DEPARTMENT OF ECONOMICS AND FINANCE

SUBJECT: MATHEMATICAL METHODS FOR ECONOMICS AND FINANCE

HEDGING STRATEGIES IN JUMP DIFFUSION PRICING MODELS

SUPERVISOR

Prof. *Fausto Gozzi*

CANDIDATE

Marco Gregnanin

Id. Number: 689941

CO-SUPERVISOR

Prof.ssa *Sara Biagini*

ACADEMIC YEAR 2018/2019

Contents

Introduction

The mathematical modeling of financial market start with Louis Bachelier, who was the first to introduce the Brownian motion as a model for the price fluctuation of a liquid traded financial asset with his doctoral thesis in 1900. In 1973 Fisher Black and Myron Scholes given a great contribution with the article "The Pricing of Option and Corporate Liabilities", which gave a new dimension to the use of probability theory in finance. The option pricing methodology introduced by Black and Scholes is unique in that distributional assumptions alone suffice to generate well-specified option pricing formulas involving mostly observable variable and parameters. One assumption is that the price of the underlying asset follows a diffusion process and an additional assumption is that the instantaneous risk-free rate is nonstochastic and constant. Under these plus other "frictionless market" assumptions, the option's payoff can be replicated by a continuously adjusted hedge portfolio composed of the underlying asset and short-term bonds. This imply that the key assumption in the Black-Scholes model is that the market is complete. In a complete market models probability does not really matter, in fact the objective evolution of the asset is only there to define the set of impossible events and serves to specify the class of equivalent measures. Hence, two statistical models with equivalent measures lead to the same option prices in a complete market setting. Therefore, the option pricing formula generated by Black and Scholes depends critically upon the distributional restriction on the volatility of the underlying asset. The result of that restriction is that the systematic risk of the option is a function of the systematic risk of the underlying asset only.
Jump diffusion process and more in general Lévy models generalize the Black and Scholes work by allowing the stock price to jump while preserving the independence and stationary of returns. Hence, the jump diffusion process described the observed reality of financial markets in a more accurate way than models based only on Brownian motion. In the real world, we observe that the asset price processes have jumps or spikes. Therefore, we can find three main reason for introducing jumps in financial modeling. First, asset price processes have jumps and some risks cannot be handled with a continuous path model but we need to study a discontinuous models. Second, the presence in the option market of the phenomenon of implied volatility smile which shows that the risk-neutral returns are non-gaussian and leptokurtic. Moreover, in continuous path models the law of returns for shorter maturities becomes closer to the Gaussian distribution, on the other hand in models with jumps returns actually become less Gaussian as the maturity becomes shorter. Finally, the jump process correspond to incomplete markets, hence we did not find a unique equivalent probability measure for the option pricing but there are many possible choice. This imply that a perfect hedge, i.e. the Black and Scholes Delta hedging, is not longer possible in jump models and the hedging in jump process achieves a trade-off between the risk due to the diffusion part and the jump risk.

This thesis is structured as follows. The first chapter contains a brief review of the main concept of probability theory and the last section gives the definition of stochastic process and, in particular, we define and explain the most well-known continuous stochastic process: the Brownian motion.
The second chapter is dedicated to the theoretical treatment of the jump diffusion process. We start the chapter introducing the Poisson process, which is the main building block of discontinuity process. Then, we talk in general about the Lévy process and we study the

main features about their distributions. And we conclude the chapter talking about the jump diffusion model and we give some example about it.
The third chapter describes the stochastic integral and the main tool to explain the time evolution of a derivative instrument. The first section is dedicated to the concept of stochastic integral and then we see how change the stochastic integral when is driven by a jump diffusion process.
The last chapter is focused on hedging strategy. We start by talking about the measure transformation, which is a key tool to find equivalent probability measure in option pricing. The second section is dedicated to the option pricing in jump diffusion model and here we see how the option pricing is different between the Black and Scholes model and a jump process. Finally, we talk about the hedging strategy. We start by describing the Delta hedging in the Black-Scholes model and then we start to talk about the hedging in discontinuous-path process. First, we introduce the Merton approach proposed in 1976, then we described the more general concept of hedging in the Lévy process: the Quadratic Hedging. We conclude the fourth chapter with a comparison between the Merton model and the Black and Scholes model.

Chapter 1

Probability Theory

This first chapter presents an introduction to probability theory and stochastic process. We start with the definition of probability space and measure, which is important in the study of stochastic process in general and, in particular, in the jump process. Then, we describe the probability law and how it converges. Moreover, we introduce the characteristic function, which is central tool in the second chapter, and we state some properties, then we give an example of characteristic function for a normal distribution.
The second part of the chapter introduce the concept of stochastic process. In particular, we talk about the construction of stochastic process and about the stopping time. Then, we define the most well-known continuous stochastic process: the Brownian motion, which is the core concept of the Black-Scholes model and it is one of the component of the jump diffusion process, which is described in the second chapter. Finally, we define the martingale and the property associate to it. We start to define the martingale in discrete time and, then, we extend the result for the continuous time case.

1.1 Probability Concept

1.1.1 Random Variable

Definition 1.1 [Probability Space] *A probability space is a triple $(\Omega, \mathcal{F}, \mathbb{P})$ where: Ω is the sample space corresponding to outcomes of some experiment; $\mathcal{F}$ is a σ-algebra of subset of Ω and $\mathbb{P}$ is a probability measure on $\mathcal{F}$ such that $\mathbb{P}(\Omega) = 1$.*

Moreover, $\mathbb{P}$ is a function with domain $\mathcal{F}$ and range $[0, 1]$ such that $\mathbb{P}(A) \geq 0$ for all $A \in \mathcal{F}$ and if $\{A_n, n \geq 1\}$ are events in $\mathcal{F}$ that are disjoint, then $\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mathbb{P}(A_n)$. In this case, we say that $\mathbb{P}$ is σ-additive. Finally, we say that an event A with probability $\mathbb{P}(A) = 1$ occurs almost surely and, on the other hand, if $\mathbb{P}(A) = 0$ the event A is impossible.
In the definition of Probability space, we saw a σ-algebra, the following two definition helps us to understand this concept.
Definition 1.2 [Algebra] (definition 1.5.2 in [12]) *A algebra or a field is a non-empty class of subsets of Ω closed under finite union, finite intersection and complements.*
Hence, a minimal set of postulates for $\mathcal{A}$ to be a field is:

1. $\Omega \in \mathcal{A}$;
2. $A \in \mathcal{A}$ implies $A^c \in \mathcal{A}$;
3. $A, B \in \mathcal{A}$ implies $A \cup B \in \mathcal{A}$.

Definition 1.3 [σ-Algebra] (definition 1.5.3 in [12]) *A σ-algebra or a σ-field is a non-empty class of subsets of Ω closed under countable union, countable intersection and complements.*
Hence, a minimal set of postulates for $\mathcal{B}$ to be a σ-field is:

1. $\Omega \in \mathcal{B}$;

2. $B \in \mathcal{B}$ implies $B^c \in \mathcal{B}$;

3. $B_i \in \mathcal{B}$, $i \geq 1$ implies $\bigcup_{i=1}^{\infty} B_i \in \mathcal{B}$ and $\bigcap_{i=1}^{\infty} B_i \in \mathcal{B}$.

Definition 1.4 [Measurable space] *A measurable space is a couple $(E, \mathcal{E})$ where: E is a space and $\mathcal{E}$ is a σ-algebra of subset of E.*
Definition 1.5 [Measure] (definition 2.1 in [2]) *A (positive) measure on $(E, \mathcal{E})$ is defined as a function*

$$\begin{aligned} \mu : &\mathcal{E} \to [0, \infty) \\ &A \mapsto \mu(A) \end{aligned}$$

such that:

- *$\mu(\emptyset) = 0$;*
- *For any sequence of disjoint sets $A_n \in \mathcal{E}$: $\mu\left(\bigcup_{n\geq 1} A_n\right) = \sum_{n\geq 1} \mu(A_n)$;*

An element $A \in \mathcal{E}$ is called a measurable set and $\mu(A)$ its measure.

Suppose $\Omega = \mathbb{R}$ and let $\mathcal{C} = \{(a, b], .\infty \leq a \leq b < \infty\}$. Then, we can define the Borel subset of $\mathbb{R}$, denoted by $\mathcal{B}(\mathbb{R})$, as:

$$\mathcal{B}(\mathbb{R}) := \sigma(\mathcal{C})$$

Thus the Borel subset of $\mathbb{R}$ are elements of the σ-field generated by intervals that are open on the left and closed on the right.
Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $(E, \mathcal{E})$ be a measurable space. Then, a random variable is a function $X : \Omega \to E$ measurable such that $X^{-1}(A) \in \mathcal{F}$ every time $A \in \mathcal{E}$. A special case occurs when $E = \mathbb{R}$ and $\mathcal{E} = \mathcal{B}(\mathbb{R})$, in this case X is a real random variable. Therefore, for a real random variable X on $(\Omega, \mathcal{F}, \mathbb{P})$, we can always define its integral.
Many integrations results are proved by first showing they hold for simple functions and then extending the result to more general functions. Recall that a function on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$: $X : \Omega \mapsto \mathbb{R}$ is simple if it has a finite range. Hence, we can assume that a simple function is measurable and this kind of function can always be written in the forms:

$$X(\omega) = \sum_{i=1}^{k} a_i \mathbf{1}_{A_i}(\omega)$$

where $a_i \in \mathbb{R}$ and $A_i \in \mathcal{F}$ and $A_1, \ldots, A_k$ are disjoint and $\sum_{i=1}^{k} A_i = \Omega$.
Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and

$$X : (\Omega, \mathcal{F}) \mapsto \tilde{\mathbb{R}}, \mathcal{B}(\tilde{\mathbb{R}})$$

where $\tilde{\mathbb{R}} = (-\infty, \infty)$. Then, we can define the expectation of X, denoted by $E(X)$, as:

$$E[X] = \int_{\Omega} X d\mathbb{P} = \int_{\Omega} X(\omega)\mathbb{P}(d\omega) \tag{1.1}$$

This integral is also known as the Lebesgue-Stieltjes integral of X with respect to $\mathbb{P}$. For example, if X is a simple random variable equal to: $X = \sum_{i=1}^{n} a_i \mathbf{1}_{A_i}$, where $|a_i| < \infty$ and $\sum_{i=1}^{k} A_i = \Omega$. The expectation of X can be defined as:

$$E[X] = \int X d\mathbb{P} =: \sum_{i=1}^{k} a_i \mathbb{P}(A_i)$$

We can note that for a simple function the expectation is computed by taking a possible value, multiplying by the probability of the possible value and then summing over all possible values. Moreover, we can see that this example define the expected value for a discrete random variables.

On the other hand, we say that X is a continuous random variable if there exists a nonnegative function f, defined for all real $x \in (-\infty, \infty)$, having the property that, for any set B of real numbers:

$$\mathbb{P}(X \in B) = \int_B f(x)dx \tag{1.2}$$

where $f(x)$ is called the probability density functions of the random variable X. Therefore, we can define the expected value of X as:

$$E[X] = \int_{-\infty}^{\infty} x f(x)dx$$

We can define the L^p spaces, with $1 \leq p \leq +\infty$, as the norm of X. More precisely, we have that $\|X\|_p = E[|X|^p]^{\frac{1}{p}}$. We say that two random variable X and Y are equivalent if $\mathbb{P}(X = Y) = 1$. We can note that if X and Y are equivalent, then we have $\|X\|_p = \|Y\|_p$ and that L^p is the set of equivalence classes of random variable X such that $\|X\|_p < +\infty$. In particular, L^p is the set of equivalence classes and it is not the set of random variable. Let X be a real random variable in L^2. Then, we can define the variance of X, denoted by $Var(X)$, as the quantity:

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

If $\alpha > 0$, the quantity $\int |x|^\alpha \mu(dx)$ is called the absolute moment of order α of μ. Moreover, if α is a positive integer, the quantity $\int x^\alpha \mu(dx)$ is called the moment of order α of μ.
Now, consider two random variable X and Y in L^2. The covariance of X and Y can be defined as:

$$Cov(X, Y) = E[(X - E[X])(Y - E[Y]) = E[XY] - E[X]E[Y]$$

We can note that if $X = Y$, then $Cov(X, Y) = Var(X)$ and that if $X \perp Y$, then $E[XY] = E[X]E[Y]$ and consequently $Cov(X, Y) = 0$. If $Cov(X, Y) = 0$, we say that X and Y are not correlated.

We need to introduce the following inequalities because they are useful in the following chapter to study the behavior of the stochastic process.

- *Jensen Inequality:* Let X be a random variable in $\mathbb{R}^m$ and let $\Phi : \mathbb{R}^m \to \mathbb{R} \cup \{+\infty\}$ be a convex function. Also suppose that X and $\Phi(X)$ are integrable. Then:

  $$E[\Phi(X)] \geq \Phi(E[X])$$

  Proof
  Suppose that Φ is twice differentiable. We know that its second derivative is always positive. Therefore, by Taylor expansion, for any x, we have:

  $$\Phi(x) \geq \Phi(E[X]) + (x - E[X])\Phi'(E[X])$$

  Putting $x = X(\omega)$ and taking expectations, we will finish the proof.

  $\square$

- *Markov Inequality:* Let $X \in L^1$ and for any $\delta > 0$, $\beta > 0$, we have that:

  $$\mathbb{P}(|X| \geq \delta) \leq \frac{E[|X|^\beta]}{\delta^\beta}$$

  Proof
  Let β is equal to 1. Then, we have:

  $$1 \cdot \mathbf{1}_{[\frac{|X|}{\delta} \geq 1]} \leq \frac{|X|}{\delta} \cdot \mathbf{1}_{[\frac{|X|}{\delta} \geq 1]} \leq \frac{|X|}{\delta}$$

□

- *Chebyshev Inequality:* Let $X \in L^2$. Then, for any $\alpha > 0$ we have:

$$\mathbb{P}(|X - E[X]| \geq \alpha) \leq \frac{Var(X)}{\alpha^2}$$

Proof

$$\begin{aligned}\mathbb{P}(|X - E[X]| \geq \alpha) &= \mathbb{P}(|X - E[X]|^2 \geq \alpha^2) \\ &\leq \frac{E[X - E[X]]^2}{\alpha^2} = \frac{Var(X)}{\alpha^2}\end{aligned}$$

where the last inequality is an application of the Markov inequality.

□

The Chebyshev inequality imply that the variance of one random variable is a quantity which is as much larger as bigger is the value of X from his mean $E[X]$.

- *Hölder Inequality:* Let Z and W be real random variable ≥ 0 and α, β real numbers ≥ 0 such that $\alpha + \beta = 1$. Then, we have:

$$E[Z^\alpha W^\beta] \leq E[Z]^\alpha E[W]^\beta$$

From the inequality above, we find that if X, Y are real random variable and p, q are numbers ≥ 0 such that $\frac{1}{q} + \frac{1}{p} = 1$, then we have the Hölder inequality:

$$E[XY] \leq E[|X|^p]^{\frac{1}{p}} E[|X|^q]^{\frac{1}{q}} = \|X\|_p \cdot \|Y\|_q$$

Proof
Let $\omega \in \Omega$ and let $x = \frac{|X(\omega)|}{\|X\|_p}$ and $y = \frac{|Y(\omega)|}{\|Y\|_q}$. Then, we consider the following inequality: $xy \leq \frac{x^p}{p} + \frac{y^q}{q}$ for x, $y > 0$ and we have:

$$\begin{aligned}xy &\leq \frac{1}{p}\left(\frac{|X|}{\|X\|_p}\right)^p + \frac{1}{q}\left(\frac{|Y|}{\|Y\|_q}\right)^q \\ E[XY] &\leq \frac{1}{p}E\left[\frac{|X|^p}{\|X\|_p^p}\right] + \frac{1}{q}E\left[\frac{|Y|^q}{\|Y\|_q^q}\right] = \frac{1}{p} + \frac{1}{q} = 1\end{aligned}$$

□

- *Minkowski Inequality:* Let $p \geq 1$ and let X, Y are real random variable such that $E[|X|^p] < \infty$ and $E[|Y|^p] < \infty$. Then:

$$\|X + Y\|_p \leq \|X\|_p + \|Y\|_p$$

Proof
If $\|X + Y\|_p = 0$ there is nothing to prove. Therefore, we prove that $\|X + Y\|_p > 0$. We can note that $(p-1)q = p$ and $\frac{p}{q} = p^{-1}$. Hence, we obtain by triangular and Hölder inequality that:

$$\begin{aligned}\|X + Y\|_p^p &= E\left[|X + Y|^{p-1}|X + Y|\right] \\ &\leq E\left[|X + Y|^{p-1}|X|\right] + E\left[|X + Y|^{p-1}|Y|\right] \\ &\leq \left\||X + Y|^{p-1}\right\|_q \|X\|_p + \left\||X + Y|^{p-1}\right\|_q \|Y\|_q \\ &= \left(E[|X + Y|^{(p-1)q}]\right)^{\frac{1}{q}} (\|X\|_p + \|Y\|_q) \\ &= \|X + Y\|_p^{p/q} (\|X\|_p + \|Y\|_q)\end{aligned}$$

Then, dividing both members by $\|X + Y\|_p^{p/q}$ we will finish the proof.

□

The Minkowski inequality imply that L^p, for $p \geq 1$, is a vector space. Moreover, if $p > q$ from the Jensen inequality with the convex function $\Phi(x) = |x|^{p/q}$, we will find:

$$\|X\|_p^p = E[|X|^p] = E[\Phi(|X|^q)] \geq \Phi(E[|X|^q]) = E[|X|^q]^{p/q}$$

and applying the p-root, we find

$$\|X\|_p \geq \|X\|_q$$

Therefore, we can conclude that if $p \geq q$ then $L^p \subset L^q$.

Let X be a random variable which takes value on the measurable space $(E, \mathcal{E})$. Then, it is easy to see that the function Q define on $\mathcal{E}$ by

$$Q(A) = \mathbb{P}(X^{-1}(A)) = \mathbb{P}([\omega : X(\omega) \in A])$$

is a probability measure, Q is also call the law of X. In words this means that we define the probability that the random variable X falls into a Borel set as the probability (on $(\Omega, \mathcal{F})$) of the inverse image of this Borel set. The following proposition is useful to compute the integral with respect to the inverse image.

Proposition 1.1 *Let $X : (\Omega, \mathcal{F}, \mathbb{P}) \to (E, \mathcal{E})$ be a random variable, $f : (E, \mathcal{E}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ be a measurable function and Q be the law of X. Then, f is Q-integrable if and only if $f \circ X$ is P-integrable. Therefore, we have*

$$\int_E f(x)Q(dx) = \int_\Omega f \circ X(\omega)P(d\omega) \tag{1.3}$$

In particular, if X is a real random variable and μ is its law, the following relation hold:

$$E[X] = \int x\mu(dx), \qquad E[|X|^\alpha] = \int |x|^\alpha \mu(dx)$$

Therefore, $X \in L^p$ if and only if its law has absolute moment of order p finite .

Let X be a real random variable. Then, it can be decomposed in a positive and in a negative part: $X = X^+ - X^-$. We call X quasi-integrable if at least one of $E[X^+]$ and $E[X^-]$ is finite. In this case, we can define the expectation of X as:

$$E[X] := E[X^+] - E[X^-]$$

If $E[X^+]$ and $E[X^-]$ are both finite, we call X integrable. If $E[X^+] < \infty$ but $E[X^-] = \infty$, then $E[X] = -\infty$. If $E[X^+] = \infty$ but $E[X^-] < \infty$, then $E[X] = \infty$. If $E[X^+] = \infty$ but $E[X^-] = \infty$, then $E[X]$ does not exist.

1.1.2 Conditional Probability and Independence

If $\mathcal{E}$ and $\mathcal{E}'$ are σ-algebra of events of E, we can denote with $\mathcal{E} \vee \mathcal{E}'$ the smallest σ-algebra, which contains $\mathcal{E}$ and $\mathcal{E}'$.

Let X be a random variable such that $(\Omega, \mathcal{F}, \mathbb{P}) \to (E, \mathcal{E})$. We denote with $\sigma(X)$ the σ-algebra generated by X, hence it is the smallest sub-σ-algebra of $\mathcal{F}$ such that X is still measurable. The following lemma describes the situation when one random variable is a function of another one by the inclusion of the σ-field generated by the random variable.

Lemma 1.1 (Doob's measurable criterion) *Let X be a random variable such that $(\Omega, \mathcal{F}, \mathbb{P}) \to (E, \mathcal{E})$. Then, all the real random variables $\sigma(X)$-measurable has the form $f(X)$, where f is a measurable operation from $(E, \mathcal{E})$ to $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.*

Consider m random variable, $X_1, \dots, X_m$, which take values respectively in $(E_1, \mathcal{E}_1), \dots, (E_m, \mathcal{E}_m)$. We say that they are independent if, for any $A'_1 \in \mathcal{E}_1, \dots, A'_m \in \mathcal{E}_m$, we have

$$\mathbb{P}(X_1 \in A'_1, \dots, X_m \in A'_m) = \mathbb{P}(X_1 \in A'_1) \dots \mathbb{P}(X_m \in A'_m)$$

The events $A_1, \dots, A_m \in \mathcal{F}$ are independent if and only if

$$\mathbb{P}(A_{i_1} \cap \dots \cap A_{i_l}) = \mathbb{P}(A_{i_1}) \dots \mathbb{P}(A_{i_l})$$

for any $1 \le l \le m$ and for any $1 \le i_1 < i_2 < \dots < i_l \le m$. This definition is equivalent to say that the random variable $\mathbf{1}_{A_1}, \dots, \mathbf{1}_{A_m}$ are independent.
If $\mathcal{F}_1, \dots, \mathcal{F}_m$ are sub-σ-algebra of $\mathcal{F}$, they are independent if, for any $A_1 \in \mathcal{F}_1, \dots, A_m \in \mathcal{F}_m$, we have:

$$\mathbb{P}(A_1 \cap \dots \cap A_m) = \mathbb{P}(A_1) \dots \mathbb{P}(A_m)$$

We can note that the random variable $X_1, \dots, X_m$ are independent if and only if also the σ-algebra generated by $\sigma(X_1), \dots, \sigma(X_m)$ are independent. Finally, we say that the random variable X is independent from the σ-algebra $\mathcal{G}$ if and only if the σ-algebra $\sigma(X)$ and $\mathcal{G}$ are independent. This happen when X is independent from each random variable W $\mathcal{G}$-measurable.

Now, we need to introduce the relationship between the independence and the law of the random variable. If μ_i is the law of X_i and we set $E = E_1 \times \dots \times E_m$, $\mathcal{E} = \mathcal{E}_1 \otimes \dots \otimes \mathcal{E}_m$ on the space $(E, \mathcal{E})$, we can consider the product law $\mu = \mu_1 \otimes \dots \otimes \mu_m$.
Proposition 1.2 *Consider tha above notation. Then, the random variable $X_1, \dots, X_m$ are independent if and only if the law of X on $(E, \mathcal{E})$ is the product law μ, with $X = (X_1, \dots, X_m)$.*
The proof of this proposition use the following theorem:
Theorem 1.1 [Carathéodory's criterion] *Let $(E, \mathcal{E})$ be a measurable space and let μ_1, μ_2 be two finite measure on $(E, \mathcal{E})$. Let $\mathcal{C}$ be a family of subset of E stable under finite intersection and which built $\mathcal{E}$. If $\mu_1(E) = \mu_2(E)$ and μ_1, μ_2 are equal on $\mathcal{C}$, then they are also equal on all $\mathcal{E}$.*
Proof (proposition 1.2)
Let ν be the law of X and let $\mathcal{C}$ be the set of parts of $\mathcal{E}$ of the form $A_1 \times \dots \times A_m$ with $A_i \in \mathcal{E}_i$, $i = 1, \dots, m$. $\mathcal{C}$ is table under finite intersection and built $\mathcal{E}$, by definition. Then, the definition of independence told us that μ and ν are equal on $\mathcal{C}$, therefore they are also equal over all $\mathcal{E}$, by the Carathéodory criterion.

□

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Then, we have the following definition for conditional probability:
Definition 1.6 [Conditional Probability] *Let $A, B \in \mathcal{F}$ be two events and suppose that $\mathbb{P}(A) > 0$. The conditional probability on $(\Omega, \mathcal{F})$ of B given A is defined as:*

$$\mathbb{P}_A(B) = \mathbb{P}(B|A) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)} \tag{1.4}$$

The conditional probability thus measures the probability of B given that we know that A has occurred.
Let X be a real random variable and Z be a random variable which takes value in a measurable set E such that $\mathbb{P}(Z = z) > 0$ for all $z \in E$. Then, for every $A \subset \mathbb{R}$ and for every $z \in E$, we have:

$$n(z, A) = \mathbb{P}(X \in A | Z = z) = \frac{\mathbb{P}(X \in A, Z = z)}{\mathbb{P}(Z = z)} \tag{1.5}$$

For every $z \in E$, $A \to n(z, A)$ is a probability on $\mathbb{R}$ and it is called the conditional law of X given $Z = z$. Hence, $A \to n(z, A)$ is the law given at the random variable X when we know

that the event $\{Z = z\}$ has occurred.
The conditional expectation of X given $Z = z$ is defined as:

$$E[X|Z = z] = \int xn(z, dx) = \frac{1}{\mathbb{P}(Z = z)} \int_{\{Z=z\}} X d\mathbb{P} = \frac{E[X\mathbf{1}_{\{Z=z\}}]}{\mathbb{P}(Z = z)}$$

Theorem 1.2 *Let X be a quasi-integrable random variable, $\mathcal{D}$ be a sub-σ-algebra of $\mathcal{F}$. We call conditional expectation of X respect to $\mathcal{D}$ and it is denoted by $E[X|\mathcal{D}]$ the class of equivalence random variable Z, $\mathcal{D}$-measurable and quasi-integrable, such that for any $B \in \mathcal{D}$:*

$$\int_B Z d\mathbb{P} = \int_B X d\mathbb{P}$$

A detailed proof can be found in chapter 3 of "Equazioni differenziali stocastiche e applicazioni" written by Baldi.

A random variable Z with these property exists always and it is unique unless there exist an equivalent probability measure.

Proposition 1.3 *Let X, X_1, X_2 be integrale random variable and α, $\beta \in \mathbb{R}$. Then:*

a) $E[\alpha X_1 + \beta X_2|\mathcal{D}] = \alpha E[X_1|\mathcal{D}] + \beta E[X_2|\mathcal{D}]$ *almost surely (a.s);*

b) if $X \geq 0$ *a.s, then* $E[X|\mathcal{D}] \geq 0$ *a.s;*

c) $E[E[X|\mathcal{D}]] = E[X]$*;*

d) if $\mathcal{D}' \subset \mathcal{D}$, $E[E[X|\mathcal{D}]|\mathcal{D}'] = E[X|\mathcal{D}']$ *a.s;*

e) if Z is $\mathcal{D}$-measurable, then $E[ZX|\mathcal{D}] = ZE[X|\mathcal{D}]$ *a.s;*

f) if X is independent from $\mathcal{D}$, $E[X|\mathcal{D}] = E[X]$ *a.s.*

Proof
we prove only the last three point because the first three are just immediate applications of the definition.
d) the random variable $E[E[X|\mathcal{D}]|\mathcal{D}']$ is $\mathcal{D}'$-measurable; moreover, if W is $\mathcal{D}'$-measurable, then it is also $\mathcal{D}$-measurable and:

$$E[WE[E[X|\mathcal{D}]|\mathcal{D}']] = E[WE[X|\mathcal{D}]] = E[WX]$$

e) $ZE[X|\mathcal{D}]$ is $\mathcal{D}$-measurable. If W is $\mathcal{D}$-measurable also ZW is $\mathcal{D}$-measurable, thanks to c) we have:

$$E[ZWE[X|\mathcal{D}]] = E[E[ZWX|\mathcal{D}]] = E[ZWX]$$

f) the random variable $\omega \to E[X]$ is constant and, hence, $\mathcal{D}$-measurable. If W is $\mathcal{D}$-measurable, it is also independent from X and

$$E[WX] = E[W]E[X] = E[WE[X]]$$

therefore $E[X] = E[X|\mathcal{D}]$ a.s

$\square$

Let $\mathcal{H}$ be a σ-algebra and X be a random variable $\mathcal{H}$-measurable. If Z is a random variable independent from $\mathcal{H}$, we know that, if X and Z are integrable,

$$E[XZ|\mathcal{H}] = XE[Z]$$

The equation above is a particular case of the following lemma:
Lemma 1.2 *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, $(E, \mathcal{E})$ be a measurable space, $\mathcal{G}$ and $\mathcal{H}$*

sub-σ-algebra of $\mathcal{F}$. Moreover, $\mathcal{G}$ is independent from $\mathcal{H}$. Let X be a random variable $\mathcal{H}$-measurable which takes value in $(E, \mathcal{E})$ and $\psi(x, \omega)$ be a function on $E \times \Omega$, $\mathcal{E} \otimes \mathcal{G}$-measurable such that $\omega \to \psi(X(\omega), \omega)$ is integrable. Then

$$E[\psi(X, \cdot)|\mathcal{H}] = \Phi(X)$$

where $\Phi(X) = E[\psi(x, \cdot)]$
A detailed proof can be found in chapter 3 of "Equazioni differenziali stocastiche e applicazioni" written by Baldi.

1.1.3 Probability Law

Let μ be a probability law on $\mathbb{R}^n$ and $\pi_i : \mathbb{R}^m \to \mathbb{R}$ his projection on the i-coordinate. Then, we call marginal law of μ the image of the law μ through π_i.
Let $X = (X_1, \ldots, X_m)$ be a random variable on $\mathbb{R}^n$ with probability law μ. We can note that the marginal law is the same of the law μ_i of X_i. Therefore, the marginal law can be expressed as:

$$\mu_i(A) = \int \mathbf{1}_A(x_i)\mu(dx_1, \ldots, dx_m)$$

The probability μ on $\mathbb{R}^m$ admits density with respect the Lebesgue measure, if there exists a borel function f, ≥ 0, such that for each $A \in B(\mathbb{R}^m)$ we have:

$$\mu(A) = \int_A f(x)dx \tag{1.6}$$

Consder two random variable X, Y on $\mathbb{R}^m$ with two independent probability law, respectively, equal to μ and ν. We call convolution product of μ and ν, denoted $\mu * \nu$, the probability law of $X + Y$. We can note that this definition depends only on μ and ν and not on X and Y. In fact, $\mu * \nu$ is the image law on $\mathbb{R}^m$ through the operation $(x, y) \to x + y$ of the law of (X, Y) on $\mathbb{R}^m \times \mathbb{R}^m$.
Proposition 1.4 *Let μ and ν be two probability measure on $\mathbb{R}^m$ with density f and g, respectively. Then, $\mu * \nu$ has density h equal to:*

$$h(x) = \int_{\mathbb{R}^m} f(z)g(x - z)dz$$

Proof

$$\begin{aligned}\mu * \nu(A) &= \int \mathbf{1}_A(z)\mu * \nu(dz) = \int \mathbf{1}_A(x + y)\mu(dx)\nu(dy) \\ &= \int \mathbf{1}_A(x + y)f(x)g(x)dxdy = \int_A dx \int_{\mathbb{R}^m} f(z)g(x - z)dz\end{aligned}$$

□

Now, we need to introduce how we can study the asymptotic behavior of a sequence of random variables. Let X_n, with $n \in \mathbb{N}$, and X be a random variable on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ which take value in $(E, \mathcal{B}(E))$. Then, there are several different notions of convergence:

i. Let $(\mu_n)_n$ be a sequence of finite measure on the measurable space $(E, \mathcal{B}(E))$. Then, μ converges strictly (or weakly) if for each continuous and bounded function f on E we have:

$$\lim_{n\to\infty} \int f d\mu_n = \int f d\mu.$$

ii. We say that $(X_n)_n$ converges almost surely (a.s) to X, written $X_n \xrightarrow{a.s} X$, if there exists an event $n \in \mathcal{F}$ with $\mathbb{P}(N) = 0$ (N is called the exception set) such that

$$\lim_{n\to\infty} X_n(\omega) = X(\omega) \qquad \forall \omega \notin N$$

iii. Let $E = \mathbb{R}^m$. We say that $(X_n)_n$ converges in L^p to X, written $X_n \xrightarrow{L^p} X$, if $X \in L^p$ and
$$\lim_{n\to\infty} E[|X_n - X|^p] = 0.$$

iv. We say that $(X_n)_n$ converges in probability (i.p) to X, written $X_n \xrightarrow{P} X$, if for any $\delta > 0$
$$\lim_{n\to\infty} \mathbb{P}(|X_n - X| > \delta) = 0.$$

v. We say that $(X_n)_n$ converges in distribution to X, written $X_n \xrightarrow{d} X$, if μ_n converges weakly to μ, where μ_n and μ are, respectively, the law of X_n and X. We can note that for this case the random variable X and X_n can be defined in different probability space.

The following proposition explain the relation among the above converges.

Proposition 1.5 *If $X_n \xrightarrow{L^p} X$, then $X_n \xrightarrow{P} X$. If $X_n \xrightarrow{a.s} X$, then $X_n \xrightarrow{P} X$. If $X_n \xrightarrow{P} X$, then $X_n \xrightarrow{d} X$. If $X_n \xrightarrow{P} X$, then exists a sub-sequences $(X_{n_k})_k$ which converges almost surely to X.*

A detailed proof can be found in chapter 7 of "A Probability Path" written by Resnick. The last sentence of the proposition 1.5 imply the uniqueness of the limit in probability.

Let X be a random variable in $\mathbb{R}^m$ and μ be its law. We can define the characteristic function (or the Fourier transform) of μ as:

$$\hat{\mu}(\theta) = \int e^{i\langle\theta,x\rangle}\mu(dx) = E\left[e^{i\langle\theta,X\rangle}\right], \qquad \theta \in \mathbb{R}^m \tag{1.7}$$

The characteristic function is defined for all probability μ on $\mathbb{R}^m$ and has the following properties:

1. $\hat{\mu}(0) = 1$ and $|\hat{\mu}(\theta)| \le 1$, $\forall\theta \in \mathbb{R}^m$;
2. $(\hat{\mu} * \hat{\nu})(\theta) = \hat{\mu}(\theta)\hat{\nu}(\theta)$;
3. $\hat{\mu}$ is uniformly continuous;
4. If μ has moment of order 1 finite, $\hat{\mu}$ admits derivatives and
$$\frac{\partial\hat{\mu}}{\partial\theta_j}(\theta) = i\int x_j e^{i\langle\theta,x\rangle}\mu(dx) \tag{1.8}$$
In particular,
$$\frac{\partial\hat{\mu}}{\partial\theta_j}(0) = i\int x_j\mu(dx)$$
hence $\hat{\mu}'(0) = iE[X]$;
5. If μ has moment of order 2 finite, $\hat{\mu}$ admits second derivatives and
$$\frac{\partial^2\hat{\mu}}{\partial\theta_j\partial\theta_k}(\theta) = -\int x_j x_k e^{i\langle\theta,x\rangle}\mu(dx) \tag{1.9}$$
In particular,
$$\frac{\partial^2\hat{\mu}}{\partial\theta_j\partial\theta_k}(0) = -\int x_j x_k\mu(dx)$$
If $m = 1$ the above equation becomes $\hat{\mu}''(0) = -E[X^2]$;
6. If $\hat{\mu}(\theta) = \hat{\nu}(\theta)$ for each $\theta \in \mathbb{R}^m$, then $\mu = \nu$;

7. Let $X_1, \ldots, X_m$ be random variable with law $\mu_1, \ldots, \mu_m$, respectively. Then, they are independent if and only if the law μ of X, defined as $X = (X_1, \ldots, X_m)$, could be expressed as:
$$\hat{\mu}(\theta_1, \ldots, \theta_m) = \hat{\mu}_1(\theta_1) \ldots \hat{\mu}_m(\theta_m);$$
8. If $\mu_n \underset{n\to\infty}{\longrightarrow} \mu$ strictly, then $\hat{\mu}_n(\theta) \underset{n\to\infty}{\longrightarrow} \hat{\mu}(\theta)$, $\forall\theta$. In fact, $x \to e^{i\langle\theta,x\rangle}$ is a continuous function and its real and imaginary part are bounded. Vice versa, if
$$\lim_{n\to\infty} \hat{\mu}_n(\theta) = \psi(\theta) \qquad \forall\theta \in \mathbb{R}^m$$
and if ψ is continuous at $\theta = 0$, then ψ is the characteristic function of the probability law μ and, moreover, $\mu_n \underset{n\to\infty}{\longrightarrow} \mu$ strictly (Lévy theorem).

A detailed proof can be found in chapter 9 of "A Probability Path" written by Resnick.

For example, consider normal distribution, denoted by $N(a, \sigma^2)$, with mean $a \in \mathbb{R}$ and standard deviation $\sigma > 0$. Its density function can be expressed as:
$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} exp\left\{-\frac{(x-a)^2}{2\sigma^2}\right\} \tag{1.10}$$
Therefore, we can easily compute its characteristic function:
$$\begin{aligned}\hat{\mu}(\theta) &= \int e^{i\langle\theta,x\rangle}\mu(dx)\\ &= \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{+\infty} e^{i\theta x} exp\left\{-\frac{(x-a)^2}{2\sigma^2}\right\} dx\\ &= \frac{e^{i\theta a}}{\sqrt{2\pi}\sigma}\int_{-\infty}^{+\infty} e^{i\theta x} e^{-\frac{x^2}{2\sigma^2}} dx\end{aligned}$$
Set $u(\theta) = \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{+\infty} e^{i\theta x} e^{-\frac{x^2}{2\sigma^2}} dx$, we find $\hat{\mu}(\theta) = u(\theta)e^{i\theta a}$ and, by integration by parts, we have:
$$\begin{aligned}u'(\theta) &= \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} ixe^{i\theta x} e^{-\frac{x^2}{2\sigma^2}} dx\\ &= \left[\frac{1}{\sqrt{2\pi}\sigma}(-i\sigma^2)e^{i\theta x} e^{-\frac{x^2}{2\sigma^2}}\right]_{-\infty}^{+\infty} - \frac{\sigma^2\theta}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} e^{i\theta x} e^{-\frac{x^2}{2\sigma^2}} dx\\ &= -\sigma^2\theta u(\theta)\end{aligned}$$
We find a differential equation of first order. If we make the integral and if we set the condition $u(0) = 1$, we will find
$$\begin{aligned}u(\theta) &= e^{-\frac{1}{2}\sigma^2\theta^2}\\ \hat{\mu}(\theta) &= e^{i\theta a}e^{-\frac{1}{2}\sigma^2\theta^2}\end{aligned}$$

1.2 Stochastic Process

1.2.1 Construction of Stochastic Process

A stochastic process is a family $(X_t)_{0\leq t\leq T}$ of random variables indexed by time. For each realization of the randomness ω, the trajectory $X(\omega) : t \to X_t(\omega)$ defines a function of time, called the sample path of the process. Thus stochastic processes can also be viewed as random functions, hence random variables raking values in function spaces.

Definition 1.7 [Stochastic process] *We call stochastic process an object with the form:*

$$X = (\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in T}, (X_t)_{t\in T}, \mathbb{P})$$

where: T is a subset of $\mathbb{R}^+$; $\mathcal{F}$ is a σ-algebra in Ω; $\mathbb{P}$ is a probability law on $(\Omega, \mathcal{F})$; $(F_t)_{t\in T}$ is a filtration, hence it is an increasing family of sub-σ-algebra of $\mathcal{F}$ in t such that if $s \leq t$, $\mathcal{F}_s \subset \mathcal{F}_t$ and $(X_t)_{t\in T}$ is a family of random variable on $(\Omega, \mathcal{F})$ which takes value in a measurable space $(E, \mathcal{E})$ such that, for any t, X_t is $\mathcal{F}_t$-measurable. In this case, we say that $(X_t)_t$ is adapted to the filtration $(F_t)_t$.

$\mathcal{F}_t$ is then interpreted as the information known at time t, which increases with time. We denoted by $\mathcal{F}_\infty$ the smallest σ-algebra of parts of Ω, which is in $\bigcup_t \mathcal{F}_t$. Moreover, an $\mathcal{F}_t$-measurable random variable is a random variable whose value will be revealed at time t. A process whose value at time t is revealed by the information $\mathcal{F}_t$ is said to be nonanticipating:
Definition 1.8 [Nonanticipating process] (definition 2.12 in [2]) *A stochastic process $(X_t)_{0\leq t\leq T}$ is said to be nonanticipating with respect to the information structure $(\mathcal{F}_t)_{0\leq t\leq T}$ or $\mathcal{F}_t$-adapted if, for each $t \in [0, T]$, the value of X_t is revealed at time t: the random variable X_t is $\mathcal{F}_t$-measurable.*
If the only observation available is the past values of a stochastic process X, then the information is represented by the history, also called the natural filtration, of X defined as follows:
Definition 1.9 [History of a process] (definition 2.13 in [2]) *The history of a process X is the information flows $(\mathcal{F}_t^X)_{0\leq t\leq T}$, where $\mathcal{F}_t^X$ is the σ-algebra generated by the past values of the process, completed by the null sets:*

$$\mathcal{F}_t^X = \sigma(X_s, s \in [0, t]) \bigvee \mathcal{N}$$

where $\mathcal{N} = \{A; A \in \mathcal{F}, \mathbb{P}(A) = 0\}$.

A process is continuous if for any ω, the trajectory $t \to X_t(\omega)$ is continuous. Moreover, X is said measurable if the operation $(t, \omega) \to X_t(\omega)$ is measurable from $(T \times \Omega, \mathcal{B}(T) \otimes \mathcal{F})$ to $(E, \mathcal{B}(E))$. In the next chapter, we will make the assumption that the processes are discontinuous functions. Therefore, we need to introduce the class of càdlàg function:
Definition 1.10 [Càdlàg function] (definition 2.10 in [2]) *A function $f : [0, T] \to \mathbb{R}^d$ is said to be càdlàg if it is right-continuous with left limits: for each $t \in [0, T]$ the limits:*

$$f(t^-) = \lim_{s\to t, s<t} f(s), \qquad f(t^+) = \lim_{s\to t, s>t} f(s)$$

exists and $f(t) = f(t^+)$.
We can note that any continuous function is càdlàg but càdlàg functions can have discontinuities. If t is a discontinuity point we denote by

$$\Delta f(t) = f(t) - f(t^-)$$

the jump of f at t. A càdlàg function f can have at most a countable number of discontinuities, therefore, $\{t \in [0, T], f(t) \neq f(t^-)\}$ is finite or countable. Also, for any $\varepsilon > 0$ the number of jumps on the interval $[0, T]$ larger than ε should be finite. Hence, a càdlàg function on $[0, T]$ has a finite number of "large jumps" and a possibly infinite but countable number of small jumps. An example of càdlàg function could be a step function having a jump at some point T_0, whose value at T_0 is defined to be the value after the jump, hence $f = \mathbf{1}_{[T_0,T)}(t)$. In this case, $f(T_0^-) = 0$, $f(T_0^+) = f(T_0) = 1$ and $\Delta f(T_0) = 1$. More generally, given a continuous function $g : [0, T] \to \mathbb{R}$ and constants f_i, $i = 0, \dots, n-1$ and $t_0 = 0 < t_1 < \dots < t_n = T$, the following function is càdlàg:

$$f(t) = g(t) + \sum_{i=0}^{n-1} f_i \mathbf{1}_{[t_1,t_{i+1})}(t) \tag{1.11}$$

The function g can be interpreted as the continuous component of t to which the jumps have been added, hence the jumps of f occur at t_i, $i \geq 1$ with $\Delta f(t_i) = f_i - f_{i-1}$. Therefore, càdlàg functions are natural model for the trajectories of processes with jumps.

Definition 1.11 [Stopping Times] *Let $(\mathcal{F}_t)_{t\in T}$ be a filtration. A random variable $\tau : \Omega \to T \cup \{+\infty\}$ is called stopping times if, for each $t \in T$, $\{\tau \leq t\} \in \mathcal{F}_t$. Moreover, we set*

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

where $\mathcal{F}_\infty = \bigvee_t \mathcal{F}_t$.
We can note that $\mathcal{F}_\tau$ is the σ-algebra of the events, which at time τ are occurred or not. The following proposition tells us some properties of the stopping time:
Proposition 1.6 *Let σ and τ be two stopping time. Then:*

a) τ is $\mathcal{F}_\tau$-measurable;

b) $\sigma \vee \tau$, $\sigma \wedge \tau$ are stopping times;

c) if $\sigma \leq \tau$, $\mathcal{F}_\sigma \subset \mathcal{F}_\tau$;

d) $\mathcal{F}_{\sigma\wedge\tau} = \mathcal{F}_\sigma \cap \mathcal{F}_\tau$.

Proof
a)
For each $s \geq 0$, $\{\tau \leq s\} \in \mathcal{F}_\tau$. Its clear that $\{\tau \leq s\} \in \mathcal{F}_\tau \subset \mathcal{F}_\infty$.
For each t, $\{\tau \leq s\} \cap \{\tau \leq t\} \in \mathcal{F}_t$.
If $t \leq s$, we have that $\{\tau \leq s\} \cap \{\tau \leq t\} = \{\tau \leq t\} \in \mathcal{F}_t$.
If $t > s$, we have that $\{\tau \leq s\} \cap \{\tau \leq t\} = \{\tau \leq s\} \in \mathcal{F}_s \subset \mathcal{F}_t$.
b)
$\{\sigma \wedge \tau \leq t\} = \{\sigma \leq t\} \cup \{\tau \leq t\} \in \mathcal{F}_t \implies$ is a stopping time.
$\{\sigma \vee \tau \leq t\} = \{\sigma \leq t\} \cap \{\tau \leq t\} \in \mathcal{F}_t \implies$ is a stopping time.
c)
If $A \in \mathcal{F}_\sigma$, then for each $t:$ $A \cap \{\sigma \leq t\} \in \mathcal{F}_t$. Therefore, $\{\tau \leq t\} \subset \{\sigma \leq t\}:$

$$A \cap \{\tau \leq t\} = \underbrace{A \cap \{\sigma \leq t\}}_{\in \mathcal{F}_t} \cap \underbrace{\{\tau \leq t\}}_{\in \mathcal{F}_t}$$

d)
Let $A \in \mathcal{F}_\sigma \cap \mathcal{F}_\tau$, then $A \in \mathcal{F}_\infty$, $A \cap \{\tau \leq t\} \in \mathcal{F}_t$ and $A \cap \{\tau \leq s\} \in \mathcal{F}_t$. Therefore, we have:

$$\begin{aligned} A \cap \{\sigma \wedge \tau \leq t\} &= A \cap (\{\sigma \leq t\} \cup \{\tau \leq t\}) \\ &= (A \cap \{\sigma \leq t\}) \cup (A \cap \{\tau \leq t\}) \in \mathcal{F}_t \end{aligned}$$

Therefore, $A \in \mathcal{F}_{\sigma\wedge\tau}$.

□

Proposition 1.7 *Let X be a measurable process and $\sigma : \Omega \to \mathbb{R}^+$ be a random variable. Then $X_\sigma : \omega \to X_{\sigma(\omega)}(\omega)$ is a random variable. If τ is a finite stopping time almost surely and X is measurable, then X_τ is $\mathcal{F}_\tau$-measurable.*
A detailed proof can be found in chapter 1 of "Equazioni differenziali stocastiche e applicazioni" written by Baldi.

1.2.2 Brownian motion

Definition 1.12 [Brownian Motion] *A real-valued process $B = (\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, (B_t)_{t\geq 0}, \mathbb{P})$ is a Brownian motion if* *i)* $B_0 = 0$ *a.s;*

ii) *for each* $0 \le s \le t$*, the random variable* $B_t - B_s$ *is independent of* $\mathcal{F}_s = \sigma(B_u, u \le s)$*;*

iii) *for each* $0 \le s \le t$*,* $B_t - B_s$ *has law* $N(0, t-s)$*.*

The Brownian motion is a continuous stochastic process, in fact, the map $s \mapsto B_s(\omega)$ is continuous. We can note that the Brownian motion is a Gaussian process and we define a Gaussian process as:

Definition 1.13 *Let* $\mathcal{I}$ *be a family of random variable in* $\mathbb{R}^d$ *defined on the probability space* $(\Omega, \mathcal{F}, \mathbb{P})$*. Then, we say that* $\mathcal{I}$ *is a Gaussian process if, for each* $X_1, \dots, X_m \in \mathcal{I}$ *and* $\gamma_1, \dots, \gamma_m \in \mathbb{R}^d$*, the random variable* $\langle \gamma_1, X_1 \rangle + \cdots + \langle \gamma_m, X_m \rangle$ *is Gaussian.*

Moreover, the point *iii)* imply that the Brownian motion has stationary increments, therefore, if $s \le t$, $B_t - B_s$ and $B_{t-s} - B_0$ have the same probability law. We shall also need a definition of a Brownian motion with respect to a filtration $(\mathcal{F}_t)$.

Definition 1.14 (definition 3.2.5 in [1]) *A real-valued, continuous stochastic process is an* $(\mathcal{F}_t)$*-Brownian motion if it satisfies:*

a) *for any* $t \ge 0$*,* B_t *is* $\mathcal{F}_t$*- measurable;*

b) *if* $s \le t$*,* $B_t - B_s$ *is independent of the* σ*-algebra* $\mathcal{F}_s$*;*

c) *if* $s \le t$*,* $B_t - B_s$ *and* $B_{t-s} - B_0$ *have the same law.*

We can note that if B is a $(\mathcal{F}_t)_t$-Brownian motion, it is also a Brownian motion with respect other filtration $(\mathcal{F}'_t)_t$ smaller than $(\mathcal{F}_t)_t$. Finally, we say that we have a natural Brownian motion when $(\mathcal{F}_t)_t$ is a natural filtration.

Proposition 1.8 *If* B *is a Brownian motion, then*

1) $B_0 = 0$ *a.s;*

2) *for each* $0 \le t_1 < \cdots < t_m$*,* $(B_{t_1}, \dots, B_{t_m})$ *is "centered" normal random variable in* $\mathbb{R}^m$*;*

3) $E[B_s B_t] = s \wedge t$*.*

Vice versa, if 1), 2) and 3) are true, then B *is a natural Brownian motion. Proof*

1)

If B is a Brownian motion, $B_0 = 0$ by definition.

2)

is a consequence of the definition 1.13 and the point *iii)* of the definition 1.12

3)

is $s \le t$

$$E[B_s B_t] = E[(B_t - B_s)B_s] + E[B_s^2] = s = s \wedge t$$

Vice versa, if B satisfy *1), 2) and 3)*, then *i)* of the definition 1.12 is clear. Moreover, if $0 \le s < t$, $B_t - B_s$ is a normal random variable as linear function of (B_s, B_t) and it is centered because B_t and B_s are centered; since

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

hence, $B_t - B_s$ has distribution $N(0, t-s)$ and *iii)* is satisfy. Finally, if $u \le s \le t$

$$E[(B_t - B_s)B_u)] = E[B_t B_u] - E[B_s B_u] = t \wedge u - s \wedge u = 0$$

hence, $B_t - B_s$ is independent of $\sigma(B_u, u \le s)$.

$\square$

We say that a Brownian motion is standard if $B_0 = 0$ and its first two moment are, respectively, equal to: $E[B_t] = 0$ and $[B_t^2] = t$.
Now, we need to study the behavior of the trajectories of the Brownian motion. Therefore, we need to introduce the following corollary:
Corollary 1.1 *Let X be a process which takes value in $\mathbb{R}^d$ such that exists an $\alpha > 0, \beta > 0, c > 0$ such that, for each s, t,*

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

Then, exist a modification Y of X which is continuous. Moreover, if for each $\gamma < \frac{\alpha}{\beta}$, the trajectories of Y has Hölder exponent γ in each bounden time interval.
We can note that a Brownian motion admits always a continuity modification of it, which still remain a Brownian motion. Let $t > s$, since $B_t - B_s \sim N(0, t-s)$, we have that $B_t - B_s = (t-s)^{1/2}Z$, where $Z \sim N(0,1)$. Therefore, for $p \geq 0$,

$$E[|B_t - B_s|^{2p}] = (t-s)^p E[|Z|^{2p}]$$

We can note that $E[|Z|^{2p}] < \infty$ for each $p > 0$, then we can apply the corollary 1.1 with $\beta = 2p$, $\alpha = p - 1$ and we find that for any Brownian motions exist always a continuity modification. Moreover, these are Hölder with exponent γ, therefore $\gamma < \frac{p-1}{2p}$ which imply that $\gamma < \frac{1}{2}$.
We have seen that a Brownian motion admits always a continuity version and it is Hölder with exponent γ, for each $\gamma < \frac{1}{2}$. Now, we need to study the behavior of the trajectories and we denote with $X = (\Omega, \mathcal{F}, (\mathcal{F}_t)_t, (X_t)_t, \mathbb{P})$ a continuous Brownian motion.
Definition 1.15 *Let $I \subset \mathbb{R}$ be an interval and $f : I \to \mathbb{R}$ be a continuous function. we call modulus of continuity of f the function, for $x, y \in I$,:*

$$w(\delta) = \sup_{|x-y| \leq \delta} |f(x) - f(y)|$$

Theorem 1.3 *(P. Lévy) For each $T > 0$*

$$\mathbb{P}\left(\lim_{\delta \to 0^+} \sup_{0 \leq s < t \leq T; t-s \leq \delta} \frac{|X_t - X_s|}{(2\delta \log \frac{1}{\delta})^{1/2}} = 1 \right) = 1$$

A detailed proof can be found in chapter 1 of "Equazioni differenziali stocastiche e applicazioni" written by Baldi.
The P-Lévy theorem say that the trajectories can not be Hölder with exponent $\gamma = \frac{1}{2}$ on the interval $[0, T]$ for each T.

Definition 1.16 [Total Variation] *Given a function $f : \mathbb{R} \to \mathbb{R}$, we call total variation of f in $[a, b]$ the quantity:*

$$V_b^a f = \sup_\pi \sum |f(x_{i+1}) - f(x_i)|$$

where π changes among all the partitions $a = x_0 < x_1 < \cdots < x_{n+1} = b$ on the interval $[a, b]$. Moreover, f is said finite if $V_b^a f < +\infty$ for each $a, b \in \mathbb{R}$.
Proposition 1.9 *Let $\pi = \{t_0, \ldots, t_m\}$ with $s = t_0 < t_1 < \cdots < t_m = t$ be a partition on the interval $[s, t]$, $|\pi| = \max_{0 \leq k \leq m-1} |t_{k+1} - t_k|$. Then, if we set*

$$S_\pi = \sum_{k=0}^{m-1} |X_{t_{k+1}} - X_{t_k}|^2$$

we have that

$$\lim_{|\pi| \to 0^+} S_\pi = t - s \qquad \in L^2 \tag{1.12}$$

Therefore, the trajectories of a Brownian motion have not finite total variation in any time interval, almost surely.

Proof
Let $\tau = t - s$, then $\tau = \sum_{k=0}^{m} t_{k+1} - t_k$ and

$$S_\pi - \tau = \sum_{k=0}^{m-1} [(X_{t_{k+1}} - X_{t_k})^2 - (t_{k+1} - t_k)]$$

since the random variable $(X_{t_{k+1}} - X_{t_k})^2 - (t_{k+1} - t_k)$ are independent and centered, we have:

$$E[(S_\pi - \tau)^2] = \sum_{k=0}^{m-1} E[\big((X_{t_{k+1}} - X_{t_k})^2 - (t_{k+1} - t_k)\big)^2]$$

Therefore, we have:

$$E[(S_\pi - \tau)^2] = \sum_{k=0}^{m-1} (t_{k+1} - t_k)^2 E\left[\left(\frac{(X_{t_{k+1}} - X_{t_k})^2}{t_{k+1} - t_k} - 1\right)^2\right]$$

but for each k the random variable $\frac{X_{t_{k+1}} - X_{t_k}}{t_{k+1} - t_k}$ is $N(0,1)$, hence the quantity

$$c = E\left[\left(\frac{(X_{t_{k+1}} - X_{t_k})^2}{t_{k+1} - t_k} - 1\right)^2\right]$$

is finite and is not depend on k. Therefore

$$E[(S_\pi - \tau)^2] = c \sum_{k=0}^{m-1} (t_{k+1} - t_k)^2 \leq c|\pi| \sum_{k=0}^{m-1} |t_{k+1} - t_k| = c|\pi|(t - s) = \longrightarrow 0$$

which proves the equation (1.12). Moreover,

$$S_\pi = \sum_{k=0}^{m-1} |X_{t_{k+1}} - X_{t_k}|^2 \leq \max_{0 \leq i \leq m-1} |X_{t_{i+1}} - X_{t_i}| \sum_{k=0}^{m-1} |X_{t_{k+1}} - X_{t_k}|$$

Therefore, the trajectories are continuous,

$$\lim_{|\pi \to 0^+} \max_{0 \leq i \leq m-1} |X_{t_{i+1}} - X_{t_i}| = 0$$

hence, if the trajectories have finite total variation on the interval $[s, t]$ for ω in a set A, then in A we can have:

$$\lim_{|\pi| \to 0^+} \sum_{k=0}^{m-1} |X_{t_{k+1}} - X_{t_k}| < +\infty$$

and, in conclusion, $\lim_{|\pi| \to 0^+} S_\pi(\omega) = 0$, which contradict the proposition 1.9

□

We can note that if a function f has finite total variation, then we can definite the integral

$$\int_0^T \phi(t) df(t)$$

for each bounded borel function ϕ. But the proposition 1.9 say that we can not make the integral ω for ω because the trajectories of the Brownian motion have not finite total variation. Therefore, in the following chapter, we need to introduce the stochastic integral:

$$\int_0^T \phi(t) dX_t(\omega)$$

1.2.3 Martingale

Definition 1.17 [Martingale] *A real-valued process $M = (\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in T}, (M_t)_{t\in T}, \mathbb{P})$ is a martingale if M_t is integrable for each $t \in T$ and*

$$E[M_t|\mathcal{F}_s] = M_s \qquad \forall s \le t \tag{1.13}$$

it is a supermartingale if

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

it is a submartingale if

$$E[M_t|\mathcal{F}_s] \ge M_s \qquad \forall s \le t$$

It is clear that linear combination of martingale are still martingale. If $(M_t)_t$ is a supermartingale, then $(-M_t)_t$ is a submartingale, and vice versa. If $(M_t)_t$ is a martingale, then $(|M_t|)_t$ is a submartingale. Moreover, if M is a martingale (respectively a submartingale) and $\Phi : \mathbb{R} \to \mathbb{R}$ is a convex function (respectively an increasing convex function) such that $\Phi(M_t)$ is integrable, then $(\Phi(M_t))_t$ is a submartingale. This is a consequence of Jensen's inequality. Finally, we say that a martingale $(M_t)_t$ is in L^p, $p \ge 1$, if $M_t \in L^p$ for each t and we say that a martingale is square integrable for $p = 2$.

