Introduction to the Mathematics of Lévy Processes

Kazuhisa Matsuda

Department of Economics
The Graduate Center, The City University of New York,
365 Fifth Avenue, New York, NY 10016-4309
Email: maxmatsuda@maxmatsuda.com
http://www.maxmatsuda.com/

February 2005

Abstract

The goal of this sequel is to provide the foundations of the mathematics of Lévy processes for the readers with undergraduate knowledge of stochastic processes as simple as possible. The simplicity is a key because, for the beginners such as finance majors without the experience in stochastic processes, some available books on Lévy processes are not accessible. Lévy processes constitute a wide class of stochastic processes whose sample paths can be continuous, continuous with occasionally discontinuous, and purely discontinuous. Traditional examples of Lévy processes include a Brownian motion with drift (i.e. the only continuous Lévy processes), a Poisson process, a compound Poisson process, a jump diffusion process, and a Cauchy process. All of these are well studied and well applied stochastic processes. We define and characterize Lévy processes using theorems such as the Lévy-Itô decomposition and the Lévy-Khinchin representation and in terms of their infinite divisibilities and the Lévy measures ℓ. In the last decade and in the field of quantitative finance, there was an explosion of literatures modeling the log asset prices using purely non-Gaussian Lévy processes which are pure jump Lévy processes with infinite activity. To raise a few examples of purely non-Gaussian Lévy processes used in finance, variance gamma processes, tempered stable processes, and generalized hyperbolic Lévy motions. We cover these purely non-Gaussian Lévy processes in the next sequel with a finance application. This is because we like to keep this sequal as simple as possible with the pourpose of providing the introductory foundations of the mathematics of Lévy processes.

Contents

[1] Introduction to the Mathematics of Lévy Processes

First of all, we have to say that Lévy processes are nothing new because their properties were originally characterized by Paul Lévy in the 1930s. Lévy processes are simply defined as stochastic processes 1) whose increments are independent and stationary, 2) they are stochastically continuous, and 3) whose sample paths are right continuous and left limit functions of time with probability 1. Thus, Lévy processes constitute a wide class of stochastic processes whose sample paths can be continuous, continuous with occasionally discontinuous, and purely discontinuous. Traditional examples of Lévy processes include a Brownian motion with drift (i.e. the only continuous Lévy processes), a Poisson process, a compound Poisson process, a jump diffusion process, and a Cauchy process. All of these are well studied and well applied stochastic processes.

The probability distributions of all Lévy processes are characterized by infinite divisibility. In other words, there is one to one correspondence between an infinitely divisible distribution and a Lévy process. For example, a Brownian motion with drift which is a Lévy process is a stochastic process generated by a normal distribution which is an infinitely divisible distribution.

Lévy-Itô decomposition states that every sample path of Lévy process can be represented as a sum of two independent processes which are each expressed using a Lévy triplet (A, ℓ, γ): One is a continuous Lévy process and the other is a compensated sum of independent jumps. Obviously, a continuous Lévy process is a Brownian motion with drift. One trick is that the jump component has to be a *compensated* sum of independent jumps because a sum of independent jumps at time t may not converge.

Following Lévy-Itô decomposition, Lévy-Khinchin representation gives the characteristic functions of all infinitely divisible distributions. In other words, it gives the characteristic functions of all processes whose increments follow infinitely divisible distributions – Lévy processes in terms of a Lévy triplet (A, ℓ, γ).

The Lévy measure ℓ of a Lévy process $(X_{t\in[0,\infty)})$ is defined as a unique positive measure on $\mathbb{R}$ which measures (counts) the expected (average) number of jumps of all sizes per unit of time. In other words, it is a unique positive measure on $\mathbb{R}$ which measures arrival rate of jumps of all sizes per unit of time. Note that by definition of the Lévy measure ℓ, all Lévy processes have finite expected number of large jumps per unit of time. If a Lévy process has finite expected number of small jumps per unit of time (i.e. a finite integral of ℓ), then, it is said to be a finite activity Lévy process such as a compound Poisson process. If a Lévy process has infinite expected number of small jumps per unit of time (i.e. an infinite integral of ℓ), then, it is said to be an infinite activity Lévy process such as a gamma process.

Another important characterization of Lévy processes is that Lévy processes are stochastically continuous Markov processes with time homogeneous and spatially homogeneous transition functions.

The goal of this sequel is to provide the foundations of the mathematics of Lévy processes for the readers with undergraduate knowledge of stochastic processes as simple as possible. The simplicity is a key because, for the beginners such as finance majors without the experience in stochastic processes, some available books on Lévy processes are not accessible (from our own experience). For those advanced readers who intend to learn Lévy processes using theorems and proofs approach, we highly recommend the excellent book by Sato (1999) which is by far the best among available books treating Lévy processes.

The structure of this sequel is as follows. Section 2 provides readers with minimal necessary knowledge of stochastic processes. We especially emphasize the sample paths properties such as the concepts of continuity and limits and the total variation of stochastic processes. Martingale and Markov properties are also introduced here. Section 3 defines Lévy processes, then, provides theorems such as the Lévy-Itô decomposition and the Lévy-Khinchin representation. Infinite divisibility and the Lévy measure ℓ of Lévy processes are discussed in detail along with various ways to categorize Lévy processes. Section 4 gives three traditional examples of Lévy processes – Brownian motion, a Poisson process, and a compound Poisson process. Section 5 deals with stable processes which are Lévy processes with the broad-sense selfsimilarity.

To gain simplicity, we needed to limit the scope of the discussion in a couple of ways. One is that we only deal with one dimensional stochastic process in this sequel, not a d-dimensional vector of stochastic processes. The other is that we only treat the traditional examples of Lévy processes, namely a Brownian motion, a Poisson process, and a compound Poisson process.

In the last decade and in the field of quantitative finance, there was an explosion of literatures modeling the log asset prices using purely non-Gaussian Lévy processes which are pure jump Lévy processes with infinite activity. To raise a few examples of purely non-Gaussian Lévy processes used in finance, variance gamma processes, tempered stable processes, and generalized hyperbolic Lévy motions. We cover these purely non-Gaussian Lévy processes in the next sequel with a finance application.

[2] Basics of Stochastic Processes

This chapter presents the fundamental concepts of functions and stochastic processes which are essential to the understanding of Lévy processes.

[2.1] Function

[2.1.1] Function

Definition 2.1 Function A function $f : a \to f(a)$ or $f : A \to B$ on $\mathbb{R}$ uniquely maps (relates) a set of input values $a \in A$ to a set of output values $f(a) \in B$. The domain of a function is the set A on which a function is defined and the set of all actual outputs $f(a) \in B$ is called the range of a function.

A function is a many-to-one mapping (i.e. not one-to-many mapping). For example, a function $f(a) = a$ is a one-to-one mapping, $f(a) = -a^2$ is a two-to-one mapping except for $a = 0$, and $f(a) = \sin(2\pi a)$ is a many-to-one mapping.

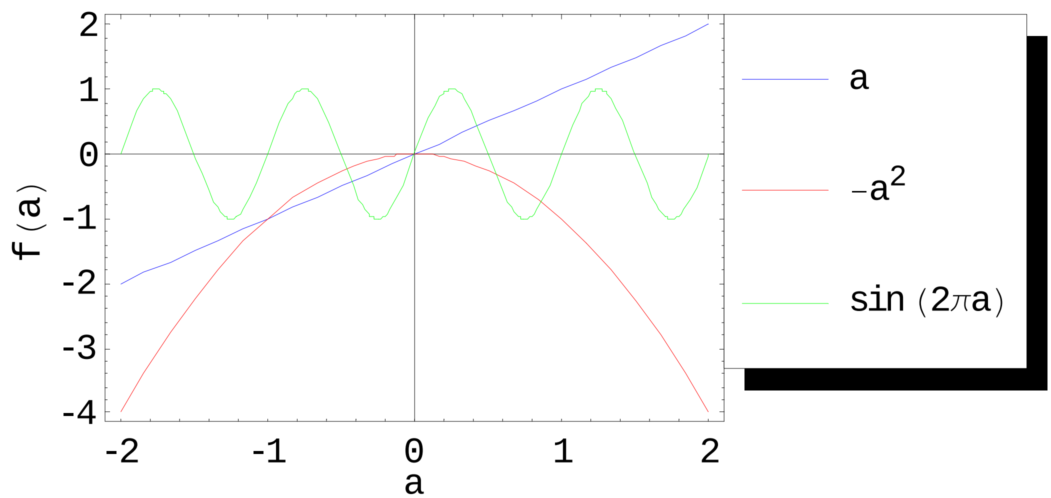


Figure 2.1: Examples of a function $f : a \to f(a)$.

[2.1.2] Left Limit and Right Limit of a Function

Definition 2.2 Left limit and Right limit of a function A function $f : a \to f(a)$ on $\mathbb{R}$ has a left limit $f(b-)$ at a point $a = b$ if $f(a)$ approaches $f(b-)$ when a approaches b from the below (the left-hand side):

$$\lim_{a \to b-} f(a) = f(b-).$$

A function $f : a \to f(a)$ on $\mathbb{R}$ has a right limit $f(b+)$ at a point $a = b$ if $f(a)$ approaches $f(b+)$ when a approaches b from the above (right-hand side):

$$\lim_{a \to b+} f(a) = f(b+).$$

[2.1.3] Right Continuous Function and Right Continuous with Left Limit (RCLL) Function

Definition 2.3 Right continuous function A function f on $\mathbb{R}$ is said to be right continuous at a point $a = b$ if it satisfies the following conditions:

(1) $f(b)$ is defined. In other words, a point b is in the domain of a function f.
(2) Right limit of the function as a approaches b from the above (right hand side) exists, i.e. $\lim_{a \to b+} f(a) = f(b+)$.
(3) $f(b+) = f(b)$.

Definition 2.4 Right continuous with left limit (rcll) function A function f on $\mathbb{R}$ is said to be right continuous with left limit at a point $a = b$ if it satisfies the following conditions:

(1) $f(b)$ is defined. In other words, a point b is in the domain of a function f.
(2) Right limit of the function as a approaches b from the above (right hand side) exists, i.e. $\lim_{a \to b+} f(a) = f(b+)$. Left limit of the function as a approaches b from the below (left hand side) exists, i.e. $\lim_{a \to b-} f(a) = f(b-)$.
(3) $f(b+) = f(b)$.

The above definitions imply that a rcll function is right continuous, but the reverse is not true. In other words, a rcll function is more restrictive than a right continuous function because a rcll function needs left limit. This point is illustrated in Figure 2.2.

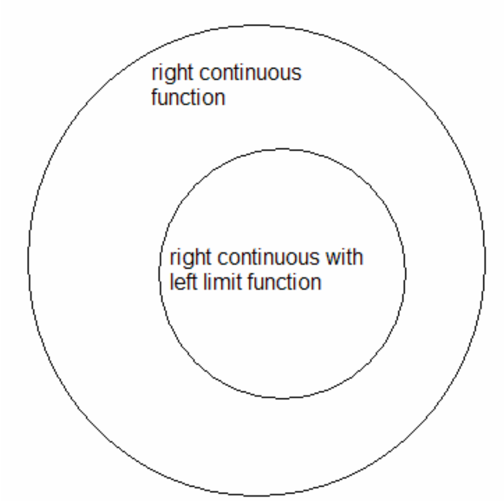


Figure 2.2: Relationship between rc function and rcll function.

Consider a piecewise constant function defined as (illustrated in Figure 2.3):

$$f(a) = \begin{cases} 0 & \text{if} \quad a < 1 \\ 1 & \text{if} \quad 1 \le a < 2 \\ 2 & \text{if} \quad 2 \le a < 3 \end{cases} . \tag{2.1}$$

The right limit at a point $a = 1$ is equal to the actual value of the function at a point $a = 1$:

$$f(1+) = f(1) = 1,$$

this means f is right continuous at a point $a = 1$. But the left limit at a point $a = 1$ is not equal to the actual value of the function at a point $a = 1$:

$$f(1-) = 0 \neq f(1) = 1,$$

this means f is not left continuous at a point $a = 1$. Therefore, this function is right continuous with left limit. And the jump size is:

$$f(1+) - f(1-) = 1 - 0 = 1.$$

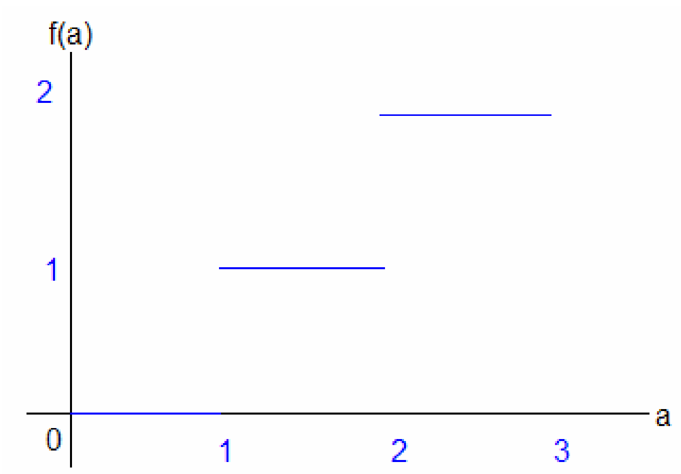


Figure 2.3: Right continuous with left limit (rcll) function.

[2.1.4] Left Continuous Function and Left Continuous with Right Limit (LCRL) Function

Definition 2.5 Left continuous function A function f on $\mathbb{R}$ is said to be left continuous at a point $a = b$ if it satisfies the following conditions:

(1) $f(b)$ is defined. In other words, a point b is in the domain of a function f.
(2) Left limit of the function as a approaches b from the below (left hand side) exists, i.e. $\lim_{a \to b-} f(a) = f(b-)$.

(3) $f(b-) = f(b)$.

Definition 2.6 Left continuous with right limit (lcrl) function A function f on $\mathbb{R}$ is said to be left continuous with right limit at a point $a = b$ if it satisfies the following conditions:

(1) $f(b)$ is defined. In other words, a point b is in the domain of a function f.
(2) Right limit of the function as a approaches b from the above (right hand side) exists, i.e. $\lim_{a \to b+} f(a) = f(b+)$. Left limit of the function as a approaches b from the below (left hand side) exists, i.e. $\lim_{a \to b-} f(a) = f(b-)$.
(3) $f(b-) = f(b)$.

Consider a piecewise constant function defined as:

$$f(a) = \begin{cases} 0 & \text{if} \quad a \le 1 \\ 1 & \text{if} \quad 1 < a \le 2 \\ 2 & \text{if} \quad 2 < a \le 3 \end{cases}. \tag{2.2}$$

The left limit at a point $a = 1$ is equal to the actual value of the function at a point $a = 1$:

$$f(1-) = f(1) = 0,$$

this means f is left continuous at a point $a = 1$. But the right limit at a point $a = 1$ is not equal to the actual value of the function at a point $a = 1$:

$$f(1+) = 1 \ne f(1) = 0,$$

this means f is not right continuous at a point $a = 1$. Therefore, this function is left continuous with right limit. And the jump size is:

$$f(1+) - f(1-) = 1 - 0 = 1.$$

[2.1.5] Continuous Function

Definition 2.7 Continuous function A function $f : a \to f(a)$ on $\mathbb{R}$ is said to be continuous at a point $a = b$ if it satisfies the following conditions:

(1) $f(b)$ is defined. In other words, a point b is in the domain of a function f.
(2) Right limit of the function as a approaches b from the above (right hand side) exists, i.e. $\lim_{a \to b+} f(a) = f(b+)$. Left limit of the function as a approaches b from the below (left hand side) exists, i.e. $\lim_{a \to b-} f(a) = f(b-)$.
(3) $f(b+) = f(b-) = f(b)$.

In other words, a continuous function is a left and right continuous function which in turn means that a continuous function is the most restrictive among rc, rcll, and continuous functions. All the functions in Figure 2.1 are continuous.

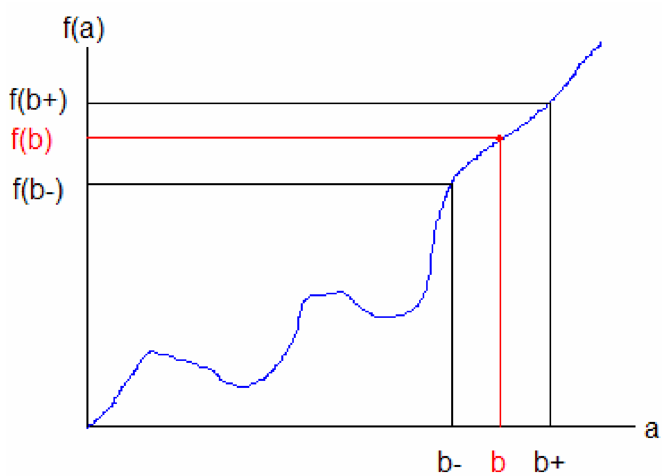


Figure 2.4: Illustration of a continuous function.

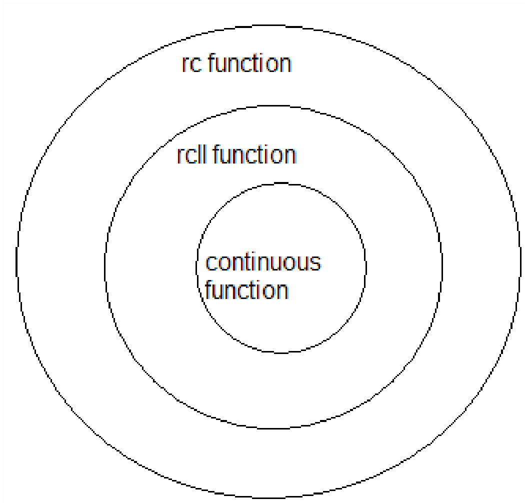


Figure 2.5: Relationship between rc, rcll, and continuous functions.

[2.1.6] Discontinuous Function

Definition 2.8 Discontinuous function A function $f : a \to f(a)$ on $\mathbb{R}$ is said to be discontinuous at a point $a = b$ (called a point of discontinuity) if it fails to satisfy being a continuous function.

There are three different categories of points of discontinuities.

Definition 2.9 A function with removable discontinuity (singularity) A function $f : a \to f(a)$ on $\mathbb{R}$ is said to have a removable discontinuity at a point $a = b$ if it satisfies the following conditions:

(1) $f(b)$ is defined or $f(b)$ is not defined.
(2) Left limit $\lim_{a \to b-} f(a) = f(b-)$ exists. Right limit $\lim_{a \to b+} f(a) = f(b+)$ exists.
(3) $f(b-) = f(b+) \neq f(b)$.

This means that a removable discontinuity at a point $a = b$ looks like a dislocated point as shown by Figure 2.6 where the example is a function:

$$f(a) = \begin{cases} -a + 5 & \text{if} \quad a \neq 3 \\ 5 \quad \text{if} \quad a = 3 \end{cases}.$$

This function has a left limit 2 which is equal to the right limit at a point $a = 3$:

$$f(3-) = f(3+) = 2,$$

but these limits are not equal to the actual value that this function takes at a point $a = 3$:

$$f(3-) = f(3+) = 2 \neq f(3) = 5,$$

which indicates that f is discontinuous at a point $a = b$.

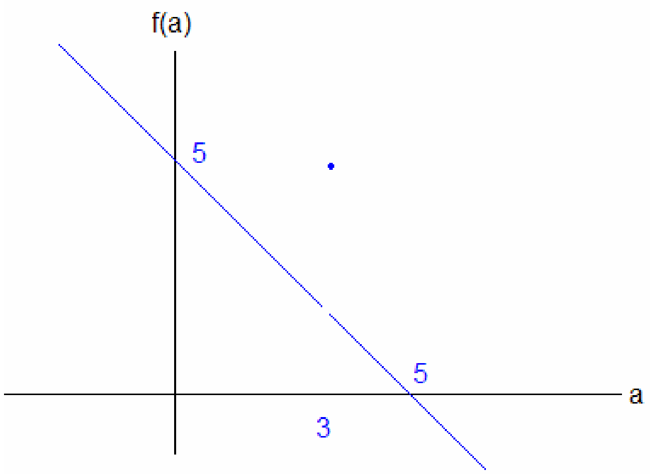


Figure 2.6: Example of a removable discontinuity with the defined discontinuity point $f(3) = 5$.

Or, consider a function:

$$f(a) = \frac{a^2 - 25}{a - 5},$$

which is undefined at a point $a = 5$. But its left limit and right limit exist and they are equal:

$$f(5-) = f(5+) = 10 \, .$$

Therefore, it is a function with removable discontinuity, too.

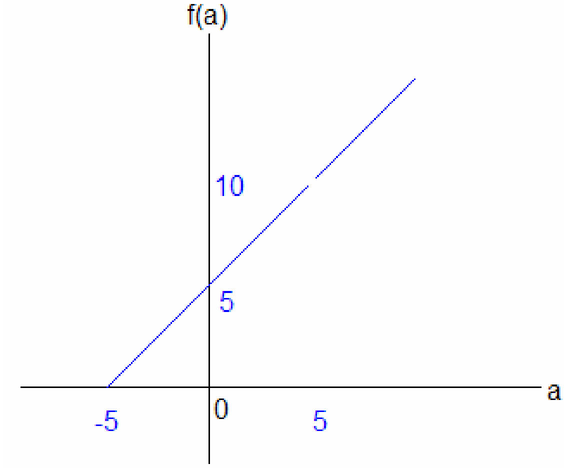


Figure 2.7: Example of a removable discontinuity with the undefined discontinuity point $f(5)$.

Definition 2.10 A function with discontinuity of the first kind (jump discontinuity)
A function $f : a \to f(a)$ on $\mathbb{R}$ is said to have a jump discontinuity at a point $a = b$ if it satisfies the following conditions:

(1) $f(b)$ is defined. In other words, a point b is in the domain of a function f .
(2) Left limit $f(b-)$ exists. Right limit $f(b+)$ exists.
(3) $f(b-) \neq f(b+)$.

Then, the jump is defined by the amount $f(b+) - f(b-)$.

Consider a function:

$$f(a) = \begin{cases} 1 & \text{if} \quad a > 1 \\ 0 & \text{if} \quad a = 1 \\ -1 & \text{if} \quad a < 1 \end{cases} \, . \tag{2.3}$$

This function has a left limit -1 which is not equal to the right limit 1 at a point $a = 1$:

$$f(1-) = -1 \neq f(1+) = 1 \, ,$$

and the jump size is:

$$f(1+) - f(1-) = 1 - (-1) = 2 \, .$$

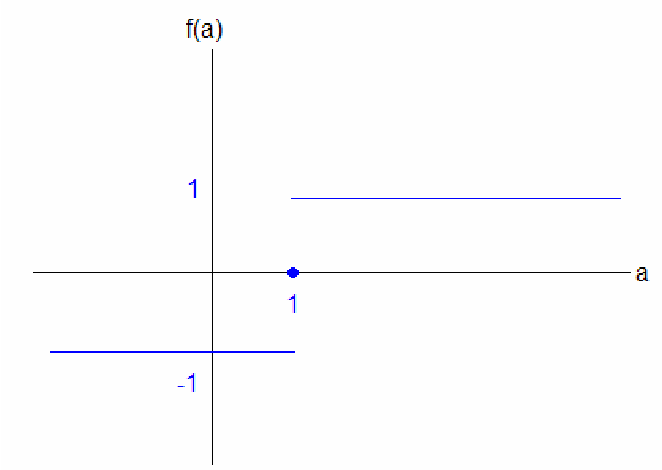


Figure 2.8: Example of a jump discontinuity.

Definition 2.11 A function with discontinuity of the second kind (essential discontinuity) A function $f : a \to f(a)$ on $\mathbb{R}$ is said to have an essential discontinuity at a point $a = b$ if either (or both) of left limit $f(b-)$ or right limit $f(b+)$ does not exist.

The typical example of an essential discontinuity given in most textbooks is the function:

$$f(a) = \begin{cases} \sin(1/a) & \text{if} \quad a \neq 0 \\ 0 \quad \text{if} \quad a = 0 \end{cases}, \tag{2.4}$$

which does not have both left limit $f(b-)$ and right limit $f(b+)$ as shown by Figure 2.9.

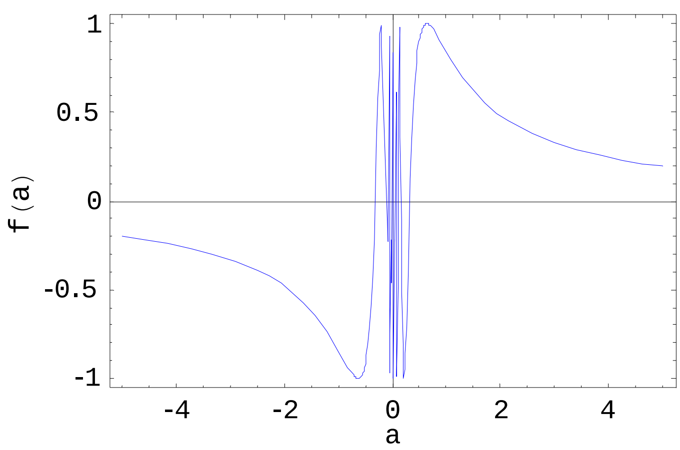


Figure 2.9: Example of an essential discontinuity.

Figure 2.10 illustrates the relationship between rcll, continuous, and discontinuous functions.

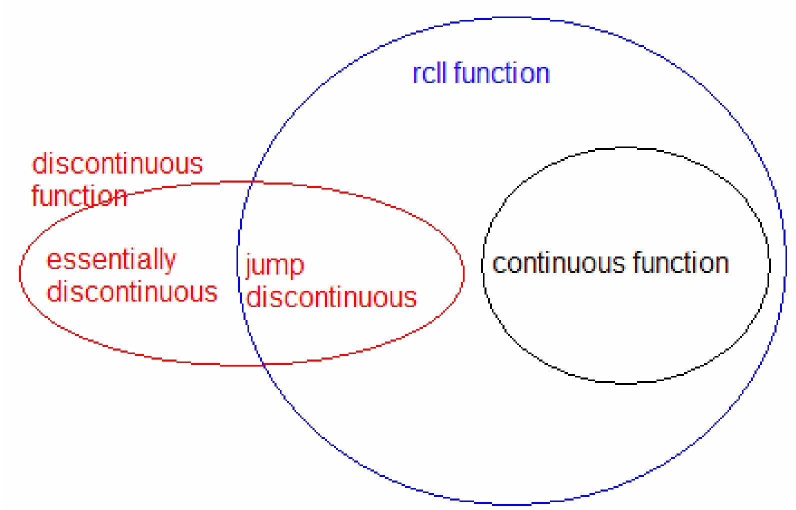


Figure 2.10: Relationship between rcll, continuous, and discontinuous functions.

[2.2] Stochastic Processes

A stochastic process is a collection of random variables:

$$(X_{t\in[0,T]}),$$

where the index denotes time. Note that we are interested in the continuous time stochastic process where the time index takes any value in the interval $t\in[0,T]$ (or it could be an infinite horizon). Discrete time stochastic process can be defined using a countable index set $t\in\mathbb{N}$:

$$(X_{t\in\mathbb{N}}).$$

A stochastic process is defined on a filtered probability space $(\Omega,\mathcal{F}_{t\in[0,T]},\mathbb{P})$ where Ω is an arbitrary set and $\mathbb{P}$ is a probability measure on $\mathcal{F}_{t\in[0,T]}$. $\mathcal{F}_{t\in[0,T]}$ is called a filtration which is an increasing family of σ-algebras of a subset of Ω which satisfy for $\forall 0\le s\le t$:

$$\mathcal{F}_s\subseteq\mathcal{F}_t.$$

Intuitively speaking, a filtration is an increasing information flow about $(X_{t\in[0,T]})$ as time progresses.

We can alternatively state that a real valued continuous time stochastic process is a random function:

$$X:[0,T]\times\Omega\to\mathbb{R}.$$

After the realization of the randomness ω, a sample path of $(X_{t\in[0,T]})$ is defined as:

$$X(\omega): t \to \mathbb{R} \text{ or } X(\omega): t \to X_t(\omega).$$

A stochastic process $(X_{t\in[0,T]})$ is said to be nonanticipating with respect to the filtration $\mathcal{F}_t$ or $\mathcal{F}_t$-adapted if the value of X_t is revealed at time t for each $t \in [0,\ T]$.

[2.2.1] Convergence of Random Variables

Definition 2.12 Pointwise convergence Let $\left(X_{n\in\mathbb{N}}(\omega)\right)$ be a sequence of real valued random variables on a space $(\Omega, \mathcal{F}, \mathbb{P})$ under a scenario (i.e. event or randomness) $\omega \in \Omega$. A sequence $\left(X_{n\in\mathbb{N}}(\omega)\right)$ is said to converge pintwisely to a random variable X if:

$$\lim_{n\to\infty} X_n(\omega) = X .$$

Pointwise convergence is the strongest notion of convergence because it requires convergence to a random variable X for all scenarios (samples) $\omega \in \Omega$, i.e. even for those scenarios with zero probability.

Definition 2.13 Almost sure convergence Let $\left(X_{n\in\mathbb{N}}(\omega)\right)$ be a sequence of real valued random variables on a space $(\Omega, \mathcal{F}, \mathbb{P})$ under a scenario $\omega \in \Omega$. A sequence $\left(X_{n\in\mathbb{N}}(\omega)\right)$ is said to converge almost surely to a random variable X if:

$$\mathbb{P}\left(\lim_{n\to\infty} X_n(\omega) = X\right) = 1.$$

Almost sure convergence is weaker than pointwise convergence since those samples $\omega \in \Omega$ with non convergence $\lim_{n\to\infty} X_n(\omega) \neq X$ have zero probability:

$$\mathbb{P}\left(\lim_{n\to\infty} X_n(\omega) = X\right) + \mathbb{P}\left(\lim_{n\to\infty} X_n(\omega) \neq X\right) = 1 + 0 = 1.$$

Almost sure convergence is used in the strong law of large numbers. Almost sure convergence implies convergence in probability which in turn implies convergence in distribution.

Definition 2.14 Convergence in probability Let $\left(X_{n\in\mathbb{N}}\right)$ be a sequence of real valued random variables on a space $(\Omega, \mathcal{F}, \mathbb{P})$. A sequence $\left(X_{n\in\mathbb{N}}\right)$ is said to converge in probability to a random variable X if for every $\varepsilon \in \mathbb{R}^+$:

$$\lim_{n\to\infty} \mathbb{P}\left(\left|X_n - X\right| > \varepsilon\right) = 0,$$

or equivalently:

$$\lim_{n\to\infty} \mathbb{P}\left(\left|X_n - X\right| \le \varepsilon\right) = 1.$$

Intuitively speaking, convergence in probability means that the probability of X_n getting closer to X rises (and eventually converges to 1) as we take n larger and larger. Convergence in probability is used in the weak law of large numbers. Convergence in probability implies convergence in distribution.

Definition 2.15 Convergence in mean square Let $\left(X_{n\in\mathbb{N}}\right)$ be a sequence of real valued random variables on a space $(\Omega, \mathcal{F}, \mathbb{P})$. A sequence $\left(X_{n\in\mathbb{N}}\right)$ is said to converge in mean square to a random variable X if for every $\varepsilon \in \mathbb{R}^+$:

$$\lim_{n\to\infty} E\left(\left|X_n - X\right|^2\right) = 0.$$

Convergence in mean square implies convergence in probability following Chebyshev's inequality.

Definition 2.16 Convergence in distribution (Weak convergence) Let $\left(X_{n\in\mathbb{N}}\right)$ be a sequence of real valued random variables on a space $(\Omega, \mathcal{F}, \mathbb{P})$. A sequence $\left(X_{n\in\mathbb{N}}\right)$ is said to converge in distribution to a random variable X if for $z \in \mathbb{R}$:

$$\lim_{n\to\infty} \mathbb{P}\left(X_n \le z\right) = \mathbb{P}\left(X \le z\right).$$

Loosely speaking, convergence in distribution means that only when we take sufficiently large n, the probability that X_n is in the interval $[a,b]$ approaches the probability that X is in the interval $[a,b]$. Convergence in distribution is the weakest definition of convergence in the sense that it does not imply any other convergence but implied by all other notions of convergence listed above.

[2.2.2] Law of Large Numbers and Central Limit Theorem

Definition 2.17 Weak law of large numbers Let $X_1, X_2, X_3 \ldots$ be *i.i.d* random variables from a distribution with mean μ and variance $\sigma^2 < \infty$. Define its sample mean as:

$$\overline{X}_n = \frac{X_1 + X_2 + \ldots + X_n}{n}.$$

Then, the sample mean $\bar{X}_n$ converges in probability to the (population) mean μ :

$$\lim_{n\to\infty}\mathbb{P}\left(\left|\bar{X}_n-\mu\right|>\varepsilon\right)=0,$$

for any $\varepsilon\in\mathbb{R}^+$. Or equivalently:

$$\lim_{n\to\infty}\mathbb{P}\left(\left|\bar{X}_n-\mu\right|<\varepsilon\right)=1.$$

Definition 2.18 Strong law of large numbers Let $X_1,X_2,X_3\ldots$ be *i.i.d* random variables from a distribution with mean $\mu<\infty$. Define its sample mean as:

$$\bar{X}_n=\frac{X_1+X_2+...+X_n}{n}.$$

Then, the sample mean $\bar{X}_n$ converges almost surely to the (population) mean μ :

$$\mathbb{P}\left(\lim_{n\to\infty}\bar{X}_n=\mu\right)=1,$$

for any $\varepsilon\in\mathbb{R}^+$.

Definition 2.19 Central limit theorem Let $X_1,X_2,X_3\ldots$ be *i.i.d* random variables from a distribution with mean $\mu<\infty$ and variance $\sigma^2<\infty$. Define the sum of a sequence of random variables as:

$$S_n=X_1+X_2+...+X_n.$$

We know the followings:

$$E\left[S_n\right]=E\left[X_1\right]+\left[X_2\right]+...\left[X_n\right]=n\mu,$$
$$Var\left[S_n\right]=Var\left[X_1\right]+Var\left[X_2\right]+...+Var\left[X_n\right]=n\sigma^2.$$

Then, informally, the sum S_n converges in distribution to a normal distribution with mean $n\mu$ and variance $n\sigma^2$ as $n\to\infty$:

$$\lim_{n\to\infty}\mathbb{P}\left(S_n\le b\right)=\mathbb{P}\left(Y\le b\right)$$
$$=\int_{-\infty}^{b}\frac{1}{\sqrt{2\pi n\sigma^2}}\exp\left\{-\frac{1}{2}\frac{\left(Y-n\mu\right)^2}{n\sigma^2}\right\}dY,$$

where Y is a normal random variable, i.e. $Y \sim N\left(n\mu, n\sigma^2\right)$.

