B. Sengul

Introduction to Lévy Processes

University of Bath
Department of Mathematical Sciences
Bath
BA2 7AY
UK
Tel: 01225 388388

bs250@bath.ac.uk

Supervisor: A.E. Kyprianou

Abstract

The aim of this paper is to introduce the reader into Lévy Processes in a formal and rigorous manner. The paper will be analysis based and no probability knowledge is required, thought it will certainly be a tough read in this case. We aim to prove some important theorems that define the structure of Lévy Processes.

The first two chapters are to reacquaint the reader with measure theory and characteristic functions, after which the topic will swiftly move on to infinitely divisible random variables. We will prove the Lévy canonical representation. Then we will go on to prove the existence of Brownian motion and some properties of it, after which we will briefly talk about Poisson processes and measures.

The final chapter is dedicated to Lévy processes in which we will prove three important theorems; Lévy-Khintchine representation, Lévy-Ito decomposition and the points of increase for Lévy processes.

KEYWORDS:*Brownian Motion, Poisson Processes, Lévy Processes, Infinitely Divisible Distributions, Lévy-Itô Decomposition, Lévy-Khintchine Representation, Points of Increase*

Contents

Acknowledgements

First and foremost, my deepest admiration and gratitude goes to Andreas Kyprianou, to whom I owe all of my current knowledge in probability. His support and enthusiasm has always been a source of inspiration for me, and I doubt I can find a measure space which his support would be in. I hope that this project has done justice to the effort he has invested in me.

I would also like to thank Juan Carlos, for being very supportive and taking time out to talk to me about his research and interests. He has pointed me towards interesting areas of probability and Lévy processes.

Also I have to thank Andrew MacPherson, Daria Gromyko and Laura Hewitt for putting up with me constantly talking about my project and giving me valuable feedback.

Last but not least I wish to express my gratitude to Akira Sakai and Adam Kinnison for being a huge source of inspiration for me. If it was not for these people, I would not have been studying probability.

Introduction

The study of Lévy processes began in 1930 though the name did not come along until later in the century. These processes are a generalisation of many stochastic processes that are around, prominent examples being Brownian motion, the Cauchy process and the compound Poisson process. These have some common features; they are all right continuous and have left limits, and they all have stationary independent increments. These properties give a rich underlying understanding of the processes and also allow very general statements to be made about many of the familiar stochastic processes.

The field owes many things to the early works of Paul Lévy, Alexander Khintchine, Kiyosi Itô and Andrey Kolmogorov. There is a lot of active research in Lévy processes, and this paper will lead naturally to subjects such as fluctuation theory, self similar Markov processes and Stable processes.

We will assume no prior knowledge of probability throughout the paper. The reader is assumed to be comfortable with analysis, and in particular L^p spaces and measure theory. The first chapter will brush over these as a reminder.

Notation

$x_n \downarrow x$ will denote a sequence $x_1 \leqslant x_2 \leqslant ...$ such that $x_n \to x$ and similarly $x_n \uparrow x$ will denote $x_1 \geqslant x_2 \geqslant ...$ with $x_n \to x$. $x+$ will be shorthand for $\lim_{y \downarrow x} y$ and $x-$ will mean $\lim_{y \uparrow x} y$. By $\mathbb{R}_+$ we mean the set of non-negative real numbers and $\overline{\mathbb{R}} = \mathbb{R} \cup \{\infty, -\infty\}$ is the extended real line. We will also be using the convention that $\inf \emptyset = \infty$.

We will denote the power set of a set Ω by $\mathscr{P}(\Omega)$. The order (or usual) topology on $\mathbb{R}$ is the topology generated by sets of the form (a, b). We will often abbreviate limit supremums, $\limsup_n A_n := \lim_{n \uparrow \infty} \sup_{k \geqslant n} A_n$. The notation ∂B where B is a set will be used to mean the boundary of B.

A càdlàg (continue à droite, limitèe à gauche) function is one that is right continuous with left limits. Unless specified otherwise, we will follow the convention that N, L, B (or W) will be Poisson, Lévy, and Wiener processes respectively. We will use X when we are talking about a general process or random variable.

Chapter 1

Preliminaries

" The theory of probability as a mathematical discipline can and should be developed from axioms in exactly the same way as geometry and algebra. "

-Andrey Kolmogorov

1.1 Measure Theory

The aim of this chapter is to familiarise the reader with the aspects of measure theory. We will not rely heavily on measure theory in this paper, it is, however, essential to get a basic grasp of the concept in order to do probability.

Definition 1.1.1. A *σ-algebra* $\mathscr{F}$ on a set Ω is a collection of subsets of Ω such that,

(i) $\emptyset \in \mathscr{F}$ and $\Omega \in \mathscr{F}$

(ii) $A \in \mathscr{F} \implies A^c \in \mathscr{F}$

(iii) $\{A_n\}_{n\in\mathbb{N}} \subset \mathscr{F} \implies \bigcup_{n\in\mathbb{N}} A_n \in \mathscr{F}$.

We call the pair $(\Omega, \mathscr{F})$ a *measurable space.*

From this we can use de Morgan's laws to deduce that a σ-algebra is also closed under countable intersection. The elements of a σ-algebra can be viewed as events, Ω being the complete event (in the sense that it is the event "something happens"). It is clear that if we have an event A, then we also have an event of A not not happening. Finite intersection and union may also be justified in terms of events, the sole reason for the countable union and intersections are however, for the purpose of analysis.

A simple question would be on how to obtain a σ-algebra from a given collection of subsets.

Proposition 1.1.2. *Let $\mathscr{T}$ be a collection of sets of Ω, then there exists a smallest σ-algebra $\mathscr{B}$ such that $\mathscr{T} \subset \mathscr{B}$.*

Proof. Take the intersection of all the σ-algebras that contain $\mathscr{T}$ (there is at least one σ-algebra, namely $\mathscr{P}(\Omega)$). This intersection is also a σ-algebra (a fact that the reader may want to confirm for themselves) and thus the smallest containing $\mathscr{T}$. □

Definition 1.1.3. A *Borel set* $B \in \mathscr{B}(X)$ is an element of the smallest σ-algebra on X, generated by a specified topology on X.

Note that we will mainly be dealing with $\mathscr{B}(\mathbb{R}^d)$ where we will take the usual order topology on $\mathbb{R}^d$. In the case of $\mathbb{R}$ we may generate the Borel sets by sets of the form $(a, b]$ or (a, b) or even $(-\infty, a)$. These will all generate the same σ-algebra due to properties (ii) and (iii) of a σ-algebra.

We wish to somehow assign a likelihood to each event. To do so we must define a map on the σ-algebra to the reals.

Definition 1.1.4. A *measure* on a measurable space $(\Omega, \mathscr{F})$ is a function $\mu : \mathscr{F} \to \overline{\mathbb{R}}_+$ such that if $A_1, A_2, ...$ are disjoint elements of $\mathscr{F}$ then [1]

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i).$$

[1] We do not exclude the possibility that some sets may have an infinite measure.

A *finite measure* is a measure μ such that $\mu(\Omega) < \infty$ and a *σ-finite measure* is a measure μ such that for each $\{\Omega_n\}_{n=1}^{\infty}$ with $\Omega_n \uparrow \Omega$, we have $\mu(\Omega_n) < \infty$ for each $n \in \mathbb{N}$.

A *probability measure* $\mathbb{P}$ is a measure with $\mathbb{P}(\Omega) = 1$.

Definition 1.1.5. A *measure space* $(\Omega, \mathscr{F}, \mu)$ is a measurable space $(\Omega, \mathscr{F})$ with a measure μ defined on it.

A *probability space* $(\Omega, \mathscr{F}, \mathbb{P})$ is a measurable space $(\Omega, \mathscr{F})$ with a probability measure $\mathbb{P}$ defined on it.

A μ-null set of a measure space is a set $A \in \mathscr{F}$ such that $\mu(A) = 0$. We will sometimes call it null sets where the measure is obvious from the context.

In a measure space a property holds *almost everywhere* if the points in which a property does not hold are the μ-null sets. In probability spaces this is also known as *almost surely* which is the same statement as saying the event happens with probability one.

Definition 1.1.6. We say that $A, B \in \mathscr{F}$ are independent on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ if $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$.

Now we look at a basic theorem about measures.

Theorem 1 (Monotone Convergence Theorem for Measures)**.** *Suppose that $(\Omega, \mathscr{F}, \mu)$ is a measure space and $\{B_n\}_{n=1}^{\infty} \subset \mathscr{F}$ is a sequence of sets that converge to B, then*

$$\mu(B) = \lim_{n\to\infty} \mu(B_n).$$

The term we shall use is infinitely often, abbreviated to i.o. This is a shorthand way of saying $\limsup$, i.e. A_n i.o $= \limsup_n A_n$. The reason for this terminology is that an element of the $\limsup$ must occur in infinitely many sets of A_n.

Using the Monotone Convergence Theorem, we will prove a very important theorem. This will be in heavy use in dealing with Brownian motion when we prove things to do with limits.

Theorem 2 (Borel-Cantelli Lemma)**.** *On a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ let $A_1, A_2, ... \in \mathscr{F}$ then if $\displaystyle\sum_{n=1}^{\infty} \mathbb{P}(A_n) < \infty$ then $\mathbb{P}(\limsup_n A_n) = 0$.*

Proof. Notice that $\limsup A_n = \cap_{i=1}^{\infty} \cup_{n=i}^{\infty} A_n$. Define $B_i = \cup_{n=i}^{\infty} A_n$. Now from the subadditivity of the measure we have that $\mathbb{P}(B_i) \leqslant \sum_{n=i}^{\infty} \mathbb{P}(A_n)$. By the assumption $\sum_{n=1}^{\infty} \mathbb{P}(A_n) < \infty$ therefore $\mathbb{P}(B_i) \to 0$ as $i \to \infty$. Hence as $n \to \infty$, $\mathbb{P}(\cap_{i=1}^{n} B_i) \to \mathbb{P}(\cap_{i=1}^{\infty} B_i) = 0$ by the Monotone Convergence Theorem. □

Now that we have a framework for probability, we need to look at more interesting things than just events. The following is a formal definition of a random variable.

Definition 1.1.7. A function f is said to be *measurable* if $f : \Omega \to Y$ where $(\Omega, \mathscr{F})$ is a measurable space, Y is a topological space and for any open set $U \subset Y$ we have that $f^{-1}(U) \in \mathscr{F}$.

Definition 1.1.8. A *random variable* X on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ is a measurable function.

Note that this is a very general definition. In all the cases, the random variables will be $\mathbb{R}^d$ valued, that is they will map to $\mathbb{R}^d$ with the usual topology. Measurability is an important concept as this allows us to assign probabilities to random variables.

Measurability is not really as strong as we like. Sets such as $(a, b]$ are not open in $\mathbb{R}$, hence we do not know if the pre-image of these are in the σ-algebra. The next definition will become very useful for us.

Definition 1.1.9. A function is said to be *Borel measurable* if $f : \Omega \to Y$ where $(\Omega, \mathscr{F})$ is a measurable space, Y is a topological space and for any Borel set $B \in \mathscr{B}(Y)$ we have that $f^{-1}(B) \in \mathscr{F}$.

We will always be assuming our random variables are Borel measurable.

Notice that a random variable X acts as a measure on $(\mathbb{R}^d, \mathscr{B}(\mathbb{R}^d))$ by the composition $\mu \circ X^{-1}$ as $X^{-1} : \mathscr{B}(\mathbb{R}^d) \to \mathscr{F}$ and $\mu : \mathscr{F} \to \mathbb{R}$. This is known as the *distribution* or *law* of X.

Now we introduce some probabilistic abuses of notation which usually is the most confusing part of probability. For a random variable X, $\mathbb{P}(X \in B)$ is shorthand for $\mathbb{P}(X^{-1}(B))$ where $B \in \mathscr{B}(\mathbb{R}^d)$. The distribution unless otherwise specified will be denoted by $\mathbb{P}(X \in dx)$.

The following are some examples of some important random variables. These will play an important role later on so it is essential to become familiar with them.

Example 1.1.10. An $\mathbb{R}^d$ valued *Normal* or *Gaussian* random variable[2] X on $(\Omega, \mathscr{F}, \mathbb{P})$ has a distribution of the form

$$\mathbb{P}(X \in dx) = \frac{1}{\sqrt{(2\pi)^{d/2}|\Sigma|^{\frac{1}{2}}}} \exp\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)\right) dx$$

where $\mu \in \mathbb{R}^d$ and Σ is a positive definite real $d \times d$ matrix. It is denoted $\mathscr{N}_d(\mu, \Sigma)$

In the case of $\mathbb{R}$ (which we will be using) it is of the form,

$$\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx$$

where $\mu, \sigma \in \mathbb{R}$. This is denoted $N(\mu, \sigma^2)$.

We can also have discrete measure spaces which gives rise to discrete random variables.

Example 1.1.11. A *Poisson* random variable N is a discrete random variable on a discrete measure space $(\Omega, \mathscr{F}, \mathbb{P})$. It can be described by,

$$\mathbb{P}(N \in \{k\}) = \frac{e^{-\lambda}\lambda^k}{k!}$$

where $\lambda > 0$ and is called the parameter. The measure of the Poisson random variable is *atomic*, that is, it assigns values to singleton sets. A Poisson random variable with parameter λ is commonly denoted $Pois(\lambda)$

We can also collect together random variables to model how something is evolving with time. This yields the next definition.

Definition 1.1.12. A *stochastic process* is a family of random variables $\{X_t, t \in I\}$.

Examples of stochastic processes will be the main concern over the next few chapters of the paper.

1.2 Integration

We will brush over some integration theory, for a detailed outline the reader is referred to Ash (1972) or Billingsley (1979) which are two of the many books that deal with this subject. The next theorem will become useful later on when we look at integration over product spaces. The theorem will not be proved, a proof can be found in any modern probability or measure theory book.

[2]This is usually called the multivariate normal distribution

Theorem 3 (Fubini's Theorem)**.** *Suppose that* $(\Omega_1, \mathscr{F}_1, \mu_1)$ *and* $(\Omega_2, \mathscr{F}_2, \mu_2)$ *are measure spaces and define a* σ*-algebra on* $\Omega = \Omega_1 \times \Omega_2$ *by* $\mathscr{F} = \mathscr{F}_1 \otimes \mathscr{F}_2$ *and a measure by* $\mu = \mu_1 \otimes \mu_2$*. If* $f : \Omega \to \overline{\mathbb{R}}_+$ *is a measurable function then the function* $F : \Omega_1 \to \overline{\mathbb{R}}_+$ *defined by*

$$F(x) = \int_{\Omega_2} f(x,s)\mu_2(ds)$$

is a measurable function and

$$\int_\Omega f d\mu = \int_{\Omega_1}\int_{\Omega_2} f(x,y)\mu_2(dy)\mu_1(dx) = \int_{\Omega_2}\int_{\Omega_1} f(x,y)\mu_1(dx)\mu_2(dy).$$

Now we define some central operators on probability spaces.

