 < \infty$. Show that

$$\mathbb{P}\left(\max_{1\le m\le n} X_m \ge x\right) \le \frac{\mathbb{E}X_n^2}{\mathbb{E}X_n^2 + x^2}, \quad \forall x > 0.$$

Hint: Use the fact that $(X_n + c)^2$ is a submartingale and optimize the inequality over c.

Theorem 2.29. *Let $(X_n)_{n=0}^\infty$ be a nonnegative submartingale satisfying $\sup_n \mathbb{E}X_n^p < \infty$ for some $p \in (1,\infty)$. Then, X_n converges a.s. and in L^p. In particular, if $(X_n)_{n=0}^\infty$ is a martingale satisfying $\sup_n \mathbb{E}|X_n|^p < \infty$ for some $p \in (1,\infty)$, then X_n converges in L^p.*

To prove this theorem, we need the following proposition.

Proposition 2.30 (L^p maximum inequality)**.** *Let $p \in (1,\infty)$ and $(X_m)_{m=0}^n$ be a nonnegative submartingale. Set $\bar{X}_n = \max_{1\le m\le n} X_m$. Then,*

$$\mathbb{E}\bar{X}_n^p \le (p/(p-1))^p \mathbb{E}X_n^p.$$

Proof of Theorem 2.29. Note that $(\mathbb{E}X_n)^p \le \mathbb{E}X_n^p$. By the martingale convergence theorem, X_n converges a.s. to some integrable random variable, say X_∞. Set $Y = \sup_n X_n$. By Proposition 2.30, one has

$$\mathbb{E}\left(\max_{1\le m\le n} X_m\right)^p \le \left(\frac{p}{p-1}\right)^p \mathbb{E}X_n^p \le \left(\frac{p}{p-1}\right)^p \sup_{n\ge 1} \mathbb{E}X_n^p.$$

Letting n tend to infinity implies $\mathbb{E}Y^p < \infty$. Observe that $|X_n - X_\infty| = \lim_m |X_n - X_m| \le 2|Y|$ with probability one. By the Lebesgue dominated convergence theorem, X_n converges to X_∞ in L^p. □

Proof of Proposition 2.30. Note that, for $K > 0$ and $x > 0$,

$$\{\bar{X}_n \wedge K > x\} = \begin{cases} \{\bar{X}_n > x\} & \text{if } K > x \\ \emptyset & \text{if } K \le x \end{cases}$$

By Proposition 2.10, one has

$$\mathbb{P}(\bar{X}_n \wedge K > x) \le x^{-1}\mathbb{E}(X_n; \bar{X}_n \wedge K > x), \quad \forall K, x > 0.$$

Consider the following computations.

$$\mathbb{E}(\bar{X}_n \wedge K)^p = \int_0^\infty \mathbb{P}((\bar{X}_n \wedge K)^p > x)dx = \int_0^\infty py^{p-1}\mathbb{P}(\bar{X}_n \wedge K > y)dy,$$

where the last equality applies the change of variables $x = y^p$. As a consequence, this implies

$$\begin{aligned}\mathbb{E}(\bar{X}_n \wedge K)^p &\le \int_0^\infty py^{p-2}\int_\Omega X_n \mathbf{1}_{\{\bar{X}_n \wedge K > y\}} d\mathbb{P} = \int_\Omega X_n \int_0^{\bar{X}_n \wedge K} py^{p-2} dy d\mathbb{P} \\ &= \frac{p}{p-1}\mathbb{E}(X_n(\bar{X}_n \wedge K)^{p-1}) \le \frac{p}{p-1}(\mathbb{E}(X_n)^p)^{1/p}(\mathbb{E}(\bar{X}_n \wedge K)^p)^{1/q}\end{aligned}$$

where q is the exponent conjugate of p and the last inequality is Hölder's inequality. Since $\mathbb{E}(\bar{X}_n \wedge K)^p < \infty$, we obtain

$$\mathbb{E}(\bar{X}_n \wedge K)^p \le \left(\frac{p}{p-1}\right)^p \mathbb{E}(X_n)^p.$$

