UiO : Department of Mathematics
University of Oslo

Thiele's partial integro-differential equation for jump diffusions

Vegard Enerstvedt

Master's Thesis, Spring 2022

This master's thesis is submitted under the master's programme *Stochastic Modelling, Statistics and Risk Analysis*, with programme option *Finance, Insurance and Risk*, at the Department of Mathematics, University of Oslo. The scope of the thesis is 30 credits.

The front page depicts a section of the root system of the exceptional Lie group E_8, projected into the plane. Lie groups were invented by the Norwegian mathematician Sophus Lie (1842–1899) to express symmetries in differential equations and today they play a central role in various parts of mathematics.

Abstract

Unit-linked policies are often complex life insurance products to price. This is due to there being more forms risk in these policies than in traditional life insurance. Unit-liked policies combine investments in financial assets with a traditional life insurance policy. In addition to modeling mortality, one also needs to model the financial investment.

The pricing of unit-linked policies can be complicated and difficult to calculate. Results like Thiele's PDE for unit-linked policies have therefore been proven for securities models like the Black and Scholes, see Aase and Persson 1993. In this thesis, we are going to generalize this to contexts where the model for the underlying securities is more complex and realistic. We find a partial integro-differential equation that models the evolution of value for a unit-linked life insurance policy when the underlying security follows a jump diffusion model. Jump diffusions are a generalization of diffusion models (like the Black and Scholes), and allow for better modeling, especially the tails of the returns. This means that large losses can be better modeled.

In addition to this we, propose some methods for numerically calculating prices and we use these methods to find the value of some example policies.

Acknowledgements

I would obviously like to thank some people, but I will keep it short. I want to thank my girlfriend Brenna and my family for their love and support, and my friends Per August, Vivian, Aurora and Eivind for providing memes and sparring on academic and other matters. I would like to thank my supervisor David Baños for his amazing help. Without him, this thesis would not have been possible.

I would like to thank Håkon Bihaug Soydan for checking the page numbers, Hazel Pessig, Glenn Palmer Smith, Cynthia Cahlor, Mariah Atkins, Johnathan Atkins, the lords Coe, Halifax and Beaverbrook, Eleanor Roosevelt, Jacques Cousteau and his wife Mimi, Søren Kierkegaard, Tom Brady, Sonny Liston, Hayden Jones and her husband Pip, Gregor Mendel the founder of the science of genetics, my tailor, Harper Lee the author of "To kill a Mockingbird", The London Symphony Orchestra woodwind section, The Duke of Wellington, Mother, Bismarck, The Royal Society for the Prevention of Birds (RSPB), Michel Palin.

Saint Francis of Assisi, Diana Ross and the Supremes, Alexander Haig, Wile E. Coyote, Mother Theresa, Herb Alpert and his Tijuana Brass, Rowan Atkinson, Dame Agatha Christie, The planet Saturn and of course all of its rings, Joan Collins, the publicity department of Turkish Airways, The Unknown Soldier, Tammy Wynette, John Cleese 2013, and last but not least God.

Contents

Abstract **i**

Acknowledgements **iii**

Contents **v**

List of Figures **vii**

1 Introduction **1**
1.1 General background 2
1.2 The structure of the thesis 3

I Theory **5**

2 Mathematical preliminaries **7**
2.1 An introduction to measure theory 7
2.2 An introduction to probability theory 10
2.3 Stochastic processes 12
2.4 Martingales 15
2.5 Lévy process 19
2.6 Markov process 22
2.7 The Poisson process 24

3 Stochastic integrals **27**
3.1 L^p−spaces 27
3.2 Convergence 28
3.3 Stochastic integral: Construction 29
3.4 Stochastic integral: Properties 31
3.5 Example: The Itô integral 33
3.6 Quadratic variation 34
3.7 Itô's formula 35
3.8 Stochastic differential equations 36
3.9 The Feynman-Kac formula 37

4 Mathematical finance **41**
4.1 The financial market 41

4.2 The portfolio and trading strategies . . . 42
4.3 Arbitrage . . . 45
4.4 The Girsanov theorem . . . 47
4.5 Complete markets . . . 49
4.6 Pricing contingent claims . . . 51

5 Insurance mathematics **55**
5.1 The insurance model . . . 55
5.2 The value of insurance policies . . . 57
5.3 Thiele's differential equation . . . 59
5.4 Examples of life insurance policies . . . 65

6 Unit-linked life insurance policies **71**
6.1 The value of a unit-linked policy . . . 71
6.2 Thiele's PIDE for unit-linked policies with jump diffusions . . . 72

II Applications **79**

7 Numerical methods **81**
7.1 The Merton jump-diffusion model . . . 81
7.2 The mortality model: K2013 . . . 83
7.3 The explicit method . . . 85
7.4 The implicit method . . . 87

8 Examples **91**
8.1 Example 1: Pure endowment policy . . . 91
8.2 Example 2: Death benefit . . . 94
8.3 Example 3: Pension . . . 95
8.4 Example 4: Commodity prices . . . 97
8.5 Sensitivity analysis . . . 99

9 Conclusion and discussions **101**
9.1 Conclusion . . . 101
9.2 Further research . . . 102
9.3 Critical remarks . . . 102

Appendices **105**

A Some proofs **107**
A.1 Chapter 2 . . . 107
A.2 Chapter 3 . . . 110
A.3 Chapter 4 . . . 116
A.4 Chapter 5 . . . 120

B Code **125**
B.1 K2013 . . . 125
B.2 Example 5.4.1 . . . 126
B.3 Example 8.1 . . . 126

Bibliography **133**

List of Figures

2.1 Sample path of a Brownian motion. 20

5.1 The evolution of value of a death benefit policy. 67
5.2 The states and possible transitions for an orphan insurance under the stated assumptions. In the interest of time we have collected all states where the child is dead in one box. 68
5.3 The value of the orphan insurance for different states of the family. 69

7.1 Sample paths of the Merton jump-diffusion. 83
7.2 Demonstrating the survival probability, $p_{\star,\star}(25, t), t \geq 25$, using K2013. 85
7.3 The grid representing the domain where we approximate the solution. Left is $t = 0$ and right is $t = T$. Blue edges are boundary conditions and red is terminal conditions, both of which are known. 85

8.1 The value of an endowment policy where the security is a Merton jump-diffusion (L) and a Black and Scholes diffusion model (R). . 92
8.2 Pure Endowment: difference between Merton jump-diffusion and Black and Scholes diffusion. 93
8.3 Value of a death benefit with jumps (R) and without jumps (L). . 95
8.4 The value of a unit-linked pension with jumps (R) and without jumps(L). 96
8.5 FPI every week from January 1. 2006 98
8.6 The value of 1 kg farm fish in Nok at every point in the growth process. 99
8.7 The value of a pure endowment for different jump intensities. . . . 100

CHAPTER 1

Introduction

Insurance is the trading, hedging and management of risk. Humans are often observed to be risk averse. This basic human trait might help explain why insurance has such a long history. We can find evidence of insurance in as far back as ancient Babylon. For example, in Hammurabi's code there are laws against making false insurance claims (Hammurabi 1904).

The history of insurance is woven into the history of the world, from ancient Greece and Rome to China. Innovations in the insurance sector were taking place during the birth of stock markets in the Netherlands and the industrial revolution (Goetzmann 2017). Today this is a huge industry going through revolutionary changes, with new technical tools and access to data like never before.

Most of the mathematical tools the industry relies on today did not exist for most of this period. In the context of the history of insurance they are about as recent as the French Revolution. The actuarial sciences made a huge leap forward in the 20th century thanks to the works of mathematicians like Andrey Kolmogorov, Harald Cramér, Paul Lévy, and Thomas Mack. Even standing on the shoulders of such giants we can still see new lands.

It is common today to divide actuarial science into life and non-life insurance. Non-life insurance is often concerned with products such as housing or travel insurance- insurance of objects or services. The tail of the claims distribution can be heavy or light, and the duration of a contract is often short. One often tries to model both the number of claims and the severity. Life insurance, on the other hand, is different. The contracts can be very long and claims are usually tied to big events in life.

For this reason, life insurance policies are essential in many people's financial planning. Such policies can provide security in old age with pension policies, or a death benefit could provide financial stability for a family, should a provider pass.

One particularly interesting kind of policy is the unit-linked life insurance policy. This is a policy that allows the traditional life insurance to be combined with investments in a financial asset. This can provide additional risk. More risk is not necessarily a bad thing. Additional risk can mean a higher payment for a lower price.

A central question with any kind of insurance is what the value should be. The valuation of a unit-linked policy is not straightforward. In addition to the mortality risk, one also needs to consider the financial risk. This require more modelling to be done. In addition to a reasonable mortality model, one also needs a model for the price of the underlying financial asset. A realistic model for financial assets needs a financial mathematical framework.

Suppose one were able to find a way of finding a value. The natural next question is then: is there a simple way of calculating this value? In short, this is the central aim of this thesis.

Thiele's differential equation is a result from classical life insurance. It explains how the value of a life insurance policy change with time. The attraction of this result is that it is simple to calculate as a backwards equation, in particular, when using numerical methods. This result has been generalized to the study of unit-linked policies in the form of Thiele's partial differential equation (PDE). However, this result only hold in situations where the underlying security are diffusion models like Black and Scholes.

We aim at finding a Thiele's partial integro-differential equation (PIDE) for a unit-linked policy where the securities model is a jump diffusion.

1.1 General background

The pricing of unit-linked polices is nothing new. The pricing of a unit linked policy and a Thiele's PDE for a unit linked policy governed by a Black and Scholes model was done in Aase and Persson 1993.

The Black and Scholes model is a standard among many practitioners and is given by the SDE

$$\frac{dS(t)}{S(t)} = \mu dt + \sigma dW(t) \tag{1.1}$$

$$S(0) = s_0 \tag{1.2}$$

Where $\mu, \sigma, s_0 \in \mathbb{R}$, dt is the time integral and $W(t)$ is a standard Brownian motion. We will define these notions in later chapters. This model's popularity can be attributed in part to its simplicity, and that it provides a closed form solution to the pricing of some contingent claims. On the other hand, there are some large drawbacks. The returns from this distribution are a log-normally distributed random variables. Some argue that this underestimates the probability of large losses. This can be catastrophic in the context of life insurance.

One way to help this is to introduce jumps into the model. The jumps have a natural interpretation in many financial markets, they can be caused by news. News about a firm can affect the price of their stocks. Suppose the firm has received a new patent. This would make the firm more valuable, and the price of its stocks should reflect that. This could cause a jump in price. On the other hand, suppose a senior manager is arrested for embezzlement. This reflects badly on the company. It signals that the internal controls might not be good enough and that there is a possibility that senior management might not be so trustworthy. This can cause the price to fall abruptly.

To do this we can generalize Equation (1.1):

$$\frac{dS(t)}{S(t)} = \mu(t, S(t))dt + \sigma(t, S(t))dW(t) + \int_{\mathbb{R}\setminus\{0\}} h(t, S(t), z)\tilde{N}(dz, dt)$$

$$S(0) = s_0$$

Here $\tilde{N}(dz, dt)$ is a compensated Poisson measure. This expression does not only generalize the process, i.e. from a diffusion to a jump diffusion, but it also allow for a dependence on t and $S(t)$. This gives us more room in which to play within our modelling.

We will show that it is possible to price unit linked policies under this kind of securities model. Moreover, we find a Thiele's PIDE for the value, thus allowing for a simpler calculation of price.

This result is useful for multiple reasons. For one it generalizes the model one can use to model the securities. This opens the door for more realistic models that take extreme events better into account. In addition, including jump processes in our model allows us to model a bigger class of Lévy processes. Lastly, this result can find broader applications in finance and in particular commodities.

1.2 The structure of the thesis

We start the next chapter by introducing the mathematical background. We start with measure theory and use this to introduce probability theory and some of the most important stochastic processes like martingales and Markov processes.

In chapter three we turn our attention to stochastic calculus. We construct the stochastic integral for semimartingales and introduce stochastic differential equations. Then we study some of the most important results from stochastic calculus like the Itô formula and Feynman-Kac formula.

In the fourth chapter we apply some of these results and build the theoretical framework we will use for mathematical finance. We discuss portfolios, trading strategies and the notion of arbitrage. Then we introduce Girsanov's theorem and discuss market completeness and pricing of contingent claims.

In the fifth chapter, we treat classical insurance mathematics. We talk about how we use Markov processes to model the states of the insured, and how to value traditional insurance contracts with deterministic payments.

In the sixth chapter we turn our attention to unit-linked policies. First we talk about how to value the policy and then derive the Thiele's PIDE for jump diffusions in Theorem 6.2.1. This is one of the main results of the thesis.

Chapter seven is the first chapter of the second part. Here we introduce the models that we will be using in the examples and some numerical methods that can be used to solve the PIDE. The models and methods are in no way exhaustive but were chosen for their simplicity and relative usefulness.

Chapter eight is devoted to examples. We look at some examples for common policies and see how Thiele's PIDE with a Merton jump diffusion model compares to a Black and Scholes model. We finish the chapter with a sensitivity test. We test how changing the jump intensity change the value of the policy.

The ninth and final chapter is a conclusion. We summarize the thesis and provide some discussion on the results. Then we provide some self-critical remarks and finally some ideas for possible future research.

Appendix A contains proofs which are not part of the main text. These are proofs that are often very long and not critical for the main results. Appendix B contains the code used for calculations and graphs.

PART I

Theory

CHAPTER 2

Mathematical preliminaries

The purpose of this chapter is to introduce some of the necessary mathematical framework.

The heavy lifting in this thesis will be done using stochastic calculus. This is a field of mathematics concerned with the study of random phenomena. With stochastic calculus we can extend the theory of calculus from real functions to stochastic processes. This is going to be essential in this thesis.

However, this means that we have to do some work developing the theory.

2.1 An introduction to measure theory

The main results of this thesis depend on stochastic calculus and probability theory, both of which are built on measure theory. Measure theory is concerned with assigning size to sets. Sets are going to be very useful when we deal with random variables and random processes.

For a more comprehensive introduction in measure theory, one can consult Lindstrøm 2017, and for an introduction in probability theory, Øksendal 2007, or Walsh 2012.

The first thing we will do is to define the concept of the $\sigma-$algebra.

Definition 2.1.1 ($\sigma-$algebra)**.** Suppose that Ω is a non-empty set. Then a collection $\mathcal{A}$ of subsets of Ω is called a *σ-algebra* if the following properties are satisfied:

- $\emptyset \in \mathcal{A}$.
- If $A \in \mathcal{A}$, then $A^c = \Omega \setminus A \in \mathcal{A}$.
- If $\{A_n\}_{n\in\mathbb{N}}$ is a sequence of sets in $\mathcal{A}$, then $\cup_{n\in\mathbb{N}} A_n \in \mathcal{A}$.

We say that the pair $(\Omega, \mathcal{A})$ is a *measure space.*

Moreover we might also want to refer to a $\sigma-$algebra generated by some event or set A.

Proposition 2.1.2 ($\sigma-$algebra generated by an event)**.** *Let A be some subset of Ω. The smallest $\sigma-$algebra containing A exist and is on the form.*

$$\sigma(A) \coloneqq \cap_{\mathcal{A}\in\Sigma}\mathcal{A},$$

where Σ is the collection of $\sigma-$algebras containing A.

Proof. We first have to prove that this is a $\sigma-$algebra. Using Definition 2.1.1.

- By definition, $\emptyset \in \mathcal{A}$ for every $\mathcal{A}$. Since all $\mathcal{A}$s are $\sigma-$algebras. This means that $\emptyset$ also must be in the intersection, Σ, of all $\mathcal{A}$.
- Suppose $B \in \sigma(A)$, then $B \in \mathcal{A}$ for all $\mathcal{A} \in \Sigma$. However, this means that B^c must also be in all $\mathcal{A} \in \Sigma$. Hence, $B^c \in \cap_{\mathcal{A}\in\Sigma}\mathcal{A} = \sigma(A)$.
- Lastly suppose $B_1, B_2, \ldots \in \sigma(A)$, let $\mathcal{A}'$ be any $\sigma-$algebra in Σ. $B_i \in \mathcal{A}'$ for all i. This implies that $\cup_{i\in\mathbb{N}}B_i \in \mathcal{A}'$. Since this holds for any $\mathcal{A}' \in \Sigma$, then $\cup_{i\in\mathbb{N}}B_i \in \cap\mathcal{A} = \sigma(A)$.

We have now proven that $\sigma(A)$ is indeed a $\sigma-$algebra. It only remains to prove that it is the smallest $\sigma-$algebra containing A.

Suppose there exists some $\sigma-$algebra $\mathcal{C}$ that contains A which is different from $\sigma(A)$. But then there exists some $\sigma-$algebra containing A that is not in the set $\sigma(A)$. However, all such sets are in $\sigma(A)$ by assumption. Thus, this is a contradiction. ■

The $\sigma-$algebra is an excellent model for information, because, it allows us to use the notion of measurability. A set being measurable means that the set is in a $\sigma-$algebra. The set could also be in the compliment of the $\sigma-$algebra, but as we saw in the above definition, this implies that the set is also in the $\sigma-$algebra. A natural interpretation of this is that something that is measurable is a regular enough to assign a reasonable size to. Something that is not measurable, on the other hand, is too irregular for this to be possible.

We will consider models with multiple time periods. Naturally, as time progresses, more information is gained, and none is lost. In real life, information might be lost due to things like burning of great Egyptian libraries and secret police, but we ignore such things in our modelling. To model the evolution of information we will use the concept of filtrations.

Definition 2.1.3 (Filtration)**.** A sequence of $\sigma-$algebras $\{\mathcal{F}_t\}_{t\in\mathbb{R}_+}$ is called a *filtration* if for any $s \leq t$, then $\mathcal{F}_s \subseteq \mathcal{F}_t$.

We say that the filtration is *right-continuous* if

$$\mathcal{F}_t = \cap_{s<t}\mathcal{F}_s$$

Now that we have the basics of the universe in which we are working down we can start populating it. The first inhabitant we need is measurability. A way of looking at measurability is how regular a set is. Formally we can state this as follows.

Definition 2.1.4 (Measurability)**.** Let $\mathcal{F}$ be a $\sigma-$algebra on the set Ω and $f\colon \mathcal{F} \to \mathbb{R}$ a function. Now assume $A \in \mathcal{F}$ such that $f(A) \in \mathbb{R}$. We say that f is *measurable (with respect to $\mathcal{F}$)* if for all $r \in \mathbb{R}$

$$f^{-1}([-\infty, r)) \in \mathcal{F}.$$

That is,

$$\{\omega \in \Omega\colon f(\omega) < r\} \in \mathcal{F}$$

Sometimes this definition can be unnecessarily rigorous, therefore it can be useful to prove some equivalent statements about measurability.

Proposition 2.1.5 (Equivalent definition for measurability)**.** *For a given function* $f\colon \Omega \to \mathbb{R}$ *the following are equivalent.*

- f *is measurable*
- $\{\omega \in \Omega \colon f(\omega) \leq r\} \in \mathcal{F}$ *for all* $r \in \mathbb{R}$
- $\{\omega \in \Omega \colon f(\omega) > r\} \in \mathcal{F}$ *for all* $r \in \mathbb{R}$
- $\{\omega \in \Omega \colon f(\omega) \geq r\} \in \mathcal{F}$ *for all* $r \in \mathbb{R}$*,*

Proof. See the proof of Lindstrøm 2017, Proposition 7.3.4 on page 254. ∎

As stated above, we want to assign size to sets. To accomplish this, we need some way of actually doing this. Therefore, this we need to define a measure. The notion of a measure can be extended to probability measure.

Definition 2.1.6 (Probability measure)**.** A *measure* on $(\Omega, \mathcal{A})$ is a function $\rho\colon \mathcal{A} \to \mathbb{R}_+$ such that

- $\rho(\emptyset) = 0$
- If $\{A_i\}_{i\geq 1}$ is a disjoint sequence of sets from $\mathcal{A}$. Then,

$$\rho(\cup_{n\geq 1} A_n) = \sum_{n\geq 1} \rho(A_n).$$

Moreover, if P is a function $\mathcal{A} \to [0, 1]$ and the following hold,

- $P(\Omega) = 1$,

then we call P a *probability measure.*

When we put the concepts together we can define the triplet, $(\Omega, \mathcal{F}, P)$, this is a probability space. Sometimes, we might also want to specify the filtration. In that case we write $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, P)$. Here, Ω is the set of all possible outcomes, ω_i. If Ω is countable, then $\Omega = \{\omega_1, \omega_2, \ldots\}$.

Suppose we have different measures on the same space $(\Omega, \mathcal{F})$, P and Q. It is useful to have some notion of equivalence, in the sense that the two measures agree on something. We first need two technical definitions.

Definition 2.1.7 (Null set)**.** Let $(\Omega, \mathcal{F}, P)$ be a measure (or probability) space. We say that a set N is a *null set* if $P(N) = 0$. The set of null sets is denoted $\mathcal{N}_P$.

Definition 2.1.8 (Absolutely continuous)**.** Let $(\Omega, \mathcal{F})$ be a measurable space. Suppose μ and ν are measures. We say that μ is *absolutely continuous with regards to* ν if for every $N \in \mathcal{N}_\mu$ we have,

$$\mu(N) = 0 \implies \nu(N) = 0.$$

We denote this $\mu \gg \nu$, or in other words $\mathcal{N}_\mu \subseteq \mathcal{N}_\nu$.

Finally, equality of measure.

Definition 2.1.9 (Equivalent measures)**.** Let μ and ν be measures on the same measurable space $(\Omega, \mathcal{F})$. We say that μ and ν are *equivalent*, denoted $\mu \sim \nu$, if $\mu \gg \nu$ and $\nu \gg \mu$.

This has been an basic introduction to measure theory. This is a versitile branch of mathematics. For example, it can be used to construct integration theory. We are, however, going to use it for probability theory. The next section is on the basics of probability theory. We introduce some basics like independence, distribution, expectation, and characteristic functions, among others.

2.2 An introduction to probability theory

We have now defined some of the basic measure-theoretical concepts and can move on to more probabilistic matters. The first thing we need to consider is an event. An event is an element of a $\sigma-$algebra that we can assign size to. For example, in the discrete case a set of elements $A = \{\omega_1, \omega_2, \dots, \omega_n\}$ is an event. Examples of events can be an insured dying or the value of some financial security being within a certain range. When we want to find the probability of an event happening, we only need to measure the size of that event.

Since we are not dealing with only one event, we also have to consider the relationship between events. Therefore, we introduce the notion of independence.

Definition 2.2.1 (Independence)**.** Two events A and B are said to be *independent* if $P(A \cap B) = P(A)P(B)$

We can now move on to the random variable.

Definition 2.2.2 (Random variable)**.** A function $X : \Omega \to \mathbb{R}$ is called an *random variable* if it is measurable with respect to $\mathcal{F}$.

A random variable is simply a function that takes some outcome ω and maps it to a real number.

We can define the law or push forward measure.

Definition 2.2.3 (The law of a random variable)**.** The *law* of a d-dimensional random variable X, denoted $\mathcal{L}(X)$, is the image measure P_X on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d)$. The image measure, also called the push forward measure, is defined as such for any Borel set B,

$$\mu_X(B) = P_X(B) = P(\{\omega \in \Omega \colon \omega \in X^{-1}(B)\}) =$$

There is also the distribution function which is defined as follows:

Definition 2.2.4 (Distribution function)**.** A function F is called a *distribution function* for a random variable X if

$$F(x) = P(\{\omega \in \Omega \colon X(\omega) \leq x\}) = P(X \leq x)$$

The next point on our plan is the notion of expectation.

Definition 2.2.5 (Expectation)**.** If,

$$\int_\Omega |X| dP < \infty.$$

Then the *expectation* of a random variable $g(X)$ is defined,

$$\mathrm{E}(X) = \int_{\mathbb{R}} x d\mu(dx) = \int_{\Omega} X(\omega) dP(\omega)$$

Note that here the integral is in a Lebesgue sense, see Lindstrøm 2017, Chapter 7 for further details.

Characteristic function

When we are dealing with random variables and processes we often wish to compare them. We want to know if they have the same distribution, if they are independent and so on. A very handy tool in that context is the characteristic function.

We start by defining the characteristic function.

Definition 2.2.6 (Characteristic function)**.** Let X be a d-dimensional random variable with distribution μ_X. Then, the characteristic function of μ_X is

$$\varphi_X(u) = \mathrm{E}[e^{i<u,x>}] = \int_{\Omega} e^{i<u,x>} \mu_X(dx), u \in \mathbb{R}^d.$$

We often refer to the characteristic function of a random variable. By that we mean the characteristic function of the distribution of that random variable.

The characteristic function has some very useful properties. We can formalize this in a proposition.

Proposition 2.2.7 (Properties of a characteristic function)**.** *Suppose $\varphi_X(u)$ is the characteristic function for a random variable with distribution μ_X. Then the following holds.*

- $|\varphi_X(u)| \leq 1$ *for all* $u \in \mathbb{R}$
- $\varphi_X(0) = 1$
- φ_X *is uniformly continuous.*
- $\overline{\varphi_x(u)} = \varphi_X(-u)$, *i.e.* φ_X *is Hermitian.*
- $\varphi_X(u)$ *is a real function if and only if X has a symmetric distribution.*

For proof see A.1.

At this point an example might be in order.

Example 2.2.8 (Normal distribution)**.** The characteristic function of the normal distribution with mean μ and variance σ^2 is,

$$\varphi(u) = e^{i\mu u - \frac{1}{2}\sigma^2 t^2}$$

Calculating this is a technical affair, but the calculation can be found in Baldi 2017, p. 16.

We also want to find out what happens when two random variables has the same characteristic function.

Proposition 2.2.9. *Let X and Y be random variables with distributions μ_X and μ_Y and characteristic functions $\varphi_X(u)$ and $\varphi_Y(u)$. X and Y has the same distribution, i.e. $\mu_X = \mu_Y$, if and only if $\varphi_X(u) = \varphi_Y(u)$.*

See the appendix for proof A.1.

We want to see what happens with the distribution of a sequence of random variables. Therefore we have the following definition.

Definition 2.2.10 (Convergence in distribution)**.** Let X and $X_n, n = 1, 2, \ldots$ random variables with distribution functions F and F_n. Then we say that X_n converges to X in distribution if F_n converge to F for all continuity points.

Another convenient use of the characteristic function is to prove that two random variables are independent.

Proposition 2.2.11 (Characteristic functions and independence)**.** *Two random variables X, Y are independent if and only if*

$$\varphi_{X+Y}(n) = \varphi_X(u)\varphi_Y(u)$$

For proof see appendix, A.1.

This has been an introduction to the foundations of probability theory. In the next section, we are going to continue with probability theory and discuss stochastic processes.

2.3 Stochastic processes

Many of the elements of the insurance and financial models we will be using is based on stochastic processes. Therefore, we need to introduce some important concepts relating to random processes. We start by giving a basic definition.

Definition 2.3.1 (Stochastic process)**.** A *stochastic process* is a parametrized family of random variables, $X = \{X_t, t \in \mathcal{I}\}$, defined on the probability space $(\Omega, \mathcal{F}, P)$ and assuming values in $\mathbb{R}^d$. Here $\mathcal{I}$ is an index set.

Remark 2.3.2. Examples of index sets can be the positive real numbers ($\mathbb{R}_+$), natural numbers ($\mathbb{N}$), or other ordered sets.

Consider the process, $\{X_t\}_{t\in\mathcal{I}}$. For any fixed t we can consider,

$$\omega \longmapsto X_t(\omega), \quad \omega \in \Omega.$$

We can think of this as individual experiments. If we on the other hand fix ω we get,

$$t \longmapsto X_t(\omega), \quad t \in \mathcal{I}$$

This is called the path of the process. From this we can see that the random process can be considered a two variable function $X(t, \omega)$. This could lead to some measurability problems. Therefore, we want to define the notion of an adapted process.

Definition 2.3.3 (Adapted process)**.** Let $\{X_t\}_{t\in\mathcal{I}}$ be a stochastic process and $\{\mathcal{F}_t\}_{t\in\mathcal{I}}$ be a filtration. Then we say that X is *adapted* to $\mathcal{F}$ if X_t is $\mathcal{F}_t-$measurable for all $t \in \mathcal{I}$.

We might be interested in comparing different stochastic processes. Therefore, we need to define some notions of relation between processes.

Definition 2.3.4 (Modification)**.** Consider two stochastic processes X and Y. We say that X is a modification of Y if

$$P(X_t = Y_t) = 1$$

for all $t \in \mathcal{I}$.

An interpretation of this is that even if there exists some outcome for which X_t and Y_t are different, the probability of this outcome is zero.

Moreover, we have the slightly more stringent condition.

Definition 2.3.5 (Indistinguishable)**.** Consider two stochastic processes X and Y. We say that X and Y are indistinguishable if

$$P(X_t = Y_t, \forall t \in \mathcal{I}) = 1$$

Stochastic processes also has a notion of distribution. This is defined as follows.

Definition 2.3.6 (Finite-dimensional distribution)**.** Let $F_1, F_2, \dots F_k$ be Borel sets in $\mathbb{R}^n$, and X be a stochastic process. Then the *finite-dimensional distribution* of X are the measures $\mu_{t_1,t_2,\dots,t_k}$ on $\mathbb{R}^{nk}$, $k = 1, 2, \dots$ such that,

$$\mu_{t_1,t_2,\dots,t_k}(F_1 \times F_2 \times \cdots \times F_k) = P(X_{t_1} \in F_1, X_{t_2} \in F_2, \dots, X_{t_k} \in F_k), t_i \in \mathcal{I}$$

This definition says that if we have a stochastic process, then we can find a measure. On the other hand, there exists a result that states the opposite: that if we have an appropriate probability measure, then we can find a stochastic process that has that distribution. This is the famous Kolmogorov's extension theorem.

Theorem 2.3.7 (Kolmogorov's extension theorem)**.** *For all $t_1, t_2, \dots, t_k \in \mathcal{I}$, let $\nu_{t_1,t_2,\dots,t_k}$ be a probability measure on $\mathbb{R}^{nk}$. such that*

1. $\nu_{t_{\sigma(1)},t_{\sigma(2)},\dots t_{\sigma(k)}}(F_1 \times F_2 \times \cdots \times F_n) = \nu_{t_1,t_2,\dots,t_k}(F_{\sigma^{-1}(1)} \times F_{\sigma^{-1}(2)} \times \cdots F_{\sigma^{-1}(k)})$

   For all permutations σ of $\{1, 2, \dots, k\}$

2. $\nu_{t_1,t_2,\dots,t_k}(F_1 \times F_2 \times \cdots \times F_k) = \nu_{t_1,t_2,\dots,t_k,t_{k+1},\dots,t_{k+m}}(F_1 \times F_2 \times \cdots \times F_k \times \mathbb{R}^n \cdots \times \mathbb{R}^n)$*, for all $m \in \mathbb{N}$. Here the number of factors on the right hand side is $k + m$.*

Then there exists a probability space $(\Omega, \mathcal{F}, P)$ and a stochastic process X such that

$$\nu_{t_1,t_2,\dots,t_k}(F_1 \times F_2 \times \cdots \times F_k) = P(X_{t_1} \in F_1, X_{t_2} \in F_2, \cdots, X_{t_k} \in F_k)$$

For Borel sets $F_1, F_2, \cdots, F_k$.

A proof of this can be found in Tao 2011, p. 197.

Since we will be dealing with calculus, it will be useful to have some result on continuity of stochastic processes.

Theorem 2.3.8 (Kolmogorov–Centsov theorem). *Suppose X is a stochastic process on* $(\Omega, \{\mathcal{F}_t\}, P)$ *such that*

$$E[|X_t - X_s|^\alpha] \leq C|t-s|^{1+\beta}$$

for all $t, s \in \mathcal{I}$*, and positive real numbers* $\alpha, \beta,\ C$*.*

Then there exists a modification $\tilde{X}_t$ *of* X_t *such that* $\tilde{X}_t$ *is locally* $\gamma-$*Hölder continuous with* $\gamma \in [0, \frac{\beta}{\alpha})$*.*

This proof can be found in Varadhan 2001, p. 51.

Remark 2.3.9. A process Y_t is $\gamma-$Hölder continuous if the following holds.

$$P\left(\sup_{0<t-s<h} \frac{|Y_t - Y_s|}{|t-s|^\gamma} \leq \delta\right) = 1$$

At this point in the thesis, we can break up the monotony of theory by introducing one of the stars of the show, the Brownian motion.

First of let us give a definition.

Definition 2.3.10. Let $(\Omega, \mathcal{F}, P)$ be a probability space equipped with the filtration $\{F_t\}_{t\in\mathcal{I}}$. A $\mathcal{F}$*-Brownian motion* with variance σ^2 is a stochastic process X, adapted to the filtration $\mathcal{F}$ and satisfying the following.

- X has continuous paths, P-a.s.
- $X_0 = 0$, P-a.s.
- For all $0 \leq s \leq t$, the random variable $X_t - X_s$ is independent of $\mathcal{F}_s$
- For all $0 \leq s \leq t$, the random variable $X_t - X_s$ is normally distributed with mean 0 and variance $\sigma^2(t-s)$

A special case of the Brownian motion which is often used in modelling is the case where $\sigma^2 = 1$. This is sometimes referred to as the standard Brownian motion and will here be denoted by W.

The Brownian motion is named after its discoverer botanist Robert Brown. In 1892 Brown was studying the movement of pollen in water and how the particles moved seemingly randomly throughout the liquid, Øksendal 2007, p. 12. The process was further studied by Albert Einstein in 1905.

It is possible to modify the Brownian motion in a plethora of ways. This can be used to create models that are useful in different applications. For example we can define a process

$$Y_t = \mu t + \sigma W_t.$$

This is the Brownian motion with drift. This can be used to model phenomenon where there is an increasing or decreasing trend.

Another example of a modification to the Brownian motion is the geometric Brownian motion. It can be described as such,

$$Y_t = e^{\mu t + \sigma W_t}$$

This has been an introduction to stochastic processes. In the next chapter we are going to be look at a special kind of stochastic process known as martingales.

2.4 Martingales

Named after a 18th-century betting strategy, the martingales is an important class of stochastic processes. They are going to be important in the construction of stochastic integrals and in the financial mathematics to come. Before we jump into the martingales, we need to introduce the foundation on which martingales are built, the conditional expectation.

Conditional expectations

Conditional expectations is essential to the martingale as we shall see. One way to look at this is the the conditional expectation is our best guess about the future, given the information we have today. Another interpretation is that the conditional expectation is the projection of a random variable from a larger to a smaller $\sigma-$algebra.

In addition to conditional expectation we also want to define the conditional probability. One can interpret the conditional probability as the probability of an event given some set of information. For example, suppose that we know that a child has a parent with a rare genetic condition. The parent having the condition can affect the probability of the child also having the condition. Let us define the conditional probability and the conditional expectation.

Definition 2.4.1 (Conditional probability)**.** Let A and B be events. Then we say that the *conditional probability* of A given B is defined

$$P(A|B) = \frac{P(A \cap B)}{P(B)}.$$

Definition 2.4.2 (Conditional expectation)**.** Let X be an integrable random variable, i.e. $\int_\Omega |X|dP < \infty$ and let $\mathcal{F} \subseteq \Omega$ be a $\sigma-$algebra on Ω. Then the *conditional expectation* of X given $\mathcal{F}$ is an equivalence class of random variable Z, which is $\mathcal{F}$-measurable, such that for any set $A \in \mathcal{F}$,

$$\int_A X dP = \int_A Z dP.$$

We use notation $Z = \mathrm{E}[X|\mathcal{F}]$.

It is useful to know that this conditional expectation exists and is unique. For a proof of this, see Baldi 2017, p. 86.

We can now note some of the properties of the conditional expectation.

Proposition 2.4.3 (Properties of the conditional expectation)**.** *Let $a, b \in \mathbb{R}$, X and Y be integrable and independent random variables, $\mathcal{F}$ and $\mathcal{G}$ be $\sigma-$algebras such that $\mathcal{G} \subseteq \mathcal{F}$. Then, the following hold*

- *Linearity:* $\mathrm{E}[aX + bY|\mathcal{F}] = a\,\mathrm{E}[X|\mathcal{F}] + b\,\mathrm{E}[Y|\mathcal{F}]$.
- *The tower property:* $\mathrm{E}[\mathrm{E}[X|\mathcal{G}]|\mathcal{F}] = \mathrm{E}[\mathrm{E}[X|\mathcal{F}]|\mathcal{G}] = E[X|\mathcal{G}]$.
- *The law of total expectation:* $\mathrm{E}[\mathrm{E}[X|\mathcal{G}]] = E[X]$
- *If X is $\mathcal{G}-$measurable, then* $\mathrm{E}[XY|\mathcal{G}] = X\,\mathrm{E}[Y|\mathcal{G}]$ *P-a.s.*

- *Suppose that X is independent of $\mathcal{F}$. Then,*

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

- *The conditional Jensen inequality: if f is a convex function, then P.a.s.*

$$\mathrm{E}[f(X)|\mathcal{F}] \geq f(\mathrm{E}[X|\mathcal{F}]).$$

For proof see A.1.

Martingales: Definition

Now that we have conditional expectations we can finally define a martingale.

Definition 2.4.4 (Martingales)**.** Let $(\Omega, \mathcal{F}, P)$ be a probability space equipped with the filtration $\mathcal{F} = \{\mathcal{F}_t\}_{t\in\mathcal{I}}$ and X be a stochastic process. We call X a *martingale* if,

- X is adapted to a filtration $\mathcal{F}$.
- $\mathrm{E}[|X_t|] < \infty$ for all $t \in \mathcal{I}$.
- For any $s, t \in \mathcal{I}$, $s \leq t$ we have,

$$\mathrm{E}[X_t|\mathcal{F}_s] = X_s.$$

How should we interpret this kind of stochastic process? A natural interpretation is found in the historic origin of the name. A martingale is a betting strategy where the expected profits are zero. In other words, the expected gain is the same as the bet. Hence the expectation of a martingale process is the same as the value today.

This makes it very useful in the study of financial markets. The efficient market hypothesis states that all information is immediately incorporated into the price of the financial assets traded in the market, Berk and DeMarzo 2017. Therefore the expected value of the discounted price process of some risky security should be the same as the value of the security today. Supposing instead that one could expect to gain from buying a security, then obviously more traders would also wish to purchase the security. This would push the price upwards and making the price too high to yield any profits from the trade. We will later formalize this such that the discounted price process for an asset has to be a martingale.

We can also define some of the other kinds of martingales. The one defined above is the strongest notion of a martingale. There are contexts where we would wish to use weaker notions of martingales. The first two are very similar.

Definition 2.4.5 ((Sub-)Supermartingale)**.** Let X be a stochastic process. We call X a *(sub)supermartingale* if,

- X is adapted to a filtration $\mathcal{F}$.
- $\mathrm{E}[|X_t|] < \infty$

- For some $s \leq t \in \mathcal{I}$,

$$\mathrm{E}[X_t|\mathcal{F}_s](\geq) \leq X_s$$

A consequence of this definition is that we can prove that something is a martingale by proving that it is a supermartingale and a submartingale.

A slightly more general notion is the concept of a local-martingale. However, before we get to that we need to define the concept of stopping times.

Definition 2.4.6 (Stopping time)**.** Let $\mathcal{F} = \{\mathcal{F}_t\}$ be a filtration. Then the random variable $\tau\colon \Omega \mapsto [0, \infty]$ is called a *stopping time* if for every $t \in \mathcal{I}$

$$\{\tau \leq t\} \in \mathcal{F}_t.$$

In addition we can define the *stopped* $\sigma-$*algebra* $\mathcal{F}_\tau$,

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

The concept of a stopping times is quite intuitive. A stopping time is random time when some event occurs. Suppose one has never been to the aquarium in Bergen. The first time one visits it to look at the penguins is a stopping time.

Stopping times will be more important later on, but for now we move on to the fourth and penultimate kind of martingale we are going to define.

Definition 2.4.7 (Local martingale)**.** Suppose there exists a sequence of stopping times $\{\tau_n\}$ such that $\tau_n \leq \tau_{n+1}$ and $\lim_{n\to\infty} \tau_n = \infty$. A process X is called a *local martingale* if the stopped process, that is if $Y_{\tau_n} = X_{t\wedge\tau_n}$ is a martingale.