Definition 1.2.1. An *expectation* of a random variable X on $\mathbb{R}^d$, denoted $\mathbb{E}[X]$ is defined by,

$$\mathbb{E}[X] = \int_{\mathbb{R}^d} x\,\mathbb{P}(X \in dx).$$

The *co-variance* of two random variables X, Y on $\mathbb{R}^d$ is defined as,

$$Cov(X,Y) = \mathbb{E}[(\mathbb{E}[X] - X)(\mathbb{E}[Y] - Y)].$$

The *variance* of X is $Var(X) = Cov(X,X)$.

Intuitively, expectation is what is usually referred to by people as average. Variance is the amount by which the random variable is spread around the mean. Low variance means that the spread is tight around the mean. Notice that $\mathbb{E}$ is a linear function and also if two random variables are independent, then they have zero covariance.

Example 1.2.2. A $N(\mu, \sigma^2)$ random variable X has $\mathbb{E}[X] = \mu$ and $Var(X) = \sigma^2$. Moreover if we have Y on the same space which is $N(0, \sigma'^2)$ and $\mu = 0$, then,

$$Cov(X,Y) = \mathbb{E}[XY] = \sigma^2 \wedge \sigma'^2.$$

An important property of the normal distribution is that two normal random variables are independent *if and only if* they have zero covariance.

1.3 Convergence

In probability we have three main modes of convergence for random variables.

Definition 1.3.1. Let $\{X\}_{n=1}^\infty$ be a sequence of random variables and X be an other random variable.

We say that X_n converges to X *almost surely* and denote $X_n \xrightarrow{a.s.} X$ if $\forall \epsilon > 0$

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

Convergence in *probability* denoted $X_n \xrightarrow{prob.} X$ is when for each $\epsilon > 0$ we have

$$\lim_{n\to\infty} \mathbb{P}(|X_n - X| > \epsilon) = 0$$

We write $X_n \xrightarrow{\mathcal{D}} X$ and say X_n converges to X in *distribution* if for each $B \in \mathscr{B}(\mathbb{R})$ with $\mathbb{P}(X \in \partial B) = 0$,

$$\lim_{n\to\infty} \mathbb{P}(X_n \in B) = \mathbb{P}(X \in B).$$

There is a subtle difference between almost sure convergence and convergence in probability, however almost sure convergence is a stronger statement than convergence in probability. The reader can verify that almost sure convergence implies convergence in probability which in turn implies convergence in distribution.

Now we define convergence of measures. This will play an important part in Chapter 3 where we discuss infinitely divisible measures.

Definition 1.3.2. Let $\{\mu_n\}_{n=1}^{\infty}$ be a sequence of measures on the same measure space $(\Omega, \mathscr{F})$, then we say that μ_n *converges weakly* to a measure μ if one of the following hold

(i) For all bounded and continuous functions $f : \Omega \to \mathbb{R}$

$$\lim_{n\to\infty} \int_\Omega f(x)\mu_n(dx) = \int_\Omega f(x)\mu(dx)$$

(ii) For each closed $F \subset \Omega$, $\limsup \mu_n(F) \leqslant \mu(F)$

(iii) For each open $U \subset \Omega$, $\liminf \mu_n(U) \geqslant \mu(U)$.

With the basic tools we have, we may begin to characterise probability spaces and random variables.

Characteristic Functions

" Pure mathematics is the world's best game. It is more absorbing than chess, more of a gamble than poker, and lasts longer than Monopoly. It's free. It can be played anywhere - Archimedes did it in a bathtub. "

-Richard J. Trudeau

2.1 Basic Properties

In this section we will be assuming that X is a random variable on $(\Omega, \mathscr{F}, \mathbb{P})$ (a probability space). The aim of this chapter is to give a basic introduction to characteristic functions. We shall not be proving most statements here. For a formal approach to this subject, we refer the reader to Lukacs (1970) or Moran (1984).

In mathematics, Fourier transforms can reduce complicated tasks into simpler ones. We can also use Fourier transforms on a distribution function to simplify the expression. For some random variables the distribution cannot be explicitly known whereas we can often know the characteristic function.

Definition 2.1.1. A *characteristic function* ψ of X is defined by,

$$\psi(\theta) = \int_{\mathbb{R}} e^{i\theta x}\mathbb{P}(X \in dx)$$

and the function $\log \psi$ is referred to as the *characteristic exponent* of X.

This next theorem will play an important role in this paper. It describes the basic properties of sequence of characteristic functions. We will be using this heavily in the forthcoming chapters so it is important to keep in mind the equivalences stated in this theorem.

Theorem 4 (Lévy Continuity Theorem)**.** *Let $\{X_n : n = 1, 2, ...\}$ be a sequence of random variables (not necessarily on the same probability space) with characteristic functions ψ_n. If $\psi_n \to \psi$ pointwise then the following are equivalent,*

(i) *ψ is the characteristic function of some random variable X*

(ii) *ψ is the characteristic function of X where $X_n \xrightarrow{\mathcal{D}} X$*

(iii) *ψ is continuous*

(iv) *ψ is continuous in some neighbourhood of 0.*

For the proof see Fristedt and Gray (1997).

To see which functions are characteristic functions we need a theorem from analysis.

Theorem 5 (Bochner)**.** *A function ψ is a characteristic function if and only if the following hold;*

(i) $\psi(0) = 1$

(ii) *$\psi(\theta)$ is continuous*

(iii) *for any $\{u_i\}_{i=1}^n \subset \mathbb{R}$ and $\{v_i\}_{i=1}^n \subset \mathbb{C}$, $\sum_{i=1}^{n}\sum_{j=1}^{n} \psi(u_i - u_j)v_i\bar{v_j} \geqslant 0$.*

A proof of this can be found in any modern analysis book. The version for characteristic functions can be found in, for example, (Moran, 1984, p. 273 Theorem 6.19).

2.2 Examples

We will be using characteristic functions in the next chapter to work on a general class of random variables. It is essential that we get familiar with some solid examples of characteristic functions beforehand. These will be of the most common random variables.

Example 2.2.1. The characteristic of a $N(\mu, \sigma^2)$ random variable X is

$$\begin{aligned}\psi(\theta) = \mathbb{E}[e^{i\theta X}] = \int_{\mathbb{R}} e^{i\theta x}\mathbb{P}(X \in dx) &= \int_{\mathbb{R}} e^{i\theta x}\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx \\ &= e^{i\theta\mu - \frac{\theta^2\sigma^2}{2}} \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-(\mu+i\sigma^2\theta))^2}{2\sigma^2}\right). \\ &= e^{i\theta\mu - \frac{\theta^2\sigma^2}{2}}\end{aligned}$$

Now we give an example of a discrete random variable, namely the Poisson random variable.

Example 2.2.2. The characteristic function of a Poisson random variable N with parameter λ is of the form,

$$\psi(\theta) = \sum_{k=0}^{\infty} \frac{e^{-\lambda}(\lambda e^{i\theta})^k}{k!} = e^{\lambda(e^{i\theta}-1)}.$$

Example 2.2.3. Suppose we have a sequence of i.i.d.[1] random variables $\{\xi_n\}_{n=1}^{\infty}$ with a common law, say F, and a Poisson random variable N with rate λ that is independent of this sequence of random variables. We can define a new random variable X by

$$X = \sum_{n=1}^{N} \xi_n.$$

To find its characteristic function, we will use the tower law which states that [2]

$$\mathbb{E}[\mathbb{E}[A|B]] = \mathbb{E}[A].$$

The proof of this is simple and is left as an exercise to the reader.

Now we can compute the characteristic function of X by

$$\mathbb{E}[e^{i\theta X}|N] = \mathbb{E}[e^{i\theta \sum_{n=1}^{N} \xi_n}|N] = \mathbb{E}[\prod_{n=1}^{N} e^{i\theta\xi_n}|N] = \prod_{n=1}^{N} \mathbb{E}[e^{i\theta\xi_n}] = \mathbb{E}[e^{i\theta\xi_n}]^N$$

and so we have that

$$\mathbb{E}[e^{i\theta X}] = \mathbb{E}[\mathbb{E}[e^{i\theta\xi_n}]^N].$$

Now we need to derive the probability generating function of the Poisson process in order to obtain an analytic expression for the characteristic function of X. The probability generating function is given by

$$\mathbb{E}[s^N] = \sum_{n=0}^{\infty} s^n \mathbb{P}(N=n) = \sum_{n=0}^{\infty} s^n \frac{e^{-\lambda}\lambda^n}{n!} = e^{-\lambda}\sum_{n=0}^{\infty}\frac{(s\lambda)^n}{n!} = e^{\lambda(s-1)}.$$

Hence the characteristic function of X is

$$E[e^{i\theta X}] = e^{\lambda(\mathbb{E}[e^{i\theta\xi_n}]-1)} = e^{\lambda\int_{\mathbb{R}}(e^{i\theta x}-1)F(dx)}.$$

This will play an important role later when we introduce the compound Poisson process.

[1] Independent, identically distributed, i.e. they all have the same distribution and are independent (pairwise).

[2] We have not yet defined conditional expectation, we will be dealing with this later. The expectation here can be defined as $\mathbb{E}[A|B] = \mathbb{E}[A\mathbf{1}_B]/\mathbb{P}(B)$.

Infinitely Divisible Random Variables

"If people do not believe that mathematics is simple, it is only because they do not realize how complicated life is. "

- John Louis von Neumann

3.1 Definitions

Many of the random variables we have met can be expressed as a sum of the same random variable with different parameters. This gives rise to a generalisation of these which are called infinite divisible random variables.

Definition 3.1.1. A random variable X is said to be *infinitely divisible* if for each $n \in \mathbb{N}$ there exists $\{X_i^{(n)}\}_{i=1}^n$ of i.i.d. random variables such that

$$X \stackrel{d}{=} X_1^{(n)} + X_2^{(n)} + ... + X_n^{(n)}$$

Alternatively one may define infinite divisibility on a measure μ of a random variable,

$$\mu = \lambda_{(n)}^n = \underbrace{\lambda_{(n)} * \lambda_{(n)} * ... * \lambda_{(n)}}_{n}$$

where $\lambda_{(n)}$ is the law of some random variable.

Most distributions one encounters in every day life are infinitely divisible. This class of random variables cover a wide range of properties. A prominent example is of the normal random variable.

Example 3.1.2. A $N(\mu, \sigma^2)$ random variable X is infinitely divisible, with the distributions $X^{(n)} = N(\mu/n, \sigma^2/n)$. This is easily seen from the characteristic function of X,

$$\exp(i\theta\mu - \theta^2\sigma^2/2) = \exp(i\theta\mu/n - \theta^2\sigma^2/2n)^n.$$

3.2 Properties

As it turns out, infinitely divisible random variables can be represented in an elegant way through their characteristic functions. The aim of this section is to establish this result. We will be combining the approaches of Moran (1984) and Lukacs (1970).

We approach the problem in the same manner as Paul Lévy did in 1934. The construction is done via a sequence of Poisson like random variables which will limit to give the characteristic function of infinitely divisible random variables

To obtain this results we wish to write the characteristic function of an infinitely divisible distribution as $\psi = e^{\log \psi}$ which is valid as long as ψ is not zero (as the log function is not defined at zero). After this we can go on to find that the characteristic function will be the limit of $e^{k(\psi^{\frac{1}{k}}-1)}$ by the definition of the logarithm.

Proposition 3.2.1. *The characteristic function of an infinitely divisible random variable has no zeros.*

Proof. Let ψ be the characteristic function of an infinitely divisible random variable, then for each $n \in \mathbb{N}$ we can find a characteristic function ψ_n such that $\psi^{\frac{1}{n}} = \psi_n$.

Now consider

$$\phi(t) = \lim_{n\to\infty} \psi(t)^{\frac{1}{n}} = \begin{cases} 1 & \text{if } \psi(t) \neq 0 \\ 0 & \text{if } \psi(t) = 0 \end{cases}$$

Now as $\psi(0) = 1$ and ψ is continuous, we use the Lévy continuity theorem (Theorem 4) to conclude that ψ is continuous in a neighborhood of 0. $\phi(0) = 1$ and thus ϕ must be continuous in a neighborhood of 0. Applying the Lévy continuity theorem again we conclude that ϕ is continuous and thus $\phi = 1$, hence ψ has no zeros. □

Now we can go on to a theorem that will be very crucial in proving the main result.

Theorem 6. *The characteristic function that is the limit of characteristic functions of infinitely divisible processes is infinitely divisible.*

Proof. Let $\{f^{(n)}\}$ be a sequence of infinitely divisible characteristic functions that converge to a characteristic function f. Then for each $n, k \in \mathbb{N}$ there exists $f_k^{(n)}$ s.t. $(f_k^{(n)})^k = f^{(n)}$. Now as each $f_k^{(n)}$ is also infinitely divisible and so by Proposition 3.2.1 has no zeros. Hence we may infer that

$$f_k^{(n)} = e^{\frac{1}{k}\log f^{(n)}}.$$

So we have that $\lim_{n\to\infty} f_k^{(n)} = e^{\frac{1}{k}\log f} = f^{\frac{1}{k}}$ is a the limit of characteristic functions and as f is continuous, so is $f^{\frac{1}{k}}$ so the Lévy Continuity Theorem tells us that $f^{\frac{1}{k}}$ is a characteristic function. Thus f is infinitely divisible. □

Theorem 7 (De Finetti's Theorem)**.** *The characteristic function ψ of a random variable is infinitely divisible if and only if*

$$\psi(\theta) = \lim_{n\to\infty} e^{p_n(g_n(\theta)-1)}$$

for some $p_n \geqslant 0$ and g_n, where g_n are characteristic functions.

