Stochastic Processes

Lecture Notes

Master M1 — 2025–2026

Yaë Ulrich Gaba

"Chance favours only the prepared mind."
— Louis Pasteur

March 25, 2026

Contents

Preface

Course Objectives

This **Stochastic Processes** course is aimed at Master's-level students in mathematics, theoretical physics, quantitative finance, and engineering sciences. It provides a rigorous, measure-theoretic presentation of the main classes of random processes and their applications.

The primary objective is to guide the student from probability foundations through to the construction of the Itô stochastic integral and the solution of stochastic differential equations (SDEs), via in-depth study of Markov chains, the Poisson process, martingales, and Brownian motion.

Upon completing this course, students will be able to:

- Master the measure-theoretic formalism of stochastic processes.
- Analyse Markov chains in discrete and continuous time: classify states, compute stationary distributions, establish ergodicity.
- Work with martingales: apply fundamental inequalities, convergence theorems, and the optional stopping theorem.
- Understand the fine properties of Brownian motion.
- Construct the Itô stochastic integral and apply Itô's formula.
- Solve stochastic differential equations and apply them in finance, physics, and biology.

Prerequisites

The essential prerequisites are:

- **Measure theory and integration**: σ-algebras, measures, Lebesgue integral, convergence theorems (dominated, monotone), L^p spaces.
- **Undergraduate probability**: random variables, expectation, variance, law of large numbers, central limit theorem.
- **Elementary functional analysis**: Banach and Hilbert spaces, modes of convergence, completeness.
- **Linear algebra**: matrices, eigenvalues, spectral decomposition.

- **Ordinary differential equations**: existence and uniqueness theorems, solution methods.

Course Organisation

The course is organised into eleven chapters following a logical progression:

1. **Probability Review**: probability spaces, conditional expectation, modes of convergence. This chapter provides the measure-theoretic foundations needed for the rest of the course.

2. **Stochastic Processes — Definitions**: formal definition of a stochastic process, filtrations, stopping times, adapted and predictable processes. These notions form the basic language.

3. **Discrete-Time Markov Chains**: Markov property, state classification, ergodic theorems, stationary measures.

4. **Continuous-Time Markov Chains**: infinitesimal generator, Kolmogorov equations, birth-and-death processes.

5. **Poisson Process**: axiomatic construction, fundamental properties, compound and non-homogeneous Poisson processes.

6. **Martingales — Discrete Time**: definition, fundamental inequalities (Doob), convergence theorems, optional stopping theorem.

7. **Martingales — Continuous Time**: path regularity, Doob–Meyer decomposition, local martingales.

8. **Brownian Motion**: construction (Lévy–Ciesielski), path properties, nowhere differentiability, law of the iterated logarithm, reflection principle.

9. **Itô Stochastic Integral**: construction via simple processes, isometric extension to L^2 processes, martingale properties.

10. **Itô's Formula and SDEs**: Itô change-of-variables formula, multidimensional Itô formula, existence and uniqueness of strong solutions.

11. **Applications**: finance (Black–Scholes model, hedging), physics (Langevin equation, diffusion), biology (stochastic population models).

Conventions and Notation

Throughout this course, we adopt the following conventions:

- $(\Omega, \mathcal{F}, \mathbb{P})$ denotes a fixed probability space.

- $\{\mathcal{F}_t\}_{t\geq 0}$ or $\{\mathcal{F}_n\}_{n\geq 0}$ denotes a filtration satisfying the usual conditions (completeness, right-continuity).

- $\mathbb{E}[X]$ denotes the expectation of the random variable X.

- $\mathbb{E}[X \mid \mathcal{G}]$ denotes the conditional expectation of X given the σ-algebra $\mathcal{G}$.
- $\mathbb{P}(A)$ denotes the probability of the event A.
- $\mathrm{Var}(X)$ and $\mathrm{Cov}(X, Y)$ denote variance and covariance.
- $\mathbb{R}$, $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{C}$ denote the sets of reals, naturals, integers, and complex numbers. We write $\mathbb{R}_+ = [0, +\infty)$.
- $\|\cdot\|$ and $|\cdot|$ denote a norm and the absolute value.
- $\mathbf{1}_A$ denotes the indicator function of the set A.
- a.s. means "almost surely", r.v. means "random variable", SDE means "stochastic differential equation".
- $\xrightarrow{\text{a.s.}}$, $\xrightarrow{L^p}$, $\xrightarrow{\mathbb{P}}$, $\xrightarrow{\mathcal{L}}$ denote almost sure, L^p, in probability, and distributional convergence.

Pedagogical Approach

Each chapter contains:

- Precise **definitions** within the measure-theoretic framework.
- **Theorems** with complete or sketched proofs.
- Detailed **examples** illustrating abstract concepts.
- **Exercises** of varying difficulty, some with hints.
- *Intuition* boxes to develop probabilistic intuition.
- *Warning* boxes to flag common pitfalls.
- *Key formulas* boxes summarising essential results.

Bibliographic References

The following texts have inspired this course:

- Brémaud, P. *Markov Chains: Gibbs Fields, Monte Carlo Simulation and Queues.* Springer, 1999.
- Durrett, R. *Probability: Theory and Examples.* Cambridge University Press, 5th ed., 2019.
- Karatzas, I. and Shreve, S.E. *Brownian Motion and Stochastic Calculus.* Springer, 2nd ed., 1991.
- Le Gall, J.-F. *Brownian Motion, Martingales, and Stochastic Calculus.* Springer, 2016.

- Norris, J.R. *Markov Chains.* Cambridge University Press, 1997.
- Øksendal, B. *Stochastic Differential Equations.* Springer, 6th ed., 2003.
- Revuz, D. and Yor, M. *Continuous Martingales and Brownian Motion.* Springer, 3rd ed., 1999.
- Shreve, S.E. *Stochastic Calculus for Finance I & II.* Springer, 2004.
- Williams, D. *Probability with Martingales.* Cambridge University Press, 1991.

Acknowledgements

The author thanks colleagues and students whose remarks and suggestions have helped improve this course over the years. Discussions with practitioners in quantitative finance and physicists have also enriched the presentation of applications.

The Author
March 2026

Chapter 1

Probability Review

Before diving into the study of stochastic processes — those random functions that evolve over time — we must ensure that our probabilistic foundations are solid. This chapter reviews the essential tools: probability spaces, random variables, conditional expectation, and modes of convergence. Readers already comfortable with these notions may skim this chapter as a reference; otherwise, this is where everything begins.

1.1 Probability Spaces

Let us recall the three fundamental ingredients of Kolmogorov's framework.

Definition 1.1 (σ-Algebra)**.** Let Ω be a non-empty set. A *σ-algebra* (or σ-field) $\mathcal{F}$ on Ω is a collection of subsets of Ω satisfying:

1. $\Omega \in \mathcal{F}$,
2. if $A \in \mathcal{F}$, then $A^c \in \mathcal{F}$ (closed under complementation),
3. if $(A_n)_{n\geq 1} \subset \mathcal{F}$, then $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$ (closed under countable unions).

The pair $(\Omega, \mathcal{F})$ is called a *measurable space*.

Definition 1.2 (Probability Space)**.** A *probability space* is a triple $(\Omega, \mathcal{F}, \mathbb{P})$ where:

- Ω is the *sample space* (set of all possible outcomes),
- $\mathcal{F}$ is a σ-algebra on Ω (the events),
- $\mathbb{P} : \mathcal{F} \to [0, 1]$ is a *probability measure*, i.e. a σ-additive measure with $\mathbb{P}(\Omega) = 1$.

Remark 1.3. The σ-additivity of $\mathbb{P}$ means: for any sequence $(A_n)_{n\geq 1}$ of pairwise disjoint events,

$$\mathbb{P}\Big(\bigcup_{n=1}^{\infty} A_n\Big) = \sum_{n=1}^{\infty} \mathbb{P}(A_n).$$

Definition 1.4 (Generated σ-Algebra)**.** Let $\mathcal{C}$ be a collection of subsets of Ω. The *σ-algebra generated* by $\mathcal{C}$, denoted $\sigma(\mathcal{C})$, is the smallest σ-algebra containing $\mathcal{C}$:

$$\sigma(\mathcal{C}) = \bigcap \big\{\mathcal{F} : \mathcal{F} \text{ is a } \sigma\text{-algebra on } \Omega,\ \mathcal{C} \subset \mathcal{F}\big\}.$$

In particular, the *Borel σ-algebra* $\mathcal{B}(\mathbb{R})$ is generated by the open sets of $\mathbb{R}$.

 CHAPTER 1. PROBABILITY REVIEW

1.2 Random Variables

Definition 1.5 (Random Variable)**.** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. A *random variable* (r.v.) is a measurable map $X : (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$, i.e.:

$$\forall B \in \mathcal{B}(\mathbb{R}), \quad X^{-1}(B) = \{\omega \in \Omega : X(\omega) \in B\} \in \mathcal{F}.$$

More generally, a *random vector* is a measurable map $X : (\Omega, \mathcal{F}) \to (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$.

Definition 1.6 (σ-Algebra Generated by a Random Variable)**.** The σ-algebra generated by a random variable X is

$$\sigma(X) = \{X^{-1}(B) : B \in \mathcal{B}(\mathbb{R})\}.$$

This is the smallest σ-algebra making X measurable. Intuitively, $\sigma(X)$ represents the *information* contained in the observation of X.

Theorem 1.7 (Doob's Measurability Lemma)**.** *Let X be a random variable and $\mathcal{G} = \sigma(X)$. A random variable Y is $\mathcal{G}$-measurable if and only if there exists a Borel measurable function $g : \mathbb{R} \to \mathbb{R}$ such that $Y = g(X)$.*

1.3 Expectation and Integration

Definition 1.8 (Expectation)**.** The *expectation* of a non-negative random variable $X \geq 0$ is

$$\mathbb{E}[X] = \int_\Omega X(\omega)\, d\mathbb{P}(\omega).$$

For a general r.v., $\mathbb{E}[X] = \mathbb{E}[X^+] - \mathbb{E}[X^-]$ when at least one of these quantities is finite, where $X^+ = \max(X, 0)$ and $X^- = \max(-X, 0)$. We say X is *integrable* if $\mathbb{E}[|X|] < \infty$.

Theorem 1.9 (Lebesgue Dominated Convergence)**.** *Let $(X_n)_{n\geq 1}$ be a sequence of r.v. converging a.s. to X. If there exists an integrable r.v. Y such that $|X_n| \leq Y$ a.s. for all n, then X is integrable and*

$$\lim_{n\to\infty} \mathbb{E}[X_n] = \mathbb{E}[X].$$

Theorem 1.10 (Monotone Convergence)**.** *Let $(X_n)_{n\geq 1}$ be an increasing sequence of non-negative r.v.: $0 \leq X_1 \leq X_2 \leq \cdots$ a.s. Then*

$$\lim_{n\to\infty} \mathbb{E}[X_n] = \mathbb{E}\big[\lim_{n\to\infty} X_n\big].$$

Theorem 1.11 (Fatou's Lemma)**.** *Let $(X_n)_{n\geq 1}$ be a sequence of non-negative r.v. Then*

$$\mathbb{E}\big[\liminf_{n\to\infty} X_n\big] \leq \liminf_{n\to\infty} \mathbb{E}[X_n].$$

1.4 L^p Spaces and Inequalities

Definition 1.12 (L^p Space)**.** For $p \in [1, \infty)$, the space $L^p(\Omega, \mathcal{F}, \mathbb{P})$ is the set of (equivalence classes of) r.v. X with $\mathbb{E}[|X|^p] < \infty$, equipped with the norm

$$\|X\|_{L^p} = \big(\mathbb{E}[|X|^p]\big)^{1/p}.$$

The space L^∞ is the set of essentially bounded r.v.

Theorem 1.13 (Hölder's Inequality). *Let $p, q \in (1, \infty)$ with $\frac{1}{p} + \frac{1}{q} = 1$. For $X \in L^p$ and $Y \in L^q$,*

$$\mathbb{E}[|XY|] \leq \|X\|_{L^p} \|Y\|_{L^q}.$$

The case $p = q = 2$ gives the Cauchy–Schwarz inequality*.*

Theorem 1.14 (Jensen's Inequality). *If X is integrable and $\varphi : \mathbb{R} \to \mathbb{R}$ is convex, then*

$$\varphi(\mathbb{E}[X]) \leq \mathbb{E}[\varphi(X)],$$

provided $\mathbb{E}[|\varphi(X)|] < \infty$.

1.5 Conditional Expectation

Definition 1.15 (Conditional Expectation). Let $X \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ and let $\mathcal{G} \subset \mathcal{F}$ be a sub-σ-algebra. The *conditional expectation* of X given $\mathcal{G}$, denoted $\mathbb{E}[X \mid \mathcal{G}]$, is the unique (a.s.) random variable satisfying:

1. $\mathbb{E}[X \mid \mathcal{G}]$ is $\mathcal{G}$-measurable,
2. for all $G \in \mathcal{G}$, $\displaystyle\int_G \mathbb{E}[X \mid \mathcal{G}]\, d\mathbb{P} = \int_G X\, d\mathbb{P}$.

Existence is guaranteed by the Radon–Nikodym theorem.

Theorem 1.16 (Properties of Conditional Expectation). *Let $X, Y \in L^1$ and let $\mathcal{G}, \mathcal{H}$ be sub-σ-algebras of $\mathcal{F}$.*

1. ***Linearity**: $\mathbb{E}[aX + bY \mid \mathcal{G}] = a\mathbb{E}[X \mid \mathcal{G}] + b\mathbb{E}[Y \mid \mathcal{G}]$ a.s.*
2. ***Positivity**: if $X \geq 0$ a.s., then $\mathbb{E}[X \mid \mathcal{G}] \geq 0$ a.s.*
3. ***Tower property (smoothing)**: if $\mathcal{H} \subset \mathcal{G}$, then $\mathbb{E}[\mathbb{E}[X \mid \mathcal{G}] \mid \mathcal{H}] = \mathbb{E}[X \mid \mathcal{H}]$ a.s.*
4. ***Pull-out property**: if Y is $\mathcal{G}$-measurable and bounded, $\mathbb{E}[XY \mid \mathcal{G}] = Y\mathbb{E}[X \mid \mathcal{G}]$ a.s.*
5. ***Independence**: if $\sigma(X)$ and $\mathcal{G}$ are independent, $\mathbb{E}[X \mid \mathcal{G}] = \mathbb{E}[X]$ a.s.*
6. ***Conditional Jensen**: if φ is convex and $\varphi(X) \in L^1$, $\varphi(\mathbb{E}[X \mid \mathcal{G}]) \leq \mathbb{E}[\varphi(X) \mid \mathcal{G}]$ a.s.*

Geometric Interpretation

In $L^2(\Omega, \mathcal{F}, \mathbb{P})$, the conditional expectation $\mathbb{E}[X \mid \mathcal{G}]$ is the *orthogonal projection* of X onto the closed subspace $L^2(\Omega, \mathcal{G}, \mathbb{P})$. This justifies:

$$\mathbb{E}\left[\left(X - \mathbb{E}[X \mid \mathcal{G}]\right)^2\right] \leq \mathbb{E}\left[(X - Z)^2\right]$$

for all $Z \in L^2(\Omega, \mathcal{G}, \mathbb{P})$.

1.6 Independence

Definition 1.17 (Independence of σ-Algebras)**.** Sub-σ-algebras $\mathcal{G}_1, \ldots, \mathcal{G}_n$ of $\mathcal{F}$ are *independent* if for every choice $A_i \in \mathcal{G}_i$:

$$\mathbb{P}(A_1 \cap \cdots \cap A_n) = \mathbb{P}(A_1) \cdots \mathbb{P}(A_n).$$

Random variables $X_1, \ldots, X_n$ are independent if $\sigma(X_1), \ldots, \sigma(X_n)$ are independent.

Definition 1.18 (Mutual Independence)**.** A family $(X_i)_{i \in I}$ of random variables is *mutually independent* if for every finite subset $J \subset I$, the random variables $(X_j)_{j \in J}$ are independent.

Theorem 1.19 (Kolmogorov's Zero–One Law)**.** *Let $(X_n)_{n \geq 1}$ be a sequence of independent r.v. The* tail σ-algebra

$$\mathcal{T} = \bigcap_{n=1}^{\infty} \sigma(X_n, X_{n+1}, \ldots)$$

is trivial*: for every $A \in \mathcal{T}$, $\mathbb{P}(A) \in \{0, 1\}$.*

1.7 Modes of Convergence

Definition 1.20 (Types of Convergence)**.** Let $(X_n)_{n \geq 1}$ be a sequence of r.v. and X a r.v. defined on $(\Omega, \mathcal{F}, \mathbb{P})$.

1. **Almost sure convergence**: $X_n \xrightarrow{\text{a.s.}} X$ if $\mathbb{P}(\{\omega : X_n(\omega) \to X(\omega)\}) = 1$.
2. L^p **convergence**: $X_n \xrightarrow{L^p} X$ if $\mathbb{E}[|X_n - X|^p] \to 0$.
3. **Convergence in probability**: $X_n \xrightarrow{\mathbb{P}} X$ if for every $\varepsilon > 0$, $\mathbb{P}(|X_n - X| > \varepsilon) \to 0$.
4. **Convergence in distribution**: $X_n \xrightarrow{\mathcal{L}} X$ if $\mathbb{E}[f(X_n)] \to \mathbb{E}[f(X)]$ for every bounded continuous $f : \mathbb{R} \to \mathbb{R}$.

