Chapter V

Martingales and Stochastics

This chapter is to introduce the vocabulary for describing the evolution of random systems over time. It will also cover the basic results of classical martingale theory and mention some basic processes such as Markov chains, Poisson processes, and Brownian motion. This chapter should be treated as a reference source for chapters to come.

We start with generalities on filtrations and stopping times, go on to martingales in discrete time, and then to finer results on martingales and filtrations in continuous time. Throughout, $(\Omega, \mathcal{H}, \mathbb{P})$ is a fixed probability space in the background, and all stochastic processes are indexed by some set $\mathbb{T}$, which is either $\mathbb{N} = \{0, 1, \ldots\}$ or $\mathbb{R}_+ = [0, \infty)$ or some other subset of $\bar{\mathbb{R}} = [-\infty, +\infty]$. We think of $\mathbb{T}$ as the time-set; its elements are called times. On a first reading, the reader should take $\mathbb{T} = \mathbb{N}$.

1 Filtrations and Stopping Times

Let $\mathbb{T}$ be a subset of $\bar{\mathbb{R}}$. A *filtration* on $\mathbb{T}$ is an increasing family of sub-σ-algebras of $\mathcal{H}$ indexed by $\mathbb{T}$; that is, $\mathcal{F} = (\mathcal{F}_t)_{t \in \mathbb{T}}$ is a *filtration* if each $\mathcal{F}_t$ is a σ-algebra on Ω, each $\mathcal{F}_t$ is a subset of $\mathcal{H}$, and $\mathcal{F}_s \subset \mathcal{F}_t$ whenever $s < t$. Given a stochastic process $X = (X_t)_{t \in \mathbb{T}}$, letting $\mathcal{F}_t = \sigma\{X_s : s \leq t\}$ for each time t, we obtain a filtration $\mathcal{F} = (\mathcal{F}_t)_{t \in \mathbb{T}}$; it is called the *filtration generated* by X.

Heuristically, we think of a filtration $\mathcal{F}$ as a flow of information, with $\mathcal{F}_t$ representing the body of information accumulated by time t by some observer of the ongoing experiment modeled by $(\Omega, \mathcal{H}, \mathbb{P})$. Or, we may think of $\mathcal{F}_t$ as the collection of $\bar{\mathbb{R}}$-valued random variables V such that the observer can tell the value $V(\omega)$ at the latest by time t, whatever the outcome ω turns out to be. Of course, it is possible to have different observers with different

E. Çınlar, *Probability and Stochastics*, Graduate Texts in Mathematics 261, DOI 10.1007/978-0-387-87859-1_5,

information flows. Given two filtrations $\mathcal{F}$ and $\mathcal{G}$, we say that $\mathcal{F}$ is *finer* than $\mathcal{G}$, or that $\mathcal{G}$ is *coarser* than $\mathcal{F}$, if $\mathcal{F}_t \supset \mathcal{G}_t$ for every time t.

Adaptedness

Let $\mathcal{F} = (\mathcal{F}_t)_{t\in\mathbb{T}}$ be a filtration. Let $X = (X_t)_{t\in\mathbb{T}}$ be a stochastic process with some state space $(E, \mathcal{E})$. Then X is said to be *adapted* to $\mathcal{F}$ if, for every time t, the variable X_t is measurable with respect to $\mathcal{F}_t$ and $\mathcal{E}$. Since $\mathcal{F}$ is increasing, this is equivalent to saying that, for each t, the numerical random variables $f \circ X_s$ belong to $\mathcal{F}_t$ for all f in $\mathcal{E}$ and all times $s \le t$.

Every stochastic process is automatically adapted to the filtration it generates. Thus, if $\mathcal{G}$ is the filtration generated by X, saying that X is adapted to $\mathcal{F}$ is the same as saying that $\mathcal{F}$ is finer than $\mathcal{G}$.

Stopping times

1.1 DEFINITION. *Let $\mathcal{F}$ be a filtration on $\mathbb{T}$. A random time $T : \Omega \mapsto \bar{\mathbb{T}} = \mathbb{T} \cup \{+\infty\}$ is called a* stopping time *of $\mathcal{F}$ if*

1.2 $$\{T \le t\} \in \mathcal{F}_t \quad \textit{for each} \quad t \in \mathbb{T}.$$

1.3 REMARKS. The condition 1.2 is equivalent to requiring that the process

1.4 $$Z_t = 1_{\{T\le t\}}, \quad t \in \mathbb{T},$$

be adapted to $\mathcal{F}$. When $\mathbb{T}$ is $\mathbb{N}$ or $\bar{\mathbb{N}}$, this is equivalent to requiring that

1.5 $$\hat{Z}_n = 1_{\{T=n\}}, \quad n \in \mathbb{N},$$

be adapted to $(\mathcal{F}_n)$; this follows from the preceding remark by noting that $\hat{Z}_n = Z_n - Z_{n-1}$.

Heuristically, a random time signals the occurrence of some physical event. The process Z defined by 1.4 is indeed the indicator of whether that event has or has not occurred: $Z_t(\omega) = 0$ if $t < T(\omega)$ and $Z_t(\omega) = 1$ if $t \ge T(\omega)$. Recalling the heuristic meaning of adaptedness, we conclude that T is a stopping time of $\mathcal{F}$ if the information flow $\mathcal{F}$ enables us to detect the occurrence of that physical event as soon as it occurs, as opposed to inferring its occurrence sometime later. In still other words, T is a stopping time of $\mathcal{F}$ if the information flow $\mathcal{F}$ is such that we can tell what $T(\omega)$ is at the time $T(\omega)$, rather than by inference at some time after $T(\omega)$. These heuristic remarks are more transparent when the time set is $\mathbb{N}$.

The following mental test incorporates all these remarks into a virtual alarm system. Imagine a computer that is being fed the flow $\mathcal{F}$ of information and that is capable of checking, at each time t, whether $\omega \in H$ for every possible ω in Ω and every event H in $\mathcal{F}_t$. If it is possible to attach to it an alarm system that sounds exactly at time T, and only at time T, then T is a stopping time of $\mathcal{F}$. This alarm test will be heard on and off below.

1.6 EXAMPLE. Let $\mathbb{T} = \mathbb{N}$, let $\mathcal{F}$ be a filtration on $\mathbb{N}$, and let X be a process with index set $\mathbb{N}$ and some state space $(E, \mathcal{E})$. Suppose that X is adapted to $\mathcal{F}$. For fixed A in $\mathcal{E}$, let

$$T(\omega) = \inf\{\ n \in \mathbb{N} : X_n(\omega) \in A\ \}, \quad \omega \in \Omega.$$

Then T is called the time of first entrance to A. (Note that $T(\omega) = +\infty$ if $X_n(\omega)$ is never in A, which is the reason for allowing $+\infty$ as a value for random times in general.) This T is a stopping time of $\mathcal{F}$: Heuristically, X is adapted to $\mathcal{F}$ means that the computer is able to check, at each time n, whether $X_n(\omega) \in A$; and it seems trivial to design an alarm system that sounds exactly at the first n such that $X_n(\omega) \in A$. More precisely, T is a stopping time because, for each n in $\mathbb{N}$,

$$\{\ T \le n\ \} = \bigcup_{k=0}^{n} \{\ X_k \in A\ \}$$

belongs to $\mathcal{F}_n$, since the events $\{X_k \in A\}$, $0 \le k \le n$, are all in $\mathcal{F}_n$. In contrast,

$$L(\omega) = 0 \vee \sup\{\ n \le 5\ :\ X_n(\omega) \in A\ \}$$

is not a stopping time (except in some special cases depending on A, for instance, if entering A means never coming back to A). Because, if the outcome ω is such that $X_4(\omega) \in A$ and $X_5(\omega) \notin A$, then $L(\omega) = 4$, but the information we had at time 4 is not sufficient to conclude that $L(\omega) = 4$. So, there can be no alarm system that will sound at exactly $L(\omega)$.

1.7 EXAMPLE. *Counting Processes.* Let $0 < T_1 < T_2 < \cdots$ be some random times taking values in $\mathbb{R}_+$ and assume that $\lim T_n = +\infty$. Define

$$N_t = \sum_{1}^{\infty} 1_{[0,t]} \circ T_n, \quad t \in \mathbb{R}_+,$$

and note that $t \mapsto N_t$ is increasing and right-continuous and increases only by jumps of size one, and $N_0 = 0$ and $N_t < \infty$ for every t in $\mathbb{R}_+$, with $\lim_{t\to\infty} N_t = +\infty$. We may regard $T_1, T_2, \ldots$ as the times of successive arrivals at a store; then N_t is the number of arrivals during $[0, t]$. Let $\mathcal{F} = (\mathcal{F}_t)_{t\in\mathbb{R}_+}$ be the filtration generated by $N = (N_t)$. Then, for each integer $k \ge 1$, the time T_k is a stopping time of $\mathcal{F}$: for every t in $\mathbb{R}_+$

$$\{T_k \le t\} = \{N_t \ge k\} \in \mathcal{F}_t$$

since N_t is in $\mathcal{F}_t$. Heuristically, T_k is a stopping time because it is possible to construct an alarm system that sounds exactly at the time of k^{th} arrival. Another stopping time is

$$T = \inf\{\ t \ge a :\ N_t = N_{t-a}\ \},$$

where $a > 0$ is fixed, that is, the first time that an interval of length a passes without an arrival. We leave the proof to Exercise 1.35, because it needs tools to be developed below. Finally, here is a random time that is not a stopping time: fix $b > 0$, and let

$$L = \inf\{\ t \in \mathbb{R}_+ :\ N_t = N_b\ \},$$

that is, L is the time of last arrival before the time b if there is one, and it is 0 if there is none.

Conventions for the end of time

Soon we shall introduce the concept of information accumulated by the time T, when the alarm sounds. Since stopping times can take $+\infty$ as a value, in case $+\infty$ is not in $\mathbb{T}$, we need to extend the definition of the filtration $\mathcal{F}$ on $\mathbb{T}$ onto $\bar{\mathbb{T}} = \mathbb{T} \cup \{+\infty\}$. We do so by letting $\mathcal{F}_\infty$, which we also denote by $\lim \mathcal{F}_t$, be defined as

1.8 $$\mathcal{F}_\infty = \lim \mathcal{F}_t = \bigvee_{t\in\mathbb{T}} \mathcal{F}_t,$$

the σ-algebra generated by the union of all the $\mathcal{F}_t$. Then, $(\mathcal{F}_t)_{t\in\bar{\mathbb{T}}}$ is a filtration on $\bar{\mathbb{T}}$, and T is a stopping time of it if and only if T is a stopping time of $(\mathcal{F}_t)_{t\in\mathbb{T}}$. Also, every adapted process X indexed by $\mathbb{T}$ can be extended onto $\bar{\mathbb{T}}$ by appending to X an arbitrary variable X_∞ picked from $\mathcal{F}_\infty$. We shall still write $\mathcal{F}$ for the extended filtration, especially since $\mathcal{F}_\infty$ has no information in it that was not in $(\mathcal{F}_t)_{t\in\mathbb{T}}$.

Past until T

Let $\mathcal{F}$ be a filtration on $\mathbb{T}$, extended to $\bar{\mathbb{T}}$ as above. Let T be a stopping time of it. Corresponding to the notion of the body of information accumulated by the time T, we define

1.9 $$\mathcal{F}_T = \{\ H \in \mathcal{H} :\ H \cap \{T \le t\} \in \mathcal{F}_t \ \text{ for each } t \text{ in } \bar{\mathbb{T}}\ \}.$$

It is easy to check that $\mathcal{F}_T$ is a σ-algebra and that $\mathcal{F}_T \subset \mathcal{F}_\infty \subset \mathcal{H}$; it is called the *past until* T.

If T is a fixed time, say $T(\omega) = t$ for all ω for some constant t in $\bar{\mathbb{T}}$, then $\mathcal{F}_T = \mathcal{F}_t$; hence, there is no ambiguity in the notation $\mathcal{F}_T$.

For an arbitrary stopping time T, note that the event $\{T \le r\}$ belongs to $\mathcal{F}_T$ for every $r \ge 0$, because

$$\{T \le r\} \cap \{T \le t\} = \{T \le r \wedge t\} \in \mathcal{F}_t$$

for each t. Thus, T is $\mathcal{F}_T$-measurable.

As usual with σ-algebras, $\mathcal{F}_T$ will also denote the collection of all $\mathcal{F}_T$-measurable random variables. Heuristically, then, 1.9 is equivalent to saying that $\mathcal{F}_T$ consists of those $\bar{\mathbb{R}}$-valued variables V such that, for every possibility ω, the value $V(\omega)$ can be told by the time $T(\omega)$, the time of the alarm sound. The following is the precise version.

1.10 THEOREM. *A random variable V belongs to $\mathcal{F}_T$ if and only if*

1.11
$$V\, 1_{\{T\le t\}} \in \mathcal{F}_t$$

for every t in $\bar{\mathbb{T}}$. In particular, if $\bar{\mathbb{T}} = \bar{\mathbb{N}}$, the condition is equivalent to requiring that, for every n in $\bar{\mathbb{N}}$,

1.12
$$V \cdot 1_{\{T=n\}} \in \mathcal{F}_n.$$

Proof. We may and do assume that V is positive and let X_t be the random variable appearing in 1.11. Then, for all r in $\mathbb{R}_+$ and t in $\mathbb{T}$,

$$\{V > r\} \cap \{T \le t\} = \{X_t > r\}.$$

Thus, by the definition 1.9, the event $\{V > r\}$ is in $\mathcal{F}_T$ for all r if and only if the event $\{X_t > r\}$ is in $\mathcal{F}_t$ for all r for every t in $\bar{\mathbb{T}}$. That is, $V \in \mathcal{F}_T$ if and only if $X_t \in \mathcal{F}_t$ for every t in $\bar{\mathbb{T}}$, which is the claim of the first statement. The particular statement for the case $\bar{\mathbb{T}} = \bar{\mathbb{N}}$ is immediate upon noting that

$$V \cdot 1_{\{T=n\}} = \begin{cases} X_n - X_{n-1} & \text{if } n \in \mathbb{N} \\ X_\infty - \sum_{n\in\mathbb{N}} (X_n - X_{n-1}) & \text{if } n = +\infty. \end{cases} \qquad \square$$

Representation of $\mathcal{F}$ and $\mathcal{F}_T$

Let $\mathbb{T} \subset \bar{\mathbb{R}}$. Let $\mathcal{F}$ be a filtration on it, extended onto $\bar{\mathbb{T}} = \mathbb{T} \cup \{+\infty\}$ if $\mathbb{T}$ does not include the point $+\infty$. We identify $\mathcal{F}$ with the collection of all right-continuous processes on $\bar{\mathbb{T}}$ that are adapted to $\mathcal{F}$. More precisely, we say that $X \in \mathcal{F}$ if

1.13 a) $X = (X_t)_{t\in\bar{\mathbb{T}}}$ is adapted to $\mathcal{F} = (\mathcal{F}_t)_{t\in\bar{\mathbb{T}}}$, and
b) the path $t \mapsto X_t(\omega)$ from $\bar{\mathbb{T}}$ into $\bar{\mathbb{R}}$ is right-continuous for each ω in Ω.

REMARK. If $\mathbb{T}$ is $\mathbb{N}$ or $\bar{\mathbb{N}}$, then the condition (b) above holds automatically, because every path $n \mapsto X_n(\omega)$ is continuous in the discrete topology on $\bar{\mathbb{N}}$, which is the topology induced on $\bar{\mathbb{N}}$ by the ordinary topology of $\bar{\mathbb{R}}_+$. Consequently, in these cases,

$$X \in \mathcal{F} \iff X_n \in \mathcal{F}_n \;\text{ for each } n \text{ in } \bar{\mathbb{N}},$$

and the notation $X \in \mathcal{F}$ is amply justified.

The following characterization theorem shows the economy of thought achieved by this device: $\mathcal{F}_T$ consists of the values X_T of processes X in $\mathcal{F}$ at the time T. For a much simpler proof in the case of discrete time, see Exercise 1.32.

1.14 THEOREM. *Let T be a stopping time of $\mathcal{F}$, Then,*

$$\mathcal{F}_T = \{ \ X_T : \ X \in \mathcal{F} \ \}.$$

Proof. a) Let $V \in \mathcal{F}_T$. Define $X_t = V \ 1_{\{T \le t\}}$, $t \in \bar{\mathbb{T}}$. Then, X is adapted to $\mathcal{F}$ by Theorem 1.10 and is obviously right-continuous, that is, $X \in \mathcal{F}$. Clearly $X_T = V$. So, $\mathcal{F}_T \subset \{ \ X_T : \ X \in \mathcal{F} \ \}$.

b) To show the converse that $\mathcal{F}_T \supset \{ \ X_T : \ X \in \mathcal{F} \ \}$, we let $X \in \mathcal{F}$, put $V = X_T$, and proceed to show that $V \in \mathcal{F}_T$. To that end, in view of Theorem 1.10, it is enough to show that $V \ 1_{\{T \le t\}} \in \mathcal{F}_t$ for every t in $\bar{\mathbb{T}}$. Fix t, and note that this is equivalent to showing that the mapping

$$h : \ \omega \mapsto V(\omega) \ \text{ from } \Omega_t = \{T \le t\} \text{ into } \bar{\mathbb{R}}$$

is $\hat{\mathcal{F}}_t$-measurable, where $\hat{\mathcal{F}}_t$ is the trace of $\mathcal{F}_t$ on Ω_t.

Let $B_s = \bar{\mathbb{T}} \cap [0, s]$ for $s \le t$ and let $\mathcal{B}_t = \mathcal{B}(B_t)$. Let f be the mapping $\omega \mapsto (T(\omega), \omega)$ from Ω_t into $B_t \times \Omega$. If $s \in B_t$ and $H \in \mathcal{F}_t$, then the inverse image of the rectangle $B_s \times H$ under f is the event $\{T \le s\} \cap H$, which event is in $\mathcal{F}_t$. Thus f is measurable with respect to $\hat{\mathcal{F}}_t$ and $\mathcal{B}_t \otimes \mathcal{F}_t$.

Let g be the mapping $(s, w) \mapsto X_s(\omega)$ from $B_t \times \Omega$ into $\bar{\mathbb{R}}$. For each s, since X is adapted to $\mathcal{F}$, the mapping $\omega \mapsto X_s(\omega)$ is in $\mathcal{F}_s$ and, therefore, is in $\mathcal{F}_t$; and for each ω, by the way X is chosen, the mapping $s \mapsto X_s(\omega)$ is right-continuous on B_t. Thus, g is $\mathcal{B}_t \otimes \mathcal{F}_t$-measurable (see Exercise I.6.31 for this). It follows that the mapping $g \circ f$ from Ω_t into $\bar{\mathbb{R}}$ is $\hat{\mathcal{F}}_t$-measurable. But, $g \circ f(\omega) = g(T(\omega), \omega) = X_{T(\omega)}(\omega) = V(\omega) = h(\omega)$ for ω in Ω_t. Thus h is $\hat{\mathcal{F}}_t$-measurable as needed to complete the proof. □

1.15 REMARK. *Progressiveness.* The preceding theorem can be rephrased: $V \in \mathcal{F}_T$ if and only if $V = X_T$ for some right-continuous process X adapted to $\mathcal{F}$. This does not exclude the possibility that there is some other process Y, not right-continuous, such that $V = Y_T$ as well. Indeed, the last paragraph of the preceding proof shows what is required of Y: For each t, the mapping $(s, \omega) \mapsto Y_s(\omega)$ from $B_t \times \Omega$ into $\bar{\mathbb{R}}$ should be $\mathcal{B}_t \otimes \mathcal{F}_t$-measurable. Such processes Y are said to be $\mathcal{F}$-*progressive.* So, in fact, $V \in \mathcal{F}_T$ if and only if $V = Y_T$ for some $\mathcal{F}$-progressive process Y. Of course, every right-continuous adapted process is progressive. In discrete time, if $\mathbb{T}$ is discrete, every process is in fact continuous and, hence, every adapted process is progressive.

Comparing different pasts

If S and T are stopping times of $\mathcal{F}$, and if S is dominated by T (that is, $S(\omega) \le T(\omega)$ for all ω), then the information accumulated by the time

S should be less than that accumulated by T. The following shows this and gives further comparisons for general S and T.

1.16 THEOREM. *Let S and T be stopping times of $\mathcal{F}$. Then,*

a) $S \wedge T$ and $S \vee T$ are stopping times of $\mathcal{F}$;
b) if $S \le T$ then $\mathcal{F}_S \subset \mathcal{F}_T$;
c) in general, $\mathcal{F}_{S\wedge T} = \mathcal{F}_S \cap \mathcal{F}_T$; and
d) if $V \in \mathcal{F}_S$ then the following are in $\mathcal{F}_{S\wedge T}$:

$$V\ 1_{\{S\le T\}},\ V\ 1_{\{S=T\}},\ V\ 1_{\{S<T\}}.$$

Proof. i) Since S and T are stopping times, the events $\{S \le t\}$ and $\{T \le t\}$ are in $\mathcal{F}_t$ for every time t. Therefore, so are the events $\{S\wedge T \le t\} = \{S \le t\} \cup \{T \le t\}$ and $\{S \vee T \le t\} = \{S \le t\} \cap \{T \le t\}$. Hence, $S \wedge T$ and $S \vee T$ are stopping times. This proves (a).

ii) Let $V \in \mathcal{F}_S$. By Theorem 1.10,

1.17 $$X_t = V\ 1_{\{S\le t\}}, \quad t \in \bar{\mathbb{T}},$$

defines a process X adapted to $\mathcal{F}$. Clearly, X is right-continuous. Hence, $X \in \mathcal{F}$ in the sense of 1.13.

iii) If $S \le T$, then $X_T = V$ by 1.17, and $X_T \in \mathcal{F}_T$ by Theorem 1.14 since $X \in \mathcal{F}$. So, if $S \le T$ then $\mathcal{F}_S \subset \mathcal{F}_T$. This proves (b).

iv) We prove (d) next. Let the stopping times S and T be arbitrary. Then $S\wedge T$ is a stopping time by part (a), and $X_{S\wedge T} \in \mathcal{F}_{S\wedge T}$ by Theorem 1.14. Hence, replacing t in 1.17 with $S \wedge T$ we see that

1.18 $$V\ 1_{\{S\le T\}} \in \mathcal{F}_{S\wedge T}.$$

In particular, taking $V = 1$ in 1.18 shows that the event $\{S \le T\}$ belongs to $\mathcal{F}_{S\wedge T}$. By symmetry, then, so does the event $\{T \le S\}$. Hence, so do the events $\{S = T\} = \{S \le T\} \cap \{T \le S\}$ and $\{S < T\} = \{S \le T\} \setminus \{S = T\}$. It follows that multiplying the left side of 1.18 with the indicator of $\{S = T\}$ or with the indicator of $\{S < T\}$ will not alter the membership in $\mathcal{F}_{S\wedge T}$. This proves (d).

v) There remains to prove (c) with S and T arbitrary. Since the stopping time $S \wedge T$ is dominated by both S and T, we have $\mathcal{F}_{S\wedge T} \subset \mathcal{F}_S$ and $\mathcal{F}_{S\wedge T} \subset \mathcal{F}_T$ by part (b) proved above. Hence $\mathcal{F}_{S\wedge T} \subset \mathcal{F}_S \cap \mathcal{F}_T$. To prove the converse containment, let H be an event in $\mathcal{F}_S \cap \mathcal{F}_T$. Then, by part (d) proved above, $H \cap \{S \le T\} \in \mathcal{F}_{S\wedge T}$ since H is in $\mathcal{F}_S$, and $H \cap \{T \le S\} \in \mathcal{F}_{S\wedge T}$ since H is in $\mathcal{F}_T$, hence, their union, which is H, belongs to $\mathcal{F}_{S\wedge T}$. So, $\mathcal{F}_S \cap \mathcal{F}_T \subset \mathcal{F}_{S\wedge T}$. □

Times foretold

Let S be a stopping time of $\mathcal{F}$. Let T be a random time such that $T \geq S$ but whose value can be told by the time S, that is, $T \in \mathcal{F}_S$. Then, T is said to be *foretold* by S. Obviously, T is again a stopping time of $\mathcal{F}$. For example, if t is deterministic, then $S + t \in \mathcal{F}_S$ and $S + t \geq S$, so $S + t$ is foretold by S and is a stopping time.

Approximation by discrete stopping times

Discrete stopping times, that is, stopping times that take values in a countable subset of $\bar{\mathbb{R}}$, are generally easier to work with. The following constructs a sequence of such times that approximates a given stopping time with values in $\bar{\mathbb{R}}_+$.

We start by defining, for each integer n in $\mathbb{N}$,

1.19 $$d_n(t) = \begin{cases} \frac{k+1}{2^n} & \text{if } \frac{k}{2^n} \leq t < \frac{k+1}{2^n} \text{ for some } k \text{ in } \mathbb{N}, \\ +\infty & \text{if } t = +\infty. \end{cases}$$

Then, $d_n : \bar{\mathbb{R}}_+ \to \bar{\mathbb{R}}_+$ is a step function, it is increasing and right-continuous, and $d_n(t) > t$ for every $t < \infty$. Further, $d_0 \geq d_1 \geq d_2 \geq \cdots$ with $\lim d_n(t) = t$ for each t in $\bar{\mathbb{R}}_+$.

1.20 PROPOSITION. *Let $\mathcal{F}$ be a filtration on $\bar{\mathbb{R}}_+$ and let T be a stopping time of it. Define*

$$T_n = d_n \circ T, \quad n \in \mathbb{N}.$$

Then (T_n) is a sequence of discrete stopping times of $\mathcal{F}$ which decreases to T.

Proof. Fix n. Being a measurable function of T, the random time T_n belongs to $\mathcal{F}_T$. Since $d_n(t) > t$ for all $t < \infty$ and $d_n(\infty) = \infty$, we have $T_n \geq T$. Thus, T_n is foretold by T and is a stopping time of $\mathcal{F}$. Obviously, it is discrete. Since $d_n(t)$ decreases to t as $n \to \infty$, the sequence (T_n) decreases to T. □

Conditioning at stopping times

This refers to conditional expectations given $\mathcal{F}_T$, where T is a stopping time of the filtration $\mathcal{F}$. Since $\mathcal{F}_T$ represents the total information by the time T, we think of $\mathbb{E}_{\mathcal{F}_T} X = \mathbb{E}(X|\mathcal{F}_T)$ as our estimate of X at time T, based on the information available then. To indicate this point of view, and also to lighten the notation somewhat, we adopt the following notational device:

1.21 CONVENTION. *We write $\mathbb{E}_T$ for $\mathbb{E}_{\mathcal{F}_T} = \mathbb{E}(\cdot|\mathcal{F}_T)$.*

In particular, every deterministic time t is a stopping time, and the notation makes sense for such t as well: $\mathbb{E}_t$ is the short notation for $\mathbb{E}_{\mathcal{F}_t} = \mathbb{E}(\cdot|\mathcal{F}_t)$. The following is a summary, in this context and notation, of the definition and various properties of the conditional expectations given $\mathcal{F}_T$.