Proof. Suppose that $\psi(\theta)$ is infinitely divisible. Let $p_n = n$ and $g_n = \psi^{\frac{1}{n}}$, then as ψ has no zeros (by Proposition 3.2.1) and it follows that

$$\psi(\theta) = e^{\log\psi(\theta)} = \lim_{n\to\infty} e^{n(\psi(\theta)^{\frac{1}{n}}-1)} = \lim_{n\to\infty} e^{p_n(g_n(\theta)-1)}.$$

Now suppose that $\psi(\theta) = \lim_{n\to\infty} e^{p_n(g_n(\theta)-1)}$, then for each $q > 0$

$$f(x) = \left(1 + \frac{p_n(g_n(\theta)-1)}{n}\right)^q.$$

is continuous and positive definite for each $n \in \mathbb{N}$, and $f(0) = 1$ so we may apply Bochner's Theorem (Theorem 5) to conclude that it is a characteristic function. Hence we have that

$$\left(1 + \frac{p_n(g_n(\theta)-1)}{n}\right)^n$$

is a characteristic function of an infinitely divisible distribution for each $n \in \mathbb{N}$.

Hence passing to the limit as n tends to infinity gives

$$\psi(\theta) = \lim_{n\to\infty} \left(1 + \frac{p_n(g_n(\theta)-1)}{n}\right)^n = \lim_{n\to\infty} e^{p_n(g_n(\theta)-1)}$$

is infinitely divisible by Lévy Continuity Theorem (Theorem 4) and Theorem 6. □

Corollary 3.2.2. *A characteristic function is of an infinitely divisible distribution if and only if it is the limit of Poisson like characteristic functions.*

Proof. From De Finetti's Theorem we have that ψ is the characteristic function of an infinitely divisible distribution if and only if

$$\psi(\theta) = \lim_{n\to\infty} e^{p_n(g_n(\theta)-1)} = \lim_{n\to\infty} e^{p_n \int_{\mathbb{R}}(e^{i\theta x}-1)G_n(dx)}$$

□

Now we can prove the main result. This will later prove to be very useful in dealing with stochastic processes that have infinitely divisible distributions.

Theorem 8 (Lévy canonical representation)**.** *A characteristic function ψ is of an infinitely divisible distribution if and only if it is of the form*

$$\log\psi(\theta) = a\theta i - \frac{1}{2}\sigma^2\theta^2 + \int_{-\infty}^{-0}\left(e^{i\theta x} - 1 - \frac{i\theta x}{1+x^2}\right)M(dx) + \int_{+0}^{\infty}\left(e^{i\theta x} - 1 - \frac{i\theta x}{1+x^2}\right)N(dx) \tag{3.2.1}$$

where M, N are non decreasing on the intervals $(-\infty, 0)$ and $(0,\infty)$ respectively such that,

$$\lim_{x\to\infty}\int_{-\infty}^{-x} M(dx) = \lim_{x\to\infty}\int_{-\infty}^{x} N(dx) = 0 \tag{3.2.2}$$

and

$$\forall\epsilon > 0 \quad \int_{-\epsilon}^{0} x^2 M(dx) < \infty \quad \int_{0}^{\epsilon} x^2 N(dx) < \infty. \tag{3.2.3}$$

Proof. We first prove the necessity of the condition. Let ψ be the characteristic function of an infinite divisible distribution then by Corollary 3.2.2 we have that,

$$\psi(\theta) = \lim_{n\to\infty} e^{n\int_{\mathbb{R}}(e^{i\theta x}-1)G_n(dx)}$$

where G_n is the measure of a Poisson like random variable.

Define $\psi_n(\theta) = e^{n\int_{\mathbb{R}}(e^{i\theta x}-1)G_n(dx)}$ and so

$$\log\psi_n(\theta) = n\int_{\mathbb{R}}(e^{i\theta x} - 1)G_n(dx) = ia_n\theta + \int_{\mathbb{R}}\left(e^{i\theta x} - 1 - \frac{i\theta x}{1+x^2}\right)\frac{1+x^2}{x^2}H_n(dx) \tag{3.2.4}$$

where

$$a_n = n\int_{\mathbb{R}}\frac{x}{1+x^2}G_n(dx)$$

and

$$H_n(dx) = n\frac{x^2}{1+x^2}G_n(dx).$$

Clearly H_n does not blow up as $x^2/(1+x^2)$ bounded on $\pm\infty$. Now we need to show that $a_n \to a$ and $H_n \to H$ weakly.

Define

$$\begin{aligned}
\lambda_n(\theta) &= \int_0^1 \log\psi_n(\theta) - \frac{\log\psi_n(\theta+h) + \log\psi_n(\theta-h)}{2}dh \\
&= \int_0^1 \int_{\mathbb{R}}\left(e^{i\theta x} - \frac{e^{i(\theta+h)x}}{2} - \frac{e^{i(\theta-h)x}}{2}\right)\frac{1+x^2}{x^2}H_n(dx) \quad dh \\
&= \int_0^1 \int_{\mathbb{R}} e^{i\theta x}\left(1 - \frac{e^{ihx}+e^{-ihx}}{2}\right)\frac{1+x^2}{x^2}H_n(dx) \quad dh \\
&= \int_0^1 \int_{\mathbb{R}} e^{i\theta x}(1-\cos xh)\frac{1+x^2}{x^2}H_n(dx) \quad dh.
\end{aligned} \tag{3.2.5}$$

Now we can reverse the order of the integration using Fubini's Theorem to get,

$$\lambda_n(\theta) = \int_{\mathbb{R}} e^{i\theta x} \left(1 - \frac{\sin x}{x}\right) \frac{1+x^2}{x^2} H_n(dx) = \int_{\mathbb{R}} e^{i\theta x} R_n(dx)$$

where,

$$R_n(dx) = \left(1 - \frac{\sin x}{x}\right) \frac{1+x^2}{x^2} H_n(dx).$$

As ψ is continuous thus we have that λ_n converges to a continuous function. Hence we can conclude[1] that R_n converges weakly to a bounded and non-decreasing function R, that is to say for all $f \in C_\#$, [2]

$$\lim_{n\to\infty} \int_{\mathbb{R}} f(x) dR_n(x) = \int_{\mathbb{R}} f(x) R(dx).$$

In particular for each $g \in C_\#$, $\left(1 - \frac{\sin x}{x}\right)^{-1} \frac{x^2}{1+x^2} g(x)$ is continuous and also vanishes at ∞ and $-\infty$.[3] Hence, H_n converges weakly to some distribution function H and by the same argument nG_n converges weakly to some G. Thus we have that,

$$\int_{\mathbb{R}} \left(e^{i\theta x} - 1 - \frac{i\theta x}{1+x^2}\right) \frac{1+x^2}{x^2} H_n(dx) \to \int_{\mathbb{R}} \left(e^{i\theta x} - 1 - \frac{i\theta x}{1+x^2}\right) \frac{1+x^2}{x^2} H(dx)$$

and

$$a_n = n \int_{\mathbb{R}} \frac{x}{1+x^2} G_n(dx) \to a.$$

Thus we have proved that any infinitely divisible distribution has a characteristic function of the form,

$$\log \psi(\theta) = ia\theta + \int_{\mathbb{R}} \left(e^{i\theta x} - 1 - \frac{i\theta x}{1+x^2}\right) \frac{1+x^2}{x^2} H(dx) \tag{3.2.6}$$

which is known as the *Lévy-Khintchine canonical representation.* Now we can define,

$$\begin{aligned} \sigma^2 &= H(+0) - H(-0) \\ M(x) &= \int_{-\infty}^{x} \frac{1+x^2}{x^2} H(dx) \quad x < 0 \\ N(x) &= -\int_{x}^{\infty} \frac{1+x^2}{x^2} H(dx) \quad x > 0. \end{aligned}$$

This satisfies (3.2.1),(3.2.2) and (3.2.3).

The sufficiency is an application of Corollary 3.2.2 to (3.2.4). □

The representation in the last theorem are unique up to distribution by the property of the Fourier transform. We will for now leave this as it is and return to it in Chapter 6 where we will be talking about Lévy processes.

[1] For proof of this see (Moran, 1984, p.252 Theorem 6.3)

[2] $C_\#$ are the set of $\mathbb{R}$-valued continuous functions that vanish at ∞ and $-\infty$

[3] Notice that $(1 - \sin x/x) \to 1$ and $x^2/(1+x^2) \to 1$ as x tends to $\pm\infty$. Thus the function defined vanishes at $\pm\infty$ as g vanishes.

Brownian Motion

" One cannot escape the feeling that these mathematical formulas have an independent existence and an intelligence of their own, that they are wiser than we are, wiser even than their discoverers... "

-Heinrich Hertz

4.1 Definition and Construction

Brownian motion is one of the most interesting stochastic processes around. It posses various amounts of properties and hence has been the focus of study for a long time. The idea of Brownian motion, sometimes known as the Wiener process, is modeled after the physical phenomena of a smoke particle moving about in air. It was Brown that discovered that it was the air particles that produced this seemingly random motion and Norbert Wiener that mathematically formalised it.

Imagine a particle of smoke being bombarded by particles of air. The seemingly random motion that this smoke particle exhibits is called Brownian motion.

Definition 4.1.1 (Brownian Motion)**.** A stochastic process W_t is called a *Brownian motion* or a *Wiener process* if it satisfies the following properties,

(i) $W_0 = 0$ almost surely

(ii) for $0 \leqslant t_1 \leqslant ... \leqslant t_n$, $W_{t_k} - W_{t_{k-1}}, ..., W_{t_2} - W_{t_1}$ are independent

(iii) for $s < t$, $W_t - W_s$ is distributed $N(0, t-s)$

(iv) $t \mapsto W_t$ is continuous almost surely.

We will be proving that a Brownian motion does indeed exist and some of the basic properties it posses. It is useful to first work in the interval $[0,1]$ as anything proved in this will be easily extended to $\mathbb{R}^d$.

4.1.1 Interval $[0,1]$

Notice that we can reformulate *(ii)* and *(iii)* as, for $s < t$ $W_t - W_s$ is independent of $\{W_u : u \leqslant s\}$ and for each $t > 0$, W_t is distributed $N(0,t)$.

The proof for existence of Brownian motion without continuity is routine. This is done by using the Kolmogorov's extension theorem. Kolmogorov's extension theorem, also known as the consistency theorem, is a theorem that allows a finite dimensional distribution to be extended to a stochastic process. It is similar to that of Carathèodory's extension theorem.[1]

Suppose we have $0 < t_1 < ... < t_n$ and a measure $\mu_{t_1,...,t_n}$ on $(R^d)^n$. If for any permutation σ on $\{1,...,n\}$, $m \in \mathbb{N}$ and $A_1,...,A_n \subset \mathbb{R}^d$,

$$\mu_{t_{\sigma(1)},t_{\sigma(2)},...,t_{\sigma(n)}}(A_1 \times A_2 \times ... \times A_n) = \mu_{t_1,t_2,...,t_n}(A_{\sigma(1)} \times A_{\sigma(2)} \times ... \times A_{\sigma(n)})$$

and

$$\mu_{t_1,..,t_n,t_{n+1},..,t_{n+m}}(A_1 \times ... \times A_n \times \mathbb{R} \times ... \times \mathbb{R}) = \mu_{t_1,..,t_n}(A_1 \times ... \times A_n)$$

then μ will extend uniquely up to a measure on some space. This theorem gives us a tool for describing a stochastic process by its finite dimensional distributions. We will describe how to construct a Brownian motion using this theorem but it will be apparent why we need more.

[1]Carathèodory's extension theorem states that a countably additive measure on a ring of sets can be extended uniquely to a measure on the σ-algebra generated by this ring.

For $0 < t_1 < ... < t_n \leqslant 1$ we define a measure μ on $\mathbb{R}^n$ as follows,

$$\mu_{t_1,...,t_n}(A_1 \times ... \times A_n) = \int_{A_1} dx_1 ... \int_{A_n} dx_n \prod_{i=1}^{n} \left(\frac{1}{\sqrt{2\pi(t_i - t_{i-1})}} \exp\left\{ -\frac{(x_i - x_{i-1})^2}{(t_i - t_{i-1})} \right\} \right).$$

The assumptions of the Kolmogorov's extension theorem can easily be verified in this instance. Therefore this extends to give a measure on $[0, 1]$. Kolmogorov's extension theorem does not guarantee the continuity of the paths. It is not at all obvious why we need the continuity in the definition. To see the importance of this, consider the following example.

Example 4.1.2. Let us for a second assume that Brownian motion exists on $[0, 1]$ call this B_t. Let U be a uniform random variable on $[0, 1]$ independent of B_t, now we define a new process B'_t by

$$B'_t = \begin{cases} B_t & \text{if} \quad t \neq U \\ 0 & \text{otherwise} \end{cases}$$

One can see that this function satisfies the first three properties of Brownian motion but it is almost surely discontinuous.

Now assured that our efforts to prove continuity are not in vain, we may continue. We will approach the problem in a much similar manner to that of Norbert Wiener. Alternative approach via the Polish space[2] $C([0, \infty), \mathbb{R}^d)$ is given in Strook and Varadhan (1979).

We will do a direct construction using the lemma below. The construction will be of a sequence of stochastic processes which are almost surely convergent uniformly.

Lemma 4.1.3. *The uniform limit of a sequence of continuous functions is continuous.*

The proof of this is just an application of the definitions which we leave to the unsure reader as an exercise.

Next we introduce *Haar functions* $\{f_0, f_{k,n}, k = 1, ..., 2^{n-1} \quad n = 1, 2, ...\}$, which form a basis for $L^2[0, 1]$.

$$f_0(t) = 1$$
$$f_{k,n}(t) = \begin{cases} 2^{\frac{n-1}{2}} & t \in \left[\frac{k-1}{2^n}, \frac{k}{2^n}\right) \\ -2^{\frac{n-1}{2}} & t \in \left[\frac{k}{2^n}, \frac{k+1}{2^n}\right) \\ 0 & \text{otherwise} \end{cases} \tag{4.1.1}$$

These are more useful for us when they are integrated indefinitely. These are called *Schauder functions* and are given by,

$$F_0(t) = t$$
$$F_{k,n}(t) = \begin{cases} 2^{\frac{n-1}{2}} \left(t - \frac{k-1}{2^n}\right) & t \in \left[\frac{k-1}{2^n}, \frac{k}{2^n}\right) \\ 2^{\frac{n-1}{2}} \left(\frac{k+1}{2^n} - t\right) & t \in \left[\frac{k}{2^n}, \frac{k+1}{2^n}\right) \\ 0 & \text{otherwise} \end{cases} \tag{4.1.2}$$

Now equipped with the Schauder functions, we may begin constructing a Brownian motion. Let $\{X_0, X_{k,n}, k = 1, ..., 2^{n-1} \quad n = 1, 2, ...\}$ be i.i.d. $N(0, 1)$ random variables. We define a sequence $\{W^n_t\}$ on $[0, 1]$ by, [3]

[2] Complete separable metric space

[3] We will be changing between notation of W^n_t, by referring to its value for a particular $\omega \in \Omega$ by $W^n(t, \omega)$

$$W^n(t,\omega) = X_0F_0(t) + \sum_{i=1}^{n} Y_i(t,\omega) \tag{4.1.3}$$

where,

$$Y_i(t,\omega) = \sum_{j=1}^{2^{i-1}} X_{j,i}(\omega)F_{j,i}(t). \tag{4.1.4}$$

In layman's terms, we construct W^n by picking dyadic rationals and assigning a normal distribution $N\left(0, \frac{k}{2^n}\right)$ for each k. Notice that the sum means that there is some dependence, but an avid reader may observe this will satisfy the independent increments on the dyadic rationals which will extend to the whole line $[0,1]$ when we take the limit.

