An Introduction to Stochastic Calculus with Applications to Finance

Ovidiu Calin
Department of Mathematics
Eastern Michigan University
Ypsilanti, MI 48197 USA
ocalin@emich.edu

Preface

The goal of this work is to introduce elementary Stochastic Calculus to senior undergraduate as well as to master students with Mathematics, Economics and Business majors. The author's goal was to capture as much as possible of the spirit of elementary Calculus, at which the students have been already exposed in the beginning of their majors. This assumes a presentation that mimics similar properties of deterministic Calculus, which facilitates the understanding of more complicated concepts of Stochastic Calculus. Since deterministic Calculus books usually start with a brief presentation of elementary functions, and then continue with limits, and other properties of functions, we employed here a similar approach, starting with elementary stochastic processes, different types of limits and pursuing with properties of stochastic processes. The chapters regarding differentiation and integration follow the same pattern. For instance, there is a product rule, a chain-type rule and an integration by parts in Stochastic Calculus, which are modifications of the well-known rules from the elementary Calculus.

Since deterministic Calculus can be used for modeling regular business problems, in the second part of the book we deal with stochastic modeling of business applications, such as Financial Derivatives, whose modeling are solely based on Stochastic Calculus.

In order to make the book available to a wider audience, we sacrificed rigor for clarity. Most of the time we assumed maximal regularity conditions for which the computations hold and the statements are valid. This will be found attractive by both Business and Economics students, who might get lost otherwise in a very profound mathematical textbook where the forest's scenary is obscured by the sight of the trees.

An important feature of this textbook is the large number of solved problems and examples from which will benefit both the beginner as well as the advanced student.

This book grew from a series of lectures and courses given by the author at the Eastern Michigan University (USA), Kuwait University (Kuwait) and Fu-Jen University (Taiwan). Several students read the first draft of these notes and provided valuable feedback, supplying a list of corrections, which is by far exhaustive. Any typos or comments regarding the present material are welcome.

The Author,
Ann Arbor, October 2012

Contents

Part I

Stochastic Calculus

Chapter 1

Basic Notions

1.1 Probability Space

The modern theory of probability stems from the work of A. N. Kolmogorov published in 1933. Kolmogorov associates a random experiment with a probability space, which is a triplet, $(\Omega, \mathcal{F}, P)$, consisting of the set of outcomes, Ω, a σ-field, $\mathcal{F}$, with Boolean algebra properties, and a probability measure, P. In the following sections, each of these elements will be discussed in more detail.

1.2 Sample Space

A *random experiment* in the theory of probability is an experiment whose outcomes cannot be determined in advance. These experiments are done mentally most of the time.

When an experiment is performed, the set of all possible outcomes is called the *sample space*, and we shall denote it by Ω. In financial markets one can regard this also as the *states of the world*, understanding by this all possible states the world might have. The number of states the world that affect the stock market is huge. These would contain all possible values for the vector parameters that describe the world, and is practically infinite.

For some simple experiments the sample space is much smaller. For instance, flipping a coin will produce the sample space with two states $\{H, T\}$, while rolling a die yields a sample space with six states. Choosing randomly a number between 0 and 1 corresponds to a sample space which is the entire segment $(0, 1)$.

All subsets of the sample space Ω form a set denoted by 2^{Ω}. The reason for this notation is that the set of parts of Ω can be put into a bijective correspondence with the set of binary functions $f : \Omega \to \{0, 1\}$. The number of elements of this set is $2^{|\Omega|}$, where $|\Omega|$ denotes the cardinal of Ω. If the set is finite, $|\Omega| = n$, then 2^{Ω} has 2^n elements. If Ω is infinitely countable (i.e. can be put into a bijective correspondence with the set of natural numbers), then $2^{|\Omega|}$ is infinite and its cardinal is the same as that of the real

number set $\mathbb{R}$. As a matter of fact, if Ω represents all possible states of the financial world, then 2^Ω describes all possible events, which might happen in the market; this is a supposed to be a fully description of the total information of the financial world.

The following couple of examples provide instances of sets 2^Ω in the finite and infinite cases.

Example 1.2.1 *Flip a coin and measure the occurrence of outcomes by 0 and 1: associate a 0 if the outcome does not occur and a 1 if the outcome occurs. We obtain the following four possible assignments:*

$$\{H \to 0, T \to 0\}, \ \ \{H \to 0, T \to 1\}, \ \ \{H \to 1, T \to 0\}, \ \ \{H \to 1, T \to 1\},$$

so the set of subsets of $\{H, T\}$ can be represented as 4 sequences of length 2 formed with 0 and 1: $\{0, 0\}$, $\{0, 1\}$, $\{1, 0\}$, $\{1, 1\}$. These correspond in order to the sets $\emptyset, \{T\}, \{H\}, \{H, T\}$, which is the set $2^{\{H,T\}}$.

Example 1.2.2 *Pick a natural number at random. Any subset of the sample space corresponds to a sequence formed with 0 and 1. For instance, the subset $\{1, 3, 5, 6\}$ corresponds to the sequence $10101100000\ldots$ having 1 on the 1st, 3rd, 5th and 6th places and 0 in rest. It is known that the number of these sequences is infinite and can be put into a bijective correspondence with the real number set $\mathbb{R}$. This can be also written as $|2^{\mathbb{N}}| = |\mathbb{R}|$, and stated by saying that the set of all subsets of natural numbers $\mathbb{N}$ has the same cardinal as the real numbers set $\mathbb{R}$.*

1.3 Events and Probability

The set of parts 2^Ω satisfies the following properties:

1. It contains the empty set $\emptyset$;
2. If it contains a set A, then it also contains its complement $\bar{A} = \Omega \backslash A$;
3. It is closed with regard to unions, i.e., if $A_1, A_2, \ldots$ is a sequence of sets, then their union $A_1 \cup A_2 \cup \cdots$ also belongs to 2^Ω.

Any subset $\mathcal{F}$ of 2^Ω that satisfies the previous three properties is called a *σ-field*. The sets belonging to $\mathcal{F}$ are called *events*. This way, the complement of an event, or the union of events is also an event. We say that an event occurs if the outcome of the experiment is an element of that subset.

The chance of occurrence of an event is measured by a probability function $P : \mathcal{F} \to [0, 1]$ which satisfies the following two properties:

1. $P(\Omega) = 1$;

2. For any mutually disjoint events $A_1, A_2, \dots \in \mathcal{F}$,

$$P(A_1 \cup A_2 \cup \cdots) = P(A_1) + P(A_2) + \cdots .$$

The triplet $(\Omega, \mathcal{F}, P)$ is called a *probability space*. This is the main setup in which the probability theory works.

Example 1.3.1 *In the case of flipping a coin, the probability space has the following elements:* $\Omega = \{H, T\}$, $\mathcal{F} = \{\emptyset, \{H\}, \{T\}, \{H, T\}\}$ *and* P *defined by* $P(\emptyset) = 0$, $P(\{H\}) = \frac{1}{2}$, $P(\{T\}) = \frac{1}{2}$, $P(\{H, T\}) = 1$.

Example 1.3.2 *Consider a finite sample space* $\Omega = \{s_1, \dots, s_n\}$, *with the* σ*-field* $\mathcal{F} = 2^{\Omega}$, *and probability given by* $P(A) = |A|/n$, $\forall A \in \mathcal{F}$. *Then* $(\Omega, 2^{\Omega}, P)$ *is called the classical probability space.*

1.4 Random Variables

Since the σ-field $\mathcal{F}$ provides the knowledge about which events are possible on the considered probability space, then $\mathcal{F}$ can be regarded as the information component of the probability space $(\Omega, \mathcal{F}, P)$. A *random variable* X is a function that assigns a numerical value to each state of the world, $X : \Omega \to \mathbb{R}$, such that the values taken by X are known to someone who has access to the information $\mathcal{F}$. More precisely, given any two numbers $a, b \in \mathbb{R}$, then all the states of the world for which X takes values between a and b forms a set that is an event (an element of $\mathcal{F}$), i.e.

$$\{\omega \in \Omega; a < X(\omega) < b\} \in \mathcal{F}.$$

Another way of saying this is that X is an $\mathcal{F}$*-measurable function.* It is worth noting that in the case of the classical field of probability the knowledge is maximal since $\mathcal{F} = 2^{\Omega}$, and hence the measurability of random variables is automatically satisfied. From now on instead of *measurable* terminology we shall use the more suggestive word *predictable*. This will make more sense in a future section when we shall introduce conditional expectations.

Example 1.4.1 *Let* $X(\omega)$ *be the number of people who want to buy houses, given the state of the market* ω. *Is* X *predictable? This would mean that given two numbers, say* $a = 10,000$ *and* $b = 50,000$, *we know all the market situations* ω *for which there are at least* $10,000$ *and at most* $50,000$ *people willing to purchase houses. Many times, in theory, it makes sense to assume that we have enough knowledge to assume* X *predictable.*

Example 1.4.2 *Consider the experiment of flipping three coins. In this case* Ω *is the set of all possible triplets. Consider the random variable* X *which gives the number of*

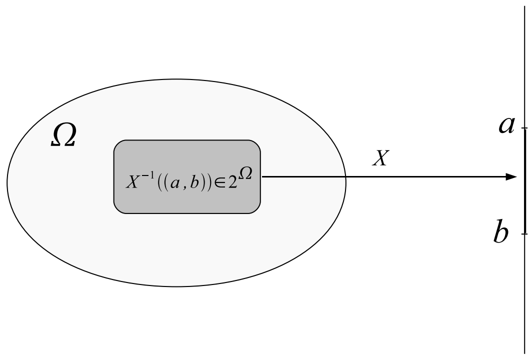


Figure 1.1: *If any pullback $X^{-1}\big((a,b)\big)$ is known, then the random variable $X : \Omega \to \mathbb{R}$ is 2^{Ω}-measurable.*

tails obtained. For instance $X(HHH) = 0$, $X(HHT) = 1$, etc. The sets

$$\{\omega; X(\omega) = 0\} = \{HHH\}, \quad \{\omega; X(\omega) = 1\} = \{HHT, HTH, THH\},$$
$$\{\omega; X(\omega) = 3\} = \{TTT\}, \quad \{\omega; X(\omega) = 2\} = \{HTT, THT, TTH\}$$

belong to 2^{Ω}, and hence X is a random variable.

Example 1.4.3 *A graph is a set of elements, called nodes, and a set of unordered pairs of nodes, called edges. Consider the set of nodes $\mathcal{N} = \{n_1, n_2, \ldots, n_k\}$ and the set of edges $\mathcal{E} = \{(n_i, n_j), 1 \le i, j \le n, i \neq j\}$. Define the probability space $(\Omega, \mathcal{F}, P)$, where*
- ◦ *the sample space is the the complete graph, $\Omega = \mathcal{N} \cup \mathcal{E}$;*
- ◦ *the σ-field $\mathcal{F}$ is the set of all subgraphs of Ω;*
- ◦ *the probability is given by $P(G) = n(G)/k$, where $n(G)$ is the number of nodes of the graph G.*

As an example of a random variable we consider $Y : \mathcal{F} \to \mathbb{R}$, $Y(G) =$ the total number of edges of the graph G. Since given $\mathcal{F}$, one can count the total number of edges of each subgraph, it follows that Y is $\mathcal{F}$-measurable, and hence it is a random variable.

1.5 Distribution Functions

Let X be a random variable on the probability space $(\Omega, \mathcal{F}, P)$. The *distribution function* of X is the function $F_X : \mathbb{R} \to [0, 1]$ defined by

$$F_X(x) = P(\omega; X(\omega) \le x).$$

It is worth observing that since X is a random variable, then the set $\{\omega; X(\omega) \le x\}$ belongs to the information set $\mathcal{F}$.

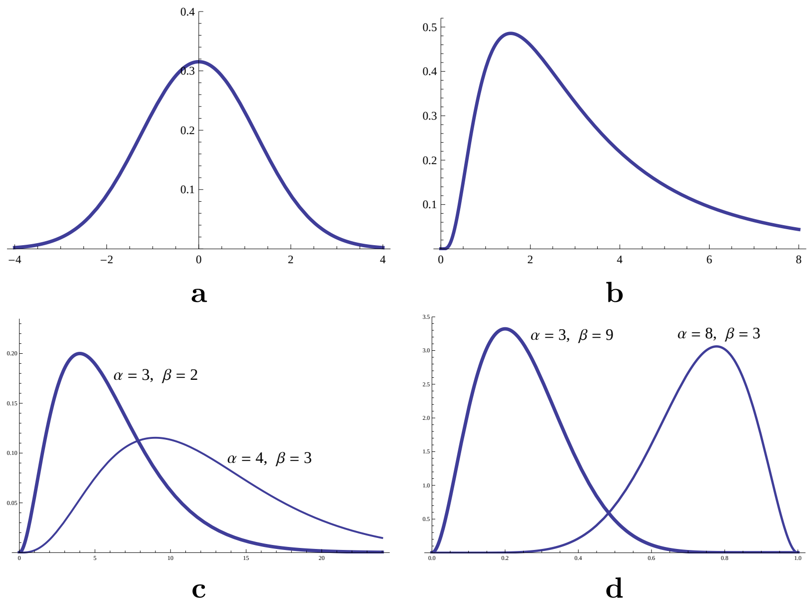


Figure 1.2: **a** *Normal distribution;* **b** *Log-normal distribution;* **c** *Gamma distributions;* **d** *Beta distributions.*

The distribution function is non-decreasing and satisfies the limits

$$\lim_{x\to-\infty} F_X(x) = 0, \qquad \lim_{x\to+\infty} F_X(x) = 1.$$

If we have

$$\frac{d}{dx}F_X(x) = p(x),$$

then we say that $p(x)$ is the *probability density function* of X. A useful property which follows from the Fundamental Theorem of Calculus is

$$P(a < X < b) = P(\omega; a < X(\omega) < b) = \int_a^b p(x)\,dx.$$

In the case of discrete random variables the aforementioned integral is replaced by the following sum

$$P(a < X < b) = \sum_{a<x<b} P(X = x).$$

For more details the reader is referred to a traditional probability book, such as Wackerly et. al. [6].

1.6 Basic Distributions

We shall recall a few basic distributions, which are most often seen in applications.

Normal distribution A random variable X is said to have a *normal distribution* if its probability density function is given by

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/(2\sigma^2)},$$

with μ and $\sigma > 0$ constant parameters, see Fig.1.2a. The mean and variance are given by

$$E[X] = \mu, \qquad Var[X] = \sigma^2.$$

If X has a normal distribution with mean μ and variance σ^2, we shall write

$$X \sim N(\mu, \sigma^2).$$

Exercise 1.6.1 *Let $\alpha, \beta \in \mathbb{R}$. Show that if X is normal distributed, with $X \sim N(\mu, \sigma^2)$, then $Y = \alpha X + \beta$ is also normal distributed, with $Y \sim N(\alpha\mu + \beta, \alpha^2\sigma^2)$.*

Log-normal distribution Let X be normally distributed with mean μ and variance σ^2. Then the random variable $Y = e^X$ is said to be *log-normal distributed*. The mean and variance of Y are given by

$$\begin{aligned} E[Y] &= e^{\mu + \frac{\sigma^2}{2}} \\ Var[Y] &= e^{2\mu+\sigma^2}(e^{\sigma^2} - 1). \end{aligned}$$

The density function of the log-normal distributed random variable Y is given by

$$p(x) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}}, \quad x > 0,$$

see Fig.1.2b.

Exercise 1.6.2 *Given that the moment generating function of a normally distributed random variable $X \sim N(\mu, \sigma^2)$ is $m(t) = E[e^{tX}] = e^{\mu t + t^2\sigma^2/2}$, show that*

(a) $E[Y^n] = e^{n\mu + n^2\sigma^2/2}$, where $Y = e^X$.

(b) Show that the mean and variance of the log-normal random variable $Y = e^X$ are

$$E[Y] = e^{\mu+\sigma^2/2}, \quad Var[Y] = e^{2\mu+\sigma^2}(e^{\sigma^2} - 1).$$

Gamma distribution A random variable X is said to have a *gamma distribution* with parameters $\alpha > 0$, $\beta > 0$ if its density function is given by

$$p(x) = \frac{x^{\alpha-1}e^{-x/\beta}}{\beta^\alpha \Gamma(\alpha)}, \quad x \geq 0,$$

where $\Gamma(\alpha)$ denotes the gamma function,[1] see Fig.1.2c. The mean and variance are

$$E[X] = \alpha\beta, \qquad Var[X] = \alpha\beta^2.$$

The case $\alpha = 1$ is known as the *exponential distribution*, see Fig.1.3a. In this case

$$p(x) = \frac{1}{\beta} e^{-x/\beta}, \quad x \geq 0.$$

The particular case when $\alpha = n/2$ and $\beta = 2$ becomes the χ^2-*distribution* with n degrees of freedom. This characterizes also a sum of n independent standard normal distributions.

Beta distribution A random variable X is said to have a *beta distribution* with parameters $\alpha > 0$, $\beta > 0$ if its probability density function is of the form

$$p(x) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)}, \quad 0 \leq x \leq 1,$$

where $B(\alpha, \beta)$ denotes the beta function.[2] See see Fig.1.2d for two particular density functions. In this case

$$E[X] = \frac{\alpha}{\alpha+\beta}, \quad Var[X] = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}.$$

Poisson distribution A discrete random variable X is said to have a *Poisson probability distribution* if

$$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}, \qquad k = 0, 1, 2, \dots,$$

with $\lambda > 0$ parameter, see Fig.1.3b. In this case $E[X] = \lambda$ and $Var[X] = \lambda$.

Pearson 5 distribution Let $\alpha, \beta > 0$. A random variable X with the density function

$$p(x) = \frac{1}{\beta\Gamma(\alpha)} \frac{e^{-\beta/x}}{(x/\beta)^{\alpha+1}}, \qquad x \geq 0$$

is said to have a Pearson 5 distribution[3] with positive parameters α and β. It can be shown that

$$E[X] = \begin{cases} \dfrac{\beta}{\alpha-1}, & \text{if } \alpha > 1 \\ \infty, & \text{otherwise,} \end{cases} \qquad Var(X) = \begin{cases} \dfrac{\beta^2}{(\alpha-1)^2(\alpha-2)}, & \text{if } \alpha > 2 \\ \infty, & \text{otherwise.} \end{cases}$$

---

[1]Recall the definition of the gamma function $\Gamma(\alpha) = \int_0^\infty y^{\alpha-1} e^{-y}\, dy$; if $\alpha = n$, integer, then $\Gamma(n) = (n-1)!$

[2]Two definition formulas for the beta functions are $B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} = \int_0^1 y^{\alpha-1}(1-y)^{\beta-1}\, dy$.

[3]The Pearson family of distributions was designed by Pearson between 1890 and 1895. There are several Pearson distributions, this one being distinguished by the number 5.

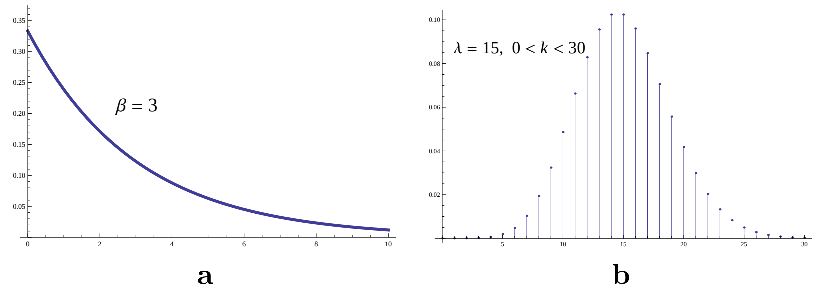


Figure 1.3: **a** *Exponential distribution;* **b** *Poisson distribution.*

The mode of this distribution is equal to $\dfrac{\beta}{\alpha+1}$.

The Inverse Gaussian distribution Let $\mu, \lambda > 0$. A random variable X has an inverse Gaussian distribution with parameters μ and λ if its density function is given by

$$p(x) = \frac{\lambda}{2\pi x^3} e^{-\frac{\lambda(x-\mu)^2}{2\mu^2 x}}, \qquad x > 0. \tag{1.6.1}$$

We shall write $X \sim IG(\mu, \lambda)$. Its mean, variance and mode are given by

$$E[X] = \mu, \qquad Var(X) = \frac{\mu^3}{\lambda}, \qquad Mode(X) = \mu\Big(\sqrt{1 + \frac{9\mu^2}{4\lambda^2}} - \frac{3\mu}{2\lambda}\Big).$$

This distribution will be used to model the time instance when a Brownian motion with drift exceeds a certain barrier for the first time.

1.7 Independent Random Variables

Roughly speaking, two random variables X and Y are independent if the occurrence of one of them does not change the probability density of the other. More precisely, if for any sets $A, B \subset \mathbb{R}$, the events

$$\{\omega; X(\omega) \in A\}, \qquad \{\omega; Y(\omega) \in B\}$$

are independent,[4] then X and Y are called *independent* random variables.

Proposition 1.7.1 *Let X and Y be independent random variables with probability density functions $p_X(x)$ and $p_Y(y)$. Then the joint probability density function of (X, Y) is given by $p_{X,Y}(x, y) = p_X(x)\, p_Y(y)$.*

[4] In Probability Theory two events A_1 and A_2 are called independent if $P(A_1 \cap A_2) = P(A_1)P(A_2)$.

Proof: Using the independence of sets, we have[5]

$$\begin{aligned} p_{X,Y}(x,y)\,dxdy &= P(x < X < x+dx,\, y < Y < y+dy) \\ &= P(x < X < x+dx)P(y < Y < y+dy) \\ &= p_X(x)\,dx\; p_Y(y)\,dy \\ &= p_X(x)p_Y(y)\,dxdy. \end{aligned}$$

Dropping the factor $dxdy$ yields the desired result. We note that the converse holds true. ■

1.8 Integration in Probability Measure

The notion of expectation is based on integration on measure spaces. In this section we recall briefly the definition of an integral with respect to the probability measure P.

Let $X : \Omega \to \mathbb{R}$ be a random variable on the probability space $(\Omega, \mathcal{F}, P)$. A partition $(\Omega_i)_{1\leq i\leq n}$ of Ω is a family of subsets $\Omega_i \subset \Omega$ satisfying

1. $\Omega_i \cap \Omega_j = \emptyset$, for $i \neq j$;
2. $\displaystyle\bigcup_i^n \Omega_i = \Omega$.

Each Ω_i is an event with the associated probability $P(\Omega_i)$. A *simple function* is a sum of characteristic functions $f = \sum_i^n c_i \chi_{\Omega_i}$. This means $f(\omega) = c_k$ for $\omega \in \Omega_k$. The integral of the simple function f is defined by

$$\boxed{\int_\Omega f\,dP = \sum_i^n c_i P(\Omega_i).}$$

If $X : \Omega \to \mathbb{R}$ is a random variable such that there is a sequence of simple functions $(f_n)_{n\geq 1}$ satisfying:

1. f_n *is fundamental in probability*: $\forall \epsilon > 0$ $\lim_{n,m\to\infty} P(\omega; |f_n(\omega) - f_m(\omega)| \geq \epsilon) \to 0$,
2. f_n *converges to* X *in probability*: $\forall \epsilon > 0$ $\lim_{n\to\infty} P(\omega; |f_n(\omega) - X(\omega)| \geq \epsilon) \to 0$

then the integral of X is defined as the following limit of integrals

$$\boxed{\int_\Omega X\,dP = \lim_{n\to\infty}\int_\Omega f_n\,dP.}$$

From now on, the integral notations $\int_\Omega X\,dP$ or $\int_\Omega X(\omega)\,dP(\omega)$ will be used interchangeably. In the rest of the chapter the integral notation will be used formally, without requiring a direct use of the previous definition.

---

[5]We are using the useful approximation $P(x < X < x+dx) = \int_x^{x+dx} p(u)\,du = p(x)dx$.

1.9 Expectation

A random variable $X : \Omega \to \mathbb{R}$ is called *integrable* if

$$\int_\Omega |X(\omega)|\, dP(\omega) = \int_{\mathbb{R}} |x| p(x)\, dx < \infty,$$

where $p(x)$ denotes the probability density function of X. The previous identity is based on changing the domain of integration from Ω to $\mathbb{R}$.

The *expectation* of an integrable random variable X is defined by

$$E[X] = \int_\Omega X(\omega)\, dP(\omega) = \int_{\mathbb{R}} x\, p(x)\, dx.$$

Customarily, the expectation of X is denoted by μ and it is also called the *mean*. In general, for any continuous[6] function $h : \mathbb{R} \to \mathbb{R}$, we have

$$E[h(X)] = \int_\Omega h\big(X(\omega)\big)\, dP(\omega) = \int_{\mathbb{R}} h(x)p(x)\, dx.$$

Proposition 1.9.1 *The expectation operator E is linear, i.e. for any integrable random variables X and Y*

1. $E[cX] = cE[X], \qquad \forall c \in \mathbb{R};$
2. $E[X + Y] = E[X] + E[Y].$

Proof: It follows from the fact that the integral is a linear operator. ∎

Proposition 1.9.2 *Let X and Y be two independent integrable random variables. Then*

$$E[XY] = E[X]E[Y].$$

Proof: This is a variant of Fubini's theorem, which in this case states that a double integral is a product of two simple integrals. Let p_X, p_Y, $p_{X,Y}$ denote the probability densities of X, Y and (X,Y), respectively. Since X and Y are independent, by Proposition 1.7.1 we have

$$E[XY] = \iint xy p_{X,Y}(x,y)\, dxdy = \int x p_X(x)\, dx \int y p_Y(y)\, dy = E[X]E[Y].$$

∎

[6]in general, measurable

1.10 Radon-Nikodym's Theorem

This section is concerned with existence and uniqueness results that will be useful later in defining conditional expectations. Since this section is rather theoretical, it can be skipped at a first reading.

Proposition 1.10.1 *Consider the probability space $(\Omega, \mathcal{F}, P)$, and let $\mathcal{G}$ be a σ-field included in $\mathcal{F}$. If X is a $\mathcal{G}$-predictable random variable such that*

$$\int_A X\,dP = 0 \qquad \forall A \in \mathcal{G},$$

then $X = 0$ a.s.

Proof: In order to show that $X = 0$ almost surely, it suffices to prove that $P\big(\omega; X(\omega) = 0\big) = 1$. We shall show first that X takes values as small as possible with probability one, i.e. $\forall \epsilon > 0$ we have $P(|X| < \epsilon) = 1$. To do this, let $A = \{\omega; X(\omega) \geq \epsilon\}$. Then

$$0 \leq P(X \geq \epsilon) = \int_A dP = \frac{1}{\epsilon}\int_A \epsilon\,dP \leq \frac{1}{\epsilon}\int_A X\,dP = 0,$$

and hence $P(X \geq \epsilon) = 0$. Similarly $P(X \leq -\epsilon) = 0$. Therefore

$$P(|X| < \epsilon) = 1 - P(X \geq \epsilon) - P(X \leq -\epsilon) = 1 - 0 - 0 = 1.$$

Taking $\epsilon \to 0$ leads to $P(|X| = 0) = 1$. This can be formalized as follows. Let $\epsilon = \frac{1}{n}$ and consider $B_n = \{\omega; |X(\omega)| \leq \epsilon\}$, with $P(B_n) = 1$. Then

$$P(X = 0) = P(|X| = 0) = P(\bigcap_{n=1}^{\infty} B_n) = \lim_{n\to\infty} P(B_n) = 1.$$

■

Corollary 1.10.2 *If X and Y are $\mathcal{G}$-predictable random variables such that*

$$\int_A X\,dP = \int_A Y\,dP \qquad \forall A \in \mathcal{G},$$

then $X = Y$ a.s.

Proof: Since $\int_A (X - Y)\,dP = 0$, $\forall A \in \mathcal{G}$, by Proposition 1.10.1 we have $X - Y = 0$ a.s. ■

Theorem 1.10.3 (Radon-Nikodym) *Let $(\Omega, \mathcal{F}, P)$ be a probability space and $\mathcal{G}$ be a σ-field included in $\mathcal{F}$. Then for any random variable X there is a $\mathcal{G}$-predictable random variable Y such that*

$$\int_A X\,dP = \int_A Y\,dP, \qquad \forall A \in \mathcal{G}. \tag{1.10.2}$$

We shall omit the proof but discuss a few aspects.

1. All σ-fields $\mathcal{G} \subset \mathcal{F}$ contain impossible and certain events $\emptyset, \Omega \in \mathcal{G}$. Making $A = \Omega$ yields

$$\int_\Omega X\, dP = \int_\Omega Y\, dP,$$

which is $E[X] = E[Y]$.

2. Radon-Nikodym's theorem states the existence of Y. In fact this is unique almost surely. In order to show that, assume there are two $\mathcal{G}$-predictable random variables Y_1 and Y_2 with the aforementioned property. Then from (1.10.2) yields

$$\int_A Y_1\, dP = \int_A Y_2\, dP, \qquad \forall A \in \mathcal{G}.$$

Applying Corollary (1.10.2) yields $Y_1 = Y_2$ a.s.

3. Since $E[X] = \int_\Omega X\, dP$ is the expectation of the random variable X, given the full knowledge $\mathcal{F}$, then the random variable Y plays the role of the expectation of X given the partial information $\mathcal{G}$. The next section will deal with this concept in detail.

1.11 Conditional Expectation

Let X be a random variable on the probability space $(\Omega, \mathcal{F}, P)$. Let $\mathcal{G}$ be a σ-field contained in $\mathcal{F}$. Since X is $\mathcal{F}$-predictable, the expectation of X, given the information $\mathcal{F}$ must be X itself. This shall be written as $E[X|\mathcal{F}] = X$ (for details see Example 1.11.3). It is natural to ask what is the expectation of X, given the information $\mathcal{G}$. This is a random variable denoted by $E[X|\mathcal{G}]$ satisfying the following properties:

1. $E[X|\mathcal{G}]$ is $\mathcal{G}$-predictable;
2. $\int_A E[X|\mathcal{G}]\, dP = \int_A X\, dP, \qquad \forall A \in \mathcal{G}.$

$E[X|\mathcal{G}]$ is called the *conditional expectation* of X given $\mathcal{G}$.

We owe a few explanations regarding the correctness of the aforementioned definition. The existence of the $\mathcal{G}$-predictable random variable $E[X|\mathcal{G}]$ is assured by the Radon-Nikodym theorem. The almost surely uniqueness is an application of Proposition (1.10.1) (see the discussion point 2 of section 1.10).

It is worth noting that the expectation of X, denoted by $E[X]$ is a number, while the conditional expectation $E[X|\mathcal{G}]$ is a random variable. When are they equal and what is their relationship? The answer is inferred by the following solved exercises.

Example 1.11.1 *Show that if $\mathcal{G} = \{\emptyset, \Omega\}$, then $E[X|\mathcal{G}] = E[X]$.*

Proof: We need to show that $E[X]$ satisfies conditions 1 and 2. The first one is obviously satisfied since any constant is $\mathcal{G}$-predictable. The latter condition is checked on each set of $\mathcal{G}$. We have

$$\begin{aligned}\int_\Omega X\,dP &= E[X] = E[X]\int_\Omega dP = \int_\Omega E[X]dP\\ \int_\emptyset X\,dP &= \int_\emptyset E[X]dP.\end{aligned}$$

■

Example 1.11.2 *Show that $E[E[X|\mathcal{G}]] = E[X]$, i.e. all conditional expectations have the same mean, which is the mean of X.*

Proof: Using the definition of expectation and taking $A = \Omega$ in the second relation of the aforementioned definition, yields

$$E[E[X|\mathcal{G}]] = \int_\Omega E[X|\mathcal{G}]\,dP = \int_\Omega X dP = E[X],$$

which ends the proof. ■

Example 1.11.3 *The conditional expectation of X given the total information $\mathcal{F}$ is the random variable X itself, i.e.*

$$E[X|\mathcal{F}] = X.$$

Proof: The random variables X and $E[X|\mathcal{F}]$ are both $\mathcal{F}$-predictable (from the definition of the random variable). From the definition of the conditional expectation we have

$$\int_A E[X|\mathcal{F}]\,dP = \int_A X\,dP, \quad \forall A \in \mathcal{F}.$$

Corollary (1.10.2) implies that $E[X|\mathcal{F}] = X$ almost surely. ■

General properties of the conditional expectation are stated below without proof. The proof involves more or less simple manipulations of integrals and can be taken as an exercise for the reader.

Proposition 1.11.4 *Let X and Y be two random variables on the probability space $(\Omega, \mathcal{F}, P)$. We have*

1. Linearity:

$$E[aX + bY|\mathcal{G}] = aE[X|\mathcal{G}] + bE[Y|\mathcal{G}], \qquad \forall a, b \in \mathbb{R};$$

2. Factoring out the predictable part:

$$E[XY|\mathcal{G}] = XE[Y|\mathcal{G}]$$

if X is $\mathcal{G}$-predictable. In particular, $E[X|\mathcal{G}] = X$.
3. Tower property:

$$E[E[X|\mathcal{G}]|\mathcal{H}] = E[X|\mathcal{H}], \textit{if } \mathcal{H} \subset \mathcal{G};$$

4. Positivity:

$$E[X|\mathcal{G}] \geq 0, \textit{if } X \geq 0;$$

5. Expectation of a constant is a constant:

$$E[c|\mathcal{G}] = c.$$

6. An independent condition drops out:

$$E[X|\mathcal{G}] = E[X],$$

if X is independent of $\mathcal{G}$.

Exercise 1.11.5 *Prove the property 3 (tower property) given in the previous proposition.*

Exercise 1.11.6 *Let X be a random variable on the probability space $(\Omega, \mathcal{F}, P)$, which is independent of the σ-field $\mathcal{G} \subset \mathcal{F}$. Consider the characteristic function of a set $A \subset \Omega$ defined by $\chi_A(\omega) = \begin{cases} 1, & \textit{if } \omega \in A \\ 0, & \textit{if } \omega \notin A \end{cases}$ Show the following:*

(*a*) χ_A *is $\mathcal{G}$-predictable for any $A \in \mathcal{G}$;*

(*b*) $P(A) = E[\chi_A]$;

(*c*) X *and* χ_A *are independent random variables*;

(*d*) $E[\chi_A X] = E[X]P(A)$ *for any* $A \in \mathcal{G}$;

(*e*) $E[X|\mathcal{G}] = E[X]$.

1.12 Inequalities of Random Variables

This section prepares the reader for the limits of sequences of random variables and limits of stochastic processes. We shall start with a classical inequality result regarding expectations:

Theorem 1.12.1 (Jensen's inequality) *Let $\varphi : \mathbb{R} \to \mathbb{R}$ be a convex function and let X be an integrable random variable on the probability space $(\Omega, \mathcal{F}, P)$. If $\varphi(X)$ is integrable, then*

$$\varphi(E[X]) \leq E[\varphi(X)]$$

almost surely (i.e. the inequality might fail on a set of probability zero).

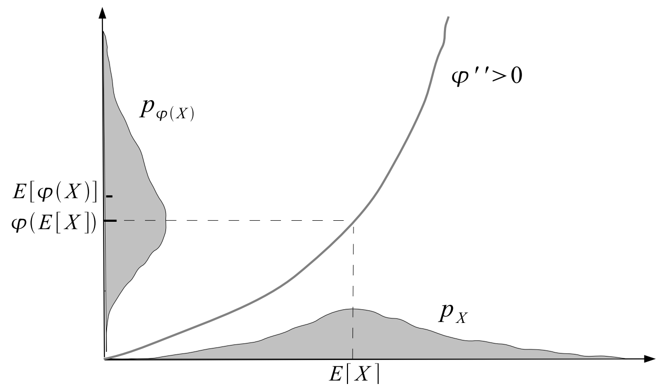


Figure 1.4: *Jensen's inequality* $\varphi(E[X]) < E[\varphi(X)]$ *for a convex function* φ.

Proof: We shall assume φ twice differentiable with φ'' continuous. Let $\mu = E[X]$. Expand φ in a Taylor series about μ and get

$$\varphi(x) = \varphi(\mu) + \varphi'(\mu)(x-\mu) + \frac{1}{2}\varphi''(\xi)(\xi-\mu)^2,$$

with ξ in between x and μ. Since φ is convex, $\varphi'' \geq 0$, and hence

$$\varphi(x) \geq \varphi(\mu) + \varphi'(\mu)(x-\mu),$$

which means the graph of $\varphi(x)$ is above the tangent line at $\big(x, \varphi(x)\big)$. Replacing x by the random variable X, and taking the expectation yields

$$\begin{aligned} E[\varphi(X)] &\geq E[\varphi(\mu) + \varphi'(\mu)(X-\mu)] = \varphi(\mu) + \varphi'(\mu)(E[X]-\mu) \\ &= \varphi(\mu) = \varphi(E[X]), \end{aligned}$$

which proves the result. ∎

Fig.1.4 provides a graphical interpretation of Jensen's inequality. If the distribution of X is symmetric, then the distribution of $\varphi(X)$ is skewed, with $\varphi(E[X]) < E[\varphi(X)]$.

It is worth noting that the inequality is reversed for φ concave. We shall next present a couple of applications.

A random variable $X : \Omega \to \mathbb{R}$ is called *square integrable* if

$$E[X^2] = \int_\Omega |X(\omega)|^2 \, dP(\omega) = \int_{\mathbb{R}} x^2 p(x) \, dx < \infty.$$

Application 1.12.2 *If* X *is a square integrable random variable, then it is integrable.*

Proof: Jensen's inequality with $\varphi(x) = x^2$ becomes

$$E[X]^2 \leq E[X^2].$$

Since the right side is finite, it follows that $E[X] < \infty$, so X is integrable. ∎

Application 1.12.3 *If $m_X(t)$ denotes the moment generating function of the random variable X with mean μ, then*

$$m_X(t) \geq e^{t\mu}.$$

Proof: Applying Jensen inequality with the convex function $\varphi(x) = e^x$ yields

$$e^{E[X]} \leq E[e^X].$$

Substituting tX for X yields

$$e^{E[tX]} \leq E[e^{tX}]. \tag{1.12.3}$$

Using the definition of the moment generating function $m_X(t) = E[e^{tX}]$ and that $E[tX] = tE[X] = t\mu$, then (1.12.3) leads to the desired inequality. ∎

The *variance* of a square integrable random variable X is defined by

$$Var(X) = E[X^2] - E[X]^2.$$

By Application 1.12.2 we have $Var(X) \geq 0$, so there is a constant $\sigma_X > 0$, called *standard deviation*, such that

$$\sigma_X^2 = Var(X).$$

Exercise 1.12.4 *Prove the following identity:*

$$Var[X] = E[(X - E[X])^2].$$

Exercise 1.12.5 *Prove that a non-constant random variable has a non-zero standard deviation.*

Exercise 1.12.6 *Prove the following extension of Jensen's inequality: If φ is a convex function, then for any σ-field $\mathcal{G} \subset \mathcal{F}$ we have*

$$\varphi(E[X|\mathcal{G}]) \leq E[\varphi(X)|\mathcal{G}].$$

Exercise 1.12.7 *Show the following:*

(*a*) $|E[X]| \leq E[|X|]$;

(*b*) $|E[X|\mathcal{G}]| \leq E[|X|\ |\mathcal{G}]$, *for any σ-field $\mathcal{G} \subset \mathcal{F}$;*

(*c*) $|E[X]|^r \leq E[|X|^r]$, *for $r \geq 1$;*

(*d*) $|E[X|\mathcal{G}]|^r \leq E[|X|^r\ |\mathcal{G}]$, *for any σ-field $\mathcal{G} \subset \mathcal{F}$ and $r \geq 1$.*

Theorem 1.12.8 (Markov's inequality) *For any $\lambda, p > 0$, we have the following inequality:*

$$P(\omega; |X(\omega)| \geq \lambda) \leq \frac{1}{\lambda^p} E[|X|^p].$$

Proof: Let $A = \{\omega; |X(\omega)| \geq \lambda\}$. Then

$$\begin{aligned} E[|X|^p] &= \int_\Omega |X(\omega)|^p \, dP(\omega) \geq \int_A |X(\omega)|^p \, dP(\omega) \geq \int_A \lambda^p \, dP(\omega) \\ &= \lambda^p \int_A dP(\omega) = \lambda^p P(A) = \lambda^p P(|X| \geq \lambda). \end{aligned}$$

Dividing by λ^p leads to the desired result. ■

Theorem 1.12.9 (Tchebychev's inequality) *If X is a random variable with mean μ and variance σ^2, then*

$$P(\omega; |X(\omega) - \mu| \geq \lambda) \leq \frac{\sigma^2}{\lambda^2}.$$

Proof: Let $A = \{\omega; |X(\omega) - \mu| \geq \lambda\}$. Then

$$\begin{aligned} \sigma^2 &= Var(X) = E[(X-\mu)^2] = \int_\Omega (X-\mu)^2 \, dP \geq \int_A (X-\mu)^2 \, dP \\ &\geq \lambda^2 \int_A dP = \lambda^2 P(A) = \lambda^2 P(\omega; |X(\omega) - \mu| \geq \lambda). \end{aligned}$$

Dividing by λ^2 leads to the desired inequality. ■

The next result deals with exponentially decreasing bounds on tail distributions.

Theorem 1.12.10 (Chernoff bounds) *Let X be a random variable. Then for any $\lambda > 0$ we have*

1. $P(X \geq \lambda) \leq \dfrac{E[e^{tX}]}{e^{\lambda t}}$, $\forall t > 0$;
2. $P(X \leq \lambda) \leq \dfrac{E[e^{tX}]}{e^{\lambda t}}$, $\forall t < 0$.

Proof: 1. Let $t > 0$ and denote $Y = e^{tX}$. By Markov's inequality

$$P(Y \geq e^{\lambda t}) \leq \frac{E[Y]}{e^{\lambda t}}.$$

Then we have

$$\begin{aligned} P(X \geq \lambda) &= P(tX \geq \lambda t) = P(e^{tX} \geq e^{\lambda t}) \\ &= P(Y \geq e^{\lambda t}) \leq \frac{E[Y]}{e^{\lambda t}} = \frac{E[e^{tX}]}{e^{\lambda t}}. \end{aligned}$$

2. The case $t < 0$ is similar. ■

In the following we shall present an application of the Chernoff bounds for the normal distributed random variables.

Let X be a random variable normally distributed with mean μ and variance σ^2. It is known that its moment generating function is given by

$$m(t) = E[e^{tX}] = e^{\mu t + \frac{1}{2}t^2\sigma^2}.$$

Using the first Chernoff bound we obtain

$$P(X \geq \lambda) \quad \leq \quad \frac{m(t)}{e^{\lambda t}} = e^{(\mu-\lambda)t + \frac{1}{2}t^2\sigma^2}, \forall t > 0,$$

which implies

$$P(X \geq \lambda) \quad \leq \quad e^{\min\limits_{t>0}[(\mu - \lambda)t + \frac{1}{2}t^2\sigma^2]}.$$

It is easy to see that the quadratic function $f(t) = (\mu - \lambda)t + \frac{1}{2}t^2\sigma^2$ has the minimum value reached for $t = \dfrac{\lambda - \mu}{\sigma^2}$. Since $t > 0$, λ needs to satisfy $\lambda > \mu$. Then

$$\min_{t>0} f(t) = f\Big(\frac{\lambda - \mu}{\sigma^2}\Big) = -\frac{(\lambda - \mu)^2}{2\sigma^2}.$$

Substituting into the previous formula, we obtain the following result:

Proposition 1.12.11 *If X is a normally distributed variable, with $X \sim N(\mu, \sigma^2)$, then for any $\lambda > \mu$*

$$P(X \geq \lambda) \leq e^{-\dfrac{(\lambda-\mu)^2}{2\sigma^2}}.$$

Exercise 1.12.12 *Let X be a Poisson random variable with mean $\lambda > 0$.*

(a) Show that the moment generating function of X is $m(t) = e^{\lambda(e^t - 1)}$;

(b) Use a Chernoff bound to show that

$$P(X \geq k) \leq e^{\lambda(e^t-1)-tk}, \qquad t > 0.$$

Markov's, Tchebychev's and Chernoff's inequalities will be useful later when computing limits of random variables.

The next inequality is called Tchebychev's inequality for monotone sequences of numbers.

Lemma 1.12.13 *Let* (a_i) *and* (b_i) *be two sequences of real numbers such that either*

$$a_1 \le a_2 \le \cdots \le a_n, \qquad b_1 \le b_2 \le \cdots \le b_n$$

or

$$a_1 \ge a_2 \ge \cdots \ge a_n, \qquad b_1 \ge b_2 \ge \cdots \ge b_n$$

If (λ_i) *is a sequence of non-negative numbers such that* $\sum_{i=1}^{n} \lambda_i = 1$*, then*

$$\Big(\sum_{i=1}^{n} \lambda_i a_i\Big)\Big(\sum_{i=1}^{n} \lambda_i b_i\Big) \le \sum_{i=1}^{n} \lambda_i a_i b_i.$$

Proof: Since the sequences (a_i) and (b_i) are either both increasing or both decreasing

$$(a_i - a_j)(b_i - b_j) \ge 0.$$

Multiplying by the positive quantity $\lambda_i \lambda_j$ and summing over i and j we get

$$\sum_{i,j} \lambda_i \lambda_j (a_i - a_j)(b_i - b_j) \ge 0.$$

Expanding yields

$$\Big(\sum_j \lambda_j\Big)\Big(\sum_i \lambda_i a_i b_i\Big) - \Big(\sum_i \lambda_i a_i\Big)\Big(\sum_j \lambda_j b_j\Big) - \Big(\sum_j \lambda_j a_j\Big)\Big(\sum_i \lambda_i b_i\Big)$$
$$+\Big(\sum_i \lambda_i\Big)\Big(\sum_j \lambda_j a_j b_j\Big) \ge 0.$$

Using $\sum_j \lambda_j = 1$ the expression becomes

$$\sum_i \lambda_i a_i b_i \ge \Big(\sum_i \lambda_i a_i\Big)\Big(\sum_j \lambda_j b_j\Big),$$

which ends the proof. ■

Next we present a meaningful application of the previous inequality.

Proposition 1.12.14 *Let* X *be a random variable and* f *and* g *be two functions, both increasing or both decreasing. Then*

$$E[f(X)g(X)] \ge E[f(X)]E[g(X)]. \qquad (1.12.4)$$

Proof: If X is a discrete random variable, with outcomes $\{x_1, \cdots, x_n\}$, inequality (1.12.4) becomes

$$\sum_j f(x_j)g(x_j)p(x_j) \geq \sum_j f(x_j)p(x_j) \sum_j g(x_j)p(x_j),$$

where $p(x_j) = P(X = x_j)$. Denoting $a_j = f(x_j)$, $b_j = g(x_j)$, and $\lambda_j = p(x_j)$, the inequality transforms into

$$\sum_j a_j b_j \lambda_j \geq \sum_j a_j \lambda_j \sum_j b_j \lambda_j,$$

which holds true by Lemma 1.12.13.

If X is a continuous random variable with the density function $p : I \to \mathbb{R}$, the inequality (1.12.4) can be written in the integral form

$$\int_I f(x)g(x)p(x)\,dx \geq \int_I f(x)p(x)\,dx \int_I g(x)p(x)\,dx. \tag{1.12.5}$$

Let $x_0 < x_1 < \cdots < x_n$ be a partition of the interval I, with $\Delta x = x_{k+1} - x_k$. Using Lemma 1.12.13 we obtain the following inequality between Riemann sums

$$\sum_j f(x_j)g(x_j)p(x_j)\Delta x \geq \Big(\sum_j f(x_j)p(x_j)\Delta x\Big)\Big(\sum_j g(x_j)p(x_j)\Delta x\Big),$$

where $a_j = f(x_j)$, $b_j = g(x_j)$, and $\lambda_j = p(x_j)\Delta x$. Taking the limit $\|\Delta x\| \to 0$ we obtain (1.12.5), which leads to the desired result. ■

Exercise 1.12.15 *Show the following inequalities:*

(*a*) $E[X^2] \geq E[X]^2$;

(*b*) $E[X \sinh(X)] \geq E[X]E[\sinh(X)]$;

(*c*) $E[X^6] \geq E[X]E[X^5]$;

(*d*) $E[X^6] \geq E[X^3]^2$.

Exercise 1.12.16 *For any $n, k \geq 1$, show that*

$$E[X^{2(n+k+1)}] \geq E[X^{2k+1}]E[X^{2n+1}].$$

1.13 Limits of Sequences of Random Variables

Consider a sequence $(X_n)_{n\geq 1}$ of random variables defined on the probability space $(\Omega, \mathcal{F}, P)$. There are several ways of making sense of the limit expression $X = \lim\limits_{n\to\infty} X_n$, and they will be discussed in the following sections.

Almost Certain Limit

The sequence X_n converges *almost certainly* to X, if for all states of the world ω, except a set of probability zero, we have

$$\lim_{n\to\infty} X_n(\omega) = X(\omega).$$

More precisely, this means

$$P\Big(\omega; \lim_{n\to\infty} X_n(\omega) = X(\omega)\Big) = 1,$$

and we shall write $\text{ac-}\lim\limits_{n\to\infty} X_n = X$. An important example where this type of limit occurs is the Strong Law of Large Numbers:

If X_n is a sequence of independent and identically distributed random variables with the same mean μ, then $ac\text{-}\lim\limits_{n\to\infty} \dfrac{X_1 + \cdots + X_n}{n} = \mu$.

It is worth noting that this type of convergence is also known under the name of *strong convergence*. This is the reason that the aforementioned theorem bares its name.

Example 1.13.1 *Let $\Omega = \{H, T\}$ be the sample space obtained when a coin is flipped. Consider the random variables $X_n : \Omega \to \{0, 1\}$, where X_n denotes the number of heads obtained at the n-th flip. Obviously, X_n are i.i.d., with the distribution given by $P(X_n = 0) = P(X_n = 1) = 1/2$, and the mean $E[X_n] = 0 \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} = \frac{1}{2}$. Then $X_1 + \cdots + X_n$ is the number of heads obtained after n flips of the coin. By the law of large numbers, $\frac{1}{n}(X_1 + \cdots + X_n)$ tends to $1/2$ strongly, as $n \to \infty$.*

Mean Square Limit

Another possibility of convergence is to look at the mean square deviation of X_n from X. We say that X_n converges to X in the *mean square* if

$$\lim_{n\to\infty} E[(X_n - X)^2] = 0.$$

More precisely, this should be interpreted as

$$\lim_{n\to\infty} \int_\Omega \big(X_n(\omega) - X(\omega)\big)^2 \, dP(\omega) = 0.$$

This limit will be abbreviated by $\text{ms-}\lim\limits_{n\to\infty} X_n = X$. The mean square convergence is useful when defining the Ito integral.

Example 1.13.1 *Consider a sequence X_n of random variables such that there is a constant k with $E[X_n] \to k$ and $Var(X_n) \to 0$ as $n \to \infty$. Show that $ms\text{-}\lim\limits_{n\to\infty} X_n = k$.*

Proof: Since we have

$$\begin{aligned} E[|X_n - k|^2] &= E[X_n^2 - 2kX_n + k^2] = E[X_n^2] - 2kE[X_n] + k^2 \\ &= \big(E[X_n^2] - E[X_n]^2\big) + \big(E[X_n]^2 - 2kE[X_n] + k^2\big) \\ &= Var(X_n) + \big(E[X_n] - k\big)^2, \end{aligned}$$

the right side tends to 0 when taking the limit $n \to \infty$. ∎

Exercise 1.13.2 *Show the following relation*

$$E[(X-Y)^2] = Var[X] + Var[Y] + \big(E[X] - E[Y]\big)^2 - 2Cov(X,Y).$$

Exercise 1.13.3 *If X_n tends to X in mean square, with $E[X^2] < \infty$, show that:*

(*a*) $E[X_n] \to E[X]$ *as* $n \to \infty$;

(*b*) $E[X_n^2] \to E[X^2]$ *as* $n \to \infty$;

(*c*) $Var[X_n] \to Var[X]$ *as* $n \to \infty$;

(*d*) $Cov(X_n, X) \to Var[X]$ *as* $n \to \infty$.

Exercise 1.13.4 *If X_n tends to X in mean square, show that $E[X_n|\mathcal{H}]$ tends to $E[X|\mathcal{H}]$ in mean square.*

Limit in Probability or Stochastic Limit

The random variable X is the *stochastic limit* of X_n if for n large enough the probability of deviation from X can be made smaller than any arbitrary ϵ. More precisely, for any $\epsilon > 0$

$$\lim_{n\to\infty} P\big(\omega; |X_n(\omega) - X(\omega)| \le \epsilon\big) = 1.$$

This can be written also as

$$\lim_{n\to\infty} P\big(\omega; |X_n(\omega) - X(\omega)| > \epsilon\big) = 0.$$

This limit is denoted by $\operatorname*{st-lim}_{n\to\infty} X_n = X$.

It is worth noting that both almost certain convergence and convergence in mean square imply the stochastic convergence. Hence, the stochastic convergence is weaker than the aforementioned two convergence cases. This is the reason that it is also called the *weak convergence*. One application is the Weak Law of Large Numbers:

If $X_1, X_2, \dots$ are identically distributed with expected value μ and if any finite number of them are independent, then $\operatorname*{st-lim}_{n\to\infty} \dfrac{X_1 + \cdots + X_n}{n} = \mu$.

Proposition 1.13.5 *The convergence in the mean square implies the stochastic convergence.*

Proof: Let $\text{ms-}\lim_{n\to\infty} Y_n = Y$. Let $\epsilon > 0$ be arbitrarily fixed. Applying Markov's inequality with $X = Y_n - Y$, $p = 2$ and $\lambda = \epsilon$, yields

$$0 \le P(|Y_n - Y| \ge \epsilon) \le \frac{1}{\epsilon^2} E[|Y_n - Y|^2].$$

The right side tends to 0 as $n \to \infty$. Applying the Squeeze Theorem we obtain

$$\lim_{n\to\infty} P(|Y_n - Y| \ge \epsilon) = 0,$$

which means that Y_n converges stochastically to Y. ■

Example 1.13.6 *Let X_n be a sequence of random variables such that $E[|X_n|] \to 0$ as $n \to \infty$. Prove that $st\text{-}\lim_{n\to\infty} X_n = 0$.*

Proof: Let $\epsilon > 0$ be arbitrarily fixed. We need to show

$$\lim_{n\to\infty} P\big(\omega; |X_n(\omega)| \ge \epsilon\big) = 0. \tag{1.13.6}$$

From Markov's inequality (see Exercise 1.12.8) we have

$$0 \le P\big(\omega; |X_n(\omega)| \ge \epsilon\big) \le \frac{E[|X_n|]}{\epsilon}.$$

Using Squeeze Theorem we obtain (1.13.6). ■

Remark 1.13.7 *The conclusion still holds true even in the case when there is a $p > 0$ such that $E[|X_n|^p] \to 0$ as $n \to \infty$.*

Limit in Distribution

We say the sequence X_n converges *in distribution* to X if for any continuous bounded function $\varphi(x)$ we have

$$\lim_{n\to\infty} \varphi(X_n) = \varphi(X).$$

This type of limit is even weaker than the stochastic convergence, i.e. it is implied by it.

An application of the limit in distribution is obtained if we consider $\varphi(x) = e^{itx}$. In this case, if X_n converges in distribution to X, then the characteristic function of X_n converges to the characteristic function of X. In particular, the probability density of X_n approaches the probability density of X.

It can be shown that the convergence in distribution is equivalent with

$$\lim_{n\to\infty} F_n(x) = F(x),$$

whenever F is continuous at x, where F_n and F denote the distribution functions of X_n and X, respectively. This is the reason that this convergence bares its name.

Remark 1.13.8 *The almost certain convergence implies the stochastic convergence, and the stochastic convergence implies the limit in distribution. The proof of these statements is beyound the goal of this book. The interested reader can consult a graduate text in probability theory.*

1.14 Properties of Limits

Lemma 1.14.1 *If* $ms\text{-}\lim\limits_{n\to\infty} X_n = 0$ *and* $ms\text{-}\lim\limits_{n\to\infty} Y_n = 0$*, then*

1. $$ms\text{-}\lim_{n\to\infty} (X_n + Y_n) = 0$$

2. $$ms\text{-}\lim_{n\to\infty} (X_n Y_n) = 0.$$

Proof: Since ms-$\lim\limits_{n\to\infty} X_n = 0$, then $\lim\limits_{n\to\infty} E[X_n^2] = 0$. Applying the Squeeze Theorem to the inequality[7]

$$0 \le E[X_n]^2 \le E[X_n^2]$$

yields $\lim\limits_{n\to\infty} E[X_n] = 0$. Then

$$\begin{aligned}
\lim_{n\to\infty} Var[X_n] &= \lim_{n\to\infty} \Big(E[X_n^2] - \lim_{n\to\infty} E[X_n]^2 \Big) \\
&= \lim_{n\to\infty} E[X_n^2] - \lim_{n\to\infty} E[X_n]^2 \\
&= 0.
\end{aligned}$$

Similarly, we have $\lim\limits_{n\to\infty} E[Y_n^2] = 0$, $\lim\limits_{n\to\infty} E[Y_n] = 0$ and $\lim\limits_{n\to\infty} Var[Y_n] = 0$. Then $\lim\limits_{n\to\infty} \sigma_{X_n} = \lim\limits_{n\to\infty} \sigma_{Y_n} = 0$. Using the correlation definition formula of two random variables X_n and Y_n

$$Corr(X_n, Y_n) = \frac{Cov(X_n, Y_n)}{\sigma_{X_n}\sigma_{Y_n}},$$

and the fact that $|Corr(X_n, Y_n)| \le 1$, yields

$$0 \le |Cov(X_n, Y_n)| \le \sigma_{X_n}\sigma_{Y_n}.$$

Since $\lim\limits_{n\to\infty} \sigma_{X_n}\sigma_{X_n} = 0$, from the Squeeze Theorem it follows that

$$\lim_{n\to\infty} Cov(X_n, Y_n) = 0.$$

Taking $n \to \infty$ in the relation

$$Cov(X_n, Y_n) = E[X_n Y_n] - E[X_n]E[Y_n]$$

[7] This follows from the fact that $Var[X_n] \ge 0$.

yields $\lim_{n\to\infty} E[X_n Y_n] = 0$. Using the previous relations, we have

$$\begin{aligned}
\lim_{n\to\infty} E[(X_n+Y_n)^2] &= \lim_{n\to\infty} E[X_n^2 + 2X_nY_n + Y_n^2] \\
&= \lim_{n\to\infty} E[X_n^2] + 2\lim_{n\to\infty} E[X_nY_n] + \lim_{n\to\infty} E[Y_n^2] \\
&= 0,
\end{aligned}$$

which means $\text{ms-}\lim_{n\to\infty}(X_n+Y_n) = 0$. ■

Proposition 1.14.2 *If the sequences of random variables X_n and Y_n converge in the mean square, then*

$$\begin{aligned}
&1. \quad ms\text{-}\lim_{n\to\infty}(X_n+Y_n) = ms\text{-}\lim_{n\to\infty} X_n + ms\text{-}\lim_{n\to\infty} Y_n \\
&2. \quad ms\text{-}\lim_{n\to\infty}(cX_n) = c\cdot ms\text{-}\lim_{n\to\infty} X_n, \quad \forall c\in\mathbb{R}.
\end{aligned}$$

Proof: 1. Let $\text{ms-}\lim_{n\to\infty} X_n = L$ and $\text{ms-}\lim_{n\to\infty} Y_n = M$. Consider the sequences $X_n' = X_n - L$ and $Y_n' = Y_n - M$. Then $\text{ms-}\lim_{n\to\infty} X_n' = 0$ and $\text{ms-}\lim_{n\to\infty} Y_n' = 0$. Applying Lemma 1.14.1 yields

$$\text{ms-}\lim_{n\to\infty}(X_n' + Y_n') = 0.$$

This is equivalent with

$$\text{ms-}\lim_{n\to\infty}(X_n - L + Y_n - M) = 0,$$

which becomes

$$\text{ms-}\lim_{n\to\infty}(X_n + Y_n) = L + M.$$

■

1.15 Stochastic Processes

A *stochastic process* on the probability space $(\Omega, \mathcal{F}, P)$ is a family of random variables X_t parameterized by $t \in \mathbf{T}$, where $\mathbf{T} \subset \mathbb{R}$. If $\mathbf{T}$ is an interval we say that X_t is a stochastic process in *continuous time*. If $\mathbf{T} = \{1, 2, 3, \dots\}$ we shall say that X_t is a stochastic process in *discrete time*. The latter case describes a sequence of random variables. The aforementioned types of convergence can be easily extended to continuous time. For instance, X_t converges in the strong sense to X as $t \to \infty$ if

$$P\Big(\omega; \lim_{t\to\infty} X_t(\omega) = X(\omega)\Big) = 1.$$

The evolution in time of a given state of the world $\omega \in \Omega$ given by the function $t \longmapsto X_t(\omega)$ is called a *path* or *realization* of X_t. The study of stochastic processes using computer simulations is based on retrieving information about the process X_t given a large number of it realizations.

Consider that all the information accumulated until time t is contained by the σ-field $\mathcal{F}_t$. This means that $\mathcal{F}_t$ contains the information of which events have already occurred until time t, and which did not. Since the information is growing in time, we have

$$\mathcal{F}_s \subset \mathcal{F}_t \subset \mathcal{F}$$

for any $s, t \in \mathbf{T}$ with $s \leq t$. The family $\mathcal{F}_t$ is called a *filtration*.

A stochastic process X_t is called *adapted* to the filtration $\mathcal{F}_t$ if X_t is $\mathcal{F}_t$- predictable, for any $t \in \mathbf{T}$.

Example 1.15.1 *Here there are a few examples of filtrations:*

1. $\mathcal{F}_t$ represents the information about the evolution of a stock until time t, with $t > 0$.

2. $\mathcal{F}_t$ represents the information about the evolution of a Black-Jack game until time t, with $t > 0$.

Example 1.15.2 *If X is a random variable, consider the conditional expectation*

$$X_t = E[X|\mathcal{F}_t].$$

From the definition of conditional expectation, the random variable X_t is $\mathcal{F}_t$-predictable, and can be regarded as the measurement of X at time t using the information $\mathcal{F}_t$. If the accumulated knowledge $\mathcal{F}_t$ increases and eventually equals the σ-field $\mathcal{F}$, then $X = E[X|\mathcal{F}]$, i.e. we obtain the entire random variable. The process X_t is adapted to $\mathcal{F}_t$.

Example 1.15.3 *Don Joe is asking a doctor how long he still has to live. The age at which he will pass away is a random variable, denoted by X. Given his medical condition today, which is contained in $\mathcal{F}_t$, the doctor infers that Mr. Joe will die at the age of $X_t = E[X|\mathcal{F}_t]$. The stochastic process X_t is adapted to the medical knowledge $\mathcal{F}_t$.*

We shall define next an important type of stochastic process.[8]

Definition 1.15.4 *A process X_t, $t \in \mathbf{T}$, is called a martingale with respect to the filtration $\mathcal{F}_t$ if*

1. X_t is integrable for each $t \in \mathbf{T}$;

2. X_t is adapted to the filtration $\mathcal{F}_t$;

3. $X_s = E[X_t|\mathcal{F}_s]$, $\forall s < t$.

Remark 1.15.5 *The first condition states that the unconditional forecast is finite $E[|X_t|] = \int_\Omega |X_t|\,dP < \infty$. Condition 2 says that the value X_t is known, given the information set $\mathcal{F}_t$. This can be also stated by saying that X_t is $\mathcal{F}_t$-predictable. The third relation asserts that the best forecast of unobserved future values is the last observation on X_t.*

[8] The concept of martingale was introduced by Lévy in 1934.

Remark 1.15.6 *If the third condition is replaced by*

$3'$. $\quad X_s \leq E[X_t|\mathcal{F}_s]$, $\forall s \leq t$

then X_t is called a submartingale; and if it is replaced by

$3''$. $\quad X_s \geq E[X_t|\mathcal{F}_s]$, $\forall s \leq t$

then X_t is called a supermartingale.

It is worth noting that X_t is a submartingale if and only if $-X_t$ is a supermartingale.

Example 1.15.1 *Let X_t denote Mr. Li Zhu's salary after t years of work at the same company. Since X_t is known at time t and it is bounded above, as all salaries are, then the first two conditions hold. Being honest, Mr. Zhu expects today that his future salary will be the same as today's, i.e. $X_s = E[X_t|\mathcal{F}_s]$, for $s < t$. This means that X_t is a martingale.*

If Mr. Zhu is optimistic and believes as of today that his future salary will increase, then X_t is a submartingale.

Exercise 1.15.7 *If X is an integrable random variable on $(\Omega, \mathcal{F}, P)$, and $\mathcal{F}_t$ is a filtration. Prove that $X_t = E[X|\mathcal{F}_t]$ is a martingale.*

Exercise 1.15.8 *Let X_t and Y_t be martingales with respect to the filtration $\mathcal{F}_t$. Show that for any $a, b, c \in \mathbb{R}$ the process $Z_t = aX_t + bY_t + c$ is a $\mathcal{F}_t$-martingale.*

Exercise 1.15.9 *Let X_t and Y_t be martingales with respect to the filtration $\mathcal{F}_t$.*

(*a*) *Is the process X_tY_t always a martingale with respect to $\mathcal{F}_t$?*

(*b*) *What about the processes X_t^2 and Y_t^2?*

Exercise 1.15.10 *Two processes X_t and Y_t are called conditionally uncorrelated, given $\mathcal{F}_t$, if*

$$E[(X_t - X_s)(Y_t - Y_s)|\mathcal{F}_s] = 0, \qquad \forall\, 0 \leq s < t < \infty.$$

Let X_t and Y_t be martingale processes. Show that the process $Z_t = X_tY_t$ is a martingale if and only if X_t and Y_t are conditionally uncorrelated. Assume that X_t, Y_t and Z_t are integrable.

In the following, if X_t is a stochastic process, the minimum amount of information resulted from knowing the process X_t until time t is denoted by $\mathcal{F}_t = \sigma(X_s; s \leq t)$. In the case of a discrete process, we have $\mathcal{F}_n = \sigma(X_k; k \leq n)$.

Exercise 1.15.11 *Let X_n, $n \geq 0$ be a sequence of integrable independent random variables, with $E[X_n] < \infty$, for all $n \geq 0$. Let $S_0 = X_0$, $S_n = X_0 + \cdots + X_n$. Show the following:*

(*a*) *$S_n - E[S_n]$ is an $\mathcal{F}_n$-martingale.*

(*b*) *If $E[X_n] = 0$ and $E[X_n^2] < \infty$, $\forall n \geq 0$, then $S_n^2 - Var(S_n)$ is an $\mathcal{F}_n$-martingale.*

(*c*) *If $E[X_n] \geq 0$, then S_n is an $\mathcal{F}_n$-submartingale.*

Exercise 1.15.12 *Let* X_n, $n \geq 0$ *be a sequence of independent, integrable random variables such that* $E[X_n] = 1$ *for* $n \geq 0$. *Prove that* $P_n = X_0 \cdot X_1 \cdot \dots \cdot X_n$ *is an* $\mathcal{F}_n$*-martingale.*

Exercise 1.15.13 (*a*) *Let* X *be a normally distributed random variable with mean* $\mu \neq 0$ *and variance* σ^2. *Prove that there is a unique* $\theta \neq 0$ *such that* $E[e^{\theta X}] = 1$.

(*b*) *Let* $(X_i)_{i \geq 0}$ *be a sequence of identically normally distributed random variables with mean* $\mu \neq 0$. *Consider the sum* $S_n = \sum_{j=0}^n X_j$. *Show that* $Z_n = e^{\theta S_n}$ *is a martingale, with* θ *defined in part* (*a*).

In section 9.1 we shall encounter several processes which are martingales.

Chapter 2

Useful Stochastic Processes

This chapter deals with the most common used stochastic processes and their basic properties. The two main basic processes are the Brownian motion and the Poisson process. The other processes described in this chapter are derived from the previous two.

2.1 The Brownian Motion

The observation made first by the botanist Robert Brown in 1827, that small pollen grains suspended in water have a very irregular and unpredictable state of motion, led to the definition of the Brownian motion, which is formalized in the following:

Definition 2.1.1 *A Brownian motion process is a stochastic process B_t, $t \geq 0$, which satisfies*

1. The process starts at the origin, $B_0 = 0$;

2. B_t has stationary, independent increments;

3. The process B_t is continuous in t;

4. The increments $B_t - B_s$ are normally distributed with mean zero and variance $|t-s|$,

$$B_t - B_s \sim N(0, |t-s|).$$

The process $X_t = x + B_t$ has all the properties of a Brownian motion that starts at x. Since $B_t - B_s$ is stationary, its distribution function depends only on the time interval $t - s$, i.e.

$$P(B_{t+s} - B_s \leq a) = P(B_t - B_0 \leq a) = P(B_t \leq a).$$

It is worth noting that even if B_t is continuous, it is nowhere differentiable. From condition 4 we get that B_t is normally distributed with mean $E[B_t] = 0$ and $Var[B_t] = t$

$$B_t \sim N(0, t).$$

This implies also that the second moment is $E[B_t^2] = t$. Let $0 < s < t$. Since the increments are independent, we can write

$$E[B_s B_t] = E[(B_s - B_0)(B_t - B_s) + B_s^2] = E[B_s - B_0]E[B_t - B_s] + E[B_s^2] = s.$$

Consequently, B_s and B_t are not independent.

Condition 4 has also a physical explanation. A pollen grain suspended in water is kicked by a very large numbers of water molecules. The influence of each molecule on the grain is independent of the other molecules. These effects are average out into a resultant increment of the grain coordinate. According to the Central Limit Theorem, this increment has to be normal distributed.

Proposition 2.1.2 *A Brownian motion process B_t is a martingale with respect to the information set $\mathcal{F}_t = \sigma(B_s; s \le t)$.*

Proof: The integrability of B_t follows from Jensen's inequality

$$E[|B_t|]^2 \le E[B_t^2] = Var(B_t) = |t| < \infty.$$

B_t is obviously $\mathcal{F}_t$-predictable. Let $s < t$ and write $B_t = B_s + (B_t - B_s)$. Then

$$\begin{aligned} E[B_t|\mathcal{F}_s] &= E[B_s + (B_t - B_s)|\mathcal{F}_s] \\ &= E[B_s|\mathcal{F}_s] + E[B_t - B_s|\mathcal{F}_s] \\ &= B_s + E[B_t - B_s] = B_s + E[B_{t-s} - B_0] = B_s, \end{aligned}$$

where we used that B_s is $\mathcal{F}_s$-predictable (from where $E[B_s|\mathcal{F}_s] = B_s$) and that the increment $B_t - B_s$ is independent of previous values of B_t contained in the information set $\mathcal{F}_t = \sigma(B_s; s \le t)$. ■

A process with similar properties as the Brownian motion was introduced by Wiener.

Definition 2.1.3 *A Wiener process W_t is a process adapted to a filtration $\mathcal{F}_t$ such that*

1. *The process starts at the origin, $W_0 = 0$;*
2. *W_t is an $\mathcal{F}_t$-martingale with $E[W_t^2] < \infty$ for all $t \ge 0$ and*

$$E[(W_t - W_s)^2] = t - s, \qquad s \le t;$$

3. *The process W_t is continuous in t.*

Since W_t is a martingale, its increments are unpredictable[1]and hence $E[W_t - W_s] = 0$; in particular $E[W_t] = 0$. It is easy to show that

$$Var[W_t - W_s] = |t - s|, \qquad Var[W_t] = t.$$

[1]This follows from $E[W_t - W_s] = E[W_t - W_s|\mathcal{F}_s] = E[W_t|\mathcal{F}_s] - W_s = W_s - W_s = 0$.

Exercise 2.1.4 *Show that a Brownian process B_t is a Winer process.*

The only property B_t has and W_t seems not to have is that the increments are normally distributed. However, there is no distinction between these two processes, as the following result states.

Theorem 2.1.5 (Lévy) *A Wiener process is a Brownian motion process.*

In stochastic calculus we often need to use infinitesimal notation and its properties. If dW_t denotes the infinitesimal increment of a Wiener process in the time interval dt, the aforementioned properties become $dW_t \sim N(0, dt)$, $E[dW_t] = 0$, and $E[(dW_t)^2] = dt$.

Proposition 2.1.6 *If W_t is a Wiener process with respect to the information set $\mathcal{F}_t$, then $Y_t = W_t^2 - t$ is a martingale.*

Proof: Y_t is integrable since

$$E[|Y_t|] \leq E[W_t^2 + t] = 2t < \infty, \qquad t > 0.$$

Let $s < t$. Using that the increments $W_t - W_s$ and $(W_t - W_s)^2$ are independent of the information set $\mathcal{F}_s$ and applying Proposition 1.11.4 yields

$$\begin{aligned}
E[W_t^2|\mathcal{F}_s] &= E[(W_s + W_t - W_s)^2|\mathcal{F}_s] \\
&= E[W_s^2 + 2W_s(W_t - W_s) + (W_t - W_s)^2|\mathcal{F}_s] \\
&= E[W_s^2|\mathcal{F}_s] + E[2W_s(W_t - W_s)|\mathcal{F}_s] + E[(W_t - W_s)^2|\mathcal{F}_s] \\
&= W_s^2 + 2W_s E[W_t - W_s|\mathcal{F}_s] + E[(W_t - W_s)^2|\mathcal{F}_s] \\
&= W_s^2 + 2W_s E[W_t - W_s] + E[(W_t - W_s)^2] \\
&= W_s^2 + t - s,
\end{aligned}$$

and hence $E[W_t^2 - t|\mathcal{F}_s] = W_s^2 - s$, for $s < t$. ■

The following result states the memoryless property of Brownian motion[2] W_t.

Proposition 2.1.7 *The conditional distribution of W_{t+s}, given the present W_t and the past W_u, $0 \leq u < t$, depends only on the present.*

Proof: Using the independent increment assumption, we have

$$\begin{aligned}
& P(W_{t+s} \leq c|W_t = x, W_u, 0 \leq u < t) \\
= \; & P(W_{t+s} - W_t \leq c - x|W_t = x, W_u, 0 \leq u < t) \\
= \; & P(W_{t+s} - W_t \leq c - x) \\
= \; & P(W_{t+s} \leq c|W_t = x).
\end{aligned}$$

[2]These type of processes are called Marcov processes.

■

Since W_t is normally distributed with mean 0 and variance t, its density function is

$$\phi_t(x) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{x^2}{2t}}.$$

Then its distribution function is

$$F_t(x) = P(W_t \le x) = \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{x} e^{-\frac{u^2}{2t}}\, du$$

The probability that W_t is between the values a and b is given by

$$P(a \le W_t \le b) = \frac{1}{\sqrt{2\pi t}} \int_{a}^{b} e^{-\frac{u^2}{2t}}\, du, \qquad a < b.$$

Even if the increments of a Brownian motion are independent, their values are still correlated.

Proposition 2.1.8 *Let* $0 \le s \le t$. *Then*

1. $Cov(W_s, W_t) = s$;

2. $Corr(W_s, W_t) = \sqrt{\frac{s}{t}}$.

Proof: 1. Using the properties of covariance

$$\begin{aligned}
Cov(W_s, W_t) &= Cov(W_s, W_s + W_t - W_s) \\
&= Cov(W_s, W_s) + Cov(W_s, W_t - W_s) \\
&= Var(W_s) + E[W_s(W_t - W_s)] - E[W_s]E[W_t - W_s] \\
&= s + E[W_s]E[W_t - W_s] \\
&= s,
\end{aligned}$$

since $E[W_s] = 0$.

We can also arrive at the same result starting from the formula

$$Cov(W_s, W_t) = E[W_s W_t] - E[W_s]E[W_t] = E[W_s W_t].$$

Using that conditional expectations have the same expectation, factoring the predictable part out, and using that W_t is a martingale, we have

$$\begin{aligned}
E[W_s W_t] &= E[E[W_s W_t | \mathcal{F}_s]] = E[W_s E[W_t | \mathcal{F}_s]] \\
&= E[W_s W_s] = E[W_s^2] = s,
\end{aligned}$$

so $Cov(W_s, W_t) = s$.

2. The correlation formula yields

$$Corr(W_s, W_t) = \frac{Cov(W_s, W_t)}{\sigma(W_t)\sigma(W_s)} = \frac{s}{\sqrt{s}\sqrt{t}} = \sqrt{\frac{s}{t}}.$$

∎

Remark 2.1.9 *Removing the order relation between s and t, the previous relations can also be stated as*

$$\begin{aligned} Cov(W_s, W_t) &= \min\{s, t\}; \\ Corr(W_s, W_t) &= \sqrt{\frac{\min\{s, t\}}{\max\{s, t\}}}. \end{aligned}$$

The following exercises state the translation and the scaling invariance of the Brownian motion.

Exercise 2.1.10 *For any $t_0 \geq 0$, show that the process $X_t = W_{t+t_0} - W_{t_0}$ is a Brownian motion. This can be also stated as saying that the Brownian motion is translation invariant.*

Exercise 2.1.11 *For any $\lambda > 0$, show that the process $X_t = \frac{1}{\sqrt{\lambda}} W_{\lambda t}$ is a Brownian motion. This says that the Brownian motion is invariant by scaling.*

Exercise 2.1.12 *Let $0 < s < t < u$. Show the following multiplicative property*

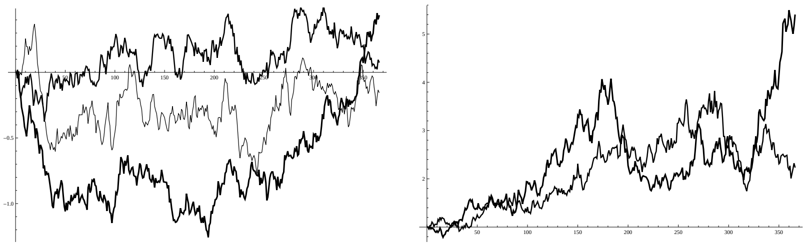
$$Corr(W_s, W_t)Corr(W_t, W_u) = Corr(W_s, W_u).$$

Exercise 2.1.13 *Find the expectations $E[W_t^3]$ and $E[W_t^4]$.*

Exercise 2.1.14 *(a) Use the martingale property of $W_t^2 - t$ to find $E[(W_t^2 - t)(W_s^2 - s)]$;*
(b) Evaluate $E[W_t^2 W_s^2]$;
(c) Compute $Cov(W_t^2, W_s^2)$;
(d) Find $Corr(W_t^2, W_s^2)$.

Exercise 2.1.15 *Consider the process $Y_t = tW_{\frac{1}{t}}$, $t > 0$, and define $Y_0 = 0$.*
(a) Find the distribution of Y_t;
(b) Find the probability density of Y_t;
(c) Find $Cov(Y_s, Y_t)$;
(d) Find $E[Y_t - Y_s]$ and $Var(Y_t - Y_s)$ for $s < t$.

It is worth noting that the process $Y_t = tW_{\frac{1}{t}}$, $t > 0$ with $Y_0 = 0$ is a Brownian motion.



a **b**

Figure 2.1: **a** *Three simulations of the Brownian motion process* W_t; **b** *Two simulations of the geometric Brownian motion process* e^{W_t}.

Exercise 2.1.16 *The process* $X_t = |W_t|$ *is called Brownian motion reflected at the origin. Show that*
(*a*) $E[|W_t|] = \sqrt{2t/\pi}$;
(*b*) $Var(|W_t|) = (1 - \frac{2}{\pi})t$.

Exercise 2.1.17 *Let* $0 < s < t$. *Find* $E[W_t^2|\mathcal{F}_s]$.

Exercise 2.1.18 *Let* $0 < s < t$. *Show that*

(*a*) $E[W_t^3|\mathcal{F}_s] = 3(t-s)W_s + W_s^3$;

(*b*) $E[W_t^4|\mathcal{F}_s] = 3(t-s)^2 + 6(t-s)W_s^2 + W_s^4$.

Exercise 2.1.19 *Show that the following processes are Brownian motions*

(*a*) $X_t = W_T - W_{T-t}$, $0 \le t \le T$;

(*b*) $Y_t = -W_t$, $t \ge 0$.

2.2 Geometric Brownian Motion

The process $X_t = e^{W_t}$, $t \ge 0$ is called *geometric Brownian motion*. A few simulations of this process are contained in Fig.2.1 b. The following result will be useful in the following.

Lemma 2.2.1 $E[e^{\alpha W_t}] = e^{\alpha^2 t/2}$, *for* $\alpha \ge 0$.

Proof: Using the definition of expectation

$$\begin{aligned} E[e^{\alpha W_t}] &= \int e^{\alpha x}\phi_t(x)\,dx = \frac{1}{\sqrt{2\pi t}}\int e^{-\frac{x^2}{2t}+\alpha x}\,dx \\ &= e^{\alpha^2 t/2}, \end{aligned}$$

where we have used the integral formula

$$\int e^{-ax^2+bx}\,dx = \sqrt{\frac{\pi}{a}}e^{\frac{b^2}{4a}}, \qquad a > 0$$

with $a = \frac{1}{2t}$ and $b = \alpha$. ■

Proposition 2.2.2 *The geometric Brownian motion $X_t = e^{W_t}$ is log-normally distributed with mean $e^{t/2}$ and variance $e^{2t} - e^t$.*

Proof: Since W_t is normally distributed, then $X_t = e^{W_t}$ will have a log-normal distribution. Using Lemma 2.2.1 we have

$$\begin{aligned} E[X_t] &= E[e^{W_t}] = e^{t/2} \\ E[X_t^2] &= E[e^{2W_t}] = e^{2t}, \end{aligned}$$

and hence the variance is

$$Var[X_t] = E[X_t^2] - E[X_t]^2 = e^{2t} - (e^{t/2})^2 = e^{2t} - e^t.$$

■

The distribution function of $X_t = e^{W_t}$ can be obtained by reducing it to the distribution function of a Brownian motion.

$$\begin{aligned} F_{X_t}(x) &= P(X_t \le x) = P(e^{W_t} \le x) \\ &= P(W_t \le \ln x) = F_{W_t}(\ln x) \\ &= \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{\ln x} e^{-\frac{u^2}{2t}}\,du. \end{aligned}$$

The density function of the geometric Brownian motion $X_t = e^{W_t}$ is given by

$$p(x) = \frac{d}{dx}F_{X_t}(x) = \begin{cases} \dfrac{1}{x\sqrt{2\pi t}}e^{-(\ln x)^2/(2t)}, & \text{if } x > 0, \\ 0, & \text{elsewhere.} \end{cases}$$

Exercise 2.2.3 *Show that*

$$E[e^{W_t - W_s}] = e^{\frac{t-s}{2}}, \qquad s < t.$$

Exercise 2.2.4 *Let $X_t = e^{W_t}$.*

(a) Show that X_t is not a martingale.

(b) Show that $e^{-\frac{t}{2}}X_t$ is a martingale.

(c) Show that for any constant $c \in \mathbb{R}$, the process $Y_t = e^{cW_t - \frac{1}{2}c^2 t}$ is a martingale.

Exercise 2.2.5 *If* $X_t = e^{W_t}$, *find* $Cov(X_s, X_t)$

(a) *by direct computation;*

(b) *by using Exercise* 2.2.4 (b).

Exercise 2.2.6 *Show that*

$$E[e^{2W_t^2}] = \begin{cases} (1-4t)^{-1/2}, & 0 \le t < 1/4 \\ \infty, & \textit{otherwise.} \end{cases}$$

2.3 Integrated Brownian Motion

The stochastic process

$$Z_t = \int_0^t W_s \, ds, \qquad t \ge 0$$

is called *integrated Brownian motion*. Obviously, $Z_0 = 0$.

Let $0 = s_0 < s_1 < \cdots < s_k < \cdots s_n = t$, with $s_k = \frac{kt}{n}$. Then Z_t can be written as a limit of Riemann sums

$$Z_t = \lim_{n\to\infty} \sum_{k=1}^{n} W_{s_k} \Delta s = t \lim_{n\to\infty} \frac{W_{s_1} + \cdots + W_{s_n}}{n},$$

where $\Delta s = s_{k+1} - s_k = \frac{t}{n}$. We are tempted to apply the Central Limit Theorem at this point, but W_{s_k} are not independent, so we first need to transform the sum into a sum of independent normally distributed random variables. A straightforward computation shows that

$$\begin{aligned} & W_{s_1} + \cdots + W_{s_n} \\ = \; & n(W_{s_1} - W_0) + (n-1)(W_{s_2} - W_{s_1}) + \cdots + (W_{s_n} - W_{s_{n-1}}) \\ = \; & X_1 + X_2 + \cdots + X_n. \end{aligned} \tag{2.3.1}$$

Since the increments of a Brownian motion are independent and normally distributed, we have

$$\begin{aligned} & X_1 \sim N\big(0, n^2 \Delta s\big) \\ & X_2 \sim N\big(0, (n-1)^2 \Delta s\big) \\ & X_3 \sim N\big(0, (n-2)^2 \Delta s\big) \\ & \vdots \\ & X_n \sim N\big(0, \Delta s\big). \end{aligned}$$

Recall now the following variant of the Central Limit Theorem:

Theorem 2.3.1 *If X_j are independent random variables normally distributed with mean μ_j and variance σ_j^2, then the sum $X_1 + \cdots + X_n$ is also normally distributed with mean $\mu_1 + \cdots + \mu_n$ and variance $\sigma_1^2 + \cdots + \sigma_n^2$.*

Then

$$X_1 + \cdots + X_n \sim N\Big(0, (1 + 2^2 + 3^2 + \cdots + n^2)\Delta s\Big) = N\Big(0, \frac{n(n+1)(2n+1)}{6}\Delta s\Big),$$

with $\Delta s = \dfrac{t}{n}$. Using (2.3.1) yields

$$t\frac{W_{s_1} + \cdots + W_{s_n}}{n} \sim N\Big(0, \frac{(n+1)(2n+1)}{6n^2}t^3\Big).$$

"Taking the limit" we get

$$Z_t \sim N\Big(0, \frac{t^3}{3}\Big).$$

Proposition 2.3.2 *The integrated Brownian motion Z_t has a normal distribution with mean 0 and variance $t^3/3$.*

Remark 2.3.3 *The aforementioned limit was taken heuristically, without specifying the type of the convergence. In order to make this to work, the following result is usually used:*

If X_n is a sequence of normal random variables that converges in mean square to X, then the limit X is normal distributed, with $E[X_n] \to E[X]$ and $Var(X_n) \to Var(X)$, as $n \to \infty$.

The mean and the variance can also be computed in a direct way as follows. By Fubini's theorem we have

$$\begin{aligned} E[Z_t] &= E[\int_0^t W_s\,ds] = \int_{\mathbb{R}} \int_0^t W_s\,ds\,dP \\ &= \int_0^t \int_{\mathbb{R}} W_s\,dP\,ds = \int_0^t E[W_s]\,ds = 0, \end{aligned}$$

since $E[W_s] = 0$. Then the variance is given by

$$\begin{aligned} Var[Z_t] &= E[Z_t^2] - E[Z_t]^2 = E[Z_t^2] \\ &= E[\int_0^t W_u\,du \cdot \int_0^t W_v\,dv] = E[\int_0^t \int_0^t W_u W_v\,dudv] \\ &= \int_0^t \int_0^t E[W_u W_v]\,dudv = \iint_{[0,t]\times[0,t]} \min\{u,v\}\,dudv \\ &= \iint_{D_1} \min\{u,v\}\,dudv + \iint_{D_2} \min\{u,v\}\,dudv, \qquad (2.3.2) \end{aligned}$$

where

$$D_1 = \{(u,v); u > v, 0 \le u \le t\}, \qquad D_2 = \{(u,v); u < v, 0 \le u \le t\}$$

The first integral can be evaluated using Fubini's theorem

$$\begin{aligned}\iint_{D_1} \min\{u,v\}\, dudv &= \iint_{D_1} v\, dudv \\ &= \int_0^t \Big(\int_0^u v\, dv\Big)\, du = \int_0^t \frac{u^2}{2}\, du = \frac{t^3}{6}.\end{aligned}$$

Similarly, the latter integral is equal to

$$\iint_{D_2} \min\{u,v\}\, dudv = \frac{t^3}{6}.$$

Substituting in (2.3.2) yields

$$Var[Z_t] = \frac{t^3}{6} + \frac{t^3}{6} = \frac{t^3}{3}.$$

Exercise 2.3.4 *(a) Prove that the moment generating function of Z_t is*

$$m(u) = e^{u^2 t^3/6}.$$

(b) Use the first part to find the mean and variance of Z_t.

Exercise 2.3.5 *Let $s < t$. Show that the covariance of the integrated Brownian motion is given by*

$$Cov\Big(Z_s, Z_t\Big) = s^2\Big(\frac{t}{2} - \frac{s}{6}\Big), \qquad s < t.$$

Exercise 2.3.6 *Show that*

(a) $Cov(Z_t, Z_t - Z_{t-h}) = \frac{1}{2}t^2 h + o(h)$, where $o(h)$ denotes a quantity such that $\lim_{h\to 0} o(h)/h = 0$;

(b) $Cov(Z_t, W_t) = \dfrac{t^2}{2}$.

Exercise 2.3.7 *Show that*

$$E[e^{W_s + W_u}] = e^{\frac{u+s}{2}} e^{\min\{s,u\}}.$$

Exercise 2.3.8 *Consider the process $X_t = \displaystyle\int_0^t e^{W_s}\, ds$.*

(a) Find the mean of X_t;
(b) Find the variance of X_t.

Exercise 2.3.9 *Consider the process $Z_t = \displaystyle\int_0^t W_u\, du$, $t > 0$.*

(a) Show that $E[Z_T|\mathcal{F}_t] = Z_t + W_t(T-t)$, for any $t < T$;
(b) Prove that the process $M_t = Z_t - tW_t$ is an $\mathcal{F}_t$-martingale.

2.4 Exponential Integrated Brownian Motion

If $Z_t = \int_0^t W_s\, ds$ denotes the integrated Brownian motion, the process

$$V_t = e^{Z_t}$$

is called *exponential integrated Brownian motion*. The process starts at $V_0 = e^0 = 1$. Since Z_t is normally distributed, then V_t is log-normally distributed. We compute the mean and the variance in a direct way. Using Exercises 2.2.5 and 2.3.4 we have

$$\begin{aligned}
E[V_t] &= E[e^{Z_t}] = m(1) = e^{\frac{t^3}{6}} \\
E[V_t^2] &= E[e^{2Z_t}] = m(2) = e^{\frac{4t^3}{6}} = e^{\frac{2t^3}{3}} \\
Var(V_t) &= E[V_t^2] - E[V_t]^2 = e^{\frac{2t^3}{3}} - e^{\frac{t^3}{3}} \\
Cov(V_s, V_t) &= e^{\frac{t+3s}{2}}.
\end{aligned}$$

Exercise 2.4.1 *Show that* $E[V_T|\mathcal{F}_t] = V_t e^{(T-t)W_t + \frac{(T-t)^3}{3}}$ *for* $t < T$.

2.5 Brownian Bridge

The process $X_t = W_t - tW_1$ is called the *Brownian bridge* fixed at both 0 and 1. Since we can also write

$$\begin{aligned}
X_t &= W_t - tW_t - tW_1 + tW_t \\
&= (1-t)(W_t - W_0) - t(W_1 - W_t),
\end{aligned}$$

using that the increments $W_t - W_0$ and $W_1 - W_t$ are independent and normally distributed, with

$$W_t - W_0 \sim N(0, t), \qquad W_1 - W_t \sim N(0, 1-t),$$

it follows that X_t is normally distributed with

$$\begin{aligned}
E[X_t] &= (1-t)E[(W_t - W_0)] - tE[(W_1 - W_t)] = 0 \\
Var[X_t] &= (1-t)^2 Var[(W_t - W_0)] + t^2 Var[(W_1 - W_t)] \\
&= (1-t)^2(t-0) + t^2(1-t) \\
&= t(1-t).
\end{aligned}$$

This can be also stated by saying that the Brownian bridge tied at 0 and 1 is a Gaussian process with mean 0 and variance $t(1-t)$, so $X_t \sim N\big(0, t(1-t)\big)$.

Exercise 2.5.1 *Let* $X_t = W_t - tW_1$, $0 \le t \le 1$ *be a Brownian bridge fixed at* 0 *and* 1. *Let* $Y_t = X_t^2$. *Show that* $Y_0 = Y_1 = 0$ *and find* $E[Y_t]$ *and* $Var(Y_t)$.

2.6 Brownian Motion with Drift

The process $Y_t = \mu t + W_t$, $t \geq 0$, is called *Brownian motion with drift*. The process Y_t tends to drift off at a rate μ. It starts at $Y_0 = 0$ and it is a Gaussian process with mean

$$E[Y_t] = \mu t + E[W_t] = \mu t$$

and variance

$$Var[Y_t] = Var[\mu t + W_t] = Var[W_t] = t.$$

Exercise 2.6.1 *Find the distribution and the density functions of the process Y_t.*

2.7 Bessel Process

This section deals with the process satisfied by the Euclidean distance from the origin to a particle following a Brownian motion in $\mathbb{R}^n$. More precisely, if $W_1(t), \cdots, W_n(t)$ are independent Brownian motions, let $W(t) = (W_1(t), \cdots, W_n(t))$ be a Brownian motion in $\mathbb{R}^n$, $n \geq 2$. The process

$$R_t = dist(O, W(t)) = \sqrt{W_1(t)^2 + \cdots + W_n(t)^2}$$

is called *n-dimensional Bessel process.*

The probability density of this process is given by the following result.

Proposition 2.7.1 *The probability density function of R_t, $t > 0$ is given by*

$$p_t(\rho) = \begin{cases} \frac{2}{(2t)^{n/2}\,\Gamma(n/2)}\,\rho^{n-1} e^{-\frac{\rho^2}{2t}}, & \rho \geq 0; \\ \\ 0, & \rho < 0 \end{cases}$$

with

$$\Gamma\Big(\frac{n}{2}\Big) = \begin{cases} (\frac{n}{2} - 1)! & \text{for } n \text{ even;} \\ \\ (\frac{n}{2} - 1)(\frac{n}{2} - 2)\cdots \frac{3}{2}\frac{1}{2}\sqrt{\pi}, & \text{for } n \text{ odd.} \end{cases}$$

Proof: Since the Brownian motions $W_1(t), \ldots, W_n(t)$ are independent, their joint density function is

$$\begin{aligned} f_{W_1 \cdots W_n}(t) &= f_{W_1}(t) \cdots f_{W_n}(t) \\ &= \frac{1}{(2\pi t)^{n/2}}\, e^{-(x_1^2 + \cdots + x_n^2)/(2t)}, \qquad t > 0. \end{aligned}$$

In the next computation we shall use the following formula of integration that follows from the use of polar coordinates

$$\int_{\{|x|\leq\rho\}} f(x)\,dx = \sigma(\mathbb{S}^{n-1})\int_0^\rho r^{n-1}g(r)\,dr, \tag{2.7.3}$$

where $f(x) = g(|x|)$ is a function on $\mathbb{R}^n$ with spherical symmetry, and where

$$\sigma(\mathbb{S}^{n-1}) = \frac{2\pi^{n/2}}{\Gamma(n/2)}$$

is the area of the $(n-1)$-dimensional sphere in $\mathbb{R}^n$.

Let $\rho \geq 0$. The distribution function of R_t is

$$\begin{aligned}
F_R(\rho) &= P(R_t \leq \rho) = \int_{\{R_t\leq\rho\}} f_{W_1\cdots W_n}(t)\,dx_1\cdots dx_n \\
&= \int_{x_1^2+\cdots+x_n^2\leq\rho^2} \frac{1}{(2\pi t)^{n/2}}\, e^{-(x_1^2+\cdots+x_n^2)/(2t)}\,dx_1\cdots dx_n \\
&= \int_0^\rho r^{n-1}\left(\int_{\mathbb{S}(0,1)} \frac{1}{(2\pi t)^{n/2}}\, e^{-(x_1^2+\cdots+x_n^2)/(2t)}\,d\sigma\right)dr \\
&= \frac{\sigma(\mathbb{S}^{n-1})}{(2\pi t)^{n/2}}\int_0^\rho r^{n-1}\,e^{-r^2/(2t)}\,dr.
\end{aligned}$$

Differentiating yields

$$\begin{aligned}
p_t(\rho) &= \frac{d}{d\rho}F_R(\rho) = \frac{\sigma(\mathbb{S}^{n-1})}{(2\pi t)^{n/2}}\,\rho^{n-1}e^{-\frac{\rho^2}{2t}} \\
&= \frac{2}{(2t)^{n/2}\Gamma(n/2)}\,\rho^{n-1}e^{-\frac{\rho^2}{2t}}, \qquad \rho > 0, t > 0.
\end{aligned}$$

■

It is worth noting that in the 2-dimensional case the aforementioned density becomes a particular case of a Weibull distribution with parameters $m = 2$ and $\alpha = 2t$, called *Wald's distribution*

$$\boxed{p_t(x) = \frac{1}{t}xe^{-\frac{x^2}{2t}}, \qquad x > 0, t > 0.}$$

Exercise 2.7.2 *Let $P(R_t \leq t)$ be the probability of a 2-dimensional Brownian motion to be inside of the disk $D(0,\rho)$ at time $t > 0$. Show that*

$$\frac{\rho^2}{2t}\Big(1 - \frac{\rho^2}{4t}\Big) < P(R_t \leq t) < \frac{\rho^2}{2t}.$$

Exercise 2.7.3 *Let R_t be a 2-dimensional Bessel process. Show that*

(a) $E[R_t] = \sqrt{2\pi t}/2$;

(b) $Var(R_t) = 2t\big(1 - \frac{\pi}{4}\big)$.

Exercise 2.7.4 *Let* $X_t = \dfrac{R_t}{t}$, $t > 0$, *where* R_t *is a 2-dimensional Bessel process. Show that* $X_t \to 0$ *as* $t \to \infty$ *in mean square.*

2.8 The Poisson Process

A *Poisson process* describes the number of occurrences of a certain event before time t, such as: the number of electrons arriving at an anode until time t; the number of cars arriving at a gas station until time t; the number of phone calls received on a certain day until time t; the number of visitors entering a museum on a certain day until time t; the number of earthquakes that occurred in Chile during the time interval $[0, t]$; the number of shocks in the stock market from the beginning of the year until time t; the number of twisters that might hit Alabama from the beginning of the century until time t.

The definition of a Poisson process is stated more precisely in the following.

Definition 2.8.1 *A Poisson process is a stochastic process* N_t, $t \geq 0$, *which satisfies*

1. The process starts at the origin, $N_0 = 0$;

2. N_t *has stationary, independent increments;*

3. The process N_t *is right continuous in* t, *with left hand limits;*

4. The increments $N_t - N_s$, *with* $0 < s < t$, *have a Poisson distribution with parameter* $\lambda(t - s)$, *i.e.*

$$P(N_t - N_s = k) = \frac{\lambda^k (t-s)^k}{k!} e^{-\lambda(t-s)}.$$

It can be shown that condition *4* in the previous definition can be replaced by the following two conditions:

$$\begin{aligned} P(N_t - N_s = 1) &= \lambda(t-s) + o(t-s) && (2.8.4) \\ P(N_t - N_s \geq 2) &= o(t-s), && (2.8.5) \end{aligned}$$

where $o(h)$ denotes a quantity such that $\lim_{h \to 0} o(h)/h = 0$. Then the probability that a jump of size 1 occurs in the infinitesimal interval dt is equal to λdt, and the probability that at least 2 events occur in the same small interval is zero. This implies that the random variable dN_t may take only two values, 0 and 1, and hence satisfies

$$\begin{aligned} P(dN_t = 1) &= \lambda\, dt && (2.8.6) \\ P(dN_t = 0) &= 1 - \lambda\, dt. && (2.8.7) \end{aligned}$$

Exercise 2.8.2 *Show that if condition* 4 *is satisfied, then conditions* $(2.8.4 - 2.8.5)$ *hold.*

Exercise 2.8.3 *Which of the following expressions are $o(h)$?*

(a) $f(h) = 3h^2 + h$;

(b) $f(h) = \sqrt{h} + 5$;

(c) $f(h) = h \ln |h|$;

(d) $f(h) = he^h$.

The fact that $N_t - N_s$ is stationary can be stated as

$$P(N_{t+s} - N_s \leq n) = P(N_t - N_0 \leq n) = P(N_t \leq n) = \sum_{k=0}^{n} \frac{(\lambda t)^k}{k!} e^{-\lambda t}.$$

From condition *4* we get the mean and variance of increments

$$E[N_t - N_s] = \lambda(t - s), \qquad Var[N_t - N_s] = \lambda(t - s).$$

In particular, the random variable N_t is Poisson distributed with $E[N_t] = \lambda t$ and $Var[N_t] = \lambda t$. The parameter λ is called the *rate* of the process. This means that the events occur at the constant rate λ.

Since the increments are independent, we have for $0 < s < t$

$$\begin{aligned} E[N_s N_t] &= E[(N_s - N_0)(N_t - N_s) + N_s^2] \\ &= E[N_s - N_0]E[N_t - N_s] + E[N_s^2] \\ &= \lambda s \cdot \lambda(t - s) + (Var[N_s] + E[N_s]^2) \\ &= \lambda^2 st + \lambda s. \end{aligned} \tag{2.8.8}$$

As a consequence we have the following result:

Proposition 2.8.4 *Let $0 \leq s \leq t$. Then*

1. $Cov(N_s, N_t) = \lambda s$;

2. $Corr(N_s, N_t) = \sqrt{\dfrac{s}{t}}$.

Proof: 1. Using (2.8.8) we have

$$\begin{aligned} Cov(N_s, N_t) &= E[N_s N_t] - E[N_s]E[N_t] \\ &= \lambda^2 st + \lambda s - \lambda s \lambda t \\ &= \lambda s. \end{aligned}$$

2. Using the formula for the correlation yields

$$Corr(N_s, N_t) = \frac{Cov(N_s, N_t)}{(Var[N_s]Var[N_t])^{1/2}} = \frac{\lambda s}{(\lambda s \lambda t)^{1/2}} = \sqrt{\frac{s}{t}}.$$

■

It worth noting the similarity with Proposition 2.1.8.

Proposition 2.8.5 *Let N_t be $\mathcal{F}_t$-adapted. Then the process $M_t = N_t - \lambda t$ is an $\mathcal{F}_t$-martingale.*

Proof: Let $s < t$ and write $N_t = N_s + (N_t - N_s)$. Then

$$\begin{aligned} E[N_t|\mathcal{F}_s] &= E[N_s + (N_t - N_s)|\mathcal{F}_s] \\ &= E[N_s|\mathcal{F}_s] + E[N_t - N_s|\mathcal{F}_s] \\ &= N_s + E[N_t - N_s] \\ &= N_s + \lambda(t - s), \end{aligned}$$

where we used that N_s is $\mathcal{F}_s$-predictable (and hence $E[N_s|\mathcal{F}_s] = N_s$) and that the increment $N_t - N_s$ is independent of previous values of N_s and the information set $\mathcal{F}_s$. Subtracting λt yields

$$E[N_t - \lambda t|\mathcal{F}_s] = N_s - \lambda s,$$

or $E[M_t|\mathcal{F}_s] = M_s$. Since it is obvious that M_t is integrable and $\mathcal{F}_t$-adapted, it follows that M_t is a martingale. ■

It is worth noting that the Poisson process N_t is not a martingale. The martingale process $M_t = N_t - \lambda t$ is called the *compensated Poisson process.*

Exercise 2.8.6 *Compute $E[N_t^2|\mathcal{F}_s]$ for $s < t$. Is the process N_t^2 an $\mathcal{F}_s$-martingale?*

Exercise 2.8.7 (*a*) *Show that the moment generating function of the random variable N_t is*

$$m_{N_t}(x) = e^{\lambda t(e^x - 1)}.$$

(*b*) *Deduct the expressions for the first few moments*

$$\begin{aligned} E[N_t] &= \lambda t \\ E[N_t^2] &= \lambda^2 t^2 + \lambda t \\ E[N_t^3] &= \lambda^3 t^3 + 3\lambda^2 t^2 + \lambda t \\ E[N_t^4] &= \lambda^4 t^4 + 6\lambda^3 t^3 + 7\lambda^2 t^2 + \lambda t. \end{aligned}$$

(*c*) *Show that the first few central moments are given by*

$$\begin{aligned} E[N_t - \lambda t] &= 0 \\ E[(N_t - \lambda t)^2] &= \lambda t \\ E[(N_t - \lambda t)^3] &= \lambda t \\ E[(N_t - \lambda t)^4] &= 3\lambda^2 t^2 + \lambda t. \end{aligned}$$

Exercise 2.8.8 *Find the mean and variance of the process $X_t = e^{N_t}$.*

Exercise 2.8.9 (*a*) *Show that the moment generating function of the random variable* M_t *is*

$$m_{M_t}(x) = e^{\lambda t(e^x - x - 1)}.$$

(*b*) *Let* $s < t$. *Verify that*

$$\begin{aligned}
E[M_t - M_s] &= 0, \\
E[(M_t - M_s)^2] &= \lambda(t-s), \\
E[(M_t - M_s)^3] &= \lambda(t-s), \\
E[(M_t - M_s)^4] &= \lambda(t-s) + 3\lambda^2(t-s)^2.
\end{aligned}$$

Exercise 2.8.10 *Let* $s < t$. *Show that*

$$Var[(M_t - M_s)^2] = \lambda(t-s) + 2\lambda^2(t-s)^2.$$

2.9 Interarrival times

For each state of the world, ω, the path $t \to N_t(\omega)$ is a step function that exhibits unit jumps. Each jump in the path corresponds to an occurrence of a new event. Let T_1 be the random variable which describes the time of the 1st jump. Let T_2 be the time between the 1st jump and the second one. In general, denote by T_n the time elapsed between the $(n-1)$th and nth jumps. The random variables T_n are called *interarrival times.*

Proposition 2.9.1 *The random variables* T_n *are independent and exponentially distributed with mean* $E[T_n] = 1/\lambda$.

