CALCULATION OF PREMIUM RESERVES IN INCOMPLETE MARKETS

by

Zhijia Zhu

THESIS

for the degree of

Master in Modelling and Data Analysis

MOD5930

(Master of Science)

Det matematisk-naturvitenskapelige fakultet

Universitetet i Oslo

May 2015

Faculty of Mathematics and Natural Sciences

University of Oslo

Acknowledgment

After spending five years in studying mathematics and statistics, this master thesis is a great opportunity for me to apply the theories in a more concrete subject, i.e. actuarial science. The writing of this thesis started in February 2015 and ended in May 2015, during which I encountered a lot of challenges and barriers. But still, after solving all those problems and completing the thesis, I would say it is an invaluable chance for me to restructure, reorganize, reevaluate and reframe all the knowledge and subjects I have learned from the Department of Mathematics, University of Oslo.

Here I must offer my sincerest gratitude to my supervisor, Prof. Frank Norbert Proske, whose patience, knowledge, expertise and encouragement have always been the greatest support and inspiration for me thoughout the course of this thesis. I truely appreciate all his contributions of time, kind assistance, wise advices and solid knowledge in actuarial science during my writing process. Without his professional supervision this thesis would not have been possible.

Those fantastic and interesting people that I have met during my postgraduate and graduate studies at the University of Oslo, who provided me with new ideas, discussions and lent their helping hand directly or indirectly in this venture, deserve my deepest thanks for keeping things in perspective.

Last but of course not least, I am deeply thankful to my beloved parents and my brothers, for their utfailing support and continuous assistance throughout my period of studies and through the course of writing this thesis.

<div align="right">

Zhijia Zhu

May 2015

Oslo, Norway

</div>

Summary and Conclusions

The thesis introduces unit-linked insurance contracts and insurance policies based on stochastic interest rates and investigates their pricing methods. As the most distinct property of such products is their connection to the financial market through which market risks are induced, stochastic analysis must be applied to explore the reasonable prices of the corresponding derivative securities those products are linked to, therefore enlarging the difficulty of the calculation of prospective reserves.

In the financial market, Black & Scholes option pricing formula has long been the classical foundation for entry-level mathematical finance course, nevertheless its assumption on complete market structure is problematic in practice. In this thesis, one step forward is taken to look at incomplete market structure, where the reason of this is assumed to be interest rate/stock price jumps. With respect to modelling such jumps Lévy processes will be used in this thesis due to their nice property of stationary and independent increments.

The main challenge under incomplete market situation is the non-uniqueness of risk-neutral probability measures. There will be in general many equivalent martingale measures (EMMs) that can be selected in the market, each of which gives a legitimate risk-neutral price for the derivative secutiry product at hand. Under such circumstances several pricing methods specific for incomplete market assumption have been developed, but throughout this thesis only the method of Esscher transform will be utlized. The reason to this is that the transformation preserves the desirable property of stationary and independent increments, in the sense that the transformed process remains a Lévy process.

The main focus of the thesis lies on the derivation of the recursive formula of prospective reserves for unit-linked products or policies where interest rates are stochastic, such that the insurance contract is linked to financial risks via the bond/stock market. The final formula is a complicated integro-partial differential equation, where analytical solutions are in general impossible. Simulation methods like Monte Carlo must be applied to calculate the corresponding reserves for the such products.

Contents

Chapter 1

Introduction

This chapter gives a general overview over the thesis, including the background underlying the current insurance and finance markets, methods and approaches that are going to be used, limitations found in the current academic literature, connection to a broader finance context, objective of the thesis, formulation of the problem and structure of the thesis. After reading the chapter, readers should have an insight on what the topic of the thesis is, how the structure of the thesis is going to be formulated and what remains to be done for further research.

1.1 Background

Currently, life insurance companies are suffering from low interest rates and increasing longevity. Many life insurance firms provide products that last decades long, and the ability of generating enough money for future payments of those products is crucial for insurance firms' stability. Interest rates play a vital role when it comes to how much money insurance firms are going to make from their investments as most of those investments are made of fixed-income products. In order to manage interest rate risks, asset-liability management (ALM) has long been the central job of life insurance firms. Due to the long-term nature of most life insurance products, fixed-income securities, i.e. bonds, are the major assets of life insurers. However, bonds are extremely sensitive to interest rate changes as their maturities are usually 5 or 10 years long, through which insurance companies are subject to risks relating to interest rate fluctuations. When interest rates have dropped significantly, mismatch of assets and liabilities appears, requiring that the insurance company uses its own capital to make up the gap of the mismatch.

One thing that complicates the situation for insurance companies is that most of their products provide a minimum guarantee of return. Normally the guarantee level is set out-of-money, i.e. the guarantee is lower than the current interest rate level, upon entering into the contract. But later if interest rates drop, the contract will become in-the-money, i.e. the return the policyholder gets is higher than what he/she would have gotten on other investment alternatives. Under such circumstances life insurance firms also have to use own assets to compensate for the difference, by which exposes them to higher uncertainties. Moreover, some products, typically annuities, allow policyholders to adjust their premium payments during the contract period. Thus, when the interest rates are high and the policy is out-of-money, the policyholder will be more willing to reduce payments or even surrender (cancel) his/her policy and receives a surrender value such that the money can be invested in other alternatives. On the other hand, if the interest rates are low and the policy is in-the-money, the policyholder will have desire to increase payments to his/her policy so as to get higher return than what he/she normally would have gotten from other alternatives. Therfore, policyholders' behaviors magnify the extent to which life insurance firms are exposed to interest rate risks.

Since annuities promise prolicyholders a predetermined level of return, life insurers can only

make money if the guaranteed returns are lower than the market interest rates. Thus, if the interest rates are significantly lower, those annuities must also be designed such that their promised returns are low as well. This will in turn make the annuities uncompetitive to other financial products.

Apart from managing the cash flows of assets and liabilities, where the duration of the assets of a life insurer are put to match the duration of the liabilities in a long term, usually decades, another way to decrease the company's exposure to interest rate risks is the purchase of interest-rate swaps. Over the last twenty years interest rate swaps have gained a huge increase in life insurance companies' assets, and have today become the major tool for life insurers to hedge interest risks.

Therefore, the main challenge life insurers are facing today can be characterized as matching the cash flows of the assets and the liabilities of the companies by statistical modelling. That is, calculating the reserves for the company when interest rate drops or increases can be taken into account. However, most models in actuarial science for calculating reserves assume constant interest rates, which is fairly unrealistic in practical usage due to the fact that over such a long period of time, variations of interest rates are huge. Simply ignoring this fact and instead assuming constant interest rates would lead to problematic results.

This thesis attempts to calculate insurance reserves where interest rates are allowed stochastic, by which the results are adjustable to interest rate changes during the course of the contract. Stochastic interest rates will enter the model through analyzing the bond prices, and then the life insurance contract will be connected to the bond market through a unit-linked scheme. The reason to this construction is due to the fact that stochastic interest rates are described by zero-coupon bonds. Modelling prices of zero-coupon bonds is the same as modelling stochastic interest rates, and the unit-linked scheme is like a socket that enables the insurance reserve formula to be connected to stochastic interest rates.

Problem Formulation

In the classical Black & Scholes option pricing model, the underlying market is assumed to be complete. The completeness of the market makes that every contingent claim in the market can be perfectly replicated, and gives the market the property that there is only one risk-neutral

probability measure in the market, such that the absence of arbitrage opportunities is equal to the existence of a risk-neutral probability measure.

However, this assumption is quite problematic in real-life applications as the market we have in practice are actually incomplete! The reasons to this phenomenon are many, but in this thesis only one of them will be investigated, namely, the discontinuities of the interest rate/stock price trajectories. In the Black & Scholes model, Brownian motion is utilized to model the diffusion part of the stock. Since Brownian motions are no where differentiable but everywhere continuous, the trajectories of the stock price is also continuous. This continuity is one of the many properties that makes the market complete. If the continuity assumption of the underlying stock price is not satisfied, the market will be incomplete.

Real-life observations tell us that interest rates often jump rapidly during even a short-time period. Modelling their trajectories as continuous is not realistic. It is therefore necessary to look for models that also can describe jumps. This thesis wil be concentrated mainly on a model called Lévy process, which is capable of modelling discontinuous trajectories. The biggest challenge of applying jump-diffusion models is that the underlying market is no longer complete, and hence absence of arbitrage opportunities is not equal to the existence of a unique risk-neutral probability measure. The absence of arbitrage opportunity property can still be constructed as before, but now there are many risk-neutral probability measures available to be chosen from. Each of these measures gives a legitimate risk-neutral price, and the problem is which one should be chosen as the best price.

There are several criteria that can help people to pick the right price. The method that is going to be applied in this thesis is the one called Esscher transform. By transforming the density function of the corresponding distribution one gets a specific risk-neutral probability measure that enables one to derive a risk-neutral price for the stock when jumps are allowed in its trajectories.

The more traditional unit-linked insurance contracts will be connected to the stock via a call option. After buying the contract, the policyholder will be given a refund guarantee such that if the life condition of him/her shifts, the payment will depend on the payoff of the call option.

The problem connected to the above unit-linked contract is that it does not have an analytical solution. So simulation must be applied to approximate the right price.

The above scheme is a more ordinary one, and will be only given for illustration purposes. This thesis will mainly concentrate on another type of unit-linked insurance contracts, where the policyholder will not be given the right to invest their paid-in premiums, and thus the payment to them is not related to the financial market. Instead, the interest rates will now be assumed to be stochastic.

The pricing of such unit-linked insurance contracts will rely on the application of Esscher transforms for selecting a specific risk-neutral probability measure. The Esscher transform is appropriate in the current model because of its desirable property that it preserves the Lévy propery of the interest rate model. Thus, after Esscher transforming the density function, the new process is still Lévy, and one has now a (local) martingale that can be used to calculate the expectation of future payments of the unit-linked contracts or policies based on stochastic interest rates.

Literature Survey

The tight connection to finance theories is the most obvious property of the pricing of unit-linked insurance contracts. Thus, this thesis will focus a lot on finance literature. Much of the finance theories of the thesis are based on the work done by e.g. [Tankov and Cont, 2003], where the introduction of Lévy processes, description of various pricing and hedging methods in incomplete markets, and Esscher transform are among the topics. While [Tankov and Cont, 2003] partially serves as a background of the thesis, [Sato, 1999] gives a more comprehensive introduction to Lévy processes and their application in finance.

Theories of insurance mathematics of the thesis are mainly based on [Møller and Steffensen, 2007] and [Koller, 2012]. Both books give a remarkable introduction to unit-linked insurance contracts.

The application of the Esscher transform in pricing security derivatives in incomplete markets will be taken from [Esscher, 1932] and [Gerber and Shiu, 1994]. Based on the work done by Esscher, Gerber and Shiu are among the first to apply the theory in pricing options where the logarithm of the underlying stock prices are governed by certain stochastic processes with stationary and independent increments in actuarial science, which has brought a hugh development to the field afterwards. Their work provides a tool in actuarial science for pricing unit-

linked insurance products where the underlying market is no longer complete. Unlike the traditional Black & Scholes option pricing model, the Esscher transform allows for selecting a specific risk-neutral probability measure in incomplete market, in such a way that it also preserves the stationary and independent increments property of the underlying stock price dynamics.

However, the thesis is devoted to pricing insurance products when interest rates are assumed stochastic, thus connecting to the financial market via bond markets. The replacement of the Brownian motion in a Vasicêk model by a Lévy process allows for interest rate jumps, making the market no longer complete.

Apart from the Esscher transform, there are also a handful of other mehods avaiable for finding an equivalent martingale measure (EMM) in order to price the underlying stock price dynamics. For example, [Fujiwara and Miyahara, 2003] proposes a method called minimal entropy martingale measure, while [Főllmer and Sondermann, 1986] suggests a method called mean-variance approach that aims to find a self-financing trading strategy minimazing the risks in the risk-neutral measure where the Galtchouk-Kunita-Watanabe decomposition is applied. [Főllmer and Schweizer, 1989] provides another method called local risk-minimization which minimizes variance based on the real-world probability measure P.

What Remains to be Done?

The pricing of insurance contracts where interest rates are stochastic involves finding the right Esscher transform parameter for transforming the corresponding density function when taking expectation. After deriving the Radon-Nikodym density, plugging it into the expectation formula and applying the Itô formula on the expectation formula, an integro-partial differential equation results. Some condition must be satisfied in order for the formula to be a (local) martingale. The final integro-partial differential equation is more difficult than the celebrated Black & Scholes formula, in that there are no analytical solution to the equation, and that certain assumptions have to be made about the Esscher transform. The application of jump diffusion processes in modelling stochastic interest rates in insurance contracts is a relatively new and small area, and hence there are few available articles and papers in academic literature. To the best of my knowledge, the proposed model of mine in this thesis on the calculation of premium reserves based on stochastic interest rates with jumps and the Esscher transform in connection

with a local martingale condition is new.

If the current low interest rate scenario lasts longer, most life insurance companies and pension funds will be affected enormously. Hence, modelling stochastic interest rates will be important for life and pension firms to manage their liabilities ans assets.

1.2   Objectives

This thesis investigates the reasons for incompleteness in the financial market, and explores various methods for pricing contingent claims under market imcompleteness. However, the main focus will be on insurance markets where the contracts are linked to the bond market via stochastic interest rates with jumps.

Apart from linking to the financial market via stochastic interest rates, unit-linked scheme can also enter into insurance contracts through direct connection, that is, the payment function of the insurance contract can be directly linked to available derivative products in the financial market, like options. This kind of unit-linked insurance contract will be given as a special example for illustration purposes. They are more traditional and more common in relation to the above one with stochastic interest rates, and there are hence more books devoted to this topic.

This thesis is an attempt to describe the methods used to price unit-linked insurance contracts. In the recent decades stochastic analysis has enjoyed huge advances in financial literature, and lots of new approaches and pricing models have been established to support the development of derivative products. Quantitative finance or computational finance, specifically, has gained more and more attention in the academics since modern finance theories keep applying more involved mathematics. On the other hand, actuarial science is still focusing on more traditional pricing methods, lacking the attention on the application of SDEs in its curriculum. This little thesis is thus aiming to give the reader a broader picture of how stochastic analysis can also be utilized in insurance industry.

1.3 Limitations

The biggest limitation when pricing contingent claims in incomplete markets is the lack of a unique risk-neutral probability measure. When market incompleteness appears, there will be a number of those probability measures to be chosen from. Different criteria give different results, and there is no a universal method practitioners agree upon. Therefore, it is not uncommon that all those methods have to be applied on the problem, and their results are compared.

The thesis will only concentrate on one of the many reasons that makes the market incomplete, that is, interest rate/stock price jumps. If the trajectories have jumps, then the market completeness assumption will be invalid. The exponential Lévy process is a good candidate for modelling trajectory jumps due to its stationary and independent increments property. Such property is desirable as it possesses mathematical tractability. Thus, in order to keep such property only Esscher transform will be used as a selection criterion to choose a specific risk-neutral probability measure.

After Esscher transforming by means of a Radon-Nikodym density, the final formula for calculating prospective reserves is an integro-partial differential equation, which does not have an analytical solution. Simulation methods like Monte Carlo must be used to model the reserves.

1.4   Approach

The main challenge in the thesis is the incompleteness of the market, resuting in several valid risk-neutral probability measures, each of which gives a risk-neutral price for the corresponding derivative security. The approach to solve the problem is the use of Esscher transform in deriving a Radon-Nikodym density which corresponds to a specific risk-neutral measure.

For modelling stochastic interest rates with jumps, a modified Vasicêk model where the Brownian motion is switched to a Lévy process will be used, hence resulting in an exponential Lévy process. The choice of the Vasicêk model is that it can be easily extended to allow for a Lévy diffusion driving part, such that it can take into account jump trajectories.

For propective reserves, a recursive formula will be constructed based on results in [Koller, 2012]. The formula can easily be extended to take into account stocastic interest rates or direct unit-linked scheme. For the former type one only needs to modify the discounting rate to be an expectation with respect to an information filtration, for the latter type one needs to equate the payment function for state shifts to the present value of a specific derivative security, like a European call option, for example.

1.5   Structure of the Thesis

Chapter 2 is a general introduction to the mathematical tools that are going to be useful in later applications, among which measure theory, probability theory, Lévy processes, Itô's formula and Black & Scholes formula are the topics. This chapter serves as a building block for the whole thesis. However, contents in the chapter are only given for illustration purposes as they can be easily found in any introduction proability and mathematical finance textbooks like [Berk and Devore, 2007] and [Pliska, 1997].

Chapter 3 first gives the several reasons to market incompleteness and then introduces various pricing methods in incomplete markets. Note that those are not the only available methods in the finance literature. There are many other useful methods as well, and new methods have been kept inventing all the time. Pricing of derivative securities in incomplete markets is an active and broad field.

Chapter 4 goes into explaining what unit-linked insurance contracts are and how they operate in the insurance market. Apart from this topic, the chapter will be devoted to formulating what Esscher transform really deals with and why they are useful in pricing derivative securities.

Chapter 5 focuses on insurance mathematics. It gives a derivation of the recursive formula for prospective reserves, and talks about its extension to allow for unit-linked scheme. The other part of the chapter concentrates on stochatic interest rate models, with special attention on the Vasicêk model.

Chapter 6 is the final chapter of the thesis, which gives the final model and its derivation and conditions. In the chapter there will be given two insurance schemes.

Chapter 2

Introduction to Some Mathematical Tools

Probability theory is a fundamental tool and a basic building block for modern stochastic analysis. In order to get a deeper insight into the core of stochastic analysis, a thorough knowledge in probability theory is a prerequisite. In this chapter, some fundamental definitions and concepts in probability theory relating closely to stochastic analysis are presented, together with explanations on why they are central in the later studies of stochasitic analysis, especially when it comes to the study of Lévy processes.

In the field of probability theory, measure theory is no doubt the cornerstone for building up the whole universe of subject. Thus, section 2.1 is devoted to introducing basic definitions and notions in measure theory, which is a comprehensive tool for understanding the concept of probability. Based on section 2.1, setion 2.2 goes one step forward and introduces the most usual member of the probabilistic family: random variables. A central notion that is going to be crucial in the later studies of Lévy processes, called characteristic function, will also be introduced here. Two classical working examples of stochastics processes, namely Brownian motion and Poisson process, will be given as well. Section 2.3 introduces the most basic element of the thesis, namely Lévy processes. Section 2.4 concentrates on Itô calculus, especially the celebrated Black & Scholes formula will be given in this section for later references.

