CHAPTER 1

Introduction

1.1. Continuous Time Processes

A stochastic process $\{x(t) : t \in \mathbb{T}\}$ is a collection $x(t) = x(t, \omega)$ of random variables defined on some probability space $(\Omega, \mathcal{F}, P)$ indexed by $t \in \mathbb{T}$. The random variable $x(t)$ can take its values in an arbitrary measurable space which is often $\mathbb{R}$ or $\mathbb{R}^d$ with its respective Borel σ-fields $\mathcal{B}(\mathbb{R})$ or $\mathcal{B}(\mathbb{R}^d)$. The variable t usually represents time. When $\mathbb{T}$ is a countable set, for instance the set of integers or a subset of it, we call the process a discrete parameter or a discrete time process. In case $\mathbb{T}$ is a finite or infinite interval in $\mathbb{R}$, it is called a continuous-time process.

There is one essential difference between the two cases. In either case, to begin with, we only know that $x(t, \cdot)$ is measurable for each t. If, for example, $x(t, \omega)$ is real valued and $\mathbb{T}$ is countable this automatically implies the measurability of sets of the form

$$A = \{\omega : x(t, \omega) \in [a, b] \ \forall\, t \in \mathbb{T}\} \tag{1.1}$$

If $\mathbb{T}$ is uncountable this is not necessarily so and for good reason. Random variables $x(t)$ are objects that are only defined almost everywhere and we are free to change them on a set of measure zero. If we do change $x(t)$ on a set $E_t \subset \Omega$ with $P(E_t) = 0$, events like (1.1) that depend on $x(t)$ for all $t \in \mathbb{T}$, could be changed on $E = \bigcup_{t \in \mathbb{T}} E_t$. If $\mathbb{T}$ is uncountable $P(E)$ need not be a set of measure zero and therefore for the set A given by (1.1), $P(A)$ may no longer be uniquely defined. Since we are often interested in calculating probabilities of events like (1.1), we need to fix this.

Let us assume that $x(t, \omega)$ is real valued. One can think of $x(t, \omega)$ as a map of $\Omega \to \mathcal{X}$ where $\mathcal{X} = \prod_{t \in \mathbb{T}} \mathbb{R}$ is the space of $\mathbb{R}$-valued of functions on $\mathbb{T}$. On $\mathcal{X}$ there is the natural σ-field $\mathcal{K}$, generated by sets of the form $\{x(\cdot) : x(t) \in A\}$ with $t \in \mathbb{T}$ and $A \in \mathcal{B}(\mathbb{R}^d)$. It is not hard to see that if $x(t, \omega)$ is measurable for each $t \in \mathbb{T}$, then $\omega \to x(\cdot, \omega)$ is a measurable map of $(\Omega, \mathcal{F}) \to (\mathcal{X}, \mathcal{K})$. Hence there is an induced measure Q on $(\mathcal{X}, \mathcal{K})$. Such measures Q are uniquely determined by the collection $\{Q_{t_1, t_2, \ldots, t_n}\}$ of joint distributions on $\mathbb{R}^n$ of $(x(t_1), x(t_2), \ldots, x(t_n))$ as n varies over all positive integers and $(t_1, t_2, \ldots, t_n)$ varies over all subsets of $\mathbb{T}$ of cardinality n.

The construction of a probability measure Q on $(\mathcal{X}, \mathcal{K})$ from a consistent family of finite-dimensional distributions $\{Q_{t_1, t_2, \ldots, t_n}\}$ when $\mathbb{T}$ is uncountable is no different from the situation when $\mathbb{T}$ is countable. After all, the crucial step is to check countable additivity on the field of finite-dimensional cylinder sets with a measurable base. If $\{A_n\}$ is from this field then each A_n, being a finite-dimensional cylinder set, depends only on the values of $x(t)$ with t from a finite subset $F_n \subset \mathbb{T}$. We can always replace $\mathbb{T}$ by the set $\mathbb{T}_0 = \bigcup_n F_n$. Since $\mathbb{T}_0$ is countable, checking the countable additivity for the sequence A_n reduces to the countable or discrete-parameter case proved in [**7**].

