Pricing Theory

Sebastian Jaimungal and Ali Al-Aradi

Contents

Introduction

This collection of notes focuses on financial theory and its application to pricing various derivative products. This text grew over a period of several years from teaching the Pricing Theory/Applied Probability for Mathematical Finance (STA2503/MMF1952/MMF1928) course at the University of Toronto. It is designed for graduate students taking a first course in mathematical finance and derivative pricing theory.

The main goals of this book are two-fold:

(i) To develop models of financial markets and describe the evolution of various processes, e.g. asset prices, interest rates, foreign exchange rates.
(ii) To price various financial derivatives written on these processes.

The book is divided into two main parts: the first introduces some of the *mathematical background* needed later in the book. It is by no means exhaustive in its treatment of the topics presented and a prior knowledge of probability theory, stochastic calculus, as well as ordinary and partial differential equations will be useful.

The second part focuses on *derivative pricing*, starting with discrete-time models and moving on to continuous-time models, with the majority of the discussion being devoted to the latter. Once again, although an effort is made to ensure that the content is as self-contained as possible, the presentation is not exhaustive and some background in financial products will be helpful to the reader.

Part I

Mathematical Background

1 Probability Theory Basics

In this chapter we discuss some preliminary aspects of probability theory and stochastic processes which will be useful for later discussions. The focus in the discussion will be on developing intuition surrounding the technical details. We will use the simple example of a six-sided die toss to better explain some of the technicalities that arise and add some intuition to the explanation.

1.1 Probability Spaces

The primary object in probability theory is that of a **probability space**. Specifically, a probability space is a triple consisting of three objects often denoted $(\Omega, \mathcal{F}, \mathbb{P})$.

1.1.1 Sample Spaces

In a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, Ω is the **sample space**, which is a set of **sample outcomes** or possible realizations ω for a given experiment or random phenomenon. Each $\omega \in \Omega$ is an "atomic" outcome in the sense that each ω is unique and different outcomes are mutually exclusive (only one outcome will occur on each trial of the experiment). This set of outcomes may be finite or infinite and countable or uncountable. In the six-sided die toss example, the sample space is the set $\Omega = \{1, 2, 3, 4, 5, 6\}$ as these are the elementary/atomic outcomes; every die toss will produce precisely one these results.

1.1.2 σ-algebras

The second element of a probability space is the **σ-algebra**, $\mathcal{F}$, which defines the collection of events that can be observed for the experiment, i.e. the information that can be derived from the experiment. Before explaining these notions we first provide the definition:

Definition 1.1 (σ-algebra). A **σ-algebra** $\mathcal{F}$ is a collection of subsets of Ω that satisfies:

(i) if $A \in \mathcal{F}$ then $A^c \in \mathcal{F}$,

(ii) for a countable sequence of sets $\{A_i\}_{i\in\mathbb{N}}$, if $A_i \in \mathcal{F}$ then $\cup_{i\in\mathbb{N}} A_i \in \mathcal{F}$.

Each member of the σ-algebra $\mathcal{F}$ is an **event**, which is itself a set of outcomes from Ω. Note that events are not synonymous with outcomes and that it is possible to define σ-algebras that do not involve any of the sample outcome singletons. The reason σ-algebras are defined in this manner is quite intuitive: if an observer can conclude that an event A has occurred they should also be able to conclude that the complement of that event A^c has not occurred as these statements are equivalent - this corresponds to the first requirement of closure under complements. Similarly, if they can conclude that an event A has occurred, they must be able to concluded that the event "at least one of A and B" i.e. $A \cup B$ has occurred. This corresponds to the second requirement of closure under unions. Note that the combination of the two statements implies closure under intersections by De Morgan's law, i.e. for $\{A_i\}_{i\in\mathbb{N}}$, if $A_i \in \mathcal{F}$ then $\cap_{i\in\mathbb{N}} A_i \in \mathcal{F}$.

To illustrate the difference between outcomes and events and to explain the relation between σ-algebras and information we return to our die toss example. For this example we can consider two separate σ-algebras on Ω: $\mathcal{F} = \mathcal{P}(\Omega)$, the power set of Ω consisting of all of subsets of Ω and $\mathcal{G} = \{\emptyset, \Omega, \{1,3,5\}, \{2,4,6\}\}$. It is easy to verify that both of these collections of subsets are indeed σ-algebras. Notice now that the σ-algebra $\mathcal{G}$ consists of only four events, none of which are the singleton sample outcomes which make up Ω. This demonstrates the difference between outcomes and events. To understand why σ-algebras are related to informational content, consider the fact that $\mathcal{G}$ contains only 4 elements while $\mathcal{F}$ contains 2^6 events. The fact that $\mathcal{G}$ groups the sample outcomes $1,3,5$ and $2,4,6$ indicates that we are effectively only capable of "identifying" if the outcome is odd or even. Meanwhile $\mathcal{F}$ can "identify" a much wider range of events, e.g. with the outcome is greater than 4 (since it contains the event $\{5,6\}$ or if the outcome is divisible by 3 (since it contains the event $\{3,6\}$). In short, the fact that $\mathcal{G} \subset \mathcal{F}$ indicates that $\mathcal{F}$ "contains more information" than $\mathcal{G}$.

1.1.3 Probability Measures

A **probability measure** is a function that assigns to each event in the σ-algebra a probability of occurring:

Definition 1.2 (Probability Measure). A **probability measure** is a function $\mathbb{P} : \mathcal{F} \to [0,1]$ that satisfies:

(i) $\mathbb{P}(A) \geq \mathbb{P}(\emptyset) = 0$ for all $A \in \mathcal{F}$,

(ii) for a countable sequence of *disjoint* sets $\{A_i\}_{i\in\mathbb{N}}$, $\mathbb{P}\left(\bigcup\limits_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(A_i)$.

Note that probabilities are assigned to events and not sample outcomes (unless the σ-algebra contains the singleton sample outcomes themselves). As an example, we will give two probability measures for the two σ-algebras defined in the previous section. For $\mathcal{F}$ we define $\mathbb{P}_{\mathcal{F}} : \mathcal{F} \to [0,1]$ to be $\mathbb{P}_{\mathcal{F}}(A) = |A|/6$ for each $A \in \mathcal{F}$ where $|A|$ denotes the number of elements in the set A. For $\mathcal{G}$ we define $\mathbb{P}_{\mathcal{G}} : \mathcal{G} \to [0,1]$ as:

$$\begin{aligned}
&\mathbb{P}_{\mathcal{G}}(\emptyset) = 0 \\
&\mathbb{P}_{\mathcal{G}}(\Omega) = 1 \\
&\mathbb{P}_{\mathcal{G}}(\{1,3,5\}) = 1/3 \\
&\mathbb{P}_{\mathcal{G}}(\{2,4,6\}) = 2/3
\end{aligned}$$

Notice that $\mathbb{P}_{\mathcal{F}}$ and $\mathbb{P}_{\mathcal{G}}$ are defined for every event in their respective σ-algebras (not for the singletons that make up the sample space Ω) and that it is easy to verify that they are indeed valid probability measures.

We have illustrated how we can define probability measures for different σ-algebras. However, it is also possible to have different probability measure defined on the same σ-algebra. For example we can define the following probability measures on the σ-algebra $\mathcal{G}$:

$A \in \mathcal{G}$	$\mathbb{P}_1(A)$	$\mathbb{P}_2(A)$	$\mathbb{P}_3(A)$	$\mathbb{P}_4(A)$
$\emptyset$	0	0	0	0
{1,3,5}	1/3	1/2	1	0
{2,4,6}	2/3	1/2	0	1
Ω	1	1	1	1

An important notion when working with multiple probability measures defined on the same space is that of **absolute continuity** and **equivalence**, which are defined as follows:

Definition 1.3 (Absolute Continuity and Equivalence). Let $\mathbb{P}$ and $\mathbb{P}^*$ be two probability measures defined on the same measurable space $(\Omega, \mathcal{F})$. $\mathbb{P}^*$ is **absolutely continuous** with respect to $\mathbb{P}$, denoted by $\mathbb{P}^* \ll \mathbb{P}$, if $\mathbb{P}(A) = 0 \implies \mathbb{P}^*(A) = 0$. $\mathbb{P}$ and $\mathbb{P}^*$ are **equivalent**, denoted $\mathbb{P} \sim \mathbb{P}^*$, if $\mathbb{P}^* \ll \mathbb{P}$ and $\mathbb{P} \ll \mathbb{P}^*$, i.e. $\mathbb{P}(A) = 0 \iff \mathbb{P}^*(A) = 0$.

Intuitively, absolute continuity and equivalence capture whether or not different probability measures agree on which events have zero probability (or equivalently, which events occur with probability 1). In the table above, we can see that $\mathbb{P}_1$ and $\mathbb{P}_2$ are equivalent but this is not the case for any other pair of probability measures. However, it is true that both $\mathbb{P}_1$ and $\mathbb{P}_2$ are absolutely continuous with respect to $\mathbb{P}_3$ and $\mathbb{P}_4$.

1.2 Random Variables

A **random variable** is a function that maps each sample outcome to an element in some target space S, that is, a random variable is a function $X : \Omega \to S$ where S is an arbitrary set. Most commonly we have $S = \mathbb{R}$, as with typical continuous random variables or $S = \mathbb{N}$, as with typical discrete random variables. This means that for every random outcome, an associated real or natural number is produced by the random variable.

Returning to our die toss example, consider the indicator function $X(\omega) = \mathbb{1}\{\omega < 4\}$. The mapping $X : \Omega \to \{0, 1\}$ identifies whether or not the die toss produced an outcome less than 4. The key point here is to realize that the random variable is in fact a *function mapping from the sample space to some target space.*

1.2.1 Measurability

The technical definition of a random variable involves the notion of **measurability**, which we explain in stages. To understand the concept of measurability, we first need to expand our notion of the random variable X. Previously, we mentioned that a random variable is a mapping for the sample space Ω to a target space S. In order to discuss the measurability of X we need to consider σ-algebras for both the domain and the target space. In the case of the domain Ω we already have the σ-algebra $\mathcal{F}$, but for the discussion to make sense we need to also consider a σ-algebra for S which we call $\mathcal{S}$. Putting this together, we will now view X as a mapping $X : (\Omega, \mathcal{F}) \to (S, \mathcal{S})$.

Next, we need to define the **pre-image** of X for a given set in the target space $B \subseteq S$. The pre-image of B under X is the set of points in the domain that X would map to an element contained in B, i.e. $X^{-1}(B) = \{\omega : X(\omega) \in B\}$. Now, we can define

Definition 1.4 (Measurablility). A function $X : (\Omega, \mathcal{F}) \to (S, \mathcal{S})$ is a **measurable** map from $(\Omega, \mathcal{F})$ to $(S, \mathcal{S})$ if, for every $B \in \mathcal{S}$, $X^{-1}(B) \in \mathcal{F}$. A function X satisfying this property is said to be **$\mathcal{F}$-measurable**, denoted $X \in \mathcal{F}$.

In words, this says that if we start from a member of the σ-algebra $\mathcal{S}$, and take its pre-image under X, we end up with a subset of elements of that is a member of the σ-algebra $\mathcal{F}$.

To illustrate this concept we return to our die toss example. First, we need to define a σ-algebra for the target space $S = \{0, 1\}$, for example $\mathcal{S} = \{\emptyset, \{0\}, \{1\}, S\}$.

Next, we need to look at the pre-image of each member of $\mathcal{S}$ under X:

$$
\begin{aligned}
X^{-1}(\emptyset) &= \emptyset \\
X^{-1}(\{0\}) &= \{1,2,3\} \\
X^{-1}(\{1\}) &= \{4,5,6\} \\
X^{-1}(S) &= \Omega
\end{aligned}
$$

Since all of these pre-images are members of $\mathcal{F}$ and $\mathcal{G}$ we can say that $X \in \mathcal{F}$ and $X \in \mathcal{G}$, i.e. X is $\mathcal{F}$-measurable and $\mathcal{G}$-measurable. An example of a σ-algebra that X would not be measurable with respect to is $\mathcal{H} = \{\emptyset, \{1,2,3,4,5\}, \{6\}, \Omega\}$, since the pre-images listed above are not members of $\mathcal{H}$. This once again ties back to the notion of σ-algebras and information: $\mathcal{F}$ and $\mathcal{G}$ have enough information to determine X while $\mathcal{H}$ does not. In fact, even though $\mathcal{G}$ and $\mathcal{H}$ have a comparable amount of information (as they contain 4 elements each), $\mathcal{H}$ does not have the *right* information needed for X.

The reason we restrict random variables to measurable functions is quite intuitive. The only subsets (events) that we have probabilities for (i.e. that we can "measure") are the members of $\mathcal{F}$, since $\mathbb{P}$ is only defined for $A \in \mathcal{F}$. On the other hand, we can define events associated with random variables in terms of subsets of Ω (the pre-images), but if those subsets are not $\mathcal{F}$ then we cannot measure "them." If X is measurable then we can guarantee that all events expressed in terms of X can be mapped back to subsets of Ω that *can be measured* because they are members of $\mathcal{F}$. An alternative way of viewing this is that measurability allows us to define a probability measure on the target space that is consistent with the original probability measure defined on the domain[1].

1.2.2 $\sigma(X)$

One σ-algebra that will prove to be particularly useful is $\sigma(X)$. This is notation used to refer to the smallest σ-algebra for which X becomes a measurable map (and hence a valid random variable). In other words, it is the smallest σ-algebra that contains the smallest/least amount of information sufficient to determine X. As we saw in our last example, $X \in \mathcal{F}$ and also $X \in \mathcal{G}$, but $\mathcal{G} = \{\emptyset, \Omega, \{1,3,5\}, \{2,4,6\}\}$ is smaller than $\mathcal{F}$. Furthermore, if we were to make $\mathcal{G}$ any smaller it would not contain the information needed to make X measurable. So, in this case $\sigma(X) = \mathcal{G}$; the smallest σ-algebra that X is measurable with respect to.

[1] The measure induced by X on the measurable space $(S, \mathcal{S})$ is defined as $\widetilde{\mathbb{P}}(B) = \mathbb{P}(X^{-1}(B))$ for all $B \in \mathcal{S}$ and is called the **distribution** of X. It is easy to see that for this measure to be well-defined we need $X^{-1}(B)) \in \mathcal{F}$ for all $B \in \mathcal{S}$, i.e. for X to be measurable.

1.3 Stochastic Processes

It is also possible for the target space S to be a *set of functions*. In other words, each random outcome is mapped to an entire function. This set of functions may be defined over a continuous domain or a discrete domain (see examples below). A **stochastic process** can be viewed precisely in this way: as a random function/sequence.

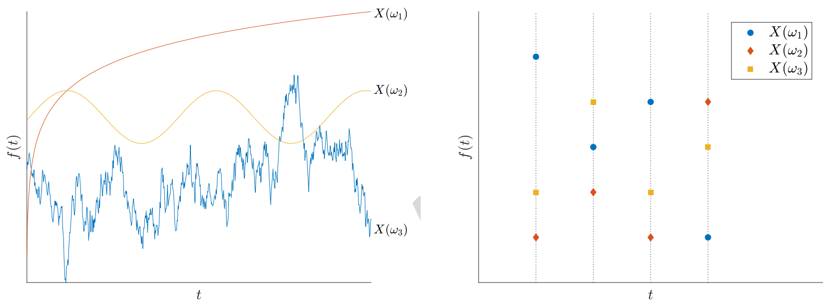


Figure 1.1 Random functions with a continuous domain (left) and discrete domain (right). Every random outcome ω_i is mapped to a certain function (left) or sequence (right).

Having established that stochastic processes are random functions we should also note that there is an additional dimension that we are interested in. Namely, it is usually not enough to know what function is realized, but what value it takes at given points in time. As such, this gives an alternative way of viewing stochastic processes: as a function of both the underlying source of randomness and of time. Concretely, we can say that a stochastic process X is a mapping from the product of the sample space Ω and a set of times (index set) $\mathcal{T}$ to the real numbers $\mathbb{R}$:

$$\begin{aligned} X : \Omega \times \mathcal{T} &\to \mathbb{R} \\ (\omega, t) &\mapsto X(\omega, t) \end{aligned}$$

In other words, given a realization of a sample outcome ω and a time point t, $X(\omega, t)$ is the value of the random function associated with ω evaluated at time t. This gives us two views of stochastic processes:

- <u>FOR FIXED ω:</u> we have a function of time (with no randomness).
- <u>FOR FIXED t:</u> we have a random variable (in the usual sense).

With this in mind, a stochastic process can also be viewed as a collection of

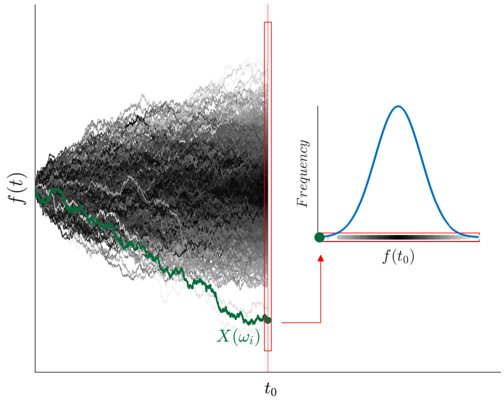


Figure 1.2 The green path is a random function of time associated with a particular sample outcome ω_i. The (random) value of different possible realized functions at fixed point t_0 gives rise to a distribution and an associated random variable, X_{t_0}.

random variables indexed by time:

$$X = \{X_t(\omega) : t \in \mathcal{T}\}$$

Here, each X_t for every $t \in \mathcal{T}$ is a random variable. Usually, we drop the dependence on the sample outcome ω and write:

$$X = \{X_t : t \in \mathcal{T}\} = \{X_t\}_{t \in \mathcal{T}}$$

1.3.1 Functions of Stochastic Processes

Often we will use functions to generate new stochastic processes, however the overloaded notation commonly used in these instances can be confusing at times. To demonstrate this with an example, let $f(t, x) = t^3 \cdot x^2$ be a real-valued function defined for each positive t and real x, i.e. $f : \mathbb{R}^+ \times \mathbb{R} \to \mathbb{R}$. Now, let $X = \{X_t\}_{t \geq 0}$ be a stochastic process defined on $\mathbb{R}$. Then $f(t, X_t) = t^3 \cdot (X_t)^2$ is a new stochastic process. It is common to use the notation $f_t = f(t, X_t)$ and denote the new process:

$$\underset{\text{process}}{\underset{\uparrow}{f}} = \{f_t\}_{t \geq 0} = \{\underset{\text{function}}{\underset{\uparrow}{f}}(t, X_t)\ \}_{t \geq 0}$$

Notice the abuse of notation here where f is used to refer to both the function and the stochastic process. It is important to distinguish which usage is being applied depending on the context. For example, when f is used in a PDE we know that we are referring to the function; when we speak of the quadratic variation of f we know that we are referring to the process. The same concept applies to the derivatives of f:

$$\partial_t f(t,x) = 3t^2 \cdot x^2 \qquad \partial_x f(t,x) = t^3 \cdot 2x$$

The two partial derivatives above are each new functions of t and x. But plugging in X_t as we have above:

$$\partial_t f(t,X_t) = 3t^2 \cdot (X_t)^2 \qquad \partial_x f(t,X_t) = t^3 \cdot 2X_t$$

gives us two new stochastic processes. Using the same notation as above we can write the process associated with the partial derivative:

$$\underset{\text{process}}{\underset{\uparrow}{\partial_x f}} = \{(\partial_x f)_t\}_{t\geq 0} = \{\underset{\text{function}}{\underset{\uparrow}{\partial_x f}}(t,X_t)\}_{t\geq 0}$$

Figure 1.3.1 further highlights the distinction between functions and processes.

1.3.2 Filtrations and Flow of Information

We have seen previously how a σ-algebra corresponds to the notion of informational content. Here we would like to extend this notion to reflect the flow of time when discussing stochastic processes. Additionally, we would like to discuss the idea of measurability in the context of processes in an analogous way as when we had a single random variable. This means that for a stochastic process, which can be viewed as a *collection* of random variables $\{X_t\}_{t\geq 0}$, we need to consider the measurability of each individual X_t. To this end, we introduce two new concepts: **filtrations** and **adapted processes**.

Definition 1.5 (Filtration). A **filtration**, $\mathfrak{F}$, is a collection of σ-algebras $\{\mathcal{F}_t\}_{t\geq 0}$ satisfying $\mathcal{F}_t \subset \mathcal{F}_u$ for all $u > t$. A probability space $(\Omega, \mathcal{F}, \mathbb{P})$ equipped with a filtration $\mathfrak{F}$ is referred to as a **filtered** probability space, denoted by $(\Omega, \mathcal{F}, \mathfrak{F}, \mathbb{P})$

In words, a filtration is simply an increasing sequence of σ-algebras. Each σ-algebra can be viewed as the information we have at that particular point in time. The requirement that the σ-algebras are contained in one another (increasing) reflects the fact that we gain new information with the passage of time and do not lose any of the past information. For an extremely simple example, we return to our die toss and the σ-algebras $\mathcal{F}$ and $\mathcal{G}$. If we define $\mathcal{F}_0 = \{\emptyset, \Omega\}, \mathcal{F}_1 = \mathcal{G}\mathcal{F}_2 = \mathcal{F}$, then by virtue of the fact that $\mathcal{F}_0 \subset \mathcal{F}_1 \subset \mathcal{F}_2$, the sequence $\mathfrak{F} = \{\mathcal{F}_0, \mathcal{F}_1, \mathcal{F}_2\}$ forms a (very small, yet valid) filtration.

Equipped the concept of a filtration we can now discuss the measurability of a stochastic process. We have seen before that a single random variable X is measurable with respect to a σ-algebra $\mathcal{F}$, i.e. $\mathcal{F}$-measurable ($X \in \mathcal{F}$) if $\mathcal{F}$ contains

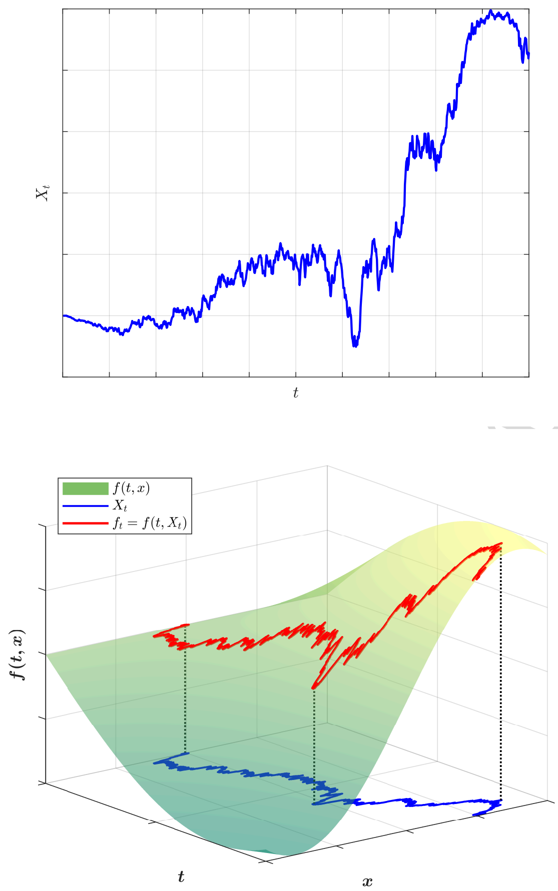


Figure 1.3 Distinguishing between functions and processes generated using functions. The top panel shows a realization of a stochastic process, X_t. The bottom panel shows the surface of the function $f(t,x) = \sin(tx)$; the blue line is the same stochastic process, X_t, drawn on the (t,x) plane; the red line is a second stochastic process $f_t = f(t, X_t)$ generated by feeding X_t into the function $f(t,x)$.

enough information to determine X. The extension for processes quite natural and involves a filtration:

Definition 1.6 (Adapted Process). A stochastic process is **adapted** to a filtration $\mathfrak{F} = \{\mathcal{F}_t\}_{t\geq 0}$, or $\mathfrak{F}$-adapted, if $X_t \in \mathcal{F}_t$ for all t.

In other words, for every fixed t, the random variable X_t (which is a member of the stochastic process) is measurable with respect to the σ-algebra $\mathcal{F}_t$ (which is a member of the filtration $\mathfrak{F}$). This means that the information contained in the filtration (or rather, the filtration's σ-algebras) is sufficient to determine the process *at each point in time.*

Finally, we can also extend the concept of $\sigma(X)$ - the smallest σ-algebra to which X is measurable - to stochastic process. We write $\sigma(\{X_t\}_{t\geq 0})$ to denote the smallest *filtration* to which the process $X = \{X_t\}_{t\geq 0}$ is adapted.

1.3.3 Martingales

One of the most important classes of stochastic processes in probability theory (and pricing theory) is that of a martingale:

DEFINITION 1.7 (Martingale). A stochastic process $M = (M_t)_{t\geq 0}$ defined on the filtered probability space $\{\Omega, \mathcal{F}, \mathfrak{F}, \mathbb{P})\}$ is a **martingale** with respect to the filtration $\mathfrak{F}$ and the probability measure $\mathbb{P}$ if:

(i) M is $\mathfrak{F}$-adapted,
(ii) $\mathbb{E}^{\mathbb{P}}[|M_t|] < \infty$ for all t,
(iii) $\mathbb{E}^{\mathbb{P}}[M_t|\mathcal{F}_s] = M_s$ for all t.

Intuitively, a martingale is a process that remains the same on average, which can be seen by writing the third condition as $\mathbb{E}^{\mathbb{P}}[M_t - M_s|\mathcal{F}_s] = 0$ to notice that the expected increment of the process (conditional on information up to that point) is zero. Note that we emphasize that the conditional expectation is taken using the probability measure $\mathbb{P}$ by including it in the superscript. This is important because a process may be a martingale with respect to one probability measure $\mathbb{P}$ but not with respect to another probability measure $\mathbb{P}^*$. To emphasize the dependence of the martingale property on the probability measure we will sometimes refer to the process as a $\mathbb{P}$-martingale.

2 Stochastic Calculus

This note presents a non-rigorous introduction to some of the key topics in stochastic calculus. Refer to Shreve (2010)[1] or Øksendal (2010)[2] for a more detailed treatment.

2.1 Random Walks and Brownian Motion

2.1.1 Random Walks

We begin with random walks, which are discrete-time stochastic processes, to build some intuition and motivate Brownian motion as their continuous-time counterpart. Assume that we are considering a fixed time interval $[0, T]$, which we divide into n equal increments of size $\Delta t = T/n$.

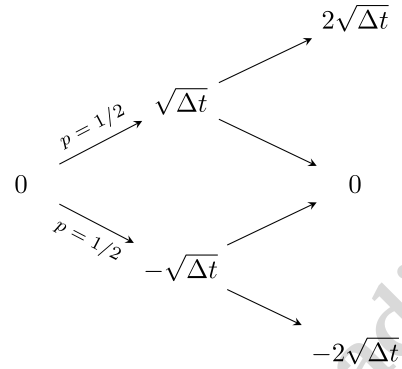
Definition 2.1 (Random walk). A **(scaled) random walk** is a discrete-time stochastic process, $X = (X_k)_{k \in \{0, \Delta t, \ldots, T\}}$ satisyfing $X_0 = 0$ and

$$X_k \;=\; X_{k-1} + \sqrt{\Delta t} \cdot x_k \;=\; \sum_{m=1}^{k} \sqrt{\Delta t} \cdot x_m$$

where x_m are i.i.d. random variables with $\mathbb{P}(x_m = \pm 1) = 1/2$.

That is, X_k is a process that is defined on the points $0, \Delta t, 2\Delta t, \ldots, n\Delta t = T$ that goes up or down by $\sqrt{\Delta t}$ in any given time step with equal probability. So, X_k evolves according to the following tree:

[1] Stochastic Calculus for Finance II: Continuous-Time Models
[2] Stochastic Differential Equations: An Introduction with Applications



Aside: more generally, random walks involve sums of i.i.d. random variables of any distribution - not necessarily Bernoulli random variables.

Next, we can find **some properties of this process**:

- The mean and variance of the process are given by:

$$\begin{aligned}
\mathbb{E}(X_n) &= \mathbb{E}\left[\sum_{m=1}^{n} \sqrt{\Delta t} \cdot x_m\right] = 0, && \text{since } \mathbb{E}(x_m) = 0 \\
\mathbb{V}(X_n) &= \mathbb{V}\left[\sum_{m=1}^{n} \sqrt{\Delta t} \cdot x_m\right] = \Delta t \cdot n = T, && \text{since } \mathbb{V}(x_m) = 1,\ x_m \text{ ind.}
\end{aligned}$$

  These same properties hold over other increments, that is for $s, t \in \{0, \Delta t, ..., T\}$ with $s < t$:

$$\begin{aligned}
\mathbb{E}(X_t - X_s) &= 0 \\
\mathbb{V}(X_t - X_s) &= t - s
\end{aligned}$$

- For non-overlapping intervals, $(s,t) \cap (u,v) \neq \emptyset$, by the independence of the process increments we have:

$$\mathbb{C}(X_v - X_u,\ X_t - X_s) = 0$$

- By the Central Limit Theorem, as the size of the time increments goes to zero, $\Delta t \downarrow 0$ (or, equivalently, as the number of increments goes to infinity, $n \to \infty$), we will arrive at a continuous-time process where the larger increments will be normally distributed, i.e. $X_t - X_s \sim N(0, t-s)$.

2.1.2 Brownian Motion

Definition 2.2 (Brownian motion). A **Brownian motion** (or **Wiener process**) is a continuous-time stochastic process, $W = (W_t)_{t \geq 0}$, that satisfies the following properties:

(i) $W_0 = 0$, almost surely

(ii) W_t is a normally distributed random variable: $W_t \sim N(0, t)$

(iii) W has stationary increments, i.e. the distribution of $W_{t+s} - W_t$ depends only on interval width, s.

(iv) W has independent increments, i.e. if $(s,t) \cap (u,v) = \emptyset$, $W_t - W_s$ is independent of $W_v - W_u$

(v) W is pathwise continuous.

We can see the parallels between this process and the process discussed in the previous section. In fact, the latter satisfies these properties in the limit as $n \to \infty$. So we may think of the Wiener process as the continuous-time version (or limiting form) of the time-scaled random walk.

2.1.3 Total Variation

One important property of Brownian motions is that they are **nowhere differentiable**. To see this let us first define the notion of a **partition**:

Definition 2.3 (Partition). A **partition** of the interval $[0, T]$, Π, is a set of numbers that divide the interval into subintervals:

$$\Pi \;=\; \{t_0, ..., t_n\} \text{ where } \; 0 = t_0 < t_1 < ... < t_n = T$$

The **norm** of the partition is the length of the largest subinterval:

$$\|\Pi\| = \sup_k \; (t_k - t_{k-1})$$

Returning to the question of differentiability - we present the notion of **total variation**:

Definition 2.4 (Total variation). The **total variation** of a stochastic process X is a process defined by:

$$TV_T^X = \lim_{\|\Pi\| \downarrow 0} \sum_{k=1}^{n} |X_{t_k} - X_{t_{k-1}}|$$

The limit in the above definition is taken over a sequence of partitions whose norms form a sequence tending to zero.

Total variation can be seen as a "measure" of differentiability. To see this, we present a *heuristic* argument to show that **total variation for a differentiable**

function is finite:

$$
\begin{aligned}
TV_T^f &= \lim_{||\Pi||\downarrow 0} \sum_{k=1}^{n} |f_{t_k} - f_{t_{k-1}}| \\
&= \lim_{||\Pi||\downarrow 0} \sum_{k=1}^{n} f'_{t_k^*} \cdot \Delta t_k, \quad \text{where } t_k^* \in (t_{k-1}, t_k), \quad \text{by the Mean Value Theorem} \\
&< \lim_{||\Pi||\downarrow 0} \sup f'_{t_k^*} \underbrace{\sum_{k=1}^{n} \Delta t_k}_{T} \\
&< \infty
\end{aligned}
$$

However, **the total variation for a Brownian motion is infinite**:

$$
\begin{aligned}
TV_T^W &= \lim_{||\Pi||\downarrow 0} \sum_{k=1}^{n} |W_{t_k} - W_{t_{k-1}}| \\
&\approx \lim_{||\Pi||\downarrow 0} \sum_{k=1}^{n} \sqrt{\Delta t}, \qquad \text{approximating B.M. increment by random walk increment} \\
&\approx \lim_{\Delta t \to 0} n\sqrt{\Delta t}, \qquad \text{taking equal partition} \\
&= \lim_{n\to\infty} \sqrt{n} \underbrace{\sqrt{n\Delta t}}_{=\sqrt{T}} \\
&= \infty
\end{aligned}
$$

Thus, **Brownian motions are not differentiable**.

2.1.4 Quadratic Variation

Another important concept is that of **quadratic variation**[3]:

DEFINITION 2.5 (Quadratic variation). The **quadratic variation** of a stochastic process X is a process defined by:

$$[X, X]_T = \lim_{\|\Pi\|\downarrow 0} \sum_{k=1}^{n} (X_{t_k} - X_{t_{k-1}})^2$$

Now, we show that **the quadratic variation for differentiable functions**

[3] Not to be confused with variance; the former is a path property while the latter is a property of a random variable at a fixed point in time.

is 0:

$$
\begin{aligned}
[f,f]_t &= \lim_{\|\Pi\|\downarrow 0} \sum_{k=1}^{n} (f_{t_k} - f_{t_{k-1}})^2 \\
&= \lim_{\|\Pi\|\downarrow 0} \sum_{k=1}^{n} (f'_{t_k^*} \cdot \Delta t_k)^2, && \text{by the Mean Value Theorem} \\
&\le \lim_{\|\Pi\|\downarrow 0} \|\Pi\| \cdot \sum_{k=1}^{n} \left(f'_{t_k^*}\right)^2 \cdot \Delta t_k, && \text{since } \|\Pi\| = \sup_i \Delta t_i \ge \Delta t_k \\
&= 0
\end{aligned}
$$

This is different, however, from the quadratic variation for a Brownian motion:

Definition 2.6 (Quadratic variation of Brownian Motion). Let $W = (W_t)_{t\ge 0}$ be a standard Brownian motion. The quadratic variation of W is:

$$
[W,W]_T = T \quad \text{a.s.}
$$

Proof We present a sketch of the proof, where we show convergence in probability rather than almost sure convergence. By definition we have that:

$$
[W,W]_T = \lim_{\|\Pi\|\downarrow 0} \sum_{k=1}^{n} \left[(W_{t_k} - W_{t_{k-1}})^2 - (t_k - t_{k-1})\right] + \sum_{k=1}^{n} (t_k - t_{k-1})
$$

Here, we have added and subtracted the suspected answer $T = \sum_{k=1}^{n}(t_k - t_{k-1})$. Next, denote the first sum above by A^{Π} and compute the mean and variance of this quantity:

$$
\begin{aligned}
\mathbb{E}\left(A^{\Pi}\right) &= \sum_{k=1}^{n} \left[\mathbb{E}\left((W_{t_k} - W_{t_{k-1}})^2\right) - (t_k - t_{k-1})\right] = 0 \\
&\quad \text{since } W_{t_k} - W_{t_{k-1}} \sim N(0, t_k - t_{k-1}) \\
&\quad \implies \mathbb{V}(W_{t_k} - W_{t_{k-1}}) = \mathbb{E}\left((W_{t_k} - W_{t_{k-1}})^2\right) = t_k - t_{k-1} \\
\mathbb{V}\left(A^{\Pi}\right) &= \sum_{k=1}^{n} \mathbb{V}\left((W_{t_k} - W_{t_{k-1}})^2\right) \\
&= \sum_{k=1}^{n} \mathbb{V}\left(\left(\sqrt{t_k - t_{k-1}}\, Z\right)^2\right) && \text{where } Z \sim N(0,1) \\
&= \sum_{k=1}^{n} \Delta t_k^2 \cdot \left(\mathbb{E}(Z^4) - \mathbb{E}(Z^2)^2\right) \\
&= \sum_{k=1}^{n} 2\Delta t_k^2 \\
&\le 2\|\Pi\| \cdot \sum_{k=1}^{n} \Delta t_k \\
&= 2\|\Pi\| \cdot T
\end{aligned}
$$

And this quantity goes to zero as $\|\Pi\| \downarrow 0$. Therefore, we have that $[W, W]_T = T$ in probability. $\square$

This is one of the most fundamental results in stochastic calculus; it is the main reason stochastic calculus differs from ordinary calculus. Similarly, we can define **quadratic covariation** of two processes:

DEFINITION 2.7 (Quadratic covariation). The **quadratic covariation** of two stochastic process X and Y is a process defined by:

$$[X, Y]_T = \lim_{\|\Pi\|\downarrow 0} \sum_{k=1}^{n} (X_{t_k} - X_{t_{k-1}})(Y_{t_k} - Y_{t_{k-1}})$$

Note: **the covariation of a stochastic process with a differentiable function is 0**, i.e. $[X, f]_t = 0$ for f differentiable. This can be shown by an argument similar to the one where we showed that $[f, f]_t = 0$.

