Chapter III
Martingales and Stopping Times

The notion of **martingale** has proven to be among the most powerful ideas to emerge in probability in the past century. This chapter provides a foundation for this theory together with some illuminating examples and applications. For a prototypical illustration of the martingale property, let $Z_1, Z_2, \ldots$ be a sequence of independent integrable random variables and let $X_n = Z_1 + \cdots + Z_n,\ n \geq 1$. If $\mathbb{E}Z_j = 0,\ j \geq 1$, then one clearly has

$$\mathbb{E}(X_{n+1}|\mathcal{F}_n) = X_n, \quad n \geq 1,$$

where $\mathcal{F}_n := \sigma(X_1, \ldots, X_n)$.

Definition 3.1 *(First Definition of Martingale)* A sequence of integrable random variables $\{X_n : n \geq 1\}$ on a probability space $(\Omega, \mathcal{F}, P)$ is said to be a **martingale** if, writing $\mathcal{F}_n := \sigma(X_1, X_2, \ldots, X_n)$,

$$\mathbb{E}(X_{n+1}|\mathcal{F}_n) = X_n \text{ a.s. } (n \geq 1). \tag{3.1}$$

This definition extends to any (finite or infinite) family of integrable random variables $\{X_t : t \in T\}$, where T is a linearly ordered set: Let $\mathcal{F}_t = \sigma(X_s : s \leq t)$. Then $\{X_t : t \in T\}$ is a **martingale** if

$$\mathbb{E}(X_t|\mathcal{F}_s) = X_s \text{ a.s} \,\forall\, s < t\ (s, t \in T). \tag{3.2}$$

In the previous case of a sequence $\{X_n : n \geq 1\}$, as one can see by taking successive conditional expectations $\mathbb{E}(X_n|\mathcal{F}_m) = \mathbb{E}[\mathbb{E}(X_n|\mathcal{F}_{n+1})|\mathcal{F}_m] = \mathbb{E}(X_{n+1}|\mathcal{F}_m) = \cdots = \mathbb{E}(X_{m+1}|\mathcal{F}_m) = X_m$, (3.1) is equivalent to

$$\mathbb{E}(X_n|\mathcal{F}_m) = X_m \quad \text{a.s.} \quad \forall\, m < n. \tag{3.3}$$

R. Bhattacharya and E.C. Waymire, *A Basic Course in Probability Theory*,
Universitext, DOI 10.1007/978-3-319-47974-3_III

Thus, (3.1) is a special case of (3.2). Most commonly, $T = \mathbb{N}$ or $\mathbb{Z}^+$, or $T = [0, \infty)$. Note that if $\{X_t : t \in T\}$ is a martingale, one has the *constant expectations property:* $\mathbb{E}X_t = \mathbb{E}X_s \; \forall\, s, t \in T$.

Remark 3.1 Let $\{X_n : n \geq 1\}$ be a martingale sequence. Define its associated **martingale difference sequence** by $Z_1 := X_1$, $Z_{n+1} := X_{n+1} - X_n$ $(n \geq 1)$. Note that for $X_n \in L^2(\Omega, \mathcal{F}, P)$, $n \geq 1$, the martingale differences are uncorrelated. In fact, for $X_n \in L^1(\Omega, \mathcal{F}, P)$, $n \geq 1$, one has

$$\begin{aligned}\mathbb{E}Z_{n+1} f(X_1, X_2, \ldots, X_n) &= \mathbb{E}[\mathbb{E}(Z_{n+1} f(X_1, \ldots, X_n)|\mathcal{F}_n)] \\ &= \mathbb{E}[f(X_1, \ldots, X_n)\mathbb{E}(Z_{n+1}|\mathcal{F}_n)] = 0 \end{aligned} \tag{3.4}$$

for all bounded $\mathcal{F}_n$ measurable functions $f(X_1, \ldots, X_n)$. If $X_n \in L^2(\Omega, \mathcal{F}, P) \,\forall\, n \geq 1$, then (3.1) implies, and is equivalent to, the fact that $Z_{n+1} \equiv X_{n+1} - X_n$ is orthogonal to $L^2(\Omega, \mathcal{F}_n, P)$. It is interesting to compare this orthogonality to that of independence of Z_{n+1} and $\{Z_m : m \leq n\}$. Recall that Z_{n+1} is independent of $\{Z_m : 1 \leq m \leq n\}$ or, equivalently, of $\mathcal{F}_n = \sigma(X_1, \ldots, X_n)$ if and only if $\overline{g}(Z_{n+1})$ is orthogonal to $L^2(\Omega, \mathcal{F}_n, P)$ for all bounded measurable $\overline{g}$ such that $\mathbb{E}\overline{g}(Z_{n+1}) = 0$. Thus independence translates as $0 = \mathbb{E}\{[g(Z_{n+1}) - \mathbb{E}g(Z_{n+1})] \cdot f(X_1, \ldots, X_n)\} = \mathbb{E}\{g(Z_{n+1}) \cdot f(X_1, \ldots, X_n)\} - \mathbb{E}g(Z_{n+1}) \cdot \mathbb{E}f(X_1, \ldots, X_n)$, for all bounded measurable g on $\mathbb{R}$ and for all bounded measurable f on $\mathbb{R}^n$.

Example 1 *(Independent Increment Process)* Let $\{Z_n : n \geq 1\}$ be an independent sequence having *zero means,* and X_0 an integrable random variable independent of $\{Z_n : n \geq 1\}$. Then

$$X_0,\; X_n := X_0 + Z_1 + \cdots + Z_n \equiv X_{n-1} + Z_n \; (n \geq 1) \tag{3.5}$$

is a martingale sequence.

Definition 3.2 If one has inequality in place of (3.1), namely,

$$\mathbb{E}(X_{n+1}|\mathcal{F}_n) \geq X_n \text{ a.s.} \quad \forall n \geq 1, \tag{3.6}$$

then $\{X_n : n \geq 1\}$ is said to be a submartingale. More generally, if the index set T is as in (3.2), then $\{X_t : t \in T\}$ is a **submartingale** if

$$\mathbb{E}(X_t|\mathcal{F}_s) \geq X_s \;\forall\, s < t \; (s, t \in T). \tag{3.7}$$

If instead of $\geq$, one has $\leq$ in (3.7) (3.8), the process $\{X_n : n \geq 1\}$ ($\{X_t : t \in T\}$) is said to be a **supermartingale**.

In Example 1, if $\mathbb{E}Z_k \geq 0 \;\forall\, k$, then the sequence $\{X_n : n \geq 1\}$ of partial sums of independent random variables is a submartingale. If $\mathbb{E}Z_k \leq 0$ for all k, then $\{X_n : n \geq 1\}$ is a supermartingale. In Example 3, it follows from the triangle inequality for conditional expectations that the sequence $\{Y_n := |X_n| : n \geq 1\}$ is a submartingale. The following proposition provides an important generalization of this latter example.

Proposition 3.1 (a) If $\{X_n : n \geq 1\}$ is a martingale and $\varphi(X_n)$ is a convex and integrable function of X_n, then $\{\varphi(X_n) : n \geq 1\}$ is a submartingale. (b) If $\{X_n\}$ is a submartingale, and $\varphi(X_n)$ is a convex and nondecreasing integrable function of X_n, then $\{\varphi(X_n) : n \geq 1\}$ is a submartingale.

Proof The proof is obtained by an application of the conditional Jensen's inequality given in Theorem 2.10. In particular, for (a) one has

$$\mathbb{E}(\varphi(X_{n+1}|\mathcal{F}_n) \geq \varphi(\mathbb{E}(X_{n+1}|\mathcal{F}_n)) = \varphi(X_n). \tag{3.8}$$

Now take the conditional expectation of both sides with respect to $\mathcal{G}_n \equiv \sigma(\varphi(X_1), \dots, \varphi(X_n)) \subset \mathcal{F}_n$, to get the martingale property of $\{\varphi(X_n) : n \geq 1\}$. Similarly, for (b), for convex and nondecreasing φ one has in the case of a submartingale that

$$\mathbb{E}(\varphi(X_{n+1}|\mathcal{F}_n) \geq \varphi(\mathbb{E}(X_{n+1}|\mathcal{F}_n)) \geq \varphi(X_n), \tag{3.9}$$

and taking conditional expectation in (3.9), the desired submartingale property follows. ■

Proposition 3.1 immediately extends to martingales and submartingales indexed by an arbitrary linearly ordered set T.

Example 2 (a) If $\{X_t : t \in T\}$ is a martingale, $\mathbb{E}|X_t|^p < \infty$ $(t \in T)$ for some $p \geq 1$, then $\{|X_t|^p : t \in T\}$ is a submartingale. (b) If $\{X_t : t \in T\}$ is a submartingale, then for every real c, $\{Y_t := \max(X_t, c)\}$ is a submartingale. In particular, $\{X_t^+ := \max(X_t, 0)\}$ is a submartingale.

Remark 3.2 It may be noted that in (3.8), (3.9), the σ-field $\mathcal{F}_n$ is $\sigma(X_1, \dots, X_n)$, and not $\sigma(\varphi(X_1), \dots, \varphi(X_n))$, as seems to be required by the first definitions in (3.1), (3.6). It is, however, more convenient to give the definition of a **martingale (or a submartingale) with respect to a filtration** $\{\mathcal{F}_n\}$ for which (3.1) holds (or respectively, (3.6) holds) assuming at the outset that X_n is $\mathcal{F}_n$-measurable $(n \geq 1)$ (or, as one often says, $\{X_n\}$ is $\{F_n\}$-**adapted**). One refers to this sequence as an $\{\mathcal{F}_n\}$**-martingale** (respectively $\{\mathcal{F}_n\}$**-submartingale**). An important example of an $\mathcal{F}_n$ larger than $\sigma(X_1, \dots, X_n)$ is given by "adding independent information" via $\mathcal{F}_n = \sigma(X_1, \dots, X_n) \vee \mathcal{G}$, where $\mathcal{G}$ is a σ-field independent of $\sigma(X_1, X_2, \dots)$, and $\mathcal{G}_1 \vee \mathcal{G}_2$ denotes the smallest σ-field containing $\mathcal{G}_1 \cup \mathcal{G}_2$. We formalize this with the following definition; also see Exercise 13.

Definition 3.3 *(Second General Definition)* Let T be an arbitrary linearly ordered set and suppose $\{X_t : t \in T\}$ is a stochastic process with (integrable) values in $\mathbb{R}$ and defined on a probability space $(\Omega, \mathcal{F}, P)$. Let $\{\mathcal{F}_t : t \in T\}$ be a nondecreasing collection of sub-σ-fields of $\mathcal{F}$, referred to as a **filtration**, i.e., $\mathcal{F}_s \subset \mathcal{F}_t$ if $s \leq t$. Assume that for each $t \in T$, X_t is **adapted** to $\mathcal{F}_t$ in the sense that X_t is $\mathcal{F}_t$-measurable. We say that $\{X_t : t \in T\}$ is a **martingale**, respectively **submartingale, supermartingale**, with respect to the filtration $\{\mathcal{F}_t\}$ if $\mathbb{E}[X_t|\mathcal{F}_s] = X_s$, $\forall s, t \in T, s \leq t$, respectively $\geq X_s, \forall s, t \in T, s \leq t$, or $\leq X_s$ $\forall s, t \in T, s \leq t$.