Proof: We start by noticing that the events $\{T_1 > t\}$ and $\{N_t = 0\}$ are the same, since both describe the situation that no events occurred until after time t. Then

$$P(T_1 > t) = P(N_t = 0) = P(N_t - N_0 = 0) = e^{-\lambda t},$$

and hence the distribution function of T_1 is

$$F_{T_1}(t) = P(T_1 \le t) = 1 - P(T_1 > t) = 1 - e^{-\lambda t}.$$

Differentiating yields the density function

$$f_{T_1}(t) = \frac{d}{dt} F_{T_1}(t) = \lambda e^{-\lambda t}.$$

It follows that T_1 is has an exponential distribution, with $E[T_1] = 1/\lambda$.
In order to show that the random variables T_1 and T_2 are independent, if suffices to show that

$$P(T_2 \le t) = P(T_2 \le t | T_1 = s),$$

i.e. the distribution function of T_2 is independent of the values of T_1. We note first that from the independent increments property

$$\begin{aligned} & P\big(0 \text{ jumps in } (s, s+t], 1 \text{ jump in } (0, s]\big) = P(N_{s+t} - N_s = 0, N_s - N_0 = 1) \\ & = P(N_{s+t} - N_s = 0) P(N_s - N_0 = 1) = P\big(0 \text{ jumps in } (s, s+t]\big) P\big(1 \text{ jump in } (0, s]\big). \end{aligned}$$

Then the conditional distribution of T_2 is

$$\begin{aligned} F(t|s) &= P(T_2 \le t | T_1 = s) = 1 - P(T_2 > t | T_1 = s) \\ &= 1 - \frac{P(T_2 > t, T_1 = s)}{P(T_1 = s)} \\ &= 1 - \frac{P\big(0 \text{ jumps in } (s, s+t], 1 \text{ jump in } (0, s]\big)}{P(T_1 = s)} \\ &= 1 - \frac{P\big(0 \text{ jumps in } (s, s+t]\big) P\big(1 \text{ jump in } (0, s]\big)}{P\big(1 \text{ jump in } (0, s]\big)} = 1 - P\big(0 \text{ jumps in } (s, s+t]\big) \\ &= 1 - P(N_{s+t} - N_s = 0) = 1 - e^{-\lambda t}, \end{aligned}$$

which is independent of s. Then T_2 is independent of T_1 and exponentially distributed. A similar argument for any T_n leads to the desired result.

■

2.10 Waiting times

The random variable $S_n = T_1 + T_2 + \cdots + T_n$ is called the *waiting time until the nth jump*. The event $\{S_n \le t\}$ means that there are n jumps that occurred before or at time t, i.e. there are at least n events that happened up to time t; the event is equal to $\{N_t \ge n\}$. Hence the distribution function of S_n is given by

$$F_{S_n}(t) = P(S_n \le t) = P(N_t \ge n) = e^{-\lambda t} \sum_{k=n}^{\infty} \frac{(\lambda t)^k}{k!}\cdot$$

Differentiating we obtain the density function of the waiting time S_n

$$f_{S_n}(t) = \frac{d}{dt} F_{S_n}(t) = \frac{\lambda e^{-\lambda t} (\lambda t)^{n-1}}{(n-1)!}\cdot$$

Writing

$$f_{S_n}(t) = \frac{t^{n-1} e^{-\lambda t}}{(1/\lambda)^n \Gamma(n)},$$

it turns out that S_n has a gamma distribution with parameters $\alpha = n$ and $\beta = 1/\lambda$. It follows that

$$E[S_n] = \frac{n}{\lambda}, \quad Var[S_n] = \frac{n}{\lambda^2}\cdot$$

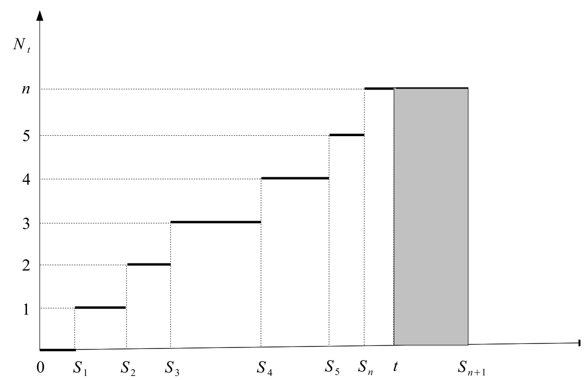


Figure 2.2: *The Poisson process N_t and the waiting times $S_1, S_2, \cdots S_n$. The shaded rectangle has area $n(S_{n+1} - t)$.*

The relation $\lim_{n\to\infty} E[S_n] = \infty$ states that the expectation of the waiting time is unbounded as $n \to \infty$.

Exercise 2.10.1 *Prove that* $\frac{d}{dt}F_{S_n}(t) = \frac{\lambda e^{-\lambda t}(\lambda t)^{n-1}}{(n-1)!}$.

Exercise 2.10.2 *Using that the interarrival times $T_1, T_2, \cdots$ are independent and exponentially distributed, compute directly the mean $E[S_n]$ and variance $Var(S_n)$.*

2.11 The Integrated Poisson Process

The function $u \to N_u$ is continuous with the exception of a set of countable jumps of size 1. It is known that such functions are Riemann integrable, so it makes sense to define the process

$$U_t = \int_0^t N_u \, du,$$

called the *integrated Poisson process.* The next result provides a relation between the process U_t and the partial sum of the waiting times S_k.

Proposition 2.11.1 *The integrated Poisson process can be expressed as*

$$U_t = tN_t - \sum_{k=1}^{N_t} S_k.$$

Let $N_t = n$. Since N_t is equal to k between the waiting times S_k and S_{k+1}, the process U_t, which is equal to the area of the subgraph of N_u between 0 and t, can be expressed as

$$U_t = \int_0^t N_u \, du = 1 \cdot (S_2 - S_1) + 2 \cdot (S_3 - S_2) + \cdots + n(S_{n+1} - S_n) - n(S_{n+1} - t).$$

Since $S_n < t < S_{n+1}$, the difference of the last two terms represents the area of last the rectangle, which has the length $t - S_n$ and the height n. Using associativity, a computation yields

$$1 \cdot (S_2 - S_1) + 2 \cdot (S_3 - S_2) + \cdots + n(S_{n+1} - S_n) = nS_{n+1} - (S_1 + S_2 + \cdots + S_n).$$

Substituting in the aforementioned relation yields

$$\begin{aligned} U_t &= nS_{n+1} - (S_1 + S_2 + \cdots + S_n) - n(S_{n+1} - t) \\ &= nt - (S_1 + S_2 + \cdots + S_n) \\ &= tN_t - \sum_{k=1}^{N_t} S_k, \end{aligned}$$

where we replaced n by N_t.

The conditional distribution of the waiting times is provided by the following useful result.

Theorem 2.11.2 *Given that $N_t = n$, the waiting times $S_1, S_2, \cdots, S_n$ have the the joint density function given by*

$$f(s_1, s_2, \cdots, s_n) = \frac{n!}{t^n}, \qquad 0 < s_1 \le s_2 \le \cdots \le s_n < t.$$

This is the same as the density of an ordered sample of size n from a uniform distribution on the interval $(0, t)$. A naive explanation of this result is as follows. If we know that there will be exactly n events during the time interval $(0, t)$, since the events can occur at any time, each of them can be considered uniformly distributed, with the density $f(s_k) = 1/t$. Since it makes sense to consider the events independent, taking into consideration all possible $n!$ permutaions, the joint density function becomes $f(s_1, \cdots, s_n) = n! f(s_1) \cdots f(s_n) = \dfrac{n!}{t^n}$.

Exercise 2.11.3 *Find the following means*

(a) $E[U_t]$.

(b) $E\Big[\sum_{k=1}^{N_t} S_k\Big]$.

Exercise 2.11.4 *Show that* $Var(U_t) = \dfrac{\lambda t^3}{3}$.

Exercise 2.11.5 *Can you apply a similar proof as in Proposition* 2.3.2 *to show that the integrated Poisson process* U_t *is also a Poisson process?*

Exercise 2.11.6 *Let* $Y : \Omega \to \mathbb{N}$ *be a discrete random variable. Then for any random variable* X *we have*

$$E[X] = \sum_{y \geq 0} E[X|Y = y]P(Y = y).$$

Exercise 2.11.7 *Use Exercise 2.11.6 to solve Exercise 2.11.3* (*b*).

Exercise 2.11.8 (*a*) *Let* T_k *be the* k*th interarrival time. Show that*

$$E[e^{-\sigma T_k}] = \frac{\lambda}{\lambda + \sigma}, \qquad \sigma > 0.$$

(*b*) *Let* $n = N_t$. *Show that*

$$U_t = nt - [nT_1 + (n-1)T_2 + \cdots + 2T_{n-1} + T_n].$$

(*c*) *Find the conditional expectation*

$$E\Big[e^{-\sigma U_t}\Big|N_t = n\Big].$$

(Hint: If know that there are exactly n *jumps in the interval* $[0, T]$, *it makes sense to consider the arrival time of the jumps* T_i *independent and uniformly distributed on* $[0, T]$*).*
(*d*) *Find the expectation*

$$E\Big[e^{-\sigma U_t}\Big].$$

2.12 Submartingales

A stochastic process X_t on the probability space $(\Omega, \mathcal{F}, P)$ is called *submartingale* with respect to the filtration $\mathcal{F}_t$ if:

(a) $\int_\Omega |X_t|\, dP < \infty$ (X_t integrable);

(b) X_t is known if $\mathcal{F}_t$ is given (X_t is adaptable to $\mathcal{F}_t$);

(c) $E[X_{t+s}|\mathcal{F}_t] \geq X_t, \forall t, s \geq 0$ (future predictions exceed the present value).

Example 2.12.1 *We shall prove that the process* $X_t = \mu t + \sigma W_t$, *with* $\mu > 0$ *is a submartingale.*

The integrability follows from the inequality $|X_t(\omega)| \leq \mu t + |W_t(\omega)|$ and integrability of W_t. The adaptability of X_t is obvious, and the last property follows from the computation:

$$\begin{aligned} E[X_{t+s}|\mathcal{F}_t] &= E[\mu t + \sigma W_{t+s}|\mathcal{F}_t] + \mu s > E[\mu t + \sigma W_{t+s}|\mathcal{F}_t] \\ &= \mu t + \sigma E[W_{t+s}|\mathcal{F}_t] = \mu t + \sigma W_t = X_t, \end{aligned}$$

where we used that W_t is a martingale.

Example 2.12.2 *We shall show that the square of the Brownian motion, W_t^2, is a submartingale.*

Using that $W_t^2 - t$ is a martingale, we have

$$\begin{aligned} E[W_{t+s}^2|\mathcal{F}_t] &= E[W_{t+s}^2 - (t+s)|\mathcal{F}_t] + t + s = W_t^2 - t + t + s \\ &= W_t^2 + s \geq W_t^2. \end{aligned}$$

The following result supplies examples of submatingales starting from martingales or submartingales.

Proposition 2.12.3 (a) *If X_t is a martingale and ϕ a convex function such that $\phi(X_t)$ is integrable, then the process $Y_t = \phi(X_t)$ is a submartingale.*

(b) *If X_t is a submartingale and ϕ an increasing convex function such that $\phi(X_t)$ is integrable, then the process $Y_t = \phi(X_t)$ is a submartingale.*

(c) *If X_t is a martingale and $f(t)$ is an increasing, integrable function, then $Y_t = X_t + f(t)$ is a submartingale.*

Proof: (a) Using Jensen's inequality for conditional probabilities, Exercise 1.12.6, we have

$$E[Y_{t+s}|\mathcal{F}_t] = E[\phi(X_{t+s})|\mathcal{F}_t] \geq \phi\Big(E[X_{t+s}|\mathcal{F}_t]\Big) = \phi(X_t) = Y_t.$$

(b) From the submatingale property and monotonicity of ϕ we have

$$\phi\Big(E[X_{t+s}|\mathcal{F}_t]\Big) \geq \phi(X_t).$$

Then apply a similar computation as in part (a).

(c) We shall check only the forcast property, since the other are obvious.

$$\begin{aligned} E[Y_{t+s}|\mathcal{F}_t] &= E[X_{t+s} + f(t+s)|\mathcal{F}_t] = E[X_{t+s}|\mathcal{F}_t] + f(t+s) \\ &= X_t + f(t+s) \geq X_t + f(t) = Y_t, \qquad \forall s, t > 0. \end{aligned}$$

■

Corollary 2.12.4 (a) *Let X_t be a martingale. Then X_t^2, $|X_t|$, e^{X_t} are submartingales.*
(b) *Let $\mu > 0$. Then $e^{\mu t+\sigma W_t}$ is a submartingale.*

Proof: (a) Results from part (a) of Proposition 2.12.3.
(b) It follows from Example 2.12 and part (b) of Proposition 2.12.3. ■

The following result provides important inequalities involving submartingales.

Proposition 2.12.5 (Doob's Submartingale Inequality) (a) *Let X_t be a non-negative submartingale. Then*

$$P(\sup_{s\le t} X_s \ge x) \le \frac{E[X_t]}{x}, \qquad \forall x > 0.$$

(b) *If X_t is a right continuous submartingale, then for any $x > 0$*

$$P(\sup_{s\le t} X_t \ge x) \le \frac{E[X_t^+]}{x},$$

where $X_t^+ = \max\{X_t, 0\}$.

Exercise 2.12.6 *Let $x > 0$. Show the inequalities:*
(a) $P(\sup_{s\le t} W_s^2 \ge x) \le \dfrac{t}{x}$.
(b) $P(\sup_{s\le t} |W_s| \ge x) \le \dfrac{\sqrt{2t/\pi}}{x}$.

Exercise 2.12.7 *Show that* $st\text{-}\lim\limits_{t\to\infty} \dfrac{\sup_{s\le t}|W_s|}{t} = 0$.

Exercise 2.12.8 *Show that for any martingale X_t we have the inequality*

$$P(\sup_{s\le t} X_t^2 > x) \le \frac{E[X_t^2]}{x}, \qquad \forall x > 0.$$

It is worth noting that Doob's inequality implies Markov's inequality. Since $\sup\limits_{s\le t} X_s \ge X_t$, then $P(X_t \ge x) \le P(\sup\limits_{s\le t} X_s \ge x)$. Then Doob's inequality

$$P(\sup_{s\le t} X_s \ge x) \le \frac{E[X_t]}{x}$$

implies Markov's inequality (see Theorem 1.12.8)

$$P(X_t \ge x) \le \frac{E[X_t]}{x}.$$

Exercise 2.12.9 *Let N_t denote the Poisson process and consider the information set $\mathcal{F}_t = \sigma\{N_s; s \leq t\}$.*

(a) *Show that N_t is a submartingale;*

(b) *Is N_t^2 a submartingale?*

Exercise 2.12.10 *It can be shown that for any $0 < \sigma < \tau$ we have the inequality*

$$E\Big[\sum_{\sigma \leq t \leq \tau} \Big(\frac{N_t}{t} - \lambda\Big)^2\Big] \leq \frac{4\tau\lambda}{\sigma^2}\cdot$$

Using this inequality prove that $ms\text{–}\lim\limits_{t\to\infty} \dfrac{N_t}{t} = \lambda.$

Chapter 3

Properties of Stochastic Processes

3.1 Stopping Times

Consider the probability space $(\Omega, \mathcal{F}, P)$ and the filtration $(\mathcal{F}_t)_{t\geq 0}$, i.e.

$$\mathcal{F}_t \subset \mathcal{F}_s \subset \mathcal{F}, \qquad \forall t < s.$$

Assume that the decision to stop playing a game before or at time t is determined by the information $\mathcal{F}_t$ available at time t. Then this decision can be modeled by a random variable $\tau : \Omega \to [0, \infty]$ which satisfies

$$\{\omega; \tau(\omega) \leq t\} \in \mathcal{F}_t.$$

This means that given the information set $\mathcal{F}_t$, we know whether the event $\{\omega; \tau(\omega) \leq t\}$ had occurred or not. We note that the possibility $\tau = \infty$ is included, since the decision to continue the game for ever is also a possible event. However, we ask the condition $P(\tau < \infty) = 1$. A random variable τ with the previous properties is called a *stopping time*.

The next example illustrates a few cases when a decision is or is not a stopping time. In order to accomplish this, think of the situation that τ is the time when some random event related to a given stochastic process occurs first.

Example 3.1.1 *Let $\mathcal{F}_t$ be the information available until time t regarding the evolution of a stock. Assume the price of the stock at time $t = 0$ is \$50 per share. The following decisions are stopping times:*

(a) *Sell the stock when it reaches for the first time the price of \$100 per share;*

(b) *Buy the stock when it reaches for the first time the price of \$10 per share;*

(c) *Sell the stock at the end of the year;*

(d) Sell the stock either when it reaches for the first time \$80 *or at the end of the year.*

(e) Keep the stock either until the initial investment doubles or until the end of the year;

The following decisions are not stopping times:

(f) Sell the stock when it reaches the maximum level it will ever be;

(g) Keep the stock until the initial investment at least doubles.

Part (f) is not a stopping time because it requires information about the future that is not contained in $\mathcal{F}_t$. For part (g), since the initial stock price is $S_0 = \$50$, the general theory of stock prices state

$$P(S_t \geq 2S_0) = P(S_t \geq 100) < 1,$$

i.e. there is a positive probability that the stock never doubles its value. This contradicts the condition $P(\tau = \infty) = 0$. In part (e) there are two conditions; the latter one has the occurring probability equal to 1.

Exercise 3.1.2 *Show that any positive constant, $\tau = c$, is a stopping time with respect to any filtration.*

Exercise 3.1.3 *Let $\tau(\omega) = \inf\{t > 0; |W_t(\omega)| > K\}$, with $K > 0$ constant. Show that τ is a stopping time with respect to the filtration $\mathcal{F}_t = \sigma(W_s; s \leq t)$.*

The random variable τ is called the *first exit time* of the Brownian motion W_t from the interval $(-K, K)$. In a similar way one can define the *first exit time* of the process X_t from the interval (a, b):

$$\tau(\omega) = \inf\{t > 0; X_t(\omega) \notin (a, b)\} = \inf\{t > 0; X_t(\omega) > b \text{ or } X_t(\omega) < a)\}.$$

Let $X_0 < a$. The *first entry time* of X_t in the interval (a, b) is defined as

$$\tau(\omega) = \inf\{t > 0; X_t(\omega) \in (a, b)\} = \inf\{t > 0; X_t(\omega) > a \text{ or } X_t(\omega) < b)\}.$$

If let $b = \infty$, we obtain the *first hitting time* of the level a

$$\tau(\omega) = \inf\{t > 0; X_t(\omega) > a)\}.$$

We shall deal with hitting times in more detail in section 3.3.

Exercise 3.1.4 *Let X_t be a continuous stochastic process. Prove that the first exit time of X_t from the interval (a, b) is a stopping time.*

We shall present in the following some properties regarding operations with stopping times. Consider the notations $\tau_1 \vee \tau_2 = \max\{\tau_1, \tau_2\}$, $\tau_1 \wedge \tau_2 = \min\{\tau_1, \tau_2\}$, $\bar{\tau}_n = \sup_{n\geq 1} \tau_n$ and $\underline{\tau}_n = \inf_{n\geq 1} \tau_n$.

Proposition 3.1.5 *Let τ_1 and τ_2 be two stopping times with respect to the filtration $\mathcal{F}_t$. Then*

1. $\tau_1 \vee \tau_2$
2. $\tau_1 \wedge \tau_2$
3. $\tau_1 + \tau_2$

are stopping times.

Proof: 1. We have

$$\{\omega; \tau_1 \vee \tau_2 \leq t\} = \{\omega; \tau_1 \leq t\} \cap \{\omega; \tau_2 \leq t\} \in \mathcal{F}_t,$$

since $\{\omega; \tau_1 \leq t\} \in \mathcal{F}_t$ and $\{\omega; \tau_2 \leq t\} \in \mathcal{F}_t$. From the subadditivity of P we get

$$\begin{aligned} P\Big(\{\omega; \tau_1 \vee \tau_2 = \infty\}\Big) &= P\Big(\{\omega; \tau_1 = \infty\} \cup \{\omega; \tau_2 = \infty\}\Big) \\ &\leq P\Big(\{\omega; \tau_1 = \infty\}\Big) + P\Big(\{\omega; \tau_2 = \infty\}\Big). \end{aligned}$$

Since

$$P\Big(\{\omega; \tau_i = \infty\}\Big) = 1 - P\Big(\{\omega; \tau_i < \infty\}\Big) = 0, \qquad i = 1, 2,$$

it follows that $P\Big(\{\omega; \tau_1 \vee \tau_2 = \infty\}\Big) = 0$ and hence $P\Big(\{\omega; \tau_1 \vee \tau_2 < \infty\}\Big) = 1$. Then $\tau_1 \vee \tau_2$ is a stopping time.

2. The event $\{\omega; \tau_1 \wedge \tau_2 \leq t\} \in \mathcal{F}_t$ if and only if $\{\omega; \tau_1 \wedge \tau_2 > t\} \in \mathcal{F}_t$.

$$\{\omega; \tau_1 \wedge \tau_2 > t\} = \{\omega; \tau_1 > t\} \cap \{\omega; \tau_2 > t\} \in \mathcal{F}_t,$$

since $\{\omega; \tau_1 > t\} \in \mathcal{F}_t$ and $\{\omega; \tau_2 > t\} \in \mathcal{F}_t$ (the σ-algebra $\mathcal{F}_t$ is closed to complements). The fact that $\tau_1 \wedge \tau_2 < \infty$ almost surely has a similar proof.

3. We note that $\tau_1 + \tau_2 \leq t$ if there is a $c \in (0, t)$ such that

$$\tau_1 \leq c, \qquad \tau_2 \leq t - c.$$

Using that the rational numbers are dense in $\mathbb{R}$, we can write

$$\{\omega; \tau_1 + \tau_2 \leq t\} = \bigcup_{0<c<t, c\in\mathbb{Q}} \Big(\{\omega; \tau_1 \leq c\} \cap \{\omega; \tau_2 \leq t - c\}\Big) \in \mathcal{F}_t,$$

since

$$\{\omega; \tau_1 \leq c\} \in \mathcal{F}_c \subset \mathcal{F}_t, \qquad \{\omega; \tau_2 \leq t - c\} \in \mathcal{F}_{t-c} \subset \mathcal{F}_t.$$

Writing

$$\{\omega; \tau_1 + \tau_2 = \infty\} = \{\omega; \tau_1 = \infty\} \cup \{\omega; \tau_2 = \infty\}$$

yields

$$P\Big(\{\omega; \tau_1 + \tau_2 = \infty\}\Big) \leq P\Big(\{\omega; \tau_1 = \infty\}\Big) + P\Big(\{\omega; \tau_2 = \infty\}\Big) = 0,$$

Hence $P\Big(\{\omega; \tau_1 + \tau_2 < \infty\}\Big) = 1$. It follows that $\tau_1 + \tau_2$ is a stopping time. ■

A filtration $\mathcal{F}_t$ is called *right-continuous* if $\mathcal{F}_t = \bigcap_{n=1}^{\infty} \mathcal{F}_{t+\frac{1}{n}}$, for $t > 0$. This means that the information available at time t is a good approximation for any future infinitesimal information $\mathcal{F}_{t+\epsilon}$; or, equivalently, nothing more can be learned by peeking infinitesimally far into the future.

Exercise 3.1.6 *(a) Let $\mathcal{F}_t = \sigma\{W_s; s \le t\}$, where W_t is a Brownian motion. Show that $\mathcal{F}_t$ is right-continuous.*

(b) Let $\mathcal{N}_t = \sigma\{N_s; s \le t\}$, where N_t is a Poisson motion. Is $\mathcal{N}_t$ right-continuous?

Proposition 3.1.7 *Let $\mathcal{F}_t$ be right-continuous and $(\tau_n)_{n\ge 1}$ be a sequence of bounded stopping times.*

(a) Then $\sup_n \tau$ and $\inf \tau_n$ are stopping times.

(b) If the sequence $(\tau_n)_{n\ge 1}$ converges to τ, $\tau \neq 0$, then τ is a stopping time.

Proof: (a) The fact that $\bar{\tau}_n$ is a stopping time follows from

$$\{\omega; \bar{\tau}_n \le t\} \subset \bigcap_{n\ge 1} \{\omega; \tau_n \le t\} \in \mathcal{F}_t,$$

and from the boundedness, which implies $P(\tau_n < \infty) = 1$.

In order to show that $\underline{\tau}_n$ is a stopping time we shall proceed as in the following. Using that $\mathcal{F}_t$ is right-continuous and closed to complements, it suffices to show that $\{\omega; \underline{\tau}_n \ge t\} \in \mathcal{F}_t$. This follows from

$$\{\omega; \underline{\tau}_n \ge t\} = \bigcap_{n\ge 1} \{\omega; \tau_n > t\} \in \mathcal{F}_t.$$

(b) Let $\lim_{n\to\infty} \tau_n = \tau$. Then there is an increasing (or decreasing) subsequence τ_{n_k} of stopping times that tends to τ, so $\sup \tau_{n_k} = \tau$ (or $\inf \tau_{n_k} = \tau$). Since τ_{n_k} are stopping times, by part (a), it follows that τ is a stopping time. ■

The condition that τ_n is bounded is significant, since if take $\tau_n = n$ as stopping times, then $\sup_n \tau_n = \infty$ with probability 1, which does not satisfy the stopping time definition.

Exercise 3.1.8 *Let τ be a stopping time.*

(a) Let $c \ge 1$ be a constant. Show that $c\tau$ is a stopping time.

(b) Let $f : [0, \infty) \to \mathbb{R}$ be a continuous, increasing function satisfying $f(t) \ge t$. Prove that $f(\tau)$ is a stopping time.

(c) Show that e^{τ} is a stopping time.

Exercise 3.1.9 *Let τ be a stopping time and $c > 0$ a constant. Prove that $\tau + c$ is a stopping time.*

Exercise 3.1.10 *Let a be a constant and define $\tau = \inf\{t \geq 0; W_t = a\}$. Is τ a stopping time?*

Exercise 3.1.11 *Let τ be a stopping time. Consider the following sequence $\tau_n = (m+1)2^{-n}$ if $m2^{-n} \leq \tau < (m+1)2^{-n}$ (stop at the first time of the form $k2^{-n}$ after τ). Prove that τ_n is a stopping time.*

3.2 Stopping Theorem for Martingales

The next result states that in a fair game, the expected final fortune of a gambler, who is using a stopping time to quit the game, is the same as the expected initial fortune. From the financially point of view, the theorem says that if you buy an asset at some initial time and adopt a strategy of deciding when to sell it, then the expected price at the selling time is the initial price; so one cannot make money by buying and selling an asset whose price is a martingale. Fortunately, the price of a stock is not a martingale, and people can still expect to make money buying and selling stocks.

If $(M_t)_{t\geq 0}$ is an $\mathcal{F}_t$-martingale, then taking the expectation in

$$E[M_t|\mathcal{F}_s] = M_s, \qquad \forall s < t$$

and using Example 1.11.2 yields

$$E[M_t] = E[M_s], \qquad \forall s < t.$$

In particular, $E[M_t] = E[M_0]$, for any $t > 0$. The next result states necessary conditions under which this identity holds if t is replaced by any stopping time τ.

Theorem 3.2.1 (Optional Stopping Theorem) *Let $(M_t)_{t\geq 0}$ be a right continuous $\mathcal{F}_t$-martingale and τ be a stopping time with respect to $\mathcal{F}_t$. If either one of the following conditions holds:*

1. *τ is bounded, i.e. $\exists N < \infty$ such that $\tau \leq N$;*
2. *$\exists c > 0$ such that $E[|M_t|] \leq c$, $\forall t > 0$,*

then $E[M_\tau] = E[M_0]$. If M_t is an $\mathcal{F}_t$-submartingale, then $E[M_\tau] \geq E[M_0]$.

Proof: We shall sketch the proof for the first case only. Taking the expectation in relation

$$M_\tau = M_{\tau\wedge t} + (M_\tau - M_t)\mathbf{1}_{\{\tau > t\}},$$

see Exercise 3.2.3, yields

$$E[M_\tau] = E[M_{\tau\wedge t}] + E[M_\tau \mathbf{1}_{\{\tau>t\}}] - E[M_t \mathbf{1}_{\{\tau>t\}}].$$

Since $M_{\tau\wedge t}$ is a martingale, see Exercise 3.2.4 (b), then $E[M_{\tau\wedge t}] = E[M_0]$. The previous relation becomes

$$E[M_\tau] = E[M_0] + E[M_\tau \mathbf{1}_{\{\tau>t\}}] - E[M_t \mathbf{1}_{\{\tau>t\}}], \quad \forall t > 0.$$

Taking the limit yields

$$E[M_\tau] = E[M_0] + \lim_{t\to\infty} E[M_\tau \mathbf{1}_{\{\tau>t\}}] - \lim_{t\to\infty} E[M_t \mathbf{1}_{\{\tau>t\}}]. \tag{3.2.1}$$

We shall show that both limits are equal to zero.

Since $|M_\tau \mathbf{1}_{\{\tau>t\}}| \leq |M_\tau|$, $\forall t > 0$, and M_τ is integrable, see Exercise 3.2.4 (a), by the dominated convergence theorem we have

$$\lim_{t\to\infty} E[M_\tau \mathbf{1}_{\{\tau>t\}}] = \lim_{t\to\infty} \int_\Omega M_\tau \mathbf{1}_{\{\tau>t\}}\, dP = \int_\Omega \lim_{t\to\infty} M_\tau \mathbf{1}_{\{\tau>t\}}\, dP = 0.$$

For the second limit

$$\lim_{t\to\infty} E[M_t \mathbf{1}_{\{\tau>t\}}] = \lim_{t\to\infty} \int_\Omega M_t \mathbf{1}_{\{\tau>t\}}\, dP = 0,$$

since for $t > N$ the integrand vanishes. Hence relation (3.2.1) yields $E[M_\tau] = E[M_0]$. ■

It is worth noting that the previous theorem is a special case of the more general Optional Stopping Theorem of Doob:

Theorem 3.2.2 *Let M_t be a right continuous martingale and σ, τ be two bounded stopping times, with $\sigma \leq \tau$. Then M_σ, M_τ are integrable and*

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

In particular, taking expectations, we have

$$E[M_\tau] = E[M_\sigma] \quad a.s.$$

In the case when M_t is a submartingale then $E[M_\tau] \geq E[M_\sigma]$ a.s.

Exercise 3.2.3 *Show that*

$$M_\tau = M_{\tau\wedge t} + (M_\tau - M_t)\mathbf{1}_{\{\tau>t\}},$$

where

$$\mathbf{1}_{\{\tau>t\}}(\omega) = \begin{cases} 1, & \tau(\omega) > t; \\ 0, & \tau(\omega) \leq t \end{cases}$$

is the indicator function of the set $\{\tau > t\}$.

Exercise 3.2.4 *Let M_t be a right continuous martingale and τ be a stopping time. Show that*

(a) *M_τ is integrable;*

(b) *$M_{\tau\wedge t}$ is a martingale.*

Exercise 3.2.5 *Show that if let $\sigma = 0$ in Theorem* 3.2.2 *yields Theorem* 3.2.1.

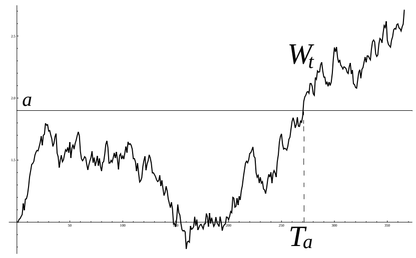


Figure 3.1: *The first hitting time T_a given by $W_{T_a} = a$.*

3.3 The First Passage of Time

The first passage of time is a particular type of hitting time, which is useful in finance when studying barrier options and lookback options. For instance, knock-in options enter into existence when the stock price hits for the first time a certain barrier before option maturity. A lookback option is priced using the maximum value of the stock until the present time. The stock price is not a Brownian motion, but it depends on one. Hence the need for studying the hitting time for the Brownian motion.

The first result deals with the first hitting time for a Brownian motion to reach the barrier $a \in \mathbb{R}$, see Fig.3.1.

Lemma 3.3.1 *Let T_a be the first time the Brownian motion W_t hits a. Then the distribution function of T_a is given by*

$$\boxed{P(T_a \le t) = \frac{2}{\sqrt{2\pi}} \int_{|a|/\sqrt{t}}^{\infty} e^{-y^2/2}\, dy.}$$

Proof: If A and B are two events, then

$$\begin{aligned} P(A) &= P(A \cap B) + P(A \cap \overline{B}) \\ &= P(A|B)P(B) + P(A|\overline{B})P(\overline{B}). \end{aligned} \tag{3.3.2}$$

Let $a > 0$. Using formula (3.3.2) for $A = \{\omega; W_t(\omega) \ge a\}$ and $B = \{\omega; T_a(\omega) \le t\}$ yields

$$\begin{aligned} P(W_t \ge a) &= P(W_t \ge a | T_a \le t) P(T_a \le t) \\ &\quad + P(W_t \ge a | T_a > t) P(T_a > t) \end{aligned} \tag{3.3.3}$$

If $T_a > t$, the Brownian motion did not reach the barrier a yet, so we must have $W_t < a$. Therefore

$$P(W_t \ge a | T_a > t) = 0.$$

If $T_a \leq t$, then $W_{T_a} = a$. Since the Brownian motion is a Markov process, it starts fresh at T_a. Due to symmetry of the density function of a normal variable, W_t has equal chances to go up or go down after the time interval $t - T_a$. It follows that

$$P(W_t \geq a | T_a \leq t) = \frac{1}{2}.$$

Substituting into (3.3.3) yields

$$\begin{aligned} P(T_a \leq t) &= 2P(W_t \geq a) \\ &= \frac{2}{\sqrt{2\pi t}} \int_a^\infty e^{-x^2/(2t)}\, dx = \frac{2}{\sqrt{2\pi}} \int_{a/\sqrt{t}}^\infty e^{-y^2/2}\, dy. \end{aligned}$$

If $a < 0$, symmetry implies that the distribution of T_a is the same as that of T_{-a}, so we get

$$P(T_a \leq t) = P(T_{-a} \leq t) = \frac{2}{\sqrt{2\pi}} \int_{-a/\sqrt{t}}^\infty e^{-y^2/2}\, dy.$$

■

Remark 3.3.2 *The previous proof is based on a more general principle called the Reflection Principle: If τ is a stopping time for the Brownian motion W_t, then the Brownian motion reflected at τ is also a Brownian motion.*

Theorem 3.3.3 *Let $a \in \mathbb{R}$ be fixed. Then the Brownian motion hits a (in a finite amount of time) with probability 1.*

Proof: The probability that W_t hits a (in a finite amount of time) is

$$\begin{aligned} P(T_a < \infty) &= \lim_{t\to\infty} P(T_a \leq t) = \lim_{t\to\infty} \frac{2}{\sqrt{2\pi}} \int_{|a|/\sqrt{t}}^\infty e^{-y^2/2}\, dy \\ &= \frac{2}{\sqrt{2\pi}} \int_0^\infty e^{-y^2/2}\, dy = 1, \end{aligned}$$

where we used the well known integral

$$\int_0^\infty e^{-y^2/2}\, dy = \frac{1}{2} \int_{-\infty}^\infty e^{-y^2/2}\, dy = \frac{1}{2}\sqrt{2\pi}.$$

■

The previous result stated that the Brownian motion hits the barrier a almost surely. The next result shows that the expected time to hit the barrier is infinite.

Proposition 3.3.4 *The random variable T_a has a Pearson 5 distribution given by*

$$\boxed{p(t) = \frac{|a|}{\sqrt{2\pi}} e^{-\frac{a^2}{2t}} t^{-\frac{3}{2}}, \qquad t > 0.}$$

It has the mean $E[T_a] = \infty$ and the mode $\dfrac{a^2}{3}$.

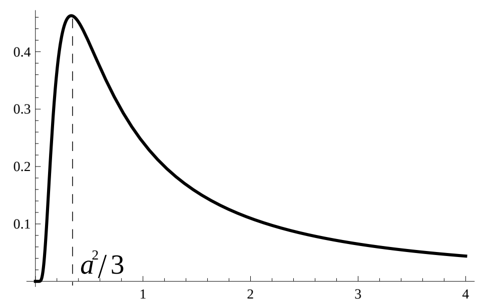


Figure 3.2: *The distribution of the first hitting time* T_a.

Proof: Differentiating in the formula of distribution function[1]

$$F_{T_a}(t) = P(T_a \le t) = \frac{2}{\sqrt{2\pi}} \int_{a/\sqrt{t}}^{\infty} e^{-y^2/2}\, dy$$

yields the following probability density function

$$p(t) = \frac{dF_{T_a}(t)}{dt} = \frac{a}{\sqrt{2\pi}} e^{-\frac{a^2}{2t}} t^{-\frac{3}{2}}, \qquad t > 0.$$

This is a Pearson 5 distribution with parameters $\alpha = 1/2$ and $\beta = a^2/2$. The expectation is

$$E[T_a] \;=\; \int_0^\infty tp(t)\, dt = \frac{a}{\sqrt{2\pi}} \int_0^\infty \frac{1}{\sqrt{t}} e^{-\frac{a^2}{2t}}\, dt.$$

Using the inequality $e^{-\frac{a^2}{2t}} > 1 - \dfrac{a^2}{2t}$, $t > 0$, we have the estimation

$$E[T_a] \;>\; \frac{a}{\sqrt{2\pi}} \int_0^\infty \frac{1}{\sqrt{t}}\, dt - \frac{a^3}{2\sqrt{2\pi}} \int_0^\infty \frac{1}{t^{3/2}}\, dt = \infty, \tag{3.3.4}$$

since $\int_0^\infty \frac{1}{\sqrt{t}}\, dt$ is divergent and $\int_0^\infty \frac{1}{t^{3/2}}\, dt$ is convergent.

The mode of T_a is given by

$$\frac{\beta}{\alpha + 1} = \frac{a^2}{2(\frac{1}{2} + 1)} = \frac{a^2}{3}.$$

■

[1]One may use Leibnitz's formula $\dfrac{d}{dt} \displaystyle\int_{\varphi(t)}^{\psi(t)} f(u)\, du = f(\psi(t))\psi'(t) - f(\varphi(t))\varphi'(t)$.

Remark 3.3.5 *The distribution has a peak at $a^2/3$. Then if we need to pick a small time interval $[t-dt, t+dt]$ in which the probability that the Brownian motion hits the barrier a is maximum, we need to choose $t = a^2/3$.*

Remark 3.3.6 *Formula (3.3.4) states that the expected waiting time for W_t to reach the barrier a is infinite. However, the expected waiting time for the Brownian motion W_t to hit either a or $-a$ is finite, see Exercise 3.3.9.*

Corollary 3.3.7 *A Brownian motion process returns to the origin in a finite amount time with probability 1.*

Proof: Choose $a = 0$ and apply Theorem 3.3.3. ∎

Exercise 3.3.8 *Try to apply the proof of Lemma 3.3.1 for the following stochastic processes*

(*a*) $X_t = \mu t + \sigma W_t$, *with* $\mu, \sigma > 0$ *constants;*

(*b*) $X_t = \displaystyle\int_0^t W_s\, ds$.

Where is the difficulty?

Exercise 3.3.9 *Let $a > 0$ and consider the hitting time*

$$\tau_a = \inf\{t > 0; W_t = a \text{ or } W_t = -a\} = \inf\{t > 0; |W_t| = a\}.$$

Prove that $E[\tau_a] = a^2$.

Exercise 3.3.10 (*a*) *Show that the distribution function of the process*

$$X_t = \max_{s\in[0,t]} W_s$$

is given by

$$P(X_t \le a) = \frac{2}{\sqrt{2\pi}} \int_0^{a/\sqrt{t}} e^{-y^2/2}\, dy.$$

(*b*) *Show that $E[X_t] = \sqrt{2t/\pi}$ and $Var(X_t) = t\Big(1 - \dfrac{2}{\pi}\Big)$.*

Exercise 3.3.11 (*a*) *Find the distribution of $Y_t = |W_t|$, $t \ge 0$;*

(*b*) *Show that $E\big[\max\limits_{0\le t\le T} |W_t|\big] = \sqrt{\dfrac{\pi T}{2}}$.*

The fact that a Brownian motion returns to the origin or hits a barrier almost surely is a property characteristic to the first dimension only. The next result states that in larger dimensions this is no longer possible.

Theorem 3.3.12 *Let $(a,b) \in \mathbb{R}^2$. The 2-dimensional Brownian motion $W(t) = \big(W_1(t), W_2(t)\big)$ (with $W_1(t)$ and $W_2(t)$ independent) hits the point (a,b) with probability zero. The same result is valid for any n-dimensional Brownian motion, with $n \geq 2$.*

However, if the point (a,b) is replaced by the disk $D_\epsilon(\mathbf{x}_0) = \{x \in \mathbb{R}^2; |\mathbf{x} - \mathbf{x}_0| \leq \epsilon\}$, there is a difference in the behavior of the Brownian motion from $n = 2$ to $n > 2$.

Theorem 3.3.13 *The 2-dimensional Brownian motion $W(t) = \big(W_1(t), W_2(t)\big)$ hits the disk $D_\epsilon(\mathbf{x}_0)$ with probability one.*

Theorem 3.3.14 *Let $n > 2$. The n-dimensional Brownian motion $W(t) = \big(W_1(t), \cdots W_n(t)\big)$ hits the ball $D_\epsilon(\mathbf{x}_0)$ with probability*

$$P = \Big(\frac{|\mathbf{x}_0|}{\epsilon}\Big)^{2-n} < 1.$$

The previous results can be stated by saying that that Brownian motion is transient in $\mathbb{R}^n$, for $n > 2$. If $n = 2$ the previous probability equals 1. We shall come back with proofs to the aforementioned results in a later chapter.

Remark 3.3.15 *If life spreads according to a Brownian motion, the aforementioned results explain why life is more extensive on earth rather than in space. The probability for a form of life to reach a planet of radius R situated at distance d is $\frac{R}{d}$. Since d is large the probability is very small, unlike in the plane, where the probability is always 1.*

Exercise 3.3.16 *Is the one-dimensional Brownian motion transient or recurrent in $\mathbb{R}$?*

3.4 The Arc-sine Laws

In this section we present a few results which provide certain probabilities related with the behavior of a Brownian motion in terms the arc-sine of a quotient of two time instances. These results are generally known under the name of *law of arc-sines.*

The following result will be used in the proof of the first law of Arc-sine.

Proposition 3.4.1 (*a*) *If $X : \Omega \to \mathbb{N}$ is a discrete random variable, then for any subset $A \subset \Omega$, we have*

$$P(A) = \sum_{x \in \mathbb{N}} P(A|X = x)P(X = x).$$

(*b*) *If $X : \Omega \to \mathbb{R}$ is a continuous random variable, then*

$$P(A) = \int P(A|X = x)dP = \int P(A|X = x)f_X(x)\,dx.$$

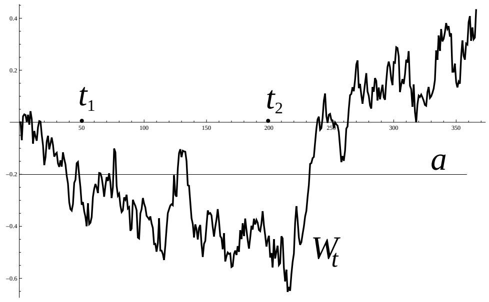


Figure 3.3: *The event $A(a; t_1, t_2)$ in the Law of Arc-sine.*

Proof: (a) The sets $X^{-1}(x) = \{X = x\} = \{\omega; X(\omega) = x\}$ form a partition of the sample space Ω, i.e.:

(i) $\Omega = \bigcup_x X^{-1}(x)$;

(ii) $X^{-1}(x) \cap X^{-1}(y) = \emptyset$ for $x \neq y$.

Then $A = \bigcup\limits_x \Big(A \cap X^{-1}(x)\Big) = \bigcup\limits_x \Big(A \cap \{X = x\}\Big)$, and hence

$$\begin{aligned} P(A) &= \sum_x P\Big(A \cap \{X = x\}\Big) \\ &= \sum_x \frac{P(A \cap \{X = x\})}{P(\{X = x\})} P(\{X = x\}) \\ &= \sum_x P(A|X = x) P(X = x). \end{aligned}$$

(b) In the case when X is continuous, the sum is replaced by an integral and the probability $P(\{X = x\})$ by $f_X(x)dx$, where f_X is the density function of X. ■

The *zero set* of a Brownian motion W_t is defined by $\{t \geq 0; W_t = 0\}$. Since W_t is continuous, the zero set is closed with no isolated points almost surely. The next result deals with the probability that the zero set does not intersect the interval (t_1, t_2).

Theorem 3.4.2 (The law of Arc-sine) *The probability that a Brownian motion W_t does not have any zeros in the interval (t_1, t_2) is equal to*

$$\boxed{P(W_t \neq 0, t_1 \leq t \leq t_2) = \frac{2}{\pi} \arcsin \sqrt{\frac{t_1}{t_2}}.}$$

Proof: Let $A(a; t_1, t_2)$ denote the event that the Brownian motion W_t takes on the value a between t_1 and t_2. In particular, $A(0; t_1, t_2)$ denotes the event that W_t has (at least) a zero between t_1 and t_2. Substituting $A = A(0; t_1, t_2)$ and $X = W_{t_1}$ into the formula provided by Proposition 3.4.1

$$P(A) = \int P(A|X = x) f_X(x)\, dx$$

yields

$$\begin{aligned} P\big(A(0;t_1,t_2)\big) &= \int P\big(A(0;t_1,t_2)|W_{t_1} = x\big) f_{W_{t_1}}(x)\, dx \\ &= \frac{1}{\sqrt{2\pi t_1}} \int_{-\infty}^{\infty} P\big(A(0;t_1,t_2)|W_{t_1} = x\big) e^{-\frac{x^2}{2t_1}}\, dx \end{aligned} \tag{3.4.5}$$

Using the properties of W_t with respect to time translation and symmetry we have

$$\begin{aligned} P\big(A(0;t_1,t_2)|W_{t_1} = x\big) &= P\big(A(0;0,t_2-t_1)|W_0 = x\big) \\ &= P\big(A(-x;0,t_2-t_1)|W_0 = 0\big) \\ &= P\big(A(|x|;0,t_2-t_1)|W_0 = 0\big) \\ &= P\big(A(|x|;0,t_2-t_1)\big) \\ &= P\big(T_{|x|} \le t_2 - t_1\big), \end{aligned}$$

the last identity stating that W_t hits $|x|$ before $t_2 - t_1$. Using Lemma 3.3.1 yields

$$P\big(A(0;t_1,t_2)|W_{t_1} = x\big) = \frac{2}{\sqrt{2\pi(t_2-t_1)}} \int_{|x|}^{\infty} e^{-\frac{y^2}{2(t_2-t_1)}}\, dy.$$

Substituting into (3.4.5) we obtain

$$\begin{aligned} P\big(A(0;t_1,t_2)\big) &= \frac{1}{\sqrt{2\pi t_1}} \int_{-\infty}^{\infty} \Big(\frac{2}{\sqrt{2\pi(t_2-t_1)}} \int_{|x|}^{\infty} e^{-\frac{y^2}{2(t_2-t_1)}}\, dy \Big) e^{-\frac{x^2}{2t_1}}\, dx \\ &= \frac{1}{\pi\sqrt{t_1(t_2-t_1)}} \int_0^{\infty} \int_{|x|}^{\infty} e^{-\frac{y^2}{2(t_2-t_1)} - \frac{x^2}{2t_1}}\, dy dx. \end{aligned}$$

The above integral can be evaluated to get (see Exercise 3.4.3)

$$P\big(A(0;t_1,t_2)\big) = 1 - \frac{2}{\pi} \arcsin\sqrt{\frac{t_1}{t_2}}.$$

Using $P(W_t \neq 0, t_1 \le t \le t_2) = 1 - P\big(A(0;t_1,t_2)\big)$ we obtain the desired result. ■

Exercise 3.4.3 *Use polar coordinates to show*

$$\frac{1}{\pi\sqrt{t_1(t_2-t_1)}} \int_0^{\infty} \int_{|x|}^{\infty} e^{-\frac{y^2}{2(t_2-t_1)} - \frac{x^2}{2t_1}}\, dy dx = 1 - \frac{2}{\pi} \arcsin\sqrt{\frac{t_1}{t_2}}.$$

Exercise 3.4.4 *Find the probability that a 2-dimensional Brownian motion $W(t) = \big(W_1(t), W_2(t)\big)$ stays in the same quadrant for the time interval $t \in (t_1, t_2)$.*

Exercise 3.4.5 *Find the probability that a Brownian motion W_t does not take the value a in the interval (t_1, t_2).*

Exercise 3.4.6 *Let $a \neq b$. Find the probability that a Brownian motion W_t does not take any of the values $\{a, b\}$ in the interval (t_1, t_2). Formulate and prove a generalization.*

We provide below without proof a few similar results dealing with arc-sine probabilities. The first result deals with the amount of time spent by a Brownian motion on the positive half-axis.

Theorem 3.4.7 (Arc-sine Law of Lévy) *Let $L_t^+ = \int_0^t sgn^+ W_s\, ds$ be the amount of time a Brownian motion W_t is positive during the time interval $[0, t]$. Then*

$$P(L_t^+ \leq \tau) = \frac{2}{\pi} \arcsin \sqrt{\frac{\tau}{t}}.$$

The next result deals with the Arc-sine law for the last exit time of a Brownian motion from 0.

Theorem 3.4.8 (Arc-sine Law of exit from 0) *Let $\gamma_t = \sup\{0 \leq s \leq t; W_s = 0\}$. Then*

$$P(\gamma_t \leq \tau) = \frac{2}{\pi} \arcsin \sqrt{\frac{\tau}{t}}, \quad 0 \leq \tau \leq t.$$

The Arc-sine law for the time the Brownian motion attains its maximum on the interval $[0, t]$ is given by the next result.

Theorem 3.4.9 (Arc-sine Law of maximum) *Let $M_t = \max\limits_{0 \leq s \leq t} W_s$ and define*

$$\theta_t = \sup\{0 \leq s \leq t; W_s = M_t\}.$$

Then

$$P(\theta_t \leq s) = \frac{2}{\pi} \arcsin \sqrt{\frac{s}{t}}, \qquad 0 \leq s \leq t, t > 0.$$

3.5 More on Hitting Times

In this section we shall deal with results regarding hitting times of Brownian motion with drift. They will be useful in the sequel when pricing barrier options.

Theorem 3.5.1 *Let $X_t = \mu t + W_t$ denote a Brownian motion with nonzero drift rate μ, and consider $\alpha, \beta > 0$. Then*

$$\boxed{P(X_t \text{ goes up to } \alpha \text{ before down to } -\beta) = \frac{e^{2\mu\beta} - 1}{e^{2\mu\beta} - e^{-2\mu\alpha}}.}$$

Proof: Let $T = \inf\{t > 0; X_t = \alpha \text{ or } X_t = -\beta\}$ be the first exit time of X_t from the interval $(-\beta, \alpha)$, which is a stopping time, see Exercise 3.1.4. The exponential process

$$M_t = e^{cW_t - \frac{c^2}{2}t}, \qquad t \geq 0$$

is a martingale, see Exercise 2.2.4(c). Then $E[M_t] = E[M_0] = 1$. By the Optional Stopping Theorem (see Theorem 3.2.1), we get $E[M_T] = 1$. This can be written as

$$1 \;=\; E[e^{cW_T - \frac{1}{2}c^2 T}] = E[e^{cX_T - (c\mu + \frac{1}{2}c^2)T}]. \tag{3.5.6}$$

Choosing $c = -2\mu$ yields $E[e^{-2\mu X_T}] = 1$. Since the random variable X_T takes only the values α and $-\beta$, if let $p_\alpha = P(X_T = \alpha)$, the previous relation becomes

$$e^{-2\mu\alpha} p_\alpha + e^{2\mu\beta}(1 - p_\alpha) = 1.$$

Solving for p_α yields

$$p_\alpha = \frac{e^{2\mu\beta} - 1}{e^{2\mu\beta} - e^{-2\mu\alpha}}. \tag{3.5.7}$$

Noting that

$$p_\alpha = P(X_t \text{ goes up to } \alpha \text{ before down to } -\beta)$$

leads to the desired answer. ∎

It is worth noting how the previous formula changes in the case when the drift rate is zero, i.e. when $\mu = 0$, and $X_t = W_t$. The previous probability is computed by taking the limit $\mu \to 0$ and using L'Hospital's rule

$$\lim_{\mu \to 0} \frac{e^{2\mu\beta} - 1}{e^{2\mu\beta} - e^{-2\mu\alpha}} = \lim_{\mu \to 0} \frac{2\beta e^{2\mu\beta}}{2\beta e^{2\mu\beta} + 2\alpha e^{-2\mu\alpha}} = \frac{\beta}{\alpha + \beta}.$$

Hence

$$\boxed{P(W_t \text{ goes up to } \alpha \text{ before down to } -\beta) = \frac{\beta}{\alpha + \beta}.}$$

Taking the limit $\beta \to \infty$ we recover the following result

$$\boxed{P(W_t \text{ hits } \alpha) = 1.}$$

If $\alpha = \beta$ we obtain

$$\boxed{P(W_t \text{ goes up to } \alpha \text{ before down to } -\alpha) = \frac{1}{2},}$$

which shows that the Brownian motion is equally likely to go up or down an amount α in a given time interval.

If T_α and T_β denote the times when the process X_t reaches α and β, respectively, then the aforementioned probabilities can be written using inequalities. For instance the first identity becomes

$$P(T_\alpha \leq T_{-\beta}) = \frac{e^{2\mu\beta} - 1}{e^{2\mu\beta} - e^{-2\mu\alpha}}.$$

Exercise 3.5.2 *Let $X_t = \mu t + W_t$ denote a Brownian motion with nonzero drift rate μ, and consider $\alpha > 0$.*
(a) If $\mu > 0$ show that

$$P(X_t \text{ goes up to } \alpha) = 1.$$

(b) If $\mu < 0$ show that

$$P(X_t \text{ goes up to } \alpha) = e^{2\mu\alpha} < 1.$$

Formula (a) can be written equivalently as

$$P(\sup_{t\geq 0}(W_t + \mu t) \geq \alpha) = 1, \qquad \mu \geq 0,$$

while formula (b) becomes

$$P(\sup_{t\geq 0}(W_t + \mu t) \geq \alpha) = e^{2\mu\alpha}, \qquad \mu < 0,$$

or

$$P(\sup_{t\geq 0}(W_t - \gamma t) \geq \alpha) = e^{-2\gamma\alpha}, \qquad \gamma > 0,$$

which is known as one of the Doob's inequalities. This can be also described in terms of stopping times as follows. Define the stopping time $\tau_\alpha = \inf\{t > 0; W_t - \gamma t \geq \alpha\}$. Using

$$P(\tau_\alpha < \infty) = P\big(\sup_{t\geq 0}(W_t - \gamma t) \geq \alpha\big)$$

yields the identities

$$\begin{aligned} P(\tau_\alpha < \infty) &= e^{-2\alpha\gamma}, \qquad \gamma > 0, \\ P(\tau_\alpha < \infty) &= 1, \qquad \gamma \leq 0. \end{aligned}$$

Exercise 3.5.3 *Let $X_t = \mu t + W_t$ denote a Brownian motion with nonzero drift rate μ, and consider $\beta > 0$. Show that the probability that X_t never hits $-\beta$ is given by*

$$\begin{cases} 1 - e^{-2\mu\beta}, & \text{if } \mu > 0 \\ 0, & \text{if } \mu < 0. \end{cases}$$

Recall that T is the first time when the process X_t hits α or $-\beta$.

Exercise 3.5.4 *(a) Show that*

$$E[X_T] = \frac{\alpha e^{2\mu\beta} + \beta e^{-2\mu\alpha} - \alpha - \beta}{e^{2\mu\beta} - e^{-2\mu\alpha}}.$$

(b) Find $E[X_T^2]$;
(c) Compute $Var(X_T)$.

The next result deals with the time one has to wait (in expectation) for the process $X_t = \mu t + W_t$ to reach either α or $-\beta$.

Proposition 3.5.5 *The expected value of T is*

$$E[T] = \frac{\alpha e^{2\mu\beta} + \beta e^{-2\mu\alpha} - \alpha - \beta}{\mu(e^{2\mu\beta} - e^{-2\mu\alpha})}.$$

Proof: Using that W_t is a martingale, with $E[W_t] = E[W_0] = 0$, applying the Optional Stopping Theorem, Theorem 3.2.1, yields

$$0 = E[W_T] = E[X_T - \mu T] = E[X_T] - \mu E[T].$$

Then by Exercise 3.5.4(a) we get

$$E[T] = \frac{E[X_T]}{\mu} = \frac{\alpha e^{2\mu\beta} + be^{-2\mu\alpha} - \alpha - \beta}{\mu(e^{2\mu\beta} - e^{-2\mu\alpha})}.$$

■

Exercise 3.5.6 *Take the limit $\mu \to 0$ in the formula provided by Proposition 3.5.5 to find the expected time for a Brownian motion to hit either α or $-\beta$.*

Exercise 3.5.7 *Find $E[T^2]$ and $Var(T)$.*

Exercise 3.5.8 (Wald's identities) *Let T be a finite stopping time for the Brownian motion W_t. Show that*

(a) $E[W_T] = 0$;

(b) $E[W_T^2] = E[T]$.

The previous techniques can be also applied to right continuous martingales. Let $a > 0$ and consider the hitting time of the Poisson process of the barrier a

$$\tau = \inf\{t > 0; N_t \geq a\}.$$

Proposition 3.5.9 *The expected waiting time for N_t to reach the barrier a is $E[\tau] = \frac{a}{\lambda}$.*

Proof: Since $M_t = N_t - \lambda t$ is a right continuous martingale, by the Optional Stopping Theorem $E[M_\tau] = E[M_0] = 0$. Then $E[N_\tau - \lambda\tau] = 0$ and hence $E[\tau] = \frac{1}{\lambda}E[N_\tau] = \frac{a}{\lambda}$.
■

3.6 The Inverse Laplace Transform Method

In this section we shall use the Optional Stopping Theorem in conjunction with the inverse Laplace transform to obtain the probability density for hitting times.

A. The case of standard Brownian motion Let $x > 0$. The first hitting time $\tau = T_x = \inf\{t > 0; W_t = x\}$ is a stopping time. Since $M_t = e^{cW_t - \frac{1}{2}c^2 t}$, $t \geq 0$, is a martingale, with $E[M_t] = E[M_0] = 1$, by the Optional Stopping Theorem, see Theorem 3.2.1, we have

$$E[M_\tau] = 1.$$

This can be written equivalently as $E[e^{cW_\tau} e^{-\frac{1}{2}c^2\tau}] = 1$. Using $W_\tau = x$, we get

$$E[e^{-\frac{1}{2}c^2\tau}] = e^{-cx}.$$

It is worth noting that $c > 0$. This is implied from the fact that $e^{-\frac{1}{2}c^2\tau} < 1$ and $\tau, x > 0$.
Substituting $s = \frac{1}{2}c^2$, the previous relation becomes

$$E[e^{-s\tau}] = e^{-\sqrt{2s}x}. \tag{3.6.8}$$

This relation has a couple of useful applications.

Proposition 3.6.1 *The moments of the first hitting time are all infinite $E[\tau^n] = \infty$, $n \geq 1$.*

Proof: The nth moment of τ can be obtained by differentiating and taking $s = 0$

$$\frac{d^n}{ds^n} E[e^{-s\tau}]\Big|_{s=0} = E[(-\tau)^n e^{-s\tau}]\Big|_{s=0} = (-1)^n E[\tau^n].$$

Using (3.6.8) yields

$$E[\tau^n] = (-1)^n \frac{d^n}{ds^n} e^{-\sqrt{2s}x}\Big|_{s=0}.$$

Since by induction we have

$$\frac{d^n}{ds^n} e^{-\sqrt{2s}x} = (-1)^n e^{-\sqrt{2s}x} \sum_{k=0}^{n-1} \frac{M_k}{2^{r_k/2}} \frac{x^{n-k}}{s^{(n+k)/2}},$$

with M_k, r_k positive integers, it easily follows that $E[\tau^n] = \infty$.

For instance, in the case $n = 1$ we have

$$E[\tau] = -\frac{d}{ds} e^{-\sqrt{2s}x}\Big|_{s=0} = \lim_{s \to 0^+} e^{-\sqrt{2s}x} \frac{x}{2\sqrt{2sx}} = +\infty.$$

■

Another application involves the inverse Laplace transform to get the probability density. This way we can retrieve the result of Proposition 3.3.4.

Proposition 3.6.2 *The probability density of the hitting time τ is given by*

$$p(t) = \frac{|x|}{\sqrt{2\pi t^3}} e^{-\frac{x^2}{2t}}, \qquad t > 0. \tag{3.6.9}$$

Proof: Let $x > 0$. The expectation

$$E[e^{-s\tau}] = \int_0^\infty e^{-s\tau} p(\tau)\, d\tau = \mathcal{L}\{p(\tau)\}(s)$$

is the Laplace transform of $p(\tau)$. Applying the inverse Laplace transform yields

$$\begin{aligned} p(\tau) &= \mathcal{L}^{-1}\{E[e^{-s\tau}]\}(\tau) = \mathcal{L}^{-1}\{e^{-\sqrt{2s}x}\}(\tau) \\ &= \frac{x}{\sqrt{2\pi\tau^3}} e^{-\frac{x^2}{2\tau}}, \qquad \tau > 0. \end{aligned}$$

In the case $x < 0$ we obtain

$$p(\tau) = \frac{-x}{\sqrt{2\pi\tau^3}} e^{-\frac{x^2}{2\tau}}, \qquad \tau > 0,$$

which leads to (3.6.9). ∎

The computation on the inverse Laplace transform $\mathcal{L}^{-1}\{e^{-\sqrt{2s}x}\}(\tau)$ is beyound the goal of this book. The reader can obtain the value of this inverse Laplace transform using the Mathematica software. However, the more mathematical interested reader is reffered to consult the method of complex integration in a book on inverse Laplace transforms.

Another application of formula (3.6.8) is the following inequality.

Proposition 3.6.3 (Chernoff bound) *Let τ denote the first hitting time when the Brownian motion W_t hits the barrier x, $x > 0$. Then*

$$P(\tau \le \lambda) \le e^{-\frac{x^2}{2\lambda}}, \qquad \forall \lambda > 0.$$

Proof: Let $s = -t$ in the part 2 of Theorem 1.12.10 and use (3.6.8) to get

$$P(\tau \le \lambda) \le \frac{E[e^{tX}]}{e^{\lambda t}} = \frac{E[e^{-sX}]}{e^{-\lambda s}} = e^{\lambda s - x\sqrt{2s}}, \qquad \forall s > 0.$$

Then $P(\tau \le \lambda) \le e^{\min_{s>0} f(s)}$, where $f(s) = \lambda s - x\sqrt{2s}$. Since $f'(s) = \lambda - \frac{x}{\sqrt{2s}}$, then $f(s)$ reaches its minimum at the critical point $s_0 = \frac{x^2}{2\lambda^2}$. The minimum value is

$$\min_{s>0} f(s) = f(s_0) = -\frac{x^2}{2\lambda}.$$

Substituting in the previous inequality leads to the required result. ∎

B. The case of Brownian motion with drift

Consider the Brownian motion with drift $X_t = \mu t + \sigma W_t$, with $\mu, \sigma > 0$. Let

$$\tau = \inf\{t > 0; X_t = x\}$$

denote the first hitting time of the barrier x, with $x > 0$. We shall compute the distribution of the random variable τ and its first two moments.

Applying the Optional Stopping Theorem (Theorem 3.2.1) to the martingale $M_t = e^{cW_t - \frac{1}{2}c^2 t}$ yields

$$E[M_\tau] = E[M_0] = 1.$$

Using that $W_\tau = \frac{1}{\sigma}(X_\tau - \mu\tau)$ and $X_\tau = x$, the previous relation becomes

$$E[e^{-(\frac{c\mu}{\sigma} + \frac{1}{2}c^2)\tau}] = e^{-\frac{c}{\sigma}x}. \tag{3.6.10}$$

Substituting $s = \dfrac{c\mu}{\sigma} + \dfrac{1}{2}c^2$ and completing to a square yields

$$2s + \frac{\mu^2}{\sigma^2} = \Big(c + \frac{\mu}{\sigma}\Big)^2.$$

Solving for c we get the solutions

$$c = -\frac{\mu}{\sigma} + \sqrt{2s + \frac{\mu^2}{\sigma^2}}, \qquad c = -\frac{\mu}{\sigma} - \sqrt{2s + \frac{\mu^2}{\sigma^2}}.$$

Assume $c < 0$. Then substituting the second solution into (3.6.10) yields

$$E[e^{-s\tau}] = e^{\frac{1}{\sigma^2}(\mu + \sqrt{2s\sigma^2 + \mu^2})x}.$$

This relation is contradictory since $e^{-s\tau} < 1$ while $e^{\frac{1}{\sigma^2}(\mu + \sqrt{2s\sigma + \mu^2})x} > 1$, where we used that $s, x, \tau > 0$. Hence it follows that $c > 0$. Substituting the first solution into (3.6.10) leads to

$$E[e^{-s\tau}] = e^{\frac{1}{\sigma^2}(\mu - \sqrt{2s\sigma^2 + \mu^2})x}.$$

We arrived at the following result:

Proposition 3.6.4 *Assume $\mu, x > 0$. Let τ be the time the process $X_t = \mu t + \sigma W_t$ hits x for the first time. Then we have*

$$E[e^{-s\tau}] = e^{\frac{1}{\sigma^2}(\mu - \sqrt{2s\sigma^2 + \mu^2})x}, \qquad s > 0. \tag{3.6.11}$$

Proposition 3.6.5 *Let τ be the time the process $X_t = \mu t + \sigma W_t$ hits x, with $x > 0$ and $\mu > 0$.*

(a) Then the density function of τ is given by

$$p(\tau) = \frac{x}{\sigma\sqrt{2\pi}\tau^{3/2}} e^{-\frac{(x-\mu\tau)^2}{2\tau\sigma^2}}, \qquad \tau > 0. \tag{3.6.12}$$

(*b*) *The mean and variance of* τ *are*

$$E[\tau] = \frac{x}{\mu}, \qquad Var(\tau) = \frac{x\sigma^2}{\mu^3}.$$

Proof: (*a*) Let $p(\tau)$ be the density function of τ. Since

$$E[e^{-s\tau}] = \int_0^\infty e^{-s\tau} p(\tau)\, d\tau = \mathcal{L}\{p(\tau)\}(s)$$

is the Laplace transform of $p(\tau)$, applying the inverse Laplace transform yields

$$\begin{aligned} p(\tau) &= \mathcal{L}^{-1}\{E[e^{-s\tau}]\} = \mathcal{L}^{-1}\{e^{\frac{1}{\sigma^2}(\mu-\sqrt{2s\sigma^2+\mu^2})x}\} \\ &= \frac{x}{\sigma\sqrt{2\pi}\tau^{3/2}} e^{-\frac{(x-\mu\tau)^2}{2\tau\sigma^2}}, \qquad \tau > 0. \end{aligned}$$

It is worth noting that the computation of the previous inverse Laplace transform is non-elementary; however, it can be easily computed using the Mathematica software.

(*b*) The moments are obtained by differentiating the moment generating function and taking the value at $s = 0$

$$\begin{aligned} E[\tau] &= -\frac{d}{ds}E[e^{-s\tau}]\Big|_{s=0} = -\frac{d}{ds}e^{\frac{1}{\sigma^2}(\mu-\sqrt{2s\sigma^2+\mu^2})x}\Big|_{s=0} \\ &= \frac{x}{\sqrt{\mu^2+2s\mu}}e^{\frac{1}{\sigma^2}(\mu-\sqrt{2s\sigma^2+\mu^2})x}\Big|_{s=0} \\ &= \frac{x}{\mu}. \end{aligned}$$

$$\begin{aligned} E[\tau^2] &= (-1)^2\frac{d^2}{ds^2}E[e^{-s\tau}]\Big|_{s=0} = \frac{d^2}{ds^2}e^{\frac{1}{\sigma^2}(\mu-\sqrt{2s\sigma^2+\mu^2})x}\Big|_{s=0} \\ &= \frac{x(\sigma^2 + x\sqrt{\mu^2+2s\sigma^2})}{(\mu^2+2s\sigma^2)^{3/2}}e^{\frac{1}{\sigma^2}(\mu-\sqrt{2s\sigma^2+\mu^2})x}\Big|_{s=0} \\ &= \frac{x\sigma^2}{\mu^3} + \frac{x^2}{\mu^2}. \end{aligned}$$

Hence

$$Var(\tau) = E[\tau^2] - E[\tau]^2 = \frac{x\sigma^2}{\mu^3}.$$

It is worth noting that we can arrive at the formula $E[\tau] = \frac{x}{\mu}$ in the following heuristic way. Taking the expectation in the equation $\mu\tau + \sigma W_\tau = x$ yields $\mu E[\tau] = x$, where we used that $E[W_\tau] = 0$ for any finite stopping time τ (see Exercise 3.5.8 (*a*)). Solving for $E[\tau]$ yields the aforementioned formula. ■

Even if the computations are more or less similar with the previous result, we shall treat next the case of the negative barrier in its full length. This is because of its particular importance in being useful in pricing perpetual American puts.

Proposition 3.6.6 *Assume $\mu, x > 0$. Let τ be the time the process $X_t = \mu t + \sigma W_t$ hits $-x$ for the first time.*
(a) We have

$$E[e^{-s\tau}] = e^{\frac{-1}{\sigma^2}(\mu + \sqrt{2s\sigma^2+\mu^2})x}, \qquad s > 0. \tag{3.6.13}$$

(b) Then the density function of τ is given by

$$p(\tau) = \frac{x}{\sigma\sqrt{2\pi}\tau^{3/2}} e^{-\frac{(x+\mu\tau)^2}{2\tau\sigma^2}}, \qquad \tau > 0. \tag{3.6.14}$$

(c) The mean of τ is

$$E[\tau] = \frac{x}{\mu} e^{-\frac{2\mu x}{\sigma^2}}.$$

Proof: (a) Consider the stopping time $\tau = \inf\{t > 0; X_t = -x\}$. By the Optional Stopping Theorem (Theorem 3.2.1) applied to the martingale $M_t = e^{cW_t - \frac{c^2}{2}t}$ yields

$$\begin{aligned} 1 &= M_0 = E[M_\tau] = E\big[e^{cW_\tau - \frac{c^2}{2}\tau}\big] = E\big[e^{\frac{c}{\sigma}(X_\tau - \mu\tau) - \frac{c^2}{2}\tau}\big] \\ &= E\big[e^{-\frac{c}{\sigma}x - \frac{c\mu}{\sigma}\tau - \frac{c^2}{2}\tau}\big] = e^{-\frac{c}{\sigma}x} E\big[e^{-(\frac{c\mu}{\sigma} + \frac{c^2}{2})\tau}\big]. \end{aligned}$$

Therefore

$$E\big[e^{-(\frac{c\mu}{\sigma} + \frac{c^2}{2})\tau}\big] = e^{\frac{c}{\sigma}x}. \tag{3.6.15}$$

If let $s = \frac{c\mu}{\sigma} + \frac{c^2}{2}$, then solving for c yields $c = -\frac{\mu}{\sigma} \pm \sqrt{2s + \frac{\mu^2}{\sigma^2}}$, but only the negative solution works out; this comes from the fact that both terms of the equation (3.6.15) have to be less than 1. Hence (3.6.15) becomes

$$E[e^{-s\tau}] = e^{\frac{-1}{\sigma^2}(\mu + \sqrt{2s\sigma^2+\mu^2})x}, \qquad s > 0.$$

(b) Relation (3.6.13) can be written equivalently as

$$\mathcal{L}\big(p(\tau)\big) = e^{\frac{-1}{\sigma^2}(\mu + \sqrt{2s\sigma^2+\mu^2})x}.$$

Taking the inverse Laplace transform, and using Mathematica software to compute it, we obtain

$$p(\tau) = \mathcal{L}^{-1}\Big(e^{\frac{-1}{\sigma^2}(\mu + \sqrt{2s\sigma^2+\mu^2})x}\Big)(\tau) = \frac{x}{\sigma\sqrt{2\pi}\tau^{3/2}} e^{-\frac{(x+\mu\tau)^2}{2\tau\sigma^2}}, \quad \tau > 0.$$

(c) Differentiating and evaluating at $s = 0$ we obtain

$$E[\tau] = -\frac{d}{ds}E[e^{-s\tau}]\Big|_{s=0} = e^{-\frac{2\mu}{\sigma^2}x}\frac{x}{\sigma^2}\frac{\sigma^2}{\sqrt{\mu^2}} = \frac{x}{\mu}e^{-\frac{2\mu x}{\sigma^2}}.$$

■

Exercise 3.6.7 *Assume the hypothesis of Proposition 3.6.6 satisfied. Find* $Var(\tau)$.

Exercise 3.6.8 *Find the modes of distributions* (3.6.12) *and* (3.6.14). *What do you notice?*

Exercise 3.6.9 *Let* $X_t = 2t + 3W_t$ *and* $Y_t = 6t + W_t$.

(*a*) *Show that the expected times for* X_t *and* Y_t *to reach any barrier* $x > 0$ *are the same.*

(*b*) *If* X_t *and* Y_t *model the prices of two stocks, which one would you like to own?*

Exercise 3.6.10 *Does* $4t + 2W_t$ *hit* 9 *faster (in expectation) than* $5t + 3W_t$ *hits* 14*?*

Exercise 3.6.11 *Let* τ *be the first time the Brownian motion with drift* $X_t = \mu t + W_t$ *hits* x, *where* $\mu, x > 0$. *Prove the inequality*

$$P(\tau \leq \lambda) \leq e^{-\frac{x^2+\lambda^2\mu^2}{2\lambda}+\mu x}, \quad \forall \lambda > 0.$$

C. The double barrier case

In the following we shall consider the case of double barrier. Consider the Brownian motion with drift $X_t = \mu t + W_t$, $\mu > 0$. Let $\alpha, \beta > 0$ and define the stopping time

$$T = \inf\{t > 0; X_t = \alpha \text{ or } X_t = -\beta\}.$$

Relation (3.5.6) states

$$E[e^{cX_T} e^{-(c\mu+\frac{1}{2}c^2)T}] = 1.$$

Since the random variables T and X_T are independent (why?), we have

$$E[e^{cX_T}]E[e^{-(c\mu+\frac{1}{2}c^2)T}] = 1.$$

Using $E[e^{cX_T}] = e^{c\alpha}p_\alpha + e^{-c\beta}(1-p_\alpha)$, with p_α given by (3.5.7), then

$$E[e^{-(c\mu+\frac{1}{2}c^2)T}] = \frac{1}{e^{c\alpha}p_\alpha + e^{-c\beta}(1-p_\alpha)}.$$

If substitute $s = c\mu + \frac{1}{2}c^2$, then

$$E[e^{-sT}] = \frac{1}{e^{(-\mu+\sqrt{2s+\mu^2})\alpha}p_\alpha + e^{-(-\mu+\sqrt{2s+\mu^2})\beta}(1-p_\alpha)}. \tag{3.6.16}$$

The probability density of the stopping time T is obtained by taking the inverse Laplace transform of the right side expression

$$p(T) = \mathcal{L}^{-1}\Big\{\frac{1}{e^{(-\mu+\sqrt{2s+\mu^2})\alpha}p_\alpha + e^{-(-\mu+\sqrt{2s+\mu^2})\beta}(1-p_\alpha)}\Big\}(\tau),$$

an expression which is not feasible for having closed form solution. However, expression (3.6.16) would be useful for computing the price for double barrier derivatives.

Exercise 3.6.12 *Use formula* (3.6.16) *to find the expectation* $E[T]$.

Exercise 3.6.13 *Denote by* $M_t = N_t - \lambda t$ *the compensated Poisson process and let* $c > 0$ *be a constant.*

(a) *Show that*

$$X_t = e^{cM_t - \lambda t(e^c - c - 1)}$$

is an $\mathcal{F}_t$*-martingale, with* $\mathcal{F}_t = \sigma(N_u; u \leq t)$.

(b) *Let* $a > 0$ *and* $T = \inf\{t > 0; M_t > a\}$ *be the first hitting time of the level* a. *Use the Optional Stopping Theorem to show that*

$$E[e^{-\lambda s T}] = e^{-\varphi(s)a}, \qquad s > 0,$$

where $\varphi : [0, \infty) \to [0, \infty)$ *is the inverse function of* $f(x) = e^x - x - 1$.

(c) *Show that* $E[T] = \infty$.

(d) *Can you use the inverse Laplace transform to find the probability density function of* T*?*

3.7 Limits of Stochastic Processes

Let $(X_t)_{t \geq 0}$ be a stochastic process. We can make sense of the limit expression $X = \lim\limits_{t \to \infty} X_t$, in a similar way as we did in section 1.13 for sequences of random variables. We shall rewrite the definitions for the continuous case.

Almost Certain Limit

The process X_t converges *almost certainly* to X, if for all states of the world ω, except a set of probability zero, we have

$$\lim_{t \to \infty} X_t(\omega) = X(\omega).$$

We shall write $\text{ac-}\lim\limits_{t \to \infty} X_t = X$. It is also sometimes called *strong convergence*.

Mean Square Limit

We say that the process X_t converges to X in the *mean square* if

$$\lim_{t \to \infty} E[(X_t - X)^2] = 0.$$

In this case we write $\text{ms-}\lim\limits_{t \to \infty} X_t = X$.

Limit in Probability or Stochastic Limit

The stochastic process X_t converges in *stochastic limit* to X if

$$\lim_{t\to\infty} P\big(\omega; |X_t(\omega) - X(\omega)| > \epsilon\big) = 0.$$

This limit is abbreviated by st-$\lim\limits_{t\to\infty} X_t = X$.

It is worth noting that, like in the case of sequences of random variables, both almost certain convergence and convergence in mean square imply the stochastic convergence, which implies the limit in distribution.

Limit in Distribution

We say that X_t converges *in distribution* to X if for any continuous bounded function $\varphi(x)$ we have

$$\lim_{t\to\infty} \varphi(X_t) = \varphi(X).$$

It is worth noting that the stochastic convergence implies the convergence in distribution.

3.8 Convergence Theorems

The following property is a reformulation of Exercise 1.13.1 in the continuous setup.

Proposition 3.8.1 *Consider a stochastic process X_t such that $E[X_t] \to k$, a constant, and $Var(X_t) \to 0$ as $t \to \infty$. Then ms-$\lim\limits_{t\to\infty} X_t = k$.*

It is worthy to note that the previous statement holds true if the limit to infinity is replaced by a limit to any other number.

Next we shall provide a few applications.

Application 3.8.2 *If $\alpha > 1/2$, then*

$$ms\text{-}\lim_{t\to\infty} \frac{W_t}{t^\alpha} = 0.$$

Proof: Let $X_t = \dfrac{W_t}{t^\alpha}$. Then $E[X_t] = \dfrac{E[W_t]}{t^\alpha} = 0$, and $Var[X_t] = \dfrac{1}{t^{2\alpha}} Var[W_t] = \dfrac{t}{t^{2\alpha}} = \dfrac{1}{t^{2\alpha-1}}$, for any $t > 0$. Since $\dfrac{1}{t^{2\alpha-1}} \to 0$ as $t \to \infty$, applying Proposition 3.8.1 yields ms-$\lim\limits_{t\to\infty} \dfrac{W_t}{t^\alpha} = 0$. ∎

Corollary 3.8.3 $ms\text{-}\lim\limits_{t\to\infty} \dfrac{W_t}{t} = 0.$

Application 3.8.4 *Let* $Z_t = \int_0^t W_s\, ds$. *If* $\beta > 3/2$, *then*

$$ms\text{-}\lim_{t\to\infty} \frac{Z_t}{t^\beta} = 0.$$

Proof: Let $X_t = \dfrac{Z_t}{t^\beta}$. Then $E[X_t] = \dfrac{E[Z_t]}{t^\beta} = 0$, and $Var[X_t] = \dfrac{1}{t^{2\beta}} Var[Z_t] = \dfrac{t^3}{3t^{2\beta}} = \dfrac{1}{3t^{2\beta-3}}$, for any $t > 0$. Since $\dfrac{1}{3t^{2\beta-3}} \to 0$ as $t \to \infty$, applying Proposition 3.8.1 leads to the desired result. ■

Application 3.8.5 *For any* $p > 0$, $c \geq 1$ *we have*

$$ms\text{-}\lim_{t\to\infty} \frac{e^{W_t - ct}}{t^p} = 0.$$

Proof: Consider the process $X_t = \dfrac{e^{W_t - ct}}{t^p} = \dfrac{e^{W_t}}{t^p e^{ct}}$. Since

$$\begin{aligned} E[X_t] &= \frac{E[e^{W_t}]}{t^p e^{ct}} = \frac{e^{t/2}}{t^p e^{ct}} = \frac{1}{e^{(c-\frac{1}{2})t}} \frac{1}{t^p} \to 0, \quad \text{as } t \to \infty \\ Var[X_t] &= \frac{Var[e^{W_t}]}{t^{2p} e^{2ct}} = \frac{e^{2t} - e^t}{t^{2p} e^{2ct}} = \frac{1}{t^{2p}} \Big(\frac{1}{e^{2t(c-1)}} - \frac{1}{e^{t(2c-1)}} \Big) \to 0, \end{aligned}$$

as $t \to \infty$, Proposition 3.8.1 leads to the desired result. ■

Application 3.8.6 *Show that*

$$ms\text{-}\lim_{t\to\infty} \frac{\max\limits_{0\leq s\leq t} W_s}{t} = 0.$$

Proof: Let $X_t = \dfrac{\max\limits_{0\leq s\leq t} W_s}{t}$. Since by Exercise 3.3.10

$$\begin{aligned} E[\max_{0\leq s\leq t} W_s] &= 0 \\ Var(\max_{0\leq s\leq t} W_s) &= 2\sqrt{t}, \end{aligned}$$

then

$$\begin{aligned} E[X_t] &= 0 \\ Var[X_t] &= \frac{2\sqrt{t}}{t^2} \to 0, \qquad t \to \infty. \end{aligned}$$

Apply Proposition 3.8.1 to get the desired result. ■

Remark 3.8.7 *One of the strongest result regarding limits of Brownian motions is called the law of iterated logarithms and was first proved by Lamperti:*

$$\lim_{t\to\infty} \sup \frac{W_t}{\sqrt{2t\ln(\ln t)}} = 1,$$

almost certainly.

Exercise 3.8.8 *Use the law of iterated logarithms to show*

$$\lim_{t\to 0^+} \frac{W_t}{\sqrt{2t\ln\ln(1/t)}} = 1.$$

Application 3.8.9 *We shall show that ac-$\lim\limits_{t\to\infty} \dfrac{W_t}{t} = 0$.*

From the law of iterated logarithms $\dfrac{W_t}{\sqrt{2t\ln(\ln t)}} < 1$ for t large. Then

$$\frac{W_t}{t} = \frac{W_t}{\sqrt{2t\ln(\ln t)}} \frac{\sqrt{2\ln(\ln t)}}{\sqrt{t}} < \frac{\sqrt{2\ln(\ln t)}}{\sqrt{t}}.$$

Let $\epsilon_t = \dfrac{\sqrt{2\ln(\ln t)}}{\sqrt{t}}$. Then $\dfrac{W_t}{t} < \epsilon_t$ for t large. As an application of the l'Hospital rule, it is not hard to see that ϵ_t satisfies the following limits

$$\begin{aligned} \epsilon_t \to 0, &\qquad t\to\infty \\ \epsilon_t\sqrt{t} \to \infty, &\qquad t\to\infty. \end{aligned}$$

In order to show that ac-$\lim\limits_{t\to\infty} \dfrac{W_t}{t} = 0$, it suffices to prove

$$P\Big(\omega; \Big|\frac{W_t(\omega)}{t}\Big| < \epsilon_t\Big) \to 1, \qquad t\to\infty. \tag{3.8.17}$$

We have

$$\begin{aligned} P\Big(\omega; \Big|\frac{W_t(\omega)}{t}\Big| < \epsilon_t\Big) &= P\Big(\omega; -t\epsilon_t < W_t(\omega) < t\epsilon_t\Big) = \int_{-t\epsilon_t}^{t\epsilon_t} \frac{1}{\sqrt{2\pi t}} e^{-\frac{u^2}{2t}}\, du \\ &= \int_{-\epsilon_t\sqrt{t}}^{\epsilon_t\sqrt{t}} \frac{1}{\sqrt{2\pi}} e^{-\frac{v^2}{2}}\, dv \to \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{v^2}{2}}\, dv = 1, \qquad t\to\infty, \end{aligned}$$

where we used $\epsilon_t\sqrt{t}\to\infty$, as $t\to\infty$, which proves (3.8.17).

Proposition 3.8.10 *Let X_t be a stochastic process. Then*

$$ms\text{-}\lim_{t\to\infty} X_t = 0 \Longleftrightarrow ms\text{-}\lim_{t\to\infty} X_t^2 = 0.$$

Proof: Left as an exercise. ∎

Exercise 3.8.11 *Let X_t be a stochastic process. Show that*

$$ms\text{-}\lim_{t\to\infty} X_t = 0 \Longleftrightarrow ms\text{-}\lim_{t\to\infty} |X_t| = 0.$$

Another convergence result can be obtained if we consider the continuous analog of Example 1.13.6:

Proposition 3.8.12 *Let X_t be a stochastic process such that there is a $p > 0$ such that $E[|X_t|^p] \to 0$ as $t \to \infty$. Then $st\text{-}\lim\limits_{t\to\infty} X_t = 0$.*

Application 3.8.13 *We shall show that for any $\alpha > 1/2$*

$$st\text{-}\lim_{t\to\infty} \frac{W_t}{t^\alpha} = 0.$$

Proof: Consider the process $X_t = \dfrac{W_t}{t^\alpha}$. By Proposition 3.8.10 it suffices to show $\text{ms-}\lim\limits_{t\to\infty} X_t^2 = 0$. Since the mean square convergence implies the stochastic convergence, we get $\text{st-}\lim\limits_{t\to\infty} X_t^2 = 0$. Since

$$E[|X_t|^2] = E[X_t^2] = E\Big[\frac{W_t^2}{t^{2\alpha}}\Big] = \frac{E[W_t^2]}{t^{2\alpha}} = \frac{t}{t^{2\alpha}} = \frac{1}{t^{2\alpha-1}} \to 0, \quad t \to \infty,$$

then Proposition 3.8.12 yields $\text{st-}\lim\limits_{t\to\infty} X_t = 0$. ∎

The following result can be regarded as the L'Hospital's rule for sequences:

Lemma 3.8.14 (Cesaró-Stoltz) *Let x_n and y_n be two sequences of real numbers, $n \geq 1$. If the limit $\lim\limits_{n\to\infty} \dfrac{x_{n+1} - x_n}{y_{n+1} - y_n}$ exists and is equal to L, then the following limit exists*

$$\lim_{n\to\infty} \frac{x_n}{y_n} = L.$$

Proof: (Sketch) Assume there are differentiable functions f and g such that $f(n) = x_n$ and $g(n) = y_n$. (How do you construct these functions?) From Cauchy's theorem[2] there is a $c_n \in (n, n+1)$ such that

$$L = \lim_{n\to\infty} \frac{x_{n+1} - x_n}{y_{n+1} - y_n} = \lim_{n\to\infty} \frac{f(n+1) - f(n)}{g(n+1) - g(n)} = \lim_{n\to\infty} \frac{f'(c_n)}{g'(c_n)}.$$

[2]This says that if f and g are differentiable on (a, b) and continuous on $[a, b]$, then there is a $c \in (a, b)$ such that $\dfrac{f(a) - f(b)}{g(a) - g(b)} = \dfrac{f'(c)}{g'(c)}$.

Since $c_n \to \infty$ as $n \to \infty$, we can also write the aforementioned limit as

$$\lim_{t\to\infty} \frac{f'(t)}{g'(t)} = L.$$

(Here one may argue against this, but we recall the freedom of choice for the functions f and g such that c_n can be any number between n and $n+1$). By l'Hospital's rule we get

$$\lim_{t\to\infty} \frac{f(t)}{g(t)} = L.$$

Making $t = n$ yields $\lim\limits_{t\to\infty} \frac{x_n}{y_n} = L$. ■

The next application states that if a sequence is convergent, then the arithmetic average of its terms is also convergent, and the sequences have the same limit.

Example 3.8.1 *Let a_n be a convergent sequence with $\lim\limits_{n\to\infty} a_n = L$. Let*

$$A_n = \frac{a_1 + a_2 + \cdots + a_n}{n}$$

be the arithmetic average of the first n terms. Then A_n is convergent and

$$\lim_{n\to\infty} A_n = L.$$

Proof: This is an application of Cesaró-Stoltz lemma. Consider the sequences $x_n = a_1 + a_2 + \cdots + a_n$ and $y_n = n$. Since

$$\frac{x_{n+1} - x_n}{y_{n+1} - y_n} = \frac{(a_1 + \cdots + a_{n+1}) - (a_1 + \cdots + a_n)}{(n+1) - n} = \frac{a_{n+1}}{1},$$

then

$$\lim_{n\to\infty} \frac{x_{n+1} - x_n}{y_{n+1} - y_n} = \lim_{n\to\infty} a_{n+1} = L.$$

Applying the Cesaró-Stoltz lemma yields

$$\lim_{n\to\infty} A_n = \lim_{n\to\infty} \frac{x_n}{y_n} = L.$$

■

Exercise 3.8.15 *Let b_n be a convergent sequence with $\lim\limits_{n\to\infty} b_n = L$. Let*

$$G_n = (b_1 \cdot b_2 \cdot \cdots \cdot b_n)^{1/n}$$

be the geometric average of the first n terms. Show that G_n is convergent and

$$\lim_{n\to\infty} G_n = L.$$

The following result extends the Cesaró-Stoltz lemma to sequences of random variables.

Proposition 3.8.16 *Let X_n be a sequence of random variables on the probability space $(\Omega, \mathcal{F}, P)$, such that*

$$ac\text{-}\lim_{n\to\infty} \frac{X_{n+1} - X_n}{Y_{n+1} - Y_n} = L.$$

Then

$$ac\text{-}\lim_{n\to\infty} \frac{X_n}{Y_n} = L.$$

Proof: Consider the sets

$$\begin{aligned} A &= \{\omega \in \Omega; \lim_{n\to\infty} \frac{X_{n+1}(\omega) - X_n(\omega)}{Y_{n+1}(\omega) - Y_n(\omega)} = L\} \\ B &= \{\omega \in \Omega; \lim_{n\to\infty} \frac{X_n(\omega)}{Y_n(\omega)} = L\}. \end{aligned}$$

Since for any given state of the world ω, the sequences $x_n = X_n(\omega)$ and $y_n = Y_n(\omega)$ are numerical sequences, Lemma 3.8.14 yields the inclusion $A \subset B$. This implies $P(A) \le P(B)$. Since $P(A) = 1$, it follows that $P(B) = 1$, which leads to the desired conclusion. ■

Example 3.8.17 *Let S_n denote the price of a stock on day n, and assume that*

$$ac\text{-}\lim_{n\to\infty} S_n = L.$$

Then

$$ac\text{-}\lim_{n\to\infty} \frac{S_1 + \cdots + S_n}{n} = L \quad \text{and} \quad ac\text{-}\lim_{n\to\infty} (S_1 \cdot \cdots \cdot S_n)^{1/n} = L.$$

This says that if almost all future simulations of the stock price approach the steady state limit L, the arithmetic and geometric averages converge to the same limit. The statement is a consequence of Proposition 3.8.16 and follows a similar proof as Example 3.8.1. Asian options have payoffs depending on these type of averages, as we shall see in Part II.

3.9 The Martingale Convergence Theorem

We state now, without proof, a result which is a powerful way of proving the almost certain convergence. We start with the discrete version:

Theorem 3.9.1 *Let X_n be a martingale with bounded means*

$$\exists M > 0 \text{ such that } E[|X_n|] \le M, \qquad \forall n \ge 1. \tag{3.9.18}$$

Then there is $L < \infty$ such that $ac\text{-}\lim_{n\to\infty} X_n = L$, i.e.

$$P\big(\omega; \lim_{n\to\infty} X_n(\omega) = L\big) = 1.$$

Since $E[|X_n|]^2 \le E[X_n^2]$, the condition (3.9.18) can be replaced by its stronger version

$$\exists M > 0 \text{ such that } E[X_n^2] \le M, \qquad \forall n \ge 1.$$

The following result deals with the continuous version of the Martingale Convergence Theorem. Denote the infinite knowledge by $\mathcal{F}_\infty = \sigma\Big(\cup_t \mathcal{F}_t\Big)$.

Theorem 3.9.2 *Let X_t be an $\mathcal{F}_t$-martingale such that*

$$\exists M > 0 \text{ such that } E[|X_t|] < M, \qquad \forall t > 0.$$

Then there is an $\mathcal{F}_\infty$-measurable random variable X_∞ such that $X_t \to X_\infty$ a.c. as $t \to \infty$.

The next exrcise deals with a process that is a.c-convergent but is not ms-convergent.

Exercise 3.9.3 *It is known that $X_t = e^{W_t - t/2}$ is a martingale. Since*

$$E[|X_t|] = E[e^{W_t - t/2}] = e^{-t/2}E[e^{W_t}] = e^{-t/2}e^{t/2} = 1,$$

by the Martingale Convergence Theorem there is a number L such that $X_t \to L$ a.c. as $t \to \infty$.

(*a*) *What is the limit L? How did you make your guess?*

(*b*) *Show that*

$$E[|X_t - 1|^2] = Var(X_t) + \Big(E(X_t) - 1\Big)^2.$$

(*c*) *Show that X_t does not converge in the mean square to 1.*

(*d*) *Prove that the sequence X_t is a.c-convergent but it is not ms-convergent.*

3.10 The Squeeze Theorem

The following result is the analog of the Squeeze Theorem from usual Calculus.

Theorem 3.10.1 *Let X_n, Y_n, Z_n be sequences of random variables on the probability space $(\Omega, \mathcal{F}, P)$ such that*

$$X_n \le Y_n \le Z_n \ \ a.s. \ \forall n \ge 1.$$

If X_n and Z_n converge to L as $n \to \infty$ almost certainly (or in mean square, or stochastic or in distribution), then Y_n converges to L in a similar mode.

Proof: For any state of the world $\omega \in \Omega$ consider the sequences $x_n = X_n(\omega)$, $y_n = Y_n(\omega)$ and $z_n = Z_n(\omega)$ and apply the usual Squeeze Theorem to them. ■

Remark 3.10.2 *The previous theorem remains valid if n is replaced by a continuous positive parameter t.*

Example 3.10.1 *Show that* $ac\text{-}\lim\limits_{t\to\infty} \dfrac{W_t \sin(W_t)}{t} = 0$.

Proof: Consider the sequences $X_t = 0$, $Y_t = \dfrac{W_t \sin(W_t)}{t}$ and $Z_t = \dfrac{W_t}{t}$. From Application 3.8.9 we have $\text{ac-}\lim\limits_{t\to\infty} Z_t = 0$. Applying the Squeeze Theorem we obtain the desired result. ■

Exercise 3.10.3 *Use the Squeeze Theorem to find the following limits:*

(a) $ac\text{-}\lim\limits_{t\to\infty} \dfrac{\sin(W_t)}{t}$;

(b) $ac\text{-}\lim\limits_{t\to 0} t \cos W_t$;

(c) $ac\text{-}\lim\limits_{t\to -\infty} e^t (\sin W_t)^2$.

3.11 Quadratic Variations

For some stochastic processes the sum of squares of consecutive increments tends in mean square to a finite number, as the norm of the partition decreases to zero. We shall encounter in the following a few important examples that will be useful when dealing with stochastic integrals.

3.11.1 The Quadratic Variation of W_t

The next result states that the *quadratic variation* of the Brownian motion W_t on the interval $[0, T]$ is T. More precisely, we have:

Proposition 3.11.1 *Let $T > 0$ and consider the equidistant partition $0 = t_0 < t_1 < \cdots t_{n-1} < t_n = T$. Then*

$$\boxed{ms\text{-}\lim_{n\to\infty} \sum_{i=0}^{n-1} (W_{t_{i+1}} - W_{t_i})^2 = T.} \tag{3.11.19}$$

Proof: Consider the random variable

$$X_n = \sum_{i=0}^{n-1} (W_{t_{i+1}} - W_{t_i})^2.$$

Since the increments of a Brownian motion are independent, Proposition 4.2.1 yields

$$\begin{aligned} E[X_n] &= \sum_{i=0}^{n-1} E[(W_{t_{i+1}} - W_{t_i})^2] = \sum_{i=0}^{n-1}(t_{i+1} - t_i) \\ &= t_n - t_0 = T; \end{aligned}$$

$$\begin{aligned} Var(X_n) &= \sum_{i=0}^{n-1} Var[(W_{t_{i+1}} - W_{t_i})^2] = \sum_{i=0}^{n-1} 2(t_{i+1} - t_i)^2 \\ &= n \cdot 2\Big(\frac{T}{n}\Big)^2 = \frac{2T^2}{n}, \end{aligned}$$

where we used that the partition is equidistant. Since X_n satisfies the conditions

$$\begin{aligned} E[X_n] &= T, \quad \forall n \geq 1; \\ Var[X_n] &\to 0, \quad n \to \infty, \end{aligned}$$

by Proposition 3.8.1 we obtain $\text{ms-}\lim\limits_{n\to\infty} X_n = T$, or

$$\text{ms-}\lim_{n\to\infty} \sum_{i=0}^{n-1}(W_{t_{i+1}} - W_{t_i})^2 = T. \tag{3.11.20}$$

■

Exercise 3.11.2 *Prove that the quadratic variation of the Brownian motion W_t on $[a, b]$ is equal to $b - a$.*

The Fundamental Relation $dW_t^2 = dt$

The relation discussed in this section can be regarded as the fundamental relation of Stochastic Calculus. We shall start by recalling relation (3.11.20)

$$\text{ms-}\lim_{n\to\infty} \sum_{i=0}^{n-1}(W_{t_{i+1}} - W_{t_i})^2 = T. \tag{3.11.21}$$

The right side can be regarded as a regular Riemann integral

$$T = \int_0^T dt,$$

while the left side can be regarded as a stochastic integral with respect to dW_t^2

$$\int_0^T (dW_t)^2 = \text{ms-}\lim_{n\to\infty} \sum_{i=0}^{n-1}(W_{t_{i+1}} - W_{t_i})^2.$$

Substituting into (3.11.21) yields

$$\int_0^T (dW_t)^2 = \int_0^T dt, \quad \forall T > 0.$$

The differential form of this integral equation is

$$\boxed{dW_t^2 = dt.}$$

In fact, this expression also holds in the mean square sense, as it can be inferred from the next exercise.

Exercise 3.11.3 *Show that*

(*a*) $E[dW_t^2 - dt] = 0$;

(*b*) $Var(dW_t^2 - dt) = o(dt)$;

(*c*) $ms\text{-}\lim_{dt \to 0} (dW_t^2 - dt) = 0.$

Roughly speaking, the process dW_t^2, which is the square of infinitesimal increments of a Brownian motion, is totally predictable. This relation plays a central role in Stochastic Calculus and will be useful when dealing with Ito's lemma.

The following exercise states that $dtdW_t = 0$, which is another important stochastic relation useful in Ito's lemma.

Exercise 3.11.4 *Consider the equidistant partition* $0 = t_0 < t_1 < \cdots t_{n-1} < t_n = T$. *Then*

$$ms\text{-}\lim_{n \to \infty} \sum_{i=0}^{n-1} (W_{t_{i+1}} - W_{t_i})(t_{i+1} - t_i) = 0. \tag{3.11.22}$$

3.11.2 The Quadratic Variation of $N_t - \lambda t$

The following result deals with the quadratic variation of the compensated Poisson process $M_t = N_t - \lambda t$.

Proposition 3.11.5 *Let* $a < b$ *and consider the partition* $a = t_0 < t_1 < \cdots < t_{n-1} < t_n = b$. *Then*

$$\boxed{ms\text{-}\lim_{\|\Delta_n\| \to 0} \sum_{k=0}^{n-1} (M_{t_{k+1}} - M_{t_k})^2 = N_b - N_a,} \tag{3.11.23}$$

where $\|\Delta_n\| := \sup_{0 \le k \le n-1} (t_{k+1} - t_k)$.

Proof: For the sake of simplicity we shall use the following notations:

$$\Delta t_k = t_{k+1} - t_k, \quad \Delta M_k = M_{t_{k+1}} - M_{t_k}, \quad \Delta N_k = N_{t_{k+1}} - N_{t_k}.$$

The relation we need to prove can also be written as

$$\text{ms-}\lim_{n\to\infty} \sum_{k=0}^{n-1} \big[(\Delta M_k)^2 - \Delta N_k\big] = 0.$$

Let

$$Y_k = (\Delta M_k)^2 - \Delta N_k = (\Delta M_k)^2 - \Delta M_k - \lambda \Delta t_k.$$

It suffices to show that

$$E\Big[\sum_{k=0}^{n-1} Y_k\Big] = 0, \tag{3.11.24}$$

$$\lim_{n\to\infty} Var\Big[\sum_{k=0}^{n-1} Y_k\Big] = 0. \tag{3.11.25}$$

The first identity follows from the properties of Poisson processes (see Exercise 2.8.9)

$$\begin{aligned} E\Big[\sum_{k=0}^{n-1} Y_k\Big] &= \sum_{k=0}^{n-1} E[Y_k] = \sum_{k=0}^{n-1} E[(\Delta M_k)^2] - E[\Delta N_k] \\ &= \sum_{k=0}^{n-1} (\lambda \Delta t_k - \lambda \Delta t_k) = 0. \end{aligned}$$

For the proof of the identity (3.11.25) we need to find first the variance of Y_k.

$$\begin{aligned} Var[Y_k] &= Var[(\Delta M_k)^2 - (\Delta M_k + \lambda \Delta t_k)] = Var[(\Delta M_k)^2 - \Delta M_k] \\ &= Var[(\Delta M_k)^2] + Var[\Delta M_k] - 2Cov[\Delta M_k^2, \Delta M_k] \\ &= \lambda \Delta t_k + 2\lambda^2 \Delta t_k^2 + \lambda \Delta t_k \\ &\quad -2\Big[E[(\Delta M_k)^3] - E[(\Delta M_k)^2] E[\Delta M_k]\Big] \\ &= 2\lambda^2 (\Delta t_k)^2, \end{aligned}$$

where we used Exercise 2.8.9 and the fact that $E[\Delta M_k] = 0$. Since M_t is a process with independent increments, then $Cov[Y_k, Y_j] = 0$ for $i \neq j$. Then

$$\begin{aligned} Var\Big[\sum_{k=0}^{n-1} Y_k\Big] &= \sum_{k=0}^{n-1} Var[Y_k] + 2\sum_{k\neq j} Cov[Y_k, Y_j] = \sum_{k=0}^{n-1} Var[Y_k] \\ &= 2\lambda^2 \sum_{k=0}^{n-1} (\Delta t_k)^2 \leq 2\lambda^2 \|\Delta_n\| \sum_{k=0}^{n-1} \Delta t_k = 2\lambda^2 (b-a)\|\Delta_n\|, \end{aligned}$$

and hence $Var\Big[\sum_{k=0}^{n-1} Y_n\Big] \to 0$ as $\|\Delta_n\| \to 0$. According to the Example 1.13.1, we obtain the desired limit in mean square. ■

The previous result states that the quadratic variation of the martingale M_t between a and b is equal to the jump of the Poisson process between a and b.

The Fundamental Relation $dM_t^2 = dN_t$

Recall relation (3.11.23)

$$\text{ms-}\lim_{n\to\infty} \sum_{k=0}^{n-1} (M_{t_{k+1}} - M_{t_k})^2 = N_b - N_a. \tag{3.11.26}$$

The right side can be regarded as a Riemann-Stieltjes integral

$$N_b - N_a = \int_a^b dN_t,$$

while the left side can be regarded as a stochastic integral with respect to $(dM_t)^2$

$$\int_a^b (dM_t)^2 := \text{ms-}\lim_{n\to\infty} \sum_{k=0}^{n-1} (M_{t_{k+1}} - M_{t_k})^2.$$

Substituting in (3.11.26) yields

$$\int_a^b (dM_t)^2 = \int_a^b dN_t,$$

for any $a < b$. The equivalent differential form is

$$\boxed{(dM_t)^2 = dN_t.} \tag{3.11.27}$$

The Relations $dt\, dM_t = 0$**,** $dW_t\, dM_t = 0$

In order to show that $dt\, dM_t = 0$ in the mean square sense, we need to prove the limit

$$\text{ms-}\lim_{n\to\infty} \sum_{k=0}^{n-1} (t_{k+1} - t_k)(M_{t_{k+1}} - M_{t_k}) = 0. \tag{3.11.28}$$

This can be thought as a vanishing integral of the increment process dM_t with respect to dt

$$\int_a^b dM_t\, dt = 0, \qquad \forall a, b \in \mathbb{R}.$$

Denote

$$X_n = \sum_{k=0}^{n-1}(t_{k+1} - t_k)(M_{t_{k+1}} - M_{t_k}) = \sum_{k=0}^{n-1} \Delta t_k \Delta M_k.$$

In order to show (3.11.28) it suffices to prove that

1. $E[X_n] = 0$;

2. $\lim\limits_{n\to\infty} Var[X_n] = 0$.