2.1 Measure Theory

2.1.1 Definitions of Measures and σ-algebras

In mathematical analysis course you might probably have heard about the term Lebesgue measure. At first glance it seems mysterious to beginners, but in fact it is a classical example of measures. A measure is a generalization of the more commonly-known concepts length, volume and area etc. In a more mathematical language, a measure is a way of assigning a non-negative real-valued number or $+\infty$ to some subsets of a set X. In particular, if μ is a measure, then $\mu(\emptyset) = 0$ where $\emptyset$ denotes an empty set. But before a rigorous definition on measures can be given, an important concept called σ-algebras must be made clear:

Definition 2.1 *(σ-algebra) Let E be some set. Then a σ-algebra on E, denoted by $\mathcal{F}$, is a non-empty collection of subsets of E, which satisfies the following three conditions:*

1. *$\emptyset \in \mathcal{F}$.*

2. *If $A \in \mathcal{F}$, then $A^C \in \mathcal{F}$.*

3. *If $A_n \in \mathcal{F}$ for each n, and $(A_n)_{n \geq 1}$ are disjoint, then $\bigcup_{n \geq 1} A_n \in \mathcal{F}$.*

σ-algebra is the foundation for extablishing the notion of measure. With this definition at hand, measures are easily defined as:

Definition 2.2 *(Measure) Let $\mathcal{F}$ be a σ-algebra on some set E. A function $\mu : \mathcal{F} \to [-\infty, +\infty]$[1] is called a measure if and only if the following conditions are satisfied:*

1. ***Non-negativity**: For all subsets $A \in \mathcal{F}, \mu(A) \geq 0$.*

2. ***Null empty set**: $\mu(\emptyset) = 0$.*

3. ***σ-additivity**: If $(A_n)_{n \geq 1}$ is a sequence of disjoint subsets in $\mathcal{F}$, then $\mu(\bigcup_{n \geq 1} A_n) = \sum_{n \geq 1} \mu(A_n)$.*

If 3 in the above definition is met, then 2 is superfluous since by σ-additivity property, one can easily derive that:

$$\mu(A) = \mu(A \cup \emptyset) = \mu(A) + \mu(\emptyset), \text{ such that } \mu(\emptyset) = \mu(A) - \mu(A) = 0.$$

---

[1] Called affinely extended real number system and denoted by $\bar{\mathbb{R}}$.

All subsets of the σ-algebra $\mathcal{F}$ are measurable, and the pair $(E, \mathcal{F})$ is called a measurable space. Also pay special attention to the above definition that the possibility of infinite measure is not excluded. For some specific choice of E, its measure could be $\mu(E) = \pm\infty$. If instead $\mu(E) < \pm\infty$, consider $A \in E$ with its complement A^C such that $A \cap A^C = E$. The σ-additivity property defined above can be used to derive the relation that:

$$\mu(A \cap A^C) = \mu(A) + \mu(A^C) = \mu(E),$$

such that:

$$\mu(A^C) = \mu(E) - \mu(A).$$

Thus, the complement of a measurable set is also measurable.

Another important concept in basic measure theory that worth mentioning is Borel σ-algebra[2]. A Borel set is a set that is related to the topological space and can be generated by open sets[3] through operations like union, intersection or complement. A measure defined on the Borel set is called Borel measure. Also, for a specific topological space, the collection of all Borel sets on this space is called Borel σ-algebra and will be denoted as $\mathcal{B}$. Borel sets are fundamental in measure theory since all measures defined on the open sets of a space must also be defined on all Borel sets of that space. That means, for a compact space E, the Borel σ-algebra $\mathcal{B}(E)$ is the smallest σ-algebra of subsets of E that contains all the open subsets of E.

Now, the Lebesgue measure mentioned at the beginning of this subsection is defined to be a measure on the space $\mathbb{R}^n$ with the corresponding Borel σ-algebra $\mathcal{B}(\mathbb{R}^n)$ that for a subset $A \in \mathcal{B}(\mathbb{R}^n)$:

$$\mu(A) = \int_A dx,$$

which is the volume of A in n-dimensional sense. Namely, it is just an extension of length, area or volume to a more complicated scenario. In the simplest setting, the Lebesgue measure on an interval is just its length. That is, for $I = [a, b]$, the Lebesgue measure on I is simply just:

$$\mu(I) = |b - a|.$$

---

[2]Named after the French mathematician Émile Borel, who made considerable contributions to measure theory and probability.

[3]Or equivalently, by closed sets.

In the later studies of Lévy processes, a measure called Dirac measure is useful. Hence, it is defined precisely her. Dirac measure is a way of indicating whether a set contains a fixed point. More mathematically, it is:

Definition 2.3 *(Dirac measure) Let $(E, \mathcal{F})$ be a measurable space, $A \subseteq E$ a measurable subset, and $x \in E$ a point. Then the Dirac measure on the point x, denoted by δ_x, is defined to be $\delta_x : \mathcal{F} \to \bar{\mathbb{R}}$ such that:*

$$\delta_x(A) = \begin{cases} 0, & \text{if } x \notin A \\ 1, & \text{if } x \in A. \end{cases}$$

In other words, Dirac measure $\delta_x(A)$ is just an indicator function indicating whether $x \in A$.

2.1.2   Measures, Functions and Continuity

In measure theory, measurable functions play an important role because they form the basis for defining integration, as they preserve structures between measurable spaces. So, what are exactly measurable functions? Explained in words, a function between measurable spaces is said to be measurable provided that the preimage of each measurable set is also measurable. In a more mathematical context, they are defined to be:

Definition 2.4 *(Measurable function) Let $(X, \mathcal{F})$ and (Y, Σ) be two measurable spaces. That is, X and Y are two measurable sets equipped with the corresponding σ-algebras $\mathcal{F}$ and Σ. For a subset $A \subseteq \Sigma$, a function $f : X \to Y$ is said to be measurable if the preimage of A through f is in $\mathcal{F}$. In other words:*

$$f^{-1}(A) = \{x \in X | f(x) = A\} \in \mathcal{F}.$$

Until now the notion of σ-algebra might still seem theoretical, but in fact, in probability theory, it just represents the collection of available information. Measurable functions can also be defined in such a way that its outcomes are known only through available information. Thus, the concept of σ-algebras is central in the whole construction of probability theory, particularly stochastic processes, which are the main topic of later chapters.

Having defined measurable functions, it's time to look at the relationship between measurability and continuity. Unfortunately, measurability and continuity are not equivalent. Measurable functions include continuous functions. That is, a continuous function connecting two

topological spaces and equipped with the corresponding Borel σ-algebra is measurable. But not all measurable functions are continuous. A more formal connection from measurability to continuity is defined by the Lusin's theorem[4]. But first the definition of Radon measure must be stated:

Definition 2.5 *(Radon measure) A measure μ on the σ-algebra of a Hausdorff topological space[5] E is called a Radon measure if and only if the following two conditions are satisfied:*

1. **Inner regularity***: For all compact subsets[6] A of a Borel set $\mathcal{B}$, $\mu(\mathcal{B})$ is always the supremum of $\mu(A)$.*

2. **Locally finity***: For all points in E there exists a neighborhood U such that $\mu(U)$ is always finite.*

With Radon measure at hand, the Lusin's theorem can be defined as:

Theorem 2.1 *(Lusin's theorem) Assume $(X, \mathcal{F}, \mu)$ is a Radon measure and Y is a second-countable topological space[7]. Assume again:*

$$f : X \to Y$$

is a measurable function. Let $\epsilon > 0$ be a small number. Then, $\forall A \in \mathcal{F}$ of finite measure there exists a closed set E satisfying $\mu(A \backslash E) < \epsilon$ such that f defined on E is a continuous function.

Since the proof of the Lusin's theorem is quite involved, it's omitted here. But for the purpose of better understanding the relationship between measurability and continuity, a deduction from Lusin' theorem on one-dimensional space is provided:

Corollary 2.1 *(Application of Lusin's theorem on one-dimensional space) Let $f : [a, b] \to \mathbb{C}$[8] be a measurable function. Then for all $\epsilon > 0$ there exists a compact set $E \in [a, b]$ such that $\mu([a, b] \backslash E) < \epsilon$. If f is restricted on E, f is continuous. Here μ is a Lebesgue measure.*

Thus, loosely speaking, a measurable function is almost continuous.

---

[4]Named after the russian mathematician Nikolai Luzin, who was active in the field mathematical analysis in the 20th century.

[5]Distinct points have disjoint neighbourhoods, named after the German mathematician Felix Hausdorff.

[6]A subset A is compact if for every open covering of A there exists a finite subcovering of A.

[7]A topological space satisfying the second axiom of countability.

[8]$\mathbb{C}$ denotes the set of complex numbers.

2.2 Preliminaries in Probability Theory

2.2.1 Probability Spaces and Random Variables

An important application of measure theory is in probability theory. In this subsection, a rigorous introduction on some basic concepts of probability theory is given. Note that the introduction is not a thorough listing of concepts, definitions and theorems in probability theory. The knowledge is without any doubt easy to be found in any textbooks. The aim here, however, is to give the readers an insight on the most basic probability theories which are to be used as building blocks for constructing stochastic processes, especially jump processes in the later chapters.

First of all some fundamental concepts.

Definition 2.6 *(Probability measure) Let $(\Omega, \mathcal{F})$ be a measurable space where $\mathcal{F}$ is the corresponding σ-algebra. A probability measure P, defined on this space, is a function $P : \mathcal{F} \to [0, 1]$ such that the following criterions are satisfied:*

1. $P(\emptyset) = 0$ and $P(\Sigma) = 1$.

2. If $A_n \in \mathcal{F}$ for each n, and $(A_n)_{n \geq 1}$ are disjoint, then $P(\bigcup_{n \geq 1} A_n) = \sum_{n \geq 1} P(A_n)$.

Then, the triple $(\Omega, \mathcal{F}, P)$ is called a probability space. The probability measure P on the measurable space $(\Omega, \mathcal{F})$ is just a measure on $[0, 1]$, i.e. it's positive, finite, and with total mass 1.

Here, the measurable space $(\Omega, \mathcal{F})$ can be given concrete meanings. Ω represents the possible scenarios that can happen in the real world, and a subset $A \in \mathcal{F}$ represents the set of events among the Ω-scenarios that can be measured by P, i.e. that can be assigned a positive number in $[0, 1]$ indicating the probability that such an event occurs. In particular, if one has $A \in \mathcal{F}$ such that $P(A) = 1$, the event A is said to occur with probability 1, or almost surely (a.s.).

Consider an $\mathcal{F}$-measurable function $X : \Omega \to \mathbb{R}^n$, then X is called a random variable. That is, for each $\omega \in \Omega, X(\omega)$ represents the realization of the scenario ω.

Next is the notion of distribution function, which incorporates random variables and produces probabilities.

Definition 2.7 *(**Distribution function**) Let $(\Omega, \mathcal{F}, P)$ be a probability space. Assume $X : \Omega \to \mathbb{R}^n$ is a random variable. Then $\forall \omega \in \Omega$, $X(\omega)$ is a realization of the random variable X under scenario ω. Therefore:*

$$\mu_X(B) = P(X(\omega) \in B)$$

is called the distribution function of X.

Distribution function is often called cumulative distribution function (CDF). $X : \Omega \to \mathbb{R}$ is called a real-valued random variable. Thus, distribution function defined on a real-valued random variable X is just a representation of the probability that X is less than or equal to a given real constant number.

Now assume that X is a read-valued random variable with probability space $(\Omega, \mathcal{F}, P)$, then under the condition that:

$$\int_{\Omega} |X(\omega)| dP(\omega) < \infty$$

the expectation of X is defined to be:

$$E[X] = \int_{\Omega} X(\omega) dP(\omega) = \int_{\mathbb{R}} x d\mu_X(x).$$

Assume now two random variables X and Y, independence of these two random variables is defined to be:

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

if the two conditions $E[|X|] < \infty$ and $E[|Y|] < \infty$ are met.

Another important concept related to distribution function is its density function. But first, in order to define properly what a density function is, some additional notions must stated:

Definition 2.8 *(**σ-finite measure**) Let E be some set and $\mathcal{F}$ be the corresponding σ-algebra of subsets of E. Then the positive measure μ defined on $\mathcal{F}$ is said to be finite if $\mu(E)$ is a finite real-valued number. In particular, μ is called σ-finite if E is generated by countable union of measurable sets with finite measures.*

Definition 2.9 *(**Absolute continuity**) Let $(E, \mathcal{F}, \mu)$ be a measurable space and f be a non-negative*

Borel function[9]*. If there exists another measure ν definied by:*

$$\forall A \in E, \nu(A) = \int_A f \, d\mu$$

such that:

$$\mu(A) = 0 \Rightarrow \nu(A) = 0,$$

then ν is said to be absolutely continuous with respect to μ, and is denoted by $\nu \ll \mu$.

If two measures μ and ν are mutually absolutely continuous to each other, they are called equivalent measures.

With the above notions at hand, the Radon-Nikodym theorem can be defined, which is the fundamental theorem in defining density functions.

Theorem 2.2 *(Radon-Nikodym theorem) Let $(E, \mathcal{F})$ be a measurable space and μ be a σ-finite measure on $(E, \mathcal{F})$. If another measure ν defined over $(E, \mathcal{F})$ has the relationship $\nu \ll \mu$, then there exists a measurable function $f : E \rightarrow [0, +\infty)$ such that for any subset $A \in E$:*

$$\nu(A) = \int_A f \, d\mu.$$

The function f is called Radon-Nikodym derivative and denoted by $\frac{d\nu}{d\mu}$.

If $f \geq 0$ and $\int f \, d\mu = 1$ for almost everythere (a.e.) μ, then ν is a probability measure and f is called its probability density function (PDF) with respect to μ.

2.2.2 Characteristic Functions

For a real-valued random variable, instead of its probability density function and cumulative distribution function, an alternative way to defining its distribution is through characteristic function, which is simply the inverse Fourier transform of its density function. The characteristic function of a distribution always exists, independent of the existence of density function.

---

[9]It's just a measurable function connecting two Borel spaces.

Since topic in this thesis is mostly concentrated on real-valued random variables, in the sequel most definitions and concepts are only given to real-valued random variables, i.e. random variables in the real line $\mathbb{R}$.

Definition 2.10 *(Characteristic function) Assume X is a real-valued random variable, then its characteristic function is defined to be:*

$$\phi_X(u) = E[e^{iuX}] = \int_{\mathbb{R}} e^{iux} d\mu_X(x) = \int_{-\infty}^{+\infty} e^{iux} f(x) dx, \tag{2.1}$$

where $u \in \mathbb{R}$ and i is an imaginary number.

The characteristic function completely defines distribution of a random variable. Thus, if two random variables have the same characteristic functions, they are said to be identically distributed.

Indenpendence of random variables can also be defined via characteristic functions. Assume two random variables X and Y. Then X and Y are said to be independent if and only if:

$$\phi_{X+Y}(u) = \phi_X(u)\phi_Y(u). \tag{2.2}$$

An important concept closely relating to characteristic function is the moment-generating function. Its definition is given to be:

Definition 2.11 *(Moment-generating function) Assume a real-valued random variable X and the existence of its expectation, then $\forall t \in \mathbb{R}$ the moment-generating function of X is defined to be:*

$$M_X(t) = E[e^{tX}].$$

Many important properties of a distribution can be characterized by its moment-generating function. For example, expectation and variance of a real-valued random variable X can be derived through its moments:

$$\mu_X = E[X] = M_X'(0)$$

$$\sigma_X^2 = V[X] = M_X''(0) - [M_X'(0)]^2.$$

If the moment-generating function of a random variable X exists, then the relationship between monent-generating function and characteristic function of X is:

$$\phi(-it) = M_X(t).$$

2.2.3   Stochastic Processes and Poisson Process

Stochastic processes are a collection of random variables evolving over time. As opposed to their deterministic counterparts, where future envolutions have already been set fixed by definition, there are infinitely many possibilities of the developments of stochastic processes over time. Hence, the art of studying stochastic processes is like falling down into a big jungel and finding way out of this gigantic forrest maze. In order to accomplish this, each step must be prudently and properly chosen. Likewise, in the jungel of stochastic processes, each concept and theory also must be defined and stated precisely.

Definition 2.12 *(Stochastic process) Let $(\Omega, \mathcal{F}, P)$ be a probability space and T be an arbitary set. If the collection of random variables $X = \{X_t : t \in T\}$ is defined on $(\Omega, \mathcal{F}, P)$, then $(X_t)_{t \geq 0}$ is called a stochastic process with $t \in T$.*

T will usually be considered as time. The time values T of a stochastic process can either be continuous or discrete, but in this thesis time will always be considered as continuous. Similarly, sample paths of stochastic processes can either be continuous or discrete as well, but this thesis mainly focuses on discontinuous sample pathes. Continuous time stochastic processes with discontinuous sample paths are called jump processes.

In order to understand theories behind continuous time jump processes, the notion of càdlàg function is crucial, which is simply a right-continuous function with left limits. Particularly, if t_0 is a jump time point of the process f, then the size of the jump at t_0 is defined to be $\triangle f(t_0) = f(t_0) - f(t_0-)$, where $f(t_0-) = \lim\limits_{t \to t_0, t < t_0} f(t)$. Thus, a càdlàg stochastic process is just a process where jumps only happen in future time and are not predictable.

The unpredictedness and randomness properties of stochastic processes raise the question of how the information about a specific stochastic process will be revealed as time goes by, as intuitively more details about the process will be known after that longer time has passed. More-

over, some quantities of the process might alter as time changes. To answer this question the concept of filtration is important, since it describes the way how available information about stochastic processes will be disclosed as time moves:

Definition 2.13 *(Filtration) Let* $(\Omega, \mathcal{F}, P)$ *be a probability space, where $\mathcal{F}$ is the corresponding σ-algebra. A filtration defined on $(\Omega, \mathcal{F}, P)$ is collection of increasing subsets $(\mathcal{F}_t)_{t>0}$ of $\mathcal{F}$ such that for $s < t$, $\mathcal{F}_s \subset \mathcal{F}_t$.*

Many important concepts like martingales, Brownian motions, Markov chains etc. are heavily relying on filtration. But a deeper insight into the notion of filtration is adaptedness:

Definition 2.14 *(Adaptedness) Let $(X_t)_{t\geq 0}$ be a stochastic process defined on the probability space* $(\Omega, \mathcal{F}, P)$*, then $(X_t)_{t>0}$ is said to be adapted to filtration $(\mathcal{F}_t)_{t>0}$ if X_t is $\mathcal{F}_t$-measurable for all $t \geq 0$.*

At each fixed time t, $\mathcal{F}_t$ represents simply the information available at time t, or the events that have already occured up to time t. Thus, $(\mathcal{F}_t)_{t\geq 0}$ are the sets of all information available at time t.

