Lectures on Lévy Processes, Stochastic Calculus and Financial Applications, Ovronnaz September 2005

David Applebaum
Probability and Statistics Department,
University of Sheffield,
Hicks Building, Hounsfield Road,
Sheffield, England, S3 7RH

e-mail: D.Applebaum@sheffield.ac.uk

Introduction

A **Lévy process** is essentially a stochastic process with stationary and independent increments. The basic theory was developed, principally by Paul Lévy in the 1930s. In the past 15 years there has been a renaissance of interest and a plethora of books, articles and conferences. Why ?

There are both theoretical and practical reasons.

Theoretical

- There are many interesting examples - Brownian motion, simple and compound Poisson processes, α-stable processes, subordinated processes, financial processes, relativistic process, Riemann zeta process ...
- Lévy processes are simplest generic class of process which have (a.s.) continuous paths interspersed with random jumps of arbitrary size occurring at random times.
- Lévy processes comprise a natural subclass of *semimartingales* and of *Markov processes of Feller type*.
- Noise. Lévy processes are a good model of "noise" in random dynamical systems.

Input + Noise = Output

Attempts to describe this differentially leads to *stochastic calculus*. A large class of Markov processes can be built as solutions of *stochastic differential equations* driven by Lévy noise.

Lévy driven *stochastic partial differential equations* are beginning to be studied with some intensity.

- Robust structure. Most applications utilise Lévy processes taking values in Euclidean space but this can be replaced by a Hilbert space, a Banach space (these are important for spdes), a locally compact group, a manifold. Quantised versions are non-commutative Lévy processes on quantum groups.

APPLICATIONS

These include:

- Turbulence via Burger's equation (Bertoin)
- New examples of quantum field theories (Albeverio, Gottshalk, Wu)
- Viscoelasticity (Bouleau)
- Time series - Lévy driven CARMA models (Brockwell)
- Finance (a cast of thousands)

The biggest explosion of activity has been in mathematical finance. Two major areas of activity are:

- option pricing in incomplete markets.
- interest rate modelling.

1 Lecture 1: Infinite Divisibility and Lévy Processes

1.1 Some Basic Ideas of Probability

Notation. Our state space is Euclidean space $\mathbb{R}^d$. The inner product between two vectors $x = (x_1, \ldots, x_d)$ and $y = (y_1, \ldots, y_d)$ is

$$(x, y) = \sum_{i=1}^{d} x_i y_i.$$

The associated norm (length of a vector) is

$$|x| = (x, x)^{\frac{1}{2}} = \left(\sum_{i=1}^{d} x_i^2 \right)^{\frac{1}{2}}.$$

Let $(\Omega, \mathcal{F}, P)$ be a probability space, so that Ω is a set, $\mathcal{F}$ is a σ-algebra of subsets of Ω and P is a probability measure defined on $(\Omega, \mathcal{F})$.

Random variables are measurable functions $X : \Omega \to \mathbb{R}^d$. The law of X is p_X, where for each $A \in \mathcal{F}, p_X(A) = P(X \in A)$.

$(X_n, n \in \mathbb{N})$ are *independent* if for all $i_1, i_2, \ldots i_r \in \mathbb{N}, A_{i_1}, A_{i_2}, \ldots, A_{i_r} \in \mathcal{B}(\mathbb{R}^d)$,

$$P(X_{i_1} \in A_1, X_{i_2} \in A_2, \ldots, X_{i_r} \in A_r) = P(X_{i_1} \in A_1) P(X_{i_2} \in A_2) \cdots P(X_{i_r} \in A_r).$$

If X and Y are independent, the law of $X + Y$ is given by *convolution of measures*

$$p_{X+Y} = p_X * p_Y, \text{ where } (p_X * p_Y)(A) = \int_{\mathbb{R}^d} p_X(A - y) p_Y(dy).$$

Equivalently

$$\int_{\mathbb{R}^d} g(y)(p_X * p_Y)(dy) = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} g(x + y) p_X(dx) p_Y(dy),$$

for all $g \in B_b(\mathbb{R}^d)$ (the bounded Borel measurable functions on $\mathbb{R}^d$).

If X and Y are independent with densities f_X and f_Y, respectively, then $X + Y$ has density f_{X+Y} given by *convolution of functions*:

$$f_{X+Y} = f_X * f_Y, \text{ where } (f_X * f_Y)(x) = \int_{\mathbb{R}^d} f_X(x - y) f_Y(y) dy.$$

The *characteristic function* of X is $\phi_X : \mathbb{R}^d \to \mathbb{C}$, where

$$\phi_X(u) = \int_{\mathbb{R}^d} e^{i(u,x)} p_X(dx).$$

Theorem 1.1 (Kac's theorem) *$X_1, \ldots, X_n$ are independent if and only if*

$$\mathbb{E}\left(\exp\left(i\sum_{j=1}^{n}(u_j, X_j)\right)\right) = \phi_{X_1}(u_1)\cdots\phi_{X_n}(u_n)$$

for all $u_1, \ldots, u_n \in \mathbb{R}^d$.

More generally, the characteristic function of a probability measure μ on $\mathbb{R}^d$ is

$$\phi_\mu(u) = \int_{\mathbb{R}^d} e^{i(u,x)} \mu(dx).$$

Important properties are:-

1. $\phi_\mu(0) = 1$.
2. ϕ_μ is *positive definite* i.e. $\sum_{i,j} c_i \bar{c_j} \phi_\mu(u_i - u_j) \geq 0$, for all $c_i \in \mathbb{C}, u_i \in \mathbb{R}^d, 1 \leq i, j \leq n, n \in \mathbb{N}$.
3. ϕ_μ is uniformly continuous - Hint: Look at $|\phi_\mu(u+h) - \phi_\mu(u)|$ and use dominated convergence)).

Conversely *Bochner's theorem* states that if $\phi : \mathbb{R}^d \to \mathbb{C}$ satisfies (1), (2) and is continuous at $u = 0$, then it is the characteristic function of some probability measure μ on $\mathbb{R}^d$.

$\psi : \mathbb{R}^d \to \mathbb{C}$ is *conditionally positive definite* if for all $n \in \mathbb{N}$ and $c_1, \ldots, c_n \in \mathbb{C}$ for which $\sum_{j=1}^{n} c_j = 0$ we have

$$\sum_{j,k=1}^{n} c_j \bar{c}_k \psi(u_j - u_k) \geq 0,$$

for all $u_1, \ldots, u_n \in \mathbb{R}^d$. $\psi : \mathbb{R}^d \to \mathbb{C}$ will be said to be *hermitian* if $\overline{\psi(u)} = \psi(-u)$, for all $u \in \mathbb{R}^d$.

Theorem 1.2 (Schoenberg correspondence) *$\psi : \mathbb{R}^d \to \mathbb{C}$ is hermitian and conditionally positive definite if and only if $e^{t\psi}$ is positive definite for each $t > 0$.*

Proof. We only give the easy part here.

Suppose that $e^{t\psi}$ is positive definite for all $t > 0$. Fix $n \in \mathbb{N}$ and choose $c_1, \ldots, c_n$ and $u_1, \ldots, u_n$ as above. We then find that for each $t > 0$,

$$\frac{1}{t} \sum_{j,k=1}^{n} c_j \bar{c_k} (e^{t\psi(u_j - u_k)} - 1) \geq 0,$$

and so

$$\sum_{j,k=1}^{n} c_j \bar{c_k} \psi(u_j - u_k) = \lim_{t \to 0} \frac{1}{t} \sum_{j,k=1}^{n} c_j \bar{c_k} (e^{t\psi(u_j - u_k)} - 1) \geq 0.$$

□

1.2 Infinite Divisibility

We study this first because a Lévy process is infinite divisibility in motion, i.e. infinite divisibility is the underlying probabilistic idea which a Lévy process embodies dynamically.

Let μ be a probability measure on $\mathbb{R}^d$. Define $\mu^{*^n} = \mu * \cdots * \mu$ (n times). We say that μ has a *convolution* n*th root*, if there exists a probability measure $\mu^{\frac{1}{n}}$ for which $(\mu^{\frac{1}{n}})^{*^n} = \mu$.

μ is *infinitely divisible* if it has a convolution nth root for all $n \in \mathbb{N}$. In this case $\mu^{\frac{1}{n}}$ is unique.