The last kind of process we will define is the semimartingale. This is going to be essential in the construction of stochastic integrals.

Definition 2.4.8 (Semimartingale)**.** Let X be an $\mathcal{F}$ adapted right continuous process with left limits, also called a càdlàg process. Then X is a *semimartingale* if

$$X_t = X_0 + A_t + M_t$$

Where A and M are càdlàg, A being a process of bounded variation and M being a local martingale.

We have now defined five different types of martingales. However martingales have interesting properties that we want to study. This is the content of the next subsection.

Martingales: Important results

The first important result we are going to study is the optional stopping theorem.

Theorem 2.4.9 (Doob's optional stopping theorem)**.** *Let X be a martingale with regards to the filtration $\mathcal{F} = \{\mathcal{F}_t\}_{t\in\mathcal{I}}$. Let τ_1 and τ_2 be bounded stopping times such that $0 \leq \tau_1 \leq \tau_2 \leq T$ for some $T \in \mathbb{R}$.*

Then,

$$\mathrm{E}[X_{\tau_2}|\mathcal{F}_{\tau_1}] = X_{\tau_1}$$

For proof, see Theorem 3.22 in Karatzas and Shreve 2014.

Let us now consider some of the properties of martingales.

Proposition 2.4.10 (properties of martingales)**.** *Let X and Y be martingales independent of each other, and $a \in \mathbb{R}$. Then the following are true,*

- $X_t + Y_t$ *is a martingale.*
- aX_t *is a martingale.*

A proof of this can be found in the appendix, A.1

Martingales: Example

In this subsection we are going to prove that the process that we introduced earlier, the Brownian motion is a martingale.

Example 2.4.11 (Brownian motion: part II)**.** Recall the Brownian motion, Definition 2.3.10. We want to prove that the Brownian motion, X, is a martingale.

- The Brownian motion is always adapted to the filtration $\mathcal{F}_t = \sigma(X_s, s \leq t)$, i.e. the filtration generated by X_s for $s \leq t$.
- $$\mathrm{E}[|X_t|] = \sqrt{\frac{2t}{\pi}}\sigma < \infty$$

  The calculation is of no particular interest so it has been omitted.
- Lastly, to the martingale property. Let $s \leq t$

  $$\begin{aligned} \mathrm{E}[X_t|\mathcal{F}_s] &= \mathrm{E}[X_t + X_s - X_s|\mathcal{F}_s] \\ &= X_s + \mathrm{E}[X_t - X_s|\mathcal{F}_s] \\ &= X_s + \mathrm{E}[X_t - X_s] \\ &= X_s. \end{aligned}$$

  Here, we have used the common trick of adding a "fruitful zero". Moreover, in the second equality we use the fact that X_s is $\mathcal{F}_s - measurable$, in the third we have used that the increment $X_t - X_s$ is independent of $\mathcal{F}_s$.

By proving this we can conclude that the Brownian motion is a martingale.

The Brownian motion is one of the most common examples of a martingale. It is also an canonical example of a Lévy process. This is the next type of process we will be looking at.

2.5 Lévy process

Lévy processes are very useful when dealing with stochastic calculus. They have properties that makes them simple to handle, yet they are a class of processes that are very versatile. We start by defining a Lévy process. Then we continue with the example of the Brownian motion. After this we introduce some important concepts like the Lévy measure, and show some important results like the Lévy-Khintchine formula and the Lévy-Itô decomposition.

Definition 2.5.1 (Lévy process)**.** A stochastic process $X = \{X_t, t \geq 0\}$ is called a *Lévy process* if the following conditions are true,

1. $X_0 = 0$, P-a.s.
2. X has independent increments.
3. X has stationary increments.
4. X is continuous in probability.

There are a few things to note with this definition.

Remark 2.5.2.

- Stationary increments are related to the distribution of the increments of the process. A process X having stationary increments is therefore equivalent to saying that for some $h > 0$ and $s \neq t$, then $X_s - X_{s+h} \stackrel{d}{=} X_t - X_{t+h}$.
- Continuous in probability means that, for some $\epsilon > 0$ and $t \in \mathbb{R}_+$, $\lim_{s \to t} P(|X_t - X_s| > \epsilon) = 0$. In other words, the probability that the increment of the Lévy process X is greater than ϵ converge to 0 when s converge to t.

Lévy processes have some useful properties that we are going to take advantage of.

Example 2.5.3 (The Brownian motion)**.** We return once again to our running example from the section on stochastic processes, the Brownian motion. Actually the Brownian motion is also a Lévy process. This can be proven fairly simply by checking the definition. Let X be a Brownian motion.

- $X_0 = 0$ a.s. follows from the definition.
- The second condition follows from Definition 2.3.10.
- The third condition also follows from Definition 2.3.10.
- Continuity of probability follows from Kolmogorov's continuity criteria, i.e. Theorem 2.3.8.

Indeed, the Brownian motion is a Lévy process! We can plot a sample path of the Brownian motion.

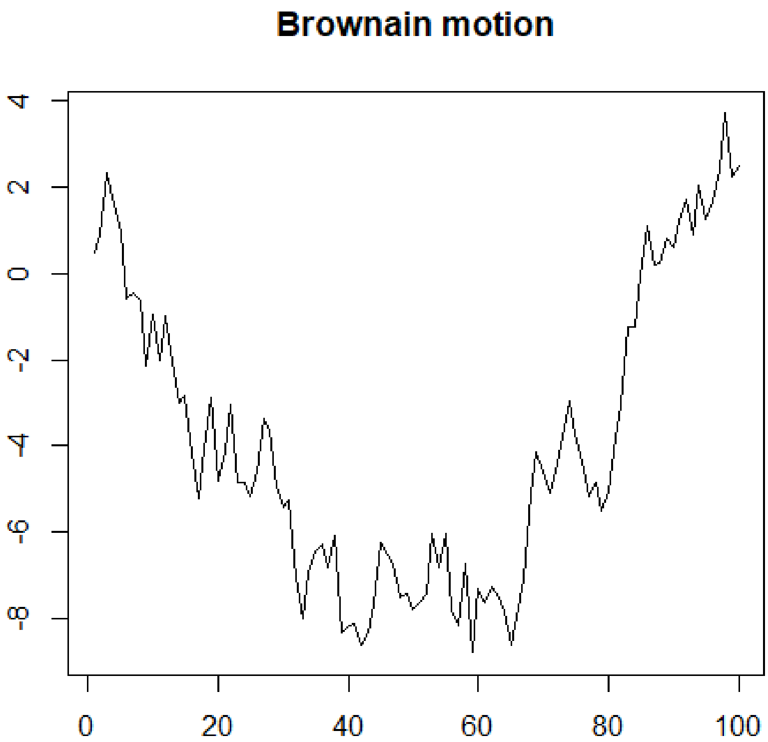

Figure 2.1: Sample path of a Brownian motion.

A very useful result concerning Lévy process is the Lévy-Khintchine formula. This formula gives us a way of calculating the characteristic function of a Lévy process. Before we get there, however, we need to define the Lévy measure and the concept of infinitely divisible distributions.

Definition 2.5.4 (Lévy measure)**.** Let ν be a Borel measure on $\mathbb{R}^d \setminus \{0\}$. We call ν a *Lévy measure* if

$$\int_{\mathbb{R}^d \setminus \{0\}} (|x|^2 \wedge 1)\nu(dx) < \infty$$

We will also need to know about infinitely divisible distributions. These kinds of distributions are closely related to Lévy processes.

Definition 2.5.5 (Infinitely divisible distributions)**.** Let X be a random variable taking values in $\mathbb{R}^d$ with law P_X. We say that X is *infinitely divisible* if for all $n \in \mathbb{N}$ there exists i.i.d. random variables Y_i such that

$$X \stackrel{d}{=} Y_1 + Y_2 + \cdots + Y_n$$

Infinitely divisible distributions and Lévy processes are closely related. Indeed, any Lévy process is infinitely divisible, see Proposition 1.3.1 in Applebaum 2009, p. 43.

Theorem 2.5.6 (The Lévy-Khintchine formula). *μ is infinitely divisible if there exist a $b \in \mathbb{R}^d$, a matrix $A \in \mathbb{R}^{d\times d}$ and a Lévy measure ν on $\mathbb{R}^d \setminus \{0\}$, such that for all $u \in \mathbb{R}^d$,*

$$\exp(\eta(u)) = \exp\bigg(i(b,u) - \frac{1}{2}(u,Au) + \int_{\mathbb{R}^d\setminus\{0\}} [e^{i(u,y)} - 1 - i(u,y)\mathbf{1}_{|y|<1}(y)]\nu(dy)\bigg)$$

On the other hand, any infinitely divisible distribution has a characteristic function of the above form.

For proof see Theorem 1.2.14 in Applebaum 2009.

We sometimes call the triplet (b, A, ν) the generating or characteristic triplet of a infinitely divisible distribution.

We stated above that any Lévy process is infinitely divisible. Hence, this formula can help us find the characteristic function of a Lévy process.

Theorem 2.5.7 (Characteristic function for a Lévy process). *If X is a Lévy process, then the characteristic function $\varphi_{X_t}(u)$ has the form*

$$\begin{aligned}\varphi_{X_t}(u) &= \exp(t\eta(u)) \\ &= \exp\bigg(t\Big[i(b,u) - \frac{1}{2}(u,Au) \\ &+ \int_{\mathbb{R}^d\setminus\{0\}} [e^{i(u,y)} - 1 - i(u,y)\mathbf{1}_{\hat{B}}(y)]\nu(dy)\Big]\bigg)\end{aligned}$$

A proof of this can also be found in Applebaum 2009, p. 44, see Theorem 1.3.3. This means that we can also use the triplet (b, A, ν) to characterize Lévy processes as well.

The last theorem we will introduce in this section is a decomposition theorem. This is very powerful when modelling a Lévy process.

Theorem 2.5.8 (Lévy-Itô decomposition). *If X is a Lévy process, then there exists a $b \in \mathbb{R}^d$, a Brownian motion B, a covariance matrix A, and an independent Poisson random measure N on $\mathbb{R}^d \times (\mathbb{R}^d \setminus \{0\})$ such that for each $t \geq 0$,*

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

Here $\tilde{N}(t,dx)$ is the compensated Poisson process as defined in Definition 2.7.5. Note that $\tilde{N}(t,dx)$ is a martingale.

See Theorem 2.4.16 Applebaum 2009, p. 126. This is a powerful result that gives us a representation for any Lévy process.

With the Lévy processes in hand, we can move on to a different type of stochastic process that will be useful, Markov processes.

2.6 Markov process

Markov processes are another class of processes that we will heavily utilize. Named after the Russian mathematician Andrey Markov, Markov processes are defined based on information, in particular, the state of the process at the present point in time.

A Markov process is a great way of modelling the state of the insured. We are therefore going to focus on the case of countable states. That is the state space $\mathcal{S}$ has a countable number of elements, or $\mathcal{S} = \{s_1, s_2, \ldots\}$.

After defining the process, we are going to prove some important results about transitional rates and transitional probabilities. We will return to Markov processes in the chapter on classical insurance.

Definition 2.6.1 (Markov process)**.** Let $X = \{X_t, t \geq 0\}$ be a stochastic process on the $(\Omega, \mathcal{F}, P)$ such that $X_t \in \mathcal{S}$. Then X is called a *Markov process* if

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

for all $t_i \in \mathcal{I}$, $t_{i_1}, t_{i_2}, \ldots t_{i_{n+1}} \in \mathcal{S}$ with $P(X_{t_1} = i_1, \ldots X_{t_n} = i_n) \neq 0$.

We sometimes refer to Markov processes as "memoryless". This is because our predictions about the future only depend on the current state of the process. Therefore, knowing what states the process has been in previously tells us nothing valuable about the future states of the process.

In order to effectively use the Markov process to model the state of the insured, we also need to have a means of modelling the probability of moving from one state to another.

Definition 2.6.2 (Transition probabilities)**.** The *transition probabilities* are the functions

$$p_{i,j}(s,t) \colon = P(X_t = j | X_s = i), \quad s \leq t, \quad i, j \in \mathcal{S}$$

This is the probability that the process is in state i at time s is in state j at time t.

This is reasonably simple and direct to calculate when dealing with discrete time. However, we are interested in continuous time. Therefore, we also need the transitional rates.

Definition 2.6.3 (Transitional rates)**.** Let $X = \{X_t, t \geq 0\}$ be a Markov process with finite state space $\mathcal{S}$. The *transitional rates* μ_i and $\mu_{j,i}$ for $i, j \in \mathcal{S}$ are the functions

$$\mu_i(t) \coloneqq \lim_{h \to 0, h > 0} \frac{1 - p_{i,i}(t, t+h)}{h}$$
$$\mu_{i,j}(t) \coloneqq \lim_{h \to 0, h > 0} \frac{p_{i,j}(t, t+h)}{h}$$

if they exist and are finite.

If in addition they are continuous in t, the process X is called *regular*.

With this in hand, we move on to some important results regarding Markov processes.

Theorem 2.6.4 (Chapman-Kolmogorov equations)**.** *Let X be a Markov process and $P(s,t)$ be a matrix of the transition probabilities. Then*

$$p_{i,j}(s,t) = \sum_{k\in\mathcal{S}} p_{i,k}(s,u)p_{k,i}(u,t)$$

For $s \leq u \leq t, i,j \in \mathcal{S}$ and $P(X_s = i), P(X_t = j) \neq 0$. In matrix notation this becomes,

$$P(s,t) = P(s,u)P(u,t)$$

Proof. We can show this directly using the laws of total probability and the definition of conditional probability.

$$\begin{aligned}
p_{i,j}(s,t) &= P(X_t = i|X_s = i) \\
&= \sum_{k\in\mathcal{S}} P(X_t = i, X_u = k|X_s = j) \\
&= \sum_{k\in\mathcal{S}} \frac{P(X_t = i, X_u = k, X_s = j)}{P(X_s = j)} \\
&= \sum_{k\in\mathcal{S}} \frac{P(X_t = i, X_u = k, X_s = j)}{P(X_s = j, X_u = k)} \frac{P(X_s = j, X_u = k)}{P(X_s = j)} \\
&= \sum_{k\in\mathcal{S}} P(X_t = i|X_u = k, X_s = j)P(X_u = k|X_s = j) \\
&= \sum_{k\in\mathcal{S}} P(X_t = i|X_u = k)P(X_u = k|X_s = j)
\end{aligned}$$

■

Another important result is the following.

Theorem 2.6.5 (Kolmogorov equation)**.** *Let $X = \{X_t, t \geq 0\}$ be a regular Markov process. Then* the backward Kolmogorov equation *is:*

$$\frac{d}{ds}p_{i,j}(s,t) = \mu_i(s)p_{i,j}(s,t) - \sum_{k\in\mathcal{S},k\neq i} \mu_{i,k}(s)p_{k,j}(s,t)$$

and the forward Kolmogorov equation *is:*

$$\frac{d}{dt}p_{i,j}(s,t) = -p_{i,j}(s,t)\mu_j(t) + \sum_{k\in\mathcal{S},k\neq j} p_{i,k}(s,t)\mu_{k,j}(t)$$

Note that by convention $\mu_i(t) = -\mu_{i,i}(t)$.

Before we look at the proofs for these equations, it can be useful to take a few moments to consider the usefulness of this result. The Kolmogorov equations give us differential equations describing the probabilities of the process moving from one state to another. This gives us a method for finding the transition probabilities through the transition rates.

Proof. We start off with the backwards equations. Let $h > 0$ such that $s \leq s + h \leq t$. We are going to find the derivative with regards to s via the definition of the derivative.

$$\begin{aligned}
\frac{d}{ds}p_{i,j}(s,t) &= \lim_{h\to 0}\frac{p_{i,j}(s+h,t)-p_{i,j}(s,t)}{h} \\
&= \lim_{h\to 0}\frac{p_{i,j}(s+h,t)-\sum_{k\in S}p_{i,k}(s,s+h)p_{k,j}(s+h,t)}{h} \\
&= \lim_{h\to 0}\frac{p_{i,j}(s+h,t)-p_{i,i}(s,s+h)p_{i,j}(s+h,t)}{h} \\
&- \lim_{h\to 0}\frac{\sum_{k\in S,k\neq i}p_{i,k}(s,s+h)p_{k,j}(s+h,t)}{h} \\
&= \lim_{h\to 0}p_{i,j}(s+h,t)\frac{(1-p_{i,i}(s,s+h)}{h} \\
&- \lim_{h\to 0}\sum_{k\in S,k\neq i}\frac{p_{i,k}(s,s+h)}{h}p_{k,j}(s+h,t) \\
&= p_{i,j}(s,t)\mu_i(s) - \sum_{k\in S,k\neq i}\mu_{i,k}(s)p_{k,j}(s,t)
\end{aligned}$$

Using similar calculations. We can find the forwards equations.

$$\begin{aligned}
\frac{d}{dt}p_{i,j}(s,t) &= \lim_{h\to 0}\frac{p_{i,j}(s,t+h)-p_{i,j}(s,t)}{h} \\
&= \lim_{h\to 0}\frac{\sum_{k\in S}p_{i,k}(s,t)p_{k,j}(t,t+h)-p_{i,j}(s,t)}{h} \\
&= \lim_{h\to 0}\frac{p_{i,j}(s,t)p_{j,j}(t,t+h)+\sum_{k\in S,k\neq j}p_{i,k}(s,t)p_{k,j}(t,t+h)-p_{i,j}(s,t)}{h} \\
&= \lim_{h\to 0}\frac{p_{i,j}(s,t)(p_{j,j}(t,t+h)-1)+\sum_{k\in S,k\neq j}p_{i,k}(s,t)p_{k,j}(t,t+h)}{h} \\
&= \lim_{h\to 0}\frac{p_{i,j}(s,t)(p_{j,j}(t,t+h)-1)+\sum_{k\in S,k\neq j}p_{i,k}(s,t)p_{k,j}(t,t+h)}{h} \\
&= -p_{i,j}(s,t)\mu_i(t) + \sum_{k\in S,k\neq i}p_{i,k}(s,t)\mu_{k,j}(t)
\end{aligned}$$

∎

Markov processes are important in a lot of applications. In this thesis Markov processes going to be used to model the state of the insured. The last section of this chapter is going to be devoted to the Poisson process.

2.7 The Poisson process

We have already introduced the Brownian motion, the canonical example of a stochastic process which is a martingale, Lévy, and so on, but this is not the only process that we will be using. Another important process is the Poisson process, and a special case of this process, the compensated Poisson process.

We will start by defining the Poisson process, and relating it to the Poisson random measure. We will introduce the compensated Poisson process.

Stochastic integrals driven by Poisson processes will be introduced in a later chapters.

Definition 2.7.1 (Homogeneous Poisson process)**.** *The homogeneous Poisson process* with intensity $\lambda > 0$ is a Lévy process N taking values in $\mathbb{N} \cup \{0\}$ such that $N(t) \sim \text{Poisson}(\lambda t)$. That is,

$$P(N(t) = n) = \frac{(\lambda t)^n}{n!} e^{-\lambda t}$$

The next stop on our journey is to relate this to the Poisson random measure. This is going to give us a much more precise grasp when we get to the stochastic integrals.

Consider first a Lévy process X. We define the jump process of X as $\Delta X(t) = X(t) - X(t-)$. How should one interpret this kind of process?

Suppose there is some Lévy process. The jump process is greater than 0 every time the process has a value at time t different from the value at the limit from the left. This means that there is a discontinuity at time t. We call a discontinuity a jump.

We are often interested in the evolution of a process over time. How does the total number of jumps over a period for such a jump process act? If we restrict ourselves a little bit, the answer is simple.

Proposition 2.7.2. *Let N be an integer valued Lévy process that is almost surely increasing. If the jump process, $\Delta N(t)$ is such that it takes values in $\{0, 1\}$, then N is a Poisson process.*

The proof of this can be found in Applebaum 2009, p. 98. The proof includes finding that the distribution of time between jumps is exponential. This is an interesting result in and of itself.

Moving on, we finally get to the Poisson random measure. Note that instead of counting every jump, we count only jumps of a certain size. This avoids the problems associated with ΔX where it is zero P-a.s. and so on. Again, for more details, see Applebaum 2009.

Definition 2.7.3 (Poisson random measure)**.** Let $B \in \mathcal{B}(\mathbb{R} \setminus \{0\})$, i.e. a Borel set, and let ΔX be a Lévy jump process. Then *the Poisson random measure on B* is defined

$$N(t, B)(\omega) = \#\{0 \leq s \leq t \colon \Delta X(s, \omega) \in B\} = \sum_{0 \leq s \leq t} \mathbf{1}_B(\Delta X(s, \omega))$$

Moreover, we define the *intensity measure over B*, $\mu(B)$,

$$\mu(B) = \int_\Omega N(1, B)(\omega) dP(\omega) = \mathrm{E}[N(1, B)].$$

A possible interpretation of $N(t, B)$ is that it is the number of jumps of a size within the interval B up to a time t. The intensity measure, on the other hand, is then the average number of jumps in one time unit, related to λ from the original definition.

The time has now come to connect the notion of a Poisson random measure and a Poisson process.

Theorem 2.7.4. *If B is bounded from below, that is $0 \notin B$, then $\{N(t,B)\}_{t\geq 0}$ is a Poisson process.*

For a proof of this, see Applebaum 2009, p. 101.

Last on the agenda for this section is to consider the compensated Poisson process. We would like to construct a stochastic integral with jump processes. This is not possible with the standard Poisson process. The Poisson process is not a martingale, therefore, we need to find some way of adjusting the Poisson process for it to become a martingale. The answer to this is the compensated Poisson process.

Definition 2.7.5 (Compensated Poisson process)**.** Let $N(t,B)$ denote a Poisson measure over a set Borel set B with an intensity measure $\mu(B)$. Then *the compensated Poisson process* is defined

$$\tilde{N}(t,B) \colon = N(t,B) - t\mu(B)$$

The final part of this section is to prove that this in fact is a martingale, whenever $\mu(B) < \infty$.

Lemma 2.7.6. *For a $B \in \mathcal{B}(\mathbb{R}\setminus\{0\})$ the compensated Poisson process $\tilde{N}(t,B)$ is a martingale under a suitable filtration.*

Proof. We need to check the definition.

- we can find an upper bound,

$$\begin{aligned} \mathrm{E}[|\tilde{N}(t,B)|] = \mathrm{E}[|\tilde{N}(t,B)|] = \mathrm{E}[|N(t,B) - t\mu(B)|] &\leq \mathrm{E}[|N(t,B)|] + \mathrm{E}[|t\mu(B)|] \\ = \mathrm{E}[N(t,B)] + t\mu(B) = 2t\mu(B) \end{aligned}$$

  The third equality follows from the fact that $N(t) \geq 0$ since it is increasing a.s. and as a Lévy process it starts at 0. Moreover, the value of the Poisson process at time t is Poisson distributed with intensity $t\mu(B)$.

- The last thing to check is the Martingale property. Let $s \leq t$,

$$\begin{aligned} \mathrm{E}[\tilde{N}(t,B)|\mathcal{F}_s] &= \mathrm{E}[N(t,B) - t\mu(B)\|mathcalF_s] \\ &= \mathrm{E}[N(t,B) - t\mu(B) + N(s,B) - N(s,B) + s\mu(B) - s\mu(B)|\mathcal{F}_s] \\ &= N(s,B) - s\mu(B) + \mathrm{E}[N(t,B) - N(s,B)|\mathcal{F}_s] - t\mu(B) + s\mu(B) \\ &= \tilde{N}(s,B) + \mathrm{E}[N(t,B) - N(s,B)|] - (t-s)\mu(B) \\ &= \tilde{N}(s,B) + \mathrm{E}[N(t-s,B)] - (t-s)\mu(B) \\ &= \tilde{N}(s,B) + (t-s)\mu(B) - (t-s)\mu(B) \\ &= \tilde{N}(s,B) \end{aligned}$$

■

With this study of the Poisson process, we are finished with the mathematical preliminaries. This has been a foundation for the following chapters. In the next chapter we will use some of these tools to construct the stochastic integral.

CHAPTER 3

Stochastic integrals

In this chapter we will develop the theory of stochastic integration and some useful results that will be used in the rest of the thesis. Since we are going to work with integrals driven by both a Brownian motion and a compound Poisson process, we need a more general integral based on semimartingales.

Before we get to that, however, we need to introduce L^p- spaces. Then we will discuss convergance, before we construct the stochastic integrals and prove some useful properties. After that we introduce some examples and one of the most important results in stochastic calculus, Itô's formula. Lastly, we will look at stochastic differential equations and the Feynman-Kac formula.

3.1 L^p-spaces

L^p-spaces are a very useful vector spaces when dealing with stochastic processes, a special case of which is the L^2-space. Convergence is the motivation for why we need L^p-spaces. We are not able to prove the convergence needed to construct stochastic integrals in any stronger sense than in L^2. It is worth noting that we are dealing with equivalence classes in L^p-spaces, and not the elements themselves.

The first thing we need to do is to define what a norm is. The norm on a vector space is how we measure distance.

Definition 3.1.1 (Norm)**.** If V is a vector space over $\mathbb{R}$, a *norm* on V is a function $||\cdot||: V \mapsto \mathbb{R}$ such that,

- $||u|| \geq 0$ with equality if and only if $u = 0$.
- $||\alpha u|| = |\alpha|||u||$ for all $\alpha \in \mathbb{R}$ and all $u \in \mathbb{R}$.
- $||u + v|| \leq ||u|| + ||v||$ for all $u, v \in V$.

The norm on theses spaces is, as the names might suggest, the L^p-norm. This is defined as such.

Definition 3.1.2 (L^p-Norm)**.** The L^p-*norm* for some constant $1 \leq p < \infty$ on the function f is defined as follows,

$$||f||_p = \left(\int_\Omega |f|^p d\mu \right)^{1/p}$$

We also have the set of L^p-functions,

$$\mathcal{L}^p = \{f : ||f||_p < \infty, f \text{ is measurable}\}$$

This integral is in the Lebesgue sense.

Example 3.1.3. Suppose X is a random variable. When dealing with random variables the above definition becomes,

$$||X||_p = \left(\int_\Omega |X|^p dP \right)^{1/p}$$

If $||X||_p < \infty$ for some p we say that $X \in L^p$.

There are some important inequalities in L^p-spaces that are going to be useful.

Theorem 3.1.4 (The Hölder inequality)**.** *Let $p, q \in (1, \infty)$ such that $\frac{1}{p} + \frac{1}{q} = 1$. Suppose that $f \in \mathcal{L}^p$ and $g \in \mathcal{L}^q$. Then, $fg \in \mathcal{L}^1$ and*

$$||fg||_1 = \int_\Omega |fg| d\mu \leq ||f||_p ||g||_q$$

Note that we can find q explicitly by solving the condition on p and q. It is $q = \frac{p}{p-1}$. This is a well know theorem, and a proof can be found in Bédos 2019 or Kuttler 2017.

Another important inequality is the Minkowski inequality. This is a corollary of the Hölder inequality.

Corollary 3.1.5 (Minkowski's inequality)**.** *Let $p \in [1, \infty)$. For all $f, g \in \mathcal{L}^p$ we have,*

$$||f + g||_p \leq ||f||_p + ||g||_p$$

For a proof see A.2

These inequalities are important in many proofs involving stochastic integrals, particularly when trying to prove convergence. We will study convergence more closely in the following section.

3.2 Convergence

Since we have now defined the L^p-norm and have yet to start on the stochastic integral part, it could be fitting to discuss convergence. There are multiple ways a sequence can converge. These modes of convergence differ from each other, although some modes imply others.

Definition 3.2.1 (Types of convergence)**.** Let X be a random variables and $\{X_n(\omega)\}_{n\in\mathbb{N}}$ be a sequence of random variables.

- If $X_n(\omega) \to X(\omega)$ for all $\omega \in \Omega$, we say that $X_n(\omega)$ *converge pointwise* to $X(\omega)$.
- If there exist an event E of probability zero such that if $\omega \notin E$ $X_n(\omega) \to X(\omega)$, we say that X_n *converge almost surely* to X.

- If $P(|X_n - X| > \epsilon) \to 0$ as $n \to \infty$, we say that X_n *converge in probability* to X.
- If $E(|X_n - X|^p) \to 0$ as $n \to \infty$, we say that X_n *converge in* L^p to X.
- A sequence of random variables $\{X_n\}$ with distribution $F_{X_n}(x)$ is said to *converge in distribution* to X if $\lim F_{X_n}(x) = F_X(x)$.

Remark 3.2.2.

- The strongest notion of convergence is pointwise. We will seldom use this one.
- L^p convergence implies convergence in probability, see Proposition 4.4 in Walsh 2012, p. 118.
- Convergence almost surely implies convergence in probability, see Proposition 4.5 in Walsh 2012, p. 118

Lastly, this leads to a couple of handy theorems about convergence and expectations.

Theorem 3.2.3 (The monotone convergence theorem)**.** *Let* $\{X_n\}_{n\in\mathbb{N}}$ *be a sequence of random variables such that* $X_n \leq X_{n+1}$ *and* $\lim_{n\to\infty} X_n = X$*, both almost surely. Then*

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

Theorem 3.2.4 (The dominated convergence theorem)**.** *Let* $\{X_n\}_{n\in\mathbb{N}}$ *be a sequence of random variables, and let* X *be a random variable. Suppose further that there exists a random variable* Y *such that* $|X_n| \leq Y$ *for all* n *almost surely,* $\lim_{n\to\infty} X_n = X$ *a.s. or in probability and that* $\mathrm{E}(Y) < \infty$*. Then,*

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

For proofs of both of these, consult Chapter 4.2 in Walsh 2012.

This has been convergence, the silence before the storm. In the next section we will use many of the concepts that we have introduced to construct stochastic integrals.

3.3 Stochastic integral: Construction

We now move over to the actual construction of the stochastic integral. We start with step functions and expand the class of integrable functions by showing how each broader class can be approximated. The problem we run into is that the approximations do not, in general, converge in any stronger sense than L^2. However, we will survive this.

The smallest class of integrands which we will define the stochastic integral for is step functions. In this section we denote the set of step functions $\mathcal{S}$.

Definition 3.3.1 (Stochastic integral for step functions)**.** Let $X = \{X_t, t \geq 0\}$ be a semimartingale and h be a random varibale. Define $Y_t = \sum_{i=1}^{n} h_i \mathbf{1}_{(T_i,T_{i+1}])}(t)$, i.e. a step function on some partition $\Pi = \{T_1, T_2, \ldots\}$. Then the *stochastic integral* is defined as

$$\int_0^t Y_s dX(s) = \sum_{i=1}^{n} h_i (X_{t\wedge T_{i+1}} - X_{t\wedge T_i})$$

This is the first baby step on the way to stochastic integration. Beginning here we can actually already prove our first important stochastic integral result.

Theorem 3.3.2 (Itô isometry)**.** *Suppose f is a step function, and X is a martingale. Then,*

$$\mathrm{E}\left[\left(\int_0^t f(s)dX_s\right)^2\right] = \mathrm{E}\left[\int_0^t |f(s)|^2 ds\right]$$

Proof. We will use the fact that for independent martingales, $\mathrm{E}[\Delta X_t \Delta X_s] = 0$ and $\mathrm{E}[\Delta X_t^2] = 1$, where we use the notation that $\Delta X_t = X_t - X_{t-1}$.

$$\begin{aligned} \mathrm{E}\left[\left(\int_0^t f(s)dX_s\right)^2\right] &= \mathrm{E}[\sum_{i=1}^{n} h_i \Delta X_i \sum_{j=1}^{n} h_j \Delta X_j] \\ &= \mathrm{E}[\sum_{i=1}^{n}\sum_{j=1}^{n} h_i h_j \Delta X_i \Delta X_j] \\ \text{(DCT)} &= \sum_{i=1}^{n}\sum_{j=1}^{n} h_i h_j \,\mathrm{E}[\Delta X_i \Delta X_j] \\ &= \begin{cases} 0, & \text{for } i \neq j \\ h_i^2, & \text{for } i = j \end{cases} \end{aligned}$$

The result follows from this last equality. ■

Note that this result has been proven for martingales, not semimartingales.

The class of step functions is not particularly useful in most applications. However we can show that many functions can be approximated by the step functions.

Consider now the following set of functions.

Definition 3.3.3. Let $B \in \mathcal{B}(\mathbb{R}^n)$. We denote $\mathcal{H}^2$ to be an equivalence class of mappings $F\colon [0,t) \times B \times \Omega \mapsto \mathbb{R}$ such that

- F is progressive, i.e. measurable on $\Omega \otimes [0,T]$.
- $\mathrm{E}[\int_0^t |F(t)|^2 dt] < \infty$.

This might seem like a cryptic definition, but the first condition ensures mesurability and the second is a boundedness condition.

Lemma 3.3.4. *$\mathcal{S}$ is dense in $\mathcal{H}^2$*

The proof of this is very involved and can be found in Applebaum 2009, p. 218.

The intuition behind this lemma is that we can approximate any function in $\mathcal{H}^2$ with a function from $\mathcal{S}$.

Since we have now learned that we can approximate any function with a sequence of step functions, we can extend the stochastic integral to also allow integrands in $\mathcal{H}^2$.

Definition 3.3.5 (Stochastic integral for $\mathcal{H}^2-$functions)**.** Suppose $X = \{X_t, t \geq 0\}$ is a semimartingale. Let $f \in \mathcal{H}^2$ and $\{f_n\}_{n\geq 1}$ be a sequence of step functions such that $f_n \to f$, i.e. $f_n \in \mathcal{S} \quad \forall n \geq 1$. Then the *stochastic integral* is defined,

$$\int_0^t f(s)dX_s = \lim_{n\to\infty} \int_0^t f_n(s)dX_s.$$

It is possible to extend this even further to the following class of functions.

Definition 3.3.6. Let $B \in \mathcal{B}(\mathbb{R}^n)$. We denote $\mathcal{P}^2$ to be an equivalence class of mappings $F\colon [0, t) \times B \times \Omega \mapsto \mathbb{R}$ such that

- F is predictable, i.e., left continous and adapted.
- $P(\int_0^t |F(t)|dt < \infty) = 1$

We also have a very similar lemma as above.

Lemma 3.3.7. *$\mathcal{H}^2$ is dense in $\mathcal{P}^2$.*

A proof of this can also be found in Applebaum 2009, p. 225.

Using a similar definition as for $\mathcal{H}^2$ functions, we can construct the integral for $\mathcal{P}^2$ functions. However, this is not requiered for this thesis.

We have now constructed the stochastic integral for a sufficiently wide class of functions or processes. In the next section, we will explore some of the properties of the stochastic integral.

3.4 Stochastic integral: Properties

We start off by proving a more general case of a previous result. We have already shown how Itô's isometry holds for step functions. However, we would like it to hold for all functions in $\mathcal{H}^2$. This presents a problem. In our framework where the integral is defined, this does not hold in general for semimartingales.

Theorem 3.4.1 (Itô isometry II)**.** *Let $f \in \mathcal{H}^2$ and X_t be a martingale. Then the following holds,*

$$\mathrm{E}\left[\left(\int_0^t f(s)dX_s\right)^2\right] = \mathrm{E}\left[\int_0^t |f(s)|^2 ds\right]$$

Proof. Let $f \in \mathcal{H}^2$ and $\{f_n\}_{n\geq 1}$ be a sequence in $\mathcal{S}$ such that $f_n \to f$ in L^2 from below. Then,

$$\begin{aligned}
\mathrm{E}\left[\left(\int_0^t f(s)dX_s\right)^2\right] &= \mathrm{E}\left[\left(\int_0^t f(s)dX_s\right)\left(\int_0^t f(s)dX_s\right)\right] \\
&= \mathrm{E}\left[\left(\lim_{n\to\infty}\int_0^t f_n(s)dX_s\right)\left(\lim_{n\to\infty}\int_0^t f_n(s)dX_s\right)\right] \\
&= \mathrm{E}\left[\left(\lim_{n\to\infty}\int_0^t f_n(s)dX_s\right)\left(\int_0^t f_n(s)dX_s\right)\right] \\
(\text{DCT}) &= \lim_{n\to\infty}\mathrm{E}\left[\left(\int_0^t f_n(s)dX_s\right)\left(\int_0^t f_n(s)dX_s\right)\right] \\
(\text{Itô isometry I}) &\leq \lim_{n\to\infty}\mathrm{E}\left[\left(\int_0^t f_n(s)^2 ds\right)\right] \\
(\text{DCT}) &= \mathrm{E}\left[\left(\lim_{n\to\infty}\int_0^t f_n(s)^2 ds\right)\right] \\
&= \mathrm{E}\left[\left(\int_0^t f(s)^2 ds\right)\right]
\end{aligned}$$

∎

We continue this section with some nice and handy properties.

Theorem 3.4.2 (Properties of the stochastic integral). *Let $f \in \mathcal{H}^2$ and let X be a semimartingale. Moreover, let $\alpha, \beta \in \mathbb{R}$ Then,*

- $\int_0^t(\alpha f + \beta g)dX_s = \alpha\int_0^t f dX_s + \beta\int_0^t g dX_s$.
- $\int_0^u f dX_s + \int_u^t f dX_s = \int_0^t f dX_s$.
- *Let* $V_t = \int_0^t Y_s dX_s$. *Then* V_t *is a semimartingale and* $\int_0^t Z_s dV_s = \int_0^t Z_s Y_s dZ_s$.
- X *is a martingale, then* $\mathrm{E}[\int_0^t f dX_t] = 0$.
- $\int_0^t f dX_s$ *is* $\mathcal{F}_t-$*measurable.*
- *If* X *is a local martingale, then* $\int_0^t f dX_s$ *is also a local martingale.*

The proof of this is very long and has been relegated to the Appendix, A.2.

There are also two important representation theorem. These allow us to represent many stochastic processes with stochastic integrals.

Theorem 3.4.3 (The Itô representation theorem). *If $X(t)$ is in $L^2(\Omega, \mathcal{F}_T, P)$, then there exists unique functions $\Phi_1(t)$ and $\Phi_2(z,t)$ such that,*

$$X(t) = \mathrm{E}[X] + \int_0^t \sigma(s)\Phi_1(t)dB(s) + \int_0^t\int_{\mathbb{R}\setminus\{0\}} h(z,t)\Phi_2(z,t)\tilde{N}(dz,dt)$$

Theorem 3.4.4 (The martingale representation theorem). *If X is a square integrable martingale, i.e.* $\mathrm{E}[X_t^2] < \infty$*, then there exists unique functions* $G(t)$ *and* $H(z,t)$ *in* $\mathcal{L}^2$ *such that,*

$$M(t) = \mathrm{E}[M(0)] + \sigma \int_0^t G(s)dB(s) + \int_0^t \int_{\mathbb{R}\setminus\{0\}} H(z,t)d\tilde{N}(dz,ds)$$

A proof of both theorems can be found in Applebaum 2009, pp. 303–304.

These have been some important results about stochastic calculus. In the next section we will look at an example of a stochastic integral.

3.5 Example: The Itô integral

We now turn our attention to a commonly used stochastic integral. The Itô integral is integrated with respect to a Brownian motion. It is a very common example of a stochastic integral, and is heavily used in practice. We usually denote the Brownian integral

$$\int_0^t f(s)dB(s)$$

where $f \in \mathcal{H}^2$ and B is a Brownian motion. The Brownian motion is a martingale, as opposed to a semimartingale. This does not pose any problems, though, as any Lévy process is also a semimartingale, see Prop. 2.7.1 in Applebaum 2009, p. 137.

