1 Introduction to Stochastic Processes

The main result that guarantees the existence of a wide class of stochastic processes is the Kolmogorov consistency theorem. Though the Kolmogorov construction of stochastic processes is set on a very large space equipped with a small σ-field, it is canonical and its applicability is wide. First, we give a proof of this result, followed by important examples that illustrate its usefulness. Next, we introduce basic terminology and notation that are useful in stochastic analysis. The chapter ends with a brief overview of stopping times, associated σ-fields, and progressive measurability.

1.1 The Kolmogorov Consistency Theorem

Throughout the book, Ω will denote an abstract space which, in probability theory, is called the sample space or the space of all outcomes. Let $\mathcal{F}$ denote a σ-field of subsets of Ω, known as the class of events or measurable sets in Ω. A measure P on $(\Omega, \mathcal{F})$ is said to be a probability measure if it is a nonnegative, countably additive set function with $P(\Omega) = 1$. The triplet $(\Omega, \mathcal{F}, P)$ is called a probability space.

In several applications, it is more natural to encounter a finitely additive probability measure, P_0, on a field $\mathcal{G}$ of sets rather than a measure on a σ-field $\mathcal{F}$. The first question that arises is to find conditions under which P_0 can be extended to a probability measure on the σ-field generated by $\mathcal{G}$. The answer is provided by a well-known theorem on extension of measures.

The following proposition is quite useful and easy to prove.

Proposition 1.1.1 *Suppose P_0 is a finitely additive probability measure defined on a field $\mathcal{G}$ of subsets of a space Ω. Let P_0 be continuous at $\emptyset$, that is, if $E_n \in \mathcal{G}$ for all n and $E_n \downarrow \emptyset$, then $P_0(E_n) \downarrow 0$. Then P_0 is a probability measure on $\mathcal{G}$.*

Proof Take any sequence $\{E_n\}$ of pairwise disjoint sets from $\mathcal{G}$ such that $\cup_{j=1}^{\infty} E_j \in \mathcal{G}$. Then for any finite n, $\cup_{j=n+1}^{\infty} E_j \in \mathcal{G}$. Also, one can easily check that $\cup_{j=n+1}^{\infty} E_j \downarrow \emptyset$ as

$n \to \infty$. Therefore $P_0\left(\cup_{j=n+1}^{\infty} E_j\right) \downarrow 0$ as $n \to \infty$. By finite additivity,

$$P_0\left(\cup_{j=1}^{\infty} E_j\right) = P_0\left(\cup_{j=1}^{n} E_j\right) + P_0\left(\cup_{j=n+1}^{\infty} E_j\right)$$

$$= \sum_{j=1}^{n} P_0\left(E_j\right) + P_0\left(\cup_{j=n+1}^{\infty} E_j\right).$$

The partial sums $\sum_{j=1}^{n} P_0\left(E_j\right)$ form a bounded increasing sequence indexed by n, and hence as $n \to \infty$, we have

$$P_0\left(\cup_{j=1}^{\infty} E_j\right) = \sum_{j=1}^{\infty} P_0\left(E_j\right).$$ ■

The theorem stated below is a standard result in measure theory. A proof of it can be found in many texts (for e.g., [56], pp. 19-23).

Theorem 1.1.2 (The Kolmogorov extension theorem) *If P_0 is a probability measure on a field $\mathcal{G}$ of subsets of Ω, and if $\mathcal{F}$ denotes the σ-field generated by $\mathcal{G}$, then P_0 can be uniquely extended to a probability measure on $\mathcal{F}$.*

A function $X : \Omega \to \mathbb{R}^1$ is called a random variable provided X is an $\mathcal{F}$-measurable function. In probability theory, what we are about to observe is a random variable. The probabilistic information on X is fully contained in the distribution of X. Recall that given a random variable X on $(\Omega, \mathcal{F}, P)$, the distribution of X is the measure μ defined on $(\mathbb{R}^1, \mathcal{B})$ by $\mu(B) = P\{X \in B\}$ for all Borel sets $B \in \mathcal{B}$.

A natural question that arises is the following converse: Given a probability measure μ on $(\mathbb{R}^1, \mathcal{B})$, can one construct a random variable X on some probability space $(\Omega, \mathcal{F}, P)$ such that the distribution of X coincides with the given measure μ? The answer to this question is quite simple. One can take $\Omega = \mathbb{R}^1$, $\mathcal{F} = \mathcal{B}$, and $P = \mu$. Define $X(\omega) = \omega$ for all $\omega \in \Omega$. Thus one can go back and forth between a random variable and its distribution.

If the range of X is $\mathbb{R}^n$, then X is said to be an n-dimensional random vector. The discussion in the above paragraph carries over for probability measures on $\mathbb{R}^n$.

In applications, one assumes the existence of a family of random variables defined on a probability space. When a finite number, n, of random variables out of the given family is considered, it is a random vector and hence gives rise to a finite-dimensional distribution which is a probability measure on $\mathbb{R}^n$. As we vary the selections of the random variables from the given family, we obtain a family of finite-dimensional measures. The assumption mentioned above can be removed, had one started with a family of finite-dimensional distributions. In this situation, it is unclear if there exists a common probability space on which a corresponding family of random variables can be defined. To get an affirmative answer, one needs a certain consistency property in the above family of finite-dimensional measures. This is the content of a famous result known as the

Kolmogorov consistency theorem. We start with a lemma on regularity of probability measures that holds in general for measures defined on the Borel σ-field of complete, separable metric spaces (see [33]). However, we prove the lemma for measures on $(\mathbb{R}^n, \mathcal{B}^n)$.

Lemma 1.1.3 *Let μ be any given probability measure on $(\mathbb{R}^n, \mathcal{B}^n)$. Given any $B \in \mathcal{B}^n$, and $\epsilon > 0$, there exist a compact set K and an open set G such that $K \subseteq B \subseteq G$ with $\mu(G \setminus K) < \epsilon$.*

Proof Let $\mathcal{A}$ denote the class of all Borel sets B that satisfy the stated property. It suffices to show that $\mathcal{A}$ contains all closed sets and is a σ-field.

Step 1 Clearly, the empty set $\emptyset \in \mathcal{A}$. Let $B_N = [-N, N]^{\times n}$, the n-dimensional closed box with origin as the center and side-length $2N$. Since $\mu(B_N) \uparrow 1$ as N increases to ∞, one can choose N large enough so that $\mu(\mathbb{R}^n \setminus B_N) < \epsilon$. Thus, the full space $\mathbb{R}^n \in \mathcal{A}$.

Step 2 We will show that $\mathcal{A}$ is closed under complements and countable unions. If $A \in \mathcal{A}$, then by the definition of $\mathcal{A}$, it is possible to choose $K \subseteq A \subseteq G$ such that $\mu(G \setminus K) < \epsilon/2$. As in Step 1, one can choose N large so that $\mu(B_N^c) < \epsilon/2$.

Clearly, $G^c \subseteq A^c \subseteq K^c$.

$$\begin{aligned}\mu\left(K^c \setminus (G^c \cap B_N)\right) &= \mu\left(K^c \cap (G \cup B_N^c)\right)\\ &\leq \mu(G \setminus K) + \mu(B_N^c)\\ &< \epsilon.\end{aligned}$$

Thus, $A^c \in \mathcal{A}$.