While in the countable case events like (1.1) are always in $\mathcal{K}$, when $\mathbb{T}$ is uncountable this is not necessarily so. One has to cut down the space Ω to a suitable subset Ω_0, such that when restricted to Ω_0 these sets are measurable, i.e. belong to the natural restriction of $\mathcal{K}$ to Ω_0. When $\mathbb{T}$ is a subset of $\mathbb{R}$, if Ω_0 were to consist of all continuous maps from $\mathbb{T} \to \mathbb{R}$, then events like (1.1) are clearly representable as countable intersections by restricting t to the rationals in $\mathbb{T}$. According to standard results in measure theory, even when $\Omega_0 \subset \Omega$ is not in $\mathcal{K}$, if the outer measure, defined for arbitrary sets E by

$$P^*(E) = \inf_{\substack{A \supset E \\ A \in \mathcal{K}}} P(A)$$

satisfies $P^*(\Omega_0) = 1$, then the restriction of P^* to the σ-field $\mathcal{K}_0$ of subsets of Ω_0 which are intersections of sets in $\mathcal{K}$ with Ω_0, i.e. sets of the form $A \cap \Omega_0$ with $A \in \mathcal{K}$, is a countably additive probability measure. With a suitable choice of Ω_0 this could provide a more convenient model $(\Omega_0, \mathcal{K}_0, P^*)$ with the same finite-dimensional distributions for $\{x(t,\omega)\}$ as the original one. The condition $P^*(\Omega_0) = 1$ is to be interpreted as P being supported on Ω_0 is not inconsistent with the given collection of finite-dimensional distributions. Of course it is often not inconsistent with P being supported on Ω_0^c either. We have the option to choose a 'good' model or a 'bad' one. We will now prove the result concerning restriction of the outer measure to full sets.

Let $(\Omega, \mathcal{F}, P)$ be a probability space. We can define for any $E \subset \Omega$ the outer measure

$$P^*(E) = \inf_{\substack{B \supset E \\ B \in \mathcal{F}}} P(B)$$

It is obvious that if $A \in \mathcal{F}$, then $P^*(A) = P(A)$. Let $\Omega_0 \subset \Omega$. Consider sets of the form $\Omega_0 \cap A$ where $A \in \mathcal{F}$. They constitute a σ-field $\mathcal{F}_0$ of subsets of Ω_0.

THEOREM 1.1 *If $P^*(\Omega_0) = 1$, then the restriction of P^* to $\mathcal{F}_0$ is a countably additive probability measure such that $P^*(\Omega_0 \cap A) = P(A)$ for $\forall A \in \mathcal{F}$.*

PROOF: Outer measures are always countably sub-additive. This is easily seen. Let $E = \bigcup_j E_j$. Let $\epsilon > 0$ be given. By definition we can find $B_j \in \mathcal{F}$ such that $B_j \supset E_j$ and $P(B_j) \le P^*(E_j) + \frac{\epsilon}{2^j}$. Clearly $B = \bigcup_j B_j \supset \bigcup_j E_j = E$ and

$$P^*(E) \le P(B) \le \sum_j P(B_j) \le \sum_j P^*(E_j) + \sum_j \frac{\epsilon}{2^j} = \sum_j P^*(E_j) + \epsilon$$

Since $\epsilon > 0$ is arbitrary the assertion follows. In order to prove that P^* is countably additive on $\mathcal{F}_0$, we need to prove that if $A_1, A_2 \in \mathcal{F}$ are such that $A_1 \cap \Omega_0$ and $A_2 \cap \Omega_0$ are disjoint then

$$P^*[(\Omega_0 \cap A_1) \cup (\Omega_0 \cap A_2)] = P^*[\Omega_0 \cap A_1] + P^*[\Omega_0 \cap A_2]$$

which amounts to proving

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) \tag{1.2}$$

Since $\Omega_0 \cap A_1$ and $\Omega_0 \cap A_2$ are disjoint, $A_1 \cap A_2 \subset \Omega_0^c$ and $P(A_1 \cap A_2) = 0$. It follows that $P(A_1 \cup A_2) = P(A_1) + P(A_2)$, i.e. (1.2) holds. □

We will not pursue this but will adopt a slightly different point of view. Most of the processes we deal with will be almost surely either continuous or continuous except for jumps. We will therefore find it more convenient to put the probability measure directly on a nice space consisting of such functions. The space $C[a, b]$ will consist of all continuous functions on the interval $[a, b]$. On the other hand the space $D[a, b]$ will consist of functions that have left and right limits at every point. These functions $x(t)$ are normalized so that the value at any t is equal to the limit $x(t + 0)$ from the right, except when the point t is the right end point b, in which case the value is equal to the left limit $x(b - 0)$. Since these functions are all determined by their values on rationals, there are no problems regarding the measurability of events like (1.1).

