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 last century. In this section some basic foundations are presented. A more comprehensive treatment of the theory and its applications is provided in our text on stochastic processes.[1] For the prototypical illustration of the martingale property, let $Z_1, Z_2, \ldots$ be an i.i.d. sequence of integrable random variables and let $X_n = Z_1 + \cdots + Z_n,\ n \geq 1$. If $\mathbb{E}Z_1 = 0$ 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

[1] Bhattacharya, R., and E. Waymire (2007): *Theory and Applications of Stochastic Processes*, Springer Graduate Texts in Mathematics.

$\{X_t : t \in T\}$ 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

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

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 $g(Z_{n+1})$ is orthogonal to $L^2(\Omega, \mathcal{F}_n, P)$ for all bounded measurable g such that $\mathbb{E}g(Z_{n+1}) = 0$. 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 with $\mathcal{F}_n = \sigma(X_1, \ldots, X_n)$ 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, with $\mathcal{F}_t$ as in Definition 3.2,

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

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 $\pm X_{n+1} \leq |X_{n+1}|$ taking conditional expectations, that the sequence $\{Y_n \equiv |X_n| : \ n \geq 1\}$ 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 2.7. 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 $\mathcal{G}_n \equiv \sigma(\varphi(X_1), \ldots, \varphi(X_n)) \subseteq \mathcal{F}_n$, 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), given $\mathcal{G}_n$, 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 $\sigma(X_1, \ldots, X_n)$, and not $\sigma(\varphi(X_1), \ldots, \varphi(X_n))$, 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 $\{\mathcal{F}_n\}$-**adapted**). One refers to this sequence as an $\{\mathcal{F}_n\}$**-martingale** (respectively $\{\mathcal{F}_n\}$**-submartingale**). An important example of $\mathcal{F}_n$ larger than $\sigma(X_1, \ldots, X_n)$ is given by "adding independent information" via $\mathcal{F}_n = \sigma(X_1, \ldots, X_n) \vee \mathcal{G}$, where $\mathcal{G}$ is a σ-field independent of $\sigma(X_1, X_2, \ldots)$, 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 10.

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., $\mathcal{F}_s \subseteq \mathcal{F}_t$ 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

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

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,

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

Remark 3.3. By an obvious change in the definition of $A_k (k = 1, \ldots, n)$, one obtains (3.11) with strict inequality $M_n > \lambda$ on both sides of the asserted inequality.

Remark 3.4. 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 $\{\mathcal{F}_k : 1 \le k \le n\}$-martingale such that $\mathbb{E}X_n^2 < \infty$. Then $\mathbb{E}M_n^2 \le 4\mathbb{E}X_n^2$.

Proof. A standard application of the Fubini–Tonelli theorem (see (2.5)) provides the second moment formula

$$\mathbb{E}M_n^2 = 2\int_0^\infty \lambda P(M_n > \lambda)d\lambda. \tag{3.12}$$

Applying the first inequality in (3.11), then another application of the Fubini–Tonelli theorem, and finally the Cauchy–Schwarz inequality, it follows that

$$\begin{aligned}\mathbb{E}M_n^2 &\le 2\int_0^\infty \mathbb{E}\left(|X_n|\mathbf{1}_{[M_n \ge \lambda]}\right)d\lambda = 2\mathbb{E}\left(|X_n|M_n\right)\\ &\le 2\sqrt{\mathbb{E}X_n^2}\sqrt{\mathbb{E}M_n^2}.\end{aligned}$$

Divide both sides by $\sqrt{\mathbb{E}M_n^2}$ to complete the proof. ■

Corollary 3.4. 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 $\lambda > 0$,

$$P(M_T > \lambda) \le \frac{1}{\lambda^p}\int_{[M_T > \lambda]} X_T^p dP \le \frac{1}{\lambda^p}\mathbb{E}X_T^p. \tag{3.13}$$

Proof. Consider the nonnegative submartingale $\{X_0, X_{T2^{-n}}, \ldots, X_{Ti2^{-n}}, \ldots, X_T\}$, for each $n = 1, 2, \ldots$, 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 $n \uparrow \infty$, (3.13) is obtained. ■

We finally come to the notion of **stopping 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 $[\tau \le t] \in \mathcal{F}_t \ \forall \ t \in T$. 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 $B \in \mathcal{S}$,

$$\tau_B := \inf\{t \ge 0 : X_t \in B\} \tag{3.14}$$

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 $t \ge 0$, and hence τ_B is an optional time.