2.1.5 Correlated Brownian Motions

DEFINITION 2.8 (Correlated Brownian motions). **Correlated Brownian motions** $(W, B) = (W_t, B_t)_{t\geq 0}$ with instantaneous correlation ρ are joint processes that satisfy:

(i) $W_0 = B_0 = 0$, almost surely
(ii) $\begin{pmatrix} W_t \\ B_t \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} t & \rho t \\ \rho t & t \end{pmatrix}\right)$
(iii) (W, B) has stationary increments
(iv) (W, B) has independent increments, i.e. if $(s, t) \cap (u, v) \neq \emptyset$, $W_t - W_s$ is independent of both $W_v - W_u$ and $B_v - B_u$; similarly, $B_t - B_s$ is independent of both $W_v - W_u$ and $B_v - B_u$
(v) Stochastic continuity of the joint process.

One useful trick is that **correlated Brownian motions can be rewritten in terms of two independent Brownian motions**:

$$\begin{cases} W_t = Z_t \\ B_t = \rho Z_t + \sqrt{(1-\rho^2)} Z_t^{\perp} \end{cases}$$

where $Z_t, Z_t^{\perp}$ are independent Brownian motions. We can check that the main properties we expect of the correlated Brownian motions are preserved:

$$\begin{aligned} \mathbb{E}[W_t] &= \mathbb{E}[Z_t] = 0 \\ \mathbb{E}[B_t] &= \mathbb{E}\left[\rho Z_t + \sqrt{(1-\rho^2)} Z_t^{\perp}\right] = 0 \\ \mathbb{V}[W_t] &= \mathbb{V}[Z_t] = t \\ \mathbb{V}[B_t] &= \mathbb{V}\left[\rho Z_t + \sqrt{(1-\rho^2)} Z_t^{\perp}\right] = \rho^2 \cdot t + (1-\rho)^2 \cdot t = t \\ \mathbb{C}[W_t, B_t] &= \mathbb{C}\left[Z_t, \rho Z_t + \sqrt{(1-\rho^2)} Z_t^{\perp}\right] = \rho \cdot t \end{aligned}$$

2.2 Stochastic Integrals and Stochastic Differential Equations

2.2.1 Itô integrals

Before we tackle stochastic integrals, let us build up some intuition from **ordinary Riemann integrals**. First, fix the interval over which we wish to integrate to be $[0, t]$. One representation of the integral of a function f is:

$$\int_0^t f_s \; ds = \lim_{\|\Pi\| \downarrow 0} \sum_{k=1}^{n} f_{t_k^*} \cdot (t_k - t_{k-1}) \qquad \text{where } t_k^* \in (t_{k-1}, t_k)$$

What we have done is taken a partition of the interval, then for each subinterval we multiply the function value at an intermediate point by the interval width, sum up all the resulting values, and take the limit over finer and finer partitions. We can define the more general **Riemann-Stieltjes integral** in a similar manner:

$$\int_0^t f_s \; dg_s = \lim_{\|\Pi\| \downarrow 0} \sum_{k=1}^{n} f_{t_k^*} \cdot (g_{t_k} - g_{t_{k-1}})$$

The same procedure as above applies, except we have replaced the time increments with increments of some deterministic function, g. One example where such integrals arise is in the context of computing probabilities or expectations, where g is taken to be the cumulative distribution function of a random variable.

An **Itô integral** is defined in a similar manner; instead of using time increments or increments of a deterministic function, we use increments of a Brownian motion:

Definition 2.9 (Itô integral). Let f_t be an adapted stochastic process satisfying $\mathbb{E}\left(\int_0^t f_s^2 \; ds\right) < \infty$. Then the integral of this process with respect to a standard Brownian motion (**Itô integral**) is defined as:

$$\int_0^t f_s \, dW_s = \lim_{\|\Pi\| \downarrow 0} \sum_{k=1}^{n} f_{t_{k-1}} \cdot (W_{t_k} - W_{t_{k-1}})$$

Notice that **for the Itô integral, we took the value of f at the leftmost point of the interval**. With Riemann-Stietljes integrals this does not make a difference, but when we integrate with respect to Brownian motion there is a difference. Mathematically, it is valid to define stochastic integrals using the rightmost point or the midpoint, but the method we chose makes "financial" sense, since portfolios are set up at the start of time intervals and held fixed until the next interval.

A few things to note:

- Since the increments we are taking are of a stochastic process, this means that the **stochastic integral itself is a random variable for a fixed t, and a stochastic process when viewed as a function of t.**

- **The integrand may be a function of a stochastic process itself**, e.g.

$$\int_0^t f(s, W_s)\; dW_s = \lim_{\|\Pi\|\downarrow 0} \sum_{k=1}^{n} f(t_{k-1}, W_{t_{k-1}}) \cdot (W_{t_k} - W_{t_{k-1}})$$

- **When the integrand f_t is deterministic, the integral is normally distributed**. Note that this does not necessarily hold when f is stochastic.

The following results are useful for **computing the mean and variance of Itô integrals**:

THEOREM 2.10 (Martingale property of Itô integrals). *Let $I_t = \int_0^t f_s\; dW_s$ be an Itô integral as defined above. Then I_t is a martingale. Moreover, since $I_0 = 0$, this implies that:*

$$\mathbb{E}\left(\int_0^t f_s\; dW_s\right) = 0$$

THEOREM 2.11 (Itô's isometry). *Let $(W, B) = (W_t, B_t)_{t\geq 0}$ be correlated Brownian motions with instantaneous correlation ρ. Then:*

$$\mathbb{E}\left[\left(\int_0^t g(s, W_s, B_s)\; dW_s\right)\left(\int_0^t h(s, W_s, B_s)\; dB_s\right)\right] \\ = \mathbb{E}\left[\int_0^t g(s, W_s, B_s) \cdot h(s, W_s, B_s)\; \rho\; ds\right]$$

In particular, Itô's isometry can be used to compute the variances (and covariances) of Itô integrals since these integrals have zero mean:

$$\mathbb{V}\left[\int_0^t f_s\; dW_s\right] = \mathbb{E}\left[\left(\int_0^t f_s\; dW_s\right)^2\right] = \mathbb{E}\left[\int_0^t f_s^2\; ds\right]$$

We can also consider the **quadratic variation of an Itô integrals**:

THEOREM 2.12 (Quadratic variation of Itô integral). *Let $I_t = \int_0^t f_s\; dW_s$ be an Itô integral as defined above. Then the quadratic variation of I is equal to:*

$$[I, I]_t = \int_0^t f_s^2\; ds$$

Finally, we note that Itô integrals satisfy the following linearity property:

THEOREM 2.13 (Linearity of Itô integrals). *Let $a, b \in \mathbb{R}$ be some constants. Then we have that:*

$$a\int_0^t f_s\; dW_s + b\int_0^t g_s\; dW_s = \int_0^t (af_s + bg_s)\; dW_s$$

2.2.2 Stochastic Differential Equations

The idea with Itô integrals is that they allow us to build processes that are driven by increments of a Brownian motion. Depending on how the underlying Brownian motion changes in any given interval, the new process will change in a certain way. Take for example a process defined in the following manner:

$$X_t = X_0 + \int_0^t \mu_s \, ds + \int_0^t \sigma_s \, dW_s$$

Clearly, this process starts at X_0 as both integrals will be equal to zero when $t = 0$. It will then evolve in a "deterministic" fashion[4] according to the function μ_t, but it will also vary according to the Brownian increments coming from the stochastic integral, which in turn are being scaled by the function σ_t.

Now, if we took X_0 to the LHS and rewrote $X_t - X_0 = \int_0^t dX_s$ (that is, the change in the process is equal to the sum of all its increments), then we could remove the integrals and present the equation above in the form of a **stochastic differential equation (SDE)**:

$$dX_t = \mu_t \, dt + \sigma_t \, dW_t$$

This statement itself is not particularly rigorous;[5] it is actually **shorthand notation for the (formal) definition given above in terms of an ordinary integral and an Itô integral**. However, it does have some intuitive appeal. If you think of dX_t as the increment of the X_t process over a very short period of time, i.e. "$dX_t = X_{t+dt} - X_t$", you can see that this small change is driven by two things:

(i) An increment due to the passage of time, which is scaled by a (possibly random) quantity μ_t;
(ii) An increment equal to a small change in the underlying Brownian motion ("$dW_t = W_{t+dt} - W_t$"), scaled by a (possibly random) quantity σ_t.

The reason we say "possibly random" is because the processes μ and σ may be stochastic or deterministic. You can think of this as a way of allowing the scaling factors to be larger when the process itself is large, for example. Processes of this form are known as an **Itô processes**:

Definition 2.14 (Itô process). Let μ_t and σ_t be adapted stochastic process and let X_0 be nonrandom. Then an **Itô process** is a process of the form:

$$X_t \;=\; X_0 + \int_0^t \mu_s \, ds + \int_0^t \sigma_s \, dW_s$$

[4] The quotation marks are there because μ may be a stochastic process, but it would be scaling *deterministic* increments of time.

[5] For one thing, dW_t is nonsensical, since Brownian motion is not differentiable.

That is, X satisfies the stochastic differential equation:

$$dX_t \; = \; \mu_t \; dt + \sigma_t \; dW_t$$

The processes μ and σ are referred to as the **drift** and **volatility** of the process X.

2.3 Itô's Lemma

Assume we have an Itô process of the form:

$$dX_t = \mu_t \; dt + \sigma_t \; dW_t$$

By taking a function of this process we can create a new process: $Y_t = f(t, X_t)$. Then we may ask the question, **what is the SDE that the new process satisfies**? The answer lies in the application of **Itô's lemma**:

THEOREM 2.15 (Itô's lemma). *Let X_t be an Itô process that satisfies the SDE:*

$$dX_t = \mu_t \; dt + \sigma_t \; dW_t$$

and define a process $Y_t = f(t, X_t)$ where $f(t, x) \in C^{1,2}$, i.e. its derivatives in t and x are defined and continuous (once in t and twice in x). Then Y_t satisfies the SDE:

$$dY_t = \left(\partial_t f(t, X_t) + \mu_t \cdot \partial_x f(t, X_t) + \frac{1}{2}\sigma_t^2 \cdot \partial_{xx} f(t, X_t)\right) dt + \sigma_t \cdot \partial_x f(t, X_t) \; dW_t$$

So, the new process, Y_t, is itself an Itô process with drift equal to $\left(\partial_t f + \mu_t \cdot \partial_x f + \frac{1}{2}\sigma_t^2 \cdot \partial_{xx} f\right)$ and volatility equal to $(\sigma_t \cdot \partial_x f)$.

We can obtain some intuition for this result by looking at Taylor expansion of f. For this, we need to refer to the following heuristic multiplication table:

$$dt \cdot dt = 0, \qquad dt \cdot dW_t = 0, \qquad dW_t \cdot dW_t = dt$$

These are non-rigorous statements related to the quadratic variation of a Brownian motion, as well as its covariation with a differentiable function. Next, consider the second order Taylor expansion of $f(t, X_t)$:

$$df_t \; = \; \partial_t f \cdot dt + \partial_x f \cdot dX_t + \frac{1}{2}\partial_{xx} f \cdot (dX_t)^2$$

Substituting in the expression for dX_t we have:

$$\begin{aligned} df_t \; &= \; \partial_t f \cdot dt + \partial_x f \cdot (\mu_t \; dt + \sigma_t \; dW_t) + \frac{1}{2}\partial_{xx} f \cdot (\mu_t \; dt + \sigma_t \; dW_t)^2 \\ &= \; \partial_t f \cdot dt + \mu_t \cdot \partial_x f \cdot dt + \sigma_t \cdot \partial_x f \; dW_t + \frac{1}{2}\partial_{xx} f \cdot \left(\mu_t^2 \; (dt)^2 + 2\mu_t\sigma_t dt dW_t + \sigma_t^2 \; (dW_t)^2\right) \\ &= \; \left(\partial_t f + \mu_t \cdot \partial_x f + \frac{1}{2}\sigma_t^2 \cdot \partial_{xx} f\right) dt + \sigma_t \cdot \partial_x f \; dW_t \end{aligned}$$

The final step follows by applying the multiplication rules above and collecting dt terms.

Itô's lemma also extends to the **two-dimensional case**, i.e. when we build processes out of multiple Itô processes:

THEOREM 2.16 (Two-dimensional Itô's lemma). *Suppose that X_t and Y_t are Itô processes satisfying the SDEs:*

$$\begin{aligned} dX_t &= \mu_t \; dt + \sigma_t \; dW_t \\ dY_t &= \nu_t \; dt + \eta_t \; dB_t \end{aligned}$$

where W_t, B_t are correlated Brownian motions with instantaneous correlation ρ. Define a process $Z_t = f(t, X_t, Y_t)$ where $f(t, x, y) \in C^{1,2,2}$. Then Z_t satisfies the SDE:

$$\begin{aligned} dZ_t = \left(\partial_t g + \mu_t \cdot \partial_x g + \frac{1}{2}\sigma_t^2 \cdot \partial_{xx} g + \nu_t \cdot \partial_y g + \frac{1}{2}\eta_t^2 \cdot \partial_{yy} g + \rho \sigma_t \eta_t \partial_{xy} g \right) dt \\ + \sigma_t \cdot \partial_x g \; dW_t + \eta_t \cdot \partial_x g \; dB_t \end{aligned}$$

Notice that in the drift we find the terms that would appear if g were a function of X_t alone or of Y_t alone, as well as a cross-term and the usual term associated with time. Naturally, we also find that the new process is driven by both Brownian motions that drive X_t and Y_t. The result can be linked to the Taylor expansion of f as we did in the one-dimensional case by adding the following to the multiplication rules:

$$dW_t \cdot dB_t \;=\; \rho \; dt$$

and writing the multidimensional version of the second order Taylor expansion:

$$df_t \;=\; \partial_t f \cdot dt + \sum_{i=1}^{n} \partial_{x_i} f \cdot dX_t^{(i)} + \frac{1}{2} \sum_{i,j=1}^{n} \partial_{x_i x_j} f \cdot \left(dX_t^{(i)} \cdot dX_t^{(j)} \right)$$

The following two corollaries are useful extensions of the theorem above:

COROLLARY 2.17 (Itô's product and quotient rule). *For two Itô processes X and Y we have:*

$$\frac{d(X_t Y_t)}{X_t Y_t} = \frac{dX_t}{X_t} + \frac{dY_t}{Y_t} + \frac{d[X,Y]_t}{X_t Y_t}; \qquad \frac{d(X_t/Y_t)}{X_t/Y_t} = \frac{dX_t}{X_t} - \frac{dY_t}{Y_t} + \frac{d[Y,Y]_t}{Y_t^2} - \frac{d[X,Y]_t}{X_t Y_t}$$

2.4 Mean-reverting processes

An important class of stochastic processes is that of mean-reverting processes. The main feature of these processes is that they have some level (which may be time varying, or even itself stochastic) to which they gravitate to.

2.4.1 Ornstein-Uhlenbeck Process

An **Ornstein-Uhlenbeck (OU) process** is an Itô process that satisfies the SDE:

$$dX_t = \kappa(\theta - X_t)\ dt + \sigma\ dW_t$$

where W is a standard Brownian motion. To see why there is mean reversion, consider the sign of the drift term when $X_t > \theta$ vs. $X_t < \theta$: in the first case the drift term is negative which pushes X_t downwards towards θ, and in the second case the drift term is positive which pushes X_t upwards towards θ. As such, θ is referred to as the **mean reversion level**. Furthermore, the magnitude of κ dictates how large the drift term is and, by extension, how quickly X_t moves towards θ. For this reason, κ is referred to as the **mean reversion rate**.

It is possible to obtain a **closed-form solution for** X_t in this case. Consider the process $Y_t = X_t \cdot e^{\kappa t}$. By applying Itô's lemma we have:

$$\begin{aligned}
dY_t &= dX_t \cdot e^{\kappa t} + X_t \cdot d\left(e^{\kappa t}\right) + d\left[X, e^{\kappa t}\right]_t && \text{by Itô's product rule} \\
&= (\kappa(\theta - X_t)\ dt + \sigma\ dW_t) \cdot e^{\kappa t} + X_t \cdot \kappa e^{\kappa t}\ dt && \text{since } e^{\kappa t} \text{ is differentiable} \\
&= \kappa\theta e^{\kappa t}\ dt + \sigma e^{\kappa t}\ dW_t
\end{aligned}$$

Integrating on both sides and substituting in $Y_t = X_t \cdot e^{\kappa t}$ we have:

$$\begin{aligned}
Y_T - Y_t &= \kappa\theta \int_t^T e^{\kappa u}\ du + \int_t^T \sigma e^{\kappa u}\ dW_u \\
\implies \quad X_T \cdot e^{\kappa T} - X_t \cdot e^{\kappa t} &= \theta\left(e^{\kappa T} - e^{\kappa t}\right) + \sigma \int_t^T e^{\kappa u}\ dW_u \\
\implies \quad X_T &= X_t \cdot e^{-\kappa(T-t)} + \theta\left(1 - e^{-\kappa(T-t)}\right) + \sigma \int_t^T e^{-\kappa(T-u)}\ dW_u
\end{aligned}$$

It follows from the properties of Itô integrals that, conditional on $\mathcal{F}_t$, X_T is normally distributed:

$$X_T | \mathcal{F}_t \sim N\left(X_t \cdot e^{-\kappa(T-t)} + \theta\left(1 - e^{-\kappa(T-t)}\right),\ \frac{\sigma^2}{2\kappa}\left(1 - e^{-2\kappa(T-t)}\right)\right)$$

Notice that as T tends to infinity, X_T is asymptotically Gaussian with mean θ (the mean reversion level) and asymptotic variance $\frac{\sigma^2}{2\kappa}$.

(i) Show that the integral of the OU process, $\int_t^T X_u\ du$, is normally distributed with:

$$\int_t^T X_u\, du \sim N\left(\theta(T-t) - \frac{1 - e^{-\kappa(T-t)}}{\kappa}(\theta - X_t),\ \frac{\sigma^2}{\kappa^2}\int_t^T \left(1 - e^{-\kappa(T-u)}\right)^2 du\right)$$

(ii) Compute $\mathbb{C}\left(X_T,\ \int_t^T X_u\ du\right)$.

2.4.2 Feller Process

A **Feller process** is a mean-reverting process that satisfies the SDE:

$$dX_t = \kappa(\theta - X_t)\ dt + \sigma\sqrt{X_t}\ dW_t$$

Similar to the OU process the Feller process reverts to the mean reversion level θ with mean reversion rate κ. However, unlike the OU process, the **Feller process remains positive almost surely**, provided that the **Feller condition** is met: $2\kappa\theta > \sigma^2$. The intuition for the process remaining positive is to notice that as X_t decreases the drift term becomes positive while the noise term tends to 0; the square root rate is the smallest rate that ensures that the noise diminishes fast enough to ensure positivity with probability one.

Also, unlike the OU process, **the Feller process does not admit a closed-form solution**. However, it is possible to show that its time-dependent mean is the same as that of the OU process:

$$\begin{aligned} & dX_t = \kappa(\theta - X_t)\ dt + \sigma\sqrt{X_t}\ dW_t \\ \Longrightarrow\quad & X_t - X_0 = \int_0^t \kappa(\theta - X_u)\ du + \sigma\int_0^t \sqrt{X_u}\ dW_u && \text{integrating from 0 to } t \\ \Longrightarrow\quad & \mathbb{E}(X_t) - X_0 = \int_0^t \kappa(\theta - \mathbb{E}(X_u))\ du + \sigma\cdot\mathbb{E}\left[\int_0^t \sqrt{X_u}\ dW_u\right] && \text{taking expectations} \\ \Longrightarrow\quad & m_t - m_0 = \int_0^t \kappa(\theta - m_u)\ du && \text{letting } \mathbb{E}(X_t) = m_t \end{aligned}$$

The last step follows since the mean of Itô integral is equal to zero. Now, we can write this last line in differential form to obtain the following ODE:

$$\begin{cases} m'(t) = \kappa(\theta - m_t) \\ m_0 = X_0 \end{cases}$$

The solution to this ODE gives us:

$$\mathbb{E}(X_t) = X_0\cdot e^{-\kappa t} + \theta\left(1 - e^{-\kappa t}\right)$$

Notice that, once again, this quantity tends to θ as t tends to infinity.

2.5 Exercises

Exercise 2.18. Suppose that $W = \{W_t\}_{t\geq 0}$ and $Z = \{Z_t\}_{t\geq 0}$ are correlated Wiener processes with instantaneous correlation ρ. Find the mean and variance of each of the following random variables:

(i) $W_1 + W_2 + \cdots + W_N$ for $N \in \mathbb{N}$
(ii) $\exp\{W_t\}$
(iii) $a\,W_t + b\,Z_t$
(iv) $\exp\{a\,W_t + b\,Z_t\}$
(v) $W_t\,Z_t$
(vi) $\int_0^t W_s\,Z_s\,ds$
(vii) $W_t W_s$ for $s < t$
(viii) $\int_0^t W_s\;ds$
(ix) $W_s Z_t$ for $s < t$
(x) $\int_0^t W_s\;dZ_s - \int_0^t Z_s\;dW_s$

Exercise 2.19. Suppose that $W = \{W_t\}_{t\geq 0}$ and $Z = \{Z_t\}_{t\geq 0}$ are standard correlated Wiener processes with instantaneous correlation ρ. Find the SDE which X_t solves:

(i) $X_t = W_t^n$
(ii) $X_t = \cos(W_t)$
(iii) $X_t = W_t\,Z_t$
(iv) $X_t = \exp\{\alpha\,Z_t + \beta\,W_t\}$

Exercise 2.20. In the following $W = \{W_t\}_{t\geq 0}$ and $Z = \{Z_t\}_{t\geq 0}$ denote correlated Wiener processes with instantaneous correlation ρ. For each of the following (i) compute the mean and variance of Y_t and (ii) derive an integration by parts formulae:

(i) $Y_t = \int_0^t W_s\,dW_s$
(ii) $Y_t = \int_0^t (W_s)^2\,dW_s$
(iii) $Y_t = \int_0^t s\,W_s\,dW_s$
(iv) $Y_t = \int_0^t s\,W_s^2\,dW_s$
(v) $Y_t = \int_0^t s^2\,W_s^2\,dW_s$
(vi) $Y_t = \int_0^t e^{-W_s}\,dW_s$
(vii) $Y_t = \int_0^t s\,e^{-W_s}\,dW_s$
(viii) $Y_t = \int_0^t W_s\,dZ_s$
(ix) $Y_t = \int_0^t s\,W_s\,dZ_s$

Exercise 2.21. Prove the following from first principles, i.e. using the fundamental definition of the Itô integral (Note: W_t and Z_t denote correlated Wiener processes with instantaneous correlation ρ.):

(i) $\int_0^t W_s^2 dW_s = \frac{1}{3}W_t^3 - \int_0^t W_s\,ds$
(ii) $\int_0^t W_s\,dZ_s + \int_0^t Z_s\,dW_s = W_t\,Z_t - \rho\,t$

EXERCISE 2.22. Assume that the prices of two stocks S_t and U_t satisfy the following coupled SDEs:

$$\frac{dU_t}{U_t} = a_t\,dt + b_t\,dW_t^U\ , \qquad \frac{dS_t}{S_t} = c_t\,dt + d_t\,dW_t^S \tag{2.1}$$

where W_t^U and W_t^S denote two correlated Wiener processes with constant instantaneous correlation ρ and a_t, b_t, c_t, d_t are deterministic functions only of time t.

(i) Compute each of the following:
  (a) $d\ln(S_t)$
  (b) $d(S_t\,U_t)$
  (c) $d(S_t/U_t)$
  (d) $d((S_t)^\alpha\,(U_t)^\beta)$ for $\alpha \neq 0; \beta \neq 0$.

(ii) Solve the system of SDEs (2.1).

(iii) What is the distribution of $Y = U_T S_T$ for a fixed $T > 0$?

EXERCISE 2.23. Suppose X_t satisfies the SDE:

$$dX_t = \kappa(\theta - X_t)\,dt + \sigma\,dW_t\ , \qquad X_0 = x\ .$$

show that the following is a solution

$$X_0 = \theta + (x - \theta)\,e^{-\kappa t} + \sigma \int_0^t e^{-\kappa(t-u)} dW_u$$

EXERCISE 2.24. Let W be a standard Brownian motion and define:

$$S_t = \exp\left[\left(\mu - \tfrac{1}{2}\sigma^2\right)t + \eta W_t\right]$$

Find the SDE that S_t satisfies.

EXERCISE 2.25. Suppose that an asset, S, satisfies:

$$\frac{dS_t}{S_t} = r\ dt + \sigma\ dW_t^{\mathbb{Q}}$$

Solve the SDE by introducing $X_t = \ln S_t$ and applying Itô's lemma.

EXERCISE 2.26. Let X_t satisfy:

$$\frac{dX_t}{X_t} = \mu_t\ dt + \sigma_t\ dW_t$$

What is the drift and volatility of X_t?

EXERCISE 2.27. Let $\ln X_t$ satisfy:

$$d\ln X_t = \mu(t)\ dt + \sigma(t)\ dW_t$$

where $\mu(t), \sigma(t)$ are deterministic functions.

(i) Solve for X_t and compute $\mathbb{E}(X_t)$ and $\mathbb{V}(X_t)$
(ii) What SDE does X_t satisfy?

EXERCISE 2.28. Let X and Y satisfy:

$$\frac{dX_t}{X_t} = \mu_t \; dt + \sigma_t \; dW_t$$
$$\frac{dY_t}{Y_t} = \nu_t \; dt + \eta_t \; dB_t$$

where $d[W,B]_t = \rho \; dt$. Find the SDEs that $X_t Y_t$ and X_t/Y_t satisfy.

EXERCISE 2.29. Determine the joint distribution of r_T and $\int_0^T r_s \; ds$ for the Vasicek model:

$$dr_t = \kappa(\theta - r_t) \; dt + \sigma \; dW_t$$

3 Feynman-Kac Theorem

The Feynman-Kac theorem establishes an interesting link between partial differential equations and stochastic processes. It can be viewed in two ways:

- It provides a solution to a certain class of PDEs, written in terms of an expectation involving a related stochastic process
- It gives a way of computing certain expectations by solving an associated PDE.

Theorem 3.1 (Feynman-Kac). *The solution to the partial differential equation*

$$\begin{cases} \partial_t h + a(t,x)\cdot\partial_x h + \frac{1}{2}b(t,x)^2\cdot\partial_{xx}h &= c(t,x)\cdot h(t,x) \\ h(T,x) &= H(x) \end{cases}$$

admits a stochastic representation given by:

$$h(t,x) \;=\; \mathbb{E}^{\mathbb{P}^*}_{t,x}\left[H(X_T)\cdot e^{-\int_t^T c(s,X_s)\;ds}\right]$$

where $\mathbb{E}_{t,x}[\;\cdot\;] = \mathbb{E}[\;\cdot\;|X_t = x]$ *and the process* $X = (X_t)_{t\geq 0}$ *satisfies the SDE:*

$$dX_t \;=\; a(t,X_t)\;dt + b(t,X_t)\;dW_t^{\mathbb{P}^*}$$

where $W^{\mathbb{P}^*} = \left(W_t^{\mathbb{P}^*}\right)_{t\geq 0}$ *is a standard Brownian motion under* $\mathbb{P}^*$*.*

Before presenting the proof, let us discuss the statement of the theorem:

- When confronted with a PDE of the form above, we can define a (fictitious) process X with drift and volatility given by the processes $a(t,X_t)$ and $b(t,X_t)$, respectively.
- Thinking of c as a "discount factor" we then consider the conditional expectation of the discounted terminal condition, $H(X_T)$, given that the value of X at time t is equal to a known value, x. Clearly, this conditional expectation is a function of t and x; for every value of t and x we have some conditional expectation value.
- This function (the conditional expectation as a function of t and x) is precisely the solution to the PDE we started with.

The alternative view of this theorem is to say that if we had a process with known dynamics, such as X, and we are interested in computing the conditional expectation above, then this can be done by solving the associated PDE.

The key to the proof of the Feynman-Kac theorem is the notion of a Doob martingale:

THEOREM 3.2 (Doob martingale). *Let X be an $\mathcal{F}_T$-measurable random variable with $\mathbb{E}[|X|] < \infty$. Define the process $M = (M_t)_{0 \le t \le T}$ as follows:*

$$M_t \;=\; \mathbb{E}[X|\mathcal{F}_t]$$

*Then M is a martingale with respect to the filtration $\mathcal{F}_t$, and the process M is referred to as a **Doob martingale**.*

Proof To show that M is a martingale with respect to $\mathcal{F}_t$ we need to show that:

(i) M is $\mathcal{F}_t$-adapted, i.e. M_t is $\mathcal{F}_t$-measurable for all t,
(ii) $\mathbb{E}[|M_t|] < \infty$ for all t,
(iii) $\mathbb{E}[M_t|\mathcal{F}_s] = M_s$ for all $s < t$.

To show (i) notice that the conditional expectation $\mathbb{E}[X|\mathcal{F}]$ is an $\mathcal{F}$-measurable random variable by definition. To show (ii) notice that:

$$\begin{aligned}
\mathbb{E}[|M_t|] &= \mathbb{E}[|\mathbb{E}[X|\mathcal{F}_t]|] && \text{by definition} \\
&\le |\mathbb{E}[\mathbb{E}[X|\mathcal{F}_t]]| && \text{by Jensen's inequality} \\
&= |\mathbb{E}[X]| && \text{by tower property of expectations} \\
&< \infty && \text{by assumption}
\end{aligned}$$

To show (iii) notice that:

$$\begin{aligned}
\mathbb{E}[M_t|\mathcal{F}_s] &= \mathbb{E}[\mathbb{E}[X|\mathcal{F}_t]|\mathcal{F}_s] \\
&= \mathbb{E}[X|\mathcal{F}_s] && \text{“smallest } \sigma\text{-algebra wins”} \\
&= M_s
\end{aligned}$$

□

Now, we can turn to the proof the main theorem:

Proof The goal is to show that the function

$$h(t,x) \;=\; \mathbb{E}^{\mathbb{P}^*}_{t,x}\left[H(X_T)\cdot e^{-\int_t^T c(s,X_s)\,ds}\right]$$

satisfies the given PDE along with the terminal condition. Notice first that the function satisfies the terminal condition $h(T,x) = H(x)$ since:

$$h(T,x) \;=\; \mathbb{E}^{\mathbb{P}^*}_{T,x}\left[H(X_T)\cdot e^{-\int_T^T c(s,X_s)\,ds}\right] \;=\; \mathbb{E}^{\mathbb{P}^*}\left[H(X_T)|X_T = x\right] \;=\; H(x)$$

Next, we need to show that the function satisfies the PDE. Define a process $h = (h_t)_{0 \le t \le T}$ in terms of the process X and the function h as follows:

$$h_t \;=\; h(t,X_t) \;=\; \mathbb{E}^{\mathbb{P}^*}_{t,X_t}\left[H(X_T)\cdot e^{-\int_t^T c(s,X_s)\,ds}\right]$$

Multiplying the expressions above by $e^{-\int_0^t c(s,X_s)\ ds}$, which is $\mathcal{F}_t$-measurable, and defining a new process, $\tilde{h}$, we have:

$$\tilde{h}_t \;=\; h_t \cdot e^{-\int_0^t c(s,X_s)\ ds} \;=\; \mathbb{E}^{\mathbb{P}^*}_{t,X_t}\left[H(X_T)\cdot e^{-\int_0^T c(s,X_s)\ ds}\right]$$

Notice now that the RHS of the equation above involves a conditional expectation of an $\mathcal{F}_t$-measurable random variable (since the term in the expectation does not involve t and both X_T and the integral appearing in the exponential are $\mathcal{F}_T$-measurable). Therefore, the term on the RHS, and hence the process $\tilde{h}$, is a Doob martingale.

Now applying Itô's lemma on either side of the equation yields:

$$\begin{aligned}
d\tilde{h}_t \;&=\; d\left(h_t \cdot e^{-\int_0^t c(s,X_s)\ ds}\right) \\
&=\; dh_t \cdot e^{-\int_0^t c(s,X_s)\ ds} + h_t \cdot d\left(e^{-\int_0^t c(s,X_s)\ ds}\right) + \underbrace{d\left[h, e^{-\int_0^t c(s,X_s)\ ds}\right]_t}_{=0,\text{ since integral is diff. in } t} \\
&\qquad\qquad\text{by Itô's product rule} \\
&=\; e^{-\int_0^t c(s,X_s)\ ds}\cdot\left(\partial_t h(t,X_t) + a(t,X_t)\cdot\partial_x h(t,X_t) + \frac{1}{2}b(t,X_t)^2\cdot\partial_{xx}h(t,X_t)\right)dt \\
&\qquad + e^{-\int_0^t c(s,X_s)\ ds}\cdot b(t,X_t)\cdot\partial_x h(t,X_t)\ dW_t^{\mathbb{P}^*} - h_t\cdot c(t,X_t)\cdot e^{-\int_0^t c(s,X_s)\ ds}\ dt \\
&\qquad\qquad\text{by Itô's lemma and differentiating the exponential term} \\
&=\; e^{-\int_0^t c(s,X_s)\ ds}\cdot\left((\partial_t + \mathcal{L})h_t - h_t\cdot c(t,X_t)\right)dt + e^{-\int_0^t c(s,X_s)\ ds}\cdot b(t,X_t)\cdot\partial_x h\ dW_t^{\mathbb{P}^*}
\end{aligned}$$

where the operator $\mathcal{L} = a(t,x)\cdot\partial_x + \frac{1}{2}b(t,x)^2\cdot\partial_{xx}$ is called the generator of the process X.

Next, we integrate both sides of the equation from t to $t+\Delta t$:

$$\begin{aligned}
\tilde{h}_{t+\Delta t} - \tilde{h}_t \;&=\; \int_t^{t+\Delta t} e^{-\int_0^u c(s,X_s)\ ds}\cdot\left((\partial_t+\mathcal{L})h_u - h_u\cdot c(u,X_u)\right)du \\
&\qquad + \int_t^{t+\Delta t} e^{-\int_0^u c(s,X_s)\ ds}\cdot b(u,X_u)\cdot\partial_x h\ dW_u^{\mathbb{P}^*}
\end{aligned}$$

Now, taking conditional expectations on both sides and recalling that $\tilde{h}$ and the

Itô integral on the RHS are martingales we have:

$$
\begin{aligned}
0 &= \mathbb{E}_t\left[\int_t^{t+\Delta t} e^{-\int_0^u c(s,X_s)\,ds}\cdot((\partial_t+\mathcal{L})h_u - h_u\cdot c(u,X_u))\,du\right] \\
&= \lim_{\Delta t\downarrow 0}\frac{1}{\Delta t}\,\mathbb{E}_t\left[\int_t^{t+\Delta t} e^{-\int_0^u c(s,X_s)\,ds}\cdot((\partial_t+\mathcal{L})h_u - h_u\cdot c(u,X_u))\,du\right] \\
&= \mathbb{E}_t\left[\lim_{\Delta t\downarrow 0}\frac{1}{\Delta t}\int_t^{t+\Delta t} e^{-\int_0^u c(s,X_s)\,ds}\cdot((\partial_t+\mathcal{L})h_u - h_u\cdot c(u,X_u))\,du\right] \\
&\qquad \text{by dominated convergence theorem} \\
&= \mathbb{E}_t\left[e^{-\int_0^t c(s,X_s)\,ds}\cdot((\partial_t+\mathcal{L})h_t - h_t\cdot c(t,X_t))\right] \\
&\qquad \text{by fundamental theorem of calculus} \\
&= e^{-\int_0^t c(s,X_s)\,ds}\cdot((\partial_t+\mathcal{L})h_t - h_t\cdot c(t,X_t)) \\
&\qquad \text{since random variable is } \mathcal{F}_t\text{-measurable} \\
&= (\partial_t+\mathcal{L})h_t - h_t\cdot c(t,X_t) \\
&= \partial_t h(t,X_t) + a(t,X_t)\cdot\partial_x h(t,X_t) + \frac{1}{2}b(t,X_t)^2\cdot\partial_{xx}h(t,X_t) - h(t,X_t)\cdot c(t,X_t)
\end{aligned}
$$

But this result must hold for every path of X_t, therefore for all (t,x) we must have

$$
0 = \partial_t h(t,x) + a(t,x)\cdot\partial_x h(t,x) + \frac{1}{2}b(t,x)^2\cdot\partial_{xx}h(t,x) - c(t,x)\cdot h(t,x)
$$

□

3.1 Exercises

Exercise 3.3. Give a Feynman-Kac representation of the solution to the following PDE:

$$
\begin{cases}
\partial_t f + \partial_x f + 2\cdot\partial_{xx} f = 3\cdot f(t,x) \\
f(1,x) = x^2
\end{cases}
$$

4 Girsanov's Theorem

Let's say that you have a Brownian motion, W, defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and that you use it to define a new process by adding a drift term:

$$W_t^* = \int_0^t \lambda(u, W_u)\, du + W_t$$

Clearly, if λ is not identically zero then W^* cannot be a Brownian motion under the measure $\mathbb{P}$ (for one thing, there will be values of t where the expectation of W^* is nonzero). The main idea with Girsanov's theorem is that it identifies a *different* probability measure under which W^* *is* a standard Brownian motion.