Theorem 1.3 *μ is infinitely divisible iff for all $n \in \mathbb{N}$, there exists a probability measure μ_n with characteristic function ϕ_n such that*

$$\phi_\mu(u) = (\phi_n(u))^n,$$

for all $u \in \mathbb{R}^d$. Moreover $\mu_n = \mu^{\frac{1}{n}}$.

Proof. If μ is infinitely divisible, take $\phi_n = \phi_{\mu^{\frac{1}{n}}}$. Conversely, for each $n \in \mathbb{N}$, by Fubini's theorem,

$$\begin{aligned} \phi_\mu(u) &= \int_{\mathbb{R}^d} \cdots \int_{\mathbb{R}^d} e^{i(u, y_1 + \cdots + y_n)} \mu_n(dy_1) \cdots \mu_n(dy_n) \\ &= \int_{\mathbb{R}^d} e^{i(u,y)} \mu_n^{*^n}(dy). \end{aligned}$$

But $\phi_\mu(u) = \int_{\mathbb{R}^d} e^{i(u,y)} \mu(dy)$ and ϕ determines μ uniquely. Hence $\mu = \mu_n^{*^n}$. □

- If μ and ν are each infinitely divisible, then so is $\mu * \nu$.

- If $(\mu_n, n \in \mathbb{N})$ are infinitely divisible and $\mu_n \stackrel{w}{\Rightarrow} \mu$, then μ is infinitely divisible.

[Note: *Weak convergence.* $\mu_n \stackrel{w}{\Rightarrow} \mu$ means

$$\lim_{n \to \infty} \int_{\mathbb{R}^d} f(x) \mu_n(dx) = \int_{\mathbb{R}^d} f(x) \mu(dx),$$

for each $f \in C_b(\mathbb{R}^d)$.]

A random variable X is *infinitely divisible* if its law p_X is infinitely divisible, e.g. $X \stackrel{d}{=} Y_1^{(n)} + \cdots + Y_n^{(n)}$, where $Y_1^{(n)}, \ldots, Y_n^{(n)}$ are i.i.d., for each $n \in \mathbb{N}$.

1.2.1 Examples of Infinite Divisibility

In the following, we will demonstrate infinite divisibility of a random variable X by finding i.i.d. $Y_1^{(n)}, \ldots, Y_n^{(n)}$ such that $X \overset{d}{=} Y_1^{(n)} + \cdots + Y_n^{(n)}$, for each $n \in \mathbb{N}$.

Example 1 - Gaussian Random Variables

Let $X = (X_1, \ldots, X_d)$ be a random vector.

We say that it is $(non-degenerate) Gaussian$ if it there exists a vector $m \in \mathbb{R}^d$ and a strictly positive-definite symmetric $d \times d$ matrix A such that X has a pdf (probability density function) of the form:-

$$f(x) = \frac{1}{(2\pi)^{\frac{d}{2}} \sqrt{\det(A)}} \exp\left(-\frac{1}{2}(x-m, A^{-1}(x-m))\right), \qquad (1.1)$$

for all $x \in \mathbb{R}^d$.

In this case we will write $X \sim N(m, A)$. The vector m is the mean of X, so $m = \mathbb{E}(X)$ and A is the covariance matrix so that $A = \mathbb{E}((X-m)(X-m)^T)$. A standard calculation yields

$$\phi_X(u) = \exp\{i(m,u) - \frac{1}{2}(u, Au)\}, \qquad (1.2)$$

and hence

$$(\phi_X(u))^{\frac{1}{n}} = \exp\left\{ i\left(\frac{m}{n}, u\right) - \frac{1}{2}\left(u, \frac{1}{n}Au\right)\right\},$$

so we see that X is infinitely divisible with each $Y_j^{(n)} \sim N(\frac{m}{n}, \frac{1}{n}A)$ for each $1 \leq j \leq n$.

We say that X is a *standard normal* whenever $X \sim N(0, \sigma^2 I)$ for some $\sigma > 0$.

We say that X is *degenerate Gaussian* if (1.2) holds with $\det(A) = 0$, and these random variables are also infinitely divisible.

Example 2 - Poisson Random Variables

In this case, we take $d = 1$ and consider a random variable X taking values in the set $n \in \mathbb{N} \cup \{0\}$. We say that is *Poisson* if there exists $c > 0$ for which

$$P(X = n) = \frac{c^n}{n!} e^{-c}.$$

In this case we will write $X \sim \pi(c)$. We have $\mathbb{E}(X) = \text{Var}(X) = c$. It is easy to verify that

$$\phi_X(u) = \exp[c(e^{iu} - 1)],$$

from which we deduce that X is infinitely divisible with each $Y_j^{(n)} \sim \pi(\frac{c}{n})$, for $1 \leq j \leq n, n \in \mathbb{N}$.

Example 3 - Compound Poisson Random Variables

Let $(Z(n), n \in \mathbb{N})$ be a sequence of i.i.d. random variables taking values in $\mathbb{R}^d$ with common law μ_Z and let $N \sim \pi(c)$ be a Poisson random variable which is independent of all the $Z(n)$'s. The *compound Poisson random variable* X is defined as follows:-

$$X = Z(1) + \cdots + Z(N).$$

Proposition 1.1 *For each* $u \in \mathbb{R}^d$,

$$\phi_X(u) = \exp\left[\int_{\mathbb{R}^d}(e^{i(u,y)} - 1)c\mu_Z(dy)\right].$$

Proof. Let ϕ_Z be the common characteristic function of the Z_n's. By conditioning and using independence we find,

$$\begin{aligned}
\phi_X(u) &= \sum_{n=0}^{\infty}\mathbb{E}(e^{i(u,Z(1)+\cdots+Z(N))}|N = n)P(N = n)\\
&= \sum_{n=0}^{\infty}\mathbb{E}(e^{i(u,Z(1))+\cdots+Z(n))})e^{-c}\frac{c^n}{n!}\\
&= e^{-c}\sum_{n=0}^{\infty}\frac{[c\phi_Z(u)]^n}{n!}\\
&= \exp[c(\phi_Z(u) - 1)],
\end{aligned}$$

and the result follows on writing $\phi_Z(u) = \int e^{i(u,y)}\mu_Z(dy)$. □

If X is compound Poisson as above, we write $X \sim \pi(c, \mu_Z)$. It is clearly infinitely divisible with each $Y_j^{(n)} \sim \pi(\frac{c}{n}, \mu_Z)$, for $1 \leq j \leq n$.

1.2.2 The Lévy-Khintchine Formula

de Finetti (1920's) suggested that the most general infinitely divisible random variable could be written $X = Y + W$, where Y and W are independent, $Y \sim N(m, A), W \sim \pi(c, \mu_Z)$. Then $\phi_X(u) = e^{\eta(u)}$, where

$$\eta(u) = i(m, u) - \frac{1}{2}(u, Au) + \int_{\mathbb{R}^d}(e^{i(u,y)} - 1)c\mu_Z(dy). \tag{1.3}$$

This is WRONG! $\nu(\cdot) = c\mu_Z(\cdot)$ is a finite measure here. Lévy and Khintchine showed that ν can be σ-finite, provided it is what is now called a *Lévy measure* on $\mathbb{R}^d - \{0\} = \{x \in \mathbb{R}^d, x \neq 0\}$, i.e.

$$\int(|y|^2 \wedge 1)\nu(dy) < \infty, \tag{1.4}$$

(where $a \wedge b := \min\{a, b\}$, for $a, b \in \mathbb{R}$). Since $|y|^2 \wedge \epsilon \leq |y|^2 \wedge 1$ whenever $0 < \epsilon \leq 1$, it follows from (1.4) that

$$\nu((-\epsilon, \epsilon)^c) < \infty \quad \text{for all } \epsilon > 0.$$

Here is the fundamental result of this lecture:-

Theorem 1.4 (Lévy-Khintchine) *A Borel probability measure μ on $\mathbb{R}^d$ is infinitely divisible if there exists a vector $b \in \mathbb{R}^d$, a non-negative symmetric $d \times d$ matrix A and a Lévy measure ν on $\mathbb{R}^d - \{0\}$ such that for all $u \in \mathbb{R}^d$,*

$$\phi_\mu(u) = \exp\left[i(b, u) - \frac{1}{2}(u, Au) + \int_{\mathbb{R}^d - \{0\}} (e^{i(u,y)} - 1 - i(u, y)\chi_{\hat{B}}(y))\nu(dy)\right]. \tag{1.5}$$

where $\hat{B} = B_1(0) = \{y \in \mathbb{R}^d; |y| < 1\}$.