Hierarchy of Convergences

$$\text{a.s.} \implies \text{in probability} \implies \text{in distribution}$$

$$L^p \implies L^q \;\; (q \leq p) \implies \text{in probability} \implies \text{in distribution}$$

The converses are false in general. However:

- $X_n \xrightarrow{\mathbb{P}} X \implies$ there exists a subsequence $X_{n_k} \xrightarrow{\text{a.s.}} X$.
- If $X_n \xrightarrow{\mathcal{L}} c$ (constant), then $X_n \xrightarrow{\mathbb{P}} c$.

1.8 Laws of Large Numbers

Theorem 1.21 (Strong Law of Large Numbers)**.** *Let $(X_n)_{n \geq 1}$ be a sequence of i.i.d. integrable r.v. with mean $\mu = \mathbb{E}[X_1]$. Then*

$$\frac{1}{n} \sum_{k=1}^{n} X_k \xrightarrow{a.s.} \mu.$$

Theorem 1.22 (Central Limit Theorem). *Let $(X_n)_{n\geq 1}$ be i.i.d. with $\mathbb{E}[X_1] = \mu$ and* $\mathrm{Var}(X_1) = \sigma^2 \in (0, \infty)$. *Then*

$$\frac{1}{\sqrt{n}} \sum_{k=1}^{n} \frac{X_k - \mu}{\sigma} \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1).$$

1.9 Borel–Cantelli Lemma

Theorem 1.23 (Borel–Cantelli Lemma). *Let $(A_n)_{n\geq 1}$ be a sequence of events.*

1. *If $\sum_{n=1}^{\infty} \mathbb{P}(A_n) < \infty$, then $\mathbb{P}(\limsup_n A_n) = 0$, i.e. a.s. only finitely many A_n occur.*
2. *If the A_n are* independent *and $\sum_{n=1}^{\infty} \mathbb{P}(A_n) = \infty$, then $\mathbb{P}(\limsup_n A_n) = 1$, i.e. a.s. infinitely many A_n occur.*

Independence Hypothesis

The second part of the Borel–Cantelli lemma requires independence of the events. Without this hypothesis, divergence of $\sum \mathbb{P}(A_n)$ does not guarantee $\mathbb{P}(\limsup A_n) = 1$.

1.10 Characteristic Functions

Definition 1.24 (Characteristic Function). The *characteristic function* of a r.v. X is

$$\varphi_X(t) = \mathbb{E}[e^{itX}], \quad t \in \mathbb{R}.$$

Proposition 1.25 (Properties).

1. $\varphi_X(0) = 1$ and $|\varphi_X(t)| \leq 1$ for all t.
2. φ_X is uniformly continuous on $\mathbb{R}$.
3. φ_X characterises the law of X: if $\varphi_X = \varphi_Y$, then $X \stackrel{\mathcal{L}}{=} Y$.
4. If X and Y are independent, $\varphi_{X+Y}(t) = \varphi_X(t)\varphi_Y(t)$.
5. If $\mathbb{E}[|X|^n] < \infty$, then $\varphi_X^{(n)}(0) = i^n \mathbb{E}[X^n]$.

Theorem 1.26 (Lévy's Continuity Theorem). *$X_n \xrightarrow{\mathcal{L}} X$ if and only if $\varphi_{X_n}(t) \to \varphi_X(t)$ for all $t \in \mathbb{R}$.*

1.11 Exercises

Exercise 1.1. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and X a non-negative r.v. Show that $\mathbb{E}[X] = \int_0^{\infty} \mathbb{P}(X > t)\, dt$.

Exercise 1.2. Show that a.s. convergence does not imply L^1 convergence by constructing a counterexample.

Exercise 1.3. Let $(X_n)_{n\geq 1}$ be independent r.v. with $\mathbb{P}(X_n = n) = 1/n$ and $\mathbb{P}(X_n = 0) = 1 - 1/n$. Show that $X_n \to 0$ a.s.

Exercise 1.4. Prove the tower property of conditional expectation (Theorem 1.16, item 3).

Exercise 1.5. Let X be integrable and $\mathcal{G}_n \uparrow \mathcal{G}_\infty$ an increasing filtration of sub-σ-algebras. Show that $\mathbb{E}[X \mid \mathcal{G}_n] \to \mathbb{E}[X \mid \mathcal{G}_\infty]$ a.s. and in L^1. *(This is Lévy's martingale convergence theorem.)*

Chapter 2

Stochastic Processes — Definitions

What exactly is a stochastic process? The question admits two equally valid answers, and the tension between them runs through the entire theory. On one hand, a stochastic process is a family of random variables indexed by time: at each instant t, we draw a value from some distribution. On the other hand, it is a single random function—a trajectory, a path, a realisation chosen at random from a space of possible histories. The first viewpoint is algebraic and distributional; the second is geometric and pathwise.

This dual perspective was clarified in the early twentieth century by Andrei Kolmogorov and Joseph Doob, who built the rigorous foundations of stochastic process theory. Kolmogorov's consistency theorem (1933) showed that specifying all finite-dimensional distributions suffices to construct the process, while Doob's work on regularity conditions revealed when one can demand that the paths be continuous or at least well-behaved. Understanding this duality—distributions versus paths—is the first step toward mastering the theory.

2.1 Definition of a Stochastic Process

Definition 2.1 (Stochastic Process)**.** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and T an index set (typically $T = \mathbb{N}$, $\mathbb{Z}_+$, or $\mathbb{R}_+$). A *stochastic process* with values in a measurable space $(E, \mathcal{E})$ is a family of random variables

$$X = (X_t)_{t \in T}, \quad X_t : \Omega \to E,$$

where each X_t is $\mathcal{F}/\mathcal{E}$-measurable.

Equivalently, X is a measurable map $X : (T \times \Omega, \mathcal{B}(T) \otimes \mathcal{F}) \to (E, \mathcal{E})$ (when joint measurability is required).

Definition 2.2 (Sample Path)**.** For a fixed $\omega \in \Omega$, the map $t \mapsto X_t(\omega)$ is called the *sample path* (or *trajectory*, or *realisation*) of the process associated with ω.

Example 2.3.

1. **Simple random walk**: let $(Y_n)_{n \geq 1}$ be i.i.d. with $\mathbb{P}(Y_n = 1) = \mathbb{P}(Y_n = -1) = 1/2$. The process $S_n = \sum_{k=1}^{n} Y_k$ for $n \geq 0$ (with $S_0 = 0$) is a discrete-time, $\mathbb{Z}$-valued stochastic process.

2. **Poisson process**: $N = (N_t)_{t \geq 0}$ with values in $\mathbb{N}$, having increasing càdlàg sample paths.

3. **Brownian motion**: $B = (B_t)_{t \geq 0}$ with values in $\mathbb{R}$, having continuous sample paths.

CHAPTER 2. STOCHASTIC PROCESSES — DEFINITIONS

2.2 Classification of Processes

Definition 2.4 (Discrete vs. Continuous Time).

- A process is *discrete-time* if T is countable (e.g. $T = \mathbb{N}$ or $T = \mathbb{Z}$).
- A process is *continuous-time* if T is an interval in $\mathbb{R}$ (e.g. $T = \mathbb{R}_+$ or $T = [0, T_0]$).

Definition 2.5 (State Space). The space $(E, \mathcal{E})$ is called the *state space* of the process.

- If E is countable (finite or infinite), we have a *discrete state space*.
- If $E = \mathbb{R}^d$, we have a *continuous state space*.

Definition 2.6 (Path Regularity). A real-valued process $(X_t)_{t\geq 0}$ is said to be:

- *continuous* if a.s. $t \mapsto X_t(\omega)$ is continuous,
- *càdlàg* (right-continuous with left limits) if a.s. $X_t = X_{t+}$ for all t and X_{t-} exists,
- *càglàd* (left-continuous with right limits) if a.s. $X_t = X_{t-}$ for all t and X_{t+} exists.

2.3 Filtrations

Definition 2.7 (Filtration). A *filtration* on $(\Omega, \mathcal{F}, \mathbb{P})$ is an increasing family of sub-σ-algebras $\mathbb{F} = (\mathcal{F}_t)_{t\in T}$:

$$s \leq t \implies \mathcal{F}_s \subset \mathcal{F}_t \subset \mathcal{F}.$$

The space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$ is called a *filtered probability space*.

Increasing Information

The filtration models the *information available over time*. The σ-algebra $\mathcal{F}_t$ represents the set of events that can be observed (decided as occurred or not) at time t. The increase reflects the fact that information is never lost: what is known at time s remains known at time $t > s$.

Definition 2.8 (Natural Filtration). The *natural filtration* (or *generated filtration*) of a process $(X_t)_{t\in T}$ is

$$\mathcal{F}_t^X = \sigma(X_s : s \leq t), \quad t \in T.$$

This is the smallest filtration with respect to which (X_t) is adapted.

Definition 2.9 (Usual Conditions). A filtration $(\mathcal{F}_t)_{t\geq 0}$ satisfies the *usual conditions* (of Dellacherie–Meyer) if:

1. **Completeness**: $\mathcal{F}_0$ contains all $\mathbb{P}$-null sets of $\mathcal{F}$.
2. **Right-continuity**: for every $t \geq 0$, $\mathcal{F}_t = \mathcal{F}_{t+} := \bigcap_{s>t} \mathcal{F}_s$.

Remark 2.10. Any filtration can be "completed" and "regularised". Let $\mathcal{N}$ be the collection of $\mathbb{P}$-null sets. Define

$$\overline{\mathcal{F}}_t = \sigma(\mathcal{F}_t \cup \mathcal{N}), \quad \widehat{\mathcal{F}}_t = \bigcap_{s>t} \overline{\mathcal{F}}_s.$$

The filtration $(\widehat{\mathcal{F}}_t)$ satisfies the usual conditions.

2.4 Adapted and Predictable Processes

Definition 2.11 (Adapted Process)**.** A process $(X_t)_{t\in T}$ is *adapted* to the filtration $(\mathcal{F}_t)_{t\in T}$ if for every $t \in T$, X_t is $\mathcal{F}_t$-measurable.

Definition 2.12 (Predictable Process — Discrete Time)**.** In discrete time, a process $(X_n)_{n\geq 0}$ is *predictable* with respect to $(\mathcal{F}_n)$ if X_0 is deterministic (or $\mathcal{F}_0$-measurable) and X_n is $\mathcal{F}_{n-1}$-measurable for all $n \geq 1$.

Definition 2.13 (Predictable σ-Algebra — Continuous Time)**.** In continuous time, the *predictable* σ*-algebra* $\mathcal{P}$ on $\mathbb{R}_+ \times \Omega$ is generated by sets of the form $(s,t] \times F$ with $0 \leq s < t$ and $F \in \mathcal{F}_s$, and sets $\{0\} \times F_0$ with $F_0 \in \mathcal{F}_0$.

A process is *predictable* if it is measurable with respect to $\mathcal{P}$.

Remark 2.14. Every adapted left-continuous process is predictable. Every predictable process is adapted. However, an adapted càdlàg process is not necessarily predictable.

2.5 Stopping Times

Definition 2.15 (Stopping Time)**.** Let $(\mathcal{F}_t)_{t\in T}$ be a filtration. A random variable $\tau : \Omega \to T \cup \{+\infty\}$ is a *stopping time* (or *Markov time*) if:

- In discrete time ($T = \mathbb{N}$): $\{\tau = n\} \in \mathcal{F}_n$ for all $n \in \mathbb{N}$.
- In continuous time ($T = \mathbb{R}_+$): $\{\tau \leq t\} \in \mathcal{F}_t$ for all $t \geq 0$.

Remark 2.16. In discrete time, $\{\tau = n\} \in \mathcal{F}_n$ is equivalent to $\{\tau \leq n\} \in \mathcal{F}_n$. In continuous time, if the filtration is right-continuous, the condition $\{\tau < t\} \in \mathcal{F}_t$ is equivalent.

Example 2.17.

1. Any constant $\tau = t_0$ is a stopping time.
2. The *first passage time* $\tau_a = \inf\{t \geq 0 : X_t = a\}$ is a stopping time if (X_t) is adapted with continuous (or càdlàg under the usual conditions) sample paths.
3. The *first entry time* into an open set G: $\tau_G = \inf\{t \geq 0 : X_t \in G\}$ is a stopping time under the usual conditions.

Definition 2.18 (Stopped σ-Algebra)**.** If τ is a stopping time, the σ*-algebra of events prior to* τ is

$$\mathcal{F}_\tau = \{A \in \mathcal{F} : \forall t \in T,\ A \cap \{\tau \leq t\} \in \mathcal{F}_t\}.$$

Proposition 2.19 (Properties of Stopping Times)**.** Let σ, τ be stopping times.

1. $\sigma \wedge \tau = \min(\sigma, \tau)$ and $\sigma \vee \tau = \max(\sigma, \tau)$ are stopping times.
2. If $\sigma \leq \tau$ a.s., then $\mathcal{F}_\sigma \subset \mathcal{F}_\tau$.
3. In discrete time, $\tau + 1$ is a stopping time.
4. $\{\sigma \leq \tau\}$, $\{\sigma < \tau\}$, $\{\sigma = \tau\}$ belong to $\mathcal{F}_\sigma \cap \mathcal{F}_\tau$.

Definition 2.20 (Stopped Process)**.** The process *stopped at* the stopping time τ is

$$X_t^\tau = X_{t\wedge\tau} = X_{\min(t,\tau)}, \quad t \in T.$$

CHAPTER 2. STOCHASTIC PROCESSES — DEFINITIONS

2.6 Finite-Dimensional Distributions

Definition 2.21 (Finite-Dimensional Distributions)**.** The *finite-dimensional distributions* of a process $(X_t)_{t\in T}$ are the laws of the random vectors $(X_{t_1}, \dots, X_{t_n})$ for every choice $t_1 < \cdots < t_n$ in T and every $n \geq 1$.

Theorem 2.22 (Kolmogorov Extension Theorem)**.** *Let $(E, \mathcal{E})$ be a Polish space. Let $(\mu_{t_1,\dots,t_n})$ be a family of probability measures on $(E^n, \mathcal{E}^{\otimes n})$ satisfying the consistency conditions:*

1. ***Permutation invariance**: for every permutation π of $\{1, \dots, n\}$,*

$$\mu_{t_{\pi(1)},\dots,t_{\pi(n)}}(B_{\pi(1)} \times \cdots \times B_{\pi(n)}) = \mu_{t_1,\dots,t_n}(B_1 \times \cdots \times B_n).$$

2. ***Marginalisation**: $\mu_{t_1,\dots,t_n}(B_1 \times \cdots \times B_n \times E) = \mu_{t_1,\dots,t_n}(B_1 \times \cdots \times B_n)$.*

Then there exists a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a process $(X_t)_{t\in T}$ whose finite-dimensional distributions are the $(\mu_{t_1,\dots,t_n})$.

Limitations of Kolmogorov's Theorem

Kolmogorov's extension theorem guarantees existence of a process with the correct finite-dimensional distributions, but *not* path regularity. Properties such as path continuity require additional results (e.g. the Kolmogorov–Čentsov theorem).

2.7 Kolmogorov–Čentsov Theorem

Theorem 2.23 (Kolmogorov–Čentsov)**.** *Let $(X_t)_{t\in[0,T]}$ be a real-valued stochastic process. If there exist constants $\alpha > 0$, $\beta > 0$, and $C > 0$ such that*

$$\mathbb{E}[|X_t - X_s|^\alpha] \leq C\,|t - s|^{1+\beta}, \quad \forall s, t \in [0, T],$$

then there exists a modification $(\tilde{X}_t)$ of (X_t) whose sample paths are a.s. Hölder continuous with exponent γ for every $\gamma < \beta/\alpha$.

Remark 2.24. Recall that a *modification* of (X_t) is a process $(\tilde{X}_t)$ defined on the same probability space with $\mathbb{P}(X_t = \tilde{X}_t) = 1$ for every t.

2.8 Gaussian Processes

Definition 2.25 (Gaussian Process)**.** A process $(X_t)_{t\in T}$ is *Gaussian* if for every choice $t_1, \dots, t_n \in T$, the vector $(X_{t_1}, \dots, X_{t_n})$ has a multivariate Gaussian distribution.

A Gaussian process is entirely determined by its *mean function* $m(t) = \mathbb{E}[X_t]$ and its *covariance function* $K(s, t) = \operatorname{Cov}(X_s, X_t)$.

Example 2.26. Standard Brownian motion is the Gaussian process with $m(t) = 0$ and $K(s, t) = \min(s, t)$. The *Brownian bridge* on $[0, 1]$ is the Gaussian process with $m(t) = 0$ and $K(s, t) = \min(s, t) - st$.

2.9 Stationarity and Ergodicity

Definition 2.27 (Strict Stationarity)**.** A process $(X_t)_{t\in T}$ is *strictly stationary* if for every $h > 0$ and every $n \geq 1$,

$$(X_{t_1+h}, \ldots, X_{t_n+h}) \stackrel{\mathcal{L}}{=} (X_{t_1}, \ldots, X_{t_n}).$$

Definition 2.28 (Weak Stationarity)**.** A square-integrable process (X_t) is *weakly stationary* (or *second-order stationary*, or *wide-sense stationary*) if:

1. $\mathbb{E}[X_t] = \mu$ is constant,
2. $\mathrm{Cov}(X_s, X_t) = R(t - s)$ depends only on $t - s$.