1.22 THEOREM. *The following hold for all positive random variables X, Y, Z and all stopping times S and T of $\mathcal{F}$:*

a) Defining property*: $\mathbb{E}_T X = Y$ if and only if $Y \in \mathcal{F}_T$ and $\mathbb{E}\ VX = \mathbb{E}\ VY$ for every positive V in $\mathcal{F}_T$.*

b) Unconditioning*: $\mathbb{E}\ \mathbb{E}_T X = \mathbb{E}\ X$.*

c) Repeated conditioning*: $\mathbb{E}_S \mathbb{E}_T X = \mathbb{E}_{S\wedge T} X$.*

d) Conditional determinism*: $\mathbb{E}_T(X + YZ) = X + Y\mathbb{E}_T Z$ if $X, Y \in \mathcal{F}_T$.*

REMARK. The positivity condition on X, Y, Z ensures that the conditional expectations are well-defined. The properties above can be extended to integrable X, Y, Z and further, once one makes sure that the conditional expectations involved do exist. Of course, in the defining property, V can be limited to indicators in $\mathcal{F}_T$.

Proof. Except for the claim on repeated conditioning, all these are no more than re-wordings of the definition of conditional expectations and Theorem IV.1.10.

To show the claim regarding repeated conditioning, we start with the following observation: If $S \le T$ then $\mathcal{F}_S \subset \mathcal{F}_T$ by Theorem 1.16 above, and Theorem IV.1.10 applies to show that $\mathbb{E}_S \mathbb{E}_T = \mathbb{E}_S$. For arbitrary stopping times S and T, the preceding observation applies with the stopping times $S \wedge T \le T$ to yield $\mathbb{E}_{S\wedge T}\mathbb{E}_T = \mathbb{E}_{S\wedge T}$. Thus, putting

1.23 $$Y = \mathbb{E}_T X,$$

we see that the claim to be proved reduces to showing that

1.24 $$\mathbb{E}_S Y = \mathbb{E}_{S\wedge T} Y.$$

The right side of 1.24 is a random variable in $\mathcal{F}_{S\wedge T}$, and $\mathcal{F}_{S\wedge T} \subset \mathcal{F}_S$ since $S \wedge T \le S$; thus, the right side is in $\mathcal{F}_S$ and, hence, has the required measurability to be a candidate for $\mathbb{E}_S Y$. To complete the proof of 1.24, there remains to show that

1.25 $$\mathbb{E}\ VY = \mathbb{E}\ V\ \mathbb{E}_{S\wedge T} Y$$

for every positive V in $\mathcal{F}_S$ (see the defining property for $\mathbb{E}_S$).

Fix V such. Then, $V\ 1_{\{S\le T\}} \in \mathcal{F}_{S\wedge T}$ by Theorem 1.16d, and the defining property for $\mathbb{E}_{S\wedge T}$ yields

1.26 $$\mathbb{E}\ V\ 1_{\{S\le T\}} Y = \mathbb{E}\ V\ 1_{\{S\le T\}}\ \mathbb{E}_{S\wedge T} Y.$$

On the other hand, since $Y \in \mathcal{F}_T$ by its definition 1.23, Theorem 1.16d shows that $Y\ 1_{\{T<S\}} \in \mathcal{F}_{S\wedge T}$, and the conditional determinism yields

1.27 $$\mathbb{E}\ VY1_{\{T<S\}} = \mathbb{E}\ V\mathbb{E}_{S\wedge T} Y 1_{\{T<S\}} = \mathbb{E}\ V1_{\{T<S\}}\mathbb{E}_{S\wedge T} Y.$$

Adding 1.26 and 1.27 side by side yields the desired equality 1.25. □

Exercises

1.28 *Galmarino's test.* Let X be a continuous stochastic process with index set $\mathbb{R}_+$ and state space $\mathbb{R}$. Let $\mathcal{F}$ be the filtration generated by X. Show that a random time T is a stopping time of $\mathcal{F}$ if and only if, for every pair of outcomes ω and ω',

$$T(\omega) = t, \quad X_s(\omega) = X_s(\omega') \text{ for all } s \le t \Rightarrow T(\omega') = t.$$

1.29 *Entrance times.* Let X and $\mathcal{F}$ be as in 1.28. For fixed $b \ge 0$, let T be the time of first entrance to $[b, \infty]$, that is,

$$T = \inf\{t \in \mathbb{R}_+ : \ X_t \ge b\}.$$

Show that T is a stopping time of $\mathcal{F}$. Show that, in general,

$$T = \inf\{t \in \mathbb{R}_+ : \ X_t > b\}$$

is *not* a stopping time of $\mathcal{F}$.

1.30 *Past until* T. Show that $\mathcal{F}_T$ defined by 1.9 is indeed a σ-algebra on Ω.

1.31 *Strict past at* T. Let T be a stopping time of $\mathcal{F} = (\mathcal{F}_t)_{t\in\mathbb{R}_+}$. Let $\hat{\mathcal{F}}_t$ be the trace of $\mathcal{F}_t$ on $\{t < T\}$, that is, $\hat{\mathcal{F}}_t$ consists of events of the form $H \cap \{t < T\}$ with H in $\mathcal{F}_t$. Let $\mathcal{F}_{T-}$ be the σ-algebra generated by $\cup_t \hat{\mathcal{F}}_t$. Unlike $\mathcal{F}_T$, events in $\mathcal{F}_{T-}$ do not have explicit representations. Show that $\mathcal{F}_{T-} \subset \mathcal{F}_T$.

1.32 *Characterization of* $\mathcal{F}_T$ *in discrete time.* Prove Theorem 1.14 directly when $\mathbb{T} = \mathbb{N}$. Hints: $V \in \mathcal{F}_T \iff V 1_{\{T=n\}} \in \mathcal{F}_n$ for every n in $\mathbb{N}$; and if $X \in \mathcal{F}$ then $X_T 1_{\{T=n\}} = X_n 1_{\{T=n\}}$.

1.33 *Stopping times foretold.* Let S and T be stopping times. Show that $S+T$ is foretold by $S \vee T$ and thus is a stopping time.

1.34 *Supremums.* Let T_n be a stopping time for each n in $\mathbb{N}^*$. Show that, then, $\sup T_n$ is again a stopping time of $\mathcal{F}$. A similar claim for $\inf T_n$ is generally false; see, however, Proposition 7.9.

1.35 *Arrival processes.* In Example 1.7, observe that, for every $t < \infty$, we have $T_k(\omega) \le t < T_{k+1}(\omega)$ for some integer k depending on t and ω. Recall that every T_k is a stopping time of $\mathcal{F}$, the filtration generated by N. Put $T_0 = 0$ for convenience. Let T be as defined in 1.7. Note that, for every ω, $T(\omega) = T_k(\omega) + a$ for some k.

a) Show that, for each k in $\mathbb{N}$,

$$\{T = T_k + a\} = \{T_1 - T_0 \le a, \ldots, T_k - T_{k-1} \le a\} \cap \{T_{k+1} > T_k + a\}.$$

Show that this event is in $\mathcal{F}_{T_k+a}$. Conclude that, for every t in $\mathbb{R}_+$,

$$\{T = T_k + a\} \cap \{T \le t\} \in \mathcal{F}_t.$$

b) Show that T is a stopping time of $\mathcal{F}$.

1.36 *Continuation.* Now we regard $0 = T_0 < T_1 < T_2 < \cdots$ as the times of successive replacements for some device. Then N_t becomes the number of replacements during $(0, t]$, and we define A_t to be the age of the unit in use at time t:

$$A_t(\omega) = t - T_k(\omega) \qquad \text{if } T_k(\omega) \le t < T_{k+1}(\omega).$$

a) Show that $t \mapsto A_t(\omega)$ is strictly increasing and continuous everywhere on $\mathbb{R}_+$ except for downward jumps to 0 at times $T_1(\omega), T_2(\omega), \ldots$. At these times, it is right-continuous.

b) Show that the process $A = (A_t)_{t\in\mathbb{R}_+}$ is adapted to $\mathcal{F}$.

c) Show that $T = \inf\{t \in \mathbb{R}_+ : \ A_t \ge a\}$, and show that T is a stopping time of $\mathcal{F}$ by a direct reasoning using this relationship to A.

2 Martingales

Martingales are the mainstay and unifying force underlying much of the theory of stochastic processes. This section is to introduce them and give some examples from Markov chains, Brownian motion, and Poisson processes.

Let $\mathbb{T}$ be a subset of $\bar{\mathbb{R}}$, let $\mathcal{F} = (\mathcal{F}_t)_{t\in\mathbb{T}}$ be a filtration over $\mathbb{T}$ extended onto $\bar{\mathbb{T}} = \mathbb{T} \cup \{+\infty\}$ by 1.8 if $+\infty$ is not in $\mathbb{T}$, and recall the notational convention 1.21 regarding conditional expectations given $\mathcal{F}_T$.

2.1 Definition. *A real-valued stochastic process $X = (X_t)_{t\in\mathbb{T}}$ is called an* $\mathcal{F}$-submartingale *if X is adapted to $\mathcal{F}$, each X_t is integrable, and*

2.2 $$\mathbb{E}_s(X_t - X_s) \ge 0$$

whenever $s < t$. It is called an $\mathcal{F}$-supermartingale *if $-X$ is an $\mathcal{F}$-submartingale, and an $\mathcal{F}$-martingale if it is both an $\mathcal{F}$-submartingale and an $\mathcal{F}$-supermartingale.*

Adaptedness and integrability are regularity conditions; they remain the same for submartingales, supermartingales, and martingales. The essential condition is 2.2: Given the information $\mathcal{F}_s$, the conditional expectation of the future increment $X_t - X_s$ is positive for submartingales, negative for supermartingales, and zero for martingales.

Indeed, since $X_s \in \mathcal{F}_s$, the conditional determinism property yields $\mathbb{E}_s X_s = X_s$, which shows that the parentheses around $X_t - X_s$ are superfluous; they are put there to make us think in terms of the increments. So, 2.2 can be re-written as

2.3 $$\mathbb{E}_s X_t \ge X_s, \qquad s < t\ ;$$

this is for submartingales. The inequality is reversed for supermartingales and becomes an equality for martingales. Thus, roughly speaking, submartingales have a systematic tendency to be increasing, supermartingales to be decreasing, and martingales to be neither increasing nor decreasing. See Theorem 3.2 below for a sharper version of this remark.

2.4 Remarks. a) Let X be an $\mathcal{F}$-submartingale. For $s < t < u$ in $\mathbb{T}$,

$$\mathbb{E}_s(X_u - X_t) = \mathbb{E}_s\mathbb{E}_t(X_u - X_t) \geq \mathbb{E}_s 0 = 0$$

by 1.22c on repeated conditioning and the submartingale inequality $\mathbb{E}_t(X_u - X_t) \geq 0$. That is, given the cumulative information $\mathcal{F}_s$ available at the present time s, the estimate of any remote future increment is positive. Obviously, if X is a martingale, the inequality becomes an equality.

b) When the index set $\mathbb{T}$ is discrete, the reasoning of the preceding remark shows that it is sufficient to check the inequality 2.2 for times s and t that are next to each other, and then 2.2 holds for arbitrary $t > s$. For instance, when $\mathbb{T} = \mathbb{N}$, the martingale equality $\mathbb{E}_s(X_t - X_s) = 0$ holds if and only if

$$\mathbb{E}_n(X_{n+1} - X_n) = 0, \qquad n \in \mathbb{N}.$$

c) Let X be an $\mathcal{F}$-submartingale. For $s < t$, the random variable $\mathbb{E}_s(X_t - X_s)$ is positive and, therefore, is almost surely zero if and only if its expectation $\mathbb{E}\ \mathbb{E}_s(X_t - X_s)$ is zero. Since $\mathbb{E}\ \mathbb{E}_s = \mathbb{E}$, it follows that the submartingale X is in fact a martingale if $\mathbb{E}X_t = \mathbb{E}X_0$ for all times t.

d) If X and Y are $\mathcal{F}$-submartingales, then so is $aX + bY$ for a and b in $\mathbb{R}_+$. If X and Y are martingales, then so is $aX + bY$ for a and b in $\mathbb{R}$.

e) If X and Y are $\mathcal{F}$-submartingales, then so is $X \vee Y$, where $X \vee Y = (X_t \vee Y_t)_{t \in \mathbb{T}}$. If X and Y are $\mathcal{F}$-supermartingales, then so is $X \wedge Y$.

f) Let f be a convex function on $\mathbb{R}$. If X is an $\mathcal{F}$-martingale and if $f \circ X_t$ is integrable for every time t, then $f \circ X$ is an $\mathcal{F}$-submartingale. This follows from Jensen's inequality for conditional expectations (see IV.1.8): for $s < t$,

$$\mathbb{E}_s\ f \circ X_t \geq f \circ (\mathbb{E}_s X_t) = f \circ X_s$$

since $\mathbb{E}_s X_t = X_s$ for martingales. In particular, if X is a martingale, then $X^+ = (X_t^+)$ and $X^- = (X_t^-)$ and $|X| = (|X_t|)$ are submartingales, and so is $|X|^p = (|X_t|^p)$ provided that $\mathbb{E}|X_t|^p < \infty$ for every time t.

g) Similarly, if f is convex and increasing, and if X is an $\mathcal{F}$-submartingale with $f \circ X_t$ integrable for all t, then $f \circ X$ is again an $\mathcal{F}$-submartingale. In particular, if X is a submartingale, so is X^+.

h) Since $\mathbb{E}_s(X_t - X_s)$ belongs to $\mathcal{F}_s$, it is positive if and only if its integral over every event H in $\mathcal{F}_s$ is positive. Thus, the submartingale inequality 2.2 is equivalent to the following:

$$\mathbb{E}(X_t - X_s)\ 1_H \geq 0, \qquad H \in \mathcal{F}_s, s < t.$$

i) Let X be an $\mathcal{F}$-submartingale. Let $\mathcal{G}$ be the filtration generated by X. Then, X is automatically adapted to $\mathcal{G}$ and is integrable, and

$$\mathbb{E}_{\mathcal{G}_s}(X_t - X_s) = \mathbb{E}_{\mathcal{G}_s}\mathbb{E}_{\mathcal{F}_s}(X_t - X_s) = \mathbb{E}_{\mathcal{G}_s}\mathbb{E}_s(X_t - X_s) \geq 0$$

by the repeated conditioning property since $\mathcal{G}_s \subset \mathcal{F}_s$. Thus, X is a $\mathcal{G}$-submartingale.

Examples of martingales

2.5 *Sums of independent variables.* Let $X_1, X_2, \ldots$ be independent random variables with mean 0. Let $S_0 = 0$ and put $S_n = S_0 + X_1 + \cdots + X_n$ for $n \geq 1$. Let $\mathcal{F} = (\mathcal{F}_n)_{n\in\mathbb{N}}$ be the filtration generated by $S = (S_n)_{n\in\mathbb{N}}$. Then S is adapted to $\mathcal{F}$ trivially, and each S_n is integrable (with mean 0), and

$$\mathbb{E}_n(S_{n+1} - S_n) = \mathbb{E}_n X_{n+1} = \mathbb{E}\ X_{n+1} = 0,$$

since X_{n+1} is independent of $\mathcal{F}_n$ and has mean 0. Thus, S is a martingale; see Remark 2.4b. Much of classical martingale theory is an extension of this case.

2.6 *Products of independent variables.* Let $R_1, R_2, \ldots$ be independent random variables with mean 1 and some finite variance. Let $M_0 = 1$ and

$$M_n = M_0 R_1 R_2 \cdots R_n, \qquad n \in \mathbb{N}.$$

Let $\mathcal{F}$ be the filtration generated by $M = (M_n)_{n\in\mathbb{N}}$. Then, M is adapted to $\mathcal{F}$ trivially, and each M_n is integrable in view of Schwartz's inequality (see Theorem II.3.6a) and the assumption that the R_n have finite variances. Also,

$$\mathbb{E}_n M_{n+1} = \mathbb{E}_n M_n R_{n+1} = M_n \mathbb{E}_n R_{n+1} = M_n$$

by the independence of R_{n+1} from $\mathcal{F}_n$ and the hypothesis that $\mathbb{E}\ R_{n+1} = 1$. Hence, M is an $\mathcal{F}$-martingale via Remarks 2.3 and 2.4b.

In the further case where the R_n are positive, the martingale M is considered to be a reasonable model for the evolution of the price of a share of stock. Then, M_n stands for the price of a share at time n, and R_{n+1} is interpreted as the return at time $n+1$ per dollar invested at time n in that stock. The economists' argument for the martingale equality is as follows (very roughly): The information $\mathcal{F}_n$ is available to the whole market. If the conditional expectation $\mathbb{E}_n(M_{n+1} - M_n)$ were strictly positive over some event H, then there would have been a rush to buy which would have forced M_n to go higher; if the expectation were strictly negative over some event, then there would have been a rush to sell and M_n would go lower; the equilibrium attains only if the conditional expectation is zero over an almost sure set.

Uniformly integrable martingales

These are martingales that are also uniformly integrable. They play the central role in martingale theory. The next proposition shows how to obtain such a martingale: take an integrable random variable and let X_t be our estimate of it at time t, given the information $\mathcal{F}_t$ accumulated until then. Conversely, it will be shown in Theorems 4.7 and 5.13 that every uniformly integrable martingale is obtained in this manner. Here, $\mathbb{T} \subset \bar{\mathbb{R}}$ is arbitrary and $\mathcal{F} = (\mathcal{F}_t)_{t\in\mathbb{T}}$ is a filtration on $\mathbb{T}$; see Definition II.3.12 *et seq.* for uniform integrability.

2.7 PROPOSITION. *Let Z be an integrable random variable. Define*

$$X_t = \mathbb{E}_t Z, \qquad t \in \mathbb{T}.$$

Then $X = (X_t)_{t\in\mathbb{T}}$ is an $\mathcal{F}$-martingale and is uniformly integrable.

Proof. Adaptedness is immediate, since $\mathbb{E}_t Z$ is in $\mathcal{F}_t$ by the definition of conditional expectations. Each X_t is integrable, because Z is so and the conditional expectation of an integrable variable is integrable. The martingale equality follows from the properties of repeated conditioning: for times $s < t$,

$$\mathbb{E}_s X_t = \mathbb{E}_s \mathbb{E}_t Z = \mathbb{E}_s Z = X_s.$$

Finally, the uniform integrability of the collection (X_t) follows from the following more general result of independent interest. □

2.8 LEMMA. *Let Z be an integrable random variable. Then,*

$$\mathcal{K} = \{X : X = \mathbb{E}_{\mathcal{G}} Z \textit{ for some sub-}\sigma\textit{-algebra } \mathcal{G} \textit{ of } \mathcal{H}\}$$

is uniformly integrable.

Proof. Since Z is integrable, the singleton $\{Z\}$ is uniformly integrable. Thus, by Theorem II.3.19, there is an increasing convex function f with $\lim_{x\to\infty} f(x)/x = +\infty$ such that $\mathbb{E}\ f\circ|Z| < \infty$. We show next that, with the same f,

2.9 $$\mathbb{E}\ f\circ|X| \le \mathbb{E}\ f\circ|Z|$$

for every X in $\mathcal{K}$, which implies, via Theorem II.3.19 again, that $\mathcal{K}$ is uniformly integrable.

Let $X = \mathbb{E}_{\mathcal{G}} Z$ for some sub-σ-algebra $\mathcal{G}$ of $\mathcal{H}$. Then, by Jensen's inequality IV.1.8,

$$|X| = |\mathbb{E}_{\mathcal{G}} Z| \le \mathbb{E}_{\mathcal{G}} |Z|.$$

Thus, since f is increasing and convex,

$$f\circ|X| \le f\circ(\mathbb{E}_{\mathcal{G}}|Z|) \le \mathbb{E}_{\mathcal{G}} f\circ|Z|,$$

where the last inequality is Jensen's again. Now, taking expectations on both sides and recalling that $\mathbb{E}\ \mathbb{E}_{\mathcal{G}} = \mathbb{E}$, we obtain the desired end 2.9. □

Markov chains

Here, the index set is $\mathbb{N}$, and $\mathcal{F}$ is a filtration over $\mathbb{N}$. Let $X = (X_n)_{n\in\mathbb{N}}$ be a stochastic process with state space $(E, \mathcal{E})$, and let P be a Markov kernel on $(E, \mathcal{E})$; see section I.6 for the latter and recall the notation

2.10 $$Pf(x) = \int_E P(x, dy)\ f(y), \qquad x \in E, f \in \mathcal{E}_+.$$

2.11 Definition. *The process X is called a Markov chain with transition kernel P, with respect to $\mathcal{F}$, if X is adapted to $\mathcal{F}$ and*

2.12 $$\mathbb{E}_n\ f\circ X_{n+1} = (Pf)\circ X_n$$

for every function f in $\mathcal{E}_+$ and time n in $\mathbb{N}$.

Markov chains have an extensive theory; see also Chapter IV, Section 5, for many variations. Much of their theory (and the theory of their continuous-time counterparts) has been influenced strongly by its connections to classical potential theory. As a result, harmonic and subharmonic and superharmonic functions of the classical theory have found their counterparts for Markov processes and, through Markov processes, have influenced the definitions of martingales and submartingales and supermartingales. Here is the connection.

Let X be a Markov chain, with respect to some filtration $\mathcal{F}$, with state space $(E, \mathcal{E})$ and transition kernel P. A bounded function f in $\mathcal{E}$ is said to be *harmonic*, *subharmonic*, and *superharmonic* if, respectively,

2.13 $$f = Pf, \quad f \le Pf, \quad f \ge Pf.$$

Put $M_n = f\circ X_n$; it is integrable since f is bounded, and it is obviously in $\mathcal{F}_n$. Indeed, $M = (M_n)_{n\in\mathbb{N}}$ is a martingale if f is harmonic, a submartingale if f is subharmonic, and a supermartingale if f is superharmonic. Here is the proof of the supermartingale inequality assuming that f is superharmonic; the other two cases can be shown similarly.

$$\mathbb{E}_n M_{n+1} = \mathbb{E}_n f\circ X_{n+1} = (Pf)\circ X_n \le f\circ X_n = M_n,$$

where we used the Markov property 2.12 to justify the second equality and the superharmonicity ($Pf \le f$) to justify the inequality.

A more recent connection is the following characterization of Markov chains in terms of martingales; this becomes a deep result in continuous-time.

2.14 Theorem. *Let X be adapted to $\mathcal{F}$. Then X is a Markov chain with transition kernel P with respect to $\mathcal{F}$ if and only if*

$$M_n = f\circ X_n - \sum_{m=0}^{n-1} (Pf - f)\circ X_m, \qquad n \in \mathbb{N},$$

is a martingale with respect to $\mathcal{F}$ for every bounded f in $\mathcal{E}_+$.

Proof. Note that

$$M_{n+1} - M_n = f\circ X_{n+1} - (Pf)\circ X_n;$$

thus, $\mathbb{E}_n M_{n+1} - M_n = 0$ if and only if X has the Markov property 2.12. □

Wiener Processes

Let $\mathcal{F}$ be a filtration over $\mathbb{R}_+$. Let $W = (W_t)_{t\in\mathbb{R}_+}$ be a continuous process with state space $(\mathbb{R}, \mathcal{B}_\mathbb{R})$ and starting point $W_0 = 0$.

2.15 Definition. *The continuous process W is called a Wiener process with respect to $\mathcal{F}$ if it is adapted to $\mathcal{F}$ and*

2.16 $$\mathbb{E}_s\ f(W_{s+t} - W_s) = \int_\mathbb{R} dx \frac{1}{\sqrt{2\pi t}} e^{-x^2/2t} f(x)$$

for all s and t in $\mathbb{R}_+$ and all positive Borel functions f on $\mathbb{R}$.

The defining relation 2.16 has three statements in it: the increment $W_{s+t} - W_s$ over the interval $(s, s+t]$ is independent of the past $\mathcal{F}_s$, the distribution of that increment is free of s, and the distribution is Gaussian with mean 0 and variance t. Indeed, 2.16 defines the probability law of W uniquely: for $0 = t_0 < t_1 < \cdots < t_n$, the probability law of $(W_{t_1}, \ldots, W_{t_n})$ is determined uniquely by the probability law of $(W_{t_1} - W_{t_0}, \ldots, W_{t_n} - W_{t_{n-1}})$, and the latter is the product of the distributions of $W_{t_1} - W_{t_0}, \ldots, W_{t_n} - W_{t_{n-1}}$ by the independence of the increments, and the distributions are further specified by 2.16 as Gaussian with mean 0 and respective variances $t_1 - t_0, \ldots, t_n - t_{n-1}$. Incidentally, we see that W has stationary and independent increments (stationarity refers to the invariance of the distribution of $W_{s+t} - W_s$ as s varies). We shall study Wiener processes in Chapter VIII. Our aim at present is to introduce three martingales related to W. First is a useful characterization.

2.17 Proposition. *The process W is a Wiener process with respect to $\mathcal{F}$ if and only if, for each r in $\mathbb{R}$,*

$$M_t = \exp(rW_t - \tfrac{1}{2}r^2 t), \qquad t \in \mathbb{R}_+,$$

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

Proof. *Necessity.* Suppose that W is Wiener. Then, a direct computation using 2.16 shows that, for $s < t$,

2.18 $$\mathbb{E}_s(M_t/M_s) = \mathbb{E}_s \exp[r(W_t - W_s) - \tfrac{1}{2}r^2(t-s)] = 1.$$

Thus, $\mathbb{E}_s M_t = M_s \mathbb{E}_s(M_t/M_s) = M_s$, which shows that M is a martingale (adaptedness and integrability being obvious).

Sufficiency. If M is a martingale, then $\mathbb{E}_s(M_t/M_s) = 1$, which means that 2.18 holds, or equivalently,

$$\mathbb{E}_s\ \exp\ r(W_{s+t} - W_s) = \exp\ \tfrac{1}{2}r^2 t.$$

This, being true for all r in $\mathbb{R}$, is equivalent to 2.16. □

It is worth noting that, for fixed r in $\mathbb{R}$, the process M of the preceding theorem is a continuous-time version of Example 2.6. Indeed, $M_{n+1} = M_n R_{n+1}$ where the random variable R_{n+1} now has a very specific distribution, namely, the distribution of $\exp[r(W_{n+1} - W_n - \frac{1}{2}r^2]$. Thus, the *exponential martingale* M is much used as a model for the evolution of stock prices. It is also the primary tool for studying Brownian motions by using results from martingale theory; see 5.20 *et seq.*

The next theorem gives the *martingale characterization of Wiener process.* We are able to prove here only the easy part, the necessity. For a proof of the sufficiency, see 6.21 to come.

2.19 THEOREM. *The continuous process W is Wiener with respect to $\mathcal{F}$ if and only if*

a) *W is an $\mathcal{F}$-martingale, and*
b) *$Y = (W_t^2 - t)_{t\in\mathbb{R}_+}$ is an $\mathcal{F}$-martingale.*

Proof of necessity. Let W be Wiener. Then, adaptedness and integrability conditions are obvious for W and Y. Now, the martingale equality for W is straightforward: for $s < t$, the increment $W_t - W_s$ is independent of $\mathcal{F}_s$ and has mean 0; thus,

$$\mathbb{E}_s\ (W_t - W_s) = \mathbb{E}\ (W_t - W_s) = 0.$$

To show the martingale equality for Y, we first note that

$$Y_t - Y_s = (W_t - W_s)^2 + 2W_s(W_t - W_s) - (t - s)$$

and then use the facts that $W_s \in \mathcal{F}_s$ and that $W_t - W_s$ is independent of $\mathcal{F}_s$ and has mean 0 and variance $t - s$. Thus, as needed,

$$\mathbb{E}_s\ (Y_t - Y_s) = \mathbb{E}\ (W_t - W_s)^2 + 2W_s\ \mathbb{E}(W_t - W_s) - (t - s) = 0. \qquad \square$$

The Wiener process is the continuous martingale par excellence. It plays the same role in stochastic analysis as the Lebesgue measure does in ordinary analysis. In particular, every continuous martingale (in continuous-time) is obtained from a Wiener process by a random time change, just as most measures on $\mathbb{R}$ are obtained from the Lebesgue measure by a time change (see Theorem I.5.4).

