1

Introduction to Lévy Processes

Lévy processes, i.e. processes in continuous time with stationary and independent increments, are named after Paul Lévy: he made the connection with infinitely divisible distributions (Lévy–Khintchine formula) and described their structure (Lévy–Itô decomposition).

I believe that their study is of particular interest today for the following reasons

- They form a subclass of general Markov processes which is large enough to include many familiar processes such as Brownian motion, the Poisson process, Stable processes, etc, but small enough that a particular member can be specified by a few quantities (the *characteristics* of a Lévy process).
- In a sense, they stand in the same relation to Brownian motion as general random walks do to the simple symmetric random walk, and their study draws on techniques from both these areas.
- Their *sample path behaviour* poses a variety of difficult and fascinating questions, some of which are not relevant for Brownian motion.
- They form a flexible class of models, which have been applied to the study of storage processes, insurance risk, queues, turbulence, laser cooling, ... and of course finance, where the feature that they include examples having "heavy tails" is particularly important.

This course will cover only a part of the theory of Lévy processes, and will not discuss applications. Even within the area of fluctuation theory, there are many recent interesting developments that I won't have time to discuss.

Almost all the material in Chapters 1–4 can be found in Bertoin [12]. For related background material, see Bingham [19], Satô [90], and Satô [91].

1.1 Notation

We will use the canonical notation, and denote by $X = (X_t, t \geq 0)$ the co-ordinate process, i.e. $X_t = X_t(\omega) = \omega(t)$, where $\omega \in \Omega$, the space of real-valued cadlag paths, augmented by a cemetery point ϑ, and endowed with the Skorohod topology. The Borel σ-field of Ω will be denoted by $\mathcal{F}$ and the lifetime by $\zeta = \zeta(\omega) = \inf\{t \geq 0 : \omega(t) = \vartheta\}$.

Definition 1. *Let $\mathbb{P}$ be a probability measure on $(\Omega, \mathcal{F})$ with $\mathbb{P}(\zeta = \infty) = 1$. We say that X is a (real-valued) Lévy process for $(\Omega, \mathcal{F}, \mathbb{P})$ if for every $t \geq s \geq 0$, the increment $X_{t+s} - X_t$ is independent of $(X_u, 0 \leq u \leq t)$ and has the same distribution as X_s.*

Note that this forces $\mathbb{P}(X_0 = 0) = 1$; we will later write $\mathbb{P}_x$ for the measure corresponding to $(x + X_t, t \geq 0)$ under $\mathbb{P}$.

(Incidentally the name Lévy process has only been the accepted terminology for approximately 20 years; prior to that the name "process with stationary and independent increments" was generally used.)

From the decomposition

$$X_1 = X_{\frac{1}{n}} + \left(X_{\frac{2}{n}} - X_{\frac{1}{n}}\right) + \cdots + \left(X_{\frac{n}{n}} - X_{\frac{n-1}{n}}\right)$$

it is apparent that X_1 has an *infinitely divisible* distribution under $\mathbb{P}$. The form of a general infinitely divisible distribution is given by the well-known Lévy–Khintchine formula, and from it we deduce easily the following result.

Theorem 1. *Let X be a Lévy process on $(\Omega, \mathcal{F}, \mathbb{P})$; then*

$$\mathbb{E}(\exp i\lambda X_t) = e^{-t\Psi(\lambda)},\ t \geq 0, \lambda \in \mathbb{R},$$

where, for some real γ, σ and measure Π on $\mathbb{R} - \{0\}$ which satisfies

$$\int_{-\infty}^{\infty} \{x^2 \wedge 1\} \Pi(dx) < \infty, \tag{1.1.1}$$

$$\Psi(\lambda) = -i\gamma\lambda + \frac{\sigma^2}{2}\lambda^2 + \int_{-\infty}^{\infty} \left\{1 - e^{i\lambda x} + i\lambda x \mathbf{1}_{(|x|<1)}\right\} \Pi(dx). \tag{1.1.2}$$

Ψ is called the **Lévy exponent** of X, and we will call the quantities γ the linear cefficient, σ the Brownian coefficient, and Π the Lévy measure of X : together they constitute the **characteristics** of X. There is an existence theorem: given real γ, any $\sigma \geq 0$ and measure Π satisfying (1.1.1) there is a measure under which X is a Lévy process with characteristics γ, σ and Π. There is also a uniqueness result, as any alteration in one or more of the characteristics results in a Lévy process with a different distribution.