The function R is called the *autocorrelation function*.

2.10 Exercises

Exercise 2.1. Show that the natural filtration of a continuous process satisfies the right-continuity condition after completion.

Exercise 2.2. Let τ be a stopping time with respect to $(\mathcal{F}_n)_{n\geq 0}$. Show that X_τ is $\mathcal{F}_\tau$-measurable when (X_n) is adapted.

Exercise 2.3. Let $(B_t)_{t\geq 0}$ be a standard Brownian motion and $a > 0$. Show that $\tau_a = \inf\{t \geq 0 : B_t = a\}$ is a stopping time with respect to the completed natural filtration of B.

Exercise 2.4. Verify that the Brownian bridge $(X_t)_{t\in[0,1]}$ defined by $X_t = B_t - tB_1$ is a Gaussian process and compute its covariance function.

Exercise 2.5. Let $(X_t)_{t\geq 0}$ satisfy the hypotheses of the Kolmogorov–Čentsov theorem with $\alpha = 4$ and $\beta = 1$. What Hölder exponents are possible for the continuous modification?

CHAPTER 2. STOCHASTIC PROCESSES — DEFINITIONS

Chapter 3

Discrete-Time Markov Chains

"The future depends on the past only through the present." This disarmingly simple statement captures the essence of the Markov property, formulated by Andrei Markov in 1906 while studying sequences of letters in Pushkin's poem *Eugene Onegin*. Markov wanted to prove that the law of large numbers could apply beyond independent variables, and in doing so he discovered one of the most universal structures in applied mathematics. From search engines to genetic models, from queuing theory to MCMC algorithms, Markov chains are everywhere.

This chapter develops the theory of Markov chains with countable state spaces and discrete time: classification of states, recurrence, ergodicity, and convergence to equilibrium.

3.1 The Markov Property

Definition 3.1 (Markov Chain)**.** Let E be a countable set (the *state space*). A process $(X_n)_{n\geq 0}$ with values in E, adapted to a filtration $(\mathcal{F}_n)_{n\geq 0}$, is a *Markov chain* if for every $n \geq 0$ and every $j \in E$:

$$\mathbb{P}(X_{n+1} = j \mid \mathcal{F}_n) = \mathbb{P}(X_{n+1} = j \mid X_n) \quad \text{a.s.}$$

Memorylessness

The Markov property means that the future of the process, given the present, is independent of the past. All relevant past information is summarised in the current state X_n.

Definition 3.2 (Time Homogeneity)**.** A Markov chain is *(time-)homogeneous* if the transition probabilities do not depend on n: there exists a *transition matrix* $P = (p_{ij})_{i,j\in E}$ with

$$p_{ij} = \mathbb{P}(X_{n+1} = j \mid X_n = i), \quad \forall n \geq 0.$$

The entries satisfy $p_{ij} \geq 0$ and $\sum_{j\in E} p_{ij} = 1$ for all i (stochastic matrix).

Definition 3.3 (n-Step Transition Matrix)**.** The n-step transition matrix is $P^n = (p_{ij}^{(n)})$ where

$$p_{ij}^{(n)} = \mathbb{P}(X_n = j \mid X_0 = i).$$

CHAPTER 3. DISCRETE-TIME MARKOV CHAINS

Theorem 3.4 (Chapman–Kolmogorov Equations)**.** *For all $m, n \geq 0$ and $i, j \in E$:*

$$p_{ij}^{(m+n)} = \sum_{k \in E} p_{ik}^{(m)} p_{kj}^{(n)}.$$

In matrix notation: $P^{m+n} = P^m P^n$.

3.2 Strong Markov Property

Theorem 3.5 (Strong Markov Property)**.** *Let $(X_n)_{n \geq 0}$ be a homogeneous Markov chain and let τ be a stopping time with $\mathbb{P}(\tau < \infty) > 0$. Then, conditionally on $\{\tau < \infty\}$ and $X_\tau = i$, the process $(X_{\tau+n})_{n \geq 0}$ is a Markov chain with the same transition matrix P, starting from i, independent of $\mathcal{F}_\tau$.*

3.3 State Classification

Definition 3.6 (Communication)**.** We say i *leads to* j, written $i \to j$, if there exists $n \geq 1$ with $p_{ij}^{(n)} > 0$. We say i and j *communicate*, written $i \leftrightarrow j$, if $i \to j$ and $j \to i$.

Proposition 3.7. The relation $\leftrightarrow$ is an equivalence relation on E. The equivalence classes are called *communicating classes.*

Definition 3.8 (Irreducible Chain)**.** A Markov chain is *irreducible* if E forms a single communicating class, i.e. $i \leftrightarrow j$ for all $i, j \in E$.

Definition 3.9 (Period)**.** The *period* of a state i is

$$d(i) = \gcd\{n \geq 1 : p_{ii}^{(n)} > 0\}.$$

If $d(i) = 1$, the state i is called *aperiodic.* If the chain is irreducible, all states have the same period d.

Definition 3.10 (Recurrence and Transience)**.** Let $\tau_i = \inf\{n \geq 1 : X_n = i\}$ be the first return time to i.

- State i is *recurrent* if $\mathbb{P}_i(\tau_i < \infty) = 1$.
- State i is *transient* if $\mathbb{P}_i(\tau_i < \infty) < 1$.

Theorem 3.11 (Recurrence Criterion)**.** *A state i is recurrent if and only if*

$$\sum_{n=0}^{\infty} p_{ii}^{(n)} = +\infty.$$

Equivalently, i is transient if and only if $\sum_{n=0}^{\infty} p_{ii}^{(n)} < \infty$.

Proof. Let $N_i = \sum_{n=1}^{\infty} \mathbf{1}_{\{X_n = i\}}$ be the number of visits to i after time 0. By the strong Markov property, $\mathbb{P}_i(N_i \geq k) = (\mathbb{P}_i(\tau_i < \infty))^k$ for all $k \geq 1$. Hence:

$$\mathbb{E}_i[N_i] = \sum_{k=1}^{\infty} \mathbb{P}_i(N_i \geq k) = \frac{\mathbb{P}_i(\tau_i < \infty)}{1 - \mathbb{P}_i(\tau_i < \infty)}$$

(with the convention $= +\infty$ if $\mathbb{P}_i(\tau_i < \infty) = 1$). Since $\mathbb{E}_i[N_i] = \sum_{n=1}^{\infty} p_{ii}^{(n)}$, the result follows. □

Definition 3.12 (Positive and Null Recurrence)**.** A recurrent state i is:

- *positive recurrent* if $\mathbb{E}_i[\tau_i] < \infty$,
- *null recurrent* if $\mathbb{E}_i[\tau_i] = +\infty$.

State Classification

$$\text{State } i \begin{cases} \text{transient} & \mathbb{P}_i(\tau_i < \infty) < 1 \\ \text{recurrent} & \mathbb{P}_i(\tau_i < \infty) = 1 \begin{cases} \text{null recurrent} & \mathbb{E}_i[\tau_i] = +\infty \\ \text{positive recurrent} & \mathbb{E}_i[\tau_i] < +\infty \end{cases} \end{cases}$$

In an irreducible chain, all states have the same type.

3.4 Stationary Measures

Definition 3.13 (Stationary Measure)**.** A measure $\pi = (\pi_j)_{j\in E}$ on E (with $\pi_j \geq 0$) is *stationary* (or *invariant*) for P if:

$$\pi_j = \sum_{i\in E} \pi_i p_{ij}, \quad \forall j \in E,$$

i.e. $\pi P = \pi$ in vector notation. If additionally $\sum_j \pi_j = 1$, it is called a *stationary distribution*.

Theorem 3.14 (Existence and Uniqueness)**.** *Let (X_n) be an irreducible Markov chain.*

1. *A stationary measure exists if and only if the chain is recurrent.*
2. *The stationary measure is unique (up to a multiplicative constant) and given by $\pi_j = 1/\mathbb{E}_j[\tau_j]$. It is normalisable (hence a distribution) if and only if the chain is positive recurrent.*

3.5 Ergodic Theorem

Theorem 3.15 (Ergodic Theorem for Markov Chains)**.** *Let $(X_n)_{n\geq 0}$ be an irreducible, aperiodic, positive recurrent Markov chain with stationary distribution π. Then for all $i \in E$ and every function $f : E \to \mathbb{R}$ with $\sum_j |f(j)| \pi_j < \infty$:*

1. ***Ergodic convergence**: $\displaystyle\frac{1}{n}\sum_{k=0}^{n-1} f(X_k) \xrightarrow{a.s.} \sum_{j\in E} f(j)\pi_j.$*
2. ***Convergence in distribution**: $p_{ij}^{(n)} \to \pi_j$ as $n \to \infty$.*

Role of Aperiodicity

Aperiodicity is needed for the convergence $p_{ij}^{(n)} \to \pi_j$. If the chain is periodic with period d, one only has Cesàro convergence: $\frac{1}{n}\sum_{k=0}^{n-1} p_{ij}^{(k)} \to \pi_j$.

CHAPTER 3. DISCRETE-TIME MARKOV CHAINS

3.6 Reversibility

Definition 3.16 (Reversibility)**.** A Markov chain with matrix P is *reversible* with respect to a distribution π if the *detailed balance equations* hold:

$$\pi_i p_{ij} = \pi_j p_{ji}, \quad \forall i, j \in E.$$

Proposition 3.17. If π satisfies detailed balance, then π is a stationary distribution.

Proof. Summing over i: $\sum_i \pi_i p_{ij} = \sum_i \pi_j p_{ji} = \pi_j \sum_i p_{ji} = \pi_j$. □

Example 3.18 (Random Walk on a Graph)**.** The random walk on a finite undirected graph $G = (V, E)$ with $p_{ij} = 1/\deg(i)$ if $(i, j) \in E$ is reversible with respect to $\pi_i = \deg(i)/(2\,|E|)$.

3.7 Fundamental Examples

Example 3.19 (Random Walk on $\mathbb{Z}$)**.** Let (X_n) be the simple random walk on $\mathbb{Z}$ with $p = \mathbb{P}(Y_n = 1)$ and $q = 1 - p = \mathbb{P}(Y_n = -1)$.

- If $p = q = 1/2$: the chain is null recurrent (no stationary distribution).
- If $p \neq 1/2$: the chain is transient.

Example 3.20 (Random Walk on $\mathbb{Z}^d$)**.** The simple symmetric random walk on $\mathbb{Z}^d$ is:

- recurrent for $d = 1$ and $d = 2$ (Pólya's theorem),
- transient for $d \geq 3$.

Example 3.21 (Ehrenfest Chain)**.** Two urns contain a total of N balls. At each step, a ball is chosen uniformly at random and moved to the other urn. If X_n is the number of balls in the first urn:

$$p_{i,i+1} = 1 - \frac{i}{N}, \quad p_{i,i-1} = \frac{i}{N}.$$

The stationary distribution is $\pi_i = \binom{N}{i} 2^{-N}$ (binomial). The chain is reversible.

3.8 Hitting Times and the Gambler's Ruin

Proposition 3.22 (Hitting Times)**.** For an irreducible chain, the mean hitting time $h_i = \mathbb{E}_i[\tau_j]$ (for fixed $j \neq i$) satisfies the system:

$$h_i = 1 + \sum_{k \neq j} p_{ik} h_k, \quad i \neq j, \qquad h_j = 0.$$

Example 3.23 (Gambler's Ruin)**.** A gambler starts with i euros and plays against a casino. At each round, he wins \$1 with probability p and loses \$1 with probability $q = 1 - p$. The game ends when the fortune reaches 0 or N. The ruin probability is:

$$\mathbb{P}_i(\text{ruin}) = \begin{cases} \dfrac{(q/p)^i - (q/p)^N}{1 - (q/p)^N} & \text{if } p \neq q, \\ 1 - i/N & \text{if } p = q = 1/2. \end{cases}$$

3.9 Exercises

Exercise 3.1. Let (X_n) be a Markov chain on $E = \{1, 2, 3\}$ with transition matrix

$$P = \begin{pmatrix} 0 & 1/2 & 1/2 \\ 1/3 & 1/3 & 1/3 \\ 1/2 & 0 & 1/2 \end{pmatrix}.$$

Determine whether the chain is irreducible, compute its period and stationary distribution.

Exercise 3.2. Show that in an irreducible chain, if one state is recurrent then all states are recurrent. *Hint: use* $i \leftrightarrow j$.

Exercise 3.3. Prove Pólya's theorem: the simple symmetric random walk on $\mathbb{Z}^2$ is recurrent. *Hint: use the recurrence criterion (Theorem 3.11) and an asymptotic expansion of* $p_{00}^{(2n)}$.

Exercise 3.4. Let (X_n) be the Markov chain on $\mathbb{N}$ with $p_{i,i+1} = p$ and $p_{i,0} = q = 1 - p$ for $i \geq 0$. Find the stationary distribution (if it exists) and conditions on p for positive recurrence.

Exercise 3.5. Prove that if (X_n) is a Markov chain reversible with respect to π, then the time-reversed chain $(X_{N-n})_{0 \leq n \leq N}$ has the same law as $(X_n)_{0 \leq n \leq N}$ when $X_0 \sim \pi$.

CHAPTER 3. DISCRETE-TIME MARKOV CHAINS

Chapter 4

Continuous-Time Markov Chains

The passage from discrete to continuous time is not a mere formality: it profoundly transforms the nature of Markov chains. In discrete time, transitions occur at regular instants. In continuous time, the process may jump from one state to another at any moment, and the sojourn time in each state follows an exponential distribution—the only continuous memoryless distribution. This link between the Markov property and the exponential law is one of the most elegant observations in the theory. Continuous-time Markov chains model queues (M/M/1 systems), stochastic chemical reactions, telecommunications networks, and many other systems where events occur randomly in time.

4.1 Definition and the Markov Property

Definition 4.1 (Continuous-Time Markov Chain)**.** Let E be a countable state space. A process $(X_t)_{t\geq 0}$ with values in E and càdlàg sample paths is a *continuous-time Markov chain* (CTMC) if for all $s, t \geq 0$ and every $j \in E$:

$$\mathbb{P}(X_{t+s} = j \mid \mathcal{F}_s) = \mathbb{P}(X_{t+s} = j \mid X_s) \quad \text{a.s.}$$

The chain is *homogeneous* if $\mathbb{P}(X_{t+s} = j \mid X_s = i) = p_{ij}(t)$ does not depend on s.

Definition 4.2 (Transition Semigroup)**.** The family of matrices $(P(t))_{t\geq 0}$ with $P(t) = (p_{ij}(t))_{i,j\in E}$ is called the *transition semigroup*. It satisfies:

1. $P(0) = I$ (identity matrix),
2. $P(s + t) = P(s)P(t)$ for all $s, t \geq 0$ (semigroup property),
3. $p_{ij}(t) \geq 0$ and $\sum_j p_{ij}(t) = 1$ for all i.

4.2 Infinitesimal Generator

Definition 4.3 (Infinitesimal Generator)**.** The *infinitesimal generator* (or *rate matrix*) of the CTMC is the matrix $Q = (q_{ij})_{i,j\in E}$ defined by:

$$Q = \lim_{t\to 0^+} \frac{P(t) - I}{t}, \quad \text{i.e.} \quad q_{ij} = \lim_{t\to 0^+} \frac{p_{ij}(t) - \delta_{ij}}{t}.$$

Proposition 4.4 (Properties of Q)**.** The matrix Q satisfies:

1. $q_{ij} \geq 0$ for $i \neq j$,
2. $q_{ii} \leq 0$ for all i,
3. $\sum_{j \in E} q_{ij} = 0$ for all i, i.e. $q_{ii} = -\sum_{j \neq i} q_{ij}$.

We write $q_i = -q_{ii} = \sum_{j \neq i} q_{ij}$ for the *exit rate* from state i. If $q_i > 0$, the holding time in i is exponentially distributed with parameter q_i.

Construction via Holding Times

A CTMC can be constructed as follows:

1. Starting from state i, the process remains there for an exponential time with parameter q_i.
2. At the end of this time, it jumps to $j \neq i$ with probability q_{ij}/q_i.
3. The process restarts from the new state.

The embedded discrete chain (the sequence of visited states) has transition matrix $\tilde{P}$ with $\tilde{p}_{ij} = q_{ij}/q_i$ for $j \neq i$.

4.3 Kolmogorov Equations

Theorem 4.5 (Kolmogorov Forward Equation). *Under regularity conditions (e.g.* $\sup_i q_i < \infty$*):*

$$P'(t) = P(t)Q, \quad i.e. \quad \frac{d}{dt}p_{ij}(t) = \sum_{k \in E} p_{ik}(t)q_{kj}.$$

Theorem 4.6 (Kolmogorov Backward Equation). *Under the same conditions:*

$$P'(t) = QP(t), \quad i.e. \quad \frac{d}{dt}p_{ij}(t) = \sum_{k \in E} q_{ik}p_{kj}(t).$$

Remark 4.7. If E is finite, the solution is $P(t) = e^{Qt}$, where the matrix exponential is defined by $e^{Qt} = \sum_{n=0}^{\infty} \frac{(Qt)^n}{n!}$.

Kolmogorov Equations

$$\begin{aligned} &\text{Backward:} && P'(t) = QP(t), \quad P(0) = I, \\ &\text{Forward:} && P'(t) = P(t)Q, \quad P(0) = I. \end{aligned}$$

Solution (finite state space): $P(t) = e^{Qt}$.