Poisson martingales

Saying that a process is a martingale amounts to stating a property of it without specifying its probability law. However, on rare occasions, martingale property specifies the probability law. We stated, without proof, one such case: if W is a continuous martingale and if $W^2 - t$ is a martingale, then W is a Wiener process. Here, we provide another such case, even sharper, this time a pure-jump process.

Here, the index set is $\mathbb{R}_+$, and $\mathcal{F}$ is a filtration over it. Let $N = (N_t)_{t\in\mathbb{R}_+}$ be a counting process: this is a process with state space $(\mathbb{N}, 2^{\mathbb{N}})$ whose every path $t \mapsto N_t(\omega)$ starts from $N_0(\omega) = 0$, is increasing and right-continuous, and increases by jumps of size one only. Therefore, $N_t(\omega)$ is equal to the number of jumps of $s \mapsto N_s(\omega)$ in the interval $(0, t]$; Example 1.7 provides the complete picture. The following definition parallels that of the Wiener processes, Definition 2.15.

2.20 Definition. *The counting process N is said to be a Poisson process with rate c with respect to $\mathcal{F}$ if it is adapted to $\mathcal{F}$ and*

2.21
$$\mathbb{E}_s\, f(N_{s+t} - N_s) = \sum_{k=0}^{\infty} \frac{e^{-ct}(ct)^k}{k!} f(k)$$

for all s and t in $\mathbb{R}_+$ and all positive functions f on $\mathbb{N}$.

The defining equation 2.21 is equivalent to saying that the increment $N_{s+t} - N_s$ is independent of $\mathcal{F}_s$ and has the Poisson distribution with mean ct. As with Wiener processes, then, N has stationary and independent increments, and its probability law is completely determined by the positive constant c. Just as W is a martingale, for the Poisson process N, we have that $M = (N_t - ct)_{t\in\mathbb{R}_+}$ is an $\mathcal{F}$-martingale; this is immediate from 2.21:

2.22
$$\mathbb{E}_s(N_{s+t} - N_s) = ct, \qquad s, t \in \mathbb{R}_+.$$

It is surprising that, as the next theorem states, the simple property 2.22 is equivalent to 2.21. This is the *martingale characterization* theorem for Poisson processes; it parallels Theorem 2.19 and is even sharper.

2.23 Theorem. *Let N be a counting process. It is a Poisson process with rate c, with respect to $\mathcal{F}$, if and only if*

$$M_t = N_t - ct, \qquad t \in \mathbb{R}_+,$$

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

The proof will be given in Section 6; see Proposition 6.13 and its proof.

Exercises and complements

2.24 *Restrictions.* Let $\mathbb{T}_0 \subset \bar{\mathbb{R}}$, and let $(X_t)_{t\in\mathbb{T}_0}$ be a martingale with respect to a filtration $(\mathcal{F}_t)_{t\in\mathbb{T}_0}$. Then, for every $\mathbb{T}_1 \subset \mathbb{T}_0$, the process $(X_t)_{t\in\mathbb{T}_1}$ is a martingale with respect to the filtration $(\mathcal{F}_t)_{t\in\mathbb{T}_1}$. The word “martingale” can be replaced with “submartingale” or with “supermartingale”.

2.25 *Markov chains.* Let $X = (X_n)$ be a Markov chain with state space $(E, \mathcal{E})$ and transition kernel P. With P^n denoting the n^{th} power of P–see I.6.6 for the definition–show that

$$\mathbb{E}_m\, f\circ X_{m+n} = (P^n f)\circ X_m, \qquad m, n \in \mathbb{N},$$

for every f in $\mathcal{E}_+$. Show that, for each fixed integer $k \geq 1$, $M_n = (P^{k-n}f)\circ X_n$ defines a martingale on $\mathbb{T} = \{0, 1, \ldots, k\}$.

2.26 *Poisson processes.* Let $N = (N_t)_{t\in\mathbb{R}_+}$ be a counting process adapted to some filtration $\mathcal{F} = (\mathcal{F}_t)_{t\in\mathbb{R}_+}$. Prove the following characterization theorem (see Proposition 2.17 for the parallel fo Wiener processes): N is a Poisson process with rate c with respect to $\mathcal{F}$ if and only if

$$M_t = \exp\ (-rN_t + ct - cte^{-r}), \qquad t \in \mathbb{R}_+,$$

is an $\mathcal{F}$-martingale for every r in $\mathbb{R}_+$.

2.27 *Averages.* Let (X_n) be adapted to some filtration $(\mathcal{F}_n)$ and suppose that each X_n is integrable. Define

$$\bar{X}_n = \frac{1}{n+1}(X_0 + \cdots + X_n), \quad n \in \mathbb{N},$$

and assume that $\mathbb{E}_n X_{n+1} = \bar{X}_n$ for all n. Show that $(\bar{X}_n)$ is an $\mathcal{F}$-martingale.

2.28 *Positive supermartingales.* Let (X_n) be a positive supermartingale with respect to some filtration $(\mathcal{F}_n)$. Then, the following holds for almost every ω: if $X_m(\omega) = 0$ for some m, then $X_n(\omega) = 0$ for all $n \geq m$. Show this. Hint: Let $H = \{X_m = 0\}$ and show that $\mathbb{E}_m\ 1_H X_n = 0$ for $n \geq m$.

2.29 *Uniform integrability.* Let Z be an integrable random variable. Let $\mathcal{F} = (\mathcal{F}_t)_{t\in\bar{\mathbb{R}}_+}$ be a filtration on $\bar{\mathbb{R}}_+$. For each stopping time T of $\mathcal{F}$, let

$$X_T = \mathbb{E}_T\ Z,$$

Show that the collection $\{X_T :\ T \text{ is a stopping time of } \mathcal{F}\}$ is uniformly integrable.

2.30 *Martingales in L^p.* For p in $[1, \infty]$, a process X is said to be a martingale in L^p if, in addition to adaptedness and the martingale equality, the integrability condition for X_t is strengthened to requiring that $X_t \in L^p$ for every time t. Note that martingales in L^1 are simply martingales. Submartingales and supermartingales in L^p are defined similarly by replacing the condition $X_t \in L^1$ with the stronger condition that $X_t \in L^p$.

2.31 *L^p-boundedness.* A process (X_t) is said to be L^p-bounded if

$$\sup_t \mathbb{E}\ |X_t|^p < \infty.$$

With the notation $\|\cdot\|_p$ for the L^p-norm, the condition means that $\|X_t\|_p \leq c$ for some constant $c < \infty$. Recall: uniform integrability implies L^1-boundedness; the converse is generally false; but L^p-boundedness for some $p > 1$ implies uniform integrability.

2.32 *Square integrable martingales.* These are martingales that are L^2-bounded. This somewhat misleading usage seems well established.

2.33 *Quadratic variation.* Let (M_n) be a martingale in L^2 adapted to some filtration $(\mathcal{F}_n)$. Define an increasing process by setting $Q_0 = 0$ and

$$Q_{n+1} - Q_n = (M_{n+1} - M_n)^2, \qquad n \geq 0.$$

Show that the process X defined by

$$M_n^2 = M_0^2 + X_n + Q_n, \qquad n \geq 0,$$

is a martingale with $X_0 = 0$. Show that

$$\text{Var}M_n = \text{Var}M_0 + \mathbb{E}Q_n = \text{Var}M_0 + \text{Var}(M_1 - M_0) + \cdots + \text{Var}(M_n - M_{n-1}).$$

The process Q is called the quadratic variation process for M.

3 Martingale Transformations and Maxima

This section contains the basic results for martingales in discrete time: integration in discrete time, Doob's stopping theorem, and inequalities for upcrossings and maxima. The index set is $\mathbb{N}$ unless stated otherwise; $\mathcal{F}$ is a filtration which we keep in the background; Convention 1.21 regarding conditional expectations is in force throughout; and all martingales, stopping times, and so on are with respect to the filtration $\mathcal{F}$.

Doob's decomposition

The object is to write a given process as the sum of a martingale and a predictable process, the latter to be defined presently. In the case of a submartingale, its predictable part turns out to be increasing, which clarifies our earlier statement that submartingales have a systematic tendency to be increasing.

3.1 Definition. *A process $F = (F_n)_{n\in\mathbb{N}}$ is said to be $\mathcal{F}$-predictable if $F_0 \in \mathcal{F}_0$ and $F_{n+1} \in \mathcal{F}_n$ for every n in $\mathbb{N}$.*

Heuristically, then, the cumulative information $\mathcal{F}_n$ available at time n determines the next value F_{n+1}, and thus, F is predictable in this dynamic sense. Note that every predictable process is adapted and more. The following is Doob's decomposition.

3.2 Theorem. *Let X be adapted and integrable. Then, it can be decomposed as*

3.3 $$X_n = X_0 + M_n + A_n, \qquad n \in \mathbb{N},$$

where M is a martingale with $M_0 = 0$, and A is predictable with $A_0 = 0$. This decomposition is unique up to equivalence. In particular, A is increasing if X is a submartingale, and decreasing if X is a supermartingale.

Proof. a) Put $M_0 = A_0 = 0$ and define M and A through their increments:

$$A_{n+1} - A_n = \mathbb{E}_n\,(X_{n+1} - X_n), \qquad M_{n+1} - M_n = (X_{n+1} - X_n) - (A_{n+1} - A_n)$$

for each $n \in \mathbb{N}$. Then, 3.3 holds, M is obviously a martingale, and A is predictable by the $\mathcal{F}_n$-measurability of the conditional expectation $\mathbb{E}_n\,(X_{n+1} - X_n)$. This proves the first statement.

b) If X is a submartingale, then 2.2 shows that $A_{n+1} - A_n \geq 0$, that is, A is increasing. If X is a supermartingale, then the inequality is reversed, and A is decreasing.

c) There remains to show the statement on uniqueness. To that end, let $X = X_0 + M' + A'$ be another such decomposition. Then $B = A - A' = M' - M$ is both predictable and a martingale. Thus,

$$B_{n+1} - B_n = \mathbb{E}_n\,(B_{n+1} - B_n) = 0, \qquad n \in \mathbb{N}\ ;$$

in other words, almost surely, $B_n = B_0 = 0$. Hence, almost surely, $A = A'$ and $M = M'$, as claimed. □

In Doob's decomposition, we have $X_{n+1} - X_n = A_{n+1} - A_n + M_{n+1} - M_n$; of these, $A_{n+1} - A_n$ is known by the time n; thus, the extra information gained by observing $X_{n+1} - X_n$ consists of the martingale increment $M_{n+1} - M_n$. For this reason, in engineering literature, A is called the prediction process, and M the *innovation* process.

Integration in discrete time

This is a resume of stochastic integration in discrete time. Let $M = (M_n)$ and $F = (F_n)$ be real-valued stochastic processes and define

3.4 $$X_n = M_0 F_0 + (M_1 - M_0) F_1 + \cdots + (M_n - M_{n-1}) F_n, \qquad n \in \mathbb{N}.$$

Then, $X = (X_n)$ is called the integral of F with respect to M, or the *transform* of M by F, and we shall write

3.5 $$X = \int F\, dM$$

to indicate it. Indeed, F is a random function on $\mathbb{N}$, and M defines a random signed-measure on $\mathbb{N}$ which puts the mass $M_n - M_{n-1}$ at n except that the mass is M_0 at $n = 0$; then 3.4 is equivalent to writing

$$X_n = \int_{[0,n]} F\, dM,$$

the Lebesgue-Stieltjes integral of F over $[0, n]$ with respect to M. Such integrals are harder to define in continuous time, because Lebesgue-Stieltjes integrals make sense for M that are of bounded variation over bounded intervals, whereas most continuous martingales (including Wiener processes) have infinite variation over every open interval. Here we are working with the straightforward case of discrete time.

3.6 Theorem. *Let F be a bounded predictable process and let $X = \int F\, dM$. If M is a martingale, then so is X. If M is a submartingale and F is positive, then X is a submartingale.*

Proof. Suppose that M is a martingale. Since $F_0, \ldots, F_n$ and $M_0, \ldots, M_n$ are in $\mathcal{F}_n$, so is X_n; that is, X is adapted to $\mathcal{F}$. Since F is bounded, say by some constant $b > 0$, we see that $|X_n|$ is bounded by b times $|M_0| + |M_1 - M_0| + \cdots + |M_n - M_{n-1}|$, which is integrable. So X is integrable. Finally,

$$\begin{aligned}\mathbb{E}_n\ (X_{n+1} - X_n) &= \mathbb{E}_n\ (M_{n+1} - M_n)F_{n+1} \\ &= F_{n+1}\ \mathbb{E}_n\ (M_{n+1} - M_n) = F_{n+1} \cdot 0 = 0,\end{aligned}$$

where the second equality uses the predictability of F to move F_{n+1} outside the conditional expectation $\mathbb{E}_n$, and the third equality is merely the martingale equality for M. Hence, X is a martingale. If M is a submartingale and F is positive, the third and fourth equalities become $\geq$, and X is a submartingale. □

3.7 Heuristics. Here is an interpretation of the preceding theorem. A person buys and sells shares of a certain stock, presumably to make a profit. Let M_n be the price of a share at time n, and let F_n denote the number of shares owned during the time interval $(n-1, n]$. Then, the profit made during $(n-1, n]$ is $(M_n - M_{n-1}) \cdot F_n$. Hence, in 3.4, X_n is the sum of the initial value $X_0 = M_0 F_0$ and the total profit made during $(0, n]$. Since the decision on how many shares to own during $(n, n+1]$ must be made at time n based on the information $\mathcal{F}_n$ available at that time, it follows that F_{n+1} be $\mathcal{F}_n$-measurable, that is, F be predictable. For reasons mentioned in Example 2.6, the price process M should be a martingale. Then, the preceding theorem shows that it is impossible to make a profit systematically (or to lose systematically); no matter what "strategy" F one uses, X has no systematic tendency to move up or down.

Predictability

To enhance the value of the preceding theorem, the following provides some examples of predictable processes. Other examples may be constructed by noting that the class of predictable processes form a linear space that is closed under all limits.

3.8 Example. Let S and T be stopping times of $\mathcal{F}$ with $S \leq T$. Let V be a random variable in $\mathcal{F}_S$. Then,

$$V\ 1_{(S,T]}, \quad V\ 1_{(S,\infty]}, \quad 1_{(S,T]}, \quad 1_{[0,T]}$$

are all predictable processes. To see these, we start with the second, with $F = V\ 1_{(S,\infty]}$, that is, $F_n = V \cdot 1_{(S,\infty]}(n)$. Note that,

$$F_{n+1} = V\ 1_{\{S<n+1\}} = V \cdot 1_{\{S\leq n\}}\ \in \mathcal{F}_n$$

by Theorem 1.10 (or Theorem 1.16d with $T = n$). Thus, $V\ 1_{(S,\infty]}$ is predictable. Since $V \in \mathcal{F}_S$, and $\mathcal{F}_S \subset \mathcal{F}_T$ by the hypothesis that $S \leq T$, we have $V \in \mathcal{F}_T$, and the preceding sentence implies that $V\ 1_{(T,\infty]}$ is predictable. Hence, the difference of the two, $V 1_{(S,T]}$, is predictable. Taking $V = 1$ shows that $1_{(S,T]}$ is predictable. Taking $T = \infty$, we see that $1_{(S,\infty]}$ is predictable, and finally, $1_{[0,S]} = 1 - 1_{(S,\infty]}$ is predictable.

Martingales stopped

Let $M = (M_n)$ be a process. Let T be a random time with values in $\bar{\mathbb{N}}$. Then, the process X defined by

3.9 $$X_n(\omega) = M_{n \wedge T(\omega)}(\omega) = \begin{cases} M_n(\omega) & \text{if } n \leq T(\omega) \\ M_{T(\omega)}(\omega) & \text{if } n > T(\omega) \end{cases}$$

is called the process M stopped at T. We observe that X is the integral 3.5 with $F = 1_{[0,T]}$. This F is bounded and positive, and further, it is predictable when T is a stopping time (see the preceding example). Hence, the following is immediate from Theorem 3.6.

3.10 THEOREM. *Let T be a stopping time. Let X be the process M stopped at T. If M is a martingale, then so is X. If M is a submartingale, then so is X.*

Doob's stopping theorem

This theorem captures the essence of the martingale property. For a martingale, given the cumulative information available at present, our estimate of any future increment is zero. Doob's theorem enables us to take the present and future times to be stopping times with some restriction (see 4.12, 4.13, and 5.8 as well), and further, it adds a simpler, more intuitive, characterization of the martingale property. The time set is still $\mathbb{N}$.

3.11 THEOREM. *Let M be adapted to $\mathcal{F}$. Then, the following are equivalent:*

a) M is a submartingale.

b) For every pair of bounded stopping times S and T with $S \leq T$, the random variables M_S and M_T are integrable and

3.12 $$\mathbb{E}_S(M_T - M_S) \geq 0.$$

c) For every pair of bounded stopping times S and T with $S \leq T$, the random variables M_S and M_T are integrable and

3.13 $$\mathbb{E}\ (M_T - M_S) \geq 0.$$

These statements remain equivalent when (a) is changed to read "M is a martingale" provided that the inequalities in 3.12 and 3.13 are changed to equalities.

3.14 REMARK. In fact, when (a) is changed to read "M is a martingale", then the inequality 3.12 needs to be replaced by equality, and (c) can be replaced with the following:

c) For every bounded stopping time T, the random variable M_T is integrable and $\mathbb{E}\, M_T = \mathbb{E}\, M_0$.

Proof. The martingale case is immediate from the submartingale case, since M is a martingale if and only if both M and $-M$ are submartingales. We shall, therefore, show that (a)⇒(b)⇒(c)⇒(a).

i) Let M be a submartingale. Let S and T be stopping times with $S(\omega) \le T(\omega) \le n$ for all ω, where n is some fixed integer. Let V be a bounded positive variable in $\mathcal{F}_S$. Putting $F = V\, 1_{(S,T]}$ in 3.4 yields a process X such that

$$X_n - X_0 = V \cdot (M_T - M_S).$$

The process F is predictable as noted in Example 3.8, and it is bounded and positive since V is so. Thus, by Theorem 3.6, the process X is a submartingale. The particular case with $V = 1$ and $S = 0$ shows that M_T is integrable (since X_n and X_0 are so), and the case with $V = 1$ and $T = n$ shows that M_S is integrable. Finally, recalling that $V \in \mathcal{F}_S$ and using the defining property for $\mathbb{E}_S$, we get

$$\mathbb{E}\, V\, \mathbb{E}_S(M_T - M_S) = \mathbb{E}\, V \cdot (M_T - M_S) = \mathbb{E}(X_n - X_0) \ge 0,$$

where the last inequality follows from the submartingale inequality for X. Since this holds for arbitrary V positive and bounded in $\mathcal{F}_S$, the random variable $\mathbb{E}_S(M_T - M_S)$ must be positive. This proves the implication (a)⇒(b).

ii) Suppose that (b) holds. Taking expectations on both sides of 3.12 yields 3.13, since $\mathbb{E}\, \mathbb{E}_S = \mathbb{E}$. So, (b)⇒(c).

iii) Suppose that (c) holds. Then, the integrability of M_n follows from that of M_T for the particular choice of $T = n$; and adaptedness of M is by hypothesis. Thus, to show (a), there remains to check the submartingale inequality $\mathbb{E}_m(M_n - M_m) \ge 0$, which is equivalent to showing that

3.15 $$\mathbb{E}\, 1_H\, \mathbb{E}_m(M_n - M_m) \ge 0, \qquad 0 \le m < n, H \in \mathcal{F}_m.$$

Fix m, n, H such. Define, for ω in Ω,

$$S(\omega) = m, \qquad T(\omega) = n 1_H(\omega) + m 1_{\Omega \setminus H}(\omega).$$

Now, S is a fixed time and is a stopping time trivially. Since $T \ge S$ and $H \in \mathcal{F}_S$, the time T is foretold at the time $S = m$; hence, T is a stopping time. Obviously, $S \le T \le n$. Finally, $M_T - M_S = 1_H \cdot (M_n - M_m)$ by the way T and S are defined. Now 3.13 shows that 3.15 holds as needed. □

Upcrossings

Let M be an adapted process. Fix a and b in $\mathbb{R}$ with $a < b$. Put $T_0 = -1$ for convenience and for each integer $k \geq 1$ define

3.16 $S_k = \inf \{ n > T_{k-1} : M_n \leq a \}, \quad T_k = \inf \{ n > S_k : M_n \geq b \}.$

Since M is adapted, $(S_1, T_1, S_2, T_2, \ldots)$ is an increasing sequence of stopping times; $S_1, S_2, \ldots$ are called the downcrossing times of the interval (a, b), and $T_1, T_2, \ldots$ are called the upcrossing times; See Figure 4 for an illustration. Then,

3.17
$$U_n(a, b) = \sum_{k=1}^{\infty} 1_{(0,n]} \circ T_k$$

is the number of upcrossings of (a, b) completed by M during $[0, n]$.

As in Heuristics 3.8, think of M_n as the price at time n of a share of some stock. Imagine someone who buys a share when the price hits a or below and sells it later when the price becomes b or above, repeating the scheme forever. Then, he buys a share at time S_1 and sells it at T_1, buys a share at time S_2 and sells it at T_2, and so on. The number of buy-sell cycles completed during $[0, n]$ is $U_n(a, b)$. The strategy employed is that of holding F_n shares during $(n-1, n]$, where

3.18
$$F_n = \sum_{k=1}^{\infty} 1_{(S_k, T_k]}(n), \quad n \geq 1,$$

and we put $F_0 = 0$ for definiteness. Now $X = \int F dM$ describes the evolution of his capital, and $X_n - X_0$ is the profit during $(0, n]$, which profit is at least $(b - a) U_n(a, b)$ assuming that the share being held at time n (if any) is worth more than what it was bought for. This heuristic observation will be of use in the proof next.

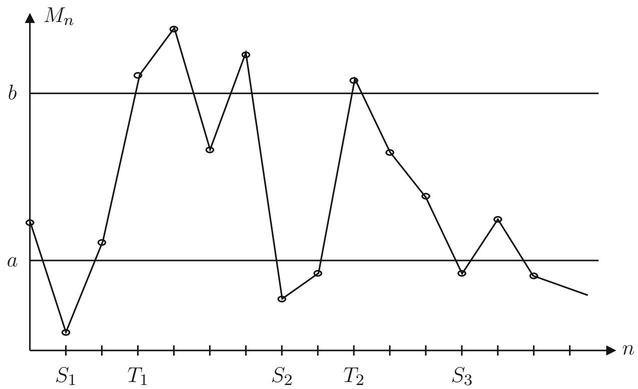

Figure 4: Upcrossing times of (a, b) are $T_1, T_2, \ldots$.

3.19 Proposition. *Suppose that M is a submartingale. Then,*

$$(b-a)\mathbb{E}\, U_n(a,b) \le \mathbb{E}\left[(M_n-a)^+ - (M_0-a)^+\right].$$

Proof. An upcrossing of (a,b) by M is the same as an upcrossing of $(0, b-a)$ by the process $(M-a)^+$, and the latter is again a submartingale by Remark 2.4g. Thus, we may and do assume that $a=0$ and $M \ge 0$.

Let $X = \int F dM$, defined by 3.4, with F given by 3.18. Note that F is predictable. Thus, $F_{k+1} \in \mathcal{F}_k$, and we have

$$\begin{aligned}\mathbb{E}_k(X_{k+1}-X_k) &= \mathbb{E}_k(M_{k+1}-M_k)F_{k+1} \\ &= F_{k+1}\mathbb{E}_k(M_{k+1}-M_k) \le \mathbb{E}_k(M_{k+1}-M_k),\end{aligned}$$

where the inequality follows from the positivity of $\mathbb{E}_k(M_{k+1}-M_k)$ and the observation that $F_{k+1} \le 1$. Taking expectations on both sides and summing over k we get

$$\mathbb{E}(X_n - X_0) \le \mathbb{E}(M_n - M_0).$$

On the other hand, as mentioned as a heuristic remark, $X_n - X_0 \ge bU_n(0,b)$ since $M_n \ge 0$ and $a = 0$. Hence,

$$b\,\mathbb{E}\, U_n(0,b) \le \mathbb{E}(X_n - X_0) \le \mathbb{E}(M_n - M_0),$$

which is the claim when $a=0$ and $M \ge 0$. □

Maxima and minima

Let $M = (M_n)$ be a process adapted to $\mathcal{F}$. For n in $\mathbb{N}$, define

3.20 $$M_n^* = \max_{k \le n} M_k, \quad m_n^* = \min_{k \le n} M_k,$$

the maxima and minima.

3.21 Theorem. *Suppose that M is a submartingale. Then, for $b > 0$,*

$$\begin{aligned}&b\,\mathbb{P}\{M_n^* \ge b\} \le \mathbb{E}\, M_n 1_{\{M_n^* \ge b\}} \le \mathbb{E}\, M_n^+, \\ &b\,\mathbb{P}\{m_n^* \le -b\} \le -\mathbb{E}\, M_0 + \mathbb{E}\, M_n 1_{\{m_n^* > -b\}} \le \mathbb{E}\, M_n^+ - \mathbb{E}\, M_0.\end{aligned}$$

3.22 Remark. It is convenient to think of these inequalities in terms of the stopping times (we suppress their dependence on b)

$$T = \inf\{n \ge 0 : M_n \ge b\}, \quad S = \inf\{n \ge 0 : M_n \le -b\},$$

that is, the time of first entrance to $[b, \infty)$ by M and the time of first entrance to $(-\infty, -b]$. Note that

3.23 $$\{M_n^* \ge b\} = \{T \le n\}, \quad \{m_n^* \le -b\} = \{S \le n\}.$$

We shall give the proof below in terms of T and S.

Proof. Fix b, fix n. Note that, on the set $\{T \le n\}$, we have $M_{T\wedge n} = M_T \ge b$. Thus,

$$b\ 1_{\{T\le n\}} \le M_{T\wedge n} 1_{\{T\le n\}} \le (\mathbb{E}_{T\wedge n} M_n) 1_{\{T\le n\}} = \mathbb{E}_{T\wedge n} M_n 1_{\{T\le n\}},$$

where the second inequality is Doob's submartingale inequality 3.12 applied with the bounded stopping times $T \wedge n$ and n, and the last equality uses Theorem 1.16d to the effect that $\{T \le n\} \in \mathcal{F}_{T\wedge n}$. Taking expectations on both sides yields the first inequality concerning M_n^* in view of Remark 3.22; the second inequality is obvious.

Similarly, on the set $\{S \le n\}$, we have $M_S \le -b$ and, hence,

$$M_{S\wedge n} = M_S\ 1_{\{S\le n\}} + M_n\ 1_{\{S>n\}} \le -b\ 1_{\{S\le n\}} + M_n\ 1_{\{S>n\}}.$$

Taking expectations, and noting that $\mathbb{E}\ M_0 \le \mathbb{E}\ M_{S\wedge n}$ by the submartingale inequality 3.13 applied with the bounded stopping times 0 and $S \wedge n$, we obtain the first inequality claimed for m_n^*. The second is obvious. □

When M is a martingale, $|M|^p$ is a submartingale for $p \ge 1$ provided that M_n be in L^p for every n. Then, applying the preceding theorem to the submartingale $|M|^p$ yields the following corollary. This is called Doob-Kolmogorov inequality; it is a generalization of Kolmogorov's inequality (Lemma III.7.1) for sums of independent variables.