Letting $K \to \infty$ gives the desired inequality. □

Exercise 2.19. Prove that if X_n is a submartingale, then

$$\mathbb{E}\bar{X}_n \le (1 - 1/e)^{-1}[1 + \mathbb{E}(X_n^+ \log^+(X_n^+))]$$

where $\bar{X}_n = \max_{1 \le m \le n} X_m^+$, $\log^+ t = \max\{\log t, 0\}$ and $0 \log 0 := 0$.
Hint: First, show that

$$\mathbb{E}(\bar{X}_n \wedge K) \le 1 + \int_\Omega X_n^+ \log(\bar{X}_n \wedge K)d\mathbb{P}$$

and then apply the following inequality

$$a \log b \le a \log a + \frac{b}{e} \le a \log^+ a + \frac{b}{e}.$$

2.8. **Other materials.**

2.8.1. *Bounded increments.*

Theorem 2.31. *Let $X_1, X_2, \ldots$ be a martingale satisfying $\mathbb{P}(|X_{n+1} - X_n| \le M) = 1$ for some constant M with $X_0 = 0$. Then, $\mathbb{P}(C \cup D) = 1$, where*

$$C = \{X_n \text{ converges to a finite limt}\}, \quad D = \left\{\limsup_{n\to\infty} X_n = \infty, \liminf_{n\to\infty} X_n = -\infty\right\}.$$

Proof. For $k \in \mathbb{N}$, $N_k = \inf\{n : X_n \le -k\}$, where $\inf \emptyset := \infty$. Clearly, N_k is a stopping time and, by Theorem 2.7, $(X_{n\wedge N_k})_{n=1}^\infty$ is a martingale. Observe that, on $\{N_k > n\}$, $X_{n\wedge N_k} = X_n > -k$ and, on $\{n \ge N_k\}$, $X_{n\wedge N_k} = X_{N_k} = X_{N_k - 1} + X_{N_k} - X_{N_k-1} > -k - M$. By Corollary 2.13, $X_{n\wedge N_k}$ converges a.s. to some integrable random variable, which implies that X_n converges a.s. on $\{N_k = \infty\}$ for all $k \ge 1$. Set $E = \bigcup_{k=1}^\infty \{N_k = \infty\}$. Clearly, $E = \{\inf_n X_n > -\infty\}$ and then $\mathbb{P}(C \cup \{\inf_n X_n = -\infty\}) = 1$. Similarly, one can show $\mathbb{P}(C \cup \{\sup_n X_n = \infty\}) = 1$. Combining both conclusions gives the desired identity. □

Exercise 2.20. Let X_n be a submartingale with $\mathbb{P}(\{\sup_n X_n < \infty\}) = 1$ and assume $\mathbb{E}(\sup_n \xi_n^+) < \infty$, where $\xi_n = X_n - X_{n-1}$. Show that X_n converges a.s. to an almost surely real-valued random variable.

Exercise 2.21. Let $\mathcal{F}_n$ be a filtration and X_n, Y_n be nonnegative random variables adapted to $\mathcal{F}_n$. Suppose $\mathbb{E}(X_{n+1}|\mathcal{F}_n) \le X_n + Y_n$ and $\mathbb{P}(\sum_n Y_n < \infty) = 1$. Prove that X_n converges a.s.. *Hint:* Consider $Z_n = X_n - (Y_1 + Y_2 + \cdots + Y_{n-1})$.

Corollary 2.32 (The second Borel-Cantelli lemma)**.** *Let $\mathcal{F}_n$ be a filtration with $\mathcal{F}_0 = \{\emptyset, \Omega\}$ and $A_n \in \mathcal{F}_n$. Then,*

$$\{A_n \text{ i.o.}\} \overset{a.s.}{=} \left\{ \sum_{n=1}^{\infty} \mathbb{P}(A_n|\mathcal{F}_{n-1}) = \infty \right\}.$$

Proof. Set $X_0 \equiv 0$ and

$$X_n = \sum_{m=1}^{n} \left(\mathbf{1}_{A_m} - \mathbb{P}(A_m|\mathcal{F}_{m-1}) \right).$$