Using the additivity of the expectation and Exercise 2.8.9, (ii)

$$E[X_n] = E\Big[\sum_{k=0}^{n-1} \Delta t_k \Delta M_k\Big] = \sum_{k=0}^{n-1} \Delta t_k E[\Delta M_k] = 0.$$

Since the Poisson process N_t has independent increments, the same property holds for the compensated Poisson process M_t. Then $\Delta t_k \Delta M_k$ and $\Delta t_j \Delta M_j$ are independent for $k \neq j$, and using the properties of variance we have

$$Var[X_n] = Var\Big[\sum_{k=0}^{n-1} \Delta t_k \Delta M_k\Big] = \sum_{k=0}^{n-1}(\Delta t_k)^2 Var[\Delta M_k] = \lambda \sum_{k=0}^{n-1}(\Delta t_k)^3,$$

where we used

$$Var[\Delta M_k] = E[(\Delta M_k)^2] - (E[\Delta M_k])^2 = \lambda \Delta t_k,$$

see Exercise 2.8.9 (ii). If we let $\|\Delta_n\| = \max\limits_k \Delta t_k$, then

$$Var[X_n] = \lambda \sum_{k=0}^{n-1}(\Delta t_k)^3 \leq \lambda \|\Delta_n\|^2 \sum_{k=0}^{n-1} \Delta t_k = \lambda(b-a)\|\Delta_n\|^2 \to 0$$

as $n \to \infty$. Hence we proved the stochastic differential relation

$$\boxed{dt\, dM_t = 0.} \tag{3.11.29}$$

For showing the relation $dW_t\, dM_t = 0$, we need to prove

$$\text{ms-}\lim_{n\to\infty} Y_n = 0, \tag{3.11.30}$$

where we have denoted

$$Y_n = \sum_{k=0}^{n-1}(W_{k+1} - W_k)(M_{t_{k+1}} - M_{t_k}) = \sum_{k=0}^{n-1} \Delta W_k \Delta M_k.$$

Since the Brownian motion W_t and the process M_t have independent increments and ΔW_k is independent of ΔM_k, we have

$$E[Y_n] = \sum_{k=0}^{n-1} E[\Delta W_k \Delta M_k] = \sum_{k=0}^{n-1} E[\Delta W_k] E[\Delta M_k] = 0,$$

where we used $E[\Delta W_k] = E[\Delta M_k] = 0$. Using also $E[(\Delta W_k)^2] = \Delta t_k$, $E[(\Delta M_k)^2] = \lambda \Delta t_k$, and invoking the independence of ΔW_k and ΔM_k, we get

$$\begin{aligned} Var[\Delta W_k \Delta M_k] &= E[(\Delta W_k)^2 (\Delta M_k)^2] - (E[\Delta W_k \Delta M_k])^2 \\ &= E[(\Delta W_k)^2] E[(\Delta M_k)^2] - E[\Delta W_k]^2 E[\Delta M_k]^2 \\ &= \lambda (\Delta t_k)^2. \end{aligned}$$

Then using the independence of the terms in the sum, we get

$$\begin{aligned} Var[Y_n] &= \sum_{k=0}^{n-1} Var[\Delta W_k \Delta M_k] = \lambda \sum_{k=0}^{n-1} (\Delta t_k)^2 \\ &\leq \lambda \|\Delta_n\| \sum_{k=0}^{n-1} \Delta t_k = \lambda (b-a) \|\Delta_n\| \to 0, \end{aligned}$$

as $n \to \infty$. Since Y_n is a random variable with mean zero and variance decreasing to zero, it follows that $Y_n \to 0$ in the mean square sense. Hence we proved that

$$\boxed{dW_t \, dM_t = 0.} \qquad (3.11.31)$$

Exercise 3.11.6 *Show the following stochastic differential relations:*

(a) $dt \, dN_t = 0$; (b) $dW_t \, dN_t = 0$; (c) $dt \, dW_t = 0$;
(d) $(dN_t)^2 = dN_t$; (e) $(dM_t)^2 = dN_t$; (f) $(dM_t)^4 = dN_t$.

The relations proved in this section will be useful in the Part II when developing the stochastic model of a stock price that exhibits jumps modeled by a Poisson process.

Chapter 4

Stochastic Integration

This chapter deals with one of the most useful stochastic integrals, called the *Ito integral*. This type of integral was introduced in 1944 by the Japanese mathematician K. Ito, and was originally motivated by a construction of diffusion processes.

4.1 Nonanticipating Processes

Consider the Brownian motion W_t. A process F_t is called a *nonanticipating process* if F_t is independent of any future increment $W_{t'} - W_t$ for any t and t' with $t < t'$. Consequently, the process F_t is independent of the behavior of the Brownian motion in the future, i.e. it cannot anticipate the future. For instance, W_t, e^{W_t}, $W_t^2 - W_t + t$ are examples of nonanticipating processes, while W_{t+1} or $\frac{1}{2}(W_{t+1} - W_t)^2$ are not.

Nonanticipating processes are important because the Ito integral concept applies only to them.

If $\mathcal{F}_t$ denotes the information known until time t, where this information is generated by the Brownian motion $\{W_s; s \leq t\}$, then any $\mathcal{F}_t$-adapted process F_t is nonanticipating.

4.2 Increments of Brownian Motions

In this section we shall discuss a few basic properties of the increments of a Brownian motion, which will be useful when computing stochastic integrals.

Proposition 4.2.1 *Let W_t be a Brownian motion. If $s < t$, we have*

1. $E[(W_t - W_s)^2] = t - s$.
2. $Var[(W_t - W_s)^2] = 2(t - s)^2$.

Proof: 1. Using that $W_t - W_s \sim N(0, t - s)$, we have

$$E[(W_t - W_s)^2] = E[(W_t - W_s)^2] - (E[W_t - W_s])^2 = Var(W_t - W_s) = t - s.$$

2. Dividing by the standard deviation yields the standard normal random variable $\frac{W_t - W_s}{\sqrt{t-s}} \sim N(0,1)$. Its square, $\frac{(W_t - W_s)^2}{t-s}$ is χ^2-distributed with 1 degree of freedom.[1] Its mean is 1 and its variance is 2. This implies

$$\begin{aligned} E\Big[\frac{(W_t - W_s)^2}{t-s}\Big] &= 1 \Longrightarrow E[(W_t - W_s)^2] = t - s; \\ Var\Big[\frac{(W_t - W_s)^2}{t-s}\Big] &= 2 \Longrightarrow Var[(W_t - W_s)^2] = 2(t-s)^2. \end{aligned}$$

■

Remark 4.2.2 *The infinitesimal version of the previous result is obtained by replacing $t - s$ with dt*

1. $E[dW_t^2] = dt$;
2. $Var[dW_t^2] = 2dt^2$.

We shall see in an upcoming section that in fact dW_t^2 and dt are equal in a mean square sense.

Exercise 4.2.3 *Show that*

(a) $E[(W_t - W_s)^4] = 3(t-s)^2$;

(b) $E[(W_t - W_s)^6] = 15(t-s)^3$.

4.3 The Ito Integral

The Ito integral is defined in a way that is similar to the Riemann integral. The Ito integral is taken with respect to infinitesimal increments of a Brownian motion, dW_t, which are random variables, while the Riemann integral considers integration with respect to the predictable infinitesimal changes dt. It is worth noting that the Ito integral is a random variable, while the Riemann integral is just a real number. Despite this fact, there are several common properties and relations between these two types of integrals.

Consider $0 \le a < b$ and let $F_t = f(W_t, t)$ be a nonanticipating process with

$$E\Big[\int_a^b F_t^2 \, dt\Big] < \infty. \tag{4.3.1}$$

The role of the previous condition will be made more clear when we shall discuss about the martingale property of a the Ito integral. Divide the interval $[a, b]$ into n subintervals using the partition points

$$a = t_0 < t_1 < \cdots < t_{n-1} < t_n = b,$$

[1] A χ^2-distributed random variable with n degrees of freedom has mean n and variance $2n$.

and consider the partial sums

$$S_n = \sum_{i=0}^{n-1} F_{t_i}(W_{t_{i+1}} - W_{t_i}).$$

We emphasize that the intermediate points are the left endpoints of each interval, and this is the way they should be always chosen. Since the process F_t is nonanticipative, the random variables F_{t_i} and $W_{t_{i+1}} - W_{t_i}$ are independent; this is an important feature in the definition of the Ito integral.

The *Ito integral* is the limit of the partial sums S_n

$$\operatorname*{ms-lim}_{n\to\infty} S_n = \int_a^b F_t\, dW_t,$$

provided the limit exists. It can be shown that the choice of partition does not influence the value of the Ito integral. This is the reason why, for practical purposes, it suffices to assume the intervals equidistant, i.e.

$$t_{i+1} - t_i = \frac{(b-a)}{n}, \qquad i = 0, 1, \cdots, n-1.$$

The previous convergence is in the mean square sense, i.e.

$$\lim_{n\to\infty} E\Big[\Big(S_n - \int_a^b F_t\, dW_t\Big)^2\Big] = 0.$$

Existence of the Ito integral

It is known that the Ito stochastic integral $\displaystyle\int_a^b F_t\, dW_t$ exists if the process $F_t = f(W_t, t)$ satisfies the following properties:

1. The paths $t \to F_t(\omega)$ are continuous on $[a, b]$ for any state of the world $\omega \in \Omega$;
2. The process F_t is nonanticipating for $t \in [a, b]$;
3. $E\Big[\displaystyle\int_a^b F_t^2\, dt\Big] < \infty$.

For instance, the following stochastic integrals exist:

$$\int_0^T W_t^2\, dW_t, \quad \int_0^T \sin(W_t)\, dW_t, \quad \int_a^b \frac{\cos(W_t)}{t}\, dW_t.$$

4.4 Examples of Ito integrals

As in the case of the Riemann integral, using the definition is not an efficient way of computing integrals. The same philosophy applies to Ito integrals. We shall compute in the following two simple Ito integrals. In later sections we shall introduce more efficient methods for computing Ito integrals.

4.4.1 The case $F_t = c$, constant

In this case the partial sums can be computed explicitly

$$\begin{aligned} S_n &= \sum_{i=0}^{n-1} F_{t_i}(W_{t_{i+1}} - W_{t_i}) = \sum_{i=0}^{n-1} c(W_{t_{i+1}} - W_{t_i}) \\ &= c(W_b - W_a), \end{aligned}$$

and since the answer does not depend on n, we have

$$\int_a^b c\, dW_t = c(W_b - W_a).$$

In particular, taking $c = 1$, $a = 0$, and $b = T$, since the Brownian motion starts at 0, we have the following formula:

$$\boxed{\int_0^T dW_t = W_T.}$$

4.4.2 The case $F_t = W_t$

We shall integrate the process W_t between 0 and T. Considering an equidistant partition, we take $t_k = \dfrac{kT}{n}$, $k = 0, 1, \cdots, n-1$. The partial sums are given by

$$S_n = \sum_{i=0}^{n-1} W_{t_i}(W_{t_{i+1}} - W_{t_i}).$$

Since

$$xy = \frac{1}{2}[(x+y)^2 - x^2 - y^2],$$

letting $x = W_{t_i}$ and $y = W_{t_{i+1}} - W_{t_i}$ yields

$$W_{t_i}(W_{t_{i+1}} - W_{t_i}) = \frac{1}{2}W_{t_{i+1}}^2 - \frac{1}{2}W_{t_i}^2 - \frac{1}{2}(W_{t_{i+1}} - W_{t_i})^2.$$

Then after pair cancelations the sum becomes

$$\begin{aligned} S_n &= \frac{1}{2}\sum_{i=0}^{n-1} W_{t_{i+1}}^2 - \frac{1}{2}\sum_{i=0}^{n-1} W_{t_i}^2 - \frac{1}{2}\sum_{i=0}^{n-1}(W_{t_{i+1}} - W_{t_i})^2 \\ &= \frac{1}{2}W_{t_n}^2 - \frac{1}{2}\sum_{i=0}^{n-1}(W_{t_{i+1}} - W_{t_i})^2. \end{aligned}$$

Using $t_n = T$, we get

$$S_n = \frac{1}{2}W_T^2 - \frac{1}{2}\sum_{i=0}^{n-1}(W_{t_{i+1}} - W_{t_i})^2.$$

Since the first term on the right side is independent of n, using Proposition 3.11.1, we have

$$\text{ms-}\lim_{n\to\infty} S_n = \frac{1}{2}W_T^2 - \text{ms-}\lim_{n\to\infty} \frac{1}{2}\sum_{i=0}^{n-1}(W_{t_{i+1}} - W_{t_i})^2 \tag{4.4.2}$$

$$= \frac{1}{2}W_T^2 - \frac{1}{2}T. \tag{4.4.3}$$

We have now obtained the following explicit formula of a stochastic integral:

$$\boxed{\int_0^T W_t\, dW_t = \frac{1}{2}W_T^2 - \frac{1}{2}T.}$$

In a similar way one can obtain

$$\boxed{\int_a^b W_t\, dW_t = \frac{1}{2}(W_b^2 - W_a^2) - \frac{1}{2}(b-a).}$$

It is worth noting that the right side contains random variables depending on the limits of integration a and b.

Exercise 4.4.1 *Show the following identities:*

(a) $E[\int_0^T dW_t] = 0$;

(b) $E[\int_0^T W_t dW_t] = 0$;

(c) $Var[\displaystyle\int_0^T W_t dW_t] = \frac{T^2}{2}$.

4.5 Properties of the Ito Integral

We shall start with some properties which are similar with those of the Riemannian integral.

Proposition 4.5.1 *Let $f(W_t, t)$, $g(W_t, t)$ be nonanticipating processes and $c \in \mathbb{R}$. Then we have*

1. Additivity:

$$\int_0^T [f(W_t,t) + g(W_t,t)]\, dW_t = \int_0^T f(W_t,t)\, dW_t + \int_0^T g(W_t,t)\, dW_t.$$

2. Homogeneity:

$$\int_0^T cf(W_t,t)\, dW_t = c\int_0^T f(W_t,t)\, dW_t.$$

3. Partition property:

$$\int_0^T f(W_t,t)\, dW_t = \int_0^u f(W_t,t)\, dW_t + \int_u^T f(W_t,t)\, dW_t, \qquad \forall 0 < u < T.$$

Proof: 1. Consider the partial sum sequences

$$X_n = \sum_{i=0}^{n-1} f(W_{t_i}, t_i)(W_{t_{i+1}} - W_{t_i})$$
$$Y_n = \sum_{i=0}^{n-1} g(W_{t_i}, t_i)(W_{t_{i+1}} - W_{t_i}).$$

Since $\text{ms-}\lim_{n\to\infty} X_n = \int_0^T f(W_t, t)\, dW_t$ and $\text{ms-}\lim_{n\to\infty} Y_n = \int_0^T g(W_t, t)\, dW_t$, using Proposition 1.14.2 yields

$$\begin{aligned}
&\int_0^T \Big(f(W_t, t) + g(W_t, t)\Big)\, dW_t \\
= \; & \text{ms-}\lim_{n\to\infty} \sum_{i=0}^{n-1} \Big(f(W_{t_i}, t_i) + g(W_{t_i}, t_i)\Big)(W_{t_{i+1}} - W_{t_i}) \\
= \; & \text{ms-}\lim_{n\to\infty} \Big[\sum_{i=0}^{n-1} \Big(f(W_{t_i}, t_i)(W_{t_{i+1}} - W_{t_i}) + \sum_{i=0}^{n-1} g(W_{t_i}, t_i)(W_{t_{i+1}} - W_{t_i})\Big] \\
= \; & \text{ms-}\lim_{n\to\infty} (X_n + Y_n) = \text{ms-}\lim_{n\to\infty} X_n + \text{ms-}\lim_{n\to\infty} Y_n \\
= \; & \int_0^T f(W_t, t)\, dW_t + \int_0^T g(W_t, t)\, dW_t.
\end{aligned}$$

The proofs of parts 2 and 3 are left as an exercise for the reader. ∎

Some other properties, such as monotonicity, do not hold in general. It is possible to have a nonnegative random variable F_t for which the random variable $\int_0^T F_t\, dW_t$ has negative values. More precisely, let $F_t = 1$. Then $F_t > 0$ but $\int_0^T 1\, dW_t = W_T$ it is not always positive. The probability to be negative is $P(W_T < 0) = 1/2$.

Some of the random variable properties of the Ito integral are given by the following result:

Proposition 4.5.2 *We have*

1. Zero mean:

$$E\Big[\int_a^b f(W_t, t)\, dW_t\Big] = 0.$$

2. Isometry:

$$E\Big[\Big(\int_a^b f(W_t, t)\, dW_t\Big)^2\Big] = E\Big[\int_a^b f(W_t, t)^2\, dt\Big].$$

3. Covariance:

$$E\Big[\Big(\int_a^b f(W_t, t)\, dW_t\Big)\Big(\int_a^b g(W_t, t)\, dW_t\Big)\Big] = E\Big[\int_a^b f(W_t, t) g(W_t, t)\, dt\Big].$$

We shall discuss the previous properties giving rough reasons why they hold true. The detailed proofs are beyond the goal of this book.

1. The Ito integral $I = \int_a^b f(W_t, t)\, dW_t$ is the mean square limit of the partial sums $S_n = \sum_{i=0}^{n-1} f_{t_i}(W_{t_{i+1}} - W_{t_i})$, where we denoted $f_{t_i} = f(W_{t_i}, t_i)$. Since $f(W_t, t)$ is a nonanticipative process, then f_{t_i} is independent of the increments $W_{t_{i+1}} - W_{t_i}$, and hence we have

$$\begin{aligned} E[S_n] &= E\Big[\sum_{i=0}^{n-1} f_{t_i}(W_{t_{i+1}} - W_{t_i})\Big] = \sum_{i=0}^{n-1} E[f_{t_i}(W_{t_{i+1}} - W_{t_i})] \\ &= \sum_{i=0}^{n-1} E[f_{t_i}]E[(W_{t_{i+1}} - W_{t_i})] = 0, \end{aligned}$$

because the increments have mean zero. Applying the Squeeze Theorem in the double inequality

$$0 \le \Big(E[S_n - I]\Big)^2 \le E[(S_n - I)^2] \to 0, \qquad n \to \infty$$

yields $E[S_n] - E[I] \to 0$. Since $E[S_n] = 0$ it follows that $E[I] = 0$, i.e. the Ito integral has zero mean.

2. Since the square of the sum of partial sums can be written as

$$\begin{aligned} S_n^2 &= \Big(\sum_{i=0}^{n-1} f_{t_i}(W_{t_{i+1}} - W_{t_i})\Big)^2 \\ &= \sum_{i=0}^{n-1} f_{t_i}^2 (W_{t_{i+1}} - W_{t_i})^2 + 2\sum_{i \ne j} f_{t_i}(W_{t_{i+1}} - W_{t_i}) f_{t_j}(W_{t_{j+1}} - W_{t_j}), \end{aligned}$$

using the independence yields

$$\begin{aligned} E[S_n^2] &= \sum_{i=0}^{n-1} E[f_{t_i}^2]E[(W_{t_{i+1}} - W_{t_i})^2] \\ &\quad +2\sum_{i \ne j} E[f_{t_i}]E[(W_{t_{i+1}} - W_{t_i})]E[f_{t_j}]E[(W_{t_{j+1}} - W_{t_j})] \\ &= \sum_{i=0}^{n-1} E[f_{t_i}^2](t_{i+1} - t_i), \end{aligned}$$

which are the Riemann sums of the integral $\displaystyle\int_a^b E[f_t^2]\, dt = E\Big[\int_a^b f_t^2\, dt\Big]$, where the last identity follows from Fubini's theorem. Hence $E[S_n^2]$ converges to the aforementioned integral.

3. Consider the partial sums

$$S_n = \sum_{i=0}^{n-1} f_{t_i}(W_{t_{i+1}} - W_{t_i}), \qquad V_n = \sum_{j=0}^{n-1} g_{t_j}(W_{t_{j+1}} - W_{t_j}).$$

Their product is

$$\begin{aligned} S_n V_n &= \Big(\sum_{i=0}^{n-1} f_{t_i}(W_{t_{i+1}} - W_{t_i})\Big)\Big(\sum_{j=0}^{n-1} g_{t_j}(W_{t_{j+1}} - W_{t_j})\Big) \\ &= \sum_{i=0}^{n-1} f_{t_i} g_{t_i}(W_{t_{i+1}} - W_{t_i})^2 + \sum_{i\neq j} f_{t_i} g_{t_j}(W_{t_{i+1}} - W_{t_i})(W_{t_{j+1}} - W_{t_j}) \end{aligned}$$

Using that f_t and g_t are nonanticipative and that

$$\begin{aligned} E[(W_{t_{i+1}} - W_{t_i})(W_{t_{j+1}} - W_{t_j})] &= E[W_{t_{i+1}} - W_{t_i}]E[W_{t_{j+1}} - W_{t_j}] = 0, \ i \neq j \\ E[(W_{t_{i+1}} - W_{t_i})^2] &= t_{i+1} - t_i, \end{aligned}$$

it follows that

$$\begin{aligned} E[S_n V_n] &= \sum_{i=0}^{n-1} E[f_{t_i} g_{t_i}] E[(W_{t_{i+1}} - W_{t_i})^2] \\ &= \sum_{i=0}^{n-1} E[f_{t_i} g_{t_i}](t_{i+1} - t_i), \end{aligned}$$

which is the Riemann sum for the integral $\int_a^b E[f_t g_t]\, dt$.

From 1 and 2 it follows that the random variable $\int_a^b f(W_t, t)\, dW_t$ has mean zero and variance

$$Var\Big[\int_a^b f(W_t, t)\, dW_t\Big] = E\Big[\int_a^b f(W_t, t)^2\, dt\Big].$$

From 1 and 3 it follows that

$$Cov\Big[\int_a^b f(W_t, t)\, dW_t, \int_a^b g(W_t, t)\, dW_t\Big] = \int_a^b E[f(W_t, t) g(W_t, t)]\, dt.$$

Corollary 4.5.3 (Cauchy's integral inequality) *Let $f_t = f(W_t, t)$ and $g_t = g(W_t, t)$. Then*

$$\Big(\int_a^b E[f_t g_t]\, dt\Big)^2 \leq \Big(\int_a^b E[f_t^2]\, dt\Big)\Big(\int_a^b E[g_t^2]\, dt\Big).$$

Proof: It follows from the previous theorem and from the correlation formula $|Corr(X,Y)| = \dfrac{|Cov(X,Y)|}{[Var(X)Var(Y)]^{1/2}} \le 1.$ ∎

Let $\mathcal{F}_t$ be the information set at time t. This implies that f_{t_i} and $W_{t_{i+1}} - W_{t_i}$ are known at time t, for any $t_{i+1} \le t$. It follows that the partial sum $S_n = \sum_{i=0}^{n-1} f_{t_i}(W_{t_{i+1}} - W_{t_i})$ is $\mathcal{F}_t$-predictable. The following result, whose proof is omited for technical reasons, states that this is also valid after taking the limit in the mean square:

Proposition 4.5.4 *The Ito integral $\int_0^t f_s\, dW_s$ is $\mathcal{F}_t$-predictable.*

The following two results state that if the upper limit of an Ito integral is replaced by the parameter t we obtain a continuous martingale.

Proposition 4.5.5 *For any $s < t$ we have*

$$E\Big[\int_0^t f(W_u,u)\, dW_u|\mathcal{F}_s\Big] = \int_0^s f(W_u,u)\, dW_u.$$

Proof: Using part 3 of Proposition 4.5.2 we get

$$\begin{aligned} & E\Big[\int_0^t f(W_u,u)\, dW_u|\mathcal{F}_s\Big] \\ = \;& E\Big[\int_0^s f(W_u,u)\, dW_u + \int_s^t f(W_u,u)\, dW_u|\mathcal{F}_s\Big] \\ = \;& E\Big[\int_0^s f(W_u,u)\, dW_u|\mathcal{F}_s\Big] + E\Big[\int_s^t f(W_u,u)\, dW_u|\mathcal{F}_s\Big]. \end{aligned} \tag{4.5.4}$$

Since $\int_0^s f(W_u,u)\, dW_u$ is $\mathcal{F}_s$-predictable (see Proposition 4.5.4), by part 2 of Proposition 1.11.4

$$E\Big[\int_0^s f(W_u,u)\, dW_u|\mathcal{F}_s\Big] = \int_0^s f(W_u,u)\, dW_u.$$

Since $\int_s^t f(W_u,u)\, dW_u$ contains only information between s and t, it is independent of the information set $\mathcal{F}_s$, so we can drop the condition in the expectation; using that Ito integrals have zero mean we obtain

$$E\Big[\int_s^t f(W_u,u)\, dW_u|\mathcal{F}_s\Big] = E\Big[\int_s^t f(W_u,u)\, dW_u\Big] = 0.$$

Substituting into (4.5.4) yields the desired result. ∎

Proposition 4.5.6 *Consider the process $X_t = \int_0^t f(W_s,s)\, dW_s$. Then X_t is continuous, i.e. for almost any state of the world $\omega \in \Omega$, the path $t \to X_t(\omega)$ is continuous.*

Proof: A rigorous proof is beyond the purpose of this book. We shall provide a rough sketch. Assume the process $f(W_t,t)$ satisfies $E[f(W_t,t)^2] < M$, for some $M > 0$. Let t_0 be fixed and consider $h > 0$. Consider the increment $Y_h = X_{t_0+h} - X_{t_0}$. Using the aforementioned properties of the Ito integral we have

$$\begin{aligned} E[Y_h] &= E[X_{t_0+h} - X_{t_0}] = E\Big[\int_{t_0}^{t_0+h} f(W_t,t)\,dW_t\Big] = 0 \\ E[Y_h^2] &= E\Big[\Big(\int_{t_0}^{t_0+h} f(W_t,t)\,dW_t\Big)^2\Big] = \int_{t_0}^{t_0+h} E[f(W_t,t)^2]\,dt \\ &< M\int_{t_0}^{t_0+h} dt = Mh. \end{aligned}$$

The process Y_h has zero mean for any $h > 0$ and its variance tends to 0 as $h \to 0$. Using a convergence theorem yields that Y_h tends to 0 in mean square, as $h \to 0$. This is equivalent with the continuity of X_t at t_0. ■

Proposition 4.5.7 *Let $X_t = \int_0^t f(W_s,s)\,dW_s$, with $E\big[\int_0^\infty f^2(s,W_s)\,ds\big] < \infty$. Then X_t is a continuous $\mathcal{F}_t$-martingale.*

Proof: We shall check in the following the properties of a martingale.
Integrability: Using properties of Ito integrals

$$E[X_t^2] = E\Big[\Big(\int_0^t f(W_s,s)\,dW_s\Big)^2\Big] = E\Big[\int_0^t f^2(W_s,s)\,ds\Big] < E\Big[\int_0^\infty f^2(W_s,s)\,ds\Big] < \infty,$$

and then from the inequality $E[|X_t|]^2 \le E[X_t^2]$ we obtain $E[|X_t|] < \infty$, for all $t \ge 0$.
Predictability: X_t is $\mathcal{F}_t$-predictable from Proposition 4.5.4.
Forecast: $E[X_t|\mathcal{F}_s] = X_s$ for $s < t$ by Proposition 4.5.5.
Continuity: See Proposition 4.5.6. ■

4.6 The Wiener Integral

The *Wiener integral* is a particular case of the Ito stochastic integral. It is obtained by replacing the nonanticipating stochastic process $f(W_t,t)$ by the deterministic function $f(t)$. The Wiener integral $\int_a^b f(t)\,dW_t$ is the mean square limit of the partial sums

$$S_n = \sum_{i=0}^{n-1} f(t_i)(W_{t_{i+1}} - W_{t_i}).$$

All properties of Ito integrals also hold for Wiener integrals. The Wiener integral is a random variable with zero mean

$$E\Big[\int_a^b f(t)\,dW_t\Big] = 0$$

and variance

$$E\Big[\Big(\int_a^b f(t)\,dW_t\Big)^2\Big] = \int_a^b f(t)^2\,dt.$$

However, in the case of Wiener integrals we can say something about their distribution.

Proposition 4.6.1 *The Wiener integral $I(f) = \int_a^b f(t)\,dW_t$ is a normal random variable with mean 0 and variance*

$$Var[I(f)] = \int_a^b f(t)^2\,dt := \|f\|^2_{L^2}.$$

Proof: Since increments $W_{t_{i+1}} - W_{t_i}$ are normally distributed with mean 0 and variance $t_{i+1} - t_i$, then

$$f(t_i)(W_{t_{i+1}} - W_{t_i}) \sim N(0, f(t_i)^2(t_{i+1} - t_i)).$$

Since these random variables are independent, by the Central Limit Theorem (see Theorem 2.3.1), their sum is also normally distributed, with

$$S_n = \sum_{i=0}^{n-1} f(t_i)(W_{t_{i+1}} - W_{t_i}) \sim N\Big(0, \sum_{i=0}^{n-1} f(t_i)^2(t_{i+1} - t_i)\Big).$$

Taking $n \to \infty$ and $\max\limits_i \|t_{i+1} - t_i\| \to 0$, the normal distribution tends to

$$N\Big(0, \int_a^b f(t)^2\,dt\Big).$$

The previous convergence holds in distribution, and it still needs to be shown in the mean square. However, we shall omit this essential proof detail. ■

Exercise 4.6.2 *Show that the random variable $X = \int_1^T \frac{1}{\sqrt{t}}\,dW_t$ is normally distributed with mean 0 and variance $\ln T$.*

Exercise 4.6.3 *Let $Y = \int_1^T \sqrt{t}\,dW_t$. Show that Y is normally distributed with mean 0 and variance $(T^2 - 1)/2$.*

Exercise 4.6.4 *Find the distribution of the integral $\int_0^t e^{t-s}\,dW_s$.*

Exercise 4.6.5 *Show that $X_t = \int_0^t (2t-u)\,dW_u$ and $Y_t = \int_0^t (3t-4u)\,dW_u$ are Gaussian processes with mean 0 and variance $\frac{7}{3}t^3$.*

Exercise 4.6.6 *Show that ms-$\lim\limits_{t\to 0} \dfrac{1}{t}\displaystyle\int_0^t u\,dW_u = 0$.*

Exercise 4.6.7 *Find all constants a, b such that $X_t = \int_0^t \Big(a + \frac{bu}{t}\Big)\,dW_u$ is normally distributed with variance t.*

Exercise 4.6.8 *Let n be a positive integer. Prove that*

$$Cov\Big(W_t, \int_0^t u^n \, dW_u\Big) = \frac{t^{n+1}}{n+1}.$$

Formulate and prove a more general result.

4.7 Poisson Integration

In this section we deal with the integration with respect to the compensated Poisson process $M_t = N_t - \lambda t$, which is a martingale. Consider $0 \leq a < b$ and let $F_t = F(t, M_t)$ be a nonanticipating process with

$$E\Big[\int_a^b F_t^2 \, dt\Big] < \infty.$$

Consider the partition

$$a = t_0 < t_1 < \cdots < t_{n-1} < t_n = b$$

of the interval $[a, b]$, and associate the partial sums

$$S_n = \sum_{i=0}^{n-1} F_{t_{i-}}(M_{t_{i+1}} - M_{t_i}),$$

where $F_{t_{i-}}$ is the left-hand limit at t_{i-}. For predictability reasons, the intermediate points are the left-handed limit to the endpoints of each interval. Since the process F_t is nonanticipative, the random variables $F_{t_{i-}}$ and $M_{t_{i+1}} - M_{t_i}$ are independent.

The integral of F_{t-} with respect to M_t is the mean square limit of the partial sum S_n

$$\text{ms-}\lim_{n\to\infty} S_n = \int_0^T F_{t-} \, dM_t,$$

provided the limit exists. More precisely, this convergence means that

$$\lim_{n\to\infty} E\Big[\Big(S_n - \int_a^b F_{t-} \, dM_t\Big)^2\Big] = 0.$$

Exercise 4.7.1 *Let c be a constant. Show that* $\displaystyle\int_a^b c \, dM_t = c(M_b - M_a)$.

4.8 An Work Out Example: the case $F_t = M_t$

We shall integrate the process M_{t-} between 0 and T with respect to M_t. Consider the equidistant partition points $t_k = \dfrac{kT}{n}$, $k = 0, 1, \cdots, n-1$. Then the partial sums are given by

$$S_n = \sum_{i=0}^{n-1} M_{t_i-}(M_{t_{i+1}} - M_{t_i}).$$

Using $xy = \frac{1}{2}[(x+y)^2 - x^2 - y^2]$, by letting $x = M_{t_i-}$ and $y = M_{t_{i+1}} - M_{t_i}$, we get

$$M_{t_i-}(M_{t_{i+1}} - M_{t_i}) \quad = \quad \frac{1}{2}(M_{t_{i+1}} - M_{t_i} + M_{t_i-})^2 - \frac{1}{2}M^2_{t_i-} - \frac{1}{2}(M_{t_{i+1}} - M_{t_i})^2.$$

Let J be the set of jump instances between 0 and T. Using that $M_{t_i} = M_{t_i-}$ for $t_i \notin J$, and $M_{t_i} = 1 + M_{t_i-}$ for $t_i \in J$ yields

$$M_{t_{i+1}} - M_{t_i} + M_{t_i-} = \begin{cases} M_{t_{i+1}}, & \text{if } t_i \notin J \\ M_{t_{i+1}} - 1, & \text{if } t_i \in J. \end{cases}$$

Splitting the sum, canceling in pairs, and applying the difference of squares formula we have

$$\begin{aligned}
S_n &= \frac{1}{2}\sum_{i=0}^{n-1}(M_{t_{i+1}} - M_{t_i} + M_{t_i-})^2 - \frac{1}{2}\sum_{i=0}^{n-1} M^2_{t_i-} - \frac{1}{2}\sum_{i=0}^{n-1}(M_{t_{i+1}} - M_{t_i})^2 \\
&= \frac{1}{2}\sum_{t_i\in J}(M_{t_{i+1}} - 1)^2 + \frac{1}{2}\sum_{t_i\notin J} M^2_{t_{i+1}} - \frac{1}{2}\sum_{t_i\notin J} M^2_{t_i} - \frac{1}{2}\sum_{t_i\in J} M^2_{t_i-} - \frac{1}{2}\sum_{i=0}^{n-1}(M_{t_{i+1}} - M_{t_i})^2 \\
&= \frac{1}{2}\sum_{t_i\in J}\Big((M_{t_{i+1}} - 1)^2 - M^2_{t_i-}\Big) + \frac{1}{2}M^2_{t_n} - \frac{1}{2}\sum_{i=0}^{n-1}(M_{t_{i+1}} - M_{t_i})^2 \\
&= \frac{1}{2}\sum_{t_i\in J}\underbrace{(M_{t_{i+1}} - 1 - M_{t_i-})}_{=0}(M_{t_{i+1}} - 1 + M_{t_i-}) + \frac{1}{2}M^2_{t_n} - \frac{1}{2}\sum_{i=0}^{n-1}(M_{t_{i+1}} - M_{t_i})^2 \\
&= \frac{1}{2}M^2_{t_n} - \frac{1}{2}\sum_{i=0}^{n-1}(M_{t_{i+1}} - M_{t_i})^2.
\end{aligned}$$

Hence we have arrived at the following formula

$$\boxed{\int_0^T M_{t-}\,dM_t = \frac{1}{2}M_T^2 - \frac{1}{2}N_T.}$$

Similarly, one can obtain

$$\boxed{\int_a^b M_{t-}\,dM_t = \frac{1}{2}(M_b^2 - M_a^2) - \frac{1}{2}(N_b - N_a).}$$

Exercise 4.8.1 (a) *Show that* $E\Big[\int_a^b M_{t-}\,dM_t\Big] = 0,$

(b) *Find* $Var\Big[\int_a^b M_t\,dM_t\Big].$

Remark 4.8.2 (a) *Let* ω *be a fixed state of the world and assume the sample path* $t \to N_t(\omega)$ *has a jump in the interval* (a,b). *Even if beyound the scope of this book, one can be shown that the integral*

$$\int_a^b N_t(\omega)\,dN_t$$

does not exist in the Riemann-Stieltjes sense.

(b) *Let* N_{t-} *denote the left-hand limit of* N_t. *It can be shown that* N_{t-} *is predictable, while* N_t *is not.*

The previous remarks provide the reason why in the following we shall work with M_{t-} *instead of* M_t*: the integral* $\int_a^b M_t\,dN_t$ *might not exist, while* $\int_a^b M_{t-}\,dN_t$ *does exist.*

Exercise 4.8.3 *Show that*

$$\int_0^T N_{t-}\,dM_t = \frac{1}{2}(N_t^2 - N_t) - \lambda\int_0^t N_t\,dt.$$

Exercise 4.8.4 *Find the variance of*

$$\int_0^T N_{t-}\,dM_t.$$

The following integrals with respect to a Poisson process N_t are considered in the Riemann-Stieltjes sense.

Proposition 4.8.5 *For any continuous function* f *we have*

(a) $E\Big[\int_0^t f(s)\,dN_s\Big] = \lambda\int_0^t f(s)\,ds;$

(b) $E\Big[\Big(\int_0^t f(s)\,dN_s\Big)^2\Big] = \lambda\int_0^t f(s)^2\,ds + \lambda^2\Big(\int_0^t f(s)\,ds\Big)^2;$

(c) $E\Big[e^{\int_0^t f(s)\,dN_s}\Big] = e^{\lambda\int_0^t (e^{f(s)}-1)\,ds}.$

Proof: (a) Consider the equidistant partition $0 = s_0 < s_1 < \cdots < s_n = t$, with $s_{k+1} - s_k = \Delta s$. Then

$$
\begin{aligned}
E\Big[\int_0^t f(s)\,dN_s\Big] &= \lim_{n\to\infty} E\Big[\sum_{i=0}^{n-1} f(s_i)(N_{s_{i+1}} - N_{s_i})\Big] = \lim_{n\to\infty} \sum_{i=0}^{n-1} f(s_i) E\Big[N_{s_{i+1}} - N_{s_i}\Big] \\
&= \lambda \lim_{n\to\infty} \sum_{i=0}^{n-1} f(s_i)(s_{i+1} - s_i) = \lambda \int_0^t f(s)\,ds.
\end{aligned}
$$

(b) Using that N_t is stationary and has independent increments, we have respectively

$$
\begin{aligned}
E[(N_{s_{i+1}} - N_{s_i})^2] &= E[N^2_{s_{i+1}-s_i}] = \lambda(s_{i+1} - s_i) + \lambda^2(s_{i+1} - s_i)^2 \\
&= \lambda\,\Delta s + \lambda^2(\Delta s)^2, \\
E[(N_{s_{i+1}} - N_{s_i})(N_{s_{j+1}} - N_{s_j})] &= E[(N_{s_{i+1}} - N_{s_i})]E[(N_{s_{j+1}} - N_{s_j})] \\
&= \lambda(s_{i+1} - s_i)\lambda(s_{j+1} - s_j) = \lambda^2(\Delta s)^2.
\end{aligned}
$$

Applying the expectation to the formula

$$
\begin{aligned}
\Big(\sum_{i=0}^{n-1} f(s_i)(N_{s_{i+1}} - N_{s_i})\Big)^2 &= \sum_{i=0}^{n-1} f(s_i)^2(N_{s_{i+1}} - N_{s_i})^2 \\
&\quad +2\sum_{i\neq j} f(s_i)f(s_j)(N_{s_{i+1}} - N_{s_i})(N_{s_{j+1}} - N_{s_j})
\end{aligned}
$$

yields

$$
\begin{aligned}
E\Big[\Big(\sum_{i=0}^{n-1} f(s_i)(N_{s_{i+1}} - N_{s_i})\Big)^2\Big] &= \sum_{i=0}^{n-1} f(s_i)^2(\lambda\Delta s + \lambda^2(\Delta s)^2) + 2\sum_{i\neq j} f(s_i)f(s_j)\lambda^2(\Delta s)^2 \\
&= \lambda\sum_{i=0}^{n-1} f(s_i)^2\Delta s + \lambda^2\Big[\sum_{i=0}^{n-1} f(s_i)^2(\Delta s)^2 + 2\sum_{i\neq j} f(s_i)f(s_j)(\Delta s)^2\Big] \\
&= \lambda\sum_{i=0}^{n-1} f(s_i)^2\Delta s + \lambda^2\Big(\sum_{i=0}^{n-1} f(s_i)\,\Delta s\Big)^2 \\
&\to \lambda\int_0^t f(s)^2\,ds + \lambda^2\Big(\int_0^t f(s)\,ds\Big)^2, \text{ as } n\to\infty.
\end{aligned}
$$

(c) Using that N_t is stationary with independent increments and has the moment

generating function $E[e^{kN_t}] = e^{\lambda(e^k-1)t}$, we have

$$\begin{aligned}
E\Big[e^{\int_0^t f(s)\,dN_s}\Big] &= \lim_{n\to\infty} E\Big[e^{\sum_{i=0}^{n-1} f(s_i)(N_{s_{i+1}}-N_{s_i})}\Big] = \lim_{n\to\infty} E\Big[\prod_{i=0}^{n-1} e^{f(s_i)(N_{s_{i+1}}-N_{s_i})}\Big] \\
&= \lim_{n\to\infty} \prod_{i=0}^{n-1} E\Big[e^{f(s_i)(N_{s_{i+1}}-N_{s_i})}\Big] = \lim_{n\to\infty} \prod_{i=0}^{n-1} E\Big[e^{f(s_i)(N_{s_{i+1}-s_i})}\Big] \\
&= \lim_{n\to\infty} \prod_{i=0}^{n-1} e^{\lambda(e^{f(s_i)}-1)(s_{i+1}-s_i)} = \lim_{n\to\infty} e^{\lambda \sum_{i=0}^{n-1}(e^{f(s_i)}-1)(s_{i+1}-s_i)} \\
&= e^{\lambda \int_0^t (e^{f(s)}-1)\,ds}.
\end{aligned}$$

■

Since f is continuous, the Poisson integral $\int_0^t f(s)\,dN_s$ can be computed in terms of the waiting times S_k

$$\int_0^t f(s)\,dN_s = \sum_{k=1}^{N_t} f(S_k).$$

This formula can be used to give a proof for the previous result. For instance, taking the expectation and using conditions over $N_t = n$, yields

$$\begin{aligned}
E\Big[\int_0^t f(s)\,dN_s\Big] &= E\Big[\sum_{k=1}^{N_t} f(S_k)\Big] = \sum_{n\geq 0} E\Big[\sum_{k=1}^{n} f(S_k)|N_t = n\Big] P(N_t = n) \\
&= \sum_{n\geq 0} \frac{n}{t} \int_0^t f(x)\,dx \,\frac{(\lambda t)^n}{n!} e^{-\lambda t} = e^{-\lambda t} \int_0^t f(x)\,dx\, \frac{1}{t} \sum_{n\geq 0} \frac{(\lambda t)^n}{(n-1)!} \\
&= e^{-\lambda t} \int_0^t f(x)\,dx\, \lambda e^{\lambda t} = \lambda \int_0^t f(x)\,dx.
\end{aligned}$$

Exercise 4.8.6 *Solve parts (b) and (c) of Proposition 4.8.5 using a similar idea with the one presented above.*

Exercise 4.8.7 *Show that*

$$E\Big[\Big(\int_0^t f(s)\,dM_s\Big)^2\Big] = \lambda \int_0^t f(s)^2\,ds,$$

where $M_t = N_t - \lambda t$ is the compensated Poisson process.

Exercise 4.8.8 *Prove that*

$$Var\Big(\int_0^t f(s)\,dN_s\Big) = \lambda \int_0^t f(s)^2\,dN_s.$$

Exercise 4.8.9 *Find*

$$E\Big[e^{\int_0^t f(s)\,dM_s}\Big].$$

Proposition 4.8.10 *Let $\mathcal{F}_t = \sigma(N_s; 0 \leq s \leq t)$. Then for any constant c, the process*

$$M_t = e^{cN_t + \lambda(1-e^c)t}, \qquad t \geq 0$$

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

Proof: Let $s < t$. Since $N_t - N_s$ is independent of $\mathcal{F}_s$ and N_t is stationary, we have

$$\begin{aligned} E[e^{c(N_t-N_s)}|\mathcal{F}_s] &= E[e^{c(N_t-N_s)}] = E[e^{cN_{t-s}}] \\ &= e^{\lambda(e^c-1)(t-s)}. \end{aligned}$$

On the other side, taking out the predictable part yields

$$E[e^{c(N_t-N_s)}|\mathcal{F}_s] = e^{-cN_s}E[e^{cN_t}|\mathcal{F}_s].$$

Equating the last two relations we arrive at

$$E[e^{cN_t+(1-e^c)t}|\mathcal{F}_s] = e^{cN_s+\lambda(1-e^c)s},$$

which is equivalent with the martingale condition $E[M_t|\mathcal{F}_s] = M_s$. ∎

We shall present an application of the previous result. Consider the waiting time until the nth jump, $S_n = \inf\{t > 0; N_t = n\}$, which is a stopping time, and the filtration $\mathcal{F}_t = \sigma(N_s; 0 \leq s \leq t)$. Since

$$M_t = e^{cN_t+\lambda(1-e^c)t}$$

is an $\mathcal{F}_t$-martingale, by the Optional Stopping Theorem (Theorem 3.2.1) we have $E[M_{S_n}] = E[M_0] = 1$, which is equivalent with $E[e^{\lambda(1-e^c)S_n}] = e^{-cn}$. Substituting $s = -\lambda(1-e^c)$, then $c = \ln(1+\frac{s}{\lambda})$. Since $s, \lambda > 0$, then $c > 0$. The previous expression becomes

$$E[e^{-sS_n}] = e^{-n\ln(1+\frac{s}{\lambda})} = \Big(\frac{\lambda}{\lambda+s}\Big)^x.$$

Since the expectation on the left side is the Laplace transform of the probability density of S_n, then

$$\begin{aligned} p(S_n) &= \mathcal{L}^{-1}\{E[e^{-sS_n}]\} = \mathcal{L}^{-1}\Big\{\Big(\frac{\lambda}{\lambda+s}\Big)^x\Big\} \\ &= \frac{e^{-t\lambda}t^{n-1}\lambda^n}{\Gamma(n)}, \end{aligned}$$

which shows that S_n has a gamma distribution.

4.9 The distribution function of $X_T = \int_0^T g(t)\,dN_t$

In this section we consider the function $g(t)$ continuous. Let $S_1 < S_2 < \cdots < S_{N_t}$ denote the waiting times until time t. Since the increments dN_t are equal to 1 at S_k and 0 otherwise, the integral can be written as

$$X_T = \int_0^T g(t)\,dN_t = g(S_1) + \cdots + g(S_{N_t}).$$

The distribution function of the random variable $X_T = \int_0^T g(t)\,dN_t$ can be obtained conditioning over the N_t

$$\begin{aligned} P(X_T \le u) &= \sum_{k\ge 0} P(X_T \le u | N_T = k)\, P(N_T = k) \\ &= \sum_{k\ge 0} P(g(S_1) + \cdots + g(S_{N_t}) \le u | N_T = k)\, P(N_T = k) \\ &= \sum_{k\ge 0} P(g(S_1) + \cdots + g(S_k) \le u)\, P(N_T = k). \end{aligned} \qquad (4.9.5)$$

Considering $S_1, S_2, \cdots, S_k$ independent and uniformly distributed over the interval $[0, T]$, we have

$$P\big(g(S_1) + \cdots + g(S_k) \le u\big) = \int_{D_k} \frac{1}{T^k}\, dx_1 \cdots dx_k = \frac{vol(D_k)}{T^k},$$

where

$$D_k = \{g(x_1) + g(x_2) + \cdots + g(x_k) \le u\} \cap \{0 \le x_1, \cdots, x_k \le T\}.$$

Substituting back in (4.9.5) yields

$$\begin{aligned} P(X_T \le u) &= \sum_{k\ge 0} P(g(S_1) + \cdots + g(S_k) \le u)\, P(N_T = k) \\ &= \sum_{k\ge 0} \frac{vol(D_k)}{T^k} \frac{\lambda^k T^k}{k!} e^{-\lambda T} = e^{-\lambda T} \sum_{k\ge 0} \frac{\lambda^k vol(D_k)}{k!}. \end{aligned} \qquad (4.9.6)$$

In general, the volume of the k-dimensional solid D_k is not obvious easy to obtain. However, there are simple cases when this can be computed explicitly.

A Particular Case. We shall do an explicit computation of the partition function of $X_T = \int_0^T s^2\,dN_s$. In this case the solid D_k is the intersection between the k-dimensional ball of radius $\sqrt{u}$ centered at the origin and the k-dimensional cube $[0, T]^k$. There are three possible shapes for D_k, which depend on the size of $\sqrt{u}$:

(a) if $0 \le \sqrt{u} < T$, then D_k is a $\frac{1}{2^k}$-part of a k-dimensional sphere;

(b) if $T \le \sqrt{u} < T\sqrt{k}$, then D_k has a complicated shape;

(c) if $T\sqrt{k} \leq \sqrt{u}$, then D_k is the entire k-dimensional cube, and then $vol(D_k) = T^k$.

Since the volume of the k-dimensional ball of radius R is given by $\dfrac{\pi^{k/2}R^k}{\Gamma(\frac{k}{2}+1)}$, then the volume of D_k in case (**a**) becomes

$$vol(D_k) = \frac{\pi^{k/2}u^{k/2}}{2^k\Gamma(\frac{k}{2}+1)}.$$

Substituting in (4.9.6) yields

$$P(X_T \leq u) = e^{-\lambda T}\sum_{k\geq 0}\frac{(\lambda^2\pi u)^{k/2}}{k!\Gamma(\frac{k}{2}+1)}, \qquad 0 \leq \sqrt{u} < T.$$

It is worth noting that for $u \to \infty$, the inequality $T\sqrt{k} \leq \sqrt{u}$ is satisfied for all $k \geq 0$; hence relation (4.9.6) yields

$$\lim_{u\to\infty} P(X_T \leq u) = e^{-\lambda T}\sum_{k\geq 0}\frac{\lambda^k T^k}{k!} = e^{-kT}e^{kT} = 1.$$

The computation in case (**b**) is more complicated and will be omitted.

Exercise 4.9.1 *Calculate the expectation* $E\Big[\int_0^T e^{ks}\,dN_s\Big]$ *and the variance* $Var\Big(\int_0^T e^{ks}\,dN_s\Big)$.

Exercise 4.9.2 *Compute the distribution function of* $X_t = \int_0^T s\,dN_s$.

Chapter 5

Stochastic Differentiation

5.1 Differentiation Rules

Most stochastic processes are not differentiable. For instance, the Brownian motion process W_t is a continuous process which is nowhere differentiable. Hence, derivatives like $\frac{dW_t}{dt}$ do not make sense in stochastic calculus. The only quantities allowed to be used are the infinitesimal changes of the process, in our case, dW_t.

The infinitesimal change of a process
The change in the process X_t between instances t and $t + \Delta t$ is given by $\Delta X_t = X_{t+\Delta t} - X_t$. When Δt is infinitesimally small, we obtain the infinitesimal change of a process X_t

$$\boxed{dX_t = X_{t+dt} - X_t.}$$

Sometimes it is useful to use the equivalent formula $X_{t+dt} = X_t + dX_t$.

5.2 Basic Rules

The following rules are the analog of some familiar differentiation rules from elementary Calculus.

The constant multiple rule
If X_t is a stochastic process and c is a constant, then

$$\boxed{d(c\,X_t) = c\,dX_t.}$$

The verification follows from a straightforward application of the infinitesimal change formula

$$d(c\,X_t) = c\,X_{t+dt} - c\,X_t = c(X_{t+dt} - X_t) = c\,dX_t.$$

The sum rule
If X_t and Y_t are two stochastic processes, then

$$\boxed{d(X_t + Y_t) = dX_t + dY_t.}$$

The verification is as in the following:

$$\begin{aligned} d(X_t + Y_t) &= (X_{t+dt} + Y_{t+dt}) - (X_t + Y_t) \\ &= (X_{t+dt} - X_t) + (Y_{t+dt} - Y_t) \\ &= dX_t + dY_t. \end{aligned}$$

The difference rule
If X_t and Y_t are two stochastic processes, then

$$\boxed{d(X_t - Y_t) = dX_t - dY_t.}$$

The proof is similar to the one for the sum rule.

The product rule
If X_t and Y_t are two stochastic processes, then

$$\boxed{d(X_t Y_t) = X_t\, dY_t + Y_t\, dX_t + dX_t\, dY_t.}$$

The proof is as follows:

$$\begin{aligned} d(X_t\, Y_t) &= X_{t+dt} Y_{t+dt} - X_t Y_t \\ &= X_t(Y_{t+dt} - Y_t) + Y_t(X_{t+dt} - X_t) + (X_{t+dt} - X_t)(Y_{t+dt} - Y_t) \\ &= X_t\, dY_t + Y_t\, dX_t + dX_t\, dY_t, \end{aligned}$$

where the second identity is verified by direct computation.

If the process X_t is replaced by the deterministic function $f(t)$, then the aforementioned formula becomes

$$\boxed{d(f(t)Y_t) = f(t)\, dY_t + Y_t\, df(t) + df(t)\, dY_t.}$$

Since in most practical cases the process Y_t is an Ito diffusion

$$dY_t = a(t, W_t)dt + b(t, W_t)dW_t,$$

using the relations $dt\, dW_t = dt^2 = 0$, the last term vanishes

$$df(t)\, dY_t = f'(t)dtdY_t = 0,$$

and hence

$$\boxed{d(f(t)Y_t) = f(t)\, dY_t + Y_t\, df(t).}$$

This relation looks like the usual product rule.

The quotient rule
If X_t and Y_t are two stochastic processes, then

$$\boxed{d\Big(\frac{X_t}{Y_t}\Big) = \frac{Y_t dX_t - X_t dY_t - dX_t dY_t}{Y_t^2} + \frac{X_t}{Y_t^3}(dY_t)^2.}$$

The proof follows from Ito's formula and will be addressed in section 5.3.3.

When the process Y_t is replaced by the deterministic function $f(t)$, and X_t is an Ito diffusion, then the previous formula becomes

$$\boxed{d\Big(\frac{X_t}{f(t)}\Big) = \frac{f(t)dX_t - X_t df(t)}{f(t)^2}.}$$

Example 5.2.1 *We shall show that*

$$\boxed{d(W_t^2) = 2W_t\, dW_t + dt.}$$

Applying the product rule and the fundamental relation $(dW_t)^2 = dt$, yields

$$d(W_t^2) = W_t\, dW_t + W_t\, dW_t + dW_t\, dW_t = 2W_t\, dW_t + dt.$$

Example 5.2.2 *Show that*

$$\boxed{d(W_t^3) = 3W_t^2\, dW_t + 3W_t dt.}$$

Applying the product rule and the previous exercise yields

$$\begin{aligned}
d(W_t^3) &= d(W_t \cdot W_t^2) = W_t d(W_t^2) + W_t^2\, dW_t + d(W_t^2)\, dW_t \\
&= W_t(2W_t\, dW_t + dt) + W_t^2\, dW_t + dW_t(2W_t\, dW_t + dt) \\
&= 2W_t^2\, dW_t + W_t\, dt + W_t^2\, dW_t + 2W_t(dW_t)^2 + dt\, dW_t \\
&= 3W_t^2\, dW_t + 3W_t\, dt,
\end{aligned}$$

where we used $(dW_t)^2 = dt$ and $dt\, dW_t = 0$.

Example 5.2.3 *Show that* $d(tW_t) = W_t\, dt + t\, dW_t$.

Using the product rule and $dt\, dW_t = 0$, we get

$$\begin{aligned}
d(tW_t) &= W_t\, dt + t\, dW_t + dt\, dW_t \\
&= W_t\, dt + t\, dW_t.
\end{aligned}$$

Example 5.2.4 *Let* $Z_t = \int_0^t W_u\, du$ *be the integrated Brownian motion. Show that*

$$\boxed{dZ_t = W_t\, dt.}$$

The infinitesimal change of Z_t is

$$dZ_t = Z_{t+dt} - Z_t = \int_t^{t+dt} W_s\, ds = W_t\, dt,$$

since W_s is a continuous function in s.

Example 5.2.5 *Let $A_t = \frac{1}{t} Z_t = \frac{1}{t}\int_0^t W_u\, du$ be the average of the Brownian motion on the time interval $[0, t]$. Show that*

$$\boxed{dA_t = \frac{1}{t}\Big(W_t - \frac{1}{t} Z_t\Big)\, dt.}$$

We have

$$\begin{aligned} dA_t &= d\Big(\frac{1}{t}\Big)\, Z_t + \frac{1}{t}\, dZ_t + d\Big(\frac{1}{t}\Big)\, dZ_t \\ &= \frac{-1}{t^2} Z_t\, dt + \frac{1}{t} W_t\, dt + \frac{-1}{t^2} W_t \underbrace{dt^2}_{=0} \\ &= \frac{1}{t}\Big(W_t - \frac{1}{t} Z_t\Big)\, dt. \end{aligned}$$

Exercise 5.2.1 *Let $G_t = \frac{1}{t}\int_0^t e^{W_u}\, du$ be the average of the geometric Brownian motion on $[0, t]$. Find dG_t.*

5.3 Ito's Formula

Ito's formula is the analog of the chain rule from elementary Calculus. We shall start by reviewing a few concepts regarding function approximations.

Let f be a differentiable function of a real variable x. Let x_0 be fixed and consider the changes $\Delta x = x - x_0$ and $\Delta f(x) = f(x) - f(x_0)$. It is known from Calculus that the following second order Taylor approximation holds

$$\Delta f(x) = f'(x)\Delta x + \frac{1}{2} f''(x)(\Delta x)^2 + O(\Delta x)^3.$$

When x is infinitesimally close to x_0, we replace Δx by the differential dx and obtain

$$df(x) = f'(x)dx + \frac{1}{2} f''(x)(dx)^2 + O(dx)^3. \tag{5.3.1}$$

In the elementary Calculus, all terms involving terms of equal or higher order to dx^2 are neglected; then the aforementioned formula becomes

$$df(x) = f'(x)dx.$$

Now, if we consider $x = x(t)$ to be a differentiable function of t, substituting into the previous formula we obtain the differential form of the well known chain rule

$$df\big(x(t)\big) = f'\big(x(t)\big)dx(t) = f'\big(x(t)\big)x'(t)\,dt.$$

We shall present a similar formula for the stochastic environment. In this case the deterministic function $x(t)$ is replaced by a stochastic process X_t. The composition between the differentiable function f and the process X_t is denoted by $F_t = f(X_t)$.

Neglecting the increment powers higher than or equal to $(dX_t)^3$, the expression (5.3.1) becomes

$$\boxed{dF_t = f'\big(X_t\big)dX_t + \frac{1}{2}f''\big(X_t\big)\big(dX_t\big)^2.} \tag{5.3.2}$$

In the computation of dX_t we may take into the account stochastic relations such as $dW_t^2 = dt$, or $dt\,dW_t = 0$.

5.3.1 Ito's formula for diffusions

The previous formula is a general case of *Ito's formula*. However, in most cases the increments dX_t are given by some particular relations. An important case is when the increment is given by

$$dX_t = a(W_t, t)dt + b(W_t, t)dW_t.$$

A process X_t satisfying this relation is called an *Ito diffusion*.

Theorem 5.3.1 (Ito's formula for diffusions) *If X_t is an Ito diffusion, and $F_t = f(X_t)$, then*

$$\boxed{dF_t = \Big[a(W_t, t)f'(X_t) + \frac{b(W_t, t)^2}{2}f''(X_t)\Big]dt + b(W_t, t)f'(X_t)\,dW_t.} \tag{5.3.3}$$

Proof: We shall provide a formal proof. Using relations $dW_t^2 = dt$ and $dt^2 = dW_t\,dt = 0$, we have

$$\begin{aligned}
(dX_t)^2 &= \Big(a(W_t, t)dt + b(W_t, t)dW_t\Big)^2 \\
&= a(W_t, t)^2dt^2 + 2a(W_t, t)b(W_t, t)dW_tdt + b(W_t, t)^2dW_t^2 \\
&= b(W_t, t)^2dt.
\end{aligned}$$

Substituting into (5.3.2) yields

$$\begin{aligned}
dF_t &= f'\big(X_t\big)dX_t + \frac{1}{2}f''\big(X_t\big)\big(dX_t\big)^2 \\
&= f'\big(X_t\big)\Big(a(W_t, t)dt + b(W_t, t)dW_t\Big) + \frac{1}{2}f''\big(X_t\big)b(W_t, t)^2dt \\
&= \Big[a(W_t, t)f'(X_t) + \frac{b(W_t, t)^2}{2}f''(X_t)\Big]dt + b(W_t, t)f'(X_t)\,dW_t.
\end{aligned}$$

■

In the case $X_t = W_t$ we obtain the following consequence:

Corollary 5.3.2 *Let $F_t = f(W_t)$. Then*

$$\boxed{dF_t = \frac{1}{2}f''(W_t)dt + f'(W_t)\,dW_t.} \tag{5.3.4}$$

Particular cases

1. If $f(x) = x^\alpha$, with α constant, then $f'(x) = \alpha x^{\alpha-1}$ and $f''(x) = \alpha(\alpha-1)x^{\alpha-2}$. Then (5.3.4) becomes the following useful formula

$$\boxed{d(W_t^\alpha) = \frac{1}{2}\alpha(\alpha-1)W_t^{\alpha-2}dt + \alpha W_t^{\alpha-1}\,dW_t.}$$

A couple of useful cases easily follow:

$$\begin{aligned} d(W_t^2) &= 2W_t\,dW_t + dt \\ d(W_t^3) &= 3W_t^2\,dW_t + 3W_t dt. \end{aligned}$$

2. If $f(x) = e^{kx}$, with k constant, $f'(x) = ke^{kx}$, $f''(x) = k^2e^{kx}$. Therefore

$$\boxed{d(e^{kW_t}) = ke^{kW_t}dW_t + \frac{1}{2}k^2e^{kW_t}\,dt.}$$

In particular, for $k = 1$, we obtain the increments of a geometric Brownian motion

$$d(e^{W_t}) = e^{W_t}dW_t + \frac{1}{2}e^{W_t}\,dt.$$

3. If $f(x) = \sin x$, then

$$\boxed{d(\sin W_t) = \cos W_t\,dW_t - \frac{1}{2}\sin W_t\,dt.}$$

Exercise 5.3.3 *Use the previous rules to find the following increments*

(a) $d(W_te^{W_t})$

(b) $d(3W_t^2 + 2e^{5W_t})$

(c) $d(e^{t+W_t^2})$

(d) $d\big((t+W_t)^n\big)$.

(e) $d\Big(\frac{1}{t}\int_0^t W_u\,du\Big)$

(f) $d\Big(\frac{1}{t^\alpha}\int_0^t e^{W_u}\,du\Big)$, *where α is a constant.*

In the case when the function $f = f(t,x)$ is also time dependent, the analog of (5.3.1) is given by

$$df(t,x) = \partial_t f(t,x)dt + \partial_x f(t,x)dx + \frac{1}{2}\partial_x^2 f(t,x)(dx)^2 + O(dx)^3 + O(dt)^2. \qquad (5.3.5)$$

Substituting $x = X_t$ yields

$$df(t,X_t) = \partial_t f(t,X_t)dt + \partial_x f(t,X_t)dX_t + \frac{1}{2}\partial_x^2 f(t,X_t)(dX_t)^2. \qquad (5.3.6)$$

If X_t is an Ito diffusion we obtain an extra-term in formula (5.3.3)

$$\begin{aligned} dF_t &= \Big[\partial_t f(t,X_t) + a(W_t,t)\partial_x f(t,X_t) + \frac{b(W_t,t)^2}{2}\partial_x^2 f(t,X_t)\Big]dt \\ &\quad + b(W_t,t)\partial_x f(t,X_t)\, dW_t. \qquad (5.3.7) \end{aligned}$$

Exercise 5.3.4 *Show that*

$$d(tW_t^2) = (t + W_t^2)dt + 2tW_t\, dW_t.$$

Exercise 5.3.5 *Find the following increments*

$$\begin{array}{ll} (a)\ \ d(tW_t) & \qquad (c)\ \ d(t^2\cos W_t) \\ (b)\ \ d(e^t W_t) & \qquad (d)\ \ d(\sin t\, W_t^2). \end{array}$$

5.3.2 Ito's formula for Poisson processes

Consider the process $F_t = F(M_t)$, where $M_t = N_t - \lambda t$ is the compensated Poisson process. Ito's formula for the process F_t takes the following integral form. For a proof the reader can consult Kuo [9].

Proposition 5.3.6 *Let F be a twice differentiable function. Then for any $a < t$ we have*

$$F_t = F_a + \int_a^t F'(M_{s-})\, dM_s + \sum_{a<s\le t}\Big(\Delta F(M_s) - F'(M_{s-})\Delta M_s\Big),$$

where $\Delta M_s = M_s - M_{s-}$ and $\Delta F(M_s) = F(M_s) - F(M_{s-})$.

We shall apply the aforementioned result for the case $F_t = F(M_t) = M_t^2$. We have

$$M_t^2 = M_a^2 + 2\int_a^t M_{s-}\, dM_s + \sum_{a<s\le t}\Big(M_s^2 - M_{s-}^2 - 2M_{s-}(M_s - M_{s-})\Big). \qquad (5.3.8)$$

Since the jumps in N_s are of size 1, we have $(\Delta N_s)^2 = \Delta N_s$. Since the difference of the processes M_s and N_s is continuous, then $\Delta M_s = \Delta N_s$. Using these formulas we have

$$\begin{aligned}\Big(M_s^2 - M_{s-}^2 - 2M_{s-}(M_s - M_{s-})\Big) &= (M_s - M_{s-})\Big(M_s + M_{s-} - 2M_{s-}\Big)\\ &= (M_s - M_{s-})^2 = (\Delta M_s)^2 = (\Delta N_s)^2\\ &= \Delta N_s = N_s - N_{s-}.\end{aligned}$$

Since the sum of the jumps between s and t is $\sum_{a<s\le t} \Delta N_s = N_t - N_a$, formula (5.3.8) becomes

$$M_t^2 = M_a^2 + 2\int_a^t M_{s-}\, dM_s + N_t - N_a. \tag{5.3.9}$$

The differential form is

$$d(M_t^2) = 2M_{t-}\, dM_t + dN_t,$$

which is equivalent with

$$d(M_t^2) = (1 + 2M_{t-})\, dM_t + \lambda dt,$$

since $dN_t = dM_t + \lambda dt$.

Exercise 5.3.7 *Show that*

$$\int_0^T M_{t-}\, dM_t = \frac{1}{2}(M_T^2 - N_T).$$

Exercise 5.3.8 *Use Ito's formula for the Poison process to find the conditional expectation $E[M_t^2|\mathcal{F}_s]$ for $s < t$.*

5.3.3 Ito's multidimensional formula

If the process F_t depends on several Ito diffusions, say $F_t = f(t, X_t, Y_t)$, then a similar formula to (5.3.7) leads to

$$\begin{aligned}dF_t &= \frac{\partial f}{\partial t}(t, X_t, Y_t)dt + \frac{\partial f}{\partial x}(t, X_t, Y_t)dX_t + \frac{\partial f}{\partial y}(t, X_t, Y_t)dY_t\\ &\quad + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(t, X_t, Y_t)(dX_t)^2 + \frac{1}{2}\frac{\partial^2 f}{\partial y^2}(t, X_t, Y_t)(dY_t)^2\\ &\quad + \frac{\partial^2 f}{\partial x \partial y}(t, X_t, Y_t)dX_t\, dY_t.\end{aligned}$$

Particular cases

In the case when $F_t = f(X_t, Y_t)$, with $X_t = W_t^1$, $Y_t = W_t^2$ independent Brownian

motions, we have

$$\begin{aligned} dF_t &= \frac{\partial f}{\partial x}dW_t^1 + \frac{\partial f}{\partial y}dW_t^2 + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(dW_t^1)^2 + \frac{1}{2}\frac{\partial^2 f}{\partial y^2}(dW_t^2)^2 \\ &\quad + \frac{1}{2}\frac{\partial^2 f}{\partial x \partial y}dW_t^1\, dW_t^2 \\ &= \frac{\partial f}{\partial x}dW_t^1 + \frac{\partial f}{\partial y}dW_t^2 + \frac{1}{2}\Big(\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}\Big)dt \end{aligned}$$

The expression

$$\Delta f = \frac{1}{2}\Big(\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}\Big)$$

is called the *Laplacian* of f. We can rewrite the previous formula as

$$dF_t = \frac{\partial f}{\partial x}dW_t^1 + \frac{\partial f}{\partial y}dW_t^2 + \Delta f\, dt$$

A function f with $\Delta f = 0$ is called *harmonic*. The aforementioned formula in the case of harmonic functions takes the simple form

$$\boxed{dF_t = \frac{\partial f}{\partial x}dW_t^1 + \frac{\partial f}{\partial y}dW_t^2.} \tag{5.3.10}$$

Exercise 5.3.9 *Let W_t^1, W_t^2 be two independent Brownian motions. If the function f is harmonic, show that $F_t = f(W_t^1, W_t^2)$ is a martingale. Is the converse true?*

Exercise 5.3.10 *Use the previous formulas to find dF_t in the following cases*

(*a*) $F_t = (W_t^1)^2 + (W_t^2)^2$

(*b*) $F_t = \ln[(W_t^1)^2 + (W_t^2)^2]$.

Exercise 5.3.11 *Consider the Bessel process $R_t = \sqrt{(W_t^1)^2 + (W_t^2)^2}$, where W_t^1 and W_t^2 are two independent Brownian motions. Prove that*

$$dR_t = \frac{1}{2R_t}dt + \frac{W_t^1}{R_t}dW_t^1 + \frac{W_t^2}{R_t}dW_t^2.$$

Example 5.3.1 (The product rule) *Let X_t and Y_t be two processes. Show that*

$$d(X_tY_t) = Y_tdX_t + X_tdY_t + dX_tdY_t.$$

Consider the function $f(x,y) = xy$. Since $\partial_x f = y$, $\partial_y f = x$, $\partial_x^2 f = \partial_y^2 f = 0$, $\partial_x\partial_y = 1$, then Ito's multidimensional formula yields

$$\begin{aligned} d(X_tY_t) &= d\big(f(X_,Y_t)\big) = \partial_x f\, dX_t + \partial_y f\, dY_t \\ &\quad + \frac{1}{2}\partial_x^2 f(dX_t)^2 + \frac{1}{2}\partial_y^2 f(dY_t)^2 + \partial_x\partial_y f\, dX_tdY_t \\ &= Y_tdX_t + X_tdY_t + dX_tdY_t. \end{aligned}$$

Example 5.3.2 (The quotient rule) *Let X_t and Y_t be two processes. Show that*

$$d\Big(\frac{X_t}{Y_t}\Big) = \frac{Y_t dX_t - X_t dY_t - dX_t dY_t}{Y_t^2} + \frac{X_t}{Y_t^2}(dY_t)^2.$$

Consider the function $f(x, y) = \frac{x}{y}$. Since $\partial_x f = \frac{1}{y}$, $\partial_y f = -\frac{x}{y^2}$, $\partial_x^2 f = 0$, $\partial_y^2 f = -\frac{x}{y^2}$, $\partial_x \partial_y = \frac{1}{y^2}$, then applying Ito's multidimensional formula yields

$$\begin{aligned}
d\Big(\frac{X_t}{Y_t}\Big) &= d\big(f(X,Y_t)\big) = \partial_x f\, dX_t + \partial_y f\, dY_t \\
&\quad + \frac{1}{2}\partial_x^2 f (dX_t)^2 + \frac{1}{2}\partial_y^2 f (dY_t)^2 + \partial_x \partial_y f\, dX_t dY_t \\
&= \frac{1}{Y_t} dX_t - \frac{X_t}{Y_t^2} dY_t - \frac{1}{Y_t^2} dX_t dY_t \\
&= \frac{Y_t dX_t - X_t dY_t - dX_t dY_t}{Y_t^2} + \frac{X_t}{Y_t^2}(dY_t)^2.
\end{aligned}$$

Chapter 6

Stochastic Integration Techniques

Computing a stochastic integral starting from the definition of the Ito integral is a quite inefficient method. Like in the elementary Calculus, several methods can be developed to compute stochastic integrals. In order to keep the analogy with the elementary Calculus, we have called them *Fundamental Theorem of Stochastic Calculus* and *Integration by Parts*. The integration by substitution is more complicated in the stochastic environment and we have considered only a particular case of it, which we called *The method of heat equation.*

6.1 Fundamental Theorem of Stochastic Calculus

Consider a process X_t whose increments satisfy the equation $dX_t = f(t, W_t)dW_t$. Integrating formally between a and t yields

$$\int_a^t dX_s = \int_a^t f(s, W_s)dW_s. \tag{6.1.1}$$

The integral on the left side can be computed as in the following. If we consider the partition $0 = t_0 < t_1 < \cdots < t_{n-1} < t_n = t$, then

$$\int_a^t dX_s = \operatorname*{ms-lim}_{n\to\infty} \sum_{j=0}^{n-1} (X_{t_{j+1}} - X_{t_j}) = X_t - X_a,$$

since we canceled the terms in pairs. Substituting into formula (6.1.1) yields $X_t = X_a + \int_a^t f(s, W_s)dW_s$, and hence $dX_t = d\Big(\int_a^t f(s, W_s)dW_s\Big)$, since X_a is a constant.

Theorem 6.1.1 (The Fundamental Theorem of Stochastic Calculus)
(i) For any $a < t$, we have

$$\boxed{d\Big(\int_a^t f(s, W_s)dW_s\Big) = f(t, W_t)dW_t.}$$

(ii) If Y_t is a stochastic process, such that $Y_t dW_t = dF_t$, then

$$\boxed{\int_a^b Y_t\, dW_t = F_b - F_a.}$$

We shall provide a few applications of the aforementioned theorem.

Example 6.1.1 *Verify the stochastic formula*

$$\int_0^t W_s\, dW_s = \frac{W_t^2}{2} - \frac{t}{2}.$$

Let $X_t = \int_0^t W_s\, dW_s$ and $Y_t = \dfrac{W_t^2}{2} - \dfrac{t}{2}$. From Ito's formula

$$dY_t = d\Big(\frac{W_t^2}{2}\Big) - d\Big(\frac{t}{2}\Big) = \frac{1}{2}(2W_t\, dW_t + dt) - \frac{1}{2}dt = W_t\, dW_t,$$

and from the Fundamental Theorem of Stochastic Calculus

$$dX_t = d\Big(\int_0^t W_s\, dW_s\Big) = W_t\, dW_t.$$

Hence $dX_t = dY_t$, or $d(X_t - Y_t) = 0$. Since the process $X_t - Y_t$ has zero increments, then $X_t - Y_t = c$, constant. Taking $t = 0$, yields

$$c = X_0 - Y_0 = \int_0^0 W_s\, dW_s - \Big(\frac{W_0^2}{2} - \frac{0}{2}\Big) = 0,$$

and hence $c = 0$. It follows that $X_t = Y_t$, which verifies the desired relation.

Example 6.1.2 *Verify the formula*

$$\int_0^t sW_s\, dW_s = \frac{t}{2}\Big(W_t^2 - \frac{t}{2}\Big) - \frac{1}{2}\int_0^t W_s^2\, ds.$$

Consider the stochastic processes $X_t = \int_0^t sW_s\, dW_s$, $Y_t = \dfrac{t}{2}\Big(W_t^2 - 1\Big)$, and $Z_t = \frac{1}{2}\int_0^t W_s^2\, ds$. The Fundamental Theorem yields

$$\begin{aligned} dX_t &= tW_t\, dW_t \\ dZ_t &= \frac{1}{2}W_t^2 dt. \end{aligned}$$

Applying Ito's formula, we get

$$\begin{aligned} dY_t &= d\Big(\frac{t}{2}\Big(W_t^2 - \frac{t}{2}\Big)\Big) = \frac{1}{2}d(tW_t^2) - d\Big(\frac{t^2}{4}\Big) \\ &= \frac{1}{2}\Big[(t + W_t^2)dt + 2tW_t\, dW_t\Big] - \frac{1}{2}tdt \\ &= \frac{1}{2}W_t^2 dt + tW_t\, dW_t. \end{aligned}$$

We can easily see that

$$dX_t = dY_t - dZ_t.$$

This implies $d(X_t - Y_t + Z_t) = 0$, i.e. $X_t - Y_t + Z_t = c$, constant. Since $X_0 = Y_0 = Z_0 = 0$, it follows that $c = 0$. This proves the desired relation.

Example 6.1.3 *Show that*

$$\int_0^t (W_s^2 - s)\, dW_s = \frac{1}{3}W_t^3 - tW_t.$$

Consider the function $f(t,x) = \frac{1}{3}x^3 - tx$, and let $F_t = f(t, W_t)$. Since $\partial_t f = -x$, $\partial_x f = x^2 - t$, and $\partial_x^2 f = 2x$, then Ito's formula provides

$$\begin{aligned} dF_t &= \partial_t f\, dt + \partial_x f\, dW_t + \frac{1}{2}\partial_x^2 f\, (dW_t)^2 \\ &= -W_t dt + (W_t^2 - t)\, dW_t + \frac{1}{2} 2W_t\, dt \\ &= (W_t^2 - t) dW_t. \end{aligned}$$

From the Fundamental Theorem we get

$$\int_0^t (W_s^2 - s)\, dW_s = \int_0^t dF_s = F_t - F_0 = F_t = \frac{1}{3}W_t^3 - tW_t.$$

6.2 Stochastic Integration by Parts

Consider the process $F_t = f(t)g(W_t)$, with f and g differentiable. Using the product rule yields

$$\begin{aligned} dF_t &= df(t)\, g(W_t) + f(t)\, dg(W_t) \\ &= f'(t)g(W_t)dt + f(t)\big(g'(W_t)dW_t + \frac{1}{2}g''(W_t)dt\big) \\ &= f'(t)g(W_t)dt + \frac{1}{2}f(t)g''(W_t)dt + f(t)g'(W_t)dW_t. \end{aligned}$$

Writing the relation in the integral form, we obtain the first integration by parts formula:

$$\boxed{\int_a^b f(t)g'(W_t)\, dW_t = f(t)g(W_t)\Big|_a^b - \int_a^b f'(t)g(W_t)\, dt - \frac{1}{2}\int_a^b f(t)g''(W_t)\, dt.}$$

This formula is to be used when integrating a product between a function of t and a function of the Brownian motion W_t, for which an antiderivative is known. The following two particular cases are important and useful in applications.

1. If $g(W_t) = W_t$, the aforementioned formula takes the simple form

$$\boxed{\int_a^b f(t)\, dW_t = f(t)W_t\Big|_{t=a}^{t=b} - \int_a^b f'(t)W_t\, dt.} \qquad (6.2.2)$$

It is worth noting that the left side is a Wiener integral.

2. If $f(t) = 1$, then the formula becomes

$$\boxed{\int_a^b g'(W_t)\, dW_t = g(W_t)\Big|_{t=a}^{t=b} - \frac{1}{2}\int_a^b g''(W_t)\, dt.} \qquad (6.2.3)$$

Application 1 Consider the Wiener integral $I_T = \displaystyle\int_0^T t\, dW_t$. From the general theory, see Proposition 4.6.1, it is known that I is a random variable normally distributed with mean 0 and variance

$$Var[I_T] = \int_0^T t^2\, dt = \frac{T^3}{3}.$$

Recall the definition of integrated Brownian motion

$$Z_t = \int_0^t W_u\, du.$$

Formula (6.2.2) yields a relationship between I and the integrated Brownian motion

$$I_T = \int_0^T t\, dW_t = TW_T - \int_0^T W_t\, dt = TW_T - Z_T,$$

and hence $I_T + Z_T = TW_T$. This relation can be used to compute the covariance between I_T and Z_T.

$$\begin{aligned} Cov(I_T + Z_T, I_T + Z_T) &= Var[TW_T] \Longleftrightarrow \\ Var[I_T] + Var[Z_T] + 2Cov(I_T, Z_T) &= T^2 Var[W_T] \Longleftrightarrow \\ T^3/3 + T^3/3 + 2Cov(I_T, Z_T) &= T^3 \Longleftrightarrow \\ Cov(I_T, Z_T) &= T^3/6, \end{aligned}$$

where we used that $Var[Z_T] = T^3/3$. The processes I_t and Z_t are not independent. Their correlation coefficient is 0.5 as the following calculation shows

$$\begin{aligned} Corr(I_T, Z_T) &= \frac{Cov(I_T, Z_T)}{\Big(Var[I_T]Var[Z_T]\Big)^{1/2}} = \frac{T^3/6}{T^3/3} \\ &= 1/2. \end{aligned}$$

Application 2 If we let $g(x) = \frac{x^2}{2}$ in formula (6.2.3), we get

$$\boxed{\int_a^b W_t\, dW_t = \frac{W_b^2 - W_a^2}{2} - \frac{1}{2}(b-a).}$$

It is worth noting that letting $a = 0$ and $b = T$, we retrieve a formula that was proved by direct methods in chapter 2

$$\boxed{\int_0^T W_t\, dW_t = \frac{W_T^2}{2} - \frac{T}{2}.}$$

Similarly, if we let $g(x) = \frac{x^3}{3}$ in (6.2.3) yields

$$\boxed{\int_a^b W_t^2\, dW_t = \frac{W_t^3}{3}\Big|_a^b - \int_a^b W_t\, dt.}$$

Application 3
Choosing $f(t) = e^{\alpha t}$ and $g(x) = \cos x$, we shall compute the stochastic integral $\int_0^T e^{\alpha t} \cos W_t\, dW_t$ using the formula of integration by parts

$$\begin{aligned}
\int_0^T e^{\alpha t} \cos W_t\, dW_t &= \int_0^T e^{\alpha t} (\sin W_t)'\, dW_t \\
&= e^{\alpha t} \sin W_t \Big|_0^T - \int_0^T (e^{\alpha t})' \sin W_t\, dt - \frac{1}{2}\int_0^T e^{\alpha t} (\cos W_t)''\, dt \\
&= e^{\alpha T} \sin W_T - \alpha \int_0^T e^{\alpha t} \sin W_t\, dt + \frac{1}{2}\int_0^T e^{\alpha t} \sin W_t\, dt \\
&= e^{\alpha T} \sin W_T - \Big(\alpha - \frac{1}{2}\Big) \int_0^T e^{\alpha t} \sin W_t\, dt.
\end{aligned}$$

The particular case $\alpha = \frac{1}{2}$ leads to the following exact formula of a stochastic integral

$$\boxed{\int_0^T e^{\frac{t}{2}} \cos W_t\, dW_t = e^{\frac{T}{2}} \sin W_T.} \tag{6.2.4}$$

In a similar way, we can obtain an exact formula for the stochastic integral $\int_0^T e^{\beta t} \sin W_t\, dW_t$ as follows

$$\begin{aligned}
\int_0^T e^{\beta t} \sin W_t\, dW_t &= -\int_0^T e^{\beta t} (\cos W_t)'\, dW_t \\
&= -e^{\beta t} \cos W_t \Big|_0^T + \beta \int_0^T e^{\beta t} \cos W_t\, dt - \frac{1}{2}\int_0^T e^{\beta t} \cos W_t\, dt.
\end{aligned}$$

Taking $\beta = \frac{1}{2}$ yields the closed form formula

$$\boxed{\int_0^T e^{\frac{t}{2}} \sin W_t \, dW_t = 1 - e^{\frac{T}{2}} \cos W_T.} \tag{6.2.5}$$

A consequence of the last two formulas and of Euler's formula

$$e^{iW_t} = \cos W_t + i \sin W_t,$$

is

$$\boxed{\int_0^T e^{\frac{t}{2}+iW_t} \, dW_t = i(1 - e^{\frac{T}{2}+iW_T}).}$$

The proof details are left to the reader.

A general form of the integration by parts formula

In general, if X_t and Y_t are two Ito diffusions, from the product formula

$$d(X_tY_t) = X_t dY_t + Y_t dX_t + dX_t \, dY_t.$$

Integrating between the limits a and b

$$\int_a^b d(X_tY_t) = \int_a^b X_t dY_t + \int_a^b Y_t dX_t + \int_a^b dX_t \, dY_t.$$

From the Fundamental Theorem

$$\int_a^b d(X_tY_t) = X_bY_b - X_aY_a,$$

so the previous formula takes the following form of integration by parts

$$\boxed{\int_a^b X_t dY_t = X_bY_b - X_aY_a - \int_a^b Y_t dX_t - \int_a^b dX_t \, dY_t.}$$

This formula is of theoretical value. In practice, the term $dX_t \, dY_t$ needs to be computed using the rules $W_t^2 = dt$, and $dt \, dW_t = 0$.

Exercise 6.2.1 (a) *Use integration by parts to get*

$$\int_0^T \frac{1}{1+W_t^2} dW_t = \tan^{-1}(W_T) + \int_0^T \frac{W_t}{(1+W_t^2)^2} \, dt, \quad T > 0.$$

(b) *Show that*

$$E[\tan^{-1}(W_T)] = -\int_0^T E\Big[\frac{W_t}{(1+W_t^2)^2}\Big] \, dt.$$

(*c*) *Prove the double inequality*

$$-\frac{3\sqrt{3}}{16} \le \frac{x}{(1+x^2)^2} \le \frac{3\sqrt{3}}{16}, \qquad \forall x \in \mathbb{R}.$$

(*d*) *Use part* (*c*) *to obtain*

$$-\frac{3\sqrt{3}}{16}T \le \int_0^T \frac{W_t}{(1+W_t^2)^2}\,dt \le \frac{3\sqrt{3}}{16}T.$$

(*e*) *Use part* (*d*) *to get*

$$-\frac{3\sqrt{3}}{16}T \le E[\tan^{-1}(W_T)] \le \frac{3\sqrt{3}}{16}T.$$

(*f*) *Does part* (*e*) *contradict the inequality*

$$-\frac{\pi}{2} < \tan^{-1}(W_T) < \frac{\pi}{2}?$$

Exercise 6.2.2 (*a*) *Show the relation*

$$\int_0^T e^{W_t}\,dW_t = e^{W_T} - 1 - \frac{1}{2}\int_0^T e^{W_t}\,dt.$$

(*b*) *Use part* (*a*) *to find* $E[e^{W_t}]$.

Exercise 6.2.3 (*a*) *Use integration by parts to show*

$$\int_0^T W_t e^{W_t}\,dW_t = 1 + W_T e^{W_T} - e^{W_T} - \frac{1}{2}\int_0^T e^{W_t}(1+W_t)\,dt;$$

(*b*) *Use part* (*a*) *to find* $E[W_t e^{W_t}]$;

(*c*) *Show that* $Cov(W_t, e^{W_t}) = te^{t/2}$;

(*d*) *Prove that* $Corr(W_t, e^{W_t}) = \sqrt{\dfrac{t}{e^t - 1}}$, *and compute the limits as* $t \to 0$ *and* $t \to \infty$.

Exercise 6.2.4 (*a*) *Let* $T > 0$. *Show the following relation using integration by parts*

$$\int_0^T \frac{2W_t}{1+W_t^2}\,dW_t = \ln(1+W_T^2) - \int_0^T \frac{1-W_t^2}{(1+W_t^2)^2}\,dt.$$

(*b*) *Show that for any real number* x *the following double inequality holds*

$$-\frac{1}{8} \le \frac{1-x^2}{(1+x^2)^2} \le 1.$$

(c) Use part (b) to show that

$$-\frac{1}{8}T \le \int_0^T \frac{1-W_t^2}{(1+W_t^2)^2}\,dt \le T.$$

(d) Use parts (a) and (c) to get

$$-\frac{T}{8} \le E[\ln(1+W_T^2)] \le T.$$

(e) Use Jensen's inequality to get

$$E[\ln(1+W_T^2)] \le \ln(1+T).$$

Does this contradict the upper bound provided in (d)?

6.3 The Heat Equation Method

In elementary Calculus, integration by substitution is the inverse application of the chain rule. In the stochastic environment, this will be the inverse application of Ito's formula. This is difficult to apply in general, but there is a particular case of great importance.

Let $\varphi(t,x)$ be a solution of the equation

$$\partial_t\varphi + \frac{1}{2}\partial_x^2\varphi = 0. \tag{6.3.6}$$

This is called the *heat equation without sources.* The non-homogeneous equation

$$\partial_t\varphi + \frac{1}{2}\partial_x^2\varphi = G(t,x) \tag{6.3.7}$$

is called *heat equation with sources.* The function $G(t,x)$ represents the density of heat sources, while the function $\varphi(t,x)$ is the temperature at point x at time t in a one-dimensional wire. If the heat source is time independent, then $G = G(x)$, i.e. G is a function of x only.

Example 6.3.1 *Find all solutions of the equation (6.3.6) of type $\varphi(t,x) = a(t) + b(x)$.*

Substituting into equation (6.3.6) yields

$$\frac{1}{2}b''(x) = -a'(t).$$

Since the left side is a function of x only, while the right side is a function of variable t, the only where the previous equation is satisfied is when both sides are equal to the

same constant C. This is called a separation constant. Therefore $a(t)$ and $b(x)$ satisfy the equations

$$a'(t) = -C, \qquad \frac{1}{2}b''(x) = C.$$

Integrating yields $a(t) = -Ct + C_0$ and $b(x) = Cx^2 + C_1x + C_2$. It follows that

$$\varphi(t,x) = C(x^2 - t) + C_1x + C_3,$$

with C_0, C_1, C_2, C_3 arbitrary constants.

Example 6.3.2 *Find all solutions of the equation (6.3.6) of the type $\varphi(t,x) = a(t)b(x)$.*

Substituting into the equation and dividing by $a(t)b(x)$ yields

$$\frac{a'(t)}{a(t)} + \frac{1}{2}\frac{b''(x)}{b(x)} = 0.$$

There is a separation constant C such that $\dfrac{a'(t)}{a(t)} = -C$ and $\dfrac{b''(x)}{b(x)} = 2C$. There are three distinct cases to discuss:

1. $C = 0$. In this case $a(t) = a_0$ and $b(x) = b_1x + b_0$, with a_0, a_1, b_0, b_1 real constants. Then

$$\varphi(t,x) = a(t)b(x) = c_1x + c_0, \qquad c_0, c_1 \in \mathbb{R}$$

is just a linear function in x.

2. $C > 0$. Let $\lambda > 0$ such that $2C = \lambda^2$. Then $a'(t) = -\frac{\lambda^2}{2}a(t)$ and $b''(x) = \lambda^2 b(x)$, with solutions

$$\begin{aligned} a(t) &= a_0 e^{-\lambda^2 t/2} \\ b(x) &= c_1 e^{\lambda x} + c_2 e^{-\lambda x}. \end{aligned}$$

The general solution of (6.3.6) is

$$\varphi(t,x) = e^{-\lambda^2 t/2}(c_1 e^{\lambda x} + c_2 e^{-\lambda x}), \quad c_1, c_2 \in \mathbb{R}.$$

3. $C < 0$. Let $\lambda > 0$ such that $2C = -\lambda^2$. Then $a'(t) = \frac{\lambda^2}{2}a(t)$ and $b''(x) = -\lambda^2 b(x)$. Solving yields

$$\begin{aligned} a(t) &= a_0 e^{\lambda^2 t/2} \\ b(x) &= c_1 \sin(\lambda x) + c_2 \cos(\lambda x). \end{aligned}$$

The general solution of (6.3.6) in this case is

$$\varphi(t,x) = e^{\lambda^2 t/2}\big(c_1 \sin(\lambda x) + c_2 \cos(\lambda x)\big), \quad c_1, c_2 \in \mathbb{R}.$$

In particular, the functions x, $x^2 - t$, $e^{x-t/2}$, $e^{-x-t/2}$, $e^{t/2}\sin x$ and $e^{t/2}\cos x$, or any linear combination of them are solutions of the heat equation (6.3.6). However, there are other solutions which are not of the previous type.

Exercise 6.3.1 *Prove that* $\varphi(t,x) = \frac{1}{3}x^3 - tx$ *is a solution of the heat equation (6.3.6).*

Exercise 6.3.2 *Show that* $\varphi(t,x) = t^{-1/2}e^{-x^2/(2t)}$ *is a solution of the heat equation (6.3.6) for* $t > 0$.

Exercise 6.3.3 *Let* $\varphi = u(\lambda)$, *with* $\lambda = \dfrac{x}{2\sqrt{t}}$, $t > 0$. *Show that* φ *satisfies the heat equation (6.3.6) if and only if* $u'' + 2\lambda u' = 0$.

Exercise 6.3.4 *Let* $erfc(x) = \dfrac{2}{\sqrt{\pi}}\displaystyle\int_x^{\infty} e^{-r^2}\,dr$. *Show that* $\varphi = erfc(x/(2\sqrt{t}))$ *is a solution of the equation (6.3.6).*

Exercise 6.3.5 (the fundamental solution) *Show that* $\varphi(t,x) = \frac{1}{\sqrt{4\pi t}}e^{-\frac{x^2}{4t}}$, $t > 0$ *satisfies the equation (6.3.6).*

Sometimes it is useful to generate new solutions for the heat equation from other solutions. Below we present a few ways to accomplish this:

(i) by linear combination: if φ_1 and φ_2 are solutions, then $a_1\varphi_1 + a_1\varphi_2$ is a solution, where a_1, a_2 constants.

(ii) by translation: if $\varphi(t,x)$ is a solution, then $\varphi(t-\tau, x-\xi)$ is a solution, where (τ, ξ) is a translation vector.

(iii) by affine transforms: if $\varphi(t,x)$ is a solution, then $\varphi(\lambda\tau, \lambda^2 x)$ is a solution, for any constant λ.

(iv) by differentiation: if $\varphi(t,x)$ is a solution, then $\dfrac{\partial^{n+m}}{\partial^n x \partial^m t}\varphi(t,x)$ is a solution.

(v) by convolution: if $\varphi(t,x)$ is a solution, then so are

$$\int_a^b \varphi(t, x-\xi) f(\xi)\,d\xi$$
$$\int_a^b \varphi(t-\tau, x) g(\tau)\,dt.$$

For more detail on the subject the reader can consult Widder [17] and Cannon [4].

Theorem 6.3.6 *Let* $\varphi(t,x)$ *be a solution of the heat equation (6.3.6) and denote* $f(t,x) = \partial_x\varphi(t,x)$. *Then*

$$\boxed{\int_a^b f(t, W_t)\,dW_t = \varphi(b, W_b) - \varphi(a, W_a).}$$

Proof: Let $F_t = \varphi(t, W_t)$. Applying Ito's formula we get

$$dF_t = \partial_x \varphi(t, W_t)\, dW_t + \Big(\partial_t \varphi + \frac{1}{2}\partial_x^2 \varphi\Big) dt.$$

Since $\partial_t \varphi + \frac{1}{2}\partial_x^2 \varphi = 0$ and $\partial_x \varphi(t, W_t) = f(t, W_t)$, we have

$$dF_t = f(t, W_t)\, dW_t.$$

Applying the Fundamental Theorem yields

$$\int_a^b f(t, W_t)\, dW_t = \int_a^b dF_t = F_b - F_a = \varphi(b, W_b) - \varphi(a, W_a).$$

∎

Application 6.3.7 *Show that*

$$\int_0^T W_t\, dW_t = \frac{1}{2}W_T^2 - \frac{1}{2}T.$$

Choose the solution of the heat equation (6.3.6) given by $\varphi(t, x) = x^2 - t$. Then $f(t, x) = \partial_x \varphi(t, x) = 2x$. Theorem 6.3.6 yields

$$\int_0^T 2W_t\, dW_t = \int_0^T f(t, W_t)\, dW_t = \varphi(t, x)\Big|_0^T = W_T^2 - T.$$

Dividing by 2 leads to the desired result.

Application 6.3.8 *Show that*

$$\int_0^T (W_t^2 - t)\, dW_t = \frac{1}{3}W_T^3 - TW_T.$$

Consider the function $\varphi(t, x) = \frac{1}{3}x^3 - tx$, which is a solution of the heat equation (6.3.6), see Exercise 6.3.1. Then $f(t, x) = \partial_x \varphi(t, x) = x^2 - t$. Applying Theorem 6.3.6 yields

$$\int_0^T (W_t^2 - t)\, dW_t = \int_0^T f(t, W_t)\, dW_t = \varphi(t, W_t)\Big|_0^T = \frac{1}{3}W_T^3 - TW_T.$$

Application 6.3.9 *Let $\lambda > 0$. Prove the identities*

$$\boxed{\int_0^T e^{-\frac{\lambda^2 t}{2} \pm \lambda W_t}\, dW_t = \frac{1}{\pm\lambda}\Big(e^{-\frac{\lambda^2 T}{2} \pm \lambda W_T} - 1\Big).}$$

Consider the function $\varphi(t,x) = e^{-\frac{\lambda^2 t}{2} \pm \lambda x}$, which is a solution of the homogeneous heat equation (6.3.6), see Example 6.3.2. Then $f(t,x) = \partial_x \varphi(t,x) = \pm\lambda e^{-\frac{\lambda^2 t}{2} \pm \lambda x}$. Apply Theorem 6.3.6 to get

$$\int_0^T \pm\lambda e^{-\frac{\lambda^2 t}{2} \pm \lambda x}\, dW_t = \int_0^T f(t,W_t)\, dW_t = \varphi(t,W_t)\Big|_0^T = e^{-\frac{\lambda^2 T}{2} \pm \lambda W_T} - 1.$$

Dividing by the constant $\pm\lambda$ ends the proof.

In particular, for $\lambda = 1$ the aforementioned formula becomes

$$\boxed{\int_0^T e^{-\frac{t}{2} + W_t}\, dW_t = e^{-\frac{T}{2} + W_T} - 1.} \tag{6.3.8}$$

Application 6.3.10 *Let $\lambda > 0$. Prove the identity*

$$\boxed{\int_0^T e^{\frac{\lambda^2 t}{2}} \cos(\lambda W_t)\, dW_t = \frac{1}{\lambda} e^{\frac{\lambda^2 T}{2}} \sin(\lambda W_T).}$$

From the Example 6.3.2 we know that $\varphi(t,x) = e^{\frac{\lambda^2 t}{2}} \sin(\lambda x)$ is a solution of the heat equation. Applying Theorem 6.3.6 to the function $f(t,x) = \partial_x \varphi(t,x) = \lambda e^{\frac{\lambda^2 t}{2}} \cos(\lambda x)$, yields

$$\begin{aligned}
\int_0^T \lambda e^{\frac{\lambda^2 t}{2}} \cos(\lambda W_t)\, dW_t &= \int_0^T f(t,W_t)\, dW_t = \varphi(t,W_t)\Big|_0^T \\
&= e^{\frac{\lambda^2 t}{2}} \sin(\lambda W_t)\Big|_0^T = e^{\frac{\lambda^2 T}{2}} \sin(\lambda W_T).
\end{aligned}$$

Divide by λ to end the proof.

If we choose $\lambda = 1$ we recover a result already familiar to the reader from section 6.2

$$\boxed{\int_0^T e^{\frac{t}{2}} \cos(W_t)\, dW_t = e^{\frac{T}{2}} \sin W_T.} \tag{6.3.9}$$

Application 6.3.11 *Let $\lambda > 0$. Show that*

$$\boxed{\int_0^T e^{\frac{\lambda^2 t}{2}} \sin(\lambda W_t)\, dW_t = \frac{1}{\lambda}\Big(1 - e^{\frac{\lambda^2 T}{2}} \cos(\lambda W_T)\Big).}$$

Choose $\varphi(t,x) = e^{\frac{\lambda^2 t}{2}} \cos(\lambda x)$ to be a solution of the heat equation. Apply Theorem 6.3.6 for the function $f(t,x) = \partial_x \varphi(t,x) = -\lambda e^{\frac{\lambda^2 t}{2}} \sin(\lambda x)$ to get

$$\begin{aligned}
\int_0^T (-\lambda) e^{\frac{\lambda^2 t}{2}} \sin(\lambda W_t)\, dW_t &= \varphi(t,W_t)\Big|_0^T \\
&= e^{\frac{\lambda^2 T}{2}} \cos(\lambda W_t)\Big|_0^T = e^{\frac{\lambda^2 T}{2}} \cos(\lambda W_T) - 1,
\end{aligned}$$

and then divide by $-\lambda$.

Application 6.3.12 *Let $0 < a < b$. Show that*

$$\boxed{\int_a^b t^{-\frac{3}{2}} W_t e^{-\frac{W_t^2}{2t}}\,dW_t = a^{-\frac{1}{2}} e^{-\frac{W_a^2}{2a}} - b^{-\frac{1}{2}} e^{-\frac{W_b^2}{2b}}.} \qquad (6.3.10)$$

From Exercise 6.3.2 we have that $\varphi(t,x) = t^{-1/2}e^{-x^2/(2t)}$ is a solution of the homogeneous heat equation. Since $f(t,x) = \partial_x \varphi(t,x) = -t^{-3/2}xe^{-x^2/(2t)}$, applying Theorem 6.3.6 yields to the desired result. The reader can easily fill in the details.

Integration techniques will be used when solving stochastic differential equations in the next chapter.

Exercise 6.3.13 *Find the value of the following stochastic integrals*

(a) $\displaystyle\int_0^1 e^t \cos(\sqrt{2}W_t)\,dW_t$

(b) $\displaystyle\int_0^3 e^{2t} \cos(2W_t)\,dW_t$

(c) $\displaystyle\int_0^4 e^{-t+\sqrt{2}W_t}\,dW_t.$

Exercise 6.3.14 *Let $\varphi(t,x)$ be a solution of the following non-homogeneous heat equation with time-dependent and uniform heat source $G(t)$*

$$\partial_t \varphi + \frac{1}{2}\partial_x^2 \varphi = G(t).$$

Denote $f(t,x) = \partial_x \varphi(t,x)$. Show that

$$\boxed{\int_a^b f(t,W_t)\,dW_t = \varphi(b,W_b) - \varphi(a,W_a) - \int_a^b G(t)\,dt.}$$

How does the formula change if the heat source G is constant?

6.4 Table of Usual Stochastic Integrals

Now we present a user-friendly table, which enlists identities that are among the most important of our standard techniques. This table is far too complicated to be memorized in full. However, the first couple of identities in this table are the most memorable, and should be remembered.

Let $a < b$ and $0 < T$. Then we have:

1. $\displaystyle\int_a^b dW_t = W_b - W_a;$

2. $\displaystyle\int_0^T W_t\,dW_t = \frac{W_T^2}{2} - \frac{T}{2};$

3. $\displaystyle\int_0^T (W_t^2 - t)\,dW_t = \frac{W_T^2}{3} - TW_T;$

4. $\displaystyle\int_0^T t\,dW_t = TW_T - \int_0^T W_t\,dt, \quad 0 < T;$

5. $\displaystyle\int_0^T W_t^2\,dW_t = \frac{W_T^3}{3} - \int_0^T W_t\,dt;$

6. $\displaystyle\int_0^T e^{\frac{t}{2}} \cos W_t\,dW_t = e^{\frac{T}{2}} \sin W_T;$

7. $\displaystyle\int_0^T e^{\frac{t}{2}} \sin W_t\,dW_t = 1 - e^{\frac{T}{2}} \cos W_T;$

8. $\displaystyle\int_0^T e^{-\frac{t}{2}+W_t} = e^{-\frac{T}{2}+W_T} - 1;$

9. $\displaystyle\int_0^T e^{\frac{\lambda^2 t}{2}} \cos(\lambda W_t)\,dW_t = \frac{1}{\lambda} e^{\frac{\lambda^2 T}{2}} \sin(\lambda W_T);$

10. $\displaystyle\int_0^T e^{\frac{\lambda^2 t}{2}} \sin(\lambda W_t)\,dW_t = \frac{1}{\lambda}\Big(1 - e^{\frac{\lambda^2 T}{2}} \cos(\lambda W_T)\Big);$

11. $\displaystyle\int_0^T e^{-\frac{\lambda^2 t}{2} \pm \lambda W_t}\,dW_t = \frac{1}{\pm\lambda}\Big(e^{-\frac{\lambda^2 T}{2} \pm \lambda W_T} - 1\Big);$

12. $\displaystyle\int_a^b t^{-\frac{3}{2}} W_t e^{-\frac{W_t^2}{2t}}\,dW_t = a^{-\frac{1}{2}} e^{-\frac{W_a^2}{2a}} - b^{-\frac{1}{2}} e^{-\frac{W_b^2}{2b}};$

13. $\displaystyle d\Big(\int_a^t f(s, W_s)\,dW_s\Big) = f(t, W_t)\,dW_t;$

14. $\displaystyle\int_a^b Y_t\,dW_t = F_b - F_a$, if $Y_t dW_t = dF_t$;

15. $\displaystyle\int_a^b f(t)\,dW_t = f(t)W_t\big|_a^b - \int_a^b f'(t) W_t\,dt;$

16. $\displaystyle\int_a^b g'(W_t)\,dW_t = g(W_t)\Big|_a^b - \frac{1}{2}\int_a^b g''(W_t)\,dt.$

Chapter 7

Stochastic Differential Equations

7.1 Definitions and Examples

Let X_t be a continuous stochastic process. If small changes in the process X_t can be written as a linear combination of small changes in t and small increments of the Brownian motion W_t, we may write

$$\boxed{dX_t = a(t, W_t, X_t)dt + b(t, W_t, X_t)\, dW_t} \tag{7.1.1}$$

and call it a *stochastic differential equation*. In fact, this differential relation has the following integral meaning:

$$\boxed{X_t = X_0 + \int_0^t a(s, W_s, X_s)\, ds + \int_0^t b(s, W_s, X_s)\, dW_s,} \tag{7.1.2}$$

where the last integral is taken in the Ito sense. Relation (7.1.2) is taken as the definition for the stochastic differential equation (7.1.1). However, since it is convenient to use stochastic differentials informally, we shall approach stochastic differential equations by analogy with the ordinary differential equations, and try to present the same methods of solving equations in the new stochastic environment.