After defining what stochastic processes are and introducing their properties, it is time to look at some specific examples of stochastic processes. The two going to be mentioned here are Brownian motion and Poisson process. They are classical, representitive and useful for later studies.

First is Brownian motion. Brownian motions are continuous time stochastic processes.

Definition 2.15 *(Brownian motion) A real-valued stochatic process $(B(t))_{t\geq 0}$ is called a Brownian motion if the following requirements are satisfied:*

1. ***Independent increments****: For increasing times $0 = t_0 < t_1 < ... < t_n = T$, the increments $B(t_n) - B(t_{n-1}), B(t_{n-1}) - B(t_{n-2}), ..., B(t_1) - B(t_0)$ are independent random variables.*

2. ***Stationary increments****: For $t, s \geq 0$, the increment $B(t + s) - B(t)$ only depends on s in distribution.*

3. ***Normal distribution****: For $t, s \geq 0$, the increment $B(t + s) - B(t)$ is normally distributed with mean 0 and variance s.*

One important feature of Brownian motions is its almost surely continuous but no where differentiable sample paths. That is, no jumps are permitted among Brownian motions. Also, Brownian motions are scale invariant, which means that no matter how one zooms in or out, the pattern of a Brownian motion remains the same. This again confirms the the fact Brownian motions are always continuous and no jumps can be found.

Another important example of stochastic processes is Poisson process, which is a continuous-time stochastic process with discontinuous trajectories. A Poisson process counts the number of events in a given time interval, and marks the time points where these events happen. But prior to introducing Poisson process, the distribution of Poisson must be clear:

Definition 2.16 *(Poisson distribution) A discrete random variable X is said to be Poisson with parameter λ if for each $x = 0, 1, 2, ...$, the distribution of X satisfies:*

$$P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!}.$$

The mean and variance of a Poisson distributed random variable X are both equal to λ. Poisson distribution is used to express the probability of a given number of events happening in a time interval, where the time interval these events occur is independent of the rest of time intervals where other events occur.

Now it is time to introduce the more important concept Poisson process:

Definition 2.17 *(Poisson process) The continuous-time couning process $\{N(t) : t \geq 0\}$ is called a Poisson process if the following requirements are satisfied:*

1. ***Starting at origin****: $N(0) = 0$.*

2. ***Independent increments****: For $0 = t_0 < t_1 < ... < t_n = T$, the increments $N(t_n) - N(t_{n-1}), N(t_{n-1}) - N(t_{n-2}), ..., N(t_1) - N(t_0)$ are independent random variables.*

3. ***Stationary increments****: For $t, s \geq 0$, the increment $N(t + s) - N(t)$ only depends on s in distribution.*

4. ***Poisson distributed****: $N(t)$ is Poisson distributed with parameter λt.*

5. ***Non-overlapping occuring times***: *No more than one event can occur in an small neighbour of an occuring time.*

One important feature of Poisson processes is their independent exponential distributed inter-arrival times. That is, the time between each consecutive events is exponential distributed with parameter λ, and is independent of the rest of the other inter-arrival times. Another important feature of Poisson processes is that within a time interval, the occurences of events are uniformly distributed.

If the parameter λ of a Poisson process is constant over time, then the Poisson process is called a homogeneous Poison process. On the other hand, if the parameter varies over time and is therefore denoted by $\lambda(t)$, then the Poisson process is called inhomogeneous. A homogeneous Poisson process is a classical example of Lévy processes.

An expansion of the Poisson process is its centered version, the so-called compensated Poisson process: $\tilde{N}(t) = N(t) - t\mu$. The compensated Poisson process is going to be important in the next subsection when it comes to defining Lévy-Itô decomposition.

A Poisson process almost always jumps at 1. Sometimes it will be practical to have another type of jump sizes, preferably modelled by a common distribution. Thus, an important generalization of Poisson processes is compound Poisson process, having the following form:

$$X(t) = \sum_{i=1}^{N(t)} Y_i.$$

Here $N(t)$ is a Poisson process with intensity λ, and $\{Y_i\}_{i\geq1}$ is a sequence of independent random variables with commen distribution function G and intensity funtion g. In such a process, the inter-arrival times of jumps are still exponential distributed, but the jump sizes are now described by the distribution G. Trajectories of compound Poisson processes are piecewise constant.

Compound Poisson process has independent and stationary increments, but tits density function at time t is unknown. However, characteristic function of a compound Poisson process can be written as:

$$E\left[e^{iuX(t)}\right] = e^{t\lambda \int_{\mathcal{R}}(e^{iux}-1)f(dx)}.$$

2.3 Lévy Processes

2.3.1 Definitions of Lévy Processes

Lévy processes[10] are one of the most important stochastic processes. It includes a broad spectrum of stochastic processes that are in active use in both engineering and finance. Yet definition of Lévy processes is straightforward and easy to understand:

Definition 2.18 *(Lévy process) Let $\{X_t : t \geq 0\}$ be a stochastic process on $\mathbb{R}^d$. Then X_t is called a Lévy process if the following requirements are met:*

1. *$\mathbf{Starting\ at\ origin}$: $X_0 = 0$ a.e.*

2. *$\mathbf{Independent\ increments}$: For $0 = t_0 < t_1 < ... < t_n = T$, the increments $X_{t_n} - X_{t_{n-1}}, X_{t_{n-1}} - X_{t_{n-2}}, ..., X_{t_1} - X_{t_0}$ are independent random variables.*

3. *$\mathbf{Stationary\ increments}$: For $t, s \geq 0$, the increment $X_{t+s} - X_t$ only depends on s.*

4. *$\mathbf{Continuity\ in\ probability}$: For all $t \geq 0$ and $\epsilon > 0$ a small number, $\lim_{s \to 0} P(|X_{t+s} - X_t| > \epsilon) = 0$.*

One of the important properties of Lévy processes is the càdlàg feature. That is, for $t \geq 0$, X_t is right continuous and has left limits when $t > 0$.

As mentioned before, Brownian motion and Poisson process are two classical examples of Lévy processes. But notice here that sample paths of Brownian motion with drift are continuous while sample paths of Poisson process are discrete. Actually, Brownian motion with drift is the only (non-deterministic) Lévy process that has continuous sample paths. The rest of Lévy processes all have discontinuous sample paths.

Another important property of Lévy processes is their infinitely divisible property, which states that an infinitely divisible probability distribution can be expressed as a sum of an arbitrary number of independent and identically distributed (i.i.d.) random variables, i.e.:

Definition 2.19 *(Infinite divisibility) Assume a random variable $X : \mathbb{R}^d \to \mathbb{R}$. Then X is called infinitely divisible if and only if:*

$$X \stackrel{d}{=} X_1 + \cdots + X_n$$

---

[10] Named after the French mathematician Paul Lévy.

for all $n \geq 2$, where X_i's are indenpendent and identically distributed random variables.

There exists a close relationship between infinite divisibility and Lévy processes. In fact, if two processes are considered the same only when they have identical distribution functions, then there exists an one-to-one correspondence between the collection of all infinitely divisible distributions and the collection of all Lévy processes [Sato, 1999].

Recall (2.1) of subsection 2.2.2 that the characteristic function of the Lévy process X_t is given by $\phi_{X_t}(u) = E[e^{iuX_t}]$. But more specifically the characteristic functions of Lévy processes are in exponential form. This is generalized in the following proposition, with proof given:

Proposition 2.1 *(Characteristic function of Lévy processes) Assume $(X_t)_{t \geq 0}$ is a Lévy process on $\mathbb{R}^d$. Then there exists a continuous function $\psi : \mathbb{R}^d \to \mathbb{R}$, called the characteristic exponent of X_t, such that the characteristic function of X_t can be written as:*

$$\phi_{X_t}(u) = E[e^{iuX_t}] = e^{t\psi(u)}, u \in \mathbb{R}^d.$$

In order to prove the proposition, the following lemma is required:

Lemma 2.1 *(Convergence of stochastic processes) A stochastic process $(X_t)_{t \geq 0} \to X$ as $n \to \infty$ if and only if $\phi_{X_n}(u) \to \phi_X(u)$ as $n \to \infty$, for all $u \in \mathbb{R}^d$.*

Proof: **(Proposition 3.1)** First, notice that $X_{t+s} = X_s + X_{t+s} - X_s$. By the stationary property of Lévy processes it is clear that $X_{t+s} - X_s$ is independent of X_s. Then, recall (2.2), the characteristic function of X_{t+s} can be written as:

$$\phi_{X_{t+s}}(u) = \phi_{X_s + X_{t+s} - X_s}(u) = \phi_{X_s}(u)\phi_{X_{t+s} - X_s}(u) \tag{2.3}$$

Again, reapplying the stationary property of Lévy processes results in $X_{t+s} - X_s \stackrel{d}{=} X_t$, such that:

$$\phi_{X_{t+s} - X_s}(u) = \phi_{X_t}(u) \tag{2.4}$$

Substituting (2.4) into (2.3) results in:

$$\phi_{X_{t+s}}(u) = \phi_{X_s}(u)\phi_{X_t}(u) \tag{2.5}$$

Thus, the characteristic function of Lévy processes is of multiplicative form. According to the continuity in probability property of Lévy processes, $X_s \xrightarrow{d} X_t$ as $s \to t$. But Lemma 3.1 implies simply that $\phi_{X_s}(u) \to \phi_{X_t}(u)$ as $s \to t$, demonstrating that $\phi_{X_t}(u)$ must be continuous. (2.5) together with the continuity property entail that $\phi_{X_t}(u)$ is an exponential function. $\square$

2.3.2   Lévy-Khintchine Formula and Lévy-Itô Decomposition

In the last subsection, it was proven that the characteristic function of a Lévy process is of exponential form. But could it be more specific? Since a Lévy process is fully determined by its characteristic function, it is of great necessity to give a more comprehensive picture of characteristic functions of Lévy processes. But first, an important concept:

Definition 2.20 *(**Lévy measure**) Assume $(X_t)_{t \geq 0}$ is a Lévy process on $\mathbb{R}^d$ and $A \in \mathcal{B}(\mathbb{R}^d)$. The measure v on $\mathbb{R}^d$ defined by:*

$$v(A) = E[\#\{t \in [0,1] : \Delta X_t \neq 0, \Delta X_t \in A\}]$$

is called the Lévy measure of X_t. $v(A)$ represents the expected number of jumps per unit time whose size belongs to A.

With Lévy measure at hand, the Lévy-Khintchine formula is given as:

Theorem 2.3 *(**Lévy-Khintchine formula**) Let $(X_t)_{t \geq 0}$ be a Lévy process on $\mathbb{R}^d$. Then, there exists a unique function $\psi : \mathbb{R}^d \to \mathbb{C}$ such that the chatacteristic function of the process can be written as:*

$$\phi_{X_t}(u) = E[e^{iuX_t}] = e^{t\psi(u)}, u \in \mathbb{R}^d, t \geq 0 \tag{2.6}$$

where $\psi(u)$ is:

$$\psi(u) = iub - \frac{1}{2}u^t \Sigma u + \int_{\mathbb{R}^d} \left(e^{iux} - 1 - \frac{iux}{1 + ||x||} \right) dv(x). \tag{2.7}$$

In particular, Σ, b and v satisfy the following conditions:

1. *$\Sigma \in \mathbb{R}^{d^2}$ and is a positive semidefinite matrix.*

2. *$b \in \mathbb{R}^d$.*

3. *v is a Lévy measure on $\mathbb{R}^d$ such that $\int_{\mathbb{R}^d} ||x||^2 \wedge 1 dv(x) < \infty$.*

(Σ, b, ν) is called a Lévy triple, and completely determines the distribution of the Lévy process X_t. More specifically, Σ is a parameter describing the covariance structure of the Brownian motion, b is a drift, and ν is the amplitude of the jumps of the process. The reason why such a Lévy process can be categorized into these three components will be made clear in the Lévy-Itô decomposition theorem that is going to be stated below.

One important result derived from the Lévy-Khintchine formula is that when there exists a Lévy triple satisfying the three conditions above, then there must exist a Lévy process satisfying (2.6) and (2.7).

Lévy-Khintchine formula can be easily derived from the so-call Lévy-Itô decomposition, which is the next topic of this subsection.

Review subsection 2.2.3, if process $\{X_t : t \geq 0\}$ is càdlàg, then its jumps are represented as $\triangle X_t = X_t - X_{t-}$. Based on the jumps of a càdlàg process X_t, the Poisson random measure can be defined as:

$$N(t, A) = \#\{\triangle X(s) \in A : s \in [0, t]\}.$$

If A is bounded below[11], then the process $\{N(t, A) : t \geq 0\}$ is a Process process with intensity given by $\mu(A) = E[N(1, A)]$.If A is bounded below, then $\mu(A) < \infty$, which implies that μ is σ-finite, and recall subsection 2.2.3, $\tilde{N}(t, A) = N(t, A) - t\mu(A)$ is a compensated Poisson process.

Now the Lévy-Itô decomposition can be stated as:

Theorem 2.4 *(Lévy-Itô decomposition) Let $\{X_t : t \geq 0\}$ be a Lévy process. Then there exists a $b \in \mathbb{R}^d$, a Brownian motion with W_A with covariance matrix $A \in \mathbb{R}^d$, and an independent Poisson random measure N on $\mathbb{R}^+ \times (\mathbb{R}^d \setminus \{0\})$ such that for each $t \geq 0$,*

$$X(t) = bt + W_A(t) + \int_{|x|<1} x\tilde{N}(t, dx) + \int_{|x|\geq 1} xN(t, dx). \tag{2.8}$$

In analogy to Lévy-Khintchine formula, the Lévy-Itô decomposition confirms that a Lévy process is completely determined by its Lévy triplet, which is (b, A, μ) in this case.

(2.8) shows that every Lévy process can be divided into three components. The first component is $X^1(t) = bt + W_A(t)$, which is a Brownian motion with drift. b is the drift and A is the covariance matrix of the Brownian motion. $X^1(t)$ represents the absolutely continuous part of

---

[11]A set is said to be bounded below if it has lower bound.

the process, capturing the drift b that draws the process up or down. The second component is

$$X^2(t) = \int_{|x| \geq 1} x N(t, dx) = \sum_{0 \leq s \leq t}^{|\triangle X_s| \geq 1} \triangle X_s,$$

which is a compound Poisson process. $X^2(t)$ represents the discontinuous part of the process, and only covers points whose absolute values are greater or equal to 1. $X^2(t)$ contains finite number of terms a.e., i.e. there are limited number of huge jumps a.e.. However, when it comes to small jumps, that is, jumps with amplitude between 0 and 1, the same result is not guaranteed. There could be infinitely many small jumps during the execution of the process, and their sum does not necassarily converge [Tankov and Cont, 2003]. Hence, the compensated version of the jump integral has to be applied for the part dealing with small jumps. This is exactly the third part, i.e. $X^3(t) = \int_{|x| < 1} x \tilde{N}(t, dx)$. Since the sum has been centered, on a finite interval the number of jumps is countable a.e..

The Lévy-Khintchine formula can be easily derived from the Lévy-Itô decomposition by independence of the decomposition and characteristic function of Lévy processes.

2.4 Itô Calculus

2.4.1 Itô's Lemma

Stochastic differential equations are building blocks of mathematical finance, especially when it comes to option pricing theory, where Itô calculus[12] plays a central role. Itô calsulus extends the Riemann-Stieltjes integral to stochastic processes, where the integrator is a non-differentiable stochatic process. Such an integral is called Itô integral. Its most important feature is that both the integrand and the integrator are stochastic processes (the integrator is often a Brownian motion), and the integral itself is another stochatic process:

$$Y(t) = \int_0^t X(s)dB(s). \tag{2.9}$$

(2.9) is called Itô integral of $X(t)$ with respect to Brownian motion $B(t)$.

The Itô integral is essential in mathematical finance since $X(s)$ can be used to represent how much stock a shareholder has at time s, and $B(s)$ represents fluctuations of the stock at time s. The integtal of $X(s)$ with respect to $B(s)$ from 0 to t then indicates how much money the shareholder has between the time interval $[0, t]$.

As stated in the previous sections, sample paths of Brownian motions are no where differentiable and their variations over every time interval are infinite. Those features contradict requirements of Riemann-Stieltjes integrals, such that integrals like (2.9) cannot be solved using standard integration methods. However, in order to solve such a mysterious integral, the integrand $X(s)$ has to be adapted, i.e. its values at time t can only depend on information available up until time t. This restriction has practical meanings in mathematical finance context, stating that stock/trading strategies $X(t)$ can only depend on information up until time t, such that it is impossible to gain money from knowiing what is going to happen in the future.

Since standard integration methods are on longer available here, what is the alternative method specific for finding solutions of Itô integrals? The answer is Itô's lemma or Itô's formula. Itô's formula is a fundamental tool in modern stochatic analysis which is capable of describing evolution of stochastic processes over time, and thus can be used to assign values to prices of

---

[12]Named after the Japanese mathematician Kiyoshi Itô.

financial derivatives like options. The formula is an extension of the classical chain rule method.

To avoid mathematical difficulties, only the 1-dimensional Itô formula will be presented here. A thorough introduction to stochastic analysis can be found in [Øksendal, 2010].

Definition 2.21 *(Itô processes) Let $B(t)$ be a Brownian motion defined on the probability space $(\Omega, \mathcal{F}, P)$. An Itô process is an adapted stochastic process defined on $(\Omega, \mathcal{F}, P)$ which can be formulated as:*

$$X(t) = X(0) + \int_0^t \sigma(s) dB(s) + \int_0^t \mu(s) ds$$

i.e. it is a sum of an integral with respect to a Brownian motion and an integral with respect to time. The condition which guarantees that $\sigma(t)$ is predictable and $B(t)$-integrable and $\mu(t)$ is predictable and Lebesgue-integrable is:

$$\int_0^t (\sigma^2(s) + |\mu(s)|) ds < 0.$$

In the rest of the thesis, Itô processes are often written as the differential form:

$$dX(t) = \sigma(t) dB(t) + \mu(t) dt.$$

Definition 2.22 *(Itô formula) Assume $X(t)$ is an Itô process having the form:*

$$dX(t) = \sigma(t) dB(t) + \mu(t) dt$$

and $f(t, x)$ is a twice continuously differentiable function defined on $[0, +\infty) \times \mathbb{R}$. Then, one has:

$$df(t, X(t)) = \frac{\partial f}{\partial t}(t, X(t)) dt + \frac{\partial f}{\partial x}(t, X(t)) dX(t) + \frac{1}{2} \frac{\partial^2 f}{\partial x^2}(t, X(t))(dX(t))^2 \tag{2.10}$$

According to the form of the Itô process and the following rules:

$$dt dt = dt dB(t) = dB(t) dt = 0, dB(t) dB(t) = dt,$$

it is straighforward to see that the Itô formula can be simplified as:

$$df(t, X(t)) = \left(\frac{\partial f}{\partial t}(t, X(t)) + \mu(t)\frac{\partial f}{\partial x}(t, X(t)) + \frac{1}{2}\sigma^2(t)\frac{\partial^2 f}{\partial x^2}(t, X(t))\right)dt + \sigma(t)\frac{\partial f}{\partial x}(t, X(t))dB(t).$$

(2.11)

(2.10) and (2.11) indicate that the Itô formula is again an Itô process.