Then, X_n is a martingale and $|X_n - X_{n-1}| \overset{a.s.}{\le} 1$ for all n. Let C and D be events in Theorem 2.31. Then, $\mathbb{P}(C \cup D) = 1$. Note that, on C,

$$\sum_{n=1}^{\infty} \mathbf{1}_{A_n} = \infty \quad \Leftrightarrow \quad \sum_{n=1}^{\infty} \mathbb{P}(A_n|\mathcal{F}_{n-1}) = \infty,$$

and, on D,

$$\sum_{n=1}^{\infty} \mathbf{1}_{A_n} = \infty, \quad \sum_{n=1}^{\infty} \mathbb{P}(A_n|\mathcal{F}_{n-1}) = \infty.$$

□

Exercise 2.22. Prove the following statements.

(1) Use the Borel-Cantelli lemma to show that, for $p_n \in [0, 1)$,

$$\prod_{n=1}^{\infty} (1 - p_n) = 0 \quad \Leftrightarrow \quad \sum_{n=1}^{\infty} p_n = \infty.$$

(2) If $\mathbb{P}(\cap_{m=1}^n A_m^c) > 0$ for all n and $\sum_{n=2}^{\infty} \mathbb{P}(A_n| \cap_{m=1}^{n-1} A_m^c) = \infty$, then $\mathbb{P}(A_n \text{ i.o.}) = 1$.

2.8.2. *Optional stopping theorem.*

Definition 2.8. Let $\mathcal{F}_n$ be a filtration on a measurable space $(\Omega, \mathcal{F})$ and N be a stopping time w.r.t. $\mathcal{F}_n$. Define $\mathcal{F}_N$ to be a collection of events $A \in \mathcal{F}$ satisfying $A \cap \{N = n\} \in \mathcal{F}_n$ for all n.

Remark 2.11. It is easy to show that $\mathcal{F}_N$ is a sub-σ-field of $\mathcal{F}$ and N is $\mathcal{F}_N$-measurable. Furthermore, $\mathcal{F}_N$ is also a collection of events $A \in \mathcal{F}$ satisfying $A \cap \{N \le n\} \in \mathcal{F}_n$ for all $n \ge 0$.

Proposition 2.33. *Let $\mathcal{F}_n$ be a filtration and M, N be stopping times w.r.t. $\mathcal{F}_n$.*

(1) *$M \vee N$ and $M \wedge N$ are stopping times.*
(2) *If $M \le N$, then $\mathcal{F}_M \subset \mathcal{F}_N$.*
(3) *If $M \le N$, then, $M\mathbf{1}_A + N\mathbf{1}_{A^c}$ is a stopping time w.r.t. $\mathcal{F}_n$ for any $A \in \mathcal{F}_M$.*
(4) *If X_n is adapted to $\mathcal{F}_n$, then $X_N \mathbf{1}_{\{N<\infty\}}$ is $\mathcal{F}_N$-measurable.*

Proof. Obvious from the definition. □

In the following, when $X_n(\omega)$ converges, we write $X_\infty(\omega)$ for its limit.

Theorem 2.34. *Let X_n be a uniformly integrable submartingale w.r.t. $\mathcal{F}_n$ and N be a stopping time for $\mathcal{F}_n$. Then, $X_{N\wedge n}$ is uniformly integrable and converges to X_N in L^1 and $\mathbb{E}X_0 \le \mathbb{E}X_N \le \mathbb{E}X_\infty < \infty$.*

Proof. Since X_n is uniform integrable, X_n converges a.s., which implies that X_N is a.s. defined. By Theorem 2.7, $X_{N\wedge n}$ is a submartingale. As X_n^+ is also a submartingale, one has

$$\mathbb{E}X_{N\wedge n}^+ = \mathbb{E}(X_N^+; N < n) + \mathbb{E}(X_n^+; N \ge n) \le \mathbb{E}X_n^+ \le \sup_n \mathbb{E}X_n^+ < \infty.$$