Formal central limit theorem is a standardization of the above informal one. Define a random variable Z_n as:

$$Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}}.$$

Then, Z_n converges in distribution to the standard normal distribution as $n \to \infty$:

$$\begin{aligned}\lim_{n\to\infty} \mathbb{P}\left(Z_n \le b\right) &= \mathbb{P}\left(Z \le b\right) \\ &= \int_{-\infty}^{b} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} Z^2\right) dZ ,\end{aligned}$$

where Z is the standard normal random variable, i.e. $Z \sim N\left(0,1\right)$.

[2.2.3] Inequalities

Definition 2.20 Markov's inequality Let X be a nonnegative random variable. Then, for any $b \in \mathbb{R}^+$:

$$\mathbb{P}\left(X \ge b\right) \le \frac{E\left[X\right]}{b}.$$

Proof

$$E\left[X\right] = \int_0^\infty X d\mathbb{P}\left(X\right) = \int_0^b X d\mathbb{P}\left(X\right) + \int_b^\infty X d\mathbb{P}\left(X\right).$$

This means:

$$E\left[X\right] \ge \int_b^\infty X d\mathbb{P}\left(X\right) \ge \int_b^\infty b d\mathbb{P}\left(X\right) = b\int_b^\infty d\mathbb{P}\left(X\right) = b\mathbb{P}\left(X \ge b\right).$$

Thus:

$$\frac{E\left[X\right]}{b} \ge \mathbb{P}\left(X \ge b\right).$$

□

Markov's inequality provides an upper bound of the probability that a nonnegative random variable is greater than an arbitrary positive constant b by relating a probability to an expectation. A variant of Markov's inequality is called Chebyshev's inequality.

Definition 2.21 Chebyshev's inequality Let X be a random variable on $\mathbb{R}$ (i.e. both $\mathbb{R}^+$ and $\mathbb{R}^-$) with mean $\mu < \infty$ and variance $\sigma^2 < \infty$. Then, for any $k \in \mathbb{R}^+$:

$$\mathbb{P}\left(\left|X-\mu\right| \geq k\right) \leq \frac{\sigma^2}{k^2} .$$

Proof

Start with Markov's inequality:

$$\mathbb{P}\left(X \geq b\right) \leq \frac{E\left[X\right]}{b} .$$

Replace a random variable X with a random variable $(X-\mu)^2$ and b with k^2 :

$$\mathbb{P}\left((X-\mu)^2 \geq k^2\right) \leq \frac{E\left[(X-\mu)^2\right]}{k^2} = \frac{\sigma^2}{k^2} ,$$

which in turn indicates:

$$\mathbb{P}\left(\left|X-\mu\right| \geq k\right) \leq \frac{\sigma^2}{k^2} ,$$
$$\mathbb{P}\left(\left|X-\mu\right| \geq k\sigma\right) \leq \frac{1}{k^2} .$$

□

Chebyshev's inequality provides bounds of random variables from any distributions as long as their means and variances are known. For example, when $k = 2$:

$$\mathbb{P}\left(\left|X-\mu\right| \geq 2\sigma\right) \leq \frac{1}{4}$$
$$\mathbb{P}\left(-X+\mu \geq 2\sigma, X-\mu \geq 2\sigma\right) \leq \frac{1}{4}$$
$$\mathbb{P}\left(X \leq \mu - 2\sigma, X \geq \mu + 2\sigma\right) \leq \frac{1}{4}$$
$$\mathbb{P}\left(\mu - 2\sigma \leq X \leq \mu + 2\sigma\right) \geq \frac{3}{4} .$$

This tells us that the probability that any random variable lies within two standard deviations is at least .75.

Definition 2.22 Cauchy-Schwarz's inequality Let X and Y be jointly distributed random variables on $\mathbb{R}$ with each having finite variance. Then:

$$\left(E[XY]\right)^2 \leq E[X^2]E[Y^2].$$

Proof

Refer to Abramowitz (1993).

[2.3] Putting Structure on Stochastic Processes

The purpose of any mathematical (statistical) modeling regardless of the field is to fit less complicated models to the highly complicated real world phenomena as accurate as possible. Mathematical models are less complicated in the sense that they make some simplifying assumptions or put some simplifying structures (restrictions) on the real world phenomena for the purpose of gaining tractability. There are some popular dependence structures put on stochastic processes which mathematicians have developed and used for years.

[2.3.1] Processes with Independent and Stationary Increments: Imposing Structure on a Probability Measure $\mathbb{P}$

Before giving the definition of processes with independent and stationary increments, we must know the basics.

Definition 2.23 Conditional probability The conditional probability of an arbitrary event A given an event with positive probability B is:

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

When $\mathbb{P}(B) = 0$, $\mathbb{P}(A|B)$ is undefined.

Definition 2.24 Statistical (Stochastic) independence Two arbitrary events A and B are said to be independent, if and only if:

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

This definition of independence has two advantages. Firstly, it is symmetric in A and B. In other words, an event A's independence of an event B implies an event B's independence of an event A. Secondly, this definition holds even when an event B has zero probability, i.e. $\mathbb{P}(B) = 0$.

When two arbitrary events A and B are independent, from the definition of a conditional probability:

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

It is important to note that this is a result of statistical independence and not the definition. This is because the above equation is not true (i.e. undefined) when $\mathbb{P}(B) = 0$ and it is not symmetric in that $\mathbb{P}(A|B) = \mathbb{P}(A)$ does not necessarily imply $\mathbb{P}(B|A) = \mathbb{P}(B)$.

Definition 2.25 Mutual statistical independence Arbitrary events A_1, A_2, …, A_n are said to be mutually independent, if and only if:

$$\mathbb{P}(A_1 \cap A_2 \cap ... \cap A_n) = \mathbb{P}(A_1)\mathbb{P}(A_2)...\mathbb{P}(A_n).$$

Definition 2.26 Processes with Independent and Stationary Increments A stochastic process $\left(X_{t\in[0,T]}\right)$ with values in $\mathbb{R}$ on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,T]}, \mathbb{P})$ is said to be a process with independent and stationary increments if it satisfies the following conditions:

(1) Its increments are independent. In other words, for $t_1 < t_2 < ... < t_n$:

$$\mathbb{P}(X_{t_2} - X_{t_1} \cap X_{t_3} - X_{t_2} \cap ... \cap X_{t_n} - X_{t_{n-1}}) = \mathbb{P}(X_{t_2} - X_{t_1})\mathbb{P}(X_{t_3} - X_{t_2})...\mathbb{P}(X_{t_n} - X_{t_{n-1}}).$$

(2) Its increments are stationary: i.e. for $\forall h \in \mathbb{R}^+$, $X_{t+h} - X_t$ has the same distribution as X_h. In other words, the distribution of increments does not depend on t (i.e. temporal homogeneity).

Consider an increasing sequence of time $0 < t_1 < t_2 < ... < t_n < t < u < \infty$ where t is the present. As a result of independent increments condition:

$$\begin{aligned}
&\mathbb{P}(X_u - X_t | X_{t_1} - X_0, X_{t_2} - X_{t_1}, ..., X_t - X_{t_n}) \\
&= \frac{\mathbb{P}(X_u - X_t \cap X_{t_1} - X_0, X_{t_2} - X_{t_1}, ..., X_t - X_{t_n})}{\mathbb{P}(X_{t_1} - X_0, X_{t_2} - X_{t_1}, ..., X_t - X_{t_n})} \\
&= \frac{\mathbb{P}(X_u - X_t)\mathbb{P}(X_{t_1} - X_0, X_{t_2} - X_{t_1}, ..., X_t - X_{t_n})}{\mathbb{P}(X_{t_1} - X_0, X_{t_2} - X_{t_1}, ..., X_t - X_{t_n})} \\
&= \mathbb{P}(X_u - X_t),
\end{aligned}$$

which means that there is no correlation (probabilistic dependence structure) on the increments among the past, the present, and the future.

For example, independent increments condition means that when modeling a log stock price $\ln S_t$ as an independent increment process, the probability distribution of a log stock price in year 2005 – 2006 is independent of the way the log stock price increment has evolved over the years (i.e. stock price dynamics), i.e. it doesn't matter if this stock crushes or soars in year 2004 – 2005):

$$\begin{aligned}&\mathbb{P}(\ln S_{2006} - \ln S_{2005} \,|\, ..., \ln S_{2003} - \ln S_{2002}, \ln S_{2004} - \ln S_{2003}, \ln S_{2005} - \ln S_{2004}) \\ &= \mathbb{P}(\ln S_{2006} - \ln S_{2005}).\end{aligned}$$

Using the simple relationship $X_u \equiv (X_u - X_t) + X_t$ for an increasing sequence of time $0 < t_1 < t_2 < ... < t_n < t < u < \infty$:

$$\begin{aligned}\mathbb{P}(X_u \,|\, X_0, X_{t_1}, X_{t_2}, ..., X_{t_n}, X_t) &= \mathbb{P}((X_u - X_t) + X_t \,|\, X_0, X_{t_1}, X_{t_2}, ..., X_{t_n}, X_t) \\ &= \mathbb{P}(X_u \,|\, X_t),\end{aligned}$$

which holds because an increment $(X_u - X_t)$ is independent of X_t by definition and the value of X_t depends on its realization $X_t(\omega)$. This is a strong probabilistic structure imposed on a stochastic process because this means that the conditional probability of the future value X_u depends only on the previous realization $X_t(\omega)$ and not on the entire past history of realizations $X_0, X_{t_1}, X_{t_2}, ..., X_{t_n}, X_t$ (i.e. called Markov property which is discussed soon).

Although this condition seems too strong, it imposes a very tractable property on the process. Because if two variables X and Y are independent:

$$\begin{aligned}&E[XY] = E[X]E[Y], \\ &Var[X+Y] = Var[X] + Var[Y], \\ &Cov[X,Y] = 0 \text{ (i.e. } Corr[X,Y] = 0\text{)}.\end{aligned}$$

Stationary increments condition means that the distributions of increments $X_{t+h} - X_t$ do not depend on the time t, but they depend on the time-distance h of two observations (i.e. interval of time). In other words, the probability density function of increments does not change over time. For example, if you model a log stock price $\ln S_t$ as a process with stationary increments, the distribution of increment in year 2005 – 2006 is the same as that in year 2050 – 2051:

$$\ln S_{2006} - \ln S_{2005} \overset{d}{=} \ln S_{2051} - \ln S_{2050}.$$

There is no doubt that the above independent and stationary increments conditions impose a strong structure on a stochastic process $\left(X_{t\in[0,T]}\right)$, as a result of these restrictions, the mean and variance of the process is tractable:

$$E[X_t] = \mu_0 + \mu_1 t\,,$$
$$Var[X_t] = \sigma_0^{\,2} + \sigma_1^{\,2} t\,,$$

where $\mu_0 = E[X_0]$, $\mu_1 = E[X_1] - \mu_0$, $\sigma_0^{\,2} = E[(X_0 - \mu_0)^2]$, and $\sigma_1^{\,2} = E[(X_1 - \mu_1)^2] - \sigma_0^{\,2}$.

[2.3.2] Martingale: Structure on Conditional Expectation

[2.3.2.1] Definition of Martingale

Originally, the word 'martingale' comes from a French acronym of a gambling strategy. Imagine a coin flip gamble in which you win if a head turns up and you lose if a tail turns up. Martingale strategy requires a gambler to double his bet after every loss. Following martingale strategy, a gambler can recover all the losses he made and end up with an initial amount of his wealth plus an initial bet. Table 2.1 gives a sample path of a martingale strategy in which a gambler initially owns \$200 of wealth, start betting with a stake of \$2, and due to his bad luck his first win comes at the seventh trial. As you can see, he basically ends up where he started, i.e. his initial wealth of \$200 (plus an initial bet of \$2). Thus, a martingale strategy tells that after gambling many hours a gambler gains nothing (loses nothing) and his wealth remains constant on average.

Table 2.1 Martingale Gambling Strategy

Trial	0	1	2	3	4	5	6	7
Result		Loss	Loss	Loss	Loss	Loss	Loss	Win
Bet		\$2	\$4	\$8	\$16	\$32	\$64	\$128
Net Gain		-\$2	-\$4	-\$8	-\$16	-\$32	-\$64	+\$128
Wealth	\$200	\$198	\$194	\$186	\$170	\$138	\$74	+\$202

In probability theory, a stochastic process is said to be a martingale if its sample path has no trend. Formally, a martingale is defined as the follows.

Definition 2.27 Martingale Consider a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,T]}, \mathbb{P})$. A rcll stochastic process $(X_t)_{t\in[0,\,T]}$ is said to be a martingale with respect to the filtration $\mathcal{F}_t$ and under the probability measure $\mathbb{P}$ if it satisfies the following conditions:

(1) X_t is nonanticipating.

(2) $E[|X_t|] < \infty$ for $\forall t \in [0,\ T]$. Finite mean condition.

(3) $E[X_u | \mathcal{F}_t] = X_t$ for $\forall u > t$.

In other words, if a stochastic process is a martingale, then, the best prediction of its future value is its present value. Note that the definition of martingale makes sense only when the underlying probability measure P and the filtration $(\mathcal{F}_t)_{t\in[0,T]}$ have been specified.

The fundamental property of a martingale process is that its future variations are completely unpredictable with the filtration $\mathcal{F}_t$:

$$\forall u > 0,\ E[x_{t+u} - x_t \,|\mathcal{F}_t] = E[x_{t+u}\,|\mathcal{F}_t] - E[x_t\,|\mathcal{F}_t] = x_t - x_t = 0 .$$

Finite mean condition is necessary to ensure the existence of the conditional expectation.

[2.3.2.2] Example of Continuous Martingale: Standard Brownian Motion

Let $(B_{t\in[0,\infty)})$ be a standard Brownian motion process defined on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,\infty)}, \mathbb{P})$. Then, $(B_{t\in[0,\infty)})$ is a continuous martingale with respect to the filtration $\mathcal{F}_{t\in[0,\infty)}$ and the probability measure $\mathbb{P}$.

Proof

By definition, $(B_{t\in[0,\infty)})$ is a nonanticipating process (i.e. $\mathcal{F}_{t\in[0,\infty)}$ - adapted process) with the finite mean $E[|B_t|] = 0 < \infty$ for $\forall t \in [0,\infty)$. For $\forall 0 \le t \le u < \infty$:

$$B_u = B_t + \int_t^u dB_v .$$

Using the fact that a Brownian motion is a nonanticipating process, i.e. $E[B_t\,|\mathcal{F}_t] = B_t$:

$$E[B_u - B_t\,|\mathcal{F}_t] = E[B_u\,|\mathcal{F}_t] - E[B_t\,|\mathcal{F}_t] = E[B_t + \int_t^u dB_v\,|\mathcal{F}_t] - B_t$$
$$E[B_u - B_t\,|\mathcal{F}_t] = E[B_t\,|\mathcal{F}_t] + E[\int_t^u dB_v\,|\mathcal{F}_t] - B_t$$
$$E[B_u - B_t\,|\mathcal{F}_t] = B_t + 0 - B_t = 0 ,$$

or in other words:

$$E[B_u\,|\mathcal{F}_t] = E[B_t + \int_t^u dB_v\,|\mathcal{F}_t] = E[B_t\,|\mathcal{F}_t] + E[\int_t^u dB_v\,|\mathcal{F}_t] = B_t + 0$$
$$E[B_u\,|\mathcal{F}_t] = B_t ,$$

which is a martingale condition.

□

Let $(B_{t\in[0,\infty)})$ be a standard Brownian motion process defined on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,\infty)}, \mathbb{P})$. Then, a Brownian motion with drift $(X_{t\in[0,\infty)}) \equiv (\mu t + \sigma B_{t\in[0,\infty)})$ is not a continuous martingale with respect to the filtration $\mathcal{F}_{t\in[0,\infty)}$ and the probability measure $\mathbb{P}$.

Proof

By definition, $(X_{t\in[0,\infty)})$ is a nonanticipating process (i.e. $\mathcal{F}_{t\in[0,\infty)}$- adapted process) with the finite mean $E[X_t] = E[\mu t + \sigma B_t] = \mu t < \infty$ for $\forall t \in [0,\infty)$ and $\mu \in \mathbb{R}$. For $\forall 0 \le t \le u < \infty$:

$$X_u = X_t + \int_t^u dX_v \, .$$

Using the fact that a Brownian motion with drift is a nonanticipating process, i.e. $E[X_t | \mathcal{F}_t] = X_t$:

$$E[X_u | \mathcal{F}_t] = E[X_t + \int_t^u dX_v | \mathcal{F}_t] = E[X_t | \mathcal{F}_t] + E[\int_t^u dX_v | \mathcal{F}_t]$$
$$E[X_u | \mathcal{F}_t] = X_t + \mu(u-t) \, ,$$

which violates a martingale condition.

□

But one way to transform nonmartingales into martingales is to make the process driftless. In other words, eliminating the trend of the process which is sometimes called a detrending. Consider the following example.

A detrended Brownian motion with drift defined as:

$$(X_{t\in[0,\infty)} - \mu t) \equiv (\mu t + \sigma B_{t\in[0,\infty)} - \mu t) \equiv (\sigma B_{t\in[0,\infty)}) \, ,$$

is a continuous martingale with respect to the filtration $\mathcal{F}_{t\in[0,\infty)}$ and the probability measure $\mathbb{P}$.

Proof

For $\forall 0 \le t \le u < \infty$:

$$E[X_u - \mu u | \mathcal{F}_t] = E[(X_t - \mu t) + (\int_t^u dX_v - \mu \int_t^u dv) | \mathcal{F}_t]$$

$$E[X_u - \mu u|\mathcal{F}_t] = E[(X_t - \mu t)|\mathcal{F}_t] + E[(\int_t^u dX_v - \mu \int_t^u dv)|\mathcal{F}_t]$$
$$E[X_u - \mu u|\mathcal{F}_t] = X_t - \mu t + \mu(u-t) - \mu(u-t)$$
$$E[X_u - \mu u|\mathcal{F}_t] = X_t - \mu t\,,$$

which satisfies a martingale condition.

□

[2.3.2.3] Martingale Asset Pricing

Most of financial asset prices are not martingales because they are not completely unpredictable and most financial time series have trends. Consider a stock price process $\{S_t; 0 \le t \le T\}$ on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,T]}, \mathbb{P})$ and let r be the risk-free interest rate. In a small time interval Δ, risk-averse investors expect S_t to grow at some positive rate. This can be written as under actual probability measure $\mathbb{P}$:

$$E^{\mathbb{P}}[S_{t+\Delta}|\mathcal{F}_t] > S_t\,.$$

This means that a stock price S_t is not martingale under $\mathbb{P}$ and with respect to $\mathcal{F}_t$. To be more precise, risk-averse investors expect S_t to grow at a rate greater than r because a stock is risky:

$$E^{\mathbb{P}}[e^{-r\Delta} S_{t+\Delta}|\mathcal{F}_t] > S_t\,.$$

The stock price discounted by the risk-free interest rate $e^{-r\Delta} S_{t+\Delta}$ is not martingale under $\mathbb{P}$ and with respect to $\mathcal{F}_t$.

How can we convert a discounted stock price $e^{-r\Delta} S_{t+\Delta}$ into a martingale? First approach is to eliminate the trend. The trend in this case is the risk premium π which risk-averse investors demand for bearing extra amount of risk. If we can estimate π correctly, a discounted stock price $e^{-r\Delta} S_{t+\Delta}$ can be converted into a martingale by detrending:

$$E^{\mathbb{P}}[e^{-\pi\Delta} e^{-r\Delta} S_{t+\Delta}|\mathcal{F}_t] = E^{\mathbb{P}}[e^{-(r+\pi)\Delta} S_{t+\Delta}|\mathcal{F}_t] = S_t\,.$$

But this approach involves the rather difficult job of estimating the risk premium π and is not used in quantitative finance. Martingale asset pricing uses the second approach to convert non-martingales into martingales by changing the probability measure. We will try to find an equivalent probability measure $\mathbb{Q}$ (called risk-neutral measure) under which a discounted stock price becomes martingale:

$$E^{\mathbb{Q}}[e^{-r\Delta} S_{t+\Delta}|\mathcal{F}_t] = S_t\,.$$

[2.3.2.4] Submartingales and Supermartingales

Definition 2.28 Submartingale Consider a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,T]}, \mathbb{P})$. A rcll stochastic process $(X_t)_{t\in[0,\,T]}$ is said to be a submartingale with respect to the filtration $\mathcal{F}_t$ and under the probability measure $\mathbb{P}$ if it satisfies the following conditions:

(1) X_t is nonanticipating.
(2) $E[|X_t|] < \infty$ for $\forall t \in [0,\ T]$. Finite mean condition.
(3) $E[X_u | \mathcal{F}_t] \geq X_t$ for $\forall u > t$.

Intuitively, a submartingale is a stochastic process with a positive (upward) trend. A submartingale gains or grows on average as time progresses.

Definition 2.29 Supermartingale Consider a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,T]}, \mathbb{P})$. A rcll stochastic process $(X_t)_{t\in[0,\,T]}$ is said to be a supermartingale with respect to the filtration $\mathcal{F}_t$ and under the probability measure $\mathbb{P}$ if it satisfies the following conditions:

(1) X_t is nonanticipating.
(2) $E[|X_t|] < \infty$ for $\forall t \in [0,\ T]$. Finite mean condition.
(3) $E[X_u | \mathcal{F}_t] \leq X_t$ for $\forall u > t$.

Intuitively, a supermartingale is a stochastic process with a negative (downward) trend. A supermartingale loses or declines on average as time progresses.

By definition, any martingale is a submartingale and a supermartingale.

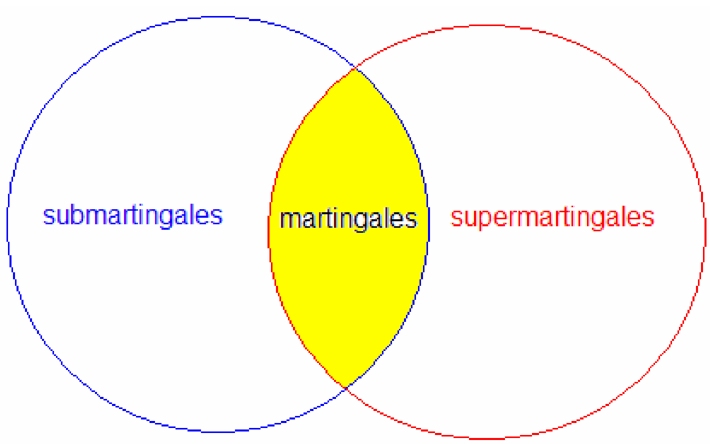


Figure 2.11: Relationship among martingales, submartingales, and supermartingales.

[2.3.3] Markov Processes: Structure on Conditional Probability

This section gives a brief introduction to a class of stochastic processes called Markov processes which impose a restriction on the conditional probabilities. This differs from martingales which impose a structure on conditional expectations.

[2.3.3.1] Discrete Time Markov Chains

Definition 2.30 Discrete time Markov chain Consider a discrete time stochastic process $(X_n)_{n\in\mathbb{N}}$ (i.e. $n = 0,1,2,...$) defined on a filtered probability space $(\Omega, \mathcal{F}_{n\in\mathbb{N}}, \mathbb{P})$ which takes values in a countable or a finite set E called a state space of the process. A realization X_n is said to be in state $i \in E$ at time n if $X_n = i$. An E-valued discrete time Markov chain is a stochastic process which satisfies for $\forall n \in \mathbb{N}$ and $\forall i, j \in E$:

$$\mathbb{P}(X_{n+1} = j \,|\, X_0, X_1, X_2, ..., X_n = i) = \mathbb{P}(X_{n+1} = j \,|\, X_n = i).$$

This is called a Markov property. Markov property means that the probability of a random variable X_{n+1} at time $n+1$ (tomorrow) being in a state j conditional on the entire history of the stochastic process $(X_0, X_1, X_2, ..., X_n)$ is equal to the probability of a random variable X_{n+1} at time $n+1$ (tomorrow) being in a state j conditional only on the value of a random variable at time n (today). In other words, the history (sample path) of the stochastic process $(X_0, X_1, X_2, ..., X_n)$ is of no importance in that the way this stochastic process evolved or the dynamics $(X_1 - X_0, X_2 - X_1, ...)$ does not mean a thing in terms of the conditional probability of the process. The only factor which influences the conditional probability of a random variable X_{n+1} at time $n+1$ (tomorrow) is the state of a random variable at time n (today).

The probability $\mathbb{P}(X_{n+1} = j \,|\, X_n = i)$ which is a conditional probability of moving from a state i at time n to a state j at time $n+1$ is called a one step transition probability. In the general cases, transition probabilities are dependent on the states and time such that $\forall m \neq n \in \mathbb{N}$:

$$\mathbb{P}(X_{n+1} = j \,|\, X_n = i) \neq \mathbb{P}(X_{m+1} = j \,|\, X_m = i).$$

When transition probabilities are independent of time n, discrete time Markov chains are said to be time homogeneous or stationary.

Definition 2.31 Time homogeneous (stationary) discrete time Markov chain Consider a discrete time stochastic process $(X_n)_{n\in\mathbb{N}}$ (i.e. $n = 0,1,2,...$) defined on a filtered probability space $(\Omega, \mathcal{F}_{n\in\mathbb{N}}, \mathbb{P})$ which takes values in a countable or a finite set E called a state space of the process. A realization X_n is said to be in state $i \in E$ at time n

if $X_n = i$. An E -valued time homogeneous discrete time Markov chain is a stochastic process which satisfies for $\forall n \in \mathbb{N}$ and $\forall i, j \in E$:

$$\begin{aligned}\mathbb{P}(X_{n+1} = j \,|\, X_0, X_1, X_2, ..., X_n = i) &= \mathbb{P}(X_{n+1} = j \,|\, X_n = i) \\ &= \mathbb{P}(X_1 = j \,|\, X_0 = i) \\ &= \mathbb{P}(j \,|\, i) .\end{aligned}$$

In other words, transition probabilities do not depend on time n and only depend on transition states from i to j . A matrix of transition probabilities $\mathbb{P} = \left\| \mathbb{P}(j \,|\, i) \right\|_{i,\, j \in E}$ is called a transition probability matrix:

$$\left\| \mathbb{P}(j \,|\, i) \right\|_{i,\, j \in E} = \left\| \begin{matrix} \mathbb{P}(0|0) & \mathbb{P}(1|0) & \mathbb{P}(2|0) & \mathbb{P}(3|0) & \cdots \\ \mathbb{P}(0|1) & \mathbb{P}(1|1) & \mathbb{P}(2|1) & \mathbb{P}(3|1) & \cdots \\ \mathbb{P}(0|2) & \mathbb{P}(1|2) & \mathbb{P}(2|2) & \mathbb{P}(3|2) & \cdots \\ \vdots & \vdots & \vdots & \vdots & \\ \mathbb{P}(0|i) & \mathbb{P}(1|i) & \mathbb{P}(2|i) & \mathbb{P}(3|i) & \cdots \\ \vdots & \vdots & \vdots & \vdots & \end{matrix} \right\| .$$

Transition probabilities $\mathbb{P}(j \,|\, i)$ satisfy the following conditions:

(1) $\mathbb{P}(j \,|\, i) \geq 0$ for $\forall i, j \in E$.
(2) $\sum_{j \in E} \mathbb{P}(j \,|\, i) = 1$ for $\forall i \in E$.

Condition (2) guarantees the occurrence of a transition including a case in which the state remains unchanged.

Proposition 2.1 Defining a discrete time Markov chain An E -valued general discrete time Markov chain $(X_n)_{n \in \mathbb{N}}$ is completely defined if it satisfies the following conditions:

(1) All transition probabilities $\mathbb{P}(X_{n+1} = i_{n+1} \,|\, X_n = i_n)$ are known.
(2) The probability distribution of the initial (i.e. time 0) state of the Markov chain $\mathbb{P}(X_0 = i_0) = \mathbb{P}_0$ is known.

Proof

Consider obtaining the joint probability distribution of an E -valued general discrete time Markov chain $(X_n)_{n \in \mathbb{N}}$. From the definition of a conditional probability:

$$\mathbb{P}(X_0 = i_0, X_1 = i_1, X_2 = i_2, ..., X_n = i_n)$$
$$= \mathbb{P}(X_n = i_n \mid X_0 = i_0, X_1 = i_1, X_2 = i_2, ..., X_{n-1} = i_{n-1})\mathbb{P}(X_0 = i_0, X_1 = i_1, X_2 = i_2, ..., X_{n-1} = i_{n-1}).$$

Since $(X_n)_{n\in\mathbb{N}}$ is a Markov chain:

$$\mathbb{P}(X_n = i_n \mid X_0 = i_0, X_1 = i_1, X_2 = i_2, ..., X_{n-1} = i_{n-1}) = \mathbb{P}(X_n = i_n \mid X_{n-1} = i_{n-1}) = \mathbb{P}(i_n \mid i_{n-1}).$$

Joint probability can be calculated as:

$$\begin{aligned}
\mathbb{P}(X_0 &= i_0, X_1 = i_1, X_2 = i_2, ..., X_n = i_n) \\
&= \mathbb{P}(i_n \mid i_{n-1})\mathbb{P}(X_0 = i_0, X_1 = i_1, X_2 = i_2, ..., X_{n-1} = i_{n-1}) \\
&= \mathbb{P}(i_n \mid i_{n-1})\mathbb{P}(X_{n-1} = i_{n-1} \mid X_0 = i_0, X_1 = i_1, X_2 = i_2, ..., X_{n-2} = i_{n-2}) \\
&\quad \times\mathbb{P}(X_0 = i_0, X_1 = i_1, X_2 = i_2, ..., X_{n-2} = i_{n-2}) \\
&= \mathbb{P}(i_n \mid i_{n-1})\mathbb{P}(i_{n-1} \mid i_{n-2})\mathbb{P}(X_0 = i_0, X_1 = i_1, X_2 = i_2, ..., X_{n-2} = i_{n-2}) \\
&= \mathbb{P}(i_n \mid i_{n-1})\mathbb{P}(i_{n-1} \mid i_{n-2})...\mathbb{P}(i_2 \mid i_1)\mathbb{P}(i_1 \mid i_0)\mathbb{P}_0
\end{aligned}$$

□

Consider a transition probability of a time homogeneous discrete time Markov chain $(X_n)_{n\in\mathbb{N}}$ from a state i at time k (i.e. $X_k = i$) to a state j at time $k+n$. This is called a n-step transition probability and expressed as:

$$\mathbb{P}(X_{k+n} = j \mid X_k = i) = \mathbb{P}(X_n = j \mid X_0 = i) = \mathbb{P}^{(n)}(j \mid i).$$

Proposition 2.2 n step transition probability matrix (a special case of Chapman-Kolmogorov equation) Consider a time homogeneous discrete time Markov chain $(X_n)_{n\in\mathbb{N}}$ defined on a filtered probability space $(\Omega, \mathcal{F}_{n\in\mathbb{N}}, \mathbb{P})$ which takes values in a countable or a finite set E called a state space of the process. Then, its n-step transition probability matrix from a state i at time k (i.e. $X_k = i$) to a state j at time $k+n$ is given by for $\forall k, n \in \mathbb{N}$ and $\forall i, j \in E$:

$$\mathbb{P}^{(n)}(j \mid i) = \mathbb{P}^{n}(j \mid i) = \sum\nolimits_{v\in E} \mathbb{P}^{(y)}(v \mid i)\mathbb{P}^{(z)}(j \mid v) = \sum\nolimits_{v\in E} \mathbb{P}^{y}(v \mid i)\mathbb{P}^{z}(j \mid v).$$

where $y + z = n$ and $\mathbb{P}^{(0)}(j \mid i)$ is defined as:

$$\mathbb{P}^{(0)}(j \mid i) = \begin{cases} 1 & \text{for} \quad i = j \\ 0 & \text{for} \quad i \neq j \end{cases}.$$

Proof

When $n = 1$:

$$\mathbb{P}^{(1)}(j|i) = \mathbb{P}(j|i).$$

When $n = 2$:

$$\mathbb{P}^{(2)}(j|i) = \sum_{v \in E} \mathbb{P}(v|i)\mathbb{P}(j|v).$$

By induction:

$$\mathbb{P}^{(n+1)}(j|i) = \sum_{v \in E} \mathbb{P}(v|i)\mathbb{P}^{n}(j|v).$$

□

One interesting topic about this n step transition probability matrix is its asymptotic behavior as $n \to \infty$. As n becomes larger, the initial state i becomes less important and in the limit as $n \to \infty$, $\mathbb{P}^{n}(j|i)$ is independent of i. We recommend Karlin and Taylor (1975) for more details.

[2.3.3.2] Markov Processes

Definition 2.32 Markov Processes (Continuous time Markov chains) Consider a continuous time stochastic process $\left(X_{t \in [0,T]}\right)$ defined on a filtered probability space $(\Omega, \mathcal{F}_{t \in [0,T]}, \mathbb{P})$ which takes values in $\mathbb{N}$ (for simplicity) called a state space of the process. $\left(X_{t \in [0,T]}\right)$ is said to be a time homogeneous Markov process if for $\forall h \in \mathbb{R}^{+}$ and $\forall i, j \in \mathbb{N}$:

$$\mathbb{P}_h(j|i) = \mathbb{P}(X_{t+h} = j | \mathcal{F}_t) = \mathbb{P}(X_{t+h} = j | X_t = i).$$

Markov property means that the probability of a random variable X_{t+h} at time $t+h$ (tomorrow) being in a state j conditional on the entire history of the stochastic process $\mathcal{F}_{[0,t]} \equiv X_{[0,t]}$ is equal to the probability of a random variable X_{t+h} at time $t+h$ (tomorrow) being in a state j conditional only on the value of a random variable at time t (today). In other words, the history (sample path) of the stochastic process $\mathcal{F}_{[0,t]}$ is of no importance in that the way this stochastic process evolved or the dynamics does not mean a thing in terms of the conditional probability of the process.

We discuss Markov processes more in detail in section 3.8.

[2.4] Sample Path Properties of Stochastic Processes

[2.4.1] Continuous Stochastic Process

In this section we give the formal definition of the continuity of a sample path of a stochastic process. Note that the continuity of path and continuity of time are different subjects. In other words, a continuous time stochastic process does not imply continuous stochastic process. For example, a Poisson process is a continuous time stochastic process, but it has discontinuous sample paths.

There are different notions of continuity of sample paths which use different notions of convergence of random variables we saw in section 2.2.1.

Definition 2.33 Continuous in mean square A real valued stochastic process $(X_{t\in[0,T]})$ on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,T]}, \mathbb{P})$ is said to be continuous in mean square if for $\forall t \in [0,\ T]$:

$$\lim_{s\to t} E[\left|X_s - X_t\right|^2] = 0 .$$

Continuity in mean square implies continuity in probability following Chebyshev's inequality.