Conversely, any mapping of the form (1.5) is the characteristic function of an infinitely divisible probability measure on $\mathbb{R}^d$.

The triple (b, A, ν) is called the *characteristics* of the infinitely divisible random variable X. Define $\eta = \log \phi_\mu$, where we take the principal part of the logarithm. η is called the *Lévy symbol* or the *characteristic exponent.*

We're not going to prove this result here. To understand it, it is instructive to let $(U_n, n \in \mathbb{N})$ be a sequence of Borel sets in $B_1(0)$ with $U_n \downarrow \{0\}$. Observe that

$$\eta(u) = \lim_{n\to\infty} \eta_n(u) \quad \text{where each}$$

$$\eta_n(u) = i\left[\left(b - \int_{U_n^c \cap \hat{B}} y\nu(dy), u\right)\right] - \frac{1}{2}(u, Au) + \int_{U_n^c} (e^{i(u,y)} - 1)\nu(dy),$$

so η is in some sense (to be made more precise later) the limit of a sequence of sums of Gaussians and independent compound Poissons. Interesting phenomena appear in the limit as we'll see below. First, we classify Lévy symbols analytically:-

Theorem 1.5 *η is a Lévy symbol if and only if it is a continuous, hermitian conditionally positive definite function for which $\eta(0) = 0$.*

1.2.3 Stable Laws

This is one of the most important subclasses of infinitely divisible laws.

We consider the general central limit problem in dimension $d = 1$, so let $(Y_n, n \in \mathbb{N})$ be a sequence of real valued i.i.d. random variables and consider the rescaled partial sums

$$S_n = \frac{Y_1 + Y_2 + \cdots + Y_n - b_n}{\sigma_n},$$

where $(b_n, n \in \mathbb{N})$ is an arbitrary sequence of real numbers and $(\sigma_n, n \in \mathbb{N})$ an arbitrary sequence of positive numbers. We are interested in the case where there exists a random variable X for which

$$\lim_{n \to \infty} P(S_n \leq x) = P(X \leq x), \tag{1.6}$$

for all $x \in \mathbb{R}$ i.e. $(S_n, n \in \mathbb{N})$ converges in distribution to X. If each $b_n = nm$ and $\sigma_n = \sqrt{n}\sigma$ for fixed $m \in R, \sigma > 0$ then $X \sim N(m, \sigma^2)$ by the usual Laplace - de-Moivre central limit theorem.

More generally a random variable is said to be *stable* if it arises as a limit as in (1.6). It is not difficult to show that (1.6) is equivalent to the following:-

There exist real valued sequences $(c_n, n \in \mathbb{N})$ and $(d_n, n \in \mathbb{N})$ with each $c_n > 0$ such that

$$X_1 + X_2 + \cdots + X_n \stackrel{d}{=} c_n X + d_n \tag{1.7}$$

where $X_1, \ldots, X_n$ are independent copies of X. X is said to be *strictly stable* if each $d_n = 0$.

To see that (1.7) $\Rightarrow$ (1.6) take each $Y_j = X_j, b_n = d_n$ and $\sigma_n = c_n$. In fact it can be shown that the only possible choice of c_n in (1.7) is $c_n = \sigma n^{\frac{1}{\alpha}}$, where $0 < \alpha \leq 2$ and $\sigma > 0$. The parameter α plays a key role in the investigation of stable random variables and is called the *index of stability.*

Note that (1.7) can also be expressed in the equivalent form

$$\phi_X(u)^n = e^{iud_n} \phi_X(c_n u),$$

for each $u \in \mathbb{R}$.

It follows immediately from (1.7) that all stable random variables are infinitely divisible and the characteristics in the Lévy-Khintchine formula are given by the following result.

Theorem 1.6 *If X is a stable real-valued random variable, then its characteristics must take one of the two following forms.*

1. *When* $\alpha = 2, \quad \nu = 0$ *(so* $X \sim N(b, A)$*).*
2. *When* $\alpha \neq 2, \quad A = 0$ *and* $\nu(dx) = \frac{c_1}{x^{1+\alpha}}\chi_{(0,\infty)}(x)dx + \frac{c_2}{|x|^{1+\alpha}}\chi_{(-\infty,0)}(x)dx$*,*

   where $c_1 \geq 0, c_2 \geq 0$ *and* $c_1 + c_2 > 0$*.*

A careful transformation of the integrals in the Lévy-Khintchine formula gives a different form for the characteristic function which is often more convenient.

Theorem 1.7 *A real-valued random variable* X *is stable if and only if there exists* $\sigma > 0, -1 \leq \beta \leq 1$ *and* $\mu \in \mathbb{R}$ *such that for all* $u \in \mathbb{R}$*,*

1.
$$\phi_X(u) = \exp\left[i\mu u - \frac{1}{2}\sigma^2 u^2\right] \quad \text{when} \quad \alpha = 2.$$

2.
$$\phi_X(u) = \exp\left[i\mu u - \sigma^\alpha |u|^\alpha (1 - i\beta sgn(u)\tan(\frac{\pi\alpha}{2}))\right] \quad \text{when} \quad \alpha \neq 1, 2.$$

3.
$$\phi_X(u) = \exp\left[i\mu u - \sigma|u|(1 + i\beta\frac{2}{\pi}sgn(u)\log(|u|))\right] \quad \text{when} \quad \alpha = 1.$$

It can be shown that $\mathbb{E}(X^2) < \infty$ if and only if $\alpha = 2$ (i.e. X is Gaussian) and $\mathbb{E}(|X|) < \infty$ if and only if $1 < \alpha \leq 2$.

All stable random variables have densities f_X, which can in general be expressed in series form. In three important cases, there are closed forms.

1. **The Normal Distribution**

$$\alpha = 2, \quad X \sim N(\mu, \sigma^2).$$

2. **The Cauchy Distribution**

$$\alpha = 1, \beta = 0 \quad f_X(x) = \frac{\sigma}{\pi[(x-\mu)^2 + \sigma^2]}.$$

3. **The Lévy Distribution**

$$\alpha = \frac{1}{2}, \beta = 1 \quad f_X(x) = \left(\frac{\sigma}{2\pi}\right)^{\frac{1}{2}} \frac{1}{(x-\mu)^{\frac{3}{2}}} \exp\left(-\frac{\sigma}{2(x-\mu)}\right), \quad \text{for } x > \mu.$$

In general the series representations are given in terms of a real valued parameter λ.

For $x > 0$ and $0 < \alpha < 1$:

$$f_X(x, \lambda) = \frac{1}{\pi x} \sum_{k=1}^{\infty} \frac{\Gamma(k\alpha + 1)}{k!} (-x^{-\alpha})^k \sin\left(\frac{k\pi}{2}(\lambda - \alpha)\right)$$

For $x > 0$ and $1 < \alpha < 2$,

$$f_X(x, \lambda) = \frac{1}{\pi x} \sum_{k=1}^{\infty} \frac{\Gamma(k\alpha^{-1} + 1)}{k!} (-x)^k \sin\left(\frac{k\pi}{2\alpha}(\lambda - \alpha)\right)$$

In each case the formula for negative x is obtained by using

$$f_X(-x, \lambda) = f_X(x, -\lambda), \quad \text{for } x > 0.$$

Note that if a stable random variable is symmetric then Theorem 1.7 yields

$$\phi_X(u) = \exp(-\rho^\alpha |u|^\alpha) \ \text{ for all } 0 < \alpha \leq 2, \tag{1.8}$$

where $\rho = \sigma(0 < \alpha < 2)$ and $\rho = \frac{\sigma}{\sqrt{2}}$, when $\alpha = 2$, and we will write $X \sim S\alpha S$ in this case.

One of the reasons why stable laws are so important in applications is the nice decay properties of the tails. The case $\alpha = 2$ is special in that we have exponential decay, indeed for a standard normal X there is the elementary estimate

$$P(X > y) \sim \frac{e^{-\frac{1}{2}y^2}}{\sqrt{2\pi}y} \ \text{ as } y \to \infty,$$

When $\alpha \neq 2$ we have the slower polynomial decay as expressed in the following,

$$\lim_{y\to\infty} y^\alpha P(X > y) = C_\alpha \frac{1+\beta}{2} \sigma^\alpha,$$

$$\lim_{y\to\infty} y^\alpha P(X < -y) = C_\alpha \frac{1-\beta}{2} \sigma^\alpha,$$

where $C_\alpha > 1$. The relatively slow decay of the tails for non-Gaussian stable laws ("heavy tails") makes them ideally suited for modelling a wide range of interesting phenomena, some of which exhibit "long-range dependence". Deeper mathematical investigations of heavy tails require the mathematical technique of *regular variation.*

The generalisation of stability to random vectors is straightforward - just replace $X_1, \ldots, X_n$, X and each d_n in (1.7) by vectors and the formula in

Theorem 1.6 extends directly. Note however that when $\alpha \neq 2$ in the random vector version of Theorem 1.6, the Lévy measure takes the form

$$\nu(dx) = \frac{c}{|x|^{d+\alpha}} dx$$

where $c > 0$.