The functions $a(t, W_t, X_t)$ and $b(t, W_t, X_t)$ are called *drift rate* and *volatility*, respectively. A process X_t is called a (strong) *solution* for the stochastic equation (7.1.1) if it satisfies the equation. We shall start with an example.

Example 7.1.1 (The Brownian bridge) *Let $a, b \in \mathbb{R}$. Show that the process*

$$X_t = a(1-t) + bt + (1-t)\int_0^t \frac{1}{1-s}\, dW_s, \quad 0 \le t < 1$$

is a solution of the stochastic differential equation

$$dX_t = \frac{b - X_t}{1-t}dt + dW_t, \qquad 0 \le t < 1, X_0 = a.$$

We shall perform a routine verification to show that X_t is a solution. First we compute the quotient $\dfrac{b-X_t}{1-t}$:

$$\begin{aligned} b - X_t &= b - a(1-t) - bt - (1-t)\int_0^t \frac{1}{1-s}\,dW_s \\ &= (b-a)(1-t) - (1-t)\int_0^t \frac{1}{1-s}\,dW_s, \end{aligned}$$

and dividing by $1-t$ yields

$$\frac{b-X_t}{1-t} = b - a - \int_0^t \frac{1}{1-s}\,dW_s. \tag{7.1.3}$$

Using

$$d\Big(\int_0^t \frac{1}{1-s}\,dW_s\Big) = \frac{1}{1-t}\,dW_t,$$

the product rule yields

$$\begin{aligned} dX_t &= a\,d(1-t) + bdt + d(1-t)\int_0^t \frac{1}{1-s}\,dW_s + (1-t)d\Big(\int_0^t \frac{1}{1-s}\,dW_s\Big) \\ &= \Big(b - a - \int_0^t \frac{1}{1-s}\,dW_s\Big)dt + dW_t \\ &= \frac{b-X_t}{1-t}dt + dW_t, \end{aligned}$$

where the last identity comes from (7.1.3). We just verified that the process X_t is a solution of the given stochastic equation. The question of *how this solution was obtained in the first place*, is the subject of study for the next few sections.

7.2 Finding Mean and Variance from the Equation

For most practical purposes, the most important information one needs to know about a process is its mean and variance. These can be found directly from the stochastic equation in some particular cases without solving explicitly the equation. We shall deal in the present section with this problem.

Taking the expectation in (7.1.2) and using the property of the Ito integral as a zero mean random variable yields

$$E[X_t] = X_0 + \int_0^t E[a(s, W_s, X_s)]\,ds. \tag{7.2.4}$$

Applying the Fundamental Theorem of Calculus we obtain

$$\frac{d}{dt}E[X_t] = E[a(t, W_t, X_t)].$$

We note that X_t is not differentiable, but its expectation $E[X_t]$ is. This equation can be solved exactly in a few particular cases.

1. If $a(t, W_t, X_t) = a(t)$, then $\frac{d}{dt}E[X_t] = a(t)$ with the exact solution $E[X_t] = X_0 + \int_0^t a(s)\, ds$.

2. If $a(t, W_t, X_t) = \alpha(t)X_t + \beta(t)$, with $\alpha(t)$ and $\beta(t)$ continuous deterministic function, then

$$\frac{d}{dt}E[X_t] = \alpha(t)E[X_t] + \beta(t),$$

which is a linear differential equation in $E[X_t]$. Its solution is given by

$$E[X_t] = e^{A(t)}\Big(X_0 + \int_0^t e^{-A(s)}\beta(s)\, ds\Big), \tag{7.2.5}$$

where $A(t) = \int_0^t \alpha(s)\, ds$. It is worth noting that the expectation $E[X_t]$ does not depend on the volatility term $b(t, W_t, X_t)$.

Exercise 7.2.1 *If $dX_t = (2X_t + e^{2t})dt + b(t, W_t, X_t)dW_t$, then*

$$E[X_t] = e^{2t}(X_0 + t).$$

Proposition 7.2.2 *Let X_t be a process satisfying the stochastic equation*

$$dX_t = \alpha(t)X_t dt + b(t)dW_t.$$

Then the mean and variance of X_t are given by

$$\begin{aligned} E[X_t] &= e^{A(t)}X_0 \\ Var[X_t] &= e^{2A(t)}\int_0^t e^{-A(s)}b^2(s)\, ds, \end{aligned}$$

where $A(t) = \int_0^t \alpha(s)\, ds$.

Proof: The expression of $E[X_t]$ follows directly from formula (7.2.5) with $\beta = 0$. In order to compute the second moment we first compute

$$\begin{aligned} (dX_t)^2 &= b^2(t)\, dt; \\ d(X_t^2) &= 2X_t dX_t + (dX_t)^2 \\ &= 2X_t\big(\alpha(t)X_t dt + b(t)dW_t\big) + b^2(t)dt \\ &= \big(2\alpha(t)X_t^2 + b^2(t)\big)dt + 2b(t)X_t dW_t, \end{aligned}$$

where we used Ito's formula. If we let $Y_t = X_t^2$, the previous equation becomes

$$dY_t = \big(2\alpha(t)Y_t + b^2(t)\big)dt + 2b(t)\sqrt{Y_t}\, dW_t.$$

Applying formula (7.2.5) with $\alpha(t)$ replaced by $2\alpha(t)$ and $\beta(t)$ by $b^2(t)$, yields

$$E[Y_t] = e^{2A(t)}\Big(Y_0 + \int_0^t e^{-2A(s)} b^2(s)\,ds\Big),$$

which is equivalent with

$$E[X_t^2] = e^{2A(t)}\Big(X_0^2 + \int_0^t e^{-2A(s)} b^2(s)\,ds\Big).$$

It follows that the variance is

$$Var[X_t] = E[X_t^2] - (E[X_t])^2 = e^{2A(t)} \int_0^t e^{-2A(s)} b^2(s)\,ds.$$

■

Remark 7.2.3 *We note that the previous equation is of linear type. This shall be solved explicitly in a future section.*

The mean and variance for a given stochastic process can be computed by working out the associated stochastic equation. We shall provide next a few examples.

Example 7.2.1 *Find the mean and variance of e^{kW_t}, with k constant.*

From Ito's formula

$$d(e^{kW_t}) = ke^{kW_t} dW_t + \frac{1}{2}k^2 e^{kW_t} dt,$$

and integrating yields

$$e^{kW_t} = 1 + k\int_0^t e^{kW_s}\,dW_s + \frac{1}{2}k^2 \int_0^t e^{kW_s}\,ds.$$

Taking the expectations we have

$$E[e^{kW_t}] = 1 + \frac{1}{2}k^2 \int_0^t E[e^{kW_s}]\,ds.$$

If we let $f(t) = E[e^{kW_t}]$, then differentiating the previous relations yields the differential equation

$$f'(t) = \frac{1}{2}k^2 f(t)$$

with the initial condition $f(0) = E[e^{kW_0}] = 1$. The solution is $f(t) = e^{k^2 t/2}$, and hence

$$\boxed{E[e^{kW_t}] = e^{k^2 t/2}.}$$

The variance is

$$\begin{aligned} Var(e^{kW_t}) &= E[e^{2kW_t}] - (E[e^{kW_t}])^2 = e^{4k^2 t/2} - e^{k^2 t} \\ &= e^{k^2 t}(e^{k^2 t} - 1). \end{aligned}$$

Example 7.2.2 *Find the mean of the process $W_t e^{W_t}$.*

We shall set up a stochastic differential equation for $W_t e^{W_t}$. Using the product formula and Ito's formula yields

$$\begin{aligned} d(W_t e^{W_t}) &= e^{W_t} dW_t + W_t d(e^{W_t}) + dW_t\, d(e^{W_t}) \\ &= e^{W_t} dW_t + (W_t + dW_t)(e^{W_t} dW_t + \frac{1}{2} e^{W_t} dt) \\ &= (\frac{1}{2} W_t e^{W_t} + e^{W_t}) dt + (e^{W_t} + W_t e^{W_t}) dW_t. \end{aligned}$$

Integrating and using that $W_0 e^{W_0} = 0$ yields

$$W_t e^{W_t} = \int_0^t (\frac{1}{2} W_s e^{W_s} + e^{W_s})\, ds + \int_0^t (e^{W_s} + W_s e^{W_s})\, dW_s.$$

Since the expectation of an Ito integral is zero, we have

$$E[W_t e^{W_t}] = \int_0^t \big(\frac{1}{2} E[W_s e^{W_s}] + E[e^{W_s}]\big)\, ds.$$

Let $f(t) = E[W_t e^{W_t}]$. Using $E[e^{W_s}] = e^{s/2}$, the previous integral equation becomes

$$f(t) = \int_0^t (\frac{1}{2} f(s) + e^{s/2})\, ds,$$

Differentiating yields the following linear differential equation

$$f'(t) = \frac{1}{2} f(t) + e^{t/2}$$

with the initial condition $f(0) = 0$. Multiplying by $e^{-t/2}$ yields the following exact equation

$$(e^{-t/2} f(t))' = 1.$$

The solution is $f(t) = t e^{t/2}$. Hence we obtained that

$$E[W_t e^{W_t}] = t e^{t/2}.$$

Exercise 7.2.4 *Find a $E[W_t^2 e^{W_t}]$; (b) $E[W_t e^{kW_t}]$.*

Example 7.2.3 *Show that for any integer $k \geq 0$ we have*

$$\boxed{E[W_t^{2k}] = \frac{(2k)!}{2^k k!} t^k, \qquad E[W_t^{2k+1}] = 0.}$$

In particular, $E[W_t^4] = 3t^2$, $E[W_t^6] = 15t^3$.

From Ito's formula we have

$$d(W_t^n) = nW_t^{n-1}dW_t + \frac{n(n-1)}{2}W_t^{n-2}dt.$$

Integrate and get

$$W_t^n = n\int_0^t W_s^{n-1}\,dW_s + \frac{n(n-1)}{2}\int_0^t W_s^{n-2}\,ds.$$

Since the expectation of the first integral on the right side is zero, taking the expectation yields the following recursive relation

$$E[W_t^n] = \frac{n(n-1)}{2}\int_0^t E[W_s^{n-2}]\,ds.$$

Using the initial values $E[W_t] = 0$ and $E[W_t^2] = t$, the method of mathematical induction implies that $E[W_t^{2k+1}] = 0$ and $E[W_t^{2k}] = \frac{(2k)!}{2^k k!}t^k$.

Exercise 7.2.5 *(a) Is* $W_t^4 - 3t^2$ *an* $\mathcal{F}_t$*-martingale?*
(b) What about W_t^3*?*

Example 7.2.4 *Find* $E[\sin W_t]$.

From Ito's formula

$$d(\sin W_t) = \cos W_t\,dW_t - \frac{1}{2}\sin W_t\,dt,$$

then integrating yields

$$\sin W_t = \int_0^t \cos W_s\,dW_s - \frac{1}{2}\int_0^t \sin W_s\,ds.$$

Taking expectations we arrive at the integral equation

$$E[\sin W_t] = -\frac{1}{2}\int_0^t E[\sin W_s]\,ds.$$

Let $f(t) = E[\sin W_t]$. Differentiating yields the equation $f'(t) = -\frac{1}{2}f(t)$ with $f(0) = E[\sin W_0] = 0$. The unique solution is $f(t) = 0$. Hence

$$E[\sin W_t] = 0.$$

Exercise 7.2.6 *Let* σ *be a constant. Show that*

(a) $E[\sin(\sigma W_t)] = 0$;
(b) $E[\cos(\sigma W_t)] = e^{-\sigma^2 t/2}$;
(c) $E[\sin(t + \sigma W_t)] = e^{-\sigma^2 t/2}\sin t$;
(d) $E[\cos(t + \sigma W_t)] = e^{-\sigma^2 t/2}\cos t$;

Exercise 7.2.7 *Use the previous exercise and the definition of expectation to show that*

$$(a) \quad \int_{-\infty}^{\infty} e^{-x^2} \cos x \, dx = \frac{\pi^{1/2}}{e^{1/4}};$$

$$(b) \quad \int_{-\infty}^{\infty} e^{-x^2/2} \cos x \, dx = \sqrt{\frac{2\pi}{e}}.$$

Exercise 7.2.8 *Using expectations show that*

$$(a) \quad \int_{-\infty}^{\infty} x e^{-ax^2+bx} \, dx = \sqrt{\frac{\pi}{a}} \Big(\frac{b}{2a}\Big) e^{b^2/(4a)};$$

$$(b) \quad \int_{-\infty}^{\infty} x^2 e^{-ax^2+bx} \, dx = \sqrt{\frac{\pi}{a}} \frac{1}{2a} \Big(1 + \frac{b^2}{2a}\Big) e^{b^2/(4a)};$$

(c) *Can you apply a similar method to find a close form expression for the integral*

$$\int_{-\infty}^{\infty} x^n e^{-ax^2+bx} \, dx?$$

Exercise 7.2.9 *Using the result given by Example 7.2.3 show that*

$(a) \quad E[\cos(tW_t)] = e^{-t^3/2};$

$(b) \quad E[\sin(tW_t)] = 0;$

$(c) \quad E[e^{tW_t}] = 0.$

For general drift rates we cannot find the mean, but in the case of concave drift rates we can find an upper bound for the expectation $E[X_t]$. The following result will be useful.

Lemma 7.2.10 (Gronwall's inequality) *Let $f(t)$ be a non-negative function satisfying the inequality*

$$f(t) \leq C + M \int_0^t f(s) \, ds$$

for $0 \leq t \leq T$, with C, M constants. Then

$$f(t) \leq Ce^{Mt}, \qquad 0 \leq t \leq T.$$

Proposition 7.2.11 *Let X_t be a continuous stochastic process such that*

$$dX_t = a(X_t)dt + b(t, W_t, X_t) \, dW_t,$$

with the function $a(\cdot)$ satisfying the following conditions

1. $a(x) \geq 0$, *for* $0 \leq x \leq T$;
2. $a''(x) < 0$, *for* $0 \leq x \leq T$;
3. $a'(0) = M$.

Then $E[X_t] \leq X_0 e^{Mt}$, for $0 \leq X_t \leq T$.

Proof: From the mean value theorem there is $\xi \in (0, x)$ such that

$$a(x) = a(x) - a(0) = (x - 0)a'(\xi) \le xa'(0) = Mx, \tag{7.2.6}$$

where we used that $a'(x)$ is a decreasing function. Applying Jensen's inequality for concave functions yields

$$E[a(X_t)] \le a(E[X_t]).$$

Combining with (7.2.6) we obtain $E[a(X_t)] \le ME[X_t]$. Substituting in the identity (7.2.4) implies

$$E[X_t] \le X_0 + M \int_0^t E[X_s]\, ds.$$

Applying Gronwall's inequality we obtain $E[X_t] \le X_0 e^{Mt}$. ∎

Exercise 7.2.12 *State the previous result in the particular case when $a(x) = \sin x$, with $0 \le x \le \pi$.*

Not in all cases can the mean and the variance be obtained directly from the stochastic equation. In these cases we need more powerful methods that produce closed form solutions. In the next sections we shall discuss several methods of solving stochastic differential equation.

7.3 The Integration Technique

We shall start with the simple case when both the drift and the volatility are just functions of time t.

Proposition 7.3.1 *The solution X_t of the stochastic differential equation*

$$dX_t = a(t)dt + b(t)dW_t$$

is Gaussian distributed with mean $X_0 + \int_0^t a(s)\, ds$ and variance $\int_0^t b^2(s)\, ds$.

Proof: Integrating in the equation yields

$$X_t - X_0 = \int_0^t dX_s = \int_0^t a(s)\, ds + \int_0^t b(s)\, dW_s.$$

Using the property of Wiener integrals, $\int_0^t b(s)\, dW_s$ is Gaussian distributed with mean 0 and variance $\int_0^t b^2(s)\, ds$. Then X_t is Gaussian (as a sum between a predictable function

and a Gaussian), with

$$\begin{aligned}
E[X_t] &= E[X_0 + \int_0^t a(s)\,ds + \int_0^t b(s)\,dW_s] \\
&= X_0 + \int_0^t a(s)\,ds + E[\int_0^t b(s)\,dW_s] \\
&= X_0 + \int_0^t a(s)\,ds, \\
Var[X_t] &= Var[X_0 + \int_0^t a(s)\,ds + \int_0^t b(s)\,dW_s] \\
&= Var\Big[\int_0^t b(s)\,dW_s\Big] \\
&= \int_0^t b^2(s)\,ds,
\end{aligned}$$

which ends the proof. ■

Exercise 7.3.2 *Solve the following stochastic differential equations for $t \geq 0$ and determine the mean and the variance of the solution*

(a) $dX_t = \cos t\,dt - \sin t\,dW_t,\ X_0 = 1.$

(b) $dX_t = e^t\,dt + \sqrt{t}\,dW_t,\ X_0 = 0.$

(c) $dX_t = \frac{t}{1+t^2}\,dt + t^{3/2}\,dW_t,\ X_0 = 1.$

If the drift and the volatility depend on both variables t and W_t, the stochastic differential equation

$$dX_t = a(t, W_t)dt + b(t, W_t)dW_t, \qquad t \geq 0$$

defines an *Ito diffusion*. Integrating yields the solution

$$X_t = X_0 + \int_0^t a(s, W_s)\,ds + \int_0^t b(s, W_s)\,dW_s.$$

There are several cases when both integrals can be computed explicitly.

Example 7.3.1 *Find the solution of the stochastic differential equation*

$$dX_t = dt + W_t\,dW_t, \qquad X_0 = 1.$$

Integrate between 0 and t and get

$$\begin{aligned}
X_t &= 1 + \int_0^t ds + \int_0^t W_s\,dW_s = t + \frac{W_t^2}{2} - \frac{t}{2} \\
&= \frac{1}{2}(W_t^2 + t).
\end{aligned}$$

Example 7.3.2 *Solve the stochastic differential equation*

$$dX_t = (W_t - 1)dt + W_t^2\, dW_t, \qquad X_0 = 0.$$

Let $Z_t = \int_0^t W_s\, ds$ denote the integrated Brownian motion process. Integrating the equation between 0 and t yields

$$\begin{aligned} X_t &= \int_0^s dX_s = \int_0^t (W_s - 1)ds + \int_0^t W_s^2\, dW_s \\ &= Z_t - t + \frac{1}{3}W_t^3 - \frac{1}{2}W_t^2 - \frac{t}{2} \\ &= Z_t + \frac{1}{3}W_t^3 - \frac{1}{2}W_t^2 - \frac{t}{2}. \end{aligned}$$

Example 7.3.3 *Solve the stochastic differential equation*

$$dX_t = t^2 dt + e^{t/2}\cos W_t\, dW_t, \qquad X_0 = 0,$$

and find $E[X_t]$ *and* $Var(X_t)$*.*

Integrating yields

$$\begin{aligned} X_t &= \int_0^t s^2\, ds + \int_0^t e^{s/2}\cos W_s\, dW_s \\ &= \frac{t^3}{3} + e^{t/2}\sin W_t, \end{aligned} \tag{7.3.7}$$

where we used (6.3.9). Even if the process X_t is not Gaussian, we can still compute its mean and variance. By Ito's formula we have

$$d(\sin W_t) = \cos W_t\, dW_t - \frac{1}{2}\sin W_t\, dt$$

Integrating between 0 and t yields

$$\sin W_t = \int_0^t \cos W_s\, dW_s - \frac{1}{2}\int_0^t \sin W_s\, ds,$$

where we used that $\sin W_0 = \sin 0 = 0$. Taking the expectation in the previous relation yields

$$E[\sin W_t] = E\Big[\int_0^t \cos W_s\, dW_s\Big] - \frac{1}{2}\int_0^t E[\sin W_s]\, ds.$$

From the properties of the Ito integral, the first expectation on the right side is zero. Denoting $\mu(t) = E[\sin W_t]$, we obtain the integral equation

$$\mu(t) = -\frac{1}{2}\int_0^t \mu(s)\, ds.$$

Differentiating yields the differential equation

$$\mu'(t) = -\frac{1}{2}\mu(t)$$

with the solution $\mu(t) = ke^{-t/2}$. Since $k = \mu(0) = E[\sin W_0] = 0$, it follows that $\mu(t) = 0$. Hence

$$E[\sin W_t] = 0.$$

Taking expectation in (7.3.7) leads to

$$E[X_t] = E\Big[\frac{t^3}{3}\Big] + e^{t/2}E[\sin W_t] = \frac{t^3}{3}.$$

Since the variance of predictable functions is zero,

$$\begin{aligned} Var[X_t] &= Var\Big[\frac{t^3}{3} + e^{t/2}\sin W_t\Big] = (e^{t/2})^2 Var[\sin W_t] \\ &= e^t E[\sin^2 W_t] = \frac{e^t}{2}(1 - E[\cos 2W_t]). \end{aligned} \tag{7.3.8}$$

In order to compute the last expectation we use Ito's formula

$$d(\cos 2W_t) = -2\sin 2W_t\, dW_t - 2\cos 2W_t\, dt$$

and integrate to get

$$\cos 2W_t = \cos 2W_0 - 2\int_0^t \sin 2W_s\, dW_s - 2\int_0^t \cos 2W_s\, ds$$

Taking the expectation and using that Ito integrals have zero expectation, yields

$$E[\cos 2W_t] = 1 - 2\int_0^t E[\cos 2W_s]\, ds.$$

If we denote $m(t) = E[\cos 2W_t]$, the previous relation becomes an integral equation

$$m(t) = 1 - 2\int_0^t m(s)\, ds.$$

Differentiate and get

$$m'(t) = -2m(t),$$

with the solution $m(t) = ke^{-2t}$. Since $k = m(0) = E[\cos 2W_0] = 1$, we have $m(t) = e^{-2t}$. Substituting into (7.3.8) yields

$$Var[X_t] = \frac{e^t}{2}(1 - e^{-2t}) = \frac{e^t - e^{-t}}{2} = \sinh t.$$

In conclusion, the solution X_t has the mean and the variance given by

$$E[X_t] = \frac{t^3}{3}, \qquad Var[X_t] = \sinh t.$$

Example 7.3.4 *Solve the following stochastic differential equation*

$$e^{t/2}dX_t = dt + e^{W_t}\,dW_t, \qquad X_0 = 0,$$

and find the distribution of the solution X_t and its mean and variance.

Dividing by $e^{t/2}$, integrating between 0 and t, and using formula (6.3.8) yields

$$\begin{aligned} X_t &= \int_0^t e^{-s/2}\,ds + \int_0^t e^{-s/2+W_s}\,dW_s \\ &= 2(1-e^{-t/2}) + e^{-t/2}e^{W_t} - 1 \\ &= 1 + e^{-t/2}(e^{W_t} - 2). \end{aligned}$$

Since e^{W_t} is a geometric Brownian motion, using Proposition 2.2.2 yields

$$\begin{aligned} E[X_t] &= E[1 + e^{-t/2}(e^{W_t} - 2)] = 1 - 2e^{-t/2} + e^{-t/2}E[e^{W}_t] \\ &= 2 - 2e^{-t/2}. \\ Var(X_t) &= Var[1 + e^{-t/2}(e^{W_t} - 2)] = Var[e^{-t/2}e^{W_t}] = e^{-t}Var[e^{W_t}] \\ &= e^{-t}(e^{2t} - e^t) = e^t - 1. \end{aligned}$$

The process X_t has the following distribution:

$$\begin{aligned} F(y) &= P(X_t \le y) = P\big(1 + e^{-t/2}(e^{W_t} - 2) \le y\big) \\ &= P\Big(W_t \le \ln\big(2 + e^{t/2}(y-1)\big)\Big) = P\Big(\frac{W_t}{\sqrt{t}} \le \frac{1}{\sqrt{t}}\ln\big(2 + e^{t/2}(y-1)\big)\Big) \\ &= N\Big(\frac{1}{\sqrt{t}}\ln\big(2 + e^{t/2}(y-1)\big)\Big), \end{aligned}$$

where $N(u) = \dfrac{1}{\sqrt{2\pi}}\displaystyle\int_{-\infty}^u e^{-s^2/2}\,ds$ is the distribution function of a standard normal distributed random variable.

Example 7.3.5 *Solve the stochastic differential equation*

$$dX_t = dt + t^{-3/2}W_t e^{-W_t^2/(2t)}\,dW_t, \qquad X_1 = 1.$$

Integrating between 1 and t and applying formula (6.3.10) yields

$$\begin{aligned} X_t &= X_1 + \int_1^t ds + \int_1^t s^{-3/2}W_s e^{-W_s^2/(2s)}\,dW_s \\ &= 1 + t - 1 - e^{-W_1^2/2} - \frac{1}{t^{1/2}}e^{-W_t^2/(2t)} \\ &= t - e^{-W_1^2/2} - \frac{1}{t^{1/2}}e^{-W_t^2/(2t)}, \qquad \forall t \ge 1. \end{aligned}$$

7.4 Exact Stochastic Equations

The stochastic differential equation

$$dX_t = a(t, W_t)dt + b(t, W_t)dW_t \tag{7.4.9}$$

is called *exact* if there is a differentiable function $f(t, x)$ such that

$$a(t, x) = \partial_t f(t, x) + \frac{1}{2}\partial_x^2 f(t, x) \tag{7.4.10}$$
$$b(t, x) = \partial_x f(t, x). \tag{7.4.11}$$

Assume the equation is exact. Then substituting in (7.4.9) yields

$$dX_t = \Big(\partial_t f(t, W_t) + \frac{1}{2}\partial_x^2 f(t, W_t)\Big)dt + \partial_x f(t, W_t)dW_t.$$

Applying Ito's formula, the previous equations becomes

$$dX_t = d\big(f(t, W_t)\big),$$

which implies $X_t = f(t, W_t) + c$, with c constant.

Solving the partial differential equations system (7.4.10)–(7.4.10) requires the following steps:

1. Integrating partially with respect to x in the second equation to obtain $f(t, x)$ up to an additive function $T(t)$;

2. Substitute into the first equation and determine the function $T(t)$;

3. The solution is $X_t = f(t, W_t) + c$, with c determined from the initial condition on X_t.

Example 7.4.1 *Solve the stochastic differential equation*

$$dX_t = e^t(1 + W_t^2)dt + (1 + 2e^t W_t)dW_t, \quad X_0 = 0.$$

In this case $a(t, x) = e^t(1 + x^2)$ and $b(t, x) = 1 + 2e^t x$. The associated system is

$$\begin{aligned} e^t(1 + x^2) &= \partial_t f(t, x) + \frac{1}{2}\partial_x^2 f(t, x) \\ 1 + 2e^t x &= \partial_x f(t, x). \end{aligned}$$

Integrate partially in x in the second equation yields

$$f(t, x) = \int (1 + 2e^t x)\, dx = x + e^t x^2 + T(t).$$

Then $\partial_t f = e^t x^2 + T'(t)$ and $\partial_x^2 f = 2e^t$. Substituting in the first equation yields

$$e^t(1 + x^2) = e^t x^2 + T'(t) + e^t.$$

This implies $T'(t) = 0$, or $T = c$ constant. Hence $f(t, x) = x + e^t x^2 + c$, and $X_t = f(t, W_t) = W_t + e^t W_t^2 + c$. Since $X_0 = 0$, it follows that $c = 0$. The solution is $X_t = W_t + e^t W_t^2$.

Example 7.4.2 *Find the solution of*

$$dX_t = \big(2tW_t^3 + 3t^2(1+W_t)\big)dt + (3t^2W_t^2 + 1)dW_t, \quad X_0 = 0.$$

The coefficient functions are $a(t,x) = 2tx^3 + 3t^2(1+x)$ and $b(t,x) = 3t^2x^2 + 1$. The associated system is given by

$$\begin{aligned} 2tx^3 + 3t^2(1+x) &= \partial_t f(t,x) + \frac{1}{2}\partial_x^2 f(t,x) \\ 3t^2x^2 + 1 &= \partial_x f(t,x). \end{aligned}$$

Integrating partially in the second equation yields

$$f(t,x) = \int (3t^2x^2 + 1)\,dx = t^2x^3 + x + T(t).$$

Then $\partial_t f = 2tx^3 + T'(t)$ and $\partial_x^2 f = 6t^2x$, and substituting into the first equation we get

$$2tx^3 + 3t^2(1+x) = 2tx^3 + T'(t) + \frac{1}{2}6t^2x.$$

After cancellations we get $T'(t) = 3t^2$, so $T(t) = t^3 + c$. Then

$$f(t,x) = t^2x^3 + x + 3t^2 = t^2(x^3+1) + x + c.$$

The solution process is given by $X_t = f(t,W_t) = t^2(W_t^3 + 1) + W_t + c$. Using $X_0 = 0$ we get $c = 0$. Hence the solution is $X_t = t^2(W_t^3+1) + W_t$.

The next result deals with a condition regarding the closeness of the stochastic differential equation.

Theorem 7.4.1 *If the stochastic differential equation (7.4.9) is exact, then the coefficient functions $a(t,x)$ and $b(t,x)$ satisfy the condition*

$$\boxed{\partial_x a = \partial_t b + \frac{1}{2}\partial_x^2 b.} \tag{7.4.12}$$

Proof: If the stochastic equation is exact, there is a function $f(t,x)$ satisfying the system (7.4.10)–(7.4.10). Differentiating the first equation of the system with respect to x yields

$$\partial_x a = \partial_t\partial_x f + \frac{1}{2}\partial_x^2\partial_x f.$$

Substituting $b = \partial_x f$ yields the desired relation. ■

Remark 7.4.2 *The equation (7.4.12) has the meaning of a heat equation. The function $b(t,x)$ represents the temperature measured at x at the instance t, while $\partial_x a$ is the density of heat sources. The function $a(t,x)$ can be regarded as the potential from which the density of heat sources is derived by taking the gradient in x.*

It is worth noting that equation (7.4.12) is a just necessary condition for exactness. This means that if this condition is not satisfied, then the equation is not exact. In that case we need to try a different method to solve the equation.

Example 7.4.3 *Is the stochastic differential equation*

$$dX_t = (1 + W_t^2)dt + (t^4 + W_t^2)dW_t$$

exact?

Collecting the coefficients, we have $a(t, x) = 1 + x^2$, $b(t, x) = t^4 + x^2$. Since $\partial_x a = 2x$, $\partial_t b = 4t^3$, and $\partial_x^2 b = 2$, the condition (7.4.12) is not satisfied, and hence the equation is not exact.

7.5 Integration by Inspection

When solving a stochastic differential equation by inspection we look for opportunities to apply the product or the quotient formulas:

$$\boxed{d(f(t)Y_t) = f(t)\, dY_t + Y_t\, df(t).}$$

$$\boxed{d\Big(\frac{X_t}{f(t)}\Big) = \frac{f(t)dX_t - X_t df(t)}{f(t)^2}.}$$

For instance, if a stochastic differential equation can be written as

$$dX_t = f'(t)W_t dt + f(t)dW_t,$$

the product rule brings the equation into the exact form

$$dX_t = d\Big(f(t)W_t\Big),$$

which after integration leads to the solution

$$X_t = X_0 + f(t)W_t.$$

Example 7.5.1 *Solve*

$$dX_t = (t + W_t^2)dt + 2tW_t dW_t, \qquad X_0 = a.$$

We can write the equation as

$$dX_t = W_t^2 dt + t(2W_t dW_t + dt),$$

which can be contracted to

$$dX_t = W_t^2 dt + td(W_t^2).$$

Using the product rule we can bring it to the exact form

$$dX_t = d(tW_t^2),$$

with the solution $X_t = tW_t^2 + a$.

Example 7.5.2 *Solve the stochastic differential equation*

$$dX_t = (W_t + 3t^2)dt + tdW_t.$$

If we rewrite the equation as

$$dX_t = 3t^2 dt + (W_t dt + tdW_t),$$

we note the exact expression formed by the last two terms $W_t dt + tdW_t = d(tW_t)$. Then

$$dX_t = d(t^3) + d(tW_t),$$

which is equivalent with $d(X_t) = d(t^3 + tW_t)$. Hence $X_t = t^3 + tW_t + c$, $c \in \mathbb{R}$.

Example 7.5.3 *Solve the stochastic differential equation*

$$e^{-2t} dX_t = (1 + 2W_t^2)dt + 2W_t dW_t.$$

Multiply by e^{2t} to get

$$dX_t = e^{2t}(1 + 2W_t^2)dt + 2e^{2t} W_t dW_t.$$

After regrouping, this becomes

$$dX_t = (2e^{2t} dt)W_t^2 + e^{2t}(2W_t dW_t + dt).$$

Since $d(e^{2t}) = 2e^{2t} dt$ and $d(W_t^2) = 2W_t dW_t + dt$, the previous relation becomes

$$dX_t = d(e^{2t})W_t^2 + e^{2t} d(W_t^2).$$

By the product rule, the right side becomes exact

$$dX_t = d(e^{2t} W_t^2),$$

and hence the solution is $X_t = e^{2t} W_t^2 + c$, $c \in \mathbb{R}$.

Example 7.5.4 *Solve the equation*

$$t^3 dX_t = (3t^2 X_t + t)dt + t^6 dW_t, \qquad X_1 = 0.$$

The equation can be written as

$$t^3 dX_t - 3X_t t^2 dt = tdt + t^6 dW_t.$$

Divide by t^6:

$$\frac{t^3 dX_t - X_t d(t^3)}{(t^3)^2} = t^{-5} dt + dW_t.$$

Applying the quotient rule yields

$$d\Big(\frac{X_t}{t^3}\Big) = -d\Big(\frac{t^{-4}}{4}\Big) + dW_t.$$

Integrating between 1 and t, yields

$$\frac{X_t}{t^3} = -\frac{t^{-4}}{4} + W_t - W_1 + c$$

so

$$X_t = ct^3 - \frac{1}{4t} + t^3(W_t - W_1), \quad c \in \mathbb{R}.$$

Using $X_1 = 0$ yields $c = 1/4$ and hence the solution is

$$X_t = \frac{1}{4}\Big(t^3 - \frac{1}{t}\Big) + t^3(W_t - W_1), \quad c \in \mathbb{R}.$$

Exercise 7.5.1 *Solve the following stochastic differential equations by the inspection method*

(a) $dX_t = (1 + W_t)dt + (t + 2W_t)dW_t, \qquad X_0 = 0;$

(b) $t^2 dX_t = (2t^3 - W_t)dt + tdW_t, \qquad X_1 = 0;$

(c) $e^{-t/2}dX_t = \frac{1}{2}W_t dt + dW_t, \qquad X_0 = 0;$

(d) $dX_t = 2tW_t dW_t + W_t^2 dt, \qquad X_0 = 0;$

(e) $dX_t = \Big(1 + \frac{1}{2\sqrt{t}}W_t\Big)dt + \sqrt{t}\,dW_t, \qquad X_1 = 0.$

7.6 Linear Stochastic Differential Equations

Consider the stochastic differential equation with drift term linear in X_t

$$dX_t = \big(\alpha(t)X_t + \beta(t)\big)dt + b(t, W_t)dW_t, \qquad t \geq 0.$$

This also can be written as

$$dX_t - \alpha(t)X_t dt = \beta(t)dt + b(t, W_t)dW_t.$$

Let $A(t) = \int_0^t \alpha(s)\,ds$. Multiplying by the integrating factor $e^{-A(t)}$, the left side of the previous equation becomes an exact expression

$$\begin{aligned} e^{-A(t)}\Big(dX_t - \alpha(t)X_t dt\Big) &= e^{-A(t)}\beta(t)dt + e^{-A(t)}b(t, W_t)dW_t \\ d\Big(e^{-A(t)}X_t\Big) &= e^{-A(t)}\beta(t)dt + e^{-A(t)}b(t, W_t)dW_t. \end{aligned}$$

Integrating yields

$$\begin{aligned} e^{-A(t)}X_t &= X_0 + \int_0^t e^{-A(s)}\beta(s)\,ds + \int_0^t e^{-A(s)}b(s,W_s)\,dW_s \\ X_t &= X_0e^{A(t)} + e^{A(t)}\Big(\int_0^t e^{-A(s)}\beta(s)\,ds + \int_0^t e^{-A(s)}b(s,W_s)\,dW_s\Big). \end{aligned}$$

The first integral within the previous parentheses is a Riemann integral, and the latter one is an Ito stochastic integral. Sometimes, in practical applications these integrals can be computed explicitly.

When $b(t,W_t) = b(t)$, the latter integral becomes a Wiener integral. In this case the solution X_t is Gaussian with mean and variance given by

$$\begin{aligned} E[X_t] &= X_0e^{A(t)} + e^{A(t)}\int_0^t e^{-A(s)}\beta(s)\,ds \\ Var[X_t] &= e^{2A(t)}\int_0^t e^{-2A(s)}b(s)^2\,ds. \end{aligned}$$

Another important particular case is when $\alpha(t) = \alpha \neq 0$, $\beta(t) = \beta$ are constants and $b(t,W_t) = b(t)$. The equation in this case is

$$dX_t = \big(\alpha X_t + \beta\big)dt + b(t)dW_t, \qquad t \geq 0,$$

and the solution takes the form

$$X_t = X_0e^{\alpha t} + \frac{\beta}{\alpha}(e^{\alpha t} - 1) + \int_0^t e^{\alpha(t-s)}b(s)\,dW_s.$$

Example 7.6.1 *Solve the linear stochastic differential equation*

$$dX_t = (2X_t + 1)dt + e^{2t}dW_t.$$

Write the equation as

$$dX_t - 2X_tdt = dt + e^{2t}dW_t$$

and multiply by the integrating factor e^{-2t} to get

$$d\Big(e^{-2t}X_t\Big) = e^{-2t}dt + dW_t.$$

Integrate between 0 and t and multiply by e^{2t}, and obtain

$$\begin{aligned} X_t &= X_0e^{2t} + e^{2t}\int_0^t e^{-2s}ds + e^{2t}\int_0^t dW_s \\ &= X_0e^{2t} + \frac{1}{2}(e^{2t} - 1) + e^{2t}W_t. \end{aligned}$$

Example 7.6.2 *Solve the linear stochastic differential equation*

$$dX_t = (2 - X_t)dt + e^{-t}W_t dW_t.$$

Multiplying by the integrating factor e^t yields

$$e^t(dX_t + X_t dt) = 2e^t dt + W_t\,dW_t.$$

Since $e^t(dX_t + X_t dt) = d(e^t X_t)$, integrating between 0 and t we get

$$e^t X_t = X_0 + \int_0^t 2e^t\,dt + \int_0^t W_s\,dW_s.$$

Dividing by e^t and performing the integration yields

$$X_t \quad = \quad X_0 e^{-t} + 2(1 - e^{-t}) + \frac{1}{2}e^{-t}(W_t^2 - t).$$

Example 7.6.3 *Solve the linear stochastic differential equation*

$$dX_t = (\frac{1}{2}X_t + 1)dt + e^t \cos W_t\,dW_t.$$

Write the equation as

$$dX_t - \frac{1}{2}X_t dt = dt + e^t \cos W_t\,dW_t$$

and multiply by the integrating factor $e^{-t/2}$ to get

$$d(e^{-t/2}X_t) = e^{-t/2}dt + e^{t/2}\cos W_t\,dW_t.$$

Integrating yields

$$e^{-t/2}X_t = X_0 + \int_0^t e^{-s/2}\,ds + \int_0^t e^{s/2}\cos W_s\,dW_s$$

Multiply by $e^{t/2}$ and use formula (6.3.9) to obtain the solution

$$X_t = X_0 e^{t/2} + 2(e^{t/2} - 1) + e^t \sin W_t.$$

Exercise 7.6.1 *Solve the following linear stochastic differential equations*

(a) $dX_t = (4X_t - 1)dt + 2dW_t$;

(b) $dX_t = (3X_t - 2)dt + e^{3t}dW_t$;

(c) $dX_t = (1 + X_t)dt + e^t W_t dW_t$;

(d) $dX_t = (4X_t + t)dt + e^{4t}dW_t$;

(e) $dX_t = \left(t + \frac{1}{2}X_t\right)dt + e^t \sin W_t\,dW_t$;

(f) $dX_t = -X_t dt + e^{-t}\,dW_t$.

In the following we present an important example of stochastic differential equations, which can be solved by the method presented in this section.

Proposition 7.6.2 (The mean-reverting Ornstein-Uhlenbeck process) *Let m and α be two constants. Then the solution X_t of the stochastic equation*

$$\boxed{dX_t = (m - X_t)dt + \alpha dW_t} \tag{7.6.13}$$

is given by

$$\boxed{X_t = m + (X_0 - m)e^{-t} + \alpha \int_0^t e^{s-t}\, dW_s.} \tag{7.6.14}$$

X_t is Gaussian with mean and variance given by

$$\begin{aligned} E[X_t] &= m + (X_0 - m)e^{-t} \\ Var(X_t) &= \frac{\alpha^2}{2}(1 - e^{-2t}). \end{aligned}$$

Proof: Adding $X_t dt$ to both sides and multiplying by the integrating factor e^t we get

$$d(e^t X_t) = me^t dt + \alpha e^t dW_t,$$

which after integration yields

$$e^t X_t = X_0 + m(e^t - 1) + \alpha \int_0^t e^s\, dW_s.$$

Hence

$$\begin{aligned} X_t &= X_0 e^{-t} + m - e^{-t} + \alpha e^{-t} \int_0^t e^s\, dW_s \\ &= m + (X_0 - m)e^{-t} + \alpha \int_0^t e^{s-t}\, dW_s. \end{aligned}$$

Since X_t is the sum between a predictable function and a Wiener integral, then we can use Proposition 4.6.1 and it follows that X_t is Gaussian, with

$$\begin{aligned} E[X_t] &= m + (X_0 - m)e^{-t} + E\Big[\alpha \int_0^t e^{s-t}\, dW_s\Big] = m + (X_0 - m)e^{-t} \\ Var(X_t) &= Var\Big[\alpha \int_0^t e^{s-t}\, dW_s\Big] = \alpha^2 e^{-2t} \int_0^t e^{2s}\, ds \\ &= \alpha^2 e^{-2t} \frac{e^{2t} - 1}{2} = \frac{1}{2}\alpha^2 (1 - e^{-2t}). \end{aligned}$$

∎

The name *mean-reverting* comes from the fact that

$$\lim_{t\to\infty} E[X_t] = m.$$

The variance also tends to zero exponentially, $\lim_{t\to\infty} Var[X_t] = 0$. According to Proposition 3.8.1, the process X_t tends to m in the mean square sense.

Proposition 7.6.3 (The Brownian bridge) *For $a, b \in \mathbb{R}$ fixed, the stochastic differential equation*

$$dX_t = \frac{b - X_t}{1-t}dt + dW_t, \quad 0 \le t < 1, X_0 = a$$

has the solution

$$\boxed{X_t = a(1-t) + bt + (1-t)\int_0^t \frac{1}{1-s}dW_s,\ 0 \le t < 1.} \tag{7.6.15}$$

The solution has the property $\lim_{t\to 1} X_t = b$, *almost certainly.*

Proof: If we let $Y_t = b - X_t$ the equation becomes linear in Y_t

$$dY_t + \frac{1}{1-t}Y_t dt = -dW_t.$$

Multiplying by the integrating factor $\rho(t) = \frac{1}{1-t}$ yields

$$d\Big(\frac{Y_t}{1-t}\Big) = -\frac{1}{1-t}dW_t,$$

which leads by integration to

$$\frac{Y_t}{1-t} = c - \int_0^t \frac{1}{1-s}\,dW_s.$$

Making $t = 0$ yields $c = a - b$, so

$$\frac{b - X_t}{1-t} = a - b - \int_0^t \frac{1}{1-s}\,dW_s.$$

Solving for X_t yields

$$X_t = a(1-t) + bt + (1-t)\int_0^t \frac{1}{1-s}dW_s,\ 0 \le t < 1.$$

Let $U_t = (1-t)\int_0^t \frac{1}{1-s}\, dW_s$. First we notice that

$$\begin{aligned}
E[U_t] &= (1-t)E\Big[\int_0^t \frac{1}{1-s}\, dW_s\Big] = 0,\\
Var(U_t) &= (1-t)^2 Var\Big[\int_0^t \frac{1}{1-s}\, dW_s\Big] = (1-t)^2\int_0^t \frac{1}{(1-s)^2}\, ds\\
&= (1-t)^2\Big(\frac{1}{1-t} - 1\Big) = t(1-t).
\end{aligned}$$

In order to show ac-$\lim_{t\to 1} X_t = b$, we need to prove

$$P\big(\omega; \lim_{t\to 1} X_t(\omega) = b\big) = 1.$$

Since $X_t = a(1-t) + bt + U_t$, it suffices to show that

$$P\big(\omega; \lim_{t\to 1} U_t(\omega) = 0\big) = 1. \tag{7.6.16}$$

We evaluate the probability of the complementary event

$$P\big(\omega; \lim_{t\to 1} U_t(\omega) \neq 0\big) = P\big(\omega; |U_t(\omega)| > \epsilon, \forall t\big),$$

for some $\epsilon > 0$. Since by Markov's inequality

$$P\big(\omega; |U_t(\omega)| > \epsilon\big) < \frac{Var(U_t)}{\epsilon^2} = \frac{t(1-t)}{\epsilon^2}$$

holds for any $0 \le t < 1$, choosing $t \to 1$ implies that

$$P\big(\omega; |U_t(\omega)| > \epsilon, \forall t\big) = 0,$$

which implies (7.6.16). ∎

The process (7.6.15) is called the *Brownian bridge* because it joins $X_0 = a$ with $X_1 = b$. Since X_t is the sum between a deterministic linear function in t and a Wiener integral, it follows that it is a Gaussian process, with mean and variance

$$\begin{aligned}
E[X_t] &= a(1-t) + bt\\
Var(X_t) &= Var(U_t) = t(1-t).
\end{aligned}$$

It is worth noting that the variance is maximum at the midpoint $t = (b-a)/2$ and zero at the end points a and b.

Example 7.6.4 *Find $Cov(X_s, X_t)$, $0 < s < t$ for the following cases:*

(a) X_t is a mean reverting Orstein-Uhlenbeck process;

(b) X_t is a Brownian bridge process.

Stochastic equations with respect to a Poisson process

Similar techniques can be applied in the case when the Brownian motion process W_t is replaced by a Poisson process N_t with constant rate λ. For instance, the stochastic differential equation

$$\begin{aligned} dX_t &= 3X_t dt + e^{3t} dN_t \\ X_0 &= 1 \end{aligned}$$

can be solved multiplying by the integrating factor e^{-3t} to obtain

$$d(e^{-3t}X_t) = dN_t.$$

Integrating yields $e^{-3t}X_t = N_t + C$. Making $t = 0$ yields $C = 1$, so the solution is given by $X_t = e^{3t}(1 + N_t)$.

The following equation

$$dX_t = (m - X_t)dt + \alpha dN_t$$

is similar with the equation defining the mean-reverting Ornstein-Uhlenbeck process. As we shall see, in this case, the process is no more mean-reverting. It reverts though to a certain constant. A similar method yields the solution

$$X_t = m + (X_0 - m)e^{-t} + \alpha e^{-t} \int_0^t e^s \, dN_s.$$

Since from Proposition 4.8.5 and Exercise 4.8.8 we have

$$\begin{aligned} E\Big[\int_0^t e^s \, dN_s\Big] &= \lambda \int_0^t e^s \, ds = \lambda(e^t - 1) \\ Var\Big(\int_0^t e^s \, dN_s\Big) &= \lambda \int_0^t e^{2s} \, ds = \frac{\lambda}{2}(e^{2t} - 1), \end{aligned}$$

it follows that

$$\begin{aligned} E[X_t] &= m + (X_0 - m)e^{-t} + \alpha\lambda(1 - e^{-t}) \\ Var(X_t) &= \frac{\lambda\alpha^2}{2}(1 - e^{-2t}). \end{aligned}$$

It is worth noting that in this case the process X_t is no more Gaussian.

7.7 The Method of Variation of Parameters

Let's start by considering the following stochastic equation

$$dX_t = \alpha X_t dW_t, \tag{7.7.17}$$

with α constant. This is the equation which, in physics is known to model the *linear noise*. Dividing by X_t yields

$$\frac{dX_t}{X_t} = \alpha\, dW_t.$$

Switch to the integral form

$$\int \frac{dX_t}{X_t} = \int \alpha\, dW_t,$$

and integrate "blindly" to get $\ln X_t = \alpha W_t + c$, with c an integration constant. This leads to the "pseudo-solution"

$$X_t = e^{\alpha W_t + c}.$$

The nomination "pseudo" stands for the fact that X_t does not satisfy the initial equation. We shall find a correct solution by letting the parameter c be a function of t. In other words, we are looking for a solution of the following type:

$$X_t = e^{\alpha W_t + c(t)}, \tag{7.7.18}$$

where the function $c(t)$ is subject to be determined. Using Ito's formula we get

$$\begin{aligned} dX_t &= d(e^{W_t + c(t)}) = e^{\alpha W_t + c(t)}(c'(t) + \alpha^2/2)dt + \alpha e^{\alpha W_t + c(t)} dW_t \\ &= X_t(c'(t) + \alpha^2/2)dt + \alpha X_t\, dW_t. \end{aligned}$$

Substituting the last term from the initial equation (7.7.17) yields

$$dX_t = X_t(c'(t) + \alpha^2/2)dt + dX_t,$$

which leads to the equation

$$c'(t) + \alpha^2/2 = 0$$

with the solution $c(t) = -\frac{\alpha^2}{2}t + k$. Substituting into (7.7.18) yields

$$X_t = e^{\alpha W_t - \frac{\alpha^2}{2}t + k}.$$

The value of the constant k is determined by taking $t = 0$. This leads to $X_0 = e^k$. Hence we have obtained the solution of the equation (7.7.17)

$$X_t = X_0 e^{\alpha W_t - \frac{\alpha^2}{2}t}.$$

Example 7.7.1 *Use the method of variation of parameters to solve the equation*

$$dX_t = X_t W_t dW_t.$$

Dividing by X_t and converting the differential equation into the equivalent integral form, we get

$$\int \frac{1}{X_t}\,dX_t = \int W_t\,dW_t.$$

The right side is a well-known stochastic integral given by

$$\int W_t\,dW_t = \frac{W_t^2}{2} - \frac{t}{2} + C.$$

The left side will be integrated "blindly" according to the rules of elementary Calculus

$$\int \frac{1}{X_t}\,dX_t = \ln X_t + C.$$

Equating the last two relations and solving for X_t we obtain the "pseudo-solution"

$$X_t = e^{\frac{W_t^2}{2} - \frac{t}{2} + c},$$

with c constant. In order to get a correct solution, we let c to depend on t and W_t. We shall assume that $c(t, W_t) = a(t) + b(W_t)$, so we are looking for a solution of the form

$$X_t = e^{\frac{W_t^2}{2} - \frac{t}{2} + a(t) + b(W_t)}.$$

Applying Ito's formula, we have

$$dX_t \;=\; X_t\big[-\frac{1}{2} + a'(t) + \frac{1}{2}(1 + b''(W_t))\big]dt + X_t\big(W_t + b'(W_t)\big)dW_t.$$

Subtracting the initial equation $dX_t = X_t W_t dW_t$ yields

$$0 \;=\; X_t\big(a'(t) + \frac{1}{2}b''(W_t)\big)dt + X_t b'(W_t)dW_t.$$

This equation is satisfied if we are able to choose the functions $a(t)$ and $b(W_t)$ such that the coefficients of dt and dW_t vanish

$$b'(W_t) = 0, \qquad a'(t) + \frac{1}{2}b''(W_t) = 0.$$

From the first equation b must be a constant. Substituting into the second equation it follows that a is also a constant. It turns out that the aforementioned "pseudo-solution" is in fact a solution. The constant $c = a + b$ is obtained letting $t = 0$. Hence the solution is given by

$$X_t = X_0 e^{\frac{W_t^2}{2} - \frac{t}{2}}.$$

Example 7.7.2 *Use the method of variation of parameters to solve the stochastic differential equation*

$$dX_t = \mu X_t dt + \sigma X_t dW_t,$$

with μ and σ constants.

After dividing by X_t we bring the equation into the equivalent integral form

$$\int \frac{dX_t}{X_t} = \int \mu\, dt + \int \sigma\, dW_t.$$

Integrate on the left "blindly" and get

$$\ln X_t = \mu t + \sigma W_t + c,$$

where c is an integration constant. We arrive at the following "pseudo-solution"

$$X_t = e^{\mu t + \sigma W_t + c}.$$

Assume the constant c is replaced by a function $c(t)$, so we are looking for a solution of the form

$$X_t = e^{\mu t + \sigma W_t + c(t)}. \tag{7.7.19}$$

Apply Ito's formula and get

$$dX_t \quad = \quad X_t\big(\mu + c'(t) + \frac{\sigma^2}{2}\big)dt + \sigma X_t dW_t.$$

Subtracting the initial equation yields

$$\big(c'(t) + \frac{\sigma^2}{2}\big)dt = 0,$$

which is satisfied for $c'(t) = -\frac{\sigma^2}{2}$, with the solution $c(t) = -\frac{\sigma^2}{2}t + k$, $k \in \mathbb{R}$. Substituting into (7.7.19) yields the solution

$$X_t = e^{\mu t + \sigma W_t - \frac{\sigma^2}{2}t + k} = e^{(\mu - \frac{\sigma^2}{2})t + \sigma W_t + k} = X_0 e^{(\mu - \frac{\sigma^2}{2})t + \sigma W_t}.$$

7.8 Integrating Factors

The method of integrating factors can be applied to a class of stochastic differential equations of the type

$$dX_t = f(t, X_t)dt + g(t)X_t dW_t, \tag{7.8.20}$$

where f and g are continuous deterministic functions. The integrating factor is given by

$$\rho_t = e^{-\int_0^t g(s)\, dW_s + \frac{1}{2}\int_0^t g^2(s)\, ds}.$$

The equation can be brought into the following exact form

$$d(\rho_t X_t) = \rho_t f(t, X_t)dt.$$

Substituting $Y_t = \rho_t X_t$, we obtain that Y_t satisfies the deterministic differential equation

$$dY_t = \rho_t f(t, Y_t/\rho_t)dt,$$

which can be solved by either integration or as an exact equation. We shall exemplify the method of integrating factors with a few examples.

Example 7.8.1 *Solve the stochastic differential equation*

$$dX_t = r dt + \alpha X_t dW_t, \tag{7.8.21}$$

with r and α constants.

The integrating factor is given by $\rho_t = e^{\frac{1}{2}\alpha^2 t - \alpha W_t}$. Using Ito's formula, we can easily check that

$$d\rho_t = \rho_t(\alpha^2 dt - \alpha dW_t).$$

Using $dt^2 = dt\, dW_t = 0$, $(dW_t)^2 = dt$ we obtain

$$dX_t d\rho_t = -\alpha^2 \rho_t X_t dt.$$

Multiplying by ρ_t, the initial equation becomes

$$\rho_t dX_t - \alpha\rho_t X_t dW_t = r\rho_t dt,$$

and adding and subtracting $\alpha^2 \rho_t X_t dt$ from the left side yields

$$\rho_t dX_t - \alpha\rho_t X_t dW_t + \alpha^2\rho_t X_t dt - \alpha^2\rho_t X_t dt = r\rho_t dt.$$

This can be written as

$$\rho_t dX_t + X_t d\rho_t + d\rho_t dX_t = r\rho_t dt,$$

which, with the virtue of the product rule, becomes

$$d(\rho_t X_t) = r\rho_t dt.$$

Integrating yields

$$\rho_t X_t = \rho_0 X_0 + r\int_0^t \rho_s\, ds$$

and hence the solution is

$$\begin{aligned} X_t &= \frac{1}{\rho_t}X_0 + \frac{r}{\rho_t}\int_0^t \rho_s\, ds \\ &= X_0 e^{\alpha W_t - \frac{1}{2}\alpha^2 t} + r\int_0^t e^{-\frac{1}{2}\alpha^2(t-s)+\alpha(W_t - W_s)}\, ds. \end{aligned}$$

Exercise 7.8.1 *Let α be a constant. Solve the following stochastic differential equations by the method of integrating factors*

(a) $dX_t = \alpha X_t dW_t$;

(b) $dX_t = X_t dt + \alpha X_t dW_t$;

(c) $dX_t = \dfrac{1}{X_t}dt + \alpha X_t dW_t$, $X_0 > 0$.

Exercise 7.8.2 *Let X_t be the solution of the stochastic equation $dX_t = \sigma X_t\, dW_t$, with σ constant. Let $A_t = \frac{1}{t}\int_0^t X_s\, dW_s$ be the stochastic average of X_t. Find the stochastic equation satisfied by A_t, the mean and variance of A_t.*

7.9 Existence and Uniqueness

An exploding solution
Consider the non-linear stochastic differential equation

$$dX_t = X_t^3 dt + X_t^2\, dW_t, \qquad X_0 = 1/a. \tag{7.9.22}$$

We shall look for a solution of the type $X_t = f(W_t)$. Ito's formula yields

$$dX_t \;=\; f'(W_t)dW_t + \frac{1}{2}f''(W_t)dt.$$

Equating the coefficients of dt and dW_t in the last two equations yields

$$f'(W_t) \;=\; X_t^2 \Longrightarrow f'(W_t) = f(W_t)^2 \tag{7.9.23}$$

$$\frac{1}{2}f''(W_t) \;=\; X_t^3 \Longrightarrow f''(W_t) = 2f(W_t)^3 \tag{7.9.24}$$

We note that equation (7.9.23) implies (7.9.24) by differentiation. So it suffices to solve only the ordinary differential equation

$$f'(x) = f(x)^2, \qquad f(0) = 1/a.$$

Separating and integrating we have

$$\int \frac{df}{f(x)^2} = \int ds \Longrightarrow f(x) = \frac{1}{a-x}.$$

Hence a solution of equation (7.9.22) is

$$X_t = \frac{1}{a - W_t}.$$

Let T_a be the first time the Brownian motion W_t hits a. Then the process X_t is defined only for $0 \le t < T_a$. T_a is a random variable with $P(T_a < \infty) = 1$ and $E[T_a] = \infty$, see section 3.3.

The following theorem is the analog of Picard's uniqueness result from ordinary differential equations:

Theorem 7.9.1 (Existence and Uniqueness) *Consider the stochastic differential equation*

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t, \qquad X_0 = c$$

where c is a constant and b and σ are continuous functions on $[0,T] \times \mathbb{R}$ satisfying

1. $|b(t,x)| + |\sigma(t,x)| \le C(1 + |x|); \qquad x \in R, t \in [0,T]$
2. $|b(t,x) - b(t,y)| + |\sigma(t,x) - \sigma(t,y)| \le K|x-y|, \qquad x, y \in R, t \in [0,T]$

with C, K positive constants. There there is a unique solution process X_t that is continuous and satisfies

$$E\Big[\int_0^T X_t^2\, dt\Big] < \infty.$$

The first condition says that the drift and volatility increase no faster than a linear function in x. The second conditions states that the functions are Lipschitz in the second argument.

Example 7.9.1 *Consider the stochastic differential equation*

$$dX_t = \Big(\sqrt{1 + X_t^2} + \frac{1}{2}X_t\Big)dt + \sqrt{1 + X_t^2}\, dW_t, \quad X_0 = x_0.$$

(a) *Solve the equation;*

(b) *Show that there a unique solution.*

Chapter 8

Applications of Brownian Motion

8.1 The Generator of an Ito Diffusion

Let $(X_t)_{t\geq 0}$ be a stochastic process with $X_0 = x_0$. In this section we shall deal with the operator associated with X_t. This is an operator which describes infinitesimally the rate of change of a function which depends smoothly on X_t.

More precisely, the *generator* of the stochastic process X_t is the second order partial differential operator A defined by

$$Af(x) = \lim_{t \searrow 0} \frac{E[f(X_t)] - f(x)}{t},$$

for any smooth function (at least of class C^2) with compact support $f : \mathbb{R}^n \to \mathbb{R}$. Here E stands for the expectation operator taken at $t = 0$, i.e.,

$$E[f(X_t)] = \int_{\mathbb{R}^n} f(y) p_t(x, y)\, dy,$$

where $p_t(x, y)$ is the transition density of X_t.

In the following we shall find the generator associated with the Ito diffusion

$$dX_t = b(X_t)dt + \sigma(X_t)dW(t), \qquad t \geq 0, X_0 = x, \tag{8.1.1}$$

where $W(t) = \big(W_1(t), \ldots, W_m(t)\big)$ is an m-dimensional Brownian motion, $b : \mathbb{R}^n \to \mathbb{R}^n$ and $\sigma : \mathbb{R}^n \to \mathbb{R}^{n\times m}$ are measurable functions.

The main tool used in deriving the formula for the generator A is Ito's formula in several variables. If let $F_t = f(X_t)$, then using Ito's formula we have

$$dF_t = \sum_i \frac{\partial f}{\partial x_i}(X_t)\, dX_t^i + \frac{1}{2} \sum_{i,j} \frac{\partial^2 f}{\partial x_i \partial x_j}(X_t)\, dX_t^i\, dX_t^j, \tag{8.1.2}$$

where $X_t = (X_t^1, \cdots, X_t^n)$ satisfies the Ito diffusion (8.1.1) on components, i.e.,

$$\begin{aligned} dX_t^i &= b_i(X_t)dt + [\sigma(X_t)\, dW(t)]_i \\ &= b_i(X_t)dt + \sum_k \sigma_{ik}\, dW_k(t). \end{aligned} \tag{8.1.3}$$

Using the stochastic relations $dt^2 = dt\, dW_k(t) = 0$ and $dW_k(t)\, dW_r(t) = \delta_{kr} dt$, a computation provides

$$\begin{aligned} dX_t^i\, dX_t^j &= \Big(b_i dt + \sum_k \sigma_{ik} dW_k(t)\Big)\Big(b_j dt + \sum_k \sigma_{jk} dW_k(t)\Big) \\ &= \Big(\sum_k \sigma_{ik} dW_k(t)\Big)\Big(\sum_r \sigma_{jr} dW_r(t)\Big) \\ &= \sum_{k,r} \sigma_{ik}\sigma_{jr}\, dW_k(t) dW_r(t) = \sum_k \sigma_{ik}\sigma_{jk}\, dt \\ &= (\sigma\sigma^T)_{ij}\, dt. \end{aligned}$$

Therefore

$$dX_t^i\, dX_t^j = (\sigma\sigma^T)_{ij}\, dt. \tag{8.1.4}$$

Substitute (8.1.3) and (8.1.4) into (8.1.2) yields

$$\begin{aligned} dF_t &= \left[\frac{1}{2}\sum_{i,j} \frac{\partial^2 f}{\partial x_i \partial x_j}(X_t)(\sigma\sigma^T)_{ij} + \sum_i b_i(X_t)\frac{\partial f}{\partial x_i}(X_t)\right] dt \\ &\quad + \sum_{i,k} \frac{\partial f}{\partial x_i}(X_t)\sigma_{ik}(X_t)\, dW_k(t). \end{aligned}$$

Integrate and obtain

$$\begin{aligned} F_t &= F_0 + \int_0^t \left[\frac{1}{2}\sum_{i,j} \frac{\partial^2 f}{\partial x_i \partial x_j}(\sigma\sigma^T)_{ij} + \sum_i b_i \frac{\partial f}{\partial x_i}\right](X_s)\, ds \\ &\quad + \sum_k \int_0^t \sum_i \sigma_{ik} \frac{\partial f}{\partial x_i}(X_s)\, dW_k(s). \end{aligned}$$

Since $F_0 = f(X_0) = f(x)$ and $E(f(x)) = f(x)$, applying the expectation operator in the previous relation we obtain

$$E(F_t) = f(x) + E\left[\int_0^t \Big(\frac{1}{2}\sum_{i,j}(\sigma\sigma^T)_{ij}\frac{\partial^2 f}{\partial x_i \partial x_j} + \sum_i b_i \frac{\partial f}{\partial x_i}\Big)(X_s)\, ds\right]. \tag{8.1.5}$$

Using the commutation between the operator E and the integral $\int_0^t$ yields

$$\lim_{t\searrow 0}\frac{E(F_t)-f(x)}{t}=\frac{1}{2}\sum_{i,j}(\sigma\sigma^T)_{ij}\frac{\partial^2 f(x)}{\partial x_i\partial x_j}+\sum_k b_k\frac{\partial f(x)}{\partial x_k}.$$

We conclude the previous computations with the following result.

Theorem 8.1.1 *The generator of the Ito diffusion (8.1.1) is given by*

$$A=\frac{1}{2}\sum_{i,j}(\sigma\sigma^T)_{ij}\frac{\partial^2}{\partial x_i\partial x_j}+\sum_k b_k\frac{\partial}{\partial x_k}. \tag{8.1.6}$$

The matrix σ is called *dispersion* and the product $\sigma\sigma^T$ is called *diffusion* matrix. These names are related with their physical significance Substituting (8.1.6) in (8.1.5) we obtain the following formula

$$E[f(X_t)]=f(x)+E\Big[\int_0^t Af(X_s)\,ds\Big], \tag{8.1.7}$$

for any $f\in C_0^2(\mathbb{R}^n)$.

Exercise 8.1.2 *Find the generator operator associated with the n-dimensional Brownian motion.*

Exercise 8.1.3 *Find the Ito diffusion corresponding to the generator $Af(x)=f''(x)+f'(x)$.*

Exercise 8.1.4 *Let $\Delta_G=\frac{1}{2}(\partial_{x_1}^2+x_1^2\partial_{x_2}^2)$ be the Grushin's operator.*

(*a*) *Find the diffusion process associated with the generator Δ_G.*

(*b*) *Find the diffusion and dispersion matrices and show that they are degenerate.*

Exercise 8.1.5 *Let X_t and Y_t be two one-dimensional independent Ito diffusions with infinitesimal generators A_X and A_Y. Let $Z_t=(X_t,Y_t)$ with the infinitesimal generator A_Z. Show that $A_Z=A_X+A_Y$.*

Exercise 8.1.6 *Let X_t be an Ito diffusions with infinitesimal generator A_X. Consider the process $Y_t=(t,X_t)$. Show that the infinitesimal generator of Y_t is given by $A_Y=\partial_t+A_X$.*

8.2 Dynkin's Formula

Formula (8.1.7) holds under more general conditions, when t is a stopping time. First we need the following result, which deals with a continuity-type property in the upper limit of a Ito integral.

Lemma 8.2.1 *Let g be a bounded measurable function and τ be a stopping time for X_t with $E[\tau] < \infty$. Then*

$$\lim_{k\to\infty} E\Big[\int_0^{\tau\wedge k} g(X_s)\,dW_s\Big] = E\Big[\int_0^{\tau} g(X_s)\,dW_s\Big]; \tag{8.2.8}$$

$$\lim_{k\to\infty} E\Big[\int_0^{\tau\wedge k} g(X_s)\,ds\Big] = E\Big[\int_0^{\tau} g(X_s)\,ds\Big]. \tag{8.2.9}$$

Proof: Let $|g| < K$. Using the properties of Ito integrals, we have

$$\begin{aligned} & E\Big[\Big(\int_0^{\tau} g(X_s)\,dW_s - \int_0^{\tau\wedge k} g(X_t)\,dW_s\Big)^2\Big] = E\Big[\Big(\int_{\tau\wedge k}^{\tau} g(X_s)\,dW_s\Big)^2\Big] \\ = \;& E\Big[\int_{\tau\wedge k}^{\tau} g^2(X_s)\,ds\Big] \le K^2 E[\tau - \tau\wedge k] \to 0, \qquad k\to\infty. \end{aligned}$$

Since $E[X^2] \le E[X]^2$, it follows that

$$E\Big[\int_0^{\tau} g(X_s)\,dW_s - \int_0^{\tau\wedge k} g(X_t)\,dW_s\Big] \to 0, \qquad k\to\infty,$$

which is equivalent with relation (8.2.8).

The second relation can be proved similar and is left as an exercise for the reader.

■

Exercise 8.2.2 *Assume the hypothesis of the previous lemma. Let $\mathbf{1}_{\{s<\tau\}}$ be the characteristic function of the interval $(-\infty, \tau)$*

$$\mathbf{1}_{\{s<\tau\}}(u) = \begin{cases} 1, & \textit{if}\, u < \tau \\ 0, & \textit{otherwise.} \end{cases}$$

Show that

(a) $\displaystyle\int_0^{\tau\wedge k} g(X_s)\,dW_s = \int_0^{k} \mathbf{1}_{\{s<\tau\}}\, g(X_s)\,dW_s,$

(b) $\displaystyle\int_0^{\tau\wedge k} g(X_s)\,ds = \int_0^{k} \mathbf{1}_{\{s<\tau\}}\, g(X_s)\,ds.$

Theorem 8.2.3 (Dynkin's formula) *Let $f \in C_0^2(\mathbb{R}^n)$, and X_t be an Ito diffusion starting at x. If τ is a stopping time with $E[\tau] < \infty$, then*

$$E[f(X_\tau)] = f(x) + E\Big[\int_0^\tau Af(X_s)\,ds\Big], \tag{8.2.10}$$

where A is the infinitesimal generator of X_t.

Proof: Replace t by k and f by $\mathbf{1}_{\{s<\tau\}}f$ in (8.1.7) and obtain

$$E[1_{s<\tau}f(X_k)] = \mathbf{1}_{\{s<\tau\}}f(x) + E\Big[\int_0^k A(\mathbf{1}_{\{s<\tau\}}f)(X_s)\,ds\Big],$$

which can be written as

$$\begin{aligned} E[f(X_{k\wedge\tau})] &= \mathbf{1}_{\{s<\tau\}}f(x) + E\Big[\int_0^k \mathbf{1}_{\{s<\tau\}}(s)A(f)(X_s)\,ds\Big] \\ &= \mathbf{1}_{\{s<\tau\}}f(x) + E\Big[\int_0^{k\wedge\tau} A(f)(X_s)\,ds\Big]. \end{aligned} \tag{8.2.11}$$

Since by Lemma 8.2.1 (b)

$$E[f(X_{k\wedge\tau})] \to E[f(X_\tau)], \qquad k \to \infty$$

$$E\Big[\int_0^{k\wedge\tau} A(f)(X_s)\,ds\Big] \to E\Big[\int_0^\tau A(f)(X_s)\,ds\Big], \qquad k \to \infty,$$

using Exercise (8.2.2) relation (8.2.11) yields (8.2.10). ■

Exercise 8.2.4 *Write Dynkin's formula for the case of a function $f(t, X_t)$. Use Exercise 8.1.6.*

In the following sections we shall present a few important results of stochastic calculus that can be obtained as direct consequences of Dynkin's formula.

8.3 Kolmogorov's Backward Equation

For any function $f \in C_0^2(\mathbb{R}^n)$ let $v(t,x) = E[f(X_t)]$, given that $X_0 = x$. As usual, E denotes the expectation at time $t = 0$. Then $v(0,x) = f(x)$, and differentiating in Dynkin's formula (8.1.7)

$$v(t,x) = f(x) + \int_0^t E[Af(X_s)]\,ds$$

yields

$$\frac{\partial v}{\partial t} = E[Af(X_t)] = AE[f(X_t)] = Av(t,x).$$

We arrived at the following result.

Theorem 8.3.1 (Kolmogorov's backward equation) *For any $f \in C_0^2(\mathbb{R}^n)$ the function $v(t,x) = E[f(X_t)]$ satisfies the following Cauchy's problem*

$$\begin{aligned} \frac{\partial v}{\partial t} &= Av, \qquad t > 0 \\ v(0,x) &= f(x), \end{aligned}$$

where A denotes the generator of the Ito's diffusion (8.1.1).

8.4 Exit Time from an Interval

Let $X_t = x_0 + W_t$ e a one-dimensional Brownian motion starting at x_0. Consider the exit time of the process X_t from the strip (a, b)

$$\tau = \inf\{t > 0; X_t \notin (a, b)\}.$$

Assuming $E[\tau] < 0$, applying Dynkin's formula yields

$$E\Big[f(X_\tau)\Big] = f(x_0) + E\Big[\int_0^\tau \frac{1}{2}\frac{d^2}{dx^2} f(X_s)\,ds\Big]. \tag{8.4.12}$$

Choosing $f(x) = x$ in (8.4.12) we obtain

$$E[X_\tau] = x_0 \tag{8.4.13}$$

Exercise 8.4.1 *Prove relation (8.4.13) using the Optional Stopping Theorem for the martingale X_t.*

Let $p_a = P(X_\tau = a)$ and $p_b = P(X_\tau = b)$ be the exit probabilities from the interval (a, b). Obviously, $p_a + p_b = 1$, since the probability that the Brownian motion will stay for ever inside the bounded interval is zero. Using the expectation definition, relation (8.4.13) yields

$$ap_a + b(1 - p_a) = x_0.$$

Solving for p_a and p_b we get the following exit probabilities

$$p_a = \frac{b - x_0}{b - a} \tag{8.4.14}$$

$$p_b = 1 - p_a = \frac{x_0 - a}{b - a}. \tag{8.4.15}$$

It is worth noting that if $b \to \infty$ then $p_a \to 1$ and if $a \to -\infty$ then $p_b \to 1$. This can be stated by saying that a Brownian motion starting at x_0 reaches any level (below or above x_0) with probability 1.

Next we shall compute the mean of the exit time, $E[\tau]$. Choosing $f(x) = x^2$ in (8.4.12) yields

$$E[(X_\tau)^2] = x_0^2 + E[\tau].$$

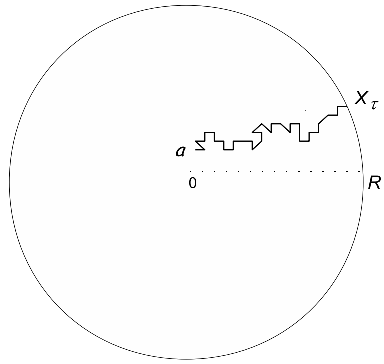


Figure 8.1: *The Brownian motion X_t in the ball $B(0,R)$.*

From the definition of the mean and formulas (8.4.14)-(8.4.15) we obtain

$$\begin{aligned} E[\tau] &= a^2 p_a + b^2 p_b - x_0^2 = a^2\frac{b-x_0}{b-a} + b^2\frac{x_0-a}{b-a} - x_0 \\ &= \frac{1}{b-a}\Big[ba^2 - ab^2 + x_0(b-a)(b+a)\Big] - x_0^2 \\ &= -ab + x_0(b+a) - x_0^2 \\ &= (b-x_0)(x_0-a). \end{aligned} \tag{8.4.16}$$

Exercise 8.4.2 (a) *Show that the equation $x^2-(b-a)x+E[\tau]=0$ cannot have complex roots;*

(b) *Prove that $E[\tau] \le \dfrac{(b-a)^2}{4}$;*

(c) *Find the point $x_0 \in (a,b)$ such that the expectation of the exit time, $E[\tau]$, is maximum.*

8.5 Transience and Recurrence of Brownian Motion

We shall consider first the expectation of the exit time from a ball. Then we shall extend it to an annulus and compute the transience probabilities.

1. Consider the process $X_t = a + W(t)$, where $W(t) = \big(W_1(t), \ldots, W_n(t)\big)$ is an n-dimensional Brownian motion, and $a = (a_1, \ldots, a_n) \in \mathbb{R}^n$ is a fixed vector, see Fig. 8.1. Let $R > 0$ such that $R > |a|$. Consider the exit time of the process X_t from the ball $B(0,R)$

$$\tau = \inf\{t > 0; |X_t| > R\}. \tag{8.5.17}$$

Assuming $E[\tau] < \infty$ and letting $f(x) = |x|^2 = x_1^2 + \cdots + x_n^2$ in Dynkin's formula

$$E\big[f(X_\tau)\big] = f(x) + E\Big[\int_0^\tau \frac{1}{2}\Delta f(X_s)\, ds\Big]$$

yields

$$R^2 = |a|^2 + E\Big[\int_0^\tau n\, ds\Big],$$

and hence

$$E[\tau] = \frac{R^2 - |a|^2}{n}. \tag{8.5.18}$$

In particular, if the Brownian motion starts from the center, i.e. $a = 0$, the expectation of the exit time is

$$E[\tau] = \frac{R^2}{n}.$$

(i) Since $R^2/2 > R^2/3$, the previous relation implies that it takes longer for a Brownian motion to exit a disk of radius R rather than a ball of the same radius.

(ii) The probability that a Brownian motion leaves the interval $(-R, R)$ is twice the probability that a 2-dimensional Brownian motion exits the disk $B(0, R)$.

Exercise 8.5.1 *Prove that $E[\tau] < \infty$, where τ is given by (8.5.17).*

Exercise 8.5.2 *Apply the Optional Stopping Theorem for the martingale $W_t = W_t^2 - t$ to show that $E[\tau] = R^2$, where*

$$\tau = \inf\{t > 0; |W_t| > R\}$$

is the first exit time of the Brownian motion from $(-R, R)$.

2. Let $b \in \mathbb{R}^n$ such that $b \notin B(0, R)$, i.e. $|b| > R$, and consider the annulus

$$A_k = \{x; R < |x| < kR\}$$

where $k > 0$ such that $b \in A_k$. Consider the process $X_t = b + W(t)$ and let

$$\tau_k = \inf\{t > 0; X_t \notin A_k\}$$

be the first exit time of X_t from the annulus A_k. Let $f : A_k \to \mathbb{R}$ be defined by

$$f(x) = \begin{cases} -\ln|x|, & \text{if } n = 2 \\ \frac{1}{|x|^{n-2}}, & \text{if } n > 2. \end{cases}$$

A straightforward computation shows that $\Delta f = 0$. Substituting into Dynkin's formula

$$E\big[f(X_{\tau_k})\big] = f(b) + E\Big[\int_0^{\tau_k} \Big(\frac{1}{2}\Delta f\Big)(X_s)\, ds\Big]$$

yields

$$E\big[f(X_{\tau_k})\big] = f(b). \tag{8.5.19}$$

This can be stated by saying that the value of f at a point b in the annulus is equal to the expected value of f at the first exit time of a Brownian motion starting at b.

Since $|X_{\tau_k}|$ is a random variable with two outcomes, we have

$$E\big[f(X_{\tau_k})\big] = p_k f(R) + q_k f(kR),$$

where $p_k = P(|X_{\tau_k}| = R)$, $q_k = P(|X_{X_{\tau_k}}|) = kR$ and $p_k + q_k = 1$. Substituting in (8.5.19) yields

$$p_k f(R) + q_k f(kR) = f(b). \tag{8.5.20}$$

There are the following two distinguished cases:

(i) If $n = 2$ we obtain

$$-p_k \ln R - q_k(\ln k + \ln R) = -\ln b.$$

Using $p_k = 1 - q_k$, solving for p_k yields

$$p_k = 1 - \frac{\ln(\frac{b}{R})}{\ln k}.$$

Hence

$$P(\tau < \infty) = \lim_{k\to\infty} p_k = 1,$$

where $\tau = \inf\{t > 0; |X_t| < R\}$ is the first time X_t hits the ball $B(0, R)$. Hence in $\mathbb{R}^2$ a Brownian motion hits with probability 1 any ball. This is stated equivalently by saying that the Brownian motion is *recurrent* in $\mathbb{R}^2$.

(ii) If $n > 2$ the equation (8.5.20) becomes

$$\frac{p_k}{R^{n-2}} + \frac{q_k}{k^{n-2}R^{n-2}} = \frac{1}{b^{n-2}}.$$

Taking the limit $k \to \infty$ yields

$$\lim_{k\to\infty} p_k = \Big(\frac{R}{b}\Big)^{n-2} < 1.$$

Then in $\mathbb{R}^n$, $n > 2$, a Brownian motion starting outside of a ball hits it with a probability less than 1. This is usually stated by saying that the Brownian motion is *transient*.

3. We shall recover the previous results using the n-dimensional Bessel process

$$\mathcal{R}_t = dist\big(0, W(t)\big) = \sqrt{W_1(t)^2 + \cdots + W_n(t)^2}.$$

Consider the process $Y_t = \alpha + \mathcal{R}_t$, with $0 \le \alpha < R$. The generator of Y_t is the Bessel operator of order n

$$A = \frac{1}{2}\frac{d^2}{dx^2} + \frac{n-1}{2x}\frac{d}{dx},$$

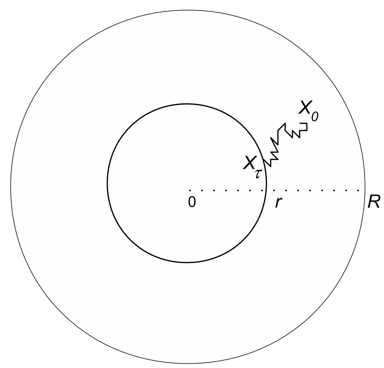


Figure 8.2: *The Brownian motion X_t in the annulus $A_{r,R}$.*

see section 2.7. Consider the exit time

$$\tau = \{t > 0; Y_t > R\}.$$

Applying Dynkin's formula

$$E\big[f(Y_\tau)\big] = f(Y_0) + E\Big[\int_0^\tau (Af)(Y_s)\,ds\Big]$$

for $f(x) = x^2$ yields $R^2 = \alpha^2 + E\big[\int_0^\tau n ds\big]$. This leads to

$$E[\tau] = \frac{R^2 - \alpha^2}{n}.$$

which recovers (8.5.18) with $\alpha = |a|$.

In the following assume $n \geq 3$ and consider the annulus

$$A_{r,R} = \{x \in \mathbb{R}^n; r < |x| < R\}.$$

Consider the stopping time $\tau = \inf\{t > 0; X_t \notin A_{r,R}\} = \inf\{t > 0; Y_t \notin (r, R)\}$, where $|Y_0| = \alpha \in (r, R)$. Applying Dynkin's formula for $f(x) = x^{2-n}$ yields $E\big[f(Y_\tau\big] = f(\alpha)$. This can be written as

$$p_r r^{2-n} + p_R R^{2-n} = \alpha^{2-n},$$

where

$$p_r = P(|X_t| = r), \qquad p_R = P(|X_t| = R), \qquad p_r + p_R = 1.$$

Solving for p_r and p_R yields

$$p_r = \frac{\left(\frac{R}{\alpha}\right)^{n-2} - 1}{\left(\frac{R}{r}\right)^{n-2} - 1}, \qquad p_R = \frac{\left(\frac{r}{\alpha}\right)^{n-2} - 1}{\left(\frac{r}{R}\right)^{n-2} - 1}.$$

The transience probability is obtained by taking the limit to infinity

$$p_r = \lim_{R\to\infty} p_{r,R} = \lim_{R\to\infty} \frac{\alpha^{2-n} R^{n-2} - 1}{r^{2-n} R^{n-2} - 1} = \left(\frac{r}{\alpha}\right)^{n-2},$$

where p_r is the probability that a Brownian motion starting outside the ball of radius r will hit the ball, see Fig. 8.2.

Exercise 8.5.3 *Solve the equation* $\frac{1}{2}f''(x) + \frac{n-1}{2x}f'(x) = 0$ *by looking for a solution of monomial type* $f(x) = x^k$.

8.6 Application to Parabolic Equations

This section deals with solving first and second order parabolic equations using the integral of the cost function along a certain *characteristic* solution. The first order equations are related to predictable characteristic curves, while the second order equations depend on stochastic characteristic curves.

1. Predictable characteristics Let $\varphi(s)$ be the solution of the following one-dimensional ODE

$$\begin{aligned} \frac{dX(s)}{ds} &= a\big(s, X(s)\big), \qquad t \le s \le T \\ X(t) &= x, \end{aligned}$$

and define the cumulative cost between t and T along the solution φ

$$u(t,x) = \int_t^T c\big(s, \varphi(s)\big)\, ds, \tag{8.6.21}$$

where c denotes a continuous cost function. Differentiate both sides with respect to t

$$\begin{aligned} \frac{\partial}{\partial t} u\big(t, \varphi(t)\big) &= \frac{\partial}{\partial t} \int_t^T c\big(s, \varphi(s)\big)\, ds \\ \partial_t u + \partial_x u\, \varphi'(t) &= -c\big(t, \varphi(t)\big). \end{aligned}$$

Hence (8.6.21) is a solution of the following final value problem

$$\begin{aligned} \partial_t u(t,x) + a(t,x)\partial_x u(t,x) &= -c(t,x) \\ u(T,x) &= 0. \end{aligned}$$

It is worth mentioning that this is a variant of the method of characteristics.[1] The curve given by the solution $\varphi(s)$ is called a characteristic curve.

[1] This is a well known method of solving linear partial differential equations.

Exercise 8.6.1 *Using the previous method solve the following final boundary problems:*
(*a*)

$$\begin{aligned} \partial_t u + x\partial_x u &= -x \\ u(T,x) &= 0. \end{aligned}$$

(*b*)

$$\begin{aligned} \partial_t u + tx\partial_x u &= \ln x, \qquad x > 0 \\ u(T,x) &= 0. \end{aligned}$$

2. Stochastic characteristics Consider the diffusion

$$\begin{aligned} dX_s &= a(s,X_s)ds + b(s,X_s)dW_s, \qquad t \le s \le T \\ X_t &= x, \end{aligned}$$

and define the stochastic cumulative cost function

$$u(t,X_t) = \int_t^T c(s,X_s)\,ds, \tag{8.6.22}$$

with the conditional expectation

$$\begin{aligned} u(t,x) &= E\Big[u(t,X_t)|X_t = x\Big] \\ &= E\Big[\int_t^T c(s,X_s)\,ds|X_t = x\Big]. \end{aligned}$$

Taking increments in both sides of (8.6.22) yields

$$du(t,X_t) = d\int_t^T c(s,X_s)\,ds.$$

Applying Ito's formula on one side and the Fundamental Theorem of Calculus on the other, we obtain

$$\partial_t u(t,x)dt + \partial_x u(t,X_t)dX_t + \frac{1}{2}\partial_x^2 u(t,t,X_t)dX_t^2 = -c(t,X_t)dt.$$

Taking the expectation $E[\,\cdot\,|X_t = x]$ on both sides yields

$$\partial_t u(t,x)dt + \partial_x u(t,x)a(t,x)dt + \frac{1}{2}\partial_x^2 u(t,x)b^2(t,x)dt = -c(t,x)dt.$$

Hence, the expected cost

$$u(t,x) = E\Big[\int_t^T c(s,X_s)\,ds|X_t = x\Big]$$

is a solution of the following second order parabolic equation

$$\begin{aligned} \partial_t u + a(t,x)\partial_x u + \frac{1}{2}b^2(t,x)\partial_x^2 u(t,x) &= -c(t,x) \\ u(T,x) &= 0. \end{aligned}$$

This represents the probabilistic interpretation of the solution of a parabolic equation.

Exercise 8.6.2 *Solve the following final boundary problems:*
(*a*)

$$\begin{aligned} \partial_t u + \partial_x u + \frac{1}{2}\partial_x^2 u &= -x \\ u(T,x) &= 0. \end{aligned}$$

(*b*)

$$\begin{aligned} \partial_t u + \partial_x u + \frac{1}{2}\partial_x^2 u &= e^x, \\ u(T,x) &= 0. \end{aligned}$$

(*c*)

$$\begin{aligned} \partial_t u + \mu x\partial_x u + \frac{1}{2}\sigma^2 x^2\partial_x^2 u &= -x, \\ u(T,x) &= 0. \end{aligned}$$

Chapter 9

Martingales and Girsanov's Theorem

9.1 Examples of Martingales

In this section we shall use the knowledge acquired in previous chapters to present a few important examples of martingales. These will be useful in the proof of Girsanov's theorem in the next section.

We start by recalling that a process X_t, $0 \le t \le T$, is an $\mathcal{F}_t$-*martingale* if

1. $E[|X_t|] < \infty$ (X_t integrable for each t);
2. X_t is $\mathcal{F}_t$-predictable;
3. the forecast of future values is the last observation: $E[X_t|\mathcal{F}_s] = X_s$, $\forall s < t$.

We shall present three important examples of martingales and some of their particular cases.

Example 9.1.1 *If $v(s)$ is a continuous function on $[0, T]$, then*

$$\boxed{X_t = \int_0^t v(s)\, dW_s}$$

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

Taking out the predictable part

$$\begin{aligned} E[X_t|\mathcal{F}_s] &= E\Big[\int_0^s v(\tau)\, dW_\tau + \int_s^t v(\tau)\, dW_\tau \Big| \mathcal{F}_s\Big] \\ &= X_s + E\Big[\int_s^t v(\tau)\, dW_\tau \Big| \mathcal{F}_s\Big] = X_s, \end{aligned}$$

where we used that $\int_s^t v(\tau)\,dW_\tau$ is independent of $\mathcal{F}_s$ and the conditional expectation equals the usual expectation

$$E\Big[\int_s^t v(\tau)\,dW_\tau\Big|\mathcal{F}_s\Big] = E\Big[\int_s^t v(\tau)\,dW_\tau\Big] = 0.$$

Example 9.1.2 *Let* $X_t = \displaystyle\int_0^t v(s)\,dW_s$ *be a process as in Example 9.1.1. Then*

$$\boxed{M_t = X_t^2 - \int_0^t v^2(s)\,ds}$$

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

The process X_t satisfies the stochastic equation $dX_t = v(t)dW_t$. By Ito's formula

$$d(X_t^2) = 2X_t dX_t + (dX_t)^2 = 2v(t)X_t dW_t + v^2(t)dt. \tag{9.1.1}$$

Integrating between s and t yields

$$X_t^2 - X_s^2 = 2\int_s^t X_\tau v(\tau)\,dW_\tau + \int_s^t v^2(\tau)\,d\tau.$$

Then separating the predictable from the unpredictable we have

$$\begin{aligned}
E[M_t|\mathcal{F}_s] &= E\Big[X_t^2 - \int_0^t v^2(\tau)\,d\tau\Big|\mathcal{F}_s\Big] \\
&= E\Big[X_t^2 - X_s^2 - \int_s^t v^2(\tau)\,d\tau + X_s^2 - \int_0^s v^2(\tau)\,d\tau\Big|\mathcal{F}_s\Big] \\
&= X_s^2 - \int_0^s v^2(\tau)\,d\tau + E\Big[X_t^2 - X_s^2 - \int_s^t v^2(\tau)\,d\tau\Big|\mathcal{F}_s\Big] \\
&= M_s + 2E\Big[\int_s^t X_\tau v(\tau)\,dW_\tau\Big|\mathcal{F}_s\Big] = M_s,
\end{aligned}$$

where we used relation (9.1.1) and that $\displaystyle\int_s^t X_\tau v(\tau)\,dW_\tau$ is totally unpredictable given the information set $\mathcal{F}_s$. In the following we shall mention a few particular cases.

1. If $v(s) = 1$, then $X_t = W_t$. In this case $M_t = W_t^2 - t$ is an $\mathcal{F}_t$-martingale.
2. If $v(s) = s$, then $X_t = \int_0^t s\,dW_s$, and hence

$$M_t = \Big(\int_0^t s\,dW_s\Big)^2 - \frac{t^3}{3}$$

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

Example 9.1.3 *Let $u : [0, T] \to \mathbb{R}$ be a continuous function. Then*

$$\boxed{M_t = e^{\int_0^t u(s)\, dW_s - \frac{1}{2}\int_0^t u^2(s)\, ds}}$$

is an $\mathcal{F}_t$-martingale for $0 \le t \le T$.

Consider the process $U_t = \displaystyle\int_0^t u(s)\, dW_s - \frac{1}{2}\int_0^t u^2(s)\, ds$. Then

$$\begin{aligned} dU_t &= u(t)dW_t - \frac{1}{2}u^2(t)dt \\ (dU_t)^2 &= u(t)dt. \end{aligned}$$

Then Ito's formula yields

$$\begin{aligned} dM_t &= d(e^{U_t}) = e^{U_t}dU_t + \frac{1}{2}e^{U_t}(dU_t)^2 \\ &= e^{U_t}\Big(u(t)dW_t - \frac{1}{2}u^2(t)dt + \frac{1}{2}u^2(t)dt\Big) \\ &= u(t)M_t dW_t. \end{aligned}$$

Integrating between s and t yields

$$M_t = M_s + \int_s^t u(\tau)M_\tau\, dW_\tau$$

Since $\displaystyle\int_s^t u(\tau)M_\tau\, dW_\tau$ is independent of $\mathcal{F}_s$, then

$$E\Big[\int_s^t u(\tau)M_\tau\, dW_\tau|\mathcal{F}_s\Big] = E\big[\int_s^t u(\tau)M_\tau\, dW_\tau\big] = 0,$$

and hence

$$E[M_t|\mathcal{F}_s] = E[M_s + \int_s^t u(\tau)M_\tau\, dW_\tau|\mathcal{F}_s] = M_s.$$

Remark 9.1.4 *The condition that $u(s)$ is continuous on $[0, T]$ can be relaxed by asking only*

$$u \in L^2[0, T] = \{u : [0, T] \to \mathbb{R};\, measurable\ and\ \int_0^t |u(s)|^2\, ds < \infty\}.$$

It is worth noting that the conclusion still holds if the function $u(s)$ is replaced by a stochastic process $u(t, \omega)$ satisfying Novikov's condition

$$E\big[e^{\frac{1}{2}\int_0^T u^2(s,\omega)\, ds}\big] < \infty.$$

The previous process has a distinguished importance in the theory of martingales and it will be useful when proving the Girsanov theorem.

Definition 9.1.5 *Let $u \in L^2[0,T]$ be a deterministic function. Then the stochastic process*

$$M_t = e^{\int_0^t u(s)\,dW_s - \frac{1}{2}\int_0^t u^2(s)\,ds}$$

is called the exponential process induced by u.

Particular cases of exponential processes.

1. Let $u(s) = \sigma$, constant, then $M_t = e^{\sigma W_t - \frac{\sigma^2}{2}t}$ is an $\mathcal{F}_t$-martingale.
2. Let $u(s) = s$. Integrating in $d(tW_t) = tdW_t - W_t dt$ yields

$$\int_0^t s\,dW_s = tW_t - \int_0^t W_s\,ds.$$

Let $Z_t = \int_0^t W_s\,ds$ be the integrated Brownian motion. Then

$$\begin{aligned} M_t &= e^{\int_0^t d\,dW_s - \frac{1}{2}\int_0^t s^2\,ds} \\ &= e^{tW_t - \frac{t^3}{6} - Z_t} \end{aligned}$$

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

Example 9.1.6 *Let X_t be a solution of $dX_t = u(t)dt + dW_t$, with $u(s)$ a bounded function. Consider the exponential process*

$$M_t = e^{-\int_0^t u(s)\,dW_s - \frac{1}{2}\int_0^t u^2(s)\,ds}. \tag{9.1.2}$$

Then $Y_t = M_t X_t$ is an $\mathcal{F}_t$-martingale.

In Example 9.1.3 we obtained $dM_t = -u(t)M_t dW_t$. Then

$$dM_t\,dX_t = -u(t)M_t dt.$$

The product rule yields

$$\begin{aligned} dY_t &= M_t dX_t + X_t dM_t + dM_t\,dX_t \\ &= M_t\big(u(t)dt + dW_t\big) - X_t u(t) M_t dW_t - u(t)M_t dt \\ &= M_t\big(1 - u(t)X_t\big)dW_t. \end{aligned}$$

Integrating between s and t yields

$$Y_t = Y_s + \int_s^t M_\tau\big(1 - u(\tau)X_\tau\big)dW_\tau.$$

Since $\int_s^t M_\tau\big(1-u(\tau)X_\tau\big)dW_\tau$ is independent of $\mathcal{F}_s$,

$$E\Big[\int_s^t M_\tau\big(1-u(\tau)X_\tau\big)dW_\tau|\mathcal{F}_s\Big] = E\Big[\int_s^t M_\tau\big(1-u(\tau)X_\tau\big)dW_\tau\Big] = 0,$$

and hence

$$E[Y_t|\mathcal{F}_s] = Y_s.$$

Exercise 9.1.7 *Prove that* $(W_t+t)e^{-W_t-\frac{1}{2}t}$ *is an* $\mathcal{F}_t$*-martingale.*

Exercise 9.1.8 *Let* h *be a continuous function. Using the properties of the Wiener integral and log-normal random variables, show that*

$$E\Big[e^{\int_0^t h(s)\,dW_s}\Big] = e^{\frac{1}{2}\int_0^t h(s)^2\,ds}.$$

Exercise 9.1.9 *Let* M_t *be the exponential process (9.1.2). Use the previous exercise to show that for any* $t>0$

$$(a)\ E[M_t]=1 \qquad\qquad (b)\ E[M_t^2] = e^{\int_0^t u(s)^2\,ds}.$$

Exercise 9.1.10 *Let* $\mathcal{F}_t = \sigma\{W_u; u\le t\}$. *Show that the following processes are* $\mathcal{F}_t$*-martingales:*

(a) $e^{t/2}\cos W_t$;

(b) $e^{t/2}\sin W_t$.

Recall that the Laplacian of a twice differentiable function f is defined by $\Delta f(x) = \sum_{j=1}^n \partial_{x_j}^2 f$.

Example 9.1.11 *Consider the smooth function* $f:\mathbb{R}^n\to R$, *such that*

(i) $\Delta f = 0$;

(ii) $E\big[|f(W_t)|\big]<\infty$, $\forall t>0$ *and* $x\in R$.

Then the process $X_t = f(W_t)$ *is an* $\mathcal{F}_t$*-martingale.*

Proof: It follows from the more general Example 9.1.13.

Exercise 9.1.12 *Let* $W_1(t)$ *and* $W_2(t)$ *be two independent Brownian motions. Show that* $X_t = e^{W_1(t)}\cos W_2(t)$ *is a martingale.*

Example 9.1.13 *Let* $f:\mathbb{R}^n\to\mathbb{R}$ *be a smooth function such that*

(i) $E\big[|f(W_t)|\big]<\infty$;

(ii) $E\Big[\int_0^t|\Delta f(W_s)|\,ds\Big]<\infty$.

Then the process $X_t = f(W_t)-\frac{1}{2}\int_0^t\Delta f(W_s)\,ds$ *is a martingale.*

Proof: For $0 \le s < t$ we have

$$\begin{aligned} E\big[X_t|\mathcal{F}_s\big] &= E\big[f(W_t)|\mathcal{F}_s\big] - E\Big[\frac{1}{2}\int_0^t \Delta f(W_u)\,du|\mathcal{F}_s\Big] \\ &= E\big[f(W_t)|\mathcal{F}_s\big] - \frac{1}{2}\int_0^s \Delta f(W_u)\,du - \int_s^t E\Big[\frac{1}{2}\Delta f(W_u)|\mathcal{F}_s\Big]\,du \quad (9.1.3) \end{aligned}$$

Let $p(t,y,x)$ be the probability density function of W_t. Integrating by parts and using that p satisfies the Kolmogorov's backward equation, we have

$$\begin{aligned} E\Big[\frac{1}{2}\Delta f(W_u)|\mathcal{F}_s\Big] &= \frac{1}{2}\int p(u-s,W_s,x)\,\Delta f(x)\,dx = \frac{1}{2}\int \Delta_x p(u-s,W_s,x) f(x)\,dx \\ &= \int \frac{\partial}{\partial u} p(u-s,W_s,x) f(x)\,dx. \end{aligned}$$

Then, using the Fundamental Theorem of Calculus, we obtain

$$\begin{aligned} \int_s^t E\Big[\frac{1}{2}\Delta f(W_u)|\mathcal{F}_s\Big]\,du &= \int_s^t \Big(\frac{\partial}{\partial u}\int p(u-s,W_s,x)f(x)\,dx\Big)\,du \\ &= \int p(t-s,W_s,x)f(x)\,dx - \lim_{\epsilon\searrow 0}\int p(\epsilon,W_s,x)f(x)\,dx \\ &= E\big[f(W_t)|\mathcal{F}_s\big] - \int \delta(x=W_s)f(x)\,dx \\ &= E\big[f(W_t)|\mathcal{F}_s\big] - f(W_s). \end{aligned}$$

Substituting in (9.1.3) yields

$$\begin{aligned} E\big[X_t|\mathcal{F}_s\big] &= E\big[f(W_t)|\mathcal{F}_s\big] - \frac{1}{2}\int_0^s \Delta f(W_u)\,du - E\big[f(W_t)|\mathcal{F}_s\big] + f(W_s) \quad (9.1.4) \\ &= f(W_s) - \frac{1}{2}\int_0^s \Delta f(W_u)\,du \quad (9.1.5) \\ &= X_s. \quad (9.1.6) \end{aligned}$$

Hence X_t is an $\mathcal{F}_t$-martingale. ■

Exercise 9.1.14 *Use the Example 9.1.13 to show that the following processes are martingales:*

(*a*) $X_t = W_t^2 - t$;

(*b*) $X_t = W_t^3 - 3\int_0^t W_s\,ds$;

(*c*) $X_t = \frac{1}{n(n-1)}W_t^n - \frac{1}{2}\int_0^t W_s^{n-2}\,ds$;

(*d*) $X_t = e^{cW_t} - \frac{1}{2}c^2\int_0^t e^{W_s}\,ds$, *with* c *constant;*

(*e*) $X_t = \sin(cW_t) + \frac{1}{2}c^2\int_0^t \sin(cW_s)\,ds$, *with* c *constant.*

Exercise 9.1.15 *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function such that*

(i) $E\big[|f(W_t)|\big] < \infty$;

(ii) $\Delta f = \lambda f$, λ *constant.*

Show that the process $X_t = f(W_t) - \frac{\lambda}{2}\int_0^t f(W_s)\,ds$ *is a martingale.*

9.2 Girsanov's Theorem

In this section we shall present and prove a particular version of Girsanov's theorem, which will suffice for the purpose of later applications. The application of Girsanov's theorem to finance is to show that in the study of security markets the differences between the mean rates of return can be removed.

We shall recall first a few basic notions. Let $(\Omega, \mathcal{F}, P)$ be a probability space. When dealing with an $\mathcal{F}_t$-martingale on the aforementioned probability space, the filtration $\mathcal{F}_t$ is considered to be the sigma-algebra generated by the given Brownian motion W_t, i.e. $\mathcal{F}_t = \sigma\{W_u; 0 \le u \le s\}$. By default, a martingale is considered with respect to the probability measure P, in the sense that the expectations involve an integration with respect to P

$$E^P[X] = \int_\Omega X(\omega)dP(\omega).$$

We have not used the upper script until now since there was no doubt which probability measure was used. In this section we shall use also another probability measure given by

$$dQ = M_T dP,$$

where M_T is an exponential process. This means that $Q : \mathcal{F} \to \mathbb{R}$ is given by

$$Q(A) = \int_A dQ = \int_A M_T dP, \qquad \forall A \in \mathcal{F}.$$

Since $M_T > 0$, $M_0 = 1$, using the martingale property of M_t yields

$$\begin{aligned} Q(A) &> 0, \qquad A \neq \emptyset; \\ Q(\Omega) &= \int_\Omega M_T dP = E^P[M_T] = E^P[M_T|\mathcal{F}_0] = M_0 = 1, \end{aligned}$$

which shows that Q is a probability on $\mathcal{F}$, and hence $(\Omega, \mathcal{F}, Q)$ becomes a probability space. The following transformation of expectation from the probability measure Q to P will be useful in part II. If X is a random variable

$$\begin{aligned} E^Q[X] &= \int_\Omega X(\omega)\,dQ(\omega) = \int_\Omega X(\omega)M_T(\omega)\,dP(\omega) \\ &= E^P[XM_T]. \end{aligned}$$

The following result will play a central role in proving Girsanov's theorem:

Lemma 9.2.1 *Let X_t be the Ito process*

$$dX_t = u(t)dt + dW_t, \qquad X_0 = 0, \ 0 \leq t \leq T,$$

with $u(s)$ a bounded function. Consider the exponential process

$$M_t = e^{-\int_0^t u(s)\,dW_s - \frac{1}{2}\int_0^t u^2(s)\,ds}.$$

Then X_t is an $\mathcal{F}_t$-martingale with respect to the measure

$$dQ(\omega) = M_T(\omega)dP(\omega).$$

Proof: We need to prove that X_t is an $\mathcal{F}_t$-martingale with respect to Q, so it suffices to show the following three properties:

1. Integrability of X_t. This part usually follows from standard manipulations of norms estimations. We shall do it here in detail, but we shall omit it in other proofs. Integrating in the equation of X_t between 0 and t yields

$$X_t = \int_0^t u(s)\,ds + W_t. \tag{9.2.7}$$

We start with an estimation of the expectation with respect to P

$$\begin{aligned} E^P[X_t^2] &= E^P\Big[\Big(\int_0^t u(s)\,ds\Big)^2 + 2\int_0^t u(s)\,ds\,W_t + W_t^2\Big] \\ &= \Big(\int_0^t u(s)\,ds\Big)^2 + 2\int_0^t u(s)\,ds\,E^P[W_t] + E^P[W_t^2] \\ &= \Big(\int_0^t u(s)\,ds\Big)^2 + t < \infty, \qquad \forall 0 \leq t \leq T \end{aligned}$$

where the last inequality follows from the norm estimation

$$\begin{aligned} \int_0^t u(s)\,ds &\leq \int_0^t |u(s)|\,ds \leq \Big[t\int_0^t |u(s)|^2\,ds\Big]^{1/2} \\ &\leq \Big[t\int_0^T |u(s)|^2\,ds\Big]^{1/2} = T^{1/2}\|u\|_{L^2[0,T]}. \end{aligned}$$

Next we obtain an estimation with respect to Q

$$\begin{aligned} E^Q[|X_t|]^2 &= \Big(\int_\Omega |X_t| M_T\,dP\Big)^2 \leq \int_\Omega |X_t|^2\,dP \int_\Omega M_T^2\,dP \\ &= E^P[X_t^2]\,E^P[M_T^2] < \infty, \end{aligned}$$

since $E^P[X_t^2] < \infty$ and $E^P[M_T^2] = e^{\int_0^T u(s)^2\,ds} = e^{\|u\|^2_{L^2[0,T]}}$, see Exercise 9.1.9.