Definition 2.34 Continuous in probability A real valued stochastic process $(X_{t\in[0,T]})$ on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,T]}, \mathbb{P})$ is said to be continuous in probability if for $\forall t \in [0,\ T]$ and every $\varepsilon \in \mathbb{R}^+$:

$$\lim_{s\to t} \mathbb{P}\left(\left|X_s - X_t\right| > \varepsilon\right) = 0 ,$$

or equivalently:

$$\lim_{s\to t} \mathbb{P}\left(\left|X_s - X_t\right| \le \varepsilon\right) = 1 .$$

Intuitively speaking, continuity in probability means that the probability of X_s getting closer to X_t rises (and eventually converges to 1) as s approaches t.

For example, a Brownian motion process is continuous in mean square and continuous in probability although its proof is not that easy. But it turns out that the above definitions of continuity are too loose because a Poisson process can be proven to be continuous in mean square and probability (consult Karlin and Taylor (1975) for details). Therefore, a more strict definition of continuity is used for the definition of a continuity of a sample path of a stochastic process.

Definition 2.35 Continuous stochastic process A real valued nonanticipating stochastic process $(X_{t\in[0,T]})$ on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,T]}, \mathbb{P})$ is said to be (almost surely) continuous if a sample path of the process $(X_{t\in[0,T]}(\omega))$ is almost surely a continuous function for $\forall t \in [0,\ T]$. In other words, a sample path of the process $(X_{t\in[0,T]}(\omega))$ satisfies for $\forall t \in [0,\ T]$:

(1) Right limit of the process as s approaches t from the above (right hand side) exists, i.e. $\lim_{s\to t, s>t} X_s = X_{t+}$. Left limit of the process as s approaches t from the below (left hand side) exists, i.e. $\lim_{s\to t, s<t} X_s = X_{t-}$.

(2) $X_{t+} = X_{t-} = X_t$.

This means that a continuous stochastic process is a right continuous and left continuous stochastic process.

[2.4.2] Right Continuous with Left Limit (RCLL) Stochastic Processes

Definition 2.36 Right continuous with left limit (rcll) stochastic processes A real valued nonanticipating stochastic process $(X_{t\in[0,T]})$ on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,T]}, \mathbb{P})$ is said to be a rcll stochastic process if for $\forall t \in [0,\ T]$:

(1) Right limit of the process as s approaches t from the above (right hand side) exists, i.e. $\lim_{s\to t, s>t} X_s = X_{t+}$. Left limit of the process as s approaches t from the below (left hand side) exists, i.e. $\lim_{s\to t, s<t} X_s = X_{t-}$.

(2) $X_{t+} = X_t$.

In other words, only the right continuity is needed (this allows jumps). Apparently, a continuous stochastic process implies a rcll stochastic process (but the reverse is not true). What we encounter in finance literatures are all rcll stochastic processes (for the modeling of stock price dynamics. Rcll processes include jump discontinuous process such as Poisson processes and infinite activity Lévy processes. Essentially discontinuous processes are useless in finance because they don’t have either (or both) of the left limit X_{t-} or the right limit X_{t+}.

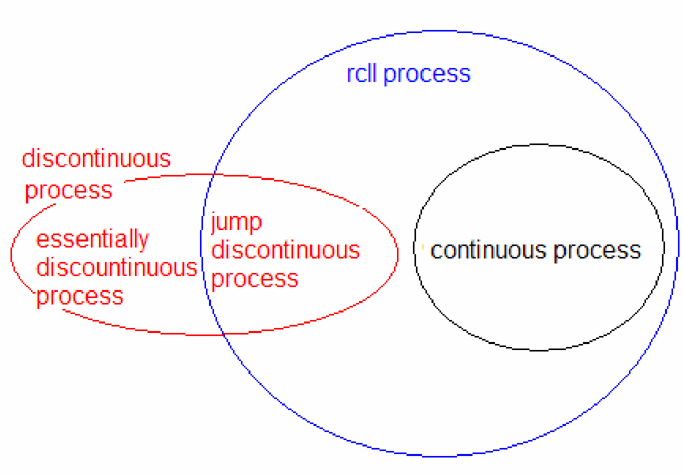
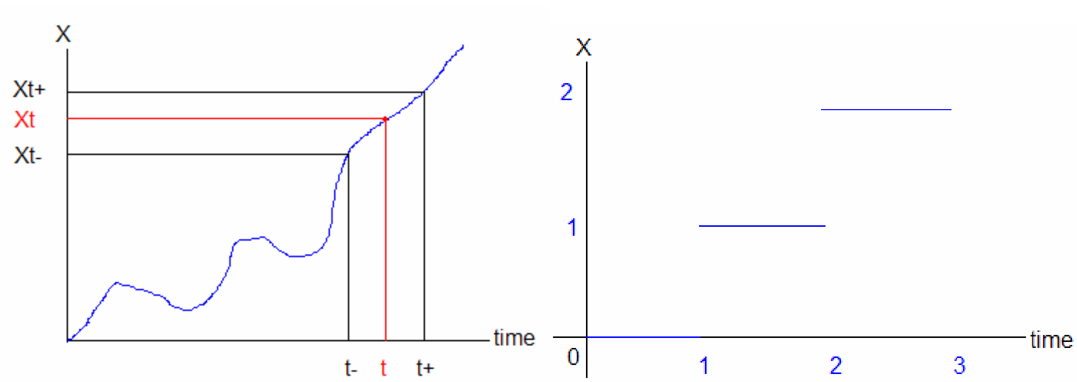


Figure 2.12: Relationship between rcll, continuous, and jump discontinuous processes.



A) A continuous stochastic process. B) A jump discontinuous stochastic process.

Figure 2.13: Examples of rcll stochastic processes.

Suppose t is a discontinuity point. The jump of the stochastic process at t is defined as:

$$\Delta X_t = X_t - X_{t-} .$$

A rcll nonanticipating stochastic process $(X_{t\in[0,T]})$ can have a finite number of large jumps and countable number (possibly infinite) of small jumps.

[2.4.3] Total Variation

Definition 2.37 Total variation of a function Let $f(x)$ be a bounded function defined in the interval $[a,b]$:

$$f(x):[a,b] \to \mathbb{R} .$$

The interval can be infinite, i.e. $[-\infty,\infty]$. Consider partitioning the interval $[a,b]$ with the points:

$$a = x_0 < x_1 < x_2 ... x_{n-1} < x_n = b .$$

Then, the total variation of a function $f(x)$ is defined by:

$$T(f) = \sup \sum_{i=1}^{n} \left| f(x_i) - f(x_{i-1}) \right| ,$$

where sup indicates a supremum (least upper bound).

Definition 2.38 Function of finite variation A function $f(x)$ on the interval $[a,b]$ is said to be a function of finite variation, if its total variation on the interval $[a,b]$ is finite:

$$T(f) = \sup \sum_{i=1}^{n} \left| f(x_i) - f(x_{i-1}) \right| < \infty .$$

Proposition 2.3 Every bounded increasing or decreasing function is of finite variation on the interval $[a,b]$.

Proof

Consider an increasing function $f(x)$ on the interval $[a,b]$. By its definition, for $\forall i$:

$$f(x_i) - f(x_{i-1}) \geq 0 .$$

And:

$$\begin{aligned} T(f) &= \sup \left\{ f(x_n) - f(x_{n-1}) + f(x_{n-1}) - f(x_{n-2}) + ... + f(x_2) - f(x_1) + f(x_1) - f(x_0) \right\} \\ T(f) &= \sup \left\{ f(x_n) - f(x_0) \right\} \\ T(f) &= \sup \left\{ f(b) - f(a) \right\} , \end{aligned}$$

which is finite because $f(x)$ is bounded:

$$-\infty < f(a), f(b) < \infty .$$

□

Definition 2.39 Total variation of a stochastic process Consider a real valued stochastic process $(X_{t \in [0,T]})$ on a filtered probability space $(\Omega, \mathcal{F}_{t \in [0,T]}, \mathbb{P})$. Partition the time interval $[0,T]$ with the points:

$$0 = t_0 < t_1 < t_2 ... < t_{n-1} < t_n = T .$$

Then, the total variation of a stochastic process $(X_{t\in[0,T]})$ on the time interval $[0,T]$ is defined by:

$$T(X) = \sup \sum_{i=1}^{n} \left| X(t_i) - X(t_{i-1}) \right|,$$

where sup indicates a supremum (least upper bound).

Definition 2.40 Stochastic process of finite variation A real valued stochastic process $(X_{t\in[0,T]})$ on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,T]}, \mathbb{P})$ on the interval $[0,T]$ is said to be a stochastic process of finite variation, if the total variation on the interval $[0,T]$ of a sample path of the process is finite with probability 1:

$$\mathbb{P}(T(X) = \sup \sum_{i=1}^{n} \left| X(t_i) - X(t_{i-1}) \right| < \infty) = 1.$$

[3] Lévy Processes

In this chapter some theorems and propositions are presented without proofs. This is obviously because it is not the goal of this sequel to prove theorems and some proofs are beyond what we need while consuming too many pages. But for those inquisitive readers, we provide the information about where to look for more details of the subjects and their proofs. Our goal is to present the foundations of the mathematics of Lévy processes as simple as possible.

[3.1] Definition of Lévy Processes

Definition 3.1 Lévy processes A real valued stochastic process $(X_{t\in[0,\infty)})$ on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,\infty)}, \mathbb{P})$ is said to be a Lévy process on $\mathbb{R}$ if it satisfies the following conditions:

(1) Its increments are independent. In other words, for $0 \le t_1 < t_2 < ... < t_n < \infty$:

$$\begin{aligned}&\mathbb{P}(X_{t_0} \cap X_{t_1} - X_{t_0} \cap X_{t_2} - X_{t_1} \cap ... \cap X_{t_n} - X_{t_{n-1}})\\&= \mathbb{P}(X_{t_0})\mathbb{P}(X_{t_1} - X_{t_0})\mathbb{P}(X_{t_2} - X_{t_1})...\mathbb{P}(X_{t_n} - X_{t_{n-1}}).\end{aligned}$$

(2) Its increments are stationary (time homogeneous): i.e. for $h \ge 0$, $X_{t+h} - X_t$ has the same distribution as X_h. In other words, the distribution of increments does not depend on t.

(3) $\mathbb{P}(X_0 = 0) = 1$. The process starts from 0 almost surely (with probability 1).

(4) The process is stochastically continuous: $\forall \varepsilon > 0$, $\lim_{h\to 0} \mathbb{P}(\left|X_{t+h} - X_t\right| \ge \varepsilon) = 0$.

(5) Its sample path (trajectory) is rcll almost surely.

Next, we describe what each condition implies. Consider an increasing sequence of time $0 < t_1 < t_2 < ... < t_n < t < u < \infty$ where t is the present. As a result of independent increments condition:

$$\begin{aligned}&\mathbb{P}(X_u - X_t \left| X_{t_1} - X_0, X_{t_2} - X_{t_1}, ..., X_t - X_{t_n}\right.)\\&= \frac{\mathbb{P}(X_u - X_t \cap X_{t_1} - X_0, X_{t_2} - X_{t_1}, ..., X_t - X_{t_n})}{\mathbb{P}(X_{t_1} - X_0, X_{t_2} - X_{t_1}, ..., X_t - X_{t_n})}\\&= \frac{\mathbb{P}(X_u - X_t)\mathbb{P}(X_{t_1} - X_0, X_{t_2} - X_{t_1}, ..., X_t - X_{t_n})}{\mathbb{P}(X_{t_1} - X_0, X_{t_2} - X_{t_1}, ..., X_t - X_{t_n})}\\&= \mathbb{P}(X_u - X_t),\end{aligned}$$

which means that there is no correlation (probabilistic dependence structure) on the increments among the past, the present, and the future.

For example, independent increments condition means that when modeling a log stock price $\ln S_t$ as an independent increment process, the probability distribution of a log stock price in year 2005 – 2006 is independent of the way the log stock price increment has evolved over the years (i.e. stock price dynamics), i.e. it doesn't matter if this stock crushes or soars in year 2004 – 2005):

$$\mathbb{P}(\ln S_{2006} - \ln S_{2005} \,|\, ..., \ln S_{2003} - \ln S_{2002}, \ln S_{2004} - \ln S_{2003}, \ln S_{2005} - \ln S_{2004}) \\ = \mathbb{P}(\ln S_{2006} - \ln S_{2005}).$$

Using the simple relationship $X_u \equiv (X_u - X_t) + X_t$ for an increasing sequence of time $0 < t_1 < t_2 < ... < t_n < t < u < \infty$:

$$\mathbb{P}(X_u \,|\, X_0, X_{t_1}, X_{t_2}, ..., X_{t_n}, X_t) = \mathbb{P}((X_u - X_t) + X_t \,|\, X_0, X_{t_1}, X_{t_2}, ..., X_{t_n}, X_t) \\ = \mathbb{P}(X_u \,|\, X_t),$$

which holds because an increment $(X_u - X_t)$ is independent of X_t by definition and the value of X_t depends on its realization $X_t(\omega)$. This is a strong probabilistic structure imposed on a stochastic process because this means that the conditional probability of the future value X_u depends only on the previous realization $X_t(\omega)$ and not on the entire past history of realizations $X_0, X_{t_1}, X_{t_2}, ..., X_{t_n}, X_t$ (i.e. called Markov property which is discussed soon).

Although this condition seems too strong, it imposes a very tractable property on the process. Because if two variables X and Y are independent:

$$E[XY] = E[X]E[Y], \\ Var[X+Y] = Var[X] + Var[Y], \\ Cov[X,Y] = 0 \text{ (i.e. } Corr[X,Y] = 0 \text{)}.$$

Stationary increments condition (2) means that the distributions of increments $X_{t+h} - X_t$ do not depend on the time t, but they depend on the time distance h of two observations (i.e. interval of time). In other words, the probability density function of increments does not change over time. For example, if you model a log stock price $\ln S_t$ as a process with stationary increments, the distribution of a log stock price increment in 2005-2006 is the same as that in 2050-2051:

$$\ln S_{2006} - \ln S_{2005} \overset{d}{=} \ln S_{2051} - \ln S_{2050}.$$

Processes satisfying the conditions (1) and (2) are called processes with independent and stationary increments. Independent increments condition is a restriction on the probabilistic dependence structure of increments among the past, present, and future. Stationary increments condition is a restriction on the shape of the distribution of increments among the past, present, and future.

The condition (4) (which is implied by the conditions (2), (3), and (5)) does not imply the continuous sample paths of the process. It means that if we are at time t, the probability of a jump at time t is zero because there is no uncertainty about the present. Jumps occur at random times. This property is called stochastic continuity or continuity in probability which we saw in section 2.4.

Rcll condition (5) does not need to be imposed. This is because a real valued Lévy process in law which is a process satisfying conditions (1), (2), (3), and (4) is modified to a Lévy process which satisfies the conditions (1), (2), (3), (4) , and (5) (theorem 11.5 of Sato (1999)). In other words, the condition (5) results from the conditions (1), (2), (3), and (4) through a theorem.

Definition 3.2 Right continuous with left limit (rcll) stochastic process A real valued nonanticipating stochastic process $(X_{t\in[0,T]})$ on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,T]}, \mathbb{P})$ is said to be a rcll stochastic process if for $\forall t \in [0, T]$:

(1) Right limit of the process as s approaches t from the above (right hand side) exists, i.e. $\lim_{s\to t, s>t} X_s = X_{t+}$. Left limit of the process as s approaches t from the below (left hand side) exists, i.e. $\lim_{s\to t, s<t} X_s = X_{t-}$.

(2) $X_{t+} = X_t$.

As you can see, the fact that left continuity is not needed allows the process to have jumps. A continuous stochastic process implies a rcll stochastic process but the reverse is not true. All stochastic processes used in finance literatures for the modeling of asset price dynamics are rcll stochastic processes. Rcll processes include jump discontinuous process such as Poisson processes and infinite activity Lévy processes. Essentially discontinuous processes are useless in finance because they don't have either (or both) of the left limit X_{t-} or the right limit X_{t+}.

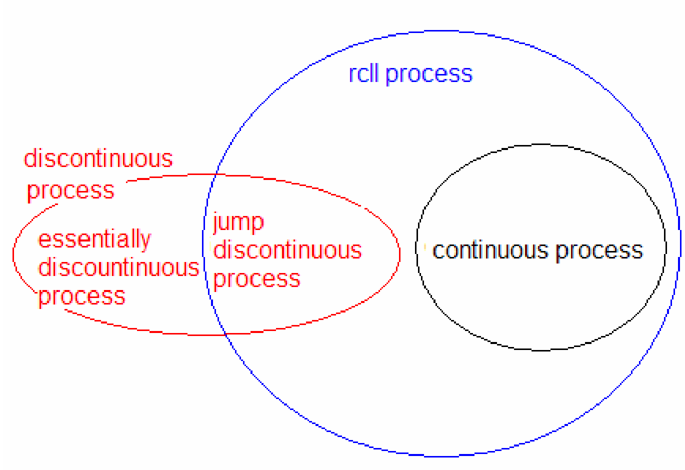
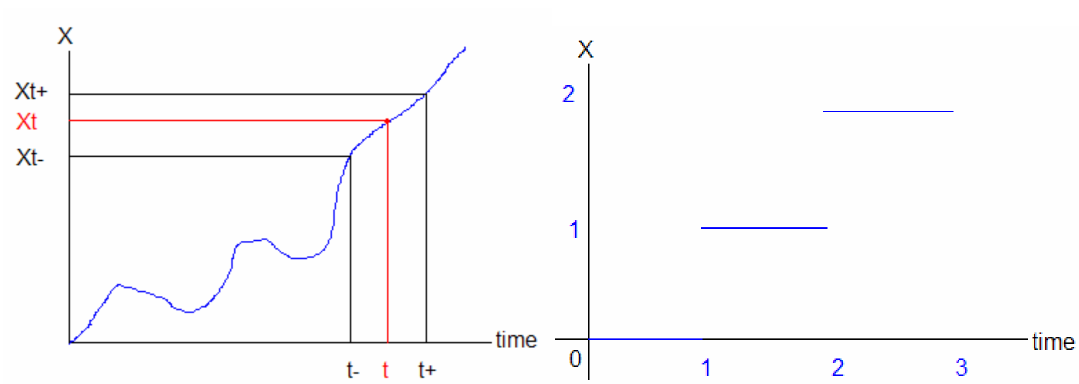


Figure 3.1: Relationship between rcll, continuous, and jump discontinuous processes.



A) A continuous stochastic process. B) A jump discontinuous stochastic process.

Figure 3.2: Examples of rcll stochastic processes.

Suppose t is a discontinuity point. The jump of the stochastic process at t is defined as:

$$\Delta X_t = X_t - X_{t-}.$$

A rcll nonanticipating stochastic process $(X_{t\in[0,T]})$ can have a finite number of large jumps and countable number (possibly infinite) of small jumps.

We saw the definition of a Lévy process. Next we discuss infinite divisibility of a distribution. It turns out that we cannot separate Lévy processes from infinitely divisible distributions because Lévy processes are generated by infinitely divisible distributions.

[3.2] Infinitely Divisible Random Variable and Distribution

Definition 3.3 Infinitely divisible random variable and distribution A real valued random variable X with the probability density function $\mathbb{P}(x)$ is said to be infinitely divisible if for $\forall n \in \mathbb{N}$ there exist $i.i.d.$ random variables $X_1, X_2, ..., X_n$ satisfying:

$$X \underset{=}{d} X_1 + X_2 + ... + X_n .$$

$\mathbb{P}(x)$ is said to be an infinitely divisible distribution.

Definition 3.4 Characteristic function Let X be a random variable with its probability density function $\mathbb{P}(x)$. A characteristic function $\phi(\omega)$ with $\omega \in \mathbb{R}$ is defined as the Fourier transform of the probability density function $\mathbb{P}(x)$ using Fourier transform parameters $(a,b) = (1,1)$:

$$\phi(\omega) \equiv \mathcal{F}[\mathbb{P}(x)] \equiv \int_{-\infty}^{\infty} e^{i\omega x} \mathbb{P}(x) dx \equiv E[e^{i\omega x}] . \tag{3.1}$$

In terms of a characteristic function, infinite divisibility is defined as follows.

Proposition 3.1 Infinitely divisible random variable and distribution A real valued random variable X with the probability density function $\mathbb{P}(x)$ and the characteristic function $\phi_X(\omega)$ is said to be infinitely divisible if for $\forall n \in \mathbb{N}$ there exist *i.i.d.* random variables $X_1, X_2, ..., X_n$ each with a characteristic function $\phi_{X_i}(\omega)$ such that:

$$\phi_X(\omega) = (\phi_{X_i}(\omega))^n \text{ or } (\phi_X(\omega))^{1/n} = \phi_{X_i}(\omega) .$$

$\mathbb{P}(x)$ is said to be an infinitely divisible distribution.

Proof

Consult Applebaum (2004) section 1.2.2.

Examples of infinitely divisible distributions include normal distributions on $\mathbb{R}$, gamma distributions on $\mathbb{R}$, α-stable distributions on $\mathbb{R}$, Poisson distributions on $\mathbb{R}$, compound Poisson distributions on $\mathbb{R}$, geometric distributions on $\mathbb{R}$, negative binomial distributions on $\mathbb{R}$, exponential distributions on $\mathbb{R}$. From Sato (1999), probability measures with bounded supports (e.g. uniform and binomial distributions) are not infinitely divisible.

Suppose that a random variable Y is drawn from a normal distribution with the mean μ and the variance σ^2:

$$\mathbb{P}(y) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{ -\frac{1}{2} \frac{(y-\mu)^2}{\sigma^2} \right\} .$$

Consider $i.i.d.$ normal random variables $Y_1, Y_2, ..., Y_n$ with the mean μ / n and the variance σ^2 / n. We learn in any undergraduate statistics courses that the sum of normal distributions is also normal with the linear mean and the linear variance:

$$E[Y_1 + Y_2 + ... + Y_n] = E[n(\mu / n)] = \mu \ ,$$
$$Var[Y_1 + Y_2 + ... + Y_n] = Var[Y_1] + Var[Y_2] + ... + Var[Y_n] = n(\sigma^2 / n) = \sigma^2 \ .$$

Therefore, a normal distribution is an infinitely divisible distribution since it satisfies:

$$Y \underline{\underline{d}} \ Y_1 + Y_2 + ... + Y_n \ .$$

This can be shown more formally using characteristic functions. The characteristic function of a normal random variable $Y \sim N(\mu, \sigma^2)$ is:

$$\phi_Y(\omega) = \mathcal{F}_Y[\mathbb{P}(y)](\omega) = \int_{-\infty}^{\infty} e^{i\omega y} \mathbb{P}(y) dy = \exp(i\mu\omega - \frac{\sigma^2\omega^2}{2}) \ .$$

The characteristic function for the $i.i.d.$ n summands of Y, Y_i, can be computed as:

$$\phi_{Y_i}(\omega) = \mathcal{F}_{Y_i}[\mathbb{P}(y_i)](\omega) = \int_{-\infty}^{\infty} e^{i\omega y_i} \mathbb{P}(y_i) dy$$
$$\phi_{Y_i}(\omega) = \int_{-\infty}^{\infty} e^{i\omega y_i} \frac{1}{\sqrt{2\pi\sigma^2 / n}} \exp\left\{ -\frac{1}{2} \frac{(y_i - \mu / n)^2}{\sigma^2 / n} \right\} dy$$
$$\phi_{Y_i}(\omega) = \exp\left\{ i(\mu / n)\omega - \frac{(\sigma^2 / n)\omega^2}{2} \right\} .$$

But, we can see the obvious relationship between $\phi_Y(\omega)$ and $\phi_{Y_i}(\omega)$:

$$\phi_{Y_i}(\omega) = (\phi_Y(\omega))^{1/n} = \left\{ \exp(i\mu\omega - \frac{\sigma^2\omega^2}{2}) \right\}^{1/n} = \exp\left\{ i(\mu / n)\omega - \frac{(\sigma^2 / n)\omega^2}{2} \right\} .$$

This proves that the $i.i.d.$ n summands of $Y \sim N(\mu, \sigma^2)$ are also normally distributed with the mean μ / n and the variance σ^2 / n (because a characteristic function uniquely determines a probability distribution) and therefore proving the infinite divisibility of a normal distribution:

$$Y \underline{\underline{d}} \ \sum_{k=0}^{n-1} Y_k \ , \ Y_k \sim i.i.d. \ N(\mu / n, \sigma^2 / n) \ .$$

Another example is a Poisson case. Suppose that Z is a Poisson random variable, i.e. $Z \sim Poisson(\lambda)$ with $\lambda \in \mathbb{R}^+$. Its characteristic function is:

$$\phi_Z(\omega) = \mathcal{F}_Z[\mathbb{P}(z)](\omega) = \sum_{z=0}^{\infty} \frac{e^{-\lambda}\lambda^z}{z!} e^{i\omega z} = \exp[\lambda(e^{i\omega} - 1)].$$

Many readers learned that the sum of Poisson random variables is also a Poisson random variable. Consider $i.i.d.$ Poisson random variables $Z_1, Z_2, ..., Z_n$ with the intensity λ / n. Its characteristic function is:

$$\phi_{Z_i}(\omega) = \mathcal{F}_{Z_i}[\mathbb{P}(z_i)](\omega) = \sum_{z_i=0}^{\infty} \frac{e^{-\lambda/n} \lambda / n^{z_i}}{z_i!} e^{i\omega z_i} = \exp[\frac{\lambda}{n}(e^{i\omega} - 1)].$$

But, we can see the obvious relationship between $\phi_Z(\omega)$ and $\phi_{Z_i}(\omega)$:

$$\phi_{Z_i}(\omega) = (\phi_Z(\omega))^{1/n} = [\exp\{\lambda(e^{i\omega} - 1)\}]^{1/n} = \exp\{\frac{\lambda}{n}(e^{i\omega} - 1)\}.$$

This proves that the $i.i.d.$ n summands of $Z \sim Poisson(\lambda)$ are also Poisson distributed with the intensity λ / n (because a characteristic function uniquely determines a probability distribution) and therefore proving the infinite divisibility of a Poisson distribution:

$$Z \underset{=}{d} \sum_{k=0}^{n-1} Z_k, Z_k \sim i.i.d. Poisson(\lambda / n).$$

[3.3] Relationship between Lévy Processes and Infinitely Divisible Distributions

Proposition 3.2 If $(X_{t\in[0,\infty)})$ is a real valued Lévy process on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,\infty)}, \mathbb{P})$, then, X_t has an infinitely divisible distribution for $\forall t \in [0, T]$. This corresponds to the corollary 11.6 of Sato (1999).

Proof

Consider a realization X_t at time t. Partition the time t into $n \in \mathbb{N}$ intervals using $t_i = i(t / n)$:

$$t_0 = 0, t_1 = t / n, t_2 = 2t / n, ..., t_n = nt / n = t.$$

A realization X_t can be considered as a sum of $n \in \mathbb{N}$ increments:

$$X_t = (X_{t_1} - X_{t_0}) + (X_{t_2} - X_{t_1}) + ... + (X_{t_n} - X_{t_{n-1}}).$$

Note that all these increments are $i.i.d.$ increments because $(X_{t\in[0,\infty)})$ is a real valued Lévy process. Let $\phi(\omega; X(t))$ be the characteristic function of X_t and $\phi(\omega; X(t_i - t_{i-1}))$ be the characteristic function of $i.i.d.$ increments. Then, we have:

$$\phi(\omega; X(t)) = \phi(\omega; X(t_i - t_{i-1}))^n , \tag{3.2}$$

which follows from the property of a characteristic function: if $\{X_k, k = 1,...,n\}$ are independent random variables, the characteristic function of their sum $X_1 + X_2 + ... + X_n$ is the product of their characteristic functions:

$$\phi_{X_1+X_2+...+X_n}(\omega) = \prod_{k=1}^{n} \phi_{X_k}(\omega) .$$

The equation (3.2) satisfies the proposition 3.1. Therefore, the distribution of Lévy process possesses infinite divisibility.

□

Thus, for every $t \in [0,\infty)$ and $h \in \mathbb{R}^+$, increments of a Lévy process $X_{t+h} - X_t$ follows an infinitely divisible law. Its converse is also true.

Proposition 3.3 For every infinitely divisible distribution $\mathbb{P}$ on $\mathbb{R}$, there exists a Lévy process $(X_{t\in[0,\infty)})$ on $\mathbb{R}$ whose distribution of increments $X_{t+h} - X_t$ is governed by $\mathbb{P}$. This is the corollary 11.6 of Sato (1999).

This proposition is extremely important because it means that there is one-to-one correspondence between an infinitely divisible distribution and a Lévy process. Some typical examples are illustrated in Table 3.1.

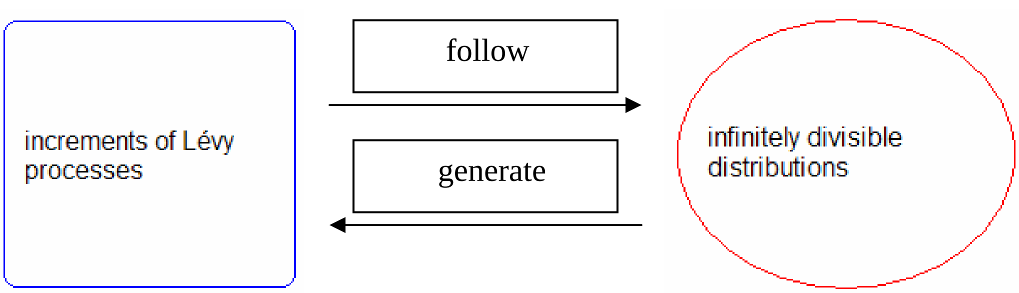


Figure 3.3: Relationship between Lévy Processes and Infinitely Divisible Distributions

Table 3.1: One-to-one correspondence between an infinitely divisible distribution and a Lévy process

Infinitely divisible probability measure $\mathbb{P}$	$\rightleftarrows$	Lévy process

Normal distribution	Brownian motion (with drift)
Poisson distribution	Poisson process
Compound Poisson distribution	Compound Poisson process
Cauchy distribution	Cauchy process
Exponential distribution	Gamma process

[3.4] Lévy-Khinchin Representation

Lévy-Khinchin representation gives the characteristic functions of all infinitely divisible distributions. In other words, it gives the characteristic functions of all processes whose increments follow infinitely divisible distributions – Lévy processes.

Theorem 3.1 General Lévy-Khinchin representation (formula) of all infinitely divisible distributions Let $\mathbb{P}(x)$ be a real valued infinitely divisible distribution. Then, its characteristic function $\phi_X(\omega)$ is given by for $\forall \omega \in \mathbb{R}$:

$$\phi_X(\omega) = \exp\left(\psi_X(\omega)\right),$$

where $\psi_X(\omega)$ called a characteristic exponent (or a log characteristic function) is given by:

$$\psi_X(\omega) = -\frac{A\omega^2}{2} + i\gamma\omega + \int_{-\infty}^{\infty} \left\{ \exp(i\omega x) - 1 - i\omega x 1_D \right\} \ell(dx), \tag{3.3}$$

where $D = \{x : |x| \le 1\}$, A is a unique nonnegative constant (i.e. $A \in \mathbb{R}^+$), γ is a unique constant on $\mathbb{R}$, and ℓ is a unique measure on $\mathbb{R}$ satisfying:

$$\ell(\{0\}) = 0 \text{ and } \int_{-\infty}^{\infty} \min\{|x|^2, 1\} \ell(dx) < \infty .$$

Proof

Consult Sato (1999) p.37-p.47.

Theorem 3.2 Converse of Theorem 3.1 Consider a characteristic function $\phi_X(\omega)$ of a probability distribution $\mathbb{P}(x)$. If there exist a unique nonnegative constant A, a unique real valued constant γ, and a unique real valued measure ℓ satisfying $\ell(\{0\}) = 0$ and $\int_{-\infty}^{\infty} \min\{|x|^2, 1\}\ell(dx) < \infty$ which yield the characteristic function of the form:

$$\phi_X(\omega) = \exp\left(\psi_X(\omega)\right),$$

where:

$$\psi_X(\omega) = -\frac{A\omega^2}{2} + i\gamma\omega + \int_{-\infty}^{\infty} \left\{\exp(i\omega x) - 1 - i\omega x 1_D\right\} \ell(dx),$$

then, $\mathbb{P}(x)$ is infinitely divisible.

Proof

Consult Sato (1999) p.37-p.47.

Theorem 3.3 General Lévy-Khinchin representation (formula) of all Lévy processes (processes whose increments follow infinitely divisible distributions) Let $(X_{t\in[0,\infty)})$ be a real valued Lévy process defined on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,\infty)}, \mathbb{P})$. Then, for any $\omega \in \mathbb{R}$, the characteristic function $\phi_X(\omega)$ of a Lévy process $(X_{t\in[0,\infty)})$ can be expressed as:

$$\phi_X(\omega) = \exp\left(t\psi_X(\omega)\right),$$

where $\psi_X(\omega)$ called a characteristic exponent (or a log characteristic function) is given by:

$$\psi_X(\omega) = -\frac{A\omega^2}{2} + i\gamma\omega + \int_{-\infty}^{\infty} \left\{\exp(i\omega x) - 1 - i\omega x 1_D\right\} \ell(dx), \tag{3.4}$$

where $D = \{x : |x| \leq 1\}$, A is a unique nonnegative constant (i.e. $A \in \mathbb{R}^+$), γ is a unique constant on $\mathbb{R}$, and ℓ is a unique measure on $\mathbb{R}$ satisfying:

$$\ell(\{0\}) = 0 \text{ and } \int_{-\infty}^{\infty} \min\{|x|^2, 1\}\ell(dx) < \infty.$$

Proof

Consult Sato (1999) p.37-p.47.

Theorem 3.4 Converse of Theorem 3.3 Consider a characteristic function $\phi_X(\omega)$ of a random variable X_t of a real valued stochastic process $(X_{t\in[0,\infty)})$ defined on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,\infty)}, \mathbb{P})$. If there exist a unique nonnegative constant A, a unique real valued constant γ, and a unique real valued measure ℓ satisfying $\ell(\{0\}) = 0$ and $\int_{-\infty}^{\infty} \min\{|x|^2, 1\}\ell(dx) < \infty$ which yield the characteristic function of the form:

$$\phi_X(\omega) = \exp\left(t\psi_X(\omega)\right),$$

where:

$$\psi_X(\omega) = -\frac{A\omega^2}{2} + i\gamma\omega + \int_{-\infty}^{\infty} \left\{\exp(i\omega x) - 1 - i\omega x 1_D\right\} \ell(dx),$$

then, $(X_{t\in[0,\infty)})$ is a real valued Lévy process (a process whose increments follow an infinitely divisible distribution).

Proof

Consult Sato (1999) p.37-p.47.

Definition 3.5 Lévy triplet (generating triplet) In theorem 3.3, a unique nonnegative constant A is called a Gaussian variance and a unique real valued measure ℓ satisfying $\ell(\{0\}) = 0$ and $\int_{-\infty}^{\infty} \min\{|x|^2, 1\}\ell(dx) < \infty$ is called a Lévy measure. A unique real valued constant γ does not have any intrinsic meaning since it depends on the behavior of a Lévy measure ℓ. It turns out that these triplets uniquely define a Lévy process as a result of Lévy-Itô decomposition. This triplet is called a Lévy triplet and compactly written as (A, ℓ, γ).