The spaces $C[a, b]$ and $D[a, b]$ can be provided with metrics turning them into nice Polish (i.e. complete separable metric) spaces. While $C[a, b]$ is a standard Banach space with its norm $\|f\| = \sup_{t\in[a,b]} |f(t)|$, the metric for $D[a, b]$ is unusual and the corresponding topology is known as the *Skorohod topology*. See [**5**] for a description and properties of $D[a, b]$ under this metric. The σ-field of Borel sets in either case is generated by $\{x(t) : t \in [a, b]\}$. Of course the same σ-field is generated by $x(t)$ with t is restricted to $\mathbb{Q} \cap [a, b]$, the rationals in $[a, b]$.

1.2. Continuous Parameter Martingales

Martingales that depend on a continuous parameter will play an important role in the study of continuous-parameter Markov processes just as discrete-parameter martingales play a very useful role in the theory of Markov chains. We start with some definitions.

DEFINITION 1.1 $(\Omega, \mathcal{F})$ is a space with a σ-field. A family $\{\mathcal{F}_t : t \geq 0\}$ of sub-σ-fields of $\mathcal{F}$ is called a *filtration* if $\mathcal{F}_s \subset \mathcal{F}_t$ if $s \leq t$.

If Ω is a space of functions on $\mathbb{T} \subset \mathbb{R}^+$, it comes with a natural filtration $\mathcal{F}_t = \sigma\{x(s) : s \leq t\}$.

DEFINITION 1.2 Given a probability space $(\Omega, \mathcal{F}, P)$ and a filtration $\mathcal{F}_t \subset \mathcal{F}$ indexed by $t \geq 0$, a family $M(t) = M(t, \omega)$ of random variables is called a *martingale relative to* $(\Omega, \mathcal{F}_t, P)$ if the following are true.

(i) For almost all ω, $M(t, \omega)$ has left and right limits at every t and is continuous from the right, i.e. $M(t, \omega) = M(t + 0, \omega)$
(ii) For each $t \geq 0$, $M(t)$ is $\mathcal{F}_t$ measurable and integrable.
(iii) For $0 \leq s \leq t$,

$$E\{M(t) \mid \mathcal{F}_s\} = M(s) \text{ a.e.} \tag{1.3}$$

Most of the results proved for discrete-parameter martingales extend to the continuous-parameter case. One notes that a continuous-parameter martingale $M(t)$ evaluated at increasing times $\{t_j\}$ yields a discrete-parameter martingale $X_j = M(t_j)$ with respect to the σ-fields $\{\mathcal{F}_{t_j}\}$.

EXERCISE 1.1. Extend to continuous-parameter martingales Doob's inequality. That is, if $M(t)$ is a martingale and E_ℓ is the event

$$E_\ell = \big\{\omega : \sup_{0\le s\le t} |M(s,\omega)| \ge \ell\big\}$$

then,

$$P[E_\ell] \le \frac{1}{\ell}\int_{E_\ell} |M(t)|\,dP \le \frac{1}{\ell}\int |M(t)|\,dP \tag{1.4}$$

$$P[E_\ell] \le \frac{1}{\ell^2}\int_{E_\ell} |M(t)|^2\,dP \le \frac{1}{\ell^2}\int |M(t)|^2\,dP \tag{1.5}$$

DEFINITION 1.3 Let $\mathcal{F}_t$ be a filtration and $x(t)$ a right continuous process with left limits, defined for t in $\mathbb{R}^+$ or an interval $[0, T]$. The process $x(t)$ is assumed to be $\mathcal{F}_t$ measurable and integrable for each t. It is called a *submartingale* if it satisfies

$$E[x(t) \mid \mathcal{F}_s] \ge x(s) \text{ a.e.} \tag{1.6}$$

and a supermartingale if it satisfies

$$E[x(t) \mid \mathcal{F}_s] \le x(s) \text{ a.e.} \tag{1.7}$$

EXERCISE 1.2. Show that if $M(t)$ is a martingale with respect to a filtration then $|M(t)|$ is a submartingale with respect to the same filtration. In fact, if $\phi(x)$ is a convex function of x and $\phi(M(T))$ is integrable, then $\phi(M(t))$ is integrable for $0 \le t \le T$ and is a submartingale.

We can define stopping times and their associated σ-fields in a manner analogous to the discrete case.

DEFINITION 1.4 A *stopping time* for the filtration $\{\mathcal{F}_t\}$ is a random variable $\tau(\omega) : \Omega \to [0, +\infty]$ that satisfies for every $t \ge 0$,

$$E_t = \{\omega : \tau(\omega) \le t\} \in \mathcal{F}_t \tag{1.8}$$

DEFINITION 1.5 Given a stopping time τ, the corresponding σ-field $\mathcal{F}_\tau$ is defined as consisting of those sets $A \in \mathcal{F}$ that satisfy for each $t \ge 0$,

$$A \cap \{\omega : \tau(\omega) \le t\} \in \mathcal{F}_t \tag{1.9}$$

The following properties of stopping times are left as exercises.