An important example of the applications of Itô formula is geometric Brownian motion. Assume now that an Itô process is of the form $dS(t) = S(t)(\sigma dB(t) + \mu dt)$, and the corresponding twice continuously differentiable function f is of the form $f(S(t)) = logS(t)$. Then, applying Itô formula on $f(S(t))$ results in:

$$\begin{aligned} df(S(t)) &= dlogS(t) \\ &= \frac{\partial f}{\partial s}(t, S(t))dS(t) + \frac{1}{2}\frac{\partial^2 f}{\partial s^2}(t, S(t))S(t)^2\sigma^2 dt \\ &= \frac{1}{S(t)}\left(\sigma S(t)dB(t) + \mu S(t)dt\right) - \frac{1}{2}\sigma^2 dt \\ &= \sigma dB(t) + \left(\mu - \frac{1}{2}\sigma^2\right)dt. \end{aligned}$$

Integrating the final equation from 0 to t, one has:

$$logS(t) - logS(0) = \left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma B(t).$$

Moving $logS(0)$ to the right hand side and taking exponential gives the following result:

$$S(t) = S(0)e^{\left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma B(t)}$$

(2.12)

(2.12) is also called lognormal process. It is the dynamical model of a stock in the celebrated Black-Scholes model, which is the topic of next subsection.

2.4.2 Black-Schloes Option Pricing Model

One of the most remarkable contributions to modern finance theory is the Black-Scholes option pricing formula, due to Fischer Black and Myron Scholes in their 1973 work. The formula is the cornerstone of modern option trading activities around the world. The core content of the for-

mula is that price of a European type option over time is governed by a partial differential equation. Solving the equation gives the so-called Black-Scholes prices of the option evolving over time. Various tests have showed the these prices are fairly close to the actual prices of an option sold in the market. The model has a parameter called volalitity, which can be calibrated according to the actual prices spread of an option traded in the market. However, the model assumes that this volatility parameter is constant, but in reality volatilities of option prices vary over time. Such drawback produces the "volatility smile" problem, and is a popular research topic in nowadays quantitative finance. Numerous corrections and alternative models have been proposed to address this problem, among which is the stochastic volatility model.

Another important assumption of Black-Scholes model is the the returns of the underlying stock of the option are normally distributed [Black and Scholes, 1973]. Hence Brownian motion is used in the formula.

Black-Scholes model is mainly applied in option pricing area. An option is a financial contract that gives the holder the right to sell or buy a certain amount of stocks, commodities, bonds, exchange rates or energies. Two classical examples are European call and European put options. For a European call option, its mathematical form can be formulated as $max(0, S(T) - K)$. Here $S(t)$ denotes the price of the underlying asset at time t, and K is called strike price. The function of the call option is that if one buys the option on a specific product at time 0, then he/she becomes an option-holder and has the right to buy the product at price K, in future time T. If the price of the product at time T is more than the strike price K, the holder exercises the option and buys the product at price K, and gains $S(T) - K$ in profit. On the other hand, if the price is less than K at time T, the holder will simply go to the market and buy the product directly, without exercising the option. In this scenario the gain is zero. Therefore, whether exercising the option depends on the product price at time T, and hence the form $max(0, S(T) - K)$. For a European call option, its mathematical form is formulated as $max(0, K - S(T))$. The option-holder has the right to sell a specific product in the future time T. If the product price at time T is more than K, the holder will simply go to the market and sell his/her product directly, without exercising the option. In this scenario, the profit gain is zero. If on the other hand the product price at time T is less than K, then the holder will exercise the option and sell the product at price K, and gain $K - S(T)$ in profit. Thus, the mathematical

representation $max(0, K - S(T))$ is straightforward.

Apart from European type options, there are also many other types of options selling in the market. But no matter which type the option is, both the buyer and the seller of the option contract have to agree on the price of the option contract at time 0. Since $S(T)$ is unknown at time 0, it is impossible to find a fixed discounted price of the underlying asset at time 0. But in order to find a reasonable price of the contract at time 0 such that both sides are satisfied one has to use a systematic way to approximate the discounted price the option back at time 0. The following is the estimated discounted price of a European call option at time 0, based on the real world probability P:

$$P_{call}(0) = e^{-rT} E[max(0, S(T) - K)]. \tag{2.13}$$

r here denotes annual risk-free rate of interest, and E denotes expectation, based on real world probability measure P.

In order to find a risk-neutral price, that is, no arbitrage opportunities[13], the real world probability P often has to be changed to a parallell universe Q, which is called risk-neutral probability, under which the expected return of the underlying stock will equal the risk free interest rates, thus the name "risk-neutral"[Benth, 2003].

To solve (2.13), expression of $S(T)$ must be known. As stated at the end of subsection 2.4.1, prices of the underlying asset under Black-Scholes model are governed by a geometric Brownian motion (2.12). Here μ denotes annual expected rate of return on the underlying asset (often assumed to be a stock in the BS model), and σ is the volatility.

In order for the market to be arbitrage-free, the expected value of the stock price $S(T)$ discounted at the risk-free interest rates should be a martingale. However this is not true under the actual probability measure P. To see the reason to this definition of martingales must be stated first:

Definition 2.23 *(Martingale) Let $\{M_t\}_{t\geq 0}$ be a stochastic process defined on probability space $(\Omega, \mathcal{F}, P)$. Then $\{M_t\}_{t\geq 0}$ is called a martingale with respect to filtration $\{\mathcal{F}_t\}_{t\geq 0}$ if and only if the following requirements are satisfied:*

1. M_t is $\mathcal{F}_t$-measurable for all t.

---

[13]Arbitrage opportunities mean earning profit from nothing, which is impossible in real market activities, thus must be discarded.

2. $E[|M_t|] < \infty$ *for all* t.

3. *For all* $s \leq t$, $E[M_t|\mathcal{F}_s] = M_s$.

Now, according to the end of the last subsection:

$$S(T) = S(0)e^{\left(\mu - \frac{1}{2}\sigma^2\right)T + \sigma B(T)}.$$

This is the stock price at exercise time T to be discounted back to starting time 0. If one wants the price to be discounted back to an arbitrary time $t < T$, one defines the following equation:

$$Z_t^{S(t)}(T) = S(t)e^{\left(\mu - \frac{1}{2}\sigma^2\right)(T-t) + \sigma(B(T)-B(t))},$$

where t controls the time the price to be discounted to.

Since $B(T) - B(t)$ is normally distributed with mean 0 and variance $T - t$. Using a standard normally distributed variable y, the above equation can be represented as:

$$Z_t^{S(t)}(T) \stackrel{d}{=} S(t)e^{\left(\mu - \frac{1}{2}\sigma^2\right)(T-t) + \sigma\sqrt{T-t}\,y}.$$

Then, the expected value of the stock price $S(T)$ discounted at the risk-free interest rate r under the actual probability measure back to time $t < T$ is given by:

$$
\begin{aligned}
E[e^{-r(T-t)}Z_t^{S(t)}(T)|\mathcal{F}_t] &= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2} e^{-r(T-t)} S(t) e^{\left(\mu - \frac{1}{2}\sigma^2\right)(T-t) + \sigma\sqrt{T-t}\,y} \, dy \\
&= S(t) \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2 + (\mu - r)(T-t) - \frac{1}{2}\sigma^2(T-t) + \sigma\sqrt{T-t}\,y} \, dy \\
&= S(t) e^{(\mu-r)(T-t)} \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}[y^2 - 2\sigma\sqrt{T-t} + \sigma^2(T-t)]} \, dy \\
&= S(t) e^{(\mu-r)(T-t)} \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}[y - \sigma\sqrt{T-t}]^2} \, dy \\
&= S(t) e^{(\mu-r)(T-t)} \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z} \, dz \\
&= S(t) e^{(\mu-r)(T-t)}
\end{aligned}
$$

where $z = y - \sigma\sqrt{T-t}$.

Since a rational investor in the market will require that the annual expected rate of return

μ on the underlying stock $S(t)$ be greater than the annual risk-free interest rate r, one has $S(t)e^{(\mu-r)(T-t)} > S(t)$, such that $e^{-r(T-t)}Z_t^{S(t)}(T)$ is not a martingale with respect to the filtration $\mathcal{F}_t$ under P. Thus, under the actual probability measure there exists possibility that arbitrage-opportunities occur.

To eliminate arbitrage-opportunities, one must have $E[e^{-r(T-t)}Z_t^{S(t)}(T)|\mathcal{F}_t] = S_t(t)$. Now define a parallell probability measure Q, under which a new Brownian motion has the relationship relating to the old one:

$$dW(t) = dB(t) + \alpha dt.$$

The return of the underlying stock $S(t)$ is defined to be:

$$\frac{dS(t)}{S(t)} = \frac{\mu S(t)dt + \sigma S(t)dB(t)}{S(t)} = \mu dt + \sigma dB(t).$$

The above return is under probability measure P. Now switching to probability measure Q gives the new return:

$$
\begin{aligned}
\frac{dS(t)}{S(t)} &= \mu dt + \sigma dB(t) \\
&= \mu dt + \sigma(dW(t) - \alpha dt) \\
&= (\mu - \sigma\alpha)dt + \sigma dW(t)
\end{aligned}
$$

Under the new measure Q, one wants the market to be arbitrage-opportunities free. This is equivalent to requiring that the expected return of the underlying stock is equal to the risk-free interest rate, i.e.:

$$E_Q\left[\frac{dS(t)}{S(t)}\right] = E[(\mu - \sigma\alpha)dt + \sigma dW(t)] = (\mu - \sigma\alpha)dt = rdt,$$

such that:

$$\alpha = \frac{\mu - r}{\sigma}.$$

Now the market is risk-neutral, hence the new probability measure Q is called risk-neutral probability measure or equivalent martingale measure, and the corresponding dynamics of the

stock under Q becomes:

$$
\begin{aligned}
dS(t) &= \mu S(t)dt + \sigma S(t)dB(t) \\
&= \mu S(t)dt + \sigma S(t)\left(dW(t) - \frac{\mu - r}{\sigma}\right) \\
&= rS(t)dt + \sigma S(t)dW(t),
\end{aligned}
$$

or equivalently,

$$
\begin{aligned}
S(t) &= S(0)e^{\left(r - \frac{1}{2}\sigma^2\right)t + \sigma W(t)} \\
&= S(0)e^{\left(r - \frac{1}{2}\sigma^2\right)t + \sigma\sqrt{t}y}.
\end{aligned}
$$

The corresponding stock price at exercise time T to be discounted back to time $t < T$ is now:

$$
Z_t^{S(t)}(T) = S(t)e^{\left(r - \frac{1}{2}\sigma^2\right)(T - t) + \sigma\sqrt{T - t}y}. \tag{2.14}
$$

The expected value of the stock price $S(T)$ discounted at the risk-free interest rate r under the risk-neutral probability measure Q back to time $t < T$ now becomes:

$$
\begin{aligned}
E_Q[e^{-r(T-t)}Z_t^{S(t)}(T)|\mathcal{F}_t] &= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2} e^{-r(T-t)}S(t)e^{\left(r-\frac{1}{2}\sigma^2\right)(T-t)+\sigma\sqrt{T-t}y}dy \\
&= S(t)\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2+(r-r)(T-t)-\frac{1}{2}\sigma^2(T-t)+\sigma\sqrt{T-t}y}dy \\
&= S(t)\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}[y^2-2\sigma\sqrt{T-t}+\sigma^2(T-t)]}dy \\
&= S(t)\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}[y-\sigma\sqrt{T-t}]^2}dy \\
&= S(t)\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z}dz \\
&= S(t).
\end{aligned}
$$

Thus, $e^{-r(T-t)}Z_t^{S(t)}(T)$ is a martingale under Q, such that the market now does not contain any arbitrage-opportunities. In fact, Q is the only equivalent martingale measure that turns $e^{-r(T-t)}Z_t^{S(t)}(T)$ into a martingale.

The change of probability measures is based on Randon-Nikodym theorem, which was stated in subsection 2.2.1. For a more thorough information in this topic, interested readers are recommended to take a look at [Benth, 2003].

Now assume one is interested in the price of a European call option at time $t < T$, i.e.:

$$P_{call}(t) = e^{-r(T-t)} E_Q \left[max\left(0, Z_t^{S(t)}(T) - K\right) | \mathcal{F}_t \right]. \tag{2.15}$$

In order to eliminate the maximum calculator, one must have:

$$Z_t^{S(t)}(T) > K,$$

that is,

$$S(t) e^{\left(r - \frac{1}{2}\sigma^2\right)(T-t) + \sigma\sqrt{T-t}y} > K.$$

After some rearrangements, this requirement becomes:

$$y > \frac{log\frac{logS(t)}{K} + (r - \frac{1}{2}\sigma^2)(T - t)}{\sigma\sqrt{T - t}} = -d_2.$$

Thus, the expectation of (2.15) can now be calculated as:

$$
\begin{aligned}
E_Q\left[max\left(0, Z_t^{S(t)}(T) - K\right) | \mathcal{F}_t\right] &= E_Q\left[max\left(0, S(t) e^{\left(r - \frac{1}{2}\sigma^2\right)(T-t) + \sigma\sqrt{T-t}y} - K\right) | \mathcal{F}_t\right] \\
&= \int_{-d_2}^{+\infty} \left(e^{logS(t) + (r - \frac{1}{2}\sigma^2)(T-t) + \sigma\sqrt{T-t}Y} - K\right) g(y) dy \\
&= \int_{-d_2}^{+\infty} e^{logS(t) + (r - \frac{1}{2}\sigma^2)(T-t) + \sigma\sqrt{T-t}y} g(y) dy - K\int_{-d_2}^{+\infty} g(y) dy \\
&= \int_{-d_2}^{+\infty} e^{logS(t) + (r - \frac{1}{2}\sigma^2)(T-t) + \sigma\sqrt{T-t}y} g(y) dy - KP(Y < d_2) \\
&= \int_{-d_2}^{+\infty} e^{logS(t) + (r - \frac{1}{2}\sigma^2)(T-t) + \sigma\sqrt{T-t}y} g(y) dy - K\Phi(d_2) \\
&= \int_{-d_2}^{+\infty} e^{logS(t) + (r - \frac{1}{2}\sigma^2)(T-t) + \sigma\sqrt{T-t}y} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2} dy - K\Phi(d_2) \\
&= \frac{1}{\sqrt{2\pi}} S(t) e^{r(T-t)} \int_{-d_2}^{+\infty} e^{-\frac{1}{2}\sigma^2(T-t) + \sigma\sqrt{T-t}y - \frac{1}{2}y^2} dy - K\Phi(d_2) \\
&= \frac{1}{\sqrt{2\pi}} S(t) e^{r(T-t)} \int_{-d_2}^{+\infty} e^{-\frac{1}{2}[\sigma^2(T-t) - 2\sigma\sqrt{T-t}y + y^2]} dy - K\Phi(d_2)
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{1}{\sqrt{2\pi}} S(t) e^{(T-t)} \int_{-d_2}^{+\infty} e^{-\frac{1}{2}(y-\sigma\sqrt{T-t})^2} dy - K\Phi(d_2) \\
&= \frac{1}{\sqrt{2\pi}} S(t) e^{r(T-t)} \int_{-d_2-\sigma\sqrt{T-t}}^{+\infty} e^{-\frac{1}{2}z^2} dz - K\Phi(d_2) \\
&= S(t) e^{r(T-t)} \Phi(d_2 + \sigma\sqrt{T-t}) - K\Phi(d_2) \\
&= S(t) e^{r(T-t)} \Phi(d_1) - K\Phi(d_2)
\end{aligned}
$$

In the derivation above, $g(\cdot)$ is the density function of the standard normally distributed random variable Y, $\Phi(\cdot)$ is its corresponding cumulative distribution function, $Z = Y - \sigma\sqrt{T-t}$, and $d_1 = d_2 + \sigma\sqrt{T-t}$.

(2.15) now becomes:

$$
\begin{aligned}
P_{call}(t) &= e^{-r(T-t)} E_Q \left[max\left(0, Z_t^{S(t)}(T) - K\right) | \mathcal{F}_t \right] \\
&= e^{-r(T-t)} [S(t) e^{r(T-t)} \Phi(d_1) - K\Phi(d_2)] \\
&= S(t)\Phi(d_1) - e^{-r(T-t)} K\Phi(d_2)
\end{aligned}
$$

This is the celebrated Black-Schole model for European call options. For a European put option $max(0, K - S(T))$, the put-call parity gives:

$$
\begin{aligned}
P_{put}(t) &= Ke^{-r(T-t)} - S(t) + P_{call}(t) \\
&= Ke^{-r(T-t)} - S(t) + S(t)\Phi(d_1) - e^{-r(T-t)} K\Phi(d_2) \\
&= Ke^{-r(T-t)} [1 - \Phi(d_2)] - S(t)[1 - \Phi(d_1)] \\
&= Ke^{-r(T-t)} \Phi(-d_2) - S(t)\Phi(-d_1)
\end{aligned}
$$

according to the symmetry of standard normal distribution.

Chapter 3

Various Pricing Methods in Incomplete Markets

The Black-Scholes model derived in the last subsection of the previous chapter has numerous shortcomings, one of which is its completeness assumption. This chapter is devoted to revealing why such an assumption is problematic in reality, and giving a handful of alternative approches that can be applied in incmplete markets to price contingent claims.

The aim of this thesis is the application of exponential Lévy jump processes in modelling stochastic interest rates in the real market. The use of jump processes will not affect the arbitrage-free assumption in the market, but the completeness assumption now will not apply, in such a way that the market has several equivalent martingale measures available. The consequence of this is that there will be many arbitrage-free prices to choose from. A natural question in this situation is how to pick the right price for the corresponding contingent claim. The main part of this chapter presents several methods for pricing derivative securities under market incompleteness assumption, based on different criteria. Readers should bear in mind that there does not exist a universal criterion that fits best under all conditions. Therefore the selection of which method to be applied is a subjective work. In later applications Esscher transform will be chosen due to its nice properties connected to Lévy processes.