Next, we present special cases of the general Lévy-Khinchin representation of all Lévy processes by restricting the behavior of the Lévy measure ℓ.

Theorem 3.5 Lévy-Khinchin representation of Lévy processes whose Lévy measure satisfies the additional condition $\int_{|x|\leq 1} |x| \ell(dx) < \infty$ Let $(X_{t\in[0,\infty)})$ be a real valued Lévy process defined on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,\infty)}, \mathbb{P})$ whose Lévy measure satisfies the additional condition $\int_{|x|\leq 1} |x| \ell(dx) < \infty$. Then, for any $\omega \in \mathbb{R}$, the characteristic function $\phi_X(\omega)$ of a Lévy process $(X_{t\in[0,\infty)})$ can be expressed as:

$$\phi_X(\omega) = \exp\left(t\psi_X(\omega)\right),$$

where $\psi_X(\omega)$ called a characteristic exponent (or a log characteristic function) is given by from the equation (3.4):

$$\psi_X(\omega) = -\frac{A\omega^2}{2} + i\gamma\omega + \int_{-\infty}^{\infty} \left\{ \exp(i\omega x) - 1 - i\omega x 1_{|x|\leq 1} \right\} \ell(dx)$$
$$\psi_X(\omega) = -\frac{A\omega^2}{2} + i\left(\gamma - \int_{|x|\leq 1} |x| \ell(dx)\right)\omega + \int_{-\infty}^{\infty} \left\{ \exp(i\omega x) - 1 \right\} \ell(dx)$$
$$\psi_X(\omega) = -\frac{A\omega^2}{2} + i\gamma_0\omega + \int_{-\infty}^{\infty} \left\{ \exp(i\omega x) - 1 \right\} \ell(dx), \tag{3.5}$$

where $\gamma_0 \equiv \gamma - \int_{|x|\leq 1} |x| \ell(dx)$ is a unique real valued constant called a drift, A is a unique nonnegative constant (i.e. $A \in \mathbb{R}^+$) called a Gaussian variance, and ℓ is a unique measure on $\mathbb{R}$ called a Lévy measure satisfying:

$$\ell(\{0\}) = 0 \ , \ \int_{-\infty}^{\infty} \min\{|x|^2, 1\} \ell(dx) < \infty \ , \ \int_{|x|\leq 1} |x| \ell(dx) < \infty \ .$$

Theorem 3.6 Lévy-Khinchin representation of Lévy processes whose Lévy measure satisfies the additional condition $\int_{|x|>1} |x| \ell(dx) < \infty$ Let $(X_{t\in[0,\infty)})$ be a real valued Lévy process defined on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,\infty)}, \mathbb{P})$ whose Lévy measure satisfies the additional condition $\int_{|x|>1} |x| \ell(dx) < \infty$. Then, for any $\omega \in \mathbb{R}$, the characteristic function $\phi_X(\omega)$ of a Lévy process $(X_{t\in[0,\infty)})$ can be expressed as:

$$\phi_X(\omega) = \exp\left(t\psi_X(\omega)\right),$$

where $\psi_X(\omega)$ called a characteristic exponent (or a log characteristic function) is given by from the equation (3.4):

$$\psi_X(\omega) = -\frac{A\omega^2}{2} + i\gamma\omega + \int_{-\infty}^{\infty} \left\{ \exp(i\omega x) - 1 - i\omega x 1_{|x|\leq 1} \right\} \ell(dx)$$
$$\psi_X(\omega) = -\frac{A\omega^2}{2} + i\gamma_1\omega + \int_{-\infty}^{\infty} \left\{ \exp(i\omega x) - 1 - i\omega x \right\} \ell(dx), \tag{3.6}$$

where γ_1 is a unique real valued constant called a center (identical to the mean), A is a unique nonnegative constant (i.e. $A \in \mathbb{R}^+$) called a Gaussian variance, and ℓ is a unique measure on $\mathbb{R}$ called a Lévy measure satisfying:

$$\ell(\{0\}) = 0 \ , \ \int_{-\infty}^{\infty} \min\{|x|^2, 1\} \ell(dx) < \infty \ , \ \int_{|x|>1} |x| \ell(dx) < \infty \ .$$

Consult Cont and Tankov (2004) page 83 for more details.

Next, we consider a Lévy-Khinchin representation for a subclass of Lévy processes called a finite variation Lévy process. For a detailed description of the concept of variation, read section 2.4.3.

Definition 3.6 Total variation of a stochastic process Consider a real valued stochastic process $(X_{t\in[0,T]})$ on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,T]}, \mathbb{P})$. Partition the time interval $[0,T]$ with the points:

$$0 = t_0 < t_1 < t_2 ... < t_{n-1} < t_n = T .$$

Then, the total variation of a stochastic process $(X_{t\in[0,T]})$ on the time interval $[0,T]$ is defined by:

$$T(X) = \sup \sum_{i=1}^{n} \left| X(t_i) - X(t_{i-1}) \right| ,$$

where sup indicates a supremum (least upper bound).

Definition 3.7 Stochastic process of finite variation A real valued stochastic process $(X_{t\in[0,T]})$ on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,T]}, \mathbb{P})$ on the interval $[0,T]$ is said to be a stochastic process of finite variation, if the total variation on the interval $[0,T]$ of a sample path of the process is finite with probability 1:

$$\mathbb{P}(T(X) = \sup \sum_{i=1}^{n} \left| X(t_i) - X(t_{i-1}) \right| < \infty) = 1 .$$

Definition 3.8 Lévy process of finite variation A real valued Lévy process $(X_{t\in[0,T]})$ on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,T]}, \mathbb{P})$ on the interval $[0,T]$ is said to be a Lévy process of finite variation, if the total variation on the interval $[0,T]$ of a sample path of the Lévy process is finite with probability 1:

$$\mathbb{P}(T(X) = \sup \sum_{i=1}^{n} \left| X(t_i) - X(t_{i-1}) \right| < \infty) = 1 .$$

Theorem 3.7 Lévy process of finite variation If $(X_{t\in[0,T]})$ is a real valued Lévy process of finite variation on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,T]}, \mathbb{P})$ on the interval $[0,T]$, then, its Lévy triplet (A, ℓ, b) satisfies:

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

This corresponds to the proposition 3.9 of Cont and Tankov (2004) where its proof is given.

Theorem 3.8 Lévy-Khinchin representation for Lévy processes of finite variation
Let $(X_{t\in[0,\infty)})$ be a real valued Lévy process of finite variation with Lévy triplet $(A=0,\ell,b)$ defined on a filtered probability space $(\Omega,\mathcal{F}_{t\in[0,\infty)},\mathbb{P})$. Then, for any $\omega\in\mathbb{R}$, its characteristic function $\phi_X(\omega)$ can be expressed as:

$$\phi_X(\omega)=\exp\left(t\psi_X(\omega)\right),$$

where $\psi_X(\omega)$ called a characteristic exponent (or a log characteristic function) is given by:

$$\psi_X(\omega)=ib\omega+\int_{-\infty}^{\infty}\left\{\exp(i\omega x)-1-i\omega x 1_{|x|\le 1}\right\}\ell(dx)$$
$$\psi_X(\omega)=i\left(b-\int_{|x|\le 1}x\ell(dx)\right)\omega+\int_{-\infty}^{\infty}\left\{\exp(i\omega x)-1\right\}\ell(dx), \tag{3.7}$$

Theorem 3.9 Approximation of Lévy processes by compound Poisson processes
Every infinitely divisible distribution can be obtained as the weak limit of a sequence of compound Poisson distributions. This means that every Lévy process can be obtained as the weak limit of a sequence of compound Poisson random variables. In other words, every Lévy process can be approximated by a compound Poisson process. This theorem corresponds to corollary 8.8 of Sato (1999) where you can find its proof.

[3.5] Lévy-Itô Decomposition of Sample Paths of Lévy Processes

Historically speaking, Paul Lévy proposed the original idea of Lévy-Itô decomposition and as a result of Lévy-Itô decomposition, Lévy-Khinchin representation was proposed. In other words, Lévy-Khinchin representation uses the results of Lévy-Itô decomposition. Later, both Lévy-Itô decomposition and Lévy-Khinchin representation were formally proven by Kiyoshi Itô.

Lévy-Itô decomposition basically states that every sample path of Lévy process can be represented as a sum of two independent processes: One is a continuous Lévy process and the other is a compensated sum of independent jumps. Obviously, a continuous Lévy process is a Brownian motion with drift. One trick is that the jump component has to be a *compensated* sum of independent jumps because a sum of independent jumps at time t may not converge.

Theorem 3.10 General Lévy-Itô Decomposition of Sample Paths of Lévy Processes
Consider a real valued Lévy process $(X_{t\in[0,\infty)})$ with Lévy triplet (A,ℓ,γ) defined on a filtered probability space $(\Omega,\mathcal{F}_{t\in[0,\infty)},\mathbb{P})$. Then, Lévy-Itô decomposition states:

(1) Let $J_X \equiv J(A,\omega)$ be the random jump measure at time t of a Lévy process $(X_{t\in[0,\infty)})$ which contains the information of the timing of jumps and the size of jumps and the jump size belongs to a Borel set, i.e. $\Delta X_t(\omega) = X_t(\omega) - X_{t-}(\omega) \in \mathcal{B}_{\mathbb{R}}$. Then, J_X has a Poisson distribution with the mean (intensity) $\ell(dx)dt$.

(2) Every sample path of Lévy process can be represented as a sum of a continuous Lévy process and a discontinuous Lévy process:

$$X_t(\omega) = X_t^C(\omega) + X_t^D(\omega).$$

(3) The continuous part $X_t^C(\omega)$ has a Lévy triplet $(A, 0, \gamma)$. This means that the continuous part is a Brownian motion with drift:

$$X_t^C(\omega) = \gamma t + \sqrt{A} B_t = \gamma t + \sigma B_t,$$

where $Var[\gamma t + \sqrt{A} B_t] = Var[\gamma t + \sigma B_t] = At = \sigma^2 t$. This is why A is called a Gaussian variance.

(4) The discontinuous part $X_t^D(\omega)$ is a Lévy process with Lévy triplet $(0, \ell < \infty, 0)$ which is the definition of a compound Poisson process (we will discuss this soon). The discontinuous part $X_t^D(\omega)$ can be decomposed into a large jumps part $X_t^L(\omega)$ and a limit as $\varepsilon \to 0$ of compensated small jumps part $\tilde{X}_t^S(\omega)$:

$$X_t^D(\omega) = X_t^L(\omega) + \lim_{\varepsilon \downarrow 0} \tilde{X}_t^S(\omega).$$

We arbitrarily define large jumps as those with absolute size greater than 1 (i.e. this is completely arbitrary) and small jumps as those with absolute size between $\varepsilon \in \mathbb{R}^+$ and 1:

$$\Delta X_t^L(\omega) \equiv \left|\Delta X_t(\omega)\right| > 1 \text{ and } \Delta X_t^S(\omega) \equiv \varepsilon < \left|\Delta X_t(\omega)\right| < 1.$$

A large jumps part $X_t^L(\omega)$ is a sum of large jumps during the time interval $(0,t]$ and $X_t^L(\omega)$ is almost surely finite because Lévy processes have finite number of large jumps by the definition of the Lévy measure (read section 3.6):

$$\mathbb{P}\left(X_t^L(\omega) < \infty\right) = \mathbb{P}\left(\sum_{0 \le r \le t} \Delta X_r^L(\omega) < \infty\right) = 1.$$

This implies that there is no convergence issue with respect to a large jumps part $X_t^L(\omega)$.

Next, consider a *non-compensated* small jumps part $X_t^S(\omega)$ which is simply a sum of small jumps during the time interval $(0,t]$:

$$X_t^S(\omega) = \sum_{0 \le r \le t} \Delta X_r^S(\omega) = \int_{s\in(0,t],\varepsilon<\Delta|x|<1} xJ_X(ds \times dx, \omega),$$

which is a compound Poisson process. The reason that we cannot use a *non-compensated* small jumps part $X_t^S(\omega)$ in Lévy-Itô decomposition is that in the limit $\varepsilon \downarrow 0$, $X_t^S(\omega)$ almost surely does not achieve convergence (i.e. $\lim_{\varepsilon\downarrow 0} X_t^S(\omega) = \infty$) because Lévy processes can have infinitely many number of small jumps (i.e. Lévy measure ℓ can have a singularity at 0). Therefore, a *non-compensated* small jumps part $X_t^S(\omega)$ must be *compensated* by its mean $\ell(dx)ds$ (i.e. because $X_t^S(\omega)$ is Poisson process) as:

$$\tilde{X}_t^S(\omega) \equiv X_t^S(\omega) - \ell(\varepsilon < dx < 1)dt,$$

which makes $\tilde{X}_t^S(\omega)$ a martingale so that $\lim_{\varepsilon\downarrow 0} \tilde{X}_t^S(\omega)$ can now achieve convergence. In other words, a compensated small jumps part $\tilde{X}_t^S(\omega)$ is a *compensated* sum of small jumps during the time interval $(0,t]$:

$$\tilde{X}_t^S = \sum_{0 \le r \le t} \Delta \tilde{X}_r^S = \int_{s\in(0,t],\varepsilon<\Delta|x|<1} \{xJ_X(\omega) - x\ell(dx)ds\},$$

which always converges.

(5) Two processes $X_t^C(\omega)$ (i.e. a Brownian motion with drift) and $X_t^D(\omega)$ (i.e. a compound Poisson process) are independent.

Proof

Consult Sato (1999) section 19.

[3.6] Lévy Measure

Definition 3.9 Lévy measure of all Lévy processes Let $(X_{t\in[0,\infty)})$ be a real valued Lévy process with Lévy triplet (A, ℓ, γ) defined on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,\infty)}, \mathbb{P})$. The Lévy measure ℓ of a Lévy process $(X_{t\in[0,\infty)})$ is defined as a unique positive measure on $\mathbb{R}$ which measures (counts) the expected (average) number of jumps per unit of time:

$$\ell(A) = E[\#\{t \in [0,1] : \Delta X_t = X_t - X_{t-} \neq 0, \Delta X_t \in A\}],$$

where $\Delta X_t \in A$ indicates that the jump size belongs to a set A and a set A is a member of Borel set. This definition is equivalent to stating that the Lévy measure ℓ of a Lévy process $(X_{t\in[0,\infty)})$ is defined as a unique positive measure on $\mathbb{R}$ which measures arrival rate of jumps per unit of time. We like to emphasize that a Lévy measure ℓ measures the expected number of jumps of all sizes (i.e. small jumps plus large jumps) per unit of time. This definition of Lévy measure ℓ is true regardless of the types of Lévy processes (i.e. it doesn't matter whether it is a finite activity Lévy processes or an infinite activity Lévy processes). For more details, consult Cont and Tankov (2004) section 3.3.

By definition, a Lévy measure ℓ satisfies the two conditions:

(1) $\ell(\{0\}) = 0$, i.e. the measure of an empty set is zero.
(2) $\int_{-\infty}^{\infty} \min\{|x|^2, 1\}\ell(dx) < \infty$, mathematicians write this as $\int_{-\infty}^{\infty} (|x|^2 \wedge 1)\ell(dx) < \infty$.

Condtion (2) is extremely important. When the jump size is greater than 1 (i.e. $|x| > 1$), the jump is defined as a large jump using the arbitrary truncation point of 1 (the choice of truncation point is of no importance), then, the condition (2) reduces to:

$$\int_{|x|>1} \ell(dx) < \infty,$$

which tells that the expected number of large jumps per unit of time is finite. When the jump size is less than 1 (i.e. $|x| < 1$), the jump is defined as a small jump, then, the condition (2) reduces to:

$$\int_{|x|<1} |x|^2 \ell(dx) < \infty,$$

which means that a Lévy measure must be square-integrable around the origin. This indicates the following definition of Lévy processes with respect to the number of small and large jumps.

Definition 3.10 Lévy processes Lévy processes have finite expected number of large jumps per unit of time. Lévy processes can have finite or infinite expected number of small jumps per unit of time.

Depending on the behavior of Lévy measures, Lévy processes can be classified into three types.

Definition 3.11 Lévy process with zero Lévy measure $\ell = 0$ Let $(X_{t\in[0,\infty)})$ be a real valued Lévy process with Lévy triplet (A, ℓ, γ) defined on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,\infty)}, \mathbb{P})$. If the Lévy measure ℓ of this Lévy process is $\ell = 0$, $(X_{t\in[0,\infty)})$ is a Brownian motion (with drift). Zero Lévy measure $\ell = 0$ means that the Lévy process

$(X_{t\in[0,\infty)})$ has no small or large jumps, in other words, sample paths of the process is continuous.

Definition 3.12 Finite Activity Lévy processes Let $(X_{t\in[0,\infty)})$ be a real valued Lévy process with Lévy triplet (A,ℓ,γ) defined on a filtered probability space $(\Omega,\mathcal{F}_{t\in[0,\infty)},\mathbb{P})$. A Lévy process $(X_{t\in[0,\infty)})$ is said to be a finite activity Lévy process if its Lévy measure ℓ has a finite integral:

$$\int_{-\infty}^{\infty}\ell(dx)<\infty .$$

It is important to emphasize that the above condition implies that a finite activity Lévy process has a finite expected number of small jumps and a finite expected number of large jumps per unit of time:

$$\int_{|x|<1}\ell(dx)<\infty \text{ and } \int_{|x|>1}\ell(dx)<\infty .$$

Definition 3.13 Infinite Activity Lévy processes Let $(X_{t\in[0,\infty)})$ be a real valued Lévy process with Lévy triplet (A,ℓ,γ) defined on a filtered probability space $(\Omega,\mathcal{F}_{t\in[0,\infty)},\mathbb{P})$. A Lévy process $(X_{t\in[0,\infty)})$ is said to be an infinite activity Lévy process if its Lévy measure ℓ has an infinite integral:

$$\int_{-\infty}^{\infty}\ell(dx)=\infty .$$

It is important to emphasize that the above condition implies that an infinite activity Lévy process has an infinite expected number of small jumps and a finite expected number of large jumps per unit of time:

$$\int_{|x|<1}\ell(dx)=\infty \text{ and } \int_{|x|>1}\ell(dx)<\infty .$$

[3.7] Classification of Lévy Processes

In this section, we present various ways to categorize Lévy processes.

[3.7.1] In Terms of Gaussian or Not

Definition 3.14 Gaussian Lévy Process A real valued Lévy process $(X_{t\in[0,T]})$ on a filtered probability space $(\Omega,\mathcal{F}_{t\in[0,T]},\mathbb{P})$ is said to be a Gaussian Lévy process if it satisfies one of the following three equivalent conditions:

(1) $(X_{t\in[0,T]})$ is a Brownian motion (with drift). In other words, for $\forall t \in [0,T)$ and $\forall h \in \mathbb{R}^+$, its increments $X_{t+h} - X_t$ follow a normal distribution.
(2) Its sample path is continuous.
(3) Its Lévy measure is zero, $\ell = 0$.

Definition 3.15 Continuous Lévy Process (same as Gaussian Lévy Process) The only continuous Lévy process is a Gaussian Lévy process which is a Brownian motion (with drift).

Definition 3.16 Non-Gaussian Lévy Processes (same as jump Lévy Processes) A real valued Lévy process $(X_{t\in[0,T]})$ on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,T]}, \mathbb{P})$ is said to be a non-Gaussian Lévy process if it satisfies one of the following three equivalent conditions:

(1) For $\forall t \in [0,T)$ and $\forall h \in \mathbb{R}^+$, its increments $X_{t+h} - X_t$ do not follow a normal distribution.
(2) Its sample path is discontinuous.
(3) Its Lévy measure is not zero, $\ell \neq 0$.

Definition 3.17 Purely non-Gaussian Lévy Processes (same as pure jump Lévy Processes) A real valued Lévy process $(X_{t\in[0,T]})$ on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,T]}, \mathbb{P})$ is said to be a purely non-Gaussian Lévy process if it satisfies either condition (1) or (2) (i.e. these conditions are equivalent):

(1) Its Gaussian variance A is zero, that is, its Lévy triplet is given by $(A = 0, \ell, b)$.
(2) The process $(X_{t\in[0,T]})$ does not contain a Brownian motion (with drift).

Consider a jump diffusion (JD) process which is a Brownian motion (with drift) plus a compound Poisson process. Firstly, a JD process obviously contains a Brownian motion (with drift), thus, it is not a purely non-Gaussian Lévy process. But the increments $X_{t+h} - X_t$ of a JD process do not follow a normal distribution (because of the addition of compound Poisson process), thus, it is not a Gaussian Lévy process. Therefore, a JD process is a non-Gaussian Lévy process. Variance process is a purely non-Gaussian Lévy process because it does not contain a Brownian motion (with drift).

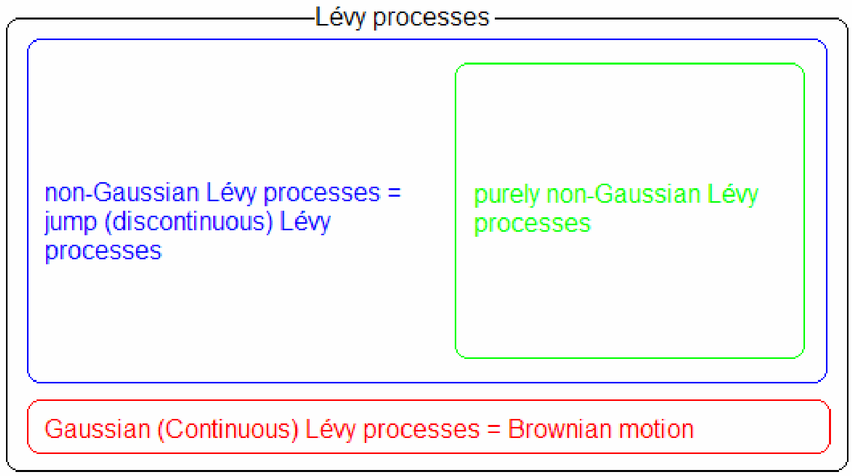


Figure 3.4 Classification of Lévy processes in terms of Gaussian or not

[3.7.2] In Terms of the Behavior of Lévy Measure ℓ

Read section 3.6 for the definition of Lévy measure ℓ.

Definition 3.18 Lévy process with zero Lévy measure $\ell = 0$ Let $(X_{t\in[0,\infty)})$ be a real valued Lévy process with Lévy triplet (A, ℓ, γ) defined on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,\infty)}, \mathbb{P})$. If the Lévy measure ℓ of this Lévy process is $\ell = 0$, $(X_{t\in[0,\infty)})$ is a Brownian motion (with drift). Zero Lévy measure $\ell = 0$ means that the Lévy process $(X_{t\in[0,\infty)})$ has no small or large jumps, in other words, sample paths of the process is continuous.

Definition 3.19 Finite Activity Lévy processes Let $(X_{t\in[0,\infty)})$ be a real valued Lévy process with Lévy triplet (A, ℓ, γ) defined on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,\infty)}, \mathbb{P})$. A Lévy process $(X_{t\in[0,\infty)})$ is said to be a finite activity Lévy process if its Lévy measure ℓ has a finite integral:

$$\int_{-\infty}^{\infty} \ell(dx) < \infty .$$

It is important to emphasize that the above condition implies that a finite activity Lévy process has a finite expected number of small jumps and a finite expected number of large jumps per unit of time:

$$\int_{|x|<1} \ell(dx) < \infty \text{ and } \int_{|x|>1} \ell(dx) < \infty .$$

Definition 3.20 Infinite Activity Lévy processes Let $(X_{t\in[0,\infty)})$ be a real valued Lévy process with Lévy triplet (A, ℓ, γ) defined on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,\infty)}, \mathbb{P})$. A

Lévy process $(X_{t\in[0,\infty)})$ is said to be an infinite activity Lévy process if its Lévy measure ℓ has an infinite integral:

$$\int_{-\infty}^{\infty} \ell(dx) = \infty .$$

It is important to emphasize that the above condition implies that an infinite activity Lévy process has an infinite expected number of small jumps and a finite expected number of large jumps per unit of time:

$$\int_{|x|<1} \ell(dx) = \infty \text{ and } \int_{|x|>1} \ell(dx) < \infty .$$

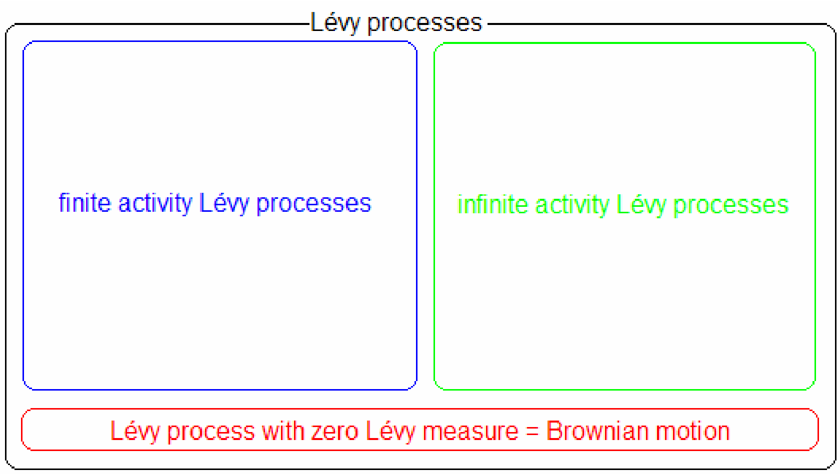


Figure 3.5 Classification of Lévy processes in terms of the behavior of Lévy measure ℓ

[3.7.3] In Terms of the Total Variation of Lévy Process

Definition 3.21 Total variation of a stochastic process Consider a real valued stochastic process $(X_{t\in[0,T]})$ on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,T]}, \mathbb{P})$. Partition the time interval $[0,T]$ with the points:

$$0 = t_0 < t_1 < t_2 ... < t_{n-1} < t_n = T .$$

Then, the total variation of a stochastic process $(X_{t\in[0,T]})$ on the time interval $[0,T]$ is defined by:

$$T(X) = \sup \sum_{i=1}^{n} \left| X(t_i) - X(t_{i-1}) \right| ,$$

where sup indicates a supremum (least upper bound).

Definition 3.22 Lévy processes of finite variation A real valued Lévy process $(X_{t\in[0,T]})$ on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,T]}, \mathbb{P})$ on the interval $[0,T]$ is said to be a Lévy process of finite variation, if the total variation on the interval $[0,T]$ of a sample path of the Lévy process is finite with probability 1:

$$\mathbb{P}\left(T(X) = \sup \sum_{i=1}^{n} \left| X(t_i) - X(t_{i-1}) \right| < \infty \right) = 1 .$$

Theorem 3.11 Lévy processes of finite variation If $(X_{t\in[0,T]})$ is a real valued Lévy process of finite variation on the interval $[0,T]$ defined on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,T]}, \mathbb{P})$, then, its Lévy triplet (A, ℓ, γ) satisfies:

$$A = 0 \text{ and } \int_{|x|<1} |x| \, \ell(dx) < \infty .$$

This corresponds to the proposition 3.9 of Cont and Tankov (2004) where its proof is given. Also consult a Sato's monograph in Barndorff-Nielsen et al (2001) page 6.

Lévy-Khinchin representation for Lévy processes of finite variation is given by theorem 3.8.

Definition 3.23 Lévy processes of infinite variation A real valued Lévy process $(X_{t\in[0,T]})$ on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,T]}, \mathbb{P})$ on the interval $[0,T]$ is said to be a Lévy process of infinite variation, if the total variation on the interval $[0,T]$ of a sample path of the Lévy process is infinite with probability 1:

$$\mathbb{P}\left(T(X) = \sup \sum_{i=1}^{n} \left| X(t_i) - X(t_{i-1}) \right| = \infty \right) = 1 .$$

Theorem 3.12 Lévy processes of infinite variation If $(X_{t\in[0,T]})$ is a real valued Lévy process of infinite variation on the interval $[0,T]$ defined on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,T]}, \mathbb{P})$, then, its Lévy triplet (A, ℓ, γ) satisfies:

$$A \neq 0 \text{ or } \int_{|x|<1} |x| \, \ell(dx) = \infty .$$

Consult a Sato's monograph in Barndorff-Nielsen et al (2001) page 6.

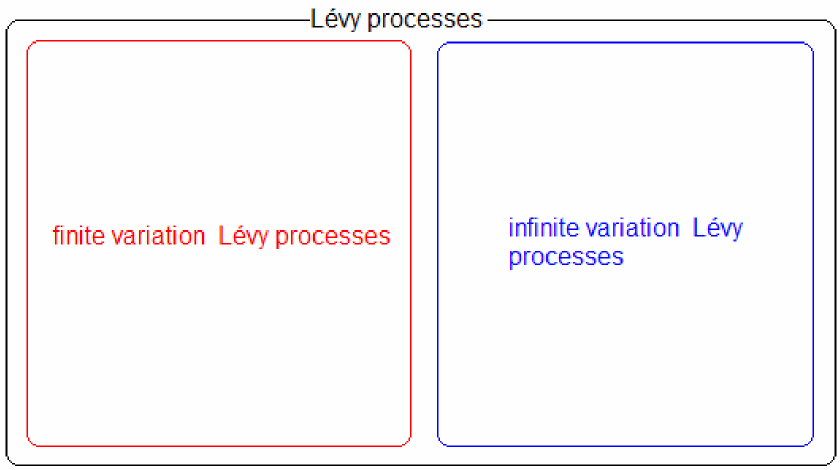


Figure 3.6 Classification of Lévy processes in terms of its total variation on the interval $[0,T]$

[3.7.4] In Terms of the Properties of Lévy Triplet (A,ℓ,γ) by Sato

Definition 3.24 Lévy processes of type A, type B, and type C Consider a real valued Lévy process $(X_{t\in[0,T]})$ with the Lévy triplet (A,ℓ,γ) on a filtered probability space $(\Omega,\mathcal{F}_{t\in[0,T]},\mathbb{P})$ on the time interval $[0,T]$.

(1) $(X_{t\in[0,T]})$ is said to be a type A Lévy process, if the Lévy triplet (A,ℓ,γ) satisfies:

$$A=0 \text{ and } \int_{-\infty}^{\infty}\ell(dx)<\infty .$$

In other words, a type A Lévy process is a purely non-Gaussian finite activity Lévy process whose sample paths have a finite number of small and large jumps in any finite time interval. A compound Poisson process is a typical example of a type A Lévy process.

(2) $(X_{t\in[0,T]})$ is said to be a type B Lévy process, if the Lévy triplet (A,ℓ,γ) satisfies:

$$A=0\,,\ \int_{-\infty}^{\infty}\ell(dx)=\infty\,,\text{ and}\int_{|x|<1}|x|\,\ell(dx)<\infty .$$

In other words, a type B Lévy process is a purely non-Gaussian infinite activity Lévy process of finite variation whose sample paths have an infinite number of small jumps and a finite number of large jumps in any finite time interval. A variance process is one example of a type B Lévy process.

(3) $(X_{t\in[0,T]})$ is said to be a type C Lévy process, if the Lévy triplet (A,ℓ,γ) satisfies:

$$A\neq 0 \text{ or } \int_{|x|<1}|x|\,\ell(dx)=\infty .$$

In other words, a type C Lévy process is a Lévy process of infinite variation. A Brownian motion (with drift) is a typical example of a type C Lévy process.

These definitions are from Sato's monograph in Barndorff-Nielsen et al (2001) page 6.

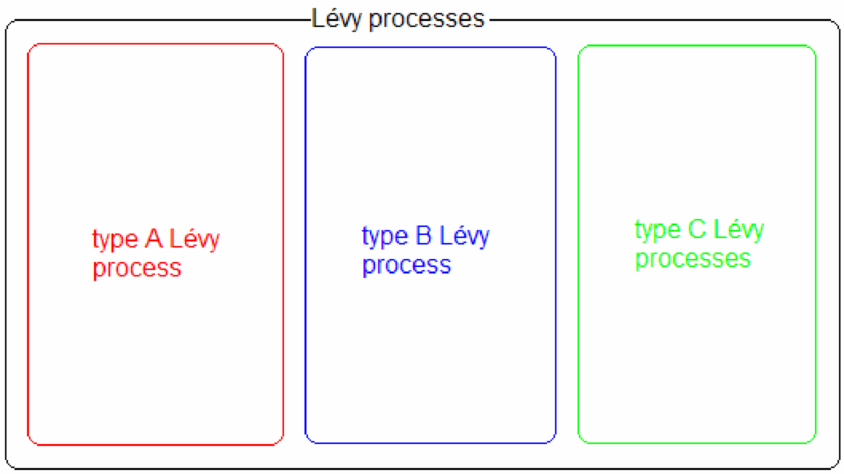


Figure 3.7 Sato's Classification of Lévy processes in terms of the properties of Lévy triplet (A, ℓ, γ)

[3.7.5] In Terms of the Sample Paths Properties of Lévy Processes

Theorem 3.13 A Lévy process with continuous sample path Let $(X_{t\in[0,\infty)})$ be a real valued Lévy process on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,\infty)}, \mathbb{P})$. If the sample paths of $(X_{t\in[0,\infty)})$ are continuous with probability 1, then, $(X_{t\in[0,\infty)})$ is a Gaussian Lévy process.

This corresponds to theorem 11.7 of Sato (1999). Theorem 3.13 together with Lévy-Itô decomposition proves that a normal distribution is the only distribution which generates a real valued Lévy process with continuous sample paths.

Theorem 3.14 A Lévy process with a piecewise constant sample path Let $(X_{t\in[0,\infty)})$ be a real valued Lévy process on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,\infty)}, \mathbb{P})$. If and only if a sample path of $(X_{t\in[0,\infty)})$ is piecewise constant with probability 1, then, $(X_{t\in[0,\infty)})$ is a compound Poisson process.

Theorem 3.15 The Converse of theorem 3.14 Let $(X_{t\in[0,\infty)})$ be a compound Poisson process on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,\infty)}, \mathbb{P})$ whose Lévy triplet (A, ℓ, γ) satisfies $A = 0$, $\int_{-\infty}^{\infty} \ell(dx) < \infty$, and $\int_{|x|<1} |x| \ell(dx) < \infty$ (by the definition of a compound Poisson process). Then, a sample path of a compound Poisson process is piecewise constant with probability 1.

Proof

Consult proposition 3.3 of Cont and Tankov (2004).