We leave the proofs of (a)–(c) as Exercise 2. Note that (b), (c) imply that under the hypothesis of (b), τ_B is an optional time if B is open or closed.

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 $A \cap [\tau \le t] \in \mathcal{F}_t$ 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 $\mathcal{F}_t = \sigma(X_s : 0 \le s \le t) \subseteq \mathcal{F}, \forall t$, where $\{X_t : t \in T\}$ is a process with values in some measurable space $(S, \mathcal{S})$, then $\mathcal{F}_\tau = \sigma(X_{\tau \wedge t} : t \ge 0)$; (Exercise 8). The stochastic process $\{X_{\tau \wedge t} : t \ge 0\}$ is referred to as the **stopped process**.

If τ_1, τ_2 are two $\{\mathcal{F}_t\}$-stopping times and $\tau_1 \le \tau_2$, then it is simple to check that (Exercise 1)

$$\mathcal{F}_{\tau_1} \subseteq \mathcal{F}_{\tau_2}. \tag{3.15}$$

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 3.5. (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 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., $[X_\tau\in B]\cap[\tau\le t]\in\tilde{\mathcal{F}}_t\subseteq\mathcal{F}_t$ 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 $\{(s,\omega)\in[0,t]\times\Omega : X_s^{(n)}(\omega)\in B\} = \cup_{j=1}^{2^n}([j-1)2^{-n}t, j2^{-n}t)\times\{\omega : X_{j2^{-n}t}(\omega)\in B\})\cup(\{t\}\times\{\omega : X_t(\omega)\in B\})\in\mathcal{B}[0,t]\otimes\mathcal{F}_t$, and $X_s^{(n)}(\omega)\to X_s(\omega)$ for all (s,ω) as $n\to\infty$, in view of the right-continuity of $s\mapsto X_s(\omega)$, it follows that $\{(s,\omega)\in[0,t]\times\Omega : X_s(\omega)\in B\}\in\mathcal{B}[0,t]\otimes\mathcal{F}_t$. ■

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 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 $B\subseteq A$ 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 3.5.

Theorem 3.6. *(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 $\tau_1\le\tau_2$ are bounded stopping times, then

$$\mathbb{E}(X_{\tau_2}|\mathcal{F}_{\tau_1}) = X_{\tau_1}. \tag{3.16}$$

(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 τ is a stopping time such that (i) $P(\tau < \infty) = 1$, and (ii) $X_{\tau\wedge t}(t \in T)$ is uniformly integrable. Then

$$\mathbb{E}X_\tau = \mathbb{E}X_0. \tag{3.17}$$

Proof. We will give a proof for the case $T = [0,\infty)$. The case $T = \mathbb{N}$ is similar and simpler (Exercise 5). Let $\tau_1 \le \tau_2 \le t_0$ a.s. The idea for the proof is 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 $\mathcal{F}_{\tau_1} \subseteq \mathcal{F}_{\tau_2}$. That is, it will then follow that

$$\mathbb{E}[X_{\tau_2}|\mathcal{F}_{\tau_1}] = \mathbb{E}[\mathbb{E}(X_{t_0}|\mathcal{F}_{\tau_2})|\mathcal{F}_{\tau_1}] = \mathbb{E}[X_{t_0}|\mathcal{F}_{\tau_1}] = X_{\tau_1}. \tag{3.18}$$

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 $\mathcal{F}_\tau \subseteq \mathcal{F}_{\tau^{(n)}}$ (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

$$\begin{aligned}
\mathbb{E}(\mathbf{1}_G X_{t_0}) &= \sum_{k=0}^{2^n-1} \mathbb{E}(\mathbf{1}_{G\cap[\tau^{(n)}=(k+1)2^{-n}t_0]} X_{t_0}) \\
&= \sum_{k=0}^{2^n-1} \mathbb{E}(\mathbf{1}_{G\cap[\tau^{(n)}=(k+1)2^{-n}t_0]} X_{(k+1)2^{-n}t_0}) \\
&= \sum_{k=0}^{2^n-1} \mathbb{E}(\mathbf{1}_{G\cap[\tau^{(n)}=(k+1)2^{-n}t_0]} X_{\tau^{(n)}}) = \mathbb{E}(\mathbf{1}_G X_{\tau^{(n)}}) \to \mathbb{E}(\mathbf{1}_G X_\tau).
\end{aligned} \tag{3.19}$$

The last convergence is due to the L^1-convergence criterion of Theorem 1.8 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 (3.19) 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.