In this chapter, a thorough introduction to market incompleteness will be given in section 3.1. Then various pricing methods for contingent claims in incomplete markets will be sketched in section 3.2.

3.1 Market Incompleteness

3.1.1 Assumptions for the Black-Scholes Model

In the Black-Scholes option pricing model derived in the last subsection of the previous chapter, there have been made several important assumptions on the market structure. The first one is the arbitrage-free assumption, meaning that no one can earn money from trading in the market without taking any risks. In the derivations of the Black-Scholes price for a European call option in subsection 2.4.2, one sees that the absence of arbitrage opportunities is closely connected to the existence of an equivalent martingale measure Q. Actually, this is one of the most important theorems in mathematical finance:

Definition 3.1 *(The First Fundamental Theorem of Asset Pricing) A discrete market is arbitrage-free if and only if there exists at least one equivalent martingale measure Q corresponding to the original measure P.*

In the Black-Scholes option pricing model, there is an equivalent martingale measure Q under which the dynamics of the underlying stock can be described as:

$$dS(t) = rS(t)dt + \sigma S(t)dW(t)$$

where $dW(t)$ is a Brownian motion under Q.

It was also proven that the price dynamics of the underlying stock becomes a martingale after discounting under Q. Due to the existence of the equivalent martingale measure Q, according to the theorem above, one knows immediately that the market on which the Black-Scholes model is based in arbitrage-free.

The second important assumption concerning the Black-Scholes market is the assumption of completetness. Market completeness means any contingent claims in the market can be exactly replicated by buying stocks or bonds. In the market where Black-Scholes model is constructed, the market is complete due to the following theorem:

Definition 3.2 *(The Second Fundamental Theorem of Asset Pricing) A discrete market is complete if and only if there exists only one equivalent martingale measure Q corresponding to the original measure P.*

For complete proofs of the above two theorems, interested readers are referred to [Delbaen and Schachermayer, 1994].

In the derivation of the Black-Scholes price for a European call option, it was already shown that there exists only one equivalent martingale measure, hence the market is complete.

Combing the two theorems, a market is both arbitrage-free and complete if and only if there is a unique equivalent martingale measure available in the market turning discounted expected stock price into a martingale. If the market has multiple stocks and the underlying stock is modelled by a multi-dimensional Brownian motion, the two theorems can be related to the number of the stocks in the market and the dimension of the Brownian motion. For example, let n denote the number of stocks in a market and m denote the number of dimensions, then the market is arbitrage-free if and only if $n \geq m$, and is complete if and only if $n \leq m$. That is, when $n = m$, the market is both arbitrage-free and complete. In the Black-Scholes model, there is only one stock $S(t)$ and the Brownian motion $B(t)$ is of one dimension, i.e. $n = m$, hence the underlying market on which the model is built is both arbitrage-free and complete.

While the first assumption of the Black-Scholes model seems reasonable, the second assumption is unrealistic in operating markets across the world. More about incomepleteness of the Black-Scholes model will be discussed in next subsection.

Although Black-Scholes model has some major drawbacks in its assumptions and built-in structures, it is still widely used everyday in the financial markets. This is due to its applicability and attainability, in the sense that it is easy to fit and its calibrated prices are quite close to real world option prices. In the recent decades, numerous efforts and corrections have been made to address the shortcomings of the model. Some of them are intended to modify its constant volatility structure, others are intended to correct its Brownian noise such that is can fit jump trajectories. The model of this thesis falls into the second category, which will be made clear in next chapter.

3.1.2 Market Imperfection and Market Imcompleteness

In the last subsection, it was proven that the market under which Black-Scholes model is defined is complete. In order for the market to be complete, there are three main requirements that must be satisfied. The first one is that transactions in the market are free. That is, buying or

selling an option is cost-free, giving the possibility that market participants can trade in options as much as they want, without costing any money. The second requirement is that the market does not contain any restrictions. Restrictions like constraint on short selling is prohibited. The last requirement is that the stock dynamics is described by a geometric Brownian motion. If the dynamics deviates from geometric Brownian motion, the market will be incomplete.

If a market is complete, then every contingent claim can be replicated exactly by entering into the stock or bond market. But since contingent claims like options can be replicated exactly, their very existence is a myestry. Option holders will never need to exercise their options in a complete market. Contrary to this, option markets are huge across the globle, indicating that they are actually incomplete in reality. This is due to the fact that both the first and the second requirement are quite unrealistic in everyday life. Trading in the option market is almost always not transaction-free. Moreover, regulators will often impose constraints in the market. If the market is highly illiquid, it will also be incomplete.

The third assumption requires that returns of the underlying stock are normally distributed, and the dynamics of the stock is governed by a geometric Brownian motion. If the dynamics is governed by a jump process like Lévy or geometric Lévy processes, then the market will be incomplete.

Incompleteness is a natural phenomena in real left markets. According to the two fundamental theorems of asset pricing, if the market is incomplete, there exist many equivalent martingale measures available to be chosen from, each giving an arbitrage-free price for the contingent claim. Therefore an incomplete market is at least arbitrage-free. However, in such a situation, a natural question to ask is how to select the right price for the discounted expected payoff of a contingent claim conditioned on the information up to time t. There are many different pricing methods to be used in an incomplete market, some of which will be summerized in next section.

In the final model of the thesis, interest rate jumps will be allowed. Thus, the market will be incomplete. The dynamics of interest rate evolutions will be modelled by a geometric Lévy process, and Esscher transform will be utilized to select an equivalent martingale measure that can be used to find the corresponding arbitrage-free price.

3.2 Pricing Methods in Incomplete Markets

3.2.1 Merton's Jump-Diffusion Model

Diffusion models like the Black-Scholes introduced in the previous sections have continuous paths. In such models the possibility that stock prices move by large amounts over a short period of time is small, making them impractical for observed prices in real markets. The correction for this drawback is to look at models with jumps, also called jump-diffusion models. In a jump-diffusion model, suden jumps of stock prices can be modelled satisfactorily, without sacrificing for a unrealistically high value of volatility.

Another important reason for the use of jump-diffusion models is its incompleteness feature. Discountinuities make the market incomplete, such that options cannot be exactly replicated by buying or selling only the underlying stock.

One of the most well-known jump-diffusion models is Merton's model, proposed by Robert Merton in [Merton, 1976]. The idea is to combine a Brownian motion with drift with a compound Poisson process:

$$X(t) = \mu t + \sigma B(t) + \sum_{i=1}^{N(t)} Y_i$$

where $B(t)$ is a standard Brownian motion, $N(t)$ is a Poisson process and $\{Y_i\}_{i \geq 1}$ are normally distributed random variables independent from $B(t)$ and $N(t)$. In such a model, changes in stock prices are consisted of a normal component that is modelled by a Brownian motion with drift and a abnormal component that is modelled by a compound Poisson process. The first part models the continuous part of the price dynamics and the second part models the discontinuous part (i.e. the jumps) of the dynamics.

The underlying stock price dynamics will be modelled by:

$$S(t) = S(0)e^{X(t)} = S(0)e^{\mu t + \sigma B(t) + \sum_{i=1}^{N(t)} Y_i}.$$

$X(t)$ is again a Lévy process. Thus the stock price dynamics is actually modelled by an exponential or geometric Lévy process.

In order to derive solutions for Vanilla option prices, assume that $N(t)$ is Poisson distributed with intensity λ, $Y_i \sim N(m, \delta^2)$, and $B(t)$, $N(t)$ and Y_i are independent from each other.

Since the market now is incomplete, there are many equivalent martingale measures. To find the corresponding risk-neutral measure Q, Merton suggests to change the drift term in the Brownian motion and leave the rest components unchanged. That is, under Q, the stock dynamics can now be written as:

$$S(t) = S(0)e^{\mu^M t + \sigma W(t) + \sum_{i=1}^{N(t)} Y_i} \tag{3.1}$$

such that $e^{-rt}S(t)$ becomes a martingale.

Merton assumes that y_i are nonnegative or absolute jump sizes. In a small time interval dt, process $S(t)$ jumps to $y_i S(t)$. Thus, the relative price jump size can described by:

$$\frac{dS(t)}{S(t)} = \frac{y_i S(t) - S(t)}{S(t)} = y_i - 1.$$

Since $log\, y_i = Y_i$ are normally distributed with mean m and variance δ^2, it is straightforward to see that y_i is lognormally distributed such that $E(y_i) = e^{m + \frac{\delta^2}{2}}$ and $V(y_i) = (e^{\delta^2} - 1)e^{2m + \delta^2}$.

Then, the dynamics of $S(t)$ can be written as:

$$\frac{dS(t)}{S(t)} = (r - \lambda k)dt + \sigma dW(t) + (y_i - 1)dN(t)$$

where $k = E[e^{Y_i} - 1] = E[e^{Y_i}] - 1 = e^{m + \frac{\delta^2}{2}} - 1$.

According to Itô's formula and (3.1), $\mu^M = r - \frac{\sigma^2}{2} - \lambda k = r - \frac{\sigma^2}{2} - \lambda(e^{m + \frac{\delta^2}{2}} - 1)$, such that under the risk neutral probability measure Q, dynamics of the underlying stock $S(t)$ is:

$$S(t) = S(0)e^{\left[r - \frac{\sigma^2}{2} - \lambda(e^{m + \frac{\delta^2}{2}} - 1)\right]t + \sigma W(t) + \sum_{i=1}^{N(t)} Y_i}.$$

Now define:

$$Z_t^{S(t)}(T) = S(t)e^{\left[r - \frac{\sigma^2}{2} - \lambda(e^{m + \frac{\delta^2}{2}} - 1)\right](T-t) + \sigma(W(T) - W(t)) + \sum_{i=1}^{N(T-t)} Y_i}.$$

The equivalent martingale measure Q is obtained by only changing the drift of the Brownian motion. Both the inter-arrival times of jumps $N(t)$ and the jump sizes Y_i are unchanged. This is

due to the assumption made by Merton that jump risks can be diversified away, hence no extra risk premiums are charged due to price jumps.

Now assume $H[Z_t^{S(t)}(T)]$ is the payoff function of a contingent claim. The discounted expected risk-neutral price of the option at time $t < T$ is given by:

$$P(t) = e^{-r(T-t)} E_Q[H[Z_t^{S(t)}(T)]|\mathcal{F}_t].$$

Let $\tau = T - t$ and $y \sim N(0,1)$. Note that since $Y_i \overset{i.i.d.}{\sim} N(m, \delta^2)$, for $j = 1, 2, \ldots, N(T-t), \sum_{i=1}^{j} Y_i \sim N(jm, j\delta^2)$. Then one has:

$$
\begin{aligned}
P(t) \quad &= \quad e^{-r(T-t)} E_Q[H[Z_t^{S(t)}(T)]|\mathcal{F}_t] \\
&= \quad e^{-r(T-t)} E_Q\left[H\left[S(t)e^{\left[r-\frac{\sigma^2}{2}-\lambda(e^{m+\frac{\delta^2}{2}}-1)\right](T-t)+\sigma W(T-t)+\sum_{i=1}^{N(T-t)} Y_i}\right]|\mathcal{F}_t = S\right] \\
&= \quad e^{-r\tau} E_Q\left[H\left[Se^{\left[r-\frac{\sigma^2}{2}-\lambda(e^{m+\frac{\delta^2}{2}}-1)\right]\tau+\sigma W(\tau)+\sum_{i=1}^{N(\tau)} Y_i}\right]\right] \\
&= \quad e^{-r\tau} \sum_{j\geq 1} Q[N(\tau)=j] E_Q\left[H\left[Se^{\left[r-\frac{\sigma^2}{2}-\lambda(e^{m+\frac{\delta^2}{2}}-1)\right]\tau+\sigma W(\tau)+\sum_{i=1}^{j} Y_i}\right]\right] \\
&= \quad e^{-r\tau} \sum_{j\geq 1} \frac{e^{-\lambda\tau}(\lambda\tau)^j}{j!} E_Q\left[H\left[Se^{\left[r-\frac{\sigma^2}{2}-\lambda(e^{m+\frac{\delta^2}{2}}-1)\right]\tau+\sigma W(\tau)+\sum_{i=1}^{j} Y_i}\right]\right] \\
&= \quad e^{-r\tau} \sum_{j\geq 1} \frac{e^{-\lambda\tau}(\lambda\tau)^j}{j!} E_Q\left[H\left[Se^{\left[r-\frac{\sigma^2}{2}-\lambda(e^{m+\frac{\delta^2}{2}}-1)\right]\tau+\sigma W(\tau)+\sum_{i=1}^{j} Y_i}\right]\right] \\
&\overset{d}{=} \quad e^{-r\tau} \sum_{j\geq 1} \frac{e^{-\lambda\tau}(\lambda\tau)^j}{j!} E_Q\left[H\left[Se^{\left[r-\frac{\sigma^2}{2}-\lambda(e^{m+\frac{\delta^2}{2}}-1)\right]\tau+jm+\sqrt{\sigma^2\tau+j\delta^2}\,y}\right]\right] \\
&= \quad e^{-r\tau} \sum_{j\geq 1} \frac{e^{-\lambda\tau}(\lambda\tau)^j}{j!} E_Q\left[H\left[Se^{\left[r-\frac{\sigma^2}{2}+\frac{j\delta^2}{2\tau}-\frac{j\delta^2}{2\tau}-\lambda(e^{m+\frac{\delta^2}{2}}-1)\right]\tau+jm+\sqrt{\sigma^2\tau+j\delta^2}\,y}\right]\right] \\
&= \quad e^{-r\tau} \sum_{j\geq 1} \frac{e^{-\lambda\tau}(\lambda\tau)^j}{j!} E_Q\left[H\left[Se^{\left[r-\frac{1}{2}(\sigma^2+\frac{j\delta^2}{\tau})+\frac{j\delta^2}{2\tau}-\lambda(e^{m+\frac{\delta^2}{2}}-1)\right]\tau+jm+\sqrt{\sigma^2\tau+j\delta^2}\,y}\right]\right] \\
&\overset{\sigma_j^2=\sigma^2+\frac{j\delta^2}{\tau}}{=} \quad e^{-r\tau} \sum_{j\geq 1} \frac{e^{-\lambda\tau}(\lambda\tau)^j}{j!} E_Q\left[H\left[Se^{\left[r-\frac{1}{2}\sigma_j^2+\frac{j\delta^2}{2\tau}-\lambda(e^{m+\frac{\delta^2}{2}}-1)\right]\tau+jm+\sigma_j\sqrt{\tau}\,y}\right]\right]
\end{aligned}
$$

$$= \quad e^{-r\tau} \sum_{j\geq 1} \frac{e^{-\lambda\tau}(\lambda\tau)^j}{j!} E_Q \left[H \left[Se^{\left[\frac{j\delta^2}{2\tau} - \lambda(e^{m+\frac{\delta^2}{2}}-1)\right]\tau + jm} e^{\left[(r-\frac{1}{2}\sigma_j^2)\tau + \sigma_j\sqrt{\tau}y\right]} \right] \right]$$

$$= \quad \sum_{j\geq 1} \frac{e^{-\lambda\tau}(\lambda\tau)^j}{j!} P_{BS}\left(\tau, Se^{\left[\frac{j\delta^2}{2\tau} - \lambda(e^{m+\frac{\delta^2}{2}}-1)\right]\tau + jm}, \sigma_j \right)$$

where $P_{BS}(T-t, S, \sigma) = e^{-r(T-t)} E_{Q_{BS}}\left[H\left(Se^{(r-\frac{1}{2}\sigma^2)(T-t)+\sigma(W(T)-W(t))}\right) | \mathcal{F}_t = S \right]$ denotes the discounted expected Black-Scholes price of payoff H at time t.

Thus, the discounted expected risk-neutral price of a contingent claim at time t up to information $\mathcal{F}_t$ under the Merton's model can be expressed as weighted sums of Black-Scholes prices. For other solutions and simulation techniques about Merton's method, interested readers are referred to [Matsuda, 2004] and [Burger and Kliaras, 2013].

3.2.2 Superhedging

In a complete market, every contingent claim H is attainable. That is, every H can be replicated by a self-financing trading strategy. A trading strategy is a dynamic portfolio represented by a pair $(x, \theta) = (x, \theta(t))_{0 \leq y \leq T}$ where x is the initial value invested in the underlying asset and $\theta(t)$ is the number of the asset at time t. Then, at any time t, the value of the portfolio is given as $V_t(x, \theta) = x + \int_0^t \theta(u) dS(u)$. A strategy is called self-financing if and only if any fluctuations of the portfolio are subject to x and θ, meaning that there is no extra deposit or withdrawal of money during the whole course of the portfolio. More precisely, for a self-financing trading strategy the initial investment in the underlying asset is equal to the value of the portfolio at time t: $V_t(V_0, \theta) = V_0 + \int_0^t \theta(u) dS(u)$. In the following if a portfolio is self-financing then denote its initial value is given by $x = V_0 = v$ and will be denoted as (v, θ). In a complete market, for every contingent claim H there will be a self-financing strategy (v, θ) such that $H = V_T(v, \theta)$ at maturity time $TP - a.s..$ Uner such circumstances, the price of the H at time 0 must be equal to v. This is exactly the case for the celebrated Black-Scholes pricing model given in 2.4.2, where $S(t)$ is formulated as a geometric Brownian motion. Then θ is the called the hedging strategy of the model.

However, in an incomplete market, contingent claims are not attainable. In order to replicate H, one can resort to a method called superhegding. Let V_T be an admissible self-financing

trading strategy such that its final capital at exercise time T is:

$$V_T(v,\theta) = v + \int_0^T \theta(u)dS(u) \geq H.$$

$V_T(v,\theta)$ here is called a superhedging strategy for the contingent claim H. The superhedging method guarantees that the replicating payoff at time T is at least the contingent claim H.

If $\mathcal{A}$ denotes the set of all admissable strategies, then the superhedging price $\Pi(H)$ is defined to be:

$$\Pi(H) = inf\left\{v \in \mathcal{R}, \exists \theta \in \mathcal{A}, V_T(v,\theta) = v + \int_0^T \theta(u)dS(u) \geq H a.s.\right\}.$$

The superhedging price is actually the upper bound of the selling price of the contingent claim H since it is defined to replicate H with 100% safety. If the seller of H is willing to take on some risk he or she can just release a little bit the requirement and accept a price lower than the superhedging price. According to [Tankov and Cont, 2003] the lower bound of the selling price is given by $-\Pi(H)$ due to nonlinearity of $\Pi(H)$. All selling prices of H must therefore lie inside the interval $[-\Pi(-H),\Pi(H)]$.