If $\{A_j\}$ is a sequence of sets in $\mathcal{A}$, one can choose $K_j \subseteq A_j \subseteq G_j$ such that $\mu(G_j \setminus K_j) < \epsilon/2^{j+1}$. Therefore,

$$\mu\left((\cup G_j) \setminus (\cup K_j)\right) \leq \sum_j \mu\left(G_j \setminus K_j\right) < \epsilon/2.$$

Let G denote $\cup_j G_j$. Then, $\lim_{N\to\infty} \mu(G \setminus (\cup_{j=1}^N K_j)) < \epsilon/2$. Therefore, for large enough N, one gets $\mu(G \setminus (\cup_{j=1}^N K_j)) < \epsilon$. Thus $\cup_{j=1}^{\infty} A_j \in \mathcal{A}$. We have thus shown that $\mathcal{A}$ is a σ-field.

Step 3 The class $\mathcal{A}$ contains all closed sets. For, if F is a closed set, then let F_δ denote the delta neighborhood of F. That is,

$$F_\delta = \left\{x : |x - a| < \delta \text{ for some } a \in F\right\}.$$

The set F_δ is an open set and decreases to F as $\delta \downarrow 0$. Therefore, one can choose δ small enough so that $\mu(F_\delta \setminus F) < \epsilon/2$.

As in Step 1, choose N large so that $\mu(B_N^c) < \epsilon/2$. Then $\mu(F_\delta \setminus (F \cap B_N)) < \epsilon$ which completes the proof. ∎

Let T denote an index set such as $[0, \infty)$.

Definition 1.1.1 *For any $n \in \mathbb{N}$, let $t_1 < \cdots < t_n$ be any selection (that is, finite sequence) of distinct elements in T. Let $\mu_{t_1,\ldots,t_n}$ be a probability measure on the Borel σ-field of $\mathbb{R}^n$ that corresponds to the selection. The family of probability measures $\{\mu_{t_1,\ldots,t_n} : t_1 < \cdots < t_n, t_i \in T \ \forall\ i, n \in \mathbb{N}\}$ is said to be a consistent family of probability measures if the following condition holds:*

Let $n \in \mathbb{N}$, and $t_1 < \cdots < t_{n+1}$ be any choice of $n+1$ distinct elements from T. Let $B_1, \ldots, B_n$ be one-dimensional Borel sets, arbitrarily chosen. Then, for any k, we have

$$\mu_{t_1,\ldots,t_{n+1}}(B_1 \times \cdots \times B_k \times \mathbb{R} \times B_{k+1} \times \cdots \times B_n) = \mu_{t_1,\ldots,t_k,t_{k+2},\ldots,t_{n+1}}(B_1 \times \cdots \times B_n).$$

Definition 1.1.2 *Let $\mathbb{R}^T$ denote the space of all real-valued functions defined on T. A subset of $\mathbb{R}^T$ is said to be a finite-dimensional* cylinder set *if it is of the form $\{x \in \mathbb{R}^T : (x_{t_1}, \ldots, x_{t_n}) \in B\}$ for any $n \in \mathbb{N}$ and distinct $t_1 < \cdots < t_n \in T$ and B, Borel in $\mathbb{R}^n$.*

The class of all cylinder sets in $\mathbb{R}^T$ will be denoted $\mathcal{C}$. The σ-field generated by $\mathcal{C}$ is denoted by $\mathcal{F}$. For $x \in \mathbb{R}^T$, the projection $x \to (x_{t_1}, \ldots, x_{t_n})$ is denoted by $\pi_{t_1,\ldots,t_n}$.

Theorem 1.1.4 (Kolmogorov consistency theorem) *Suppose the family of probability measures $\{\mu_{t_1,\ldots,t_n} : t_1 < \cdots < t_n, n \in \mathbb{N}\}$ is consistent.*

(i) Then there exists a probability measure μ on $(\mathbb{R}^T, \mathcal{F})$ such that $\mu\pi^{-1}_{t_1,\ldots,t_n} = \mu_{t_1,\ldots,t_n}$ for any $n \in \mathbb{N}$ and distinct $t_1 < \cdots < t_n$.

(ii) If $\Omega := \mathbb{R}^T$ and $P := \mu$, then on $(\Omega, \mathcal{F}, P)$, there exist random variables $\{X_t : t \in T\}$ such that for any distinct $t_1 < \cdots < t_n$, the joint distribution of $X_{t_1}, \ldots, X_{t_n}$ coincides with $\mu_{t_1,\ldots,t_n}$.

Proof

Step 1 The class $\mathcal{C}$ is a field of subsets of $\mathbb{R}^T$. For each $C \in \mathcal{C}$, define a set function

$$P_0(C) = \mu_{t_1,\cdots,t_n}(B)$$

if $C = \{x \in \mathbb{R}^T : (x_{t_1}, \ldots, x_{t_n}) \in B\}$. By using the consistency condition, it follows that P_0 is well defined. It is easy to verify that P_0 is a finitely additive probability measure on $\mathcal{C}$. If P_0 were countably additive on $\mathcal{C}$, we can use the extension theorem to obtain a unique extension of P_0 to a measure on $\mathcal{F}$.

Step 2 By the proposition, it suffices to show that P_0 is continuous at $\emptyset$. We will use the method of contradiction. Let $A_n \in \mathcal{C} \ \ \forall \ \ n$, and $A_n \downarrow \emptyset$. Suppose there exists $\delta > 0$ such that $\lim_{n\to\infty} P_0(A_n) \geq \delta$.

Let $(\{t_1, \ldots, t_{k_n}\}, B_n)$ be a representation in finite dimensions for A_n for each n. Then,

$$P_0(A_n) = \mu_{t_1,\ldots,t_{k_n}}(B_n).$$

By inserting as many sets as needed between each A_j and A_{j+1}, we can assume without loss of generality that $k_n = n$ for all n.

Given any $0 < \epsilon < \delta$, by Lemma 1.1.3, there exists a compact set F_n such that $F_n \subseteq B_n$ and

$$\mu_{t_1,\ldots,t_n}\left(B_n \setminus F_n\right) < \epsilon/2^n.$$

Let us denote by E_n the cylinder set represented by $(\{t_1, \ldots, t_n\}, F_n)$.

Then, $P_0\left(A_n \setminus E_n\right) < \epsilon/2^n$. Let Λ_n denote $\cap_{j=1}^n E_j$. Then,

$$P_0\left(A_n - \Lambda_n\right) = P_0\left(\cup_{j=1}^n \left(A_n - E_j\right)\right) \leq \sum_{j=1}^n P_0\left(A_j \setminus E_j\right) < \epsilon$$

since $A_n \subseteq A_j$ for $j \leq n$. Using $\Lambda_n \subseteq E_n \subseteq A_n$, we get

$$P_0\left(\Lambda_n\right) \geq P_0\left(A_n\right) - \epsilon \geq \delta - \epsilon > 0.$$

Each Λ_n is nonempty, since $P_0(\Lambda_n) > 0$. Therefore, for each n, there exists a $\mathbf{x}^{(n)} \in \Lambda_n$. By using the definition of Λ_n, it follows that for each k, $(x_{t_1}^{(n)}, \ldots, x_{t_k}^{(n)}) \in F_k$ for all $n \geq k$.