3.24 COROLLARY. *Let M be a martingale in L^p for some p in $[1, \infty)$. Then, for $b > 0$,*

$$b^p\ \mathbb{P}\{\ \max_{k\le n} |M_k| > b\ \} \le \mathbb{E}\ |M_n|^p.$$

Another corollary, this time about the submartingale M directly, can be obtained by combining the two statements of Theorem 3.21:

3.25 $$b\,\mathbb{P}\{\ \max_{k\le n} |M_k| > b\ \} \le 2\,\mathbb{E}\,M_n^+ - \mathbb{E}\,M_0\ \le\ 3\ \max_{k\le n} \mathbb{E}\,|M_k|\,.$$

The following gives a bound on the expected value of the maxima of $|M|$ when M is a martingale. It is called *Doob's norm inequality.*

3.26 THEOREM. *Let M be a martingale in L^p for some $p > 1$. Let q be the exponent conjugate to p, that is, $1/p + 1/q = 1$. Then,*

$$\mathbb{E}\ \max_{k\le n} |M_k|^p \le q^p\ \mathbb{E}\ |M_n|^p.$$

Proof. Fix n, and introduce $Z = \max_{k\le n} |M_k|$ for typographical ease. We are to show that

3.27 $$\mathbb{E}\ Z^p \le q^p\ \mathbb{E}\ |M_n|^p.$$

We start by noting that

$$Z^p = \int_0^Z dx\ px^{p-1} = \int_0^\infty dx\ px^{p-2} x\ 1_{\{Z \geq x\}},$$

and

$$\mathbb{E}\ x\ 1_{\{Z\geq x\}} = x\ \mathbb{P}\{\max_{k\leq n} |M_k| \geq x\} \leq \mathbb{E}\ |M_n| \cdot 1_{\{Z\geq x\}}$$

by Theorem 3.21 applied to the submartingale $|M|$. Thus,

$$\begin{aligned} Ebb\ Z^p &\leq \mathbb{E}\ |M_n| \int_0^\infty dx\ px^{p-2}\ 1_{\{Z\geq x\}} = \mathbb{E}\ |M_n|\ q\ Z^{p-1} \\ &\leq q\ (\mathbb{E}\ |M_n|^p)^{1/p}\ (\mathbb{E}\ Z^p)^{1/q}, \end{aligned}$$

where the last inequality follows from Hölder's, II.3.6a. Solving this for $\mathbb{E}\ Z^p$ yields the desired bound 3.27. □

Exercises

3.28 *Doob's Decomposition.* Let $X = (X_n)$ be a submartingale and let

$$X = X_0 + M + A$$

be its Doob decomposition as in Theorem 3.2. Show that X is L^1-bounded if and only if both M and A are L^1-bounded.

3.29 *Martingales in L^2.* Let M be a martingale in L^2 and let Q by its quadratic variation process; see 2.30 and 2.33. Show that the martingale $X = M^2 - M_0^2 - Q$ has the form (see 3.5)

$$X = \int F\ dM$$

with $F_0 = 0$ and $F_n = 2M_{n-1}$, $n \geq 1$.

3.30 *Continuation.* Note that Q is a submartingale with $Q_0 = 0$. Let $Q = Y + A$ be its Doob decomposition with Y a martingale and A an increasing predictable process. Describe Y and A. Show that, with $N = X + Y$,

$$M^2 = M_0^2 + N + A$$

is Doob's decomposition for the submartingale M^2.

3.31 *Upcrossings.* Recall the definitions 3.16-3.18. Show that F can be obtained recursively by, starting with $F_0 = 0$,

$$F_{n+1} = F_n\ 1_{\{M_n < b\}} + (1 - F_n)\ 1_{\{M_n \leq a\}}, \quad n \geq 0.$$

Define the stopping times $S_1, T_1, S_2, T_2, \ldots$ in terms of the F_n. Show that

$$U_n(a, b) = \sum_{k=1}^n\ F_k \cdot (1 - F_{k+1}) = \sum_{k=1}^n\ F_k\ 1_{\{M_k \geq b\}}.$$

4 Martingale Convergence

This section is on the fundamental results of the classical theory of martingales. We give the basic convergence theorems, characterization of uniformly integrable martingales, and a sample of applications: Hunt's extension of the dominated convergence theorem, Lévy's extension of the Borel-Cantelli lemma, new proofs of Kolmogorov's 0-1 law and the strong law of large numbers, and a constructive proof of the Radon-Nikodym theorem.

As usual, $(\Omega, \mathcal{H}, \mathbb{P})$ is the probability space in the background. The index set is $\mathbb{N}$ unless stated otherwise. The filtration $\mathcal{F}$ is kept in the background as well, and Convention 1.21 on conditional expectations is in force throughout. All martingales, stopping times, and so on are relative to the filtration $\mathcal{F}$.

Martingale convergence theorem

The next theorem is basic. Let X be a submartingale. Doob's decomposition of it shows that it has a tendency to increase. If that tendency can be curbed, then it should be convergent.

4.1 Theorem. *Let X be a submartingale. Suppose that*

4.2 $$\sup_n \mathbb{E}\, X_n^+ < \infty.$$

Then, the sequence (X_n) converges almost surely to an integrable random variable.

4.3 Remarks. a) The condition 4.2 is that the process X^+ be L^1-bounded. Since X is a submartingale, so is X^+ by Remark 2.4g, and $\mathbb{E}\, X_n^+$ is increasing in n as a result. The condition 4.2 delimits the upward tendency of X, and the convergence of X becomes intuitive.

b) If X is a negative submartingale, then 4.2 is automatic and X converges.

c) If X is a positive supermartingale, then the preceding remark applies to $-X$ and, hence, X converges.

d) Since every martingale is a submartingale and a supermartingale, the preceding theorem and remarks apply: If X is a martingale, and if X is positive or negative or bounded from below by an integrable random variable or bounded from above similarly, then X converges almost surely to an integrable random variable X_∞.

e) When X is a submartingale, $\mathbb{E}\, X_n \geq \mathbb{E}\, X_0$ for every n, and

$$\mathbb{E}\, X_n^+ \leq \mathbb{E}\, |X_n| = 2\, \mathbb{E}\, X_n^+ - \mathbb{E}\, X_n \leq 2\, \mathbb{E}\, X_n^+ - \mathbb{E}\, X_0.$$

Hence, the condition that X^+ be L^1-bounded is equivalent to requiring that X be L^1-bounded, that is, 4.2 holds if and only if

4.4 $$\sup_n \mathbb{E}\, |X_n| < \infty.$$

Proof. Pick an outcome ω, suppose that the sequence of numbers $X_n(\omega)$ does not have a limit; then its limit inferior is strictly less than its limit superior, in which case there are at least two rationals a and b with $a < b$ that can be inserted between the two limits, and which in turn implies that the sequence upcrosses the interval (a, b) infinitely often. The set of all such ω is the union, over all rationals a and b with $a < b$, of the sets $\{U(a, b) = +\infty\}$, where $U(a, b) = \lim_n U_n(a, b)$, the total number of upcrossings of (a, b). Thus, to show that $\lim X_n$ exists almost surely, it is enough to show that for every pair of rationals a and b with $a < b$ we have $U(a, b) < \infty$ almost surely.

Fix $a < b$ such. Since $U_n(a, b)$ is increasing in n,

$$(b-a)\mathbb{E}\,U(a,b) = (b-a)\lim\mathbb{E}\,U_n(a,b) \le \sup\mathbb{E}(X_n-a)^+ \le \sup\mathbb{E}X_n^+ + |a| < \infty,$$

where we used, in succession, the monotone convergence theorem, Proposition 3.19 on upcrossings, the observation that $(x-a)^+ \le x^+ + |a|$, and the condition 4.2. Thus, $U(a, b) < \infty$ almost surely.

It follows that $X_\infty = \lim X_n$ exists almost surely. By Fatou's lemma and Remark 4.3e,

$$\mathbb{E}\,|X_\infty| = \mathbb{E}\,\liminf|X_n| \le \liminf\mathbb{E}|X_n| \le 2\sup_n \mathbb{E}\,X_n^+ - \mathbb{E}\,X_0 < \infty,$$

which shows that the limit is integrable (and thus real-valued), and hence, X is convergent. □

Convergence and uniform integrability

The following improves upon the preceding theorem in the presence of uniform integrability. Recall 1.8 *et seq.* on extending the filtration $\mathcal{F}$ onto $\bar{\mathbb{N}}$ by setting $\mathcal{F}_\infty = \lim \mathcal{F}_n = \vee_n \mathcal{F}_n$.

4.5 THEOREM. *Let X be a submartingale. Then, X converges almost surely and in L^1 if and only if it is uniformly integrable. Moreover, if it is so, setting $X_\infty = \lim X_n$ extends X to a submartingale $\bar{X} = (X_n)_{n\in\bar{\mathbb{N}}}$.*

Proof. If X converges almost surely and in L^1, then it must be uniformly integrable; see Theorem III.4.6.

If the submartingale X is uniformly integrable, then it is L^1-bounded by Remark II.3.13c and the condition 4.2 follows from Remark 4.3e. Thus, X converges almost surely by Theorem 4.1, and also in L^1 by Theorem III.4.6. Moreover, then, the limit X_∞ is integrable by 4.1 and belongs to $\mathcal{F}_\infty$ since all the X_n belong to $\mathcal{F}_\infty$. To show that $\bar{X}$ is a submartingale over $\bar{\mathbb{N}}$, there remains to show that, for every m in $\mathbb{N}$,

4.6 $$\mathbb{E}_m(X_\infty - X_m) \ge 0.$$

Fix m. Fix H in $\mathcal{F}_m$. The submartingale inequality for X implies that

$$\mathbb{E}\,1_H\cdot(X_n - X_m) = \mathbb{E}\,1_H\,\mathbb{E}_m(X_n - X_m) \ge 0$$

for every $n \geq m$. Thus, since $X_n - X_m$ goes to $X_\infty - X_m$ in L^1 as $n \to \infty$,

$$\mathbb{E}\ 1_H(X_\infty - X_m) = \lim_n\ \mathbb{E}\ 1_H(X_n - X_m) \geq 0$$

by Proposition III.4.7. Since H in $\mathcal{F}_m$ is arbitrary, this implies 4.6. □

Uniformly integrable martingales

The following theorem characterizes uniformly integrable martingales and identifies their limits. Its proof is nearly immediate from Proposition 2.7 and the preceding theorem.

4.7 THEOREM. *A process $M = (M_n)_{n\in\mathbb{N}}$ is a uniformly integrable martingale if and only if*

4.8 $$M_n = \mathbb{E}_n\ Z, \qquad n \in \mathbb{N},$$

for some integrable random variable Z. If so, it converges almost surely and in L^1 to the integrable random variable

4.9 $$M_\infty = \mathbb{E}_\infty\ Z,$$

and, moreover, $\bar{M} = (M_n)_{n\in\bar{\mathbb{N}}}$ is again a uniformly integrable martingale.

Proof. If M has the form 4.8, then it is a uniformly integrable martingale as was shown in Proposition 2.7. If M is a uniformly integrable martingale, then the preceding theorem shows that it converges almost surely and in L^1 to some integrable random variable M_∞ and that $\bar{M} = (M_n)_{n\in\bar{\mathbb{N}}}$ is again a martingale; it follows from the martingale property for $\bar{M}$ that M has the form 4.8 with $Z = M_\infty$. This completes the proof of the first statement and much of the second.

To complete the proof, there remains to show that if 4.8 holds then 4.9 holds as well, which amounts to showing that

4.10 $$\mathbb{E}\ M_\infty\ 1_H = \mathbb{E}\ Z\ 1_H$$

for every H in $\mathcal{F}_\infty$. Let $\mathcal{D}$ be the collection of all H in $\mathcal{F}_\infty$ for which 4.10 holds. Then $\mathcal{D} \supset \mathcal{F}_n$ for each n since $\mathbb{E}_n\ M_\infty = M_n = \mathbb{E}_n\ Z$; thus, $\mathcal{D} \supset \cup_n \mathcal{F}_n$. Since $\mathcal{D}$ is clearly a d-system, and since it contains the p-system $\cup_n \mathcal{F}_n$, it follows that $\mathcal{D} \supset \mathcal{F}_\infty = \sigma(\cup_n \mathcal{F}_n)$; this is by the monotone class theorem. So, 4.10 holds for every H in $\mathcal{F}_\infty$. □

4.11 COROLLARY. *For every integrable variable Z,*

$$\mathbb{E}_n\ Z \longrightarrow \mathbb{E}_\infty\ Z$$

almost surely and in L^1.

Proof is not needed; this is a partial re-statement of the preceding theorem. Note that, in particular, if $Z \in \mathcal{F}_\infty$ then $\mathbb{E}_n\, Z \to Z$; that is, if Z is revealed by the end of time, then our estimate of it at time n converges to it as $n \to \infty$.

The following supplements the preceding results and removes the boundedness condition in Doob's stopping theorem 3.11. In view of Theorem 4.7 above, the condition of the next theorem is equivalent to saying that $(M_n)_{n\in\mathbb{N}}$ is a uniformly integrable martingale and $M_\infty = \lim M_n$.

4.12 THEOREM. *Suppose that, for some integrable random variable Z,*

$$M_n = \mathbb{E}_n\; Z, \quad n \in \bar{\mathbb{N}}.$$

Then, for every stopping time T,

$$M_T = \mathbb{E}_T\; Z.$$

Moreover, for arbitrary stopping times S and T,

$$\mathbb{E}_S\; M_T = M_{S\wedge T}.$$

4.13 REMARKS. a) *On the meaning of M_T*: Since $M_n(\omega)$ is well-defined for every integer n and $n = +\infty$, the random variable M_T is well-defined even for ω with $T(\omega) = +\infty$.

b) *Doob's stopping theorem.* According to the first claim, M_T is the conditional expectation of Z given $\mathcal{F}_T$. Since Z is integrable, this implies that M_T is integrable. So, if S and T are arbitrary stopping times (taking values in $\bar{\mathbb{N}}$) with $S \le T$, the random variables M_S and M_T are integrable and

$$\mathbb{E}_S\; M_T = M_S$$

by the second claim. Thus, for uniformly integrable martingales, Doob's stopping theorem 3.11 remains true without the condition of boundedness on the stopping times.

Proof. We shall be using the repeated conditioning property, $\mathbb{E}_S\mathbb{E}_T = \mathbb{E}_{S\wedge T}$, a number of times without further comment. To prove the first claim, we start by noting that, for each n in $\mathbb{N}$,

$$M_{T\wedge n} = \mathbb{E}_{T\wedge n}\; M_n;$$

this follows from Doob's stopping theorem 3.11 for the martingale M used with the bounded stopping times $T\wedge n$ and n. Replacing M_n by $\mathbb{E}_n Z$, noting that $\mathbb{E}_{T\wedge n}\mathbb{E}_n = \mathbb{E}_{T\wedge n} = \mathbb{E}_n\mathbb{E}_T$, we get

$$M_{T\wedge n} = \mathbb{E}_n\mathbb{E}_T Z, \quad n \in \mathbb{N}.$$

As $n \to \infty$, the left side converges to M_T almost surely, whereas the right side converges to $\mathbb{E}_\infty\mathbb{E}_T Z = \mathbb{E}_T Z$ by Corollary 4.11 applied to the integrable random variable $\mathbb{E}_T Z$. Thus, $M_T = \mathbb{E}_T Z$ as claimed. The second claim is immediate:

$$\mathbb{E}_S M_T = \mathbb{E}_S \mathbb{E}_T Z = \mathbb{E}_{S \wedge T} Z = M_{S \wedge T}. \qquad \square$$

The following corollary is immediate from the preceding theorem and Lemma 2.8.

4.14 COROLLARY. *If* $(M_n)_{n \in \bar{\mathbb{N}}}$ *is a uniformly integrable martingale, then the collection*

$$\{M_T : \ T \textit{ is a stopping time}\}$$

is uniformly integrable. $\square$

Within the proof of the last theorem, we have shown that

4.15
$$\mathbb{E}_{T \wedge n} \, Z \longrightarrow \mathbb{E}_T \, Z$$

almost surely and in L^1; this follows from applying Corollary 4.11 to $\mathbb{E}_n \mathbb{E}_T Z = \mathbb{E}_{T \wedge n} Z$. Here is a consequence of this useful fact.

4.16 PROPOSITION. *Suppose that* $(\Omega, \mathcal{H}, \mathbb{P})$ *is complete, and all negligible events belong to* $\mathcal{F}_0$ *(and therefore to all the* $\mathcal{F}_n$*). Then, for every stopping time* T *of the filtration* $(\mathcal{F}_n)$*,*

$$\mathcal{F}_T = \lim_n \mathcal{F}_{T \wedge n} = \vee_n \mathcal{F}_{T \wedge n}.$$

Proof. Let $\hat{\mathcal{F}}_n = \mathcal{F}_{T \wedge n}$; we are to show that $\hat{\mathcal{F}}_\infty = \mathcal{F}_T$. Since $\hat{\mathcal{F}}_n \subset \mathcal{F}_T$ for every n, we have $\hat{\mathcal{F}}_\infty \subset \mathcal{F}_T$. To show the converse containment, let Z be a bounded variable in $\mathcal{F}_T$. Then, $Z = \mathbb{E}_T Z$ by definition and, thus, is the almost sure limit of $\mathbb{E}_{T \wedge n} Z \in \hat{\mathcal{F}}_n \subset \hat{\mathcal{F}}_\infty$. Since $\hat{\mathcal{F}}_\infty \supset \mathcal{F}_0$ and $\mathcal{F}_0$ includes every negligible event, it follows that $Z \in \hat{\mathcal{F}}_\infty$. So, $\mathcal{F}_T \subset \hat{\mathcal{F}}_\infty$ as well. $\square$

Convergence in reversed time

In this subsection, the index set is $\mathbb{T} = \{\ldots, -2, -1, 0\}$, and $\mathcal{F}$ is a filtration on $\mathbb{T}$, that is, $\mathcal{F}_m \subset \mathcal{F}_n$ for $m < n$ as before but for m and n in $\mathbb{T}$.

4.17 THEOREM. *Let* $X = (X_n)_{n \in \mathbb{T}}$ *be a martingale relative to* $\mathcal{F}$*. Then,* X *is uniformly integrable and, as* $n \to -\infty$*, it converges almost surely and in* L^1 *to the integrable random variable* $X_{-\infty} = \mathbb{E}_{-\infty} X_0$*, where* $\mathcal{F}_{-\infty} = \cap_{n \in \mathbb{T}} \mathcal{F}_n$*.*

Proof. i) The martingale property for X implies that

$$X_n = \mathbb{E}_n \, X_0, \qquad n \in \mathbb{T}.$$

Thus, X has the same form as in Proposition 2.7 and is uniformly integrable as shown there.

ii) Let a and b be real numbers with $a < b$. By Proposition 3.19 on upcrossings applied to the martingale $(X_n, X_{n+1}, \ldots, X_0, X_0, X_0, \ldots)$, the expected number of upcrossings of (a, b) by X during $[n, 0]$ is bounded by

$$\frac{1}{b-a}\mathbb{E}\ [(X_0 - a)^+ - (X_n - a)^+] \le \frac{1}{b-a}\ \mathbb{E}\ (X_0 - a)^+ < \infty.$$

Thus, the number of upcrossings of (a, b) by X over $(-\infty, 0]$ is almost surely finite, just as in the proof of the martingale convergence theorem 4.1. Hence, as in 4.1 again, X converges almost surely to some integrable random variable $X_{-\infty}$ as $n \to -\infty$. By the uniform integrability of X, the convergence is in L^1 as well. Clearly, the limit belongs to $\mathcal{F}_n$ for every n in $\mathbb{T}$, and hence, is in $\mathcal{F}_{-\infty}$. □

The next corollary mirrors Corollary 4.11. No proof is needed.

4.18 Corollary. *For every integrable random variable Z,*

$$\mathbb{E}_n\ Z \longrightarrow \mathbb{E}_{-\infty}\ Z$$

almost surely and in L^1 as $n \to -\infty$.

The preceding proof extends to submartingales on $\mathbb{T}$, but requires a condition to check the downward tendency of the submartingale as n goes toward $-\infty$.

4.19 Theorem. *Let $X = (X_n)_{n \in \mathbb{T}}$ be a submartingale relative to $\mathcal{F}$. Suppose that*

4.20 $$\inf_n\ \mathbb{E}\ X_n > -\infty.$$

Then, X is uniformly integrable and, as $n \to -\infty$, converges almost surely and in L^1 to an integrable random variable $X_{-\infty}$.

Proof. i) Since X is a submartingale, $\mathbb{E}\ X_n$ decreases as n decreases, and 4.20 implies that $a = \lim\ \mathbb{E}\ X_n$ is finite. Fix $\varepsilon > 0$, take the negative integer m such that $\mathbb{E}\ X_n \le a + \varepsilon$ for all $n \le m$, and recall that X^+ is again a submartingale. Thus, for $n \le m$ and $H \in \mathcal{F}_n$,

$$-\mathbb{E}\ X_n \le -\mathbb{E}\ X_m + \varepsilon, \quad \mathbb{E}\ X_n^+\ 1_H \le \mathbb{E}\ X_m^+\ 1_H, \quad \mathbb{E}\ X_n\ 1_{\Omega\setminus H} \le \mathbb{E}\ X_m\ 1_{\Omega\setminus H}.$$

These inequalities imply, since $|Z|\ 1_H = -Z + 2Z^+ 1_H + Z\ 1_{\Omega\setminus H}$ for all variables Z, that

4.21 $$\mathbb{E}\ |X_n|\ 1_H \le \mathbb{E}\ |X_m|\ 1_H + \varepsilon$$

for all $n \le m$ and H in $\mathcal{F}_n$.

Next, fix $b > 0$, take $H = \{\ |X_n| > b\}$, and use Markov's inequality and the submartingale property for X^+ and the fact that $\mathbb{E}\ X_n \ge a$ for all n. We get

$$b\ \mathbb{P}(H) \le \mathbb{E}\ |X_n| \le 2\mathbb{E}\ X_n^+ - \mathbb{E}\ X_n \le 2\mathbb{E}\ X_0^+ - a,$$

which shows that the probability here can be made as small as desired by taking b large enough. Using this in 4.21 and recalling the integrability of X_m, we deduce that $(X_n)_{n\le m}$ is uniformly integrable. Since adding the integrable random variables $X_{m+1}, X_{m+2}, \ldots, X_0$ does not affect the uniform integrability, we conclude that X is uniformly integrable.

ii) The remainder of the proof follows the part (ii) of the proof of 4.17 word for word. □

In the remainder of this section we give a sample of applications and extensions of the convergence theorems above.

Hunt's dominated convergence theorem

This is a useful extension of the dominated convergence theorem. Note that there is no assumption of adaptedness for the sequence.

4.22 THEOREM. *Let (X_n) be dominated by an integrable random variable and suppose that $X_\infty = \lim X_n$ exists almost surely. Then, the sequence $(\mathbb{E}_n\ X_n)$ converges to $\mathbb{E}_\infty\ X_\infty$ almost surely and in L^1.*

Proof. Suppose that $|X_n| \le Z$ for every n, where Z is integrable. Then, (X_n) is uniformly integrable, its limit X_∞ is integrable, and Corollary 4.11 implies that $\mathbb{E}_n X_\infty \to \mathbb{E}_\infty X_\infty$ almost surely and in L^1. Thus, the proof is reduced to showing that, as $n \to \infty$,

$$|\mathbb{E}_n X_n - \mathbb{E}_n X_\infty| \to 0$$

almost surely and in L^1. Convergence in L^1 is easy:

$$\mathbb{E}\ |\mathbb{E}_n X_n - \mathbb{E}_n X_\infty| \le \mathbb{E}\ \mathbb{E}_n\ |X_n - X_\infty| = \mathbb{E}\ |X_n - X_\infty| \to 0$$

since $X_n \to X_\infty$ in L^1 as well. To show the almost sure convergence, let $Z_m = \sup_{n\ge m} |X_n - X_\infty|$. Observe that, almost surely,

$$\limsup_{n\to\infty} |\mathbb{E}_n X_n - \mathbb{E}_n X_\infty| \le \limsup_{n\to\infty} \mathbb{E}_n Z_m = \mathbb{E}_\infty Z_m,$$

where the last equality follows from Corollary 4.11 after noting that Z_m is integrable, in fact, $|Z_m| \le 2Z$. This completes the proof since $\mathbb{E}_\infty Z_m \to \mathbb{E}_\infty \lim Z_m = 0$ as $m \to \infty$ by the dominated convergence property IV.1.8 for conditional expectations. □

An extension of the Borel-Cantelli lemma

Let $X = (X_n)$ be a sequence, adapted to $(\mathcal{F}_n)$, of positive integrable random variables. Put $S_0 = 0$ and

$$S_n = S_0 + X_1 + \cdots + X_n, \quad n \in \mathbb{N}.$$

Then, $S = (S_n)_{n\in\mathbb{N}}$ is an increasing integrable process adapted to $\mathcal{F}$, that is, an increasing submartingale. Let

4.23 $$S = M + A$$

be its Doob decomposition: M is a martingale and A is an increasing predictable process. Since both S and A are increasing, the limits $S_\infty = \lim S_n$ and $A_\infty = \lim A_n$ are well-defined.

4.24 Proposition. *For almost every ω,*

$$A_\infty(\omega) < \infty \Rightarrow S_\infty(\omega) < \infty.$$

If X is bounded by some constant, then the reverse implication holds as well.

4.25 Remark. Recall that, in Doob's decomposition 4.23, we have

$$A_n = \mathbb{E}_0 X_1 + \mathbb{E}_1 X_2 + \cdots + \mathbb{E}_{n-1} X_n.$$

If the X_n are independent, then $A_n = \mathbb{E}S_n$, and the preceding proposition relates the convergence of S_n to the convergence of its mean, which relation is what the Borel-Cantelli lemma is about.

Proof. i) For b in $(0, \infty)$ let

$$T = \inf\{n : A_{n+1} > b\},$$

and let N be the martingale M stopped at T (see 3.9). Predictability of A implies that T is a stopping time, and Theorem 3.10 shows that N is a martingale. Since $M = S - A \geq -A$ and since $A_n \leq b$ on $\{n \leq T\}$, the martingale $N + b$ is positive, and, hence, it is almost surely convergent by Remark 4.3d. Hence, N is convergent almost surely.

ii) Let Ω_b be the almost sure event on which the limit N_∞ exists and is finite, and let $H_b = \Omega_b \cap \{A_\infty \leq b\}$. For every ω in H_b, we have $T(\omega) = \infty$, which means that $M_n(\omega) = N_n(\omega)$ for every integer n, which implies that the limit $M_\infty(\omega)$ of $M_n(\omega)$ exists and is finite, which in turn allows us to conclude that $S_\infty(\omega) = M_\infty(\omega) + A_\infty(\omega) < \infty$. Hence, on the event $H = H_1 \cup H_2 \cup \cdots$, we have $A_\infty < \infty$ and $S_\infty < \infty$, and noting that $\{A_\infty < \infty\} \setminus H$ is negligible completes the proof of the first statement.

iii) Next, suppose that the sequence X is bounded by some constant c. For b in $(0, \infty)$ define

$$T = \inf\{n : S_n > b\},$$

and let N be the martingale M stopped at T. Since $X_n \leq c$ for every n, the process N^+ is bounded by $b + c$. So, N converges almost surely to some finite random variable N_∞; this is by the martingale convergence theorem 4.1. The remainder of the proof of the second statement follows the part (ii) above with the letter A replaced by S, and S by A. □

Kolmogorov's 0-1 law

This is to illustrate the power of the martingale machinery by giving a short proof of Kolmogorov's 0-1 law. Let $X_1, X_2, \ldots$ be independent random variables. Suppose that $\mathcal{F}_n = \sigma(X_1, \ldots, X_n)$. Put $\mathcal{T}_n = \sigma(X_{n+1}, X_{n+2}, \ldots)$, and let $\mathcal{T} = \cap_n \mathcal{T}_n$, the tail σ-algebra.