We can generalise the definition of stable random variables if we weaken the conditions on the random variables $(Y(n), n \in \mathbb{N})$ in the general central limit problem by requiring these to be independent, but no longer necessarily identically distributed. In this case (subject to a technical growth restriction) the limiting random variables are called *self-decomposable* (or of *class L*) and they are also infinitely divisible. Alternatively a random variable X is self-decomposable if and only if for each $0 < a < 1$, there exists a random variable Y_a which is independent of X such that

$$X \stackrel{d}{=} aX + Y_a \Leftrightarrow \phi_X(u) = \phi_X(au)\phi_{Y_a}(u),$$

for all $u \in \mathbb{R}^d$. An infinitely divisible law is self-decomposable if and only if the Lévy measure is of the form:

$$\nu(dx) = \frac{k(x)}{|x|} dx,$$

where k is decreasing on $(0, \infty)$ and increasing on $(-\infty, 0)$. There has recently been increasing interest in these distributions both from a theoretical and applied perspective. Examples include gamma, Pareto, Student-t, F and log-normal distributions.

2 Lévy Processes

Let $X = (X(t), t \geq 0)$ be a stochastic process defined on a probability space $(\Omega, \mathcal{F}, P)$. We say that it has *independent increments* if for each $n \in \mathbb{N}$ and each $0 \leq t_1 < t_2 < \cdots < t_{n+1} < \infty$, the random variables $(X(t_{j+1}) - X(t_j), 1 \leq j \leq n)$ are independent and it has *stationary increments* if each $X(t_{j+1}) - X(t_j) \stackrel{d}{=} X(t_{j+1} - t_j) - X(0)$.

We say that X is a *Lévy process* if

(L1) Each $X(0) = 0$ (a.s),

(L2) X has independent and stationary increments, (L3) X is *stochastically continuous* i.e. for all $a > 0$ and for all $s \geq 0$,

$$\lim_{t \to s} P(|X(t) - X(s)| > a) = 0.$$

Note that in the presence of (L1) and (L2), (L3) is equivalent to the condition

$$\lim_{t \downarrow 0} P(|X(t)| > a) = 0.$$

The *sample paths* of a process are the maps $t \to X(t)(\omega)$ from $\mathbb{R}^+$ to $\mathbb{R}^d$, for each $\omega \in \Omega$.

We are now going to explore the relationship between Lévy processes and infinite divisibility.

Proposition 2.1 *If X is a Lévy process, then $X(t)$ is infinitely divisible for each $t \geq 0$.*

Proof. For each $n \in \mathbb{N}$, we can write

$$X(t) = Y_1^{(n)}(t) + \cdots + Y_n^{(n)}(t)$$

where each $Y_k^{(n)}(t) = X(\frac{kt}{n}) - X(\frac{(k-1)t}{n})$. The $Y_k^{(n)}(t)$'s are i.i.d. by (L2). □

By Proposition 2.1, we can write $\phi_{X(t)}(u) = e^{\eta(t,u)}$ for each $t \geq 0, u \in \mathbb{R}^d$, where each $\eta(t, \cdot)$ is a Lévy symbol.

Theorem 2.1 *If X is a Lévy process, then*

$$\phi_{X(t)}(u) = e^{t\eta(u)},$$

for each $u \in \mathbb{R}^d, t \geq 0$, where η is the Lévy symbol of $X(1)$.

Proof. Suppose that X is a Lévy process and for each $u \in \mathbb{R}^d, t \geq 0$, define $\phi_u(t) = \phi_{X(t)}(u)$ then by (L2) we have for all $s \geq 0$,

$$\begin{aligned}
\phi_u(t+s) &= \mathbb{E}(e^{i(u,X(t+s))}) \\
&= \mathbb{E}(e^{i(u,X(t+s)-X(s))}e^{i(u,X(s))}) \\
&= \mathbb{E}(e^{i(u,X(t+s)-X(s))})\mathbb{E}(e^{i(u,X(s))}) \\
&= \phi_u(t)\phi_u(s) \ldots \text{(i)}
\end{aligned}$$

Now $\phi_u(0) = 1 \ldots$ (ii) by (L1), and the map $t \to \phi_u(t)$ is continuous. However the unique continuous solution to (i) and (ii) is given by $\phi_u(t) =$

$e^{t\alpha(u)}$, where $\alpha : \mathbb{R}^d \to \mathbb{C}$. Now by Proposition 2.1, $X(1)$ is infinitely divisible, hence α is a Lévy symbol and the result follows. $\square$

We now have the Lévy-Khinchine formula for a Lévy process $X = (X(t), t \geq 0)$:-

$$\mathbb{E}(e^{i(u,X(t))}) = \exp\left(t\left[i(b,u) - \frac{1}{2}(u, Au) + \int_{\mathbb{R}^d - \{0\}} (e^{i(u,y)} - 1 - i(u,y)\chi_{\hat{B}}(y))\nu(dy)\right]\right), \tag{2.9}$$

for each $t \geq 0, u \in \mathbb{R}^d$, where (b, A, ν) are the characteristics of $X(1)$.

We will define the Lévy symbol and the characteristics of a Lévy process X to be those of the random variable $X(1)$. We will sometimes write the former as η_X when we want to emphasise that it belongs to the process X.

Let p_t be the law of $X(t)$, for each $t \geq 0$. By (L2), we have for all $s, t \geq 0$ that:

$$p_{t+s} = p_t * p_s.$$

By (L3), we have $p_t \stackrel{w}{\to} \delta_0$ as $t \to 0$, i.e. $\lim_{t \to 0} f(x)p_t(dx) = f(0)$.

$(p_t, t \geq 0)$ is a *weakly continuous convolution semigroup of probability measures* on $\mathbb{R}^d$. Conversely, given any such semigroup, we can always construct a Lévy process on path space via Kolmogorov's construction.

Informally, we have the following asymptotic relationship between the law of a Lévy process and its Lévy measure:

$$\nu = \lim_{t \downarrow 0} \frac{p_t}{t}.$$

More precisely

$$\lim_{t \downarrow 0} \frac{1}{t} \int_{\mathbb{R}^d} f(x)p_t(dx) = \int_{\mathbb{R}^d} f(x)\nu(dx), \tag{2.10}$$

for bounded, continuous functions f which vanish in some neighborhood of the origin.

2.1 Examples of Lévy Processes

Example 1, Brownian Motion and Gaussian Processes

A *(standard) Brownian motion* in $\mathbb{R}^d$ is a Lévy process $B = (B(t), t \geq 0)$ for which

(B1) $B(t) \sim N(0, tI)$ for each $t \geq 0$,
(B2) B has continuous sample paths.

It follows immediately from (B1) that if B is a standard Brownian motion, then its characteristic function is given by

$$\phi_{B(t)}(u) = \exp\{-\frac{1}{2}t|u|^2\},$$

for each $u \in \mathbb{R}^d, t \geq 0$.

We introduce the marginal processes $B_i = (B_i(t), t \geq 0)$ where each $B_i(t)$ is the ith component of $B(t)$, then it is not difficult to verify that the B_i's are mutually independent Brownian motions in $\mathbb{R}$. We will call these *one-dimensional Brownian motions* in the sequel.

Brownian motion has been the most intensively studied Lévy process. In the early years of the twentieth century, it was introduced as a model for the physical phenomenon of Brownian motion by Einstein and Smoluchowski and as a description of the dynamical evolution of stock prices by Bachelier. The theory was placed on a rigorous mathematical basis by Norbert Wiener in the 1920's.