Definition 3.25 Increasing Lévy processes (Subordinators) Consider a real valued Lévy process $(X_{t\in[0,\infty)})$ with the Lévy triplet (A,ℓ,γ) on a filtered probability space $(\Omega,\mathcal{F}_{t\in[0,\infty)},\mathbb{P})$. A Lévy process $(X_{t\in[0,\infty)})$ is said to be an increasing Lévy process if sample paths of the process $(X_{t\in[0,\infty)})$ are non-decreasing with probability 1:

$$\mathbb{P}(X_{t+h}\geq X_t : h\in\mathbb{R}^+)=1.$$

Theorem 3.16 Increasing Lévy processes (Subordinators) Let $(X_{t\in[0,\infty)})$ be an increasing Lévy process on a filtered probability space $(\Omega,\mathcal{F}_{t\in[0,\infty)},\mathbb{P})$. Then, its Lévy triplet (A,ℓ,γ) satisfies the following conditions:

(1) $A=0$. Gaussian variance is zero. In other words, sample paths of $(X_{t\in[0,\infty)})$ do not contain Brownian motion (with drift) parts (Lévy-Itô decomposition).
(2) $\int_{-\infty}^{0}\ell(dx)=0$. The Lévy measure of an increasing Lévy process is concentrated on positive half axis. In other words, an increasing Lévy process has no negative jumps.
(3) $\int_{|x|<1}|x|\,\ell(dx)<\infty$. But because we always have positive jumps, this becomes $\int_{(0,1)}x\ell(dx)<\infty$. This means that an increasing Lévy process is a Lévy process of finite variation (theorem 3.11).
(4) $\gamma_0\geq 0$. Sample paths of $(X_{t\in[0,\infty)})$ have nonnegative drift (Lévy-Itô decomposition).

For more details and the proofs, consult Sato (1999) page 137-142.

Definition 3.26 Symmetric Lévy processes Consider a real valued Lévy process $(X_{t\in[0,\infty)})$ with the Lévy triplet (A,ℓ,γ) on a filtered probability space $(\Omega,\mathcal{F}_{t\in[0,\infty)},\mathbb{P})$. A Lévy process $(X_{t\in[0,\infty)})$ is said to be symmetric if sample paths of the process $(X_{t\in[0,\infty)})$ satisfy:

$$(X_t)\ \underline{\underline{d}}\ (-X_t),$$

which means that (X_t) and $(-X_t)$ have identical distributions.

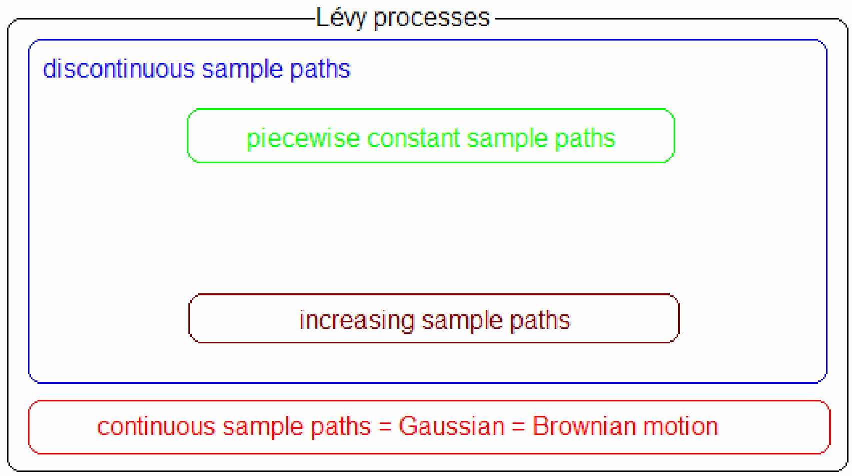


Figure 3.8 Classification of Lévy processes in terms of Sample Paths Properties

[3.8] Lévy Processes as a Subclass of Markov Processes

For the background knowledge of Markov processes, read section 2.3.3.

Definition 3.27 Transition function Consider a continuous time nonanticipating stochastic process $(X_{t\in[0,\infty)})$ defined on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,\infty)}, \mathbb{P})$ which takes values in a measurable space $(B, \mathcal{B})$ (i.e. $B \in \mathcal{B}(\mathbb{R})$). $(B, \mathcal{B})$ is called a state space of the process and the process is said to be B - valued. Consider an increasing sequence of time $0 \le t \le u \le v < \infty$. A real valued transition function $\mathbb{P}_{t,v}(x, B)$ with $x \in \mathbb{R}$ and $B \in \mathcal{B}(\mathbb{R})$ is a mapping which satisfies the following conditions:

(1) $\mathbb{P}_{t,v}(x, B)$ is a probability measure which maps every fixed x into B.

(2) $\mathbb{P}_{t,v}(x, B)$ is $\mathcal{B}$ - measurable for every $B \in \mathcal{B}(\mathbb{R})$.

(3) $\mathbb{P}_{t,t}(x, B) = \delta(B)$.

(4) $\mathbb{P}_{t,v}(x, B) = \int_{\mathbb{R}} \mathbb{P}_{t,u}(x, dy)\mathbb{P}_{u,v}(y, B)$.

The condition (4) is called the Chapman-Kolmogorov identity. Chapman-Kolmogorov identity means that the transition probability $\mathbb{P}_{t,v}(x, B)$ of moving from a state x at time t to a state B at time v can be calculated as a sum (i.e. integral) of the product of the transition probabilities via an intermediate time $t \le u \le v$, i.e. $\mathbb{P}_{t,u}(x, dy)$ and $\mathbb{P}_{u,v}(y, B)$. In the general cases, transition functions are dependent on the states and time.

Definition 3.28 Time homogeneous (temporary homogeneous or stationary) transition function Consider an increasing sequence of time $0 \le t \le u \le v < \infty$. A real valued transition function $\mathbb{P}_{t,v}(x, B)$ with $x \in \mathbb{R}$ and $B \in \mathcal{B}(\mathbb{R})$ is said to be time homogeneous if it satisfies:

$$\mathbb{P}_{t,v}(x,B) = \mathbb{P}_{0,v-t}(x,B) = \mathbb{P}_{v-t}(x,B),$$

which indicates that the transition function $\mathbb{P}_{t,v}(x,B)$ of moving from a state x at time t to a state B at time v is equivalent to the transition function $\mathbb{P}_{0,v-t}(x,B)$ of moving from a state x at time 0 to a state B at time $v-t$. In other words, the transition function is independent of the time t and depends only on the interval of time $v-t$.

Definition 3.29 Chapman-Kolmogorov identity for the time homogeneous transition function Consider an increasing sequence of time $0 \le t \le u < \infty$. Chapman-Kolmogorov identity for the time homogeneous transition function is:

$$\int_{\mathbb{R}} \mathbb{P}_{0,t}(x,dy)\mathbb{P}_{0,u}(y,B) = \int_{\mathbb{R}} \mathbb{P}_{t}(x,dy)\mathbb{P}_{u}(y,B) = \mathbb{P}_{t+u}(x,B).$$

Definition 3.30 Markov Processes (less formal) Consider a continuous time nonanticipating stochastic process $(X_{t\in[0,\infty)})$ defined on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,\infty)}, \mathbb{P})$. Then, the process $(X_{t\in[0,\infty)})$ is said to be a Markov process if it satisfies, for every increasing sequence of time $0 < t_1 \le t_2 \le ... \le t_n \le t \le u < \infty$:

$$\mathbb{P}(X_u \,|\, \mathcal{F}_t) = \mathbb{P}(X_u \,|\, X_0, X_{t_1}, X_{t_2}, ..., X_{t_n}, X_t) = \mathbb{P}(X_u \,|\, X_t),$$

Informally, Markov property means that the probability of a random variable X_u at time $u \ge t$ (tomorrow) conditional on the entire history of the stochastic process $\mathcal{F}_{[0,t]} \equiv X_{[0,t]}$ is equal to the probability of a random variable X_u at time $u \ge t$ (tomorrow) conditional only on the value of a random variable at time t (today). In other words, the history (sample path) of the stochastic process $\mathcal{F}_{[0,t]}$ is of no importance in that the way this stochastic process evolved or the dynamics does not mean a thing in terms of the conditional probability of the process. This property is sometimes called a memoryless property.

Definition 3.31 Markov Processes (formal) Consider a continuous time nonanticipating stochastic process $(X_{t\in[0,\infty)})$ defined on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,\infty)}, \mathbb{P})$ which takes values in a measurable space $(E, \mathcal{E})$. $(E, \mathcal{E})$ is called a state space of the process and the process is said to be E - valued. Then, the process $(X_{t\in[0,\infty)})$ is said to be a Markov process if it satisfies, for an increasing sequence of time $0 \le t \le u \le v < \infty$ $0 < t \le u < \infty$:

$$E[X_v \,|\, \mathcal{F}_t] = E[X_v \,|\, X_t],$$

with the transition function (defined by the definition 3.27):

$$\mathbb{P}_{t,v}(x,B) = \int_{\mathbb{R}} \mathbb{P}_{t,u}(x,dy)\mathbb{P}_{u,v}(y,B).$$

Definition 3.32 Spatially homogeneous transition function Consider an increasing sequence of time $0 \le t \le u \le v < \infty$. A real valued transition function $\mathbb{P}_{t,v}(x,B)$ with $x \in \mathbb{R}$ and $B \in \mathcal{B}(\mathbb{R})$ is said to be spatially homogeneous if it satisfies, for $\forall t, v, x, B, B - x \in \{y - x : y \in B\}$:

$$\mathbb{P}_{t,v}(x,B) = \mathbb{P}_{t,v}(0, B - x).$$

Theorem 3.17 Lévy processes Let $(X_{t\in[0,\infty)})$ be a real valued Lévy process with Lévy triplet (A, ℓ, γ) defined on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,\infty)}, \mathbb{P})$. Then, the transition functions of Lévy processes satisfy, for $0 \le t \le u \le v < \infty$:

$$\mathbb{P}_{t,v}(x,B) = \mathbb{P}_{0,v-t}(0, B - x),$$

in other words, Lévy processes are a subclass of Markov processes with the time homogeneous and spatially homogeneous transition functions. Its converse is also true.

Theorem 3.18 Relationship between Markov processes and Lévy processes Let $(X_{t\in[0,\infty)})$ be a real valued Markov process defined on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,\infty)}, \mathbb{P})$. Then, a Markov process becomes a Lévy process if it satisfies the following additional conditions:

(1) The process is stochastically continuous: $\forall \varepsilon > 0$, $\lim_{h\to 0} \mathbb{P}(|X_{t+h} - X_t| \ge \varepsilon) = 0$.

(2) Its transition functions are time homogeneous and spatially homogeneous:

$$\mathbb{P}_{t,v}(x,B) = \mathbb{P}_{0,v-t}(0, B - x).$$

Proof

For the proofs of theorem 17 and 18 and for the more details about Lévy processes as a subclass of Markov processes, consult Sato (1999) section 10.

Therefore, we can state that all Lévy processes are Markov process, but the converse is not true.

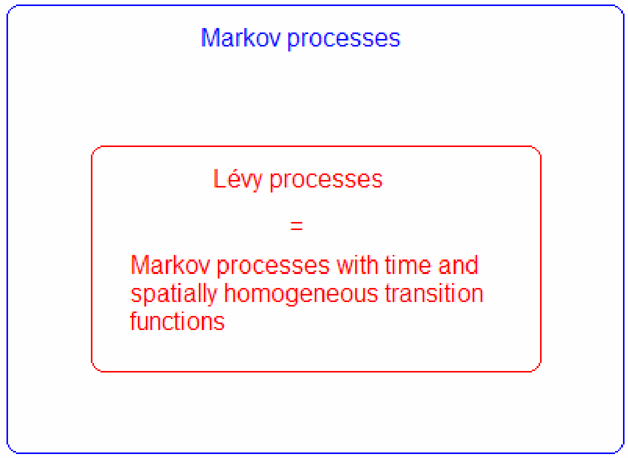


Figure 3.9 Lévy processes as a subclass of Markov processes

It turns out that Lévy processes satisfy not only Markov property, but also strong Markov property.

Theorem 3.19 Strong Markov property of Lévy processes Let $(X_{t\in[0,\infty)})$ be a real valued Lévy process with Lévy triplet (A,ℓ,γ) defined on a filtered probability space $(\Omega,\mathcal{F}_{t\in[0,\infty)},\mathbb{P})$. Define a new stochastic process as:

$$(X_{t+h}-X_h : t\in[0,\infty), h\geq 0).$$

Then, a process $(X_{t+h}-X_h)$ is a Lévy process satisfying:

(1) $(X_{t+h}-X_h : t\in[0,\infty), h\geq 0) \underset{=}{\underline{d}} (X_{t\in[0,\infty)})$.
(2) $(X_{t+h}-X_h : t\in[0,\infty), h\geq 0)$ and $(X_t : 0\leq t\leq h)$ are independent.

Proof

Consult Sato (1999) section 10.

[3.9] Other Important Properties of Lévy Processes

Theorem 3.20 Linear transformation of Lévy processes Consider a real valued Lévy process $(X_{t\in[0,T]})$ on a filtered probability space $(\Omega,\mathcal{F}_{t\in[0,T]},\mathbb{P})$ whose Lévy triplet is (A,ℓ,b). Let c be a constant on $\mathbb{R}$. Then, a linear transformation of the original Lévy Process $(cX_{t\in[0,T]})$ is also a real valued Lévy process whose Lévy triplet (A_c,ℓ_c,b_c) is given by:

$$A_c = c^2 A,$$

$$\ell_c = \left[\frac{\ell}{c}\right]_{\mathbb{R}\setminus\{0\}} \text{ (i.e. a restriction of a measure } \ell c^{-1} \text{ to } \mathbb{R}\setminus\{0\}\text{)},$$

$$b_c = cb + \int_{-\infty}^{\infty} cx\left(cx1_{|x|\le 1} - x1_{|x|\le 1}\right)\ell(dx) .$$

This corresponds to proposition 11.10 of Sato (1999) where its proof is given.

Theorem 3.21 Independence of Lévy processes

Part 1) Consider a two dimensional real valued non-Gaussian Lévy process $(X_{t\in[0,T]}, Y_{t\in[0,T]})$ on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,T]}, \mathbb{P})$ whose Lévy triplet is $(A = 0, \ell, b)$. Then, X_t and Y_t are independent (i.e. $\mathbb{P}(X_t | Y_t) = \mathbb{P}(X_t)$), if and only if the set $\{(x, y) : xy = 0\}$ contains the support of its Lévy measure ℓ . In other words, X_t and Y_t are independent, if and only if X_t and Y_t do not jump together.

Part 2) If X_t and Y_t are independent, then, for any arbitrary set E , the Lévy measure of X_t can be expressed as (i.e. when X_t jumps, Y_t never jumps):

$$\ell_X(E_X) = \ell_X(\{x : (x,0) \in E\}) ,$$

and the Lévy measure of Y_t can be expressed as:

$$\ell_Y(E_Y) = \ell_Y(\{y : (0, y) \in E\}) .$$

And, the Lévy measure of a two dimensional independent non-Gaussian Lévy process $(X_{t\in[0,T]}, Y_{t\in[0,T]})$ is given by the sum of these two Lévy measures:

$$\ell(E) = \ell_X(E_X) + \ell_Y(E_Y) .$$

This corresponds to proposition 5.3 of Cont and Tankov (2004) where its proof is given.

Theorem 3.22 Sums of independent Lévy processes If two Lévy processes $(X_{t\in[0,T]})$ and $(Y_{t\in[0,T]})$ with Lévy triplets (A_X, ℓ_X, b_X) and (A_Y, ℓ_Y, b_Y) are independent, then, their sum $X_t + Y_t$ is also a Lévy process. For more details and its proof, consult an example 4.1 of Cont and Tankov (2004).

[4] Examples of Lévy Processes

In this section, we present some building blocks of Lévy processes.

[4.1] Brownian Motion: The only Lévy Process with Continuous Sample Paths

Some people (non-mathematicians) believe that all Lévy processes have discontinuous sample paths, in other words, Lévy processes are equivalent to jump processes. But this belief is wrong. It is true to state that most (but one) Lévy processes have discontinuous sample paths. But you can carefully go through the definition of Lévy processes again and you'll notice that the only sample paths condition which Lévy processes must satisfy is being a right continuous with left limit (i.e. rcll) stochastic process. This in turn indicates that a Lévy process can have continuous sample paths because all continuous processes are rcll processes by definition (read section 2). It turns out that the only Lévy process having continuous sample paths is a Brownian motion.

[4.1.1] Definition of a Brownian Motion

Definition 4.1 Standard Brownian motion (Standard Wiener process) A standard Brownian motion $(B_{t\in[0,\infty)})$ is a real valued stochastic process defined on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,\infty)}, \mathbb{P})$ satisfying:

(1) Its increments are independent. In other words, for $0 \le t_1 < t_2 < ... < t_n < \infty$:

$$\begin{aligned} &\mathbb{P}(B_{t_0} \cap B_{t_1} - B_{t_0} \cap B_{t_2} - B_{t_1} \cap ... \cap B_{t_n} - B_{t_{n-1}}) \\ &= \mathbb{P}(B_{t_0})\mathbb{P}(B_{t_1} - B_{t_0})\mathbb{P}(B_{t_2} - B_{t_1})...\mathbb{P}(B_{t_n} - B_{t_{n-1}}). \end{aligned}$$

(2) Its increments are stationary (time homogeneous): i.e. for $h \ge 0$, $B_{t+h} - B_t$ has the same distribution as B_h . In other words, the distribution of increments does not depend on t .

(3) $\mathbb{P}(B_0 = 0) = 1$. The process starts from 0 almost surely (with probability 1).

(4) $B_t \sim Normal(0,t)$. Its increments follow a Gaussian distribution with the mean 0 and the variance t .

Theorem 4.1 Standard Brownian motion process (Standard Wiener process) A standard Brownian motion process $(B_{t\in[0,\infty)})$ defined on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,\infty)}, \mathbb{P})$ satisfies the following conditions:

(1) The process is stochastically continuous: $\forall \varepsilon > 0$, $\lim_{h\to 0} \mathbb{P}(|X_{t+h} - X_t| \ge \varepsilon) = 0$.

(2) Its sample path (trajectory) is continuous in t (i.e. continuous $\in$ rcll) almost surely.

Proof

Consult Karlin (1975). We have to remind you that this proof is not that simple.

As many of readers realize, the above definitions and conditions all appear to define Lévy processes in definition 3.1 other than (4) of definition 4.1. Let us put this into another words.

Definition 4.2 Standard Brownian motion as a Lévy process A standard Brownian motion $(B_{t\in[0,\infty)})$ defined on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,\infty)}, \mathbb{P})$ is a Lévy process satisfying:

(1) $B_t \sim Normal(0,t)$. Its increments follow a Gaussian distribution with the mean 0 and the variance t.
(2) Its sample path (trajectory) is continuous in t (i.e. continuous $\in$ rcll) almost surely.

The condition (1) implies the condition (2).

Or more generally, a Brownian motion with drift is an only Lévy process with continuous sample paths.

Definition 4.3 Brownian motion with drift Let $(B_{t\in[0,\infty)})$ be a standard Brownian motion process defined on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,\infty)}, \mathbb{P})$. Then, a Brownian motion with drift is a real valued stochastic process defined on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,\infty)}, \mathbb{P})$ as:

$$(X_{t\in[0,\infty)}) \equiv (\mu t + \sigma B_{t\in[0,\infty)}),$$

where $\mu \in \mathbb{R}$ is called a drift and $\sigma \in \mathbb{R}^+$ is called a diffusion (volatility) parameter. A Brownian motion with drift satisfies the following conditions:

(1) Its increments are independent. In other words, for $0 \le t_1 < t_2 < ... < t_n < \infty$:

$$\mathbb{P}(X_{t_0} \cap X_{t_1} - X_{t_0} \cap X_{t_2} - X_{t_1} \cap ... \cap X_{t_n} - X_{t_{n-1}})$$
$$= \mathbb{P}(X_{t_0})\mathbb{P}(X_{t_1} - X_{t_0})\mathbb{P}(X_{t_2} - X_{t_1})...\mathbb{P}(X_{t_n} - X_{t_{n-1}}).$$

(2) Its increments are stationary (time homogeneous): i.e. for $h \ge 0$, $X_{t+h} - X_t$ has the same distribution as X_h. In other words, the distribution of increments does not depend on t.
(3) $X_t \equiv \mu t + \sigma B_t \sim Normal(\mu t, \sigma^2 t)$. Its increments follow a Gaussian distribution with the mean μt and the variance $\sigma^2 t$.
(4) Its sample path (trajectory) is continuous in t (i.e. continuous $\in$ rcll) almost surely.

Definition 4.4 Brownian motion with drift as a Lévy process A Brownian motion with drift $(X_{t\in[0,\infty)}) \equiv (\mu t + \sigma B_{t\in[0,\infty)})$ defined on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,\infty)}, \mathbb{P})$ is a Lévy process satisfying:

(1) $X_t \equiv \mu t + \sigma B_t \sim Normal(\mu t, \sigma^2 t)$. Its increments follow a Gaussian distribution with the mean μt and the variance $\sigma^2 t$.
(2) Its sample path (trajectory) is continuous in t (i.e. continuous $\in$ rcll) almost surely.

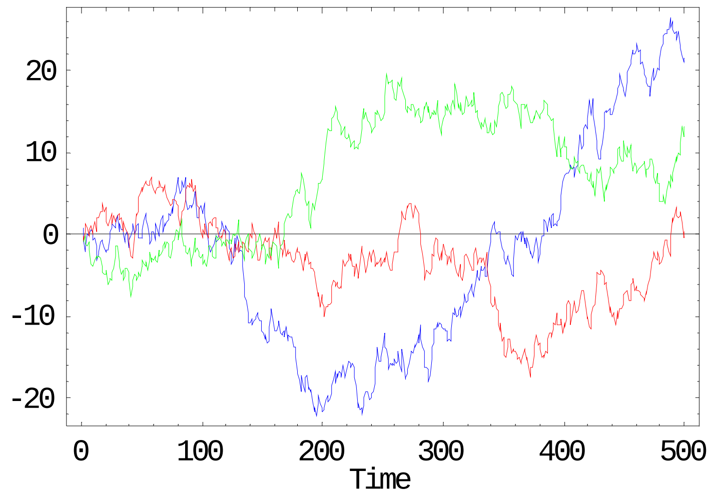
Next, we define a Brownian motion with drift process in terms of the properties of its Lévy triplet (A, ℓ, γ) and sample paths.

Definition 4.5 Brownian motion with drift in terms of Lévy triplet (A, ℓ, γ) Let $(X_{t\in[0,\infty)})$ be a real valued Lévy process with Lévy triplet (A, ℓ, γ) defined on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,\infty)}, \mathbb{P})$. Then, $(X_{t\in[0,\infty)})$ is a Brownian motion with drift, if it satisfies one of the following conditions:

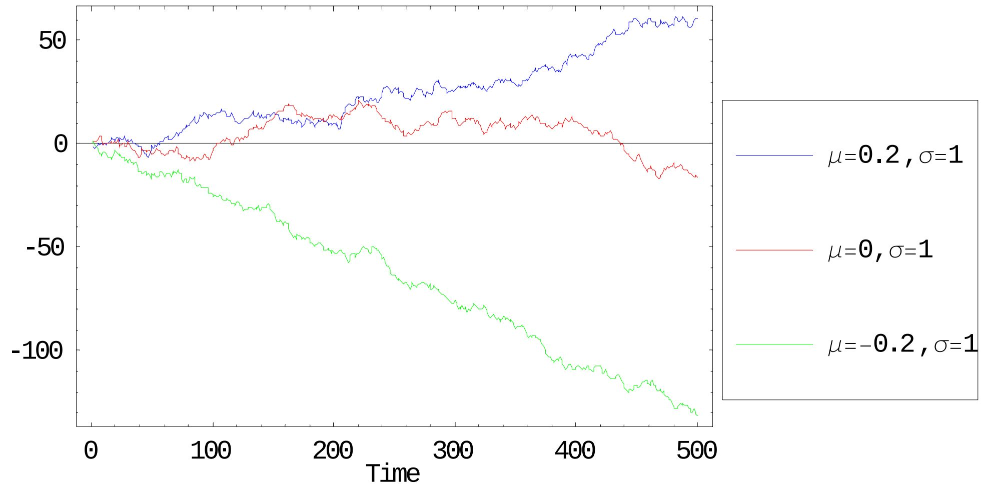
(1) $\ell = 0$. The Lévy measure ℓ of this Lévy process is zero. Zero Lévy measure $\ell = 0$ means that the Lévy process $(X_{t\in[0,\infty)})$ has no small or large jumps (i.e. no jumps at all), in other words, sample paths of the process is continuous with probability 1.
(2) Its increments follow a Gaussian distribution.

[4.1.2] Sample Paths Properties of a Brownian Motion

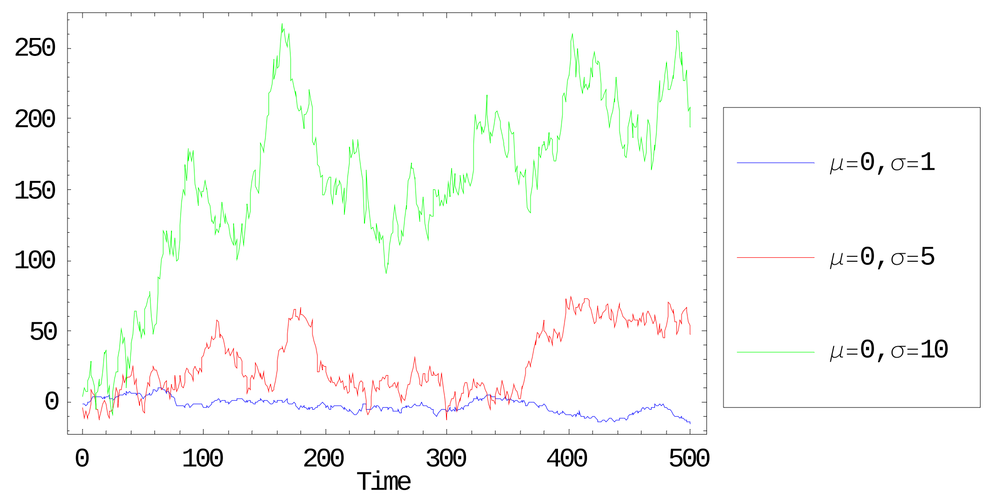
Before discussing the sample paths properties of Brownian motion, take a look at simulated sample paths of a standard Brownian motion on Panel (A) in Figure 1.1 and those of a Brownian motion with drift on Panel (B) and (C).



A) Sample Paths of Standard Brownian Motions.



B) Sample Paths of Brownian Motions with Drift. Different drifts and same diffusion parameters.



C) Sample Paths of Brownian Motions with Drift. Zero drifts and different diffusion parameters.

Figure 4.1 Simulated Sample Paths of Brownian Motions with Drift

Theorem 4.2 Sample paths properties of Brownian motion with drift Consider a real valued Brownian motion with drift process $(X_{t\in[0,\infty)}) \equiv (\mu t + \sigma B_{t\in[0,\infty)})$ defined on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,\infty)}, \mathbb{P})$ and let (A, ℓ, γ) be its Lévy triplet. Then, the sample paths of $(X_{t\in[0,\infty)})$ possess following properties:

(1) Sample paths are continuous with probability 1. In other words, the Lévy measure ℓ of the process is zero. Zero Lévy measure $\ell = 0$ means that a Brownian motion with drift process has no small or large jumps, therefore, no jumps at all.

(2) Sample paths are of infinite variation on any finite interval $[0,t]$. In other words, the total variation on any finite interval $[0,t]$ of a sample path of a Brownian motion with drift is infinite with probability 1 in the limit $n \to \infty$ (as the partition becomes finer and finer):

$$\mathbb{P}\left(\lim_{n\to\infty} T(X) = \lim_{n\to\infty} \sup \sum_{i=1}^{n} \left| X(t_i) - X(t_{i-1}) \right| = \infty \right) = 1.$$

As we saw in section 3.7.3, the infinite variation property is equivalent to the nonzero Gaussian variance in terms of the Lévy triplet (A, ℓ, γ):

$$A \neq 0 .$$

Intuitively speaking, the infinite variation property means highly oscillatory sample paths.
(3) The quadratic variations of sample paths of Brownian motions with drift $(X_{t\in[0,\infty)})$ are finite on any finite interval $[0,t]$ and converge to $\sigma^2 t$ with probability 1 in the limit $n \to \infty$ (as the partition becomes finer and finer):

$$\mathbb{P}\left(\lim_{n\to\infty} T^2(X) = \lim_{n\to\infty} \sup \sum_{i=1}^{n} \left| X(t_i) - X(t_{i-1}) \right|^2 = \sigma^2 t < \infty \right) = 1.$$

For more details and proofs about theorem 4.2, consult Sato (1999) page 22 – 28 and Karatzas and Shreve (1991) section 1.5 and 2.9. We also recommend Rogers and Williams (2000) chapter 1.

[4.1.3] Equivalent Transformations of Standard Brownian Motion

Theorem 4.3 Equivalent transformations of Standard Brownian motion If $(B_{t\in[0,\infty)})$ is a real valued standard Brownian motion defined on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,\infty)}, \mathbb{P})$, then, it satisfies the four conditions:

(1) A standard Brownian motion $(B_{t\in[0,\infty)})$ is symmetric. In other words, the process $(-B_{t\in[0,\infty)})$ is also a standard Brownian motion:

$$(B_{t\in[0,\infty)}) \underline{\underline{d}} (-B_{t\in[0,\infty)}) .$$

(2) A standard Brownian motion $(B_{t\in[0,\infty)})$ has a time shifting property. In other words, the process $(B_{t+A} - B_A)$ is also a standard Brownian motion for $\forall A \in \mathbb{R}^+$:

$$(B_{t+A} - B_A) \underline{\underline{d}} (B_{t\in[0,\infty)}) .$$

(3) Time scaling property of a standard Brownian motion. For any nonzero $c \in \mathbb{R}$, the process $(\sqrt{c} B_{t/c})$ or $(\frac{1}{\sqrt{c}} B_{ct})$ is also a standard Brownian motion:

$$(\frac{1}{\sqrt{c}} B_{ct}) \underset{=}{d} (\sqrt{c} B_{t/c}) \underset{=}{d} (B_{t \in [0,\infty)}).$$

(4) Time inversion property of a standard Brownian motion (i.e. a variant of (3)). The process defined as:

$$(\tilde{B}_{t \in [0,\infty)}) = \begin{cases} 0 \quad \text{if} \quad t = 0 \\ (tB_{1/t}) \quad \text{if } 0 < t < \infty \end{cases},$$

is also a standard Brownian motion:

$$(\tilde{B}_{t \in [0,\infty)}) \underset{=}{d} (B_{t \in [0,\infty)}).$$

Proof

These are easy exercises for readers. For the proof of the continuity of $(\tilde{B}_{t \in [0,\infty)})$ at 0, consult Rogers and Williams (2000) page 4.

[4.1.4] Characteristic Function of Brownian Motion

Consider a real valued Brownian motion with drift process $(X_{t \in [0,\infty)}) \equiv (\mu t + \sigma B_{t \in [0,\infty)})$ defined on a filtered probability space $(\Omega, \mathcal{F}_{t \in [0,\infty)}, \mathbb{P})$ and let (A, ℓ, γ) be its Lévy triplet. Its characteristic function can be obtained by two approaches. First approach is the direct use of the definition of a characteristic function (i.e. Fourier transform of the probability density function with Fourier transform parameters $(1,1)$):

$$\phi_{X_t}(\omega) \equiv \mathcal{F}\left[\mathbb{P}(x)\right] \equiv \int_{-\infty}^{\infty} e^{i\omega x} \mathbb{P}(x) dx$$
$$\phi_{X_t}(\omega) = \int_{-\infty}^{\infty} e^{i\omega x} \frac{1}{\sqrt{2\pi\sigma^2 t}} \exp\left\{ -\frac{\left(x - \mu t\right)^2}{2\sigma^2 t} \right\} dx$$
$$\phi_{X_t}(\omega) = \exp(i\mu t\omega - \frac{\sigma^2 t \omega^2}{2}).$$

Second approach is the use of Lévy-Khinchin representation (theorem 3.3). Because we know the Lévy triplet of a Brownian motion with drift is given as $(A = \sigma^2, \ell = 0, \gamma = \mu)$, its characteristic exponent is given by:

$$\psi_X(\omega) = -\frac{A\omega^2}{2} + i\gamma\omega + \int_{-\infty}^{\infty} \left\{ \exp(i\omega x) - 1 - i\omega x 1_D \right\} \ell(dx)$$

$$\psi_X(\omega) = -\frac{\sigma^2\omega^2}{2} + i\mu\omega \,,$$

thus, its characteristic function $\phi_X(\omega)$ is expressed as:

$$\phi_X(\omega) = \exp\left(t\psi_X(\omega)\right),$$

$$\phi_{X_t}(\omega) = \exp(-\frac{\sigma^2 t\omega^2}{2} + i\mu t\omega) \,.$$

Table 4.1 summarizes the properties of a Brownian motion with drift.

Table 4.1 Brownian motion with drift

Lévy process	Gaussian variance A	Lévy measure ℓ	drift γ	variation	sample path
standard Brownian motion B_t	$A = 1$	$\ell = 0$	$\gamma = \mu = 0$	infinite	continuous
Brownian motion with drift $\mu t + \sigma B_t$	$A = \sigma^2$	$\ell = 0$	$\gamma = \mu \neq 0$	infinite	continuous

[4.1.5] Brownian Motion as a Subclass of Continuous Martingale

For the background knowledge of martingales, read section 2.3.2.

Theorem 4.4 Standard Brownian motion is a continuous martingale Let $(B_{t\in[0,\infty)})$ be a standard Brownian motion process defined on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,\infty)}, \mathbb{P})$. Then, $(B_{t\in[0,\infty)})$ is a continuous martingale with respect to the filtration $\mathcal{F}_{t\in[0,\infty)}$ and the probability measure $\mathbb{P}$.

Proof

By definition, $(B_{t\in[0,\infty)})$ is a nonanticipating process (i.e. $\mathcal{F}_{t\in[0,\infty)}$ - adapted process) with the finite mean $E[|B_t|] = 0 < \infty$ for $\forall t \in [0,\infty)$. For $\forall 0 \le t \le u < \infty$:

$$B_u = B_t + \int_t^u dB_v \,. \tag{1}$$

Using the equation (1) and the fact that a Brownian motion is a nonanticipating process, i.e. $E[B_t | \mathcal{F}_t] = B_t$:

$$E[B_u - B_t \,|\mathcal{F}_t] = E[B_u \,|\mathcal{F}_t] - E[B_t \,|\mathcal{F}_t] = E[B_t + \int_t^u dB_v \,|\mathcal{F}_t] - B_t$$
$$E[B_u - B_t \,|\mathcal{F}_t] = E[B_t \,|\mathcal{F}_t] + E[\int_t^u dB_v \,|\mathcal{F}_t] - B_t$$
$$E[B_u - B_t \,|\mathcal{F}_t] = B_t + 0 - B_t = 0\,,$$

or in other words:

$$E[B_u \,|\mathcal{F}_t] = E[B_t + \int_t^u dB_v \,|\mathcal{F}_t] = E[B_t \,|\mathcal{F}_t] + E[\int_t^u dB_v \,|\mathcal{F}_t] = B_t + 0$$
$$E[B_u \,|\mathcal{F}_t] = B_t\,,$$

which is a martingale condition.