Now, we give some result for martingale in discrete time and then we extend it to the continuous one. Let $T = \mathbb{N}$. A process $(A_n)_n$ adapted to the filtration $(\mathcal{F}_n)_n$ is called increasing predictable process if $A_0 = 0$, $A_n \le A_{n+1}$ and A_{n+1} is $\mathcal{F}_n$ measurable. Let $(X_n)_n$ be a $(\mathcal{F}_n)_n$-submartingale and we set

$$A_0 = 0; \qquad\qquad A_{n+1} = A_n + E[X_{n+1}|\mathcal{F}_n] - X_n$$

By construction $(A_n)_n$ is an increasing predictable process. If $M_n = X_n - A_n$, then

$$E[M_{n+1}|\mathcal{F}_n] = E[X_{n+1}|\mathcal{F}_n] - A_{n+1} = X_n - A_n = M_n$$

Therefore $(M_n)_n$ is a martingale. Consider another decomposition of $(X_n)_n$: $X_n = M'_n + A'_n$, where M' is the martingale part and A' is the increasing predictable process. Then, we have $A_0 = A'_0 = 0$ and

$$A'_{n+1} - A'_n = X_{n+1} - X_n - (M'_{n+1} - M'_n)$$

If we take the conditional expectation respect to $\mathcal{F}_n$, we find:

$$A'_{n+1} - A'_n = E[X_{n+1}|\mathcal{F}_n] - X_n$$

therefore $A'_n = A_n$ and $M'_n = M_n$. This result is called Doob's decomposition, which shows that each submartingale $(X_n)_n$ can be decomposed in a sum of predictable increasing process $(A_n)_n$ and martingale part $(M_n)_n$.

If $(X_n)_n$ is a martingale, also the stopping process $X^\tau_n = X_{n\wedge\tau}$ is a martingale; in this notation τ is a stopping time of the filtration $(\mathcal{F}_n)_n$. In fact, by definition of stopping time $\{\tau \ge n+1\} = \{\tau \le n\}^c \in \mathcal{F}_n$ and, since, $X^\tau_{n+1} = X^\tau_n$ on $\{\tau \le n\}$, we have

$$\begin{aligned} E[X^\tau_{n+1} - X^\tau_n|\mathcal{F}_n] &= E[(X_{n+1} - X_n)\mathbf{1}_{\{\tau \ge n+1\}}|\mathcal{F}_n] \\ &= \mathbf{1}_{\{\tau\ge n+1\}} E[X_{n+1} - X_n|\mathcal{F}_n] \\ &= 0 \end{aligned}$$

Theorem 1.4 (Sampling Theorem) *Let $X = (\Omega, \mathcal{F}, (\mathcal{F}_n)_n, (X_n)_n, \mathbb{P})$ be a supermartingale and let τ_1 and τ_2 be two stopping time associated to the filtration $(\mathcal{F}_n)_n$, bounded a.s and such that $\tau_1 \le \tau_2$ a.s. Then, the random variable X_{τ_1} and X_{τ_2} are integrable and*

$$E[X_{\tau_2}|\mathcal{F}_{\tau_1}] \le X_{\tau_1} \tag{1.14}$$

Proof

The integrability of the random variable X_{τ_1} and X_{τ_2} are obviuos because, for $i = 2, 2$, $|X_{\tau_i}| \le$

$\sum_{j=1}^{k} |X_j|$, where k is a number which is added to τ_2.
Let $\tau_2 \equiv k \in \mathbb{N}$ and let $A \in \mathcal{F}_{\tau_1}$. Since $A \cap \{\tau_1 = j\} = \mathcal{F}_j$, we have, for $j \leq k$,

$$\int_{A\cap\{\tau_1=j\}} X_{\tau_1} d\mathbb{P} = \int_{A\cap\{\tau_1=j\}} X_j d\mathbb{P} \geq \int_{A\cap\{\tau_1=j\}} X_k d\mathbb{P}$$

and, making the sum respect to j, with $0 \leq j \leq k$, we find

$$\int_A X_{\tau_1} d\mathbb{P} = \sum_{j=0}^{k} \int_{A\cap\{\tau_1=j\}} X_j d\mathbb{P} \geq \sum_{j=0}^{k} \int_{A\cap\{\tau_1=j\}} X_k d\mathbb{P} = \int_A X_{\tau_2} d\mathbb{P}$$

Hence, we have shown that the theorem hold if τ_2 is a constant stopping time. Now, we change the hypothesis above with $\tau_2 \leq k$. If we apply the result find in the first part of this proof at the martingale $(X_n^{\tau_2})_n$, at the stopping time τ_1 and at k, we find

$$\int_A X_{\tau_1} d\mathbb{P} = \int_A X_{\tau_1}^{\tau_2} d\mathbb{P} \geq \int_A X_k^{\tau_2} d\mathbb{P} = \int_A X_{\tau_2} d\mathbb{P}$$

which ends the proof.

□

Corollary 1.2 *Consider the hypothesis of theorem 1.4. If X is a martingale, then we have*

$$E[X_{\tau_2} | \mathcal{F}_{\tau_1}] = X_{\tau_1} \tag{1.15}$$

A process $(X_t)_{t\in[0,T]}$ is called a local martingale if there exists a sequence of stopping times (τ_n) with $\tau_n \to \infty$ a.s such that $(X_{t\wedge\tau_n})_{t\in[0,T]}$ is a martingale. Thus a local martingale behaves like a martingale up to some stopping time τ_n, which can be chosen as large as one wants. Moreover, any martingales are a local martingale but there exists local martingales which are not martingales.

A martingale M is bounded in L^p if $\sup_n E[|M_n|^p] < +\infty$. Then, we have the following theorem:
Theorem 1.5 (Doob's inequality) *Let $X = (\Omega, \mathcal{F}, (\mathcal{F}_n)_n, (M_n)_n, \mathbb{P})$ be a bounded martingale in L^p, $p > 1$. Then, let $M^* = \sup_n |M_n| \in L^p$ and*

$$\|M^*\|_p \leq q \sup_n \|M_n\|_p$$

where $q = \frac{p}{p-1}$.
This theorem is a consequence of the following lemma:
Lemma 1.3 *If X is a positive submartingale, then for each $\alpha > 1$ and $n \in \mathbb{N}$:*

$$E\left[\max_{0<\leq i\leq n} X_i^\alpha\right] \leq \left(\frac{\alpha}{\alpha-1}\right)^\alpha E[X_n^\alpha]$$

A detailed proof can be found in chapter 4 of "Equazioni differenziali stocastiche e applicazioni" written by Baldi.
Theorem 1.6 *Let X be a supermartingale such that $sup_{n\geq 0} E[X_n^-] < +\infty$. Then, X converges a.s and it has finite limit.*
A detailed proof can be found in chapter 4 of "Equazioni differenziali stocastiche e applicazioni" written by Baldi.
If M is a martingale bounded in L^p, then it has $\sup_{n\geq 0} M_n^- \leq M^*$. Therefore, the martingale M converges a.s to a random variable M_∞ such that $|M_\infty| \leq M^*$. Since $|M_n - M_\infty|^p \leq 2^{p-1}(|M_n|^p + |M_\infty|^p) \leq 2^p M^{*p}$, we can apply the Lebesgue theorem and we find

$$\lim_{n\to\infty} E[|M_n - M_\infty|^p] = 0$$

Hence, for $p > 1$, the following theorem told us the behavior of a martingale bounded in L^p:
Theorem 1.7 *If $p > 1$, a martingale bounded in L^p converges a.s and in L^p.*

Now, we need to study the converges of martingale in L^1 and we need to introduce the definition of uniformly integrable:
Definition 1.18 *Let $\mathcal{H}$ be a family of random variable in $\mathbb{R}$. We say that $\mathcal{H}$ is uniformly integrable if:*

$$\lim_{c\to+\infty}\sup_{Y\in\mathcal{H}}\int_{\{|Y|>c\}}|Y|d\mathbb{P}=0$$

Consider a set built with one random variable, this is the easiest example of set uniformly integrable. Therefore, we have $\lim_{c\to+\infty}|Y|\mathbf{1}_{\{|Y|>c\}}=0$ a.s and, since $|Y|\mathbf{1}_{\{|Y|>c\}}\leq|Y|$ and for the Lebesgue theorem, we have

$$\lim_{c\to+\infty}\int_{\{|Y|>c\}}|Y|d\mathbb{P}=0$$

Hence, $\mathcal{H}$ is uniformly integrable if there exists a real integrable random variable Z such that $Z\geq|Y|$ for each $Y\in\mathcal{H}$. Therefore, in this case we have:

$$\int_{\{|Y|>c\}}|Y|d\mathbb{P}\leq\int_{\{Z>c\}}Zd\mathbb{P}$$

Then the following theorem is an extension of the Lebesgue theorem:
Theorem 1.8 *Let $(Y_n)_n$ be a sequence of random variable convergent a.s to Y. A necessary and sufficient condition for Y to be integrable and to be convergent in L^1 is that $(Y_n)_n$ is uniformly integrable.*
A detailed proof can be found in chapter 6 of "A Probability Path" written by Resnick.
In any case, a family $\mathcal{H}$ uniformly integrable is bounded in L^1. In fact, let $c>0$ such that $\sup_{Y\in\mathcal{H}}\int_{\{|Y|>c\}}|Y|d\mathbb{P}\leq1$, then we have, for each $Y\in\mathcal{H}$,

$$E[|Y|]=\int_{\{|Y|>c\}}|Y|d\mathbb{P}+\int_{\{|Y|\leq c\}}|Y|d\mathbb{P}\leq c+1$$

We can note that if $(\mathcal{F}_n)_n$ is a filtration on the probability space $(\Omega,\mathcal{F},\mathbb{P})$ and $Y\in L^1$, $(E[Y|\mathcal{F}_n])_n$ is a uniform integrable martingale. Vice versa, if $M=(M_n)_n$ is a uniform integrable martingale, then it is bounded in L^1. Therefore, the theorem 1.6 is satisfy and M converges a.s to a random variable Y. Moreover, by theorem 1.8, $Y\in L^1$ and it converges in L^1. Hence, we have

$$M_m=E[M_n|\mathcal{F}_m]\underset{n\to\infty}{\longrightarrow}E[Y|\mathcal{F}_m]\qquad\in L^1$$

Therefore, we have already proved the following theorem:
Theorem 1.9 *A martingale $(M_n)_n$ is uniformly integrable if and only it it has the form $M_n=E[Y|\mathcal{F}_n]$, where $Y\in L^1(\Omega,\mathcal{F},\mathbb{P})$. In this case, $(M_n)_n$ converges a.s and in L^1.*

Finally, we change the assumption from $T\in\mathbb{R}$ to $T\in\mathbb{R}^+$. Hence, we consider the martingale in continuous time. The theorem above are still valid for the continuous time case but, in this case, we require that the martingale is right continuous. We can rewrite the theorem 1.4 for the continuous martingale as:
Theorem 1.10 (Optimal Sampling theorem) (theorem 3.3.4 in [1]) *If $(M_t)_{t\geq0}$ is a continuous martingale with respect to the filtration $(\mathcal{F}_t)_{t\geq0}$, and if τ_1 and τ_2 are two stopping time such that $\tau_1\leq\tau_2\leq K$, where K is a finite real number, then M_{τ_2} is integrable and*

$$E[M_{\tau_2}|\mathcal{F}_{\tau_1}]=M_{\tau_1}\quad a.s\tag{1.16}$$

This result implies that if τ is a bounded stopping time, then $E[M_\tau]=E[M_0]$. Moreover, let $M=(\Omega,,(\mathcal{F}_t)_t,(M_t)_t,\mathbb{P})$ be a right continuous martingale and let τ be a stopping time. Then $(M_{t\wedge\tau})_t$ is a martingale respect to the filtration $(\mathcal{F}_t)_t$.
We need to study if the Doob's decomposition is true for submartingale in continuous time. Therefore, we introduce the following theorem (we did not prove the theorem but a detailed

proof can be found in chapter 1 of "Brownian Motion and Stochastic Calculus, 2^{nd} edition" written by Karatzas and Shreve):
Theorem 1.11 *Let M be a continuous square integrable martingale respect to the filtration $(\mathcal{F}_t)_t$ completed with the null set. Then, exists an unique continuous increasing process A, with $A_0 = 0$, such that $(M_t^2 - A_t)_t$ is a martingale. If $\pi = \{0 = t_0 < t_1 < \cdots < t_m = t\}$ is a partition on the interval $[0,t]$, then we have*

$$A_t = \lim_{|\pi| \to 0} \sum_{k=0}^{m-1} |M_{t_{k+1}} - M_{t_k}|^2$$

We call the process $(A_t)_t$ of the theorem 1.11 an increasing process respect to the square integrable martingale M and we use the notation $\langle M \rangle_t$. We can note that the increasing process respect to a continuous martingale does not depend on the filtration of the martingale. Hence, if we have two filtration completed with the null set, $(\mathcal{F}_t)_t$ and $(\tilde{\mathcal{F}}_t)_t$, then $(M_t)_t$ is a martingale respect both filtration and the increasing process respect the two filtration $(\mathcal{F}_t)_t$ and $(\tilde{\mathcal{F}}_t)_t$ are equal. Moreover, the increasing process A respect to a Brownian motion is equal to $A_t = t$.
Theorem 1.11 imply that all the continuous square integrable martingale have not finite total variation. Therefore, we have the following proposition:
Proposition 1.10 *Let M be a continuous square integrable martingale, then on $\{\langle M \rangle_t > 0\}$ M has not finite total variation on the interval $[0,t]$ a.s. On the other hand, on $\{\langle M \rangle_t = 0\}$ M is constant on the interval $[0,t]$ a.s.*
Proof
The first part of the proof is exactly the same of preposition 1.9. We have, for each partition $\pi = \{0 = t_0 < t_1 < \cdots < t_m = t\}$

$$\sum_{k=0}^{m-1} |M_{t_{k+1}} - M_{t_k}|^2 \leq \sup_{1 \leq i \leq m-1} |M_{t_{i+1}} - M_{t_i}| \sum_{k=0}^{m-1} |M_{t_{k+1}} - M_{t_k}|$$

On $\{\langle M \rangle_t > 0\}$, the left term converges to $\langle M \rangle_t > 0$ and the right term, if $t \to M_t(\omega)$ has total variation, converges to 0 for $|\pi| \to 0$. This conclude the first part of the proof.
For the second part of the proof we can suppose that $M_0 = 0$ because $(M_t - M_0)_t$ is still a martingale with the same increasing process. Let $\tau = \inf\{t; \langle M \rangle_t > 0\}$ be a stopping time and, since $\langle M \rangle_s = 0$ for $s \leq \tau$ and $X_t = M_t^2 - \langle M \rangle_t$ is a null martingale in 0, we have

$$E[M_{t\wedge\tau}^2] = E[M_{t\wedge\tau}^2 - \langle M \rangle_{t\wedge\tau}] = E[X_{t\wedge\tau}] = E[X_0] = 0$$

If we apply the following inequality:

$$\lambda \mathbb{P}\left(\inf_{0 \leq t \leq T} M_t \leq -\lambda\right) \leq \int_{\{\inf_{0 \leq t \leq T} X_t \leq -\lambda\}} -M_T d\mathbb{P} \leq E[|M_T|]$$

at the negative supermartingale $(-M_{t\wedge\tau}^2)_t$, we find that the supermartingale is equal to 0 a.s. This conclude the second part of the proof.

□

Chapter 2

Jump Process

In the previous chapter we have seen the most well-known continuous process: the Brownian motion. Here, we introduce and explain a family of discontinuous process called Lévy processes. We begin with the definition of a Poisson process, which is the main building block for stochastic process with discontinuous trajectories. Then, we talk about compound Poisson process, which is use to built a jump-diffusion model, and we study its property. The second section of the chapter, starts with the definition of Lévy process, then we discuss its infinitely divisible distribution and we present the Lévy-Khintchine formula, which links processes to distributions. The opposite way, from distribution to processes, is the subject of the Lévy-Ito decomposition of a Lévy process. The Lévy measure, which is responsable for the richness of the class of Lévy processes, is studied in some detail and we use it to draw some conclusions about the path and the moment of a Lévy process.
The last section uses the Lévy processes and its properties to built a model for financial applications, which can be decomposed in two main categories: the jump diffusion model and the infinite activity models. Here, we give some example of jump diffusion model and we explain the properties and the relationship between the ordinary and stochastic exponential models.

2.1 Poisson Process

2.1.1 Definition and Properties

Definition 2.1 [Poisson Process] (definition 7.1.1 in [1]) *Let $(T_i)_{i\geq 1}$ be a sequence of independent, identically, exponentially distributed random variables* [1] *with parameters $\lambda(\lambda > 0)$ and let $\tau_n = \sum_{i=1}^n T_i$. We call Poisson process with intensity λ the process N_t defined by:*

$$N_t = \sum_{n\geq 1} \mathbf{1}_{\{\tau_n \leq t\}} = \sum_{n\geq 1} n\mathbf{1}_{\{\tau_n \leq t < \tau_{n+1}\}}$$

Where N_t indicates the number of points of the sequence $(\tau_n)_{n\geq 1}$ which are smaller than or equal to t. Moreover, a poisson process can be described as a counting process. Given an increasing sequence of random times $\{\tau_n, n \geq 1\}$ with $\mathbb{P}(\tau_n \to \infty) = 1$, we can define the associated counting process $(X_t)_{t\geq 0}$ with

$$X_t = \sum_{n\geq 1} \mathbf{1}_{\{\tau_n \leq t\}} = \#\{n \geq 1, \tau_n \geq t\}$$

The condition $\mathbb{P}(\tau_n \to \infty) = 1$ told that X_t is well-defined, hence finite, for any $t \geq 0$ with probability 1. Therefore, the Poisson process counts the number of random times (τ_n) which

[1] A positive random variable X follow an exponential distribution with parameter $\lambda > 0$ if it has a probability density function equal to

$$\lambda e^{-\lambda x}\mathbf{1}_{\{x>0\}}$$

occur between 0 and t, where $(\tau_n - \tau_{n-1})_{n\geq 1}$ is an independent and identically distributed (i.i.d.) sequence of exponential variables.
Let $(N_t)_{t\geq 0}$ be a Poisson process and it has the following properties:

1) For all $t \geq 0$, N_t is almost surely (a.s.) finite;
2) The trajectories of N (in other words: $\forall\omega$, the sample path $t \mapsto N_t(\omega)$) are piecewise constant with jumps of size 1;
3) The trajectories are right continuous with left limit (càdlàg[2]);
4) $\forall t > 0$, $N_{t^-} = N_t$ with probability 1;
5) (N_t) is continuous in probability:
$$\forall t > 0, N_s \xrightarrow[s\to t]{\mathbb{P}} N_t;$$
6) $\forall t > 0$, N_t follows a Poisson distribution with parameter λt:
$$\forall n \in \mathbb{N},\, \mathbb{P}(N_t = n) = e^{-\lambda t}\frac{(\lambda t)^n}{n!}$$
7) The characteristic function of N_t is
$$E[e^{iuN_t}] = exp\left\{\lambda t(e^{iu} - 1)\right\},\, \forall u \in \mathbb{R};$$
8) Independence of increments: for all $0 \leq t_0 < t_1 < \cdots < t_n$ and $n \geq 1$ the increments
$$N_{t_1} - N_{t_0}, \ldots, N_{t_n} - N_{t_{n-1}}$$
are mutually independent random variables. In other words, if $s > 0$, $N_{t_n} - N_{t_{n-1}}$ is independent of the σ-algebra $\mathcal{F}_t$;
9) Stationarity of increments: $N_{t+h} - N_{s+h}$ has the same distribution as $N_t - N_s$ for all $h > 0$ and $0 \leq s \leq t$. Hence, the law of $N_{t+s} - N_t$ is identical to the law of $N_s - N_0 = N_s$;
10) (N_t) has the Markov property:
$$\forall t > s,\, E[f(N_t)|N_u, u \leq s] = E[f(N_t)|N_s];$$
11) The Poisson process is a Lévy process.

A detailed proof of this property can be found in chapter 2 of "Financial Modeling with Jump Process" written by Cont and Tankov.

The right continuity, càdlàg property, of the Poisson process is not really a "property". In fact, we have defined N_t in such a way that at a discontinuity point $N_t = N_{t^+}$ but a function could be càglàd (left continuous with right limit, in this case we have $f(t) = f(t^-)$ and $N_t = N_{t^-}$). There is a difference between a càdlàg and a càglàg process especially in the context of financial modeling. In fact, if a right continuous function has a jump at time t, then the value $f(t)$ is not predictable by following the trajectory up to time t and the discontinuity is seen as a sudden event. On the other hand, if the function was left continuous, an observer approaching t along the path could predict the value at t. Hence, jumps represent unexpected, unforeseeable events and the assumption of right-continuity is natural. By contrast, we should use a càglàd process if we want to model a discontinuous process whose values are predictable. This will be the case when we want built trading strategies.

Theorem 2.1 *Assume that the counting process $(N_t)_{t\in\mathbb{R}_+}$ satisfies the independence and the stationary of increments property. Then for all fixed $0 \leq s \leq t$ we have:*

[2] Defined in chapter 1

$$\mathbb{P}(N_t - N_s = k) = e^{-\lambda(t-s)}\frac{(\lambda(t-s))^k}{k!}, \qquad k \in \mathbb{N},$$

for some constant $\lambda > 0$.
The parameter $\lambda > 0$ is called the intensity of the Poisson process $(N_t)_{t\in\mathbb{R}_+}$ and can be found as

$$\lambda := \lim_{h\to 0}\frac{1}{h}\mathbb{P}(N_h = 1)$$

There are other two important properties of Poisson process: the superposition property and the thinning property. The superposition property said that a sum of independent Poisson process is again a Poisson process. Hence, let $(N_t^1)_{t\geq 0}$ and $(N_t^2)_{t\geq 0}$ are two independent Poisson processes with intensities λ_1, λ_2, then $(N_t^1 + N_t^2)_{t\geq 0}$ is a Poisson process with intensity $\lambda_1 + \lambda_2$ [3]. The other property define a new process X_t by "thinning" N_t, which is a Poisson process with intensity λ. In particular, it takes all the jump events $(\tau_n, n \geq 1)$ corresponding to N and it keeps them with probability $0 < p < 1$ or delete them with probability $1-p$, independently from each other. Therefore, we can collect and order all the points which have not been deleted: $\tau_1', \ldots, \tau_n', \ldots$ and we can define the new process as:

$$X_t = \sum_{n\geq 1}\mathbf{1}_{\{\tau_n'\geq t\}}$$

Then the new process X is still a Poisson process but it has intensity equal to $p\lambda$ [4]. In other words, if the arrival τ_n of each event in the Poisson process N has probability p, independently from event to event, then the process of joint events thus obtained is again a Poisson process whose intensity is equal to the intensity of N but it is decreased by the marking probability: $\lambda_X = p\lambda$.

2.1.2 Compensated Poisson Processes

The compensated Poisson process define the "centered" version of the Poisson process N_t by

$$\tilde{N}_t = N_t - \lambda t.$$

where λt is the expected value of the Poisson process[5]. $(\tilde{N}_t)$ has centered increments because it has the expected value equal to zero. Moreover, $(\tilde{N})$ follows a centered version of the Poisson law with characteristic function:

$$\psi_{\tilde{N}_t}(u) = exp[\lambda t(e^{iu} - 1 - iu)]$$

In addition, the compensated Poisson process $(N_t - \lambda t)_{t\in\mathbb{R}_+}$ has independent increments and we can show that:

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

so $(\tilde{N}_t)$ is a martingale with respect to it's generated filtration $\mathcal{F}_t$ ($\mathcal{F}_t := \sigma(N_s : s \in [0,t])$, $t \in \mathbb{R}_+$):

$$E[\tilde{N}_t|\tilde{N}_s] = \tilde{N}_s \qquad \forall t > s$$

proof

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

[3] See the appendix A.1 for the proof.
[4] See the appendix A.1 for the proof.
[5] The expected value is computed in the appendix A.1.

□

$(\tilde{N}_t)_{t\geq 0}$ is called a compensated Poisson process and $(\lambda t)_{t\geq 0}$ is called the compensator of $(N_t)_{t\geq 0}$ and it is the quantity which has to be subtracted from N_t in order to obtain a martingale. Moreover, the compensated Poisson process is no longer integer valued because it is not a counting process unlike the Poisson process.
The rescaled version of the compensated Poisson process, i.e. $\frac{\tilde{N}_t}{\lambda}$, has the same first two moments as a standard Brownian motion:

$$E\left[\frac{\tilde{N}_t}{\lambda}\right] = 0 \qquad Var\left[\frac{\tilde{N}_t}{\lambda}\right] = t$$

Moreover, when the intensity of the jumps increases the interpolated compensated Poisson process converges in distribution to a Wiener process:

$$\left(\frac{\tilde{N}_t}{\lambda}\right)_{t\in[0,T]} \overset{\lambda\to\infty}{\Longrightarrow} (W_t)_{t\in[0,T]}$$

This result is a consequence of the Donsker invariance principle[6] and it can be seen as a "functional" central limit theorem.

2.1.3 Compound Poisson Process

The Poisson process itself appears to be too limited to develop realistic asset prices model because the assumption that the jumps size are always equal to 1 is too restrictive, but it can be used as building block to built richer models. Therefore, there is some interest in considering jump processes that can have random jump sizes.

Definition 2.2 [Compound Poisson Process] *The compound Poisson process with jump intensity λ and jump size distribution μ is a stochastic process $(X_t)_{t\geq 0}$ defined by:*

$$X_t = \sum_{i=1}^{N_t} Y_i,$$

where $(Y_i)_{i\geq 1}$ is a sequence of independent random variable with law μ and N_t is a Poisson process with intensity λ independent from $(Y_i)_{i\geq 1}$.

This definition means that a compound Poisson process is a piecewise constant process which jumps at jump times of a standard Poisson process and whose jump size are i.i.d random variables with a given law.
Proposition 2.1 (Characteristic function of the compound Poisson process) (proposition 3.4 in [2]) *Let $(X_t)_{t\geq 0}$ be a compound Poisson process with jump intensity λ and jump size distribution μ. Then X is a piecewise constant Lévy process and its characteristic function is given by:*

$$E[e^{iuX_t}] = exp\left\{\lambda t \int_{-\infty}^{\infty} (e^{iux} - 1)\mu(dx)\right\}. \tag{2.1}$$

A detailed proof can be found in chapter 3 of "Financial Modeling with Jump Process" written by Cont and Tankov.

Let X_{t^-} denote the left limit, i.e. $X_{t^-} := \lim_{s\nearrow t} X_s$ with $t > 0$, then we note that the jump size ($\Delta X_t := X_t - X_{t^-}$,with $t \in \mathbb{R}_+$) of $(X_t)_{t\in\mathbb{R}_+}$ at time t is equal to:

$$\Delta X_t = Y_{N_t}\Delta N_t \qquad t \in \mathbb{R}_+$$

where ΔN_t ($:= N_t - N_{t^-} \in [0,1]$ and $t \in \mathbb{R}_+$) denotes the jump size of the standard Poisson process N_t and N_{t^-} is the left limit.
We know that the n ($N_T = n$) jump sizes of $(X_t)_{t\in\mathbb{R}_+}$ on $[0,T]$ are independent random variables which are distributed on $\mathbb{R}$ according to $\nu(dx)$. Therefore, we can compute the

[6]The Donsker invariance principle is described in the appendix A.2.

moment generating function of the increments $X_T - X_t$ with the following proposition:

Proposition 2.2 *For any $t \in [0,T]$ we have:*

$$E[exp(\alpha(X_T - X_t)] = exp\left(\lambda(T-t)\int_{-\infty}^{\infty}(e^{\alpha x}-1)\nu(dx)\right) \qquad \alpha \in \mathbb{R}.$$

Proof:

$$\begin{aligned}
E[exp(\alpha(X_T - X_t)] &= E\left[exp\left(\alpha\sum_{i=N_t+1}^{N_T} Y_i\right)\right] = E\left[exp\left(\alpha\sum_{i=1}^{N_T-N_t} Y_i\right)\right] \\
&= \sum_{n=0}^{\infty} E\left[exp\left(\alpha\sum_{i=1}^{n} Y_i\right) \mid N_T - N_t = n\right]\mathbb{P}(N_T - N_t = n) \\
&= e^{-\lambda(T-t)}\sum_{n=0}^{\infty}(T-t)^n\frac{\lambda^n}{n!}E\left[exp\left(\alpha\sum_{i=1}^{n} Y_i\right)\right] \\
&= e^{-\lambda(T-t)}\sum_{n=0}^{\infty}(T-t)^n\frac{\lambda^n}{n!}\prod_{i=1}^{n}E[exp(\alpha Y_i)] \\
&= e^{-\lambda(T-t)}\sum_{n=0}^{\infty}(T-t)^n\frac{\lambda^n}{n!}(E[exp(\alpha Y_i)])^n \\
&= exp(\lambda(T-t)(E[exp(\alpha Y)]-1)) \\
&= exp\left(\lambda(T-t)\int_{-\infty}^{\infty}e^{\alpha y}\nu(dy) - \lambda(T-t)\int_{-\infty}^{\infty}\nu(dy)\right) \\
&= exp\left(\lambda(T-t)\int_{-\infty}^{\infty}(e^{\alpha y}-1)\nu(dy)\right)
\end{aligned}$$

Since the probability distribution $\nu(dy)$ of Y satisfies:

$$E[exp(\alpha Y)] = \int_{-\infty}^{\infty}e^{\alpha y}\nu(dy) \text{ and } \int_{-\infty}^{\infty}\nu(dy) = 1$$

□

Now, we can compute the expectation of X_t, for fixed t, as the product of the mean number of jump times ($E[N_t] = \lambda t$) and the mean jump size ($E[Y]$). This is equal to:

$$E[X_t] = \frac{\partial}{\partial\alpha}E[e^{\alpha X_t}]_{|\alpha=0} = \lambda t\int_{-\infty}^{\infty}x\nu(dx) = E[N_t]E[Y] = \lambda t E[Y]$$

The equation above make the assumption that the moment generation function takes finite values for all α in a certain neighborhood $(-\varepsilon,\varepsilon)$ of 0 because so it is possible to exchange the differentiation and the expectation operators. On the other hand, the variance is equal to:

$$Var(X_t) = \lambda t\int_{-\infty}^{\infty}x^2\nu(dx) = \lambda t E[|Y|^2] = E[N_t]E[|Y|^2]$$

Moreover, the compound Poisson process has independent increments if for any finite sequence of times $t_0 < t_1 < \cdots < t_n$, the increments:

$$X_{t_1} - X_{t_0}, X_{t_2} - X_{t_1}, \ldots, X_{t_n} - X_{t_{n-1}}$$

are mutually independent random variables.
By construction, the compound Poisson processes only have a finite number of jumps on any interval, therefore, they belong to the family of Lévy processes which may have an infinite number of jumps on any finite time interval.

2.1.4 Poisson Random Measures

The definition of the Poisson random measure is a key point for the theory of Lévy processes, which are described in the next section of this chapter.
Definition 2.3 [Random measure] *Let $(\Omega, \mathbb{P}, \mathcal{F})$ be a probability space and let $(E, \mathcal{E})$ be a measurable space. Then $M : \Omega \times \mathcal{E} \to \mathbb{R}$ is a random measure if:*

- *for every $\omega \in \Omega$, $M(\omega, \cdot)$ is a measure on $\mathcal{E}$;*
- *for every $A \in \mathcal{E}$, $M(\cdot, A)$ is measurable.*

We can express a Poisson process in terms of the random measure M in the following way:

$$N_t(\omega) = M(\omega, [0,t]) = \int_{[0,t]} M(\omega, ds)$$

where M is called the random jump measure associated to the Poisson process N. Furthermore, the properties of the Poisson process must be translated to fit the measure M. Therefore, we have the following properties for disjoint intervals $[t_1, t_1'], \ldots, [t_n, t_n']$:

1) $M([t_n, t_n'])$ is the number of jumps of the Poisson process in $[t_n, t_n']$: it is a Poisson random variable with parameter $\lambda(t_k' - t_k)$;
2) for two disjoint intervals $j \neq k$, $M([t_j, t_j'])$ ans $M([t_k, t_k'])$ are independent random variables;
3) for any measurable set A, $M(A)$ follows a Poisson distribution with parameter $\lambda|A|$ where $|A| = \int_A dx$ is the Lebesgue measure of A.

We can give another interpretation to the random measure M which is the "derivative" of the Poisson process. Hence, its derivative (in the sense of distributions) is a positive measure because each trajectory $t \mapsto N_t(\omega)$ of a Poisson process is an increasing step function. In fact, it is simply the superposition of Dirac masses located at the jump times:

$$\tfrac{\partial}{\partial t} N_t(\omega) = M(\omega, [0,t]) \text{ where } M = \sum_{i \geq 1} \delta_{T_i(\omega)}$$

Definition 2.4 [Radon measure] (definition 2.2 in [2]) *Let $E \subset \mathbb{R}^d$. A Radon measure on $(E, \mathcal{B})$ is a measure μ such that for every compact measurable set $B \in \mathcal{B}$, $\mu(B) < \infty$*

Definition 2.5 [Poisson random measure] (definition 2.18 in [2]) *Let $(\Omega, \mathbb{P}, \mathcal{F})$ be a probability space, $(E, \mathcal{E})$ be a measurable space and μ a measure on $(E, \mathcal{E})$. Then*

$$\begin{aligned} M : \Omega \times \mathcal{E} &\to \mathbb{R} \\ (\omega, A) &\mapsto M(\omega, A), \end{aligned}$$

is a Poisson random measure with intensity μ if:

1. *for (almost all) $\omega \in \Omega$, $M(\omega, \cdot)$ is an integer-valued Radon measure on E: for any bounded measurable $A \subset E$, $M(A) < \infty$ is an integer valued random variable;*
2. *for all $A \in \mathcal{E}$ with $\mu(A) < \infty$, $M(A)$ follows the Poisson law with parameter $E[M(A)] = \mu(A)$:*

$$\forall k \in \mathbb{N}, \qquad \mathbb{P}(M(A) = k) = e^{-\mu(A)} \tfrac{(\mu(A))^k}{k!}$$

3. *for any disjoint measurable sets $A_1, \ldots, A_n \in \mathcal{E}$, the variables $M(A_1), \ldots, M(A_n)$ are independent.*

In particular, the following preposition show as the Poisson random measure can be constructed as the counting measure of randomly scattered points.

Proposition 2.3 (Construction of Poisson random measures) (proposition 2.14 in [2]) *Let μ be a σ-finite measure on a measurable subset E of $\mathbb{R}^d$. Then, there exists a Poisson random measure M on E with intensity μ.*

Proof:

1. Assume that $\mu(E) < \infty$. Let $(X_i)_{i\geq 1}$ be a sequence of independent random variables such that $\mathbb{P}(X_i \in A) = \frac{\mu(A)}{\mu(E)}$, $\forall i$ and $\forall A \in \mathcal{B}(E)$, and let $M(E)$ be a Poisson random variables with intensity $\mu(E)$ independent from $(X_i)_{i\geq 1}$. Then, it is easy to see that the random measure M defined by:

$$M(A) := \sum_{i=1}^{M(E)} \mathbf{1}_A(X_i), \quad \forall A \in \mathcal{B}(E)$$

is a Poisson random measure on E with intensity μ.

2. Assume that $\mu(E) = \infty$. Then, we choose a sequence of disjoint measurable sets $(E_i)_{i\geq 1}$ such that $\mu(E_i) < \infty$, $\forall i$ and $\bigcup_i E_i = E$. We can built a Poisson random measure M_i on each E_i as described above and define:

$$M(A) := \sum_{i=1}^{\infty} M_i(A), \qquad \forall A \in \mathcal{B}(E)$$

$\square$

The following proposition is useful to study the convergence of Poisson random measures:

Proposition 2.4 (Convergence of Poisson random measures) (proposition 2.15 in [2]) *Let $(M_n)_{n\geq 1}$ be a sequance of Poisson random measure on $E \subset \mathbb{R}^d$ with intensities $(\mu_n)_{n\geq 1}$. Then, $(M_n)_{n\geq 1}$ converges in distribution if and only if the intensities (μ_n) converge to a Radon measure μ. Hence, $M_n \Rightarrow M$, where M is a Poisson random measure with intensity μ.*

In the same way as we have defined the compensated Poisson process, we can construct the compensated Poisson random measure $\tilde{M}$ by subtracting from M its intensity measure:

$$\tilde{M}(A) = M(A) - \mu(A)$$

Moreover, from the definition of Poisson random measures, we can note that for disjoint compact sets $(A_1, \ldots, A_n \in \mathcal{E})$ the variables $\tilde{M}(A_1), \ldots, \tilde{M}(A_n)$ are independent and have the following two moments:

$$E[\tilde{M}(A_i)] = 0 \qquad Var[\tilde{M}(A_i)] = \mu(A_i).$$

Corollary 2.1 (Exponential formula) *Let M be a Poisson Random measure on $(E, \mathcal{E})$ with intensity μ, $B \in \mathcal{E}$ and let f be a measurable function with $\int_B |e^{f(x)} - 1|\mu(dx) < \infty$. Then:*

$$E\left[e^{\int_B f(x)M(dx)}\right] = exp\left[\int_B (e^{f(x)} - 1)\mu(dx)\right]$$

Definition 2.6 [Jump measure] *Let X be a $\mathbb{R}^d-$valued càdlàg process. The jump measure of X is a random measure on $\mathcal{B}([0, \infty) \times \mathbb{R}^d)$ defined by*

$$J_X(A) = \#\left\{t : \Delta X_t \neq 0 \quad and \quad (t, \Delta X_t) \in A\right\}. \tag{2.2}$$

This definition means that the jump measure of a set of the form $[s, t] \times A$ counts the number of jumps of X between s and t such that their amplitude belongs to A. In other words, J_X contains all the information about the discontinuities, i.e. jumps, of the process X. It tells us when the jumps occur and how big they are. Therefore, J_X does not tell us anything about the continuous component of X, which has continuous sample path if and only if $J_X = 0$ almost surely. This means that there are no jumps in the process.

For a counting process, since the jumps size is always equal to 1, the jump measure can be seen as a random measure on $[0, \infty)$.
Proposition 2.4 *Let X be a Poisson process with intensity λ. Then, J_X is a Poisson random measure on $[0, \infty)$ with intensity $\lambda \times dt$.*
Proposition 2.5 (Jump measure of a compound Poisson process) (proposition 3.5 in [2]) *Let $(X_t)_{t\geq 0}$ be a compound Poisson process with intensity λ and jump size distribution f. Its jump measure J_X is a Poisson random measure on $\mathbb{R}^d \times [0, \infty)$ with intensity measure $\mu(dx \times dt) = \nu(dx)dt = \lambda f(dx)dt$.*
A detailed proof can be found in chapter 3 of "Financial Modelling with Jump Process" written by Cont and Tankov.
This proposition implies that every compound Poisson process can be represented in the following form:

$$X_t = \sum_{s\in[0,t]} \Delta X_s = \int_{[0,t]\times\mathbb{R}^d} x J_X(ds \times dx)$$

where J_X is a Poisson random measure with intensity measure $\nu(dx)dt$. In this equation, we have rewritten the process X as the sum of its jumps and since it is a compound Poisson process, it has almost surely a finite number of jumps in the interval $[0, t]$. Moreover, the stochastic integral in the equation is a finite sum, hence there are no convergence problems.

2.2 Lévy Processes

2.2.1 Definition and Properties

Definition 2.7 [Lévy process] (definition 3.1 in [2]) *A càdlàg stochastic process $(X_t)_{t\geq 0}$ on $(\Omega, \mathcal{F}, \mathbb{P})$ with values in $\mathbb{R}^d$ such that $X_0 = 0$ is called a Lèvy process if it possesses the following properties:*

1) Independent increments: for every increasing sequence of times $t_0, \dots, t_n$, the random variables $X_{t_0}, X_{t_1} - X_{t_0}, \dots, X_{t_n} - X_{t_{n-1}}$ are independent;

2) Stationary increments: the law of $X_{t+h} - X_t$ does not depend on t;

3) Stochastic continuity: $\forall \varepsilon > 0, \quad \lim_{h\to 0} \mathbb{P}(|X_{t+h} - X_t| \geq \varepsilon) = 0.$

The last properties does not imply that the sample path are continuous as it is in the Poisson process. The stochastic continuity serves to exclude processes with jump at fixed times, which can be regarded as "calendar effects". Therefore, it means that for given time t, the probability of seeing a jump at t is zero, discontinuities occur at random times:

$$\forall t, \quad \mathbb{P}(X_{t^-} = X_t) = 1.$$

The simplest Lévy process is the linear process is the linear drift, a deterministic process. Brownian motion is the only (non-deterministic) Lévy process with continuous sample paths. Other examples of Lévy processes are the Poisson and the compound Poisson processes. Moreover, the sum of a linear drift, a Brownian motion and a compound Poisson process is again a Lévy process and it is called a "jump-diffusion process."

We say that a probability distribution F on $\mathbb{R}^d$ is infinitely divisible if for any integer $n \geq 2$, there exists n i.i.d. random variables $Y_1, \dots, Y_n$ such that $Y_1 + \dots + Y_n$ has distribution F. Hence, if X is a Lévy process, the distribution of X_t is infinitely divisible for any $t > 0$. Therefore, the distribution of increments of a Lévy process has to be infinitely divisible, this puts a constraint on the possible choices of distribution for X_t. Gaussian distribution, Gamma distribution and Poisson distribution are common examples of infinitely divisible laws. A random variable having any of these distributions can be decomposed into a sum of n i.i.d. parts having the same distribution but with modified parameters.
Proposition 2.6 (proposition 3.1 in [2]) *Let $(X_t)_{t\geq 0}$ be a Lévy process. Then, X_t has*

an infinitely divisible distribution for every t. Conversely, if F is an infinitely divisible distribution then there exists a Lévy process (X_t) such that the distribution of X_1 is a given by F.

Proposition 2.7 (Characteristic function of a Lévy process) (proposition 3.2 in [2]) *Let $(X_t)_{t\geq 0}$ be a Lévy process on $\mathbb{R}^d$. There exists a continuous function $\psi : \mathbb{R}^d \mapsto \mathbb{R}$ called the characteristic exponent of X, such that:*

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

Where ψ is the cumulant generating function of X_1. The cumulant generating function $\psi(t)$ is the natural logarithm of the moment generating function:

$$\psi(t) = \log E[e^{tX}]$$

The law of X_t is determined by the knowledge of the law of X_1 because the cumulant generating function of X_t varies linearly in t. Therefore, the only degree of freedom that we have to specify is the distribution of X_t for a single time ($t = 1$).

The proposition regarding the Jump measure of a compound Poisson process can be used to define the Lévy measure for all the Lévy process. Therefore, we give the following definition:

Definition 2.8 [Lévy measure] (definition 3.4 in [2]) *Let $(X_t)_{t\geq 0}$ be a Lévy process on $\mathbb{R}^d$. The measure ν on $\mathbb{R}^d$ defined by:*

$$\nu(A) = E[\#\{t \in [0,1] : \Delta X_t \neq 0,\, \Delta X_t \in A\}], \quad A \in \mathcal{B}(\mathbb{R}^d)$$

is called the Lévy measure of X. $\nu(A)$ is the expected number, per unit time, of jumps whose size belongs to A.

Lévy processes are basically processes with jumps. In fact, it can be shown that any Lévy process which has continuous trajectories is a Brownian motion with drift a.s.

Proposition 2.8 *Let X be a continuous Lévy process. Then, there exists $\gamma \in \mathbb{R}^d$ and a symmetric positive definite matrix A such that:*

$$X_t = \gamma t + W_t$$

where W is a Brownian motion with covariance matrix A.

This preposition is important for understand the Lévy processes. Hence, we give a proof for the one-dimensional case.

Proof

We need to show that X_1 has Gaussian law because the rest will follow from the stationary and independence of increments.

a). Let $\xi_n^k := X_{\frac{k}{n}} - X_{\frac{k-1}{n}}$ and $b_n = \mathbb{P}(|\xi_n^1| > \epsilon)$. The continuity of X implies that:

$$\lim_{n\to\infty} \mathbb{P}(\sup_k |\xi_n^k| > \epsilon) = 0, \quad \forall \epsilon$$

Since

$$\mathbb{P}(\sup_k |\xi_n^k| > \epsilon) = 1 - [1 - \mathbb{P}(|\xi_n^1| > \epsilon)]^n$$

We find that $\lim\limits_{n\to\infty} (1 - b_n)^n = 1$, from which it follows that $\lim\limits_{n\to\infty} n \log(1 - b_n) = 0$. But $n\log(1 - b_n) \leq -nb_n \leq 0$. Therefore, we have:

$$\lim_{n\to\infty} n\mathbb{P}(|X_{\frac{1}{2}}| > \epsilon) = 0. \tag{2.3}$$

b). We use the property of the independence and stationary of increments to show that:

$$\lim_{n\to\infty} nE[\cos X_{\frac{1}{n}} - 1] = \frac{1}{2}\big\{\log E[e^{iX_1}] + \log E[e^{-iX_1}]\big\} := -A; \tag{2.4}$$

$$\lim_{n\to\infty} nE[\sin X_{\frac{1}{n}}] = \frac{1}{2i}\big\{\log E[e^{iX_1}] - \log E[e^{-iX_1}]\big\} := \gamma. \tag{2.5}$$

The equation (2.3) and (2.4) allow to prove that for every function f such that $f(x) = o(|x|^2)$ in a neighborhood of 0, $\lim_{n\to\infty} nE[f(X_{\frac{1}{n}})] = 0$, which implies that $\epsilon > 0$:

$$\lim_{n\to\infty} nE[X_{\frac{1}{n}} \mathbf{1}_{|X_{\frac{1}{n}}|\le\epsilon}] = \gamma,$$
$$\lim_{n\to\infty} nE[X^2_{\frac{1}{n}} \mathbf{1}_{|X_{\frac{1}{n}}|\le\epsilon}] = A,$$
$$\lim_{n\to\infty} nE[|X_{\frac{1}{n}}|^3 \mathbf{1}_{|X_{\frac{1}{n}}|\le\epsilon}] = 0.$$

c). Putting together the different equations, we find:

$$\begin{aligned}
\log E[e^{iuX_1}] &= n \log E[e^{iuX_{\frac{1}{n}}} \mathbf{1}_{X_{\frac{1}{n}}\le\epsilon}] + o(1) \\
&= n \log \left\{ 1 + iuE[X_{\frac{1}{n}} \mathbf{1}_{X_{\frac{1}{n}}\le\epsilon}] - \frac{u^2}{2} E[X^2_{\frac{1}{n}} \mathbf{1}_{X_{\frac{1}{n}}\le\epsilon}] + o(\frac{1}{n}) \right\} + o(1) \\
&= iu\gamma - \frac{Au^2}{2} + o(1) \xrightarrow[n\to\infty]{} iu\gamma - \frac{Au^2}{2}
\end{aligned}$$

where $o(1)$ denotes a quantity which tends to 0 as $n \to \infty$.

□

Now, consider a Brownian motion with drift $\gamma t + W_t$, independent from X^0, the sum $X_t = X^0 + \gamma t + W_t$ defines another Lévy process, which can be decomposed as:

$$X_t = \gamma t + W_t + \sum_{s\in[0,t]} \Delta S_s = \gamma t + W_t + \int_{[0,t]\times\mathbb{R}^d} x J_X(ds \times dx)$$

where J_X is a Poisson random measure on $[0,\infty) \times \mathbb{R}^d$ with intensity $\nu(dx)dt$, where ν is a finite measure defined by:

$$\nu(A) = E[\#\left\{t \in [0,1] : \Delta X^0_t \neq 0,\ \Delta X^0_t \in A\right\}], \quad A \in \mathcal{B}(\mathbb{R}^d).$$

For every Lévy process X_t we can define its Lévy measure ν as above. For any compact set A such that $0 \notin A$, $\nu(A)$ is still finite. Otherwise, the process would have an infinite number of jumps of finite size on $[0,T]$, which contradicts the càdlàg property. Hence, ν defines a Radon measure on $\mathbb{R}^d \setminus \{0\}$. On the contrary, ν is not necessarily a finite measure, the above restriction still allows it to blow up at zero and X may have an infinite number of small jumps on $[0,T]$. In this case, the sum of the jumps becomes an infinite series and its convergence imposes some conditions on the measure ν, under which which we obtain a decomposition of X given by the following proposition.

Proposition 2.9 (Lévy-Ito decomposition) (proposition 3.7 in [2]) *Let $(X_t)_{t\ge0}$ be a Lèvy process on $\mathbb{R}^d$ and ν its Lèvy measure. Then:*

- *the Lévy measure ν satisfies the integrability condition:*

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

- *the jump measure J_X of X is a Poisson random measure on $[0,\infty) \times \mathbb{R}^d$ with intensity $dt \times \nu = \nu(dx)dt$;*

- *there exists $\gamma \in \mathbb{R}^d$ and a d-dimensional Brownian motion $(B_t)_{t\ge0}$ with covariance matrix A such that:*

$$\begin{aligned}
X_t &= \gamma t + B_t + N_t + M_t, \quad \textit{where} \qquad (2.6)\\
N_t &= \int_{|x|>1, s\in[0,t]} x J_X(ds \times dx) \\
M_t &= \int_{0<|x|\le1, s\in[0,t]} x\{J_X(ds \times dx) - \nu(dx)ds\} \\
&\equiv \int_{0<|x|\le1, s\in[0,t]} x\tilde{J}_X(ds \times dx)
\end{aligned}$$

The three terms in (2.6) *are independent and the convergence in the last term is almost sure and uniform in* t *on* $[0, T]$.
A detailed proof can be found in chapter 3 of "Financial Modeling with Jump Process" written by Cont and Tankov.

The Lévy-Ito decomposition say that for every Lévy process there exist a vector γ (drift), a positive definite matrix A and a positive measure ν that uniquely determine its distribution. We call the triplet (A, ν, γ) characteristic triplet or Lévy triplet of the process X_t.
The term in the equation (2.6) have the following meaning: $\gamma t + B_t$ is a continuous Gaussian Lévy process and every Gaussian Lévy process is continuous and it can be written in this form Moreover, it can be described by the drift γ and the covariance matrix of the Brownian motion A. The other two terms, $N_t + M_t$, are discontinuous processes incorporating the jumps of X_t and they are described by the Lévy measure ν. The integrability condition can be also written as:

$$\nu \textit{ is a Radon measure on } \mathbb{R}^d \setminus \{0\} \textit{ and verify: } \int_{|x|\leq 1} |x|^2 \nu(dx) < \infty, \quad \int_{|x|\geq 1} \nu(dx) < \infty$$

The integral $\displaystyle\int_{|x|\geq 1} \nu(dx) < \infty$ means that X has a finite number of jumps with absolute value greater or equal to 1. So, the sum:

$$N_t = \sum_{0\leq s\leq t}^{|\Delta X_s|\geq 1} \Delta X_s$$

contains a finite number of terms and N_t is a compound Poisson process almost surely. ν can have a singularity in zero, which means that there can be infinitely many small jumps and that their sum does not necessarily converge. In order to obtain convergence, we replace the jump integral by its compensated version, which is a martingale, and it is equal to:

$$M_t = \textstyle\int_{0<|x|\leq 1, s\in[0,t]} x \tilde{J}_X(ds \times dx).$$

M_t can be seen as an infinite superposition of independent compensated.
An important result of the Lévy-Ito decomposition is that every Lévy process is a combination of a Brownian motion with drift and a possibile infinite sum of independent compound Poisson process. Therefore, every Lévy process can be approximated by a jump-diffusion process, which is equal to the sum of Brownian motion with drift and a compound Poisson process.

Proposition 2.10 *Let* (X_t, Y_t) *be a Lévy process such that* Y *is a piecewise constant and* $\Delta X_t \Delta Y_t = 0$ *for all* t *a.s. Then,* X *and* Y *are independent.*
Proof
It is enough show that X_1 and Y_1 are independent due to the independence and the stationary of increments. Let $M_t = \frac{e^{iuX_t}}{E[e^{iuX_t}]}$ and $N_t = \frac{e^{iuY_t}}{E[e^{iuY_t}]}$. Then, M and N are martingales on $[0, 1]$. From the independence and stationary of increments, we know that for every Lévy process Z:

$$E[e^{iuZ_t}] = E[e^{iuZ_1}]^t \quad \textit{and} \quad E[e^{iuZ_1}] \neq 0, \forall u.$$

This means that M is bounded. Since (N_t) is a Lévy process and a counting process then (N_t) is a Poisson process, therefore, the number of jumps of Y on $[0, 1]$ is a Poisson random variable. Hence, N has integrable variation on this interval and by the martingale property and by the dominated convergence, we find:

$$\begin{aligned} E[M_1 N_1] - 1 &= E\left[\sum_{i=1}^{n} (M_{\frac{i}{n}} - M_{\frac{(i-1)}{n}})(N_{\frac{i}{n}} - N_{\frac{(i-1)}{n}})\right] \\ &\to E\left[\sum_{0\leq t\leq 1} \Delta M_t \Delta N_t\right] \end{aligned}$$

which implies that $E[e^{iuX_1 + ivY_1}] = E[e^{iuX_1}]E[e^{ivY_1}]$.

□

The following theorem give to us the second fundamental result of the structure of the path of Poisson process and it announces the expression of the characteristic function of a Lévy process in terms of its characteristic triplet (A, ν, γ) :

Theorem 2.2 [Lévy-Khinchin representation] (theorem3.1 in [2]) *Let $(X_t)_{t\geq 0}$ be a Lévy process on $\mathbb{R}^d$ with characteristic triplet (A, ν, γ). Then:*

$$\begin{aligned} &E[e^{i\langle u, X_t\rangle}] = e^{t\psi(u)}, \quad u \in \mathbb{R}^d && (2.7)\\ &\textit{with} \quad \psi(u) = i\langle \gamma, u\rangle - \frac{1}{2}\langle u, Au\rangle + \int_{\mathbb{R}^d} (e^{i\langle u,x\rangle} - 1 - i\langle u, x\rangle \mathbf{1}_{|x|\leq 1})\nu(dx). \end{aligned}$$

A detailed proof can be found in chapter 3 of "Financial Modeling with Jump Process" written by Cont and Tankov.

For real-valued Lévy processes, the equation (2.7) becomes:

$$\begin{aligned} &E[e^{iuX_t}] = e^{t\psi(u)}, \quad u \in \mathbb{R}^d && (2.8)\\ &\textit{with} \quad \psi(u) = i\gamma u - \frac{1}{2}Au^2 + \int_{\mathbb{R}^d} (e^{iux} - 1 - iux\mathbf{1}_{|x|\leq 1})\nu(dx). \end{aligned}$$

When $\nu(\mathbb{R}^d) = \infty$, we are in the infinite activity case and the set of jumps times of every trajectory of the Lévy process is countably infinite and dense in $[0, \infty)$. The countably follows directly from the fact that the path are càdlàg. The following theorem gives the characteristic function of infinitely divisible distributions:

Theorem 2.3 [Characteristic function of infinitely divisible distributions] (theorem3.2 in [2]) *Let F be an infinitely divisible distribution on $\mathbb{R}^d$. Its characteristic function can be represented as:*

$$\begin{aligned} &\Phi_F(u) = e^{t\psi(u)}, \quad u \in \mathbb{R}^d\\ &\psi(u) = i\langle \gamma, u\rangle - \frac{1}{2}\langle u, Au\rangle + \int_{\mathbb{R}^d} (e^{i\langle u,x\rangle} - 1 - i\langle u, x\rangle \mathbf{1}_{|x|\leq 1})\nu(dx) && (2.9) \end{aligned}$$

where A is a symmetric positive $n \times n$ matrix, $\gamma \in \mathbb{R}^d$ and ν is a positive Radon measure on $\mathbb{R}^d \setminus \{0\}$ verifying:

$$\int_{|x|\leq 1} |x|^2\nu(dx) < \infty, \qquad \int_{|x|\geq 1} \nu(dx) < \infty$$

where ν is called the Lévy measure of the distribution F.

This theorem imply that since X has stationary and independent increments, we have that $E\left[e^{i\langle u,X_t\rangle}\right] = \left\{E\left[e^{i\langle u,X_1\rangle}\right]\right\}^t$, $\forall t \in \mathbb{R}$ and by the right continuity of X, $\forall t$. Moreover, the exponent (8) is called the Lévy exponent of the Lévy process $(X_t)_{t\geq 0}$. Note that the first term is the Lévy exponent of the Lévy process γt. The second term, is the Lévy exponent of the Lévy process ΣB_t, where B_t are d-independent Brownian processes and Σ is a $d \times d$ lower triangular matrix in the Cholesky decomposition $A = \Sigma\Sigma^T$. The last term in the Lévy exponent can be decomposed into two terms:

$$\begin{aligned} \psi^{cp}(u) &= \int_{|x|>1} \left(e^{i\langle u,x\rangle} - 1\right)\nu(dx),\\ \psi^{lccp}(u) &= \int_{|x|\leq 1} \left(e^{i\langle u,x\rangle} - 1 - i\langle u, x\rangle\right)\nu(dx). \end{aligned}$$

The first equation above is the Lévy exponent of a compound Poisson process (indicated with "cp") X^{cp} with Lévy measure: $\nu_1(dx) := \mathbf{1}_{|x|>1}\nu(dx)$. The second term corresponds to the limit in distribution of compensated compound Poisson process (indicated with "lccp"). Suppose that $X^{(\varepsilon)}$ is a compound Poisson process with Lévy measure $\nu_\varepsilon(dx) := \mathbf{1}_{\varepsilon<|x|\leq 1}\nu(dx)$, then the process $X_t^\varepsilon - E\left[X_t^\varepsilon\right]$ converges in distribution to a process with characteristic function $exp\left\{t\psi^{lccp}\right\}$. The Lévy-Khinchin representation implies that, in distribution, X is the superposition of four independent Lévy processes:

$$X_t \overset{\mathcal{D}}{\Rightarrow} \underbrace{\gamma t}_{Drift} + \underbrace{\Sigma B_t}_{Brownian\,component}$$

$$+ \underbrace{X_t^{cp}}_{Compounded\,Poisson} + \underbrace{\lim_{\varepsilon\to 0}\left(X_t^\varepsilon - E\left[X_t^\varepsilon\right]\right)}_{Limit\,of\,compensated\,compounded\,Poisson}.$$

The condition in the theorem on ν of the characteristic function of infinitely divisible distribution guarantees that the X^{cp} is indeed well defined and the compensated compound Poisson process converges in distribution.