We could try to use the Kolmogorov existence theorem to construct one-dimensional Brownian motion from the following prescription on cylinder sets of the form $I^H_{t_1,\ldots,t_n} = \{\omega \in \Omega; \omega(t_1) \in [a_1, b_1], \ldots, \omega(t_n) \in [a_n, b_n]\}$ where $H = [a_1, b_1] \times \cdots [a_n, b_n]$ and we have taken Ω to be the set of all mappings from $\mathbb{R}^+$ to $\mathbb{R}$:

$$\begin{aligned} & P(I^H_{t_1,\ldots,t_n}) \\ = & \int_H \frac{1}{(2\pi)^{\frac{n}{2}}\sqrt{t_1(t_2 - t_1)\ldots(t_n - t_{n-1})}} \exp\left(-\frac{1}{2}\left(\frac{x_1^2}{t_1} + \frac{(x_2 - x_1)^2}{t_2 - t_1} + \cdots \right.\right. \\ + & \left.\left. \frac{(x_n - x_{n-1})^2}{t_n - t_{n-1}}\right)\right) dx_1 \cdots dx_n. \end{aligned}$$

However there there is then no guarantee that the paths are continuous. The literature contains a number of ingenious methods for constructing Brownian motion. One of the most delightful of these (originally due to Paley and Wiener) obtains this, in the case $d = 1$, as a random Fourier series

$$B(t) = \frac{\sqrt{2}}{\pi}\sum_{n=0}^{\infty}\frac{\sin(\pi t(n + \frac{1}{2}))}{n + \frac{1}{2}}\xi(n),$$

for each $t \geq 0$, where $(\xi(n), n \in \mathbb{N})$ is a sequence of i.i.d. $N(0,1)$ random variables.

We list a number of useful properties of Brownian motion in the case $d = 1$.

- Brownian motion is locally Hölder continuous with exponent α for every $0 < \alpha < \frac{1}{2}$ i.e. for every $T > 0, \omega \in \Omega$ there exists $K = K(T, \omega)$ such that
$$|B(t)(\omega) - B(s)(\omega)| \leq K|t - s|^{\alpha},$$
for all $0 \leq s < t \leq T$.

- The sample paths $t \to B(t)(\omega)$ are almost surely nowhere differentiable.

- For any sequence, $(t_n, n \in \mathbb{N})$ in $\mathbb{R}^+$ with $t_n \uparrow \infty$,
$$\liminf_{n\to\infty} B(t_n) = -\infty \quad \text{a.s.} \qquad \limsup_{n\to\infty} B(t_n) = \infty \quad \text{a.s.}$$

- The law of the iterated logarithm:-
$$P\left(\limsup_{t\downarrow 0} \frac{B(t)}{(2t\log(\log(\frac{1}{t})))^{\frac{1}{2}}} = 1\right) = 1.$$

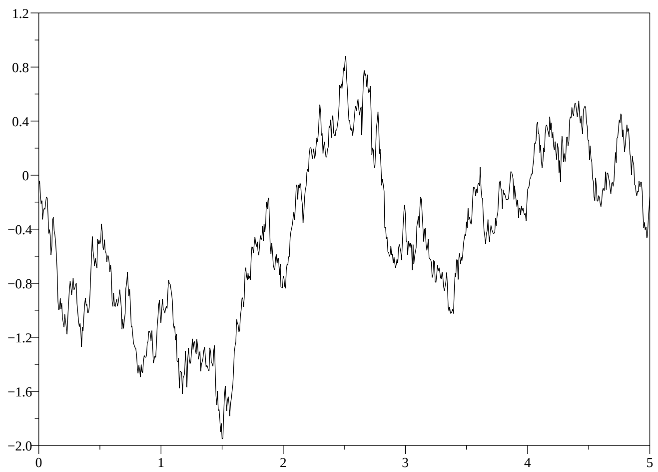

Figure 1 Simulation of standard Brownian motion

Let A be a non-negative symmetric $d \times d$ matrix and let σ be a square root of A so that σ is a $d \times m$ matrix for which $\sigma\sigma^T = A$. Now let $b \in \mathbb{R}^d$ and let B be a Brownian motion in R^m. We construct a process $C = (C(t), t \geq 0)$ in $\mathbb{R}^d$ by

$$C(t) = bt + \sigma B(t), \tag{2.11}$$

then C is a Lévy process with each $C(t) \sim N(tb, tA)$. It is not difficult to verify that C is also a Gaussian process, i.e. all its finite dimensional distributions are Gaussian. It is sometimes called *Brownian motion with drift*. The Lévy symbol of C is

$$\eta_C(u) = i(b, u) - \frac{1}{2}(u, Au).$$

In fact a Lévy process has continuous sample paths if and only if it is of the form (2.11).

Example 2 - The Poisson Process

The Poisson process of intensity $\lambda > 0$ is a Lévy process N taking values in $\mathbb{N} \cup \{0\}$ wherein each $N(t) \sim \pi(\lambda t)$ so we have

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

for each $n = 0, 1, 2, \ldots$. The Poisson process is widely used in applications and there is a wealth of literature concerning it and its generalisations. We define non-negative random variables $(T_n, \mathbb{N} \cup \{0\})$ (usually called waiting times) by $T_0 = 0$ and for $n \in \mathbb{N}$,

$$T_n = \inf\{t \geq 0; N(t) = n\},$$

then it is well known that the T_n's are gamma distributed. Moreover, the inter-arrival times $T_n - T_{n-1}$ for $n \in \mathbb{N}$ are i.i.d. and each has exponential distribution with mean $\frac{1}{\lambda}$. The sample paths of N are clearly piecewise constant with "jump" discontinuities of size 1 at each of the random times $(T_n, n \in \mathbb{N})$.

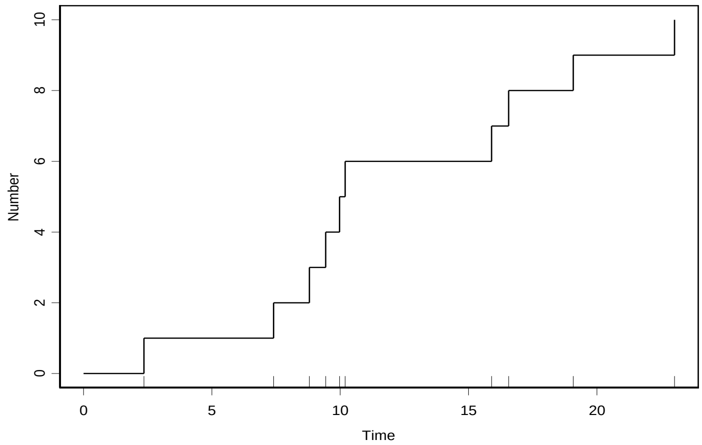

Figure 2. Simulation of a Poisson process ($\lambda = 0.5$)

For later work it is useful to introduce the *compensated Poisson process* $\tilde{N} = (\tilde{N}(t), t \geq 0)$ where each $\tilde{N}(t) = N(t) - \lambda t$. Note that $\mathbb{E}(\tilde{N}(t)) = 0$ and $\mathbb{E}(\tilde{N}(t)^2) = \lambda t$ for each $t \geq 0$.

Example 3 - The Compound Poisson Process

Let $(Z(n), n \in \mathbb{N})$ be a sequence of i.i.d. random variables taking values in $\mathbb{R}^d$ with common law μ_Z and let N be a Poisson process of intensity λ which is independent of all the $Z(n)$'s. The *compound Poisson process* Y is defined as follows:-

$$Y(t) = Z(1) + \ldots + Z(N(t)), \tag{2.12}$$

for each $t \geq 0$, so each $Y(t) \sim \pi(\lambda t, \mu_Z)$.

By Proposition 1.1 we see that Y has Lévy symbol

$$\eta_Y(u) = \left[\int (e^{i(u,y)} - 1)\lambda \mu_Z(dy) \right].$$

Again the sample paths of Y are piecewise constant with "jump discontinuities" at the random times $(T(n), n \in \mathbb{N})$, however this time the size of the jumps is itself random, and the jump at $T(n)$ can be any value in the range of the random variable $Z(n)$.