Therefore, as a sequence in n, $\{x_{t_1}^{(n)}\}$ is in a closed bounded set F_1 of the real line so that there exists a convergent subsequence $\{x_{t_1}^{(1,n)}\}$ with limit denoted by $\alpha_1 \in F_1$.

The sequence $\{x_{t_1}^{(1,n)}, x_{t_2}^{(1,n)}\}$ is a sequence in F_2 and has a convergent subsequence whose limit is denoted by $(\alpha_1, \alpha_2) \in F_2$. This procedure can be continued to obtain an element $(\alpha_1, \ldots, \alpha_k) \in F_k$ for any finite k.

It is clear that there exist elements y of $\mathbb{R}^T$ such that $y_{t_k} = \alpha_k$ for all k. Any such element y has the property that $(y_{t_1}, \ldots, y_{t_k}) \in F_k$ for any finite k. Therefore, $y \in E_k \subseteq A_k$ for all k, which implies that $\cap_{n=1}^\infty A_n$ is non-empty. This contradicts the assumption that $\cap_{n=1}^\infty A_n$ is empty. We have thus proved that P_0 on $\mathcal{C}$ is countably additive. Invoking the Kolmogorov extension theorem, there exists a unique probability measure μ on $(\mathbb{R}^T, \mathcal{F})$ that extends P_0. The proof of part (i) is over.

Step 3 Set $\Omega = \mathbb{R}^T$ and $P = \mu$ on $(\Omega, \mathcal{F})$. Let ω denote a generic element of Ω. Define the functions $X_t(\omega) = \omega_t$ for all $t \in T$ and $\omega \in \Omega$.

Clearly, $\{X_t \leq x\} = \{\omega : \omega_t \leq x\} \in \mathcal{C} \subseteq \mathcal{F}$, so that X_t is a random variable. Further,

$$P\left\{\cap_{j=1}^n \left(X_{t_j} \leq x_j\right)\right\} = \mu\left\{\omega : \cap_{j=1}^n \left(\omega_{t_j} \leq x_j\right)\right\} = \mu_{t_1,\ldots,t_n}\left\{\times_{j=1}^n (-\infty, x_j]\right\},$$

so that the finite-dimensional distribution of the vector random variable $X_{t_1}, \ldots, X_{t_n}$ is given by $\mu_{t_1,\ldots,t_n}$. ■

It is possible to reformulate the Kolmogorov consistency theorem in terms of characteristic functions. Toward building such a statement, consider a probability measure P

on $(\mathbb{R}^T, \mathcal{F})$. Let F^T denote the collection of all points λ in $\mathbb{R}^T$ such that all but a finite number of coordinates of λ are zero. Define for any $\lambda \in F^T$,

$$\phi(\lambda) = \int_{\mathbb{R}^T} e^{i(\lambda,x)} P(dx) \tag{1.1.1}$$

where $(\lambda, x) = \sum_j \lambda_{t_j} x_{t_j}$ is a finite sum.

Let $(t_1 < \cdots < t_n)$ be a fixed finite number of distinct elements of T, and let λ be any element of F^T with $\lambda_t = 0$ if $t \neq t_j$ for any $j = 1, \ldots, n$. The function $\phi(\lambda)$ restricted to such λ is known as the *section* of ϕ determined by $(t_1, \ldots, t_n)$.

Theorem 1.1.5 (A reformulation of the Kolmogorov consistency theorem) *Let ϕ be a given complex-valued function on F^T. If any arbitrary section of a function ϕ is a characteristic function, then there exists a probability measure P on $(\mathbb{R}^T, \mathcal{F})$ such that* (1.1.1) *holds for all $\lambda \in F^T$.*

We call the family of random variables $\{X_t : t \in T\}$ obtained in the consistency theorem the canonical process. It should be observed that the space Ω in part (ii) of the above theorem is too large while the σ-field $\mathcal{F}$ is too small. In fact, Ω is the space of all real-valued functions defined on T. The smallness of $\mathcal{F}$ can be inferred from the following result and examples. We will assume that T is nondenumerable. In what follows, the symbols $\omega(t_j)$ and ω_{t_j} are synonymous. Let A be a subset in Ω. If there exist a countable set $\{t_1, t_2, \ldots\}$ in T and a Borel set $B \in \mathcal{B}(\mathbb{R}^T)$ such that $A = \{\omega : (\omega(t_1), \omega(t_2), \ldots) \in B\}$, then A will be called as a **set with a countable base**.

The following theorem is due to Dynkin, who extracted a particularly useful part of an otherwise general, result due to Sierpinski. The theorem is given without a proof. Before stating it, let us recall that a collection $\mathcal{A}$ of subsets of a space E is known as a **Dynkin class** or a λ-system if $E \in \mathcal{A}$ and if $\mathcal{A}$ is closed under proper differences and increasing unions.

Theorem 1.1.6 (The Dynkin class theorem) *Let $\mathcal{S}$ be a π system, that is, a collection of subsets of E that is closed under finite intersections. If $\mathcal{A}$ is a λ-system with $\mathcal{S} \subseteq \mathcal{A}$, then $\sigma(\mathcal{S}) \subseteq \mathcal{A}$.*

Lemma 1.1.7 *Let $\mathcal{D}$ denote the class of all sets with a countable base. Then $\mathcal{F} := \sigma(\mathcal{C}) \subseteq \mathcal{D}$.*

Proof It is clear that $\mathcal{C} \subseteq \mathcal{D}$. The class $\mathcal{C}$, being a field, is a π-system.

By taking any countable set $\{t_1, t_2, \ldots\}$ in T and $B = \mathbb{R}^T$, we get the full space $\Omega \in \mathcal{D}$. Let $A_1, A_2 \in \mathcal{D}$ and $A_1 \subseteq A_2$. If A_1 is represented as $A_1 = \{\omega : (\omega(t_1), \omega(t_2), \ldots) \in B_1\}$, then there exists a Borel set B_2 such that $A_2 = \{\omega : (\omega(t_1), \omega(t_2), \ldots) \in B_2\}$ with $B_1 \subseteq B_2$. Therefore, $A_2 \backslash A_1 = \{\omega : (\omega(t_1), \omega(t_2), \ldots) \in B_2 \backslash B_1\}$, so that $\mathcal{D}$ is closed under proper differences.

If $\{A_j\}$ is an increasing sequence in $\mathcal{D}$ and if $A_1 = \{\omega : (\omega_{t_1}, \omega_{t_2}, \ldots) \in B_1\}$, then each A_j can be represented by using the same set $\{t_1, t_2, \ldots\}$. In fact, $A_j = \{\omega : (\omega_{t_1}, \omega_{t_2}, \ldots) \in B_j\}$ with $B_i \subseteq B_j$ for all $i \leq j$. Thus, $\cup A_j = \{\omega : (\omega_{t_1}, \omega_{t_2}, \ldots) \in \cup B_j\}$, which shows that $\mathcal{D}$ is closed under increasing unions.