Example 3 Let X be an integrable random variable on $(\Omega, \mathcal{F}, P)$ and let $\{\mathcal{F}_n : n \geq 1\}$ be a filtration of $\mathcal{F}$. One may check that the stochastic process defined by

$$X_n := \mathbb{E}(X|\mathcal{F}_n) \ (n \geq 1) \tag{3.10}$$

is an $\{\mathcal{F}_n\}$-martingale.

Note that for submartingales the expected values are nondecreasing, while those of supermartingales are nonincreasing. Of course, martingales continue to have constant expected values under this more general definition.

Theorem 3.2 *(Doob's Maximal Inequality)* Let $\{X_1, X_2, \ldots, X_n\}$ be an $\{\mathcal{F}_k : 1 \leq k \leq n\}$-martingale, or a nonnegative submartingale, and $\mathbb{E}|X_n|^p < \infty$ for some $p \geq 1$. Then, for all $\lambda > 0$, $M_n := \max\{|X_1|, \ldots, |X_n|\}$ satisfies

$$P(M_n \geq \lambda) \leq \frac{1}{\lambda^p} \int_{[M_n > \lambda]} |X_n|^p dP \leq \frac{1}{\lambda^p} \mathbb{E}|X_n|^p. \tag{3.11}$$

Proof Let $A_1 = [|X_1| \geq \lambda]$, $A_k = [|X_1| < \lambda, \ldots, |X_{k-1}| < \lambda, \ |X_k| \geq \lambda]$ $(2 \leq k \leq n)$. Then $A_k \in \mathcal{F}_k$ and $[A_k : 1 \leq k \leq n]$ is a (disjoint) partition of $[M_n \geq \lambda]$. Therefore,

$$\begin{aligned} P(M_n \geq \lambda) &= \sum_{k=1}^{n} P(A_k) \leq \sum_{k=1}^{n} \frac{1}{\lambda^p} \mathbb{E}(\mathbf{1}_{A_k} |X_k|^p) \leq \sum_{k=1}^{n} \frac{1}{\lambda^p} E(\mathbf{1}_{A_k} |X_n|^p) \\ &= \frac{1}{\lambda^p} \int_{[M_n \geq \lambda]} |X_n|^p dP \leq \frac{\mathbb{E}|X_n|^p}{\lambda^p}. \end{aligned}$$

■

Remark 3.3 The classical **Kolmogorov maximal inequality** for sums of i.i.d. mean zero, square-integrable random variables is a special case of **Doob's maximal inequality** obtained by taking $p = 2$ for the martingales of Example 1 having square-integrable increments.

Corollary 3.3 Let $\{X_1, X_2, \ldots, X_n\}$ be an $\{\mathcal{F}_k : 1 \leq k \leq n\}$-martingale such that $\mathbb{E}|X_n|^p < \infty$ for some $p \geq 2$, and $M_n = \max\{|X_1|, \ldots, |X_n|\}$. Then $\mathbb{E}M_n^p \leq p^q \mathbb{E}|X_n|^p$.

Proof A standard application of the Fubini–Tonelli theorem (see (1.10)) provides the second moment formula

$$\mathbb{E}M_n^p = p \int_0^{\infty} x^{p-1} P(M_n > x) dx.$$

Noting that $p - 1 \geq 1$ to first apply the Doob maximal inequality (3.11), one then makes another application of the Fubini–Tonelli theorem, and finally the Hölder inequality, noting $pq - q = p$ for the conjugacy $\frac{1}{p} + \frac{1}{q} = 1$, to obtain

$$\begin{aligned}\mathbb{E}M_n^p &\leq p\int_0^\infty \mathbb{E}\left(|X_n|^{p-1}\mathbf{1}_{[M_n\geq x]}\right)dx = p\mathbb{E}\left(|X_n|^{p-1}M_n\right)\\ &\leq p(\mathbb{E}|X_n|^{(p-1)q})^{\frac{1}{q}}(\mathbb{E}M_n^p)^{\frac{1}{p}}.\end{aligned}$$

Divide both sides by $(\mathbb{E}|M_n|^p)^{\frac{1}{p}}$ and use monotonicity of $x \to x^{\frac{1}{q}}, x \geq 0$, to complete the proof. ■

Doob also obtained a bound of this type for $p > 1$ with a smaller constant $q^p \leq p^q$ when $p \geq 2$, but it also requires a bit more clever estimation than in the above proof. Doob's statement and proof are as follows.

Theorem 3.4 *(Doob's Maximal Inequality for Moments)* Suppose that $\{X_1, X_2, \ldots, X_n\}$ is an $\{\mathcal{F}_k : 1 \leq k \leq n\}$-martingale, or a nonnegative submartingale, and let $M_n = \max\{|X_1|, \ldots, |X_n|\}$. Then

1. $\mathbb{E}M_n \leq \frac{e}{e-1}\left(1 + \mathbb{E}|X_n|\log^+|X_n|\right)$.
2. If $\mathbb{E}|X_n|^p < \infty$ for some $p > 1$, then $\mathbb{E}M_n^p \leq q^p\mathbb{E}|X_n|^p$, where q is the conjugate exponent defined by $\frac{1}{q} + \frac{1}{p} = 1$, i.e., $q = \frac{p}{p-1}$.

Proof For any nondecreasing function F_1 on $[0,\infty)$ with $F_1(0) = 0$, one may define a corresponding Lebesgue–Stieltjes measure $\mu_1(dy)$. Use the integration by parts Proposition 1.4, to get

$$\begin{aligned}\mathbb{E}F_1(M_n) &= \int_{[0,\infty)} P(M_n \geq y)F_1(dy)\\ &\leq \int_{[0,\infty)} \Big[\frac{1}{y}\int_{[M_n\geq y]} |X_n|dP\Big]F_1(dy)\\ &= \int_\Omega |X_n|\Big(\int_{[0,M_n]}\frac{1}{y}F_1(dy)\Big)dP, \qquad (3.12)\end{aligned}$$

where the integrability follows from Theorem 3.2 (with $p = 1$). For the first part, consider the function $F_1(y) = y\mathbf{1}_{[1,\infty)}(y)$. Then $y - 1 \leq F_1(y)$, and one gets

$$\begin{aligned}\mathbb{E}(M_n - 1) \leq \mathbb{E}F_1(M_n) &\leq \int_\Omega \big|X_n\big|\int_{[1,\max\{1,M_n\}]} y\frac{1}{y}dy\big)dP\\ &= \int_\Omega |X_n|\log(\max\{1, M_n\})dP\\ &= \int_{[M_n\geq 1]} |X_n|\log M_n dP. \qquad (3.13)\end{aligned}$$

Now use the inequality (proved in the remark below)

$$a\log b \leq a\log^+ a + \frac{b}{e}, \quad a, b \geq 0, \qquad (3.14)$$

to further arrive at

$$\mathbb{E}M_n \le 1 + \mathbb{E}|X_n|\log^+|X_n| + \frac{\mathbb{E}M_n}{e}. \tag{3.15}$$

This establishes the inequality for the case $p = 1$. For $p > 1$ take $F_1(y) = y^p$. Then

$$\begin{aligned}\mathbb{E}M_n^p &\le \mathbb{E}\big(|X_n|\int_{[0,M_n]} py^{p-2}dy\big)\\ &= \mathbb{E}\big(|X_n|\frac{p}{p-1}M_n^{p-1}\big)\\ &\le \frac{p}{p-1}(\mathbb{E}|X_n|^p)^{\frac{1}{p}}(\mathbb{E}M_n^{(p-1)q})^{\frac{1}{q}}\\ &= q(\mathbb{E}|X_n|^p)^{\frac{1}{p}}(\mathbb{E}M_n^p)^{\frac{1}{q}}. \end{aligned}\tag{3.16}$$

The bound for $p > 1$ now follows by dividing by $(\mathbb{E}M_n^p)^{\frac{1}{q}}$ and a little algebra. ■

Remark 3.4 To prove the inequality (3.14) it is sufficient to consider the case $1 < a < b$, since it obviously holds otherwise. In this case it may be expressed as

$$\log b \le \log a + \frac{b}{ae},$$

or

$$\log\frac{b}{a} \le \frac{b}{ae}.$$

But this follows from the fact that $f(x) = \frac{\log x}{x}$, $x > 1$, has a maximum value $\frac{1}{e}$.

Corollary 3.5 Let $\{X_t : t \in [0,T]\}$ be a right-continuous nonnegative $\{\mathcal{F}_t\}$-submartingale with $\mathbb{E}|X_T|^p < \infty$ for some $p \ge 1$. Then $M_T := \sup\{X_s : 0 \le s \le T\}$ is $\mathcal{F}_T$-measurable and, for all $\lambda > 0$,

$$P(M_T > \lambda) \le \frac{1}{\lambda^p}\int_{[M_T>\lambda]} X_T^p dP \le \frac{1}{\lambda^p}\mathbb{E}X_T^p. \tag{3.17}$$

Proof Consider the nonnegative submartingale $\{X_0, \dots, X_{Ti2^{-n}}, \dots, X_T\}$, for each $n = 1, 2, \dots,$ and let $M_n := \max\{X_{iT2^{-n}} : 0 \le i \le 2^n\}$. For $\lambda > 0$, $[M_n > \lambda] \uparrow [M_T > \lambda]$ as $n \uparrow \infty$. In particular, M_T is $\mathcal{F}_T$-measurable. By Theorem 3.2,

$$P(M_n > \lambda) \le \frac{1}{\lambda^p}\int_{[M_n>\lambda]} X_T^p dP \le \frac{1}{\lambda^p}\mathbb{E}X_T^p.$$

Letting $n \uparrow \infty$, (3.17) is obtained. ■

We finally come to the notions of **stopping times**, and **optional times** which provide a powerful probabilistic tool to analyze processes by viewing them at appropriate random times.

Definition 3.4 Let $\{\mathcal{F}_t : t \in T\}$ be a filtration on a probability space $(\Omega, \mathcal{F}, P)$, with T a linearly ordered index set to which one may adjoin, if necessary, a point '∞' as the largest point of $T \cup \{\infty\}$. A random variable $\tau : \Omega \to T \cup \{\infty\}$ is an $\{\mathcal{F}_t\}$-**stopping time** if $[\tau \leq t] \in \mathcal{F}_t \ \forall\, t \in T$. If $[\tau < t] \in \mathcal{F}_t$ for all $t \in T$ then τ is called an **optional time**.

Most commonly, T in this definition is $\mathbb{N}$ or $\mathbb{Z}^+$, or $[0, \infty)$, and τ is related to an $\{\mathcal{F}_t\}$-adapted process $\{X_t : t \in T\}$.