By the martingale convergence theorem, $X_{N\wedge n}$ converges to X_N a.s. and $\mathbb{E}|X_N| < \infty$. This leads to

$$\begin{aligned} \mathbb{E}(|X_{N\wedge n}|; |X_{N\wedge n}| > x) &= \mathbb{E}(|X_N|; |X_N| > x,\ N \le n) + \mathbb{E}(|X_n|; |X_n| > x,\ N > n) \\ &\le \mathbb{E}(|X_N|; |X_N| > x) + \sup_n \mathbb{E}(|X_n|; |X_n| > x) \end{aligned} \tag{2.11}$$

Letting $x \to \infty$ proves that $X_{N\wedge n}$ is uniformly integrable.

For the inequalities, since X_n is a submartingale, one has $\mathbb{E}X_0 \le \mathbb{E}X_{N\wedge n} \le \mathbb{E}X_n$. As $X_{N\wedge n}$ and X_n converge to X_N and X_∞ in L^1, we obtain

$$\mathbb{E}X_0 \le \lim_{n\to\infty} \mathbb{E}X_{N\wedge n} = \mathbb{E}X_N \le \lim_{n\to\infty} \mathbb{E}X_n = \mathbb{E}X_\infty.$$

□

Remark 2.12. It follows immediately from (2.11) that, for any process $(X_n)_{n=0}^\infty$, if N is a stopping time for X_n such that $\mathbb{E}(|X_N|; N < \infty) < \infty$ and $X_n \mathbf{1}_{\{N>n\}}$ is uniformly integrable, then $X_{N\wedge n}$ is uniformly integrable.

Theorem 2.35 (The optional stopping theorem)**.** *Let X_n be a submartingale with respect to $\mathcal{F}_n$ and M, N be stopping times for $\mathcal{F}_n$. Suppose $M \le N$ and X_n is uniformly integrable. Then, $X_M \le \mathbb{E}(X_N|\mathcal{F}_M)$. In particular, $\mathbb{E}X_M \le \mathbb{E}X_N$.*

Proof of the optional stopping theorem. By Theorem 2.34, $X_{N\wedge n}$ is a uniformly integrable submartingale w.r.t. $\mathcal{F}_n$ and $\mathbb{E}|X_N| < \infty$. Again, by applying Theorem 2.34 with $X_{N\wedge n}$ and M, we have $\mathbb{E}|X_M| < \infty$ and $\mathbb{E}X_M \le \mathbb{E}X_N$. Let $A \in \mathcal{F}_M$ and set $L = M\mathbf{1}_A + N\mathbf{1}_{A^c}$. Clearly, $L \le N$. By Proposition 2.33, L is a stopping time for $\mathcal{F}_n$. Replacing M with L in the above discussion then yields $\mathbb{E}X_L \le \mathbb{E}X_N$ or equivalently $\mathbb{E}(X_M; A) \le \mathbb{E}(X_N; A)$. □

Theorem 2.36. *Let X_n be a submartingale w.r.t. $\mathcal{F}_n$ and N be a stopping time for $\mathcal{F}_n$. Suppose $\mathbb{E}N < \infty$ and there exists a constant $c > 0$ such that*

$$\mathbb{E}(|X_{n+1} - X_n||\mathcal{F}_n) \le c \quad \text{on } \{N > n\}, \quad \forall n. \tag{2.12}$$

Then, $X_{N\wedge n}$ is uniformly integrable and $\mathbb{E}X_N \ge \mathbb{E}X_0$.