First, we need to define the notion of a **Radon-Nikodym derivative**, which is an object that relates two probability measures:

THEOREM 4.1 (Radon-Nikodym). *Let $\mathbb{P}$ and $\mathbb{P}^*$ be two probability measures defined on the same measurable space $(\Omega, \mathcal{F})$ such that $\mathbb{P}^*$ is absolutely continuous with respect to $\mathbb{P}$, i.e. $\mathbb{P}(A) = 0 \implies \mathbb{P}^*(A) = 0$. Then, there exists a random variable, $\frac{d\mathbb{P}^*}{d\mathbb{P}}$, such that for any random variable X:*

$$\mathbb{E}^{\mathbb{P}^*}[X] \;=\; \mathbb{E}^{\mathbb{P}}\left[X \cdot \frac{d\mathbb{P}^*}{d\mathbb{P}}\right]$$

*This random variable is known as a **Radon-Nikodym derivative** and must satisfy:*

(i) $\frac{d\mathbb{P}^*}{d\mathbb{P}} > 0$, $\mathbb{P}$-*a.s.*

(ii) $\mathbb{E}^{\mathbb{P}}\left[\frac{d\mathbb{P}^*}{d\mathbb{P}}\right] = 1$

Furthermore, if $\mathbb{P}$ is also absolutely continuous with respect to $\mathbb{P}^*$, i.e. the two probability measures are equivalent, $\mathbb{P} \sim \mathbb{P}^*$, then there exists another random variable, $\frac{d\mathbb{P}}{d\mathbb{P}^*}$, such that:

$$\mathbb{E}^{\mathbb{P}}[X] = \mathbb{E}^{\mathbb{P}^*}\left[X \cdot \frac{d\mathbb{P}}{d\mathbb{P}^*}\right] \qquad \textit{and} \qquad \frac{d\mathbb{P}}{d\mathbb{P}^*} = \left(\frac{d\mathbb{P}^*}{d\mathbb{P}}\right)^{-1}$$

We omit the proof of the main result, but we can justify the two properties:

(i) By taking X to be the indicator of some event A, $X = \mathbb{1}_A$, we have that

$$\mathbb{E}^{\mathbb{P}^*}[\mathbb{1}_A] \;=\; \mathbb{P}^*(A) \;=\; \mathbb{E}^{\mathbb{P}}\left[\mathbb{1}_A \cdot \frac{d\mathbb{P}^*}{d\mathbb{P}}\right]$$

Since probabilities must be nonnegative and $\mathbb{1}_A \geq 0$, it follows that $\frac{d\mathbb{P}^*}{d\mathbb{P}} > 0$, $\mathbb{P}$-a.s.

(ii) Taking $X = 1$ we have:

$$\mathbb{E}^{\mathbb{P}^*}[1] \;=\; 1 \;=\; \mathbb{E}^{\mathbb{P}}\left[\frac{d\mathbb{P}^*}{d\mathbb{P}}\right]$$

So, Radon-Nikodym derivatives allow us to compute expectations of random variables under one measure by computing expectations of a modified version of the random variables under a different measure. One reason we may want to do this is if the computation is easier under one measure compared to another, or if we don't have a complete characterization of a certain measure but we can relate it to another measure.

We can also define measure changes in the context of stochastic processes - in this case we will have a **Radon-Nikdoym process**:

Theorem 4.2 (Radon-Nikdoym process). *Let $\mathbb{P}$ and $\mathbb{P}^*$ be two probability measures defined on the same filtered measurable space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0})$ such that $\mathbb{P}^*$ is absolutely continuous with respect to $\mathbb{P}$. Then, there exists a stochastic process, $\eta = (\eta_t)_{t\geq 0}$ also denoted $\eta_t = \left(\frac{d\mathbb{P}^*}{d\mathbb{P}}\right)_t$, such that for any process $X = (X_t)_{t\geq 0}$ we have:*

$$\mathbb{E}^{\mathbb{P}^*}[X_t] \;=\; \mathbb{E}^{\mathbb{P}}[X_t \cdot \eta_t] \qquad \textit{and} \qquad \mathbb{E}_t^{\mathbb{P}^*}[X_u] \;=\; \frac{\mathbb{E}_t^{\mathbb{P}}[X_u \cdot \eta_u]}{\mathbb{E}_t^{\mathbb{P}}[\eta_u]} \qquad \textit{for } u > t$$

*where $\mathbb{E}_t[\;\cdot\;] = \mathbb{E}[\;\cdot\;|\mathcal{F}_t]$. Furthermore, the **Radon-Nikodym process** η must satisfy:*

(i) $\eta_0 = 1$
(ii) $\eta_t > 0$, *$\mathbb{P}$-a.s. for all $t \geq 0$*
(iii) *η is a $\mathbb{P}$-martingale.*

This can be viewed as a conditional version or a process-based version of the first measure change result.

Before tackle the main theorem in this note, we introduce the **Dolèans-Dade exponential**, which will be used to define the Radon-Nikodym derivative in Girsanov's theorem:

Definition 4.3 (Dolèans-Dade exponential). The **Dolèans-Dade exponential** (or **stochastic exponential**) associated with a process $X = (X_t)_{t\geq 0}$ is defined to be the solution to the SDE:

$$\frac{dY_t}{Y_t} = dX_t, \qquad Y_0 = 1$$

$$\text{i.e. } Y_t = \exp\left(X_t - X_0 - \frac{1}{2}[X,X]_t\right)$$

The process $Y = (Y_t)_{t\geq 0}$ is often denoted by $\mathcal{E}(X) = (\mathcal{E}(X)_t)_{t\geq 0}$. Note that this process is a martingale.

Now, we give the main theorem of this note:

THEOREM 4.4 (Girsanov's Theorem). *Let $W = (W_t)_{t\geq 0}$ be a $\mathbb{P}$-Brownian motion and define a new process, $W^* = (W_t^*)_{t\geq 0}$ by:*

$$W_t^* = \int_0^t \lambda_u \, du + W_t$$

where λ_t is an adapted process. In differential notation this can be written as:

$$dW_t^* = \lambda_t \, dt + dW_t$$

Then, there exists a probability measure, $\mathbb{P}^*$, such that W^* is a $\mathbb{P}^*$-Brownian motion, and $\mathbb{P}^*$ is defined through the Radon-Nikodym derivative process:

$$\begin{aligned}\eta_t = \left(\frac{d\mathbb{P}^*}{d\mathbb{P}}\right)_t &= \mathcal{E}\left(-\int_0^t \lambda_u \, dW_u\right) \\ &= \exp\left(-\int_0^t \lambda_u \, dW_u - \frac{1}{2}\int_0^t \lambda_u^2 \, du\right)\end{aligned}$$

Furthermore, the process η satisfies the SDE:

$$\frac{d\eta_t}{\eta_t} = -\lambda_t \, dW_t, \qquad \eta_0 = 1$$

Proof First, we need to ensure that η_t is a valid Radon-Nikodym derivative process. Clearly, $\eta_0 = 1$ and $\eta_t > 0$ for all t. To show that η is a $\mathbb{P}$-martingale notice that:

$$\ln \eta_t = -\int_0^t \lambda_u \, dW_u - \frac{1}{2}\int_0^t \lambda_u^2 \, du$$

or in differential notation:

$$d(\ln \eta_t) = -\frac{1}{2}\lambda_t^2 \, dt - \lambda_t \, dW_t$$

Applying Itô's lemma we have that:

$$\begin{aligned} d\eta_t &= -\lambda_t \eta_t \, dW_t \\ \implies \quad \eta_t &= -\int_0^t \lambda_u \eta_u \, dW_u \end{aligned}$$

which is an Itô integral under $\mathbb{P}$ and hence a $\mathbb{P}$-martingale. Next, we find that the distribution of W_t^* under $\mathbb{P}^*$ by computing its characteristic function using the measure change given by η_t:

$$\begin{aligned}
\mathbb{E}^{\mathbb{P}^*}\left[e^{ivW_t^*}\right] &= \mathbb{E}^{\mathbb{P}}\left[e^{ivW_t^*}\cdot\eta_t\right] \\
&= \mathbb{E}^{\mathbb{P}}\left[\exp\left[iv\left(\int_0^t \lambda_u\ du + W_t\right)\right]\cdot\exp\left[-\int_0^t \lambda_u\ dW_u - \frac{1}{2}\int_0^t \lambda_u^2\ du\right]\right] \\
&= \mathbb{E}^{\mathbb{P}}\left[\exp\left[iv\left(\int_0^t \lambda_u\ du + \int_0^t dW_u\right) - \int_0^t \lambda_u\ dW_u - \frac{1}{2}\int_0^t \lambda_u^2\ du\right]\right] \\
&= \mathbb{E}^{\mathbb{P}}\left[\exp\left[\int_0^t (iv\lambda_u - \frac{1}{2}\lambda_u^2)\ du - \int_0^t (iv-\lambda_u)\ dW_u\right]\right]
\end{aligned}$$

Now, let $b_t = iv - \lambda_t$ so that $b_t^2 = \lambda_t^2 - 2iv\lambda_t - v^2$. Subsitituting into the expression above:

$$\begin{aligned}
\mathbb{E}^{\mathbb{P}^*}\left[e^{ivW_t^*}\right] &= \mathbb{E}^{\mathbb{P}}\left[\exp\left[-\frac{1}{2}\int_0^t (b_u^2+v^2)\ du - \int_0^t b_u\ dW_u\right]\right] \\
&= e^{-\frac{1}{2}v^2t}\cdot\mathbb{E}^{\mathbb{P}}\left[\exp\left(-\frac{1}{2}\int_0^t b_u^2\ du - \int_0^t b_u\ dW_u\right)\right] \\
&= e^{-\frac{1}{2}v^2t}
\end{aligned}$$

and this is the characteristic function of a normal random variable with mean zero and variance t. Therefore, $W_t^* \sim N(0,t)$ under $\mathbb{P}^*$. We can show the independence and stationarity of increments under $\mathbb{P}^*$ using the characteristic function as well. □

This result also extends to the multidimensional case:

THEOREM 4.5 (Multidimensional Girsanov's theorem). *Let $\mathbf{W} = (\mathbf{W}_t)_{t\geq 0}$ be an n-dimensional vector of independent $\mathbb{P}$-Brownian motions, i.e. $\mathbf{W}_t = (W_t^1, ..., W_t^n)$, and define $\boldsymbol{\lambda} = (\boldsymbol{\lambda}_t)_{t\geq 0}$ as a vector-valued process of the same dimension, i.e. $\boldsymbol{\lambda}_t = (\lambda_t^1, ..., \lambda_t^n)$. Also, define a new process, $\mathbf{W}^* = (\mathbf{W}_t^*)_{t\geq 0}$ by:*

$$\mathbf{W}_t^* = \int_0^t \boldsymbol{\lambda}_u\ du + \mathbf{W}_t$$

In differential notation this can be written as:

$$d\mathbf{W}_t^* = \boldsymbol{\lambda}_t\ dt + d\mathbf{W}_t$$

i.e. λ^i is the drift adjustment for W^i:

$$W_t^{*,i} = \int_0^t \lambda_u^i\ du + W_t^i \qquad , \qquad dW_t^{*,i} = \lambda_t^i\ dt + dW_t^i$$

Then, there exists a probability measure, $\mathbb{P}^*$, such that W^* is a $\mathbb{P}^*$-Brownian motion, and $\mathbb{P}^*$ is defined through the Radon-Nikodym derivative process:

$$\begin{aligned}
\eta_t = \left(\frac{d\mathbb{P}^*}{d\mathbb{P}}\right)_t &= \mathcal{E}\left(-\int_0^t \boldsymbol{\lambda}_u'\ d\mathbf{W}_u\right) \\
&= \exp\left(-\int_0^t \boldsymbol{\lambda}_u'\ d\mathbf{W}_u - \frac{1}{2}\int_0^t \boldsymbol{\lambda}_u'\boldsymbol{\lambda}_u\ du\right)
\end{aligned}$$

Furthermore, the process η satisfies the SDE:

$$\begin{aligned}\frac{d\eta_t}{\eta_t} &= -\boldsymbol{\lambda}_t'\, d\mathbf{W}_t, \qquad \eta_0 = 1 \\ &= -\left(\lambda_t^1\, dW_t^1 + \lambda_t^2\, dW_t^2 + \;\cdots\; + \lambda_t^n\, dW_t^n\right)\end{aligned}$$

The result also extends to the case of correlated Brownian motions:

THEOREM 4.6 (Correlated multidimensional Girsanov's theorem). *Let* $\mathbf{W} = (\mathbf{W}_t)_{t\geq 0}$ *be an* n*-dimensional vector of* <u>*correlated*</u> $\mathbb{P}$*-Brownian motions with correlation matrix* $\boldsymbol{\rho}$*, i.e.* $\mathbf{W}_t = (W_t^1, ..., W_t^n)$ *with* $d[W^i, W^j]_t = \boldsymbol{\rho}_{ij} dt$*, and define* $\boldsymbol{\lambda} = (\boldsymbol{\lambda}_t)_{t\geq 0}$ *as a vector-valued process of the same dimension, i.e.* $\boldsymbol{\lambda}_t = (\lambda_t^1, ..., \lambda_t^n)$*. Also, define a new process,* $\mathbf{W}^* = (\mathbf{W}_t^*)_{t\geq 0}$ *by:*

$$\mathbf{W}_t^* = \int_0^t \boldsymbol{\rho}\boldsymbol{\lambda}_u \, du + \mathbf{W}_t$$

In differential notation this can be written as:

$$d\mathbf{W}_t^* = \boldsymbol{\rho}\boldsymbol{\lambda}_t \, dt + d\mathbf{W}_t$$

Then, there exists a probability measure, $\mathbb{P}^*$*, such that* W^* *is a* $\mathbb{P}^*$*-Brownian motion, and* $\mathbb{P}^*$ *is defined through the Radon-Nikodym derivative process:*

$$\begin{aligned}\eta_t = \left(\frac{d\mathbb{P}^*}{d\mathbb{P}}\right)_t &= \mathcal{E}\left(-\int_0^t \boldsymbol{\lambda}_u' \, d\mathbf{W}_u\right) \\ &= \exp\left(-\int_0^t \boldsymbol{\lambda}_u' \, d\mathbf{W}_u - \frac{1}{2}\int_0^t \boldsymbol{\lambda}_u'\boldsymbol{\rho}\boldsymbol{\lambda}_u \, du\right)\end{aligned}$$

Furthermore, the process η *satisfies the SDE:*

$$\begin{aligned}\frac{d\eta_t}{\eta_t} &= -\boldsymbol{\lambda}_t' \, d\mathbf{W}_t, \qquad \eta_0 = 1 \\ &= -\left(\lambda_t^1 \, dW_t^1 + \lambda_t^2 \, dW_t^2 + \quad \cdots \quad + \lambda_t^n \, dW_t^n\right)\end{aligned}$$

4.1 Exercises

EXERCISE 4.7. Assume an asset S has $\mathbb{Q}$-dynamics:

$$\frac{dS_t}{S_t} = r_t \, dt + \sigma_t \, dW_t^{\mathbb{Q}}$$

and $\mathbb{Q}^S$-dynamics:

$$\frac{dS_t}{S_t} = (r_t + \sigma_t^2) \, dt + \sigma_t \, dW_t^{\mathbb{Q}^S}$$

(i) Write $W_t^{\mathbb{Q}^S}$ in terms of $W_t^{\mathbb{Q}}$, i.e. find the relevant drift correction.
(ii) What is the Radon-Nikdoym derivative, $\frac{d\mathbb{Q}^S}{d\mathbb{Q}}$, connecting these two measures?

EXERCISE 4.8. Assume that:

$$\frac{dN_t}{N_t} = r_t \, dt + \sigma_t^N \, dW_t^{\mathbb{Q}} \,, \qquad N_t > 0 \;\; \forall t \;\; \text{a.s.}$$

(i) Show that $\eta_t = \frac{N_t/N_0}{M_t/M_0}$ is a Radon-Nikodym process, where M is the bank account.

(ii) Show that if V_t is the price process of a traded asset, then

$$\mathbb{E}^{\mathbb{Q}^N}\left[\frac{V_T}{N_T}\,\middle|\,\mathcal{F}_t\right] = \frac{V_t}{N_t}$$

Part II

Derivative Pricing

5 Introduction to Financial Derivatives

This chapter introduces some basic concepts and terminology that we will encounter when tackling problems involving the pricing of financial derivatives along with a few common examples.

5.1 Financial Derivatives

Financial derivatives, also referred to as **contingent claims**, are financial instruments that *derive* their value from one or more *external sources*, referred to as **underlying process(es)**. These underlying processes may or may not be traded assets. For example, a contingent claim can be *written* - a term used to describe the selling of a claim - on a traded stock that pays the holder \$1 if the stock price is above a certain value on a given date. Another contingent claim can entitle the holder to a payment of \$1 if the cumulative snowfall between November and February is over a certain level (this is an example of a weather derivative). In the first case the underlying is a traded asset while in the latter case the underlying is the weather which is clearly not tradeable.

Any financial derivative requires two counterparties: the **buyer** or **holder** of the claim, who is said to be in a **long position** with respect to the claim, receives the payoff from the **seller** or **writer** of the claim, who is said to be in a **short position**. Mathematically, we identify financial derivatives by their **payoff function**, which describes the payment that the claim seller will make to the buyer in terms of the underlying as well as the conditions under which the payment is to be made.

For example, let $A = (A_t)_{t\geq 0}$ and $B = (B_t)_{t\geq 0}$ be stochastic processes representing the prices of two assets. Suppose we have a claim that pays A_T^2 at time T. This means that at time T the claim seller must pay the buyer the square of the prevailing asset price, i.e. the payoff function is $f(a) = a^2$. This payment is made regardless of the value of A_T. Another example is a claim that pays $(A_T - B_T)_+ = \max(A_T - B_T, 0)$ at time T in which case the payoff function is $g(a, b) = (a - b)_+$ and for which a payment is made only if the price of asset A is greater than the price of asset B at time T.

The time at which the payment occurs is known as the **maturity date**, which in the previous two examples was T. It is possible for a derivative to involve multiple payments made by either party, though typically those different streams of payments would be stripped into individual components that are priced separately. It is also possible for the time of the payment to differ from the time used to determine the amount being paid, for example a claim that pays $(A_{T_1} - K)_+$ at time T_2 where K is a fixed constant and $T_1 < T_2$. A contingent claim can also depend on the value of the underlying process at multiple points in time, e.g. a claim that pays the average price of an asset taken at N different time points, $\frac{1}{N}\sum_{i=1}^{N} A_{T_i}$, at time $T_N > T_{N-1} > \cdots > T_1$, or even a continuum of times, e.g. a claim that pays $\frac{1}{T}\int_0^T A_t \, dt$ at time T.

The common thread between all of these examples is that the (random) amount being paid at a given time must be based on information up to and including the time of the payment. In other words, the payment amount may be unknown to the parties at the time the contract is entered and remain unknown through the life of the contract, but it must be determined with full certainty at the time the payment is due. This somewhat obvious statement can be expressed mathematically as follows:

DEFINITION 5.1. A **payoff function** for a payment of a financial derivative made at time T is an $\mathcal{F}_T$-measurable random variable, where $\mathcal{F} = \{\mathcal{F}_t\}_{t \geq 0}$ is the filtration generated by the underlying processes in the derivative.

In all of the examples above, and for any given financial derivative with some payoff function, the main point of interest is: *what amount should be paid by the buyer of the claim to the seller in order to enter the contract?* This will be the main question we address in the second part of the book.

Next, we introduce some common financial derivatives: call and put options.

5.1.1 European Call and Put Options

Call and put options are among the simplest and most commonly traded financial derivatives. A **European call option** with strike price K gives the option holder the right (but not the obligation) to *buy* the underlying asset at the strike price K at some future maturity time T. That is, the payoff function at time T is given by $(A_T - K)_+ = \max(A_T - K,\ 0)$. Similarly, a **European put option** gives the option holder the right (but not the obligation) to *sell* the underlying asset at the strike price K at some future maturity time T. That is, the payoff function at time T is given by $(K - A_T)_+ = \max(K - A_T,\ 0)$. An option is said to be **in-the-money** (ITM), resp. **out-of-the-money** (OTM), if the current price of the underlying asset is at a level where the payoff would be positive, resp. zero. Clearly, these regions are different for call and put options. If the current price of the underlying is equal to the strike price, the option is said to be **at-the-money** (ATM). A common use of terminology is to refer to "writing an at-the-money option," meaning an option whose strike price is set to be equal to the prevailing price of the underlying asset the option is written on.

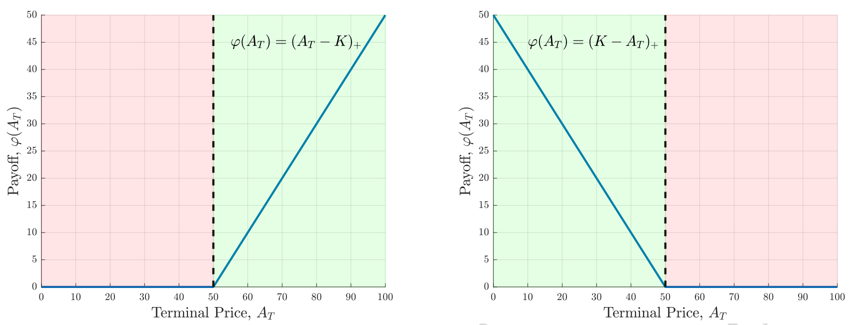


Figure 5.1 Payoff function for a European call option (left) and European put option (right) with a strike price of $K = \$50$ as a function of the underlying asset price. The green (red) shaded area indicates where the option is ITM (OTM); the dashed black line indicates where the option is ATM.

5.1.2 Early Exercise and American Options

Some options allow for **early exercise**, where the option holder forces the seller to make the payoff at a time of their choosing. Such options are referred to as **American-style options**. For example, in an American put option, the option holder can *exercise* the option at any time up to and including the maturity date T. Thus, if they decide to exercise at some time $\tau < T$ then they will receive $(K - S_\tau)_+$ from the option seller at time τ. Note that although the payment time τ is now random, the payment amount is still fully determined by information up to this time (that is, the payoff function is $\mathcal{F}_\tau$-measurable).

6 Fundamental Theorem of Asset Pricing

The notions of arbitrage and the Fundamental Theorem of Asset Pricing (FTAP) are among the most important concepts in mathematical finance, and form the foundation for derivative pricing. In this chapter, we explore these two notions in a fairly general context.

6.1 Arbitrage

Colloquially, arbitrage is often referred to as a "riskless profit." In our context, we have a precise mathematical definition for the notion of arbitrage. First, we will assume that our economy can be described by a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\in\mathcal{T}}, \mathbb{P})$, where:

- Ω is the set of possible outcomes,
- $\mathcal{F}$ is the σ-algebra describing the set of possible events,
- $\mathcal{F}_t$ is a filtration describing the flow of information
- $\mathbb{P}$ is the *physical* or *real-world* probability measure.

We will further assume that the economy consists of n assets whose values at time t are given by $A_t^{(i)}$ for $i = 1, ..., n$.

Before we can define the notion of arbitrage, we must introduce the concept of a portfolio (or strategy) and, more specifically, a *self-financing* portfolio.

Definition 6.1 (Portfolio). A **portfolio** (or **strategy**) is an $\mathcal{F}_t$-adapted vector-valued process $\alpha_t = \left(\alpha_t^{(1)}, ..., \alpha_t^{(n)}\right)$, where $\alpha_t^{(i)}$ represents the holdings in asset i at time t. The value of the portfolio at time t is given by:

$$V_t = \sum_{i=1}^{n} \alpha_t^{(i)} A_t^{(i)}$$

The requirement that α_t be $\mathcal{F}_t$-adapted corresponds to the fact that we must be able to know what we are investing in at any given time.

Definition 6.2 (Self-financing portfolio). A **self-financing portfolio** (or **strat-**

egy) is a portfolio that satisfies:

$$dV_t = \sum_{i=1}^{n} \alpha_t^{(i)} dA_t^{(i)}$$

Intuitively, this means that changes to the portfolio value over very small time intervals can only be due to changes in the asset values. Alternatively, we can say that the purchase of new shares of any asset during rebalancing is *financed* by sales of shares already held in the portfolio and that there are no infusions (or extractions) of cash at any point in time.

Now, we are ready to define the notion of arbitrage:

DEFINITION 6.3 (Arbitrage portfolio). A portfolio V is an **arbitrage portfolio** if it is self-financing and its value process satisfies:

(i) $V_0 = 0$
(ii) There exists a time $t > 0$ such that:
  (a) $\mathbb{P}(V_t \geq 0) = 1$
  (b) $\mathbb{P}(V_t > 0) > 0$.

Examining the definition above we find that an arbitrage portfolio is a costless portfolio that we can hold up to a certain time t, at which point we are guaranteed to not suffer a loss *and* have a positive probability of making a profit.

The definition above is valid in continuous-time and discrete-time settings. In cases where we only have two periods the definition reduces to:

(i)	$V_0 = 0$	(the portfolio costs nothing initially)
(ii)	$\mathbb{P}(V_1 \geq 0) = 1$	(the portfolio never loses in any scenario)
(iii)	$\mathbb{P}(V_1 > 0) > 0$	(the portfolio can win in at least one scenario)

6.2 Fundamental Theorem of Asset Pricing

Of particular interest is whether or not an economy admits any arbitrage opportunities. This is precisely what is addressed by the FTAP:

THEOREM 6.4 (Fundamental Theorem of Asset Pricing). *An economy does not admit any arbitrage opportunities if and only if there exists:*

(i) *a **numeraire** asset, i.e. a traded asset B satisfying $\mathbb{P}(B_t > 0) = 1$ for all t;*
(ii) *an associated probability measure $\mathbb{Q}^B$ equivalent to the physical measure $\mathbb{P}$,*

such that the prices of all traded assets relative to B are martingales under $\mathbb{Q}^B$, i.e. for any traded asset A we have:

$$\frac{A_t}{B_t} = \mathbb{E}^{\mathbb{Q}^B}\left[\frac{A_u}{B_u}\middle|\mathcal{F}_t\right] \quad \textit{for } u > t$$

where $\mathcal{F}_t$ is the σ-algebra generated by all asset paths up to time t.

Proof ($\Leftarrow$) Assume there exists a self-financing strategy with corresponding value process $(V_t)_{t\geq 0}$ and a time t such that:

(i) $\mathbb{P}(V_t \geq 0) = 1 \underset{\text{since } B_t > 0 \text{ a.s.}}{\Longrightarrow} \mathbb{P}\left(\frac{V_t}{B_t} \geq 0\right) = 1 \underset{\text{since } \mathbb{Q}^B \sim \mathbb{P}}{\Longrightarrow} \mathbb{Q}^B\left(\frac{V_t}{B_t} \geq 0\right) = 1$

(ii) $\mathbb{P}(V_t > 0) > 0 \underset{\text{since } B_t > 0 \text{ a.s.}}{\Longrightarrow} \mathbb{P}\left(\frac{V_t}{B_t} > 0\right) > 0 \underset{\text{since } \mathbb{Q}^B \sim \mathbb{P}}{\Longrightarrow} \mathbb{Q}^B\left(\frac{V_t}{B_t} > 0\right) > 0$

The two statements above imply that $\mathbb{E}^{\mathbb{Q}^B}\left[\frac{V_t}{B_t}\right] > 0$, since the relative process is nonnegative almost surely with positive mass on the positive reals.

Next, applying the martingale property on the portfolio (since it is a traded asset), we find that:

$$\frac{V_0}{B_0} = \mathbb{E}^{\mathbb{Q}^B}\left[\frac{V_t}{B_t}\middle|\mathcal{F}_0\right]$$

Rearranging, we have:

$$V_0 = B_0 \cdot \mathbb{E}^{\mathbb{Q}^B}\left[\frac{V_t}{B_t}\middle|\mathcal{F}_0\right] > 0$$

The positivity follows since we have shown the expectation to be positive, and since $B_t > 0$ a.s. Since $V_0 > 0$, V cannot be an arbitrage portfolio as that would require $V_0 = 0$. Therefore, no arbitrage opportunities exist in the economy.

($\Rightarrow$) Omitted. □

A few notes on the above theorem:

- The probability measure $\mathbb{Q}^B$ is referred to as an **equivalent martingale measure** (EMM) or a **risk-neutral measure**; the B superscript signifies that the measure is induced by using B a numeraire.
- EMMs induced by certain assets have specific names, e.g. the EMM associated with using the money market account as a numeraire is referred to the **risk-neutral measure**[1] often denoted $\mathbb{Q}$; the EMM induced by the T-maturity bond is referred to as the **T-forward neutral measure**.
- The equivalence of the probability measures $\mathbb{P}$ and $\mathbb{Q}^B$ means that the measures agree on sets of measure zero, i.e. for a given event A we have: $\mathbb{P}(A) = 0 \iff \mathbb{Q}^B(A) = 0$. Intuitively, this means that events that have some probability of occurring in the real world cannot be ignored by the risk-neutral measure, and vice versa.
- It is possible for a single numeraire to give rise to multiple EMMs; the theorem only requires existence, not uniqueness, of the EMM. The uniqueness of the EMM is related to the notion of market completeness, which will be discussed later.

[1] Note the abuse of terminology here: "risk-neutral measure" is sometimes used to refer to a general EMM, and sometimes to the specific EMM induced by the money market account.

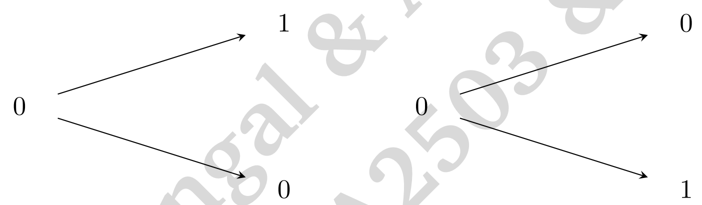
6.3 Complete and Incomplete Markets

Definition 6.5 (Complete market). A **complete market** is one in which any contingent claim or cash-flow stream can be perfectly replicated/hedged by other assets in the economy at each point in time.

The significance of a complete market is that it implies that every contingent claim has a unique no-arbitrage price (equal to the replicating portfolio) at every state and time. This is implied by the law of one price. The main theorem regarding market completeness - sometimes referred to as the second FTAP - relates the notion to the EMM from the first FTAP:

Theorem 6.6 (Completeness). *A market is complete if and only if the equivalent martingale measure* $\mathbb{Q}$ *is unique.*

On the other hand, **incomplete markets allow for a range**[2] **of no-arbitrage prices** - one corresponding to each EMM. Moreover, it is possible for a market to be complete without the existence of an EMM. For example, consider the following one-period, two-asset economy:



Clearly, this market is complete since the payoff vectors of the two assets form a basis for $\mathbb{R}^2$. However, since the economy admits arbitrage opportunities, an EMM does not exist.

One final note to consider is that if we have multiple numeraires, the fact that each induces a different measure does not violate market completeness; the "uniqueness" in the theorem refers to the uniqueness of the EMM induced by *each* numeraire, i.e. for the theorem to hold each numeraire must give rise to a single EMM.

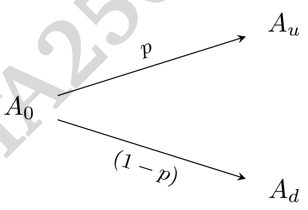
[2] Setting a single price requires the use of approaches such as minimum variance hedging, indifference pricing, etc.

7 Binomial Tree Models

In this note, we investigate a simple class of models for an economy known as *tree models*. These can be single-period or multi-period models in which, given a starting value, asset prices can take on one of a finite number possible future states. Our focus will be on *binomial tree models* in which, given a starting value, asset prices can take on one of two possible future states.

7.1 Single-Asset Binomial Model:

Consider an asset, A, whose price at time t we will denote by A_t. Assume that our economy has a single future period and that the price evolves according to the binomial model:



In this setup, the asset's current price is A_0, and its price at time $t = 1$, A_1, is a random variable; at time $t = 1$ the price of the asset will become A_u with probability p or A_d with probability $1 - p$. Notice that the value of p describes the physical probability measure $\mathbb{P}$ in this simple setting. Also, we assume that $p \in (0, 1)$ to avoid the trivial case where one one future state is possible.

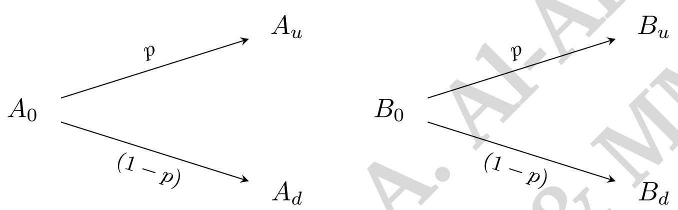
How can we determine current price of the asset, A_0, if its future (random) payoffs, A_u and A_d, are known? One seemingly sensible approach would be to compute the expectation of the future price under the physical measure $\mathbb{P}$, $\mathbb{E}^{\mathbb{P}}[A_1]$, and then discount it at some rate, r, to take into account the passage of time:

$$A_0 = \frac{1}{1+r} \cdot \mathbb{E}^{\mathbb{P}}[A_1] = \frac{1}{1+r} \left(A_u \cdot p + A_d \cdot (1-p) \right)$$

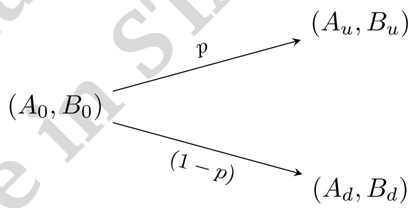
However this does not take into account the *utility* of investors. For example, they may be more concerned about a drop in prices than they would be pleased with a price increase (i.e. they are risk averse). An alternative approach would be to use certainty equivalence or indifference pricing[1] - however, we are more interested in economies with multiple assets.

7.2 Two-Asset Binomial Model:

Let us introduce a second asset, B, to the economy that behaves in a similar way to asset A:



Notice that the outcomes here are paired, meaning that there are two states of the world: one where the assets A and B have values A_u and B_u at time $t = 1$, and one where their values are A_d and B_d. An alternative way of drawing the above trees would be:



Now we can define portfolios in this economy:

Definition 7.1 (Portfolio). A **portfolio** (α, β) is an asset whose payoffs/prices are identical to holding α units in asset A and β units in asset B.

In other words, portfolios are linear combinations of the two assets. If we

[1] It is interesting to note that for investors with exponential utility, i.e. their utility function is given by $u(w) = -\frac{1}{\gamma}e^{-\gamma w}$ where w is wealth and γ is the risk aversion parameter, the price implied by certainty equivalence tends to $(A_u \cdot p + A_d \cdot (1 - p))$ as γ tends to 0. In other words, investors that have this utility function and are risk-neutral will price the asset as the expected future price.

denote the value of this portfolio at time t by V_t, then we can represent the behavior of the portfolio value in a binomial tree as follows:

$$\begin{tikzcd} & V_u = \alpha A_u + \beta B_u \\ V_0 = \alpha A_0 + \beta B_0 \arrow[ur, "p"] \arrow[dr, "(1-p)"'] & \\ & V_d = \alpha A_d + \beta B_d \end{tikzcd}$$

7.2.1 No-arbitrage conditions

We are now interested in deriving the conditions under which there is no arbitrage in the economy. Recall that in a one-period model an **arbitrage portfolio** is a portfolio that satisfies:

(i)	$V_0 = 0$	(the portfolio costs nothing initially)
(ii)	$\mathbb{P}(V_1 \geq 0) = 1$	(the portfolio never loses in any state)
(iii)	$\mathbb{P}(V_1 > 0) > 0$	(the portfolio can win in at least one state)

For example, of the three portfolios shown below, only the second satisfies all the conditions of an arbitrage portfolio:

$$\begin{tikzcd} & 2 & & & 1 \\ 1 \arrow[ur] \arrow[dr] & & & 0 \arrow[ur] \arrow[dr] & \\ & -1 & & & 1 \end{tikzcd}$$

$$\begin{tikzcd} & 2 \\ 1 \arrow[ur] \arrow[dr] & \\ & 2 \end{tikzcd}$$

The question of interest now is: **what conditions would A_0, A_u, A_d and B_0, B_u, B_d have to satisfy to ensure that the economy is arbitrage-free?** To answer this question we first construct a zero-cost portfolio, and then determine the conditions for the present and future asset values which ensure that the remaining two conditions for an arbitrage portfolio *cannot* be satisfied.

First, we need to ensure that the initial cost of our portfolio is zero, i.e. $V_0 = 0$:

$$V_0 \;=\; \alpha A_0 + \beta B_0 = 0 \quad \Longleftrightarrow \quad \beta \;=\; -\,\alpha \cdot \frac{A_0}{B_0}$$

So, given a choice of α, the no-cost condition imposes a value for β. Next, let us consider the binomial tree for the zero-cost portfolio $\left(\alpha, -\alpha \cdot \frac{A_0}{B_0}\right)$:

$$\begin{tikzcd}
 & V_u = \alpha A_u - \left(\alpha \cdot \frac{A_0}{B_0}\right) \cdot B_u \\
\underbrace{V_0 = 0}_{\text{by construction}} \arrow[ur] \arrow[dr] & \\
 & V_d = \alpha A_d - \left(\alpha \cdot \frac{A_0}{B_0}\right) \cdot B_d
\end{tikzcd}$$

A less cluttered version:

$$\begin{tikzcd}
 & \alpha\left(A_u - \frac{A_0}{B_0} \cdot B_u\right) \\
0 \arrow[ur] \arrow[dr] & \\
 & \alpha\left(A_d - \frac{A_0}{B_0} \cdot B_d\right)
\end{tikzcd}$$

To make sure that it is *not* possible for this to be an arbitrage portfolio we need one of the future states to be strictly positive and the other to be strictly negative.[2] We can see why this requirement is necessary by investigating the alternative: if both states are strictly positive, we immediately have an arbitrage opportunity. Likewise, if both are strictly negative, an arbitrage opportunity is available by shorting the portfolio at hand. If we assume[3] that $\alpha \neq 0$ then, after some rearranging, the above conditions reduce to:

Theorem 7.2 (No-arbitrage conditions for binomial model). *The one-period, two-asset binomial model economy described above admits no arbitrage opportunities if and only if*

$$\frac{A_d}{B_d} \;<\; \frac{A_0}{B_0} \;<\; \frac{A_u}{B_u} \qquad \mathit{OR} \qquad \frac{A_d}{B_d} \;>\; \frac{A_0}{B_0} \;>\; \frac{A_u}{B_u}$$

[2] Intuitively, this says that all zero-cost portfolios have some chance of winning and some chance of losing.