4.4 Stationary Distributions and Convergence

Definition 4.8 (Stationary Distribution). A probability vector $\pi = (\pi_i)_{i \in E}$ is a *stationary distribution* for the CTMC if:

$$\pi Q = 0, \quad \text{i.e.} \quad \sum_{i \in E} \pi_i q_{ij} = 0 \quad \forall j \in E.$$

Equivalently, $\pi P(t) = \pi$ for all $t \geq 0$.

Theorem 4.9 (Convergence to Equilibrium)**.** *If the CTMC is irreducible and positive recurrent, there exists a unique stationary distribution π, and for all $i, j \in E$:*

$$\lim_{t \to \infty} p_{ij}(t) = \pi_j.$$

4.5 Birth-and-Death Processes

Definition 4.10 (Birth-and-Death Process)**.** A birth-and-death process on $E = \mathbb{N}$ is a CTMC with generator:

$$q_{i,i+1} = \lambda_i \quad \text{(birth rate)}, \quad q_{i,i-1} = \mu_i \quad \text{(death rate)}, \quad i \geq 0,$$

with $\mu_0 = 0$, $\lambda_i > 0$ for $i \geq 0$, $\mu_i > 0$ for $i \geq 1$, and $q_{ij} = 0$ for $|i - j| > 1$.

Theorem 4.11 (Stationary Distribution)**.** *The birth-and-death process admits a stationary distribution if and only if*

$$S = \sum_{n=1}^{\infty} \frac{\lambda_0 \lambda_1 \cdots \lambda_{n-1}}{\mu_1 \mu_2 \cdots \mu_n} < \infty.$$

In this case, $\pi_0 = (1 + S)^{-1}$ and $\pi_n = \pi_0 \frac{\lambda_0 \cdots \lambda_{n-1}}{\mu_1 \cdots \mu_n}$.

Proof. The detailed balance equations $\pi_i q_{i,i+1} = \pi_{i+1} q_{i+1,i}$ give $\pi_i \lambda_i = \pi_{i+1} \mu_{i+1}$, whence the formula by induction. Normalisability is equivalent to $S < \infty$. □

Example 4.12 ($M/M/1$ Queue)**.** Poisson arrivals at rate λ, exponential service at rate μ: $\lambda_i = \lambda$, $\mu_i = \mu$ for all i. The stability condition is $\rho = \lambda/\mu < 1$, and the stationary distribution is geometric: $\pi_n = (1 - \rho)\rho^n$.

Example 4.13 ($M/M/c$ Queue)**.** With c servers: $\lambda_i = \lambda$ for all i and $\mu_i = \min(i, c)\mu$. The stability condition is $\rho = \lambda/(c\mu) < 1$.

Example 4.14 ($M/M/\infty$ Queue)**.** Infinitely many servers: $\lambda_i = \lambda$, $\mu_i = i\mu$. The stationary distribution is Poisson: $\pi_n = e^{-\lambda/\mu} \frac{(\lambda/\mu)^n}{n!}$.

4.6 Explosion and Regularity

Definition 4.15 (Explosion)**.** A CTMC *explodes* if the number of jumps in finite time is infinite, i.e. $T_\infty = \sum_{n=0}^{\infty} S_n < \infty$ a.s., where S_n is the n-th holding time.

Theorem 4.16 (Non-Explosion Criterion)**.** *If $\sup_{i \in E} q_i < \infty$ (bounded exit rates), the CTMC does not explode. More generally, the chain does not explode if and only if the minimal solution of the backward Kolmogorov equations is stochastic (i.e. $\sum_j p_{ij}(t) = 1$).*

Example 4.17 (Yule Process)**.** The pure birth process with $\lambda_n = n\lambda$ (each individual reproduces independently at rate λ). We have $\mathbb{E}[X_t] = X_0 e^{\lambda t}$, and the chain does not explode since $\sum_n \frac{1}{q_n} = \sum_n \frac{1}{n\lambda} = +\infty$.

CHAPTER 4. CONTINUOUS-TIME MARKOV CHAINS

4.7 Strong Markov Property

Theorem 4.18 (Strong Markov Property — Continuous Time)**.** *Let $(X_t)_{t\geq 0}$ be a regular (non-explosive) CTMC and τ an a.s. finite stopping time. Then, conditionally on $X_\tau = i$, the process $(X_{\tau+t})_{t\geq 0}$ is a CTMC with the same generator, starting from i, independent of $\mathcal{F}_\tau$.*

4.8 Reversibility in Continuous Time

Definition 4.19 (Reversibility)**.** A CTMC with generator Q is reversible with respect to a distribution π if:

$$\pi_i q_{ij} = \pi_j q_{ji}, \quad \forall i, j \in E.$$

Theorem 4.20 (Burke's Theorem)**.** *In an $M/M/1$ queue in steady state ($X_0 \sim \pi$), the departure process is a Poisson process with rate λ.*

4.9 Exercises

Exercise 4.1. Consider a CTMC on $E = \{1, 2, 3\}$ with generator

$$Q = \begin{pmatrix} -3 & 2 & 1 \\ 1 & -2 & 1 \\ 2 & 1 & -3 \end{pmatrix}.$$

Compute the stationary distribution. Check whether the chain is reversible.

Exercise 4.2. For a birth-and-death process with $\lambda_n = \lambda$ and $\mu_n = n\mu$ ($n \geq 1$), find the stationary distribution and show that the chain does not explode.

Exercise 4.3. Show that for a CTMC on a finite state space, the matrix exponential $P(t) = e^{Qt}$ satisfies both the forward and backward Kolmogorov equations.

Exercise 4.4. Consider an $M/M/1$ queue with $\lambda = 2$ and $\mu = 3$. Compute the steady-state probability that the system is empty, the average number of customers, and the average sojourn time (Little's law).

Exercise 4.5. Show that a pure birth process with $\lambda_n = n^2$ explodes in finite time. *Hint: show that $\sum_n 1/\lambda_n < \infty$.*

Exercise 4.6. For the continuous-time Ehrenfest chain (each ball is moved independently at rate 1), write down the generator Q and find the stationary distribution.

Chapter 5

Poisson Process

How many calls arrive at a call centre between 2pm and 3pm? How many radioactive decays occur in one second? How many customers enter a shop in an hour? In each of these scenarios, events occur randomly in time, independently and at a constant average rate. The process that models exactly this situation is the *Poisson process*, one of the most fundamental objects in probability theory. Discovered by Siméon Denis Poisson in 1837 as a probability distribution, it was elevated to the status of a stochastic process by the work of Palm, Feller, and Khintchine in the 1940s.

5.1 Axiomatic Definition

Definition 5.1 (Poisson Process)**.** A process $(N_t)_{t\geq 0}$ with values in $\mathbb{N}$ is a *Poisson process* with intensity $\lambda > 0$ if:

1. $N_0 = 0$,
2. (N_t) has *independent increments*: for all $0 \leq t_1 < t_2 < \cdots < t_n$, the r.v. $N_{t_2} - N_{t_1}, N_{t_3} - N_{t_2}, \ldots, N_{t_n} - N_{t_{n-1}}$ are independent,
3. (N_t) has *stationary increments*: $N_{t+s} - N_s \sim \text{Poisson}(\lambda t)$ for all $s, t \geq 0$,
4. the sample paths of (N_t) are càdlàg (right-continuous, increasing, with jumps of size 1).

Remark 5.2. Conditions (1)–(3) suffice to characterise the law of the process. Condition (4) concerns path regularity and can be obtained by construction.

Theorem 5.3 (Existence)**.** *A Poisson process with intensity λ exists for every $\lambda > 0$.*

Proof sketch. Construct (N_t) from i.i.d. exponential inter-arrival times: let $(T_n)_{n\geq 1}$ be i.i.d. $\mathcal{E}(\lambda)$. Set $S_n = T_1 + \cdots + T_n$ (the n-th arrival time) and $N_t = \max\{n \geq 0 : S_n \leq t\}$. One verifies that (N_t) satisfies the axioms of Definition 5.1. □

5.2 Fundamental Properties

Proposition 5.4 (Moments)**.** For all $t \geq 0$:

$$\mathbb{E}[N_t] = \lambda t, \quad \text{Var}(N_t) = \lambda t.$$

Proposition 5.5 (Inter-Arrival Times)**.** The times between consecutive jumps $T_n = S_n - S_{n-1}$ (with $S_0 = 0$) are i.i.d. $\mathcal{E}(\lambda)$.

Proposition 5.6 (Arrival Times)**.** The n-th arrival time $S_n = T_1 + \cdots + T_n$ has a Gamma distribution: $S_n \sim \text{Gamma}(n, \lambda)$, with density:

$$f_{S_n}(t) = \frac{\lambda^n t^{n-1}}{(n-1)!} e^{-\lambda t}, \quad t > 0.$$

Poisson Process Properties

$$\begin{aligned} N_{t+s} - N_s &\sim \text{Poisson}(\lambda t) \\ \mathbb{E}[N_t] &= \lambda t, \quad \text{Var}(N_t) = \lambda t \\ T_n &\sim \mathcal{E}(\lambda) \text{ (i.i.d.)} \\ S_n &\sim \text{Gamma}(n, \lambda) \\ \mathbb{P}(N_t = k) &= \frac{(\lambda t)^k}{k!} e^{-\lambda t} \end{aligned}$$

5.3 Construction as a CTMC

Remark 5.7. The Poisson process is a special case of a pure birth process: it is a CTMC on $\mathbb{N}$ with $q_{n,n+1} = \lambda$ for all n and $q_{ij} = 0$ for $j \neq i + 1$. The generator is:

$$(Qf)(n) = \lambda[f(n+1) - f(n)].$$

5.4 Markov Property and Memorylessness

Theorem 5.8 (Strong Markov Property)**.** *The Poisson process satisfies the strong Markov property: for every a.s. finite stopping time τ, the process $(N_{\tau+t} - N_\tau)_{t \geq 0}$ is a Poisson process with intensity λ, independent of $\mathcal{F}_\tau$.*

Proposition 5.9 (Memorylessness of the Exponential)**.** The exponential distribution is the unique continuous distribution with the memoryless property: if $T \sim \mathcal{E}(\lambda)$, then

$$\mathbb{P}(T > t + s \mid T > t) = \mathbb{P}(T > s) = e^{-\lambda s}, \quad \forall s, t \geq 0.$$

5.5 Operations on the Poisson Process

5.5.1 Superposition

Theorem 5.10 (Superposition)**.** *Let $(N_t^{(1)})$ and $(N_t^{(2)})$ be independent Poisson processes with intensities λ_1 and λ_2. Then $(N_t^{(1)} + N_t^{(2)})$ is a Poisson process with intensity $\lambda_1 + \lambda_2$.*

5.5.2 Thinning

Theorem 5.11 (Thinning). *Let (N_t) be a Poisson process with intensity λ. Each jump is independently classified as type 1 with probability p and type 2 with probability $1-p$. Then the counting processes $N_t^{(1)}$ and $N_t^{(2)}$ of types 1 and 2 are independent Poisson processes with intensities λp and $\lambda(1-p)$.*

Proof. For any t, conditionally on $N_t = n$, the n jumps are independently classified. Hence $N_t^{(1)} \mid (N_t = n) \sim \mathrm{Bin}(n, p)$. We compute:

$$\mathbb{P}(N_t^{(1)} = k) = \sum_{n=k}^{\infty} \binom{n}{k} p^k (1-p)^{n-k} \frac{(\lambda t)^n}{n!} e^{-\lambda t} = \frac{(\lambda p t)^k}{k!} e^{-\lambda p t}.$$

Independence of $N_t^{(1)}$ and $N_t^{(2)}$ follows similarly. □

5.5.3 Order Statistics

Theorem 5.12 (Order Statistics). *Conditionally on $N_t = n$, the n jump times $(S_1, \dots, S_n)$ have the same distribution as the order statistics of n i.i.d. uniform random variables on $[0, t]$.*

5.6 Compound Poisson Process

Definition 5.13 (Compound Poisson Process). Let $(N_t)_{t \geq 0}$ be a Poisson process with intensity λ and $(Y_k)_{k \geq 1}$ an i.i.d. sequence independent of (N_t). The *compound Poisson process* is

$$Z_t = \sum_{k=1}^{N_t} Y_k, \quad t \geq 0.$$

Proposition 5.14 (Moments of the Compound Process).

$$\mathbb{E}[Z_t] = \lambda t \, \mathbb{E}[Y_1], \quad \mathrm{Var}(Z_t) = \lambda t \, \mathbb{E}[Y_1^2].$$

Example 5.15. An insurance claims model: claims arrive according to a Poisson process with rate λ, each claim having a random amount Y_k (i.i.d.). The total claims amount up to time t is $Z_t = \sum_{k=1}^{N_t} Y_k$.

5.7 Non-Homogeneous Poisson Process

Definition 5.16 (Non-Homogeneous Poisson Process). A process $(N_t)_{t \geq 0}$ is a *non-homogeneous Poisson process* with *intensity function* $\lambda(t) \geq 0$ if:

1. $N_0 = 0$,
2. (N_t) has independent increments,
3. $N_{t+s} - N_s \sim \mathrm{Poisson}\left(\int_s^{t+s} \lambda(u) \, du\right)$.

Proposition 5.17 (Time Change). Let $\Lambda(t) = \int_0^t \lambda(u) \, du$ be the *cumulative intensity function*. If $(M_s)_{s \geq 0}$ is a standard Poisson process (intensity 1), then $N_t = M_{\Lambda(t)}$ is a non-homogeneous Poisson process with intensity function $\lambda(t)$.

 CHAPTER 5. POISSON PROCESS

5.8 The Poisson Process as a Martingale

Proposition 5.18 (Compensated Martingale)**.** If (N_t) is a Poisson process with intensity λ, then:

1. $M_t = N_t - \lambda t$ is a martingale (the *compensated Poisson process*).
2. $M_t^2 - \lambda t$ is a martingale.
3. $e^{\theta N_t - \lambda t(e^\theta - 1)}$ is a martingale for all $\theta \in \mathbb{R}$.

Proof. For item (1):

$$\begin{aligned}\mathbb{E}[M_{t+s} - M_s \mid \mathcal{F}_s] &= \mathbb{E}[N_{t+s} - N_s - \lambda t \mid \mathcal{F}_s] \\ &= \mathbb{E}[N_{t+s} - N_s] - \lambda t = \lambda t - \lambda t = 0.\end{aligned}$$

□

5.9 Exercises

Exercise 5.1. Show that if (N_t) is a Poisson process with intensity λ, then for fixed $s < t$, $\mathbb{E}[N_s \mid N_t] = N_t \cdot s/t$.

Exercise 5.2. Telephone calls arrive according to a Poisson process with rate $\lambda = 5$ per hour. Compute the probability of having at least 3 calls in 30 minutes.

Exercise 5.3. Let (N_t) be a Poisson process with intensity λ and $T > 0$. Show that $\mathbb{E}[S_1 \mid N_T = n] = T/(n+1)$ where S_1 is the first jump time.

Exercise 5.4. Let $Z_t = \sum_{k=1}^{N_t} Y_k$ be a compound Poisson process with $Y_k \sim \mathcal{E}(\mu)$ and N_t Poisson with intensity λ. Compute $\mathbb{E}[e^{-\alpha Z_t}]$ for $\alpha > 0$.

Exercise 5.5. Prove the superposition theorem (Theorem 5.10) using probability generating functions.

Exercise 5.6. Buses arrive at a stop according to a Poisson process with intensity λ. A passenger arrives at a deterministic time $t_0 > 0$. Show that the passenger's waiting time is exponential $\mathcal{E}(\lambda)$, independent of the past.

Chapter 6

Martingales — Discrete Time

The term "martingale" comes from the world of gambling: it is the strategy of doubling one's bet after each loss, hoping to recover everything at once. Joseph Leo Doob borrowed this word to designate a far deeper mathematical concept: a process whose best prediction of the future, given the past, is the present value. Martingales are the central tool of modern probability theory: they appear in convergence theorems, stochastic integration, mathematical finance (discounted prices are martingales under the risk-neutral measure), and optimal stopping theory.

6.1 Definitions

Definition 6.1 (Discrete-Time Martingale)**.** Let $(\Omega, \mathcal{F}, (\mathcal{F}_n)_{n\geq 0}, \mathbb{P})$ be a filtered probability space. An adapted process $(M_n)_{n\geq 0}$ is a *martingale* if:

1. $\mathbb{E}[|M_n|] < \infty$ for all $n \geq 0$,
2. $\mathbb{E}[M_{n+1} \mid \mathcal{F}_n] = M_n$ a.s. for all $n \geq 0$.

Analogously:

- *supermartingale*: $\mathbb{E}[M_{n+1} \mid \mathcal{F}_n] \leq M_n$ a.s.
- *submartingale*: $\mathbb{E}[M_{n+1} \mid \mathcal{F}_n] \geq M_n$ a.s.

Remark 6.2. By the tower property, the martingale condition is equivalent to: for all $m \leq n$, $\mathbb{E}[M_n \mid \mathcal{F}_m] = M_m$ a.s. In particular, $\mathbb{E}[M_n] = \mathbb{E}[M_0]$ for all n.

Example 6.3.