(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 3.6. If $\{X_t : t \in T\}$ in Theorem 3.6 is taken to be a submartingale, then instead of the equality sign "=" in (3.16), (3.17), one gets "$\le$."

The following proposition and its corollary are often useful for verifying the hypothesis of Theorem 3.6 in examples.

Proposition 3.7. Let $\{Z_n : n \in \mathbb{N}\}$ be real-valued random variables such that for some $\varepsilon > 0$, $\delta > 0$, one has

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

or

$$P(Z_{n+1} < -\varepsilon \mid \mathcal{G}_n) \ge \delta \text{ a.s.} \quad \forall\ n = 0, 1, 2, \ldots, \tag{3.20}$$

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

$$\sup_{\{x: a<x<b\}} \mathbb{E}e^{\tau z} < \infty \quad \text{for} \quad -\infty < z < \frac{1}{n_0}\left(\log \frac{1}{1-\delta_0}\right), \tag{3.21}$$

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

$$n_0 = \left[\frac{b-a}{\varepsilon}\right] + 1, \qquad \delta_0 = \delta^{n_0}. \tag{3.22}$$

Proof. Suppose the first relation in (3.20) holds. Clearly, if $Z_j > \varepsilon\ \forall\ j = 1, 2, \ldots, n_0$, then $S_{n_0}^x > b$, so that $\tau \le n_0$. Therefore, $P(\tau \le n_0) \ge P(Z_1 > \varepsilon, \ldots, Z_{n_0} > \varepsilon) \ge \delta^{n_0}$, by taking successive conditional expectations (given $\mathcal{G}_{n_0-1}, \mathcal{G}_{n_0-2}, \ldots, \mathcal{G}_0$, in that order). Hence $P(\tau > n_0) \le 1 - \delta^{n_0} = 1 - \delta_0$. For every integer $k \ge 2$, $P(\tau > kn_0) = P(\tau > (k-1)n_0, \tau > kn_0) = \mathbb{E}[\mathbf{1}_{[\tau>(k-1)n_0]} P(\tau > kn_0 | \mathcal{G}_{(k-1)n_0})] \le (1-\delta_0) P(\tau > (k-1)n_0)$, since, on the set $[\tau > (k-1)n_0]$, $P(\tau \le kn_0 | \mathcal{G}_{(k-1)n_0}) \ge P(Z_{(k-1)n_0+1} > \varepsilon, \ldots, Z_{kn_0} > \varepsilon | \mathcal{G}_{(k-1)n_0}) \ge \delta^{n_0} = \delta_0$. By induction, $P(\tau > kn_0) \le (1-\delta_0)^k$.

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 (3.20) holds. ■

The following corollary immediately follows from Proposition 3.7.

Corollary 3.8. Let $\{Z_n : n = 1, 2, \cdots\}$ 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 3.6 (c) (Exercise 7). 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

$$P(\tau_b < \tau_a) = \frac{x-a}{b-a}, \quad P(\tau_a < \tau_b) = \frac{b-x}{b-a}, \quad a \leq x \leq b. \tag{3.23}$$

To illustrate the importance of the hypothesis imposed on τ in Theorem 3.6 (c), one may naively try to apply (3.17) to τ_b (see Exercise 7) and arrive at the silly conclusion $x = b$!

Example 6. One may apply Theorem 3.6 (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

$$x \equiv \mathbb{E}X_0^x = \mathbb{E}X_\tau^x = \mathbb{E}S_\tau^x - (2p-1)\mathbb{E}\tau = a + (b-a)P(\tau_b < \tau_a) - (2p-1)\mathbb{E}\tau. \tag{3.24}$$

Since we do not know $\mathbb{E}\tau$ yet, we can not quite solve (3.24). 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 (3.17) we get

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

or

$$P(\tau_b < \tau_a) = \frac{(q/p)^x - (q/p)^a}{(q/p)^b - (q/p)^a} \equiv \varphi(x) \quad (a \leq x \leq b). \tag{3.25}$$

Using this in (3.24) we get

$$\mathbb{E}\tau \equiv \mathbb{E}\tau_a \wedge \tau_b = \frac{x - a - (b-a)\varphi(x)}{1-2p}, \quad a \leq x \leq b. \tag{3.26}$$

EXERCISES

Exercise Set III

1. Prove (3.15). Also prove that an $\{\mathcal{F}_t\}$-stopping time is an $\{\mathcal{F}_t\}$-optional time.

2. (i) Prove that τ_B defined by (3.14) 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 $B \in \mathcal{S}$.

3. (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 $c < 0$.