Proof. First, let's write

$$X_{N\wedge n} = \sum_{k=0}^{n} \mathbf{1}_{\{N=k\}} \left(X_0 + \sum_{i=0}^{k-1} (X_{i+1} - X_i) \right) + \sum_{k>n} \mathbf{1}_{\{N=k\}} \left(X_0 + \sum_{i=0}^{n-1} (X_{i+1} - X_i) \right).$$

By the triangle inequality, this implies

$$|X_{N\wedge n}| \le |X_0| + X, \quad X := \sum_{i=0}^{\infty} |X_{i+1} - X_i| \mathbf{1}_{\{N>i\}}.$$

Consequently, (2.12) implies $\mathbb{E}(|X_{i+1} - X_i|; N > i) \le c\mathbb{P}(N > i)$, which leads to

$$\mathbb{E}X = \sum_{i=0}^{\infty} \mathbb{E}(|X_{i+1} - X_i|; N > i) \le c \sum_{i=0}^{\infty} \mathbb{P}(N > i) = c\mathbb{E}N < \infty.$$

By Exercise 2.8(2), $(X_{N\wedge n})_{n=0}^{\infty}$ is uniformly integrable. □

Theorem 2.37. *If X_n is a nonnegative supermartingale w.r.t. $\mathcal{F}_n$ and N is a stopping time for $\mathcal{F}_n$, then $\mathbb{E}X_0 \geq \mathbb{E}X_N$.*

Proof. By Theorem 2.7, $X_{N\wedge n}$ is a nonnegative supermartingale w.r.t. $\mathcal{F}_n$ and thus $\mathbb{E}X_{N\wedge n} \leq \mathbb{E}X_0$. By Corollary 2.13, X_n and $X_{N\wedge n}$ converge a.s. and this implies X_N is almost surely defined and $\mathbb{E}|X_N| < \infty$. Applying the monotone convergence theorem and Fatou's lemma, we obtain

$$\mathbb{E}(X_N; N < \infty) = \lim_{n\to\infty} \mathbb{E}(X_N; N \leq n), \quad \mathbb{E}(X_N; N = \infty) \leq \liminf_{n\to\infty} \mathbb{E}(X_n; N > n).$$

Consequently, this leads to $\mathbb{E}X_N \leq \liminf_n \mathbb{E}X_{N\wedge n} \leq \mathbb{E}X_0$. □

Exercise 2.23. Let X_n be a nonnegative supermartingale. Show that $\mathbb{P}(\sup_n X_n > \epsilon) \leq \mathbb{E}X_0/\epsilon$.

In the following, $(X_n)_{n=1}^{\infty}$ refers to a sequence of i.i.d. random variables, $S_n = X_1 + \cdots + X_n$ and $\mathcal{F}_n = \mathcal{F}(X_1, X_2, ..., X_n)$.

Theorem 2.38 (Wald's identity)**.** *Let N be a stopping time for $\mathcal{F}_n$ and $\varphi(\lambda) = \mathbb{E}e^{\lambda X_1}$. Assume that $\mathbb{E}N < \infty$, $1 \leq \varphi(\lambda_0) < \infty$ for some real $\lambda_0 \neq 0$ and there is $c > 0$ such that $|S_n| < c$ on $\{N > n\}$ for all n. Then, $\mathbb{E}(e^{\lambda_0 S_N}/\varphi(\lambda_0)^N) = 1$.*

Proof. Set $Y_n = e^{\lambda_0 S_n}/\varphi(\lambda_0)^n$. Clearly, Y_n is a martingale w.r.t. to $\mathcal{F}_n$. Note that

$$\mathbb{E}(|Y_{n+1} - Y_n||\mathcal{F}_n) = \frac{e^{\lambda_0 S_n}}{\varphi(\lambda_0)^n}\mathbb{E}\left|\frac{e^{\lambda_0 X_{n+1}}}{\varphi(\lambda_0)} - 1\right| \leq 2e^{c|\lambda_0|} \quad \text{on } \{N > n\}.$$

By Theorem 2.36, $\mathbb{E}Y_N = \mathbb{E}Y_0 = 1$. □

Example 2.9. Consider the random walk on $\mathbb{Z}$ with $\mathbb{P}(X_1 = 1) = p$ and $\mathbb{P}(X_1 = -1) = 1 - p$. For $i \in \mathbb{Z}$, set $T_i = \min\{n \geq 0 | S_n = i\}$. Fix $a < 0 < b$ and let $N = T_a \wedge T_b$. Clearly, T_a, T_b and N are stopping times for $\mathcal{F}_n$.
Claim 1: $\mathbb{P}(N < \infty) = 1$.