EXERCISE 1.3. If τ_1 and τ_2 are stopping times so are $\tau_1 \wedge \tau_2$ and $\tau_1 \vee \tau_2$.

EXERCISE 1.4. If $f(\cdot)$ is a measurable function and $f(t) \ge t$ for all t, then for any stopping time τ, $f(\tau)$ is a stopping time as well. In particular, if τ is a stopping time so is $\tau_k = \frac{[k\tau]+1}{k}$ for any positive integer k, where $[y]$ is the largest integer not exceeding y. This shows that any bounded stopping time τ can be approximated by a sequence $\{\tau_k\}$ of stopping times with $\tau_k \geq \tau$ and each τ_k taking only a finite set of values.

EXERCISE 1.5. If $0 \leq \tau_1 \leq \tau_2$ are two stopping times then $\mathcal{F}_{\tau_1} \subset \mathcal{F}_{\tau_2}$.

We can extend Doob's optional stopping theorem to continuous-time martingales.

THEOREM 1.2 *If $0 \leq \tau_1 \leq \tau_2 \leq C$ are two bounded stopping times and $M(t)$ is a martingale then*

$$E[M(\tau_2) \mid \mathcal{F}_{\tau_1}] = M(\tau_1) \; a.e. \tag{1.10}$$

First we establish a lemma that we will need.

LEMMA 1.3 *$X(\omega)$ is an integrable random variable on a probability space $(\Omega, \mathcal{F}, P)$ and $X_\Sigma = E[X \mid \Sigma]$ where $\Sigma \subset \mathcal{F}$ is a sub-σ-field of $\mathcal{F}$. The collection $\{X_\Sigma\}$ as Σ varies over all sub-σ-fields of $\mathcal{F}$ is uniformly integrable.*

PROOF: From Jensen's inequality

$$E[|X| \mid \Sigma] \geq |X_\Sigma| \text{ a.e.}$$

and therefore

$$P[|X_\Sigma| \geq \ell] \leq \frac{E[|X_\Sigma|]}{\ell} \leq \frac{E[|X|]}{\ell}$$

On the other hand the set $\{\omega : X_\Sigma(\omega) \geq \ell\} \in \Sigma$ and

$$\int_{|X_\Sigma| \geq \ell} |X_\Sigma| dP \leq \int_{|X_\Sigma| \geq \ell} |X| \, dP$$

can be made small if $P[|X_\Sigma| \geq \ell]$ is small, because $|X|$ is integrable. □

PROOF: (of Theorem 1.2). It is enough to prove that for any stopping time τ bounded by C

$$E[M(C) \mid \mathcal{F}_\tau] = M(\tau) \tag{1.11}$$

It would then follow that

$$E[M(\tau_2) \mid \mathcal{F}_{\tau_1}] = E[E[M(C) \mid \mathcal{F}_{\tau_2}] \mid \mathcal{F}_{\tau_1}] = E[M(C) \mid \mathcal{F}_{\tau_1}] = M(\tau_1)$$

To prove (1.11) we note that it suffices to prove it for some approximating sequence $\{\tau_k\}$ that satisfies $\tau_k \downarrow \tau$ and $\tau_k \leq C$. Then if $A \in \mathcal{F}_\tau$, it will also be true that $A \in \mathcal{F}_{\tau_k}$ and

$$\int_A M(\tau_k) dP = \int_A M(C) dP$$

We can let $k \to \infty$. By lemma 1.3 $\{M(\tau_k)\}$ is uniformly integrable and it follows that

$$\int_A M(\tau) dP = \int_A M(C) dP$$

We can take $\tau_k = \frac{[k\tau]+1}{k}$ where $[x]$ is the largest integer not exceeding x. Each stopping time τ_k in our approximating sequence takes only a finite set of possible values. For τ that takes only a finite set $t_1 < t_2 < \cdots < t_k$ of values, if $E_i = \{\omega : \tau(\omega) = t_i\}$ and $A \in \mathcal{F}_\tau$, then $A \cap E_i \in \mathcal{F}_{t_i}$ and

$$\int_{A\cap E_i} M(\tau)dP = \int_{A\cap E_i} M(t_i)dP = \int_{A\cap E_i} M(C)dP$$

Summing over i gives

$$\int_A M(\tau)dP = \int_A M(C)dP$$

□

COROLLARY 1.4 *In particular, if τ is any stopping time and $M(t)$ is a martingale with respect to $(\Omega, \mathcal{F}_t, P)$, then so is $M(\tau \wedge t)$.*

REMARK 1.1. Often it is necessary to apply Doob's optional stopping theorem to unbounded stopping times. We can always approximate τ by $\tau \wedge T$ which is bounded and then let $T \to \infty$. It is now a question of establishing the uniform integrability of $\{M(t \wedge T)\}$. If we have a martingale that is uniformly bounded this is not a problem.