2.2.2 Pathwise properties

We know that almost all the trajectories of a Lévy process are piecewise constant if and only if it is of compound Poisson type. Combining this with the characteristic function of a compound Poisson process (equation (2.1)), we obtain the following proposition:
Proposition 2.11 (proposition 3.8 in [2]) *A Lévy process has piecewise constant trajectories if and only if its characteristic triplet satisfies the following condition:*

- $A = 0,$
- $\displaystyle\int_{\mathbb{R}^d} \nu(dx) < \infty,$
- $\displaystyle\gamma = \int_{|x|\leq 1} x\nu(dx) < \infty$

or, equivalently, if its characteristic exponent is equal to:

$$\psi(u) = \int_{-\infty}^{\infty} (e^{iux} - 1)\nu(dx) \quad with\ \nu(\mathbb{R}) < \infty.$$

The meaning of the condition above is that the Lévy process has covariance matrix of the Brownian motion (A) equal to 0, the drift parameter (γ) must be finite and the second condition told us that the process has a finite number of jumps.
Moreover, a Lévy process is said to be of finite variation[7] if its trajectories are functions of finite variation with probability 1. Therefore, we have the following proposition for finite variation Lévy processes:
Proposition 2.12 (Finite variation Lévy processes) (proposition 3.9 in [2]) *A Lévy process is of finite variation if and only if its characteristic triplet* (A, ν, γ) *satisfies:*

$$A = 0 \quad and \quad \int_{|x|\leq 1} |x|\nu(dx) < \infty.$$

A detailed proof can be found in chapter 3 of "Financial Modeling with Jump Process" written by Cont and Tankov.
The preposition above said that in the finite variation case the Lévy-Ito decomposition and the Lévy-Khinchin representation can be simplified with the following corollary:
Corollary 2.2 (Lévy-Ito decomposition and Lévy-Khinchin representation in the finite variation case) (corollary 3.1 in [2]) *Let* $(X_t)_{t\geq 0}$ *be a Lévy process of finite variation with characteristic triplet given by* $(0, \nu, \gamma)$*. Then, X can be expressed as the sum of its jumps between 0 and t and a linear drift term. So, we find:*

$$X_t = bt + \int_{[0,t]\times\mathbb{R}^d} xJ_X(ds \times dx) = bt + \sum_{s\in[0,t]}^{\Delta X_s \neq 0} \Delta X_s \tag{2.10}$$

[7] Defined in chapter 1.

and its characteristic function can be expressed as:

$$E\left[e^{i\langle u,X_t\rangle}\right] = exp\,t\left\{i\langle b,u\rangle + \int_{\mathbb{R}^d}\left(e^{i\langle u,x\rangle}\right)\nu(dx)\right\}, \qquad (2.11)$$

where $b = \gamma - \int_{|x|\leq 1} x\nu(dx)$ *is such that* $\mathbb{P}\left(\lim_{t\to 0}\frac{X_t}{t} = b\right) = 1.$

We can highlight that the Lévy triplet of X in not given by $(0,\nu,b)$ instead by $(0,\nu,\gamma)$. Indeed, γ is not an intrinsic quantity and depends on the truncation function used in the Lévy-Khinchin representation while bt has an intrinsic interpretation as the continuous part of X. For every bounded measurable function $g:\mathbb{R}^d\to\mathbb{R}$ satisying $g(x) = 1 + o(|x|)$ as $x\to 0$ and $g(x) = o(\frac{1}{|x|})$ as $x\to\infty$, we can write the Lévy-Khinchin representation as:

$$\begin{aligned} &E[e^{i\langle u,X_t\rangle}] = e^{t\psi(u)}, \quad u\in\mathbb{R}^d \\ &\textit{with}\quad \psi(u) = i\langle\gamma^g,u\rangle - \frac{1}{2}\langle u,Au\rangle + \int_{\mathbb{R}^d}\left(e^{i\langle u,x\rangle} - 1 - i\langle u,xg(x)\rangle\right)\nu(dx). \end{aligned}$$

the function g is called the truncation function and the Lévy triplet (A,ν,γ^g) is called the characteristic triplet of X with respect to the truncation function g. Different choices of g do not affect A and ν which are intrinsic parameters of the Lévy process, but γ depends on the choice of truncation function.

Proposition 2.13 *Let* $(X_t)_{t\geq 0}$ *be a Lévy process on* $\mathbb{R}$*. The following conditions are equivalent:*

a. $X_t \geq 0$ *a.s.for some* $t > 0$*;*

b. $X_t \geq 0$ *a.s.for every* $t > 0$*;*

c. Sample path of (X_t) *are almost surely nondecreasing:* $t\geq s \Rightarrow X_t \geq X_s$*;*

d. The characteristic triplet of (X_t) *satisfies* $A = 0$, $\nu((-\infty,0]) = 0$, $\int_0^\infty (x\wedge 1)\nu(dx) < \infty$ *and* $b\geq 0$*. Then,* (X_t) *has no diffusion component, only positive jumps of finite variation and positive drift.*

Proof

Here, we give only a short proof of the equivalence between the condition c and d, for the rest point look in chapter 3 of "Financial Modeling with Jump Process" written by Cont and Tankov.

Since, the trajectories are nondecreasing they are of finite variation, so $A = 0$ and $\int_{-\infty}^\infty (x\wedge 1)\nu(dx) < \infty$. On the other hand, the trajectories are nonincreasing if there will be no negative jumps, therefore $\nu((-\infty,0]) = 0$. If a function is nondecresing then after removing some of its jumps, we obtain another nondecreasing function. When we remove all jumps from a trajectory of X_t, we obtain a deterministic function bt, which must be nondecreasing. This allows to conclude that $b\geq 0$.

□

Increasing Lévy processes are called subordinators because they can be used as time changes for other Lévy process. The following proposition gives an important example of subordinator:

Proposition 2.14 (proposition 3.11 in [2]) *Let* $(X_t)_{t\geq 0}$ *be a Lévy process on* $\mathbb{R}^d$ *and let* $f:\mathbb{R}^d\to[0,\infty)$ *be a positive function such that* $f(x) = o(|x|^2)$ *when* $x\to 0$*. Then, the process* $(S_t)_{t\geq 0}$ *is a subornitator and is defined by:*

$$S_t = \sum_{s\leq t,\Delta X_s\neq 0} f(\Delta X_s).$$

If there exists a Lévy processes without diffusion and with negative jumps, it will satisfy the condition: $\int_0^1 |x|\nu(dx) = \infty$. The above proposition entails that these processes cannot have increasing trajectories. whatever drift coefficient they may have.

2.2.3 Distribution Properties

Let $(X_t)_{t\geq 0}$ be a Lévy process then the distribution of X_t is infinitely divisible and has a characteristic function as in equation (2.7) for any $t > 0$. However, the Lévy process X_t does not always have a density, in fact, if we have a compound Poisson process we find:

$$\mathbb{P}(X_t = 0) = e^{-\lambda t} > 0.$$

Hence, the probability distribution of X_t has an atom at zero for all t. On the other hand, if X_t is not a compound Poisson process, then it has a continuous density.
Proposition 2.15 (Existence of a smooth density) (proposition 3.12 in [2]) *Let X be a real-valued Lévy process with characteristic triplet (σ^2, ν, γ). Then:*

i. If $\sigma > 0$ or $\nu(\mathbb{R}) = \infty$, X_t has a continuous density $p_t(.)$ on $\mathbb{R}^d$;

ii. If the Lévy measure ν verifies: $\exists \beta \in (0,2)$, $\liminf_{\varepsilon\to 0} \varepsilon^{-\beta} \int_{-\varepsilon}^{\varepsilon} |x|^2 d\nu(x) > 0$, then for each $t > 0$, X_t has a smooth density $p_t(.)$ on $\mathbb{R}^d$ such that:

$$p_t(.) \in C^\infty(\mathbb{R}) \, \forall n \geq 1, \quad \tfrac{\partial^n p_t}{\partial x^n}(t,x) \xrightarrow[|x|\to\infty]{} 0.$$

Now, we focus on the relation between probability density function and the Lévy density. In particular, in the compound Poisson process there is a simple relation between probability distribution at time t and the jump size distribution or the Lévy measure. Let $(X_t)_{t\geq 0}$ be a compounded Poisson process with intensity λ and jump size distribution f and $(N_t)_{t\geq 0}$ be the number of jumps of X on $[0,t]$. Then:

$$\begin{aligned}\mathbb{P}(X_t \in A) &= \sum_{n=0}^{\infty} \mathbb{P}(X_t \in A | N_t = n) \frac{e^{-\lambda t}(\lambda t)^n}{n!} \\ &= e^{-\lambda t}\delta_0 + \sum_{n=1}^{\infty} f^{*n}(A) \frac{e^{-\lambda t}(\lambda t)^n}{n!},\end{aligned}$$

where f^{*n} denotes the n-th convolution power of f and δ_0 is the Dirac measure concentrated at 0.[8]
As note before, this probability measure has not a density because $\mathbb{P}(X_t = 0) > 0$. Recall that a Lebesgue measure λ is a measure on $(\mathbb{R}, \mathcal{R})$, satisfying: $\lambda((a,b]) = b - a \quad forall \quad a < b,\ a, b \in \mathbb{R}$. Therefore, if we consider jump size distribution with Lebesgue measure and if it has a density, then the law of X_t is absolutely continuous except at zero. Therefore, the law of X_t can be decomposed as:

$$\mathbb{P}(X_t \in A) = e^{-\lambda t}\mathbf{1}_{0\in A} + \int_A p_t^{ac}(x)dx$$

where

$$p_t^{ac} = \sum_{n=1}^{\infty} f^{*n}(x) \frac{e^{-\lambda t}(\lambda t)^n}{n!} \quad \forall x \neq 0,$$

and we denote the jump size density by $f(x)$ and by p_t^{ac} the density, which is conditioned on the fact that the process has jumped at least once. Then, we can conclude with the following asymptotic relation:

$$\lim_{t\to 0} \frac{1}{t} p_t^{ac}(x) = \lambda f(x) = \nu(x), \quad \forall x \neq 0$$

where $\nu(x)$ is the Lévy density, which describes the small time behavior of the probability density. Moreover, from this relation we can find the small time behavior for expectation of function X_t, given any bounded measurable function of f such that $f(0) = 0$ we have:

[8] The convolution power and the Dirac measure are explain in the appendix A.3 and A.4.

$$\lim_{t\to 0}\frac{1}{t}E[f(X_t)] = \lim_{t\to 0}\frac{1}{t}\int_{\mathbb{R}^d} f(x)p_t(dx) = \int_{\mathbb{R}^d} f(x)\nu(dx).$$

Another important property to say is what is the tail behavior of the distribution of the Lévy process and how its moments are determined by the Lévy measure. Therefore, we consider the following proposition:

Proposition 2.16 (Moments and cumulants of a Lévy process) (proposition 3.13 in [2]) *Let $(X_t)_{t\geq 0}$ be a Lévy process on $\mathbb{R}$ with characteristic triplet (A,ν,γ). The n-th absolute moment of X_t, $E[|X_t|^n]$ is finite for some t or, equivalently, for every $t>0$ if and only if $\int_{|x|\geq 1}|x|^n\nu(dx)<\infty$. In this case, moments of X_t can be computed from its characteristic function by differentiation. Therefore, using the cumulants[9] of X_t we have:*

$$E[X_t] = t\left(\gamma + \int_{|x|\geq 1} x\nu(dx)\right)$$

$$c_2(X_t) = Var(X_t) = t\left(A + \int_{-\infty}^{\infty} x^2\nu(dx)\right)$$

$$c_n(X_t) = t\int_{-\infty}^{\infty} x^n\nu(dx), \quad for\, n\geq 3.$$

We can note that all the infinitely distribution are leptokurtic[10] since $c_4(X_t)>0$ and also the cumulants of the distributions of X_t increase linearly with t. In particular, the kurtosis and skewness of the increments $X_{t+\Delta}-X_t$ or X_Δ are given by:

$$s(X_\Delta) = \frac{c_3(X)}{\sqrt{c_2(X)^3}} = \frac{s(X_1)}{\sqrt{\Delta}}, \qquad \kappa(X_\Delta) = \frac{c_4(X_\Delta)}{c_2(X_\Delta)^2} = \frac{\kappa(X_1)}{\Delta}.$$

Therefore, the increments of a Lévy process or of all infinitely divisible distributions are always leptokurtic but the skewness (if there is any) and the kurtosis decreases with the time scale over which increments are computed by $\frac{1}{\sqrt{\Delta}}$ and $\frac{1}{\Delta}$, respectively.

With the following proposition we define the exponential moments of a Lévy process:

Proposition 2.17 (Exponential moments) (proposition 3.14 in [2]) *Let $(X_t)_{t\geq 0}$ be a Lévy process on $\mathbb{R}$ with characteristic triplet (A,ν,γ) and let $u\in\mathbb{R}$. We call $E\left[e^{uX_t}\right]$ the exponential moment which is finite for some t or for all $t>0$ if and only if $\int_{|x|\geq 1} e^{ux}\nu(dx)<\infty$. In this case, we have:*

$$E\left[e^{uX_t}\right] = e^{t\psi(-iu)}$$

where ψ is the characteristic exponent of the Lévy process defined in equation (2.7).

For a detailed proof theorem 25.17 in "Lévy Process and Infinitely Divisible Distribution" written by Stato.

One last thing to note is how the Lévy measure ν can be inferred from the characteristic function $\Phi_F(u)$ of the Lévy process (equation (2.9)). First, we need to say that the uniqueness of the matrix A, given by the Lévy triplet (A,ν,γ), is a consequence of the following equation:

$$\lim_{h\to 0} h\log\Phi_{X_t}\left(\frac{u}{\sqrt{h}}\right) = -\frac{t}{2}\langle u, Au\rangle. \tag{2.12}$$

In term of the process X, this result implies that:

$$\left\{\frac{1}{\sqrt{h}X_{ht}}\right\}_{t\geq 0} \xrightarrow[h\to 0]{d} \{\Sigma W_t\}_{t\geq 0}$$

where W is a d-dimensional Wiener processes and Σ is a lower triangular matrix such that $A=\Sigma\Sigma^T$. This means that the short-term increments $(X_{(k+1)h}-X_{kh})_{k=1}^n$, properly scaled,

[9] Defined in the appendix A.5.

[10] X is said to be leptokurtic or "fat-tailed" if $\kappa(X)>0$ ($\kappa(X)$ is called the excess kurtosis of X).

behave like the increments of a Wiener process, when $\Sigma \neq 0$.
Now, we recover $\langle u, Au \rangle$ from equation (2.12) and we can find:

$$\Upsilon(u) := \log \Phi_{X_1}(u) + \frac{1}{2}\langle u, Au \rangle$$

Then, it turns out that:

$$\int_{[-1,1]^d} (\Upsilon(u) - \Upsilon(u+w))dw = \int_{\mathbb{R}^d} e^{i\langle u,x \rangle} \underbrace{2^d \left(1 - \prod_{j=1}^{d} \frac{\sin x_j}{x_j}\right) \nu(dx)}_{=:\tilde{\nu}(dx)} \tag{2.13}$$

where $\frac{\sin x_j}{x_j}$ is equal to 1 when $x_j = 0$ and $\tilde{\nu}$ is a finite measure which can be recovered from the inverse Fourier transform[11] of the left-hand side of the above equation.
Proof
We can note that:

$$(\Upsilon(u) - \Upsilon(u+w)) = \int_{\mathbb{R}^d} \left(e^{i\langle u,x\rangle} - e^{i\langle u+w,x\rangle} + i \langle w, x\rangle \mathbf{1}_{|x|\leq 1}\right) \nu(dx) - i \langle \gamma, w\rangle$$

moreover, we know that the argument inside the integral has the following relation:

$$\begin{aligned} |e^{i\langle u,x\rangle} - e^{i\langle u+w,x\rangle} + i \langle w, x\rangle | &\leq |1 - e^{i\langle u+w,x\rangle} + i \langle w, x\rangle | + | \langle w, x\rangle ||1 - e^{\langle u,x\rangle}| \\ &\leq \frac{1}{2}|w|^2|x|^2 + |w||x|^2|u| \end{aligned}$$

Therefore, we can use Fubini Theorem and we get:

$$\int_{[-1,1]^d} (\Upsilon(u) - \Upsilon(u+w))dw = \int_{\mathbb{R}^d} e^{i\langle u,x\rangle}\nu(dx) \int_{[-1,1]^d} \left(1 - e^{i\langle w,x\rangle}\right) dw$$

which shows equation (2.13).
Now, let show that $\tilde{\nu}(dx)$ is finite. It is finite since:

$$\prod_{j=1}^{d} \frac{\sin x_j}{x_j} = 1 - \frac{1}{6}|x^2| + O(|x|^4) \quad as \quad |x| \to 0$$

□

2.2.4 Lévy processes as Markov process and martingales

From the independent of increments property of the Lévy process, we can built different martingales. Therefore, we can introduce the following proposition:
Proposition 2.18 (proposition 3.17 in [2]) *Let $(X_t)_{t\geq 0}$ be a real-valued process with independent increments. Then:*

1) $\left(\frac{e^{iuX_t}}{E[e^{iuX_t}]}\right)_{t\geq 0}$ *is a martingale,* $\forall u \in \mathbb{R}$;

2) if for some $u \in \mathbb{R}$, $E\left[e^{uX_t}\right] < \infty, \quad \forall t \geq 0$, *then* $\left(\frac{e^{iuX_t}}{E[e^{iuX_t}]}\right)_{t\geq 0}$ *is a martingale;*

3) if $E[X_t] < \infty, \quad \forall t \geq 0$, *then* $M_t = X_t - E[X_t]$ *is a martingale and also a process with independent increments;*

4) if $Var(X_t) < \infty, \quad \forall t \geq 0$, *then* $(M_t)^2 - E[(M_t)^2]$ *is a martingale and* M *is defined as in the point* 3).

[11]Described in the appendix A.6.

Let (X_t) is a Lévy process. Therefore, it is a martingale if its corresponding moments are finite for one value of t.
The proposition 2.18 with the Lévy-Khinchin formula imply the following proposition:
Proposition 2.19 (proposition 3.18 in [2]) *Let $(X_t)_{t\geq 0}$ be a Lévy process on $\mathbb{R}$ with characteristic triplet (A, ν, γ). Then:*

1. *(X_t) is a martingale if and only if $\int_{|x|\geq 1} |x|\nu(dx) < \infty$ and*
$$\gamma + \int_{|x|\geq 1} x\nu(dx) = 0$$
2. *$exp(X_t)$ is a martingale if and only if $\int_{|x|\geq 1} e^x\nu(dx) < \infty$ and*
$$\frac{A}{2} + \gamma + \int_{-\infty}^{\infty} \left(e^x - 1 - x\mathbf{1}_{|x|\leq 1}\right)\nu(dx) = 0$$

The above proposition told us the necessary and sufficient conditions for a Lévy process or its exponential to be a martingale.

Definition 2.9 [Semimartingale] *A semimartingale is a stochastic process $(X_t)_{0\leq t\leq T}$ which admits the decomposition:*

$$X = X_0 + M + A \tag{2.14}$$

where X_0 is finite and $\mathcal{F}_0$-measurable, M is a local martingale with $M_0 = 0$ and A is a finite variation process with $A_0 = 0$.
If A is predictable, then X is a special semimartingale and all special semimartingale have a "canonical decomposition" equal to:

$$X = X_0 + B + X^c + x(J_X - \nu^X) \tag{2.15}$$

where X^c is the continuous martingale part of X and $x(J_X - \nu^X)$ is the purely discontinuous martingale part of X. In particular, J_X is the jump measure of X (defined in equation (2.2)) and ν^X is called the compensator of J_X.
Every Lévy processes are also a samimartingale, which follows from the definition of semimartingale (equation (2.14) and from the Lévy-Ito decomposition (equation (2.6)). On the other hand, every Lévy processes with finite first moment are also special samimartingale and all the Lévy processes which are a special semimartingale, have a finite first moment. Therefore, we have the following lemma:
Lemma 2.1 *Let $(X_t)_{t\geq 0}$ be a Lévy process with Lévy triplet (A, ν, γ). Then, the following conditions are equivalent:*

- *X is a special semimartingale;*
- $\int_{\mathbb{R}^d} (|x| \wedge |x|^2)\nu(dx) < \infty$;
- $\int_{\mathbb{R}^d} (|x|\mathbf{1}_{|x|\geq 1})\nu(dx) < \infty$.

Another important property of the Lévy process is the Markov property. It states that an $\mathcal{F}_t$-adapted process $(X_t)_{t\geq 0}$ satisfies the Markov property if, for any bounded Borel function f and for any s and t, such that $s \leq t$, we have:

$$E[f(X_t)|\mathcal{F}_s] = E[f(X_t)|X_s]$$

In other words, the meaning of the Markov property is that the future behavior of the process $(X_t)_{t\geq 0}$ after t depends only on the value X_t and is not influenced by the history of

the process before t.
We can define the transitional kernel of the process X_t as:

$$P_{s,t}(x, B) = \mathbb{P}(X_t \in B | X_s = x) \quad \forall B \in \mathcal{B}. \tag{2.16}$$

Moreover, the markov property implies the Champam-Kolmogorov equations:

$$P_{s,u}(x, B) = \int_{\mathbb{R}^d} P_{s,t}(x, dy) P_{t,u}(y, B)$$

An important result from these condition is that the Lévy processes are the only Markov processes which are homogeneous in space and in time. In fact, the Lévy processes satisfy a stronger version of the Markov property: for all t, the increments $(X_{t+s} - X_t)_{s \geq 0}$ has the same law as the process $(X_s)_{s \geq 0}$ and is independent from $(X_s)_{a \leq s \leq t}$. Therefore, the strong Markov property of Lévy processes allows to replace the nonrandom time t by any random time which is nonanticipating with respect to the history of X. If τ is a nonanticipating random time, then the process $Y_t = X_{t+\tau} - X_\tau$ is again a Lévy process, independent from $\mathcal{F}_\tau$ and with the same law as $(X_t)_{t \geq 0}$.

2.3 Jump-Diffusion Model

The financial models with jumps can be decomposed in two main categories: the jump-diffusion model and the infinite activity models. We focus only in the first category but we make a short description also of the second type.
The infinite activity models consists in a model with infinite number of jumps in every interval, therefore we did not need to introduce a Brownian component since the process moves essentially by jumps. This imply that the distribution of the jump size does not exist because jumps arrive infinitely often. The infinite activity model gives a more realistic description of the historical price process.
On the other hand, in the jump-diffusion model the evolution of prices are given by a diffusion process which has jumps at random intervals. Here, the jumps represent rare events such as crashes and large drawdown. Since the distribution of jump sizes is known, the dynamic structure of the jump process is easy to understand and describe. The jump-diffusion models perform well for implied volatility smile interpolation.

2.3.1 Exponential Lévy Models

In order to construct an exponential Lévy model for the process X, we need to start from the Black-Scholes model and how it describes the evolution of an asset price. Here, the asset price (S_t) follow a geometric Brownian motion:

$$S_t = S_0 e^{\mu t + \sigma W_t} \tag{2.17}$$

If we replace $\mu t + \sigma W_t$ by a Lévy process X_t, we obtain the class of the exponential Lévy models:

$$S_t = S_0 e^{X_t} \tag{2.18}$$

Now, consider a Lévy process of jump-diffusion type with the following form:

$$X_t = \gamma t + \sigma W_t + \sum_{i=1}^{N_t} Y_1 \tag{2.19}$$

where $(N_t)_{t \geq 0}$ is the Poisson process which counting the jumps of X and Y_i are the jump sizes, which are i.i.d. variables. Therefore, the evolution of the asset price becomes:

$$S_t = S_0 e^{\gamma t + \sigma W_t + \sum_{i=1}^{N_t} Y_1} \tag{2.20}$$

We need to specify the distribution of jump sizes $\nu_0(x)$ in order to define the parametric model completely. Is is important to specify the tail behavior of $\nu_0(x)$ correctly because the tail behavior of the jump measure determines the tail behavior of the probability density of the process.

In the *Merton model* (introduced by Merton in 1973 with the article *"Option pricing when underlying stock return are discontinuous"*) we have that the process is equal to the equation (2.20) and the jumps are assumed to have a Gaussian distribution, therefore $Y_i \sim N(\mu, \delta^2)$. This allows to obtain the probability density of X_t as a quickly converging series. In fact,

$$\mathbb{P}(X_t \in A) = \sum_{k=0}^{\infty} \mathbb{P}(X_t \in A | N_t = k)\mathbb{P}(N_t = k)$$

then the probability density of X_t satisfies the equation:

$$p_t(x) = e^{-\lambda t} \sum_{k=0}^{\infty} \frac{(\lambda t)^k exp\left\{-\frac{(x-\gamma t-k\mu)^2}{2(\sigma^2 t+k\delta^2)}\right\}}{k!\sqrt{2\pi(\sigma^2 t + k\delta^2)}}$$

The Lévy density of the model is equal to:

$$\nu(x) = \frac{\lambda}{\delta\sqrt{2\pi}} exp\left\{-\frac{(x-\mu)^2}{2\delta^2}\right\}$$

One last thing to note is the moment of the process in the Merton model. Hence, we have that the characteristic exponent of the characteristic function is equal to:

$$\psi(u) = i\gamma u - \frac{\sigma^2 u^2}{2} + \lambda\left\{e^{-\frac{\delta^2 u^2}{2} + i\mu u} - 1\right\}$$

It follows that: $E[X_t] = t(\gamma + \lambda\mu)$ and $Var(X_t) = t(\sigma^2 + \lambda\gamma^2 + \lambda\mu^2)$. If we analyze the moment, we can note the tail behavior of the probability density, which are heavier than Gaussian but all the exponential moments are finite.

In the *Kou model* (introduced by Kou in 2002 with the article *"A jump-diffusion model for option pricing"*) we have that the process X_t is equal as in the Merton model but the distribution of jumps sizes is an asymmetric exponential (i.e. has a double exponential distribution, therefore $Y_i \sim DbExp(p, \theta_1, \theta_2)$) with a density of the form:

$$\nu_0(dx) = \left[p\theta_1 e^{-\theta_1 x}\mathbf{1}_{x>0} + (1-p)\theta_2 e^{-\theta_2|x|}\mathbf{1}_{x<0}\right] dx \tag{2.21}$$

where $\theta_1 > 0$, $\theta_2 > 0$ represent the decay of the tails for the distribution of positive and negative jump sizes, respectively, and $p \in [0, 1]$ represent the probability of an upward jump. Therefore, we can easily find the Lévy measure of the process:

$$\nu(x) = p\lambda\theta_1 e^{-\theta_1 x}\mathbf{1}_{x>0} + (1-p)\lambda\theta_2 e^{-\theta_2|x|}\mathbf{1}_{x<0}$$

The first two moments of the process are equal to: $E[X_t] = t\left(\gamma + \frac{\lambda t}{\theta_1} - \frac{\lambda(1-p)}{\theta_2}\right)$ and $Var(X_t) = t\left(\sigma^2 + \frac{\lambda t}{\theta_1^2} - \frac{\lambda(1-p)}{\theta_2^2}\right)$. We find these two result from the characteristic function of the process, which has characteristic exponent equal to:

$$\psi(u) = i\gamma u - \frac{\sigma^2 u^2}{2} + iu\lambda\left\{\frac{p}{\theta_1 - iu} - \frac{1-p}{\theta_2 + iu}\right\}$$

In this case, the probability distribution of returns has semi-heavy exponential tails. On one hand, we have that $p(x) \sim e^{-\theta_1 x}$ when $x \to +\infty$, on the other hand, we have that $p(x) \sim e^{-\theta_2|x|}$ when $x \to -\infty$.

The advantage of the Kou model compared to the Merton model is that analytical expressions for expectations involving first passage times may be obtained due to the memoryless property of exponential random variables.

In the Kou and in the Merton model the intensity of the jump is assume to be constant. If we want to implement the model building of the price evolution of an asset, we can consider the Doubly stochastic Poisson process, also known as *Cox processes*. Here, we want that the intensity λ of the counting process N is stochastic, therefore we want that $\lambda = (\lambda)_{0 \leq t \leq T}$. One approach is to computing the probability that an event arrives at time t given the information we have at time s, hence is to define $\mathbb{P}(N_t - N_s = n|\mathcal{F})$ where $\mathcal{F}$ is the natural filtration generated by (N, λ). Therefore, we have the following property:

$$\mathbb{P}\left(N_t - N_s = n|\mathcal{F}_s \vee \sigma((\lambda_u)_{s \leq u \leq t})\right) = exp\left\{-\int_s^t \lambda_u du\right\} \frac{\left(\int_s^t \lambda_u du\right)^n}{n!}$$

so that

$$\mathbb{P}\left(N_t - N_s = 0|\mathcal{F}_s\right) = E\left[exp\left\{-\int_s^t \lambda_u du\right\} \frac{\left(\int_s^t \lambda_u du\right)^n}{n!} \middle| \mathcal{F}_s\right]$$

where $\sigma((\lambda_u)_{s \leq u \leq t})$ denotes the smallest $\sigma-$algebra generated by the intensity process λ over the time interval $[s, t]$, and $\mathcal{F}_s \vee \sigma((\lambda_u)_{s \leq u \leq t})$ represents the information contained on the entire path of λ up to time t, but excluding the information on the N process on the interval $(s, t]$. Therefore, we can note that the Cox process is conditionally (conditioned on $\sigma((\lambda_u)_{s \leq u \leq t})$) an inhomogeneous Poisson process with the conditionally known intensity. Some examples of driver of the intensity process might be an independent diffusion, an independent jump process or a counting process itself.
Now, we can give few examples of the intensity process λ:

1. *Feller process:*
$$d\lambda_t = k(\theta - \lambda_t)dt + \eta\sqrt{\lambda_t}dW_t$$
2. *Ornstein-Uhlenbeck process:*
$$d\lambda_t = -k\lambda_t dt + \gamma dJ_t$$
3. *Jump-diffusion:*
$$d\lambda_t = k(\theta - \lambda_t)dt + \eta\sqrt{\lambda_t}dW_t + \gamma dJ_t$$
4. *Hawkes process:*
$$\lambda_t = \int_0^t g(t-s)dN_s$$

where W is an independent Brownian motion, θ is the long-run mean, k is the rate of mean reversion and J is an independent compound Poisson process with non-negative jumps and with intensity λ_j and i.i.d. jumps ε with distribution function F. The first three processes exhibit mean-reversion. The second process mean-reverts to 0, while the first and the third mean revert to θ. However, if we consider the jump in the process, the mean-reversion level does not reflect the long-run behavior. Hence, we should rewrite the process in terms of their compensated version and we introduce the following proposition:
Proposition 2.20 (Compensated Doubly stochastic Poisson process) $\tilde{N} = (\tilde{N}_t)_{0 \leq t \leq T}$ *is a martingale if* $\tilde{N}_t = N_t - \int_0^t \lambda_s ds$.
Therefore, in terms of the compensated Doubly stochastic Poisson process we have:

1. *Ornstein-Uhlenbeck process:*
$$d\lambda_t = k\left(\frac{\gamma\lambda_j}{k}E[\varepsilon] - \lambda_t\right)dt + \gamma d\tilde{J}_t$$
2. *Jump-diffusion:*
$$d\lambda_t = k\left(\theta + \frac{\gamma\lambda_j}{k}E[\varepsilon] - \lambda_t\right)dt + \eta\sqrt{\lambda_t}dW_t + \gamma d\tilde{J}_t$$

Then, we can see that the expected average intensities in the long run are $\frac{\gamma\lambda_j}{k}E[\varepsilon]$ and $\theta + \frac{\gamma\lambda_j}{k}E[\varepsilon]$, respectively. Therefore, the expected long run intensity is the mean-reversion level plus the jump correction terms $\frac{\gamma\lambda_j}{k}E[\varepsilon]$.

The following proposition told us the condition of the exponential Lévy process to be a semimartingale:

Proposition 2.21 (Exponential Lévy process) (proposition 8.20 in [2]) *Let $(X_t)_{t\geq 0}$ be a Lévy process with Lévy triplet (σ^2, ν, γ) verifying $\int_{|y|\geq 1} e^y \nu(dy) < \infty$. Then, $Y_t = e^{X_t}$ is a semimartingale with decomposition $Y_t = M_t + A_t$ where the martingale part is given by:*

$$M_t = 1 + \int_0^t Y_{s-}\sigma dW_s + \int_{[0,t]\times\mathbb{R}} Y_{s-}\left(e^z - 1\right)\tilde{J}_X(ds \times dz)$$

and the continuous finite variation drift part is given by:

$$A_t = \int_0^t Y_{s-}\left[\gamma + \frac{\sigma^2}{2} + \int_{-\infty}^{\infty}\left(e^z - 1 - z\mathbf{1}_{|z|\geq 1}\right)\nu(dz)\right]ds.$$

(Y_t) is a martingale if and only if

$$\gamma + \frac{\sigma^2}{2} + \int_{-\infty}^{\infty}\left(e^z - 1 - z\mathbf{1}_{|z|\geq 1}\right)\nu(dz) = 0$$

Proof

Let $(X_t)_{t\geq 0}$ be a Lévy process with jump measure J_X and let $Y_t = e^{X_t}$. Then, we apply the Ito formula to Y_t and we find:

$$\begin{aligned} Y_t &= 1 + \int_0^t Y_{s-}\sigma dX_s + \frac{\sigma^2}{2}\int_0^t Y_{s-}ds + \sum_{0\leq s\leq t;\Delta X_s\neq 0}\left(e^{X_{s-}+\Delta X_s} - e^{X_{s-}} - \Delta X_s e^{X_{s-}}\right) \\ &= 1 + \int_0^t Y_{s-}\sigma dX_s + \frac{\sigma^2}{2}\int_0^t Y_{s-}ds + \int_{[0,t]\times\mathbb{R}} Y_{s-}\left(e^z - 1 - z\right)\tilde{J}_X(ds \times dz) \end{aligned}$$

We can make the assumption that $E[|Y_t|] = E[e^{X_t}] < \infty$ which is equivalent to $\int_{|y|\geq 1} e^y \nu(dy) < \infty$. Therefore, we can decompose Y_t into a martingale part and a drift part, where the martingale part is the sum of an integral with respect to the Brownian component of X and a compensated sum of jump terms:

$$1 + \int_0^t Y_{s-}\sigma dW_s + \int_{[0,t]\times\mathbb{R}} Y_{s-}\left(e^z - 1\right)\tilde{J}_X(ds \times dz)$$

while the drift term is given by:

$$\int_0^t Y_{s-}\left[\gamma + \frac{\sigma^2}{2} + \int_{-\infty}^{\infty}\left(e^z - 1 - z\mathbf{1}_{|z|\geq 1}\right)\nu(dz)\right]ds$$

Then, Y is a martingale if and only if $E[e^{X_t}] = E[Y_t] = 1$ but $E[e^{X_t}] = e^{t\psi_X(-i)}$, where ψ_X is the characteristic exponent of X. Hence, we obtain that:

$$\psi_X(-i) = \gamma + \frac{\sigma^2}{2} + \int_{-\infty}^{\infty}\left(e^z - 1 - z\mathbf{1}_{|z|\geq 1}\right)\nu(dz) = 0$$

□

2.3.2 Stochastic exponential of Jump process

The stochastic exponential was introduced by Doléans-Dade and it can be found using the Ito formula in the geometric Brownian motion (equation (2.17)) and substituting a Lévy process. Hence, if we apply the Ito formula in (2.17) we obtain:

$$\frac{dS_t}{S_t} = \left(\mu + \frac{\sigma^2}{2}\right) dt + \sigma dW_t$$

Then, we can define $B_t^1 = (\mu + \frac{\sigma^2}{2})t + \sigma W_t$ and the above equation becomes:

$$\frac{dS_t}{S_t} = dB_t^1 \tag{2.22}$$

If we substitute B_t^1 by a Lévy process X, we obtain the stochastic exponential. Therefore, with the following proposition we can introduce a generic stochastic exponential for a process $(Z_t)_{t\geq 0}$.

Proposition 2.22 (Stochastic exponential) (proposition 8.21 in [1]) *Let $(X_t)_{t\geq 0}$ be a Lévy process with Lévy triplet (σ^2, ν, γ). Then, there exists a unique càdlàg process $(Z_t)_{t\geq 0}$ such that:*

$$\begin{cases} dZ_t = Z_{t^-} dX_t \\ Z_0 = 1 \end{cases} \tag{2.23}$$

Where Z is given by:

$$Z_t = e^{X_t - \frac{1}{2}\int_0^t \sigma_s^2 ds} \prod_{0\leq s\leq t} (1 + \Delta X_s) e^{-\Delta X_s} \tag{2.24}$$

If $\int_{-1}^{1} |x|\nu(dx) < \infty$, then the jumps of X have finite variation and the stochastic exponential of X can be expressed as:

$$Z_t = e^{\sigma W_t + \gamma_0 t - \frac{\sigma^2 t}{2}} \prod_{0\leq s\leq t} (1 + \Delta X_s)$$

where $\gamma_0 = \gamma - \int_{-1}^{1} x\nu(dx)$.

Z is called the stochastic exponential of X and is denoted by $Z = \mathcal{E}(X)$.

Proof

The first step is to show that the following process exists and is of finite variation:

$$V_t = \prod_{0\leq s\leq t; \Delta X_s \neq 0} (1 + \Delta X_s) e^{-\Delta X_s}$$

So, we decompose the process V_t as the product of two terms:

$$V_t = V_t^a V_t^b$$

where:

$$V_t^a = \prod_{0\leq s\leq t; |\Delta X_s|\leq 1/2} (1 + \Delta X_s) e^{-\Delta X_s}$$
$$V_t^b = \prod_{0\leq s\leq t; |\Delta X_s|> 1/2} (1 + \Delta X_s) e^{-\Delta X_s}$$

Let's start to analyze V_t^b. Since, for every t, it is a product of finite number of factors follows that it is of finite variation. Now, we look to V_t^a. We consider its logarithm because it is positive and we have:

$$\ln V_t^a = \sum_{0\leq s\leq t; |\Delta X_s|\leq 1/2} (\ln(1 + \Delta X_s) - \Delta X_s)$$

Each terms of this sum satisfies: $0 > \ln(1+\Delta X_s) - \Delta X_s > -\Delta X_s^2$. Therefore, the series is decreasing and bounded fro, below by $-\sum_{0\le s\le t}\Delta X_s^2$, which is finite for every Lévy process. Hence, $\ln V_t^a$ exists and is a decreasing process. Finally, we can say that V_t exists and has trajectories of finite variation.
In the second step, we consider:

$$Z_t = e^{X_t - \frac{1}{2}\int_0^t \sigma_s^2 ds} V_t$$

now, if we apply the Ito formula at the equation define above, we find in differential form:

$$\begin{aligned} dZ_t = & -\frac{\sigma_t^2}{2} e^{X_t - \frac{1}{2}\int_0^t \sigma_s^2 ds} V_t dt + e^{X_{t-} - \frac{1}{2}\int_0^t \sigma_s^2 ds} V_{t-} dX_t \\ & + \frac{\sigma_t^2}{2} e^{X_t - \frac{1}{2}\int_0^t \sigma_s^2 ds} V_t dt + e^{X_t - \frac{1}{2}\int_0^t \sigma_s^2 ds} V_t \\ & - e^{X_{t-} - \frac{1}{2}\int_0^t \sigma_s^2 ds} V_{t-} - e^{X_{t-} - \frac{1}{2}\int_0^t \sigma_s^2 ds} V_{t-} \Delta X_t - e^{X_{t-} - \frac{1}{2}\int_0^t \sigma_s^2 ds} \Delta V_t \end{aligned}$$

Since V_t is a pure jump process we have that: $dV_t \equiv \Delta V_t = V_{t-}\left(e^{\Delta X_t}(1+\Delta X_t) - 1\right)$. Then, substituting into the above equation and make some calculus, we find the equation (2.23).

□

We can note that the stochastic exponential is always nonnegative if all the jumps of X_t are greater than -1, i.e. $\nu((-\infty,-1]) = 0$.
Goll and Kallsen have shown that the stochastic exponential is equivalent to the ordinary exponential. In fact, if $Z > 0$ is the stochastic exponential of a Lévy process, it is also the ordinary exponential of another Lévy process (it is also true the opposite case). Therefore, the two exponential end up by giving us the same class of positive processes. The following proposition shows the relation between ordinary and stochastic exponential:
Proposition 2.23 (Relation between ordinary and stochastic exponential) (proposition 8.22 in [2])

1. *Let $(X_t)_{t\ge0}$ be a real valued Lévy process with Lévy triplet (σ^2,ν,γ) and $Z = \mathcal{E}(X)$ its stochastic exponential. If $Z > 0$ almost surely, then there exists another Lévy process $(L_t)_{t\ge0}$ with triplet $(\sigma_L^2,\nu_L,\gamma_L)$ such that $Z_t = e^{L_t}$ where:*

$$\begin{aligned} L_t = \ln Z_t = X_t - \frac{\sigma^2 t}{2} + \sum_{0\ge s\ge t} (\ln(1+\Delta X_s) - \Delta X_s \\ \sigma_L = \sigma \\ \nu_L(A) = \nu(\{x : \ln(1+x) \in A\}) = \int \mathbf{1}_A(\ln(1+x))\nu(dx) \\ \gamma_L = \gamma - \frac{\sigma^2}{2} + \int [\ln(1+x)\mathbf{1}_{[-1,1]}(\ln(1+x)) - x\mathbf{1}_{[-1,1]}(x)]\nu(dx) \end{aligned}$$

2. *Let $(L_t)_{t\ge0}$ be a real valued Lévy process with Lévy triplet $(\sigma_L^2,\nu_L,\gamma_L)$ and $S_t = e^{L_t}$ its exponential. Then, there exists a Lévy process $(X_t)_{t\ge0}$ such that S_t is the stochastic exponential of X : $S = \mathcal{E}(X)$ where:*

$$X_t = L_t + \frac{\sigma^2 t}{2} + \sum_{0\le s\le t}\left(e^{\Delta L_s} - 1 - \Delta L_s\right)$$

Therefore, the Lévy triplet (σ^2,ν,γ) of X is given by:

$$\begin{aligned} \sigma = \sigma_L \\ \nu(A) = \nu_L(\{x : (e^x - 1) \in A\}) = \int \mathbf{1}_A(e^x - 1)\nu_L(dx) \\ \gamma = \gamma_L - \frac{\sigma_L^2}{2} + \int [(e^x-1)\mathbf{1}_{[-1,1]}(e^x-1) - x\mathbf{1}_{[-1,1]}(x)]\nu_L(dx) \end{aligned}$$

3. *Let* $(X_t)_{t\geq 0}$ *be a Lévy process and a martingale. Then, its stochastic exponential* $Z = \mathcal{E}(X)$ *is also a martingale. Therefore, for every Lévy process* X *with* $E[|X_t|] < \infty$ *we have:*

$$E[\mathcal{E}(X_t)] = e^{E[X_t]} \quad t > 0$$

   This property is also known as Martingale preserving property.

A detailed proof can be found in chapter 8 of "Financial Modeling with Jump Process" written by Cont and Tankov.

Chapter 3

Stochastic Calculus for Jump Process

In financial market there are two type of strategies related to the price of a financial asset: trading strategies and hedging strategies.
If we want to describe a trading strategy, we need to consider a dynamic portfolio resulting from buying and selling the assets which satisfies the non-arbitrage assumption. We can define an arbitrage strategy as a self-financing strategy ϕ with zero initial value and non-zero final value with probability equal to 1. Moreover, a strategy is called self-financing if the following equation is satisfied for all $t : \langle \phi_t, S_t \rangle = \langle \phi_{t+1}, S_t \rangle$. Therefore, we can consider an investor who trades at times $T_0 = 0 < T_1 < \cdots < T_n < T_{n+1} = T$ and detaining a quantity ϕ_i of an asset whose price is S during the period $(T_i, T_{i+1}]$. Then, we can definite the capital gain $G_t(\phi)$ as:

$$G_t(\phi) := \sum_{i=0}^{n} \phi_i (S_{T_{i+1}} - S_{T_i}) \tag{3.1}$$

We can write the quantity which represents the capital gain of the investors following the strategy ϕ as :

$$G_t(\phi) = \sum_{i=0}^{n} \phi_i (S_{T_{i+1}} - S_{T_i}) = \int_0^T \phi_t dS_t \tag{3.2}$$

where the last term in equation (3.2) represent the stochastic integral ϕ with respect to S.

In this chapter, we describe the stochastic integral and the main tools to explain the time evolution of a derivative instrument. The first section introduces the concept of stochastic integral. We describe its properties in the case is built with a semimartingale of respect to a Brownian motion. Then, we give the definition of quadratic variation and covariation for the stochastic integral. The second part of the chapter is entirely focused on the stochastic integral with jump and, in particular, we talk about the stochastic integral with respect to a Poisson process and to a Poisson Random measure. The last section talks about the Ito's formula, which is the key tool to describe the time evolution of a derivative instruments. Before we define the Ito's formula for a jump-diffusion and, in general for a Lévy process, then the introduce the Ito's formula for martingale and semimartingale.

3.1 Stochastic integral

Consider a vector of asset whose price S is described by a stochastic process, i.e. $S_t = (S_t^1, S_t^2, \ldots S_t^d)$ and a portfolio $\phi = (\phi^1, \phi^2, \ldots, \phi^d)$ which describes the amount of each asset held by the investor. Therefore, the value of such portfolio at time t is equal to:

$$V_t(\phi) = \sum_{k=1}^{d} \phi^d S_t^k \equiv \langle \phi_t, S_t \rangle \tag{3.3}$$

We also assume a dynamic trading strategy, which consist in buying and selling assets at different dates, and we consider an investor who trades at times $T_0 = 0 < T_1 < \cdots < T_n < T_{n+1} = T$. We also assume that the strategy is self-financing and that between two transaction dates T_i and T_{i+1} the portfolio remains unchanged. The meaning of the self-financing assumption is that at time t the investors readjusts his position from ϕ_t to ϕ_{t+1} without bringing or consuming any wealth. Moreover, if we dropped this assumption, we would had arbitrage opportunities because a portfolio which is empty at time 0 but to which cash (> 0) is added, without any liability, would trivially be an arbitrage portfolio. The second assumption told us that the investor did not know in advance the transaction dates but he will decide to buy or sell at T_{i+1} depending on the information revealed before T_{i+1}. Hence, the transaction date T_{i+1} is a stopping time. In the first chapter, we assume that the processes are càdlàg (i.e. right continuous with left limits), whereas here we have that the trading strategy is càglàd (i.e. left continuous with right limits). We have the left continuity in the process because if the investor decides to make a transaction at $t = T_i$, the portfolio will take the new value at ϕ_i before that the value of the portfolio is still described by ϕ_{i-1}. Therefore, we have that $(\phi_t)_{t\in[0,T]}$ is a predictable process and we have the following definition:
Definition 3.1 [Simple Predictable Process] (definition 8.1 in [2]) *A stochastic process $(\phi_t)_{t\in[0,T]}$ is called a simple predictable process if it can be represented as:*

$$\phi_t = \phi_0 \mathbf{1}_{t=0} + \sum_{i=0}^{n} \phi_i \mathbf{1}_{(T_i, T_{i+1}]}(t) \tag{3.4}$$

where $T_0 = 0 < T_1 < \cdots < T_n < T_{n+1} = T$ are nonanticipating random times and each ϕ_i is bounded random variable whose value is revealed at T_i (i.e. $\mathcal{F}_{T_i}$-measurable).
We can define the gain process of the strategy ϕ followed by an investor as the stochastic process $(G_t(\phi))_{t\in[0,T]}$ equal to:

$$G_t(\phi) = \langle \phi_0, S_0 \rangle + \sum_{i=0}^{j-1} \big\langle \phi_i, (S_{T_{i+1}} - S_{T_i}) \big\rangle + \big\langle \phi_j, (S_t - S_{T_j}) \big\rangle \quad for \quad T_j < t \leq T_{j+1}$$

where $(S_{T_{i+1}} - S_{T_i})$ represent the asset price movement between time T_{i+1} and T_i. We can write the equation above with stopping time notation, therefore we find a more compact equation:

$$G_t(\phi) = \langle \phi_0, S_0 \rangle + \sum_{i=0}^{n} \big\langle \phi_i, (S_{T_{i+1}\wedge t} - S_{T_i \wedge t}) \big\rangle \tag{3.5}$$

where $T_{i+1} \wedge t$ represent the minimum between T_{i+1} and t. Hence, the stochastic process $G_t(\phi)$ can be expressed as the stochastic integral of the simple predictable process ϕ with respect to S and it is equal to:

$$\int_0^t \phi_u dS_u := \langle \phi_0, S_0 \rangle + \sum_{i=0}^{n} \big\langle \phi_i, (S_{T_{i+1}\wedge t} - S_{T_i \wedge t}) \big\rangle \tag{3.6}$$

Since the self-financing assumption imply that the cost of the process (defined as $C_t(\phi) = V_t(\phi) - G_t(\phi) = \langle \phi_0, S_t \rangle - \int_0^t \phi_u dS_u$) is equal to zero, we have that the value of the portfolio, $V_t(\phi)$, is equal to:

$$V_t(\phi) = \int_0^t \phi_u dS_u = \phi_0 S_0 + \int_{0^+}^t \phi_u dS_u$$

where the first term is the initial value of the portfolio and the second term is the capital gain between 0 and t. Therefore, for an investors the only source of variation of the portfolio's value is the variation of the asset values.

Proposition 3.1 (Martingale preserving property) (proposition 8.1 in [2]) *If $(S_t)_{t\in[0,T]}$ is a martingale, then for any predictable process ϕ the stochastic integral $G_t = \int_0^t \phi dS$ is*

also a martingale.
Proof
Consider the process defined in the equation (3.4). We need to prove that $E[G_T|\mathcal{F}_s] = G_t$ but is sufficient to show that $E[\phi_i(S_{T_{i+1}} - S_{T_i})|\mathcal{F}_t] = \phi_i(S_{T_{i+1}\wedge\tau} - S_{T_i\wedge\tau})$ for each i. Therefore, we have that:

$$\begin{aligned}E[\phi_i(S_{T_{i+1}} - S_{T_i})|\mathcal{F}_t] &= E[\mathbf{1}_{t>T_{i+1}}\phi_i(S_{T_{i+1}} - S_{T_i})|\mathcal{F}_t] \\ &+ E[\mathbf{1}_{(T_i,T_{i+1}]}(t)\phi_i(S_{T_{i+1}} - S_{T_i})|\mathcal{F}_t] \\ &+ E[\mathbf{1}_{t\leq T_i}\phi_i(S_{T_{i+1}} - S_{T_i})|\mathcal{F}_t]\end{aligned}$$

Since $\mathbf{1}_{t>T_{i+1}}$, $\mathbf{1}_{(T_i,T_{i+1}]}$ and $\mathbf{1}_{t\leq T_i}$ are $\mathcal{F}_t$-measurable because T_i and T_{i+1} are stopping times, we can bring out the indicator function from the conditional expectation and, in the third term, we need to use the law of iterated expectation. Hence, the three terms in the above equation become:

$$\begin{aligned}E[\mathbf{1}_{t>T_{i+1}}\phi_i(S_{T_{i+1}} - S_{T_i})|\mathcal{F}_t] &= \mathbf{1}_{t>T_{i+1}}\phi_i(S_{T_{i+1}} - S_{T_i}) \\ E[\mathbf{1}_{(T_i,T_{i+1}]}(t)\phi_i(S_{T_{i+1}} - S_{T_i})|\mathcal{F}_t] &= \mathbf{1}_{(T_i,T_{i+1}]}(t)\phi_i E[(S_{T_{i+1}} - S_{T_i})|\mathcal{F}_t] \\ &= \mathbf{1}_{(T_i,T_{i+1}]}(t)\phi_i(S_t - S_{T_i}) \\ E[\mathbf{1}_{t\leq T_i}\phi_i(S_{T_{i+1}} - S_{T_i})|\mathcal{F}_t] &= \mathbf{1}_{t\leq T_i}E[E[\phi_i(S_{T_{i+1}} - S_{T_i})|\mathcal{F}_{T_i}]|\mathcal{F}_t] \\ &= \mathbf{1}_{t\leq T_i}E[\phi_i E[(S_{T_{i+1}} - S_{T_i})|\mathcal{F}_{T_i}]|\mathcal{F}_t] \\ &= \mathbf{1}_{t\leq T_i}E[\phi_i(\underbrace{E[S_{T_{i+1}}|\mathcal{F}_{T_i}]}_{=S_{T_i}} - \underbrace{E[S_{T_i}|\mathcal{F}_{T_i}]}_{=S_{T_i}})|\mathcal{F}_t] \\ &= 0\end{aligned}$$

Therefore, putting all together we have:

$$\begin{aligned}E[\phi_i(S_{T_{i+1}} - S_{T_i})|\mathcal{F}_t] &= \mathbf{1}_{t>T_{i+1}}\phi_i(S_{T_{i+1}} - S_{T_i}) + \mathbf{1}_{(T_i,T_{i+1}]}(t)\phi_i(S_t - S_{T_i}) \\ &= \phi_i(S_{T_{i+1}\wedge\tau} - S_{T_i\wedge\tau})\end{aligned}$$

□

This proposition imply that if the asset follows a martingale then the the value of any self-financing strategy is a martingale.

Now, consider a nonanticipating càdlàg process $(X_t)_{t\in[0,T]}$, we can built a new stochastic processes by choosing various simple predictable processes $(\sigma_t)_{t\in[0,T]}$, hence the new process is equal to:

$$\int_0^t \sigma_u dX_u$$

where X_t represents the "source of randomness" and σ_t is the "volatility coefficient". Therefore, we have the following proposition:
Proposition 3.2 (Associativity) (proposition 8.2 in [2]) *Let $(X_t)_{t\in[0,T]}$ be a real-valued nonanticipating càdlàag process and $(\sigma_t)_{t\geq 0}$ and $(\phi_t)_{t\geq 0}$ be a real-valued predictable processes. Then, $S_t = \int_0^t \sigma dX$ (which can be written in differential notation as $dS_t = \sigma_t dX_t$) is a nonanticipating càdlàg process and*

$$\int_0^t \phi_u dS_u = \int_0^t \phi_u \sigma_u dX_u. \tag{3.7}$$

The associativity proposition means that the gain process of any strategy involving S, defined as stochastic integral with respect to a source of randomness (i.e. $S_t = \int_0^t \sigma dX$), can be expressed as a stochastic integral with respect to X.

3.1.1 Semimartingale

Since a Lévy process X is not stable under stochastic integration or non-linear transformations, we need to consider the class of samimartingales, which are a larger class of stochastic processes. These kind of class are both stable under stochastic integration and non-linear transformation. Moreover, they are also stable under other operation such as change of filtration and change of measure. We have already given the definition of semimartingale (definition 2.9) but now, we give the definition of semimartingale with respect a simple predictable process.

Definition 3.2 [Semimartingale] (definition 8.2 in [2]) *A nonanticipating càdlàg process S is called a semimartingale if the stochastic integral of simple predictable process with respect to S:*

$$\phi = \phi_0 \mathbf{1}_{t=0} + \sum_{i=0}^{n} \phi_i \mathbf{1}_{(T_i, T_{i+1}]} \mapsto \int_0^T \phi dS = \phi_0 S_0 + \sum_{i=0}^{n} \phi_i (S_{T_{i+1}} - S_{T_i})$$

verifies the following continuity property: for every ϕ^n, $\phi \in \mathbb{S}([0,T])$ if:

$$\sup_{(t,\omega)\in[0,T]\times\Omega} |\phi_t^n(\omega) - \phi_t(\omega)| \underset{n\to\infty}{\to} 0$$

then

$$\int_0^T \phi^n dS \xrightarrow[n\to\infty]{\mathbb{P}} \int_0^T \phi dS \tag{3.8}$$

where $\mathbb{S}([0,T])$ is a set of simple predictable processes on $[0,T]$.

The class of semimartingales satisfy the stability property: a small change in the portfolio should lead to a small change in the gain process. If this property does not hold, it means that a small change in the portfolio can lead to large change in the gain process. Therefore, we need to use stochastic processes which are semimartingales and the following proposition shows that the stability property holds for process defined by stochastic integral:

Proposition 3.3 (proposition 8.3 in [2]) *If $(S_t)_{t\in[0,T]}$ is a semimartingale then for every ϕ^n, $\phi \in \mathcal{S}([0,T])$:*

$$if \quad \sup_{(t,\omega)\in[0,T]\times\Omega} |\phi_t^n(\omega) - \phi_t(\omega)| \underset{n\to\infty}{\to} 0 \tag{3.9}$$

$$then \quad \sup_{t\in[0,T]} \left| \int_0^t \phi^n dS - \int_0^t \phi dS \right| \xrightarrow[n\to\infty]{\mathbb{P}} 0. \tag{3.10}$$

Moreover, we have that any linear combination of a finite number of semimartingales is a semimartingales. In fact, all the Lévy processes are semimartingale because it can be decomposed into a sum of square integrable martingale (the Wiener process) and a finite variation process (the Poisson process). We can easily see that for a finite variation process S, we always have:

$$\sup_{t\in[0,T]} \int_0^t \phi dS \leq TV(S) \sup_{(t\omega)\in[0,T]\times\Omega} |\phi_t(\omega)|$$

where $TV(S)$ is the total variation[1] of S on $[0,T]$. Then, for a square integrable martingale

[1] Defined in chapter 1

M we have that:

$$
\begin{aligned}
E\left[\left(\int_0^t \phi dM\right)^2\right] &= E\left[\left(\phi_0 M_0 + \sum_{i=0}^{n} \phi_i (M_{T_{i+1}\wedge\tau} - M_{T_i\wedge\tau})\right)^2\right] \\
&= E\left[\phi_0^2 M_0^2 + \sum_{i=0}^{n} \phi_i^2 (M_{T_{i+1}\wedge\tau} - M_{T_i\wedge\tau})^2\right] \\
&\leq \sup_{t,\omega} |\phi_s(\omega)| E\left[M_0^2 + \sum_{i=0}^{n} (M_{T_{i+1}\wedge\tau} - M_{T_i\wedge\tau})^2\right] \\
&\leq \sup_{t,\omega} |\phi_s(\omega)| E\left[M_0^2 + \sum_{i=0}^{n} (M^2_{T_{i+1}\wedge\tau} - M^2_{T_i\wedge\tau})\right] \\
&\leq \sup_{t.\omega} |\phi_s(\omega)| \sup_s E[M_s^2]
\end{aligned}
$$

To show this result we have used the Optional Sampling Theorem (Theorem 1.10). Moreover, the above inequality implies the convergence in probability because the stochastic integrals converge in L^2, uniformly in t.