4.26 Proposition. *If $H \in \mathcal{T}$ then $\mathbb{P}(H)$ is either 0 or 1.*

Proof. By Corollary 4.11, for every event H,

$$\mathbb{E}_n\, 1_H \to \mathbb{E}_\infty\, 1_H$$

almost surely. When $H \in \mathcal{T}$, since $\mathcal{T}$ is independent of $\mathcal{F}_n$, we have $\mathbb{E}_n\, 1_H = \mathbb{E}\, 1_H = \mathbb{P}(H)$. On the other hand, since $\mathcal{T}_n \subset \mathcal{F}_\infty$ for every n, we have $\mathcal{T} \subset \mathcal{F}_\infty$, which implies that $\mathbb{E}_\infty\, 1_H = 1_H$. Thus, $1_H(\omega)$ is equal to the number $\mathbb{P}(H)$ for almost every ω, which makes the latter either 0 or 1. □

Of course, consequently, for every random variable X in $\mathcal{T}$, there is constant c in $[-\infty, +\infty]$ such that $X(\omega) = c$ for almost every ω.

Strong law of large numbers

Let $X_1, X_2, \ldots$ be independent and identically distributed real-valued random variables with finite mean a. Then, by the strong law of large numbers,

4.27 $$\bar{X}_n = \frac{1}{n}(X_1 + \cdots + X_n)$$

converges almost surely to the constant a. Here is a martingale proof of this, which shows that the convergence is in L^1 as well.

Let $\mathcal{F}_{-n} = \sigma(\bar{X}_n, \bar{X}_{n+1}, \ldots)$, which is the same as the σ-algebra generated by $\bar{X}_n$ and the tail $\mathcal{T}_n = \sigma(X_{n+1}, X_{n+2}, \ldots)$. Since the vector $(X_1, \ldots, X_n)$ is independent of $\mathcal{T}_n$, the conditional expectation $\mathbb{E}_{-n} X_k$ of X_k given $\mathcal{F}_{-n}$ depends only on $\bar{X}_n$. Since the distribution of that vector remains invariant under permutations of its entries, $\mathbb{E}_{-n} X_1 = \cdots = \mathbb{E}_{-n} X_n$. But the sum of these n things is equal to $\mathbb{E}_{-n}(X_1 + \cdots + X_n) = n\bar{X}_n$. Hence,

4.28 $$\bar{X}_n = \mathbb{E}_{-n} X_1, \quad n = 1, 2, \ldots.$$

Corollary 4.18 on reversed martingales applies to the right side:

4.29 $$\bar{X}_\infty = \lim \bar{X}_n$$

exists almost surely and in L^1. Convergence in L^1 helps to see that

$$\mathbb{E}\, \bar{X}_\infty = \lim \mathbb{E}\, \bar{X}_n = a.$$

On the other hand, 4.27 shows that

$$\bar{X}_\infty = \lim_n \frac{1}{n}(X_{k+1} + \cdots + X_{k+n}),$$

which shows that $\bar{X}_\infty$ belongs to $\mathcal{T}_k$ for every k and, hence, to the tail σ-algebra $\cap_k \mathcal{T}_k$. By Kolmogorov's 0-1 law, $\bar{X}_\infty$ is a constant over an almost sure event, and obviously that constant is its mean a.

Radon-Nikodym theorem

Our aim here is to give a constructive, and intuitive, proof of the Radon-Nikodym theorem and a very useful extension of it due to Doob.

First, a definition: A σ-algebra $\mathcal{G}$ on Ω is said to be *separable* if it is generated by some sequence (H_n) of subsets of Ω. Then, letting $\mathcal{F}_n = \sigma(H_1, \ldots, H_n)$, we obtain a filtration $(\mathcal{F}_n)$ such that

4.30 $$\mathcal{F}_\infty = \lim \mathcal{F}_n = \vee_n \mathcal{F}_n = \mathcal{G}.$$

Indeed, each $\mathcal{F}_n$ has only a finite number of elements, and we can easily find a finite partition $\mathcal{D}_n$ of Ω such that $\mathcal{F}_n = \sigma(\mathcal{D}_n)$, that is, each set H in $\mathcal{F}_n$ is the union of some finite number of elements of $\mathcal{D}_n$. Obviously, $\mathcal{D}_n$ gets more and more refined as n increases. So, $\mathcal{G}$ is a separable σ-algebra if and only if it is generated by a sequence $(\mathcal{D}_n)$ of finite partitions of Ω.

For example, if $\Omega = [0,1]$, its Borel σ-algebra $\mathcal{G}$ is separable: for $\mathcal{D}_n$ take the partition whose elements are $[0,a]$, $(a, 2a]$, $(2a, 3a]$, $\ldots$, $(1-a, 1]$ with $a = 1/2^n$. This example is worth keeping in mind.

4.31 THEOREM. *Let $\mathcal{G}$ be a separable sub-σ-algebra of $\mathcal{H}$. Let Q be a finite measure on $(\Omega, \mathcal{G})$. Suppose that Q is absolutely continuous with respect to P, the latter being the restriction of $\mathbb{P}$ to $\mathcal{G}$. Then, there exists a positive random variable Z in $\mathcal{G}$ such that*

4.32 $$Q(H) = \int_H \mathbb{P}(d\omega) Z(\omega), \quad H \in \mathcal{G}.$$

4.33 REMARKS. a) Of course, the conclusion is that

$$Z = \frac{dQ}{dP},$$

that is, Z is a version of the Radon-Nikodym derivative of Q with respect to P.

b) If $\mathcal{H}$ is separable, or if $\mathcal{H}$ differs from a separable σ-algebra by a collection of negligible events, then the theorem remains true with $\mathcal{G} = \mathcal{H}$. In fact, in most situations in probability theory, this remark is applicable to $\mathcal{H}$.

c) In fact, the separability condition can be dropped: The claim of the theorem is true for arbitrary sub-σ-algebra $\mathcal{G}$ of $\mathcal{H}$ (See the notes for references).

Proof. We start by constructing a sequence of random variables (this is the intuitive part) and give the proof through a series of lemmas.

4.34 CONSTRUCTION. For each n, let $\mathcal{F}_n$ be the σ-algebra generated by a finite partition $\mathcal{D}_n$ of Ω such that the sequence $(\mathcal{F}_n)$ is a filtration and 4.30 holds.

For each ω in Ω, there is a unique element H of $\mathcal{D}_n$ such that $\omega \in H$, and then we define $X_n(\omega)$ to be the ratio $Q(H)/P(H)$; in other words,

4.35 $$X_n(\omega) = \sum_{H \in \mathcal{D}_n} \frac{Q(H)}{P(H)} 1_H(\omega), \qquad n \in \mathbb{N},\ \omega \in \Omega,$$

with the convention that $0/0 = 0$. Obviously, each X_n is positive and is in $\mathcal{F}_n$ and takes finitely many values, and

4.36 $$Q(H) = \mathbb{E}\ 1_H X_n, \quad H \in \mathcal{F}_n.$$

4.37 LEMMA. *The process (X_n) is a positive martingale with respect to the filtration $(\mathcal{F}_n)$; it converges almost surely to an integrable positive random variable Z in $\mathcal{G}$.*

Proof. The positivity and adaptedness are obvious. Taking $H = \Omega$ in 4.36 shows that $\mathbb{E}\ X_n = Q(\Omega) < \infty$. To see the martingale property, let $H \in \mathcal{F}_n$. Then, $H \in \mathcal{F}_{n+1}$ as well, and 4.36 shows that

4.38 $$\mathbb{E}\ 1_H X_n = Q(H) = \mathbb{E}\ 1_H X_{n+1}.$$

This is another way of saying that $\mathbb{E}_n\ X_{n+1} = X_n$. Thus, X is a positive martingale. The remaining claim is immediate from the convergence theorem 4.1 and Remark 4.3d. □

4.39 LEMMA. *For every $\varepsilon > 0$ there is $\delta > 0$ such that, for every event H in $\mathcal{G}$,*

$$P(H) \leq \delta \Rightarrow Q(H) \leq \varepsilon.$$

Proof. This is by the assumed absolute continuity of Q with respect to P. We show it by contradiction. Suppose that for some $\varepsilon > 0$ there is no such δ. Then, there must exist H_n in $\mathcal{G}$ such that

$$P(H_n) \leq 1/2^n, \quad Q(H_n) > \varepsilon.$$

Define $H = \limsup H_n$, that is, 1_H is the limit superior of the indicators of the H_n. By Borel-Cantelli for the probability measure P we have $P(H) = 0$, whereas

$$Q(H) \geq \limsup Q(H_n) \geq \varepsilon$$

by Fatou's lemma applied with the finite measure Q. This contradicts the absolute continuity ($P(H) = 0 \Rightarrow Q(H) = 0$). □

The following lemma completes the proof of Theorem 4.31.

4.40 LEMMA. *The martingale (X_n) is uniformly integrable, and its limit Z is in $\mathcal{G}$ and satisfies 4.32.*

Proof. Pick $\varepsilon > 0$ and choose $\delta > 0$ as in the preceding lemma. Let $b = Q(\Omega)/\delta$ and $H = \{X_n > b\}$. Since

$$P(H) \le \frac{1}{b}\ \mathbb{E}\ X_n = \frac{1}{b}\ Q(\Omega) = \delta,$$

we have (recall 4.36)

$$\mathbb{E}\ X_n\ 1_{\{X_n > b\}} = \mathbb{E}\ 1_H\ X_n = Q(H) \le \varepsilon.$$

Thus, (X_n) is uniformly integrable. By Lemma 4.37 it is a positive martingale and converges to Z in $\mathcal{G}$ almost surely. Hence, it converges to Z in L^1 as well.

Define $\hat{Q}(H)$ to be the integral on the right side of 4.32. Convergence in L^1 allows us to write

$$\hat{Q}(H) = \mathbb{E}\ 1_H\ Z = \lim_n\ \mathbb{E}\ 1_H\ X_n$$

for every event H. But in view of 4.36, $\hat{Q}(H) = Q(H)$ for every H in $\mathcal{F}_n$; that is, $Q = \hat{Q}$ on $\mathcal{F}_n$ for each n. Hence, Q and $\hat{Q}$ coincide on the p-system $\cup_n \mathcal{F}_n$ and, therefore, on the σ-algebra $\mathcal{G}$ generated by that p-system; see Proposition 3.8 of Chapter I. □

4.41 REMARK. *Singularity.* Suppose that Q on $(\Omega, \mathcal{G})$ is singular with respect to P. Lemma 4.37 still holds, the almost sure limit Z is positive and integrable, and the right side of 4.32 defines a finite measure $\hat{Q}$. Using Fatou's lemma with 4.36 shows that $\hat{Q}(H) \le Q(H)$, which means that $\hat{Q}$ puts all its mass on a set of zero P-measure. But, obviously, $\hat{Q}$ is absolutely continuous with respect to P. It follows that $\hat{Q} = 0$, which means that $Z = 0$ almost surely.

4.42 REMARK. *Lebesgue's decomposition.* Let the measure Q be an arbitrary finite measure on $(\Omega, \mathcal{G})$. We may normalize it to make it a probability measure, and we assume so. Then, $\hat{P} = {}^1\!/_2(P + Q)$ is a probability measure, and both P and Q are absolutely continuous with respect to $\hat{P}$. Thus, by the preceding theorem, there exists a positive X in $\mathcal{G}$ such that

$$P(H) = \int_H \hat{P}(d\omega) X(\omega), \quad Q(H) = \int_H \hat{P}(d\omega)\,(2 - X(\omega))$$

and we may assume that $0 \le X \le 2$. Thus, with $Z = \frac{2}{X} - 1$ and $\Omega_0 = \{X = 0\}$,

4.43 $$Q(H) = \int_H P(d\omega)\, Z(\omega) + Q(H \cap \Omega_0), \quad H \in \mathcal{G}.$$

On the right side, the integral defines a measure Q_c which is absolutely continuous with respect to P, and the second term defines a measure Q_s which is singular with respect to P. This decomposition $Q = Q_c + Q_s$ is called the Lebesgue decomposition of Q.

Doob's theorem for families of measures

This is immensely useful in the theory of Markov processes. The separability condition cannot be removed. In applications, E is sometimes a "space", sometimes the time set, and sometimes is the space-time product. The point of the theorem is the joint measurability of Z.

4.44 THEOREM. *Let Ω be a set and $\mathcal{G}$ a separable σ-algebra on it. Let $(E, \mathcal{E})$ be an arbitrary measurable space. Let Q be a bounded transition kernel, and P a probability kernel, both from $(E, \mathcal{E})$ into $(\Omega, \mathcal{G})$. Suppose that, for each x in E, the measure $H \mapsto Q(x, H)$ is absolutely continuous with respect to the measure $H \mapsto P(x, H)$. Then, there exists a positive Z in $\mathcal{E} \otimes \mathcal{G}$ such that*

4.45
$$Q(x, H) = \int_H P(x, d\omega) Z(x, \omega), \quad x \in E, H \in \mathcal{G}.$$

Proof. For each x, define $X_n(x, \omega)$ by 4.35 from the measures $Q(x, \cdot)$ and $P(x, \cdot)$. Measurability of $Q(x, H)$ and $P(x, H)$ in x shows that, for each r in $\mathbb{R}_+$,

$$\{(x, \omega) \in E \times \Omega : \quad X_n(x, \omega) \le r\}$$

is a finite union of rectangles $A \times H$ in $\mathcal{E} \otimes \mathcal{F}_n$. Thus, all the X_n are in $\mathcal{E} \otimes \mathcal{G}$ (this is joint measurability), which implies that Z defined next is in $\mathcal{E} \otimes \mathcal{G}$: For x in E and ω in Ω, let

$$Z(x, \omega) = \lim_n X_n(x, \omega)$$

if the limit exists and is finite, and otherwise, put $Z(x, \omega) = 0$. Now, for each x, Theorem 4.31 shows that 4.45 holds. □

Exercises and complements

4.46 *Doob's decomposition.* Let X be an L^1-bounded submartingale, and let $X = X_0 + M + A$ be its Doob decomposition. Show that the martingale M and the increasing process A are both convergent and their limits M_∞ and A_∞ are integrable.

4.47 *Convergence in L^p.* If $Z \in L^p$ for some p in $[1, \infty]$, then the martingale $(\mathbb{E}_n Z)$ is L^p-bounded and converges to $\mathbb{E}_\infty Z$ almost surely and in L^p.

4.48 *Dominated convergence in L^p.* Let X be a martingale that is L^p-bounded for some $p > 1$. Define

$$X^* = \sup_n |X_n|.$$

a) Show that $X^* \in L^p$ and that (X_n) is dominated by X^*.

b) Show that X converges almost surely and in $L^p$to a random variable X_∞, and $|X_\infty| \le X^*$.

4.49 *Convergence in reversed time.* In Theorem 4.17, suppose further that $X_0 \in L^p$ for some p in $[1, \infty]$. Show that, then, X converges to $X_{-\infty}$ in L^pas well. In particular, show that $X^* = \sup_n |X_n|$ is in L^p.

4.50 *Markov chains started at $-\infty$.* Let $X = (X_n)_{n \in \mathbb{T}}$ be a Markov chain with state space $(E, \mathcal{E})$ and transition kernel P over the time-set $\mathbb{T} = \{\dots, -2, -1, 0\}$; see Definition 2.11 and take $\mathcal{F} = (\mathcal{F}_n)_{n \in \mathbb{T}}$ and require 2.12 for n in $\mathbb{T}$. Let f be a bounded function in $\mathcal{E}_+$, and set

$$M_n f = (P^{-n} f) \circ X_n, \quad n \in \mathbb{T}.$$

a) Show that $M_n f = \mathbb{E}_n\, f \circ X_0$ for $n \in \mathbb{T}$. Show that it converges, as $n \to -\infty$, to a bounded random variable mf almost surely and in L^1.

b) Suppose that there is a probability measure μ on $(E, \mathcal{E})$ such that $P^n f \to \mu f$ as $n \to +\infty$ for every bounded f in $\mathcal{E}_+$. Show that, then, the random variable mf is equal to the constant μf, the integral of f with respect to the measure μ. Thus, we have shown that

$$\lim_{n \to -\infty} \mathbb{E}_n\, f \circ X_0 = \mu f$$

for every bounded f in $\mathcal{E}_+$, which enables us to interpret μ as "the distribution of X_0 assuming that the chain is started at $-\infty$".

4.51 *Submartingales majorized by martingales.* Let $X = (X_n)_{n \in \mathbb{N}}$ be a uniformly integrable submartingale and let $X_\infty = \lim X_n$. Define

$$M_n = \mathbb{E}_n\, X_\infty, \quad n \in \mathbb{N},$$

Then, M is a uniformly integrable martingale. Show that $X_n \le M_n$ almost surely for all n.

4.52 *Decomposition of submartingales.* Let X be an L^1-bounded submartingale and let $X_\infty = \lim X_n$. Define

$$M_n = \mathbb{E}_n X_\infty, \quad V_n = X_n - M_n, \quad n \in \mathbb{N}.$$

This yields a decomposition $X = M + V$, where M is a uniformly integrable martingale and V is a submartingale with $\lim V_n = 0$ almost surely. Show these, and show that this decomposition is unique.

4.53 *Potentials.* Let X be a positive supermartingale. (Then, it is convergent almost surely.) If $\lim X_n = 0$ almost surely, then X is called a *potential.* Show that a positive supermartingale X is a potential if $\lim \mathbb{E}\, X_n = 0$.

4.54 *Decomposition of supermartingales.* Let X be a supermartingale with $\sup \mathbb{E}\, X_n^- < \infty$. Show that, then, X converges almost surely to an integrable random variable X_∞. Show that there is a unique decomposition

$$X = M + V$$

where M is a uniformly integrable martingale and V is a potential.

4.55 *Riesz decomposition.* Every positive supermartingale X has a decomposition

$$X = Y + Z$$

where Y is a positive martingale and Z is a potential. This is called the Riesz decomposition of X. Show this by following the steps below:

a) Show that the limit $Y_m = \lim_{n\to\infty} \mathbb{E}_m X_{m+n}$ exists almost surely.

b) Show that $Y = (Y_m)$ is a positive martingale.

c) Show that $Z = X - Y$ is a positive supermartingale and use 4.53 to conclude that Z is a potential.

4.56 *Continuation.* The martingale Y in the preceding decomposition is the maximal submartingale majorized by X, that is, if W is a submartingale and $W_n \le X_n$ for every n then $W_n \le Y_n$ for every n.

4.57 *Another decomposition for supermartingales.* Let X be a supermartingale with $\sup \mathbb{E}X_n^- < \infty$. Write $X = M+V$ as in Exercise 4.54. Let $V = N+Z$ be the Riesz decomposition of V. Then,

$$X = M + N + Z,$$

where M is a uniformly integrable martingale, N is a martingale potential, and Z is a potential.

4.58 *Krickeberg decomposition.* Let X be an L^1-bounded martingale. Then $X^+ = (X_n^+)$ and $X^- = (X_n^-)$ are positive submartingales. Show that

$$Y_n = \lim_m \mathbb{E}_n\, X_{n+m}^+, \quad Z_n = \lim_m \mathbb{E}_n\, X_{n+m}^-$$

exist. Show that Y and Z are positive and L^1-bounded martingales. Show that

$$X = Y - Z.$$

This is called the Krickeberg decomposition.

4.59 *Continuation.* A martingale is L^1-bounded if and only if it is the difference of two positive L^1-bounded martingales. Show.

4.60 *Continuation.* In the Krickeberg decomposition of an L^1-bounded martingale X, the process Y is the minimal positive martingale majorizing X, and the process Z is the minimal positive martingale majorizing $-X$ (see 4.56 for the meaning). Show this.

5 Martingales in Continuous Time

Throughout this section, $(\Omega, \mathcal{H}, \mathbb{P})$ is a complete probability space in the background, and $\mathcal{F}$ is a filtration over $\mathbb{R}_+$ which is extended onto $\bar{\mathbb{R}}_+$ as usual by setting $\mathcal{F}_\infty = \lim \mathcal{F}_t = \vee_t \mathcal{F}_t$. We shall assume that $\mathcal{F}$ satisfies two technical conditions. First, we assume that $\mathcal{F}$ is *right-continuous*, that is,

$$\mathcal{F}_t = \cap_{\varepsilon>0} \mathcal{F}_{t+\varepsilon}, \quad t \in \mathbb{R}_+.$$

Heuristically, this means that $\mathcal{F}_t$ includes all events that can be told by an "infinitesimal peek" into the future. Second, we assume that $\mathcal{F}$ is *augmented*, which means that $(\Omega, \mathcal{H}, \mathbb{P})$ is complete and that $\mathcal{F}_0$ (and therefore every $\mathcal{F}_t$) contains the collection of all negligible events in $\mathcal{H}$. These two conditions are considered harmless and are generally easy to fulfill; we shall clarify these concepts in Section 7 along with their ramifications.

Throughout, X is a real-valued stochastic process indexed by $\mathbb{R}_+$ or $\bar{\mathbb{R}}_+$ and adapted to $\mathcal{F}$. We shall assume that it is right-continuous and has left-limits on $\mathbb{R}_+$. We shall show in Section 7 that all martingales can be modified to have such regularity properties; thus, these assumptions on X are harmless.

All martingales, stopping times, etc. will be relative to the filtration $\mathcal{F}$ unless stated otherwise. As always, Convention 1.21 is in force: $\mathbb{E}_T$ denotes the conditional expectation operator given $\mathcal{F}_T$. Since X is adapted and right-continuous, Theorem 1.14 characterizing $\mathcal{F}_T$ shows that X_T belongs to $\mathcal{F}_T$, but one should take care that X_T be well-defined (for ω with $T(\omega) = +\infty$).

Doob martingales

For martingales in discrete time, Doob's stopping theorem 3.11 extends the martingale property at fixed times to random times that are bounded stopping times; see also Theorem 4.12 and Remark 4.13b. The resulting "strong" martingale equality (3.12 with the equality sign) captures the essence of martingales. We isolate this and incorporate it into the following definitions.

5.1 Definition. *The process X is said to have the* Doob property for (S,T) *provided that S and T be stopping times with $S \le T$, and X_S and X_T be well-defined and integrable, and*

$$X_S = \mathbb{E}_S\ X_T.$$

5.2 Definition. *The process X is said to be a* Doob martingale *on* $[0, \zeta]$ *if ζ is a stopping time and X has the Doob property for (S,T) for all stopping times S and T with $0 \le S \le T \le \zeta$.*

5.3 Remarks. a) For the Doob property, the condition that X_S and X_T be well-defined is needed only when X is indexed by $\mathbb{R}_+$ and S or T is allowed to take $+\infty$ as a value. If X is a Doob martingale on $[0, \zeta]$, there is the implicit assumption that X_ζ is well-defined and integrable and has expectation equal to $\mathbb{E}\ X_0$; these follow from the assumed Doob property for (S,T) with $S = 0$ and $T = \zeta$.

b) Note that Doob martingales are defined for closed intervals $[0, \zeta]$. Being closed on the right plays a significant role in the treatment below.

c) Suppose that X is a Doob martingale on $[0, \zeta]$. Then, the Doob property for $(t \wedge \zeta, \zeta)$ implies that

$$\hat{X}_t = X_{\zeta \wedge t} = \mathbb{E}_{\zeta \wedge t} X_\zeta = \mathbb{E}_t \mathbb{E}_\zeta X_\zeta = \mathbb{E}_t X_\zeta$$

for every t in $\mathbb{R}_+$. Thus, $(\hat{X}_t)_{t\in\mathbb{R}_+}$ is a uniformly integrable martingale; see Proposition 2.7. In other words, if X is a Doob martingale on $[0, \zeta]$ then the process $\hat{X}$ obtained by stopping X at ζ is a uniformly integrable martingale. We shall show below, in Theorem 5.14, that the converse is true as well.

In the following, we characterize Doob martingales in terms of simpler looking conditions, show their intimate connections to uniform integrability, and discuss some of their uses on Brownian motions. We start with the following characterization; see Remark 3.15 for the discrete-time source of the ideas.

5.4 THEOREM. *Let ζ be a stopping time. Then, the following are equivalent:*

a) The process X is a Doob martingale on $[0, \zeta]$.
b) For every stopping time T majorized by ζ,

$$X_T = \mathbb{E}_T X_\zeta .$$

c) For every stopping time T majorized by ζ,

$$\mathbb{E}\ X_T = \mathbb{E}\ X_0 .$$

Proof. Clearly (a)⇒(b): the latter is the Doob property for (T, ζ). If (b) holds, then $\mathbb{E}\ X_T = \mathbb{E}\ X_\zeta$ and taking $T = 0$ we get $\mathbb{E}\ X_\zeta = \mathbb{E}\ X_0$; so, (b)⇒(c).

To show that (c)⇒(b), assume (c). Let T be a stopping time majorized by ζ, that is, $T \le \zeta$. Take an event H in $\mathcal{F}_T$ and define

$$S = T\ 1_H + \zeta(1 - 1_H).$$

Then, S is a stopping time majorized by ζ, and

$$X_\zeta - X_S = (X_\zeta - X_T)\ 1_H .$$

The expectation of the left side is 0 since $\mathbb{E}\ X_\zeta = \mathbb{E}\ X_0 = \mathbb{E}\ X_S$ by the assumed property (c). Thus, the expectation of the right side is 0, and this is for arbitrary H in $\mathcal{F}_T$; hence, (b) holds.

Finally, (b)⇒(a): If S and T are stopping times with $S \le T \le \zeta$, then

$$\mathbb{E}_S X_T = \mathbb{E}_S \mathbb{E}_T X_\zeta = \mathbb{E}_S X_\zeta = X_S ,$$

where we used (b) to justify the first and the last equalities; this shows that Doob property holds for (S, T). □

5.5 COROLLARY. *If X is a Doob martingale on $[0, \zeta]$, then*

$$\{X_T :\ T \text{ is a stopping time, } T \le \zeta\} = \{X_{T\wedge\zeta} :\ T \text{ is a stopping time}\}$$

is uniformly integrable.

Proof. The proof is immediate from the statement (b) of the preceding theorem and Lemma 2.8. □

Doob's stopping theorem

The statement (c) of the next theorem is the classical version of Doob's stopping theorem for martingales served up in the language of Definition 5.1.

5.6 THEOREM. *The following are equivalent for X:*

a) It is a martingale on $\mathbb{R}_+$.
b) It is a Doob martingale on $[0, b]$ for every b in $\mathbb{R}_+$.
c) It has the Doob property for (S, T) whenever S and T are bounded stopping times with $S \le T$.