The Brownian motion driven integral does have some attractive properties. For example, if the integrand f is deterministic, then the integral is a random variable with a normal distribution. We find this situation in many applications, for example the Black–Scholes model. To see this consider first the step function f.

$$\int_0^t f(s)dB(s) = \sum_i^n f(t_i)\Delta B_{t_i}$$

Here we see that we have a sum of something deterministic multiplied by an increment of the Brownian motion. The increments of the Brownian motion is a normal random variable. Moreover, the sum of independent normal random variables are also normal. The Brownian motion has independent increments, since it is a Lévy process. Let σ be the standard deviation of the Brownian motion. Then the CF of the integral is,

$$\varphi_n(u) = \mathrm{E}[\exp(iu \int_0^t f_n(s)dB(s))] = \mathrm{E}[\exp(iu \sum_i^n f_n(t_i)\Delta B_{t_i})]$$

$$\text{(independent increments)} = \prod_i^n \mathrm{E}[\exp(iuf_n(t_i)\Delta B_{t_i})]$$

$$\text{(increment normal r.v.)} = \prod_i^n \exp(\frac{1}{2}(\sigma(t_i - t_{i-1})f_n(t_i)u)^2]$$

$$= \exp(\frac{1}{2}(\sum_i^n \sigma(t_i - t_{i-1})f_n(t_i))^2u^2]$$

This is the CF of a normal random variable. Letting n go to infinity gives us the CF of the integral with deterministic $f \in \mathcal{H}^2$. Hence the integral is a normal random variable.

The Itô integral is very useful indeed. We are not done with stochastic calculus. For future results, we need the concept of quadratic variation. This is the topic of the next section.

3.6 Quadratic variation

One way to consider the "smoothness" of a function is with notion of quadratic variation. This concept is important in the Itô formula. A result we will become very familiar with later. However, the first thing we will do is give a more general formal definition for quadratic variation.

Definition 3.6.1 (CovariationI)**.** Let X and Y be semimartingales and $\Pi = \{\tau_0, \tau_1, \dots, \tau_n\}$ be a partition of stopping times. We define the *covariation of X and Y* as

$$[X,Y]_t := \lim_{|\Pi|\to 0} V_t^2(\Pi) = \lim_{|\Pi|\to 0} \sum_{j=1}^{n} |X_{t\wedge\tau_{i+1}^n} - X_{t\wedge\tau_i^n}||Y_{t\wedge\tau_{i+1}^n} - Y_{t\wedge\tau_i^n}|, \quad (3.1)$$

here the limits are in the sense that the distance between points in the partition goes to zero.

When we are dealing only with a single process we have the following.

Definition 3.6.2 (Quadratic variation)**.** Let X be a semimartingale, then X's covariation with itself is called the *quadratic variation* and is denoted,

$$[X,X]_t := [X]_t.$$

An equivalent definition can be made as follows.

Definition 3.6.3 (Covariation II)**.** Let X and Y be semimartingales. Then the *covariation of X and Y* is defined as the process,

$$[X,Y]_t = X_tY_t - \int_0^t X_{s-}dY_s - \int_0^t Y_{s-}dX_s, t \geq 0. \quad (3.2)$$

This is, in fact, an equivalent definition to Definition 3.6.1, see Theorem 23(ii) in part II chapter 6 of Protter 2005.

In some contexts, it can be usefull to separate the continuous part and the discontinuous part of the quadratic variation. We do this via the following notation,

$$[X,Y]_t = X_0Y_0 + [X,Y]_t^c + \sum_{0\leq s\leq t} \Delta X_s \Delta Y_s,$$

where $[X,Y]_t^c$ is the continuous part and $\Delta X_s \Delta Y_s$ is the discontinuous part.

The smoothness of a process is important in many applications. We need it for the next section about Itô's formula.

3.7 Itô's formula

In this section we are going to turn our eyes to one of the most important results in stochastic calculus. Itô's formula gives us a method to calculate some stochastic integral, but it can also help us find the dynamics of a stochastic process. First, however, we wish to introduce some notation.

Definition 3.7.1 (Stochastic differential). Let μ, σ and h be functions. A *stochastic differential* is on the form

$$dX(t) = \mu(t, X_t)dt + \sigma(t, X_t)dW(t) + \int_{\mathbb{R}\setminus\{0\}} h(t, X_t, z)\tilde{N}(dt, dz). \tag{3.3}$$

Where $dW(t)$ is the stochastic integral with regards to the standard Brownian motion, $\tilde{N}(dt, dz)$ is with regards to the compensated Poisson process, and dt is the time integral.

Theorem 3.7.2 (Multidimensional Itô's formula). *Let $f \in C^{1,2}([0,T] \times \mathbb{R}^d)$ and let $\mathbf{X}(t) = (X_1(t), X_2(t), \ldots, X_d(t))$ denote a d-dimensional semimartingale, if $\mathbf{X}(t)$ is on the form* (3.3) *then,*

$$\begin{aligned} f(t, \mathbf{X}_t) &= f(0, \mathbf{X}_0) + \sum_{i=1}^{d} \int_0^t \frac{\partial f}{\partial x_i}(s-, X_i(s-))dX_s^{i,c} \\ &+ \frac{1}{2} \sum_{1 \le i,j \le d} \int_0^t \frac{\partial^2 f}{\partial x_i \partial x_j} d[X_i, X_j]_s^c \\ &+ \int_0^t \left(\int_{|x|<1} f(s-, X_{s-} + h(s,x)) - f(s-, X_{s-}) \right) \tilde{N}(ds, dx) \\ &+ \int_0^t \left(\int_{|x|<1} [f(s-, X_{s-} + h(s,x)) - f(s-, X_{s-}) \right. \\ &\left. - \sum_{i=1}^{d} h_i(s,x) \frac{\partial f}{\partial x_i}(s-, X_{s-})\nu(dx) \right) ds, \end{aligned}$$

where $X_s^{i,c}$ is the continous part of X_i.

For proof see Theorem 4.4.7 in Applebaum 2009, p. 251.

A special case of this is the one-dimensional Itô formula. The proof for this one-dimentional case is a simple application of Theorem 3.7.2 with $d = 1$.

Theorem 3.7.3 (Itô's formula). *Suppose $f \in C^{1,2}([0,T] \times \mathbb{R})$. If X_t is a semimartingale on the form* (3.3) *then,*

$$\begin{aligned} f(t, X_t) &= f(0, X_0) + \int_0^t \frac{\partial f}{\partial x}(s - X_{s-})dX_s^c + \frac{1}{2}\int_0^t \frac{\partial^2 f}{\partial^2 x} d[X,X]_s^c \\ &+ \int_0^t \left(\int_{|x|<1} f(s-, X_{s-} + h(s,x)) - f(s-, X_{s-}) \right) \tilde{N}(ds, dx) \\ &+ \int_0^t \left(\int_{|x|<1} [f(s-, X_{s-} + h(s,x)) \right. \\ &\left. - f(s-, X_{s-}) - h(s,x)\frac{\partial f}{\partial x}(s-, X_{s-})\nu(dx)] \right) ds \end{aligned}$$

In addition to the Itô formula we also have a result relating to the product of two different processes.

Corollary 3.7.4 (Itô's product rule)**.** *Let X and Y be semimartingales. Then the following holds,*

$$X_tY_t = X_0Y_0 + \int_0^t X_{s-}dY_s + \int_0^t Y_{s-}dX_s + [X,Y]_t$$

Proof. This can be proven via the Itô formula for the function $f(x,y) = xy$. ∎

Itô's formula helps us in many different ways, one of which is when dealing with stochastic differential equations. It is often easier to model the change in some phenomenon than to model the object itself. This is also the case when what we want to model is random. For this purpose we use stochastic differential equations.

3.8 Stochastic differential equations

In this section we are going to develop a little bit of the theory surrounding stochastic differential equations, or SDEs. These are important tools in the modelling of many real world phenomena. One of the first things we need to consider is whether solutions to these kinds of equations exist. If the solution exists, we want to know if the solutions unique.

These questions of existence and uniqueness have been studied extensively, and luckily, well-known conditions exist to ensure this.

Consider the following SDE:

$$\mathbf{dX}(t) = \mu(s,\mathbf{X}(t))dt + \sigma(s,\mathbf{X}(t))\mathbf{dW}(t) + \int_{\mathbb{R}\setminus\{0\}} h(t, X_i(t), z)\tilde{N}(dt, dz)$$

Here $\mathbf{X}(t)$ is a multi-dimensional vector of independent semimartingales $X_i(t)$. In addition, μ and σ are d-dimensional vectors of functions.

We can formalize existence and uniqueness conditions in the following theorem.

Theorem 3.8.1 (Existence and uniqueness of SDEs)**.** *Let $T > 0$, $\mu(\cdot,\cdot)\colon [0,T] \times \mathbb{R}^d \mapsto \mathbb{R}^d$, $\sigma\colon [0,T]\times\mathbb{R}^d \mapsto \mathbb{R}^{d\times m}$ be measurable functions such that the following holds,*

- *Linear growth:* $|\mu(t,x)| + |\sigma(t,x)| + \int_{\mathbb{R}\setminus\{0\}} |h(t,x)|^2\nu(dx) \leq C(1+|x|)$, $x \in \mathbb{R}^d$, $t \in [0,T]$ *and* $C \in \mathbb{R}$.
- *Lipschitz continuity:* $|\mu(t,x)-\mu(t,y)|+|\sigma(t,x)-\sigma(t,y)|+\int_{\mathbb{R}\setminus\{0\}} |h(t,x)-h(t,y)|\nu(dx) \leq D|x-y|$, $x, y \in \mathbb{R}^d$, $t \in [0,T]$ *and* $D \in \mathbb{R}$.

Moreover, suppose Z is a random variable measurable for $\mathcal{F}_\infty$ such that $\mathrm{E}[|Z|^2] < \infty$.

Then the SDE,

$$dX(t) = \mu(s,X(t))dt + \sigma(s,X(t))dW(t) + \int_{\mathbb{R}^d\setminus\{0\}} h(z,t,X(t))\tilde{N}(dz,dt)$$

$$X_0 = Z$$

has a unique solution $X(t)$ *that is adapted to the filtration generated by* Z, $W(t)$ *and* $\tilde{N}(t,z)$ *and with* $\mathrm{E}[\int_0^T |X(t)|^2 dt]$.

This is a well known result, and a proof of a special case, where $h = 0$, can be found in Øksendal 2007, p. 70. Another proof can be found in Applebaum 2009, p. 374. We have given our own proof in Appendix, A.2.

SDEs wil be heavily featured in later chapter as models, e.g. for securities. One of the most fasinating results in stochastic calculus, in the author's humble opinion, is the subject of the next section.

3.9 The Feynman-Kac formula

We finish up this chapter by introducing the Feynman-Kac formula. The theorem was derived by the famous Richard Feynman and Mark Kac (pronounced "Kahts" Raimi 1984) while both were working at Cornell in the 40's, Kac 1985, p. 115. The formula relates certain partial integro-differential equations (PIDEs) with stochastic processes. This allows us to find a numerical solution to a PIDE by taking the pathwise average of a stochastic process. It also allows us to calculate a conditional expectation by solving a PIDE.

It is common in the literature to treat the case where the stochastic process is driven by a Brownian motion. However, we also need jump-processes. Therefore, we are going to prove a slightly generalized version of the Feynman-Kac formula.

Theorem 3.9.1 (The Feynman-Kac formula)**.** *Consider the PIDE defined,*

$$0 = \frac{\partial u(x,t)}{\partial t} + \mu(x,t)\frac{\partial u(x,t)}{\partial x} + \frac{1}{2}\sigma^2(x,t)\frac{\partial^2 u(x,t)}{\partial^2 x} \tag{3.4}$$

$$+ \int_{\mathbb{R}\setminus\{0\}} \left[u(x + h(x,t),t) - u(x,t) - h(x,t)\frac{\partial u(x,t)}{\partial x} \right] \nu(dx) \tag{3.5}$$

$$- V(x,t)u(x,t) + f(x,t) \tag{3.6}$$

Defined on all $x \in \mathbb{R}$, $t \in [0,T]$ *and with terminal condition* $u(x,T) = \psi(x)$. ψ *is a function. Then the solution* $u(x,t)$, *is given by,*

$$u(t,x) = \mathrm{E}\left[\int_t^T e^{-\int_t^u V(X_r,r)dr} f(X_u,u)du + e^{\int_t^T V(X_r,r)dr}\psi(X_T)|X_t = x \right]$$

Under the probability measure P *such that,*

$$dX_t = \mu(X_t,t)dt + \sigma(X_t,t)dB_t + \int_{\mathbb{R}\setminus\{0\}} h(t,X_t,z)\tilde{N}(dt,dz)$$

where the Brownian motion and the compensated Poisson process are pairwise independent and with initial condition $X_t = x$

The proof for existence of solution has been omitted here, but can be found in Baldi 2017, p. 317, or in Friedman 1975.

Proof. Suppose $u(x,t)$ is the solution to (3.4). Consider the process,

$$Y(t) = e^{-\int_s^u V(X_u,u)du} u(X_t,t) + \int_s^t e^{-\int_s^u V(X_r,r)dr} f(X_u,u)du$$

By Itô's product rule,

$$
\begin{aligned}
dY(t) &= d\Big(e^{-\int_s^u V(X_u,u)du}\Big)u(X_t,t) + d\Big(\int_s^t e^{-\int_s^u V(X_r,r)dr} f(X_u,u)du\Big) \\
&\quad + e^{-\int_s^u V(X_u,u)du} du(X_t,t) \\
&\quad = -V(X_t,t)e^{-\int_s^u V(X_u,u)du}u(X_t,t) + e^{-\int_s^t V(X_r,r)dr} f(X_t,t)dt \\
&\quad + e^{-\int_s^u V(X_u,u)du}\Big[\mu(X_t,t)\frac{\partial u}{\partial t}dt + \sigma(X_t,t)\frac{\partial u}{\partial t}dB(t) + \frac{1}{2}\sigma(X_t,t)^2\frac{\partial^2 u}{\partial^2 x}dt \\
&\quad + \int_{\mathbb{R}\setminus\{0\}} u(X_{s-}+h) - u(X_{s-})d\tilde{N}(dz,dt) \\
&\quad + \int_{\mathbb{R}\setminus\{0\}} (u(X_t + h(X_t,t,z),t) - u(X_s,t)) - h(X_t,t,z)\frac{\partial u}{\partial x}\nu(dz)\Big] \\
&\quad = e^{-\int_s^u V(X_u,u)du}\Big[- V(X_t,t)u(X_t,t) + f(X_t,t) + \mu(X_t,t)\frac{\partial u}{\partial t} + \frac{1}{2}\sigma^2\frac{\partial^2 u}{\partial^2 x} \\
&\quad + \int_{\mathbb{R}\setminus\{0\}} (u(X_t + h(X_t,t,z),t) - u(X_s,t)) - h(X_t,t,z)\frac{\partial u}{\partial x}\nu(dz)\Big]dt \\
&\quad + e^{-\int_s^u V(X_u,u)du}\sigma\frac{\partial u}{\partial t}dB(t) \\
&\quad + e^{-\int_s^u V(X_u,u)du}\int_{\mathbb{R}\setminus\{0\}} u(X_{s-}+h(X_t,t,z)) - u(X_{s-})d\tilde{N}(dz,dt) \\
&\quad = e^{-\int_s^u V(X_u,u)du}\sigma(X_t,t)\frac{\partial u}{\partial t}dB(t) \\
&\quad + e^{-\int_s^u V(X_u,u)du}\int_{\mathbb{R}\setminus\{0\}} u(X_{s-}+h(X_t,t,z)) - u(X_{s-})d\tilde{N}(dz,dt)
\end{aligned}
$$

The expression inside the square brackets is equal to 0 by assumption, since it is the solution to the PIDE. This leaves us with two stochastic integrals, one driven by Brownian motion, and one by a compensated Poisson. This means that we are left with a sum of two independent martingales, which obviously is also a martingale, see Proposition 2.4.10.

Now taking the integral on both sides yields,

$$
\begin{aligned}
Y(T) - Y(t) &= \int_t^T e^{-\int_s^r V(X_u,u)du}\sigma(X_r,r)\frac{\partial u}{\partial t}dB(r) \\
&\quad + \int_t^T e^{-\int_s^r V(X_u,u)du}\int_{\mathbb{R}\setminus\{0\}} u(X_{s-}+h(X_s,s,z)) - u(X_{s-})d\tilde{N}(dz,ds)
\end{aligned}
$$

Since the right side is a martingale with value at time zero of 0, we have that the conditional expectation with $X_s = x$ is,

$$
\mathrm{E}[Y(T)|X_s = x] = \mathrm{E}[Y(s)|X_s = x] = u(s,x),
$$

which means that,

$$
u(s,x) = \mathrm{E}[e^{-\int_s^T V(X_u,u)du}u(x,T) + \int_s^T e^{-\int_s^r V(X_u,u)du} f(X_r,r)dr|X_s = x]
$$

which proves the result. ■

We have now developed some of the theory of stochastic calculus and proven some of the most important results like the existence of unique solutions to stochastic differential, the Itô and the Feynman-Kac formulas. These are going to be central in the study of financial markets and insurance contracts in the coming chapters.

CHAPTER 4

Mathematical finance

The two previous chapters have been spent developing a solid theoretical framework. The examples have been limited and not particularly applied. In this chapter, however, we start applying the theory to concrete problems from the financial field. We are going to dip our toes into problems like arbitrage, portfolios, and option pricing. The subject matter could fill an entire library, so we are sadly forced to keep it short and sweet.

To avoid repetition, we note here that for the entirety of the chapter we are dealing with the filtration generated by the securities' price process. This means that we are dealing with a filtered probability space $(\Omega, \mathcal{F}_T, P)$ equipped with the filtration $\{\mathcal{F}_t\}_{t\in\mathcal{I}}$.

Moreover, we are interested in dealing with securities that are modeled by SDEs on the form:

$$dS(t) = \mu(t, S(t))S(t)dt + \sigma(t, S(t))S(t)dW(t) + S(t)\int_{\mathbb{R}\setminus\{0\}} h(t, S(t), z)\tilde{N}(dt, dz)$$

Assume that the coefficients are Lipschitz continuous and of linear growth, such that a unique solution exists.

This inclusion of jumps means that we cannot use standard or generalized Black-Scholes markets. We have to venture into Lévy-Itô markets, where valuing contingent claims might not yield a single price. We forget these problems for now and focus on the basics.

We start this chapter by introducing one of the basic components of a financial market, the security. Then we talk about portfolios and trading strategies before we venture into arbitrage theory and Girsanov's theorem. After Girsanov we talk briefly about complete markets before we finish with the pricing of contingent claims.

4.1 The financial market

A financial market is a place where market actors (often called investors) meet to exchange financial goods. These goods can often be stock or bonds, or they can be financial contracts, commodities or even weather, for more on weather derivatives see F. E. Benth, J. S. Benth, and Koekebakker 2008. We can often call the goods traded in the markets securities, assets, or funds. Financial markets are vital in a modern economy. They allow businesses, governments and even private individuals to raise cash, manage of risk, and even move income and consumption in time.

We are not generally concerned with the people in the market, but rather the goods exchanged in it. We call these good securities. We say that a market has n risky securities and a risk free security. The risk-free security, usually indexed by 0, is governed by the following dynamics,

$$dS_0(t) = r(t)S_0(t)dt$$
$$S_0(0) = 1,$$

where $r(t)$ is a deterministic rate of risk-free return. We sometimes call this security the "bank account" since there is assumed no risk of default. The bank account process is sometimes denoted $B(t)$ instead of $S_0(t)$.

The risky assets, on the other hand, are a bit more complicated. This does not mean that the risk-free security is trivial; it can be very complex. We say that our risky assets are governed by the following dynamics under P,

$$\frac{dS_i(t)}{S_i(t)} = \mu(t, S_i(t))dt + \sigma(t, S_i(t))dW(t) + \int_{\mathbb{R}\setminus\{0\}} h(t, S_i(t), z)\tilde{N}(dt, dz) \quad (4.1)$$
$$S_i(0) = S_i(0) \quad (4.2)$$

Where μ is a d-dimensional vector, σ and h are $d \times m$ matrices, $dW = (dW_1, dW_2, \ldots, dW_m)^T$ is a vector of m independent standard Brownian motions and $\tilde{N}(dt, dz) = (\tilde{N}_1(dt, dz), \tilde{N}_2(dt, dz), \ldots, \tilde{N}_d(dt, dz))^T$ is a vector of d independent Poisson processes.

In addition, we need some conditions on μ, σ and h in order to ensure that this is well defined. Since they are deterministic, this is not a problem.

The securities are not the only thing that are traded in the financial markets. A separate class of financial assets are derivatives, often called options or contingent claims. These are financial contracts that promise the owner a payment by the issuer if certain conditions apply. This is usually tied to the price of some underlying security at a certain time or period in the future. For example, these financial contracts can be used by for example insurance providers in order to be sure that they can raise enough money to pay pension obligations. An important example is the European call option.

Example 4.1.1 (The European call option)**.** Consider the i'th financial asset, a time to maturity T and an amount $K > 0$. Then the pay off of an European option written on the asset i is,

$$\max(S_i(T) - K, 0) = (S_i(T) - K)_+$$

We will try to find a price for such a contract at the end of the chapter. In order to price contingent claims, we need to consider portfolios and trading strategies. This is the next topic.

4.2 The portfolio and trading strategies

We stated in the introduction of this chapter that financial securities are bought and sold, but where did the investors keep track of these assets? What happened to them after the sale was done? The answer to this is the portfolio.

Definition 4.2.1 (Portfolio)**.** The process $\eta(t) = (\eta_0(t), \eta_1(t), \ldots, \eta_n(t))$ is called a *portfolio* or *trading strategy* if it is $\mathcal{F}$- adapted and the following hold,

- $\mathrm{E}\left[\int_0^T |\eta_0(t)r(t)S_0(t)|dt\right] < \infty.$
- $\mathrm{E}\left[\int_0^T |\eta_i \mu(t, S_i(t))S_i(t)|dt\right] < \infty$, for all $i \geq 1$.
- $\mathrm{E}\left[\int_0^T |\eta_i \sigma(t, S_i(t))S_i(t)|^2 dt\right] < \infty$, for all $i \geq 1$.
- $\mathrm{E}\left[\int_0^T \int_{\mathbb{R}\setminus\{0\}} |\eta_i h(z, t, S_i(t))S_i(t)|\nu(dz)dt\right] < \infty$, for all $i \geq 1$.

Remark 4.2.2. An assumption that has been implicit until now has been that we are allowed to trade fractional securities. This can seem unrealistic but what we are doing in reality is specifying the composition of a portfolio without specifying the size.

Another thing to note is that there is no condition of positivity on any η_i. This means that short selling and borrowing is allowed. Hence, the investor is allowed to have a negative position in a security.

The latter four conditions are rather technical, but they ensure that the stochastic calculus works out later. In particular, the third condition ensures that the Itô differential in (4.1) gives rise to a martingale. The trading strategy describes how many units of a security are held at what time.

What is the value of a portfolio?

This is a question of accounting, and we elegantly sidestep the complexities of the accounting world and assume that we can know the value of security in continuous time. Then the value of the trading strategy at time t is,

$$V_\eta(t) := \eta_0(t)S_0(t) + \sum_{i=1}^{n} \eta_i(t)S_i(t)$$

The set of trading strategies is quite extensive. Therefore, we need to narrow it down a bit. To do this we introduce some properties that we would like our trading strategies to have.

Definition 4.2.3 (Admissible trading strategy)**.** We say that a trading strategy is *admissible* if there exists some $K \in \mathbb{R}$, $K \leq 0$ such that

$$V_\eta(t) \geq K \text{a.s.}$$

The interpretation here is that the investor has a limit on how large losses can be. This can be because the person is insolvent, or that no financial institution want to extend more credit. In other words, cash is finite.

The other property that we want our trading strategy to have is related to what makes the value change. There are two ways of making the value of a trading strategy change- either the value of the security changes, or the amount invested changes. We would like the change in value to only reflect a change in value of the underlying asset. Indeed, consider what happens when we use Itô's product formula on the value of the portfolio.

$$\begin{aligned} d(V_\eta(t)) &= d(\eta_0(t)S_0(t) + \sum_{i=1}^{n} \eta_i(t)S_i(t)) \\ &= d\eta_0(t)S_0(t) + \eta_0(t)dS_0(t) + \sum_{i=1}^{n} d\eta_i(t)S_i(t) + \sum_{i=1}^{n} \eta_i(t)dS_i(t) \\ &= d\eta_0(t)S_0(t) + \eta_0(t)r_tS_0(t) + \sum_{i=1}^{n} d\eta_i(t)S_i(t) + \sum_{i=1}^{n} \eta_i(t)dS_i(t) \end{aligned}$$

The value changes if either the value of the security changes or the amount invested changes. This gives us some of the motivation for the definition of a self-financing trading strategy.

Definition 4.2.4 (Self-financing trading strategy)**.** We call a trading strategy *self-financing* if,

$$d\eta_i(t)S_i(t) = 0, \quad \forall i \geq 0$$

The practical implication of this is that cash cannot be deposited or withdrawn from the portfolio.

Moving forward, we assume that our trading strategies are self-financing and admissible. The last problem to handle in this section is the time value of money.

The reason we like to have money is that we would like to buy things from time to time, like books, beer and electricity. However, the price of the things we would like to buy changes over time, both relative to the value of money, and relative to each other. This change in the value of money has to be taken into account in our model. Therefore, we introduce the concept of discounting.

When we discount the value of something, we adjust the value of future payments such that they are comparable to payments we receive today.

A common way to do this adjustment is with the bank account. The bank account process is a magical time machine that transports our money from the world of today and into the world of tomorrow. To motivate this, suppose we are given a choice between getting a payment of a Norwegian krone today or in a year. Which has the highest value? Obviously the krone today, because we could put the krone in the bank and earn some interest on it. Thus, the value today of the payment in a year has to be the amount of money we need to put in the bank to get 1 NOK in a year. Hence the discounting factor needs to be the bank account.

It is in principle possible to use any security as a numéraire. However, the bank account process does has an advantage. It is the only process that is risk-free. For example, if one were to use the shares of a company as a numéraire, one could risk the company going bankrupt. Then its shares would no longer be traded in the markets and the numéraire would no longer exist.

We define the discounted price and value processes.

Definition 4.2.5 (Discounted Price and Value process)**.** We define the *discounted price process* for security i,

$$\tilde{S}_i(t) := e^{-\int_0^t r_s ds} S_i(t).$$

Moreover, we define the *discounted value process* for security i,

$$\tilde{V}_\eta(t) := e^{-\int_0^t r_s ds} V_\eta(t).$$

Remark 4.2.6. If the numéraire is the risk-free asset S_0, then $e^{-\int_0^t r_s ds} S_i(t) = \frac{S_i(t)}{S_0(t)}$.

This has been a basic introduction to some of the most important elements in mathematical finance. In the next section, we will tackle the problem of arbitrage.

4.3 Arbitrage

One view of a financial market is that it is a market for risk. The investors in the market buy and sell risk to each other. In this view, the only fair way to make money is by taking on a risk. For stocks, this means that one makes money by buying the stock and taking on the risk that the company will go bankrupt. If you did not buy the stock, you would instead be sitting with cash that you could spend.

In light of this opportunity cost, we would like our models to reflect the fact that one cannot make money without taking on some risk. Profit without risk is called arbitrage and we can define it mathematically.

Definition 4.3.1 (Arbitrage)**.** A self-financing trading strategy η is said to an *arbitrage strategy* if the associated value process $V_\eta(t)$ has the following properties,

- $V_\eta(0) = 0$
- $V_\eta(t) \geq 0$ a.s. for every $t \in [0, t]$.
- $P(V_\eta(T) > 0) > 0$.

There are other notions of arbitrage, but this one suffices for us.

Remark 4.3.2. One can interpret the conditions as follows:

- The value of the portfolio is zero in the beginning. This means that the purchase of securities has been wholly financed by debt.
- The value of the portfolio is never negative.
- It has to be possible for the portfolio to have a positive value at time T.
- This notion of arbitrage is a very strong one. A less strict notion is called "no free lunch with vanishing risk" (NFLVR). This is a more mathematically complicated definition, see Definition 6.1.6 in Bingham and Kiesel 1998.

In reality, arbitrage exist in financial markets, but they often only exist for a very limited time. In principle, non-arbitrage is not realistic, but for all practical purposes it does make sense. Assuming non-arbitrage also means that we can take advantage of non-arbitrage pricing.

We can formalize this kind of market.

Definition 4.3.3. A market model is called *arbitrage-free* if there does not exist any admissible self-financing trading strategy that is also an arbitrage strategy.

How do we know if this is the case? It is not generally possible to test every possible trading strategy, so what kind of condition would ensure that we are in a arbitrage free-market?

To motivate the condition, consider one of the fundamental examples of arbitrage, different prices for the same security. If a stock is traded for two different prices at different exchanges, then obviously we could make a risk-free profit by buying the stock at the lower price and sell it again at the other exchange for a higher price.

What causes a difference in price? Different investors have different view of securities. Some are optimistic and some are pessimistic, and this is a reflection of risk. If then the optimists were allowed to trade with other optimists in one exchange and pessimists with pessimists in a different exchanges, but not with each other, then obviously different prices would prevail.

The idea is to remove the different risk assessments and risk appetites, and find some probability measure for the entire market, the risk neutral probability. This will solve our problem of arbitrage.

Hopefully this story gives some kind of motivation to why we move in the direction of probability measures to solve our problems.

We start with the end and introduce the probability measure we are eventually going to use.

Definition 4.3.4 (Risk neutral probability measure)**.** We say that a probability measure, Q on $(\Omega, \mathcal{F})$ is a *risk-neutral probability measure* or alternatively *equivalent local martingale measure* (ELMM) if,

- Q is equivalent to P.
- The discounted price process, $\tilde{S}_i(t)$ is a local martingale for all securities.

If the discounted process is a martingale, then we call Q a *strong risk neutral probability measure* or *equivalent martingale measure* (EMM).

Recall the definition of equivalent measures Definition 2.1.9.

We can now use this to prove that there is no arbitrage in the market.

Theorem 4.3.5. *If there exist an ELMM, then there is no arbitrage in the market.*

This is one implication of the famous *first fundamental theorem of asset pricing*. The following proof follows the proof in Bingham and Kiesel 1998, p. 175, but this is a slightly more general setting, so a proof is included.

Proof. Let η be any trading strategy that is self-financing and admissible.

We first want to show that then the $\tilde{V}_\eta(t)$ is a supermartingale. Since $\tilde{S}$ is a local martingale, then the process $G_eta(t) = \int_0^t \eta d\tilde{S}(u)$ is also a local martingale. This process is sometimes referred to as the gains process. The value process can be written $\tilde{V}_\eta(t) = V_\eta(0) + \tilde{G}_\eta(t)$. Since η is an admissible process we have that $\tilde{V}_\eta(t) \geq K$ for some $K \in \mathbb{R}$. Since $\tilde{V}_\eta(t)$ is bounded from below it has to be a supermartingale. Indeed, let $s \leq t$, and note that $\tilde{V}_\eta(t \wedge T_n)$ is a martingale for an increasing sequence of stopping times $\{T_n\}_{n\geq 0}$ such that $T_n \to \infty$.

$$\begin{aligned}
\mathrm{E}[\tilde{V}_\eta(t)|\mathcal{F}_s] &= \mathrm{E}[\liminf_{n\to\infty} \tilde{V}_\eta(t\wedge T_n)|\mathcal{F}_s] \\
\text{(Fatou's lemma)} &\leq \liminf_{n\to\infty} \mathrm{E}[\tilde{V}_\eta(t\wedge T_n)|\mathcal{F}_s] \\
&= \liminf_{n\to\infty} \tilde{V}_\eta(s\wedge T_n) \\
&= \tilde{V}_\eta(s)
\end{aligned}$$

Here we have used Fatou's lemma, see Lindstrøm 2017, p. 266.

Suppose η is an arbitrage strategy. Then, together with the first part we have that,

$$E_Q[\tilde{V}_\eta(T)] = E_Q[\tilde{V}_\eta(T)|\mathcal{F}_0] \leq \tilde{V}_\eta(0) = 0.$$

Since η is admissible, $K \leq \mathrm{E}_Q[\tilde{V}_\eta(t)] \leq 0$. However, for η to be an arbitrage opportunity $P(\tilde{V}_\eta(T) > 0) > 0$ which is clearly not the case. ∎

As mentioned, this is one implication of the first fundamental theorem of asset pricing. The proof of the other implication is very long an complicated, but can be found in Delbaen 2006, p. 164. For the sake of completeness we also state the full theorem.

Theorem 4.3.6 (The first fundamental theorem of asset pricing)**.** *There exists an equivalent local martingale measure if and only if NFLVR holds.*

This result ensures not only that we can know that there is no arbitrate, but it also gives us a nice way of pricing cash flows. We can simply use the expected discounted cash flow. However, it does not state how we can go about finding such a probability measure. That is the goal of the next section.

4.4 The Girsanov theorem

The Girsanov theorem is fundamental to finding changes of measures that make pricing of portfolios and derivatives possible. The theorem gives us a way to check whether there is an equivalent local martingale measure and allows us to construct a dynamic for our security or portfolio under the ELMM. This in turn allows us to find the price. The Girsanov theorem is a generalization of the Cameron-Martin theorem.

In this section we are first going to introduce the stochastic exponential. Then we are going to study when the stochastic exponential is a martingale. We shall see that this is essential. Lastly, we will prove Girsanov's theorem.

The first part of our little program is to define the stochastic exponential. Catherine Doléans-Dade defined the following stochastic exponential in her 1970 paper Doléans-Dade 1970.

Definition 4.4.1 (Doléans-Dade exponential)**.** The *Doléans-Dade exponential* of a semimartingale $X(t)$, denoted $\mathcal{E}(X)(t)$ is the unique strong solution to the the SDE,

$$\begin{aligned}
dZ(t) &= Z(t-)dX(t) \\
Z(0) &= 1
\end{aligned}$$

The nice thing about these Doléan-Dade exponentials is that they work very well as Radon-Nikodym derivatives. In order to use the Doléan-Dade exponential as a change of measure, we need them to be martingales. What kind of conditions do we need for this to be the case?

We can see that, by definition, it also has the martingale property. Since X is a semimartingale by definition, the stochastic integral is also a semimartingale. In some applications, we are more interested in local martingales. Since a semimartingale is defined as the sum of a local martingale and a càdlàg process, any local martingale is also a semimartingale.

Suppose that X is a local martingale. This means that the process has the martingale property. Hence, if it is adapted and we have bounded expectation, then the Doléans-Dade exponential is a local martingale.

To find the expectation of a Doléan-Dade exponential is often difficult. The Novikov condition is a sufficient condition. However, a different strategy can also be used. Consider an SDE such as (4.1). What conditions do we need for $\mathcal{E}(S)(t)$ to be a martingale?

Proposition 4.4.2 (Martingale conditions)**.** *Suppose S_t is a well-defined SDE on the form* (4.1)*. $\mathcal{E}(S)(t)$ is a martingale if and only if,*

$$\mathrm{E}[\mathcal{E}(S)(t)] = 1$$

A proof can be found in Appendix A.3

With this in hand, we can state and prove Girsanov's theorem.

Theorem 4.4.3 (Girsanov's theorem)**.** *Let S be a Itô-Lévy process on the form* (4.1)*. Assume there exists $\theta_0(t)$, $\theta_1(t,z)$ and a predictable process $dX(t) = \theta_0(t)dW(t) + \int_{\mathbb{R}\setminus\{0\}} \theta_1(t,z)\tilde{N}(dz,dt)$ such that $dZ(t) = Z(t-)(-\theta_0(t)dW(t) - \int_{\mathbb{R}\setminus\{0\}} \theta_1(z,t)\tilde{N}(dz,dt))$ is a P-martingale.*

If in addition the following holds,

$$\sigma(t)\theta_0(t) + \int_{\mathbb{R}\setminus\{0\}} h(z,t)\theta_1(t,z)\nu(dz)dt = \mu(t),$$

then $\frac{dQ}{dP}\Big|_{\mathcal{F}_t} = Z(t)$ defines a new probability measure Q, such that $Q \sim P$ and $S(t)$ is a Q-local martingale.

In addition,

$$dW_t^Q = \theta_0(t)dt + dW_t$$
$$\tilde{N}^Q(dt,dz) = \theta_1(t,z)\nu(dz)dt + \tilde{N}(dt,dz)$$

dW_t^Q is a Q-Brownian motion and $\tilde{N}^Q(dt,dz)$ is a compensated Poisson random measure under Q.

The proof of this is very long and contains a lot of algebra. Hence, it has been relegated to the Appendix, A.3.

Remark 4.4.4. In one dimension Z_t takes the form,

$$Z_t = \exp\Bigg(- \theta_0(t)dW_t - \frac{1}{2}\theta_0(s)^2 dt$$

$$+ \int_{\mathbb{R}\setminus\{0\}} \log(1-\theta_1(t,z))\tilde{N}(dt,dz)$$
$$+ \int_{\mathbb{R}\setminus\{0\}} [\log(1-\theta_1(t,z)) + \theta_1(t,z)]\nu(dz)dt \Big).$$

One can prove this via Itô's formula.

These results are going to be very useful when we price contingent claims. What if there is is a probability measure Q, but it is not unique? The next section deals with this problem and presents the second fundamental theorem of asset pricing.

4.5 Complete markets

Until now we have only discussed whether or not there exists an equivalent martingale measure. However, what happens when there is more than one?

In order to motivate the further discussion, we need to discuss contingent claims. What should the price of such a contract or cash flow be? We already decided that we do not want arbitrage in our models. Suppose that we can replicate the cash flow by buying the underlying security. Then obviously the price of the contract needs to be the same as the value of that replicating portfolio. Otherwise we could make an arbitrage profit by buying the contract and "selling" the portfolio, or selling the contract and buying the portfolio. We will return to the question of pricing contingent claims in the next section.

In this section we will focus on the question of whether it is possible to replicate a cash flow. In furtherance of this, we define some notions regarding contingent claims.

Definition 4.5.1 (Contingent claim)**.** We say that a *contingent claim at time T* is a random variable $X > K$ such that $\mathrm{E}[X^2] < \infty$ and X is measurable with respect to $\mathcal{F}_T$, for some $K \in \mathbb{R}$. Moreover, we say that the contingent claim X is *replicable* if there exists an admissible and self-financing trading strategy $\eta(t)$ and $x \in \mathbb{R}$ such that,

$$X = X_T^{\eta,x} = x + \int_0^T \eta(t)dS(t).$$

and

$$\tilde{X}_t^{\eta,x} = x + \int_0^t \eta(u)d\tilde{S}(u)$$

is a Q-martingale.

We can note the similarities between the form of the replicating portfolio and the martingale representation Theorem (3.4.4). With this vocabulary we can define a complete market.

Definition 4.5.2 (Complete market)**.** We say that a market is complete if all contingent claims are replicable.

The second fundamental theorem of asset pricing connects these two definitions.

Theorem 4.5.3 (The second fundamental theorem of asset pricing)**.** *The following is equivalent:*

- *A market is complete*
- *Every contingent claim, which is a Q-local martingale, has the following representation,*

$$\tilde{X} = x + \int_0^T \eta(t) d\tilde{S}(t)$$

- *The ELMM Q is unique.*

A proof of this can be found in Harrison and Pliska 1983.

This result not only connects the important bits, uniqueness and completeness, but it also gives us a way to check if our market is complete. The second condition is the ticket. Using the representation we can find conditions on the market models that ensures completeness.

Theorem 4.5.4 (Conditions of market completeness)**.** *Suppose X is a contingent claim with maturity T and that the discounted claim has representation,*

$$\tilde{X} = \mathrm{E}_Q[\tilde{X}] + \int_0^T \psi(t) dB^Q(t) + \int_0^T \int_{\mathbb{R}\setminus\{0\}} \gamma(z,t) \tilde{N}^Q(dz, dt).$$

where,

$$\mathrm{E}_Q\left[\int_0^T \psi(t)^2 dt\right] < \infty$$

and

$$\mathrm{E}_Q\left[\int_0^T \int_{\mathbb{R}\setminus\{0\}} \gamma(z,t)^2 (1 - \theta_1(z,t)) \nu(dz) dt\right] < \infty.$$

Then X is replicable if and only if there exists some predictable process η such that,

$$\gamma(z,t) = e^{\int_0^t r_s ds} \eta(t) h(z,t)$$
$$\psi(t) = e^{\int_0^t r_s ds} \eta(t) \sigma(t)$$

The proof of this can be found in Appendix A.3.

From this result one could see the connection between the completeness of the market and the solvability of a set of equations. A good rule of thumb is therefore that if there are more processes than there are underlying noise, then the market is probably not complete. Lévy-Itô markets are often incomplete.

We finish this chapter with a section the pricing of contingent claims.