Finally, we can note that all the new processes constructed from semimartingales using stochastic integration are again semimartingales due to associativity property, which helps us to show that a stochastic integral with respect to a semimartingale is a semimartingale. And that every semimartingale is the sum of a finite variation process and a local martingale, which can be defined as the process (X_t) in which there exists a sequence of stopping times $(\tau_i)_{i\geq 1}$ such that $\tau_i \to \infty$ when $i \to \infty$ and for each i, $(X_{\tau_i \wedge t})$ is a martingale.

3.1.2 Stochastic integral with respect to Brownian motion

Consider the simple predictable process ϕ defines in equation (3.4). Then, we can define the Brownian stochastic integral as:

$$\int_0^T \phi_t dW_t = \sum_{=0}^{n} \phi_i (W_{T_{i+1}} - W_{T_i}) \tag{3.11}$$

Proposition 3.4 (Isometry formula) (proposition 8.5 in [2]) *Let $(\phi_t)_{0\leq t\leq T}$ be a simple predictable process and $(W_t)_{0\leq t\leq T}$ be a Wiener process. Then:*

i.

$$E\left[\int_0^T \phi_t dW_t\right] = 0, \tag{3.12}$$

ii.

$$E\left[\left|\int_0^T \phi_t dW_t\right|^2\right] = E\left[\int_0^T |\phi_t|^2 dt\right] \tag{3.13}$$

Proof

The first equation is easy to prove since W_t is a martingale, then also $\int_0^t \phi dW$ is a martingale. Therefore, $E[\int_0^t \phi dW] = 0$.

To show the second equation, we compute the second moment due to the independent of increments in W. Therefore, we have:

$$
\begin{aligned}
E\left[\left|\int_0^T \phi_t dW_t\right|^2\right] &= Var\left(\sum_{=0}^{n} \phi_i (W_{T_{i+1}} - W_{T_i})\right) \\
&= \sum_{=0}^{n} E[\phi_i^2 (W_{T_{i+1}} - W_{T_i})^2] \\
&\quad + 2\sum_{>j} Cov(\phi_i (W_{T_{i+1}} - W_{T_i}), \phi_j (W_{T_{j+1}} - W_{T_j})) \\
&= \sum_{=0}^{n} E[E[\phi_i^2 (W_{T_{i+1}} - W_{T_i})^2 | \mathcal{F}_{T_i}]] \\
&\quad + 2\sum_{>j} E[E[\phi_i \phi_j (W_{T_{i+1}} - W_{T_i})(W_{T_{j+1}} - W_{T_j}) | \mathcal{F}_{T_j}]] \\
&= \sum_{=0}^{n} E[E[\phi_i^2 (W_{T_{i+1}} - W_{T_i})^2 | \mathcal{F}_{T_i}]] + 0 \\
&= \sum_{i=0}^{n} E[\phi_i^2](T_{i+1} - T_i) = E\left[\int_0^t \phi_t^2 dt\right]
\end{aligned}
$$

□

We can use the isometry formula to built stochastic integrals with respect to the Wiener process for predictable processes. We need that the predictable processes $(\phi_t)_{t\in[0,T]}$ verify:

$$E\left[\int_0^T |\phi_t|^2 dt\right] < \infty$$

$$E\left[\int_0^T |\phi_t^n - \phi_t|^2 dt\right] \underset{n\to\infty}{\longrightarrow} 0.$$

Therefore, we have the following proposition for Brownian integrals:

Proposition 3.5 (Isometry formula for Brownian integrals) (proposition 8.6 in [2]) *Let $(\phi_t)_{0\leq t\leq T}$ be a predictable process which satisfy:*

$$E\left[\int_0^T |\phi_t|^2 dt\right] < \infty$$

Then, $\int_0^t \phi dW_t$ is a square integrable martingale and

i.

$$E\left[\int_0^T \phi_t dW_t\right] = 0,$$

ii.

$$E\left[\left|\int_0^T \phi_t dW_t\right|^2\right] = E\left[\int_0^T |\phi_t|^2 dt\right]$$

We can note that ϕ can not be interpreted as a "trading strategy" even if $\int_0^t \phi dW_t$ is a well-defined random variable. Moreover, its integral can not be represented as a limit of Riemann sums, which can be defined with the following proposition:

Proposition 3.6 (Stochastic integral via Riemann sums) (proposition 8.4 in [2]) *Let S be a semimartingale, ϕ be a càdlàg process and $\pi^n = (T_0^n = 0 < T_1^n < \cdots < T_{n+1}^n = T)$*

a sequence of random partitions of $[0,T]$ *such that* $|\pi^n| = \sup_k |T_k^n - T_{k-1}^n| \to 0$ *a.s when* $n \to \infty$. *Then:*

$$\phi_0 S_0 + \sum_{k=0}^{n} \phi_{T_k}(S_{T_{k+1}\wedge t} - S_{T_k\wedge t}) \xrightarrow[n\to\infty]{\mathbb{P}} \int_0^t \phi_{u^-} dS_u \tag{3.14}$$

uniformly in t *on* $[0,T]$.
We can note that in the sum the variation of S is multiplied by the value of ϕ at the left endpoint of the interval. We use the stochastic integrals via Riemann sums when we want make a stochastic integrals for càglàd processes.

3.1.3 Quadratic variation and covariation

Consider a process observed on a time grid $\pi = (t_0 = 0 < t_1 < \cdots < t_{n+1} = T)$, then we can define the realized variance as:

$$V_X(\pi) = \sum_{t_i \in \pi} (X_{t_{i+1}} - X_{t_i})^2$$

We can rewrite the realized variance as a Riemann sum:

$$V_X(\pi) = X_T^2 - X_0^2 - 2\sum_{t_i\in\pi} X_{t_i}(X_{t_{i+1}} - X_{t_i})$$

If X is a semimartingale with $X_0 = 0$, it will be a nonanticipating right-continuous process with left limits. Therefore, we can define the càdlàg process $X_- = (X_{t^-})_{t\in[0,T]}$. We can note that the Riemann sum defined above converge in probability to a random variable: the quadratic variation. Hence, we can give the following definition of the quadratic variation process:
Definition 3.3 [Quadratic Variation] (definition 8.3 in [2]) *The quadratic variation process of a semimartingale* X *is the nonanticipating càdlàg process defined by:*

$$[X,X]_t = |X_t|^2 - 2\int_0^t X_{u^-} dX_u \tag{3.15}$$

Is important to specify that the quadratic variation is a random variable and not a number. Moreover, if $\pi^n = (t_0^n = 0 < t_1^n < \cdots < t_{n+1}^n = T)$ is a sequence of partitions of $[0,T]$ such that $|\pi^n| = \sup_k |t_k^n - t_{k-1}^n| \to 0$ as $n \to \infty$, then

$$\sum_{t_i \pi^n}^{0\le t_i<t} (X_{t_{i+1}} - X_{t_i})^2 \xrightarrow[n\to\infty]{\mathbb{P}} [X,X]_t$$

where the convergence is uniform in t. Then, the following proposition summarizes some properties of the quadratic variation:
Proposition 3.7 (Properties of quadratic variation) *Consider* $[X,X]_t = |X_t|^2 - 2\int_0^t X_{u^-} dX_u$. *Then, we have the following properties:*

a) $([X,X]_t)_{t\in[0,T]}$ *is an increasing process. This allows to define integrals* $\int_0^t \phi d[X,X]$;

b) the jumps of $[X,X]$ *are related to the jumps of* X *by:* $\Delta[X,X]_t = |\Delta X_t|^2$. *In particular,* $[X,X]$ *has continuous sample paths if and only if* X *does;*

c) if X *is continuous and has finite variation, then* $[X,X] = 0$;

d) if X *is a martingale and* $[X,X] = 0$ *then* $X = X_0$ *almost surely.*

Proof
We give a proof for the property a) and c).
a). Since $[X,X]$ is defined as a limit of a positive sum $[X,X]_t \geq 0$, for $t > s$ and since $[X,X]_t - [X,X]_s$ is again a limit of positive sums $[X,X]_t \geq [X,X]_s$. We can conclude that

$[X, X]$ is an increasing process.
c). If X is continuous and has paths of finite variation, we obtain:

$$\begin{aligned}
\sum_{t_i \pi^n} (X_{t_{i+1}} - X_{t_i})^2 &\le \sup_i |X_{t_{i+1}} - X_{t_i}| \sum_{t_i \in \pi} |X_{t_{i+1}} - X_{t_i}| \\
&\le \sup_i |X_{t_{i+1}} - X_{t_i}| TV(X) \underset{|\pi| \to 0}{\longrightarrow} 0
\end{aligned}$$

where $TV(X)$ is the total variation of X on $[0, T]$. Therefore, $[X, X] = 0$. In particular, for a smooth (C^1) function, $[f, f] = 0$. Moreover, this result is no longer true for processes with discontinuous sample paths since, in this case, $|X_{t_{i+1}} - X_{t_i}|$ will not go to zero when $|X_{t_{i+1}} - X_{t_i}| \to 0$.

□

The property d) imply that if we have a continuous square-integrable martingale with path of finite variation, it will be constant with probability 1. This implication allows to say that if a process is decomposed into the sum of a square-integrable martingale term and a continuous process with finite variation (i.e. $X_t = M_t + \int_0^t a(t)dt$), then this decomposition is unique.

Consider a Brownian motion $B_t = \sigma W_t$, where W is a standard Wiener process, then the quadratic variation of the Brownian motion is equal to $[B, B]_t = \sigma^2 t$. We can note that the quadratic variation is equal to the variance of the process in the Brownian motion. To prove this statement, consider a sequence of partitions of $[0, T]$, i.e. $\pi^n = (t_0^n = 0 < t_1^n < \cdots < t_{n+1}^n = T)$, such that $|\pi^n| = \sup_k |t_k^n - t_{k-1}^n| \to 0$. We can see that $V_B(\pi^n) - \sigma^2 T = \sum_{\pi^n} (B_{t_{i+1}} - B_{t_i})^2 - \sigma^2(t_{i+1} - t_i)$ is a sum of independent terms with mean zero. Therefore, we have:

$$\begin{aligned}
E\left[|V_B(\pi^n) - \sigma^2 T|^2\right] &= \sum_{\pi^n} E\left[(B_{t_{i+1}} - B_{t_i})^2 - \sigma^2(t_{i+1} - t_i)\right]^2 \\
&= \sum_{\pi^n} \sigma^4 |t_{i+1} - t_i|^2 E\left[\left(\frac{(B_{t_{i+1}} - B_{t_i})^2}{\sigma^2(t_{i+1} - t_i)} - 1\right)^2\right] \\
&= \sigma^4 \sum_{\pi^n} |t_{i+1} - t_i|^2 E[(Z^2 - 1)^2] \quad where\, Z \sim N(0, 1) \\
&\le E[(Z^2 - 1)^2 \sigma^4 T |\pi^n|] \to 0
\end{aligned}$$

Hence, $E\left[|V_B(\pi^n) - \sigma^2 T|^2\right] \to 0$ implies convergence in probability of $V_B(\pi^n)$ to $\sigma^2 T$.
On the other hand, if we consider a Lévy process X with characteristic triplet (σ^2, ν, γ), the quadratic variation is equal to:

$$\begin{aligned}
[X, X]_t &= \sigma^2 t + \sum_{s \in [0,t], \Delta X_s \neq 0} |\Delta X_s|^2 \\
&= \sigma^2 t + \int_{[0,t]} \int_{\mathbb{R}} y^2 J_X(ds \times dy)
\end{aligned}$$

In the quadratic variation, we consider only one process X but, in the reality, we can see more stochastic process. Therefore, we need to introduce the multidimensional counterpart of the realized volatility: the realized covariance. Consider a time grid $\pi = (t_0 = 0 < t_1 < \cdots < t_{n+1} = T)$ and two process X and Y. Then, we can define the realized covariance as:

$$\sum_{t_i \in \pi} (X_{t_{i+1}} - X_{t_i})(Y_{t_{i+1}} - Y_{t_i}) \tag{3.16}$$

As we do for the realized variance, we can rewrite the sum above as a Riemann sum and we find:

$$X_T Y_T - X_0 Y_0 - \sum_{t_i \in \pi} \left(Y_{t_i}(X_{t_{i+1}} - X_{t_i}) + X_{t_i}(Y_{t_{i+1}} - Y_{t_i})\right)$$

If X, Y are semimartingale, the expression above convergence in probability to the random variable called quadratic covariation, which can be defined as:
Definition 3.4 [Quadratic Covariation] (definition 8.4 in [2]) *Given two semimartingales X, Y. The quadratic covariation process $[X, Y]$ is the semimartingale defined by:*

$$[X,Y]_t = X_t Y_t - X_0 Y_0 - \int_0^t X_{s^-} dY_s - \int_0^t Y_{s^-} dX_s \tag{3.17}$$

Consider the quadratic covariance defined in the expression (3.16). It discrete approximations converge in probability to $[X, Y]$ uniformly on $[0, T]$. Therefore, we have that:

$$\sum_{t_i \in \pi}^{t_i < t} (X_{t_{i+1}} - X_{t_i})(Y_{t_{i+1}} - Y_{t_i}) \xrightarrow[|\pi| \to 0]{\mathbb{P}} [X,Y]_t \tag{3.18}$$

The following proposition summarizes some important properties of the quadratic covariation:
Proposition 3.8 (Properties of the quadratic covariation) *Consider the quadratic covariation $[X,Y]_t = X_t Y_t - X_0 Y_0 - \int_0^t X_{s^-} dY_s - \int_0^t Y_{s^-} dX_s$. Then, we have the following properties:*

a) $[X, Y]$ is a nonanticipating càdlàg process with path of finite variation;

b) Polarization identity: $[X, Y] = \frac{1}{4}([X + Y, X + Y] - [X - Y, X - Y])$;

c) The quadratic covariation $[X, Y]$ is not modified if we add to X or Y continuous processes with finite variation, i.e random drift terms. It is only sensitive to the martingale parts, i.e. noise terms, or jumps in X and Y;

d) If X, Y are semimartingales and ϕ, ψ are integrable predictable processes then:

$$\left[\int \phi dX, \int \psi dY\right]_t = \int_0^t \phi\psi d[X,Y];$$

e) Product differentiation rule: if X, Y are semimartingales, then:

$$X_t Y_t = X_0 Y_0 + \int_0^t X_{s^-} dY_s + \int_0^t Y_{s^-} dX_s + [X,Y]_t$$

Consider two Brownian motion: $B_t^1 = \sigma_1 W_t^1$ and $B_t^2 = \sigma_2 W_t^2$, where W^1, W^2 are two standard wiener processes with correlation ρ (typically, with differential notation we define the correlation between two standard Wiener process as: $dW^1 dW^2 = \rho dt$. Hence, we can show that:

$$\begin{aligned}
[B_t^1, B_t^2]_t &= \frac{1}{4}\left([B_t^1 + B_t^2, B_t^1 + B_t^2] - [B_t^1 - B_t^2, B_t^1 - B_t^2]\right) \\
&= \frac{1}{4}\left([\sigma_1 W_t^1 + \sigma_2 W_t^2, \sigma_1 W_t^1 + \sigma_2 W_t^2] - [\sigma_1 W_t^1 - \sigma_2 W_t^2, \sigma_1 W_t^1 - \sigma_2 W_t^2]\right) \\
&= \frac{1}{4}\left((\sigma_1 W_t^1)^2 + 2\sigma_1\sigma_2 W_t^1 W_t^2 + (\sigma_2 W_t^2)^2 - (\sigma_1 W_t^1)^2 + 2\sigma_1\sigma_2 W_t^1 W_t^2 - (\sigma_2 W_t^2)^2\right) \\
&= \frac{1}{4}\left(4\sigma_1\sigma_2 W_t^1 W_t^2\right) \\
&= \sigma_1\sigma_2 W_t^1 W_t^2 \\
&= \rho\sigma_1\sigma_2 t
\end{aligned}$$

3.2 Stochastic Integral with Jumps

3.2.1 Stochastic Integral with respect to Poisson process

Consider the relation, for the Poisson process, described in chapter 2, i.e. $\Delta X_t = Y_{N_t} \Delta N_t$. Then, we can define the stochastic integral of a stochastic process $(\phi_t)_{t \geq 0}$ with respect to

$(X_t)_{t\geq 0}$ by:

$$\int_0^T \phi_t dX_t = \int_0^T \phi_t Y_{N_t} dN_t := \sum_{k=1}^{N_T} \phi_{T_k} Y_k \tag{3.19}$$

The meaning of the above equation is that the value at time T of a portfolio containing a quantity ϕ_t of an asset at time t, whose price evolves according to random returns Y_k, generate capital gain or losses equal to $\phi_{T_k} Y_k$ at random times T_k.
Consider a compound Poisson process (Definition 2.2) $(X_t)_{t\geq 0}$, it admits stochastic integral representation equal to:

$$X_t = X_0 + \sum_{k=1}^{N_t} Y_k = Y_0 + \int_0^t Y_{N_s} dN_s$$

Proposition 3.9 (Smoothing formula) (proposition 15.9 in [6]) *Let $(\phi_t)_{t\geq 0}$ be a process adapted to the filtration generated by $(X_t)_{t\geq 0}$ and such that:*

$$E\left[\int_0^T |\phi_t| dt\right] < \infty$$

Then, the expected value of the compensated Poisson stochastic integral is equal to:

$$E\left[\int_0^T \phi_{t^-} dX_t\right] = E\left[\int_0^T \phi_{t^-} Y_t dN_t\right] = \lambda E[Y] E\left[\int_0^T \phi_t dt\right] \tag{3.20}$$

where ϕ_t is the left limit of the process.
The equation (3.20) holds only for the left limit of the process ϕ, otherwise if we consider the full process ϕ we can have arbitrage opportunities.
Proof
From chapter 2, we already known that the compensated compound Poisson process is a martingale, therefore the stochastic integral is also a martingale due to the adaptedness of $(\phi_t)_{t\geq 0}$ to the filtration generated by $(X_t)_{t\geq 0}$, which makes the process $(\phi_{t^-})_{t\geq 0}$ predictable (i.e. $\mathcal{F}_{t^-} := \sigma(X_s : s \in [0,t))$). In fact, we have:

$$E\left[\int_0^T \phi_{t^-} dX_t\right] = \int_0^T \phi_{t^-} d(X_s - \lambda E[Y]s) = \int_0^T \phi_{t^-} (Y_{N_s} dN_s - \lambda E[Y] ds)$$

Now, we need to show that the expectation of a martingale remains constant over time. Therefore:

$$\begin{aligned} 0 &= E\left[\int_0^T \phi_{t^-} d(X_t - \lambda E[Y]t)\right] \\ &= E\left[\int_0^T \phi_{t^-} dX_t\right] - \lambda E[Y]\left[\int_0^T \phi_{t^-} dt\right] \\ &= E\left[\int_0^T \phi_{t^-} dX_t\right] - \lambda E[Y]\left[\int_0^T \phi_t dt\right] \end{aligned}$$

□

Proposition 3.10 *Let $(\phi_t)_{t\geq 0}$ be a process adapted to the filtration generated by $(X_t)_{t\geq 0}$ and such that*

$$E\left[\int_0^T |\phi_t|^2 dt\right] < \infty$$

Then, the expected value of the squared compensated Poisson stochastic integral is equal to:

$$E\left[\left(\int_0^T \phi_{t^-}(dX_t - \lambda E[Y]dt)\right)^2\right] = \lambda E[|Y|^2]E\left[\int_0^T |\phi_t|^2 dt\right] \tag{3.21}$$

We can see that only the generic jump size Y is squared whereas the intensity of the jump, i.e. λ, is not.

Consider a counting process N_t with jump times T_i and with random variables observed at T_i describe by Y_i. Let X_t be a process defined by $X_t = \sum_{i=1}^{N_t} Y_i$, hence the quadratic variation of the process is equal to:

$$[X,X]_t = \sum_{s\le t}(\Delta X_s)^2 = \sum_{i=1}^{N_t} Y_i^2$$

We can note that the same formula holds for every finite variation process X. Moreover, the predictable quadratic variation of the process (i.e. "angle bracket") is the compensator of $[X,X]$, namely:

$$\langle X,X\rangle_t = \lambda t E[Y_1^2]$$

For the quadratic covariation we need to consider another counting process N_t^b, which has jump times T_j and random variables observed at T_j, described by Y_j^b. Then, we consider the process $Z_t = \sum_{j=1}^{N_t^b} Y_j^b$. Now, we make the assumption that X and Z have finite variation processes whose jumps times are almost surely disjoint, hence they did not jump at the same time, therefore the quadratic covariation is equal to:

$$[X,Z]_t = \sum_{s\le t}\Delta X_s \Delta Z_s = 0$$

The assumption of disjoint jumps is a strong assumption and we consider it only for the stock price behavior. In fact, if we consider the exchange rate we drop this assumption and we consider correlated jumps between the rate.

3.2.2 Stochastic Integral with respect to Poisson random measure

Consider a Poisson random measure [2] M on $[0,T]\times\mathbb{R}^d$ with intensity $\mu(dt\times dx)$. Let $\tilde{M}$ be the compensated random measure defined as the centered version of M: $\tilde{M}(A) = M(A) - \mu(A) = M(A) - E[M(A)]$, where $A \subset \mathbb{R}^d$.
We can define the simple predictable process with respect to the Poisson random measure as:

$$\phi(t,y) = \sum_{i=1}^{n}\sum_{j=1}^{m}\phi_{ij}\mathbf{1}_{(T_i,T_{i+1}]}(t)\mathbf{1}_{A_j}(y) \tag{3.22}$$

where $\phi : \Omega\times[0,T]\times\mathbb{R}^d \to \mathbb{R}$ is a simple predictable functions, $(\phi_{ij})_{j=1,\dots,m}$ are bounded $\mathcal{F}_{T_i}$-measurable random variables, $T_1 \le T_2 \le \cdots \le T_n$ are non anticipating random times and $(A_j)_{j=1,\dots,m}$ are disjoint subsets of $\mathbb{R}^d$ with $\mu([0,T]\times A_j) < \infty$. The disjoint subset implies that the compensated random measure is a martingale with respect to A_j and that if $A\cap B = \emptyset$, then $M_t(A)$ and $M_t(B)$ are independent.
Now, we can define the stochastic integral with respect to Poisson process as the random variable:

$$\begin{aligned}\int_0^T\int_{\mathbb{R}^d}\phi(s,y)M(ds\times dy) &= \sum_{i,j=1}^{n,m}\phi_{ij}M\left((T_i,T_{i+1}]\times A_j\right)\\ &= \sum_{i=1}^{n}\sum_{j=1}^{m}\phi_{ij}[M_{T_{i+1}}(A_j) - M_{T_i}(A_j)]\end{aligned} \tag{3.23}$$

[2]Defined in chapter 2

If we want that the the stochastic integral is a càdlàg, nonanticipating process, we need to define the process $t \mapsto \int_0^t \int_{\mathbb{R}^d} \phi(t,y)M(dt \times dy)$ by:

$$\int_0^t \int_{\mathbb{R}^d} \phi(s,y)M(ds \times dy) = \sum_{i,j=1}^{n,m} \phi_{ij}[M_{T_{i+1}\wedge t}(A_j) - M_{T_i \wedge t}(A_j)]$$

Similarly, we can define the stochastic integral with respect to the compensated Poisson process, which can be defined as the random variable:

$$\begin{aligned}\int_0^T \int_{\mathbb{R}^d} \phi(s,y)\tilde{M}(ds \times dy) &= \sum_{i,j=1}^{n,m} \phi_{ij}\tilde{M}\left((T_i, T_{i+1}] \times A_j\right) \\ &= \sum_{i=1}^{n}\sum_{j=1}^{m} \phi_{ij}[M\left((T_i, T_{i+1}] \times A_j\right) - \mu\left((T_i, T_{i+1}] \times A_j\right)] \quad (3.24)\end{aligned}$$

As we do for the equation (3.23), the equation (3.24) can be written in stopping times notation by restricting the terms with $T_i \leq t$, therefore we obtain a stochastic process equal to:

$$\int_0^t \int_{\mathbb{R}^d} \phi(s,y)\tilde{M}(ds \times dy) = \sum_{i,j=1}^{n,m} \phi_{ij}[\tilde{M}_{T_{i+1}\wedge t}(A_j) - \tilde{M}_{T_i \wedge t}(A_j)] \quad (3.25)$$

We have introduced the stochastic integral with respect to the compensated Poisson process because we use it to show the martingale preserving property.

Proposition 3.11 (Martingale preserving property) (proposition 8.7 in [2]) *For any simple predictable function $\phi : \Omega \times [0,T] \times \mathbb{R}^d \to \mathbb{R}$ the process $(X_t)_{t\in[0,T]}$ defined by the compensated integral:*

$$X_t = \int_0^t \int_{\mathbb{R}^d} \phi(s,y)\tilde{M}(ds \times dy)$$

is a square integrable martingale and verifies the isometry formula:

$$E[|X_t|^2] = E\left[\int_0^t \int_{\mathbb{R}^d} |\phi(s,y)|^2 \mu(ds \times dy)\right] \quad (3.26)$$

Proof

First, define a process Y_t^j, with $j = 1, \ldots, m$, as $Y_t^j = \tilde{M}\left((0,t] \times A_j\right) = \tilde{M}_t(A_j)$. We know that $(Y_t^j)_{t\in[0,T]}$ is a martingale with independent increments and that the process Y^j are mutually independent since A_j are disjoint. Therefore, we can write the argument inside the sum in the equation (3.25), (i.e.$[\tilde{M}_{T_{i+1}\wedge t}(A_j) - \tilde{M}_{T_i\wedge t}(A_j)]$), as $Y^j_{T_{i+1}\wedge t} - Y^j_{T_i \wedge t}$. Hence, we find that the compensated integral X_t is equal to:

$$\begin{aligned}X_t &= \sum_{i,j=1}^{n,m} \phi_{ij}(Y^j_{T_{i+1}\wedge t} - Y^j_{T_i\wedge t}) \\ &= \sum_{j=1}^{m}\sum_{i=1}^{n} \phi_{ij}(Y^j_{T_{i+1}\wedge t} - Y^j_{T_i\wedge t}) \\ &= \sum_{j=1}^{m} \int_0^t \phi^j dY^j\end{aligned}$$

where $\phi^j = \sum_{i=1}^n \phi_{ij}\mathbf{1}_{(T_i, T_{i+1}]}$. We can note that ϕ^j is a simple predictable process, therefore its stochastic integral $(\int_0^t \phi^j dY^j)$ is a martingale by the martingale preserving property (proposition 3.1), which allow us to conclude that also X_t is a martingale. Now, we need to compute the first two moment of X_t: the mean and the variance.

$$\begin{aligned}
E[X_t] &= E\left[\sum_{j=1}^{m}\int_0^t \phi^j dY^j\right] \\
&= \sum_{j=1}^{m} E\left[\int_0^t \phi^j dY^j\right] \\
&= \sum_{j=1}^{m} E\left[E\left[\int_0^t \phi^j dY^j \,\middle|\, \mathcal{F}_{T_i}\right]\right] \\
&= 0
\end{aligned}$$

$$\begin{aligned}
E[|X_t|^2] &= Var\left(\int_0^t \int_{\mathbb{R}^d} \phi(s,y)\tilde{M}(ds \times dy)\right) \\
&= \sum_{i,j=1}^{n,m} E\left[|\phi_{ij}|^2 (Y^j_{T_{i+1}\wedge t} - Y^j_{T_i \wedge t})^2\right] \\
&= \sum_{i,j=1}^{n,m} E\left[E\left[|\phi_{ij}|^2 (Y^j_{T_{i+1}\wedge t} - Y^j_{T_i \wedge t})^2 |\mathcal{F}_{T_i}\right]\right] \\
&= \sum_{i,j=1}^{n,m} E\left[|\phi_{ij}|^2 E\left[(Y^j_{T_{i+1}\wedge t} - Y^j_{T_i \wedge t})^2 |\mathcal{F}_{T_i}\right]\right] \\
&= \sum_{i,j=1}^{n,m} E\left[|\phi_{ij}|^2 \mu\left([T_i, T_{i+1}] \times A_j\right)\right]
\end{aligned}$$

which yields (3.26). Finally, we can say that X_t is a square integrable martingale because $E[|X_t|^2] \leq E[|X_T|^2] \leq \infty$.

□

We can extend the isometry formula (equation (3.26)) to square integrable predictable functions and we have the following proposition:

Proposition 3.12 (Compensated Poisson integrals) (proposition 8.8 in [2]) *For any predictable random function $\phi : \Omega \times [0,T] \times \mathbb{R}^d \to \mathbb{R}$ verifying*

$$E\left[\int_0^t \int_{\mathbb{R}^d} |\phi(s,y)|^2 \mu(ds \times dy)\right] < \infty$$

the following property hold:

- $t \mapsto \int_0^t \int_{\mathbb{R}^d} \phi(s,y)\tilde{M}(ds \times dy)$ *is a square integrable martingale;*
- $$E\left[\left|\int_0^t \int_{\mathbb{R}^d} \phi(s,y)\mu(ds \times dy)\right|^2\right] = E\left[\int_0^t \int_{\mathbb{R}^d} |\phi(s,y)|^2 \mu(ds \times dy)\right] \qquad (3.27)$$

Consider a Lévy process $(X_t)_{t\geq 0}$ with Lévy measure ν and a Poisson random measure J_X with intensity $\mu(dt \times dx) = dt\nu(dx)$. Then, for a predictable random function ϕ the integral in equation (3.23) is equal to:

$$\int_0^T \int_{\mathbb{R}^d} \phi(s,y)M(ds \times dy) = \sum_{t\in[0,T]}^{\Delta X_t \neq 0} \phi(t, \Delta X_t)$$

The meaning of the above equation is that the integral of the predictable function is a sum of terms involving jump times (t) and jump sizes (ΔX_t).

Let M be a Poisson random measure on $[0,T] \times \mathbb{R}^d$ with intensity $\mu(ds \times dy)$ and let $\phi : [0,T] \times \mathbb{R}^d \mapsto \mathbb{R}$. Then, we can define the process X as the integral of ϕ with respect to M:

$$X_t = \int_0^t \int_{\mathbb{R}^d} \phi(s,y) M(ds \times dy)$$

Therefore, the quadratic variation of X is equal to:

$$[X,X]_t = \int_0^t \int_{\mathbb{R}^d} |\phi(s,y)|^2 M(ds \times dy)$$

Now, consider a Wiener process $(W_t)_{t\in[0,T]}$ independent from M and two process X. Hence, the process X_t can be written as

$$X_t^i = X_0^i + \int_0^t \sigma_s^i dW_s + \int_0^t \int_{\mathbb{R}^d} \phi^i(s,y) M(ds \times dy) \qquad i = 1,2$$

Hence, the quadratic covariation of the two process is equal to:

$$[X^1, X^2]_t = \int_0^t \sigma_s^1 \sigma_s^2 ds + \int_0^t \int_{\mathbb{R}^d} \phi^1(s,y)\phi^2(s,y) M(ds \times dy) \qquad i = 1,2$$

3.3 Change of variable formula

In this section we talk about the change of variable formula for the jump processes. But first, we need to remind the change of variables formula for smooth functions and the Ito formula for Brownian integrals. Let $f : \mathbb{R} \to \mathbb{R}$ be a C^2 function and let $g : [0,T] \to \mathbb{R}$ be a C^1 function. Then, the change of variables formula for smooth function is:

$$f(g(t)) - f(g(0)) = \int_0^t f'(g(s))g'(s)ds = \int_0^t f'(g(s))dg(s) \tag{3.28}$$

Now, we can consider the Brownian integral defined as: $X_t = \int_0^t \sigma_s dW_s$ and the function f defined as above. Then, if we apply the Ito formula at X_t we find:

$$f(X_t) = f(0) + \int_0^t f'(X_s)\sigma_s dW_s + \frac{1}{2}\int_0^t f''(X_s)ds \tag{3.29}$$

3.3.1 Calculus for finite jump processes

Let $x : [0,T] \to \mathbb{R}$ be a function with a finite number of discontinuities at time $0 = T_0 \leq T_1 \leq T_2 \leq \cdots \leq T_n \leq T_{n+1} = T$ and the function x is smooth on each interval, defines as (T_i, T_{i+1}). Moreover, we can define $x(T_i) := x(T_i^+)$, which means that x is càdlàg at the discontinuity points. Let $f : \mathbb{R} \to \mathbb{R}$ be a C^1 function. Since x is smooth on each interval (T_i, T_{i+1}), $f(x(t))$ is also smooth. Then, the change of variable formula for piecewise smooth functions is given by the following proposition:

Proposition 3.13 (Change of variable formula for piecewise smooth functions) (proposition 8.12 in [2]) *If x is a piecewise C^1 function given by:*

$$x(t) = \int_0^t b(s)ds + \sum_{\{i.T_i \leq t\}} \Delta x_i \quad with\, i = 1, \ldots, n+1$$

where $\Delta x_i = x(T_i) - x(T_i^-)$. Then, for every C^1 function $f : \mathbb{R} \to \mathbb{R}$ we have:

$$\begin{aligned} f(x(T)) - f(x(0)) &= \int_0^T b(t) f'(x(t^-))dt \\ &+ \sum_{i=1}^{n+1} \left(f(x(T_i^-) + \Delta x_i) - f(x(T_i^-)) \right) \end{aligned} \tag{3.30}$$

Proof
Consider the function x represented by: $x(t) = \int_0^t b(s)ds + \sum_{\{i.T_i \leq t\}} \Delta x_i$. We apply the change of variable formula for smooth function, ie. equation (3.28), and we find:

$$f(x(T_{i+1}^-)) - f(x(T_i)) = \int_{T_i}^{T_{i+1}^-} f'(x(t))x'(t)dt = \int_{T_i}^{T_{i+1}^-} f'(x(t))b(t)dt$$

Now, we need to study what happen at each discontinuity point at the function $f(x(t))$, which has jumps equal to:

$$f(x(T_i)) - f(x(T_i^-) = f(x(T_i^-) + \Delta X_i) - f(x(T_i^-))$$

Therefore, we can write the variation of f between 0 and t as:

$$\begin{aligned}
f(x(T)) - f(x(0)) &= \sum_{i=0}^{n} (f(x(T_{i+1})) - f(x(T_i))) \\
&= \sum_{i=0}^{n} \big(f(x(T_{i+1})) + f(x(T_{i+1}^-)) + f(x(T_{i+1}^-)) - f(x(T_i))\big) \\
&= \sum_{i=0}^{n+1} \left(\underbrace{f(x(T_i^-) + \Delta X_i) - f(x(T_i^-))}_{Jump\, part} \right) + \sum_{i=0}^{n} \int_{T_i}^{T_{i+1}^-} b(t)f'(x(t))dt \\
&= \int_0^T b(t)f'(x(t^-))dt + \sum_{i=1}^{n+1} f(x(T_i^-) + \Delta x_i) - f(x(T_i^-))
\end{aligned}$$

□

We can note that if b is continuous, then x is piecewise C^1 and if $b = 0$, then x is piecewise constant and the integral term is equal to zero.
Now, consider a stochastic process $(X_t)_{t\in[0,T]}$ defined by:

$$X_t = \int_0^t b_s ds + \sum_{i=1}^{N_t} \Delta X_i$$

where $\Delta X_i := X(T_i) - X(T_i^-)$ represent the jump size and N_t is the random number of jumps. Then by the proposition above, the following change of variable formula holds almost surely:

$$\begin{aligned}
f(X_T) - f(X_0) &= \int_0^T b(t)f'(X_t)dt + \sum_{i,T_i \leq t} \Big(f(X_{T_i^-} + \Delta X_i) - f(X_{T_i^-})\Big) \\
&= \int_0^T b(t^-)f'(X_{t^-})dt + \sum_{0\leq t \leq T}^{\Delta X_t \neq 0} \Big(f(x(T_i^-) + \Delta X_t) - f(X_{t_i^-})\Big)
\end{aligned}$$

We can note that this change of variable formula is valid independently of the probabilistic structure of the process X. Moreover, the following proposition summarized the Ito formula for finite activity jump process where the counting process N_t is a martingale:
Proposition 3.14 (Ito formula for finite activity jump processes) (proposition 8.13 in [2]) *Let X be a jump process with values in $\mathbb{R}$ defined by:*

$$X_t = \int_0^t b_s ds + \sum_{i=1}^{N_t} Y_i$$

where b_s is a nonanticipating càdlàg process, N_t is a counting process representing the number of jumps between 0 and t and Y_i is the size of the i-th jump. Denote by $(T_n)_{n\geq 1}$ the jump *times of X_t and J_X the random measure on $[0,T]\times\mathbb{R}$ associated to the jumps of X*[3] *Then, for any measurable function $f:[0,T]\times\mathbb{R}\to\mathbb{R}$ we have:*

$$
\begin{aligned}
f(t,X_t)-f(0,X_0) &= \int_0^t \left(\frac{\partial f}{\partial s}(s,X_{s^-})+b_s\frac{\partial f}{\partial x}(s,X_{s^-})ds\right)\\
&+\sum_{n\geq 1, T_n\leq T}\left(f(s,X_{s^-}+\Delta X_s)-f(s,X_{s^-})\right)\\
&= \int_0^t \left(\frac{\partial f}{\partial s}(s,X_{s^-})+b_s\frac{\partial f}{\partial x}(s,X_{s^-})ds\right)\\
&+\int_0^t\int_{-\infty}^{\infty}\left(f(s,X_{s^-}+y)-f(s,X_{s^-})\right)J_X(ds\times dy) \qquad (3.31)
\end{aligned}
$$

Moreover, if N_t is a Poisson process with $E[N_t]=\lambda t$, with $Y_i\sim F$ are i.i.d. and f is boundend, then $Y_t=f(t,Y_t)=V_t+M_t$, where M is the martingale or noise component and V is the continuous finite variation drift. This two component are respectively equal to:

$$
\begin{aligned}
M_t &= \int_0^t\int_{-\infty}^{\infty}\left(f(s,X_{s^-}+y)-f(s,X_{s^-})\right)\tilde{J}_X(ds\times dy) \qquad (3.32)\\
& \text{where } \tilde{J}_X(ds\times dy)=J_X(dt\times dy)-\lambda F(dy)dt\\
V_t &= \int_0^t \left(\frac{\partial f}{\partial s}(s,X_{s^-})+b_s\frac{\partial f}{\partial x}(s,X_{s^-})ds\right)\\
&+\int_0^t ds\int_{\mathbb{R}^d}F(dy)\left(f(s,X_{s^-}+y)-f(s,X_{s^-})\right) \qquad (3.33)
\end{aligned}
$$

3.3.2 Ito formula for jump-diffusion and Lévy process

Consider a jump-diffusion process defined in chapter 2 by the equation (2.19) (i.e. $X_t=\gamma t+\sigma W_t+\sum_{i=1}^{N_t}Y_i$). We can write this process with a different notation:

$$
X_t=\underbrace{\gamma t+\sigma W_t}_{X^c(t)}+\underbrace{\sum_{i=1}^{N_t}\Delta X_i}_{J_t}
$$

where $\Delta X_i := X(T_i)-X(T_i^-)$. Therefore, the equation above can be write as:

$$
X_t=X^c(t)+J_t \qquad (3.34)
$$

Let f be a C^2 function on $\mathbb{R}$ and let T_i, $i=1,\dots,N_T$, be the jump times of X. Then, we can define $Y_t=f(X_t)$ and we can say that X, between T_i and T_{i+1}, evolves according to the differential equation equal to:

$$
dX_t=dX_t^c=\gamma dt+\sigma dW_t
$$

therefore, by applying the Ito formula in the Brownian case, which is described in equation (3.29), we find:

$$
\begin{aligned}
Y_{T_{i+1}^-}-Y_{T_i} &= \int_{T_i}^{T_{i+1}^-}\frac{\sigma^2}{2}f''(X_t)dt+\int_{T_i}^{T_{i+1}^-}f'(X_t)dX_t\\
&= \int_{T_i}^{T_{i+1}^-}\left(\frac{\sigma^2}{2}f''(X_t)dt+f'(X_t)dX_t^c\right)
\end{aligned}
$$

[3] J_X can be defined as: $J_X=\sum_{n\geq 1,T_n\leq T}\delta_{(T_n,Y_n)}$, where δ is the dirac measure (see appendix A.4)

since we consider the behavior of the function inside the interval (T_i, T_{i+1}), we have that $dX_t = dX_t^c$. Now, we need to analyze what happen at Y_t when a jump of size ΔX_t occurs. Hence, the change in Y_t is equal to: $f(X_{t^-} + \Delta X_t) - f(X_{t^-})$. Therefore, if we add these two contributions, we will find the total change in Y_t:

$$\begin{aligned} f(X_t) - f(X_0) = &\int_0^t f'(X_s)dX_s^c + \int_0^t \frac{\sigma^2}{2} f''(X_s)ds \\ &+ \sum_{0\le s\le t}^{\Delta X_s \neq 0} (f(X_{s^-} + \Delta X_s) - f(X_{s^-})) \end{aligned} \tag{3.35}$$

The equation (3.35) can be rewritten in a more general form. If we replace dX_s^c by $dX_s - \Delta X_s$, we find:

$$\begin{aligned} f(X_t) - f(X_0) = &\int_0^t f'(X_s)dX_s + \int_0^t \frac{\sigma^2}{2} f''(X_s)ds \\ &+ \sum_{0\le s\le t}^{\Delta X_s \neq 0} (f(X_{s^-} + \Delta X_s) - f(X_{s^-}) - \Delta X_s f'(X_{s^-})) \end{aligned} \tag{3.36}$$

This equation becomes equivalent to the equation (3.35), when the number of jumps is finite. Moreover, in (3.36) the stochastic integral and the sum over the jumps are well-defined for any semimartingale, even if we have an infinite number of jumps. Instead, the equation (3.35) could not converge if jumps have an infinite variation.
The following proposition summarized the result for the jump-diffusion processes when σ is a nonanticipating square-integrable process:
Proposition 3.15 (Ito formula for jump-diffusion processes) (proposition 8.14 in [2]) *Let X be a diffusion process with jumps defined as:*

$$X_t = X_0 + \int_0^t b_s ds + \int_0^t \sigma_s dW_s + \sum_{i=1}^{N_t} \Delta X_i$$

where $\sum_{i=1}^{N_t} \Delta X_i$ is a compound Poisson process and b_t and σ_t are continuous nonanticipating processes with satisfy the condition:

$$E\left[\int_0^t \sigma_t dt\right]^2 = E\left[\int_0^t \sigma_t^2 dt\right] < \infty$$

Then, for any $C^{1,2}$ function $f : [0,T] \times \mathbb{R} \to \mathbb{R}$, the process $Y_t = f(t, X_t)$ can be represented as:

$$\begin{aligned} f(t, X_t) - f(0, X_0) = &\int_0^t \left[\frac{\partial f}{\partial s}(s, X_s) + \frac{\partial f}{\partial x}(s, X_s) b_s\right] ds \\ &+ \frac{1}{2}\int_0^t \sigma_s^2 \frac{\partial^2 f}{\partial x^2}(s, X_s)ds + \int_0^t \frac{\partial f}{\partial x}(s, X_s)\sigma_s dW_s \\ &+ \sum_{i\ge 1, T_1 \le t} \left(f(X_{T_i^-} + \Delta X_i) - f(X_{T_i^-})\right) \end{aligned} \tag{3.37}$$

The equation (3.37) can be written in differential notation as:

$$\begin{aligned} dY_t = &\frac{\partial f}{\partial t}(t, X_t)dt + b_t \frac{\partial f}{\partial x}(t, X_t)dt + \frac{\sigma_t^2}{2}\frac{\partial^2 f}{\partial x^2}(t, X_t)dt \\ &+ \frac{\partial f}{\partial x}(t.X_t)\sigma_t dW_t + (f(X_{t^-} + \Delta X_t) - f(X_{t^-})) \end{aligned} \tag{3.38}$$

More in general, when we consider a Lévy process we can have an infinite number of jumps in each interval, which imply that the above result could be not true. Therefore, we need to study the conditions under which the sum in the equation (3.36) converges. Suppose that f and its two derivatives are bounded by a constant C, hence we can see that the sum in the equation (3.36) using proposition 2.14 is equal to:

$$|\left(f(X_{s^-}+\Delta X_s)-f(X_{s^-})-\Delta X_s f'(X_{s^-})\right)| \leq C\Delta X_s^2$$

which means that the sum in the equation (3.36) is finite. Therefore, we have the following proposition for the Lévy processes:
Proposition 3.16 (Ito formula for scalar Lévy processes) (proposition 8.15 in [2]) *Let $(X_t)_{t\geq 0}$ be a Lévy process with Lévy triplet (σ^2,ν,γ) and $f:\mathbb{R}\to\mathbb{R}$ a C^2 function. Then:*

$$\begin{aligned} f(X_t) = f(0) &+ \int_0^t f'(X_s)dX_s + \int_0^t \frac{\sigma^2}{2} f''(X_s)ds \\ &+ \sum_{0\leq s\leq t}^{\Delta X_s\neq 0} \left(f(X_{s^-}+\Delta X_s)-f(X_{s^-})-\Delta X_s f'(X_{s^-})\right) \end{aligned} \tag{3.39}$$

The equation (3.39) *can be written in differential notation as:*

$$df(X_t) = \frac{\sigma^2}{2} f''(X_t)dt + f'(X_{t^-})dX_t + f(X_t) - f(X_{t^-}) - \Delta X_t f'(X_{t^-})$$

If we have that the Lévy process is of finite variation, we do not need to subtract $\Delta X_s f'(X_{s^-})$ from each term of the sum in the equation (3.39). In this case, we have the following proposition that summarizes the Ito formula for Lévy processes with finite variation:
Proposition 3.17 (Ito formula for Lévy processes with finite variation jumps) (proposition 8.17 in [2]) *Let X be a finite variation Lévy process with characteristic exponent equal to:*

$$\psi_X(u) = ibu + \int_{-\infty}^{\infty} \left(e^{iuy}-1\right)\nu(dy)$$

where the Lévy measure ν verifies $\int |y|\nu(dy) < \infty$. Then, for any C^1 function $f:[0,T]\times\mathbb{R}\to\mathbb{R}$ we have:

$$\begin{aligned} f(t,X_t)-f(0,X_0) = &\int_0^t \left(\frac{\partial f}{\partial s}(s,X_{s^-}) + b\frac{\partial f}{\partial x}(s,X_{s^-})\right)ds \\ &+ \sum_{0\leq s\leq t}^{\Delta X_s\neq 0} \left(f(X_{s^-}+\Delta X_s)-f(X_{s^-})\right) \end{aligned} \tag{3.40}$$

If f and its first derivative in x are bounded, then $Y_t = f(t,Y_t)$ is the sum of a martingale part and a drift part. This two component are respectively equal to:

$$M_t = \int_0^t \int_{-\infty}^{\infty} \left(f(s,X_{s^-}+y)-f(s,X_{s^-})\right)\tilde{J}_X(ds\times dy) \tag{3.41}$$

$$\begin{aligned} V_t = &\int_0^t \left(\frac{\partial f}{\partial s}(s,X_{s^-}) + b_s\frac{\partial f}{\partial x}(s,X_{s^-})ds\right) \\ &+ \int_0^t ds \int_{\mathbb{R}} \nu(dy)\left(f(s,X_{s^-}+y)-f(s,X_{s^-})\right) \end{aligned} \tag{3.42}$$

3.3.3 Ito formula for martingale and semimartingale

If we use the change of variable formula in $Y_t = f(t,X_t)$, we find that the process Y_t is not anymore a Lévy process even if X_t is defined as Lévy process. One solution could be written

Y_t as a stochastic integral, which imply that the process Y_t is a semimartingale. Therefore, our new problem is find a change of variable formula for semimartingale. The following proposition shows us the Ito formula for semimartingales:

Proposition 3.18 (Ito formula for semimartingale) (proposition 8.19 in [2]) *Let $(X_t)_{t\geq 0}$ be a martingale and let $f : [0,T] \times \mathbb{R} \to \mathbb{R}$ be a C^2 function. Then, the Ito formula is equal to:*

$$\begin{aligned} f(t, X_t) - f(0, X_0) = &\int_0^t \frac{\partial f}{\partial s}(s, X_s)ds + \int_0^t \frac{\partial f}{\partial x}(s, X_{s^-})dX_s \\ &+ \frac{1}{2}\int_0^t \frac{\partial^2 f}{\partial x^2}(s, X_{s^-})d[X,X]^c_s \\ &+ \sum_{0\leq s\leq t}^{\Delta X_s \neq 0} \left(f(s, X_s) - f(s, X_{s^-}) - \Delta X_s \frac{\partial f}{\partial x}(s, X_{s^-})\right) \end{aligned} \tag{3.43}$$

where $[X,X]^c$ is the continuous part of the quadratic variation $[X,X]$, which it can be split into a jump part and a continuous part because it is an increasing process.

Proof

Consider a partition $T_0 = 0 < T_1 < \cdots < T_n < T_{n+1} = t$ and consider a second order Taylor expansion[4]. Then, we can write $f(X_t)$ as the sum of increments and we find:

$$f(X_t) - f(X_0) = \sum_{i=0}^{n} (f(X_{T_{i+1}}) - f(X_{T_i}))$$

then, we apply the second order Taylor expansion at the equation above and we find:

$$\begin{aligned} &\sum_{i=0}^{n} (f(X_{T_{i+1}}) - f(X_{T_i})) = \\ &\qquad = \sum_{i=0}^{n} f'(X_{t_i})(X_{T_{i+1}} - X_{T_i}) + \frac{1}{2}\sum_{i=0}^{n} f''(X_{t_i})(X_{T_{i+1}} - X_{T_i})^2 \\ &\qquad + \sum_{i=0}^{n} r(X_{T_{i+1}}, X_{T_i}) \end{aligned}$$

We can note that X has a well defined quadratic variation, hence $\sum \Delta X_s^2$ converges almost surely. Then, let $A \subset [0,T] \times \Omega$ such that $\sum_{0\leq s\leq t} \Delta X_s^2 < \varepsilon$ on A for $\varepsilon > 0$ and $B\{(s,\omega) \notin A, \Delta X_s \neq 0\}$. Therefore, the sum above can be rewritten as:

$$\begin{aligned} f(X_t) - f(X_0) = &\sum_{i=0}^{n} f'(X_{t_i})(X_{T_{i+1}} - X_{T_i}) + \frac{1}{2}\sum_{i=0}^{n} f''(X_{t_i})(X_{T_{i+1}} - X_{T_i})^2 \\ &+ \sum_{B\cap(T_i,T_{i+1})\neq 0} (f(X_{T_{i+1}}) - f(X_{T_i}) - f'(X_{t_i})(X_{T_{i+1}} - X_{T_i}) \\ &- \frac{1}{2} f''(X_{t_i})(X_{T_{i+1}} - X_{T_i})^2) + \sum_{B\cap(T_i,T_{i+1})\neq 0} r(X_{T_{i+1}}, X_{T_i}) \end{aligned}$$

Now, let $\sup |T_{i+1} - T_i| \to 0$ a.s, then the first three terms converge to the following expressions:

a)
$$\sum_{i=0}^{n} f'(X_{t_i})(X_{T_{i+1}} - X_{T_i}) \to \int_0^t f'(X_s)dX_s$$

[4] A second order Taylor expansion is equal to: $f(y) = f(x) + f'(x)(y-x) + \frac{f''(x)(y-x)^2}{2} + r(x,y)$

b)

$$\frac{1}{2}\sum_{i=0}^{n} f''(X_{t_i})(X_{T_{i+1}} - X_{T_i})^2 \to \frac{1}{2}\int_0^t f''(X_s)d[X,X]_s$$

c)

$$\sum_{B\cap(T_i,T_{i+1})\neq 0} (f(X_{T_{i+1}}) - f(X_{T_i}) - f'(X_{t_i})(X_{T_{i+1}} - X_{T_i}) - \frac{1}{2}f''(X_{t_i})(X_{T_{i+1}} - X_{T_i})^2)$$
$$\to \sum_B (f(X_s) - f(X_{s^-}) - \Delta X_s f'(X_{s^-}) - f''(X_{s^-})|\Delta X_s|^2)$$

the first two term (i.e a) and b)) are Riemann sum. Now, we need to analyze the last term: the remainder. We can note that the remainder verifies:

$$r(x,y) \leq (y-x)^2\alpha(|x-y|) \quad with \ \alpha(u) \to 0 \ as \ u \to 0$$

Moreover, the remainder's sum only contain term with $B \cap (T_i, T_{i+1}) \neq 0$, then we have that $|X_{T_{i+1}} - X_{T_i}| \leq \varepsilon$ when $\sup |T_{i+1} - T_i| \to 0$. Therefore, we have that:

$$\begin{aligned}\sum_{B\cap(T_i,T_{i+1})\neq 0} |r(X_{T_{i+1}}, X_{T_i})| &\leq \alpha(2\varepsilon)\sum_i (X_{T_{i+1}} - X_{T_i})^2 \\ &\leq \alpha(2\varepsilon)[X,X]_t \to 0 \quad as \ \sup|T_{i+1} - T_i| \to 0\end{aligned}$$

Hence, we sum all the terms together and we find:

$$\begin{aligned}f(X_t) - f(X_0) &= \int_0^t f'(X_s)dX_s + \frac{1}{2}\int_0^t f''(X_s)d[X,X]_s \\ &+ \sum_{0\leq s\leq t} (f(X_s) - f(X_{s^-}) - \Delta X_s f'(X_{s^-}) - f''(X_{s^-})|\Delta X_s|^2)\end{aligned}$$

and since:

$$\int_0^t f''(X_s)d[X,X]_s = \int_0^t f''(X_s)d[X,X]_s^c + \sum_{0\leq s\leq t} f''(X_{s^-})|\Delta X_s|^2$$

we finally obtain the equation (3.43).

$\square$

Chapter 4

Hedging Strategy

This chapter describe how to compute the option price in an exponential-Lévy model. The first section talk about the measure transformation, which represent the main tool to find the risk-neutral probability to compute the option pricing. The second part of the chapter introduces the concept of option and, in particular, of European call option. We show how is built the pricing of European option in the Black-Scholes model and, then, we define the pricing of European option on exponential-Lévy model. Moreover, we give an example of option pricing for a jump diffusion process. Then, we briefly describe the concept of implied volatility and we describe some feature. Finally, we will see how to use an European call option for hedging purpose. We start by describing the hedging in the Black-Scholes model, in particular, we will talk about the delta hedging. Then, we will introduce the Merton hedging for a jump diffusion process and we will compare the result with the hedging in the Black-Scholes model. Finally, we will generalize the hedging strategy for jump diffusion process, therefore we will introduce and explain the quadratic hedging in jump diffusion process. The last part of this chapter is entirely focused on empirical result and we compare the hedging in the Black and Scholes model with the hedging in the Merton model for the jump diffusion process

4.1 Measure Transformation

One normal assumption in each model built in finance is that the market is complete, which means that every contingent claim in the market is attainable. This imply that there is only one arbitrage-free way to value an option, which is a linear combination between a risky asset and a riskless one (i.e. is called the replicating portfolio). Hence, there exist only one risk-neutral probability in the market. Unfortunately, the complete market assumption is not true in the real market because the asset prices have jumps, which imply that there is not a unique risk-neutral probability but we can find a much greater variety of equivalent measure by changing the distribution of jumps. Therefore, the perfect hedges do not exists (i.e. the delta hedging in the Black-Scholes model) since it is impossible replicate an option by trading in the underlying asset due to the presence of jumps in the price behavior.

In the Black-Scholes model to find the equivalent measure we use the Radon-Nikodym theorem. Hence, we need to introduce the concempt of equivalent measure. Let $(\Omega, \mathcal{F})$ be a measurable space and let $\mathbb{Q}$, $\mathbb{P}$ be two probability measure on $\mathcal{F}$. Then, we say that $\mathbb{Q}$ is absolutely continuous respect to $\mathbb{P}$ ($\mathbb{P} \gg \mathbb{Q}$) if:

$$\forall A \in \mathcal{F} \qquad \mathbb{P}(A) = 0 \Rightarrow \mathbb{Q}(A) = 0$$

Therefore, we can say that two probability measure $\mathbb{Q}$, $\mathbb{P}$ on $\mathcal{F}$ are equivalent ($\mathbb{P} \sim \mathbb{Q}$) if $\mathbb{P} \gg \mathbb{Q}$ and $\mathbb{Q} \gg \mathbb{P}$, hence if $\mathbb{Q}$ and $\mathbb{P}$ define the same set of impossible events:

$$\forall A \in \mathcal{F} \qquad \mathbb{Q}(A) = 0 \Leftrightarrow \mathbb{P}(A) = 0 \tag{4.1}$$

Therefore, we have the following theorem:
Theorem 4.1 [Radon-Nikodym Theorem] *Let $\mathbb{P} \gg \mathbb{Q}$, then exist a random variable Λ, $\mathcal{F}$-measurable, with non-negative value such that for every random variable X ($\mathcal{F}$-measurable) integrable under $\mathbb{P}$ the following relation is true:*

$$E_{\mathbb{Q}}[X] = E_{\mathbb{P}}[\Lambda X] = \int_A \Lambda d\mathbb{P}$$

In particular:

$$\forall A \in \mathcal{F} \qquad \mathbb{Q}(A) = E_{\mathbb{P}}(\Lambda \mathbf{1}_A).$$

Λ is called the Radon-Nikodym derivative and, usually, it is written as:

$$\Lambda = \frac{d\mathbb{Q}}{d\mathbb{P}}$$

We can note that the variable Λ is unique. In fact, if Λ' is another Radon-Nikodym derivative, then we have the following equality:

$$E_{\mathbb{P}}[\Lambda - \Lambda'] = 0$$

Moreover, the Radon-Nikodym derivative satisfy the following relation:

$$E_{\mathbb{P}}[\Lambda] = E_{\mathbb{P}}\left[\frac{d\mathbb{Q}}{d\mathbb{P}}\right] = 1$$

Instead, $\mathbb{P}$ and $\mathbb{Q}$ are orthogonal is there exist an event A such that $\mathbb{P}(A) = 1$ and $\mathbb{Q}(A) = 0$. This imply that if $\mathbb{P}$ and $\mathbb{Q}$ are orthogonal, it will not be possible find that one probability measure is absolutely continuous respect to the other one.