The class $\mathcal{D}$ is thus a Dynkin class, and hence, $\mathcal{F} = \sigma(\mathcal{C}) \subseteq \mathcal{D}$. ∎

Example 1.1.8 *Let Ω_c be the set of all continuous functions in Ω. Then Ω_c is not in $\mathcal{F}\,\sigma(\mathcal{C})$.*

Proof If Ω_c were in $\mathcal{F}$, then by the above lemma, Ω_c is a set with a countable base so that

$$\Omega_c = \{\omega : (\omega_{t_1}, \omega_{t_2}, \ldots) \in B\}.$$

However, this is not possible since discontinuous functions are also included in the set on the right side. ■

Next, we show that for any probability measure P on $(\Omega, \mathcal{F})$, the inner measure $\underline{P}$ of Ω_c is zero. Recall that if A is any subset of Ω, then the outer measure of A is defined by

$$\bar{P}(A) = \inf\left\{P(E) : A \subseteq E = \cup_j E_j \quad \text{where} \quad E_j \in \mathcal{C}\right\}.$$

The inner measure $\underline{P}(A)$ is defined by $\underline{P}(A) = 1 - \bar{P}(A^c)$.

Proposition 1.1.9 $\underline{P}(\Omega_c) = 0$.

Proof We will show that $\bar{P}(\Omega_c^c) = 1$. Let $(\Omega_c^c) \subseteq E = \cup_j E_j$ where $E_j \in \mathcal{C}$ for all j. Let $S = \{s_1, s_2, \ldots\} \subseteq T$ be a countable base for E so that there exists a Borel $B \in \mathcal{B}(\mathbb{R}^\infty)$ such that

$$E = \{\omega : (\omega_{s_1}, \omega_{s_2}, \ldots) \in B\}. \tag{1.1.2}$$

If $\omega* \in \Omega$ is any continuous, real-valued function on T, then let ω_0 be a *discontinuous* function on T whose values coincide with the values of $\omega*$ on S. It follows that $\omega* \in E$, since the right side of (1.1.2) depends only on the function values on the countable set S. Thus, $E = \Omega$, and hence $P(E) = 1$, which yields that, $\bar{P}(\Omega_c^c) = 1$. ■

Another example of a set with zero inner measure is given by

$$\Omega_m = \{\omega \in \Omega : \omega(t) \quad \text{is a Lebesgue measurable function of } t\}.$$

The proof is similar to the one given above and consists in showing that $\bar{P}(\Omega_m^c) = 1$.

We will now construct a rather important measure P on $(\Omega, \mathcal{F})$ when the index set $T = [0, \infty)$ by specifying the family of finite-dimensional measures as follows:

$$\mu_{t_1, t_2, \ldots, t_n}(B_1 \times B_2 \cdots \times B_n) = \int_{B_1} \cdots \int_{B_n} \Pi_{j=1}^n p\left(t_j - t_{j-1}; x_j - x_{j-1}\right)\, dx_n \cdots dx_1 \tag{1.1.3}$$

where $0 < t_1 < t_2 \cdots < t_n$, B_j are any Borel sets in $\mathbb{R}$, $t_0 = 0$, $x_0 = 0$, and

$$p(t; x) = \frac{1}{\sqrt{2\pi t}}\, exp\left\{-\frac{x^2}{2t}\right\}$$

for all $t > 0$ and any real number x. When $t = 0$, $\mu_0(B) = \delta_0(B)$ for any Borel set B.

It is easy to show that the above family of probability measures satisfies the Kolmogorov consistency condition so that there exists a family of random variables $\{X_t\}$ on the probability space $(\Omega, \mathcal{F}, P)$ such that their finite-dimensional distributions are given by $\{\mu_{t_1,t_2,\ldots,t_n}\}$. Recall that the process was defined by $X_t(\omega) = \omega(t)$ for all $\omega \in \Omega$. By the specification of measures (1.1.3), the following properties of $\{X_t\}$ follow easily:

(i) The random variable X_0 is identically equal to 0.
(ii) For any $t > 0$, X_t is an $N(0, t)$ random variable.
(iii) For any $0 < s < t$, $X_t - X_s$ is an $N(0, t-s)$ random variable.
(iv) If $0 < s_1 < t_1 \leq s_2 < t_2 \cdots \leq s_n < t_n$, the random variables $\{X_{t_j} - X_{s_j} : j = 1, \ldots n\}$ are independent.

We will proceed to show that under this measure P, $\bar{P}(\Omega_c) = 1$.

In terms of the canonical process $\{X_t\}$, this assertion would translate to $P(\omega : \{X_t(\omega)\} \in E) = 1$ for any set $E \in \mathcal{F}$ that contains Ω_c. Such sets E have a countable base and yet must successfully diagnose the property of continuity in t of ω. In other words, E must detect the uniform continuity of ω in the interval $[0, T]$ for all $T > 0$. Toward this goal, we prove below the following propositions.

Proposition 1.1.10 *Let* $T := \{t_j : j = 0, 1, \ldots, n\}$ *arranged in increasing order with* $t_0 = 0$ *and* $t_n = 1$. *Then, for any* x,

(i) $P\{\max_{t\in T} X_t > x\} \leq 2P\{X_1 > x\}$, *and*
(ii) $P\{\max_{t\in T} |X_t| > x\} \leq 2P\{|X_1| > x\}$.

Proof Define $Y_j = X_{t_j} - X_{t_{j-1}}$, and $\tau = \inf\{j : X_{t_j} > x\}$. Then $P\{\max_{tinT} X_t > x\} = \cup_{j=1}^{n} p\{\tau = j\}$.

By the symmetry of the distribution of $X_1 - X_{t_j}$, we can write

$$\frac{1}{2} P\left\{\max_{t\in T} X_t > x\right\} = \sum_{j=1}^{n} P\{\tau = j\} P\left\{X_1 - X_{t_j} \geq 0\right\}.$$

The event $\{\tau = j\}$ belongs to the σ-field generated by $\{X_{t_1}, \ldots, X_{t_j}\}$, namely $\sigma(X_{t_1}, \ldots, X_{t_j})$ which is the same as

$$\sigma\left(X_{t_1}, X_{t_2} - X_{t_1}, \ldots, X_{t_j} - X_{t_{j-1}}\right).$$

Therefore, the events $\{\tau = j\}$ and $\{X_1 - X_{t_j} > 0\}$ are independent so that

$$\begin{aligned}\frac{1}{2} P\left\{\max_{t\in T} X_t > x\right\} &= \sum_{j=1}^{n} P\left\{\tau = j, X_1 - X_{t_j} \geq 0\right\} \\ &\leq \sum_{j=1}^{n} P\{\tau = j, X_1 > x\} \\ &\leq P\{X_1 > x\}.\end{aligned}$$

The first inequality in the proposition is thus proved. To prove the second inequality, take any $x > 0$ and note that

$$\begin{aligned}\left\{\max_{t\in T_n} |X_t| > x\right\} &= \left\{\max_{t\in T_n} X_t > x\right\} \cup \left\{\min_{t\in T_n} X_t < -x\right\} \\ &= \left\{\max_{t\in T_n} X_t > x\right\} \cup \left\{\max_{t\in T_n} (-X_t) > x\right\}.\end{aligned}$$

The finite-dimensional distribution of $\{-X_t\}$ is the same as that of $\{X_t\}$. Using the inequality (i),

$$\begin{aligned}P\left\{\max_{t\in T_n} |X_t| > x\right\} &= P\left\{\max_{t\in T_n} X_t > x\right\} + P\left\{\max_{t\in T_n} (-X_t) > x\right\} \\ &\leq 2\left(P\{X_1 > x\} + P\{-X_1 > x\}\right) \\ &= 2P\left\{|X_1| > x\right\}.\end{aligned}$$