□

Theorem 4.5 Brownian motion with drift is not a continuous martingale Let $(B_{t\in[0,\infty)})$ be a standard Brownian motion process defined on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,\infty)}, \mathbb{P})$. Then, a Brownian motion with drift $(X_{t\in[0,\infty)}) \equiv (\mu t + \sigma B_{t\in[0,\infty)})$ is not a continuous martingale with respect to the filtration $\mathcal{F}_{t\in[0,\infty)}$ and the probability measure $\mathbb{P}$.

Proof

By definition, $(X_{t\in[0,\infty)})$ is a nonanticipating process (i.e. $\mathcal{F}_{t\in[0,\infty)}$ - adapted process) with the finite mean $E[X_t] = E[\mu t + \sigma B_t] = \mu t < \infty$ for $\forall t \in [0,\infty)$ and $\mu \in \mathbb{R}$. For $\forall 0 \le t \le u < \infty$:

$$X_u = X_t + \int_t^u dX_v\,. \tag{2}$$

Using the equation (2) and the fact that a Brownian motion with drift is a nonanticipating process, i.e. $E[X_t \,|\mathcal{F}_t] = X_t$:

$$E[X_u \,|\mathcal{F}_t] = E[X_t + \int_t^u dX_v \,|\mathcal{F}_t] = E[X_t \,|\mathcal{F}_t] + E[\int_t^u dX_v \,|\mathcal{F}_t]$$
$$E[X_u \,|\mathcal{F}_t] = X_t + \mu(u - t)\,,$$

which violates a martingale condition.

□

Theorem 4.6 Detrended Brownian motion with drift is a continuous martingale Let $(B_{t\in[0,\infty)})$ be a standard Brownian motion process defined on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,\infty)}, \mathbb{P})$. Then, a detrended Brownian motion with drift defined as:

$$(X_{t\in[0,\infty)} - \mu t) \equiv (\mu t + \sigma B_{t\in[0,\infty)} - \mu t) \equiv (\sigma B_{t\in[0,\infty)}),$$

is a continuous martingale with respect to the filtration $\mathcal{F}_{t\in[0,\infty)}$ and the probability measure $\mathbb{P}$.

Proof

For $\forall 0 \le t \le u < \infty$:

$$E[X_u - \mu u | \mathcal{F}_t] = E[(X_t - \mu t) + (\int_t^u dX_v - \mu \int_t^u dv) | \mathcal{F}_t]$$
$$E[X_u - \mu u | \mathcal{F}_t] = E[(X_t - \mu t) | \mathcal{F}_t] + E[(\int_t^u dX_v - \mu \int_t^u dv) | \mathcal{F}_t]$$
$$E[X_u - \mu u | \mathcal{F}_t] = X_t - \mu t + \mu(u-t) - \mu(u-t)$$
$$E[X_u - \mu u | \mathcal{F}_t] = X_t - \mu t,$$

which satisfies a martingale condition.

□

Theorem 4.7 Exponential of a standard Brownian motion is a continuous martingale
Let $(B_{t\in[0,\infty)})$ be a standard Brownian motion process defined on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,\infty)}, \mathbb{P})$. Then, for any $\theta \in \mathbb{R}$, the exponential of a standard Brownian motion defined as:

$$Z_t = \exp(\theta B_t - \frac{1}{2}\theta^2 t), \tag{3}$$

is a continuous martingale with respect to the filtration $\mathcal{F}_{t\in[0,\infty)}$ and the probability measure $\mathbb{P}$.

Proof

We first prove the often used proposition.

Proposition 4.1 Note that if $X \sim Normal(\mu t, \sigma^2 t)$, then for any $\theta \in \mathbb{R}$:

$$E[\exp(\theta X)] = \exp(\theta \mu t + \frac{1}{2}\theta^2 \sigma^2 t). \tag{4}$$

Proof

$$
\begin{aligned}
E[\exp(\theta X)] &= \int_{-\infty}^{\infty} \exp(\theta X)\frac{1}{\sqrt{2\pi\sigma^2 t}}\exp\{-\frac{(X-\mu t)^2}{2\sigma^2 t}\}dX \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2 t}}\exp\{-\frac{-\theta X\, 2\sigma^2 t + X^2 - 2X\mu t + \mu^2 t^2}{2\sigma^2 t}\}dX \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2 t}}\exp\{-\frac{X^2 - 2(\theta\sigma^2 t + \mu t)X + \mu^2 t^2}{2\sigma^2 t}\}dX \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2 t}}\exp\{-\frac{(X-(\theta\sigma^2 t + \mu t))^2 - (\theta\sigma^2 t + \mu t)^2 + \mu^2 t^2}{2\sigma^2 t}\}dX \\
&= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2 t}}\exp\{-\frac{(X-(\theta\sigma^2 t + \mu t))^2}{2\sigma^2 t}\}\exp\{\frac{(\theta\sigma^2 t + \mu t)^2 - \mu^2 t^2}{2\sigma^2 t}\}dX \\
&= \exp\{\frac{(\theta\sigma^2 t + \mu t)^2 - \mu^2 t^2}{2\sigma^2 t}\}\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2 t}}\exp\{-\frac{(X-(\theta\sigma^2 t + \mu t))^2}{2\sigma^2 t}\}dX \\
&= \exp\{\frac{(\theta\sigma^2 t + \mu t)^2 - \mu^2 t^2}{2\sigma^2 t}\} = \exp\{\frac{\theta^2\sigma^4 t^2 + 2\theta\sigma^2 t\mu t}{2\sigma^2 t}\} \\
&= \exp\{\theta\mu t + \frac{1}{2}\theta^2\sigma^2 t\}
\end{aligned}
$$

$\square$

Now we are ready to prove the Brownian exponential defined by the equation (3) is a martingale.

Firstly, the process $(Z_{t\in[0,\infty)})$ is nonanticipating because a standard Brownian motion $(B_{t\in[0,\infty)})$ is nonanticipating.

Secondly, it satisfies the finite mean condition, since $E[Z_t] = 1 < \infty$:

$$
\begin{aligned}
E[Z_t] &= E[\exp(\theta B_t - \frac{1}{2}\theta^2 t)] \\
E[Z_t] &= E[\exp(\theta B_t)\exp(-\frac{1}{2}\theta^2 t)] \\
E[Z_t] &= \exp(-\frac{1}{2}\theta^2 t)E[\exp(\theta B_t)],
\end{aligned}
$$

using the proposition 4.1, $E[\exp(\theta B_t)] = \exp(\frac{1}{2}\theta^2 t)$:

$$
E[Z_t] = \exp(-\frac{1}{2}\theta^2 t)\exp(\frac{1}{2}\theta^2 t) = 1 .
$$

For $\forall 0 \le t \le t+h < \infty$, by the definition of Z_t:

$$E[Z_{t+h} | \mathcal{F}_t] = E[\exp\{\theta B_{t+h} - \frac{1}{2}\theta^2(t+h)\} | \mathcal{F}_t].$$

The trick is to multiply $\exp(\theta B_t - \theta B_t) = e^0 = 1$ inside the expectation operator:

$$E[Z_{t+h} | \mathcal{F}_t] = E[\exp(\theta B_t - \theta B_t)\exp\{\theta B_{t+h} - \frac{1}{2}\theta^2(t+h)\} | \mathcal{F}_t]$$
$$E[Z_{t+h} | \mathcal{F}_t] = E[\exp(\theta B_t)\exp(-\theta B_t)\exp(\theta B_{t+h})\exp(-\frac{1}{2}\theta^2 t)\exp(-\frac{1}{2}\theta^2 h) | \mathcal{F}_t]$$
$$E[Z_{t+h} | \mathcal{F}_t] = E[\exp(\theta B_t - \frac{1}{2}\theta^2 t)\exp\{\theta(B_{t+h} - B_t) - \frac{1}{2}\theta^2 h\} | \mathcal{F}_t].$$

Since Brownian increments are independent:

$$E[Z_{t+h} | \mathcal{F}_t] = E[\exp(\theta B_t - \frac{1}{2}\theta^2 t) | \mathcal{F}_t] E[\exp\{\theta(B_{t+h} - B_t) - \frac{1}{2}\theta^2 h\} | \mathcal{F}_t],$$

and since B_t is $\mathcal{F}_t$-adapted:

$$E[Z_{t+h} | \mathcal{F}_t] = \exp(\theta B_t - \frac{1}{2}\theta^2 t) E[\exp\{\theta(B_{t+h} - B_t) - \frac{1}{2}\theta^2 h\}].$$

By the definition of Z_t:

$$E[Z_{t+h} | \mathcal{F}_t] = Z_t E[\exp\{\theta(B_{t+h} - B_t)\}\exp(-\frac{1}{2}\theta^2 h)],$$

and since $\exp(-\frac{1}{2}\theta^2 h)$ is a constant:

$$E[Z_{t+h} | \mathcal{F}_t] = Z_t \exp(-\frac{1}{2}\theta^2 h) E[\exp\{\theta(B_{t+h} - B_t)\}].$$

Use the proposition 2.1 because $B_{t+h} - B_t \sim Normal(0, h)$:

$$E[Z_{t+h} | \mathcal{F}_t] = Z_t \exp(-\frac{1}{2}\theta^2 h)\exp(\frac{1}{2}\theta^2 h)$$
$$E[Z_{t+h} | \mathcal{F}_t] = Z_t.$$

□

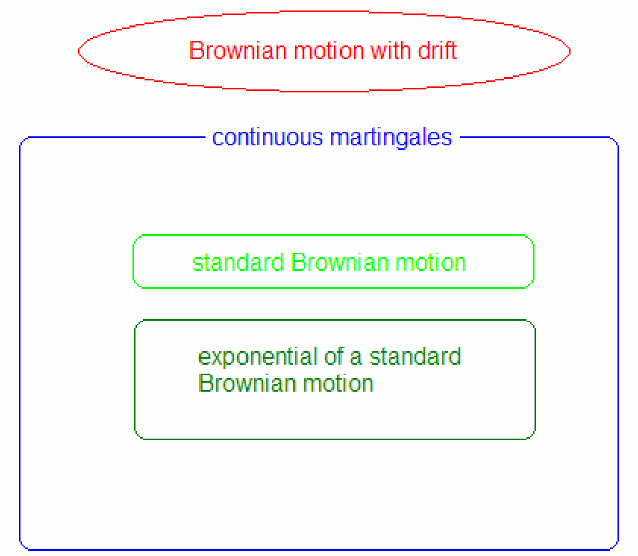


Figure 4.2 Brownian motion as a subclass of continuous martingales

[4.1.6] Brownian Motion as a Subclass of Markov Processes

For the background knowledge of Markov processes, read section 2.3.3 and 3.8.

Theorem 4.8 A standard Brownian motion process A standard Brownian motion $(B_{t\in[0,\infty)})$ defined on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,\infty)}, \mathbb{P})$ satisfies the followings:

(1) It is a time homogeneous Markov process. In other words, for any bounded Borel function $f:\mathbb{R}\to\mathbb{R}$ and for $\forall 0\le t\le u<\infty$:

$$E[f(B_u)|\mathcal{F}_t] = \mathbb{P}_{0,u-t} f(B_t) = \mathbb{P}_{u-t} f(B_t).$$

(2) Its transition function $\mathbb{P}_{u-t} \equiv \mathbb{P}_h$ is given by:

$$\mathbb{P}_h(x,y) = \frac{1}{\sqrt{2\pi h}} \exp\left\{-\frac{(x-y)^2}{2h}\right\}.$$

(3) $\mathbb{P}_h f(x) = \begin{cases} f(x) & \text{if} \quad h=0 \\ \int_{-\infty}^{\infty} \mathbb{P}_h(x,y) f(y)dy & \text{if} \quad h>0 \end{cases}.$

Proof

Markov property is a result of independent increments property of Brownian motion. Let $(B_{t\in[0,\infty)})$ be a standard Brownian motion defined on a filtered probability space

$(\Omega, \mathcal{F}_{t\in[0,\infty)}, \mathbb{P})$. Consider an increasing sequence of time $0 < t_1 < t_2 < ... < t_n < t < u < \infty$ where t is the present. As a result of independent increments condition:

$$\begin{aligned}
&\mathbb{P}(X_u - X_t \left| X_{t_1} - X_0, X_{t_2} - X_{t_1}, ..., X_t - X_{t_n} \right.) \\
&= \frac{\mathbb{P}(X_u - X_t \cap X_{t_1} - X_0, X_{t_2} - X_{t_1}, ..., X_t - X_{t_n})}{\mathbb{P}(X_{t_1} - X_0, X_{t_2} - X_{t_1}, ..., X_t - X_{t_n})} \\
&= \frac{\mathbb{P}(X_u - X_t)\mathbb{P}(X_{t_1} - X_0, X_{t_2} - X_{t_1}, ..., X_t - X_{t_n})}{\mathbb{P}(X_{t_1} - X_0, X_{t_2} - X_{t_1}, ..., X_t - X_{t_n})} \\
&= \mathbb{P}(X_u - X_t),
\end{aligned}$$

which means that there is no correlation (probabilistic dependence structure) on the increments among the past, the present, and the future.

Using the simple relationship $X_u \equiv (X_u - X_t) + X_t$ for an increasing sequence of time $0 < t_1 < t_2 < ... < t_n < t < u < \infty$:

$$\begin{aligned}
\mathbb{P}(X_u \left| X_0, X_{t_1}, X_{t_2}, ..., X_{t_n}, X_t \right.) &= \mathbb{P}((X_u - X_t) + X_t \left| X_0, X_{t_1}, X_{t_2}, ..., X_{t_n}, X_t \right.) \\
&= \mathbb{P}(X_u \left| X_t \right.),
\end{aligned}$$

which holds because an increment $(X_u - X_t)$ is independent of X_t by definition and the value of X_t depends on its realization $X_t(\omega)$.

□

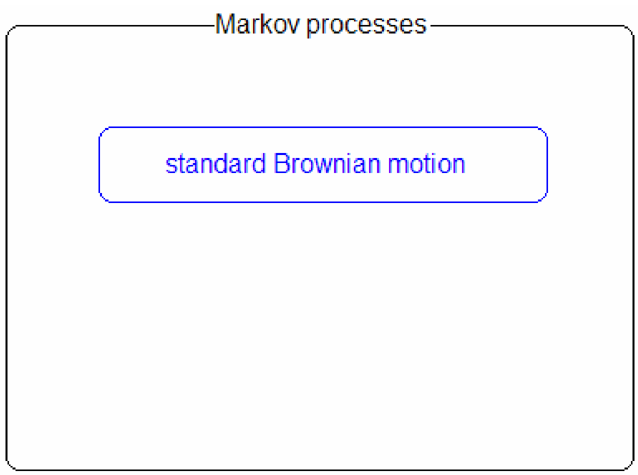


Figure 4.3 Brownian motion as a subclass of Markov processes

[4.2] Poisson Process

A Poisson process is a continuous time stochastic process with discontinuous sample paths. To be more precise, the sample paths of Poisson process is right continuous with left limit (i.e. rcll) step functions of jump size 1. It can be used as a building block for all Lévy processes.

[4.2.1] Exponential Random Variable

Definition 4.6 Exponential random variable An exponential random variable X with a parameter $\lambda \in \mathbb{R}^+$ is a positive random variable whose probability density function is given, for $x \in \mathbb{R}^+$, by:

$$f_X(x) = \lambda e^{-\lambda x}. \tag{5}$$

Its distribution function is for $x \in \mathbb{R}^+$:

$$F_X(x) = \Pr(X \le x) = 1 - e^{-\lambda x}. \tag{6}$$

Its mean and variance are:

$$E[X] = \frac{1}{\lambda} \text{ and } Var[X] = \frac{1}{\lambda^2}.$$

For example, the probability density function of an exponential random variable X with $\lambda = 0.01$ is plotted below.

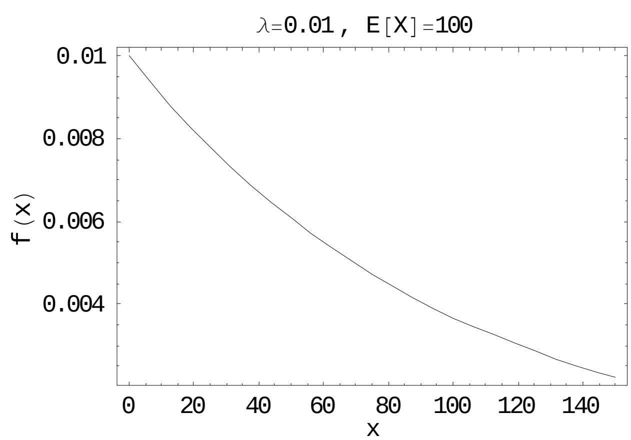


Figure 4.4 Plot of the probability density function of an exponential random variable with $\lambda = 0.01$

Theorem 4.9 Lack of memory of an exponential random variable If a random variable X is an exponential random variable, then, for $\forall a, b \in \mathbb{R}^+$:

$$\Pr\{X > a + b \mid X > b\} = \Pr\{X > a\}.$$

If X is a random arrival time of an event, the probability of $X > a+b$ given $X > b$ is the same as the probability of $X > a$. This concept of lack of memory has nothing to do with the concept of statistical independence. Let $A \equiv a+b$, $B \equiv X > b$, and $C \equiv X > a$. Then, the lack of memory property becomes:

$$\Pr\{A | B\} = \Pr\{C\}.$$

If two events A and B are independent:

$$\Pr\{A | B\} = \frac{\Pr\{A \cap B\}}{\Pr\{B\}} = \frac{\Pr\{A\}\Pr\{B\}}{\Pr\{B\}} = \Pr\{A\}.$$

Proof

$$\Pr\{X > a+b | X > b\} = \frac{\Pr\{X > a+b \cap X > b\}}{\Pr\{X > b\}} = \frac{\Pr\{X > a+b\}}{\Pr\{X > b\}}$$
$$\Pr\{X > a+b | X > b\} = \frac{1-F_X(a+b)}{1-F_X(b)} = \frac{1-(1-e^{-\lambda(a+b)})}{1-F_X(b)} = \frac{e^{-\lambda a}e^{-\lambda b}}{1-F_X(b)}$$
$$\Pr\{X > a+b | X > b\} = \frac{\{1-F_X(a)\}\{1-F_X(b)\}}{1-F_X(b)} = 1-F_X(a)$$
$$\Pr\{X > a+b | X > b\} = \Pr\{X > a\}$$

□

We will give one illustrating example. Suppose that the first arrival time of a large earthquake X is modeled as an exponential random variable with the mean of $1/\lambda = 100$ years. Using $a = 10$ years and $b = 40$ years, the lack of memory property becomes:

$$\Pr\{X > 50 | X > 40\} = \Pr\{X > 10\}.$$

This means that the probability that 40 years has past since the last large earthquake and there will be more than 10 years for the next large earthquake to hit (i.e. which is $\Pr\{X > 50 | X > 40\}$) is equal to the unconditional probability that there will be more than 10 years for the next large earthquake to hit (i.e. which is $\Pr\{X > 10\}$). In other words, an exponential random variable X does not remember the fact that the 40 years has past since the last large earthquake.

[4.2.2] Poisson Random Variable and Poisson Distribution

Definition 4.7 Poisson random variable A Poisson random variable N is a discrete random variable with a parameter $\lambda \in \mathbb{R}^+$ called an intensity whose probability mass function is given, for $\forall k \in \mathbb{N}$ (i.e. $k = 0,1,2,...$), by:

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

Its mean and variance are:

$$E[N]=\lambda \text{ and } Var[X]=\lambda.$$

A Poisson random variable N is used for the purpose of counting the number of arrivals of an event in unit time interval. For example, if we model the number of arrivals of small earthquakes in one year as a Poisson random variable with the intensity $\lambda=10$, its probability mass function is given by:

$$\mathbb{P}(N=k)=\frac{e^{-10}10^k}{k!},$$

which is plotted in Figure 4.5.

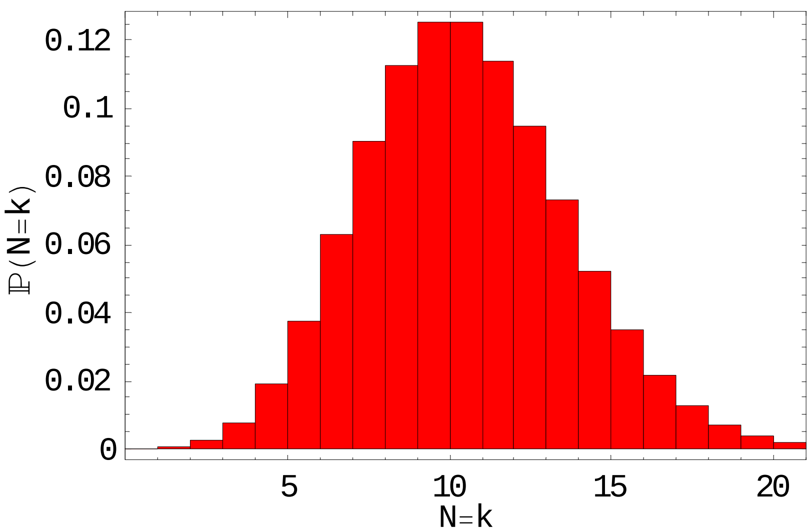


Figure 4.5 Plot of the probability mass function of a Poisson random variable with $\lambda=10$

[4.2.3] Relationship between the Sums of Independent Exponential Random Variables and Poisson Distribution

Suppose we model the number of dogs we see in one day N as a Poisson random variable with the intensity $\lambda\in\mathbb{R}^+$. Then, its probability mass function is:

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

Next, the number of dogs we see in t days N_t can be models as a Poisson random variable with the intensity $\lambda t \in \mathbb{R}^+$. Then, its probability mass function is:

$$\mathbb{P}(N_t = k) = \frac{e^{-\lambda t}(\lambda t)^k}{k!}. \qquad (7)$$

The intensity λt can be interpreted as the average number of arrivals of an event (i.e. seeing a dog in our case) in the time interval of $[0,t]$. Let T be the waiting time until the first arrival of an event. Then, the probability of zero arrival of an event in $[0,t]$ can be calculated as:

$$\mathbb{P}(T > t) = \mathbb{P}(N_t = 0) = e^{-\lambda t}. \qquad (8)$$

Figure 4.6 Zero arrival of an event in $[0,t]$

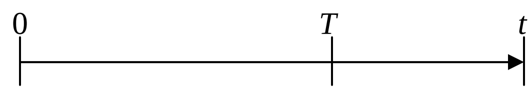


Figure 4.7 One arrival of an event in $[0,t]$

Using the equation (8), the distribution function of a waiting time T of the first arrival of an event in the interval $[0,t]$ can be expressed as:

$$F_T(t) = \mathbb{P}(T < t) = 1 - \mathbb{P}(T > t) = 1 - e^{-\lambda t}. \qquad (9)$$

As you can see, the equation (9) is identical to the equation (6) which means that a waiting time T of the first arrival of an event in the interval $[0,t]$ is an exponential random variable with the parameter λ.

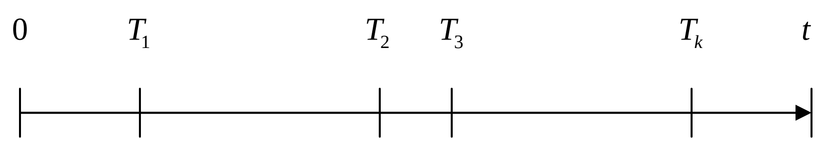


Figure 4.8 k **arrivals of an event in** $[0,t]$

This idea can be extended. Suppose that an event arrives k times in the interval $[0,t]$. Note in this case, T_i indicates the moment when i th event arrives. Waiting time of the arrival of the first event is $T_1 - 0 = T_1$ which means that the waiting time between the arrival of the second event and the first event can be expressed as $T_2 - T_1$ and the waiting time between the arrival of the i th event and the $(i-1)$ th event can be expressed as $T_i - T_{i-1}$. In this case, a Poisson random variable N_t with the intensity λt counts the number of arrival of an event in the interval $[0,t]$ which means that a Poison random variable is a discrete random variable or has discontinuous sample paths (i.e. $N_t \in \mathbb{N}$). And, each waiting time $T_i - T_{i-1}$ is an *i.i.d* exponential random variable with the parameter λ :

$$\mathbb{P}(T_i - T_{i-1}) = \lambda e^{-\lambda(T_i - T_{i-1})},$$

which indicates that $T_i - T_{i-1}$ is a continuous random variable because an event can arrive at any moment.

For example, suppose that the average number of dogs we see in one day (i.e. the intensity) is $\lambda = 12$:

$$E[N_1] = \lambda = 12 .$$

Then, each waiting time $T_i - T_{i-1}$ is an *i.i.d* exponential random variable with the parameter $\lambda = 12$ which means that the average waiting time to see a dog is 1/12 day which is equal to 2 hours:

$$E[T_i - T_{i-1}] = \frac{1}{\lambda} = \frac{1 \text{ day}}{12\text{dogs}} = \frac{2 \text{ hours}}{1 \text{ dog}} .$$

[4.2.4] Poisson Process

Definition 4.8 Poisson process A Poisson process $(N_{t\in[0,\infty]})$ with the intensity $\lambda \in \mathbb{R}^+$ is a stochastic process which counts the number of random times T_k (i.e. $T_1, T_2, ..., T_k$) of the arrival of an event in the time interval $[0,t]$ defined as:

$$N_t = \sum_{k\geq 1} 1_{t \geq T_k} ,$$

which means that the sample paths of $(N_{t\in[0,\infty]})$ are right continuous step functions of jump size equal to one (i.e. discontinuous sample paths). Its probability mass function follows a Poisson distribution with the parameter λt :

$$\mathbb{P}(N_t = k) = \frac{e^{-\lambda t}(\lambda t)^k}{k!} .$$

And, each waiting time $T_i - T_{i-1}$ is an *i.i.d* exponential random variable with the parameter λ:

$$\mathbb{P}(T_i - T_{i-1}) = \lambda \exp\{-\lambda(T_i - T_{i-1})\},$$

which indicates that a Poisson process $(N_{t\in[0,\infty]})$ is a continuous time stochastic process because an event can arrive at any moment.

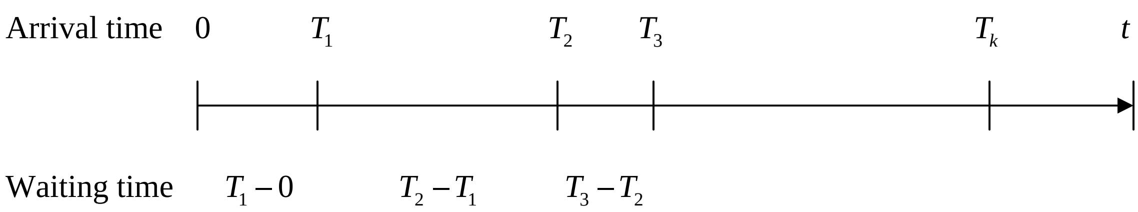


Figure 4.9 An event arrivals k times in the interval $[0,t]$

For more details, consult Karlin and Taylor (1975) pages 22-26.

[4.2.5] Properties of Poisson Process

We present the important properties of a Poisson process in this section. Some of which can be easily seen by the simulated sample paths of a Poisson process which is shown in Figure 4.10. Note that a path 1 has 10 jumps, a path 2 has 7 jumps, and a path 3 has 13 jumps.

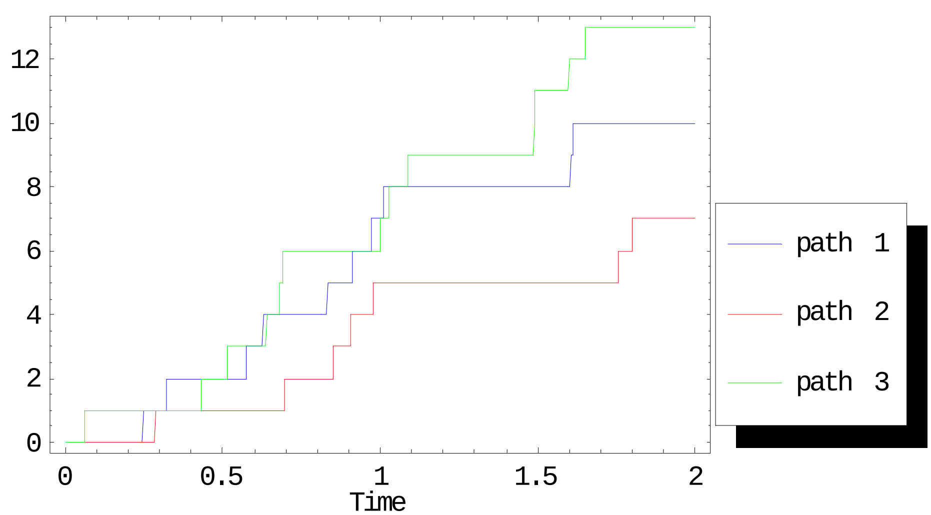


Figure 4.10 Simulated sample paths of Poisson processes with the intensity λ = 5 and $t \in [0, 2]$ **.**

Theorem 4.10 Fundamental properties of Poisson processes A Poisson process $(N_{t\in[0,\infty]})$ with the intensity $\lambda \in \mathbb{R}^+$ defined on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,\infty]}, \mathbb{P})$ satisfies the following conditions:

(1) Its increments are independent. In other words, for any increasing sequence of time $0 \le t_1 < t_2 < ... < t_n$:

$$\mathbb{P}(N_0 \cap N_{t_1} - N_0 \cap N_{t_2} - N_{t_1} \cap ... \cap N_{t_n} - N_{t_{n-1}})$$
$$= \mathbb{P}(N_0)\mathbb{P}(N_{t_1} - N_0)\mathbb{P}(N_{t_2} - N_{t_1})...\mathbb{P}(N_{t_n} - N_{t_{n-1}}) .$$

(2) Its increments are stationary (time homogeneous): i.e. for $h \ge 0$, $N_{t+h} - N_t$ has the same distribution as N_h . In other words, the distribution of increments does not depend on t .
(3) $\mathbb{P}(N_0 = 0) = 1$. The process starts from 0 almost surely (with probability 1).
(4) The process is stochastically continuous: $\forall \varepsilon > 0$, $\lim_{h \to 0} \mathbb{P}(|N_{t+h} - N_t| \ge \varepsilon) = 0$.
(5) Its sample paths are 1) non-decreasing functions, 2) right continuous with left limit step functions in $\forall t \in [0, \infty]$, and 3) its jump (step) size is 1. Obviously, these sample paths properties are true almost surely.
(6) For $\forall t \in (0, \infty]$, $\mathbb{P}(N_t = k < \infty) = 1$. In other words, the number of arrivals of an event is almost surely finite for any $t > 0$ including an infinite time horizon (i.e. $t = \infty$).

Proof

Theorems (1), (2), and (5) are true by definition. For more details, consult Ross (1983) chapter2 and Cont and Tankov (2004) pages 48-52.

[4.2.6] Poisson Process as a Lévy Process

As you may notice, it is very obvious that a Poisson process is a Lévy process because the condition (1) – (5) of theorem 4.10 is the definition of Lévy process (i.e. the definition 3.1).

Definition 4.9 Poisson process as a Lévy process A Poisson process $(N_{t\in[0,\infty]})$ with the intensity $\lambda \in \mathbb{R}^+$ defined on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,\infty]}, \mathbb{P})$ is a Lévy process satisfying:

(1) N_t follows a Poisson distribution with the intensity λt :

$$\mathbb{P}(N_t = k) = \frac{e^{-\lambda t}(\lambda t)^k}{k!}.$$

(2) Its sample paths are non-decreasing right continuous with left limit step functions of step size 1 in $\forall t \in [0,\infty]$ with probability 1.

The condition (1) implies the condition (2).

Theorem 4.11 Lévy process whose sample paths are non-decreasing right continuous with left limit step functions of step size 1 Let $(X_{t\in[0,\infty)})$ be a real valued Lévy process on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,\infty)}, \mathbb{P})$. If and only if the sample paths of $(X_{t\in[0,\infty)})$ are non-decreasing right continuous with left limit step functions of step size 1 in $\forall t \in [0,\infty]$ with probability 1, then, $(X_{t\in[0,\infty)})$ is a Poisson process.

Proof

Consult proposition 3.3 of Cont and Tankov (2004).

Definition 4.10 Poisson process in terms of Lévy triplet (A, ℓ, γ) Let $(X_{t\in[0,\infty)})$ be a real valued Lévy process with Lévy triplet (A, ℓ, γ) defined on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,\infty)}, \mathbb{P})$. Then, $(X_{t\in[0,\infty)})$ is a Poisson process with the intensity $\lambda \in \mathbb{R}^+$, if it satisfies the following conditions:

(1) $A = 0$. Its Gaussian variance is zero (i.e. Lévy-Itô decomposition).
(2) $\gamma = 0$. Its drift is zero (i.e. Lévy-Itô decomposition).
(3) Its Lévy measure is given by:

$$\ell(x) = \lambda\delta(x-1),$$

where $\delta(x-1)$ is a Dirac's delta function (see Appendix 1) satisfying:

$$\delta(x-1) = \begin{cases} \delta(0) & \text{if} \quad x = 1 \\ 0 & \text{otherwise} \end{cases}.$$

$\delta(x-1)$ is a pulse of unbounded height and zero width with a unit integral:

$$\int_{-\infty}^{\infty} \delta(x-1)dx = 1.$$

Note that a Dirac's delta function $\delta(x-1)$ is a jump size probability density function for Poisson processes because Poisson processes have only one type of jumps which are jumps of size 1. And not surprisingly, the integral of the Lévy measure of a Poisson process is the intensity parameter $\lambda \in \mathbb{R}^+$ because a Lévy measure $\ell(x)$ measures the arrival rate of jumps:

$$\int_{-\infty}^{\infty} \ell(x)dx = \int_{-\infty}^{\infty} \lambda\delta(x-1)dx = \lambda\int_{-\infty}^{\infty} \delta(x-1)dx = \lambda < \infty,$$

which is finite because the number of arrivals of an event is almost surely finite for any $t > 0$ including an infinite time horizon $t = \infty$ (the number (6) of theorem 4.10).