Examples

- The characteristics of standard Brownian motion are $\gamma = 0, \sigma = 1, \Pi \equiv 0$, and $\Psi(\lambda) = \frac{\lambda^2}{2}$.
- The characteristics of a compound Poisson process with jump rate c and step distribution F are

$$\gamma = c \int_{\{|x|<1\}} xF(dx), \sigma = 0, \Pi(dx) = cF(dx),$$

  and $\Psi(\lambda) = c(1 - \phi(\lambda))$, where $\phi(\theta) = \int_{-\infty}^{\infty} e^{i\lambda x} dF(x)$.

- The characteristics of a Gamma process are

$$\gamma = c(1 - e^{-1}), \sigma = 0, \Pi(dx) = cx^{-1}e^{-x}\mathbf{1}_{\{x>0\}}dx,$$

and $\Psi(\lambda) = c\log(1 - i\lambda)$.
- The characteristics of a strictly stable process of index $\alpha \in (0,1)\cup(1,2)$ are

$$\gamma \text{ arbitrary}, \sigma = 0, \Pi(dx) = \begin{matrix} c_+ x^{-\alpha-1}dx & \text{if } x > 0, \\ c_-|x|^{-\alpha-1}dx & \text{if } x < 0. \end{matrix}$$

If $\alpha \neq 1$, $c_+ \geq 0$ and $c_- \geq 0$ are arbitrary, and

$$\Psi(\lambda) = c|\lambda|^\alpha \{1 - i\beta sgn(\lambda)\tan(\pi\alpha/2)\} - i\gamma\lambda.$$

If $\alpha = 1$, $c_+ = c_- > 0$, and $\Psi(\lambda) = c|\lambda| - i\gamma\lambda$; this is a Cauchy process with drift.

Note that there is a fairly obvious generalisation of Theorem 1 to $\mathbb{R}^d$, but we will stick, almost exclusively, to the 1-dimensional case.

The first step to getting a probabilistic interpretation of Theorem 1 is to realise that the process of jumps,

$$\Delta = (\Delta_t, t \geq 0) \text{ where } \Delta_t = X_t - X_{t-},$$

is a Poisson point process, but first we need some background material.

1.2 Poisson Point Processes

A random measure ϕ on a Polish space E (this means it is metric-complete and separable) is called a Poisson measure with intensity ν if

1. ν is a σ-finite measure on E;
2. for every Borel subset B of E with $0 < \nu(B) < \infty$, $\phi(B)$ has a Poisson distribution with parameter $\nu(B)$; in particular $\phi(B)$ has mean $\nu(B)$;
3. for disjoint Borel subsets $B_1, \cdots B_n$ of E, the random variables $\phi(B_1), \cdots, \phi(B_n)$ are independent.

In the case that $c := \nu(E) < \infty$, it is clear that we can represent ϕ as a sum of Dirac point masses as follows. Let $y_1, y_2, \cdots$ be a sequence of independent and identically distributed E-valued random variables with distribution $c^{-1}\nu$, and N an independent Poisson-distributed random variable with parameter c; then we can represent ϕ as

$$\phi = \sum_1^N \delta_{y_j},$$

where δ_y denotes the Dirac point mass at $y \in E$. If $\nu(E) = \infty$, there is a decomposition of E into disjoint Borel sets $E_1, E_2, \cdots$, each having $\nu(E_j)$

finite, and we can represent ϕ as the sum of independent Poisson measures ϕ_j having intensities $\nu \mathbf{1}_{E_j}$, each having the above representation, so again ϕ can be represented as the sum of Dirac point masses.