Proof. Consider the process $S_{N\wedge n}$. Note that $S_{N\wedge n}$ is a uniformly bounded martingale if $p = 1/2$ (resp. submartingale if $p > 1/2$ and supermartingale if $p < 1/2$). The martingale convergence theorem implies that $S_{N\wedge n}$ converges a.s. and in L^1. Since $S_{N\wedge n}$ is integer-valued, $S_{N\wedge n}$ converges a.s. if and only if $S_{N\wedge n} \in \{a, b\}$ for n large enough. This implies $\mathbb{P}(N < \infty) = 1$. □

Claim 2: If $p = 1/2$, then $\mathbb{P}(S_N = a) = b/(b - a)$.

Proof. By Claim 1, $S_{N\wedge n}$ converges a.s. and in L^1 to S_N. As $S_{N\wedge n}$ is a martingale, we have $0 = \mathbb{E}S_0 = \mathbb{E}S_N = b\mathbb{P}(T_b < T_a) + a\mathbb{P}(T_a < T_b)$. The desired probability is given by $\mathbb{P}(T_a > T_b) = 1 - \mathbb{P}(T_a < T_b) = 1 - \mathbb{P}(S_N = a)$. □

For Claims 3 and 4, we assume $p \in (1/2, 1)$ and set $\varphi(x) = ((1 - p)/p)^x$.
Claim 3: $\varphi(S_{N\wedge n})$ is a martingale.

Proof. It is easy to check that $\varphi(S_n)$ is a martingale. By Theorem 2.7, $\varphi(S_{N\wedge n})$ is a martingale. □

Claim 4: $\mathbb{P}(T_a < T_b) = \frac{\varphi(b)-1}{\varphi(b)-\varphi(a)}$.

Proof. Since $\varphi(S_{N\wedge n})$ is a martingale taking values on a finite set, $\varphi(S_{N\wedge n})$ converges a.s. and in L^1 to $\varphi(S_N)$. Thus, we have

$$1 = \mathbb{E}\varphi(S_0) = \mathbb{E}\varphi(S_N) = \varphi(b)\mathbb{P}(T_b < T_a) + \varphi(a)\mathbb{P}(T_b > T_a).$$

A similar computation as in Claim 2 yields the desired probability. □

Claim 5: If $p = 1/2$, then $\mathbb{P}(T_a < \infty) = \mathbb{P}(T_b < \infty) = 1$. If $p \in (1/2, 1)$, then $\mathbb{P}(T_b < \infty) = 1$ and $\mathbb{P}(T_a < \infty) = (p/(1-p))^a$.

Proof. Note that $T_b \to \infty$ pointwise as $b \to \infty$. This implies $\{T_a < \infty\} = \bigcup_{b=1}^{\infty}\{T_a < T_b\}$. By Claim 2, when $p = 1/2$,

$$\mathbb{P}(T_a < \infty) = \lim_{b\to\infty} \mathbb{P}(T_a < T_b) = 1,$$

while the symmetry of the random walk implies $\mathbb{P}(T_b < \infty) = 1$. When $p \in (1/2, 1)$, Claim 4 implies

$$\mathbb{P}(T_a < \infty) = \lim_{b\to\infty} \mathbb{P}(T_a < T_b) = 1/\varphi(a), \quad \mathbb{P}(T_b < \infty) = \lim_{a\to-\infty} \mathbb{P}(T_b < T_a) = 1.$$

□

2.8.3. *Orthogonality of martingale increments.*

Proposition 2.39. *Let X_n be a martingale w.r.t. $\mathcal{F}_n$ satisfying $\mathbb{E}X_n^2 < \infty$. Then, $\mathbb{E}((X_n - X_m)Y) = 0$ for $m \le n$, where Y is any $\mathcal{F}_m$-measurable random variable satisfying $\mathbb{E}Y^2 < \infty$. In particular, $\mathbb{E}(X_n X_m) = \mathbb{E}X_m^2$ for $m \le n$.*