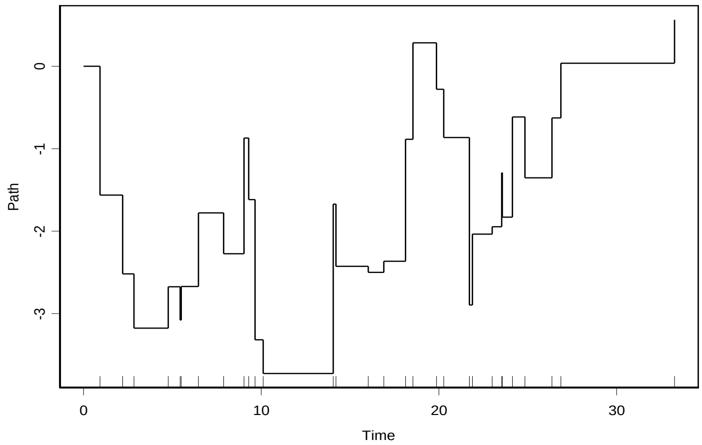

Figure 3. Simulation of a compound Poisson process with $N(0,1)$ summands($\lambda = 1$).

Example 4 - Interlacing Processes

Let C be a Gaussian Lévy process as in Example 1 and Y be a compound Poisson process as in Example 3, which is independent of C. Define a new process X by

$$X(t) = C(t) + Y(t),$$

for all $t \geq 0$, then it is not difficult to verify that X is a Lévy process with Lévy symbol of the form (1.3). Using the notation of Examples 2 and 3, we see that the paths of X have jumps of random size occurring at random times. In fact we have,

$$\begin{aligned} X(t) &= C(t) \quad \text{for } 0 \leq t < T_1, \\ &= C(T_1) + Z_1 \quad \text{when } t = T_1, \\ &= X(T_1) + C(t) - C(T_1) \quad \text{for } T_1 < t < T_2, \\ &= X(T_2-) + Z_2 \quad \text{when } t = T_2, \end{aligned}$$

and so on recursively. We call this procedure an *interlacing* as a continuous path process is “interlaced” with random jumps. From the remarks after Theorem 1.4, it seems reasonable that the most general Lévy process might arise as the limit of a sequence of such interlacings, and this can be established rigorously.

Example 5 - Stable Lévy Processes

A *stable Lévy process* is a Lévy process X in which the Lévy symbol is given by theorem 1.6. So, in particular, each $X(t)$ is a stable random variable.

Of particular interest is the rotationally invariant case whose Lévy symbol is given by

$$\eta(u) = -\sigma^{\alpha}|u|^{\alpha},$$

where α is the index of stability ($0 < \alpha \leq 2$). One of the reasons why these are important in applications is that they display self-similarity. In general, a stochastic process $Y = (Y(t), t \geq 0)$ is *self-similar with Hurst index* $H > 0$ if the two processes $(Y(at), t \geq 0)$ and $(a^H Y(t), t \geq 0)$ have the same finite-dimensional distributions for all $a \geq 0$. By examining characteristic functions, it is easily verified that a rotationally invariant stable Lévy process is self-similar with Hurst index $H = \frac{1}{\alpha}$, so that e.g. Brownian motion is self-similar with $H = \frac{1}{2}$. A Lévy process X is self-similar if and only if each $X(t)$ is strictly stable.

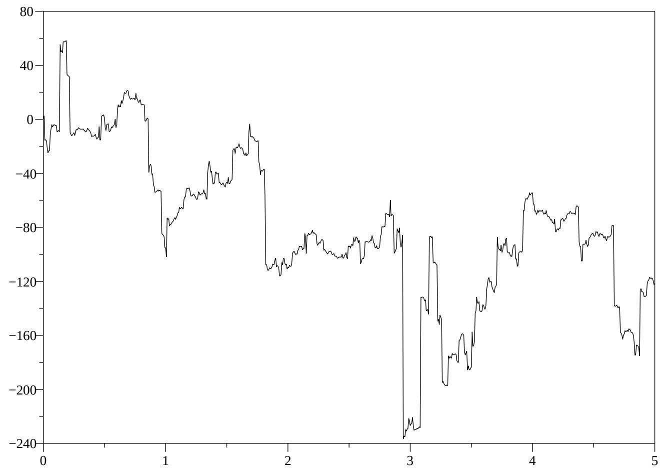

Figure 4 Simulation of the Cauchy process.

2.2 Densities of Lévy Processes

Question: When does a Lévy process have a density f_t for all $t > 0$ so that for all Borel sets B:

$$P(X_t \in B) = p_t(B) = \int_B f_t(x)dx.$$

In general, a random variable has a continuous density if its characteristic function is integrable and in this case, the density is the Fourier transform of the characteristic function. So for Lévy processes, if for all $t > 0$,

$$\int_{\mathbb{R}^d} |e^{t\eta(u)}| du = \int_{\mathbb{R}^d} e^{t\Re(\eta(u))} du < \infty$$

we have

$$f_t(x) = (2\pi)^{-d} \int_{\mathbb{R}^d} e^{t\eta(u) - i(u,x)} du.$$

For example ($d = 1$) if X is α-stable, it has a density since for all $1 \leq \alpha \leq 2$:

$$\int_{\mathbb{R}} e^{-t|u|^\alpha} du \leq \int_{\mathbb{R}} e^{-t|u|} du < \infty,$$

and for $0 \leq \alpha < 1$:

$$\int_{\mathbb{R}} e^{-t|u|^\alpha} du = \frac{2}{\alpha} \int_0^\infty e^{-ty} y^{\frac{1}{\alpha}-1} dy < \infty.$$

In $d = 1$ the following result giving a condition to have a density in terms of the Lévy measure is due to Orey:

Theorem 2.2 *A Lévy process X has a smooth density f_t for all $t > 0$ if*

$$\liminf_{r \downarrow 0} \frac{1}{r^{2-\beta}} \int_{-r}^{r} x^2 \nu(dx) > 0,$$

for some $0 < \beta < 2$.

A Lévy process has a *Lévy density* g_ν if its Lévy measure ν is absolutely continuous with respect to Lebesgue measure, then g_ν is defined to be the Radon-Nikodym derivative $\frac{d\nu}{dx}$.

Example. Let X be a compound Poisson process with each $X(t) = Y_1 + Y_2 + \cdots + Y_{N(t)}$ wherein each Y_j has a density f_Y, then $g_\nu = \lambda f_Y$ is the Lévy density.

We have $p_t(A) = e^{-\lambda t}\delta_0(A) + \int_A f_t^{ac}(x)dx$, where for $x \neq 0$

$$f_t^{ac}(x) = e^{-\lambda t} \sum_{n=1}^{\infty} \frac{(\lambda t)^n}{n!} f_Y^{*n}(x).$$

$f^{ac}(x)$ is the density of X conditioned on the fact that it jumps at least once between 0 and t. In this case, (2.10) takes the precise form (for $x \neq 0$)

$$g_\nu(x) = \lim_{t \downarrow 0} \frac{f_t^{ac}(x)}{t}.$$

2.3 Subordinators

A *subordinator* is a one-dimensional Lévy process which is increasing a.s. Such processes can be thought of as a random model of time evolution, since if $T = (T(t), t \geq 0)$ is a subordinator we have

$$T(t) \geq 0 \text{ for each } t > 0 \text{ a.s. } \quad \text{and } T(t_1) \leq T(t_2) \text{ whenever } t_1 \leq t_2 \text{ a.s.}$$

Now since for $X(t) \sim N(0, At)$ we have $P(X(t) \geq 0) = P(X(t) \leq 0) = \frac{1}{2}$, it is clear that such a process cannot be a subordinator. More generally we have

Theorem 2.3 *If T is a subordinator then its Lévy symbol takes the form*

$$\eta(u) = ibu + \int_{(0,\infty)} (e^{iuy} - 1)\lambda(dy), \tag{2.13}$$

where $b \geq 0$, and the Lévy measure λ satisfies the additional requirements

$$\lambda(-\infty, 0) = 0 \quad \textit{and} \quad \int_{(0,\infty)} (y \wedge 1)\lambda(dy) < \infty.$$

Conversely, any mapping from $\mathbb{R}^d \to \mathbb{C}$ of the form (2.13) is the Lévy symbol of a subordinator.

We call the pair (b, λ), the *characteristics* of the subordinator T.

Now for each $t \geq 0$, the map $u \to \mathbb{E}(e^{iuT(t)})$ can be analytically continued to the region $\{iu, u > 0\}$ and we then obtain the following expression for the Laplace transform of the distribution

$$\mathbb{E}(e^{-uT(t)}) = e^{-t\psi(u)},$$

$$\text{where } \psi(u) = -\eta(iu) = bu + \int_{(0,\infty)} (1 - e^{-uy})\lambda(dy) \tag{2.14}$$

for each $t, u \geq 0$. We note that this is much more useful for both theoretical and practical application than the characteristic function.

The function ψ is usually called the *Laplace exponent* of the subordinator.