[3] We ignore the case where $\alpha = 0$ since it leads to $\beta = 0$ which gives a trivial portfolio that has zero value in all states, and is clearly not an arbitrage opportunity.

7.2.2 Relation to the Fundamental Theorem of Asset Pricing

The conditions for no-arbitrage derived in the previous section can be related to the FTAP. Consider a binomial tree for the relative price process:

$$
\begin{tikzcd}[column sep=large]
 & \frac{A_u}{B_u} \\
\frac{A_0}{B_0} \arrow[ur] \arrow[dr] & \\
 & \frac{A_d}{B_d}
\end{tikzcd}
$$

From the previous discussion, we know that for the economy to be arbitrage-free $\frac{A_0}{B_0}$ must lie between $\frac{A_d}{B_d}$ and $\frac{A_u}{B_u}$. Another way of stating this is to say that there exists $q^B \in (0,1)$ such that:

$$\frac{A_0}{B_0} \;=\; q^B \cdot \frac{A_u}{B_u} + (1-q^B) \cdot \frac{A_d}{B_d}$$

But given that q^B lies between 0 and 1, we can interpret q^B and $1-q^B$ as a probability[4] measure, $\mathbb{Q}^B$, defined on the two outcomes. Moreover, we can then think of the RHS of the last equation as an expectation under this new measure:

$$\frac{A_0}{B_0} \;=\; q^B \cdot \frac{A_u}{B_u} + (1-q^B) \cdot \frac{A_d}{B_d} \;=\; \mathbb{E}^{\mathbb{Q}^B}\left[\frac{A_1}{B_1}\right]$$

Also note that since $p, q^B \in (0,1)$ the probability measures $\mathbb{P}$ and $\mathbb{Q}^B$ are equivalent. Putting this information together we find that in this economy:

$$
\begin{aligned}
\nexists \text{ arbitrage} \quad &\iff \quad \frac{A_d}{B_d} < \frac{A_0}{B_0} < \frac{A_u}{B_u} \quad \text{OR} \quad \frac{A_d}{B_d} > \frac{A_0}{B_0} > \frac{A_u}{B_u} \\
&\iff \quad \exists\ \mathbb{Q}^B \sim \mathbb{P} \ \text{ such that } \frac{A_0}{B_0} \;=\; \mathbb{E}^{\mathbb{Q}^B}\left[\frac{A_1}{B_1}\right]
\end{aligned}
$$

This demonstrates how the FTAP (which applies in general cases) applies to the case of binomial models. In the discussion above, the probability measure $\mathbb{Q}^B$ is an equivalent martingale measure (or risk-neutral measure) whose existence guarantees that the economy is arbitrage-free. In particular, $\mathbb{Q}^B$ is the risk-neutral measure associated with the use of asset B as a numeraire.

Furthermore, we can **explicitly solve for the risk-neutral probability** q^B:

[4] Note that the probabilities q^B and $1-q^B$ are NOT the real-world probabilities associated with assets A and B going up and down; the probabilities for those events are still p and $1-p$. $(q^B, 1-q^B)$ is considered a probability measure since the values can be assigned to the up and down events, the values are nonnegative and sum up to 1.

Theorem 7.3 (Risk-neutral probabilities). *The risk-neutral probability associated with using B as the numeraire is given by:*

$$q^B \;=\; \frac{A_0/B_0 - A_d/B_d}{A_u/B_u - A_d/B_d}$$

Aside: notice that when we write the risk-neutral probabilities in this form, we can see immediately that the ordering conditions for no-arbitrage, $\frac{A_d}{B_d} < \frac{A_0}{B_0} < \frac{A_u}{B_u}$ OR $\frac{A_d}{B_d} > \frac{A_0}{B_0} > \frac{A_u}{B_u}$, would imply that $q^B \in (0,1)$ - as both numerator and denominator would have the same sign - and how all of these facts correspond to absence of arbitrage.

7.2.3 Change of Numeraire

We can repeat the entire exercise with A as the numeraire. Consider the other relative price process, $\frac{B_t}{A_t}$, and its corresponding binomial tree:

$$\begin{tikzcd} & \frac{B_u}{A_u} \\ \frac{B_0}{A_0} \arrow[ur] \arrow[dr] & \\ & \frac{B_d}{A_d} \end{tikzcd}$$

By the same reasoning as above, there would no arbitrage in the economy if and only if there exists $q^A \in (0,1)$ such that:

$$\frac{B_0}{A_0} \;=\; q^A \cdot \frac{B_u}{A_u} + (1-q^A)\cdot\frac{B_d}{A_d} \;=\; \mathbb{E}^{\mathbb{Q}^A}\left[\frac{B_1}{A_1}\right]$$

Here, $\mathbb{Q}^A$ is the risk-neutral measure associated with asset A, i.e. when A is the numeraire. Note that q^A and q^B are not necessarily equal; the superscripts are used to denote which asset we used as the numeraire. For the purposes of the FTAP we are only interested in the existence of at least one of these two measures to show that there is no arbitrage. However, we can ask the question: **how are q^A and q^B related?** Recall that q^A and q^B are given by:

$$q^A \;=\; \frac{B_0/A_0 - B_d/A_d}{B_u/A_u - B_d/A_d} \qquad\qquad q^B \;=\; \frac{A_0/B_0 - A_d/B_d}{A_u/B_u - A_d/B_d}$$

Now, it is possible to verify that (with some algebra and a little bit of patience...) that:

$$q^A \;=\; \frac{A_u/A_0}{B_u/B_0}\cdot q^B \qquad\qquad 1-q^A \;=\; \frac{A_d/A_0}{B_d/B_0}\cdot(1-q^B)$$

If we consider the two pre-multipliers from these two cases, $\frac{A_d/A_0}{B_d/B_0}$ and $\frac{A_u/A_0}{B_u/B_0}$, then we can think of these as two possible realizations of the random variable $\frac{A_1/A_0}{B_1/B_0}$, which we denote as:

$$\frac{d\mathbb{Q}^A}{d\mathbb{Q}^B} = \frac{A_1/A_0}{B_1/B_0}$$

This object is called a **Radon-Nikodym derivative** and it allows us to compute expectations under a different probability measure than our starting one:

$$\mathbb{E}^{\mathbb{Q}^A}[\ \cdot\ ] = \mathbb{E}^{\mathbb{Q}^B}\left[\ \cdot\ \left(\frac{d\mathbb{Q}^A}{d\mathbb{Q}^B}\right)\right]$$

In our case we would have:

$$\frac{B_0}{A_0} = \mathbb{E}^{\mathbb{Q}^A}\left[\frac{B_1}{A_1}\right] = \mathbb{E}^{\mathbb{Q}^B}\left[\frac{B_1}{A_1}\cdot\left(\frac{d\mathbb{Q}^A}{d\mathbb{Q}^B}\right)\right]$$

The first equality holds by the FTAP since A is the numeraire and our expectation is computed under $\mathbb{Q}^A$; the second equality holds since we are taking the expectation under $\mathbb{Q}^B$ of $\frac{B_1}{A_1}\cdot\left(\frac{d\mathbb{Q}^A}{d\mathbb{Q}^B}\right) = \frac{B_1}{A_1}\cdot\frac{A_1/A_0}{B_1/B_0} = B_0/A_0$, which is a constant (and equals the LHS of the equation). We will return to this concept later in the course.

7.3 Pricing Contingent Claims

Assume that we are interested in the value of a derivative[5]**/contingent claim C written on one of the two assets** that satisfies the following binomial tree:

$$\begin{tikzcd}[row sep=small] & C_u \\ C_0 \arrow[ur] \arrow[dr] & \\ & C_d \end{tikzcd}$$

There are two ways of pricing this contingent claim:

(i) Risk-neutral pricing:
If the economy is arbitrage-free, then the relative price process of the contingent claim - just like any asset in the economy - would have to satisfy the martingale condition as given in the FTAP:

$$\frac{C_0}{B_0} = \mathbb{E}^{\mathbb{Q}^B}\left[\frac{C_1}{B_1}\right] = q^B \cdot \frac{C_u}{B_u} + (1-q^B) \cdot \frac{C_d}{B_d}$$

It is also possible to use A as the numeraire:

$$\frac{C_0}{A_0} = \mathbb{E}^{\mathbb{Q}^A}\left[\frac{C_1}{A_1}\right] = q^A \cdot \frac{C_u}{A_u} + (1-q^A) \cdot \frac{C_d}{A_d}$$

In practice, one would use the values of A and B to compute the risk-neutral probabilities q^A or[6] q^B, and then use the appropriate equation from above to solve for the unknown price of the contingent claim, C_0.

(ii) Replication:
The replication approach involves constructing a portfolio (α, β) in asset A and B such that $C_1 = \alpha A_1 + \beta B_1$, i.e.

$$\begin{cases} C_d = \alpha A_d + \beta B_d \\ C_u = \alpha A_u + \beta B_u \end{cases}$$

Solving this system of two equations for (α, β) implies that this portfolio has the same payoff as C in all states of the world. Thus, to avoid arbitrage, it must be the case that $C_0 = \alpha A_0 + \beta B_0$. If C_0 differs from $\alpha A_0 + \beta B_0$ then we

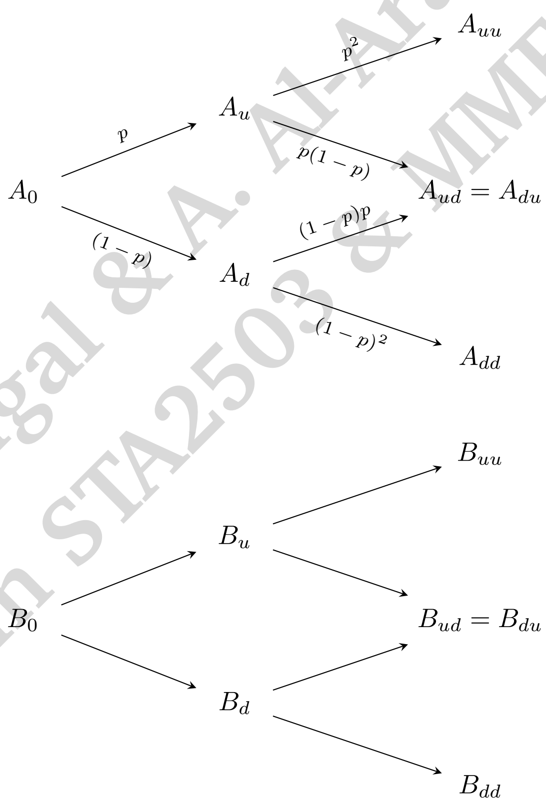
[5] A derivative is named so because it *derives* its value from the other assets in the economy

[6] Using either A or B as a numeraire must result in the same valuation if the economy is arbitrage-free. Also, choosing one of these as the numeraire would mean that the chosen asset would be in the denominator of the relative price process.

could construct an arbitrage portfolio consisting of the replicating portfolio and asset C itself (long the one with larger value and short the one with smaller value).

7.4 Multiple Assets, States and Periods

All the notions discussed in this note extend naturally to the case where we increase the number of assets and states. **When we are interested in valuation in a multi-period setting, we can use backward induction**. If the assets behave as follows:



we can focus on one branch at time t, find implied risk-neutral probabilities in that branch then use those probabilities to price the claim we are interested in.[7] For example, we can use A_u, A_{uu}, A_{ud} and B_u, B_{uu}, B_{ud} to compute risk neutral probabilities for the upper-right branch and use A_d, A_{dd}, A_{ud} and B_d, B_{dd}, B_{ud} to find the risk neutral probabilities for the lower-right branch.

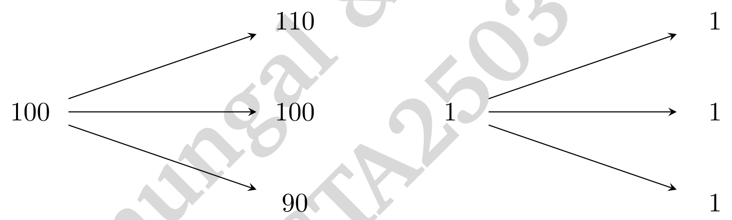
[7] This assumes, as usual, that we know the terminal payoff of the contingent claim in all states.

If we are interested in pricing a claim C and we know the 3 payoffs of a claim at time $t = 2$, we can use the risk-neutral probabilities from the previous step to back out the two prices of the claim at time $t = 1$, and then repeat the process once more to find the price of the claim at $t = 0$.

7.5 Exercises

EXERCISE 7.4. Consider the following economy:

Asset A **Asset B**



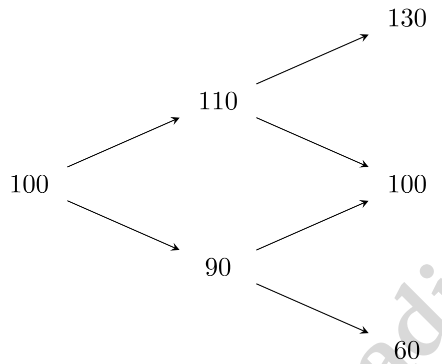
(i) Does this economy admit any arbitrage opportunities?

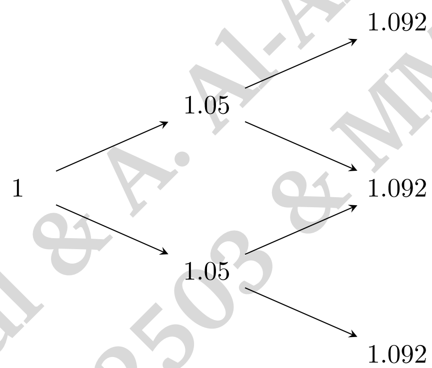
(ii) What is the value of a call option struck at 100 on the risky asset, C_0?

(iii) Show that if $C_0 = 6$ then there is an arbitrage.

EXERCISE 7.5. Suppose that two assets over two periods follow:

Asset A **Asset B**

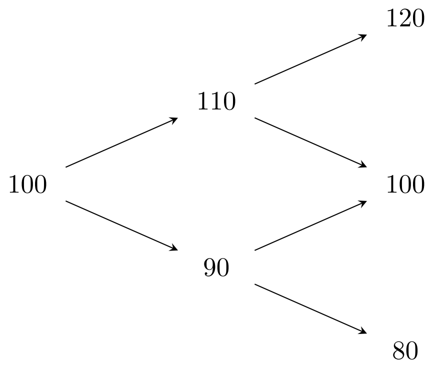


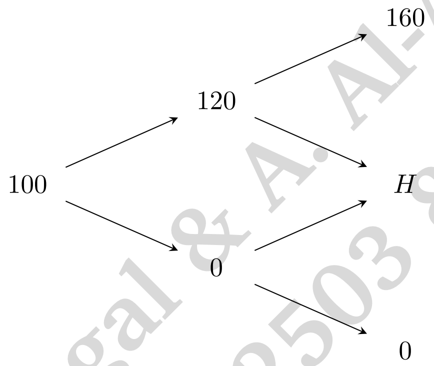


Value a put option on the risky asset struck at 110 with maturity $T = 2$ using risk-neutral probabilities.

Exercise 7.6. The following two assets are traded in a binomial economy.

Asset A **Asset B**





Asset A represents a typical stock dynamics, while asset B depicts a stock which may default then possibly recover.

(i) Determine the restrictions on H such that there is no arbitrage in this economy. Any comments?

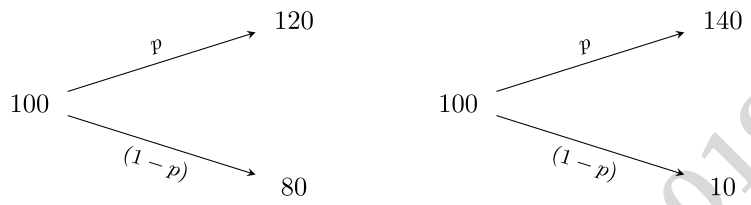
(ii) Construct an arbitrage strategy when $H = 50$.

(iii) Assuming that $H = 0$. Determine the branching probabilities induced by using asset A as a numeraire asset and determine the bounds for the price of a call option written on asset A maturing at $t = 2$ struck at (i) $K = 90$ and (ii) $K = 110$. Any comments?

(iv) What hedging strategy (by trading in asset A and B) would you use to hedge the two options in part (c)?

EXERCISE 7.7. Suppose two risky assets are traded and their price dynamics follow the following binomial trees:

Asset A **Asset B**



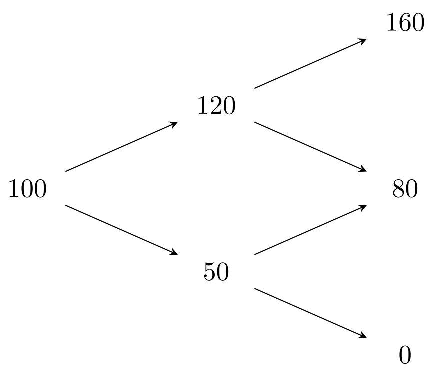
(i) Show that this is an arbitrage-free economy.

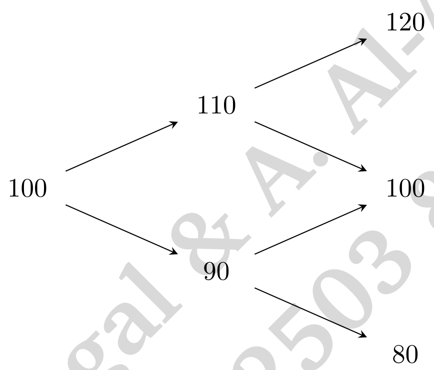
(ii) What is the arbitrage-free price of an option which allows the exchange of A for B (i.e. it allows you to give up one unit of asset A to receive one unit of asset B)?

(iii) What are the implied (i) risk-neutral branching probabilities, (ii) risk-free rate?

Exercise 7.8. The following two assets are being actively traded in a two-period binomial market economy. Asset A behaves like a stock which may default, while asset B behaves "normally".

Asset A **Asset B**





(i) Determine all relevant risk-neutral probabilities and short rates of interest.

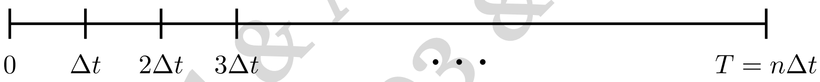
(ii) Using risk-neutral valuation, compute the price and replication strategy for a two-period American put option on asset A struck at 90.

8 Cox-Ross-Rubinstein Model

In this note, we discuss the Cox-Ross-Rubinstein (CRR) model, which is a multi-period binomial tree model that may be thought of as the discrete-time counterpart of the Black-Scholes model.

8.1 Model Setup

We consider a fixed time horizon, T, divided into n subperiods of length Δt each:



We assume that there are two assets in this economy:

(i) A **risk-free asset**, B_t, which is a money market account where the per-period interest rate is given by a constant $r > 0$. That is, the asset evolves according to:

$$\begin{aligned} B_{m\Delta t} &= B_{(m-1)\Delta t} \cdot (1 + r\Delta t) \approx B_{(m-1)\Delta t} \cdot e^{r\Delta t} \\ &= B_0 \cdot (1 + r\Delta t)^n \approx B_0 \cdot e^{r \cdot m\Delta t} \end{aligned}$$

For simplicity, we assume that $B_0 = 1$.

(ii) A **risky asset** which evolves according to the following model:

$$\begin{aligned} A_{m\Delta t} &= A_{(m-1)\Delta t} \cdot e^{\sigma\sqrt{\Delta t}\; x_m} \\ &= A_0 \cdot \exp\left[\sigma\sqrt{\Delta t}\; \sum_{i=1}^{m} x_i\right] \end{aligned}$$

where x_i are i.i.d. random variables with $\mathbb{P}(x_i = 1) = p$ and $\mathbb{P}(x_i = -1) = 1-p$ and $\sigma > 0$ is a known constant.

Thus, the assets evolve according to a binomial tree model with multiple periods, similar to the ones discussed previously. The trees produced by the CRR model begin as follows:

$$
\begin{tikzcd}
 & & A_0\ e^{2\sigma\sqrt{\Delta t}} \\
 & A_0\ e^{\sigma\sqrt{\Delta t}} \arrow[ur] \arrow[dr] & \\
A_0 \arrow[ur, "p"] \arrow[dr, "(1-p)"'] & & A_0 \\
 & A_0\ e^{-\sigma\sqrt{\Delta t}} \arrow[ur] \arrow[dr] & \\
 & & A_0\ e^{-2\sigma\sqrt{\Delta t}}
\end{tikzcd}
$$

$$
\begin{tikzcd}
 & & e^{r\cdot 2\Delta t} \\
 & e^{r\Delta t} \arrow[ur] \arrow[dr] & \\
B_0 = 1 \arrow[ur] \arrow[dr] & & e^{r\cdot 2\Delta t} \\
 & e^{r\Delta t} \arrow[ur] \arrow[dr] & \\
 & & e^{r\cdot 2\Delta t}
\end{tikzcd}
$$

Thus, at any given step m the prices of the two assets evolve as follows:

$$
\begin{tikzcd}
 & A_{m-1} \cdot e^{\sigma\sqrt{\Delta t}} \\
A_{m-1} \arrow[ur, "p"] \arrow[dr, "(1-p)"'] & \\
 & A_{m-1} \cdot e^{-\sigma\sqrt{\Delta t}}
\end{tikzcd}
$$

$$
\begin{tikzcd}
 & B_{m-1} \cdot e^{r\Delta t} \\
B_{m-1} \arrow[ur, "p"] \arrow[dr, "(1-p)"'] & \\
 & B_{m-1} \cdot e^{r\Delta t}
\end{tikzcd}
$$

We will further assume that the **physical probability of an up-move**, p, is given by:

$$p \;=\; \frac{1}{2}\left(1+\frac{\mu-\frac{1}{2}\sigma^2}{\sigma}\sqrt{\Delta t}\right)$$

This is a convenient reparametrization that will make later steps more manageable.

8.2 Risk neutral probabilities and no-arbitrage conditions

First, we can **derive the risk-neutral probabilities in this model**. Using the money market as the numeraire asset produces the following relative price process:

$$\begin{array}{ccl} & \nearrow & \dfrac{A_{m-1}}{B_{m-1}}\dfrac{e^{\sigma\sqrt{\Delta t}}}{e^{r\Delta t}} = \dfrac{A_{m-1}}{B_{m-1}}\, e^{\sigma\sqrt{\Delta t}-r\Delta t} \\ \dfrac{A_{m-1}}{B_{m-1}} & & \\ & \searrow & \dfrac{A_{m-1}}{B_{m-1}}\dfrac{e^{-\sigma\sqrt{\Delta t}}}{e^{r\Delta t}} = \dfrac{A_{m-1}}{B_{m-1}}\, e^{-\sigma\sqrt{\Delta t}-r\Delta t} \end{array}$$

Recall that the risk-neutral probability of an up-move (when using B as the numeraire) in a binomial tree is given by:

$$q \;=\; \frac{A_0/B_0 - A_d/B_d}{A_u/B_u - A_d/B_d}$$

<u>Note:</u> here we are dropping the B superscript since B is the money market account, and it is common to refer to the EMM associated with this asset as *the* risk-neutral measure, often denoted $\mathbb{Q}$.

Applying this formula to the CRR model gives the following risk-neutral probability:

THEOREM 8.1 (CRR risk-neutral measure). *In the CRR model the risk-neutral measure, $\mathbb{Q}$, is characterized by the up-move probability:*

$$q \;=\; \frac{e^{r\Delta t}-e^{-\sigma\sqrt{\Delta t}}}{e^{\sigma\sqrt{\Delta t}}-e^{-\sigma\sqrt{\Delta t}}}$$

Next, we are interested in the no-arbitrage condition for the CRR model. Given the above expression for q, and since we have assumed[1] that

[1] This is a natural assumption since interest rates are typically (though not necessarily) positive, and changing the sign of σ only causes the tree to flip vertically - moreover, σ is associated with the volatility of the asset, which is a positive quantity.

$\sigma, r > 0$, we see that $q \in (0,1)$ - and hence the economy is arbitrage-free - if and only if $r\Delta t < \sigma\sqrt{\Delta t}$

THEOREM 8.2 (No-arbitrage condition in the CRR model). *The CRR model is arbitrage-free if and only if*

$$r \; < \; \frac{\sigma}{\sqrt{\Delta t}}$$

8.3 Continuous-Time Limit of CRR Model

8.3.1 Limiting Distribution Under $\mathbb{P}$

We are now interested in the limiting distribution of asset prices under $\mathbb{P}$ in the limit as the number of subperiods approaches infinity. Intuitively, we are trying to deduce the continuous-time behavior of the asset, as the up and down moves occur more frequently and become smaller in magnitude. Note that as $n \to \infty$, $\Delta t \to 0$ and vice versa. In other words, as we increase the number of subperiods, the length each subperiod tends to 0 since the horizon, T, is fixed. Recall that the asset price at time T is given by:

$$A_T = A_{n\Delta t} = A_0 \cdot \exp\left[\sigma\sqrt{\Delta t} \cdot \sum_{i=1}^{n} x_i\right]$$

and denote the term in the exponent by $X_n = \sigma\sqrt{\Delta t} \cdot \sum_{i=1}^{n} x_i$. Let us **derive the characteristic function of X_n**:

$$\begin{aligned}
\psi_n(u) &= \mathbb{E}^{\mathbb{P}}\left[e^{iuX_n}\right] \\
&= \left(\mathbb{E}^{\mathbb{P}}\left[e^{iu\sigma\sqrt{\Delta t}\cdot x_1}\right]\right)^n && \text{since } x_i \text{ are i.i.d.} \\
&= \left(\mathbb{E}^{\mathbb{P}}\left[1 + iu\sigma\sqrt{\Delta t}\cdot x_1 + \frac{1}{2}(iu\sigma\sqrt{\Delta t}\cdot x_1)^2 + o(\Delta t)\right]\right)^n && \text{Taylor expansion of } \exp(x) \\
&= \left(1 + iu\sigma\sqrt{\Delta t}\cdot(2p-1) - \frac{u^2}{2}\sigma^2\Delta t + o(\Delta t)\right)^n && \mathbb{E}(x_1) = 2p-1,\ \mathbb{E}(x_1^2) = 1 \\
&= \left(1 + iu\left(\mu - \frac{1}{2}\sigma^2\right)\Delta t - \frac{u^2}{2}\sigma^2\Delta t + o(\Delta t)\right)^n \\
&= \left(1 + \frac{\left(iu\left(\mu - \frac{1}{2}\sigma^2\right) - \frac{u^2}{2}\sigma^2\right)T}{n} + o\left(\frac{1}{n}\right)\right)^n && \text{since } \Delta t = T/n
\end{aligned}$$

Taking the limit as n tends to infinity we find that:

$$\psi_n(u) \underset{n\to\infty}{\longrightarrow} \exp\left[iu\left(\mu - \frac{1}{2}\sigma^2\right)T - \frac{u^2}{2}\sigma^2 T\right]$$

Notice now that the expression on the RHS is the characteristic function of a normal random variable with mean $\left(\mu - \frac{1}{2}\sigma^2\right)T$ and variance $\sigma^2 T$. Since this function is continuous, we can invoke Levy's continuity theorem to conclude that X_n weakly converges to this distribution as the number of subperiods tends to infinity. From this we can conclude that the limiting distirbution of the asset value is lognormal:

THEOREM 8.3 (Limiting distribution of A_T under $\mathbb{P}$). *In the CRR model, as $n \to \infty$, the asset value A_T converges in distribution:*

$$A_T \xrightarrow{d} A_0 \cdot \exp\left[\left(\mu - \frac{1}{2}\sigma^2\right) T + \sigma\sqrt{T} Z^{\mathbb{P}}\right]$$

where $Z^{\mathbb{P}}$ is a standard normal random variable under $\mathbb{P}$.

8.3.2 Limiting Distribution Under $\mathbb{Q}$

Next, we turn to the limiting distribution under $\mathbb{Q}$. Returning to the risk-neutral probability, q, and using the fact that $e^x = \sum_{k=0}^{n} \frac{x^k}{k!} + o(x^n)$, we can write:

$$\begin{aligned}
q &= \frac{e^{r\Delta t} - e^{-\sigma\sqrt{\Delta t}}}{e^{\sigma\sqrt{\Delta t}} - e^{-\sigma\sqrt{\Delta t}}} \\
&= \frac{\left(1 + r\Delta t + o(\Delta t^2)\right) - \left(1 - \sigma\sqrt{\Delta t} + \frac{1}{2}\sigma^2\Delta t + o(\Delta t^{3/2})\right)}{\left(1 + \sigma\sqrt{\Delta t} + \frac{1}{2}\sigma^2\Delta t + o(\Delta t^{3/2})\right) - \left(1 - \sigma\sqrt{\Delta t} + \frac{1}{2}\sigma^2\Delta t + o(\Delta t^{3/2})\right)} \\
&= \frac{\sigma\sqrt{\Delta t} + \left(r - \frac{1}{2}\sigma^2\right)\Delta t + o(\Delta t^{3/2})}{2\sigma\sqrt{\Delta t} + o(\Delta t^{3/2})} \\
&= \frac{1}{2}\left(1 + \frac{r - \frac{1}{2}\sigma^2}{\sigma}\sqrt{\Delta t}\right) + o(\Delta t)
\end{aligned}$$

Notice that this expression is identical (up to terms of order Δt) to the physical probability p, but with μ being replaced by the risk-free rate r. Therefore, if we repeat the same steps as in the previous section using (the approximation of) q instead of p we can derive that the limiting distribution of A_T under $\mathbb{Q}$ is:

THEOREM 8.4 (Limiting distribution of A_T under $\mathbb{Q}$). *In the CRR model, as $n \to \infty$, the asset value A_T converges in distribution:*

$$A_T \xrightarrow{d} A_0 \cdot \exp\left[\left(r - \frac{1}{2}\sigma^2\right) T + \sigma\sqrt{T} Z^{\mathbb{Q}}\right]$$

where $Z^{\mathbb{Q}}$ is a standard normal random variable under $\mathbb{Q}$.

8.3.3 Limiting Distribution Under $\mathbb{Q}^A$

The analysis in the previous section can be repeated with the risky asset as the numeraire to derive an analogous result involving the limiting distribution of A_T under $\mathbb{Q}^A$.

THEOREM 8.5 (Limiting distribution of A_T under $\mathbb{Q}^A$). *In the CRR model, as $n \to \infty$, the asset value A_T converges in distribution:*

$$A_T \xrightarrow{d} A_0 \cdot \exp\left[\left(r + \frac{1}{2}\sigma^2\right)T + \sigma\sqrt{T}Z^{\mathbb{Q}^A}\right]$$

where $Z^{\mathbb{Q}^A}$ is a standard normal random variable under $\mathbb{Q}^A$.

Proof When using A as the numeraire the risk-neutral probability is equal to:

$$\begin{aligned}
q^A &= \frac{B_0/A_0 - B_d/A_d}{B_u/A_u - B_d/A_d} \\
&= \frac{e^{-r\Delta t} - e^{\sigma\sqrt{\Delta t}}}{e^{-\sigma\sqrt{\Delta t}} - e^{\sigma\sqrt{\Delta t}}} \\
&= \frac{\left(1 - r\Delta t + o(\Delta t^2)\right) - \left(1 + \sigma\sqrt{\Delta t} + \frac{1}{2}\sigma^2\Delta t + o(\Delta t^{3/2})\right)}{\left(1 - \sigma\sqrt{\Delta t} + \frac{1}{2}\sigma^2\Delta t + o(\Delta t^{3/2})\right) - \left(1 + \sigma\sqrt{\Delta t} + \frac{1}{2}\sigma^2\Delta t + o(\Delta t^{3/2})\right)} \\
&= \frac{-\sigma\sqrt{\Delta t} - \left(r + \frac{1}{2}\sigma^2\right)\Delta t + o(\Delta t^{3/2})}{-2\sigma\sqrt{\Delta t} + o(\Delta t^{3/2})} \\
&= \frac{1}{2}\left(1 + \frac{r + \frac{1}{2}\sigma^2}{\sigma}\sqrt{\Delta t}\right) + o(\Delta t)
\end{aligned}$$

Comparing q^A to the form of p or q we arrive at the desired result. □

8.4 Black-Scholes Option Pricing Formula

In this section we derive the Black-Scholes option pricing formula for European call and put options. The Black-Scholes model can be viewed as the continuous-time counterpart of the CRR model. The key to pricing derivatives under the assumptions of the Black-Scholes model (or any modeling assumptions) is to invoke the FTAP.

A **European call option** with strike price K gives the option holder the right (but not the obligation) to *buy* the underlying asset at the strike price K at some future maturity time T. That is, the payoff function at time T is given by $(A_T - K)_+ = \max(A_T - K,\ 0)$. Similarly, a **European put option** gives the option holder the right (but not the obligation) to *sell* the underlying asset at the strike price K at some future maturity time T. That is, the payoff function at time T is given by $(K - A_T)_+ = \max(K - A_T,\ 0)$. We will use two different approaches to compute the price of the option at time 0, V_0^{call} (the price of the put can be obtained in an analogous manner):

(i) Direct approach using $\mathbb{Q}$:
We can use the money market asset as the numeraire and the fact that relative

prices are martingales under the risk-neutral measure:

$$\frac{V_0^{call}}{B_0} \;=\; \mathbb{E}^{\mathbb{Q}}\left[\frac{V_T^{call}}{B_T}\middle|\mathcal{F}_0\right] \;=\; \mathbb{E}_0^{\mathbb{Q}}\left[\frac{V_T^{call}}{B_T}\right]$$

Now, we assumed that $B_0 = 1$ and we have that $B_T = e^{rT}$ from previous sections. Also, the price of the option at the exercise date, T, is simply the payoff function. Therefore, we have:

$$V_0^{call} \;=\; e^{-rT}\;\mathbb{E}_0^{\mathbb{Q}}\left[(A_T - K)_+\right]$$

We also know the distribution of A_T under $\mathbb{Q}$, so we can use that information to compute the expectation:

$$\begin{aligned} V_0^{call} &= e^{-rT}\;\mathbb{E}_0^{\mathbb{Q}}\left[\left(A_0\; e^{(r-\frac{1}{2}\sigma^2)T+\sigma\sqrt{T}Z} - K\right)_+\right] \qquad \text{where } Z \sim N(0,1) \\ &= e^{-rT}\int_{-\infty}^{\infty}\left(A_0\; e^{(r-\frac{1}{2}\sigma^2)T+\sigma\sqrt{T}z} - K\right)_+ \frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}}\,dz \end{aligned}$$

Note that the term under the integral is 0 when $A_0\; e^{(r-\frac{1}{2}\sigma^2)T+\sigma\sqrt{T}z} < K$, i.e. when $z < z^* = -\frac{\ln(A_0/K)+(r-\frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}$. Using this, we can integrate from z^* and remove the max function:

$$\begin{aligned} V_0^{call} &= e^{-rT}\int_{z^*}^{\infty}\left(A_0\; e^{(r-\frac{1}{2}\sigma^2)T+\sigma\sqrt{T}z} - K\right)\cdot\frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}}\,dz \\ &= e^{-rT}\int_{z^*}^{\infty} A_0\; e^{(r-\frac{1}{2}\sigma^2)T+\sigma\sqrt{T}z}\cdot\frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}}\,dz \;-\; e^{-rT}\int_{z^*}^{\infty} K\cdot\frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}}\,dz \\ &= A_0\; e^{-\frac{1}{2}\sigma^2 T}\int_{z^*}^{\infty} e^{\sigma\sqrt{T}z}\cdot\frac{e^{-\frac{z^2}{2}}}{\sqrt{2\pi}}\,dz \;-\; K\; e^{-rT}\Phi(-z^*) \end{aligned}$$

Use the substitution $z' = z - \sigma\sqrt{T}$ and complete the square in the exponent to find:

$$\begin{aligned} V_0^{call} &= A_0\int_{z^*-\sigma\sqrt{T}}^{\infty}\frac{e^{-\frac{(z')^2}{2}}}{\sqrt{2\pi}}\,dz' - K\; e^{-rT}\Phi(-z^*) \\ &= A_0\;\Phi(-z^* + \sigma\sqrt{T}) - K\; e^{-rT}\Phi(-z^*) \end{aligned}$$

Now denote:

$$\begin{aligned} d_- &= \frac{\ln(A_0/K) + (r - \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}} \;=\; -\,z^* \\ d_+ &= \frac{\ln(A_0/K) + (r + \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}} \;=\; -\,z^* + \sigma\sqrt{T} \end{aligned}$$

and we can write the **Black-Scholes option pricing formula** for European call options:

$$V_0^{call} \;=\; A_0\;\Phi(d_+) - Ke^{-rT}\Phi(d_-)$$

where Φ is the cdf of a standard normal random variable.