Proof. Since X_n and Y are in L^2, $\mathbb{E}|(X_n - X_m)Y| < \infty$. This implies

$$\mathbb{E}((X_n - X_m)Y) = \mathbb{E}[\mathbb{E}((X_n - X_m)Y|\mathcal{F}_m)] = \mathbb{E}[Y\mathbb{E}(X_n - X_m|\mathcal{F}_m)] = 0.$$

□

Corollary 2.40. *If X_n is a martingale w.r.t. $\mathcal{F}_n$ satisfying $\mathbb{E}X_n^2 < \infty$, then for $m \le n$,*

$$\mathbb{E}((X_n - X_m)^2|\mathcal{F}_m) = \mathbb{E}(X_n^2|\mathcal{F}_m) - X_m^2.$$

The above corollary can be regarded as a formula on the conditional variance of X_n given $\mathcal{F}_m$.

Exercise 2.24. Let X_n and Y_n be L^2 martingales w.r.t. $\mathcal{F}_n$. Show that

$$\mathbb{E}(X_nY_n) - \mathbb{E}(X_0Y_0) = \sum_{m=1}^{n} \mathbb{E}[(X_m - X_{m-1})(Y_m - Y_{m-1})].$$

Exercise 2.25. Let X_n be a martingale and $\xi_n = X_n - X_{n-1}$. Show that:

(1) If $\mathbb{E}X_0^2 < \infty$ and $\sum_{n=1}^{\infty} \mathbb{E}\xi_n^2 < \infty$, then X_n converges a.s. and in L^2.
(2) If $b_n \uparrow \infty$ and $\sum_{n=1}^{\infty} \mathbb{E}\xi_n^2/b_n^2 < \infty$, then $X_n/b_n \to 0$ a.s..

Example 2.10. Recall the branching process introduced before. Let $\{\xi_i^n : i, n \ge 0\}$ be a family of i.i.d. non-negative integer-valued random variables and $Z_{n+1} = \sum_{i=1}^{Z_n} \xi_i^{n+1}$. Let $\mu = \mathbb{E}\xi_i^n$ and $X_n = Z_n/\mu^n$. It was proved that X_n is a martingale and $\mathbb{E}X_n = 1$ for all $n \ge 1$. Assume in the following that $\mathrm{Var}(\xi_i^n) = \sigma^2 < \infty$. By Corollary 2.40, one has

$$\mathbb{E}X_n^2 = \mathbb{E}X_{n-1}^2 + \mathbb{E}(X_n - X_{n-1})^2.$$

Note that

$$\mathbb{E}[(X_n - X_{n-1})^2; Z_{n-1} = i] = \mu^{-2n}\mathbb{E}\left[\left(\sum_{j=1}^{i} \xi_j^n - i\mu\right)^2; Z_{n-1} = i\right]$$

$$= \mu^{-2n}\mathbb{E}\left[\left(\sum_{j=1}^{i}(\xi_j^n - \mu)\right)^2\right]\mathbb{P}(Z_{n-1} = i) = \mu^{-2n}\sigma^2 i\mathbb{P}(Z_{n-1} = i).$$

Summing up i gives $\mathbb{E}(X_n - X_{n-1})^2 = \mu^{-n-1}\sigma^2$. By induction, this implies $\mathbb{E}X_n^2 = 1 + \sigma^2\sum_{i=1}^n \mu^{-k-1}$. Assume that $\mu > 1$. Clearly, the above computation leads to $\sup_n \mathbb{E}X_n^2 < \infty$. By Theorem 2.29, X_n converges in L^2 to some random variable X. Applying the fact of $\mathbb{E}X_n = 1$, we obtain $\mathbb{E}X = 1$, which yields $\mathbb{P}(Z_n > 0 \text{ for all } n) \geq \mathbb{P}(X > 0) > 0$.

2.9. **Some other exercises.**

Exercise 2.26. If X_n and Y_n are submartingales w.r.t. $\mathcal{F}_n$, then $X_n \vee Y_n$ is a submartingale w.r.t. $\mathcal{F}_n$.

Exercise 2.27. Prove the following statements.