REMARK 1.2. In a similar manner one can establish analogues of the upcrossing inequality and martingale convergence theorem for the continuous case without much additional work.

Doob's optional stopping theorem works for sub and supermartingales too.

THEOREM 1.5 *If $M(t)$ is a submartingale and $\tau_1 \le \tau_2 \le C$ are two bounded stopping times then*

$$E[M(\tau_2) \mid \mathcal{F}_{\tau_1}] \ge M(\tau_1) \tag{1.12}$$

The proof of Theorem 1.2 will not work. It will only show that

$$E[M(\tau_C) \mid \mathcal{F}_{\tau_i}] \ge M(\tau_i)$$

for $i = 1, 2$, which will not do.

Actually a slightly different proof of Theorem 1.2 is more useful. Suppose $\tau_1 \le \tau_2 \le C$ are two bounded stopping times that take only a finite set $t_1 < t_2 < \ldots < t_k$ of values. Let us denote by $E_i = \{\omega : \tau_1 = t_i\}$ and $F_i = \{\omega : \tau_2 = t_i\}$. It is clear that $E_i \cap F_j = \varnothing$ unless $i \le j$. We can then write E_i as $E_i = \bigcup_{j \ge i}(E_i \cap F_j)$. Now if $A \in \mathcal{F}_{\tau_1}$, for any i we begin by writing E_i as the disjoint union of $E_i \cap F_i$ and $E_i \cap F_i^{\mathsf{c}}$.

$$\begin{aligned}
\int_{A\cap E_i} M(\tau_1)dP &= \int_{A\cap E_i} M(t_i)dP \\
&= \int_{A\cap E_i\cap F_i} M(t_i)dP + \int_{A\cap E_i\cap F_i^{\mathsf{c}}} M(t_i)dP \\
&= \int_{A\cap E_i\cap F_i} M(t_i)dP + \int_{A\cap E_i\cap F_i^{\mathsf{c}}} M(t_{i+1})dP
\end{aligned}$$

Here we have used the fact that $A \cap E_i \cap F_i^{\mathsf{c}} \in \mathcal{F}_{t_i}$. The process can be repeated to yield

$$\int_{A\cap E_i} M(\tau_1)dP = \sum_{j\geq i} \int_{A\cap E_i \cap F_j} M(t_j)dP = \int_{A\cap E_i} M(\tau_2)dP$$

and summing over i we get

$$\int_A M(\tau_1)dP = \int_A M(\tau_2)dP$$

or

$$E[M(\tau_2) \mid \mathcal{F}_{\tau_1}] = M(\tau_1) \text{ a.e.}$$

The advantage of this proof is that if $M(t)$ were a submartingale instead of a martingale we would have the inequality at every step and end up with

$$E[M(\tau_2) \mid \mathcal{F}_{\tau_1}] \geq M(\tau_1) \text{ a.e.}$$

for stopping times $\tau_1 \leq \tau_2$ that take only a finite set of values. It is then easy to extend this inequality to arbitrary bounded stopping times and thereby prove Theorem 1.5.

PROOF: If $\tau_1 \leq \tau_2 \leq C$ are bounded stopping times, replacing τ_2 by $\tau_2+\delta$ one can assume that we have approximations $\tau_{1,r}$, $\tau_{2,k}$ with the following properties: for any two sufficiently large choices of k and r we have $\tau_{2,k} \geq \tau_{1,r}$ and $\tau_{1,r}$, $\tau_{2,k}$ are nonincreasing and tend to τ_1, τ_2, respectively. From the inequality

$$E[M(\tau_{2,k}) \mid \mathcal{F}_{\tau_{1,r}}] \geq M(\tau_{1,r}) \text{ a.e.}$$

by letting $r \to \infty$ while keeping k fixed, using the standard martingale convergence theorem, we can deduce that

$$E[M(\tau_{2,k}) \mid \mathcal{F}_{\tau_1}] \geq M(\tau_1) \text{ a.e.}$$

In order to let $k \to \infty$ and conclude that

$$E[M(\tau_2) \mid \mathcal{F}_{\tau_1}] \geq M(\tau_1) \text{ a.e.}$$

we will need the uniform integrability of the positive parts $\{M(\tau_{2,k})^+\}$ which would imply that for $A \in \mathcal{F}_{\tau_1}$,

$$\int_A M(\tau_2)dP \geq \limsup_{k\to\infty} \int_A M(\tau_{2,k})dP \geq \int_A M(\tau_1)dP$$