Examples

(1) **The Poisson Case**

Poisson processes are clearly subordinators. More generally a compound Poisson process will be a subordinator if and only if the $Z(n)$'s in equation (2.12) are all $\mathbb{R}^+$ valued.

(2) **α-Stable Subordinators**

Using straightforward calculus, we find that for $0 < \alpha < 1$, $u \geq 0$,

$$u^\alpha = \frac{\alpha}{\Gamma(1-\alpha)} \int_0^\infty (1 - e^{-ux}) \frac{dx}{x^{1+\alpha}}.$$

Hence by (2.14), Theorem 2.3 and Theorem 1.6, we see that for each $0 < \alpha < 1$ there exists an α-stable subordinator T with Laplace exponent

$$\psi(u) = u^\alpha.$$

and the characteristics of T are $(0, \lambda)$ where $\lambda(dx) = \frac{\alpha}{\Gamma(1-\alpha)} \frac{dx}{x^{1+\alpha}}$.

Note that when we analytically continue this to obtain the Lévy symbol we obtain the form given in Theorem 1.7(2) with $\mu = 0, \beta = 1$ and $\sigma^\alpha = \cos\left(\frac{\alpha\pi}{2}\right)$.

(3) **The Lévy Subordinator**

The $\frac{1}{2}$-stable subordinator has a density given by the Lévy distribution (with $\mu = 0$ and $\sigma = \frac{t^2}{2}$)

$$f_{T(t)}(s) = \left(\frac{t}{2\sqrt{\pi}}\right) s^{-\frac{3}{2}} e^{-\frac{t^2}{4s}},$$

for $s \geq 0$. The Lévy subordinator has a nice probabilistic interpretation as a first hitting time for one-dimensional standard Brownian motion $(B(t), t \geq 0)$, more precisely

$$T(t) = \inf\left\{s > 0; B(s) = \frac{t}{\sqrt{2}}\right\}. \tag{2.15}$$

To show directly that for each $t \geq 0$,

$$\mathbb{E}(e^{-uT(t)}) = \int_0^\infty e^{-us} f_{T(t)}(s) ds = e^{-tu^{\frac{1}{2}}},$$

write $g_t(u) = \mathbb{E}(e^{-uT(t)})$. Differentiate with respect to u and make the substitution $x = \frac{t^2}{4us}$ to obtain the differential equation $g_t'(u) = -\frac{t}{2\sqrt{u}} g_t(u)$. Via the substitution $y = \frac{t}{2\sqrt{s}}$ we see that $g_t(0) = 1$ and the result follows.

(4) **Inverse Gaussian Subordinators**

We generalise the Lévy subordinator by replacing Brownian motion by the Gaussian process $C = (C(t), t \geq 0)$ where each $C(t) = B(t) + \mu t$ and $\mu \in \mathbb{R}$. The *inverse Gaussian subordinator* is defined by

$$T(t) = \inf\{s > 0; C(s) = \delta t\}$$

where $\delta > 0$ and is so-called since $t \to T(t)$ is the generalised inverse of a Gaussian process.

Using martingale methods, we can show that for each $t, u > 0$,

$$\mathbb{E}(e^{-uT(t)}) = e^{-t\delta(\sqrt{2u+\mu^2}-\mu)}, \tag{2.16}$$

In fact each $T(t)$ has a density:-

$$f_{T(t)}(s) = \frac{\delta t}{\sqrt{2\pi}} e^{\delta t \mu} s^{-\frac{3}{2}} \exp\left\{-\frac{1}{2}(t^2\delta^2 s^{-1} + \mu^2 s)\right\}, \tag{2.17}$$

for each $s, t \geq 0$.

In general any random variable with density $f_{T(1)}$ is called an *inverse Gaussian* and denoted as $\mathrm{IG}(\delta, \mu)$.

(5) **Gamma Subordinators**

Let $(T(t), t \geq 0)$ be a *gamma process* with parameters $a, b > 0$ so that each $T(t)$ has density

$$f_{T(t)}(x) = \frac{b^{at}}{\Gamma(at)} x^{at-1} e^{-bx},$$

for $x \geq 0$; then it is easy to verify that for each $u \geq 0$,

$$\int_0^\infty e^{-ux} f_{T(t)}(x)dx = \left(1 + \frac{u}{b}\right)^{-at} = \exp\left(-ta\log\left(1 + \frac{u}{b}\right)\right).$$

From here it is a straightforward exercise in calculus to show that

$$\int_0^\infty e^{-ux} f_{T(t)}(x)dx = \exp\left[-t\int_0^\infty (1 - e^{-ux})ax^{-1}e^{-bx}dx\right].$$

From this we see that $(T(t), t \geq 0)$ is a subordinator with $b = 0$ and $\lambda(dx) = ax^{-1}e^{-bx}dx$. Moreover $\psi(u) = a\log\left(1 + \frac{u}{b}\right)$ is the associated Bernstein function (see below).

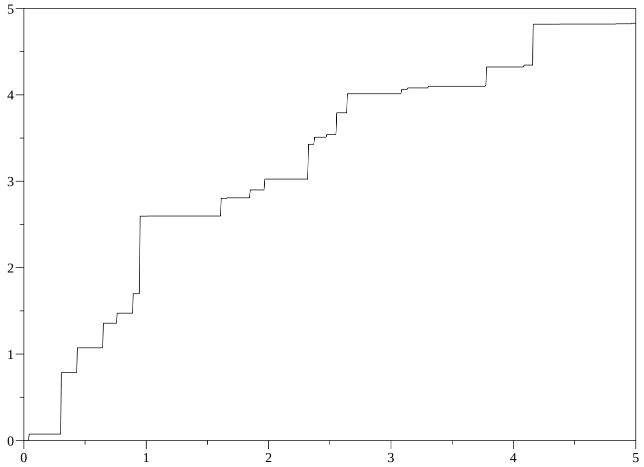

Figure 5 Simulation of a gamma subordinator.

Before we go further into the probabilistic properties of subordinators we'll make a quick diversion into analysis.

Let $f \in C^{\infty}((0,\infty))$. We say it is *completely monotone* if $(-1)^n f^{(n)} \geq 0$ for all $n \in \mathbb{N}$, and a *Bernstein function* if $f \geq 0$ and $(-1)^n f^{(n)} \leq 0$ for all $n \in \mathbb{N}$. We then have the following

Theorem 2.4 *1. f is a Bernstein function if and only if the mapping $x \to e^{-tf(x)}$ is completely monotone for all $t \geq 0$.*

2. f is a Bernstein function if and only if it has the representation

$$f(x) = a + bx + \int_0^\infty (1 - e^{-yx})\lambda(dy),$$

for all $x > 0$ where $a, b \geq 0$ and $\int_0^\infty (y \wedge 1)\lambda(dy) < \infty$.

3. g is completely monotone if and only if there exists a measure μ on $[0,\infty)$ for which

$$g(x) = \int_0^\infty e^{-xy}\mu(dy).$$

To interpret this theorem, first consider the case $a = 0$. In this case, if we compare the statement of Theorem 2.4 with equation (2.14), we see that there is a one to one correspondence between Bernstein functions for which

$\lim_{x\to 0} f(x) = 0$ and Laplace exponents of subordinators. The Laplace transforms of the laws of subordinators are always completely monotone functions and a subclass of all possible measures μ appearing in Theorem 2.4 (3) is given by all possible laws $p_{T(t)}$ associated to subordinators. A general Bernstein function with $a > 0$ can be given a probabilistic interpretation by means of "killing".

One of the most important probabilistic applications of subordinators is to "time change". Let X be an arbitrary Lévy process and let T be a subordinator defined on the same probability space as X such that X and T are independent. We define a new stochastic process $Z = (Z(t), t \geq 0)$ by the prescription

$$Z(t) = X(T(t)),$$

for each $t \geq 0$ so that for each $\omega \in \Omega, Z(t)(\omega) = X(T(t)(\omega))(\omega)$. The key result is then the following.

Theorem 2.5 *Z is a Lévy process.*

We compute the Lévy symbol of the subordinated process Z.