It is immediately obvious from equations (4.1.4) and (4.1.2) that each W^n is continuous for all $\omega \in \Omega$. The following theorem will provide the most useful information.

Theorem 9. *The sequence of $\{W^n\}$ defined above almost surely converges uniformly to a stochastic process W.*

Proof. First we see from (4.1.3) that for each $\omega \in \Omega$, we first need to analyse $Y_n(t,\omega)$ for each $n \in \mathbb{N}$. Notice that by (4.1.4) and the fact that $F_{k,n}$ is maximum at $t = \frac{k}{2^n}$,

$$\max_{t\in[0,1]} |Y_n(t,\omega)| = 2^{\frac{-n-1}{2}} \max_{1\leqslant i\leqslant 2^{n-1}} |X_{i,n}| \tag{4.1.5}$$

So for each $c_n \in \mathbb{R}$

$$\begin{aligned}
\mathbb{P}(\max_{t\in[0,1]} |Y_n(t,\omega)| > 2^{\frac{-n-1}{2}} c_n) &= \mathbb{P}(\max_{1\leqslant i\leqslant 2^{n-1}} |X_{i,n}| > c_n) \\
&\leqslant \sum_{i=1}^{2^{n-1}} \mathbb{P}(|X_{i,n}| > c_n) \\
&= 2\sum_{i=1}^{2^{n-1}} \mathbb{P}(X_{i,n} > c_n) \quad \text{by symmetry} \\
&= \sum_{i=1}^{2^{n-1}} \frac{2}{\sqrt{2\pi}} \int_{c_n}^{\infty} e^{-\frac{x^2}{2}}\,dx \\
&\leqslant \sum_{i=1}^{2^{n-1}} \frac{2}{\sqrt{2\pi}} \int_{c_n}^{\infty} \frac{x}{c_n} e^{-\frac{x^2}{2}}\,dx \\
&= \frac{2^n}{c_n\sqrt{2\pi}} e^{-\frac{c_n^2}{2}}.
\end{aligned}$$

Now picking $c_n = 2\sqrt{2n\log n}$ gives,

$$\mathbb{P}(\max_{t\in[0,1]} |Y_n(t,\omega)| > 2^{\frac{1-n}{2}} \sqrt{2n\log 2}) \leqslant C\frac{2^{-3n}}{\sqrt{n}} \tag{4.1.6}$$

for some constant C.

Notice that applying the ratio test to $\frac{2^{-3n}}{\sqrt{n}}$,

$$\frac{2^{-3n-3}\sqrt{n}}{2^{-3n}\sqrt{n+1}} \leqslant 2^{-3} < 1$$

gives that

$$\sum_{n=1}^{\infty} \mathbb{P}(\max_{t\in[0,1]} |Y_n(t,\omega)| > 2^{\frac{1-n}{2}}\sqrt{2n\log 2}) \leqslant \sum_{n=1}^{\infty} C\frac{2^{-3n}}{\sqrt{n}} < \infty.$$

Application of the Borel-Cantelli Lemma shows,

$$\mathbb{P}(\max_{t\in[0,1]} |Y_n(t,\omega)| > 2^{\frac{1-n}{2}}\sqrt{2n\log 2} \quad \text{i.o.}) = 0. \tag{4.1.7}$$

Let $A_n = \{\omega \in \Omega : \max_{t\in[0,1]} |Y_n(t,\omega)| > 2^{\frac{1-n}{2}}\sqrt{2n\log 2}\}$ then the statement above means that

$$\mathbb{P}(\omega \in A_n \quad \text{i.o.}) = 0.$$

Denote $\limsup_n A_n = A$ and notice that $2^{\frac{1-n}{2}}\sqrt{2n\log 2} \to 0$ as $n \to \infty$. So we have that if $\omega \notin A$ then by the definition of A, $W^n(t,\omega)$ is uniformly convergent. If $\omega \in A$ then we have that with probability one, ω is in finitely many A_n. Hence with probability one we can pick an $N \in \mathbb{N}$ such that,

$$\max_{t\in[0,1]} |Y_n(t,\omega)| \leqslant 2^{\frac{1-n}{2}}\sqrt{2n\log 2} \quad \text{for } n > N.$$

Hence we have that Y_n is almost surely convergent.

Now we apply the ratio test once again,

$$\frac{2^{-\frac{n}{2}}\sqrt{2n\log n + 2\log 2}}{2^{\frac{1}{2}}2^{-\frac{n}{2}}\sqrt{2n\log 2}} \to 2^{-\frac{1}{2}} < 1$$

gives that $\sum_{n=1}^{\infty} Y_n(t,\omega) < \infty$ almost surely, which completes the proof. □

There is still the outstanding issue of what to do with the set of ω that do not converge. We can set these points to 0 as these are the $\mathbb{P}$-null sets, it will not effect the probabilities nor any of the properties that the limiting stochastic process has.

Now we go on to prove that this W is actually a Brownian motion.

Theorem 10. *The process W given above is a Brownian motion.*

Proof. We will be checking the conditions given by the definition of Brownian motion at the start of the chapter.

(i) Notice that for each $n \in \mathbb{N}$ we have $W^n(0,\omega) = 0$, hence $W_0 = 0$ almost surely.

(ii) The independence of the increments follow directly from the construction. The sequence of normal random variables are independent.

(iii) Let us relabel the Haar functions defined on (4.1.1) as $\{f_i, i = 1, 2, ...\}$ for convenience. Define

$$I_t(s) = \begin{cases} 1 & \text{if } s < t \\ 0 & \text{otherwise} \end{cases}$$

As the Haar functions form a complete orthonormal basis over L^2 we have that,

$$I_t = \sum_{i=1}^{\infty} \langle I_t, f_i\rangle f_i.$$

Now we have that

$$t = ||I_t||^2 = \langle I_t, I_t\rangle = \left\langle I_t, \sum_{i=1}^{\infty} \langle I_t, f_i\rangle f_i \right\rangle = \sum_{i=1}^{\infty} \langle I_t, f_i\rangle^2,$$

We can use the fact that the $\{X_0, X_{l,k}\}$ defined by (4.1.4) are i.i.d. to observe that for each $n \in \mathbb{N}$, W_t^n is distributed $N(0,1)$ times the sum

$$\sum_{i=1}^{k(n)} \int_0^t f_i(x)dx = \sum_{i=1}^{k(n)} \langle f_i, I_t \rangle$$

where $k(n)$ is an increasing function of n such that $k(0) = 0$.[4] Thus each W_t^n is distributed $N(0, \sum_{i=1}^{k(n)} \langle f_i, I_t \rangle^2)$ and thus has characteristic function of the form $\exp(-\frac{\sum_{i=1}^{k(n)} \langle f_i, I_t \rangle^2}{2})$.[5] This converges to $\exp(-\frac{\sum_{i=1}^{\infty} \langle f_i, I_t \rangle^2}{2}) = \exp(-\frac{t}{2})$. Hence by Lévy Continuity Theorem, W_t is distributed $N(0,t)$ and coupled with *(i)* and *(ii)* gives the result.

(iv) Continuity follows from Theorem 9. □

4.1.2 Extension to $[0, \infty)^d$

First let us extend the Brownian motion to $[0, \infty)$.

Theorem 11. *Let W_t, W_t' be two independent copies of a Brownian motion on $[0,1]$, then B_t defined as*

$$B_t = \begin{cases} W_t & \text{if } t \in [0,1] \\ tW'_{1/t} & \text{if } t \in (1,\infty) \end{cases} \tag{4.1.8}$$

is a Brownian motion on $[0, \infty)$.

Proof. It is easy to see that on $[0,1]$, B_t is a Brownian motion. Now *(i)*, of the definition of Brownian motion, is satisfied by this property.

For *(ii)* we prove that for $s < t$, $W_t - W_s$ is independent of $\{W_u : u \leqslant s\}$.

Case $s, t \in [0,1]$: Holds by the construction.

Case $s, t, \in (1, \infty)$: Consider $Cov(B_t - B_s, B_u) = \mathbb{E}[(\mathbb{E}[B_t] - \mathbb{E}[B_s] - (B_t - B_s))(\mathbb{E}[B_u] - B_u)] = \mathbb{E}[(sW'_{1/s} - tW'_{1/t})uW'_{1/u}]$ for $u \in (1, \infty)$. Thus we have that

$$Cov(B_t - B_s, B_u) = su\mathbb{E}[W'_{1/s}W'_{1/u}] - tu\mathbb{E}[W'_{1/t}W'_{1/u}] = su\left(\frac{1}{s} \wedge \frac{1}{u}\right) - tu\left(\frac{1}{t} \wedge \frac{1}{u}\right) = 0.$$

Hence $B_t - B_s$ is independent of B_u. If $u \in [0,1]$ then the result is obvious.

Case $s \in [0,1]$, $t \in (1,\infty)$: Immediate from the construction (W and W' are independent).

(iii) follows from the fact that $tW'_{1/t}$ is distributed $N(0, t^2/t)$.

(iv) is immediately obvious, except at the point $t = 1$. Intuitively this should hold, but it does no harm to check that this is indeed true.

Define $A_n = \{\omega : |B_{1-\frac{1}{2n^2}} - B_{1+\frac{1}{2n^2}}| \geqslant c_n\}$, it suffices to show that $\mathbb{P}(A_n \text{ i.o.}) = 0$, as we can set all the null-sets that are not continuous to be 0. We shall apply the Borel-Cantelli Lemma, so first notice that

$$\begin{aligned}\mathbb{P}(A_n) &= \mathbb{P}\left(|B_{1-\frac{1}{2n^2}} - B_{1+\frac{1}{2n^2}}| \geqslant c_n\right) = \mathbb{P}\left(|B_{\frac{1}{n^2}}| \geqslant c_n\right) = 2\mathbb{P}\left(B_{\frac{1}{n^2}} \geqslant c_n\right) = 2\mathbb{P}(Z \geqslant nc_n) \\ &= \int_{nc_n}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \leqslant \int_{nc_n}^{\infty} \frac{1}{\sqrt{2\pi}} x e^{-\frac{x^2}{2}} dx = \frac{1}{\sqrt{2\pi}} e^{-n^2 c_n^2}.\end{aligned}$$

[4] This formalism is due to the fact that we have relabeled the Haar functions and hence we may be summing more than n of them.

[5] To see this let $\{X_i\}_{i=1}^{k(n)}$ be a sequence of normal random variables, then $\sum_{i=1}^{k(n)} X_i \langle f_i, I_t \rangle = X_1 \langle f_1, I_t \rangle + ... + X_{k(n)} \langle f_{k(n)}, I_t \rangle$, with each $X_i \langle f_i, I_t \rangle$ distributed $N(0, \langle f_i, I_t \rangle^2)$.

Now we pick $c_n = 1/\sqrt{n}$ and we have that A_n i.o. describes the sample paths that are discontinuous at 1. By noting that

$$\sum_{n=1}^{\infty} \mathbb{P}(A_n) \leqslant \sum_{n=1}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-n^2 c_n^2} = \frac{1}{\sqrt{2\pi}} \sum_{n=1}^{\infty} (e^{-1})^n < \infty$$

we may apply the Borel-Cantelli Lemma to deduce that $\mathbb{P}(A_n \text{ i.o.}) = 0$, hence B_t is almost surely continuous at $t = 1$. □

Now we can extend Brownian motion to $[0, \infty)^d$ in an obvious way.

Theorem 12. *Suppose* $\{B^{(i)}\}_{i=1}^d$ *is a sequence of independent Brownian motions on* $[0, \infty)$, *then* $\mathbf{B} = \{\mathbf{B_t} : t \geqslant 0\}$ *defined by*

$$\mathbf{B_t} = (B_t^{(1)}, ..., B_t^{(d)})$$

is a Brownian motion on $[0, \infty)^d$.

The proof of the theorem is rather tedious checking of the definition, which we shall leave out. Unsure reader is advised to check for themselves that this indeed does satisfy the definition of a Brownian motion.

4.2 Properties

Brownian motion has been subject to a considerable amount of study. The main reason for this is that it satisfies a lot of nice properties which we shall devote some time to in this section. Interested reader is referred to Rogers and Williams (1988) for a full study of the properties.

A corollary to Baire's category theorem states that the set of nowhere differentiable functions are dense in the continuous functions. The first function that was known to be nowhere differentiable but continuous was the Weierstrass function. Now we can give an example of an almost surely continuous but nowhere differentiable function.

Theorem 13. *Brownian motion is nowhere differentiable almost everywhere.*

Intuitively this is pretty clear. If we take a Newton quotient of Brownian motion $(W(t+h) - W(t))/h$, we find that a derivative at any point is distributed $N\left(0, \frac{1}{h}\right)$. As $h \downarrow 0$ this clearly does not converge. We will use the Borel-Cantelli lemma to verify this in what follows.