(ii) Using both $\mathbb{Q}$ and $\mathbb{Q}^A$:
A simpler approach involves breaking the option payoff into two parts:

$$V_T^{call} = (A_T - K)_+ = \underbrace{A_T \, \mathbb{1}_{\{A_T > K\}}}_{\text{asset-or-nothing option}} - \underbrace{K \, \mathbb{1}_{\{A_T > K\}}}_{\text{digital option}}$$

Now, each part can be priced using different numeraires/measures that make the computation easier.[2] For the asset-or-nothing option we can use A as the numeraire:

$$\begin{aligned} \frac{V_0^1}{A_0} &= \mathbb{E}_0^{\mathbb{Q}^A}\left[\frac{V_T^1}{A_T}\right] && \text{by FTAP} \\ &= \mathbb{E}_0^{\mathbb{Q}^A}\left[\frac{A_T \, \mathbb{1}_{\{A_T>K\}}}{A_T}\right] \\ &= \mathbb{Q}_0^A\left(A_T > K\right) \\ &= \mathbb{Q}_0^A\left(A_0 \, e^{(r+\frac{1}{2}\sigma^2)T+\sigma\sqrt{T}Z^A} > K\right) && \text{where } Z^A \sim N(0,1) \text{ under } \mathbb{Q}^A \\ &= \mathbb{Q}_0^A\left(Z^A < d_+\right) \end{aligned}$$

Thus, the value of the asset-or-nothing option is: $V_0^1 = A_0\Phi(d_+)$.

Next, we value the digital option using the money market as the numeraire:

$$\begin{aligned} \frac{V_0^2}{B_0} &= \mathbb{E}_0^{\mathbb{Q}}\left[\frac{V_T^2}{B_T}\right] \\ &= \mathbb{E}_0^{\mathbb{Q}}\left[\frac{K \, \mathbb{1}_{\{A_T>K\}}}{B_T}\right] \\ &= Ke^{-rT}\mathbb{Q}_0\left(A_T > K\right) \\ &= \mathbb{Q}_0\left(A_0 \, e^{(r-\frac{1}{2}\sigma^2)T+\sigma\sqrt{T}Z} > K\right) && \text{where } Z \sim N(0,1) \text{ under } \mathbb{Q} \\ &= \mathbb{Q}_0\left(Z < d_-\right) \end{aligned}$$

and so the value of the digital option is $V_0^2 = Ke^{-rT}\Phi(d_-)$, since $B_0 = 1$. Combining the two values yields the same formula as before.

We summarize the main result and extend to the value of the options at any given time $t < T$:

Theorem 8.6 (Black-Scholes Option Pricing Formula). *In the Black-Scholes model, the price of call and put options struck on an asset with value S_t at time*

[2] The use of carefully chosen numeraires and measure changes is a very useful technique in derivative pricing problems.

t with strike price K is given by:

$$V_t^{call} = S_t\ \Phi(d_+) - Ke^{-r(T-t)}\Phi(d_-)$$
$$V_t^{put} = Ke^{-r(T-t)}\Phi(-d_-) - S_t\ \Phi(-d_+)$$

where

$$d_\pm = \frac{\ln(S_t/K) + (r \pm \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}$$

8.5 Exercieses

EXERCISE 8.7. Consider a call option written on a put option. The call matures at $T = 1$ and has strike 10, while the put matures at $T = 2$ and has strike 100. The underlying for the put has $S = 100$, $\sigma = 20\%$ and $r = 5\%$. Price the call option using $\Delta T = 1$.

EXERCISE 8.8. Suppose that the CRR tree is modified so that

$$S_n = S_{n-1} \cdot e^{\left(r - \frac{1}{2}\sigma^2\right)\Delta t + \sigma\sqrt{\Delta t}x_n}$$

where $x_1, x_2, ...$ are i.i.d. Bernoulli random variables. Compute the risk-neutral branching probabilities to order Δt (not $\sqrt{\Delta t}$).

EXERCISE 8.9. This question concerns the CRR binomial tree model for stock price dynamics. Assume $S_0 = 100$, $\sigma = 60\%$, $\mu = 15\%$, $r = 10\%$ and $\Delta t = \frac{1}{12}$.

(i) Determine the price and the hedging strategy for each of the following options maturing in 3-months:
  (a) a call option struck at \$100.
  (b) a digital call option struck at \$100.
  (c) a strangle option with $K_1 = \$95$ and $K_2 = 105$.

(ii) Suppose that the market price for the strangle option in a.iii is 10% higher than the no arbitrage price. Construct an arbitrage strategy using only the money market and the underlying stock.

EXERCISE 8.10. Using a CRR tree, with $S = 100$, $\sigma = 50\%$, $r = 5\%$ and $\Delta t = \frac{1}{12}$, determine the value and hedging strategy for each of the following 3-month European options:

(i) digital call struck at 100
(ii) digital put struck at 100
(iii) put struck at 100

(iv) call struck at 100
(v) straddle struck at 100
(vi) strangle with $K_1 = 95$, $K_2 = 115$
(vii) bull spread with $K_1 = 95$, $K_2 = 115$

Exercise 8.11. **Two-Dimensional CRR Model.** You are modeling the evolution of two correlated stocks as follows:

$$S_1(i\,\Delta t) = S_1((i-1)\,\Delta t)\,e^{\alpha_1\,x_i} \qquad S_2(i\,\Delta t) = S_2((i-1)\,\Delta t)\,e^{\alpha_2\,y_i}$$

where $(x_1, y_1), \dots, (x_n, y_n)$ are pairs of i.i.d random variables with the following real-world probabilities:

$$\mathbb{P}(x_i = +1,\; y_i = +1) = p_1 \qquad \mathbb{P}(x_i = +1,\; y_i = -1) = p_2$$
$$\mathbb{P}(x_i = -1,\; y_i = +1) = p_3 \qquad \mathbb{P}(x_i = -1,\; y_i = -1) = p_4$$

This corresponds to a tree with four nodes emanating from every point and does in fact recombine in two-dimensions.

You wish to ensure that in the limit $n \to \infty$ with $\Delta t = T/n$ (T fixed and finite), the following means, variances and correlation match the market values:

$$\mathbb{E}^{\mathbb{P}}\left[\ln(S_1(T)/S_1(0))\right] = (\mu_1^* - \tfrac{1}{2}(\sigma_1^*)^2)T \qquad \mathbb{V}^{\mathbb{P}}\left[\ln(S_1(T)/S_1(0))\right] = (\sigma_1^*)^2\,T$$
$$\mathbb{E}^{\mathbb{P}}\left[\ln(S_2(T)/S_2(0))\right] = (\mu_2^* - \tfrac{1}{2}(\sigma_2^*)^2)T \qquad \mathbb{V}^{\mathbb{P}}\left[\ln(S_2(T)/S_2(0))\right] = (\sigma_2^*)^2\,T$$
$$\mathbb{C}^{\mathbb{P}}(\ln(S_1(T)/S_1(0)), \ln(S_2(T)/S_2(0))) = \rho^*\sigma_1^*\sigma_2^*$$

(i) Determine the probabilities $\{p_j : j = 1, 2, 3, 4\}$ and $\alpha_{1,2}$ in terms of the model parameters and Δt to lowest order in Δt.
(ii) To determine the risk-neutral probabilities you must make both relative prices of S_1 and S_2 (w.r.t. the money-market account) martingales. Since you have only three assets, but four outcomes, the risk-neutral branching probabilities are not unique. Nonetheless, show that under the risk-neutral measure we have the following:

$$\mathbb{E}^{\mathbb{Q}}\left[\ln(S_1(T)/S_1(0))\right] = (r - \tfrac{1}{2}(\sigma_1^*)^2)T \qquad \mathbb{V}^{\mathbb{Q}}\left[\ln(S_1(T)/S_1(0))\right] = (\sigma_1^*)^2\,T$$
$$\mathbb{E}^{\mathbb{Q}}\left[\ln(S_2(T)/S_2(0))\right] = (r - \tfrac{1}{2}(\sigma_2^*)^2)T \qquad \mathbb{V}^{\mathbb{Q}}\left[\ln(S_2(T)/S_2(0))\right] = (\sigma_2^*)^2\,T$$

Furthermore, determine the no arbitrage bounds (to lowest order in Δt) for the correlation under $\mathbb{Q}$.
(iii) Implement the two-asset tree in MATLAB to determine the no-arbitrage bounds on a 1-year European exchange option. This option lets you exchange asset 1 for asset 2 – if it is favorable to do so. Assume that $\sigma_1^* = 20\%$, $\sigma_2^* = 30\%$, $r = 5\%$, and $\rho^* = 0.5$. How does correlation affect the bounds/prices? Explain.

Exercise 8.12. Too similar to Q10? You are modeling the evolution of two correlated stocks as follows:

$$S_1(i\,\Delta t) = S_1((i-1)\,\Delta t)\, e^{\alpha_1\, x_i} \qquad S_2(i\,\Delta t) = S_2((i-1)\,\Delta t)\, e^{\alpha_2\, y_i}$$

where $(x_1, y_1), \dots, (x_n, y_n)$ are pairwise i.i.d Bernoulli random variables with the following real-world probabilities:

$$\begin{aligned} \mathbb{P}(x_i = +1,\ y_i = +1) &= p_1 \qquad & \mathbb{P}(x_i = +1,\ y_i = -1) &= p_2 \\ \mathbb{P}(x_i = -1,\ y_i = +1) &= p_3 \qquad & \mathbb{P}(x_i = -1,\ y_i = -1) &= p_4 \end{aligned}$$

You wish to ensure that in the limit in which $n \to \infty$ with $\Delta t = T/n$ (T fixed and finite), the joint distribution of $S_1(T)$ and $S_2(T)$ is a joint log-normal with the following properties

$$\begin{aligned} \mathbb{E}^{\mathbb{P}}\left[S_1(T)\right] &= e^{\mu_1 T} S_1(0) \\ \mathbb{V}^{\mathbb{P}}\left[\ln(S_1(T)/S_1(0))\right] &= \sigma_1^2\, T \\ \mathbb{E}^{\mathbb{P}}\left[S_2(T)\right] &= e^{\mu_2 T} S_2(0) \\ \mathbb{V}^{\mathbb{P}}\left[\ln(S_2(T)/S_2(0))\right] &= \sigma_2^2\, T \\ \mathbb{C}^{\mathbb{P}}(\ln(S_1(T)/S_1(0)), \ln(S_2(T)/S_2(0))) &= \rho\sigma_1\sigma_2 \end{aligned}$$

Determine the probabilities $\{p_j : j = 1, 2, 3, 4\}$ and $\alpha_{1,2}$ in terms of the model parameters and Δt to lowest order in Δt.

9 Dynamic Hedging and the Generalized Black-Scholes Equation

In the simple 2-period binomial tree setting, one approach for valuing derivatives was via replication. That is, when we wanted to price a claim we would find a portfolio of other assets whose payoffs matched that of the claim in every state of the world. If there is no arbitrage in the economy, then the value of the replicating portfolio must be equal to the value of the derivative. Here, we **extend this general notion to the continuous-time setting**.

9.1 Model Setup

We begin with a **model of the economy**. We assume that we have:

(i) **A source of uncertainty** given by the process $X = (X_t)_{t\geq 0}$. We will assume that X is an Itô process that satisfies the SDE:

$$\frac{dX_t}{X_t} = \mu^X(t, X_t)\, dt + \sigma^X(t, X_t)\, dW_t$$

where $\mu^X : \mathbb{R}^+ \times \mathbb{R} \to \mathbb{R}$ and $\sigma^X : \mathbb{R}^+ \times \mathbb{R} \to \mathbb{R}^+ \backslash \{0\}$ are functions and $W = (W_t)_{t\geq 0}$ is a $\mathbb{P}$-Brownian motion. For brevity, we will use the following notation for the drift and volatility processes[1]: $\mu_t^X = \mu^X(t, X_t)$ and $\sigma_t^X = \sigma^X(t, X_t)$.

(ii) A **bank account** (process), $B = (B_t)_{t\geq 0}$, satisfying:

$$\frac{dB_t}{B_t} = r(t, X_t)\, dt$$

where $r : \mathbb{R}^+ \times \mathbb{R} \to \mathbb{R}$ is a function and $r = (r_t)_{t\geq 0} = (r(t, X_t))_{t\geq 0}$ is the **short rate process**. Notice that this is an instantaneously risk-free asset, as it is not driven by a Brownian motion.

[1] Note the abuse of notation and the difference between processes and functions: μ^X is a function that maps $(t, x) \mapsto \mu^X(t, x)$, but also $\mu^X = (\mu_t^X)_{t\geq 0} = (\mu^X(t, X_t))_{t\geq 0}$ is a process.

(iii) A **traded claim written on the source of uncertainty** whose price process, $f = (f_t)_{t \geq 0}$, satisfies:

$$\frac{df_t}{f_t} = \mu^f(t, X_t)\, dt + \sigma^f(t, X_t)\, dW_t$$

Again, we will use the notation $\mu_t^f = \mu^f(t, X_t)$ and $\sigma_t^f = \sigma^f(t, X_t)$.

Now, **we are interested in the price of a second contingent claim** written on X, whose price process we will denote by $g = (g_t)_{0 \leq t \leq T}$ and that pays $G(X_T)$ at time T. Furthermore, we will assume that g is Markovian in X, i.e. that there exists a function $g : \mathbb{R}^+ \times \mathbb{R} \to \mathbb{R}$ such that $g_t = g(t, X_t)$. Intuitively, this means that the price of the claim depends only on the prevailing level of the underlying process it is written on. If we further assume that $g \in C^{1,2}$, then **we can apply Itô's lemma to find the dynamics of** g. In general form the SDE that g_t satisfies is:

$$\frac{dg_t}{g_t} = \mu^g(t, X_t)\, dt + \sigma^g(t, X_t)\, dW_t$$

or applying Itô's lemma we can also write:

$$dg_t = \left(\partial_t g + \mu_t^x\, X_t \cdot \partial_x g + \frac{1}{2}(\sigma_t^x)^2\, X_t^2 \cdot \partial_{xx} g \right) dt + (\sigma_t^x\, X_t \cdot \partial_x g)\, dW_t$$

Note: X may or may not be a traded asset. For example, if we are valuing a call option written on a stock, then X is the price process of a traded asset. But we can also write derivatives on non-traded processes, e.g. consider a derivative that pays \$1 if the temperature on a certain date is below some level; here X_t would be the temperature at time t which is clearly non-traded.

9.2 Dynamic Hedging Argument

The goal of the dynamic hedging argument is to construct a replicating portfolio for the claim g and invoke a no-arbitrage argument to find the PDE that the function g satisfies. Specifically, we will **build an instantaneously risk-free portfolio trading in f and B and short the claim g**:

- **Set up a zero cost, self-financing portfolio** with $(\alpha, \beta) = (\alpha_t, \beta_t)_{t \geq 0}$ in f and B and short the claim we wish to price, g. Thus, the value of the portfolio satisfies:

$$\begin{aligned} V_0 &= 0 \\ V_t &= \alpha_t f_t + \beta_t B_t - g_t \\ dV_t &= \alpha_t \cdot df_t + \beta_t \cdot dB_t - dg_t \qquad \text{(self-financing assumption)} \end{aligned}$$

- **Substitute the expressions for df_t, dB_t and dg_t** into the portfolio value SDE:

$$\begin{aligned} dV_t &= \alpha_t \underbrace{\left(f_t \cdot \mu_t^f \; dt + f_t \cdot \sigma_t^f \; dW_t\right)}_{df_t} + \beta_t \underbrace{\left(B_t \cdot r_t \; dt\right)}_{dB_t} - \underbrace{\left(g_t \cdot \mu_t^g \; dt + g_t \cdot \sigma_t^g \; dW_t\right)}_{dg_t} \\ &= \left(\alpha_t f_t \mu_t^f + \beta_t B_t r_t - g_t \mu_t^g\right) dt + \left(\alpha_t f_t \sigma_t^f - g_t \sigma_t^g\right) dW_t \end{aligned}$$

- **To achieve an instantaneously risk-free portfolio, we must set the coefficient of dW_t to be zero**, since this is the source of all the noise in the portfolio process. This can be achieved with a particular choice of α_t:

$$\begin{aligned} \alpha_t^* f_t \sigma_t^f - g_t \sigma_t^g &= 0 \\ \implies \quad \alpha_t^* &= \frac{g_t \; \sigma_t^g}{f_t \; \sigma_t^f} \end{aligned}$$

  This choice of α_t eliminates local stochasticity from the evolution of the portfolio value. The resulting portfolio will have value process with previsible drift:

$$\begin{aligned} dV_t &= \left(\alpha_t^* f_t \mu_t^f + \beta_t B_t r_t - g_t \mu_t^g\right) dt + \left(\alpha_t^* f_t \sigma_t^f - g_t \sigma_t^g\right) dW_t \\ &= \left(\frac{g_t \; \sigma_t^g}{f_t \; \sigma_t^f} \cdot f_t \mu_t^f + \beta_t B_t r_t - g_t \mu_t^g\right) dt + \left(\frac{g_t \; \sigma_t^g}{f_t \; \sigma_t^f} \cdot f_t \sigma_t^f - g_t \sigma_t^g\right) dW_t \\ &= \left(\frac{\sigma_t^g}{\sigma_t^f} \cdot g_t \mu_t^f + \beta_t B_t r_t - g_t \mu_t^g\right) dt \end{aligned}$$

- **The drift term for the value process must be identically zero to avoid arbitrage**, since the portfolio value at the outset is zero, $V_0 = 0$. That is, the no-arbitrage assumption forces the drift of the process above to be equal to zero:

$$\frac{\sigma_t^g}{\sigma_t^f} \cdot g_t \mu_t^f + \beta_t B_t r_t - g_t \cdot \mu_t^g = 0$$

- **The no-arbitrage constraint implies two important facts about the economy:**

  (i) All traded assets have equal market price of risk:

  Since $V_0 = 0$, the choice of α along with the no-arbitrage constraint above force $V_t = 0$ for all t (since the process has no randomness and zero drift).

Recalling that we can write the portfolio value in terms of the holdings and their values, we have:

$$
\begin{aligned}
V_t &= \alpha_t f_t + \beta_t B_t - g_t = 0 \\
\implies \beta_t B_t &= g_t - \alpha_t f_t \\
\implies \beta_t B_t &= g_t \left(1 - \frac{\sigma_t^g}{\sigma_t^f}\right) \qquad \text{(substituting in } \alpha_t^* \text{ and collecting terms)}
\end{aligned}
$$

Substituting this expression for $\beta_t B_t$ into the drift constraint above yields:

$$
\begin{aligned}
\frac{\sigma_t^g}{\sigma_t^f} \cdot g_t \mu_t^f + \underbrace{g_t \left(1 - \frac{\sigma_t^g}{\sigma_t^f}\right)}_{=\beta_t B_t} r_t &= g_t \cdot \mu_t^g \\
\implies \frac{\mu_t^f - r_t}{\sigma_t^f} &= \frac{\mu_t^g - r_t}{\sigma_t^g}
\end{aligned}
$$

The quantity on either side of the equation is referred to as the **market price of risk (process)** of f and g. Since, there were no special assumptions made on the claims f and g that gave rise to this relation, it follows that **any two contingent claims written on X in this economy must have the same market price of risk**. So, we can denote the process by $\lambda_t = \lambda(t, X_t)$.

(ii) All claims written on X satisfy the generalized Black-Scholes PDE:
Notice from the definition of the market price of risk for g that:

$$
\begin{aligned}
\lambda_t &= \frac{\mu_t^g - r_t}{\sigma_t^g} = \frac{g_t \mu_t^g - g_t r_t}{g_t \sigma_t^g} \\
\implies g_t \mu_t^g - \lambda_t \cdot g_t \sigma_t^g &= r_t g_t
\end{aligned}
$$

Also, recall that from the dynamics of the claim we are interested in pricing, g, we find that its drift and volatility are given by:

$$
\begin{aligned}
g_t \mu_t^g &= \partial_t g + \mu_t^x \; X_t \cdot \partial_x g + \frac{1}{2} (\sigma_t^x)^2 \; X_t^2 \cdot \partial_{xx} g \\
g_t \sigma_t^g &= \sigma_t^x \; X_t \cdot \partial_x g
\end{aligned}
$$

Substituting these expressions into the equation above we have:

$$
\partial_t g + \left(\mu_t^X - \lambda_t \cdot \sigma_t^X\right) X_t \cdot \partial_x g + \frac{1}{2} \left(\sigma_t^X\right)^2 X_t^2 \cdot \partial_{xx} g = r_t g_t
$$

Here the processes can be written out in terms of the arguments of the functions:

$$
\begin{aligned}
&\partial_t g(t, X_t) + \left[\mu^X(t, X_t) - \lambda(t, X_t) \cdot \sigma^X(t, X_t)\right] X_t \cdot \partial_x g(t, X_t) \\
&\qquad + \frac{1}{2} \left[\sigma^X(t, X_t)\right]^2 X_t^2 \cdot \partial_{xx} g(t, X_t) = r(t, X_t) \cdot g(t, X_t)
\end{aligned}
$$

Since this relationship must hold for all possible paths of X_t it must hold for all values of $(t, x) \in \mathbb{R}^+ \times \mathbb{R}$, which implies that the function g must satisfy the following PDE:

$$\partial_t g(t,x) + \left[\mu^X(t,x) - \lambda(t,x)\cdot\sigma^X(t,x)\right] x \cdot \partial_x g(t,x) \\ + \frac{1}{2}\left[\sigma^X(t,x)\right]^2 x^2 \cdot \partial_{xx} g(t,x) \;=\; r(t,x)\cdot g(t,x)$$

This is the **generalized Black-Scholes partial differential equation. This PDE requires a boundary condition which we can find from the payoff function for the claim**. Recall that $g_t = g(t, X_t)$ is the price process of the claim, and at T the price is equal to payoff function, i.e. $g_T = g(T, X_T) = G(X_T)$. Therefore, the boundary condition for the function g is given by the payoff function G, namely that $g(T,x) = G(x)$.

The following theorems summarize the main results we have discussed thus far:

Theorem 9.1 (Generalized Black-Scholes PDE). *Assume that a source of uncertainty given by an Itô process $X = (X_t)_{t\geq 0}$ satisfies the SDE:*

$$\frac{dX_t}{X_t} \;=\; \mu^X(t,X_t)\;dt + \sigma^X(t,X_t)\;dW_t$$

*where $\mu^X : \mathbb{R}^+ \times \mathbb{R} \to \mathbb{R}$ and $\sigma^X : \mathbb{R}^+ \times \mathbb{R} \to \mathbb{R}^+\backslash\{0\}$ are functions and $W = (W_t)_{t\geq 0}$ is a $\mathbb{P}$-Brownian motion. Further assume that a claim is written on X that pays $G(X_T)$ at time T whose price process, $g = (g_t)_{0\leq t\leq T}$, is Markovian in X, i.e. there exists a function $g : \mathbb{R}^+ \times \mathbb{R} \to \mathbb{R}$ such that $g_t = g(t, X_t)$. Then the function g satisfies the **generalized Black-Scholes PDE**:*

$$\begin{cases} \left(\partial_t + \mathcal{L}_t^X\right) g(t,x) \;=\; r(t,x)\cdot g(t,x) \\ \qquad\qquad\qquad g(T,x) \;=\; G(x) \end{cases}$$

where $\mathcal{L}_t^X = \left(\mu^X(t,x) - \lambda(t,x)\cdot\sigma^X(t,x)\right) x\cdot\partial_x + \frac{1}{2}\left(\sigma^X(t,x)\right)^2 x^2 \cdot \partial_{xx}$

Theorem 9.2 (Market Price of Risk). *Given the assumptions of the previous theorem, all claims written on X have the same **market price of risk**, i.e. if the dynamics of two claims written on X are given by:*

$$\frac{df_t}{f_t} \;=\; \mu^f(t,X_t)\;dt + \sigma^f(t,X_t)\;dW_t$$

$$\frac{dg_t}{g_t} \;=\; \mu^g(t,X_t)\;dt + \sigma^g(t,X_t)\;dW_t$$

Then it follows that:

$$\lambda(t,X_t) \;=\; \frac{\mu^f(t,X_t) - r(t,X_t)}{\sigma^f(t,X_t)} \;=\; \frac{\mu^g(t,X_t) - r(t,X_t)}{\sigma^g(t,X_t)}$$

Furthermore, if X itself is a traded asset then we must have that:

$$\lambda(t,X_t) \;=\; \frac{\mu^X(t,X_t) - r(t,X_t)}{\sigma^X(t,X_t)}$$

and the generalized Black-Scholes PDE reduces to:

$$\begin{cases} \left(\partial_t + \mathcal{L}_t^X\right) g(t,x) &= r(t,x) \cdot g(t,x) \\ g(T,x) &= G(x) \end{cases}$$

where $\mathcal{L}_t^X = r(t,x)x \cdot \partial_x + \frac{1}{2}\left(\sigma^X(t,x)\right)^2 x^2 \cdot \partial_{xx}$

Let us **summarize** the steps in the dynamic hedging argument:

- We start with a set of SDEs describing the evolution of: the source of uncertainty on which a claim is written, X, the value of the claim f, and a risk-free asset, B.
- We are interested in pricing a second claim, g. This is achieved by hedging a sold unit of the claim g by trading in f and the money market account. The hedging portfolio is set up to have zero cost in the outset.
- We determine the SDE that the portfolio value satisfies while imposing a self-financing constraint and force the process to have no randomness by our choice of holdings in f. A second constraint ensures that the known, deterministic increments of the process are equal to zero to avoid arbitrage.
- The conditions we impose and the portfolio we construct leads us to the fact that the market price of risk is equal for all traded assets. Thus, market price of risk does not depend on claims, but rather is a property of the market.[2]
- We use the fact that g can be viewed in two ways: as a process with its drift and volatility, and as a function of the underlying source of uncertainty so that its drift and volatility can be derived from the underlying process using Itô's lemma. Equating these two approaches, and plugging the expressions into the market price of risk allows us to arrive at the Black-Scholes PDE.
- The payoff function for the claim as a function of the underlying gives the PDE a boundary condition.
- If the source of uncertainty X is a traded asset, then we can simplify the PDE by replacing the term $\mu^X(t,x) - \lambda(t,x)\sigma^X(t,x)$ with the short rate process $r(t,x)$.

9.3 Solving the Generalized Black-Scholes PDE

Recall that the unknown function g in the generalized Black-Scholes PDE gives us the price of a claim at every time/state combination, i.e. $g(t,x)$ is the price of the claim at time t when the source of uncertainty is at the level $X_t = x$. If we enter the source of uncertainty itself as the second argument, then we would generate a new process, $g_t = g(t, X_t)$, which is the claim price process.

[2] The theory does not specify λ_t, only that it is the same for all claims; λ_t is given by model assumptions.

We are interested in solving the PDE above for the unknown function g, so that we may compute prices at various times and levels of X_t. Although different claims written on X satisfy the same PDE, they will have different price functions by virtue of having different payoff functions appearing in the boundary condition.

We can solve the generalized Black-Scholes PDE by invoking the Feynman-Kac theorem. In this case the PDE is given by:

$$\begin{cases} \partial_t g(t,x) + \underbrace{\left(\mu^X(t,x) - \lambda(t,x)\cdot\sigma^X(t,x)\right)x}_{\text{F-K drift}} \cdot \partial_x g(t,x) \\ \qquad + \frac{1}{2}\underbrace{\left(\sigma^X(t,x)\right)^2 x^2}_{\text{F-K volatility}} \cdot \partial_{xx} g(t,x) \;=\; \underbrace{r(t,x)}_{\text{"discount rate"}} \cdot g(t,x) \\ \qquad g(T,x) \;=\; G(x) \end{cases}$$

According to the Feynman-Kac theorem, a stochastic representation of the solution to this PDE is given by:

$$g(t,x) \;=\; \mathbb{E}^{\mathbb{P}^*}_{t,x}\left[G(X_T)\cdot e^{-\int_t^T r(s,X_s)\ ds}\right]$$

where X satisfies the SDE:

$$\frac{dX_t}{X_t} = \left(\mu_t^X - \lambda_t \sigma_t^X\right)\ dt + \sigma_t^X\ dW_t^*$$

and where W^* is a $\mathbb{P}^*$-Brownian motion.

At this point, we haven't discussed the nature of this new measure $\mathbb{P}^*$, aside from the fact that it is used in the stochastic representation of the solution to the PDE. However, notice that the exponential term can be written in terms of the bank account asset:

$$\frac{dB_t}{B_t} = r(t,X_t)\ dt \quad \Longrightarrow \quad B_t = e^{\int_0^t r(s,X_s)\ ds}$$
$$\Longrightarrow \quad e^{-\int_t^T r(s,X_s)\ ds} = \frac{B_t}{B_T}$$

Substituting this into the equation for above:

$$\begin{aligned} g(t,x) &= \mathbb{E}^{\mathbb{P}^*}_{t,x}\left[G(X_T)\cdot\frac{B_t}{B_T}\right] \\ \Longrightarrow \quad g(t,X_t) &= \mathbb{E}^{\mathbb{P}^*}_{t,X_t}\left[G(X_T)\cdot\frac{B_t}{B_T}\right] \qquad \text{going from function to process} \\ \Longrightarrow \quad g_t &= \mathbb{E}^{\mathbb{P}^*}\left[g_T\cdot\frac{B_t}{B_T}\;\middle|\;\mathcal{F}_t\right] \\ \Longrightarrow \quad \frac{g_t}{B_t} &= \mathbb{E}^{\mathbb{P}^*}\left[\frac{g_T}{B_T}\;\middle|\;\mathcal{F}_t\right] \qquad \text{since } B_t \text{ is } \mathcal{F}_t\text{-measurable} \end{aligned}$$

Notice that the last statement resembles the fundamental theorem of asset pricing when B plays the role of the numeraire. This implies that $\mathbb{P}^*$ **is in fact the risk-neutral measure invoked by the bank account as numeraire (or simply, the risk-neutral measure), which we denote by** $\mathbb{Q}$**.**

THEOREM 9.3 (Risk-neutral pricing). *Let g_t be the price at time t of a claim written on a source of uncertainty, X, with payoff function $G(x)$, where X has $\mathbb{P}$-dynamics given by the SDE:*

$$\frac{dX_t}{X_t} = \mu_t^X \ dt + \sigma_t^X \ dW_t^{\mathbb{P}}$$

where $W^{\mathbb{P}}$ is a $\mathbb{P}$-Brownian motion. Then g_t is given by:

$$g_t \ = \ \mathbb{E}_t^{\mathbb{Q}}\left[G(X_T)\cdot e^{-\int_t^T r(s,X_s)\ ds}\right]$$

where the $\mathbb{Q}$-dynamics of X are given by the SDE:

$$\frac{dX_t}{X_t} = \left(\mu_t^X - \lambda_t\sigma_t^X\right) \ dt + \sigma_t^X \ dW_t^{\mathbb{Q}}$$

Furthermore, if X is a traded asset, then the $\mathbb{Q}$-dynamics of X are given by:

$$\frac{dX_t}{X_t} = r_t \ dt + \sigma_t^X \ dW_t^{\mathbb{Q}}$$

9.4 Relating $\mathbb{P}$ and $\mathbb{Q}$ via Girsanov's Theorem

Comparing the two SDEs in the previous theorem (dynamics of X under $\mathbb{P}$ and $\mathbb{Q}$), it appears that the two Brownian motions are related:

$$\begin{aligned} \frac{dX_t}{X_t} &= \left(\mu_t^X - \lambda_t\sigma_t^X\right) dt + \sigma_t^X \ dW_t^{\mathbb{Q}} \\ &= \mu_t^X + \sigma_t^X\left(dW_t^{\mathbb{Q}} - \lambda_t \ dt\right) \\ \text{and} \quad \frac{dX_t}{X_t} &= \mu_t^X \ dt + \sigma_t^X \ dW_t^{\mathbb{P}} \\ \overset{?}{\Longrightarrow} \quad dW_t^{\mathbb{Q}} &= \lambda_t \ dt + dW_t^{\mathbb{P}} \end{aligned}$$

This relationship is in fact true, and is justified by Girsanov's theorem. We can view $W^{\mathbb{Q}}$ as a new process defined as a drifted version of $W^{\mathbb{P}}$, where the drift is precisely the market price of risk process λ_t. Therefore, by Girsanov's theorem, there exists a different measure under which $W^{\mathbb{Q}}$ is a standard Brownian motion. That measure, which we will denote $\mathbb{Q}$, is characterized by the Radon-Nikodym derivative process:

$$\eta_t = \left(\frac{d\mathbb{Q}}{d\mathbb{P}}\right)_t \ = \ \exp\left(-\int_0^t \lambda_u \ dW_u - \frac{1}{2}\int_0^t \lambda_u^2 \ du\right)$$

which satisfies the SDE :

$$\frac{d\eta_t}{\eta_t} = -\lambda_t \, dW_t^{\mathbb{P}}$$

9.5 The Black-Scholes Model \ Black-Scholes PDE

The **Black-Scholes model** refers to the case where risky assets have constant drift and volatility parameters (i.e. they are modeled by constant parameter geometric Brownian motions), and where the short rate of interest is a constant. That is, the $\mathbb{P}$-dynamics of a risky asset S and the bank account B are given by:

$$\begin{aligned} \frac{dS_t}{S_t} &= \mu \, dt + \sigma \, dW_t^{\mathbb{P}} \\ \frac{dB_t}{B_t} &= r \, dt \end{aligned}$$

where $W^{\mathbb{P}}$ is a $\mathbb{P}$-Brownian motion. In this case, the market price of risk is also a constant:

$$\lambda = \frac{\mu - r}{\sigma}$$

Moreover, a claim g written on the asset S with payoff function $G(S)$ must satisfy the **Black-Scholes PDE** (a specific version of the generalized Black-Scholes PDE):

$$\begin{cases} \partial_t g(t,S) + rS \cdot \partial_S g(t,S) + \frac{1}{2}\sigma^2 S^2 \cdot \partial_{SS} g(t,S) \;=\; r \cdot g(t,S) \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad g(T,S) \;=\; G(S) \end{cases}$$

Notice that the term $\mu_t^X - \lambda_t^X \sigma_t^X$ was replaced by r in this case where all the parameters are constants since S is traded. A Feynman-Kac representation of the solution to this PDE is given by:

$$g(t,S) \;=\; e^{-r(T-t)} \cdot \mathbb{E}_{t,S}^{\mathbb{Q}} \left[G(S_T) \right]$$

where S satisfies the SDE:

$$\frac{dS_t}{S_t} \;=\; r \, dt + \sigma \, dW_t^{\mathbb{Q}}$$

Notice that the usual integral involving the discount rate in the exponential reduces to $r(T-t)$ since r is constant.

9.6 Exercises

Exercise 9.4. Assume the Black-Scholes model. A claim written on a traded asset, X, has the payoff function $F(x) = x^2$ and the claim price process is

$f_t = f(t, X_t) = h(t)X_t^2$. Find $h(t)$. *Hint: the claim price must satisfy the Black-Scholes PDE.*

EXERCISE 9.5. Suppose S is the price of a traded asset and that

$$\frac{dS_t}{S_t} = \mu\ dt + \sigma\ dW_t$$

where μ, σ are constants. Show that:

(i) $V_t = S_t$ satisfies the Black-Scholes PDE.
(ii) $V_t = S_t^n \cdot \exp\left[\left(r + \frac{1}{2}n\sigma^2\right)(n-1)(T-t)\right]$ satisfies the Black-Scholes PDE.

EXERCISE 9.6. Assuming the Black-Scholes model, determine the value and Delta (for all times $0 < t \le T$) of contingent claims having the following payoffs at time T:

(i) $\varphi(S) = \mathbb{1}_{\{S>K\}}$ (digital option)
(ii) $\varphi(S) = S \cdot \mathbb{1}_{\{S>K\}}$ (asset-or-nothing option)
(iii) $\varphi(S) = S^n$
(iv) $\varphi(S) = (S-K)_+^m$, $m \in \mathbb{N}$ (power call option)
(v) $\min(S_T\ ;\ k\ S_U)$ where $U < T$

EXERCISE 9.7. Assuming the Black-Scholes model, use Excel to plot the price versus spot level for each option in Question 26 using the following sets of parameters (put each parameter set on a single plot):

(i) $T = \{\frac{1}{4}, \frac{1}{2}, 1\}$; $\sigma = 20\%$; $r = 5\%$; $\delta = 3\%$
(ii) $T = 1$; $\sigma = \{10\%, 20\%, 30\%\}$; $r = 5\%$; $\delta = 3\%$
(iii) $T = 1$; $\sigma = 20\%$; $r = \{0\%, 5\%, 10\%\}$; $\delta = 3\%$
(iv) $T = 1$; $\sigma = 20\%$; $r = 5\%$; $\delta = \{0\%, 3\%, 6\%\}$

EXERCISE 9.8. Assuming the Black-Scholes model, run a Monte Carlo simulation and obtain the present value profit and loss histogram of the options in Question 26 for each of the parameter sets in Question 27 assuming that $S = 100$ and $\mu = 10\%$. Use 50,000 simulations to obtain the PnL.