4.6 Pricing contingent claims

This chapter has been dedicated to building a theoretical framework for mathematical finance. We are now going to put this framework to use and price an option. In particular we are interested in finding a price for the European call option, $\Phi = \max(S(T) - K)_+$ and the corresponding hedging portfolio. This hedging portfolio can be obtained by solving a PDE.

We are operating under the probability space $(\Omega, F_T, \{\mathcal{F}_t\}, P)$, where $\{\mathcal{F}_t\}$ is the filtration generated by the Brownian motion W and the compensated Poisson process $\tilde{N}$.

In the previous section we motivated why the price of a contingent claim should be the same as the value of the replicating portfolio. The value of the contingent claim at time t needs to be the same as the conditional expectation of the discounted payoff under the risk neutral measure.

$$\pi(t) = \mathrm{E}_Q\left[\frac{S_0(t)}{S_0(T)}\Phi(S(T))\middle|\mathcal{F}_t\right]$$

Where $S_0(t)$ is the value of the bank account at time t, Φ is the payoff function of the contingent claim, and $S(T)$ is the value of the security at time T. We assume that $\mathrm{E}[\Phi(S(T))^2] < \infty$.

To keep things simple, we will consider a replicable contingent claim. It is possible to price contingent claims that are not replicable, but this requires more advanced theory. One can for example find an interval of prices. Alternatively, one can look for a portfolio that does not replicate the claim perfectly, but that comes close, either in that it minimizes risk, or it minimizes cost. For more on this, consult for example Eberlein and Kallsen 2019.

The hedging portfolio

We can first find the PDE for the hedging portfolio. This can be done via the Feynmann-Kac formula. Consider the model in (4.1) and a contingent claim with the payoff Φ and maturity T. The hedging portfolio for such a contingent claim can be found. We denote $V(t,x) = S_0(t)\,\mathrm{E}_Q[\frac{\Phi(S(T))}{S_0(T)}|\mathcal{F}_t] = S_0(t)F_T(S(t),t)$. Note first that by the Itô formula,

$$\begin{aligned} dF(t,S(t)) = {} & \frac{\partial F}{\partial t}dt + \frac{\partial F}{\partial S(t)}S(t)\mu(t,S(t))dt + \frac{1}{2}S(t)^2\sigma^2(t,S(t))\frac{\partial^2 F}{\partial^2 S_t}dt \\ & + \int_{\mathbb{R}\setminus\{0\}}(F(t,S(t)+h(z,t)) - F(t,S(t)) - h(z,t)\frac{\partial F}{\partial S(t)}\nu(dz)dt \\ & + \sigma(t,S(t))\frac{\partial F}{\partial S(t)}dW_t + \int_{\mathbb{R}\setminus\{0\}} h(z,t)S(t)\frac{\partial F}{\partial S(t)}\tilde{N}(dz,dt). \end{aligned}$$

However, this is under P, but we want the dynamics under Q. By Girsanov's theorem, $dW_t = dW_t^Q - \theta_0(t)dt$ and $\tilde{N}(dz,dt) = \tilde{N}^Q(dz,dt) - \theta_1(z,t)\nu(dz)dt$.

$$
\begin{aligned}
dF(t,S_t) &= \frac{\partial F}{\partial t}dt + \frac{\partial F}{\partial S_t}S_t r(,t)dt + \frac{1}{2}S_t^2\sigma^2(S_t,t)\frac{\partial^2 F}{\partial^2 S_t}dt \\
&+ \int_{\mathbb{R}\setminus\{0\}} (F(S_t + h(z,t),t) - F(S_t,t) - h(z,t))\frac{\partial F}{\partial S_t}S_t\nu(dz)dt \\
&- \int_{\mathbb{R}\setminus\{0\}} h(z,t)S_t\frac{\partial F}{\partial S_t}\theta_1(z,t)\nu(dz)dt - \sigma(S_t,t)\frac{\partial F}{\partial S_t}S_t\theta_0(t)dt \\
&+ \sigma(S_t,t)\frac{\partial F}{\partial S_t}dW_t^Q + \int_{\mathbb{R}\setminus\{0\}} h(z,t)S_t\frac{\partial F}{\partial S_t}\tilde{N}^Q(dz,dt) \\
&= \Bigg(\frac{\partial F}{\partial t} + \frac{\partial F}{\partial S_t}S_t \\
&\Bigg(r(t) + \int_{\mathbb{R}\setminus\{0\}} (F(S_t + h(z,t),t) - F(S_t,t) \\
&- h(z,t) - \theta_1(z,t)h(z,t))\nu(dz) - \sigma(S_t,t)\theta_0(t)\Bigg) + \frac{1}{2}S_t^2\sigma^2(S_t,t)\frac{\partial^2 F}{\partial^2 S_t}\Bigg)dt \\
&+ \sigma(S_t,t)\frac{\partial F}{\partial S_t}dW_t^Q + \int_{\mathbb{R}\setminus\{0\}} h(z,t)S_t\frac{\partial F}{\partial S_t}\tilde{N}^Q(dz,dt) \\
&= \Bigg(\frac{\partial F}{\partial t} + r_t S_t\frac{\partial F}{\partial x} + \frac{1}{2}S_t^2\sigma^2(S_t,t)\frac{\partial^2 F}{\partial^2 S_t} \\
&+ \int_{\mathbb{R}\setminus\{0\}} [F(x + h(x,t),t) - F(x,t) - S_t h(x,t)\frac{\partial F(x,t)}{\partial x}]\nu(dx)\Bigg)dt \\
&+ \sigma(S_t,t)\frac{\partial F}{\partial S_t}dW_t^Q + \int_{\mathbb{R}\setminus\{0\}} h(z,t)S_t\frac{\partial F}{\partial S_t}\tilde{N}^Q(dz,dt)
\end{aligned}
$$

Note that F is a martingale, and thus the sum of all dt terms has to be zero. The last equality follows from the fact that under the risk neutral measure, the drift of the security has to be r_t, i.e. the risk free interest. The hedging portfolio becomes,

$$
\begin{aligned}
&= \frac{\partial V}{\partial t} + r_t x\frac{\partial V}{\partial x} + \frac{1}{2}x^2\sigma(x,t)^2\frac{\partial^2 F}{\partial^2 x} \\
&+ \int_{\mathbb{R}\setminus\{0\}} [F(x + h(z,t),t) - F(x,t) - xh(z,t)\frac{\partial F(x,t)}{\partial x}]\nu(dz) \\
&= r_t V(t,x)
\end{aligned}
$$

The boundary conditions has to be $V(T,x) = \Phi(x)$. This kind of PDE can be solved by numerical methods.

The Black-Scholes formula

Most kinds of contingent claims in our framework do not give a closed form solution. They demand numerical or Monte Carlo methods in order to find a price, one special case where this is possible is the famed Black-Scholes model. The Black-Scholes model does not have jumps, and assumes a constant drift, volatility and interest rate. The SDE governing this model is,

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

$$S(0) = S_0$$

Under the ELMM this becomes,

$$\frac{dS(t)}{S(t)} = rdt + \sigma dW(t)^Q$$

We can solve this,

$$S(t) = S(0) \exp\left(\left(r - \frac{1}{2}\sigma^2\right)t - \sigma dW(t)^Q\right)$$

We can then price a call option at time t with maturity T and strike $K \in \mathbb{R}$ as such,

$$\begin{aligned}
\pi(t) &= \frac{S_0(t)}{S_0(T)} \operatorname{E}_Q\left[\Phi(S(T))|\mathcal{F}_t\right] = \frac{S_0(t)}{S_0(T)} \operatorname{E}_Q\left[(S(T)-K)_+|\mathcal{F}_t\right] \\
&= \frac{S_0(t)}{S_0(T)}\left(\operatorname{E}_Q\left[S(T)\mathbf{1}_{S(T)>K}|\mathcal{F}_t\right] - \operatorname{E}_Q\left[K\mathbf{1}_{S(T)\leq K}|\mathcal{F}_t\right]\right)
\end{aligned}$$

Consider first $\operatorname{E}_Q\left[K\mathbf{1}_{S(T)\leq K}|\mathcal{F}_t\right]$. Since the expectation of an indicator function is just the probability of that event, and the Brownian motion is a Lévy process (and thus has independent increments).

$$\begin{aligned}
\operatorname{E}_Q\left[K\mathbf{1}_{S(T)<K}|\mathcal{F}_t\right] &= K \operatorname{E}_Q\left[\mathbf{1}_{S(T)<K}|\mathcal{F}_t\right] \\
&= KQ(S(T) \leq K|\mathcal{F}_t) \\
&= KQ(S(t)\exp\left(\left(r - \frac{1}{2}\sigma^2\right)(T-t) + \int_t^T \sigma dW(s)^Q\right) \leq K)|\mathcal{F}_t) \\
&= KQ(\exp\left(\left(r - \frac{1}{2}\sigma^2\right)(T-t) + \int_t^T \sigma dW(s)^Q\right) \leq \frac{K}{S(t)})|\mathcal{F}_t) \\
&= KQ\left(\left(r - \frac{1}{2}\sigma^2\right)(T-t) + \int_t^T \sigma dW(s)^Q \leq \log\left(\frac{K}{S(t)}\right)\right)|\mathcal{F}_t\Big) \\
&= KQ\left(\left(r - \frac{1}{2}\sigma^2\right)(T-t) + \sigma\sqrt{T-t}Z \leq \log\left(\frac{K}{S(t)}\right)\right)|\mathcal{F}_t\Big) \\
&= KQ\left(\sigma\sqrt{T-t}Z \leq \log\left(\frac{K}{S(t)}\right)\right) - \left(r - \frac{1}{2}\sigma^2\right)(T-t)|\mathcal{F}_t\Big) \\
&= KQ\left(Z \leq \frac{\log\left(\frac{K}{S(t)}\right) - \left(r - \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}\right) \\
&= K\Phi\left(\frac{\log\left(\frac{K}{S(t)}\right) - \left(r - \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}\right) \\
&= K\Phi(d_1)
\end{aligned}$$

Where

$$d_1 = \frac{\log\left(\frac{K}{S(t)}\right)\Big) - \left(r - \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}},$$

$Z \sim N(0,1)$ and Φ is the cumulative distribution function for the standard normal distribution. The conditioning disappeared because the Brownian motion has independent increments, and hence the increment from t to T is independent of $\mathcal{F}_t$.

On the other hand, a similar calculation can be done for $\mathrm{E}_Q\left[S_T \mathbf{1}_{S_T>K} | \mathcal{F}_t\right]$ which gives

$$\begin{aligned}\mathrm{E}_Q\left[S(T)\mathbf{1}_{S_T>K}|\mathcal{F}_t\right] &= S(t)\Phi\left(\frac{\log\left(\frac{K}{S(t)}\right)\Big) + \left(r - \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}\right) \\ &= S(t)\Phi(d_2)\end{aligned}$$

Where,

$$d_2 = \frac{\log\left(\frac{K}{S(t)}\right)\Big) + \left(r - \frac{1}{2}\sigma^2\right)(T-t)}{\sigma\sqrt{T-t}}.$$

Hence we get the closed form formula,

$$\pi(t) = S(t)\Phi(d_2) - Ke^{-r(T-t)}\Phi(d_1).$$

This concludes the chapter on mathematical finance. There is a lot to untangle in this field, but this is hopefully a sufficient introduction to be able to use in the following chapter on insurance mathematics.

CHAPTER 5

Insurance mathematics

One of the basic functions of a financial market is the trade of risk. Nothing exemplifies this more than insurance. An insurance policy is usually a contract between two parties, the insurer and the insured. The insurer agrees that that they will pay the insured some amount of money if a specific event were to occur to the insured. This event is typically something that is disadvantageous to the insured.

Traditionally, one divides insurance into life and non-life. We are going to focus on life insurance. In life insurance, payments to the insured generally happen either when the insured is in some state of health, or when the insured moves between states. A simple example of this is a policy where the insurer pays if the insured dies. Then the states are alive and dead, and the payment occurs if the insured moves between states. We are going to model the movement between states with Markov processes.

It is possible to derive a theory for life insurance in discrete time. However, we will focus on models in continuous time.

We are going to start this chapter by introducing the basic life insurance model. We will then show how one can find the value a policy before we derive Thiele's differential equation.

5.1 The insurance model

Consider a time period from 0 to T and let $(\Omega, \mathcal{F}_T, P)$ be a complete probability space, equipped with a filtration $\mathcal{F} = \{\mathcal{F}_t\}_{t\in[0,T]}$. Consider the finite state space $\mathcal{S}$, and let $X = \{X_t, t \in [0,T]\}$ be a Markov process that describes the state of the insured. For the previous mentioned filtration, we have that $\mathcal{F}_t = \sigma(X_s, s \geq 0)$.

X_t is a Markov process in continous time. Hence recall the the transition probabilities,

$$p_{i,j}(s,t) = P(X_t = j | X_s = i)$$

and the transition rates,

$$\mu_{i,j}(t) = \lim_{h\to 0, h>0} \frac{p_{i,j}(t,t+h)}{h}, \quad i \neq j$$

$$\mu_i(t) = \lim_{h\to 0, h>0} \frac{1 - p_{i,i}(t,t+h)}{h},$$

where $s \leq t$ and $i, j \in \mathcal{S}$.

This is going to be the foundations for our mortality model. However, we also need to describe the policy and the cash flow. The eventual goal of this chapter is to find a value of the insurance contracts and how much the premiums should be.

We can define the stochastic cash flow.

Definition 5.1.1 (Stochastic cash flow)**.** A *stochastic cash flow* is a process $A = \{A_t, t \in [0, T]\}$ with paths of bounded variation P-a.s.

The idea is that the stochastic cash flow is a process that can be integrated over. The cash flow depends on the policy. We can describe the terms of functions.

Definition 5.1.2 (Policy functions)**.** Let $a_i, a_{i,j} \colon [0, \infty) \mapsto \mathbb{R}$, where $i, j \in \mathcal{S}, i \neq j$. We say that a_i and $a_{i,j}$ are *policy functions* if they model the following,

- $a_i(t)$ is the accumulated payments to the insured up to time t, given that the insured has always been in state i.
- $a_{i,j}(t)$ is the payment to the insured when the insured moves from state i to state j at time t.

We assume for now that the policy functions are deterministic.

Combining policy functions and the underlying markov chain, we can describe the policy cash flow.

Definition 5.1.3 (Policy cash flow)**.** Let $a_i(t), a_{i,j}(t) \geq 0$ be policy functions. $i, j \in \mathcal{S}$ and $i \neq j$. The *(stochastic) policy cash flow* at time t is defined:

$$A(t) := \sum_{i \in \mathcal{S}} A_i(t) + \sum_{i \neq j, i, j \in \mathcal{S}} A_{i,j}(t).$$

Where

$$\begin{aligned} A_i(t) &= \int_0^t \mathbf{1}_{\{X_s = i\}} da_i(s) \\ A_{i,j} &= \int_0^t a_{i,j}(s) dN_{i,j}^X(s) \end{aligned}$$

where $N_{i,j}^X(t) := \#\{s \in (0, t) \colon X_{s-} = i, X_s = j\}$.

The integrals are in the Riemann–Stieltjes sense.

We finish this introduction by defining the prospective value of the policy cash flow. This is a kind of present value of the stochastic cash flow.

Definition 5.1.4 (Prospective value of a stochastic cash flow)**.** Let the discount factor be defined $v(t) = 1/S_0(t)$, where S_0 is the numéraire. Moreover, let $\mathbf{1}_i^X(s) = \mathbf{1}_{\{X_s = i\}}$. We denote the *Prospective value of a (stochastic) cash flow*, $V^+(t, A)$ and define it by,

$$\begin{aligned} V^+(t, A) :=& \frac{1}{v(t)} \int_t^\infty v(s) dA(s) \\ =& \frac{1}{v(t)} \Bigg[\sum_{i \in \mathcal{S}} \int_t^\infty v(s) \mathbf{1}_i^X(s) da_i(s) \end{aligned}$$

$$+ \sum_{i,j\in\mathcal{S},j\neq i} \int_t^\infty v(s)a_{i,j}(s)dN_{i,j}^X(s)\Bigg], \quad t \geq 0.$$

The next section deals with methods for calculating this value. We are going to find both a direct method and a differential equation.

5.2 The value of insurance policies

As stated above, we want to find the value of insurance policies. The policy is described by the policy functions, and from the policy functions we get a prospective cash flow. Thus, the value of the policy should be the expected value of this cash flow given some knowledge of the state of the insured.

First some technical results.

Lemma 5.2.1. *Let $i, j, k \in \mathcal{S}$. Then if $b \in L^1$,*

$$\mathrm{E}[\int_t^\infty b(s)dN_{j,k}^X(s)|X_t = i] = \int_t^\infty b(s)p_{i,j}(t,s)\mu_{j,k}(s)ds.$$

If c is of bounded variation then,

$$\mathrm{E}[\int_t^\infty \mathbf{1}_{\{X_s=j\}}dc(s)|X_t = i] = \int_t^\infty p_{i,j}(t,s)dc(s)$$

The proof follows the same basic strategy as Koller 2012, Theorem 4.6.3. and can be found in Appendix A.4.

We can now find the explicit formula for the value of a stochastic cash flow.

It is useful to simplify notation a little bit. If x is the age of the insured at the start of the contract, then we write $p_{i,j}(x+s, x+t) = p_{i,j}^x(s,t)$.

Theorem 5.2.2. *Let x be the age of the insured at the start of the contract. Then the value if the cash flow A_j and $A_{j,k}$ is,*

$$V_i(t, A_j) = \frac{1}{v(t)} \int_t^\infty v(s)p_{i,j}^x(t,s)da_j(s)$$
$$V_i(t, A_{j,k}) = \frac{1}{v(t)} \int_t^\infty v(s)p_{i,j}^x(t,s)\mu_{j,k}^x(s)a_{j,k}(s)ds$$

Note that the subscript i in $V_i^+(t, A_i)$ indicates that the insured is in state i.

Proof. Taking the conditional expectation of the terms from Definition 5.1.4. The results follows from Lemma 5.2.1

$$\begin{aligned}
V_i(t, A_j) &= \mathrm{E}\left[\frac{1}{v(t)} \sum_{i\in\mathcal{S}} \int_t^\infty v(s)\mathbf{1}_{\{X_s=i}\}(s)da_i(s)|X_t = i\right] \\
&= \frac{1}{v(t)}\,\mathrm{E}\left[\int_t^\infty \mathbf{1}_{X_t=i}^X(s)v(s)da_i(s)|X_t = i\right] \\
&= \frac{1}{v(t)}\,\mathrm{E}\left[\int_t^\infty \mathbf{1}_{\{X_s=i}\}(s)dc(s)|X_t = i\right] \\
&= \frac{1}{v(t)} \int_t^\infty p_{i,j}(t,s)v(s)da_j(s)
\end{aligned}$$

Here we have used the fact that the bank account is deterministic and let $c(t) = \int_0^t v(s) da_j(s)$.

On the other hand we have,

$$\begin{aligned} V_i(t, A_{j,k}) &= \mathrm{E}\left[\frac{1}{v(t)} \int_t^\infty v(s) a_{j,k}(s) dN_{j,k}^X(s) | X_t = i\right] \\ &= \frac{1}{v(t)} \mathrm{E}\left[\int_t^\infty v(s) a_{j,k}(s) dN_{j,k}^X(s) | X_t = i\right] \\ &= \frac{1}{v(t)} \mathrm{E}\left[\int_t^\infty v(s) a_{j,k}(s) dN_{j,k}^X(s) | X_t = i\right] \\ &= \frac{1}{v(t)} \int_t^\infty v(s) a_{j,k}(s) p_{i,j}(t,s) \mu_{j,k}^x(s) ds \\ &= \frac{1}{v(t)} \int_t^\infty v(s) a_{j,k}(s) p_{i,j}(t,s) \mu_{j,k}^x(s) ds. \end{aligned}$$

■

This result only tells us the value of a stochastic cash flow for each state and each transition, but not for the entire insurance policy. However, the value of the policy is the sum of the value of the cash flows from all states. Hence the value of an insurance policy is just the sum of the above cash flows.

We can formalize this as a theorem.

Theorem 5.2.3 (Explicit formula for a policy)**.** *Suppose x is the age of the insured at the start of the contract. Moreover, suppose that the contract has cash flow A associated with policy functions $a_i, a_{i,j}$ for $i, j \in S, j \neq i$. Given that the insured is in state i, then the value at time t of the future payments is given by,*

$$\begin{aligned} V_i^+(t, A) = &\frac{1}{v(t)} \sum_{j \in \mathcal{S}} \int_t^\infty v(s) p_{i,j}^x(t,s) da_j(s) \\ &+ \frac{1}{v(t)} \sum_{j,k \in \mathcal{S}, k \neq j} \int_t^\infty v(s) p_{i,j}^x(t,s) \mu_{j,k}^x(s) a_{j,k}(s) ds \end{aligned}$$

Proof. This proof follows the fact that the value of a sum of cash flows is the same as the combined cash flow.

$$\begin{aligned} V_i^+(t, A) &= V_i^+\left(t, \sum_{j \in \mathcal{S}} A_j + \sum_{i,j \in \mathcal{S}, i \neq j} A_{i,j}\right) \\ &= \sum_{j \in \mathcal{S}} V_j^+(t, A_j) + \sum_{i,j, i \neq j} V_i^+(t, A_{j,k}) \\ &= \sum_{j \in \mathcal{S}} \frac{1}{v(t)} \int_t^\infty v(s) p_{i,j}^x(t,s) da_j(s) \\ &+ \sum_{i,j \in \mathcal{S}, i \neq j} \frac{1}{v(t)} \int_t^\infty v(s) p_{i,j}^x(t,s) \mu_{j,k}^x(s) a_{j,k}(s) ds \end{aligned}$$

The last equality follow from Theorem 5.2.2. ■

It is in generally nice to know if a value exist, and it is even better if there is a way to find the value. Here we have both. However, it is often hard to calculate the above expressions, because of the dependence on $p_{i,j}(t, s)$. If there are multiple states, for example, there might not exist an explicit expression for the transition probabilities. In that case, we might rely on numerical methods. However, we are going to choose a different strategy. In order to get an expression that is easier to calculate, we are going to introduce Thiele's differential equation.

5.3 Thiele's differential equation

Finding explicit values of cash flows at a time t is a very powerful tool. However, this power is diminished by the computational difficulties associated with the transition probabilities. It is possible to use numerical methods to solve this, but it is complicated.

Hence, if would be preferable if we were to find a way to calculate the value of the cash flows using only the transition rates. This is what Thiele's differential equation gives us.

The basic idea is that we want to find a differential equation to model the value of a policy. When dealing with differential equations, we also need to find the boundary conditions. Since we are interested in finding the value today, but often know the value of the policy at the end, we need to find a backward equation. Mathematically, this is not less desirable, although the intuition might be a little harder. This means that we are going to calculate the value of a policy today by starting at the end of the contract and working our way forwards. In the end we arrive at the value today.

We first need a technical lemma.

Lemma 5.3.1. *Let us denote $W_i^+(t) = v(t)V_i^+(t)$. Then for $t \leq u$,*

$$\begin{aligned} W_i^+(t) = \sum_{j \in \mathcal{S}} \Bigg[& p_{i,j}(t,u) W_j^+(u) \\ & + \int_t^u v(s) p_{i,j}(t,s) da_j(s) \\ & + \int_t^u v(s) p_{i,j}^x(t,s) \sum_{k \in \mathcal{S}, k \neq j} \mu_{j,k}^x(s) a_{j,k}(s) ds \Bigg] \end{aligned}$$

The proof of this has been relegated to Appendix A.4.

Using this lemma, we can prove Thiele's equation. We are going to prove two versions of the equation.

Theorem 5.3.2 (Thiele's differential equation I)**.** *Assume that $a_i(t) = \int_0^t a_i'(s)ds$. Let x be the age of the insured at the start of the contract and that time t has passed since the start of the contract. Then,*

$$\begin{aligned} \frac{d}{dt} W_j^+(t) = & -v(t) \Bigg[a_j'(t) + \sum_{k \in \mathcal{S}, k \neq j} \mu_{j,k}(t) a_{j,k}(t) \Bigg] + \mu_j(t) W_j^+(t) \\ & - \sum_{k \in \mathcal{S}, k \neq j} \mu_{j,k}(t) W_k^+(t) \end{aligned}$$

Moreover, the differential equation for $V_i^+(t)$ is

$$\frac{d}{dt}V_i^+(t) = r(t)V_i^+(t) - \left[a_i'(t) + \sum_{k\in\mathcal{S},k\neq j} \mu_{j,k}(t)a_{j,k}(t)\right] + \mu_j(t)V_i^+(t) - \sum_{k\in\mathcal{S},k\neq j} \mu_{j,k}(t)V_k^+(t)$$

This proof follow the proof found in, Sande 2021, p. 67

Proof. Since $a_i(t) = \int_0^t a_i'(s)ds$, W_i^+ is continuous. Recall Lemma 5.3.1 and let $u = t+h$.

$$\begin{aligned} W_i^+(t) = &\sum_{j\in\mathcal{S}} p_{i,j}^x(t,t+h)W_j^+(t+h) \\ &+ \sum_{j\in\mathcal{S}}\Bigg[\int_t^{t+h} v(s)p_{i,j}^x(t,s)da_j(s) \\ &+ \int_t^{t+h} v(s)p_{i,j}^x(t,s)\sum_{k\in\mathcal{S},k\neq j}\mu_{j,k}^x(s)a_{j,k}(s)ds\Bigg]. \end{aligned}$$

Taking the difference $W_i^+(t+h) - W_i^+(t)$ gives us,

$$\begin{aligned} W_i^+(t+h) - W_i^+(t) &= W_i^+(t+h) - \sum_{j\in\mathcal{S}} p_{i,j}^x(t,t+h)W_j^+(t+h) \\ &- \sum_{j\in\mathcal{S}}\Bigg[\int_t^{t+h} v(s)p_{i,j}^x(t,s)da_j(s) \\ &+ \int_t^{t+h} v(s)p_{i,j}^x(t,s)\sum_{k\in\mathcal{S},k\neq j}\mu_{j,k}^x(s)a_{j,k}(s)ds\Bigg] \\ &= W_i^+(t+h)(1 - p_{i,i}^x(t,t+h)) \\ &- \sum_{j\in\mathcal{S},j\neq i} p_{i,j}^x(t,t+h)W_j^+(t+h) \\ &- \sum_{j\in\mathcal{S}}\Bigg[\int_t^{t+h} v(s)p_{i,j}^x(t,s)da_j(s) \\ &+ \int_t^{t+h} v(s)p_{i,j}^x(t,s)\sum_{k\in\mathcal{S},k\neq j}\mu_{j,k}^x(s)a_{j,k}(s)ds\Bigg]. \end{aligned}$$

We want to find the derivative of this expression. Hence, we divide this expression by h and find the limit when $h \to 0$.

$$\begin{aligned} \frac{W_i^+(t+h) - W_i^+(t)}{h} = &\frac{W_i^+(t+h)(1 - p_{i,i}^x(t,t+h))}{h} \\ &- \frac{\sum_{j\in\mathcal{S},j\neq i} p_{i,j}^x(t,t+h)W_j^+(t+h)}{h} \\ &- \frac{\sum_{j\in\mathcal{S}}\int_t^{t+h} v(s)p_{i,j}^x(t,s)da_j(s)}{h} \end{aligned}$$

$$
\begin{aligned}
&-\frac{\sum_{j\in\mathcal{S}}\int_t^{t+h} v(s)p_{i,j}^x(t,s)\sum_{k\in\mathcal{S},k\neq j}\mu_{j,k}^x(s)a_{j,k}(s)ds}{h}\\
&=\frac{d}{dt}W_i^+(t)=W_i^+(t)\mu_i^x(t)-\sum_{j\in\mathcal{S},j\neq i}\mu_{i,j}^x(t)W_j^+(t)\\
&-v(t)\left[a_i'(t)+\sum_{k\in\mathcal{S},k\neq i}\mu_{i,k}^x(t)a_{i,k}(t)\right]
\end{aligned}
$$

This follows from the definition of $\mu_{i,i}(t)$, the fundamental theorem of calculus, and the fact that,

$$
p_{i,j}(t,t+h)\to\begin{cases}0, \text{if } i\neq j\\ 1, \text{if } i=j\end{cases}
$$

when $h\to 0$.

As for the last expression,

$$
\begin{aligned}
\frac{d}{dt}W_i^+(t)&=\frac{d}{dt}v(t)V_i^+(t)+v(t)\frac{d}{dt}V_i^+(t)\\
v(t)\frac{d}{dt}V_i^+(t)&=\frac{d}{dt}W_i^+(t)+r(t)v(t)V_i^+(t)\\
\frac{d}{dt}V_i^+(t)&=r(t)V_i^+(t)+V_i^+(t)\mu_i^x(t)-\sum_{j\in\mathcal{S},j\neq i}\mu_{i,j}^x(t)V_j^+(t)-a_i'(t)\\
&-\sum_{k\in\mathcal{S},k\neq i}\mu_{i,k}^x(t)a_{i,k}(t)
\end{aligned}
$$

■

We can generalize this result a little bit. In particular, many common insurance policies have cash flows that are such that $a_i(s)$ has a discontinuity. An example of this is an endowment policy where the insured receives a lump sum of money when they reach a certain age, if they are still alive. To handle this kind of policy, we have the following theorem.

Theorem 5.3.3 (Thiele's differential equation II)**.** *Let x be the aged of the insured at the start of the contract, t be the time the contract has been in force and T be the time the contract ends. Assume that a_i is differentiable a.e. with at most one discontinuity at time T. Then,*

$$
\frac{d}{dt}V_i^+(t)=r(t)V_i^+(t)-a_i'(t)-\sum_{j\in\mathcal{S},j\neq i}\mu_{i,j}^x(t)(a_{i,j}(t)+V_j^+(t)-V_i^+(t))
$$

with $V_i(T)^+=a_i(T)-a_i(T-)$.

This proof follows the proof of Theorem 5.6 in Sande 2021.

Proof. Since we accept a.e. differentiation with a discontinuity at the end point $t=T$, we have that for all integrable f.

$$
\int_{(t,T]} f(s)da_i(s)=f(T)(a_i(T)-a_i(T-))+\int_t^T f(s)a_i'(s)ds.
$$

Now define,

$$\theta_j^x(s) = a_j'(a) + \sum_{k\in\mathcal{S},k\neq j} \mu_{j,k}^x(s)a_{j,k}(s).$$

The explicit expression for the prospective reserves gives,

$$\begin{aligned}
V_i^+(t) &= \frac{1}{v(t)}\Bigg[\sum_{j\in\mathcal{S}} p_{i,j}^x(t,T)(a(T)-a(T-)) \\
&+ \sum_{j\in\mathcal{S}}\int_t^T v(s)p_{i,j}^x(t,s)(a_j'(s) + \sum_{k\in\mathcal{S},k\neq j}\mu_{j,k}^x(s)a_{j,k}(s))ds\Bigg] \\
&= \frac{1}{v(t)}\Bigg[\sum_{j\in\mathcal{S}} p_{i,j}^x(t,T)(a(T)-a(T-)) \\
&+ \sum_{j\in\mathcal{S}}\int_t^T v(s)p_{i,j}^x(t,s)\theta_j^x(s)ds\Bigg]
\end{aligned}$$

We can introduce the following notation,

$$\begin{aligned}
G_i^T(t) &= v(T)\sum_{j\in\mathcal{S}} p_{i,j}^x(t,T)(a_j(T)-a_j(T-)) \\
F_i^s(t) &= v(t)\sum_{j\in\mathcal{S}} p_{i,j}(t,s)\theta_j^x(s)
\end{aligned}$$

which means that,

$$v(t)V_i^+(t) = G_i^T(t) + \int_t^T F_i^s(t)ds$$

We can differentiate both G and F.

$$\begin{aligned}
\frac{d}{dt}G_i^T &= v(T)\sum_{j\in\mathcal{S}} \frac{d}{dt}p_{i,j}^x(t,T)(a_j(T)-a_j(T-)) \\
\frac{d}{dt}F_i^s(t) &= v(t)\sum_{j\in\mathcal{S}} \frac{d}{dt}p_{i,j}^x(t,s)\theta_j^x(s).
\end{aligned}$$

By the Kolmogorov backward equation from Theorem 2.6.5,

$$\begin{aligned}
\frac{d}{dt}p_{i,j}^x(t,s) &= \mu_i^x(t)p_{i,j}^x(t,s) - \sum_{k\in\mathcal{S},k\neq i}\mu_{i,k}^x(t)p_{k,j}^x(t,s) \\
&= \sum_{k\in\mathcal{S},k\neq i}\mu_{i,k}^x(t)(p_{i,j}^x(t,s) - p_{k,j}^x(t,s))
\end{aligned}$$

This is because $\mu_i^x(t) = -\mu_{i,i}^x(t) = \sum_{k\neq i}\mu_{i,k}^x(t)$.

However this means that,

$$\frac{d}{dt}F_i^s(t) = v(t)\sum_{j\in\mathcal{S}}\Bigg(\sum_{k\in\mathcal{S},k\neq i}\mu_{i,k}^x(t)(p_{i,j}^x(t,s) - p_{k,j}^x(t,s)\Bigg)\theta_j^x(s)$$

$$
\begin{aligned}
&= v(t)\sum_{j\in\mathcal{S}}\sum_{k\in\mathcal{S},k\neq i}\mu_{i,k}^x(t)p_{i,j}^x(t,s)\theta_j^x(s) \\
&- v(t)\sum_{j\in\mathcal{S}}\sum_{k\in\mathcal{S},k\neq i}\mu_{i,k}^x(t)p_{k,j}^x(t,s)\theta_j^x(s) \\
&= \sum_{k\in\mathcal{S},k\neq i}\mu_{i,k}^x(t)\Bigg(v(t)\sum_{j\in\mathcal{S}}p_{i,j}^x(t,s)\theta_j^x(s) - v(t)\sum_{j\in\mathcal{S}}p_{k,j}^x(t,s)\theta_j^x(s)\Bigg) \\
&= \sum_{k\in\mathcal{S},k\neq i}\mu_{i,k}^x(t)\Bigg(F_i^s(t) - F_k^s(t)\Bigg). \qquad (5.1)
\end{aligned}
$$

Moreover,

$$
\begin{aligned}
\frac{d}{dt}G_i^T &= v(T)\sum_{j\in\mathcal{S}}\Bigg(\sum_{k\in\mathcal{S},k\neq i}\mu_{i,k}^x(t)(p_{i,j}^x(t,T) - p_{k,j}^x(t,T))\Bigg)(a_j(T) - a_j(T-)) \\
&= \sum_{k\in\mathcal{S},k\neq i}\mu_{i,k}^x(t)\Bigg(v(T)\sum_{j\in\mathcal{S}}p_{i,j}^x(t,T)(a_j(T) - a_j(T-)) \\
&- v(T)\sum_{j\in\mathcal{S}}p_{k,j}^x(t,T))(a_j(T) - a_j(T-))\Bigg) \\
&= \sum_{k\in\mathcal{S},k\neq i}\mu_{i,k}^x(t)(G_i^T(t) - G_k^T(t)). \qquad (5.2)
\end{aligned}
$$

But now consider the function

$$
H(x,y) = \int_x^T F_i^s(y)ds.
$$

Taking the partial derivatives,

$$
\begin{aligned}
\frac{\partial}{\partial x}H(x,y) &= -F_i^x(y), \\
\frac{\partial}{\partial y}H(x,y) &= \int_x^T \frac{d}{dy}F_i^s(y)ds.
\end{aligned}
$$

In this proof we are interested in $\frac{d}{dt}H(t,t)$. Hence, by the chain rule.

$$
\begin{aligned}
\frac{d}{dt}H(t,t) &= \frac{\partial}{\partial x}H(x,y)\bigg|_{(x,y)=(t,t)} + \frac{\partial}{\partial y}H(x,y)\bigg|_{(x,y)=(t,t)} \\
&= -F_i^t(t) + \int_t^T \frac{d}{dy}F_i^s(t)ds.
\end{aligned}
$$

Letting $s \to t$ from above we can note that,

$$
\begin{aligned}
F_i^t(t) &= \lim_{s\to t}F_i^s(t) = v(t)\sum_{j\in\mathcal{S}}p_{i,j}^x(t,t)\theta_j^x(t) \\
&= v(t)\theta_i^x(t) = v(t)a_i'(t) + v(t)\sum_{k\in\mathcal{S},k\neq i}\mu_{i,k}^x(t)a_{i,k}(t).
\end{aligned}
$$

Now on one hand, by differentiation,

$$\begin{aligned}\frac{d}{dt}(v(t)V_i^+(t)) &= \frac{d}{dt}(v(t))V_i^+(t) + v(t)\frac{d}{dt}(V_i^+(t)) \\ &= -r(t)v(t)V_i^+(t) + v(t)\frac{d}{dt}(V_i^+(t)).\end{aligned}$$

On the other hand,

$$\begin{aligned}\frac{d}{dt}(v(t)V_i^+(t)) &= \frac{d}{dt}\left(G_i^T(t) + \int_t^T F_i^s(t)ds\right) \\ &= \frac{d}{dt}G_i^T(t) - v(t)a_i'(t) - v(t)\sum_{k\in\mathcal{S},k\neq i}\mu_{i,k}^x(t)a_{i,k}(t) \\ &+ \int_t^T \frac{d}{dt}F_i^s(t)ds.\end{aligned}$$

These two equations have to be equal and hence rearranging gives,

$$\begin{aligned}\int_t^T \frac{d}{dt}F_i^s(t)ds &= -r(t)v(t)V_i^+(t) + v(t)\frac{d}{dt}(V_i^+(t)) \\ &- \frac{d}{dt}G_i^T(t) + v(t)a_i'(t) + v(t)\sum_{k\in\mathcal{S},k\neq i}\mu_{i,k}^x(t)a_{i,k}(t).\end{aligned}$$

Integrating (5.1) and inserting (5.2) gives the equivalent expression.

$$\begin{aligned}\int_t^T \frac{d}{dt}F_i^s(t)ds &= \int_t^T \sum_{k\in\mathcal{S},k\neq i}\mu_{i,k}^x(t)\Big(F_i^s(t) - F_k^s(t)\Big)ds \\ &= \int_t^T v(t)\sum_{k\in\mathcal{S},k\neq i}\mu_{i,k}^x(t)\Big(F_i^s(t) - F_k^s(t) + G_i^T(t) - G_i^T(t) \\ &+ G_k^T(t) - G_k^T(t)\Big)ds \\ &= v(t)\sum_{k\in\mathcal{S},k\neq i}\mu_{i,k}^x(t)\Big(\int_t^T F_i^s(t)ds \\ &- \int_t^T F_k^s(t)ds + G_i^T(t) - G_i^T(t) + G_k^T(t) - G_k^T(t)\Big) \\ &= v(t)\sum_{k\in\mathcal{S},k\neq i}\mu_{i,k}^x(t)\Big(V_i^+(t) - V_k^x(t) - G_i^T(t) + G_k^T(t)\Big) \\ &= v(t)\sum_{k\in\mathcal{S},k\neq i}\mu_{i,k}^x(t)\Big(V_i^+(t) - V_k^+(t)\Big) \\ &- v(t)\sum_{k\in\mathcal{S},k\neq i}\mu_{i,k}^x(t)\Big(G_i^T(t) - G_k^T(t)\Big) \\ &= v(t)\sum_{k\in\mathcal{S},k\neq i}\mu_{i,k}^x(t)\Big(V_i^+(t) - V_k^+(t)\Big) - \frac{d}{dt}G_i^T.\end{aligned}$$

Hence, we get the equation,

$$\int_t^T \frac{d}{dt}F_i^s(t)ds = B(t)\sum_{k\in\mathcal{S},k\neq i}\mu_{i,k}^x(t)\Big(V_i^+(t)-V_k^+(t)\Big)-\frac{d}{dt}G_i^T$$

$$= -r(t)B(t)V_i^+(t)+B(t)\frac{d}{dt}(V_i^+(t))$$

$$-\frac{d}{dt}G_i^T(t)+B(t)a_i'(t)+B(t)\sum_{k\in\mathcal{S},k\neq i}\mu_{i,k}^x(t)a_{i,k}(t).$$

Since $B(t)\neq 0$ for all $t\geq 0$ we have that,

$$\sum_{k\in\mathcal{S},k\neq i}\mu_{i,k}^x(t)\Big(V_i^x(t)-V_k^x(t)\Big)$$

$$= -r(t)V_i^+(t)+\frac{d}{dt}V_i^+(t)$$

$$+a_i'(t)+\sum_{k\in\mathcal{S},k\neq i}\mu_{i,k}^x(t)a_{i,k}(t).$$

Rearranging this we get that,

$$\frac{d}{dt}(V_i^+(t)) = r(t)V_i^+(t)-a_i'(t)-\sum_{k\in\mathcal{S},k\neq i}\mu_{i,k}^x(t)a_{i,k}(t)$$

$$+\sum_{k\in\mathcal{S},k\neq i}\mu_{i,k}^x(t)\Big(V_i^+(t)-V_k^+(t)\Big)$$

$$= r(t)V_i^+(t)-a_i'(t)-\sum_{k\in\mathcal{S},k\neq i}\mu_{i,k}^x(t)\Big(a_{i,k}(t)+V_k^x(t)-V_i^x(t)\Big)$$

This is the equation that we want.