Proof. Notice that as W_t is distributed as $N(0, t)$, $\frac{W_{t+h} - W_t}{h}$ is distributed as $N\left(0, \frac{1}{h}\right)$ for each $h > 0$. Fix $t \geqslant 0$ and let

$$K_n := \left\{ \omega : \inf_{x \in \mathbb{R}} \left| \frac{W_{t+\frac{1}{n}}(\omega) - W_t(\omega)}{\frac{1}{n}} - x \right| \leqslant \frac{1}{n} \right\}.$$

For each $x \in \mathbb{R}$, $\left\{ \frac{W_{t+\frac{1}{n}}(\omega) - W_t(\omega)}{\frac{1}{n}} - x \right\}$ has the distribution $N(-x, n)$. Thus for each $x, c_n \in \mathbb{R}$, we have

$$\begin{aligned} \mathbb{P}\left(\left| \frac{W_{t+\frac{1}{n}}(\omega) - W_t(\omega)}{\frac{1}{n}} - x \right| \leqslant c_n \right) &= 2 \int_0^{c_n} \frac{1}{\sqrt{2\pi n}} \exp\left(-\frac{(s+x)^2}{2n} \right) ds \\ &\leqslant \frac{2c_n}{\sqrt{2\pi n}}. \end{aligned}$$

An obvious choice of c_n is $\frac{1}{\sqrt{n^3}}$ which gives

$$\sum_{n=1}^{\infty} \mathbb{P}\left(\left|\frac{W_{t+\frac{1}{n}}(\omega) - W_t(\omega)}{\frac{1}{n}} - x\right| \leqslant \frac{1}{\sqrt{n^3}}\right) \leqslant \frac{2}{\sqrt{2\pi}} \sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{4}{\sqrt{2\pi}}.$$

Note that the above holds for each $x \in \mathbb{R}$ and in particular $\displaystyle\sum_{n=1}^{\infty} \mathbb{P}(K_n) \leqslant \frac{4}{\sqrt{2\pi}}$. So by the Borel-Cantelli Lemma

$$\mathbb{P}\left(\inf_{x\in\mathbb{R}} \left|\frac{W_{t+\frac{1}{n}}(\omega) - W_t(\omega)}{\frac{1}{n}} - x\right| \leqslant \frac{1}{n} \text{ i.o.}\right) = 0.$$

Hence W_t is almost surely not differentiable at t. Our choice of t was arbitrary. □

While proving the existence of Brownian motion, we came across some properties that are listed below.

Theorem 14. *Brownian motion is time invertible. That is,* $B = \{tW_{1/t}, t \in [0, \infty)\}$ *(where* $B_0 = 0$*) is a Brownian motion.*

Proof. See Theorem 11. □

Theorem 15. *Brownian motion is time reversible. That is for any fixed* $T > 0$*,* $B = \{W_T - W_{T-t}, t \in [0, T]\}$ *is a Brownian motion.*

Proof. Let W_t be a Brownian motion on $[0, T]$, then we show that $B_t = W_T - W_{T-t}$ is a Brownian motion on $[0, T]$. $B_t = 0$ a.s. so *(i)* is satisfied.

Suppose that $0 \leqslant s < t \leqslant T$, then $B_t - B_s = W_T - W_{T-t} - W_T - W_{T-s} = W_{T-s} - W_{T-t}$. As $s < t$, $W_{T-s} - W_{T-t}$ is independent of all W_{T-u} such that $T - u \leqslant T - t$, i.e. of all B_u. This proves *(ii)*.

(iii) follows from the properties of normal distributions and *(iv)* follows from the construction. □

Poisson Processes

" We've all heard that a million monkeys banging on a million typewriters will eventually reproduce the entire works of Shakespeare. Now, thanks to the internet, we know this is not true. "

-Robert Silensky

5.1 Poisson Processes

In this chapter we will be working with an important class of stochastic processes, namely the Poisson processes. These have a lot of applications in finance and queuing theory.

The idea behind the Poisson process is to model arrivals with a Poisson random variable at some rate λ. Say we count the amount of cars at a street and the arrival rate of the cars is distributed as a Poisson distribution with parameter λ, then the number of cars N_t at time t would be a Poisson process. This leads to the following definition.

Definition 5.1.1. A *Poisson process* N_t is a stochastic process that satisfies the following,

(i) $N_0 = 0$ almost surely

(ii) for $0 \leqslant t_1 \leqslant ... \leqslant t_n$, $N_{t_k} - N_{t_{k-1}}, ..., N_{t_2} - N_{t_1}$ are independent

(iii) for $s < t$, $N_t - N_s$ is distributed $Pois((t-s)\lambda)$

(iv) $t \mapsto N_t$ is almost surely right continuous with left limits.

We may construct a Poisson process in the following way. Let $\{X_n\}_{n=0}^{\infty}$ be a random walk with X_1 distributed as an exponential of rate λ. Now we can define N_t as follows,

$$N_t(\omega) = n \quad \text{if and only if} \quad X_n(\omega) \leqslant t < X_{n+1}(\omega).$$

Intuitive argument to why this is a Poisson process is due to the holding times at each state being distributed exponentially with parameter λ. A rigorous proof of this can be found in Sato (1999).

We will now introduce the compound Poisson process. Carrying on with our example, suppose that we only measure the red cars passing by and that the distribution of the cars being red is ξ. If we let $\{\xi_i\}_{i=1}^{\infty}$ be an i.i.d. sequence of random variables, then clearly the number of cars that are red at time t is given by $\sum_{i=1}^{N_t} \xi_i$. This leads on to the definition of a compound Poisson process.

Definition 5.1.2. Let N_t be a Poisson process and $\{\xi_i\}_{i=1}^{\infty}$ be a set of i.i.d. random variables independent of N_t. Then a *compound Poisson process* Y_t is defined by,

$$Y_t = \sum_{i=1}^{N_t} \xi_i.$$

5.2 Poisson Measures

Recall that a Polish space E is a complete separable metric space. The aim of this section is to introduce the concept of Poisson measures. We will not be approaching this matter formally, as we will not be using it in the rest of the paper. A formal approach to this can be found in Kyprianou (2006) and a more relaxed approach can be found in Bertoin (1996).

Definition 5.2.1. Let E be a Polish space and ν a σ-finite measure on E. We call a random measure K a *Poisson measure* with intensity ν if for each disjoint $B_1, ...B_n \subset \mathscr{B}(E)$, $K(B_1), ..., K(B_n)$ are independent and have Poisson distributions with parameters $\nu(B_1), ..., \nu(B_n)$ respectively.

We can construct the Poisson measure in the following way, first suppose that the measure ν is finite, then let $\{\xi_i\}_{i=1}^{\infty}$ be a sequence of i.i.d. random variables with the common law $\nu/\nu(E)$, and N be a Poisson random variable with parameter $\nu(E)$, independent of the ξ_is. Then we can define the random measure K as, [1]

$$K = \sum_{i=1}^{N} \delta_{\xi_i}$$

where δ is the Dirac delta function.[2]

If ν is σ-finite we do the same construction for $\{E_n\}_{n=1}^{\infty}$ where $E_n \uparrow E$ to obtain a corresponding K_n, then we set

$$K = \sum_{n=1}^{\infty} K_n.$$

Let Δ_i be the jump times of N_t, i.e. $N_{\Delta_i} - N_{\Delta_i -} > 0$. We see that the measure K on sets $A \in \mathscr{B}[0, \infty) \times \mathscr{B}(\mathbb{R}\backslash\{0\})$ can also be expressed as,

$$K(A) = \sum_{i=1}^{\infty} \mathbf{1}_{(\Delta_i, \xi_i) \in A}. \tag{5.2.1}$$

This construction gives that for any $B \in \mathscr{B}[0, t) \times \mathscr{B}(\mathbb{R}\backslash\{0\})$, $t > 0$, $K(B) < \infty$ as the Poisson process has finite jumps in any finite time interval. The latter fact can be verified by the construction given above.

These measures play an important role in describing jumps of Lévy processes but we shall not be using them in what follows. The reader is encouraged to think about the proofs in the forthcoming chapter in a Poisson measure way.

[1] c.f. Bertoin (1996).

[2] Recall that Dirac delta function δ_x is zero everywhere but x, where it attains the value ∞.

Lévy Processes

" Mathematicians are like Frenchmen: whatever you say to them they translate into their own language and forthwith, it is something entirely different. "

-Johann Wolfgang von Goethe

6.1 Definitions

All the processes we have met in the previous chapter share some common ground. All of them have stationary, independent increments. Notice that even thought Brownian motion is continuous and Poisson processes are not continuous, however, they are all right continuous with left limits. This gives rise to a very general class of processes whose name is attributed to Paul Lévy.

The analysis of these processes will give a rich understanding of the underlying structure of most of the stochastic processes that we may encounter. Some books that deal with these processes are Kyprianou (2006), Sato (1999) and Bertoin (1996).

Definition 6.1.1 (Lévy Process)**.** A stochastic process L_t is said to be a *Lévy process* if it satisfies the following

(i) $L_0 = 0$ almost surely

(ii) for $0 \leqslant t_1 \leqslant ... \leqslant t_n$, $L_{t_k} - L_{t_{k-1}}, ..., L_{t_2} - L_{t_1}$ are independent

(iii) for $s < t$, $L_t - L_s$ is equal in distribution to L_{t-s}

(iv) $t \mapsto L_t$ is almost surely right continuous with left limits.

Any process satisfying *(i)*, *(ii)* and *(iii)* is called a *Lévy process in law.*

As it turns out, there is a deep connection between Lévy processes and infinitely divisible random variables. The next lemma will be a starting point of this connection. We will see in this chapter that we can, in some sense, establish a one-to-one correspondence between Lévy processes and infinitely divisible random variables.

Lemma 6.1.2. *A Lévy process L_t is infinitely divisible for each $t \geqslant 0$.*

Proof. Let L_t be a Lévy process, then for each $n \in \mathbb{N}$,

$$L_t \stackrel{d}{=} L_{\frac{t}{n}} + \left(L_{\frac{2t}{n}} - L_{\frac{t}{n}}\right) ... + \left(L_{\frac{nt}{n}} - L_{\frac{(n-1)t}{n}}\right)$$

□

6.2 Representations

6.2.1 Lévy-Khintchine representation

This next theorem is a reformulation of the Lévy-Khintchine canonical representation. This was presented by Paul Lévy in Lévy (1934) and later a much simpler proof was given by Khintchine (1937). This gives a simple and elegant way of working with Lévy processes, it allows us to have a general form of a characteristic function which is as good as having the law of the process.

Theorem 16 (Lévy-Khintchine representation)**.** *Let ψ_t be the characteristic function of a Lévy process then it is of the form*

$$\psi_t(\theta) = e^{t\Psi(\theta)} \tag{6.2.1}$$

where

$$\Psi(\theta) = \gamma\theta i - \frac{1}{2}\sigma^2\theta^2 + \int_{\mathbb{R}\backslash\{0\}} (e^{i\theta x} - 1 - i\theta x \mathbf{1}_{|x|<1})\Pi(dx) \tag{6.2.2}$$

where $\gamma \in \mathbb{R}$, $\sigma \geqslant 0$ and Π is a measure that satisfies

$$\int_{\mathbb{R}\backslash\{0\}} (x^2 \wedge 1) \quad \Pi(dx) < \infty.$$

Moreover for any characteristic function of this form, there exists a Lévy process in law such that it obtains the given characteristic function.

Proof. Let L_t be a Lévy process and let $n, m \in \mathbb{N}$, then we have that $L_n = L_{n/m} + (L_{2n/m} - L_{n/m}) + ... + (L_n - L_{(m-1)n/m})$. Hence we see that $n\psi_1 = \psi_n = m\psi_{n/m}$, and so for any $t > 0$ let $a_i \downarrow t$ where $\{a_i\}_{i=1}^{\infty} \subset \mathbb{Q}$, then $\psi_t = \lim_{i\to\infty} \psi_{a_i} = \lim_{i\to\infty} a_i\psi_1 = t\psi_1$ which gives (6.2.1).

As ψ_t is infinitely divisible, from Lévy canonical representation we know that, [1]

$$\Psi(\theta) = i\gamma\theta + \int_{\mathbb{R}} \left(e^{i\theta x} - 1 - \frac{i\theta x}{1+x^2}\right) \frac{1+x^2}{x^2} H(dx)$$

We define a measure Π on $\mathbb{R}\backslash\{0\}$ by,

$$\Pi(dx) = \frac{1+x^2}{x^2} H(dx).$$

Notice that this is well defined as a distribution function, as H is a distribution function so the limits are finite as $x \to \pm\infty$. Also for the same reasons as in the proof of Lévy canonical representation (Theorem 8), Π has an atom at $x = 0$. Recall that $F(dx) = \left(1 - \frac{\sin x}{x}\right) \frac{1+x^2}{x^2} H(dx)$ and so rearranging the Lévy-Khintchine canonical representation gives,

$$\begin{aligned} \Psi(\theta) &= i\theta a + \int_{\mathbb{R}} \left(e^{i\theta x} - 1 - \frac{i\theta x}{1+x^2}\right) \Pi(dx) \\ &= i\theta\gamma - \frac{1}{2}\sigma^2\theta^2 + \int_{\mathbb{R}\backslash\{0\}} (e^{i\theta x} - 1 - i\theta x \mathbf{1}_{|x|<1})\Pi(dx) \end{aligned}$$

where

$$\begin{aligned} \gamma &= a + \int_{\mathbb{R}\backslash\{0\}} \left(\mathbf{1}_{|x|\geqslant 1} \frac{1+x^2}{x} - x\right) F(dx) \\ \sigma^2 &= \Pi(\{0\}) \end{aligned}.$$

Note that

$$\int_{\mathbb{R}\backslash\{0\}} \left(\mathbf{1}_{|x|\geqslant 1} \frac{1+x^2}{x} - x\right) F(dx) < \infty$$

as we have $(1+x^2)/x - x \to 0$ as $x \to \pm\infty$ and as F is a finite measure.

For the second part take $0 \leqslant t_1 \leqslant t_2 \leqslant ... \leqslant t_n < \infty$ and consider the characteristic functions $\psi_i = e^{t_i\Psi}$ for each $i = 1, ..., n$, where Ψ is the characteristic exponent given by (6.2.2). We can reverse the procedure of the first half of the proof to obtain the Lévy-Khintchine canonical

[1] We use the Lévy-Khintchine canonical form (3.2.6)

representation,[2] thus we see that for each $i = 1, ..., n$, ψ_i is the characteristic function of some infinitely divisible random variable by the Lévy canonical representation.

Now we have specified the finite dimensional distributions of the a stochastic process. The reason for this is that the characteristic function is unique up to distribution,[3] which is a well known fact from Fourier transforms. Hence we can use the Kolmogorov's extension theorem to make this into a stochastic process on some probability space. It can be easily checked that this construction satisfies the conditions of the extension theorem.

We can also see that this satisfies stationary independent increments and hence it is a Lévy process in law. □

Remark 6.2.1. Notice that Kolmogorov's extension theorem does not give càdlàg paths. A proposition in Rogers and Williams (1988) tells us that a function $y : \mathbb{Q} \to \mathbb{R}$ has limits on $t \in \mathbb{R}$ if on any finite time interval $\sup |y| < \infty$ and the number of upper crossing are finite. By restricting the process we get from Kolmogorov's extension theorem to $\mathbb{Q}$, we can prove that these conditions hold almost surely, thus we can in fact construct a càdlàg process, by taking $X_t = \lim_{\mathbb{Q} \ni s \downarrow t} X'_s$, that has the characteristic function given by (6.2.2). The interested reader is referred to Protter (2005) for a formal approach using martingales.

Notation 6.2.2. A Lévy process is characterised by the triplet (γ, σ^2, Π) given by the Lévy-Khintchine representation in (6.2.2). The measure Π is called the Lévy measure.