The intuitive idea of τ as a stopping-time strategy is that to "stop by time t, or not," according to τ, is determined by the knowledge of the past up to time t, and does not require "a peek into the future."

Example 4 Let $\{X_t : t \in T\}$ be an $\{\mathcal{F}_t\}$-adapted process with values in a measurable space $(S, \mathcal{S})$, with a linearly ordered index set. (a) If $T = \mathbb{N}$ or $\mathbb{Z}^+$, then for every $B \in \mathcal{S}$,

$$\tau_B := \inf\{t \geq 0 : X_t \in B\} \tag{3.18}$$

is an $\{\mathcal{F}_t\}$-stopping time. (b) If $T = \mathbb{R}_+ \equiv [0, \infty)$, S is a metric space $\mathcal{S} = \mathcal{B}(S)$, and B is *closed,* $t \mapsto X_t$ is continuous, then τ_B is an $\{\mathcal{F}_t\}$-stopping time. (c) If $T = \mathbb{R}_+$, S is a topological space, $t \mapsto X_t$ is right-continuous, and B is *open*, then $[\tau_B < t] \in \mathcal{F}_t$ for all $t \geq 0$, and hence τ_B is an optional time; see Definition 3.4.

We leave the proofs of (a)–(c) as Exercise 11. Note that (b), (c) imply that under the hypothesis of (b), τ_B is an optional time if B is open or closed; recall Definition 3.4.

Definition 3.5 Let $\{\mathcal{F}_t : t \in T\}$ be a filtration on $(\Omega, \mathcal{F})$. Suppose that τ is a $\{\mathcal{F}_t\}$-stopping time. The **pre-τ σ-field** $\mathcal{F}_\tau$ comprises all $A \in \mathcal{F}$ such that $A \cap [\tau \leq t] \in \mathcal{F}_t$ for all $t \in T$.

Heuristically, $\mathcal{F}_\tau$ comprises events determined by information available only up to time τ. For example, if T is discrete with elements $t_1 < t_2 < \cdots$, and $\mathcal{F}_t = \sigma(X_s : 0 \leq s \leq t) \subset \mathcal{F}, \forall t$, where $\{X_t : t \in T\}$ is a process with values in some measurable space $(S, \mathcal{S})$, then $\mathcal{F}_\tau = \sigma(X_{\tau \wedge t} : t \geq 0)$; (Exercise 9). The stochastic process $\{X_{\tau \wedge t} : t \geq 0\}$ is referred to as the **stopped process**. The notation $\wedge$ is defined by $a \wedge b = \min\{a, b\}$. Similarly $\vee$ is defined by $a \vee b = \max\{a, b\}$.

If τ_1, τ_2 are two $\{\mathcal{F}_t\}$-stopping times and $\tau_1 \leq \tau_2$, then it is simple to check that

$$\mathcal{F}_{\tau_1} \subset \mathcal{F}_{\tau_2}. \tag{3.19}$$

Suppose $\{X_t\}$ is an $\{\mathcal{F}_t\}$-adapted process with values in a measurable space $(S, \mathcal{S})$, and τ is an $\{\mathcal{F}_t\}$-stopping time. For many purposes the following notion of adapted joint measurability of $(t, \omega) \mapsto X_t(\omega)$ is important.

Definition 3.6 Let $T = [0, \infty)$ or $T = [0, t_0]$ for some $t_0 < \infty$. A stochastic process $\{X_t : t \in T\}$ with values in a measurable space $(S, \mathcal{S})$ is **progressively measurable** with respect to $\{\mathcal{F}_t\}$ if for each $t \in T$, the map $(s, \omega) \mapsto X_s(\omega)$, from $[0, t] \times \Omega$ to

S is measurable with respect to the σ-fields $\mathcal{B}[0,t] \otimes \mathcal{F}_t$ (on $[0,t] \times \Omega$) and $\mathcal{S}$ (on S). Here $\mathcal{B}[0,t]$ is the Borel σ-field on $[0,t]$, and $\mathcal{B}[0,t] \otimes \mathcal{F}_t$ is the usual product σ-field.

Proposition 3.6 (a) Suppose $\{X_t : t \in T\}$ is progressively measurable, and τ is a stopping time. Then X_τ is $\mathcal{F}_\tau$-measurable, i.e., $[X_\tau \in B] \cap [\tau \le t] \in \mathcal{F}_t$ for each $B \in \mathcal{S}$ and each $t \in T$. (b) Suppose $\mathcal{S}$ is a metric space and $\mathcal{S}$ its Borel σ-field. If $\{X_t : t \in T\}$ is right-continuous, then it is progressively measurable.

Proof (a) Fix $t \in T$. On the set $\Omega_t := [\tau \le t]$, X_τ is the composition of the maps (i) $f(\omega) := (\tau(\omega), \omega)$, from $\omega \in \Omega_t$ into $[0,t] \times \Omega_t$, and (ii) $g(s,\omega) = X_s(\omega)$ on $[0,t] \times \Omega_t$ into $\mathcal{S}$. Now f is $\tilde{\mathcal{F}}_t$-measurable on Ω_t, where $\tilde{\mathcal{F}}_t := \{A \cap \Omega_t : A \in \mathcal{F}_t\}$ is the **trace σ-field** on Ω_t, and $\mathcal{B}[0,t] \otimes \tilde{\mathcal{F}}_t$ is the σ-field on $[0,t] \times \Omega_t$. Next the map $g(s,\omega) = X_s(\omega)$ on $[0,t] \times \Omega$ into S is $\mathcal{B}[0,t] \otimes \mathcal{F}_t$-measurable. Therefore, the restriction of this map to the measurable subset $[0,t] \times \Omega_t$ is measurable on the trace σ-field $\{A \cap ([0,t] \times \Omega_t) : A \in \mathcal{B}[0,t] \otimes \mathcal{F}_t\}$. Therefore, the composition X_τ is $\tilde{\mathcal{F}}_t$-measurable on Ω_t, i.e., $[X_\tau \in B] \cap [\tau \le t] \in \tilde{\mathcal{F}}_t \subset \mathcal{F}_t$ and hence $[X_\tau \in B] \in \mathcal{F}_\tau$, for $B \in \mathcal{S}$.

(b) Fix $t \in T$. Define, for each positive integer n, the stochastic process $\{X_s^{(n)} : 0 \le s \le t\}$ by

$$X_s^{(n)} := X_{j2^{-n}t} \text{ for } (j-1)2^{-n}t \le s < j2^{-n}t \quad (1 \le j \le 2^n),\ X_t^{(n)} = X_t.$$

Since $\{(s,\omega) \in [0,t] \times \Omega : X_s^{(n)}(\omega) \in B\} = \cup_{j=1}^{2^n}([j-1)2^{-n}t, j2^{-n}t) \times \{\omega : X_{j2^{-n}t}(\omega) \in B\}) \cup (\{t\} \times \{\omega : X_t(\omega) \in B\}) \in \mathcal{B}[0,t] \otimes \mathcal{F}_t$, it follows that $\{X_t^{(n)}\}$ is progressively measurable. Now $X_t^{(n)}(\omega) \to X_t(\omega)$ for all (t,ω) as $n \to \infty$, in view of the right-continuity of $t \mapsto X_t(\omega)$. Hence $\{X_t : t \in T\}$ is progressively measurable. ■

Remark 3.5 It is often important to relax the assumption of 'right-continuity' of $\{X_t : t \in T\}$ to "a.s. right-continuity." To ensure progressive measurability in this case, it is convenient to take $\mathcal{F}, \mathcal{F}_t$ to be ***P*-complete,** i.e., if $P(A) = 0$ and $B \subset A$ then $B \in \mathcal{F}$ and $B \in \mathcal{F}_t\ \forall\, t$. Then modify X_t to equal $X_0\ \forall\, t$ on the P-null set $N = \{\omega : t \to X_t(\omega) \text{ is not right-continuous}\}$. This modified $\{X_t : t \in T\}$, together with $\{\mathcal{F}_t : t \in T\}$ satisfy the hypothesis of part (b) of Proposition 3.6.

The following proposition is distinguished as a characterization of the uniformly integrable martingales as conditional expectations.

Proposition 3.7 (a) Let Y be integrable and $\mathcal{F}_n (n = 1, 2, \ldots)$ a filtration. Then the martingale $Y_n = \mathbb{E}(Y|\mathcal{F}_n)$ is uniformly integrable. (b) Suppose Y_n is a $\mathcal{F}_n$-martingale $(n = 1, 2, \ldots)$ such that $Y_n \to Y$ in L^1. Then $Y_n = \mathbb{E}(Y|\mathcal{F}_n)$, $n \ge 1$.

Proof (a) Note that $P(|Y_n| > \lambda) \le \frac{\mathbb{E}|Y_n|}{\lambda} \le \frac{\mathbb{E}|Y|}{\lambda}$ for $\lambda > 0$. Hence, given $\varepsilon > 0$, there exist $\lambda > 0$ such that $P(|Y_n| > \lambda) < \varepsilon$ for all n. Therefore, $\mathbb{E}(\mathbf{1}_{[|Y_n|>\lambda]}|Y_n|) \le \mathbb{E}(\mathbf{1}_{[|Y_n|>\lambda]}\mathbb{E}(|Y||\mathcal{F}_n)) = \mathbb{E}(\mathbf{1}_{[|Y_n|>\lambda]}|Y|) \to 0$ as $\lambda \to \infty$ uniformly in n (see Exercise 14). (b) Let $A \in \mathcal{F}_m$. Then $\mathbb{E}(\mathbf{1}_A Y_m) = \mathbb{E}(\mathbf{1}_A Y_n)$ for all $n \ge m$. Taking the limit as $n \to \infty$, we have $\mathbb{E}(1_A Y_m) = \mathbb{E}(1_A Y)$ for all $A \in \mathcal{F}_m$. That is, $Y_m = \mathbb{E}(Y|\mathcal{F}_m)$. ■

Remark 3.6 One can show that a $\mathcal{F}_n$-martingale $\{Z_n : n \geq 1\}$ has the representation $Z_n = \mathbb{E}(Z|\mathcal{F}_n)$ iff it is uniformly integrable, and then $Z_n \to Z$ a.s. and in L^1. Indeed, a uniformly integrable martingale converges a.s. and in L^1 by the martingale convergence theorem; see Theorems 1.10 and 3.12.