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space which describe a market between 0 and T. Then, we can define the underlying asset S by a nonanticipating (càdlàg) process:

$$\begin{aligned} S : [0,T] \times \Omega &\mapsto \mathbb{R}^{d+1} \\ (t, \omega) &\mapsto (S_t^0(\omega), S_t^1(\omega), S_t^2(\omega), \dots, S_t^d(\omega)) \end{aligned}$$

where $S_t^i(\omega)$ represent the value of the asset i in the market scenario ω and $S_t^0(\omega)$ is a numeraire (we define it as $S_t^0(\omega) = e^{rt}$, where r is the interest rate) A self-financing strategy (ϕ_t^0, ϕ_t), in the Black-Scholes model, is said to be a perfect hedge or a replication strategy for a contingent claim H, if we have the following:

$$H = V_0 + \int_0^T \phi_t dS_t + \int_0^T \phi_t^0 dS_t^0 \qquad \mathbb{P} - a.s. \tag{4.2}$$

where S_t is the asset price. Moreover, we can say that a market is complete if any contingent claim H, admits a replicating portfolio which means that for any $H \in \mathcal{H}$ there exists a self-financing strategy (ϕ_t^0, ϕ_t) such that the equation (4.2) holds with probability 1 under $\mathbb{P}$. If the equation (4.2) holds with probability 1, it also holds with probability 1 under any equivalent martingale measure $\mathbb{Q} \sim \mathbb{P}$. Therefore, we find the following proposition:
Proposition 4.1 *A market defined by the asset $(S_t^0, S_t^1, \dots, S_t^d)_{0 \le t \le T}$ described as stochastic processes on $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ is complete if and only if there is a unique martingale measure $\mathbb{Q}$ equivalent to $\mathbb{P}$.*

If we consider a discount factor equal to $B(t,T) = e^{-r(T-t)}$, then we can write the discounted value of H (equation (4.2)) as:

$$\hat{H} = V_0 + \int_0^T \phi_t d\hat{S}_t \qquad \mathbb{Q} - a.s. \tag{4.3}$$

We can note that $V_0 = E_{\mathbb{Q}}[\hat{H}]$. If the above equation holds for all payoff with finite variance (i.e. $H \in L^2(\mathcal{F}_t, \mathbb{Q})$), then we can represented the above process as:

$$\hat{H} = E[H] + \int_0^T \phi_t d\hat{S}_t$$

for some predictable process ϕ, moreover the martingale $(\hat{S}_t)_{0 \leq t \leq T}$ is said to have the predicatable representation property. But this property did not hold for most discontinuous model used in finance. Therefore, we need to introduce a representation of $\hat{H}$ in terms of a stochastic integral with respect to $\hat{S}$, which is called the predictable representation with respect to W, M or predicatable representation property:

Proposition 4.2 (Predictable representation property) (proposition 9.4 in [2]) *Let $(W_t)_{0 \leq t \leq T}$ be a d-dimensional Wiener process and M a Poisson random measure on $[0, T] \times \mathbb{R}^d$, independent from W. Then, any random variable $\hat{H}$ with finite variance depending on the filtration $(\mathcal{F}_t)_{0 \leq t \leq T}$ of W and M between 0 and T can be represented as the sum of a constant, a stochastic integral with respect to W and a compensated Poisson integral with respect to M. There exists a predictable process $\phi : \Omega \times [0, T] \mapsto \mathbb{R}$ and a predictable random function $\psi : \Omega \times [0, T] \times \mathbb{R}^d \mapsto \mathbb{R}$ such that:*

$$\hat{H} = E[\hat{H}] + \int_0^t \phi_s dW_s + \int_0^t \int_{\mathbb{R}^d} \psi(s, y) \tilde{M}(ds \times dy) \tag{4.4}$$

In particular, equation (4.4) is important when we talk about hedging strategies.

4.1.1 Risk-neutral pricing

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \leq t \leq T}, \mathbb{P})$ be a probability space, where $(\mathcal{F}_t)_{0 \leq t \leq T}$ denote the set of information generated by the history of assets up to t and let H be a contingent claim with maturity T, which can be represented by defining its terminal payoff $H(\omega)$ in each scenario. We can note that H is a $\mathcal{F}_T$-measurable map $H : \Omega \to \mathbb{R}$ since H is known at T. We denote the set of contingent claims of interest by $\mathcal{H}$ and we can assume that $S_T^i \in \mathcal{H}$, which means that the underlying assets themselves can be seen as particular contingent claims whose payoff is given by the terminal value S_T^i. Then, we can define a pricing rule as the procedure which attributes to each contingent claim $H \in \mathcal{H}$ a value $\Pi_t(H)$ at each point in time. If $\Pi_t(H)$ is a pricing rule, it satisfy the following requirements:

1. $\Pi_t(H)$ should be a nonanticipating process, therefore we can compute the value of $\Pi_t(H)$ with the information given at t;

2. $\Pi_t(H)$ should be positive, which means that a claim with positive payoff should naturally have a positive value:

$$\forall \omega \in \Omega,\ H(\omega) \geq 0 \Rightarrow \forall t \in [0, T],\ \Pi_t(H) \geq 0;$$

3. $\Pi_t(H)$ should be linear, which means that the value of a portfolio is given by the sum of the value of its components:

$$\Pi_t\left(\sum_{j=1}^{J} H_j\right) = \sum_{j=1}^{J} \Pi_t(H_j)$$

   This requirement may fail when we consider large portfolio due to the discount prices on the market.

Consider an event $A \in \mathcal{F}$ and a random variable $\mathbf{1}_A$, which represents the payoff of a contingent claim which pays 1 at T if A occurs and zero otherwise, moreover, we will assume that $\mathbf{1}_A \in \mathcal{H}$. On one hand, we can start from a pricing rule Π and then we can construct a probability measure $\mathbb{Q}$. On the other hand, Π can be found from $\mathbb{Q}$: consider a random payoff of the form $H = \sum c_i \mathbf{1}_{A_i}$, then by the linearity property of Π we have that $\Pi_0(H) = E_{\mathbb{Q}}[H]$. Moreover, if the dominated convergence theorem holds on $\mathcal{H}$, we can conclude that for any random payoff $H \in \mathcal{H}$ we have:

$$\Pi_0(H) = e^{-rT} E_{\mathbb{Q}}[H]$$

Therefore, there is a one-to-one correspondence between linear valuation rules Π (verifying the properties above) and probability measure $\mathbb{Q}$, in fact we have:

$$\Pi_t(H) = e^{-r(T-t)} E_{\mathbb{Q}}[H] \tag{4.5}$$

$$\mathbb{Q}(A) = e^{r(T-t)} \Pi_0(\mathbf{1}_A) \tag{4.6}$$

The equation (4.5) is called a risk-neutral pricing formula and means that the value of a random payoff is given by its discountend expectation under the probability $\mathbb{Q}$. We can note that the probability $\mathbb{Q}$ has nothing to do with the actual or objective probability. Moreover, if we want that the pricing rule Π is constant over time (the value at 0 of the payoff H at T is equal to the value at 0 of the payoff $\Pi_t(H)$ at t), then $\mathbb{Q}$ should be given by the restriction of $\mathbb{Q}_t$ to the filtration $\mathcal{F}_t$. Hence, we have that:

$$\Pi_t(H) = e^{-r(T-t)} E_{\mathbb{Q}}[H|\mathcal{F}_t]$$

The equation above means that the pricing rule $\Pi_t(H)$ is equal to the discounted condition expectation under the probability $\mathbb{Q}$.

Now, consider the objective probability $\mathbb{P}$, which represents the probability of future scenario, defined in the probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\leq t \leq T}, \mathbb{P})$. Then, a fundamental requirement for a pricing rule is that it does not generate arbitrage opportunities. Recall that an arbitrage opportunity (in probability meaning) is a self-financing strategy ϕ, which can lead to a positive terminal gain without any probability of intermediate loss, therefore:

$$\mathbb{P}(\forall t \in [0,T].\, V_t(\phi) \geq 0) = 1, \qquad \mathbb{P}(V_T(\phi) > V_0(\phi)) \neq 0$$

A consequence of the arbitrage-free assumption is the "Law of One Price" which say that in the absence of trade frictions (such as transaction cost) two self-financing strategies with the same terminal payoff must have the same value at all times, otherwise the difference would generate an arbitrage opportunity.
Now, recall the definition of equivalent probability measure given by the equation (4.1). Then, we have the following two proposition, where the first one is also known as the "Fundamental theorem of asset pricing":
Proposition 4.3 *Let $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ be a probability space which defined a market model and let $(S_t)_{0\leq t\leq T}$ be the asset price in the time interval $[0,T]$. The asset price is arbitrage-free if and only if there exists an equivalent probability measure $\mathbb{Q} \sim \mathbb{P}$ such that the discounted asset price $(\hat{S}_t)_{0\leq t\leq T}$ are martingale with respect to $\mathbb{Q}$.*
Proposition 4.4 (Risk-neutral pricing) *Let $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ be a probability space which defined a market model under the probability $\mathbb{P}$. Then, any arbitrage-free pricing rule Π can be expressed as:*

$$\Pi_t(H) = e^{-r(T-t)} E_{\mathbb{Q}}[H|\mathcal{F}_t]$$

where $\mathbb{Q}$ is an equivalent martingale measure of $\mathbb{P}$.
The following two example can help us to understand the above proposition. Consider a market with two assets S^0 and S^1, where S^0 is the numeraire. Then, the two asset have price equal to S^1_t and $S^0_t = e^{rt}$. Then, consider a buy and hold strategies for the asset S^1: hold the asset until T, which generate a terminal payoff of S^1_T, or sold the asset at time t at the price S^1_t and invest the sum at the interest rate r until time T, which generate a terminal payoff equal to $e^{r(T-t)}S^1_t$. We can note that these two strategies are self-financing and that they have the same terminal payoff, therefore, by the "Law of One Price", they must have the same value at all time t:

$$E_{\mathbb{Q}}[S^1_T|\mathcal{F}_t] = E_{\mathbb{Q}}[e^{r(T-t)}S^1_t|\mathcal{F}_t] = e^{r(T-t)}S^1_t$$

Then, we divide for the numeraire, in particular we consider the time T, and we find that the discounted asset price is:

$$E_{\mathbb{Q}}\left[\frac{S_T^1}{S_T^0}\,\middle|\,\mathcal{F}_t\right] = \frac{e^{r(T-t)}S_t^1}{e^{rT}} = \frac{S_t^1}{S_t^0} = \hat{S}_t^1$$

Therefore, the absence of arbitrage implies that the discount value ($\hat{S}_t^1 = e^{-rt}S_1^1$) of the asset is a martingale with respect to the probability measure $\mathbb{Q}$, which is called equivalent martingale measure. This result can be generalized of all the traded assets.
Consider a self-financing strategy $(\phi_t)_{0\leq t\leq T}$ and let $\mathbb{Q}$ be a martingale measure. Then, $\hat{S}_t$ is a martingale under $\mathbb{Q}$. Hence, the value of the portfolio $V_t(\phi)$ ($V_t(\phi) = V_0 + \int_0^t \phi dS$) is a martingale and, in particular, $E_{\mathbb{Q}}[\int_0^t \phi dS] = 0$. Therefore, the random variable $\int_0^t \phi dS$ can take both positive and negative value, this imply that: $\mathbb{Q}\left(V_T(\phi) - V_0 = \int_0^T \phi_t dS_t \geq 0\right) \neq 1$. Hence, we can conclude that ϕ cannot be an arbitrage strategy since $\mathbb{P} \sim \mathbb{Q}$, which entails that $\mathbb{P}\left(\int_0^T \phi_t dS_t \geq 0\right) \neq 1$

4.1.2 Equivalence measures in Lévy processes

We have seen how much important is the equivalent change of measure in defining arbitrage-free pricing models in the Black-Scholes model, now we will study such changes of measure in the Lévy process. When we consider Lévy process the equivalence of their measures, gives relations between their parameters.

Consider two Poisson process defined by jump size, respectively, equal to a_1, a_2 and jump intensity, respectively, equal to λ_1, λ_2. Then, the following proposition shows the equivalence of measure for Poisson processes:
Proposition 4.5 (Equivalence of measure for Poisson processes) (proposition 9.5 in [2]) *Let $(N, \mathbb{P}_{\lambda_1})$ and $(N, \mathbb{P}_{\lambda_2})$ be Poisson process on $(\Omega, \mathcal{F}_t)$ with intensities λ_1 and λ_2 and jump sizes a_1 and a_2. Then, we have:*

1. *if $a_1 = a_2$, then $\mathbb{P}_{\lambda_1}$ is equivalent to $\mathbb{P}_{\lambda_2}$ with Radon-Nikodym density:*

$$\frac{d\mathbb{P}_{\lambda_1}}{d\mathbb{P}_{\lambda_2}} = exp\left[(\lambda_2 - \lambda_1)T - N_T \ln\frac{\lambda_2}{\lambda 1}\right] \tag{4.7}$$

2. *if $a_1 \neq a_2$, then $\mathbb{P}_{\lambda_1}$ is not equivalent to $\mathbb{P}_{\lambda_2}$.*

Proof
1. Let $A \in \mathcal{F}_t$. We need to show that, under the Radon-Nikodym derivative given by equation (7), the following equality is satisfy:

$$\mathbb{P}_{\lambda_1}(A) = E_{\mathbb{P}_{\lambda_2}}\left[\mathbf{1}_A \frac{d\mathbb{P}_{\lambda_1}}{d\mathbb{P}_{\lambda_2}}\right]$$

We can note that the left-hand side of the above equation could be rewritten as:

$$\mathbb{P}_{\lambda_1}(A) = \sum_{k=0}^{\infty} \frac{e^{-\lambda_1 T}(\lambda_1 T)^k}{k!} E_{\mathbb{P}_{\lambda_1}}[\mathbf{1}_A | N_T = k]$$

on the other hand, the right-hand side of the equation is equal to:

$$\begin{aligned} E_{\mathbb{P}_{\lambda_2}}\left[\mathbf{1}_A \frac{d\mathbb{P}_{\lambda_1}}{d\mathbb{P}_{\lambda_2}}\right] &= \sum_{k=0}^{\infty} \frac{e^{-\lambda_2 T}(\lambda_2 T)^k}{k!} \left(\frac{\lambda_1}{\lambda_2}\right)^k e^{(\lambda_2-\lambda_1)T} E_{\mathbb{P}_{\lambda_2}}[\mathbf{1}_A | N_T = k] \\ &= \sum_{k=0}^{\infty} \frac{e^{-\lambda_1 T}(\lambda_1 T)^k}{k!} E_{\mathbb{P}_{\lambda_2}}[\mathbf{1}_A | N_T = k] \end{aligned}$$

Then, putting all together we find:

$$\sum_{k=0}^{\infty}\frac{e^{-\lambda_1 T}(\lambda_1 T)^k}{k!}E_{\mathbb{P}_{\lambda_1}}[\mathbf{1}_A|N_T=k]=\sum_{k=0}^{\infty}\frac{e^{-\lambda_1 T}(\lambda_1 T)^k}{k!}E_{\mathbb{P}_{\lambda_2}}[\mathbf{1}_A|N_T=k]$$
$$E_{\mathbb{P}_{\lambda_1}}[\mathbf{1}_A|N_T=k]=E_{\mathbb{P}_{\lambda_2}}[\mathbf{1}_A|N_T=k]$$

We can note that the jump times of a Poisson process are uniformly distributed on this interval, therefore $E_{\mathbb{P}_\lambda}[\mathbf{1}_A|N_T=k]$ does not depend on λ.

□

This proposition told us that if we want the equivalence measure of two Poisson process, we can freely change the intensity of the jumps but the jump size must remain the same. In other word, the intensity of a Poisson process can be modified without changing the jump size of the process, but with changing the size of the jumps, which generates a new measure. This new measure assigns nonzero probability to some events which otherwise were impossible under the old one. We can note that two Poisson processes with different intensities define equivalent measures only on a finite time interval. In fact, if T is infinity, the Radon-Nikodym derivative (equation (4.7)) is either zero or infinity when the intensity of two Poisson process are different. This result is due to the fact that the intensity cannot be find from a trajectory of finite length but it can be estimated in an almost sure way from an infinite trajectory.

Now, consider two compound Poisson process and the following proposition gives us the equivalence of measure in this case:
Proposition 4.6 (Equivalence of measure for compound Poisson processes) (proposition 9.6 in [2]) *Let $(X,\mathbb{P})$ and $(X,\mathbb{Q})$ be compound Poisson processes on $(\Omega,\mathcal{F}_t)$ with Lévy measure $\nu_\mathbb{P}$ and $\nu_\mathbb{Q}$. The probability $\mathbb{P}$ and $\mathbb{Q}$ are equivalent if and only if $\nu_\mathbb{P}$ and $\nu_\mathbb{Q}$ are equivalent. In this case, the Radon-Nikodym derivative is:*

$$D_T=\frac{d\mathbb{Q}}{d\mathbb{P}}=exp\left(T(\lambda_\mathbb{P}-\lambda_\mathbb{Q})+\sum_{s\leq T}\phi(\Delta X_s)\right)\tag{4.8}$$

where $\lambda_\mathbb{P}\equiv\nu_\mathbb{P}(\mathbb{R})$ and $\lambda_\mathbb{Q}\equiv\nu_\mathbb{Q}(\mathbb{R})$ are the jumps intensities of the two processes and $\phi\equiv\ln\left(\frac{d\nu_\mathbb{Q}}{d\nu_\mathbb{P}}\right)$.
Proof
This proposition is a "if and only if" statement, therefore we need to show first the "if" part and then the "only if".
Hence, we start with the "if" part and we suppose that $\nu_\mathbb{P}$ and $\nu_\mathbb{Q}$ are equivalent. Now, conditioning the trajectory of X on the number of jumps on $[0,T]$, we find:

$$\begin{aligned}E_\mathbb{P}[D_T]&=E_\mathbb{P}\left[exp\left(T(\lambda_\mathbb{P}-\lambda_\mathbb{Q})+\sum_{s\leq T}\phi(\Delta X_s)\right)\right]\\&=e^{-\lambda_\mathbb{Q}T}\sum_{k=0}^{\infty}\frac{(\lambda_\mathbb{P}T)^k}{k!}E_\mathbb{P}\left[e^{\phi(\Delta X)}\right]^k=1\end{aligned}$$

Hence, $E_\mathbb{P}[D_T]$ is a probability measure. Therefore, we need to show that if X is a compound Poisson process under $\mathbb{P}$ is also a compound Poisson process under $\mathbb{Q}$ with Lévy measure $\nu_\mathbb{Q}$. The first step is check that X has $\mathbb{Q}$-independent increments and then check if X_T under the probability $\mathbb{Q}$ is a compound Poisson process with Lévy measure $T\nu_\mathbb{Q}$. To prove the independence of increments, consider two bounded measurable functions f and g and let $s<t\leq T$. We can note that X and $\ln D$ are $\mathbb{P}$-Lévy processes and that D is also a

$\mathbb{P}$-martingale. Hence, we have:

$$\begin{aligned}
E_{\mathbb{Q}}[f(X_s)g(X_t - X_s)] &= E_{\mathbb{P}}[f(X_s)g(X_t - X_s)D_t] \\
&= E_{\mathbb{P}}[f(X_s)D_s]E_{\mathbb{P}}\left[g(X_t - X_s)\frac{D_t}{D_s}\right] \\
&= E_{\mathbb{P}}[f(X_s)D_s]E_{\mathbb{P}}\left[g(X_t - X_s)D_t\right] \\
&= E_{\mathbb{Q}}[f(X_s)]E_{\mathbb{Q}}\left[g(X_t - X_s)\right]
\end{aligned}$$

which proves the $\mathbb{Q}$-independent increments. Then, if we use the characteristic function on X, we will find:

$$\begin{aligned}
E_{\mathbb{P}}\left[exp\left(T(\lambda_{\mathbb{P}} - \lambda_{\mathbb{Q}}) + \sum_{s \le T} \phi(\Delta X_s) + iuX_t\right)\right] &= \\
&= e^{-\lambda_{\mathbb{Q}} T} \sum_{k=0}^{\infty} \frac{(\lambda_{\mathbb{P}} T)^k}{k!} E_{\mathbb{P}}\left[e^{iu\Delta X + \phi(\Delta X)}\right]^k \\
&= exp\left(T \int (e^{iux} - 1)\nu_{\mathbb{Q}}(dx)\right)
\end{aligned}$$

Now, we prove the "only if" part and we assume the opposite of the "if" part, therefore we consider the case when $\nu_{\mathbb{P}}$ and $\nu_{\mathbb{Q}}$ are not equivalent. Then, we can find either a set A such that $\nu_{\mathbb{P}}(A) > 0$ and $\nu_{\mathbb{Q}}(A) = 0$ or a set A' such that$\nu_{\mathbb{P}}(A') = 0$ and $\nu_{\mathbb{Q}}(A') > 0$. Assume that we are in the second case, therefore we find that the set of trajectories having at least one jump size in A' has positive $\mathbb{Q}$-probability and zero $\mathbb{P}$-probability. Hence, we have shown that these two measures are not equivalent.

□

Before talk about the change of measure for general Lévy processes, we need to introduce the last important change of measure with respect the Brownian motion with drift. The following proposition gives us the equivalence of measure in this case:

Proposition 4.7 (Equivalence of measure for Brownian motion with drift) (proposition 9.7 in [2]) *Let $(X, \mathbb{P})$ and $(X, \mathbb{Q})$ be two Brownian motion on $(\Omega, \mathcal{F}_t)$ with volatilities $\sigma_{\mathbb{P}} > 0$ and $\sigma_{\mathbb{Q}} > 0$ and drift $\mu_{\mathbb{P}}$ and $\mu_{\mathbb{Q}}$. The probability $\mathbb{P}$ and $\mathbb{Q}$ are equivalent if and only if $\sigma_{\mathbb{P}} = \sigma_{\mathbb{Q}} > 0$ and singular otherwise. Then, when they are equivalent the Radon-Nikodym derivative is:*

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = exp\left(\frac{\mu_{\mathbb{Q}} - \mu_{\mathbb{P}}}{\sigma^2} X_T - \frac{1}{2}\frac{(\mu_{\mathbb{Q}} - \mu_{\mathbb{P}})^2}{\sigma^2} T\right) \tag{4.9}$$

With the Cameron-Martin theorem can rewrite the equation (4.9) as an exponential martingale equal to:

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = exp\left(\frac{\mu_{\mathbb{Q}} - \mu_{\mathbb{P}}}{\sigma} W_T - \frac{1}{2}\frac{(\mu_{\mathbb{Q}} - \mu_{\mathbb{P}})^2}{\sigma^2} T\right)$$

where $W_T = \frac{X_t - \mu_{\mathbb{P}} t}{\sigma}$ is a standard Brownian motion under the probability $\mathbb{P}$. This result shows that the drift and the volatility play a crucial roles in defining a diffusion model. On one hand, if we modify the drift, we will reweight the paths of X (the scenario); on the other hand, if we change the volatility, we will find a completely different process, leading to a new scenarios which were initially impossible.

After this introduction about the change of measure of the Poisson process and the Brownian motion, we can give a general result of equivalence of measure for Lévy processes. We can already say that in presence of jumps the class of probabilities equivalent to a given one is large even if we restrict our attention to structure preserving measures. The following proposition describes the possible measure changes under which a Lévy process remains a

Lévy process.

Proposition 4.8 (proposition 9.8 in [2]) *Let $(X_t, \mathbb{P})$ and $(X_t, \mathbb{Q})$ be two Lévy processes on $\mathbb{R}$ with characteristic triplets $(\sigma^2_{\mathbb{P}}, \nu_{\mathbb{P}}, \gamma_{\mathbb{P}})$ and $(\sigma^2_{\mathbb{Q}}, \nu_{\mathbb{Q}}, \gamma_{\mathbb{Q}})$. Then, $\mathbb{P}|_{\mathcal{F}_t}$ and $\mathbb{Q}|_{\mathcal{F}_t}$ [1] are equivalent for all t, or equivalently for one t >0, if and only if the following three conditions are satisfied:*

1) $\sigma_{\mathbb{P}} = \sigma_{\mathbb{Q}}$;

2) The Lévy measures are equivalent with

$$\int_{-\infty}^{\infty} \left(e^{\frac{\phi(x)}{2}} - 1\right)^2 \nu(dx) < \infty \tag{4.10}$$

where $\phi(x) = \ln\left(\frac{d\nu_{\mathbb{Q}}}{d\nu_{\mathbb{P}}}\right)$;

3) If $\sigma = 0$ *then in addition we have:*

$$\gamma_{\mathbb{Q}} - \gamma_{\mathbb{P}} = \int_{-1}^{1} x(\nu_{\mathbb{Q}} - \nu_{\mathbb{P}})dx \tag{4.11}$$

When $\mathbb{P}$ and $\mathbb{Q}$ are equivalent, the Radon-Nikodym derivative is:

$$\frac{d\mathbb{Q}|_{\mathcal{F}_t}}{d\mathbb{P}|_{\mathcal{F}_t}} = e^{U_t} \tag{4.12}$$

with

$$\begin{aligned} U_t &= \eta X_t^c - \frac{\eta^2\sigma^2 t}{2} - \eta\gamma t + \lim_{\varepsilon\to 0}\left(\sum_{s\le t, |\Delta X_s|>\varepsilon} \phi(\Delta X_s) - t\int_{|x|>\varepsilon} (e^{\phi(x)} - 1)\nu(dx)\right) && (4.13)\\ &= \langle \eta, X_t^c\rangle + \int_0^t \int_{\mathbb{R}^d} (e^{\phi(x)} - 1)\tilde{J}_X(ds\times dx) && (4.14)\end{aligned}$$

where X_t^c denotes the continuous martingale (Brownian motion) part of X_t and η is such that

$$\gamma_{\mathbb{Q}} - \gamma_{\mathbb{P}} - \int_{-1}^{1} x(\nu_{\mathbb{Q}} - \nu_{\mathbb{P}})dx = \sigma^2\eta$$

if $\sigma > 0$ and zero if $\sigma = 0$. Finally, U_t is a Lévy process with characteristic triplet $(\sigma_U^2, \nu_U, \gamma_U)$ given by:

a.

$$\sigma_U^2 = \sigma^2\eta^2$$

b.

$$\nu_U = \nu\phi^{-1}$$

c.

$$\gamma_U = -\frac{1}{2}\sigma_U\eta^2 - \int_{-\infty}^{\infty} (e^y - 1 - y\mathbf{1}_{|y|\le 1})(\nu\phi^{-1})dy$$

If we want that the equivalence measure hold, we cannot freely change the drift of the process unless a diffusion component is present. Moreover, we can note that we have freedom to change the distribution of large jumps as long as the Lévy measure is absolutely continuous with respect to the old one. This is an important result since only the large jumps are important in option pricing because they affect the tail of the return distribution. On the other hand, we cannot freely change the distribution of the small jumps because they depend on the behavior of the Lévy measure near zero.

[1] $\mathbb{P}|_{\mathcal{F}_t}$ represents the restriction of the probability measure $\mathbb{P}$ to $\mathcal{F}_t$ and it is a probability measure on $\mathcal{F}_t$ which assigns to all events in $\mathcal{F}_t$ the same probability $\mathbb{P}$ (the same is true for $\mathbb{Q}$).

4.1.3 Esscher transform and Relative entropy

When we consider model with jumps, if the Gaussian component is absent, we can find a much variety of equivalent measures by changing the distribution of jumps. Instead, in the Black-Scholes model we find equivalent measure by changing the drift.

One of the main tool to find the equivalence measure is used the Esscher transform, which constructs equivalent martingale measure in exponential-Lévy models[2]. To find this transformation consider a Lévy process X with triplet equal to (σ^2, ν, γ), a real number $\theta \in \mathbb{R}^d$ and a the Lévy measure ν, which satisfy $\int_{|x|\geq 1} e^{\theta x}\nu(dx) < \infty$. Then, we apply the measure transformation of proposition 4.8 with $\phi(x) = \theta X$ and we obtain an equivalent probability under which X is a Lévy process with: zero Gaussian component, Lévy measure $\tilde{\nu}(dx)$ and drift $\tilde{\gamma}$, respectively equal to $\tilde{\nu}(dx) = e^{\theta x}\nu(dx)$ and $\tilde{\gamma} = \gamma + \int_{-1}^{1} x(e^{\theta x} - 1)\nu(dx)$. Hence, using the proposition 4.8, we can find that the Radon-Nikodym derivative corresponding to the Esscher transform is equal to:

$$\frac{d\mathbb{Q}|_{\mathcal{F}_t}}{d\mathbb{P}|_{\mathcal{F}_t}} = \frac{e^{\theta X_t}}{E[e^{\theta X_t}]} = e^{\theta X_t + \gamma(\theta)t}$$

where $\gamma(\theta) = -\ln E[e^{\theta X_1}]$ is the log of the moment generating function of X_1 which, up to the change of variable $\theta \leftrightarrow -i\theta$, is given by the characteristic exponent of the Lévy process X.
In the Esscher transform we consider the exponential-Lévy model via ordinary exponential to find the equivalent measure but we can also consider the exponential-Lévy model via the stochastic exponential [3] since these two definition are equivalent. However, the set of Lévy processes that lead to arbitrage-free models with the form of ordinary exponential could be different from the set of arbitrage-free models defined with the stochastic exponential. Therefore, it will be convenient find the arbitrage-free condition using the stochastic exponential and then find the condition for the ordinary exponential using the transformation $X_t := \ln \mathcal{E}(Y_t)$. The exponential-Lévy model is arbitrage-free if we have one of the following case:

i. X has a nonzero Gaussian component: $\sigma > 0$;

ii. X has infinite variation: $\int_{-1}^{1} |x|\nu(dx) = \infty$;

iii. X has both negative and positive jumps;

iv. X has positve jumps and negative drift or negative jumps and positive drift.

Therefore, the above four case could be written as the following proposition:
Proposition 4.9 (Absence of arbitrage in exponential-Lévy models) *Let $(X, \mathbb{P})$ be a Lévy process. If the trajectories of X are neither almost surely increasing or decreasing, then the exponential-Lévy model given by $S_t = e^{rt+X_t}$ is arbitrage-free. Therefore, there exists a probability measure $\mathbb{Q}$ equivalent to $\mathbb{P}$ under which $(e^{-rt}S_t)_{0\leq t\leq T}$ is a $\mathbb{Q}$-martingale (where r is the interest rate).*
Proof

i. (X has a nonzero Gaussian component) Let X be a Lévy process with characteristic triplet (σ^2, ν, γ). As in the Black-Scholes model, an equivalence martingale measure can be found by changing the drift and without changing the Lévy measure when $\sigma > 0$. Therefore, we consider the case when $\sigma = 0$ and the function $\phi(x) = -x^2$. Then, we apply the measure transformation of the proposition 4.8 at the function $\phi(x)$ and we find an equivalent probability measure under which X is a Lévy process with zero Gaussian component, the same drift coefficient and Lévy measure equal to $\tilde{\nu}(dx) = e^{-x^2}\nu(dx)$. Hence, we can assume that

[2] Defined in chapter 2.
[3] Defined in chapter 2.

ν has exponential moments of all orders since $\tilde{\nu}$ has exponential moments of all orders.

iii. (X has both negative and positive jumps) Consider the Lévy triplet find by the Esscher transform is $(0, \tilde{\nu}, \tilde{\gamma})$, with $\tilde{\nu}(dx) = e^{\theta x}\nu(dx)$ and $\tilde{\gamma} = \gamma + \int_{-1}^{1} x(e^{\theta x} - 1)\nu(dx)$. For e^X to be a martingale under the new probability, the following equation must be satisfied:

$$\tilde{\gamma} + \int_{-\infty}^{\infty} (e^x - 1 - x\mathbf{1}_{|x|\leq 1})\tilde{\nu}(dx) = 0$$

Therefore, we substitute $\tilde{\nu}$ and $\tilde{\gamma}$ and we find:

$$\gamma + \int_{-1}^{1} x(e^{\theta x} - 1)\nu(dx) + \int_{-\infty}^{\infty} (e^x - 1 - x\mathbf{1}_{|x|\leq 1})e^{\theta x}\nu(dx) = 0$$

$$\underbrace{\int_{-1}^{1} x(e^{\theta x} - 1)\nu(dx) + \int_{-\infty}^{\infty} (e^x - 1 - x\mathbf{1}_{|x|\leq 1})e^{\theta x}\nu(dx)}_{=f(\theta)} = -\gamma \tag{4.15}$$

$$-\gamma = f(\theta)$$

By the dominated convergence, we have that f is continuous and that its first derivative is greater or equal to 0: $f'(\theta) = \int_{-\infty}^{\infty} x(e^x - 1)e^{\theta x}\nu(dx) \geq 0$. This imply that $f(\theta)$ is an increasing functions. Moreover, if $\nu((0,\infty)) > 0$ and $\nu((-\infty, 0)) > 0$, then f' is everywhere bounded from below by a positive number. Hence, $f(+\infty) = +\infty$, $f(-\infty) = -\infty$ and we have a solution.

iv. (X has positive jumps and negative drift or viceversa) We consider only the case when $\nu((-\infty, 0)) = 0$, since the proposition is symmetric. In this case, we still have that $f(+\infty) = +\infty$ but we need that $\lim_{\theta\to-\infty} f(\theta) \neq -\infty$, hence:

$$\lim_{\theta\to-\infty} \int_{0}^{\infty} (e^x - 1 - x\mathbf{1}_{|x|\leq 1})e^{\theta x}\nu(dx) \to 0$$

$$\lim_{\theta\to-\infty} \int_{0}^{\infty} x(e^{\theta x} - 1)\nu(dx) = \int_{0}^{\infty} x\nu(dx)$$

The first limit always converges to zero. On the other hand, the second limit above has a solution if $\int_0^{\infty} x\nu(dx) = \infty$ (it goes to $-\infty$ as $\theta \to \infty$), otherwise it converges to $-\int_0^{\infty} x\nu(dx)$ which is the difference between γ and the drift of X in the finite variation case. Therefore, in the finite variation case a solution exists if X has negative drift. Hence, we can conclude that a solution exists unless $\nu((-\infty, 0)) = 0$, $\int_0^{\infty} x\nu(dx) < \infty$ and the drift is positive.

□

When we find two equivalent probability measure and we want to measure the proximity of them, we can use the Relative entropy. Let $(\Omega, \mathcal{F})$ be the space of real-valued discontinuous càdlàg function defined on $[0, T]$, $\mathcal{F}_t$ be the history of path up to t and $\mathbb{P}$ and $\mathbb{Q}$ be two equivalent probability measure on $(\Omega, \mathcal{F})$. Then, we can define the relative entropy of $\mathbb{Q}$ with respect to $\mathbb{P}$ as:

$$\mathcal{E}(\mathbb{Q}, \mathbb{P}) = E_{\mathbb{Q}}\left[\ln \frac{d\mathbb{Q}}{d\mathbb{P}}\right] = E_{\mathbb{P}}\left[\frac{d\mathbb{Q}}{d\mathbb{P}} \ln \frac{d\mathbb{Q}}{d\mathbb{P}}\right] \tag{4.16}$$

If we consider a strictly convex function: $f(x) = x \ln x$, we can write the equation (16) as:

$$\mathcal{E}(\mathbb{Q}, \mathbb{P}) = E_{\mathbb{P}}\left[f\left(\frac{d\mathbb{Q}}{d\mathbb{P}}\right)\right]$$

We can note that the relative entropy is a convex function of $\mathbb{Q}$ and that it is always nonnegative $\mathcal{E}(\mathbb{Q}, \mathbb{P}) \geq 0$ due to the Jensen's inequality [4]. In fact, $\mathcal{E}(\mathbb{Q}, \mathbb{P}) = 0$ if and only if

[4] Defined in chapter 1.

$\frac{d\mathbb{Q}}{d\mathbb{P}} = 1$, almost surely.
The following proposition shows the relative entropy of Lévy processes when the measures are generated by the exponential-Lévy models:
Proposition 4.10 (Relative entropy of Lévy processes) (Proposition 9.10 in [2]) *Let $\mathbb{Q}$ and $\mathbb{P}$ be equivalent measure on $(\Omega, \mathcal{F})$ generated by exponential-Lévy models with Lévy triplet $(\sigma^2, \nu_{\mathbb{P}}.\gamma_{\mathbb{P}})$ and $(\sigma^2, \nu_{\mathbb{Q}}.\gamma_{\mathbb{Q}})$. Assume $\sigma > 0$. Then, the relative entrapy $\mathcal{E}(\mathbb{Q}, \mathbb{P})$ is given by:*

$$\begin{aligned}\mathcal{E}(\mathbb{Q},\mathbb{P}) = \frac{T}{2\sigma^2}&\left\{\gamma_{\mathbb{Q}} - \gamma_{\mathbb{P}} - \int_{-1}^{1} x(\nu_{\mathbb{Q}} - \nu_{\mathbb{Q}})(dx)\right\}^2 + \\ T&\int_{-\infty}^{\infty}\left(\frac{d\nu_{\mathbb{Q}}}{d\nu_{\mathbb{P}}}\ln\frac{d\nu_{\mathbb{Q}}}{d\nu_{\mathbb{P}}} + 1 - \frac{d\nu_{\mathbb{Q}}}{d\nu_{\mathbb{P}}}\right)\nu_{\mathbb{P}}(dx)\end{aligned} \tag{4.17}$$

If $\mathbb{Q}$ and $\mathbb{P}$ correspond ro risk-neutral exponential-Lévy models, the equation (4.17) *becomes:*

$$\begin{aligned}\mathcal{E}(\mathbb{Q},\mathbb{P}) = \frac{T}{2\sigma^2}&\left\{\int_{-\infty}^{\infty} (e^x - 1)(\nu_{\mathbb{Q}} - \nu_{\mathbb{Q}})(dx)\right\}^2 + \\ T&\int_{-\infty}^{\infty}\left(\frac{d\nu_{\mathbb{Q}}}{d\nu_{\mathbb{P}}}\ln\frac{d\nu_{\mathbb{Q}}}{d\nu_{\mathbb{P}}} + 1 - \frac{d\nu_{\mathbb{Q}}}{d\nu_{\mathbb{P}}}\right)\nu_{\mathbb{P}}(dx)\end{aligned} \tag{4.18}$$

Proof
Let (X_t) be a Lévy process and we can define $S_t = e^{X_t}$. We can note that the filtration generated by X_t and S_t are the same. Therefore, we can compute the relative entropy of the log-prices processes. Hence, we use the Radon-Nikodym derivative define in equation (9) and we find:

$$\mathcal{E} = \int \frac{d\mathbb{Q}}{d\mathbb{P}} \ln \frac{d\mathbb{Q}}{d\mathbb{P}} d\mathbb{P} = E_{\mathbb{P}}\left[U_T e^{U_T}\right]$$

where U_T is a Lévy process with Lévy triplet $(\sigma_U^2, \nu_U, \gamma_U)$, which are defined by the point *a. b. c.* of the proposition 4.6. Now, consider the characteristic function $\phi_t(z)$ equal to $\phi_t(z) = E_{\mathbb{P}}\left[U_T e^{izU_T}\right] = e^{t\psi(z)}$, where $\psi(z)$ is the characteristic exponent. Then, we can write the expectation above as:

$$\begin{aligned}E_{\mathbb{P}}\left[U_T e^{U_T}\right] &= -i\frac{d}{dz}\phi_T(-i) = -iTe^{T\psi(-i)}\psi'(-i) \\ &= -iT\psi'(-i)E_{\mathbb{P}}[e^{U_T}] = -iT\psi'(-i)\end{aligned}$$

From the Lévy-Khinchin formula [5] we now that:

$$\psi'(z) = -a_U z + i\gamma_U + \int_{-\infty}^{\infty}\left(ixe^{izx} - ix\mathbb{1}_{|x|\le 1}\right)\nu_U(dx)$$

Then, we substitute and we find the relative entropy as follow:

$$\begin{aligned}\mathcal{E} &= -a_U T + \gamma_U T + T\int_{-\infty}^{\infty}\left(ixe^{izx} - ix\mathbf{1}_{|x|\le 1}\right)\nu_U(dx) \\ &= \frac{\sigma^2 T}{2}\eta^2 + T\int (ye^y - e^y + 1)(\nu_{\mathbb{P}}\phi^{-1})(dy) \\ &= \frac{\sigma^2 T}{2}\eta^2 + T\int\left(\frac{d\nu_{\mathbb{Q}}}{d\nu_{\mathbb{P}}}\ln\frac{d\nu_{\mathbb{Q}}}{d\nu_{\mathbb{P}}} + 1 - \frac{d\nu_{\mathbb{Q}}}{d\nu_{\mathbb{P}}}\right)\nu_{\mathbb{P}}(dx)\end{aligned}$$

where η is chosen such that: $\gamma_{\mathbb{Q}} - \gamma_{\mathbb{P}} - \int_{-1}^{1} x(\nu_{\mathbb{Q}} - \nu_{\mathbb{P}})(dx) = \sigma^2\eta$. Since, we assume that $\sigma > 0$ we can write the first term on the above equation as:

$$\frac{T}{2}\sigma^2\eta^2 = \frac{T}{2\sigma^2}\left\{\gamma_{\mathbb{Q}} - \gamma_{\mathbb{P}} - \int_{-1}^{1} x(\nu_{\mathbb{Q}} - \nu_{\mathbb{Q}})(dx)\right\}^2$$

[5] Theorem 2.2 in chapter 2

which leads to the equation (4.17).
Now, consider the case where $\mathbb{P}$ and $\mathbb{Q}$ are martingale measures. Then, we can express the drift γ using σ and ν as:

$$\frac{T}{2}\sigma^2\eta^2 = \frac{T}{2\sigma^2}\left\{\int_{-\infty}^{\infty}(e^x-1)(\nu_{\mathbb{Q}}-\nu_{\mathbb{Q}})(dx)\right\}^2$$

which leads to the equation (4.18).

□

4.2 Option Pricing

The modern finance is centered on the pricing of derivative instruments, which are instruments whose payoff is a function of the value of another financial instruments (such as commodities, currency, bond, stock), also called underlying asset. The derivative itself is a contract between two or more counterparties and the derivatives traded directly between two counterparties are called over-the-counter (OTC) derivatives, which contrast with exchange-traded derivatives where an exchange matches buyers and sellers and each counterparty faces the exchange on the contract.
One of the most popular derivative contract in the world is the option contract. An option is a contract between a buyer and a seller that gives at the purchaser of the option the right, but not the obligation, to buy or to sell a particular asset at an exercise date at an agreed price (exercise price). Later in this chapter, we denote with K the strike or exercise price, with T the exercise date or maturity and with S_T the value of the asset at the maturity. On one hand, we have a call option when we have the right to buy an asset S for K at time T and we can represent its payout at time T as:

$$C_T = \max(S_T - K, 0) = (S_T - K)^+ \tag{4.19}$$

on the other hand, we have a put option when we have the right to sell an asset S for K at time T and we can represent its payout at time T as:

$$P_T = \max(K - S_T, 0) = (K - S_T)^+ \tag{4.20}$$

In the market we can find two type of option contract: European option and American option. In the European option, we can exercise the option only at the maturity, instead in the American one, we can exercise the option at any time t, with $t \leq T$.

Consider an asset S and an European call option (for the rest of the chapter we consider only the European option type) with underlying asset S and maturity T. Let $C_t(T,K)$ be the price of the call option on S with strike K, $\forall t \in [0,T]$. Then, follow Merton, we can decompose the price of the call option in the following way:

$$C_t(T,K) = \underbrace{\max(S_t - K, 0)}_{intrinsic\ value} + \underbrace{C_t(T,K) - \max(S_t - K, 0)}_{extrinsic\ value}$$

We use the notation of $C_t^I(T,K)$ to identify the intrinsic value and $C_t^E(T,K)$ to identify the extrinsic value. We can note that for every time t such that $0 \leq t < T$, the extrinsic value $C_t^E(T,K)$ is always positive and the intrinsic value $C_t^I(T,K)$ is never negative. If we consider the case when $t = T$, the extrinsic value is zero and the value of the call option is equal to the intrinsic one:

$$C_T(T,K) = \max(S_T - K, 0) = C_T^I(T,K)$$

Therefore, we can say that a call option at time t is:

- in the money (ITM) if $S_t > K$;

- at the money (ATM) if $S_t = K$;
- out of the money (OTM) if $S_t < K$;
- deep in the money if $S_t \gg K$;
- deep out of the money if $S_t \ll K$;
- just in the money or just out of the money if $S_t \cong K$;

We can note that for the ATM and OTM call option we have only the extrinsic value and that this terminology is still valid for the put option with a little bit of change in the notation.

Consider a call and put option at time T. It easy to note that the two payoff are never negative since we want the maximum between zero and the difference between the underlying asset and the strike price in the call option (vice versa for the put option):

$$C_T(T, K) = \max(S_T - K, 0) \geq 0 \qquad \textit{and} \qquad P_T(T, K) = \max(K - S_T, 0) \geq 0$$

Hence: $C_t(T, K) \geq 0$ and $P_t(T, K) \geq 0$. Moreover, if we study the behavior of the strike price, which can go either to zero or infinity, we will find how the put and call option behave:

	$\lim_{k\to 0}$	$\lim_{k\to\infty}$
Call	$\max(S_T - K, 0) = S_T$	$\max(S_T - K, 0) = 0$
Put	$\max(K - S_T, 0) = 0$	$\max(K - S_T, 0) = \infty$

Let $V_t(T, K)$ be the value at time t of the forward contract [6] with delivery price K. Then, we can define the forward price $F(t, T)$ at current time $t \leq T$ to be the delivery price K such that $V_t(T, K) = 0$, in other words, such that the forward contract has zero value at time t. Therefore, we find the following relationship:

$$V_t(T, K) = (F(t, T) - K)e^{-r(T-t)}$$

Now, we can find how the price of a call and a put of the same strike are related with the value of the forward:

$$C_t(T, K) - P_t(T, K) = V_t(T, K) \tag{4.21}$$

The above equation is called Put-Call Parity, which states that long one call and short one put is equal to go long to one forward. After some transformation, the Put-Call Parity can be written as:

$$C_t(T, K) - P_t(T, K) = S_t - Ke^{-r(T-t)} \tag{4.22}$$

The Put-Call Parity is important for three reason. First, it is an arbitrage-free condition. In fact, any violation of the Put-Call Parity leads to an arbitrage opportunity. Second, when we want pricing an option, we can focus only in a call (for example) and then find the price of the put using the Put-Call Parity. Third, the Put-Call Parity is model-independent, which means that this parity relationship between the values of put and call options holds, regardless of the model assumed for the evolution of the price of the underlying asset or arbitrage opportunities occur.

Consider two call option with two different strike price: $C_t(T, K_1)$ and $C_t(T, K_2)$. If $K_1 < K_2$, then we have that $C_t(T, K_1) \geq C_t(T, K_2)$ and $P_t(T, K_1) \leq P_t(T, K_2)$. This result follows from the monotonicity theorem. Moreover, if $K_1 < K_2$, we will find that $C_t(T, K_1) - C_t(T, K_2) \leq e^{-r(T-t)}(K_2 - K_1)$ and $P_t(T, K_2) - P_t(T, K_1) \leq e^{-r(T-t)}(K_2 - K_1)$. Therefore, combining this two result, it is easy to see that:

$$C_t(T, K_2) \leq C_t(T, K_1) \leq C_t(T, K_2) + e^{-r(T-t)}(K_2 - K_1)$$

[6] A forward contract (or forward) is an agreement between two counterparties to trade a specific asset at a certain future time T, called maturity, and at a certain price K, called delivery price.

Since $e^{-r(T-t)}$ does not depend on K_1 or K_2, we can say that $C_t(T,K)$ is a Lipschitz continuous[7] function of K with Lipschitz constant $e^{-r(T-t)}$. Moreover, if $K_1 \neq K_2$, we have the following arbitrage relation for the call option:

$$-e^{-r(T-t)} < \frac{C_t(T,K_2) - C_t(T,K_1)}{K_2 - K_1} < 0$$

Finally, consider an european call option $C_t(T,K)$ with underlying S and let $K_1 < K_2$, $K^* = \lambda K_1 + (1-\lambda)K_2$ and let $\lambda \in [0,1]$. We can show that the price of the call option is a convex function of the strike price K; therefore, for each $\lambda \in [0,1]$, we have:

$$C_t(T,K^*) \leq \lambda C_t(T,K_1) + (1-\lambda)C_t(T,K_2)$$

We can note that the function $x^+ = \max(0,x)$ is convex, hence for each $x, y \in \mathbb{R}$ and for each $\lambda \in [0,1]$, we have that:

$$\lambda \max(0,x) + (1-\lambda)\max(0,y) \geq \max(0, \lambda x + (1-\lambda)y)$$

therefore, for each S_T we have the following inequality:

$$\lambda(S_T - K_1)^+ + (1-\lambda)(S_T - K_1)^+ \geq (S_T - \lambda K_1 - (1-\lambda)K_2)^+$$

We see later that when there is a jump in the asset trajectories, the call option in the money can become out of the money and that the jump could transform the convex payoff into a concave one. This imply that the delta-hedging is not possible in reality when the jump occurs.

4.2.1 Pricing European Option in Black-Scholes model

In the Black and Scholes model the behavior of prices is a continuous time model with the assumption of one risky asset (denoted by S_t at time t) and a riskless asset (denoted by S_t^0 at time t). Moreover, we assume that the risky asset will not pay dividend and that the behavior of the riskless asset is expressed by the following ordinary differential equation:

$$dS_t^0 = rS_t^0 dt \tag{4.23}$$

where r is an instantaneous interest rate and it is a non-negative constant. We also set that S_0^0, which imply that $S_t^0 = e^{rt}$ for $t \geq 0$. On the other hand, the behavior of the risky asset is determined by the following stochastic differential equation:

$$dS_t = S_t(\mu dt + \sigma dB_t) \tag{4.24}$$

where B_t is a standard Brownian motion defined in the probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \leq t \leq T}, \mathbb{P})$ and μ and σ (called the volatility of the asset) are two constant, which are bounded and locally Lipschitz continuous. We consider the model valid for the time interval $[0,T]$, where T is the maturity date of the option. Equation (4.24) has a closed-form solution equal to:

$$S_t = S_0 e^{(\mu - \frac{\sigma^2}{2})t + \sigma B_t} \tag{4.25}$$

where S_0 is the spot price at time 0. Moreover, we can note that S_t has a lognormal law, which imply that (S_t) is a solution of an equation of the type (4.24) if and only if the process $(\log(S_t))$ is a Brownian motion. Therefore, we can find three properties, which can express in concrete terms the hypotheses of Black and Scholes on the behavior of the stock price. Hence, the risky asset has the following properties:

i. continuity of the sample paths;

ii. independence of the relative increments: if $u \leq t$, the relative increments $(S_t - S_u)/S_u$ is independent of the σ-algebra $\sigma(S_v, v \leq u)$;

[7] Defined in the appendix A.7

iii. stationary of the relative increments: if $u \leq t$, the law of the relative increments $(S_t - S_u)/S_u$ is identical to the law of $(S_{t-u} - S_0)/S_0$.

Now, we need to show that there exist a probability equivalent to $\mathbb{P}$ under which the discounted stock price is a martingale. Therefore, we need to introduce the following theorem, called Girsanov theorem:
Theorem 4.2 (theorem 4.2.2 in [1]) *Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0\leq t\leq T}, \mathbb{P})$ be a filtered probability space and $(B_t)_{0\leq t\leq T}$ an $\mathcal{F}_t$-standard Brownian motion. Let $(\theta_t)_{0\leq t\leq T}$ be an adapted process satisfying $\int_0^T \theta_s^2 ds < \infty$ a.s. and such that the process $(L_t)_{0\leq t\leq T}$ defined by*

$$L_t = exp\left(-\int_0^t \theta_s dB_s - \frac{1}{2}\int_0^t \theta_s^2 ds\right) \tag{4.26}$$

is a martingale. Then, under the probability $\mathbb{P}^{(L)}$ with density L_T with respect to $\mathbb{P}$, the process $(W_t)_{0\leq t\leq T}$ defined by $W_t = B_t + \int_0^t \theta_s ds$ is an $\mathcal{F}_t$-standard Brownian motion.
A detailed proof can be found in chapter 5 of "Stochastic Calculus for Finance II" written by Shreve

If we define the discounted value as $\tilde{S}_t = e^{-rt}S_t$, we find:

$$d\tilde{S}_t = -re^{-rt}S_t dt + e^{-rt}dS_t$$

We can substitute inside the above equation the equation (4.24) and we find:

$$\begin{aligned} d\tilde{S}_t &= -re^{-rt}S_t dt + e^{-rt}(S_t(\mu dt + \sigma dB_t)) \\ &= -re^{-rt}S_t dt + e^{-rt}S_t\mu dt + e^{-rt}S_t\sigma dB_t \\ &= e^{-rt}S_t(-rdt + \mu dt\sigma dB_t) \\ &= \tilde{S}_t((\mu - r)dt + \sigma dB_t) \end{aligned}$$

If we set $W_t = B_t(\mu - r)t/\sigma$, we can rewrite the above result as:

$$d\tilde{S}_t = \tilde{S}_t\sigma dW_t \tag{4.27}$$

Now, if we apply the theorem 4.2 with $\theta = (\mu - r)/\sigma$, we will find the probability $\mathbb{Q}$ equivalent to $\mathbb{P}$ under which $(W_t)_{0\leq t\leq T}$ is a standard Brownian motion. Then, under the probability $\mathbb{Q}$, we can note from (4.27) that $(\tilde{S}_t)$ is a martingale and that:

$$\tilde{S}_t = \tilde{S}_0 e^{\sigma W_t - \frac{\sigma^2 t}{2}}$$

Finally, we can price an option and, in particular, we will focus on European option and for simplicity we use the notation of $h = f(x) = (x - K)^+$ for the call option. Moreover, we will focus only on admissible strategies defined as:
Definition 4.1 (definition 4.3.1 in [1]) *A strategy $\phi = (H_t^0, H_t)_{0\leq t\leq T}$ is admissible if it is self-financing and if the discounted value $\tilde{V}_t(\phi) = H_t^0 + H_t\tilde{S}_t$ of the corresponding portfolio is, for all t, non-negative, and such that $\sup_{t\in[0,T]} \tilde{V}_t$ is square-integrable under $\mathbb{Q}$.*
Hence, an option is said to be replicable if its payoff at maturity is equal to the final value of an admissible strategy. It easy to note that an option defined by h is replicable, if it is square-integralbe under $\mathbb{Q}$. In particular, when we consider a call option, this property hold since $E_{\mathbb{Q}}[S_T^2]$; on the other hand, if we consider a put option, h is bounded.
Theorem 4.3 (theorem 4.3.2 in [1]) *In the Black-Scholes model, any option defined by a non-negative, $\mathcal{F}_T$-measurable random variable h, which is square-integrable under the probability $\mathbb{Q}$, is replicable and the value at time t of any replicating portfolio is given by:*

$$V_t = E_{\mathbb{Q}}[e^{-r(T-t)}h|\mathcal{F}_t]$$

A detailed proof can be found in chapter 4 of "Introduction to Stochastic Calculus applied to Finance" written by Lamberton and Lapeyre.