Calculations of the superhedging price $\Pi(H)$ and the corresponding strategy $V_T(v,\theta)$ can be summerized in the following theorem:

Theorem 3.1 *(Super-Hedging Theorem) Let $\mathcal{Q}$ denote the set of all equivalent martingale measures. The the superhedging price $\Pi(H)$ can be calculated as:*

$$\Pi(H) = \sup_{Q \in \mathcal{Q}} E_Q[e^{-rT}H].$$

On the other hand, the self-financing strategy $V_t(x,\theta)$ at time $t < T$ can be calculated as:

$$V_t(v,\theta) = ess\sup_{Q \in \mathcal{Q}} E_Q\left[e^{-r(T-t)}H|\mathcal{F}_t\right]$$

where ess sup denotes essential supremum.

In most cases the superhedging price and the corresponding self-financing strategy are diffiult to calculate. Moreover, the superhedging price is too high, making it impractical in reality. For more information about superhedging, special interested readers are referred to [Karoui

and Quenez, 1997] and [Kramkov, 1996].

3.2.3 Utility Maximization

The superhedging pricing approach is both impractical and unrealistic. One of the many reasons to this is that it gives equal weighting to each scenario having non-zero probability [Tankov and Cont, 2003]. Thus, an alternative approach to this problem is to look at the maximum of expected utility function, derived from classical economic theories:

$$\max_{Z} E[U(Z)] \tag{3.2}$$

where U is called the utility function and Z is the payoff function.

Maximizing utility is also called indifference pricing. In the option pricing context, indifference price is the price that the expected utility of an agent will be indifferent between exercising his option and not exercising.

As before, assume $V_T(v, \theta) = v + \int_0^T \theta(u) dS(u)$ is the final capital of a self-financing trading strategy θ at maturity time T, and define in addition that $\mathcal{S}$ is the set of all self-financing trading strategies at time T, H is the payoff of a contingent claim at T, U is a specific utility function. Then the indifference pricing problem can be formulated mathematically as:

$$
\begin{aligned}
V(v, H) &= \sup_{\theta \in \mathcal{S}} E[U(V_T + H)] \\
&= \sup_{\theta \in \mathcal{S}} E\left[U\left(v + \int_0^T \theta(u) dS(u) + H \right) \right].
\end{aligned}
$$

The indifference buying price is then defined to be the price $\pi_{buy}(H)$ such that:

$$V[v - \pi_{buy}(H), H] = V(v, 0),$$

and the indifference selling price $\pi_{sell}(H)$ is defined to be:

$$V(v, 0) = V[v + \pi_{sell}(H), -H].$$

[Carmona, 2009] states that all indifference prices will lie inside the interval $[\pi_{buy}(H), \pi_{sell}(H)]$.

By [Tankov and Cont, 2003] one major drawback of using the indifferent pricing method is the choice of initial capital. For two participants with same utility functions, the final derived prices of both sides will be different provided that they have different initial capitals. To overcome this shortcoming the utility function must be chosen carefully. The following is an example of using an exponential utility function, which makes both sides are indifferent about how much initial capitals they have prior to purchasing the option:

$$U(x) = 1 - e^{-ax}, a \in \mathcal{R}.$$

When $a \to \infty$, the indifference price converges to the superhedging price.

However, as behavior of the indifference price depends on a, the choice of this parameter now is a major problem. This also raises the question about whether this pricing method is robust.

3.2.4 Quadratic Hedging

As already been noted, in an incomplete market it is impossible to find a self-financing porfolio $V_t(v, \theta)$ such that its terminal value equals the contingent claims H:

$$V_T(v, \theta) = H.$$

The method of quadratic hedging provides two ways of addressing this problem. The first one is to release the self-financing requirement but insist on $V_T(x, \theta) = H$, while the second one is to release requirement that $V_T(v, \theta)$ equals H but insist on the self-financing condition. The first approach can be formulated as a locally risk minimization problem and the second one a mean-variance optimality problem. In this subsecion both approaches will be briefly introduced, with focus on application of the mean-variance hedging on exponential Lévy processes.

Under the mean-variance hedging approach, though the self-financing condition is retained, the terminal value of the portfolio $V_t(v, \theta)$ is no longer equal to the contingent claim H. Thus, the task now is to find a self-financing strategy $V_t(v, \theta)$ under a risk-neutral probability measure

Q such that the following loss/error is minimized:

$$\hat{H} - V_T(v,\theta) = \hat{H} - v - \int_0^T \theta(u) d\hat{S}(u), \tag{3.3}$$

where $\hat{S}(t) = \frac{S(t)}{S(0)}$ denotes discounted stock price at time t and $\hat{H} = \frac{H}{S(0)}$ denotes the discounted congingent claim.

(3.3) must be minimized in mean-square sense, i.e. the minimization problem is formulated as:

$$\inf_\theta E_Q[|\hat{H} - V_T(v,\theta)|^2] = \inf_\theta E_Q[|\hat{H} - v - \int_0^T \theta(u) d\hat{S}(u)|^2]. \tag{3.4}$$

The expectation in (3.4) is to be minimized in L^2[1]-norm for mathematical convenience. For probability space $(\Omega, \mathcal{F}, Q)$ the corresponding L^2 norm can be written as $L^2(\Omega, \mathcal{F}, Q)$.

Now define that Θ_{all} denotes the set of all θ such that $\int_0^T \theta(u) d\hat{S}(u) \in L^2(\Omega, \mathcal{F}, Q)$. Then, if Θ is a linear subspace of Θ_{all}, any portfolio strategy $(v,\theta) \in \mathcal{R} \times \Theta$ gives that its payoff $v + \int_0^T \theta(u) d\hat{S}(u)$ will have finite variance. The minimization problem (3.4) can now be formulated as:

Theorem 3.2 *(Mean-Variance Hedging) Assume $H \in (\Omega, \mathcal{F}, Q)$ is a contingent claim with finite variance and $\Theta \subset \Theta_{all}$. Then, a strategy $(v_{optimal}, \theta_{optimal}) \in \Theta$ is called Θ-mean-variance optimal if and only if it minimizes $||\hat{H} - v - \int_0^T \theta(u) d\hat{S}(u)||_{L^2(\Omega, \mathcal{F}, Q)}$ over all strategies in Θ.*

If $(v_{optimal}, \theta_{optimal})$ exists, $v_{optimal}$ is called the Θ-optimal price for the contingent claim H.

As stated in subsection 3.2.2, $\mathcal{A}$ denotes the set of all admissible strategies. That is, it is the set of all contingent claims which can be replicated by self-financing strategies taken from Θ_{all}. To be more specific, $\mathcal{A}$ can be written as:

$$\mathcal{A} = \left\{ v + \int_0^T \theta(u) d\hat{S}(u) \,\middle|\, (v,\theta) \in \mathcal{R} \times \Theta \right\}.$$

The above theorem can be seen as decomposing the minimization problem into finding the orthogonal projection of $\hat{H}$ on $\mathcal{A}$ in L^2-sense. This is due to the classical Galtchouk-Kunita-Watanabe decomposition theorem:

Proposition 3.1 *(Galtchouk-Kunita-Watanabe Decomposition) With $\mathcal{A}$ defined as above. Assume under the risk-neutral probability measure Q, $\{\hat{S}(t)\}_{t \geq 0}$ is a square-integrable martingale*

---

[1]A set of square-integrable functions.

and $\hat{H}$ is a congingent claims with finite variance and depending on the history generated by $\hat{S}(t)$.

Then, $\hat{H}$ can be decomposed as:

$$\hat{H} = E_Q[\hat{H}] + \int_0^T \theta(u) d\hat{S}(u) + L^H, \mathcal{Q} - a.s.$$

where $\theta(t) \in \Theta$ and $L_t^H = E_Q[L^H | \mathcal{F}_t]$ is a local martingale orthogonal to $\mathcal{A}$.

According to the decomposition, if a strategy minimizing (3.4) is available, then it is given by $\theta_{optimial}$ and must be a cáglád[2] process.

This thesis is devoted to application of exponential Lévy processes. Thus it would be of interest to see how mean-variance quadratic hedging can be utilized in exponential Lévy model. In the following only the result of the trading strategies derived from this method will be presented. For the complete derivation and proof of this result, it is recommended to take a look at [Tankov and Cont, 2003].

Assume $\{X(t)\}_{t \geq 0}$ is a Lévy process and $S(t) = S(0)e^{rt + X(t)}$ is the corresponding exponential Lévy process. Assume now that $S(t)$ is square-integrable martingale and $X(t)$ has finite variance. Thus, under Q $\hat{S}(t)$ is also a square-integrable martingale. As before let $(v, \theta) \in \mathcal{R} \times \Theta$, such that $\int_0^T \theta(u) d\hat{S}(u)$ becomes a square-integrable martingale under Q as well. As a result, the solution to (3.4) can be summarized into the following proposition:

Proposition 3.2 *(Mean-Variance Quadratic Hedging) Assume $H(\hat{S}(t))$ is the payoff function of a European option with underlying asset governed by the discounted process $\hat{S}(t)$. Let the discounted expected price of the option at time $t < T$ under the risk-neutral probability measure Q be denoted by:*

$$P(t, S) = e^{-r(T-t)} E_Q[H(\hat{S}(T)) | \hat{S}(T) = S].$$

Assume in addition that $S(t)$ is an exponential Lévy process given by $S(t) = S(0)e^{rt + X(t)}$ where $X(t)$ is a Lévy process with measure v and diffusion σ. For the strategy to exist, the following requirement must be satisfied:

$$|H(x) - H(y)| \leq K|x - y|$$

---

[2]i.e. left continuous with right limits.

for x, y ≥ 0 and K > 0. Then the trading strategy minimizing (3.4) *is given as:*

$$\theta(t) = \frac{\sigma^2 \frac{\partial P}{\partial S(t-)}(t, S(t-)) + \frac{1}{S(t-)} \int v(dy)(e^v - 1)[P(t, S(t-)e^v) - P(t, S(t-))]}{\sigma^2 + \int (e^v - 1)^2 v(dy)}.$$

Note special that both the European call and put options $H_{call}[S(T)] = max((S(T) - K, 0)$ and $H[S(T)]_{put} = max(K - S(T), 0)$ satisfy the condition $|H(x) - H(y)| \le K|x - y|$. Thus the above result works well for these options and any combinations of them.

The above mean-variance quadratic hedging approach is dependent on the choice of the risk-neutral probability measure Q. Since the market is incomplete, different choices of Q give different results. Futhermore, the orthogonal projection is not invariant under measure transformation. These facts force people to consider the other approach, where the trading strategy is no longer self-financing, but instead $V_T(x, \theta) = H$. This method is called risk minimization.

The essence of the method is the use of the so-called Főllmer and Schweizer decomposition [Főllmer and Schweizer, 1989]. The decomposition is a generalization of the previous Galtchouk-Kunita-Watanabe decomposition in that N_T^H now is a stochastic variable. The detail of the method is to find a risk-neutral probability measure Q^{FS} such that for $\hat{S}(t) = M(t) + A(t)$ where $M(t)$ is the orthogonal component of $\hat{S}(t)$ under the real world probability measure P, any maringale orthogonal to $M(t)$ under P will remain orthogonal to $\hat{S}(t)$ under Q^{FS}. Such a martingale measure Q^{FS} is called a minimal martingale measure, which turns $\hat{S}(t)$ into a martingale in such a way that Galtchouk-Kunita-Watanabe decomposition is eligible to be applied.

Since the preservation of orthogonality property of the minimal martingale measure Q^{FS}, L_T^H is now orthogonal to the martingale part of the discounted underlying asset $\hat{S}(t)$ at time T, under the real world probability P. This structure brings the pricing method a great advantage in the sense that its result is closer to the market mechanism, hence more robust.

Under real world probability P, L_T^H is a mean-zero martingale, and the strategy θ derived from the solution of the Főllmer and Schweizer decomposition is called a mean self-financing trading strategy.

Again, this method will be tested on exponential Lévy processes, in order to see its stability and robustness in connection with the model in this thesis. However, according to [Zhang, 1994], the existence of Q^{FS} depends on the structure of the underlying exponential L'evy pro-

cess. To see this, define $S(t) = e^{\mu t + \sigma W(t) + \sum_{i=1}^{N(t)} Y_i}$ where $N(t)$ is a Poisson process with $E[N(t)] = \lambda t$, Y_i are *i.i.d.* with distribution function given by F with mean m and variance δ^2. Then, the condition is formulated as:

$$-1 \le \frac{\mu + \lambda m - r}{\sigma^2 + \lambda(\delta^2 + m^2)} \le 0. \tag{3.5}$$

When the above condition is satisfied, the risk minimization hedge is given:

Proposition 3.3 *(Risk-Minimization Hedging) Assume one has an exponential Lévy process having the form:*

$$S(t) = e^{\mu t + \sigma W(t) + \sum_{i=1}^{N(t)} Y_i}$$

with the corresponding parameters given as above. Let:

$$\eta = \frac{\mu + \lambda m - r}{\sigma^2 + \lambda(\delta^2 + m^2)}$$

verifying condition (3.5). *Let the discounted expected price of the option at time $t < T$ under the risk-neutral probability measure Q^{FS} be denoted by:*

$$P(t, S) = e^{-r(T-t)} E_{Q^{FS}}[H(S(T))|S(T) = S].$$

Then the risk-minimization hedging position on a European-type option is:

$$\theta(t) = \frac{\sigma^2 \frac{\partial P}{\partial S(t-)}(t, S(t-)) + \frac{\lambda}{S(t-)} \int F(dy) y(1 - \eta y)[P(t, S(t-)(1+y)) - P(t, S(t-))]}{\sigma^2 + \lambda \int y^2(1 - \eta y)^2 F(dy)}.$$

More details about the derivation and the proof of the above proposition are in [Tankov and Cont, 2003].

The problems with risk minimization are that trading strategies involved are not self-financing any more, and that the existence of Q^{FS} depends on the structure of the undelying exponential Lévy model $S(t)$.

Chapter 4

Unit-Linked Insurance Policies and Esscher Transform

As stated before, the selection of which method to be applied in pricing derivative securities in incomplete markets is a subjective work. Different practitioners will have different considerations when they choose a specific method. In this thesis, Esscher transform will be chosen to do the job mainly due to its ability to preserve the stationary and independent increments property of Lévy processes. That is, after Esscher transforming the corresponding density function when taking expectation of future payments, the transformed process remains a Lévy process, giving a great mathematical tractability for solving the problem.

In section 4.1, an introduction to unit-linked insurance scheme is given. The scheme is a system, and there are several different unit-linked products in everyday usage. Section 4.2 demonstrates the Esscher transform method and its application in finance theory.

4.1   A Brief Introduction to Unit-Linked Insurance Products

Various unit-linked insurance products were introduced in the 1950s and since then they have gained increasing popularities accross the globe. The main feature of unit-linked insurance policies is that insurance and investment are combined together, in that certain amount of premiums to life insurance products is invested in products in the financial market, while the rest is kept traditionally as technical provisions to cover future losses. Policyholders will be given the right to choose which financial product their paid-in premiums will be invested in, and they get returns based on performances of those products. This scheme gives policyholders more flexibility to allocate their investment needs, and hence making insurance companies more competitive in relation to other financial institutions.

Common investment products in the financial market include stock indices, certain assets and investment portfolios. Returns on those investments are linked to their performances in the market, hence financial risk will be generated. On the other hand, a certain portion of the paid-in premiums is kept to form a pool aiming to cover future claim losses. The pool is called technical provisions and its amount is directly connected to insurance risk, i.e. risk that depends on the mortality of policyholders. Thus, unit-linked insurance policies have two kinds of risk, one is financial and the other one is insurance risk. Those two kinds of risk together are called integrated risk, and unit-linked insurance contracts that depend on both insurance risk and financial risk are called integrated contingent claims. Previous chapters demonstrated that financial risk can be modelled by stochastic analysis, and products with purely financial risk can be hedged and priced accordingly. However, mortality risk is typical unique to insurance market and cannot be hedged away by financial instruments. Hence, in the financial market mortality risk is not hedgable and cannot be priced as their financial risk counterpart. That is, mortality risk is not diversifiable in the financial market, making integrated contingent claims unattainable. Thus, the market of integrated contingent claims is not complete, and various pricing methods introduced in previous chapters can hence be applied.

Unit-linked insurance constracts are ususally financed by a single premium due management convenience. In the market there are several such unit-linked products. If the amount payable to the policyholder upon his survival to maturity time T depends only on the under-

lying asset at T, the contract is called a pure unit-linked life insurance contract. On the other hand, if the amount payable upon time T is the maximum of the underlying asset at T and a predefined amount K, the contract is called a unit-linked life insurance contract with terminal guarantee. The amount K is called a refund guarantee, and it can be a constant or dependent on the amount of paid-in premiums.

The thesis is aiming to price those integrated contingent claims in continuous time. For pricing integrated contingent claims in discrete time horizon by different pricing methods, [Møller and Steffensen, 2007] is a good source of information. In the thesis, however, Esscher transform will be applied to price those contracts in continuous time, which is the topic of the next subsection.

4.2   Esscher Transform

One of the advantages of the Black-Scholes market is that the equivalent martingale measure Q can be obtained by simply changing the drift of the geometric Brownian motion and leaving the Gaussian process untouched. However, when it comes to the market of integrated contingent claims, the market is incomplete and one cannot obtain equivalent martingale measures by simply changing the drift of the process of the underlying asset. In the exponential Lévy model, Gaussian process is swiched to a Lévy process with jumps. In order to obain equivalent martingale measures such that arbitrage property follows, the distribution of Lévy process now has to be altered, by what is called Esscher transform [Esscher, 1932]. This method was proposed by [Gerber and Shiu, 1994], and has a long history of application in actuarial science.

The idea of Esscher is to transform a probability density to another probability density in order to derive some favorable features. The method proposed by Esscher is as follows:

Definition 4.1 *(Esscher Transform) Assume $f(x)$ is a valid probability density function. Then, the Esshcer transform is defined to be:*

$$f(x;h) = \frac{e^{hx}f(x)}{\int_{-\infty}^{+\infty} e^{hx}f(x)dx}$$

where h is a parameter. In a more general setting, the Esscher transfrom can be defined directly on probability measures. Assume now μ is a probability measure, and $E_h(\mu)$ denotes the Esscher transform of the probability measure μ, which again is a probability measure. Then the probability density function corresponding to the new measure is given as:

$$f_{E_h(\mu)}(x;h) = \frac{e^{hx}}{\int_{-\infty}^{+\infty} e^{hx}d\mu(x)}.$$

The Esscher transform of Lévy processes is often conduted via transforming their Lévy measures. Two typical examples are Lévy tilting and Lévy tempering.