■

Remark

1. The inequalities in the above proposition are valid even if T is a countable set in $[0, 1]$. Indeed, write T as $\cup T_n$ where T_n is an increasing sequence of finite ordered subsets of T such that $0, 1 \in T_n$ for each n. For each T_n, the inequalities hold. Let $n \to \infty$ to complete the proof.
2. For simplicity, we have taken $t_0 = 0$ and $t_n = 1$. For general $t_0 = a$ and $t_n = b$ the inequalities would appear as follows:

$$P\left\{\max_{t\in T} X_t - X_a > x\right\} \leq 2P\{X_b - X_a > x\}$$

$$P\left\{\max_{t\in T} |X_t - X_a| > x\right\} \leq 2P\left\{|X_b - X_a| > x\right\}.$$

Proposition 1.1.11 *On any countable set D in $[0, \infty)$ and any for any finite $T > 0$, the process $\{X_t\}$ restricted to $D \cap [0, T]$ is uniformly continuous with probability one.*

Proof Without loss of generality, let $T = 1$. Let us enlarge D by including the set of points $\{k/2^n : k = 1, \ldots, 2^n\}$ for all $n \geq 1$, and still denote it as D. Define

$$M_n = \sup_{\{t_i, t_j \in D : |t_j - t_i| \leq 1/2^n\}} |X_{t_j} - X_{t_i}|.$$

We need to show that $\lim_{n\to\infty} M_n = 0$ almost surely (a.s.). Since M_n is a decreasing sequence, it suffices to show that for any $\epsilon > 0$,

$$\lim_{n\to\infty} P\{M_n > \epsilon\} = 0.$$

Let $I_k = [k/2^n, (k+1)/2^n]$. Define

$$Z_k = \sup_{t\in I_k \cap D} |X_t - X_{k/2^n}| \quad \forall \quad k = 0, 1, \ldots, 2^n - 1.$$

By the triangle inequality, $M_n \leq 3 \max_k Z_k$. It is enough to show that

$$P\left\{\max_k Z_k > \epsilon\right\} \to 0$$

as $n \to \infty$. Clearly,

$$P\left\{\max_k Z_k > \epsilon\right\} = P\left(\cup\{Z_k > \epsilon\}\right) \leq \sum_{k=0}^{2^n-1} P\{Z_k > \epsilon\}. \tag{1.1.4}$$

By the remark following the above proposition,

$$\begin{aligned} P\left\{\max_{t\in I_k\cap D} |X_t - X_{k/2^n}| > \epsilon\right\} &\leq 2P\left\{|X_{(k+1)/2^n} - X_{k/2^n}| > \epsilon\right\} \\ &= 2P\left\{|X_{1/2^n}| > \epsilon\right\}. \end{aligned}$$

By using this estimate in (1.1.4),

$$P\left\{\max_k Z_k > \epsilon\right\} \leq 2^{n+1} P\left\{|X_{1/2^n}| > \epsilon\right\}$$

which tends to zero as $n \to \infty$. In fact, as n tends to ∞, it is easily seen that $2^n P\{|X_{1/2^n}| > \epsilon\} \to 0$ by the Markov inequality. The proof is thus completed. ■

Theorem 1.1.12 *The outer measure $\bar{P}$ of the set of continuous functions in Ω is one.*

Proof If E is any set in $\mathcal{F}$, then E is countably based, and we can augment the countable base by all the dyadic rationals. Let us call the enlarged countable base D.

Set $D_n = D \cap [0, n]$ for all integers $n \geq 1$.

Define $U = \cap_n \{\omega : \omega|_{D_n} \text{ is uniformly continuous}\}$. If $\Omega_c \subseteq E$, then $U \subseteq E$. By the above proposition, for each n,

$$P\left\{\omega : \omega|_{D_n} \text{ is uniformly continuous}\right\} = 1.$$

Therefore, $P(E) = 1$. We conclude that $\bar{P}(\Omega_c) = 1$. ■

Remark To sum up, we have proved in Proposition 1.1.9 and in Theorem 1.1.12 that $\bar{P}\{(\Omega_c)^c\} = 1$ and $\bar{P}(\Omega_c) = 1$. This is possible since the outer measure is not an additive set function.

The following proposition allows us to localize P to Ω_c and construct a probability measure on $(\Omega_c, \mathcal{F} \cap \Omega_c)$.

Proposition 1.1.13 *Let $(\Omega, \mathcal{F}, P)$ be any probability space, and let $A \subseteq \Omega$ be a subset having P-outer measure one. Then, there exists a unique probability measure Q on $(A, \mathcal{F} \cap A)$ with the property that $Q(E \cap A) = P(E)$ for all $E \in \mathcal{F}$.*

Proof We have to show that the measure Q is well defined. If E_1 and E_2 are in $\mathcal{F}$, and $E_1 \cap A = E_2 \cap A$, we need to prove that $P(E_1) = P(E_2)$. The symmetric difference $E_1 \Delta E_2$ satisfies

$$\begin{aligned}
I_{E_1 \Delta E_2} &= I_{E_1} + I_{E_2} \quad \pmod 2 \\
&= I_{\{E_1 \backslash A\}} + I_{\{E_2 \backslash A\}} \pmod 2 \\
&= I_{\{E_1 \backslash A\} \Delta \{E_2 \backslash A\}} \\
&= I_{\{(E_1 \Delta E_2) \backslash A\}} \\
&\leq I_{\{\Omega \backslash A\}}
\end{aligned}$$

where the second equality is due to $E_1 \cap A = E_2 \cap A$.

Since A has outer measure one, $P(E_1 \Delta E_2) = 0$. Therefore, $P(E_1) = P(E_2)$.

Thus Q is well defined. The uniqueness of Q is easily seen. ■

The measure P that we constructed on $(\Omega, \mathcal{F})$ can thus be restricted to Ω_c, and named as Q. The canonical process $\{Y_t\}$ on $(\Omega_c, \mathcal{F} \cap \Omega_c, Q)$ has the same finite-dimensional distributions as the canonical process $\{X_t\}$ on $(\Omega, \mathcal{F}, P)$. In addition, the process $\{Y_t\}$ has continuous paths with Q measure one.

1.2 The Language of Stochastic Processes

Let C denote the space of all continuous, real-valued functions defined on $[0, \infty)$. Let C be equipped with the metric

$$d(f, g) = \sum_{n=1}^{\infty} \frac{1}{2^n} \left(1 \wedge \sup_{t \in [0,n]} |f(t) - g(t)|\right)$$

for any $f, g \in C$. Under this topology of uniform convergence on compact subsets of $[0, \infty)$, the space C is a complete separable metric space. Let the Borel σ-field of C be denoted by $\mathcal{B}$. Let $\mathcal{G}$ denote the σ-field in C generated by the class of all finite-dimensional cylinder sets.