■

We have now developed some theory on life insurance mathematics. This theory is going to be crucial in the next chapter when we tackle unit-linked policies. Before we get there, through, we are going to take a look at some examples of life insurance policies.

5.4 Examples of life insurance policies

In this section we are going to tackle some examples. One example is very simple, and the other a bit more advanced.

We have not discussed mortality models directly, but they are essential in applications. Modeling the mortality of the insured is a huge topic in itself, but will not be given much space in this thesis. To put it shortly, the mortality model tells us about $\mu_{i,k}$, i.e. it governs the transition rates, and hence also the transition probabilities and the prospective reserves.

In these examples we will use "K2013" as our mortality model, Finanstilsynet 2013. This is a standard model developed by the Financial Supervisory Authority of Norway (Finanstilsynet), the purpose of which is to give a minimum level of mortality. This means that no insurers can have lower reserves than what K2013 gives. We will return to K2013 in the chapter on applications.

Example 5.4.1 (Death benefit)**.** A death benefit is a life insurance policy that pays out if the insured dies during the contract. This is a very common form of insurance. Examples of uses can be to secure a family's finances if the insured is the breadwinner. Another common use is for business partners having a policy on each other. If one of them dies, then the other would have the cash to buy the other's part of a company from the heirs.

In this context we only have two states $S = \{\star, \dagger\}$, where $\star$ indicates that the insured is alive and $\dagger$ indicates that the insured is dead. Obviously the person can move from $\star$ to $\dagger$, but not from $\dagger$ to $\star$. Therefore, $\mu_{\dagger,\star} = 0$. On the other hand, $\mu_{\star,\dagger}$ is determined by K2013.

Suppose the insured is a man at the age of 24. The death benefit is $B = 100\,000$ NOK if the insured dies before the age of 70. This means that we have the following policy functions,

$$a_{\star,\dagger}(t) = \begin{cases} 100\,000 NOK, \text{ if } t < 70 \\ 0, \text{ else} \end{cases}$$

All other policy functions are zero.

Suppose that the risk free interest rate $r = 3\%$ and note that $V_{\dagger}^{+}(t) \equiv 0$. This means that we get the following Thiele's differential equation for the state $\star$,

$$\frac{V_{\star}^{+}(t) - V_{\star}^{+}(t-h)}{h} \approx r(t)V_{\star}^{+}(t) - \mu_{\star,\dagger}^{24}(t)(B - V_{\star}^{+}(t)).$$

In order to solve this we can use Euler's method, see Mørken 2015, p. 332. This is not the most accurate method, but it is computationally and conceptually simple. Let $h = 1/12$, $T = 114 - 24 = 90$. Then,

$$\begin{aligned} V_{\star}^{+}(t) &\approx V_{\star}^{+}(t+h) - h\frac{d}{dt}V_{\star}^{+}(t+h) \\ &\approx V_{\star}^{+}(t+h) - h(r(t+h)V_{\star}^{+}(t+h) - \mu_{\star,\dagger}^{24}(t+h)(B - V_{\star}^{+}(t+h))) \\ V_{\star}^{+}(T) &= 0 \end{aligned}$$

The terminal condition has to be zero, as there is no payment if the insured lives to 114. 114 is a common actuarial choice. In addition, we want to find the premium. To do this we first consider a policy where the insured pays 1 NOK per year until 70, hence giving us the present value of such a cash flow. Then we can find the yearly premium by dividing present value of policy by present value of the 1 NOK payments.

The calculations have been implemented in Appendix B.2. Note that the mortality $\mu_{\star\dagger}$ is implemented in Appendix B.1.

The one time premium can be calculated as 4 211.38 NOK, and the yearly premium is 173.43 NOK. Based on what the author pays for a similar policy, this seems entirely reasonable. The graph of the value is also reasonable. The value increases as the probability of death increases. On the other hand, when the person reaches the age of ca. 55, the short time left in the policy start to dominate as the probability of getting a payout over the last years of the contract decreases fast.

Example 5.4.2 (Orphan insurance)**.** Orphan insurance can be an important peace of mind for many families. How can one ensure that a child is provided

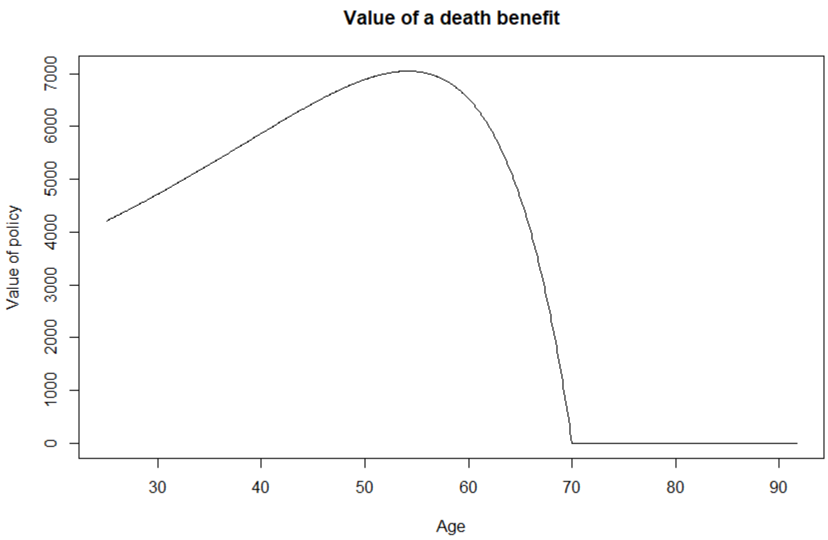

Figure 5.1: The evolution of value of a death benefit policy.

for if one or both parents die? The answer to this is orphan insurance. This policy pays out a yearly sum to the insured if one or both parents die up to and including the year the child turns 25 years.

Suppose a baby boy has just been born, that the mother is 25 years and the father 27 years. This insurance would have eight states. We use the following convention $\{child, mother, father\}$. The policy is described as such,

$$a_{\star,\dagger,\star}(t) = a_{\star,\star,\dagger}(t) = \begin{cases} 50\,000t, & \text{if t} \leq 25 \\ 1\,250\,000, & \text{if t} > 25, \end{cases}$$

$$a_{\star,\dagger,\dagger}(t) = \begin{cases} 100\,000t, & \text{if t} \leq 25 \\ 2\,500\,000, & \text{if t} > 25. \end{cases}$$

We assume that the lifetime of the members of the family is independent. This is unrealistic, but the alternative forces us to model the co-mortality which is outside the scope of this thesis. Moreover, we assume that no member of the family can perish at the same time. This simplifies the transitions between states. There are no payments if the parents outlive the child. Therefore all states where the child is dead has value zero. The same is true for the value of the policy if the child is older than 25. The states and possible transitions are shown in Figure 5.4.2.

We are then left with the following Thiele's equation:

$$\begin{aligned} \frac{d}{dt}V_{\star,\star,\star}(t) = r(t)V_{\star,\star,\star}(t) \\ &- \mu_{\star\star\star,\star\star\dagger}(t)(V_{\star,\star,\dagger}(t) - V_{\star\star\star}(t)) \\ &- \mu_{\star\star\star,\star\dagger\star}(t)(V_{\star\dagger\star}(t) - V_{\star\star\star}(t)) \\ &+ \mu_{\star\star\star,\dagger\star\star}(t)V_{\star,\star,\star}(t) \end{aligned}$$

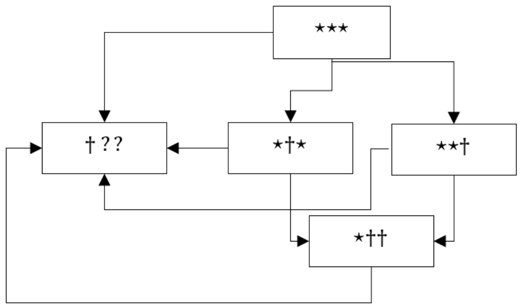

Figure 5.2: The states and possible transitions for an orphan insurance under the stated assumptions. In the interest of time we have collected all states where the child is dead in one box.

$$\begin{aligned}
\frac{d}{dt}V_{\star\star\dagger}(t) &= r(t)V_{\star\star\dagger}(t) - 50\,000 \\
&\quad - \mu_{\star\star\dagger,\star\dagger\dagger}(t)(V_{\star\dagger\dagger}(t) - V_{\star\star\dagger}(t)) \\
&\quad + \mu_{\star\star\dagger,\dagger\star\dagger}(t)V_{\star\star\dagger}(t) \\
\frac{d}{dt}V_{\star\dagger\star}(t) &= r(t)V_{\star\dagger\star}(t) - 50\,000 \\
&\quad - \mu_{\star\dagger\star,\star\dagger\dagger}(t)(V_{\star\dagger\dagger}(t) - V_{\star\dagger\star}(t)) \\
&\quad + \mu_{\star\dagger\star,\dagger\dagger\star}(t)V_{\star\dagger\star}(t) \\
\frac{d}{dt}V_{\star\dagger\dagger}(t) &= r(t)V_{\star\dagger\dagger}(t) - 100\,000 \\
&\quad + \mu_{\star\dagger\dagger,\dagger\dagger\dagger}(t)V_{\star\dagger\dagger}(t)
\end{aligned}$$

The final value of the insurance for all states is zero. Using a similar set-up as above, we find that the one-time premium of the insurance is 6 498.48 NOK. The value evolves for the different states are displayed in Figure 5.4.2.

This concludes the example.

In this chapter we have introduced the basics of insurance mathematics and defined the basic insurance model with stochastic cash flows and policy functions. Then we found the value of a policy and derived a differential equation for the value. The payouts in this chapter have all been deterministic. A natural question to ask is: what happens when the payment is stochastic and tied to a financial asset?

These kinds of policies are called unit-linked life insurance policies, and these are is the topic of the next chapter.

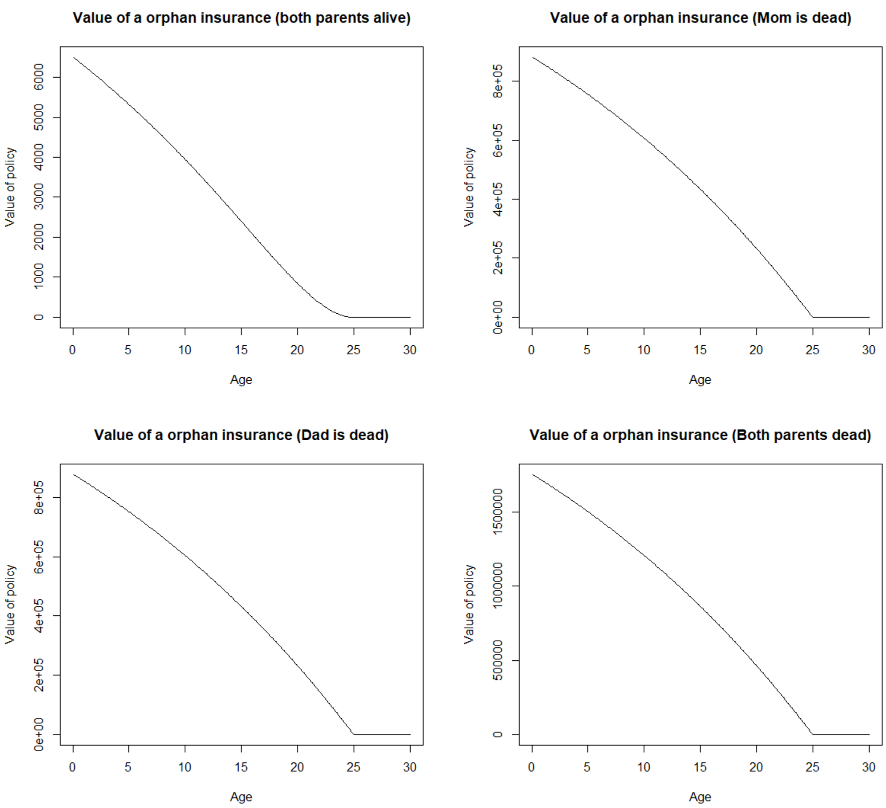

Figure 5.3: The value of the orphan insurance for different states of the family.

CHAPTER 6

Unit-linked life insurance policies

In this chapter we will discuss the unit-linked life insurance policy (ULLIP). This is a form of insurance policy that introduces risk to the payoff of the contract. Hence one gets two sources of risk, the state of the insured and the payment to the insured. This means that this kind of contract combines classical insurance mathematics with financial mathematics. In particular we will study contracts where the payment to the insured is based on the value of some risky asset. The most natural application is that this risky asset is some form of stock fund.

In the first section of this chapter we will find the explicit value of a general unit-linked policy. The value of a unit-linked policy depends on the underlying model of the fund. The second section is therefore going to be devoted to finding a Thiele's equation for a unit-linked policy governed by a jump diffusion model like the one in (4.1). This is a new result from this thesis. Lastly, we will show that this is a more general case of the Thiele's PDE for a Black and Scholes market from Aase and Persson 1993.

6.1 The value of a unit-linked policy

First, assume that we are in a complete filtered probability space $(\Omega, \mathcal{F}_T, \{\mathcal{F}_t\}, P)$, where the filtration $\mathcal{F}$ is the filtration generated by the security S and the state of the insured X. I.e. $\mathcal{F}_t = \sigma(S_1(t)) \vee \sigma(X_t)$. A common unit-linked contract is a version of the call derivative, that is $\max(G, S_T)_+$. This means that the payoff is the greater of some constant G, called a guarantee, and the value of the fund at time T.

What should the value of a unit-linked contract be? Assume that the policy function $a_i(t)$ is almost everywhere differentiable with at most one discontinuity at $t = T$, where $\Delta a_i = a_i(T) - a_i(T-)$. Suppose there are functions f_i, g_i and $h_{i,j} \colon [0,T] \times \mathbb{R} \mapsto \mathbb{R}$ such that,

$$\begin{aligned} f_i(T, S_T) &= \Delta a_i(T) = a_i(T) - a_i(T-), \\ g_i(t, S_t) &= a_i'(t), \\ h_{i,j}(t, S_t) &= a_{i,j}(t). \end{aligned}$$

Then, using this notation we find that the value of an insurance policy at time t, with a fund of value S and the insured in state $i \in \mathcal{S}$ can be written,

$$\begin{aligned}
V_i^+(t, S_t) &= V_i^+(t, S_t, A) \\
&= \sum_{j \in \mathcal{S}} V_i(t, S_t, A_j) + \sum_{j \in S} V_i(t, S_t, A_{j,k}) \\
&= \frac{1}{v(t)} \sum_{j \in \mathcal{S}} v(T) p_{i,j}^x \, \mathrm{E}_Q[f_j(T, S_T)|\mathcal{F}_t] \\
&+ \frac{1}{v(t)} \sum_{j \in \mathcal{S}} \int_t^T v(s) p_{i,j}^x(t,s) \, \mathrm{E}_Q[g_j(s, S_s)|\mathcal{F}_t] ds \\
&+ \frac{1}{v(t)} \sum_{j \in \mathcal{S}, i \neq j} \int_t^T v(s) p_{i,j}^x(t,s) \mu_{j,k}^x(s) \, \mathrm{E}_Q[h_{j,k}(s, S_s)|\mathcal{F}_t] ds.
\end{aligned}$$

The second equality follows from the fact that $A = \sum_{i \in \mathcal{S}} A_i + \sum_{j,k \in \mathcal{S}} A_{j,k}$. The third equality follows from Theorem 5.2.3 with appropriate values for the policy functions.

Note that in this expression we can find the value of European derivatives,

$$\begin{aligned}
&\mathrm{E}_Q\left[\frac{v(T)}{v(t)} f_j(T, S_T)|\mathcal{F}_t\right], \\
&\mathrm{E}_Q\left[\frac{v(s)}{v(t)} g_j(s, S_s)|\mathcal{F}_t\right], \\
&\mathrm{E}_Q\left[\frac{v(s)}{v(t)} h_{j,k}(s, S_s)|\mathcal{F}_t\right].
\end{aligned}$$

To simplify later expressions we can introduce the following notation,

$$U_s^\phi = \mathrm{E}_Q\left[\frac{v(s)}{v(t)} \phi(s, S_s)|\sigma(S_t)\right].$$

where ϕ is the payoff function for some contingent claim.

We now have a value for the explicit value of a unit-linked policy. However, as with the explicit value of a deterministic policy, we would prefer to have a way of calculating this without having to calculate the probability $p_{i,j}^x(s,t)$. This is the topic for the next section.

6.2 Thiele's PIDE for unit-linked policies with jump diffusions

In this section we are going to derive Thiele's equation for a jump diffusion model. This is the main new result in this thesis. Since we use S to denote the fund, we can use $\mathcal{S}$ to denote the state space of the insured. The jump diffusion model under P is as follows,

$$\begin{aligned}
\frac{dS(t)}{S(t)} &= \mu(t, S(t))dt + \sigma(t, S(t))dW(t) + \int_{\mathbb{R}\setminus\{0\}} h(t, S(t), z)\tilde{N}(dt, dz), \\
S(0) &= S_0.
\end{aligned}$$

where S_0 is a constant. We also assume that the coefficients are Lipschitz continous and have, at most, linear growth.

We will derive a partial integro-differential equation (PIDE) that describes the changes in value for the contract. If we had only been looking for the change in value with a change in time, we would have gotten an ordinary differential equation. However, not only does the change in value depend on the starting value of the fund, $S(0)$, but we are also getting the integral over a Lévy measure, ν. This means that we get a partial integro-differential equation.

First, however, we want to simplify the notation.

$$\begin{aligned}
V_i(t,S_t) &= \sum_{j\in} p^x_{i,j}(t,T)U^{f_j}_T(t,S_t) + \sum_{j\in\mathcal{S}}\int_t^T p^x_{i,j}(t,s)U^{g_j}_s(t,S_t)ds \\
&+ \sum_{j,k\in\mathcal{S},k\neq j}\int_t^T p^x_{i,j}(t,s)\mu^x_{j,k}(s)U^{h_{j,k}}_s(t,S_t)ds \\
&= \sum_{j\in\mathcal{S}} p^x_{i,j}(t,T)U^{f_j}_T(t,S_t) + \int_t^T\Bigg[\sum_{j\in\mathcal{S}} p^x_{i,j}(t,s)U^{g_j}_s(t,S_t) \\
&+ \sum_{j,k\in\mathcal{S},k\neq j} p^x_{i,j}(t,s)\mu^x_{j,k}(s)U^{h_{j,k}}_s(t,S_t)\Bigg]ds \\
&= \sum_{j\in\mathcal{S}} p^x_{i,j}(t,T)U^{f_j}_T(t,S_t) + \int_t^T\sum_{j\in\mathcal{S}} p^x_{i,j}(t,s)\Bigg[U^{g_j}_s(t,S_t) \\
&+ \sum_{k\in\mathcal{S},k\neq j}\mu^x_{j,k}(s)U^{h_{j,k}}_s(t,S_t)\Bigg]ds \\
&= \sum_{j\in\mathcal{S}} p^x_{i,j}(t,T)U^{f_j}_T(t,S_t) + \int_t^T\sum_{j\in\mathcal{S}} p^x_{i,j}(t,s)\,\mathrm{E}_Q\left[\frac{v(s)}{v(t)}\Bigg(g_j(t,S_t)\right. \\
&+ \left.\sum_{k\in\mathcal{S},k\neq j}\mu^x_{j,k}(s)h_{j,k}(t,S_t)\Bigg)|\sigma(S_t)\right]ds \\
&= \sum_{j\in\mathcal{S}} p^x_{i,j}(t,T)U^{f_j}_T(t,S_t) + \int_t^T\sum_{j\in\mathcal{S}} p^x_{i,j}(t,s)\,\mathrm{E}_Q\left[\frac{v(s)}{v(t)}\theta^s_j(y)|\sigma(S_t)\right]ds \\
&= \sum_{j\in\mathcal{S}} p^x_{i,j}(t,T)U^{f_j}_T(t,S_t) + \int_t^T\sum_{j\in\mathcal{S}} p^x_{i,j}(t,s)U^{\theta^s_j}_s(t,S_t)dt \\
&= G^T_i(t,S_t) + \int_t^T F^s_i(t,S_t)ds
\end{aligned}$$

Hence,

$$G^T_i(t,S_t) = \sum_{j\in\mathcal{S}} p^x_{i,j}(t,T)U^{f_j}_T(t,S_t) \tag{6.1}$$

$$F^s_i(t,S_t)ds = \sum_{j\in\mathcal{S}} p^x_{i,j}(t,s)U^{\theta^s_j}_s(t,S_t) \tag{6.2}$$

Note that $U^\phi_s(t,S_t)$ is the value of a European contingent claim with payoff $\phi(s,S_s)$ at time $t\leq s$. Thus, by the Feynman-Kac formula (3.4), the function

$(t,x) \mapsto U_s^{\phi}(t,S_t)$ satisfies the PIDE,

$$\frac{\partial}{\partial t}U_s^{\theta}(t,x) + r_t x\frac{\partial}{\partial x}U_s^{\theta}(t,x) + \frac{1}{2}\sigma^2(t,x)x^2\frac{\partial^2}{\partial^2 x}U_s^{\theta}(t,x)$$
$$+ \int_{\mathbb{R}\setminus\{0\}}\left[U_s^{\theta}(t,x+h(t,x),t) - U_s^{\theta}(t,x) - h(t,x)x\frac{\partial}{\partial x}U_s^{\theta}(t,x)\right]\nu(dx) = r_t U_s^{\theta}(t,x).$$

We are now ready for the main new result of this thesis.

Theorem 6.2.1 (Thiele's PIDE for jump diffusion)**.** *Let A be the cashflow associated with the policy functions, f_i, g_i and $h_{i,j}, i,j \in \mathcal{S}$. We denote the value of an insurance contract at time $t \in [0,T]$ with the insured in state $i \in \mathcal{S}$ and given the market information $\mathcal{F}$ as $V^+_{\mathcal{F},i}(t,A)$. Then,*

$$V^+_{\mathcal{F},i}(t,A) = V_i(t,S_t),$$

where the function $(t,x) \mapsto V_i(t,x)$ is the solution to the PIDE,

$$\frac{\partial}{\partial t}V_i = r_t V_i - g_i(t,x) - \sum_{j\neq i}\mu_{i,j}(h_{i,j}(t,x) + V_j - V_i) - LV_i,$$

where the boundary conditions are $V_i(T,x) = f_i(T,x)$, and

$$LV_i(t,x) = r_t x\frac{\partial}{\partial x}V_i(t,x) + \frac{1}{2}x^2\sigma(t,x)^2\frac{\partial^2}{\partial^2 x}V_i(t,x)$$
$$+ \int_{\mathbb{R}\setminus\{0\}}\left[V_i(t,x+h(t,x),t) - V_i(t,x) - h(t,x)x\frac{\partial}{\partial x}V_i(t,x)\right]\nu(dx).$$

Proof. This is a second order parabolic PIDE and if the coefficients are Lipschitz, there exists a unique solution.

Moreover, suppose that ϕ_1 and ϕ_2 are final conditions to the PIDE U, with U^{ϕ_1} and U^{ϕ_2} denoting their solutions. Then, by the uniqueness of PIDEs and the linearity of the expectation U in ϕ, thus $U^{\phi_1+\phi_2}$ is the solution to the final condition $\phi_1 + \phi_2$.

As per the previous notation and using (6.1):

$$V^+_{\mathcal{F},i}(t,A) = V_i(t,x)\Big|_{x=S_t} = G_i^T(t,x)\Big|_{x=S_t} + \int_t^T F_i^s(t,x)ds\Big|_{x=S_t}$$
$$= G_i^T(t,S_t) + \int_t^T F_i^s(t,S_t)ds.$$

Fix T and let $s \geq t$. Then $U_T^{f_j^T}$ and $U_s^{\theta_j^s}$ are well defined.

By Kolmogorov's backward equation (Theorem 2.6.5) we have that,

$$\frac{\partial}{\partial t}p_{i,j}(t,s) = -\mu_{i,i}(t)p_{i,j}(t,s) - \sum_{k\in S,k\neq i}\mu_{i,k}(s)p_{k,j}(t,s)$$
$$= \sum_{k\in\mathcal{S},k\neq i}\mu_{i,k}(t)p_{i,j}(t,s) - \sum_{k\in S,k\neq i}\mu_{i,k}(s)p_{k,j}(t,s)$$
$$= \sum_{k\in\mathcal{S},k\neq i}\mu_{i,k}(t)(p_{i,j}(t,s) - p_{k,j}(t,s)).$$

We can use this to find the time derivative of F_i^s.

$$\begin{aligned}
\frac{\partial}{\partial t}F_i^s &= \frac{\partial}{\partial t}\sum_{j\in\mathcal{S}} p_{i,j}^x(t,s)U_s^{\theta_j^s}(t,S_t) \\
&= \sum_{j\in\mathcal{S}} \frac{\partial}{\partial t}\left(p_{i,j}^x(t,s)\right)U_s^{\theta_j^s}(t,S_t) + \sum_{j\in S} p_{i,j}^x(t,s)\frac{\partial}{\partial t}\left(U_s^{\theta_j^s}(t,S_t)\right) \\
&= \sum_{j\in\mathcal{S}}\sum_{k\in S,k\neq i} \mu_{i,k}(t)(p_{i,j}(t,s)-p_{k,j}(t,s))U_s^{\theta_j^s}(t,S_t) \\
&+ \sum_{j\in\mathcal{S}} p_{i,j}^x(t,s)\frac{\partial}{\partial t}\left(U_s^{\theta_j^s}(t,S_t)\right) \\
&= \sum_{j\in\mathcal{S}}\sum_{k\in S,k\neq i} \mu_{i,k}(t)(p_{i,j}(t,s)U_s^{\theta_j^s}(t,S_t)-p_{k,j}(t,s)U_s^{\theta_j^s}(t,S_t)) \\
&+ \sum_{j\in\mathcal{S}} p_{i,j}^x(t,s)\frac{\partial}{\partial t}\left(U_s^{\theta_j^s}(t,S_t)\right) \\
&= \sum_{j,k\in S,k\neq i} \mu_{i,k}(t)(F_i^s - F_k^s) \\
&+ \sum_{j\in\mathcal{S}} p_{i,j}^x(t,s)\frac{\partial}{\partial t}\left(U_s^{\theta_j^s}(t,S_t)\right).
\end{aligned}$$

The third equality above follows from the previous calculation with Kolmogorov's backward equation.

Next we want to prove that L is linear. Let $\alpha, \beta \in \mathbb{R}$ and U, V be functions.

$$\begin{aligned}
L(\alpha V+\beta U) &= r_t x\frac{\partial}{\partial x}(\alpha V+\beta U) + \frac{1}{2}x^2\sigma(t,x)^2\frac{\partial^2}{\partial^2 x}(\alpha V+\beta U) \\
&+ \int_{\mathbb{R}\setminus\{0\}}\Big[(\alpha V+\beta U)(t,x+h(t,x),t) - (\alpha V+\beta U)(t,x) \\
&- h(t,x)x\frac{\partial}{\partial x}(\alpha V+\beta U)\Big]\nu(dx) \\
&= \alpha\bigg(r_t x\frac{\partial}{\partial x}V_i + \frac{1}{2}x^2\sigma(t,x)^2\frac{\partial^2}{\partial^2 x}V_i \\
&+ \int_{\mathbb{R}\setminus\{0\}}\Big[V(t,x+h(t,x),t) - V(t,x) - h(t,x)x\frac{\partial}{\partial x}V\Big]\nu(dx)\bigg) \\
&+ \beta\bigg(r_t x\frac{\partial}{\partial x}U + \frac{1}{2}x^2\sigma(t,x)^2\frac{\partial^2}{\partial^2 x}U \\
&+ \int_{\mathbb{R}\setminus\{0\}}\Big[U(t,x+h(t,x),t) - U(t,x) - h(t,x)x\frac{\partial}{\partial x}U\Big]\nu(dx)\bigg) \\
&= \alpha LV + \beta LU.
\end{aligned}$$

Hence, L is a linear operator. Thus,

$$LF_i^s = \sum_j p_{i,j}(t,s)LU_s^{\theta_j^s}.$$

Next we want to calculate the dynamics of $F_i^s(t, S_t)$ under some equivalent local martingale measure Q. This might not be a unique measure, since the jump diffusion model need not be complete.

One way to find the dynamics is by direct definition,

$$
\begin{aligned}
dF_i^s(t,S_t) &= \frac{\partial}{\partial t}F_i^s(t,S_t)dt + r_t S_t \frac{\partial}{\partial x}F_i^s(t,S_t)dt \\
&+ \frac{1}{2}\sigma(t,S_t)^2 S_t^2 \frac{\partial^2}{\partial^2 x}F_i^s(t,S_t)dt \\
&+ \int_{\mathbb{R}\setminus\{0\}} \left[F_i^s(t,S_t+h(t,z)) - F_i^s(t,S_t) - S_t h(t,z)\frac{\partial}{\partial t}U_s^{\theta}(t,S_t)\right]\nu(dz) \\
&+ \sigma(t,S_t)\frac{\partial}{\partial x}F_i^s(t,S_t)dW_t^Q \\
&+ \int_{\mathbb{R}\setminus\{0\}} \frac{\partial}{\partial x}F_i^s(t,S_t)h(t,S_t,z)\tilde{N}^Q(dt,dz) \\
&= \left(\frac{\partial}{\partial t}F_i^s(t,S_t) + LF_i^s(t,S_t)\right)dt \\
&+ \sigma(t,S_t)\frac{\partial}{\partial x}F_i^s(t,S_t)dW_t^Q \\
&+ \int_{\mathbb{R}\setminus\{0\}} \frac{\partial}{\partial x}F_i^s(t,S_t)h(t,S_t,z)\tilde{N}^Q(dt,dz).
\end{aligned}
$$

On the other hand,

$$
\begin{aligned}
dF_i^s(t,S_t) &= \Bigg(\sum_{j,k\in\mathcal{S},k\neq i} \mu_{i,k}(t)(F_i^s(t,S_t) - F_k^s(t,S_t)) \\
&+ \sum_{j\in\mathcal{S}} p_{i,j}(t,s)\frac{\partial}{\partial t}U_s^{\theta_j^s}(t,S_t) + \sum_{j\in S} p_{i,j} LU_s^{\theta_j^s}\Bigg)dt \\
&+ \sigma(t,S_t)\frac{\partial}{\partial x}F_i^s(t,S_t)dW_t^Q \\
&+ \int_{\mathbb{R}\setminus\{0\}} \frac{\partial}{\partial x}F_i^s(t,S_t)h(t,S_t,z)\tilde{N}^Q(dt,dz) \\
&= \Bigg(\sum_{j,k\in\mathcal{S},k\neq i} \mu_{i,k}(t)(F_i^s(t,S_t) - F_k^s(t,S_t)) \\
&+ \sum_{j\in\mathcal{S}} p_{i,j}(t,s)(\frac{\partial}{\partial t}U_s^{\theta_j^s}(t,S_t) + LU_s^{\theta_j^s})\Bigg)dt \\
&+ \sigma(t,S_t)\frac{\partial}{\partial x}F_i^s(t,S_t)dW_t^Q \\
&+ \int_{\mathbb{R}\setminus\{0\}} \frac{\partial}{\partial x}F_i^s(t,S_t)h(t,S_t,z)\tilde{N}^Q(dt,dz).
\end{aligned}
$$

Since $U_s^{\theta_j^s}$ is the solution to $\frac{\partial}{\partial t}U_s^{\theta_j^s} + LU_s^{\theta_j^s} = r_t U_s^{\theta_j^s}$.

$$
dF_i^s(t,S_t) = \left(\sum_{k\neq i} \mu_{i,j}(t)(F_i^s(t,S_t) - F_k^s(t,S_t)) + r_t F_i^s(t,S_t)\right)dt
$$

$$
\begin{aligned}
&+ \sigma(t,S_t)\frac{\partial}{\partial x}F_i^s(t,S_t)dW_t^Q \\
&+ \int_{\mathbb{R}\setminus\{0\}} \frac{\partial}{\partial x}F_i^s(t,S_t)h(t,S_t,z)\tilde{N}^Q(dt,dz).
\end{aligned}
$$

Here we have used the fact that $LF_i^s = \sum_{j\in S} p_{i,j}(t,s)LU_s^{\theta_j^s}$.

We can see that these are two expressions that are equivalent. This means that,

$$
\sum_{k\neq i}\mu_{i,j}(t)(F_i^s(t,S_t) - F_k^s(t,S_t)) + r_tF_i^s(t,S_t) = \frac{\partial}{\partial t}U_s^{\theta_j^s}(t,S_t) + LU_s^{\theta_j^s}.
$$

This is a PIDE in S_t and t for a fixed s.

Since $V_i(t,S_t) = V_i(t,x)\Big|_{(t,x)=(t,S_t)}$ and $V_i(t,x) = G_i^T(t,x) + \int_t^T F_i^s(t,x)ds$ we can use Lebesgue's dominated convergence theorem and the fundamental theorem of calculus to get,

$$
\frac{\partial}{\partial t}V_i(t,x) = \frac{\partial}{\partial t}G_i^T(t,x) + \int_t^T \frac{\partial}{\partial t}F_i^s(t,x)ds - \lim_{s\to t}F_i^s(t,x).
$$

Finding the limit,

$$
\begin{aligned}
\lim_{s\to t}F_i^s(t,x) &= \lim_{s\to t}\sum_{j\in\mathcal{S}}p_{i,j}(t,s)U_s^{\theta_j^s} = \lim_{s\to t}\sum_{j\in\mathcal{S}}p_{i,j}(t,t)U_t^{\theta_j^t} \\
&= \lim_{s\to t}\sum_{j\in S}U_t^{\theta_j^t} = U_t^{\theta_i^t} = \theta_i^t \\
&= g_i(t,x) + \sum_{k\neq i}\mu_{i,k}(t)h_{i,k}(t,x).
\end{aligned}
$$

Inserting this into the previous expression yields,

$$
\frac{\partial}{\partial t}V_i(t,x) = \frac{\partial}{\partial t}G_i^T(t,x) + \int_t^T \frac{\partial}{\partial t}F_i^s(t,x)ds - g_i(t,x) + \sum_{k\neq i}\mu_{i,k}(t)h_{i,k}(t,x).
$$

This is equivalent to,

$$
\int_t^T \frac{\partial}{\partial t}F_i^s(t,x)ds = \frac{\partial}{\partial t}V_i(t,x) + g_i(t,x) - \sum_{k\neq i}\mu_{i,k}(t)h_{i,k}(t,x).
$$

We now return to the equality

$$
\frac{\partial}{\partial t}F_i^s + LF_i^s = \sum_{k\neq i}\mu_{i,k}(t)(F_i^s - F_k^s) + r_tF_i^s. \tag{6.3}
$$

Integrating this with respect to s over the interval $[t,T]$.

$$
\begin{aligned}
&\int_t^T \frac{\partial}{\partial t}F_i^s ds + \int_t^T LF_i^s ds = \int_t^T \left(\sum_{k\neq i}\mu_{i,k}(t)(F_i^s - F_k^s) + r_tF_i^s\right)ds \\
&\int_t^T \frac{\partial}{\partial t}F_i^s ds + L\int_t^T F_i^s ds = \sum_{k\neq i}\mu_{i,k}(t)\left(\int_t^T F_i^s ds - \int_t^T F_k^s ds\right) + r_t\int_t^T F_i^s ds,
\end{aligned}
$$

which is equivalent to

$$\begin{aligned}&\frac{\partial}{\partial t}(V_i - G_i^T) + g_i - \sum_{k\neq i}\mu_{i,k}(t)h_{i,k} + L(V_i - G_i^T)\\&= \sum_{k\neq i}\mu_{i,k}(t)\Big(V_i - V_k + G_k^T - G_i^T\Big) + r_t(V_i - G_i^T).\end{aligned}$$

Now consider the following

$$\begin{aligned}\frac{\partial}{\partial t}G_i^T &= \sum_{j\in\mathcal{S}}\frac{\partial}{\partial t}p_{i,j}(t,T)U_T^{f_j} + \sum_{j\in\mathcal{S}}p_{i,j}(t,T)\frac{\partial}{\partial t}U_T^{f_j}\\&= \sum_{j\in\mathcal{S},k\neq i}p_{i,j}(t,T)(G_i^T - G_k^T) + \sum_{j\in S}p_{i,j}(t,T)\frac{\partial}{\partial t}U_T^{f_j}.\end{aligned}$$

In addition,

$$\begin{aligned}\frac{\partial}{\partial t}G_i^T + LG_i^T &= \sum_{k\neq i}\mu_{i,k}(t)(G_i^T - G_k^T) + \sum_{j\in\mathcal{S}}p_{i,j}(t,T)\bigg(\frac{\partial}{\partial t}U_T^{f_j} + LU_T^{f_j}\bigg)\\&= \sum_{k\neq i}p_{i,j}(t,T)\bigg(\frac{\partial}{\partial t}U_T^{f_j} + r_tG_i^T\bigg)\end{aligned}$$

This last equality follows from the fact that $\frac{\partial}{\partial t}U_T^{f_j} + LU_T^{f_j} = r_tU_T^{f_j}$ which is a PIDE we got from the Feynman-Kac formula. Finally,

$$\frac{\partial}{\partial t}V_i(t,x) = r_tV_i - g_i(t,x) - \sum_{k\neq i}\mu_{i,k}(t)h_{i,k}(t,x) + \sum_{k\neq i}\mu_{i,k}(t)(V_i - V_k) - LV_i.$$

■

We have now proven the main theoretical result of this thesis, giving us a simpler way to find the value of an insurance contract where the payment depends on the value of a financial asset modeled with jumps.

One interesting special case of this is when $\mu, \sigma \in \mathbb{R}$ and $h = 0$. This is the Black and Scholes case. In this case we can see that the differential operator becomes,

$$LV_i = r_tx\frac{\partial}{\partial x}V_i + \frac{1}{2}x^2\sigma(t,S(t))^2\frac{\partial}{\partial x}V_i.$$

This is a known result and can be found in Aase and Persson 1993, p. 97.

In the next chapter we will explore some numerical methods for estimating the solution to the PIDE.

PART II

Applications

CHAPTER 7

Numerical methods

In the theoretical part of this thesis, we derived a partial integro-differential equation (PIDE). However, there might not be a simple and efficient way to find an analytic solution. Therefore we need numeric methods.

In this chapter we will explore some of the methods that can be used to solve such PIDEs. We introduce two methods. Both are based on backwards integration on a bounded domain, but differ in computational cost and in numerical stability. The bounded domain is a limitation we place on the solution for computational reasons. We will refer to the first method as the *explicit method* and the second method as the *implicit method*.

In order to calculate reserves and prices we need to specify the context in which we are working. To that end we need to introduce a jump diffusion that models the securities and a model for the mortality of the insured. We will model the security with the Merton jump-diffusion model. The Merton jump-diffusion model is an extension of the Black and Scholes model, but with jumps. Merton eventually shared the 1997 Nobel prize with Myron Scholes (and would probably have shared with Fischer Black, had he not abruptly died in 1995). The mortality model we are going to use is "K2013". This is a standard method used by the Norwegian financial supervisory authority. Norwegian law mandated that life insurance companies can not use models for mortality that gives less conservative reserves than K2013. This makes it a natural model to use.