EXERCISE 9.9. Using the Black-Scholes model, determine the price, the delta and the gamma for all times $t \in [0, T)$ of the following European options with payoffs at time $T > 0$:

(i) Digital put option. That is, $\varphi = \mathbb{1}\{S_T < K\}$.
(ii) Digital call option. That is, $\varphi = \mathbb{1}\{S_T > K\}$.

(iii) A forward-start digital call option, which pays 1 at T if the asset price at maturity is above a percentage α of the asset price at time U (where $t < U < T$). That is, $\varphi = \mathbb{1}\{S_T > \alpha S_U\}$.

(iv) A forward-start asset-or-nothing option which pays the asset at T if the asset price at maturity is above a percentage α of the asset price at time U (where $t < U < T$). That is, $\varphi = S_T\, \mathbb{1}\{S_T > \alpha S_U\}$.

(v) A call option (maturing at V) on a forward-start asset-or-nothing option. The embedded forward-start asset-or-nothing option pays the asset at $T > V$ if the asset price at T is above a percentage α of the asset price at time U (where $V < U < T$). The strike of the call option is K.

Exercise 9.10. Repeat the dynamic hedging argument assuming that the source of uncertainty is a traded dividend-paying stock where holding 1 unit of the asset pays $\delta_t S_t dt$ over a small period of time dt. *Hint: how does this asset affect dV_t?*

Exercise 9.11. Let X and Y be two sources of uncertainty satisfying:

$$\frac{dX_t}{X_t} = \mu^X(t, X_t)\; dt + \sigma^X(t, X_t)\; dW_t^X$$
$$\frac{dY_t}{Y_t} = \mu^Y(t, Y_t)\; dt + \sigma^Y(t, Y_t)\; dW_t^Y$$

Assume two claims are written on X and Y and that their price processes are Markovian, i.e. $\exists\, f, g$ such that $f_t = f(t, X_t)$ and $g_t = g(t, Y_t)$ with

$$\frac{df_t}{f_t} = \mu^f(t, f_t)\; dt + \sigma^f(t, f_t)\; dW_t^X$$
$$\frac{dg_t}{g_t} = \mu^g(t, g_t)\; dt + \sigma^g(t, g_t)\; dW_t^Y$$

Assume further a money market account, B, satisfying:

$$\frac{dB_t}{B_t} = r(t, X_t, Y_t)\; dt$$

Let $h_t = h(t, X_t, Y_t)$ be the price of a claim written on X and Y with terminal payoff $H(X_T, Y_T)$.

(i) Determine the PDE that h must satisfy.
(ii) Give a Feynman-Kac representation of the solution.
(iii) How does the solution change if both X and Y are traded assets?

Exercise 9.12. Suppose that two indices (not necessarily traded) have processes X_t and Y_t and that they are jointly GBMs, i.e.

$$\frac{dX_t}{X_t} = \mu_x\, dt + \sigma_x\, dW_t^x\;, \qquad \frac{dY_t}{Y_t} = \mu_y\, dt + \sigma_y\, dW_t^y\;, \tag{9.1}$$

where W_t^X and W_t^Y are correlated standard Brownian motions under the $\mathbb{P}$-measure with correlation ρ. Consider a contingent claim f written on the two indices with payoff $\varphi(X_T, Y_T)$ at time T.

(i) ** Use a dynamic hedging argument to demonstrate that to avoid arbitrage, the price of f must satisfy the following PDE:

$$\begin{cases} \big(\partial_t + (\mu_x - \lambda_x\sigma_x)\, x\partial_x + (\mu_y - \lambda_y\sigma_y)\, y\partial_y \\ \qquad + \tfrac{1}{2}\sigma_x^2\, x^2\, \partial_{xx} + \tfrac{1}{2}\sigma_y^2\, y^2\, \partial_{yy} + \rho\sigma_x\sigma_y\, x\, y\, \partial_{xy}\big)\, f &= r\, f \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad f(T, x, y) &= \varphi(x, y)\ , \end{cases} \tag{9.2}$$

where λ_x and λ_y are market prices of risk for the two risk factors.

(ii) Suppose that the payoff is homogenous, so that $\varphi(x, y) = y\, g(x/y)$ for some function g. An example of such a payoff is the payoff from an exchange option which would have $\varphi(x, y) = (x - y)_+$. By assuming that $f(t, x, y) = y\, b(t)\, h(t, a(t)\, x/y)$, and that λ_x and λ_y are constants, find the PDE which h satisfies and show that the price f can be written in the form

$$f(X_t, Y_t) = \mathbb{E}_T^{\mathbb{Q}}[Y_t]\; \mathbb{E}_t^{\mathbb{Q}^*}\left[g(U_T)\right] \tag{9.3}$$

where, $U_t = \mathbb{E}_T^{\mathbb{Q}}[X_t]/\mathbb{E}_T^{\mathbb{Q}}[Y_t]$ and U_t satisfies an SDE of the form

$$\frac{dU_t}{U_t} = \sigma_U\, dW_t^*\ ,$$

for some constant σ_U and W_t^* a $\mathbb{Q}^*$ Brownian motion.

(iii) Obtain the expectation result of part (b) by using a measure change – note that since X_t and Y_t are not traded assets, they cannot be used as numeraires. However, a measure change using one of those processes is still possible.

10 Dynamic Hedging Implementation

In this note we consider how dynamic hedging strategies can be implemented in practice. The main issue to consider is the fact that the dynamic hedging argument was developed in continuous-time, however , in practice, portfolios can only be rebalanced at discrete times.

10.1 Delta-Hedging

Recall the **dynamic hedging argument** where we begin by assuming the following model for the economy:

(i) A **source of uncertainty** given by the Itô process $X = (X_t)_{t\geq 0}$ satisfying the SDE:

$$\frac{dX_t}{X_t} = \mu_t^X\ dt + \sigma_t^X\ dW_t$$

(ii) A **bank account** (process), $B = (B_t)_{t\geq 0}$, satisfying:

$$\frac{dB_t}{B_t} = r_t\ dt$$

(iii) A **traded claim written on the source of uncertainty** whose price process, $f = (f_t)_{t\geq 0}$, satisfies:

$$\frac{df_t}{f_t} = \mu_t^f\ dt + \sigma_t^f\ dW_t$$

(iv) A **second contingent claim** written on X that we are interested in pricing, whose price process we will denote by $g = (g_t)_{0\leq t\leq T}$ and that pays $G(X_T)$ at time T. Recall that we assume that g is Markovian in X, i.e. that there exists a function $g : \mathbb{R}^+ \times \mathbb{R} \to \mathbb{R}$ such that $g_t = g(t, X_t)$ and g_t satisfies:

$$\frac{dg_t}{g_t} = \mu_t^g\ dt + \sigma_t^g\ dW_t$$

The dynamic hedging argument involves constructing a zero cost, self-financing portfolio with $(\alpha, \beta) = (\alpha_t, \beta_t)_{t\geq 0}$ units invested in f and B and short one unit of g. This leads to the portfolio value dynamics:

$$dV_t = \left(\alpha_t f_t \mu_t^f + \beta_t B_t r_t - g_t \mu_t^g\right) dt + \left(\alpha_t f_t \sigma_t^f - g_t \sigma_t^g\right) dW_t$$

We choose α_t to achieve an instantaneously risk-free portfolio, by setting the coefficient of dW_t to be zero:

$$\begin{aligned} \alpha_t^* f_t \sigma_t^f - g_t \sigma_t^g &= 0 \\ \implies \quad \alpha_t^* &= \frac{g_t \ \sigma_t^g}{f_t \ \sigma_t^f} \end{aligned}$$

Now, if X is traded, then it can be used as the hedging instrument, i.e. taking $f_t = X_t$. In this case α_t becomes:

$$\alpha_t^* = \frac{g_t \ \sigma_t^g}{X_t \ \sigma_t^X}$$

Furthermore, recall from Itô's lemma that the volatility of the claim price process g is given by $g_t \sigma_t^g = \partial_x g(t, X_t) \cdot \sigma_t^X X_t$. This implies that:

$$\alpha_t^* = \partial_x g(t, X_t)$$

The quantity $\partial_x g(t, X_t)$ is known as the **delta** of the option g (denoted Δ^g), and represents the change in the option value due to an increase in the value of the underlying asset on which the option is written. The delta of the portfolio is defined analogously:

$$\begin{aligned} V_t &= \alpha_t^* X_t + \beta_t B_t - g_t \\ \implies \quad \Delta^V &= \underbrace{\alpha_t^*}_{=\Delta^g} \underbrace{\Delta^X}_{=1} + \beta_t \underbrace{\Delta^B}_{=0} - \Delta^g \\ \implies \quad \Delta^V &= 0 \end{aligned}$$

In other words, the dynamic hedging portfolio is a **delta-neutral** portfolio whose value is unaffected by (first-order) changes in the value of the underlying asset X.

In practice, the question we are interested in is that of a trader that sells an option and wishes to hedge their short position. Ideally, implementing the dynamic hedging leads to zero profit-and-loss (PnL) as we would have $V_0 = 0$ and $V_t = 0$ for all t, whoever this would require trading continuously. In practice, trading to rebalance the hedging portfolio is done discretely which leads to variance in the trader's PnL.

Next, **we analyze the evolution of the trader's PnL when the hedging portfolio is rebalanced discretely** at an arbitrary set of rebalancing times $\{t_0, t_1, ..., t_n\}$, where $t_0 = 0$ and $t_n = T$ and $\Delta t_i = t_i - t_{i-1}$. We will consider the value of the portfolio prior to rebalancing, i.e. at t_i^-, and after rebalancing.

10.2 Delta-Gamma-Hedging

As we have discussed in the previous section, delta-hedging immunizes a portfolio from first-order changes in the underlying asset's value. A natural question that arises is: **can we do better than delta-hedging**? A simple extension of delta-hedging is to estimate the slope using a quadratic function instead of a straight line. This is the basis for **delta-gamma-hedging**.

To understand the intuition let us consider how the value of a contigent claim changes if the price of the underlying asset changes. To this end, consider the following second order Taylor expansion:

$$\begin{aligned} g(t, X_{t+\Delta t}) &= g(t, X_t + \Delta X_t) \\ &= g(t, X_t) + \partial_x g(t, X_t) \cdot \Delta X_t + \frac{1}{2} \cdot \partial_{xx} g(t, X_t) \cdot (\Delta X_t)^2 \end{aligned}$$

The second-order derivative $\partial_{xx} g(t, X_t)$ is known as the **gamma** of the option g, denoted Γ^g. The idea behind delta-gamma-hedging is to construct a portfolio that has zero net delta and zero net gamma. This extends the notion of delta-hedging where the only requirement was to have zero net delta.

Now, notice that since the underlying asset has a delta of 1 and a gamma of 0, it is not possible to construct a delta-gamma neutral portfolio using only the underlying asset and the bank account. To achieve a delta-gamma neutral portfolio we need to introduce a hedging option with value $h(t, X_t)$. If we hold $(\alpha, \beta, \eta) = (\alpha_t, \beta_t, \eta_t)_{t\geq 0}$ units in (X, B, h) to hedge our unit short exposure in g then the value of our portfolio is given by:

$$V_t = \alpha_t X_t + \beta_t B_t + \eta_t h_t - g_t$$

The delta and gamma of this portfolio is equal to:

$$\begin{aligned} \Delta_t^V &= \alpha_t + \eta_t \Delta_t^h - \Delta_t^g \\ \Gamma_t^V &= \eta_t \Gamma_t^h - \Gamma_t^g \end{aligned}$$

Setting the two equations above to zero to achieve delta-gamma-neutrality allows us to solve for the required number of units we need to hold in the underlying asset and the hedging option:

$$\alpha_t^* = \Delta_t^g - \frac{\Gamma_t^g}{\Gamma_t^h} \Delta_t^g \qquad\qquad \eta_t^* = \frac{\Gamma_t^g}{\Gamma_t^h}$$

Time	Evolution of Hedging Portfolio	
$t = t_0$	*sell option:*	receive $g_{t_0} = g(t_0, X_{t_0})$
	buy α_{t_0} units of X:	pay $\alpha_{t_0} X_{t_0}$
	bank account position:	$B_{t_0} = g_{t_0} - \alpha_{t_0} X_{t_0}$
	(if $B_{t_0} < 0$ we are borrowing to buy units of X, otherwise we are placing the cash from the short sale of X into the bank account)	
$t = t_1^-$	*option value:*	$g_{t_1^-} = g\left(t_1^-, X_{t_1^-}\right)$
	asset holdings value:	$\alpha_{t_0} X_{t_1^-}$
	bank account position:	$B_{t_1^-} = (g_{t_0} - \alpha_{t_0} X_{t_0})\, e^{r\Delta t_1}$
$t = t_1$	*option value:*	$g_{t_1} = g\,(t_1, X_{t_1})$
	rebalance to α_{t_1} in X:	sell $\alpha_{t_0} X$, buy $\alpha_{t_1} X$
		$\Longrightarrow$ **cash flow** $= (\alpha_{t_0} - \alpha_{t_1}) X_{t_1}$
	bank account position:	$B_{t_1} = (g_{t_0} - \alpha_{t_0} X_{t_0})\, e^{r\Delta t_1} + (\alpha_{t_0} - \alpha_{t_1}) X_{t_1}$
	(purchase/sale of X is financed by borrowing/lending from bank account)	
	$\vdots$	
$t = t_k^-$	*option value:*	$g_{t_k^-} = g\left(t_k^-, X_{t_k^-}\right)$
	asset holdings value:	$\alpha_{t_{k-1}} X_{t_k^-}$
	bank account position:	$B_{t_k^-} = B_{t_{k-1}} e^{r\Delta t_k}$
$t = t_k$	*option value:*	$g_{t_k} = g\,(t_k, X_{t_k})$
	rebalance to α_{t_k} in X:	sell $\alpha_{t_{k-1}} X$, buy $\alpha_{t_k} X$
		$\Longrightarrow$ **cash flow** $= (\alpha_{t_{k-1}} - \alpha_{t_k}) X_{t_k}$
	bank account position:	$B_{t_k} = B_{t_{k-1}} e^{r\Delta t_k} + (\alpha_{t_{k-1}} - \alpha_{t_k}) X_{t_k}$
	(purchase/sale of X is financed by borrowing/lending from bank account)	
	$\vdots$	
$t = t_n$	*option payoff owed:*	$G\,(X_{t_n})$
	asset holdings value:	$\alpha_{t_{n-1}} X_{t_n}$
	bank account position:	$B_{t_n} = B_{t_{n-1}} e^{r\Delta t_n}$

$$\text{PnL} = \alpha_{t_{n-1}} X_{t_n} + B_{t_{n-1}} e^{r\Delta t_n} - G\,(X_{t_n})$$

11 Forward and Futures Contracts

The focus of this note will be to develop a dynamic hedging argument involving futures contracts and claims written on futures prices. The first step will be to define futures contracts. For this, we will first need to define *forward* contracts and futures *prices* - both of which distinct from futures *contracts*.

11.1 Forward and Futures Contracts

Forward contracts are agreements to buy (for a long position) or sell (for a short position) an asset at a fixed date T for a delivery price K which is set when the contract is written. In other words, a forward contract is a claim that has the payoff function $\varphi(S_T) = S_T - K$. To see why this is, notice that if you have a long position in a forward contract, then you must buy the asset at a price of K at time T. If you were to then immediately sell the asset you would receive its prevailing price, S_T, so that your total payoff is $S_T - K$.

The value of a forward contract is given by:

$$\begin{aligned} g(t, S_t) &= \mathbb{E}_t^{\mathbb{Q}}\left[e^{-\int_t^T r_s\, ds} \cdot (S_T - K)\right] \\ &= \mathbb{E}_t^{\mathbb{Q}}\left[e^{-\int_t^T r_s\, ds} \cdot S_T\right] - K \cdot \mathbb{E}_t^{\mathbb{Q}}\left[e^{-\int_t^T r_s ds}\right] \\ &= S_t - K \cdot P_t(T) \end{aligned}$$

where $P_t(T)$ is the price of a T-maturity bond at time t.

The **forward price** is the delivery price K that sets this value of the contract to zero at its inception. Thus, the forward price is given by:

$$F_t(T) = \frac{S_t}{P_t(T)}$$

Notice that this is a $\mathbb{Q}^T$-martingale, where $\mathbb{Q}^T$ is the T-forward neutral measure, i.e. the EMM induced by using the T-maturity bond as numeraire. Notice that due to the FTAP we also have that:

$$\begin{aligned} F_t(T) &= \frac{S_t}{P_t(T)} \\ &= \mathbb{E}_t^{\mathbb{Q}^T}\left[\frac{S_T}{P_T(T)}\right] \\ &= \mathbb{E}_t^{\mathbb{Q}^T}[S_T] \end{aligned}$$

which highlights the fact that the forward price is a $\mathbb{Q}^T$-martingale as it can be viewed as a Doob martingale.

Finally, a **futures contract** gives the holder the obligation to buy/sell an asset at time T at the prevailing futures price $F_t(T)$. The mechanics behind futures contracts differ from forward contracts. With the former, the delivery price is fixed at the outset. With futures contracts, however, the delivery price (futures price) fluctuates through time. To account for this fluctuation, **the holder of the futures contract is paid the change in the futures price**. Since the contract is costless to enter, **it has zero value at the outset**, and since the holder is compensated with a cash flow through time, **the futures contract always has zero value**.

To summarize, a futures contract is a position that is entered at zero cost, and that provides the holder with a stochastic cash flow stream driven by the change in the futures price.[1]

11.2 Dynamic Hedging with Futures Contracts

Now, **we develop a dynamic hedging argument to value a claim written on a futures price**. To this end, let $F_t(T)$ denote the futures price[2] at time t, and assume that it satisfies the SDE:

$$\frac{dF_t}{F_t} = \mu_t^F\ dt + \sigma_t^F\ dW_t^{\mathbb{P}}$$

Assume also that there is a money market account with the usual (constant interest rate) dynamics:

$$\frac{dB_t}{B_t} = r\ dt$$

Assume now that **we wish to value a claim on the futures price**, g_t, that pays $G(F_{T_0})$ at maturity date $T_0 \le T$. Here T_0 is the maturity date of the claim, while T is the delivery date for the futures contract. Since the claim is a function

[1] An alternative view is to think of a futures price as a fictitious price (since it is NOT the price of a traded asset) that is a proxy for the underlying asset to be bought at at a future date. The futures contract can be viewed as providing a stochastic dividend through time based on the futures price.

[2] Note that this is NOT the value of the futures contract (which is always equal to 0).

of time and the underlying futures price we can write $g_t = g(t, F_t)$ and, by Itô's lemma, we can deduce that the claim satisfies the SDE:

$$dg_t = (\partial_t + \mathcal{L})g_t \; dt + \sigma_t^F F_t \cdot \partial_F g_t \; dW_t^{\mathbb{P}}$$

where $\mathcal{L}$ is the infinitesimal generator of the process: $\mathcal{L} = \mu_t F_t \cdot \partial_F g + \frac{1}{2}(\sigma_t^F F_t)^2 \cdot \partial_{FF} g$.

Next, we **set up a hedging portfolio with a starting value of zero**, $V_0 = 0$, with $(\alpha_t, \beta_t, -1)$ in (f_t, B_t, g_t), where f_t is the *value*[3] of the futures contract. However, since this value is always equal to zero, the value of the portfolio is given by:

$$V_t = \beta_t B_t - g_t$$

Next, we **derive the dynamics of the portfolio value**. For this, we must bear in mind that although the futures contract does not have any value, it *does* provide a cash flow which influences the evolution of the portfolio's value. The cash flow is equal to the number of units held times the change in the futures price. Combining this with the usual self-financing constraint we find that the portfolio value satisfies the SDE:

$$dV_t = \alpha_t \; dF_t + \beta_t \; dB_t - dg_t$$

Substituting in the dynamics of the futures price, the bank account and g_t, this becomes:

$$\begin{aligned} dV_t &= \alpha_t \; \left(\mu_t^F F_t \; dt + \sigma_t^F F_t \; dW_t^{\mathbb{P}}\right) + \beta_t r B_t \; dt - (\partial_t + \mathcal{L})g_t \; dt - \sigma_t^F F_t \cdot \partial_F g_t \; dW_t^{\mathbb{P}} \\ &= \left(\alpha_t \mu_t^F F_t + \beta_t r B_t - (\partial_t + \mathcal{L})g_t\right) \; dt + \left(\alpha_t \sigma_t^F F_t - \sigma_t^F F_t \cdot \partial_F g_t\right) \; dW_t^{\mathbb{P}} \end{aligned}$$

To **locally remove risk** we set our position in the futures contract to be $\alpha_t = \partial_F g_t$, which forces the coefficient of the stochastic term to be zero and updates the drift term:

$$\begin{aligned} dV_t &= \left(\mu_t^F F_t \cdot \partial_F g + \beta_t r B_t - (\partial_t + \mathcal{L})g_t\right) \; dt \\ &= \left(\mu_t^F F_t \cdot \partial_F g + \beta_t r B_t - \partial_t g - \mu_t F_t \cdot \partial_F g - \frac{1}{2}(\sigma_t^F F_t)^2 \cdot \partial_{FF} g\right) \; dt \\ &= \left(\beta_t r B_t - \partial_t g - \frac{1}{2}(\sigma_t^F F_t)^2 \cdot \partial_{FF} g\right) \; dt \end{aligned}$$

Since $V_0 = 0$, **the no-arbitrage condition forces the drift term to also be equal to zero**. As a consequence, it follows that $V_t = 0$ for all t. But since $V_t = \beta_t B_t - g_t$, it must be the case that $\beta_t B_t = g_t$. Setting the drift term equal to zero and substituting in this last equation we finally have:

$$r g_t - \partial_t g - \frac{1}{2}(\sigma_t^F F_t)^2 \cdot \partial_{FF} g = 0$$

[3] Again, not to be confused with the futures price.

Combining this PDE with the boundary condition given by the claim's payoff at maturity we have:

$$\begin{cases} \partial_t g + \frac{1}{2}(\sigma_t^F F_t)^2 \cdot \partial_{FF} g = r_t g \\ \qquad\qquad g(T_0, F) = G(F) \end{cases}$$

We can **invoke the Feynman-Kac theorem to obtain a stochastic representation of the solution** to this PDE:

$$g(t, F) = \mathbb{E}^{\mathbb{Q}}\left[e^{-(r(T_0-t))} \cdot G(F_{T_0}) \;\middle|\; F_t = F\right]$$
$$\text{where} \quad dF_t = \sigma_t^F F_t \; dW_t^{\mathbb{Q}}$$

A few comments on this result:

- Similar to the previous case, the price of the claim is given as the expected discounted payoff computed under the risk-neutral measure $\mathbb{Q}$. The fact that we are using the bank account as a numeraire justifies the use of "risk-neutral measure" when referring to $\mathbb{Q}$ in this case.
- Under the risk-neutral measure, futures prices are martingales (as there is no drift in the above SDE). This is in contrast to traded assets whose drift is equal to the risk-free rate under $\mathbb{Q}$.[4]
- Traded assets that pay dividends follow the SDE under risk-neutral measure:

  $$\frac{dS_t}{S_t} = (r - \delta)\; dt + \sigma \; dW_t^{\mathbb{Q}}$$

  That is, they grow at a rate of $r - \delta$ under the risk-neutral measure. Notice that futures prices behave in the same way as prices of dividend-paying assets with $\delta = r$.

[4] As reminder, recall that this implies that traded assets are not martingales under $\mathbb{Q}$ - however, *discounted* traded assets are martingales under $\mathbb{Q}$

12 Numeraire Changes

Numeraire changes can be used to simplify the computation of expectations in derivative valuation problems. In this note we will develop techniques for changing numeraires and associated measures to solve such problems.

12.1 Motivating Example: Exchange Option

Assume we have two risky assets and one riskless asset in our economy with the following $\mathbb{P}$-dynamics:

$$\begin{aligned}\frac{dA_t}{A_t} &= \mu^A(t, A_t, B_t)\,dt + \sigma^A(t, A_t, B_t)\,dW_t^{\mathbb{P},A}\\ \frac{dB_t}{B_t} &= \mu^B(t, A_t, B_t)\,dt + \sigma^B(t, A_t, B_t)\,dW_t^{\mathbb{P},B}\\ \frac{dM_t}{M_t} &= r(t, A_t, B_t)\,dt\end{aligned}$$

where $W^{\mathbb{P},A}$ and $W^{\mathbb{P},B}$ are standard Brownian motions under the physical measure, $\mathbb{P}$, with instantaneous correlation ρ that drive assets A and B respectively.

Now, suppose that we are interested in valuing an **exchange option** (that is, an option to exchange one risky asset for another risky asset at maturity) written on A for B. This claim has the payoff function $G(A_T, B_T) = (A_T - B_T)_+$, and we know, based on previous discussions, that **its value can be computed in one of two ways**:

(i) PDE approach:

Following a dynamic hedging argument, it can be shown that if the price of the claim is given by $g_t = g(t, A_t, B_t)$, then the function g must satisfy the PDE:

$$\begin{cases}\left(\partial_t + \mathcal{L}^{A,B}\right) g(t,a,b) = r(t,a,b)\cdot g(t,a,b)\\ g(T,a,b) = (a-b)_+\end{cases}$$

where the infinitesimal generator $\mathcal{L}^{A,B}$ is obtained through the application of the 2-dimensional version of Itô's lemma:

$$\begin{aligned}\mathcal{L}^{A,B} = \partial_t &+ (r(t,a,b)\cdot a)\cdot\partial_a + \frac{1}{2}\left(\sigma^A(t,a,b)\cdot a\right)^2\cdot\partial_{aa} \\ &+ (r(t,a,b)\cdot b)\cdot\partial_b + \frac{1}{2}\left(\sigma^B(t,a,b)\cdot b\right)^2\cdot\partial_{bb} \\ &+ \rho\,\sigma^A(t,a,b)\sigma^B(t,a,b)\cdot ab\cdot\partial_{ab}\end{aligned}$$

(ii) Expectation approach:

By invoking the Feynman-Kac theorem, we find that the solution to the PDE above admits the following stochastic representation:

$$g(t,a,b) = \mathbb{E}^{\mathbb{Q}}_{t,a,b}\left[e^{-\int_t^T r(s,A_s,B_s)ds}\cdot(A_T-B_T)_+\right]$$

where
$$\frac{dA_t}{A_t} = r(t,A_t,B_t)\;dt + \sigma^A(t,A_t,B_t)\;dW_t^{\mathbb{Q},A}$$
$$\frac{dB_t}{B_t} = r(t,A_t,B_t)\;dt + \sigma^B(t,A_t,B_t)\;dW_t^{\mathbb{Q},B}\quad;\quad d\left[W^{\mathbb{Q},A},W^{\mathbb{Q},B}\right]_t = \rho\;dt$$

Alternatively, we can arrive at this solution directly by invoking the Fundamental Theorem of Asset Pricing using the riskless asset M_t as the numeraire[1] and noting that assets have a drift of r_t under the risk-neutral measure, $\mathbb{Q}$, induced by using the bank account as a numeriare.

The problem we face at this point is that **the PDE is difficult to solve and the expectation is difficult to compute**. We will focus on the expectation approach, but rather than tackling the problem directly, **we will sidestep the difficulties by using a different numeraire**. To see where the simplification comes from, assume that B were the numeraire. Then by the FTAP we have:

$$\begin{aligned}\frac{g_t}{B_t} &= \mathbb{E}_t^{\mathbb{Q}^B}\left[\frac{g_T}{B_T}\right] \\ &= \mathbb{E}_t^{\mathbb{Q}^B}\left[\frac{(A_T-B_T)_+}{B_T}\right] \\ &= \mathbb{E}_t^{\mathbb{Q}^B}\left[\left(\frac{A_T}{B_T}-1\right)_+\right]\end{aligned}$$

This expectation is much simpler to compute than the previous one. Firstly, it only involves the process $Z_t = \frac{A_t}{B_t}$, which happens to be a martingale under $\mathbb{Q}^B$ measure since it is a relative price process with the numeraire in the denominator. This observation will simplify any calculations we need to perform, since we can ignore the drift term of Z_t as we know it to be zero. Moreover, the payoff function now resembles a vanilla call option with a strike price of 1, which is an object that we have come across in past discussions. The crucial point now is that **we require the dynamics of A_t and B_t under the new measure** $\mathbb{Q}^B$ in order to compute the expectation.

[1] Notice that the terms above can be re-arranged to obtain relative prices on both sides of the equation - with the numeraire in the denominator - and a Doob martingale on the RHS.

12.2 Numeraire and Measure Changes

We are interested in finding the Radon-Nikodym derivatives that allow us to change between EMMs induced by different numeraires. Assume that an economy admits no arbitrage opportunities. Then, by FTAP, there exists a numeraire asset B ($B_t > 0$ a.s. for all t) and a probability measure $\mathbb{Q}^B \sim \mathbb{P}$ such that for every traded asset A we have that $\frac{A_t}{B_t} = \mathbb{E}_t^{\mathbb{Q}^B}\left[\frac{A_u}{B_u}\right]$ for $u > t$. If we assume further that there exists another asset, C, with $C_t > 0$ a.s. for all t then there *must* exist a probability measure $\mathbb{Q}^C \sim \mathbb{P}$ such that for every traded asset A we have that $\frac{A_t}{C_t} = \mathbb{E}_t^{\mathbb{Q}^C}\left[\frac{A_u}{C_u}\right]$ for $u > t$. Since $\mathbb{Q}^B, \mathbb{Q}^C \sim \mathbb{P}$, it follows that $\mathbb{Q}^B \sim \mathbb{Q}^C$ and therefore there exists a Radon-Nikdoym derivative, $\frac{d\mathbb{Q}^C}{d\mathbb{Q}^B}$, such that for all $\mathcal{F}_u$-measurable random variables:

$$\mathbb{E}_t^{\mathbb{Q}^C}[X_u] = \frac{\mathbb{E}_t^{\mathbb{Q}^B}\left[X_u \cdot \frac{d\mathbb{Q}^C}{d\mathbb{Q}^B}\right]}{\mathbb{E}_t^{\mathbb{Q}^B}\left[\frac{d\mathbb{Q}^C}{d\mathbb{Q}^B}\right]} \qquad \text{for } u > t$$

The goal is to identify the Radon-Nikdoym derivative $\frac{d\mathbb{Q}^C}{d\mathbb{Q}^B}$. By FTAP, it follows that $A_t = B_t \cdot \mathbb{E}_t^{\mathbb{Q}^B}\left[\frac{A_u}{B_u}\right]$ and $A_t = C_t \cdot \mathbb{E}_t^{\mathbb{Q}^C}\left[\frac{A_u}{C_u}\right]$. Equating these two terms and rearranging we have:

$$\begin{aligned}
\mathbb{E}_t^{\mathbb{Q}^C}\left[\frac{A_u}{C_u}\right] &= \frac{\mathbb{E}_t^{\mathbb{Q}^B}[A_u/B_u]}{C_t/B_t} \\
&= \frac{\mathbb{E}_t^{\mathbb{Q}^B}[A_u/C_u \cdot C_u/B_u]}{C_t/B_t} && \text{multiply and divide by } C_u \\
&= \frac{\mathbb{E}_t^{\mathbb{Q}^B}\left[A_u/C_u \cdot \mathbb{E}_u^{\mathbb{Q}^B}[C_T/B_T]\right]}{\mathbb{E}_t^{\mathbb{Q}^B}[C_T/B_T]} && \text{by FTAP where } t < u < T \\
&= \frac{\mathbb{E}_t^{\mathbb{Q}^B}\left[\mathbb{E}_u^{\mathbb{Q}^B}[A_u/C_u \cdot C_T/B_T]\right]}{\mathbb{E}_t^{\mathbb{Q}^B}[C_T/B_T]} && \text{since } A_u/C_u \text{ is } \mathcal{F}_u\text{-measurable} \\
&= \frac{\mathbb{E}_t^{\mathbb{Q}^B}[A_u/C_u \cdot C_T/B_T]}{\mathbb{E}_t^{\mathbb{Q}^B}[C_T/B_T]} && \text{“smaller } \sigma\text{-alebgra wins”}
\end{aligned}$$

Comparing the last expression to the required measure change property given above it appears that C_T/B_T is the Radon-Nikodym derivative we seek. This is *almost* correct; it is true that $C_T/B_T > 0$ a.s. but since the Radon-Nikodym derivative must have expected value equal to 1 under the original measure and since $\mathbb{E}^{\mathbb{Q}^B}\left[\frac{C_T}{B_T}\right] = \frac{C_0}{B_0}$ by FTAP, it follows that the required Radon-Nikodym derivative is:

$$\eta_T = \frac{d\mathbb{Q}^C}{d\mathbb{Q}^B} = \frac{C_T/C_0}{B_T/B_0}$$

We can define the associated Radon-Nikodym process:

$$\eta_t = \mathbb{E}_t^{\mathbb{Q}^B}\left[\frac{d\mathbb{Q}^C}{d\mathbb{Q}^B}\right] = \mathbb{E}_t^{\mathbb{Q}^B}\left[\frac{C_T/C_0}{B_T/B_0}\right] = \frac{C_t/C_0}{B_t/B_0}$$

The last equality follows from FTAP. Notice now that the process η_t satisfies:

(i) η_t is a $\mathbb{Q}^B$-martingale, since it is equal to a relative price process with the numeraire in the denominator (a martingale by FTAP) multiplied by a positive constant.

(ii) $\mathbb{E}_0[\eta_t] = \eta_0 = 1$

(iii) $\eta_t > 0$ a.s. for all t

Therefore, we can conclude that η_t **is a valid Radon-Nikodym process, which may be used to perform a measure change from the EMM induced by B to the EMM induced by C.**

Next, we **deduce the dynamics of the Radon-Nikodym process η under $\mathbb{Q}^B$.** For this we need to assume some dynamics for the two assets under $\mathbb{Q}^B$:

$$\frac{dB_t}{B_t} = \mu_t^B\ dt + \sigma_t^B\ dW_t^{\mathbb{Q}^B,B}$$
$$\frac{dC_t}{C_t} = \mu_t^C\ dt + \sigma_t^C\ dW_t^{\mathbb{Q}^B,C}$$

By Itô's quotient rule, it follows that

$$\begin{aligned}
\frac{d\eta_t}{\eta_t} &= \frac{d(C_t/B_t)}{C_t/B_t} \\
&= \frac{dC_t}{C_t} - \frac{dB_t}{B_t} + \frac{d[B,B]_t}{B_t^2} - \frac{d[B,C]_t}{B_tC_t} \\
&= \mu_t^C\ dt + \sigma_t^C\ dW_t^{\mathbb{Q}^B,C} - (\mu_t^B\ dt + \sigma_t^B\ dW_t^{\mathbb{Q}^B,B}) + \left(\sigma_t^B\right)^2 dt - \rho\sigma_t^B\sigma_t^C\ dt \\
&= -\sigma_t^B\ dW_t^{\mathbb{Q}^B,B} + \sigma_t^C\ dW_t^{\mathbb{Q}^B,C}
\end{aligned}$$

In the last line above, we can drop all the dt terms since we know that η_t is a $\mathbb{Q}^B$-martingale so its drift must be zero. Now, we can write the SDE for η in the following manner:

$$\frac{d\eta_t}{\eta_t} = -\boldsymbol{\lambda}_t'\ d\mathbf{W}_t^{\mathbb{Q}^B}$$

where $\boldsymbol{\lambda}_t = \left(\sigma_t^B,\ -\sigma_t^C\right)'$. Writing the SDE this way lets us **use Girsanov's theorem to deduce the drift correction that relates Brownian motions in $\mathbb{Q}^B$ and $\mathbb{Q}^C$.** In particular, Girsanov's theorem states that the Brownian motions are related as follows:

$$d\mathbf{W}_t^{\mathbb{Q}^C} = \boldsymbol{\rho}\boldsymbol{\lambda}_t\ dt + d\mathbf{W}_t^{\mathbb{Q}^B}$$

We can write this in more detail:

$$\begin{aligned}\begin{pmatrix} dW_t^{\mathbb{Q}^C,B} \\ dW_t^{\mathbb{Q}^C,C} \end{pmatrix} &= \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \begin{pmatrix} \sigma_t^B \\ -\sigma_t^C \end{pmatrix} dt + \begin{pmatrix} dW_t^{\mathbb{Q}^B,B} \\ dW_t^{\mathbb{Q}^B,C} \end{pmatrix} \\ &= \begin{pmatrix} \sigma_t^B - \rho\sigma_t^C \\ \rho\sigma_t^B - \sigma_t^C \end{pmatrix} dt + \begin{pmatrix} dW_t^{\mathbb{Q}^B,B} \\ dW_t^{\mathbb{Q}^B,C} \end{pmatrix}\end{aligned}$$

And, so the Brownian motions are related as follows:

$$\begin{aligned} dW_t^{\mathbb{Q}^C,B} &= \left(\sigma_t^B - \rho\sigma_t^C\right) dt + dW_t^{\mathbb{Q}^B,B} \\ dW_t^{\mathbb{Q}^C,C} &= \left(\rho\sigma_t^B - \sigma_t^C\right) dt + dW_t^{\mathbb{Q}^B,C} \end{aligned}$$

We summarize the main result in the following theorem:

Theorem 12.1 (Numeraire changes). *Assume two assets, $B = (B_t)_{t\geq 0}$ and $C = (C_t)_{t\geq 0}$, satisfy the SDEs:*

$$\begin{aligned} \frac{dB_t}{B_t} &= \mu_t^B \; dt + \sigma_t^B \; dW_t^{\mathbb{Q}^B,B} \\ \frac{dC_t}{C_t} &= \mu_t^C \; dt + \sigma_t^C \; dW_t^{\mathbb{Q}^B,C} \end{aligned}$$

where $W^{\mathbb{Q}^B,B}$ and $W^{\mathbb{Q}^C,C}$ are correlated $\mathbb{Q}^B$-Brownian motions with instantaneous correlation ρ and $\mathbb{Q}^B$ is the EMM induced by B as a numeraire. Furthermore, let $\mathbb{Q}^C$ be the EMM induced by C as a numeraire. Then the Radon-Nikodym derivative process relating the two measures is given by:

$$\eta_t = \left(\frac{d\mathbb{Q}^C}{d\mathbb{Q}^B}\right)_t = \frac{C_t/C_0}{B_t/B_0}$$

which satisfies the SDE:

$$\frac{d\eta_t}{\eta_t} = -\sigma_t^B \; dW_t^{\mathbb{Q}^B,B} + \sigma_t^C \; dW_t^{\mathbb{Q}^B,C}$$

Then, by Girsanov's theorem we have that:

$$\begin{aligned} dW_t^{\mathbb{Q}^C,B} &= \left(\sigma_t^B - \rho\sigma_t^C\right) dt + dW_t^{\mathbb{Q}^B,B} \\ dW_t^{\mathbb{Q}^C,C} &= \left(\rho\sigma_t^B - \sigma_t^C\right) dt + dW_t^{\mathbb{Q}^B,C} \end{aligned}$$

are correlated $\mathbb{Q}^C$-Brownian motions with instantaneous correlation ρ. Finally, the $\mathbb{Q}^C$-dynamics of the assets are given by the SDEs:

$$\begin{aligned} \frac{dB_t}{B_t} &= \left(\mu_t^B - \sigma_t^B\left(\sigma_t^B - \rho\sigma_t^C\right)\right) dt + \sigma_t^B \; dW_t^{\mathbb{Q}^C,B} \\ \frac{dC_t}{C_t} &= \left(\mu_t^C - \sigma_t^C\left(\rho\sigma_t^B - \sigma_t^C\right)\right) dt + \sigma_t^C \; dW_t^{\mathbb{Q}^C,C} \end{aligned}$$

The most common numeraire change involves switching from the bank account to another traded asset as numeraire. This is because working with the bank account as numeraire is a natural starting point, since we know that all traded assets (including the new numeraire) must have a drift equal to the short rate of interest. The following corollary formalizes this, and is simply a special case of the theorem above where $\sigma_t^B = 0$.