Lemma 1.2.1 *The σ-field $\mathcal{G}$ is equal to the Borel σ-field $\mathcal{B}$.*

Proof Let G be an open subset of C. Then, G is a countable union of open balls. Consider $B_\epsilon(g)$, the ϵ ball in C centered at g. Clearly,

$$B_\epsilon(g) = \left\{ f \in C : \sum_{n=1}^{\infty} \frac{1}{2^n} 1 \left(\wedge \sup_{t \in [0,n] \cap \mathbb{Q}} |f(t) - g(t)| < \epsilon \right) \right\}.$$

The set on the right side is in $\mathcal{G}$ by Lemma (1.1.7). Hence we can conclude that, $\mathcal{B} \subseteq \mathcal{G}$.

To prove the converse, consider the set $A = \{ f \in C : f(t) \in O \}$ for any fixed $t > 0$ and any open set O in $\mathbb{R}^1$. If h is any element in A, then there exists $\epsilon > 0$ such that the interval $(h(t) - \epsilon, h(t) + \epsilon) \in O$. It follows that $B_{\epsilon/2}(h) \subseteq A$ so that A is an open set in C. Hence, $A \in \mathcal{B}$. The proof is completed upon noting that sets of the form A generate $\mathcal{G}$. ∎

It is clear from the above lemma that the probability measure Q constructed in the previous section can indeed be considered as a measure on $(C, \mathcal{B})$ since $\Omega_c = C$ and $\mathcal{F} \cap \Omega_c = \mathcal{B}$. The measure Q is known as the **Wiener measure**.

Thus far, the family of random variables that we have encountered has been the canonical (or the coordinate) process, which is just one example of what are known as stochastic processes. Before we proceed further, a few basic definitions are needed and given below.

Definition 1.2.1 *Let $(\Omega, \mathcal{F}, P)$ be any complete probability space, and S, a complete separable metric space. Let $\mathcal{B}$ denote the σ-field of Borel sets of S. Then, $X = \{X_t : t \in T\}$ is called a stochastic process if, for each $t \in T$, X_t is an S-valued random variable. The space S is called the state space of X.*

The index set T, from now on, is either a finite interval, $[0, \infty)$, or $(-\infty, \infty)$. It usually represents time. For each fixed $\omega \in \Omega$, the function $t \to X_t(\omega)$ is known as the **sample path** or **trajectory** associated with ω. A process $X = \{X_t\}$ is said to be continuous, right-continuous, or left-continuous if for P-almost all ω, the trajectory of ω has this property.

Definition 1.2.2 *A stochastic process $X = \{X_t\}$, $t \in T$, is a measurable (or jointly measurable) process if the map $(t, \omega) \to X_t(\omega)$ is measurable with respect to $\mathcal{B}(T) \times \mathcal{F}$.*

The basic regularity that one would expect in a stochastic process is its measurability. *We will assume that all processes considered in this book are measurable.*

Definition 1.2.3 *Let $X = \{X_t\}$ and $Y = \{Y_t\}$ be two stochastic processes defined on $(\Omega, \mathcal{F}, P)$. They are said to be* **versions** *or* **modifications of** *each other if*

$$P\{\omega : X_t(\omega) = Y_t(\omega)\} = 1 \text{ for each } t.$$

If X and Y are continuous (or right-continuous, or left-continuous) processes, and if X is a version of Y, then the following stronger statement holds:

$$P\{\omega : X_t(\omega) = Y_t(\omega) \ \ \forall t \in \mathbb{R}^+\}\, 1.$$

The above equality is called **indistinguishability** of the processes X and Y. A weaker notion of modification is the concept of **equivalence** of processes given below. It makes sense even when two processes are defined on different probability spaces.

Definition 1.2.4 *Let $X = \{X_t\}$ and $Y = \{Y_t\}$ be two stochastic processes defined on $(\Omega, \mathcal{F}, P)$ and $(\Omega', \mathcal{F}', P')$ respectively. The processes are said to be* **equivalent** *if for every $\{t_1, \ldots, t_n\} \subseteq T$ and $B_j \in \mathcal{B}$ for $j = 1, \ldots, n$*

$$P\{\omega : X_{t_1}(\omega) \in B_1, \ldots, X_{t_n}(\omega) \in B_n\} = P'\left\{\omega' : X'_{t_1}(\omega') \in B_1, \ldots, X'_{t_n}(\omega') \in B_n\right\}.$$

Using the concept of equivalence of processes, we have:

Definition 1.2.5 *Let $X = \{X_t\}$ be a given stochastic process on $\left(\mathbb{R}^T, \mathcal{F}, P\right)$. Let $A \subseteq \mathbb{R}^T$ and $\mathcal{A}$ be the σ-field generated by the finite-dimensional Borel cylinder subsets of A. Then, X is said to have a* **realization** *in A if there exists a probability measure P' on $(A, \mathcal{A})$ such that the coordinate process $Y_t(\omega) = \omega(t)$ for $\omega \in A$ is equivalent to X.*

For instance, the process $\{X_t\}$ (with finite-dimensional distributions specified by equation (1.1.3)) defined on $\left(\mathbb{R}^T, \mathcal{F}, P\right)$ is equivalent to the canonical process $\{Y_t\}$ on $(C, \mathcal{B}, Q)$ and hence has a realization in the space of continuous functions. It is worthwhile to note that such a property does not hold even for the simplest process $X = \{X_t\}$ of mutually independent and identically distributed (iid) random variables as shown in [39]:

Example 1.2.2 *Let $T = [0, 1]$. Let $\{X_t\}$ be iid random variables with range of X_t containing at least two distinct values. Then, the process X doesn't have a realization in $C[0, 1]$.*

Proof Assume the contrary. Suppose μ is the probability measure on $C = C[0, 1]$ induced by the distribution of $\{X_t\}$. Let a be chosen such that $\mu\{x \in C : x(1) > a\}$ is strictly between 0 and 1. Call it δ. Define

$$F_k = \{x \in C : x(1 - 1/n) > a \text{ for } n = k + 1, \ldots, 2k\}.$$

Since $\mu(F_k) = \delta^k$, we get $\sum_{k=1}^{\infty} \mu(F_k) < \infty$. By the Borel-Cantelli lemma, $\mu\{\limsup_{k\to\infty} F_k\} = 0$. This implies that $\mu\{x \in C : x(1) > a\} = 0$, a contradiction since $\delta > 0$. ■

In fact, there is no measurable process equivalent to X as the following example illustrates. However, such processes arise as models for white noise, and are treated in a different framework.

Example 1.2.3 *Let $X = \{X_t\}$ be iid random variables as in the previous example, where $t \in [0, 1]$. Then, there is no measurable process equivalent to X.*

Proof Assume the contrary. Choose a large enough K, and define the process $Y_t = X_t 1_{(X_t \le K)}$ where Y_t takes at least two distinct values. Let $E(Y_t) = m$ and variance $V(Y_t) = \sigma^2$ so that $Z_t := \dfrac{Y_t - m}{\sigma}$ is a process consisting of iid random variables with zero expectation and unit variance. Our supposition would imply that $\{Z_t\}$ is a measurable process. We will establish a contradiction.

If I is any subinterval of $[0, 1]$, then $E \int_I \int_I |Z_t Z_s| \, dt \, ds < \infty$.