7.1 The Merton jump-diffusion model

Robert Merton introduced the model in his 1976 paper "Option pricing when underlying stock returns are discontinuous", see Merton 1976. The model is an extension of the Black and Scholes model where jumps are introduced. The jumping process is at time t a homogeneous compensated compound Poisson process, where the jump intensity is a constant λ and the jumps are i.i.d. random variables Z_i. Using Merton's own notation one can state the dynamics under P as follows,

$$\frac{dS_t}{S_t} = (\mu - \lambda E(Z))dt + \sigma dW_t + dN_t \tag{7.1}$$

Here $dN_t = Y_t - 1$ if a jump occurs and 0 otherwise.

Indeed this is a jump diffusion of the kind we have been studying when a proper adjustment to the drift is made.

We can find an expression for the solution to this SDE.

$$S_t = S_0 \exp\left((\mu - \sigma^2/2 - \lambda\,\mathrm{E}(Z))t + \sigma B_t + \sum_{i=1}^{N_t} Z_i\right)$$

One can prove this by using Itô's formula, however, since this is not a very central result, we can refer the reader to Merton 1976.

The Merton model does not produce a complete market. Since a single security is driven by two noisy processes, our rule of thumb states that the market is not complete and that there is no unique local martingale measure. More formally, no predictable process η exists such that,

$$\psi(t) = e^{\int_0^t r_s ds}\eta(t)\sigma(t)$$
$$\gamma(z,t) = e^{\int_0^t r_s ds}\eta(t)h(z,t).$$

Thus market incompleteness follows from theorem 4.5.4.

This means that when we go from the dynamics under P to the dynamics under Q, we have to make a choice of which Q we use.

We choose Q like in Merton 1976 and get the following dynamics,

$$\frac{dS_t}{S_{t-}} = (r - \lambda\,\mathrm{E}[Z])dt + \sigma dB_t^Q + \sum_{k=i}^{N_t} Z_i$$

As for the jumps, one could in principle use many distributions, but a careful analysis of the returns of the asset should be conducted. In Merton 1976 a second independent log-normal distribution is used. We will also use this here. This choice of distribution means that jumps are positive. Using a log-normal distribution means that we avoid the problems associated with having negative prices. This also means that we can find an explicit expression for the Lévy measure.

$$\nu(x) = \lambda f(x) = \frac{\lambda}{\delta\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-m}{\delta}\right)^2\right)$$

Here δ is the standard deviations of the jumps, m is the mean of the jumps and λ is the jump intensity. The source of this result is F. E. Benth, J. S. Benth, and Koekebakker 2008, p. 50.

One can prove this by finding the characteristic function of the compound Poisson and using the Lévy-Khintchine formula to identify the correct term.

Moreover we can find the PIDE for a European option with Merton's model as the underlying model. Simply consider Theorem 3.7.3, where all dt terms disappear. Alternatively, see Cheang and Chiarella 2012.

$$\frac{\partial}{\partial t}V(t,x) = r_t V(t,x) - r_t x\frac{\partial}{\partial x}V(t,x) - \frac{1}{2}\sigma^2 x^2 \frac{\partial^2}{\partial^2 x}V(t,x) \tag{7.2}$$
$$- \int_{\mathbb{R}\setminus\{0\}} V(t,x+y) - V(t,x) - xy\frac{\partial}{\partial x}(t,x)\nu(dy) \tag{7.3}$$

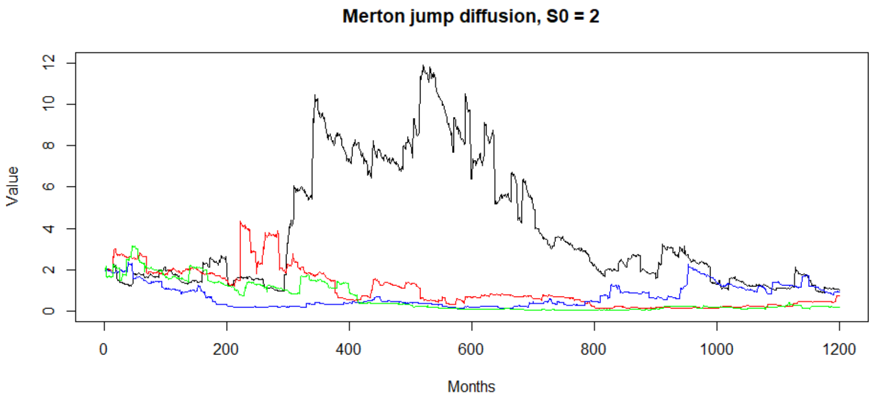

Figure 7.1: Sample paths of the Merton jump-diffusion.

This is going to be the generator L for our Thiele's PIDE.

Using parts of the code in B.3 we can generate Figure 7.1 showing some sample paths of Merton's jump-diffusion.

Now that we have explored the securities model we can move on to the mortality model.

7.2 The mortality model: K2013

Another essential ingredient in life insurance is the modeling of the state of the insured. Recall that the state of the insured is modeled by a Markov process. In this thesis we will be looking at contracts where the insured is either alive or dead. Therefore, we will use a model that only describes mortality. If one were to look at a contract with multiple states of the insured, such as disability insurance, one would have to model the movement between these states. These kinds of models are slightly more complicated, but do not change the mathematics of Thiele's PIDE in any meaningful way, except that we would need one PIDE for each state.

Since we are only looking at the case where the insured is either alive or dead, let us derive a useful result connecting the transition rate with the transition probability.

Proposition 7.2.1. *Consider a Markov process, $X = \{X_t, t \geq 0\}$ with state space $S = \{\star, \dagger\}$ and suppose that $\mu(t)$ describes the transition rate between state $\star$ and $\dagger$. Introduce the notation $p_{\star,\star}(s,t) = P(X_t = \star | X_s = \star)$. Then,*

$$p_{\star,\star}(s,t) = \exp\left(- \int_s^t \mu_{\star,\dagger}(s) ds \right).$$

Proof. The idea is to find an ODE that gives us the solution. Let $s \leq t$ and $h > 0$.

$$\begin{aligned} p_{\star,\star}(s,t) - p_{\star,\star}(s,t+h) &= p_{\star,\star}(s,t)(1 - p_{\star,\star}(t,t+h)) \\ &= p_{\star,\star}(s,t)p_{\star,\dagger}(t,t+h) \end{aligned}$$

Dividing this by h and taking the limit as $h \to 0$ thus gives,

$$-\frac{d}{dt}p_{\star,\star}(s,t) = p_{\star,\star}(s,t)\mu_{\star,\dagger}(t)$$

This is an ODE with solution,

$$p_{\star,\star}(s,t) = \exp\bigg(-\int_s^t \mu_{\star,\dagger}(s)ds\bigg).$$

with boundary conditions $p_{\star,\star}(t,t) = 1$. ■

This result is very useful because it gives us a quick way to find the probability of moving between states in a given time without having to solve systems of ordinary differential equations (ODEs). If the state space had more elements with more transition probabilities, we would have one ODE for each state, all of them interconnected. This would not yield a nice formula like this, but would probably require numerical methods.

One might wonder why we needed to introduce this result. This is because our mortality model, K2013, is a transition rate μ. Using this result we can then easily find the probabilities of survival. The model is as follows,

$$\mu(x,t) = \mu_{kol}(x,2013)\bigg(1 + \frac{w(x)}{100}\bigg)^{t-2013},$$

where

$$w(x) = \begin{cases} \min(2.671548 - 0.172480x + 0.0014285x^2, 0), \text{for men} \\ \min(1.287968 - 0.101090x + 0.000814x^2, 0), \text{for women} \end{cases}$$

and

$$\mu_{kol}(x,t) = \begin{cases} (0.241752 + 0.004536 * 10^{0.051x})/1000, \text{for men} \\ (0.085411 + 0.003114 * 10^{0.051x})/1000, \text{for women} \end{cases}$$

Here x is the current year and t is the age of the insured.

We choose this method primary for two reasons. Firstly, it is a standard model in the Norwegian industry. Hence, by using K2013 the mortality in the examples are realistic and familiar to the industry. Secondly, the model is simple and transparent. We implemented K2013 in Appendix B.1.

To conclude this section, the author has calculated his own survival probability based on K2013. Life expectancy was 82 years. This is in line with demographic expectations.

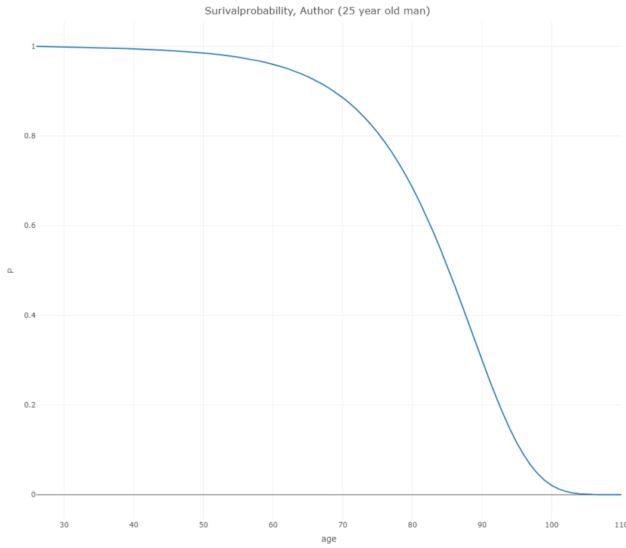

Figure 7.2: Demonstrating the survival probability, $p_{\star,\star}(25, t), t \geq 25$, using K2013.

7.3 The explicit method

The explicit method is based on Euler's scheme for solving ODEs. We divide the domain into a grid. Since we know the value of three border regions we can calculate each of the border points systematically by taking the know values and interpolating the interior point. A different introduction to this can be found in Tankov 2003, p. 424.

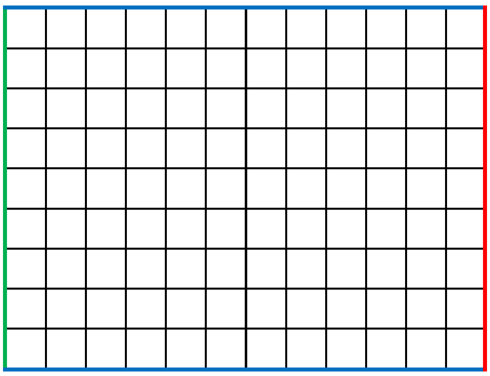

Figure 7.3: The grid representing the domain where we approximate the solution. Left is $t = 0$ and right is $t = T$. Blue edges are boundary conditions and red is terminal conditions, both of which are known.

Consider the grid in Figure 7.3. Suppose the x-axis represents the partition in time and the y-axis represent a partition in space, i.e. in value of the fund.

The points on the red line are the terminal conditions, f_i, and thus known. The blue lines represents the boundary conditions we have imposed. Hence, the points on the red and the blue lines are known. We seek to interpolate each of the points in the intersection based on the known information. This means that we start with the rightmost black line, and begin at the top. This point has known neighboring points, both in the space, and the time direction. Using Euler's method we can interpolate it. The next point on the same line also has known neighbors in the time and space directions, one of which is the previous point we interpolated. We continue this for all points on the rightmost black line. Once finished, we do the same thing to the second rightmost black line, then the third rightmost, and so on.

This is the gist of the method. However, we can formulate this more mathematically.

Consider the collection of points (x_i, t_j) for $i = 0, 1, \dots n$ and $j = 0, 1, \dots, m$. Introduce the following simplified notation:

$$\begin{aligned} r_i &= r(t_i), \\ g_i^j &= g(t_i, x_j), \\ \mu_i &= \mu^{x_{\text{start}}}(t_i), \\ f_i^j &= f_k(t_i, x_j), \\ h_i^j &= h_{k,l}(t_i, x_j), \\ \sigma_i^j &= \sigma(t_i, x_i). \end{aligned}$$

Note that we assume that $k, l \in \mathcal{S}$. The method is easily extended to the case where there are multiple states the insured can be in.

Next consider the Taylor approximations for the derivatives,

$$\begin{aligned} \frac{\partial}{\partial t} V(t,x) &\approx \frac{V(t+\Delta t, x) - V(t,x)}{\Delta t} \\ \frac{\partial}{\partial x} V(t,x) &\approx \frac{V(t, x+\Delta x) - V(t,x)}{\Delta x} \\ \frac{\partial^2}{\partial^2 x} V(t,x) &\approx \frac{V(t, x+\Delta x) - 2V(t,x) + V(t, x-\Delta x)}{(\Delta x)^2} \end{aligned}$$

Define $V_i^j = V(t_i, x_j)$ and let x_{start} be the age of the insured at the start of the contract. This means that we can rewrite Thiele's PIDE 6.2.1 as

$$\begin{aligned} \frac{V_i^j - V_{i-1}^j}{\Delta t} &= r_t V_i^j - g_{t_i} - \mu(x_{\text{start}} + t_i)(h - V) - r_t x_j \frac{V_i^{j+1} - V_i^j}{\Delta x} \\ &- \frac{1}{2}\sigma^2 x_j^2 \frac{V_i^{jj+1} - V_i^j + V_i^{j-1}}{(\Delta x)^2} - \int_{R\setminus\{0\}} V_i^{x_j+y} - V_i^j - x_j e^{y_k} \frac{V_i^{j+1} - V_i^j}{\Delta x} \nu(dy) \end{aligned}$$

Suppose that we bound the domain to some interval $[B_l, B_r]$. Cont and Voltchkova 2003 also suggest truncating the jump integral to an interval $[K_l, K_r]$ such that $[B_l, B_r] \subset [(K_l - 0.5)\Delta x, (K_r + 0.5)\Delta x]$. Hence the expression above can be written.

$$\frac{V_j^i - V_{j-1}^i}{\Delta t} = r_t V_j^i - g_{t_i} - \mu(x_{\text{start}} + t_i)(h - V) - r_t x_j \frac{V_j^{i+1} - V_j^i}{\Delta x}$$

$$- \frac{1}{2}\sigma^2 x_j^2 \frac{V_j^{i+1} - 2V_j^i + V_j^{i-1}}{(\Delta x)^2} - \sum_{k=K_l}^{K_r} V_j^{x_j+y_k}\nu_j + V_j^i \sum_{k=K_l}^{K_r} \nu_j$$
$$+ x\frac{V_j^{i+1} - V_j^i}{\Delta x} \sum_{k=K_l}^{K_r} e^{y_k}\nu_j$$

Here $\nu_j = \int_{(j-0.5)\Delta x}^{(j+0.5)\Delta x} \nu(dy)$. Moreover, we can note the following approximations. Firstly, $\sum_{k=K_l}^{K_r} \nu_j = \hat{\lambda}$. This is the sum of the increments of the Lévy measures over the domain of the jumps, i.e. an estimate of the mean of the Lévy measure. Secondly, $\sum_{k=K_l}^{K_r} e^y \nu_j = \hat{m}$, since the jumps are assumed to be log-normal this becomes an estimate of the mean of the jumps.

Rearranging this we can approximate the value at time $t - \Delta t$,

$$V_{j-1}^i = V_j^i - \Delta t \Bigg(r_t V_j^i - r_t x \frac{V_j^{i+1} - V_j^i}{\Delta x} - g_{t_i} - \mu(x_{\text{start}} + t_j)(h - V)$$
$$- \frac{1}{2}\sigma^2 x_j^2 \frac{V_j^{i+1} - 2V_j^i + V_j^{i-1}}{(\Delta x)^2}$$
$$- \sum_{k=K_l}^{K_r} V_j^{x_j+y_k}\nu_j + V_j^x \hat{\lambda} + x_j \hat{m} \frac{V_j^{i+1} - V_j^i}{\Delta x} \Bigg)$$

This can be implemented and solved with the terminal conditions being that of the payoff function at time T. The main drawback of this method is that it is unstable. Therefore, we have preferred to use the implicit method in this thesis.

7.4 The implicit method

The idea behind the implicit method is to express the discretized expression as a system of equations that can then be solved. Consider again the discretized expression for Thiele's PIDE,

$$\frac{V_{i+1}^j - V_i^j}{\Delta t} = r_t V_i^j - g_{t_i} - \mu(x_{\text{start}} + t)(h_i^j - V) - r_t x_j \frac{V_i^{j+1} - V_i^j}{\Delta x}$$
$$- \frac{1}{2}\sigma^2 x_j^2 \frac{V_i^{j+1} - 2V_t^j + V_i^{j-1}}{(\Delta x)^2} - \sum_{k=K_l}^{K_r} V_{i+1}^{x_j+y_k}\nu_j + V_i^j \sum_{k=K_l}^{K_r} \nu_j$$
$$+ x_j \frac{V_i^{j+1} - V_i^j}{\Delta x} \sum_{k=K_l}^{K_r} e^{y_k}\nu_j.$$

If we write this out and group the terms together we can show that this is equivalent to,

$$\begin{aligned}
\frac{V_{i+1}^j}{\Delta t} - g_{t_i} - \mu(x_{\text{start}}+t)h = V_i^j &\left(\frac{1}{\Delta t} + r + \mu + \frac{r_t x_j}{\Delta x} + \hat{\lambda} - \frac{x_j \hat{m}}{\Delta x} + \frac{1}{2}\left(\frac{x_j\sigma}{\Delta x}\right)^2\right) \\
&+ V_i^{j+1}\left(r x_j + \frac{1}{2}\frac{\sigma^2 x_j^2}{\Delta x^2} + \frac{x_j}{\Delta x}\hat{m}\right) \\
&- V_i^{j-1}\frac{1}{2}\frac{\sigma^2 x_j^2}{\Delta x^2} \\
&- \sum_{k=-Kl}^{Kr} V_{i+1}^{x_j+y_k}\nu_j.
\end{aligned}$$

One can simplify the notation further by introducing the following notation.

$$\begin{aligned}
A_i^j &= -\frac{1}{2}\frac{\sigma^2 x_j^2 \Delta t}{\Delta x^2}, \\
B_i^j &= 1 + \Delta t\left(r_t + \mu_t + \frac{r_t x_j}{\Delta x} + \left(\frac{x_j \sigma_t}{\Delta x}\right)^2 - \hat{\lambda} + \frac{x_j \hat{m}}{\Delta x}\right), \\
C_i^j &= \Delta t\left(\frac{r_t x_j}{\Delta x} + \frac{1}{2}\frac{\sigma_t^2 x_j^2}{\Delta x^2} + \frac{x_j}{\Delta x}\hat{m}\right), \\
D_i^j &= V_{i+1}^j + \Delta t(g_{t_i} + \mu(x_{\text{start}}+t)h(t_i,x_j) + \sum_{k=-Kl}^{Kr} V_{i+1}^{x_j+y_k}\nu_j).
\end{aligned}$$

Here V_t denotes the vector of value for all x in the domain at time t. Inserting this notation yields the following system of equations,

$$A_i^j V_i^{j-1} + B_i^j V_i^j + C_i^j V_i^{j+1} = D_i^j$$

Note that the jump matrix depends, as with D, on time $t+\Delta t$. Thus, we can set this up as matrices.

$$\begin{pmatrix}
A_t^1 & B_t^1 & C_t 1 & 0 & \cdots & 0 & 0 & 0 & 0 \\
0 & A_t^2 & B_t^2 & C_t^2 & \cdots & 0 & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & 0 & \cdots & A_t^{m-2} & B_t^{m-2} & C_t^{m-2} & 0 \\
0 & 0 & 0 & 0 & \cdots & 0 & A_t^{m-1} & B_t^{m-1} & C_t^{m-1}
\end{pmatrix}
\begin{pmatrix} V_t^0 \\ V_t^1 \\ \vdots \\ V_t^{m-1} \\ V_t^m \end{pmatrix}
+ \begin{pmatrix} A_t^1 V_i^0 \\ \vdots \\ C_t^{m-1} V_i^m \end{pmatrix}
= \begin{pmatrix} D_t^1 \\ D_t^2 \\ \vdots \\ D_t^{m-2} \\ D_t^{m-1} \end{pmatrix}.$$

Denote the first matrix containing A, B and C as M. This set of equations can be expressed,

$$M_t V_t + E_t = D_t$$

which can be solved,

$$V_t = (M_t)^{-1}(D_t - E_t)$$

An alternative scheme for Merton's PIDE can be found in Cont and Voltchkova 2003. It does not include the insurance part of the PIDE. However, one can easily extend this implementation to also cover the insurance part. This alternative scheme has been implemented in Cantarutti 2022. This code has been adapted to fit the insurance context we here are working with.

We will use this method when we look at some examples in the next chapter.

CHAPTER 8

Examples

In this chapter we will look at some examples. These examples will highlight some of the differences between the jump-diffusion model and the classic Black and Scholes model. We will study three unit-linked life insurance contracts this is a pure endowment policy, a death benefit, and a pension. Moreover, we are also going to look at a non-insurance example, fish farming.

For the following examples suppose that we have the probability space $(\Omega, \mathcal{F}_T, P)$ equipped with the filtration $\mathcal{F} = \{\mathcal{F}_t\}_{t\geq 0}$ generated by the fund S and the Markov chain X.

8.1 Example 1: Pure endowment policy

A pure endowment policy is one of the simplest life insurance policies in the market. The policy pays a lump sum when the insured reaches a certain age, but only if the insured is alive. This means that the state space is $\mathcal{S} = \{\star, \dagger\}$, where $\star$ signifies *alive* and † signifies *dead*. Suppose the contract is such that at the age of 65 the insured is to get a sum of G = 3 (million NOK) guaranteed or the value of a fund S if this is greater than G. We assume that the fund is modeled by a Merton model as described. We assume that the insured is a 55 year old man. From this we can summarize the following policy functions,

$$\begin{aligned} f_\star &= \max(S_T, G), \\ g_\star(t) &= 0, \\ h_{\star,\dagger} &= 0. \end{aligned}$$

All functions describing the state † have to be zero. From this we can derive the following PIDE,

$$\begin{aligned} \frac{d}{dt}V_\star &= rV_\star - \mu_{\star,\dagger}V_\star - LV_\star \\ &= rV_\star - \mu_{\star,\dagger}V_\star - r_t x\frac{\partial}{\partial x}V_\star - \frac{1}{2}x^2\sigma^2\frac{\partial^2}{\partial^2 x}V_\star \\ &- \int_{\mathbb{R}\setminus\{0\}}\left[V_\star(t, x+h(t,x),t) - V_\star - xy\frac{\partial}{\partial x}V_\star\right]\nu(dy). \end{aligned}$$

Note that since there is no payment whatsoever when the insured is in state † the value of the insurance has to be zero for this state.

Parameter estimation for the securities model is outside the scope of this thesis. Therefore, the values chosen are reasonable, but not based on any data set. Methods for this do exist, such as in Tang 2018 or Meyer-Brandis and Tankov 2008. The yield of the 10 year US treasury bond (as of 01.04.2022) has been chosen as the risk-free interest rate. Hence, the security model has the following parameters:

Parameter	Value
σ_d	0.19
λ	10
μ_J	0.025
σ_J	0.25
r	0.023

With these assumptions in place, the first thing to consider is the boundary conditions. If the insured is alive at this point, the payoff of the contract is $(S_T - G)_+ + G$. Here $(S_T - G)_+ = (S_T - G)\mathbf{1}_{(S_T - G) > 0}$. This means that the terminal condition, i.e. the value at the last point in time, has to have this form. Consider now the situation where the value of the fund is low. Obviously, the insured would get the guaranteed amount, but only if he survived to collect it. Hence, the lower bound should be $Gp_{\star,\star}(s,T)e^{-r(T-s)}$, where $s \in [0,T]$. On the other hand, what payment would the insured get if the value of the fund is high? The value of the fund should, in this context, be the upper bound on the value of the fund multiplied by the survival probability, $X_{max}p_{\star,\star}(s,T)$.

Running the code B.3 we get the figure on the left of Figure 8.1. It can be of interest to compare the shape of the value to a policy with the same terms but with a securities model where there are no jumps. The figure on the right of Figure 8.1.

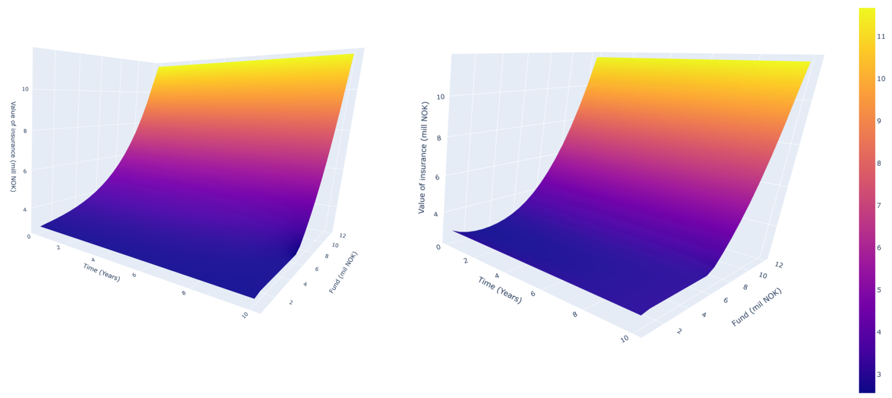

Figure 8.1: The value of an endowment policy where the security is a Merton jump-diffusion (L) and a Black and Scholes diffusion model (R).

We can see that the shape of the surface is slightly different. The jump-diffusion model gives a less convex graph. This makes sense. Since the value of the fund is subject to jumps, there is a greater possibility that the final value of the fund is different than the initial value. Moreover, there is a lower bound to the payments, but not an upper limit. The jumps therefore make a high payment more likely and thus makes the contract more valuable. This conclusion might be due to the choice of Merton's jump-diffusion.

We can find the difference between value of the insurance with and without jumps.

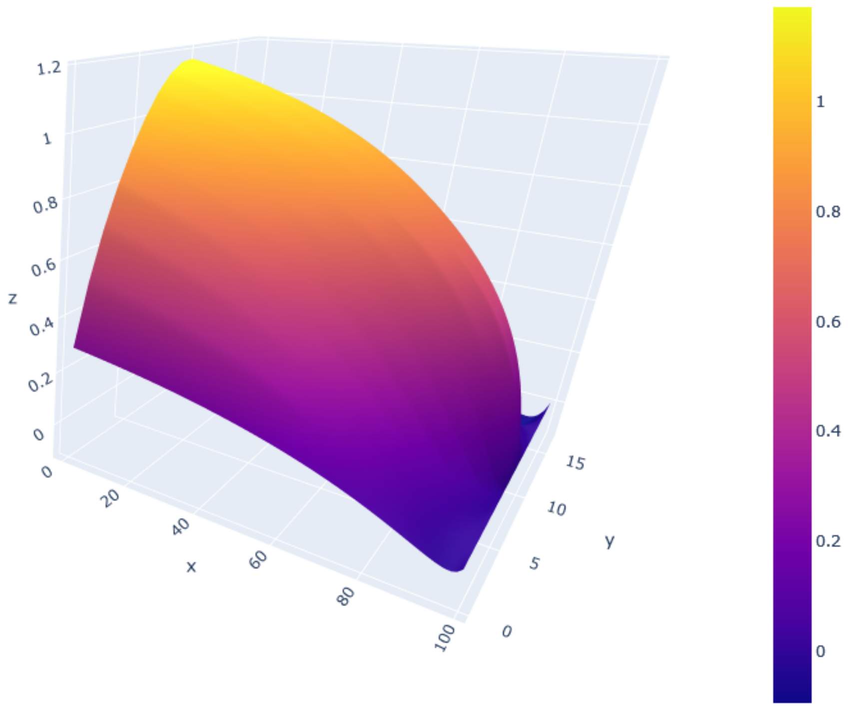

Figure 8.2: Pure Endowment: difference between Merton jump-diffusion and Black and Scholes diffusion.

S_0	Merton Jump-diffusion	Black and Scholes diffusion
1	3.1	2.8
2	3.8	2.7
5	4.7	4.5

Note that there is some numerical error in all of these calculations. However, as we can see for most initial values the jump-diffusion gives a more valuable contract.

To confirm these findings, a Monte Carlo simulation was done for this example with the same securities and mortality model. This was implemented in B.3. We simulated 10000 lifetimes and paths for the security and found the sample mean and discounted. For $S_0 = 1$ the sample mean was 3.1, the same as we found in using the implicit method.

8.2 Example 2: Death benefit

A death benefit contract is a very typcal life insurance policy previously described in more detail in example 5.4.1.

Consider a 40 year old woman. The contract has a duration of 20 years, and if the insured dies the contract pays the maximum of 1 (million) or the value of the fund at the time of death. This means that we have the following policy functions

$$\begin{aligned} f_\star &= 0, \\ g_\star &= 0, \\ h_{\star,\dagger}(S,\tau) &= \max(S_\tau, G). \end{aligned}$$

The function $f_\star$ is zero because the contract is worthless when it expires. There is also no continuous payment, and thus $g_\star$ is also zero. Since the only payment that can occur is the insured being payed when moving from state $\star$ to state $\dagger$ the value of the insurance is zero when the insured is in state $\dagger$. This means that we have the following Thiele's PIDE,

$$\begin{aligned} \frac{\partial}{\partial t} V_\star &= r_t V_\star - \mu_{\star,\dagger}(h_{\star,\dagger} - V_\star) - L V_\star \\ &= r_t V_\star - \mu_{\star,\dagger}(h_{\star,\dagger} - V_\star) - r_t x \frac{\partial}{\partial x} V_\star - \frac{1}{2} x^2 \sigma^2 \frac{\partial^2}{\partial^2 x} V_\star \\ &- \int_{\mathbb{R}\setminus\{0\}} \left[V_\star(t, x+y) - V_\star - xy \frac{\partial}{\partial x} V_\star \right] \nu(dy). \end{aligned}$$

Here the generator L is the same as in the above example. Calculating the value of insurances with a similar script as Appendix B.3, we get the following graph. The right graph is with $\lambda = 10$, i.e. with vigorous jump activity, the left is the case $\lambda = 0$, i.e. no jumps.

As we can see, the jumps "smooth out" the curve. The reason for this might be that the jumps make the fund more volatile. The concave shape seems to be made by the cash flow in the event of a death. The more volatile fund might overpower this and mute the effect. Hence creating a less concave graph.

We can compare the value of the policy at time 0 with jumps and without jumps.

S_0	Merton Jump-diffusion	Black and Scholes diffusion
1	0.0037	0.0041
2	0.0054	0.0064
3	0.0074	0.0084

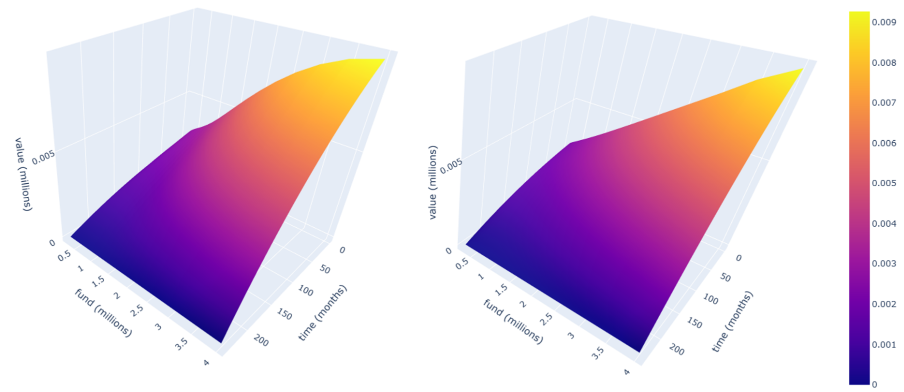

Figure 8.3: Value of a death benefit with jumps (R) and without jumps (L).

8.3 Example 3: Pension

A pension is a third very common life insurance policy. Most Norwegians have at least one pension policy: one from the government, often from an employer, and some also have private plans. Pensions generally give the insured some financial security during old age. This allows them to work less or not at all.

A pension plan generally works by having the insured pay a premium every period (every year or every month). Then after reaching an age determined by the contract, the insured receives some amount every period. The value of having a unit-linked policy as oppose to a regular fixed payment policy is that the insured can take on some of the financial risk, and thus might receive a cheaper policy, and higher returns.

In this examples we consider a 60 year old man. His policy stretches over 20 years, i.e. until he is 80. The policy has a retirement age of 70. When the insured turns 70 he will receive 1 (million) or the value of the fund, whichever is the highest, every year. When he reaches 80 the payments will cease.

The terminal condition therefore is 0. No matter how much the fund is worth there is not going to be any payments after 20 years. The upper boundary condition shows the value of the policy increasing until the age of 70, and decreasing after. The same is the case for lower boundary condition. This is because the insured only gets the pension if he survives. The closer to retirement age he is, the higher the probability of getting a pension is, and the more valuable the policy becomes. After the age of 70, there is only a limited number of payments, since the policy is of limited duration. Hence, the value should decrease for every time period. In addition to this, the probability of dying during the next period increases, making the policy worth less.

Consider first the case where the value of the fund is very high. Then the probability of receiving the value of the fund is also high. Hence, the upper boundary condition is based on the present value of receiving the value of the fund every year. On the other hand, the lower boundary condition is based on the present value of receiving 1 (million) every year. Both the upper and lower boundary have the behavior described above in regards to the value increasing until 70 and decreasing afterwards.

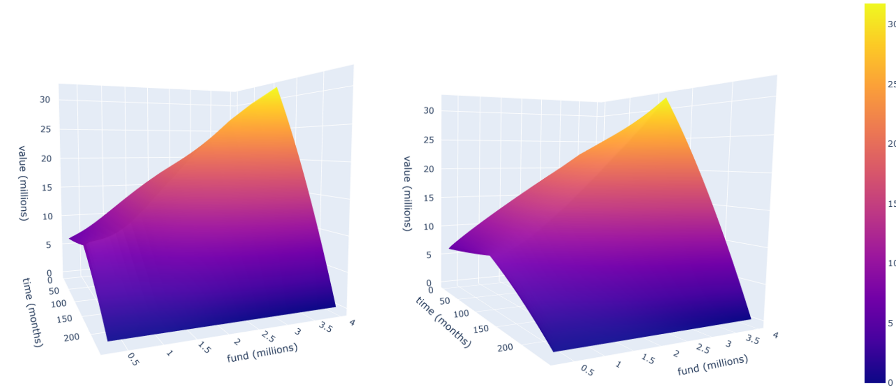

Figure 8.4: The value of a unit-linked pension with jumps (R) and without jumps(L).

This policy has the following policy functions.

$$\begin{aligned} f_\star &= 0, \\ g_\star(t, S_t) &= \max(S_t, 1)\mathbf{1}_{t\geq 70}, \\ h_{\star,\dagger} &= 0. \end{aligned}$$

As for the other policies discussed, there are no payments happening when the insured is in state †. Thus the value of the policy and the policy functions for this state is 0. From this we can derive the corresponding PIDE for the pension policy,

$$\begin{aligned} \frac{\partial}{\partial t}V_\star = r_t V_\star - g_\star(t,x) + \mu_{\star,\dagger}(t)V_\star - r_t x\frac{\partial}{\partial x}V_\star - \frac{1}{2}x^2\sigma^2\frac{\partial^2}{\partial^2 x}V_\star \\ + \int_{\mathbb{R}\setminus\{0\}}\left[V_\star(t,x+y) - V_\star(t,x) - xy\frac{\partial}{\partial x}V_\star\right]\nu(dy) \end{aligned}$$

Using a similar script as in Appendix B.3 we get Figure 8.3.

As we can see, both the model with and without jumps produce the typical "tent shape" in time as we would expect, where the value of the policy gradually increases until retirement and decreases afterwards. As the fund increases in value the model without jumps is more S-shaped, whereas the model with jumps has a flatter shape.

A possible explanation of this can be that when the fund is low, the risk in the model with jumps makes the policy more valuable. Since the jumps make the value more volatile, the probability of the fund increasing in value is higher. Since there is a lower bound on the payments the downside risk is bounded, but the upside risk is not, hence the value of the policy is higher. This would mean that we do not see the valley in the model with jumps as we see in the one without jumps.

When the fund is higher we see the opposite. The downside risk in the model with jumps is higher than in the model with jumps. Moreover, they have a similar upside risk. This means that the value of the more risky policy is lower. Hence, "cutting the mountain".

Lastly, we can compare the value of the insurance at time 0.

S_0	Merton Jump-diffusion	Black and Scholes diffusion
1	10.84	10.32
2	15.19	15.53
3	20.21	20.64

The table of initial values tells the same story.

8.4 Example 4: Commodity prices

The aim of this thesis has been to develop a PIDE for pricing insurance. However, if we remove our insurance hat, and change our perspective we might be able to apply this PIDE to a different context. A natural situation in which this could be appropriate is Norwegian fish farming. Fish have a limited time span where they grow before being harvested and sold. Here we have the same two sources of risk, mortality risk and risk surrounding payment. In reality there is also some risk as to how fast the fish grows, but we will not look at this here.

The first thing to consider is the mortality model. Obviously the fish does not follow the same mortality as humans. Hence, K2013 is not appropriate for this context. The public availability of data is a problem here. However, if one were to get mortality data one could use standard actuarial methods to estimate the mortality rate. Thus, we need to consult literature. Estimates vary greatly, but a few papers point to 5-10 death per 1000 fish each week, Oliveira et al. 2021. This changes based on time of year, temperature, and the geographic region. Going forward we assume that the mortality rate for a fish is $\mu_{\star,\dagger} = 0.05$. A more accurate model is possible, but this one is sufficient for this demonstration.

The next model we need is a model for the security. Luckily, there are some data on the price. The Fish pool index (FPI) is a weekly index used for settling the financial contracts at the Fish Pool exchange, *Fish Pool Index* 2022. Fish pool is an international exchange for fish and seafood commodities and related financial contracts. Using the FPI as a proxy for the value of the salmon at the time of sales, we can try to estimate the parameters of the Merton jump-diffusion. Using the code found Appendix B.3 we can find the following estimates.

μ	0.0093
σ^2	0.0047
μ_{jump}	-0.0752
σ_{jump}	0.7592
λ	0.8713

With all the modeling finnished we can move on to specifying the contract. Suppose the salmon will be harvested in 2 years if it survives. At the point of

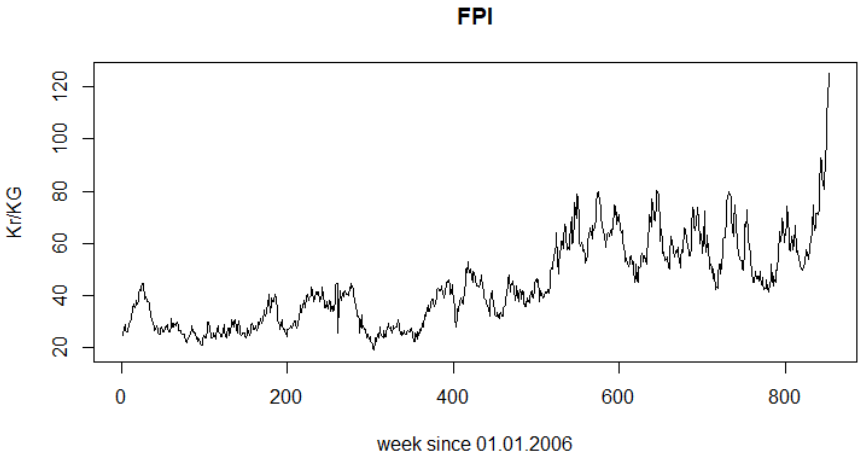

Figure 8.5: FPI every week from January 1. 2006

harvest the fish is 3 kg. If the fish is harvested then the payment is going to be random and decided by the Merton jump-diffusion. We suppose that the there is a contract where the fish is either sold for 50 or at the market price. This is not unreasonable when a fish farm is part of a bigger company. This means that the terminal condition is that of a call option. On the other hand, the lower boundary has to be equal to the future price adjusted for the probability of delivery, and the upper boundary has to have the imposed maximum value discounted and multiplied by the probability of payments.