One of the important implications of the martingale property is that of *constant expected values*. Let us consider a substantially stronger property. Consider a discrete parameter martingale sequence $X_0, X_1, \ldots$, and stopping time τ with respect to some filtration $\mathcal{F}_n$, $n \geq 0$. Let m be an integer and suppose that $\tau \leq m$. For $G \in \mathcal{F}_\tau$, write $g = \mathbf{1}_G$. One has that $G \cap [\tau = k] = (G \cap [\tau \leq k]) \backslash (G \cap [\tau \leq k-1]) \in \mathcal{F}_k$ from the definition of $\mathcal{F}_\tau$. It follows from the martingale property $\mathbb{E}(X_m|\mathcal{F}_k) = X_m$, one has

$$\begin{aligned}\mathbb{E}(gX_\tau) &= \sum_{k=0}^{m} \mathbb{E}(g\mathbf{1}_{[\tau=k]}X_k) \\ &= \sum_{k=0}^{m} \mathbb{E}(g\mathbf{1}_{[\tau=k]}\mathbb{E}(X_m|\mathcal{F}_k)) \\ &= \sum_{k=0}^{m} \mathbb{E}(g\mathbf{1}_{[\tau=k]}X_m) = \mathbb{E}(gX_m).\end{aligned} \tag{3.20}$$

Thus the constancy of expectations $\mathbb{E}X_n = \mathbb{E}X_0$ property of martingales extends to certain stopping times τ in place of n. However, as illustrated in Example 5, below, this requires some further conditions on τ than merely being a stopping time to extend to unbounded cases. The following theorem provides precisely such conditions.

Theorem 3.8 *(Optional Stopping)* Let $\{X_t : t \in T\}$ be a right-continuous $\{\mathcal{F}_t\}$-martingale, where $T = \mathbb{N}$ or $T = [0, \infty)$. (a) If $\tau_1 \leq \tau_2$ are bounded stopping times, then

$$\mathbb{E}(X_{\tau_2}|\mathcal{F}_{\tau_1}) = X_{\tau_1}. \tag{3.21}$$

(b) *(Optional Sampling).* If τ is a stopping time (not necessarily finite), then $\{X_{\tau\wedge t} : t \in T\}$ is an $\{\mathcal{F}_{\tau\wedge t}\}_{t\in T}$-martingale.

(c) Suppose $\tau_1 \leq \tau_2$ are stopping times such that (i) $P(\tau_2 < \infty) = 1$, and (ii) $X_{\tau_2\wedge t} (t \in T)$ is uniformly integrable. Then

$$\mathbb{E}(X_{\tau_2}|\mathcal{F}_{\tau_1}) = X_{\tau_1}. \tag{3.22}$$

In particular,

$$\mathbb{E}(X_{\tau_2}) = \mathbb{E}(X_{\tau_1}) = \mathbb{E}(X_0).$$

Proof First we consider the case $T = \mathbb{N}$. In the case that τ_2 is bounded by some positive integer m, it follows from the above calculation (3.20) that

$$\mathbb{E}(X_m|\mathcal{F}_{\tau_2}) = X_{\tau_2}.$$

Thus if $\tau_1 \le \tau_2 \le m$, then one also has

$$\mathbb{E}(X_m|\mathcal{F}_{\tau_1}) = X_{\tau_1}.$$

In other words, the two term sequence X_{τ_1}, X_{τ_2} is a martingale with respect to $\mathcal{F}_{\tau_1}, \mathcal{F}_{\tau_2}$. Hence it follows from the "smoothing property"of conditional expectation that $X_{\tau_1} = \mathbb{E}(X_m|\mathcal{F}_{\tau_1}) = \mathbb{E}(\mathbb{E}(X_m|\mathcal{F}_{\tau_2})|\mathcal{F}_{\tau_1}) = \mathbb{E}(X_{\tau_2}|\mathcal{F}_{\tau_1})$. This proves (a) in the discrete parameter case. Part (b) follows directly from (a) since $\tau_1 := \tau \wedge n \le \tau_2 := n < \infty$ for any n, and both of these τ_1, τ_2, so-defined, are stopping times. For (c), let $G \in \mathcal{F}_{\tau_1}$. Then, $G \cap [\tau_1 \le n] \in \mathcal{F}_n$. Also $G \cap [\tau_1 \le n] \in \mathcal{F}_{\tau_1 \wedge n}$. To see this, for $m \ge n$, $G \cap [\tau_1 \le n]) \cap [\tau_1 \wedge n \le m] = G \cap [\tau_1 \le n] \in \mathcal{F}_n \subset \mathcal{F}_m$, and if $m < n$ then $(G \cap [\tau_1 \le n]) \cap [\tau_1 \wedge n \le m] = G \cap [\tau_1 \le m] \in \mathcal{F}_m$. So in either case $G \cap [\tau_1 \le n] \in \mathcal{F}_{\tau_1 \wedge n}$. Also $\tau_1 \wedge n \le \tau_2 \wedge n$. By part (a), $\mathbb{E}(X_{\tau_2 \wedge n}|\mathcal{F}_{\tau_1 \wedge n}) = X_{\tau_1 \wedge n}$. Thus

$$\mathbb{E}(g\mathbf{1}_{[\tau_1 \le n]}X_{\tau_2 \wedge n}) = \mathbb{E}(g\mathbf{1}_{\tau_1 \le n]}X_{\tau_1 \wedge n}), \quad g = \mathbf{1}_G. \tag{3.23}$$

So, by the uniform integrability of $X_{\tau_2 \wedge n}$, and the fact that $X_{\tau_2 \wedge n} \to X_{\tau_2}$ a.s. as $n \to \infty$, one has $X_{\tau_2 \wedge n} \to X_{\tau_2}$ in L^1. Now observe that the uniform integrability of $X_{\tau_2 \wedge n}$ implies that of $X_{\tau_1 \wedge n}, n \ge 1$, as follows: Since $X_{\tau_2 \wedge n}, n \ge 1$, is uniformly integrable, it converges in L^1 (and a.s.) to X_{τ_2}, and $X_{\tau_2 \wedge n} = \mathbb{E}(X_{\tau_2}|\mathcal{F}_{\tau_2 \wedge n})$. Therefore $X_{\tau_1 \wedge n} = \mathbb{E}(X_{\tau_2 \wedge n}|\mathcal{F}_{\tau_1 \wedge n}) = \mathbb{E}[\mathbb{E}(X_{\tau_2}|\mathcal{F}_{\tau_2 \wedge n})|\mathcal{F}_{\tau_1 \wedge n}] = \mathbb{E}(X_{\tau_2}|\mathcal{F}_{\tau_1 \wedge n})$. Uniform integrability of $X_{\tau_1 \wedge n}, n \ge 1$, now follows from Proposition 3.7(a). Putting this uniform integrability together, it follows that the left side of (3.23) converges to $\mathbb{E}(gX_{\tau_2})$ and the right side to $\mathbb{E}(gX_{\tau_1})$. Since this is for any $g = \mathbf{1}_G$, $G \in \mathcal{F}_{\tau_1}$, the proof of (b) follows.

Next we consider the case $T = [0, \infty)$. Let $\tau_1 \le \tau_2 \le t_0$ a.s. The idea for the proof is, as above, to check that $\mathbb{E}[X_{t_0}|\mathcal{F}_{\tau_i}] = X_{\tau_i}$, for each of the stopping times $(i = 1, 2)$ simply by virtue of their being bounded. Once this is established, the result (a) follows by smoothing of conditional expectation, since $\mathcal{F}_{\tau_1} \subset \mathcal{F}_{\tau_2}$. That is, it will then follow that

$$\mathbb{E}[X_{\tau_2}|\mathcal{F}_{\tau_1}] = \mathbb{E}[\mathbb{E}(X_{t_0}|\mathcal{F}_{\tau_2})|\mathcal{F}_{\tau_1}] = \mathbb{E}[X_{t_0}|\mathcal{F}_{\tau_1}] = X_{\tau_1}.$$

So let τ denote either of τ_i, $i = 1, 2$, and consider $\mathbb{E}[X_{t_0}|\mathcal{F}_\tau]$. For each $n \ge 1$ consider the nth dyadic subdivision of $[0, t_0]$ and define $\tau^{(n)} = (k+1)2^{-n}t_0$ if $\tau \in [k2^{-n}t_0,\ (k+1)2^{-n}t_0)(k = 0, 1, \dots, 2^n - 1)$, and $\tau^{(n)} = t_0$ if $\tau = t_0$. Then $\tau^{(n)}$ is a stopping time and $\mathcal{F}_\tau \subset \mathcal{F}_{\tau^{(n)}}$ (since $\tau \le \tau^{(n)}$). For $G \in \mathcal{F}_\tau$, exploiting the martingale property $\mathbb{E}[X_{t_0}|\mathcal{F}_{(k+1)2^{-n}t_0}] = X_{t_{(k+1)2^{-n}t_0}}$, one has

$$\begin{aligned}\mathbb{E}(\mathbf{1}_G X_{t_0}) &= \sum_{k=0}^{2^n-1} \mathbb{E}(\mathbf{1}_{G\cap[\tau^{(n)}=(k+1)2^{-n}t_0]} X_{t_0}) \\ &= \sum_{k=0}^{2^n-1} \mathbb{E}(\mathbf{1}_{G\cap[\tau^{(n)}=(k+1)2^{-n}t_0]} X_{(k+1)2^{-n}t_0}) \\ &= \sum_{k=0}^{2^n-1} \mathbb{E}(\mathbf{1}_{G\cap[\tau^{(n)}=(k+1)2^{-n}t_0]} X_{\tau^{(n)}}) \\ &= \mathbb{E}(\mathbf{1}_G X_{\tau^{(n)}}) \to \mathbb{E}(\mathbf{1}_G X_\tau). \end{aligned} \tag{3.24}$$

The last convergence is due to the L^1-convergence criterion of Theorem 1.10 in view of the following checks: (1) X_t is right-continuous (and $\tau^{(n)} \downarrow \tau$), so that $X_{\tau^{(n)}} \to X_\tau$ a.s., and (2) $X_{\tau^{(n)}}$ is uniformly integrable, since by the submartingale property of $\{|X_t| : t \in T\}$,

$$\begin{aligned}\mathbb{E}(\mathbf{1}_{[|X_{\tau^{(n)}}|>\lambda]}|X_{\tau^{(n)}}|) &= \sum_{k=0}^{2^n-1} \mathbb{E}(\mathbf{1}_{[\tau^{(n)}=(k+1)2^{-n}t_0]\cap[|X_{\tau^{(n)}}|>\lambda]}|X_{(k+1)2^{-n}t_0}|) \\ &\le \sum_{k=0}^{2^n-1} \mathbb{E}(\mathbf{1}_{[\tau^{(n)}=(k+1)2^{-n}t_0]\cap[|X_{\tau^{(n)}}|>\lambda]}|X_{t_0}|) \\ &= \mathbb{E}(\mathbf{1}_{[|X_{\tau^{(n)}}|>\lambda]}|X_{t_0}|) \to \mathbb{E}(\mathbf{1}_{[|X_\tau|>\lambda]}|X_{t_0}|).\end{aligned}$$

Since the left side of (3.24) does not depend on n, it follows that

$$\mathbb{E}(\mathbf{1}_G X_{t_0}) = \mathbb{E}(\mathbf{1}_G X_\tau) \quad \forall\, G \in \mathcal{F}_\tau,$$

i.e., $\mathbb{E}(X_{t_0}|\mathcal{F}_\tau) = X_\tau$ applies to both $\tau = \tau_1$ and $\tau = \tau_2$. The result (a) therefore follows by the smoothing property of conditional expectations noted at the start of the proof.