Corollary 12.2 (Measure changes from $\mathbb{Q}$). *Assume $A = (A_t)_{t\geq 0}$ is a numeraire asset satisfying:*

$$\frac{dA_t}{A_t} = r_t\; dt + \sigma_t\; dW_t^{\mathbb{Q}}$$

where r_t is the short rate of interest and $\mathbb{Q}$ is the EMM induced by using the money market account, M, as the numeraire asset. Furthermore, let $B^{\mathbb{Q}}$ be a correlated $\mathbb{Q}$-Brownian motion with instantaneous correlation ρ. If $\mathbb{Q}^A$ is the EMM induced by using A as a numeraire then

$$\begin{aligned} dW_t^{\mathbb{Q}^A} &= -\sigma_t\; dt + dW_t^{\mathbb{Q}} \\ dB_t^{\mathbb{Q}^A} &= -\rho\sigma_t\; dt + dB_t^{\mathbb{Q}} \end{aligned}$$

are correlated $\mathbb{Q}^A$-Brownian motions with instantaneous correlation ρ, and the Radon-Nikodym process for the measure change is given by:

$$\eta_t = \left(\frac{d\mathbb{Q}^A}{d\mathbb{Q}}\right)_t = \frac{A_t/A_0}{M_t/M_0}$$

which satisfies the SDE:

$$\frac{d\eta_t}{\eta_t} = \sigma_t\; dW_t^{\mathbb{Q}}$$

12.3 Valuing the Exchange Option

Now, we can return to the original problem of valuing the exchange option. For simplicity, assume that the parameters are all constant, that is:

$$\begin{aligned} \frac{dA_t}{A_t} &= \mu_A\; dt + \sigma_A\; dW_t^{\mathbb{P},A} \\ \frac{dB_t}{B_t} &= \mu_B\; dt + \sigma_B\; dW_t^{\mathbb{P},B} \\ \frac{dM_t}{M_t} &= r\; dt \end{aligned}$$

where $W^{\mathbb{P},A}$ and $W^{\mathbb{P},B}$ are standard Brownian motions under the physical measure, $\mathbb{P}$, with instantaneous correlation ρ. Recall also that the payoff function for the exchange option is $G(A_T, B_T) = (A_T - B_T)_+$.

Show that the price of the exchange option, $g_t = g(t, A_t, B_t)$ is given by:

$$g_t = A_t \cdot \Phi(d_+) - B_t \cdot \Phi(d_-) \qquad \text{where} \quad d_\pm = \frac{\ln\left(\frac{A_t}{B_t}\right) \pm \frac{1}{2}\bar{\sigma}^2(T-t)}{\bar{\sigma}\sqrt{T-t}}$$
$$\bar{\sigma}^2 = \sigma_A^2 + \sigma_B^2 - 2\rho\sigma_A\sigma_B$$

Extra Practice: Solve for the option price in the case where the parameters are deterministic functions of time.

13 Stochastic Volatility

In this note we investigate some of the shortcomings of the Black-Scholes model by introducing the notion of implied volatility, and then we discuss how these shortcomings are overcome through the use of local and stochastic volatility models.

13.1 Implied Volatility

Recall that in the Black-Scholes model interest rates are assumed to be constant and asset prices are assumed to satisfy the following $\mathbb{P}$-dynamics:

$$\frac{dS_t}{S_t} = \mu\ dt + \sigma\ dW_t^{\mathbb{P}}$$

where $W^{\mathbb{P}}$ is a standard Brownian motion under the physical measure $\mathbb{P}$. Furthermore, the price of a claim written on S with payoff function $G(S_T)$ is equal to:

$$g_t \;=\; e^{-r(T-t)} \cdot \mathbb{E}_t^{\mathbb{Q}}\left[G(S_T)\right]$$

where $\mathbb{Q}$ is the risk-neutral measure and S satisfies the SDE:

$$\frac{dS_t}{S_t} \;=\; r\ dt + \sigma\ dW_t^{\mathbb{Q}}$$

In particular, we can also show that the prices of call and put options in this model are given by:

$$\begin{aligned} V_t^{call} &= S_t\ \Phi(d_+) - Ke^{-r(T-t)}\Phi(d_-) \\ V_t^{put} &= Ke^{-r(T-t)}\Phi(-d_-) - S_t\ \Phi(-d_+) \end{aligned}$$

where

$$d_\pm \;=\; \frac{\ln(S_t/K) + (r \pm \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}$$

A natural question to ask is: **how well does this model reflect reality**? One way of answering this question is to **compare model-implied option prices with option prices observed in the market**. It turns out that the model is not a good representation of reality nor does it capture market prices well. To see why this is the case we can look at the **implied volatility surface**.

Definition 13.1 (Implied volatility). Let $V^{BS}(K, T, S_0, r, \sigma)$ be the Black-Scholes model-implied option price for a given set of parameters, and let $V^*(K, T; S_0)$ be the market price of the same option, i.e. the market price of the option written on the same asset with the same strike and maturity. **Implied volatility**, σ_{imp}, is the value of σ that equates these two quantities, that is:

$$V^*(K, T; S_0) \;=\; V^{BS}(K, T, S_0, r, \sigma_{\text{imp}}(K, T))$$

Notice there is an implied volatility for every value of K and T, i.e. for every strike and maturity combination there is some value of σ which, when used as the input to the Black-Scholes option price, produces the market price of the same option. As such, implied volatility can be viewed as a function of strike and maturity. Varying K and T and plotting the resulting implied volatility value yields an **implied volatility surface**.

One shortcoming of the Black-Scholes model is that the volatility parameter, σ, is assumed to be constant, which implies that the **Black-Scholes implied volatility surface is flat**. In particular, **the model fails to capture the so-called volatility smile observed in option prices**. Moreover, the Black-Scholes model assumes that the $\mathbb{P}$-measure volatility of assets is constant, which is not an accurate representation of reality as **empirical evidence suggests that volatility is time-varying and has properties of mean reversion and clustering**.

There are two main approaches for addressing these issues:

(i) Local volatility models
(ii) Stochastic volatility models

13.2 Local Volatility Models

Local volatility models involve modifying the Black-Scholes model by **making volatility a function of time and spot price** rather than a constant:

$$\frac{dS_t}{S_t} = \mu \; dt + \sigma(t, S_t) \; dW_t^{\mathbb{P}}$$

$$\frac{dS_t}{S_t} = r \; dt + \sigma(t, S_t) \; dW_t^{\mathbb{Q}}$$

Typically the function is chosen such that $\sigma(t, S)$ *increases* as S decreases; this reflects the fact that assets prices become more volatile as they become more distressed and is known as the **leverage effect**. Additionally, the **generalized Black-Scholes PDE** still holds in this case:

$$\begin{cases} \partial_t g(t,S) + rS \cdot \partial_S g(t,S) + \frac{1}{2}\sigma^2(t,S)S^2 \cdot \partial_{SS} g(t,S) \;=\; r \cdot g(t,S) \\ \hfill g(T,S) \;=\; G(S) \end{cases}$$

Local volatility models involve choosing the function $\sigma(t, S)$ such that the model-implied option prices match the observed market option prices. In fact, the main result surrounding local volatility models is:

THEOREM 13.2. *For any implied volatility surface $\sigma^*_{imp}(K, T)$ there exists a local volatility function, $\sigma_{loc}(t, S)$, such that the implied volatility surface given by the corresponding local volatility model matches the market-observed implied volatility surface, i.e.*

$$\forall \; \sigma^*_{imp}(K,T) \quad \exists \; \sigma_{loc}(t,S) \quad \text{such that} \quad \sigma^{loc}_{imp}(K,T) = \sigma^*_{imp}(K,T)$$

*This process involves the use of **Dupire's formula**.*

One issue with local volatility models is that they treat volatility as static; once the local volatility function is set the only source of stochasticity is derived from the evolution of the spot price S_t. In other words, there is no inherent stochasticity in the volatility process. This can be remedied through the use of stochastic volatility models such as the Heston model.

13.3 Heston Model

13.3.1 Model Setup

The **Heston model**, which belongs to the class of **stochastic volatility** models, is a generalization of the Black-Scholes model where the variance of asset returns, $\nu = (\nu_t)_{t \geq 0}$, is itself a stochastic process known as the **variance process**; $\sqrt{\nu_t}$ is referred to as the **volatility process**. In particular, the variance process is modeled using a **Feller process**. That is, the asset price and variance process satisfy the following $\mathbb{P}$-dynamics:

$$\begin{aligned}\frac{dS_t}{S_t} &= \mu\ dt + \sqrt{\nu_t}\ dW_t^{\mathbb{P}} \\ d\nu_t &= \kappa^{\mathbb{P}}\left(\theta^{\mathbb{P}} - \nu_t\right)\ dt + \eta\sqrt{\nu_t}\ dB_t^{\mathbb{P}}\end{aligned}$$

where $W^{\mathbb{P}}$ and $B^{\mathbb{P}}$ are correlated $\mathbb{P}$-Brownian motions with $d\left[W_t^{\mathbb{P}}, B_t^{\mathbb{P}}\right]_t = \rho\ dt$. The parameter η is referred to as the **vol-vol**. The **variance process remains positive** if the Feller condition $\left(2\kappa^{\mathbb{P}}\theta^{\mathbb{P}} > \eta^2\right)$ is met. Additionally, **the Feller process ensures that variance is mean reverting** with mean-reversion level and rate $\theta^{\mathbb{P}}$ and $\kappa^{\mathbb{P}}$, respectively.

13.3.2 Risk-Neutral Dynamics

The model above gives the $\mathbb{P}$-dynamics of the asset price and its variance. But **in order to price claims we require the $\mathbb{Q}$-dynamics of these processes**. In other words, we require drifted versions of $W_t^{\mathbb{P}}$ and $B_t^{\mathbb{P}}$:

$$\begin{aligned}dW_t^{\mathbb{Q}} &= \lambda_t^S\ dt + dW_t^{\mathbb{P}} \\ dB_t^{\mathbb{Q}} &= \lambda_t^\nu\ dt + dB_t^{\mathbb{P}}\end{aligned}$$

What are the correct drift corrections that we must choose for $W_t^{\mathbb{P}}$ and $B_t^{\mathbb{P}}$?

- λ^S must be chosen to ensure that the drift of the asset under the measure $\mathbb{Q}$ is equal to the risk-free rate r. Since the $\mathbb{Q}$-dynamics are given by:

  $$\frac{dS_t}{S_t} = \left(\mu - \lambda_t^S\sqrt{\nu_t}\right)\ dt + \sqrt{\nu_t}\ dW_t^{\mathbb{Q}}$$

  Setting the drift term equal to r, it follows that the required drift correction must be

  $$\lambda_t^S = \frac{\mu - r}{\sqrt{\nu_t}}$$

- **Since ν is not a traded asset, there are no additional no-arbitrage requirements that need to be satisfied**. Thus, λ^ν can be chosen freely. Given a choice of λ^ν the $\mathbb{Q}$-dynamics of the variance process will be:

  $$d\nu_t = \left(\kappa^{\mathbb{P}}\left(\theta^{\mathbb{P}} - \nu_t\right) - \lambda_t^\nu\eta\sqrt{\nu_t}\right)\ dt + \eta\sqrt{\nu_t}\ dB_t^{\mathbb{Q}}$$

  One approach is to **choose λ^ν such that the functional form of the process is maintained**. That is, the $\mathbb{Q}$-dynamics of the variance process still follow a Feller process. For example, we can choose:

$$(1)\quad \lambda_t^\nu = a\sqrt{\nu_t} \qquad \Longrightarrow \qquad d\nu_t = \kappa^{\mathbb{Q}}\left(\theta^{\mathbb{Q}} - \nu_t\right)\,dt + \eta\sqrt{\nu_t}\,dB_t^{\mathbb{Q}}$$

$$\kappa^{\mathbb{Q}} = \kappa^{\mathbb{P}} + a\eta, \qquad \theta^{\mathbb{Q}} = \frac{\kappa^{\mathbb{P}}\theta^{\mathbb{P}}}{\kappa^{\mathbb{P}} + a\eta}$$

$$(2)\quad \lambda_t^\nu = a\sqrt{\nu_t} + \frac{b}{\sqrt{\nu_t}} \qquad \Longrightarrow \qquad d\nu_t = \kappa^{\mathbb{Q}}\left(\theta^{\mathbb{Q}} - \nu_t\right)\,dt + \eta\sqrt{\nu_t}\,dB_t^{\mathbb{Q}}$$

$$\kappa^{\mathbb{Q}} = \kappa^{\mathbb{P}} + a\eta, \qquad \theta^{\mathbb{Q}} = \frac{\kappa^{\mathbb{P}}\theta^{\mathbb{P}} - b\eta}{\kappa^{\mathbb{P}} + a\eta}$$

Proceeding with either option for λ^ν, the $\mathbb{Q}$-dynamics of the processes are given by:

$$\frac{dS_t}{S_t} = r\,dt + \sqrt{\nu_t}\,dW_t$$
$$d\nu_t = \kappa\left(\theta - \nu_t\right)\,dt + \eta\sqrt{\nu_t}\,dB_t$$

where $d\left[W, B\right]_t = \rho\,dt$ and the $\mathbb{Q}$ superscripts are dropped for brevity.

A few notes on this model:

- The parameters $\mu, \kappa^{\mathbb{P}}, \theta^{\mathbb{P}}, \eta$ describe the **real-world dynamics of the asset price and variance processes**, and are estimated using historical data.
- The parameters a and b are degrees of freedom that **allow us to calibrate the model to observed market prices**. That is, every choice of a and b will imply a different set of option prices. Calibration involves finding the a and b values that minimize the difference between market prices and model-implied prices.
- Once a and b are determined the parameters $\kappa^{\mathbb{Q}}, \theta^{\mathbb{Q}}$ can be computed. These values represent the **market-implied mean reversion level and rate**, i.e. a forward-looking measure of mean reversion level and rate (as opposed to the backward-looking historical values).
- Using alternative drift adjustments leads to valid no-arbitrage prices, but prices will be different based on the choice of λ^ν. This is an example of an **incomplete market**.

13.3.3 Derivative Valuation under the Heston Model

To price derivatives under the Heston model, we can proceed in the usual manner, via an expectation or a PDE-based approach. Let $g = (g_t)_{t\geq 0}$ be the price of a claim written on S and ν with payoff function $G(S_T, \nu_T)$ and assume that the process is Markovian in S and ν, i.e. there exists a function g such that $g_t = g(t, S_t, \nu_t)$. Then g_t is given by:

$$g_t = \mathbb{E}_t^{\mathbb{Q}}\left[G(S_T, \nu_T)\cdot e^{-r(T-t)}\right]$$

where the $\mathbb{Q}$-dynamics of S and ν are given by the SDEs:

$$\frac{dS_t}{S_t} = r\ dt + \sqrt{\nu_t}\ dW_t$$
$$d\nu_t = \kappa\left(\theta - \nu_t\right)\ dt + \eta\sqrt{\nu_t}\ dB_t$$

Additionally, by the Feynman-Kac theorem, we know that g must satisfy the PDE:

$$\begin{cases} (\partial_t + \mathcal{L}^{S,\nu})g \ = \ rg \\ \quad g(T,S,\nu) = G(S,\nu) \end{cases}$$

where $\mathcal{L}^{S,\nu} = rS \cdot \partial_S + \frac{1}{2}\nu S^2 \cdot \partial_{SS} + \kappa(\theta - \nu) \cdot \partial_\nu + \frac{1}{2}\eta^2\nu \cdot \partial_{\nu\nu} + \rho\eta\nu S \cdot \partial_{S\nu}$ is the infinitesimal generator of the process.

Now, **we can linearize the PDE above by reparameterizing the problem in terms of the log-asset price**, $X_t = \log S_t$. That is, by substituting we have:

$$g_t = \mathbb{E}_t^{\mathbb{Q}}\left[G\left(e^{X_T}, \nu_T\right) \cdot e^{-r(T-t)}\right]$$

where the $\mathbb{Q}$-dynamics of X and ν, using Itô's lemma, are given by the SDEs:

$$dX_t = \left(r - \frac{1}{2}\nu_t\right)dt + \sqrt{\nu_t}\ dW_t$$
$$d\nu_t = \kappa\left(\theta - \nu_t\right)\ dt + \eta\sqrt{\nu_t}\ dB_t$$

The corresponding PDE in terms of the state variable $x = \log S$ is then:

$$\begin{cases} (\partial_t + \mathcal{L}^{X,\nu})g \ = \ rg \\ \quad g\left(T,x,\nu\right) = G\left(e^x,\nu\right) \end{cases}$$

where $\mathcal{L}^{X,\nu} = \left(r - \frac{1}{2}\nu\right) \cdot \partial_x + \frac{1}{2}\nu \cdot \partial_{xx} + \kappa(\theta - \nu) \cdot \partial_\nu + \frac{1}{2}\eta^2\nu \cdot \partial_{\nu\nu} + \rho\eta\nu \cdot \partial_{x\nu}$ is the infinitesimal generator of the process. The advantage of formulating the problem in this manner is that it leads to an **affine PDE**, i.e. the PDE has coefficients that are at most linear in the state variables. These models are typically easier to work with and allow us to arrive at closed-form solutions.

One way these affine models are used is in **computing the $\mathbb{Q}$-characteristic function of the log-asset price**. To understand the usefulness of this, notice that if a payoff function can be written in terms of its inverse Fourier transform:

$$G(x) = \int_{-\infty}^{\infty} e^{i\omega x} \cdot \hat{G}(\omega)\ d\omega$$

then being able to price a claim that pays $e^{i\omega X_T}$ is equivalent to being able to price a claim with payoff function $G(X_T)$. Now, the price of the former is equal to:

$$h(t,x,\nu) = \mathbb{E}_t^{\mathbb{Q}}\left[e^{i\omega X_T}\right]$$

which is the characteristic function of the log-asset price at maturity, X_T, under the risk-neutral measure $\mathbb{Q}$. By the Feynman-Kac theorem, h must satisfy the PDE:

$$\begin{cases} (\partial_t + \mathcal{L}^{X,\nu})h \;=\; 0 \\ \qquad h(T,x,\nu) = e^{i\omega x} \end{cases}$$

Since the PDE is affine and the boundary condition is exponential affine, the solution is of the form:

$$h(t,x,\nu) = e^{A(t)+B(t)\nu+C(t)x}$$

Using the PDE and ansatz above, find the ODE system that the functions A, B and C must satisfy (including the appropriate boundary conditions).

Hint: differentiate the given form of h and substitute into the PDE and note that the coefficients of all state variables must be identically zero since the equation must hold for all values of x and ν.

13.4 Variance Swaps

Variance swaps are financial derivatives that are used to speculate on and hedge risks associated with an asset's volatility. They depend on the **realized variance** of a process:

DEFINITION 13.3 (Realized variance). Let $S = (S_t)_{t\geq 0}$ be a process and let $\{t_k = k\Delta t\}_{k=0,\dots,n}$ with $\Delta t = T/n$ denote a set of discrete equally spaced times. The **realized variance** of S over $[0,T]$ is defined as:

$$R_T = \sum_{k=1}^{n}\left(\frac{S_{t_k}}{S_{t_{k-1}}} - 1\right)^2$$

A **variance swap with strike** K is a claim that pays $R_T - K$ at maturity T. Notice that this is a path-dependent claim due to the way R_T is defined. The **value of the claim** is:

$$V_0 \;=\; e^{-rT}\cdot \mathbb{E}^{\mathbb{Q}}\left[R_T - K\right] \;=\; e^{-rT}\cdot \left(\mathbb{E}^{\mathbb{Q}}\left[R_T\right] - K\right)$$

The **fair variance swap strike** is the value of K that sets the value of the variance swap equal to zero. It follows from the expression above that this quantity is equal to:

$$K = \mathbb{E}^{\mathbb{Q}}\left[R_T\right]$$

Assume the asset follows a stochastic volatility model of the form:

$$\frac{dS_t}{S_t} = r_t\; dt + \sqrt{\nu_t}\; dW_t$$

where W is a $\mathbb{Q}$-Brownian motion and $\nu = (\nu_t)_{t\geq 0}$ is a stochastic process. **We can approximate realized variance** as follows:

$$\begin{aligned} R_T \;&=\; \sum_{k=1}^{n}\left(\frac{S_{t_k} - S_{t_{k-1}}}{S_{t_{k-1}}}\right)^2 \\ &\approx\; \sum_{k=1}^{n}\left(\frac{dS_t}{S_t}\right)^2 \\ &\approx\; \sum_{k=1}^{n}\nu_{t_{k-1}}\Delta t \\ &\underset{n\to\infty}{\longrightarrow}\; \int_0^T \nu_t\; dt \end{aligned}$$

In particular, the $\mathbb{Q}$-expectation of R_T, which also gives the fair variance swap strike is approximated by:

$$\mathbb{E}^{\mathbb{Q}}\left[R_T\right] \;=\; \int_0^T \mathbb{E}^{\mathbb{Q}}[\nu_t]\; dt$$

Recall that **in the Heston model** the variance process follows a Feller process:

$$d\nu_t = \kappa(\theta - \nu_t)\; dt + \eta\sqrt{\nu_t}\; dB_t$$

where B is a $\mathbb{Q}$-Brownian motion correlated with W with correlation ρ. As such we can solve for the expected value of the variance process:

$$\mathbb{E}^{\mathbb{Q}}[\nu_t] = \nu_0 \cdot e^{-\kappa t} + \theta\left(1 - e^{-\kappa t}\right)$$

Which leads to a closed-form solution of the fair variance swap strike:

$$\mathbb{E}^{\mathbb{Q}}\left[R_T\right] = \int_0^T \nu_0 \cdot e^{-\kappa t} + \theta\left(1 - e^{-\kappa t}\right)\, dt$$

More generally, we can solve for the fair variance swap strike by expressing the claim as a collection of calls and puts, which is valid regardless of the underlying stochastic volatility model. Assume for simplicity that interest rates are equal to zero, $r_t = 0$ (the analysis for non-zero interest rates is similar). Then we have:

$$\begin{aligned}
& \frac{dS_t}{S_t} = \sqrt{\nu_t}\, dW_t \\
\implies\ & d\ln S_t = -\frac{1}{2}\nu_t\, dt + \sqrt{\nu_t}\, dW_t \\
\implies\ & \ln S_T - \ln S_0 = -\frac{1}{2}\int_0^T \nu_t\, dt + \int_0^T \sqrt{\nu_t}\, dW_t \\
\implies\ & \mathbb{E}^{\mathbb{Q}}\left[\int_0^T \nu_t\, dt\right] = -2\,\mathbb{E}^{\mathbb{Q}}\left[\ln\left(\frac{S_T}{S_0}\right)\right]
\end{aligned}$$

In other words, **we have written the fair strike** (LHS of the equation) **as the price of a claim that pays** $\ln(S_T/S_0)$ (RHS of the equation). Next, we focus on the term on the RHS and we use the fact (deferring the proof to the end) that the value of any claim with payoff function $g(S)$ where $g \in C^2$ can be written in terms of a sum of puts and calls:

$$V_0^g = g(S^*) + (S_0 - S^*)g'(S^*) + \int_0^{S^*} V^{put}(S_0, K)g''(K)\, dK + \int_{S^*}^{\infty} V^{call}(S_0, K)g''(K)\, dK$$

where S_0 is the current spot price and $S^* < S_0$. The formula above says that the price of the claim with payoff function g is equal to the sum of the values of:

(i) A bond paying $g(S^*)$
(ii) $g'(S^*)$ units of forward contracts with strike S^*
(iii) $g''(K)$ units of puts with strike K for strikes between 0 and S^*
(iv) $g''(K)$ units of calls with strike K for strikes between S^* and infinity.

Applying this formula to the payoff function $\ln(S_T/S_0)$ we have:

$$\begin{aligned}
\mathbb{E}^{\mathbb{Q}}\left[\ln\left(\frac{S_T}{S_0}\right)\right] = \ & \ln\left(\frac{S^*}{S_0}\right) + \frac{S^* - S_0}{S^*} + \int_0^{S^*} V^{put}(S_0, K)\left(-\frac{1}{K^2}\right)\, dK \\
& + \int_{S^*}^{\infty} V^{call}(S_0, K)\left(-\frac{1}{K^2}\right)\, dK
\end{aligned}$$

In practice, S^* is taken to be the largest available strike below the current spot, that is $S^* = \max\{K_i : K_i < S_0\}$. The integrals are discretized using the available strikes in the market:

$$\int_0^{S^*} V^{put}(S_0, K)\left(-\frac{1}{K^2}\right) dK \approx -\sum_{i:K_i<S_0} \frac{V^{put}(S_0,K_i)}{K_i^2}\cdot \Delta K_i$$
$$\int_{S^*}^{\infty} V^{call}(S_0, K)\left(-\frac{1}{K^2}\right) dK \approx -\sum_{i:K_i\geq S_0} \frac{V^{put}(S_0,K_i)}{K_i^2}\cdot \Delta K_i$$

Futhermore, we can approximate the first two terms in expression as follows:

$$\ln\left(\frac{S^*}{S_0}\right) + \frac{S^*-S_0}{S^*} = -\ln\left(\frac{S_0}{S^*}\right) + \left(\frac{S_0}{S^*}-1\right)$$
$$\approx \frac{1}{2}\left(\frac{S_0}{S^*}-1\right)^2$$

The last step follows from using the Taylor expansion $\ln(1+x) \approx x - \frac{1}{2}x^2$:

$$\ln\left(\frac{S_0}{S^*}\right) = \ln\left(1+\left(\frac{S_0}{S^*}-1\right)\right) \approx \left(\frac{S_0}{S^*}-1\right) - \frac{1}{2}\left(\frac{S_0}{S^*}-1\right)^2$$

Combining all of the terms above we can write the fair variance swap strike as:

$$\mathbb{E}^{\mathbb{Q}}\left[\int_0^T \nu_s\, ds\right] = 2\left(\sum_i \frac{V^{otm}(S_0,K_i)}{K_i^2}\cdot \Delta K_i - \frac{1}{2}\left(\frac{S_0}{S^*}-1\right)^2\right)$$

where V^{otm} refers to the call or put option that is out-of-the-money at the given strike.

Now, we return to the deferred proof of writing a claim in terms of a sum of puts and calls. The payoff functions of a put and call option along with their derivatives with respect to strike prices are given by:

$$\phi_P = (K-S)_+ \qquad \phi_C = (S-K)_+$$
$$\partial_K \phi_P = \mathbb{1}_{\{K>S\}} \qquad \partial_K \phi_C = 1 - \mathbb{1}_{\{K>S\}}$$
$$\partial_{KK} \phi_P = \delta(S-K) \qquad \partial_{KK} \phi_C = \delta(S-K)$$

where $\delta(x)$ is the Dirac delta function, which can be loosely thought of as “equal” to infinity at $x=0$ and 0 elsewhere. The main property of delta functions is that we can use them to “extract” a function's value at $x=0$:

$$\int_{-\infty}^{\infty} g(x)\delta(x)\, dx = g(0)$$

If $g \in C^2$ is a payoff function for a given claim, we can write this as:

$$g(S) = \int_0^{\infty} g(K)\delta(S-K)\, dK$$

Writing the delta function using the call option payoff derivatives given above we have:

$$g(S) = \int_0^{S^*} g(K) \cdot \partial_{KK}\phi_P(K)\ dK + \int_{S^*}^{\infty} g(K) \cdot \partial_{KK}\phi_C(K)\ dK$$

where $S^* < S$ is a fixed value. The rationale for choosing such a value is related to simplifications that occur in the following steps. Next, we perform an integration by parts on each integral twice:

$$\begin{aligned}
\int_0^{S^*} g(K) \cdot \partial_{KK}\phi_P(K)\ dK &= -\int_0^{S^*} g'(K) \cdot \partial_K\phi_P(K)\ dK + \underbrace{g(K) \cdot \partial_K\phi_P(K)\Big|_0^{S^*}}_{=\,0,\text{ by choice of } S^*} \\
&= \int_0^{S^*} g''(K) \cdot \phi_P(K)\ dK + \underbrace{g'(K) \cdot \phi_P(K)\Big|_0^{S^*}}_{=\,0} \\
\int_{S^*}^{\infty} g(K) \cdot \partial_{KK}\phi_C(K)\ dK &= -\int_{S^*}^{\infty} g'(K) \cdot \partial_K\phi_C(K)\ dK + \underbrace{g(K) \cdot \partial_K\phi_C(K)\Big|_{S^*}^{\infty}}_{=\,g(S^*)} \\
&= g(S^*) + \int_{S^*}^{\infty} g''(K) \cdot \phi_C(K)\ dK + \underbrace{g'(K) \cdot \phi_C(K)\Big|_{S^*}^{\infty}}_{=\,(S-S^*)g'(S^*)}
\end{aligned}$$

Combining these terms we have:

$$g(S) \;=\; g(S^*) + (S - S^*)g'(S^*) + \int_0^{S^*} \phi_P(K)g''(K)\ dK + \int_{S^*}^{\infty} \phi_C(K)g''(K)\ dK$$

Assuming interest rates are zero, taking risk-neutral expectations on both sides and noting that $S_0 = \mathbb{E}^{\mathbb{Q}}[S_T]$ we arrive at the desired result:

$$V_0^g \;=\; g(S^*) + (S_0 - S^*)g'(S^*) + \int_0^{S^*} V^{put}(S_0, K)g''(K)\ dK + \int_{S^*}^{\infty} V^{call}(S_0, K)g''(K)\ dK$$

14 Interest Rate Derivatives

14.1 Introduction

In the dynamic hedging argument claims were written on a source of uncertainty X_t. Naturally, this source of uncertainty can include interest rate processes, r_t. Such claims are referred to as **interest rate derivatives**.

Assume that the short rate process, $r = (r_t)_{t\geq 0}$, satisfies the SDE:

$$dr_t = \mu^r(t, r_t)\ dt + \sigma^r(t, r_t)\ dW_t^{\mathbb{P}}$$

where $W^{\mathbb{P}}$ is a standard Brownian motion under the physical measure $\mathbb{P}$ and μ^r and σ^r are the drift and volatility functions of the short rate process. Furthermore, assume that the price process of a claim written on r with payoff function $G(r)$, $g = (g_t)_{t\geq 0}$, is Markovian in r, i.e. there exists a function g such that $g_t = g(t, r_t)$. Then by the dynamic hedging argument, **the function g satisfies the PDE**:

$$\begin{cases} \partial_t g(t,r) + (\mu^r(t,r) - \lambda(t,r)\sigma^r(t,r)) \cdot \partial_r g(t,r) + \frac{1}{2}(\sigma^r(t,r))^2 \cdot \partial_{rr} g(t,r) = r \cdot g(t,r) \\ \hfill g(T,r) = G(r) \end{cases}$$

As usual, this admits the following **stochastic representation**:

$$\begin{aligned} g(t,r) &= \mathbb{E}_t^{\mathbb{Q}}\left[e^{-\int_t^T r_s ds} \cdot G(r_T)\right] \\ \text{where} \qquad dr_t &= (\mu^r(t,r_t) - \lambda(t,r_t)\sigma^r(t,r_t))\ dt + \sigma^r(t,r)\ dW_t^{\mathbb{Q}} \end{aligned}$$

and $W^{\mathbb{Q}}$ is a standard Brownian motion under the risk-neutral measure $\mathbb{Q}$.

A few notes on this:

- Given a specific interest rate model, i.e. specific μ^r and σ^r, the price of the claim can be obtained by solving the PDE or computing the expectation.
- Since r_t is not a traded asset there are no restrictions on the choice of the drift adjustment λ. Typically, it is chosen to maintain the functional form of the short rate model.
- The expectation approach requires the joint distribution of r_T and $\int_t^T r_s\ ds$.

14.2 Interest Rate Models

14.2.1 Vasicek Model

In the Vasicek model, the short rate is modeled as an Ornstein-Uhlenbeck process:

$$dr_t = \kappa(\theta - r_t)\ dt + \sigma\ dW_t^{\mathbb{P}}$$

Recall that this implies that **the short rate process is mean-reverting** with mean reversion rate and level given by κ and θ, respectively. **The solution to the Vasicek model and the distribution of** r_T are given by:

$$\begin{aligned} r_T &= r_t e^{-\kappa(T-t)} + \theta\left(1 - e^{-\kappa(T-t)}\right) + \sigma \int_t^T e^{-\kappa(T-s)}\ dW_s^{\mathbb{P}} \\ r_T &\sim N\left(r_t e^{-\kappa(T-t)} + \theta\left(1 - e^{-\kappa(T-t)}\right),\ \frac{\sigma^2}{2\kappa}\left(1 - e^{-2\kappa(T-t)}\right)\right) \end{aligned}$$

Moreover, integrating the SDE directly allows us to obtain **an expression for the integral** $\int_t^T r_s ds$ **and find its distribution**:

$$\begin{aligned} \int_t^T r_s\ ds &= \theta(T-t) - \frac{1 - e^{-\kappa(T-t)}}{\kappa}(\theta - r_t) + \frac{\sigma}{\kappa}\int_t^T 1 - e^{-\kappa(T-s)}\ dW_s^{\mathbb{P}} \\ \int_t^T r_s\ ds &\sim N\left(\theta(T-t) - \frac{1 - e^{-\kappa(T-t)}}{\kappa}(\theta - r_t),\ \frac{\sigma^2}{\kappa^2}\int_t^T \left(1 - e^{-\kappa(T-s)}\right)^2 ds\right) \end{aligned}$$

Note that we can also show that r_t and $\int_t^T r_s\ ds$ are jointly normally distributed.