Proposition 2.2

$$\eta_Z = -\psi_T \circ (-\eta_X).$$

Proof. For each $u \in \mathbb{R}^d, t \geq 0$,

$$\begin{aligned}
\mathbb{E}(e^{i\eta_{Z(t)}(u)}) &= \mathbb{E}(e^{i(u,X(T(t)))}) \\
&= \int \mathbb{E}(e^{i(u,X(s))}) p_{T(t)}(ds) \\
&= \int e^{s\eta_X(u)} p_{T(t)}(ds) \\
&= \mathbb{E}(e^{-(-\eta_X(u))T(t)}) \\
&= e^{-t\psi_T(-\eta_X(u))}.
\end{aligned}$$
□

Example : From Brownian Motion to 2α-stable Processes

Let T be an α-stable subordinator (with $0 < \alpha < 1$) and X be a d-dimensional Brownian motion with covariance $A = 2I$, which is independent of T. Then for each $s \geq 0, u \in \mathbb{R}^d, \psi_T(s) = s^\alpha$ and $\eta_X(u) = -|u|^2$, and hence $\eta_Z(u) = -|u|^{2\alpha}$, i.e. Z is a rotationally invariant 2α-stable process.

In particular, if $d = 1$ and T is the Lévy subordinator, then Z is the *Cauchy process*, so each $Z(t)$ has a symmetric Cauchy distribution with parameters $\mu = 0$ and $\sigma = 1$. It is interesting to observe from (2.15) that Z is constructed from two independent standard Brownian motions.

Examples of subordinated processes have recently found useful applications in mathematical finance. We briefly mention two interesting cases:-

(i) **The Variance Gamma Process**

In this case $Z(t) = B(T(t))$, for each $t \geq 0$, where B is a standard Brownian motion and T is an independent gamma subordinator. The name derives from the fact that, in a formal sense, each $Z(t)$ arises by replacing the variance of a normal random variable by a gamma random variable. Using Proposition 2.2, a simple calculation yields

$$\Phi_{Z(t)}(u) = \left(1 + \frac{u^2}{2b}\right)^{-at},$$

for each $t \geq 0, u \in \mathbb{R}$, where a and b are the usual parameters which determine the gamma process. It is an easy exercise in manipulating characteristic functions to compute the alternative representation:

$$Z(t) = G(t) - L(t),$$

where G and L are independent gamma subordinators each with parameters $\sqrt{2b}$ and a. This yields a nice financial representation of Z as a difference of independent "gains" and "losses". From this representation, we can compute that Z has a Lévy density

$$g_\nu(x) = \frac{a}{|x|}(e^{\sqrt{2b}x}\chi_{(-\infty,0)}(x) + e^{-\sqrt{2b}x}\chi_{(0,\infty)}(x)).$$

The *CGMY processes* are a generalisation of the variance-gamma processes due to Carr, Geman, Madan and Yor. They are characterised by their Lévy density:

$$g_\nu(x) = \frac{a}{|x|^{1+\alpha}}(e^{b_1 x}\chi_{(-\infty,0)}(x) + e^{-b_2 x}\chi_{(0,\infty)}(x)),$$

where $a > 0, 0 \leq \alpha < 2$ and $b_1, b_2 \geq 0$. We obtain stable Lévy processes when $b_1 = b_2 = 0$. In fact, the CGMY processes are a subclass of the *tempered stable processes*. Note how the exponential dampens the effects of large jumps.

(ii) **The Normal Inverse Gaussian Process**

In this case $Z(t) = C(T(t)) + \mu t$ for each $t \geq 0$ where each $C(t) = B(t) + \beta t$, with $\beta \in \mathbb{R}$. Here T is an inverse Gaussian subordinator, which is independent of B, and in which we write the parameter $\gamma = \sqrt{\alpha^2 - \beta^2}$, where $\alpha \in \mathbb{R}$ with $\alpha^2 \geq \beta^2$. Z depends on four parameters and has characteristic function

$$\Phi_{Z(t)}(\alpha, \beta, \delta, \mu)(u) = \exp\{\delta t(\sqrt{\alpha^2 - \beta^2} - \sqrt{\alpha^2 - (\beta + iu)^2}) + i\mu t u\}$$

for each $u \in \mathbb{R}, t \geq 0$. Here $\delta > 0$ is as in (2.16).

Each $Z(t)$ has a density given by

$$f_{Z(t)}(x) = C(\alpha, \beta, \delta, \mu; t) q\left(\frac{x - \mu t}{\delta t}\right)^{-1} K_1\left(\delta t \alpha q\left(\frac{x - \mu t}{\delta t}\right)\right) e^{\beta x},$$

for each $x \in \mathbb{R}$, where $q(x) = \sqrt{1 + x^2}$, $C(\alpha, \beta, \delta, \mu; t) = \pi^{-1}\alpha e^{\delta t\sqrt{\alpha^2 - \beta^2} - \beta\mu t}$ and K_1 is a Bessel function of the third kind.

2.4 Filtrations, Markov Processes and Martingales

We recall the probability space $(\Omega, \mathcal{F}, P)$ which underlies our investigations. $\mathcal{F}$ contains all possible events in Ω. When we introduce the arrow of time, its convenient to be able to consider only those events which can occur up to and including time t. We denote by $\mathcal{F}_t$ this sub-σ-algebra of $\mathcal{F}$. To be able to consider all time instants on an equal footing, we define a *filtration* to be an increasing family $(\mathcal{F}_t, t \geq 0)$ of sub-σ-algebras of $\mathcal{F}$, , i.e.

$$0 \leq s \leq t < \infty \Rightarrow \mathcal{F}_s \subseteq \mathcal{F}_t.$$

A stochastic process $X = (X(t), t \geq 0)$ is *adapted* to the given filtration if each $X(t)$ is $\mathcal{F}_t$-measurable.

e.g. any process is adapted to its *natural filtration*,
$\mathcal{F}_t^X = \sigma\{X(s); 0 \leq s \leq t\}$.

An adapted process $X = (X(t), t \geq 0)$ is a *Markov process* if for all $f \in B_b(\mathbb{R}^d), 0 \leq s \leq t < \infty$,

$$\mathbb{E}(f(X(t))|\mathcal{F}_s) = \mathbb{E}(f(X(t))|X(s)) \quad (a.s.). \tag{2.18}$$

(i.e. “past” and “future” are independent, given the present).

The *transition probabilities* of a Markov process are $p_{s,t}(x, A) = P(X(t) \in A|X(s) = x)$, i.e. the probability that the process is in the Borel set A at time t given that it is at the point x at the earlier time s.

Theorem 2.6 *If X is an adapted Lévy process wherein each $X(t)$ has law q_t, then it is a Markov process with transition probabilities $p_{s,t}(x, A) = q_{t-s}(A - x)$.*

Proof. This essentially follows from

$$\begin{aligned} \mathbb{E}(f(X(t))|\mathcal{F}_s) &= \mathbb{E}(f(X(s) + X(t) - X(s))|\mathcal{F}_s) \\ &= \int_{\mathbb{R}^d} f(X(s) + y) q_{t-s}(dy). \end{aligned}$$ □

Now let X be an adapted process defined on a filtered probability space which also satisfies the integrability requirement $\mathbb{E}(|X(t)|) < \infty$ for all $t \geq 0$. We say that it is a *martingale* if for all $0 \leq s < t < \infty$,

$$\mathbb{E}(X(t)|\mathcal{F}_s) = X(s) \quad \text{a.s.}$$

Note that if X is a martingale, then the map $t \to \mathbb{E}(X(t))$ is constant.

An adapted Lévy process with zero mean is a martingale (with respect to its natural filtration) since in this case, for $0 \leq s \leq t < \infty$ and using the convenient notation $\mathbb{E}_s(\cdot) = \mathbb{E}(\cdot|\mathcal{F}_s)$:

$$\begin{aligned} \mathbb{E}_s(X(t)) &= \mathbb{E}_s(X(s) + X(t) - X(s)) \\ &= X(s) + \mathbb{E}(X(t) - X(s)) = X(s) \end{aligned}$$

Although there is no good reason why a generic Lévy process should be a martingale (or even have finite mean), there are some important examples:

e.g. the processes whose values at time t are

- $\sigma B(t)$ where $B(t)$ is a standard Brownian motion, and σ is an $r \times d$ matrix.
- $\tilde{N}(t)$ where $\tilde{N}$ is a compensated Poisson process with intensity λ.

Some important martingales associated to Lévy processes include:

- $\exp\{i(u, X(t)) - t\eta(u)\}$, where $u \in \mathbb{R}^d$ is fixed.
- $|\sigma B(t)|^2 - \text{tr}(A)t$ where $A = \sigma^T \sigma$.
- $\tilde{N}(t)^2 - \lambda t$.