To set up a Poisson point process we consider the product space $E \times [0, \infty)$, the measure $\mu = \nu \times dx$, and a Poisson measure ϕ on $E \times [0, \infty)$ with intensity μ. It is easy to check that a.s. $\phi(E \times \{t\}) = 1$ or 0 for all $t \geq 0$, so we can introduce a process $(e(t), t \geq 0)$ by letting $(e(t), t)$ denote the position of the point mass on $E \times \{t\}$ in the first case, and in the second case put $e(t) = \xi$, where ξ is an additional isolated point. Then we can write

$$\phi = \sum_{t \geq 0} \delta_{(e(t),t)}.$$

The process $e = (e(t), t \geq 0)$ is called a Poisson point process with characteristic measure ν.

The basic properties of a Poisson point process are stated in the next result.

Proposition 1. *Let B be a Borel set with $\nu(B) < \infty$, and define its counting process by*

$$N_t^B = \#\{s \leq t : e(s) \in B\} = \phi(B \times [0, t]),\ \ t \geq 0,$$

and its entrance time by

$$T_B = \inf\{t \geq 0 : e(t) \in B\}.$$

Then

(i) N^B is a Poisson process of parameter $\nu(B)$, which is adapted to the filtration $\mathcal{G}$ of e.
(ii) T_B is a $(\mathcal{G}_t)$-stopping time which has an exponential distribution with parameter $\nu(B)$.
(iii) $e(T_B)$ and T_B are independent, and for any Borel set A

$$\mathbb{P}(e(T_B) \in A) = \frac{\nu(A \cap B)}{\nu(B)}.$$

(iv) The process e' defined by $e'(t) = \xi$ if $e(t) \in B$ and $e'(t) = e(t)$ otherwise is a Poisson point process with characteristic measure $\nu \mathbf{1}_{B^c}$, and it is independent of $(T_B, e(T_B))$.

The process $(e(t), 0 \leq t \leq T_B)$ is called the process **stopped** at the first point in B; its law is characterized by Proposition 1.

If we define a deterministic function on $E \times [0, \infty)$ by $H_t(y) = \mathbf{1}_{B \times (t_1, t_2]}(y, t)$ it is clear that

$$\mathbb{E}\left(\sum_{0 \leq t < \infty} H_t(e(t))\right) = (t_2 - t_1)\nu(B);$$

this is the building block on which the following important result is based.

Proposition 2. ***(The compensation formula)*** *Let $H = (H_t, t \geq 0)$ be a predictable process taking values in the space of nonnegative measurable functions on $E_\cup\{\xi\}$ and having $H_t(\xi) \equiv 0$. Then*

$$\mathbb{E}\left(\sum_{0\leq t<\infty} H_t(e(t))\right) = \mathbb{E}\left(\int_0^\infty dt \int_E H_t(y)\nu(dy)\right).$$

A second important result is called **the exponential formula**;

Proposition 3. *Let f be a complex-valued Borel function on $E_\cup\{\xi\}$ with $f(\xi) = 0$ and*

$$\int_E |1 - e^{f(y)}|\nu(dy) < \infty.$$

Then for any $t \geq 0$

$$\mathbb{E}\left(\exp\left\{\sum_{0\leq s\leq t} f(e(s))\right\}\right) = \exp\left\{-t\int_E (1 - e^{f(y)})\nu(dy)\right\}.$$