The PIDE has the same general shape as in the case of the pure endowment,

$$\begin{aligned}\frac{d}{dt}V_\star &= rV_\star - \mu_{\star,\dagger}V_\star - LV_\star \\ &= rV_\star - \mu_{\star,\dagger}V_\star - r_t x\frac{\partial}{\partial x}V_\star - \frac{1}{2}x^2\sigma^2\frac{\partial^2}{\partial^2 x}V_\star \\ &- \int_{\mathbb{R}\backslash\{0\}}\Big[V_\star(t, x + h(t,x), t) - V_\star\Big]\end{aligned}$$

Running a script similar to the one we used for the pure endowment policy we get Figure 8.4. Due to limitations in the software used for plotting (or the use of said software) the values on the spot price-axis have disappeared.

However, the shape is the most important part. We can see that the lower the initial price, the lower the value of the fish. The initial value of the fish increases a lot with an increase in spot price. This means that the young fish have a more volatile value than older fish. On the other hand, when the spot price is high, then the fish has a very stable value.

The spot price as of time of writing is 119.69Nok/kg. This means that the value of a fish that is going to become 3 kg when harvested is today 296.22 Nok. We can summarize the initial value of fish with different spot prices.

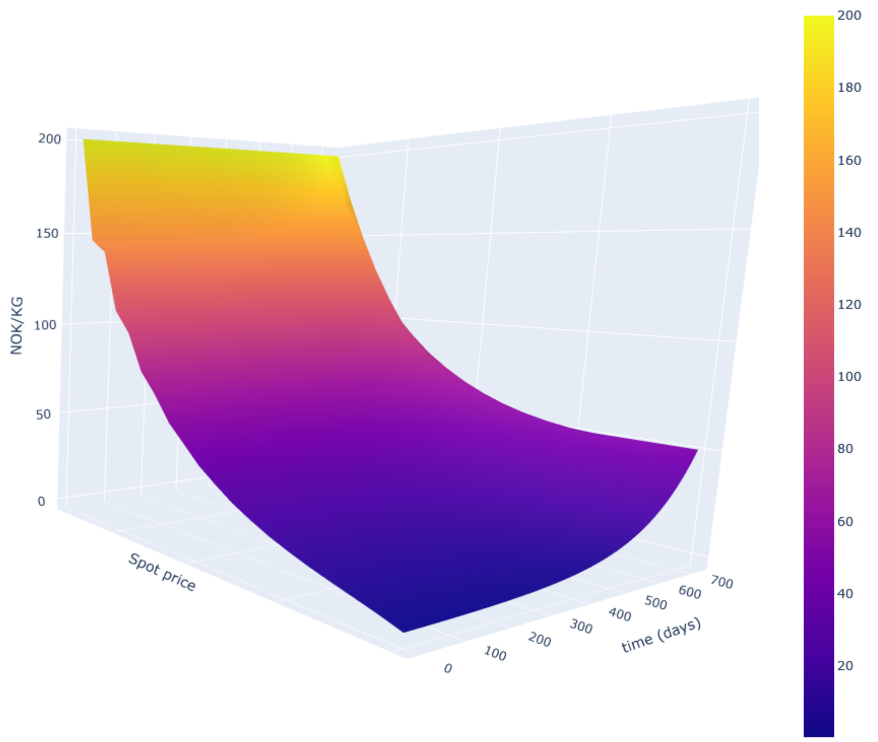

Figure 8.6: The value of 1 kg farm fish in Nok at every point in the growth process.

Spot price	Initial Value
12.5	0.27
30	23.09
110	98.70

We finish out this chapter by doing a simple sensitivity analysis on the value of the jump intensity parameter, λ.

8.5 Sensitivity analysis

In this section we are interested in understanding how the jump parameter λ can affect the value of a policy. We return to the first examples of this chapter, the pure endowment policy. We want to compare the value at time 0 of the policy when we keep all parameters equal except the jump intensity parameter λ.

If we consider this question intuitively we would expect a higher λ to make the unbderlying fund more risky. Since the jumps are positive and thus to the advantage of the insured, a more volatile fund makes the policy more valuable. Moreover, the downside risk is insured with a guarantee and the upside risk is not limited. Since the added risk is to the advantage of the insured, one would expect the policy to be more valuable when λ is increasing.

In order to test this hypotheses, we run the same script as we did in the example, but with different value for λ. We test from the no jump case of $\lambda = 0$ to the case where $\lambda = 20$. We get the following result.

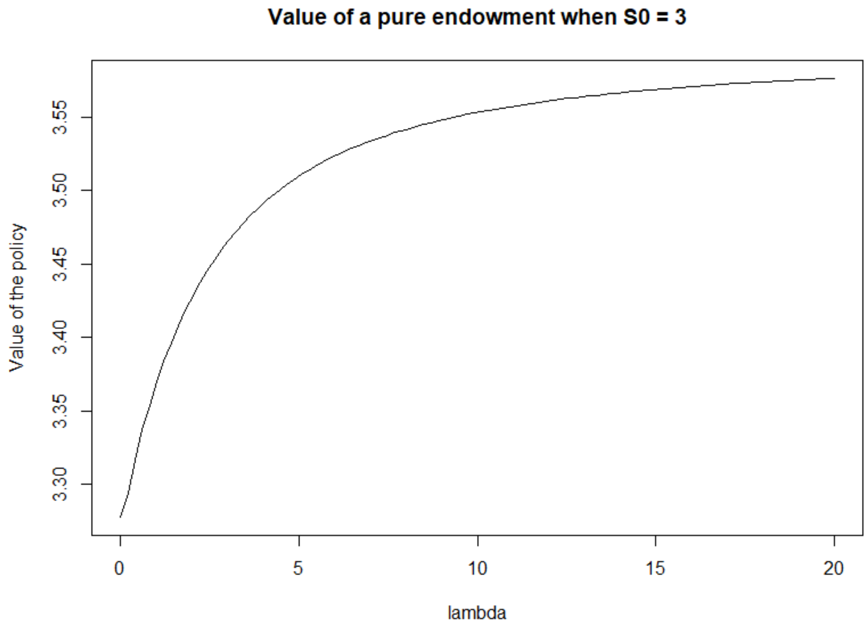

Figure 8.7: The value of a pure endowment for different jump intensities.

We see that it is in fact the case that the value at time 0 increases when the jump intensity increases. Whether this is for the intuitive reason as above stated need not be the case but it is a plausible explanation that would be in line with financial theory.

CHAPTER 9

Conclusion and discussions

In this chapter we will take a step back and discuss the results of the thesis. We will look at some possible further improvements and critical remarks.

9.1 Conclusion

The main goal of this thesis has been to generalize Thiele's differential equation for jump diffusion securities models. We accomplished this in Theorem 6.2.1. What is the value of such a result?

There is a previously known result that assumes the underlying security follows a Black and Scholes model, see Aase and Persson 1993. The Black and Scholes model is a standard model used by many practitioners. It is simple and fairly accurate in many situations which has helped keep it popular. However, there are some major problems. One of the most damning issues with the Black and Scholes model is that the model predict a distribution of the returns that have too thin tails. This means that major movements, gains or losses, have a vastly underestimated probability. This is a major problem when large losses can have devastating effects on a portfolio.

We can alleviate this problem by generalizing the security models to a jump diffusion. A jump diffusion is a model in which, in addition to a Brownian motion term, there is a Poisson random measure. The Poisson process can be thought of, for example, as a process that models big events like elections, pandemics, etc. The added jumps makes the tails of the distribution of the returns thicker. This makes the models more realistic.

In addition, by using a Brownian motion and a Poisson random measure, one can expand the set of possible models. This means that one can use a lot more realistic models.

For our insurance context, having a more realistic model for the underlying security means that we can improve the pricing of our policies. A more accurate model for the fund means a more accurate model of the cash flow to the insured. This allows us to place a more accurate value on the insurance. In particular, we saw in the previous chapter that when the jump intensity increased, the value of the pure endowment increased. In general, when the risk of the underlying fund increases, the value of the policy should change. Thus, a more realistic securities model makes the pricing better.

9.2 Further research

There are multiple ways to extend this research. In this section we will do some some preliminary brainstorming.

The most immediate avenue of research in this direction is to generalize even further. It should be possible to find even more general security models. One of the insights in this thesis is that having non-unique equivalent martingale measures does not pose a big problem. Thus, we should be able to find a Thiele's PIDE for even more general security models than a jump diffusion.

One interesting possible research project in this general theme is to study other kinds of models. One example could be the model proposed in Mandelbrot, Fisher, and Calvet 1997, and is based on the works of Benoit Mandelbrot. Mandelbrot is primary known for his work in fractals, and is possibly most known for the Mandelbrot set. However, much of his work has been in finance, notably in commodity pricing. The Multifractal Model of Asset Returns (MMAR) is based on a fractional Brownian motion with an subordinator, i.e. time is modeled by an increasing stochastic process. Since the fractional Brownian motion does not have independent increments, it cannot be a Lévy process. Thus, we could not use the main result of this thesis with such a model of the fund.

To take the possible research in a different direction, one could also extend this method to different kinds of financial contracts. What differentiates unit-linked contracts from other European contingent claims is the state of the insured and mortality risk. Consider on the other hand a ship. The ship can be in four different states: operating, in for repairs, in lay up, or sold. In each of these different states the ship has a cash flow. The cash flow when the ship is working is based on the rates. When the ship is in for repairs or in lay-up the cash flow is fixed. When the ship is scraped the cash flow is based on the value of the metal in the ship.

It could be interesting to try to model the movement between the states. This would requiere dealing with the interplay between multiple securties and possibly even weather. In the end this could lead to new methods of valuing ships. This could also be a tool for ship-owners to better plan the use of their vessels.

Lastly, a direct application of Theorem 6.2.1 is in commodities markets where the value of the commodity is modeled by a Lévy process, and there is a risk of delivery. If there is a risk that the delivery will not happen, the value of the commodity contract in the market should reflect this. The agricultural and aquaculture sectors are the most immediate possible uses for this result. Since what they produce are living thing that might not survive long enough to be sold. In addition, the sales price is subject to market fluctuations. It would allow the producers and the purchasers to factor in the mortality risk inherent in food production more accurately.

9.3 Critical remarks

We finish this thesis by making some critical remarks.

The first issue we want to raise is the fact that by introducing another stochastic process we also introduce new parameters. If we wanted to estimate the value of a unit-linked policy in the wild, it should be based on real securities. In order to do this we would have to estimate the parameters of the security model. In general, the more parameters, the harder the estimations are going to be.

In addition to this general difficulty, we also have introduced jumps. This raises another estimation problem. Namely, how to separate jumps from non-jumps. In order to do this one might want to resort to stochastic filtering techniques or solid threshold methods. This could make the estimation task more difficult: neither of these problems are insurmountable, but because of this we have not spent time doing very accurate estimates of parameters in this thesis. The work required could probably fill a second master thesis.

As a result of this statistical issue, we have not done parameter estimation in most of our examples. Since we used Merton's jump diffusion, it is possible to find an explicit likelihood-function one could optimize numerically. However, as this was not central to the thesis, it was not prioritized.

A different critical remark also concerning the examples is the choice of securities model. The Merton jump diffusion model might not have shown the extent of the usefulness of the PIDE. The choice was made on the basis of simplicity and the fact that the jump distribution was non-negative. Otherwise one would have to do a lot of extra work modeling the probability of positive and negative jumps.

The numerical methods can also be criticized. As with parameter estimation, the numerical methods are not the focus of the thesis and were thus not a priority. The explicit method proved to be very unstable. Some work could possibly be done to improve stability, but the jumps made the stability problems too extreme to be fixed in the short amount of time available. The implicit method proved a lot better on the stability front. However, there are also some stability problems here. The ideal choice in retrospect would have been a different method for numerical solution to PIDEs. The reason for not doing this was that the time it would take to implement and test adequately would have been too great.

A general critique is the accuracy in the calculations. This is a problem with any numerical method, and is only compounded if one were to use estimates based on real life data. Hence, the calculations do all have some degree of error, although it should not be great enough to change the results significantly.

There is one last criticism to add. No matter how much we improve the models, they can never be perfect. The actual value of an insurance model need the individual mortality of the insured and the true distribution of the fund. This is impossible with the tools available to us today. Even the insured does not know their own mortality, and when it comes to unit-lined policies, the financial markets are remarkably difficult to model. To this end we might conclude this thesis with the knowledge that we might have made some progress towards better understanding and valuation of insurance policies, but there is still a long way to go.

The words of George Box still plague us. All models are wrong, but some are useful.

Appendices

APPENDIX A

Some proofs

This appendix encompasses a lot of the proofs omitted in the main part of the thesis.

A.1 Chapter 2

Properties of a characteristic function

Proposition 2.2.7

Proof. We have to prove each point.

First, we want to find $|\varphi_X(u)|$.

$$|\varphi_X(u)| \leq \mathrm{E}[|e^{i<u,X>}|] = 1$$

this follows from Jensen's inequality and the fact that $e^{i<u,X>}$ is a compelx number of modulus 1.

Second point,

$$\varphi_X(0) = \mathrm{E}[e^0]) = 1$$

Third point follows from the fact that the exponential function is uniformly continuous, and thus the integral also has to be continuous.

Point four follows from the properties of the inner product.

$$\varphi_X(-u) = \mathrm{E}[e^{i<-u,X>}] = \mathrm{E}[e^{i<-1u,X>}] = \mathrm{E}[e^{-1i<u,X>}] = \mathrm{E}[e^{-i<u,X>}] = \overline{\varphi_X(u)}$$

Fifth point is a little bit harder. Suppose that μ_X is symmetric. Then X and $-X$ has the same distribution. However, this means that $\varphi_X(u) = \varphi_X(-u) = \overline{\varphi_X(u)}$. This is only possible if φ_X is a real function.

On the other hand, suppose that $\varphi_X(u)$ is a real function. Then, $\overline{\varphi_X(u)} = \varphi_X(u)$. However, the following also holds, $\overline{\varphi_X(u)} = \varphi_X(-u)$. This means that $\varphi_X(u) = \varphi_X(-u)$, implying that X and $-X$ has the same distribution.

■

Proof of Proposition 2.2.9

Proof. Suppose first that X and Y has the same distribution. That is, $\mu_X = \mu_Y$. Then,

$$\varphi_X(u) = \int_\Omega e^{i<u,x>}\mu_X(dx) = \int_\Omega e^{i<u,x>}\mu_Y(dx) = \varphi_Y(u)$$

The other implication follows from Lévy's inversion theorem Walsh 2012, p.159.

■

Proof of Proposition 2.2.11

Proof. First note that if $X_1, X_2, \dots, X_n$ are independent random variables with distributions $\mu_{X_1}, \mu_{X_2}, \dots, \mu_{X_n}$. Then,

$$\mu_{(X_1,X_2,\dots X_n)} = \mu_{X_1}\mu_{X_2}, \dots \mu_{X_n}$$

This is a generalization of Theorem 8.7.2 in Lindstrøm 2017, p. 315 and involves product measures but is a standard result from measure theory.

Then suppose first that $X_1, X_2, \dots, X_n$ are independent random variables. Then,

$$\begin{aligned}
\varphi_{X_1+X_2+\cdots X_n}(u) &= \mathrm{E}[e^{i<X_1+X_2+\cdots+X_n,u>}] \\
&= \int_{\mathbb{R}^d} e^{i<x_1+x_2+\cdots+x_n,u>}\mu_{X_1,X_2,\dots X_n}(dx) \\
&= \int_{\mathbb{R}^d} e^{i<x_1+x_2+\cdots+x_n,u>}\mu_{X_1}\mu_{X_2}, \dots \mu_{X_n}(dx) \\
&= \int_{\mathbb{R}^d} e^{i<x_1,u>}\mu_{X_1}e^{i<x_2,u>}\mu_{X_2}\cdots e^{i<x_n,u>}\mu_{X_n}(dx) \\
&= \mathrm{E}[e^{i<X_1,u>}]\,\mathrm{E}[e^{i<X_2,u>}]\cdots\mathrm{E}[e^{i<X_n,u>}] \\
&= \varphi_{X_1}(u)\varphi_{X_2}(u)\cdots\varphi_{X_n}(u).
\end{aligned}$$

Now on the other hand suppose that

$$\varphi_{X_1+X_2+\cdots X_n}(u) = \varphi_{X_1}(u)\varphi_{X_2}(u)\cdots\varphi_{X_n}(u).$$

By the calculation above, this implies that

$$\begin{aligned}
&\int_{\mathbb{R}^d} e^{i<x_1+x_2+\cdots+x_n,u>}\mu_{X_1,X_2,\dots X_n}(dx) \\
&= \int_{R^d} e^{i<x_1+x_2+\cdots+x_n,u>}\mu_{X_1}\mu_{X_2}, \dots \mu_{X_n}(dx)
\end{aligned}$$

which means that

$$\mu_{(X_1,X_2,\dots X_n)} = \mu_{X_1}\mu_{X_2}, \dots \mu_{X_n}.$$

This is equivalent with $X_1, X_2, \dots X_n$ being independent.

■

Proof of Proposition 2.4.3

Proof. We will prove the statements in the same order as in the proposition.

- This can be showed via the definition of the conditional expectation. Let A be any set in the $\sigma-$algebra $\mathcal{F}$.
$$\int_A \mathrm{E}[aX+bY|\mathcal{F}]dP = \int_A (aX+bY)dP = a\int_A XdP + b\int_A YdP$$
$$= a\int_A \mathrm{E}[X|\mathcal{F}]dP + b\int_A \mathrm{E}[Y|\mathcal{F}]dP.$$
- The tower property:
$$\mathrm{E}[\mathrm{E}[X|\mathcal{G}]|\mathcal{F}] = \mathrm{E}[X|\mathcal{F}].$$
The first equality follows from the fact that anything $\mathcal{G}$-measurable is also $\mathcal{F}$.
On the other hand, suppose that A is any set in $\mathcal{G}$, then A is also in $\mathcal{F}$. Hence,
$$\mathrm{E}[\mathrm{E}[X|\mathcal{F}]|\mathcal{G}] = \mathrm{E}[\mathrm{E}[X|\mathcal{F}]\mathbf{1}_A] = \mathrm{E}[\mathrm{E}[X\mathbf{1}_A|\mathcal{F}]] = \mathrm{E}[\mathrm{E}[X|\mathcal{G}]|\mathcal{F}].$$
- This is also called the law of iterative expectation. Let A be any set in $\mathcal{F}$. Then by definition,
$$\int_A \mathrm{E}[X|\mathcal{F}]dP = \int_A XdP.$$
The result follows by taking $A = \Omega$.
- Suppose that Y is a random variable. Let $B, A \in \mathcal{F}$. Now assume that X is $\mathcal{G}$-measurable and that $X = \mathbf{1}_A$ for any $A \in \mathcal{F}$,
$$\int_B \mathrm{E}[XY|\mathcal{F}]dP = \mathrm{E}[\mathrm{E}[1_A Y|\mathcal{F}]\mathbf{1}_B] = \mathrm{E}[\mathrm{E}[Y|\mathcal{F}]\mathbf{1}_{B\cap A}] = \mathrm{E}[\mathrm{E}[Y|\mathcal{F}]\mathbf{1}_A\mathbf{1}_B]$$
$$= \mathrm{E}[\mathrm{E}[Y|\mathcal{F}]X\mathbf{1}_B] = \int_B X\,\mathrm{E}[Y|\mathcal{F}]dP.$$
Now suppose that $X = \sum_{i\geq n}\alpha_i 1_{B_i}$, where $B_i \in \mathcal{F}$. Then the result follow from the linearity of conditional expectations.
Lastly, suppose X_i is a sequence of step functions converging to X from below. Then the result follows from the monotone convergance theorem and Proposition 7.5.3. in Lindstrøm 2017.
- Suppose that X is independent of $\mathcal{F}$. Suppose first that $X = \mathbf{1}_A$ and that B is any set in $\mathcal{F}$. Then,
$$\mathrm{E}[X|\mathcal{F}] = \mathrm{E}[\mathbf{1}_A 1\mathbf{1}_B] = \mathrm{E}[\mathbf{1}_A]\,\mathrm{E}[1\mathbf{1}_B] = \mathrm{E}[\mathbf{1}_A]\,\mathrm{E}[1|\mathcal{F}] = \mathrm{E}[X].$$
If one assumes that $X = \sum_{i\geq n}\alpha_i 1_{B_i}$, where $B_i \in \Omega$, then the result follows from linearity. Lastly, one can extend this to any integrable X with the monotone convergance theorem and Proposition 7.5.3. in Lindstrøm 2017.

- Let f be a convex function. Then there exists sequences of real numbers $\{a_n\}_{n\in\mathbb{N}}$ and $\{b_n\}_{n\in\mathbb{N}}$ such that $f(x) = \sup_{n\in\mathbb{N}}(a_n + b_n x)$.

$$\mathrm{E}[f(X)|\mathcal{F}] \geq \mathrm{E}\left[\sup_{n\in\mathbb{N}}(a_n + b_n X)|\mathcal{F}\right] = \sup_{n\in\mathbb{N}} a_n + b_n \mathrm{E}\left[X|\mathcal{F}\right] = f(\mathrm{E}[X|\mathcal{F}]).$$

■

Proof of Proposition 2.4.10

Proof. We can consider the first two properties in one $aX_t + Y_t$. X_t and Y_t are adapted by definition. Since a is a constant is has to be adapted to any $\sigma-$algebra. The boundedness follows from Jensen's inequality. Thus the last thing to check is the martingale condition. Let $s \leq t$,

$$\begin{aligned} \mathrm{E}[aX_t + Y_t|\mathcal{F}_s] &= a\,\mathrm{E}[X_t|\mathcal{F}_s] + \mathrm{E}[Y_t|\mathcal{F}_s] \\ &= aX_s + Y_s. \end{aligned}$$

■

A.2 Chapter 3

Proof of Corollary 3.1.5

This proof follow very closely the proof for Corollary 2.1.3 in Bédos 2019.

Proof. The case where $p = 1$ is simple, and it follows from the linearity of the integral.

Now assume that $p \in (1, \infty)$. The proof hinges on the Hölder inequality. Therefore, we first need to show that $|f + g|^{p-1} \in \mathcal{L}^q$. Note first that by the condition on p and q the following two things hold,

$$\begin{aligned} (p-1)q &= p, \\ \frac{p}{q} &= p - 1, \end{aligned}$$

$$|||f+g|^{p-1}||_q = \left(\int_\Omega |f+g|^p d\mu\right)^{1/q} = ||f+g||_p^{q/p} = ||f+g||_p^{p-1}.$$

Hence, since $f, g \in \mathcal{L}^p$, $|f+g|^{p-1} \in \mathcal{L}^q$.

Finally,

$$\begin{aligned} ||f+g||_p^p &= \int_\Omega |f+g|^p dP \\ &= \int_\Omega |f+g||f+g|^{p-1} dP \\ &= \int_\Omega |f||f+g|^{p-1} dP + \int_\Omega |g||f+g|^{p-1} dP \\ \text{Hölders inequality} &\leq ||f||_p|||f+g|^{p-1}||_q + ||g||_p|||f+g|^{p-1}||_q \end{aligned}$$

$$= (||f||_p||g||_p)|||f+g|^{p-1}||_q$$
$$= (||f||_p||g||_p)||f+g||_q|^{p-1}.$$

This implies that $||f+g||_p \leq ||f||_p||g||_p$.

■

Proof of Theorem: 3.4.2

Proof. The basic strategy used is to prove that the statement holds for all step functions and then taking the limit.

- Suppose first that $f, g \in \mathcal{S}$ with some partition $t_1, t_2, \ldots$. Then,

$$\begin{aligned}\int_0^t (\alpha f + \beta g) dX_t &= \sum_{i=1}^{n} (\alpha f + \beta g)\Delta X_i \\ &= \sum_{i=1}^{n} (\alpha f \Delta X_i + \beta g \Delta X_i) \\ &= \alpha \sum_{i=1}^{n} f(t_i)\Delta X_i + \beta \sum_{i=1}^{n} g(t_i)\Delta X_i \\ &= \alpha \int_0^t f dX_s + \beta \int_0^t g dX_s.\end{aligned}$$

  Since this holds for the case where $f, g \in \mathcal{S}$ we can show by a limit argument that it also holds for $f, g \in \mathcal{H}^2$. Suppose $f_n \to f$ and $g_n \to g$. Then,

$$\begin{aligned}\int_0^t (\alpha f + \beta g) dX_t &= \lim_{n\to\infty} \int_0^t (\alpha f_n + \beta g_n) dX_t \\ &= \lim_{n\to\infty} \alpha \int_0^t f_n dX_s + \beta \int_0^t g_n dX_s \\ &= \alpha \int_0^t f dX_s + \beta \int_0^t g dX_s.\end{aligned}$$

- Now suppose that $0 \leq u < t$, and $f_n \in \mathcal{S}$ such that $f_n \to f$ where $f \in \mathcal{H}^2$.

$$\begin{aligned}\int_0^u f dX_s + \int_u^t f dX_s &= \lim_{n\to\infty} \int_0^u f_n dX_s + \int_u^t f_n dX_s \\ &= \lim_{n\to\infty} \sum_{i=1}^{m} f_n(t_i)\Delta X_i + \sum_{i=m+1}^{k} f_n(t_i)\Delta X_i \\ &= \lim_{n\to\infty} \sum_{i=1}^{k} f_n(t_i)\Delta X_i \\ &= \lim_{n\to\infty} \int_0^t f_n dX_s \\ &= \int_0^t f dX_s.\end{aligned}$$

Here we did both steps of the proof in one, but the general strategy is the same as in the first point.

- See Theorem 19 in chapter II.5 of Protter 2005
- Suppose that X is a martingale. Suppose that $f \in \mathcal{S}$.

$$
\begin{aligned}
\mathrm{E}\left[\int_0^t f(s) dX_s\right] &= \mathrm{E}\left[\sum_{i\geq 1} f_{t_i}(X_{t_{i+1}} - X_{t_i})\right] \\
&= \sum_{i\geq 1} f_i \, \mathrm{E}\left[(X_{t_{i+1}} - X_{t_i})\right] \\
&= 0.
\end{aligned}
$$

Now let $\{f_n\}_{n\geq 1}$ be a sequence of step functions converging to f from below. Then,

$$
\begin{aligned}
\mathrm{E}\left[\int_0^t f(s) dX_s\right] &= \mathrm{E}\left[\int_0^t \lim_{n\to\infty} f_n(s) dX_s\right] \\
\text{DCT} &= \lim_{n\to\infty} \mathrm{E}\left[\int_0^t f_n(s) dX_s\right] \\
&= 0.
\end{aligned}
$$

- The mesurability follows from the fact that f is progressive and hence measurable on $\mathcal{F}_t$ and the fact that since X is a martingale it has to be adapted.
- Let $\{\tau_i\}_{i\in\mathbb{N}}$ be the sequence for which $X_{t\wedge\tau}$ is a martingale. We check the martingale definition.

  1. Adaptability follows from an earlier property.
  2. Next comes finite absolute expectation:

$$
\begin{aligned}
\mathrm{E}\left[\left|\int_0^t f(s) dX_s\right|\right] &= \mathrm{E}\left[\left|\lim_{n\to\infty}\int_0^t f_n(s) dX_s\right|\right] \\
\text{(DCT)} &= \lim_{n\to\infty} \mathrm{E}\left[\left|\int_0^t f_n(s) dX_s\right|\right] \\
&= \lim_{n\to\infty} \mathrm{E}\left[\left|\sum_{i=1}^m h_i \Delta X_i\right|\right] \\
\text{(TI)} &\leq \lim_{n\to\infty} \sum_{i=1}^m h_i \, \mathrm{E}\left[\left|\Delta X_i\right|\right] \\
&< \infty.
\end{aligned}
$$

3. Lastly we have the martingale property. Let $s \leq t$ and $s, t \in \tau$.

$$\begin{aligned}
\mathrm{E}\left[\int_0^t f(u)dX_u|\mathcal{F}_s\right] &= \mathrm{E}\left[\int_0^s f(u)dX_u + \int_s^t f(u)dX_u|\mathcal{F}_s\right] \\
&= \int_0^s f(u)dX_u + \mathrm{E}\left[\int_s^t f(u)dX_u|\mathcal{F}_s\right] \\
&= \int_0^s f(u)dX_u + \mathrm{E}\left[\int_s^t f(u)dX_u\right] \\
&= \int_0^s f(u)dX_u.
\end{aligned}$$

■

Proof of Theorem 3.8.1

Proof. The first claim we are going to prove is the uniqueness.

Suppose first that there are two different solutions to the SDE X and $\tilde{X}$ with initial condition $Z = \tilde{Z}$. Let, $a(s,\omega) = \mu(s, X(s,\omega)) - \tilde{\mu}(s, \tilde{X}(s,\omega))$, $b(s,\omega) = \sigma(s, \tilde{X}(s,\omega)) - \sigma(s, X(s,\omega))$ and $c(\omega, s, z) = h(\tilde{X}(s,\omega), s, z) - h(X(s,\omega), z)$. Then,

$$\begin{aligned}
\mathrm{E}[|X(t) - \tilde{X}(t)|^2] &= \mathrm{E}\Bigg[\Bigg(Z - \tilde{Z} + \int_0^T a(s)ds + \int_0^T b(s)dW(s) \\
&+ \int_0^T \int_{\mathbb{R}\setminus\{0\}} c(t,z)\tilde{N}(dt,dz)\Bigg)^2\Bigg] \\
&= \mathrm{E}[(Z - \tilde{Z})^2 + 2(Z - \tilde{Z})\int_0^T a(s)ds \\
&+ 2(Z - \tilde{Z})\int_0^T b(s)dW(s)) \\
&+ 2(Z - \tilde{Z})\int_0^T \int_{\mathbb{R}\setminus\{0\}} c(t,z)\tilde{N}(dt,dz) \\
&+ 2\int_0^T a(s)ds \int_0^T \int_{\mathbb{R}\setminus\{0\}} c(t,z)\tilde{N}(dt,dz) \\
&+ 2\int_0^T b(s)dW(s) \int_0^T \int_{\mathbb{R}\setminus\{0\}} c(t,z)\tilde{N}(dt,dz) \\
&+ \left(\int_{\mathbb{R}\setminus\{0\}} c(t,z)\tilde{N}(dt,dz)\right)^2 \\
&+ (\int_0^T a(s)ds)^2 + \int_0^T a(s)ds \int_0^T b(s)dW(s) \\
&+ \left(\int_0^T b(s)dW(s)\right)^2\Bigg] \\
&= \mathrm{E}[(Z - \tilde{Z})^2] + 2\,\mathrm{E}[(Z - \tilde{Z})\int_0^T a(s)ds] \\
&+ 2\,\mathrm{E}[(Z - \tilde{Z})\int_0^T b(s)dW(s))]
\end{aligned}$$

$$
\begin{aligned}
&+ 2\,\mathrm{E}[(Z-\tilde{Z}) \int_0^T \int_{\mathbb{R}\setminus\{0\}} c(t,z)\tilde{N}(dt,dz)] \\
&+ 2\,\mathrm{E}[\int_0^T a(s)ds \int_0^T \int_{\mathbb{R}\setminus\{0\}} c(t,z)\tilde{N}(dt,dz)] \\
&+ 2\,\mathrm{E}[\int_0^T b(s)dW(s) \int_0^T \int_{\mathbb{R}\setminus\{0\}} c(t,z)\tilde{N}(dt,dz)] \\
&+ \mathrm{E}[\left(\int_{\mathbb{R}\setminus\{0\}} c(t,z)\tilde{N}(dt,dz)\right)^2] \\
&+ \mathrm{E}[(\int_0^T a(s)ds)^2] + \mathrm{E}[\int_0^T a(s)ds \int_0^T b(s)dW(s)] \\
&+ \mathrm{E}[(\int_0^T b(s)dW(s))^2] \\
&\leq 3\,\mathrm{E}[(Z-\tilde{Z})^2] + 3\,\mathrm{E}[(\int_0^T a(s)ds)^2] + 3\,\mathrm{E}[(\int_0^T b(s)dW(s))^2] \\
&+ 3\,\mathrm{E}[\left(\int_{\mathbb{R}\setminus\{0\}} c(t,z)\tilde{N}(dt,dz)\right)^2] \\
&\leq 3\,\mathrm{E}[(Z-\tilde{Z})^2] + 3(1+t)D^2 \int_0^T \mathrm{E}[|X_s - \tilde{X}_s|^2]ds \\
&\leq 0
\end{aligned}
$$

Remark A.2.1.

- Three cross-terms disappear in the third equality because the covariance is zero. The product of the expectations in all cases is also zero.
- The inequality happens because we multiply all therm by 3. This seems strange, but is needed for Gronwall's inequality.
- In the penultimate inequality we just rewrite $3\,\mathrm{E}[(\int_0^T a(s)ds)^2] + 3\,\mathrm{E}[(\int_0^T b(s)dX^c(s))^2] + \mathrm{E}\left[\left(\int_{\mathbb{R}\setminus\{0\}} c(t,z)\tilde{N}(dt,dz)\right)^2\right] = 3(1+t)D^2 \int_0^T \mathrm{E}[|X_s - \tilde{X}_s|^2]ds$. This is not hard but tedious and follows from the linear growth condition and the definitions of a, b and c.
- The last inequality follows form Gronwall's lemma which stated that if $u(t) \leq F + A\int_0^T u(t)ds$, then $u(t) \leq Fe^{At}$. Here $u(t) = \mathrm{E}[|X(t) - \tilde{X}(t)|^2]$, $F = 3\,\mathrm{E}[(Z-\tilde{Z})^2]$ and $A = 3(1+t)D^2 \int_0^T \mathrm{E}[|X_s - \tilde{X}_s|^2]ds$. Since $Z = \tilde{Z}$ we see that $F = 0$ and the result follows.

This proves the uniqueness condition. One could also argue that this gives a method for calculating the distance between two solutions given different initial conditions.

We can now move on to existence. This will be done via Picard iterations. Let $Y^0(t) = X_0$ and define recursively the sequence of $Y^{(k)}(t)$ as follows,

$$
Y^{(k)}(t) = X(0) + \int_0^t \mu(s, Y^{(k-1)}(s))ds + \int_0^t \sigma(s, Y^{(k-1)}(s))dW(s)
$$

$$+\int_0^t \int_{\mathbb{R}\setminus\{0\}} h(z, Y^{(k-1)}(s))\tilde{N}(dz, dt).$$

By a similar calculation as above we can see that,

$$\mathrm{E}[|Y^{(k)}(t) - Y^{(k-1)}(t)|^2] \leq (1+T)3D^2 \int_0^t \mathrm{E}[|Y^{(k)}(s) - Y^{(k-1)}(s)|^2]ds.$$

One can see this by defining a and b in the same way as above with $Z - \tilde{Z} = X_0 - \tilde{X}_0 = 0$.

We now want to show that the distance between each iteration is sufficiently small to create a contraction mapping. To do this we use induction. Consider the difference between the 0th and first iteration

$$\begin{aligned}
\mathrm{E}[|Y^{(1)}(t) - Y^{(0)}(t)|^2] &= \mathrm{E}[|\int_0^t \mu(s, X_0)ds + \int_0^t \sigma(s, X_0)dW(s)|2] \\
&\leq 2\,\mathrm{E}[t\int_0^t |\mu(s, X_0)|^2 ds + \int_0^t |\sigma(s, X_0)|^2 dW(s)] \\
&\leq 2\,\mathrm{E}[(t+1)\int_0^t C(1+|X_0|^2)ds] \\
&\leq 2(t+1)tC(1+\mathrm{E}[|X_0|^2])] \\
\leq A_1 t.
\end{aligned}$$

Denote this $A_1 t$. Next by induction,

$$\mathrm{E}[|Y^{(k+1)}(t) - Y^{(k)}(t)|^2] \leq \frac{A_2^{k+1} t^{k+1}}{(k+1)!}.$$

This calculation is similar to the zero case.

Let $||\cdot||$ be the L^2 norm, then for any $n, m \in \mathbb{N}$ we have that,

$$\begin{aligned}
||Y^{(m)}(t) - Y^{(n)}(t)|| &= ||\sum_{k=n}^{m-1}(Y^{k+1}(t) - Y^k(t))|| \\
&\leq \sum_{k=n}^{m-1} ||(Y^{k+1}(t) - Y^k(t))|| \\
&= \sum_{k=n}^{m-1} ||\,\mathrm{E}[\int_0^T |Y^{k+1}(s) - Y^k(s)|^2 ds]^{1/2} \\
&= \sum_{k=n}^{m-1} (\int_0^T \frac{A_2^{k+1} t^{k+1}}{(k+1)!})^{1/2} \\
&= \sum_{k=n}^{m-1} (\frac{A_2^{k+1} T^{k+1}}{(k+2)!})^{1/2}
\end{aligned}$$

The faculty term explodes and this series converges to zero as $n, m \to \infty$. However, this means that $\{Y^{(k)}(t)\}_{k\in\mathbb{N}}$ forms a Cauchy sequence in L^2. This means that it converges in L^2 since L^2 is complete. We then define

$$X(t) \coloneqq \lim_{k\to\infty} Y^{(k)}(t).$$

Now we only need to show that this does indeed satisfy the SDE.

$$\begin{aligned}Y^{(n+1)}(t) = X_0 + \int_0^t \mu(s, Y^{(n)}(s))ds + \int_0^t \sigma(s, Y^{(n)}(s))dW(s) \\ + \int_0^t h(s, Y^{(n)}(s), z)\tilde{N}(dt, dz).\end{aligned}$$

Note that by Hölder's inequality and the Itô isometry,

$$\begin{aligned}\int_0^t \mu(s, Y^{(n)}(s))ds &\to \int_0^t \mu(s, X(s))ds, \\ \int_0^t \sigma(s, Y^{(n)}(s))dW(s) &\to \int_0^t \sigma(s, X(s))dW(s), \\ \int_0^t h(s, Y^{(n)}(s), z)\tilde{N}(dt, dz) &\to \int_0^t h(s, X(s), z)\tilde{N}(dt, dz).\end{aligned}$$

Hence,

$$\begin{aligned}Y^{(n+1)}(t) = X_0 + \int_0^t \mu(s, Y^{(n)}(s))ds + \int_0^t \sigma(s, Y^{(n)}(s))dW(s) \\ + \int_0^t h(s, Y^{(n)}(s), z)\tilde{N}(dt, dz),\end{aligned}$$

which converge to

$$\begin{aligned}&X_0 + \int_0^t \mu(s, X(s))ds + \int_0^t \sigma(s, X(s))dW(s) \\ &+ \int_0^t h(s, X(s), z)\tilde{N}(dt, dz) \\ &= X(t)\end{aligned}$$

■

A.3 Chapter 4

Proof of Proposition 4.4.2

Proof. We want to prove that this condition is sufficient to ensure that the stochastic exponential is a martingale.

Consider fist the implication that if Z is a martingale, then $\mathrm{E}(\mathcal{E}(S)(t)) = 1$. This is almost trivial.

$$\mathrm{E}[\mathcal{E}(S)(t)] = \mathrm{E}[\mathcal{E}(S)(t)|\mathcal{F}_0] = \mathcal{E}(S)(0) = Z_0 = 1.$$

Now consider the other implication. Let $\{T_n\}$ be a sequence of times such that $\mathcal{E}(S)(t \wedge T_n)$ is a local martingale. Then by Fatou's lemma,

$$\mathrm{E}[\mathcal{E}(S)(t)|\mathcal{F}_s] = \mathrm{E}[\liminf_{n\to\infty} \mathcal{E}(S)(t \wedge T_n)|\mathcal{F}_s]$$

$$\leq \liminf_{n\to\infty} \mathrm{E}[\mathcal{E}(S)(t \wedge T_n)|\mathcal{F}_s]$$
$$= \liminf_{n\to\infty} \mathcal{E}(S)(s \wedge T_n).$$

This means that $\mathcal{E}(S)(s \wedge T_n)$ is a supermartingale.

We now need to prove that if in addition $\mathrm{E}[\mathcal{E}(S)(t)] = 1$, then $\mathcal{E}(S)(t)$ is martingale. Consider $\mathcal{E}(S)(s) - \mathrm{E}[\mathcal{E}(S)(t)|\mathcal{F}_s]$. We can take the expectation of this.

$$\mathrm{E}[\mathcal{E}(S)(s)] - \mathrm{E}[\mathrm{E}[\mathcal{E}(S)(t)|\mathcal{F}_s]] = 1 - \mathrm{E}[\mathcal{E}(S)(t)] = 0$$

But this imples that,

$$\mathcal{E}(S)(s) = \mathrm{E}[\mathcal{E}(S)(t)|\mathcal{F}_s]$$

This proves the martingality. Note that $\mathcal{E}(S)(t)$ is adapted and has finite expectation because of supermartingality.