Theorem 4.12 Finite variation property of Poisson process If $(X_{t\in[0,T]})$ is a Poisson process defined on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,T]}, \mathbb{P})$, then, it is a real valued Lévy process of finite variation on the interval $[0,T]$. And, its Lévy triplet (A, ℓ, γ) satisfies:

$$A = 0 \text{ and } \int_{|x|<1} |x| \ell(dx) < \infty,$$

which are finite variation conditions for Lévy processes (i.e. theorem 3.7). Zero Gaussian variance $A = 0$ of a Poisson process is obvious and the second condition for a Poisson process is satisfied because there is no need to truncate jumps and jump sizes are all 1:

$$\int_{|x|<1} |x| \ell(dx) = \int_{x=1} x\lambda\delta(x-1)dx = \lambda\int_{x=1} x\delta(x-1)dx$$
$$\int_{|x|<1} |x| \ell(dx) = \lambda < \infty.$$

Finite variation property of a Poisson process can be easily guessed from its sample paths behavior. Stochastic processes of infinite variation have highly oscillatory sample paths such as a Brownian motion with drift. But, the sample paths of a Poisson process are rcll step functions of step size 1 which implies that a Poisson process is of finite variation:

$$\mathbb{P}(T(X) = \sup \sum_{i=1}^{n} \left| X(t_i) - X(t_{i-1}) \right| < \infty) = 1.$$

Table 4.2 summarizes the properties of a Poisson process.

Table 4.2 Poisson process

Lévy process	Gaussian variance	Lévy measure	drift	variation	sample path
Poisson process	$A = 0$	$\ell = \lambda\delta(x-1)$	$\gamma = 0$	finite	rcll step functions of step size 1

[4.2.7] Characteristic Function of Poisson Process

Consider a Poisson process $(X_{t\in[0,\infty]})$ defined on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,\infty]}, \mathbb{P})$ and let (A, ℓ, γ) be its Lévy triplet. Its characteristic function can be obtained by two approaches. First approach is the direct use of the definition of a characteristic function (i.e. Fourier transform of the probability density function with Fourier transform parameters $(1,1)$):

$$\phi_X(\omega) \equiv \mathcal{F}\left[\mathbb{P}(X_t = k)\right] \equiv \sum_{k=0}^{\infty} \frac{e^{-\lambda t}(\lambda t)^k}{k!} e^{i\omega k}$$
$$\phi_X(\omega) = \exp[t\lambda(e^{i\omega} - 1)] .$$

Note that in the above, a series expansion is used instead of the Fourier integral $\int_{-\infty}^{\infty} e^{i\omega x}\mathbb{P}(x)dx$ because a Poisson distribution is discrete.

Second approach is the use of Lévy-Khinchin representation (theorem 3.3). Because we know the Lévy triplet of a Poisson process is given by:

$$\left(A = 0, \ell = \lambda\delta(x-1), \gamma = 0\right),$$

its characteristic exponent is given by:

$$\psi_X(\omega) = -\frac{A\omega^2}{2} + i\gamma\omega + \int_{-\infty}^{\infty}\left\{\exp(i\omega x) - 1 - i\omega x 1_D\right\}\ell(dx)$$
$$\psi_X(\omega) = \int_{-\infty}^{\infty}\left\{\exp(i\omega x) - 1\right\}\lambda\delta(x-1)dx$$
$$\psi_X(\omega) = \lambda(\mathrm{e}^{i\omega} - 1) .$$

Note that the term $-i\omega x 1_D$ in the Lévy-Khinchin representation drops out for Poisson processes because there is no need to distinguish between large and small jumps for Poisson processes (i.e. all jumps sizes are 1). Thus, its characteristic function $\phi_X(\omega)$ is expressed as:

$$\phi_X(\omega) = \exp\left(t\psi_X(\omega)\right),$$
$$\phi_X(\omega) = \exp\{t\lambda(\mathrm{e}^{i\omega} - 1)\} .$$

[4.2.8] Lévy Measure of Poisson Process

Theorem 4.13 Lévy measure of Poisson process A Lévy measure of Poisson process $(X_{t\in[0,\infty]})$ with the intensity $\lambda \in \mathbb{R}^+$ defined on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,\infty]}, \mathbb{P})$ is given by:

$$\ell(x) = \lambda\delta(x-1) .$$

And, a Poisson process is a finite activity Lévy process because its Lévy measure satisfies:

$$\int_{-\infty}^{\infty} \ell(x)dx = \lambda < \infty .$$

Proof

The Lévy measure $\ell(x)$ of a Poisson process must satisfy two important conditions by the definition. One is that its jump size density is concentrated at $x=1$ and the other is that the average number of jumps per unit of time must equal to the intensity λ of a Poisson process. Thus, the only Lévy measure satisfying these two conditions is given by:

$$\ell(x) = \lambda\delta(x-1) ,$$

where $\delta(x-1)$ is a Dirac's delta function (see Appendix 1) satisfying:

$$\delta(x-1) = \begin{cases} \delta(0) & \text{if} \quad x=1 \\ 0 & \text{otherwise} \end{cases} .$$

$\delta(x-1)$ is a pulse of unbounded height and zero width with a unit integral:

$$\int_{-\infty}^{\infty} \delta(x-1)dx = 1 .$$

Note that a Dirac's delta function $\delta(x-1)$ is a jump size probability density function for Poisson processes because Poisson processes have only one type of jumps which are jumps of size 1. The integral of the Lévy measure of a Poisson process is the intensity parameter $\lambda \in \mathbb{R}^+$ because a Lévy measure $\ell(x)$ measures the arrival rate of jumps:

$$\int_{-\infty}^{\infty} \ell(x)dx = \int_{-\infty}^{\infty} \lambda\delta(x-1)dx = \lambda\int_{x=1} \delta(x-1)dx$$
$$\int_{-\infty}^{\infty} \ell(x)dx = \lambda < \infty ,$$

which is finite because the number of arrivals of an event is almost surely finite for any $t > 0$ including an infinite time horizon $t = \infty$ (the number (6) of theorem 4.10).

[4.2.9] Poisson Process as a Subclass of Markov Processes

Theorem 4.14 Poisson process as a time homogeneous and spatially homogeneous Markov process A Poisson process $(X_{t\in[0,\infty]})$ with the intensity $\lambda \in \mathbb{R}^+$ defined on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,\infty]}, \mathbb{P})$ is a Markov process with a time homogeneous and spatially homogeneous transition function.

Proof

For an increasing sequence of time $0 \le t_1 \le t_2 \le ... \le t \le t+h$:

$$\mathbb{P}(X_{t+h} - X_t = 1 \big| \mathcal{F}_t) = \mathbb{P}(X_{t+h} - X_t = 1 \big| X_0 = k_0, X_{t_1} = k_{t_1}, ..., X_t = k_t)$$
$$\mathbb{P}(X_{t+h} - X_t = 1 \big| \mathcal{F}_t) = \mathbb{P}(X_{t+h} - X_t = 1) = \mathbb{P}(X_h = 1) = \frac{e^{-\lambda h}(\lambda h)^1}{1!}$$
$$\mathbb{P}(X_{t+h} - X_t = 1 \big| \mathcal{F}_t) = e^{-\lambda h}\lambda h .$$

In the limit $h \downarrow 0$:

$$\lim_{h \downarrow 0} \mathbb{P}(X_{t+h} - X_t = 1 \big| \mathcal{F}_t) = \lambda h .$$

In other words, the transition function of a Poisson process is time homogeneous and spatially homogeneous:

$$\mathbb{P}_{t,t+h}(x_t, x_t + 1) = \mathbb{P}_{0,h}(0,1) = \lambda h .$$

□

[4.2.10] Poisson Process and Martingales: Compensated Poisson Process

Theorem 4.15 Nonmartingale property of Poisson process A Poisson process $(X_{t \in [0,\infty]})$ with the intensity $\lambda \in \mathbb{R}^+$ defined on a filtered probability space $(\Omega, \mathcal{F}_{t \in [0,\infty]}, \mathbb{P})$ is not a martingale.

Proof

A Poisson process has a nonanticipating X_t (i.e. X_t is $\mathcal{F}_t$-adapted) by definition. A Poisson process has a finite mean by definition, for $\forall t \in [0, \infty]$:

$$E[|X_t|] = E[X_t] = \lambda t < \infty ,$$

which is the number (6) of theorem 4.10. For $\forall 0 \le t \le u \le \infty$:

$$E[X_u \big| \mathcal{F}_t] = E[X_t + X_{u-t} \big| \mathcal{F}_t] = E[X_t \big| \mathcal{F}_t] + E[X_{u-t} \big| \mathcal{F}_t]$$
$$E[X_u \big| \mathcal{F}_t] = X_t + E[X_{u-t} \big| \mathcal{F}_t] = X_t + \lambda(u-t)$$
$$E[X_u \big| \mathcal{F}_t] \ne X_t .$$

□

Theorem 4.16 Martingale property of Compensated Poisson process A compensated Poisson process $(\tilde{X}_{t\in[0,\infty]}) \equiv (X_{t\in[0,\infty]} - \lambda t)$ with the intensity $\lambda \in \mathbb{R}^+$ defined on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,\infty]}, \mathbb{P})$ is a martingale.

Proof

A compensated Poisson process has a nonanticipating $\tilde{X}_t \equiv X_t - \lambda t$ (i.e. $\tilde{X}_t$ is $\mathcal{F}_t$-adapted) by definition. A compensated Poisson process has a finite mean by definition, for $\forall t \in [0, \infty]$:

$$E[|\tilde{X}_t|] = E[X_t - \lambda t] = E[X_t] - \lambda t = \lambda t - \lambda t = 0 < \infty .$$

For $\forall 0 \le t \le u \le \infty$:

$$\begin{aligned}
E[\tilde{X}_u | \mathcal{F}_t] &= E[\tilde{X}_t + \tilde{X}_{u-t} | \mathcal{F}_t] = E[\tilde{X}_t | \mathcal{F}_t] + E[\tilde{X}_{u-t} | \mathcal{F}_t] \\
E[\tilde{X}_u | \mathcal{F}_t] &= \tilde{X}_t + E[\tilde{X}_{u-t} | \mathcal{F}_t] = \tilde{X}_t + E[X_{u-t} - \lambda(u-t) | \mathcal{F}_t] \\
E[\tilde{X}_u | \mathcal{F}_t] &= \tilde{X}_t + E[X_{u-t} | \mathcal{F}_t] - \lambda(u-t) \\
E[\tilde{X}_u | \mathcal{F}_t] &= \tilde{X}_t + \lambda(u-t) - \lambda(u-t) = \tilde{X}_t .
\end{aligned}$$

□

A compensated Poisson process is not integer-valued and not a counting process because of the compensator λt. It behaves like a standard Brownian motion after rescaling it by $1/\sqrt{\lambda}$ since $\frac{1}{\sqrt{\lambda}}(\tilde{X}_{t\in[0,\infty]})$ satisfies:

$$E[\frac{\tilde{X}_t}{\sqrt{\lambda}}] = \frac{1}{\sqrt{\lambda}} E[X_t - \lambda t] = \frac{1}{\sqrt{\lambda}}(\lambda t - \lambda t) = 0$$

and:

$$\begin{aligned}
Var[\frac{1}{\sqrt{\lambda}} \tilde{X}_t] &= \frac{1}{\lambda} Var[\tilde{X}_t] = \frac{1}{\lambda} Var[X_t - \lambda t] = \frac{1}{\lambda} Var[X_t] \\
Var[\frac{1}{\sqrt{\lambda}} \tilde{X}_t] &= \frac{1}{\lambda} \lambda t = t .
\end{aligned}$$

Theorem 4.17 Property of Compensated Poisson process Consider a finite time horizon $t \in [0, T]$. In the limit as the intensity of a compensated Poisson process $(\tilde{X}_{t\in[0,T]}) \equiv (X_{t\in[0,T]} - \lambda t)$ approaches infinity:

$$\lambda \in \mathbb{R}^+ \uparrow \infty ,$$

a compensated Poisson process has an identical distribution to a standard Brownain motion:

$$\lim_{\lambda\uparrow\infty}(\tilde{X}_{t\in[0,T]}) \equiv (X_{t\in[0,T]} - \lambda t) \underset{=}{d} (B_{t\in[0,T]}) .$$

For more details, consult Cont and Tankov (2004) page 53.

[4.3] Compound Poisson Process

A compound Poisson process is a general case of a Poisson process. It is a continuous time stochastic process with discontinuous sample paths. But, unlike a Poisson process, a compound Poisson process is not necessarily an increasing process and jump size density can be of any type (thus, it is more general).

[4.3.1] Compound Poisson Process

Definition 4.11 Compound Poisson process A compound Poisson process $(X_{t\in[0,\infty]})$ with the intensity $\lambda \in \mathbb{R}^+$ is a stochastic process defined on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,\infty]}, \mathbb{P})$ which is the sum of $i.i.d.$ jumps X_i from the jump size density $f(x)$:

$$X_t = \sum_{i=1}^{N_t} X_i \text{ with } X_i \sim i.i.d. f(x) ,$$

where a Poisson process $(N_{t\in[0,\infty]})$ with the intensity $\lambda \in \mathbb{R}^+$ counts the number of random times T_k (i.e. $T_1, T_2, ..., T_k$) of the arrival of an event in the time interval $[0,t]$ defined as:

$$N_t = \sum_{k\geq 1} 1_{t\geq T_k} .$$

Note that a Poisson process $(N_{t\in[0,\infty]})$ and the jumps sizes $(X_i)_{i\geq 1}$ are assumed to be independent.

This means that if an event arrives $k \in \mathbb{R}^+$ times in the time interval $\forall t \in (0,\infty]$, i.e. $N_t = k$, then, a compound Poisson process is the sum of k $i.i.d.$ jumps X_i from the jump size density $f(x)$:

$$X_t = \sum_{i=1}^{N_t = k} X_i = X_1 + X_2 + ... + X_k .$$

Another point is that a Poisson process is considered as a compound Poisson process with the unit jump size $X_i = 1$, since:

$$X_t = \sum_{i=1}^{N_t} X_i \text{ with } X_i = 1$$
$$X_t = \sum_{i=1}^{N_t} 1 = N_t = \sum_{k \geq 1} 1_{t \geq T_k} .$$

[4.3.2] Properties of Compound Poisson Process

We present the important properties of a compound Poisson process in this section. Some of which can be easily seen by its simulated sample paths illustrated by Figure 4.11. In this figure, the jump size density $f(x)$ is a standard normal:

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{x^2}{2}).$$

Note that a path 1 has 7 jumps, a path 2 has 10 jumps, and a path 3 has 11 jumps.

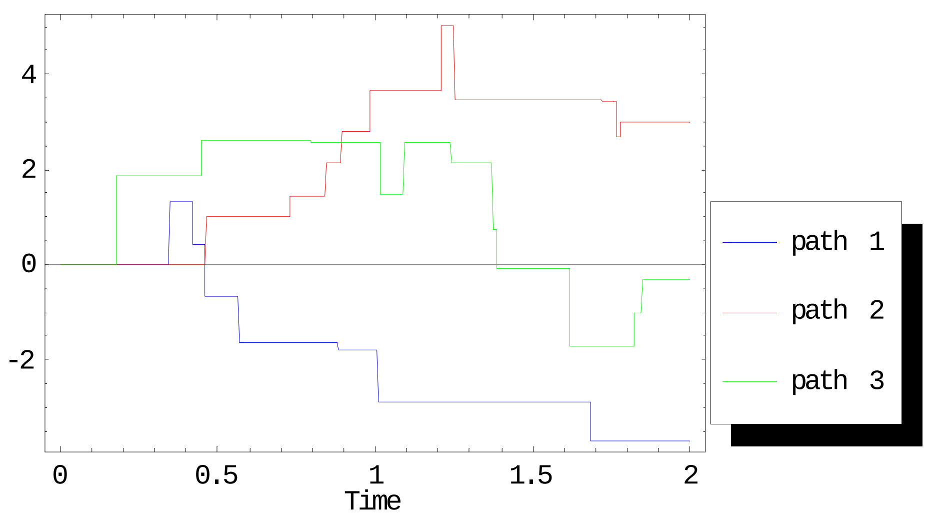


Figure 4.11 Simulated sample paths of a compound Poisson process with the intensity λ = 5 and $t \in [0, 2]$.

Theorem 4.18 Fundamental properties of compound Poisson processes A compound Poisson process $(X_{t\in[0,\infty]})$ with the intensity $\lambda \in \mathbb{R}^+$ defined on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,\infty]}, \mathbb{P})$ satisfies the following conditions:

(1) Its increments are independent. In other words, for any increasing sequence of time $0 \le t_1 < t_2 < ... < t_n$:

$$\mathbb{P}(X_0 \cap X_{t_1} - X_0 \cap X_{t_2} - X_{t_1} \cap ... \cap X_{t_n} - X_{t_{n-1}})$$
$$= \mathbb{P}(X_0)\mathbb{P}(X_{t_1} - X_0)\mathbb{P}(X_{t_2} - X_{t_1})...\mathbb{P}(X_{t_n} - X_{t_{n-1}}).$$

(2) Its increments are stationary (time homogeneous): i.e. for $h \ge 0$, $X_{t+h} - X_t$ has the same distribution as X_h. In other words, the distribution of increments does not depend on t.

(3) $\mathbb{P}(X_0 = 0) = 1$. The process starts from 0 almost surely (with probability 1).

(4) The process is stochastically continuous: $\forall \varepsilon > 0$, $\lim_{h \to 0} \mathbb{P}(|X_{t+h} - X_t| \ge \varepsilon) = 0$.

(5) Its sample paths are right continuous with left limit step functions in $\forall t \in [0, \infty]$ and the jump sizes X_i are $i.i.d.$ random variables from a density $f(x)$.

(6) For $\forall t \in (0, \infty]$, $\mathbb{P}(N_t = k < \infty) = 1$. In other words, the number of arrivals of an event is almost surely finite for any $t > 0$ including an infinite time horizon (i.e. $t = \infty$).

Proof

Similar to the proof of theorem 4.10.

[4.3.3] Compound Poisson Process as a Lévy Process

As you may notice, it is very obvious that a compound Poisson process is a Lévy process because the condition (1) – (5) of theorem 4.18 is the definition of Lévy process (i.e. the definition 3.1).

Definition 4.12 Compound Poisson process as a Lévy process A Compound Poisson process $(X_{t \in [0,\infty]}) \equiv (\sum_{i=1}^{N_t} X_i)$ with the intensity $\lambda \in \mathbb{R}^+$ defined on a filtered probability space $(\Omega, \mathcal{F}_{t \in [0,\infty]}, \mathbb{P})$ is a Lévy process satisfying:

(1) A counter N_t follows a Poisson distribution with the intensity λt:

$$\mathbb{P}(N_t = k) = \frac{e^{-\lambda t}(\lambda t)^k}{k!}.$$

(2) The jump sizes are $i.i.d.$ random variables from a density $f(x)$:

$$X_i \sim i.i.d. f(x).$$

(3) Its sample paths are right continuous with left limit step functions.

The condition (3) is implied by the conditions (2) and (3).

Theorem 4.19 Lévy process whose sample paths are right continuous with left limit step functions Let $(X_{t\in[0,\infty)})$ be a real valued Lévy process on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,\infty)}, \mathbb{P})$. If and only if the sample paths of $(X_{t\in[0,\infty)})$ are right continuous with left limit step functions in $\forall t \in [0,\infty]$ with probability 1, then, $(X_{t\in[0,\infty)})$ is a compound Poisson process.

Proof

Consult proposition 3.3 of Cont and Tankov (2004).

Definition 4.13 Compound Poisson process in terms of Lévy triplet (A, ℓ, γ) Let $(X_{t\in[0,\infty)})$ be a real valued Lévy process with Lévy triplet (A, ℓ, γ) defined on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,\infty)}, \mathbb{P})$. Then, $(X_{t\in[0,\infty)})$ is a compound Poisson process with the intensity $\lambda \in \mathbb{R}^+$, if it satisfies the following conditions:

(1) $A = 0$. Its Gaussian variance is zero (i.e. Lévy-Itô decomposition).
(2) $\gamma = 0$. Its drift is zero (i.e. Lévy-Itô decomposition).
(3) Its Lévy measure is given by:

$$\ell(x) = \lambda f(x),$$

where $f(x)$ is a jump size density satisfying:

$$\int_{-\infty}^{\infty} f(x)dx = 1,$$

which holds because $f(x)$ is a probability density function. The integral of the Lévy measure of a compound Poisson process is the intensity parameter $\lambda \in \mathbb{R}^+$ because a Lévy measure $\ell(x)$ measures the arrival rate of jumps:

$$\int_{-\infty}^{\infty} \ell(x)dx = \int_{-\infty}^{\infty} \lambda f(x)dx = \lambda \int_{-\infty}^{\infty} f(x)dx = \lambda < \infty,$$

which is finite because the number of arrivals of an event is almost surely finite for any $t > 0$ including an infinite time horizon $t = \infty$ (the number (6) of theorem 4.10).

Theorem 4.20 Finite variation property of Compound Poisson process If $(X_{t\in[0,T]})$ is a compound Poisson process defined on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,T]}, \mathbb{P})$, then, it is a real valued Lévy process of finite variation on the interval $[0,T]$. And, its Lévy triplet (A,ℓ,γ) satisfies:

$$A=0 \text{ and } \int_{|x|<1} |x|\,\ell(dx)<\infty,$$

which are finite variation conditions for Lévy processes (i.e. theorem 3.7). Finite variation property of a compound Poisson process can be easily guessed from its sample paths behavior. Stochastic processes of infinite variation have highly oscillatory sample paths such as a Brownian motion with drift. But, the sample paths of a compound Poisson process are rcll step functions which imply that a compound Poisson process is of finite variation:

$$\mathbb{P}(T(X)=\sup\sum_{i=1}^{n}\left|X(t_i)-X(t_{i-1})\right|<\infty)=1.$$

Table 4.3 summarizes the properties of a compound Poisson process.

Table 4.3 Compound Poisson process

Lévy process	Gaussian variance	Lévy measure	drift	variation	sample path
Compound Poisson process	$A=0$	$\ell=\lambda f(x)$	$\gamma=0$	finite	rcll step functions

[4.3.4] Characteristic Function of Compound Poisson Process

Theorem 4.21 (with the proof) Characteristic function of Compound Poisson process Consider a compound Poisson process $(X_{t\in[0,\infty]})$ with the intensity $\lambda\in\mathbb{R}^+$ defined on a filtered probability space $(\Omega,\mathcal{F}_{t\in[0,\infty]},\mathbb{P})$ and let (A,ℓ,γ) be its Lévy triplet. Individual jumps X_i of a compound Poisson process are $i.i.d.$ random variables from a density $f(x)$, i.e. $X_i \sim i.i.d. f(x)$. Let ϕ_f be a characteristic function of a jump size density:

$$\phi_f(\omega)\equiv\mathcal{F}\left[f(x)\right]\equiv\int_{-\infty}^{\infty}e^{i\omega x}f(x)dx=E[e^{i\omega x}].$$

Then, using the Lévy-Khinchin representation (theorem 3.3) with the Lévy triplet of a compound Poisson process $\left(A=0,\ell=\lambda f(x),\gamma=0\right)$, the characteristic exponent of a compound Poisson process is given by:

$$\psi_X(\omega)=\int_{-\infty}^{\infty}\left\{\exp(i\omega x)-1\right\}\lambda f(x)dx=\lambda\int_{-\infty}^{\infty}\left\{f(x)\exp(i\omega x)-f(x)\right\}dx$$
$$\psi_X(\omega)=\lambda\{\int_{-\infty}^{\infty}f(x)\exp(i\omega x)dx-\int_{-\infty}^{\infty}f(x)dx\}$$
$$\psi_X(\omega)=\lambda\{\phi_f(\omega)-1\}.$$

Thus, the characteristic function of a compound Poisson process $\phi_X(\omega)$ is expressed as:

$$\phi_X(\omega) = \exp\left(t\psi_X(\omega)\right)$$
$$\phi_X(\omega) = \exp\left[t\lambda\{\phi_f(\omega) - 1\}\right].$$

For more information regarding the characteristic function of a compound Poisson process, consult Cont and Tankov (2004) pages 74-75 and Sato (1999) pages 18-21:

[4.3.5] Lévy Measure of a Compound Poisson Process

Theorem 4.22 Lévy measure of compound Poisson process A Lévy measure of a compound Poisson process $(X_{t\in[0,\infty]})$ with the intensity $\lambda \in \mathbb{R}^+$ defined on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,\infty]}, \mathbb{P})$ is given by:

$$\ell(x) = \lambda f(x),$$

where individual jumps X_i of a compound Poisson process are $i.i.d.$ random variables from a density $f(x)$, i.e. $X_i \sim i.i.d. f(x)$.

A compound Poisson process is a finite activity Lévy process because the integral of the Lévy measure of a compound Poisson process is the intensity parameter $\lambda \in \mathbb{R}^+$:

$$\int_{-\infty}^{\infty} \ell(x)dx = \int_{-\infty}^{\infty} \lambda f(x)dx = \lambda \int_{-\infty}^{\infty} f(x)dx = \lambda < \infty,$$

which is finite because the number of arrivals of an event is almost surely finite for any $t > 0$ including an infinite time horizon $t = \infty$ (the number (6) of theorem 4.10).

[5] Stable Processes

In this section we present stable processes which are a subclass of Lévy processes. But since stable processes are very important, we decided to give it an independent section.

[5.1] Stable Distributions and Stable Processes

Definition 5.1 Strictly stable distribution Let X be an infinitely divisible random variable on $\mathbb{R}$ and $\phi_X(\omega)$ be its characteristic function. Then, X is said to have a strictly stable distribution, if its characteristic function satisfies, for $\forall a > 0$, $\forall b(a) > 0$, and $\forall \omega \in \mathbb{R}$:

$$\phi_X(\omega)^a = \phi_X(\omega b(a)) .$$

Definition 5.2 Stable distribution Let X be an infinitely divisible random variable on $\mathbb{R}$ and $\phi_X(\omega)$ be its characteristic function. Then, X is said to have a stable distribution, if its characteristic function satisfies, for $\forall a > 0$, $\forall b(a) > 0$, $c(a) \in \mathbb{R}$, and $\forall \omega \in \mathbb{R}$:

$$\phi_X(\omega)^a = \phi_X\left(\omega b(a)\right)\exp\left(ic(a)\omega\right) .$$

Example 5.1 Normal distribution as a stable distribution Consider a normal random variable Y with the mean $\forall \mu \in \mathbb{R}$ and variance $\forall \sigma^2 \in \mathbb{R}^+$, i.e. $Y \sim N(\mu, \sigma^2)$.
Then, Y has a stable distribution with $b(a) = a^{1/2}$ and $c(a) = (-a^{1/2} + a)\mu$.

Proof

The characteristic function of a normal random variable Y is, for $\forall \omega \in \mathbb{R}$:

$$\phi_Y(\omega) = \exp(i\mu\omega - \sigma^2\omega^2 / 2) .$$

Then, Y has a stable distribution since it satisfies the definition 5.2 with $b(a) = a^{1/2}$ and $c(a) = (-a^{1/2} + a)\mu$:

$$\begin{aligned}
\phi_Y(\omega)^a &= \{\exp(i\mu\omega - \sigma^2\omega^2 / 2)\}^a = \exp(i\mu\omega a - \sigma^2\omega^2 a / 2) \\
\phi_Y(\omega)^a &= \exp(i\mu\omega a - \sigma^2\omega^2 a / 2)\exp(i\mu\omega\sqrt{a} - i\mu\omega\sqrt{a}) \\
\phi_Y(\omega)^a &= \exp\{i\mu\omega\sqrt{a} - \sigma^2(\omega\sqrt{a})^2 / 2\}\exp(i\mu\omega a - i\mu\omega\sqrt{a}) \\
\phi_Y(\omega)^a &= \phi_X(\omega\sqrt{a})\exp\{i(a - \sqrt{a})\mu\omega\} .
\end{aligned}$$

□

Example 5.2 Zero Mean normal distribution as a strictly stable distribution
Consider a normal random variable Y with the mean $\mu = 0$ and variance $\forall \sigma^2 \in \mathbb{R}^+$, i.e. $Y \sim N(0, \sigma^2)$. Then, Y has a strictly stable distribution with $b(a) = a^{1/2}$.

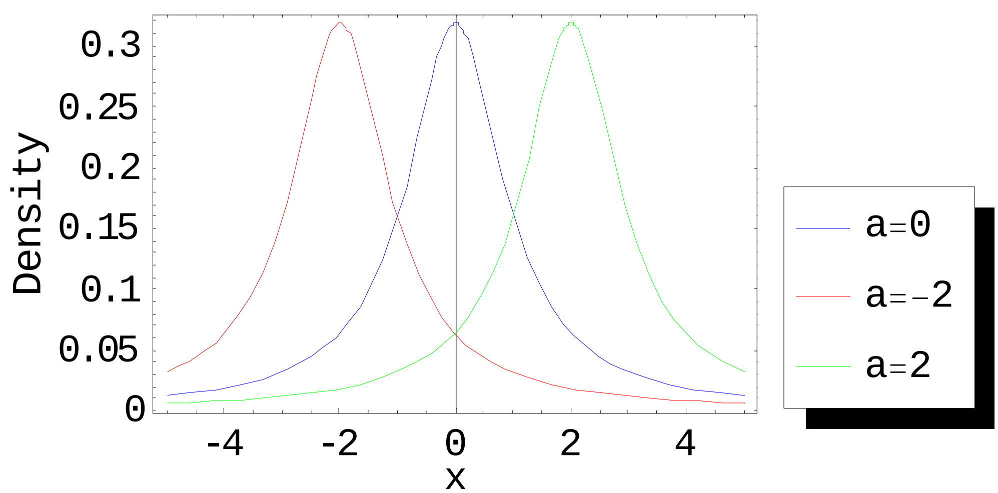
Proof

Apply the previous proof.

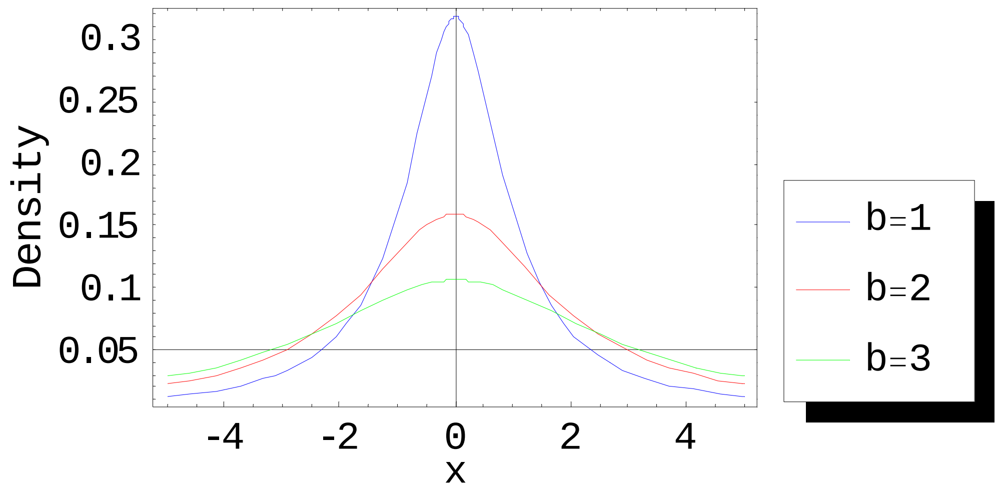
Definition 5.3 Cauchy distribution A Cauchy distribution is a continuous probability distribution whose probability density function is given by, for $\forall x \in \mathbb{R}$:

$$f(x) = \frac{1}{\pi} \frac{b}{b^2 + (x-a)^2},$$

where $a \in \mathbb{R}$ is called a location parameter which determines the location of the peak of the density as illustrated by the Panel A of Figure 5.1 and $b \in \mathbb{R}^+$ is called a scale parameter which influences the fatness of the tail of the density as illustrated by the Panel B of Figure 5.1.



A) Role of a location parameter $a \in \mathbb{R}$. A scale parameter b is set to 1.



B) Role of a scale parameter $b \in \mathbb{R}^+$. A location parameter a is set to 0.

Figure 5.1 Plot of Cauchy distribution

When $a = 0$ and $b = 1$, the Cauchy probability density function becomes:

$$f(x) = \frac{1}{\pi} \frac{1}{1 + x^2},$$

which is called a standard Cauchy distribution.

Theorem 5.1 Properties of Cauchy distribution Consider a Cauchy distribution whose probability distribution function is given by, for $\forall a \in \mathbb{R}$, $\forall b \in \mathbb{R}^+$, and $\forall x \in \mathbb{R}$:

$$f(x) = \frac{1}{\pi} \frac{b}{b^2 + (x - a)^2}.$$

$f(x)$ has the following properties:

(1) It has a unit integral because it is a probability measure:

$$\int_{-\infty}^{\infty} f(x) dx = 1.$$

(2) $f(x)$ is a positive measure for $\forall x \in \mathbb{R}$ because it is a probability measure:

$$f(x) \geq 0.$$

(3) Its mean is undefined and its variance is infinite. Thus, higher order moments are undefined as well.
(4) Its mode and median are equal to a.
(5) Its characteristic function is calculated as:

$$\phi_X(\omega) \equiv \int_{-\infty}^{\infty} e^{i\omega x} f(x) dx = \int_{-\infty}^{\infty} e^{i\omega x} \frac{1}{\pi} \frac{b}{b^2 + (x-a)^2} dx$$
$$\phi_X(\omega) = \exp(ia\omega - b|\omega|) .$$

(6) It is an infinitely divisible distribution.

Proof of (6)

$$\phi_X(\omega)^{1/n} = \left(\exp(ia\omega - b|\omega|)\right)^{1/n} = \exp\left(\frac{ia\omega}{n} - \frac{b|\omega|}{n}\right)$$
$$\phi_X(\omega)^{1/n} = \exp\left(i(\frac{a}{n})\omega - (\frac{b}{n})|\omega|\right).$$

This means that a Cauchy random variable X with parameters a and b has an identical distribution as a sum of n $i.i.d.$ Cauchy random variables each with the location parameter a/n and the scale parameter b/n. Take a look at proposition 3.1.

□

Example 5.3 Cauchy distribution as a strictly stable distribution Consider a Cauchy random variable Y with the location parameter $c \in \mathbb{R}$ and the scale parameter $d \in \mathbb{R}^+$ whose probability density function is given by:

$$f(x) = \frac{1}{\pi} \frac{d}{d^2 + (x-c)^2} .$$

Then, Y has a strictly stable distribution with $b = a$.

Proof

The characteristic function of a Cauchy random variable Y is, for $\forall \omega \in \mathbb{R}$:

$$\phi_Y(\omega) = \exp(ic\omega - d|\omega|) .$$

Then, Y has a stable distribution since it satisfies the definition 5.2 with $b = a$:

$$\phi_Y(\omega)^a = \left(\exp(ic\omega - d|\omega|)\right)^a = \exp(ic\omega a - d|\omega| a)$$
$$\phi_Y(\omega)^a = \exp\left(ic(\omega a) - d(|\omega| a)\right)$$

$$\phi_Y(\omega)^a = \phi_Y(\omega a)$$

□

We can extend this idea from single infinitely divisible random variable X on $\mathbb{R}$ to a real valued Lévy process $(X_{t\in[0,\infty)})$ defined on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,\infty)}, \mathbb{P})$ because of the proposition 3.2 and 3.3 which proposes one to one relationship between an infinitely divisible distribution and a Lévy process.