1.3 The Lévy–Itô Decomposition

It is important to get a probabilistic interpretation of the Lévy–Khintchine formula, and this is what this decomposition does. Fundamentally, it describes the way that the measure Π determines the structure of the jumps in the process. Specifically it states that X can be written in the form

$$X_t = \gamma t + \sigma B_t + Y_t,$$

where B is a standard Brownian motion, and Y is a Lévy process which is independent of B, and is "determined by its jumps", in the following sense. Let $\Delta = \{\Delta_t, t \geq 0\}$ be a Poisson point process on $\mathbb{R} \times [0, \infty)$ with characteristic measure Π, and note that since $\Pi\{x : |x| \geq 1\} < \infty$, then $\sum_{s\leq t} 1_{\{|\Delta_s|\geq 1\}}|\Delta_s| < \infty$ a.s. Moreover if we define

$$Y_t^{(2)} = \sum_{s\leq t} 1_{\{|\Delta_s|\geq 1\}}\Delta_s, \ t \geq 0$$

then it is easy to see that, provided $c = \Pi\{x : |x| \geq 1\} > 0$, $(Y_t^{(2)}, t \geq 0)$ is a compound Poisson process with jump rate c, step distribution $F(dx) = c^{-1}\Pi(dx)\mathbf{1}_{\{|x|\geq 1\}}$ and, by the exponential formula, Lévy exponent

$$\Psi^{(2)}(\lambda) = \int_{|x|\geq 1} \{1 - e^{i\lambda x}\}\Pi(dx).$$

If

$$I = \int (1 \wedge |x|)\Pi(dx) < \infty, \tag{1.3.1}$$

then, by considering the limit of $\sum_{s\le t} 1_{\{\varepsilon<|\Delta_s|<1\}} |\Delta_s|$ as $\varepsilon \downarrow 0$, we see that

$$\sum_{s\le t} 1_{\{|\Delta_s|<1\}} |\Delta_s| < \infty \text{ a.s. for each } t < \infty,$$

and in this case we set $Y_t = Y_t^{(1)} + Y_t^{(2)}$, where

$$Y_t^{(1)} = \sum_{s\le t} \Delta_s 1_{\{|\Delta_s|<1\}},\ t \ge 0,$$

is independent of $Y^{(2)}$. Clearly, in this case Y has bounded variation (on each finite time interval), and it's exponent is

$$\Psi^{(1)}(\lambda) = \int_{|x|<1} \{1 - e^{i\lambda x}\}\Pi(dx).$$

In this case we can rewrite the Lévy–Khintchine formula as

$$\Psi(\lambda) = -i\delta\lambda + \frac{\sigma^2}{2}\lambda^2 + \Psi^{(1)}(\lambda) + \Psi^{(2)}(\lambda),$$

where $\delta = \gamma - \int_{|x|<1} x\Pi(dx)$ is finite, and the Lévy–Itô decomposition takes the form

$$X_t = \delta t + \sigma B_t + Y_t^{(1)} + Y_t^{(2)},\ t \ge 0, \tag{1.3.2}$$

where the processes $B, Y^{(1)}$ and $Y^{(2)}$ are independent. The constant δ is called the **drift coefficient** of X.

However, if $I = \infty$ then a.s. $\sum_{s\le t} |\Delta_s| = \infty$ for each $t > 0$, and in this case we need to define $Y^{(1)}$ differently: in fact as the a.s. limit as $\varepsilon \downarrow 0$ of the compensated partial sums,

$$Y_{\varepsilon,t}^{(1)} = \sum_{s\le t} 1_{\{\varepsilon<|\Delta_s|\le 1\}}\Delta_s - t\int_{\varepsilon<|x|\le 1} x\Pi(dx).$$