Hence, the option value at time t can be defined by the expression $E_{\mathbb{Q}}\left[e^{-r(T-t)}h \,\middle|\, \mathcal{F}_t\right]$. When the random variable h can be written as $h = f(S_T)$, we can express the option value V_t at time t as a function of t and S_t. Then, we have:

$$\begin{aligned} V_t &= E_{\mathbb{Q}}\left[e^{-r(T-t)}f(S_T) \,\middle|\, \mathcal{F}_t]\right] \\ &= E_{\mathbb{Q}}\left[e^{-r(T-t)}f\left(S_t e^{r(T-t)}e^{\sigma(W_T-W_t)-\frac{\sigma^2}{2}(T-t)}\right) \,\middle|\, \mathcal{F}_t\right] \end{aligned}$$

We can note that the random variable S_t is $\mathcal{F}_t$-measurable and, under the probability $\mathbb{Q}$, $W_T - W_t$ is independent of $\mathcal{F}_t$. Therefore, we conclude that:

$$V_t = F(t, S_t)$$

where

$$F(t,x) = E_{\mathbb{Q}}\left[e^{-r(T-t)}f\left(xe^{r(T-t)}e^{\sigma(W_T-W_t)-\frac{\sigma^2}{2}(T-t)}\right)\right] \tag{4.28}$$

Since, under$\mathbb{Q}$, $W_T - W_t$ is a zero-mean normal variable with variance $T - t$, we have:

$$F(t,x) = \frac{e^{-r(T-t)}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f\left(xe^{(r-\frac{\sigma^2}{2})(T-t)+\sigma y\sqrt{T-t}}\right) e^{-\frac{y^2}{2}} dy$$

Consider a call option, where $F(x) = (x - K)^+$, then the equation (4.28) can be rewritten:

$$\begin{aligned} F(t,x) &= E_{\mathbb{Q}}\left[e^{-r(T-t)}f\left(xe^{\sigma(W_T-W_t)+(r-\frac{\sigma^2}{2})(T-t)} - K\right)^+\right] \\ &= E_{\mathbb{Q}}\left[\left(xe^{\sigma\sqrt{\theta}g-\frac{\sigma^2\theta}{2}} - Ke^{-r\theta}\right)^+\right] \end{aligned}$$

where g is a standard Gaussian variable and $\theta = T - t$. Then, we can set:

$$d_1 = \frac{\ln\left(\frac{x}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)\theta}{\sigma\sqrt{\theta}} \text{ and } d_2 = d_1 - \sigma\sqrt{\theta}$$

Therefore, we find with this notation

$$\begin{aligned} F(t,x) &= E\left[\left(xe^{\sigma\sqrt{\theta}g-\frac{\sigma^2\theta}{2}} - Ke^{-r\theta}\right)\mathbf{1}_{g+d_2\geq 0}\right] \\ &= \int_{-d_2}^{+\infty} \left(xe^{\sigma\sqrt{\theta}g-\frac{\sigma^2\theta}{2}} - Ke^{-r\theta}\right)\frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}}dy \\ &= \int_{-\infty}^{d_2} \left(xe^{\sigma\sqrt{\theta}g-\frac{\sigma^2\theta}{2}} - Ke^{-r\theta}\right)\frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}}dy \\ &= \int_{-\infty}^{d_2} \left(xe^{\sigma\sqrt{\theta}g-\frac{\sigma^2\theta}{2}}\right)\frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}} - \int_{-\infty}^{d_2} \left(Ke^{-r\theta}\right)\frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}}dy \end{aligned}$$

Now, in the first integral we use a change of variable with $z = y + \sigma\sqrt{\theta}$ and the last equation above become:

$$F(t,x) = xN(d_1) - Ke^{-r\theta}N(d_2) \tag{4.29}$$

where $N(d) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{d} e^{-x^2/2}dx$ is the Gaussian cumulative distribution function. The equation (4.29) is the price of the call option in the Black-Scholes model. On the other hand, the price of a put in the Black-Scholes model is equal to:

$$F(t,x) = Ke^{-r\theta}N(-d_2) - xN(-d_1) \tag{4.30}$$

4.2.2 Pricing European Option in exponential-Lévy models

Let $(S_t)_{0\le t\le T}$ be a stochastic process, which describe the asset price behavior and let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_t, \mathbb{P})$ be a probability space, where $\mathbb{P}_t$ represents the history of the asset price. We saw in the previous section that, in the Black-Scholes model, the dynamic of an asset price is given by equation (4.25), which can be rewritten as:

$$S_t = S_0 e^{B_t^0} \tag{4.31}$$

where $B_t^0 = \left(\mu - \frac{\sigma^2}{2}\right)t + \sigma W_t$. If we apply the Ito formula, we will find:

$$\frac{dS_t}{S_t} = \mu dt + \sigma dW_t = dB_t^1 \tag{4.32}$$

where $B_t^1 = \mu t + \sigma W_t$. Therefore, we find two ways to define the risk neutral dynamics: select the exponential as in (4.31) or select the stochastic exponential as in (4.32). If we replace the Brownian motion with drift by a Lévy process, we will find a class of risk neutral models with jumps. Hence, if we make this substitution in (4.31), we will find:

$$S_t = S_0 e^{X_t}$$

This model is called exponential-Lévy model. In order to use this model to price an option, we want that the discounted value of stock price is a martingale. Hence, we need to impose an additional restrictions on the characteristic triplet (σ^2, ν, γ) of X:

$$\int_{|x|\ge 1} e^x \nu(dx) < +\infty$$
$$\gamma + \frac{\sigma^2}{2} + \int (e^y - 1 - y\mathbf{1}_{|y|\le 1})\nu(dy) = 0$$

Therefore, we can conclude that $(X_t)_{t\ge 0}$ is a Lévy process such that $E_{\mathbb{Q}}[e^{X_t}] = 1$ for all t. On the other hand, we can replace B_t^1 in (4.32) by a Lévy process Z_t and we find:

$$dS_t = rS_{t^-}dt + S_{t^-}dZ_t$$

then S_t corresponds to the stochastic exponential of Z. If we want that the discounted stock price is a martingale, we need that the Lévy process Z_t satisfy $E[Z_1] = 1$.

The price of a call option can be expressed as the risk-neutral conditional expectation of the payoff:

$$C_t(T, K) = e^{-r(T-t)} E_{\mathbb{Q}}[(S_T - K)^+ | \mathcal{F}_t] \tag{4.33}$$

In an exponential-Lévy model with the stationary and independence of increments property, the equation (4.33) could be rewritten as an expectation of the process at time $\theta = T - t$:

$$\begin{aligned} C_t(T = t + \theta, K) &= e^{-r\theta} E[(S_T - K)^+ | S_t = S] \\ &= e^{-r\theta} E[(Se^{r\theta + X_\theta} - K)^+] \\ &= Ke^{-r\theta} E[(e^{x + X_\theta} - 1)^+] \end{aligned} \tag{4.34}$$

where $x = \ln\left(\frac{S}{K}\right) + r\theta$ is the log-forward-moneyness. When the option is at the money, x is equal to 0. Therefore, we can note that the call option price in the exponential-Lévy model depends on the time remaining until maturity and it is a homogeneous function of order 1 of S and K.

Now, if we define the relative-forward option price, we can see that the structure of the option price in exponential-Lévy models is parametrized by only two variables. Therefore, we define the relative-forward price in terms of the relative variables $u(\tau, x)$:

$$u(\theta, x) = \frac{e^{r\theta} C_t(T = t + \theta, K)}{K} \tag{4.35}$$

and if we substitute in the above equation the result in (4.34), we can conclude that the structure of the option price is:

$$u(\theta, x) = E[(e^{x+X_\theta} - 1)^+]$$

This result is a consequence of temporal and spatial homogeneity of Lévy processes. Moreover, we can rewritten $u(\theta, \cdot)$ as the convolution product between the payoff function h and the transition density of the Lévy process ρ_θ: $u(\theta, \cdot) = \rho_\theta * h$. Thus, if the process has smooth transition densities, $u(\theta, \cdot)$ will be smooth, even if the payoff function h is not.

Now, consider an exponential price process of the form:

$$S_t := S_0 e^{\mu t + \sigma W_t + Y_t}$$

where Y_t is compound Poisson process, defined in chapter 2 (definition 2.2). Therefore, the process S_t can be written as:

$$\begin{aligned} S_t &= S_0 exp\left(\mu t + \sigma W_t + \sum_{i=1}^{N_t} Z_i\right) \\ &= S_0 e^{\mu t + \sigma W_t} \prod_{i=1}^{N_t} e^{Z_i} \\ &= S_0 e^{\mu t + \sigma W_t} \prod_{0 \le s \le t} e^{\Delta Y_t}, \qquad t \in \mathbb{R}^+ \end{aligned}$$

from relation $\Delta Y_t = Z_{N_t} \Delta N_t$ (defined in chapter 2). The process $(S_t)_{t\in\mathbb{R}}$ is equivalently given by the log-returns dynamics:

$$d \log S_t = \mu dt + \sigma dW_t + dY_t \qquad t \in \mathbb{R}^+$$

Then, in exponential model we have:

$$S_t = S_0 e^{\left(\mu \frac{\sigma^2}{2}\right)t + \sigma W_t - \frac{t\sigma^2}{2} + Y_t}$$

and the process S_t satisfies the stochastic differential equation:

$$\begin{aligned} dS_t &= \left(\mu + \frac{\sigma^2}{2}\right) S_t dt + \sigma S_t dW_t + S_{t^-}(e^{\Delta Y_t} - 1) dN_t \\ &= \left(\mu + \frac{\sigma^2}{2}\right) S_t dt + \sigma S_t dW_t + S_{t^-}(e^{Z_{N_t}} - 1) dN_t \end{aligned}$$

We can see that the process S_t has jump size equal to: $S_{t^-}(e^{Z_{N_t}} - 1)$. In order for the discounted price process $(e^{-rt}S_t)_{t\in\mathbb{R}}$ to be a martingale, we need to choose a drift parameter $\tilde{\mu} \in \mathbb{R}$, intensity $\tilde{\lambda} > 0$ and jump distribution $\tilde{\nu}$ satisfying the equation:

$$\mu - r = \sigma\tilde{\mu} - \tilde{\lambda} E_{\tilde{\nu}}[Z]$$

Therefore:

$$\mu + \frac{\sigma^2}{2} - r = \sigma\tilde{\mu} - \tilde{\lambda} E_{\tilde{\nu}}[e^Z - 1]$$

under this condition we can choose a risk-neutral probability $\mathbb{P}_{\tilde{\mu},\tilde{\lambda},\tilde{\nu}}$ under which $(e^{-rt}S_t)_{t\in\mathbb{R}}$ is a martingale, for simplicity of notation we denoted the probability $\mathbb{P}_{\tilde{\mu},\tilde{\lambda},\tilde{\nu}}$ with $\mathbb{Q}$. Then, the discounted expected value with respect the new probability measure represent a non-unique arbitrage price at time $t \in [0, T]$ for the contingent claim with payoff $f(S_T)$, hence we have

$$e^{-r(T-t)} E_{\mathbb{Q}}[f(S_T)|\mathcal{F}_t]$$

Set $\theta = T - t$. Then, we can express this arbitrage price as:

$$\begin{aligned}
e^{-r(T-t)} E_{\mathbb{Q}}[f(S_T)|\mathcal{F}_t] &= e^{-r\theta} E_{\mathbb{Q}}[f(S_0 e^{\mu T + \sigma W_T + Y_T})|\mathcal{F}_t] \\
&= e^{-r\theta} E_{\mathbb{Q}}[f(S_0 e^{\mu(T-t)+\sigma(W_T - W_t) + W_T - Y_t})|\mathcal{F}_t] \\
&= e^{-r\theta} E_{\mathbb{Q}}\left[f\left(x \exp\left(\mu\theta + \sigma(W_T - W_t) + \sum_{i=N_t+1}^{N_t} Z_i \right)\right)\right]_{x=S_t} \\
&= e^{-r\theta - \theta\tilde{\lambda}} \sum_{n=0}^{\infty} \frac{(\tilde{\lambda}(T-t))^n}{n!} E_{\mathbb{Q}}\left[f\left(x e^{\mu\theta + \sigma(W_T - W_t)} \exp\left(\sum_{i=1}^{n} Z_i \right)\right)\right]_{x=S_t} \\
&= e^{-\theta(r+\tilde{\lambda})} \sum_{n=0}^{\infty} \frac{(\tilde{\lambda}\theta)^n}{n!} E_{\mathbb{Q}}\left[f\left(x e^{\mu\theta + \sigma(W_T - W_t)} \exp\left(\sum_{i=1}^{n} Z_i \right)\right)\right]_{x=S_t}
\end{aligned}$$

4.2.3 Implied Volatility

One of the main advantages of the Black-Scholes formula is the fact that the pricing formula, as well as the hedging formula, depend only on one non-observable parameter: the volatility σ. In fact, the drift parameter μ disappears by changing the probability measure. In the Black-Scholes model $\nu = 0$ and the call option prices are uniquely given by the equation (4.29):

$$F(t, x) = C^{BS} = xN(d_1) - Ke^{-r\theta} N(d_2)$$

If we fixed all the parameters of the equation (4.29), we see that the value of the call in the Black-Scholes model is an increasing continuous function of σ, mapping $]0, \infty[$ into $](S_t - Ke^{-r\theta})^+, S_t[$. The last interval represent an arbitrage bound for a call option prices. Therefore, we can defined the Black-Scholes implied volatility of the option, denoted by $\sigma_t^{IV}(T, K)$, as the value of the volatility of the underlying instrument, which when substituted into the Black-Scholes formula, will return the correct option prices, denoted by $C_t^*(T, K)$:

$$\exists!\, \sigma_t^{IV}(T, K) > 0, \qquad C^{BS}(S_t, K, \theta, \sigma_t^{IV}(T, K)) = C_t^*(K, T)$$

We can note that, for fixed (T, K), the implied volatility is in general a stochastic process. Furthermore, if we fixed t, we will find the implied volatility surface at date t, which is equal at the function $\sigma_t^{IV} : (T, K) \to \sigma_t^{IV}(T, K)$. This means that, for fixed t, the implied volatility value depends on the characteristics of the option such as the maturity and the strike price, respectively equal to T and K. Moreover, if we substitute the moneyness m (i.e $m = \frac{K}{S_t}$) into the implied volatility surface, it can be represented as a function of moneyness and time to maturity: $I_t(\theta, m) = \sigma_t^{IV}(t + \theta, mS(t))$. In general, the implied volatility surface $I_t(\theta, m)$ may depend not only on the maturity of options but also on the current date or the spot price. However, in the exponential-Lévy models the evolution in time of implied volatilities is particularly simple, as shown by the following proposition:

Proposition 4.11 (Proposition 11.1 in [2]) *When the risk neutral dynamics is given by an exponential-Lévy process, the implied volatility for a given meneyness level $m = \frac{K}{S_t}$ and time to maturity θ, i.e $\theta = T - t$, does not depend on time:*

$$\forall t \geq 0, \quad I_t(\theta, m) = I_0(\theta, m)$$

Proof

Consider the value of a call in an exponential-Lévy model, given by the equation (4.34):

$$C_t(T = t + \theta, K) = Ke^{-r\theta} E[(e^{x + X_\theta} - 1)^+]$$

If we divided both term by S_t and we substituted the moneyness and the log-forward-moneyness, we find:

$$\begin{aligned}\frac{C_t(T=t+\theta,K)}{S_t} &= \frac{Ke^{-r\theta}E[(e^{x+X_\theta}-1)^+]}{S_t}\\ &= me^{-r\theta}E[(e^{\ln\left(\frac{S}{K}\right)+r\theta+X_\theta}-1)^+]\\ &= me^{-r\theta}E[(m^{-1}e^{r\theta+X_\theta}-1)^+]\\ &= g(\theta,m)\end{aligned}$$

that is, the ratio of option price to the underlying which depends only on the moneyness and time to maturity. We can do the same for the price of the call option in the Black-Scholes model and we find:

$$\frac{C^{BS}}{S_t} = g^{BS}(\theta,m,\sigma)$$

this is true because the Black-Scholes model is a particular case of the exponential-Lévy model. Therefore, the implied volatility $I_t(\theta,m)$ is defined by solving the equation:

$$C^{BS} = S_t g(\theta,m) \iff g^{BS}(\theta,m,I_t(\theta,m)) = g(\theta,m)$$

Since each side does not depend on t but depends only on (θ,m), we can conclude that the implied volatility for a given time to maturity θ and moneyness m does not evolve in time:

$$\forall t\geq 0,\quad I_t(\theta,m) = I_0(\theta,m)$$

□

However, we can note that the implied volatility for a given strike price, K, is not constant in time. In fact, it evolves stochastically according to:

$$\sigma_t^{IV} = I_0\left(\frac{K}{S_t},T-t\right)$$

We can note that the implied volatility surface I_t does not vary with t, therefore we can study only the case in which $t=0$. This study explain some features of the implied volatility surface in the exponential-Lévy model. First, a negative skewed jump distribution give rise to a skew in implied volatility, hence the skew decrease characteristic with respect to moneyness. On the other hand, a strong variance of jumps leads to a curvature in the implied volatility, hence we can see smile pattern. Second, exponential-Lévy models and, in general, model with jumps in the price lead to a strong short term skew contrarily to diffusion models which have small skew for short maturities. Finally, in a Lévy process with finite variance we can see the effect called aggregation normality, which is when long maturity prices of options will be cloeser to Black-Scholes price and the implied volatility smile will become flat as $T\to\infty$. In particular, the central limit theorem shows that when the maturity T is large, the distribution of $(X_T - E[X_T])/\sqrt{T}$ becomes approximately Gaussian. This effect is more pronounced in exponential-Lévy models respect to the actual market prices.

4.3 Hedging Strategy

Consider an asset prices $(S_t)_{t\in[0,T]}$ and a market described by a filtered probability space $(\Omega,\mathcal{F},(\mathcal{F}_t)_{t\in[0,T]},\mathbb{P})$, where $(\mathcal{F}_t)_{t\in[0,T]}$ is the history of the assets, $\mathbb{P}$ represents the so-called real-word measure and S_t will be one dimensional. We assume that there are two assets in the market: a riskless asset, described by the following differential equation $dS_t^0 = rS_t^0dt$, and a risky asset, S_t. Let $S_t^0 = e^{rt}$ be a numeraire. Then, we denoted by V_t the value of a portfolio and by $\tilde{V}_t$ its discounted value, which is equal to $\tilde{V}_t = V_t/S_t^0$.

4.3.1 Black-Scholes Hedging

Consider the Black-Scholes model, which is described above (4.2.1). Therefore, the behavior of the stock price is represent by equation (4.25) :

$$S_t = S_0 e^{(\mu - \frac{\sigma^2}{2})t + \sigma W_t}$$

and the price of call option in the Black-Scholes model are equal to the equation (4.29):

$$C^{BS}(t,S) = SN(d_1) - Ke^{-r\theta}N(d_2) \tag{4.36}$$

where $\theta = T - t$ and $d_1 = \dfrac{\ln\left(\frac{S}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)\theta}{\sigma\sqrt{\theta}}$ and $d_2 = d_1 - \sigma\sqrt{\theta}$.
Let V be the value of a portfolio of derivative securities on one underlying asset. The rates of change of the value of the portfolio with respect to the spot price S of the underlying asset is important for hedging purpose. This change is called "Delta" and is equal to:

$$\Delta(V) = \frac{\partial V}{\partial S}$$

Then, the delta of the call option described in equation (4.36) is equal to:

$$\begin{aligned}\Delta(C^{BS}) &= \frac{\partial C^{BS}}{\partial S} \\ &= N(d_1) + S\frac{\partial}{\partial S}(N(d_1)) - Ke^{-r\theta}\frac{\partial}{\partial S}(N(d_2))\end{aligned} \tag{4.37}$$

If we apply the chain rule, we obtain that:

$$\begin{aligned}\frac{\partial}{\partial S}(N(d_1)) &= N'(d_1)\frac{\partial d_1}{\partial S} \\ \frac{\partial}{\partial S}(N(d_2)) &= N'(d_2)\frac{\partial d_2}{\partial S}\end{aligned}$$

Therefore, we can write equation (4.37) as:

$$\Delta(C^{BS}) = N(d_1) + SN'(d_1)\frac{\partial d_1}{\partial S} - Ke^{-r\theta}N'(d_2)\frac{\partial d_2}{\partial S} \tag{4.38}$$

Now, recall that d_1 and d_2 are equal to:

$$d_1 = \frac{\ln\left(\frac{S}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)\theta}{\sigma\sqrt{\theta}} = \frac{\ln\left(\frac{S}{K}\right) + r\theta}{\sigma\sqrt{\theta}} + \frac{\sigma\sqrt{\theta}}{2} \tag{4.39}$$

$$d_2 = d_1 - \sigma\sqrt{\theta} = \frac{\ln\left(\frac{S}{K}\right) + r\theta}{\sigma\sqrt{\theta}} - \frac{\sigma\sqrt{\theta}}{2} \tag{4.40}$$

Lemma 4.1 *Let d_1 and d_2 be given by equation* (5.4) *and* (5.5). *Then, we have the following result:*

$$SN'(d_1) = Ke^{-r\theta}N'(d_2)$$

Proof

Recall that $N(d) = \dfrac{1}{\sqrt{2\pi}}\displaystyle\int_{-\infty}^{d} e^{-x^2/2}dx$ is the Gaussian cumulative distribution function. It is easy to see that $N'(d) = \frac{1}{\sqrt{2\pi}}e^{-d^2/2}$, hence we find

$$\begin{aligned}N'(d_1) &= \frac{1}{\sqrt{2\pi}}e^{-d_1^2/2} \\ N'(d_2) &= \frac{1}{\sqrt{2\pi}}e^{-d_2^2/2}\end{aligned}$$

Therefore, we need to show that the following formula is true:

$$Se^{-d_1^2/2} = Ke^{-r\theta}e^{-d_2^2/2}$$

can also be written as

$$\frac{Se^{r\theta}}{K} = e^{(d_1^2-d_2^2)/2}$$

Now, if we substitute equation (4.39) and (4.40) in the left term of the above equation, we find:

$$\begin{aligned}
d_1^2 - d_2^2 &= d_1^2 - \left(d_1 - \sigma\sqrt{\theta}\right)^2 \\
&= d_1^2 - d_1^2 + \sigma^2\theta + 2d_1\sigma\sqrt{\theta} \\
&= 2\left(\frac{\ln\left(\frac{S}{K}\right) + r\theta}{\sigma\sqrt{\theta}} + \frac{\sigma\sqrt{\theta}}{2}\right)\sigma\sqrt{\theta} - \sigma^2\theta \\
&= 2\left(\ln\left(\frac{S}{K}\right) + r\theta\right)
\end{aligned}$$

Therefore, we have

$$\begin{aligned}
\frac{Se^{r\theta}}{K} &= e^{(d_1^2-d_2^2)/2} \\
&= exp\left(\ln\left(\frac{S}{K}\right) + r\theta\right) \\
&= \frac{Se^{r\theta}}{K}
\end{aligned}$$

□

From (4.39) and (4.40), we find that

$$\frac{\partial d_1}{\partial S} = \frac{\partial d_2}{\partial S} = \frac{1}{\sigma S\sqrt{\theta}} \tag{4.41}$$

Using equation (4.41) and lemma 4.1 in the equation (4.38), we find that

$$\begin{aligned}
\Delta(C^{BS}) &= N(d_1) + SN'(d_1)\frac{\partial d_1}{\partial S} - Ke^{-r\theta}N'(d_2)\frac{\partial d_2}{\partial S} \\
&= N(d_1) + SN'(d_1)\left(\frac{\partial d_1}{\partial S} - \frac{\partial d_2}{\partial S}\right) \\
&= N(d_1) \qquad (4.42)
\end{aligned}$$

Hence, we have found that the delta for an European call option in the Black-Scholes model is equal to the cumulative distribution function of a standard normal variable evaluated in d_1.

The delta in an option is important because helps to build the so-called "delta hedging". Assume that we go long in one call option. If the price of the underlying asset declines, the value of the call decreases and the long call position loses money. To protect against a downturn in the price of the underlying asset, we can sell short Δ units of the underlying asset. The goal of the delta hedging is to choose Δ in such a way that the value of the portfolio is not sensitive to small changes in the price of the underlying asset. Therefore, if V is the value of the portfolio, the value of the hedge portfolio is

$$V = C(t, S) - \Delta S_t$$

We can note that a portfolio is delta-neutral only over a short period of time. When the price of the underlying asset changes, the porfolio might become unbalanced.

4.3.2 Merton Approach

The delta hedging in the Black-Scholes model is always possible since the market is complete and, therefore, exists only one equivalent risk neutral probability. This is the main assumption in the Black-Scholes model. Unfortunately, the market is not complete and there is not a unique risk neutral probability because the asset has discontinuities, i.e jumps, in their paths.

The first application of jump process in option pricing was introduced by Merton[8]. Merton considered the following jump diffusion model defined in the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$:

$$S_t = S_0 \exp\left(\mu t + \sigma W_t + \sum_{i=1}^{N_t} Y_i\right) \tag{4.43}$$

where W_t is a standard Wiener process, N_t is a Poisson process with intensity λ independent from W and $Y_i \sim N(m, \delta^2)$ are i.i.d. random variables independent from W and N. Since the model is incomplete, there exists many possible choices for a risk-neutral measure and Merton proposed to change the drift of the Wiener process and keep the other variable unchanged. Therefore, μ^M is chosen such that $\hat{S}_t = S_t e^{-rt}$ is a martingale under the new probability measure $\mathbb{Q}$, which is the equivalent probability measure to $\mathbb{P}$, and is equal to

$$\mu^M = r - \frac{\sigma^2}{2} - \lambda E[e^{Y_i} - 1] = r - \frac{\sigma^2}{2} - \lambda\left[\exp\left(m + \frac{\delta^2}{2}\right) - 1\right]$$

The equivalent martingale measure is obtained by shifting the drift of the Brownian motion but leaving the jump part unchanged. Merton justified this choice by assuming that the jump risk is diversificable and, therefore, no risk premium is attached to it. Then, an European call option with payoff $f(S_T)$ can be priced according to:

$$C^M(t, S_t) = e^{-r(T-t)} E_{\mathbb{Q}}[f(S_T)|\mathcal{F}_t] \tag{4.44}$$

Set $\theta = T - t$. Then, we can express this arbitrage price as:

$$\begin{aligned} C^M(t, S_t) &= e^{-r(T-t)} E_{\mathbb{Q}}[f(S_T - K)^+ | S_t = S] \\ &= e^{-r\theta} E[f\left(e^{\mu^M\theta + \sigma W^M_{T-t} + \sum_{i=1}^{N_{T-t}} Y_i}\right)] \end{aligned}$$

By conditioning on the number of jumps N_t, we can express the value of the call option as a weighted sum of Black-Scholes price, therefore we find:

$$\begin{aligned} C^M(t, S_t) &= e^{-r\theta} \sum_{n\geq 0} \mathbb{Q}(N_t = n) E_{\mathbb{Q}}\left[f\left(S \exp\left(\mu^M\theta + \sigma W^M_\theta + \sum_{i=1}^{n} Y_i\right)\right)\right] \\ &= e^{-r\theta} \sum_{n\geq 0} \frac{e^{-\lambda\theta}(\lambda\theta)^n}{n!} E_{\mathbb{Q}}\left[f\left(S e^{nm + \frac{n\delta^2}{2} - \lambda\exp(m+\frac{\delta^2}{2}) + \lambda\theta} e^{r\theta - \frac{\sigma_n^2\theta}{2} + \sigma_n W_\theta}\right)\right] \\ &= e^{-r\theta} \sum_{n\geq 0} \frac{e^{-\lambda\theta}(\lambda\theta)^n}{n!} C^{BS}(\theta, S_n, \sigma_n) \end{aligned} \tag{4.45}$$

where

i. $\sum_{i=1}^n Y_i \sim N(nm, n\delta^2)$;

ii. $\sigma_n^2 = \sigma^2 + \frac{n\delta^2}{\theta}$;

iii. $S_n = S\exp\left(nm + \frac{n\delta^2}{2} - \lambda\exp(m + \frac{\delta^2}{2}) + \lambda\theta\right)$;

[8] A brief introduction about the Merton model is given in chapter 2

iv. $C^{BS}(\theta, S, \sigma) = e^{-r\theta} E\left[f\left(Se^{(r-\frac{\sigma^2}{2})\theta+\sigma W_\theta}\right)\right]$.

The point *iv.* is the value of the European call option with time to maturity θ and payoff f in a Black-Scholes model with volatility σ. We can note that if $\lambda = 0$ then $C^M(t, S) = C^{BS}(t, S)$, indeed all the terms appearing in the sum (4.45) are equal to 0, except for $j = 0$, when $S_0 = S$ and $\sigma_0 = \sigma$.
The hedging portfolio proposed by Merton is the self-financing strategy (ϕ_t^0, ϕ_t) given by:

$$\phi_t = \frac{\partial C^M}{\partial S}(t, S_{t^-}) \tag{4.46}$$

$$\phi_t^0 = \phi_t S_t - \int_0^t \phi dS \tag{4.47}$$

which means that we choose to hedge only the risk represented by the diffusion part. This approach is justified if we assume that the investor holds a portfolio with many assets for which the diffusion components may be correlated but the jumps components are independent across assets. This hypothesis would imply that in a large market a diversified portfolio such as S&P500 would not have jumps. Finally, the assumption of diversifiability of jump risk is not justifiable if we are pricing index options, in fact a jump in the index is not diversificable.
We can note that in model with jumps, contrarily to diffusion models, a pricing measure cannot be simply obtained by adjusting the drift coefficient.

4.3.3 Quadratic Hedging

We can define the quadratic hedging as the choice of a hedging strategy which minimizes the hedging error in a mean square losses. This imply that losses and gains are treated in a symmetric manner, therefore we measure the risk in terms of variance.
Consider a risk-neutral model $(S_t)_{t\in[0,T]}$ given by $S_t = e^{rt+X_t}$, where X_t is a Lévy process on the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{Q})$. We assume that S is a square integrable martingale, therefore the following condition is satisfied:

$$\int_{|y|\geq 1} e^{2y}\nu(dy) < \infty$$

Moreover, we assume that X_t has finite variance and its characteristic function can be expressed as:

$$E\left[e^{iuX_t}\right] = \exp\left\{t\left[-\frac{\sigma^2 u^2}{2} + b_X t + \int \nu_X(dy)\left(e^{iuy} - 1 - iuy\right)\right]\right\}$$

with b_X chosen such that $\hat{S} = e^X$ is a martingale. As we have seen in the previous chapter, $\hat{S}_t$ can also be written as a stochastic exponential of another Lévy process (Z_t):

$$d\hat{S}_t = \hat{S}_t dZ_t$$

where Z is a martingale with jumps size greater than -1 and it is also a Lévy process. Let $(\phi_t^0, \phi_t)_{t\in[0,T]}$ be a self-financing strategy. In order to apply the quadratic hedging criteria, we need to find portfolio such that its terminal value has a well-defined variance. Therefore, we want that the asset S is in the set of all the admissible strategies defined as:

$$\mathcal{S} = \left\{\phi \text{ caglad predictable and } E\left[\left|\int_0^T \phi_t d\hat{S}_t\right|\right]^2 < \infty\right\}$$

Using preposition 3.5 and the proposition 3.11, the above condition is equivalent to:

$$E\left[\int_0^T |\phi_t \hat{S}_t|^2 dt + \int_0^T \int_{\mathbb{R}} z^2 |\phi_t \hat{S}_t|^2 dt \nu(dz)\right] < \infty \tag{4.48}$$

Let $L^2(\mathcal{S})$ be the set of process ϕ which verify the above condition (4.48). Therefore, the terminal payoff of such strategy is equal to:

$$G_T(\phi) = \int_0^T r\phi_t^0 dt + \int_0^T \phi_t S_{t^-} dZ_u$$

We can note that $\hat{S}_t$ is a martingale under the probability measure $\mathbb{Q}$ and that $\phi \in L^2(\hat{S})$, therefore the discounted gain process, equal to $\hat{G}_T(\phi) = \int_0^t \phi d\hat{S}$, is also a square integrable martingale. Using proposition 2.23 we find that $\hat{G}_T(\phi)$ is given by the martingale part of the above equation:

$$\begin{aligned}
\hat{G}_T(\phi) &= \int_0^T \phi_t S_{t^-} \sigma dW_t + \int_0^T \int_{\mathbb{R}} \tilde{J}_X(dt \times dx) x \phi_t S_{t^-} \\
&= \int_0^T \phi_t S_{t^-} \sigma dW_t + \int_0^T \int_{\mathbb{R}} \tilde{J}_Z(dt \times dz) \phi_t S_{t^-} (e^z - 1)
\end{aligned}$$

where J is the jump measure[9]. Now, we can written the quadratic hedging problem as:

$$\inf_{\phi \in L^2(\hat{S})} E_{\mathbb{Q}} \left[|\hat{G}_T(\phi) + V_0 - \hat{H}|^2 \right] \tag{4.49}$$

where $\hat{H}$ is defined by the equation (4.3), i.e $\hat{H} = V_0 + \int_0^T \phi_t d\hat{S}_t$ $\mathbb{Q}$-a.s. We can note that the expectation of the hedging error is equal to $V_0 - E_{\mathbb{Q}}[\hat{H}]$, therefore if we decomposed the above equation into

$$E_{\mathbb{Q}} \left[|V_0 - E_{\mathbb{Q}}[\hat{H}]|^2 \right] + Var_{\mathbb{Q}} \left(\hat{G}_T(\phi) - \hat{H} \right)$$

we will se that the optimal value for the initial capital is: $V_0 = E_{\mathbb{Q}}[f(S_T)]$.
Proposition 4.12 (proposition 10.5 in [2]) *Consider the risk neutral dynamics*

$$\mathbb{Q}: \qquad d\hat{S}_t = \hat{S}_{t^-} dZ_t \tag{4.50}$$

where Z is a Lévy process with Lévy measure ν_Z and diffusion coefficient $\sigma > 0$. For a European option with payoff $f(S_T)$ where $f : \mathbb{R}^+ \to \mathbb{R}$ verifies

$$\exists K > 0, \qquad |f(x) - f(y)| \leq K|x - y|$$

the risk minimizing hedge, solution of (4.49)*, amounts to holding a position in the underlying equal to $\phi_t = \Delta(t, S_{t^-})$ where:*

$$\Delta(t, S_{t^-}) = \frac{\sigma^2 \frac{\partial C}{\partial S}(t, S) + \frac{1}{S} \int \nu_Z(dy) z [C(t, S(1+z)) - C(t, S)]}{\sigma^2 + \int z^2 \nu_Z(dy)} \tag{4.51}$$

with $C(t, S) = e^{-r(T-t)} E_{\mathbb{Q}}[f(S_T)|S_t = S]$.

Proof
We know that the discounted price $\hat{S}_t$ is a martingale under the risk-neutral measure $\mathbb{Q}$. Consider a self-financing strategy given by a nonanticipating càglàd process (ϕ_t^0, ϕ_t) with $\phi \in L^2(\hat{S})$. The discountend value of the portfolio $(\hat{V}_t)$ is then a martingale whose terminal value is given by:

$$\begin{aligned}
\hat{V}_T &= \int_0^T \phi_t d\hat{S}_t = \int_0^T \phi_t \hat{S}_{t^-} dZ_t \\
&= \int_0^T \phi_t \hat{S}_t \sigma dW_t + \int_0^T \int_{\mathbb{R}} \phi_t \hat{S}_t z \tilde{J}_Z(dt \times dz)
\end{aligned} \tag{4.52}$$

[9]Defined in chapter 2

Now, we can define the function

$$C(t,S) = e^{-r(T-t)}E_{\mathbb{Q}}[f(S_T)|\mathcal{F}_t] = e^{-r(T-t)}E_{\mathbb{Q}}[f(S_T)|S_t = S]$$

and its discountend value by $\hat{C}(t,S) = e^{-rt}C(t,S)$. We can note that $\hat{C}(t,S) = e^{-rt}E_{\mathbb{Q}}[f(S_T)|\mathcal{F}_t]$ is a square integrable martingale by construction and that $C(t,S)$ is continuously differentiable with respect to t ans twice continuously differentiable with respect to S. Therefore, we can applied the Ito formula to $\hat{C}(t,S_t) = e^{-rt}C(t,S_t)$ in the interval $[0,t]$, and we find:

$$\begin{aligned}
\hat{C}(t,S_t) - \hat{C}(0,S_0) &= \int_0^t \frac{\partial C}{\partial S}(u,S_{u^-})\hat{S}_{u^-}\sigma dW_u \\
&\quad + \int_0^t \int_{\mathbb{R}} [C(u,S_{u^-}(1+z)) - C(u,S_{u^-})]\tilde{J}_Z(du \times dz) \\
&= \int_0^t \frac{\partial C}{\partial S}(u,S_{u^-})\hat{S}_{u^-}\sigma dW_u \\
&\quad + \int_0^t \int_{\mathbb{R}} [C(u,S_{u^-}e^x) - C(u,S_{u^-})]\tilde{J}_X(du \times dx) \qquad (4.53)
\end{aligned}$$

where (X_t) is a Lévy process such that $\hat{S}_t = e^{X_t}$ for all t. The payoff function $f(S_T)$ is Lipschitz continuous, this imply that also C is Lipschizt continuous with respect to the second variable:

$$\begin{aligned}
C(t,x) - C(t,y) &= e^{-r(T-t)}E\left[f\left(xe^{r(T-t)+X_{T-t}}\right) - f\left(ye^{r(T-t)+X_{T-t}}\right)\right] \\
&\le K|x-y|E\left[e^{X_{T-t}}\right] = K|x-y|
\end{aligned}$$

since e^{X_t} is a martingale. Therefore, the predictable random function $\psi(t,z) = [C(u,S_{u^-}(1+z)) - C(u,S_{u^-})]$ verifies

$$\begin{aligned}
E\left[\int_0^T dt \int_{\mathbb{R}} \nu_Z(dz)|\psi(t,z)|^2\right] = \\
&= E\left[\int_0^T dt \int_{\mathbb{R}} \nu_Z(dz)|[C(u,S_{u^-}(1+z)) - C(u,S_{u^-})]|^2\right] \\
&\le E\left[\int_0^T dt \int_{\mathbb{R}} z^2 S_{t^-}^2 \nu(dz)\right] < \infty
\end{aligned}$$

so from proposition 3.11, the compensated Poisson integral in (4.53) is a square integral martingale. Then, if we subtract the equation (4.52) from the equation (4.53), we will find the hedging error:

$$\begin{aligned}
\varepsilon(V_0,\phi) &= \int_0^T \left(\phi_t\hat{S}_{t^-} - \hat{S}_{t^-}\frac{\partial C}{\partial S}(t,S_{t^-})\right)\sigma dW_t \\
&\quad + \int_0^T dt \int_{\mathbb{R}} \tilde{J}_Z(dt \times dz)\left[z\phi_t\hat{S}_{t^-} - (C(t,S_{t^-}(1+z)) - C(t,S_{t^-}))\right]
\end{aligned}$$

where each stochastic integral is a well-defined, zero-mean random variable with finite variance. Finally, we can compute the variance of the hedging error thanks to the isometry formula given by the preposition 3.5 and 3.11:

$$\begin{aligned}
E\left[|\varepsilon(\phi)|^2\right] &= E\left[\int_0^T dt \int_{\mathbb{R}} \nu_Z(dz)|C(t,S_{t^-}(1+z)) - C(t,S_{t^-}) - z\phi_t\hat{S}_{t^-}|^2\right] \\
&\quad + E\left[\int_0^T \hat{S}_{t^-}^2\left(\phi_t - \frac{\partial C}{\partial S}(t,S_{t^-})\right)^2 \sigma^2 dt\right] \qquad (4.54)
\end{aligned}$$

We can note that the terms under the integral in the equation (4.54) are positive process which are a quadratic function of ϕ_t with coefficients depending on (t, S_{t^-}). Therefore, the optimal risk-minimizing hedge is obtained by minimizing this equation respect to ϕ_t, which means that we find the first order condition:

$$\hat{S}^2_{t^-}\left(\phi_t - \frac{\partial C}{\partial S}(t, S_{t^-})\right) + \int_{\mathbb{R}} \nu_Z(dz) z \hat{S}_{t^-}\left[z\phi_t \hat{S}_{t^-} - C(t, S_{t^-}(1+z)) - C(t, S_{t^-})\right] = 0$$

whose solution is given by the equation (4.51).

□

If we consider an exponential-Lévy model, i.e. $S_t = S_0 e^{rt+X_t}$, the optimal quadratic hedge can be expressed in terms of the Lévy measure ν_X of X as

$$\Delta(t, S_{t^-}) = \frac{\sigma^2 \frac{\partial C}{\partial S}(t, S) + \frac{1}{S}\int \nu_X(dx)\,(e^x - 1)\,[C(t, Se^x) - C(t, S)]}{\sigma^2 + \int (e^x - 1)^2 \nu_X(dx)}$$

We can note that we have also found an expression for the residual risk of a hedging strategy (ϕ_t^0, ϕ_t):

$$R_T(\phi) = E\left[\int_0^T \left|\phi_t - \frac{\partial C}{\partial S}(t, S_{t^-})\right|^2 \hat{S}^2_{t^-} dt\right]$$
$$+ E\left[\int_0^T dt \int_{\mathbb{R}} \nu(dz) |C(t, S_{t^-}(1+z)) - C(t, S_{t^-}) - z\phi_t \hat{S}_{t^-}|^2\right]$$

The residual risk allows us to examine whether there are any cases where the hedging error can be reduced to zero, hence where we can achieve a perfect hedge. We find that in only two case is possible achieve a perfect hedge. The first one, is when there are no jumps, i.e $\nu = 0$. In this case, the residual risk is equal to:

$$\varepsilon(\phi) = E\left[\int_0^T \left(\phi_t S_{t^-} - S_{t^-}\frac{\partial C}{\partial S}(t, S_{t^-})\right)^2 dt\right]$$

and we find that $\varepsilon(\phi) = 0$ a.s when ϕ_t is equal to the Black-Scholes delta hedging. The second case, is when $\sigma = 0$ and there is a single jump size $\nu = \delta_a : X_t = aN_t$, where N is a Poisson process. In this case

$$R_T(\phi) = E\left[\int_0^T dt S^2_{t^-} |C(t, S_{t^-}(1+a)) - C(t, S_{t^-}) - \phi_t|^2\right]$$

if we choose $\phi_t = \frac{C(t, S_{t^-}(1+a)) - C(t, S_{t^-})}{aS_{t^-}}$ and $\phi_t^0 = e^{rt} S_t \phi_t - e^{rt}\int_0^t \phi_t dS_t$, we will obtain a self-financing strategy $(\phi, \phi^0$ which is a replication strategy:

$$f(S_T) = V_0 + \int_0^T \frac{C(t, S_{t^-}(1+a)) - C(t, S_{t^-})}{aS_{t^-}} dS_t + \int_0^T r\phi_t^0 dt$$

We can note that the quadratic hedge achieves a mean-variance trade-off between the risk due to the diffusion part and the jump risk.

Another solution proposed by Föller and Schweizer was to find a new martingale measure $\mathbb{Q}^{FS}$ which is orthogonal to the probability $\mathbb{P}$. If $\hat{S}_t = M_t + A_t$ where M is the martingale component of S under $\mathbb{P}$, any martingale (N_t) which is orthogonal to (M_t) under $\mathbb{P}$ should remain a martingale orthogonal to $\hat{S}$ under $\mathbb{Q}^{FS}$. Such probability $\mathbb{Q}^{FS}$ is called minimal martingale measure.
Consider now a jump diffusion model:

$$\frac{dS_t}{S_{t^-}} \qquad \text{where} \quad Z_t = \mu t + \sigma W_t + \sum_{i=1}^{N_t} Y_i \tag{4.55}$$

with $Y_i \sim F$ i.i.d random variable and N a Poisson process with intensity λ, which implies that $E[N_t] = \lambda t$. Moreover, we assume that $E[Y_i] = m$ and $Var(Y_i) = \delta^2$. In this case, the minimal martingale measure exists if and only if

$$-1 \leq \eta = \frac{\mu + \lambda m - r}{\sigma^2 + \lambda(\delta^2 + m^2)} \leq 0$$

Zhang show this result in 1994[10].This assumption means that the risk premium in the asset return should be negative. When this condition is verified, the minimal martingale measure $\mathbb{Q}^{FS}$ is equal to

$$\frac{d\mathbb{Q}^{FS}}{d\mathbb{P}} = e^{-\sigma\eta W_T + \lambda\eta m T - \frac{\sigma^2\eta^2}{2}T} \prod_{j=1}^{N_t} (1 - \eta U_j)^{N_T}$$

Therefore, the risk-neutral dynamics of the asset under $\mathbb{Q}^{FS}$ can be expressed as:

$$\frac{dS_t}{S_{t^-}} = rdt + dU_t$$

$$U_t = \lambda[\eta(m^2 + \delta^2) - m]t + \sigma W_t' + \sum_{i=1}^{N_t'} \Delta U_i \qquad (4.56)$$

where under $\mathbb{Q}^{FS}$: W_t' is a standard Wiener process, N_t' is a Poisson process with intensity $\lambda' = \lambda(1 - \eta m)$ and the jump sizes (ΔU_i) are i.i.d with distribution F_U where $dF_U = \frac{1-\eta x}{1-\eta m} dF(x)$. Then, the locally risk minimizing hedge for a European option with payoff $f(S_T)$, which verify the Lipschitz continuous property, for this jump diffusion model (4.55) is given by $\phi_t = \Delta(t, S_{t^-})$ where

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

with $C(t, S) = e^{-r(T-t)} E_{\mathbb{Q}^{FS}}[f(S_T)|S_t = S]$ is the expected discounted payoff taken with respect to (4.56).

[10]in his PhD thesis "Analyse Numérique des Options Américaines dans un Modèle de Diffusion avev Sauts"

4.4 Comparison

We want to show that the hedging in the Merton model outperforms the hedging in the Black and Scholes model, which are described in the section above. Before we talk about the hedging strategy, we show that the Marton model also outperform the Black-Scholes model to replicate the stock behavior from historical data. We consider the daily log-returns of the Standard & Poor's 500 Index (S&P500) in the period from 31-12-2009 to 29-01-2009. There is a total of 2273 daily closing price and we have to deal with n=2272 log-returns. Moreover, from the S&P500 data it is possible to find the following information:

$$\begin{aligned} E^{SP} &\simeq 0.00036; \\ M_2^{SP} &\simeq 0.0095; \\ s^{SP} &\simeq -0.4666 < 0; \\ k^{SP} &\simeq 7.5614 > 3; \end{aligned}$$

where E is the mean, s is the skewness and k is the kurtosis. In order to find a relationship among the two models and the statistical result of the S&P500, we will work with an interval of amplitude Δt, which can be defined as $\Delta t = 1/252 \simeq 0.004$ where the denominator 252 represent the trading days in a year. Therefore, we can write the Black-Scholes model as

$$\Delta \log S_t = \mu_{BS} \Delta T + \sigma_{BS} \Delta W_t \tag{4.57}$$

where $\Delta W_t \sim N(0, \Delta t)$. While the Merton model can be written as:

$$\Delta \log S_t = \mu_M \Delta T + \sigma_M \Delta W_t + V \Delta N_t \tag{4.58}$$

where V is the price ratio (> 0) associated with the i-th jump along the path of the stock price and is equal to $V = \log\left(\frac{S_{T_i}}{S_{T_{i-}}}\right) \sim N(m, \delta^2)$ and $\Delta N_t \sim Po(\lambda \Delta t)$. Then, the following theorem described the relation among the parameter of the two model:

Theorem 4.4 (theorem 1 in [12]) *Consider the equation* (4.57)*, we find the following relation:*

$$\begin{aligned} E^{BS} &= \mu_{BS} \Delta t; \\ M_2^{BS} &= \sigma_{BS}^2 \Delta t; \\ M_3^{BS} &= 0 \Longrightarrow s^{BS} = \frac{M_3^{BS}}{\left(M_2^{BS}\right)^{3/2}}; \\ M_4^{BS} &= 3\sigma_{BS}^4 \Delta t^2 \Longrightarrow k^{BS} = \frac{M_4^{BS}}{\left(M_2^{BS}\right)^2}. \end{aligned}$$

while for the equation (4.58) *we have*

$$\begin{aligned} E^M &= \mu_M \Delta t + m\lambda \Delta t; \\ M_2^M &= \sigma_M^2 \Delta t + \left(\delta^2(1 + \lambda \Delta t) + m^2\right) \lambda \Delta t; \\ M_3^M &= m(3\delta^2 + m^2)\lambda \Delta t + 6m\delta^2(\lambda \Delta t)^2 \Longrightarrow s^M = \frac{M_3^M}{\left(M_2^M\right)^{3/2}}; \\ M_4^M &= 3(\sigma_M^2 \Delta t)^2 + (m^4 + 3\delta^4 + 6m^2\delta^2)\lambda \Delta t + (3m^4 + 21\delta^4 + 30m^2\delta^2)(\lambda \Delta t)^2 \\ &\quad + 6\sigma_M^2 \Delta t(\delta^2 + m^2)\lambda \Delta t + (18\delta^2 + 6m^2\delta^2)(\lambda \Delta t)^3 + 6\sigma_M^2 \delta^2 \Delta t(\lambda \Delta t)^2 \\ &\quad + 3\delta^4(\lambda \Delta t)^4 \Longrightarrow k^M = \frac{M_4^M}{\left(M_2^M\right)^2}. \end{aligned}$$

Proof

The quantities E^{BS} and M_2^{BS} can be derived immediately applying the properties of the

Brownian motion[11]. For the central moments M_4^{BS} we recall that

$$M_i^{BS} = E\left[(\mu_{BS}\Delta t + \sigma_{BS}^2 \Delta W_t - \mu_{BS}\Delta t)^i\right] = \sigma_{BS}^i E[\Delta W_t^i]$$

since $E[\Delta W_t^i]$ can be computed using the characteristic function of the normal random variable ΔW_t, i.e $\phi_{\Delta W_t}(y) = e^{-y^2 \frac{\Delta t}{2}}$, we have

$$E[\Delta W_t^3] = 0, \quad E[\Delta W_t^4] = 3\Delta t^2$$

this prove also the skewness and the kurtosis for the Black and Scholes model.
Switching to the Merton model, the mean term E^M can be obtained recalling that $\Delta N_t \sim Po(\lambda \Delta t)$, while the central moments in M_2^M, M_3^M and M_4^M are derived applying the following formula

$$\begin{aligned} M_i^M &= E\left[(\mu_M \Delta t + \sigma_M^2 \Delta W_t + V\Delta N_t - \mu_M \Delta t - m\lambda\Delta t)^i\right] \\ &= E\left[(\sigma_M^2 \Delta W_t + V\Delta N_t - m\lambda\Delta t)^i\right] \end{aligned}$$

We omit the detailed computation term by term for the above equation, however we can note that we assumed that every variable is independent of the other. Moreover, the moments of V and ΔN_t are derived using the characteristic functions, i.e $\phi_V(y) = e^{imy - \frac{\delta^2 y^2}{2}}$ and $\phi_{\Delta N_t}(y) = e^{\lambda \Delta t (e^{iy} - 1)}$, from which

$$\begin{aligned} E[V^2] &= m^2 + \delta^2 \\ E[V^3] &= 3(m^2 + delta^2)m - 2m^3 \\ E[V^4] &= 3\delta^4 + 6m^2(m^2 + delta^2) - 5m^4 \end{aligned}$$

and

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

□

Therefore, if we apply the theorem 4.4, we can find the vector of parameter for the Black and Scholes model (μ_{BS}, σ_{BS}) and for the Merton model $(\mu_{BS}, \sigma_{BS}, \lambda, m, \delta)$. For the Black-Scholes model we assume:

$$E^{BS} = E^{SP} \;\; M_2^{BS} = M_2^{SP}$$

so that

$$\begin{aligned} \mu_{BS} &= \frac{E^{SP}}{\Delta t} \simeq 0.0922 \\ \sigma_{BS} &= \sqrt{\frac{M_2^{SP}}{\Delta t}} \simeq 0.1507 \end{aligned}$$

we can recall that a normal distribution is completely determined by its mean and variance. On the other hand, in the Merton model we have 5 parameters to estimate. We can reduce this set assuming that

$$E^M = E^{SP} \;\; M_2^M = M_2^{SP}$$

which implies that

$$\begin{aligned} \mu_M &= \frac{E^{SP} - m\lambda\Delta t}{\Delta t} \\ \sigma_M &= \sqrt{\frac{M_2^{SP} - (\delta^2(1 + \lambda\Delta t) + m^2)\,\lambda\Delta t}{\Delta t}} \end{aligned}$$

[11] Described in chapter 1

hence, the diffusion paramters are expressed as function of the jumps ones and we have only 3 parameters to estimate. We use the Multinomial Maximum Likelihood approach to estimate this 3 parameters, which can be represented as a 3-dimensional vectors $\eta \doteq (\lambda, m, \delta)$. The step of the Multinomial Maximum Likelihood approach can be summarized as follows:

1. sort empirical data into $\tilde{n} < n$ bins, in order to get a computationally tractable problem. Then, for each of these bins, extract the sample frequency f_i^{SP}, $i = i, \dots, \tilde{n}$;
2. construct the theoretical jump diffusion frequency function
$$f_i^M(\eta) \doteq n \int_{B_i} \psi_{\Delta t}(y; \eta) dy \quad i = 1, \dots, \tilde{n}$$
where B_i is the i-th bin and $\psi_{\Delta t}(y; \eta)$ is the log-return probability density function for the Merton model (described in chapter 2), i.e.
$$\psi_t(y) = e^{-\lambda t} \sum_{j=0}^{\infty} \frac{(\lambda t)^j \exp\left\{-\frac{(y-(\mu-\lambda k-\frac{1}{2}\sigma)t-jm)^2}{2(\sigma^2 t+j\delta^2)}\right\}}{j!\sqrt{2\pi(\sigma^2 t + j\delta^2)}};$$
3. minimize the objective function
$$l(\eta) \doteq -\sum_{i=1}^{\tilde{n}} f_i^{SP} \log\left(f_i^M(\eta)\right).$$

Therefore, by the Multinomial Maximum Likelihood algorithm we obtain that

$$\lambda \simeq 62.752; \; m \simeq -0.006323; \; \delta \simeq 0.006291$$

hence, in the Merton model μ, σ, skewness and kurtosis are equal to:

$$\begin{aligned}
\mu_M &= \frac{E^{SP} - m\lambda\Delta t}{\Delta t} \simeq 0.48678 \\
\sigma_M &= \sqrt{\frac{M_2^{SP} - (\delta^2(1 + \lambda\Delta t) + m^2)\,\lambda\Delta t}{\Delta t}} \simeq 0.1301 \\
s^M &\simeq 1.4261 \\
k^M &\simeq 7.9952
\end{aligned}$$

We can note that the skewness is bigger than the one obtained using the real S&P500 data, i.e. $1.4261 > -0.4666$, but, unlike in the Black and Scholes model where $s^{BS} = 0$, the Merton approach tents to capture a clear absence of symmetry with the same sign. Moreover, the kurtosis in the Merton model is very close to the one obtained using the real S&P500 data, while the Black and Scholes model provides poor result. Hence, we can conclude that the log-normal jump diffusion model represents a substantial and concrete improvement when compared to the Black and Scholes model.

Now, we compare the Black and Scholes hedging strategy, i.e. Delta hedging, with the Merton hedging for the jump diffusion process. We consider the closing price of the S&P500 from 29-12-2017 to 29-01-2019 and we consider a call option with underlying the S&P500, strike price equal to 2700 and maturity at 01-02-2019. Moreover, we assume that the risk-free rate is equal to 2,98%, denoted by r.
For these period, we have a total of 272 daily closing price and we have to deal with $n = 271$ log-returns. From the log-returns we find the following information from the S&P500 data:

$$\begin{aligned}
E^{SP} &\simeq -0.0000468; \\
M_2^{SP} &\simeq 0.0109; \\
s^{SP} &\simeq -0.4333 < 0; \\
k^{SP} &\simeq 5.9362 > 3;
\end{aligned}$$

Therefore, the Black and Scholes parameters can be estimated as follow:

$$\mu_{BS} = \frac{E^{SP}}{\Delta t} \simeq -0.0118$$
$$\sigma_{BS} = \sqrt{\frac{M_2^{SP}}{\Delta t}} \simeq 0.1723$$

The Merton parameter can be estimated using the Multinomial Maximum Likelihood algorithm and we obtain that

$$\lambda \simeq 3.1596; \ m \simeq -0.04942; \ \delta \simeq 0.0076$$

hence, in the Merton model μ, σ, skewness and kurtosis are equal to:

$$\begin{aligned} \mu_M &= \frac{E^{SP} - m\lambda\Delta t}{\Delta t} \simeq 0.1444 \\ \sigma_M &= \sqrt{\frac{M_2^{SP} - \left(\delta^2(1+\lambda\Delta t) + m^2\right)\lambda\Delta t}{\Delta t}} \simeq 0.1476 \\ s^M &\simeq 1.5929 \\ k^M &\simeq 7.8115 \end{aligned}$$

We can note that also in these case the Merton model represents a substantial and concrete improvement when compared to the Black and Scholes model. Therefore, we can expect that the hedging in the Merton model perform better than the delta hedging in the Black and Scholes.

Then, consider the following hedging strategy for the Black-Scholes model: we assume that we go long in the call option and to protect against a downturn in the price of the underlying asset we will sell short Δ unit of the underlying asset. The goal is to choose Δ in such a way that the value of the portfolio is not sensitive to small changes in the price of the underlying asset. If we denoted with Π the value of the portfolio, then $\Pi = C - \Delta S$ or, equivalently, $\Pi(S) = C(S) - \Delta S$. To implement the Delta hedging we assume that if the Delta is negative we will go long on the asset and short the call option. We can note that a portfolio is Delta neutral only over a short period of time. We recall that the Delta of a call option is equal to the equation (4.42):

$$\Delta(C^{BS}) = N(d_1)$$

where $d_1 = \frac{\ln\left(\frac{S}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)\theta}{\sigma\sqrt{\theta}}$ and $N(d) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{d} e^{-x^2/2} dx$ is the Gaussian cumulative distribution function. To implement the Black-Scholes formula, equation (4.36), the cumulative distribution $N(d)$ of the standard normal variable x must be estimated numerically and we use the algorithm proposed by Abramowitz and Stegun in 1970 which has an approximation error smaller than $7.5 \cdot 10^{-7}$ at any point on the real axis. Abramowitz and Stegun proposed:

$$N(d) = \mathbb{P}(G \leq d)$$

where G is a real Gaussian random variable with mean 0 and variance 1. Therefore $N(d) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{d} e^{-x^2/2} dx$. If $d > 0$ we have that

$$N(d) \approx 1 - \frac{1}{\sqrt{2\pi}} e^{-\frac{d^2}{2}} \left(b_1 t + b_2 t^2 + b_3 t^3 + b_4 t^4 + b_5 t^5\right)$$

where

$$
\begin{aligned}
p &= 0.231641900 \\
b_1 &= 0.319381530 \\
b_2 &= -0.356563782 \\
b_3 &= 1.781477937 \\
b_4 &= -1.821255978 \\
b_5 &= 1.330274429 \\
t &= \frac{1}{1+pd}
\end{aligned}
$$

On the other hand, for the Merton jump diffusion model we consider the hedging prosed by Merton. Therefore, we find that the price of the call option in this model is equal to the equation (4.45)

$$C^M(t, S_t) = e^{-r\theta} \sum_{n \geq 0} \frac{e^{-\lambda\theta}(\lambda\theta)^n}{n!} C^{BS}(\theta, S_n, \sigma_n)$$

where

i. $\sum_{i=1}^{n} Y_i \sim N(nm, n\delta^2)$;

ii. $\sigma_n^2 = \sigma^2 + \frac{n\delta^2}{\theta}$;

iii. $S_n = S \exp\left(nm + \frac{n\delta^2}{2} - \lambda \exp(m + \frac{\delta^2}{2}) + \lambda\theta\right)$;

iv. $C^{BS}(\theta, S, \sigma) = e^{-r\theta} E\left[f\left(Se^{(r-\frac{\sigma^2}{2})\theta + \sigma W_\theta}\right)\right]$;

v. $\theta = T - t$.

The point *iv.* is the value of the European call option with time to maturity θ and payoff f in a Black-Scholes model with volatility σ. The hedging portfolio proposed by Merton is the self-financing strategy (ϕ_t^0, ϕ_t) given by:

$$
\begin{aligned}
\phi_t &= \frac{\partial C^M}{\partial S}(t, S_{t-}) \\
\phi_t^0 &= \phi_t S_t - \int_0^t \phi dS
\end{aligned}
$$

which means that we choose to hedge only the risk represented by the diffusion part.
The result of these two hedging strategy can be seen in the table below which report the return and the variance:

	Return	Variance (σ^2)
Black-Scholes Hedging	6.40%	0.1589%
Merton Hedging	6.49%	0.1589%

We can see that the two hedging strategy have the same variance and the return are more or less the same, the Merton return is a greater only of 0.09 respect to the Black-Scholes return. One possible explanation is that we consider a trading strategy only for one year and, as said before, a portfolio is Delta neutral over a short period of time. Despite this, we can say that the hedging strategy also confirms the above: the Merton model represents a substantial and concrete improvement when compared to the Black and Scholes model. In fact, we can safely say that no-one would choose the Delta hedging compared to Merton hedging as the second has a bigger return, even if small, with the same variance. Therefore, the Merton hedging dominates the Black-Scholes hedging since Merton considers in the stock process a jump component.