This uniform integrability follows from the inequality

$$M(\tau_{2,k}) \leq E[M(C) \mid \mathcal{F}_{\tau_{2,k}}]$$

and Lemma 1.3. The positive parts of $M(\tau_{2,k})$ are dominated by a uniformly integrable family and hence they are themselves uniformly integrable. A similar argument allows us to let $\delta \to 0$ at the end. □

1.3. Semimartingales

Let $(\Omega, \mathcal{F}, P)$ be a probability space and $\mathcal{F}_j$ a filtration. If $\{X_j : j \geq 0\}$ is any discrete-parameter process, X_j being $\mathcal{F}_j$ measurable and integrable, we can define

$$a_j(\omega) = E[X_j - X_{j-1} \mid \mathcal{F}_{j-1}]$$

and for $n \geq 1$

$$A_n(\omega) = \sum_{j=1}^{n} a_j(\omega)$$

with $A_0(\omega) = 0$. It is easy to verify that

$$Y_n = X_n - A_n$$

is $\mathcal{F}_n$ measurable and is a martingale. The process X_n is called a semimartingale (any process of integrable random variables is a semimartingale if time is discrete) and A_n is called the compensator. The decomposition

$$X_n = Y_n + A_n$$

is called the Doob decomposition. Note that A_n is $\mathcal{F}_{n-1}$ measurable. This is important. Otherwise one can take $A_n = X_n$ and $Y_n = 0$. If X_n is a submartingale, then $a_n \geq 0$ or A_n is increasing. For supermartingales X_n we will have $a_n \leq 0$ and A_n decreasing.

EXERCISE 1.6. Let $(\Omega, \mathcal{F}_j, P)$ be a filtration and $\{X_j : j \geq 0\}$ any integrable $\mathcal{F}_j$ measurable process. Show that a decomposition of the form $X_j = Y_j + A_j$, with $A_0 = 0$ and A_j being $\mathcal{F}_{j-1}$ measurable, is unique.

The continuous analogue of this result which is called the Doob-Meyer decomposition has several problems. Most but not all submartingales and supermartingales can be written as $X(t) = M(t) + A(t)$, with $A(t)$ being either increasing or decreasing. The natural definition of a semimartingale would be that the process can be written as $X(t) = M(t) + A(t)$ where $A(t)$ is a process which is of bounded variation. But in the discrete case A_n was $\mathcal{F}_{n-1}$ measurable and the difference was important. While it is natural for $A(t)$ to be $\mathcal{F}_t$ measurable what is the analogue of $\mathcal{F}_{n-1}$ in the continuous case?

This is not a simple question and we will not try to answer it here. However if we can find a decomposition of the form $X(t) = M(t) + A(t)$, with all processes adapted to the filtration $(\Omega, \mathcal{F}_t, P)$, where $M(t)$ is a martingale and $A(t)$ is an almost surely *continuous process of bounded variation*, then such a decomposition is unique. A continuous version of the compensator may not always exist. But, if it exists, it is unique. Proving uniqueness is essentially the following lemma.

LEMMA 1.6 *If* $\{A(t) : t \in [0, T]\}$ *is continuous-parameter martingale satisfying*

(i) *It is almost surely continuous,*
(ii) *it is almost surely of bounded variation, and*
(iii) $A(0) = 0$,

then $A(t) \equiv 0$.

PROOF: Without loss of generality we can assume that the total variation of $A(t)$ is uniformly bounded by a constant C. Otherwise, if $A^*(t)$ is the variation of $A(\cdot)$ on $[0, t]$, $A^*(t)$ is continuous and we can define the stopping time

$$\tau = \inf\{t : A^*(t) \geq C\}$$

and by Corollary 1.4, $A(\tau \wedge t)$ would be a continuous martingale with total variation bounded uniformly by C. This would make $A(\tau \wedge t) \equiv 0$ and we can let $C \to \infty$ to establish that $A(t) \equiv 0$.