■

Proof of Theorem 4.4.3

Proof. We here only prove the one dimensional case. First, we need to prove that Z_t defines an equivalent probability measure Q. By the Radon-Nikodym theorem $Q(A) = \int_A Z_T dP$.

Suppose first that A is such that $P(A) = 0$. Then, $Q(A) = \int_A dQ = \int_A Z_T dP = 0$. On the other hand, let $Q(A) = 0$. $0 = \int_A (Z_T)^{-1} dQ = \int_A dP = P(A)$. Thus, we have proven that $Q \sim P$.

We now need to show that $Y(t) = S(t)$ is a Q-local martingale. To do this it suffices to show that $Y(t)Z(t)$ is a P-local martingale.

By Itô's formula,

$$\begin{aligned}
d(S(t)Z(t)) &= Z(t-)dS(t) + S(t-)dZ(t) + d[S,Z](t) \\
&= Z(t-)S(t)(\mu(t,S(t))dt + \sigma(t,S(t))dW(t) \\
&+ \int_{\mathbb{R}\setminus\{0\}} h(t,S(t),z)\tilde{N}(dt,dz)) \\
&+ S(t-)Z(t)(-\theta_0(t)dW(t) - \int_{\mathbb{R}\setminus\{0\}} \theta_1(z,t)\tilde{N}(dz,dt)) \\
&- Z(t-)\sigma(t,S(t))\theta_0(t)dt - Z(t-)h(t,S(t),z)\theta_1(z,t)N(dt,dz) \\
&= Z(t-)(\mu(t,S(t)) - \sigma(t,S(t))\theta_0(t) - h(t,S(t),z)\theta_1(z,t)\nu(dz))dt \\
&+ Z(t-)(\mu(t,S(t)) - \sigma(t,S(t))\theta_0(t))dW_t \\
&+ Z(t-)\int_{\mathbb{R}\setminus\{0\}} (h(t,S(t),z) - \theta_1(t,z)S(t-) \\
&- h(t,S(t),z)\theta_1(t,z))\tilde{N}(dt,dz) \\
&= Z(t-)(\mu(t,S(t)) - \sigma(t,S(t))\theta_0(t))dW_t \\
&+ Z(t-)\int_{\mathbb{R}\setminus\{0\}} (h(t,S(t),z) - \theta_1(t,z)S(t-) \\
&- h(t,S(t),z)\theta_1(t,z))\tilde{N}(dt,dz)
\end{aligned}$$

This expression only has local martingale terms, and is therefore a local martingale. Since this product is a P-local martingale then $X(t)$ has to be a Q-local martingale.

To check that the product $S(t)Z(t)$ has bounded absolute moments we can note that they both are sums of well defined stochastic integrals (with a few bounded deterministic integrals), hence by the triangle inequality the expected absolute value also needs to be bounded.

We now need to check the second part, that dW_t^Q is a Q-Brownian motion and $\tilde{N}^Q(dt, dz)$ is a compensated random measure of N.

Consider first dW_t^Q. A proof of this can be found in Baldi 2017, p. 368.

We now need to prove that $\tilde{N}^Q(dt, dz)$ is a compensated random measure of N under Q. First let,

$$da_t = \int_{\mathbb{R}\setminus\{0\}} \gamma(z,t)\tilde{N}(dz,dt).$$
$$db_t = \int_{\mathbb{R}\setminus\{0\}} \gamma(z,t)\theta_1(z,t)\nu(dz)dt.$$

Let $M_t = a_t + b_t$. Using Itô's formula,

$$\begin{aligned}
dM_t^2 &= d((a_t + b_t)^2) \\
&= 2(a_{t-} + b_{t-})(da_t + db_t) + d[M, M]_t \\
&= 2(a_{t-} + b_{t-})da_t + (a_{t-} + b_{t-})db_t + d[a+b, a+b]_t \\
&= 2(a_{t-} + b_{t-}) \int_{\mathbb{R}\setminus\{0\}} \gamma(z,t)\tilde{N}(dz,dt) \\
&+ 2(a_{t-} + b_{t-}) \int_{\mathbb{R}\setminus\{0\}} \gamma(z,t)\theta_1(z,t)\nu(dz)dt \\
&+ d[a, a]_t \\
&= 2(a_{t-} + b_{t-}) \int_{\mathbb{R}\setminus\{0\}} \gamma(z,t)\tilde{N}(dz,dt) \\
&+ 2(a_{t-} + b_{t-}) \int_{\mathbb{R}\setminus\{0\}} \gamma(z,t)\theta_1(z,t)\nu(dz)dt \\
&+ \int_{\mathbb{R}\setminus\{0\}} \gamma(t,z)^2\tilde{N}(dz,dt) + \int_{\mathbb{R}\setminus\{0\}} \gamma(t,z)^2\nu(dz)dt \\
&= \int_{\mathbb{R}\setminus\{0\}} [2(a_{t-} + b_{t-})\gamma(z,t) + \gamma(t,z)^2]\tilde{N}(dz,dt) \\
&+ [2(a_{t-} + b_{t-}) \int_{\mathbb{R}\setminus\{0\}} (\gamma(z,t)\theta_1(z,t)\nu(dz) + \int_{\mathbb{R}\setminus\{0\}} \gamma(t,z)^2\nu(dz)]dt.
\end{aligned}$$

Now denote the process R_t,

$$R_t = M_t^2 - \int_0^t \int_{\mathbb{R}\setminus\{0\}} \gamma(z,s)^2(1 - \theta_1(s,x))\nu(dz)ds.$$

We can find the dynamics of R_t,

$$dR_t = dM_t^2 - \int_{\mathbb{R}\setminus\{0\}} \gamma(z,t)^2(1 - \theta_1(t,z))\nu(dz)dt$$

$$= \int_{\mathbb{R}\backslash\{0\}} [2(a_{t-} + b_{t-})\gamma(z,t) + \gamma(t,z)^2]\tilde{N}(dz,dt)$$
$$+ [2(a_{t-} + b_{t-}) \int_{\mathbb{R}\backslash\{0\}} (\gamma(z,t)\theta_1(z,t)\nu(dz) + \int_{\mathbb{R}\backslash\{0\}} \gamma(t,z)^2\nu(dz)]dt$$
$$- \int_{\mathbb{R}\backslash\{0\}} \gamma(z,t)^2(1-\theta_1(t,z))\nu(dz)dt$$
$$= \int_{\mathbb{R}\backslash\{0\}} [2(a_{t-} + b_{t-})\gamma(z,t) + \gamma(t,z)^2]\tilde{N}(dz,dt)$$
$$+ [2(a_{t-} + b_{t-}) \int_{\mathbb{R}\backslash\{0\}} (\gamma(z,t)\theta_1(z,t)\nu(dz) - \int_{\mathbb{R}\backslash\{0\}} \gamma(t,z)^2\theta_1(z,t\nu(dz)]dt.$$

Lastly, we turn our eyes to Z_t and seek to find the dynamics of Z_tR_t.

$$d(Z_tR_t) = Z_{t-}dR_t + R_t dZ_{t-} + d[Z,R]_t$$
$$= Z_{t-} \int_{\mathbb{R}\backslash\{0\}} [2(a_{t-} + b_{t-})\gamma(z,t) + \gamma(t,z)^2]\tilde{N}(dz,dt)$$
$$+ Z_{t-}[2(a_{t-} + b_{t-}) \int_{\mathbb{R}\backslash\{0\}} (\gamma(z,t)\theta_1(z,t)\nu(dz) - \int_{\mathbb{R}\backslash\{0\}} \gamma(t,z)^2\theta_1(z,t\nu(dz)]dt$$
$$- R_{t-}Z_t(\theta_0 dW_t + \int_{\mathbb{R}\backslash\{0\}} \theta_1(t,z)\tilde{N}(dz,dt))$$
$$- Z_{t-} \int_{\mathbb{R}\backslash\{0\}} (2(a_t + b_t)\gamma(t,z) + \gamma(t,z)^2)\theta_1(t,z)N(dt,dz)$$
$$= Z_{t-} \int_{\mathbb{R}\backslash\{0\}} [2(a_{t-} + b_{t-})\gamma(z,t) + \gamma(t,z)^2]\tilde{N}(dz,dt)$$
$$- R_{t-}Z_t\theta_0 dW_t - R_{t-}Z_t \int_{\mathbb{R}\backslash\{0\}} \theta_1(t,z)\tilde{N}(dz,dt)$$
$$- Z_{t-} \int_{\mathbb{R}\backslash\{0\}} (2(a_t + b_t)\gamma(t,z) + \gamma(t,z)^2)\theta_1(t,z)\tilde{N}(dt,dz).$$

■

Proof of Theorem 4.5.4

Proof. Let us suppose first that X is replicable. Then there exists an admissible and self-financing portfolio such that,

$$\tilde{X} = \mathrm{E}_Q[\tilde{X}] + \int_0^T \eta(t)d\tilde{S}(t)$$
$$= \mathrm{E}_Q[\tilde{X}] + \int_0^T \eta(t)e^{\int_0^t r_s ds}\left(\sigma(t)dB^Q(t) + \int_{\mathbb{R}\backslash\{0\}} h(z,t)\tilde{N}^Q(dz,dt)\right)$$
$$= \mathrm{E}_Q[\tilde{X}] + \int_0^T \eta(t)e^{\int_0^t r_s ds}\sigma(t)dB^Q(t) + \int_0^T \int_{\mathbb{R}\backslash\{0\}} \eta(t)e^{\int_0^t r_s ds}h(z,t)\tilde{N}^Q(dz,dt)$$
$$= \mathrm{E}_Q[\tilde{X}] + \int_0^T \psi(t)dB^Q(t) + \int_0^T \int_{\mathbb{R}\backslash\{0\}} \gamma(z,t)\tilde{N}^Q(dz,dt).$$

Here we have used the martingale representation theorem. On the other hand, assume now that that there exists some η under the conditions of the theorem. Let $\tilde{S}$ be a martingale. Then by definition of X,

$$\begin{aligned}
\tilde{X} &= \mathrm{E}_Q[\tilde{X}] + \int_0^T \psi(t)dB^Q(t) + \int_0^T \int_{\mathbb{R}\setminus\{0\}} \gamma(z,t)\tilde{N}^Q(dz,dt) \\
&= \mathrm{E}_Q[\tilde{X}] + \int_0^T e^{\int_0^t r_s ds}\eta(t)\sigma(t)dB^Q(t) + \int_0^T \int_{\mathbb{R}\setminus\{0\}} e^{\int_0^t r_s ds}\eta(t)h(z,t)\tilde{N}^Q(dz,dt) \\
&= \mathrm{E}_Q[\tilde{X}] + \int_0^T e^{\int_0^t r_s ds}\eta(t)\sigma(t)dB^Q(t) + \int_0^T \int_{\mathbb{R}\setminus\{0\}} e^{\int_0^t r_s ds}\eta(t)h(z,t)\tilde{N}^Q(dz,dt) \\
&= \mathrm{E}_Q[\tilde{X}] + \int_0^T e^{\int_0^t r_s ds}\eta(t)[\sigma(t)dB^Q(t) + \int_{\mathbb{R}\setminus\{0\}} h(z,t)\tilde{N}^Q(dz,dt)] \\
&= \mathrm{E}_Q[\tilde{X}] + \int_0^T e^{\int_0^t r_s ds}\eta(t)dS(t) \\
&= \mathrm{E}_Q[\tilde{X}] + \int_0^T \eta(t)d\tilde{S}(t).
\end{aligned}$$

This also uses the martingale representation theorem. Hence there is a replicating portfolio. ■

A.4 Chapter 5

Proof of Lemma 5.2.1

Proof. The second expression follows from Fubini's theorem.

$$\mathrm{E}[\int_t^\infty \mathbf{1}_{\{X_s=j\}}dc(s)|X_t=i] = \int_t^\infty \mathrm{E}[\mathbf{1}_{\{X_s=j\}}dc(s)|X_t=i] = \int_t^\infty p_{i,j}(t,s)dc(s).$$

For a stochastic version of Fubini, see Filipović 2009, p. 99.

The first expression is a bit more work. Suppose first that b is a step function such that $b(s) = \mathbf{1}_{[\alpha,\beta)}(s)$. Define,

$$g(s) \coloneqq \mathrm{E}[N_{j,k}^X(s)|X_t=i], \quad s \geq t.$$

Now consider the increment from s to $s+h$. Then by the Markov property,

$$\begin{aligned}
g(s+h) - g(s) &= \mathrm{E}[N_{j,k}^X(s+h) - N_{j,k}^X(s)|X_t=i] \\
&= \sum_{l\in S} \mathrm{E}[\mathbf{1}_{X_s=l}(N_{j,k}^X(s+h) - N_{j,k}^X(s))|X_t=i] \\
&= \sum_{l\in S} \frac{1}{P(X_t=i)} \mathrm{E}[\mathbf{1}_{X_s=l}\mathbf{1}_{X_t=i}(N_{j,k}^X(s+h) - N_{j,k}^X(s))] \\
&= \sum_{l\in S} \frac{P(X_s=l)}{P(X_t=i)} \mathrm{E}[\mathbf{1}_{X_t=i}(N_{j,k}^X(s+h) - N_{j,k}^X(s))|X_s=k].
\end{aligned}$$

Moreover, observe that

$$N^X_{j,k}(s+h) - N^X_{j,k}(s) = \sum_{s\le u<s+h} \mathbf{1}_{\{}X_{u-} = j, X_u = k\} = f(X_{s+\xi}, \xi \ge 0).$$

For some cadlàg Borel function and for all $\xi \ge 0$.

If $X_s = k$ and $0 \le u < s$, then $X_{s+\xi}$ is independent of X_u. This means that $\mathbf{1}_{X_t=i}$ is independent given $X_s = k$. Hence,

$$\begin{aligned} g(s+h) - g(s) &= \sum_{l\in S} \frac{P(X_s = k)}{P(X_t = i)} \operatorname{E}[N^X_{j,k}(s+h) - N^X_{j,k}(s)\mathbf{1}_{X_t=i}|X_s = l] \\ &= \sum_{l\in S} \frac{P(X_s = k)}{P(X_t = i)} \operatorname{E}[N^X_{j,k}(s+h) - N^X_{j,k}(s)|X_s = k] \operatorname{E}[\mathbf{1}_{X_t=i}|X_s = l] \\ &= \sum_{l\in S} \operatorname{E}[N^X_{j,k}(s+h) - N^X_{j,k}(s)|X_s = l] p_{i,l}(s,t) \end{aligned}$$

Let, $Z(h) = \operatorname{E}[N^X_{j,k}(s+h) - N^X_{j,k}(s)|X_s = k] = o(h)$. Since X_s is Cadlàg,

$$\lim_{h\to 0, h\ge 0} \frac{Z(h)}{h} = \begin{cases} \mu_{l,k}(s), \text{if } k = j \\ 0, \text{else} \end{cases}$$

Hence, we can find the derivative for g.

$$\begin{aligned} g'(s) &= \lim_{h\to 0} \frac{g(s+h) - g(s)}{h} \\ &= \lim_{h\to 0} \frac{\sum_{l\in S} \operatorname{E}[N^X_{j,k}(s+h) - N^X_{j,k}(s)|X_s = l] p_{i,l}(s,t)}{h} \\ &= \lim_{h\to 0} \frac{\sum_{l\in S} Z(h) p_{i,l}(s,t)}{h} \\ &= p_{i,j}(t,s)\mu_{j,k}(s). \end{aligned}$$

By the fundamental theorem of calculus we have,

$$\begin{aligned} g(\beta) - g(\alpha) &= \int_\alpha^\beta g'(s) ds \\ &= \int_\alpha^\beta p_{i,j}(t,s)\mu_{j,k}(s) ds \\ &= \int_\alpha^\beta b(s)\alpha p_{i,j}(t,s)\mu_{j,k}(s) ds. \end{aligned}$$

On the other hand by definition of g we have,

$$\begin{aligned} g(\beta) - g(\alpha) &= \operatorname{E}[N^X_{j,k}(\beta) - N^X_{j,k}(\alpha)|X_t = i] \\ &= \operatorname{E}[\int_\alpha^\beta dN^X_{j,k}(s)|X_t = i] \\ &= \operatorname{E}[\int_\alpha^\beta \mathbf{1}_{(\alpha,\beta]}(s) dN^X_{j,k}(s)|X_t = i] \\ &= \operatorname{E}[\int_\alpha^\beta b(s) dN^X_{j,k}(s)|X_t = i]. \end{aligned}$$

This proves the statement for an indicator function. Since the set of simple functions is dense in $\mathcal{H}^2$ the result follows by the linearity of integrals and expectations. ■

Proof of Lemma 5.3.1

Proof. We are going to use the Chapman-Kolmogorov equations in (2.6.4) and the explicit expression in (5.2.3).

Use the following simplifying notation,

$$\begin{aligned}
A &:= B(s)p^x_{i,j}(t,s)\\
D &:= B(s)p^x_{i,j}(t,s)\sum_{k\in S, k\neq j}\mu^x_{j,k}(s)a_{j,k}(s)\\
C &:= \sum_{k\in S, k\neq j}\mu^x_{j,k}(s)a_{j,k}(s).
\end{aligned}$$

From (5.2.3) we have,

$$\begin{aligned}
W_i^+(t) &= \sum_{j\in S}\int_t^\infty B(s)p^x_{i,j}(t,s)da_j(s)\\
&+\sum_{j\in S}B(s)p^x_{i,j}(t,s)\sum_{k\in S,k\neq j}\mu^x_{j,k}(s)a_{j,k}(s)\\
&=\sum_{j\in S}\int_t^\infty Ada_j(s)+\sum_{j\in S}\int_t^\infty Dds\\
&=\sum_{j\in S}\int_t^u Ada_j(s)+\int_u^\infty Ada_j(s)\\
&+\sum_{j\in S}\int_t^u Dds+\int_u^\infty Dds\\
\text{(Chapman-Kolmogorov)} &=\sum_{j\in S}\int_t^\infty Ada_j(s)+\sum_{j\in S}\int_t^\infty Dds\\
&+\sum_{j\in S}B(s)\sum_{l\in S}p^x_{i,l}(t,u)p^x_{l,j}(u,s)da_j(s)\\
&+\sum_{j\in S}\int_u^\infty B(s)\sum_{l\in S}p^x_{i,l}(t,u)p^x_{l,j}(u,s)Cds\\
&=\sum_{j\in S}\int_t^u Ada_j(s)+\sum_{j\in S}\int_t^u Dds\\
&+\sum_{l\in S}p^x_{i,l}(t,u)\left[\sum_{j\in S}\int_u^\infty B(s)p^x_{l,j}(u,s)da_j(s)+\sum_{j\in S}\int_u^\infty B(s)p^x_{l,j}(u,s)Cds\right]\\
&=\sum_{j\in S}\int_t^u B(s)p^x_{i,j}(t,s)\\
&+\sum_{j\in S}\int_t^u B(s)p^x_{i,j}(t,s)\sum_{k\in S,k\neq j}\mu^x_{j,k}(s)a_{j,k}(s)ds
\end{aligned}$$

$$+ \sum_{l \in S} p^x_{i,j}(t,u) W^x_l(u).$$

■

APPENDIX B

Code

This appendix contains the code that has been used in thesis. Most of the code is in R or Python.

B.1 K2013

K2013 is the Financial Supervisory of Norway's model for mortalities specified in Finanstilsynet 2013.

```
        #Implementation of K2013
w <- function(x, man){
  if(man){
    return(min(c(2.671548-0.172480*x + 0.0014285*x^2,0)))
  }
  if(!man){
    return(min(c(1.287968-0.101090*x+0.000814*x^2,0)))
  }
}

mu_2013 <-function(x, man){
  if(man){
    return((0.241752+0.004536*10^(0.051*x))/1000)
  }
  if(!man){
    return((0.085411+0.003114*10^0.051*x)/1000)
  }
}

mu <- function(t,x, man){
  return(mu_2013(x,man)*(1+w(x,man)/100)^(t-2013))
}

#numerical integration
integral <- function(f,s,t,man,year, age, h=1/100){
  if(s==t){return(0)}
  else{
    N <- (t-s)/h
    bin <- seq(s,t*N)/N
    X <- f(year,age+bin,man)
    return(sum(X)/N)
  }

}
```

B.2 Example 5.4.1

This is the code used in Example 5.4.1. It calculates the value of a death benefit, plots this value in time and finds the premium.

```
        #Parameters
Tmax <- 120  #max age
age <- 25   #start age
Year <- 2022  #Current year

r <- 0.03   #rate
#Indicator function, saves some programming
less_than_70<- function(t){
  if(t<70){
    return(1)
  }
  else{
    return(0)
  }
}
B <- 100000


h <- 1/12   #Time steps
N <- (Tmax-age)/h   #Number of steps

V_T <- 0  #Terminal value
V_t <- rep(0,N) #Value of policy
V_p <- rep(0,N) #Value of premiums

#Loop to calculate
for(i in (N-1):1){
  V_t[i] <- V_t[i+1]-h*(r*V_t[i+1] - mu(year, age+i*h,T)*(B*less_than_70(age+i*h)-V_t[i+1]))
  V_p[i] <- V_p[i+1] - h*(r*V_p[i+1] +less_than_70(age+i*h) + mu(year, age+i*h,T)*V_p[i+1])
}

#One time premium
V_t[1]

#Yearly premium
-V_t[1]/V_p[1]

#plot
plot(((1:N)*h+age)[1:800],V_t[1:800], type ="l", xlab = "Age", ylab = "Value of policy", main = "Value
```

B.3 Example 8.1

The code is used to calculate the value of the policy.

```
from scipy import sparse
from scipy.sparse.linalg import splu

import numpy as np
import scipy as scp
import scipy.stats as ss
import sympy; sympy.init_printing()
from scipy import signal
from scipy.integrate import quad, quad_vec
import matplotlib.pyplot as plt
import plotly.graph_objects as go
```

```
#mortality model, K2013
def w(x, male):
    if male:
        return(np.minimum(2.671548-0.172480*x + 0.0014285**(x**2),0))
    else:
        return(np.minimum(1.287968-0.101090*x+0.000814**(x**2),0))

def mu_2013(x, male):
    if male:
        return((0.241752+0.004536*10**(0.051*x))/1000)
    else:
        return((0.085411+0.003114*10**0.051*x)/1000)

def mu(x):
    """
    The mortality model. This is the method where sex and starting age is determined
    """
    man = False
    t = 2022
    return(mu_2013(x,man)*(1+w(x,man)/100)**(t-2013))


#insurance functions
def f(x, G):
    """
    Policy function f
    """
    return(np.maximum(np.exp(x),G))
    #return(0)


def g(x, t):
    """
    Policy function g, sustained payments
    """
    #if (t>60):
    #    return(np.exp(x))
    #else:
    return(0)


def h(x, t):
    """
    Policy function h.
    """
    #B = 3
    #return(np.maximum(np.exp(x),B))
    return(0)

def survival_probability(s,t):
    """
    probability of surviving from age s to age t
    """
    return(np.exp(-quad_vec(mu, s,t)[0]))


# #Test mortality, Gompertz-Makeham law
# def mu(t):
#     a = -9.13275
#     b = 0.0809438
#     c = 0.000011018
```

```
#     return(np.exp(a+b*t-c*t**2))

# def survival_probability(s,t):
#     a = -9.13275
#     b = 0.0809438
#     c = 0.000011018
#     mu = b/(2*c)
#     sigma = np.sqrt(1/(2*c))
#     val = np.exp(-sigma*np.exp(a+(b*b)/(4*c))*np.sqrt(2*np.pi)*(scp.stats.norm.cdf((t-mu)/sigma,0,1)-
#     return(val)




#Setting up the parameters of the model
#financial parameters
r = 0.023
sig = 0.19
S0 = 1
X0 = np.log(S0)

lam = 0  # lambda
muJ = 0.025     # (or alpha) is the mean of the jump size
sigJ = 0.25  # (or xi) is the standard deviation of the jump size

#insurance parameters
G = np.exp(3)
B = 3
Texpir = 10
start_age = 55



S_max = 4*float(G)
S_min = float(G)/4
x_max = np.log(S_max)  # A2
x_min = np.log(S_min)  # A1

Nspace = 20   # M space steps
Ntime = Texpir*12   # N time steps

#discretization
dev_X = np.sqrt(lam * sigJ**2 + lam * muJ**2)   # std dev of the jump component
dx = (x_max - x_min)/(Nspace-1)
extraP = int(np.floor(3*dev_X/dx))  # extra points
x = np.linspace(x_min-extraP*dx, x_max+extraP*dx, Nspace + 2*extraP)   # space discretization
T, dt = np.linspace(0, Texpir, Ntime, retstep=True)       # time discretization

#Boundry conditions
Payoff = f(x,G)            # Call payoff
V = np.zeros((Nspace + 2*extraP, Ntime))       # grid initialization
offset = np.zeros(Nspace-2)        # vector to be used for the boundary terms

# terminal conditions
V[:,-1] = Payoff

test_death_benefit = lambda t: 0 if(t>60) else 1
for i in range(len(T)):
    V[:extraP+1,i] = G*survival_probability(start_age+T[i], start_age+Texpir) #lower consdition
    V[-extraP-1:,:] = np.exp(x[-extraP-1:]).reshape(extraP+1,1) * np.ones((extraP+1,Ntime))*survival_pr
#upper condition
```

```
    #Test Death benefit
    #V[:extraP+1,i] = test_death_benefit(start_age+T[i])*G*survival_probability(start_age,start_age+T[i])
    #V[-extraP-1:,:] = test_death_benefit(start_age+T[i])*np.exp(x[-extraP-1:]).reshape(extraP+1,1) * np.ones((extr



#setting up the jump levy measure
cdf = ss.norm.cdf([np.linspace(-(extraP+1+0.5)*dx, (extraP+1+0.5)*dx, 2*(extraP+2) )], loc=muJ, scale=sigJ)[0]
nu = lam * (cdf[1:] - cdf[:-1])

#finding lambda hat and m hat
lam_appr = sum(nu)    # sum of the components of nu

m = lam * (np.exp(muJ + (sigJ**2)/2) -1)     # coefficient m
m_int = quad(lambda z: lam * (np.exp(z)-1) * ss.norm.pdf(z,muJ,sigJ), -(extraP+1.5)*dx, (extraP+1.5)*dx )[0]
m_appr = np.array([ np.exp(i*dx)-1 for i in range(-(extraP+1), extraP+2)]) @ nu

#constructing the diffusion matrix D
sig2 = sig*sig
dxx = dx * dx
def get_a(t, dt, dx, sig2, m_appr, lam_appr, mu):
    dxx = dx * dx
    return( (dt/2) * ( (r-m_appr-0.5*sig2)/dx - sig2/dxx ) )
def get_b(t, dt, dx, sig2, m_appr, lam_appr, mu):
    dxx = dx * dx
    return( 1 + dt * ( sig2/dxx + r +mu(t)+ lam_appr) )
def get_c(t, dt, dx, sig2, m_appr, lam_appr, mu):
    dxx = dx * dx
    return(-(dt/2) * ( (r-m_appr-0.5*sig2)/dx + sig2/dxx ) )


#Constructing the Jump matrix J
J = np.zeros((Nspace-2, Nspace + 2*extraP))
for i in range(Nspace-2):
    J[i, i:(len(nu)+i)] = nu



# Backward iteration
for i in range(Ntime-2,-1,-1):
    a = get_a(start_age+T[i],dt, dx, sig2, m_appr, lam_appr, mu)
    b = get_b(start_age+T[i],dt, dx, sig2, m_appr, lam_appr, mu)
    c = get_c(start_age+T[i],dt, dx, sig2, m_appr, lam_appr, mu)
    D = sparse.diags([a, b, c], [-1, 0, 1], shape=(Nspace-2, Nspace-2)).tocsc()
    DD = splu(D)
    offset[0] = get_a(start_age+T[i],dt, dx, sig2, m_appr, lam_appr, mu) * V[extraP,i]
    offset[-1] = get_c(start_age+T[i],dt, dx, sig2, m_appr, lam_appr, mu) * V[-1-extraP,i]
    V_jump = V[extraP+1 : -extraP-1, i+1] + dt * (J @ V[:,i+1]) -dt*(g(x[extraP+1 : -extraP-1],start_age + T[i])-mu
    V[extraP+1 : -extraP-1, i] = DD.solve( V_jump - offset )
    #print(i)

#plotting
#fig = go.Figure(data=[go.Surface(z=V[extraP:-extraP,:])])
fig = go.Figure(data=[go.Surface(y = x,z=np.log(V))])
fig.update_layout(scene = dict(
    xaxis_title = "time (month)",
    yaxis_title = "fund (million)",
    zaxis_title = "value (million)"
    ))
fig.show()
#print(x)
#print(survival_probability(30,40)
```

Monte Carlo simulation for pure endowment with jump diffusions

```
#Simulating unit linked polcies
start.time <- Sys.time()

set.seed(100)
#K2013
library(plotly)


#Implementation of K2013
w <- function(x, man){
  if(man){
    return(min(c(2.671548-0.172480*x + 0.0014285*x^2,0)))
  }
  if(!man){
    return(min(c(1.287968-0.101090*x+0.000814*x^2,0)))
  }
}

mu_2013 <-function(x, man){
  if(man){
    return((0.241752+0.004536*10^(0.051*x))/1000)
  }
  if(!man){
    return((0.085411+0.003114*10^0.051*x)/1000)
  }
}

mu <- function(t,x, man){
  return(mu_2013(x,man)*(1+w(x,man)/100)^(t-2013))
}

#numerical integration
integral <- function(f,s,t,man,year, age, h=1/100){
  if(s==t){return(0)}
  else{
    N <- (t-s)/h
    bin <- seq(s,t*N)/N
    X <- f(year,age+bin,man)
    return(sum(X)/N)
  }

}


p_ss <- function(x,t, year, man){
  return(exp(-integral(mu,0,t,man,year,x)))
}


#draw random life
drawlife <- function(start_age, n){
  findzero <- function(z){
    uniroot((function(y) 1-p_ss(start_age, y, 2022, T)-z), c(0, 120))$root
  }
  u <- runif(n)
  life <- u
  for(i in 1:n){
    print(i)
    life[i]<- floor(findzero(u[i]))
  }
```

```
  return(start_age+life)

}



life.gen <- function(age,num){
  life.inv <- function(y, lower=0, upper=120)
    uniroot( (function(x) 1-surv_prob(age,age+x)-y), lower=lower, upper=upper )$root

  u <- runif(num,0,1)
  lives <- u
  for(i in 1:num){ lives[i] <- floor(life.inv(u[i])) }
  return(age+lives)
}



#parameters
n <- 10000
Tmax <- 10

h <- 1/120

r <- 0.023
sig <- 0.19
lambda <- 10
muJ <- 0.025
sigJ <- 0.25

S0 <- 1





#Simulatng sample paths of jump diffusion
paths <- matrix(rep(0, n*(Tmax/h)), nrow = n)
for(i in 1:n){
  print(i)
  path <- rep(S0,Tmax/h)
  bm <- rnorm(Tmax/h, r*h, sig*sqrt(h))
  jump_numbers <- rpois(Tmax/h, lambda*h)
  sum(jump_numbers)
  jump <- rep(0,Tmax/h)

  for(j in 1:(Tmax/h)){
    #I <- 0
    #for(k in 1:jump_numbers[j]){
    #  I = I + rnorm(1, muJ*h,sigJ*sqrt(h))
    #}
    #path[j] <-bm[j] + I
    jump[j] <- sum(rnorm(jump_numbers[j], muJ,sigJ))
  }
  geo_bm <- (r-0.5*sig^2 -lambda*(muJ+0.5*sigJ^2))*h + sig*sqrt(h)*rnorm(Tmax/h)

  paths[i,] <- S0*exp(cumsum(geo_bm)+cumsum(jump))
}
```

```
#plot(paths[1,], pch = 1,type = "l", main = "Merton jump diffusion, S0 = 2", xlab = "Months", ylab = "V
#lines(paths[2,], pch = 2,  col = "red")
#lines(paths[3,], pch = 3,  col = "blue")
#lines(paths[4,], pch = 4,  col = "green")



#call option
G = 3
#mean(pmax(paths[,Tmax/h],G))*exp(-r*Tmax)

life_generated <- drawlife(55,n)
payoff <- rep(0,n)
for(i in 1:n){
  if(life_generated[i]>=65){
    payoff[i]<-pmax(paths[i,Tmax/h],G)
  }

}
mean(payoff)*exp(-r*Tmax)
end.time <- Sys.time()
time.taken <- end.time - start.time
time.taken


#Runtime: 52 min
#Results: 3.1287.
```

FPI model calibration

```
#importing data
library(readxl)
library(BaPreStoPro)
Salmon_price <- read_excel("C:/Users/vegar/Desktop/Masters thesis/Code/Salmon_price.xls",
                           skip = 1)
View(Salmon_price)
plot(seq(2006,2022,by = 1/52), Salmon_price$‘NOK/kg‘, type = "l", main = "FPI", ylab = "Kr/KG", xlab =


#Estimating Merton jump diffusion
fit <- set.to.class("Merton", parameter = list(thetaT = 0.1, phi = 0.05, gamma2 = 0.1, xi = 0.5))
t <- seq(1,length(Salmon_price$Week))
dat <- Salmon_price$‘NOK/kg‘
est <- estimate(fit, t, dat, 1000)
plot(est@thetaT, type = "l")
plot(est@phi, type = "l")
plot(est@gamma2, type = "l")
plot(est@xi, type = "l")


plot(log(dat[2:length(dat)]/dat[1:(length(dat)-1)]), type = "l")
boxplot(log(dat[2:length(dat)]/dat[1:(length(dat)-1)]))
hist(log(dat[2:length(dat)]/dat[1:(length(dat)-1)]),100)
```

Bibliography

Aase, K. K. and Persson, S.-A. (1993). "Pricing of unit-linked life insurance policies". eng. In: Working paper (Norges handelshøyskole. Institutt for foretaksøkonomi : trykt utg.) Vol. 2,1993.

Applebaum, D. (2009). *Lévy processes and stochastic calculus*. eng. Second edition. Vol. 116. Cambridge studies in advanced mathematics ; Cambridge: Cambridge University Press.

Baldi, P. (2017). *Stochastic Calculus: An Introduction Through Theory and Exercises*. eng. Universitext. Cham: Springer International Publishing.

Bédos, E. (2019). *Notes on Elementary Linear Analysis*.

Benth, F. E., Benth, J. S., and Koekebakker, S. (2008). *Stochastic modelling of electricity and related markets*. Vol. 11. World Scientific.

Berk, J. and DeMarzo, P. M. (2017). *Corporate finance*. eng. Harlow.

Bingham, N. H. and Kiesel, R. (1998). *Risk-neutral valuation: pricing and hedging of financial derivatives*. eng. Springer finance. Springer.

Cantarutti, N. (Mar. 2022). *Financial-Models-Numerical-Methods*. original-date: 2019-09-10T17:35:01Z. URL: https://github.com/cantaro86/Financial-Models-Numerical-Methods/blob/7a7f2b2ff3510ae686d1e5432afbb70bc7c77024/3.1%20Merton%20jump-diffusion%2C%20PIDE%20method.ipynb (visited on 03/29/2022).

Cheang, G. H. and Chiarella, C. (2012). "A modern view on Merton's jump-diffusion model". In: *Stochastic Processes, Finance and Control: A Festschrift in Honor of Robert J Elliott*. World Scientific, pp. 217–234.

Cleese, J. (2013). *John Cleese thanks everyone on the planet for his award*. Youtube. URL: https://www.youtube.com/watch?v=7_JUBgPHYmY.

Cont, R. and Voltchkova, E. (Sept. 2003). "A Finite Difference Scheme for Option Pricing in Jump-Diffusion and Exponential Levy Models". en. In: no. ID 458200.

Delbaen, F. (2006). *The Mathematics of Arbitrage*. eng. 1st ed. 2006. Springer Finance. Berlin, Heidelberg and Schachermayer, Walter: Springer Berlin Heidelberg : Imprint: Springer.

Doléans-Dade, C. (1970). "Quelques applications de la formule de changement de variables pour les semimartingales". fre. In: *Zeitschrift f r Wahrscheinlichkeitstheorie und Verwandte Gebiete* vol. 16, no. 3, pp. 181–194.

Eberlein, E. and Kallsen, J. (2019). *Mathematical Finance*. eng. Cham.

Filipović, D. (2009). *Term-structure models : a graduate course*. eng. Berlin.

Finanstilsynet (2013). *Nytt beregningsgrunnlag for kollektiv pensjonsforsikring.* www.finansnorge.no. URL: https://www.finansnorge.no/aktuelt/nyheter/2013/06/k2013-minimum-beregningsgrunnlag-for-premieberegninger/ (visited on 02/14/2022).

Fish Pool Index (2022). Fish Pool. URL: https://fishpool.eu/fish-pool-index/ (visited on 05/04/2022).

Friedman, A. (1975). *Stochastic differential equations and applications : 1 and 2*. eng. New York.

Goetzmann, W. N. (2017). *Money Changes Everything : How Finance Made Civilization Possible*. eng. Princeton, NJ: Princeton University Press.

Hammurabi, K. o. B. (1904). *Code of Hammurabi*. eng. Translation into english. Chicago.

Harrison, J. M. and Pliska, S. R. (1983). "A stochastic calculus model of continuous trading: Complete markets". In: *Stochastic Processes and their Applications* vol. 15, no. 3, pp. 313–316.

Kac, M. (1985). *Enigmas of chance : an autobiography*. eng. Alfred P. Sloan Foundation series. New York: Harper & Row.

Karatzas, I. and Shreve, S. (2014). *Brownian Motion and Stochastic Calculus*. eng. Springer study edition. New York, NY: Springer.

Koller, M. (2012). *Stochastic models in life insurance*. eng. Berlin.

Kuttler, K. (2017). *Linear Algebra, Theory And Applications*.

Lindstrøm, T. L. (2017). *Spaces: An Introduction to Real Analysis*. Vol. 29. American Mathematical Soc.

Mandelbrot, B. B., Fisher, A. J., and Calvet, L. E. (1997). "A multifractal model of asset returns". In.

Merton, R. C. (1976). "Option pricing when underlying stock returns are discontinuous". eng. In: *Journal of financial economics*. Journal of Financial Economics vol. 3, no. 1, pp. 125–144.

Meyer-Brandis, T. and Tankov, P. (2008). "Multi-factor jump-diffusion models of electricity prices". In: *International Journal of Theoretical and Applied Finance* vol. 11, no. 05, pp. 503–528.

Mørken, K. (2015). *Numerical algorithms and digital representation*. eng. Oslo.

Øksendal, B. (2007). *Stochastic differential equations : an introduction with applications*. eng. 6th ed. Universitext. Berlin: Springer.

Oliveira, V. H. et al. (2021). "Factors associated with baseline mortality in Norwegian Atlantic salmon farming". In: *Scientific Reports* vol. 11, no. 1, pp. 1–14.

Protter, P. (2005). *Stochastic Integration and Differential Equations*. eng. Berlin, Heidelberg.

Raimi, R. A. (Nov. 1984). *The world should have known him better*.

Sande, Å. H. (2021). *Prospective Reserves of Life Insurance Policies under the Heath-Jarrow-Morton Framework*. eng.

Tang, F. (2018). "Merton Jump-Diffusion Modeling of Stock Price Data". Bachelor's Thesis. Linnaeus University.

Tankov, P. (2003). *Financial Modelling with Jump Processes*. eng. Chapman & Hall/CRC financial mathematics series. London: Chapman and Hall/CRC.

Tao, T. (2011). *An introduction to measure theory*. eng. Vol. vol. 126. Graduate studies in mathematics. Providence, R.I: American Mathematical Society.

Varadhan, S. R. S. (2001). *Probability theory*. eng. New York.

Walsh, J. B. (2012). *Knowing the odds : an introduction to probability.* eng. Vol. vol. 139. Graduate studies in mathematics. Providence, R.I: American Mathematical Society.