2. $\mathcal{F}_t$-predictability of X_t. This follows from equation (9.2.7) and the fact that W_t is $\mathcal{F}_t$-predictable.

3. Conditional expectation of X_t. From Examples 9.1.3 and 9.1.6 recall that for any $0 \le t \le T$

M_t is an $\mathcal{F}_t$-martingale with respect to probability measure P;

$X_t M_t$ is an $\mathcal{F}_t$-martingale with respect to probability measure P.

We need to verify that

$$E^Q[X_t|\mathcal{F}_s] = X_s, \qquad \forall s \le t.$$

which can be written as

$$\int_A X_t\, dQ = \int_A X_s\, dQ, \qquad \forall A \in \mathcal{F}_s.$$

Since $dQ = M_T dP$, the previous relation becomes

$$\int_A X_t M_T\, dP = \int_A X_s M_T\, dP, \qquad \forall A \in \mathcal{F}_s.$$

This can be written in terms of conditional expectation as

$$E^P[X_t M_T|\mathcal{F}_s] = E^P[X_s M_T|\mathcal{F}_s]. \tag{9.2.8}$$

We shall prove this identity by showing that both terms are equal to $X_s M_s$. Since X_s is $\mathcal{F}_s$-predictable and M_t is a martingale, the right side term becomes

$$E^P[X_s M_T|\mathcal{F}_s] = X_s E^P[M_T|\mathcal{F}_s] = X_s M_s, \qquad \forall s \le T.$$

Let $s < t$. Using the tower property (see Proposition 1.11.4, part 3), the left side term becomes

$$\begin{aligned} E^P[X_t M_T|\mathcal{F}_s] &= E^P\big[E^P[X_t M_T|\mathcal{F}_t]|\mathcal{F}_s\big] = E^P\big[X_t E^P[M_T|\mathcal{F}_t]|\mathcal{F}_s\big] \\ &= E^P\big[X_t M_t|\mathcal{F}_s\big] = X_s M_s, \end{aligned}$$

where we used that M_t and $X_t M_t$ are martingales and X_t is $\mathcal{F}_t$-predictable. Hence (9.2.8) holds and X_t is an $\mathcal{F}_t$-martingale with respect to the probability measure Q. ■

Lemma 9.2.2 *Consider the process*

$$X_t = \int_0^t u(s)\, ds + W_t, \qquad 0 \le t \le T,$$

with $u \in L^2[0,T]$ a deterministic function, and let $dQ = M_T dP$. Then

$$E^Q[X_t^2] = t.$$

Proof: Denote $U(t) = \int_0^t u(s)\,ds$. Then

$$\begin{aligned} E^Q[X_t^2] &= E^P[X_t^2 M_T] = E^P[U^2(t)M_T + 2U(t)W_t M_T + W_t^2 M_T] \\ &= U^2(t)E^P[M_T] + 2U(t)E^P[W_t M_T] + E^P[W_t^2 M_T]. \end{aligned} \tag{9.2.9}$$

From Exercise 9.1.9 (a) we have $E^P[M_T] = 1$. In order to compute $E^P[W_t M_T]$ we use the tower property and the martingale property of M_t

$$\begin{aligned} E^P[W_t M_T] &= E[E^P[W_t M_T|\mathcal{F}_t]] = E[W_t E^P[M_T|\mathcal{F}_t]] \\ &= E[W_t M_t]. \end{aligned} \tag{9.2.10}$$

Using the product rule

$$\begin{aligned} d(W_t M_t) &= M_t dW_t + W_t dM_t + dW_t dM_t \\ &= \big(M_t - u(t)M_t W_t\big)dW_t - u(t)M_t dt, \end{aligned}$$

where we used $dM_t = -u(t)M_t dW_t$. Integrating between 0 and t yields

$$W_t M_t = \int_0^t \big(M_s - u(s)M_s W_s\big)dW_s - \int_0^t u(s)M_s\,ds.$$

Taking the expectation and using the property of Ito integrals we have

$$E[W_t M_t] = -\int_0^t u(s)E[M_s]\,ds = -\int_0^t u(s)\,ds = -U(t). \tag{9.2.11}$$

Substituting into (9.2.10) yields

$$E^P[W_t M_T] = -U(t). \tag{9.2.12}$$

For computing $E^P[W_t^2 M_T]$ we proceed in a similar way

$$\begin{aligned} E^P[W_t^2 M_T] &= E[E^P[W_t^2 M_T|\mathcal{F}_t]] = E[W_t^2 E^P[M_T|\mathcal{F}_t]] \\ &= E[W_t^2 M_t]. \end{aligned} \tag{9.2.13}$$

Using the product rule yields

$$\begin{aligned} d(W_t^2 M_t) &= M_t d(W_t^2) + W_t^2 dM_t + d(W_t^2)dM_t \\ &= M_t(2W_t dW_t + dt) - W_t^2\big(u(t)M_t dW_t\big) \\ &\qquad\qquad -(2W_t dW_t + dt)\big(u(t)M_t dW_t\big) \\ &= M_t W_t\big(2 - u(t)W_t\big)dW_t + \big(M_t - 2u(t)W_t M_t\big)dt. \end{aligned}$$

Integrate between 0 and t

$$W_t^2 M_t = \int_0^t [M_s W_s\big(2 - u(s)W_s\big)]\,dW_s + \int_0^t \big(M_s - 2u(s)W_s M_s\big)\,ds,$$

and take the expected value to get

$$\begin{aligned}E[W_t^2 M_t] &= \int_0^t \big(E[M_s] - 2u(s)E[W_s M_s]\big)\,ds \\ &= \int_0^t \big(1 + 2u(s)U(s)\big)\,ds \\ &= t + U^2(t),\end{aligned}$$

where we used (9.2.11). Substituting into (9.2.13) yields

$$E^P[W_t^2 M_T] = t + U^2(t). \tag{9.2.14}$$

Substituting (9.2.12) and (9.2.14) into relation (9.2.9) yields

$$E^Q[X_t^2] = U^2(t) - 2U(t)^2 + t + U^2(t) = t, \tag{9.2.15}$$

which ends the proof of the lemma. ■

Now we are prepared to prove one of the most important results of stochastic calculus.

Theorem 9.2.3 (Girsanov Theorem) *Let $u \in L^2[0,T]$ be a deterministic function. Then the process*

$$X_t = \int_0^t u(s)\,ds + W_t, \qquad 0 \le t \le T$$

is a Brownian motion with respect to the probability measure Q given by

$$dQ = e^{-\int_0^T u(s)\,dW_s - \frac{1}{2}\int_0^T u(s)^2\,ds}dP.$$

Proof: In order to prove that X_t is a Brownian motion on the probability space $(\Omega, \mathcal{F}, Q)$ we shall apply Lévy's characterization theorem, see Theorem 2.1.5. Hence it suffices to show that X_t is a Wiener process. Lemma 9.2.1 implies that the process X_t satisfies the following properties:

1. $X_0 = 0$;
2. X_t is continuous in t;
3. X_t is a square integrable $\mathcal{F}_t$-martingale on the space $(\Omega, \mathcal{F}, Q)$.

The only property we still need to show is

4. $E^Q[(X_t - X_s)^2] = t - s, \qquad s < t.$

Using Lemma 9.2.2, the martingale property of W_t, the additivity and the tower property of expectations, we have

$$\begin{aligned}E^Q[(X_t - X_s)^2] &= E^Q[X_t^2] - 2E^Q[X_t X_s] + E^Q[X_s^2] \\ &= t - 2E^Q[X_t X_s] + s \\ &= t - 2E^Q[E^Q[X_t X_s|\mathcal{F}_s]] + s \\ &= t - 2E^Q[X_s E^Q[X_t|\mathcal{F}_s]] + s \\ &= t - 2E^Q[X_s^2] + s \\ &= t - 2s + s = t - s,\end{aligned}$$

which proves relation 4. ■

Choosing $u(s) = \lambda$, constant, we obtain the following consequence that will be useful later in finance applications.

Corollary 9.2.4 *Let W_t be a Brownian motion on the probability space $(\Omega, \mathcal{F}, P)$. Then the process*

$$X_t = \lambda t + W_t, \qquad 0 \le t \le T$$

is a Brownian motion on the probability space $(\Omega, \mathcal{F}, Q)$, where

$$dQ = e^{-\frac{1}{2}\lambda^2 T - \lambda W_T}\, dP.$$

Part II

Applications to Finance

Chapter 10

Modeling Stochastic Rates

Elementary Calculus provides powerful methods of modeling phenomena from the real world. However, the real world is imperfect, and in order to study it, one needs to employ methods of Stochastic Calculus.

10.1 An Introductory Problem

The model in an ideal world

Consider the amount of money $M(t)$ at time t invested in a bank account that pays interest at a constant rate r. The differential equation which models this problem is

$$\boxed{dM(t) = rM(t)dt.} \tag{10.1.1}$$

Given the initial investment $M(0) = M_0$, the account balance at time t is given by the solution of the equation, $M(t) = M_0 e^{rt}$.

The model in the real world

In the real world the interest rate r is not constant. It may be assumed constant only for a very small amount of time, such as one day or one week. The interest rate changes unpredictably in time, which means that it is a stochastic process. This can be modeled in several different ways. For instance, we may assume that the interest rate at time t is given by the continuous stochastic process $r_t = r + \sigma W_t$, where $\sigma > 0$ is a constant that controls the volatility of the rate, and W_t is a Brownian motion process. The process r_t represents a diffusion that starts at $r_0 = r$, with constant mean $E[r_t] = r$ and variance proportional with the time elapsed, $Var[r_t] = \sigma^2 t$. With this change in the model, the account balance at time t becomes a stochastic process M_t that satisfies the stochastic equation

$$\boxed{dM_t = (r + \sigma W_t)M_t dt, \qquad t \geq 0.} \tag{10.1.2}$$

Solving the equation
In order to solve this equation, we write it as $dM_t - r_t M_t dt = 0$ and multiply by the integrating factor $e^{-\int_0^t r_s\,ds}$. We can check that

$$\begin{aligned} d\Big(e^{-\int_0^t r_s\,ds}\Big) &= -e^{-\int_0^t r_s\,ds} r_t dt \\ dM_t\, d\Big(e^{-\int_0^t r_s\,ds}\Big) &= 0, \end{aligned}$$

since $dt^2 = dt\,dW_t = 0$. Using the product rule, the equation becomes exact

$$d\Big(M_t\, e^{-\int_0^t r_s\,ds}\Big) = 0.$$

Integrating yields the solution

$$\begin{aligned} M_t &= M_0 e^{\int_0^t r_s\,ds} = M_0 e^{\int_0^t (r+\sigma W_s)\,ds} \\ &= M_0 e^{rt+\sigma Z_t}, \end{aligned}$$

where $Z_t = \int_0^t W_s\,ds$ is the integrated Brownian motion process. Since the moment generating function of Z_t is $m(\sigma) = E[e^{\sigma Z_t}] = e^{\sigma^2 t^3/6}$ (see Exercise 2.3.4), we obtain

$$\begin{aligned} E[e^{\sigma W_s}] &= e^{\sigma^2 t^3/6}; \\ Var[e^{\sigma W_s}] &= m(2\sigma) - m(\sigma) = e^{\sigma^2 t^3/3}(e^{\sigma^2 t^3/3} - 1). \end{aligned}$$

Then the mean and variance of the solution $M_t = M_0 e^{rt+\sigma Z_t}$ are

$$\begin{aligned} E[M_t] &= M_0 e^{rt} E[e^{\sigma Z_t}] = M_0 e^{rt+\sigma^2 t^3/6}; \\ Var[M_t] &= M_0^2 e^{2rt} Var[e^{\sigma Z_t}] = M_0^2 e^{2rt+\sigma^2 t^3/3}(e^{\sigma^2 t^3/3} - 1). \end{aligned}$$

Conclusion
We shall make a few interesting remarks. If $M(t)$ and M_t represent the balance at time t in the ideal and real worlds, respectively, then

$$E[M_t] = M_0 e^{rt} e^{\sigma^2 t^3/6} > M_0 e^{rt} = M(t).$$

This means that we expect to have more money in the account of an ideal world rather than in the real world account. Similarly, a bank can expect to make more money when lending at a stochastic interest rate than at a constant interest rate. This inequality is due to the convexity of the exponential function. If $X_t = rt + \sigma Z_t$, then Jensen's inequality yields

$$E[e^{X_t}] \geq e^{E[X_t]} = e^{rt}.$$

10.2 Langevin's Equation

We shall consider another stochastic extension of the equation (10.1.1). We shall allow for continuously random deposits and withdrawals which can be modeled by an unpredictable term, given by αdW_t, with α constant. The obtained equation

$$\boxed{dM_t = rM_t dt + \alpha dW_t, \quad t \geq 0} \tag{10.2.3}$$

is called *Langevin's equation.*

Solving the equation
We shall solve it as a linear stochastic equation. Multiplying by the integrating factor e^{-rt} yields

$$d(e^{-rt}M_t) = \alpha e^{-rt} dW_t.$$

Integrating we obtain

$$e^{-rt}M_t = M_0 + \alpha \int_0^t e^{-rs}\, dW_s.$$

Hence the solution is

$$\boxed{M_t = M_0 e^{rt} + \alpha \int_0^t e^{r(t-s)}\, dW_s.} \tag{10.2.4}$$

This is called the *Ornstein-Uhlenbeck process*. Since the last term is a Wiener integral, by Proposition (7.3.1) we have that M_t is Gaussian with the mean

$$E[M_t] = M_0 e^{rt} + E[\alpha \int_0^t e^{r(t-s)}\, dW_s] = M_0 e^{rt}$$

and variance

$$Var[M_t] = Var\big[\alpha \int_0^t e^{r(t-s)}\, dW_s\big] = \frac{\alpha^2}{2r}(e^{2rt} - 1).$$

It is worth noting that the expected balance is equal to the ideal world balance $M_0 e^{rt}$. The variance for t small is approximately equal to $\alpha^2 t$, which is the variance of αW_t.

If the constant α is replaced by an unpredictable function $\alpha(t, W_t)$, the equation becomes

$$dM_t = rM_t dt + \alpha(t, W_t) dW_t, \quad t \geq 0.$$

Using a similar argument we arrive at the following solution:

$$\boxed{M_t = M_0 e^{rt} + \int_0^t e^{r(t-s)} \alpha(t, W_t)\, dW_s.} \tag{10.2.5}$$

This process is not Gaussian. Its mean and variance are given by

$$\begin{aligned} E[M_t] &= M_0 e^{rt} \\ Var[M_t] &= \int_0^t e^{2r(t-s)} E[\alpha^2(t, W_t)]\, ds. \end{aligned}$$

In the particular case when $\alpha(t, W_t) = e^{\sqrt{2r}W_t}$, using Application 6.3.9 with $\lambda = \sqrt{2r}$, we can work out for (10.2.5) an explicit form of the solution

$$\begin{aligned} M_t &= M_0 e^{rt} + \int_0^t e^{r(t-s)} e^{\sqrt{2r}W_t}\, dW_s \\ &= M_0 e^{rt} + e^{rt} \int_0^t e^{-rs} e^{\sqrt{2r}W_t}\, dW_s \\ &= M_0 e^{rt} + e^{rt} \frac{1}{\sqrt{2r}} \big(e^{-rs+\sqrt{2r}W_t} - 1 \big) \\ &= M_0 e^{rt} + \frac{1}{\sqrt{2r}} \big(e^{\sqrt{2r}W_t} - e^{rt} \big). \end{aligned}$$

The previous model for the interest rate is not a realistic one because it allows for negative rates. The Brownian motion hits $-r$ after a finite time with probability 1. However, it might be a good stochastic model for something such as the evolution of population, since in this case the rate can be negative.

10.3 Equilibrium Models

Let r_t denote the spot rate at time t. This is the rate at which one can invest for the shortest period of time.[1] For the sake of simplicity, we assume the interest rate r_t satisfies an equation with one source of uncertainty

$$\boxed{dr_t = m(r_t)dt + \sigma(r_t)dW_t}. \tag{10.3.6}$$

The drift rate and volatility of the spot rate r_t do not depend explicitly on the time t. There are several classical choices for $m(r_t)$ and $\sigma(r_t)$ that will be discussed in the following sections.

10.3.1 The Rendleman and Bartter Model

The model introduced in 1986 by Rendleman and Bartter [14] assumes that the short-time rate satisfies the process

$$\boxed{dr_t = \mu r_t dt + \sigma r_t dW_t.}$$

The growth rate μ and the volatility σ are considered constants. This equation has been solved in Example 7.7.2 and its solution is given by

$$r_t = r_0 e^{(\mu - \frac{\sigma^2}{2})t + \sigma W_t}.$$

[1]The longer the investment period, the larger the rate. The instantaneous rate or spot rate is the rate at which one can invest for a short period of time, such as one day or one second.

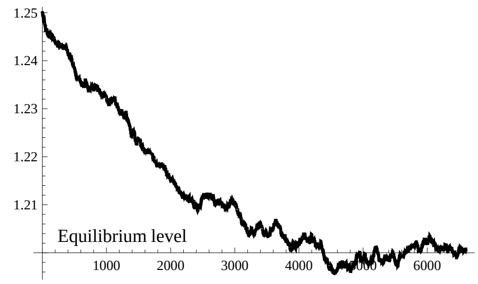


Figure 10.1: *A simulation of* $dr_t = a(b - r_t)dt + \sigma dW_t$, *with* $r_0 = 1.25$, $a = 3$, $\sigma = 1\%$, $b = 1.2$.

The distribution of r_t is log-normal. Using Example 7.2.1, the mean and variance become

$$\begin{aligned} E[r_t] &= r_0 e^{(\mu - \frac{\sigma^2}{2})t} E[e^{\sigma W_t}] = r_0 e^{(\mu - \frac{\sigma^2}{2})t} e^{\sigma^2 t/2} = r_0 e^{\mu t}. \\ Var[r_t] &= r_0^2 e^{2(\mu - \frac{\sigma^2}{2})t} Var[e^{\sigma W_t}] = r_0^2 e^{2(\mu - \frac{\sigma^2}{2})t} e^{\sigma^2 t}(e^{\sigma^2 t} - 1) \\ &= r_0^2 e^{2\mu t}(e^{\sigma^2 t} - 1). \end{aligned}$$

The next two models incorporate the *mean reverting* phenomenon of interest rates. This means that in the long run the rate converges towards an average level. These models are more realistic and are based on economic arguments.

10.3.2 The Vasicek Model

Vasicek's assumption is that the short-term interest rates should satisfy the mean reverting stochastic differential equation

$$\boxed{dr_t = a(b - r_t)dt + \sigma dW_t,} \qquad (10.3.7)$$

with a, b, σ positive constants, see Vasicek [16].

Assuming the spot rates are deterministic, we take $\sigma = 0$ and obtain the ODE

$$dr_t = a(b - r_t)dt.$$

Solving it as a linear equation yields the solution

$$r_t = b + (r_0 - b)e^{-at}.$$

This implies that the rate r_t is pulled towards level b at the rate a. This means that, if $r_0 > b$, then r_t is decreasing towards b, and if $r_0 < b$, then r_t is increasing towards the horizontal asymptote b. The term σdW_t in Vasicek's model adds some "white noise" to the process. In the following we shall find an explicit formula for the spot rate r_t in the stochastic case.

Proposition 10.3.1 *The solution of the equation (11.4.5) is given by*

$$\boxed{r_t = b + (r_0 - b)e^{-at} + \sigma e^{-at} \int_0^t e^{as}\, dW_s.}$$

The process r_t is Gaussian with mean and variance

$$\begin{aligned} E[r_t] &= b + (r_0 - b)e^{-at}; \\ Var[r_t] &= \frac{\sigma^2}{2a}(1 - e^{-2at}). \end{aligned}$$

Proof: Multiplying by the integrating factor e^{at} yields

$$d\Big(e^{at} r_t\Big) = abe^{at}\, dt + \sigma e^{at}\, dW_t.$$

Integrating between 0 and t and dividing by e^{at} we get

$$\begin{aligned} r_t &= r_0 e^{-at} + be^{-at}(e^{at} - 1) + \sigma e^{-at} \int_0^t e^{as}\, dW_s \\ &= b + (r_0 - b)e^{-at} + \sigma e^{-at} \int_0^t e^{as}\, dW_s. \end{aligned}$$

Since the spot rate r_t is the sum between the predictable function $r_0 e^{-at} + be^{-at}(e^{at} - 1)$ and a multiple of a Wiener integral, from Proposition (4.6.1) it follows that r_t is Gaussian, with

$$\begin{aligned} E[r_t] &= b + (r_0 - b)e^{-at} \\ Var(r_t) &= Var\Big[\sigma e^{-at} \int_0^t e^{as}\, dW_s\Big] = \sigma^2 e^{-2at} \int_0^t e^{2as}\, ds \\ &= \frac{\sigma^2}{2a}(1 - e^{-2at}). \end{aligned}$$

■

The following consequence explains the name of *mean reverting rate.*

Remark 10.3.2 *Since $\lim_{t\to\infty} E[r_t] = \lim_{t\to\infty}(b + (r_0 - b)e^{-at}) = b$, it follows that the spot rate r_t tends to b as $t \to \infty$. The variance tends in the long run to $\frac{\sigma^2}{2a}$.*

Since r_t is normally distributed, the Vasicek model has been criticized because it allows for negative interest rates and unbounded large rates. See Fig.10.1 for a simulation of the short-term interest rates for the Vasicek model.

Exercise 10.3.3 *Let $0 \le s < t$. Find the following expectations*
(a) $E[W_t \int_0^s W_u e^{au}\, du]$.
(b) $E[\int_0^t W_u e^{au}\, du \int_0^s W_v e^{av}\, dv]$.

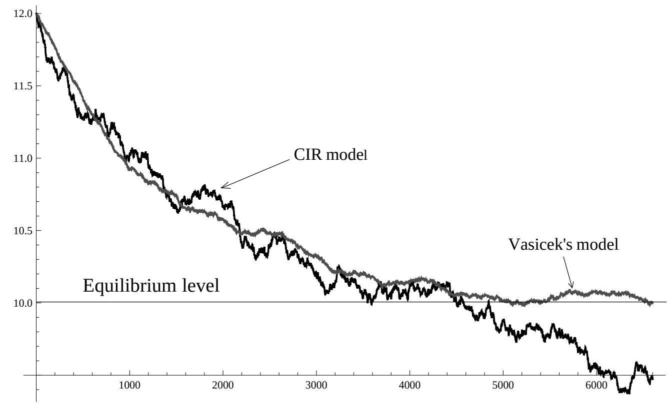


Figure 10.2: *Comparison between a simulation in CIR and Vasicek models, with parameter values $a = 3$, $\sigma = 15\%$, $r_0 = 12$, $b = 10$. Note that the CIR process tends to be more volatile than Vasicek's.*

Exercise 10.3.4 (*a*) *Find the probability that r_t is negative.*
(*b*) *What happens with this probability when $t \to \infty$?*
(*c*) *Find the rate of change of this probability.*
(*d*) *Compute $Cov(r_s, r_t)$.*

10.3.3 The Cox-Ingersoll-Ross Model

The Cox-Ingersoll-Ross (CIR) model assumes that the spot rates verify the stochastic equation

$$\boxed{dr_t = a(b - r_t)dt + \sigma\sqrt{r_t}dW_t,} \tag{10.3.8}$$

with a, b, σ constants, see [5]. Two main advantages of this model are
• the process exhibits mean reversion.
• it is not possible for the interest rates to become negative.

A process that satisfies equation (10.3.8) is called a *CIR process*. This is not a Gaussian process. In the following we shall compute its first two moments.

To get the first moment, integrating the equation (10.3.8) between 0 and t which yields

$$r_t = r_0 + abt - a\int_0^t r_s\,ds + \sigma\int_0^t \sqrt{r_s}\,dW_s.$$

Take the expectation we obtain

$$E[r_t] = r_0 + abt - a\int_0^t E[r_s]\,ds.$$

Denote $\mu(t) = E[r_t]$. Then differentiate with respect to t in

$$\mu(t) = r_0 + abt - a\int_0^t \mu(s)\,ds,$$

which yields the differential equation $\mu'(t) = ab - a\mu(t)$. Multiply by the integrating factor e^{at} and obtain

$$d(e^{at}\mu(t)) = abe^{at}.$$

Integrate and use that $\mu(0) = r_0$, which provides the solution

$$\boxed{\mu(t) = b + e^{-at}(r_0 - b).}$$

Hence

$$\lim_{t\to\infty} E[r_t] = \lim_{t\to\infty}\big(b + e^{-at}(r_0 - b)\big) = b,$$

which shows that the process is mean reverting.

We compute in the following the second moment $\mu_2(t) = E[r_t^2]$. By Ito's formula we have

$$\begin{aligned} d(r_t^2) &= 2r_t\,dr_t + (dr_t)^2 = 2r_t\,dr_t + \sigma^2 r_t dt \\ &= 2r_t\,[a(b - r_t)dt + \sigma\sqrt{r_t}dW_t] + \sigma^2 r_t dt \\ &= [(2ab + \sigma^2)r_t - 2ar_t^2]dt + 2\sigma r_t^{3/2} dW_t. \end{aligned}$$

Integrating we get

$$r_t^2 = r_0^2 + \int_0^t [(2ab + \sigma^2)r_s - 2ar_s^2]\,ds + 2\sigma\int_0^t r_s^{3/2}\,dW_s.$$

Taking the expectation yields

$$\mu_2(t) = r_0^2 + \int_0^t [(2ab + \sigma^2)\mu(s) - 2a\mu_2(s)]\,ds.$$

Differentiating again:

$$\mu_2'(t) = (2ab + \sigma^2)\mu(t) - 2a\mu_2(t).$$

Solving as a linear differential equation in $\mu_2(t)$ yields

$$d(\mu_2(t)e^{2at}) = (2ab + \sigma^2)e^{2at}\mu(t).$$

Substituting the value of $\mu(t)$ and integrating yields

$$\mu_2(t)e^{2at} = r_0^2 + (2ab + \sigma^2)\Big[\frac{b}{2a}(e^{2at} - 1) + \frac{r_0 - b}{a}(e^{at} - 1)\Big].$$

Hence the second moment has the formula

$$\boxed{\mu_2(t) = r_0^2 e^{-2at} + (2ab + \sigma^2)\Big[\frac{b}{2a}(1 - e^{-2at}) + \frac{r_0 - b}{a}(1 - e^{-at})e^{-at}\Big].}$$

Exercise 10.3.5 *Use a similar method to find a recursive formula for the moments of a CIR process.*

10.4 No-arbitrage Models

In the following models the drift rate is a function of time, which is chosen such that the model is consistent with the term structure.

10.4.1 The Ho and Lee Model

The first no-arbitrage model was proposed in 1986 by Ho and Lee [7]. The model was presented initially in the form of a binomial tree. The continuous time-limit of this model is

$$\boxed{dr_t = \theta(t)dt + \sigma dW_t.}$$

In this model $\theta(t)$ is the average direction in which r_t moves and it is considered independent of r_t, while σ is the standard deviation of the short rate. The solution process is Gaussian and is given by

$$r_t = r_0 + \int_0^t \theta(s)\, ds + \sigma W_t.$$

If $F(0,t)$ denotes the forward rate at time t as seen at time 0, it is known that $\theta(t) = F_t(0,t) + \sigma^2 t$. In this case the solution becomes

$$r_t = r_0 + F(0,t) + \frac{\sigma^2}{2}t^2 + \sigma W_t.$$

10.4.2 The Hull and White Model

The model proposed by Hull and White [8] is an extension of the Ho and Lee model model that incorporates mean reversion

$$\boxed{dr_t = (\theta(t) - ar_t)dt + \sigma dW_t,}$$

with a and σ constants. We can solve the equation by multiplying by the integrating factor e^{at}

$$d(e^{at}r_t) = \theta(t)e^{at}dt + \sigma e^{at}dW_t.$$

Integrating between 0 and t yields

$$r_t = r_0 e^{-at} + e^{-at}\int_0^t \theta(s)e^{as}\, ds + \sigma e^{-at}\int_0^t e^{as}\, dW_s. \tag{10.4.9}$$

Since the first two terms are deterministic and the last is a Wiener integral, the process r_t is Gaussian.

The function $\theta(t)$ can be calculated from the term structure $F(0,t)$ as

$$\theta(t) = \partial_t F(0,t) + aF(0,t) + \frac{\sigma^2}{2a}(1 - e^{-2at}).$$

Then

$$\begin{aligned}
\int_0^t \theta(s)e^{as}\,ds &= \int_0^t \partial_s F(0,s)e^{as}\,ds + a\int_0^t F(0,s)e^{as}\,ds + \frac{\sigma^2}{2a}\int_0^t (1-e^{-2as})e^{as}\,ds \\
&= F(0,t)e^{at} - r_0 + \frac{\sigma^2}{a^2}\big(\cosh(at)-1\big),
\end{aligned}$$

where we used that $F(0,0)=r_0$. The deterministic part of r_t becomes

$$r_0e^{-at} + e^{-at}\int_0^t \theta(s)e^{as}\,ds = F(0,t) + \frac{\sigma^2}{a^2}e^{-at}\big(\cosh(at)-1\big).$$

An algebraic manipulation shows that

$$e^{-at}\big(\cosh(at)-1\big) = \frac{1}{2}(1-e^{at})^2.$$

Substituting into (10.4.9) yields

$$\boxed{r_t = F(0,t) + \frac{\sigma^2}{2a^2}(1-e^{at})^2 + \sigma e^{-at}\int_0^t e^{as}\,dW_s.}$$

The mean and variance are:

$$\begin{aligned}
E[r_t] &= F(0,t) + \frac{\sigma^2}{2a^2}(1-e^{at})^2 \\
Var(r_t) &= \sigma^2e^{-2at}Var\Big[\int_0^t e^{as}\,dW_s\Big] = \sigma^2e^{-2at}\int_0^t e^{2as}\,ds \\
&= \frac{\sigma^2}{2a}(1-e^{-2at}).
\end{aligned}$$

10.5 Nonstationary Models

These models assume both θ and σ as functions of time. In the following we shall discuss two models with this property.

10.5.1 Black, Derman and Toy Model

The binomial tree of Black, Derman and Toy [2] is equivalent with the following continuous model of short-time rate

$$\boxed{d(\ln r_t) = \Big[\theta(t) + \frac{\sigma'(t)}{\sigma(t)}\ln r_t\Big]dt + \sigma(t)dW_t.}$$

Making the substitution $u_t = \ln r_t$, we obtain a linear equation in u_t

$$du_t = \Big[\theta(t) + \frac{\sigma'(t)}{\sigma(t)}u_t\Big]dt + \sigma(t)dW_t.$$

The equation can be written equivalently as

$$\frac{\sigma(t)du_t - d\sigma(t)\,u_t}{\sigma^2(t)} = \frac{\theta(t)}{\sigma(t)}dt + dW_t,$$

which after using the quotient rule becomes

$$d\Big[\frac{u_t}{\sigma(t)}\Big] = \frac{\theta(t)}{\sigma(t)}dt + dW_t.$$

Integrating and solving for u_t leads to

$$u_t = \frac{u_0}{\sigma(0)}\sigma(t) + \sigma(t)\int_0^t \frac{\theta(s)}{\sigma(s)}\,ds + \sigma(t)W_t.$$

This implies that u_t is Gaussian and hence $r_t = e^{u_t}$ is log-normal for each t. Using $u_0 = \ln r_0$ and

$$e^{\frac{u_0}{\sigma(0)}\sigma(t)} = e^{\frac{\sigma(t)}{\sigma(0)}\ln r_0} = r_0^{\frac{\sigma(t)}{\sigma(0)}},$$

we obtain the following explicit formula for the spot rate

$$\boxed{r_t = r_0^{\frac{\sigma(t)}{\sigma(0)}}\, e^{\sigma(t)\int_0^t \frac{\theta(s)}{\sigma(s)}\,ds} e^{\sigma(t)W_t}.} \qquad (10.5.10)$$

Since $\sigma(t)W_t$ is normally distributed with mean 0 and variance $\sigma^2(t)t$, the log-normal variable $e^{\sigma(t)W_t}$ has

$$\begin{aligned} E[e^{\sigma(t)W_t}] &= e^{\sigma^2(t)t/2} \\ Var[e^{\sigma(t)W_t}] &= e^{\sigma(t)^2 t}(e^{\sigma(t)^2 t} - 1). \end{aligned}$$

Hence

$$\begin{aligned} E[r_t] &= r_0^{\frac{\sigma(t)}{\sigma(0)}}\, e^{\sigma(t)\int_0^t \frac{\theta(s)}{\sigma(s)}\,ds} e^{\sigma^2(t)t/2} \\ Var[r_t] &= r_0^{\frac{2\sigma(t)}{\sigma(0)}}\, e^{2\sigma(t)\int_0^t \frac{\theta(s)}{\sigma(s)}\,ds} e^{\sigma(t)^2 t}(e^{\sigma(t)^2 t} - 1). \end{aligned}$$

Exercise 10.5.1 (a) *Solve the Black, Derman and Toy model in the case when σ is constant.*
(b) *Show that in this case the spot rate r_t is log-normally distributed.*
(c) *Find the mean and the variance of r_t.*

10.5.2 Black and Karasinski Model

In 1991 Black and Karasinski [3] proposed the following more general model for the spot rates:

$$\boxed{d(\ln r_t) = \Big[\theta(t) - a(t)\ln r_t\Big]dt + \sigma(t)dW_t}$$

Exercise 10.5.2 *Find the explicit formula for r_t in this case.*

Chapter 11

Bond Valuation and Yield Curves

A *bond* is a contract between a buyer and a financial institution (bank, government, etc) by which the financial institution agrees to pay a certain principal to the buyer at a determined time T in the future, plus some periodic coupon payments done during the life time of the contract. If there are no coupons to be payed, the bond is called *a zero coupon bond* or a *discount bond*. The price of a bond at any time $0 \leq t \leq T$ is denoted by $P(t,T)$.

If the spot rates are constant, $r_t = r$, then the price at time t of a discount bond that pays off \$1 at time T is given by

$$P(t,T) = e^{-r(T-t)}.$$

If the spot interest rates depend deterministic on time, $r_t = r(t)$, the formula becomes

$$P(t,T) = e^{-\int_t^T r(s)\,ds}.$$

However, the spot rates r_t are in general stochastic. In this case, the formula $e^{-\int_t^T r_s\,ds}$ is a random variable, and the price of the bond in this case can be calculated as the expectation

$$P(t,T) = E_t[e^{-\int_t^T r_s\,ds}],$$

where E_t is the expectation as of time t. If $\bar{r} = \dfrac{1}{T-t}\displaystyle\int_t^T r_s\,ds$ denotes the average value of the spot rate in the time interval between t and T, the previous formula can be written in the equivalent form

$$P(t,T) = E_t[e^{-\bar{r}(T-t)}].$$

11.1 The Case of a Brownian Motion Spot Rate

Assume the spot rate r_t satisfies the stochastic equation

$$dr_t = \sigma dW_t,$$

with $\sigma > 0$, constant. We shall show that the price of the zero-coupon bond that pays off \$1 at time T is given by

$$P(t, T) = e^{-r_t(T-t)+\frac{1}{6}\sigma^2(T-t)^3}.$$

Let $0 < t < T$ be fixed. Solving the stochastic differential equation, we have for any $t < s < T$

$$r_s = r_t + \sigma(W_s - W_t).$$

Integrating yields

$$\int_t^T r_s\, ds = r_t(T-s) + \sigma \int_t^T (W_s - W_t)\, ds.$$

Then taking the exponential, we obtain

$$e^{-\int_t^T r_s\, ds} = e^{-r_t(T-s)} e^{-\sigma \int_t^T (W_s - W_t)\, ds}.$$

The price of the bond at time t is given by

$$\begin{aligned}
P(t, T) &= E\Big[e^{-\int_t^T r_s\, ds}\Big|\mathcal{F}_t\Big] \\
&= e^{-r_t(T-t)} E\Big[e^{-\sigma \int_t^T (W_s - W_t)\, ds}\Big|\mathcal{F}_t\Big] \\
&= e^{-r_t(T-t)} E\Big[e^{-\sigma \int_t^T (W_s - W_t)\, ds}\Big] \\
&= e^{-r_t(T-t)} E\Big[e^{-\sigma \int_0^{T-t} W_s\, ds}\Big] \\
&= e^{-r_t(T-t)} e^{\frac{\sigma^2}{2}\frac{(T-t)^3}{3}}.
\end{aligned}$$

In the second identity we took the $\mathcal{F}_t$-predictable part out of the expectation, while in the third identity we dropped the condition since $W_s - W_t$ is independent of the information set $\mathcal{F}_t$ for any $t < s$. The fourth identity invoked the stationarity. The last identity follows from the Exercise 1.6.2 (b) and from the fact that $-\sigma \int_0^{T-t} W_s\, ds$ is normal distributed with mean 0 and variance $\frac{\sigma^2(T-t)^3}{3}$.

It is worth noting that the price of the bond, $P(t, T)$, depends only the spot rate at time t, r_t, and the time difference $T - t$.

Exercise 11.1.1 *Spot rates which exhibit positive jumps of size $\sigma > 0$ satisfy the stochastic equation*

$$dr_t = \sigma dN_t,$$

where N_s is the Poisson process. Find the price of the zero-coupon bond that pays off \$1 at time T.

11.2 The Case of Vasicek's Model

The next result regarding bond pricing is due to Vasicek [16]. Its initial proof is based on partial differential equations, while here we provide an approach solely based on expectatations.

Proposition 11.2.1 *Assume the spot rate r_t satisfies Vasicek's mean reverting model*

$$dr_t = a(b - r_t)dt + \sigma dW_t.$$

We shall show that the price of the zero-coupon bond that pays off \$1 at time T is given by

$$P(t,T) = A(t,T)e^{-B(t,T)r_t}, \tag{11.2.1}$$

where

$$\begin{aligned} B(t,T) &= \frac{1-e^{-a(T-t)}}{a} \\ A(t,T) &= e^{\frac{(B(t,T)-T+t)(a^2b-\sigma^2/2)}{a^2} - \frac{\sigma^2 B(t,T)^2}{4a}}. \end{aligned}$$

We shall start the deduction of the previous formula by integrating the stochastic differential equation between t and T yields

$$-a\int_t^T r_s\, ds = r_T - r_t - \sigma(W_T - W_t) - ab(T-t).$$

Taking the exponential we obtain

$$e^{-\int_t^T r_s\, ds} = e^{\frac{1}{a}(r_T - r_t) - \frac{\sigma}{a}(W_T - W_t) - b(T-t)}. \tag{11.2.2}$$

Multiplying in the stochastic differential equation by e^{as} yields the exact equation

$$d(e^{as}r_s) = abe^{as}ds + \sigma e^{as}dW_s.$$

Integrating between t and T to get

$$e^{aT}r_T - e^{at}r_t = b(e^{aT} - e^{at}) + \sigma\int_t^T e^{as}\, dW_s.$$

Solving for r_T and subtracting r_t yields

$$r_T - r_t = -r_t(1 - e^{a(T-t)}) + b(1 - e^{a(T-t)}) + \sigma e^{-aT}\int_t^T e^{as}\, dW_s.$$

Dividing by a and using the notation for $B(t,T)$ yields

$$\frac{1}{a}(r_T - r_t) = -r_t B(t,T) + bB(t,T) + \frac{\sigma}{a}e^{-aT}\int_t^T e^{as}\, dW_s.$$

Substituting into (11.4.7) and using that $W_T - W_t = \int_t^T dW_s$, we get

$$e^{-\int_t^T r_s\,ds} = e^{-r_t B(t,T)} e^{bB(t,T)-b(T-t)} e^{\frac{\sigma}{a}\int_t^T (e^{as-aT}-1)\,dW_s}. \qquad (11.2.3)$$

Taking the predictable part out and dropping the independent condition, the price of the zero-coupon bond at time t becomes

$$\begin{aligned} P(t,T) &= E\big[e^{-\int_t^T r_s\,ds}|\mathcal{F}_t\big] = e^{-r_t B(t,T)} e^{bB(t,T)-b(T-t)} E\big[e^{\frac{\sigma}{a}\int_t^T (e^{as-aT}-1)\,dW_s}|\mathcal{F}_t\big] \\ &= e^{-r_t B(t,T)} e^{bB(t,T)-b(T-t)} E\big[e^{\frac{\sigma}{a}\int_t^T (e^{as-aT}-1)\,dW_s}\big] \\ &= e^{-r_t B(t,T)} A(t,T), \end{aligned}$$

where

$$A(t,T) = e^{bB(t,T)-b(T-t)} E\big[e^{\frac{\sigma}{a}\int_t^T (e^{as-aT}-1)\,dW_s}\big]. \qquad (11.2.4)$$

As a Wiener integral, $\displaystyle\int_t^T (e^{as-aT}-1)\,dW_s$ is normally distributed with mean zero and variance

$$\begin{aligned} \int_t^T (e^{as-aT}-1)^2\,ds &= e^{-2aT}\frac{e^{2aT}-e^{2at}}{a} - 2e^{-aT}\frac{e^{aT}-e^{at}}{a} + (T-t) \\ &= B(t,T)\frac{2-aB(t,T)}{2} - 2B(t,T) + (T-t) \\ &= -\frac{a}{2}B(t,T)^2 - B(t,T) + (T-t). \end{aligned}$$

Then the log-normal variable $e^{\frac{\sigma}{a}\int_t^T (e^{as-aT}-1)\,dW_s}$ has the mean

$$E\big[e^{\frac{\sigma}{a}\int_t^T (e^{as-aT}-1)\,dW_s}\big] = e^{\frac{1}{2}\frac{\sigma^2}{a^2}[-\frac{a}{2}B(t,T)^2 - B(t,T) + (T-t)]}.$$

Substituting into (11.4.9) and using that

$$\begin{aligned} & bB(t,T) - b(T-t) + \frac{1}{2}\frac{\sigma^2}{a^2}[-\frac{a}{2}B(t,T)^2 - B(t,T) + (T-t)] \\ =& -\frac{\sigma^2}{4a}B(t,T)^2 + (T-t)\Big(\frac{\sigma^2}{2a^2} - b\Big) + \Big(b - \frac{\sigma^2}{2a^2}\Big)B(t,T) \\ =& -\frac{\sigma^2}{4a}B(t,T)^2 + \frac{1}{a^2}(a^2 b - \sigma^2/2)\big(B(t,T) - T + t\big) \end{aligned}$$

we obtain

$$A(t,T) = e^{\frac{(B(t,T)-T+t)(a^2 b - \sigma^2/2)}{a^2} - \frac{\sigma^2 B(t,T)^2}{4a}}.$$

It is worth noting that $P(t,T)$ depends on the time to maturity, $T-t$, and the spot rate r_t.

Exercise 11.2.2 *Find the price of an infinitely lived bond in the case when spot rates satisfy Vasicek's model.*

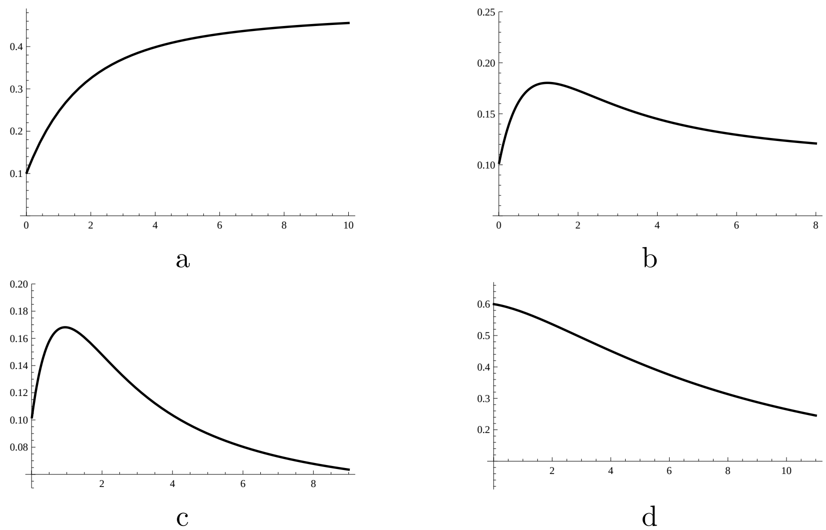


Figure 11.1: *The yield on zero-coupon bond versus maturity time $T-t$ in the case of Vasicek's model: a. $a=1$, $b=0.5$, $\sigma=0.6$, $r_t=0.1$; b. $a=1$, $b=0.5$, $\sigma=0.8$, $r_t=0.1$; c. $a=1$, $b=0.5$, $\sigma=0.97$, $r_t=0.1$; d. $a=0.3$, $b=0.5$, $\sigma=0.3$, $r_t=0.r$.*

The term structure Let $R(t,T)$ be the continuously compounded interest rate at time t for a term of $T-t$. Using the formula for the bond price $B(t,T)=e^{-R(t,T)(T-t)}$ we get

$$R(t,T)=-\frac{1}{T-t}\ln B(t,T).$$

Using formula for the bond price (11.2.1) yields

$$R(t,T)=-\frac{1}{T-t}\ln A(t,T)+\frac{1}{T-t}B(t,T)r_t.$$

A few possible shapes of the term structure $R(t,T)$ in the case of Vasicek's model are given in Fig. 11.1.

11.3 The Case of CIR's Model

In the case when r_t satisfy the CIR model (10.3.8), the zero-coupon bond price has a similar form as in the case of Vasicek's model

$$P(t,T)=A(t,T)e^{-B(t,T)r_t}.$$

The functions $A(t,T)$ and $B(t,T)$ are given in this case by

$$\begin{aligned} B(t,T) &= \frac{2(e^{\gamma(T-t)}-1)}{(\gamma+a)(e^{\gamma(T-t)}-1)+2\gamma} \\ A(t,T) &= \left(\frac{2\gamma e^{(a+\gamma)(T-t)/2}}{(\gamma+a)(e^{\gamma(T-t)}-1)+2\gamma}\right)^{2ab/\sigma^2}, \end{aligned}$$

where $\gamma = (a^2+2\sigma^2)^{1/2}$. For details the reader can consult [5].

11.4 The Case of a Mean Reverting Model with Jumps

We shall consider a model for the short-term interest rate r_t that is mean reverting and incorporates jumps. Let N_t be the Poisson process of constant rate λ and $M_t = N_t - \lambda t$ denote the compensated Poisson process, which is a martingale.

Consider the following model for the spot rate

$$dr_t = a(b - r_t)dt + \sigma dM_t, \tag{11.4.5}$$

with a, b, σ positive constants. It is worth noting the similarity with the Vasicek's model, which is obtained by replacing the process dM_t by dW_t, where W_t is a one-dimensional Brownian motion.

Making abstraction of the uncertainty source σdM_t, this implies that the rate r_t is pulled towards level b at the rate a. This means that, if for instance $r_0 > b$, then r_t is decreasing towards b. The term σdM_t adds jumps of size σ to the process. A few realizations of the process r_t are given in Fig. 11.2.

The stochastic differential equation (11.4.5) can be solved explicitly using the method of integrating factor, see [13]. Multiplying by e^{at} we get an exact equation; integrating between 0 and t yields the following closed form formula for the spot rate

$$r_t = b + (r_0 - b)e^{-at} + \sigma e^{-at}\int_0^t e^{as}\, dM_s. \tag{11.4.6}$$

Since the integral with respect to dM_t is a martingale, the expectation of the spot rate is

$$E[r_t] = b + (r_0 - b)e^{-at}.$$

In long run $E[r_t]$ tends to b, which shows that the process is mean reverting. Using the formula

$$E\Big[\Big(\int_0^t f(s)\, dM_s\Big)^2\Big] = \lambda\int_0^t f(s)^2\, ds$$

the variance of the spot rate can be computed as follows

$$\begin{aligned} Var(r_t) &= \sigma^2 e^{-2at} Var\Big(\int_0^t e^{as}\, dM_s\Big) = \sigma^2 e^{-2at} E\Big[\Big(\int_0^t e^{as}\, dM_s\Big)^2\Big] \\ &= \frac{\lambda\sigma^2}{2a}(1 - e^{-2at}). \end{aligned}$$

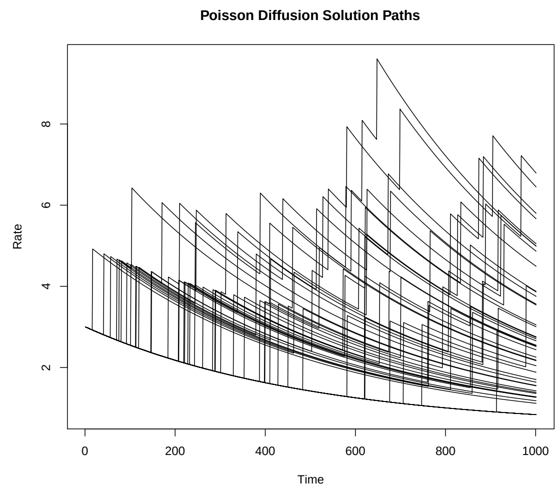


Figure 11.2: *Realizations of the spot rate r_t.*

It is worth noting that in long run the variance tends to the constant $\frac{\lambda\sigma^2}{2a}$. Another feature implied by this formula is that the variance is proportional with the frequency of jumps λ.

In the following we shall valuate a zero-coupon bond. Let $\mathcal{F}_t$ be the σ-algebra generated by the random variables $\{N_s; s \le t\}$. This contains all information about jumps occurred until time t. It is known that the price at time t of a zero-coupon bond with the expiration time T is given by

$$P(t,T) = E\big[e^{-\int_t^T r_s\,ds}|\mathcal{F}_t\big].$$

One of the main results is given in the following.

Proposition 11.4.1 *Assume the spot rate r_t satisfies the mean reverting model with jumps*

$$dr_t = a(b - r_t)dt + \sigma dM_t.$$

Then the price of the zero-coupon bond that pays off \$1 at time T is given by

$$P(t,T) = A(t,T)e^{-B(t,T)r_t},$$

where

$$\begin{aligned}
B(t,T) &= \frac{1 - e^{-a(T-t)}}{a} \\
A(t,T) &= \exp\{(b + \frac{\lambda\sigma}{a})B(t,T) + [\lambda(1 - \frac{\sigma}{a}) - b](T-t) - \lambda e^{-\sigma a}\int_0^{T-t} e^{\frac{\sigma}{a}e^{-ax}}\,dx\}.
\end{aligned}$$

Proof: Integrating the stochastic differential equation between t and T yields

$$-a\int_t^T r_s\,ds = r_T - r_t - \sigma(M_T - M_t) - ab(T-t).$$

Taking the exponential we obtain

$$e^{-\int_t^T r_s\,ds} = e^{\frac{1}{a}(r_T - r_t) - \frac{\sigma}{a}(M_T - M_t) - b(T-t)}. \tag{11.4.7}$$

Multiplying in the stochastic differential equation by e^{as} yields the exact equation

$$d(e^{as} r_s) = abe^{as}ds + \sigma e^{as} dM_s.$$

Integrate between t and T to get

$$e^{aT} r_T - e^{at} r_t = b(e^{aT} - e^{at}) + \sigma\int_t^T e^{as}\,dM_s.$$

Solving for r_T and subtracting r_t yields

$$r_T - r_t = -r_t(1 - e^{a(T-t)}) + b(1 - e^{a(T-t)}) + \sigma e^{-aT}\int_t^T e^{as}\,dM_s.$$

Dividing by a and using the notation for $B(t,T)$ yields

$$\frac{1}{a}(r_T - r_t) = -r_t B(t,T) + bB(t,T) + \frac{\sigma}{a}e^{-aT}\int_t^T e^{as}\,dM_s.$$

Substituting into (11.4.7) and using that $M_T - M_t = \int_t^T dM_s$, we get

$$e^{-\int_t^T r_s\,ds} = e^{-r_t B(t,T)} e^{bB(t,T) - b(T-t)} e^{\frac{\sigma}{a}\int_t^T (e^{as-aT} - 1)\,dM_s}. \tag{11.4.8}$$

Taking the predictable part out and dropping the independent condition, the price of the zero-coupon bond at time t becomes

$$\begin{aligned}
P(t,T) &= E\big[e^{-\int_t^T r_s\,ds}|\mathcal{F}_t\big] = e^{-r_t B(t,T)} e^{bB(t,T)-b(T-t)} E\big[e^{\frac{\sigma}{a}\int_t^T (e^{as-aT}-1)\,dM_s}|\mathcal{F}_t\big] \\
&= e^{-r_t B(t,T)} e^{bB(t,T)-b(T-t)} E\big[e^{\frac{\sigma}{a}\int_t^T (e^{as-aT}-1)\,dM_s}|\mathcal{F}_t\big] \\
&= e^{-r_t B(t,T)} A(t,T),
\end{aligned}$$

where

$$A(t,T) = e^{bB(t,T)-b(T-t)} E\big[e^{\frac{\sigma}{a}\int_t^T (e^{as-aT}-1)\,dM_s}|\mathcal{F}_t\big]. \tag{11.4.9}$$

We shall compute in the following the right side expectation. From Bertoin [1], p. 8, the exponential process

$$e^{\int_0^t u(s)\,dN_s + \lambda\int_0^t (1 - e^{u(s)})\,ds}$$

is an $\mathcal{F}_t$-martingale, with $\mathcal{F}_t = \sigma(N_u; u \le t)$. This implies

$$E\Big[e^{\int_t^T u(s)\,dN_s}|\mathcal{F}_t\Big] = e^{-\lambda\int_t^T (1-e^{u(s)})\,ds}.$$

Using $dM_t = dN_t - \lambda dt$ yields

$$E\Big[e^{\int_t^T u(s)\,dM_s}|\mathcal{F}_t\Big] = E\Big[e^{\int_t^T u(s)\,dN_s}|\mathcal{F}_t\Big]e^{-\lambda\int_t^T u(s)\,ds} = e^{-\lambda\int_t^T (1+u(s)-e^{u(s)})\,ds}.$$

Let $u(s) = \frac{\sigma}{a}(e^{a(s-T)} - 1)$ and substitute in (11.4.9); then after changing the variable of integration we obtain

$$\begin{aligned}
A(t,T) &= e^{bB(t,T)-b(T-t)}E\big[e^{\frac{\sigma}{a}\int_t^T (e^{as-aT}-1)\,dM_s}|\mathcal{F}_t\big] \\
&= e^{bB(t,T)-b(T-t)}e^{\lambda\int_0^{T-t}[1+\frac{\sigma}{a}(e^{-ax}-1)-e^{\frac{\sigma}{a}(e^{-ax}-1)}]\,dx} \\
&= e^{bB(t,T)-b(T-t)}e^{\lambda(T-t)+\frac{\lambda\sigma}{a}\Big(B(t,T)-(T-t)\Big)-\lambda\int_0^{T-t}e^{\frac{\sigma}{a}(e^{-ax}-1)}\,dx} \\
&= e^{(b+\frac{\lambda\sigma}{a})B(t,T)}e^{[\lambda(1-\frac{\sigma}{a})-b](T-t)}e^{-\lambda\int_0^{T-t}e^{\frac{\sigma}{a}(e^{-ax}-1)}\,dx} \\
&= e^{(b+\frac{\lambda\sigma}{a})B(t,T)}e^{[\lambda(1-\frac{\sigma}{a})-b](T-t)}e^{-\lambda e^{-\sigma a}\int_0^{T-t}e^{\frac{\sigma}{a}e^{-ax}}\,dx} \\
&= \exp\{(b+\frac{\lambda\sigma}{a})B(t,T) + [\lambda(1-\frac{\sigma}{a}) - b](T-t) - \lambda e^{-\sigma a}\int_0^{T-t} e^{\frac{\sigma}{a}e^{-ax}}\,dx\}.
\end{aligned} \tag{11.4.10}$$

■

Evaluation using Special Functions The solution of the initial value problem

$$\begin{aligned}
f'(x) &= \frac{e^x}{x},\ x > 0 \\
f(0) &= -\infty
\end{aligned}$$

is the exponential integral function $f(x) = Ei(x)$, $x > 0$. This is a special function that can be evaluated numerically in MATHEMATICA by calling the function ExpIntegralEi[x]. For instance, for any $0 < \alpha < \beta$

$$\int_\alpha^\beta \frac{e^t}{t}\,dt = Ei(\beta) - Ei(\alpha).$$

The reader can find more details regarding the exponential integral function in Abramovitz and Stegun [10]. The last integral in the expression of $A(t,T)$ can be evaluated using this special function. Substituting $t = \frac{\sigma}{a}e^{-ax}$ we have

$$\begin{aligned}
\int_0^{T-t} e^{\frac{\sigma}{a}e^{-ax}}\,dx &= \frac{1}{a}\int_{\frac{\sigma}{a}e^{-a(T-t)}}^{\frac{\sigma}{a}} \frac{e^t}{t}\,dt \\
&= \frac{1}{a}\Big[Ei\Big(\frac{\sigma}{a}\Big) - Ei\Big(\frac{\sigma}{a}e^{-a(T-t)}\Big)\Big].
\end{aligned}$$

11.5 The Case of a Model with pure Jumps

Consider the spot rate r_t satisfying the stochastic differential equation

$$dr_t = \sigma dM_t, \tag{11.5.11}$$

where σ is a positive constant denoting the volatility of the rate. This model is obtained when the rate at which r_t is pulled toward b is $a = 0$, so there is no mean reverting effect. This type of behavior can be noticed during a short time in a highly volatile market; in this case the behavior of r_t is most influenced by the jumps.

The solution of (11.5.11) is

$$r_t = r_0 + \sigma M_t = r_0 - \lambda\sigma t + \sigma N_t,$$

which is an $\mathcal{F}_t$-martingale. The rate r_t has jumps of size σ that occur at the arrival times t_k, which are exponentially distributed.

Next we shall compute the value $P(t,T)$ at time t of a zero-coupon bond that pays the amount of \$1 at maturity T. This is given by the conditional expectation

$$P(t,T) = E[e^{-\int_t^T r_s\,ds}|\mathcal{F}_t].$$

Integrating between t and s in equation (11.5.11) yields

$$r_s = r_t + \sigma(M_s - M_t), \qquad t < s < T.$$

And then

$$\int_t^T r_s\,ds = r_t(T-t) + \sigma\int_t^T (M_s - M_t)\,ds.$$

Taking out the predictable part and dropping the independent condition yields

$$\begin{aligned}
P(t,T) &= E[e^{-\int_t^T r_s\,ds}|\mathcal{F}_t] \\
&= e^{-r_t(T-t)}E[e^{-\sigma\int_t^T (M_s-M_t)\,ds}|\mathcal{F}_t] \\
&= e^{-r_t(T-t)}E[e^{-\sigma\int_0^{T-t} M_\tau\,d\tau}] \\
&= e^{-r_t(T-t)}E[e^{-\sigma\int_0^{T-t} (N_\tau-\lambda\tau)\,d\tau}] \\
&= e^{-r_t(T-t)+\frac{1}{2}\sigma\lambda(T-t)^2}E[e^{-\sigma\int_0^{T-t} N_\tau\,d\tau}].
\end{aligned} \tag{11.5.12}$$

We need to work out the expectation. Using integration by parts,

$$\begin{aligned}
-\sigma\int_0^T N_t\,dt &= -\sigma\Big(TN_T - \int_0^T t\,dN_t\Big) \\
&= -\sigma\Big(T\int_0^T dN_t - \int_0^T t\,dN_t\Big) \\
&= -\sigma\int_0^T (T-t)\,dN_t,
\end{aligned}$$

so

$$e^{-\sigma\int_0^T N_t\,dt} = e^{-\sigma\int_0^T (T-t)\,dN_t}.$$

Using the formula

$$E\Big[e^{\int_0^T u(t)\,dN_t}\Big] = e^{-\lambda\int_0^T (1-e^{u(t)})\,dt}$$

we have

$$\begin{aligned}
E\Big[e^{-\sigma\int_0^T N_t\,dt}\Big] &= E\Big[e^{-\sigma\int_0^T (T-t)\,dN_t}\Big] = e^{-\lambda\int_0^T (1-e^{-\sigma(T-t)})\,dt} \\
&= e^{-\lambda T}e^{-\frac{\lambda}{\sigma}(e^{-\sigma T}-1)} \\
&= e^{-\lambda\left(T+\frac{1}{\sigma}(e^{-\sigma T}-1)\right)}.
\end{aligned}$$

Replacing T by $T-t$ and t by τ yields

$$E\Big[e^{-\sigma\int_0^{T-t} N_\tau\,d\tau}\Big] = e^{-\lambda\left(T-t+\frac{1}{\sigma}(e^{-\sigma(T-t)}-1)\right)}.$$

Substituting in (11.5.12) yields the formula for the bond price

$$\begin{aligned}
P(t,T) &= e^{-r_t(T-t)+\frac{1}{2}\sigma\lambda(T-t)^2}e^{-\lambda\left(T-t+\frac{1}{\sigma}(e^{-\sigma(T-t)}-1)\right)} \\
&= \exp\{-(\lambda+r_t)(T-t)+\frac{1}{2}\sigma\lambda(T-t)^2-\frac{\lambda}{\sigma}\big(e^{-\sigma(T-t)}-1\big)\}.
\end{aligned}$$

Proposition 11.5.1 *Assume the spot rate r_t satisfies $dr_t = \sigma dM_t$, with $\sigma > 0$ constant. Then the price of the zero-coupon bond that pays off \$1 at time T is given by*

$$P(t,T) = \exp\{-(\lambda+r_t)(T-t)+\frac{1}{2}\sigma\lambda(T-t)^2-\frac{\lambda}{\sigma}\big(e^{-\sigma(T-t)}-1\big)\}.$$

The yield curve is given by

$$\begin{aligned}
R(t,T) &= -\frac{1}{T-t}\ln P(t,T) \\
&= r_t+\lambda-\frac{\lambda\sigma}{2}(T-t)+\frac{\lambda}{\sigma(T-t)}(e^{-\sigma(T-t)}-1).
\end{aligned}$$

Exercise 11.5.2 *The price of an interest rate derivative security with maturity time T and payoff $P_T(r_T)$ has the price at time $t=0$ given by $P_0 = E_0[e^{-\int_0^T r_s\,ds}P_T(r_T)]$. (The zero-coupon bond is obtained for $P_T(r_T)=1$). Find the price of an interest rate derivative with the payoff $P_T(r_T)=r_T$.*

Exercise 11.5.3 *Assume the spot rate satisfies the equation $dr_t = \sigma r_t dM_t$, where M_t is the compensated Poisson process. Find a solution of the form $r_t = r_0 e^{\phi(t)}(1+\sigma)^{N_t}$, where N_t denotes the Poisson process.*

Chapter 12

Modeling Stock Prices

The price of a stock can be modeled by a continuous stochastic process which is the sum of a predictable and an unpredictable part. However, this type of model does not take into account market crashes. If those are to be taken into consideration, the stock price needs to contain a third component which models unexpected jumps. We shall discuss these models in the present chapter.

12.1 Constant Drift and Volatility Model

Let S_t denote the price of a stock at time t. If $\mathcal{F}_t$ denotes the information set at time t, then S_t is a continuous process that is $\mathcal{F}_t$-adapted. The return on the stock during the time interval Δt measures the percentage increase in the stock price between instances t and $t + \Delta t$ and is given by $\dfrac{S_{t+\Delta t} - S_t}{S_t}$. When Δt is infinitesimally small, we obtain the instantaneous return, $\dfrac{dS_t}{S_t}$. This is supposed to be the sum of two components:

- the predictable part μdt
- the noisy part due to unexpected news σdW_t.

Adding these parts yields

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t,$$

which leads to the stochastic equation

$$\boxed{dS_t = \mu S_t dt + \sigma S_t dW_t.} \qquad (12.1.1)$$

The parameters μ and σ are positive constants which represent the *drift* and *volatility* of the stock. This equation has been solved in Example 7.7.2 by applying the method of variation of parameters. The solution is

$$\boxed{S_t = S_0 e^{(\mu - \frac{\sigma^2}{2})t + \sigma W_t},} \qquad (12.1.2)$$

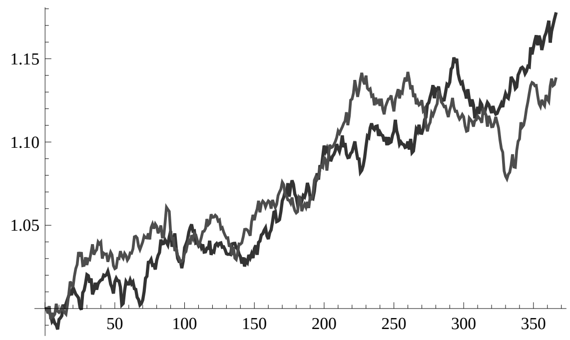


Figure 12.1: *Two distinct simulations for the stochastic equation* $dS_t = 0.15 S_t dt + 0.07 S_t dW_t$*, with* $S_0 = 1$.

where S_0 denotes the price of the stock at time $t = 0$. It is worth noting that the stock price is $\mathcal{F}_t$-adapted, positive, and it has a log-normal distribution. Using Exercise 1.6.2, the mean and variance are

$$E[S_t] = S_0 e^{\mu t} \tag{12.1.3}$$
$$Var[S_t] = S_0^2 e^{2\mu t}(e^{\sigma^2 t} - 1). \tag{12.1.4}$$

See Fig.12.1 for two simulations of the stock price.

Exercise 12.1.1 *Let* $\mathcal{F}_u$ *be the information set at time* u*. Find* $E[S_t|\mathcal{F}_u]$ *and* $Var[S_t|\mathcal{F}_u]$ *for* $u \leq t$.[1] *How these formulas become in the case* $s = t$*?*

Exercise 12.1.2 *Find the stochastic process followed by* $\ln S_t$*. What are the values of* $E[\ln(S_t)]$ *and* $Var[\ln(S_t)]$*?*

Exercise 12.1.3 *Find the stochastic differential equations associated with the following processes*

$$(a)\ \frac{1}{S_t} \qquad (b)\ S_t^n \qquad (c)\ (S_t - 1)^2.$$

Exercise 12.1.4 (a) *Show that* $E[S_t^2] = S_0^2 e^{(2\mu+\sigma^2)t}$.
(b) *Find a similar formula for* $E[S_t^n]$*, with* n *positive integer.*

Exercise 12.1.5 (a) *Find the expectation* $E[S_t W_t]$.
(b) *Find the correlation function* $\rho_t = Corr(S_t, W_t)$*. What happens for* t *large?*

[1] The conditional variance is defined by $Var(X|\mathcal{F}) = E[X^2|\mathcal{F}] - E[X|\mathcal{F}]^2$.

The next result deals with the probability that the stock price reaches a certain barrier before another barrier.

Theorem 12.1.6 *Let S_u and S_d be fixed, such that $S_d < S_0 < S_u$. The probability that the stock price S_t hits the upper value S_u before the lower value S_d is*

$$p = \frac{d^\gamma - 1}{d^\gamma - u^\gamma},$$

where $S_u/S_0 = u$, $S_d/S_0 = d$, and $\gamma = 1 - 2\mu/\sigma^2$.

Proof: Let $X_t = mt + W_t$. Theorem 3.5.1 provides

$$P(X_t \text{ goes up to } \alpha \text{ before down to } -\beta) = \frac{e^{2m\beta} - 1}{e^{2m\beta} - e^{-2m\alpha}}. \tag{12.1.5}$$

Choosing the following values for the parameters

$$m = \frac{\mu}{\sigma} - \frac{\sigma}{2}, \quad \alpha = \frac{\ln u}{\sigma}, \quad \beta = -\frac{\ln d}{\sigma},$$

we have the sequence of identities

$$\begin{aligned}
& P(X_t \text{ goes up to } \alpha \text{ before down to } -\beta) \\
= \; & P(\sigma X_t \text{ goes up to } \sigma\alpha \text{ before down to } -\sigma\beta) \\
= \; & P(S_0 e^{\sigma X_t} \text{ goes up to } S_0 e^{\sigma\alpha} \text{ before down to } S_0 e^{-\sigma\beta}) \\
= \; & P(S_t \text{ goes up to } S_u \text{ before down to } S_d).
\end{aligned}$$

Using (12.1.5) yields

$$P(S_t \text{ goes up to } S_u \text{ before down to } S_d) = \frac{e^{2m\beta} - 1}{e^{2m\beta} - e^{-2m\alpha}}. \tag{12.1.6}$$

Since a computation shows that

$$\begin{aligned}
e^{2m\beta} &= e^{-(\frac{2\mu}{\sigma^2} - 1)\ln d} = d^{1 - \frac{2\mu}{\sigma^2}} = d^\gamma \\
e^{-2m\alpha} &= e^{(-\frac{2\mu}{\sigma^2} + 1)\ln u} = u^{1 - \frac{2\mu}{\sigma^2}} = u^\gamma,
\end{aligned}$$

formula (12.1.6) becomes

$$\boxed{P(S_t \text{ goes up to } S_u \text{ before down to } S_d) = \frac{d^\gamma - 1}{d^\gamma - u^\gamma},}$$

which ends the proof. ∎

Corollary 12.1.7 *Let $S_u > S_0 > 0$ be fixed. Then*

$$\boxed{P(S_t \text{ hits } S_u) = \Big(\frac{S_0}{S_u}\Big)^{1-\frac{2\mu}{\sigma^2}} \quad \text{for some } \; t > 0.}$$

Proof: Taking $d = 0$ implies $S_d = 0$. Since S_t never reaches zero,

$$\begin{aligned} P(S_t \text{ hits } S_u) &= P(S_t \text{ goes up to } S_u \text{ before down to } S_d = 0) \\ &= \frac{d^\gamma - 1}{d^\gamma - u^\gamma}\bigg|_{d=0} = \frac{1}{u^\gamma} = \Big(\frac{S_0}{S_u}\Big)^{1-\frac{2\mu}{\sigma^2}}. \end{aligned}$$

■

Exercise 12.1.8 *A stock has $S_0 = \$10$, $\sigma = 0.15$, $\mu = 0.20$. What is the probability that the stock goes up to $\$15$ before it goes down to $\$5$?*

Exercise 12.1.9 *Let $0 < S_0 < S_u$. What is the probability that S_t hits S_u for some time $t > 0$?*

The next result deals with the probability of a stock to reach a certain barrier.

Proposition 12.1.10 *Let $b > 0$. The probability that the stock S_t will ever reach the level b is*

$$P(\sup_{t\geq 0} S_t \geq b) = \begin{cases} \Big(\dfrac{b}{S_0}\Big)^{\frac{2\mu}{\sigma^2}-1}, & \text{if } r < \frac{\sigma^2}{2} \\ \\ 1, & \text{if } r \geq \frac{\sigma^2}{2} \end{cases} \tag{12.1.7}$$

Proof: Using $S_t = S_0 e^{(r-\frac{\sigma^2}{2})t + \sigma W_t}$ we have

$$\begin{aligned} & P(S_t \geq b \text{ for some } t \geq 0) = P\Big(S_0 e^{(r-\frac{\sigma^2}{2})t+\sigma W_t} \geq b, \text{ for some } t \geq 0\Big) \\ = \;& P\Big(\sigma W_t + (r - \frac{\sigma^2}{2})t \geq \ln \frac{b}{S_0}, \text{ for some } t \geq 0\Big) \\ = \;& P\Big(W_t - \gamma t \geq \alpha, \text{ for some } t \geq 0\Big), \end{aligned}$$

where $\alpha = \dfrac{1}{\sigma}\ln\dfrac{b}{S_0}$ and $\gamma = \dfrac{\sigma}{2} - \dfrac{r}{\sigma}$. Using Exercise 3.5.2 the previous probability is equal to

$$P\Big(W_t - \gamma t \geq \alpha, \text{ for some } t \geq 0\Big) = \begin{cases} 1, & \text{if } \gamma \leq 0 \\ e^{-2\alpha\gamma}, & \text{if } \gamma > 0, \end{cases}$$

which is equivalent to (12.1.7). ■

12.2 When Does the Stock Reach a Certain Barrier?

Let $a > S_0$ and consider the first time when the stock reaches the barrier a

$$T_a = \inf\{t > 0; S_t = a\}.$$

Since we have

$$\begin{aligned}
\{S_t = a\} &= \{S_0 e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma W_t} = a\} \\
&= \{(\mu - \frac{1}{2}\sigma^2)t + \sigma W_t = \ln(a/S_0)\} \\
&= \{\alpha t + \sigma W_t = x\},
\end{aligned}$$

with $x = \ln(a/S_0)$ and $\alpha = \mu - \frac{1}{2}\sigma^2$, using Proposition 3.6.5 (a) it follows that

$$T_a = \inf\{t > 0; \alpha t + \sigma W_t = x\}$$

has the inverse gaussian probability density

$$\begin{aligned}
p(\tau) &= \frac{x}{\sqrt{2\pi\sigma^3\tau^3}} e^{-\frac{(x-\alpha\sigma\tau)^2}{2\tau\sigma^3}} \\
&= \frac{\ln(a/S_0)}{\sqrt{2\pi\sigma^3\tau^3}} e^{-\left(\ln(a/S_0) - \alpha\sigma\tau\right)^2/(2\tau\sigma^3)}.
\end{aligned}$$

The mean and the variance of the hitting time T_a are given by Proposition 3.6.5 (b)

$$\begin{aligned}
E[T_a] &= \frac{x}{\sigma\alpha} = \frac{\ln(a/S_0)}{\sigma(\mu - \frac{1}{2}\sigma^2)} \\
Var(T_a) &= \frac{x}{\alpha^3} = \frac{\ln(a/S_0)}{(\mu - \frac{1}{2}\sigma^2)^3}.
\end{aligned}$$

Exercise 12.2.1 *Consider the doubling time of a stock* $T_2 = \inf\{t > 0; S_t = 2S_0\}$.

(a) *Find* $E[T_2]$ *and* $Var(T_2)$. *Do these values depend of* S_0*?*

(b) *The expected return of a stock is* $\mu = 0.15$ *and its volatility* $\sigma = 0.20$. *Find the expected time when the stock dubles its value.*

Exercise 12.2.2 *Let* $\overline{S}_t = \max\limits_{u \le t} S_u$ *and* $\underline{S}_t = \min\limits_{u \le t} S_u$ *be the running maximum and minimum of the stock. Find the distribution functions of* $\overline{S}_t$ *and* $\underline{S}_t$.

Exercise 12.2.3 (a) *What is the probability that the stock* S_t *reaches level* a, $a > S_0$, *before time* T*?*

(b) *What is the probability that the stock* S_t *reaches level* a, $a > S_0$, *before time* T_2 *and after time* T_1*?*

We shall use the properties of the hitting time later when pricing perpetual lookback options and rebate options.

12.3 Time-dependent Drift and Volatility Model

This model considers the drift $\mu = \mu(t)$ and volatility $\sigma = \sigma(t)$ to be deterministic functions of time. In this case the equation (12.1.1) becomes

$$\boxed{dS_t = \mu(t)S_t dt + \sigma(t)S_t dW_t.} \tag{12.3.8}$$

We shall solve the equation using the method of integrating factors presented in section 7.8. Multiplying by the integrating factor

$$\rho_t = e^{-\int_0^t \sigma(s)\,dW_s + \frac{1}{2}\int_0^t \sigma^2(s)\,ds}$$

the equation (12.3.8) becomes $d(\rho_t S_t) = \rho_t \mu(t) S_t\,dt$. Substituting $Y_t = \rho_t S_t$ yields the deterministic equation $dY_t = \mu(t)Y_t\,dt$ with the solution

$$Y_t = Y_0 e^{\int_0^t \mu(s)\,ds}.$$

Substituting back $S_t = \rho_t^{-1} Y_t$, we obtain the closed-form solution of equation (12.3.8)

$$\boxed{S_t = S_0 e^{\int_0^t (\mu(s) - \frac{1}{2}\sigma^2(s))\,ds + \int_0^t \sigma(s)\,dW_s}.}$$

Proposition 12.3.1 *The solution S_t is $\mathcal{F}_t$-adapted and log-normally distributed, with mean and variance given by*

$$\begin{aligned} E[S_t] &= S_0 e^{\int_0^t \mu(s)\,ds} \\ Var[S_t] &= S_0^2 e^{2\int_0^t \mu(s)\,ds}\big(e^{\int_0^t \sigma^2(s)\,ds} - 1\big). \end{aligned}$$

Proof: Let $X_t = \int_0^t (\mu(s) - \frac{1}{2}\sigma^2(s))\,ds + \int_0^t \sigma(s)\,dW_s$. Since X_t is a sum of a predictable integral function and a Wiener integral, it is normally distributed, see Proposition 4.6.1, with

$$\begin{aligned} E[X_t] &= \int_0^t \big(\mu(s) - \frac{1}{2}\sigma^2(s)\big)\,ds \\ Var[X_t] &= Var\Big[\int_0^t \sigma(s)\,dW_s\Big] = \int_0^t \sigma^2(s)\,ds. \end{aligned}$$

Using Exercise 1.6.2, the mean and variance of the log-normal random variable $S_t = S_0 e^{X_t}$ are given by

$$\begin{aligned} E[S_t] &= S_0 e^{\int_0^t (\mu - \frac{\sigma^2}{2})\,ds + \frac{1}{2}\int_0^t \sigma^2\,ds} = S_0 e^{\int_0^t \mu(s)\,ds} \\ Var[S_t] &= S_0^2 e^{2\int_0^t (\mu - \frac{1}{2}\sigma^2)\,ds + \int_0^t \sigma^2\,ds}\big(e^{\int_0^t \sigma^2} - 1\big) \\ &= S_0^2 e^{2\int_0^t \mu(s)\,ds}\big(e^{\int_0^t \sigma^2(s)\,ds} - 1\big). \end{aligned}$$

$\blacksquare$

If the average drift and average squared volatility are defined as

$$\begin{aligned}\overline{\mu} &= \frac{1}{t}\int_0^t \mu(s)\,ds \\ \overline{\sigma}^2 &= \frac{1}{t}\int_0^t \sigma^2(s)\,ds,\end{aligned}$$

the aforementioned formulas can also be written as

$$\begin{aligned}E[S_t] &= S_0 e^{\overline{\mu}t} \\ Var[S_t] &= S_0^2 e^{2\overline{\mu}t}(e^{\overline{\sigma}^2 t} - 1).\end{aligned}$$

It is worth noting that we have obtained formulas similar to (12.1.3)–(12.1.4).

12.4 Models for Stock Price Averages

In this section we shall provide stochastic differential equations for several types of averages on stocks. These averages are used as underlying assets in the case of Asian options. The most common type of average is the arithmetic one, which is used in the definition of the stock index.

Let $S_{t_1}, S_{t_2}, \cdots, S_{t_n}$ be a sample of stock prices at n instances of time $t_1 < t_2 < \cdots < t_n$. The most common types of discrete averages are:

- The *arithmetic average*

$$A(t_1, t_2, \cdots, t_n) = \frac{1}{n}\sum_{k=1}^{n} S_{t_k}.$$

- The *geometric average*

$$G(t_1, t_2, \cdots, t_n) = \Big(\prod_{k=1}^{n} S_{t_k}\Big)^{\frac{1}{n}}.$$

- The *harmonic average*

$$H(t_1, t_2, \cdots, t_n) = \frac{n}{\displaystyle\sum_{k=1}^{n} \frac{1}{S_{t_k}}}.$$

The well-known inequality of means states that we always have

$$H(t_1, t_2, \cdots, t_n) \le G(t_1, t_2, \cdots, t_n) \le A(t_1, t_2, \cdots, t_n), \tag{12.4.9}$$

with identity in the case of constant stock prices.

In the following we shall obtain expressions for continuous sampled averages and find their associated stochastic equations.

The continuously sampled arithmetic average
Let $t_n = t$ and assume $t_{k+1} - t_k = \frac{t}{n}$. Using the definition of the integral as a limit of Riemann sums, we have

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} S_{t_k} = \lim_{n\to\infty} \frac{1}{t} \sum_{k=1}^{n} S_{t_k} \frac{t}{n} = \frac{1}{t} \int_0^t S_u \, du.$$

It follows that the *continuously sampled arithmetic average* of stock prices between 0 and t is given by

$$\boxed{A_t = \frac{1}{t} \int_0^t S_u \, du.}$$

Using the formula for S_t we obtain

$$A_t = \frac{S_0}{t} \int_0^t e^{(\mu - \sigma^2/2)u + \sigma W_u} \, du.$$

This integral can be computed explicitly only in the case $\sigma = 0$. It is worth noting that A_t is neither normal nor log-normal, a fact that makes the price of Asian options on arithmetic averages hard to evaluate.

Let $I_t = \int_0^t S_u \, du$. The Fundamental Theorem of Calculus implies $dI_t = S_t \, dt$. Then the quotient rule yields

$$dA_t = d\Big(\frac{I_t}{t}\Big) = \frac{dI_t \, t - I_t \, dt}{t^2} = \frac{S_t t \, dt - I_t dt}{t^2} = \frac{1}{t}(S_t - A_t)dt,$$

i.e. the continuous arithmetic average A_t satisfies

$$\boxed{dA_t = \frac{1}{t}(S_t - A_t)dt.}$$

If $A_t < S_t$, the right side is positive and hence $dA_t > 0$, i.e. the average A_t goes up. Similarly, if $A_t > S_t$, then the average A_t goes down. This shows that the average A_t tends to trace the stock values S_t.

By l'Hospital's rule we have

$$A_0 = \lim_{t\to 0} \frac{I_t}{t} = \lim_{t\to 0} S_t = S_0.$$

Using that the expectation commutes with integrals, we have

$$E[A_t] = \frac{1}{t} \int_0^t E[S_u] \, du = \frac{1}{t} \int_0^t S_0 e^{\mu u} \, du = S_0 \frac{e^{\mu t} - 1}{\mu t}.$$

Hence

$$E[A_t] = \begin{cases} S_0 \frac{e^{\mu t} - 1}{\mu t}, & \text{if } t > 0 \\ S_0, & \text{if } t = 0. \end{cases} \tag{12.4.10}$$

In the following we shall compute the variance $Var[A_t]$. Since

$$Var[A_t] = \frac{1}{t^2} E[I_t^2] - E[A_t]^2, \tag{12.4.11}$$

it suffices to find $E[I_t^2]$. We need first the following result:

Lemma 12.4.1 *(i) We have*

$$E[I_t S_t] = \frac{S_0^2}{\mu + \sigma^2} [e^{(2\mu+\sigma^2)t} - e^{\mu t}].$$

(ii) The processes A_t and S_t are not independent.

Proof: (i) Using Ito's formula

$$\begin{aligned} d(I_t S_t) &= dI_t\, S_t + I_t\, dS_t + dI_t\, dS_t \\ &= S_t^2 dt + I_t(\mu S_t dt + \sigma S_t dW_t) + \underbrace{S_t dt\, dS_t}_{=0} \\ &= (S_t^2 + \mu I_t S_t) dt + \sigma I_t S_t dW_t. \end{aligned}$$

Using $I_0 S_0 = 0$, integrating between 0 and t yields

$$I_t S_t = \int_0^t (S_u^2 + \mu I_u S_u)\, du + \sigma \int_0^t I_u S_u\, dW_u.$$

Since the expectation of the Ito integral is zero, we have

$$E[I_t S_t] = \int_0^t (E[S_u^2] + \mu E[I_u S_u])\, du.$$

Using Exercise 12.1.4 this becomes

$$E[I_t S_t] = \int_0^t \Big(S_0^2 e^{(2\mu+\sigma^2)u} + \mu E[I_u S_u] \Big)\, du.$$

If we denote $g(t) = E[I_t S_t]$, differentiating yields the ODE

$$g'(t) = S_0^2 e^{(2\mu+\sigma^2)t} + \mu g(t),$$

with the initial condition $g(0) = 0$. This can be solved as a linear differential equation in $g(t)$ by multiplying by the integrating factor $e^{-\mu t}$. The solution is

$$g(t) = \frac{S_0^2}{\mu + \sigma^2} [e^{(2\mu+\sigma^2)t} - e^{\mu t}].$$

(ii) Since A_t and I_t are proportional, it suffices to show that I_t and S_t are not independent. This follows from part (i) and the fact that

$$E[I_t S_t] \neq E[I_t]E[S_t] = \frac{S_0^2}{\mu}(e^{\mu t} - 1)e^{\mu t}.$$

■

Next we shall find $E[I_t^2]$. Using $dI_t = S_t dt$, then $(dI_t)^2 = 0$ and hence Ito's formula yields

$$d(I_t^2) = 2I_t\, dI_t + (dI_t)^2 = 2I_t S_t\, dt.$$

Integrating between 0 and t and using $I_0 = 0$ leads to

$$I_t^2 = 2\int_0^t I_u S_u\, du.$$

Taking the expectation and using Lemma 12.4.1 we obtain

$$E[I_t^2] = 2\int_0^t E[I_u S_u]\, du = \frac{2S_0^2}{\mu + \sigma^2}\Big[\frac{e^{(2\mu+\sigma^2)t} - 1}{2\mu + \sigma^2} - \frac{e^{\mu t} - 1}{\mu}\Big]. \tag{12.4.12}$$

Substituting into (12.4.11) yields

$$Var[A_t] = \frac{S_0^2}{t^2}\Big\{\frac{2}{\mu + \sigma^2}\Big[\frac{e^{(2\mu+\sigma^2)t} - 1}{2\mu + \sigma^2} - \frac{e^{\mu t} - 1}{\mu}\Big] - \frac{(e^{\mu t} - 1)^2}{\mu^2}\Big\}.$$

Concluding the previous calculations, we have the following result:

Proposition 12.4.2 *The arithmetic average A_t satisfies the stochastic equation*

$$dA_t = \frac{1}{t}(S_t - A_t)dt, \quad A_0 = S_0.$$

Its mean and variance are given by

$$\begin{aligned} E[A_t] &= S_0\frac{e^{\mu t} - 1}{\mu t}, \quad t > 0 \\ Var[A_t] &= \frac{S_0^2}{t^2}\Big\{\frac{2}{\mu + \sigma^2}\Big[\frac{e^{(2\mu+\sigma^2)t} - 1}{2\mu + \sigma^2} - \frac{e^{\mu t} - 1}{\mu}\Big] - \frac{(e^{\mu t} - 1)^2}{\mu^2}\Big\}. \end{aligned}$$

Exercise 12.4.3 *Find approximative formulas for $E[A_t]$ and $Var[A_t]$ for t small, up to the order $O(t^2)$. (Recall that $f(t) = O(t^2)$ if $\lim_{t\to\infty} \frac{f(t)}{t^2} = c < \infty$.)*

The continuously sampled geometric average
Dividing the interval $(0,t)$ into equal subintervals of length $t_{k+1}-t_k=\frac{t}{n}$, we have

$$\begin{aligned}G(t_1,\dots,t_n) &= \Big(\prod_{k=1}^n S_{t_k}\Big)^{1/n} = e^{\ln\Big(\prod_{k=1}^n S_{t_k}\Big)^{1/n}}\\ &= e^{\frac{1}{n}\sum_{k=1}^n \ln S_{t_k}} = e^{\frac{1}{t}\sum_{k=1}^n \ln S_{t_k}\frac{t}{n}}.\end{aligned}$$

Using the definition of the integral as a limit of Riemann sums

$$G_t = \lim_{n\to\infty}\Big(\prod_{k=1}^n S_{t_k}\Big)^{1/n} = \lim_{n\to\infty} e^{\frac{1}{t}\sum_{k=1}^n \ln S_{t_k}\frac{t}{n}} = e^{\frac{1}{t}\int_0^t \ln S_u\,du}.$$

Therefore, the *continuously sampled geometric average* of stock prices between instances 0 and t is given by

$$\boxed{G_t = e^{\frac{1}{t}\int_0^t \ln S_u\,du}.} \tag{12.4.13}$$

Theorem 12.4.4 *G_t has a log-normal distribution, with the mean and variance given by*

$$\begin{aligned}E[G_t] &= S_0 e^{(\mu-\frac{\sigma^2}{6})\frac{t}{2}}\\ Var[G_t] &= S_0^2 e^{(\mu-\frac{\sigma^2}{6})t}\big(e^{\frac{\sigma^2 t}{3}}-1\big).\end{aligned}$$

Proof: Using

$$\ln S_u = \ln\Big(S_0 e^{(\mu-\frac{\sigma^2}{2})u+\sigma W_u}\Big) = \ln S_0 + (\mu-\frac{\sigma^2}{2})u + \sigma W_u,$$

then taking the logarithm yields

$$\ln G_t = \frac{1}{t}\int_0^t \Big[\ln S_0+(\mu-\frac{\sigma^2}{2})u+\sigma W_u\Big]\,du = \ln S_0+(\mu-\frac{\sigma^2}{2})\frac{t}{2}+\frac{\sigma}{t}\int_0^t W_u\,du. \tag{12.4.14}$$

Since the integrated Brownian motion $Z_t=\int_0^t W_u\,du$ is Gaussian with $Z_t\sim N(0,t^3/3)$, it follows that $\ln G_t$ has a normal distribution

$$\ln G_t \sim N\Big(\ln S_0+(\mu-\frac{\sigma^2}{2})\frac{t}{2},\ \frac{\sigma^2 t}{3}\Big). \tag{12.4.15}$$

This implies that G_t has a log-normal distribution. Using Exercise 1.6.2, we obtain

$$\begin{aligned}E[G_t] &= e^{E[\ln G_t]+\frac{1}{2}Var[\ln G_t]} = e^{\ln S_0+(\mu-\frac{\sigma^2}{2})\frac{t}{2}+\frac{\sigma^2 t}{6}} = S_0 e^{(\mu-\frac{\sigma^2}{6})\frac{t}{2}}.\\ Var[G_t] &= e^{2E[\ln G_t]+Var[\ln G_t]}\big(e^{Var[\ln G_t]}-1\big)\\ &= e^{2\ln S_0+(\mu-\frac{\sigma^2}{6})t}\big(e^{\frac{\sigma^2 t}{3}}-1\big) = S_0^2 e^{(\mu-\frac{\sigma^2}{6})t}\big(e^{\frac{\sigma^2 t}{3}}-1\big).\end{aligned}$$

$\blacksquare$

Corollary 12.4.5 *The geometric average G_t is given by the closed-form formula*

$$\boxed{G_t = S_0 e^{(\mu - \frac{\sigma^2}{2})\frac{t}{2} + \frac{\sigma}{t}\int_0^t W_u\,du}.}$$

Proof: Take the exponential in the formula (12.4.14). ■

An important consequence of the fact that G_t is log-normal is that Asian options on geometric averages have closed-form solutions.

Exercise 12.4.6 (*a*) *Show that* $\ln G_t$ *satisfies the stochastic differential equation*

$$d(\ln G_t) = \frac{1}{t}\Big(\ln S_t - \ln G_t\Big)\,dt.$$

(*b*) *Show that* G_t *satisfies*

$$dG_t = \frac{1}{t}G_t\Big(\ln S_t - \ln G_t\Big)\,dt.$$

The continuously sampled harmonic average

Let S_{t_k} be values of a stock evaluated at the sampling dates t_k, $i = 1, \dots, n$. Their *harmonic average* is defined by

$$H(t_1, \cdots, t_n) = \frac{n}{\displaystyle\sum_{k=1}^{n} \frac{1}{S(t_k)}}.$$

Consider $t_k = \frac{kt}{n}$. Then the continuously sampled harmonic average is obtained by taking the limit as $n \to \infty$ in the aforementioned relation

$$\lim_{n\to\infty} \frac{n}{\displaystyle\sum_{k=1}^{n} \frac{1}{S(t_k)}} = \lim_{n\to\infty} \frac{t}{\displaystyle\sum_{k=1}^{n} \frac{1}{S(t_k)}\frac{t}{n}} = \frac{t}{\displaystyle\int_0^t \frac{1}{S_u}\,du}.$$

Hence, the *continuously sampled harmonic average* is defined by

$$\boxed{H_t = \frac{t}{\displaystyle\int_0^t \frac{1}{S_u}\,du}.}$$

We may also write $H_t = \dfrac{t}{I_t}$, where $I_t = \displaystyle\int_0^t \frac{1}{S_u}\,du$ satisfies

$$dI_t = \frac{1}{S_t}\,dt, \qquad I_0 = 0, \qquad d\Big(\frac{1}{I_t}\Big) = -\frac{1}{S_t I_t^2}\,dt.$$

From the l'Hospital's rule we get

$$H_0 = \lim_{t\searrow 0} H_t = \lim_{t\searrow 0} \frac{t}{I_t} = S_0.$$

Using the product rule we obtain the following:

$$\begin{aligned} dH_t &= t\, d\Big(\frac{1}{I_t}\Big) + \frac{1}{I_t}\, dt + dt\, d\Big(\frac{1}{I_t}\Big) \\ &= \frac{1}{I_t}\Big(1 - \frac{t}{S_t I_t}\Big)\, dt = \frac{1}{t} H_t \Big(1 - \frac{H_t}{S_t}\Big)\, dt, \end{aligned}$$

so

$$\boxed{dH_t = \frac{1}{t} H_t \Big(1 - \frac{H_t}{S_t}\Big)\, dt.} \qquad (12.4.16)$$

If at the instance t we have $H_t < S_t$, it follows from the equation that $dH_t > 0$, i.e. the harmonic average increases. Similarly, if $H_t > S_t$, then $dH_t < 0$, i.e H_t decreases. It is worth noting that the converses are also true. The random variable H_t is not normally distributed nor log-normally distributed.

Exercise 12.4.7 *Show that* $\dfrac{H_t}{t}$ *is a decreasing function of* t. *What is its limit as* $t \to \infty$*?*

Exercise 12.4.8 *Show that the continuous analog of inequality (12.4.9) is*

$$H_t \le G_t \le A_t.$$

Exercise 12.4.9 *Let* S_{t_k} *be the stock price at time* t_k. *Consider the power* α *of the arithmetic average of* $S_{t_k}^{\alpha}$

$$A^{\alpha}(t_1, \cdots, t_n) = \left[\frac{\sum_{k=1}^{n} S_{t_k}^{\alpha}}{n}\right]^{\alpha}.$$

(*a*) *Show that the aforementioned expression tends to*

$$A_t^{\alpha} = \Big[\frac{1}{t}\int_0^t S_u^{\alpha}\, du\Big]^{\alpha},$$

as $n \to \infty$.

(*b*) *Find the stochastic differential equation satisfied by* A_t^{α}.

(*c*) *What does* A_t^{α} *become in the particular cases* $\alpha = \pm 1$*?*

Exercise 12.4.10 *The stochastic average of stock prices between* 0 *and* t *is defined by*

$$X_t = \frac{1}{t}\int_0^t S_u\, dW_u,$$

where W_u *is a Brownian motion process.*

(*a*) *Find* dX_t, $E[X_t]$ *and* $Var[X_t]$.

(*b*) *Show that* $\sigma X_t = R_t - \mu A_t$, *where* $R_t = \dfrac{S_t - S_0}{t}$ *is the "raw average" of the stock price and* $A_t = \frac{1}{t}\int_0^t S_u\, du$ *is the continuous arithmetic average.*

12.5 Stock Prices with Rare Events

In order to model the stock price when rare events are taken into account, we shall combine the effect of two stochastic processes:

- the Brownian motion process W_t, which models regular events given by infinitesimal changes in the price, and which is a continuous process;

- the Poisson process N_t, which is discontinuous and models sporadic jumps in the stock price that corresponds to shocks in the market.

Since $E[dN_t] = \lambda dt$, the Poisson process N_t has a positive drift and we need to "compensate" by subtracting λt from N_t. The resulting process $M_t = N_t - \lambda t$ is a martingale, called the compensated Poisson process, that models unpredictable jumps of size 1 at a constant rate λ. It is worth noting that the processes W_t and M_t involved in modeling the stock price are assumed to be independent.
Let $S_{t-} = \lim\limits_{u \nearrow t} S_u$ denote the value of the stock before a possible jump occurs at time t. To set up the model, we assume the instantaneous return on the stock, $\frac{dS_t}{S_{t-}}$, to be the sum of the following three components:

- the predictable part μdt;
- the noisy part due to unexpected news σdW_t;
- the rare events part due to unexpected jumps ρdM_t,

where μ, σ and ρ are constants, corresponding to the drift rate of the stock, volatility and instantaneous return jump size.[2]

Adding yields

$$\frac{dS_t}{S_{t-}} = \mu dt + \sigma dW_t + \rho dM_t.$$

[2] In this model the jump size is constant; there are models where the jump size is a random variable, see Merton [11].

Hence, the dynamics of a stock price, subject to rare events, are modeled by the following stochastic differential equation

$$dS_t = \mu S_{t-}dt + \sigma S_{t-}dW_t + \rho S_{t-}dM_t. \tag{12.5.17}$$

It is worth noting that in the case of zero jumps, $\rho = 0$, the previous equation becomes the classical stochastic equation (12.1.1).

Using that W_t and M_t are martingales, we have

$$\begin{aligned} E[\rho S_t dM_t | \mathcal{F}_t] &= \rho S_t E[dM_t | \mathcal{F}_t] = 0, \\ E[\sigma S_t dW_t | \mathcal{F}_t] &= \sigma S_t E[dW_t | \mathcal{F}_t] = 0. \end{aligned}$$

This shows the unpredictability of the last two terms, i.e. given the information set $\mathcal{F}_t$ at time t, it is not possible to predict any future increments in the next interval of time dt. The term $\sigma S_t dW_t$ captures regular events of insignificant size, while $\rho S_t dM_t$ captures rare events of large size. The "rare events" term, $\rho S_t dM_t$, incorporates jumps proportional to the stock price and is given in terms of the Poisson process N_t as

$$\rho S_{t-}dM_t = \rho S_{t-}d(N_t - \lambda t) = \rho S_{t-}dN_t - \lambda\rho S_{t-}dt.$$

Substituting into equation (12.5.17) yields

$$dS_t = (\mu - \lambda\rho)S_{t-}dt + \sigma S_{t-}dW_t + \rho S_{t-}dN_t. \tag{12.5.18}$$

The constant λ represents the rate at which the jumps of the Poisson process N_t occur. This is the same as the rate of rare events in the market, and can be determined from historical data.

The following result provides an explicit solution for the stock price when rare events are taken into account.

Proposition 12.5.1 *The solution of the stochastic equation (12.5.18) is given by*

$$S_t = S_0 e^{(\mu - \lambda\rho - \frac{\sigma^2}{2})t + \sigma W_t}(1+\rho)^{N_t}, \tag{12.5.19}$$

where

μ *is the stock price drift rate;*
σ *is the volatility of the stock;*
λ *is the rate at which rare events occur;*
ρ *is the size of jump in the expected return when a rare event occurs.*

Proof: We shall construct first the solution and then show that it verifies the equation (12.5.18). If t_k denotes the kth jump time, then $N_{t_k} = k$. Since there are no jumps before t_1, the stock price just before this time is satisfying the stochastic differential equation

$$dS_t = (\mu - \lambda\rho)S_t dt + \sigma S_t dW_t$$

with the solution given by the usual formula

$$S_{t_1-} = S_0 e^{(\mu-\lambda\rho-\frac{\sigma^2}{2})t_1+\sigma W_{t_1}}.$$

Since $\dfrac{dS_{t_1}}{S_{t_1-}} = \dfrac{S_{t_1} - S_{t_1-}}{S_{t_1-}} = \rho$, then $S_{t_1} = (1+\rho)S_{t_1-}$. Substituting in the aforementioned formula yields

$$S_{t_1} = S_{t_1-}(1+\rho) = S_0 e^{(\mu-\lambda\rho-\frac{\sigma^2}{2})t_1+\sigma W_{t_1}}(1+\rho).$$

Since there is no jump between t_1 and t_2, a similar procedure leads to

$$\begin{aligned} S_{t_2} &= S_{t_2-}(1+\rho) = S_{t_1} e^{(\mu-\lambda\rho-\frac{\sigma^2}{2})(t_2-t_1)+\sigma(W_{t_2}-W_{t_1})}(1+\rho) \\ &= S_0 e^{(\mu-\lambda\rho-\frac{\sigma^2}{2})t_2+\sigma W_{t_2}}(1+\rho)^2. \end{aligned}$$

Inductively, we arrive at

$$S_{t_k} = S_0 e^{(\mu-\frac{\sigma^2}{2})t_k+\sigma W_{t_k}}(1+\rho)^k.$$

This formula holds for any $t \in [t_k, t_{k+1})$; replacing k by N_t yields the desired formula (12.5.19).

In the following we shall show that (12.5.19) is a solution of the stochastic differential equation (12.5.18). If denote

$$U_t = S_0 e^{(\mu-\lambda\rho-\frac{\sigma^2}{2})t+\sigma W_t}, \qquad V_t = (1+\rho)^{N_t},$$

we have $S_t = U_t V_t$ and hence

$$dS_t = V_t dU_t + U_t dV_t + dU_t dV_t. \tag{12.5.20}$$

We already know that U_t verifies

$$\begin{aligned} dU_t &= (\mu-\lambda\rho)U_t dt + \sigma U_t dW_t \\ &= (\mu-\lambda\rho)U_{t-} dt + \sigma U_{t-} dW_t, \end{aligned}$$

since U_t is a continuous process, i.e., $U_t = U_{t-}$. Then the first term of (12.5.20) becomes

$$V_t dU_t = (\mu-\lambda\rho)S_{t-} dt + \sigma S_{t-} dW_t.$$

In order to compute the second term of (12.5.20) we write

$$\begin{aligned} dV_t = V_t - V_{t-} &= (1+\rho)^{N_t} - (1+\rho)^{N_{t-}} = \begin{cases} (1+\rho)^{1+N_{t-}} - (1+\rho)^{N_{t-}}, & \text{if } t = t_k \\ (1+\rho)^{N_t} - (1+\rho)^{N_t}, & \text{if } t \neq t_k \end{cases} \\ &= \begin{cases} \rho(1+\rho)^{N_{t-}}, & \text{if } t = t_k \\ 0, & \text{if } t \neq t_k \end{cases} \\ &= \rho(1+\rho)^{N_{t-}} \begin{cases} 1, & \text{if } t = t_k \\ 0, & \text{if } t \neq t_k \end{cases} \\ &= \rho(1+\rho)^{N_{t-}} dN_t, \end{aligned}$$

so $U_t dV_t = \rho U_t (1+\rho)^{N_{t-}} dN_t = \rho S_{t-} dN_t$. Since $dt dN_t = dN_t dW_t = 0$, the last term of (12.5.20) becomes $dU_t dV_t = 0$. Substituting back into (12.5.20) yields the equation

$$dS_t = (\mu - \lambda\rho) S_{t-} dt + \sigma S_{t-} dW_t + \rho S_{t-} dN_t.$$

■

Formula (12.5.19) provides the stock price at time t if exactly N_t jumps have occurred and all jumps in the return of the stock are equal to ρ.

It is worth noting that if ρ_k denotes the jump of the instantaneous return at time t_k, a similar proof leads to the formula

$$S_t = S_0 e^{(\mu - \lambda\rho \frac{\sigma^2}{2})t + \sigma W_t} \prod_{k=1}^{N_t} (1 + \rho_k),$$

where $\rho = E[\rho_k]$. The random variables ρ_k are assumed independent and identically distributed. They are also independent of W_t and N_t. For more details the reader is referred to Merton [11].

For the following exercises S_t is given by (12.5.19).

Exercise 12.5.2 *Find $E[S_t]$ and $Var[S_t]$.*

Exercise 12.5.3 *Find $E[\ln S_t]$ and $Var[\ln S_t]$.*

Exercise 12.5.4 *Compute the conditional expectation $E[S_t|\mathcal{F}_u]$ for $u < t$.*

Remark 12.5.5 *Besides stock, the underlying asset of a derivative can be also a stock index, or foreign currency. When use the risk neutral valuation for derivatives on a stock index that pays a continuous dividend yield at a rate q, the drift rate μ is replace by $r - q$.*
In the case of foreign currency that pays interest at the foreign interest rate r_f, the drift rate μ is replace by $r - r_f$.

Chapter 13

Risk-Neutral Valuation

13.1 The Method of Risk-Neutral Valuation

This valuation method is based on the *risk-neutral valuation principle*, which states that the price of a derivative on an asset S_t is not affected by the risk preference of the market participants; so we may assume they have the same risk aversion. In this case the valuation of the derivative price f_t at time t is done as in the following:

1. Assume the expected return of the asset S_t is the risk-free rate, $\mu = r$.
2. Calculate the expected payoff of the derivative as of time t, under condition 1.
3. Discount at the risk-free rate from time T to time t.

The first two steps require considering the expectation as of time t in a risk-neutral world. This expectation is denoted by $\widehat{E}_t[\,\cdot\,]$ and has the meaning of a conditional expectation given by $E[\,\cdot\,|\mathcal{F}_t, \mu = r]$. The method states that if a derivative has the payoff f_T, its price at any time t prior to maturity T is given by

$$\boxed{f_t = e^{-r(T-t)}\hat{E}_t[f_T].}$$

The rate r is considered constant, but the method can be easily adapted for time dependent rates.

In the following we shall present explicit computations for the most common European type[1] derivative prices using the risk-neutral valuation method.

13.2 Call Option

A *call option* is a contract which gives the buyer the right of buying the stock at time T for the price K. The time T is called *maturity time* or *expiration date* and K is called the *strike price*. It is worth noting that a call option is a *right* (not an obligation!) of

[1] A derivative is of European type if can be exercised only at the expiration time.

the buyer, which means that if the price of the stock at maturity, S_T, is less than the strike price, K, then the buyer may choose not to exercise the option. If the price S_T exceeds the strike price K, then the buyer exercises the right to buy the stock, since he pays K dollars for something which worth S_T in the market. Buying at price K and selling at S_T yields a profit of $S_T - K$.