By Fubini's theorem,

$$E\left[\left(\int_I Z_t \, dt\right)^2\right] = E\left(\int_I \int_I Z_t Z_s \, dt \, ds\right) = \int_I \int_I E\left(Z_t Z_s\right) \, dt \, ds = 0 \tag{1.2.1}$$

since $E\left(Z_t Z_s\right) = 0$ if $t \neq s$, and is $= 1$ if $t = s$. Hence, $\int_I Z_t\left(\omega\right) \, dt = 0$ P-a.a. ω. We thus obtain a P-null set N_I such that $\int_I Z_t\left(\omega\right) \, dt = 0$ if ω is not in N_I. Let $\mathcal{I}$ be the class of all intervals in $[0, 1]$ with rational endpoints. Define $N = \cup_{I \in \mathcal{I}} N_I$. Clearly, $P\left(N\right) = 0$, and for $\omega \in N^c$, $\int_a^b Z_t\left(\omega\right) \, dt = 0$ for any $[a, b] \subseteq [0, 1]$.

Thus, for any $\omega \in N^c$, $Z_t\left(\omega\right) = 0$, except possibly for a set of Lebesgue measure zero. Therefore, by Fubini,

$$\int_0^1 \int_\Omega Z_t^2\left(\omega\right) P\left(d\omega\right) dt = 0,$$

a contradiction since $E(Z_t^2) = 1$, and hence the left side in the above equation is equal to 1. ■

1.3 Sigma Fields, Measurability, and Stopping Times

Let $(\Omega, \mathcal{F}, P)$ be a given probability space. An increasing family of sub σ-fields $(\mathcal{F}_t : t \in \mathbb{R}^+)$ of $\mathcal{F}$, that is, $\mathcal{F}_s \subseteq \mathcal{F}_t \subseteq \mathcal{F}$ for any $s \le t$, is known as a **filtration**. A stochastic process $X = \{X_t\}$ is said to be $\mathcal{F}_t$-**adapted** if each X_t is measurable with respect to $\mathcal{F}_t$.

The sub σ-field $\mathcal{F}_t$ provides us information about the way in which the full space Ω is partitioned at time t. As time increases, the partitions become finer. One refers to $\mathcal{F}_t$ as the information available up to time t.

If $X = \{X_t\}$ is a given process, let $\sigma\left(X_s : 0 \le s \le t\right)$ be the smallest σ-field with respect to which the random variables X_s are measurable for all $s \le t$. Let us denote this σ-field by $\mathcal{F}_t^{X,0}$. The family $\left\{\mathcal{F}_t^{X,0} : t \ge 0\right\}$ is known as the **natural filtration** for the process X. Let us also define

$$\mathcal{F}_t^X = \mathcal{F}_t^{X,0} \vee \{P - \text{null sets of } \mathcal{F}\}$$

that is, the augmentation of $\mathcal{F}_t^{X,0}$ by the P-null sets of $\mathcal{F}$.

We will *assume* in this book that

(i) *The space* $(\Omega, \mathcal{F}, P)$ *is a complete probability space,* and
(ii) *The* σ *field* $\mathcal{F}_0$ *contains all P-null sets in* $\mathcal{F}$.

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$ be the given probability space with a **filtration**.

Definition 1.3.1 *A random variable* $\tau : \Omega \to [0, \infty]$ *is called a* **stopping time** *if the event* $\{\tau \leq t\} \in \mathcal{F}_t$ *for any* $t \in \mathbb{R}^+$.

Note that a stopping time is allowed to take the value ∞.

Example 1.3.1 *Let X be an adapted continuous process. Let F be a closed set. Then,* $\tau = \{t > 0 : X_t(\omega) \in F\}$ *is a stopping time.*

Proof Define $G_n = \{x : d(x, F) < 1/n\}$, where $d(x, F)$ denotes the distance from the point x to the set F. Then, G_n is an open set. By continuity of the process,

$$\{\tau \leq t\} = \{X_t \in F\} \cup \left\{\cap_n \cup_{s \in \mathbb{Q} \cap [0,t)} \{X_s \in G_n\}\right\}.$$

The expression on the right side belongs to $\mathcal{F}_t$. ■

The stopping time τ given in the above example is known as the **hitting time** of F for X.

If τ_1 and τ_2 are stopping times, then $\tau_1 \wedge \tau_2 := \min\{\tau_1, \tau_2\}$, $\tau_1 \vee \tau_2 := \max\{\tau_1, \tau_2\}$, and $\tau_1 + \tau_2$ are all stopping times. If $\{\tau_n\}$ is a sequence of stopping times, then $\sup \tau_n$ is a stopping time. These statements are easy to prove, and hence left as exercises.

Definition 1.3.2 *Let* τ *be a stopping time. Then, the* σ*-field* $\mathcal{F}_\tau$ *is defined to be*

$$\{A \in \mathcal{F} : A \cap \{\tau \leq t\} \in \mathcal{F}_t \textit{ for all } t \geq 0\}.$$

It is easy to verify that $\mathcal{F}_\tau$ defined above is indeed a σ-field and that τ is $\mathcal{F}_\tau$ measurable. If τ_1 and τ_2 are two stopping times with $\tau_1 \leq \tau_2$, then $\mathcal{F}_{\tau_1} \subseteq \mathcal{F}_{\tau_2}$.

Advantages of a Right Continuous Filtration

A filtration $\{\mathcal{F}_t\}$ is said to be *right-continuous* if $\mathcal{F}_{t^+} = \mathcal{F}_t$ for each t where $\mathcal{F}_{t^+} = \cap_{u>t}\mathcal{F}_u$.

When a given filtration satisfies the additional requirement of right-continuity, the space $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$ is said to satisfy the **usual conditions**. Some of the niceties that accrue as a result are listed below:

(i) If $\{\tau_n\}$ is a sequence of stopping times, $\inf_n \tau_n$, $\liminf \tau_n$, and $\limsup \tau_n$ are stopping times.
(ii) If $\tau = \inf_n \tau_n$, then $\mathcal{F}_\tau = \cap_n \mathcal{F}_{\tau_n}$.
(iii) Let X be an adapted, right-continuous process. The hitting time of an open set is a stopping time.

The statements above are quite easy to prove and are left to the reader. Whenever right-continuity of a filtration is used, it would be expressly mentioned.

Progressive Measurability

Definition 1.3.3 *Let $X = \{X_t\}$ be a stochastic process on $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$. We call X* **progressively measurable** *with respect to $\mathcal{F}_t$ if for each $t \in \mathbb{R}^+$, the map*

$$(s, \omega) \to X_s(\omega)$$

from $[0, t] \times \Omega$ to $\mathbb{R}$ is measurable with respect to $\mathcal{B}[0, t] \times \mathcal{F}_t$.

The following proposition shows that the class of progressively measurable processes is huge and contains several processes of interest. If a filtration $(\mathcal{F}_t)$ is given and doesn't change, we will simply write "progressively measurable" instead of "progressively measurable with respect to $\mathcal{F}_t$".