1. **Random walk**: if $(Y_k)_{k\geq 1}$ are i.i.d. centred and $S_n = \sum_{k=1}^n Y_k$, then (S_n) is a martingale with respect to $\mathcal{F}_n = \sigma(Y_1, \ldots, Y_n)$.
2. **Martingale transform**: if (M_n) is a martingale and (H_n) is predictable and bounded, then $(H \cdot M)_n = \sum_{k=1}^n H_k(M_k - M_{k-1})$ is a martingale.
3. **Exponential martingale**: if (Y_k) are i.i.d. and $\psi(\theta) = \log \mathbb{E}[e^{\theta Y_1}] < \infty$, then $\exp(\theta S_n - n\psi(\theta))$ is a martingale.
4. **Lévy martingale**: if $X \in L^1$ and $\mathcal{F}_n$ is a filtration, then $M_n = \mathbb{E}[X \mid \mathcal{F}_n]$ is a martingale.

CHAPTER 6. MARTINGALES — DISCRETE TIME

6.2 Martingale Transform

Definition 6.4 (Martingale Transform)**.** Let $(M_n)_{n\geq 0}$ be a martingale and $(H_n)_{n\geq 1}$ a predictable process. The *martingale transform* is

$$(H\cdot M)_n = \sum_{k=1}^{n} H_k(M_k - M_{k-1}), \quad (H\cdot M)_0 = 0.$$

Theorem 6.5 (Stability)**.** *If (M_n) is a martingale (resp. supermartingale) and (H_n) is predictable with $0 \leq H_n \leq C$ for a constant C, then $(H\cdot M)_n$ is a martingale (resp. supermartingale).*

Gambling Strategy

The martingale transform models the cumulative gains of a gambler who bets H_n at step n in a fair game (M_n). Predictability means the bet is decided before observing the outcome. The stability theorem says that no bounded betting strategy can turn a fair game into a favourable one.

6.3 Fundamental Inequalities

Theorem 6.6 (Doob's Maximal Inequality)**.** *Let $(M_n)_{0\leq n\leq N}$ be a submartingale. Then for all $\lambda > 0$:*

$$\lambda\,\mathbb{P}\Big(\max_{0\leq n\leq N} M_n \geq \lambda\Big) \leq \mathbb{E}[M_N^+].$$

Theorem 6.7 (Doob's L^p Inequality)**.** *Let $(M_n)_{0\leq n\leq N}$ be a martingale and $p > 1$. Then*

$$\mathbb{E}\Big[\max_{0\leq n\leq N} |M_n|^p\Big] \leq \Big(\frac{p}{p-1}\Big)^p \mathbb{E}[|M_N|^p].$$

Theorem 6.8 (Doob's L^2 Inequality)**.** *For a square-integrable martingale (M_n):*

$$\mathbb{E}\Big[\max_{0\leq n\leq N} M_n^2\Big] \leq 4\mathbb{E}[M_N^2].$$

6.4 Doob Decomposition

Theorem 6.9 (Doob Decomposition)**.** *Every adapted integrable process $(X_n)_{n\geq 0}$ admits a unique decomposition:*

$$X_n = M_n + A_n,$$

where (M_n) is a martingale and (A_n) is a predictable process with $A_0 = 0$. The process A is given by

$$A_n = \sum_{k=1}^{n}\big(\mathbb{E}[X_k \mid \mathcal{F}_{k-1}] - X_{k-1}\big).$$

If (X_n) is a submartingale, then (A_n) is increasing.

6.5 Convergence Theorems

Theorem 6.10 (Martingale Convergence — Doob's Theorem). *Let $(M_n)_{n\geq 0}$ be a submartingale with* $\sup_n \mathbb{E}[M_n^+] < \infty$. *Then there exists a r.v. $M_\infty \in L^1$ such that $M_n \to M_\infty$ a.s.*

Proof sketch. We use Doob's *upcrossing lemma*: if $U_N([a,b])$ denotes the number of upcrossings of $(M_n)_{0\leq n\leq N}$ from a to b (with $a<b$), then

$$(b-a)\mathbb{E}[U_N([a,b])] \leq \mathbb{E}[(M_N-a)^+].$$

Under $\sup_n \mathbb{E}[M_n^+] < \infty$, we get $\mathbb{E}[U_\infty([a,b])] < \infty$ for all rational $a<b$, which implies a.s. convergence. □

A.s. vs. L^1 Convergence

Doob's theorem gives a.s. convergence, but *not* L^1 convergence in general. For $M_n \to M_\infty$ in L^1, an additional condition (uniform integrability) is needed.

Definition 6.11 (Uniform Integrability). A family $(X_i)_{i\in I}$ of r.v. is *uniformly integrable* (UI) if

$$\lim_{a\to\infty} \sup_{i\in I} \mathbb{E}[|X_i| \mathbf{1}_{\{|X_i|>a\}}] = 0.$$

Theorem 6.12 (L^1 Convergence). *Let (M_n) be a martingale. The following are equivalent:*

1. *(M_n) is uniformly integrable,*
2. *M_n converges a.s. and in L^1 to M_∞,*
3. *there exists $X \in L^1$ such that $M_n = \mathbb{E}[X \mid \mathcal{F}_n]$ for all n.*

In this case, $M_\infty = X$ a.s. and the martingale is called closed.

Corollary 6.13 (L^p Convergence for $p>1$). *If (M_n) is a martingale bounded in L^p for $p>1$ (i.e.* $\sup_n \mathbb{E}[|M_n|^p] < \infty$*), then (M_n) converges a.s. and in L^p.*

6.6 Optional Stopping Theorem

Theorem 6.14 (Optional Stopping — Bounded Version). *Let $(M_n)_{n\geq 0}$ be a martingale and τ a* bounded *stopping time ($\tau \leq N$ a.s. for some $N<\infty$). Then*

$$\mathbb{E}[M_\tau] = \mathbb{E}[M_0].$$

Proof. Write $M_\tau = M_0 + \sum_{k=1}^N H_k(M_k - M_{k-1})$ where $H_k = \mathbf{1}_{\{\tau\geq k\}}$ is predictable (since $\{\tau \geq k\} = \{\tau \leq k-1\}^c \in \mathcal{F}_{k-1}$). Since the martingale transform is a martingale, $\mathbb{E}[M_\tau] = \mathbb{E}[M_0]$. □

Theorem 6.15 (Optional Stopping — General Version). *Let $(M_n)_{n\geq 0}$ be a uniformly integrable martingale and σ,τ stopping times with $\sigma\leq\tau$ a.s. Then*

$$\mathbb{E}[M_\tau \mid \mathcal{F}_\sigma] = M_\sigma \quad a.s.$$

In particular, $\mathbb{E}[M_\tau] = \mathbb{E}[M_0]$.

CHAPTER 6. MARTINGALES — DISCRETE TIME

Conditions for Optional Stopping

The optional stopping theorem does *not* apply to an arbitrary stopping time without additional conditions. Classic counterexample: the symmetric random walk (S_n) with $\tau = \inf\{n : S_n = 1\}$. We have $\mathbb{E}[S_0] = 0$ but $S_\tau = 1$, since τ is unbounded and (S_n) is not UI.

6.7 Applications

6.7.1 Gambler's Ruin Revisited

Example 6.16. Random walk (S_n) with $S_0 = i$ and absorbing barriers at 0 and N. Set $\tau = \inf\{n : S_n \in \{0, N\}\}$.

If $p \neq 1/2$, the exponential martingale $M_n = (q/p)^{S_n}$ is bounded between $(q/p)^0 = 1$ and $(q/p)^N$, so τ is a bounded stopping time for $(M_{n\wedge\tau})$. By optional stopping:

$$(q/p)^i = \mathbb{E}[(q/p)^{S_\tau}] = (q/p)^0 \mathbb{P}_i(\text{ruin}) + (q/p)^N (1 - \mathbb{P}_i(\text{ruin})),$$

whence $\mathbb{P}_i(\text{ruin}) = \frac{(q/p)^i - (q/p)^N}{1-(q/p)^N}$.

6.7.2 Law of the Iterated Logarithm — First Step

Proposition 6.17. If (M_n) is a square-integrable martingale with $\mathbb{E}[(M_n - M_{n-1})^2 \mid \mathcal{F}_{n-1}] = \sigma^2$ a.s., then $\langle M \rangle_n = n\sigma^2$ and $M_n^2 - n\sigma^2$ is a martingale.

6.8 Predictable Quadratic Variation

Definition 6.18 (Predictable Quadratic Variation)**.** Let (M_n) be a square-integrable martingale. The *predictable quadratic variation* $(\langle M \rangle_n)$ is the unique predictable increasing process such that $(M_n^2 - \langle M \rangle_n)$ is a martingale. It is given by:

$$\langle M \rangle_n = \sum_{k=1}^{n} \mathbb{E}[(M_k - M_{k-1})^2 \mid \mathcal{F}_{k-1}].$$

Theorem 6.19 (Convergence via Quadratic Variation)**.** *Let (M_n) be a square-integrable martingale. If $\langle M \rangle_\infty = \lim_n \langle M \rangle_n < \infty$ a.s., then (M_n) converges a.s. and in L^2.*

6.9 Exercises

Exercise 6.1. Show that if (M_n) is a martingale and φ is convex, then $(\varphi(M_n))$ is a submartingale (assuming integrability).

Exercise 6.2. Let $(X_n)_{n\geq 0}$ be a simple symmetric random walk on $\mathbb{Z}$ starting from 0. Show that $X_n^2 - n$ is a martingale.

Exercise 6.3. Let $\tau = \inf\{n \geq 0 : |S_n| = a\}$ where (S_n) is the symmetric random walk. Use the optional stopping theorem to compute $\mathbb{E}[\tau]$.

Exercise 6.4. Prove Doob's maximal inequality (Theorem 6.6). *Hint: use* $M_N^+ \geq M_\sigma^+$ *where* $\sigma = \inf\{n : M_n \geq \lambda\} \wedge N$*.*

Exercise 6.5. Let (M_n) be a positive martingale with $M_0 = 1$. Show that $\mathbb{P}(\sup_n M_n \geq a) \leq 1/a$ for all $a > 0$.

Exercise 6.6. Let (M_n) be a martingale bounded in L^2. Show that it is uniformly integrable.

CHAPTER 6. MARTINGALES — DISCRETE TIME

Chapter 7

Martingales — Continuous Time

The passage from discrete-time to continuous-time martingales is one of the most delicate transitions in stochastic process theory. In discrete time, one conditions on a finite past; in continuous time, the filtration $(\mathcal{F}_t)_{t\geq 0}$ forms a continuum of information, and subtle phenomena emerge: path regularity, uncountable stopping times, maximal inequalities. Joseph Doob, in the 1950s, showed that every continuous-time martingale admits a càdlàg modification (right-continuous with left limits), a technical but essential result. Doob's inequalities, the optional stopping theorem, and the Doob–Meyer decomposition form the foundation on which all of stochastic calculus rests — and, by extension, modern mathematical finance.

7.1 Definitions

Definition 7.1 (Continuous-Time Martingale)**.** Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ be a filtered probability space satisfying the usual conditions. An adapted process $(M_t)_{t\geq 0}$ is a *martingale* (in continuous time) if:

1. $\mathbb{E}[|M_t|] < \infty$ for all $t \geq 0$,
2. $\mathbb{E}[M_t \mid \mathcal{F}_s] = M_s$ a.s. for all $0 \leq s \leq t$.

Definition 7.2 (Local Martingale)**.** An adapted continuous process $(M_t)_{t\geq 0}$ is a *local martingale* if there exists an increasing sequence of stopping times $(\tau_n)_{n\geq 1}$ with $\tau_n \uparrow \infty$ a.s. such that for each n, the stopped process $(M_{t\wedge\tau_n})_{t\geq 0}$ is a martingale.

The sequence (τ_n) is called a *localising sequence*.

Local Martingale $\neq$ Martingale

Every martingale is a local martingale, but the converse is false. A positive local martingale is a supermartingale (by Fatou's lemma), but not necessarily a martingale.

7.2 Path Regularity

Theorem 7.3 (Càdlàg Modification)**.** *Let $(M_t)_{t\geq 0}$ be a real-valued martingale (or supermartingale). If $t \mapsto \mathbb{E}[M_t]$ is right-continuous, then (M_t) admits a càdlàg modification.*

CHAPTER 7. MARTINGALES — CONTINUOUS TIME

Remark 7.4. Under the usual conditions on the filtration, the càdlàg modification is still a martingale. Henceforth, we assume continuous-time martingales are càdlàg.

Theorem 7.5 (Continuity of L^2 Martingales). *If (M_t) is a square-integrable martingale with continuous sample paths, then the quadratic variation process $\langle M\rangle_t$ is continuous and $(M_t^2 - \langle M\rangle_t)$ is a martingale.*

7.3 Doob's Inequalities in Continuous Time

Theorem 7.6 (Doob's Maximal Inequality). *Let $(M_t)_{0\le t\le T}$ be a càdlàg martingale. For all $\lambda > 0$:*

$$\mathbb{P}\Big(\sup_{0\le t\le T} |M_t| \ge \lambda\Big) \le \frac{\mathbb{E}[|M_T|]}{\lambda}.$$

Theorem 7.7 (Doob's L^p Inequality — Continuous Time). *For $p > 1$ and a càdlàg martingale (M_t):*

$$\mathbb{E}\Big[\sup_{0\le t\le T} |M_t|^p\Big] \le \Big(\frac{p}{p-1}\Big)^p \mathbb{E}[|M_T|^p].$$

Theorem 7.8 (Burkholder–Davis–Gundy Inequality). *There exist universal constants $c_p, C_p > 0$ such that for every continuous local martingale (M_t) with $M_0 = 0$ and for all $p > 0$:*

$$c_p\,\mathbb{E}[\langle M\rangle_T^{p/2}] \le \mathbb{E}\Big[\sup_{0\le t\le T} |M_t|^p\Big] \le C_p\,\mathbb{E}[\langle M\rangle_T^{p/2}].$$

For $p = 2$, one may take $c_2 = 1$ and $C_2 = 4$ (Doob's inequality).

7.4 Quadratic Variation

Definition 7.9 (Quadratic Variation). Let $(M_t)_{t\ge 0}$ be a continuous local martingale. The *quadratic variation* (or *bracket*) of M is the unique continuous increasing process $(\langle M\rangle_t)_{t\ge 0}$ with $\langle M\rangle_0 = 0$ such that $(M_t^2 - \langle M\rangle_t)$ is a local martingale.

It has the representation:

$$\langle M\rangle_t = \lim_{|\Pi|\to 0} \sum_{k=0}^{n-1} (M_{t_{k+1}} - M_{t_k})^2 \quad \text{(limit in probability)},$$

where $\Pi = \{0 = t_0 < t_1 < \cdots < t_n = t\}$ is a partition of $[0, t]$.

Definition 7.10 (Cross Variation). For two continuous local martingales M and N, the *cross variation* (or *cross bracket*) is

$$\langle M, N\rangle_t = \frac{1}{4}\big(\langle M+N\rangle_t - \langle M-N\rangle_t\big),$$

and is the unique continuous finite-variation process such that $(M_tN_t - \langle M, N\rangle_t)$ is a local martingale.

Fundamental Quadratic Variations

For Brownian motion (B_t):

$$\langle B\rangle_t = t$$
$$\langle B^{(i)}, B^{(j)}\rangle_t = \delta_{ij} t \quad \text{(multidimensional Brownian motion)}$$

7.5 Doob–Meyer Decomposition

Theorem 7.11 (Doob–Meyer Decomposition)**.** *Let $(X_t)_{t\geq 0}$ be a càdlàg supermartingale of class (D). Then there exists a unique decomposition:*

$$X_t = M_t - A_t,$$

where (M_t) is a càdlàg martingale and (A_t) is an increasing predictable process with $A_0 = 0$.

Definition 7.12 (Class (D))**.** A supermartingale (X_t) is of *class (D)* if the family $\{X_\tau : \tau \text{ finite stopping time}\}$ is uniformly integrable.

Remark 7.13. The Doob–Meyer decomposition is the continuous-time analogue of the discrete Doob decomposition. It is fundamental for constructing the stochastic integral.

7.6 Optional Stopping in Continuous Time

Theorem 7.14 (Optional Stopping — Continuous Time)**.** *Let $(M_t)_{t\geq 0}$ be a continuous uniformly integrable martingale and τ a stopping time (possibly infinite). Then*

$$\mathbb{E}[M_\tau] = \mathbb{E}[M_0].$$

More generally, if $\sigma \leq \tau$ are stopping times:

$$\mathbb{E}[M_\tau \mid \mathcal{F}_\sigma] = M_\sigma \quad a.s.$$

Theorem 7.15 (Stopping for Positive Supermartingales)**.** *If $(M_t)_{t\geq 0}$ is a positive continuous supermartingale and τ is a stopping time, then*

$$\mathbb{E}[M_\tau] \leq \mathbb{E}[M_0].$$

7.7 Martingale Representation

Theorem 7.16 (Brownian Martingale Representation)**.** *Let $(B_t)_{t\geq 0}$ be a Brownian motion and $(\mathcal{F}_t)$ its completed natural filtration. Every continuous local martingale (M_t) with respect to $(\mathcal{F}_t)$ admits the representation:*

$$M_t = M_0 + \int_0^t H_s \, dB_s,$$

where (H_t) is a predictable process with $\int_0^t H_s^2 \, ds < \infty$ a.s. for all t.

CHAPTER 7. MARTINGALES — CONTINUOUS TIME

Completeness of the Brownian Filtration

This theorem says that the Brownian filtration is "complete" in the sense that every martingale can be written as a stochastic integral with respect to Brownian motion. This is a central result for applications in finance (market completeness theorem).