6.2.2 Lévy-Itô decomposition

Next we prove a theorem that gives deep insight into the workings of a Lévy process. This was first proved by Itô (1942). It states that a Lévy process can be thought of a combination of Brownian motion, compound Poisson and a pure jump process. We will be following a similar path to that of Kyprianou (2007). An alternative, nevertheless very similar, proof using Poisson random measures can be found in Kyprianou (2006).

The Lévy-Khintchine representation gives us a characteristic function for any given Lévy process. This section is dedicated to the converse of this. Given a characteristic function of the form (6.2.2), we wish to obtain a Lévy process that satisfies this. We so far can only obtain a Lévy process in law. We will prove the converse of this via the Lévy-Itô decomposition, which will give more insight then we require into the composition of Lévy processes.

Square Integrable Martingales

Looking at the Lévy-Khintchine representation, we see that it separates into three parts. First of these is a Brownian motion with drift, which has characteristic exponent of the form $\gamma\theta i - \frac{1}{2}\sigma^2\theta^2$. The integral gives us two process, one of which is a compound Poisson process on $\mathbb{R}\backslash(-1,1)$ with characteristic exponent $\int_{\mathbb{R}\backslash(-1,1)}(e^{i\theta x} - 1)\Pi(dx)$. So all that is left is the mysterious process with the characteristic exponent $\int_{(-1,1)\backslash\{0\}}(e^{i\theta x} - 1 - i\theta x)\Pi(dx)$. As it will be shown, this too is a Lévy process, moreover it is a square integrable Martingale with finite jumps on any finite interval.

We will construct this process as a limit of compound Poisson process. The study of martingales is first required in order to be able to construct this.

By a *filtration* on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$, we mean a sequence of σ-subalgebras $\{\mathscr{F}_t\}_{t\in\mathbb{R}}$ such that $\mathscr{F}_s \subset \mathscr{F}_t$ whenever $s \leqslant t$. A stochastic process X is said to be $\mathscr{F}_t$ adapted if for each $t \geqslant 0$, X_t is $\mathscr{F}_t$-measurable. We will always be assuming that the process we are dealing with is adapted to the filtration. A natural way to construct a filtration with a stochastic process is

[2] Here we absorb the t_i into the characteristic function, i.e. we have $\gamma' = t_i\gamma$, $\sigma'^2 = t_i\sigma^2$ and $\Pi' = t_i\Pi$.

[3] By this we mean two random variables have the same characteristic function if and only if they have the same distribution.

$\mathscr{F}_t = \sigma(X_s : s \leqslant t)$.[4] We define conditional expectation $\mathbb{E}[Y|\mathscr{F}_s]$, where Y is a random variable, to be the random variable Z such that for each $B \in \mathscr{F}_s$, $\int_B Z(\omega)\mathbb{P}(d\omega) = \int_B \mathbb{E}[Y|\mathscr{F}_s](\omega)\mathbb{P}(d\omega)$.

A *martingale* M with respect to a filtration $\mathscr{F}_t$ is a stochastic process with $\mathbb{E}[|M_t|] < \infty$ for each $t \in \mathbb{R}$ and $\mathbb{E}[M_t|\mathscr{F}_s] = M_s$ whenever $s \leqslant t$.

We can now begin by proving a crucial result. The processes that we wish to take the limit of are Cauchy, thus we need to define a Hilbert space to conclude that they converge.

Theorem 17. *The space of square integrable martingales* $\mathbb{M}_T^2 = \mathbb{M}_T^2(\Omega, \mathscr{F}, \{\mathscr{F}_t : t \in [0,T]\}, \mathbb{P})$ *is a Hilbert space under the scalar product*

$$\langle M, N\rangle = \mathbb{E}[M_T N_T] \quad M, N \in \mathbb{M}_T^2. \tag{6.2.3}$$

Proof. First we prove that (6.2.3) induces a scalar product on $\mathbb{M}_T^2$. Let $M, N, H \in \mathbb{M}_T^2$ and $\alpha, \beta \in \mathbb{R}$ then

$$\begin{aligned}\langle \alpha N + \beta M, H\rangle = \mathbb{E}[(\alpha N_T + \beta M_T)H_T] &= \mathbb{E}[\alpha N_T H_T + \beta M_T H_T] \\ &= \alpha\mathbb{E}[N_T H_T] + \beta\mathbb{E}[M_T H_T] \\ &= \alpha \langle N, H\rangle + \beta \langle M, H\rangle .\end{aligned}$$

Also

$$\langle M, M\rangle = \mathbb{E}[M_T^2] \geqslant 0.$$

When $\mathbb{E}[M_T^2] = 0$ then we can use Doob's Maximal Inequality,

$$\mathbb{E}\left[\sup_{0\leqslant t\leqslant T} M_t^2\right] \leqslant 4\mathbb{E}[M_T^2] = 0$$

thus we have that $M = 0$ almost surely. Thus we have an scalar product on $\mathbb{M}_T^2$.

Now take a Cauchy sequence $\{M^{(n)}\}_{n=1}^\infty$ in $\mathbb{M}_T^2$ then

$$||M^{(n)} - M^{(m)}|| = \mathbb{E}[(M_T^{(n)} - M_T^{(m)})^2]^{\frac{1}{2}}.$$

Hence $\{M_T^{(n)}\}_{n=1}^\infty$ is a Cauchy sequence in the Hilbert space of square integrable random variables $L^2(\Omega, \mathscr{F}_T, \mathbb{P})$ and hence converges.[5] Thus we have that $\{M^{(n)}\}_{n=1}^\infty$ converges to some $M \in \mathbb{M}_T^2$ which completes the proof. □

With a Hilbert space, we may begin constructing a sequence that will turn out to be a Cauchy sequence and hence converge. Firstly we need some preliminary results about the compound Poisson process. This next Lemma may seem obvious but it does no harm for us to check it.

Lemma 6.2.3. *Suppose that* $\int_{\mathbb{R}} |x|F(dx) < \infty$*, then the process* $\{M_t : t \geqslant 0\}$ *defined by*

$$M_t = \sum_{i=1}^{N_t} \xi_i - \lambda t \int_{\mathbb{R}} xF(dx)$$

is a martingale with the natural filtration.[6]

Moreover if $\int_{\mathbb{R}} x^2 F(dx) < \infty$ *then* M *is a square integrable martingale with,*

$$\mathbb{E}[M_t^2] = \lambda t \int_{\mathbb{R}} x^2 F(dx).$$

[4] Henceforth we will be assuming that the filtration is complete and right continuous. We say that a σ-algebra is complete (with respect to $\mathbb{P}$) if for each $\mathbb{P}$-null set $B \in \mathscr{F}$, the subsets $A \subset B$ are also in $\mathscr{F}$. We can complete a filtration by taking all the $\mathbb{P}$-null sets and adding them on to each $\mathscr{F}_t$. We can make any filtration $\{\mathscr{F}_t : t \geqslant 0\}$ right continuous by making a new filtration $\mathscr{F}_t^* := \cap_{s>t}\mathscr{F}_s$. Thus these two assumptions do not restrict us.

[5] Notice that square integrable random variables are closed in L^2 thus they form a Hilbert space in their own right.

[6] Here we take N_t to be a Poisson process with rate λ and F to be the measure of the i.i.d. random variables $\{\xi_i\}_{i=1}^\infty$.

Proof. Notice that M is a càdlàg process with independent, stationary increments, as it is a compound Poisson process with drift. Thus we have that

$$\mathbb{E}[M_t|\mathscr{F}_s] = M_s + \mathbb{E}[M_t - M_s|\mathscr{F}_s] = M_s + \mathbb{E}[M_{t-s}]. \tag{6.2.4}$$

We will be done once we can show that $\mathbb{E}[M_{t-s}] = 0$. Note that for any $u \geqslant 0$,

$$\mathbb{E}[M_u|N_u] = \sum_{i=1}^{N_u} \mathbb{E}[\xi_i|N_u] - \lambda t \int_{\mathbb{R}} xF(dx) = N_u \mathbb{E}[\xi_1] - \lambda t \int_{\mathbb{R}} xF(dx).$$

Hence we can use the tower property to deduce that,

$$\mathbb{E}[M_u] = \lambda t \mathbb{E}[\xi_1] - \lambda t \int_{\mathbb{R}} xF(dx).$$

Note that $\lambda t \int_{\mathbb{R}} |x| F(dx) < \infty$, thus we can deduce that $\mathbb{E}[M_u] = 0$ and hence by plugging this in to (6.2.4) we see that M_t is a martingale with respect to the natural filtration.

Now we prove the second part of the Lemma. Suppose that $\int_{\mathbb{R}} x^2 F(dx) < \infty$, then

$$\begin{aligned}
\mathbb{E}[M_t^2] &= \mathbb{E}\left[\left(\sum_{i=1}^{N_t} \xi_i\right)^2\right] - 2\lambda t \mathbb{E}\left[\sum_{i=1}^{N_t} \xi_i\right] \int_{\mathbb{R}} xF(dx) + \lambda^2 t^2 \left(\int_{\mathbb{R}} xF(dx)\right)^2 \\
&= \mathbb{E}\left[\left(\sum_{i=1}^{N_t} \xi_i\right)^2\right] - \lambda^2 t^2 \left(\int_{\mathbb{R}} xF(dx)\right)^2 \qquad (6.2.5) \\
&= \mathbb{E}\left[\sum_{i=1}^{N_t} \xi_i^2\right] + \mathbb{E}\left[\sum_{k=1}^{N_t} \sum_{l=1, l\neq k}^{N_t} \xi_k \xi_l\right] - \lambda^2 t^2 \left(\int_{\mathbb{R}} xF(dx)\right)^2.
\end{aligned}$$

Notice that conditioning on N_t we get that,

$$\begin{aligned}
\mathbb{E}\left[\sum_{k=1}^{N_t} \sum_{l=1, l\neq k}^{N_t} \xi_k \xi_l | N_t\right] &= \sum_{k=1}^{N_t} \sum_{l=1, l\neq k}^{N_t} \mathbb{E}[\xi_1]^2 \\
&= (N_t^2 - N_t)\mathbb{E}[\xi_1]^2.
\end{aligned}$$

Hence we obtain

$$\mathbb{E}\left[\sum_{k=1}^{N_t} \sum_{l=1, l\neq k}^{N_t} \xi_k \xi_l\right] = \mathbb{E}[N_t^2 - N_t] \left(\int_{\mathbb{R}} xF(dx)\right)^2 = \lambda^2 t^2 \left(\int_{\mathbb{R}} xF(dx)\right)^2.$$

Plugging this in (6.2.5) gives that,

$$\mathbb{E}[M_t^2] = \mathbb{E}\left[\sum_{i=1}^{N_t} \xi_i^2\right] = \lambda t \int_{\mathbb{R}} x^2 F(dx).$$

□

For the next theorem let us define $\{N_t^{(n)}\}_{n=1}^{\infty}$ to be mutually independent Poisson processes with rate λ_n.[7] For $n \in \mathbb{N}$ let $\{\xi_i^{(n)}\}_{i=1}^{\infty}$ be i.i.d. random variables with common distribution F_n which does not assign a mass to the origin. Suppose further that,

$$\int_{\mathbb{R}} x^2 F_n(dx) < \infty \quad n \in \mathbb{N}.$$

[7] If $\lambda_n = 0$ then we take $N = 0$.

Let $M^{(n)}$ be constructed as in the previous Lemma with the pair $(N^{(n)}, F_n)$, then we may obtain a common filtration by

$$\mathscr{F}_t = \sigma\left(\bigcup_{n\geqslant 1} \mathscr{F}_t^{(n)}\right)$$

where $\mathscr{F}_t^{(n)}$ is the natural filtration generated by $M^{(n)}$.

Theorem 18. *If*

$$\sum_{n=1}^{\infty} \lambda_n \int_{\mathbb{R}} x^2 F_n(dx) < \infty \tag{6.2.6}$$

then there exists a Lévy Process $L = \{L_t : t \geqslant 0\}$ that is a square integrable martingale on the same probability space as $\{M^{(n)} : n \geqslant 1\}$ which has a characteristic exponent of the form

$$\Psi(\theta) = \int_{\mathbb{R}} (e^{i\theta x} - 1 - i\theta x) \sum_{n=1}^{\infty} \lambda_n F_n(dx). \tag{6.2.7}$$

Moreover for each $\theta \in \mathbb{R}$ such that for each fixed $T > 0$ we have,

$$\lim_{k\uparrow\infty} \mathbb{E}\left[\sup_{t\leqslant T}\left(L_t - \sum_{n=1}^{k} M_t^{(n)}\right)^2\right] = 0. \tag{6.2.8}$$

Proof. Notice that by the linearity of the expectation (more precisely the conditional expectation), we have that any sum of the form $\sum_n M^{(n)}$ is also a martingale. Moreover by independence and the martingale property $\mathbb{E}[M_t^{(i)} M_t^{(j)}] = \mathbb{E}[M_t^{(i)}]\mathbb{E}[M_t^{(j)}] = 0$ for $i \neq j$. Thus we have that

$$\mathbb{E}\left[\left(\sum_{n=1}^{k} M_t^{(n)}\right)^2\right] = \sum_{n=1}^{k} \mathbb{E}[(\mathbb{M}_t^{(n)})^2] = t\sum_{n=1}^{k} \lambda_n \int_{\mathbb{R}} x^2 F_n(dx) < \infty. \tag{6.2.9}$$

We have now that for each $k \in \mathbb{N}$, $\displaystyle\sum_{n=1}^{k} M_t^{(n)} \in \mathbb{M}_T^2$ for a fixed $T > 0$. We wish to prove now that the sequence $\{L^{(k)}\}_{k=1}^{\infty}$ defined by,

$$L^{(k)} = \sum_{n=1}^{k} M_t^{(n)}$$

is a Cauchy sequence in $\mathbb{M}_T^2$, where we will assume that $T > 0$ is fixed. Now we have that for $m > n$,

$$||L^{(m)} - L^{(n)}|| = \mathbb{E}\left[\left(\sum_{k=1}^{m} M_t^{(k)} - \sum_{k=1}^{n} M_t^{(k)}\right)^2\right] = \sum_{k=n}^{m} \mathbb{E}[(M_t^{(k)})^2] = \sum_{k=n}^{m} \lambda_n \int_{\mathbb{R}} x^2 F_n(dx)$$

using (6.2.9). Now as $\sum_{n=1}^{\infty} \lambda_n \int_{\mathbb{R}} x^2 F_n(dx) < \infty$, we see that $\{L^{(k)}\}_{k=1}^{\infty}$ is a Cauchy sequence in $\mathbb{M}_T^2$ and by Theorem 17 converges to some $L \in \mathbb{M}_T^2$. We can now apply Doob's Maximal Inequality to obtain,

$$\lim_{n\uparrow\infty} \mathbb{E}\left[\sup_{0\leqslant t\leqslant T} (L_t^{(n)} - L_t)^2\right] = 0.$$