Recall the Lévy-Itô decomposition in subsection 2.3.2, an important property of the decomposition is that if $\{X(t)\}_{t\geq 0}$ is a Lévy process with its Lévy measure denoted by $v \in \mathcal{R}^d \setminus \{0\}$. Then,

v must satisfy the condition:

$$\int_{|x|\leq 1} |x|^2 v(dx) < \infty \text{ and } \int_{|x|\geq 1} v(dx) < \infty.$$

Then, Lévy tilting and Lévy tempering are defined to be:

Proposition 4.1 *(**Lévy tiltiing**) If there exists a parameter $\theta \in \mathcal{R}^d$ such that:*

$$\int_{|x|\geq 1} e^{\theta x} v(dx) < \infty$$

then:

$$\tilde{v}(dx) := e^{\theta x} v(dx)$$

is a new Lévy measure. The new Lévy process with $\tilde{v}$ is called Esscher transform of $\{X(t)\}_{t\geq 0}$.

Proposition 4.2 *(**Lévy tempering**) Assume $d = 1$ and $v \in \mathcal{R}$ is a Lévy measure. Then for λ_+ and λ_- as two positive parameters:*

$$\tilde{v}(dx) = v(dx)\left(1_{x>0}e^{-\lambda_+ x} + 1_{x<0}e^{-\lambda_-|x|}\right)$$

is a new Lévy measure. The new Lévy process with $\tilde{v}$ is called Esscher transform of $\{X(t)\}_{t\geq 0}$.

In the thesis the Esscher transform will be applied to derive arbitrage-free risk-neutral probability measures for exponential Lévy prcesses that can be used to generate fair prices of unit-linked insurance contracts, in an incomplete market. The existence of such risk-neutral probability measures follows as:

Proposition 4.3 *(**Esscher transform for exponential Lévy processes**) Let $\{X(t)\}_{t\geq 0}$ be a Lévy process. If there are both positive and negative jumps of the process, then there exists a risk-neutral probability measure Q resulted from Esscher transforming its jump distribution. That is, there exists an arbitrage-free probability measure Q such that exponential Lévy process is arbitrage-free and such that $\{e^{-rt}S(t)\}_{t\in[0,T]}$ is a Q-martingale.*

There are two kinds of exponential Lévy processes. One is called stochastic exponential Lévy

process and is defined to be:

$$\frac{dS(t)}{S(t-)} = r\,dt + dX(t)$$

where $X(t)$ is a Lévy process and r is the corresponding interest rate.

The other one is called ordinary exponential Lévy process (It is simply called exponential Lévy process throughout the thesis) and is defined by:

$$S(t) = S(0)e^{rt+X(t)}.$$

[Esche and Schweizer, 2005] and [Chan, 1999] justified the use of Esscher transform for stochastic exponential Lévy processes in that the equivalent martingale measure Q derived from the Esscher transform minimizes the relative entropy[1] between Q and P over all other equivalent martingale measures. [Raible, 2000] justified the use of Esscher transform for ordinary exponential Lévy process that the density function resulted from the transform is the only density function depending only on the current stock price $S(t)$. All the other equivalent martingale measures have density functions depending on the whole history of the underlying stock price.

---

[1]It is the Kullback-Leibler distance measuring the closeness of two equivalent martingale measures.

Chapter 5

Insurance Reserving with Stochastic Interest Rates

In the insurance industry, insurance companies must set aside certain amounts of capital to cover future payments to the insured. The money saved by the insurance companies for this purpose is called insurance reserves. How much money each insurance company has to reserve depending on its scale and its business area is an important regulation target of many supervision authorities across the world. Inside the EEA[1] area the EU has placed a new insurance regulation directive called Solvency II, aiming to consolidate and harmonize insurance markets across EEA countries. One of the major regulation requirements is that technical provisions[2] of insurance companies in the whole EEA area are required to meet the MCR[3] and the SCR[4] capital levels. If the company fails the MCR, it will be deprived from running business in the European market, while if it fails the SCR, the company will be given a certain amount of period to restore its capital levels. The reason to this new solidarity regulation setup is that insurance reserves are vital to both the company's solvency situation and to the market's stability. Hence, insurance reserving is a fundamental job carried out by actuaries in each insurance company, and the calculations must be based on well-defined models and mathematical reasoning. In this chapter methods for calculating those insurance reserves will be presented.

---

[1] European Economic Area.
[2] Just another fancy word for insurance reserves.
[3] Minimal Capital Requirement.
[4] Solvency Capital Requirement.

Section 5.1 introduces some basic concepts in insurance mathematics. Section 5.2 is focusing on insurance reserving and the corresponding derivation leading to the formula of prospective reserves. Section 5.3 gives a brief introduction to interest rate theory, where the Vasicêk interest rate model will be given as an example for illustrating how to model interest rates mathematically. The Vasicêk model is chosen here is due to its special ability of extension for later application of Lévy processes.

5.1 Primineries in Insurance Mathematics

In an insurance policy, at the final time T the insured will be in certain conditions, called insurance states. Let S denote the state space of an insurance policy. For a life insurance, for example, S will typically be {died, alive} at time T. That is, the insured will either be dead or alive, leading to different payment schemes. For a P&C[5] insurance contracts, S could be {partial damaged, ruined, intact}. Through the thesis S will be assumed to be finite.

Different states correspond to different payment schemes at time T. In order to denote this mechanism, payment functions are needed. $a_i(t)$ denotes the sum of the payments to the insured up to time t, provided that the insured is always in state i. $a_{ij}(t)$ denotes the payment at time t when the state of the insured shifts from state i to state j at time t. Note that $a_i(t)$ is accumulated payments while $a_{ij}(t)$ is a single payment.

To model an insurance policy, a model for the different states of the insured and their interaction is needed. In the final model of this thesis, a Markov chain model will be fitted to describe the states.

Definition 5.1 *(**Markov Chains**) Assume $(\Omega, \mathcal{A}, P)$ is a probability space on which a stochastic process $\{X(t)\}_{t \geq 0}$ is defined. Let S denote a state space and T denote the final time. If for all:*

$$n \geq 1, t_1 < t_2 < \cdots < t_{n+1} \in T, i_1, i_2, \ldots, i_{n+1} \in S,$$

and $P(X_{t_1} = i_1, X_{t_2} = i_2, \ldots, X_{t_n} = i_n) > 0$, the following condition is satisfied:

$$P(X_{t_{n+1}} = i_{n+1} | X_{t_1} = i_1, X_{t_2} = i_2, \ldots, X_{t_n} = i_n) = P(X_{t_{n+1}} = i_{n+1} | X_{t_n} = i_n),$$

then $\{X(t)\}_{t \geq 0}$ is called a Markov chain.

Markov chains are just processes that the future state only depend on the current state. Past states that led the process into the current state are ignored.

Interaction between states can be represented by transitional probabilities. That is, they are the probabilities that describing transitions between different states, accross time horizon.

---

[5]Property and casualty insurance.

Definition 5.2 *(Transitional Probabilities) Let $\{X(t)\}_{t \geq 0}$ be a stochastic process defined on the probability space $(\Omega, \mathcal{A}, P)$, and let S be the state space of the process and T is the final time of the process. Then:*

$$p_{ij}(s, t) = P(X_t = j | X_s = i), \text{ for } s \leq t \text{ and } i, j \in S$$

is called the transitional probability of the process from state i to state j from time s to time t.

One important theorem in conjunction with transitional probabilities is the Chapman-Kolmogorov equation:

Theorem 5.1 *(Chapman-Kolmogorov Equation) Let $\{X(t)\}_{t \geq 0}$ be a Markov chain defined on the probability space $(\Omega, \mathcal{A}, P)$, and let S be the state space of the process and $i, j \in S$. Let T be the final time of the process and assume $s \leq t \leq u \in T$, then the Chapman-Kolmogorov equation states that:*

$$p_{ij}(s, u) = \sum_{k \in S} p_{ik}(s, t) p_{kj}(t, u)$$

where $P(X_i) > 0$ for all $i \in S$.

Another concept that is closely related to transitional probabilities is transitional rates, which are just instantaneous transitional probabilities from t to $t + \triangle t$, and can be defined as:

$$\mu_i(t) = \lim_{\triangle t \to 0} \frac{1 - p_{ii}(t, t + \triangle t)}{\triangle t} \text{ for } i \in S \tag{5.1}$$

$$\mu_{ij}(t) = \lim_{\triangle t \to 0} \frac{p_{ii}(t, t + \triangle t)}{\triangle t} \text{ for } i, j \in S \tag{5.2}$$

$\triangle t$ denotes a small positive increment of time.

If (5.1) and (5.2) exist, the underlying Markov chain is said to be regular. Note special that $\mu_{ii}(t) = -\mu_i(t)$. From now on, all Markov chains in the thesis will be assumed to be regular.

With transitional probabilities and transitional rates at hand, the powerful forward and backward Kolmogorov theorem can be defined as:

Theorem 5.2 *(Kolmogorov Differential Equations) Let $\{X(t)\}_{t \geq 0}$ be a Markov chain defined on the probability space $(\Omega, \mathcal{A}, P)$, and let S be the state space of the process and $i, j \in S$. Let T be the*

final time of the process and assume $s \leq t \leq u \in T$, then the following equations hold:

$$\frac{d}{ds}p_{ij}(s,t) = \mu_i(s)p_{ij}(s,t) - \sum_{k \neq i} \mu_{ik}(s)p_{kj}(s,t) \text{ (Backward Kolmogorov equation)}$$

$$\frac{d}{dt}p_{ij}(s,t) = -p_{ij}(s,t)\mu_j(t) + \sum_{k \neq j} p_{ik}(s,t)\mu_{kj}(t) \text{ (Forward Kolmogorov equation)}.$$

The Kolmogorov backward and forward equations can be used to calculated transitional probabilities between different states. However, for the probabilities of staying in the same state during a time interval, one can apply the following equation:

$$\bar{p}_{jj}(s,t) = e^{-\sum_{k \neq j} \int_s^t \mu_{jk}(u)du}.$$

Proofs of the theorems stated in this subsection can be found in [Koller, 2012].

5.2 Reserving for Insurance Claims

In actuarial science, the money the insurance companies has to be set aside in order to cover future payments to the insured are called mathematical reserves. For the purpose of illustrating the essence of calculating mathematical reserves in a simpler setting, now assume the interest rate is constant. Later this assumption will be released and the interest rate will be allowed to vary over time.

An important concept in relation to interest rate is interest intensity, which is denoted by $\delta(t)$. In the interest rate is a constant r, then then interest intensity is also a constant given to be $\delta = ln(1+r)$.

Pricing of insurance contracts involves discounting the cash flows paid by the contract over time. The so-called discount rate is central when discounting cash flows, and is defined to be $v(t) = e^{-\int_0^t \delta(u)du}$, for continuous time model.

In the final model, cash flow will be defined to be random, i.e. they are stochastic processes of bounded variation that vary over time. Bounded variation is a desirable feasure when working with models in reality. Processes with bounded variation mean that all sample paths will not blow up during the whole time horizon. Mathematically, for a probability space $(\Omega, \mathcal{A}, P)$ on which $\{X(t)\}_{t \geq 0}$ is defined and a state space S, stochastic cash flows can be defined to be:

$$dA_{ij}(t, \omega) = a_{ij}(t)dN_{ij}(t, \omega) \tag{5.3}$$

$$dA_i(t, \omega) = I_i(t, \omega)da_i(t), \tag{5.4}$$

where $N_{ij}(t, \omega)$ denotes the number of individuals changing from state i to state j up to time t when scenario ω occurs, $a_{ij}(t)$ and $a_i(t)$ are payout functions defined earlier, and $I_i(t, \omega)$ is an indicator function which is 1 when $X(t, \omega) = i$ and 0 otherwise.

(5.3) signifies cash flows when transaction from state i to state j occurs while (5.4) corresponds to cash flows generated by staying in the same state i. Thus, total cash flows up to time t when scenarios ω occurs is given by:

$$dA(t, \omega) \quad = \quad \sum_{i \in S} dA_i(t, \omega) + \sum_{(i,j) \in S \times S, i \neq j} dA_{ij}(t, \omega)$$

$$= \sum_{i \in S} I_i(t,\omega) \, da_i(t) + \sum_{(i,j) \in S \times S, i \neq j} a_{ij}(t) \, dN_{ij}(t,\omega)$$

The value of the cash flows $A(t)$ at time t can then be defined to be:

$$
\begin{aligned}
V^+(t, A) \ &= \ \frac{1}{v(t)} \int_t^T v(u) \, dA(u,\omega) \\
&= \ \frac{1}{v(t)} \int_t^T v(u) \left[\sum_{i \in S} I_i(u,\omega) \, da_i(u) + \sum_{(i,j) \in S \times S, i \neq j} a_{ij}(u) \, dN_{ij}(u,\omega) \right] \\
&= \ \frac{1}{v(t)} \left[\sum_{i \in S} \int_t^T v(u) I_i(u,\omega) \, da_i(u) + \sum_{(i,j) \in S \times S, i \neq j} \int_t^T v(u) a_{ij}(u) \, dN_{ij}(u,\omega) \right] \quad (5.5)
\end{aligned}
$$

Assume that the stochastic cash flows $\{A(t)\}_{t \geq 0}$ defined on the probability space $(\Omega, \mathcal{A}, P)$ are adapted to the information flows $\{\mathcal{F}(t)\}_{t \geq 0}$. That is, up to time t all information about $A(t)$ is contained in $\mathcal{F}(t)$ such that it is impossible to predict future scenarios of $A(t)$. In the current Markov chain model, all information contained in $\mathcal{F}(t)$ is nothing but the process $X(t)$. Hence, by plugging in (5.5), prospective reserves[6] of the insurance contract $X(t)$ can be defined as:

$$
\begin{aligned}
V_k^+(t, A) \ &= \ E[V^+(t, A) | \mathcal{F}(t)] \\
&= \ E[V^+(t, A) | X(t) = k] \\
&= \ E\left(\frac{1}{v(t)} \left[\sum_{i \in S} \int_t^T v(u) I_i(u,\omega) \, da_i(u) + \sum_{(i,j) \in S \times S, i \neq j} \int_t^T v(u) a_{ij}(u) \, dN_{ij}(u,\omega) \right] \, \middle| \, x(t) = k \right) \\
&= \ \frac{1}{v(t)} \sum_{i \in S} E\left[\int_t^T v(u) I_i(u,\omega) \, da_i(u) \, \middle| \, X(t) = k \right] \\
&\quad + \frac{1}{v(t)} \sum_{(i,j) \in S \times S, i \neq j} E\left[\int_t^T v(u) a_{ij}(u) \, dN_{ij}(u,\omega) \, \middle| \, X(t) = k \right] \\
&= \ \frac{1}{v(t)} \sum_{i \in S} \int_t^T v(u) p_{ki}(t,u) \, da_i(u) + \frac{1}{v(t)} \sum_{(i,j) \in S \times S, i \neq j} \int_t^T v(u) a_{ij}(u) p_{ki}(t,u) \mu_{ij}(u) \, du \\
&= \ \frac{1}{v(t)} \left[\int_t^T v(u) \sum_{i \in S} p_{ki}(t,u) \, da_i(u) + \int_t^T v(u) \sum_{i \in S} p_{ki}(t,u) \sum_{j \in S, j \neq i} a_{ij}(u) \mu_{ij}(u) \, du \right] \\
&= \ \frac{1}{v(t)} \int_t^T v(u) \sum_{i \in S} p_{ki}(t,u) \left(da_i(u) + \sum_{j \in S, j \neq i} a_{ij}(u) \mu_{ij}(u) \, du \right). \quad (5.6)
\end{aligned}
$$

The derivation above is based on the assumption that $a_i(t), a_{ij}(t) \in L^1(\mathcal{R})$. A more detailed

---

[6] Present value of future cash flows conditional on current state of the process.

proof of the derivation can be found at [Koller, 2012].

In (5.6) for calculating prospective reserves, the interest rate is assumed to be constant for simplicity. However, stochastic interest rates are becoming more and more popular in both academic world and practical application. In the next chapter, unit-linked insurance contracts with stochastic interest rates will be given as a illustrating example on how they look like. Exponential Lévy model is a good candidate to be plugged into (5.6). Next section is a brief introduction to interest rate theory.

5.3 Building the Jump-Diffusion Process for Interest Rates

5.3.1 Basic Concepts of Interest Rate Theory

In the derivation of (5.6), the yearly interest rate was assumed to be a constant. However, in the following this condition will be released and a stochastic interest rate model is required. But first, two most important concepts in interest rate theory must be made clear:

Short Rate: *Interest rates for a specified time interval (often one year) offered at different time points.*

Spot Rate: *Interest rates quoted immediately on certain zero-coupon bonds with different maturities.*

Spot rates, denoted by $r(t)$, are also called instantaneous spot rates. They are the basic state variables used in short-rate models, and are the interest rates we usually understand in financial activities.

Zero-coupon bonds are fundamental in interest rate theory. In words, a zero-coupon bond is a bond that does not generate any periodic interest payments/coupons during its existence, with its face value paid only at the pre-defined maturity time T. Zero-coupon bonds with different maturities have different yields[7]. Zero-coupon bonds are closely related to spot rates. Let $r(t)$ denotes a stochastic spot-rate model adapted to information flows $\mathcal{F}_t$ and Q is the corresponding risk-neutral probability measure for $r(t)$. Then, the underlying zero-coupon bond price at time t with maturity T is given by:

$$B(t,T) = E_Q[e^{-\int_t^T r(s)ds}|\mathcal{F}_t].$$

The yield curve produced by the zero-coupon bond prices makes future interest rates of the bond available. The future interest rates of the zero-coupon bond can then be used to define the so-called instantaneous forward rates:

$$f(t,T) = -\frac{\partial}{\partial T}lnB(t,T).$$

---

[7]An annually-reported figure indicating how much dividend, interest or return an investor will get from a security, similar to interest rates but not exactly the same.

Straightforwardly, a forward rate is simply the future yield on the underlying zero-coupon bond.

5.3.2 Vasiček Interest Rate Model

There are today many stochastic interest models, but here the Vasiček model will be introduced due to its property of easy extension to other processes.

Let $r(t)$ denote the spot rate at time t, then the dynamics of the Vasiček model is given as:

$$dr(t) = a(b - r(t))dt + \sigma dW(t), \tag{5.7}$$

where $W(t)$ is a Brownian motion and $a > 0$. The process has initial assumption $r(0) = r_0$.