7.8 Girsanov's Theorem

Theorem 7.17 (Girsanov)**.** *Let $(B_t)_{t\geq 0}$ be a Brownian motion under $\mathbb{P}$ and $(\theta_t)_{t\geq 0}$ an adapted process with $\int_0^T \theta_s^2\,ds < \infty$ a.s. Define the Doléans–Dade exponential:*

$$Z_t = \exp\Big(-\int_0^t \theta_s\,dB_s - \frac{1}{2}\int_0^t \theta_s^2\,ds\Big).$$

If (Z_t) is a martingale (Novikov's condition: $\mathbb{E}[\exp(\frac{1}{2}\int_0^T \theta_s^2\,ds)] < \infty$), then under the measure $\mathbb{Q}$ defined by $d\mathbb{Q}/d\mathbb{P}\big|_{\mathcal{F}_T} = Z_T$, the process

$$\tilde{B}_t = B_t + \int_0^t \theta_s\,ds$$

is a Brownian motion.

7.9 Convergence in Continuous Time

Theorem 7.18 (Continuous Martingale Convergence)**.** *Let $(M_t)_{t\geq 0}$ be a right-continuous supermartingale with $\sup_{t\geq 0}\mathbb{E}[M_t^+] < \infty$. Then $M_\infty = \lim_{t\to\infty} M_t$ exists a.s. and $M_\infty \in L^1$.*

7.10 Exercises

Exercise 7.1. Let $(B_t)_{t\geq 0}$ be a Brownian motion. Verify that $M_t = \exp(\theta B_t - \theta^2 t/2)$ is a martingale for all $\theta \in \mathbb{R}$.

Exercise 7.2. Prove that if (M_t) is a continuous positive local martingale, then it is a supermartingale. *Hint: use Fatou's lemma.*

Exercise 7.3. Let (M_t) be a continuous square-integrable martingale. Show that $\mathbb{E}[M_t^2] = \mathbb{E}[M_0^2] + \mathbb{E}[\langle M\rangle_t]$.

Exercise 7.4. Compute $\langle M\rangle_t$ for $M_t = B_t^2 - t$ where (B_t) is a Brownian motion.

Exercise 7.5. Let (B_t) be a Brownian motion and $\tau_a = \inf\{t \geq 0 : B_t = a\}$. Use optional stopping with the exponential martingale to compute $\mathbb{E}[e^{-\alpha\tau_a}]$ for $\alpha > 0$ and $a > 0$.

Exercise 7.6. State and prove Girsanov's theorem in the case where $\theta_t = \theta$ is constant. Verify Novikov's condition.

Exercise 7.7. Show that the Burkholder–Davis–Gundy inequality (Theorem 7.8) implies Doob's L^2 inequality for continuous martingales.

Chapter 8

Brownian Motion

In 1827, the botanist Robert Brown observed pollen grains suspended in water through a microscope. They trembled, jiggled, changed direction ceaselessly — and nobody understood why. It would take until 1905 and a clerk at the Bern patent office, a certain Albert Einstein, for the mystery to begin to clear: these grains were being bombarded by billions of water molecules, and their erratic trajectory reflected the thermal agitation of the fluid. Meanwhile, the mathematician Louis Bachelier had already, in his 1900 thesis, modelled stock market fluctuations using a process with remarkably similar properties. Norbert Wiener, in the 1920s, gave the first rigorous construction. Today, Brownian motion is the cornerstone of stochastic calculus, mathematical finance, and statistical physics — the continuous-time process from which all others are built.

8.1 Definition

Definition 8.1 (Standard Brownian Motion)**.** A real-valued stochastic process $(B_t)_{t\geq 0}$ is a *standard Brownian motion* (or *Wiener process*) if:

1. $B_0 = 0$ a.s.,
2. (B_t) has *independent increments*: for all $0 \leq t_1 < t_2 < \cdots < t_n$, the r.v. $B_{t_2} - B_{t_1}, B_{t_3} - B_{t_2}, \ldots, B_{t_n} - B_{t_{n-1}}$ are independent,
3. (B_t) has *stationary Gaussian increments*: $B_{t+s} - B_s \sim \mathcal{N}(0, t)$ for all $s, t \geq 0$,
4. the sample paths $t \mapsto B_t(\omega)$ are a.s. *continuous*.

Remark 8.2. Brownian motion is a centred Gaussian process with covariance function $\mathrm{Cov}(B_s, B_t) = \min(s, t)$.

Definition 8.3 (Multidimensional Brownian Motion)**.** A d-dimensional Brownian motion is $B_t = (B_t^{(1)}, \ldots, B_t^{(d)})$ where $B^{(1)}, \ldots, B^{(d)}$ are independent standard Brownian motions.

8.2 Existence: Lévy–Ciesielski Construction

Theorem 8.4 (Existence of Brownian Motion)**.** *There exists a stochastic process satisfying the four properties of Definition 8.1.*

 CHAPTER 8. BROWNIAN MOTION

Lévy–Ciesielski construction (sketch). We use the Schauder basis of $C([0,1])$. Let $(Z_n)_{n\geq 0}$ be i.i.d. $\mathcal{N}(0,1)$. Define:

$$B_t = \sum_{n=0}^{\infty} Z_n \, s_n(t),$$

where the s_n are the Schauder functions (dyadic hat functions). The series converges uniformly a.s. in $C([0,1])$ and the limit is a Brownian motion on $[0,1]$. Extension to $\mathbb{R}_+$ is by concatenation. □

8.3 Fundamental Properties

Proposition 8.5 (Martingale Properties)**.** Brownian motion satisfies:

1. (B_t) is a martingale with respect to its natural filtration.
2. $(B_t^2 - t)$ is a martingale.
3. $\exp(\theta B_t - \theta^2 t/2)$ is a martingale for all $\theta \in \mathbb{R}$.

Proposition 8.6 (Invariance Properties)**.** The following processes are also standard Brownian motions:

1. **Symmetry**: $(-B_t)_{t\geq 0}$.
2. **Scaling**: $(c^{-1/2} B_{ct})_{t\geq 0}$ for any $c > 0$.
3. **Time translation**: $(B_{t+s} - B_s)_{t\geq 0}$ for any $s \geq 0$.
4. **Time inversion**: $(tB_{1/t})_{t>0}$ (with $B_{1/t}\big|_{t=0} = 0$).

Proposition 8.7 (Strong Markov Property)**.** For every a.s. finite stopping time τ, the process $(B_{\tau+t} - B_\tau)_{t\geq 0}$ is a Brownian motion independent of $\mathcal{F}_\tau$.

8.4 Path Regularity

Theorem 8.8 (Hölder Continuity)**.** *The sample paths of Brownian motion are a.s. Hölder continuous with exponent γ for every $\gamma < 1/2$, but not for $\gamma = 1/2$.*

Theorem 8.9 (Nowhere Differentiability)**.** *The sample paths of Brownian motion are a.s.* nowhere differentiable*. More precisely, for almost every ω, there is no $t \geq 0$ such that $\lim_{h\to 0}(B_{t+h}(\omega) - B_t(\omega))/h$ exists (even as an infinite limit).*

Theorem 8.10 (Quadratic Variation)**.** *For any sequence of partitions $\Pi_n = \{0 = t_0^n < t_1^n < \cdots < t_{k_n}^n = t\}$ of $[0,t]$ with $|\Pi_n| \to 0$:*

$$\sum_{i=0}^{k_n-1} (B_{t_{i+1}^n} - B_{t_i^n})^2 \xrightarrow{L^2} t.$$

That is, $\langle B \rangle_t = t$.

Proof. Set $Q_n = \sum_{i=0}^{k_n-1}(B_{t_{i+1}} - B_{t_i})^2$. Then $\mathbb{E}[Q_n] = \sum_i (t_{i+1} - t_i) = t$. Moreover:

$$\begin{aligned}\mathrm{Var}(Q_n) &= \sum_i \mathrm{Var}\big((B_{t_{i+1}} - B_{t_i})^2\big) = \sum_i 2(t_{i+1} - t_i)^2 \\ &\le 2\,|\Pi_n| \sum_i (t_{i+1} - t_i) = 2\,|\Pi_n| \cdot t \to 0.\end{aligned}$$

Hence $Q_n \to t$ in L^2. □

Theorem 8.11 (Infinite Total Variation). *The sample paths of Brownian motion a.s. have* infinite total variation *on every interval* $[0, t]$ *with* $t > 0$*:*

$$\sup_{\Pi} \sum_i \left|B_{t_{i+1}} - B_{t_i}\right| = +\infty \quad a.s.$$

Finite Quadratic Variation, Infinite Total Variation

Brownian motion has finite quadratic variation ($= t$) but infinite total variation. This property prevents defining the stochastic integral $\int_0^t f\,dB$ in the Stieltjes sense and necessitates Itô's special construction.

8.5 Law of the Iterated Logarithm

Theorem 8.12 (Law of the Iterated Logarithm).

$$\limsup_{t\to\infty} \frac{B_t}{\sqrt{2t\log\log t}} = 1 \quad a.s.$$

$$\liminf_{t\to\infty} \frac{B_t}{\sqrt{2t\log\log t}} = -1 \quad a.s.$$

There is an analogous result as $t \to 0$*:*

$$\limsup_{t\to 0^+} \frac{B_t}{\sqrt{2t\log\log(1/t)}} = 1 \quad a.s.$$

8.6 Reflection Principle and Law of the Maximum

Theorem 8.13 (Reflection Principle). *Let* $\tau_a = \inf\{t \ge 0 : B_t = a\}$ *for* $a > 0$*. For all* $b \le a$*:*

$$\mathbb{P}(\tau_a \le t,\ B_t \le b) = \mathbb{P}(B_t \ge 2a - b).$$

Corollary 8.14 (Law of the Maximum). *Let* $M_t = \max_{0\le s\le t} B_s$*. Then:*

$$\mathbb{P}(M_t \ge a) = 2\mathbb{P}(B_t \ge a) = 2\big(1 - \Phi(a/\sqrt{t})\big), \quad a \ge 0,$$

where Φ *is the c.d.f. of* $\mathcal{N}(0,1)$*. Moreover,* $M_t \overset{\mathcal{L}}{=} |B_t|$*.*

Corollary 8.15 (Distribution of τ_a). *For* $a > 0$*:*

$$\mathbb{P}(\tau_a \le t) = 2\big(1 - \Phi(a/\sqrt{t})\big), \quad f_{\tau_a}(t) = \frac{a}{\sqrt{2\pi t^3}} \exp\Big(-\frac{a^2}{2t}\Big), \quad t > 0.$$

In particular, $\mathbb{E}[\tau_a] = +\infty$*.*

8.7 Brownian Motion and PDEs

Theorem 8.16 (Feynman–Kac Formula (Simple Case))**.** *Let $g : \mathbb{R} \to \mathbb{R}$ be bounded and continuous. The function*

$$u(t,x) = \mathbb{E}_x[g(B_T)] = \mathbb{E}[g(x + B_{T-t})], \quad 0 \le t \le T,$$

is the unique bounded solution of the heat equation:

$$\frac{\partial u}{\partial t} + \frac{1}{2}\frac{\partial^2 u}{\partial x^2} = 0, \quad u(T,x) = g(x).$$

8.8 Local Times

Definition 8.17 (Local Time)**.** The *local time* of Brownian motion at level $a \in \mathbb{R}$ is the process $(L_t^a)_{t\ge 0}$ defined by the *occupation density formula*:

$$\int_0^t f(B_s)\,ds = \int_{-\infty}^{+\infty} f(a) L_t^a\,da$$

for every Borel function $f \ge 0$. Equivalently:

$$L_t^a = \lim_{\varepsilon \to 0} \frac{1}{2\varepsilon} \int_0^t \mathbf{1}_{\{|B_s - a| < \varepsilon\}}\,ds.$$

Theorem 8.18 (Tanaka's Formula)**.** *For all $a \in \mathbb{R}$:*

$$|B_t - a| = |a| + \int_0^t sgn(B_s - a)\,dB_s + L_t^a,$$

where $sgn(x) = \mathbf{1}_{\{x>0\}} - \mathbf{1}_{\{x\le 0\}}$.

8.9 Exercises

Exercise 8.1. Show that Brownian motion is a.s. non-monotone on every interval.

Exercise 8.2. Compute $\mathbb{E}[B_s B_t]$ for $s, t \ge 0$ directly from the definition.

Exercise 8.3. Verify that $(tB_{1/t})_{t>0}$ is a Brownian motion (with the convention $= 0$ at $t = 0$).

Exercise 8.4. Use the reflection principle to compute $\mathbb{P}(\tau_1 \le 4)$.

Exercise 8.5. Let (B_t) be a Brownian motion and $f \in C^2(\mathbb{R})$ with $f'' \le 0$. Show that $(f(B_t))$ is a local supermartingale.

Exercise 8.6. Prove that $\mathbb{E}[\tau_a] = +\infty$ for all $a \ne 0$ using the density of τ_a.

Exercise 8.7. Show that the Brownian bridge $X_t = B_t - tB_1$ $(0 \le t \le 1)$ is a Gaussian process and compute its covariance function. Verify that $X_0 = X_1 = 0$ a.s.

Chapter 9

Itô Stochastic Integral

9.1 Motivation

In 1942, in Kyoto, the mathematician Kiyosi Itô tackled a fundamental problem: how does one define an integral with respect to Brownian motion? Brownian paths are continuous but nowhere differentiable, and their total variation is infinite on every interval. The classical Stieltjes integral simply does not apply. Itô circumvented the obstacle through an ingenious construction: he approximated the integrand by step processes and showed that the limit exists in L^2, thanks to the Itô isometry. The choice of evaluating at the *left endpoint* of each subinterval (unlike Stratonovich, who uses the midpoint) gives the Itô integral the property of being a martingale — a property crucial for finance and probability.

The goal of this chapter is to give rigorous meaning to the integral

$$\int_0^t H_s \, dB_s,$$

where (B_t) is a Brownian motion and (H_t) an adapted process. Since Brownian sample paths have infinite total variation (Theorem 8.11), this integral cannot be defined in the Stieltjes sense. Itô's construction proceeds by approximation.

9.2 Construction for Simple Processes

Definition 9.1 (Simple Process)**.** A process $(H_t)_{t\in[0,T]}$ is *simple* (or *elementary*) if there exist a partition $0 = t_0 < t_1 < \cdots < t_n = T$ and bounded $\mathcal{F}_{t_k}$-measurable r.v. H_k such that:

$$H_t = H_0 \mathbf{1}_{\{0\}}(t) + \sum_{k=0}^{n-1} H_k \mathbf{1}_{(t_k, t_{k+1}]}(t).$$

We write $\mathcal{S}$ for the space of simple processes.

Definition 9.2 (Itô Integral for Simple Processes)**.** For $H \in \mathcal{S}$, we define:

$$I(H) = \int_0^T H_s \, dB_s := \sum_{k=0}^{n-1} H_k (B_{t_{k+1}} - B_{t_k}).$$

Left-Point Evaluation

In Itô's definition, the integrand H_k is evaluated at the *left* endpoint t_k of the interval $(t_k, t_{k+1}]$. This choice is crucial for the martingale property. A different choice (e.g. the midpoint, as in the Stratonovich integral) gives a different result.

Proposition 9.3 (Properties for Simple Processes)**.** For $H, G \in \mathcal{S}$ and $a, b \in \mathbb{R}$:

1. **Linearity**: $I(aH + bG) = aI(H) + bI(G)$.
2. **Zero expectation**: $\mathbb{E}[I(H)] = 0$.
3. **Itô isometry**: $\mathbb{E}[I(H)^2] = \mathbb{E}\big[\int_0^T H_s^2\, ds\big]$.
4. **Martingale property**: the process $t \mapsto \int_0^t H_s\, dB_s$ is a martingale.

Proof of Itô isometry.

$$\begin{aligned}\mathbb{E}[I(H)^2] &= \mathbb{E}\Big[\Big(\sum_{k=0}^{n-1} H_k \Delta B_k\Big)^2\Big] \\ &= \sum_{k=0}^{n-1} \mathbb{E}[H_k^2(\Delta B_k)^2] + 2\sum_{j<k} \mathbb{E}[H_j \Delta B_j H_k \Delta B_k].\end{aligned}$$

The cross terms vanish since, for $j < k$:

$$\mathbb{E}[H_j \Delta B_j H_k \Delta B_k] = \mathbb{E}[\underbrace{H_j \Delta B_j H_k}_{\mathcal{F}_{t_k}\text{-measurable}} \underbrace{\mathbb{E}[\Delta B_k \mid \mathcal{F}_{t_k}]}_{=0}] = 0.$$

For the diagonal terms:

$$\mathbb{E}[H_k^2(\Delta B_k)^2] = \mathbb{E}[H_k^2 \mathbb{E}[(\Delta B_k)^2 \mid \mathcal{F}_{t_k}]] = \mathbb{E}[H_k^2(t_{k+1} - t_k)].$$

Hence $\mathbb{E}[I(H)^2] = \sum_k \mathbb{E}[H_k^2(t_{k+1} - t_k)] = \mathbb{E}[\int_0^T H_s^2\, ds]$. □

9.3 Extension to L^2 Processes

Definition 9.4 (Space $\mathcal{H}^2$)**.** The space of admissible integrands is:

$$\mathcal{H}^2 = \mathcal{H}^2([0,T]) = \Big\{H \text{ adapted} : \mathbb{E}\Big[\int_0^T H_s^2\, ds\Big] < \infty\Big\}.$$

Equipped with the norm $\|H\|_{\mathcal{H}^2} = \big(\mathbb{E}[\int_0^T H_s^2\, ds]\big)^{1/2}$, this is a Hilbert space.