Now we can use the Lévy Continuity Theorem to get that

$$\mathbb{E}[e^{i\theta(L_t - L_s)}] = \lim_{n\uparrow\infty} \mathbb{E}[e^{i\theta(L_t^{(n)} - L_s^{(n)})}] = \lim_{n\uparrow\infty} \mathbb{E}[e^{i\theta L_{t-s}^{(n)}}] = \mathbb{E}[e^{i\theta L_{t-s}}]$$

which shows the independent stationary increments. Using the fact that the $M^{(n)}$s are independent we obtain that,

$$\begin{aligned}
\mathbb{E}[e^{i\theta L_t^{(k)}}] &= \prod_{n=1}^{k} \mathbb{E}[e^{i\theta M_t^{(n)}}] \\
&= \prod_{n=1}^{n} \mathbb{E}[e^{\sum_{j=1}^{N_t} \xi_j - i\theta\lambda_n t \int_{\mathbb{R}} x F_n(dx)}] \\
&= \prod_{n=1}^{n} \mathbb{E}[e^{-\lambda_n t \int_{\mathbb{R}} (1-e^{i\theta x}) F_n(dx) - i\theta\lambda_n t \int_{\mathbb{R}} x F_n(dx)}] \\
&= \prod_{n=1}^{n} \mathbb{E}[e^{-\lambda_n t \int_{\mathbb{R}} (1-e^{i\theta x} + i\theta x) F_n(dx)}] \\
&= \mathbb{E}[e^{t \int_{\mathbb{R}} (e^{i\theta x} - 1 - i\theta x) \sum_{n=1}^{k} \lambda_n F_n(dx)}].
\end{aligned}$$

We can use the Lévy Continuity Theorem and (6.2.6) we can see that,

$$\mathbb{E}[e^{i\theta L_t}] = e^{\int_{\mathbb{R}} (e^{i\theta x} - 1 - i\theta x) \sum_{n=1}^{\infty} \lambda_n F_n(dx)}.$$

All that is left to prove L is a Lévy process is that L has càdlàg paths.

Consider the space of functions $f : [0,T] \to \mathbb{R}$ under the supremum metric $d(f,g) = \sup_{t\in[0,T]} |f(t) - g(t)|$. Take a sequence f_n of càdlàg functions that converge to f pointwise. Fix $\epsilon > 0$, then by convergence we can pick $N \in \mathbb{N}$ s.t. $d(f_n(x) - f(x)) \leqslant \epsilon/2$ for $n > N$, and for each $n \in \mathbb{N}$ $d(f_n(x+\epsilon) - f(x)) \to 0$ as, hence

$$\begin{aligned}
d(f(x+\epsilon) - f(x)) &\leqslant d(f(x+\epsilon) - f_{N+1}(x+\epsilon)) + d(f_{N+1}(x+\epsilon) - f_{N+1}(x)) + d(f(x) - f_{N+1}(x)) \\
&\leqslant \epsilon + d(f_{N+1}(x+\epsilon) - f_{N+1}(x)).
\end{aligned}$$

Taking the limit as $\epsilon \to 0$ gives the result that f is right continuous. Similarly

$$\begin{aligned}
d(f(x-\epsilon) - f(x)) &\leqslant d(f(x-\epsilon) - f_{N+1}(x-\epsilon)) + d(f_{N+1}(x-\epsilon) - f_{N+1}(x)) + d(f(x) - f_{N+1}(x)) \\
&\leqslant \epsilon + d(f_{N+1}(x-\epsilon) - f_{N+1}(x))
\end{aligned}$$

and again by letting $\epsilon \to 0$ we have that f has left limits (as f_{N+1} have left limits). Hence the space of càdlàg functions is closed.

As L is the limit of càdlàg functions, this shows that L is càdlàg almost surely. So L must be a Lévy process.

We have one outstanding issue, that is the process L depends on T. This may be problematic if the limit changed when we changed T. For the proof to work we need the processes to agree on the same time horizons. We will now confirm this fact.

Suppose that we have two time horizons $T_1 \geqslant T_2$ and label L^T as the process L with the time horizon T. Using the triangle inequality of the supremum and Minkovski's inequality[8] we obtain

$$\mathbb{E}\left[\sup_{t\in T_1} (L_t^{T_1} - L_t^{T_2})^2\right]^{\frac{1}{2}} \leqslant \mathbb{E}\left[\sup_{t\in T_1} (L_t^{T_1} - L_t^{(n)})^2\right]^{\frac{1}{2}} + \mathbb{E}\left[\sup_{t\in T_1} (L_t^{(n)} - L_t^{T_2})^2\right]^{\frac{1}{2}}.$$

Letting $n \to \infty$ and using (6.2.8) that we proved earlier the expectation tends to zero, hence we see that the two processes agree almost surely on the time horizon T_1. Thus the limit does not depend on T. □

[8]The triangle inequality on L^p spaces

Lévy-Itô decomposition

We are now in a position to prove the main result that we seek.

Theorem 19 (Lévy-Itô decomposition). *Let (a, σ^2, Π) be a Lévy triplet, then there exists a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ on which three processes $L^{(1)}$, $L^{(2)}$ and $L^{(3)}$ exist, where $L^{(1)}$ is a Brownian motion with drift, $L^{(2)}$ is a compound Poisson process and $L^{(3)}$ is a square integrable pure jump martingale that almost surely has a countable number of jumps on each finite interval.*

L defined by $L = L^{(1)} + L^{(2)} + L^{(3)}$ is a Lévy process.

Proof. Decompose the Lévy process L as given; by Lévy-Khintchine we have that the characteristic exponent Ψ of L is given by,

$$\Psi = \Psi_1 + \Psi_2 + \Psi_3$$

where Ψ_1 is the characteristic exponent of a Brownian motion with drift, Ψ_2 is the characteristic exponent of a compound Poisson process on $\mathbb{R}\backslash(1, -1)$ and

$$\Psi_3(\theta) = \int_{(-1,1)\backslash\{0\}} (e^{i\theta x} - 1 - i\theta x)\Pi(dx).$$

The existence of $L^{(1)}$ and $L^{(2)}$ have been shown in previous chapters. We wish to show the existence of $L^{(3)}$. Take

$$\lambda_n = \Pi(\{x | 2^{-(n+1)} \leqslant |x| < 2^{-n}\})$$

and

$$F_n(dx) = \lambda_n^{-1}\Pi(dx)|_{\{x|2^{-(n+1)} \leqslant |x| < 2^{-n}\}}.$$

We verify that the assumptions of Theorem 18 hold as

$$\sum_{n=1}^{\infty} \lambda_n \int x^2 F_n(dx) = \int_{(-1,1)\backslash\{0\}} x^2 \Pi(dx) < \infty.$$

So we may conclude that a Lévy process $L^{(3)}$ exists and has characteristic exponent given by Ψ_3. We can take a common probability space (e.g. the product space) where each of these processes exist on and hence conclude that L exists. □

6.3 Strong Markov Property

In this section we will assume to be working on a filtered probability space $(\Omega, \mathscr{F}, \{\mathscr{F}_t : t \geqslant 0\}, \mathbb{P})$, where we will assume that the filtration $\{\mathscr{F}_t : t \geqslant 0\}$ is right continuous and complete.

Markov property is one of the most famous properties of stochastic processes. It was formulated by Andrey Markov. Recall that we say a stochastic process $X = \{X_t : t \geqslant 0\}$ possesses the *Markov property* if for any $B \in \mathscr{B}(\mathbb{R})$ and $0 \leqslant s < t$ we have that

$$\mathbb{P}(X_t \in B | \mathscr{F}_s) = \mathbb{P}(X_t \in B | \sigma(X_s)).$$

In a way, this formula tells us that the stochastic process is memoryless up to some extent. If we know where we are, the probability of where we are going does not change if we know where we came from.

It is simple to check that any Lévy process satisfies the Markov property. We see that they posses the property that increments are independent. This is a much stronger statement than the Markov property. The aim of this section is to establish this.

First we need some basic definitions of stopping times.

Definition 6.3.1. A non-negative random variable τ is called a *stopping time* if

$$\{\tau < t\} \in \mathscr{F}_t.$$

Mostly we will be working with first hitting times $\tau_x := \inf\{t > 0 : X_t > x\}$. These, as the name suggest, give the first time that a process attains a value greater than x.

Now we need a notion of a random filtration. Take the natural filtration on a process X, it is natural to define a random filtration by $\mathscr{F}_\tau = \{B \in \mathscr{F}_t : B \cap \{\tau < t\} \in \mathscr{F}_t\}$. We are in a position now to define the strong Markov property.

Definition 6.3.2. We say that a stochastic process X satisfies the *strong Markov property* if

$$\mathbb{P}(X_t \in B|\mathscr{F}_\tau) = \mathbb{P}(X_t \in B|\sigma(X_\tau))$$

for any stopping time $\tau < \infty$ a.s.

Equivalently $(X_{\tau+t} - X_\tau)$ conditioned on $\{\tau < \infty\}$, where $\mathbb{P}(\tau < \infty) > 0$ is independent of $\mathscr{F}_\tau$ and has law $\mathbb{P}$.

Now we can state the main result of this section.

Theorem 20. *A Lévy process L satisfies the strong Markov property.*

Proof. If the stopping time is deterministic then the result follows from the Markov property so let τ be a non-deterministic stopping time with $\tau < \infty$ a.s. We prove this for two cases, first we consider τ taking values in a discrete set, then we have

$$\tau = \sum_{n=1}^{\infty} t_n \mathbf{1}_{\tau = t_n}$$

for some $0 < t_1 < t_2 < ....$ Thus we have for any $B \subset \mathscr{B}(\mathbb{R})$,

$$\begin{aligned}\mathbb{P}(L_{\tau+t} - L_\tau \in B|\mathscr{F}_\tau) &= \sum_{k=\sum t_n} \mathbb{P}(L_{t_n+t} - L_{t_n} \in B|\mathscr{F}_k)\mathbb{P}(\tau = k|\mathscr{F}_\tau)\\ &= \mathbb{P}(L_t \in B) \sum_{k=\sum t_n} \mathbb{P}(\tau = k) = \mathbb{P}(L_t \in B).\end{aligned}$$

Now suppose τ is not discrete, then we construct $\tau_n = 2^{-n}\lfloor 2^n\tau + 1\rfloor$ where $\lfloor x\rfloor$ is x rounded down to the closest integer. Now it is obvious that $\tau_n \downarrow \tau$ and so by right continuity we have $L_{\tau_n} \to L_\tau$ and hence the result follows from the first part. □

6.4 Points of Increase

Denote by I, the set of all t such that,

$$X_s \leqslant X_t \quad s \in [t - \delta_1, t] \qquad X_s \geqslant X_t \quad s \in [t, t + \delta_2] \tag{6.4.1}$$

for some $\delta_1, \delta_2 > 0$, where X is a stochastic process (not necessarily a Lévy process). We call the set I points of increase because they describe the points in which a stochastic process is lower than, in some interval below, and higher than in some other interval above. The question is, what are the conditions under which $I \neq \emptyset$? By Kolmogorov's 0-1 Law we can deduce that $\mathbb{P}(I \neq \emptyset) = 0$ or 1. In the case that we almost surely have $I \neq \emptyset$ we say that X has points of increase.

Burdzy (1990) proved in the case of Brownian motion we do not have any points of increase. This is intuitively clear as the Brownian motion is nowhere differentiable. It is also clear to see that a Poisson process (or indeed a compound Poisson process) has points of increase. We will give sufficient and necessary conditions of a Lévy processes having points of increase. We will be describing the steps in the paper by Doney (1996), and expanding out the paper by explicitly proving the claimed obvious statements.

Let $L = \{L_t : t \geqslant 0\}$ be a Lévy process and $\mathbf{e}_q$ be an exponential random variable with parameter $q > 0$ that is independent of L. Define $\overline{L}$ and $\underline{L}$ as

$$\overline{L}_t = \sup_{0 \leqslant s \leqslant t} L_s \quad \underline{L}_t = \inf_{0 \leqslant s \leqslant t} L_s$$

and suppose that $\overline{L}_{\mathbf{e}_q}$, $\underline{L}_{\mathbf{e}_q}$ have distribution functions $\overline{F}$ and $\underline{F}$ respectively. Define the first passage times

$$\overline{\tau}_x = \inf\{t : \overline{L}_t > x\} \quad \underline{\tau}_x = \inf\{t : \underline{L}_t < -x\}$$

for $x \geqslant 0$ and define

$$R_\epsilon = \begin{cases} \overline{L}_{\underline{\tau}_\epsilon} - L_{\underline{\tau}_\epsilon} & \text{if } \underline{\tau}_\epsilon \leqslant \mathbf{e}_q \\ \infty & \text{if } \underline{\tau}_\epsilon > \mathbf{e}_q \end{cases} .$$

We will require some more terminology before we can state the main theorem. We say $x \in \mathbb{R}$ is regular for a closed or open subset $B \subset \mathbb{R}$,

$$\mathbb{P}(\tau^B = 0 | L_0 = x) = 1$$

where $\tau^B = \inf\{t > 0 : L_t \in B\}$. Informally, the process hits B straight after starting at x. Now we can prepare some preliminary results. These can be found in Rogers (1984). First of these is the so called duality lemma. We will not be utilising the full potential of this result here, however an important result that follows is the Wiener-Hopf factorizations. Interested reader is referred to chapter 6 of Kyprianou (2006).

Lemma 6.4.1 (Duality Lemma)**.** *Suppose that L is a Lévy process and fix $T > 0$, then the following have the same laws;*

$$\{L_t : 0 \leqslant t \leqslant T\} \qquad \{L_{T-} - L_{(T-t)-} : 0 \leqslant t \leqslant T\}.$$

Proof. First let $L_t^* = L_{T-} - L_{(T-t)-}$, now it is clear that both L_t and L_t^* start at 0 and are càdlàg. Now we will prove that the characteristic functions of L^* and L coincide and that L^* has independent increments, which will complete the proof.