It is clear that $\{Y_{\varepsilon,t}^{(1)}, t \ge 0\}$ is a Lévy process, in fact a compensated compound Poisson process with exponent

$$\Psi_\varepsilon^{(1)}(\lambda) = \int_{-\infty}^{\infty} \{1 - e^{i\lambda x} + i\lambda x\}\mathbf{1}_{(\varepsilon<|x|<1)}\Pi(dx),$$

and hence a martingale. The key point, (see e.g. [12] p14), is that the basic assumption that $\int(1 \wedge x^2)\Pi(dx) < \infty$ allows us to use a version of Doob's maximal inequality for martingales to show that the limit as $\varepsilon \downarrow 0$ exists, has stationary and independent increments, and is a Lévy process with exponent

$$\Psi^{(1)}(\lambda) = \int_{-\infty}^{\infty} \{1 - e^{i\lambda x} + i\lambda x\}\mathbf{1}_{(|x|<1)}\Pi(dx).$$

In this case the Lévy–Itô decomposition takes the form

$$X_t = \gamma t + \sigma B_t + Y_t^{(1)} + Y_t^{(2)},\ t \geq 0, \tag{1.3.3}$$

where again the processes $B, Y^{(1)}$ and $Y^{(2)}$ are independent.

Since $Y^{(2)}$ has unbounded variation we see that X has bounded variation $\Longleftrightarrow \sigma = 0$ and $I < \infty$. All the examples we have discussed have bounded variation, except for Brownian motion and stable processes with index $\in (1,2)$.

To conclude this section, we record some information about the asymptotic behaviour of the Lévy exponent.

Proposition 4. *(i) In all cases we have*

$$\lim_{|\lambda|\to\infty} \frac{\Psi(\lambda)}{\lambda^2} = \frac{\sigma^2}{2}.$$

(ii) If X has bounded variation and drift coefficient δ,

$$\lim_{|\lambda|\to\infty} \frac{\Psi(\lambda)}{\lambda} = -i\delta.$$

(iii) X is a compound Poisson process if and only if Ψ is bounded.

(Note that we reserve the name compound Poisson process for a Lévy process with a finite Lévy measure, no Brownian component and drift coefficient zero.)

1.4 Lévy Processes as Markov Processes

It is clear that any Lévy process has the simple Markov property in the stronger, spatially homogeneous form that, given $X_t = x$, the process $\{X_{t+s}, s \geq 0\}$ is independent of $\{X_u, u < t\}$ and has the law of $\{x + X_s, s \geq 0\}$. In fact

- a similar form of the strong Markov property also holds. In particular this means that the above is valid if the fixed time t is replaced by a *first passage time*
$$T_B = \inf\{t \geq 0 : X_t \in B\}$$
whenever B is either open or closed.
- It is also the case that the semi-group of X has the Feller property and it turns out that the strong Feller property holds in the important special case that the law of X_t is absolutely continuous with respect to Lebesgue measure.
- In these, and some other circumstances, the resolvent kernel is absolutely continuous, i.e. there exists a non-negative measurable function $u^{(q)}$ such that

$$U^{(q)}f(x) := \int_0^\infty e^{-qt} P_t f(x) dt = \int_{-\infty}^{\infty} f(x+y) u^{(q)}(y) dy,$$

where

$$P_t f(x) = \mathbb{E}_x(f(X_t)).$$

- The associated potential theory requires no additional hypotheses; in particular if we write $X^* = -X$ for the dual of X we have the following duality relations. Let f and g be non-negative; then

$$\int_{\mathbb{R}} P_t f(x) g(x) dx = \int_{\mathbb{R}} f(x) P_t^* g(x) dx, \ t > 0,$$

and

$$\int_{\mathbb{R}} U^{(q)} f(x) g(x) dx = \int_{\mathbb{R}} f(x) U^{*(q)} g(x) dx, \ t > 0,$$

- The relation between X and X^* via time-reversal is also simple; *for each fixed* $t > 0$, the reversed process $\{X_{(t-s)-} - X_t, 0 \le s \le t\}$ and the dual process $\{X^*_s, 0 \le s \le t\}$ have the same law under $\mathbb{P}$.

In summary; X is a “nice” Markov process, and many of technical problems which appear in the general theory are simplified for Lévy processes.