Theorem 9.5 (Extension of the Itô Integral)**.** *The linear isometric map $I : \mathcal{S} \to L^2(\Omega)$ extends uniquely to an isometry $I : \mathcal{H}^2 \to L^2(\Omega)$. The extended integral, denoted*

$$\int_0^T H_s\, dB_s = I(H),$$

satisfies the same properties as for simple processes: linearity, zero expectation, Itô isometry.

Proof sketch. The space $\mathcal{S}$ is dense in $\mathcal{H}^2$ (any adapted square-integrable process can be approximated by simple processes). The Itô isometry guarantees that I is an isometry from $(\mathcal{S}, \|\cdot\|_{\mathcal{H}^2})$ into $(L^2(\Omega), \|\cdot\|_{L^2})$. By completeness of $L^2(\Omega)$, I extends uniquely. □

9.4 Properties of the Itô Integral

Theorem 9.6 (Fundamental Properties)**.** *Let $H \in \mathcal{H}^2([0,T])$. The process $M_t = \int_0^t H_s \, dB_s$ satisfies:*

1. ***Continuity**: (M_t) has a version with continuous sample paths.*
2. ***Martingale**: $(M_t)_{0 \le t \le T}$ is a square-integrable martingale with respect to $(\mathcal{F}_t)$.*
3. ***Itô isometry**: $\mathbb{E}[M_t^2] = \mathbb{E}[\int_0^t H_s^2 \, ds]$.*
4. ***Quadratic variation**: $\langle M \rangle_t = \int_0^t H_s^2 \, ds$.*

Key Properties of the Itô Integral

For $M_t = \int_0^t H_s \, dB_s$ with $H \in \mathcal{H}^2$:

$$\begin{aligned} \mathbb{E}[M_t] &= 0 \\ \mathbb{E}[M_t^2] &= \mathbb{E}\left[\int_0^t H_s^2 \, ds\right] \quad \text{(Itô isometry)} \\ \langle M \rangle_t &= \int_0^t H_s^2 \, ds \\ dM_t &= H_t \, dB_t \end{aligned}$$

9.5 Extension to Locally L^2 Integrands

Definition 9.7 (Generalised Itô Integral)**.** For an adapted process H with $\int_0^T H_s^2 \, ds < \infty$ a.s. (but not necessarily $\mathbb{E}[\int_0^T H_s^2 \, ds] < \infty$), the Itô integral is defined by localisation: set $\tau_n = \inf\{t : \int_0^t H_s^2 \, ds \ge n\}$ and $\int_0^t H_s \, dB_s = \lim_{n\to\infty} \int_0^{t \wedge \tau_n} H_s \, dB_s$.

The resulting process is a continuous *local martingale*.

9.6 Stratonovich Integral

Definition 9.8 (Stratonovich Integral)**.** The *Stratonovich integral* is defined by:

$$\int_0^T H_s \circ dB_s = \lim_{|\Pi| \to 0} \sum_{k=0}^{n-1} \frac{H_{t_k} + H_{t_{k+1}}}{2} (B_{t_{k+1}} - B_{t_k}).$$

Proposition 9.9 (Itô–Stratonovich Relation)**.** If $H_t = f(B_t)$ with $f \in C^1(\mathbb{R})$, then:

$$\int_0^T f(B_s) \circ dB_s = \int_0^T f(B_s) \, dB_s + \frac{1}{2} \int_0^T f'(B_s) \, ds.$$

Remark 9.10. The Stratonovich integral obeys the classical rules of differential calculus (no Itô correction term) but is not a martingale in general. The Itô integral is preferred in finance (martingale property), while the Stratonovich integral is natural in physics (coordinate invariance).

9.7 Computational Examples

Example 9.11 ($\int_0^t B_s\,dB_s$)**.** For a partition $\Pi = \{0 = t_0 < t_1 < \cdots < t_n = t\}$:

$$\sum_{k=0}^{n-1} B_{t_k}(B_{t_{k+1}} - B_{t_k}) = \sum_{k=0}^{n-1}\left[\frac{1}{2}(B_{t_{k+1}}^2 - B_{t_k}^2) - \frac{1}{2}(B_{t_{k+1}} - B_{t_k})^2\right]$$
$$= \frac{1}{2}B_t^2 - \frac{1}{2}\sum_{k=0}^{n-1}(\Delta B_k)^2.$$

Taking the limit and using $\sum(\Delta B_k)^2 \to t$ (quadratic variation):

$$\int_0^t B_s\,dB_s = \frac{1}{2}B_t^2 - \frac{1}{2}t.$$

The $-t/2$ Term

The result $\int_0^t B_s\,dB_s = \frac{1}{2}B_t^2 - \frac{1}{2}t$ shows that stochastic calculus differs from ordinary calculus. In ordinary calculus, one would have $\int_0^t x\,dx = x^2/2$. The term $-t/2$ is the "Itô correction term", a consequence of the non-zero quadratic variation of Brownian motion.

Example 9.12 ($\int_0^t e^{B_s - s/2}\,dB_s$)**.** Set $M_t = e^{B_t - t/2}$. One verifies in the next chapter (Itô's formula) that $dM_t = M_t\,dB_t$, hence:

$$\int_0^t e^{B_s - s/2}\,dB_s = M_t - M_0 = e^{B_t - t/2} - 1.$$

9.8 Exercises

Exercise 9.1. Compute $\mathbb{E}[(\int_0^1 B_s\,dB_s)^2]$ using the Itô isometry.

Exercise 9.2. Let $f \in C^1(\mathbb{R})$. Show that $\int_0^t f(s)\,dB_s$ is a Gaussian r.v. and compute its law.

Exercise 9.3. Compute $\int_0^t s\,dB_s$ and verify that it is a Gaussian r.v. with variance $t^3/3$.

Exercise 9.4. Show that $\int_0^t B_s^2\,dB_s = \frac{1}{3}B_t^3 - \int_0^t B_s\,ds$ using Itô's formula (anticipate the next chapter, or prove directly by approximation).

Exercise 9.5. Let $H_t = \operatorname{sgn}(B_t)$ (sign of Brownian motion). Verify that $H \in \mathcal{H}^2$ and that $\int_0^t \operatorname{sgn}(B_s)\,dB_s = |B_t| - L_t^0$ (Tanaka's formula).

Exercise 9.6. Compare the Itô and Stratonovich integrals of $\int_0^t B_s\,dB_s$ and verify the correction formula.

Chapter 10

Itô Calculus and Stochastic Differential Equations

"Stochastic calculus is the art of giving rigorous meaning to integrals that classical analysis refuses."

This chapter develops the theory of the Itô integral, establishes Itô's formula (the stochastic chain rule), and studies existence and uniqueness of solutions to stochastic differential equations (SDEs). We treat in detail geometric Brownian motion and the Ornstein–Uhlenbeck process.

10.1 Construction of the Itô Integral

Definition 10.1 (Itô Integral for Simple Processes)**.** Let $(B_t)_{t\geq 0}$ be a standard Brownian motion on $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$. For a simple (step) process $H_t = \sum_{i=0}^{n-1} h_i \mathbb{1}_{(t_i,t_{i+1}]}(t)$ where each h_i is $\mathcal{F}_{t_i}$-measurable and bounded, define:

$$\int_0^T H_s\, dB_s = \sum_{i=0}^{n-1} h_i\, (B_{t_{i+1}} - B_{t_i}).$$

Theorem 10.2 (Itô Isometry)**.** *For any adapted simple process H:*

$$\mathbb{E}\left[\left(\int_0^T H_s\, dB_s\right)^2\right] = \mathbb{E}\left[\int_0^T H_s^2\, ds\right].$$

Proof. Expand the square:

$$\mathbb{E}\left[\left(\sum_i h_i\, \Delta B_i\right)^2\right] = \sum_i \mathbb{E}[h_i^2]\, \mathbb{E}[(\Delta B_i)^2] + 2\sum_{i<j} \mathbb{E}[h_i\, \Delta B_i\, h_j\, \Delta B_j].$$

The first sum gives $\sum_i \mathbb{E}[h_i^2]\,(t_{i+1} - t_i)$. For $i < j$, $h_j\, \Delta B_j$ is independent of $\mathcal{F}_{t_j}$ and $\mathbb{E}[\Delta B_j] = 0$, so the cross terms vanish. Hence $\mathbb{E}[(\int_0^T H\, dB)^2] = \mathbb{E}[\int_0^T H_s^2\, ds]$. $\square$

Intuition

The Itô isometry is the stochastic analogue of Parseval's identity in Fourier analysis. It allows extending the Itô integral by continuity to all adapted square-integrable processes, using completeness of L^2.

Definition 10.3 (Extension to L^2)**.** Let $\mathcal{H}^2_T$ denote the space of adapted processes $(H_t)_{0\le t\le T}$ with $\mathbb{E}[\int_0^T H_s^2\,ds] < \infty$. The Itô integral extends uniquely to $\mathcal{H}^2_T$ by density of simple processes and the Itô isometry.

Proposition 10.4 (Fundamental Properties)**.** The Itô integral $M_t = \int_0^t H_s\,dB_s$ satisfies:

1. **Martingale:** (M_t) is a square-integrable martingale.
2. **Linearity:** $\int_0^t(\alpha H_s + \beta K_s)\,dB_s = \alpha\int_0^t H_s\,dB_s + \beta\int_0^t K_s\,dB_s$.
3. **Continuous paths:** $t \mapsto M_t$ is a.s. continuous.

Attention

The Itô integral is *not* a Riemann–Stieltjes integral because Brownian paths have infinite variation. Evaluating the integrand at the *left* endpoint of each subinterval is crucial: choosing the midpoint would give the Stratonovich integral, which obeys different calculus rules.

10.2 Itô's Formula

Theorem 10.5 (Itô's Formula, One Dimension)**.** *Let $X_t = X_0 + \int_0^t b_s\,ds + \int_0^t \sigma_s\,dB_s$ be an Itô process and $f \in C^2(\mathbb{R})$. Then:*

$$f(X_t) = f(X_0) + \int_0^t f'(X_s)\,b_s\,ds + \int_0^t f'(X_s)\,\sigma_s\,dB_s + \frac{1}{2}\int_0^t f''(X_s)\,\sigma_s^2\,ds.$$

Proof sketch. Write the Taylor expansion to second order: $f(X_{t+dt}) - f(X_t) \approx f'(X_t)\,dX_t + \frac{1}{2}f''(X_t)\,(dX_t)^2$. The key rule is $(dB_t)^2 = dt$ (in mean square), while $dt\,dB_t = 0$ and $(dt)^2 = 0$. Therefore $(dX_t)^2 = \sigma_t^2\,(dB_t)^2 = \sigma_t^2\,dt$. Integration yields the result. □

Remark 10.6. The extra term $\frac{1}{2}f''(X_s)\,\sigma_s^2\,ds$, absent in classical calculus, is called the *Itô correction term*. It arises from the non-zero quadratic variation of Brownian motion.

Key Formulas

Itô Calculus Rules

$$(dB_t)^2 = dt, \qquad dt\cdot dB_t = 0, \qquad (dt)^2 = 0$$

$$d(X_tY_t) = X_t\,dY_t + Y_t\,dX_t + dX_t\,dY_t \quad \text{(product rule)}$$

$$df(X_t) = f'(X_t)\,dX_t + \tfrac{1}{2}f''(X_t)\,(dX_t)^2$$

Corollary 10.7 (Multidimensional Itô's Formula)**.** *Let* $\mathbf{X}_t = (X_t^1, \ldots, X_t^d)$ *be a vector of Itô processes and* $f \in C^2(\mathbb{R}^d)$*. Then:*

$$df(\mathbf{X}_t) = \sum_{i=1}^{d} \frac{\partial f}{\partial x_i}\, dX_t^i + \frac{1}{2}\sum_{i,j=1}^{d} \frac{\partial^2 f}{\partial x_i\, \partial x_j}\, d\langle X^i, X^j\rangle_t.$$

Example 10.8. Compute $d(B_t^2)$ where B_t is a standard Brownian motion. With $f(x) = x^2$, $f'(x) = 2x$, $f''(x) = 2$:

$$d(B_t^2) = 2B_t\, dB_t + dt.$$

Thus $\int_0^t B_s\, dB_s = \frac{1}{2}(B_t^2 - t)$, showing that the Itô integral differs from the naive Riemann–Stieltjes result.

10.3 Stochastic Differential Equations

Definition 10.9 (Stochastic Differential Equation)**.** An **SDE** is an equation of the form:

$$dX_t = b(t, X_t)\, dt + \sigma(t, X_t)\, dB_t, \qquad X_0 = x_0,$$

where $b : [0,T] \times \mathbb{R} \to \mathbb{R}$ is the *drift* coefficient and $\sigma : [0,T] \times \mathbb{R} \to \mathbb{R}$ the *diffusion* coefficient.

Theorem 10.10 (Existence and Uniqueness (Itô–Lipschitz))**.** *If* b *and* σ *are Lipschitz in* x *uniformly in* t*:*

$$|b(t,x) - b(t,y)| + |\sigma(t,x) - \sigma(t,y)| \leq K\, |x - y|\,,$$

and have linear growth: $|b(t,x)| + |\sigma(t,x)| \leq C(1 + |x|)$*, then the SDE admits a unique strong solution* $(X_t)_{0\leq t\leq T}$ *that is adapted, has continuous paths, and satisfies* $\mathbb{E}[\sup_{0\leq t\leq T} X_t^2] < \infty$*.*

Proof sketch. Use Picard iteration. Set $X_t^{(0)} = x_0$ and:

$$X_t^{(n+1)} = x_0 + \int_0^t b(s, X_s^{(n)})\, ds + \int_0^t \sigma(s, X_s^{(n)})\, dB_s.$$

The Lipschitz condition and Itô isometry give $\mathbb{E}[\sup_{s\leq t} \left|X_s^{(n+1)} - X_s^{(n)}\right|^2] \leq (Ct)^n/n!$ (stochastic Gronwall lemma), yielding uniform L^2 convergence. ☐

10.4 Geometric Brownian Motion

Definition 10.11 (Geometric Brownian Motion)**. Geometric Brownian motion** (GBM) is the solution of:

$$dS_t = \mu\, S_t\, dt + \sigma\, S_t\, dB_t, \qquad S_0 > 0.$$

Theorem 10.12 (Explicit Solution of GBM)**.** *The solution of GBM is:*

$$S_t = S_0 \exp\left[\left(\mu - \frac{\sigma^2}{2}\right) t + \sigma\, B_t\right].$$

Proof. Set $Y_t = \ln S_t$. By Itô's formula with $f(x) = \ln x$:

$$dY_t = \frac{1}{S_t}\, dS_t - \frac{1}{2S_t^2}\,(dS_t)^2 = \mu\, dt + \sigma\, dB_t - \frac{\sigma^2}{2}\, dt = \left(\mu - \frac{\sigma^2}{2}\right) dt + \sigma\, dB_t.$$

Integrating: $Y_t = Y_0 + (\mu - \sigma^2/2)\, t + \sigma B_t$. ☐

CHAPTER 10. ITÔ CALCULUS AND STOCHASTIC DIFFERENTIAL EQUATIONS

Intuition

GBM is the canonical model for financial asset prices (Black–Scholes model). The term $-\sigma^2/2$ in the exponent is the Itô correction: without it, one would get $\mathbb{E}[S_t] = S_0 e^{(\mu+\sigma^2/2)t}$ instead of the correct $\mathbb{E}[S_t] = S_0 e^{\mu t}$.

10.5 The Ornstein–Uhlenbeck Process

Definition 10.13 (Ornstein–Uhlenbeck Process)**.** The **Ornstein–Uhlenbeck** process solves:

$$dX_t = -\theta\,(X_t - \mu)\,dt + \sigma\,dB_t, \qquad X_0 = x_0,$$

where $\theta > 0$ is the mean-reversion rate, μ the long-run mean, and $\sigma > 0$ the volatility.

Proposition 10.14 (Explicit Solution)**.** The solution is:

$$X_t = \mu + (x_0 - \mu)\,e^{-\theta t} + \sigma \int_0^t e^{-\theta(t-s)}\,dB_s.$$

In particular, X_t is Gaussian with $\mathbb{E}[X_t] = \mu + (x_0 - \mu)e^{-\theta t}$ and $\mathrm{Var}(X_t) = \frac{\sigma^2}{2\theta}(1 - e^{-2\theta t})$.

Proof. Set $Y_t = e^{\theta t}(X_t - \mu)$. By Itô's formula, $dY_t = \theta e^{\theta t}(X_t - \mu)\,dt + e^{\theta t}\,dX_t = \sigma e^{\theta t}\,dB_t$. Integrating: $Y_t = Y_0 + \sigma \int_0^t e^{\theta s}\,dB_s$, so $X_t - \mu = (x_0 - \mu)e^{-\theta t} + \sigma \int_0^t e^{-\theta(t-s)}\,dB_s$. The variance follows from the Itô isometry. □

Remark 10.15. The Ornstein–Uhlenbeck process is the unique Gaussian process that is simultaneously stationary (as $t \to \infty$), Markovian, and has continuous paths. Its invariant measure is $\mathcal{N}(\mu, \sigma^2/(2\theta))$.