Proof. All the implications are immediate from the definitions except for the implication (a)⇒(b). To show it, we use the preceding characterization theorem. Accordingly, assume (a), fix b in $\mathbb{R}_+$, and let T be a stopping time bounded by b. Then, b is a fixed stopping time, X_T is well-defined, and we are to show that $\mathbb{E}\, X_T = \mathbb{E}\, X_0$.

Let (T_n) be as in Proposition 1.20. For fixed n, the stopping time T_n is bounded by $b+1$ and takes values in the discrete set $\mathbb{T}$ consisting of the numbers $b+1$ and $k/2^n$ with k in $\mathbb{N}$. By Doob's stopping theorem 3.11 for the discrete-time martingale $(X_t)_{t\in\mathbb{T}}$ applied at the bounded times T_n and $b+1$,

5.7 $$X_{T_n} = \mathbb{E}_{T_n} X_{b+1}.$$

Recalling that $\cdots \le T_2 \le T_1 \le T_0$, this means that $(X_{T_n})_{n\in\mathbb{N}}$ is a reversed-time martingale relative to the filtration $(\mathcal{F}_{T_n})$. Thus, by Theorem 4.17, it converges almost surely and in L^1 to an integrable random variable. But, since (T_n) decreases to T and X is right-continuous, that limit is X_T. Since convergence in L^1 implies the convergence of expectations, and in view of 5.7,

$$\mathbb{E}\, X_T = \lim\ \mathbb{E}\, X_{T_n} = \mathbb{E}\, X_{b+1} = \mathbb{E}\, X_0. \qquad \square$$

The preceding theorem shows, in particular, that X is a Doob martingale on $[0, b]$ if and only if it is a martingale on $[0, b]$: in the proof, replace T_n by $T_n \wedge b$ and replace $\mathbb{T}$ with $\mathbb{T} \wedge b = \{t \wedge b : t \in \mathbb{T}\}$. This remains true when b is replaced by $+\infty$, as the next theorem shows. Note that the second assertion is Doob's stopping theorem, for this case, extended to arbitrary stopping times. See 4.12 for the discrete-time version.

5.8 THEOREM. *The process X is a Doob martingale on $\bar{\mathbb{R}}_+$ if and only if it is a martingale on $\bar{\mathbb{R}}_+$. If so, then*

$$\mathbb{E}_S X_T = X_{S\wedge T}$$

for arbitrary stopping times S and T.

Proof. Assuming that X is a martingale on $\bar{\mathbb{R}}_+$, we shall show that

5.9 $$X_T = \mathbb{E}_T X_\infty$$

for every stopping time T. This will prove the first claim via Theorem 5.4 with $\zeta = +\infty$. The second claim follows from 5.9 since $\mathbb{E}_S\mathbb{E}_T = \mathbb{E}_{S\wedge T}$. So, assume X is a martingale on $\bar{\mathbb{R}}_+$ and let T be a stopping time.

For each n in $\mathbb{N}$, Theorem 5.4 implies that X has the Doob property for the bounded stopping times $T \wedge n$ and n:

$$X_{T\wedge n} = \mathbb{E}_{T\wedge n} X_n.$$

Since X is a martingale on $\bar{\mathbb{R}}_+$, the process $(X_n)_{n\in\bar{\mathbb{N}}}$ is a martingale relative to $(\mathcal{F}_n)_{n\in\bar{\mathbb{N}}}$; thus, $X_n = \mathbb{E}_n X_\infty$ for each n. This implies, together with the observations $\mathbb{E}_{T\wedge n}\mathbb{E}_n = \mathbb{E}_{T\wedge n} = \mathbb{E}_n\mathbb{E}_T$, that

5.10 $$X_{T\wedge n} = \mathbb{E}_n\mathbb{E}_T X_\infty, \quad n \in \mathbb{N}.$$

At this point, we remark that

5.11 $$\hat{\mathcal{F}}_\infty = \vee_{n\in\mathbb{N}}\mathcal{F}_n = \vee_{t\in\mathbb{R}_+}\mathcal{F}_t = \mathcal{F}_\infty.$$

This is because every $\mathcal{F}_n$ is contained in some $\mathcal{F}_t$ and vice versa.

The right side of 5.10 converges almost surely, by Corollary 4.11, to the conditional expectation of $\mathbb{E}_T X_\infty$ given $\hat{\mathcal{F}}_\infty$, which is the same as $\mathbb{E}_\infty\mathbb{E}_T X_\infty = \mathbb{E}_T X_\infty$ in view of 5.11. Whereas, the left side of 5.10 converges almost surely to X_T: if $T(\omega) < \infty$ then $X_{T\wedge n}(\omega) = X_T(\omega)$ for all n large enough, and if $T(\omega) = \infty$ then $X_{T\wedge n}(\omega) = X_n(\omega)$ for every n, which converges to $X_\infty(\omega)$ for almost every ω by the fact that $(X_n)_{n\in\bar{\mathbb{N}}}$ is a martingale (see Theorem 4.7). Hence, 5.9 holds. □

Uniform integrability

The best one can ask of a martingale is that it be a Doob martingale on $\bar{\mathbb{R}}_+$. Often, however, one starts with a martingale on $\mathbb{R}_+$.

5.12 Theorem. *Suppose that X is a martingale on $\mathbb{R}_+$. Then, it can be extended to a Doob martingale on $\bar{\mathbb{R}}_+$ if and only if it is uniformly integrable.*

Proof. Suppose that the martingale can be extended to a Doob martingale $\bar{X}$ on $\bar{\mathbb{R}}_+$, that is, there exists a random variable X_∞ in $\mathcal{F}_\infty$ such that $\bar{X} = (X_t)_{t\in\bar{\mathbb{R}}_+}$ is a Doob martingale on $[0, +\infty]$. Then, Corollary 5.5 implies that X is uniformly integrable.

Conversely, suppose that X is a uniformly integrable martingale on $\mathbb{R}_+$. Then, in particular, $(X_n)_{n\in\mathbb{N}}$ is a uniformly integrable martingale. By Theorem 4.7, it converges almost surely and in L^1 to an integrable random variable X_∞ in $\mathcal{F}_\infty$ (see 5.11 to the effect that $\mathcal{F}_\infty$ is the limit of $(\mathcal{F}_n)_{n\in\mathbb{N}}$ as well as of $(\mathcal{F}_t)_{t\in\mathbb{R}_+}$), and $X_n = \mathbb{E}_n X_\infty$ for every n. For t in $\mathbb{R}_+$, choose n in $\mathbb{N}$ so that $t < n$; then, since X is a martingale on $\mathbb{R}_+$,

$$X_t = \mathbb{E}_t X_n = \mathbb{E}_t\mathbb{E}_n X_\infty = \mathbb{E}_t X_\infty.$$

This shows that $\bar{X} = (X_t)_{t\in\bar{\mathbb{R}}_+}$ is a martingale and, therefore, is a Doob Martingale on $\bar{\mathbb{R}}_+$ in view of Theorem 5.8 above. □

Together with Proposition 2.7, the preceding proof shows the following as well.

5.13 THEOREM. *The process X is a uniformly integrable martingale on $\mathbb{R}_+$ if and only if*

$$X_t = \mathbb{E}_t Z, \quad t \in \mathbb{R}_+,$$

for some integrable random variable Z. Moreover, then, $X_\infty = \lim X_t$ exists almost surely and in L^1 and satisfies $X_\infty = \mathbb{E}_\infty Z$, and $\bar{X} = (X_t)_{t \in \bar{\mathbb{R}}_+}$ is a Doob martingale on $\bar{\mathbb{R}}_+$.

Stopped martingales

5.14 THEOREM. *Let ζ be a stopping time. Let $\hat{X}$ be the process X stopped at ζ, that is, $\hat{X}_t = X_{t\wedge\zeta}$ for t in $\mathbb{R}_+$.*

a) If X is a martingale, $\hat{X}$ is a martingale.

b) The process X is a Doob martingale on $[0, \zeta]$ if and only if $\hat{X}$ is a uniformly integrable martingale.

Proof. a) Suppose that X is a martingale on $\mathbb{R}_+$. Let T be a stopping time bounded by some b in $\mathbb{R}_+$. Then, $T \wedge \zeta$ is a stopping time bounded by b, and X is a Doob martingale on $[0, b]$ by Theorem 5.6, which together yield $\mathbb{E}\ X_{T\wedge\zeta} = \mathbb{E}\ X_0$ via Theorem 5.4. But, $X_{T\wedge\zeta} = \hat{X}_T$ and $X_0 = \hat{X}_0$. So,

$$\mathbb{E}\ \hat{X}_T = \mathbb{E}\ \hat{X}_0$$

for every stopping time bounded by some $b < \infty$, which implies via Theorem 5.4 that $\hat{X}$ is a Doob martingale on $[0, b]$ for every b in $\mathbb{R}_+$, which in turn implies via Theorem 5.6 that $\hat{X}$ is a martingale.

b) Necessity part of the statement (b) was shown in Remark 5.3c. To show the sufficiency part, suppose that $\hat{X}$ is a uniformly integrable martingale on $\mathbb{R}_+$. By Theorem 5.13, we can extend it to a Doob martingale on $\bar{\mathbb{R}}_+$ by defining $\hat{X}_\infty = \lim \hat{X}_t$. Then, for every stopping time T majorized by ζ, we have $X_T = \hat{X}_T$ and

$$\mathbb{E}\ X_T = \mathbb{E}\ \hat{X}_T = \mathbb{E}\ \hat{X}_0 = \mathbb{E}\ X_0$$

by Theorem 5.4 for $\hat{X}$. Thus, by 5.4 again, X is a Doob martingale on $[0, \zeta]$. □

Criteria for being Doob

The following criterion is easy to fulfill in many applications.

5.15 PROPOSITION. *Suppose that X is a martingale on $\mathbb{R}_+$. Let ζ be a stopping time. Suppose that X is dominated on $[0, \zeta] \cap \mathbb{R}_+$ by an integrable random variable. Then, the almost sure limit $X_\zeta = \lim_{t\to\infty} X_{\zeta\wedge t}$ exists and is integrable, and X is a Doob martingale on $[0, \zeta]$.*

Proof. Let X and ζ be such. Let Z be an integrable random variable such that, for almost every ω, $|X_t(\omega)| \le Z(\omega)$ for every t in $\mathbb{R}_+$ with $t \le \zeta(\omega)$. Define $\hat{X}$ to be X stopped at ζ.

By Theorem 5.14a, then, $\hat{X}$ is a martingale on $\mathbb{R}_+$. By assumption, $\hat{X}$ is dominated by the integrable random variable Z almost surely, which implies that $\hat{X}$ is uniformly integrable. Thus, the almost sure (and in L^1) limit

$$\lim_{t\to\infty} \hat{X}_t = \lim_{t\to\infty} X_{\zeta\wedge t} = X_\zeta$$

exists and is integrable. It follows from Theorem 5.14b that X is a Doob martingale on $[0, \zeta]$. □

5.16 Example. Let X be a continuous martingale. For fixed integer $n \ge 1$, let

$$\zeta_n = \inf\{t \ge 0 : |X_t| \ge n\}.$$

Then, ζ_n is a stopping time, and X is dominated by the constant n on $[0, \zeta_n] \cap \mathbb{R}_+$. The preceding theorem implies that X is a Doob martingale on $[0, \zeta_n]$.

Local martingales and semimartingales

The modern theory of stochastic analysis is built around these objects. Our aim is to provide a bridge to it by introducing the terms.

5.17 Definition. *Let ζ be a stopping time. The process X is called a* local martingale *on* $[0, \zeta)$ *if there exists an increasing sequence of stopping times ζ_n with limit ζ such that $(X_t - X_0)_{t\in\mathbb{R}_+}$ is a Doob martingale on $[0, \zeta_n]$ for every n. If it is a local martingale on $\mathbb{R}_+ = [0, \infty)$, then it is simply called a local martingale.*

Every martingale is a local martingale, because, if X is a martingale, then it is a Doob martingale on $[0, n]$ and the definition is satisfied with $\zeta_n = n$ and $\zeta = +\infty$. Of course, if X is a Doob martingale on $[0, \zeta]$, then it is a local martingale on $[0, \zeta)$ trivially (take $\zeta_n = \zeta$ for all n).

In the definition above, the sequence (ζ_n) is called a *localizing sequence*. In general, there are many localizing sequences for the same local martingale. Choosing the correct one is an art and depends on the application at hand. For example, if X is a continuous martingale as in Example 5.16, one localizing sequence is given by $\zeta_n = n$, another by the ζ_n defined there; the latter has the advantage of making X a bounded (Doob) martingale on $[0, \zeta_n]$ for each n. In general, it is worth noting that, if (ζ'_n) and (ζ''_n) are localizing sequences, then so is $(\zeta'_n \wedge \zeta''_n)$.

Localization is used in other contexts as well. For instance, a process (V_t) is said to be *locally of finite variation* if there exists a sequence (ζ_n) of stopping times increasing to $+\infty$ almost surely such that, for every ω, the path $t \mapsto V_t(\omega)$ has finite variation over the interval $[0, \zeta_n(\omega)]$ for every n.

5.18 Definition. *The process X is called a semimartingale if it can be decomposed as*

$$X = L + V$$

where L is a local martingale and V is locally of finite variation.

In the definition, it is understood that both L and V are to be adapted to the same filtration $\mathcal{F}$ to which X is adapted. The localizing sequences for L and V can differ, but it is always possible to find a sequence that localizes both.

Applications to Brownian motion

Our aim is to illustrate some uses of the foregoing with a sustained example or two. Many other problems can be solved by similar techniques. We start with the more delicate problem of showing that most hitting times of Brownian motion are almost surely finite.

Let $W = (W_t)_{t\in\mathbb{R}_+}$ be a Wiener process with respect to the filtration $\mathcal{F}$ on $\mathbb{R}_+$; see Definition 2.15. Define

5.19 $$T_a = \inf\{t > 0 : W_t \geq a\}, \quad a > 0,$$

the time W enters $[a, \infty)$ for the first time. It is a stopping time for each a.

5.20 Proposition. *For each a in $(0, \infty)$ the stopping time T_a is almost surely finite, its expected value is $+\infty$, and its distribution and the corresponding Laplace transform are as follows:*

$$\mathbb{P}\{T_a \in B\} = \int_B dt \frac{ae^{-a^2/2t}}{\sqrt{2\pi t^3}}, \; B \in \mathcal{B}_{\mathbb{R}_+}; \quad \mathbb{E}\; e^{-pT_a} = e^{-a\sqrt{2p}}, \; p \in \mathbb{R}_+.$$

Proof. Fix a, write T for T_a. Fix $p > 0$, put $r = \sqrt{2p}$, and note that $p = r^2/2$. It was shown in Proposition 2.17 that

5.21 $$X_t = \exp(rW_t - pt), \quad t \in \mathbb{R}_+,$$

defines a martingale. For arbitrary ω, if $T(\omega) < \infty$ and $t \leq T(\omega)$, then $W_t(\omega) \leq a$ by 5.19 and the continuity of W, which implies that $X_t(\omega) \leq e^{ra}$, and the same holds for $t < \infty$ when $T(\omega) = +\infty$. In other words, X is bounded by the constant e^{ra} on $[0, T] \cap \mathbb{R}_+$. Thus, by Proposition 5.15, X_T is well-defined and integrable, and X is a Doob martingale on $[0, T]$. By Theorem 5.4, then,

5.22 $$\mathbb{E}\; X_T = \mathbb{E}\; X_0 = 1.$$

On the event $\{T = +\infty\}$, we have $W_t \leq a$ for all t, which implies that $X_T = \lim X_t = 0$. Whereas, on the event $\{T < \infty\}$, we have $W_T = a$ and $X_T = \exp(ra - pT)$. Hence,

$$e^{ra-pT} = e^{ra-pT} 1_{\{T<\infty\}} = X_T 1_{\{T<\infty\}} = X_T,$$

which yields, by 5.22,

$$\mathbb{E}\ e^{-pT} = e^{-ra} = e^{-a\sqrt{2p}}$$

(recall that $r = \sqrt{2p}$). Since $p > 0$ was arbitrary, this shows that the Laplace transform for T is as claimed. Letting $p \to 0$, we see that

$$\mathbb{P}\{T < \infty\} = \lim_{p\to 0} \mathbb{E}\ e^{-pT} = 1.$$

Inverting the Laplace transform yields the claimed distribution for T. Using the distribution to compute the expectation we get $\mathbb{E}\ T = +\infty$. □

5.23 COROLLARY. *Almost surely,*

$$T_0 = \inf\{t > 0 : W_t > 0\} = 0.$$

Proof. Clearly, $0 \le T_0 \le T_a$ for every $a > 0$. Thus, for $p > 0$,

$$\mathbb{E}\ e^{-pT_0} \ge \mathbb{E}\ e^{-pT_a} = e^{-a\sqrt{2p}}$$

for every $a > 0$. This shows that the left side is equal to 1, which implies in turn that $T_0 = 0$ almost surely. □

Similarly to 5.19, we define the entrance times

5.24 $$T_a = \inf\{t > 0 : W_t \le a\}, \quad a < 0.$$

Since $(-W_t)$ is again a Wiener process, T_a and T_{-a} have the same distribution for every a, and the distribution is given by 5.20 for $a > 0$. We state this next and add a remark or two whose proofs are left as exercises.

5.25 COROLLARY. *For every non-zero a in $\mathbb{R}$, the stopping time T_a is almost surely finite, has expected value $+\infty$, and has the same distribution as a^2/Z^2 where Z is a standard Gaussian variable.*

For $a > 0$ for instance, T_a is the amount of time W spends in the interval $(-\infty, a)$ before exiting it. The interval being unbounded, $\mathbb{E}\ T_a = +\infty$. Otherwise, W exits every bounded interval in finite time with finite expected value. We show this next along with related results. Define

5.26 $$T = \inf\{t : W_t \notin (a, b)\}, \quad a < 0 < b,$$

that is, T is the time of exit from (a, b); recall that $W_0 = 0$. Obviously,

5.27 $$T = T_a \wedge T_b,$$

and we have shown above that the entrance times T_a and T_b are almost surely finite. Thus, $T < \infty$ almost surely, and W_T is either a or b, with some probabilities p_a and p_b respectively, $p_a + p_b = 1$. Since the martingale W is bounded on the time interval $[0, T]$, it is a Doob martingale on $[0, T]$. It follows that $\mathbb{E}\ W_T = \mathbb{E}\ W_0 = 0$; in other words, $ap_a + bp_b = 0$. So,

5.28 $$p_a = \mathbb{P}\{W_T = a\} = \frac{b}{b-a}, \quad p_b = \mathbb{P}\{W_T = b\} = \frac{-a}{b-a}.$$

In order to compute the expected value of the exit time T, we consider the martingale Y defined in 2.19, that is, $Y_t = W_t^2 - t$, $t \in \mathbb{R}_+$. By Theorem 5.6, it has the Doob property for the bounded stopping times 0 and $T \wedge t$, that is, $\mathbb{E}_0 Y_{T\wedge t} = Y_0 = 0$. Hence,

$$\mathbb{E}\ (T \wedge t) = \mathbb{E}\ (W_{T\wedge t})^2, \quad t \in \mathbb{R}_+.$$

Since $T \wedge t$ increases to T as $t \to \infty$, the left side increases to $\mathbb{E}\ T$ by the monotone convergence theorem. Since $(W_{T\wedge t})^2$ is bounded by $a^2 \vee b^2$, and converges to W_T^2, the right side goes to $\mathbb{E}\ W_T^2$ by the bounded convergence theorem. Hence,

5.29 $$\mathbb{E}\ T = \mathbb{E}\ W_T^2 = (-a) \cdot b, \quad a < 0 < b,$$

in view of 5.28. Incidentally, we have also shown that Y is a Doob martingale on $[0, T]$.

Finally, we specify the distribution of the time of exit from a symmetric interval by means of Laplace transforms.

5.30 Proposition. *Let T be the first time that W exits the interval $(-a, a)$, where $a > 0$ is a fixed constant. Then, $\mathbb{E}\ T = a^2$ and*

$$\mathbb{E}\ e^{-pT} = 2/\left(e^{a\sqrt{2p}} + e^{-a\sqrt{2p}}\right), \quad p \in \mathbb{R}_+.$$

Proof. Fix $p > 0$, put $r = \sqrt{2p}$, and let X be as in 5.21. Then, X is a positive martingale bounded by e^{ra} in $[0, T]$, and $T < \infty$ almost surely, and W_T is well-defined and bounded. So, X is a Doob martingale on $[0, T]$, and $\mathbb{E}\ X_T = \mathbb{E}\ X_0 = 1$ by 5.4, that is,

$$\mathbb{E}\ \exp(rW_T - pT) = 1.$$

Note that T is also the exit time from $(-a, a)$ by the process $(-W_T)$; this is because the interval is symmetric. Hence,

$$\mathbb{E}\ \exp(-rW_T - pT) = 1.$$

Adding the last two equations side by side, we get

$$\mathbb{E}\ [\exp(rW_T) + \exp(-rW_T)]\,[\exp(-pT)] = 2.$$

Whether W_T is equal to a or $-a$, the first factor inside the expectation is equal to $e^{ra} + e^{-ra}$, which constant can come out of the expectation. So,

$$(e^{ra} + e^{-ra})\mathbb{E}\ e^{-pT} = 2,$$

which yields the claimed Laplace transform once we put $r = \sqrt{2p}$. □

Exercises and Complements

5.31 *Stopped martingales.* Let S be a stopping time and let Y be the process X stopped at S.

a) If X is a Doob martingale on $[0, \zeta]$, then Y is a Doob martingale on $[0, \zeta]$. Show.

b) Use (a) to prove that if X is a martingale then Y is a martingale.

5.32 *Continuation.* If X is a local martingale on $[0, \zeta)$, then Y is a local martingale on $[0, \zeta)$. Show.

5.33 *Doob's maximal inequalities.* Let $X = (X_t)$ be a submartingale that is positive and continuous. Let

$$M_t = \max_{0 \leq r \leq t} X_s.$$

Show that, for $p \geq 1$ and $b \geq 0$,

$$b^p \, \mathbb{P}\{M_t > b\} \;\leq\; \mathbb{E}\, X_t^p.$$

Show that, if $X_t \in L^p$ for some $p > 1$, then, with $\frac{1}{p} + \frac{1}{q} = 1$,

$$\mathbb{E}\, M_t^p \;\leq\; q^p \, \mathbb{E}\, X_t^p.$$

Hint: Let $D_n = \{tk/2^n : k = 0, 1, \ldots, 2^n\}$; note that M_t is the limit, as $n \to \infty$, of $\max_{s \in D_n} X_s$; Use the discrete time results for the latter maxima.

5.34 *Convergence theorem in continuous-time.* Let X be a right-continuous submartingale. Suppose that it is L^1-bounded, that is, $\sup_{t \in \mathbb{R}_+} \mathbb{E}\, |X_t| < \infty$, which condition is equivalent to having $\lim_{t \to \infty} \mathbb{E}\, X_t^+ < \infty$. Then, the almost sure limit $X_\infty = \lim_{t \to \infty} X_t$ exists and is integrable. If X is uniformly integrable, then the convergence is in L^1 as well and $\bar{X} = (X_t)_{t \in \bar{\mathbb{R}}_+}$ is a submartingale.

5.35 *Reverse-time convergence.* Let $X = (X_t)_{t>0}$ be a right-continuous submartingale. Suppose that $\sup_{t \leq 1} \mathbb{E}|X_t| < \infty$.

a) Show that the condition is equivalent to $\lim_{t \to 0} \mathbb{E}\, X_t < \infty$.

b) Show that $\lim_{t \to 0} X_t$ exists almost surely and in L^1.

Supplements for Brownian motion

Throughout the following exercises, W is a Wiener process with respect to the filtration $\mathcal{F}$.

5.36 *Distribution of* T_a. Let T_a be as defined by 5.19. Show that T_a has the same distribution as a^2/Z^2, where Z is a standard Gaussian variable. Note that the same is true of T_a with $a < 0$ as well.

5.37 *Continuation.* Show that $\mathbb{P}\{T_a \leq t\} = \mathbb{P}\{|W_t| > a\}$.

5.38 *Maxima and minima.* Define

$$M_t = \max_{s \le t} W_s, \quad m_t = \min_{s \le t} W_s.$$

a) Show that $M_t(\omega) \ge a$ if and only if $T_a(\omega) \le t$, this being true for all $a > 0$, $t > 0$, and ω in Ω.

b) Show that M_t has the same distribution as $|W_t|$, and m_t the same distribution as $-|W_t|$.

5.39 *Continuation.* Show that the following are true for almost every ω:

a) $T_a(\omega) < \infty$ for every a in $\mathbb{R}$,
b) $t \mapsto M_t(\omega)$ is continuous, real-valued, and increasing with limit $+\infty$,
c) $t \mapsto m_t(\omega)$ is continuous, real-valued, and decreasing with limit $-\infty$.
d) $\liminf W_t(\omega) = -\infty$, $\limsup W_t(\omega) = +\infty$.
e) The set $\{t \in \mathbb{R}_+ : W_t(\omega) = 0\}$ is unbounded, that is, for every $b < \infty$ there is $t > b$ such that $W_t(\omega) = 0$. Consequently, there is a sequence (t_n), depending on ω, such that $t_n \nearrow +\infty$ and $W_{t_n}(\omega) = 0$ for every n.

5.40 *Exponential martingale.* Let $X_t = \exp(rW_t - \frac{1}{2}r^2 t)$ where r is a fixed real number. Since X is a positive martingale, $X_\infty = \lim X_t$ exists almost surely. Identify the limit. Is X uniformly integrable?

5.41 *Brownian motion.* Let $B_t = at + bW_t$, $t \ge 0$, where a and b are fixed numbers. Then, B is called a Brownian motion with drift rate a and volatility b and with $B_0 = 0$. Suppose that $a > 0$, $b > 0$, and fix $x > 0$. Show that

$$T = \inf\{t : B_t \ge x\}$$

is finite almost surely. Use the exponential martingale with $p > 0$ and $r = -\frac{a}{b} + \sqrt{\frac{a^2}{b^2} + 2p}$ to get

$$\mathbb{E}\, e^{-pT} = e^{-xr/b}, \quad \mathbb{E}\, T = x/a, \quad \text{Var}\, T_x = xb^2/a^3.$$

5.42 *Brownian motion with negative drift.* Let $a > 0$ and put $B_t = W_t - at$, $t \in \mathbb{R}_+$. For $x > 0$, let $T = \inf\{t : B_t \ge x\}$.

a) Show that

$$\mathbb{E}\, e^{-pT} = \exp(-xa - x\sqrt{a^2 + 2p}), \quad p > 0.$$

Conclude, in particular, that $\mathbb{P}\{T < \infty\} = e^{-2ax}$.

b) Let $M = \sup_{t \in \mathbb{R}_+} B_t$. Show that M has the exponential distribution with parameter $2a$.

5.43 *Exit from an interval.* With $a > 0$ and $b > 0$, put $B_t = at + bW_t$, $t \in \mathbb{R}_+$. For $x < 0 < y$, let

$$T = \inf\{t > 0 : B_t \notin (x, y)\}.$$

Show that T is almost surely finite. Compute the distribution of B_T. Compute the mean and variance of T.