One of the most important features of the Vasiček model is its mean reversion property, which means that in the long run trajectories of $r(t)$ will be fluctuating around b. Without the stochastic noise induced by the Brownian motion, the trajectories will converge exactly to b in the long run.

The process was first studied in physics and is called Ornstein-Uhlenbeck process. Its solution has closed form, making it desirable in practice. Meanings to the several parameters can be explained as:

a: *controls how strong the reversion effect is around b.*

σ: *volatility of r(t), measuring the spread of the trajectories of r(t).*

The closed-form solution to (5.7) is found to be:

$$r(t) = r_0 e^{-at} + b(1 - e^{-at}) + \sigma e^{-at} \int_0^t e^{as} dW(s).$$

Since $W(t)$ is a Brownian motion, $E[W(t)] = 0$ and $V[W(t)] = t$. Thus, the mean and variance of $r(t)$ are given as:

$$E[r(t)] = r_0 e^{-at} + b(1 - e^{-at})$$

$$V[r(t)] = \frac{\sigma^2}{2a}(1 - e^{-2at}).$$

By definition $a > 0$, thus $E[r(t)] \to b$ and $V[r(t)] \to \frac{\sigma^2}{2a}$ as $t \to \infty$. The convergence of the

mean of $r(t)$ is called mean reversion property of short rates. However, random fluctuations induced by the Brownian motion $W(t)$ are of magnitude $\frac{\sigma}{\sqrt{2a}}$, which converges to 0 as $a \to \infty$ or $\sigma \to 0$. This proves that a controls how strong the reversion is and that strong reversion corresponds to low volatility.

Though having nice tractability properties, there are several significant shortcomings of the Vasicěk model. One of which is its positive probabilities for negative interest rates, for example. However, the model is introduced here is due to its ability for extension. In the next chapter, there will be given an example where the Brownian motion term in (5.7) will be replaced by a Lévy process, such that the new interest rate model is capable of describing interest rate jumps. Thus, in contrast to the Vasicěk model, the new model will no longer be continuous.

Chapter 6

The Final Model

In this chapter all the results in the previous chapters will serve as the foundation for the final model. Derivations, limitations and difficulties of the final model will be properly stated, aiming to give the readers a as clear formulation as possible.

The final model of the thesis involves stochastic interest rates, and is one of the many forms of (unit-linked) insurance contracts. Another example, where interest rates are assumed constant but the payment function of the prospective reserves is directly linked to a derivative security will also be given in section 1 so as to make the picture more complete. The pricing of this more primitive unit-linked insurance contract is easier than in the final model. In the academic literature there are several materials specializing in pricing of derivative securities in incomplete markets.

The form of the final model reveals that there is in general no analytical solutions to the integro-partial differential equation, such that simulation techniques must be applied to derive prices of the (unit-linked) contracts.

6.1   An Example of Unit-Linked Insurance Contract with Constant Interest Rates

In contrast to the final model, where interest rates are assumed to be stochastic, a more traditional and more obvious example of unit-linked life insurance contract is where interest rates are assumed to be constant, but the payment function corresponding to state shift will be directly connected to a derivative security which depends on the dynamics of the underlying stock. In this example a call option will be linked to the insurance contract, where the underlying stock dynamics is modelled by an exponential Lévy process, meaning that stock price jumps are allowed in the model.

Recall again (5.6), where interest rates are assumed constant. The unit-linked scheme is entered into the model via $a_{ij}(u)$, i.e. when the state to the insured shifts from state i to state j at time t, such that its form now becomes:

$$\tilde{a}_{ij}(t) = E_Q\left[e^{-rT}a_{ij}(T)|\mathcal{G}_t\right],$$

and (5.6) is changes to:

$$V_k^+(t,A) = \frac{1}{v(t)}\int_t^T v(u)\sum_{i\in S} p_{ki}(t,u)\left(da_i(u) + \sum_{j\in S, j\neq i}\tilde{a}_{ij}(u)\mu_{ij}(u)du\right).$$

For a life insurance contract typically, $i = *$ and $j = †$ where $*$ denotes alive and $†$ denotes death, such that one has:

$$a_{*†}(t) = max(S(t), G(t)).$$

$G(t)$ is a deterministic refund guarantee or premium compensation whose value depends on time t, whereas $S(t)$ is the stock price at time t modelled by the exponential Lévy process:

$$S(t) = S(0)e^{rt+L(t)}$$

where $L(t)$ is a Lévy proess.

$\tilde{a}_{ij}(t)$ can now be written as:

$$
\begin{aligned}
\tilde{a}_{ij}(t) &= E_Q\left[e^{-rT}a_{*\dagger}(T)|\mathcal{G}_t\right] \\
&= E_Q\left[e^{-rT}max(S(T),G(T))|\mathcal{G}_t\right] \\
&= E_Q\left[e^{-rT}max(S(0)e^{rT+L(T)},G(T))|\mathcal{G}_t\right] \\
&= E_Q\left[e^{-rT}\left(S(0)e^{rT+L(T)}-G(T)\right)^+|\mathcal{G}_t\right] \\
&= e^{-rT}\int_\tau^{+\infty}\frac{e^{hl}f(l)}{\int_{-\infty}^{+\infty}e^{hl}f(l)dl}\left(S(0)e^{rT+L(T)}-G(T)\right)dl
\end{aligned}
\tag{6.1}
$$

where $\tau = log\left[\frac{G(T)}{S(0)}\right] - rT$ and $f(l)$ is the density of the Lévy process $L(T)$.

In the derivation of (6.1) above, the Esscher transform was applied to derive the corresponding risk-neutral probability measure Q. The only job remains to be done in (6.1) is the determination of the parameter h.

There is no analytical solution to (6.1). Simulation methods like Monte Carlo must be used to model the dynamics of the equation.

6.2 The Final Model

In (5.7), Brownian motion was used as the diffusion part. As noted in 2.2.3, Brownian paths are continuous but no where differentiable. Thus, together with the drift part, Vasiček model is used for modelling continuous interest rates. However, the continuous interest rate assumption is not reasonable in practice, as interest rates determined by various central banks around the world often jump up or down frequently. Hence, an interest rate model taking into account interest rate jumps will be more realistic.

In (5.7), the Brownian motion will be replaced by a Lévy process, which, as stated in 2.3.1, allows for jumps in its trajectories. The form of the new interest rate model is:

$$dr(t) = a(b - r(t))dt + dL(t),$$

where $L(t)$ denotes Lévy process at time t.

As before assume that the initial value of the process is $r(0) = r_0$, then the dynamics of $r(t)$ can be written as:

$$r(t) = r_0 + \int_0^t a(b - r(s))ds + L(t), 0 \le t \le T. \tag{6.2}$$

(6.2) is the final model for interest rates. But due to the presence of jumps, the market is in general incomplete, having the consequence that there may exist infinitely many risk-neutral martingale measures available for no-arbitrage pricing. As explained before, the Esscher transform can be used to choose a unique probability measure in such a case, enabling pricing method as in the complete market scenario can be used as usual.

But first, note that the discount function from s to t based on continuous-time stochastic interest rate model is given as:

$$B(s, t) = E_Q[e^{-\int_s^t r(u)du}|\mathcal{G}_s], \tag{6.3}$$

where $r(t)$ is assumed to be adapted to the information flows $\mathcal{G}_t$, and Q is a risk-neutral probability measure, resulted from the use of Esscher transform.

(6.3) is the same as the price at s of a zero-coupon bond paying 1 at maturity time t, and it will be used to discount the cash flows of an insurance contract, aiming to calculate the reserves

of the contract at a fixed time point.

Recall (5.6) for calculating the prospective reserves, where $v(t)$ is the discount factor based on constant interest rate assumption. Now plugging (6.3) for stochastic interest rates (5.6) results in:

$$
\begin{aligned}
V_k^+(t, A) &= \int_t^T B(t, u) \sum_{i \in S} p_{ki}(t, u) \left(da_i(u) + \sum_{j \in S, j \neq i} a_{ij}(u) \mu_{ij}(u) du \right) \\
&= \int_t^T E_Q \left[e^{-\int_t^u r(s) ds} \Big| \mathcal{G}_t \right] \sum_{i \in S} p_{ki}(t, u) \left(da_i(u) + \sum_{j \in S, j \neq i} a_{ij}(u) \mu_{ij}(u) du \right) \quad (6.4)
\end{aligned}
$$

for calculating the prospective reserves at time t under stochastic interest rate assumption.

(6.4) calculates the prospective reserves at time t of an insurance contract with stochastic interest rates modelled by means of a Lévy process, where it is given that the current state of the insured at time t is k.

Finally we want to find conditions under which the discounted bond prices become a local martingale with respect to the Esscher transform.

Now assume that the Lévy process $L(t)$ can be written as:

$$
L(t) = \sigma B(t) + \int_0^{t^+} \int_{\mathcal{R}_0} z \mathbf{1}_{|z| < |}(z) \tilde{N}(ds, dz)
$$

where $\mathcal{R}_0 = \mathcal{R} - \{0\}, \sigma > 0, B(t), 0 \leq t \leq T$ is a Brownian motion, and $\tilde{N}(ds, dz) = N(ds, dz) - v(dz) ds$ is the compensated Poisson random measure associated with the Lévy process and where v is the corresponding Lévy measure.

The Radon-Nikodym density with respect to the Esscher transform can now be represented as:

$$
f = e^{\theta L(T) + \gamma(\theta) T}
$$

where $\theta, \gamma(\theta)$ are real numbers. See [Tankov and Cont, 2003].

Let $\tilde{P}$ denote the corresponding Esscher transform applied in the model, then under $\tilde{P}$ (6.2) can be written as:

$$
r(t) = r_0 + \int_0^t a(b - r(s)) ds + \tilde{L}(t), 0 \leq t \leq T, \quad (6.5)
$$

where

$$\tilde{L}(t) = \tilde{\gamma} t + \int_0^{t^+} \int_{\mathcal{R}_0} z \mathbf{1}_{|z|<1}(z) \tilde{N}(ds, dz), 0 \le t \le T \tag{6.6}$$

is a new Lévy process with the corresponding Lévy measure given by $\tilde{v}(dz) = e^{\theta z} v(dz)$. In (6.6), the new parameter $\tilde{\gamma}$ is defined to be:

$$\tilde{\gamma} = \int_1^{-1} z(e^{\theta z} - 1) v(dz).$$

Assume that g is a C^2 function and that that zero-coupon bond price takes the form:

$$B(t, T) = g(t, r(t)), 0 \le t \le T$$

and $r(t)$ is given by (6.5).

Applying Itô's formula (see [Tankov and Cont, 2003]) on g results in:

$$
\begin{aligned}
B(t, T) &= g(r, r(t)) \\
&= g(0, r(s)) + \int_0^t \frac{\partial g}{\partial s}(s, r(s)) ds + \int_0^t \frac{\partial g}{\partial r(s)}(s, r(s)) \tilde{\gamma} ds \\
&\quad + \int_{0_t}^t \int_{\mathbf{R}_0} \left[g\left(s, r(s^-) + z\mathbf{1}_{|z|<1}(z)\right) - g(s, r(s^-)) \right] \tilde{N}(ds, dz) \\
&\quad + \int_0^t \int_{\mathbf{R}_0} \left[g\left(s, r(s) + z\mathbf{1}_{|z|<1}(z)\right) - g(s, r(s)) - z\mathbf{1}_{|z|<1}(z) \frac{\partial g}{\partial r(s)}(s, r(s)) \right] \tilde{v}(dz) ds
\end{aligned}
\tag{6.7}
$$

under $\tilde{P}$.

Define $R(t) = e^{-\int_0^t r(s) ds}, 0 \le t \le T$, such that by applying integration by parts the discounted bond prices at time 0 can be written as:

$$
\begin{aligned}
\tilde{B}(t, T) &= R(t) B(t, T) \\
&= g(0, r(s)) + \int_0^t \tilde{P}(s, T)(-1) r(s) ds + \int_0^t R(s) \frac{\partial g}{\partial s}(s, r(s)) ds + \int_0^t R(s) \frac{\partial g}{\partial r(s)}(s, r(s)) \tilde{\gamma} ds \\
&\quad + \int_{0_t}^t \int_{\mathbf{R}_0} R(s) \left[g\left(s, r(s^-) + z\mathbf{1}_{|z|<1}(z)\right) - g(s, r(s^-)) \right] \tilde{N}(ds, dz) \\
&\quad + \int_0^t R(s) \int_{\mathbf{R}_0} \left[g\left(s, r(s) + z\mathbf{1}_{|z|<1}(z)\right) - g(s, r(s)) - z\mathbf{1}_{|z|<1}(z) \frac{\partial g}{\partial r(s)}(s, r(s)) \right] \tilde{v}(dz) ds
\end{aligned}
$$

under $\tilde{P}$.

According to the above relation, it is clear that $\tilde{B}(t, T)$ is a $\tilde{P}$-(local) martingale if and only if:

$$\tilde{P}(s, T)(-1)r(s) + R(s)\frac{\partial g}{\partial s}(s, r(s)) + R(s)\frac{\partial g}{\partial r(s)}(s, r(s))\tilde{\gamma}$$
$$+ R(s)\int_{\mathbf{R}_0}\left[g\left(s, r(s) + z\mathbf{1}_{|z|<1}(z)\right) - g(s, r(s)) - z\mathbf{1}_{|z|<1}(z)\frac{\partial g}{\partial r(s)}(s, r(s))\right]\tilde{v}(dz) = 0 \text{ a.e.}.$$

Thus if $r(t)$ has a probability density for all $t > 0$, one has the following condition for the Esscher transform $\tilde{p}$ to ba a (local) martingale measure (or risk-neutral pricing measure), given by the integro-partial differential equation for g:

$$-yg(s, y) + \frac{\partial g}{\partial s}(s, y) + \frac{\partial g}{\partial y}g(s, y)\tilde{\gamma} + \int_{\mathbf{R}_0}\left[g(s, y + z\mathbf{1}_{|z|<1}(z) - g(s, y)\right]e^{\theta z}v(dz) = 0$$

with boundary condition for all $g(T, y) = 1$ for all y.

Finally, plugging (6.8) into (6.4) results in the the model we have mentioned, that is, the prospective reserves of a life insurance contract based on stochastic interest rates.

There is in general no analytical solution with respect to our model, thus simulation techniques like Monte Carlo method may be applied in order to calculate the prospective reserves.

Appendix A

Acronyms

1. *i.i.d.: independent and identically distributed random variables.*

2. *EMM: equivalent martingale measure.*

3. SDE: stochastic differential equation.

4. *a.e.: almost everywhere.*

5. *a.s.: almost surely.*

6. *lim: limit of a sequence or of a function.*

7. *log: logarithm.*

8. *max: maximum of a set.*

9. *min: minimum of a set.*

10. *pdf: probability density function.*

11. *mgf: moment-generating function.*

12. *cdf: cumulative distribution function.*

13. *sup: supremum.*

14. *inf: infimum.*

Bibliography

Benth, F. E. (2003). *Option Theory with Stochastic Analysis*. Universitext. Springer.

Berk, K. N. and Devore, J. L. (2007). *Modern Mathematical Statistics with Applications*. Cengage Learning, 1 edition.

Black, F. and Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3):637–654.

Burger, P. and Kliaras, M. (2013). Jump diffusion models for option pricing vs. the black scholes model. Working Paper Series by the University of Applied Sciences bfi Vienna.

Carmona, R., editor (2009). *Indifference Pricing: Theory and Applications*. Princeton Series in Financial Engineering. Princeton University Press.

Chan, T. (1999). Pricing contingent claims on stocks driven by Lévy processes. *The Annals of Applied Probability*, 9(2):pp. 504–528.

Delbaen, F. and Schachermayer, W. (1994). A general version of the fundamental theorem of asset pricing. *Mathematlsche Annalen*, (300):463–520.

Esche, F. and Schweizer, M. (2005). Minimal entropy preserves the lévy property: how and why. *Stochastic Processes and their Applications*, 115(2):299 – 327.

Esscher, F. (1932). On the probability function in the collective theory of risk. *Skandinavisk Aktuarietidskrift*, 15(3):175–195.

Főllmer, H. and Schweizer, M. (1989). Hedging by sequential regression an introduction to the mathematics of option trading. *ASTIN Bulletin*, (18):147–160.

Főllmer, H. and Sondermann, D. (1986). Hedging of non-redundant contingent claims. In Hildenbrand, W. and Mas-Colell, A., editors, *Contributions to Mathematical Economics*, In Honor of Gérard Debreu, pages 205–223, Amsterdam. North-Holland Press.

Fujiwara, T. and Miyahara, Y. (2003). The minimal entropy martingale measures for geometric lévy processes. *Finance and Statistics*, (7):509–531.

Gerber, H. U. and Shiu, E. S. (1994). Option pricing by esscher transforms. *Transactions of Society of Actuaries*, 46:99–191.

Karoui, N. E. and Quenez, M. (1997). Imperfect markets and backward stochastic differential equations. In Rogers, L. C. G. and Talay, D., editors, *Numerical Methods in Finance*, pages 181–214. Cambridge University Press. Cambridge Books Online.

Koller, M. (2012). *Stochastic Models in Life Insurance*. EEA Series. Springer.

Kramkov, D. (1996). Optional decomposition of supermartingales and hedging contingent claims in incomplete security markets. *Probability Theory and Related Fields*, 105(4):459–479.

Matsuda, K. (2004). Introduction to merton jump diffusion model. Working Paper, The Graduate Center, The University of New York.

Merton, R. C. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, (3):125–144.

Møller, T. and Steffensen, M. (2007). *Market-Valuation Methods in Life and Pension Insurance*. International Series on Actuarial Science. Cambridge University Press, 1st edition.

Øksendal, B. (2010). *Stochastic Differential Equations*. Universitext. Springer, 6th edition.

Pliska, S. R. (1997). *Introduction to Mathematical Finance: Discrete Time Models*. Wiley, 1 edition.

Raible, S. (2000). *Lévy Processes in Finance: Theory, Numerics, and Empirical Facts*. PhD thesis, Albert-Ludwigs-Universität Freiburg.

Sato, K.-i. (1999). *Lévy Processes and Infinitey Divisible Distributions.* Cambridge Studies in Advanced Mathematics. Cambridge University Press.

Tankov, P. and Cont, R. (2003). *Financial Modelling with Jump Processes.* Financial Mathematics Series. Chapman and Hall/CRC, 1st edition.

Zhang, X. (1994). *Analyse Numérique des Options Américaines dans un Modèle de Diffusion avec Sauts.* PhD thesis, Ecole Nationale des Ponts et Chaussées.