4. 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 τ is an $\{\mathcal{F}_t\}_{t\in T}$-stopping time if and only if $[\tau = t] \in \mathcal{F}_t \ \forall\, t \in T$.

5. Let $\{X_n : n = 0, 1, 2, \ldots\}$ be an $\{\mathcal{F}_n\}$-martingale, and τ an $\{\mathcal{F}_n\}$-stopping time. Give simple direct proofs of the following: (i) $\mathbb{E}X_\tau = \mathbb{E}X_0$ if τ is bounded. [*Hint*: Let $\tau \leq m$ a.s. Then $\mathbb{E}X_\tau = \sum_{n=0}^m \mathbb{E}X_n \mathbf{1}_{[\tau=n]} = \sum_{n=0}^m \mathbb{E}X_m \mathbf{1}_{[\tau=n]} = \mathbb{E}X_m \mathbf{1}_{[\tau\leq m]} = \mathbb{E}X_m = \mathbb{E}X_0$.] (ii) If $\mathbb{E}\tau < \infty$ and $\mathbb{E}|X_{\tau\wedge m} - X_\tau| \to 0$ as $m \to \infty$, then $\mathbb{E}X_\tau = \mathbb{X}_0$.

6. (*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*: Apply Theorem 3.6(c) to the martingale $\{S_n - n\mu : n \geq 0\}$.]

7. In Example 5 for $\tau = \tau_a \wedge \tau_b$, show that (i) $\mathbb{E}\tau < \infty \ \forall \ a \le x \le b$, and $|S_{\tau\wedge n}| \le \max\{|a|,|b|\} \ \forall \ n \ge 0$, is uniformly integrable, (ii) $P(\tau_a < \infty) = 1 \ \forall \ x, a$, but $\{S_{\tau_a \wedge n} : n \ge 0\}$ is not uniformly integrable. (iii) For Example 5 also show that $Y_n := S_n^2 - n, n \ge 0$, is a martingale and $\{Y_{\tau\wedge n} : n \ge 0\}$ is uniformly integrable. Use this to calculate $\mathbb{E}\tau$. [*Hint*: Use triangle inequality estimates on $|Y_{\tau\wedge n}| \le |S_{\tau\wedge n}|^2 + \tau \wedge n$.]

8. 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 $\mathcal{F}_t = \sigma(X_s : 0 \le s \le t) \subseteq \mathcal{F}$, $t \in T$. 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 $\{X_{\tau\wedge t} : t \in T\}$.

9. Prove that if τ is an optional time with respect to a filtration $\{\mathcal{F}_t : 0 \le t < \infty\}$, then τ is a stopping time with respect to $\{\mathcal{F}_{t+} : 0 \le t < \infty\}$, 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 $\{\mathcal{F}_{t+} : 0 \le t < \infty\}$.

10. 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 $\mathcal{F}_t \subseteq \mathcal{G}_t, \forall t \in T$. Show that if $\{X_t : t \in T\}$ is a $\{\mathcal{G}_t\}$-martingale (or sub or super) then it is an $\{\mathcal{F}_t\}$-martingale (or respectively sub or super).

11. Let $Z_1, Z_2, \ldots$ be i.i.d. ± 1-valued Bernoulli random variables with $P(Z_n = 1) = p, P(Z_n = -1) = 1 - p, n \ge 1$, where $0 < p < 1/2$. Let $S_n = Z_1 + \cdots + Z_n, n \ge 1, S_0 = 0$.
   (i) Show that $P(\sup_{n\ge 0} S_n > y) \le (\frac{p}{q})^y, y \ge 0$. [*Hint*: Apply a maximal inequality to $X_n = (q/p)^{S_n}$.]
   (ii) Show for $p < 1/2$ that $\mathbb{E}\sup_{n\ge 0} S_n \le \frac{p}{q-p}$. [*Hint*: Use (2.5).]

12. 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. [*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 $Y_n := \max_{k,j\ge n} |S_k - S_j|$ is a.s. a decreasing sequence and hence has a limit a.s. Apply Kolmogorov's maximal inequality to $\max_{n\le j\le N} |S_j - S_n|$ to show that the limit in probability is zero, and hence a.s. zero.]
   (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) Suppose each Z_n is symmetric Bernoulli ± 1-valued. Show that the series $\sum_{n=1}^\infty Z_n a_n$ converges with probability one if $\{a_n\}$ is any square-summable sequence of real numbers.
   (iii) Show that $\sum_{n=1}^\infty Z_n \sin(n\pi t)/n$ converges a.s. for each t if the Z_n's are i.i.d. standard normal.