As in the discrete parameter case, (b) follows immediately from (a). For if $s < t$ are given, then $\tau \wedge s$ and $\tau \wedge t$ are both bounded by t, and $\tau \wedge s \le \tau \wedge t$.

(c) Since $\tau < \infty$ a.s., $\tau \wedge t$ equals τ for sufficiently large t (depending on ω), outside a P-null set. Therefore, $X_{\tau\wedge t} \to X_\tau$ a.s. as $t \to \infty$. By assumption (ii), $X_{\tau\wedge t}$ $(t \ge 0)$ is uniformly integrable. Hence $X_{\tau\wedge t} \to X_\tau$ in L^1. In particular, $\mathbb{E}(X_{\tau\wedge t}) \to \mathbb{E}(X_\tau)$ as $t \to \infty$. But $\mathbb{E}X_{\tau\wedge t} = \mathbb{E}X_0\ \forall\, t$, by (b). ■

Remark 3.7 If $\{X_t : t \in T\}$ in Theorem 3.8 is taken to be a submartingale, then instead of the equality sign "=" in (3.21), (3.22), one gets "≤."

Remark 3.8 The *stopping time approximation technique* used in the proof of Theorem 3.8, to obtain a decreasing sequence $\tau^{(1)} \ge \tau^{(2)} \ge \cdots$ of discrete stopping times converging to τ, is adaptable to any number of situations involving the analysis of processes having right-continuous sample paths.

The following proposition and its corollary are often useful for verifying the hypothesis of Theorem 3.8 in examples.

Proposition 3.9 Let $\{Z_n : n \in \mathbb{N}\}$ be real-valued random variables such that for some $\varepsilon > 0$, $\delta > 0$, one has

$$P(Z_{n+1} > \varepsilon \mid \mathcal{G}_n) \geq \delta \text{ a.s.} \quad \forall\, n = 0, 1, 2, \ldots$$

or

$$P(Z_{n+1} < -\varepsilon \mid \mathcal{G}_n) \geq \delta \text{ a.s.} \quad \forall\, n = 0, 1, 2, \ldots, \tag{3.25}$$

where $\mathcal{G}_n = \sigma\{Z_1, \ldots, Z_n\}$ $(n \geq 1)$, $\mathcal{G}_0 = \{\emptyset, \Omega\}$. Let $S_n^x = x + Z_1 + \cdots + Z_n$ $(n \geq 1)$, $S_0^x = x$, and let $a < x < b$. Let τ be the first escape time of $\{S_n^x\}$ from (a, b), i.e., $\tau = \tau^x = \inf\{n \geq 1 : S_n^x \in (a, b)^c\}$. Then $\tau < \infty$ a.s. and

$$\sup_{\{x: a < x < b\}} \mathbb{E}e^{\tau z} < \infty \quad \text{for} \quad -\infty < z < \frac{1}{n_0}\left(\log \frac{1}{1 - \delta_0}\right), \tag{3.26}$$

writing $[y]$ for the integer part of y,

$$n_0 = \left[\frac{b - a}{\varepsilon}\right] + 1, \qquad \delta_0 = \delta^{n_0}. \tag{3.27}$$

Proof Suppose the first relation in (3.25) holds. Clearly, if $Z_j > \varepsilon\, \forall\, j = 1, 2, \ldots, n_0$, then $S_{n_0}^x > b$, so that $\tau \leq n_0$. Therefore, $P(\tau \leq n_0) \geq P(Z_1 > \varepsilon, \ldots, Z_{n_0} > \varepsilon) \geq \delta^{n_0}$, by taking successive conditional expectations (given $\mathcal{G}_{n_0 - 1}, \mathcal{G}_{n_0 - 2}, \ldots, \mathcal{G}_0$, in that order). Hence $P(\tau > n_0) \leq 1 - \delta^{n_0} = 1 - \delta_0$. For every integer $k \geq 2$, $P(\tau > kn_0) = P(\tau > (k-1)n_0, \tau > kn_0) = \mathbb{E}[\mathbf{1}_{[\tau > (k-1)n_0]} P(\tau > kn_0 | \mathcal{G}_{(k-1)n_0})] \leq (1 - \delta_0) P(\tau > (k-1)n_0)$, since, on the set $[\tau > (k-1)n_0]$, $P(\tau \leq kn_0 | \mathcal{G}_{(k-1)n_0}) \geq P(Z_{(k-1)n_0 + 1} > \varepsilon, \ldots, Z_{kn_0} > \varepsilon | \mathcal{G}_{(k-1)n_0}) \geq \delta^{n_0} = \delta_0$. Hence, by induction, $P(\tau > kn_0) \leq (1 - \delta_0)^k$. Hence $P(\tau = \infty) = 0$, and for all $z > 0$,

$$\begin{aligned}
\mathbb{E}e^{z\tau} &= \sum_{r=1}^{\infty} e^{zr} P(\tau = r) \leq \sum_{k=1}^{\infty} e^{zkn_0} \sum_{r=(k-1)n_0+1}^{kn_0} P(\tau = r) \\
&\leq \sum_{k=1}^{\infty} e^{zkn_0} P(\tau > (k-1)n_0) \leq \sum_{k=1}^{\infty} e^{zkn_0} (1 - \delta_0)^{k-1} \\
&= e^{zn_0}(1 - (1 - \delta_0)e^{zn_0})^{-1} \quad \text{if} \quad e^{zn_0}(1 - \delta_0) < 1.
\end{aligned}$$

An entirely analogous argument holds if the second relation in (3.25) holds. ■

The following corollary immediately follows from Proposition 3.9.

Corollary 3.10 Let $\{Z_n : n = 1, 2, \ldots\}$ be an i.i.d. sequence such that $P(Z_1 = 0) < 1$. Let $S_n^n = x + Z_1 + \cdots + Z_n$ $(n \geq 1)$, $S_0^x = x$, and $a < x < b$. Then the first escape time τ of the random walk from the interval (a, b) has a finite moment generating function in a neighborhood of 0.

Example 5 Let $Z_n (n \geq 1)$ be i.i.d. symmetric Bernoulli, $P(Z_i = +1) = P(Z_i = -1) = \frac{1}{2}$, and let $S_n^x = x + Z_1 + \cdots + Z_n (n \geq 1)$, $S_0^x = x$, be the simple symmetric random walk on the state space $\mathbb{Z}$, starting at x. Let $a \leq x \leq b$ be integers, $\tau_y := \inf\{n \geq 0 : S_n^x = y\}$, $\tau = \tau_a \wedge \tau_b = \inf\{n \geq 0 : S_n^x \in \{a, b\}\}$. Then $\{S_n^x : n \geq 0\}$ is a martingale and τ satisfies the hypothesis of Theorem 3.8(c) (Exercise 4). Hence

$$x \equiv \mathbb{E}S_0^x = \mathbb{E}S_\tau^x = aP(\tau_a < \tau_b) + bP(\tau_b < \tau_a) = a + (b - a)P(\tau_b < \tau_a),$$

so that

$$P(\tau_b < \tau_a) = \frac{x - a}{b - a}, \quad P(\tau_a < \tau_b) = \frac{b - x}{b - a}, \quad a \leq x \leq b. \tag{3.28}$$

Letting $a \downarrow -\infty$ in the first relation, and letting $b \uparrow \infty$ in the second, one arrives at the conclusion that the simple symmetric random walk reaches every state with probability one, no mater where it starts. This property is referred to as *recurrence*; also see Example 9, of Chapter II. To illustrate the importance of the hypothesis imposed on τ in Theorem 3.8(c), one may naively try to apply (3.22) to τ_b (see Exercise 4) and arrive at the silly conclusion $x = b$!

Example 6 One may apply Theorem 3.8(c) to a simple asymmetric random walk with $P(Z_i = 1) = p$, $P(Z_i = -1) = q \equiv 1 - p (0 < p < 1, \ p \neq 1/2)$, so that $X_n^x := S_n^x - (2p - 1)n \ (n \geq 1)$, $X_0^x \equiv x$, is a martingale. Then with $\tau_a, \tau_b, \ \tau = \tau_a \wedge \tau_b$ as above, one gets

$$x \equiv \mathbb{E}X_0^x = \mathbb{E}X_\tau^x = \mathbb{E}S_\tau^x - (2p - 1)\mathbb{E}\tau = a + (b - a)P(\tau_b < \tau_a) - (2p - 1)\mathbb{E}\tau. \tag{3.29}$$

Since we do not know $\mathbb{E}\tau$ yet, we can not quite solve (3.29). We therefore use a second martingale $(q/p)^{S_n^x}$ $(n \geq 0)$. Note that $\mathbb{E}[(q/p)^{S_{n+1}^x} | \sigma\{Z_1, \ldots, Z_n\}] = (q/p)^{S_n^x} \cdot \mathbb{E}[(q/p)^{Z_{n+1}}] = (q/p)^{S_n^x}[(q/p)p + (q/p)^{-1}q] = (q/p)^{S_n^x} \cdot 1 = (q/p)^{S_n^x}$, proving the martingale property of the "exponential process" $Y_n := (q/p)^{S_n^x} = \exp(cS_n^x)$, $c = \ln(q/p), n \geq 0$. Note that $(q/p)^{S_{\tau \wedge n}^x} \leq \max\{(q/p)^y : a \leq y \leq b\}$, which is a finite number. Hence the hypothesis of uniform integrability holds. Applying (3.22) we get

$$(q/p)^x = (q/p)^a \cdot P(\tau_a < \tau_b) + (q/p)^b P(\tau_b < \tau_a),$$

or

$$P(\tau_b < \tau_a) = \frac{(q/p)^x - (q/p)^a}{(q/p)^b - (q/p)^a} \equiv \varphi(x) \quad (a \leq x \leq b). \tag{3.30}$$

Using this in (3.29) we get

$$\mathbb{E}\tau \equiv \mathbb{E}\tau_a \wedge \tau_b = \frac{x - a - (b - a)\varphi(x)}{1 - 2p}, \quad a \leq x \leq b. \tag{3.31}$$

Suppose now that $p < q$, i.e., $p < \frac{1}{2}$. Letting $a \downarrow -\infty$ in (3.30), one sees that the probability of ever reaching b starting from $x < b$ is $(\frac{p}{q})^{b-x} < 1$. Similarly if $p > \frac{1}{2}$, i.e., $q < p$, then the probability of ever reaching a starting from $x > a$ is $(\frac{q}{p})^{x-a} < 1$ (Exercise 6).

Very loosely speaking the submartingale and and supermartingale properties convey a sense of "monotonicity"in predictions of successive terms based on the past. This is so much so that the expected values comprise a monotone sequence of numbers. Recall from calculus that every sequence of real numbers bounded above (or below) must have a limit. Perhaps some form of "boundedness"at least seems worthy of consideration in the context of martingale convergence ? Indeed, as we now see, the implications are striking!