As mentioned above, the drift adjustment is usually chosen to maintain the functional form of the model. In the case of the Vasicek model, assuming $\mathbb{P}$-dynamics are given by:

$$dr_t = \kappa^{\mathbb{P}}\left(\theta^{\mathbb{P}} - r_t\right)\ dt + \sigma\ dW_t^{\mathbb{P}}$$

the $\mathbb{Q}$-dynamics of the short rate process process will be:

$$dr_t = \left(\kappa^{\mathbb{P}}\left(\theta^{\mathbb{P}} - r_t\right) - \lambda_t\sigma\right)\ dt + \sigma\ dW_t^{\mathbb{Q}}$$

We can then **choose λ such that the functional form of the process is maintained**. That is, the $\mathbb{Q}$-dynamics of the short rate process still follow an Ornstein-Uhlenbeck process. For example, we can choose:

$$(1)\quad \lambda_t = a \implies dr_t = \kappa^{\mathbb{Q}}\left(\theta^{\mathbb{Q}} - r_t\right)\ dt + \sigma\ dW_t^{\mathbb{Q}}, \quad \kappa^{\mathbb{Q}} = \kappa^{\mathbb{P}}, \quad \theta^{\mathbb{Q}} = \theta^{\mathbb{P}} - \frac{a\sigma}{\kappa^{\mathbb{P}}}$$

$$(2)\quad \lambda_t = a + br_t \implies dr_t = \kappa^{\mathbb{Q}}\left(\theta^{\mathbb{Q}} - r_t\right)\ dt + \sigma\ dW_t^{\mathbb{Q}}, \quad \kappa^{\mathbb{Q}} = \kappa^{\mathbb{P}} + b\sigma, \quad \theta^{\mathbb{Q}} = \theta^{\mathbb{P}} - \frac{a\sigma}{\kappa^{\mathbb{P}} + b\sigma}$$

A few notes on this model:

- The parameters $\kappa^{\mathbb{P}}, \theta^{\mathbb{P}}, \sigma$ describe the **real-world dynamics of the asset price and variance processes**, and are estimated using historical data.
- The parameters a and b are degrees of freedom that **allow us to calibrate the model to observed market prices**. That is, every choice of a and b will imply a different set of derivative prices. Calibration involves finding the a and b values that minimize the difference between market prices and model-implied prices.
- Once a and b are determined the parameters $\kappa^{\mathbb{Q}}, \theta^{\mathbb{Q}}$ can be computed. These values represent the **market-implied mean reversion level and rate**, i.e. a forward-looking measure of mean reversion level and rate (as opposed to the backward-looking historical values).
- Derivative valuation is done using the $\mathbb{Q}$-dynamics of r_t.

14.2.2 Cox-Ingersoll-Ross Model

In the Cox-Ingersoll-Ross model, the short rate is modeled as a Feller process:

$$dr_t = \kappa(\theta - r_t)\ dt + \sigma\sqrt{r_t}\ dW_t^{\mathbb{P}}$$

Like the Vasicek model, **rates in the Cox-Ingersoll-Ross model are mean-reverting** with mean reversion rate and level given by κ and θ, respectively. However, unlike the Vasicek model, **rates in the Cox-Ingersoll-Ross model remain positive** almost surely provided that the Feller condition ($2\kappa\theta > \sigma^2$) is met.

Once again, the drift adjustment in this case is usually chosen to maintain the functional form of the model. Assuming $\mathbb{P}$-dynamics are given by:

$$dr_t = \kappa^{\mathbb{P}}\left(\theta^{\mathbb{P}} - r_t\right)\ dt + \sigma\sqrt{r_t}\ dW_t^{\mathbb{P}}$$

the $\mathbb{Q}$-dynamics of the short rate process process will be:

$$dr_t = \left(\kappa^{\mathbb{P}}\left(\theta^{\mathbb{P}} - r_t\right) - \lambda_t\sigma\sqrt{r_t}\right)\ dt + \sigma\sqrt{r_t}\ dW_t^{\mathbb{Q}}$$

Then we can choose λ to be:

$$(1)\quad \lambda_t = a\sqrt{r_t} \quad\Longrightarrow\quad dr_t = \kappa^{\mathbb{Q}}\left(\theta^{\mathbb{Q}} - r_t\right)\ dt + \eta\sqrt{r_t}\ dW_t^{\mathbb{Q}}$$

$$\kappa^{\mathbb{Q}} = \kappa^{\mathbb{P}} + a\sigma, \qquad \theta^{\mathbb{Q}} = \frac{\kappa^{\mathbb{P}}\theta^{\mathbb{P}}}{\kappa^{\mathbb{P}} + a\sigma}$$

$$(2)\quad \lambda_t = a\sqrt{r_t} + \frac{b}{\sqrt{r_t}} \quad\Longrightarrow\quad dr_t = \kappa^{\mathbb{Q}}\left(\theta^{\mathbb{Q}} - r_t\right)\ dt + \eta\sqrt{r_t}\ dW_t^{\mathbb{Q}}$$

$$\kappa^{\mathbb{Q}} = \kappa^{\mathbb{P}} + a\sigma, \qquad \theta^{\mathbb{Q}} = \frac{\kappa^{\mathbb{P}}\theta^{\mathbb{P}} - b\sigma}{\kappa^{\mathbb{P}} + a\sigma}$$

The interpretation of the $\mathbb{P}$ and $\mathbb{Q}$ parameters is the same as in the Vasicek model.

14.3 Zero-Coupon Bonds

The simplest interest rate derivative is a **T-maturity zero-coupon bond**. This is an instrument that pays 1 when it matures at time T. We can compute the price of such bonds using the fundamental theorem of finance, with the money market account as the numeraire. **The price at time t of a zero-coupon maturing at time T** is given by:

$$P_t(T) = \mathbb{E}_t^{\mathbb{Q}}\left[e^{-\int_t^T r_s ds}\right]$$

When interest rates are deterministic the expectation operator is not required; when interest rates are constant then the price of the bond is given by $P_t(T) = e^{-r(T-t)}$, which is the usual discount factor we have used. However, when rates are stochastic the expectation above depends on the model we use for interest rates. Moreover, the price of equity derivatives must also be adjusted to incorporate the uncertainty in interest rates when rates are assumed to be stochastic.

Assuming the Vasicek model for interest rates, the price of the T-maturity zero-coupon bond is given by:

$$P_t(T) = \exp\left(A_t(T) - B_t(T) \cdot r_t\right)$$

where

$$\begin{aligned} A_t(T) &= \left(\frac{1 - e^{-\kappa(T-t)}}{\kappa} - (T-t)\right)\theta + \frac{\sigma^2}{2\kappa^2}\int_t^T \left(1 - e^{-\kappa(T-s)}\right)^2 ds \\ B_t(T) &= \frac{1 - e^{-\kappa(T-t)}}{\kappa} \end{aligned}$$

Verify the zero-coupon bond price given above. Hint: recall that the integral that appears in the expectation is normally distributed under the Vasicek model.

<u>Note:</u> if the short rate model is affine, i.e. the coefficients of the generator are at most linear in r, then bond prices will be exponentially affine in r, i.e. of the form $P_t(T) = \exp\left(A_t(T) - B_t(T) \cdot r_t\right)$. This is true of the Vasicek and CIR models.

An alternative way of arriving at this solution is to solve the PDE satisfied by the zero-coupon bond. Since the bond is an interest rate derivative, it satisfies the PDE given above (where μ^r and σ^r are given by the Vasicek model in our case) and the appropriate payoff function:

$$\begin{cases} \partial_t g + \kappa(\theta - r) \cdot \partial_r g + \frac{1}{2}\sigma^2 \cdot \partial_{rr} g = rg \\ g(T, r) = 1 \end{cases}$$

Since the PDE is affine with an exponentially affine terminal condition, the solution to this PDE is of the form:

$$g(t, r) = \exp(A_t(T) - B_t(T)r)$$

Show that the price given by solving the PDE above is the same as the one found using the expectation approach. Hint: use the ansatz above to derive an ODE system that A and B satisfy (along with appropriate terminal conditions) and solve.

Using the PDE approach, show that bond prices under the CIR model are of the form $P_t(T) = \exp\left(A_t(T) - B_t(T) \cdot r_t\right)$ and characterize the functions A and B in terms of a system of ODEs.

Since bonds are traded assets they can be used as numeraires and in some cases this leads to considerable simplifications. **The EMM induced by a T-maturity zero-coupon bond is referred to as the T-forward neutral measure**, denoted by $\mathbb{Q}^T$. This then raises the question of the drift correction that links $\mathbb{Q}$ to $\mathbb{Q}^T$.

Assuming the Vasicek model, the dynamics of the zero-coupon bond under the risk neutral measure $\mathbb{Q}$ are given by:

$$\frac{dP_t(T)}{P_t(T)} = r_t\ dt - \sigma B_t(T)\ dW_t^{\mathbb{Q}}$$

The drift follows from the fact that all traded assests must have a drift equal to the risk-free rate under risk-neutral measure. The volatility follows by applying Itô's lemma to the bond price $P_t(T) = \exp\left(A_t(T) - B_t(T) \cdot r_t\right)$.

To use the T-maturity bond as a numeraire, we define the Radon-Nikodym derivative in the usual manner:

$$\left(\frac{d\mathbb{Q}^T}{d\mathbb{Q}}\right)_t = \frac{P_t(T)/P_0(T)}{M_t/M_0}$$

The drift correction associated with linking Brownian motions under the two measures is equal to the negative of the desired numariare's volatility (in this case, the volatility is $-\sigma B_t(T)$):

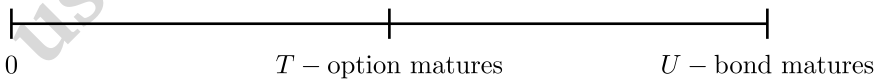
$$dW_t^{\mathbb{Q}} = \sigma B_t(T)\ dt + dW_t^{\mathbb{Q}}$$

Substituting into the bond price dynamics under $\mathbb{Q}$ we find **the dynamics of the bond price under** $\mathbb{Q}^T$:

$$\frac{dP_t(T)}{P_t(T)} = \left(r_t + \sigma^2 B_t(T)^2\right)\ dt - \sigma B_t(T)\ dW_t^{\mathbb{Q}^T}$$

14.4 Bond Options

One application of the techniques outlined above is to **compute the price of a bond option**. This derivative gives the holder the option to purchase a U-maturity bond at time T for the strike price K according to the plot below:



The **payoff function of the bond option** is then given by: $\varphi = (P_T(U) - K)_+$. Notice that we are comparing the price of the bond at time T, however the bond matures at time U. We can compute the price of this option as the expected discounted payoff:

$$g_t = \mathbb{E}_t^{\mathbb{Q}}\left[e^{-\int_t^T r_s ds} \cdot (P_T(U) - K)_+\right]$$

However, this computation can be fairly complicated. It is considerably simpler if we use the T-bond as a numeraire:

$$\frac{g_t}{P_t(T)} = \mathbb{E}_t^{\mathbb{Q}^T}\left[\frac{(P_T(U)-K)_+}{P_T(T)}\right] \qquad \Longrightarrow \qquad g_t = P_t(T)\cdot\mathbb{E}_t^{\mathbb{Q}^T}\left[(X_T-K)_+\right]$$

where $X_t = \frac{P_t(U)}{P_t(T)}$. It is clear by defining this process that X_T is equal to one part of the term in the max function at time T; the other term is unaffected since $P_T(T) = 1$ and $K/P_T(T) = K$.

The next step is to find the dynamics of X_t under $\mathbb{Q}^T$. Since X_t is a relative price process with the numeraire in the denominator, the process X_t must be a martingale. The volatility can be computed by an application of Itô's quotient rule, and turns out to be the difference in the two bond's volatilities:

$$\frac{dX_t}{X_t} = \sigma(B_t(T)-B_t(U))\,dW_t^{\mathbb{Q}^T} \qquad \Longrightarrow \qquad X_T \overset{d}{=} X_t\cdot\exp\left[-\frac{1}{2}\Sigma^2+\Sigma\,Z\right]$$

where $\Sigma^2 = \sigma^2\int_t^T(B_s(T)-B_s(U))^2\,ds$ and Z is a standard normal random variable under $\mathbb{Q}^T$. Since X_t is lognormal, we can simply apply the Black-Scholes option pricing formula with a strike price of K to obtain the bond option price:

$$g_t = P_t(U)\cdot\Phi(d_+) - KP_t(T)\cdot\Phi(d_-)\quad , \qquad d_\pm = \frac{\ln\left(\frac{P_t(U)}{KP_t(T)}\pm\frac{1}{2}\Sigma^2\right)}{\Sigma}$$

14.5 Yields, Spot Rates and Forward Rates

Associated with every zero-coupon bond price for a given term is a bond yield and a spot rate, which are defined as follows:

$$P_t(T) = e^{-y_t(T)(T-t)} = \left[1+\ell_t^T\cdot(T-t)\right]^{-1}$$

where $y_t(T)$ is the bond yield and ℓ_t^T the spot rate associated with this bond. Of course, these are functions of time and maturity; at time t a zero-coupon bond that matures at time U will have an associated spot rate ℓ_t^U which satisfies $P_t(U) = \left[1+\ell_t^U\cdot(U-t)\right]^{-1}$. By rearranging the terms, **we can write spot rates in terms of bond prices** as follows:

$$\ell_t^T = \frac{1}{T-t}\left[\frac{1}{P_t(T)}-1\right]$$

Forward rates relate spot rates of bond prices with two different maturities:

$$\left(1+\ell_t^U(U-t)\right) = \left(1+\ell_t^T(T-t)\right)\cdot\left(1+\ell_t^{T,U}(U-T)\right)$$

where $\ell_t^{T,U}$ is the forward rate for the period T to U. We can express the forward rate in terms of bond prices as well:

$$\begin{aligned}\left(1+\ell_t^U(U-t)\right)^{-1} &= \left(1+\ell_t^T(T-t)\right)^{-1}\cdot\Big(1+\ell_t^{T,U}(U-T)\Big)^{-1}\\ \implies \qquad \ell_t^{T,U} &= \frac{1}{\cancel{T-t}}\left[\frac{P_t(T)}{P_t(U)}-1\right]\end{aligned}$$

Note that since the forward rate involves a relative price process with a U-maturity bond in the denominator, **the forward rate $\ell_t^{T,U}$ is a martingale under the U-forward neutral measure,** $\mathbb{Q}^U$. This is the basis of the **log-normal forward rate model**, in which forward rates are modelled as follows:

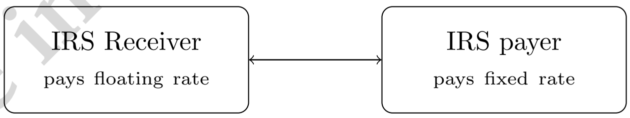
$$\frac{d\ell_t^{T,U}}{\ell_t^{T,U}} = \sigma_t^{T,U}\, dW_t^{\mathbb{Q}^U}$$

This model is often used to price caplets are floorlets, which are the equivalent of call and put options written on forward rates directly. As the name suggests, they place a cap/floor on the maximum/minimum rate that the holder has to pay/receive.

When we have a partition for time, we let T_k denote the various time points and let $\tau_k = T_K - T_{k-1}$ and denote $\ell_t^{T_{k-1},T_k} = \ell_t^{(k)}$.

14.6 Interest Rate Swaps and Swaptions

An **interest rate swap** is a derivative in which one party exchanges a set of fixed cash flows (fixed leg) with another set of cash flows that are based on a floating interest rate (floating leg). The IRS payer is the party that pays the fixed amounts, while the IRS receiver receives the fixed amounts (and pays the floating amount).



The payments are all proportional to a notional amount N and the elapsed time τ_k. The **fixed leg** payments are also proportional to a fixed rate F, while the **floating leg** payments are proportional to the prevailing forward rate. That is, for a payment at time T_k, the payment is proportional to the forward rate at the time of the previous payment $\ell_{T_{k-1}}^{(k)}$.

The **swap rate** is the value of F such that the floating leg and fixed leg of the IRS have equal value.

Fixed leg:
The payments in the fixed leg are of the amount $N \cdot F \cdot \tau_k$. The value of the

fixed leg is found by discounting all the cashflows to the present time using bond prices:

$$\begin{aligned} V_t^{\text{fixed}} &= \sum_{k=1}^{n} N \cdot F \cdot \tau_k \cdot P_t(T_k) \\ &= N \cdot F \cdot \sum_{k=1}^{n} \tau_k \cdot P_t(T_k) \end{aligned}$$

Floating leg:
The k-th payment in the floating leg is of the amount $N \cdot \tau_k \cdot \ell_{T_{k-1}}^{(k)}$. The value of the floating leg is obtained by computing the value of each payment individually using the $\mathbb{Q}^{T_k}$ measure:

$$\begin{aligned} \frac{V_t^{\text{floating},k}}{P_t(T_k)} &= \mathbb{E}_t^{\mathbb{Q}^{T_k}} \left[\frac{N \cdot \tau_k \cdot \ell_{T_{k-1}}^{(k)}}{P_{T_k}(T_k)} \right] \\ &= N \cdot \tau_k \cdot \mathbb{E}_t^{\mathbb{Q}^{T_k}} \left[\ell_{T_{k-1}}^{(k)} \right] && \text{since } P_{T_k}(T_k) = 1 \\ &= N \cdot \tau_k \cdot \ell_t^{(k)} && \text{since } \ell^{(k)} \text{ is a } \mathbb{Q}^{T_k}\text{-martingale} \end{aligned}$$

Using the relation between forward rates and prices we can write the value of the k-th payment of the floating leg as:

$$V_t^{\text{floating},k} = N \cdot (P_t(T_{k-1}) - P_t(T_k))$$

The value of the flaoting leg is then the sum of the individual cash flows:

$$V_t^{\text{floating}} = N \cdot \sum_{k=1}^{n} (P_t(T_{k-1}) - P_t(T_k)) = N \cdot (P_t(T_0) - P_t(T_n))$$

The last step follows since the sum is a telescoping sum. Setting the value of the two legs equal we the **value of the swap rate at time** t:

$$\mathcal{S}_t = \frac{(P_t(T_0) - P_t(T_n))}{\sum_{k=1}^{n} P_t(T_k)\tau_k}$$

A **payer/receiver swaption** is an instrument that gives the holder right the right to enter into the payer/receiver leg of an interest rate swap at time T and make/receive payments based on the rate K. The payoff function of a payer swaption (where the holder will pay the fixed rate) is then given by:

$$\varphi = \mathbb{1}\{\mathcal{S}_T > K\} \cdot \left(V_T^{\text{floating}} - V_t^{\text{fixed}} \right)$$

The indicator corresponds to the fact that the option will only be exercised if the fixed rate K is lower than the prevailing swap rate S_T at the time the option matures. It then makes sense to exercise the option as it will mean that the payer will have to pay smaller amounts. The second term corresponds to the difference between present value of the cash flow received (floating leg) and the cash flow paid (fixed leg). Using the relations derived earlier, and assuming a notional of 1 and equally spaced payments of with time increments ΔT we can write the payoff as:

$$\begin{aligned}
\varphi &= \mathbb{1}\{\mathcal{S}_T > K\} \cdot \left(P_T(T_0) - P_T(T_n) \; - \; K \cdot \Delta T \sum_{k=1}^{n} P_T(T_k) \right) \\
&= \left(\Delta T \cdot \sum_{k=1}^{n} P_T(T_k) \right) \cdot \mathbb{1}\{\mathcal{S}_T > K\} \cdot \left(\frac{P_T(T_0) - P_T(T_n)}{\Delta T \cdot \sum_{k=1}^{n} P_T(T_k)} - K \right) \\
&= A_T \cdot (\mathcal{S}_T - K)_+
\end{aligned}$$

where $A_T = \Delta T \cdot \sum_{k=1}^{n} P_T(T_k)$ is the value at time T of an annuity paying 1 at regular intervals. Since this is a tradeable asset it can be used as a numeraire. If we use the annuity as a numeraire we can write the value of a swaption as:

$$h_t = A_t \cdot \mathbb{E}_t^{\mathbb{Q}^A}\left[(\mathcal{S}_T - K)_+ \right]$$

and since $\mathcal{S}_T = \frac{P_T(T_0) - P_T(T_n)}{A_T}$ is a relative price with respect to the numeraire asset, it is a martingale under this measure. **The lognormal swap rate model is based on modeling swap rates as a lognormal process under $\mathbb{Q}^A$:**

$$\frac{d\mathcal{S}_t}{\mathcal{S}_t} = \sigma_t^{\mathcal{S}} \, dW_t^{\mathbb{Q}^A}$$

Note that there is no drift term to ensure that the swap rate is a martingale under the $\mathbb{Q}^A$. Now, if $\sigma_t^{\mathcal{S}}$ is deterministic, the swap rate is lognormally distributed and the swaption value can be computed using the Black-Scholes option price formula.

14.7 Exercises

Exercise 14.1. Show that call prices can always be written as follows:

$$C_0 = S_0 \cdot \widehat{\pi} - \frac{K}{1+r} \cdot \pi$$

where $\pi = \mathbb{Q}(S_1 > K)$, $\widehat{\pi} = \widehat{\mathbb{Q}}(S_1 > K)$ and $\widehat{\mathbb{Q}}$ is the measure induced by using S as a numeraire. *Hint: write* $C_1 = \varphi(S_1) = (S_1 - K)_+ = (S_1 - K)\mathbb{1}_{\{S_1 > K\}}$.

Exercise 14.2. **Path-dependent options.** Define $\bar{S}(n)$ as the geometric average of the asset's price over the n ordered times $t_0 = 0 < t_1 < t_2 < \cdots < t_n = T$. That is, $\bar{S}(n) := \left(\prod_{j=1}^{n} S(t_j) \right)^{1/n}$. An Arithmetic Asian Call option pays the following at the maturity date T

$$\varphi(S(t_1),\dots,S(t_n)) = \left(\frac{1}{n}\sum_{m=1}^{n} S(t_m) - K\right)_+ .$$

A Geometric Asian Call option pays the following at the maturity date T

$$\varphi(S(t_1),\dots,S(t_n)) = \left(\left(\prod_{m=1}^{n} S(t_m)\right)^{1/n} - K\right)_+ .$$

(i) There is no closed form equation for the price of an Arithmetic Asian option; however, the Geometric Asian can be obtained exactly – derive it!
[Hint: What is the distribution of $\overline{S}(n)$?]

(ii) Determine the probability that $\overline{S}(n) > K$ given that $\overline{S}(m) > K$, where $m < n$. Write the result in terms of the normal distribution function $\Phi(a) := \mathbb{P}(N_1 < a)$ and bivariate normal distribution function $\Phi(a,b;\rho) := \mathbb{P}(N_1 < a, N_2 < b)$ where N_1 and N_2 are standard normal random variables with correlation ρ.

(iii) Suppose you approximate the following sum (which appears in the arithmetic Asian option payoff) by a log-normal distribution:

$$\frac{1}{n}\sum_{m=1}^{n} S(t_m) \overset{d}{\approx} e^X$$

with X normally distributed. Based on this approximation, derive an analytical expression to approximate arithmetic Asian option prices.

EXERCISE 14.3. Suppose that the price of a stock is modeled as follows:

$$\frac{dS_t}{S_t} = \mu_t\,dt + \sigma_t\,dW_t$$

where μ_t and σ_t are functions only of time and where W_t is a $\mathbb{P}$-Wiener process. Furthermore, assume that the risk-free interest rate r_t is function only of time. Determine the price, the delta and the gamma for each of the following options:

(i) call option maturing at T strike of K.

(ii) forward starting put option with strike set to αS_U at time U and maturing at T.

EXERCISE 14.4. **Multi-factor Interest Rate Model.** Suppose that the short rate of interest follows a two-factor mean-reverting Vasicek model:

$$\begin{aligned} dr_t &= \alpha(\theta_t - r_t)\,dt + \sigma\,dW_t^r \\ d\theta_t &= \beta(\phi - \theta_t)\,dt + \eta\,dW_t^\theta \end{aligned}$$

where W_t^r and W_t^θ are $\mathbb{Q}$-Wiener processes.

(i) ** Solve the SDE for r_t and determine its distribution.

(ii) Find the distribution of $I_t = \int_0^t r_s\, ds$ and use it to determine the price of bonds in this model.

(iii) Write down the PDE which bond prices satisfy. Assuming that the model is affine, i.e. $P_t(T) = \exp\{A_t(T) - B_t(T)r_t - C_t(T)\theta_t\}$, reduce the PDE to a system of ODEs and solve for A,B and C.

(iv) Determine the SDE for the T-maturity bond price under the measures $\mathbb{Q}$ and $\mathbb{Q}^T$.

(v) Determine the SDE for the T-maturity bond yield under the measures $\mathbb{Q}$ and $\mathbb{Q}^T$.

(vi) ** Determine the SDE for the forward LIBOR rate process

$$l_t(T_1, T_2) = \frac{1}{T_2 - T - 1}\left(\frac{P_t(T_1)}{P_t(T_2)} - 1\right)$$

under the measures $\mathbb{Q}$ and $\mathbb{Q}^{T_1}$ and $\mathbb{Q}^{T_2}$.

Exercise 14.5. **Multiple Economies.** In a three economy system A, B, and C, let S_t^A denote a traded stock in A, S_t^B denote a traded stock in B, S_t^C denote a traded stock in C. Let X_t^{BA} denote the exchange rate for B's currency into A's currency, etc... Let r_t^A, r_t^B, r_t^C denote the three economies interest rates. Suppose that,

$$\frac{dS_t^i}{S_t^i} = \mu^i\, dt + \sigma^i \cdot dW_t$$
$$\frac{dX_t^i}{X_t^i} = \alpha^i\, dt + \eta^i \cdot dW_t$$
$$dr_t^i = \theta^i\, dt + \gamma^i \cdot dW_t$$

where W_t is a d-dimensional vector of independent Brownian motions, $\mu^i, \alpha^i, \gamma^i$ are scalar constants, $\sigma^i, \eta^i, \gamma^i$ are constant d-vectors.

(i) Show that the SDE coefficients for X_t^{BA} and X_t^{CB} determine the SDE coefficients for X_t^{AC} uniquely and determine them.

(ii) Determine the price of a digital call on the exchange rate X_t^{BA} for an A investor and a B investor. Note the A investor receives 1 unit of A currency, while the B investor receives 1 unit of B currency if the option is in-the-money at maturity.

(iii) Determine the price of a digital call on the exchange rate X_t^{AB} for an A investor and a B investor. Note the A investor receives 1 unit of A currency, while the B investor receives 1 unit of B currency if the option is in-the-money at maturity.

(iv) Determine the price of a foreign digital call on the exchange rate X_t^{AB} for an A investor. This option pays 1 unit of B currency if the options is in-the-money at maturity.

(v) Determine the price, for an A investor, of a European contingent claim on S_t^B with strike equal to K (in B's currency).

(vi) Determine the price, for an A investor, of a European contingent claim on S_t^B with strike equal to K (in A's currency).

EXERCISE 14.6. (i) **Random delay.** Consider a CDS under constant interest rate r and constant hazard rate λ. When a default occurs, the protection buyer ceases to make payments, however the protection buyer does not receive a recovery payment (of $1 - R$) immediately; instead, the investor must wait a random amount of time $T \geq 0$ (exponentially distributed with rate κ) independent of the default time from the instance of default.

Derive an expression for the CDS rate, provide intuition for the result and describe how it differs from the case without random delay.

(ii) **Counterparty risk.** You are given the following:

- yields of risk-free bonds: $r(1) = 1\%$, $r(3) = 2\%$, $r(3) = 2.5\%$
- yields of corporate bonds issued by B: $\overline{r}(1) = 1.25\%$, $\overline{r}(2) = 2.5\%$, $\overline{r}(3) = 2.75\%$
- yields of corporate bonds issued by C: $\overline{r}(1) = 2\%$, $\overline{r}(2) = 3.5\%$, $\overline{r}(3) = 4\%$

You wish to purchase 3-year credit protection from counter-party B on reference entity C with yearly payments.

(a) Ignoring counterparty risk, and assuming the recovery rate is 0, determine the CDS spread.
(b) Now take into account counterparty risk, assuming XYZ and B default independently. If B defaults, then all cash-flows stop. Determine the CDS spread.

EXERCISE 14.7. **Defaultable Bonds in Continuous Time.** In this question, default is modeled by a doubly stochastic Poisson process with hazard rate λ_t, and the default time is the first arrival time of the counting process. The short rate and the activity rate risk-neutral dynamics are:

$$dr_t = \alpha(\theta - r_t)\,dt + \sigma_r\,dW_t^r\ ,$$

$$d\lambda_t = \beta(\phi - \lambda_t)\,dt + \sigma_\lambda\,dW_t^\lambda\ ,$$

where W_t^r and W_t^λ are correlated $\mathbb{Q}$-Wiener processes, $d[W^r, W^\lambda] = \rho\,dt$.

The defaultable bond price (conditional on the company not in default) can be written as

$$\overline{P}_t(T) = \mathbb{E}_t^{\mathbb{Q}}\left[e^{-\int_t^T (r_s+\lambda_s)\,ds}\right]\ .$$

Furthermore, the defaultable bond price can be written in terms of the risk-free bond price and the forward neutral probability of survival: $\overline{P}_t(T) = P_t(T)\,S_t^T(T)$ where the forward-neutral survival probabilities are defined as $S_t^T(u) := E_t^{\mathbb{Q}^T}\left[e^{-\int_t^u \lambda_s\,ds}\right]$.

(i) Derive the PDE which $\overline{P}_t(T)$ satisfies.

(ii) Derive the PDE which $S_t^T(u)$ satisfies and **solve it**.
Hint: You need to find the dynamics of λ in the forward neutral measure first – it is still affine under $\mathbb{Q}^T$!

EXERCISE 14.8. **Jump Models.** You are using the risk-neutral dynamics for two equity price processes A_t and B_t:

$$A_t = e^{X_t} \qquad B_t = e^{Y_t} \qquad \text{where,} \tag{14.1}$$
$$dX_t = \alpha\, dt + \sigma\, dW_t + \delta\, dN_t\ , \tag{14.2}$$
$$dY_t = \beta\, dt + \eta\, dZ_t + \gamma\, dN_t\ . \tag{14.3}$$

Here, W, Z are correlated $\mathbb{Q}$-Wiener processes (correlation ρ) and N_t is a homogenous $\mathbb{Q}$- Poisson process with intensity λ – independent of the Wiener processes.

(i) Find α and β in terms of the remaining parameters – recall the model is already under risk-neutral measure.
(ii) What SDE do the processes $U_t = A_t B_t$ and $V_t = A_t/B_t$ satisfy?
(iii) Find the price of an exchange option which pays $(A_T - B_T)_+$ at maturity T using (i) conditioning on the number of jumps (ii) transform methods.
(iv) Use simulation to confirm your pricing equations. Use the following parameters: $A_0 = B_0 = 100$, $r = 5\%$, $\sigma = 20\%$, $\eta = 15\%$, $\delta = 0.05$, $\gamma = -0.05$, $\lambda = 4$, $T = 1$.

EXERCISE 14.9. **Markov-Chain Stochastic Volatility.** Let Z_t denote a K-state continuous time Markov chain with constant $\mathbb{Q}$-generator matrix A.

(i) Let $X_t = \int_0^t \sigma(Z_{s-})dW_s$ where W_t is an independent Brownian motion. What does the expectation

$$h_t = \mathbb{E}_t[g(X_T)\mathbb{I}_{Z_t=k}]$$

correspond to? Find h_t explicitly. *Hint: write down the PIDE which $h(t, X, Z)$ must satisfy and solve using the method developed in class.*
(ii) Assume that an asset price S_t evolves according to the SDE:

$$\frac{dS_t}{S_{t-}} = \nu(Z_{t-})\, dt + \sigma\, dW_t + \gamma(Z_{t-})\, dN_t$$

where W_t is a risk-neutral Brownian motion, N_t is a risk-neutral Poisson process with activity rate λ, $\gamma(k)$ is a regime dependent jump multiplier and $\nu(k)$ is a regime dependent drift. Interest rates are assume constant and equal to r in all regimes.
(a) What is $\nu(k)$ – recall the model is already under risk-neutral measure.?
(b) Solve the SDE for the asset price by introducing $\ln S_t$ and perform the resulting stochastic integral.
(c) Find the PIDE which an option which pays $\varphi(S_T)$ at T satisfies.

(d) Solve the PIDE using the methods developed in class.

Exercise 14.10. **SDE Simulation.** Use the (i) Euler and (ii) Milstein discretizations to simulate sample paths (using 252 steps per year) and obtain terminal distributions at $t = 1$ year for the following SDEs:

(i) $dX_t = W_t\, dW_t$.
(ii) $dX_t = W_t^2\, dW_t$.
(iii) $dX_t = X_t(0.1\, dt + 0.2\, dW_t)$.
(iv) $d\ln(X_t) = (0.1 - \frac{1}{2}\, v_t)\, dt + \sqrt{v_t}\, dW_t^{(1)}$ and $dv_t = 10\,(0.2^2 - v_t)\, dt + 0.01\sqrt{v_t}\, dW_t^{(2)}$, where $W_t^{(1)}$ and $W_t^{(2)}$ have instantaneous correlation of -0.5. Note: when simulating v_t always replace v_t with $|v_t|$ if v_t becomes negative (this is called a reflecting boundary condition).

Exercise 14.11. Needs fix! Now assume that the drifts and volatilities are constants.

(i) Consider a T-maturity call option struck at K written on the geometric average of the two stocks. That is, the option pays at maturity:

$$\varphi(S_T, U_T) = \Big((S_T\, U_T)^{1/2} - K\Big)_+ \;.$$

Show that the price of this option is

$$V_t = e^{-(r-\tilde{\mu})(T-t)}\,(S_t\, U_t)^{1/2}\,\Phi(d_+) - e^{-r(T-t)}\, K\,\Phi(d_-), \qquad \text{where ,}$$
$$d_\pm = \frac{\frac{1}{2}\ln(S_t U_t / K^2) + (\tilde{\mu} - \frac{1}{2}\tilde{\lambda}^2)(T-t)}{\sqrt{\tilde{\lambda}^2 (T-t)}}$$
$$\tilde{\mu} = r - \frac{1}{8}(\sigma_S^2 - 2\rho\sigma_S\sigma_U + \sigma_U^2) \qquad , \; \tilde{\lambda} = \frac{1}{2}(\sigma_S^2 + 2\rho\sigma_S\sigma_U + \sigma_U^2)^{1/2}$$

(ii) Consider a T-maturity call option struck at K written on the arithmetic average of the two stocks. That is, the option pays at maturity:

$$\varphi(S_T, U_T) = \big(\tfrac{1}{2}(S_T + U_T) - K\big)_+ \;.$$

Use Monte Carlo simulation with the geometric average option as a control variate to price this option maturing in 6 months, with spot of 100, strike of 100, expected return of 10%, volatility of 20% and assume the risk-free rate is 5%. Use five runs, with 1000 simulations per run, to obtain the confidence intervals.

Exercise 14.12. Needs fix! Assume that the drifts and volatilities are functions of time and the driving Wiener process.

(i) Use the Black-Scholes dynamic hedging argument to show that the price function $f_t(S_t, U_t)$ of an option which pays $\varphi(S_T, U_T)$ at the maturity date T satisfies the following PDE:

$$r\,f = \left(\partial_t + r\,\left(S\,\partial_S + U\,\partial_U\right) + \frac{1}{2}\left((\sigma^S)^2\,S^2\,\partial_{SS} + 2\,S\,U\,\sigma^S\,\sigma^U\,\rho\,\partial_{SU} + (\sigma^U)^2\,U^2\,\partial_{UU}\right)\right) f$$

subject to $f_T(S_T, U_T) = \varphi(S_T, U_T)$. Make sure that you report the hedge position for both S_t and U_t. *Note I have suppressed the explicit dependence on t, U_t and S_t in the PDE to shorten the notation.*

(ii) Suppose that $\varphi(S_T, U_T) = S_T\,g(U_T/S_T)$. By assuming that $f_t(S_t, U_t) = S_t\,h_t(U_t/S_t)$, find the PDE which $h_t(z)$ satisfies. Compare this new PDE to the one-dimensional Black-Scholes PDE, solve for h using the Feynman-Kac Theorem, and show that

$$f_t(S_t, U_t) = S_t\,\mathbb{E}^{\mathbb{Q}^*}\left[g(Z_T)\right]\,,$$

where the process $Z_t := U_t/S_t$ is a $\mathbb{Q}^*$-martingale, and explicitly Z_t satisfies the following SDE:

$$\frac{dZ_t}{Z_t} = \sqrt{(\sigma^U)^2 - 2\rho\sigma^U\sigma^S + (\sigma^S)^2}\,d\overline{W}_t\,.$$

where $\overline{W}$ is a $\mathbb{Q}^*$-Wiener process.

Exercise 14.13. **Interest Rate Trees.** Assuming that interest rates are model as follows:

$$r_{n\Delta t} = r_{(n-1)\Delta t}\,e^{\sigma\sqrt{\Delta t}\,X_n}$$

where $X_1, X_2, \dots$ are iid Bernoulli random variables with $\mathbb{Q}(X_j = +1) = \frac{1}{2}$ and $\mathbb{Q}(X_j = -1) = \frac{1}{2}$. Assume that $r_0 = 5\%$, $\sigma = 10\%$ and $\Delta t = \frac{1}{2}$. Use Excel to:

(i) Determine the price of a 4-year zero coupon bond on a notional of 1 million.
(ii) Determine the price of a call option maturing in 2-years on the bond in part(ii) with a strike equal to 0.95 million.
(iii) Determine the replication strategy for the call option in part (ii) using the 4-year zero bond and the 3-year zero bond as replication instruments (both on notionals of 1 million).

Exercise 14.14. **Ho-Lee Model.** Suppose that interest rates follow the following dynamics:

$$dr_t = \alpha_t\,dt + \sigma\,dW_t$$

where α_t is a deterministic function of time and W_t is a $\mathbb{Q}$-Wiener process. Determine each of the following:

(i) Bond price at time t of maturity T.
(ii) The SDE which the bond price satisfies in terms of W_t.
(iii) The choice of α_t which makes the model prices equal the market prices $P_t^*(T)$.

Subject index