Take $t_n \uparrow T$ and $s_n \uparrow T - t$, now for each $n \in \mathbb{N}$ $L_{t_n} - L_{s_n}$ has the same distribution as $L_{t_n - s_n}$ as it is a Lévy process. For the same reason, the characteristic function is given by the Lévy-Khintchine formula and is of the form

$$\mathbb{E}[e^{i\theta L_{t_n - s_n}}] = e^{(t_n - s_n)\Psi(\theta)}$$

where Ψ is the characteristic exponent of L. Now by using the continuity of the exponential we can rid of left limits and as $n \to \infty$ we have

$$e^{(t_n - s_n)\Psi(\theta)} \to e^{(T-(T+t))\Psi(\theta)} = e^{t\Psi(\theta)}.$$

As $e^{t\Psi(\theta)}$ is a characteristic function, namely of L_t, we can use the Lévy Continuity Theorem to conclude that $L_{t_n} - L_{s_n}$ converges in distribution to L_t^* and thus,

$$\mathbb{E}[e^{i\theta L_t^*}] = e^{t\Psi(\theta)} = \mathbb{E}[e^{i\theta L_t}].$$

Characteristic functions are unique up to distribution, so we can conclude that L_t^* has the same law as L_t.

To show stationary independent increments, take $t, s \geqslant 0$ and let $k_n \uparrow T - s$ and $p_n \uparrow T - t - s$. Now we have that $L_{t+s}^* - L_s^* = L_{(T-s)-} - L_{(T-t-s)-}$. Using the same argument as above we see that the characteristic function of $L_{t+s}^* - L_s^*$ is

$$\lim_{n \to \infty} e^{(k_n - p_n)\Psi(\theta)} = e^{t\Psi(\theta)}.$$

□

Now we prove a simple corollary of this lemma which we need.

Corollary 6.4.2. *Suppose* 0 *is regular for* $(-\infty, 0)$ *then*

(i) $\mathbb{P}(\{\exists t : L_t > L_{t-} = \overline{L}_{t-}\}) = 0$

(ii) $\mathbb{P}(\{\exists t : L_t < L_{t-} = \overline{L}_{t-}\}) = 0.$

Proof. We will be proving the two statements in one go. Fix $T > 0$, $\epsilon > 0$ and consider

$$\{\exists 0 \leqslant t \leqslant T : |L_t - L_{t-}| > \epsilon, L_{t-} = \overline{L}_{t-}\}. \tag{6.4.2}$$

We wish to show this event has probability zero, and we will be done. By the Lemma we have just proved, this event has the same law as

$$\{\exists 0 \leqslant t \leqslant T : |L^*_t - L^*_{t-}| > \epsilon, L^*_t \leqslant L^*_u \quad \forall u \geqslant t\} \tag{6.4.3}$$

where $L^*_t = L_{T-} - L_{(T-t)-}$.

Now notice that for each bounded stopping time τ_i, $L^*_{\tau_i} - L^*_{\tau_i - t}$ has the same distribution as L^*_t by the strong Markov property. As 0 is regular for $(-\infty, 0)$ we have that $L^*_{\tau_i+u} < L^*_{\tau_i}$ for some $u > 0$ so (6.4.3), and consequently (6.4.2), are null sets. □

We will now require the next lemma in order to prove the theorem by Doney.

Lemma 6.4.3. *Let* $\tilde{I}$ *be the set of points* t *of the form*

$$L_s \leqslant L_t \quad s \in [0, t] \qquad L_s \geqslant L_t \quad s \in [t, \mathbf{e}_q]$$

called the global increase points. Then $\mathbb{P}(I \neq \emptyset) = 1$ *if and only if* $\mathbb{P}(\tilde{I} \neq \emptyset) > 0$.

Before we begin the proof, let us try to see why this should be true. If we have a point of increase, then there is a positive probability of the exponential $\mathbf{e}_q$ taking the value $t + \delta_2$. We could just re-shift the axis to $(t - \delta_1, L_{t-\delta_1})$ and then the points of increase would be at 0. This is the intuition behind the lemma.

Proof. Suppose $I \neq \emptyset$ almost surely, then we pick $t \in I$. Notice that $L_s - L_{t-\delta_1}$ has the same distribution as $L_{s-t+\delta_1}$, hence by subtracting $L_{s-\delta_1}$ off both sides we obtain $L_{s-t+\delta_1} \leqslant L_{\delta_1}$ and as $s - t + \delta_1 \geqslant 0$, after relabeling $u = s - t + \delta_1$ we obtain

$$L_u \leqslant L_{\delta_1} \quad u \in [0, \delta_1].$$

Taking away $L_{s-\delta_1}$ and using the same substitution from the second part we obtain

$$L_u \geqslant L_{\delta_1} \quad u \in [\delta_1, \delta_1 + \delta_2].$$

Hence by picking $\mathbf{e}_q$ such that $\mathbf{e}_q \leqslant \delta_1 + \delta_2$ with strictly positive probability we obtain the result.

Now suppose that $\tilde{I} \neq \emptyset$ with strictly positive probability. It suffices to show that $I \neq \emptyset$ with strictly positive probability (as we may apply the Kolmogorov 0-1 Law, as we did above). The case when $q = 0$, i.e. $\mathbf{e}_q = \infty$, the result follows immediately. If $q > 0$, then we pick a T such that $T \leqslant \mathbf{e}_q$ with positive probability. This gives us the result that $T \in I$ with strictly positive probability and hence the result follows. □

In what follows we will be ignoring the case when L (or $-L$) is a subordinator,[9] when L is a compound Poisson process and when 0 is irregular for $(-\infty, 0)$. It is obvious in each of these cases that the process will have increasing paths.

The theorem by Doney can then be stated as follows.

[9] Almost surely non-decreasing paths.

Theorem 21. *Let L be a Lévy process such that 0 is regular for $(-\infty,0)$, then L has points of increase if and only if*

$$\lim_{\epsilon\downarrow 0}\frac{\overline{F}(\epsilon)+\int_\epsilon^\infty \mathbb{P}(y<R_\epsilon<\infty)d\overline{F}(y)}{\underline{F}(\epsilon)}<\infty. \tag{6.4.4}$$

Proof. Let us first define $\widehat{L}$ by killing L after time $\mathbf{e}_q$. That is we add an other state, say Δ, such that $\widehat{L}_t = L_t$ for $t\leqslant \mathbf{e}_q$ and $\widehat{L}_t=\Delta$ for $t>\mathbf{e}_q$. We also define $\widehat{\underline{\tau}}$ and $\widehat{\overline{\tau}}$ for $\widehat{L}$ as we did above for L.

Now fix $\epsilon>0$. We define two sequence of random variables $\{W_n : n\in\mathbb{Z}_+\}$ and $\{Z_n : n\in\mathbb{N}\}$ by setting $W_0=0$ and $Z_1=\widehat{\underline{\tau}}_\epsilon$,

$$W_1=\begin{cases}\inf\{t>Z_1:\widehat{L}_t>\overline{L}_{Z_1}\} & \text{if } Z_1<\infty\\ \infty & \text{otherwise.}\end{cases}$$

For $n\geqslant 1$ define inductively,

$$W_{n+1}=\begin{cases}W_n+L_{W_n+W_1}-L_{W_n} & \text{if } W_n<\infty\\ \infty & \text{otherwise}\end{cases}$$

and

$$Z_{n+1}=\begin{cases}W_n+L_{W_n+Z_1}-L_{W_n} & \text{if } W_n<\infty\\ \infty & \text{otherwise.}\end{cases}$$

Define

$$A_n^{(\epsilon)}=\{W_{n-1}<\infty, Z_n=\infty\}$$

and let $A^{(\epsilon)}=\cup_{n\in\mathbb{N}}A_n^{(\epsilon)}$.

Let us now stand back and try to see what $A^{(\epsilon)}$ is telling us. If $Z_1<\infty$ then we have a time t such that $t\leqslant\mathbf{e}_q$ and $L_t\geqslant -\epsilon$. What this is trying to tell us is that we have $L_{\mathbf{e}_q}\geqslant L_{\mathbf{e}_q-t}-\epsilon$. The event W_1 is trying to squeeze a point s_1 between $\mathbf{e}_q-t$ and $\mathbf{e}_q$ such that $L_{\mathbf{e}_q}\geqslant L_s$.

All this time we have to check if the exponential clock has run out. If it hasn't by the memoryless property, we have an exponential time left. So given that the exponential clock has not run out, move our axis to (W_n, L_{W_n}) and repeat the process of squeezing in points. Now it is clear that

$$A^{(\epsilon)}=\{\exists t: 0\leqslant t\leqslant \mathbf{e}_q,\quad L_s\leqslant L_t\quad s\in[0,t]\qquad L_s\geqslant L_t-\epsilon\quad s\in[t,\mathbf{e}_q]\}.$$

So the the set that is the limit of $A^{(\epsilon)}$ as $\epsilon\downarrow 0$ is the set of global increase points. The proof will be complete once we derive a condition that gives this set a strictly possitive probability.

Note that from Corollary 6.4.2 we have that L does not jump up at any time t with $L_t=\overline{L}_t$, $\underline{L}_{\mathbf{e}_q}\neq 0$ and by time reversal $\overline{L}_{\mathbf{e}_q}\neq L_{\mathbf{e}_q}$. If we label A as the limit of $A^{(\epsilon)}$ as $\epsilon\downarrow 0$, then we see that $\mathbb{P}(A)=\mathbb{P}(\tilde{I}\neq\emptyset)$. For each $n\in\mathbb{N}$ we have that

$$\mathbb{P}(A_n^{(\epsilon)})=\mathbb{P}(W_{n-1}<\infty)\mathbb{P}(Z_1=\infty)=\mathbb{P}(W_1<\infty)^{n-1}\mathbb{P}(Z_1=\infty)$$

using the strong Markov property and the memoryless property of the exponential distribution. Hence

$$\mathbb{P}(A^{(\epsilon)})=\mathbb{P}(Z_1=\infty)\sum_{n=1}^\infty \mathbb{P}(W_1<\infty)^n=\frac{\mathbb{P}(Z_1=\infty)}{1-\mathbb{P}(W_1<\infty)}=\frac{\mathbb{P}(Z_1=\infty)}{\mathbb{P}(W_1=\infty)}.$$

We know that $\mathbb{P}(Z_1 = \infty) = \mathbb{P}(\widehat{\underline{\tau}}_\epsilon = \infty) = \mathbb{P}(-\underline{L}_{\mathbf{e}_q} \leqslant \epsilon) = \underline{F}(\epsilon)$, so we only need to evaluate the event $\{W_1 = \infty\}$. The strong Markov property on Z_1 gives,

$$\begin{aligned}
\mathbb{P}(W_1 = \infty) &= \mathbb{P}(Z_1 = \infty) + \mathbb{P}(\widehat{L}_t \leqslant \overline{L}_{Z_1} \text{ for } t > Z_1, Z_1 < \infty) \\
&= \underline{F}(\epsilon) + \mathbb{P}\left(Z_1 \leqslant \mathbf{e}_q, \sup_{0 \leqslant s \leqslant \mathbf{e}_q - Z_1} \{L_{Z_1+s} - L_{Z_1}\} \leqslant R_\epsilon \right) \\
&= \underline{F}(\epsilon) + \mathbb{E}[\overline{F}(R_\epsilon)\mathbf{1}_{R_\epsilon < \infty}] \\
&= \underline{F}(\epsilon) + \overline{F}(\epsilon)(1 - \underline{F}(\epsilon)) + \int_\epsilon^\infty \mathbb{P}(y < R_\epsilon < \infty) d\overline{F}(y).
\end{aligned}$$

Hence

$$\mathbb{P}(A) = \lim_{\epsilon \downarrow 0} \frac{\underline{F}(\epsilon)}{\underline{F}(\epsilon) + \overline{F}(\epsilon)(1 - \underline{F}(\epsilon)) + \int_\epsilon^\infty \mathbb{P}(y < R_\epsilon < \infty) d\overline{F}(y)}$$

and so $\mathbb{P}(A) > 0$ if and only if

$$\begin{aligned}
&\lim_{\epsilon \downarrow 0} \frac{\underline{F}(\epsilon) + \overline{F}(\epsilon)(1 - \underline{F}(\epsilon)) + \int_\epsilon^\infty \mathbb{P}(y < R_\epsilon < \infty) d\overline{F}(y)}{\underline{F}(\epsilon)} \\
&= 1 + \lim_{\epsilon \downarrow 0} \frac{\overline{F}(\epsilon) + \int_\epsilon^\infty \mathbb{P}(y < R_\epsilon < \infty) d\overline{F}(y)}{\underline{F}(\epsilon)} - \overline{F}(\epsilon) < \infty
\end{aligned}$$

and the result follows.

□

Bibliography

R. B. Ash. *Real Analysis and Probability.* Academic Press, 1972.

J. Bertoin. *Lévy Processes.* Cambridge University Press, 1996.

P. Billingsley. *Probability and Measure.* Wiley, 1979.

K. Burdzy. On nonincrease of brownian motion. *Annals of Probability*, 18:978–980, 1990.

R. Doney. Increase of Lévy processes. *Annals of Probability*, 24(2):961–970, 1996.

B. Fristedt and L. Gray. *A Modern Approach to Probability Theory.* Birkhauser Boston, 1997.

K. Itô. On stochastic processes. *Japan J. Math.*, 18:261–301, 1942.

A. Khintchine. A new derivation of one formula by Lévy P. *Bull. Moscow State Univ.*, 1(1):1–5, 1937.

A. E. Kyprianou. *Introductory Lectures on Fluctuations of Lévy Processes with Applications.* Springer, 2006.

A. E. Kyprianou. Lecture notes on Lévy processes from Sonderborg. Website, 2007. `http://www.maths.bath.ac.uk/~ak257/Levy-sonderborg.pdf`.

P. Lévy. Sur les integrales dont les elements sont des variables aleatoires independentes. *Ann. R. Scuola Norm. Super. Pisa*, 3:337–366, 1934.

E. Lukacs. *Characteristic Functions.* Oxford University Press, 2nd edition, 1970.

P. A. P. Moran. *An Introduction to Probability Theory.* Oxford University Press, 2nd edition, 1984.

P. E. Protter. *Stochastic Integration and Differential Equations.* Springer-Varlag, 2nd edition, 2005.

L. C. G. Rogers. A new identity for real levy processes. *Annales de l'I. H. P.*, 20(1):21–34, 1984.

L. C. G. Rogers and D. Williams. *Diffusions, Markov Processes and Martingales*, volume 1. Wiley, 1988.

K. Sato. *Lévy Processes and Infinitely Divisible Distributions.* Cambridge University Press, 1999.

D. W. Strook and S. R. S. Varadhan. *Multidimensional Diffusion Processes.* Springer-Varlag, 1979.