Definition 5.3 Strictly stable process A strictly stable process $(X_{t\in[0,\infty)})$ is a real valued stochastic process defined on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,\infty)}, \mathbb{P})$ satisfying the following conditions:

(1) $(X_{t\in[0,\infty)})$ is a Lévy process.
(2) The distribution of X_t at $t=1$ is a strictly stable distribution.

Definition 5.4 Stable process A stable process $(X_{t\in[0,\infty)})$ is a real valued stochastic process defined on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,\infty)}, \mathbb{P})$ satisfying the following conditions:

(1) $(X_{t\in[0,\infty)})$ is a Lévy process.
(2) The distribution of X_t at $t=1$ is a stable distribution.

It is apparent, but a standard Brownian motion is an example of a strictly stable process and a Brownian motion with drift is an example of a stable process.

[5.2] Selfsimilar and Broad-Sense Selfsimilar Stochastic Processes

Definition 5.5 Selfsimilar stochastic processes A real valued stochastic process $(X_{t\in[0,\infty)})$ defined on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,\infty)}, \mathbb{P})$ is said to be selfsimilar if, for $\forall a \in \mathbb{R}^+$ and $\forall b \in \mathbb{R}^+$, the process satisfies:

$$(X_{at\in[0,\infty)}) \underline{\underline{d}} (bX_{t\in[0,\infty)}) .$$

This means that for selfsimilar stochastic processes, a change in the time domain is equivalent to a change in the spatial domain in terms of the distributional property.

Definition 5.6 Broad-sense selfsimilar stochastic processes A real valued stochastic process $(X_{t\in[0,\infty)})$ defined on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,\infty)}, \mathbb{P})$ is said to be broad-sense selfsimilar if, for $\forall a \in \mathbb{R}^+$, $\forall b \in \mathbb{R}^+$, and $c(t):[0,\infty) \to \mathbb{R}$, the process satisfies:

$$(X_{at\in[0,\infty)}) \underset{=}{d} (bX_{t\in[0,\infty)} + c(t)) .$$

Example 5.4 Standard Brownian motion as a selfsimilar stochastic process A standard Brownian motion $(B_{t\in[0,\infty)})$ defined on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,\infty)}, \mathbb{P})$ is a selfsimilar stochastic process with $b = a^{1/2}$.

Proof

We know:

$$B_{at} \sim Normal(0, at)$$
$$\sqrt{a} B_t \sim Normal(0, at) .$$

Therefore:

$$(B_{at\in[0,\infty)}) \underset{=}{d} (\sqrt{a} B_{t\in[0,\infty)}) .$$

□

Example 5.5 Brownian motion with drift as a broad-sense selfsimilar stochastic process A Brownian motion with drift $(X_{t\in[0,\infty)}) \equiv (\mu t + \sigma B_{t\in[0,\infty)})$ defined on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,\infty)}, \mathbb{P})$ is a broad-sense selfsimilar stochastic process with $b = \sigma a^{1/2}$ and $c(t) = -\mu t \sigma \sqrt{a} + \mu a t$.

Proof

We know:

$$X_t \sim Normal(\mu t, \sigma^2 t)$$
$$X_{at} \sim Normal(\mu a t, \sigma^2 a t)$$
$$\sigma \sqrt{a} X_t \sim Normal(\mu t \sigma \sqrt{a}, \sigma^2 a t) .$$

Therefore:

$$(X_{at\in[0,\infty)}) \underset{=}{d} (\sigma \sqrt{a} X_{t\in[0,\infty)} - \mu t \sigma \sqrt{a} + \mu a t) .$$

□

[5.3] Relationship between Stability and Broad-Sense Selfsimilarity for Lévy Processes

Theorem 5.2 Strict stability and selfsimilarity for Lévy Processes A real valued Lévy process $(X_{t\in[0,\infty)})$ defined on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,\infty)}, \mathbb{P})$ is selfsimilar, if and only if $(X_{t\in[0,\infty)})$ is strictly stable. Its converse is also true. A real valued Lévy process $(X_{t\in[0,\infty)})$ is strictly stable, if and only if $(X_{t\in[0,\infty)})$ is selfsimilar.

Theorem 5.3 Stability and broad-sense selfsimilarity for Lévy Processes A real valued Lévy process $(X_{t\in[0,\infty)})$ defined on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,\infty)}, \mathbb{P})$ is broad-sense selfsimilar, if and only if $(X_{t\in[0,\infty)})$ is stable. Its converse is also true. A real valued Lévy process $(X_{t\in[0,\infty)})$ is stable, if and only if $(X_{t\in[0,\infty)})$ is broad-sense selfsimilar.

Proof

Consult Sato (1999) page 71.

Apparently, theorems 5.2 and 5.3 indicate that for Lévy processes, selfsimilarity is identical to strict stability and broad-sense selfsimilarity is identical to stability.

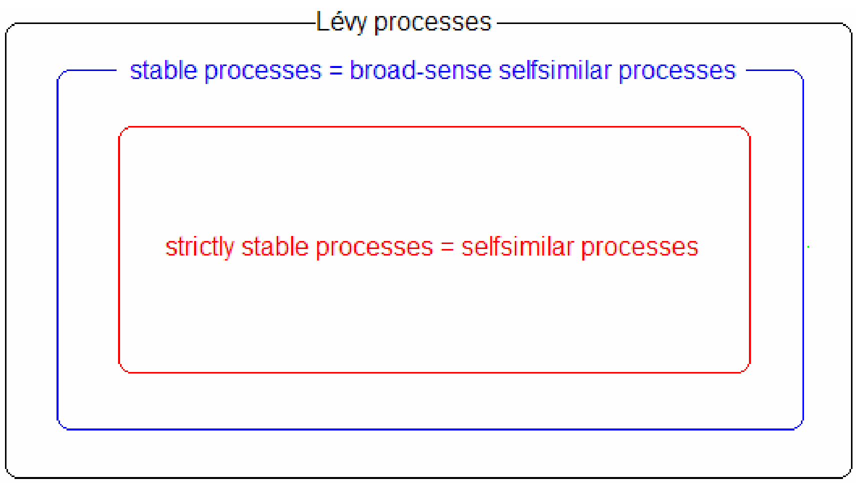


Figure 5.1 Relationship among Lévy processes, stable processes, and broad-sense selfsimilar processes

[5.4] More on Stable Processes

[5.4.1] Stability Index α

Theorem 5.4 Stability Index α
(1) For every stable distribution, there exist a constant called a stability index $\alpha \in (0,2]$ satisfying, for $\forall a > 0$, $\forall b(a) > 0$, $c(a) \in \mathbb{R}$, and $\forall \omega \in \mathbb{R}$:

$$\phi_X(\omega)^a = \phi_X\left(\omega b(a)\right)\exp\left(ic(a)\omega\right)$$
$$\phi_X(\omega)^a = \phi_X(\omega a^{1/\alpha})\exp\left(ic(a)\omega\right),$$

in other words:

$$b(a) = a^{1/\alpha} .$$

A stable distribution with the stability index $\alpha \in (0,2]$ is said to be α -stable distribution.
(2) For every stable process $(X_{t\in[0,\infty)})$, there exist a constant called a stability index $\alpha \in (0,2]$ satisfying, for $\forall a \in \mathbb{R}^{+}$, $\forall b(a) \in \mathbb{R}^{+}$, and $c(t):[0,\infty) \to \mathbb{R}$, and $\forall \omega \in \mathbb{R}$:

$$(X_{at\in[0,\infty)}) \underline{\underline{d}} \ (b(a)X_{t\in[0,\infty)} + c(t))$$
$$(X_{at\in[0,\infty)}) \underline{\underline{d}} \ (a^{1/\alpha} X_{t\in[0,\infty)} + c(t)) ,$$

in other words:

$$b(a) = a^{1/\alpha} .$$

A stable process $(X_{t\in[0,\infty)})$ with the stability index $\alpha \in (0,2]$ is said to be α -stable process.

Proof

Consult Sato (1999) pages 75-76.

Example 5.6 Normal distribution as the only 2-stable distribution Consider a normal random variable Y with the mean $\forall \mu \in \mathbb{R}$ and variance $\forall \sigma^2 \in \mathbb{R}^{+}$, i.e. $Y \sim N(\mu, \sigma^2)$. Then, Y is the only 2-stable distribution with $b(a) = a^{1/2}$ and $c(a) = (-a^{1/2} + a)\mu$.

Proof

The characteristic function of a normal random variable Y is, for $\forall \omega \in \mathbb{R}$:

$$\phi_Y(\omega) = \exp(i\mu\omega - \sigma^2\omega^2 / 2) .$$

Then, Y is a stable distribution with the stability index $\alpha = 2$, since it satisfies the theorem 5.4 with $b(a) = a^{1/2}$ and $c(a) = (-a^{1/2} + a)\mu$:

$$\phi_Y(\omega)^a = \{\exp(i\mu\omega - \sigma^2\omega^2 / 2)\}^a = \exp(i\mu\omega a - \sigma^2\omega^2 a / 2)$$
$$\phi_Y(\omega)^a = \exp(i\mu\omega a - \sigma^2\omega^2 a / 2)\exp(i\mu\omega\sqrt{a} - i\mu\omega\sqrt{a})$$
$$\phi_Y(\omega)^a = \exp\{i\mu\omega\sqrt{a} - \sigma^2(\omega\sqrt{a})^2 / 2\}\exp(i\mu\omega a - i\mu\omega\sqrt{a})$$
$$\phi_Y(\omega)^a = \phi_X(\omega\sqrt{a})\exp\{i(a - \sqrt{a})\mu\omega\} .$$

□

Theorem 5.5 Normal distribution as the only 2-stable distribution A real valued infinitely divisible probability density function $\mathbb{P}(x)$ is normal, if $\mathbb{P}(x)$ is 2-stable. Its converse is also true. A real valued infinitely divisible probability density function $\mathbb{P}(x)$ is 2-stable, if $\mathbb{P}(x)$ is normal.

Theorem 5.6 Zero Mean Normal distribution as the only strictly 2-stable distribution A real valued infinitely divisible probability density function $\mathbb{P}(x)$ is normal with zero mean, if $\mathbb{P}(x)$ is strictly 2-stable. Its converse is also true. A real valued infinitely divisible probability density function $\mathbb{P}(x)$ is strictly 2-stable, if $\mathbb{P}(x)$ is normal with zero mean.

Proof

These correspond to theorems 14.1 and 14.2 of Sato (1999) where proofs are provided.

Example 5.7 Brownian motion with drift as the only 2-stable process A Brownian motion with drift $(X_{t\in[0,\infty)}) \equiv (\mu t + \sigma B_{t\in[0,\infty)})$ defined on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,\infty)}, \mathbb{P})$ is the only 2-stable process with $b(a) = \sigma a^{1/2}$ and $c(t) = -\mu t\sigma\sqrt{a} + \mu at$.

Proof

We know:

$$X_t \sim Normal(\mu t, \sigma^2 t)$$
$$X_{at} \sim Normal(\mu at, \sigma^2 at)$$
$$\sigma\sqrt{a}X_t \sim Normal(\mu t\sigma\sqrt{a}, \sigma^2 at).$$

Therefore:

$$(X_{at\in[0,\infty)}) \underset{=}{\overset{d}{}} (\sigma\sqrt{a}X_{t\in[0,\infty)} - \mu t\sigma\sqrt{a} + \mu at).$$

□

Theorem 5.7 Brownian motion with drift as the only 2-stable process A real valued Lévy process $(X_{t\in[0,\infty)})$ (by definition its increments possess infinite divisibility) defined on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,\infty)}, \mathbb{P})$ is a Brownian motion with drift, if $(X_{t\in[0,\infty)})$ is a 2-stable process. Its converse is also true. A real valued Lévy process $(X_{t\in[0,\infty)})$ is a 2-stable process, if $(X_{t\in[0,\infty)})$ is a Brownian motion with drift.

Theorem 5.7 Standard Brownian motion as the only strictly 2-stable process A real valued Lévy process $(X_{t\in[0,\infty)})$ (by definition its increments possess infinite divisibility) defined on a filtered probability space $(\Omega, \mathcal{F}_{t\in[0,\infty)}, \mathbb{P})$ is a standard Brownian motion, if $(X_{t\in[0,\infty)})$ is a strictly 2-stable process. Its converse is also true. A real valued Lévy process $(X_{t\in[0,\infty)})$ is a strictly 2-stable process, if $(X_{t\in[0,\infty)})$ is a standard Brownian motion.

Example 5.8 Cauchy distribution as a strictly 1-stable distribution Consider a Cauchy random variable Y with the location parameter $c \in \mathbb{R}$ and the scale parameter $d \in \mathbb{R}^+$ whose probability density function is given by:

$$f(x) = \frac{1}{\pi}\frac{d}{d^2 + (x-c)^2}.$$

Then, Y is a strictly 1-stable distribution with $b(a) = a$.

Proof

The characteristic function of a Cauchy random variable Y is, for $\forall \omega \in \mathbb{R}$:

$$\phi_Y(\omega) = \exp(ic\omega - d\left|\omega\right|).$$

Then, Y is a strictly 1-stable distribution with $b(a) = a$, since:

$$\phi_Y(\omega)^a = \left(\exp(ic\omega - d\left|\omega\right|)\right)^a = \exp(ic\omega a - d\left|\omega\right| a)$$
$$\phi_Y(\omega)^a = \exp\left(ic(\omega a) - d(\left|\omega\right| a)\right)$$
$$\phi_Y(\omega)^a = \phi_Y(\omega a)$$

□

[5.4.2] Properties of Stable Distributions and Stable Processes with the Stability Index $0 < \alpha < 2$

Theorem 5.7 Properties of stable distributions with the stability index $0 < \alpha < 2$
Let X be a real-valued random variable from a stable distribution $\mathbb{P}(x)$ with the stability index $0 < \alpha < 2$ (this excludes the Gaussian). Let (A, ℓ, γ) be its Lévy triplet. Then, X has the following properties:

(1) $\mathbb{P}(x)$ is an infinitely divisible distribution (by definition).
(2) $A = 0$. Gaussian variance is zero (Lévy-Itô decomposition).
(3) Its Lévy measure $\ell(x)$ is absolutely continuous and given by, for $c_1 \geq 0$, $c_2 \geq 0$, $c_1 + c_2 > 0$:

$$\ell(x) = \frac{c_1}{x^{1+\alpha}} 1_{x>0} + \frac{c_2}{|x|^{1+\alpha}} 1_{x<0} .$$

(4) The total mass of the Lévy measure $\ell(x)$ is infinite:

$$\int_{-\infty}^{\infty} \ell(x) dx = \infty .$$

For more details and proofs, consult Sato (1999) section 14.

Theorem 5.8 Properties of stable processes with the stability index $0 < \alpha < 2$ Let $(X_{t\in[0,\infty)})$ be a real-valued stable process with the stability index $0 < \alpha < 2$ (this excludes the Brownian motion with drift). Let (A, ℓ, γ) be its Lévy triplet. Then, $(X_{t\in[0,\infty)})$ has the following properties:

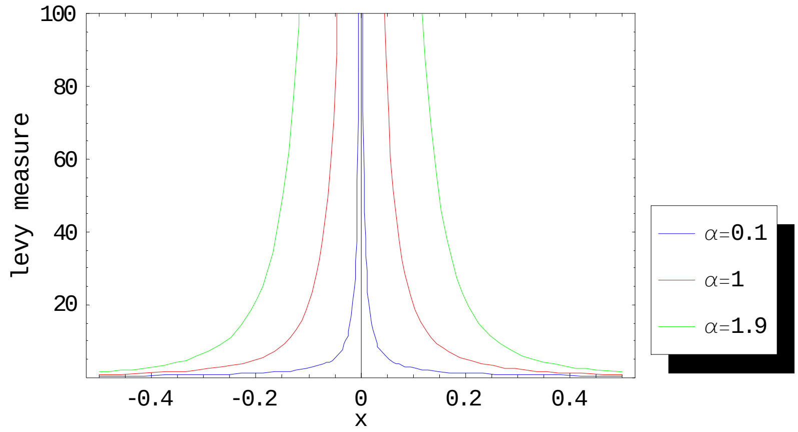
(1) $(X_{t\in[0,\infty)})$ is a Lévy process by definition.
(2) $A = 0$. Gaussian variance is zero (Lévy-Itô decomposition). This means that stable processes with the stability index $0 < \alpha < 2$ are pure jump processes (i.e. purely non-Gaussian processes).
(3) Its Lévy measure $\ell(x)$ is absolutely continuous and given by, for $c_1 \geq 0$, $c_2 \geq 0$, $c_1 + c_2 > 0$:

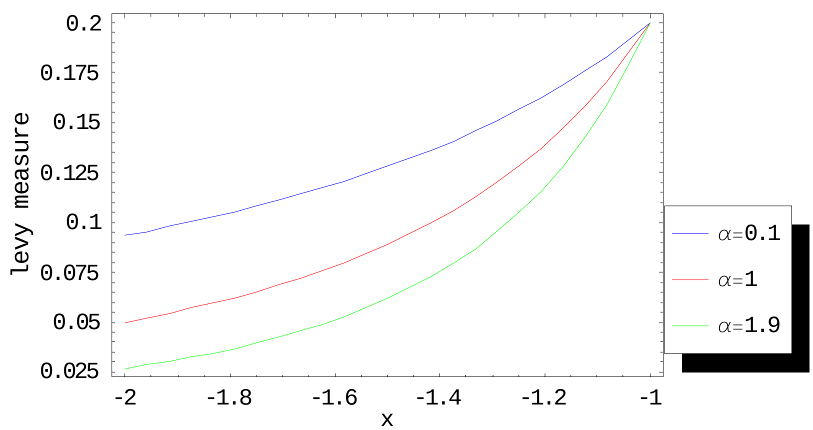
$$\ell(x) = \frac{c_1}{x^{1+\alpha}} 1_{x>0} + \frac{c_2}{|x|^{1+\alpha}} 1_{x<0} .$$

(4) The total mass of the Lévy measure $\ell(x)$ is infinite:

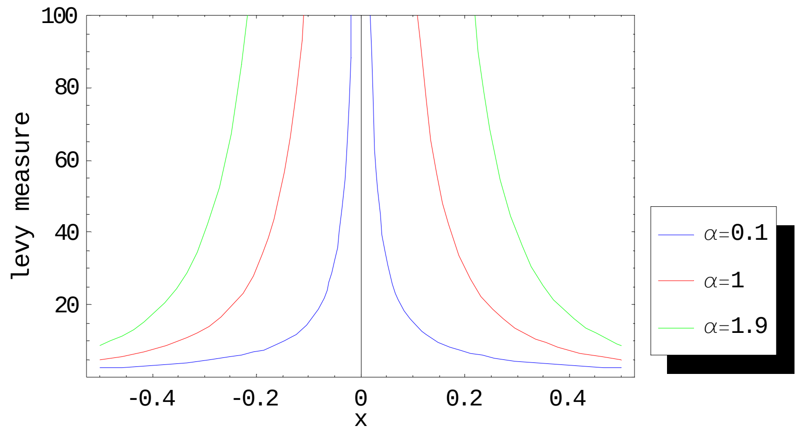
$$\int_{-\infty}^{\infty} \ell(x) dx = \infty .$$

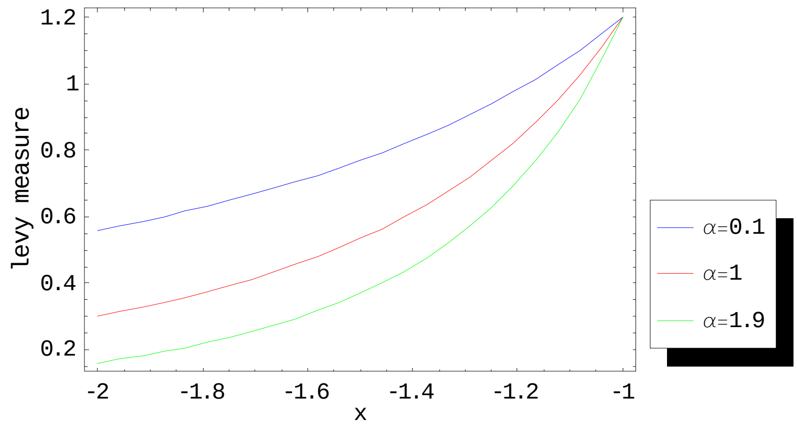
In other words, all stable processes with the stability index $0 < \alpha < 2$ are infinite activity Lévy processes which has a finite number of large jumps and an infinite number of small jumps.
(5) As the stability index gets closer to zero, i.e. $\alpha \to 0$, the Lévy measure $\ell(x)$ becomes less concentrated at zero and its tails become fatter which means that the frequency of arrivals of large jumps increases. Therefore, in this case, large jumps drive the process $(X_{t\in[0,\infty)})$. As the stability index gets closer to 2, i.e. $\alpha \to 2$, the Lévy measure $\ell(x)$ becomes more concentrated at zero and its tails become thinner which means that the frequency of arrivals of large jumps decreases. Therefore, in this case, small jumps drive the process $(X_{t\in[0,\infty)})$. This point is illustrated by Figure 5.2 where the nonnegative constants are set as $c_1 = c_2 = 1.2$ (this implies a symmetric Lévy measure).





A) Three different values for the stability index $0 < \alpha < 2$ with nonnegative constants are all set $c_1 = c_2 = 0.2$.





B) Three different values for the stability index $0<\alpha<2$ with nonnegative constants are all set $c_1 = c_2 = 1.2$.

Figure 5.2 Plot of Lévy measure $\ell(x)$ of stable processes with the stability index $0<\alpha<2$

Theorem 5.9 Characteristic function of stable processes with the stability index $0<\alpha<2$ Let $(X_{t\in[0,\infty)})$ be a real valued stable process with the stability index $0<\alpha<2$ (this excludes the Brownian motion with drift). Then, for any $\omega \in \mathbb{R}$, the characteristic function $\phi_X(\omega)$ of $(X_{t\in[0,\infty)})$ can be expressed as:

$$\phi_X(\omega) = \exp\left(t\psi_X(\omega)\right),$$

where $\psi_X(\omega)$ called a characteristic exponent is given by:

(1) $\psi_X(\omega) = i\tau\omega - c\left|\omega\right|^{\alpha}\left(1 - i\beta \operatorname{sgn}\omega \tan\dfrac{\pi\alpha}{2}\right)$, when $\alpha \neq 1$,

(2) $\psi_X(\omega) = i\tau\omega - c\left|\omega\right|\left(1 + i\beta\dfrac{2}{\pi}\operatorname{sgn}\omega \ln\left|\omega\right|\right)$, when $\alpha = 1$,

where $c>0$, $\beta \in [-1,1]$ and $\tau \in \mathbb{R}$. A parameter τ is called a shift parameter which equals the drift γ_0 of the Lévy-Itô decomposition when $0<\alpha<1$ and equals the center γ_1 when $1<\alpha<2$. A parameter c is a scale parameter. A parameter β determines the skewness of the Lévy measure $\ell(x)$. As a compact notation, $S_\alpha(c,\beta,\tau)$ denotes a real valued stable distribution with the stability index $\alpha \in (0,2]$ and with parameters $c>0$, $\beta \in [-1,1]$ and $\tau \in \mathbb{R}$.

Proof

This corresponds to theorem 14.10 and 14.15 of Sato (1999) where the proofs are given.

Theorem 5.10 A necessary and sufficient condition of a strictly stable distribution
A real valued stable distribution $S_\alpha(c,\beta,\tau)$ with the stability index $0<\alpha<2$ (this excludes the Gaussian) is strictly stable, if it satisfies the following conditions:

(1) $\tau=0$, when $\alpha\neq 1$.
(2) $\beta=0$, when $\alpha=1$.

Proof

This corresponds to theorem 14.10 and 14.15 of Sato (1999) where the proofs are given.

Theorem 5.11 Role of skewness parameter β of stable distribution The Lévy measure $\ell(x)$ of a real valued stable distribution $S_\alpha(c,\beta,\tau)$ with the stability index $0<\alpha<2$ (this excludes the Gaussian) is:

(1) symmetric, if and only if $\beta=0$.
(2) concentrated on the positive half axis, i.e. $\int_{-\infty}^{0}\ell(x)dx=0$, if and only if $\beta=1$. This means that the process has no negative jumps.
(3) concentrated on the negative half axis, i.e. $\int_{0}^{\infty}\ell(x)dx=0$, if and only if $\beta=-1$. This means that the process has no positive jumps.

This corresponds to definition 14.16 of Sato (1999).

Definition 5.7 Symmetric stable distribution A real valued stable distribution $S_\alpha(c,\beta,\tau)$ with the stability index $0<\alpha<2$ (this excludes the Gaussian) is said to be a symmetric stable distribution if $S_\alpha(c,\beta,\tau)$ satisfies the following condition:

(1) $\beta=0$. This means that the Lévy measure $\ell(x)$ is symmetric.
(2) $\tau=0$. Zero shift parameter.
(3) Its characteristic function is of the form, from theorem 5.9:

$$\phi_X(\omega)=\exp\left(-c\left|\omega\right|^\alpha\right).$$

Note the condition (3) is implied by the conditions (1) and (2).

Theorem 5.12 (with Proof) Moments of stable distribution with the stability index $0<\alpha<2$ Consider a real valued stable distribution $S_\alpha(c,\beta,\tau)$ with the stability index $0<\alpha<2$ (this excludes the Gaussian). Then, $S_\alpha(c,\beta,\tau)$ admits a first moment which is equal to zero only if $\alpha>1$ and never admits a second moment plus higher order moments due to the explicit form of the Lévy measure given by theorem 5.8 or that of the characteristic function given by theorem 5.9. For the ease of illustration, consider a symmetric stable random variable X with the characteristic function (i.e. definition 5.7):

$$\phi_X(\omega) = \exp\left(-c|\omega|^{\alpha}\right).$$

From its characteristic exponent $\psi_X(\omega) = -c|\omega|^{\alpha}$, its cumulants are calculated as:

$$\begin{aligned}
c_1 &= 0^{\alpha-1} i\alpha c \\
c_2 &= 0^{\alpha-2}(\alpha-1)\alpha c \\
c_3 &= 0^{\alpha-3}(-i)(\alpha-1)(\alpha-2)\alpha c \\
c_4 &= -0^{\alpha-4}(\alpha-1)(\alpha-2)(\alpha-3)\alpha c\,.
\end{aligned}$$

It is obvious that X admits a first moment which equals zero if $\alpha > 1$, and never admits a second moment because $0 < \alpha < 2$ makes $0^{\alpha-2}$ a complex infinity.

Theorem 5.13 Three closed form probability density functions of stable distributions
The probability density functions of real valued stable distributions $S_{\alpha}(c, \beta, \tau)$ with the stability index $0 < \alpha \le 2$ are known in closed form for only three following cases:

(1) Gaussian distribution $S_2(c, 0, \mu)$:

$$S_2(c, 0, \mu) = \frac{1}{\sqrt{2\pi(2c)}} \exp\left(-\frac{1}{2}\frac{(x-\mu)^2}{(2c)}\right).$$

(2) Cauchy distribution $S_1(c, 0, \mu)$:

$$S_1(c, 0, \mu) = \frac{1}{\pi}\frac{c}{c^2 + (x-\mu)^2}\,.$$

(3) Lévy distribution $S_{1/2}(c, 1, \mu)$:

$$S_{1/2}(c, 1, \mu) = \begin{cases} \dfrac{c}{\sqrt{2\pi}}(x-\mu)^{-(1+\frac{1}{2})} \exp\left\{-\dfrac{c^2}{2(x-\mu)}\right\} & if \quad x > \mu \\ 0 & \text{otherwise} \end{cases}.$$

Appendix

[A.1] Dirac's Delta Function (Impulse Function)

Consider a function of the form with $n \in \mathbb{R}^+$:

$$h(x) = \frac{n}{\sqrt{\pi}} \exp\left(-n^2 x^2\right).$$

This function is plotted in Figure A.1.1 for three different values for n. The function $h(x)$ becomes more and more concentrated around zero as the value of n increases. The function $h(x)$ has a unit integral:

$$\int_{-\infty}^{\infty} h(x)dx = \int_{-\infty}^{\infty} \frac{n}{\sqrt{\pi}} \exp\left(-n^2 x^2\right) dx = 1.$$

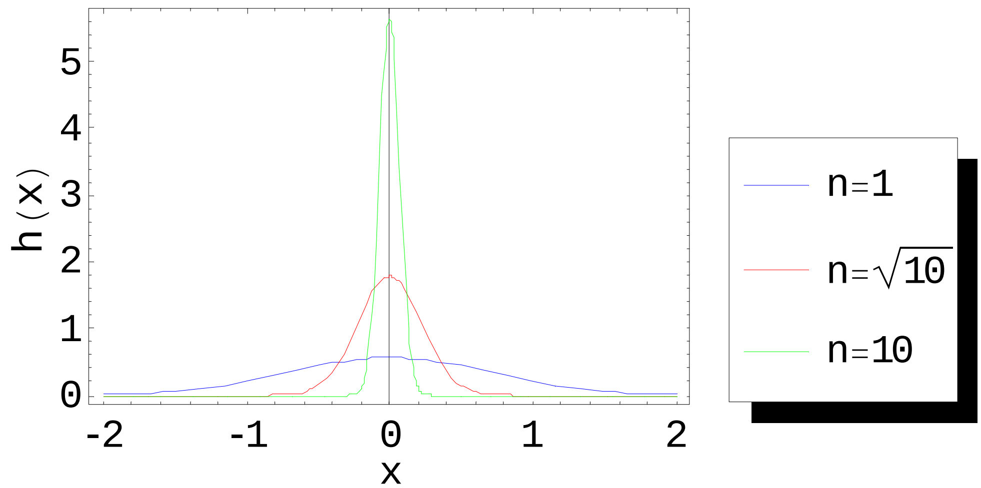


Figure A.1.1 Plot of a function h(x) for n = 1, n = $10^{1/2}$, and n = 10.

Dirac's delta function denoted by $\delta(x)$ can be considered as a limit of $h(x)$ when $n \to \infty$. In other words, $\delta(x)$ is a pulse of unbounded height and zero width with a unit integral:

$$\int_{-\infty}^{\infty} \delta(x)dx = 1.$$

Dirac's delta function $\delta(x)$ evaluates to 0 at all $x \in \mathbb{R}$ other than $x = 0$:

$$\delta(x) = \begin{cases} \delta(0) & \text{if} \quad x = 0 \\ 0 & \text{otherwise} \end{cases}$$

where $\delta(0)$ is undefined. $\delta(x)$ is called a generalized function not a function because of undefined $\delta(0)$. Therefore, $\delta(x)$ is a distribution with compact support $\{0\}$ meaning that $\delta(x)$ does not occur alone but occurs combined with any continuous functions $f(x)$ and is well defined only when it is integrated.

Dirac's delta function can be defined more generally by its sampling property. Suppose that a function $f(x)$ is defined at $x = 0$. Applying $\delta(x)$ to $f(x)$ yields $f(0)$:

$$\int_{-\infty}^{\infty} f(x)\delta(x)dx = f(0).$$

This is why Dirac's delta function $\delta(x)$ is called a functional because the use of $\delta(x)$ assigns a number $f(0)$ to a function $f(x)$. More generally for $a \in \mathbb{R}$:

$$\delta(x-a) = \begin{cases} \delta(0) & \text{if} \quad x = a \\ 0 & \text{otherwise} \end{cases}$$

and:

$$\int_{-\infty}^{\infty} f(x)\delta(x-a)dx = f(a),$$

or for $\varepsilon > 0$:

$$\int_{a-\varepsilon}^{a+\varepsilon} f(x)\delta(x-a)dx = f(a).$$

$\delta(x)$ has identities such as:

$$\delta(ax) = \frac{1}{|a|}\delta(x),$$
$$\delta(x^2 - a^2) = \frac{1}{2|a|}\left[\delta(x+a) + \delta(x-a)\right].$$

Dirac's delta function $\delta(x)$ can be defined as the limit $n \to \infty$ of a class of delta sequences:

$$\delta(x) = \lim_{n\to\infty} \delta_n(x),$$

such that:

$$\lim_{n\to\infty} \int_{-\infty}^{\infty} \delta_n(x) f(x) dx = f(0),$$

where $\delta_n(x)$ is a class of delta sequences. Examples of $\delta_n(x)$ other than (3.8) are:

$$\delta_n(x) = \begin{cases} \text{n} & \text{if} \quad \text{-}1/2n < x < 1/2n \\ 0 & \text{otherwise} \end{cases},$$

$$\delta_n(x) = \frac{1}{2\pi} \int_{-n}^{n} \exp(iux) du,$$

$$\delta_n(x) = \frac{1}{\pi x} \frac{e^{inx} - e^{-inx}}{2i},$$

$$\delta_n(x) = \frac{1}{2\pi} \frac{\sin\left[\left(n + 1/2\right) x\right]}{\sin\left(x/2\right)}.$$

Bibliography

Applebaum, D. 2004, Lévy Processes and Stochastic Calculus, Cambridge University Press.

Barndorff-Nielsen, O. et al. 2001, Lévy Processes: Theory and Applications, Birkhäuser.

Brzezniak, Z. and Zastawniak, T. 1999, Basic Stochastic Processes, Springer.

Capinski, M., Kopp, E., and Kopp, P. E., 2004, Measure, Integral and Probability, Springer-Verlag.

Cont, R. and Tankov, P., 2004, Financial Modelling with Jump Processes, Chapman & Hall/CRC Financial Mathematics Series.

Feller, W., 1968, An Introduction to Probability Theory and Its Applications, John Wiley & Sons.

Harrison, J. M., and Pliska, S.R., 1981, "Martingales and Stochastic Integrals In the Theory of Continuous Trading," Stochastic Processes and Applications 11, 215-260.

Harrison, J. M. and Pliska, S.R., 1983, "A Stochastic Calculous Model of Continuous Trading: Complete Markets," Stochastic Processes and Applications 15, 313-316.

Karatzas, Ioannis., and Shreve, S. E., 1991, Brownian Motion and Stochastic Calculus, Springer-Verlag.

Karlin, S., and Taylor, H. M., 1975, A First Course in Stochastic Process, Academic Press.

Neftci, S. N., 2000, An Introduction to the Mathematics of Financial Derivatives, Academic Press.

Oksendal, B., 2003, Stochastic Differential Equations: An Introduction with Applications, Springer.

Rogers, L.C.G., and Williams, D., 2000, Diffusions, Markov Processes and Martingales, Cambridge University Press.

Ross, S. M., 1983, Stochastic Processes, John Wiley & Sons.

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