If the total variation of $A(\cdot)$ is bounded by C on $[0, T]$, we divide $[0, T]$ into n equal parts $0 = t_0 < t_1 < \cdots t_n = T$, with $t_j = \frac{j}{n}$.

$$\sum_{j=1}^{n}[A(t_j) - A(t_{j-1}]^2 \leq \Big[\sup_{1\leq j\leq n} |A(t_j) - A(t_{j-1})|\Big]\Bigg[\sum_{j=1}^{n} |A(t_j) - A(t_{j-1})|\Bigg]$$
$$\leq C\Big[\sup_{1\leq j\leq n} |A(t_j) - A(t_{j-1})|\Big]$$

We note that $A(\cdot)$ is bounded by C and continuous on $[0, T]$. Taking limits as $n \to \infty$ we conclude

$$\lim_{n\to\infty} E\Bigg[\sum_{j=1}^{n}[A(t_j) - A(t_{j-1})]^2\Bigg] = 0$$

On the other, because $A(\cdot)$ is a martingale, for any n,

$$\begin{aligned} E[A^2(T)] &= E\Bigg[\sum_{j=1}^{n}[A^2(t_j) - A^2(t_{j-1})]\Bigg] \\ &= E\Bigg[\sum_{j=1}^{n}[A(t_j) + A(t_{j-1})][A(t_j) - A(t_{j-1})]\Bigg] \\ &= E\Bigg[2\sum_{j=1}^{n} -2A(t_{j-1})[A(t_j) - A(t_{j-1})]\Bigg] \\ &\quad + E\Bigg[\sum_{j=1}^{n}[A(t_j) + A(t_{j-1})][A(t_j) - A(t_{j-1})]\Bigg] \\ &= E\Bigg[\sum_{j=1}^{n}[A(t_j) - A(t_{j-1})]^2\Bigg] \end{aligned}$$

If we let $n \to \infty$ we conclude that $E[A^2(T)] = 0$. □

REMARK 1.3. There are martingales $M(t)$ that are almost surely of bounded variation. They need not be constant if they are not almost surely continuous. The natural decomposition for them of course is to take the compensator as 0. We could take $M(t)$ itself as compensator and we need to rule it out. Continuity will rule it out. While compensators need not be continuous, since they are of bounded variation they can have simple jumps. The time of first jump of a certain size both for a compensator and a martingale are stopping times. But they are qualitatively different. The jump times for martingales are like accidents without warning. You only notice them after they have occurred. They are never foreseen. On the other hand jump times of compensators while still random and unpredictable are nevertheless accessible in the sense that one can see them about to happen like an inevitable accident on the street.

While we will not appeal to the existence of the decomposition, we will see many instances of it. The semimartingales may have jumps, but the compensators will be always almost surely continuous functions of bounded variation. The following lemma is often useful.

LEMMA 1.7 *Let $M(t)$ be a martingale with respect to $(\Omega, \mathcal{F}_t, P)$. Let $A(t)$ be an adapted process which is almost surely a continuous function of bounded variation of t in any interval $[0, T]$ with a uniform bound $C(T)$ on the total variation. Then*

$$X(t) = M(t)A(t) - M(0)A(0) - \int_0^t M(s)dA(s)$$

is a martingale with respect to $(\Omega, \mathcal{F}_t, P)$

PROOF: Since $A(\cdot)$ has no jumps we can approximate

$$\begin{aligned}\int_s^t M(\sigma)dA(\sigma) &= \lim_{n\to\infty}\sum_{j=1}^n M(t_j)[A(t_j) - A(t_{j-1})] \\ &= M(t)A(t) - M(s)A(s) \\ &\quad - \lim_{n\to\infty}\sum_{j=1}^n A(t_{j-1})[M(t_j) - M(t_{j-1})]\end{aligned}$$

where $s = t_0 < t_1 < \cdots < t_n = t$ is a partition of $[s, t]$ into n equal parts. The rest is an elementary calculation if we use

$$E\big[A(t_{j-1})[M(t_j) - M(t_{j-1})] \mid \mathcal{F}_{t_{j-1}}\big] = 0$$

□

The simple intuition behind the lemma is that

$$dX(t) = d[M(t)A(t)] = M(t)dA(t) + A(t)dM(t)$$

and $\int_0^t A(s)dM(s)$ is a martingale.

1.4. Martingales and Stochastic Integrals

If $\{X_j : j \geq 1\}$ is a sequence of independent random variables on $(\Omega, \mathcal{F}, P)$ with mean 0 and variance $\{\sigma_j^2\}$, in order to sum the series

$$S = \sum_j a_j X_j$$

in $L_2(P)$, where $\{a_j\}$ is a sequence of constants, we only need the summability of the series $\sum_j |a_j|^2\sigma_j^2$. Then S will be well defined as a square integrable random variable with mean 0 and variance $\sigma^2 = \sum_j |a_j|^2 \sigma_j^2$. The infinite sum S will be the limit of the partial sums S_n, in mean square and by a theorem of Kolmogorov almost surely as well. In fact, $\{a_j\}$ do not have to be constants. They can be random variables. While a_0 has to be a constant, for $j \geq 1$, a_j can be a function of $X_1, X_2, \ldots, X_j$. The partial sums $S_0 = 0$ and for $n \geq 1$,