Let $\{Z_n : n = 1, 2, \ldots\}$ be a $\{\mathcal{F}_n\}_{n=1}^{\infty}$-submartingale, and $a < b$ arbitrary real numbers. Recursively define successive *crossing times* of (a, b) by $\eta_1 = 1, \eta_2 = \inf\{n \geq 1 : Z_n \geq b\}$, $\eta_{2k-1} = \inf\{n \geq \eta_{2k-2} : Z_n \leq a\}$, $\eta_{2k} = \inf\{n \geq \eta_{2k-1} : Z_n \geq b\}$. In particular η_{2k} is the time of the k-th upcrossing of the interval (a, b) by the sequence $\{Z_n : n = 1, 2, \ldots\}$. η_{2k} is also the time of the k-th upcrossing of $(0, b - a)$ by the sequence $X_n = \max(Z_n - a, 0) = (Z_n - a)^+, n \geq 1$. Note that these crossing times are in fact stopping times. Also, X_n is *nonnegative* and $X_n = 0$ if $Z_n \leq a$, and $X_n \geq b - a$ if $Z_n \geq b$.

For a positive integer N, consider their truncations $\tau_k = \eta_k \wedge N$, which are also $\{\mathcal{F}_n\}$-stopping times, in fact, *bounded* stopping times. Let $U_N = \max\{k : \eta_{2k} \leq N\}$ denote the number of upcrossings of (a, b) by $\{Z_n : n = 1, 2 \ldots\}$ by the time N. Then U_N may also be viewed as the number of upcrossings of the interval $(0, b - a)$ by the submartingale $X_n (n = 1, 2, \ldots)$.

Theorem 3.11 *(Doob's Upcrossing Inequality)* Let $\{Z_n : n \geq 1\}$ be a $\{\mathcal{F}_n\}$-submartingale, and $a < b$ arbitrary real numbers. Then the number U_N of upcrossings of (a, b) by time N satisfies

$$\mathbb{E}U_N \leq \frac{\mathbb{E}(Z_N - a)^+ - \mathbb{E}(Z_1 - a)^+}{b - a} \leq \frac{\mathbb{E}|Z_N| + |a|}{b - a}.$$

Proof For $k > U_N + 1$, $\eta_{2k} > N$, and $\eta_{2k-1} > N$ so that $\tau_{2k} = N$ and $\tau_{2k-1} = N$. Hence, $X_{\tau_{2k}} = X_N = X_{\tau_{2k-1}}$. If $k \leq U_N$ then $\eta_{2k} \leq N$, and $\eta_{2k-1} \leq N$ so that $X_{\tau_{2k}} \geq b - a, 0 = X_{\tau_{2k-1}} = X_{\eta_{2k-1}}$. Now suppose $k = U_N + 1$. Then $\eta_{2k} > N$ and $X_{\tau_{2k}} = X_N$. Also, either $\eta_{2k-1} \geq N$ so that $\tau_{2k-1} = N$, and $X_{\tau_{2k-1}} = X_N$, $X_{\tau_{2k}} - X_{\tau_{2k-1}} = 0$, or $\eta_{2k-1} < N$, in which case $\eta_{2k-1} = \tau_{2k-1}$ and $X_{\tau_{2k-1}} = 0$, so that $X_{\tau_{2k}} - X_{\tau_{2k-1}} = X_N \geq 0$. Thus, in any case, if $k = U_N + 1$, $X_{\tau_{2k}} - X_{\tau_{2k-1}} \geq 0$. Now choose a (nonrandom) integer $m > \frac{N}{2} + 2$. Then $m > U_N + 1$ and one has

$$
\begin{aligned}
X_N - X_1 = X_{\tau_{2m}} - X_1 &= \sum_{k=1}^{m}(X_{\tau_{2k}} - X_{\tau_{2k-1}}) + \sum_{k=2}^{m}(X_{\tau_{2k-1}} - X_{\tau_{2k-2}}) \\
&= \sum_{k=1}^{U_N+1}(X_{\tau_{2k}} - X_{\tau_{2k-1}}) + \sum_{k=2}^{m}(X_{\tau_{2k-1}} - X_{\tau_{2k-2}}) \\
&\geq (b-a)U_N + \sum_{k=2}^{m}(X_{\tau_{2k-1}} - X_{\tau_{2k-2}}). \qquad (3.32)
\end{aligned}
$$

Taking expected values and using the fact from the optional sampling theorem that $\{X_{\tau_k} : k \geq 1\}$ is a submartingale, one has

$$\mathbb{E}X_N - \mathbb{E}X_1 \geq (b-a)\mathbb{E}U_N.$$

■

Remark 3.9 Observe that the relations (3.32) do not require the submartingale assumption on $\{Z_n : n \geq 1\}$. It is merely a relationship among a sequence of numbers.

One of the most significant consequences of the uncrossing inequality is the following.

Theorem 3.12 *(Submartingale Convergence Theorem)* Let $\{Z_n : n \geq 1\}$ be a submartingale such that $\mathbb{E}(Z_n^+)$ is a bounded sequence. Then $\{Z_n : n \geq 1\}$ converges a.s. to a limit Z_∞. If $M := \sup_n \mathbb{E}|Z_n| < \infty$, then Z_∞ is a.s. finite and $\mathbb{E}|Z_\infty| \leq M$.

Proof Let $U(a, b)$ denote the total number of upcrossings of (a, b) by $\{Z_n : n \geq 1\}$. Then $U_N(a, b) \uparrow U(a, b)$ as $N \uparrow \infty$. Therefore, by the monotone convergence theorem

$$\mathbb{E}U(a, b) = \lim_{N\uparrow\infty} \mathbb{E}U_N(a, b) \leq \sup_N \frac{\mathbb{E}Z_N^+ + |a|}{b-a} < \infty. \qquad (3.33)$$

In particular $U(a, b) < \infty$ almost surely, so that

$$P\,(\liminf Z_n < a < b < \limsup Z_n) = 0. \qquad (3.34)$$

Since this holds for every pair $a, b = a + \frac{1}{m}$ with $a \varepsilon Q$ and m a positive integer, and the set of all such pairs is countable, one must have $\liminf Z_n = \limsup Z_n$ almost surely. Let Z_∞ denote the a.s. limit. By Fatou's lemma, $\mathbb{E}|Z_\infty| \leq \lim \mathbb{E}|Z_n|$. ■

An immediate consequence of Theorem 3.12 is

Corollary 3.13 A nonnegative martingale $\{Z_n : n \geq 1\}$ converges almost surely to a finite limit Z_∞. Also, $\mathbb{E}Z_\infty \leq \mathbb{E}Z_1$.

Proof For a nonnegative martingale $\{Z_n : n \geq 1\}$, $|Z_n| = Z_n$ and therefore, $\sup \mathbb{E}|Z_n| = \sup \mathbb{E}Z_n = \mathbb{E}Z_1 < \infty$. Hence the Corollary follows from Theorem 3.12. ■

The following Corollary provides an illustrative application of this theory.

Corollary 3.14 Suppose $X_1, X_2, \ldots$ is a sequence of independent, nonnegative random variables such that $\sum_{n=1}^{\infty} \mathbb{E}X_n < \infty$. Then $\sum_{n=1}^{\infty} X_n$ converges almost surely.

Proof Since $Z_n = \sum_{j=1}^{n}(X_j - \mathbb{E}X_j), n \geq 1$, is a martingale with $\mathbb{E}Z_n^+ \leq 2\sum_{j=1}^{\infty} \mathbb{E}X_j < \infty$ for all n, one has that $Z_\infty = \lim_{n\to\infty} Z_n$ exists. Thus $\sum_{j=1}^{n} X_j = Z_n + \sum_{j=1}^{n} \mathbb{E}X_j$ has an a.s. limit as $n \to \infty$. ■

Doob's upcrossing inequality (Theorem 3.11) also applies to the so-called *reverse martingales, submartingales* defined as follows.

Definition 3.7 Let $\mathcal{F}_n, n \geq 1$, be a decreasing sequence of sub-sigmafields of $\mathcal{F}$, i.e., $\mathcal{F} \supset \mathcal{F}_n \supset \mathcal{F}_{n+1}, n = 1, 2, \ldots$. A sequence $\{X_n : n \geq 1\}$ of integrable random variables on $(\Omega, \mathcal{F}, P)$ is said to be a ***reverse submartingale*** with respect to $\mathcal{F}_n, n \geq 1$, if X_n is $\mathcal{F}_n$-measurable and $\mathbb{E}(X_n|\mathcal{F}_{n+1}) \geq X_{n+1}, \forall n$. If one has equality for each n then the sequence is called a ***reverse martingale*** with respect to $\mathcal{F}_n, n \geq 1$.

Theorem 3.15 *(Reverse submartingale convergence theorem)* Let $\{X_n : n \geq 1\}$ be a reverse submartingale with respect to a decreasing sequence $\mathcal{F}_n, n \geq 1$. Then X_n converges almost surely to an integrable random variable Z as $n \to \infty$.

Proof For each $N > 1$, $\{X_N, X_{N-1}, \ldots, X_1\}$ is a submartingale with respect to the filtration $\{\mathcal{F}_N, \mathcal{F}_{N-1}, \ldots, \mathcal{F}_1\}$. Thus, with U_N denoting the number of up crossings of (a, b) by $\{X_1, \ldots, X_N\}$, Doob's inequality yields $\mathbb{E}U_N \leq \frac{\mathbb{E}|X_1|+|a|}{b-a}$. Arguing as in the proof of of the submartingale convergence theorem (Theorem 3.12), the desired result follows. ■

Remark 3.10 The martingale proof of the strong law of large numbers provides a beautiful illustration of Theorem 3.15 that will be given in Chapter V. Viewed this way, it will follow easily from the reverse martingale convergence theorem that the limit of the sample averages of an i.i.d. sequence of integrable random variables exists. However something more is needed to identify the limit (as the expected value).

Example 7 Let X be an integrable random variable on $(\Omega, \mathcal{F}, P)$ and $\mathcal{F}_n$, $(n \geq 1)$, a decreasing sequence of sigmafields $\mathcal{F}_n \subset \mathcal{F}, n \geq 1$. Then $X_n = \mathbb{E}(X|\mathcal{F}_n), n \geq 1$, is a reverse martingale. Thus $X_n \to Z$ a.s. as $n \to \infty$, for some integrable random variable Z. Note that $\{X_n : n \geq 1\}$ is uniformly integrable since $\int_{[|X_n|>\lambda]} |X_n| dP \leq \int_{[|X_n|>\lambda]} \mathbb{E}(|X||\mathcal{F}_n) dP = \int_{[|X_n|>\lambda]} \mathbb{E}(|X||\mathcal{F}_1) dP$. Hence $X_n \to Z$ in L^1 as well.