Conclusion

In this dissertation we have introduced and explained the jump diffusion process, which are obtained from the Black and Scholes model by adding a compensated compound Poisson process. Moreover, we have seen that the Poisson process produce discontinuities in the stock process and this imply that the market is not complete. Therefore, in the market did not exists a unique risk neutral probability for the option pricing as it was assumed in the Black-Scholes model. In chapter 4, we have seen how pricing an option when the underlying asset is driven by a jump diffusion process and we have seen the impossibility to completely hedge the risk carried by the introduction of sudden and unpredictable moves in the in the stock price, i.e the presence of the random jump component. Finally, we compare the Black and Scholes model to the Merton approach to the jump diffusion process. The results show us that the Merton model turns out to outperform the Black-Scholes one, when we take into account the performances of the two with respect to real financial data. We have also seen that the addition of the jump parameters brings a great improvement in option pricing and hedging. In future work, we will drop the assumption of independence in jump and we will study how correlated jump affect the option pricing and the hedging in the jump diffusion process.

Appendix

A.1 Poisson property

The superposition property say that if $(N_t^1)_{t\geq 0}$ and $(N_t^2)_{t\geq 0}$ are two independent Poisson processes with intensities λ_1, λ_2, then $(N_t^1 + N_t^2)_{t\geq 0}$ is a Poisson process with intensity $\lambda_1 + \lambda_2$.
Proof:
Let $N^1 \sim po(\lambda_1)$ and $N^2 \sim po(\lambda_2)$. Consider the characteristic function for N^1 and N^2:

$$\varphi_{N_1}(t) = E[e^{iuN^1}] = exp\left\{\lambda_1(e^{iu} - 1)\right\}$$
$$\varphi_{N_2}(t) = E[e^{iuN^2}] = exp\left\{\lambda_2(e^{iu} - 1)\right\}$$

Since N^1 and N^2 are two independent Poisson process we have that:

$$\varphi_{N_1+N_2}(t) = \varphi_{N_1}(t)\varphi_{N_2}(t) = exp\left\{\lambda_1(e^{iu} - 1)\right\} exp\left\{\lambda_2(e^{iu} - 1)\right\} = exp\left\{(\lambda_1 + \lambda_2)(e^{iu} - 1)\right\}$$

As the characteristic function completely determines the distribution, we can conclude that $N_1 + N_2 \sim po(\lambda_1 + \lambda_2)$.

□

The following is the proof of the thinning property:
Proof:
First of all, we recall that the generating function of the binomial distribution with parameter n and p and the generating function of the Poisson distribution with parameter λ are equal to:

$$\Pi_X(t) = (q + pt)^n \qquad \text{where } q = 1 - p$$
$$\Pi_X(t) = \sum_{k\geq 0} \frac{e^{-\lambda}\lambda^k}{k!} t^k = e^{-\lambda(1-t)}$$

Therefore

$$\begin{aligned} E[t^{X_N}] &= \sum_{n\geq 0} \frac{e^{-\lambda}\lambda^n}{n!}(q + pt)^n = \\ &= exp\left\{(\lambda q + \lambda pt) - \lambda\right\} \\ &= exp\left\{\lambda pt - \lambda p\right\} \end{aligned}$$

Hence, it follow that X has the Poisson distribution with parameter λp

□

The expected value of N_t can be computed as:

$$\begin{aligned} E[N_t] &= \sum_{k=0}^{\infty} k\mathbb{P}(N_t = k) = e^{-\lambda t} \sum_{k=0}^{\infty} k\frac{(\lambda t)^k}{k!} \\ &= e^{-\lambda t} \sum_{k=1}^{\infty} \frac{(\lambda t)^k}{(k-1)!} = \lambda t e^{-\lambda t} \sum_{k=0}^{\infty} k\frac{(\lambda t)^k}{k!} \\ &= \lambda t \end{aligned}$$

A.2 Donsker's invariance principle

Let $(\xi_n)_{n\in\mathbb{N}}$ are i.i.d. with $E[\xi_i]=0$ and $E[\xi_i^2]=1$. Then, we define $X_n := \sum_{i=1}^n \xi_i$ with $X_0 := 0$ and it satisfies the Central Limit Theorem (CLT), i.e. $\frac{X_n}{\sqrt{n}} \Rightarrow N(0,1)$. We can extend the weak convergence result to the continuous process $(X_t)_{t\geq 0}$ defined by:

$$X_t := X_{\lfloor t\rfloor} + (t-\lfloor t\rfloor)(X_{\lfloor t\rfloor+1} - X_{\lfloor t\rfloor})$$

which linearly interpolates $(X_n)_{n\in\mathbb{N}\cup\{0\}}$ between integer times. Then, we have the following theorem which is an extension of the CLT to the path level:

Theorem: [Donsker's invariance principle] *Let $(\xi_n)_{n\in\mathbb{N}}$ and $(X_t)_{t\geq 0}$ be defined as above. Then $(\frac{1}{\sqrt{n}}X_{nt})_{0\leq t\leq 1} \Rightarrow (B_t)_{0\leq t\leq 1}$ as $C([0,1],\mathbb{R})$-valued random variables, where B_t is a standard Brownian motion.*

A.3 Convolution power

The convolution power is the n-fold iteration of the convolution (is a mathematical operation of two functions to produce a third function that expresses how the shapre of one is modified by the other) with itself. Thus, if x is a function on $\mathbb{R}^d$ and $n \in \mathbb{N}$, then the convolution power is defined by:

$$x^{*n} = \underbrace{x * x * \cdots * x}_{n}, \quad x^{*0} = \delta_0$$

where $*$ denotes the convolution operation of functions on $\mathbb{R}^d$ and δ_0 is the Dirac delta distribution, which is a linear function that maps every function to its value at zero.

A.4 Dirac measure

A Dirac measure is a measure δ_x on a set X (with any $\sigma-$ algebra of subsets of X) defined for a given $x \in X$ and any set $A \subseteq X$ by:

$$\delta_x(A) = \mathbf{1}_A(x) = \begin{cases} 0 & x \notin A \\ 1 & x \in A \end{cases}$$

Therefore, the Dirac measure is a probability measure, and in terms of probability it represents the almost sure outcome x in the sample space X. Moreover, we have: $\displaystyle\int_X f(y)d\delta_x(y) = \int_X f(y)\delta_x(y)dy = f(x)$.

A.5 Cumulants

The cumulants of X are defined by:

$$c_n(X) = \frac{1}{i^n}\frac{\partial^n \Psi_X}{\partial u^n}(0)$$

where Ψ_X is the cumulant generating function or log-characteristic function of X and is defined in a neighbothood of zero such that: $\Psi_X(0) = 0$ and $\psi_X(u) = exp\left[\Psi_X(u)\right]$.

A.6 Fourier transform

The Fourier transform decomposes a function of time into the frequencies that it make it up. Let $f : \mathbb{R} \mapsto \mathbb{C}$, the Fourier transform is equal to:

$$\hat{f}(\varepsilon) = \int_{-\infty}^{\infty} f(x)e^{-2\pi i x\varepsilon}dx \quad \forall\varepsilon \in \mathbb{R}$$

Under suitable condition, f is determined by $\hat{f}$ via the inverse transform:

$$f(x) = \int_{-\infty}^{\infty} \hat{f}(\varepsilon)e^{2\pi i x\varepsilon}d\varepsilon \quad \forall x \in \mathbb{R}$$

A.7 Lipschitz Continuous

***Definition* [Lipschitz continuous]** *A function $\boldsymbol{h} : X \to Y$, X, Y (usually, but not necessarily, subset of $\mathbb{R}^n$) is Lipschitz continuous in X if there exists a constant $M > 0$ such that:*

$$d_Y(\boldsymbol{h}(x_1) - \boldsymbol{h}(x_2)) \leq M d_X(x_1 - x_2), \qquad \forall x_1, x_2 \in X$$

$\boldsymbol{h}$ is locally Lipschitz continuous in X if for every $x \in X$ there exists a neighborhood of it where $\boldsymbol{h}$ is Lipschitz continuous

Bibliography

[1] Damien Lamberton, Bernard Lapeyre, *Introduction to Stochastic Calculus Applied to Finance*, Second edition, Chapman and Hall/CRC, UK.

[2] Rama Cont, Peter Tankov, *Financial Modelling with Jump Processes*, Second edition, Chapman and Hall/CRC, UK (2004).

[3] José E. Figueroa-Lopez, *Jump-Diffusion models driven by Lévy processes*,

[4] Antonis Papapantoleon *An Introduction to Lévy processes with application in Finance*, Lectures Notes, TU Vienna (2008).

[5] Ken-Iti Sato, *Lévy Processes and Infinitely Divisible Distributions*, Cambridge University Press, UK (1994).

[6] Andrea Pascucci, *PDE and Martingale Methods in Option Pricing*, Springer, Bocconi university press, Milano (2010).

[7] Stephen Blyth, *An Introduction to Quantitative Finance*, Oxford university press (2014).

[8] Giuseppe Curci, Mario Dell'Era, *Finanza quantitativa e Modelli Matematici*, Seconda edizione, Pisa university press (2016).

[9] Paolo Baldi, *Equazioni Differenziali Stocastiche e Applicazioni*, Pitagora Editrice, Bologna (2000).

[10] Dan Stefanica, *A Primer for the Mathematics of Financial Engineering*, Second edition, FE Press, New York (2011).

[11] Steven E. Shreve, *Stochastic Calculus fof Finance II, a continuous-time models*, Springer, USA (2004).

[12] Resnik, *A Probability path*,

[13] Allan Gut, *Probability: a graduate course*,

[14] Nicola Gugole, *Merton Jump-diffusion model versus the Black and Scholes approach for the log-returns and volatility smile fitting*, International Journal of Pure and Applied Mathematics, volume 109, pp 719-736.

[15] Barbara Grunewald, Siegfried Tratmann, *Option Hedging in the Presence of Jump Risk*, Department of Law and Economics, Johannes Gutenberg-Universitat Mainz, (1996).

DEPARTMENT OF ECONOMICS AND FINANCE

SUBJECT: MATHEMATICAL METHODS FOR ECONOMICS AND FINANCE

HEDGING STRATEGIES IN JUMP DIFFUSION PRICING MODELS

SUMMARY

SUPERVISOR

Prof. *Fausto Gozzi*

CANDIDATE

Marco Gregnanin

Id. Number: 689941

CO-SUPERVISOR

Prof.ssa *Sara Biagini*

ACADEMIC YEAR 2018/2019

Contents

Introduction

The mathematical modeling of financial market start with Louis Bachelier, who was the first to introduce the Brownian motion as a model for the price fluctuation of a liquid traded financial asset with his doctoral thesis in 1900. In 1973 Fisher Black and Myron Scholes given a great contribution with the article "The Pricing of Option and Corporate Liabilities", which gave a new dimension to the use of probability theory in finance. The option pricing methodology introduced by Black and Scholes is unique in that distributional assumptions alone suffice to generate well-specified option pricing formulas involving mostly observable variable and parameters. One assumption is that the price of the underlying asset follows a diffusion process and an additional assumption is that the instantaneous risk-free rate is nonstochastic and constant. Under these plus other "frictionless market" assumptions, the option's payoff can be replicated by a continuously adjusted hedge portfolio composed of the underlying asset and short-term bonds. This imply that the key assumption in the Black-Scholes model is that the market is complete. In a complete market models probability does not really matter, in fact the objective evolution of the asset is only there to define the set of impossible events and serves to specify the class of equivalent measures. Hence, two statistical models with equivalent measures lead to the same option prices in a complete market setting. Therefore, the option pricing formula generated by Black and Scholes depends critically upon the distributional restriction on the volatility of the underlying asset. The result of that restriction is that the systematic risk of the option is a function of the systematic risk of the underlying asset only.

Jump diffusion process and more in general Lévy models generalize the Black and Scholes work by allowing the stock price to jump while preserving the independence and stationary of returns. Hence, the jump diffusion process described the observed reality of financial markets in a more accurate way than models based only on Brownian motion. In the real world, we observe that the asset price processes have jumps or spikes. Therefore, we can find three main reason for introducing jumps in financial modeling. First, asset price processes have jumps and some risks cannot be handled with a continuous path model but we need to study a discontinuous models. Second, the presence in the option market of the phenomenon of implied volatility smile which shows that the risk-neutral returns are non-gaussian and leptokurtic. Moreover, in continuous path models the law of returns for shorter maturities becomes closer to the Gaussian distribution, on the other hand in models with jumps returns actually become less Gaussian as the maturity becomes shorter. Finally, the jump process correspond to incomplete markets, hence we did not find a unique equivalent probability measure for the option pricing but there are many possible choice. This imply that a perfect hedge, i.e. the Black and Scholes Delta hedging, is not longer possible in jump models and the hedging in jump process achieves a trade-off between the risk due to the diffusion part and the jump risk.

1 Jump Process

In this chapter we introduce and explain a family of discontinuous process called Lévy processes. We begin with the definition of a Poisson process, which is the main building block for stochastic process with discontinuous trajectories. Then, we talk about compound Poisson process, which is use to built a jump-diffusion model, and we study its property. The second section of the chapter, start with the definition of Lévy process, then we discuss its infinitely divisible distribution and we present the Lévy-Khintchine formula, which links processes to distributions.
The last section uses the Lévy processes and its properties to built a model for financial applications, which can be decomposed in two main categories: the jump diffusion model and the infinite activity models. Here, we give some example of jump diffusion model and we explain the properties and the relationship between the ordinary and stochastic exponential models.

1.1 Poisson Process

1.1.1 Definition and Properties

Definition 1.1 [Poisson Process] (definition 7.1.1 in [1]) *Let $(T_i)_{i\geq 1}$ be a sequence of independent, identically, exponentially distributed random variables with parameters $\lambda(\lambda > 0)$ and let $\tau_n = \sum_{i=1}^{n} T_i$. We call Poisson process with intensity λ the process N_t defined by:*

$$N_t = \sum_{n\geq 1} \mathbf{1}_{\{\tau_n \leq t\}} = \sum_{n\geq 1} n\mathbf{1}_{\{\tau_n \leq t < \tau_{n+1}\}}$$

Where N_t indicates the number of points of the sequence $(\tau_n)_{n\geq 1}$ which are smaller than or equal to t. Let $(N_t)_{t\geq 0}$ be a Poisson process and it has the following properties:

1) For all $t \geq 0$, N_t is almost surely (a.s.) finite;

2) The trajectories of N (in other words: $\forall \omega$, the sample path $t \mapsto N_t(\omega)$) are piecewise constant with jumps of size 1;

3) The trajectories are right continuous with left limit (càdlàg);

4) $\forall t > 0$, $N_{t^-} = N_t$ with probability 1;

5) $\forall t > 0$, N_t follows a Poisson distribution with parameter λt:

$$\forall n \in \mathbb{N},\, \mathbb{P}(N_t = n) = e^{-\lambda t}\tfrac{(\lambda t)^n}{n!}$$

6) The characteristic function of N_t is

$$E[e^{iuN_t}] = exp\left\{\lambda t(e^{iu} - 1)\right\},\, \forall u \in \mathbb{R};$$

7) Independence of increments: for all $0 \leq t_0 < t_1 < \cdots < t_n$ and $n \geq 1$ the increments

$$N_{t_1} - N_{t_0}, \ldots, N_{t_n} - N_{t_{n-1}}$$

are mutually independent random variables. In other words, if $s > 0$, $N_{t_n} - N_{t_{n-1}}$ is independent of the σ-algebra $\mathcal{F}_t$;

8) Stationarity of increments: $N_{t+h} - N_{s+h}$ has the same distribution as $N_t - N_s$ for all $h > 0$ and $0 \leq s \leq t$. Hence, the law of $N_{t+s} - N_t$ is identical to the law of $N_s - N_0 = N_s$;

The right continuity, càdlàg property, of the Poisson process is not really a "property". In fact, we have defined N_t in such a way that at a discontinuity point $N_t = N_{t+}$ but a function could be càglàd (left continuous with right limit, in this case we have $f(t) = f(t^-)$ and $N_t = N_{t^-}$). There is a difference between a càdlàg and a càglàg process especially in the context of financial modeling. In fact, if a right continuous function has a jump at time t, then the value $f(t)$ is not predictable by following the trajectory up to time t and the discontinuity is seen as a sudden event. On the other hand, if the function was left continuous, an observer approaching t along the path could predict the value at t. Hence, jumps represent unexpected, unforeseeable events and the assumption of right-continuity is natural. By contrast, we should use a càglàd process if we want to model a discontinuous process whose values are predictable. This will be the case when we want built trading strategies.

1.1.2 Compensated Poisson Processes

The compensated Poisson process define the "centered" version of the Poisson process N_t by

$$\tilde{N}_t = N_t - \lambda t.$$

where λt is the expected value of the Poisson process. $(\tilde{N}_t)$ has centered increments because it has the expected value equal to zero. Moreover, $(\tilde{N})$ follows a centered version of the Poisson law with characteristic function:

$$\psi_{\tilde{N}_t}(u) = exp[\lambda t(e^{iu} - 1 - iu)]$$

$(\tilde{N}_t)_{t\geq 0}$ is called a compensated Poisson process and $(\lambda t)_{t\geq 0}$ is called the compensator of $(N_t)_{t\geq 0}$ and it is the quantity which has to be subtracted from N_t in order to obtain a martingale. Moreover, the compensated Poisson process is no longer integer valued because it is not a counting process unlike the Poisson process.

1.1.3 Compound Poisson Process

The Poisson process itself appears to be too limited to develop realistic asset prices model because the assumption that the jumps size are always equal to 1 is too restrictive, but it can be used as building block to built richer models. Therefore, there is some interest in considering jump processes that can have random jump sizes.

Definition 1.2 [Compound Poisson Process] *The compound Poisson process with jump intensity λ and jump size distribution μ is a stochastic process $(X_t)_{t\geq 0}$ defined by:*

$$X_t = \sum_{i=1}^{N_t} Y_i,$$

where $(Y_i)_{i\geq 1}$ is a sequence of independent random variable with law μ and N_t is a Poisson process with intensity λ independent from $(Y_i)_{i\geq 1}$.

This definition means that a compound Poisson process is a piecewise constant process which jumps at jump times of a standard Poisson process and whose jump size are i.i.d random variables with a given law.

Proposition 1.1 (Characteristic function of the compound Poisson process) (proposition 3.4 in [2]) *Let $(X_t)_{t\geq 0}$ be a compound Poisson process with jump intensity λ and jump size distribution μ. Then X is a piecewise constant Lévy process and its characteristic function is given by:*

$$E[e^{iuX_t}] = exp\left\{\lambda t \int_{-\infty}^{\infty} (e^{iux} - 1)\mu(dx)\right\}. \tag{1}$$

1.1.4 Poisson Random Measures

The definition of the Poisson random measure is a key point for the theory of Lévy processes, which are described in the next section of this chapter.
Definition 1.3 [Random measure] *Let $(\Omega, \mathbb{P}, \mathcal{F})$ be a probability space and let $(E, \mathcal{E})$ be a measurable space. Then $M : \Omega \times \mathcal{E} \to \mathbb{R}$ is a random measure if:*

- *for every $\omega \in \Omega$, $M(\omega, \cdot)$ is a measure on $\mathcal{E}$;*
- *for every $A \in \mathcal{E}$, $M(\cdot, A)$ is measurable.*

We can express a Poisson process in terms of the random measure M in the following way:

$$N_t(\omega) = M(\omega, [0, t]) = \int_{[0,t]} M(\omega, ds)$$

where M is called the random jump measure associated to the Poisson process N.
Definition 1.4 [Jump measure] *Let X be a $\mathbb{R}^d-$valued càdlàg process. The jump measure of X is a random measure on $\mathcal{B}([0, \infty) \times \mathbb{R}^d)$ defined by*

$$J_X(A) = \# \{ t : \Delta X_t \neq 0 \quad \textit{and} \quad (t, \Delta X_t) \in A \} . \tag{2}$$

This definition means that the jump measure of a set of the form $[s, t] \times A$ counts the number of jumps of X between s and t such that their amplitude belongs to A. In other words, J_X contains all the information about the discontinuities, i.e. jumps, of the process X. It tells us when the jumps occur and how big they are. Therefore, J_X does not tell us anything about the continuous component of X, which has continuous sample path if and only if $J_X = 0$ almost surely. This means that there are no jumps in the process.
For a counting process, since the jumps size is always equal to 1, the jump measure can be seen as a random measure on $[0, \infty)$.
Proposition 1.2 *Let X be a Poisson process with intensity λ. Then, J_X is a Poisson random measure on $[0, \infty)$ with intensity $\lambda \times dt$.*
Proposition 1.3 (Jump measure of a compound Poisson process) (proposition 3.5 in [2]) *Let $(X_t)_{t \geq 0}$ be a compound Poisson process with intensity λ and jump size distribution f. Its jump measure J_X is a Poisson random measure on $\mathbb{R}^d \times [0, \infty)$ with intensity measure $\mu(dx \times dt) = \nu(dx)dt = \lambda f(dx)dt$.*
This proposition implies that every compound Poisson process can be represented in the following form:

$$X_t = \sum_{s \in [0,t]} \Delta X_s = \int_{[0,t] \times \mathbb{R}^d} x J_X(ds \times dx)$$

where J_X is a Poisson random measure with intensity measure $\nu(dx)dt$. In this equation, we have rewritten the process X as the sum of its jumps and since it is a compound Poisson process, it has almost surely a finite number of jumps in the interval $[0, t]$. Moreover, the stochastic integral in the equation is a finite sum, hence there are no convergence problems.

1.2 Lévy Processes

1.2.1 Definition and Properties

Definition 1.5 [Lévy process] (definition 3.1 in [2]) *A càdlàg stochastic process $(X_t)_{t \geq 0}$ on $(\Omega, \mathcal{F}, \mathbb{P})$ with values in $\mathbb{R}^d$ such that $X_0 = 0$ is called a Lèvy process if it possesses the following properties:*

1) *Independent increments: for every increasing sequence of times $t_0, \ldots, t_n$, the random variables $X_{t_0}, X_{t_1} - X_{t_0}, \ldots, X_{t_n} - X_{t_{n-1}}$ are independent;*
2) *Stationary increments: the law of $X_{t+h} - X_t$ does not depend on t;*
3) *Stochastic continuity: $\forall \varepsilon > 0, \quad \lim_{h \to 0} \mathbb{P}(|X_{t+h} - X_t| \geq \varepsilon) = 0$.*

Proposition 1.4 (Characteristic function of a Lévy process) (proposition 3.2 in [2]) *Let $(X_t)_{t\geq 0}$ be a Lévy process on $\mathbb{R}^d$. There exists a continuous function $\psi : \mathbb{R}^d \mapsto \mathbb{R}$ called the characteristic exponent of X, such that:*

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

Where ψ is the cumulant generating function of X_1. The cumulant generating function $\psi(t)$ is the natural logarithm of the moment generating function:

$$\psi(t) = \log E[e^{tX}]$$

The proposition regarding the Jump measure of a compound Poisson process can be used to define the Lévy measure for all the Lévy process. Therefore, we give the following definition:
Definition 1.6 [Lévy measure] (definition 3.4 in [2]) *Let $(X_t)_{t\geq 0}$ be a Lévy process on $\mathbb{R}^d$. The measure ν on $\mathbb{R}^d$ defined by:*

$$\nu(A) = E[\#\{t \in [0,1] : \Delta X_t \neq 0,\ \Delta X_t \in A\}], \quad A \in \mathcal{B}(\mathbb{R}^d)$$

is called the Lévy measure of X. $\nu(A)$ is the expected number, per unit time, of jumps whose size belongs to A.
Proposition 1.5 (Lévy-Ito decomposition) (proposition 3.7 in [2]) *Let $(X_t)_{t\geq 0}$ be a Lèvy process on $\mathbb{R}^d$ and ν its Lèvy measure. Then:*

- *the Lévy measure ν satisfies the integrability condition:*

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

- *the jump measure J_X of X is a Poisson random measure on $[0,\infty) \times \mathbb{R}^d$ with intensity $dt \times \nu = \nu(dx)dt$;*

- *there exists $\gamma \in \mathbb{R}^d$ and a d-dimensional Brownian motion $(B_t)_{t\geq 0}$ with covariance matrix A such that:*

$$\begin{aligned} X_t &= \gamma t + B_t + N_t + M_t, \quad \textit{where} \qquad (3)\\ N_t &= \int_{|x|>1, s\in[0,t]} xJ_X(ds \times dx)\\ M_t &= \int_{0<|x|\leq 1, s\in[0,t]} x\{J_X(ds \times dx) - \nu(dx)ds\}\\ &\equiv \int_{0<|x|\leq 1, s\in[0,t]} x\tilde{J}_X(ds \times dx) \end{aligned}$$

The three terms in (1.3) *are independent and the convergence in the last term is almost sure and uniform in t on $[0,T]$.*
The Lévy-Ito decomposition say that for every Lévy process there exist a vector γ (drift), a positive definite matrix A and a positive measure ν that uniquely determine its distribution. We call the triplet (A,ν,γ) characteristic triplet or Lévy triplet of the process X_t.
The following theorem give to us the second fundamental result of the structure of the path of Poisson process and it announces the expression of the characteristic function of a Lévy process in terms of its characteristic triplet (A,ν,γ) :
Theorem 1.1 [Lévy-Khinchin representation] (theorem3.1 in [2]) *Let $(X_t)_{t\geq 0}$ be a Lévy process on $\mathbb{R}^d$ with characteristic triplet (A,ν,γ). Then:*

$$\begin{aligned} &E[e^{i\langle u, X_t\rangle}] = e^{t\psi(u)}, \quad u \in \mathbb{R}^d \qquad (4)\\ &\textit{with} \quad \psi(u) = i\langle \gamma, u\rangle - \frac{1}{2}\langle u, Au\rangle + \int_{\mathbb{R}^d} (e^{i\langle u,x\rangle} - 1 - i\langle u,x\rangle \mathbf{1}_{|x|\leq 1})\nu(dx). \end{aligned}$$

This theorem imply that since X has stationary and independent increments, we have that $E\left[e^{i\langle u,X_t\rangle}\right] = \left\{E\left[e^{i\langle u,X_1\rangle}\right]\right\}^t$, $\forall t \in \mathbb{R}$ and by the right continuity of X, $\forall t$.

Definition 1.7 [Semimartingale] *A semimartingale is a stochastic process* $(X_t)_{0\leq t\leq T}$ *which admits the decomposition:*

$$X = X_0 + M + A \tag{5}$$

where X_0 *is finite and* $\mathcal{F}_0$*-measurable,* M *is a local martingale with* $M_0 = 0$ *and* A *is a finite variation process with* $A_0 = 0$.
If A is predictable, then X is a special semimartingale and all special semimartingale have a "canonical decomposition" equal to:

$$X = X_0 + B + X^c + x(J_X - \nu^X) \tag{6}$$

where X^c is the continuous martingale part of X and $x(J_X - \nu^X)$ is the purely discontinuous martingale part of X. In particular, J_X is the jump measure of X and ν^X is called the compensator of J_X.

1.3 Jump Diffusion Model

The financial models with jumps can be decomposed in two main categories: the jump-diffusion model and the infinite activity models. The jump-diffusion model the evolution of prices are given by a diffusion process which has jumps at random intervals. Here, the jumps represent rare events such as crashes and large drawdown. Since the distribution of jump sizes is known, the dynamic structure of the jump process is easy to understand and describe. The jump-diffusion models perform well for implied volatility smile interpolation.

1.3.1 Exponential Lévy Models

In order to construct an exponential Lévy model for the process X, we need to start from the Black-Scholes model and how it describes the evolution of an asset price. Here, the asset price (S_t) follow a geometric Brownian motion:

$$S_t = S_0 e^{\mu t + \sigma W_t} \tag{7}$$

If we replace $\mu t + \sigma W_t$ by a Lévy process X_t, we obtain the class of the exponential Lévy models:

$$S_t = S_0 e^{X_t} \tag{8}$$

Now, consider a Lévy process of jump-diffusion type with the following form:

$$X_t = \gamma t + \sigma W_t + \sum_{i=1}^{N_t} Y_1 \tag{9}$$

where $(N_t)_{t\geq 0}$ is the Poisson process which counting the jumps of X and Y_i are the jump sizes, which are i.i.d. variables. Therefore, the evolution of the asset price becomes:

$$S_t = S_0 e^{\gamma t + \sigma W_t + \sum_{i=1}^{N_t} Y_1} \tag{10}$$

We need to specify the distribution of jump sizes $\nu_0(x)$ in order to define the parametric model completely. Is is important to specify the tail behavior of $\nu_0(x)$ correctly because the tail behavior of the jump measure determines the tail behavior of the probability density of the process.

In the *Merton model* (introduced by Merton in 1973 with the article *"Option pricing when underlying stock return are discontinuous"*) we have that the process is equal to the equation (1.10) and the jumps are assumed to have a Gaussian distribution, therefore $Y_i \sim N(\mu, \delta^2)$. This allows to obtain the probability density of X_t as a quickly converging series. In fact,

$$\mathbb{P}(X_t \in A) = \sum_{k=0}^{\infty} \mathbb{P}(X_t \in A | N_t = k)\mathbb{P}(N_t = k)$$

then the probability density of X_t satisfies the equation:

$$p_t(x) = e^{-\lambda t} \sum_{k=0}^{\infty} \frac{(\lambda t)^k exp\left\{-\frac{(x-\gamma t-k\mu)^2}{2(\sigma^2 t+k\delta^2)}\right\}}{k!\sqrt{2\pi(\sigma^2 t + k\delta^2)}}$$

The Lévy density of the model is equal to:

$$\nu(x) = \frac{\lambda}{\delta\sqrt{2\pi}} exp\left\{-\frac{(x-\mu)^2}{2\delta^2}\right\}$$

One last thing to note is the moment of the process in the Merton model. Hence, we have that the characteristic exponent of the characteristic function is equal to:

$$\psi(u) = i\gamma u - \frac{\sigma^2 u^2}{2} + \lambda\left\{e^{-\frac{\delta^2 u^2}{2}+i\mu u} - 1\right\}$$

It follows that: $E[X_t] = t(\gamma + \lambda\mu)$ and $Var(X_t) = t(\sigma^2 + \lambda\gamma^2 + \lambda\mu^2)$. If we analyze the moment, we can note the tail behavior of the probability density, which are heavier than Gaussian but all the exponential moments are finite.
In the *Kou model* (introduced by Kou in 2002 with the article *"A jump-diffusion model for option pricing"*) we have that the process X_t is equal as in the Merton model but the distribution of jumps sizes is an asymmetric exponential (i.e. has a double exponential distribution, therefore $Y_i \sim DbExp(p, \theta_1, \theta_2)$) with a density of the form:

$$\nu_0(dx) = \left[p\theta_1 e^{-\theta_1 x}\mathbf{1}_{x>0} + (1-p)\theta_2 e^{-\theta_2|x|}\mathbf{1}_{x<0}\right] dx \tag{11}$$

where $\theta_1 > 0$, $\theta_2 > 0$ represent the decay of the tails for the distribution of positive and negative jump sizes, respectively, and $p \in [0, 1]$ represent the probability of an upward jump. Therefore, we can easily find the Lévy measure of the process:

$$\nu(x) = p\lambda\theta_1 e^{-\theta_1 x}\mathbf{1}_{x>0} + (1-p)\lambda\theta_2 e^{-\theta_2|x|}\mathbf{1}_{x<0}$$

The first two moments of the process are equal to: $E[X_t] = t\left(\gamma + \frac{\lambda t}{\theta_1} - \frac{\lambda(1-p)}{\theta_2}\right)$ and $Var(X_t) = t\left(\sigma^2 + \frac{\lambda t}{\theta_1^2} - \frac{\lambda(1-p)}{\theta_2^2}\right)$. We find these two result from the characteristic function of the process, which has characteristic exponent equal to:

$$\psi(u) = i\gamma u - \frac{\sigma^2 u^2}{2} + iu\lambda\left\{\frac{p}{\theta_1 - iu} - \frac{1-p}{\theta_2 + iu}\right\}$$

In this case, the probability distribution of returns has semi-heavy exponential tails. On one hand, we have that $p(x) \sim e^{-\theta_1 x}$ when $x \to +\infty$, on the other hand, we have that $p(x) \sim e^{-\theta_2|x|}$ when $x \to -\infty$.
The advantage of the Kou model compared to the Merton model is that analytical expressions for expectations involving first passage times may be obtained due to the memoryless property of exponential random variables.
The following proposition told us the condition of the exponential Lévy process to be a semimartingale:
Proposition 1.6 (Exponential Lévy process) (proposition 8.20 in [2]) *Let $(X_t)_{t\geq 0}$ be a Lévy process with Lévy triplet (σ^2, ν, γ) verifying $\int_{|y|\geq 1} e^y \nu(dy) < \infty$. Then, $Y_t = e^{X_t}$ is a semimartingale with decomposition $Y_t = M_t + A_t$ where the martingale part is given by:*

$$M_t = 1 + \int_0^t Y_{s^-}\sigma dW_s + \int_{[0,t]\times\mathbb{R}} Y_{s^-}\left(e^z - 1\right)\tilde{J}_X(ds \times dz)$$

and the continuous finite variation drift part is given by:

$$A_t = \int_0^t Y_{s^-}\left[\gamma + \frac{\sigma^2}{2} + \int_{-\infty}^{\infty}\left(e^z - 1 - z\mathbf{1}_{|z|\geq 1}\right)\nu(dz)\right] ds.$$

(Y_t) *is a martingale if and only if*

$$\gamma + \frac{\sigma^2}{2} + \int_{-\infty}^{\infty} \left(e^z - 1 - z\mathbf{1}_{|z|\geq 1}\right) \nu(dz) = 0$$

1.3.2 Stochastic exponential of Jump process

The stochastic exponential was introduced by Doléans-Dade and it can be found using the Ito formula in the geometric Brownian motion (equation (1.7)) and substituting a Lévy process. Hence, if we apply the Ito formula in (1.7) we obtain:

$$\frac{dS_t}{S_t} = \left(\mu + \frac{\sigma^2}{2}\right) dt + \sigma dW_t$$

Then, we can define $B_t^1 = (\mu + \frac{\sigma^2}{2})t + \sigma W_t$ and the above equation becomes:

$$\frac{dS_t}{S_t} = dB_t^1 \tag{12}$$

If we substitute B_t^1 by a Lévy process X, we obtain the stochastic exponential. Therefore, with the following proposition we can introduce a generic stochastic exponential for a process $(Z_t)_{t\geq 0}$.
Proposition 1.7 (Stochastic exponential) (proposition 8.21 in [1]) *Let $(X_t)_{t\geq 0}$ be a Lévy process with Lévy triplet (σ^2, ν, γ). Then, there exists a unique càdlàg process $(Z_t)_{t\geq 0}$ such that:*

$$\begin{cases} dZ_t = Z_{t-} dX_t \\ Z_0 = 1 \end{cases} \tag{13}$$

Where Z is given by:

$$Z_t = e^{X_t - \frac{1}{2}\int_0^t \sigma_s^2 ds} \prod_{0\leq s\leq t} (1 + \Delta X_s) e^{-\Delta X_s} \tag{14}$$

If $\int_{-1}^{1} |x|\nu(dx) < \infty$, then the jumps of X have finite variation and the stochastic exponential of X can be expressed as:

$$Z_t = e^{\sigma W_t + \gamma_0 t - \frac{\sigma^2 t}{2}} \prod_{0\leq s\leq t} (1 + \Delta X_s)$$

where $\gamma_0 = \gamma - \int_{-1}^{1} x\nu(dx)$.
Z is called the stochastic exponential of X and is denoted by $Z = \mathcal{E}(X)$.

We can note that the stochastic exponential is always nonnegative if all the jumps of X_t are greater than -1, i.e. $\nu((-\infty, -1]) = 0$.
Goll and Kallsen have shown that the stochastic exponential is equivalent to the ordinary exponential. In fact, if $Z > 0$ is the stochastic exponential of a Lévy process, it is also the ordinary exponential of another Lévy process (it is also true the opposite case). Therefore, the two exponential end up by giving us the same class of positive processes. The following proposition shows the relation between ordinary and stochastic exponential:
Proposition 1.8 (Relation between ordinary and stochastic exponential) (proposition 8.22 in [2])

1. *Let $(X_t)_{t\geq 0}$ be a real valued Lévy process with Lévy triplet (σ^2, ν, γ) and $Z = \mathcal{E}(X)$ its stochastic exponential. If $Z > 0$ almost surely, then there exists another Lévy process*

$(L_t)_{t\geq 0}$ with triplet $(\sigma_L^2, \nu_L, \gamma_L)$ such that $Z_t = e^{L_t}$ where:

$$L_t = \ln Z_t = X_t - \frac{\sigma^2 t}{2} + \sum_{0\geq s\geq t} (\ln(1+\Delta X_s) - \Delta X_s$$
$$\sigma_L = \sigma$$
$$\nu_L(A) = \nu(\{x : \ln(1+x) \in A\}) = \int \mathbf{1}_A(\ln(1+x))\nu(dx)$$
$$\gamma_L = \gamma - \frac{\sigma^2}{2} + \int [\ln(1+x)\mathbf{1}_{[-1,1]}(\ln(1+x)) - x\mathbf{1}_{[-1,1]}(x)]\nu(dx)$$

2. *Let $(L_t)_{t\geq 0}$ be a real valued Lévy process with Lévy triplet $(\sigma_L^2, \nu_L, \gamma_L)$ and $S_t = e^{L_t}$ its exponential. Then, there exists a Lévy process $(X_t)_{t\geq 0}$ such that S_t is the stochastic exponential of X : $S = \mathcal{E}(X)$ where:*

$$X_t = L_t + \frac{\sigma^2 t}{2} + \sum_{0\leq s\leq t} \left(e^{\Delta L_s} - 1 - \Delta L_s\right)$$

Therefore, the Lévy triplet (σ^2, ν, γ) of X is given by:

$$\sigma = \sigma_L$$
$$\nu(A) = \nu_L(\{x : (e^x - 1) \in A\}) = \int \mathbf{1}_A(e^x - 1)\nu_L(dx)$$
$$\gamma = \gamma_L - \frac{\sigma_L^2}{2} + \int [(e^x - 1)\mathbf{1}_{[-1,1]}(e^x - 1) - x\mathbf{1}_{[-1,1]}(x)]\nu_L(dx)$$

3. *Let $(X_t)_{t\geq 0}$ be a Lévy process and a martingale. Then, its stochastic exponential $Z = \mathcal{E}(X)$ is also a martingale. Therefore, for every Lévy process X with $E[|X_t|] < \infty$ we have:*

$$E[\mathcal{E}(X_t)] = e^{E[X_t]} \quad t > 0$$

This property is also known as Martingale preserving property.

2 Stochastic Calculus for Jump Process

We can define an arbitrage strategy as a self-financing strategy ϕ with zero initial value and non-zero final value with probability equal to 1. Moreover, a strategy is called self-financing if the following equation is satisfied for all t : $\langle \phi_t, S_t \rangle = \langle \phi_{t+1}, S_t \rangle$. Therefore, we can consider an investor who trades at times $T_0 = 0 < T_1 < \cdots < T_n < T_{n+1} = T$ and detaining a quantity ϕ_i of an asset whose price is S during the period $(T_i, T_{i+1}]$. Then, we can definite the capital gain $G_t(\phi)$ as:

$$G_t(\phi) := \sum_{i=0}^{n} \phi_i(S_{T_{i+1}} - S_{T_i}) \tag{15}$$

We can write the quantity which represents the capital gain of the investors following the strategy ϕ as :

$$G_t(\phi) = \sum_{i=0}^{n} \phi_i(S_{T_{i+1}} - S_{T_i}) = \int_0^T \phi_t dS_t \tag{16}$$

where the last term in equation (3.2) represent the stochastic integral ϕ with respect to S.

2.1 Stochastic integral

Consider a vector of asset whose price S is described by a stochastic process, i.e. $S_t = (S_t^1, S_t^2, \dots S_t^d)$ and a portfolio $\phi = (\phi^1, \phi^2, \dots, \phi^d)$ which describes the amount of each asset held by the investor. Therefore, the value of such portfolio at time t is equal to:

$$V_t(\phi) = \sum_{k=1}^{d} \phi^d S_t^k \equiv \langle \phi_t, S_t \rangle \tag{17}$$

We also assume a dynamic trading strategy, which consist in buying and selling assets at different dates, and we consider an investor who trades at times $T_0 = 0 < T_1 < \cdots < T_n < T_{n+1} = T$. We also assume that the strategy is self-financing and that between two transaction dates T_i and T_{i+1} the portfolio remains unchanged. The meaning of the self-financing assumption is that at time t the investors readjusts his position from ϕ_t to ϕ_{t+1} without bringing or consuming any wealth. Moreover, if we dropped this assumption, we would had arbitrage opportunities because a portfolio which is empty at time 0 but to which cash (> 0) is added, without any liability, would trivially be an arbitrage portfolio. The second assumption told us that the investor did not know in advance the transaction dates but he will decide to buy or sell at T_{i+1} depending on the information revealed before T_{i+1}. Hence, the transaction date T_{i+1} is a stopping time. In the first chapter, we assume that the processes are càdlàg (i.e. right continuous with left limits), whereas here we have that the trading strategy is càglàd (i.e. left continuous with right limits). We have the left continuity in the process because if the investor decides to make a transaction at $t = T_i$, the portfolio will take the new value at ϕ_i before that the value of the portfolio is still described by ϕ_{i-1}. Therefore, we have that $(\phi_t)_{t\in[0,T]}$ is a predictable process and we have the following definition:
Definition 2.1 [Simple Predictable Process] (definition 8.1 in [2]) *A stochastic process $(\phi_t)_{t\in[0,T]}$ is called a simple predictable process if it can be represented as:*

$$\phi_t = \phi_0 \mathbf{1}_{t=0} + \sum_{i=0}^{n} \phi_i \mathbf{1}_{(T_i,T_{i+1}]}(t) \tag{18}$$

where $T_0 = 0 < T_1 < \cdots < T_n < T_{n+1} = T$ are nonanticipating random times and each ϕ_i is bounded random variable whose value is revealed at T_i (i.e. $\mathcal{F}_{T_i}$-measurable).
The stochastic process $G_t(\phi)$ can be expressed as the stochastic integral of the simple predictable process ϕ with respect to S and it is equal to:

$$\int_0^t \phi_u dS_u := \langle \phi_0, S_0 \rangle + \sum_{i=0}^{n} \langle \phi_i, (S_{T_{i+1}\wedge t} - S_{T_i \wedge t}) \rangle \tag{19}$$

Since the self-financing assumption imply that the cost of the process is equal to zero, we have that the value of the portfolio, $V_t(\phi)$, is equal to:

$$V_t(\phi) = \int_0^t \phi_u dS_u = \phi_0 S_0 + \int_{0^+}^t \phi_u dS_u$$

where the first term is the initial value of the portfolio and the second term is the capital gain between 0 and t. Therefore, for an investors the only source of variation of the portfolio's value is the variation of the asset values.

Proposition 2.1 (Martingale preserving property) (proposition 8.1 in [2]) *If $(S_t)_{t\in[0,T]}$ is a martingale, then for any predictable process ϕ the stochastic integral $G_t = \int_0^t \phi dS$ is also a martingale.*
This proposition imply that if the asset follows a martingale then the the value of any self-financing strategy is a martingale.

2.1.1 Semimartingale

Since a Lévy process X is not stable under stochastic integration or non-linear transformations, we need to consider the class of samimartingales, which are a larger class of stochastic processes. These kind of class are both stable under stochastic integration and non-linear transformation. Moreover, they are also stable under other operation such as change of filtration and change of measure. Now, we give the definition of semimartingale with respect a simple predictable process.

Definition 2.2 [Semimartingale] (definition 8.2 in [2]) *A nonanticipating càdlàg process S is called a semimartingale if the stochastic integral of simple predictable process with respect to S:*

$$\phi = \phi_0 \mathbf{1}_{t=0} + \sum_{i=0}^{n} \phi_i \mathbf{1}_{(T_i, T_{i+1}]} \mapsto \int_0^T \phi dS = \phi_0 S_0 + \sum_{i=0}^{n} \phi_i (S_{T_{i+1}} - S_{T_i})$$

verifies the following continuity property: for every ϕ^n, $\phi \in \mathbb{S}([0,T])$ if:

$$\sup_{(t,\omega)\in[0,T]\times\Omega} |\phi_t^n(\omega) - \phi_t(\omega)| \underset{n\to\infty}{\to} 0$$

then

$$\int_0^T \phi^n dS \xrightarrow[n\to\infty]{\mathbb{P}} \int_0^T \phi dS \tag{20}$$

where $\mathbb{S}([0,T])$ is a set of simple predictable processes on $[0,T]$.

The class of semimartingales satisfy the stability property: a small change in the portfolio should lead to a small change in the gain process.

Proposition 2.2 (proposition 8.3 in [2]) *If $(S_t)_{t\in[0,T]}$ is a semimartingale then for every ϕ^n, $\phi \in \mathcal{S}([0,T])$:*

$$if \quad \sup_{(t,\omega)\in[0,T]\times\Omega} |\phi_t^n(\omega) - \phi_t(\omega)| \underset{n\to\infty}{\to} 0 \tag{21}$$

$$then \quad \sup_{t\in[0,T]} \left| \int_0^t \phi^n dS - \int_0^t \phi dS \right| \xrightarrow[n\to\infty]{\mathbb{P}} 0. \tag{22}$$

Moreover, we have that any linear combination of a finite number of semimartingales is a semimartingales. In fact, all the Lévy processes are semimartingale because it can be decomposed into a sum of square integrable martingale (the Wiener process) and a finite variation process (the Poisson process). Finally, we can note that all the new processes constructed from semimartingales using stochastic integration are again semimartingales due to associativity property, which helps us to show that a stochastic integral with respect to a semimartingale is a semimartingale. And that every semimartingale is the sum of a finite variation process and a local martingale, which can be defined as the process (X_t) in which there exists a sequence of stopping times $(\tau_i)_{i\geq 1}$ such that $\tau_i \to \infty$ when $i \to \infty$ and for each i, $(X_{\tau_i \wedge t})$ is a martingale.

2.1.2 Stochastic integral with respect to Brownian motion

Consider the simple predictable process ϕ defines in equation (2.4). Then, we can define the Brownian stochastic integral as:

$$\int_0^T \phi_t dW_t = \sum_{=0}^{n} \phi_i (W_{T_{i+1}} - W_{T_i}) \tag{23}$$

Proposition 2.3 (Isometry formula) (proposition 8.5 in [2]) *Let $(\phi_t)_{0\leq t\leq T}$ be a simple predictable process and $(W_t)_{0\leq t\leq T}$ be a Wiener process. Then:*

i.

$$E\left[\int_0^T \phi_t dW_t\right] = 0, \tag{24}$$

ii.

$$E\left[\left|\int_0^T \phi_t dW_t\right|^2\right] = E\left[\int_0^T |\phi_t|^2 dt\right] \tag{25}$$

We can use the isometry formula to built stochastic integrals with respect to the Wiener process for predictable processes. We need that the predictable processes $(\phi_t)_{t\in[0,T]}$ verify:

$$E\left[\int_0^T |\phi_t|^2 dt\right] < \infty$$

$$E\left[\int_0^T |\phi_t^n - \phi_t|^2 dt\right] \underset{n\to\infty}{\longrightarrow} 0.$$

Therefore, we have the following proposition for Brownian integrals:

Proposition 2.4 (Isometry formula for Brownian integrals) (proposition 8.6 in [2]) *Let $(\phi_t)_{0\le t\le T}$ be a predictable process which satisfy:*

$$E\left[\int_0^T |\phi_t|^2 dt\right] < \infty$$

Then, $\int_0^t \phi dW_t$ is a square integrable martingale and

i.

$$E\left[\int_0^T \phi_t dW_t\right] = 0,$$

ii.

$$E\left[\left|\int_0^T \phi_t dW_t\right|^2\right] = E\left[\int_0^T |\phi_t|^2 dt\right]$$

2.1.3 Quadratic variation and covariation

Consider a process observed on a time grid $\pi = (t_0 = 0 < t_1 < \cdots < t_{n+1} = T)$, then we can define the realized variance as:

$$V_X(\pi) = \sum_{t_i\in\pi} (X_{t_{i+1}} - X_{t_i})^2$$

We can define the quadratic of the variation process:

Definition 2.3 [Quadratic Variation] (definition 8.3 in [2]) *The quadratic variation process of a semimartingale X is the nonanticipating càdlàg process defined by:*

$$[X,X]_t = |X_t|^2 - 2\int_0^t X_{u^-} dX_u \tag{26}$$

Is important to specify that the quadratic variation is a random variable and not a number. Moreover, if $\pi^n = (t_0^n = 0 < t_1^n < \cdots < t_{n+1}^n = T)$ is a sequence of partitions of $[0,T]$ such that $|\pi^n| = \sup_k |t_k^n - t_{k-1}^n| \to 0$ as $n\to\infty$, then

$$\sum_{t_i\pi^n}^{0\le t_i<t} (X_{t_{i+1}} - X_{t_i})^2 \xrightarrow[n\to\infty]{\mathbb{P}} [X,X]_t$$

where the convergence is uniform in t.

Consider a Brownian motion $B_t = \sigma W_t$, where W is a standard Wiener process, then the quadratic variation of the Brownian motion is equal to $[B,B]_t = \sigma^2 t$.

In the quadratic variation, we consider only one process X but, in the reality, we can see more stochastic process. Therefore, we need to introduce the multidimensional counterpart of the quadratic variation: the quadratic covariation.

Definition 2.4 [Quadratic Covariation] (definition 8.4 in [2]) *Given two semimartingales X,Y. The quadratic covariation process $[X,Y]$ is the semimartingale defined by:*

$$[X,Y]_t = X_tY_t - X_0Y_0 - \int_0^t X_{s^-} dY_s - \int_0^t Y_{s^-} dX_s \tag{27}$$

Consider two Brownian motion: $B_t^1 = \sigma_1 W_t^1$ and $B_t^2 = \sigma_2 W_t^2$, where W^1, W^2 are two standard wiener processes with correlation ρ (typically, with differential notation we define the correlation between two standard Wiener process as: $dW^1 dW^2 = \rho dt$.

2.2 Stochastic Integral with Jumps

2.2.1 Stochastic Integral with respect to Poisson process

Consider the relation, for the Poisson process: $\Delta X_t = Y_{N_t} \Delta N_t$. Then, we can define the stochastic integral of a stochastic process $(\phi_t)_{t\geq 0}$ with respect to $(X_t)_{t\geq 0}$ by:

$$\int_0^T \phi_t dX_t = \int_0^T \phi_t Y_{N_t} dN_t := \sum_{k=1}^{N_T} \phi_{T_k} Y_k \tag{28}$$

The meaning of the above equation is that the value at time T of a portfolio containing a quantity ϕ_t of an asset at time t, whose price evolves according to random returns Y_k, generate capital gain or losses equal to $\phi_{T_k} Y_k$ at random times T_k.
Consider a compound Poisson process $(X_t)_{t\geq 0}$, it admits stochastic integral representation equal to:

$$X_t = X_0 + \sum_{k=1}^{N_t} Y_k = Y_0 + \int_0^t Y_{N_s} dN_s$$

Consider a counting process N_t with jump times T_i and with random variables observed at T_i describe by Y_i. Let X_t be a process defined by $X_t = \sum_{i=1}^{N_t} Y_i$, hence the quadratic variation of the process is equal to:

$$[X,X]_t = \sum_{s\leq t} (\Delta X_s)^2 = \sum_{i=1}^{N_t} Y_i^2$$

We can note that the same formula holds for every finite variation process X. Moreover, the predictable quadratic variation of the process (i.e. "angle bracket") is the compensator of $[X,X]$, namely:

$$\langle X, X\rangle_t = \lambda t E[Y_1^2]$$

For the quadratic covariation we need to consider another counting process N_t^b, which has jump times T_j and random variables observed at T_j, described by Y_j^b. Then, we consider the process $Z_t = \sum_{j=1}^{N_t^b} Y_j^b$. Now, we make the assumption that X and Z have finite variation processes whose jumps times are almost surely disjoint, hence they did not jump at the same time, therefore the quadratic covariation is equal to:

$$[X,Z]_t = \sum_{s\leq t} \Delta X_s \Delta Z_s = 0$$

The assumption of disjoint jumps is a strong assumption and we consider it only for the stock price behavior. In fact, if we consider the exchange rate we drop this assumption and we consider correlated jumps between the rate.

2.2.2 Stochastic Integral with respect to Poisson random measure

Consider a Poisson random measure M on $[0,T]\times\mathbb{R}^d$ with intensity $\mu(dt\times dx)$. Let $\tilde{M}$ be the compensated random measure defined as the centered version of M: $\tilde{M}(A) = M(A) - \mu(A) = M(A) - E[M(A)]$, where $A \subset \mathbb{R}^d$.
We can define the simple predictable process with respect to the Poisson random measure as:

$$\phi(t,y) = \sum_{i=1}^{n}\sum_{j=1}^{m} \phi_{ij} \mathbf{1}_{(T_i,T_{i+1}]}(t)\mathbf{1}_{A_j}(y) \tag{29}$$

where $\phi : \Omega \times [0,T] \times \mathbb{R}^d \to \mathbb{R}$ is a simple predictable functions, $(\phi_{ij})_{j=1,\dots,m}$ are bounded $\mathcal{F}_{T_i}$-measurable random variables, $T_1 \leq T_2 \leq \cdots \leq T_n$ are non anticipating random times and $(A_j)_{j=1,\dots,m}$ are disjoint subsets of $\mathbb{R}^d$ with $\mu([0,T] \times A_j) < \infty$. The disjoint subset implies that the compensated random measure is a martingale with respect to A_j and that if $A \cap B = \emptyset$, then $M_t(A)$ and $M_t(B)$ are independent.

Proposition 2.5 (Martingale preserving property) (proposition 8.7 in [2]) *For any simple predictable function $\phi : \Omega \times [0,T] \times \mathbb{R}^d \to \mathbb{R}$ the process $(X_t)_{t\in[0,T]}$ defined by the compensated integral:*

$$X_t = \int_0^t \int_{\mathbb{R}^d} \phi(s,y)\tilde{M}(ds \times dy)$$

is a square integrable martingale and verifies the isometry formula:

$$E[|X_t|^2] = E\left[\int_0^t \int_{\mathbb{R}^d} |\phi(s,y)|^2 \mu(ds \times dy)\right] \tag{30}$$

We can extend the isometry formula to square integrable predictable functions and we have the following proposition:

Proposition 2.6 (Compensated Poisson integrals) (proposition 8.8 in [2]) *For any predictable random function $\phi : \Omega \times [0,T] \times \mathbb{R}^d \to \mathbb{R}$ verifying*

$$E\left[\int_0^t \int_{\mathbb{R}^d} |\phi(s,y)|^2 \mu(ds \times dy)\right] < \infty$$

the following property hold:

- $t \mapsto \int_0^t \int_{\mathbb{R}^d} \phi(s,y)\tilde{M}(ds \times dy)$ *is a square integrable martingale;*
- $$E\left[\left| \int_0^t \int_{\mathbb{R}^d} \phi(s,y)\mu(ds \times dy) \right|^2\right] = E\left[\int_0^t \int_{\mathbb{R}^d} |\phi(s,y)|^2 \mu(ds \times dy)\right] \tag{31}$$

Consider a Lévy process $(X_t)_{t\geq 0}$ with Lévy measure ν and a Poisson random measure J_X with intensity $\mu(dt \times dx) = dt\nu(dx)$.

2.3 Change of variable formula

Let $f : \mathbb{R} \to \mathbb{R}$ be a C^2 function and let $g : [0,T] \to \mathbb{R}$ be a C^1 function. Then, the change of variables formula for smooth function is:

$$f(g(t)) - f(g(0)) = \int_0^t f'(g(s))g'(s)ds = \int_0^t f'(g(s))dg(s) \tag{32}$$

Now, we can consider the Brownian integral defined as: $X_t = \int_0^t \sigma_s dW_s$ and the function f defined as above. Then, if we apply the Ito formula at X_t we find:

$$f(X_t) = f(0) + \int_0^t f'(X_s)\sigma_s dW_s + \frac{1}{2}\int_0^t f''(X_s)ds \tag{33}$$

2.3.1 Calculus for finite jump processes

Let $x : [0,T] \to \mathbb{R}$ be a function with a finite number of discontinuities at time $0 = T_0 \leq T_1 \leq T_2 \leq \cdots \leq T_n \leq T_{n+1} = T$ and the function x is smooth on each interval, defines as (T_i, T_{i+1}). Moreover, we can define $x(T_i) := x(T_i^+)$, which means that x is càdlàg at the discontinuity points. Let $f : \mathbb{R} \to \mathbb{R}$ be a C^1 function. Since x is smooth on each interval (T_i, T_{i+1}), $f(x(t))$ is also smooth. Then, the following proposition summarized the Ito formula for finite activity jump process where the counting process N_t is a martingale:

Proposition 2.7 (Ito formula for finite activity jump processes) (proposition 8.13 in [2]) *Let X be a jump process with values in $\mathbb{R}$ defined by:*

$$X_t = \int_0^t b_s ds + \sum_{i=1}^{N_t} Y_i$$

where b_s is a nonanticipating càdlàg process, N_t is a counting process representing the number of jumps between 0 and t and Y_i is the size of the i-th jump. Denote by $(T_n)_{n\geq 1}$ the jump times of X_t and J_X the random measure on $[0,T]\times\mathbb{R}$ associated to the jumps of X[1]*. Then, for any measurable function $f:[0,T]\times\mathbb{R}\to\mathbb{R}$ we have:*

$$\begin{aligned}
f(t,X_t) - f(0,X_0) &= \int_0^t \left(\frac{\partial f}{\partial s}(s,X_{s^-}) + b_s\frac{\partial f}{\partial x}(s,X_{s^-})ds\right) \\
&\quad + \sum_{n\geq 1, T_n\leq T} \left(f(s,X_{s^-}+\Delta X_s) - f(s,X_{s^-})\right) \\
&= \int_0^t \left(\frac{\partial f}{\partial s}(s,X_{s^-}) + b_s\frac{\partial f}{\partial x}(s,X_{s^-})ds\right) \\
&\quad + \int_0^t\int_{-\infty}^{\infty} \left(f(s,X_{s^-}+y) - f(s,X_{s^-})\right) J_X(ds\times dy) \qquad (34)
\end{aligned}$$

where $\Delta x_i = x(T_i) - x(T_i^-)$. Moreover, if N_t is a Poisson process with $E[N_t]=\lambda t$, with $Y_i\sim F$ are i.i.d. and f is boundend, then $Y_t = f(t,Y_t) = V_t + M_t$, where M is the martingale or noise component and V is the continuous finite variation drift. This two component are respectively equal to:

$$\begin{aligned}
M_t &= \int_0^t\int_{-\infty}^{\infty} \left(f(s,X_{s^-}+y) - f(s,X_{s^-})\right) \tilde{J}_X(ds\times dy) \qquad (35) \\
&\quad \textit{where } \tilde{J}_X(ds\times dy) = J_X(dt\times dy) - \lambda F(dy)dt \\
V_t &= \int_0^t \left(\frac{\partial f}{\partial s}(s,X_{s^-}) + b_s\frac{\partial f}{\partial x}(s,X_{s^-})ds\right) \\
&\quad + \int_0^t ds\int_{\mathbb{R}^d} F(dy)\left(f(s,X_{s^-}+y) - f(s,X_{s^-})\right) \qquad (36)
\end{aligned}$$

2.3.2 Ito formula for jump diffusion and Lévy process

Consider a jump-diffusion process defined in chapter 1, i.e. $X_t = \gamma t + \sigma W_t + \sum_{i=1}^{N_t} Y_i$. We can write this process with a different notation:

$$X_t = \underbrace{\gamma t + \sigma W_t}_{X^c(t)} + \underbrace{\sum_{i=1}^{N_t}\Delta X_i}_{J_t} = X^c(t) + J_t \qquad (37)$$

where $\Delta X_i := X(T_i) - X(T_i^-)$. Let f be a C^2 function on $\mathbb{R}$ and let T_i, $i=1,\dots,N_T$, be the jump times of X. Then, we can define $Y_t = f(X_t)$ and we can say that X, between T_i and T_{i+1}, evolves according to the differential equation equal to:

$$dX_t = dX_t^c = \gamma dt + \sigma dW_t$$

The following proposition summarized the result for the jump-diffusion processes when σ is a nonanticipating square-integrable process:

[1] J_X can be defined as: $J_X = \sum_{n\geq 1, T_n\leq T}\delta_{(T_n,Y_n)}$, where δ is the dirac measure.