Consider a call option with maturity date T, strike price K with the underlying stock price S_t having constant volatility $\sigma > 0$. The payoff at maturity time is $f_T = \max(S_T - K, 0)$, see Fig.13.1 a. The price of the call at any prior time $0 \le t \le T$ is given by the expectation in a risk-neutral world

$$c(t) \;=\; \hat{E}_t[e^{-r(T-t)} f_T] = e^{-r(T-t)} \hat{E}_t[f_T]. \tag{13.2.1}$$

If we let $x = \ln(S_T/S_t)$, using the log-normality of the stock price in a risk-neutral world

$$S_T = S_t e^{(r-\frac{\sigma^2}{2})(T-t)+\sigma(W_T - W_t)},$$

and it follows that x has the normal distribution

$$x \sim N\Big((r - \frac{\sigma^2}{2})(T-t),\, \sigma^2(T-t)\Big).$$

Then the density function of x is

$$p(x) = \frac{1}{\sigma\sqrt{2\pi(T-t)}} e^{-\frac{\left(x-(r-\frac{\sigma^2}{2})(T-t)\right)^2}{2\sigma^2(T-t)}}.$$

We can write the expectation as

$$\begin{aligned}
\hat{E}_t[f_T] &= \hat{E}_t[\max(S_T - K, 0)] = \hat{E}_t[\max(S_t e^x - K, 0)] \\
&= \int_{-\infty}^{\infty} \max(S_t e^x - K, 0) p(x)\, dx = \int_{\ln(K/S_t)}^{\infty} (S_t e^x - K) p(x)\, dx \\
&= I_2 - I_1,
\end{aligned} \tag{13.2.2}$$

with notations

$$I_1 \;=\; \int_{\ln(K/S_t)}^{\infty} K p(x)\, dx, \qquad I_2 \;=\; \int_{\ln(K/S_t)}^{\infty} S_t e^x p(x)\, dx$$

With the substitution $y = \frac{x-(r-\frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}}$, the first integral becomes

$$\begin{aligned}
I_1 &= K \int_{\ln(K/S_t)}^{\infty} p(x)\, dx = K \int_{-d_2}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}\, dy \\
&= K \int_{-\infty}^{d_2} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}\, dy = KN(d_2),
\end{aligned}$$

where

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

and

$$N(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u} e^{-z^2/2}\, dz$$

denotes the standard normal distribution function.

Using the aforementioned substitution the second integral can be computed by completing the square

$$\begin{aligned}
I_2 &= S_t \int_{\ln(K/S_t)}^{\infty} e^x p(x)\, dx = S_t \int_{-d_2}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2 + y\sigma\sqrt{T-t} + (r - \frac{\sigma^2}{2})(T-t)}\, dx \\
&= S_t \int_{-d_2}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(y-\sigma\sqrt{T-t})^2} e^{r(T-t)}\, dy \qquad (\text{let } z = y - \sigma\sqrt{T-t}) \\
&= S_t e^{r(T-t)} \int_{-d_2-\sigma\sqrt{T-t}}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}\, dz = S_t e^{r(T-t)} \int_{-\infty}^{d_1} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}\, dz \\
&= S_t e^{r(T-t)} N(d_1),
\end{aligned}$$

where

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

Substituting back into (13.2.2) and then into (13.2.1) yields

$$\begin{aligned}
c(t) &= e^{-r(T-t)}(I_2 - I_1) = e^{-r(T-t)}[S_t e^{r(T-t)} N(d_1) - KN(d_2)] \\
&= S_t N(d_1) - K e^{-r(T-t)} N(d_2).
\end{aligned}$$

We have obtained the well known formula of Black and Scholes:

Proposition 13.2.1 *The price of a European call option at time t is given by*

$$\boxed{c(t) = S_t N(d_1) - K e^{-r(T-t)} N(d_2).}$$

Exercise 13.2.2 *Show that*

(*a*) $\displaystyle\lim_{S_t \to \infty} \frac{c(t)}{S_t} = 1;$

(*b*) $\displaystyle\lim_{S_t \to 0} c(t) = 0;$

(*c*) $c(t) \sim S_t - K e^{-r(T-t)}$ *for S_t large.*

Exercise 13.2.3 *Show that* $\dfrac{dc(t)}{dS_t} = N(d_1)$. *This expression is called the delta of a call option.*

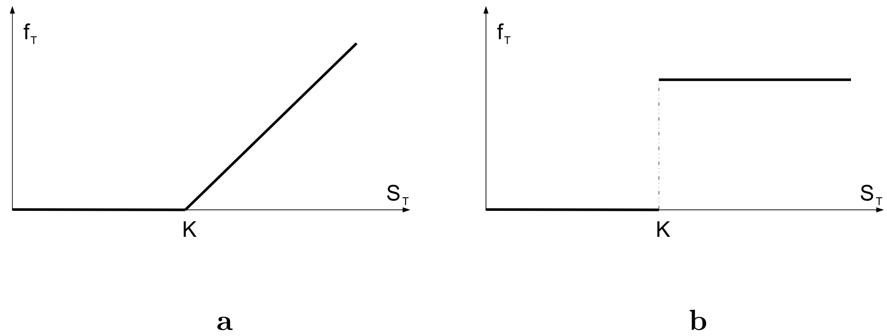


Figure 13.1: **a** *The payoff of a call option;* **b** *The payoff of a cash-or-nothing contract.*

13.3 Cash-or-nothing Contract

A financial security that pays 1 dollar if the stock price $S_T \geq K$ and 0 otherwise, is called a *bet contract*, or *cash-or-nothing contract*, see Fig.13.1 b. The payoff can be written as

$$f_T = \begin{cases} 1, & \text{if } S_T \geq K \\ 0, & \text{if } S_T < K. \end{cases}$$

Substituting $S_t = e^{X_t}$, the payoff becomes

$$f_T = \begin{cases} 1, & \text{if } X_T \geq \ln K \\ 0, & \text{if } X_T < \ln K, \end{cases}$$

where X_T has the normal distribution

$$X_T \sim N\Big(\ln S_t + (\mu - \frac{\sigma^2}{2})(T-t), \sigma^2(T-t)\Big).$$

The expectation in the risk-neutral world as of time t is

$$\begin{aligned} \widehat{E}_t[f_T] &= E[f_T|\mathcal{F}_t, \mu = r] = \int_{-\infty}^{\infty} f_T(x)p(x)\,dx \\ &= \int_{\ln K}^{+\infty} \frac{1}{\sqrt{2\pi}\sigma\sqrt{T-t}} e^{-\frac{[x - \ln S_t - (r - \frac{\sigma^2}{2})(T-t)]^2}{2\sigma^2(T-t)}}\,dx \\ &= \int_{-d_2}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}\,dy = \int_{-\infty}^{d_2} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}\,dy = N(d_2), \end{aligned}$$

where we used the substitution $y = \frac{x - \ln S_t - (r - \frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}}$ and the notation

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

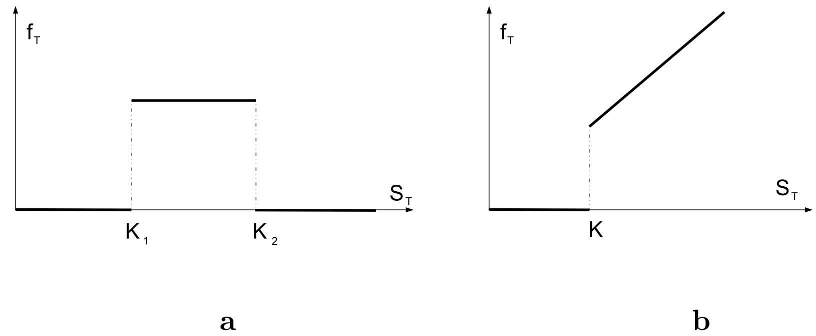


Figure 13.2: **a** *The payoff of a box-bet;* **b** *The payoff of an asset-or-nothing contract.*

The price at time t of a bet contract is

$$\boxed{f_t = e^{-r(T-t)}\widehat{E}_t[f_T] = e^{-r(T-t)}N(d_2).} \tag{13.3.3}$$

Exercise 13.3.1 *Let* $0 < K_1 < K_2$*. Find the price of a financial derivative which pays at maturity* \$1 *if* $K_1 \leq S_t \leq K_2$ *and zero otherwise, see Fig.*13.2 *a. This is a "box-bet" and its payoff is given by*

$$f_T = \begin{cases} 1, & \text{if } K_1 \leq S_T \leq K_2 \\ 0, & \text{otherwise.} \end{cases}$$

Exercise 13.3.2 *An asset-or-nothing contract pays* S_T *if* $S_T > K$ *at maturity time* T*, and pays* 0 *otherwise, see Fig.*13.2 *b. Show that the price of the contract at time* t *is* $f_t = S_t N(d_1)$.

Exercise 13.3.3 (a) *Find the price at time* t *of a derivative which pays at maturity*

$$f_T = \begin{cases} S_T^n, & \text{if } S_T \geq K \\ 0, & \text{otherwise.} \end{cases}$$

(b) *Show that the value of the contract can be written as* $f_t = g_t N(d_2 + n\sigma\sqrt{T-t})$*, where* g_t *is the value at time* t *of a power contract at time* t *given by* (13.5.5).
(c) *Recover the result of Exercise* 13.3.2 *in the case* $n = 1$.

13.4 Log-contract

A financial security that pays at maturity $f_T = \ln S_T$ is called a *log-contract*. Since the stock is log-normally distributed,

$$\ln S_T \sim N\Big(\ln S_t + (\mu - \frac{\sigma^2}{2})(T-t), \sigma^2(T-t)\Big),$$

the risk-neutral expectation at time t is

$$\widehat{E}_t[f_T] = E[\ln S_T|\mathcal{F}_t, \mu = r] = \ln S_t + (r - \frac{\sigma^2}{2})(T-t),$$

and hence the price of the log-contract is given by

$$\boxed{f_t = e^{-r(T-t)}\widehat{E}_t[f_T] = e^{-r(T-t)}\Big(\ln S_t + (r - \frac{\sigma^2}{2})(T-t)\Big).} \qquad (13.4.4)$$

Exercise 13.4.1 *Find the price at time t of a square log-contract whose payoff is given by $f_T = (\ln S_T)^2$.*

13.5 Power-contract

The financial derivative which pays at maturity the nth power of the stock price, S_T^n, is called a *power contract*. Since S_T has has a log-normal distribution with

$$S_T = S_t e^{(\mu - \frac{\sigma^2}{2})(T-t) + \sigma(W_T - W_t)},$$

the nth power of the stock, S_T^n, is also log-normally distributed, with

$$g_T = S_T^n = S_t^n e^{n(\mu - \frac{\sigma^2}{2})(T-t) + n\sigma(W_T - W_t)}.$$

Then the expectation at time t in the risk-neutral world is

$$\begin{aligned}
\widehat{E}_t[g_T] &= E[g_T|\mathcal{F}_t, \mu = r] = S_t^n e^{n(r - \frac{\sigma^2}{2})(T-t)} E[e^{n\sigma(W_T - W_t)}|\mathcal{F}_t] \\
&= S_t^n e^{n(r - \frac{\sigma^2}{2})(T-t)} E[e^{n\sigma(W_T - W_t)}] = S_t^n e^{n(r - \frac{\sigma^2}{2})(T-t)} E[e^{n\sigma W_{T-t}}] \\
&= S_t^n e^{n(r - \frac{\sigma^2}{2})(T-t)} e^{\frac{1}{2}n^2\sigma^2(T-t)}.
\end{aligned}$$

The price of the power-contract is obtained by discounting to time t

$$\begin{aligned}
g_t &= e^{-r(T-t)} E[g_T|\mathcal{F}_t, \mu = r] = S_t^n e^{-r(T-t)} e^{n(r - \frac{\sigma^2}{2})(T-t)} e^{\frac{1}{2}n^2\sigma^2(T-t)} \\
&= S_t^n e^{(n-1)(r + \frac{n\sigma^2}{2})(T-t)}.
\end{aligned}$$

Hence the value of a power contract at time t is given by

$$\boxed{g_t = S_t^n e^{(n-1)(r + \frac{n\sigma^2}{2})(T-t)}.} \qquad (13.5.5)$$

It is worth noting that if $n = 1$, i.e. if the payoff is $g_T = S_T$, then the price of the contract at any time $t \le T$ is $g_t = S_t$, i.e. the stock price itself. This will be used shortly when valuing forward contracts.
In the case $n = 2$, i.e. if the contract pays S_T^2 at maturity, then the price is $g_t = S_t^2 e^{(r+\sigma^2)(T-t)}$.

Exercise 13.5.1 *Let $n \ge 1$ be an integer. Find the price of a power call option whose payoff is given by $f_T = \max(S_T^n - K, 0)$.*

13.6 Forward Contract on the Stock

A forward contract pays at maturity the difference between the stock price, S_T, and the delivery price of the asset, K. The price at time t is

$$f_t = e^{-r(T-t)}\widehat{E}_t[S_T - K] = e^{-r(T-t)}\widehat{E}_t[S_T] - e^{-r(T-t)}K = S_t - e^{-r(T-t)}K,$$

where we used that K is a constant and that $\widehat{E}_t[S_T] = S_t e^{r(T-t)}$.

Exercise 13.6.1 *Let $n \in \{2, 3\}$. Find the price of the power forward contract that pays at maturity $f_T = (S_T - K)^n$.*

13.7 The Superposition Principle

If the payoff of a derivative, f_T, can be written as a linear combination of payoffs

$$f_T = \sum_{i=1}^{n} c_i h_{i,T}$$

with c_i constants, then the price at time t is given by

$$f_t = \sum_{i=1}^{n} c_i h_{i,t}$$

where $h_{i,t}$ is the price at time t of a derivative that pays at maturity $h_{i,T}$. We shall successfully use this method in the situation when the payoff f_T can be decomposed into simpler payoffs, for which we can evaluate directly the price of the associate derivative. In this case the price of the initial derivative, f_t, is obtained as a combination of the prices of the easier to valuate derivatives.

The reason underlying the aforementioned superposition principle is the linearity of the expectation operator $\hat{E}$

$$\begin{aligned} f_t &= e^{-r(T-t)}\widehat{E}[f_T] = e^{-r(T-t)}\widehat{E}[\sum_{i=1}^{n} c_i h_{i,T}] \\ &= e^{-r(T-t)}\sum_{i=1}^{n} c_i \widehat{E}[h_{i,T}] = \sum_{i=1}^{n} c_i h_{i,t}. \end{aligned}$$

This principle is also connected with the absence of arbitrage opportunities[2] in the market. Consider two portfolios of derivatives with equal values at the maturity time T

$$\sum_{i=1}^{n} c_i h_{i,T} = \sum_{j=1}^{m} a_j g_{j,T}.$$

[2]An arbitrage opportunity deals with the practice of making profits by taking simultaneous long and short positions in the market

If we take this common value to be the payoff of a derivative, f_T, then by the aforementioned principle, the portfolios have the same value at any prior time $t \leq T$

$$\sum_{i=1}^{n} c_i h_{i,t} = \sum_{j=1}^{m} a_j g_{j,t}.$$

The last identity can also result from the absence of arbitrage opportunities in the market. If there is a time t at which the identity fails, then buying the cheaper portfolio and selling the more expensive one will lead to an arbitrage profit.

The superposition principle can be used to price *package derivatives* such as spreads, straddles, strips, straps and strangles. We shall deal with these type of derivatives in proposed exercises.

13.8 General Contract on the Stock

By a *general contract* on the stock we mean a derivative that pays at maturity the amount $g_T = G(S_T)$, where G is a given analytic function. In the case $G(S_T) = S_T^n$, we have a power contract. If $G(S_T) = \ln S_T$, we have a log-contract. If choose $G(S_T) = S_T - K$, we obtain a forward contract. Since G is analytic we shall write $G(x) = \sum_{n\geq 0} c_n x^n$, where $c_n = G^{(n)}(0)/n!$.
Decomposing into power contracts and using the superposition principle we have

$$\begin{aligned}
\widehat{E}_t[g_T] &= \widehat{E}_t[G(S_T)] = \widehat{E}_t\big[\sum_{n\geq 0} c_n S_T^n\big] \\
&= \sum_{n\geq 0} c_n \widehat{E}_t\big[S_T^n\big] \\
&= \sum_{n\geq 0} c_n S_t^n e^{n(r-\frac{\sigma^2}{2})(T-t)} e^{\frac{1}{2}n^2\sigma^2(T-t)} \\
&= \sum_{n\geq 0} c_n \Big[S_t e^{n(r-\frac{\sigma^2}{2})(T-t)}\Big]^n \Big[e^{\frac{1}{2}\sigma^2(T-t)}\Big]^{n^2}.
\end{aligned}$$

Hence the value at time t of a general contract is

$$\boxed{g_t = e^{-r(T-t)} \sum_{n\geq 0} c_n \Big[S_t e^{n(r-\frac{\sigma^2}{2})(T-t)}\Big]^n \Big[e^{\frac{1}{2}\sigma^2(T-t)}\Big]^{n^2}.}$$

Exercise 13.8.1 *Find the value at time t of an exponential contract that pays at maturity the amount e^{S_T}.*

13.9 Call Option

In the following we price a European call using the superposition principle. The payoff of a call option can be decomposed as

$$c_T = \max(S_T - K, 0) = h_{1,T} - Kh_{2,T},$$

with

$$h_{1,T} = \begin{cases} S_T, & \text{if } S_T \geq K \\ 0, & \text{if } S_T < K, \end{cases} \qquad h_{2,T} = \begin{cases} 1, & \text{if } S_T \geq K \\ 0, & \text{if } S_T < K. \end{cases}$$

These are the payoffs of *asset-or-nothing* and of *cash-or-nothing* derivatives. From section 13.3 and Exercise 13.3.2 we have $h_{1,t} = S_t N(d_1)$, $h_{2,t} = e^{-r(T-t)}N(d_2)$. By superposition we get the price of a call at time t

$$c_t = h_{1,t} - Kh_{2,t} = S_t N(d_1) - Ke^{-r(T-t)}N(d_2).$$

Exercise 13.9.1 (Put option) (*a*) *Consider the payoff* $h_{1,T} = \begin{cases} 1, & \text{if } S_T \leq K \\ 0, & \text{if } S_T > K. \end{cases}$ *Show that*

$$h_{1,t} = e^{-r(T-t)}N(-d_2), \qquad t \leq T.$$

(*b*) *Consider the payoff* $h_{2,T} = \begin{cases} S_T, & \text{if } S_T \leq K \\ 0, & \text{if } S_T > K. \end{cases}$ *Show that*

$$h_{2,t} = S_t N(-d_1), \qquad t \leq T.$$

(*c*) *The payoff of a put is* $p_T = \max(K - S_T, 0)$. *Verify that*

$$p_T = Kh_{1,T} - h_{2,T}$$

and use the superposition principle to find the price p_t *of a put.*

13.10 General Options on the Stock

Let G be an increasing analytic function with the inverse G^{-1} and K be a positive constant. A *general call option* is a contract with the payoff

$$f_T = \begin{cases} G(S_T) - K, & \text{if } S_T \geq G^{-1}(K) \\ 0, & \text{otherwise.} \end{cases} \tag{13.10.6}$$

We note the payoff function f_T is continuous. We shall work out the value of the contract at time t, f_t, using the superposition method. Since the payoff can be written as the linear combination

$$f_T = h_{1,T} - Kh_{2,T},$$

with

$$h_{1,T} = \begin{cases} G(S_T), & \text{if } S_T \geq G^{-1}(K) \\ 0, & \text{otherwise,} \end{cases} \qquad h_{2,T} = \begin{cases} 1, & \text{if } S_T \geq G^{-1}(K) \\ 0, & \text{otherwise,} \end{cases}$$

then

$$f_t = h_{1,t} - K h_{2,t}. \tag{13.10.7}$$

We had already computed the value $h_{2,t}$. In this case we have $h_{2,t} = e^{-r(T-t)}N(d_2^G)$, where d_2^G is obtained by replacing K with $G^{-1}(K)$ in the formula of d_2

$$d_2^G = \frac{\ln S_t - \ln\big(G^{-1}(K)\big) + (r - \frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}}. \tag{13.10.8}$$

We shall compute in the following $h_{1,t}$. Let $G(S_T) = \sum_{n\geq 0} c_n S_T^n$ with $c_n = G^{(n)}(0)/n!$. Then $h_{1,T} = \sum_{n\geq 0} c_n f_T^{(n)}$, where

$$f_T^{(n)} = \begin{cases} S_T^n, & \text{if } S_T \geq G^{-1}(K) \\ 0, & \text{otherwise.} \end{cases}$$

By Exercise 13.3.3 the price at time t for a contract with the payoff $f_T^{(n)}$ is

$$f_t^{(n)} = S_t^n e^{(n-1)(r+\frac{n\sigma^2}{2})(T-t)} N(d_2^G + n\sigma\sqrt{T-t}).$$

The value at time t of $h_{1,T}$ is given by

$$\begin{aligned} h_{1,t} &= \sum_{n\geq 0} c_n f_t^{(n)} \\ &= \sum_{n\geq 0} c_n S_t^n e^{(n-1)(r+\frac{n\sigma^2}{2})(T-t)} N(d_2^G + n\sigma\sqrt{T-t}). \end{aligned}$$

Substituting in (13.10.7) we obtain the following value at time t of a general call option with the payoff (13.10.6)

$$\boxed{f_t = \sum_{n\geq 0} c_n S_t^n e^{(n-1)(r+\frac{n\sigma^2}{2})(T-t)} N(d_2^G + n\sigma\sqrt{T-t}) - K e^{-r(T-t)} N(d_2^G),} \tag{13.10.9}$$

with $c_n = \frac{G^{(n)}}{n!}$ and d_2^G given by (13.10.8).

It is worth noting that in the case $G(S_T) = S_T$ we have $n = 1$, $d_2 = d_2^G$, $d_1 = d_2 + \sigma(T-t)$ formula (13.10.9) becomes the value of a plain vanilla call option

$$f_t = S_t N(d_1) - K e^{-r(T-t)}.$$

13.11 Packages

Packages are derivatives whose payoffs are linear combinations of payoffs of options, cash and underlying asset. They can be priced using the superposition principle. Some of these packages are used in hedging techniques.

The Bull Spread Let $0 < K_1 < K_2$. A derivative with the payoff

$$f_T = \begin{cases} 0, & \text{if } S_T \le K_1 \\ S_T - K_1, & \text{if } K_1 < S_T \le K_2 \\ K_2 - K_1, & \text{if } K_2 < S_T \end{cases}$$

is called a bull spread, see Fig.13.3 a. A market participant enters a bull spread position when the stock price is expected to increase. The payoff f_T can be written as the difference of the payoffs of two calls with strike prices K_1 and K_2:

$$f_T = c_1(T) - c_2(T),$$

$$c_1(T) = \begin{cases} 0, & \text{if } S_T \le K_1 \\ S_T - K_1, & \text{if } K_1 < S_T \end{cases} \qquad c_2(T) = \begin{cases} 0, & \text{if } S_T \le K_2 \\ S_T - K_2, & \text{if } K_2 < S_T \end{cases}$$

Using the superposition principle, the price of a bull spread at time t is

$$\begin{aligned} f_t &= c_1(t) - c_2(t) \\ &= S_t N\big(d_1(K_1)\big) - K_1 e^{-r(T-t)} N\big(d_2(K_1)\big) - \Big(S_t N\big(d_1(K_2)\big) - K_2 e^{-r(T-t)} N\big(d_2(K_2)\big)\Big) \\ &= S_t[N\big(d_1(K_1)\big) - N\big(d_1(K_2)\big)] - e^{-r(T-t)}[N\big(d_2(K_1)\big) - N\big(d_2(K_2)\big)], \end{aligned}$$

with

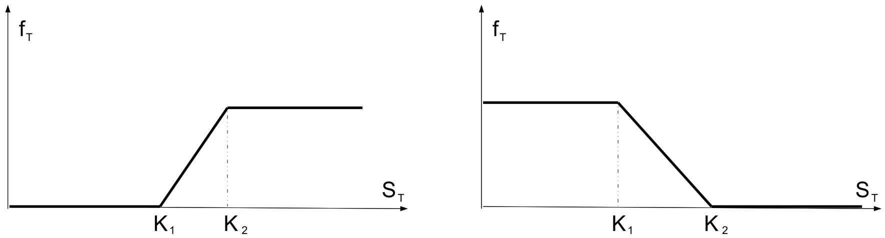
$$d_2(K_i) = \frac{\ln S_t - \ln K_i + (r - \frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}}, \qquad d_1(K_i) = d_2(K_i) + \sigma\sqrt{T-t},\ i = 1, 2.$$

The Bear Spread Let $0 < K_1 < K_2$. A derivative with the payoff

$$f_T = \begin{cases} K_2 - K_1, & \text{if } S_T \le K_1 \\ K_2 - S_T, & \text{if } K_1 < S_T \le K_2 \\ 0, & \text{if } K_2 < S_T \end{cases}$$

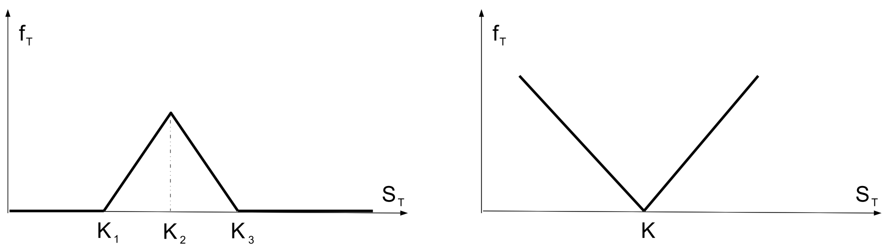
is called a bear spread, see Fig.13.3 b. A long position in this derivative leads to profits when the stock price is expected to decrease.

Exercise 13.11.1 *Find the price of a bear spread at time t, with $t < T$.*



a **b**

Figure 13.3: **a** *The payoff of a bull spread;* **b** *The payoff of a bear spread.*



a **b**

Figure 13.4: **a** *The payoff of a butterfly spread;* **b** *The payoff of a straddle.*

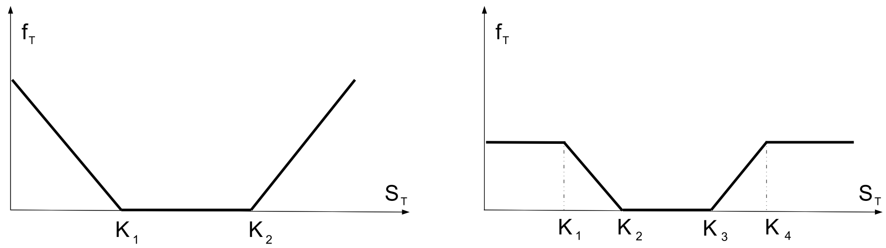
The Butterfly Spread Let $0 < K_1 < K_2 < K_3$, with $K_2 = (K_1 + K_3)/2$. A butterfly spread is a derivative with the payoff given by

$$f_T = \begin{cases} 0, & \text{if } S_T \leq K_1 \\ S_T - K_1, & \text{if } K_1 < S_T \leq K_2 \\ K_3 - S_T, & \text{if } K_2 \leq S_T < K_3 \\ 0, & \text{if } K_3 \leq S_T. \end{cases}$$

A short position in a butterfly spread leads to profits when a small move in the stock price occurs, see Fig.13.4 a.

Exercise 13.11.2 *Find the price of a butterfly spread at time t, with $t < T$.*

Straddles A derivative with the payoff $f_T = |S_T - K|$ is called a straddle, see see Fig.13.4 b. A long position in a straddle leads to profits when a move in any direction of the stock price occurs.



a **b**

Figure 13.5: **a** *The payoff for a strangle;* **b** *The payoff for a condor.*

Exercise 13.11.3 (a) *Show that the payoff of a straddle can be written as*

$$f_T = \begin{cases} K - S_T, & \text{if } S_T \leq K \\ S_T - K, & \text{if } K < S_T. \end{cases}$$

(b) *Find the price of a straddle at time t, with $t < T$.*

Strangles Let $0 < K_1 < K_2$ and $K = (K_2 + K_1)/2$, $K' = (K_2 - K_1)/2$. A derivative with the payoff $f_T = \max(|S_T - K| - K', 0)$ is called a strangle, see Fig.13.5 a. A long position in a strangle leads to profits when a small move in any direction of the stock price occurs.

Exercise 13.11.4 (a) *Show that the payoff of a strangle can be written as*

$$f_T = \begin{cases} K_1 - S_T, & \text{if } S_T \leq K_1 \\ 0, & \text{if } K_1 < S_T \leq K_2 \\ S_T - K_2, & \text{if } K_2 \leq S_T. \end{cases}$$

(b) *Find the price of a strangle at time t, with $t < T$;*
(c) *Which is cheaper: a straddle or a strangle?*

Exercise 13.11.5 *Let $0 < K_1 < K_2 < K_3 < K_4$, with $K_4 - K_3 = K_2 - K_1$. Find the value at time t of a derivative with the payoff f_T given in Fig.*13.5 *b.*

13.12 Asian Forward Contracts

Forward Contracts on the Arithmetic Average Let $A_T = \dfrac{1}{T}\displaystyle\int_0^t S_u\,du$ denote the continuous arithmetic average of the asset price between 0 and T. It sometimes makes

sense for two parties to make a contract in which one party pays the other at maturity time T the difference between the average price of the asset, A_T, and the fixed delivery price K. The payoff of a forward contract on arithmetic average is

$$f_T = A_T - K.$$

For instance, if the asset is natural gas, it makes sense to make a deal on the average price of the asset, since the price is volatile and can become expensive during the winter season.

Since the risk-neutral expectation at time $t = 0$, $\widehat{E}_0[A_T]$, is given by $E[A_T]$ where μ is replaced by r, the price of the forward contract at $t = 0$ is

$$\begin{aligned} f_0 &= e^{-rT}\widehat{E}_0[f_T] = e^{-rT}(\widehat{E}_0[A_T] - K) \\ &= e^{-rT}\Big(S_0 \frac{e^{rT} - 1}{rT} - K\Big) = S_0 \frac{1 - e^{-rT}}{rT} - e^{-rT}K \\ &= \frac{S_0}{rT} - e^{-rT}\Big(\frac{S_0}{rT} + K\Big). \end{aligned}$$

Hence the value of a forward contract on the arithmetic average at time $t = 0$ is given by

$$\boxed{f_0 = \frac{S_0}{rT} - e^{-rT}\Big(\frac{S_0}{rT} + K\Big).} \tag{13.12.10}$$

It is worth noting that the price of a forward contract on the arithmetic average is cheaper than the price of a usual forward contract on the asset. To see that, we substitute $x = rT$ in the inequality $e^{-x} > 1 - x$, $x > 0$, to get

$$\frac{1 - e^{-rT}}{rT} < 1.$$

This implies the inequality

$$S_0 \frac{1 - e^{-rT}}{rT} - e^{-rT}K < S_0 - e^{-rT}K.$$

Since the left side is the price of an Asian forward contract on the arithmetic average, while the right side is the price of a usual forward contract, we obtain the desired inequality.

Formula (13.12.10) provides the price of the contract at time $t = 0$. What is the price at any time $t < T$? One might be tempted to say that replacing T by $T - t$ and S_0 by S_t in the formula of f_0 leads to the corresponding formula for f_t. However, this does not hold true, as the next result shows:

Proposition 13.12.1 *The value at time t of a contract that pays at maturity $f_T = A_T - K$ is given by*

$$\boxed{f_t = e^{-r(T-t)}\Big(\frac{t}{T}A_t - K\Big) + \frac{1}{rT}S_t\Big(1 - e^{-r(T-t)}\Big).} \tag{13.12.11}$$

Proof: We start by computing the risk-neutral expectation $\widehat{E}_t[A_T]$. Splitting the integral into a predictable and an unpredictable part, we have

$$\begin{aligned}
\widehat{E}_t[A_T] &= \widehat{E}\Big[\frac{1}{T}\int_0^T S_u\,du|\mathcal{F}_t\Big] \\
&= \widehat{E}\Big[\frac{1}{T}\int_0^t S_u\,du + \frac{1}{T}\int_t^T S_u\,du|\mathcal{F}_t\Big] \\
&= \frac{1}{T}\int_0^t S_u\,du + \widehat{E}\Big[\frac{1}{T}\int_t^T S_u\,du|\mathcal{F}_t\Big] \\
&= \frac{1}{T}\int_0^t S_u\,du + \frac{1}{T}\int_t^T \widehat{E}[S_u|\mathcal{F}_t]\,du \\
&= \frac{1}{T}\int_0^t S_u\,du + \frac{1}{T}\int_t^T S_t e^{r(u-t)}\,du,
\end{aligned}$$

where we replaced μ by r in the formula $\widehat{E}[S_u|\mathcal{F}_t] = S_t e^{\mu(u-t)}$. Integrating we obtain

$$\begin{aligned}
\widehat{E}_t[A_T] &= \frac{1}{T}\int_0^t S_u\,du + \frac{1}{T}\int_t^T S_t e^{r(u-t)}\,du \\
&= \frac{t}{T}A_t + \frac{1}{T}S_t e^{-rt}\frac{e^{rT} - e^{rt}}{r} \\
&= \frac{t}{T}A_t + \frac{1}{rT}S_t\Big(e^{r(T-t)} - 1\Big).
\end{aligned} \tag{13.12.12}$$

Using (13.12.12), the risk-neutral valuation provides

$$\begin{aligned}
f_t &= e^{-r(T-t)}\widehat{E}_t[A_T - K] = e^{-r(T-t)}\widehat{E}_t[A_T] - e^{-r(T-t)}K \\
&= e^{-r(T-t)}\Big(\frac{t}{T}A_t - K\Big) + \frac{1}{rT}S_t\Big(1 - e^{-r(T-t)}\Big).
\end{aligned}$$

■

Exercise 13.12.2 *Show that $\lim_{t\to 0} f_t = f_0$, where f_t is given by* (13.12.11) *and f_0 by* (13.12.10).

Exercise 13.12.3 *Find the value at time t of a contract that pays at maturity date the difference between the asset price and its arithmetic average, $f_T = S_T - A_T$.*

Forward Contracts on the Geometric Average We shall consider in the following Asian forward contracts on the geometric average. This is a derivative that pays at maturity the difference $f_T = G_T - K$, where G_T is the continuous geometric average of the asset price between 0 and T and K is a fixed delivery price.

We shall work out first the value of the contract at $t = 0$. Substituting $\mu = r$ in the first relation provided by Theorem 12.4.4, the risk-neutral expectation of G_T as of time $t = 0$ is

$$\widehat{E}_0[G_T] = S_0 e^{\frac{1}{2}(r-\frac{\sigma^2}{6})T}.$$

Then

$$\begin{aligned} f_0 &= e^{-rT}\widehat{E}_0[G_T - K] = e^{-rT}\widehat{E}_0[G_T] - e^{-rT}K \\ &= e^{-rT}S_0 e^{\frac{1}{2}(r-\frac{\sigma^2}{6})T} - e^{-rT}K \\ &= S_0 e^{-\frac{1}{2}(r+\frac{\sigma^2}{6})T} - e^{-rT}K. \end{aligned}$$

Thus, the price of a forward contract on geometric average at $t = 0$ is given by

$$\boxed{f_0 = S_0 e^{-\frac{1}{2}(r+\frac{\sigma^2}{6})T} - e^{-rT}K.} \tag{13.12.13}$$

As in the case of forward contracts on arithmetic average, the value at time $0 < t < T$ cannot be obtain from (13.12.13) by replacing blindly T and S_0 by $T - t$ and S_t, respectively. The correct relation is given by the following result:

Proposition 13.12.4 *The value at time t of a contract which pays at maturity $G_T - K$ is*

$$\boxed{f_t = G_t^{\frac{t}{T}} S_t^{1-\frac{t}{T}} e^{-r(T-t)+(r-\frac{\sigma^2}{2})\frac{(T-t)^2}{2T}+\frac{\sigma^2}{T^2}\frac{(T-t)^3}{6}} - e^{-r(T-t)}K.} \tag{13.12.14}$$

Proof: Since for $t < u$

$$\ln S_u = \ln S_t + (\mu - \frac{\sigma^2}{2})(u-t) + \sigma(W_u - W_t),$$

we have

$$\begin{aligned} \int_0^T \ln S_u\,du &= \int_0^t \ln S_u\,du + \int_t^T \ln S_u\,du \\ &= \int_0^t \ln S_u\,du + \int_t^T \Big(\ln S_t + (\mu - \frac{\sigma^2}{2})(u-t) + \sigma(W_u - W_t)\Big)\,du \\ &= \int_0^t \ln S_u\,du + (T-t)\ln S_t + (\mu - \frac{\sigma^2}{2})\Big(\frac{T^2 - t^2}{2} - t(T-t)\Big) \\ &\quad +\sigma \int_t^T (W_u - W_t)\,du. \end{aligned}$$

The geometric average becomes

$$\begin{aligned} G_T &= e^{\frac{1}{T}\int_0^T \ln S_u\,du} \\ &= e^{\frac{1}{T}\int_0^t \ln S_u\,du}\, S_t^{1-\frac{t}{T}} e^{\frac{1}{T}(\mu-\frac{\sigma^2}{2})(\frac{T+t}{2}-t)(T-t)} e^{\frac{\sigma}{T}\int_t^T (W_u-W_t)\,du} \\ &= G_t^{\frac{t}{T}}\, S_t^{1-\frac{t}{T}}\, e^{(\mu-\frac{\sigma^2}{2})\frac{(T-t)^2}{2T}} e^{\frac{\sigma}{T}\int_t^T (W_u-W_t)\,du}, \end{aligned} \tag{13.12.15}$$

where we used that

$$e^{\frac{1}{T}\int_0^T \ln S_u\,du} = e^{\frac{t}{T}\ln G_t} = G_t^{\frac{t}{T}}.$$

Relation (13.12.15) provides G_T in terms of G_t and S_t. Taking the predictable part out and replacing μ by r we have

$$\widehat{E}_t[G_T] = G_t^{\frac{t}{T}}\, S_t^{1-\frac{t}{T}}\, e^{(r-\frac{\sigma^2}{2})\frac{(T-t)^2}{2T}} E\Big[e^{\frac{\sigma}{T}\int_t^T (W_u-W_t)\,du}|\mathcal{F}_t\Big]. \tag{13.12.16}$$

Since the jump $W_u - W_t$ is independent of the information set $\mathcal{F}_t$, the condition can be dropped

$$E\Big[e^{\frac{\sigma}{T}\int_t^T (W_u-W_t)\,du}|\mathcal{F}_t\Big] = E\Big[e^{\frac{\sigma}{T}\int_t^T (W_u-W_t)\,du}\Big].$$

Integrating by parts yields

$$\begin{aligned} \int_t^T (W_u - W_t)\,du &= \int_t^T W_u\,du - (T-t)W_t \\ &= TW_T - tW_t - \int_t^T u\,dW_u - TW_t + tW_t \\ &= T(W_T - W_t) - \int_t^T u\,dW_u = \int_t^T T\,dW_u - \int_t^T u\,dW_u \\ &= \int_t^T (T-u)\,dW_u, \end{aligned}$$

which is a Wiener integral. This is normal distributed with mean 0 and variance

$$\int_t^T (T-u)^2\,du = \frac{(T-t)^3}{3}.$$

Then $\displaystyle\int_t^T (W_u - W_t)\,du \sim N\Big(0, \frac{(T-t)^3}{3}\Big)$ and hence

$$E\Big[e^{\frac{\sigma}{T}\int_t^T (W_u-W_t)\,du}\Big] = E\Big[e^{\frac{\sigma}{T}\int_t^T (T-u)\,dW_u}\Big] = e^{\frac{1}{2}\frac{\sigma^2}{T^2}\frac{(T-t)^3}{3}} = e^{\frac{\sigma^2}{T^2}\frac{(T-t)^3}{6}}.$$

Substituting into (13.12.16) yields

$$\widehat{E}_t[G_T] = G_t^{\frac{t}{T}}\, S_t^{1-\frac{t}{T}}\, e^{(r-\frac{\sigma^2}{2})\frac{(T-t)^2}{2T}} e^{\frac{\sigma^2}{T^2}\frac{(T-t)^3}{6}}. \tag{13.12.17}$$

Hence the value of the contract at time t is given by

$$f_t = e^{-r(T-t)}\widehat{E}_t[G_T - K] = G_t^{\frac{t}{T}} S_t^{1-\frac{t}{T}} e^{-r(T-t)+(r-\frac{\sigma^2}{2})\frac{(T-t)^2}{2T}+\frac{\sigma^2}{T^2}\frac{(T-t)^3}{6}} - e^{-r(T-t)}K.$$

$\blacksquare$

Exercise 13.12.5 *Show that* $\lim_{t\to 0} f_t = f_0$, *where* f_t *is given by* (13.12.14) *and* f_0 *by* (13.12.13).

Exercise 13.12.6 *Which is cheaper: an Asian forward contract on* A_t *or an Asian forward contract on* G_t*?*

Exercise 13.12.7 *Using Corollary* 12.4.5 *find a formula of* G_T *in terms of* G_t, *and then compute the risk-neutral world expectation* $\widehat{E}_t[G_T]$.

13.13 Asian Options

There are several types of Asian options depending on how the payoff is related to the average stock price:

- Average Price options:
  - Call: $f_T = \max(S_{ave} - K, 0)$
  - Put: $f_T = \max(K - S_{ave}, 0)$.
- Average Strike options:
  - Call: $f_T = \max(S_T - S_{ave}, 0)$
  - Put: $f_T = \max(S_{ave} - S_T, 0)$.

The average asset price S_{ave} can be either the arithmetic or the geometric average of the asset price between 0 and T.

Geometric average price options When the asset is the geometric average, G_T, we shall obtain closed form formulas for average price options. Since G_T is log-normally distributed, the pricing procedure is similar with the one used for the usual options on stock. We shall do this by using the superposition principle and the following two results. The first one is a cash-or-nothing type contract where the underlying asset is the geometric mean of the stock between 0 and T.

Lemma 13.13.1 *The value at time* $t = 0$ *of a derivative, which pays at maturity* \$1 *if the geometric average* $G_T \geq K$ *and* 0 *otherwise, is given by*

$$h_0 = e^{-rT} N(\widetilde{d}_2),$$

where

$$\widetilde{d_2} = \frac{\ln S_0 - \ln K + (\mu - \frac{\sigma^2}{2})\frac{T}{2}}{\sigma\sqrt{T/3}}.$$

Proof: The payoff can be written as

$$h_T = \begin{cases} 1, & \text{if } G_T \geq K \\ 0, & \text{if } G_T < K \end{cases} = \begin{cases} 1, & \text{if } X_T \geq \ln K \\ 0, & \text{if } X_T < \ln K, \end{cases}$$

where $X_T = \ln G_T$ has the normal distribution

$$X_T \sim N\Big[\ln S_0 + (\mu - \frac{\sigma^2}{2})\frac{T}{2}, \frac{\sigma^2 T}{3}\Big],$$

see formula (12.4.15). Let $p(x)$ be the probability density of the random variable X_T

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma\sqrt{T/3}} e^{-[x - \ln S_0 - (\mu - \frac{\sigma^2}{2})\frac{T}{2}]^2/(\frac{2\sigma^2 T}{3})}. \tag{13.13.18}$$

The risk neutral expectation of the payoff at time $t = 0$ is

$$\begin{aligned} \widehat{E}_0[h_T] &= \int h_T(x)p(x)\,dx = \int_{\ln K}^{\infty} p(x)\,dx \\ &= \frac{1}{\sqrt{2\pi}\sigma\sqrt{T/3}} \int_{\ln K}^{\infty} e^{-[x - \ln S_0 - (r - \frac{\sigma^2}{2})\frac{T}{2}]^2/(\frac{2\sigma^2 T}{3})}\,dx, \end{aligned}$$

where μ was replaced by r. Substituting

$$y = \frac{x - \ln S_0 - (r - \frac{\sigma^2}{2})\frac{T}{2}}{\sigma\sqrt{T/3}}, \tag{13.13.19}$$

yields

$$\begin{aligned} \widehat{E}_0[h_T] &= \frac{1}{\sqrt{2\pi}} \int_{-\widetilde{d_2}}^{\infty} e^{-y^2/2}\,dy = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\widetilde{d_2}} e^{-y^2/2}\,dy \\ &= N(\widetilde{d_2}), \end{aligned}$$

where

$$\widetilde{d_2} = \frac{\ln S_0 - \ln K + (r - \frac{\sigma^2}{2})\frac{T}{2}}{\sigma\sqrt{T/3}}.$$

Discounting to the free interest rate yields the price at time $t = 0$

$$h_0 = e^{-rT}\widehat{E}_0[h_T] = e^{-rT}N(\widetilde{d_2}).$$

■

The following result deals with the price of an average-or-nothing derivative on the geometric average.

Lemma 13.13.2 *The value at time $t = 0$ of a derivative, which pays at maturity G_T if $G_T \geq K$ and 0 otherwise, is given by the formula*

$$g_0 = e^{-\frac{1}{2}(r+\frac{\sigma^2}{6})T} S_t N(\widetilde{d}_1),$$

where

$$\widetilde{d}_1 = \frac{\ln S_0 - \ln K + (r + \frac{\sigma^2}{6})\frac{T}{2}}{\sigma\sqrt{T/3}}.$$

Proof: Since the payoff can be written as

$$g_T = \begin{cases} G_T, & \text{if } G_T \geq K \\ 0, & \text{if } G_T < K \end{cases} = \begin{cases} e^{X_T}, & \text{if } X_T \geq \ln K \\ 0, & \text{if } X_T < \ln K, \end{cases}$$

with $X_T = \ln G_T$, the risk neutral expectation at the time $t = 0$ of the payoff is

$$\widehat{E}_0[g_T] = \int_{-\infty}^{\infty} g_T(x)p(x)\,dx = \int_{\ln K}^{\infty} e^x p(x)\,dx,$$

where $p(x)$ is given by (13.13.18), with μ replaced by r. Using the substitution (13.13.19) and completing the square yields

$$\begin{aligned}
\widehat{E}_0[g_T] &= \frac{1}{\sqrt{2\pi}\sigma\sqrt{T/3}} \int_{\ln K}^{\infty} e^x e^{-[x-\ln S_0 - (r-\frac{\sigma^2}{2})\frac{T}{2}]^2/(\frac{2\sigma^2 T}{3})}\,dx \\
&= \frac{1}{\sqrt{2\pi}} S_0 e^{\frac{1}{2}(r-\frac{\sigma^2}{6})T} \int_{-\widetilde{d}_2}^{\infty} e^{-\frac{1}{2}[y-\sigma\sqrt{T/3}]^2}\,dy
\end{aligned}$$

If we let

$$\begin{aligned}
\widetilde{d}_1 &= \widetilde{d}_2 + \sigma\sqrt{T/3} \\
&= \frac{\ln S_0 - \ln K + (r - \frac{\sigma^2}{2})\frac{T}{2}}{\sigma\sqrt{T/3}} + \sigma\sqrt{T/3} \\
&= \frac{\ln S_0 - \ln K + (r + \frac{\sigma^2}{6})\frac{T}{2}}{\sigma\sqrt{T/3}},
\end{aligned}$$

the previous integral becomes, after substituting $z = y - \sigma\sqrt{T/3}$,

$$\frac{1}{\sqrt{2\pi}} S_0 e^{\frac{1}{2}(r-\frac{\sigma^2}{6})T} \int_{-\widetilde{d}_1}^{\infty} e^{-\frac{1}{2}z^2}\,dz = S_0 e^{\frac{1}{2}(r-\frac{\sigma^2}{6})T} N(\widetilde{d}_1).$$

Then the risk neutral expectation of the payoff is

$$\widehat{E}_0[g_T] = S_0 e^{\frac{1}{2}(r-\frac{\sigma^2}{6})T} N(\widetilde{d}_1).$$

The value of the derivative at time t is obtained by discounting at the interest rate r

$$\begin{aligned} g_0 &= e^{-rT}\widehat{E}_0[g_T] = e^{-rT}S_0 e^{\frac{1}{2}(r-\frac{\sigma^2}{6})T}N(\widetilde{d}_1) \\ &= e^{-\frac{1}{2}(r+\frac{\sigma^2}{6})T}S_0 N(\widetilde{d}_1). \end{aligned}$$

■

Proposition 13.13.3 *The value at time t of a geometric average price call option is*

$$\boxed{f_0 = e^{-\frac{1}{2}(r+\frac{\sigma^2}{6})T}S_0 N(\widetilde{d}_1) - Ke^{-rT}N(\widetilde{d}_2).}$$

Proof: Since the payoff $f_T = \max(G_T - K, 0)$ can be decomposed as

$$f_T = g_T - Kh_T,$$

with

$$g_T = \begin{cases} G_T, & \text{if } G_T \geq K \\ 0, & \text{if } G_T < K, \end{cases} \qquad h_T = \begin{cases} 1, & \text{if } G_T \geq K \\ 0, & \text{if } G_T < K, \end{cases}$$

applying the superposition principle and Lemmas 13.13.1 and 13.13.2 yields

$$\begin{aligned} f_0 &= g_0 - Kh_0 \\ &= e^{-\frac{1}{2}(r+\frac{\sigma^2}{6})T}S_0 N(\widetilde{d}_1) - Ke^{-rT}N(\widetilde{d}_2). \end{aligned}$$

■

Exercise 13.13.4 *Find the value at time $t = 0$ of a price put option on a geometric average, i.e. a derivative with the payoff $f_T = \max(K - G_T, 0)$.*

Arithmetic average price options There is no simple closed-form solution for a call or for a put on the arithmetic average A_t. However, there is an approximate solution based on computing exactly the first two moments of the distribution of A_t, and applying the risk-neutral valuation assuming that the distribution is log-normally with the same two moments. This idea was developed by Turnbull and Wakeman [15], and works pretty well for volatilities up to about 20%.

The following result provides the mean and variance of a normal distribution in terms of the first two moments of the associated log-normally distribution.

Proposition 13.13.5 *Let Y be a log-normally distributed random variable, having the first two moments given by*

$$m_1 = E[Y], \qquad m_2 = E[Y^2].$$

Then $\ln Y$ has the normal distribution $\ln Y \sim N(\mu, \sigma^2)$, with the mean and variance given respectively by

$$\mu = \ln \frac{m_1^2}{\sqrt{m_2}}, \qquad \sigma^2 = \ln \frac{m_2}{m_1^2}\cdot \tag{13.13.20}$$

Proof: Using Exercise 1.6.2 we have

$$m_1 = E[Y] = e^{\mu+\frac{\sigma^2}{2}}, \qquad m_2 = E[Y^2] = e^{2\mu+2\sigma^2}.$$

Taking a logarithm yields

$$\mu + \frac{\sigma^2}{2} = \ln m_1, \qquad 2\mu + 2\sigma^2 = \ln m_2.$$

Solving for μ and σ yields (13.13.20). ■

Assume the arithmetic average $A_t = \frac{I_t}{t}$ has a log-normally distribution. Then $\ln A_t = \ln I_t - \ln t$ is normal, so $\ln I_t$ is normal, and hence I_t is log-normally distributed. Since $I_T = \int_0^T S_u\, du$, using (12.4.12) yields

$$\begin{aligned} m_1 = E[I_T] &= S_0 \frac{e^{\mu T} - 1}{\mu}; \\ m_2 = E[I_T^2] &= \frac{2S_0^2}{\mu + \sigma^2}\Big[\frac{e^{(2\mu+\sigma^2)T} - 1}{2\mu + \sigma^2} - \frac{e^{\mu T} - 1}{\mu}\Big]. \end{aligned}$$

Using Proposition 13.13.5 it follows that $\ln A_t$ is normally distributed, with

$$\boxed{\ln A_T \sim N\Big(\ln \frac{m_1^2}{\sqrt{m_2}} - \ln t, \ln \frac{m_2}{m_1^2}\Big).} \qquad (13.13.21)$$

Relation (13.13.21) represents the normal approximation of $\ln A_t$. We shall price the arithmetic average price call under this condition.
In the next two exercises we shall assume that the distribution of A_T is given by the log-normal distribution (13.13.21).

Exercise 13.13.6 *Using a method similar to the one used in Lemma* 13.13.1*, show that an approximate value at time* 0 *of a derivative, which pays at maturity* \$1 *if the arithmetic average* $A_T \geq K$ *and* 0 *otherwise, is given by*

$$h_0 = e^{-rT} N(\check{d}_2),$$

with

$$\check{d}_2 = \frac{\ln(m_1^2/\sqrt{m_2}) - \ln K - \ln t}{\sqrt{\ln(m_2/m_1^2)}}, \qquad (13.13.22)$$

where in the expressions of m_1 *and* m_2 *we replaced* μ *by* r*.*

Exercise 13.13.7 *Using a method similar to the one used in Lemma* 13.13.2*, show that the approximate value at time* 0 *of a derivative, which pays at maturity* A_T *if* $A_T \geq K$ *and* 0 *otherwise, is given by the formula*

$$a_0 = S_0 \frac{1 - e^{-rT}}{rt} N(\check{d}_1),$$

where (CHECK THIS AGAIN!)

$$\check{d}_1 = \ln(m_2/m_1^2) + \frac{\ln \frac{m_1^2}{\sqrt{m_2}} + \ln t - \ln K}{\sqrt{\ln(m_2/m_1^2)}}, \tag{13.13.23}$$

where in the expressions of m_1 and m_2 we replaced μ by r.

Proposition 13.13.8 *The approximate value at $t = 0$ of an arithmetic average price call is given by*

$$f_0 = \frac{S_0(1 - e^{-rT})}{r} N(\check{d}_1) - K e^{-rT} N(\check{d}_2),$$

with $\check{d}_1$ and $\check{d}_2$ given by formulas (13.13.22) *and* (13.13.23).

Exercise 13.13.9 (*a*) *Prove Proposition 13.13.8.*
(*b*) *How does the formula change if the value is taken at time t instead of time 0?*

13.14 Forward Contracts with Rare Events

We shall evaluate the price of a forward contract on a stock which follows a stochastic process with rare events, where the number of events $n = N_T$ until time T is assumed Poisson distributed. As usual, T denotes the maturity of the forward contract.

Let the jump ratios be $Y_j = S_{t_j}/S_{t_j-}$, where the events occur at times $0 < t_1 < t_2 < \cdots < t_n$, where $n = N_T$. The stock price at maturity is given by Merton's formula xxx

$$S_T = S_0 e^{(\mu - \lambda\rho - \frac{1}{2}\sigma^2)T + \sigma W_T} \prod_{j=1}^{N_T} Y_j,$$

where $\rho = E[Y_j]$ is the expected jump size, and $Y_1, \cdots, Y_n$ are considered independent among themselves and also with respect to W_t. Conditioning over $N_T = n$ yields

$$\begin{aligned} E\big[\prod_{j=1}^{N_T} Y_j\big] &= \sum_{n \geq 0} E\big[\prod_{j=1}^{N_T} |N_T = n\big] P(N_T = n) \\ &= \sum_{n \geq 0} \prod_{j=1}^{n} E[Y_j] P(N_T = n) \\ &= \sum_{n \geq 0} \rho^n \frac{(\lambda T)^n}{n!} e^{-\lambda T} = e^{\lambda \rho T} e^{-\lambda T}. \end{aligned}$$

Since W_T is independent of N_T and Y_j we have

$$\begin{aligned} E[S_T] &= S_0 e^{(\mu - \lambda\rho - \frac{\sigma^2}{2})T} E[e^{\sigma W_T}] E[\prod_{j=0}^{N_T} Y_j] \\ &= S_0 e^{(\mu - \lambda\rho)T} e^{\lambda\rho T} e^{-\lambda T} \\ &= S_0 e^{(\mu - \lambda)T}. \end{aligned}$$

Since the payoff at maturity of the forward contract is $f_T = S_T - K$, with K delivery price, the value of the contract at time $t = 0$ is obtained by the method of risk neutral valuation

$$\begin{aligned} f_0 &= e^{rT}\Big(\widehat{E}_0[S_T - K]\Big) = e^{-rT}\Big(S_0 e^{(r-\lambda)T} - K\Big) \\ &= S_0 e^{-\lambda T} - K e^{-rT}. \end{aligned}$$

Replacing T with $T - t$ and S_0 with S_t yields the value of the contract time t

$$\boxed{f_t = S_t e^{-\lambda(T-t)} - K e^{-r(T-t)}, \qquad 0 \le t \le T.} \tag{13.14.24}$$

It is worth noting that if the rate of jumps ocurrence is $\lambda = 0$, we obtain the familiar result

$$f_t = S_t - e^{-r(T-t)}K.$$

13.15 All-or-Nothing Lookback Options (Needs work!)

Consider a contract that pays the cash amount K at time T if the stock price S_t did ever reach or exceed level z until time T, and the amount 0 otherwise. The payoff is

$$V_T = \begin{cases} K, & \text{if } \overline{S}_T \ge z \\ 0, & \text{otherwise,} \end{cases}$$

where $\overline{S}_T = \max\limits_{t \le T} S_t$ is the running maximum of the stock.

In order to compute the exact value of the option we need to find the probability density of the running maximum

$$X_t = \overline{S}_t = \max_{s \le t} S_s = S_0 e^{\max\limits_{s\le t}(\mu s + \sigma W_s)}$$

where $\mu = r - \frac{\sigma^2}{2}$. Let $Y_t = \max\limits_{s \le t}(\mu s + \sigma W_s)$.

Let T_x be the first time the process $\mu t + \sigma W_t$ reaches level x, with $x > 0$. The probability density of T_x is given by Proposition 3.6.5

$$p(\tau) = \frac{x}{\sigma\sqrt{2\pi\tau^3}} e^{-\frac{(x-\mu\tau)^2}{2\tau\sigma^2}}.$$

The probability function of Y_t is given by

$$\begin{aligned} P(Y_t \le x) &= 1 - P(Y_t > x) = 1 - P\big(\max_{0\le s\le t}(\mu s + \sigma W_s) > x\big) \\ &= 1 - P(T_x \le t) = 1 - \int_0^t \frac{x}{\sigma\sqrt{2\pi\tau^3}} e^{-\frac{(x-\mu\tau)^2}{2\tau\sigma^2}}\,d\tau \\ &= \int_t^\infty \frac{x}{\sigma\sqrt{2\pi\tau^3}} e^{-\frac{(x-\mu\tau)^2}{2\tau\sigma^2}}\,d\tau. \end{aligned}$$

Then the probability function of X_t becomes

$$\begin{aligned} P(X_t \le u) &= P(e^{Y_t} \le u/S_0) = P(Y_t \le \ln(u/S_0)) \\ &= \int_t^\infty \frac{\ln(u/S_0)}{\sigma\sqrt{2\pi\tau^3}} e^{-\frac{(\ln(u/S_0)-\mu\tau)^2}{2\tau\sigma^2}}\, d\tau. \end{aligned}$$

Let $S_0 < z$. What is the probability that the stock S_t hits the barrier z before time T? This can be formalized as the probability $P(\max\limits_{t\le T} S_t > z)$, which can be computed using the previous probability function:

$$\begin{aligned} P(\overline{S}_T > z) &= 1 - P(\max_{t\le T} S_t \le z) = 1 - P(X_T \le z) \\ &= 1 - \int_T^\infty \frac{\ln(z/S_0)}{\sigma\sqrt{2\pi\tau^3}} e^{-\frac{(\ln(z/S_0)-\mu\tau)^2}{2\tau\sigma^2}}\, d\tau \\ &= \int_0^T \frac{\ln(z/S_0)}{\sigma\sqrt{2\pi\tau^3}} e^{-\frac{(\ln(z/S_0)-\mu\tau)^2}{2\tau\sigma^2}}\, d\tau, \end{aligned}$$

where $\mu = r - \sigma^2/2$. The exact value of the all-or-nothing lookback option at time $t = 0$ is

$$\begin{aligned} V_0 &= e^{-rT} E[V_T] \\ &= E\big[V_T | \max_{t\le T} S_t \ge z\big] P(\max_{t\le T} S_t \ge z) e^{-rT} \\ &\quad + \underbrace{E\big[V_T | \max_{t\le T} S_t < z\big]}_{=0} P(\max_{t\le T} S_t < z) e^{-rT} \\ &= e^{-rT} K P(\max_{t\le T} S_t \ge z) \\ &= e^{-rT} K \int_0^T \frac{\ln(z/S_0)}{\sigma\sqrt{2\pi\tau^3}} e^{-\frac{(\ln(z/S_0)-\mu\tau)^2}{2\tau\sigma^2}}\, d\tau. \end{aligned} \tag{13.15.25}$$

Since the previous integral does not have an elementary expression, we shall work out some lower and upper bounds.

Proposition 13.15.1 (Typos in the proof) *We have*

$$e^{-rT} K\Big(1 - \sqrt{\frac{2}{\pi\sigma^3 T}} \ln\frac{z}{S_0}\Big) < V_0 < \frac{K S_0}{z}.$$

Proof: First, we shall find an upper bound for the option price using Doob's inequality. The stock price at time t is given by

$$S_t = S_0 e^{(r-\frac{\sigma^2}{2})t + \sigma W_t}.$$

Under normal market conditions we have $r > \frac{\sigma^2}{2}$, and by Corollary 2.12.4, part (b) it follows that S_t is a submartingale (with respect to the filtration induced by W_t). Applying Doob's submartingale inequality, Proposition 2.12.5, we have

$$P(\max_{t \le T} S_t \ge z) \le \frac{1}{z} E[S_T] = \frac{S_0}{z} e^{rT}.$$

Then the expected value of V_T can be estimated as

$$E[V_T] = K \cdot P(\max_{t \le T} S_t \ge z) + 0 \cdot P(\max_{t \le T} S_t < z) \le \frac{KS_0}{z} e^{rT}.$$

Discounting, we obtain an upper bound for the value of the option at time $t = 0$

$$V_0 = e^{-rT} E[V_T] \le \frac{KS_0}{z}\cdot$$

In order to obtain a lower bound, we need to write the integral away from the singularity at $\tau = 0$, and use that $e^{-x} > 1 - x$ for $x > 0$:

$$\begin{aligned} V_0 &= e^{-rT} K - e^{-rT} K \int_T^\infty \frac{\ln(z/S_0)}{\sqrt{2\pi\sigma^3\tau^3}} e^{-\frac{(\ln(z/S_0) - \mu\sigma\tau)^2}{2\tau\sigma^3}} \, d\tau \\ &< e^{-rT} K - e^{-rT} K \int_T^\infty \frac{\ln(z/S_0)}{\sqrt{2\pi\sigma^3\tau^3}} \, d\tau \\ &= e^{-rT} K \Big(1 - \sqrt{\frac{2}{\pi\sigma^3 T}} \ln \frac{z}{S_0} \Big). \end{aligned}$$

■

Exercise 13.15.2 *Find the value of an asset-or-nothing look-back option whose payoff is*

$$V_T = \begin{cases} \overline{S}_T, & \textit{if } S_T \ge K \\ 0, & \textit{otherwise.} \end{cases}$$

K denotes the strike price.

Exercise 13.15.3 *Find the value of a price call look-back option with the payoff given by*

$$V_T = \begin{cases} \overline{S}_T - K, & \textit{if } S_T \ge K \\ 0, & \textit{otherwise.} \end{cases}$$

Exercise 13.15.4 *Find the value of a price put look-back option whose payoff is given by*

$$V_T = \begin{cases} K - \overline{S}_T, & \textit{if } S_T < K \\ 0, & \textit{otherwise.} \end{cases}$$

Exercise 13.15.5 *Evaluate the following strike look-back options*

(a) $V_T = S_T - \underline{S}_T$ *(call)*

(b) $V_T = \overline{S}_T - S_T$ *(put)*

It is worth noting that $(S_T - \underline{S}_T)^+ = S_T - \underline{S}_T$; *this explains why the payoff is not given as a piece-wise function.*

Hint: see p. 739 of Mc Donald for closed form formula.

Exercise 13.15.6 *Find a put-call parity for look-back options.*

Exercise 13.15.7 *Starting from the formula (see K&S p. 265)*

$$P\Big(\sup_{t\le T}(W_t - \gamma t) \ge \beta\Big) = 1 - N\Big(\gamma\sqrt{T} + \frac{\beta}{\sqrt{T}}\Big) + e^{-2\beta\gamma} N\Big(\gamma\sqrt{T} - \frac{\beta}{\sqrt{T}}\Big),$$

with $\beta > 0$, $\gamma \in \mathbb{R}$ *and* $N(x) = \dfrac{1}{\sqrt{2\pi}} \displaystyle\int_{-\infty}^{\infty} e^{-z^2/2}\,dz$, *calculate the price at* $t = 0$ *of a contract that pays* K *at time* T *if the stock price* S_t *did ever reach or exceed level* z *before time* T.

13.16 Perpetual Look-back Options

Consider a contract that pays K at the time the stock price S_t reaches level b for the first time. What is the price of such contract at any time $t \ge 0$?

This type of contract is called *perpetual* look-back because the time horizon is infinite, i.e. there is no expiration date for the contract.

The contract pays K at time τ_b, where $\tau_b = \inf\{t > 0; S_t \ge b\}$. The value at time $t = 0$ is obtained by discounting at the risk-free interest rate r

$$V_0 = E\big[Ke^{-r\tau_b}\big] = KE\big[e^{-r\tau_b}\big]. \qquad (13.16.26)$$

This expectation will be worked out using formula (3.6.11). First, we notice that the time τ_b at which $S_t = b$ is the same as the time for which $S_0 e^{(r-\sigma^2/2)t + \sigma W_t} = b$, or equivalently,

$$\Big(r - \frac{\sigma^2}{2}\Big)t + \sigma W_t = \ln\frac{b}{S_0}.$$

Applying Proposition 3.6.4 with $\mu = r - \sigma^2/2$, $s = r$, and $x = \ln(b/S_0)$ yields

$$\begin{aligned} E[e^{-r\tau_b}] &= e^{\frac{1}{\sigma^2}\left(r - \frac{\sigma^2}{2} - \sqrt{2r\sigma + (r-\frac{\sigma^2}{2})^2}\right)\ln\frac{b}{S_0}} \\ &= \Big(\frac{b}{S_0}\Big)^{\left(r - \frac{\sigma^2}{2} - \sqrt{2r\sigma + (r-\frac{\sigma^2}{2})^2}\right)/\sigma^2}. \end{aligned}$$

Substituting in (13.16.26) yields the price of the perpetual look-back option at $t = 0$

$$\boxed{V_0 = K\Big(\frac{b}{S_0}\Big)^{\left(r-\frac{\sigma^2}{2}-\sqrt{2r\sigma+(r-\frac{\sigma^2}{2})^2}\right)/\sigma^2}.} \tag{13.16.27}$$

Exercise 13.16.1 *Find the price of a contract that pays K when the initial stock price doubles its value.*

13.17 Immediate Rebate Options

This contract pays at the time the barrier is hit, T_a, the amount K. The discounted value at $t = 0$ is

$$f_0 = \begin{cases} e^{-rT_a}, & \text{if } T_a \le T \\ 0, & \text{otherwise.} \end{cases}$$

The price of the contract at $t = 0$ is

$$\begin{aligned} V_0 &= E[f_0] = E[f_0|T_a < T]P(T_a < T) + \underbrace{E[f_0|T_a \ge T]}_{=0} P(T_a \ge T) \\ &= KE[e^{-rT_a}]P(T_a < T) \\ &= K\int_0^\infty e^{-r\tau}p_a(\tau)\,d\tau \int_0^T p_a(\tau)\,d\tau, \end{aligned}$$

where

$$p_a(\tau) = \frac{\ln\frac{a}{S_0}}{\sqrt{2\pi\sigma^3\tau^3}}e^{-(\ln\frac{a}{S_0}-\alpha\sigma\tau)^2/(2\tau\sigma^3)},$$

and $\alpha = r - \dfrac{\sigma^2}{2}$. Question: Why in McDonald (22.20) the formula is different. Are they equivalent?

13.18 Deferred Rebate Options

The payoff of this contract pays K as long as a certain barrier has been hit. It is called *deferred* because the payment K is made at the expiration time T. If T_a is the first time the stock reaches the barrier a, the payoff can be formalized as

$$V_T = \begin{cases} K, & \text{if } T_a < T \\ 0, & \text{if } T_a \ge T. \end{cases}$$

The value of the contract at time $t = 0$ is

$$\begin{aligned} V_0 &= e^{-rT}E[V_T] = e^{-rT}KP(T_a < T) \\ &= Ke^{-rT}\ln\frac{a}{S_0}\int_0^T \frac{1}{\sigma\sqrt{2\pi\tau^3}}e^{(\ln\frac{a}{S_0}-\alpha\tau)^2/(2\tau\sigma^2)}\,d\tau, \end{aligned}$$

with $\alpha = r - \dfrac{\sigma^2}{2}$.

Chapter 14

Martingale Measures

14.1 Martingale Measures

An $\mathcal{F}_t$-predictable stochastic process X_t on the probability space $(\Omega, \mathcal{F}, P)$ is not always a martingale. However, it might become a martingale with respect to another probability measure Q on $\mathcal{F}$. This is called a *martingale measure*. The main result of this section is finding a martingale measure with respect to which the discounted stock price is a martingale. This measure plays an important role in the mathematical explanation of the risk-neutral valuation.

14.1.1 Is the stock price S_t a martingale?

Since the stock price S_t is an $\mathcal{F}_t$-predictable and non-explosive process, the only condition which needs to be satisfied to be a $\mathcal{F}_t$-martingale is

$$E[S_t|\mathcal{F}_u] = S_u, \qquad \forall u < t. \tag{14.1.1}$$

Heuristically speaking, this means that given all information in the market at time u, $\mathcal{F}_u$, the expected price of any future stock price is the price of the stock at time u, i.e S_u. This does not make sense, since in this case the investor would prefer investing the money in a bank at a risk-free interest rate, rather than buying a stock with zero return. Then (14.1.1) does not hold. The next result shows how to fix this problem.

Proposition 14.1.1 *Let μ be the rate of return of the stock S_t. Then*

$$E[e^{-\mu t}S_t|\mathcal{F}_u] = e^{-\mu u}S_u, \qquad \forall u < t, \tag{14.1.2}$$

i.e. $e^{-\mu t}S_t$ is an $\mathcal{F}_t$-martingale.

Proof: The process $e^{-\mu t}S_t$ is non-explosive since

$$E[|e^{-\mu t}S_t|] = e^{-\mu t}E[S_t] = e^{-\mu t}S_0 e^{\mu t} = S_0 < \infty.$$

Since S_t is $\mathcal{F}_t$-predictable, so is $e^{-\mu t}S_t$. Using formula (12.1.2) and taking out the predictable part yields

$$\begin{aligned}
E[S_t|\mathcal{F}_u] &= E[S_0 e^{(\mu-\frac{1}{2}\sigma^2)t+\sigma W_t}|\mathcal{F}_u] \\
&= E[S_0 e^{(\mu-\frac{1}{2}\sigma^2)u+\sigma W_u} e^{(\mu-\frac{1}{2}\sigma^2)(t-u)+\sigma(W_t-W_u)}|\mathcal{F}_u] \\
&= E[S_u e^{(\mu-\frac{1}{2}\sigma^2)(t-u)+\sigma(W_t-W_u)}|\mathcal{F}_u] \\
&= S_u e^{(\mu-\frac{1}{2}\sigma^2)(t-u)} E[e^{\sigma(W_t-W_u)}|\mathcal{F}_u]. \qquad (14.1.3)
\end{aligned}$$

Since the increment $W_t - W_u$ is independent of all values W_s, $s \le u$, then it will also be independent of $\mathcal{F}_u$. By Proposition 1.11.4, part 6, the conditional expectation becomes the usual expectation

$$E[e^{\sigma(W_t-W_u)}|\mathcal{F}_u] = E[e^{\sigma(W_t-W_u)}].$$

Since $\sigma(W_t - W_u) \sim N\big(0, \sigma^2(t-u)\big)$, from Exercise 1.6.2 (b) we get

$$E[e^{\sigma(W_t-W_u)}] = e^{\frac{1}{2}\sigma^2(t-u)}.$$

Substituting back into (14.1.3) yields

$$E[S_t|\mathcal{F}_u] = S_u e^{(\mu-\frac{1}{2}\sigma^2)(t-u)} e^{\frac{1}{2}\sigma^2(t-u)} = S_u e^{\mu(t-u)},$$

which is equivalent to

$$E[e^{-\mu t}S_t|\mathcal{F}_u] = e^{-\mu u}S_u.$$

■

The conditional expectation $E[S_t|\mathcal{F}_u]$ can be expressed in terms of the conditional density function as

$$E[S_t|\mathcal{F}_u] = \int S_t\, p(S_t|\mathcal{F}_u)\, dS_t, \qquad (14.1.4)$$

where S_t is taken as an integration variable.

Exercise 14.1.2 (*a*) *Find the formula for conditional density function, $p(S_t|\mathcal{F}_u)$, defined by* (14.1.4).
(*b*) *Verify the formula*

$$E[S_t|\mathcal{F}_0] = E[S_t]$$

in two different ways, either by using part (*a*)*, or by using the independence of S_t with respect to $\mathcal{F}_0$.*

The martingale relation (14.1.2) can be written equivalently as

$$\int e^{-\mu t} S_t\, p(S_t|\mathcal{F}_u)\, dS_t = e^{-\mu u}S_u, \qquad u < t.$$

This way, $dP(x) = p(x|\mathcal{F}_u)\, dx$ becomes a *martingale measure* for $e^{-\mu t}S_t$.

14.1.2 Risk-neutral World and Martingale Measure

Since the rate of return μ might not be known from the beginning, and it depends on each particular stock, a meaningful question would be:

Under what martingale measure does the discounted stock price, $M_t = e^{-rt}S_t$, become a martingale?

The constant r denotes, as usual, the risk-free interest rate. Assume such a martingale measure exists. Then we must have

$$\widehat{E}_u[e^{-rt}S_t] = \widehat{E}[e^{-rt}S_t|\mathcal{F}_u] = e^{-ru}S_u,$$

where $\widehat{E}$ denotes the expectation with respect to the requested martingale measure. The previous relation can also be written as

$$e^{-r(t-u)}\widehat{E}[S_t|\mathcal{F}_u] = S_u, \qquad u < t.$$

This states that the discounted expectation at the risk-free interest rate for the time interval $t-u$ is the price of the stock, S_u. Since this does not involve any of the riskiness of the stock, we might think of it as an expectation in the risk-neutral world. The aforementioned formula can be written in the compound mode as

$$\widehat{E}[S_t|\mathcal{F}_u] = S_u e^{r(t-u)}, \qquad u < t. \tag{14.1.5}$$

This formula can be obtained from the conditional expectation $E[S_t|\mathcal{F}_u] = S_u e^{\mu(t-u)}$ by substituting $\mu = r$ and replacing E by $\widehat{E}$, which corresponds to the definition of the expectation in a risk-neutral world. Therefore, the evaluation of derivatives in section 13 is done by using the aforementioned martingale measure under which $e^{-rt}S_t$ is a martingale. In the next section we shall determine this measure explicitly.

Exercise 14.1.3 *Consider the following two games that consit in flipping a fair coin and taking decisions:*

A. If the coin lands Heads, you win \$2; otherwise you loose \$1.

B. If the coin lands Heads, you win \$20,000; otherwise you loose \$10,000.

(a) Which game involves more risk? Explain your answer.

(b) Which game would you choose to play, and why?

(c) Are you risk-neutral in your decision?

The risk neutral measure is the measure with respect to which investors are not risk-averse. In the case of the previous exercise, it is the measure with respect to which all players are indiferent whether they choose option A or B.

14.1.3 Finding the Risk-Neutral Measure

The solution of the stochastic differential equation of the stock price

$$dS_t = \mu S_t dt + \sigma S_t dW_t$$

can be written as

$$S_t = S_0 e^{\mu t} e^{\sigma W_t - \frac{1}{2}\sigma^2 t}.$$

Then $e^{-\mu t} S_t = S_0 e^{\sigma W_t - \frac{1}{2}\sigma^2 t}$ is an exponential process. By Example 9.1.3, particular case 1, this process is an $\mathcal{F}_t$-martingale, where $\mathcal{F}_t = \sigma\{W_u; u \leq t\}$ is the information available in the market until time t. Hence $e^{-\mu t} S_t$ is a martingale, which is a result proved also by Proposition 14.1.1. The probability space where this martingale exists is $(\Omega, \mathcal{F}, P)$.

In the following we shall change the rate of return μ into the risk-free rate r and change the probability measure such that the discounted stock price becomes a martingale. The discounted stock price can be expressed in terms of the Brownian motion with drift (a hat was added in order to distinguish it from the former Brownian motion W_t)

$$\widehat{W}_t = \frac{\mu - r}{\sigma} t + W_t \tag{14.1.6}$$

as in the following

$$e^{-rt} S_t = e^{-rt} S_0 e^{\mu t} e^{\sigma W_t - \frac{1}{2}\sigma^2 t} = e^{\sigma \widehat{W}_t - \frac{1}{2}\sigma^2 t}.$$

If we let $\lambda = \dfrac{\mu - r}{\sigma}$ in Corollary 9.2.4 of Girsanov's theorem, it follows that $\widehat{W}_t$ is a Brownian motion on the probability space $(\Omega, \mathcal{F}, Q)$, where

$$dQ = e^{-\frac{1}{2}(\frac{\mu - r}{\sigma})^2 T - \lambda W_T}\, dP.$$

As an exponential process, $e^{\sigma \widehat{W}_t - \frac{1}{2}\sigma^2 t}$ becomes a martingale on this space. Consequently $e^{-rt} S_t$ is a martingale process w.r.t. the probability measure Q. This means

$$E^Q[e^{-rt} S_t | \mathcal{F}_u] = e^{-ru} S_u, \qquad u < t.$$

where $E^Q[\,\cdot\,|\mathcal{F}_u]$ denotes the conditional expectation in the measure Q, and it is given by

$$E^Q[X_t | \mathcal{F}_u] = E^P[X_t e^{-\frac{1}{2}(\frac{\mu - r}{\sigma})^2 T - \lambda W_T} | \mathcal{F}_u].$$

The measure Q is called the *equivalent martingale measure*, or the *risk-neutral measure*. The expectation taken with respect to this measure is called the *expectation in the risk-neutral world*. Customarily we shall use the notations

$$\begin{aligned} \widehat{E}[e^{-rt} S_t] &= E^Q[e^{-rt} S_t] \\ \widehat{E}_u[e^{-rt} S_t] &= E^Q[e^{-rt} S_t | \mathcal{F}_u] \end{aligned}$$

It is worth noting that $\widehat{E}[e^{-rt}S_t] = \widehat{E}_0[e^{-rt}S_t]$, since $e^{-rt}S_t$ is independent of the initial information set $\mathcal{F}_0$.

The importance of the process $\widehat{W}_t$ is contained in the following useful result.

Proposition 14.1.4 *The probability measure that makes the discounted stock price, $e^{-rt}S_t$, a martingale changes the rate of return μ into the risk-free interest rate r, i.e*

$$dS_t = rS_t dt + \sigma S_t d\widehat{W}_t.$$

Proof: The proof is a straightforward verification using (14.1.6)

$$\begin{aligned} dS_t &= \mu S_t dt + \sigma S_t dW_t = rS_t dt + (\mu - r)S_t dt + \sigma S_t dW_t \\ &= rS_t dt + \sigma S_t \Big(\frac{\mu - r}{\sigma} dt + dW_t\Big) \\ &= rS_t dt + \sigma S_t d\widehat{W}_t \end{aligned}$$

We note that the solution of the previous stochastic equation is

$$S_t = S_0 e^{rt} e^{\sigma \widehat{W}_t - \frac{1}{2}\sigma^2 t}.$$

■

Exercise 14.1.5 *Assume $\mu \neq r$ and let $u < t$.*

(a) Find $E^P[e^{-rt}S_t|\mathcal{F}_u]$ and show that $e^{-rt}S_t$ is not a martingale w.r.t. the probability measure P.

(b) Find $E^Q[e^{-\mu t}S_t|\mathcal{F}_u]$ and show that $e^{-\mu t}S_t$ is not a martingale w.r.t. the probability measure Q.

14.2 Risk-neutral World Density Functions

The purpose of this section is to establish formulas for the densities of Brownian motions W_t and $\widehat{W}_t$ with respect to both probability measures P and Q, and discuss their relationship. This will clear some confusions that appear in practical applications when we need to choose the right probability density.

The densities of W_t and $\widehat{W}_t$ with respect to P and Q will be denoted respectively by p_P, p_Q and $\widehat{p}_P$, $\widehat{p}_Q$. Since W_t and $\widehat{W}_t$ are Brownian motions on the spaces $(\Omega, \mathcal{F}, P)$ and $(\Omega, \mathcal{F}, Q)$, respectively, they have the following normal probability densities

$$\begin{aligned} p_P(x) &= \frac{1}{\sqrt{2\pi t}} e^{-\frac{x^2}{2t}} = p(x); \\ \widehat{p}_Q(x) &= \frac{1}{\sqrt{2\pi t}} e^{-\frac{x^2}{2t}} = p(x). \end{aligned}$$

The associated distribution functions are

$$\begin{aligned}
F^{P}_{W_t}(x) &= P(W_t \le x) = \int_{\{W_t \le x\}} dP(\omega) = \int_{-\infty}^{x} p(u)\,du; \\
F^{Q}_{\widehat{W}_t}(x) &= Q(\widehat{W}_t \le x) = \int_{\{\widehat{W}_t \le x\}} dQ(\omega) = \int_{-\infty}^{x} p(u)\,du.
\end{aligned}$$

Expressing W_t in terms of $\widehat{W}_t$ and using that $\widehat{W}_t$ is normally distributed with respect to Q we get the distribution function of W_t with respect to Q as

$$\begin{aligned}
F^{Q}_{W_t}(x) &= Q(W_t \le x) = Q(\widehat{W}_t - \eta t \le x) \\
&= Q(\widehat{W}_t \le x + \eta t) = \int_{\{\widehat{W}_t \le x + \eta t\}} dQ(\omega) \\
&= \int_{-\infty}^{x+\eta t} p(y)\,dy = \int_{-\infty}^{x+\eta t} \frac{1}{\sqrt{2\pi t}} e^{-\frac{y^2}{2t}}\,dy.
\end{aligned}$$

Differentiating yields the density function

$$\boxed{p_{_Q}(x) = \frac{d}{dx} F^{Q}_{W_t}(x) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{1}{2t}(y+\eta t)^2}.}$$

It is worth noting that $p_{_Q}(x)$ can be decomposed as

$$\boxed{p_{_Q}(x) = e^{-\eta x - \frac{1}{2}\eta^2 t} p(x),}$$

which makes the connection with the Girsanov theorem.

The distribution function of $\widehat{W}_t$ with respect to P can be worked out in a similar way

$$\begin{aligned}
F^{P}_{\widehat{W}_t}(x) &= P(\widehat{W}_t \le x) = P(W_t + \eta t \le x) \\
&= P(W_t \le x - \eta t) = \int_{\{W_t \le x - \eta t\}} dP(\omega) \\
&= \int_{-\infty}^{x-\eta t} p(y)\,dy = \int_{-\infty}^{x-\eta t} \frac{1}{\sqrt{2\pi t}} e^{-\frac{y^2}{2t}}\,dy,
\end{aligned}$$

so the density function is

$$\boxed{p_{_P}(x) = \frac{d}{dx} F^{P}_{\widehat{W}_t}(x) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{1}{2t}(y-\eta t)^2}.}$$

14.3 Correlation of Stocks

Consider two stock prices driven by the same novelty term

$$\begin{aligned} dS_1 &= \mu_1 S_1 dt + \sigma_1 S_1 dW_t &\quad (14.3.7)\\ dS_2 &= \mu_2 S_2 dt + \sigma_2 S_2 dW_t. &\quad (14.3.8)\end{aligned}$$

Since the underlying Brownian motions are perfectly correlated, one may be tempted to think that the stock prices $S_1$and S_2 are also correlated. The following result shows that in general the stock prices are positively correlated:

Proposition 14.3.1 *The correlation coefficient between the stock prices S_1 and S_2 driven by the same Brownian motion is*

$$Corr(S_1, S_2) = \frac{e^{\sigma_1\sigma_2 t} - 1}{(e^{\sigma_1^2 t} - 1)^{1/2}(e^{\sigma_2^2 t} - 1)^{1/2}} > 0.$$

In particular, if $\sigma_1 = \sigma_2$, then $Corr(S_1, S_2) = 1$.

Proof: Since

$$S_1(t) = S_1(0)e^{\mu_1 t - \frac{1}{2}\sigma_1^2 t}e^{\sigma_1 W_t}, \quad S_2(t) = S_2(0)e^{\mu_2 t - \frac{1}{2}\sigma_2^2 t}, e^{\sigma_2 W_t},$$

from Exercise 14.3.4 and formula $E[e^{kW_t}] = e^{k^2 t/2}$ we have

$$\begin{aligned} Corr(S_1, S_2) &= Corr(e^{\sigma_1 W_t}, e^{\sigma_2 W_t}) = \frac{Cov(e^{\sigma_1 W_t}, e^{\sigma_2 W_t})}{\sqrt{Var(e^{\sigma_1 W_t})Var(e^{\sigma_2 W_t})}}\\ &= \frac{E[e^{(\sigma_1+\sigma_2)W_t}] - E[e^{\sigma_1 W_t}]E[e^{\sigma_2 W_t}]}{\sqrt{Var(e^{\sigma_1 W_t})Var(e^{\sigma_2 W_t})}}\\ &= \frac{e^{\frac{1}{2}(\sigma_1+\sigma_2)^2 t} - e^{\frac{1}{2}\sigma_1^2 t}e^{\frac{1}{2}\sigma_2^2 t}}{[e^{\sigma_1^2 t}(e^{\sigma_1^2 t} - 1)e^{\sigma_2^2 t}(e^{\sigma_2^2 t} - 1)]^{1/2}}\\ &= \frac{e^{\sigma_1\sigma_2 t} - 1}{(e^{\sigma_1^2 t} - 1)^{1/2}(e^{\sigma_2^2 t} - 1)^{1/2}}.\end{aligned}$$

If $\sigma_1 = \sigma_2 = \sigma$ then the previous formula provides

$$Corr(S_1, S_2) = \frac{e^{\sigma^2 t} - 1}{e^{\sigma^2 t} - 1} = 1,$$

i.e. the stocks are perfectly correlated if they have the same volatility. ∎

Corollary 14.3.2 *The stock prices S_1 and S_2 are positively strongly correlated for small values of t:*

$$Corr(S_1, S_2) \to 1 \ \text{ as } \ t \to 0.$$

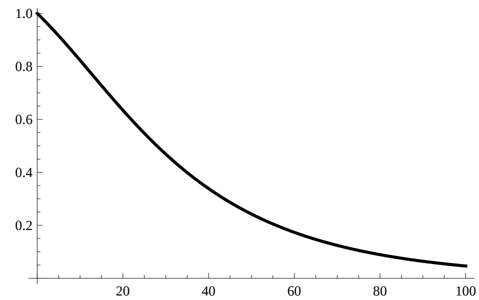


Figure 14.1: *The correlation function* $f(t) = \frac{e^{\sigma_1\sigma_2 t}-1}{(e^{\sigma_1^2 t}-1)^{1/2}(e^{\sigma_2^2 t}-1)^{1/2}}$ *in the case* $\sigma_1 = 0.15$, $\sigma_2 = 0.40$.

This fact has the following financial interpretation. If some stocks are driven by the same unpredictable news, when one stock increases, then the other one tends to increase too, at least for a small amount of time. In the case when some bad news affects an entire financial market, the risk becomes systemic, and hence if one stock fails, all the others tend to decrease as well, leading to a severe strain on the financial market.

Corollary 14.3.3 *The stock prices correlation gets weak as* t *gets large:*

$$Corr(S_1, S_2) \to 0 \quad as \quad t \to \infty.$$

Proof: It follows from the asymptotic correspondence

$$\frac{e^{\sigma_1\sigma_2 t} - 1}{(e^{\sigma_1^2 t} - 1)^{1/2}(e^{\sigma_2^2 t} - 1)^{1/2}} \sim \frac{e^{\sigma_1\sigma_2 t}}{e^{\frac{\sigma_1^2+\sigma_2^2}{2}t}} = e^{-\frac{(\sigma_1-\sigma_2)^2}{2}t} \to 0, \quad t \to 0.$$

■

It follows that in the long run any two stocks tend to become uncorrelated, see Fig.14.1.

Exercise 14.3.4 *If* X *and* Y *are random variables and* $\alpha, \beta \in \mathbb{R}$*, show that*

$$Corr(\alpha X, \beta Y) = \begin{cases} Corr(X, Y), & if\, \alpha\beta > 0 \\ -Corr(X, Y), & if\, \alpha\beta < 0. \end{cases}$$

Exercise 14.3.5 *Find the following*

(*a*) $Cov\big(dS_1(t), dS_2(t)\big)$;

(*b*) $Corr\big(dS_1(t), dS_2(t)\big)$.

14.4 Self-financing Portfolios

The value of a portfolio that contains $\theta_j(t)$ units of stock $S_j(t)$ at time t is given by

$$V(t) = \sum_{j=1}^{n} \theta_j(t) S_j(t).$$

The portfolio is called *self-financing* if

$$dV(t) = \sum_{j=1}^{n} \theta_j(t) dS_j(t).$$

This means that an infinitesinal change in the value of the portfolio is due to infinitesimal changes in stock values. All portfolios will be assumed self-financing, if otherwise stated.

14.5 The Sharpe Ratio

If μ is the expected return on the stock S_t, the *risk premium* is defined as the difference $\mu - r$, where r is the risk-free interest rate. The *Sharpe ratio*, η, is the quotient between the risk premium and stock price volatility

$$\boxed{\eta = \frac{\mu - r}{\sigma}.}$$

The following result shows that the Sharpe ratio is an important invariant for the family of stocks driven by the same uncertain source. It is also known under the name of the *market price of risk* for assets.

Proposition 14.5.1 *Let S_1 and S_2 be two stocks satisfying equations* (14.3.7)–(14.3.8). *Then their Sharpe ratio are equal*

$$\frac{\mu_1 - r}{\sigma_1} = \frac{\mu_2 - r}{\sigma_2}. \tag{14.5.9}$$

Proof: Eliminating the term dW_t from equations (14.3.7) – (14.3.8) yields

$$\frac{\sigma_2}{S_1} dS_1 - \frac{\sigma_1}{S_2} dS_2 = (\mu_1 \sigma_2 - \mu_2 \sigma_1) dt. \tag{14.5.10}$$

Consider the portfolio $P(t) = \theta_1(t) S_1(t) - \theta_2(t) S_2(t)$, with $\theta_1(t) = \dfrac{\sigma_2(t)}{S_1(t)}$ and $\theta_2(t) = \dfrac{\sigma_1(t)}{S_2(t)}$. Using the properties of self-financing portfolios,, we have

$$\begin{aligned} dP(t) &= \theta_1(t) dS_1(t) - \theta_2(t) dS_2(t) \\ &= \frac{\sigma_2}{S_1} dS_1 - \frac{\sigma_1}{S_2} dS_2. \end{aligned}$$

Substituting in (14.5.10) yields $dP = (\mu_1\sigma_2 - \mu_2\sigma_1)dt$, i.e. P is a risk-less portfolio. Since the portfolio earns interest at the risk-free interest rate, we have $dP = rPdt$. Then equating the coefficients of dt yields

$$\mu_1\sigma_2 - \mu_2\sigma_1 = rP(t).$$

Using the definition of $P(t)$, the previous relation becomes

$$\mu_1\sigma_2 - \mu_2\sigma_1 = r\theta_1 S_1 - r\theta_2 S_2,$$

that can be transformed to

$$\mu_1\sigma_2 - \mu_2\sigma_1 = r\sigma_2 - r\sigma_1,$$

which is equivalent with (14.5.9). ■

Using Proposition 14.1.4, relations (14.3.7) – (14.3.8) can be written as

$$\begin{aligned} dS_1 &= rS_1dt + \sigma_1 S_1 d\widehat{W}_t \\ dS_2 &= rS_2dt + \sigma_2 S_2 d\widehat{W}_t, \end{aligned}$$

where the risk-neutral process $d\widehat{W}_t$ is the same in both equations

$$d\widehat{W}_t = \frac{\mu_1 - r}{\sigma_1}dt + dW_t = \frac{\mu_2 - r}{\sigma_2}dt + dW_t.$$

This shows that the process $\widehat{W}_t$ is a Brownian motion with drift, where the drift is the Sharpe ratio.

14.6 Risk-neutral Valuation for Derivatives

The risk-neutral process $d\widehat{W}_t$ plays an important role in the risk neutral valuation of derivatives. In this section we shall prove that if f_T is the price of a derivative at the maturity time, then $f_t = \widehat{E}[e^{-r(T-t)}f_T|\mathcal{F}_t]$ is the price of the derivative at the time t, for any $t < T$.

In other words, the discounted price of a derivative in the risk-neutral world is the price of the derivative at the new instance of time. This is based on the fact that $e^{-rt}f_t$ is an $\mathcal{F}_t$-martingale with respect to the risk-neutral measure Q introduced previously.

In particular, the idea of the proof can be applied for the stock S_t. Applying the product rule

$$\begin{aligned} d(e^{-rt}S_t) &= d(e^{-rt})S_t + e^{-rt}dS_t + \underbrace{d(e^{-rt})dS_t}_{=0} \\ &= -re^{-rt}S_tdt + e^{-rt}(rS_tdt + \sigma S_t d\widehat{W}_t) \\ &= e^{-rt}(rS_tdt + \sigma S_t d\widehat{W}_t). \end{aligned}$$

If $u < t$, integrating between u and t

$$e^{-rt}S_t = e^{-ru}S_u + \int_u^t \sigma e^{-rs} S_s d\widehat{W}_s,$$

and taking the risk-neutral expectation with respect to the information set $\mathcal{F}_u$ yields

$$\begin{aligned}
\widehat{E}[e^{-rt}S_t|\mathcal{F}_u] &= \widehat{E}[e^{-ru}S_u + \int_u^t \sigma e^{-rs} S_s d\widehat{W}_s \,|\, \mathcal{F}_u] \\
&= e^{-ru}S_u + \widehat{E}\Big[\int_u^t \sigma e^{-rs} S_s d\widehat{W}_s|\mathcal{F}_u\Big] \\
&= e^{-ru}S_u + \widehat{E}\Big[\int_u^t \sigma e^{-rs} S_s d\widehat{W}_s\Big] \\
&= e^{-ru}S_u,
\end{aligned}$$

since $\displaystyle\int_u^t \sigma e^{-rs} S_s d\widehat{W}_s$ is independent of $\mathcal{F}_u$. It follows that $e^{-rt}S_t$ is an $\mathcal{F}_t$-martingale in the risk-neutral world. The following fundamental result can be shown using a similar proof as the one encountered previously:

Theorem 14.6.1 *If $f_t = f(t, S_t)$ is the price of a derivative at time t, then $e^{-rt}f_t$ is an $\mathcal{F}_t$-martingale in the risk-neutral world, i.e.*

$$\widehat{E}[e^{-rt}f_t|\mathcal{F}_u] = e^{-ru}f_u, \qquad \forall 0 < u < t.$$

Proof: Using Ito's formula and the risk neutral process $dS = rS + \sigma S d\widehat{W}_t$, the process followed by f_t is

$$\begin{aligned}
df_t &= \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial S}dS + \frac{1}{2}\frac{\partial^2 f}{\partial S^2}(dS)^2 \\
&= \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial S}(rSdt + \sigma S d\widehat{W}_t) + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 f}{\partial S^2}dt \\
&= \Big(\frac{\partial f}{\partial t} + rS\frac{\partial f}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 f}{\partial S^2}\Big)dt + \sigma S \frac{\partial f}{\partial S} d\widehat{W}_t \\
&= rf dt + \sigma S \frac{\partial f}{\partial S} d\widehat{W}_t,
\end{aligned}$$

where in the last identity we used that f satisfies the Black-Scholes equation. Applying the product rule we obtain

$$\begin{aligned}
d(e^{-rt}f_t) &= d(e^{-rt})f_t + e^{-rt}df_t + \underbrace{d(e^{-rt})df_t}_{=0} \\
&= -re^{-rt}f_t dt + e^{-rt}(rf_t dt + \sigma S \frac{\partial f}{\partial S} d\widehat{W}_t) \\
&= e^{-rt}\sigma S \frac{\partial f}{\partial S} d\widehat{W}_t.
\end{aligned}$$

Integrating between u and t we get

$$e^{-rt}f_t = e^{-ru}f_u + \int_u^t e^{-rs}\sigma S\frac{\partial f_s}{\partial S}d\widehat{W}_s,$$

which assures that $e^{-rt}f_t$ is a martingale, since $\widehat{W}_s$ is a Brownian motion process. Using that $e^{-ru}f_u$ is $\mathcal{F}_u$-predictable, and $\int_u^t e^{-rs}\sigma S\frac{\partial f_s}{\partial S}d\widehat{W}_s$ is independent of the information set $\mathcal{F}_u$, we have

$$\widehat{E}[e^{-rt}f_t|\mathcal{F}_u] = e^{-ru}f_u + \widehat{E}\big[\int_u^t e^{-rs}\sigma S\frac{\partial f_s}{\partial S}d\widehat{W}_s\big] = e^{-ru}f_u.$$

■

Exercise 14.6.2 *Show the following:*

(*a*) $\widehat{E}[e^{\sigma(W_t-W_u)}|\mathcal{F}_u] = e^{(r-\mu+\frac{1}{2}\sigma^2)(t-u)}, \quad u < t;$

(*b*) $\widehat{E}\Big[\frac{S_t}{S_u}|\mathcal{F}_u\Big] = e^{(\mu-\frac{1}{2}\sigma^2)(t-u)}\widehat{E}\big[e^{\sigma(W_t-W_u)}|\mathcal{F}_u\big], \quad u < t;$

(*c*) $\widehat{E}\Big[\frac{S_t}{S_u}|\mathcal{F}_u\Big] = e^{r(t-u)}, \quad u < t.$

Exercise 14.6.3 *Find the following risk-neutral world conditional expectations:*

(*a*) $\widehat{E}[\int_0^t S_u\,du|\mathcal{F}_s], \quad s < t;$

(*b*) $\widehat{E}[S_t\int_0^t S_u\,du|\mathcal{F}_s], \ \ s < t;$

(*c*) $\widehat{E}[\int_0^t S_u\,dW_u|\mathcal{F}_s], \ \ s < t;$

(*d*) $\widehat{E}[S_t\int_0^t S_u\,dW_u|\mathcal{F}_s], \ \ s < t;$

(*e*) $\widehat{E}[\big(\int_0^t S_u\,du\big)^2|\mathcal{F}_s], \quad s < t;$

Exercise 14.6.4 *Use risk-neutral valuation to find the price of a derivative that pays at maturity the following payoffs:*

(*a*) $f_T = TS_T;$

(*b*) $f_T = \int_0^T S_u\,du;$

(*c*) $f_T = \int_0^T S_u\,dW_u.$

Chapter 15

Black-Scholes Analysis

15.1 Heat Equation

This section is devoted to a basic discussion on the heat equation. Its importance resides in the remarkable fact that the Black-Scholes equation, which is the main equation of derivatives calculus, can be reduced to this type of equation.

Let $u(\tau, x)$ denote the temperature in an infinite rod at point x and time τ. In the absence of exterior heat sources the heat diffuses according to the following parabolic differential equation

$$\frac{\partial u}{\partial \tau} - \frac{\partial^2 u}{\partial x^2} = 0, \tag{15.1.1}$$

called *the heat equation*. If the initial heat distribution is known and is given by $u(0, x) = f(x)$, then we have an initial value problem for the heat equation.

Solving this equation involves a convolution between the initial temperature $f(x)$ and the fundamental solution of the heat equation $G(\tau, x)$, which will be defined shortly.

Definition 15.1.1 *The function*

$$G(\tau, x) = \frac{1}{\sqrt{4\pi\tau}} e^{-\frac{x^2}{4\tau}}, \qquad \tau > 0,$$

is called the fundamental solution of the heat equation (15.1.1).

We recall the most important properties of the function $G(\tau, x)$.

- $G(\tau, x)$ has the properties of a probability density[1], i.e.

    1. $G(\tau, x) > 0, \qquad \forall x \in \mathbb{R}, \tau > 0;$

[1] In fact it is a Gaussian probability density.

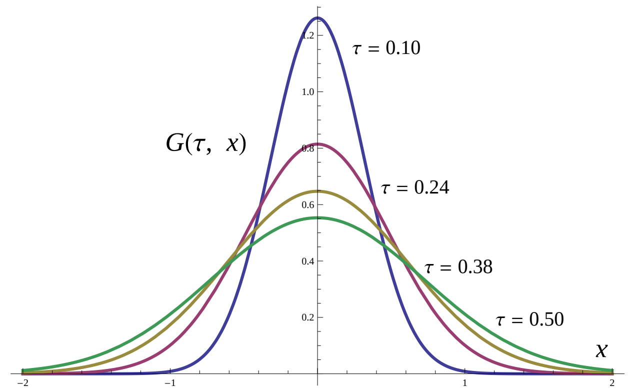


Figure 15.1: *The function $G(\tau, x)$ tends to the Dirac measure $\delta(x)$ as $\tau \searrow 0$, and flattens out as $\tau \to \infty$.*

2. $\displaystyle\int_{\mathbb{R}} G(\tau, x)\, dx = 1, \qquad \forall \tau > 0.$

- it satisfies the heat equation

$$\frac{\partial G}{\partial \tau} - \frac{\partial^2 G}{\partial x^2} = 0, \quad \tau > 0.$$

- G tends to the *Dirac measure* $\delta(x)$ as τ gets closer to the initial time

$$\lim_{\tau \searrow 0} G(\tau, x) = \delta(x),$$

where the Dirac measure can be defined using integration as

$$\int_{\mathbb{R}} \varphi(x)\delta(x)dx = \varphi(0),$$

for any smooth function with compact support φ. Consequently, we also have

$$\int_{\mathbb{R}} \varphi(x)\delta(x-y)dx = \varphi(y).$$

One can think of $\delta(x)$ as a measure with infinite value at $x = 0$ and zero for the rest of the values, and with the integral equal to 1, see Fig.15.1.

The physical significance of the fundamental solution $G(\tau, x)$ is that it describes the heat evolution in the infinite rod after an initial heat impulse of infinite size is applied at $x = 0$.

Proposition 15.1.2 *The solution of the initial value heat equation*

$$\begin{aligned}\frac{\partial u}{\partial \tau} - \frac{\partial^2 u}{\partial x^2} &= 0\\ u(0,x) &= f(x)\end{aligned}$$

is given by the convolution between the fundamental solution and the initial temperature

$$u(\tau, x) = \int_{\mathbb{R}} G(\tau, y-x) f(y)\, dy, \qquad \tau > 0.$$

Proof: Substituting $z = y - x$, the solution can be written as

$$u(\tau, x) = \int_{\mathbb{R}} G(\tau, z) f(x+z)\, dz. \tag{15.1.2}$$

Differentiating under the integral yields

$$\begin{aligned}\frac{\partial u}{\partial \tau} &= \int_{\mathbb{R}} \frac{\partial G(\tau, z)}{\partial \tau} f(x+z)\, dz,\\ \frac{\partial^2 u}{\partial x^2} &= \int_{\mathbb{R}} G(\tau, z) \frac{\partial^2 f(x+z)}{\partial x^2}\, dz = \int_{\mathbb{R}} G(\tau, z) \frac{\partial^2 f(x+z)}{\partial z^2}\, dz\\ &= \int_{\mathbb{R}} \frac{\partial^2 G(\tau, z)}{\partial z^2} f(x+z)\, dz,\end{aligned}$$

where we applied integration by parts twice and the fact that

$$\lim_{z\to\infty} G(\tau, z) = \lim_{z\to\infty} \frac{\partial G(\tau, z)}{\partial z} = 0.$$

Since G satisfies the heat equation,

$$\frac{\partial u}{\partial \tau} - \frac{\partial^2 u}{\partial x^2} = \int_{\mathbb{R}} \Big[\frac{\partial G(\tau, z)}{\partial \tau} - \frac{\partial^2 G(\tau, z)}{\partial z^2}\Big] f(x+z)\, dz = 0.$$

Since the limit and the integral commute[2], using the properties of the Dirac measure, we have

$$\begin{aligned}u(0,x) &= \lim_{\tau\searrow 0} u(\tau, x) = \lim_{\tau \searrow 0} \int_{\mathbb{R}} G(\tau, z) f(x+z)\, dz\\ &= \int_{\mathbb{R}} \delta(z) f(x+z)\, dz = f(x).\end{aligned}$$

Hence (15.1.2) satisfies the initial value heat equation. ■

It is worth noting that the solution $u(\tau, x) = \int_{\mathbb{R}} G(y-x, \tau) f(y)\, dy$ provides the temperature at any point in the rod for any time $\tau > 0$, but it cannot provide the

[2]This is allowed by the dominated convergence theorem.

temperature for $\tau < 0$, because of the singularity the fundamental solution exhibits at $\tau = 0$. We can reformulate this by saying that the heat equation is *semi-deterministic*, in the sense that given the present, we can know the future but not the past.

The semi-deterministic character of diffusion phenomena can be exemplified with a drop of ink which starts diffusing in a bucket of water at time $t = 0$. We can determine the density of the ink at any time $t > 0$ at any point x in the bucket. However, given the density of ink at a time $t > 0$, it is not possible to trace back in time the ink distribution density and to find the initial point where the drop started its diffusion.

The semi-deterministic behavior occurs in the study of derivatives too. In the case of the Black-Scholes equation, which is a backwards heat equation[3], given the present value of the derivative, we can find the past values but not the future ones. This is the capital difficulty in foreseeing the prices of stock market instruments from the present prices. This difficulty will be overcome by working the price from the given final condition, which is the payoff at maturity.