Example 8 It follows from the Corollary that the martingales $\{Z_n := \prod_{j=1}^{n} X_j\}$ converge almost surely to an integrable random variable Z_∞, if $\{X_n\}_{n=1}^{\infty}$ is an independent nonnegative sequence with $\mathbb{E}X_n = 1$ for all n. In the case $\{X_n\}_{n=1}^{\infty}$ is *i.i.d.* and $P(X_1 = 1) < 1$, it is an interesting fact that the limit of $\{Z_n : n \geq 1\}$ is 0 a.s., as shown by the following proposition.

Proposition 3.16 Let $\{X_n : n \geq 1\}$ be an i.i.d. sequence of nonnegative random variables with $\mathbb{E}X_1 = 1$. Then $\{Y_n := \prod_{j=1}^n X_j\}$ converges almost surely to 0, provided $P(X_1 = 1) < 1$.

Proof First assume $P(X_1 = 0) > 0$. Then $P(X_n = 0 \text{ for some } n) = 1 - P(X_n > 0 \text{ for all } n) = 0$, since $P(X_j > 0 \text{ for } 1 \leq j \leq n) = (P(X_1 > 0))^n$. But if $X_m = 0$ then $Z_n = 0$ for all $n \geq m$. Therefore, $P(Z_n = 0 \text{ for all sufficiently large } n) = 1$.

Assume now $P(X_1 > 0) = 1$. Consider the *i.i.d.* sequence $\{\log X_n\}_{n=1}^\infty$. Since $x \to \log x$ is concave one has, by Jensen's inequality, $\mathbb{E}\log X_1 \leq \log \mathbb{E}X_1 = 0$. Since $P(X_1 = 1) < 1$, for any $0 < h < 1$, X_1^h is not degenerate (i.e., not almost surely the constant 1.). Hence the Jensen inequality is *strict*. Therefore, $\mathbb{E}X_1^h < 1$. Thus, using Fatou's lemma,

$$0 \leq \mathbb{E}Z_\infty^h \leq \liminf_{n\to\infty} \mathbb{E}Z_n^h = \liminf_{n\to\infty} (\mathbb{E}X_1^h)^n = 0.$$

It follows that $Z_\infty^h = 0$ a.s. ■

Example 9 *(Binary Multiplicative Cascade Measure)* Suppose that one is given a countable collection $\{X_v : v \in \cup_{n=1}^\infty \{0,1\}^n\}$ of positive, mean one random variables indexed by the set of vertices $\partial T = \cup_{n=1}^\infty \{0,1\}^n$ of a binary tree. For $v = (v_1, \dots, v_n) \in \{0,1\}^n$ we write $|v| = n$. For a given "generation"$n \geq 1$, one may consider a corresponding partition of the unit interval $[0, 1)$ into 2^n subintervals $[\frac{k-1}{2^n}, \frac{k}{2^n})$, $k = 1, \dots 2^n$, and assign mass (area) $(\prod_{j=1}^n X_{v|j})2^{-n}$ to the interval indexed by v of length 2^{-n}, where $v|j = (v_1, \dots, v_j)$, $v = (v_1, \dots, v_n) \in \{0,1\}^n$, to create a *random bar graph*. The total area in the graph is then given by

$$Z_n = \sum_{|v|=n} \prod_{j=1}^n X_{v|j} 2^{-n}, n = 1, 2, \dots. \tag{3.35}$$

One may check that $\{Z_n : n \geq 1\}$ is a positive martingale. Thus $\lim_{n\to\infty} Z_n = Z_\infty$ exists almost surely. Moreover, Z_∞ satisfies the recursion

$$Z_\infty = X_0 Z_\infty(0)\frac{1}{2} + X_0 Z_\infty(1)\frac{1}{2}, \tag{3.36}$$

where $Z_\infty(0), Z_\infty(1)$ are mutually independent, and independent of X_0, X_1, and have the same distribution as Z_∞. Let $0 < h \leq 1$. Then by sub-linearity of $z \to z^h$, $z \geq 0$, one has, for a generic random variable X distributed as an X_v,

$$\mathbb{E}Z_\infty^h \geq 2^{1-h}\mathbb{E}X^h \mathbb{E}Z_\infty^h.$$

Thus, if $2^{1-h}\mathbb{E}X^h > 1$ for some $0 < h \leq 1$ then $Z_\infty = 0$ a.s; for otherwise $\mathbb{E}Z_\infty^h > 0$ and one gets the reverse inequality $2^{1-h}\mathbb{E}X^h \leq 1$, or $\chi(h) := \log \mathbb{E}X^h - (h-1)\log 2 \leq 0$. Since $h \to \chi(h), 0 < h \leq 1$ is convex, this is equivalent to $\chi'(1^-) = \mathbb{E}X \log X \leq \log 2$. Thus if $\mathbb{E}X \log X > \log 2$ then $Z_\infty = 0$ a.s. Of course if $X = 1$ a.s. then $Z_\infty = 1$ a.s. as well. In some contexts, the quantity $\mathbb{E}X \log X$ is referred to as a *disorder parameter* and $\log 2$ is a *branching rate*. The heuristic condition for a nonzero limit is that the branching rate be sufficiently large relative to the disorder. This will be confirmed in Chapter V. Let us consider the case in which X is uniform on $[0, 2]$. In this case, as will be verified in Chapter V, one can solve the recursion (3.36), to obtain that Z_∞ has a Gamma distribution with density $ze^{-z}, z \geq 0$. As an alternative for now, we will apply the Chebyshev method from (Chapter I, Example 5) to derive lower bound estimates on $P(Z_\infty \leq z), z \geq 0$. One may check that for X uniformly distributed on $[0, 2]$, $\mathbb{E}X^k = \frac{2^k}{k+1}, k = 1, 2, \ldots$. Moreover, using induction (on k) one sees that the unique positive solution to the equation of moments corresponding to (3.36), namely

$$\begin{aligned} \mathbb{E}Z_\infty^k &= 2^{-k} \sum_{j=0}^{k} \binom{k}{j} \frac{2^j}{j+1} \mathbb{E}Z_\infty^j \frac{2^{k-j}}{k-j+1} \mathbb{E}Z_\infty^{k-j} \\ &= \sum_{j=0}^{k} \binom{k}{j} \frac{\mathbb{E}Z_\infty^j}{1+j} \frac{\mathbb{E}Z_\infty^{k-j}}{1+k-j}, \end{aligned} \tag{3.37}$$

is given by $\mathbb{E}Z_\infty^k = (k+1)!$. Thus the Chebyshev method yields

$$P(Z_\infty \leq x) \geq \begin{cases} 0 & \text{if } x \leq 2, \\ 1 - \frac{6}{x^2} & \text{if } 2 < x \leq 3, \\ \cdots & \\ 1 - \frac{(k+1)!}{x^k} & \text{if } k+1 < x \leq k+2, k = 2, 3, \ldots. \end{cases} \tag{3.38}$$

Example 10 *(Ruin Probability in Insurance Risk)* The *Cramér–Lundberg*, and more generally *Sparre Andersen*, models of insurance markets involve insurance claims of strictly positive random amounts $X_1, X_2, \ldots$ arriving at random time times $T_1, T_2, \ldots$, together with a constant premium rate $c > 0$ per unit time. The two sequences $\{X_n : n \geq 1\}$ of claim sizes and arrival times $\{T_n : n \geq 1\}$ are assumed to be independent. Moreover, the inter-arrival times $A_n = T_n - T_{n-1}, n \geq 1, T_0 = 0$, are assumed to be i.i.d. positive random variables with $\mathbb{E}A_1 = \lambda < \infty$. For a company with initial capital reserves $u > 0$, the probability of ruin is defined by

$$\psi(u) = P(\cup_{n=1}^\infty [\sum_{j=1}^n X_j > u + c \sum_{j=1}^n A_j]) = P(\cup_{n=1}^\infty [\sum_{j=1}^n Y_j > u]), \tag{3.39}$$

where $Y_j := X_j - cA_j$. The common distribution of the i.i.d. sequence $\{Y_j : j \geq 1\}$ is assumed to satisfy the so-called *Net Profit Condition* (NPC)

$$\mathbb{E}Y_1 = \mathbb{E}X_1 - c\mathbb{A}_1 > 0. \tag{3.40}$$

Observe that if $\mathbb{E}Y_1$ is nonnegative and finite then, by the strong law of large numbers (SLLN), one has

$$\psi(u) \equiv 1 \quad \forall\, u. \tag{3.41}$$

To avoid the trivial case in which $\psi(u) = 0\ \forall\, u > 0$, one may assume

$$P(Y_1 > 0) > 0. \tag{3.42}$$

The Cramér–Lundberg model refers to the case in which the $A_n, n \geq 1$, are i.i.d. *exponentially* distributed, while the more general model described above is referred to as the Sparre Andersen model. For the present let us assume that the claim size distribution is *light tailed* in the sense that

$$\mathbb{E}e^{qX_1} < \infty \quad \text{for some } q > 0. \tag{3.43}$$

With this one obtains the following bound on the ruin probability as a function of the initial capital.

Proposition 3.17 *(Lundberg Inequality)* In the non-degenerate case (3.42), the Sparre Andersen model satisfying the NPC (3.40), and the light-tailed claim size distribution condition (3.43), there is a unique parameter $q = R > 0$ such that $\mathbb{E}e^{qY_1} = 1$. Moreover

$$\psi(u) \leq \exp(-Ru), \quad \forall u > 0. \tag{3.44}$$

Proof Observe that the light-tailed condition (3.43) implies that there is an $h, 0 < h \leq \infty$ such that

$$0 < m(q) := \mathbb{E}e^{qY_1} < \infty, \quad \text{for } 0 \leq q < h, \quad \lim_{q \downarrow h} m(q) = \infty.$$

Also $m(0) = 1, m'(0) = \mathbb{E}Y_1 < 0$ (or $m'(0^+) < 0$ if $m(q) = \infty \forall q < 0$), and $m''(q) = \mathbb{E}Y_1^2 \exp(qY_1) > 0, \forall q > 0$, with $m(q) \to \infty$ as $q \uparrow h$. Thus $m(q)$ decreases from $m(0) = 1$ to a minimum in $(0, 1)$ at some $\tilde{q}$ before increasing without bound as $q \uparrow h$. It follows that there is a unique $q = R > 0$ such that $m(q) = 1$. To prove the asserted Lundberg bound, let $\tau = \inf\{n \geq 1 : S_n > u\}$, where $S_n = Y_1 + \cdots + Y_n, n \geq 1, S_0 = 0$. Then τ is a stopping time with respect to the filtration $\mathcal{F}_n = \sigma(Y_1, \ldots, Y_n), n \geq 1, \mathcal{F}_0 = \{\Omega, \emptyset\}$. Then

$$\psi(u) = P(\tau < \infty).$$

Next write $W_n = u - S_n, n \geq 1, W_0 = u$. Then $M_n = \exp\{-RW_n\}, n \geq 0$, is an $\mathcal{F}_n$-martingale since,

$$\mathbb{E}(M_{n+1}|\mathcal{F}_n) = \mathbb{E}(e^{RY_{n+1}}M_n|\mathcal{F}_n) = M_n\mathbb{E}e^{RY_{n+1}} = M_n m(R) = M_n.$$