(1) If $y_n > -1$ satisfies $\sum_n |y_n| < \infty$, then $\prod_n(1 + y_n) < \infty$.
(2) Let X_n, Y_n be positive random variables adapted to $\mathcal{F}_n$. Assume that $\mathbb{E}X_n < \infty$, $\mathbb{E}(X_{n+1}|\mathcal{F}_n) \leq X_n(1 + Y_n)$ and $\sum_n Y_n$ converges a.s. Then, X_n converges a.s..

Exercise 2.28 (The switching principle)**.** Assume that X_n and Y_n be supermartingales w.r.t. $\mathcal{F}_n$ and N is a stopping time for $\mathcal{F}_n$ such that $X_N \geq Y_N$ on $\{N < \infty\}$. Set

$$Z_n = X_n\mathbf{1}_{\{N>n\}} + Y_n\mathbf{1}_{\{N\leq n\}}, \quad W_n = X_n\mathbf{1}_{\{N\geq n\}} + Y_n\mathbf{1}_{\{N<n\}}.$$

Then, Z_n and W_n are supermartingales w.r.t. $\mathcal{F}_n$.

Exercise 2.29 (Dubins' inequality)**.** Let X_n be a nonnegative supermartingale. Fix $0 < a < b$ and set $N_0 = -1$ and, for $k \geq 1$,

$$N_{2k-1} = \inf\{j > N_{2k-2} : X_j \leq a\}, \quad N_{2k} = \inf\{j > N_{2k-1} : X_j \geq b\}.$$

Let $U = \sup\{k : N_{2k} < \infty\}$. Prove that

$$\mathbb{P}(U \geq k) \leq \left(\frac{a}{b}\right)^k \mathbb{E}\left(\frac{X_0}{a} \wedge 1\right).$$

Hint: Let $Y_n = 1$ for $0 \leq n < N_1$ and for $j \geq 1$,

$$Y_n = \begin{cases} \left(\frac{b}{a}\right)^{j-1}\frac{X_n}{a} & \text{for } N_{2j-1} \leq n < N_{2j} \\ \left(\frac{b}{a}\right)^j & \text{for } N_{2j} \leq n < N_{2j+1} \end{cases}$$

Prove by induction (on j) that $Y_{n\wedge N_j}$ is a supermartingale and apply the fact of $\mathbb{E}Y_{n\wedge N_{2k}} \leq \mathbb{E}Y_0$ with $n \to \infty$.

Exercise 2.30. Let $Z_1, Z_2, ...$ be i.i.d. random variables with $\mathbb{E}|Z_1| < \infty$. Let θ be another random variable which has finite mean and is independent of Z_n. For $n \geq 1$, set $Y_n = \theta + Z_n$. Prove that $\mathbb{E}(\theta|Y_1, Y_2, ..., Y_n)$ converges to θ a.s. as $n \to \infty$. In statistics, if Z_1 is of standard normal distribution, then the distribution of θ is called the *prior distribution* and $\mathbb{P}(\theta \in \cdot|Y_1, ..., Y_n)$ is called the *posterior distribution.*

Exercise 2.31. Let Z_n be the Branching process with offspring distribution p_k. Prove that if $p_0 > 0$, then $\mathbb{P}(\lim_n Z_n \in \{0, \infty\}) = 1$. *Hint: Use Exercise 2.12.*

Exercise 2.32. Let $\mathcal{F}_n$ be a filtration and X_n be random variables adapted to $\mathcal{F}_n$ and taking values on $[0, 1]$. Let $\alpha > 0, \beta > 0$ be such that $\alpha + \beta = 1$ and assume that

$$\mathbb{P}(X_{n+1} = \alpha + \beta X_n | \mathcal{F}_n) = X_n, \quad \mathbb{P}(X_{n+1} = \beta X_n | \mathcal{F}_n) = 1 - X_n. \tag{2.13}$$

Then, X_n converges a.s. with $\mathbb{P}(\lim_n X_n = 1) = \mathbb{E}X_0$ and $\mathbb{P}(\lim_n X_n = 0) = 1 - \mathbb{E}X_0$.