15.2 What is a Portfolio?

A *portfolio* is a position in the market that consists in long and short positions in one or more stocks and other securities. The value of a portfolio can be represented algebraically as a linear combination of stock prices and other securities' values:

$$P = \sum_{j=1}^{n} a_j S_j + \sum_{k=1}^{m} b_k F_k.$$

The market participant holds a_j units of stock S_j and b_k units in derivative F_k. The coefficients are positive for long positions and negative for short positions. For instance, a portfolio given by $P = 2F - 3S$ means that we buy 2 securities and sell 3 units of stock (a position with 2 securities long and 3 stocks short).

The portfolio is self-financing if

$$dP = \sum_{j=1}^{n} a_j dS_j + \sum_{k=1}^{m} b_k dF_k.$$

15.3 Risk-less Portfolios

A portfolio P is called *risk-less* if the increments dP are completely predictable. In this case the increments' value dP should equal the interest earned in the time interval dt on the portfolio P. This can be written as

$$dP = rPdt, \tag{15.3.3}$$

[3]This comes from the fact that at some point τ becomes $-\tau$ due to a substitution

where r denotes the risk-free interest rate. For the sake of simplicity the rate r will be assumed constant throughout this section.

Let's assume now that the portfolio P depends on only one stock S and one derivative F, whose underlying asset is S. The portfolio depends also on time t, so

$$P = P(t, S, F).$$

We are interested in deriving the stochastic differential equation followed by the portfolio P. We note that at this moment the portfolio is not assumed risk-less. By Ito's formula we get

$$dP = \frac{\partial P}{\partial t}dt + \frac{\partial P}{\partial S}dS + \frac{\partial P}{\partial F}dF + \frac{1}{2}\frac{\partial^2 P}{\partial S^2}dS^2 + \frac{1}{2}\frac{\partial^2 P}{\partial F^2}(dF)^2. \tag{15.3.4}$$

The stock S is assumed to follow the geometric Brownian motion

$$dS = \mu S dt + \sigma S dW_t, \tag{15.3.5}$$

where the expected return rate on the stock μ and the stock's volatility σ are constants. Since the derivative F depends on time and underlying stock, we can write $F = F(t, S)$. Applying Ito's formula, yields

$$\begin{aligned} dF &= \frac{\partial F}{\partial t}dt + \frac{\partial F}{\partial S}dS + \frac{1}{2}\frac{\partial^2 F}{\partial S^2}(dS)^2 \\ &= \Big(\frac{\partial F}{\partial t} + \mu S\frac{\partial F}{\partial S} + \frac{1}{2}\sigma^2 S^2\frac{\partial^2 F}{\partial S^2}\Big)dt + \sigma S\frac{\partial F}{\partial S}dW_t, \end{aligned} \tag{15.3.6}$$

where we have used (15.3.5). Taking the squares in relations (15.3.5) and (15.3.6), and using the stochastic relations $(dW_t)^2 = dt$ and $dt^2 = dW_t dt = 0$, we get

$$\begin{aligned} (dS)^2 &= \sigma^2 S^2 dt \\ (dF)^2 &= \sigma^2 S^2 \Big(\frac{\partial F}{\partial S}\Big)^2 dt. \end{aligned}$$

Substituting back in (15.3.4), and collecting the predictable and unpredictable parts, yields

$$\begin{aligned} dP = \; & \Big[\frac{\partial P}{\partial t} + \mu S\Big(\frac{\partial P}{\partial S} + \frac{\partial P}{\partial F}\frac{\partial F}{\partial S}\Big) + \frac{\partial P}{\partial F}\Big(\frac{\partial F}{\partial t} + \frac{1}{2}\sigma^2 S^2\frac{\partial^2 F}{\partial S^2}\Big) \\ & + \frac{1}{2}\sigma^2 S^2\Big(\frac{\partial^2 P}{\partial S^2} + \frac{\partial^2 P}{\partial F^2}\Big(\frac{\partial F}{\partial S}\Big)^2\Big)\Big]dt \\ & + \sigma S\Big(\frac{\partial P}{\partial S} + \frac{\partial P}{\partial F}\frac{\partial F}{\partial S}\Big)dW_t. \end{aligned} \tag{15.3.7}$$

Looking at the unpredictable component, we have the following result:

Proposition 15.3.1 *The portfolio P is risk-less if and only if* $\dfrac{dP}{dS} = 0$.

Proof: A portfolio P is risk-less if and only if its unpredictable component is identically zero, i.e.

$$\frac{\partial P}{\partial S} + \frac{\partial P}{\partial F}\frac{\partial F}{\partial S} = 0.$$

Since the total derivative of P is given by

$$\frac{dP}{dS} = \frac{\partial P}{\partial S} + \frac{\partial P}{\partial F}\frac{\partial F}{\partial S},$$

the previous relation becomes $\dfrac{dP}{dS} = 0$. ■

Definition 15.3.2 *The amount* $\Delta_P = \dfrac{dP}{dS}$ *is called the delta of the portfolio P.*

The previous result can be reformulated by saying that a portfolio is risk-less if and only if its delta vanishes. In practice this can hold only for a short amount of time, so the portfolio needs to be re-balanced periodically. The process of making a portfolio risk-less involves a procedure called delta hedging, through which the portfolio's delta becomes zero or very close to this value.

Assume P is a risk-less portfolio, so

$$\frac{dP}{dS} = \frac{\partial P}{\partial S} + \frac{\partial P}{\partial F}\frac{\partial F}{\partial S} = 0. \tag{15.3.8}$$

Then equation (15.3.7) simplifies to

$$\begin{aligned} dP \;=\; & \Big[\frac{\partial P}{\partial t} + \frac{\partial P}{\partial F}\Big(\frac{\partial F}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 F}{\partial S^2}\Big) \\ & + \frac{1}{2}\sigma^2 S^2 \Big(\frac{\partial^2 P}{\partial S^2} + \frac{\partial^2 P}{\partial F^2}\Big(\frac{\partial F}{\partial S}\Big)^2\Big)\Big] dt. \end{aligned} \tag{15.3.9}$$

Comparing with (15.3.3) yields

$$\boxed{\frac{\partial P}{\partial t} + \frac{\partial P}{\partial F}\Big(\frac{\partial F}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 F}{\partial S^2}\Big) + \frac{1}{2}\sigma^2 S^2 \Big(\frac{\partial^2 P}{\partial S^2} + \frac{\partial^2 P}{\partial F^2}\Big(\frac{\partial F}{\partial S}\Big)^2\Big) = rP.} \tag{15.3.10}$$

This is the equation satisfied by a risk-free financial instrument, $P = P(t, S, F)$, that depends on time t, stock S and derivative price F.

15.4 Black-Scholes Equation

This section deals with a parabolic partial differential equation satisfied by all European-type securities, called the *Black-Scholes* equation. This was initially used by Black and Scholes to find the value of options. This is a deterministic equation obtained by eliminating the unpredictable component of the derivative by making a risk-less portfolio. The main reason for this being possible is the fact that both the derivative F and the stock S are driven by the same source of uncertainty.

The next result holds in a market with the following restrictive conditions:

- the risk-free rate r and stock volatility σ are constant.
- there are no arbitrage opportunities.
- no transaction costs.

Proposition 15.4.1 *If $F(t,S)$ is a derivative defined for $t \in [0,T]$, then*

$$\boxed{\frac{\partial F}{\partial t} + rS\frac{\partial F}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 F}{\partial S^2} = rF.} \tag{15.4.11}$$

Proof: The equation (15.3.10) works under the general hypothesis that $P = P(t,S,F)$ is a risk-free financial instrument that depends on time t, stock S and derivative F. We shall consider P to be the following particular portfolio

$$P = F - \lambda S.$$

This means to take a long position in derivative and a short position in λ units of stock (assuming λ positive). The partial derivatives in this case are

$$\begin{aligned} \frac{\partial P}{\partial t} &= 0, \qquad \frac{\partial P}{\partial F} = 1, \qquad \frac{\partial P}{\partial S} = -\lambda, \\ \frac{\partial^2 P}{\partial F^2} &= 0, \qquad \frac{\partial^2 P}{\partial S^2} = 0. \end{aligned}$$

From the risk-less property (15.3.8) we get $\lambda = \dfrac{\partial F}{\partial S}$. Substituting into equation (15.3.10) yields

$$\frac{\partial F}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 F}{\partial S^2} = rF - rS\frac{\partial F}{\partial S},$$

which is equivalent to the desired equation. ∎

However, the Black-Scholes equation is derived most often in a less rigorous way. It is based on the assumption that the number $\lambda = \frac{\partial F}{\partial S}$, which appears in the formula

of the risk-less portfolio $P = F - \lambda S$, is considered constant for the time interval Δt. If we consider the increments over the time interval Δt

$$\begin{aligned} \Delta W_t &= W_{t+\Delta t} - W_t \\ \Delta S &= S_{t+\Delta t} - S_t \\ \Delta F &= F(t+\Delta t, S_t + \Delta S) - F(t,S), \end{aligned}$$

then Ito's formula yields

$$\begin{aligned} \Delta F &= \Big(\frac{\partial F}{\partial t}(t,S) + \mu S \frac{\partial F}{\partial S}(t,S) + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 F}{\partial S^2}(t,S)\Big)\Delta t \\ &\quad + \sigma S \frac{\partial F}{\partial S}(t,S)\Delta W_t. \end{aligned}$$

On the other side, the increments in the stock are given by

$$\Delta S = \mu S \Delta t + \sigma S \Delta W_t.$$

Since both increments ΔF and ΔS are driven by the same uncertainly source, ΔW_t, we can eliminate it by multiplying the latter equation by $\frac{\partial F}{\partial S}$ and subtract it from the former

$$\Delta F - \frac{\partial F}{\partial S}(t,S)\Delta S = \Big(\frac{\partial F}{\partial t}(t,S) + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 F}{\partial S^2}(t,S)\Big)\Delta t.$$

The left side can be regarded as the increment ΔP, of the portfolio

$$P = F - \frac{\partial F}{\partial S} S.$$

This portfolio is risk-less because its increments are totally deterministic, so it must also satisfy $\Delta P = rP\Delta t$. The number $\frac{\partial F}{\partial S}$ is assumed constant for small intervals of time Δt. Even if this assumption is not rigorous enough, the procedure still leads to the right equation. This is obtained by equating the coefficients of Δt in the last two equations

$$\frac{\partial F}{\partial t}(t,S) + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 F}{\partial S^2}(t,S) = r\Big(F - \frac{\partial F}{\partial S} S\Big),$$

which is equivalent to the Black-Scholes equation.

15.5 Delta Hedging

The proof for the Black-Scholes' equation is based on the fact that the portfolio $P = F - \dfrac{\partial F}{\partial S} S$ is risk-less. Since the delta of the derivative F is

$$\Delta_F = \frac{dF}{dS} = \frac{\partial F}{\partial S},$$

then the portfolio $P = F - \Delta_F S$ is risk-less. This leads to the *delta-hedging* procedure, by which selling Δ_F units of the underlying stock S yields a risk-less investment.

Exercise 15.5.1 *Find the value of the portfolio $P = F - \Delta_F S$ in th ecase when F is a call option.*

$$P = F - \Delta_F S = c - N(d_1)S = -Ke^{-r(T-t)}.$$

15.6 Tradeable securities

A derivative $F(t, S)$ that is a solution of the Black-Scholes equation is called *tradeable*. Its name comes from the fact that it can be traded (either on an exchange or over-the-counter). The Black-Scholes equation constitutes the equilibrium relation that provides the traded price of the derivative. We shall deal next with a few examples of tradeable securities.

Example 15.6.1 (i) *It is easy to show that $F = S$ is a solution of the Black-Scholes equation. Hence the stock is a tradeable security.*
(ii) *If K is a constant, then $F = e^{rt}K$ is a tradeable security.*
(iii) *If S is the stock price, then $F = e^S$ is not a tradeable security, since F does not satisfy equation (15.4.11).*

Exercise 15.6.1 *Show that $F = \ln S$ is not a tradeable security.*

Exercise 15.6.2 *Find all constants α such that S^α is tradeable.*

Substituting $F = S^\alpha$ into equation (15.4.11) we obtain

$$rS\alpha S^{\alpha-1} + \frac{1}{2}\sigma^2 S^2 \alpha(\alpha - 1)S^{\alpha-2} = rS^\alpha.$$

Dividing by S^α yields $r\alpha + \dfrac{1}{2}\sigma^2\alpha(\alpha - 1) = r$. This can be factorized as

$$\frac{1}{2}\sigma^2(\alpha - 1)(\alpha + \frac{2r}{\sigma^2}) = 0,$$

with two distinct solutions $\alpha_1 = 1$ and $\alpha_2 = -\dfrac{2r}{\sigma^2}$. Hence there are only two tradeable securities that are powers of the stock: the stock itself, S, and S^{-2r/σ^2}. In particular, S^2 is not tradeable, since $-2r/\sigma^2 \neq 2$ (the left side is negative). The role of these two cases will be clarified by the next result.

Proposition 15.6.3 *The general form of a traded derivative, which does not depend explicitly on time, is given by*

$$F(S) = C_1 S + C_2 S^{-2r/\sigma^2}, \tag{15.6.12}$$

with C_1, C_2 constants.

Proof: If the derivative depends solely on the stock, $F = F(S)$, then the Black-Scholes equation becomes the ordinary differential equation

$$rS\frac{dF}{dS} + \frac{1}{2}\sigma^2 S^2 \frac{d^2F}{dS^2} = rF. \tag{15.6.13}$$

This is an Euler-type equation, which can be solved by using the substitution $S = e^x$. The derivatives $\dfrac{d}{dS}$ and $\dfrac{d}{dx}$ are related by the chain rule

$$\frac{d}{dx} = \frac{dS}{dx}\frac{d}{dS}.$$

Since $\dfrac{dS}{dx} = \dfrac{de^x}{dx} = e^x = S$, it follows that $\dfrac{d}{dx} = S\dfrac{d}{dS}$. Using the product rule,

$$\frac{d^2}{dx^2} = S\frac{d}{dS}\big(S\frac{d}{dS}\big) = S\frac{d}{dS} + S^2\frac{d^2}{dS^2},$$

and hence

$$S^2\frac{d^2}{dS^2} = \frac{d^2}{dx^2} - \frac{d}{dx}.$$

Substituting into (15.6.13) yields

$$\frac{1}{2}\sigma^2\frac{d^2G(x)}{dx^2} + (r - \frac{1}{2}\sigma^2)\frac{dG(x)}{dx} = rG(x).$$

where $G(x) = G(e^x) = F(S)$. The associated indicial equation

$$\frac{1}{2}\sigma^2\alpha^2 - (r - \frac{1}{2}\sigma^2)\alpha = r$$

has solutions $\alpha_1 = 1$, $\alpha_2 = -r/\sigma^2$, so the general solution has the form

$$G(x) = C_1 e^x + C_2 e^{-\frac{r}{\sigma^2}x},$$

which is equivalent with (15.6.12). ■

Exercise 15.6.4 *Show that the price of a forward contract, which is given by $F(t, S) = S - Ke^{-r(T-t)}$, satisfies the Black-Scholes equation, i.e. a forward contract is a tradeable derivative.*

Exercise 15.6.5 *Show that the bond $F(t) = e^{-r(T-t)}K$ is a tradeable security.*

Exercise 15.6.6 *Let d_1 and d_2 be given by*

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

Show that the following functions satisfy the Black-Scholes equation:

(*a*) $F_1(t,S) = SN(d_1)$

(*b*) $F_2(t,S) = e^{-r(T-t)}N(d_2)$

(*c*) $F_2(t,S) = SN(d_1) - Ke^{-r(T-t)}N(d_2)$.

To which well-known derivatives do these formulas correspond?

15.7 Risk-less investment revised

A risk-less investment, $P(t,S,F)$, which depends on time t, stock price S and derivative F, and has S as underlying asset, satisfies equation (15.3.10). Using the Black-Scholes equation satisfied by the derivative F

$$\frac{\partial F}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 F}{\partial S^2} = rF - rS\frac{\partial F}{\partial S},$$

equation (15.3.10) becomes

$$\frac{\partial P}{\partial t} + \frac{\partial P}{\partial F}\Big(rF - rS\frac{\partial F}{\partial S}\Big) + \frac{1}{2}\sigma^2 S^2\Big(\frac{\partial^2 P}{\partial S^2} + \frac{\partial^2 P}{\partial F^2}\Big(\frac{\partial F}{\partial S}\Big)^2\Big) = rP.$$

Using the risk-less condition (15.3.8)

$$\frac{\partial P}{\partial S} = -\frac{\partial P}{\partial F}\frac{\partial F}{\partial S}, \tag{15.7.14}$$

the previous equation becomes

$$\boxed{\frac{\partial P}{\partial t} + rS\frac{\partial P}{\partial S} + rF\frac{\partial P}{\partial F} + \frac{1}{2}\sigma^2 S^2\Big[\frac{\partial^2 P}{\partial S^2} + \frac{\partial^2 P}{\partial F^2}\Big(\frac{\partial F}{\partial S}\Big)^2\Big] = rP.} \tag{15.7.15}$$

In the following we shall find an equivalent expression for the last term on the left side. Differentiating in (15.7.14) with respect to F yields

$$\begin{aligned}\frac{\partial^2 P}{\partial F \partial S} &= -\frac{\partial^2 P}{\partial F^2}\frac{\partial F}{\partial S} - \frac{\partial P}{\partial F}\frac{\partial^2 F}{\partial F \partial S} \\ &= -\frac{\partial^2 P}{\partial F^2}\frac{\partial F}{\partial S},\end{aligned}$$

where we used

$$\frac{\partial^2 F}{\partial F \partial S} = \frac{\partial}{\partial S}\frac{\partial F}{\partial F} = 1.$$

Multiplying by $\dfrac{\partial F}{\partial S}$ implies

$$\frac{\partial^2 P}{\partial F^2}\Big(\frac{\partial F}{\partial S}\Big)^2 = -\frac{\partial^2 P}{\partial F \partial S}\frac{\partial F}{\partial S}.$$

Substituting in the aforementioned equation yields the Black-Scholes equation for portfolios

$$\boxed{\frac{\partial P}{\partial t}+rS\frac{\partial P}{\partial S}+rF\frac{\partial P}{\partial F}+\frac{1}{2}\sigma^2S^2\Big[\frac{\partial^2 P}{\partial S^2}-\frac{\partial^2 P}{\partial F\partial S}\frac{\partial F}{\partial S}\Big]=rP.} \tag{15.7.16}$$

We have seen in section 15.4 that $P=F-\frac{\partial F}{\partial S}S$ is a risk-less investment, in fact a risk-less portfolio. We shall discuss in the following another risk-less investment.

Application 15.7.1 *If a risk-less investment P has the variable S and F separable, i.e. it is the sum $P(S,F)=f(F)+g(S)$, with f and g smooth functions, then*

$$P(S,F)=F+c_1S+c_2S^{-2r/\sigma^2},$$

with c_1, c_2 constants. The derivative F is given by the formula

$$F(t,S)=-c_1S-c_2S^{-2r/\sigma^2}+c_3e^{rt}, \qquad c_3\in\mathbb{R}.$$

Since P has separable variables, the mixed derivative term vanishes, and the equation (15.7.16) becomes

$$Sg'(S)+\frac{\sigma^2}{2r}S^2g''(S)-g(S)=f(F)-Ff'(F).$$

There is a separation constant C such that

$$\begin{aligned} f(F)-Ff'(F)&=C\\ Sg'(S)+\frac{\sigma^2}{2r}S^2g''(S)-g(S)&=C.\end{aligned}$$

Dividing the first equation by F^2 yields the exact equation

$$\Big(\frac{1}{F}f(F)\Big)'=-\frac{C}{F^2},$$

with the solution $f(F)=c_0F+C$. To solve the second equation, let $\kappa=\frac{\sigma^2}{2r}$. Then the substitution $S=e^x$ leads to the ordinary differential equation with constant coefficients

$$\kappa h''(x)+(1-\kappa)h'(x)-h(x)=C,$$

where $h(x)=g(e^x)=g(S)$. The associated indicial equation

$$\kappa\lambda^2+(1-\kappa)\lambda-1=0$$

has the solutions $\lambda_1=1$, $\lambda_2=-\frac{1}{\kappa}$. The general solution is the sum between the particular solution $h_p(x)=-C$ and the solution of the associated homogeneous equation, which is $h_0(x)=c_1e^x+c_2e^{-\frac{1}{\kappa}x}$. Then

$$h(x)=c_1e^x+c_2e^{-\frac{1}{\kappa}x}-C.$$

Going back to the variable S, we get the general form of $g(S)$

$$g(S) = c_1 S + c_2 S^{-2r/\sigma^2} - C,$$

with c_1, c_2 constants. Since the constant C cancels by addition, we have the following formula for the risk-less investment with separable variables F and S:

$$P(S,F) = f(F) + g(S) = c_0 F + c_1 S + c_2 S^{-2r/\sigma^2}.$$

Dividing by c_0, we may assume $c_0 = 1$. We shall find the derivative $F(t,S)$ which enters the previous formula. Substituting in (15.7.14) yields

$$-\frac{\partial F}{\partial S} = c_1 - \frac{2r}{\sigma^2} c_2 S^{-1-2r/\sigma^2},$$

which after partial integration in S gives

$$F(t,S) = -c_1 S - c_2 S^{-2r/\sigma^2} + \phi(t),$$

where the integration constant $\phi(t)$ is a function of t. The sum of the first two terms is the derivative given by formula (15.6.12). The remaining function $\phi(t)$ has also to satisfy the Black-Scholes equation, and hence it is of the form $\phi(t) = c_3 e^{rt}$, with c_3 constant. Then the derivative F is given by

$$F(t,S) = -c_1 S - c_2 S^{-2r/\sigma^2} + c_3 e^{rt}.$$

It is worth noting that substituting in the formula of P yields $P = c_3 e^{rt}$, which agrees with the formula of a risk-less investment.

Example 15.7.1 *Find the function $g(S)$ such that the product $P = Fg(S)$ is a risk-less investment, with $F = F(t,S)$ derivative. Find the expression of the derivative F in terms of S and t.*

Proof: Substituting $P = Fg(S)$ into equation (15.7.15) and simplifying by rF yields

$$S\frac{dg(S)}{dS} + \frac{\sigma^2}{2r} S^2 \frac{d^2 g(S)}{dS^2} = 0.$$

Substituting $S = e^x$, and $h(x) = g(e^x) = g(S)$ yields

$$h''(x) + \Big(\frac{2r}{\sigma^2} - 1\Big) h'(x) = 0.$$

Integrating leads to the solution

$$h(x) = C_1 + C_2 e^{(1-\frac{2r}{\sigma^2})x}.$$

Going back to variable S

$$g(S) = h(\ln S) = C_1 + C_2 e^{(1-\frac{2r}{\sigma^2})\ln S} = C_1 + C_2 S^{1-\frac{2r}{\sigma^2}}.$$

■

15.8 Solving the Black-Scholes

In this section we shall solve the Black-Scholes equation and show that its solution coincides with the one provided by the risk-neutral evaluation in section 13. This way, the Black-Scholes equation provides a variant approach for European-type derivatives by using partial differential equations instead of expectations.

Consider a European-type derivative F, with the payoff at maturity T given by f_T, which is a function of the stock price at maturity, S_T. Then $F(t, S)$ satisfies the following final condition partial differential equation

$$\begin{aligned} \frac{\partial F}{\partial t} + rS\frac{\partial F}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 F}{\partial S^2} &= rF \\ F(T, S_T) &= f_T(S_T). \end{aligned}$$

This means the solution is known at the final time T and we need to find its expression at any time t prior to T, i.e.

$$f_t = F(t, S_t), \qquad 0 \le t < T.$$

First we shall transform the equation into an equation with constant coefficients. Substituting $S = e^x$, and using the identities

$$S\frac{\partial}{\partial S} = \frac{\partial}{\partial x}, \qquad S^2\frac{\partial^2}{\partial S^2} = \frac{\partial^2}{\partial x^2} - \frac{\partial}{\partial x}$$

the equation becomes

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2\frac{\partial^2 V}{\partial x^2} + (r - \frac{1}{2}\sigma^2)\frac{\partial V}{\partial x} = rV,$$

where $V(t, x) = F(t, e^x)$. Using the time scaling $\tau = \frac{1}{2}\sigma^2(T - t)$, the chain rule provides

$$\frac{\partial}{\partial t} = \frac{\partial \tau}{\partial t}\frac{\partial}{\partial \tau} = -\frac{1}{2}\sigma^2\frac{\partial}{\partial \tau}.$$

Denote $k = \dfrac{2r}{\sigma^2}$. Substituting in the aforementioned equation yields

$$\frac{\partial W}{\partial \tau} = \frac{\partial^2 W}{\partial x^2} + (k - 1)\frac{\partial W}{\partial x} - kW, \tag{15.8.17}$$

where $W(\tau, x) = V(t, x)$. Next we shall get rid of the last two terms on the right side of the equation by using a crafted substitution.

Consider $W(\tau, x) = e^{\varphi}u(\tau, x)$, where $\varphi = \alpha x + \beta\tau$, with α, β constants that will be determined such that the equation satisfied by $u(\tau, x)$ has on the right side only the

second derivative in x. Since

$$\begin{aligned}\frac{\partial W}{\partial x} &= e^{\varphi}\Big(\alpha u + \frac{\partial u}{\partial x}\Big)\\ \frac{\partial^2 W}{\partial x^2} &= e^{\varphi}\Big(\alpha^2 u + 2\alpha\frac{\partial u}{\partial x} + \frac{\partial u}{\partial x^2}\Big)\\ \frac{\partial W}{\partial \tau} &= e^{\varphi}\Big(\beta u + \frac{\partial u}{\partial \tau}\Big),\end{aligned}$$

substituting in (15.8.17), dividing by e^{φ} and collecting the derivatives yields

$$\frac{\partial u}{\partial \tau} = \frac{\partial^2 u}{\partial x^2} + \big(2\alpha + k - 1\big)\frac{\partial u}{\partial x} + \big(\alpha^2 + \alpha(k-1) - k - \beta\big)u = 0$$

The constants α and β are chosen such that the coefficients of $\frac{\partial u}{\partial x}$ and u vanish

$$\begin{aligned}2\alpha + k - 1 &= 0\\ \alpha^2 + \alpha(k-1) - k - \beta &= 0.\end{aligned}$$

Solving yields

$$\begin{aligned}\alpha &= -\frac{k-1}{2}\\ \beta &= \alpha^2 + \alpha(k-1) - k = -\frac{(k+1)^2}{4}.\end{aligned}$$

The function $u(\tau, x)$ satisfies the heat equation

$$\frac{\partial u}{\partial \tau} = \frac{\partial^2 u}{\partial x^2}$$

with the initial condition expressible in terms of f_T

$$\begin{aligned}u(0,x) &= e^{-\varphi(0,x)}W(0,x) = e^{-\alpha x}V(T,x)\\ &= e^{-\alpha x}F(T,e^x) = e^{-\alpha x}f_T(e^x).\end{aligned}$$

From the general theory of heat equation, the solution can be expressed as the convolution between the fundamental solution and the initial condition

$$u(\tau,x) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{4\pi\tau}} e^{-\frac{(y-x)^2}{4\tau}} u(0,y)\,dy$$

The previous substitutions yield the following relation between F and u

$$F(t,S) = F(t,e^x) = V(t,x) = W(\tau,x) = e^{\varphi(\tau,x)}u(\tau,x),$$

so $F(T,e^x) = e^{\alpha x}u(0,x)$. This implies

$$
\begin{aligned}
F(t,e^x) &= e^{\varphi(\tau,x)}u(\tau,x) = e^{\varphi(\tau,x)}\int_{-\infty}^{\infty}\frac{1}{\sqrt{4\pi\tau}}e^{-\frac{(y-x)^2}{4\tau}}u(0,y)\,dy \\
&= e^{\varphi(\tau,x)}\int_{-\infty}^{\infty}\frac{1}{\sqrt{4\pi\tau}}e^{-\frac{(y-x)^2}{4\tau}}e^{-\alpha y}F(T,e^y)\,dy.
\end{aligned}
$$

With the substitution $y = x = s\sqrt{2\tau}$ this becomes

$$
F(t,e^x) = e^{\varphi(\tau,x)}\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{s^2}{2}-\alpha(x+s\sqrt{2\tau})}F(T,e^{x+s\sqrt{2\tau}})\,dy.
$$

Completing the square as

$$
-\frac{s^2}{2}-\alpha(x+s\sqrt{2\tau}) = \frac{1}{2}\Big(s-\frac{k-1}{2}\sqrt{2\tau}\Big)^2+\frac{(k-1)^2\tau}{4}+\frac{k-1}{2}x,
$$

after cancelations, the previous integral becomes

$$
F(t,e^x) = e^{-\frac{(k+1)^2}{4}\tau}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{-\frac{1}{2}\left(s-\frac{k-1}{2}\sqrt{2\tau}\right)^2}e^{\frac{(k-1)^2}{4}\tau}F(T,e^{x+s\sqrt{2\tau}})\,ds.
$$

Using

$$
\begin{aligned}
e^{-\frac{(k+1)^2}{4}\tau}e^{\frac{(k-1)^2}{4}\tau} &= e^{-k\tau} = e^{-r(T-t)}, \\
(k-1)\tau &= (r-\frac{1}{2}\sigma^2)(T-t),
\end{aligned}
$$

after the substitution $z = x + s\sqrt{2\tau}$ we get

$$
\begin{aligned}
F(t,e^x) &= e^{-r(T-t)}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{-\frac{1}{2}\frac{(z-x-(k-1)\tau)^2}{2\tau}}F(T,e^z)\frac{1}{\sqrt{2\tau}}\,dz \\
&= e^{-r(T-t)}\frac{1}{\sqrt{2\pi\sigma^2(T-t)}}\int_{-\infty}^{\infty}e^{-\frac{[z-x-(r-\frac{1}{2}\sigma^2)(T-t)]^2}{2\sigma^2(T-t)}}F(T,e^z)\,dz.
\end{aligned}
$$

Since $e^x = S_t$, considering the probability density

$$
p(z) = \frac{1}{\sqrt{2\pi\sigma^2(T-t)}}e^{-\frac{[z-\ln S_t-(r-\frac{1}{2}\sigma^2)(T-t)]^2}{2\sigma^2(T-t)}},
$$

the previous expression becomes

$$
F(t,S_t) = e^{-r(T-t)}\int_{-\infty}^{\infty}p(z)f_T(e^z)\,dz = e^{-r(T-t)}\widehat{E}_t[f_T],
$$

with $f_T(S_T) = F(T,S_T)$ and $\widehat{E}_t$ the risk-neutral expectation operator as of time t, which was introduced and used in section 13.

15.9 Black-Scholes and Risk-neutral Valuation

The conclusion of the computation of the last section is of capital importance for derivatives calculus. It shows the equivalence between the Black-Scholes equation and the risk-neutral evaluation. It turns out that instead of computing the risk-neutral expectation of the payoff, as in the case of the risk-neutral evaluation, we may have the choice to solve the Black-Scholes equation directly, and impose the final condition to be the payoff.

In many cases solving a partial differential equation is simpler than evaluating the expectation integral. This is due to the fact that we may look for a solution dictated by the particular form of the payoff f_T. We shall apply that in finding put-call parities for different types of derivatives.

Consequently, all derivatives evaluated by the risk-neutral valuation are solutions of the Black-Scholes equation. The only distinction is their payoff. A few of them are given in the next example.

Example 15.9.1 (a) *The price of a European call option is the solution $F(t, S)$ of the Black-Scholes equation satisfying*

$$f_T(S_T) = \max(S_T - K, 0).$$

(b) *The price of a European put option is the solution $F(t, S)$ of the Black-Scholes equation with the final condition*

$$f_T(S_T) = \max(K - S_T, 0).$$

(c) *The value of a forward contract is the solution the Black-Scholes equation with the final condition*

$$f_T(S_T) = S_T - K.$$

It is worth noting that the *superposition principle* discussed in section 13 can be explained now by the fact that the solution space of the Black-Scholes equation is a linear space. This means that a linear combination of solutions is also a solution.

Another interesting feature of the Black-Scholes equation is its independence of the stock drift rate μ. Then its solutions must have the same property. This explains why, in the risk-neutral valuation, the value of μ does not appear explicitly in the solution.

Asian options satisfy similar Black-Scholes equations, with small differences, as we shall see in the next section.

15.10 Boundary Conditions

We have solved the Black-Scholes equation for a call option, under the assumption that there is a unique solution. The Black-Scholes equation is of first order in the time

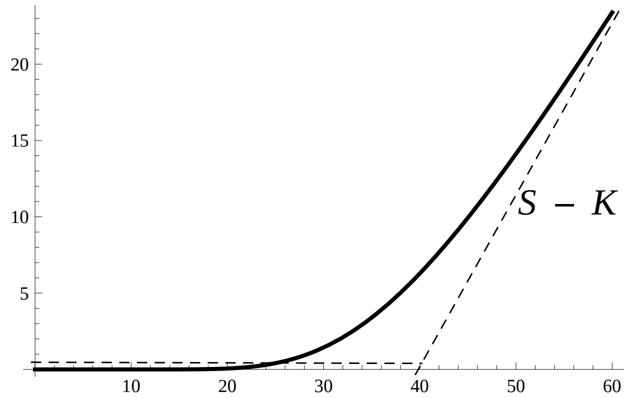


Figure 15.2: *NEEDS TO BE REDONE The graph of the option price before maturity in the case* $K = 40$, $\sigma = 30\%$, $r = 8\%$, *and* $T - t = 1$.

variable t and of second order in the stock variable S, so it needs one final condition at $t = T$ and two boundary conditions for $S = 0$ and $S \to \infty$.

In the case of a call option, the final condition is given by the following payoff:

$$F(T, S_T) = \max\{S_T - K, 0\}.$$

When $S \to 0$, the option does not get exercised, so the initial boundary condition is

$$F(t, 0) = 0.$$

When $S \to \infty$ the price becomes linear

$$F(t, S) \sim S - K,$$

the graph of $F(\,\cdot\,, S)$ having a slant asymptote, see Fig.15.2.

15.11 Risk-less Portfolios for Rare Events

Consider the derivative $P = P(t, S, F)$, which depends on the time t, stock price S and the derivative F, whose underlying asset is S. We shall find the stochastic differential equation followed by P, under the hypothesis that the stock exhibits rare events, i.e.

$$dS = \mu S dt + \sigma S dW_t + \rho S dM_t, \qquad (15.11.18)$$

where the constant μ, σ, ρ denote the drift rate, volatility and jump in the stock price in the case of a rare event. The processes W_t and $M_t = N_t - \lambda t$ denote the Brownian motion and the compensated Poisson process, respectively. The constant $\lambda > 0$ denotes the rate of occurrence of the rare events in the market.

By Ito's formula we get

$$dP = \frac{\partial P}{\partial t}dt + \frac{\partial P}{\partial S}dS + \frac{\partial P}{\partial F}dF + \frac{1}{2}\frac{\partial^2 P}{\partial S^2}dS^2 + \frac{1}{2}\frac{\partial^2 P}{\partial F^2}(dF)^2. \qquad (15.11.19)$$

We shall use the following stochastic relations:

$$(dW_t)^2 = dt, \quad (dM_t)^2 = dN_t, \quad dt^2 = dt\,dW_t = dt\,dM_t = dW_t\,dM_t = 0,$$

see sections 3.11.2 and 3.11.2. Then

$$\begin{aligned} (dS)^2 &= \sigma^2 S^2 dt + \rho^2 S^2 dN_t \\ &= (\sigma^2 + \lambda\rho^2)S^2 dt + \rho^2 S^2 dM_t, \end{aligned} \qquad (15.11.20)$$

where we used $dM_t = dN_t - \lambda dt$. It is worth noting that the unpredictable part of $(dS)^2$ depends only on the rare events, and does not depend on the regular daily events.

Exercise 15.11.1 *If S satisfies (15.11.18), find the following*

(a) $E[(dS)^2]$ $\qquad$ (b) $E[dS]$ $\qquad$ (c) $Var[dS]$.

Using Ito's formula, the infinitesimal change in the value of the derivative $F = F(t, S)$ is given by

$$\begin{aligned} dF &= \frac{\partial F}{\partial t}dt + \frac{\partial F}{\partial S}dS + \frac{1}{2}\frac{\partial^2 F}{\partial S^2}(dS)^2 \\ &= \Big(\frac{\partial F}{\partial t} + \mu S\frac{\partial F}{\partial S} + \frac{1}{2}(\sigma^2 + \lambda\rho^2)S^2\frac{\partial^2 F}{\partial S^2}\Big)dt \\ &\quad + \sigma S\frac{\partial F}{\partial S}dW_t \\ &\quad + \Big(\frac{1}{2}\rho^2 S^2\frac{\partial^2 F}{\partial S^2} + \rho S\frac{\partial F}{\partial S}\Big)dM_t, \end{aligned} \qquad (15.11.21)$$

where we have used (15.11.18) and (15.11.20). The increment dF has two independent sources of uncertainty: dW_t and dM_t, both with mean equal to 0.

Taking the square, yields

$$(dF)^2 = \sigma^2 S^2\Big(\frac{\partial F}{\partial S}\Big)^2 dt + \Big(\frac{1}{2}\rho^2 S^2\frac{\partial^2 F}{\partial S^2} + \rho S\frac{\partial F}{\partial S}\Big)dN_t. \qquad (15.11.22)$$

Substituting back in (15.11.19), we obtain the unpredictable part of dP as the sum of two components

$$\begin{aligned} &\sigma S\Big(\frac{\partial P}{\partial S} + \frac{\partial P}{\partial F}\frac{\partial F}{\partial S}\Big)dW_t \\ &+\rho S\Big(\frac{\partial P}{\partial S} + \frac{\partial P}{\partial F}\frac{\partial F}{\partial S}\Big) + \frac{1}{2}\rho^2 S^2\Big(\frac{\partial P}{\partial F}\frac{\partial^2 F}{\partial S^2} + \frac{\partial^2 P}{\partial S^2} + \frac{\partial^2 P}{\partial F^2}\frac{\partial^2 F}{\partial S^2}\Big)dM_t. \end{aligned}$$

The risk-less condition for the portfolio P is obtained when the coefficients of dW_t and dM_t vanish

$$\begin{aligned} \frac{\partial P}{\partial S} + \frac{\partial P}{\partial F}\frac{\partial F}{\partial S} &= 0 && (15.11.23) \\ \frac{\partial P}{\partial F}\frac{\partial^2 F}{\partial S^2} + \frac{\partial^2 P}{\partial S^2} + \frac{\partial^2 P}{\partial F^2}\frac{\partial^2 F}{\partial S^2} &= 0. && (15.11.24) \end{aligned}$$

These relations can be further simplified. If we differentiate in (15.11.23) with respect to S

$$\frac{\partial^2 P}{\partial S^2} + \frac{\partial P}{\partial F}\frac{\partial^2 F}{\partial S^2} = -\frac{\partial^2 P}{\partial S \partial F}\frac{\partial F}{\partial S}.$$

Substituting into (15.11.24) yields

$$\frac{\partial^2 P}{\partial F^2}\frac{\partial^2 F}{\partial S^2} = \frac{\partial^2 P}{\partial S \partial F}\frac{\partial F}{\partial S}. \qquad (15.11.25)$$

Differentiating in (15.11.23) with respect to F we get

$$\begin{aligned} \frac{\partial^2 P}{\partial F \partial S} &= -\frac{\partial^2 P}{\partial F^2}\frac{\partial F}{\partial S} - \frac{\partial P}{\partial F}\frac{\partial^2 F}{\partial F \partial S} \\ &= -\frac{\partial^2 P}{\partial F^2}\frac{\partial F}{\partial S}, && (15.11.26) \end{aligned}$$

since

$$\frac{\partial^2 F}{\partial F \partial S} = \frac{\partial}{\partial S}\Big(\frac{\partial F}{\partial F}\Big).$$

Multiplying (15.11.26) by $\dfrac{\partial F}{\partial S}$ yields

$$\frac{\partial^2 P}{\partial F \partial S}\frac{\partial F}{\partial S} = -\frac{\partial^2 P}{\partial F^2}\Big(\frac{\partial F}{\partial S}\Big)^2,$$

and substituting in the right side of (15.11.25) leads to the equation

$$\frac{\partial^2 P}{\partial F^2}\Big[\frac{\partial^2 F}{\partial S^2} + \Big(\frac{\partial F}{\partial S}\Big)^2\Big] = 0.$$

We have arrived at the following result:

Proposition 15.11.2 *Let $F = F(t, S)$ be a derivative with the underlying asset S. The investment $P = P(t, S, F)$ is risk-less if and only if*

$$\begin{aligned} \frac{\partial P}{\partial S} + \frac{\partial P}{\partial F}\frac{\partial F}{\partial S} &= 0 \\ \frac{\partial^2 P}{\partial F^2}\Big[\frac{\partial^2 F}{\partial S^2} + \Big(\frac{\partial F}{\partial S}\Big)^2\Big] &= 0. \end{aligned}$$

There are two risk-less conditions because there are two unpredictable components in the increments of dP, one due to regular changes and the other due to rare events. The first condition is equivalent with the vanishing total derivative, $\dfrac{dP}{dS} = 0$, and corresponds to offsetting the regular risk.

The second condition vanishes either if $\dfrac{\partial^2 P}{\partial F^2} = 0$ or if $\dfrac{\partial^2 F}{\partial S^2} + \left(\dfrac{\partial F}{\partial S}\right)^2 = 0$. In the first case P is linear in F. For instance, if $P = F - f(S)$, from the first condition yields $f'(S) = \dfrac{\partial F}{\partial S}$. In the second case, denote $U(t,S) = \frac{\partial F}{\partial S}$. Then we need to solve the partial differential equation

$$\frac{\partial U}{\partial t} + U^2 = 0.$$

Future research directions:

1. Solve the above equation.
2. Find the predictable part of dP.
3. Get an analog of the Black-Scholes in this case.
4. Evaluate a call option in this case.
5. Is the risk-neutral valuation still working and why?

Chapter 16

Black-Scholes for Asian Derivatives

In this chapter we shall develop the Black-Scholes equation in the case of Asian derivatives and we shall discuss the particular cases of options and forward contracts on weighted averages. In the case of the latter contracts we obtain closed form solutions, while for the former ones we apply the reduction variable method to decrease the number of variables and discuss the solution.

16.1 Weighted averages

In many practical problems the asset price needs to be considered with a certain weight. For instance, when computing car insurance, more weight is assumed for recent accidents than for accidents that occurred 10 years ago.

In the following we shall define the weight function and provide several examples.

Let $\rho : [0, T] \to \mathbb{R}$ be a *weight function*, i.e. a function satisfying

1. $\rho > 0$;
2. $\int_0^T \rho(t)\,dt = 1$.

The *stock weighted average* with respect to the weight ρ is defined as

$$\boxed{S_{ave} = \int_0^T \rho(t) S_t\,dt.}$$

Example 16.1.1 (a) *The uniform weight is obtained for* $\rho(t) = \dfrac{1}{T}$. *In this case*

$$\boxed{S_{ave} = \frac{1}{T}\int_0^T S_t\,dt}$$

is the continuous arithmetic average of the stock on the time interval $[0, T]$.

(*b*) *The linear weight is obtained if* $\rho(t) = \dfrac{2t}{T^2}$. *In this case the weight is the time*

$$\boxed{S_{ave} = \frac{2}{T^2}\int_0^T tS_t\,dt.}$$

(*c*) *The exponential weight is obtained for* $\rho(t) = \dfrac{ke^{kt}}{e^{kT}-1}$. *If* $k > 0$, *the weight is increasing, so recent data are weighted more than old data; if* $k < 0$, *the weight is decreasing. The exponential weighted average is given by*

$$\boxed{S_{ave} = \frac{k}{e^{kT}-1}\int_0^T e^{kt}S_t\,dt.}$$

Exercise 16.1.2 *Consider the polynomial weighted average*

$$S_{ave}^{(n)} = \frac{n+1}{T^{n+1}}\int_0^T t^n S_t\,dt.$$

Find the limit $\lim\limits_{n\to\infty} S_{ave}^{(n)}$ *in the cases* $0 < T < 1$, $T = 1$, *and* $T > 1$.

In all previous examples $\rho(t) = \rho(t,T) = \dfrac{f(t)}{g(T)}$, with $\int_0^T f(t)\,dt = g(T)$, so $g'(T) = f(T)$ and $g(0) = 0$. The average becomes

$$S_{ave}(T) = \frac{1}{g(T)}\int_0^T f(u)S_u\,du = \frac{I_T}{g(T)},$$

with $I_t = \int_0^t f(u)S_u\,du$ satisfying $dI_t = f(t)S_t dt$. From the product rule we get

$$\begin{aligned}
dS_{ave}(t) &= \frac{dI_t\, g(t) - I_t dg(t)}{g(t)^2} = \left(\frac{f(t)}{g(t)}S_t - \frac{g'(t)}{g(t)}\frac{I_t}{g(t)}\right)dt \\
&= \frac{f(t)}{g(t)}\Big(S_t - \frac{g'(t)}{f(t)}S_{ave}(t)\Big)dt \\
&= \frac{f(t)}{g(t)}\Big(S_t - S_{ave}(t)\Big)dt,
\end{aligned}$$

since $g'(t) = f(t)$. The initial condition is

$$S_{ave}(0) = \lim_{t\searrow 0} S_{ave}(t) = \lim_{t\searrow 0}\frac{I_t}{g(t)} = \lim_{t\searrow 0}\frac{f(t)S_t}{g'(t)} = S_0\lim_{t\searrow 0}\frac{f(t)}{g'(t)} = S_0,$$

Proposition 16.1.3 *The weighted average $S_{ave}(t)$ satisfies the stochastic differential equation*

$$\begin{aligned} dX_t &= \frac{f(t)}{g(t)}(S_t - X_t)dt \\ X_0 &= S_0. \end{aligned}$$

Exercise 16.1.4 *Let $x(t) = E[S_{ave}(t)]$.*
(a) Show that $x(t)$ satisfies the ordinary differential equation

$$\begin{aligned} x'(t) &= \frac{f(t)}{g(t)}\big(S_0 e^{\mu t} - x(t)\big) \\ x(0) &= S_0. \end{aligned}$$

(b) Find $x(t)$.

Exercise 16.1.5 *Let $y(t) = E[S_{ave}^2(t)]$.*
(a) Find the stochastic differential equation satisfied by $S_{ave}^2(t)$;
(b) Find the ordinary differential equation satisfied by $y(t)$;
(c) Solve the previous equation to get $y(t)$ and compute $Var[S_{ave}]$.

16.2 Setting up the Black-Scholes Equation

Consider an Asian derivative whose value at time t, $F(t, S_t, S_{ave}(t))$, depends on variables t, S_t, and $S_{ave}(t)$. Using the stochastic process of S_t

$$dS_t = \mu S_t dt + \sigma S_t dW_t$$

and Proposition 16.1.3, an application of Ito's formula together with the stochastic formulas

$$dt^2 = 0, \quad (dW_t)^2 = 0, \quad (dS_t)^2 = \sigma^2 S^2 dt, \quad (dS_{ave})^2 = 0$$

yields

$$\begin{aligned} dF &= \frac{\partial F}{\partial t}dt + \frac{\partial F}{\partial S_t}dS_t + \frac{1}{2}\frac{\partial^2 F}{\partial S_t^2}(dS_t)^2 + \frac{\partial F}{\partial S_{ave}}dS_{ave} \\ &= \Big(\frac{\partial F}{\partial t} + \mu S_t \frac{\partial F}{\partial S_t} + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 F}{\partial S_t^2} + \frac{f(t)}{g(t)}(S_t - S_{ave})\frac{\partial F}{\partial S_{ave}}\Big)dt \\ &\quad + \sigma S_t \frac{\partial F}{\partial S_t} dW_t. \end{aligned}$$

Let $\Delta_F = \frac{\partial F}{\partial S_t}$. Consider the following portfolio at time t

$$P(t) = F - \Delta_F S_t,$$

obtained by buying one derivative F and selling Δ_F units of stock. The change in the portfolio value during the time dt does not depend on W_t

$$\begin{aligned} dP &= dF - \Delta_F dS_t \\ &= \Big(\frac{\partial F}{\partial t} + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 F}{\partial S_t^2} + \frac{f(t)}{g(t)}(S_t - S_{ave})\frac{\partial F}{\partial S_{ave}}\Big)dt \end{aligned} \tag{16.2.1}$$

so the portfolio P is risk-less. Since no arbitrage opportunities are allowed, investing a value P at time t in a bank at the risk-free rate r for the time interval dt yields

$$dP = rPdt = \Big(rF - rS_t \frac{\partial F}{\partial S_t}\Big)dt. \tag{16.2.2}$$

Equating (16.2.1) and (16.2.2) yields the following form of the Black-Scholes equation for Asian derivatives on weighted averages

$$\boxed{\frac{\partial F}{\partial t} + rS_t \frac{\partial F}{\partial S_t} + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 F}{\partial S_t^2} + \frac{f(t)}{g(t)}(S_t - S_{ave})\frac{\partial F}{\partial S_{ave}} = rF.}$$

16.3 Weighted Average Strike Call Option

In this section we shall use the reduction variable method to decrease the number of variables from three to two. Since $S_{ave}(t) = \dfrac{I_t}{g(t)}$, it is convenient to consider the derivative as a function of t, S_t and I_t

$$V(t, S_t, I_t) = F(t, S_t, S_{ave}).$$

A computation similar to the previous one yields the simpler equation

$$\boxed{\frac{\partial V}{\partial t} + rS_t \frac{\partial V}{\partial S_t} + \frac{1}{2}\sigma^2 S_t^2 \frac{\partial^2 V}{\partial S_t^2} + f(t)S_t \frac{\partial V}{\partial I_t} = rV.} \tag{16.3.3}$$

The payoff at maturity of an average strike call option can be written in the following form

$$\begin{aligned} V_T &= V(T, S_T, I_T) = \max\{S_T - S_{ave}(T), 0\} \\ &= \max\{S_T - \frac{I_T}{g(T)}, 0\} = S_T \max\{1 - \frac{1}{g(T)}\frac{I_T}{S_T}, 0\} \\ &= S_T L(T, R_T), \end{aligned}$$

where

$$R_t = \frac{I_t}{S_t}, \qquad L(t, R) = \max\{1 - \frac{1}{g(t)}R, 0\}.$$

Since at maturity the variable S_T is separated from T and R_T, we shall look for a solution of equation (16.3.3) of the same type for any $t \le T$, i.e. $V(t,S,I) = SG(t,R)$. Since

$$\begin{aligned}
\frac{\partial V}{\partial t} &= S\frac{\partial G}{\partial t}, \qquad \frac{\partial V}{\partial I} = S\frac{\partial G}{\partial R}\frac{1}{S} = \frac{\partial G}{\partial R};\\
\frac{\partial V}{\partial S} &= G + S\frac{\partial G}{\partial R}\frac{\partial R}{\partial S} = G - R\frac{\partial G}{\partial R};\\
\frac{\partial^2 V}{\partial S^2} &= \frac{\partial}{\partial S}(G - R\frac{\partial G}{\partial R}) = \frac{\partial G}{\partial R}\frac{\partial R}{\partial S} - \frac{\partial R}{\partial S}\frac{\partial G}{\partial R} - R\frac{\partial^2 G}{\partial R^2}\frac{\partial R}{\partial S};\\
&= R\frac{\partial^2 G}{\partial R^2}\frac{I}{S};\\
S^2\frac{\partial^2 V}{\partial S^2} &= RI\frac{\partial^2 G}{\partial R^2},
\end{aligned}$$

substituting in (16.3.3) and using that $\dfrac{RI}{S} = R^2$, after cancelations yields

$$\boxed{\frac{\partial G}{\partial t} + \frac{1}{2}\sigma^2 R^2 \frac{\partial^2 G}{\partial R^2} + (f(t) - rR)\frac{\partial G}{\partial R} = 0.} \tag{16.3.4}$$

This is a partial differential equation in only two variables, t and R. It can be solved explicitly sometimes, depending on the form of the final condition $G(T, R_T)$ and expression of the function $f(t)$.

In the case of a weighted average strike call option the final condition is

$$\boxed{G(T, R_T) = \max\{1 - \frac{R_T}{g(T)}, 0\}.} \tag{16.3.5}$$

Example 16.3.1 *In the case of the arithmetic average the function $G(t,R)$ satisfies the partial differential equation*

$$\frac{\partial G}{\partial t} + \frac{1}{2}\sigma^2 R^2 \frac{\partial^2 G}{\partial R^2} + (1 - rR)\frac{\partial G}{\partial R} = 0$$

with the final condition $G(T, R_T) = \max\{1 - \dfrac{R_T}{T}, 0\}$.

Example 16.3.2 *In the case of the exponential average the function $G(t,R)$ satisfies the equation*

$$\frac{\partial G}{\partial t} + \frac{1}{2}\sigma^2 R^2 \frac{\partial^2 G}{\partial R^2} + (ke^{kt} - rR)\frac{\partial G}{\partial R} = 0 \tag{16.3.6}$$

with the final condition $G(T, R_T) = \max\{1 - \dfrac{R_T}{e^{kT} - 1}, 0\}$.

Neither of the previous two final condition problems can be solved explicitly.

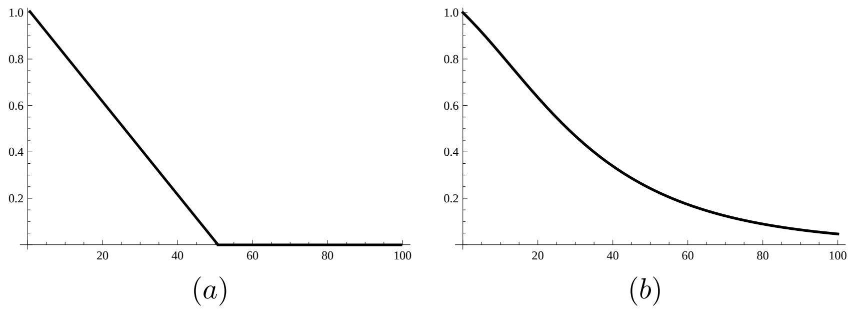


Figure 16.1: *The profile of the solution $H(t, R)$: (a) at expiration; (b) when there is $T - t$ time left before expiration.*

16.4 Boundary Conditions

The partial differential equation (16.3.4) is of first order in t and second order in R. We need to specify one condition at $t = T$ (the payoff at maturity), which is given by (16.3.5), and two conditions for $R = 0$ and $R \to \infty$, which specify the behavior of solution $G(t, R)$ at two limiting positions of the variable R.

Taking $R \to 0$ in equation (16.3.4) and using Exercise 16.4.1 yields the first boundary condition for $G(t, R)$

$$\boxed{\Big(\frac{\partial G}{\partial t} + f\,\frac{\partial G}{\partial R}\Big)\Big|_{R=0} = 0.} \tag{16.4.7}$$

The term $\frac{\partial G}{\partial R}\big|_{R=0}$ represents the slope of $G(t, R)$ with respect to R at $R = 0$, while $\frac{\partial G}{\partial t}\big|_{R=0}$ is the variation of the price G with respect to time t when $R = 0$.

Another boundary condition is obtained by specifying the behavior of $G(t, R)$ for large values of R. If $R_t \to \infty$, we must have $S_t \to 0$, because

$$R_t = \frac{1}{S_t}\int_0^t f(u)S_u\,du$$

and $\int_0^t f(u)S_u\,du > 0$ for $t > 0$. In this case we are better off not exercising the option (since otherwise we get a negative payoff), so the boundary condition is

$$\boxed{\lim_{R\to\infty} G(R, t) = 0.} \tag{16.4.8}$$

It can be shown in the theory of partial differential equations that equation (16.3.4) together with the final condition (16.3.5), see Fig.16.1(a), and boundary conditions (16.4.7) and (16.4.8) has a unique solution $G(t, R)$, see Fig.16.1(b).

Exercise 16.4.1 *Let f be a bounded differentiable function. Show that*

(*a*) $\lim_{x\to 0} xf'(x) = 0$;

(*b*) $\lim_{x\to 0} x^2 f''(x) = 0$.

There is no close form solution for the weighted average strike call option. Even in the simplest case, when the average is arithmetic, the solution is just approximative, see section 13.13. In real life the price is worked out using the *Monte-Carlo simulation*. This is based on averaging a large number, n, of simulations of the process R_t in the risk-neutral world, i.e. assuming $\mu = r$. For each realization, the associated payoff $G_{T,j} = \max\{1 - \frac{R_{T,j}}{g(T)}\}$ is computed, with $j \le n$. Here $R_{T,j}$ represents the value of R at time T in the jth realization. The average

$$\frac{1}{n}\sum_{j=1}^{n} G_{T,j}$$

is a good approximation of the payoff expectation $E[G_T]$. Discounting under the risk-free rate we get the price at time t

$$G(t,R) = e^{-r(T-t)}\Big(\frac{1}{n}\sum_{j=1}^{n} G_{T,j}\Big).$$

It is worth noting that the term on the right is an approximation of the risk neutral conditional expectation $\widehat{E}[G_T|\mathcal{F}_t]$.

When simulating the process R_t, it is convenient to know its stochastic differential equation. Using

$$dI_t = f(t)S_t dt, \qquad d\Big(\frac{1}{S_t}\Big) = \frac{1}{S_t}\Big((\sigma^2-\mu)dt - \sigma dW_t\Big)dt,$$

the product rule yields

$$\begin{aligned} dR_t &= d\Big(\frac{I_t}{S_t}\Big) = d\Big(I_t\frac{1}{S_t}\Big) \\ &= dI_t\,\frac{1}{S_t} + I_t d\Big(\frac{1}{S_t}\Big) + dI_t d\Big(\frac{1}{S_t}\Big) \\ &= f(t)dt + R_t\Big((\sigma^2-\mu)dt - \sigma dW_t\Big). \end{aligned}$$

Collecting terms yields the following stochastic differential equation for R_t:

$$\boxed{dR_t = -\sigma R_t dW_t + \big(f(t) + (\sigma^2-\mu)R_t\big)dt.} \qquad (16.4.9)$$

The initial condition is $R_0 = \frac{I_0}{S_0} = 0$, since $I_0 = 0$.

Can we solve explicitly this equation? Can we find the mean and variance of R_t?

We shall start by finding the mean $E[R_t]$. The equation can be written as

$$dR_t - (\sigma^2 - \mu)R_t dt = f(t)dt - \sigma R_t dW_t.$$

Multiplying by $e^{-(\sigma^2-\mu)t}$ yields the exact equation

$$d\Big(e^{-(\sigma^2-\mu)t}R_t\Big) = e^{-(\sigma^2-\mu)t}f(t)dt - \sigma e^{-(\sigma^2-\mu)t}R_t dW_t.$$

Integrating yields

$$e^{-(\sigma^2-\mu)t}R_t = \int_0^t e^{-(\sigma^2-\mu)u}f(u)\,du - \int_0^t \sigma e^{-(\sigma^2-\mu)u}R_u\,dW_u.$$

The first integral is deterministic while the second is an Ito integral. Using that the expectations of Ito integrals vanish, we get

$$E[e^{-(\sigma^2-\mu)t}R_t] = \int_0^t e^{-(\sigma^2-\mu)u}f(u)\,du$$

and hence

$$\boxed{E[R_t] = e^{(\sigma^2-\mu)t}\int_0^t e^{-(\sigma^2-\mu)u}f(u)\,du.}$$

Exercise 16.4.2 *Find $E[R_t^2]$ and $Var[R_t]$.*

Equation (16.4.9) is a linear equation of the type discussed in section 7.8. Multiplying by the integrating factor

$$\rho_t = e^{\sigma W_t + \frac{1}{2}\sigma^2 t}$$

the equation is transformed into an exact equation

$$d(\rho_t R_t) = \big(\rho_t f(t) + (\sigma^2 - \mu)\rho_t R_t\big)dt.$$

Substituting $Y_t = \rho_t R_t$ yields

$$dY_t = \big(\rho_t f(t) + (\sigma^2 - \mu)Y_t\big)dt,$$

which can be written as

$$dY_t - (\sigma^2 - \mu)Y_t dt = \rho_t f(t)dt.$$

Multiplying by $e^{-(\sigma^2-\mu)t}$ yields the exact equation

$$d(e^{-(\sigma^2-\mu)t}Y_t) = e^{-(\sigma^2-\mu)t}\rho_t f(t)dt,$$

which can be solved by integration

$$e^{-(\sigma^2-\mu)t}Y_t = \int_0^t e^{-(\sigma^2-\mu)u}\rho_u f(u)\,du.$$

Going back to the variable $R_t = Y_t/\rho_t$, we obtain the following closed form expression

$$\boxed{R_t = \int_0^t e^{(\mu-\frac{1}{2}\sigma^2)(u-t)+\sigma(W_u-W_t)} f(u)\,du.} \tag{16.4.10}$$

Exercise 16.4.3 *Find $E[R_t]$ by taking the expectation in formula* (16.4.10).

It is worth noting that we can arrive at formula (16.4.10) directly, without going through solving a stochastic differential equation. We shall show this procedure in the following.

Using the well-known formulas for the stock price

$$S_u = S_0 e^{(\mu-\frac{1}{2}\sigma^2)u+\sigma W_u}, \qquad S_t = S_0 e^{(\mu-\frac{1}{2}\sigma^2)t+\sigma W_t},$$

and dividing, yields

$$\frac{S_u}{S_t} = e^{(\mu-\frac{1}{2}\sigma^2)(u-t)+\sigma(W_u-W_t)}.$$

Then we get

$$\begin{aligned} R_t &= \frac{I_t}{S_t} = \frac{1}{S_t}\int_0^t S_u f(u)\,du \\ &= \int_0^t \frac{S_u}{S_t} f(u)\,du = \int_0^t e^{(\mu-\frac{1}{2}\sigma^2)(u-t)+\sigma(W_u-W_t)} f(u)\,du, \end{aligned}$$

which is formula (16.4.10).

Exercise 16.4.4 *Find an explicit formula for R_t in terms of the integrated Brownian motion $Z_t^{(\sigma)} = \int_0^t e^{\sigma W_u}\,du$, in the case of an exponential weight with $k = \frac{1}{2}\sigma^2 - \mu$, see Example* 16.1.1(*c*).

Exercise 16.4.5 (*a*) *Find the price of a derivative G which satisfies*

$$\frac{\partial G}{\partial t} + \frac{1}{2}\sigma^2 R^2 \frac{\partial^2 G}{\partial R^2} + (1 - rR)\frac{\partial G}{\partial R} = 0$$

with the payoff $G(T, R_T) = R_T^2$.

(*b*) *Find the value of an Asian derivative V_t on the arithmetic average, that has the payoff*

$$V_T = V(T, S_T, I_T) = \frac{I_T^2}{S_T},$$

where $I_T = \int_0^T S_t\,dt$.

Exercise 16.4.6 *Use a computer simulation to find the value of an Asian arithmetic average strike option with $r = 4\%$, $\sigma = 50\%$, $S_0 = \$40$, and $T = 0.5$ years.*

16.5 Asian Forward Contracts on Weighted Averages

Since the payoff of this derivative is given by

$$V_T = S_T - S_{ave}(T) = S_T\Big(1 - \frac{R_T}{g(T)}\Big),$$

the reduction variable method suggests considering a solution of the type $V(t, S_t, I_t) = S_t G(t, R_t)$, where $G(t, T)$ satisfies equation (16.3.4) with the final condition $G(T, R_T) = 1 - \dfrac{R_T}{g(T)}$. Since this is linear in R_T, this implies looking for a solution $G(t, R_t)$ in the following form

$$G(t, R_t) = a(t)R_t + b(t), \tag{16.5.11}$$

with functions $a(t)$ and $b(t)$ subject to be determined. Substituting into (16.3.4) and collecting R_t yields

$$(a'(t) - ra(t))R_t + b'(t) + f(t)a(t) = 0.$$

Since this polynomial in R_t vanishes for all values of R_t, then its coefficients are identically zero, so

$$a'(t) - ra(t) = 0, \qquad b'(t) + f(t)a(t) = 0.$$

When $t = T$ we have

$$G(T, R_T) = a(T)R_T + b(T) = 1 - \frac{R_T}{g(T)}.$$

Equating the coefficients of R_T yields the final conditions

$$a(T) = -\frac{1}{g(T)}, \qquad b(T) = 1.$$

The coefficient $a(t)$ satisfies the ordinary differential equation

$$\begin{aligned} a'(t) &= ra(t) \\ a(T) &= -\frac{1}{g(T)} \end{aligned}$$

which has the solution

$$a(t) = -\frac{1}{g(T)}e^{-r(T-t)}.$$

The coefficient $b(t)$ satisfies the equation

$$\begin{aligned} b'(t) &= -f(t)a(t) \\ b(T) &= 1 \end{aligned}$$

with the solution

$$b(t) = 1 + \int_t^T f(u)a(u)\,du.$$

Substituting in (16.5.11) yields

$$\begin{aligned} G(t,R) &= -\frac{1}{g(T)}e^{-r(T-t)}R_t + 1 + \int_t^T f(u)a(u)\,du \\ &= 1 - \frac{1}{g(T)}\big[R_t e^{-r(T-t)} + \int_t^T f(u)e^{-r(T-u)}\,du\big]. \end{aligned}$$

Then going back into the variable $I_t = S_t R_t$ yields

$$\begin{aligned} V(t,S_t,I_t) &= S_t G(t,R_t) \\ &= S_t - \frac{1}{g(T)}\big[I_t e^{-r(T-t)} + S_t \int_t^T f(u)e^{-r(T-u)}\,du\big]. \end{aligned}$$

Using that $\rho(u) = \dfrac{f(u)}{g(T)}$ and going back to the initial variable $S_{ave}(t) = I_t/g(t)$ yields

$$\begin{aligned} F(t,S_t,S_{ave}(t)) &= V(t,S_t,I_t) \\ &= S_t - \frac{g(t)}{g(T)}S_{ave}(t)e^{-r(T-t)} - S_t\int_t^T \rho(u)e^{-r(T-u)}\,du. \end{aligned}$$

We have arrived at the following result:

Proposition 16.5.1 *The value at time t of an Asian forward contract on a weighted average with the weight function $\rho(t)$, i.e. an Asian derivative with the payoff $F_T = S_T - S_{ave}(T)$, is given by*

$$\boxed{F(t,S_t,S_{ave}(t)) = S_t\Big(1 - \int_t^T \rho(u)e^{-r(T-u)}\,du\Big) - \frac{g(t)}{g(T)}e^{-r(T-t)}S_{ave}(t).}$$

It is worth noting that the previous price can be written as a linear combination of S_t and $S_{ave}(t)$

$$F(t,S_t,S_{ave}(t)) = \alpha(t)S_t + \beta(t)S_{ave}(t),$$

where

$$\begin{aligned} \alpha(t) &= 1 - \int_t^T \rho(u)e^{-r(T-u)}\,du \\ \beta(t) &= -\frac{g(t)}{g(T)}e^{-r(T-t)} = -\frac{\int_0^t f(u)\,du}{\int_0^T f(u)\,du}e^{-r(T-t)}. \end{aligned}$$

In the first formula $\rho(u)e^{-r(T-u)}$ is the discounted weight at time u, and $\alpha(t)$ is 1 minus the total discounted weight between t and T. One can easily check that $\alpha(T) = 1$ and $\beta(T) = -1$.

Exercise 16.5.2 *Find the value at time t of an Asian forward contract on an arithmetic average $A_t = \int_0^t S_u\, du$.*

Exercise 16.5.3 (*a*) *Find the value at time t of an Asian forward contract on an exponential weighted average with the weight given by Example* 16.1.1 (*c*).

(*b*) *What happens if $k = -r$? Why?*

Exercise 16.5.4 *Find the value at time t of an Asian power contract with the payoff $F_T = \Big(\int_0^T S_u\, du \Big)^n$.*

Chapter 17

American Options

American options are options that are allowed to be execised at any time before maturity. Because of this advantage, they tend to be more expensive than the European counterparts. Exact pricing formulas exist just for perpetuities, while they are missing for finitely lived American options.

17.1 Perpetual American Options

A *perpetual American option* is an American option that never expires. These contracts can be exercised at any time t, $0 \leq t \leq \infty$. Even if finding the optimal exercise time for finite maturity American options is a delicate matter, in the case of perpetual American calls and puts there is always possible to find the optimal exercise time and to derive a close form pricing formula (see Merton, [12]).

17.1.1 Present Value of Barriers

Our goal in this section will be to learn how to compute the present value of a contract that pays a fixed cash amount at a stochastic time defined by the first passage of time of a stock. These formulas will be the main ingredient in pricing perpetual American options over the next couple of sections.

Reaching the barrier from below Let S_t denote the stock price with initial value S_0 and consider a positive number b such that $b > S_0$. We recall the first passage of time τ_b when the stock S_t hits for the first time the barrier b, see Fig. 17.1

$$\tau_b = \inf\{t > 0; S_t = b\}.$$

Consider a contract that pays \$1 at the time when the stock reaches the barrier b for the first time. Under the constant interest rate assumption, the value of the contract at time $t = 0$ is obtained by discounting the value of \$1 at the rate r for the period τ_b and taking the expectation in the risk neutral world

$$f_0 = \widehat{E}[e^{-r\tau_b}].$$

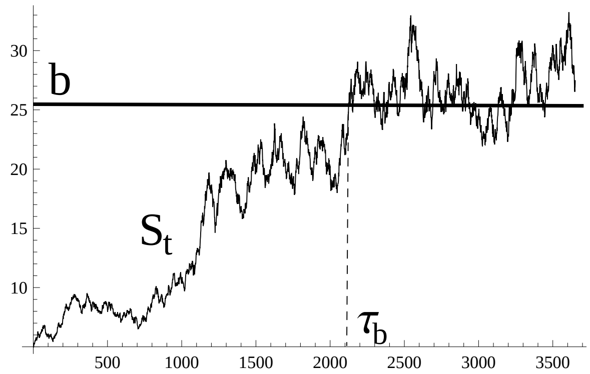


Figure 17.1: *The first passage of time when S_t hits the level b from below.*

In the following we shall compute the right side of the previous expression using two different approaches. Using that the stock price in the risk-neutral world is given by the expression $S_t = S_0 e^{(r-\frac{\sigma^2}{2})t+\sigma W_t}$, with $r > \frac{\sigma^2}{2}$, then

$$M_t = e^{-rt} S_t = S_0 e^{\sigma W_t - \frac{\sigma^2}{2}t}, \qquad t \geq 0$$

is a martingale. Applying the Optional Stopping Theorem (Theorem 3.2.1) yields $E[M_{\tau_b}] = E[M_0]$, which is equivalent to

$$E[e^{-r\tau_b} S_{\tau_b}] = S_0.$$

Since $S_{\tau_b} = b$, the previous relation implies

$$E[e^{-r\tau_b}] = \frac{S_0}{b},$$

where the expectation is taken in the risk-neutral world. Hence, we arrived at the following result:

Proposition 17.1.1 *The value at time $t = 0$ of \$1 received at the time when the stock reaches level b from below is*

$$\boxed{f_0 = \frac{S_0}{b}.}$$

Exercise 17.1.2 *In the previous proof we had applied the Optional Stopping Theorem (Theorem 3.2.1). Show that the hypothesis of the theorem are satisfied.*

Exercise 17.1.3 *Let $0 < S_0 < b$ and assume $r > \frac{\sigma^2}{2}$.*
(a) Show that $P(S_t \text{ reaches } b) = 1$. Compare with Exercise 3.5.2.
(b) Prove the identity $P(\omega; \tau_b(\omega) < \infty) = 1$.

The result of Proposition 17.1.1 can be also obtained directly as a consequence of Proposition 3.6.4. Using the expression of the stock price in the risk-neutral world, $S_t = S_0 e^{(r-\frac{\sigma^2}{2})t+\sigma W_t}$, we have

$$\begin{aligned}\tau_b &= \inf\{t>0; S_t = b\} = \inf\{t>0; (r-\frac{\sigma^2}{2})t + \sigma W_t = \ln\frac{b}{S_0}\}\\ &= \inf\{t>0; \mu t + \sigma W_t = x\},\end{aligned}$$

where $x = \ln\frac{b}{S_0} > 0$ and $\mu = r - \frac{\sigma^2}{2} > 0$. Choosing $s = r$ in Proposition 3.6.4 yields

$$\begin{aligned}E[e^{-r\tau_b}] &= e^{\frac{1}{\sigma^2}(\mu-\sqrt{2s\sigma^2+\mu^2})x} = e^{\frac{1}{\sigma^2}\left(r-\frac{\sigma^2}{2}-\sqrt{2r\sigma^2+(r-\frac{\sigma^2}{2})^2}\right)\ln\frac{b}{S_0}}\\ &= e^{-\ln\frac{b}{S_0}} = e^{\ln\frac{S_0}{b}} = \frac{S_0}{b},\end{aligned}$$

where we used that

$$r-\frac{\sigma^2}{2}-\sqrt{2r\sigma^2+\Big(r-\frac{\sigma^2}{2}\Big)^2} = r-\frac{\sigma^2}{2}-\sqrt{\Big(r+\frac{\sigma^2}{2}\Big)^2} = r-\frac{\sigma^2}{2}-r-\frac{\sigma^2}{2} = -\sigma^2.$$

Exercise 17.1.4 *Let $S_0 < b$. Find the probability density function of the hitting time τ_b.*

Exercise 17.1.5 *Assume the stock pays continuous dividends at the constant rate $\delta > 0$ and let $b > 0$ such that $S_0 < b$.*

(a) Prove that the value at time $t = 0$ of \$1 received at the time when the stock reaches level b from below is

$$f_0 = \Big(\frac{S_0}{b}\Big)^{h_1},$$

where

$$h_1 = \frac{1}{2} - \frac{r-\delta}{\sigma^2} + \sqrt{\Big(\frac{r-\delta}{\sigma^2}-\frac{1}{2}\Big)^2 + \frac{2r}{\sigma^2}}.$$

(b) Show that $h_1(0) = 1$ and $h_1(\delta)$ is an increasing function for $\delta > 0$.

(c) Find the limit of the rate of change $\lim_{\delta\to\infty} h_1'(\delta)$.

(d) Work out a formula for the sensitivity of the value f_0 with respect to the dividend rate δ and compute the long run value of this rate.

Reaching the barrier from above Sometimes a stock can reach a barrier b from above. Let S_0 be the initial value of the stock S_t and assume the inequality $b < S_0$. Consider again the first passage of time $\tau_b = \inf\{t > 0; S_t = b\}$, see Fig. 17.2.

In this paragraph we compute the value of a contract that pays \$1 at the time when the stock reaches the barrier b for the first time, which is given by $f_0 = \widehat{E}[e^{-r\tau_b}]$. We shall keep the assumption that the interest rate r is constant and $r > \frac{\sigma^2}{2}$.

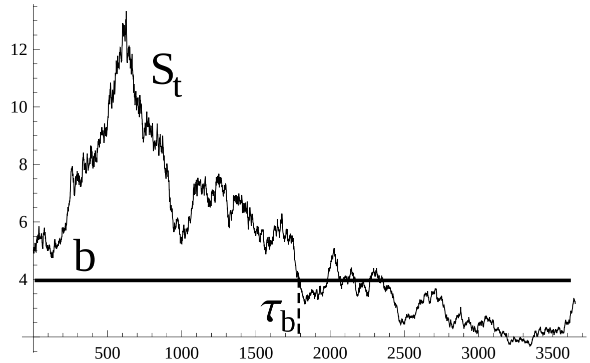


Figure 17.2: *The first passage of time when S_t hits the level b from above.*

Proposition 17.1.6 *The value at time $t = 0$ of \$1 received at the time when the stock reaches level b from above is*

$$\boxed{f_0 = \Big(\frac{S_0}{b}\Big)^{\frac{-2r}{\sigma^2}}.}$$

Proof: The reader might be tempted to use the Optional Stopping Theorem, but we refrain from applying it in this case (Why?) We should rather use a technique which reduces the problem to Proposition 3.6.6. Following this idea, we write

$$\begin{aligned} \tau_b &= \inf\{t > 0; S_t = b\} = \inf\{t > 0; \big(r - \frac{\sigma^2}{2}\big)t + \sigma W_t = \ln \frac{b}{S_0}\} \\ &= \inf\{t > 0; \mu t + \sigma W_t = -x\}, \end{aligned}$$

where $x = \ln \frac{S_0}{b} > 0$, $\mu = r - \frac{\sigma^2}{2}$. Choosing $s = r$ in Proposition 3.6.6 yields

$$\begin{aligned} E[e^{-r\tau_b}] &= e^{\frac{-1}{\sigma^2}(\mu + \sqrt{2r\sigma^2 + \mu^2})x} = e^{\frac{-1}{\sigma^2}\big(r - \frac{\sigma^2}{2} + \sqrt{2r\sigma^2 + (r - \frac{\sigma^2}{2})^2}\big)x} \\ &= e^{-\frac{2r}{\sigma^2}x} = e^{-\frac{2r}{\sigma^2}\ln\frac{S_0}{b}} = \Big(\frac{S_0}{b}\Big)^{-\frac{2r}{\sigma^2}}. \end{aligned}$$

In the previous computation we used that

$$\begin{aligned} & r - \frac{\sigma^2}{2} + \sqrt{2r\sigma^2 + \big(r - \frac{\sigma^2}{2}\big)^2} = r - \frac{\sigma^2}{2} + \sqrt{\big(r + \frac{\sigma^2}{2}\big)^2} \\ & = r - \frac{\sigma^2}{2} + r + \frac{\sigma^2}{2} = 2r. \end{aligned}$$

■

Exercise 17.1.7 *Assume the stock pays continuous dividends at the constant rate $\delta > 0$ and let $b > 0$ such that $b < S_0$.*

(a) *Use a similar method as in the proof of Proposition 17.1.6 to prove that the value at time* $t = 0$ *of $1 received at the time when the stock reaches level* b *from above is*

$$f_0 = \Big(\frac{S_0}{b}\Big)^{h_2},$$

where

$$h_2 = \frac{1}{2} - \frac{r-\delta}{\sigma^2} - \sqrt{\Big(\frac{r-\delta}{\sigma^2} - \frac{1}{2}\Big)^2 + \frac{2r}{\sigma^2}}.$$

(b) *What is the value of the contract at any time* t*, with* $0 \le t < \tau_b$*?*

Perpetual American options have simple exact pricing formulas. This is because of the "time invariance" property of their values. Since the time to expiration for these type of options is the same (i.e infinity), the option exercise problem looks the same at every instance of time. Consequently, their value do not depend on the time to expiration.

17.1.2 Perpetual American Calls

A *perpetual American call* is a call option that never expires, i.e. is a contract that gives the holder the right to buy the stock for the price K at any instance of time $0 \le t \le +\infty$. The infinity is included to cover the case when the option is never exercised.

When the call is exercised the holder receives $S_\tau - K$, where τ denotes the exercise time. Assume the holder has the strategy to exercise the call whenever the stock S_t reaches the barrier b, with $b > K$ subject to be determined later. Then at exercise time τ_b the payoff is $b - K$, where

$$\tau_b = \inf\{t > 0; S_t = b\}.$$

We note that it makes sense to choose the barrier such that $S_0 < b$. The value of this amount at time $t = 0$ is obtained discounting at the interest rate r and using Proposition 17.1.1

$$f(b) = E[(b-K)e^{-r\tau_b}] = (b-K)\frac{S_0}{b} = \Big(1 - \frac{K}{b}\Big)S_0.$$

We need to choose the value of the barrier $b > 0$ for which $f(b)$ reaches its maximum. Since $1 - \frac{K}{b}$ is an increasing function of b, the optimum value can be evaluated as

$$\max_{b>0} f(b) = \max_{b>0}\Big(1 - \frac{K}{b}\Big)S_0 = \lim_{b\to\infty}\Big(1 - \frac{K}{b}\Big)S_0 = S_0.$$

This is reached for the optimal barrier $b^* = \infty$, which corresponds to the infinite exercise time $\tau_{b^*} = \infty$. *Hence, it is never optimal to exercise a perpetual call option on a nondividend paying stock.*

The next exercise covers the case of the dividend paying stock. The method is similar with the one described previously.

Exercise 17.1.8 *Consider a stock that pays continuous dividends at rate $\delta > 0$.*

(*a*) *Assume a perpetual call is exercised whenever the stock reaches the barrier b from below. Show that the discounted value at time $t = 0$ is*

$$f(b) = (b - K)\Big(\frac{S_0}{h}\Big)^{h_1},$$

where

$$h_1 = \frac{1}{2} - \frac{r-\delta}{\sigma^2} + \sqrt{\Big(\frac{r-\delta}{\sigma^2} - \frac{1}{2}\Big)^2 + \frac{2r}{\sigma^2}}.$$

(*b*) *Use differentiation to show that the maximum value of $f(b)$ is realized for*

$$b^* = K\frac{h_1}{h_1 - 1}.$$

(*c*) *Prove the price of perpetual call*

$$f(b^*) = \max_{b>0} f(b) = \frac{K}{h_1 - 1}\Big(\frac{h_1 - 1}{h_1}\frac{S_0}{K}\Big)^{h_1}.$$

(*d*) *Let τ_{b^*} be the exercise time of the perpetual call. When do you expect to exercise the call? (Find $E[\tau_{b^*}]$).*

17.1.3 Perpetual American Puts

A *perpetual American put* is a put option that never expires, i.e. is a contract that gives the holder the right to sell the stock for the price K at any instance of time $0 \leq t < \infty$.

Assume the put is exercised when S_t reaches the barrier b. Then its payoff, $K - S_t = K - b$, has a couple of noteworthy features. First, if we choose b too large, we loose option value, which eventually vanishes for $b \geq K$. Second, if we pick b too small, the chances that the stock S_t will hit b are also small (see Exercise 17.1.9), fact that diminishes the put value. It follows that the optimum exercise barrier, b^*, is somewhere in between the two previous extreme values.

Exercise 17.1.9 *Let $0 < b < S_0$ and let $t > 0$ fixed.*

(*a*) *Show that the following inequality holds in the risk neutral world*

$$P(S_t < b) \leq e^{-\frac{1}{2\sigma^2 t}[\ln(S_0/b) + (r - \sigma^2/2)t]^2}.$$

(*b*) *Use the Squeeze Theorem to show that $\lim\limits_{b \to 0^+} P(S_t < b) = 0$.*

Since a put is an insurance that gets exercised when the the stock declines, it makes sense to assume that at the exercise time, τ_b, the stock reaches the barrier b from above,

i.e $0 < b < S_0$. Using Proposition 17.1.6 we obtain the value of the contract that pays $K - b$ at time τ_b

$$f(b) \quad = \quad E[(K-b)e^{-r\tau_b}] = (K-b)E[e^{-r\tau_b}] = (K-b)\Big(\frac{S_0}{b}\Big)^{-\frac{2r}{\sigma^2}}.$$

We need to pick the optimal value b^* for which $f(b)$ is maximum

$$f(b^*) = \max_{0<b<S_0} f(b).$$

It is useful to notice that the functions $f(b)$ and $g(b) = (K-b)b^{\frac{2r}{\sigma^2}}$ reach the maximum for the same value of b. For the same of simplicity, denote $\alpha = \frac{2r}{\sigma^2}$. Then

$$\begin{aligned} g(b) &= Kb^{\alpha} - b^{\alpha+1} \\ g'(b) &= \alpha K b^{\alpha-1} - (\alpha+1)b^{\alpha} = b^{\alpha-1}[\alpha K - (\alpha+1)b], \end{aligned}$$

and the equation $g'(b) = 0$ has the solution $b^* = \frac{\alpha}{\alpha+1}K$. Since $g'(b) > 0$ for $b < b^*$ and $g'(b) < 0$ for $b > b^*$, it follows that b^* is a maximum point for the function $g(b)$ and hence for the function $f(b)$. Substituting for the value of α the optimal value of the barrier becomes

$$b^* = \frac{2r/\sigma^2}{2r/\sigma^2 + 1}K = \frac{K}{1 + \frac{\sigma^2}{2r}}. \tag{17.1.1}$$

The condition $b^* < K$ is obviously satisfied, while the condition $b^* < S_0$ is equivalent with

$$K < \Big(1 + \frac{\sigma^2}{2r}\Big)S_0.$$

The value of the perpetual put is obtained computing the value at b^*

$$\begin{aligned} f(b^*) &= \max f(b) = (K - b^*)\Big(\frac{S_0}{b^*}\Big)^{-\frac{2r}{\sigma^2}} = \Big(K - \frac{K}{1+\frac{\sigma^2}{2r}}\Big)\Big[\frac{S_0}{K}\Big(1+\frac{\sigma^2}{2r}\Big)\Big]^{-\frac{2r}{\sigma^2}} \\ &= \frac{K}{1+\frac{2r}{\sigma^2}}\Big[\frac{S_0}{K}\Big(1+\frac{\sigma^2}{2r}\Big)\Big]^{-\frac{2r}{\sigma^2}}. \end{aligned}$$

Hence the price of a perpetual put is

$$\boxed{\frac{K}{1+\frac{2r}{\sigma^2}}\Big[\frac{S_0}{K}\Big(1+\frac{\sigma^2}{2r}\Big)\Big]^{-\frac{2r}{\sigma^2}}.}$$

The optimal exercise time of the put, τ_{b^*}, is when the stock hits the optimal barrier

$$\tau_{b^*} = \inf\{t > 0; S_t = b^*\} = \inf\{t > 0; S_t = K/\big(1 + \frac{\sigma^2}{2r}\big)\}.$$

But when is the expected exercise time of a perpetual American put? To answer this question we need to compute $E[\tau_{b^*}]$. Substituting

$$\tau = \tau_{b^*}, \quad x = \ln \frac{S_0}{b}, \quad \mu = r - \frac{\sigma^2}{2} \tag{17.1.2}$$

in Proposition 3.6.6 (c) yields

$$\begin{aligned} E[\tau_{b^*}] &= \frac{x}{\mu} e^{-\frac{2\mu x}{\sigma^2}} = \frac{\ln \frac{S_0}{b^*}}{r - \frac{\sigma^2}{2}} e^{-\frac{2}{\sigma^2}(r - \frac{\sigma^2}{2}) \ln \frac{S_0}{b^*}} \\ &= \ln \Big[\Big(\frac{S_0}{b^*} \Big)^{\frac{1}{r - \frac{\sigma^2}{2}}} \Big] e^{(1 - \frac{2r}{\sigma^2}) \ln \frac{S_0}{b^*}} = \ln \Big[\Big(\frac{S_0}{b^*} \Big)^{\frac{1}{r - \frac{\sigma^2}{2}}} \Big] \Big(\frac{S_0}{b^*} \Big)^{1 - \frac{2r}{\sigma^2}}. \end{aligned}$$

Hence the expected time when the holder should exercise the put is given by the exact formula

$$\boxed{E[\tau_{b^*}] = \ln \Big[\Big(\frac{S_0}{b^*} \Big)^{\frac{1}{r - \frac{\sigma^2}{2}}} \Big] \Big(\frac{S_0}{b^*} \Big)^{1 - \frac{2r}{\sigma^2}},}$$

with b^* given by (17.1.1). The probability density function of the optimal exercise time τ_{b^*} can be found from Proposition 3.6.6 (b) using substitutions (17.1.2)

$$p(\tau) = \frac{\ln \frac{S_0}{b^*}}{\sigma \sqrt{2\pi} \tau^{3/2}} e^{-\frac{\left(\ln \frac{S_0}{b^*} + \tau (r - \frac{\sigma^2}{2}) \right)^2}{2\tau\sigma^2}}, \qquad \tau > 0. \tag{17.1.3}$$

Exercise 17.1.10 *Consider a stock that pays continuous dividends at rate $\delta > 0$.*

(a) *Assume a perpetual put is exercised whenever the stock reaches the barrier b from above. Show that the discounted value at time $t = 0$ is*

$$g(b) = (K - b) \Big(\frac{S_0}{h} \Big)^{h_2},$$

where

$$h_2 = \frac{1}{2} - \frac{r - \delta}{\sigma^2} - \sqrt{\Big(\frac{r - \delta}{\sigma^2} - \frac{1}{2} \Big)^2 + \frac{2r}{\sigma^2}}.$$

(b) *Use differentiation to show that the maximum value of $g(b)$ is realized for*

$$b^* = K \frac{h_2}{h_2 - 1}.$$

(c) *Prove the price of perpetual put*

$$g(b^*) = \max_{b>0} g(b) = \frac{K}{1 - h_2} \Big(\frac{h_2 - 1}{h_2} \frac{S_0}{K} \Big)^{h_2}.$$

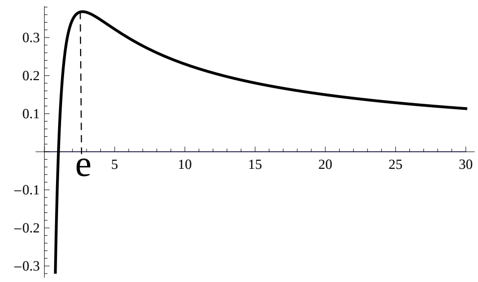


Figure 17.3: *The graph of the function* $g(b) = \frac{\ln b}{b}$, $b > 0$.

17.2 Perpetual American Log Contract

A *perpetual American log contract* is a contract that never expires and can be exercised at any time, providing the holder the log of the value of the stock, $\ln S_t$, at the exercise time t. It is interesting to note that these type of contracts are always optimal to be exercised, and their pricing formula is fairly uncomplicated.

Assume the contract is exercised when the stock S_t reaches the barrier b, with $S_0 < b$. If the hitting time of the barrier b is τ_b, then its payoff is $\ln S_\tau = \ln b$. Discounting at the risk free interest rate, the value of the contract at time $t = 0$ is

$$f(b) = E[e^{-r\tau_b} \ln S_\tau] = E[e^{-r\tau_b} \ln b] = \frac{\ln b}{b} S_0,$$

since the barrier is assumed to be reached from below, see Proposition 17.1.1.

The function $g(b) = \dfrac{\ln b}{b}$, $b > 0$, has the derivative $g'(b) = \dfrac{1 - \ln b}{b^2}$, so $b^* = e$ is a global maximum point, see Fig. 17.3. The maximum value is $g(b^*) = 1/e$. Then the optimal value of the barrier is $b^* = e$, and the price of the contract at $t = 0$ is

$$f_0 = \max_{b>0} f(b) = S_0 \max_{b>0} g(b) = \frac{S_0}{e}.$$

In order for the stock to reach the optimum barrier b^* from below we need to require the condition $S_0 < e$. Hence we arrived at the following result:

Proposition 17.2.1 *Let $S_0 < e$. Then the optimal exercise price of a perpetual American log contract is*

$$\tau = \inf\{t > 0; S_t = e\},$$

and its value at $t = 0$ is $\dfrac{S_0}{e}$.

Remark 17.2.2 *If $S_0 > e$, then it is optimal to exercise the perpetual log contract as soon as possible.*

Exercise 17.2.3 *Consider a stock that pays continuous dividends at a rate* $\delta > 0$, *and assume that* $S_0 < e^{1/h_1}$, *where*

$$h_1 = \frac{1}{2} - \frac{r-\delta}{\sigma^2} + \sqrt{\Big(\frac{r-\delta}{\sigma^2} - \frac{1}{2}\Big)^2 + \frac{2r}{\sigma^2}}.$$

(*a*) *Assume a perpetual log contract is exercised whenever the stock reaches the barrier* b *from below. Show that the discounted value at time* $t = 0$ *is*

$$f(b) = \ln b\Big(\frac{S_0}{b}\Big)^{h_1}.$$

(*b*) *Use differentiation to show that the maximum value of* $f(b)$ *is realized for*

$$b^* = e^{1/h_1}.$$

(*c*) *Prove the price of perpetual log contract*

$$f(b^*) = \max_{b>0} f(b) = \frac{S_0^{h_1}}{h_1 e}.$$

(*d*) *Show that the higher the dividend rate* δ, *the lower the optimal exercise time is.*

17.3 Perpetual American Power Contract

A *perpetual American power contract* is a contract that never expires and can be exercised at any time, providing the holder the n-th power of the value of the stock, $(S_t)^\alpha$, at the exercise time t, where $\alpha \neq 0$. (If $\alpha = 0$ the payoff is a constant, which is equal to 1).

1. Case $\alpha > 0$. Since we expect the value of the payoff to increase over time, we assume the contract is exercised when the stock S_t reaches the barrier b, from below. If the hitting time of the barrier b is τ_b, then its payoff is $(S_\tau)^\alpha$. Discounting at the risk free interest rate, the value of the contract at time $t = 0$ is

$$f(b) = E[e^{-r\tau_b}(S_\tau)^\alpha] = E[e^{-r\tau_b}b^\alpha] = b^\alpha\frac{S_0}{b} = b^{\alpha-1}S_0,$$

where we used Proposition 17.1.1. We shall discuss the following cases:

(*i*) If $\alpha > 1$, then the optimal barrier is $b^* = \infty$, and hence, it is never optimal to exercise the contract in this case.

(*ii*) If $0 < \alpha < 1$, the function $f(b)$ is decreasing, so its maximum is reached for $b^* = S_0$, which corresponds to $\tau_{b^*} = 0$. Hence is is optimal to exercise the contract as soon as possible.

(*iii*) In the case $\alpha = 1$, the value of $f(b)$ is constant, and the contract can be exercised at any time.

2. Case $\alpha < 0$. The payoff value, $(S_t)^\alpha$, is expected to decrease, so we assume the exercise occurs when the stock reaches the barrier b from above. Discounting to initial time $t = 0$ and using Proposition 17.1.6 yields

$$f(b) = E[e^{-r\tau_b}(S_\tau)^\alpha] = E[e^{-r\tau_b} b^\alpha] = b^\alpha \Big(\frac{S_0}{b}\Big)^{-\frac{2r}{\sigma^2}} = b^{\alpha + \frac{2r}{\sigma^2}} S_0^{-\frac{2r}{\sigma^2}}.$$

(i) If $\alpha < -\frac{2r}{\sigma^2}$, then $f(b)$ is decreasing, so its maximum is reached for $b^* = S_0$, which corresponds to $\tau_{b^*} = 0$. Hence is is optimal to exercise the contract as soon as possible.

(ii) If $-\frac{2r}{\sigma^2} < \alpha < 0$, then the maximum of $f(b)$ occurs for $b^* = \infty$. Hence it is never optimal to exercise the contract.

Exercise 17.3.1 *Consider a stock that pays continuous dividends at a rate* $\delta > 0$, *and assume* $\alpha > h_1$, *with* $h_1 = \frac{1}{2} - \frac{r-\delta}{\sigma^2} + \sqrt{\Big(\frac{r-\delta}{\sigma^2} - \frac{1}{2}\Big)^2 + \frac{2r}{\sigma^2}}$. *Show that the perpetual power contract with payoff* S_t^α *is never optimal to be exercised.*

Exercise 17.3.2 (Perpetual American power put) *Consider an perpetual American-type contract with the payoff* $(K - S_t)^2$, *where* $K > 0$ *is the strike price. Find the optimal exercise time and the contract value at* $t = 0$.

17.4 Finitely Lived American Options

Exact pricing formulas are great when they exist and can be easily implemented. Even if we cherish all closed form pricing formulas we can get, there is also a time when exact formulas are not possible and approximations are in order. If we run into a problem whose solution cannot be found explicitly, it would still be very valuable to know something about its approximate quantitative behavior. This will be the case of finitely lived American options.

17.4.1 American Call

The case of Non-dividend Paying Stock

The holder has the right to buy a stock for the price K at any time before or at the expiration T. The strike price K and expiration T are specified at the beginning of the contract. The payoff at time t is $(S_t - K)^+ = \max\{S_t - K, 0\}$. The price of the American call at time $t = 0$ is given by

$$f_0 = \max_{0 \le \tau \le T} \widehat{E}[e^{-r\tau}(S_\tau - K)^+],$$

where the maximum is taken over all stopping times τ less than or equal to T.

Theorem 17.4.1 *It is not optimal to exercise an American call on a non-dividend paying stock early. It is optimal to exercise the call at maturity, T, if at all. Consequently, the price of an American call is equal to the price of the corresponding European call.*

Proof: The heuristic idea of the proof is based on the observation that the difference $S_t - K$ tends to be larger as time goes on. As a result, there is always hope for a larger payoff and the later we exercise the better. In the following we shall formalize this idea mathematically using the submartingale property of the stock together with the Optional Stopping Theorem.

Let $X_t = e^{-rt}S_t$ and $f(t) = -Ke^{-rt}$. Since X_t is a martingale in the risk-neutral-world, see Proposition 14.1.1, and $f(t)$ is an increasing, integrable function, applying Proposition 2.12.3 (c) it follows that

$$Y_t = X_t + f(t) = e^{-rt}(S_t - K)$$

is an $\mathcal{F}_t$-submartingale, where $\mathcal{F}_t$ is the information set provided by the underlying Brownian motion W_t.

Since the hokey-stick function $\phi(x) = x^+ = \max\{x, 0\}$ is convex, then by Proposition 2.12.3 (b), the process $Z_t = \phi(Y_t) = e^{-rt}(S_t - K)^+$ is a submartingale. Applying Doob's stopping theorem (see Theorem 3.2.2) for stopping times τ and T, with $\tau \le T$, we obtain $\widehat{E}[Z_\tau] \le \widehat{E}[Z_T]$. This means

$$\widehat{E}[e^{-r\tau}(S_\tau - K)^+] \le \widehat{E}[e^{-rT}(S_T - K)^+],$$

i.e. the maximum of the American call price is realized for the optimum exercise time $\tau^* = T$. The maximum value is given by the right side, which denotes the price of an European call option.

■

With a slight modification in the proof we can treat the problem of American power contract.

Proposition 17.4.2 (American power contract) *Consider a contract with maturity date T, which pays, when exercised, S_t^n, where $n > 1$, and $t \le T$. Then it is not optimal to exercise this contract early.*

Proof: Using that $M_t = e^{-rt}S_t$ is a martingale (in the risk neutral world), then $X_t = M_t^n$ is a submartingale. Since $Y_t = e^{-rt}S_t^n = \left(e^{-rt}S_t\right)^n e^{(n-1)rt} = X_t e^{(n-1)rt}$, then for $s < t$

$$\begin{aligned} E[Y_t|\mathcal{F}_s] &= E[X_t e^{(n-1)rt}|\mathcal{F}_s] > E[X_t e^{(n-1)rs}|\mathcal{F}_s] \\ &= e^{(n-1)rs}E[X_t|\mathcal{F}_s] \ge e^{(n-1)rs}X_s = Y_s, \end{aligned}$$

so Y_t is a submartingale. Applying Doob's stopping theorem (see Theorem 3.2.2) for stopping times τ and T, with $\tau \le T$, we obtain $\widehat{E}[Y_\tau] \le \widehat{E}[Y_T]$, or equivalently, $\widehat{E}[e^{-r\tau}S_\tau^n] \le \widehat{E}[e^{-rT}S_T^n]$, which implies

$$\max_{\tau \le T} \widehat{E}[e^{-r\tau}S_\tau^n] = \widehat{E}[e^{-rT}S_T^n].$$

Then it is optimal to exercise the contract at maturity T. ∎

Exercise 17.4.3 *Consider an American future contract with maturity date T and delivery price K, i.e. a contract with payoff at maturity $S_T - K$, which can be exercised at any time $t \le T$. Show that it is not optimal to exercise this contract early.*

Exercise 17.4.4 *Consider an American option contract with maturity date T, and time-dependent strike price $K(t)$, i.e. a contract with payoff at maturity $S_T - K(T)$, which can be exercised at any time.*

(*a*) *Show that it is not optimal to exercise this contract early in the following two cases:*
   (*i*) *If $K(t)$ is a decreasing function;*
   (*ii*) *If $K(t)$ is an increasing function with $K(t) < e^{rt}$.*

(*b*) *What happens if $K(t)$ is increasing and $K(t) > e^{rt}$?*

The case of Dividend Paying Stock

When the stock pays dividends it is optimal to exercise the American call early. An exact solution of this problem is hard to get explicitly, or might not exist. However, there are some asymptotic solutions that are valid close to expiration (see Wilmott) and analytic approximations given by MacMillan, Barone-Adesi and Whaley.

In the following we shall discuss why it is difficult to find an exact optimal exercise time for an American call on a dividend paying stock. First, consider two contracts:

1. Consider a contract by which one can acquire a stock, S_t, at any time before or at time T. This can be seen as a contract with expiration T, that pays when exercised the stock price, S_t. Assume the stock pays dividends at a continuous rate $\delta > 0$. When should the contract be exercised in order to maximize its value?

The value of the contract at time $t = 0$ is $\max\limits_{\tau \le T}\{e^{-r\tau}S_\tau\}$, where the maximum is taken over all stopping times τ less than or equal to T. Since the stock price, which pays dividends at rate δ, is given by

$$S_t = S_0 e^{(r-\delta-\frac{\sigma^2}{2})t + \sigma W_t},$$

then $M_t = e^{-(r-\delta)t}S_t$ is a martingale (in the risk neutral world). Therefore $e^{-rt}S_t = e^{-\delta t}M_t$. Let $X_t = e^{-rt}S_t$. Then for $0 < s < t$

$$\begin{aligned} \widehat{E}[X_t|\mathcal{F}_s] &= \widehat{E}[e^{-\delta t}M_t|\mathcal{F}_s] < \widehat{E}[e^{-\delta s}M_t|\mathcal{F}_s] \\ &= e^{-\delta s}\widehat{E}[M_t|\mathcal{F}_s] = e^{-\delta s}M_s = e^{-rs}S_s = M_s, \end{aligned}$$

so X_t is a supermartingale (i.e. $-X_t$ is a submartingale). Applying the Optional Stopping Theorem for the stopping time τ we obtain

$$\widehat{E}[X_\tau] \leq \widehat{E}[X_0] = S_0.$$

Hence it is optimal to exercise the contract at the initial time $t = 0$. This makes sense, since in this case we have a longer period of time during which dividends are collected.

2. Consider a contract by which one has to pay the amount of cash K at any time before or at time T. Given the time value of money

$$e^{-rt}K > e^{-rT}K, \qquad t < T,$$

it is always optimal to defer the payment until time T.

3. Now consider a combination of the previous two contracts. This new contract pays $S_t - K$ and can be exercised at any time t, with $t \leq T$. Since it is not clear when is optimal to exercise this contract, we shall consider two limiting cases. Let τ^* denote its optimal exercise time.

(i) When $K \to 0^+$, then $\tau^* \to 0^+$, because we approach the conditions of case 1, and also assume continuity conditions on the price.

(ii) If $K \to \infty$, then the latter the pay day, the better, i.e. $\tau^* \to T^-$.

The optimal exercise time, τ^*, is *somewhere* between 0 and T, with the tendency of moving towards T as K gets large.

17.4.2 American Put

17.4.3 Mac Millan-Barone-Adesi-Whaley Approximation

TO COME

17.4.4 Black's Approximation

TO COME

17.4.5 Roll-Geske-Whaley Approximation

TO COME

17.4.6 Other Approximations

Chapter 18

Hints and Solutions

Chapter 1

1.6.1 Let $X \sim N(\mu, \sigma^2)$. Then the distribution function of Y is

$$\begin{aligned} F_Y(y) &= P(Y < y) = P(\alpha X + \beta < y) = P\Big(X < \frac{y-\beta}{\alpha}\Big) \\ &= \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\frac{y-\beta}{\alpha}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\, dx = \frac{1}{\sqrt{2\pi}\alpha\sigma} \int_{-\infty}^{y} e^{-\frac{\left(z-(\alpha\mu+\beta)\right)^2}{2\alpha^2\sigma^2}}\, dz \\ &= \frac{1}{\sqrt{2\pi}\sigma'} \int_{-\infty}^{y} e^{-\frac{(z-\mu')^2}{2(\sigma')^2}}\, dz, \end{aligned}$$

with $\mu' = \alpha\mu + \beta$, $\sigma' = \alpha\sigma$.

1.6.2 (a) Making $t = n$ yields $E[Y^n] = E[e^{nX}] = e^{\mu n + n^2\sigma^2/2}$.

(b) Let $n = 1$ and $n = 2$ in (a) to get the first two moments and then use the formula of variance.

1.11.5 The tower property

$$E\big[E[X|\mathcal{G}]|\mathcal{H}\big] = E[X|\mathcal{H}], \qquad \mathcal{H} \subset \mathcal{G}$$

is equivalent with

$$\int_A E[X|\mathcal{G}]\, dP = \int X\, dP, \qquad \forall A \in \mathcal{H}.$$

Since $A \in \mathcal{G}$, the previous relation holds by the definition on $E[X|\mathcal{G}]$.

1.11.6 (a) Direct application of the definition.

(b) $P(A) = \int_A dP = \int_\Omega \chi_A(\omega)\, dP(\omega) = E[\chi_A]$.

(d) $E[\chi_A X] = E[\chi_A]E[X] = P(A)E[X]$.

(e) We have the sequence of equivalencies

$$\begin{aligned} E[X|\mathcal{G}] &= E[X] \Leftrightarrow \int_A E[X]\,dP = \int_A X\,dP, \forall A \in \mathcal{G} \Leftrightarrow \\ E[X]P(A) &= \int_A X\,dP \Leftrightarrow E[X]P(A) = E[\chi_A], \end{aligned}$$

which follows from (d).

1.12.4 If $\mu = E[X]$ then

$$\begin{aligned} E[(X - E[X])^2] &= E[X^2 - 2\mu X + \mu^2] = E[X^2] - 2\mu^2 + \mu^2 \\ &= E[X^2] - E[X]^2 = Var[X]. \end{aligned}$$

1.12.5 From **1.12.4** we have $Var(X) = 0 \Leftrightarrow X = E[X]$, i.e. X is a constant.

1.12.6 The same proof as the one in Jensen's inequality.

1.12.7 It follows from Jensen's inequality or using properties of integrals.

1.12.12 (a) $m(t) = E[e^{tX}] = \sum_k e^{tk} \frac{\lambda^k e^{-\lambda}}{k!} = e^{\lambda(e^t - 1)}$;
(b) It follows from the first Chernoff bound.