Proposition 2.8 (Ito formula for jump-diffusion processes) (proposition 8.14 in [2]) *Let X be a diffusion process with jumps defined as:*

$$X_t = X_0 + \int_0^t b_s ds + \int_0^t \sigma_s dW_s + \sum_{i=1}^{N_t} \Delta X_i$$

where $\sum_{i=1}^{N_t} \Delta X_i$ is a compound Poisson process and b_t and σ_t are continuous nonanticipating processes with satisfy the condition:

$$E\left[\int_0^t \sigma_t dt\right]^2 = E\left[\int_0^t \sigma_t^2 dt\right] < \infty$$

Then, for any $C^{1,2}$ function $f : [0,T] \times \mathbb{R} \to \mathbb{R}$, the process $Y_t = f(t, X_t)$ can be represented as:

$$\begin{aligned} f(t, X_t) - f(0, X_0) = & \int_0^t \left[\frac{\partial f}{\partial s}(s, X_s) + \frac{\partial f}{\partial x}(s, X_s) b_s\right] ds \\ & + \frac{1}{2}\int_0^t \sigma_s^2 \frac{\partial^2 f}{\partial x^2}(s, X_s) ds + \int_0^t \frac{\partial f}{\partial x}(s, X_s)\sigma_s dW_s \\ & + \sum_{i \geq 1, T_1 \leq t} \Big(f(X_{T_i^-} + \Delta X_i) - f(X_{T_i^-})\Big) \end{aligned} \tag{38}$$

The equation (2.24) *can be written in differential notation as:*

$$\begin{aligned} dY_t = & \frac{\partial f}{\partial t}(t, X_t) dt + b_t \frac{\partial f}{\partial x}(t, X_t) dt + \frac{\sigma_t^2}{2}\frac{\partial^2 f}{\partial x^2}(t, X_t) dt \\ & + \frac{\partial f}{\partial x}(t.X_t)\sigma_t dW_t + (f(X_{t^-} + \Delta X_t) - f(X_{t^-})) \end{aligned} \tag{39}$$

3 Hedging Strategy

This chapter describe how to compute the option price in an jump diffusion model. The first section talk about the measure transformation, which represent the main tool to find the risk-neutral probability to compute the option pricing. The second part of the chapter introduces the concept of option and, in particular, of European call option. Then, we will see how to use an European call option for hedging purpose. The last part of this chapter is entirely focused on compare the hedging in the Black and Scholes model with the hedging in the Merton model for the jump diffusion process

3.1 Measure Transformation

One normal assumption in each model built in finance is that the market is complete, which means that every contingent claim in the market is attainable. Hence, there exist only one risk-neutral probability in the market. Unfortunately, the complete market assumption is not true in the real market because the asset prices have jumps, which imply that there is not a unique risk-neutral probability but we can find a much greater variety of equivalent measure by changing the distribution of jumps.
In the Black-Scholes model to find the equivalent measure we use the Radon-Nikodym theorem. Hence, we need to introduce the concempt of equivalent measure. Let $(\Omega, \mathcal{F})$ be a measurable space and let $\mathbb{Q}$, $\mathbb{P}$ be two probability measure on $\mathcal{F}$. Then, we say that $\mathbb{Q}$ is absolutely continuous respect to $\mathbb{P}$ ($\mathbb{P} \gg \mathbb{Q}$) if:

$$\forall A \in \mathcal{F} \qquad \mathbb{P}(A) = 0 \Rightarrow \mathbb{Q}(A) = 0$$

Therefore, we can say that two probability measure $\mathbb{Q}, \mathbb{P}$ on $\mathcal{F}$ are equivalent ($\mathbb{P} \sim \mathbb{Q}$) if $\mathbb{P} \gg \mathbb{Q}$ and $\mathbb{Q} \gg \mathbb{P}$, hence if $\mathbb{Q}$ and $\mathbb{P}$ define the same set of impossible events:

$$\forall A \in \mathcal{F} \qquad \mathbb{Q}(A) = 0 \Leftrightarrow \mathbb{P}(A) = 0 \tag{40}$$

Therefore, we have the following theorem:
Theorem 3.1 [Radon-Nikodym Theorem] *Let $\mathbb{P} \gg \mathbb{Q}$, then exist a random variable Λ, $\mathcal{F}$-measurable, with non-negative value such that for every random variable X ($\mathcal{F}$-measurable) integrable under $\mathbb{P}$ the following relation is true:*

$$E_{\mathbb{Q}}[X] = E_{\mathbb{P}}[\Lambda X] = \int_A \Lambda d\mathbb{P}$$

In particular:

$$\forall A \in \mathcal{F} \qquad \mathbb{Q}(A) = E_{\mathbb{P}}(\Lambda \mathbf{1}_A).$$

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space which describe a market between 0 and T. Then, we can define the underlying asset S by a nonanticipating (càdlàg) process:

$$\begin{aligned} S : [0,T] \times \Omega &\mapsto \mathbb{R}^{d+1} \\ (t, \omega) &\mapsto (S_t^0(\omega), S_t^1(\omega), S_t^2(\omega), \ldots, S_t^d(\omega)) \end{aligned}$$

where $S_t^i(\omega)$ represent the value of the asset i in the market scenario ω and $S_t^0(\omega)$ is a numeraire (we define it as $S_t^0(\omega) = e^{rt}$, where r is the interest rate) A self-financing strategy (ϕ_t^0, ϕ_t), in the Black-Scholes model, is said to be a perfect hedge or a replication strategy for a contingent claim H, if we have the following:

$$H = V_0 + \int_0^T \phi_t dS_t + \int_0^T \phi_t^0 dS_t^0 \qquad \mathbb{P} - a.s. \tag{41}$$

where S_t is the asset price. Moreover, we can say that a market is complete if any contingent claim H, admits a replicating portfolio which means that for any $H \in \mathcal{H}$ there exists a self-financing strategy (ϕ_t^0, ϕ_t) such that the equation (3.2) holds with probability 1 under $\mathbb{P}$. If the equation (3.2) holds with probability 1, it also holds with probability 1 under any equivalent martingale measure $\mathbb{Q} \sim \mathbb{P}$. Therefore, we find the following proposition:
Proposition 3.1 *A market defined by the asset $(S_t^0, S_t^1, \ldots, S_t^d)_{0 \leq t \leq T}$ described as stochastic processes on $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ is complete if and only if there is a unique martingale measure $\mathbb{Q}$ equivalent to $\mathbb{P}$.*
If we consider a discount factor equal to $B(t,T) = e^{-r(T-t)}$, then we can write the discounted value of H (equation (4.2)) as:

$$\hat{H} = V_0 + \int_0^T \phi_t d\hat{S}_t \qquad \mathbb{Q} - a.s. \tag{42}$$

3.1.1 Equivalence measures in jump processes

We will study such changes of measure in the jump process. When we consider Lévy process the equivalence of their measures, gives relations between their parameters.
Consider two Poisson process defined by jump size, respectively, equal to a_1, a_2 and jump intensity, respectively, equal to λ_1, λ_2. Then, the following proposition shows the equivalence of measure for Poisson processes:
Proposition 3.2 (Equivalence of measure for Poisson processes) (proposition 9.5 in [2]) *Let $(N, \mathbb{P}_{\lambda_1})$ and $(N, \mathbb{P}_{\lambda_2})$ be Poisson process on $(\Omega, \mathcal{F}_t)$ with intensities λ_1 and λ_2 and jump sizes a_1 and a_2. Then, we have:*

1. *if $a_1 = a_2$, then $\mathbb{P}_{\lambda_1}$ is equivalent to $\mathbb{P}_{\lambda_2}$ with Radon-Nikodym density:*

$$\frac{d\mathbb{P}_{\lambda_1}}{d\mathbb{P}_{\lambda_2}} = exp\left[(\lambda_2 - \lambda_1)T - N_T \ln \frac{\lambda_2}{\lambda 1}\right] \tag{43}$$

2. if $a_1 \neq a_2$, then $\mathbb{P}_{\lambda_1}$ is not equivalent to $\mathbb{P}_{\lambda_2}$.

This proposition told us that if we want the equivalence measure of two Poisson process, we can freely change the intensity of the jumps but the jump size must remain the same. In other word, the intensity of a Poisson process can be modified without changing the jump size of the process, but with changing the size of the jumps, which generates a new measure. This new measure assigns nonzero probability to some events which otherwise were impossible under the old one. We can note that two Poisson processes with different intensities define equivalent measures only on a finite time interval.
Now, consider two compound Poisson process and the following proposition gives us the equivalence of measure in this case:
Proposition 3.3 (Equivalence of measure for compound Poisson processes) (proposition 9.6 in [2]) *Let $(X, \mathbb{P})$ and $(X, \mathbb{Q})$ be compound Poisson processes on $(\Omega, \mathcal{F}_t)$ with Lévy measure $\nu_{\mathbb{P}}$ and $\nu_{\mathbb{Q}}$. The probability $\mathbb{P}$ and $\mathbb{Q}$ are equivalent if and only if $\nu_{\mathbb{P}}$ and $\nu_{\mathbb{Q}}$ are equivalent. In this case, the Radon-Nikodym derivative is:*

$$D_T = \frac{d\mathbb{Q}}{d\mathbb{P}} = exp\left(T(\lambda_{\mathbb{P}} - \lambda_{\mathbb{Q}}) + \sum_{s \leq T} \phi(\Delta X_s)\right) \tag{44}$$

where $\lambda_{\mathbb{P}} \equiv \nu_{\mathbb{P}}(\mathbb{R})$ and $\lambda_{\mathbb{Q}} \equiv \nu_{\mathbb{Q}}(\mathbb{R})$ are the jumps intensities of the two processes and $\phi \equiv \ln\left(\frac{d\nu_{\mathbb{Q}}}{d\nu_{\mathbb{P}}}\right)$.
The last important change of measure with respect the Brownian motion with drift and the following proposition gives us the equivalence of measure in this case:
Proposition 3.4 (Equivalence of measure for Brownian motion with drift) (proposition 9.7 in [2]) *Let $(X, \mathbb{P})$ and $(X, \mathbb{Q})$ be two Brownian motion on $(\Omega, \mathcal{F}_t)$ with volatilities $\sigma_{\mathbb{P}} > 0$ and $\sigma_{\mathbb{Q}} > 0$ and drift $\mu_{\mathbb{P}}$ and $\mu_{\mathbb{Q}}$. The probability $\mathbb{P}$ and $\mathbb{Q}$ are equivalent if and only if $\sigma_{\mathbb{P}} = \sigma_{\mathbb{Q}} > 0$ and singular otherwise. Then, when they are equivalent the Radon-Nikodym derivative is:*

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = exp\left(\frac{\mu_{\mathbb{Q}} - \mu_{\mathbb{P}}}{\sigma^2} X_T - \frac{1}{2}\frac{(\mu_{\mathbb{Q}} - \mu_{\mathbb{P}})^2}{\sigma^2} T\right) \tag{45}$$

With the Cameron-Martin theorem can rewrite the equation (4.9) as an exponential martingale equal to:

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = exp\left(\frac{\mu_{\mathbb{Q}} - \mu_{\mathbb{P}}}{\sigma} W_T - \frac{1}{2}\frac{(\mu_{\mathbb{Q}} - \mu_{\mathbb{P}})^2}{\sigma^2} T\right)$$

where $W_T = \frac{X_t - \mu_{\mathbb{P}} t}{\sigma}$ is a standard Brownian motion under the probability $\mathbb{P}$. This result shows that the drift and the volatility play a crucial roles in defining a diffusion model.

3.2 Option Pricing

The modern finance is centered on the pricing of derivative instruments, which are instruments whose payoff is a function of the value of another financial instruments (such as commodities, currency, bond, stock), also called underlying asset. One of the most popular derivative contract in the world is the option contract. An option is a contract between a buyer and a seller that gives at the purchaser of the option the right, but not the obligation, to buy or to sell a particular asset at an exercise date at an agreed price (exercise price). Later in this chapter, we denote with K the strike or exercise price, with T the exercise date or maturity and with S_T the value of the asset at the maturity. On one hand, we have a call option when we have the right to buy an asset S for K at time T and we can represent its payout at time T as:

$$C_T = \max(S_T - K, 0) = (S_T - K)^+ \tag{46}$$

on the other hand, we have a put option when we have the right to sell an asset S for K at time T and we can represent its payout at time T as:

$$P_T = \max(K - S_T, 0) = (K - S_T)^+ \tag{47}$$

In the market we can find two type of option contract: European option and American option. In the European option, we can exercise the option only at the maturity, instead in the American one, we can exercise the option at any time t, with $t \leq T$.
Let $V_t(T, K)$ be the value at time t of the forward contract [2] with delivery price K. Then, we can define the forward price $F(t, T)$ at current time $t \leq T$ to be the delivery price K such that $V_t(T, K) = 0$, in other words, such that the forward contract has zero value at time t. Therefore, we find the following relationship:

$$V_t(T, K) = (F(t, T) - K)e^{-r(T-t)}$$

Now, we can find how the price of a call and a put of the same strike are related with the value of the forward:

$$C_t(T, K) - P_t(T, K) = V_t(T, K) \tag{48}$$

The above equation is called Put-Call Parity, which states that long one call and short one put is equal to go long to one forward. After some transformation, the Put-Call Parity can be written as:

$$C_t(T, K) - P_t(T, K) = S_t - Ke^{-r(T-t)} \tag{49}$$

The Put-Call Parity is important for three reason. First, it is an arbitrage-free condition. In fact, any violation of the Put-Call Parity leads to an arbitrage opportunity. Second, when we want pricing an option, we can focus only in a call (for example) and then find the price of the put using the Put-Call Parity. Third, the Put-Call Parity is model-independent, which means that this parity relationship between the values of put and call options holds, regardless of the model assumed for the evolution of the price of the underlying asset or arbitrage opportunities occur.

3.2.1 Pricing European Option in Black-Scholes model

In the Black and Scholes model the behavior of prices is a continuous time model with the assumption of one risky asset (denoted by S_t at time t) and a riskless asset (denoted by S_t^0 at time t). Moreover, we assume that the risky asset will not pay dividend and that the behavior of the riskless asset is expressed by the following ordinary differential equation:

$$dS_t^0 = rS_t^0 dt \tag{50}$$

where r is an instantaneous interest rate and it is a non-negative constant. We also set that S_0^0, which imply that $S_t^0 = e^{rt}$ for $t \geq 0$. On the other hand, the behavior of the risky asset is determined by the following stochastic differential equation:

$$dS_t = S_t(\mu dt + \sigma dB_t) \tag{51}$$

where B_t is a standard Brownian motion defined in the probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \leq t \leq T}, \mathbb{P})$ and μ and σ (called the volatility of the asset) are two constant, which are bounded and locally Lipschitz continuous. We consider the model valid for the time interval $[0, T]$, where T is the maturity date of the option. Equation (3.12) has a closed-form solution equal to:

$$S_t = S_0 e^{(\mu - \frac{\sigma^2}{2})t + \sigma B_t} \tag{52}$$

where S_0 is the spot price at time 0. Now, we need to show that there exist a probability equivalent to $\mathbb{P}$ under which the discounted stock price is a martingale. Therefore, we need to introduce the following theorem, called Girsanov theorem:
Theorem 3.2 (theorem 4.2.2 in [1]) *Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{0 \leq t \leq T}, \mathbb{P})$ be a filtered probability space and $(B_t)_{0 \leq t \leq T}$ an $\mathcal{F}_t$-standard Brownian motion. Let $(\theta_t)_{0 \leq t \leq T}$ be an adapted process satisfying $\int_0^T \theta_s^2 ds < \infty$ a.s. and such that the process $(L_t)_{0 \leq t \leq T}$ defined by*

$$L_t = exp\left(-\int_0^t \theta_s dB_s - \frac{1}{2}\int_0^t \theta_s^2 ds\right) \tag{53}$$

[2] A forward contract (or forward) is an agreement between two counterparties to trade a specific asset at a certain future time T, called maturity, and at a certain price K, called delivery price.

is a martingale. Then, under the probability $\mathbb{P}^{(L)}$ with density L_T with respect to $\mathbb{P}$, the process $(W_t)_{0\leq t\leq T}$ defined by $W_t = B_t + \int_0^t \theta_s ds$ is an $\mathcal{F}_t$-standard Brownian motion.
If we define the discounted value as $\tilde{S}_t = e^{-rt}S_t$, we find:

$$d\tilde{S}_t = -re^{-rt}S_t dt + e^{-rt}dS_t$$

We can substitute inside the above equation the equation (3.13) and we find:

$$\begin{aligned} d\tilde{S}_t &= -re^{-rt}S_t dt + e^{-rt}(S_t(\mu dt + \sigma dB_t)) \\ &= \tilde{S}_t((\mu - r)dt + \sigma dB_t) \end{aligned}$$

If we set $W_t = B_t(\mu - r)t/\sigma$, we can rewrite the above result as: $d\tilde{S}_t = \tilde{S}_t\sigma dW_t$. Now, if we apply the theorem 3.2 with $\theta = (\mu - r)/\sigma$, we will find the probability $\mathbb{Q}$ equivalent to $\mathbb{P}$ under which $(W_t)_{0\leq t\leq T}$ is a standard Brownian motion. Then, under the probability $\mathbb{Q}$, $(\tilde{S}_t)$ is a martingale and is equal to: $\tilde{S}_t = \tilde{S}_0 e^{\sigma W_t - \frac{\sigma^2 t}{2}}$. Finally, we can price an option and, in particular, we will focus on European option and for simplicity we use the notation of $h = f(x) = (x - K)^+$ for the call option.
Theorem 3.3 (theorem 4.3.2 in [1]) *In the Black-Scholes model, any option defined by a non-negative, $\mathcal{F}_T$-measurable random variable h, which is square-integrable under the probability $\mathbb{Q}$, is replicable and the value at time t of any replicating portfolio is given by:*

$$V_t = E_{\mathbb{Q}}[e^{-r(T-t)}h|\mathcal{F}_t]$$

Hence, the option value at time t can be defined by the expression $E_{\mathbb{Q}}\left[e^{-r(T-t)}h \,\middle|\, \mathcal{F}_t\right]$. When the random variable h can be written as $h = f(S_T)$, we can express the option value V_t at time t as a function of t and S_t. Then, we have:

$$\begin{aligned} V_t &= E_{\mathbb{Q}}\left[e^{-r(T-t)}f(S_T) \,\middle|\, \mathcal{F}_t]\right] \\ &= E_{\mathbb{Q}}\left[e^{-r(T-t)}f\left(S_t e^{r(T-t)}e^{\sigma(W_T - W_t) - \frac{\sigma^2}{2}(T-t)}\right) \,\middle|\, \mathcal{F}_t\right] \end{aligned}$$

We can note that the random variable S_t is $\mathcal{F}_t$-measurable and, under the probability $\mathbb{Q}$, $W_T - W_t$ is independent of $\mathcal{F}_t$. Therefore, we conclude that: $V_t = F(t, S_t)$ where

$$F(t,x) = E_{\mathbb{Q}}\left[e^{-r(T-t)}f\left(xe^{r(T-t)}e^{\sigma(W_T - W_t) - \frac{\sigma^2}{2}(T-t)}\right)\right] \tag{54}$$

Since, under$\mathbb{Q}$, $W_T - W_t$ is a zero-mean normal random variable with variance $T - t$ and if we consider a call option, where $F(x) = (x - K)^+$, then the equation above can be written as:

$$\begin{aligned} F(t,x) &= E_{\mathbb{Q}}\left[e^{-r(T-t)}f\left(xe^{\sigma(W_T - W_t) + (r - \frac{\sigma^2}{2})(T-t)} - K\right)^+\right] \\ &= E_{\mathbb{Q}}\left[\left(xe^{\sigma\sqrt{\theta}g - \frac{\sigma^2\theta}{2}} - Ke^{-r\theta}\right)^+\right] \\ &= E\left[\left(xe^{\sigma\sqrt{\theta}g - \frac{\sigma^2\theta}{2}} - Ke^{-r\theta}\right)\mathbf{1}_{g+d_2\geq 0}\right] \\ &= \int_{-\infty}^{d_2}\left(xe^{\sigma\sqrt{\theta}g - \frac{\sigma^2\theta}{2}}\right)\frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}} - \int_{-\infty}^{d_2}\left(Ke^{-r\theta}\right)\frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}}dy \end{aligned}$$

where g is a standard Gaussian variable, $\theta = T - t$, $d_1 = \dfrac{\ln\left(\frac{x}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)\theta}{\sigma\sqrt{\theta}}$ and $d_2 = d_1 - \sigma\sqrt{\theta}$. Now, in the first integral we use a change of variable with $z = y + \sigma\sqrt{\theta}$ and the last equation above become:

$$F(t,x) = xN(d_1) - Ke^{-r\theta}N(d_2) \tag{55}$$

where $N(d) = \dfrac{1}{\sqrt{2\pi}}\displaystyle\int_{-\infty}^{d} e^{-x^2/2}dx$ is the Gaussian cumulative distribution function. The equation (4.29) is the price of the call option in the Black-Scholes model. On the other hand, the price of a put in the Black-Scholes model is equal to: $F(t,x) = Ke^{-r\theta}N(-d_2) - xN(-d_1)$

3.2.2 Pricing European Option in jump diffusion process

Let $(S_t)_{0\leq t\leq T}$ be a stochastic process, which describe the asset price behavior and let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_t, \mathbb{P})$ be a probability space, where $\mathbb{P}_t$ represents the history of the asset price. We saw in the previous section that the dynamic of an asset price can be written as: $S_t = S_0 e^{B_t^0}$, where $B_t^0 = \left(\mu - \frac{\sigma^2}{2}\right) t + \sigma W_t$. Consider an exponential price process of the form:

$$S_t := S_0 e^{\mu t + \sigma W_t + Y_t}$$

where Y_t is compound Poisson process. Therefore, the process S_t can be written as:

$$\begin{aligned}
S_t &= S_0 exp\left(\mu t + \sigma W_t + \sum_{i=1}^{N_t} Z_i\right) \\
&= S_0 e^{\mu t + \sigma W_t} \prod_{i=1}^{N_t} e^{Z_i} \\
&= S_0 e^{\mu t + \sigma W_t} \prod_{0\leq s\leq t} e^{\Delta Y_t}, \qquad t \in \mathbb{R}^+
\end{aligned}$$

from relation $\Delta Y_t = Z_{N_t} \Delta N_t$. The process $(S_t)_{t\in\mathbb{R}}$ is equivalently given by the log-returns dynamics: $d\log S_t = \mu dt + \sigma dW_t + dY_t$ with $t \in \mathbb{R}^+$. Then, in exponential model we have: $S_t = S_0 e^{\left(\mu \frac{\sigma^2}{2}\right)t + \sigma W_t - \frac{t\sigma^2}{2} + Y_t}$ and the process S_t satisfies the stochastic differential equation:

$$\begin{aligned}
dS_t &= \left(\mu + \frac{\sigma^2}{2}\right) S_t dt + \sigma S_t dW_t + S_{t^-}(e^{\Delta Y_t} - 1) dN_t \\
&= \left(\mu + \frac{\sigma^2}{2}\right) S_t dt + \sigma S_t dW_t + S_{t^-}(e^{Z_{N_t}} - 1) dN_t
\end{aligned}$$

We can see that the process S_t has jump size equal to: $S_{t^-}(e^{Z_{N_t}} - 1)$. In order for the discounted price process $(e^{-rt}S_t)_{t\in\mathbb{R}}$ to be a martingale, we need to choose a drift parameter $\tilde{\mu} \in \mathbb{R}$, intensity $\tilde{\lambda} > 0$ and jump distribution $\tilde{\nu}$ satisfying the equation:

$$\mu - r = \sigma\tilde{\mu} - \tilde{\lambda} E_{\tilde{\nu}}[Z]$$

Therefore:

$$\mu + \frac{\sigma^2}{2} - r = \sigma\tilde{\mu} - \tilde{\lambda} E_{\tilde{\nu}}[e^Z - 1]$$

under this condition we can choose a risk-neutral probability $\mathbb{P}_{\tilde{\mu},\tilde{\lambda},\tilde{\nu}}$ under which $(e^{-rt}S_t)_{t\in\mathbb{R}}$ is a martingale, for simplicity of notation we denoted the probability $\mathbb{P}_{\tilde{\mu},\tilde{\lambda},\tilde{\nu}}$ with $\mathbb{Q}$. Then, the discounted expected value with respect the new probability measure represent a non-unique arbitrage price at time $t \in [0, T]$ for the contingent claim with payoff $f(S_T)$, hence we have

$$e^{-r(T-t)} E_{\mathbb{Q}}[f(S_T)|\mathcal{F}_t]$$

Set $\theta = T - t$. Then, we can express this arbitrage price as:

$$\begin{aligned}
e^{-r(T-t)} E_{\mathbb{Q}}[f(S_T)|\mathcal{F}_t] &= e^{-r\theta} E_{\mathbb{Q}}[f(S_0 e^{\mu T + \sigma W_T + Y_T})|\mathcal{F}_t] \\
&= e^{-r\theta} E_{\mathbb{Q}}[f(S_0 e^{\mu(T-t) + \sigma(W_T - W_t) + W_T - Y_t})|\mathcal{F}_t] \\
&= e^{-r\theta} E_{\mathbb{Q}}\left[f\left(x \exp\left(\mu\theta + \sigma(W_T - W_t) + \sum_{i=N_t+1}^{N_t} Z_i\right)\right)\right]_{x=S_t} \\
&= e^{-\theta(r+\tilde{\lambda})} \sum_{n=0}^{\infty} \frac{(\tilde{\lambda}\theta)^n}{n!} E_{\mathbb{Q}}\left[f\left(x e^{\mu\theta + \sigma(W_T - W_t)} \exp\left(\sum_{i=1}^{n} Z_i\right)\right)\right]_{x=S_t}
\end{aligned}$$

3.2.3 Implied Volatility

One of the main advantages of the Black-Scholes formula is the fact that the pricing formula, as well as the hedging formula, depend only on one non-observable parameter: the volatility σ. In fact, the drift parameter μ disappears by changing the probability measure. In the Black-Scholes model $\nu = 0$ and the call option prices are uniquely given by the equation:

$$F(t, x) = C^{BS} = xN(d_1) - Ke^{-r\theta}N(d_2)$$

If we fixed all the parameters of the equation (3.16), we see that the value of the call in the Black-Scholes model is an increasing continuous function of σ, mapping $]0, \infty[$ into $](S_t - Ke^{-r\theta})^+, S_t[$. The last interval represent an arbitrage bound for a call option prices. Therefore, we can defined the Black-Scholes implied volatility of the option, denoted by $\sigma_t^{IV}(T, K)$, as the value of the volatility of the underlying instrument, which when substituted into the Black-Scholes formula, will return the correct option prices, denoted by $C_t^*(T, K)$:

$$\exists!\, \sigma_t^{IV}(T, K) > 0, \qquad C^{BS}(S_t, K, \theta, \sigma_t^{IV}(T, K)) = C_t^*(K, T)$$

We can note that, for fixed (T, K), the implied volatility is in general a stochastic process. Furthermore, if we fixed t, we will find the implied volatility surface at date t, which is equal at the function $\sigma_t^{IV} : (T, K) \to \sigma_t^{IV}(T, K)$. This means that, for fixed t, the implied volatility value depends on the characteristics of the option such as the maturity and the strike price, respectively equal to T and K. Moreover, if we substitute the moneyness m (i.e $m = \frac{K}{S_t}$) into the implied volatility surface, it can be represented as a function of moneyness and time to maturity: $I_t(\theta, m) = \sigma_t^{IV}(t+\theta, mS(t))$. In general, the implied volatility surface $I_t(\theta, m)$ may depend not only on the maturity of options but also on the current date or the spot price. However, in the exponential-Lévy models the evolution in time of implied volatilities is particularly simple, as shown by the following proposition:
Proposition 3.5 (Proposition 11.1 in [2]) *When the risk neutral dynamics is given by an exponential-Lévy process, the implied volatility for a given meneyness level $m = \frac{K}{S_t}$ and time to maturity θ, i.e $\theta = T - t$, does not depend on time:*

$$\forall t \geq 0, \quad I_t(\theta, m) = I_0(\theta, m)$$

However, we can note that the implied volatility for a given strike price, K, is not constant in time. In fact, it evolves stochastically according to: $\sigma_t^{IV} = I_0\left(\frac{K}{S_t}, T - t\right)$. We can note that the implied volatility surface I_t does not vary with t, therefore we can study only the case in which $t = 0$. This study explain some features of the implied volatility surface in the exponential-Lévy model. First, a negative skewed jump distribution give rise to a skew in implied volatility, hence the skew decrease characteristic with respect to moneyness. On the other hand, a strong variance of jumps leads to a curvature in the implied volatility, hence we can see smile pattern. Second, exponential-Lévy models and, in general, model with jumps in the price lead to a strong short term skew contrarily to diffusion models which have small skew for short maturities. Finally, in a Lévy process with finite variance we can see the effect called aggregation normality, which is when long maturity prices of options will be cloeser to Black-Scholes price and the implied volatility smile will become flat as $T \to \infty$. In particular, the central limit theorem shows that when the maturity T is large, the distribution of $(X_T - E[X_T])/\sqrt{T}$ becomes approximately Gaussian. This effect is more pronounced in exponential-Lévy models respect to the actual market prices.

3.3 Hedging Strategy

Consider an asset prices $(S_t)_{t\in[0,T]}$ and a market described by a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$, where $(\mathcal{F}_t)_{t\in[0,T]}$ is the history of the assets, $\mathbb{P}$ represents the so-called real-word measure and S_t will be one dimensional. We assume that there are two assets in the market: a riskless asset, described by the following differential equation $dS_t^0 = rS_t^0 dt$, and a risky asset, S_t. Let $S_t^0 = e^{rt}$ be a numeraire. Then, we denoted by V_t the value of a portfolio and by $\tilde{V}_t$ its discounted value, which is equal to $\tilde{V}_t = V_t/S_t^0$.

3.3.1 Black-Scholes Hedging

Consider the Black-Scholes model. The behavior of the stock price is represent by equation: $S_t = S_0 e^{(\mu - \frac{\sigma^2}{2})t + \sigma W_t}$ and the price of call option in the Black-Scholes model are equal to the equation:

$$C^{BS}(t, S) = SN(d_1) - Ke^{-r\theta} N(d_2) \tag{56}$$

where $\theta = T - t$ and $d_1 = \dfrac{\ln\left(\frac{S}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)\theta}{\sigma\sqrt{\theta}}$ and $d_2 = d_1 - \sigma\sqrt{\theta}$.

Let V be the value of a portfolio of derivative securities on one underlying asset. The rates of change of the value of the portfolio with respect to the spot price S of the underlying asset is important for hedging purpose. This change is called "Delta" and is equal to:

$$\Delta(V) = \frac{\partial V}{\partial S}$$

Then, the delta of the call option described in equation (3.17) is equal to:

$$\Delta(C^{BS}) = \frac{\partial C^{BS}}{\partial S} = N(d_1) \tag{57}$$

Hence, we have found that the delta for an European call option in the Black-Scholes model is equal to the cumulative distribution function of a standard normal variable evaluated in d_1. The delta in an option is important because helps to build the so-called "delta hedging". Assume that we go long in one call option. If the price of the underlying asset declines, the value of the call decreases and the long call position loses money. To protect against a downturn in the price of the underlying asset, we can sell short Δ units of the underlying asset. The goal of the delta hedging is to choose Δ in such a way that the value of the portfolio is not sensitive to small changes in the price of the underlying asset. Therefore, if V is the value of the portfolio, the value of the hedge portfolio is equal to: $V = C(t, S) - \Delta S_t$.

3.3.2 Merton Approach

The delta hedging in the Black-Scholes model is always possible since the market is complete and, therefore, exists only one equivalent risk neutral probability. This is the main assumption in the Black-Scholes model. Unfortunately, the market is not complete and there is not a unique risk neutral probability because the asset has discontinuities, i.e jumps, in their paths.

The first application of jump process in option pricing was introduced by Merton. He considered the following jump diffusion model defined in the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$:

$$S_t = S_0 \exp\left(\mu t + \sigma W_t + \sum_{i=1}^{N_t} Y_i\right) \tag{58}$$

where W_t is a standard Wiener process, N_t is a Poisson process with intensity λ independent from W and $Y_i \sim N(m, \delta^2)$ are i.i.d. random variables independent from W and N. Since the model is incomplete, there exists many possible choices for a risk-neutral measure and Merton proposed to change the drift of the Wiener process and keep the other variable unchanged. Therefore, μ^M is chosen such that $\hat{S}_t = S_t e^{-rt}$ is a martingale under the new probability measure $\mathbb{Q}$, which is the equivalent probability measure to $\mathbb{P}$, and is equal to

$$\mu^M = r - \frac{\sigma^2}{2} - \lambda E[e^{Y_i} - 1] = r - \frac{\sigma^2}{2} - \lambda\left[\exp\left(m + \frac{\delta^2}{2}\right) - 1\right]$$

The equivalent martingale measure is obtained by shifting the drift of the Brownian motion but leaving the jump part unchanged. Merton justified this choice by assuming that the jump risk is diversifiable and, therefore, no risk premium is attached to it. Then, an european call option with payoff $f(S_T)$ can be priced according to:

$$C^M(t, S_t) = e^{-r(T-t)} E_{\mathbb{Q}}[f(S_T)|\mathcal{F}_t] \tag{59}$$

Set $\theta = T - t$ and by conditioning on the number of jumps N_t, we can express the value of the call option as a weighted sum of Black-Scholes price, therefore we find:. Then, we can express this arbitrage price as:

$$\begin{aligned}
C^M(t, S_t) &= e^{-r(T-t)} E_{\mathbb{Q}}[f(S_T - K)^+ | S_t = S] \\
&= e^{-r\theta} E[f\left(e^{\mu^M \theta + \sigma W^M_{T-t} + \sum_{i=1}^{N_{T-t}} Y_i}\right)] \\
&= e^{-r\theta} \sum_{n \geq 0} \mathbb{Q}(N_t = n) E_{\mathbb{Q}} \left[f\left(S \exp\left(\mu^M \theta + \sigma W^M_\theta + \sum_{i=1}^n Y_i \right)\right)\right] \\
&= e^{-r\theta} \sum_{n \geq 0} \frac{e^{-\lambda\theta}(\lambda\theta)^n}{n!} C^{BS}(\theta, S_n, \sigma_n) \qquad (60)
\end{aligned}$$

where: $\sum_{i=1}^n Y_i \sim N(nm, n\delta^2)$; $\sigma_n^2 = \sigma^2 + \frac{n\delta^2}{\theta}$; $S_n = S \exp\left(nm + \frac{n\delta^2}{2} - \lambda \exp(m + \frac{\delta^2}{2}) + \lambda\theta\right)$ and $C^{BS}(\theta, S, \sigma) = e^{-r\theta} E\left[f\left(Se^{(r - \frac{\sigma^2}{2})\theta + \sigma W_\theta}\right)\right]$. We can note that the last condition is the value of the European call option with time to maturity θ and payoff f in a Black-Scholes model with volatility σ. We can note that if $\lambda = 0$ then $C^M(t, S) = C^{BS}(t, S)$, indeed all the terms appearing in the sum (3.21) are equal to 0, except for $j = 0$, when $S_0 = S$ and $\sigma_0 = \sigma$.
The hedging portfolio proposed by Merton is the self-financing strategy (ϕ_t^0, ϕ_t) given by: $\phi_t = \frac{\partial C^M}{\partial S}(t, S_{t-})$ and $\phi_t^0 = \phi_t S_t - \int_0^t \phi dS$. This means that we choose to hedge only the risk represented by the diffusion part. This approach is justified if we assume that the investor holds a portfolio with many assets for which the diffusion components may be correlated but the jumps components are independent across assets. This hypothesis would imply that in a large market a diversified portfolio such as S&P500 would not have jumps. Finally, the assumption of diversifiability of jump risk is not justifiable if we are pricing index options, in fact a jump in the index is not diversifiable.

3.3.3 Quadratic Hedging

We can define the quadratic hedging as the choice of a hedging strategy which minimizes the hedging error in a mean square losses. This imply that losses and gains are treated in a symmetric manner, therefore we measure the risk in terms of variance.
Consider a risk-neutral model $(S_t)_{t\in[0,T]}$ given by $S_t = e^{rt + X_t}$, where X_t is a Lévy process on the filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{Q})$. We assume that S is a square integrable martingale, therefore the following condition is satisfied:

$$\int_{|y|\geq 1} e^{2y} \nu(dy) < \infty$$

Moreover, we assume that X_t has finite variance and its characteristic function can be expressed as:

$$E\left[e^{iuX_t}\right] = \exp\left\{ t \left[-\frac{\sigma^2 u^2}{2} + b_X t + \int \nu_X(dy) \left(e^{iuy} - 1 - iuy\right)\right]\right\}$$

with b_X chosen such that $\hat{S} = e^X$ is a martingale. As we have seen in the previous chapter, $\hat{S}_t$ can also be written as a stochastic exponential of another Lévy process (Z_t):

$$d\hat{S}_t = \hat{S}_t dZ_t$$

where Z is a martingale with jumps size greater than -1 and it is also a Lévy process. Let $(\phi_t^0, \phi_t)_{t\in[0,T]}$ be a self-financing strategy. In order to apply the quadratic hedging criteria, we need to find portfolio such that its terminal value has a well-defined variance. Therefore, we want that the asset S is in the set of all the admissible strategies defined as:

$$\mathcal{S} = \left\{ \phi \text{ caglad predictable and } E\left[\left|\int_0^T \phi_t d\hat{S}_t\right|\right]^2 < \infty \right\}$$

Using preposition 2.4 and the proposition 2.5, the above condition is equivalent to:

$$E\left[\int_0^T |\phi_t\hat{S}_t|^2 dt + \int_0^T \int_{\mathbb{R}} z^2 |\phi_t\hat{S}_t|^2 dt\nu(dz)\right] < \infty \tag{61}$$

Let $L^2(\mathcal{S})$ be the set of process ϕ which verify the above condition (3.22). Therefore, the terminal payoff of such strategy is equal to:

$$G_T(\phi) = \int_0^T r\phi_t^0 dt + \int_0^T \phi_t S_{t^-} dZ_u$$

We can note that $\hat{S}_t$ is a martingale under the probability measure $\mathbb{Q}$ and that $\phi \in L^2(\hat{S})$, therefore the discounted gain process, equal to $\hat{G}_T(\phi) = \int_0^t \phi d\hat{S}$, is also a square integrable martingale. Using proposition 1.8 we find that $\hat{G}_T(\phi)$ is given by the martingale part of the above equation:

$$\begin{aligned}\hat{G}_T(\phi) &= \int_0^T \phi_t S_{t^-}\sigma dW_t + \int_0^T \int_{\mathbb{R}} \tilde{J}_X(dt \times dx) x \phi_t S_{t^-} \\ &= \int_0^T \phi_t S_{t^-}\sigma dW_t + \int_0^T \int_{\mathbb{R}} \tilde{J}_Z(dt \times dz)\phi_t S_{t^-}\,(e^z - 1)\end{aligned}$$

where J is the jump measure. Now, we can written the quadratic hedging problem as:

$$\inf_{\phi \in L^2(\hat{S})} E_{\mathbb{Q}}\left[|\hat{G}_T(\phi) + V_0 - \hat{H}|^2\right] \tag{62}$$

where $\hat{H}$ is defined by the equation: $\hat{H} = V_0 + \int_0^T \phi_t d\hat{S}_t$ $\mathbb{Q}$-a.s.
Proposition 3.6 (proposition 10.5 in [2]) *Consider the risk neutral dynamics*

$$\mathbb{Q}: \qquad d\hat{S}_t = \hat{S}_{t^-} dZ_t \tag{63}$$

where Z is a Lévy process with Lévy measure ν_Z and diffusion coefficient $\sigma > 0$. For a European option with payoff $f(S_T)$ where $f : \mathbb{R}^+ \to \mathbb{R}$ verifies

$$\exists K > 0, \qquad |f(x) - f(y)| \leq K|x - y|$$

the risk minimizing hedge, solution of (3.23), *amounts to holding a position in the underlying equal to $\phi_t = \Delta(t, S_{t^-})$ where:*

$$\Delta(t, S_{t^-}) = \frac{\sigma^2 \frac{\partial C}{\partial S}(t, S) + \frac{1}{S}\int \nu_Z(dy) z[C(t, S(1+z)) - C(t, S)]}{\sigma^2 + \int z^2 \nu_Z(dy)} \tag{64}$$

with $C(t, S) = e^{-r(T-t)} E_{\mathbb{Q}}[f(S_T)|S_t = S]$.

If we consider an exponential-Lévy model, i.e. $S_t = S_0 e^{rt + X_t}$, the optimal quadratic hedge can be expressed in terms of the Lévy measure ν_X of X as

$$\Delta(t, S_{t^-}) = \frac{\sigma^2 \frac{\partial C}{\partial S}(t, S) + \frac{1}{S}\int \nu_X(dx)\,(e^x - 1)\,[C(t, Se^x) - C(t, S)]}{\sigma^2 + \int (e^x - 1)^2 \nu_X(dx)}$$

We can note that we have also found an expression for the residual risk of a hedging strategy (ϕ_t^0, ϕ_t):

$$\begin{aligned}R_T(\phi) = E&\left[\int_0^T \left|\phi_t - \frac{\partial C}{\partial S}(t, S_{t^-})\right|^2 \hat{S}_{t^-}^2 dt\right] \\ &+ E\left[\int_0^T dt \int_{\mathbb{R}} \nu(dz)|C(t, S_{t^-}(1+z)) - C(t, S_{t^-}) - z\phi_t \hat{S}_{t^-}|^2\right]\end{aligned}$$

The residual risk allows us to examine whether there are any cases where the hedging error can be reduced to zero, hence where we can achieve a perfect hedge. We find that in only two case is possible achieve a perfect hedge. The first one, is when there are no jumps, i.e $\nu = 0$. In this case, the residual risk is equal to:

$$\varepsilon(\phi) = E\left[\int_0^T \left(\phi_t S_{t^-} - S_{t^-}\frac{\partial C}{\partial S}(t, S_{t^-})\right)^2 dt\right]$$

and we find that $\varepsilon(\phi) = 0$ a.s when ϕ_t is equal to the Black-Scholes delta hedging. The second case, is when $\sigma = 0$ and there is a single jump size $\nu = \delta_a : X_t = aN_t$, where N is a Poisson process. In this case

$$R_T(\phi) = E\left[\int_0^T dt S_{t^-}^2 |C(t, S_{t^-}(1+a)) - C(t, S_{t^-}) - \phi_t|^2\right]$$

if we choose $\phi_t = \frac{C(t,S_{t^-}(1+a)) - C(t,S_{t^-})}{aS_{t^-}}$ and $\phi_t^0 = e^{rt}S_t\phi_t - e^{rt}\int_0^t \phi_t dS_t$, we will obtain a self-financing strategy $(\phi, \phi^0$ which is a replication strategy:

$$f(S_T) = V_0 + \int_0^T \frac{C(t, S_{t^-}(1+a)) - C(t, S_{t^-})}{aS_{t^-}} dS_t + \int_0^T r\phi_t^0 dt$$

We can note that the quadratic hedge achieves a mean-variance trade-off between the risk due to the diffusion part and the jump risk.

3.4 Comparison

We want to show that the hedging in the Merton model outperforms the hedging in the Black and Scholes model, which are described in the section above. Before we talk about the hedging strategy, we show that the Marton model also outperform the Black-Scholes model to replicate the stock behavior from historical data. We consider the daily log-returns of the Standard & Poor's 500 Index (S&P500) in the period from 31-12-2009 to 29-01-2009. There is a total of 2273 daily closing price and we have to deal with n=2272 log-returns. Moreover, from the S&P500 data it is possible to find the following information: $E^{SP} \simeq 0.00036$; $M_2^{SP} \simeq 0.0095$; $s^{SP} \simeq -0.4666 < 0$ and $k^{SP} \simeq 7.5614 > 3$; where E is the mean, s is the skewness and k is the kurtosis. In order to find a relationship among the two model and the statistical result of the S&P500, we will work with an interval of amplitude Δt, which can be defined as $\Delta t = 1/252 \simeq 0.004$ where the denominator 252 represent the trading days in a year. Therefore, we can write the Black-Scholes model as

$$\Delta \log S_t = \mu_{BS}\Delta T + \sigma_{BS}\Delta W_t \tag{65}$$

where $\Delta W_t \sim N(0, \Delta t)$. While the Merton model can be written as:

$$\Delta \log S_t = \mu_M \Delta T + \sigma_M \Delta W_t + V\Delta N_t \tag{66}$$

where V is the price ratio (> 0) associated with the i-th jump along the path of the stock price and is equal to $V = \log\left(\frac{S_{T_i}}{S_{T_{i^-}}}\right) \sim N(m, \delta^2)$ and $\Delta N_t \sim Po(\lambda\Delta t)$. Then, the following theorem described the relation among the parameter of the two model:

Theorem 3.4 (theorem 1 in [12]) *Consider the equation* (3.26)*, we find the following relation:*

$$\begin{aligned} E^{BS} &= \mu_{BS}\Delta t; \quad M_2^{BS} = \sigma_{BS}^2\Delta t; \\ M_3^{BS} &= 0; \quad M_4^{BS} = 3\sigma_{BS}^4\Delta t^2 \\ \implies s^{BS} &= \frac{M_3^{BS}}{(M_2^{BS})^{3/2}}; \quad k^{BS} = \frac{M_4^{BS}}{(M_2^{BS})^2}. \end{aligned}$$

while for the equation (3.27) *we have*

$$
\begin{aligned}
E^M &= \mu_M \Delta t + m\lambda\Delta t; \quad M_2^M = \sigma_M^2 \Delta t + \left(\delta^2(1+\lambda\Delta t) + m^2\right)\lambda\Delta t; \\
M_3^M &= m(3\delta^2 + m^2)\lambda\Delta t + 6m\delta^2(\lambda\Delta t)^2 \Longrightarrow s^M = \frac{M_3^M}{\left(M_2^M\right)^{3/2}}; \\
M_4^M &= 3(\sigma_M^2\Delta t)^2 + (m^4 + 3\delta^4 + 6m^2\delta^2)\lambda\Delta t + (3m^4 + 21\delta^4 + 30m^2\delta^2)(\lambda\Delta t)^2 \\
&\quad + 6\sigma_M^2\Delta t(\delta^2 + m^2)\lambda\Delta t + (18\delta^2 + 6m^2\delta^2)(\lambda\Delta t)^3 + 6\sigma_M^2\delta^2\Delta t(\lambda\Delta t)^2 \\
&\quad + 3\delta^4(\lambda\Delta t)^4 \Longrightarrow k^M = \frac{M_4^M}{\left(M_2^M\right)^2}.
\end{aligned}
$$

Therefore, if we apply the theorem 3.4, we can find the vector of parameter for the Black and Scholes model (μ_{BS}, σ_{BS}) and for the Merton model $(\mu_{BS}, \sigma_{BS}, \lambda, m, \delta)$. For the Black-Scholes model we assume: $E^{BS} = E^{SP}$ and $M_2^{BS} = M_2^{SP}$ we find that

$$\mu_{BS} = \frac{E^{SP}}{\Delta t} \simeq 0.0922, \qquad \sigma_{BS} = \sqrt{\frac{M_2^{SP}}{\Delta t}} \simeq 0.1507$$

we can recall that a normal distribution is completely determined by its mean and variance. On the other hand, in the Merton model we have 5 parameters to estimate. We can reduce this set assuming that $E^M = E^{SP}$ and $M_2^M = M_2^{SP}$ which implies that

$$\mu_M = \frac{E^{SP} - m\lambda\Delta t}{\Delta t}, \qquad \sigma_M = \sqrt{\frac{M_2^{SP} - \left(\delta^2(1+\lambda\Delta t) + m^2\right)\lambda\Delta t}{\Delta t}}$$

hence, the diffusion paramters are expressed as function of the jumps ones and we have only 3 parameters to estimate. We use the Multinomial Maximum Likelihood approach to estimate this 3 parameters, which can be represented as a 3-dimensional vectors $\eta \doteq (\lambda, m, \delta)$. The step of the Multinomial Maximum Likelihood approach can be summarized as follows: first, sort empirical data into $\tilde{n} < n$ bins, in order to get a computationally tractable problem. Then, for each of these bins, extract the sample frequency f_i^{SP}, $i = i, \ldots, \tilde{n}$. Second, construct the theoretical jump diffusion frequency function

$$f_i^M(\eta) \doteq n \int_{B_i} \psi_{\Delta t}(y;\eta)dy \quad i = 1, \ldots, \tilde{n}$$

where B_i is the i-th bin and $\psi_{\Delta t}(y;\eta)$ is the log-return probability density function for the Merton model (described in chapter 1), i.e. $\psi_t(y) = e^{-\lambda t}\sum_{j=0}^{\infty} \frac{(\lambda t)^j \exp\left\{-\frac{(y-(\mu-\lambda k-\frac{1}{2}\sigma)t-jm)^2}{2(\sigma^2 t + j\delta^2)}\right\}}{j!\sqrt{2\pi(\sigma^2 t + j\delta^2)}}$.
Third, minimize the objective function: $l(\eta) \doteq -\sum_{i=1}^{\tilde{n}} f_i^{SP} \log\left(f_i^M(\eta)\right)$.
Therefore, by the Multinomial Maximum Likelihood algorithm we obtain that

$$\lambda \simeq 62.752; \; m \simeq -0.006323; \; \delta \simeq 0.006291$$

hence, in the Merton model μ, σ, skewness and kurtosis are equal to:

$$\mu_M \simeq 0.48678, \quad \sigma_M \simeq 0.1301, \quad s^M \simeq 1.4261 \quad k^M \simeq 7.9952$$

We can note that the skewness is bigger than the one obtained using the real S&P500 data, i.e. $1.4261 > -0.4666$, but, unlike in the Black and Scholes model where $s^{BS} = 0$, the Merton approach tents to capture a clear absence of symmetry with the same sign. Moreover, the kurtosis in the Merton model is very close to the one obtained using the real S&P500 data, while the Black and Scholes model provides poor result. Hence, we can conclude that the log-normal jump diffusion model represents a substantial and concrete improvement when compared to the Black and Scholes model.

Now, we compare the Black and Scholes hedging strategy, i.e. Delta hedging, with the Merton hedging for the jump diffusion process. We consider the closing price of the S&P500 from 29-12-2017 to 29-01-2019 and we consider a call option with underlying the S&P500, strike price equal to 2700 and maturity at 01-02-2019. Moreover, we assume that the risk-free rate is equal to 2,98%, denoted by r.
For these period, we have a total of 272 daily closing price and we have to deal with $n = 271$ log-returns. From the log-returns we find the following information from the S&P500 data: $E^{SP} \simeq -0.0000468$, $M_2^{SP} \simeq 0.0109$, $s^{SP} \simeq -0.4333 < 0$ and $k^{SP} \simeq 5.9362 > 3$. Therefore, the Black and Scholes parameters can be estimated as follow:

$$\mu_{BS} = \frac{E^{SP}}{\Delta t} \simeq -0.0118, \qquad \sigma_{BS} = \sqrt{\frac{M_2^{SP}}{\Delta t}} \simeq 0.1723$$

The Merton parameter can be estimated using the Multinomial Maximum Likelihood algorithm and we obtain that

$$\lambda \simeq 3.1596; \; m \simeq -0.04942; \; \delta \simeq 0.0076$$

hence, in the Merton model μ, σ, skewness and kurtosis are equal to:

$$\mu_M \simeq 0.1444, \quad \sigma_M \simeq 0.1476, \quad s^M \simeq 1.5929, \quad k^M \simeq 7.8115$$

We can note that also in these case the Merton model represents a substantial and concrete improvement when compared to the Black and Scholes model. Therefore, we can expect that the hedging in the Merton model perform better than the delta hedging in the Black and Scholes.

Then, consider the following hedging strategy for the Black-Scholes model: we assume that we go long in the call option and to protect against a downturn in the price of the underlying asset we will sell short Δ unit of the underlying asset. The goal is to choose Δ in such a way that the value of the portfolio is not sensitive to small changes in the price of the underlying asset. If we denoted with Π the value of the portfolio, then $\Pi = C - \Delta S$ or, equivalently, $\Pi(S) = C(S) - \Delta S$. To implement the Delta hedging we assume that if the Delta is negative we will go long on the asset and short the call option. We can note that a portfolio is Delta neutral only over a short period of time. We recall that the Delta of a call option is equal to the equation (3.18). To implement the Black-Scholes formula the cumulative distribution $N(d)$ of the standard normal variable x must be estimated numerically and we use the algorithm proposed by Abramowitz and Stegun in 1970 which has an approximation error smaller than $7.5 \cdot 10^{-7}$ at any point on the real axis.
On the other hand, for the Merton jump diffusion model we consider the hedging prosed by Merton. Therefore, we find that the price of the call option in this model is equal to the equation (3.21) and then we use the hedging portfolio proposed by Merton is the self-financing strategy (ϕ_t^0, ϕ_t) given by: $\phi_t = \frac{\partial C^M}{\partial S}(t, S_{t^-})$ and $\phi_t^0 = \phi_t S_t - \int_0^t \phi dS$; which means that we choose to hedge only the risk represented by the diffusion part.
The result of these two hedging strategy can be seen in the table below which report the return and the variance:

	Return	Variance (σ^2)
Black-Scholes Hedging	6.40%	0.1589 %
Merton Hedging	6.49%	0.1589 %

We can see that the two hedging strategy have the same variance and the return are more or less the same, the Merton return is a greater only of 0.09 respect to the Black-Scholes return. One possible explanation is that we consider a trading strategy only for one year and, as said before, a portfolio is Delta neutral over a short period of time. Despite this, we can say that the hedging strategy also confirms the above: the Merton model represents a substantial and concrete improvement when compared to the Black and Scholes model. In fact, we can safely say that no-one would choose the Delta hedging compared to Merton hedging as the second has a bigger return, even if small, with the same variance. Therefore, the Merton hedging dominates the Black-Scholes hedging since Merton considers in the stock process a jump component.

References

[1] Damien Lamberton, Bernard Lapeyre, *Introduction to Stochastic Calculus Applied to Finance*, Second edition, Chapman and Hall/CRC, UK.

[2] Rama Cont, Peter Tankov, *Financial Modelling with Jump Processes*, Second edition, Chapman and Hall/CRC, UK (2004).

[3] José E. Figueroa-Lopez, *Jump-Diffusion models driven by Lévy processes*,

[4] Antonis Papapantoleon *An Introduction to Lévy processes with application in Finance*, Lectures Notes, TU Vienna (2008).

[5] Ken-Iti Sato, *Lévy Processes and Infinitely Divisible Distributions*, Cambridge University Press, UK (1994).

[6] Andrea Pascucci, *PDE and Martingale Methods in Option Pricing*, Springer, Bocconi university press, Milano (2010).

[7] Stephen Blyth, *An Introduction to Quantitative Finance*, Oxford university press (2014).

[8] Giuseppe Curci, Mario Dell'Era, *Finanza quantitativa e Modelli Matematici*, Seconda edizione, Pisa university press (2016).

[9] Paolo Baldi, *Equazioni Differenziali Stocastiche e Applicazioni*, Pitagora Editrice, Bologna (2000).

[10] Dan Stefanica, *A Primer for the Mathematics of Financial Engineering*, Second edition, FE Press, New York (2011).

[11] Steven E. Shreve, *Stochastic Calculus fof Finance II, a continuous-time models*, Springer, USA (2004).

[12] Resnik, *A Probability path*,

[13] Allan Gut, *Probability: a graduate course*,

[14] Nicola Gugole, *Merton Jump-diffusion model versus the Black and Scholes approach for the log-returns and volatility smile fitting*, International Journal of Pure and Applied Mathematics, volume 109, pp 719-736.

[15] Barbara Grunewald, Siegfried Tratmann, *Option Hedging in the Presence of Jump Risk*, Department of Law and Economics, Johannes Gutenberg-Universitat Mainz, (1996).