5.44 *Multi-dimensional Wiener*. Let W be an n-dimensional Wiener process, that is, $W = (W^{(1)}, \dots, W^{(n)})$ where the components are independent Wiener processes. Then, $R = |W|$ is called the *radial Brownian motion*, or *Bessel process* of index n; for v in $\mathbb{R}^n$ we write $|v|$ for the length of v. For fixed $r > 0$, let

$$T = \inf\{t : |W_t| \geq r\},$$

the time of exit from the open ball of radius r centered at the origin. Show that $T < \infty$ almost surely. Show that $\mathbb{E}\, T = r^2/n$.

5.45 *Behavior near zero*. Returning back to one-dimension, show that

$$T_0 = \inf\{t > 0 : W_t > 0\} = \inf\{t > 0 : W_t < 0\} = 0$$

almost surely. Show that the following are true for almost every ω:

a) For every $\varepsilon > 0$ there is u in $(0, \varepsilon)$ such that $W_u(\omega) > 0$ and there is s in $(0, \varepsilon)$ such that $W_s(\omega) < 0$.

b) There exist strictly positive sequences (s_n), (t_n), (u_n) depending on ω such that

$$u_1 > t_1 > s_1 > u_2 > t_2 > s_2 > \dots, \quad \lim u_n = \lim t_n = \lim s_n = 0$$

and

$$W_{u_n}(\omega) > 0, \quad W_{t_n}(\omega) = 0, \quad W_{s_n}(\omega) < 0, \quad n \geq 1.$$

6 Martingale Characterizations for Wiener and Poisson

Our primary aim is to complete the proofs of Theorems 2.19 and 2.23, the characterizations of Wiener and Poisson processes in terms of martingales. We start with the Poisson case, because the needed preliminaries are of independent interest.

In this section, $\mathcal{F} = (\mathcal{F}_t)$ is a filtration on $\mathbb{R}_+$, not necessarily augmented or right-continuous. All processes are indexed by $\mathbb{R}_+$ and adapted to $\mathcal{F}$, all with state space $\mathbb{R}$. Considering a process $F = (F_t)$, we shall generally think of it as the mapping $(\omega, t) \mapsto F_t(\omega)$ from $\Omega \times \mathbb{R}_+$ into $\mathbb{R}$, and we may use the phrase "the process F on $\Omega \times \mathbb{R}_+$" to indicate that thought.

Predictability

This is the continuous-time analog of the concept introduced by Definition 3.1. We shall develop it briefly, just enough to serve our present needs.

6.1 Definition. *The σ-algebra on $\Omega \times \mathbb{R}_+$ generated by the collection*

$$\mathcal{F}^{pp} = \{H \times (s, t] : \ 0 \leq s < t < \infty, H \in \mathcal{F}_s\} \cup \{H \times \{0\} : \ H \in \mathcal{F}_0\}$$

is called the $\mathcal{F}$-predictable σ-algebra *and is denoted by* $\mathcal{F}^p$. *A stochastic process* $F = (F_t)$ *is said to be* $\mathcal{F}$-predictable *if the mapping* $(\omega, t) \mapsto F_t(\omega)$ *is* $\mathcal{F}^p$*-measurable.*

It is usual to simply say predictable instead of $\mathcal{F}$-predictable when there can be no confusion over the filtration involved, which is our present situation. The elements of $\mathcal{F}^{pp}$ are called the primitive sets, and their indicators are called *primitives*. The following proposition implies, in particular, that every adapted left-continuous process is predictable. In the converse direction, every predictable process is adapted, but might fail to be left-continuous.

6.2 PROPOSITION. *The predictable σ-algebra $\mathcal{F}^p$ is also generated by the collection $\mathcal{G}$ of all adapted left-continuous processes on $\Omega \times \mathbb{R}_+$.*

Proof. Every primitive process is adapted and left-continuous; thus, $\mathcal{F}^p \subset \sigma\mathcal{G}$. To show the converse, that $\sigma\mathcal{G} \subset \mathcal{F}^p$, we need to show that every G in $\mathcal{G}$ is predictable. Let G be in $\mathcal{G}$. Then, for every time t and outcome ω, the value $G_t(\omega)$ is the limit, as $n \to \infty$, of

$$G_t^{(n)}(\omega) = G_0(\omega) 1_{\{0\}}(t) + \sum G_a(\omega) 1_{(a,b]}(t),$$

where the sum is over all intervals $(a, b]$ with $a = k/2^n$ and $b = (k+1)/2^n$, $k \in \mathbb{N}$; this is by the left-continuity of $t \mapsto G_t(\omega)$. Thus, to show that G is predictable, it is enough to show that each $G^{(n)}$ is predictable, which in turn reduces to showing that every term of $G^{(n)}$ is predictable. But, for fixed $(a, b]$, the process $(\omega, t) \mapsto G_a(\omega) 1_{(a,b]}(t)$ is clearly predictable, since $G_a \in \mathcal{F}_a$ by the adaptedness of G, and $\mathcal{F}^p$ is generated by the primitive processes; similarly, the process $(\omega, t) \mapsto G_0(\omega) 1_{\{0\}}(t)$ is predictable. □

Martingales associated with some increasing processes

Let $N = (N_t)$ be an increasing right-continuous process adapted to the filtration $\mathcal{F}$. Let $\nu_t = \mathbb{E}\, N_t$ be finite for every t, and suppose that

6.3 $$\mathbb{E}_s (N_t - N_s) = \nu_t - \nu_s, \quad 0 \le s < t < \infty,$$

where $\mathbb{E}_s$ denotes the conditional expectation given $\mathcal{F}_s$ as usual. This is equivalent to assuming that

6.4 $$\tilde{N}_t = N_t - \nu_t, \quad t \in \mathbb{R}_+,$$

is an $\mathcal{F}$-martingale. In particular, these conditions are fulfilled when N has independent increments and $\mathbb{E}\, N_t < \infty$.

6.5 THEOREM. *For every positive predictable process F,*

6.6 $$\mathbb{E} \int_{\mathbb{R}_+} F_t \, dN_t = \mathbb{E} \int_{\mathbb{R}_+} F_t \, d\nu_t.$$

Remark. The integrals above are Stieltjes integrals; for instance, the one on the left defines a random variable V where $V(\omega)$ is the integral of $t \mapsto F_t(\omega)$ with respect to the measure on $\mathbb{R}_+$ defined by the increasing right-continuous function $t \mapsto N_t(\omega)$.

Proof. Consider the collection of all positive predictable processes F for which 6.6 holds. That collection is a monotone class: it includes the constants, it is a linear space, and it is closed under increasing limits, the last being the result of the monotone convergence theorem applied twice on the left side of 6.6 and twice on the right. Thus, the monotone class theorem will conclude the proof once we show that 6.6 is true for primitive processes, that is, the indicators of the sets in $\mathcal{F}^{pp}$.

Let F be the indicator of $H \times (a, b]$ with H in $\mathcal{F}_a$. Then, the left side of 6.6 is equal to

$$\mathbb{E}\ 1_H \cdot (N_b - N_a) = \mathbb{E}\ 1_H\ \mathbb{E}_a(N_b - N_a) = \mathbb{E}\ 1_H(\nu_b - \nu_a),$$

where the first equality uses the assumption that $H \in \mathcal{F}_a$ and the second equality uses 6.3. The last member is equal to the right side of 6.6 for the present F. Similarly, 6.6 holds when F is the indicator of $H \times \{0\}$ with H in $\mathcal{F}_0$. □

The following corollary enhances the preceding theorem and provides an example with further uses.

6.7 Corollary. *Let F be a positive predictable process. Let S and T be stopping times with $S \leq T$. Then,*

6.8
$$\mathbb{E}_S \int_{(S,T]} F_t\ dN_t = \mathbb{E}_S \int_{(S,T]} F_t\ d\nu_t.$$

Proof. It is enough to show that, for V in $\mathcal{F}_S$ and positive,

$$\mathbb{E}\ V \int_{(S,T]} F_t\ dN_t = \mathbb{E}\ V \int_{(S,T]} F_t\ d\nu_t,$$

which is in turn equivalent to showing that

6.9
$$\mathbb{E} \int_{\mathbb{R}_+} G_t F_t\ dN_t = \mathbb{E} \int_{\mathbb{R}_+} G_t F_t\ d\nu_t,$$

where

$$G_t = V\ 1_{(S,T]}(t), \quad t \in \mathbb{R}_+.$$

The process G is obviously left-continuous, and each G_t is in $\mathcal{F}_t$ by Theorem 1.16d applied to the variables $V 1_{\{S<t\}}$ and $1_{\{t \leq T\}}$ separately. It follows from Proposition 6.2 that G is predictable, and thus, so is the product GF. Hence, 6.9 follows from the preceding theorem applied with the positive predictable process GF. □

The preceding theorem and corollary are in fact equivalent to the following theorem about the martingale $\tilde{N}$ defined by 6.4.

6.10 THEOREM. *Let F be a bounded predictable process. Then,*

$$M_t = \int_{[0,t]} F_s \; d\tilde{N}_s, \quad t \in \mathbb{R}_+,$$

is a martingale.

Proof. It is obvious that M is adapted. Each M_t is integrable since $|M_t| \leq (N_t + \nu_t)b$ if b is a bound for $|F|$. To show the martingale property that $\mathbb{E}_s(M_t - M_s) = 0$ for $s < t$, it is sufficient to show that

6.11 $$\mathbb{E}_s \int_{(s,t]} F_u \; dN_u = \mathbb{E}_s \int_{(s,t]} F_u \; d\nu_u \; ;$$

this is because $\tilde{N} = N - \nu$ and both sides of 6.11 are real-valued. Now, 6.11 is immediate from Corollary 6.7 applied first with F^+ and then with F^-. □

REMARK. *Stochastic integrals.* The preceding theorem remains true for arbitrary martingales $\tilde{N}$, except that the proof above is no longer valid and, worse, the integral defining M has to be given a new meaning. Above, since $\tilde{N}$ has paths of finite variation over bounded intervals, the integral defining M makes sense ω by ω, that is,

6.12 $$M_t(\omega) = \int_{[0,t]} F_s(\omega) \; d\tilde{N}_s(\omega).$$

But, if $\tilde{N}$ were an arbitrary martingale or, more specifically, if $\tilde{N}$ were a Wiener process, then the paths $s \mapsto \tilde{N}_s(\omega)$ would necessarily have infinite variation over most intervals and, hence, the integral 6.12 has no meaning as a Stieltjes integral for general F. Stochastic calculus goes around the problem by defining M as the limit in probability of a sum of primitive integrals. With this new meaning for the integral M, the conclusion of the last theorem remains true. The interested reader should see a book on stochastic calculus.

Martingale characterization of Poisson processes

Here, we prove Theorem 2.23. The necessity part was already done preceding 2.23. For easy reference, we list next what is to be proved.

6.13 PROPOSITION. *Let N be a counting process adapted to $\mathcal{F}$. Suppose that, for some constant c in $(0, \infty)$,*

$$\tilde{N}_t = N_t - ct, \quad t \in \mathbb{R}_+,$$

is a martingale. Then, N is a Poisson process with rate c.

We start with a lemma of independent interest; it exploits the counting nature of N. Here, $N_{s-} = \lim_{r \uparrow s} N_r$ as usual.

6.14 LEMMA. *Let f be a bounded function on $\mathbb{N}$. Then,*

$$f(N_t) = f(0) + \int_{(0,t]} [f(N_{s-}+1) - f(N_{s-})]\ dN_s.$$

Proof. Fix an ω. If $N_t(\omega) = 0$ then the claim is obvious. If $N_t(\omega) = n \geq 1$, let $t_1, \ldots, t_n$ be the successive jump times of $s \mapsto N_s(\omega)$ during $(0, t]$, suppressing their dependence on ω. At t_i, the counting function $s \mapsto N_s(\omega)$ jumps from the left-hand limit $i-1$ to the right-hand value i. Thus, the right side of the claimed equation is equal to, for this ω,

$$f(0) + \sum_{i=1}^{n} [f(i-1+1) - f(i-1)] = f(n) = f(N_t(\omega)).$$

□

Proof of Proposition 6.13. This is an application of Corollary 6.7 with a carefully chosen F. Fix times $s < t$, fix r in $\mathbb{R}_+$, and let

$$f(n) = 1 - e^{-rn}, \quad n \in \mathbb{N}; \quad F_u = f(N_{u-}+1) - f(N_{u-}), \quad u \in \mathbb{R}_+.$$

Since $u \mapsto N_{u-}$ is adapted and left-continuous, so is F. Thus, F is bounded, positive, and predictable, the last following from Proposition 6.2. Hence, by Corollary 6.7 applied with the current N and $\nu_t = ct$,

$$\mathbb{E}_s \int_{(s,t]} F_u\ dN_u = c\ \mathbb{E}_s \int_{(s,t]} F_u\ du.$$

The integral on the left is equal to $f(N_t) - f(N_s)$ by Lemma 6.14. As to the Lebesgue integral on the right side, replacing F_u with $F_{u+} = f(N_u + 1) - f(N_u)$ will not change the integral since $F_u = F_{u+}$ for all u in $(s, t]$ except for finitely many u. Hence,

$$\mathbb{E}_s[f(N_t) - f(N_s)] = c\ \mathbb{E}_s \int_{(s,t]} [f(N_u+1) - f(N_u)]\ du.$$

Now we replace t with $s + t$, recall that $f(n) = 1 - e^{-rn}$, and multiply both sides by $\exp rN_s$. The result is (writing $\exp_{\_}\, x$ for e^{-x})

6.15 $$\mathbb{E}_s\ \exp_{\_}\, r(N_{s+t} - N_s) = 1 - c(1-e^{-r})\mathbb{E}_s \int_0^t du\ \exp_{\_}\, r(N_{s+u} - N_s).$$

Let the left side be denoted by $g(t)$, suppressing its dependence on r, s, ω. We have

$$g(t) = 1 - c(1 - e^{-r}) \int_0^t g(u)\ du,$$

whose only solution is

$$g(t) = \exp\_ \ ct(1 - e^{-r}) = \sum_{k=0}^{\infty} \frac{e^{-ct}(ct)^k}{k!} \ e^{-rk}$$

totally free of s and ω. We have shown that

$$\mathbb{E}_s \exp\_ \ r(N_{s+t} - N_s) = \sum_{k=0}^{\infty} \frac{e^{-ct}(ct)^k}{k!} \ e^{-rk}.$$

Since this is true for arbitrary r in $\mathbb{R}_+$, we conclude that $N_{s+t} - N_s$ is independent of $\mathcal{F}_s$ and has the Poisson distribution with mean ct. This concludes the proof hat N is a Poisson process with rate c. □

6.16 Remark. *Strong Markov property.* The preceding proof can be modified to show that, for the process N,

$$\mathbb{E}_S \ \exp\_ \ r \cdot (N_{S+t} - N_S) = \sum_{k=0}^{\infty} \frac{e^{-ct}(ct)^k}{k!} \ e^{-rk},$$

that is, for every finite stopping time S, the future increment $N_{S+t} - N_S$ is independent of $\mathcal{F}_S$ and has the same Poisson distribution as N_t has. This is called the *strong Markov property* for N. To show it, replace s by S and t by $S + t$ from the beginning of the proof of 6.13, and note that Corollary 6.7 applies full force. This brings us to 6.15 with s replaced by S; and the rest is exactly the same but with s replaced by S.

Non-stationary Poisson processes

These are defined just as Poisson processes except that the distribution of $N_{s+t} - N_s$ has the Poisson distribution with mean $\nu_{s+t} - \nu_s$, where ν is an arbitrary continuous increasing function (the stationary case is where $\nu_t = ct$). In other words, a counting process N adapted to $\mathcal{F}$ is said to be a (*non-stationary*) *Poisson process with mean* ν if ν is continuous increasing real-valued, and, for every positive function f on $\mathbb{N}$,

6.17
$$\mathbb{E}_s f(N_t - N_s) = \sum_{k=0}^{\infty} \frac{e^{-a} a^k}{k!} \ f(k)$$

with $a = \nu_t - \nu_s$; compare this with 6.16 and 2.20. Of course, then, $N - \nu$ is a martingale. The following states this and adds a converse.

6.18 Theorem. *Let N be a counting process adapted to $\mathcal{F}$, and let ν be a (deterministic) increasing continuous real-valued function on $\mathbb{R}_+$ with $\nu_0 = 0$. Then, N is a Poisson process with mean function ν if and only if $N - \nu$ is a martingale.*

Proof. Necessity is trivial. We prove the sufficiency. Given ν, let $\nu_\infty = \lim_{t\to\infty} \nu_t$ and let τ be the functional inverse of ν, that is,

6.19 $$\tau_u = \inf\{t > 0 : \nu_t > u\}, \quad u < \nu_\infty;$$

see Exercise 5.13 of Chapter I. Then, τ is right-continuous and strictly increasing on $[0, \nu_\infty)$, and $\nu_{\tau_u} = u$ by the continuity of ν. Clearly, (N_{τ_u}) is adapted to the filtration $(\mathcal{F}_{\tau_u})$ and is again a counting process. Since $N - \nu$ is assumed to be an $\mathcal{F}$-martingale, the process $(N_{\tau_u} - u)$ is an $(\mathcal{F}_{\tau_u})$-martingale on $[0, \nu_\infty)$. By Proposition 6.13, then, the process N_τ is a Poisson process with rate 1 on the interval $[0, \nu_\infty)$, that is, for every positive function f on $\mathbb{N}$,

6.20 $$\mathbb{E}_{\tau_u} f(N_{\tau_v} - N_{\tau_u}) = \sum_{k=0}^{\infty} \frac{e^{-(v-u)}(v-u)^k}{k!}\ f(k)$$

for $0 \le u < v < \nu_\infty$. There remains to undo the time change 6.19.

We start by observing that, if $\nu_s = \nu_t$ for some $s < t$, then $\mathbb{E}\,(N_t - N_s) = \nu_t - \nu_s = 0$ and thus $N_t - N_s = 0$ almost surely. It follows that, for $0 \le s < t$,

$$N_s = N(\tau(\nu_s)), \quad N_t = N(\tau(\nu_t))$$

almost surely. Hence, taking $u = \nu_s$ and $v = \nu_t$ in 6.20, and putting $a = v - u = \nu_t - \nu_s$, we get

$$\mathbb{E}_{\tau(\nu_s)} f(N_t - N_s) = \sum_{k=0}^{\infty} \frac{e^{-a} a^k}{k!}\ f(k).$$

Finally, apply the conditional expectation operator $\mathbb{E}_s$ on both sides; observing that $\tau(\nu_s) \ge s$ necessarily by the definition 6.19, we get 6.17 with $a = \nu_t - \nu_s$, which completes the proof that N is Poisson with mean ν. □

Martingale characterization of Wiener processes

This is to give the sufficiency part of the proof of Theorem 2.19; recall that the proof of necessity was already given. We list what is to be proved for convenience.

6.21 PROPOSITION. *Let X be a continuous $\mathcal{F}$-martingale with $X_0 = 0$. Suppose that $Y = (X_t^2 - t)_{t\in\mathbb{R}_+}$ is again an $\mathcal{F}$-martingale. Then, X is a Wiener process with respect to $\mathcal{F}$.*

We start by listing a lemma, whose highly technical proof will be given below, after the proof of 6.21.

6.22 LEMMA. *Let X be as in 6.21. Let f be a twice differentiable function on $\mathbb{R}$ and suppose that f is bounded along with its derivative f' and its second derivative f''. Then,*

6.23 $$M_t = f\circ X_t - \tfrac{1}{2}\int_0^t ds\ f''\circ X_s, \quad t \in \mathbb{R}_+,$$

defines a martingale M.

Proof of Proposition 6.21. We use the preceding lemma first with $f(x) = \cos rx$ and then with $f(x) = \sin rx$ to conclude that M defined by 6.23 with $f(x) = e^{irx} = \cos rx + i \sin rx$ is a complex-valued martingale. In other words, since $f''(x) = -r^2 f(x)$ when $f(x) = e^{irx}$,

$$\mathbb{E}_s \left[f \circ X_{s+t} - f \circ X_s + \tfrac{1}{2} r^2 \int_s^{s+t} f \circ X_u \, du \right] = 0.$$

Replacing $f(x)$ with e^{irx}, multiplying both sides by $\exp(-irX_s)$, and rearranging the result, we obtain

$$\mathbb{E}_s \exp ir(X_{s+t} - X_s) = 1 - \tfrac{1}{2} r^2 \int_0^t du \; \mathbb{E}_s \exp ir(X_{s+u} - X_s).$$

Let $g(t)$ be defined to be the left side, ignoring its dependence on s and r and ω. We get

$$g(t) = 1 - \tfrac{1}{2} r^2 \int_0^t du \; g(u),$$

whose only solution is $g(t) = \exp(-r^2 t/2)$ independent of s and ω. So, for every r in $\mathbb{R}$,

$$\mathbb{E}_s \exp ir(X_{s+t} - X_s) = \exp(-r^2 t/2),$$

which shows that the increment $X_{s+t} - X_s$ is independent of $\mathcal{F}_s$ and has the Gaussian distribution with mean 0 and variance t; in short, X is Wiener. □

We turn to proving Lemma 6.22. The sophisticated reader will have noticed that it is a simple consequence of Itô's lemma, but we do not have access to such advanced machinery at this point. Instead, the proof is by necessity of the bare-hands type. It is well-worth ignoring it, except for purposes of admiring Lévy and Doob for their power and ingenuity.

Proof of Lemma 6.22. a) We shall eventually show that $\mathbb{E}_a(M_b - M_a) = 0$ for $0 \le a < b < \infty$. Fix a and b such, fix $\varepsilon > 0$, take an integer $n \ge 1$, and let $\delta = (b - a)/n$. Define

$$T = b \wedge \inf\{t \ge a : \max_{a \le p,q \le t,\ |q-p| \le \delta} |X_q - X_p| = \varepsilon\}.$$

Since $t \mapsto X_t(\omega)$ is continuous, it is uniformly continuous on $[a, b]$, and hence, $T(\omega)$ is equal to b for all n large enough, depending on ε and the outcome ω. By the continuity of X,

$$\{T > t\} = \bigcup_m \bigcap_{p,q} \left\{ |X_q - X_p| < \varepsilon - \frac{1}{m} \right\},$$

where the union is over all the integers $m \ge 1$ and the intersection is over all rationals p and q in $[a, t]$ with $|q - p| \le \delta$. Since X is adapted, this shows that T is a stopping time. Consequently, since X and Y are martingales by hypothesis, so are the processes obtained from them by stopping at T, denoted by

6.24 $$Z = (Z_t) = (X_{t\wedge T}), \quad \bar{Z} = (\bar{Z}_t) = (Z_t^2 - t\wedge T).$$

It follows that, for s and t in $[a,b]$ with $0 < t-s \le \delta$,

6.25 $$\mathbb{E}_s(Z_t - Z_s) = 0, \quad |Z_t - Z_s| \le \varepsilon,$$

6.26 $$\begin{aligned}\mathbb{E}_s(Z_t - Z_s)^2 &= \mathbb{E}_s[Z_t^2 - 2Z_s(Z_t - Z_s) - Z_s^2] \\ &= \mathbb{E}_s(Z_t^2 - 0 - Z_s^2) = \mathbb{E}_s(t\wedge T - s\wedge T).\end{aligned}$$

b) Lef f be as described. We shall use Taylor's theorem in the following form

6.27 $$f(y) - f(x) = f'(x)(y-x) + \frac{1}{2}f''(x)(y-x)^2 + r(x,y),$$

where the remainder term is such that, for some continuous increasing function h on $\mathbb{R}_+$ with $h(0) = 0$,

6.28 $$|r(x,y)| \le (y-x)^2 h(|y-x|).$$

c) Keeping $a, b, \varepsilon, n, \delta$ as before, let $\mathcal{D}$ be the subdivision of the interval $(a,b]$ into n disjoint intervals of equal length and of form $(s,t]$, that is,

$$\mathcal{D} = \{(s,t]: \quad s = a + k\delta, t = s+\delta, k = 0,1,\ldots,n-1\}.$$

Using 6.27, with $\sum_{\mathcal{D}}$ indicating summation over all $(s,t]$ in $\mathcal{D}$,

6.29 $$\begin{aligned} f\circ Z_b - f\circ Z_a &= \sum_{\mathcal{D}} [f\circ Z_t - f\circ Z_s] \\ &= \sum_{\mathcal{D}} [(f'\circ Z_s)(Z_t - Z_s) + \frac{1}{2}(f''\circ Z_s)(Z_t - Z_s)^2] + R,\end{aligned}$$

where the remainder term R satisfies, in view of 6.25 and 6.28,

6.30 $$|R| \le \sum_{\mathcal{D}} (Z_t - Z_s)^2 \; h\circ|Z_t - Z_s| \le h(\varepsilon)\sum_{\mathcal{D}}(Z_t - Z_s)^2.$$

We now apply $\mathbb{E}_a$ to both sides of 6.29 and 6.30, using $\mathbb{E}_a\mathbb{E}_s = \mathbb{E}_a$ repeatedly for $a \le s$ and using the equalities in 6.25 and 6.26. We get

6.31 $$\mathbb{E}_a f\circ Z_b - f\circ Z_a = \tfrac{1}{2}\mathbb{E}_a \textstyle\sum_{\mathcal{D}}(f''\circ Z_s)(t\wedge T - s\wedge T) + \mathbb{E}_a R,$$

6.32 $$|\mathbb{E}_a R| \le h(\varepsilon)\mathbb{E}_a \textstyle\sum_{\mathcal{D}}(t\wedge T - s\wedge T) \le h(\varepsilon)\mathbb{E}_a(b-a) \le h(\varepsilon)(b-a).$$

Consider the sum over $(s,t]$ in $\mathcal{D}$ on the right side of 6.31. For $(q,r]$ in $\mathcal{D}$, on the event $\{q < T \le r\}$, we have $Z_s = X_s$ for $s \le q$ and the sum is equal to

$$\delta \sum_{\mathcal{D}} f''\circ X_s - Q$$

where

$$|Q| = |(f'' \circ X_q)(r - T) + \delta \sum_{s>q,\mathcal{D}} f'' \circ X_s| \le \|f''\| \cdot (b - T)$$

with $\|f''\| = \sup_x |f''(x)| < \infty$ by assumption. Hence, recalling that $a \le T \le b$ we can re-write 6.31 and 6.32 as follows:

6.33 $$\mathbb{E}_a f \circ X_{T\wedge b} - f \circ X_a = \tfrac{1}{2}\mathbb{E}_a \textstyle\sum_{\mathcal{D}} \delta f'' \circ X_s - \tfrac{1}{2}\mathbb{E}_a Q + \mathbb{E}_a R,$$

6.34 $$|\mathbb{E}_a Q| \le \|f''\| \mathbb{E}_a (b - T), \quad |\mathbb{E}_a R| \le (b - a) h(\varepsilon).$$

d) Keeping a, b, ε as before, we now let $n \to \infty$. Then, T increases to b as mentioned earlier. So, $T \wedge b \to b$, and $X_{T\wedge b} \to X_b$ by the continuity of X, which implies that $f \circ X_{T\wedge b} \to f \circ X_b$ by the continuity of f, and hence,

6.35 $$\mathbb{E}_a f \circ X_{T\wedge b} \to \mathbb{E}_a f \circ X_b$$

by the bounded convergence theorem (recall that f is bounded). Again as $n \to \infty$, on the right side of 6.33, the sum over $\mathcal{D}$ converges to the Riemann integral of $f'' \circ X_s$ over $[a, b]$, and

6.36 $$\mathbb{E}_a \sum_{\mathcal{D}} \delta f'' \circ X_s \to \mathbb{E}_a \int_a^b f'' \circ X_u \, du$$

by the bounded convergence theorem, since the sum remains dominated by $\|f''\| \cdot (b - a)$ for all n. Finally, since T increases to b, $\mathbb{E}_a (b - T) \to 0$ by the bounded convergence theorem, which yields

6.37 $$|\mathbb{E}_a Q| \to 0$$

in view of 6.34. Putting 6.35, 6.36, 6.37 into 6.33 and noting the bound for $\mathbb{E}_a R$ in 6.34, we obtain

$$|\mathbb{E}_a (M_b - M_a)| = \left| \mathbb{E}_a f \circ X_b - f \circ X_a - \frac{1}{2}\mathbb{E}_a \int_a^b du \; f'' \circ X_u \right| \le (b - a) h(\varepsilon).$$

This shows that M is a martingale, since a, b, ε are arbitrary and $h(\varepsilon)$ decreases to 0 as $\varepsilon \to 0$. □

7 Standard Filtrations and Modifications of Martingales

This section is to supplement the chapter by discussing the right-continuity and augmentation of filtrations, and the beneficial consequences of such properties on stopping times and martingales. Throughout, $(\Omega, \mathcal{H}, \mathbb{P})$ is the probability space in the background, the time-set is $\mathbb{R}_+$ unless specified otherwise, and all filtrations and processes are indexed by $\mathbb{R}_+$.