By the optional sampling theorem one then has

$$e^{-Ru} = \mathbb{E}M_0 = \mathbb{E}M_{\tau\wedge n} \geq \mathbb{E}M_{\tau\wedge n}\mathbf{1}_{[\tau\leq n]} = \mathbb{E}M_\tau \mathbf{1}_{[\tau\leq n]}, \forall n. \quad (3.45)$$

Noting that $M_\tau > 1$ on $[\tau < \infty]$, it follows from (3.45) that $e^{-Ru} \geq \mathbb{E}\mathbf{1}_{[\tau\leq n]} = P(\tau \leq n)$ for all n. Let $n \uparrow \infty$ to obtain the asserted Lundberg bound. ■

Remark 3.11 The parameter R is generally referred to as the *Lundberg coefficient*, or *adjustment coefficient*. It can be shown that the exponential decay rate provided by the Lundberg inequality cannot be improved under the conditions of the theorem. In the Cramér–Lundberg model the true asymptotic rate is given by $\psi(u) \sim ce^{-Ru}$, as $u \to \infty$, for a constant $c < 1$; here $\sim$ demotes asymptotic equality in the sense that the ratio of the two sides converges to one as $u \to \infty$.[1]

Exercise Set III

1. (i) If τ_1 and τ_2 are $\{\mathcal{F}_t\}$-stopping times, then show that so are $\tau_1 \wedge \tau_2$ and $\tau_1 \vee \tau_2$. (ii) Show that $\tau + c$ is an $\{\mathcal{F}_t\}$-stopping time if τ is an $\{\mathcal{F}_t\}$-stopping time, $c > 0$, and $\tau + c \in T \cup \{\infty\}$. (iii) Show that (ii) is false if $c < 0$.
2. If τ is a discrete random variable with values $t_1 < t_2 < \cdots$ in a finite or countable set T in $\mathbb{R}^+$, then (i) τ is an $\{\mathcal{F}_t\}_{t\in T}$-stopping time if and only if $[\tau = t] \in \mathcal{F}_t \ \forall\, t \in T$; (ii) τ is an $\{\mathcal{F}_t\}$-stopping time if and only if it is an $\{\mathcal{F}_t\}$-optional time.
3. (*Wald's Identity*) Let $\{Y_j : j \geq 1\}$ be an i.i.d. sequence with finite mean μ, and take $Y_0 = 0$, $a.s.$ Let τ be an $\{\mathcal{F}_n\}$-stopping time, where $\mathcal{F}_n = \sigma(Y_j : j \leq n)$. Write $S_n = \sum_{j=0}^n Y_j$. If $\mathbb{E}\tau < \infty$ and $\mathbb{E}|S_\tau - S_{\tau\wedge m}| \to 0$ as $m \to \infty$, prove that $\mathbb{E}S_\tau = \mu\mathbb{E}\tau$. [*Hint*: $\{S_n - n\mu : n \geq 0\}$ is a martingale.]
4. In Example 5 for $\tau = \tau_a \wedge \tau_b$, show that (i) $\mathbb{E}\tau < \infty \ \forall\, a \leq x \leq b$, and $|S_{(\tau)\wedge n}| \leq \max\{|a|, |b|\} \ \forall\, n \geq 0$, is uniformly integrable, (ii) $P(\tau_a < \infty) = 1 \ \forall\, x, a$, but $\{S_{\tau_a\wedge n} : n \geq 0\}$ is not uniformly integrable. (iii) For Example 5 also show that $Y_n := S_n^2 - n$, $n \geq 0$, is a martingale and $\{Y_{\tau\wedge n} : n \geq 0\}$ is uniformly integrable. Use this to calculate $\mathbb{E}\tau$. [*Hint*: Use triangle inequality estimates on $|Y_{\tau\wedge n}| \leq |S_{\tau\wedge n}|^2 + \tau \wedge n|$.]
5. (*A cautionary example*) Let $\Omega = \{1, 2\}^3$, and assume all outcomes equally likely. For $\omega = (\omega_1, \omega_2, \omega_3) \in \Omega$, let $Y_i(\omega_1, \omega_2, \omega_3) = \delta_{\omega_i,\omega_{i+1}}$, $(i = 1, 2)$, and $X(\omega_1, \omega_2, \omega_3) = \delta_{\omega_3,\omega_1}$. Define $J(\omega) = 2$ if $Y_1(\omega) = 1$, and $J(\omega) = 1$ if $Y_1(\omega) = 0$, $\omega \in \Omega$. Then show that Y_1 and X are independent, as are Y_2 and X, and J and X. However Y_J is not independent of X.
6. (*Transience of asymmetric simple random walk*) Let $\theta(c|x)$ denote the probability that the simple random walk starting at x ever reaches c. Use (3.30) to prove (i) $\theta(b|x) = (\frac{p}{q})^{b-x}$ for $x < b$ if $p < 1/2$, and (ii) $\theta(a|x) = (\frac{q}{p})^{x-a}$ for $x > a$ if $p > 1/2$.
7. Let $Z_1, Z_2, \ldots$ be i.i.d. ± 1-valued Bernoulli random variables with $P(Z_n = 1) = p$, $P(Z_n = -1) = 1 - p$, $n \geq 1$, where $0 < p < 1/2$. Let $S_n = Z_1 + \cdots + Z_n$, $n \geq 1$, $S_0 = 0$.

[1] See Theorem 5.12 of S. Ramasubramanian (2009) for the asymptotic equality in the case of the Cramér-Lundberg model.

(i) Show that $P(\sup_{n\geq 0} S_n > y) \leq (\frac{p}{q})^y$, $y \geq 0$. [*Hint*: Apply a maximal inequality to $X_n = (q/p)^{S_n}$.]

(ii) Show for $p < 1/2$ that $\mathbb{E}\sup_{n\geq 0} S_n \leq \frac{1}{q-p}$. [*Hint*: Use (1.10), noting that the distribution function is a step function. Also see Exercise 30 of Chapter I.]

8. Suppose that $Z_1, Z_2, \ldots$ is a sequence of independent random variables with $\mathbb{E}Z_n = 0$ such that $\sum_n \mathbb{E}Z_n^2 < \infty$. Show that $\sum_{n=1}^{\infty} Z_n := \lim_N \sum_{n=1}^{N} Z_n$ exists a.s.[2] [*Hint*: Let $S_j = \sum_{k=1}^{j} Z_k$ and show that $\{S_j\}$ is a.s. a Cauchy sequence. For this note that $Y_n := \max_{k,j\geq n} |S_k - S_j|$ is a.s. a decreasing sequence and hence has a limit a.s. Apply Kolmogorov's maximal inequality to $\max_{n\leq j\leq N} |S_j - S_n|$ to show that the limit in probability is zero, and hence a.s. zero; also see Chapter I, Exercise 34.]

(i) For what values of θ will $\sum_{n=1}^{\infty} Z_n$ converge a.s. if $P(Z_n = n^{-\theta}) = P(Z_n = -n^{-\theta}) = 1/2$?

(ii) (Random Signs[3]) Suppose each X_n is symmetric Bernoulli ± 1-valued. Show that the series $\sum_{n=1}^{\infty} X_n a_n$ converges with probability one if $\{a_n\}$ is any square-summable sequence of real numbers.

(iii) Show that $\sum_{n=1}^{\infty} X_n \sin(n\pi t)/n$ converges a.s. for each t if the X_n's are i.i.d. standard normal.

9. Let $\{X_t : t \in T\}$ be a stochastic process on $(\Omega, \mathcal{F})$ with values in some measurable space $(S, \mathcal{S})$, T a discrete set with elements $t_1 < t_2 < \cdots$. Define $\mathcal{F}_t = \sigma(X_s : 0 \leq s \leq t) \subset \mathcal{F}$, $t \in T$. Assume that τ is an $\{\mathcal{F}_t\}$-stopping time and show that $\mathcal{F}_\tau = \sigma(X_{\tau\wedge t} : t \in T)$; i.e., $\mathcal{F}_\tau$ is the σ-field generated by the stopped process $\{X_{\tau\wedge t} : t \in T\}$.

10. Prove (3.19). Also prove that an $\{\mathcal{F}_t\}$-stopping time is an $\{\mathcal{F}_t\}$-optional time; recall Definition 3.4.

11. (i) Prove that τ_B defined by (3.18) is an $\{\mathcal{F}_t\}$-stopping time if B is closed and $t \mapsto X_t$ is continuous with values in a metric space (S, ρ). [*Hint*: For $t > 0$, B closed, $[\tau_B \leq t] = \cap_{n\in\mathbb{N}} \cup_{r\in Q\cap[0,t]} [\rho(X_r, B) \leq \frac{1}{n}]$, where Q is the set of rationals.] (ii) Prove that if $t \mapsto X_t$ is right-continuous, τ_B is an optional time for B open. [*Hint*: For B open, $t > 0$, $[\tau_B < t] = \cup_{r\in Q\cap(0,t)}[X_r \in B]$.] (iii) If $T = \mathbb{N}$ or $\mathbb{Z}^+$, prove that τ_B is a stopping time for all $B \in \mathcal{S}$.

12. Prove that if τ is an optional time with respect to a filtration $\{\mathcal{F}_t : 0 \leq t < \infty\}$, then τ is an optional time with respect to $\{\mathcal{F}_{t+} : 0 \leq t < \infty\}$, where $\mathcal{F}_{t+} := \cap_{\varepsilon>0}\mathcal{F}_{t+\varepsilon}$. Deduce that under the hypothesis of Example 4(b), if B is open or closed, then τ_B is a stopping time with respect to $\{\mathcal{F}_{t+} : 0 \leq t < \infty\}$.

13. Let $\{\mathcal{F}_t : t \in T\}$ and $\{\mathcal{G}_t : t \in T\}$ be two filtrations of $(\Omega, \mathcal{F})$, each adapted to $\{X_t : t \in T\}$, and assume $\mathcal{F}_t \subset \mathcal{G}_t, \forall t \in T$. Show that if $\{X_t : t \in T\}$ is a

[2] A more comprehensive treatment of this class of problems is given in Chapter VIII.

[3] Historically this is the problem that lead Hugo Steinhaus to develop an axiomatic theory of repeated coin tossing based on his reading of Lebesgue's newly developed integral and measure on the real number line. The problem is revisited in Chapter VIII.

$\{\mathcal{G}_t\}$-martingale (or sub or super) then it is an $\{\mathcal{F}_t\}$-martingale (or respectively sub or super).

14. (*Uniform absolute continuity*) Let Y be an integrable random variable. Prove that, given $\varepsilon > 0$, there is a $\delta > 0$ such that $\int_A |Y| dP < \varepsilon$ for every A with $P(A) < \delta$. [*Hint*: Prove by contradiction: There cannot exist a sequence A_n, $P(A_n) < \frac{1}{n}$, and $\int_{A_n} |Y| dP > \varepsilon (n = 1, 2, \ldots)$.]