1.12.16 Choose $f(x) = x^{2k+1}$ and $g(x) = x^{2n+1}$.

1.13.2 By direct computation we have

$$\begin{aligned} E[(X - Y)^2] &= E[X^2] + E[Y^2] - 2E[XY] \\ &= Var(X) + E[X]^2 + Var[Y] + E[Y]^2 - 2E[X]E[Y] \\ &\quad +2E[X]E[Y] - 2E[XY] \\ &= Var(X) + Var[Y] + (E[X] - E[Y])^2 - 2Cov(X, Y). \end{aligned}$$

1.13.3 (a) Since

$$E[(X - X_n)^2] \geq E[X - X_n]^2 = (E[X_n] - E[X])^2 \geq 0,$$

the Squeeze theorem yields $\lim_{n\to\infty}(E[X_n] - E[X]) = 0$.
(b) Writing

$$X_n^2 - X^2 = (X_n - X)^2 - 2X(X - X_n),$$

and taking the expectation we get

$$E[X_n^2] - E[X^2] = E[(X_n - X)^2] - 2E[X(X - X_n)].$$

The right side tends to zero since

$$
\begin{aligned}
E[(X_n - X)^2] &\to 0 \\
|E[X(X - X_n)]| &\le \int_\Omega |X(X - X_n)|\, dP \\
&\le \Big(\int_\Omega X^2\, dP\Big)^{1/2}\Big(\int_\Omega (X - X_n)^2\, dP\Big)^{1/2} \\
&= \sqrt{E[X^2]E[(X - X_n)^2]} \to 0.
\end{aligned}
$$

(c) It follows from part (b).

(d) Apply Exercise **1.13.2**.

1.13.4 Using Jensen's inequality we have

$$
\begin{aligned}
E\big[\big(E[X_n|\mathcal{H}] - E[X|\mathcal{H}]\big)^2\big] &= E\big[\big(E[X_n - X|\mathcal{H}]\big)^2\big] \\
&\le E\big[E[(X_n - X)^2|\mathcal{H}]\big] \\
&= E[(X_n - X)^2] \to 0,
\end{aligned}
$$

as $n \to 0$.

1.15.7 The integrability of X_t follows from

$$
E[|X_t|] = E\big[|E[X|\mathcal{F}_t]|\big] \le E\big[E[|X|\,|\mathcal{F}_t]\big] = E[|X|] < \infty.
$$

X_t is $\mathcal{F}_t$-predictable by the definition of the conditional expectation. Using the tower property yields

$$
E[X_t|\mathcal{F}_s] = E\big[E[X|\mathcal{F}_t]|\mathcal{F}_s\big] = E[X|\mathcal{F}_s] = X_s, \qquad s < t.
$$

1.15.8 Since

$$
E[|Z_t|] = E[|aX_t + bY_t + c|] \le |a|E[|X_t|] + |b|E[|Y_t|] + |c| < \infty
$$

then Z_t is integrable. For $s < t$, using the martingale property of X_t and Y_t we have

$$
E[Z_t|\mathcal{F}_s] = aE[X_t|\mathcal{F}_s] + bE[Y_t|\mathcal{F}_s] + c = aX_s + bY_s + c = Z_s.
$$

1.15.9 In general the answer is no. For instance, if $X_t = Y_t$ the process X_t^2 is not a martingale, since the Jensen's inequality

$$
E[X_t^2|\mathcal{F}_s] \ge \Big(E[X_t|\mathcal{F}_s]\Big)^2 = X_s^2
$$

is not necessarily an identity. For instance B_t^2 is not a martingale, with B_t the Brownian motion process.

1.15.10 It follows from the identity

$$E[(X_t - X_s)(Y_t - Y_s)|\mathcal{F}_s] = E[X_tY_t - X_sY_s|\mathcal{F}_s].$$

1.15.11 (a) Let $Y_n = S_n - E[S_n]$. We have

$$\begin{aligned} Y_{n+k} &= S_{n+k} - E[S_{n+k}] \\ &= Y_n + \sum_{j=1}^{k} X_{n+j} - \sum_{j=1}^{k} E[X_{n+j}]. \end{aligned}$$

Using the properties of expectation we have

$$\begin{aligned} E[Y_{n+k}|\mathcal{F}_n] &= Y_n + \sum_{j=1}^{k} E[X_{n+j}|\mathcal{F}_n] - \sum_{j=1}^{k} E[E[X_{n+j}]|\mathcal{F}_n] \\ &= Y_n + \sum_{j=1}^{k} E[X_{n+j}] - \sum_{j=1}^{k} E[X_{n+j}] \\ &= Y_n. \end{aligned}$$

(b) Let $Z_n = S_n^2 - Var(S_n)$. The process Z_n is an $\mathcal{F}_n$-martingale iff

$$E[Z_{n+k} - Z_n|\mathcal{F}_n] = 0.$$

Let $U = S_{n+k} - S_n$. Using the independence we have

$$\begin{aligned} Z_{n+k} - Z_n &= (S_{n+k}^2 - S_n^2) - \big(Var(S_{n+k} - Var(S_n)\big) \\ &= (S_n + U)^2 - S_n^2 - \big(Var(S_{n+k} - Var(S_n)\big) \\ &= U^2 + 2US_n - Var(U), \end{aligned}$$

so

$$\begin{aligned} E[Z_{n+k} - Z_n|\mathcal{F}_n] &= E[U^2] + 2S_nE[U] - Var(U) \\ &= E[U^2] - (E[U^2] - E[U]^2) \\ &= 0, \end{aligned}$$

since $E[U] = 0$.

1.15.12 Let $\mathcal{F}_n = \sigma(X_k; k \le n)$. Using the independence

$$E[|P_n|] = E[|X_0|] \cdots E[|X_n|] < \infty,$$

so $|P_n|$ integrable. Taking out the predictable part we have

$$\begin{aligned} E[P_{n+k}|\mathcal{F}_n] &= E[P_nX_{n+1} \cdots X_{n+k}|\mathcal{F}_n] = P_nE[X_{n+1} \cdots X_{n+k}|\mathcal{F}_n] \\ &= P_nE[X_{n+1}] \cdots E[X_{n+k}] = P_n. \end{aligned}$$

1.15.13 (*a*) Since the random variable $Y = \theta X$ is normally distributed with mean $\theta\mu$ and variance $\theta^2\sigma^2$, then

$$E[e^{\theta X}] = e^{\theta\mu + \frac{1}{2}\theta^2\sigma^2}.$$

Hence $E[e^{\theta X}] = 1$ iff $\theta\mu + \frac{1}{2}\theta^2\sigma^2 = 0$ which has the nonzero solution $\theta = -2\mu/\sigma^2$.

(*b*) Since $e^{\theta X_i}$ are independent, integrable and satisfy $E[e^{\theta X_i}] = 1$, by Exercise **1.15.12** we get that the product $Z_n = e^{\theta S_n} = e^{\theta X_1} \cdots e^{\theta X_n}$ is a martingale.

Chapter 2

2.1.4 B_t starts at 0 and is continuous in t. By Proposition 2.1.2 B_t is a martingale with $E[B_t^2] = t < \infty$. Since $B_t - B_s \sim N(0, |t-s|)$, then $E[(B_t - B_s)^2] = |t-s|$.

2.1.10 It is obvious that $X_t = W_{t+t_0} - W_{t_0}$ satisfies $X_0 = 0$ and that X_t is continuous in t. The increments are normal distributed $X_t - X_s = W_{t+t_0} - W_{s+t_0} \sim N(0, |t-s|)$. If $0 < t_1 < \cdots < t_n$, then $0 < t_0 < t_1 + t_0 < \cdots < t_n + t_0$. The increments $X_{t_{k+1}} - X_{t_k} = W_{t_{k+1}+t_0} - W_{t_k+t_0}$ are obviously independent and stationary.

2.1.11 For any $\lambda > 0$, show that the process $X_t = \frac{1}{\sqrt{\lambda}} W_{\lambda t}$ is a Brownian motion. This says that the Brownian motion is invariant by scaling. Let $s < t$. Then $X_t - X_s = \frac{1}{\sqrt{\lambda}}(W_{\lambda t} - W_{\lambda s}) \sim \frac{1}{\sqrt{\lambda}} N\big(0, \lambda(t-s)\big) = N(0, t-s)$. The other properties are obvious.

2.1.12 Apply Property 2.1.8.

2.1.13 Using the moment generating function, we get $E[W_t^3] = 0$, $E[W_t^4] = 3t^2$.

2.1.14 (*a*) Let $s < t$. Then

$$\begin{aligned} E[(W_t^2 - t)(W_s^2 - s)] &= E\Big[E[(W_t^2 - t)(W_s^2 - s)]|\mathcal{F}_s\Big] = E\Big[(W_s^2 - s)E[(W_t^2 - t)]|\mathcal{F}_s\Big] \\ &= E\Big[(W_s^2 - s)^2\Big] = E\Big[W_s^4 - 2sW_s^2 + s^2\Big] \\ &= E[W_s^4] - 2sE[W_s^2] + s^2 = 3s^2 - 2s^2 + s^2 = 2s^2. \end{aligned}$$

(*b*) Using part (*a*) we have

$$\begin{aligned} 2s^2 = E[(W_t^2 - t)(W_s^2 - s)] &= E[W_s^2 W_t^2] - sE[W_t^2] - tE[W_s^2] + ts \\ &= E[W_s^2 W_t^2] - st. \end{aligned}$$

Therefore $E[W_s^2 W_t^2] = ts + 2s^2$.

(*c*) $Cov(W_t^2, W_s^2) = E[W_s^2 W_t^2] - E[W_s^2]E[W_t^2] = ts + 2s^2 - ts = 2s^2$.

(*d*) $Corr(W_t^2, W_s^2) = \frac{2s^2}{2ts} = \frac{s}{t}$, where we used

$$Var(W_t^2) = E[W_t^4] - E[W_t^2]^2 = 3t^2 - t^2 = 2t^2.$$

2.1.15 (*a*) The distribution function of Y_t is given by

$$\begin{aligned} F(x) &= P(Y_t \le x) = P(tW_{1/t} \le x) = P(W_{1/t} \le x/t) \\ &= \int_0^{x/t} \phi_{1/t}(y)\, dy = \int_0^{x/t} \sqrt{t/(2\pi)} e^{ty^2/2}\, dy \\ &= \int_0^{x/\sqrt{t}} \frac{1}{\sqrt{2\pi}} e^{-u^2/2}\, du. \end{aligned}$$

(*b*) The probability density of Y_t is obtained by differentiating $F(x)$

$$p(x) = F'(x) = \frac{d}{dx} \int_0^{x/\sqrt{t}} \frac{1}{\sqrt{2\pi}} e^{-u^2/2}\, du = \frac{1}{\sqrt{2\pi}} e^{-x^2/2},$$

and hence $Y_t \sim N(0,t)$.
(*c*) Using that Y_t has independent increments we have

$$\begin{aligned} Cov(Y_s, Y_t) &= E[Y_s Y_t] - E[Y_s]E[Y_t] = E[Y_s Y_t] \\ &= E\Big[Y_s(Y_t - Y_s) + Y_s^2\Big] = E[Y_s]E[Y_t - Y_s] + E[Y_s^2] \\ &= 0 + s = s. \end{aligned}$$

(*d*) Since

$$\begin{aligned} Y_t - Y_s &= (t-s)(W_{1/t} - W_0) - s(W_{1/s} - W_{1/t}) \\ E[Y_t - Y_s] &= (t-s)E[W_{1/t}] - sE[W_{1/s} - W_{1/t}] = 0, \end{aligned}$$

and

$$\begin{aligned} Var(Y_t - Y_s) &= E[(Y_t - Y_s)^2] = (t-s)^2 \frac{1}{t} + s^2(\frac{1}{s} - \frac{1}{t}) \\ &= \frac{(t-s)^2 + s(t-s)}{t} = t - s. \end{aligned}$$

2.1.16 (*a*) Applying the definition of expectation we have

$$\begin{aligned} E[|W_t|] &= \int_{-\infty}^{\infty} |x| \frac{1}{\sqrt{2\pi t}} e^{-\frac{x^2}{2t}}\, dx = \int_0^{\infty} 2x \frac{1}{\sqrt{2\pi t}} e^{-\frac{x^2}{2t}}\, dx \\ &= \frac{1}{\sqrt{2\pi t}} \int_0^{\infty} e^{-\frac{y}{2t}}\, dy = \sqrt{2t/\pi}. \end{aligned}$$

(*b*) Since $E[|W_t|^2] = E[W_t^2] = t$, we have

$$Var(|W_t|) = E[|W_t|^2] - E[|W_t|]^2 = t - \frac{2t}{\pi} = t(1 - \frac{2}{\pi}).$$

2.1.17 By the martingale property of $W_t^2 - t$ we have

$$E[W_t^2 | \mathcal{F}_s] = E[W_t^2 - t | \mathcal{F}_s] + t = W_s^2 + t - s.$$

2.1.18 (a) Expanding

$$(W_t - W_s)^3 = W_t^3 - 3W_t^2 W_s + 3W_t W_s^2 - W_s^3$$

and taking the expectation

$$\begin{aligned} E[(W_t - W_s)^3|\mathcal{F}_s] &= E[W_t^3|\mathcal{F}_s] - 3W_s E[W_t^2] + 3W_s^2 E[W_t|\mathcal{F}_s] - W_s^3 \\ &= E[W_t^3|\mathcal{F}_s] - 3(t-s)W_s - W_s^3, \end{aligned}$$

so

$$E[W_t^3|\mathcal{F}_s] = 3(t-s)W_s + W_s^3,$$

since

$$E[(W_t - W_s)^3|\mathcal{F}_s] = E[(W_t - W_s)^3] = E[W_{t-s}^3] = 0.$$

(b) Hint: Start from the expansion of $(W_t - W_s)^4$.

2.2.3 Using that $e^{W_t - W_s}$ is stationary, we have

$$E[e^{W_t - W_s}] = E[e^{W_{t-s}}] = e^{\frac{1}{2}(t-s)}.$$

2.2.4 (a)

$$\begin{aligned} E[X_t|\mathcal{F}_s] &= E[e^{W_t}|\mathcal{F}_s] = E[e^{W_t - W_s} e^{W_s}|\mathcal{F}_s] \\ &= e^{W_s} E[e^{W_t - W_s}|\mathcal{F}_s] = e^{W_s} E[e^{W_t - W_s}] \\ &= e^{W_s} e^{t/2} e^{-s/2}. \end{aligned}$$

(b) This can be written also as

$$E[e^{-t/2} e^{W_t}|\mathcal{F}_s] = e^{-s/2} e^{W_s},$$

which shows that $e^{-t/2} e^{W_t}$ is a martingale.
(c) From the stationarity we have

$$E[e^{cW_t - cW_s}] = E[e^{c(W_t - W_s)}] = E[e^{cW_{t-s}}] = e^{\frac{1}{2}c^2(t-s)}.$$

Then for any $s < t$ we have

$$\begin{aligned} E[e^{cW_t}|\mathcal{F}_s] &= E[e^{c(W_t - W_s)} e^{cW_s}|\mathcal{F}_s] = e^{cW_s} E[e^{c(W_t - W_s)}|\mathcal{F}_s] \\ &= e^{cW_s} E[e^{c(W_t - W_s)}] = e^{cW_s} e^{\frac{1}{2}c^2(t-s)} = Y_s e^{\frac{1}{2}c^2 t}. \end{aligned}$$

Multiplying by $e^{-\frac{1}{2}c^2 t}$ yields the desired result.

2.2.5 (a) Using Exercise **2.2.3** we have

$$\begin{aligned} Cov(X_s, X_t) &= E[X_s X_t] - E[X_s]E[X_t] = E[X_s X_t] - e^{t/2} e^{s/2} \\ &= E[e^{W_s + W_t}] - e^{t/2} e^{s/2} = E[e^{W_t - W_s} e^{2(W_s - W_0)}] - e^{t/2} e^{s/2} \\ &= E[e^{W_t - W_s}] E[e^{2(W_s - W_0)}] - e^{t/2} e^{s/2} = e^{\frac{t-s}{2}} e^{2s} - e^{t/2} e^{s/2} \\ &= e^{\frac{t+3s}{2}} - e^{(t+s)/2}. \end{aligned}$$

(b) Using Exercise **2.2.4** (b), we have

$$\begin{aligned} E[X_sX_t] &= E\Big[E[X_sX_t|\mathcal{F}_s]\Big] = E\Big[X_sE[X_t|\mathcal{F}_s]\Big] \\ &= e^{t/2}E\Big[X_sE[e^{-t/2}X_t|\mathcal{F}_s]\Big] = e^{t/2}E\Big[X_s\,e^{-s/2}X_s]\Big] \\ &= e^{(t-s)/2}E[X_s^2] = e^{(t-s)/2}E[e^{2W_s}] \\ &= e^{(t-s)/2}e^{2s} = e^{\frac{t+3s}{2}}, \end{aligned}$$

and continue like in part (a).

2.2.6 Using the definition of the expectation we have

$$\begin{aligned} E[e^{2W_t^2}] &= \int e^{2x^2}\phi_t(x)\,dx = \frac{1}{\sqrt{2\pi t}}\int e^{2x^2}e^{-\frac{x^2}{2t}}\,dx \\ &= \frac{1}{\sqrt{2\pi t}}\int e^{-\frac{1-4t}{2t}x^2}\,dx = \frac{1}{\sqrt{1-4t}}, \end{aligned}$$

if $1-4t>0$. Otherwise, the integral is infinite. We used the standard integral $\displaystyle\int e^{-ax^2} = \sqrt{\pi/a}$, $a>0$.

2.3.4 It follows from the fact that Z_t is normally distributed.

2.3.5 Using the definition of covariance we have

$$\begin{aligned} Cov\Big(Z_s, Z_t\Big) &= E[Z_sZ_t] - E[Z_s]E[Z_t] = E[Z_sZ_t] \\ &= E\Big[\int_0^s W_u\,du\cdot\int_0^t W_v\,dv\Big] = E\Big[\int_0^s\int_0^t W_uW_v\,dudv\Big] \\ &= \int_0^s\int_0^t E[W_uW_v]\,dudv = \int_0^s\int_0^t \min\{u,v\}\,dudv \\ &= s^2\Big(\frac{t}{2}-\frac{s}{6}\Big), \qquad s<t. \end{aligned}$$

2.3.6 (a) Using Exercise **2.3.5**

$$\begin{aligned} Cov(Z_t, Z_t - Z_{t-h}) &= Cov(Z_t, Z_t) - Cov(Z_t, Z_{t-h}) \\ &= \frac{t^3}{3} - (t-h)^2\Big(\frac{t}{2} - \frac{t-h}{6}\Big) \\ &= \frac{1}{2}t^2h + o(h). \end{aligned}$$

(b) Using $Z_t - Z_{t-h} = \int_{t-h}^t W_u\,du = hW_t + o(h)$,

$$\begin{aligned} Cov(Z_t, W_t) &= \frac{1}{h}Cov(Z_t, Z_t - Z_{t-h}) \\ &= \frac{1}{h}\Big(\frac{1}{2}t^2h + o(h)\Big) = \frac{1}{2}t^2. \end{aligned}$$

2.3.7 Let $s < u$. Since W_t has independent increments, taking the expectation in

$$e^{W_s+W_t} = e^{W_t-W_s}e^{2(W_s-W_0)}$$

we obtain

$$\begin{aligned} E[e^{W_s+W_t}] &= E[e^{W_t-W_s}]E[e^{2(W_s-W_0)}] = e^{\frac{u-s}{2}}e^{2s} \\ &= e^{\frac{u+s}{2}}e^{s} = e^{\frac{u+s}{2}}e^{\min\{s,t\}}. \end{aligned}$$

2.3.8 (a) $E[X_t] = \int_0^t E[e^{W_s}]\,ds = \int_0^t E[e^{s/2}]\,ds = 2(e^{t/2} - 1)$
(b) Since $Var(X_t) = E[X_t^2] - E[X_t]^2$, it suffices to compute $E[X_t^2]$. Using Exercise **2.3.7** we have

$$\begin{aligned} E[X_t^2] &= E\Big[\int_0^t e^{W_t}\,ds \cdot \int_0^t e^{W_u}\,du\Big] = E\Big[\int_0^t \int_0^t e^{W_s}e^{W_u}\,dsdu\Big] \\ &= \int_0^t \int_0^t E[e^{W_s+W_u}]\,dsdu = \int_0^t \int_0^t e^{\frac{u+s}{2}}e^{\min\{s,t\}}\,dsdu \\ &= \iint_{D_1} e^{\frac{u+s}{2}}e^{s}duds + \iint_{D_2} e^{\frac{u+s}{2}}e^{u}duds \\ &= 2\iint_{D_2} e^{\frac{u+s}{2}}e^{u}duds = \frac{4}{3}\Big(\frac{1}{2}e^{2t} - 2e^{t/2} + \frac{3}{2}\Big), \end{aligned}$$

where $D_1\{0 \le s < u \le t\}$ and $D_2\{0 \le u < s \le t\}$. In the last identity we applied Fubini's theorem. For the variance use the formula $Var(X_t) = E[X_t^2] - E[X_t]^2$.

2.3.9 (a) Splitting the integral at t and taking out the predictable part, we have

$$\begin{aligned} E[Z_T|\mathcal{F}_t] &= E[\int_0^T W_u\,du|\mathcal{F}_t] = E[\int_0^t W_u\,du|\mathcal{F}_t] + E[\int_t^T W_u\,du|\mathcal{F}_t] \\ &= Z_t + E[\int_t^T W_u\,du|\mathcal{F}_t] \\ &= Z_t + E[\int_t^T (W_u - W_t + W_t)\,du|\mathcal{F}_t] \\ &= Z_t + E[\int_t^T (W_u - W_t)\,du|\mathcal{F}_t] + W_t(T-t) \\ &= Z_t + E[\int_t^T (W_u - W_t)\,du] + W_t(T-t) \\ &= Z_t + \int_t^T E[W_u - W_t]\,du + W_t(T-t) \\ &= Z_t + W_t(T-t), \end{aligned}$$

since $E[W_u - W_t] = 0$.
(b) Let $0 < t < T$. Using (a) we have

$$\begin{aligned} E[Z_T - TW_T|\mathcal{F}_t] &= E[Z_T|\mathcal{F}_t] - TE[W_T|\mathcal{F}_t] \\ &= Z_t + W_t(T-t) - TW_t \\ &= Z_t - tW_t. \end{aligned}$$

2.4.1

$$\begin{aligned} E[V_T|\mathcal{F}_t] &= E\Big[e^{\int_0^t W_u\,du + \int_t^T W_u\,du}|\mathcal{F}_t\Big] \\ &= e^{\int_0^t W_u\,du} E\Big[e^{\int_t^T W_u\,du}|\mathcal{F}_t\Big] \\ &= e^{\int_0^t W_u\,du} E\Big[e^{\int_t^T (W_u - W_t)\,du + (T-t)W_t}|\mathcal{F}_t\Big] \\ &= V_t e^{(T-t)W_t} E\Big[e^{\int_t^T (W_u - W_t)\,du}|\mathcal{F}_t\Big] \\ &= V_t e^{(T-t)W_t} E\Big[e^{\int_t^T (W_u - W_t)\,du}\Big] \\ &= V_t e^{(T-t)W_t} E\Big[e^{\int_0^{T-t} W_\tau\,d\tau}\Big] \\ &= V_t e^{(T-t)W_t} e^{\frac{1}{2}\frac{(T-t)^3}{3}}. \end{aligned}$$

2.6.1

$$\begin{aligned} F(x) &= P(Y_t \le x) = P(\mu t + W_t \le x) = P(W_t \le x - \mu t) \\ &= \int_0^{x-\mu t} \frac{1}{\sqrt{2\pi t}} e^{-\frac{u^2}{2t}}\,du; \\ f(x) &= F'(x) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{(x-\mu t)^2}{2t}}. \end{aligned}$$

2.7.2 Since

$$P(R_t \le \rho) = \int_0^\rho \frac{1}{t} x e^{-\frac{x^2}{2t}}\,dx,$$

use the inequality

$$1 - \frac{x^2}{2t} < e^{-\frac{x^2}{2t}} < 1$$

to get the desired result.

2.7.3

$$\begin{aligned} E[R_t] &= \int_0^\infty x p_t(x)\,dx = \int_0^\infty \frac{1}{t} x^2 e^{-\frac{x^2}{2t}}\,dx \\ &= \frac{1}{2t}\int_0^\infty y^{1/2} e^{-\frac{y}{2t}}\,dy = \sqrt{2t}\int_0^\infty z^{\frac{3}{2}-1} e^{-z}\,dz \\ &= \sqrt{2t}\,\Gamma\big(\frac{3}{2}\big) = \frac{\sqrt{2\pi t}}{2}. \end{aligned}$$

Since $E[R_t^2] = E[W_1(t)^2 + W_2(t)^2] = 2t$, then

$$Var(R_t) = 2t - \frac{2\pi t}{4} = 2t\big(1 - \frac{\pi}{4}\big).$$

2.7.4

$$\begin{aligned} E[X_t] &= \frac{E[X_t]}{t} = \frac{\sqrt{2\pi t}}{2t} = \sqrt{\frac{\pi}{2t}} \to 0, \qquad t \to \infty; \\ Var(X_t) &= \frac{1}{t^2} Var(R_t) = \frac{2}{t}(1 - \frac{\pi}{4}) \to 0, \qquad t \to \infty. \end{aligned}$$

By Example 1.13.1 we get $X_t \to 0$, $t \to \infty$ in mean square.

2.8.2

$$\begin{aligned} P(N_t - N_s = 1) &= \lambda(t-s)e^{-\lambda(t-s)} \\ &= \lambda(t-s)\big(1 - \lambda(t-s) + o(t-s)\big) \\ &= \lambda(t-s) + o(t-s). \end{aligned}$$

$$\begin{aligned} P(N_t - N_s \geq 1) &= 1 - P(N_t - N_s = 0) + P(N_t - N_s = 1) \\ &= 1 - e^{-\lambda(t-s)} - \lambda(t-s)e^{-\lambda(t-s)} \\ &= 1 - \Big(1 - \lambda(t-s) + o(t-s)\Big) \\ &\quad -\lambda(t-s)\Big(1 - \lambda(t-s) + o(t-s)\Big) \\ &= \lambda^2(t-s)^2 = o(t-s). \end{aligned}$$

2.8.6 Write first as

$$\begin{aligned} N_t^2 &= N_t(N_t - N_s) + N_t N_s \\ &= (N_t - N_s)^2 + N_s(N_t - N_s) + (N_t - \lambda t)N_s + \lambda t N_s, \end{aligned}$$

then

$$\begin{aligned} E[N_t^2|\mathcal{F}_s] &= E[(N_t - N_s)^2|\mathcal{F}_s] + N_s E[N_t - N_s|\mathcal{F}_s] + E[N_t - \lambda t|\mathcal{F}_s]N_s + \lambda t N_s \\ &= E[(N_t - N_s)^2] + N_s E[N_t - N_s] + E[N_t - \lambda t]N_s + \lambda t N_s \\ &= \lambda(t-s) + \lambda^2(t-s)^2 + \lambda t N_s + N_s^2 - \lambda s N_s + \lambda t N_s \\ &= \lambda(t-s) + \lambda^2(t-s)^2 + 2\lambda(t-s)N_s + N_s^2 \\ &= \lambda(t-s) + [N_s + \lambda(t-s)]^2. \end{aligned}$$

Hence $E[N_t^2|\mathcal{F}_s] \neq N_s^2$ and hence the process N_t^2 is not an $\mathcal{F}_s$-martingale.

2.8.7 (a)

$$\begin{aligned} m_{N_t}(x) &= E[e^{xN_t}] = \sum_{k\geq 0} e^{xh} P(N_t = k) \\ &= \sum_{k\geq 0} e^{xk} e^{-\lambda t} \frac{\lambda^k t^k}{k!} \\ &= e^{-\lambda t} e^{\lambda t e^x} = e^{\lambda t(e^x - 1)}. \end{aligned}$$

(b) $E[N_t^2] = m''_{N_t}(0) = \lambda^2 t^2 + \lambda t$. Similarly for the other relations.

2.8.8 $E[X_t] = E[e^{N_t}] = m_{N_t}(1) = e^{\lambda t(e-1)}$.

2.8.9 (a) Since $e^{xM_t} = e^{x(N_t - \lambda t)} = e^{-\lambda t x} e^{xN_t}$, the moment generating function is

$$\begin{aligned} m_{M_t}(x) &= E[e^{xM_t}] = e^{-\lambda t x} E[e^{xN_t}] \\ &= e^{-\lambda t x} e^{\lambda t(e^x - 1)} = e^{\lambda t(e^x - x - 1)}. \end{aligned}$$

(b) For instance

$$E[M_t^3] = m'''_{M_t}(0) = \lambda t.$$

Since M_t is a stationary process, $E[(M_t - M_s)^3] = \lambda(t - s)$.

2.8.10

$$\begin{aligned} Var[(M_t - M_s)^2] &= E[(M_t - M_s)^4] - E[(M_t - M_s)^2]^2 \\ &= \lambda(t-s) + 3\lambda^2(t-s)^2 - \lambda^2(t-s)^2 \\ &= \lambda(t-s) + 2\lambda^2(t-s)^2. \end{aligned}$$

2.11.3 (a) $\quad E[U_t] = E\Big[\int_0^t N_u\, du\Big] = \int_0^t E[N_u]\, du = \int_0^t \lambda u\, du = \frac{\lambda t^2}{2}$.

(b) $\quad E\Big[\sum_{k=1}^{N_t} S_k\Big] = E[tN_t - U_t] = t\lambda t - \dfrac{\lambda t^2}{2} = \dfrac{\lambda t^2}{2}$.

2.11.5 The proof cannot go through because a product between a constant and a Poisson process is not a Poisson process.

2.11.6 Let $p_X(x)$ be the probability density of X. If $p(x, y)$ is the joint probability density of X and Y, then $p_X(x) = \sum_y p(x, y)$. We have

$$\begin{aligned} \sum_{y\geq 0} E[X|Y = y] P(Y = y) &= \sum_{y\geq 0} \int x p_{X|Y=y}(x|y) P(Y = y)\, dx \\ &= \sum_{y\geq 0} \int \frac{xp(x,y)}{P(Y = y)} P(Y = y)\, dx = \int x \sum_{y\geq 0} p(x,y)\, dx \\ &= \int x p_X(x)\, dx = E[X]. \end{aligned}$$

2.11.8 (a) Since T_k has an exponential distribution with parameter λ

$$E[e^{-\sigma T_k}] \;=\; \int_0^\infty e^{-\sigma x}\lambda e^{-\lambda x}\,dx = \frac{\lambda}{\lambda+\sigma}.$$

(b) We have

$$\begin{aligned} U_t &= T_2 + 2T_3 + 3T_4 + \cdots + (n-2)T_{n-1} + (n-1)T_n + (t - S_n)n \\ &= T_2 + 2T_3 + 3T_4 + \cdots + (n-2)T_{n-1} + (n-1)T_n + nt - n(T_1 + T_2 + \cdots + T_n) \\ &= nt - [nT_1 + (n-1)T_2 + \cdots + T_n]. \end{aligned}$$

(c) Using that the arrival times S_k, $k = 1, 2, \ldots n$, have the same distribution as the order statistics $U_{(k)}$ corresponding to n independent random variables uniformly distributed on the interval $(0, t)$, we get

$$\begin{aligned} E\Big[e^{-\sigma U_t}\Big|N_t = n\Big] &= E[e^{-\sigma(tN_t - \sum_{k=1}^{N_t} S_k)}|N_t = n] \\ &= e^{-n\sigma t}E[e^{\sigma\sum_{i=1}^n U_{(i)}}] = e^{-n\sigma t}E[e^{\sigma\sum_{i=1}^n U_i}] \\ &= e^{-n\sigma t}E[e^{\sigma U_1}]\cdots E[e^{\sigma U_n}] \\ &= e^{-n\sigma t}\frac{1}{t}\int_0^t e^{\sigma x_1}\,dx_1 \cdots \frac{1}{t}\int_0^t e^{\sigma x_n}\,dx_n \\ &= \frac{(1-e^{-\sigma t})^n}{\sigma^n t^n}. \end{aligned}$$

(d) Using Exercise **2.11.6** we have

$$\begin{aligned} E\Big[e^{-\sigma U_t}\Big] &= \sum_{n\ge 0} P(N_t = n)E\Big[e^{-\sigma U_t}\Big|N_t = n\Big] \\ &= \sum_{n\ge 0}\frac{e^{-\lambda}t^n\lambda^n}{n!}\frac{(1-e^{-\sigma t})^n}{\sigma^n t^n} \\ &= e^{\lambda(1-e^{-\sigma t})/\sigma - \lambda}. \end{aligned}$$

3.11.6 (a) $dt\,dN_t = dt(dM_t + \lambda dt) = dt\,dM_t + \lambda dt^2 = 0$

(b) $dW_t\,dN_t = dW_t(dM_t + \lambda dt) = dW_t dM_t + \lambda dW_t dt = 0$

2.12.6 Use Doob's inequality for the submartingales W_t^2 and $|W_t|$, and use that $E[W_t^2] = t$ and $E[|W_t|] = \sqrt{2t/\pi}$, see Exercise 2.1.16 (a).

2.12.7 Divide by t in the inequality from Exercise 2.12.6 part (b).

2.12.10 Let $\sigma = n$ and $\tau = n + 1$. Then

$$E\Big[\sup_{n\le t\le n=1}\Big(\frac{N_t}{t} - \lambda\Big)^2\Big] \le \frac{4\lambda(n+1)}{n^2}.$$

The result follows by taking $n \to \infty$ in the sequence of inequalities

$$0 \leq E\Big[\Big(\frac{N_t}{t} - \lambda\Big)^2\Big] \leq E\Big[\sup_{n \leq t \leq n=1} \Big(\frac{N_t}{t} - \lambda\Big)^2\Big] \leq \frac{4\lambda(n+1)}{n^2}.$$

Chapter 3

3.1.2 We have

$$\{\omega; \tau(\omega) \leq t\} = \begin{cases} \Omega, & \text{if } c \leq t \\ \emptyset, & \text{if } c > t \end{cases}$$

and use that $\emptyset, \Omega \in \mathcal{F}_t$.

3.1.3 First we note that

$$\{\omega; \tau(\omega) < t\} = \bigcup_{0<s<t} \{\omega; |W_s(\omega)| > K\}. \tag{18.0.1}$$

This can be shown by double inclusion. Let $A_s = \{\omega; |W_s(\omega)| > K\}$.
"$\subset$" Let $\omega \in \{\omega; \tau(\omega) < t\}$, so $\inf\{s > 0; |W_s(\omega)| > K\} < t$. Then exists $\tau < u < t$ such that $|W_u(\omega)| > K$, and hence $\omega \in A_u$.
"$\supset$" Let $\omega \in \bigcup_{0<s<t}\{\omega; |W_s(\omega)| > K\}$. Then there is $0 < s < t$ such that $|W_s(\omega)| > K$. This implies $\tau(\omega) < s$ and since $s < t$ it follows that $\tau(\omega) < t$.
Since W_t is continuous, then (18.0.1) can be also written as

$$\{\omega; \tau(\omega) < t\} = \bigcup_{0<r<t, r \in \mathbb{Q}} \{\omega; |W_r(\omega)| > K\},$$

which implies $\{\omega; \tau(\omega) < t\} \in \mathcal{F}_t$ since $\{\omega; |W_r(\omega)| > K\} \in \mathcal{F}_t$, for $0 < r < t$. Next we shall show that $P(\tau < \infty) = 1$.

$$\begin{aligned} P(\{\omega; \tau(\omega) < \infty\}) &= P\Big(\bigcup_{0<s}\{\omega; |W_s(\omega)| > K\}\Big) > P(\{\omega; |W_s(\omega)| > K\}) \\ &= 1 - \int_{|x|<K} \frac{1}{\sqrt{2\pi s}} e^{-\frac{y^2}{2s}}\, dy > 1 - \frac{2K}{\sqrt{2\pi s}} \to 1, \quad s \to \infty. \end{aligned}$$

Hence τ is a stopping time.

3.1.4 Let $K_m = [a + \frac{1}{m}, b - \frac{1}{m}]$. We can write

$$\{\omega; \tau \leq t\} = \bigcap_{m \geq 1} \bigcup_{r<t, r \in \mathbb{Q}} \{\omega; X_r \notin K_m\} \in \mathcal{F}_t,$$

since $\{\omega; X_r \notin K_m\} = \{\omega; X_r \in \overline{K}_m\} \in \mathcal{F}_r \subset \mathcal{F}_t$.

3.1.8 (a) We have $\{\omega; c\tau \leq t\} = \{\omega; \tau \leq t/c\} \in \mathcal{F}_{t/c} \subset \mathcal{F}_t$. And $P(c\tau < \infty) = P(\tau < \infty) = 1$.

(b) $\{\omega; f(\tau) \le t\} = \{\omega; \tau \le f^{-1}(t)\} = \mathcal{F}_{f^{-1}(t)} \subset \mathcal{F}_t$, since $f^{-1}(t) \le t$. If f is bounded, then it is obvious that $P(f(\tau) < \infty) = 1$. If $\lim_{t\to\infty} f(t) = \infty$, then $P(f(\tau) < \infty) = P\big(\tau < f^{-1}(\infty)\big) = P(\tau < \infty) = 1$.

(c) Apply (b) with $f(x) = e^x$.

3.1.10 If let $G(n) = \{x; |x - a| < \frac{1}{n}\}$, then $\{a\} = \bigcap_{n\ge1} G(n)$. Then $\tau_n = \inf\{t \ge 0; W_t \in G(n)\}$ are stopping times. Since $\sup_n \tau_n = \tau$, then τ is a stopping time.

3.2.3 The relation is proved by verifying two cases:
(i) If $\omega \in \{\omega; \tau > t\}$ then $(\tau \wedge t)(\omega) = t$ and the relation becomes

$$M_\tau(\omega) = M_t(\omega) + M_\tau(\omega) - M_t(\omega).$$

(ii) If $\omega \in \{\omega; \tau \le t\}$ then $(\tau \wedge t)(\omega) = \tau(\omega)$ and the relation is equivalent with the obvious relation

$$M_\tau = M_\tau.$$

3.2.5 Taking the expectation in $E[M_\tau|\mathcal{F}_\sigma] = M_\sigma$ yields $E[M_\tau] = E[M_\sigma]$, and then make $\sigma = 0$.

3.3.9 Since $M_t = W_T^2 - t$ is a martingale with $E[M_t] = 0$, by the Optional Stopping Theorem we get $E[M_{\tau_a}] = E[M_0] = 0$, so $E[W_{\tau_a}^2 - \tau_a] = 0$, from where $E[\tau_a] = E[W_{\tau_a}^2] = a^2$, since $W_{\tau_a} = a$.

3.3.10 (a)

$$\begin{aligned}
F(a) = P(X_t \le a) &= 1 - P(X_t > a) = 1 - P(\max_{0\le s\le t} W_s > a)\\
&= 1 - P(T_a \le t) = 1 - \frac{2}{\sqrt{2\pi}} \int_{|a|/\sqrt{t}}^{\infty} e^{-y^2/2}\, dy\\
&= \frac{2}{\sqrt{2\pi}} \int_0^\infty e^{-y^2/2}\, dy - \frac{2}{\sqrt{2\pi}} \int_{|a|/\sqrt{t}}^{\infty} e^{-y^2/2}\, dy\\
&= \frac{2}{\sqrt{2\pi}} \int_0^{|a|/\sqrt{t}} e^{-y^2/2}\, dy.
\end{aligned}$$

(b) The density function is $p(a) = F'(a) = \frac{2}{\sqrt{2\pi t}} e^{-a^2/(2t)}$, $a > 0$. Then

$$\begin{aligned}
E[X_t] &= \int_0^\infty x p(x)\, dx = \frac{2}{\sqrt{2\pi t}} \int_0^\infty x e^{-x^2/(2t)}\, dy = \sqrt{\frac{2t}{\pi}}\\
E[X_t^2] &= \frac{2}{\sqrt{2\pi t}} \int_0^\infty x^2 e^{-x^2/(2t)}\, dx = \frac{4t}{\sqrt{\pi}} \int_0^\infty y^2 e^{-y^2}\, dy\\
&= \frac{2}{\sqrt{2\pi t}} \frac{1}{2} \int_0^\infty u^{1/2} e^{-u}\, du = \frac{1}{\sqrt{2\pi t}} \Gamma(3/2) = t\\
Var(X_t) &= E[X_t^2] - E[X_t]^2 = t\Big(1 - \frac{2}{\pi}\Big).
\end{aligned}$$

3.3.16 It is recurrent since $P(\exists t > 0 : a < W_t < b) = 1$.

3.4.4 Since

$$P(W_t > 0; t_1 \le t \le t_2) = \frac{1}{2} P(W_t \neq 0; t_1 \le t \le t_2) = \frac{1}{\pi} \arcsin \sqrt{\frac{t_1}{t_2}},$$

using the independence

$$P(W_t^1 > 0, W_t^2) = P(W_t^1 > 0) P(W_t^2 > 0) = \frac{1}{\pi^2} \Big(\arcsin \sqrt{\frac{t_1}{t_2}} \Big)^2.$$

The probability for $W_t = (W_t^1, W_t^2)$ to be in one of the quadrants is $\frac{4}{\pi^2} \Big(\arcsin \sqrt{\frac{t_1}{t_2}} \Big)^2$.

3.5.2 (a) We have

$$P(X_t \text{ goes up to } \alpha) = P(X_t \text{ goes up to } \alpha \text{ before down to } -\infty) = \lim_{\beta \to \infty} \frac{e^{2\mu\beta} - 1}{e^{2\mu\beta} - e^{-2\mu\alpha}} = 1.$$

3.5.3

$$P(X_t \text{ never hits } -\beta) = P(X_t \text{ goes up to } \infty \text{ before down to } -\beta) = \lim_{\alpha \to \infty} \frac{e^{2\mu\beta} - 1}{e^{2\mu\beta} - e^{-2\mu\alpha}}.$$

3.5.4 (a) Use that $E[X_T] = \alpha p_\alpha - \beta(1 - p_\alpha)$; (b) $E[X_T^2] = \alpha^2 p_\alpha + \beta^2 (1 - p_\alpha)$, with $p_\alpha = \frac{e^{2\mu\beta} - 1}{e^{2\mu\beta} - e^{-2\mu\alpha}}$; (c) Use $Var(T) = E[T^2] - E[T]^2$.

3.5.7 Since $M_t = W_t^2 - t$ is a martingale, with $E[M_t] = 0$, by the Optional Stopping Theorem we get $E[W_T^2 - T] = 0$. Using $W_T = X_T - \mu T$ yields

$$E[X_T^2 - 2\mu T X_T + \mu^2 T^2] = E[T].$$

Then

$$E[T^2] = \frac{E[T](1 + 2\mu E[X_T]) - E[X_T^2]}{\mu^2}.$$

Substitute $E[X_T]$ and $E[X_T^2]$ from Exercise **3.5.4** and $E[T]$ from Proposition 3.5.5.

3.6.11 See the proof of Proposition 3.6.3.

3.6.12

$$E[T] = -\frac{d}{ds} E[e^{-sT}]_{|s=0} = \frac{1}{\mu}(\alpha p_\alpha + \beta(1 - p_\beta)) = \frac{E[X_T]}{\mu}.$$

3.6.13 (b) Applying the Optional Stopping Theorem

$$\begin{aligned} E[e^{cM_T - \lambda T(e^c - c - 1)}] = E[X_0] &= 1 \\ E[e^{ca - \lambda T f(c)}] &= 1 \\ E[e^{-\lambda T f(c)}] &= e^{-ac}. \end{aligned}$$

Let $s = f(c)$, so $c = \varphi(s)$. Then $E[e^{-\lambda s T}] = e^{-a\varphi(s)}$.
(c) Differentiating and taking $s = 0$ yields

$$\begin{aligned} -\lambda E[T] &= -ae^{-a\varphi(0)}\varphi'(0) \\ &= -a\frac{1}{f'(0)} = -\infty, \end{aligned}$$

so $E[T] = \infty$.
(d) The inverse Laplace transform $\mathcal{L}^{-1}\Big(e^{-a\varphi(s)}\Big)$ cannot be represented by elementary functions.

3.8.8 Use that if W_t is a Brownian motion then also $tW_{1/t}$ is a Brownian motion.

3.8.11 Use $E[(|X_t| - 0)^2] = E[|X_t|^2] = E[(X_t - 0)^2]$.

3.8.15 Let $a_n = \ln b_n$. Then $G_n = e^{\ln G_n} = e^{\frac{a_1+\cdots+a_n}{n}} \to e^{\ln L} = L$.

3.9.3 (a) $L = 1$. (b) A computation shows

$$\begin{aligned} E[(X_t - 1)^2] &= E[X_t^2 - 2X_t + 1] = E[X_t^2] - 2E[X_t] + 1 \\ &= Var(X_t) + (E[X_t] - 1)^2. \end{aligned}$$

(c) Since $E[X_t] = 1$, we have $E[(X_t - 1)^2] = Var(X_t)$. Since

$$Var(X_t) = e^{-t}E[e^{W_t}] = e^{-t}(e^{2t} - e^t) = e^t - 1,$$

then $E[(X_t - 1)^2]$ does not tend to 0 as $t \to \infty$.

3.11.3 (a) $E[(dW_t)^2 - dt^2] = E[(dW_t)^2] - dt^2 = 0$.

$$(b) \qquad \begin{aligned} Var((dW_t)^2 - dt) &= E[(dW_t^2 - dt)^2] = E[(dW_t)^4 - 2dtdW_t + dt^2] \\ &= 3st^2 - 2dt \cdot 0 + dt^2 = 4dt^2. \end{aligned}$$

Chapter 4

4.2.3 (a) Use either the definition or the moment generation function to show that $E[W_t^4] = 3t^2$. Using stationarity, $E[(W_t - W_s)^4] = E[W_{t-s}^4] = 3(t-s)^2$.

4.4.1 (a) $E[\int_0^T dW_t] = E[W_T] = 0$.
(b) $E[\int_0^T W_t\, dW_t] = E[\frac{1}{2}W_T^2 - \frac{1}{2}T] = 0$.
(c) $Var(\int_0^T W_t\, dW_t) = E[(\int_0^T W_t\, dW_t)^2] = E[\frac{1}{4}W_T^2 + \frac{1}{4}T^2 - \frac{1}{2}TW_T^2] = \frac{T^2}{2}$.

4.6.2 $X \sim N(0, \int_1^T \frac{1}{t}dt) = N(0, \ln T)$.

4.6.3 $Y \sim N(0, \int_1^T t dt) = N\Big(0, \frac{1}{2}(T^2-1)\Big)$

4.6.4 Normally distributed with zero mean and variance $\displaystyle\int_0^t e^{2(t-s)}\,ds = \frac{1}{2}(e^{2t}-1)$.

4.6.5 Using the property of Wiener integrals, both integrals have zero mean and variance $\dfrac{7t^3}{3}$.

4.6.6 The mean is zero and the variance is $t/3 \to 0$ as $t \to 0$.

4.6.7 Since it is a Wiener integral, X_t is normally distributed with zero mean and variance

$$\int_0^t \Big(a + \frac{bu}{t}\Big)^2 du = (a^2 + \frac{b^2}{3} + ab)t.$$

Hence $a^2 + \frac{b^2}{3} + ab = 1$.

4.6.8 Since both W_t and $\int_0^t f(s)\,dW_s$ have the mean equal to zero,

$$\begin{aligned} Cov\Big(W_t, \int_0^t f(s)\,dW_s\Big) &= E[W_t, \int_0^t f(s)\,dW_s] = E[\int_0^t dW_s \int_0^t f(s)\,dW_s] \\ &= E[\int_0^t f(u)\,ds] = \int_0^t f(u)\,ds. \end{aligned}$$

The general result is

$$Cov\Big(W_t, \int_0^t f(s)\,dW_s\Big) = \int_0^t f(s)\,ds.$$

Choosing $f(u) = u^n$ yields the desired identity.

4.8.6 Apply the expectation to

$$\Big(\sum_{k=1}^{N_t} f(S_k)\Big)^2 = \sum_{k=1}^{N_t} f^2(S_k) + 2\sum_{k\neq j}^{N_t} f(S_k)f(S_j).$$

4.9.1

$$\begin{aligned} E\Big[\int_0^T e^{ks}\,dN_s\Big] &= \frac{\lambda}{k}(e^{kT}-1) \\ Var\Big(\int_0^T e^{ks}\,dN_s\Big) &= \frac{\lambda}{2k}(e^{2kT}-1). \end{aligned}$$

Chapter 5

5.2.1 Let $X_t = \int_0^t e^{W_u}\,du$. Then

$$dG_t = d(\frac{X_t}{t}) = \frac{tdX_t - X_t dt}{t^2} = \frac{te^{W_t}dt - X_t dt}{t^2} = \frac{1}{t}\Big(e^{W_t} - G_t\Big)dt.$$

5.3.3

(a) $e^{W_t}(1+\frac{1}{2}W_t)dt + e^{W_t}(1+W_t)dW_t$;

(b) $(6W_t + 10e^{5W_t})dW_t + (3+25e^{5W_t})dt$;

(c) $2e^{t+W_t^2}(1+W_t^2)dt + 2e^{t+W_t^2}W_t dW_t$;

(d) $n(t+W_t)^{n-2}\Big((t+W_t+\frac{n-1}{2})dt + (t+W_t)dW_t\Big)$;

(e) $\displaystyle \frac{1}{t}\left(W_t - \frac{1}{t}\int_0^t W_u\,du\right)dt$;

(f) $\displaystyle \frac{1}{t^\alpha}\left(e^{W_t} - \frac{\alpha}{t}\int_0^t e^{W_u}\,du\right)dt.$

5.3.4

$$d(tW_t^2) = td(W_t^2) + W_t^2 dt = t(2W_t dW_t + dt) + W_t^2 dt = (t+W_t^2)dt + 2tW_t dW_t.$$

5.3.5 (a) $tdW_t + W_t dt$;

(b) $e^t(W_t dt + dW_t)$;

(c) $(2-t/2)t\cos W_t\,dt - t^2\sin W_t\,dW_t$;

(d) $(\sin t + W_t^2\cos t)dt + 2\sin t\,W_t\,dW_t$;

5.3.7 It follows from (5.3.9).

5.3.8 Take the conditional expectation in

$$M_t^2 = M_s^2 + 2\int_s^t M_{u-}\,dM_u + N_t - N_s$$

and obtain

$$\begin{aligned}
E[M_t^2|\mathcal{F}_s] &= M_s^2 + 2E[\int_s^t M_{u-}\,dM_u|\mathcal{F}_s] + E[N_t|\mathcal{F}_s] - N_s \\
&= M_s^2 + E[M_t + \lambda t|\mathcal{F}_s] - N_s \\
&= M_s^2 + M_s + \lambda t - N_s \\
&= M_s^2 + \lambda(t-s).
\end{aligned}$$

5.3.9 Integrating in (5.3.10) yields

$$F_t = F_s + \int_s^t \frac{\partial f}{\partial x}\,dW_t^1 + \int_s^t \frac{\partial f}{\partial y}\,dW_t^2.$$

One can check that $E[F_t|\mathcal{F}_s] = F_s$.

5.3.10 (a) $\quad dF_t = 2W_t^1 dW_t^1 + 2W_t^2 dW_t^2 + 2dt$; (b) $\quad dF_t = \dfrac{2W_t^1 dW_t^1 + 2W_t^2 dW_t^2}{(W_t^1)^2 + (W_t^2)^2}$.

5.3.11 Consider the function $f(x,y) = \sqrt{x^2+y^2}$. Since $\dfrac{\partial f}{\partial x} = \dfrac{x}{\sqrt{x^2+y^2}}$, $\dfrac{\partial f}{\partial y} = \dfrac{y}{\sqrt{x^2+y^2}}$, $\Delta f = \dfrac{1}{2\sqrt{x^2+y^2}}$, we get

$$dR_t = \frac{\partial f}{\partial x} dW_t^1 + \frac{\partial f}{\partial y} dW_t^2 + \Delta f dt = \frac{W_t^1}{R_t} dW_t^1 + \frac{W_t^2}{R_t} dW_t^2 + \frac{1}{2R_t} dt.$$

Chapter 6

6.2.1 (*a*) Use integration formula with $g(x) = \tan^{-1}(x)$.

$$\int_0^T \frac{1}{1+W_t^2}\, dW_t = \int_0^T (\tan^{-1})'(W_t)\, dW_t = \tan^{-1} W_T + \frac{1}{2}\int_0^T \frac{2W_t}{(1+W_t)^2}\, dt.$$

(*b*) Use $E\Big[\displaystyle\int_0^T \frac{1}{1+W_t^2}\, dW_t\Big] = 0$.

(*c*) Use Calculus to find minima and maxima of the function $\varphi(x) = \dfrac{x}{(1+x^2)^2}$.

6.2.2 (*a*) Use integration by parts with $g(x) = e^x$ and get

$$\int_0^T e^{W_t}\, dW_t = e^{W_T} - 1 - \frac{1}{2}\int_0^T e^{W_t}\, dt.$$

(*b*) Applying the expectation we obtain

$$E[e^{W_T}] = 1 + \frac{1}{2}\int_0^T E[e^{W_t}]\, dt.$$

If let $\phi(T) = E[e^{W_T}]$, then ϕ satisfies the integral equation

$$\phi(T) = 1 + \frac{1}{2}\int_0^T \phi(t)\, dt.$$

Differentiating yields the ODE $\phi'(T) = \frac{1}{2}\phi(T)$, with $\phi(0) = 1$. Solving yields $\phi(T) = e^{T/2}$.

6.2.3 (*a*) Apply integration by parts with $g(x) = (x-1)e^x$ to get

$$\int_0^T W_t e^{W_t}\, dW_t = \int_0^T g'(W_t)\, dW_t = g(W_T) - g(0) - \frac{1}{2}\int_0^T g''(W_t)\, dt.$$

(*b*) Applying the expectation yields

$$\begin{aligned} E[W_T e^{W_T}] &= E[e^{W_T}] - 1 + \frac{1}{2}\int_0^T \Big[E[e^{W_t}] + E[W_t e^{W_t}] \Big]\, dt \\ &= e^{T/2} - 1 + \frac{1}{2}\int_0^T \Big(e^{t/2} + E[W_t e^{W_t}] \Big)\, dt. \end{aligned}$$

Then $\phi(T) = E[W_T e^{W_T}]$ satisfies the ODE $\phi'(T) - \phi(T) = e^{T/2}$ with $\phi(0) = 0$.

6.2.4 (a) Use integration by parts with $g(x) = \ln(1 + x^2)$.
(e) Since $\ln(1 + T) \leq T$, the upper bound obtained in (e) is better than the one in (d), without contradicting it.

6.3.1 By straightforward computation.

6.3.2 By computation.

6.3.13 (a) $\frac{e}{\sqrt{2}}\sin(\sqrt{2}W_1)$; (b) $\frac{1}{2}e^{2T}\sin(2W_3)$; (c) $\frac{1}{\sqrt{2}}(e^{\sqrt{2}W_4 - 4} - 1)$.

6.3.14 Apply Ito's formula to get

$$\begin{aligned} d\varphi(t, W_t) &= (\partial_t \varphi(t, W_t) + \frac{1}{2}\partial_x^2 \varphi(t, W_t))dt + \partial_x \varphi(t, W_t) dW_t \\ &= G(t)dt + f(t, W_t)\, dW_t. \end{aligned}$$

Integrating between a and b yields

$$\varphi(t, W_t)\big|_a^b = \int_a^b G(t)\, dt + \int_a^b f(t, W_t)\, dW_t.$$

Chapter 7

7.2.1 Integrating yields $X_t = X_0 + \int_0^t (2X_s + e^{2s})\, ds + \int_0^t b\, dW_s$. Taking the expectation we get

$$E[X_t] = X_0 + \int_0^t (2E[X_s] + e^{2s})\, ds.$$

Differentiating we obtain $f'(t) = 2f(t) + e^{2t}$, where $f(t) = E[X_t]$, with $f(0) = X_0$. Multiplying by the integrating factor e^{-2t} yields $(e^{-2t} f(t))' = 1$. Integrating yields $f(t) = e^{2t}(t + X_0)$.

7.2.4 (a) Using product rule and Ito's formula, we get

$$d(W_t^2 e^{W_t}) = e^{W_t}(1 + 2W_t + \frac{1}{2}W_t^2)dt + e^{W_t}(2W_t + W_t^2)dW_t.$$

Integrating and taking expectations yields

$$E[W_t^2 e^{W_t}] = \int_0^t \Big(E[e^{W_s}] + 2E[W_s e^{W_s}] + \frac{1}{2}E[W_s^2 e^{W_s}] \Big)\, ds.$$

Since $E[e^{W_s}] = e^{t/2}$, $E[W_s e^{W_s}] = te^{t/2}$, if let $f(t) = E[W_t^2 e^{W_t}]$, we get by differentiation

$$f'(t) = e^{t/2} + 2te^{t/2} + \frac{1}{2}f(t), \qquad f(0) = 0.$$

Multiplying by the integrating factor $e^{-t/2}$ yields $(f(t)e^{-t/2})' = 1 + 2t$. Integrating yields the solution $f(t) = t(1+t)e^{t/2}$. (b) Similar method.

7.2.5 (a) Using Exercise 2.1.18

$$\begin{aligned} E[W_t^4 - 3t^2|\mathcal{F}_t] &= E[W_t^4|\mathcal{F}_t] - 3t^2 \\ &= 3(t-s)^2 + 6(t-s)W_s^2 + W_s^4 - 3t^2 \\ &= (W_s^4 - 3s^2) + 6s^2 - 6ts + 6(t-s)W_s^2 \neq W_s^4 - 3s^2 \end{aligned}$$

Hence $W_t^4 - 3t^2$ is not a martingale. (b)

$$E[W_t^3|\mathcal{F}_s] \;=\; W_s^3 + 3(t-s)W_s \neq W_s^3,$$

and hence W_t^3 is not a martingale.

7.2.6 (a) Similar method as in Example 7.2.4.
(b) Applying Ito's formula

$$d(\cos(\sigma W_t)) = -\sigma \sin(\sigma W_t)dW_t - \frac{1}{2}\sigma^2 \cos(\sigma W_t)dt.$$

Let $f(t) = E[\cos(\sigma W_t)]$. Then $f'(t) = -\frac{\sigma^2}{2}f(t)$, $f(0) = 1$. The solution is $f(t) = e^{-\frac{\sigma^2}{2}t}$.
(c) Since $\sin(t+\sigma W_t) = \sin t \cos(\sigma W_t) + \cos t \sin(\sigma W_t)$, taking the expectation and using (a) and (b) yields

$$E[\sin(t + \sigma W_t)] = \sin t E[\cos(\sigma W_t)] = e^{-\frac{\sigma^2}{2}t} \sin t.$$

(d) Similarly starting from $\cos(t + \sigma W_t) = \cos t \cos(\sigma W_t) - \sin t \sin(\sigma W_t)$.

7.2.7 From Exercise 7.2.6 (b) we have $E[\cos(W_t)] = e^{-t/2}$. From the definition of expectation

$$E[\cos(W_t)] = \int_{-\infty}^{\infty} \cos x \frac{1}{\sqrt{2\pi t}} e^{-\frac{x^2}{2t}}\, dx.$$

Then choose $t = 1/2$ and $t = 1$ to get (a) and (b), respectively.

7.2.8 (a) Using a standard method involving Ito's formula we can get $E(W_t e^{bW_t}) = bte^{b^2t/2}$. Let $a = 1/(2t)$. We can write

$$\begin{aligned} \int xe^{-ax^2+bx}\, dx &= \sqrt{2\pi t} \int xe^{bx} \frac{1}{\sqrt{2\pi t}} e^{-\frac{x^2}{2t}}\, dx \\ &= \sqrt{2\pi t} E(W_t e^{bW_t}) = \sqrt{2\pi t}\, bte^{b^2t/2} = \sqrt{\frac{\pi}{a}} \Big(\frac{b}{2a}\Big) e^{b^2/(4a)}. \end{aligned}$$

The same method for (b) and (c).

7.2.9 (a) We have

$$\begin{aligned} E[\cos(tW_t)] &= E\big[\sum_{n\geq 0}(-1)^n \frac{W_t^{2n}t^{2n}}{(2n)!}\big] = \sum_{n\geq 0}(-1)^n \frac{E[W_t^{2n}]t^{2n}}{(2n)!} \\ &= \sum_{n\geq 0}(-1)^n \frac{t^{2n}}{(2n)!}\frac{(2n)!t^n}{2^n n!} = \sum_{n\geq 0}(-1)^n \frac{t^{3n}}{2^n n!} \\ &= e^{-t^3/2}. \end{aligned}$$

(b) Similar computation using $E[W_t^{2n+1}] = 0$.

7.3.2 (a) $X_t = 1 + \sin t - \int_0^t \sin s\, dW_s$, $E[X_t] = 1 + \sin t$, $Var[X_t] = \int_0^t (\sin s)^2\, ds = \frac{t}{2} - \frac{1}{4}\sin(2t)$;
(b) $X_t = e^t - 1 + \int_0^t \sqrt{s}\, dW_s$, $E[X_t] = e^t - 1$, $Var[X_t] = \frac{t^2}{2}$;
(c) $X_t = 1 + \frac{1}{2}\ln(1+t^2) + \int_0^t s^{3/2}\, dW_s$, $E[X_t] = 1 + \frac{1}{2}\ln(1+t^2)$, $Var(X_t) = \frac{t^4}{4}$.

7.3.4

$$\begin{aligned} X_t &= 2(1 - e^{-t/2}) + \int_0^t e^{-\frac{s}{2}+W_s}\, dW_s = 1 - 2e^{-t/2} + e^{-t/2+W_t} \\ &= 1 + e^{t/2}(e^{W_t} - 2); \end{aligned}$$

Its distribution function is given by

$$\begin{aligned} F(y) &= P(X_t \leq y) = P(1 + e^{t/2}(e^{W_t} - 2) \leq y) = P\big(W_t \leq \ln(2 + (y-1)e^{t/2})\big) \\ &= \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{\ln(2+(y-1)e^{t/2})} e^{-\frac{x^2}{2t}}\, dx. \end{aligned}$$

$E[X_t] = 2(1 - e^{-t/2})$, $Var(X_t) = Var(e^{-t/2}e^{W_t}) = e^{-t}Var(e^{W_t}) = e^t - 1$.

7.5.1 (a)

$$\begin{aligned} dX_t &= (2W_t dW_t + dt) + W_t dt + t dW_t \\ &= d(W_t^2) + d(tW_t) = d(tW_t + W_t^2) \end{aligned}$$

so $X_t = tW_t + W_t^2$. (b) We have

$$\begin{aligned} dX_t &= (2t - \frac{1}{t^2}W_t)dt + \frac{1}{t}dW_t \\ &= 2t dt + d(\frac{1}{t}W_t) = d(t^2 + \frac{1}{t}W_t), \end{aligned}$$

so $X_t = t^2 + \frac{1}{t}W_t - 1 - W_1$.
(c) $dX_t = \dfrac{1}{2}e^{t/2}W_t dt + e^{t/2}dW_t = d(e^{t/2}W_t)$, so $X_t = e^{t/2}W_t$. (d) We have

$$\begin{aligned} dX_t &= t(2W_t dW_t + dt) - tdt + W_t^2 dt \\ &= td(W_t^2) + W_t^2 dt - \frac{1}{2}d(t^2) \\ &= d(tW_t^2 - \frac{t^2}{2}), \end{aligned}$$

so $X_t = tW_t - \frac{t^2}{2}$. (e) $dX_t = dt + d(\sqrt{t}W_t) = d(t + \sqrt{t}W_t)$, so $X_t = t + \sqrt{t}W_t - W_1$.

7.6.1 (a) $X_t = X_0 e^{4t} + \dfrac{1}{4}(1 - e^{4t}) + 2\displaystyle\int_0^t e^{4(t-s)}\, dW_s$;
(b) $X_t = X_0 e^{3t} + \dfrac{2}{3}(1 - e^{3t}) + e^{3t}W_t$;
(c) $X_t = e^t(X_0 + 1 + \dfrac{1}{2}W_t^2 - \dfrac{t}{2}) - 1$;
(d) $X_t = X_0 e^{4t} - \dfrac{t}{4} - \dfrac{1}{16}(1 - e^{4t}) + e^{4t}W_t$;
(e) $X_t = X_0 e^{t/2} - 2t - 4 + 5e^{t/2} - e^t \cos W_t$;
(f) $X_t = X_0 e^{-t} + e^{-t}W_t$.

7.8.1 (a) The integrating factor is $\rho_t = e^{-\int_0^t \alpha\, dW_s + \frac{1}{2}\int_0^t \alpha^2 ds} = e^{-\alpha W_t + \frac{\alpha^2}{2}t}$, which transforms the equation in the exact form $d(\rho_t X_t) = 0$. Then $\rho_t X_t = X_0$ and hence $X_t = X_0 e^{\alpha W_t - \frac{\alpha^2}{2}t}$.
(b) $\rho_t = e^{-\alpha W_t + \frac{\alpha^2}{2}t}$, $d(\rho_t X_t) = \rho_t X_t dt$, $dY_t = Y_t dt$, $Y_t = Y_0 e^t$, $\rho_t X_t = X_0 e^t$, $X_t = X_0 e^{(1-\frac{\alpha^2}{2})t + \alpha W_t}$.

7.8.2 $\sigma t\, dA_t = dX_t - \sigma A_t\, dt$, $E[A_t] = 0$, $Var(A_t) = E[A_t^2] = \frac{1}{t^2}\int_0^t E[X_s]^2\, ds = \frac{X_0^2}{t}$.

Chapter 8

8.1.2 $A = \frac{1}{2}\Delta = \frac{1}{2}\sum_{k=1}^n \partial_k^2$.

8.1.4

$$\begin{aligned} X_1(t) &= x_1^0 + W_1(t) \\ X_2(t) &= x_2^0 + \int_0^t X_1(s)\, dW_2(s) \\ &= x_2^0 + x_1^0 W_2(t) + \int_0^t W_1(s)\, dW_2(s). \end{aligned}$$

8.4.2 $E[\tau]$ is maximum if $b - x_0 = x_0 - a$, i.e. when $x_0 = (a+b)/2$. The maximum value is $(b-a)^2/4$.

8.5.1 Let $\tau_k = \min(k, \tau) \nearrow \tau$ and $k \to \infty$. Apply Dynkin's formula for τ_k to show that

$$E[\tau_k] \leq \frac{1}{n}(R^2 - |a|^2),$$

and take $k \to \infty$.

8.5.3 x^0 and x^{2-n}.

8.6.1 (a) We have $a(t,x) = x$, $c(t,x) = x$, $\varphi(s) = xe^{s-t}$ and $u(t,x) = \int_t^T xe^{s-t}\,ds = x(e^{T-t} - 1)$.
(b) $a(t,x) = tx$, $c(t,x) = -\ln x$, $\varphi(s) = xe^{(s^2-t^2)/2}$ and $u(t,x) = -\int_t^T \ln\big(xe^{(s^2-t^2)/2}\big)\,ds = -(T-t)[\ln x + \frac{T}{6}(T+t) - \frac{t^2}{3}]$.

8.6.2 (a) $u(t,x) = x(T-t) + \frac{1}{2}(T-t)^2$. (b) $u(t,x) = \frac{2}{3}e^x\big(e^{\frac{3}{2}(T-t)} - 1\big)$. (c) We have $a(t,x) = \mu x$, $b(t,x) = \sigma x$, $c(t,x) = x$. The associated diffusion is $dX_s = \mu X_s ds + \sigma X_s dW_s$, $X_t = x$, which is the geometric Brownian motion

$$X_s = xe^{(\mu - \frac{1}{2}\sigma^2)(s-t) + \sigma(W_s - W_t)}, \qquad s \geq t.$$

The solution is

$$\begin{aligned}
u(t,x) &= E\Big[\int_t^T xe^{(\mu-\frac{1}{2}\sigma^2)(s-t)+\sigma(W_s-W_t)}\,ds\Big] \\
&= x\int_t^T e^{(\mu-\frac{1}{2}(s-t))(s-t)} E[e^{\sigma(s-t)}]\,ds \\
&= x\int_t^T e^{(\mu-\frac{1}{2}(s-t))(s-t)} e^{\sigma^2(s-t)/2}\,ds \\
&= x\int_t^T e^{\mu(s-t)}\,ds = \frac{x}{\mu}\Big(e^{\mu(T-t)} - 1\Big).
\end{aligned}$$

Chapter 9

9.1.7 Apply Example 9.1.6 with $u = 1$.

9.1.8 $X_t = \int_0^t h(s)\,dW_s \sim N(0, \int_0^t h^2(s)\,ds)$. Then e^{X_t} is log-normal with $E[e^{X_t}] = e^{\frac{1}{2}Var(X_t)} = e^{\frac{1}{2}\int_0^t h(s)^2\,ds}$.

9.1.9 (a) Using Exercise 9.1.8 we have

$$\begin{aligned}
E[M_t] &= E[e^{-\int_0^t u(s)\,dW_s} e^{-\frac{1}{2}\int_0^t u(s)^2\,ds}] \\
&= e^{-\frac{1}{2}\int_0^t u(s)^2\,ds} E[e^{-\int_0^t u(s)\,dW_s}] = e^{-\frac{1}{2}\int_0^t u(s)^2\,ds} e^{\frac{1}{2}\int_0^t u(s)^2\,ds} = 1.
\end{aligned}$$

(b) Similar computation with (a).

9.1.10 (a) Applying the product and Ito's formulas we get

$$d(e^{t/2}\cos W_t) = -e^{-t/2}\sin W_t\,dW_t.$$

Integrating yields

$$e^{t/2}\cos W_t = 1 - \int_0^t e^{-s/2}\sin W_s\,dW_s,$$

which is an Ito integral, and hence a martingale; (b) Similarly.

9.1.12 Use that the function $f(x_1, x_2) = e^{x_1}\cos x_2$ satisfies $\Delta f = 0$.

9.1.14 (a) $f(x) = x^2$; (b) $f(x) = x^3$; (c) $f(x) = x^n/(n(n-1))$; (d) $f(x) = e^{cx}$; (e) $f(x) = \sin(cx)$.

Chapter 10

10.3.4 (a) Since $r_t \sim N(\mu, s^2)$, with $\mu = b + (r_0 - b)e^{-at}$ and $s^2 = \frac{\sigma^2}{2a}(1 - e^{-2at})$. Then

$$P(r_t < 0) \;=\; \int_{-\infty}^{0} \frac{1}{\sqrt{2\pi}s} e^{-\frac{(x-\mu)^2}{2s^2}}\,dx = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{-\mu/s} e^{-v^2/2}\,dv = N(-\mu/s),$$

where by computation

$$-\frac{\mu}{s} = -\frac{1}{\sigma}\sqrt{\frac{2a}{e^{2at}-1}}\big(r_0 + b(e^{at}-1)\big).$$

(b) Since

$$\lim_{t\to\infty}\frac{\mu}{s} = \frac{\sqrt{2a}}{\sigma}\lim_{t\to\infty}\frac{r_0 + b(e^{at}-1)}{\sqrt{e^{2at}-1}} = \frac{b\sqrt{2a}}{\sigma},$$

then

$$\lim_{t\to\infty} P(r_t < 0) = \lim_{t\to\infty} N(-\mu/s) = N\Big(-\frac{b\sqrt{2a}}{\sigma}\Big).$$

It is worth noting that the previous probability is less than 0.5.
(c) The rate of change is

$$\begin{aligned}
\frac{d}{dt}P(r_t < 0) \;&=\; -\frac{1}{\sqrt{2\pi}}e^{-\frac{\mu^2}{2s^2}}\frac{d}{ds}\Big(\frac{\mu}{s}\Big)\\
&=\; -\frac{1}{\sqrt{2\pi}}e^{-\frac{\mu^2}{2s^2}}\frac{ae^{2at}[b(e^{2at}-1)(e^{-at}-1) - r_0]}{(e^{2at}-1)^{3/2}}.
\end{aligned}$$

10.3.5 By Ito's formula

$$d(r_t^n) = \Big(na(b - r_t)r_t^{n-1} + \frac{1}{2}n(n-1)\sigma^2 r_t^{n-1}\Big)dt + n\sigma r_t^{n-\frac{1}{2}}dW_t.$$

Integrating between 0 and t and taking the expectation yields

$$\mu_n(t) = r_0^n + \int_0^t [nab\mu_{n-1}(s) - na\mu_n(s) + \frac{n(n-1)}{2}\sigma^2\mu_{n-1}(s)]\,ds$$

Differentiating yields

$$\mu_n'(t) + na\mu_n(t) = (nab + \frac{n(n-1)}{2}\sigma^2)\mu_{n-1}(t).$$

Multiplying by the integrating factor e^{nat} yields the exact equation

$$[e^{nat}\mu_n(t)]' = e^{nat}(nab + \frac{n(n-1)}{2}\sigma^2)\mu_{n-1}(t).$$

Integrating yields the following recursive formula for moments

$$\mu_n(t) = r_0^n e^{-nat} + (nab + \frac{n(n-1)}{2}\sigma^2)\int_0^t e^{-na(t-s)\mu_{n-1}(s)\,ds}.$$

10.5.1 (a) The spot rate r_t follows the process

$$d(\ln r_t) = \theta(t)dt + \sigma dW_t.$$

Integrating yields

$$\ln r_t = \ln r_0 + \int_0^t \theta(u)\,du + \sigma W_t,$$

which is normally distributed. (b) Then $r_t = r_0 e^{\int_0^t \theta(u)\,du + \sigma W_t}$ is log-normally distributed. (c) The mean and variance are

$$E[r_t] = r_0 e^{\int_0^t \theta(u)\,du} e^{\frac{1}{2}\sigma^2 t}, \quad Var(r_t) = r_0^2 e^{2\int_0^t \theta(u)\,du} e^{\sigma^2 t}(e^{\sigma^2 t} - 1).$$

10.5.2 Substitute $u_t = \ln r_t$ and obtain the linear equation

$$du_t + a(t)u_t dt = \theta(t)dt + \sigma(t)dW_t.$$

The equation can be solved multiplying by the integrating factor $e^{\int_0^t a(s)\,ds}$.

Chapter 11

11.1.1 Apply the same method used in the Application 11.1. The price of the bond is: $P(t,T) = e^{-r_t(T-t)}E\big[e^{-\sigma\int_0^{T-t} N_s\,ds}\big]$.

11.2.2 The price of an infinitely lived bond at time t is given by $\lim_{T\to\infty} P(t,T)$. Using $\lim\limits_{T\to\infty} B(t,T) = \dfrac{1}{a}$, and

$$\lim_{T\to\infty} A(t,T) = \begin{cases} +\infty, & \text{if } b < \sigma^2/(2a) \\ 1, & \text{if } b > \sigma^2/(2a) \\ (\frac{1}{a} + t)(b - \frac{\sigma^2}{2a^2}) - \frac{\sigma^2}{4a^3}, & \text{if } b = \sigma^2/(2a), \end{cases}$$

we can get the price of the bond in all three cases.

Chapter 12

12.1.1 Dividing the equations

$$\begin{aligned} S_t &= S_0 e^{(\mu-\frac{\sigma^2}{2})t+\sigma W_t} \\ S_u &= S_0 e^{(\mu-\frac{\sigma^2}{2})u+\sigma W_u} \end{aligned}$$

yields

$$S_t = S_u e^{(\mu-\frac{\sigma^2}{2})(t-u)+\sigma(W_t-W_u)}.$$

Taking the predictable part out and dropping the independent condition we obtain

$$\begin{aligned} E[S_t|\mathcal{F}_u] &= S_u e^{(\mu-\frac{\sigma^2}{2})(t-u)} E[e^{\sigma(W_t-W_u)}|\mathcal{F}_u] \\ &= S_u e^{(\mu-\frac{\sigma^2}{2})(t-u)} E[e^{\sigma(W_t-W_u)}] \\ &= S_u e^{(\mu-\frac{\sigma^2}{2})(t-u)} E[e^{\sigma W_{t-u}}] \\ &= S_u e^{(\mu-\frac{\sigma^2}{2})(t-u)} e^{\frac{1}{2}\sigma^2(t-u)} \\ &= S_u e^{\mu(t-u)} \end{aligned}$$

Similarly we obtain

$$\begin{aligned} E[S_t^2|\mathcal{F}_u] &= S_u^2 e^{2(\mu-\frac{\sigma^2}{2})(t-u)} E[e^{2\sigma(W_t-W_u)}|\mathcal{F}_u] \\ &= S_u^2 e^{2\mu(t-u)} e^{\sigma^2(t-u)}. \end{aligned}$$

Then

$$\begin{aligned} Var(S_t|\mathcal{F}_u) &= E[S_t^2|\mathcal{F}_u] - E[S_t|\mathcal{F}_u]^2 \\ &= S_u^2 e^{2\mu(t-u)} e^{\sigma^2(t-u)} - S_u^2 e^{2\mu(t-u)} \\ &= S_u^2 e^{2\mu(t-u)} \Big(e^{\sigma^2(t-u)} - 1\Big). \end{aligned}$$

When $s = t$ we get $E[S_t|\mathcal{F}_t] = S_t$ and $Var(S_t|\mathcal{F}_t) = 0$.

12.1.2 By Ito's formula

$$\begin{aligned} d(\ln S_t) &= \frac{1}{S_t} dS_t - \frac{1}{2}\frac{1}{S_t^2}(dS_t)^2 \\ &= (\mu - \frac{\sigma^2}{2})dt + \sigma dW_t, \end{aligned}$$

so $\ln S_t = \ln S_0 + (\mu - \frac{\sigma^2}{2})t + \sigma W_t$, and hence $\ln S_t$ is normally distributed with $E[\ln S_t] = \ln S_0 + (\mu - \frac{\sigma^2}{2})t$ and $Var(\ln S_t) = \sigma^2 t$.

12.1.3 (a) $d\Big(\frac{1}{S_t}\Big) = (\sigma^2-\mu)\frac{1}{S_t}dt - \frac{\sigma}{S_t}dW_t$;
(b) $d\big(S_t^n\big) = n(\mu + \frac{n-1}{2}\sigma^2)dt + n\sigma S_t^n dW_t$;
(c)

$$\begin{aligned} d\big((S_t-1)^2\big) &= d(S_t^2) - 2dS_t = 2S_t dS_t + (dS_t)^2 - 2dS_t \\ &= \Big((2\mu+\sigma^2)S_t^2 - 2\mu S_t\Big)dt + 2\sigma S_t(S_t-1)dW_t. \end{aligned}$$

12.1.4 (b) Since $S_t = S_0 e^{(\mu-\frac{\sigma^2}{2})t+\sigma W_t}$, then $S_t^n = S_0^n e^{(n\mu-\frac{\sigma^2}{2})t+n\sigma W_t}$ and hence

$$\begin{aligned} E[S_t^n] &= S_0^n e^{(n\mu-\frac{\sigma^2}{2})t} E[e^{n\sigma W_t}] \\ &= S_0^n e^{(n\mu-\frac{\sigma^2}{2})t} e^{\frac{1}{2}n^2\sigma^2 t} \\ &= S_0^n e^{(n\mu+\frac{n(n-1)}{2}\sigma^2)t}. \end{aligned}$$

(a) Let $n=2$ in (b) and obtain $E[S_t^2] = S_0 e^{(2\mu+\sigma^2)t}$.

12.1.5 (a) Let $X_t = S_t W_t$. The product rule yields

$$\begin{aligned} d(X_t) &= W_t dS_t + S_t dW_t + dS_t dW_t \\ &= W_t(\mu S_t dt + \sigma S_t dW_t) + S_t dW_t + (\mu S_t dt + \sigma S_t dW_t)dW_t \\ &= S_t(\mu W_t + \sigma)dt + (1+\sigma W_t)S_t dW_t. \end{aligned}$$

Integrating

$$X_s = \int_0^s S_t(\mu W_t + \sigma)dt + \int_0^s (1+\sigma W_t)S_t dW_t.$$

Let $f(s) = E[X_s]$. Then

$$f(s) = \int_0^s \big(\mu f(t) + \sigma S_0 e^{\mu t}\big)\, dt.$$

Differentiating yields

$$f'(s) = \mu f(s) + \sigma S_0 e^{\mu s}, \qquad f(0) = 0,$$

with the solution $f(s) = \sigma S_0 s e^{\mu s}$. Hence $E[W_t S_t] = \sigma S_0 t e^{\mu t}$.
(b) $Cov(S_t, W_t) = E[S_t W_t] - E[S_t]E[W_t] = E[S_t W_t] = \sigma S_0 t e^{\mu t}$. Then

$$\begin{aligned} Corr(S_t, W_t) &= \frac{Cov(S_t, W_t)}{\sigma_{S_t}\sigma_{W_t}} = \frac{\sigma S_0 t e^{\mu t}}{S_0 e^{\mu t}\sqrt{e^{\sigma^2 t}-1}\sqrt{t}} \\ &= \sigma\sqrt{\frac{t}{e^{\sigma^2 t}-1}} \to 0, \quad t\to\infty. \end{aligned}$$

12.1.8 $d = S_d/S_0 = 0.5$, $u = S_u/S_0 = 1.5$, $\gamma = -6.5$, $p = 0.98$.

12.2.2 Let T_a be the time S_t reaches level a for the first time. Then

$$\begin{aligned} F(a) &= P(\overline{S}_t < a) = P(T_a > t) \\ &= \ln \frac{a}{S_0} \int_t^\infty \frac{1}{\sqrt{2\pi\sigma^3\tau^3}} e^{-\frac{(\ln \frac{a}{S_0} - \alpha\sigma\tau)^2}{2\tau\sigma^3}} \, d\tau, \end{aligned}$$

where $\alpha = \mu - \frac{1}{2}\sigma^2$.

12.2.3 (a) $P(T_a < T) = \ln \frac{a}{S_0} \int_0^T \frac{1}{\sqrt{2\pi\sigma^3\tau^3}} e^{-(\ln \frac{a}{S_0} - \alpha\sigma\tau)^2/(2\tau\sigma^3)} \, d\tau$.

(b) $\int_{T_1}^{T_2} p(\tau) \, d\tau$.

12.4.3 $E[A_t] = S_0\big(1 + \frac{\mu t}{2}\big) + O(t^2)$, $Var(A_t) = \frac{S_0}{3}\sigma^2 t + O(t^2)$.

12.4.6 (a) Since $\ln G_t = \frac{1}{t}\int_0^t \ln S_u \, du$, using the product rule yields

$$\begin{aligned} d(\ln G_t) &= d\Big(\frac{1}{t}\Big) \int_0^t \ln S_u \, du + \frac{1}{t} d\Big(\int_0^t \ln S_u \, du\Big) \\ &= -\frac{1}{t^2}\Big(\int_0^t \ln S_u \, du\Big) \, dt + \frac{1}{t} \ln S_t \, dt \\ &= \frac{1}{t}(\ln S_t - \ln G_t) dt. \end{aligned}$$

(b) Let $X_t = \ln G_t$. Then $G_t = e^{X_t}$ and Ito's formula yields

$$\begin{aligned} dG_t &= de^{X_t} = e^{X_t} dX_t + \frac{1}{2} e^{X_t} (dX_t)^2 = e^{X_t} dX_t \\ &= G_t d(\ln G_t) = \frac{G_t}{t}(\ln S_t - \ln G_t) dt. \end{aligned}$$

12.4.7 Using the product rule we have

$$\begin{aligned} d\Big(\frac{H_t}{t}\Big) &= d\Big(\frac{1}{t}\Big) H_t + \frac{1}{t} dH_t \\ &= -\frac{1}{t^2} H_t dt + \frac{1}{t^2} H_t \Big(1 - \frac{H_t}{S_t}\Big) dt \\ &= -\frac{1}{t^2} \frac{H_t^2}{S_t} dt, \end{aligned}$$

so, for $dt > 0$, then $d\Big(\frac{H_t}{t}\Big) < 0$ and hence $\frac{H_t}{t}$ is decreasing. For the second part, try to apply l'Hospital rule.

12.4.8 By continuity, the inequality (12.4.9) is preserved under the limit.

12.4.9 (a) Use that $\frac{1}{n}\sum S_{t_k}^\alpha = \frac{1}{t}\sum S_{t_k}\frac{t}{n} \to \frac{1}{t}\int_0^t S_u^\alpha \, du$. (b) Let $I_t = \int_0^t S_u^\alpha \, du$. Then

$$d\Big(\frac{1}{t} I_t\Big) = d\Big(\frac{1}{t}\Big) I_t + \frac{1}{t} dI_t = \frac{1}{t}\Big(S_t^\alpha - \frac{I_t}{t}\Big) dt.$$

Let $X_t = A_t^\alpha$. Then by Ito's formula we get

$$
\begin{aligned}
dX_t &= \alpha\Big(\frac{1}{t}I_t\Big)^{\alpha-1} d\Big(\frac{1}{t}I_t\Big) + \frac{1}{2}\alpha(\alpha-1)\Big(\frac{1}{t}I_t\Big)^{\alpha-2}\bigg(d\Big(\frac{1}{t}I_t\Big)\bigg)^2 \\
&= \alpha\Big(\frac{1}{t}I_t\Big)^{\alpha-1} d\Big(\frac{1}{t}I_t\Big) = \alpha\Big(\frac{1}{t}I_t\Big)^{\alpha-1}\frac{1}{t}\Big(S_t^\alpha - \frac{I_t}{t}\Big)dt \\
&= \frac{\alpha}{t}\Big(S_t(A_t^\alpha)^{1-1/\alpha} - A_t^\alpha\Big)dt.
\end{aligned}
$$

(c) If $\alpha = 1$ we get the continuous arithmetic mean, $A_t^1 = \frac{1}{t}\int_0^t S_u\,du$. If $\alpha = -1$ we obtain the continuous harmonic mean, $A_t^{-1} = \frac{t}{\int_0^t \frac{1}{S_u}\,du}$.

12.4.10 (a) By the product rule we have

$$
\begin{aligned}
dX_t &= d\Big(\frac{1}{t}\Big)\int_0^t S_u dW_u + \frac{1}{t}d\Big(\int_0^t S_u\,dW_u\Big) \\
&= -\frac{1}{t^2}\Big(\int_0^t S_u\,dW_u\Big)dt + \frac{1}{t}S_t dW_t \\
&= = -\frac{1}{t}X_t dt + \frac{1}{t}S_t dW_t.
\end{aligned}
$$

Using the properties of Ito integrals, we have

$$
\begin{aligned}
E[X_t] &= \frac{1}{t}E[\int_0^t S_u\,dW_u] = 0 \\
Var(X_t) &= E[X_t^2] - E[X_t]^2 = E[X_t^2] \\
&= \frac{1}{t^2}E\Big[\Big(\int_0^t S_u\,dW_u\Big)\Big(\int_0^t S_u\,dW_u\Big)\Big] \\
&= \frac{1}{t^2}\int_0^t E[S_u^2]\,du = \frac{1}{t^2}\int_0^t S_0 e^{(2\mu+\sigma^2)u}\,du \\
&= \frac{S_0^2}{t^2}\frac{e^{(2\mu+\sigma^2)t} - 1}{2\mu+\sigma^2}.
\end{aligned}
$$

(b) The stochastic differential equation of the stock price can be written in the form

$$
\sigma S_t dW_t = dS_t - \mu S_t dt.
$$

Integrating between 0 and t yields

$$
\sigma\int_0^t S_u\,dW_u = S_t - S_0 - \mu\int_0^t S_u\,du.
$$

Dividing by t yields the desired relation.

12.5.2 Using independence $E[S_t] = S_0 e^{\mu t} E[(1+\rho)^{N_t}]$. Then use

$$E[(1+\rho)^{N_t}] = \sum_{n\geq 0} E[(1+\rho)^n | N_t = n] P(N_t = n) = \sum_{n\geq 0} (1+\rho)^n \frac{(\lambda t)^n}{n!} = e^{\lambda(1+\rho)t}.$$

12.5.3 $E[\ln S_t] = \ln S_0 \big(\mu - \lambda\rho - \frac{\sigma^2}{2} + \lambda \ln(\rho+1)\big)t.$

12.5.4 $E[S_t|\mathcal{F}_u] = S_u e^{(\mu-\lambda\rho-\frac{\sigma^2}{2})(t-u)}(1+\rho)^{N_{t-u}}.$

Chapter 13

13.2.2 (a) Substitute $\lim_{S_t\to\infty} N(d_1) = \lim_{S_t\to\infty} N(d_2) = 1$ and $\lim_{S_t\to\infty} \frac{K}{S_t} = 0$ in

$$\frac{c(t)}{S_t} = N(d_1) - \frac{K}{S_t} e^{-r(T-t)} N(d_2).$$

(b) Use $\lim_{S_t\to 0} N(d_1) = \lim_{S_t\to 0} N(d_2) = 0$. (c) It comes from an analysis similar to the one used at (a).

13.2.3 Differentiating in $c(t) = S_t N(d_1) - K e^{-r(T-t)} N(d_2)$ yields

$$\begin{aligned}
\frac{dc(t)}{dS_t} &= N(d_1) + S_t \frac{dN(d_1)}{dS_t} - K e^{-r(T-t)} \frac{dN(d_2)}{dS_t} \\
&= N(d_1) + S_t N'(d_1) \frac{\partial d_1}{\partial S_t} - K e^{-r(T-t)} N'(d_2) \frac{\partial d_2}{\partial S_t} \\
&= N(d_1) + S_t \frac{1}{\sqrt{2\pi}} e^{-d_1^2/2} \frac{1}{\sigma\sqrt{T-t}} \frac{1}{S_t} - K e^{-r(T-t)} \frac{1}{\sqrt{2\pi}} e^{-d_2^2/2} \frac{1}{\sigma\sqrt{T-t}} \frac{1}{S_t} \\
&= N(d_1) + \frac{1}{\sqrt{2\pi}} \frac{1}{\sigma\sqrt{T-t}} \frac{1}{S_t} e^{-d_2^2/2} \Big[S_t e^{(d_2^2-d_1^2)/2} - K e^{-r(T-t)} \Big].
\end{aligned}$$

It suffices to show

$$S_t e^{(d_2^2-d_1^2)/2} - K e^{-r(T-t)} = 0.$$

Since

$$\begin{aligned}
d_2^2 - d_1^2 &= (d_2 - d_1)(d_2 + d_1) = -\sigma\sqrt{T-t} \frac{2\ln(S_t/K) + 2r(T-t)}{\sigma\sqrt{T-t}} \\
&= -2\Big(\ln S_t - \ln K + r(T-t) \Big),
\end{aligned}$$

then we have

$$\begin{aligned}
e^{(d_2^2-d_1^2)/2} &= e^{-\ln S_t + \ln K - r(T-t)} = e^{\ln \frac{1}{S_t}} e^{\ln K} e^{-r(T-t)} \\
&= \frac{1}{S_t} K e^{-r(T-t)}.
\end{aligned}$$

Therefore

$$S_t e^{(d_2^2 - d_1^2)/2} - Ke^{-r(T-t)} = S_t \frac{1}{S_t} Ke^{-r(T-t)} - Ke^{-r(T-t)} = 0,$$

which shows that $\frac{dc(t)}{dS_t} = N(d_1)$.

13.3.1 The payoff is

$$f_T = \begin{cases} 1, & \text{if } \ln K_1 \le X_T \le \ln K2 \\ 0, & \text{otherwise.} \end{cases}$$

where $X_T \sim N\Big(\ln S_t + (\mu - \frac{\sigma^2}{2})(T-t), \sigma^2(T-t)\Big)$.

$$\begin{aligned}
\widehat{E}_t[f_T] &= E[f_T|\mathcal{F}_t, \mu = r] = \int_{-\infty}^{\infty} f_T(x) p(x)\, dx \\
&= \int_{\ln K_1}^{\ln K_2} \frac{1}{\sigma\sqrt{2\pi}\sqrt{T-t}} e^{-\frac{[x - \ln S_t - (r - \frac{\sigma^2}{2})(T-t)]^2}{2\sigma^2(T-t)}}\, dx \\
&= \int_{-d_2(K_1)}^{-d_2(K_2)} \frac{1}{\sqrt{2\pi}} e^{-y^2/2}\, dy = \int_{d_2(K_2)}^{d_2(K_1)} \frac{1}{\sqrt{2\pi}} e^{-y^2/2}\, dy \\
&= N\big(d_2(K_1)\big) - N\big(d_2(K_2)\big),
\end{aligned}$$

where

$$\begin{aligned}
d_2(K_1) &= \frac{\ln S_t - \ln K_1 + (r - \frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}} \\
d_2(K_2) &= \frac{\ln S_t - \ln K_2 + (r - \frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}}.
\end{aligned}$$

Hence the value of a box-contract at time t is $f_t = e^{-r(T-t)}[N\big(d_2(K_1)\big) - N\big(d_2(K_2)\big)]$.

13.3.2 The payoff is

$$f_T = \begin{cases} e^{X_T}, & \text{if } X_T > \ln K \\ 0, & \text{otherwise.} \end{cases}$$

$$f_t = e^{-r(T-t)} \widehat{E}_t[f_T] = e^{-r(T-t)} \int_{\ln K}^{\infty} e^x p(x)\, dx.$$

The computation is similar with the one of the integral I_2 from Proposition 13.2.1.

13.3.3 (a) The payoff is

$$f_T = \begin{cases} e^{nX_T}, & \text{if } X_T > \ln K \\ 0, & \text{otherwise.} \end{cases}$$

$$
\begin{aligned}
\widehat{E}_t[f_T] &= \int_{\ln K}^{\infty} e^{nx} p(x)\,dx = \frac{1}{\sigma\sqrt{2\pi(T-t)}}\int_{\ln K}^{\infty} e^{nx} e^{-\frac{1}{2}\frac{[x-\ln S_t-(r-\sigma^2/2)(T-t)]^2}{\sigma^2(t-t)}}\,dx \\
&= \frac{1}{\sqrt{2\pi}}\int_{-d_2}^{\infty} e^{n\sigma\sqrt{T-t}y} e^{\ln S_t^n} e^{n(r-\frac{\sigma^2}{2})(T-t)} e^{-\frac{1}{2}y^2}\,dy \\
&= \frac{1}{\sqrt{2\pi}} S_t^n e^{n(r-\frac{\sigma^2}{2})(T-t)} \int_{-d_2}^{\infty} e^{-\frac{1}{2}y^2+n\sigma\sqrt{T-t}y}\,dy \\
&= \frac{1}{\sqrt{2\pi}} S_t^n e^{n(r-\frac{\sigma^2}{2})(T-t)} e^{\frac{1}{2}n^2\sigma^2(T-t)} \int_{-d_2}^{\infty} e^{-\frac{1}{2}(y-n\sigma\sqrt{T-t})^2}\,dy \\
&= \frac{1}{\sqrt{2\pi}} S_t^n e^{(nr+n(n-1)\frac{\sigma^2}{2})(T-t)} \int_{-d_2-n\sigma\sqrt{T-t}}^{\infty} e^{-\frac{z^2}{2}}\,dz \\
&= S_t^n e^{(nr+n(n-1)\frac{\sigma^2}{2})(T-t)} N(d_2+n\sigma\sqrt{T-t}).
\end{aligned}
$$

The value of the contract at time t is

$$
f_t = e^{-r(T-t)}\widehat{E}_t[f_T] = S_t^n e^{(n-1)(r+\frac{n\sigma^2}{2})(T-t)} N(d_2+n\sigma\sqrt{T-t}).
$$

(b) Using $g_t = S_t^n e^{(n-1)(r+\frac{n\sigma^2}{2})(T-t)}$, we get $f_t = g_t N(d_2+n\sigma\sqrt{T-t})$.
(c) In the particular case $n=1$, we have $N(d_2+\sigma\sqrt{T-t}) = N(d_1)$ and the value takes the form $f_t = S_t N(d_1)$.

13.4.1 Since $\ln S_T \sim N\Big(\ln S_t + \big(\mu - \frac{\sigma^2}{2}(T-t)\big), \sigma^2(T-t)\Big)$, we have

$$
\begin{aligned}
\widehat{E}_t[f_T] &= E\big[(\ln S_T)^2|\mathcal{F}_t, \mu=r\big] \\
&= Var\Big(\ln S_T|\mathcal{F}_t, \mu=r\Big) + E[\ln S_T|\mathcal{F}_t, \mu=r]^2 \\
&= \sigma^2(T-t) + \big[\ln S_t + (r-\frac{\sigma^2}{2})(T-t)\big]^2,
\end{aligned}
$$

so

$$
f_t = e^{-r(T-t)}\Big[\sigma^2(T-t) + \big[\ln S_t + (r-\frac{\sigma^2}{2})(T-t)\big]^2\Big].
$$

13.5.1 $\ln S_T^n$ is normally distributed with

$$
\ln S_T^n = n\ln S_T \sim N\Big(n\ln S_t + n\big(\mu - \frac{\sigma^2}{2}(T-t)\big), n^2\sigma^2(T-t)\Big).
$$

Redo the computation of Proposition 13.2.1 in this case.

13.6.1 Let $n=2$.

$$
\begin{aligned}
f_t &= e^{-r(T-t)}\widehat{E}_t[(S_T-K)^2] \\
&= e^{-r(T-t)}\widehat{E}_t[S_T^2] + e^{-r(T-t)}\Big(K^2 - 2K\widehat{E}_t[S_T]\Big) \\
&= S_t^2 e^{(r+\sigma^2)(T-t)} + e^{-r(T-t)}K^2 - 2KS_t.
\end{aligned}
$$

Let $n = 3$. Then

$$\begin{aligned} f_t &= e^{-r(T-t)}\widehat{E}_t[(S_T - K)^3] \\ &= e^{-r(T-t)}\widehat{E}_t[S_T^3 - 3KS_T^2 + 3K^2S_T + K^3] \\ &= e^{-r(T-t)}\widehat{E}_t[S_T^3] - 3Ke^{-r(T-t)}\widehat{E}_t[S_T^2] + e^{-r(T-t)}(3K^2\widehat{E}_t[S_T] + K^3) \\ &= e^{2(r+3\sigma^2/2)(T-t)}S_t^3 - 3Ke^{(r+\sigma^2)(T-t)}S_t^2 + 3K^2S_t + e^{-r(T-t)}K^3. \end{aligned}$$

13.8.1 Choose $c_n = 1/n!$ in the general contract formula.

13.9.1 Similar computation with the one done for the call option.

13.11.1 Write the payoff as a difference of two puts, one with strike price K_1 and the other with strike price K_2. Then apply the superposition principle.

13.11.2 Write the payoff as $f_T = c_1 + c_3 - 2c_2$, where c_i is a call with strike price K_i. Apply the superposition principle.

13.11.4 (a) By inspection. (b) Write the payoff as the sum between a call and a put both with strike price K, and then apply the superposition principle.

13.11.4 (a) By inspection. (b) Write the payoff as the sum between a call with strike price K_2 and a put with strike price K_1, and then apply the superposition principle. (c) The strangle is cheaper.

13.11.5 The payoff can be written as a sum between the payoffs of a (K_3, K_4)-bull spread and a (K_1, K_2)-bear spread. Apply the superposition principle.

13.12.2 By computation.

13.12.3 The risk-neutral valuation yields

$$\begin{aligned} f_t &= e^{-r(T-t)}\widehat{E}_t[S_T - A_T] = e^{-r(T-t)}\widehat{E}_t[S_T] - e^{-r(T-t)}\widehat{E}_t[A_T] \\ &= S_t - e^{-r(T-t)}\frac{t}{T}A_t + \frac{1}{rT}S_t\Big(1 - e^{-r(T-t)}\Big) \\ &= S_t\Big(1 + \frac{1}{rT} - \frac{1}{rT}e^{-r(T-t)}\Big) - e^{-r(T-t)}\frac{t}{T}A_t. \end{aligned}$$

13.12.5 We have

$$\begin{aligned} \lim_{t\to 0} f_t &= S_0 e^{-rT + (r - \frac{\sigma^2}{2})\frac{T}{2} + \frac{\sigma^2 T}{6}} - e^{-rT}K \\ &= S_0 e^{-\frac{1}{2}(r + \frac{\sigma^2}{6})T} - e^{-rT}K. \end{aligned}$$

13.12.6 Since $G_T < A_T$, then $\widehat{E}_t[G_T] < \widehat{E}_t[A_T]$ and hence

$$e^{-r(T-t)}\widehat{E}_t[G_T - K] < e^{-r(T-t)}\widehat{E}_t[A_T - K],$$

so the forward contract on the geometric average is cheaper.

13.12.7 Dividing

$$\begin{aligned} G_t &= S_0 e^{(\mu - \frac{\sigma^2}{2})\frac{t}{2}} e^{\frac{\sigma}{t}\int_0^t W_u\,du} \\ G_T &= S_0 e^{(\mu - \frac{\sigma^2}{2})\frac{T}{2}} e^{\frac{\sigma}{T}\int_0^t W_u\,du} \end{aligned}$$

yields

$$\frac{G_T}{G_t} = e^{(\mu - \frac{T-t}{2})\frac{T-t}{2}} e^{\frac{\sigma}{T}\int_0^T W_u\,du - \frac{\sigma}{t}\int_0^t W_u\,du}.$$

An algebraic manipulation leads to

$$G_T = G_t e^{(\mu - \frac{\sigma^2}{2})\frac{T-t}{2}} e^{\sigma(\frac{1}{T} - \frac{1}{t})\int_0^t W_u\,du} e^{\frac{\sigma}{T}\int_t^T W_u\,du}.$$

13.15.7 Use that

$$\begin{aligned} P(\max_{t\le T} S_t \ge z) &= P\Big(\sup_{t\le T}[(\mu - \frac{\sigma^2}{2})t + \sigma W_t] \ge \ln\frac{z}{S_0}\Big) \\ &= P\Big(\sup_{t\le T}[\frac{1}{\sigma}(\mu - \frac{\sigma^2}{2})t + W_t] \ge \frac{1}{\sigma}\ln\frac{z}{S_0}\Big). \end{aligned}$$

Chapter 15

15.5.1 $P = F - \Delta_F S = c - N(d_1)S = -Ke^{-r(T-t)}$.

Chapter 17

17.1.9

$$\begin{aligned} P(S_t < b) &= P\Big((r - \frac{\sigma^2}{2})t + \sigma W_t < \ln\frac{b}{S_0}\Big) = P\Big(\sigma W_t < \ln\frac{b}{S_0} - (r - \frac{\sigma^2}{2})t\Big) \\ &= P\big(\sigma W_t < -\lambda\big) = P\big(\sigma W_t > \lambda\big) \le e^{-\frac{-\lambda^2}{2\sigma^2 t}} \\ &= e^{-\frac{1}{2\sigma^2 t}[\ln(S_0/b) + (r - \sigma^2/2)t]^2}, \end{aligned}$$

where we used Proposition 1.12.11 with $\mu = 0$.

Bibliography

[1] J. Bertoin. *Lévy Processes*. Cambridge University Press, 121, 1996.

[2] F. Black, E. Derman, and W. Toy. *A One-Factor Model of Interest Rates and Its Application to Treasury Bond Options*. Financial Analysts Journal, Jan.-Febr., 1990, pp.33–39.

[3] F. Black and P. Karasinski. *Bond and Option Pricing when Short Rates are Lognormal*. Financial Analysts Journal, Jul.-Aug., 1991, pp.52–59.

[4] J. R. Cannon. *The One-dimensional Heat Equation*. Encyclopedia of Mathematics and Its Applications, 23, Cambridge, 2008.

[5] J. C. Cox, J. E. Ingersoll, and S. A. Ross. *A Theory of the term Structure of Interest Rates*. Econometrica, 53, 1985, pp.385–407.

[6] W. D. D. Wackerly, Mendenhall, and R. L. Scheaffer. *Mathematical Statistics with Applications, 7th ed.* Brooks/Cole, 2008.

[7] T.S.Y. Ho and S.B. Lee. *Term Structure Movements and Pricing Interest Rate Contingent Claims*. Journal of Finance, 41, 1986, pp.1011–1029.

[8] J. Hull and A. White. *Pricing Interest Rates Derivative Securities*. Review of Financial Studies, 3, 4, 1990, pp.573–592.

[9] H.H. Kuo. *Introduction to Stochastic Integration*. Springer, Universitext, 2006.

[10] M. Abramovitz M. and I. Stegun. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. New York: Dover, 1965.

[11] R.C. Merton. *Option Pricing with Discontinuous Returns*. Journal of Financial Economics, 3, 1976, pp.125–144.

[12] R.C. Merton. *Theory of Rational Option Pricing*. The Bell Journal of Economics and Management Science, Vol. 4, No. 1, (Spring, 1973), pp. 141-183.

[13] B. Øksendal. *Stochastic Differential Equations, An Introduction with Applications, 6th ed.* Springer-Verlag Berlin Heidelberg New-York, 2003.

[14] R. Rendleman and B. Bartter. *The Pricing of Options on Debt Securities*. Journal of Financial and Quantitative Analysis, 15, 1980, pp.11–24.

[15] S.M. Turnbull and L. M. Wakeman. *A quick algorithm for pricing European average options*. Journal of Financial and Quantitative Analysis, 26, 1991, p. 377.

[16] O. A. Vasicek. *An Equilibrium Characterization of the Term Structure*. Journal of Financial Economics, 5, 1977, pp.177–188.

[17] D.V. Widder. *The Heat Equation*. Academic, London, 1975.

Index