$$S_n = \sum_{j=0}^{n-1} a_j(X_1, X_2, \ldots, X_j) X_{j+1}$$

define a martingale with respect to the filtration $\mathcal{F}_n = \sigma(X_1, \ldots, X_n)$. A simple calculation yields

$$E[S_n^2] = E\Big[\sum_{j=0}^{n-1} \sigma_{j+1}^2 |a_j(\omega)|^2\Big]$$

The convergence of the series

$$\sum_{j=1}^{\infty} \sigma_{j+1}^2 E\big[|a_j(\omega)|^2\big]$$

is enough to guarantee both mean square and almost sure convergence of the series

$$S = \sum_{j=0}^{\infty} a_j(X_1, X_2, \ldots, X_j) X_{j+1}$$

In fact one does not need the independence of $\{X_j\}$. It is enough if

$$M_n = \sum_{j=1}^{n} X_j$$

is a martingale. We state all of this as a theorem. The proof is elementary and will be omitted.

THEOREM 1.8 *Let $\{\mathcal{F}_n : n \geq 0\}$ be a filtration on some probability space $(\Omega, \mathcal{F}, P)$ and M_n a square integrable martingale with respect to this filtration, with $M_0 = 0$. Let $\sigma_n^2(\omega) = E[(M_{n+1} - M_n)^2 \mid \mathcal{F}_n]$. Let $\{a_n(\omega)\}$ be a sequence with a_n being measurable with respect to $\mathcal{F}_n$ and satisfying $E[|a_n(\omega)|^2 \sigma_n^2(\omega)] < \infty$. Then $\{S_n : n \geq 0\}$, defined by $S_0 = 0$, and*

$$S_n = \sum_{j=0}^{n-1} a_j(\omega)[M_{j+1} - M_j]$$

for $n \geq 1$, is again a square integrable martingale with

$$E[S_n^2] = E\Big[\sum_{j=0}^{n-1} |a_j(\omega)|^2 \sigma_j^2(\omega)\Big]$$

There is of course an analogue in the continuous case. Let $M(t)$ be a square integrable martingale. By Jensen's inequality $M^2(t)$ is a submartingale. We assume that there is a continuous increasing function $A(t, \omega)$ such that

$$M^2(t) - A(t)$$

is a martingale with respect to $(\Omega, \mathcal{F}_t, P)$. We assume that $E[M(t)^2] = E[A(t)] < \infty$ for every t. Our aim is to define the stochastic integral

$$X(t) = \int_0^t c(s, \omega) dM(s)$$

as another square integrable martingale in a manner quite parallel to Theorem 1.8. Let us assume that $c(s, \omega)$ is piece-wise constant i.e. $c(s, \omega) = c_i(\omega)$ for $s \in [t_i, t_{i+1})$ with $c_i(\omega)$ being uniformly bounded and $\mathcal{F}_{t_i}$ measurable. Then it is not hard to see that the natural definition of the stochastic integral

$$X(t) = \int_0^t c(s, \omega) dM(s)$$

is

$$X(t) = \sum_{i=0}^{k-1} c(t_i, \omega)[M(t_{i+1}) - M(t_i)] + c(t_k, \omega)[M(t) - M(t_k)] \tag{1.13}$$

for $t_k \le t < t_{k+1}$ and with this definition $X(t)$ is again a martingale. Moreover

$$X^2(t) - \int_0^t c^2(s, \omega) dA(s)$$

is a martingale as well. We can complete the space of integrands $\{c(s, \omega)\}$ with respect to the norms

$$E\left[\int_0^t |c(s, \omega)|^2 \, dA(s)\right]$$

and the square integrable martingales

$$X(t) = \int_0^t c(s, \omega) dM(s)$$

can be defined for $c(\cdot, \cdot)$ in the completion.

$$X^2(t) - \int |c(s, \omega)|^2 \, dA(s)$$

will again be a martingale. If the original martingale $M(t)$ is almost surely continuous in t so will $X(t)$ be. This is obvious for c that are piecewise constant. For general c it follows from Doob's inequality

$$E\Big[\sup_{0 \le t \le T} |X(t)|^2\Big] \le 4E\big[|X(T)|^2\big]$$

which ensures that the convergence of the martingales is almost surely uniform in t. The question of describing the class of functions $c(s, \omega)$ that can be approximated in this manner remains to be answered. We will revisit this issue when we discuss stochastic integrals with respect to Brownian motion later in Chapter 5.