Proposition 1.3.2 *Let $X = \{X_t\}$ be an $\mathcal{F}_t$-adapted, right-continuous (or left-continuous) process. Then, X is progressively measurable.*

Proof Suppose X has right-continuous sample paths. Fix any $t > 0$. For each n, define the stepwise approximation process X^n on $[0, t]$ by

$$X_s^n := \sum_{j=0}^{n-1} X_{(j+1)t/n} 1_{(jt/n,(j+1)t/n]}(s)$$

with $X_0^n = X_0$. Clearly, X^n is $\mathcal{B}([0, t]) \times \mathcal{F}_t$-measurable. By right-continuity of X, we get $\lim_{n\to\infty} X_s^n(\omega) = X_s(\omega)$ for all $(s, \omega) \in [0, t] \times \Omega$. Therefore, X as a function on $[0, t] \times \Omega$ is also $\mathcal{B}([0, t]) \times \mathcal{F}_t$-measurable.

One can likewise prove it when X has left-continuous paths. ■

A nice feature of progressively measurable processes is given in the next proposition.

Proposition 1.3.3 *Let $X = \{X_t\}$ be progressively measurable, and let τ be a finite stopping time. Then, the random variable X_τ is $\mathcal{F}_\tau$-measurable.*

Proof We will first show that the process $\{X_{\tau\wedge t}\}$ is progressively measurable.

Indeed, the process $\{\tau \wedge t\}$ is continuous and adapted to $(\mathcal{F}_t)$ and hence is progressively measurable by the previous proposition. Therefore, the map

$$(s, \omega) \to (\tau(\omega) \wedge s, \omega)$$

is measurable from $([0, t] \times \Omega, \mathcal{B}([0, t]) \times \mathcal{F}_t)$ into itself. By progressive measurability of X, we thus obtain that the map

$$(s, \omega) \to X_{\tau(\omega)\wedge s}(\omega)$$

is $\mathcal{B}([0, t]) \times \mathcal{F}_t$-measurable as a map from $[0, t] \times \Omega$ to $\mathbb{R}$.

Thus, $\{X_{\tau\wedge t}\}$ is a progressively measurable process.

$X_{\tau\wedge t}$ is $\mathcal{F}_t$-measurable for any fixed t. For any Borel set B in $\mathbb{R}$, it follows that for any t,

$$\{X_\tau \in B\} \cap \{\tau \leq t\} = \{X_{\tau\wedge t} \in B\} \cap \{\tau \leq t\} \in \mathcal{F}_t.$$

In other words, X_τ is $\mathcal{F}_\tau$-measurable. ■

Some Deep Results for Stochastic Processes

We list below a few results that are powerful, and are very difficult to prove. It is important to know them, though we do not prove these theorems in this book. The first result is due to Dellacherie [12].

Theorem 1.3.4 *Let $X = \{X_t\}$ be any progressively measurable process on $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$ where $(\mathcal{F}_t)$ satisfies the usual conditions. For any Borel set B in $\mathbb{R}$, the hitting time τ_B is a stopping time.*

The next result is due to Chung and Doob [5] (also see [53] p. 68). It is quite clear that a progressively measurable process is measurable and adapted. The converse is almost (but not quite!) true, and is given by this famous result.

Theorem 1.3.5 *If a stochastic process $X = \{X_t\}$ is measurable and adapted to a filtration $(\mathcal{F}_t)$, then it has a progressively measurable modification.*

Many processes of interest happen to be right-continuous and adapted and are therefore progressively measurable by Proposition 1.3.2. It is for this reason that we do not need Theorem 1.3.5 in many instances.

Exercises

1. Let $\Omega = (0, 1]$ equipped with its Borel σ-field. For any $n \in \mathbb{N}$, define the random variables

$$X_n = 1_{(0,\frac{1}{n})}.$$

Let $\mathcal{A} = \cup_{n=1}^{\infty} \sigma(X_1, \ldots, X_n)$, and $\mathcal{F} = \sigma(\mathcal{A})$.

(i) Characterize the sets in $\mathcal{A}$ and in $\mathcal{F}$.

(ii) For any Borel set $B \in \mathcal{B}(\mathbb{R}^n)$, assign

$$P\left\{(X_1, \ldots, X_n) \in B\right\} = 1 \text{ if } (1, \ldots, 1) \in B$$

and equal to zero otherwise. Prove that P is additive on $\mathcal{A}$ but there is no extension of P to a probability measure on $\mathcal{F}$.

2. Let P and Q be two probability measures on $(\Omega, \mathcal{F})$. Let $\mathcal{S}$ be a π system such that $\mathcal{F} = \sigma(\mathcal{S})$. If $P = Q$ on $\mathcal{S}$, show that $P = Q$ on $\mathcal{F}$.
3. Let $\mathcal{S}$ be as in the previous problem. Suppose L is a space of $\mathcal{F}$-measurable functions such that
   (i) $1 \in L$; $I_A \in L \ \forall A \in \mathcal{S}$,
   (ii) $f, g \in L$, then $af + bg \in L$ for all nonnegative constants a, b, and
   (iii) If f_n is a non-decreasing sequence of nonnegative functions in L such that $\lim_{n\to\infty} f_n = f$, then $f \in L$.

   Show that L contains all nonnegative $\mathcal{F}$-measurable functions. This is the *monotone class* theorem for measurable functions.
4. Let $(\mathcal{F}_t : 0 \leq t \leq \infty)$ be a filtration on $(\Omega, \mathcal{F})$. Let τ be a stopping time with respect to $(\mathcal{F}_t)$.
   (i) Show that $\mathcal{F}_\tau$ is a σ-field.
   (ii) Prove that τ is $\mathcal{F}_\tau$ measurable.
5. Let $(\mathcal{F}_t : 0 \leq t \leq \infty)$ be a filtration on $(\Omega, \mathcal{F})$. Let τ_1, τ_2 be two stopping times with respect to $(\mathcal{F}_t)$. Show that $\tau_1 \wedge \tau_2$, $\tau_1 \vee \tau_2$, and $\tau_1 + \tau_2$ are stopping times.
6. Let S, T be two stopping times with $S \leq T$ a.s. Show that
   (i) $\mathcal{F}_S \subset \mathcal{F}_T$, and
   (ii) $\mathcal{F}_S \cap \mathcal{F}_T = \mathcal{F}_{S\wedge T}$.
7. Suppose that a filtration $(\mathcal{F}_t)$ is right-continuous. If $\{\tau_n\}$ is a sequence of stopping times with respect to $(\mathcal{F}_t)$, prove that $\inf \tau_n$, $\liminf_{n\to\infty} \tau_n$, and $\limsup_{n\to\infty} \tau_n$ are stopping times. Note that $\sup \tau_n$ is a stopping time even without right-continuity of the filtration.
8. Given any filtration $(\mathcal{F}_t)$, show that $(\mathcal{F}_{t+})$ is a right-continuous filtration.
9. Consider a probability space $(\Omega, \mathcal{F}, P)$ with a filtration $(\mathcal{F}_t)$. Let $\mathcal{P}$ be the σ-field on $\mathbb{R}^+ \times \Omega$ generated by all continuous, adapted, real-valued processes. Show that $\mathcal{P}$ is generated by the family of sets

$$\mathcal{R} = \{(s, t] \times F : s \leq t, F \in \mathcal{F}_s\} \cup \{0 \times F : F \in \mathcal{F}_0\}.$$

   Conclude that $\mathcal{P}$ is generated by all adapted, left-continuous real-valued processes. $\mathcal{P}$ is known as the *predictable* σ-field.
10. Suppose that X and Y are equivalent processes and X is right-continuous. Show that there exists a modification of Y which is right-continuous.