Augmentation

Let $\mathcal{F} = (\mathcal{F}_t)$ be a filtration. It is said to be *augmented* if $(\Omega, \mathcal{H}, \mathbb{P})$ is complete and all the negligible events in $\mathcal{H}$ are also in $\mathcal{F}_0$ (and, therefore, in every $\mathcal{F}_t$).

Suppose that $(\Omega, \mathcal{H}, \mathbb{P})$ is complete. Let $\mathcal{F}$ be an arbitrary filtration. Let $\mathcal{N}$ be the collection of all negligible events in $\mathcal{H}$ and let $\bar{\mathcal{F}}_t$ be the σ-algebra generated by $\mathcal{F}_t \cup \mathcal{N}$. Then, $\bar{\mathcal{F}} = (\bar{\mathcal{F}}_t)$ is an augmented filtration and is called the *augmentation* of $\mathcal{F}$. Obviously, $\mathcal{F}$ is augmented if and only if $\mathcal{F} = \bar{\mathcal{F}}$.

Right-continuity

Let $\mathcal{F}$ be a filtration. We define

7.1 $$\mathcal{F}_{t+} = \cap_{\varepsilon>0} \mathcal{F}_{t+\varepsilon}, \quad t \in \mathbb{R}_+.$$

Then, $(\mathcal{F}_{t+})$ is again a filtration and is finer than $(\mathcal{F}_t)$. The filtration $\mathcal{F}$ is said to be right-continuous if

7.2 $$\mathcal{F}_t = \mathcal{F}_{t+}$$

for every t in $\mathbb{R}_+$. Note that $(\mathcal{F}_{t+})$ is itself a right-continuous filtration; it is the coarsest right-continuous filtration that is finer than $\mathcal{F}$.

Heuristically, $\mathcal{F}_{t+}$ has the same information as $\mathcal{F}_t$ plus the information gained by an "infinitesimal peek" into the future. For instance if $t \mapsto X_t$ depicts a smooth motion and $\mathcal{F}$ is the filtration generated by $X = (X_t)$, then $\mathcal{F}_t$ has all the information regarding the past of X and the present position X_t, whereas $\mathcal{F}_{t+}$ has all that information plus the velocity $V_t = \lim_{\varepsilon\to 0}(X_{t+\varepsilon} - X_t)/\varepsilon$ and acceleration at t and so on.

Sometimes, the difference between $(\mathcal{F}_{t+})$ and $(\mathcal{F}_t)$ is so slight that the augmentation of $(\mathcal{F}_t)$ is right-continuous and therefore finer than $(\mathcal{F}_{t+})$. We shall see several instances of it, especially with Lévy processes and Brownian motion. The following illustrates this in a simple case.

7.3 Example. Let T and V be positive random variables with diffuse distributions on $\mathbb{R}_+$. Define

$$X_t(\omega) = \begin{cases} V(\omega)t & \text{if } t < T(\omega), \\ V(\omega)T(\omega) & \text{if } t \geq T(\omega). \end{cases}$$

The process $X = (X_t)$ describes the motion of a particle that starts at the origin at time 0, moves with speed V until the time T, and stops at time T. Let $\mathcal{F}$ be the filtration generated by X. Note that T is not a stopping time of $\mathcal{F}$, the reason being that knowing $X_s(\omega) = V(\omega)s$ for all $s \leq t$ is insufficient to tell whether $T(\omega) = t$ or $T(\omega) > t$. But, the event $\{T \leq t\}$ is definitely in $\mathcal{F}_{t+\varepsilon}$ for every $\varepsilon > 0$, and thus, is in $\mathcal{F}_{t+}$; in other words, T is a stopping time of $(\mathcal{F}_{t+})$. The failure of T to be a stopping time of $(\mathcal{F}_t)$ is due to a negligible cause: the event $\{T = t\}$ is negligible by our assumption that the distribution of T is diffuse. Hence, letting $(\bar{\mathcal{F}}_t)$ be the augmentation of $(\mathcal{F}_t)$, we conclude that T is a stopping time of $(\bar{\mathcal{F}}_t)$ and that $\bar{\mathcal{F}}_t \supset \mathcal{F}_{t+}$ for all t except $t = 0$.

Stopping times and augmentation

Let $(\mathcal{F}_t)$ be a filtration and let T be a random time, a mapping from Ω into $\bar{\mathbb{R}}_+$. Suppose that $\mathcal{F}$ is augmented and S is a stopping time of it. If $T = S$ almost surely, then T is a stopping time of $\mathcal{F}$: For each time t, the event $\{S \le t\}$ belongs to $\mathcal{F}_t$, and $\{T \le t\}$ and $\{S \le t\}$ differ from each other by negligible events, and those negligible events are in $\mathcal{F}_t$ by our assumption that $\mathcal{F}$ is augmented.

Stopping times and right-continuity

Right-continuity of a filtration simplifies the tests for its stopping times. This is a corollary of the following.

7.4 THEOREM. *Let $\mathcal{F}$ be a filtration, and T a random time. Then, T is a stopping time of $(\mathcal{F}_{t+})$ if and only if*

7.5 $$\{T < t\} \in \mathcal{F}_t \text{ for every } t \text{ in } \mathbb{R}_+.$$

Proof. Let $\varepsilon_n = 1/n$, $n \ge 1$. If 7.5 holds, then

$$\{T \le t\} = \cap_n \{T < t + \varepsilon_n\} \in \cap_n \mathcal{F}_{t+\varepsilon_n} = \mathcal{F}_{t+}$$

for every t, which means that T is a stopping time of $(\mathcal{F}_{t+})$. Conversely, if T is a stopping times of $(\mathcal{F}_{t+})$, then $\{T \le s\} \in \mathcal{F}_{s+} \subset \mathcal{F}_t$ for all $s < t$, and hence,

$$\{T < t\} = \cup_n \{T \le t - \varepsilon_n\} \in \mathcal{F}_t. \qquad \square$$

If $\mathcal{F}$ is right-continuous, then $\mathcal{F}_t = \mathcal{F}_{t+}$ for all t, and the preceding theorem shows that T is a stopping time of $\mathcal{F}$ if and only if 7.5 holds.

7.6 EXAMPLE. Let W be a Wiener process and let $\mathcal{F}$ be the filtration generated by it. For fixed $a > 0$, let

$$T = \inf\{t > 0 : \; W_t > a\}.$$

Then, T is not a stopping time of $(\mathcal{F}_t)$, but it is a stopping time of $(\mathcal{F}_{t+})$. The latter assertion follows from the preceding theorem, because $T(\omega) < t$ if and only if $W_r(\omega) > a$ for some positive rational number $r < t$. The former assertion follows from observing that $T(\omega) = t$ if and only if $W_r(\omega) \le a$ for all rationals $r < t$ and $r = t$ and $W_r(\omega) > a$ for some rational $r > t$, and the last part with $r > t$ cannot be told at the time t.

Past until T

Let $\mathcal{F}$ be a filtration. Let T be a stopping time of the filtration $(\mathcal{F}_{t+})$, and let $\mathcal{F}_{T+}$ denote the corresponding past until T, that is, $\mathcal{F}_{T+} = \mathcal{G}_T$ where

$(\mathcal{G}_t) = (\mathcal{F}_{t+})$. More explicitly , recalling the definition 1.9 with the filtration $(\mathcal{G}_t)$,

7.7 $$\mathcal{F}_{T+} = \{H \in \mathcal{H} : \ H \cap \{T \le t\} \in \mathcal{F}_{t+} \text{ for every } t \text{ in } \mathbb{R}_+\}.$$

In fact, the arguments of the proof of 7.4 shows that

7.8 $$\mathcal{F}_{T+} = \{H \in \mathcal{H} : \ H \cap \{T < t\} \in \mathcal{F}_{t} \text{ for every } t \text{ in } \mathbb{R}_+\}.$$

Of course, if $\mathcal{F}$ is right-continuous, 7.7 shows that $\mathcal{F}_{T+} = \mathcal{F}_T$, and 7.8 becomes another characterization for $\mathcal{F}_T$.

Sequences of stopping times

Let (T_n) be a sequence of stopping times of a filtration $\mathcal{F}$. If the sequence is increasing, its limit is a stopping time of $\mathcal{F}$; see Exercise 1.34. The following contains the best that can be said about decreasing sequences.

7.9 PROPOSITION. *Let (T_n) be a sequence of stopping times of $(\mathcal{F}_t)$ or of $(\mathcal{F}_{t+})$. Then, $T = \inf T_n$ is a stopping time of $(\mathcal{F}_{t+})$, and*

$$\mathcal{F}_{T+} = \cap_n \mathcal{F}_{T_n+}.$$

Proof. Since $(\mathcal{F}_t)$ is coarser than $(\mathcal{F}_{t+})$, every stopping time of the former is a stopping time of the latter. So, the T_n are stopping times of $(\mathcal{F}_{t+})$ in either case. By Theorem 7.4, the event $\{T_n < t\}$ is in $\mathcal{F}_t$ for every n and every t. It follows that

$$\{T < t\} = \cup_n \{T_n < t\} \in \mathcal{F}_t$$

for every t, that is, T is a stopping time of $(\mathcal{F}_{t+})$ in view of 7.4.

Since $T \le T_n$ for every n, Theorem 1.16b applied with the filtration $(\mathcal{F}_{t+})$ shows that $\mathcal{F}_{T+} \subset \mathcal{F}_{T_n+}$ for every n. Hence, $\mathcal{F}_{T+} \subset \cap_n \mathcal{F}_{T_n+}$. To show the converse, let H be an event that belongs to $\mathcal{F}_{T_n+}$ for every n. Then,

$$H \cap \{T < t\} = \cup_n (H \cap \{T_n < t\}) \in \mathcal{F}_t$$

in view of 7.8 applied with T_n. Thus, by 7.8 again, $H \in \mathcal{F}_{T+}$. □

If $\mathcal{F}$ is right-continuous, and (T_n) is a sequence of stopping times of it, then the infimum and supremum and limit inferior and limit superior of the sequence are all stopping times.

Times foretold

Let $\mathcal{F}$ be a filtration. Recall that a random time T is said to be *foretold* by a stopping time S of $\mathcal{F}$ if $T \ge S$ and $T \in \mathcal{F}_S$. Obviously, if S is a stopping time of $(\mathcal{F}_{t+})$ and $T \ge S$ and $T \in \mathcal{F}_{S+}$, then T is a stopping time of $(\mathcal{F}_{t+})$. The following is a refinement.

7.10 PROPOSITION. *Let S be a stopping time of $(\mathcal{F}_{t+})$. Let T be a random time such that $T \in \mathcal{F}_{S+}$ and $T \geq S$, with strict inequality $T > S$ on the event $\{S < \infty\}$. Then, T is a stopping time of $(\mathcal{F}_t)$ and $\mathcal{F}_{S+} \subset \mathcal{F}_T$.*

Proof. Let S and T be as described. For every outcome ω and time t, if $T(\omega) \leq t$ then $S(\omega) < T(\omega)$ and $S(\omega) < t$. Thus, for H in $\mathcal{H}$,

7.11 $$H \cap \{T \leq t\} = H \cap \{T \leq t\} \cap \{S < t\}.$$

Suppose that $H \in \mathcal{F}_{S+}$. Since $T \in \mathcal{F}_{S+}$ by assumption, then, the left side is in $\mathcal{F}_{S+}$, which implies that the right side is in $\mathcal{F}_t$ in view of 7.8 for $\mathcal{F}_{S+}$. Thus, the left side of 7.11 is in $\mathcal{F}_t$ whenever $H \in \mathcal{F}_{S+}$. Taking $H = \Omega$ shows that T is a stopping time of $(\mathcal{F}_t)$, and we conclude that every H in $\mathcal{F}_{S+}$ is in $\mathcal{F}_T$. □

Approximation of stopping times

The following shows that Proposition 1.20 remains true for stopping times T of $(\mathcal{F}_{t+})$. Here, the d_n are as defined by 1.19.

7.12 PROPOSITION. *Let T be a stopping time of $(\mathcal{F}_{t+})$. Let $T_n = d_n \circ T$ for each n in $\mathbb{N}$. Then, each T_n is a stopping time of $(\mathcal{F}_t)$, each T_n is discrete, and the sequence (T_n) decreases to T and decreases strictly on the set $\{T < \infty\}$. Moreover, $\mathcal{F}_{T+} = \cap_n \mathcal{F}_{T_n}$.*

Proof. The proof is immediate from Propositions 7.9 and 7.10 once we note that each T_n is foretold by T. □

Hitting times

Augmented right-continuous filtrations are desirable for the simplifications noted above and for the following important result, which we list here without proof.

Let $X = (X_t)$ be a stochastic process with state space $(E, \mathcal{E})$, where E is topological and $\mathcal{E}$ is the Borel σ-algebra on E. Let $\mathcal{F}$ be a filtration.

7.13 THEOREM. *Suppose that $\mathcal{F}$ is right-continuous and augmented. Suppose that X is right-continuous and is adapted to $\mathcal{F}$. Then, for every Borel subset B of E, the hitting time*

$$T_B = \inf\{t \in \mathbb{R}_+ : \ X_t \in B\}$$

is a stopping time of $\mathcal{F}$.

Regularity of martingales

We start with a filtration $\mathcal{F}$ on $\mathbb{R}_+$. In the following, D is an arbitrary countable subset of $\mathbb{R}_+$ which is dense in $\mathbb{R}_+$, for example, one can take D to be the set of all rationals in $\mathbb{R}_+$.

7.14 PROPOSITION. *Let X be an $\mathcal{F}$-submartingale on $\mathbb{R}_+$. For almost every ω, the following limits exist and are in $\mathbb{R}$:*

7.15 $$X_{t+}(\omega) = \lim_{r\in D,\ r\downarrow t} X_r(\omega), \quad t \geq 0,$$

7.16 $$X_{t-}(\omega) = \lim_{r\in D,\ r\uparrow t} X_r(\omega), \quad t > 0.$$

Proof. Fix s in D. Let a and b be rational numbers with $a < b$. Let B be a finite subset of $D \cap [0, s]$ that includes the end point s. Then, $(X_r)_{r\in B}$ is a submartingale with respect to $(\mathcal{F}_r)_{r\in B}$ with a discrete-time set B. By Theorem 3.21 applied to the submartingale X on B,

7.17 $$c\,\mathbb{P}\{\max_{r\in B} X_r \geq c\} \leq \mathbb{E}\,|X_s|.$$

Next, let $U_B(a, b)$ be the number of upcrossings of the interval (a, b) by the process $(X_r)_{r\in B}$. By Proposition 3.19,

7.18 $$(b - a)\mathbb{E}\,U_B(a, b) \leq \mathbb{E}\,(X_s - a)^+ < \infty.$$

Note that the right sides of 7.17 and 7.18 are free of B. Thus, by taking supremums over all finite sets B that include s and are contained in $D\cap[0, s]$, we see that the same inequalities hold for

$$M_s = \sup_{r\in D\cap[0,s]} |X_r|, \quad U_s(a, b) = \sup_B U_B(a, b)$$

respectively. It follows that $M_s < \infty$ and $U_s(a, b) < \infty$ almost surely.

For s in D, let Ω_s be the collection of all ω for which the limits $X_{t-}(\omega)$ exist and are finite for all t in $(0, s]$ and the limits $X_{t+}(\omega)$ exist and are finite for all t in $[0, s)$. Observe that

$$\Omega_s \supset \cap_{a,b}\{M_s < \infty, U_s(a, b) < \infty\},$$

where the intersection is over all pairs (a, b) of rationals with $a < b$. For each s in D, this shows that Ω_s contains an almost sure event. Hence, $\Omega_0 = \cap_{s\in D}\Omega_s$ contains an almost sure event, and for every ω in Ω_0 the limits indicated exist and are in $\mathbb{R}$. □

7.19 PROPOSITION. *Suppose that $\mathcal{F}$ is right-continuous and augmented, and let X be a $\mathcal{F}$-submartingale. Let Ω_0 be the almost sure set of all ω for which the limits 7.15 and 7.16 exist in $\mathbb{R}$, and set $X_{t-}(\omega) = X_{t+}(\omega) = 0$ for every ω outside Ω_0.*

a) For each t in $\mathbb{R}_+$, the random variable X_{t+} is integrable and

$$X_t \leq X_{t+}$$

almost surely. Here, the equality holds almost surely if and only if $s \mapsto \mathbb{E}\,X_s$ is right-continuous at t (in particular, if X is a martingale).

b) *The process* $(X_{t+})_{r\in\mathbb{R}_+}$ *is a submartingale with respect to* $\mathcal{F}$, *and it is a martingale if* X *is so. Moreover, for every outcome* ω, *the trajectory* $t \mapsto X_{t+}(\omega)$ *is right-continuous and left-limited.*

Proof. Since $\mathcal{F}$ is augmented, the set Ω_0 belongs to $\mathcal{F}_0$ and to $\mathcal{F}_t$ for all t; thus the alteration outside Ω_0 does not change the adaptedness of X to $\mathcal{F}$, and the altered X is still an $\mathcal{F}$-submartingale.

a) Fix t in $\mathbb{R}_+$. Let (r_n) be a sequence in D decreasing strictly to t. Then, (X_{r_n}) is a reversed time submartingale, and $\mathbb{E}\ X_t \le \mathbb{E}\ X_{r_n}$ for every n. By Theorem 4.19, the sequence (X_{r_n}) is uniformly integrable and converges to X_{t+} almost surely and in L^1. It follows that X_{t+} is integrable and, for every event H in $\mathcal{F}_t$,

7.20 $$\mathbb{E}\ X_{t+}1_H = \lim\ \mathbb{E}\ X_{r_n}1_H \ge \mathbb{E}\ X_t 1_H,$$

the inequality being through the submartingale inequality for $t < r_n$. Thus,

7.21 $$\mathbb{E}_t(X_{t+} - X_t) \ge 0.$$

Since X_{r_n} belongs to $\mathcal{F}_{t+\varepsilon}$ for every $\varepsilon > 0$ and all n large enough, the limit X_{t+} belongs to $\mathcal{F}_{t+}$, and $\mathcal{F}_{t+} = \mathcal{F}_t$ by the assumed right-continuity for $\mathcal{F}$. Thus, the left side of 7.21 is equal to $X_{t+} - X_t$, which proves the claim that $X_{t+} \ge X_t$ almost surely. The equality would hold almost surely if and only if $\mathbb{E}\ X_{t+} = \mathbb{E}\ X_t$, which in turn holds if and only if $s \mapsto \mathbb{E}\ X_s$ is right-continuous at t (in which case $\mathbb{E}\ X_t = \lim\ \mathbb{E}\ X_{r_n} = \mathbb{E}\ X_{t+}$).

b) The paths $t \mapsto X_{t+}(\omega)$ are right-continuous and left-limited by the way they are defined. To see that (X_{t+}) is an $\mathcal{F}$-submartingale, take $s < t$, choose (r_n) in D strictly decreasing to t, and (q_n) in D strictly decreasing to s, ensuring that $s < q_n < t < r_n$ for every n. Then, for H in $\mathcal{F}_s$, using 7.20 twice, we get

$$\mathbb{E}\ X_{s+}1_H = \lim\ \mathbb{E}\ X_{q_n}1_H \le \lim\ \mathbb{E}\ X_{r_n}1_H = \mathbb{E}\ X_{t+}1_H,$$

where the inequality is through the submartingale property of X. This completes the proof. □

Modifications of martingales

The following is an immediate corollary of the last theorem: put $Y_t(\omega) = X_{t+}(\omega)$ for every t and every ω.

7.22 THEOREM. *Suppose that* $\mathcal{F}$ *is right-continuous and augmented. Suppose that* X *is an* $\mathcal{F}$*-submartingale and* $t \mapsto \mathbb{E}\ X_t$ *is right-continuous. Then, there exists a process* Y *such that*

a) *for every* ω, *the path* $t \mapsto Y_t(\omega)$ *is right-continuous and left-limited,*

b) *Y is an $\mathcal{F}$-submartingale,*
c) *for every t in $\mathbb{R}_+$, we have $\mathbb{P}\{X_t = Y_t\} = 1$.*

The preceding theorem is the justification for limiting our treatment in Section 5 to right-continuous processes. Note that, if X is a martingale, the right-continuity of $\mathbb{E}\ X_t$ in t is immediate, since $\mathbb{E}\ X_t = \mathbb{E}\ X_0$ for all t. The process Y is said to be a *modification* of X because of the statement (c), which matter we clarify next.

Modifications and indistinguishability

Let $\mathbb{T}$ be some index set. Let $X = (X_t)_{t\in\mathbb{T}}$ and $Y = (Y_t)_{t\in\mathbb{T}}$ be stochastic processes with the same state space. Then, one is said to be a *modification* of the other if, for each t in $\mathbb{T}$,

$$\mathbb{P}\{X_t = Y_t\} = 1.$$

They are said to be *indistinguishable* if

$$\mathbb{P}\{X_t = Y_t \text{ for all } t \text{ in } \mathbb{T}\} = 1.$$

For example, in Theorem 7.22, the last assertion is that Y is a modification of X; it does not mean that they are indistinguishable.

Suppose that X is a modification of Y. Then, for every integer $n < \infty$ and indices $t_1, \ldots, t_n$ in $\mathbb{T}$, the vectors $(X_{t_1}, \ldots, X_{t_n})$ and $(Y_{t_1}, \ldots, Y_{t_n})$ are almost surely equal. It follows that the two vectors have the same distribution. In other words, X and Y have the same finite-dimensional distributions, and therefore, they have the same probability law.

If X and Y are indistinguishable, then they are modifications of each other. If they are modifications of each other, and if the index set $\mathbb{T}$ is countable, then they are indistinguishable. Otherwise, in general, indistinguishability requires more than being modifications of each other.

For instance, suppose $\mathbb{T} = \mathbb{R}_+$ and the state space is $\mathbb{R}^d$. Suppose that X and Y are modifications of each other and are both right-continuous. Then they are indistinguishable.

Exercises

7.23 *Right-continuity.* In Example 7.3, describe $\mathcal{F}_0$ and $\mathcal{F}_{0+}$. Let $\bar{\mathcal{F}}$ be the augmentation of $\mathcal{F}$. Show that $\bar{\mathcal{F}}_{t+} = \bar{\mathcal{F}}_t$ for $t > 0$.

7.24 *Stopping times.* Show that a random time T is a stopping time of $(\mathcal{F}_{t+})$ if and only if the process $(T \wedge t)_{t\in\mathbb{R}_+}$ is adapted to $(\mathcal{F}_t)$.

7.25 *Strict past at T.* Recall from Exercise 1.31 that, for a random time T, the strict past at T is defined to be the σ-algebra generated by events of the form $H \cap \{t < T\}$ with H in $\mathcal{F}_t$ and $t \in \mathbb{R}_+$. Then, T belongs to $\mathcal{F}_{T-}$ and $\mathcal{F}_{T-} \subset \mathcal{F}_T$.

7.26 *Continuation.* Suppose that $\mathcal{F}$ is right-continuous and augmented. Let S and T be stopping times of it. Show that, for every H in $\mathcal{F}_S$, the event $H \cap \{S < T\}$ belongs to $\mathcal{F}_{T-}$. In particular, $\{S < T\} \in \mathcal{F}_{T-}$. Show that $\{S < T\}$ belongs to $\mathcal{F}_S \cap \mathcal{F}_{T-}$.

7.27 *Debuts and stopping times.* Let $\mathcal{F}$ be right-continuous and augmented as in Theorem 7.13. For $A \subset \mathbb{R}_+ \times \Omega$, let

$$D_A(\omega) = \inf\{t \in \mathbb{R}_+ : \ (t, \omega) \in A\}, \quad \omega \in \Omega.$$

If the process (X_t) defined by $X_t(\omega) = 1_A(t, \omega)$ is progressive in the sense of 1.15, then D_A is a stopping time of $\mathcal{F}$. Theorem 7.13 is a special case of this remark.

7.28 *Hitting times.* For fixed $a > 0$, let T be defined as in Example 7.6. Show that $T = T_a$ almost surely, where T_a is as defined by 5.19.

7.29 *Continuation.* Let T_a be defined by 5.19 and let S_a be the T defined in 7.6. Show that $a \mapsto T_a$ is left-continuous and $a \mapsto S_a$ is right-continuous. In fact, $T_a = S_{a-}$, the left-limit at a of S; Show this. The process (S_a) is a right-continuous modification of (T_a). They are eminently distinguishable.

7.30 *Predictable stopping times.* Let $\mathcal{F}$ be right-continuous and augmented. Let T be a stopping time of it. Then, T is said to be predictable if the set $\{(\omega, t) : \ t \geq T(\omega)\}$ belongs to the predictable σ-algebra; see 6.1 for the latter. Equivalently, T is said to be *predictable* if there exists a sequence (T_n) of stopping times such that, for every ω for which $T(\omega) > 0$, the sequence of numbers $T_n(\omega)$ increases strictly to $T(\omega)$. If W is a Wiener process and $\mathcal{F}$ is generated by it, then every stopping time of $\mathcal{F}$ is predictable.

7.31 *Classification of stopping times.* In addition, a stopping time T is said to be *σ-predictable* if there exists a sequence of predictable stopping times T_n such that, for every ω, we have $T(\omega) = T_n(\omega)$ for some n. Finally, T is said to be *totally unpredictable* if $\mathbb{P}\{T = S\} = 0$ for every predictable stopping time S. In Example 1.7, suppose that N is a Poisson process. Then, $T_1, T_2, \ldots$ are all totally unpredictable, the time T is predictable. The stopping time $T \wedge T_1$ is neither predictable nor totally unpredictable; it is equal to the predictable time T on the event $\{T < T_1\}$ and to the totally unpredictable time T_1 on the event $\{T > T_1\}$. This example is instructive. In general, for every stopping time T there exist a σ-predictable stopping time R and a totally unpredictable stopping time S such that $T = R \wedge S$. In the case of standard Markov processes, S is a jump time and R is a time of continuity for the process.