10.6 The Langevin Equation

Definition 10.16 (Langevin Equation)**.** The **Langevin equation** for a particle of mass m subject to friction $-\gamma v$ and thermal noise is:

$$m\,dv_t = -\gamma\,v_t\,dt + \sqrt{2\gamma k_B T}\,dB_t,$$

where k_B is Boltzmann's constant and T the temperature.

Proposition 10.17 (Fluctuation-Dissipation Relation)**.** At thermodynamic equilibrium, the velocity variance satisfies:

$$\mathrm{Var}(v_\infty) = \frac{k_B T}{m},$$

consistent with the equipartition theorem. This relation links the noise intensity ($2\gamma k_B T$) to the friction coefficient (γ).

Exercise 10.1. Show that $\int_0^t B_s\,dB_s = \frac{1}{2}(B_t^2 - t)$ using Itô's formula. Interpret the difference with the deterministic result $\int_0^t x\,dx = x^2/2$.

Exercise 10.2. Solve the SDE $dX_t = X_t\,dt + X_t\,dB_t$ with $X_0 = 1$. Compute $\mathbb{E}[X_t]$ and $\mathrm{Var}(X_t)$.

Exercise 10.3. Simulate the Ornstein–Uhlenbeck process numerically using the Euler–Maruyama scheme:

$$X_{n+1} = X_n - \theta(X_n - \mu)\,\Delta t + \sigma\sqrt{\Delta t}\,Z_n, \quad Z_n \sim \mathcal{N}(0,1).$$

Verify that the stationary distribution is $\mathcal{N}(\mu, \sigma^2/(2\theta))$.

Exercise 10.4. Let $f(t,x) = e^{-rt}\,x^\alpha$ and X_t be a GBM with parameters (μ,σ). Apply Itô's formula to compute $df(t,X_t)$ and determine the value of α for which $f(t,X_t)$ is a martingale.

Exercise 10.5 (Stratonovich–Itô conversion)**.** Consider the Stratonovich SDE:

$$dX_t = X_t \circ dB_t.$$

1. Recall the general relationship between the Stratonovich integral $\int_0^t \sigma(X_s) \circ dB_s$ and the Itô integral, involving the correction term $\frac{1}{2}\int_0^t \sigma'(X_s)\,\sigma(X_s)\,ds$.
2. Convert the above SDE into an Itô SDE.
3. Solve the resulting Itô SDE and verify that the solution is $X_t = X_0\,e^{B_t}$. Compare with GBM having parameters $\mu = \sigma^2/2$, $\sigma = 1$.

Exercise 10.6 (Girsanov's theorem)**.** Let $(B_t)_{0\le t\le T}$ be a standard Brownian motion under $\mathbb{P}$ and let $\theta \in \mathbb{R}$ be a constant. Define the exponential martingale:

$$Z_t = \exp\left(\theta B_t - \frac{\theta^2}{2}\,t\right).$$

1. Verify using Itô's formula that Z_t is a martingale under $\mathbb{P}$.
2. Define the measure $\mathbb{Q}$ by $d\mathbb{Q} = Z_T\,d\mathbb{P}$. Show that $\widetilde{B}_t = B_t - \theta t$ is a standard Brownian motion under $\mathbb{Q}$ (Girsanov's theorem).
3. Application: let X_t solve $dX_t = \mu\,dt + \sigma\,dB_t$ under $\mathbb{P}$. Find a measure $\mathbb{Q}$ under which X_t is a (driftless) Brownian motion. Give $d\mathbb{Q}/d\mathbb{P}$ explicitly.

Exercise 10.7 (Strong vs. weak solutions)**.** Consider Tanaka's SDE:

$$dX_t = \operatorname{sign}(X_t)\,dB_t, \qquad X_0 = 0,$$

where $\operatorname{sign}(x) = \nVdash_{x>0} - \nVdash_{x\le 0}$.

1. Show that the diffusion coefficient $\sigma(x) = \operatorname{sign}(x)$ fails to be Lipschitz at $x = 0$. Why does the Itô–Lipschitz strong existence theorem not apply?
2. Show that if X_t is a solution, then $Y_t = |X_t|$ satisfies $Y_t = |B_t|$ in law (you may use Tanaka's formula). Deduce that a *weak* solution exists.
3. Explain why no *strong* solution exists (i.e., no solution measurable with respect to the filtration generated by B). *Hint*: show that $\operatorname{sign}(X_t)$ cannot be determined from the Brownian motion B alone.

CHAPTER 10. ITÔ CALCULUS AND STOCHASTIC DIFFERENTIAL EQUATIONS

Chapter 11

Applications

"Theory is when you know everything but nothing works. Practice is when everything works but nobody knows why."

— attributed to Albert Einstein

This chapter presents major applications of stochastic calculus: the Black–Scholes formula in finance, the Feynman–Kac theorem linking PDEs and SDEs, Kalman filtering, applications in biology, and Monte Carlo methods for SDEs.

11.1 The Black–Scholes Formula

Definition 11.1 (Black–Scholes Model)**.** In the **Black–Scholes model**, the price of a risky asset follows geometric Brownian motion:

$$dS_t = \mu\, S_t\, dt + \sigma\, S_t\, dB_t,$$

and there exists a risk-free asset with return r. One assumes no arbitrage opportunities and continuous frictionless trading.

Theorem 11.2 (Black–Scholes Formula)**.** *The price at time t of a European call option with strike K and maturity T is:*

$$C(t, S_t) = S_t\, \Phi(d_1) - K\, e^{-r(T-t)}\, \Phi(d_2),$$

where Φ is the standard normal CDF and:

$$d_1 = \frac{\ln(S_t/K) + (r + \sigma^2/2)(T-t)}{\sigma\sqrt{T-t}}, \qquad d_2 = d_1 - \sigma\sqrt{T-t}.$$

Proof. Under the risk-neutral measure $\mathbb{Q}$, the discounted price is a martingale: $C(t, S_t) = e^{-r(T-t)}\, \mathbb{E}^{\mathbb{Q}}[(S_T - K)^+ \mid \mathcal{F}_t]$. Under $\mathbb{Q}$, $S_T = S_t \exp[(r - \sigma^2/2)(T-t) + \sigma(B_T^{\mathbb{Q}} - B_t^{\mathbb{Q}})]$ where $B^{\mathbb{Q}}$ is a $\mathbb{Q}$-Brownian motion (Girsanov's theorem). Setting $Z = (B_T^{\mathbb{Q}} - B_t^{\mathbb{Q}})/\sqrt{T-t} \sim \mathcal{N}(0,1)$:

$$C = e^{-r(T-t)} \int_{-\infty}^{\infty} \left(S_t\, e^{(r-\sigma^2/2)(T-t)+\sigma\sqrt{T-t}\,z} - K\right)^+ \frac{e^{-z^2/2}}{\sqrt{2\pi}}\, dz.$$

The integral splits over $\{z > -d_2\}$. Substituting $z' = z - \sigma\sqrt{T-t}$ in the first term yields the stated expressions. $\square$

Intuition

The Black–Scholes formula shows that the option price does *not* depend on the expected return μ of the asset, only on its volatility σ. This is a consequence of no-arbitrage: the option can be perfectly hedged by dynamically adjusting a portfolio of the stock and the risk-free asset.

Definition 11.3 (Black–Scholes PDE)**.** The pricing function $V(t, S)$ satisfies the PDE:

$$\frac{\partial V}{\partial t} + r\,S\,\frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2 S^2\,\frac{\partial^2 V}{\partial S^2} - r\,V = 0,$$

with terminal condition $V(T, S) = (S - K)^+$ for a call.

Key Formulas

Black–Scholes Formulas and Greeks

$$C = S\,\Phi(d_1) - Ke^{-r\tau}\Phi(d_2), \quad \tau = T - t$$

$$\Delta = \frac{\partial C}{\partial S} = \Phi(d_1), \qquad \Gamma = \frac{\partial^2 C}{\partial S^2} = \frac{\varphi(d_1)}{S\sigma\sqrt{\tau}}$$

$$\Theta = \frac{\partial C}{\partial t} = -\frac{S\sigma\varphi(d_1)}{2\sqrt{\tau}} - rKe^{-r\tau}\Phi(d_2), \qquad \mathcal{V} = \frac{\partial C}{\partial \sigma} = S\sqrt{\tau}\,\varphi(d_1)$$

11.2 The Feynman–Kac Theorem

Theorem 11.4 (Feynman–Kac)**.** *Let $u(t, x)$ solve the PDE:*

$$\frac{\partial u}{\partial t} + b(x)\,\frac{\partial u}{\partial x} + \frac{1}{2}\sigma^2(x)\,\frac{\partial^2 u}{\partial x^2} - r\,u = 0, \qquad u(T, x) = g(x).$$

Then u admits the probabilistic representation:

$$u(t, x) = \mathbb{E}\big[e^{-r(T-t)}\,g(X_T) \mid X_t = x\big],$$

where X solves $dX_s = b(X_s)\,ds + \sigma(X_s)\,dB_s$.

Proof sketch. Apply Itô's formula to $e^{-rs}\,u(s, X_s)$ on $[t, T]$. The ds term vanishes because u satisfies the PDE, leaving a stochastic integral (a martingale). Taking the conditional expectation, the stochastic integral drops out, giving $e^{-rt}\,u(t, x) = \mathbb{E}[e^{-rT}\,g(X_T) \mid X_t = x]$. □

Remark 11.5. The Feynman–Kac theorem establishes a fundamental link between parabolic PDEs and diffusion processes. It allows solving PDEs via Monte Carlo simulation, and conversely, analyzing stochastic problems using PDE techniques.

11.3 Kalman Filtering

Definition 11.6 (Linear Filtering Model)**.** Consider the continuous-time linear system:

$$\begin{aligned} dX_t &= A\,X_t\,dt + \sigma_X\,dB_t^{(1)} \quad \text{(signal)},\\ dY_t &= H\,X_t\,dt + \sigma_Y\,dB_t^{(2)} \quad \text{(observation)}, \end{aligned}$$

where $B^{(1)}$ and $B^{(2)}$ are independent Brownian motions. The *filtering problem* is to estimate X_t from the observations $(Y_s)_{0\le s\le t}$.

Theorem 11.7 (Kalman–Bucy Filter)**.** *The optimal estimator $\hat{X}_t = \mathbb{E}[X_t \mid \mathcal{F}_t^Y]$ satisfies the SDE:*

$$d\hat{X}_t = A\,\hat{X}_t\,dt + \frac{P_t\,H^T}{\sigma_Y^2}\big(dY_t - H\,\hat{X}_t\,dt\big),$$

where $P_t = \mathbb{E}[(X_t - \hat{X}_t)^2 \mid \mathcal{F}_t^Y]$ is the error variance, solving the Riccati equation:

$$\frac{dP_t}{dt} = 2A\,P_t + \sigma_X^2 - \frac{P_t^2\,H^2}{\sigma_Y^2}.$$

Intuition

The Kalman filter achieves an optimal trade-off between the dynamic model prediction ($A\,\hat{X}_t\,dt$) and the innovation from observations ($dY_t - H\,\hat{X}_t\,dt$). The Kalman gain $K_t = P_tH^T/\sigma_Y^2$ weighs these two sources: the larger the uncertainty P_t, the more weight is given to observations.

11.4 Applications in Biology

Definition 11.8 (Stochastic Gene Expression)**.** Gene expression is an intrinsically stochastic process. The protein count P_t can be modeled by the Langevin-type SDE:

$$dP_t = (k_{\text{prod}} - k_{\text{deg}}\,P_t)\,dt + \sqrt{k_{\text{prod}} + k_{\text{deg}}\,P_t}\,dB_t,$$

where k_{prod} is the production rate and k_{deg} the degradation rate.

Proposition 11.9 (Gene Expression Noise)**.** At equilibrium, the mean protein count is $\bar{P} = k_{\text{prod}}/k_{\text{deg}}$ and the *coefficient of variation* (noise measure) is:

$$\text{CV}^2 = \frac{\text{Var}(P)}{\mathbb{E}[P]^2} \approx \frac{1}{\bar{P}}.$$

The noise is thus more significant when the mean molecule count is low.

Remark 11.10. This relation $\text{CV}^2 \approx 1/\bar{P}$ is known as *Poissonian noise.* It was experimentally verified in bacterial cells by Elowitz et al. (2002).

11.5 Monte Carlo Methods for SDEs

Definition 11.11 (Euler–Maruyama Scheme)**.** The **Euler–Maruyama scheme** for the SDE $dX_t = b(X_t)\,dt + \sigma(X_t)\,dB_t$ is:

$$X_{n+1} = X_n + b(X_n)\,\Delta t + \sigma(X_n)\,\sqrt{\Delta t}\,Z_n, \qquad Z_n \sim \mathcal{N}(0,1) \text{ i.i.d.}$$

Theorem 11.12 (Strong Convergence of Euler–Maruyama)**.** *Under Lipschitz conditions on b and σ, the Euler–Maruyama scheme converges with strong order $1/2$:*

$$\mathbb{E}\left[\sup_{0\le t\le T}\left|X_t - X_t^{(\Delta t)}\right|^2\right]^{1/2} \le C\,\sqrt{\Delta t}.$$

Definition 11.13 (Milstein Scheme)**.** The **Milstein scheme** improves Euler–Maruyama by adding the Itô–Taylor correction:

$$X_{n+1} = X_n + b(X_n)\,\Delta t + \sigma(X_n)\,\sqrt{\Delta t}\,Z_n + \frac{1}{2}\,\sigma(X_n)\,\sigma'(X_n)\,(Z_n^2 - 1)\,\Delta t.$$

This scheme achieves strong order 1: $(\mathbb{E}[\sup_t \left|X_t - X_t^{(\Delta t)}\right|^2])^{1/2} \le C\,\Delta t$.

Attention

For option pricing (weak convergence), Euler–Maruyama already has weak order 1, so Milstein offers no improvement for computing $\mathbb{E}[g(X_T)]$. The distinction between strong and weak convergence is crucial in practice.

Key Formulas

Numerical Schemes for SDEs

Euler–Maruyama:	$X_{n+1} = X_n + b\,\Delta t + \sigma\sqrt{\Delta t}\,Z_n$	(strong order 1/2)
Milstein:	$X_{n+1} = X_n + b\,\Delta t + \sigma\sqrt{\Delta t}\,Z_n + \frac{1}{2}\sigma\sigma'(Z_n^2-1)\Delta t$	(strong order 1)
Monte Carlo:	$\mathbb{E}[g(X_T)] \approx \dfrac{1}{M}\displaystyle\sum_{m=1}^{M} g(X_T^{(m)})$	(error $O(1/\sqrt{M})$)

```python
import numpy as np

def euler_maruyama(b, sigma, x0, T, n_steps, n_paths):
    """Euler-Maruyama scheme for a scalar SDE."""
    dt = T / n_steps
    sqrt_dt = np.sqrt(dt)
    X = np.zeros((n_paths, n_steps + 1))
    X[:, 0] = x0
    for k in range(n_steps):
        Z = np.random.randn(n_paths)
        X[:, k+1] = (X[:, k] + b(X[:, k]) * dt
                     + sigma(X[:, k]) * sqrt_dt * Z)
```

```
    return X

# Example: geometric Brownian motion
mu_val, sigma_val, S0 = 0.05, 0.2, 100.0
paths = euler_maruyama(
    b=lambda x: mu_val * x,
    sigma=lambda x: sigma_val * x,
    x0=S0, T=1.0, n_steps=1000, n_paths=10000
)
# Call option price by Monte Carlo
K = 105.0
r = 0.03
call_price = np.exp(-r) * np.mean(np.maximum(paths[:, -1] - K, 0))
```

Exercise 11.1. Derive the Black–Scholes formula for a put option using put-call parity: $C - P = S - Ke^{-r(T-t)}$. Verify that $P = Ke^{-r\tau}\Phi(-d_2) - S\Phi(-d_1)$.

Exercise 11.2. Apply the Feynman–Kac theorem to show that the price of a knock-out barrier option satisfies the Black–Scholes PDE with boundary condition $V(t, B) = 0$ on the barrier $S = B$.

Exercise 11.3. Implement the discrete Kalman filter to track a moving object observed with noise. Compare performance with and without a velocity model.

Exercise 11.4. Compare the Euler–Maruyama and Milstein schemes for GBM: simulate $M = 10{,}000$ paths, estimate $\mathbb{E}[S_T]$, and compare the numerical error for different step sizes Δt. Verify the theoretical convergence orders.

Exercise 11.5. Simulate a stochastic gene expression model with $k_{\text{prod}} = 10$, $k_{\text{deg}} = 0.1$ using the Gillespie algorithm. Verify that the stationary distribution is approximately Poisson and compute the coefficient of variation.

CHAPTER 11. APPLICATIONS

Bibliography

[1] Revuz, D. and Yor, M., *Continuous Martingales and Brownian Motion*, 3rd ed., Springer, 1999.

[2] Karatzas, I. and Shreve, S.E., *Brownian Motion and Stochastic Calculus*, 2nd ed., Springer, 1991.

[3] Durrett, R., *Probability: Theory and Examples*, 5th ed., Cambridge, 2019.

[4] Norris, J.R., *Markov Chains*, Cambridge, 1997.

[5] Protter, P.E., *Stochastic Integration and Differential Equations*, 2nd ed., Springer, 2005.